Title

A population-scale atlas of blood and tissue in lupus nephritis

Author list

Siddarth Gurajala[1,2,3,4]*, Nicholas Sugiarto[1,2,3,4]*, Michelle Curtis[1,2,3,4]*, Tom Eisenhaure[3]*, Arnon Arazi[5], Andrea Fava[6], Qian Xiao[1,2,3,4], Joseph Mears[1,2,3,4], Brad Rovin[7], Celine C. Berthier[8], Peter Izmirly[9], Jennifer Barnas[10], Paul J. Hoover[3], Michael Peters[3], Raktima Raychowdhury[3], Alice Horsberger[2], Saori Sakaue[1,2,3,4], Yu Zhao[1,2,3,4], Richard A. Furie[11], Michael Belmont[9], Dave Hildeman[12,13], E. Steve Woodle[14], Maria Dall'Era[15], Chaim Putterman[16], Diane L. Kamen[17], Maureen A. McMahon[18], Jennifer Grossman[18], Ken Kalunian[19], Jeffrey Hodgin[20], Fernanda Payan-Schober[21], William Apruzzese[22], Harris Perlman[23], Carla M. Cuda[23], David Wofsy[15], Joel M. Guthridge[24], Jennifer H. Anolik[10], Judith A. James[24], Accelerating Medicines Partnerships Rheumatoid Arthritis/Systemic Lupus Erythematosus, Deepak A. Rao[2], Anne Davidson[5], Michelle A. Petri[6], Jill P. Buyon[9], Nir Hacohen[3], Betty Diamond[5]†, Soumya Raychaudhuri[1,2,3,4,25]†

* These authors contributed equally
†Co-senior, co-corresponding authorship

Affiliations

[1]Center for Data Sciences, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA. [2]Division of Rheumatology, Inflammation, and Immunity, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA. [3]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. [4]Division of Genetics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA. [5]Center for Autoimmune, Musculoskeletal and Hematopoietic Diseases, The Feinstein Institute for Medical Research, Northwell Health, Manhasset, NY, USA. [6]Division of Rheumatology, Johns Hopkins University, Baltimore, MD, USA. [7]Ohio State University Wexner Medical Center, Columbus, OH, 43210, USA. [8]Internal Medicine, Department of Nephrology, University of Michigan, Ann Arbor, MI, USA. [9]Department of Medicine, Division of Rheumatology, New York University School of Medicine, New York, NY, USA. [10]Department of Medicine, Division of Allergy, Immunology, and Rheumatology, University of Rochester Medical Center, Rochester, NY, USA. [11]Division of Rheumatology, Northwell Health, Great Neck, NY, USA. [12]Department of Pediatrics, University of Cincinnati, Cincinnati, OH, USA. [13]Division of Immunobiology, Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, USA. [14]Division of Transplantation, Department of Surgery, University of Cincinnati College of Medicine, Cincinnati, OH, USA. [15]Rheumatology Division, University of California San Francisco, San Francisco, CA, USA. [16]Division of Rheumatology and Department of Microbiology and Immunology, Albert Einstein College of Medicine and Montefiore Medical Center, Bronx, NY, USA. [17]Division of Rheumatology and Immunology, Medical University of South Carolina, Charleston, SC, USA. [18]Department of Medicine, University of California Los Angeles, Los Angeles, CA, USA. [19]School of Medicine, University of California San Diego, La Jolla, California, USA. [20]Department of Pathology, University of Michigan, Ann Arbor, MI, USA. [21]Department of Medicine, Paul L. Foster School of Medicine, Texas Tech University Health Sciences Center, El Paso, TX, USA. [22]Pfizer Inc., New York, NY, USA. [23]Division of Rheumatology, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA. [24]Arthritis

and Clinical Immunology, Oklahoma Medical Research Foundation, Oklahoma City, OK, USA.
[25]Department of Biomedical Informatics, Harvard Medical School, Boston, MA 02115, USA.

Consortia authors

Accelerating Medicines Partnerships Rheumatoid Arthritis/Systemic Lupus Erythematosus

Nicole Bornkamp, Michael B. Brenner, Robert M. Clancy, Evan Der, Eugene Drokhlyansky, Derek Fine, Beatrice Goilav, V. Michael Holers, Mariko Ishimori, Nicole Jordan, Chun-Hao Lee, Matthias Kretzler, Blue Lake, Rajasree Menon, James Pullman, Amit Saxena, Kamil Slowikowski, Manny Monroy-Trujillo, Michael Peters, Daniel Schwartz, Jennifer Seifert, Ummara Shah, Dawn Smilek, Patti Tosta, Thomas Tuschl, Paul J. Utz, Michael Weisman, Ming Wu, Hasret Gunduz, Yemil Atisha-Fregoso

Abstract

Lupus nephritis (LN), a severe manifestation of Systemic lupus erythematosus (SLE), is a heterogeneous disease driven by diverse immune and tissue cell types. Current treatments of LN are non-specific, increase risk of infection, and have high rates of relapse. To better understand LN, we obtained 538K single-cell and 143K single-nuclear profiles from kidney biopsies of 156 LN patients and 30 pre-implantation transplant biopsy controls, along with 327K single-cell blood profiles from many of these patients. We built an atlas of kidney tissue cell types in LN encompassing 55 immune and 36 tissue states. Integrating renal immune single cell data with blood data enabled the identification of tissue-specific immune cell states. For example, within the myeloid compartment, plasmacytoid dendritic cells (DC20) were shared across tissue and blood, while macrophage cell states were tissue-specific (M5-14). We identified LN pathological features associated with cell states. Specifically, we examined the chronicity index, which quantifies irreversible tissue damage, and the activity index, which measures active reversible inflammation. Increasing chronicity index tracked with the most dramatic changes in both tissue and immune cell states. The proportion of injured or degenerating proximal tubule cells expanded with chronicity, reflecting ongoing tissue damage. Increasing chronicity index was also associated with an expansion of tissue-specific CD56^{bright} NK cells (NK1), GPNMB^{high}NUPR1^{high} Macrophages (M5), and CLEC10A^{low} cDC2 (DC17). Importantly, most expanded immune populations associated with chronicity index were tissue-specific states. After adjusting for chronicity effects, we identified specific cell states in the myeloid compartment that associated with activity. Highly tissue-specific macrophage populations expressing known Scar Associated Macrophage (SAM) gene signatures were particularly expanded, including M5, GPNMBhigh LYVE1low (M11), C1Qlow SPP1high, SPP1highFABP5high (M7), and MERTKhighFABP5high (M9). We hypothesize that M5 may be a profibrotic pathogenic population that could be targeted for therapeutic benefit. We further demonstrate that blood only partially reflects renal tissue. Our results define key renal cell states

in the tissue and immune compartments and argue that profound changes in both compartments are associated with increasing chronicity, while specific changes in the myeloid compartment track with activity. In aggregate, these observations strongly support that therapeutic targeting of myeloid populations, in addition to B cells, may offer an unproved strategy to prevent progression of renal inflammation and ongoing kidney damage in LN.

Introduction

Systemic lupus erythematosus (SLE) is a chronic autoimmune disease characterized by autoantibodies against nuclear components and heterogenous clinical manifestations across multiple organ systems[1]. Affecting 0.1% of the population, SLE predominantly affects female and non-white individuals[2]. Lupus nephritis (LN) presents in 30-50% of patients and results from inflammation-induced kidney damage[3,4]. Immunosuppressive therapies including prednisone, mycophenolate mofetil, cyclophosphamide, rituximab, belimumab, and voclosporin, are inconsistently effective and can increase infection risk. Fewer than 50% of individuals achieve complete renal response[5], and relapse is common[6]. Approximately 15% progress to end-stage renal disease (ESRD) despite therapy[7] resulting in morbidity and mortality[8,9].

Kidney biopsy remains essential for LN diagnosis and treatment despite efforts to identify serum biomarkers. The International Society of Nephrology (ISN) classifies LN into six classes. Class I-IV glomerular pathology denotes increasing proliferative disease. Class V designates membranous disease, alone or in conjunction with proliferative disease. The rare Class VI reflects severe scarring with globally sclerotic glomeruli. Clinicians treat class III, IV and V aggressively, although emerging data supports treating less advanced disease forms as well[11,12]. Renal biopsies are also scored using the NIH activity and chronicity indices[13,14]. The activity index captures active and potentially reversible inflammation, but high sustained scores predict poor survival[15]. It reflects accumulation of immune cells in the glomerular capillaries (glomerular endocapillary hypercellularity), neutrophilic infiltration, fibrinoid necrosis, large

subendothelial immune complexes (hyaline deposits), cellular and fibrocellular crescents, and cortical tubulointerstitial inflammation[14]. The chronicity index captures irreversible tissue damage, which increases with each LN flare[16] and correlates with reduced kidney survival. It includes glomerulosclerosis, fibrous crescents, tubular atrophy, and interstitial fibrosis. While imperfect predictors, biopsy characteristics outperform other clinical metrics. Indicators such as low serum albumin and persistent proteinuria (>800mg/g creatinine) suggest poor long-term prognosis[17–19], but may not be concordant with the underlying histology[20]. Thus, clinicians variably place weight on ISN class, activity, and chronicity in therapeutic decisions.

The goal of the Accelerating Medicines Partnership in SLE is to profile single cells from human patient tissues, such as the nephritic kidney, to advance understanding of LN pathology at cellular resolution[17,21]. We applied single-cell (scRNAseq) and single-nuclear (snRNAseq) RNA-sequencing to profile immune and non-immune cell populations, respectively. These methods recovered different proportions of cell types, as non-immune cells were more sensitive to enzymatic disruption of cryopreserved tissue. With these data, we defined key tissue-specific cell states and identified those linked to case/control status, chronicity, and activity.

Results

Cell State Definition and Clustering

We obtained kidney biopsies and peripheral blood mononuclear cell (PBMC) samples from patients with LN and healthy controls (**Methods**). After quality control (**Methods**), the final dataset included tissue and clinical data from 156 LN patients and 30 healthy controls, and matched PBMC samples from 119 of the LN patients and 19 healthy controls (**Figure 1a, Supplementary Table 1**). The kidney tissue data included 538,194 scRNA-seq profiles and 142,881 snRNA-seq profiles from a subset of 50 patients, the latter preserving disaggregation-sensitive cell types. We applied weighted principal components analysis[22]

followed by Harmony[23] to ensure both modalities contributed equally (**Methods**). Clustering and annotating cells together ensured that shared transcriptional signals represented biology and not modality-specific technical artifacts. (**Figure 1b**). The PBMC data included 327,326 scRNA-seq profiles, which we integrated with Harmony (**Methods**). We also integrated PBMC and kidney tissue data (**Methods**) to identify tissue-specific cell states.

We associated cell states with case/control status and histopathological metrics using Covarying Neighborhood Analysis (CNA)[24] (**Figure 1c-d**). Rather than relying on discrete clusters, CNA overcomes multiple hypothesis testing power limitations by examining groups of similar individual cells, known as neighborhoods, and identifying those significantly associated with a sample-level variable of interest. We interpret the associated neighborhoods within the context of annotated clusters.

Clustering transcriptionally similar cells in renal tissue revealed 3 immune (T/NK, B/Plasma, Myeloid) and 5 renal tissue (glomeruli, proximal tubule, loop of Henle, distal Nephron, endothelial/stromal) populations (**Figure 1e-f**), with consistent marker genes across integrated sc and sn data (**Supplementary Note**). The immune compartment included 71,859 single-cells and 4,652 single-nuclei. The tissue compartment included 466,335 single-cells and 138,229 single-nuclei. Prior studies show related immune cell enrichment in sc data, likely due to destruction of tissue cells during tissue dissociation[25,26]. Consistently, 93.9% of immune cell profiles were from sc rather than sn data.

Subclustering each broad cell type revealed 55 fine-grained renal immune cell states including 22 T/NK (**Figure 2a**), 21 Myeloid (**Figure 2b**), and 9 B/Plasma clusters (**Figure 2c**). Within the renal parenchymal tissue cell types, we obtained 36 states, including 8 Proximal Tubule (PT, **Figure 2d**), 5 Loop of Henle (LOH, **Figure 2e**), 8 Distal Nephron (DN, **Figure 2f**), 10 Endothelial/Stromal (EC, **Figure 2g**), and 5 Glomerular clusters (GLOM, **Figure 2h**). Marker genes and cell state frequencies were concordant across sc and sn data (**Supplementary Note, Supplementary Figures 1-8**).

Renal Immune Cell States

Of the 22 T/NK clusters, two were NK clusters with high (NK0) or low (NK1) *NCAM1* (CD56) expression (**Supplementary Table 2**). CD8 T cell clusters included central memory (*SELL*+) and effector (*GZMK+/GZMB*+) subsets (**Supplementary Figure 1**), with nine total CD8 states.This included two *CD8+GZMB*+ cytotoxic lymphocyte (CTL) clusters (T2, T3) and a *CD69+ITGAE+CCR7-SELL-S1PR1-CD8+* resident memory-like cluster (T9). We classified CD4 clusters into memory and naive subsets using *CD28*, *CD69*, *SELL*, *CCR7*, *ITGAE*, and *S1PR1*. We identified two CD4 *CD28+CD69+ITGAE+CCR7-SELL-S1PR1-* resident memory clusters (T13, T14), along with a *CD28+CCR7-ITGAE-SELL-* effector memory-like cluster (T11), and specialized subsets, including *IL2RA+FOXP3+* Tregs (T19), *PDCD1+CXCR5+* T Follicular helper (TFH/TPH) (T21), and *RORC+KLRB1+CCR6+* Th17 (T17).

We identified 6 B cell states, expressing MS4A1, BACH2, and IL4R, and 3 Plasma cell states, expressing SDC1, XBP1, and PRDM1(**Supplementary Table 2**) We identified a *FOXO1+BCL6+* germinal center (GC)-like B cell cluster (B0) and a *FCRL5+ITGAX+TBX21+* autoimmune-associated B cell (ABC) cluster (B5)[27], with high and low expression of *IGHG2* and *IGHD*, respectively. We identified a *CD38+CD27+MKI67+* plasmablast-like cluster with low expression of *IGHD* (P8). Plasma cells clustered by *IGG* (P6) or *IGA* (P7) heavy chain immunoglobulin genes.

Within the 21 myeloid clusters, we distinguished macrophages, monocytes, and dendritic cells with marker genes including *APOE* and *C1Q* (**Supplementary Table 2**, **Supplementary Figure 3**). Four monocyte clusters (M0, M1, M2, M3) spanned the CD14+ (classical) to CD16+ (non-classical) gradient. Several macrophage clusters coexpressed *SPP1, LPL, CD9, TREM2, CTSB,* and *CTSD*. Within this group, we identified two *GPNMB*^{high} macrophage clusters: *GPNMB*^{high}*NUPR1*^{high} (M5) and *GPNMB*^{high}*NUPR1*^{low} (M11). We also identified three *FABP5*^{high}

clusters: *SPP1*^{high}*FABP5*^{high} (M7), SPP1^{low}FABP5^{high} (M8), *MERTK*^{high}*FABP5*^{high} (M9).We distinguished three macrophage clusters by *SELENOP* and *LYVE1* expression levels: *SELENOP*^{inter}ISG^{high} (M6), *SELENOP*^{inter}*LYVE1*^{inter} (M10) and *SELENOP*^{high}*LYVE1*^{high} resident macrophages (M12). Dendritic cell subsets included two cDC2 clusters (DC15, DC17), *CCR7*+*LAMP3*+ DCs (DC16), pDCs (DC20), and a cDC1 *XCR1*+*CLEC9A*+ cluster (DC19).

Renal Parenchymal Cell States

We identified 8 proximal tubule (PT) clusters, using *SLC34A1, SLC22A8,* and *MOGAT1* to delineate anatomical segments (S1, S2, S3) (**Supplementary Figure 4**, **Supplementary Table 2**). (PT1-4) reflected anatomical segmentation, while PT0 and PT5-7 reflected injured populations. Anatomical clusters included two S1/S2 (PT1, PT4) and one S2/S3 cluster (PT3). S1/S2 coexpressed S1 (*PRODH2*, *SLC5A2*, *SLC22A8*) and S2 genes (*SLC34A1, SLC22A7*) while S2/S3 coexpressed S2 and S3 (*MOGAT1, SATB2, ABCC3*) genes. Injury clusters PT5 and PT6 indicated successful redifferentiation into S1/S2 and S2/S3 subtypes, respectively, marked by *LRP2*, *MAF* in PT5 and *LRP2*, *PAX2*, and *HNF4A* in PT6[28]. PT0 expressed *VCAM1*, *TPM1*, and *DCDC2*[29,30] indicating maladaptive redifferentiation following injury[31]. PT7 expressed *FOS*, *JUN*, and *EGR1,* associated with a degenerative injury response[32,33].

We identified 8 fine-grain distal nephron (DN) clusters (**Supplementary Table 2, Supplementary Figure 6**), grouped into distal convoluted tubule (DCT), intercalated cells (IC), connecting tubule (CNT), and collecting duct principal cells (CD-PC) based on expression of *CLNK*, *SLC12A3*, *AQP2*, and *SLC8A1* (**Supplementary Figure 6**). Two CNT clusters coexpressed *SLC8A1*, *CALB1,* and *SC2NA*, and were distinguished by *DEFB1* (DN3) or *SPP1* (DN4). Two CD-PC clusters (DN1, DN7), expressed *GATA3*, *SCNN1G, SCNN1B* and *FXYD4*. IC subtypes included IC-beta-like CD-PCs (DN7) expressing *INSRR*, *SLC35F3*, and *TLDC2* and IC-alpha cell states resembling CNT cells (DN5), expressing *SC2NA* and *CALB1,* or cortical collecting duct cells (DN2), expressing *ADGRF5* and *LEF1*.

Renal endothelium and stroma (EC/STROMA) yielded 10 cell states: fibroblasts (FIB), endothelial (EC), vascular smooth muscle cells (VSMC), and mesangial cells, distinguished by *PECAM1*, *COL1A1*, *PLVAP*, and *NOTCH3* expression (**Supplementary Figure 7**, **Supplementary Table 2**). Two *PLVAP*^{high} fenestrated EC clusters (EC0, EC1), represented either fenestrated cells of unknown origin or from the ascending vasa recta (EC-AVR). Other EC subsets included descending vasa recta (EC2), ascendant efferent arteriole (AEA) (EC3), glomerular *EMCN*^{high} (EC4), and lymphatic ECs (EC6). VSMCs and pericytes (STR9, EC5) shared markers *MCAM, NTRK3,* and *RGS5.* FIBs (STR8) expressed *DCN, COL1A2,* and *COL1A1*. Mesangial cells (STR7) expressed *POSTN, GATA3,* and *DAAM3*.

We identified 5 loop of Henle (LOH) clusters (**Supplementary Figure 5, Supplementary Table 2**) including descending thin limb (DTL), ascending thin limb (ATL) cells, thick ascending limb (TAL), and macula densa cells, defined by *SLC12A1, CRYAB*, *TMEM25B*, and *EGF* expression (**Supplementary Figure 5**). The TAL cluster (LOH0) expressed *CYFIP2*, *MFSD4A, RAP1GAP,* and *ANK2*. Cortical TAL (C-TAL) (LOH2) expressed *TMEM52B, CLDN16,* and *WNK1.* A mixed DTL/ATL cluster (LOH1) expressed *CRYAB, TACSTD2,* and *SLC44A5.* Macula Densa (LOH4) expressed *PAPPA2* and *NOS1*. A low-quality cluster (LOH3) showed high mitochondrial and ribosomal expression.

We identified 5 glomerular (GLOM) clusters, including podocytes (GLOM2), expressing *PODXL*, NPHS2, and WT1, and parietal epithelial cells (PECs), expressing *VCAM1* and *CFH* (**Supplementary Figure 8**, **Supplementary Table 2**). The three PEC clusters included *BCAM*^{high}*AKAP12*^{high} PEC (GLOM0), *SDK1*^{high}HDAC9^{high} PEC (GLOM1) with high TNC, and *CXCL14*^{high}*MP7*^{high} PEC (GLOM3). A proliferating cluster (GLOM4) expressed *HELLS* and *CPNK*.

Comparing Immune Populations in Blood and Tissue

We profiled overlapping PBMC samples from LN patients and healthy controls, and annotated cell states with canonical marker genes (**Extended Data Figure 1**, **Supplementary Note**). To identify immune cell states shared between kidney and blood, or distinct to either tissue type, we integrated both kidney and PBMC immune populations into a joint nearest neighbor embedding (**Figure 3**, **Methods**). We developed a per-cell tissue specificity metric (TSM), defined as the proportion of a given cell's 50 nearest neighbors derived from renal tissue, weighted for differences in dataset size (**Methods**). A TSM of 1 indicated a tissue-specific cell state, 0 indicated a blood-specific state, and 0.5 reflected equal representation in tissue and blood.

Plasma cell and plasmablast clusters (P6-P8) were tissue-specific, with TSM scores near 1 (**Figure 3a-c**). The ABC−like cluster from tissue (B5) co-embedded with blood-derived ABCs (bl-B2) (**Figure 3a, b**), both with a median TSM ~0.5 (**Fig. 3c**), consistent with previous findings that ABCs spanned both blood and tissue[27,34].

In myeloid cells, macrophage populations (M5-12, M14, M16) were tissue-specific (**Figure 3d-f**). Conversely, monocyte populations (M0-2) were shared, likely reflecting migration from the blood. Classical DCs (DC13, DC15, DC17) were tissue-specific, while plasmacytoid DCs (DC20) were shared with blood, consistent with prior observations[35,36]. Non-classical monocytes (bl-M3, bl-M11) were shared between tissue and blood, while classical monocytes (bl-M0-2, bl-M5) were blood-specific.

T/NK cell TSM scores (**Figure 3g, h**) had fewer strictly blood- or tissue-specific populations. Tissue-skewed clusters included CD8+GZMK+ populations (T5, T6, T12) (**Figure 3i**), consistent with previous work[37]. TFH/TPH populations (T21), implicated in autoimmune rheumatic disease[38], and resident memory populations (T9, T10, T13-14) were tissue-skewed. Both CTL clusters (T2, 3) were shared. Some central and effector memory T cells were tissue-skewed (T11, T20), but most (T7, T15, T18) were shared, likely reflecting tissue

infiltration from blood. We observed a shared blood regulatory T cell population (bl-T16), and a tissue-skewed Treg cluster (T19), consistent with our previous analyses[39].

Differential gene expression between tissue and blood for each cell type reflected these same cell type differences (**Supplementary Note**, **Extended Data Figure 2**).

Kidney Case Control Analysis

To understand key differences between LN and unaffected kidney, we performed case-control comparisons with CNA. Examining all kidney cell types together, we observed immune population expansion, reflecting immune infiltration ($p<1x10^{-4}$ for both sc and sn) (**Extended Data Figure 3a-b**). Parenchymal compartment differences between LN patients and living donors may reflect tissue damage or collection differences.

Case-control analysis in sc data within each cell type revealed fine-grained cell state shifts (**Supplementary Table 2**); we detected shifts in all cell types (p<1e-04), except in B/Plasma cells, where low control cell counts limited power (**Supplementary Table 1**). We did not adjust for age and sex in our analysis (**Figure 4**) as doing so had minimal impact (**Extended Data Figure 4**). Within T/NK cells, the greatest disease shifts were in tissue-specific populations (**Figure 4a**) including enrichment of *CD4+IL2RA++FOXP3+* Tregs (T19), *CD4+FOXP3+* Central Memory/Naïve (T20), TFH/TPHs (T21), and *CD56*^{bright} NKs (NK1) and contraction of resident memory populations *GZMK+CD8+* Resident Memory (T9), *CD8+GZMK+CD69+* cells (T12), and *CD4+JUN*^{high} Resident Memory cells (T14). We also observed tissue-specific myeloid population shifts, including enrichment of five macrophage populations: SELENOP^{inter}LYVE1^{inter} Resident Macrophage (M10), GPMNB^{high}NUPR1^{low} Macrophage (M11), *GPNMB*^{high}*NUPR1*^{high} Macrophages (M5), SELENOP^{inter}ISG^{high} Macrophage (M6), and *SPP1*^{high}*FABP5*^{high} Macrophages (M7) and contraction of two: *APOC*^{high}*C3*^{high} Macrophages (M13) and *APOC*^{low}*C3*^{high} Macrophages (M14) (**Figure 4b**).

We identified shifts in the renal parenchymal populations. Unsurprisingly, activated/injury-associated PECs (GLOM1, SDK1highHDAC9high) were expanded in LN, while homeostatic PECs (GLOM0, BCAMhighAKAP12high) contracted (**Figure 4c**)[10,56]. Podocytes (GLOM2) were unexpectedly expanded in LN.. In Endothelial/Stromal cells, fibroblasts (STR8), EC-AVR (EC1), and EC-AEA (EC3) expanded, while VSMC/pericytes (STR9) contracted (**Figure 4d**). In the Loop of Henle, TAL (LOH0) and C-TAL (LOH2) contracted while DTL/ATL (LOH1) and Macula Densa (LOH4) expanded in LN samples (**Figure 4e**). In the Distal Nephron, *SPP1*^{high} CNT (DN4) and CD-PC (DN1) expanded (**Figure 4f**). Among Proximal Tubule cells, *PKHD1*^{high}*SLC3A1*^{high} S2/S3 (PT6) was expanded, while SLC7A7^{high}ALB^{high} S1/S2 (PT1) was contracted (**Figure 4g**). Contraction of specific states may reflect the loss of vulnerable cells in disease; for example, TAL cells have high metabolic demand and are sensitive to hypoxia[40].

Differential gene expression in LN and control samples (**Methods**, **Supplementary Table 3**) revealed widespread upregulation of genes across cell types (**Extended Data Figure 3c**). Interferon-stimulated genes (ISGs) were enriched (*IFI44L*, *IF16*, *MX1*) across cell types, as were genes related to immune activation (*STAT1*, *SPP1*)[41,42]. Several genes were broadly downregulated in LN, including innate and innate-like lymphoid cell regulators[43] (*ZBTB16*, *NFKBIA*, *PDK4*, *KLF9*, *FKBP5*). Individual cell types revealed enrichment for additional disease-relevant programs, including extracellular matrix (ECM) remodelling (*COL1A1*, *MMP7* in parenchymal cell types), fibrosis and wound healing (*TIMP1* in parenchymal cell types, *CDH11* in Endothelial/Stromal, *HBEGF* in Glomerular), lysosomal activity (*CTSD*, *LGMN*, *PCCA* in Myeloid), class II HLA expression (*HLA-DMB*, *HLA-DRB1* in parenchymal cell types and T/NK), and immune activation and regulation (*TIGIT*, *IL27RA* in T/NK, *LCN2* in Loop of Henle) (**Extended Data Figure 3c**). Gene set enrichment (GSEA) confirmed upregulation of the well-described interferon response pathway (**Methods**, **Extended Data Figure 3d, Supplementary Table 4**)[44,45].

Defining the effect of sample level covariates on LN tissue cell states

After noting case-control differences, we wanted to understand how cell states changed across LN samples. To understand this, we applied CNA to test clinical variables and pathological variables, including seven key variables capturing LN tissue state. (5 ISN classes, activity, chronicity). We observed that chronicity consistently had the most dramatic effects; changes were significant across all cell types, except B cells, where we were likely underpowered to detect any effects (**Figure 5a**). We speculated that many alterations due to increased chronicity reflect worsening tissue damage to the kidney.

Defining cell state correlations with chronicity

Chronicity index reflects irreversible damage to the kidney and is associated with reduced treatment response[17,46]. Differential expression analysis revealed gene programs associated with chronicity index across cell types (**Supplementary Table 3**). This included genes related to ECM remodeling or fibrosis (*VCAN*, *MMP7*, *TUBB3*), inflammation and immune response (*S100A9*, *GPNMB*, *C3*), and apoptosis (*CASP3*). The metabolic enzyme *NNMT*, previously linked to renal injury[47,48], was highly upregulated in multiple renal cell types. While the most significant changes occurred in parenchymal cell types, immune populations also showed upregulation of genes related to immune response (**Supplementary Table 3**). Chronicity index was also correlated with total immune cell proportion in a given sample (sc: R=0.31, P=1.2×10^{-4}; sn: R=0.41, P=0.015), suggesting chronic inflammation with increased disease damage (**Supplementary Fig. 9**). GSEA revealed enrichment for ECM remodeling, epithelial-mesenchymal transition, and translation pathways (**Figure 5b, Supplementary Table 4**). ECM remodeling was enriched in endothelial/stromal (nominally) and proximal tubule cells, consistent with prior reports of profibrotic genes in these cell types[49], as well as in distal nephron and loop of Henle cells. Interferon response pathways were downregulated in T/NK, B, Myeloid, Interstitial/Stromal, and Glomerular

cells (**Figure 5b**) but weakly upregulated in loop of Henle and proximal tubule cells, suggesting cell type-specific effects.

To understand the relationship with interferon response, we constructed an ISG score from a curated list of interferon-stimulated genes[44] (**Methods**). LN cases had higher ISG scores than controls in both tissue and immune populations (**Extended Data Figure 5a**). Among LN cases, the ISG score declined with increasing chronicity but remained stable across activity index levels (**Extended Data Figure 5b**).

We noted that chronicity was correlated with collection variables, including non-first biopsy status and biopsy site (**Extended Data Figure 5d-e**). Since chronicity increases with ongoing disease, chronicity index was lower in the first biopsy (p=1.14e-06) (**Extended Data Figure 5f**). Patient recruitment site was correlated with many clinical variables, including first biopsy status (**Extended Data Fig. 5d**); we suspected site-specific differences reflect both demographic and technical factors. To account for this, we adjusted for both first biopsy status and major recruitment sites as potential confounders in all chronicity analyses. Consistent with prior reports[17], chronicity was lower in complete responders than in non-responders (p=0.002) and partial responders (p=0.04) (**Extended Data Figure 5f**).

Broad cell types associated with increasing chronicity (FDR<0.05, p<9.99e-05) included expansion of T/NK, myeloid, B/plasma, and glomerular cells, with variable shifts in proximal tubule and loop of Henle cells (**Figure 5c**). CNA within each cell type revealed fine cell state shifts. In T/NK Cells (FDR<0.05, p=0.006), tissue-specific CD56^{bright} NK (NK1) and some CD4+ Memory/naive clusters (T11, T15) were strongly expanded (**Figure 5d**). Tissue-specific cell states GZMK+CD8+ITGAE+ (T10) and ILCs (T16) were weakly expanded. Conversely, most contracted cell states were shared with blood, including CTL (T2, T3), CD56^{dim} NK (NK0), and proliferating T cells (T4). In Myeloid cells (FDR<0.05, p=2e-4), increasing chronicity tracked with a contraction of monocyte populations shared in the blood and tissue (M0, M1, M3) as well as two tissue resident macrophages (M6, M12) (**Figure 5e**). Conversely, tissue-specific

macrophages (M5, M14) and cDCs (DC18) were expanded with increasing chronicity. In both immune populations, cell populations contracted with increased chronicity had a lower TSM score (less tissue-specific) than those that were not.

Increased chronicity was associated with contraction of glomerular endothelial cells (EC4) (FDR<0.05, p=1.1e-3) and expansion of fibroblasts (STR8) in the interstitial compartment (**Extended Data Figure 6a**). In glomeruli, podocytes were contracted (GLOM2) (FDR<0.05, p<9.99e-05) while injury-associated PEC subsets (GLOM1, GLOM3) were expanded (**Extended Data Figure 6b**). In the Proximal Tubule (FDR<0.05, p<9.99e-05), the late-injury PT0 state was highly expanded, while healthy PT1/PT2 states were contracted (**Figure 5f**). Cell states in the Loop of Henle and Distal Nephron showed both expansions and contractions with chronicity (**Extended Data Figure 6c**, FDR<0.05, p<9.99e-05 both). Adjusting for treatment response had minimal impact (**Supplementary Table 5, Extended Data Figure 6e**).

Proximal Tubule Injury Drives Chronicity-Associated Cell State Shifts

The proximal tubule was the most significantly associated population with chronicity, with the highest correlation with chronicity and largest number of local neighborhood shifts (**Supplementary Table 6**). A key epithelial injury state, PT0 late injury^{high} (R=0.542, P=4.025e-12) was strongly correlated with increasing chronicity but only weakly associated with treatment response and not associated with ISN class (**Extended Data Figure 7a-b**).

CNA within each cell type, adjusting for recruitment site and first biopsy status, revealed that increasing PT0 proportion was significantly associated with cell state shifts in the loop of Henle (p=9.99e-05), distal nephron (p=9.99e-05), endothelial/stromal (p=9.99e-05), glomerular (p=9.99e-05), T/NK (p=2.6e-03), and myeloid compartments (p=2.00e-04) (**Extended Data Figure 7c**). The PT0-associated shifts were concordant with those associated with increasing chronicity (**Extended Data Figure 7d-i**), and adjusting for PT0 proportion reduced the significance of cell state associations with chronicity. However, PT0-associated cell state shifts

remained significant even after correcting for chronicity (**Extended Data Figure 7c**). For example, Glomerular shifts strongly associated with chronicity ($p<2e-05$) became less significant after adjusting for PT0 ($p<1e-2$), while Glomerular shifts associated with PT0 remained significant even after adjusting for chronicity ($p<1e-05$ vs $p<7e-05$). Accordingly, we suspect that proximal tubule injury, marked by PT0, is a proxy for the cellular effects of higher chronicity index and may better reflect tissue damage than chronicity index in LN.

Activity is associated with expansion of myeloid and glomerular subpopulations

The activity index is a dynamic pathological measure that reflects inflammation and reversible damage in the kidney[50], and cell states associated with this index may represent therapeutic targets. To identify cell state shifts associated with activity, we adjusted for first biopsy status, recruitment site, and chronicity due to the broad effects of these variables on cell states (**Extended Data Figure 5b**). Across cell types, we observed significant shifts ($p<4.5e-04$), driven by shifts in the myeloid ($p<1e-5$) and glomerular ($p<2e-5$) populations (**Figure 6a**). Differential gene expression further revealed the largest number of genes associated with activity index in these same compartments (**Supplementary Table 3**).

Although activity index correlates with ISN class and prednisone use (**Extended Data Figure 8a**) and has been associated with treatment response in previous studies[51], adjusting for these variables did not obviate associations (**Extended Data Figure 8b-c**), suggesting that activity independently drives these cell state shifts.

Among glomerular cells, cell states associated with increasing activity index were similar to those associated with chronicity (**Extended Data Figure 8g,** R=0.487), including strong expansion of SDK1^{high}HDAC9^{high} PEC (GLOM1) (**Figure 6b-c**). GLOM1 displayed a transcriptional program consistent with activated, injury-associated PECs, characterized by the expression of genes involved in extracellular matrix production (FN1, COL8A1, COL6A2,

LOXL1), epithelial-to-mesenchymal transition (SOX4, CDH6), cytoskeletal remodeling (ACTN1, TPM1), and matrix degradation/remodeling (MMP7, TNC); reflecting a pro-fibrotic phenotype observed in glomerular injury and crescentic glomerulonephritis[10,56,101,102]. Notably, GLOM1 selectively expressed TNC (Tenascin C), an extracellular matrix glycoprotein that is transiently induced during tissue injury and repair [doi:10.1002/path.1415]. We have previously shown that urinary Tenascin C levels correlate with histological activity—and, to a lesser extent, chronicity—in patients with LN[66], mirroring the associations observed here.

Differential expression in glomerular cells revealed enrichment of genes associated with ECM remodeling and fibrosis (*FBN1*, *PLOD2*) and inflammation and interferon signaling (*SPP1*, *STAT2*) (**Extended Data Figure 8h**). *MFGE8*, previously implicated in SLE disease activity and risk[54,55], was also enriched. All are markers of the activity-associated cluster GLOM1.

CNA expands on differential expression, showing expansion of tissue-specific macrophages (**Figure 6d-e**), notably *GPNMB*^{high}*NUPR1*^{high} (M5), *GPNMB*^{high}*LYVE1*^{low} (M11), *C1Q*^{low} *SPP1*^{high}, SPP1^{high}FABP5^{high} (M7), and MERTK^{high}FABP5^{high} (M9) (**Figure 6e**). Populations shared with blood, including TPSB2+ MAST cells (M4) and CD16+CXC3CR1+ Monocytes (M0), were contracted. While M5 was also expanded with increasing chronicity, all other activity-expanded populations were not (**Extended Data Figure 8g**). For instance, CCL2+CCL3+ Monocytes (M3) and CLEC10Ahigh cDC (DC18) were depleted with increasing chronicity, but expanded with increasing activity. This suggests that activity-associated shifts are primarily driven by tissue-specific macrophages with effects independent of chronicity.

CD9 and *GPNMB* were highly correlated (R>0.7) with CNA activity associations in myeloid cells, were differentially expressed with activity index, and were highly expressed in M5 (**Extended Data Fig. 8h**). Cells invading the glomeruli in patients with focal segmental glomerulosclerosis (FSGS) and crescentic glomerulonephritis express CD9 and the prevalence of such cells reflect different degrees of podocyte injury[56]. *GPNMB* (R=0.713) is a known biomarker of SLE previously found in the urine and exosomes in a patient with acute LN[57]. Both

genes are known markers for Scar Associated Macrophages (SAMs)[58,59] which were detected in patients with liver or lung fibrosis. Other SAM genes, including *FABP5* (R=0.68), *TREM2* (R=0.63), and *SPP1* (R=0.55) were also highly correlated and were highly expressed in 4 out of 10 macrophage clusters (M5, M7, M9, M11) (**Extended Data Figure 8d**). SAM marker macrophages were somewhat positively correlated with neighborhood correlations with chronicity (**Extended Data Figure 8e)**, but very strongly correlated with activity (**Figure 6f**). In order to test whether the expansion of M5 or other SAMs was driven by tissue-resident populations, we developed a proliferation score from 93 cell cycle genes but found no evidence of proliferation (**Extended Data Figure 8f**).

Additionally, differential gene expression showed activity index was associated with lipid metabolism (*APOC1* - enriched in M5/M11, *FABP5*, *PLIN2* - enriched in M5) and phagocytic/lysosomal activity genes (*TREM2* - enriched in M5/M7/M9/M11, *CTSD* - enriched in M5, *CTSB* - enriched in M5/M7/M9). Cytoskeleton and motility-related genes (*TUBB6* - enriched in M7, *VIM*) were also enriched with respect to activity index.

Activity Associations are Independent of ISN Classification

We were interested to see if these activity changes could be better explained as a difference between membranous vs proliferative nephritis since membranous nephritis, that is ISN class V, is clinically defined as having low activity.

Class V was significantly associated with Myeloid ($p<9.99e-05$) and glomerular ($p=0.0029$) populations (**Extended Data Figure 9a-b**), even after adjusting for first biopsy status, chronicity, and collection site. In both compartments, populations expanded in ISN Class V were contracted with higher activity. Expanded populations included TPSB2+ MAST cells (M4) and CD16+CXC3CR1+ Monocytes (M0), while contracted populations includedSPP1lowFABP5high (M8), GPNMBhighLYVE1low (M11), SELENOPinterLYVE1inter Resident (M10), SPP1highFABP5high (M7), and MERTKhighFABP5high (M9) macrophages

(**Extended Data Figure 9a**). In the Glomeruli, SDK1^{high}HDAC9^{high} PEC (GLOM1) were also contracted (**Extended Data Figure 9b**). Results were unchanged after adjustment for prednisone use and Black race, both of which correlated with ISN Class V (**Extended Data Figure 5d**, **Extended Data Figure 9c**).

However, after adjusting for activity index, glomerular associations were no longer globally significant, and myeloid associations were much less significant. While activity-associated CNA results remained consistent after correcting for ISN V, ISN V associations did not (**Extended Data Figure 9d-e**). These findings suggest that cell state shifts observed in Class V can be explained by its reduced activity index.

Clinical Phenotype Associations with Immune Cells from Blood

Given the easier access to blood than kidney tissue, we examined blood cell state changes associated with case-control status, chronicity, and activity indices. After correcting for age and sex, we identified cell state shifts in B ($p<7e-5$), monocytes and DCs ($p<3e-4$), and T/NK cells ($p<5e-5$) (**Extended Data Figure 10a-c**). LN cases had expansion of ISG-high monocytes (bl-M7), ISG-high naive B cells (bl-B4), and ISG-high T cells (bl-T17), corroborating previous reports of interferon-inducible gene upregulation in LN blood[60]. Dendritic cells (bl-DC6, bl-DC9, bl-DC12) were contracted, consistent with previous studies of SLE blood[61,62]. ABCs (bl-B2) were expanded, aligning with prior human patient and murine lupus models[34,63]. CXCR5^{high} Naive (bl-B0) and CRIP1+ Mature B cells (bl-B1) were contracted. The T/NK compartment showed reduced CD4+IL7R^{high}VIM^{high} (bl-T3), CD4+ Central Memory/Naive (bl-T10), and CD8+GZMK+ TEMRA T cells (bl-T7), with strong depletion of TRGC1+ Gamma/Delta T cells (bl-T12). Differential expression further showed enrichment of interferon-stimulated genes and depletion of ribosomal genes in LN (**Extended Data Figure 10d**).

Given the strong association between renal pathology indices and tissue cell states, we examined whether there were corresponding associations with blood cell states. After adjusting for age and sex, we observed only modest associations with chronicity in Myeloid (p=0.013) and T/NK cells (p=0.0019) (**Supplementary Table 7**). In Myeloid cells, few neighborhoods (n=46/106,436) were significant (FDR<0.10). In T/NK cells, CD8+GZMK+ TEMRA T cells (bl-T7) were expanded with increasing chronicity but were depleted in cases over controls (**Extended Data Figure 10e**). Notably, after correcting for age, sex, there were no significant associations with activity. The limited activity index associations in blood aligns with a recent immunophenotyping study on the same cohort, where associations were confined to B cells[64].

Similarly, differential expression analyses revealed only weak gene associations with chronicity index (**Extended Data Fig. 10f-g**), including downregulation of interferon-stimulated genes in T/NK and B/Plasma cells. No strong gene associations with activity index were observed. Together, these results suggest that, while blood signals may help identify SLE cases and modestly reflect chronicity, blood cell state populations are unlikely to reliably capture renal disease activity.

Discussion

Clinicians have long recognized the heterogeneity of clinical manifestations in LN and the clinical value of renal biopsy. Here, we present a large cross-sectional study comparing kidney biopsies from LN patients and healthy controls, with matched PBMCs for many samples. We included both single-cell and single-nucleus data to robustly define tissue cell states independent of technical effects. This publicly available dataset is a valuable resource for future integration efforts into lupus and renal biology. Our goal was to characterize key cell states in renal tissue and blood, and how these states shift across pathological subsets. We also assessed whether those tissue state shifts manifested in the circulation.

We identified a spectrum of immune, epithelial, and stromal populations relevant to LN, including autoimmune-associated B cells in both blood and tissue (bl-B2, B5), consistent with previous findings in SLE[65], as well as Proximal Tubule epithelial cells with a degenerative injury profile (PT0). We also observed tissue-specific immune states such as GZMK+CD8+ T cells, , GPNMB^{high} macrophages (M5, M11), and SPP1^{high}FABP5^{high} macrophages (M7). To characterize how these populations changed with clinical subphenotypes, we adjusted for potential confounders such as recruitment sites and whether the biopsy was the first one obtained from a patient, since subsequent biopsies often reflect greater damage due to a greater disease burden. Our results include key observations related to LN that will inform future studies and guide therapeutic decisions.

The high interferon signature in SLE has driven drug development for both LN and non-LN manifestations. While its link to SLE in blood is well established[44], associations with clinical disease activity and tissue inflammation are less well understood. In our data, interferon signatures were clearly enriched in LN compared to controls, across blood and tissue. However, connections between interferon signature and tissue inflammation states were more subtle, with interferon signature showing a modest decline with increasing chronicity in some cell types, and no association with activity.

We observed modest shifts, such as expansion of GZMK+CD8+ T cells in both blood and tissue with increasing chronicity. However, overall, blood expression and cell states were limited associated with respect to chronicity or activity. This does not preclude the possibility that alternative profiling technologies may reveal cellular or serum biomarkers reflecting LN renal heterogeneity[10]. In contrast, non-invasive biomarkers, including macrophage-derived proteins in urine, have been associated with proliferative disease and decrease in treatment response[66].

We were particularly interested in the activity index, which reflects ongoing renal inflammation and is thought to be modifiable with therapies. Even brief reductions in activity index has been associated with reduced organ damage[68]. After adjusting for the substantive

effects of chronicity, we observed that key, tissue-specific cell states within the myeloid compartment increased with activity, including GPNMB^{high} (M5, M11)[57] and SPP1^{high}FABP5^{high} (M7) macrophages. These populations expressed a canonical *SPP1*, *CD9*, *GPNMB*, *FABP5*, and *TREM2* macrophage gene profile, known to occur across multiple chronic inflammatory diseases, often adjacent to fibrotic regions[69–71]. While these macrophages are considered reparative, they may acquire pro-fibrotic features over time[72–74]. The persistence of the M5 subset in association with both disease activity and chronicity may suggest a pathogenic role. Although this data does not definitively establish whether these states drive disease activity, mediate tissue repair, or promote fibrosis, it highlights distinct, disease-associated states with therapeutic potential. Furthermore, while we suggest that SAMs may proliferate in the tissue, past work in the brain and liver have both suggested that SAMs may have different origins based on their subset, with some being derived from infiltrating monocytes and others differentiating from tissue-resident macrophage populations[75,76]. Current SLE therapies primarily target B and T cells[5,77]. While B cell targeting has been effective, our study was underpowered to detect activity associations with B cell states. These data support the potential for targeting myeloid cells in LN in conjunction with B cells.

Renal biopsy chronicity index scores are thought to reflect cumulative, irreversible kidney damage from ongoing disease. Consistent with this, we observed that chronicity closely tracked with a shift from healthy to injured proximal tubule cells. These injured epithelial cells show cellular dedifferentiation, senescence, and a canonical transcriptional signature linked to tubulointerstitial fibrosis and chronic kidney disease (CKD) progression across diverse renal diseases[67]. These differences may not be specific to LN. Our results suggest that these changes may very closely reflect the level of chronic kidney injury, and may serve as a molecular proxy for chronicity.

Within renal tissue, chronicity had the most dramatic impact on cell types and states, including immune populations, compared to other clinical parameters. Notably, with increasing

chronicity, SPP1^{high}FABP5^{high} macrophages (M7) and GZMK+ T cells were greatly expanded within the myeloid and T cell compartments, respectively. Most immune cell states that expand with increasing chronicity are tissue-specific and largely absent in the blood, including GPNMB^{high}NUPR1^{high} Macrophage (M5) and a CD4+ Effector Memory T cell state (T11). These findings demonstrate the importance of accounting for chronicity in molecular studies of LN. For instance, associations between immune and tissue cell states and treatment response disappear after adjusting for chronicity, a well-known predictor of response[17].

A key finding is that differences in renal tissue cell states are better explained by activity and chronicity indices than by ISN classes, which are commonly used for informing therapies. For instance, ISN class V patients by definition have a decreased activity index (which is focused on lesions specific to class III and IV), it is the low activity index index and not the ISN class that explains the cell state shifts.

This study has several limitations. B cells were relatively rare in the tissue for our study, despite their relevance as SLE drug targets. So while we could identify key B cell states in tissue and blood, we did not have the power to identify cell state shifts with disease and tissue state. It is possible that B cells are pathogenically active in other tissues, such as lymph nodes, and are less active in LN kidneys. Additionally, renal immune cell states were defined transcriptionally, without surface protein markers, which can improve immune cell state annotation accuracy[78]. Our study design was also highly heterogeneous, including a large number of samples from a range of recruitment sites but also individuals at different disease and biopsy points, and with different treatment histories. While the diversity of the cohort is a strength, its heterogeneity complicates disentangling medications and cohort variables on biopsy metrics. Longitudinal followup samples were also not performed. Finally, while our tissue-specificity metric (TSM) identifies tissue and blood resident populations, it cannot distinguish whether tissue specificity is due to migration from blood or development within the tissue.

In conclusion, this comprehensive study defines the key tissue and blood cell states in SLE, and demonstrates how they shift with the development of LN, and within the renal tissue of LN cases. This useful reference will serve as a key resource in understanding the pathogenesis of LN, and devising therapies to mitigate its long-term effects on patients.

Code availability

The code for analyzing this dataset can be found on our GitHub: https://github.com/immunogenomics/AMP_SLE_Kidney.

Data availability

The processed count matrices are available on Synapse under accession code syn64064827. Interactive visualization of marker genes in our data is available on our website: https://immunogenomics.io/ampsle2/app/?ds=b_pbmcs.

Methods

Patient Recruitment and Biopsy

We have described the details of enrollment in the AMP SLE cohort elsewhere[17,79,80]. In brief, patients with lupus undergoing kidney biopsies as part of standard of care were eligible to enroll in the prospective AMP LN study. The decision to biopsy was at the discretion of the treating physician to confirm suspected LN *de novo*, assess persistent disease activity not responding to treatment, or diagnose disease recurrence. Inclusion in AMP required: 1) age ≥ 18; 2) fulfilment of the revised American College of Rheumatology[81] or the Systemic Lupus Erythematosus International Cooperating Clinics[82] classification criteria for SLE; 3) a urine protein/creatinine ratio (UPCR) >0.5 at the time of biopsy. This study began prior to the publication of the ACR/EULAR criteria. Analyses regarding responder status were restricted to patients with baseline random or 24-hour Urine Protein Creatine Ratio (UPCR) ≥ 1.0 since for

patients with UPCRs between 0.5 and 0.999, we did not define proteinuria response. We only considered patients with renal biopsies that demonstrated International Society of Nephrology/Renal Pathology Society (ISN/RPS) classes III, IV, V or combined III or IV with V read in this analysis[14,83]. Where possible, ISN Class and histopathological indices were sourced from the central pathologist scoring. For patients without a central read, we used the read from the pathologist at the recruitment site. Exclusion criteria included: 1) a history of kidney transplant; 2) rituximab treatment within 6 months of biopsy; 3) pregnancy at the time of biopsy. Informed consent was obtained from all participants, and the study protocol was approved by the institutional review boards and ethics committees of participating sites in adherence with the Declaration of Helsinki.

Baseline demographics from a predetermined set of categories, including self-reported race (Asian, Black, White, Other)/ethnicity (Hispanic, non-Hispanic) as required for NIH-funded studies, and clinical characteristics were recorded at the time of biopsy. Laboratory tests and medications were documented at each visit (baseline, week 12, week 26 and week 52) and were performed at the participating sites. For steroids, the higher dose at either baseline or week 12 was considered the induction dose for similar reasons. Pulse steroids were also captured separately.

Outcomes

Complete response (CR) required: 1) UPCR<0.5; and 2) normal creatinine (≤ 1.3 mg/dL) or, if abnormal, ≤ 125% of baseline; and 3) prednisone ≤ 10 mg/day at the time of the study visit. Partial response required: 1) >50% reduction in UPCR; and 2) normal creatinine (≤ 1.3 mg/dL) or, if abnormal, ≤ 125% of baseline; and 3) prednisone dose ≤ 15 mg/day at the time of the study visit. Patients who did not achieve a CR or PR at the specific timepoints were considered non-responders (NR) or not determined (ND) if data were missing. These response definitions were based on the ACCESS Trial [7] In agreement with the ACCESS trial, we specifically

decided not to include the microscopic review of the urine sediment given the absence of uniformity across sites in in assessing urinary sediment and the challenge of attribution especially in a population of young women. The prednisone threshold for CR at ≤ 10 mg prednisone was also based on the ACCESS trial. However, the ≤ 15 mg prednisone maximum for defining PR was agreed upon unanimously by the site investigators.

Although proteinuria was measured by either a UPCR on a spot urine or a timed urine collection, consistency of the method across the study for an individual was required. While determination from a timed urine collection was preferred, if this method was not performed at all time points for an individual participant, calculations from a spot urine were utilized.

Human Kidney Tissue Disassociation

We processed cryopreserved kidney biopsies in batches of 3 following a previously described protocol[44] with modifications. Cryovials containing tissue samples were warmed in a 37 C water bath until partially thawed. The contents of each vial were then decanted into one well of a 24 well plate containing 500uL warm RPMI (ThermoFisher) supplemented with 10% FBS (Avantor) and 10mM HEPES (ThermoFisher). Using tweezers, each biopsy was then transferred to a second well in a 24 well plate containing 2mL warm RPMI/FBS/HEPES and allowed to rest for 1 minute, after which it was placed in 100uL of RPMI supplemented with 10 mM HEPES and 0.04% BSA (Millipore Sigma) in a Petri dish. The biopsies were cut into segments approximately 2 mm in length using a clean razor blade. For biopsies analyzed by single nucleus RNA-seq, two noncontiguous segments were removed for nuclei isolation. The remainder were enzymatically dissociated as follows. Biopsy segments, along with the 100uL RPMI/BSA/HEPES, were transferred to a 1.5 mL conical tube containing DMEM/F12 (Corning) and 0.5 mg/mL Liberase TL (Roche) pre-warmed to 37 C. Samples in digestion medium were then agitated for 6 minutes using an orbital shaker set to 300 rpm and 37 C. Samples were then

gently mixed 5 times by pipetting with a large-bore tip (Rainin), after which they were returned to the shaker for a second 6 minute 300 rpm/37 C incubation. After the second incubation, samples were placed on ice. 600 uL cold RPMI/FBS/HEPES was added to stop digestion. The remaining tissue fragments were gently resuspended, and then decanted onto a 70 um strainer (Miltenyi) sitting on a 15 mL conical tube on ice. 4 mL cold RPMI/BSA/HEPES was used to rinse the digestion tube and was strained through the 70 um filter to collect single cells released during digestion. The biopsy fragments were then mechanically dissociated by pushing them against the strainer mesh with a cell strainer pestle (CELLTREAT). The pestle and strainer were then rinsed with a further 5 mL cold RPMI/BSA/HEPES to collect dissociated cells. The strained samples were centrifuged at 200 rcf for 8 minutes at 4C, after which cell pellets were resuspended in 50 uL cold RPMI/BSA/HEPES and filtered again through a 40 um strainer (Pluriselect). Cells in the resulting suspension were counted by trypan blue exclusion.

Single cell RNA-seq Library Preparation and Sequencing

Single cell RNAseq was performed on the Chromium platform (10x Genomics), using the single cell gene expression 3' V3 kit. Up to 10,000 live cells from each kidney sample were loaded into separate 10x channels. Complementary DNA amplification and library construction were carried out by the Broad Institute Genomics Platform according to the manufacturer's instructions.

To reduce sequencing requirements, we removed DNA fragments derived from mitochondrial transcripts using CRISPR-Cas9. We modified the DASH[84] method for compatibility with 10x libraries. We selected a panel of 61 guide RNA sequences (supp table) to tile regions of the mitochondrial genome heavily represented in 3' 10x libraries from kidney samples (supp figure?) using on-target and off-target predictions from the Broad Institute Genetic Perturbation Platform algorithm[85,86]. T7 reverse transcriptase template was generated from oligonucleotides

(IDT) as previously described[84]. Transcription was performed with 100U T7 RNA polymerase (New England Biolabs) in a 30 uL reaction containing 1X RNAPol Reaction Buffer (NEB), 75ng template DNA, and 0.83 mM each of rATP, rCTP, rGTP and rUTP (NEB). T7 reactions were incubated for 4 hours at 37C. Template DNA was then removed by digestion with 4U DNAse I (NEB) for 15 minutes at 37C. RNA was then purified using a Monarch® RNA Cleanup Kit (NEB), quantified by absorbance of 260 nm light, and aliquoted prior to storage at -80C. Cas9:sgRNA complexes were formed by mixing 1.5 pmoles Cas9 nuclease (NEB) with 1.5 pmoles of the sgRNA pool in a 15 uL solution containing 2uL 10 X NEBuffer™ r3.1 (NEB) and incubating at 25C for 15 minutes. 10ng of individual samples of sequencing-ready 10x gene expression libraries in 5uL water were then added to the ribonucleoprotein complexes and incubated for 2h at 37C to make double strand breaks.

To enable direct amplification of the digested libraries[87], the samples were then incubated with 1ug RNAseA (Ambion) and 0.8U Proteinase K (NEB) for 15 minutes at 37C to remove RNP complexes, followed by a second incubation for 15 minutes at 95C to inactivate the enzymes. The libraries were then amplified by 6 PCR cycles using P5/P7 oligonucleotides (AATGATACGGCGACCACCGA, CAAGCAGAAGACGGCATACGA) and KAPA HiFi HotStart ReadyMix (Roche). This step enriched the samples for intact uncut DNA fragments with non-mitochondrial inserts. The PCR products were purified using SPRIselect paramagnetic beads (Beckman Coulter) at a 0.8x volumetric ratio. Finally, the purified libraries were quantified and sequenced according to manufacturer's guidelines on a Novaseq S4 (Illumina).

Single nuclear RNA-seq Library Preparation and Sequencing

Nuclei were isolated from a subset of the cryopreserved human kidney biopsies used for single cell RNA-seq using a protocol previously described[88], with modifications. We found that ~4mm from each 18G biopsy provided enough cells or nuclei for transcriptional profiling.

Therefore we selected biopsies that were >7mm in length to perform both single cell and single nucleus sequencing. Two non-contiguous ~2mm sections of each biopsy were placed in a 2mL dounce tissue grinder (Kimble) containing cold lysis buffer (20mM Tris pH7.5, 1M sucrose, 5mM CaCl2, 3mM MgCl2, 1mM EDTA, 0.1% Triton-X100). Biopsy fragments were homogenized with 5 strokes using pestle A and 20 strokes with pestle B, then incubated for 10 minutes on ice. The sample was then strained through a 30um mesh (Celltrix) and diluted with 6mL cold PBS with 1mM EGTA. Nuclei were then pelleted by centrifugation at 500g for 5 minutes at 4C. Supernatant was removed, and the pellet resuspended in 100uL cold PBS + 1% BSA (manufacturer) and a 1:1,000 dilution of RNase inhibitor (Takara). The resulting nuclei suspension was then strained through a 10um mesh (Pluriselect) before loading up to 8,000 nuclei for gene expression profiling using the Chromium 3' V3 kit (10x genomics).

PBMC Collection and Processing

Blood was collected in heparin tubes and processed within 2 hours. PBMC were isolated by centrifugation through a density gradient medium in 50mL SepMate tubes (Stemcell Technologies) following manufacturer guidelines, and then cryopreserved in CryoStor CS10 (Stemcell Technologies). Frozen samples were shipped to Oklahoma Medical Research Foundation Biorepository for centralized storage. Samples were randomized and shipped frozen to the Broad Institute where they were processed in batches of up to 16 to generate single cell RNAseq data. Samples were thawed at 37 degrees C, then diluted in cold RPMI 1640 (Gibco) supplemented with 0.5% BSA (Miltenyi Biotec), 1X Glutamax (Gibco), and 10mM HEPES (Corning). Samples were thereafter maintained on ice or at 4 degrees C whenever possible. Cells were filtered through a 40um strainer (Pluriselect), then pelleted by centrifugation. Pellets were resuspended in autoMACS Running Buffer (Miltenyi Biotec) containing human FcR Blocking Reagent (Miltenyi Biotec) and fluorescent antibodies (Biolegend) targeting CD15 (clone W6D3 conjugated to BV510) and CD45 (clone HI30 conjugated to one of: FITC, PE,

PE/Cy7, AF700), then incubated for 20 minutes on ice. Samples were then washed, and live cells were counted using ViaStain AOPI Solution (Nexcelom Bioscience) and a Cellaca MX instrument (Nexcelom Bioscience). Pools of 4 samples were created, with the CD45 fluorochrome unique to a single sample in each pool. Each pool contained 600,000 live cells. Pools were incubated for 30 minutes at 4 degrees C in a volume of 50uL Cell Staining Buffer (Biolegend) containing oligo barcode-tagged antibodies for surface protein detection (Totalseq-A Human Universal Cocktail, V1.0 from Biolegend; ¼ test). Pools were then washed twice and resuspended in autoMACS Running Buffer containing 1mg/mL DAPI (Biolegend) before filtration through a 40um strainer. Equal numbers of live, CD45 positive, CD15 negative cells from each PBMC donor were FACS sorted (Sony MA900) into RPMI 1640 (Gibco) supplemented with 0.04% BSA (Miltenyi Biotec), 1X Glutamax (Gibco), and 10mM HEPES (Corning). Cells derived from each donor in the pool were distinguished by the CD45 antibody fluorochrome. After sorting, all samples in the batch were pooled, and 32,000 cells were loaded in each of 3 Chromium chips (10x Genomics) to generate single cell gene expression libraries using manufacturer's protocols. The libraries were sequenced on Novaseq S4 flow cells (Illumina) to an average depth of 60,000 reads per cell.

Processing and Quality Control for Kidney Transcriptomics Data

We converted raw sequencing reads to count matrices with cell ranger (version 5.0.1), first converting sequencing output to FASTQ files using the cellranger mkfastq function, We aligned reads to GRCh38 and generated count matrices with cellranger count. Quality control for the scRNAseq and snRNAseq data involved several common steps in single cell/nuclear transcriptomic sequencing analysis. Initially, we filtered cells removing cells and nuclei with less than 1000 reads per cell or nucleus, and less than 500 unique genes detected. Since we employed DASH to reduce mitochondrial read contribution to the sequencing, we could not define mitochondrial read content per cell or nucleus with the entire list of mitochondrial genes.

However, we used two non-dashed mitochondrial genes: *MT-ND5* and *MT-ND6* to estimate the mitochondrial content per cell and nucleus. We removed cells with more than 3% read contribution total reads by these two mitochondrial genes. For single nuclei, we used a stricter threshold and removed nuclei with higher than 1% contribution to total reads by these two mitochondrial genes.

Defining Broad Renal Cell Types and Doublet Removal

We employed standard scRNAseq and snRNAseq preprocessing steps. We first transformed the raw counts matrix by multiplying with a factor of 10^4 and log normalizing the count. We then scaled the raw counts to have mean zero and variance 1. We were concerned that the list of variable genes would be dominated by the scRNAseq modality, which had many more samples. Additionally, we did not want to obtain variable genes from samples with low cell or nuclear count. We therefore selected the top 500 most variable genes based on dispersion from each sample with more than 25 cells or nuclei within each technology. The final list of variable genes was obtained by taking the union of these sets obtained from these patients. We were additionally concerned about uneven contributions to low dimensional embeddings from each modality due to imbalanced sample numbers. We therefore adapted the weighted PCA approach described in Zhang. et al[89] to balance the contribution from single cell and single nuclear cells when computing principal components. We then applied harmony[24] to these principal components to correct for non-biological variation from processing batch, sample collection site, sequencing modality, and samples.

Using the harmonized principal components, we created reduced dimensional representations for visualization using the UMAP algorithm[90]. We additionally employed density-based Louvain clustering on fuzzy shared nearest neighbor graphs computed using the harmonized principal components. We clustered at low resolution to obtain broad cell type

groupings. We were concerned about the possibility of doublets formed by multiple cell types, so we employed Scrublet[91] to identify possible sources of doublets. After manual inspection, we removed clusters of cells expressing markers for multiple cell types and possessing a qualitatively higher-than-average Scrublet score. After manual inspection, we also removed clusters of low-quality cells, marked by higher-than-average mitochondrial content and a lower number of genes and reads per cell.

To define major immune and tissue cell types, we employed markers identified by several previous works, including the Kidney Precision Medicine Project[30], the AMP Phase II Rheumatoid Arthritis publication[92], and the AMP Phase I LN publication[43]. Broadly we employed known markers for immune cells, including: *CD3, CD4,* and *CD8* for T cells, *CD56* and *TYROBP* for NK cells, *CD14* and *CD16* for macrophages, and *XCR1, SDC1, MS4A1,* for B and Plasma cells. For tissue cell types, we used kidney specific marker genes such as: *CUBN* and *SLC34A1* for proximal tubules, *SLC12A1* and *CRYAB* for cells from the loop of henle, *CLNK, SLC12A3, AQP2,* and *SLC8A1* for cells from the distal nephron, *PECAM1, COL1A1,* and *NOTCH3* for cells from the Interstitial/Stromal compartment, and *CFH* and *PODXL* from cells from the glomerulus.

Defining Fine grain Renal Cell Tissue and Immune Cell States

Within each cell type identified (see above), we employed a fine grain clustering approach consistent with the methods used in broad cell type identification. Briefly, we normalized raw counts, scaled genes to have a mean of zero and variance of one across cells, and employed a weighted PCA approach within each cell type to balance principal component contribution from each modality. We applied an additional level of quality control within each cell type and removed cells of exceptionally low quality (high mitochondrial count and low total read and unique gene counts). We additionally again removed doublets expressing known marker genes from other cell types with a high scrublet score. We identified fine grain cell states using several clustering at high resolution, and identified cell state specific genes using the presto

package (https://www.biorxiv.org/content/10.1101/653253v1.abstract). We determined Immune cell subtypes using marker gene lists from the AMP Phase I LN study, the AMP Phase I and II Rheumatoid Arthritis studies, the KPMP project, and a prior in-depth study of memory T cells in tuberculosis patients[93]. We determined fine grain renal cell types using marker genes sourced primarily from the KPMP project, and supplemented cell state markers using the AMP Phase II Rheumatoid arthritis study as well as sources describing the proximal tubule injury molecular phenotype[29,94].

PBMC processing, quality control, and profiling

We processed PBMCs from raw 10X read data with cellranger pipeline (version XXX). Briefly, we generated fastq files from the bcl file output from the 10X sequencer using the cellranger makefastq function and generated bam files from fastq files using the cellranger count function. We employed cellranger version 6.1.1 for this preprocessing. PBMC samples were not hashed for sample of origin and required computational matching. We employed demuxlet[95] to match samples to donors using genotyping information collected at time of enrollment to the AMP study. We removed cells which could not be attributed to a single donor confidently via demuxlet's functionalities. Next, we imported the data into the Seurat package, and using Seurat quality control functionalities we removed: cells with less than 500 genes per cell and over 20% contribution to total reads from mitochondrial reads.

We then normalized and scaled the PBMC dataset using the Seurat pipeline's inbuilt functions and performed PCA. We used the Seurat implementation of harmony to mitigate non-biological effects from the batch of origin and sample. We employed the Seurat implementations of the UMAP algorithm for low dimensional visualization and clustered the harmonized pcs using Louvain clustering at low resolution. We further employed Scrublet to identify inter cell-type doublets and removed clusters with high scrublet scores and expression

of marker genes from multiple cell types. We repeated the PCA, harmonization, and clustering procedures post doublet removal.

We used commonly identified markers of broad immune cell types to define PBMC transcriptomic profiles post clustering in the whole PBMC space. We used *CD4, CD8*, *and CD3* for T Cellls, *CD56* and *TYROBP* for NK Cells, *CD14, CD16,* and *CCR2/L2* for monocytes. We further used *MS4A1* and *CD19* for B cell identification. We sourced markers for cell states from a previously described study of memory t cells in tuberculosis patients, the AMP Phase I LN and RA studies, the AMP Phase II RA study, and the KPMP project. Within cell types, we applied the same strategy used for tissue cells to eliminate likely doublet and low-quality clusters, and defined cell states with the sources and software packages described above.

Defining PBMC Fine Grain Cell States

We processed data within each PBMC each cell types through the steps described above for PBMC profiling. We performed an additional round of doublet removal within each cell type via a combination of high-resolution Louvain clustering, analysis of Scrublet score distribution across clusters, and identification of cluster-specific genes from other cell types. We applied log-normalization to read counts, filtered genes consistently variable in both sc and sn data, and then scaled the data. We then applied Harmony[24] to principal components weighted by sc to sn ratio to correct for processing batch, site, sample, and differences in sc and sn technologies. We then regenerated harmonized principal components and employed Louvain clustering at several different resolutions. We defined fine grain cell states in T, B, and myeloid cells using markers sourced from the previously described tuberculosis memory t cell study, the AMP Phase 1 SLE study, the AMP Phase 1 and 2 RA studies, Villani et. al[96], and Perez et. al[97]. These markers were matched with clusters via identification of cluster specific gene expression sourced from the presto package (https://www.biorxiv.org/content/10.1101/653253v1.abstract).

Blood-Tissue Immune Integration

Tissue and blood immune cells were integrated into a single embedding using weighted PCA to account for the larger amount of tissue-origin immune cells than PBMC. Variable genes were found independently between the tissue and the blood and the union was used to compute the principal components. Further batch correction was done through Harmony with theta=0 with a maximum of 50 iterations to mitigate non-biological variation arising from factors such as individual sample differences, processing batches, and the original source of the cells (tissue or blood). Only samples present in the blood and tissue were taken.

The Tissue Specificity Metric (TSM) was calculated per-cell on the basis that, for a given cell, what fraction of its 50 nearest neighbors are from the tissue over the total number of cells. The numerator of the TSM calculation represents the number of these 50 nearest neighbors that originate from the same tissue as the index cell. The denominator of the TSM equation is designed to normalize this count using a weighted factor that accounts for the inherent differences in the abundance of different cell types, particularly the often disproportionately high number of immune cells present in blood compared to tissue cells. A high TSM value for a given cell indicates that its local microenvironment is predominantly composed of cells from the same tissue, suggesting a strong tissue-specific identity. Conversely, a low TSM value suggests that the cell's neighborhood is more blood-like in nature. A TSM score of 0.5 indicates that a cell's neighborhood reflects what the expected proportion of blood to tissue cells should equal mixing occur. The TSM can be given through the equation:

$$tissueScore = \frac{\#\ of\ tissueNeighbors}{\#\ of\ tissueNeighbors + \frac{Total\ \#\ of\ tissue\ cells}{Total\ \#\ of\ blood\ cells}(\#\ of\ bloodNeighbors)}$$

Case-Control Statistical Associations

To understand cell states enriched and depleted in cases, we utilized Covarying Neighborhood Analysis (CNA)[24]. In brief, CNA divides a given embedding space into granular transcriptional neighborhoods, quantifies the abundance of these neighborhoods per sample, and assesses if the abundance of neighborhoods is associated with sample-level phenotypes. We utilized harmonized principal components to generate CNA neighborhoods. CNA produces a global significance p-value indicating that some neighborhoods are non-randomly associated with the variable being tested. If that value was significant, we visualized CNA outputs by highlighting cells that are at the center of significantly associated neighborhoods (FDR<5%). However, in certain instances where we observed global significance, but did not identify neighborhoods with FDR<5%, we displayed FDR<10% neighborhoods. We applied CNA separately to single-cell and single-nuclear embeddings. When reporting expanded and contracted cell states found by CNA of scRNAseq, we confirmed observed trends with snRNAseq CNA and reported cell states shifts found via both methods. For cell types with less than 10,000 single-nuclei, we confirmed expanded and contracted scRNAseq cell states with MASC and again reported cell state shifts found via both methods. For all PBMC analyses we adjusted for age and sex as covariates.

Within Case Statistical Associations

We wanted to identify the key clinical variables driving single cell phenotypes within our case samples to better understand heterogeneity within SLE tissues. We selected a mixture of demographic and clinical variables: age, activity index, chronicity index, first biopsy status, ISN Class, prednisone use, race, collection site, treatment response, and sex. Across multiple cell types we found associations across multiple cell types with the key clinical variables of activity index, chronicity index, and ISN Class. However, we were concerned about confounding

variables in the within case analysis. For chronicity index, we computed R-squared coefficients with: age, activity index, biopsy status, ISN Classes, prednisone use, race, treatment response, sex, and site using the lm() functionality in R. For activity index, we computed the R-squared coefficients similarly. For both these analyses we transformed multilevel categorical variables with one-hot encoding. We reported adjusted R-squared coefficients via the lm() functionality in R by fitting the continuous outcome to all one-hot encoded levels but one (as a reference level). For ISN Class we transformed class IV and V into one-hot encoded outcome variables. We employed logistic regression and reported the McFadden's pseudo-R-squared. For multilevel categorical variables correlated with Class IV and V ISN status, we conducted multiple logistic regression and reported adjusted McFadden's pseudo R-squared.

To adjust for multiple testing burden, we adjusted correlation p-values with Bonferroni adjustment. However, to ensure robustness of an observed association, we adjusted for variables nominally ($p<0.05$) associated with the variable of interest. Analysis of differential cell-state abundance and differential gene expression was conducted similarly to steps outlined for case-control analysis above. Differential cell-state abundance testing via CNA and MASC, however, was adjusted for covariates identified through the procedure outlined above.

Pseudobulk Differential Gene Expression Analysis

To identify differential genes associated with case-control status, we constructed pseudobulk profiles from single-cell RNA profiles within each cell type[98]. Differential expression analyses were performed separately for tissue and blood single-cell profiles. Pseudo-bulking was conducted via summing the raw gene counts per individual. To reduce false identification of differentially expressed due to sparsity, only genes with non-zero expression in 90% of samples were included. We used two negative binomial models to model the expression of each gene. In the null model for tissue, we modeled gene expression with an intercept and offset for sample library size. In the full model, we performed an identical model with the addition of case-control

status as a covariate. All p-values described were calculated using a likelihood ratio test (LRT) between the full and null model. We further corrected for multiple hypothesis testing by adjusting p-values with the Bonferroni method. Identical models were applied to the blood data, with the addition of covariates for age and sex added to both models.

To test for genes differentially expressed with chronicity index, we constructed similar pseudobulk profiles within only case samples, performed gene filtering as in our case-control analysis, and utilized two negative binomial models to model the expression of each gene. Within tissue, we applied a null model for gene expression with an intercept and offset for sample library size and covariates correcting for first biopsy status and site. In the full model, we applied an identical model with the addition of chronicity index as a covariate. Within blood samples, we applied similar null and full models as in tissue, with additional covariates for age and sex added to both models. P-values were calculated with a LRT and adjusted with Bonferroni correction.

To test for genes differentially expressed with activity index, we also constructed pseudobulk profiles within only case samples, performed gene filtering as in our case-control analysis, and utilized two negative binomial models to model the expression of each gene. Within tissue, we applied a null model for gene expression with an intercept and offset for sample library size and covariates correcting for first biopsy status, site, and chronicity index. In the full model, we applied an identical model with the addition of activity index as a covariate. Within blood samples, we applied similar null and full models as in tissue, with additional covariates for age and sex added to both models. P-values were calculated with a LRT and adjusted with Bonferroni correction.

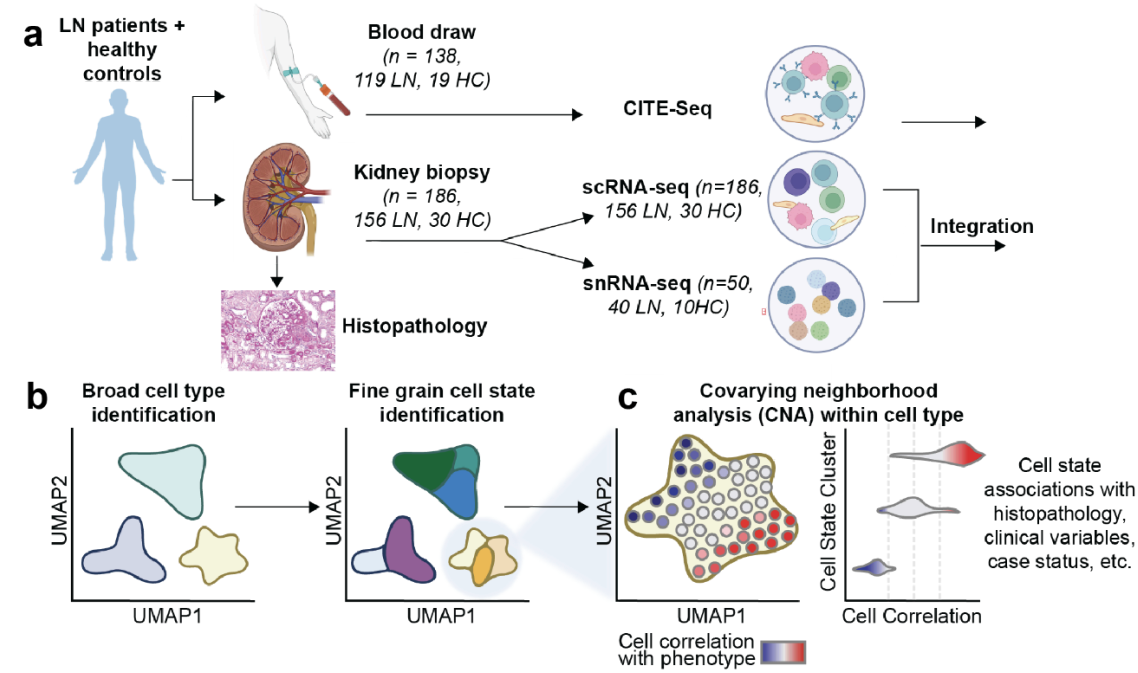
To test for genes differentially expressed between tissue and blood, we subset to only individuals with matched tissue and blood samples. We then constructed pseudobulk profiles from single-cell transcriptomes for each sample. As the sequencing depth between the tissue and blood is very different, we performed stricter quality control prior to modeling and utilized a

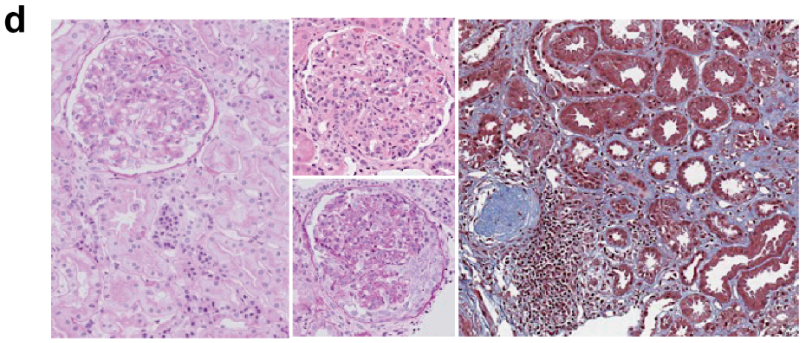
linear model rather than negative binomial. In addition to including only genes with non-zero expression in 90% of samples, we also filtered out any samples with zero expression of any of these common genes. We then performed library size normalization (TPM) and log10+1 transformation of sample counts. We used two linear models for modeling the expression of each gene. In the null model, we modeled gene expression with an intercept and covariate for individual. In the full model, we performed an identical model with the addition of tissue source as a covariate. P-values were calculated with a LRT and adjusted with Bonferroni correction.

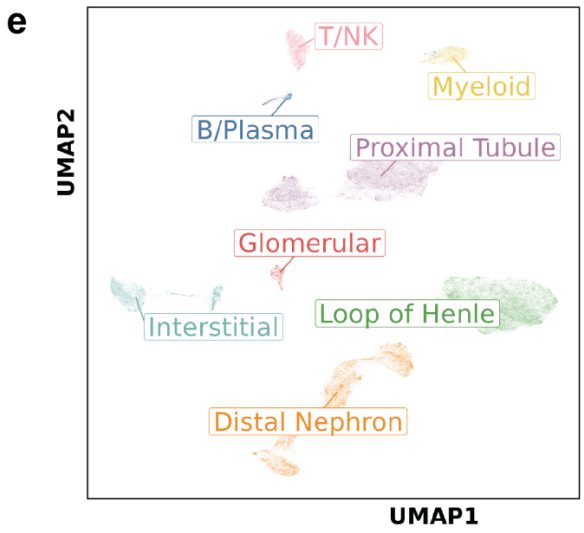
Constructing an Interferon Stimulated Gene Score

We constructed a per-cell score reflecting relative expression of interferon stimulated genes using a previously published gene list[44] and the commonly used method for single-cell scoring with a gene set used in Seurat's AddModuleScore and Scanpy's score_genes functions [99,100]. Briefly, within each cell type for sc tissue data, we performed library size normalization (TPM) and log+1 transformation of the single cell raw counts matrices, where log refers to the natural log. We then performed z-score scaling of each gene and applied Seurat's AddModuleScore() function using a set of known interferon stimulated genes. For calculating sample level correlation with chronicity and activity index, we averaged the calculated score across all immune (B/Plasma, Myeloid, and T/NK) or all tissue (Glomerular, Proximal Tubule, Distal Nephron, Loop of Henle, Interstitial/Stromal) cell types.

Figures / Legends







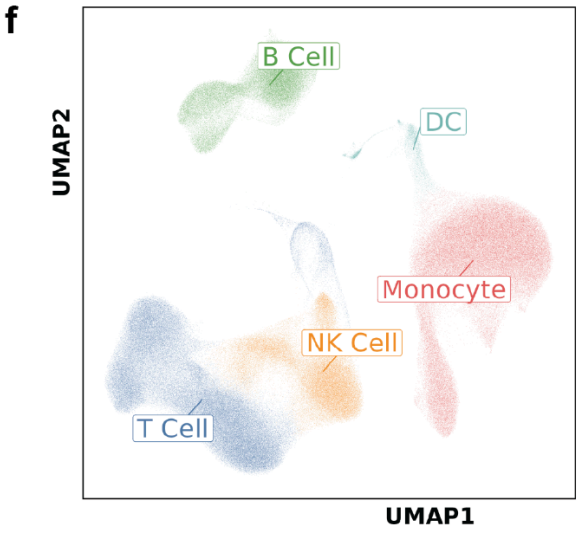
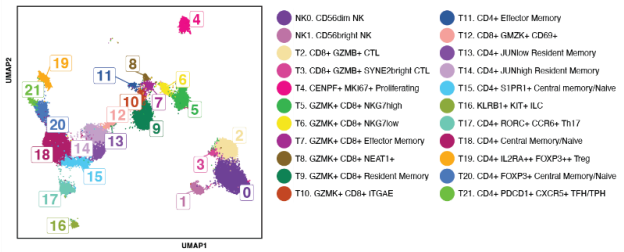
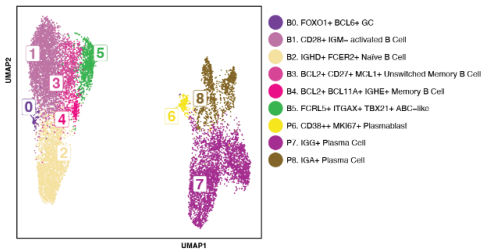


Figure 1. Dataset overview. a-c, Schematic overview of the data collection and analysis pipeline. (**a**) Blood and kidney samples were collected from LN patients and healthy controls (HC). Histopathology and clinical scoring was performed on renal biopsies. CITE-seq was performed on blood samples, and scRNA-seq and snRNA-seq were performed on tissue samples. Both tissue modalities were integrated with Harmony[24]. (**b**) Broad cell types were annotated separately in tissue and blood. Fine grain clustering for cell state annotation was performed on each broad cell type. (**c**) For each tissue, we performed CNA within each cell type to define cell state associations with various clinical variables. **e,** Left: Normal glomerulus and tubulointerstitium (PAS stain). An H&E image of the same region is preferred and will be obtained. Middle: Two glomeruli showing active lesions. Top: H&E stain demonstrating global endocapillary hypercellularity. Bottom: PAS stain showing a cellular crescent.
Right: Low-power trichrome stain highlighting chronic/fibrotic changes in blue, including diffuse interstitial fibrosis, tubular atrophy, and a globally sclerotic glomerulus. **f-g,** UMAPs of all single cell or single nuclear annotations in tissue (**f**) and blood (**g**), colored by broad cell type annotation.

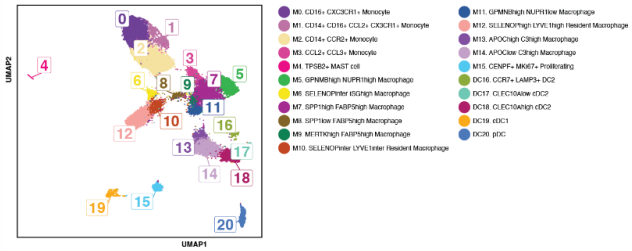
a **T/NK** (*n = 35,131 cells, 2,379 nuclei*)



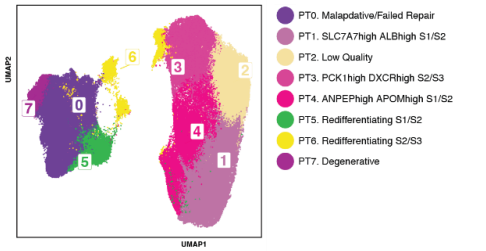
b **B/Plasma** (*n = 12,644 cells, 779 nuclei*)



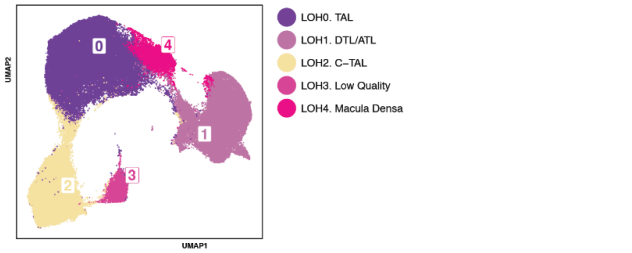
c **Myeloid** (*n = 24,084 cells, 1,494 nuclei*)



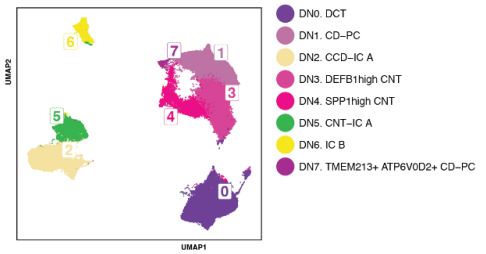
d **PT** (*n = 159,761 cells, 52,464 nuclei*)



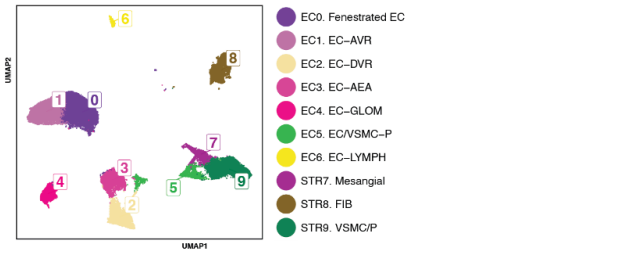
e **LOH** (*n = 130,521 cells, 42,726 nuclei*)



f **DN**(*n = 118,642 cells, 32,466 nuclei*)



g **VSL/STROMAL** (*n = 48,435 cells, 7,237 nuclei*)



h **GLOM** (*n = 8,976 cells, 3,336 nuclei*)

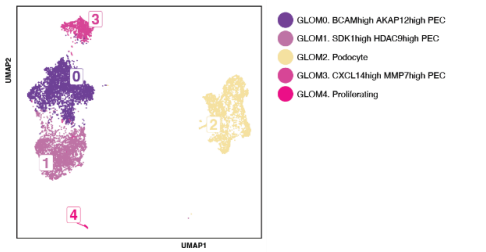


Figure 2. Cell state annotations in tissue. a, UMAP of T/NK cell states in tissue, colored by fine grain cell state annotations. Annotation boxes are located at the centroid (mean) of each cluster. **b-h,** Same as **a** for B/Plasma, Myeloid, Proximal Tubule, Loop of Henle, Distal Nephron, Interstitial/Stromal, and Glomerular cells, respectively.

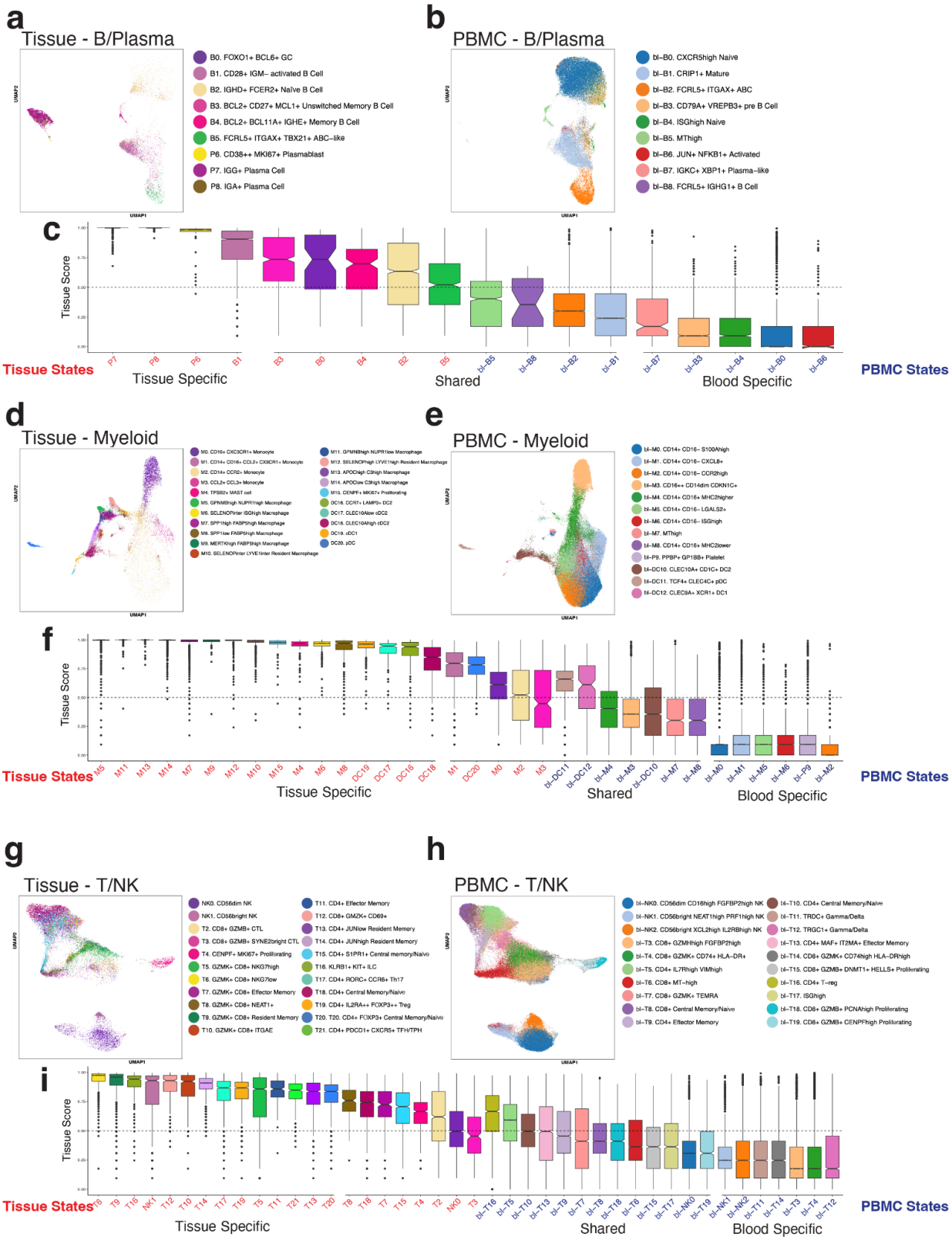

a
Tissue - B/Plasma
B0. FOXO1+ BCL6+ GC
B1. CD28+ IGM− activated B Cell
B2. IGHD+ FCER2+ Naïve B Cell
B3. BCL2+ CD27+ MCL1+ Unswitched Memory B Cell
B4. BCL2+ BCL11A+ IGHE+ Memory B Cell
B5. FCRL5+ ITGAX+ TBX21+ ABC−like
P6. CD38++ MKI67+ Plasmablast
P7. IGG+ Plasma Cell
P8. IGA+ Plasma Cell
b
PBMC - B/Plasma
bl−B0. CXCR5high Naive
bl−B1. CRIP1+ Mature
bl−B2. FCRL5+ ITGAX+ ABC
bl−B3. CD79A+ VREPB3+ pre B Cell
bl−B4. ISGhigh Naive
bl−B5. MThigh
bl−B6. JUN+ NFKB1+ Activated
bl−B7. IGKC+ XBP1+ Plasma−like
bl−B8. FCRL5+ IGHG1+ B Cell
c
Tissue Score
Tissue States
P7 P8 P6 B1 B3 B0 B4 B2 B5 bl−B5 bl−B8 bl−B2 bl−B1 bl−B7 bl−B3 bl−B4 bl−B0 bl−B6
Tissue Specific
Shared
Blood Specific
PBMC States
d
Tissue - Myeloid
M0. CD16+ CX3CR1+ Monocyte
M1. CD14+ CD16+ CCL2+ CX3CR1+ Monocyte
M2. CD14+ CCR2+ Monocyte
M3. CCL2+ CCL3+ Monocyte
M4. TPSB2+ MAST cell
M5. GPNMBhigh NUPR1high Macrophage
M6. SELENOPinter ISGHigh Macrophage
M7. SPP1high FABP5high Macrophage
M8. SPP1low FABP5high Macrophage
M9. MERTKhigh FABP5high Macrophage
M10. SELENOPinter LYVE1inter Resident Macrophage
M11. GPNMBhigh NUPR1low Macrophage
M12. SELENOPhigh LYVE1high Resident Macrophage
M13. APOChigh C9high Macrophage
M14. APOClow C9high Macrophage
M15. CENPF+ MKI67+ Proliferating
DC16. CCR7+ LAMP3+ DC2
DC17. CLEC10Alow cDC2
DC18. CLEC10Ahigh cDC2
DC19. cDC1
DC20. pDC
e
PBMC - Myeloid
bl−M0. CD14+ CD16− S100Ahigh
bl−M1. CD14+ CD16− CXCL8+
bl−M2. CD14+ CD16− CCR2high
bl−M3. CD16++ CD14dim CDKN1C+
bl−M4. CD14+ CD16+ MHC2higher
bl−M5. CD14+ CD16− LGALS2+
bl−M6. CD14+ CD16− ISGhigh
bl−M7. MThigh
bl−M8. CD14+ CD16+ MHC2lower
bl−P9. PPBP+ GP1BB+ Platelet
bl−DC10. CLEC10A+ CD1C+ DC2
bl−DC11. TCF4+ CLEC4C+ pDC
bl−DC12. CLEC9A+ XCR1+ DC1
f
Tissue Score
Tissue States
M5 M11 M13 M14 M7 M9 M12 M10 M15 M4 M6 M8 DC19 DC17 DC16 DC18 M1 DC20 M0 M2 M3 bl−DC11 bl−DC12 bl−M4 bl−M3 bl−DC10 bl−M7 bl−M8 bl−M0 bl−M1 bl−M5 bl−M6 bl−P9 bl−M2
Tissue Specific
Shared
Blood Specific
PBMC States
g
Tissue - T/NK
NK0. CD56dim NK
NK1. CD56bright NK
T2. CD8+ GZMB+ CTL
T3. CD8+ GZMB+ SYNE2bright CTL
T4. CENPF+ MKI67+ Proliferating
T5. GZMK+ CD8+ NKG7high
T6. GZMK+ CD8+ NKG7low
T7. GZMK+ CD8+ Effector Memory
T8. GZMK+ CD8+ NEAT1+
T9. GZMK+ CD8+ Resident Memory
T10. GZMK+ CD8+ ITGAE
T11. CD4+ Effector Memory
T12. CD8+ GMZK+ CD69+
T13. CD4+ JUNlow Resident Memory
T14. CD4+ JUNhigh Resident Memory
T15. CD4+ S1PR1+ Central memory/Naive
T16. KLRB1+ KIT+ ILC
T17. CD4+ RORC+ CCR6+ Th17
T18. CD4+ Central Memory/Naive
T19. CD4+ IL2RA++ FOXP3++ Treg
T20. T20. CD4+ FOXP3+ Central Memory/Naive
T21. CD4+ PDCD1+ CXCR5+ TFH/TPH
h
PBMC - T/NK
bl−NK0. CD56dim CD16high FGFBP2high NK
bl−NK1. CD56bright NEAT1high PRF1high NK
bl−NK2. CD56bright XCL2high IL2RBhigh NK
bl−T3. CD8+ GZMHhigh FGFBP2high
bl−T4. CD8+ GZMK+ CD74+ HLA−DR+
bl−T5. CD4+ IL7Rhigh VIMhigh
bl−T6. CD8+ MT−high
bl−T7. CD8+ GZMK+ TEMRA
bl−T8. CD8+ Central Memory/Naive
bl−T9. CD4+ Effector Memory
bl−T10. CD4+ Central Memory/Naive
bl−T11. TRDC+ Gamma/Delta
bl−T12. TRGC1+ Gamma/Delta
bl−T13. CD4+ MAF+ IT2MA+ Effector Memory
bl−T14. CD8+ GZMK+ CD74high HLA−DRhigh
bl−T15. CD8+ GZMB+ DNMT1+ HELLS+ Proliferating
bl−T16. CD4+ T−reg
bl−T17. ISGhigh
bl−T18. CD8+ GZMB+ PCNAhigh Proliferating
bl−T19. CD8+ GZMB+ CENPFhigh Proliferating
i
Tissue Score
Tissue States
T9 T9 T16 NK1 T12 T10 T14 T17 T19 T5 T11 T21 T13 T20 T8 T16 T7 T15 T4 T2 NK0 T3 bl−T16 bl−T5 bl−T10 bl−T13 bl−T9 bl−T7 bl−T8 bl−T18 bl−T6 bl−T15 bl−T17 bl−NK0 bl−T19 bl−NK1 bl−NK2 bl−T11 bl−T14 bl−T3 bl−T4 bl−T12
Tissue Specific
Shared
Blood Specific
PBMC States

Figure 3. Tissue specific cell states. a-b, Joint embedding of tissue (**a**) and blood (**b**) B single-cell populations, colored by their original fine-grain annotations obtained in a separate embedding. **c**, Tissue specificity metric (TSM) scores stratified by cell state annotation. TSM was calculated as the proportion of a given cell's 50 nearest neighbors that are from tissue, weighted to account for larger amount of total myeloid blood than tissue myeloid cells. The median of the distribution is at the center line and the outer bounds of the box portray the first and third quartiles. Whiskers represent the upper and lower ranges of the data. **d-f**, Same as **a-c** in Myeloid cells. **g-i**, Same as **a-c** in T/NK cells.

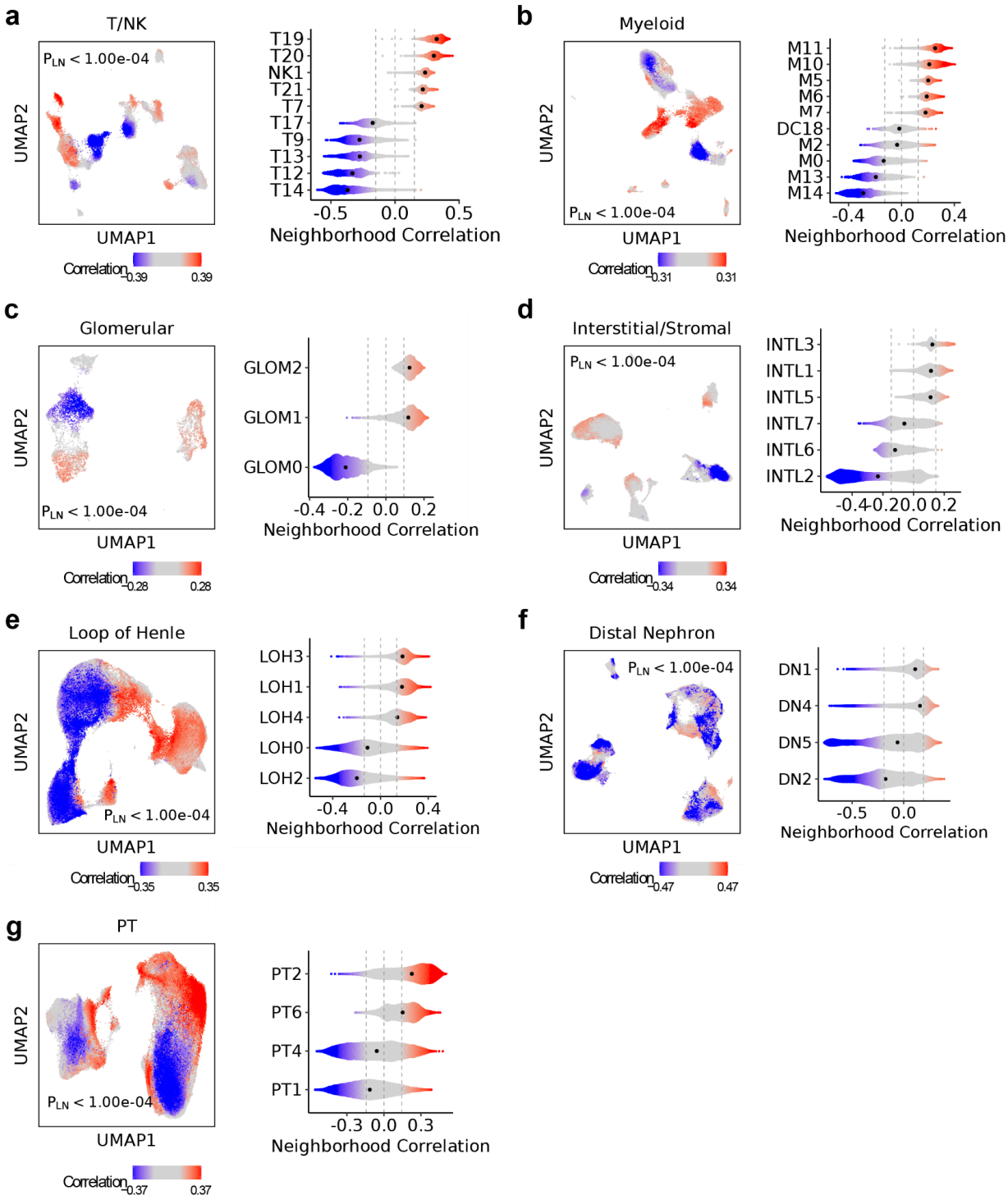

a
T/NK
P_{LN} < 1.00e-04
UMAP2
UMAP1
Correlation
-0.39
0.39
T19
T20
NK1
T21
T7
T17
T9
T13
T12
T14
-0.5
0.0
0.5
Neighborhood Correlation
b
Myeloid
P_{LN} < 1.00e-04
UMAP2
UMAP1
Correlation
-0.31
0.31
M11
M10
M5
M6
M7
DC18
M2
M0
M13
M14
-0.4
0.0
0.4
Neighborhood Correlation
c
Glomerular
P_{LN} < 1.00e-04
UMAP2
UMAP1
Correlation
-0.28
0.28
GLOM2
GLOM1
GLOM0
-0.4
-0.2
0.0
0.2
Neighborhood Correlation
d
Interstitial/Stromal
P_{LN} < 1.00e-04
UMAP2
UMAP1
Correlation
-0.34
0.34
INTL3
INTL1
INTL5
INTL7
INTL6
INTL2
-0.4
-0.2
0.0
0.2
Neighborhood Correlation
e
Loop of Henle
P_{LN} < 1.00e-04
UMAP2
UMAP1
Correlation
-0.35
0.35
LOH3
LOH1
LOH4
LOH0
LOH2
-0.4
0.0
0.4
Neighborhood Correlation
f
Distal Nephron
P_{LN} < 1.00e-04
UMAP2
UMAP1
Correlation
-0.47
0.47
DN1
DN4
DN5
DN2
-0.5
0.0
Neighborhood Correlation
g
PT
P_{LN} < 1.00e-04
UMAP2
UMAP1
Correlation
-0.37
0.37
PT2
PT6
PT4
PT1
-0.3
0.0
0.3
Neighborhood Correlation

Figure 4. Case-control associations within cell types. a-g CNA results for case-control association with no covariate corrections, in scRNA-seq data. Left - UMAP displaying significant per-cell associations with LN, with FDR cutoff of 0.05. Blue and red denote contracted and expanded neighborhoods that pass FDR, respectively. Grey represents neighborhoods that do not pass FDR. Non-significant associations are colored in grey. P-value is the global P-value for cell state associations with LN phenotype. Right - violin plots of clusters containing cells passing FDR significance for LN association. The median is denoted with a black dot. Analysis was repeated for **a,** T/NK cells, **b,** Myeloid cells, **c,** B/Plasma cells, **d,** Interstitial/Stromal cells, **e,** Loop of Henle cells, **f,** Distal Nephron, **g,** Proximal Tubule cells.

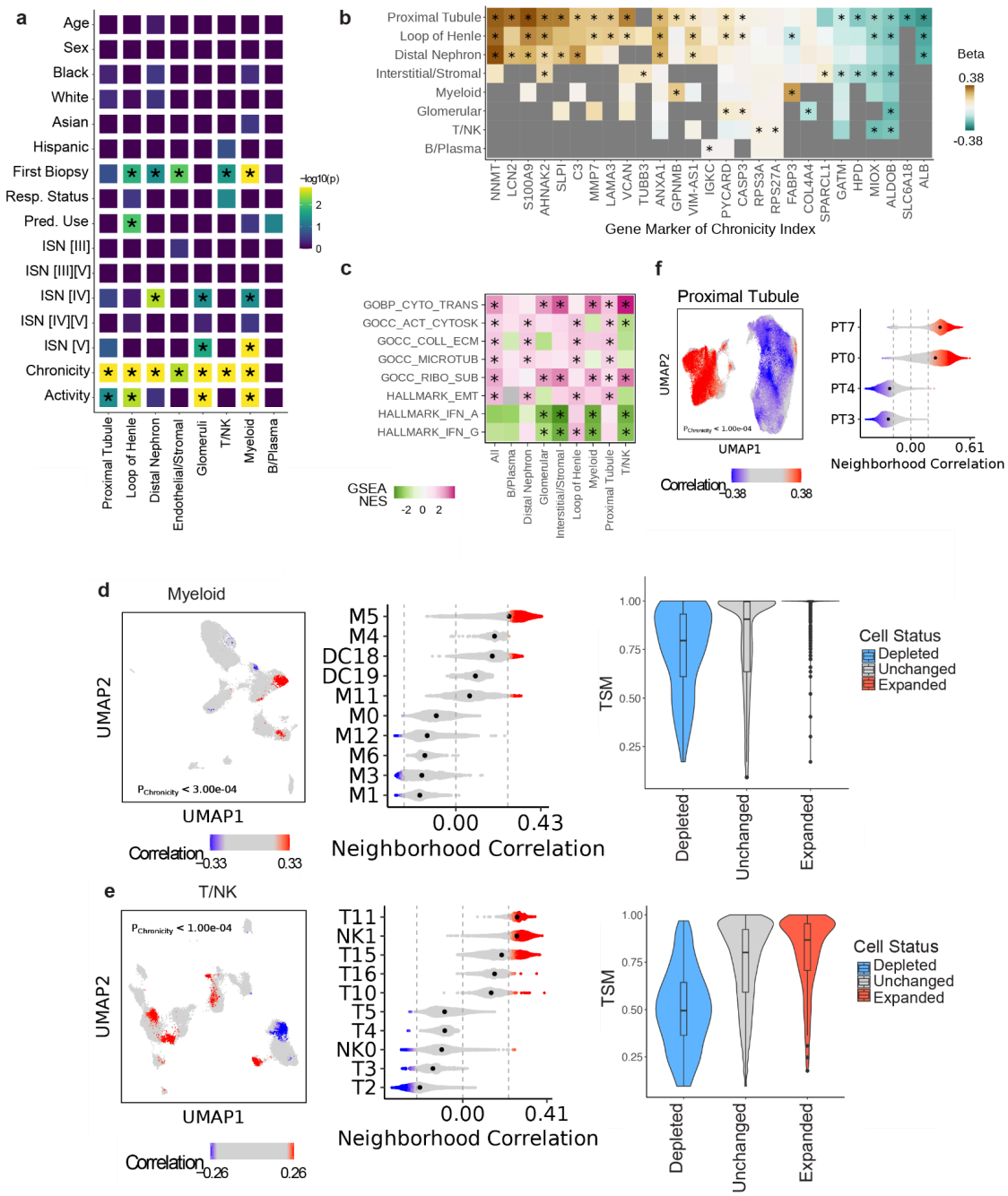

a
Age
Sex
Black
White
Asian
Hispanic
First Biopsy
Resp. Status
Pred. Use
ISN [III]
ISN [III][V]
ISN [IV]
ISN [IV][V]
ISN [V]
Chronicity
Activity
Proximal Tubule
Loop of Henle
Distal Nephron
Endothelial/Stromal
Glomeruli
T/NK
Myeloid
B/Plasma
-log10(p)
b
Proximal Tubule
Loop of Henle
Distal Nephron
Interstitial/Stromal
Myeloid
Glomerular
T/NK
B/Plasma
NNMT
LCN2
S100A9
AHNAK2
SLPI
C3
MMP7
LAMA3
VCAN
TUBB3
ANXA1
GPNMB
VIM-AS1
IGKC
PYCARD
CASP3
RPS3A
RPS27A
FABP3
COL4A4
SPARCL1
GATM
HPD
MIOX
ALDOB
SLC6A18
ALB
Gene Marker of Chronicity Index
Beta
0.38
-0.38
c
GOBP_CYTO_TRANS
GOCC_ACT_CYTOSK
GOCC_COLL_ECM
GOCC_MICROTUB
GOCC_RIBO_SUB
HALLMARK_EMT
HALLMARK_IFN_A
HALLMARK_IFN_G
All
B/Plasma
Distal Nephron
Glomerular
Interstitial/Stromal
Loop of Henle
Myeloid
Proximal Tubule
T/NK
GSEA NES
-2 0 2
f
Proximal Tubule
UMAP2
UMAP1
PT7
PT0
PT4
PT3
0.00
0.61
Neighborhood Correlation
Correlation
-0.38
0.38
d
Myeloid
UMAP2
UMAP1
M5
M4
DC18
DC19
M11
M0
M12
M6
M3
M1
0.00
0.43
Neighborhood Correlation
Correlation
-0.33
0.33
TSM
Cell Status
Depleted
Unchanged
Expanded
e
T/NK
UMAP2
UMAP1
T11
NK1
T15
T16
T10
T5
T4
NK0
T3
T2
0.00
0.41
Neighborhood Correlation
Correlation
-0.26
0.26
TSM
Cell Status
Depleted
Unchanged
Expanded

Figure 5. Cell states associated with chronicity index. a, Univariate CNA of each cell type testing for relevant clinical and demographic variables. * indicates a global p-value < 0.05. **b,** For each cell type, differential expression for genes associated with chronicity index. * indicates a Bonferroni-adjusted p-value less than 0.05. Grey color indicates the gene was not tested due to low expression in that cell type. **c,** Normalized enrichment scores (NES) from pathway enrichment of differential gene expression results, using Hallmark and Ontology pathways. * indicates a Bonferroni-adjusted p-value less than 0.05. Grey color indicates the pathway was not tested due to low expression of pathway genes in that cell type. **d,** Across Myeloid cell types, CNA results for chronicity index association while adjusting for first biopsy and site, in scRNA-seq data. Left - UMAP displaying significant per-cell associations with chronicity index, with FDR cutoff of 0.05. Non-significant associations are colored in grey. P-value is the global P-value for cell state associations with chronicity index phenotype. Middle - violin plots of clusters containing cells passing FDR significance for chronicity index association. Right - Within Myeloid cells, violin plots displaying per-cell TSM scores for cells found to be expanded, depleted, or neither with chronicity index, as identified using CNA. **e,** Same as **d** for T/NK cells. **f,** Same as **d** for PT cells, without TSM score derivations.

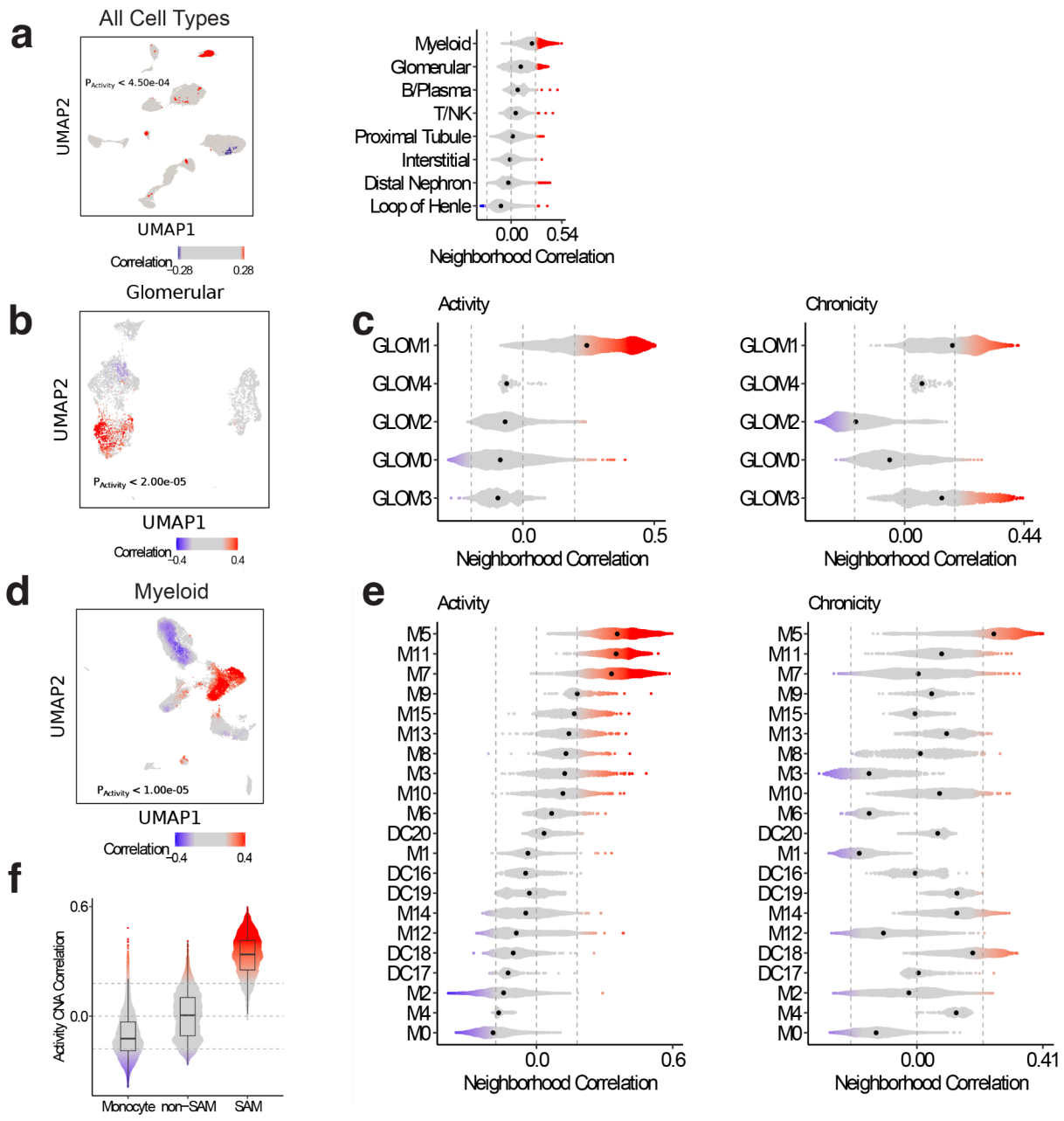
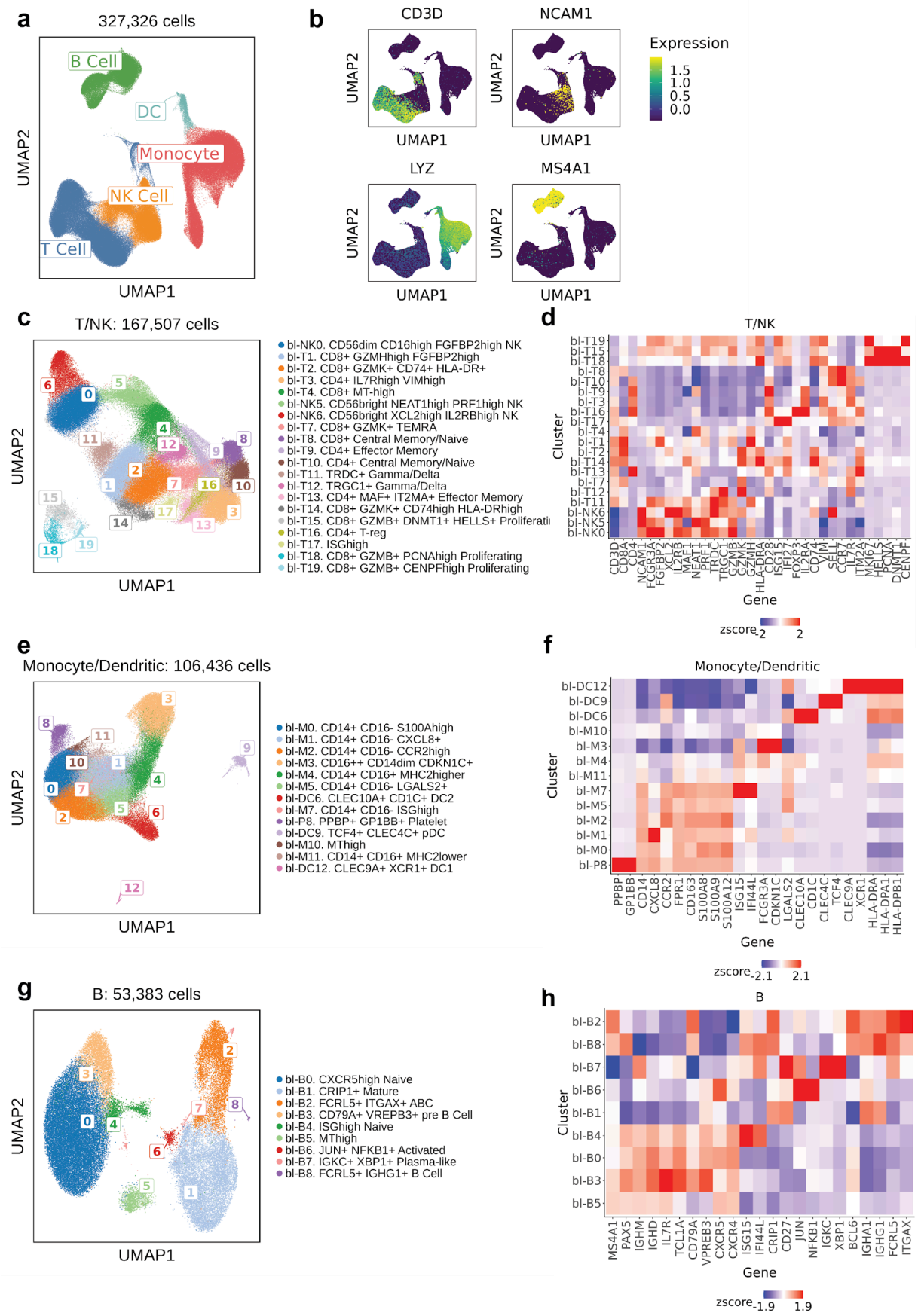


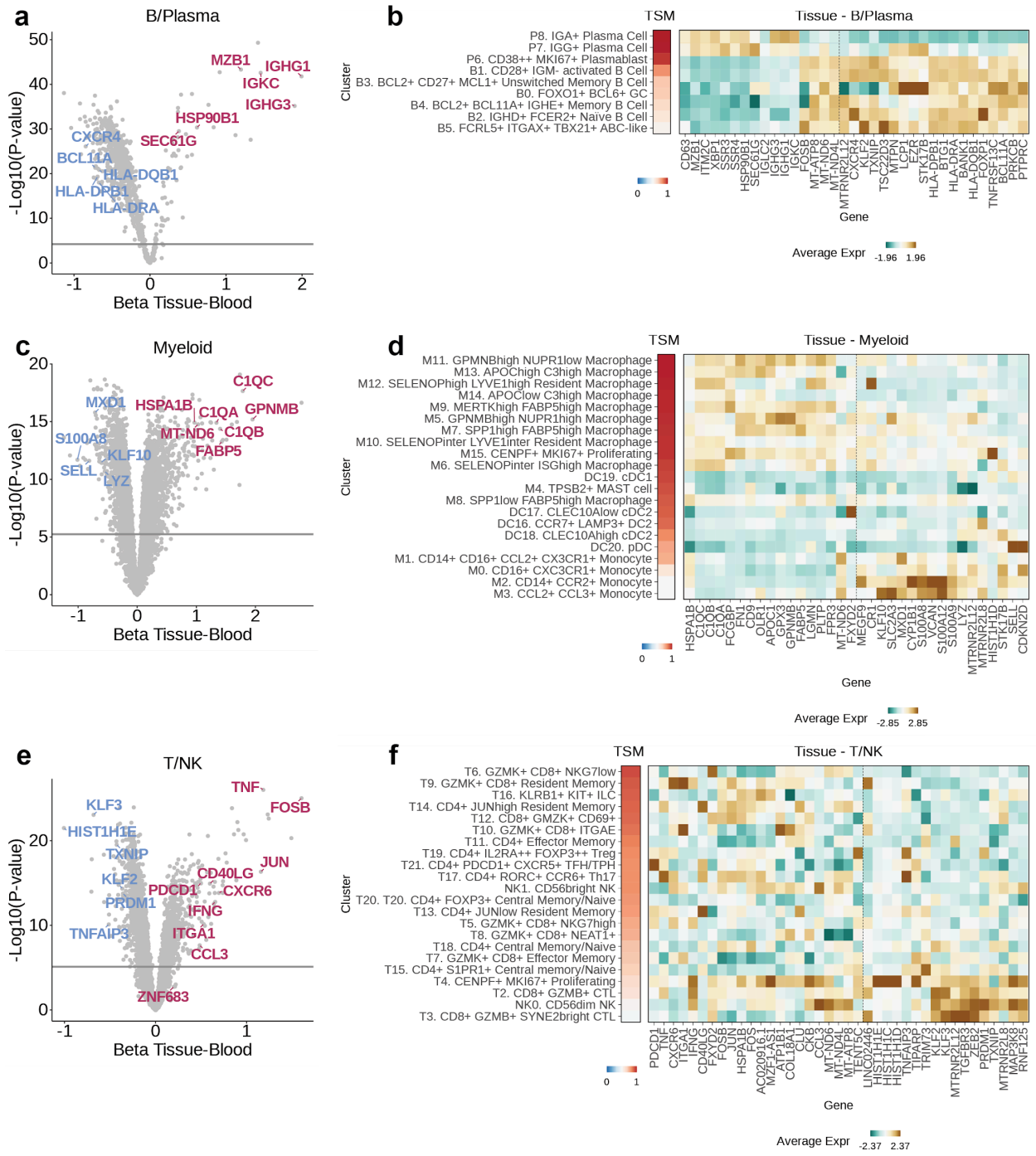
Figure 6. Cell states associated with activity index. a, Across all cell types, CNA results for activity index association while adjusting for first biopsy, site, and chronicity index, in scRNA-seq data. Left - UMAP displaying significant per-cell associations with activity index, with FDR cutoff of 0.05. Non-significant associations are colored in grey. P-value is the global P-value for cell state associations with activity index phenotype. Right - violin plots of clusters

containing cells passing FDR significance for activity index association. **b,** Same as **a** Left for Glomerular cells. **c,** Left - Same as **a** Right for Glomerular cells. Right - Within Glomerular cells, violin plots of clusters containing cells passing FDR significance for chronicity index association (adjusting for first biopsy and site). **d,** Same as **a** Left for Myeloid cells. **e,** Within Myeloid cells, violinplots displaying per-cell TSM scores for cells found to be expanded, depleted, or neither with activity index, as identified using CNA. **f,** Left - Same as **a** Right for Myeloid cells. Right - Within Myeloid cells, violin plots of clusters containing cells passing FDR significance for chronicity index association (adjusting for first biopsy and site).

Extended Data Figures / Legends

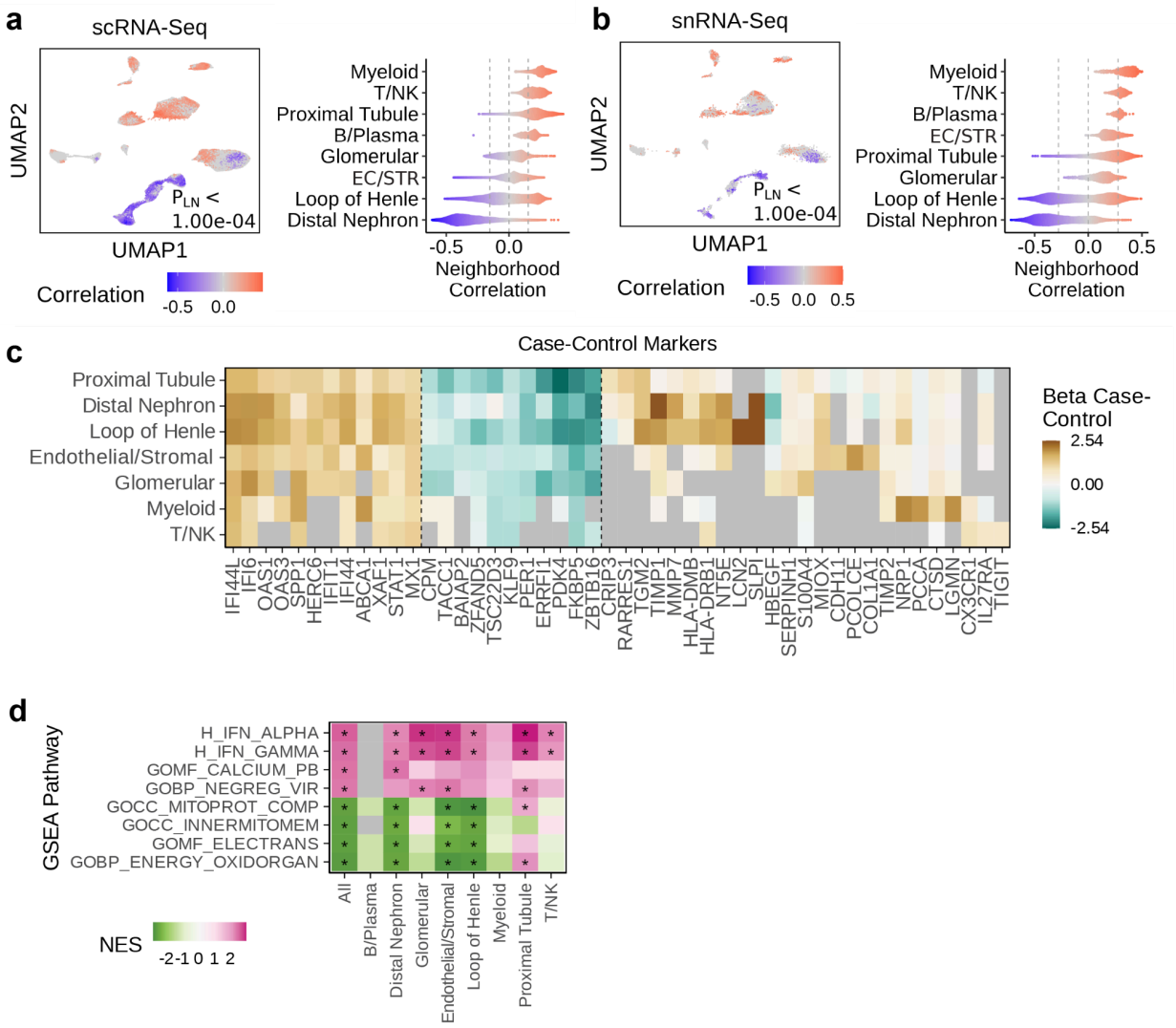

a
327,326 cells
B Cell
DC
Monocyte
NK Cell
T Cell
UMAP2
UMAP1
b
CD3D
NCAM1
LYZ
MS4A1
Expression
1.5
1.0
0.5
0.0
c
T/NK: 167,507 cells
bl-NK0. CD56dim CD16high FGFBP2high NK
bl-T1. CD8+ GZMHhigh FGFBP2high
bl-T2. CD8+ GZMK+ CD74+ HLA-DR+
bl-T3. CD4+ IL7Rhigh VIMhigh
bl-T4. CD8+ MT-high
bl-NK5. CD56bright NEAT1high PRF1high NK
bl-NK6. CD56bright XCL2high IL2RBhigh NK
bl-T7. CD8+ GZMK+ TEMRA
bl-T8. CD8+ Central Memory/Naive
bl-T9. CD4+ Effector Memory
bl-T10. CD4+ Central Memory/Naive
bl-T11. TRDC+ Gamma/Delta
bl-T12. TRGC1+ Gamma/Delta
bl-T13. CD4+ MAF+ IT2MA+ Effector Memory
bl-T14. CD8+ GZMK+ CD74high HLA-DRhigh
bl-T15. CD8+ GZMB+ DNMT1+ HELLS+ Proliferati
bl-T16. CD4+ T-reg
bl-T17. ISGhigh
bl-T18. CD8+ GZMB+ PCNAhigh Proliferating
bl-T19. CD8+ GZMB+ CENPFhigh Proliferating
d
T/NK
Cluster
Gene
zscore
-2
2
e
Monocyte/Dendritic: 106,436 cells
bl-M0. CD14+ CD16- S100Ahigh
bl-M1. CD14+ CD16- CXCL8+
bl-M2. CD14+ CD16- CCR2high
bl-M3. CD16++ CD14dim CDKN1C+
bl-M4. CD14+ CD16+ MHC2higher
bl-M5. CD14+ CD16- LGALS2+
bl-DC6. CLEC10A+ CD1C+ DC2
bl-M7. CD14+ CD16- ISGhigh
bl-P8. PPBP+ GP1BB+ Platelet
bl-DC9. TCF4+ CLEC4C+ pDC
bl-M10. MThigh
bl-M11. CD14+ CD16+ MHC2lower
bl-DC12. CLEC9A+ XCR1+ DC1
f
Monocyte/Dendritic
Cluster
Gene
zscore
-2.1
2.1
g
B: 53,383 cells
bl-B0. CXCR5high Naive
bl-B1. CRIP1+ Mature
bl-B2. FCRL5+ ITGAX+ ABC
bl-B3. CD79A+ VREPB3+ pre B Cell
bl-B4. ISGhigh Naive
bl-B5. MThigh
bl-B6. JUN+ NFKB1+ Activated
bl-B7. IGKC+ XBP1+ Plasma-like
bl-B8. FCRL5+ IGHG1+ B Cell
h
B
Cluster
Gene
zscore
-1.9
1.9

Extended Data Figure 1. a, UMAP of all single cells annotated in blood, colored by broad cell type annotation. **b,** Per-cell gene expression for various broad cell type markers. Gene values are log(TPM+1) normalized and scaled for visualization. **c,** UMAP of T/NK cell states in blood, colored by fine grain cell state annotations. Annotation boxes are located at the centroid (mean) of each cluster. **d,** Heatmap depicting marker gene expression in each T/NK cell state cluster. Expression values the log(TPM+1) cluster average of a marker, and z-score scaled by column for visualization. **e-f,** Same as **c-d** for Myeloid cells. **g-h,** Same as **c-d** for B/Plasma cells.



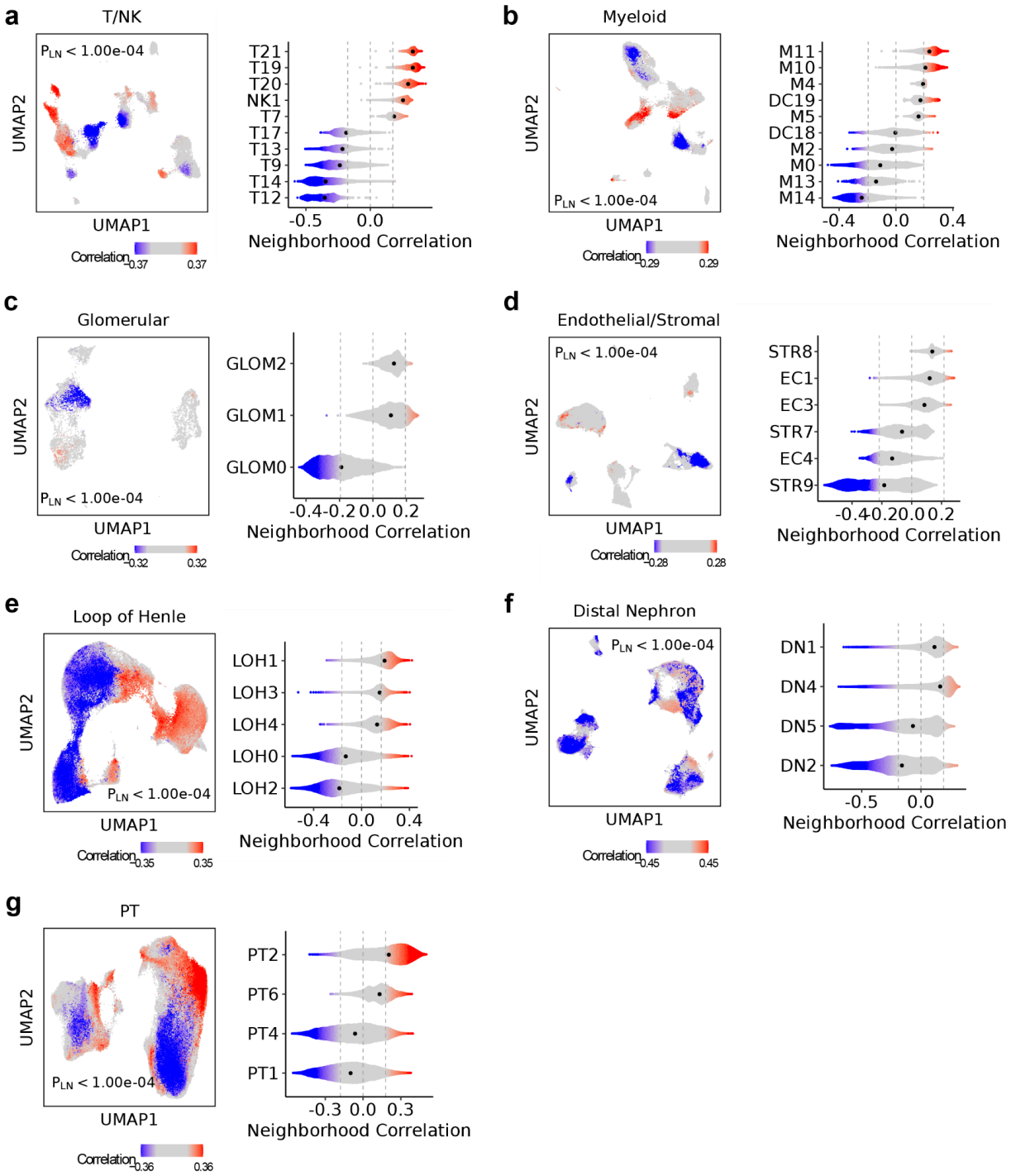
Extended Data Figure 2. a, In B/plasma cells, differential expression results between tissue and blood samples. Betas and p-values are derived from pseudobulk linear modeling (**Methods**). Genes upregulated with respect to tissue are displayed in red, downregulated genes are displayed in blue. Horizontal lines indicate the Bonferonni corrected significance

threshold for $p<0.05$. **b,** Heatmap depicting average log(CPM+1) expression of marker genes for each B/plasma cluster in tissue. For visualization, expression values are z-scored by column. Dashed line separates genes identified as enriched (left) and downregulated (right) in tissue via differential expression. Clusters are ordered by TSM score. **c-d,** Same as **a-b** but in Myeloid cells. **e-f,** Same as **a-b** but in T/NK cells.

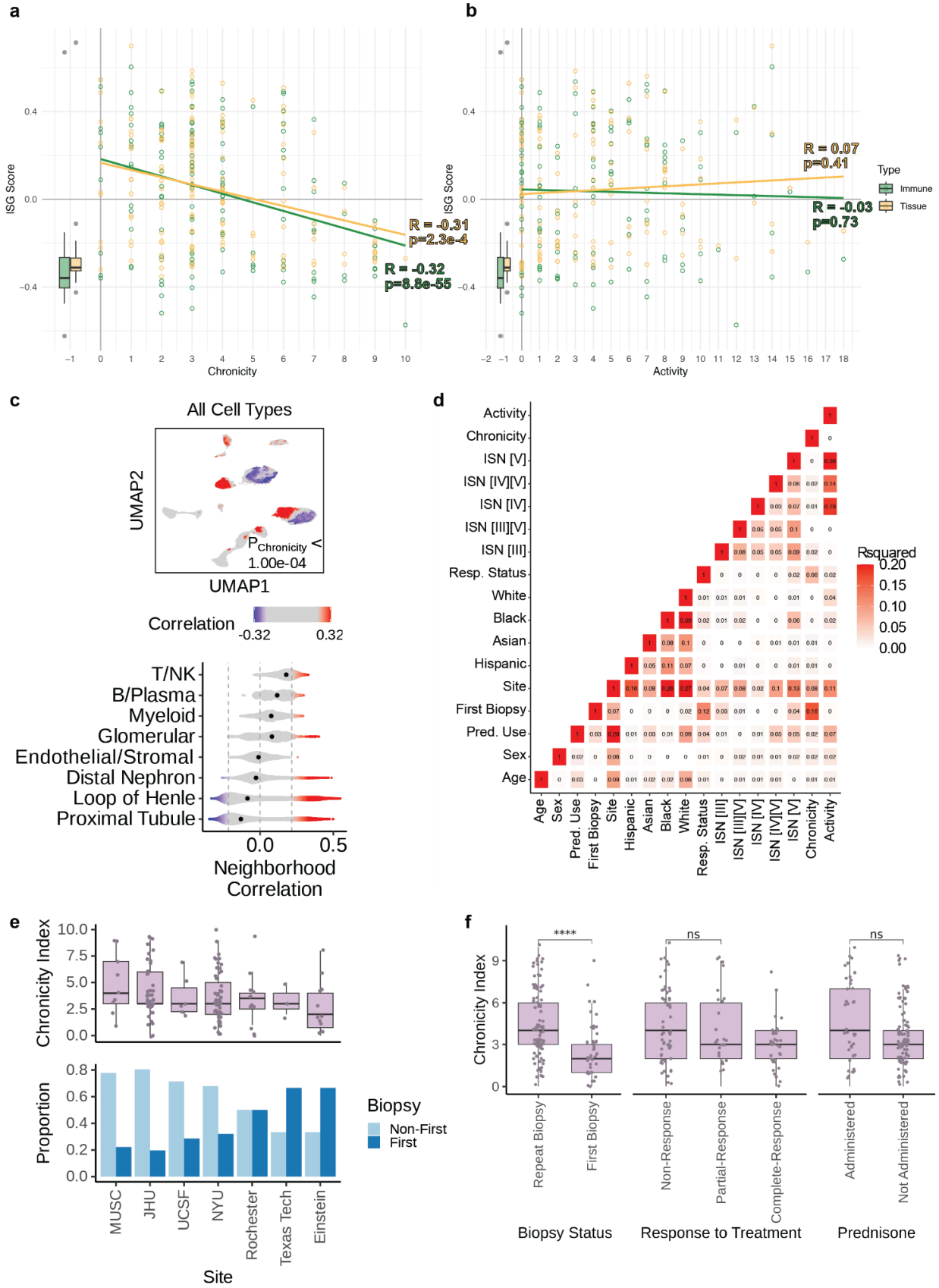


Extended Data Figure 3. a, Across all cell types, CNA results for case-control association with no covariate corrections, in scRNA-seq data. Left - UMAP displaying significant per-cell associations with LN, with FDR cutoff of 0.05. Non-significant associations are colored in grey. P-value is the global P-value for cell state associations with LN phenotype. Right - violin plots of clusters containing cells passing FDR significance for LN association. **b,** Same as **a** in snRNA-seq data. **c,** For each cell type, differential expression for genes associated with case-control status. Dotted lines separate genes enriched in most cell types and depleted in

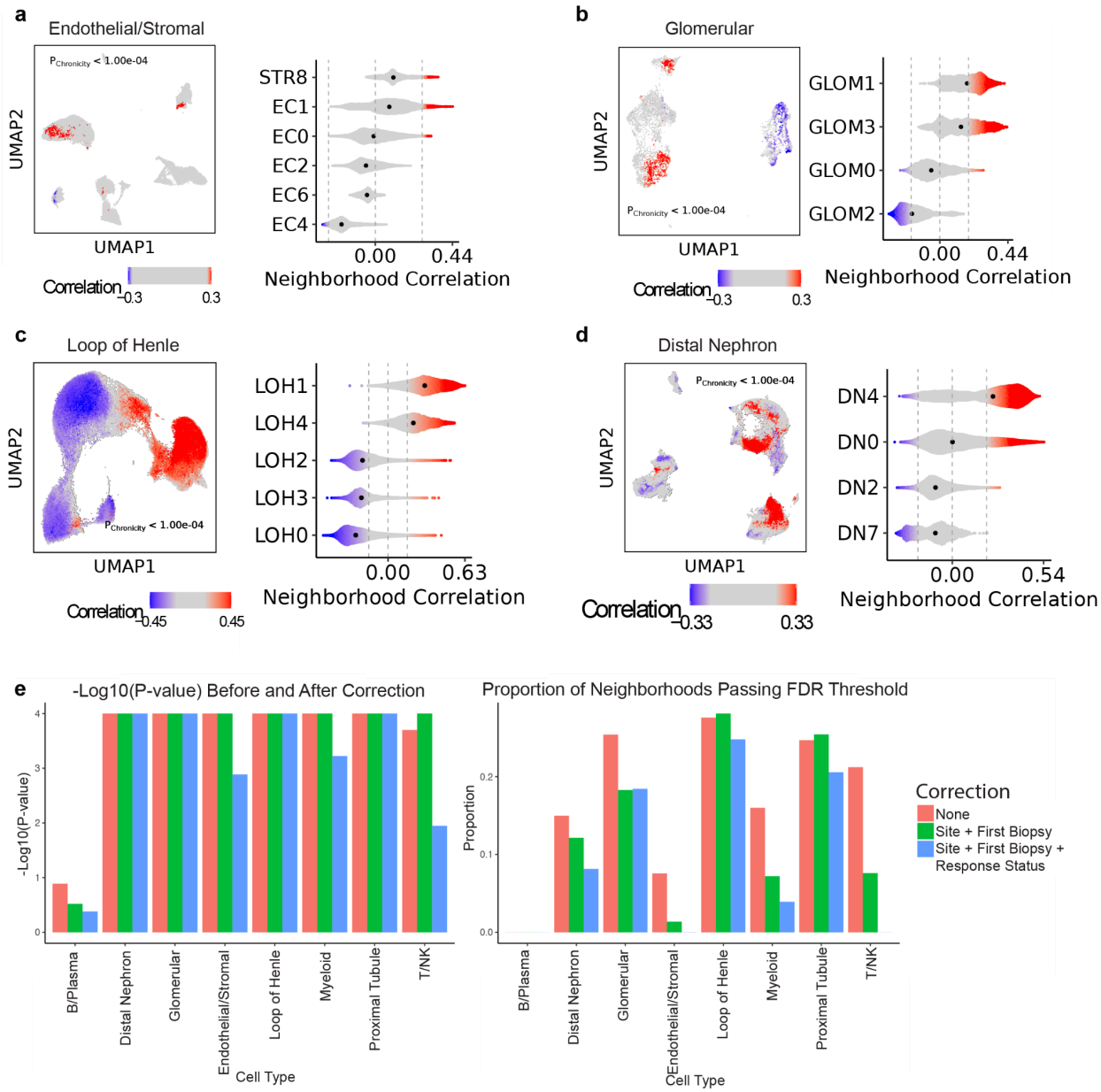
most cell types. **d,** Normalized enrichment scores (NES) from pathway enrichment of differential gene expression results, using Hallmark and Ontology pathways. * indicates a Bonferroni-adjusted p-value less than 0.05. Grey color indicates the pathway was not tested due to low expression of pathway genes in that cell type. Full pathway names, in order of the figure labels, can be found here: . **e,** Same as **d,** where pathway enrichment is on adjusted differential gene expression results, where results have been corrected for all cross-cell-type differential expression signals (**Methods**).


a
T/NK
P_{LN} < 1.00e-04
UMAP2
UMAP1
Correlation
-0.37
0.37
T21
T19
T20
NK1
T7
T17
T13
T9
T14
T12
-0.5
0.0
Neighborhood Correlation
b
Myeloid
P_{LN} < 1.00e-04
UMAP2
UMAP1
Correlation
-0.29
0.29
M11
M10
M4
DC19
M5
DC18
M2
M0
M13
M14
-0.4
0.0
0.4
Neighborhood Correlation
c
Glomerular
P_{LN} < 1.00e-04
UMAP2
UMAP1
Correlation
-0.32
0.32
GLOM2
GLOM1
GLOM0
-0.4 -0.2 0.0 0.2
Neighborhood Correlation
d
Endothelial/Stromal
P_{LN} < 1.00e-04
UMAP2
UMAP1
Correlation
-0.28
0.28
STR8
EC1
EC3
STR7
EC4
STR9
-0.4 -0.2 0.0 0.2
Neighborhood Correlation
e
Loop of Henle
P_{LN} < 1.00e-04
UMAP2
UMAP1
Correlation
-0.35
0.35
LOH1
LOH3
LOH4
LOH0
LOH2
-0.4
0.0
0.4
Neighborhood Correlation
f
Distal Nephron
P_{LN} < 1.00e-04
UMAP2
UMAP1
Correlation
-0.45
0.45
DN1
DN4
DN5
DN2
-0.5
0.0
Neighborhood Correlation
g
PT
P_{LN} < 1.00e-04
UMAP2
UMAP1
Correlation
-0.36
0.36
PT2
PT6
PT4
PT1
-0.3 0.0 0.3
Neighborhood Correlation

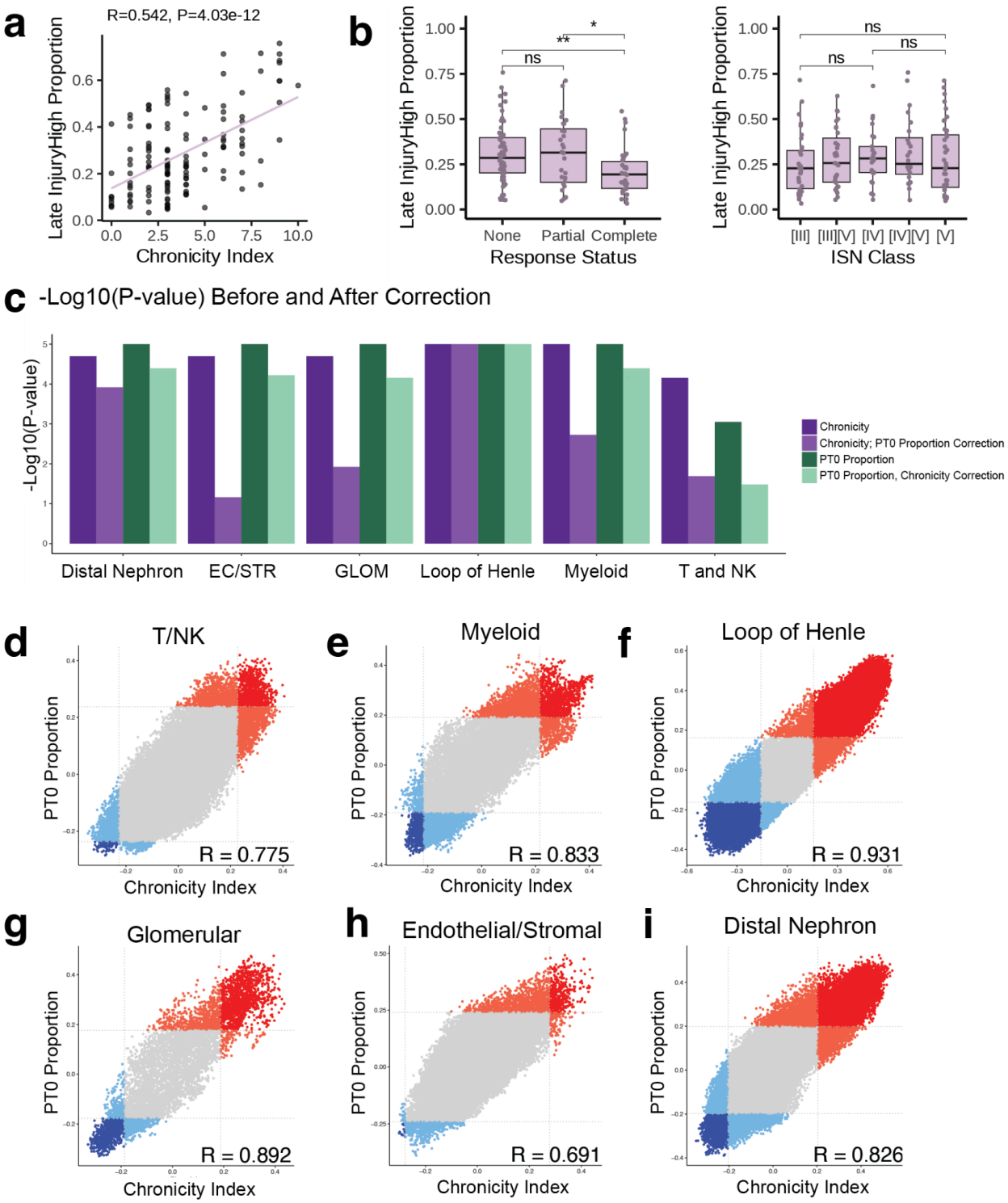
Extended Data Figure 4. a, Within T/NK cells, CNA results for case-control association with covariate corrections for age and sex, in scRNA-seq data. Left - UMAP displaying significant per-cell associations with LN, with FDR cutoff of 0.05. Non-significant associations are colored in grey. P-value is the global P-value for cell state associations with LN phenotype. Right - violin plots of clusters containing cells passing FDR significance for LN association. **b,** Same as **a** for Myeloid cells. **c,** Same as (**a)** for B/Plasma cells. **d,** Same as (**a)** for Interstitial/Stromal cells. **e,** Same as (**a)** for Loop of Henle cells. **f,** Same as (**a)** for Distal Nephron cells. **g,** Same as (**a)** for Proximal Tubule cells.


a
ISG Score
Chronicity
R = -0.31
p=2.3e-4
R = -0.32
p=8.8e-55
b
ISG Score
Activity
R = 0.07
p=0.41
R = -0.03
p=0.73
Type
Immune
Tissue
c
All Cell Types
UMAP2
UMAP1
P Chronicity < 1.00e-04
Correlation
-0.32
0.32
T/NK
B/Plasma
Myeloid
Glomerular
Endothelial/Stromal
Distal Nephron
Loop of Henle
Proximal Tubule
Neighborhood Correlation
d
Rsquared
Activity
Chronicity
ISN [V]
ISN [IV][V]
ISN [IV]
ISN [III][V]
ISN [III]
Resp. Status
White
Black
Asian
Hispanic
Site
First Biopsy
Pred. Use
Sex
Age
e
Chronicity Index
Proportion
Biopsy
Non-First
First
MUSC
JHU
UCSF
NYU
Rochester
Texas Tech
Einstein
Site
f
Chronicity Index
****
ns
ns
Repeat Biopsy
First Biopsy
Non-Response
Partial-Response
Complete-Response
Administered
Not Administered
Biopsy Status
Response to Treatment
Prednisone

Extended Data Figure 5. a-b, Correlation between ISG score (averaged per sample) and Chronicity **(a)**, or Activity Index **(b)**. Controls are displayed as boxplots on the left. Data is stratified between cells originating from immune (green) and tissue (yellow) populations. **c,** Across all cell types, CNA results for chronicity index association while adjusting for first biopsy and site, in scRNA-seq data. Top - UMAP displaying significant per-cell associations with chronicity index, with FDR cutoff of 0.05. Non-significant associations are colored in grey. P-value is the global P-value for cell state associations with chronicity index phenotype. Bottom - violin plots of clusters containing cells passing FDR significance for chronicity index association. **d,** Sample-level correlations of potentially confounding clinical variables. **e,** Top: Boxplots displaying per-sample chronicity index stratified by patient recruitment sites. Box hinges indicate the first and third quartile of the data, with the center line indicating the median. Whiskers indicate 1.5 x the interquartile range, or the furthest outlier, whichever is less. Bottom: Barplot reflecting the proportion of samples, stratified by site, corresponding to a first biopsy sample or a later (non-first) sample. Color indicates sample first biopsy status. Only sites with a chronicity index measured for more than 1 sample are displayed. **f,** Boxplots displaying per-sample chronicity index stratified by various relevant clinical categories. Box hinges indicate the first and third quartile of the data, with the center line indicating the median. Whiskers indicate 1.5 x the interquartile range, or the furthest outlier, whichever is less. **** indicates $p<0.0001$ and ns indicates $p>0.05$.

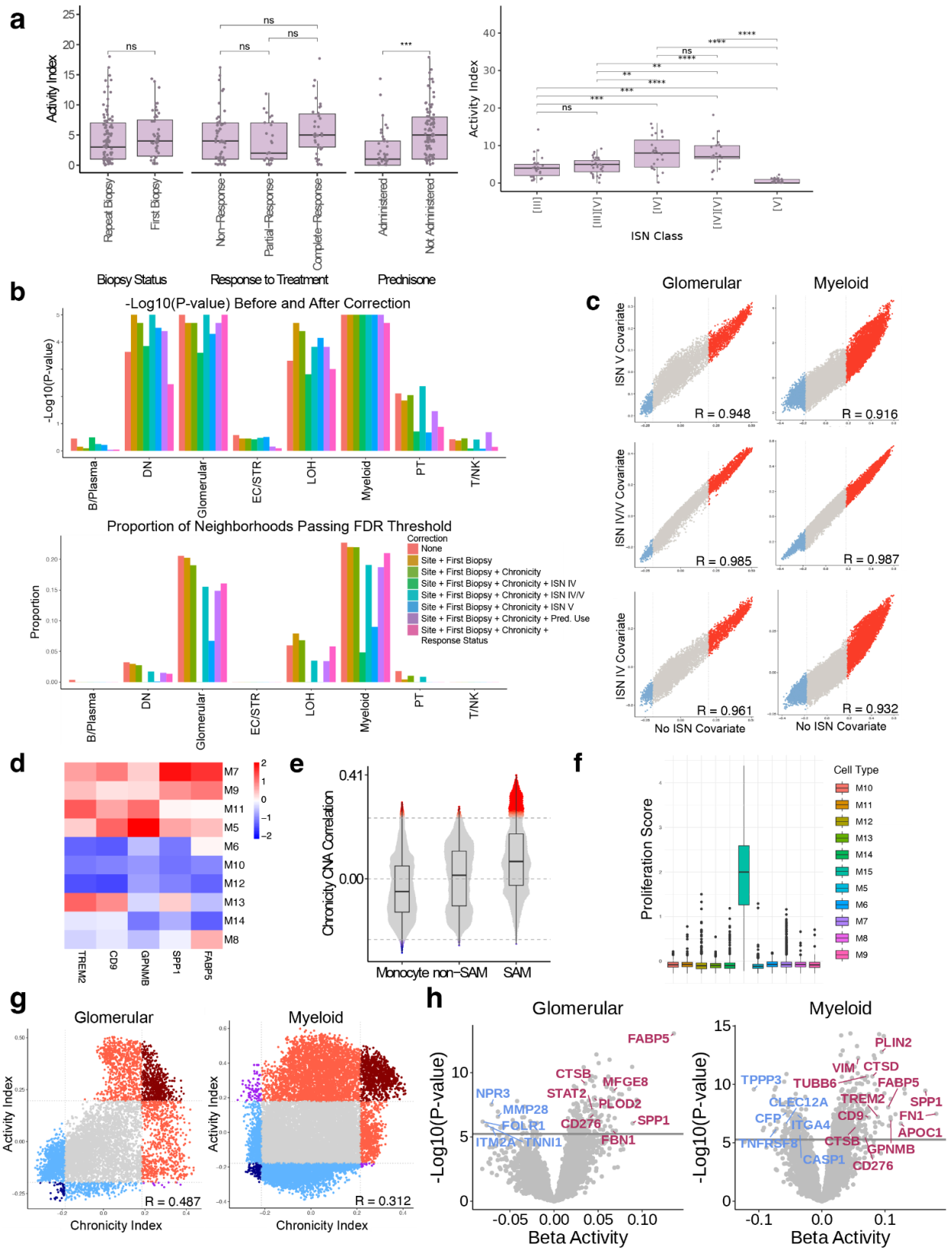


Extended Data Figure 6. a-d, CNA testing correcting for major collection sites and whether the sample is from the initial biopsy following diagnosis for the interstitium **(a)**, glomerulus **(b)**, Loop of Henle **(c)**, or the distal nephron **(d)**. **e,** Global p values (left) and the percent of neighborhoods that pass FDR<0.15 (right) for CNA testing of chronicity accounting for either no or additional relevant clinical covariates.

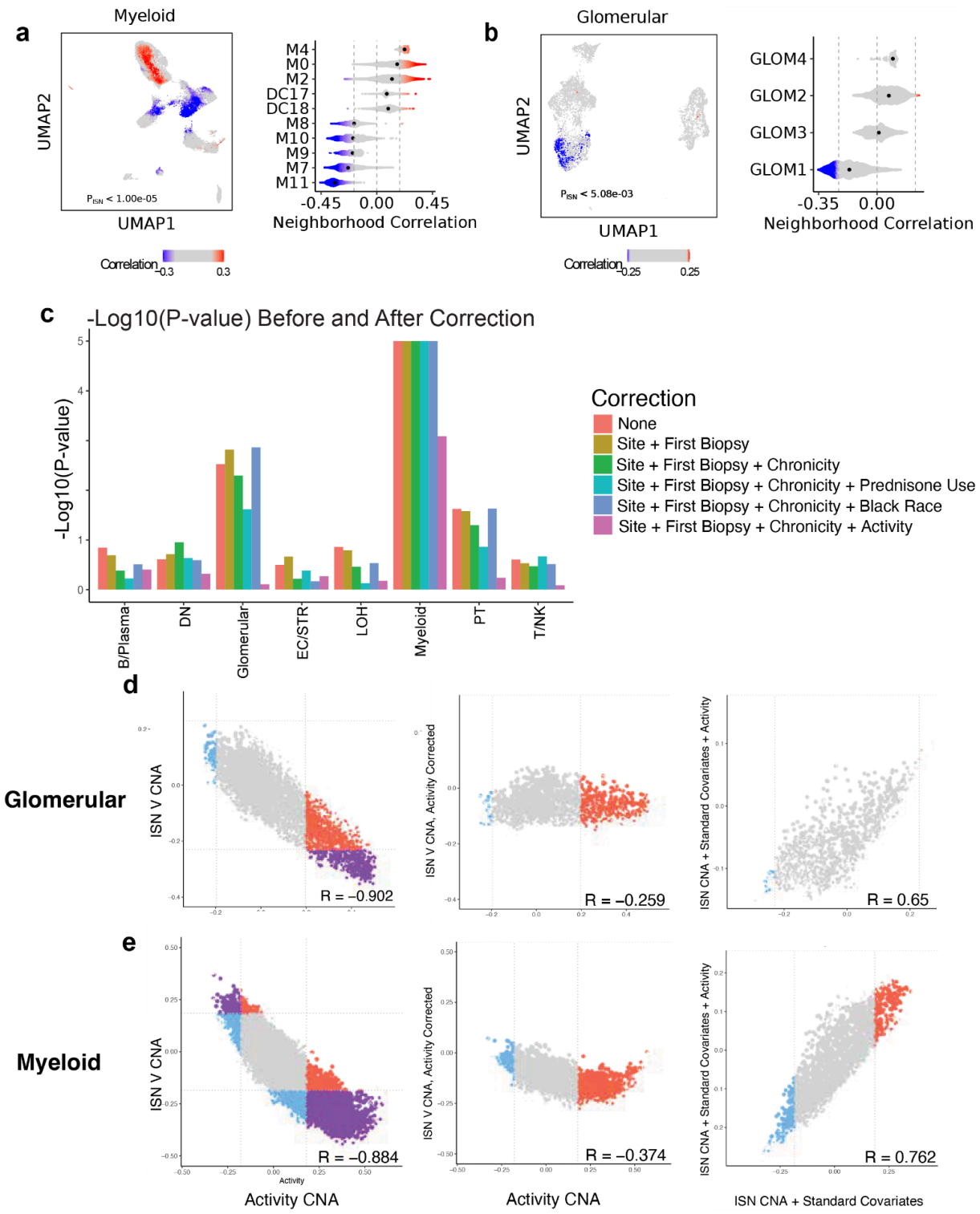


Extended Data Figure 7. a, Scatterplot of the per-sample correlation of the proportion of maladaptive proximal tubule cells and chronicity index. **b,** Boxplots indicating differences of the proportion of maladaptive proximal tubule cells between various response statuses (left) and

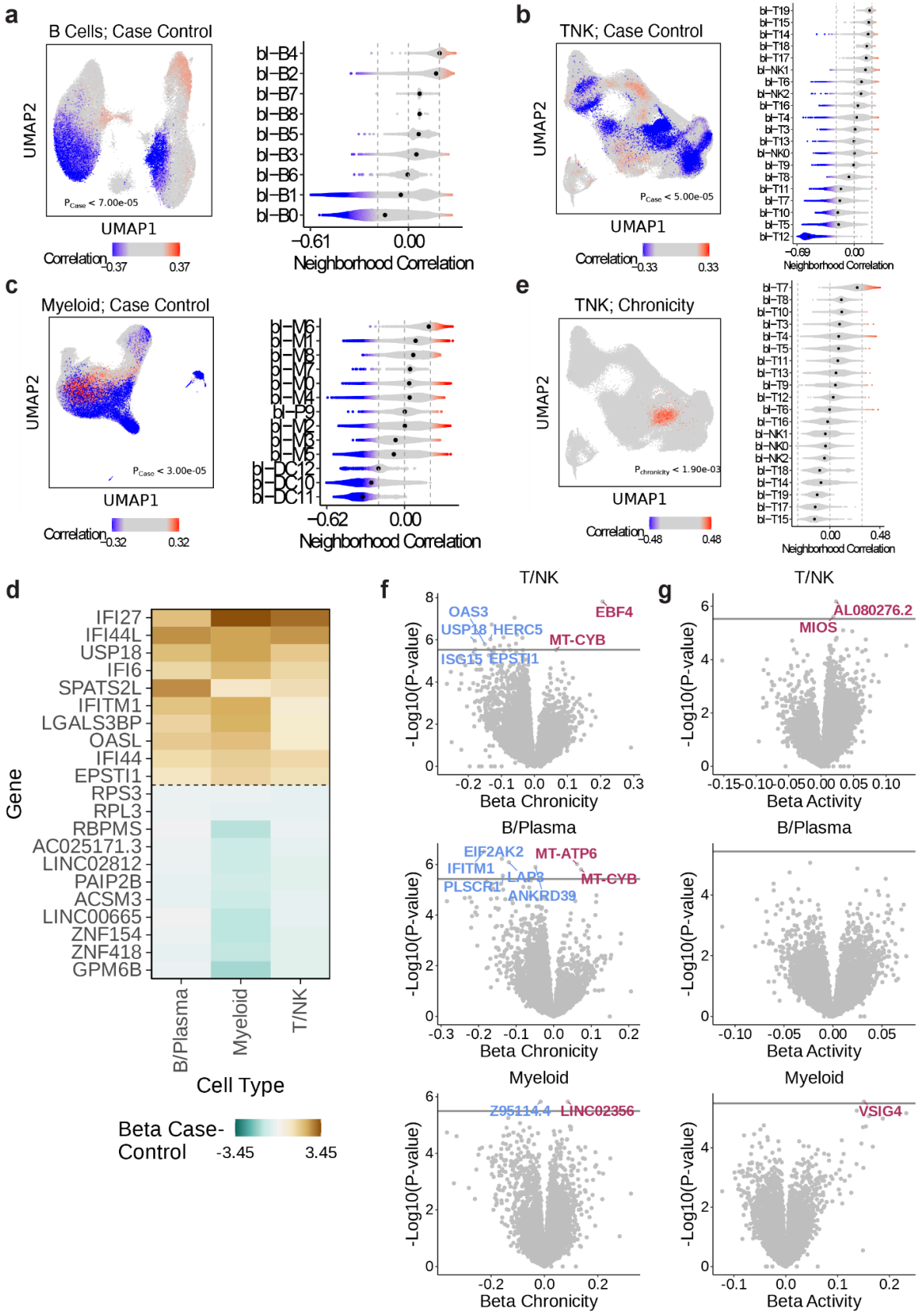
ISN class (right). * indicates $p<0.05$ and ** indicates $p<0.01$. “ns” indicates $p>0.05$**. c,** Global p values of CNA testing for either chronicity or maladaptive proximal tubule cell proportion, correcting for either no additional covariate or the opposite measure. **d-i,** Scatterplot of per-neighborhood correlations of CNA testing of Chronicity index (x-axis) and proportion of maladaptive proximal tubule cells (y-axis) for cells from either the T and NK **(d)**, Myeloid **(e)**, Loop of Henle **(f)**, glomeruli **(g)**, interstitium **(h)**, or distal nephron **(i)** compartments.


a
Activity Index
Repeat Biopsy
First Biopsy
Biopsy Status
Non−Response
Partial−Response
Complete−Response
Response to Treatment
Administered
Not Administered
Prednisone
ISN Class
b
-Log10(P-value) Before and After Correction
-Log10(P-value)
B/Plasma
DN
Glomerular
EC/STR
LOH
Myeloid
PT
T/NK
Proportion of Neighborhoods Passing FDR Threshold
Proportion
Correction
None
Site + First Biopsy
Site + First Biopsy + Chronicity
Site + First Biopsy + Chronicity + ISN IV
Site + First Biopsy + Chronicity + ISN IV/V
Site + First Biopsy + Chronicity + ISN V
Site + First Biopsy + Chronicity + Pred. Use
Site + First Biopsy + Chronicity + Response Status
c
Glomerular
Myeloid
ISN V Covariate
ISN IV/V Covariate
ISN IV Covariate
No ISN Covariate
R = 0.948
R = 0.916
R = 0.985
R = 0.987
R = 0.961
R = 0.932
d
M7
M9
M11
M5
M6
M10
M12
M13
M14
M8
TREM2
CD9
GPNMB
SPP1
FABP5
e
Chronicity CNA Correlation
0.41
0.00
Monocyte
non−SAM
SAM
f
Proliferation Score
Cell Type
M10
M11
M12
M13
M14
M15
M5
M6
M7
M8
M9
g
Glomerular
Myeloid
Activity Index
Chronicity Index
R = 0.487
R = 0.312
h
Glomerular
Myeloid
-Log10(P-value)
Beta Activity
FABP5
CTSB
MFGE8
NPR3
STAT2
PLOD2
MMP28
FOLR1
CD276
SPP1
ITM2A
TNNI1
FBN1
PLIN2
VIM
CTSD
TPPP3
TUBB6
FABP5
CLEC12A
TREM2
SPP1
CFP
ITGA4
CD9
FN1
APOC1
TNFRSF8
CTSB
GPNMB
CASP1
CD276

Extended Data Figure 8. a, Boxplots indicating differences in sample-level differences in Activity Index for relevant clinical variables (left) and ISN class (right). **b,** global p values (top) and the percent of neighborhoods that pass FDR<0.15 (bottom) for CNA testing of activity accounting for either no or additional relevant clinical covariates. **c,** Scatterplot of per-neighborhood correlations for glomeruli and myeloid populations between CNA testing of activity (x-axis) and ISN class IV, IV/V, or V (y-axis). **d,** Gene expression for SAM markers in tissue macrophage populations. **e,** Violin plot of Chronicity CNA correlations for monocyte, non-SAM, and SAM populations. Significant neighborhoods are colored red and blue for expanded and contracted populations, respectively. **f,** Box and whisker plots of proliferation scores stratified to macrophage cell populations and proliferating myeloid cells **g,** Scatterplot of per-neighborhood correlations for glomeruli and myeloid populations between CNA testing of chronicity index (x-axis) and activity index (y-axis) for Glomerular (left) and Myeloid (right) cells. **h,** In Glomerular (left) and Myeloid (right) cells, differential expression results for genes associated with activity index. Betas and p-values are derived from pseudobulk linear modeling (**Methods**). Genes upregulated with respect to activity index are displayed in red, downregulated genes are displayed in blue. Horizontal lines indicate the Bonferonni corrected significance threshold for $p<0.05$.


a
Myeloid
UMAP2
UMAP1
P_ISN < 1.00e-05
Correlation
-0.3
0.3
M4
M0
M2
DC17
DC18
M8
M10
M9
M7
M11
-0.45
0.00
0.45
Neighborhood Correlation
b
Glomerular
UMAP2
UMAP1
P_ISN < 5.08e-03
Correlation
-0.25
0.25
GLOM4
GLOM2
GLOM3
GLOM1
-0.35
0.00
Neighborhood Correlation
c
-Log10(P-value) Before and After Correction
-Log10(P-value)
B/Plasma
DN
Glomerular
EC/STR
LOH
Myeloid
PT
T/NK
Correction
None
Site + First Biopsy
Site + First Biopsy + Chronicity
Site + First Biopsy + Chronicity + Prednisone Use
Site + First Biopsy + Chronicity + Black Race
Site + First Biopsy + Chronicity + Activity
d
Glomerular
ISN V CNA
R = −0.902
ISN V CNA, Activity Corrected
R = −0.259
ISN CNA + Standard Covariates + Activity
R = 0.65
e
Myeloid
ISN V CNA
R = −0.884
Activity CNA
ISN V CNA, Activity Corrected
R = −0.374
Activity CNA
ISN CNA + Standard Covariates + Activity
R = 0.762
ISN CNA + Standard Covariates

Extended Data Figure 9. a-b, CNA testing for associations with ISN class V, correcting for site, first biopsy, and chronicity index for myeloid **(a)** and glomerular **(b)** compartments. **c,** Global p values for CNA testing of ISN class V accounting for either no or additional relevant clinical covariates. **d-e,** Scatterplots showing per-neighborhood correlations between activity index CNA and either ISN class V CNA either without (left column) or with (middle column) correcting for activity index. Additional scatterplots plot the correlation between ISN class V CNA without correcting for activity (x-axis) or with correction (y-axis). These plots were done for both the glomeruli **(d)** and myeloid **(e)** compartments. All CNA for d and e include major collection site, first biopsy status, and chronicity as covariates, along with any other specified covariates.


a
B Cells; Case Control
UMAP2
UMAP1
P_Case < 7.00e-05
Correlation
-0.37
0.37
bl−B4
bl−B2
bl−B7
bl−B8
bl−B5
bl−B3
bl−B6
bl−B1
bl−B0
−0.61
0.00
Neighborhood Correlation
b
TNK; Case Control
UMAP2
UMAP1
P_Case < 5.00e-05
Correlation
-0.33
0.33
bl−T19
bl−T15
bl−T14
bl−T18
bl−T17
bl−NK1
bl−T6
bl−NK2
bl−T16
bl−T4
bl−T3
bl−T13
bl−NK0
bl−T9
bl−T8
bl−T11
bl−T7
bl−T10
bl−T5
bl−T12
−0.69
0.00
Neighborhood Correlation
c
Myeloid; Case Control
UMAP2
UMAP1
P_Case < 3.00e-05
Correlation
-0.32
0.32
bl−M6
bl−M1
bl−M8
bl−M7
bl−M0
bl−M4
bl−P9
bl−M2
bl−M3
bl−M5
bl−DC12
bl−DC10
bl−DC11
−0.62
0.00
Neighborhood Correlation
e
TNK; Chronicity
UMAP2
UMAP1
P_chronicity < 1.90e-03
Correlation
-0.48
0.48
bl−T7
bl−T8
bl−T10
bl−T3
bl−T4
bl−T5
bl−T11
bl−T13
bl−T9
bl−T12
bl−T6
bl−T16
bl−NK1
bl−NK0
bl−NK2
bl−T18
bl−T14
bl−T19
bl−T17
bl−T15
0.00
0.48
Neighborhood Correlation
d
Gene
IFI27
IFI44L
USP18
IFI6
SPATS2L
IFITM1
LGALS3BP
OASL
IFI44
EPSTI1
RPS3
RPL3
RBPMS
AC025171.3
LINC02812
PAIP2B
ACSM3
LINC00665
ZNF154
ZNF418
GPM6B
B/Plasma
Myeloid
T/NK
Cell Type
Beta Case-Control
-3.45
3.45
f
T/NK
OAS3
USP18
HERC5
ISG15
EPSTI1
EBF4
MT-CYB
-Log10(P-value)
Beta Chronicity
B/Plasma
EIF2AK2
IFITM1
PLSCR1
LAP3
ANKRD39
MT-ATP6
MT-CYB
Beta Chronicity
Myeloid
Z95114.4
LINC02356
Beta Chronicity
g
T/NK
AL080276.2
MIOS
Beta Activity
B/Plasma
Beta Activity
Myeloid
VSIG4
Beta Activity

Extended Data Figure 10. a-c, CNA testing for case-control status for B **(a)**, T and NK **(b)**, and Myeloid **(c)** cells, accounting for age and sex as potential covariates. **d,** For each cell type, differential expression for genes associated with case-control status in blood. Dotted lines separate genes enriched in most cell types and depleted in most cell types. **e,** CNA testing for chronicity index in the T and NK compartment, accounting for site and first biopsy site as covariates. **f-g,** Differentially expressed genes for each cell type with respect to chronicity index **(f)** and activity index **(g)**. Genes upregulated with respect to each index are displayed in red, downregulated genes are displayed in blue. Horizontal lines indicate the Bonferonni corrected significance threshold for $p<0.05$. The top 5 most significant genes in each direction are labeled for each cell type.

Supplementary Note

Tissue Cell Type Classification. The Supplementary Note contains detailed figures for each of the eight broad cell types identified in sequenced renal tissue: (1) T/NK (**Supplementary Fig. 1**), (2) B/Plasma (**Supplementary Fig. 2**), (3) Myeloid (**Supplementary Fig. 3**), (4) Proximal Tubule (**Supplementary Fig. 4**), (5) Loop of Henle (**Supplementary Fig. 5**), (6) Distal Nephron (**Supplementary Fig. 6**), (7) Interstitial/Stromal (**Supplementary Fig. 7**), (8) Glomerular (**Supplementary Fig. 8**). For each figure, we show the integrated analysis of sc and sn data. We also include depiction of selected key marker genes across the different cell states. We present analyses depicting cell state concordance between sc and sn data modalities. In each figure, panels **a** and **b** depict cell states of cells from sc and sn modalities respectively, as visualized via UMAP. Similar coverage of cell states across modalities indicates cell states that are shared between data modalities (i.e. **Supplementary Fig. 1a-b**), while differential coverage of cell states between modalities may either reflect differences in processing of sc and sn data, or biological differences between cytosolic and nuclear cell states. Similarly, panel **c** reflects concordance in the proportions of each cell state between the two modalities. For each cell type, we calculated the proportions of each cell state between the two modalities, in samples with at least X cells. In panel **d**, we highlighted the expression of key cell type markers driving transcriptional clustering on the UMAP. Finally, we plotted key cell state markers for fine grain clusters for both scRNA-seq (**e**) and snRNA-seq (**f**). Generally, we observed that cell states have similar marker genes in sc and sn data. Similar marker expression patterns between sc and sn suggest concordance between annotated cell states in both modalities.

PBMC Classification. To understand the relationship between tissue and blood immune cell states, we also sequenced 327,326 PBMCs from 119 LN patients and 19 healthy controls (**Extended Data Fig. 1a**). We defined broad cell types using marker genes such as *CD3D* for T cells, *CD14* for monocytes, *MS4A1* for B cells, *NCAM1* for NK cells, and *CD1C* for dendritic

cells (**Extended Data Fig 1b**). We obtained 114,970 T cells, 52,537 NK Cells, 100,928 monocytes, 5,508 dendritic cells, and 53,383 B Cells (**Extended Data Fig. 1a**). Clustering within T and NK Cells obtained 20 clusters (**Supplementary Table 3**) spanning CD56^{bright} and CD56^{dim} NK Cells (bl-NK0, bl-NK5, bl-NK6), GZMK+CD8+ (bl-T1, bl-T2, bl-T14, bl-T15), and gamma/delta T cells (bl-T11, bl-T12). We additionally used key marker genes to designate central memory/naive populations (bl-T8, bl-T10), effector memory populations (bl-T9, bl-T13), GZMK+TEMRA (bl-T7), CD8+GZMB+proliferating CTLs (bl-T15, bl-T18, bl-T19), T^{reg} (bl-T16), and an interferon-stimulated gene (ISG) high cell state (bl-T17) (**Extended Data Fig. 1c-d, Supplementary Table 3**). Similarly, we clustered within monocytes and dendritic cells and found 13 clusters capturing populations of CD14+CD16- conventional monocytes (bl-M0, bl-M1, bl-M2, bl-M5), CD14+ CD16+ intermediate monocytes (bl-M4, bl-M11), and CD14^{dim} CD16++ non-conventional monocytes (bl-M3) (**Extended Data Fig. 1e-f, Supplementary Table 3**). We also found a cluster of ISG^{high} monocytes (bl-M7), CD1C+ DC2 dendritic cells (bl-DC6), CLEC9A+ XCR1+ DC1 dendritic cells (bl-DC12), TCF4+ CLEC4C+ pDC dendritic cells (bl-DC9), and platelet cells (bl-P8). We finally clustered B Cells to obtain 9 specialized substates, including *IGHM*+ or *IGHD*+ naïve B Cells (bl-B0, bl-B4) and *FCRL5*+ *ITGAX*+ auto-immune associated B Cells (ABCs) (bl-B2, bl-B8). (**Extended Data Fig. 1g-h, Supplementary Table 3**). We also found a population of mature B cells expressing low *IGHM* and *IGHD* but high *CRIP1* (bl-B1), a population of *CD79*+*VPREB3*+ pre–B Cells (bl-B3), and *JUN*+*NKFB1*+activated B cells (bl-B6).

Differential gene expression between blood and tissue.

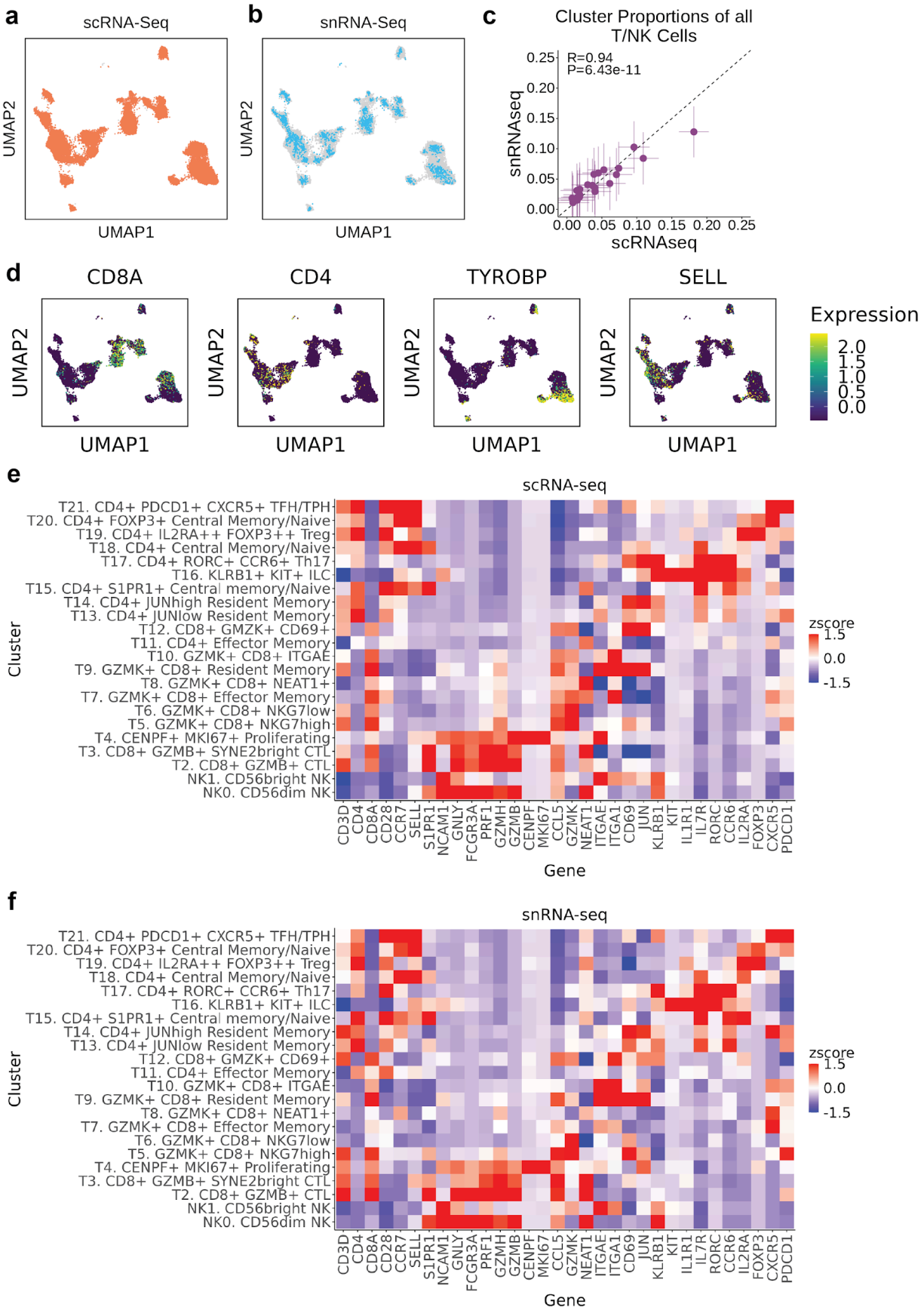
Differential gene expression between donor-matched tissue and PBMC samples revealed an upregulation of plasma cell markers (*IGH3*, *IGLC2*, *IGHCHG1*) in tissue (**Extended Data Fig. 2a, b**). In the B/plasma clusters identified in tissue, the tissue-specific (high TSM) plasma clusters were enriched for these differentially expressed tissue markers. Conversely, the shared

cell states (TSM around 0.5) had comparatively higher expression of genes differentially expressed by PBMCs.

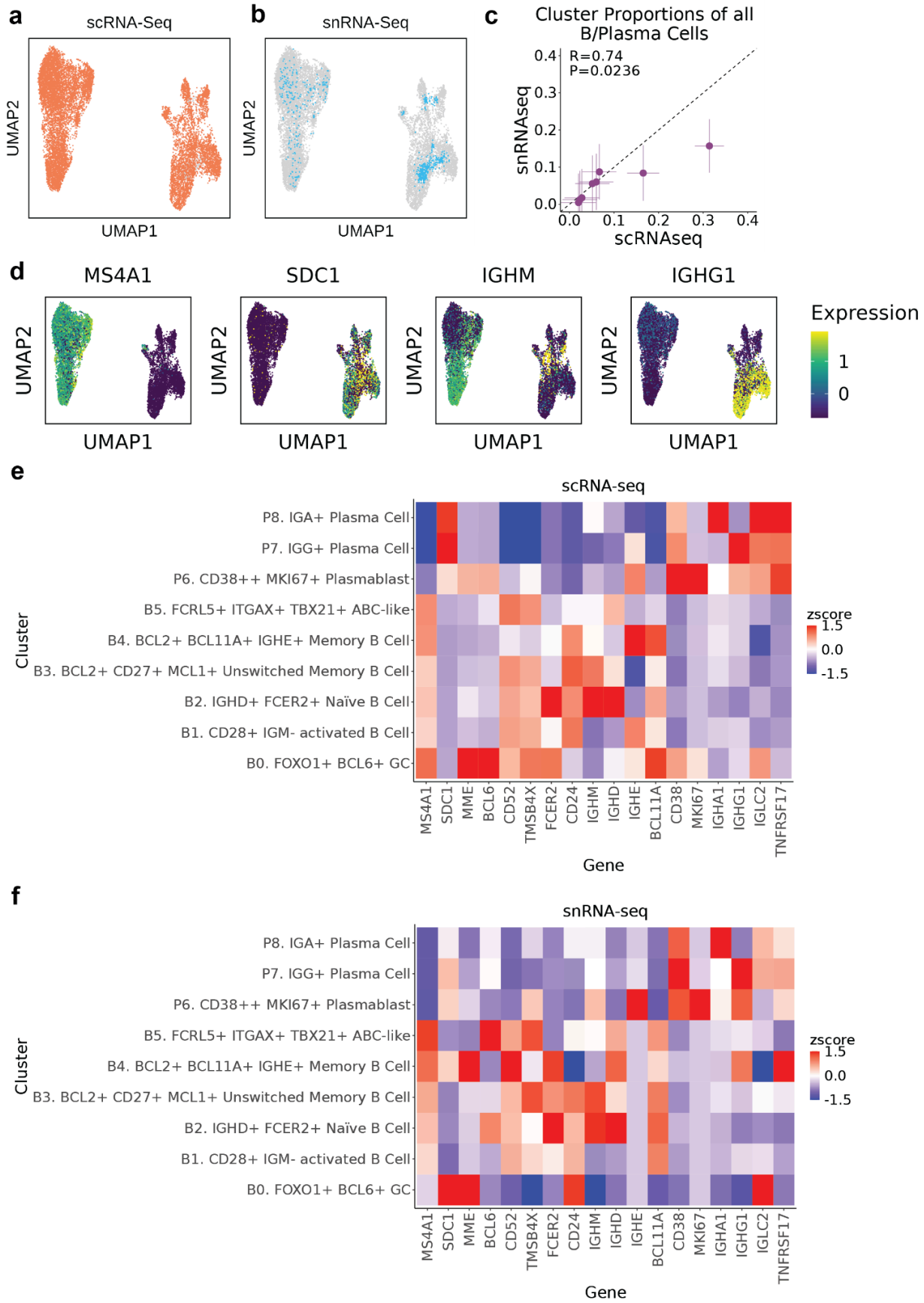
Differential gene expression in Myeloid cells revealed similar patterns. In tissue, genes associated with macrophages (*C1QA*, *C1QB*, *C1QC*) and anti-inflammatory properties (*GPNMB*) were upregulated (**Extended Data Fig. 2c, d**). Markers high in blood monocytes (*SELL*, *LYZ*) and pro-inflammatory properties (*S100A8*, *S100A9*) were downregulated. Within the tissue compartment, tissue-specific Myeloid clusters, including all macrophage states, were relatively enriched for the differentially expressed tissue markers. Shared monocyte states were relatively enriched for differentially expressed blood markers.

Differential gene expression in T/NK cells revealed increased T cell activation in tissue, indicated by upregulation of activation markers (*CD40LG*, *PDCD1*) and cytokine and chemokine secretion (*IFNG*, *TNF*, *CCL3*) (**Extended Data Fig. 2e, f**). T cells from tissue were also significantly enriched for markers of tissue-residence, including *ITGA1*, *CXCR6*, and *ZNF683* (nominally). While tissue markers were expressed across most renal T cell states, differentially expressed blood markers were relatively enriched in shared T cell states, including CD56dim NK (NK0) and CD8+GZMB+SYNE2bright CTL (T3) cells.

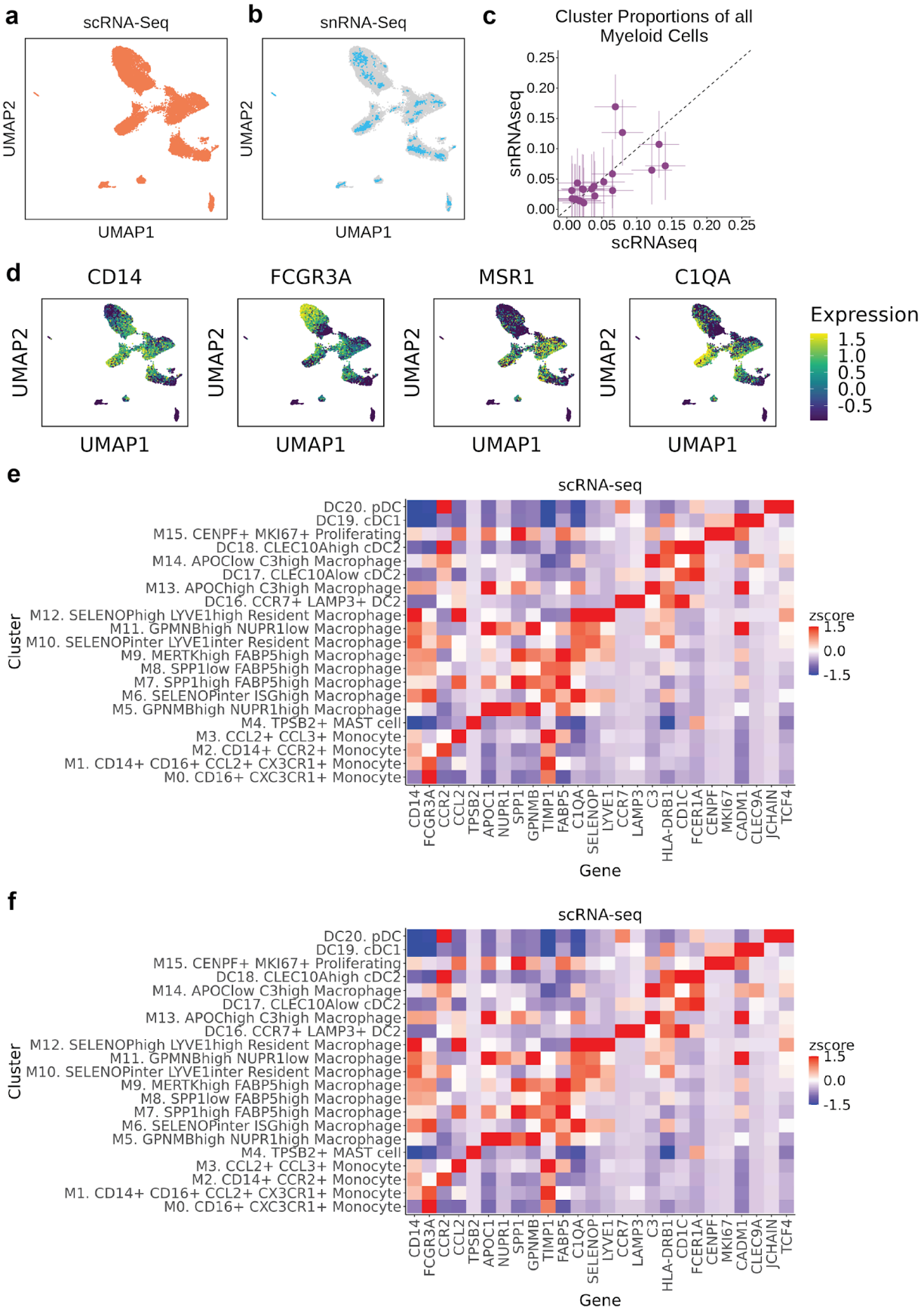
Supplemental Figures / Legends


a
scRNA-Seq
UMAP2
UMAP1
b
snRNA-Seq
UMAP2
UMAP1
c
Cluster Proportions of all T/NK Cells
R=0.94
P=6.43e-11
snRNAseq
scRNAseq
0.00
0.05
0.10
0.15
0.20
0.25
d
CD8A
CD4
TYROBP
SELL
UMAP2
UMAP1
Expression
2.0
1.5
1.0
0.5
0.0
e
scRNA-seq
Cluster
T21. CD4+ PDCD1+ CXCR5+ TFH/TPH
T20. CD4+ FOXP3+ Central Memory/Naive
T19. CD4+ IL2RA++ FOXP3++ Treg
T18. CD4+ Central Memory/Naive
T17. CD4+ RORC+ CCR6+ Th17
T16. KLRB1+ KIT+ ILC
T15. CD4+ S1PR1+ Central memory/Naive
T14. CD4+ JUNhigh Resident Memory
T13. CD4+ JUNlow Resident Memory
T12. CD8+ GMZK+ CD69+
T11. CD4+ Effector Memory
T10. GZMK+ CD8+ ITGAE
T9. GZMK+ CD8+ Resident Memory
T8. GZMK+ CD8+ NEAT1+
T7. GZMK+ CD8+ Effector Memory
T6. GZMK+ CD8+ NKG7low
T5. GZMK+ CD8+ NKG7high
T4. CENPF+ MKI67+ Proliferating
T3. CD8+ GZMB+ SYNE2bright CTL
T2. CD8+ GZMB+ CTL
NK1. CD56bright NK
NK0. CD56dim NK
CD3D
CD4
CD8A
CD28
CCR7
SELL
S1PR1
NCAM1
GNLY
FCGR3A
PRF1
GZMH
GZMB
CENPF
MKI67
CCL5
GZMK
NEAT1
ITGAE
ITGA1
CD69
JUN
KLRB1
KIT
IL1R1
IL7R
RORC
CCR6
IL2RA
FOXP3
CXCR5
PDCD1
Gene
zscore
1.5
0.0
-1.5
f
snRNA-seq
Cluster
T21. CD4+ PDCD1+ CXCR5+ TFH/TPH
T20. CD4+ FOXP3+ Central Memory/Naive
T19. CD4+ IL2RA++ FOXP3++ Treg
T18. CD4+ Central Memory/Naive
T17. CD4+ RORC+ CCR6+ Th17
T16. KLRB1+ KIT+ ILC
T15. CD4+ S1PR1+ Central memory/Naive
T14. CD4+ JUNhigh Resident Memory
T13. CD4+ JUNlow Resident Memory
T12. CD8+ GMZK+ CD69+
T11. CD4+ Effector Memory
T10. GZMK+ CD8+ ITGAE
T9. GZMK+ CD8+ Resident Memory
T8. GZMK+ CD8+ NEAT1+
T7. GZMK+ CD8+ Effector Memory
T6. GZMK+ CD8+ NKG7low
T5. GZMK+ CD8+ NKG7high
T4. CENPF+ MKI67+ Proliferating
T3. CD8+ GZMB+ SYNE2bright CTL
T2. CD8+ GZMB+ CTL
NK1. CD56bright NK
NK0. CD56dim NK
CD3D
CD4
CD8A
CD28
CCR7
SELL
S1PR1
NCAM1
GNLY
FCGR3A
PRF1
GZMH
GZMB
CENPF
MKI67
CCL5
GZMK
NEAT1
ITGAE
ITGA1
CD69
JUN
KLRB1
KIT
IL1R1
IL7R
RORC
CCR6
IL2RA
FOXP3
CXCR5
PDCD1
Gene
zscore
1.5
0.0
-1.5

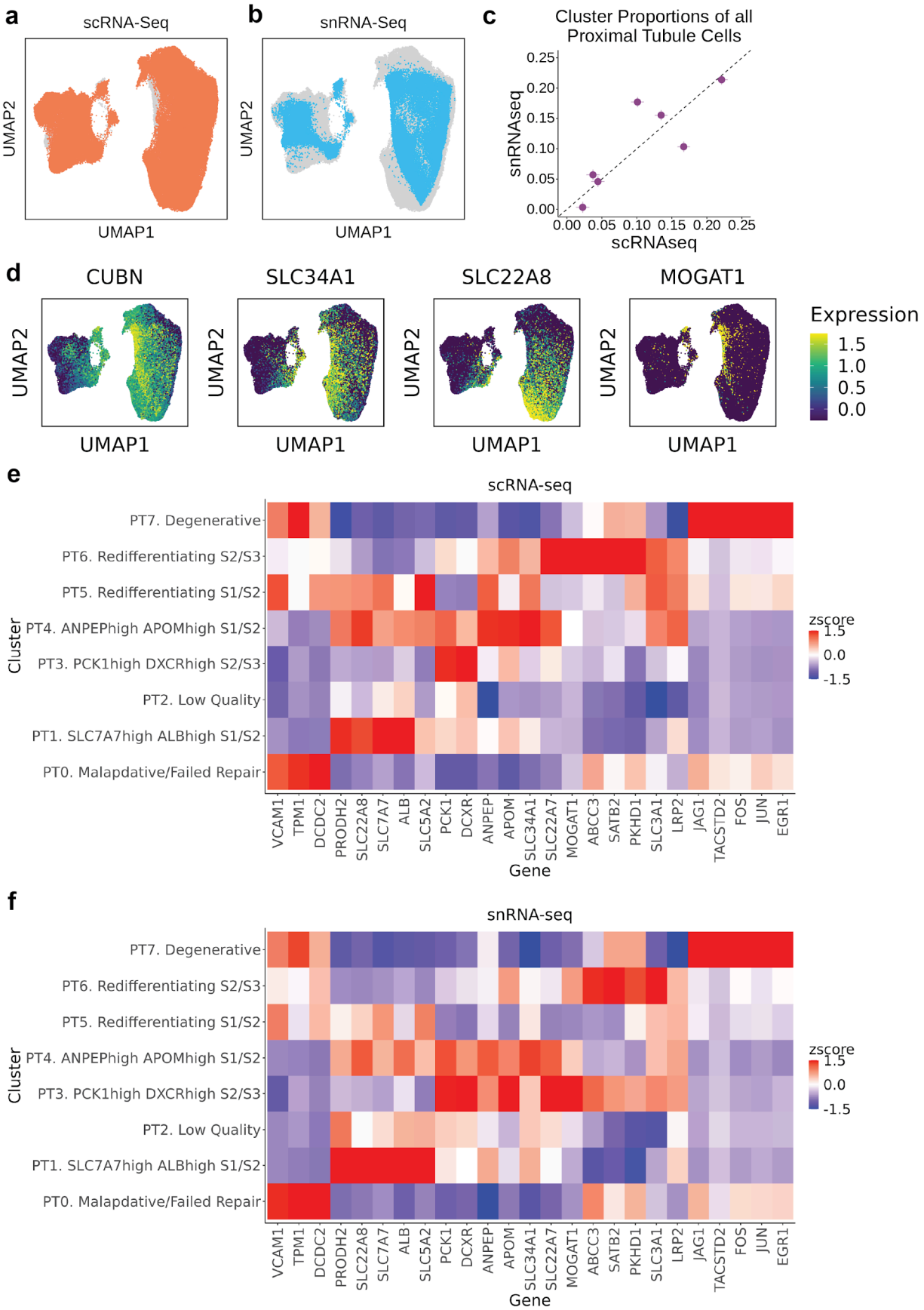
Supplementary Figure 1. T/NK. a, UMAP depiction of cell states captured from scRNA-seq. **b,** UMAP depiction of cells from snRNA-seq. **c,** Scatterplot displaying each cell state cluster's proportion out of total T/NK cells. X-axis reflects cell state proportions calculated in scRNA-seq and Y-axis reflects snRNA-seq. Proportions were calculated as the proportion of total cell state annotations across all donors with both matched scRNA-seq and snRNA-seq (n=22). Error bars reflect the standard error of that cluster's cell state proportion across donors, where horizontal lines reflect per-sample variability in scRNA-seq proportions, and vertical lines reflect per-sample variability in snRNA-seq proportions. **d,** Expression of marker genes (log-normalized, Z-scored) across all T/NK cells. **e-f,** Heatmaps depicting average expression of marker genes (log-normalized, Z-scored after averaging) within each cell state annotation. Shown in scRNA-seq **e** and snRNA-seq **f**.


a
scRNA-Seq
UMAP2
UMAP1
b
snRNA-Seq
UMAP2
UMAP1
c
Cluster Proportions of all B/Plasma Cells
R=0.74
P=0.0236
snRNAseq
scRNAseq
0.0
0.1
0.2
0.3
0.4
d
MS4A1
SDC1
IGHM
IGHG1
UMAP2
UMAP1
Expression
1
0
e
scRNA-seq
Cluster
P8. IGA+ Plasma Cell
P7. IGG+ Plasma Cell
P6. CD38++ MKI67+ Plasmablast
B5. FCRL5+ ITGAX+ TBX21+ ABC-like
B4. BCL2+ BCL11A+ IGHE+ Memory B Cell
B3. BCL2+ CD27+ MCL1+ Unswitched Memory B Cell
B2. IGHD+ FCER2+ Naïve B Cell
B1. CD28+ IGM- activated B Cell
B0. FOXO1+ BCL6+ GC
MS4A1
SDC1
MME
BCL6
CD52
TMSB4X
FCER2
CD24
IGHM
IGHD
IGHE
BCL11A
CD38
MKI67
IGHA1
IGHG1
IGLC2
TNFRSF17
Gene
zscore
1.5
0.0
-1.5
f
snRNA-seq
Cluster
P8. IGA+ Plasma Cell
P7. IGG+ Plasma Cell
P6. CD38++ MKI67+ Plasmablast
B5. FCRL5+ ITGAX+ TBX21+ ABC-like
B4. BCL2+ BCL11A+ IGHE+ Memory B Cell
B3. BCL2+ CD27+ MCL1+ Unswitched Memory B Cell
B2. IGHD+ FCER2+ Naïve B Cell
B1. CD28+ IGM- activated B Cell
B0. FOXO1+ BCL6+ GC
MS4A1
SDC1
MME
BCL6
CD52
TMSB4X
FCER2
CD24
IGHM
IGHD
IGHE
BCL11A
CD38
MKI67
IGHA1
IGHG1
IGLC2
TNFRSF17
Gene
zscore
1.5
0.0
-1.5

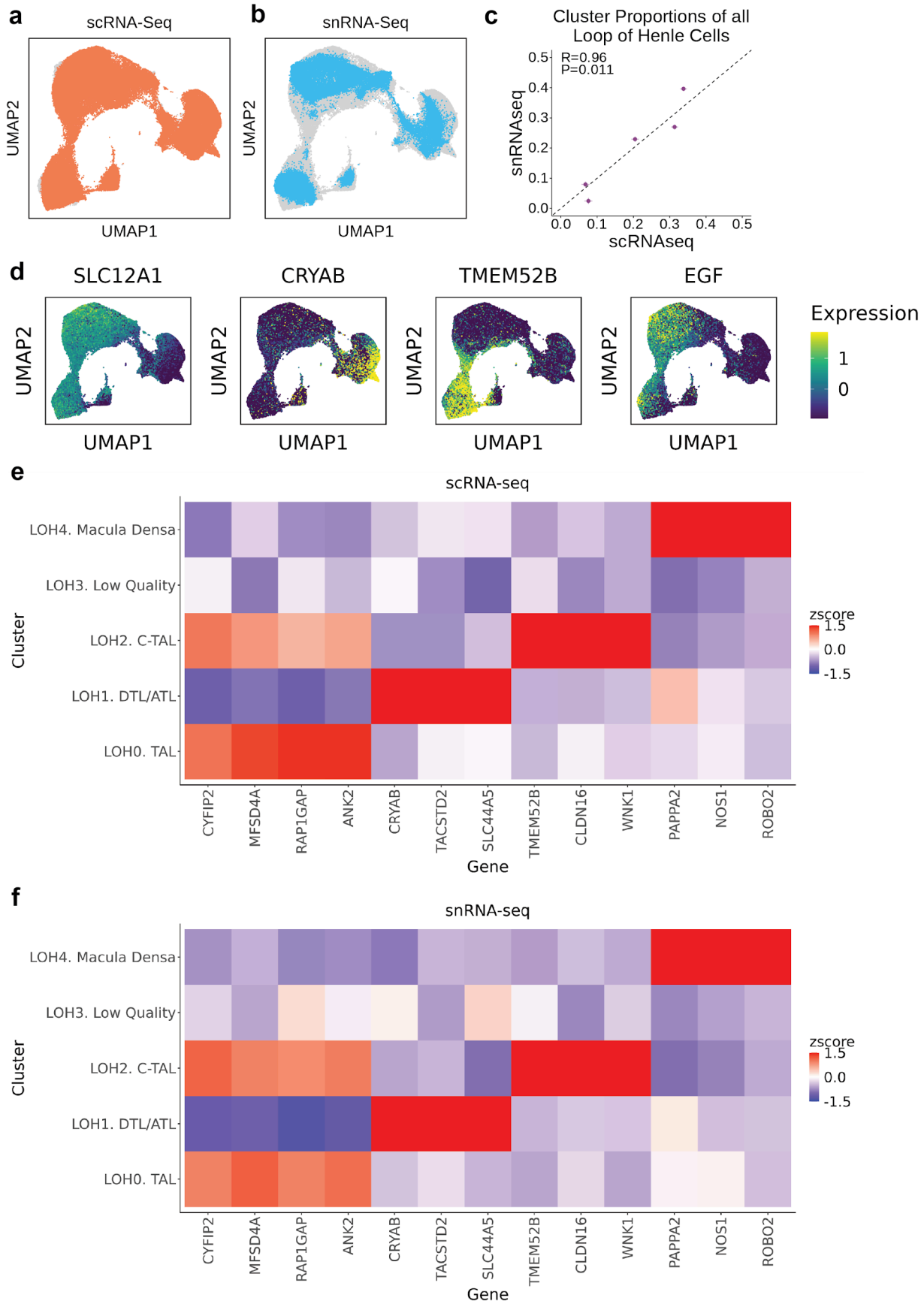
Supplementary Figure 2. B/Plasma. a, UMAP depiction of cell states captured from scRNA-seq. **b,** UMAP depiction of cells from snRNA-seq. **c,** Scatterplot displaying each cell state cluster's proportion out of total B/Plasma cells. X-axis reflects cell state proportions calculated in scRNA-seq and Y-axis reflects snRNA-seq. Proportions were calculated as the proportion of total cell state annotations across all donors with both matched scRNA-seq and snRNA-seq (n=22). Error bars reflect the standard error of that cluster's cell state proportion across donors, where horizontal lines reflect per-sample variability in scRNA-seq proportions, and vertical lines reflect per-sample variability in snRNA-seq proportions. **d,** Expression of marker genes (log-normalized, Z-scored) across all B/Plasma cells. **e-f,** Heatmaps depicting average expression of marker genes (log-normalized, Z-scored after averaging) within each cell state annotation. Shown in scRNA-seq **e** and snRNA-seq **f**.


a
scRNA-Seq
UMAP2
UMAP1
b
snRNA-Seq
UMAP2
UMAP1
c
Cluster Proportions of all Myeloid Cells
snRNAseq
scRNAseq
0.00
0.05
0.10
0.15
0.20
0.25
d
CD14
FCGR3A
MSR1
C1QA
UMAP2
UMAP1
Expression
1.5
1.0
0.5
0.0
-0.5
e
scRNA-seq
Cluster
DC20. pDC
DC19. cDC1
M15. CENPF+ MKI67+ Proliferating
DC18. CLEC10Ahigh cDC2
M14. APOClow C3high Macrophage
DC17. CLEC10Alow cDC2
M13. APOChigh C3high Macrophage
DC16. CCR7+ LAMP3+ DC2
M12. SELENOPhigh LYVE1high Resident Macrophage
M11. GPMNBhigh NUPR1low Macrophage
M10. SELENOPinter LYVE1inter Resident Macrophage
M9. MERTKhigh FABP5high Macrophage
M8. SPP1low FABP5high Macrophage
M7. SPP1high FABP5high Macrophage
M6. SELENOPinter ISGhigh Macrophage
M5. GPNMBhigh NUPR1high Macrophage
M4. TPSB2+ MAST cell
M3. CCL2+ CCL3+ Monocyte
M2. CD14+ CCR2+ Monocyte
M1. CD14+ CD16+ CCL2+ CX3CR1+ Monocyte
M0. CD16+ CXC3CR1+ Monocyte
CD14
FCGR3A
CCR2
CCL2
TPSB2
APOC1
NUPR1
SPP1
GPNMB
TIMP1
FABP5
C1QA
SELENOP
LYVE1
CCR7
LAMP3
C3
HLA-DRB1
CD1C
FCER1A
CENPF
MKI67
CADM1
CLEC9A
JCHAIN
TCF4
Gene
zscore
1.5
0.0
-1.5
f
scRNA-seq
Cluster
DC20. pDC
DC19. cDC1
M15. CENPF+ MKI67+ Proliferating
DC18. CLEC10Ahigh cDC2
M14. APOClow C3high Macrophage
DC17. CLEC10Alow cDC2
M13. APOChigh C3high Macrophage
DC16. CCR7+ LAMP3+ DC2
M12. SELENOPhigh LYVE1high Resident Macrophage
M11. GPMNBhigh NUPR1low Macrophage
M10. SELENOPinter LYVE1inter Resident Macrophage
M9. MERTKhigh FABP5high Macrophage
M8. SPP1low FABP5high Macrophage
M7. SPP1high FABP5high Macrophage
M6. SELENOPinter ISGhigh Macrophage
M5. GPNMBhigh NUPR1high Macrophage
M4. TPSB2+ MAST cell
M3. CCL2+ CCL3+ Monocyte
M2. CD14+ CCR2+ Monocyte
M1. CD14+ CD16+ CCL2+ CX3CR1+ Monocyte
M0. CD16+ CXC3CR1+ Monocyte
CD14
FCGR3A
CCR2
CCL2
TPSB2
APOC1
NUPR1
SPP1
GPNMB
TIMP1
FABP5
C1QA
SELENOP
LYVE1
CCR7
LAMP3
C3
HLA-DRB1
CD1C
FCER1A
CENPF
MKI67
CADM1
CLEC9A
JCHAIN
TCF4
Gene
zscore
1.5
0.0
-1.5

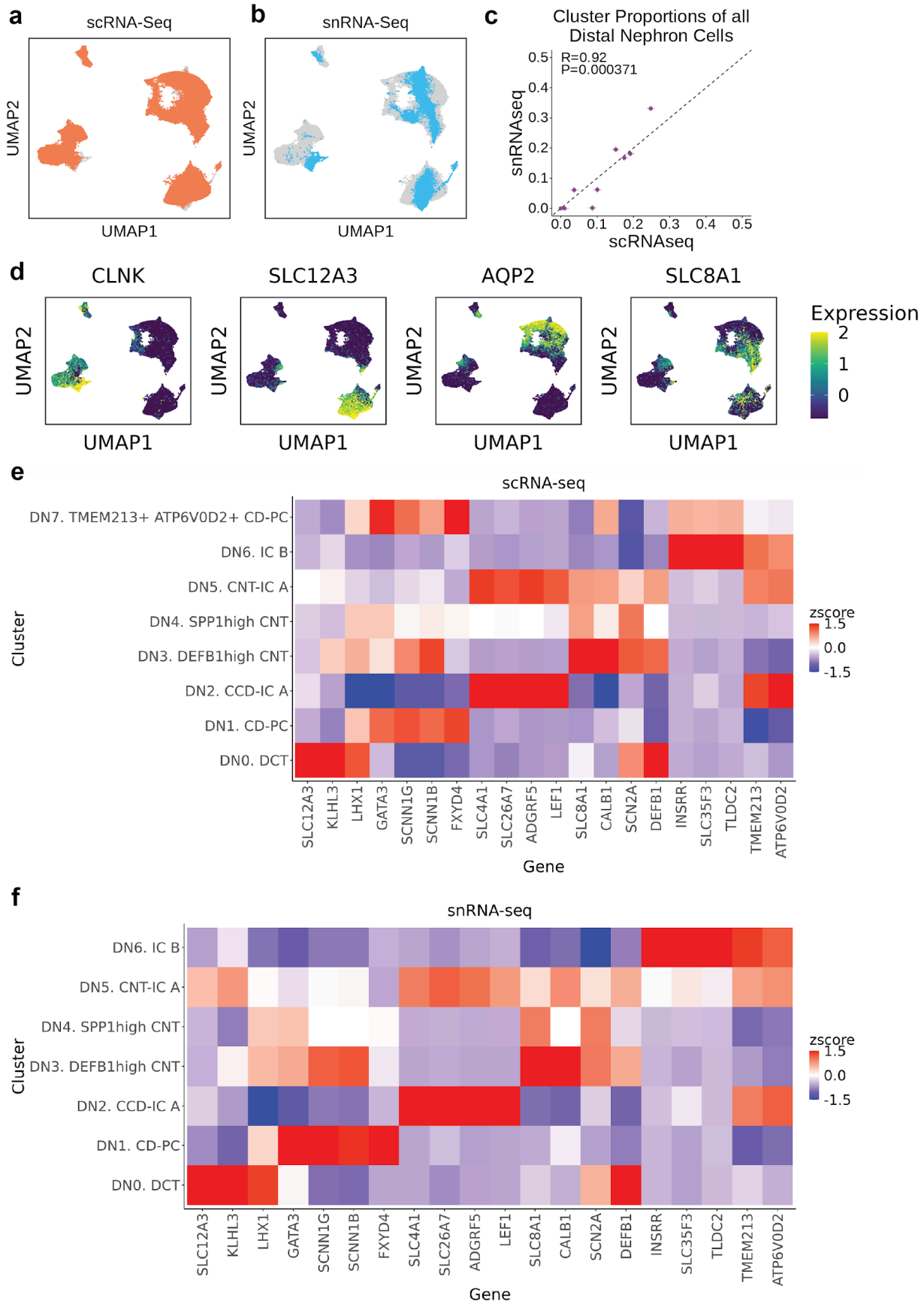
Supplementary Figure 3. Myeloid. a, UMAP depiction of cell states captured from scRNA-seq. **b,** UMAP depiction of cells from snRNA-seq. **c,** Scatterplot displaying each cell state cluster's proportion out of total Myeloid cells. X-axis reflects cell state proportions calculated in scRNA-seq and Y-axis reflects snRNA-seq. Proportions were calculated as the proportion of total cell state annotations across all donors with both matched scRNA-seq and snRNA-seq (n=22). Error bars reflect the standard error of that cluster's cell state proportion across donors, where horizontal lines reflect per-sample variability in scRNA-seq proportions, and vertical lines reflect per-sample variability in snRNA-seq proportions. **d,** Expression of marker genes (log-normalized, Z-scored) across all Myeloid cells. **e-f,** Heatmaps depicting average expression of marker genes (log-normalized, Z-scored after averaging) within each cell state annotation. Shown in scRNA-seq **e** and snRNA-seq **f**.


a
scRNA-Seq
UMAP2
UMAP1
b
snRNA-Seq
UMAP2
UMAP1
c
Cluster Proportions of all Proximal Tubule Cells
snRNAseq
scRNAseq
0.00
0.05
0.10
0.15
0.20
0.25
d
CUBN
SLC34A1
SLC22A8
MOGAT1
UMAP2
UMAP1
Expression
1.5
1.0
0.5
0.0
e
scRNA-seq
Cluster
PT7. Degenerative
PT6. Redifferentiating S2/S3
PT5. Redifferentiating S1/S2
PT4. ANPEPhigh APOMhigh S1/S2
PT3. PCK1high DXCRhigh S2/S3
PT2. Low Quality
PT1. SLC7A7high ALBhigh S1/S2
PT0. Malapdative/Failed Repair
VCAM1
TPM1
DCDC2
PRODH2
SLC22A8
SLC7A7
ALB
SLC5A2
PCK1
DCXR
ANPEP
APOM
SLC34A1
SLC22A7
MOGAT1
ABCC3
SATB2
PKHD1
SLC3A1
LRP2
JAG1
TACSTD2
FOS
JUN
EGR1
Gene
zscore
1.5
0.0
-1.5
f
snRNA-seq
Cluster
PT7. Degenerative
PT6. Redifferentiating S2/S3
PT5. Redifferentiating S1/S2
PT4. ANPEPhigh APOMhigh S1/S2
PT3. PCK1high DXCRhigh S2/S3
PT2. Low Quality
PT1. SLC7A7high ALBhigh S1/S2
PT0. Malapdative/Failed Repair
VCAM1
TPM1
DCDC2
PRODH2
SLC22A8
SLC7A7
ALB
SLC5A2
PCK1
DCXR
ANPEP
APOM
SLC34A1
SLC22A7
MOGAT1
ABCC3
SATB2
PKHD1
SLC3A1
LRP2
JAG1
TACSTD2
FOS
JUN
EGR1
Gene
zscore
1.5
0.0
-1.5

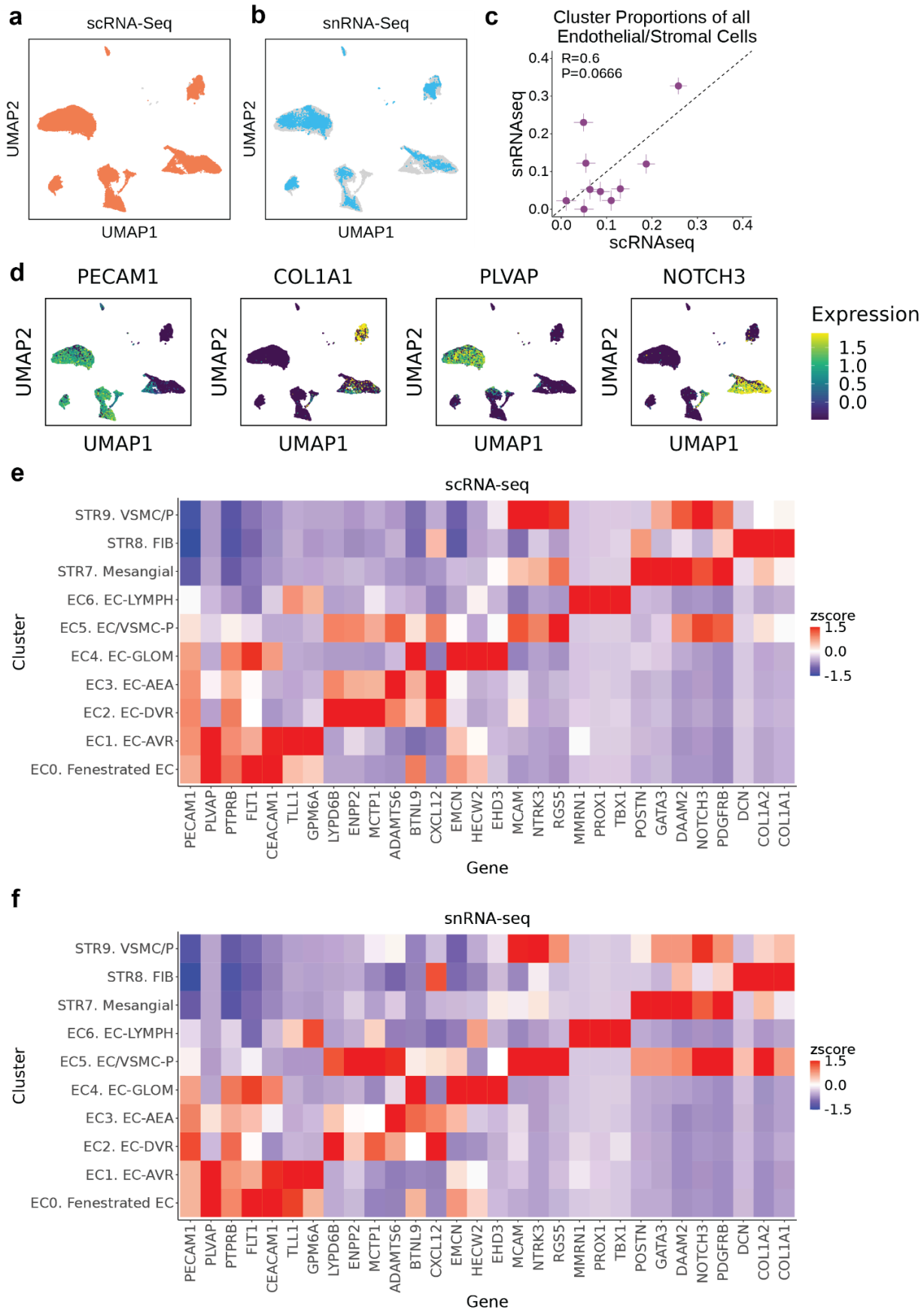
Supplementary Figure 4. Proximal Tubule. a, UMAP depiction of cell states captured from scRNA-seq. **b,** UMAP depiction of cells from snRNA-seq. **c,** Scatterplot displaying each cell state cluster's proportion out of total Proximal Tubule cells. X-axis reflects cell state proportions calculated in scRNA-seq and Y-axis reflects snRNA-seq. Proportions were calculated as the proportion of total cell state annotations across all donors with both matched scRNA-seq and snRNA-seq (n=22). Error bars reflect the standard error of that cluster's cell state proportion across donors, where horizontal lines reflect per-sample variability in scRNA-seq proportions, and vertical lines reflect per-sample variability in snRNA-seq proportions. **d,** Expression of marker genes (log-normalized, Z-scored) across all Proximal Tubule cells. **e-f,** Heatmaps depicting average expression of marker genes (log-normalized, Z-scored after averaging) within each cell state annotation. Shown in scRNA-seq **e** and snRNA-seq **f**.


a
scRNA-Seq
UMAP2
UMAP1
b
snRNA-Seq
UMAP2
UMAP1
c
Cluster Proportions of all Loop of Henle Cells
R=0.96
P=0.011
snRNAseq
scRNAseq
0.0
0.1
0.2
0.3
0.4
0.5
d
SLC12A1
CRYAB
TMEM52B
EGF
UMAP2
UMAP1
Expression
1
0
e
scRNA-seq
LOH4. Macula Densa
LOH3. Low Quality
LOH2. C-TAL
LOH1. DTL/ATL
LOH0. TAL
Cluster
zscore
1.5
0.0
-1.5
CYFIP2
MFSD4A
RAP1GAP
ANK2
CRYAB
TACSTD2
SLC44A5
TMEM52B
CLDN16
WNK1
PAPPA2
NOS1
ROBO2
Gene
f
snRNA-seq
LOH4. Macula Densa
LOH3. Low Quality
LOH2. C-TAL
LOH1. DTL/ATL
LOH0. TAL
Cluster
zscore
1.5
0.0
-1.5
CYFIP2
MFSD4A
RAP1GAP
ANK2
CRYAB
TACSTD2
SLC44A5
TMEM52B
CLDN16
WNK1
PAPPA2
NOS1
ROBO2
Gene

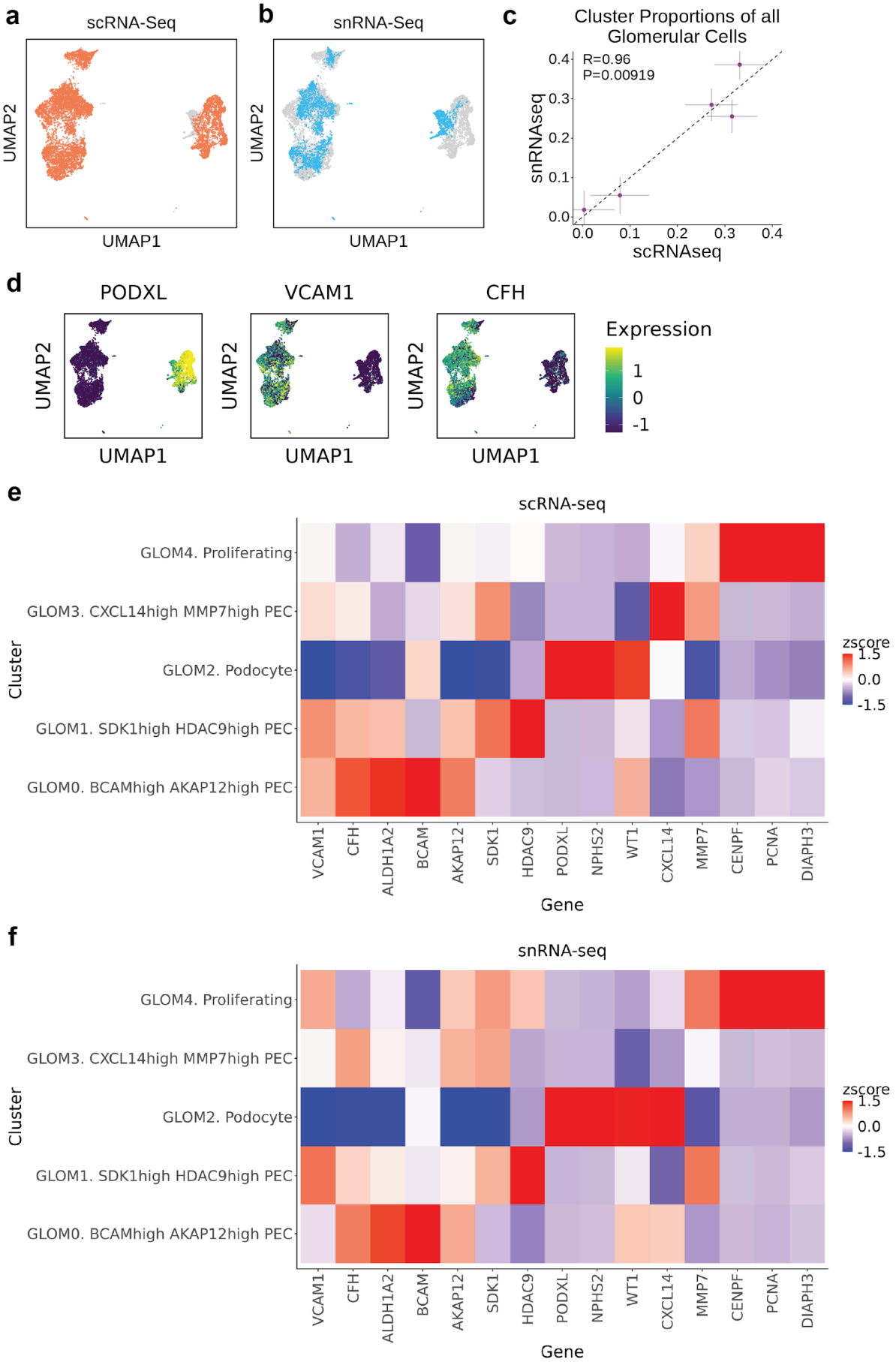
Supplementary Figure 5. Loop of Henle. a, UMAP depiction of cell states captured from scRNA-seq. **b,** UMAP depiction of cells from snRNA-seq. **c,** Scatterplot displaying each cell state cluster's proportion out of total Loop of Henle cells. X-axis reflects cell state proportions calculated in scRNA-seq and Y-axis reflects snRNA-seq. Proportions were calculated as the proportion of total cell state annotations across all donors with both matched scRNA-seq and snRNA-seq (n=22). Error bars reflect the standard error of that cluster's cell state proportion across donors, where horizontal lines reflect per-sample variability in scRNA-seq proportions, and vertical lines reflect per-sample variability in snRNA-seq proportions. **d,** Expression of marker genes (log-normalized, Z-scored) across all Loop of Henle cells. **e-f,** Heatmaps depicting average expression of marker genes (log-normalized, Z-scored after averaging) within each cell state annotation. Shown in scRNA-seq **e** and snRNA-seq **f**.


a
scRNA-Seq
UMAP2
UMAP1
b
snRNA-Seq
UMAP2
UMAP1
c
Cluster Proportions of all Distal Nephron Cells
R=0.92
P=0.000371
snRNAseq
scRNAseq
0.0
0.1
0.2
0.3
0.4
0.5
d
CLNK
SLC12A3
AQP2
SLC8A1
UMAP2
UMAP1
Expression
2
1
0
e
scRNA-seq
Cluster
DN7. TMEM213+ ATP6V0D2+ CD-PC
DN6. IC B
DN5. CNT-IC A
DN4. SPP1high CNT
DN3. DEFB1high CNT
DN2. CCD-IC A
DN1. CD-PC
DN0. DCT
SLC12A3
KLHL3
LHX1
GATA3
SCNN1G
SCNN1B
FXYD4
SLC4A1
SLC26A7
ADGRF5
LEF1
SLC8A1
CALB1
SCN2A
DEFB1
INSRR
SLC35F3
TLDC2
TMEM213
ATP6V0D2
Gene
zscore
1.5
0.0
-1.5
f
snRNA-seq
Cluster
DN6. IC B
DN5. CNT-IC A
DN4. SPP1high CNT
DN3. DEFB1high CNT
DN2. CCD-IC A
DN1. CD-PC
DN0. DCT
SLC12A3
KLHL3
LHX1
GATA3
SCNN1G
SCNN1B
FXYD4
SLC4A1
SLC26A7
ADGRF5
LEF1
SLC8A1
CALB1
SCN2A
DEFB1
INSRR
SLC35F3
TLDC2
TMEM213
ATP6V0D2
Gene
zscore
1.5
0.0
-1.5

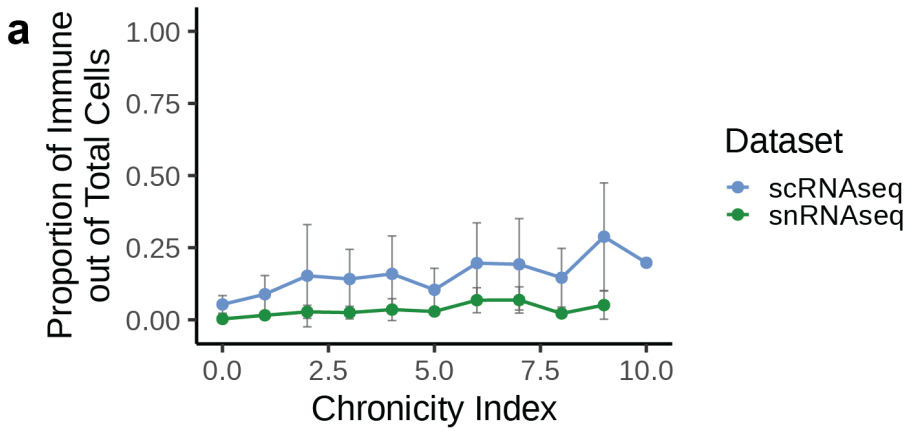
Supplementary Figure 6. Distal Nephron. a, UMAP depiction of cell states captured from scRNA-seq. **b,** UMAP depiction of cells from snRNA-seq. **c,** Scatterplot displaying each cell state cluster's proportion out of total Distal Nephron cells. X-axis reflects cell state proportions calculated in scRNA-seq and Y-axis reflects snRNA-seq. Proportions were calculated as the proportion of total cell state annotations across all donors with both matched scRNA-seq and snRNA-seq (n=22). Error bars reflect the standard error of that cluster's cell state proportion across donors, where horizontal lines reflect per-sample variability in scRNA-seq proportions, and vertical lines reflect per-sample variability in snRNA-seq proportions. **d,** Expression of marker genes (log-normalized, Z-scored) across all Distal Nephron cells. **e-f,** Heatmaps depicting average expression of marker genes (log-normalized, Z-scored after averaging) within each cell state annotation. Shown in scRNA-seq **e** and snRNA-seq **f**.


a
scRNA-Seq
UMAP2
UMAP1
b
snRNA-Seq
UMAP2
UMAP1
c
Cluster Proportions of all Endothelial/Stromal Cells
R=0.6
P=0.0666
snRNAseq
scRNAseq
0.0
0.1
0.2
0.3
0.4
d
PECAM1
COL1A1
PLVAP
NOTCH3
UMAP2
UMAP1
Expression
1.5
1.0
0.5
0.0
e
scRNA-seq
Cluster
STR9. VSMC/P
STR8. FIB
STR7. Mesangial
EC6. EC-LYMPH
EC5. EC/VSMC-P
EC4. EC-GLOM
EC3. EC-AEA
EC2. EC-DVR
EC1. EC-AVR
EC0. Fenestrated EC
PECAM1 PLVAP PTPRB FLT1 CEACAM1 TLL1 GPM6A LYPD6B ENPP2 MCTP1 ADAMTS6 BTNL9 CXCL12 EMCN HECW2 EHD3 MCAM NTRK3 RGS5 MMRN1 PROX1 TBX1 POSTN GATA3 DAAM2 NOTCH3 PDGFRB DCN COL1A2 COL1A1
Gene
zscore
1.5
0.0
-1.5
f
snRNA-seq
Cluster
STR9. VSMC/P
STR8. FIB
STR7. Mesangial
EC6. EC-LYMPH
EC5. EC/VSMC-P
EC4. EC-GLOM
EC3. EC-AEA
EC2. EC-DVR
EC1. EC-AVR
EC0. Fenestrated EC
PECAM1 PLVAP PTPRB FLT1 CEACAM1 TLL1 GPM6A LYPD6B ENPP2 MCTP1 ADAMTS6 BTNL9 CXCL12 EMCN HECW2 EHD3 MCAM NTRK3 RGS5 MMRN1 PROX1 TBX1 POSTN GATA3 DAAM2 NOTCH3 PDGFRB DCN COL1A2 COL1A1
Gene
zscore
1.5
0.0
-1.5

Supplementary Figure 7. Interstitial/Stromal. a, UMAP depiction of cell states captured from scRNA-seq. **b,** UMAP depiction of cells from snRNA-seq. **c,** Scatterplot displaying each cell state cluster's proportion out of total Interstitial/Stromal cells. X-axis reflects cell state proportions calculated in scRNA-seq and Y-axis reflects snRNA-seq. Proportions were calculated as the proportion of total cell state annotations across all donors with both matched scRNA-seq and snRNA-seq (n=22). Error bars reflect the standard error of that cluster's cell state proportion across donors, where horizontal lines reflect per-sample variability in scRNA-seq proportions, and vertical lines reflect per-sample variability in snRNA-seq proportions. **d,** Expression of marker genes (log-normalized, Z-scored) across all Interstitial/Stromal cells. **e-f,** Heatmaps depicting average expression of marker genes (log-normalized, Z-scored after averaging) within each cell state annotation. Shown in scRNA-seq **e** and snRNA-seq **f**.


a
scRNA-Seq
UMAP2
UMAP1
b
snRNA-Seq
UMAP2
UMAP1
c
Cluster Proportions of all Glomerular Cells
R=0.96
P=0.00919
snRNAseq
scRNAseq
0.0
0.1
0.2
0.3
0.4
d
PODXL
VCAM1
CFH
UMAP2
UMAP1
Expression
1
0
-1
e
scRNA-seq
Cluster
GLOM4. Proliferating
GLOM3. CXCL14high MMP7high PEC
GLOM2. Podocyte
GLOM1. SDK1high HDAC9high PEC
GLOM0. BCAMhigh AKAP12high PEC
VCAM1
CFH
ALDH1A2
BCAM
AKAP12
SDK1
HDAC9
PODXL
NPHS2
WT1
CXCL14
MMP7
CENPF
PCNA
DIAPH3
Gene
zscore
1.5
0.0
-1.5
f
snRNA-seq
Cluster
GLOM4. Proliferating
GLOM3. CXCL14high MMP7high PEC
GLOM2. Podocyte
GLOM1. SDK1high HDAC9high PEC
GLOM0. BCAMhigh AKAP12high PEC
VCAM1
CFH
ALDH1A2
BCAM
AKAP12
SDK1
HDAC9
PODXL
NPHS2
WT1
CXCL14
MMP7
CENPF
PCNA
DIAPH3
Gene
zscore
1.5
0.0
-1.5

Supplementary Figure 8. Glomerular. a, UMAP depiction of cell states captured from scRNA-seq. **b,** UMAP depiction of cells from snRNA-seq. **c,** Scatterplot displaying each cell state cluster's proportion out of total Glomerular cells. X-axis reflects cell state proportions calculated in scRNA-seq and Y-axis reflects snRNA-seq. Proportions were calculated as the proportion of total cell state annotations across all donors with both matched scRNA-seq and snRNA-seq (n=22). Error bars reflect the standard error of that cluster's cell state proportion across donors, where horizontal lines reflect per-sample variability in scRNA-seq proportions, and vertical lines reflect per-sample variability in snRNA-seq proportions. **d,** Expression of marker genes (log-normalized, Z-scored) across all Glomerular cells. **e-f,** Heatmaps depicting average expression of marker genes (log-normalized, Z-scored after averaging) within each cell state annotation. Shown in scRNA-seq **e** and snRNA-seq **f**.



Supplementary Figure 9. a, Per-sample proportions of immune cells (T/NK, B/Plasma, Myeloid) out of all cells in tissue, subset by data modality. Points represent the mean proportion across samples for a given chronicity index value. Error bars represent the standard deviation across samples at that chronicity index value.

Supplementary Tables

Supplementary Table 1. Sample demographics and cell counts for data modalities

Supplementary Table 2. Tissue cell state statistics

Supplementary Table 3. Differential gene expression results

Supplementary Table 4. Gene set enrichment results

Supplementary Table 5. Chronicity CNA p-values in tissue

Supplementary Table 6. Cell state correlations with chronicity index

Supplementary Table 7. Chronicity and activity CNA p-values in blood

References

1. Wallace, D. J. Progress toward better treatment of lupus. *N. Engl. J. Med.* **387**, 939–940 (2022).
2. Izmirly, P. M. *et al.* Prevalence of systemic lupus erythematosus in the United States: Estimates from a meta-analysis of the Centers for Disease Control and Prevention national lupus registries. *Arthritis Rheumatol.* **73**, 991–996 (2021).
3. Izmirly, P. M. *et al.* The incidence and prevalence of systemic lupus erythematosus in New York County (Manhattan), New York: The Manhattan Lupus Surveillance Program. *Arthritis Rheumatol.* **69**, 2006–2017 (2017).
4. Dall'Era, M. *et al.* The incidence and prevalence of systemic lupus erythematosus in San Francisco County, California: The California Lupus Surveillance Project. *Arthritis Rheumatol.* **69**, 1996–2005 (2017).
5. Rovin, B. H. *et al.* Efficacy and safety of voclosporin versus placebo for lupus nephritis (AURORA 1): a double-blind, randomised, multicentre, placebo-controlled, phase 3 trial. *Lancet* **397**, 2070–2080 (2021).
6. Furie, R. *et al.* Two-year, randomized, controlled trial of belimumab in lupus nephritis. *N. Engl. J. Med.* **383**, 1117–1128 (2020).
7. Gibson, K. L. *et al.* Predictors of relapse and end stage kidney disease in proliferative lupus nephritis: focus on children, adolescents, and young adults. *Clin. J. Am. Soc. Nephrol.* **4**, 1962–1967 (2009).
8. Plantinga, L. *et al.* Incidence of end-stage renal disease among newly diagnosed systemic lupus erythematosus patients: The Georgia lupus registry. *Arthritis Care Res. (Hoboken)* **68**, 357–365 (2016).
9. Mok, C. C., Kwok, R. C. L. & Yip, P. S. F. Effect of renal disease on the standardized mortality ratio and life expectancy of patients with systemic lupus erythematosus. *Arthritis*

Rheum. **65**, 2154–2160 (2013).

10. Liu, C.-C., Kao, A. H., Manzi, S. & Ahearn, J. M. Biomarkers in systemic lupus erythematosus: challenges and prospects for the future. *Ther. Adv. Musculoskelet. Dis.* **5**, 210–233 (2013).
11. Yap, D. Y. H., Tang, C. S. O., Ma, M. K. M., Lam, M. F. & Chan, T. M. Survival analysis and causes of mortality in patients with lupus nephritis. *Nephrol. Dial. Transplant* **27**, 3248–3254 (2012).
12. Anders, H.-J. *et al.* Lupus nephritis. *Nat. Rev. Dis. Primers* **6**, 7 (2020).
13. Alsuwaida, A. O. *et al.* The long-term outcomes and histological transformation in class II lupus nephritis. *Saudi Med. J.* **39**, 990–993 (2018).
14. Bajema, I. M. *et al.* Revision of the International Society of Nephrology/Renal Pathology Society classification for lupus nephritis: clarification of definitions, and modified National Institutes of Health activity and chronicity indices. *Kidney Int.* **93**, 789–796 (2018).
15. Kidney Disease: Improving Global Outcomes (KDIGO) Lupus Nephritis Work Group. KDIGO 2024 clinical practice guideline for the management of LUPUS NEPHRITIS. *Kidney Int.* **105**, S1–S69 (2024).
16. Ugarte-Gil, M. F. *et al.* The number of flares patients experience impacts on damage accrual in systemic lupus erythematosus: data from a multiethnic Latin American cohort. *Ann. Rheum. Dis.* **74**, 1019–1023 (2015).
17. Izmirly, P. M. *et al.* Longitudinal patterns and predictors of response to standard-of-care therapy in lupus nephritis: data from the Accelerating Medicines Partnership Lupus Network. *Arthritis Res. Ther.* **26**, 54 (2024).
18. Domingues, V. *et al.* Serum albumin at 1 year predicts long-term renal outcome in lupus nephritis. *Lupus Sci. Med.* **5**, e000271 (2018).
19. Dall'Era, M. *et al.* Predictors of long-term renal outcome in lupus nephritis trials: lessons learned from the Euro-Lupus Nephritis cohort. *Arthritis Rheumatol.* **67**, 1305–1313 (2015).

20. Malvar, A. *et al.* Kidney biopsy-based management of maintenance immunosuppression is safe and may ameliorate flare rate in lupus nephritis. *Kidney Int.* **97**, 156–162 (2020).
21. Hoover, P. *et al.* Accelerating Medicines Partnership: Organizational structure and preliminary data from the phase 1 studies of lupus nephritis. *Arthritis Care Res. (Hoboken)* **72**, 233–242 (2020).
22. Korsunsky, I. *et al.* Cross-tissue, single-cell stromal atlas identifies shared pathological fibroblast phenotypes in four chronic inflammatory diseases. *Med (N. Y.)* **3**, 481–518.e14 (2022).
23. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
24. Reshef, Y. A. *et al.* Co-varying neighborhood analysis identifies cell populations associated with phenotypes of interest from single-cell transcriptomics. *Nat. Biotechnol.* **40**, 355–363 (2022).
25. Oh, J.-M. *et al.* Comparison of cell type distribution between single-cell and single-nucleus RNA sequencing: enrichment of adherent cell types in single-nucleus RNA sequencing. *Exp. Mol. Med.* **54**, 2128–2134 (2022).
26. Wu, H., Kirita, Y., Donnelly, E. L. & Humphreys, B. D. Advantages of single-nucleus over single-cell RNA sequencing of adult kidney: Rare cell types and novel cell states revealed in fibrosis. *J. Am. Soc. Nephrol.* **30**, 23–32 (2019).
27. Zhang, F. *et al.* Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry. *Nat. Immunol.* **20**, 928–942 (2019).
28. Kirita, Y., Wu, H., Uchimura, K., Wilson, P. C. & Humphreys, B. D. Cell profiling of mouse acute kidney injury reveals conserved cellular responses to injury. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 15874–15883 (2020).
29. Gerhardt, L. M. S., Liu, J., Koppitch, K., Cippà, 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. U. S. A.* **118**, e2026684118 (2021).

30. Lake, B. B. *et al.* An atlas of healthy and injured cell states and niches in the human kidney. *Nature* **619**, 585–594 (2023).
31. Balzer, M. S. *et al.* Single-cell analysis highlights differences in druggable pathways underlying adaptive or fibrotic kidney regeneration. *Nat. Commun.* **13**, 4018 (2022).
32. Gray, R. & Al-Ani, Z. Risk management in clinical practice. Part 8. Temporomandibular disorders. *Br. Dent. J.* **209**, 433–449 (2010).
33. Jeong, K., Je, J., Dusabimana, T., Kim, H. & Park, S. W. Early growth response 1 contributes to renal IR injury by inducing proximal tubular cell apoptosis. *Int. J. Mol. Sci.* **24**, (2023).
34. Qin, Y. *et al.* Age-associated B cells contribute to the pathogenesis of rheumatoid arthritis by inducing activation of fibroblast-like synoviocytes via TNF-α-mediated ERK1/2 and JAK-STAT1 pathways. *Ann. Rheum. Dis.* **81**, 1504–1514 (2022).
35. Sichien, D., Lambrecht, B. N., Guilliams, M. & Scott, C. L. Development of conventional dendritic cells: from common bone marrow progenitors to multiple subsets in peripheral tissues. *Mucosal Immunol.* **10**, 831–844 (2017).
36. Siegal, F. P. *et al.* The nature of the principal type 1 interferon-producing cells in human blood. *Science* **284**, 1835–1837 (1999).
37. Jonsson, A. H. *et al.* Granzyme K+ CD8 T cells form a core population in inflamed human tissue. *Sci. Transl. Med.* **14**, eabo0686 (2022).
38. Rao, D. A. *et al.* Pathologically expanded peripheral T helper cell subset drives B cells in rheumatoid arthritis. *Nature* **542**, 110–114 (2017).
39. Gupta, A. *et al.* Dynamic regulatory elements in single-cell multimodal data implicate key immune cell states enriched for autoimmune disease heritability. *Nat. Genet.* **55**, 2200–2210 (2023).

40. Evans, R. G., Smith, D. W., Lee, C.-J., Ngo, J. P. & Gardiner, B. S. What makes the kidney susceptible to hypoxia? *Anat. Rec. (Hoboken)* **303**, 2544–2552 (2020).
41. Gao, W. *et al.* SPP1 is a prognostic related biomarker and correlated with tumor-infiltrating immune cells in ovarian cancer. *BMC Cancer* **22**, 1367 (2022).
42. Bill, R. *et al.* CXCL9:SPP1 macrophage polarity identifies a network of cellular programs that control human cancers. *Science* **381**, 515–524 (2023).
43. Mao, A.-P., Ishizuka, I. E., Kasal, D. N., Mandal, M. & Bendelac, A. A shared Runx1-bound Zbtb16 enhancer directs innate and innate-like lymphoid lineage development. *Nat. Commun.* **8**, 863 (2017).
44. Arazi, A. *et al.* The immune cell landscape in kidneys of patients with lupus nephritis. *Nat. Immunol.* **20**, 902–914 (2019).
45. Banchereau, J. & Pascual, V. Type I interferon in systemic lupus erythematosus and other autoimmune diseases. *Immunity* **25**, 383–392 (2006).
46. Singh, S. & Sreenidhi, H. C. Prognostic value of modified National Institute of Health activity and chronicity scoring in determining complete renal response in newly diagnosed lupus nephritis: a retrospective single centre study. *Int. Urol. Nephrol.* **54**, 2075–2082 (2022).
47. Yang, Y. *et al.* Nicotinamide n-methyltransferase inhibitor synergizes with sodium-glucose cotransporter 2 inhibitor to protect renal tubular epithelium in experimental models of type 2 diabetes mellitus. *J. Diabetes Complications* **39**, 108952 (2025).
48. Xu, W. & Hou, L. Knockdown of nicotinamide N-methyltransferase ameliorates renal fibrosis caused by ischemia-reperfusion injury and remodels sphingosine metabolism. *Clin. Exp. Nephrol.* **28**, 1241–1253 (2024).
49. Gewin, L. S. Renal fibrosis: Primacy of the proximal tubule. *Matrix Biol.* **68-69**, 248–262 (2018).
50. van Vollenhoven, R. F. *et al.* Treat-to-target in systemic lupus erythematosus: recommendations from an international task force. *Ann. Rheum. Dis.* **73**, 958–967 (2014).

51. Alsuwaida, A. *et al.* Strategy for second kidney biopsy in patients with lupus nephritis. *Nephrol. Dial. Transplant* **27**, 1472–1478 (2012).
52. Kaufman, L. *et al.* Up-regulation of the homophilic adhesion molecule sidekick-1 in podocytes contributes to glomerulosclerosis. *J. Biol. Chem.* **285**, 25677–25685 (2010).
53. Tan, R. J. *et al.* Tubular injury triggers podocyte dysfunction by β-catenin-driven release of MMP-7. *JCI Insight* **4**, (2019).
54. Ushikubo, M. *et al.* Milk fat globule epidermal growth factor 8 (MFG-E8) on monocytes is a novel biomarker of disease activity in systemic lupus erythematosus. *Lupus* **30**, 61–69 (2021).
55. Baek, W.-Y., Woo, J.-M., Kim, H.-A., Jung, J.-Y. & Suh, C.-H. Polymorphisms of MFGE8 are associated with susceptibility and clinical manifestations through gene expression modulation in Koreans with systemic lupus erythematosus. *Sci. Rep.* **9**, 18565 (2019).
56. Lazareth, H. *et al.* The tetraspanin CD9 controls migration and proliferation of parietal epithelial cells and glomerular disease progression. *Nat. Commun.* **10**, 3303 (2019).
57. Tsou, P.-S. & Sawalha, A. H. Glycoprotein nonmetastatic melanoma protein B: A key mediator and an emerging therapeutic target in autoimmune diseases. *FASEB J.* **34**, 8810–8823 (2020).
58. Ramachandran, P. *et al.* Resolving the fibrotic niche of human liver cirrhosis at single-cell level. *Nature* **575**, 512–518 (2019).
59. Fabre, T. *et al.* Identification of a broadly fibrogenic macrophage subset induced by type 3 inflammation. *Sci. Immunol.* **8**, eadd8945 (2023).
60. Baechler, E. C. *et al.* Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 2610–2615 (2003).
61. Migita, K. *et al.* Reduced blood BDCA-2+ (lymphoid) and CD11c+ (myeloid) dendritic cells in systemic lupus erythematosus. *Clin. Exp. Immunol.* **142**, 84–91 (2005).
62. Blomberg, S., Eloranta, M.-L., Magnusson, M., Alm, G. V. & Rönnblom, L. Expression of the

markers BDCA-2 and BDCA-4 and production of interferon-alpha by plasmacytoid dendritic cells in systemic lupus erythematosus. *Arthritis Rheum.* **48**, 2524–2532 (2003).

63. Zhang, W. *et al.* Excessive CD11c+Tbet+ B cells promote aberrant TFH differentiation and affinity-based germinal center selection in lupus. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 18550–18560 (2019).
64. Horisberger, A. *et al.* Blood immunophenotyping identifies distinct kidney histopathology and outcomes in patients with lupus nephritis. *J. Clin. Invest.* (2025) doi:10.1172/JCI181034.
65. Mouat, I. C., Goldberg, E. & Horwitz, M. S. Age-associated B cells in autoimmune diseases. *Cell. Mol. Life Sci.* **79**, 402 (2022).
66. Fava, A. *et al.* Urine proteomic signatures of histological class, activity, chronicity, and treatment response in lupus nephritis. *JCI Insight* **9**, (2024).
67. Janosevic, D., De Luca, T., Kidney Precision Medicine Project & Eadon, M. T. The Kidney Precision Medicine Project and single-cell biology of the injured proximal tubule. *Am. J. Pathol.* **195**, 7–22 (2025).
68. Hunnicutt, J. *et al.* Patients achieving low lupus disease activity state, systemic lupus erythematosus disease control or remission showed lower rates of organ damage during longitudinal follow-up: analysis of the Hopkins Lupus Cohort. *Lupus Sci. Med.* **11**, (2024).
69. Subramanian, A. *et al.* Protective role for kidney TREM2high macrophages in obesity- and diabetes-induced kidney injury. *Cell Rep.* **43**, 114253 (2024).
70. Xu, R. *et al.* Lipid-associated macrophages between aggravation and alleviation of metabolic diseases. *Trends Endocrinol. Metab.* **35**, 981–995 (2024).
71. Reggio, A., Fuoco, C., Deodati, R. & Palma, A. SPP1 macrophages across diseases: A call for reclassification? *FASEB J.* **39**, e70448 (2025).
72. Cruz Tleugabulova, M. *et al.* Induction of a distinct macrophage population and protection from lung injury and fibrosis by Notch2 blockade. *Nat. Commun.* **15**, 9575 (2024).

73. King, E. M. *et al.* Gpnmb and Spp1 mark a conserved macrophage injury response masking fibrosis-specific programming in the lung. *JCI Insight* **9**, (2024).
74. Biniaris-Georgallis, S.-I. *et al.* Amplification of autoimmune organ damage by NKp46-activated ILC1s. *Nature* **634**, 952–960 (2024).
75. De Ponti, F. F. *et al.* Spatially restricted and ontogenically distinct hepatic macrophages are required for tissue repair. *Immunity* **58**, 362–380.e10 (2025).
76. Silvin, A. *et al.* Dual ontogeny of disease-associated microglia and disease inflammatory macrophages in aging and neurodegeneration. *Immunity* **55**, 1448–1465.e6 (2022).
77. Furie, R. *et al.* A phase III, randomized, placebo-controlled study of belimumab, a monoclonal antibody that inhibits B lymphocyte stimulator, in patients with systemic lupus erythematosus. *Arthritis Rheum.* **63**, 3918–3930 (2011).
78. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).
79. Carlucci, P. M. *et al.* High incidence of proliferative and membranous nephritis in SLE patients with low proteinuria in the Accelerating Medicines Partnership. *Rheumatology (Oxford)* **61**, 4335–4343 (2022).
80. Deonaraine, K. K. *et al.* Safety of procuring research tissue during a clinically indicated kidney biopsy from patients with lupus: data from the Accelerating Medicines Partnership RA/SLE Network. *Lupus Sci. Med.* **8**, e000522 (2021).
81. Hochberg, M. C. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum.* **40**, 1725 (1997).
82. Petri, M. *et al.* Derivation and validation of the Systemic Lupus International Collaborating Clinics classification criteria for systemic lupus erythematosus. *Arthritis Rheum.* **64**, 2677–2686 (2012).
83. Weening, J. J. *et al.* The classification of glomerulonephritis in systemic lupus erythematosus revisited. *Kidney Int.* **65**, 521–530 (2004).

84. Chatelle, C. *et al.* Detection and interpretation of impossible and improbable Coma Recovery Scale-Revised scores. *Arch. Phys. Med. Rehabil.* **97**, 1295–1300.e4 (2016).
85. Doench, J. G. *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* **34**, 184–191 (2016).
86. Sanson, K. R. *et al.* Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. *Nat. Commun.* **9**, 5416 (2018).
87. Hardigan, A. A. *et al.* CRISPR/Cas9-targeted removal of unwanted sequences from small-RNA sequencing libraries. *Nucleic Acids Res.* **47**, e84 (2019).
88. Lake, B. B. *et al.* A single-nucleus RNA-sequencing pipeline to decipher the molecular anatomy and pathophysiology of human kidneys. *Nat. Commun.* **10**, 2832 (2019).
89. Zhang, F. *et al.* IFN-γ and TNF-α drive a CXCL10+ CCL2+ macrophage phenotype expanded in severe COVID-19 lungs and inflammatory diseases with tissue inflammation. *Genome Med.* **13**, 64 (2021).
90. Becht, E. *et al.* Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* **37**, 38–44 (2018).
91. Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: Computational identification of cell Doublets in Single-cell transcriptomic data. *Cell Syst.* **8**, 281–291.e9 (2019).
92. Zhang, F. *et al.* Deconstruction of rheumatoid arthritis synovium defines inflammatory subtypes. *Nature* **623**, 616–624 (2023).
93. Nathan, A. *et al.* Multimodally profiling memory T cells from a tuberculosis cohort identifies cell state associations with demographics, environment and disease. *Nat. Immunol.* **22**, 781–793 (2021).
94. Alli, A. A. *et al.* Kidney tubular epithelial cell ferroptosis links glomerular injury to tubulointerstitial pathology in lupus nephritis. *Clin. Immunol.* **248**, 109213 (2023).
95. Kang, H. M. *et al.* Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nat. Biotechnol.* **36**, 89–94 (2018).

96. Villani, A.-C. *et al.* Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356**, (2017).

97. Perez, R. K. *et al.* Single-cell RNA-seq reveals cell type-specific molecular and genetic associations to lupus. *Science* **376**, eabf1970 (2022).

98. Squair, J. W. *et al.* Confronting false discoveries in single-cell differential expression. *Nat. Commun.* **12**, 5692 (2021).

99. Tirosh, I. *et al.* Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).

100. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).

101. Cantero-Navarro, E. *et al.* Role of macrophages and related cytokines in kidney disease. *Front. Med. (Lausanne)* **8**, 688060 (2021).

102. Hadpech, S. & Thongboonkerd, V. Epithelial-mesenchymal plasticity in kidney fibrosis. *Genesis* **62**, e23529 (2024).