Title:

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

One Sentence Summary:

A single-cell atlas of paired blood and tissue samples from Lupus Nephritis patients and healthy controls identified stromal and immune populations within renal tissue, including the scar-associated macrophage populations, which correlate with and may drive renal disease activity.

Authors:

Nicholas W. Sugiarto[1,2,3,4],*Siddarth Gurajala[1,2,3,4]*, Michelle Curtis[1,2,3,4]*, Thomas M. 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], Yu Zhao[1,2,3,4], Peter M. Izmirly[9], Jennifer L. Barnas[10], Paul J. Hoover[3], Michael Peters[3], Raktima Raychowdhury[3], Alice Horisberger[2], Saori Sakaue[1,2,3,4], Richard A. Furie[11], H. Michael Belmont[9], David A. 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], Kenneth C. Kalunian[19], Jeffrey B. 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.[26]Division of Rheumatology, The University of Colorado Anschutz Medical Campus, Aurora, CO, USA. [27]Kao Autoimmunity Institute, Cedars-Sinai Medical Center, Los Angeles, California. [28]Laboratory for RNA Molecular Biology, The Rockefeller University, 1230 York Ave, Box 186, New York, NY 10065, USA.[29]Stanford University School of Medicine, Stanford, California.

Accelerating Medicines Partnerships Rheumatoid Arthritis/Systemic Lupus Erythematosus
Nicole Bornkamp[9], Michael B. Brenner[2], Robert M. Clancy[9], Evan Der[16], Eugene Drokhlyansky[3], Derek Fine[6], Beatrice Goilav[16], V. Michael Holers[26], Mariko Ishimori[27], Nicole Jordan[16], Chun-Hao Lee[6], Matthias Kretzler[8], Blue Lake[19], Rajasree Menon[8], James Pullman[16], Amit Saxena[9], Kamil Slowikowski[1,2,3,4], Manny Monroy-Trujillo[6], Michael Peters[3], Daniel Schwartz[16], Jennifer Seifert[26], Ummara Shah[10], Dawn Smilek[15], Patti Tosta[15], Thomas Tuschl[28], Paul J. Utz[29], Michael Weisman[27], Ming Wu[9], Hasret Gunduz[5], Yemil Atisha-Fregoso[5]

Abstract:

Lupus nephritis (LN), a severe manifestation of Systemic Lupus Erythematosus (SLE), is a heterogeneous disease driven by diverse immune and tissue cell types. We obtained 538K single-cell and 140K single-nuclear profiles from kidney biopsies of 155 LN patients and 30 pre-implantation transplant biopsy controls, along with 325K single-cell blood profiles overlapping many of these patients. We identified key tissue cell types and cell states, and immune cell states; we were able to determine cell states that were tissue specific, and those that were present in the blood. We observed that LN pathological features are significantly associated with cell states using differential gene expression and Covarying Neighborhood Analysis (CNA). These analyses revealed broad changes in cell states associated with irreversible chronic tissue damage. After controlling for the effects of ongoing tissue damage, we observed that expansion of key glomerular and Scar Associated Macrophages (SAMs) populations tracked with increasing inflammatory disease activity. SAMs appear to drive LN fibrosis and, in active disease, infiltrate the glomeruli more than other myeloid cells. These observations strongly support that therapeutic targeting of myeloid populations may offer an as-of-yet unproven strategy to prevent 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 support treating less advanced disease forms as well(*10*, *11*). Renal biopsies are also scored using the NIH activity and chronicity indices(*12*, *13*). The activity index captures active and potentially reversible inflammation, but high sustained scores predict poor survival(*14*). It reflects the 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(*13*). The chronicity index captures irreversible tissue damage, which increases with each LN flare(*15*) 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(*16*–*18*), but may not be concordant with the underlying histology(*19*). Thus, clinicians variably place weight on ISN class, activity, and chronicity in therapeutic decisions.

There has been great interest in using high-dimensional genomics to understand LN. The largest studies have been in bulk PBMCs and have confirmed previously identified Interferon Stimulated Gene (ISG) signatures upregulated in LN cases(*20*). Single-cell data have allowed us to localize these interferon signals to particular cell types by identifying the myeloid compartment as the largest expressor of ISGs(*21*). Recent efforts to integrate multiple public PBMC LN datasets have also pinpointed *GZMK*$^{+}$*GZMH*$^{+}$*HLA-DR*$^{+}$ effector memory CD8+ T cells as a major expressor of pro-inflammatory cytokines(*22*). However, renal tissue investigations of LN may be particularly valuable. Thus far, these

primary tissue studies have been limited in sample size(*23*) or confined to histopathology(*24*). Spatial datasets of LN tissue are now emerging(*25*, *26*). While these studies have led to valuable insights, they are currently limited by the number of samples profiled. Furthermore, these atlases are limited by the number of genes they can incorporate into their panels and therefore often require external single-cell datasets to accurately call cell types. In many instances, they lack the necessary number of transcripts in single cell profiles to resolve specific cell states and their associations to LN subphenotypes.

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(*16*, *27*). We applied single-cell (scRNAseq) and single-nuclear (snRNAseq) RNA-sequencing to profile both immune and non-immune cell populations. These methods recovered different proportions of cell types, as non-immune cells were more sensitive to the enzymatic disruption of cryopreserved tissue used to generate scRNAseq data. With these data, we defined key tissue-specific cell states and identified those linked to case/control status, chronicity, activity, and other SLE parameters. We observed that the chronicity and activity better explain effects on cell states than ISN class.

Main Text:

Results:

Cell State Definition and Clustering

We obtained kidney biopsies and peripheral blood mononuclear cell (PBMC) samples from patients with LN and healthy controls. Our cohort was 87.0% female in case samples and 76.7% female in control samples **(Table S1, Table S2C)**. 32.9% of case samples and 20.0% of control samples had a prior history of biopsy prior to enrollment. Each donor was only sampled once in our cohort. After quality control, the final dataset included tissue and clinical data from 155 LN patients and 30 healthy controls, and matched PBMC samples from 118 of the LN patients and 19 healthy controls (**Fig. 1A, Table S2A**). Following quality control to remove poor-quality cells based on low UMI counts, low gene features, and

high mitochondrial reads **(Table S3, Fig. S1A-C)**, 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. Per-biopsy cell yields were variable across donors with 95.7% of scRNAseq donors and 95.9% of snRNAseq donors contributing >500 cells **(Fig. S2B).** We applied weighted principal components analysis(*28*) followed by Harmony(*29*) to ensure both modalities contributed equally. Clustering and annotating cells together ensured that shared transcriptional signals represented biology and not modality-specific technical artifacts. (**Fig. 1B**). The PBMC data included 327,326 scRNA-seq profiles, which we integrated with Harmony. We also integrated PBMC and kidney tissue data to identify tissue-specific cell states.

We associated cell states with case/control status and histopathological metrics using Covarying Neighborhood Analysis (CNA)(*30*) (**Fig. 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 (**Fig. 1E, F**), with consistent marker genes across integrated scRNAseq and snRNAseq data **(Fig. S3-10)**. 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 scRNAseq data, likely due to destruction of tissue cells during tissue dissociation(*31*, *32*). Consistent with this, 93.9% of immune cell profiles were from scRNAseq rather than snRNAseq data **(Table S2B**).

Subclustering each broad cell type revealed 55 fine-grained renal immune cell states including 22 T/NK (**Fig. 2A**), 21 Myeloid (**Fig. 2B**), and 9 B/Plasma clusters (**Fig. 2C**). Within the renal parenchymal tissue cell types, we obtained 36 states, including 8 proximal tubule (PT, **Fig. 2D**), 5 loop of Henle (LOH, **Fig. 2E**), 8 distal nephron (DN, **Fig. 2F**), 10 endothelial/stromal (EC, **Fig. 2G**), and 5 glomerular clusters

(GLOM, **Fig. 2H**). Marker genes and cell state frequencies were concordant across scRNAseq and snRNAseq data **(Fig. S3-10)**. Read distributions across parenchymal cell types were largely consistent between each other and between modalities **(Fig. S2A)**. Immune cell read distributions were consistent with each other but captured less transcripts in snRNAseq than scRNAseq **(Fig. S2A)**

Renal Immune Cell States

Of the 22 T/NK clusters, two were NK clusters with high (NK0) or low (NK1) *NCAM1* (*CD56*) expression (**Table S4**). *CD8* T cell clusters included central memory (*SELL+*) and effector (*GZMK+/GZMB+*) subsets (**Fig. S3**), 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+SELL-ITGAE+CCR7-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* (**Table S3, Fig. S4)** 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)(*33*), 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* (**Table S3**, **Fig. S5**). 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 5 glomerular (GLOM) clusters, including podocytes (GLOM2), expressing *PODXL*, *NPHS2*, and *WT1*, and parietal epithelial cells (PECs), expressing *VCAM1* and *CFH* (**Fig. S6**, **Table S4**). 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*. GLOM1 notably expresses Epithelial-to-Mesenchymal Transition markers over other PEC subtypes (**Fig. S30A, B)**.

We identified 8 proximal tubule (PT) clusters, using *SLC34A1, SLC22A8,* and *MOGAT1* to delineate anatomical segments (S1, S2, S3) (**Fig. S7**, **Table S4**). (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(*34*). PT0 expressed *VCAM1*, *TPM1*, and *DCDC2(35, 36)* indicating maladaptive redifferentiation following injury(*37*). PT7 expressed *FOS*, *JUN*, and *EGR1,* associated with a degenerative injury response(*38*, *39*).

We identified 5 loop of Henle (LOH) clusters (**Fig. S8, Table S4**) 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 (**Fig. S8**). 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 8 fine-grain distal nephron (DN) clusters (**Fig. S9, Table S4**), 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* (**Fig. S9**). 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 (**Fig. S10**, **Table S4**). 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*. Globally, we find that transcript counts per gene between scRNAseq and snRNAseq are consistent **(Fig. S11A-H)**.

Comparing Population Frequencies in Single Cell and Spatial Transcriptomic Data

We noted the low frequency of some of the immune populations in our data, most notably B cells. Since emerging spatial transcriptomic technologies do not require disaggregation, we consider cell-type frequencies from such data to likely be accurate. We obtained a spatial dataset of childhood-onset SLE

(cSLE) consisting of 7 case biopsies and 4 controls profiled by a 980-gene CosMX panel(*25*). However, spatial transcriptomic data often lacks the high per-cell information content of single cell data: scRNAseq and snRNAseq data capture a median 5,906 transcripts per cell compared to 168 in spatial transcriptomics. **(Fig. S12A, B)**. Spatial transcriptomics however captures more cells per biopsy (median of 1,190 immune cells in spatial data compared to 189 in scRNAseq and snRNAseq). Although we have fewer immune cells collected per biopsy **(Fig. S12C)**, immune cells constitute a greater proportion of scRNAseq than of spatial methods, particularly in case samples (14.5 vs 4.9%, **Fig. S12D)**; in contrast snRNA-seq had a somewhat lower proportion of immune cells (3.6% of case samples, **Fig S12D).** Within the immune compartment, T and NK cells are relatively enriched in scRNAseq data compared to spatial data, whereas B and myeloid proportions remain similar **(Fig. S12E).** In B cells, the sparsest of our immune cell types, the proportions between spatial and scRNAseq data are consistent with the proportions in spatial data **(Fig. S12C, D, Fig. S13A)**. This concordance suggests that the low abundance of B cells (2.2% of total cells and nuclei in LN samples) reflects limited biological infiltration rather than a technical limitation of single-cell profiling approaches. Glomerular cells made up a lower proportion of total cells collected compared to the spatial dataset (12.6% vs 1.7%, **Fig. S13G**), which may reflect the possibility of lower glomerular cell yield in single cell data than other proportions. Other parenchymal cell types were concordant between our data and spatial data **(Fig. S13A, B)** except for loop of Henle cells, which were overrepresented in scRNAseq and snRNAseq relative to spatial data, while endothelial/stromal cells were underrepresented.

Defining Shared and Distinct Immune Populations in Blood and Tissue

We profiled 324,995 PBMCs from 118 LN patients and 19 healthy controls and annotated cell states with canonical marker genes (**Fig. S14**). We obtained 114,376 T cells, 52,324 NK Cells, 99,878 monocytes, 5,450 dendritic cells, and 52,967 B Cells (**Fig. S14A**). 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 (**Fig. 3**). In this joint embedding, we preserved the

original cell annotations from the respective separate tissue and blood cluster analyses. We use this embedding to illustrate the correspondence between tissue and blood clusters (**Fig. S15D-F**). Using this joint embedding, we developed a per-cell tissue specificity metric (TSM), defined as the proportion of a given cell's k=50 nearest neighbors derived from renal tissue, weighted for differences in dataset size (**Fig. 3A**). 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. TSM remains consistent across all cell types for varying choices of k (**Fig. S15A-C**). We find that per-cluster median TSM rankings remain stable following downsampling the highly abundant blood single cell data **(Fig. S16A-C)**

Plasma cell and plasmablast clusters (P6-P8) were tissue-specific, with TSM scores near 1 (**Fig. 3A, B**). The ABC−like cluster from tissue (B5) co-embedded with blood-derived ABCs (bl-B2) (**Fig. S15D**), both with a median TSM ~0.5, consistent with previous findings that ABCs spanned both blood and tissue(*33*, *40*). We wanted to understand B cell differentiation in our data. Hence, using our matched PBMC and tissue modalities, we re-embedded B cells in a joint Destiny(*41*) space and applied Slingshot(*42*) to infer trajectory pseudotimes to reflect B cell lineage. We specified blood-specific naive B cells (bl-B0. CXCR5^{high} Naive) as the starting point in our trajectories. Within this trajectory, pseudotime progresses through memory states toward B5. FCRL5+ ITGAX+ TBX21+ ABC-like and bl-B2 FCRL5^{+} ITGAX^{+} ABCs **(Fig. S17A).** Genes correlated with pseudotime include ITGAX (**Fig. S17B**), indicating that the trajectory may capture continuous differentiation of the ABC cell state.

In myeloid cells, macrophage populations (M5-12, M14, M16) were tissue-specific (**Fig. 3C**). Conversely, monocyte populations (M0-2) were shared, likely reflecting migration from the blood. Classical DCs (DC16-19) were tissue-specific, while plasmacytoid DCs (DC20) were shared with blood (**Fig. 3C**), consistent with prior observations(*43*, *44*). Non-classical monocytes (bl-M3, bl-M11) were shared between tissue and blood, while classical monocytes (bl-M0-2, bl-M5) were blood-specific. To understand myeloid cell infiltration and differentiation we applied trajectory analysis. Pseudotime inference on Destiny re-embeddings identified two trajectories. The first trajectory proceeds from blood monocytes through tissue macrophages to *GPNMB*^{high} macrophage clusters M5, M7, and M9 **(Fig. S17C)**.

This trajectory may indicate that *GPNMB*^{high} macrophage are derived from blood monocytes, with a tissue monocyte and other macrophage cell states as intermediaries. Future fate mapping experiments may be more informative as to the direct lineage of these populations. The second trajectory involves only blood and inflammatory tissue monocytes. Both monocyte-to-*GPNMB*^{high} macrophage and within-monocyte trajectories correlate with our TSM metric (R = 0.506 and R = 0.511, respectively).

T/NK cell TSM scores had fewer strictly blood- or tissue-specific populations. Tissue-skewed clusters included *CD8+GZMK+* populations (T5, T6, T12) (**Fig. 3D**), consistent with previous work(*45*). TFH/TPH populations (T21), implicated in autoimmune rheumatic disease(*46*), 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(*47*). In trajectory analysis, we observed that blood and tissue CD4+ **(Fig. S17D)** and CD8+ **(Fig. S17E)** T cells have a shared naive memory trajectory. Genes correlated with pseudotime are enriched for interferon alpha and gamma response pathways.

In B/plasma cells, differential gene expression between donor-matched tissue and PBMC samples revealed an upregulation of plasma cell markers (*IGH3*, *IGLC2*, *IGHCHG1*) in tissue (**Fig. S18A, 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 (**Fig. S18C, 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*) (**Fig. S18E, 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 *CD56*^{dim} NK (NK0) and *CD8+GZMB+SYNE2*^{bright} CTL (T3) cells.

Kidney Case Control Analysis

To understand key differences between LN and unaffected kidney, we performed case-control comparisons with CNA. When we examined all kidney cell types together, we observed expansion of immune populations, reflecting immune infiltration ($p<1x10^{-4}$ for both scRNAseq and snRNAseq) (**Fig. S19A, B**). Parenchymal compartment differences between LN patients and living donors may reflect tissue damage or differences in biopsy tissue procurement (see **Methods)**.

Case-control analysis in scRNAseq data within each cell type revealed fine-grained cell state shifts (**Fig. 4**). We detected shifts in all cell types (p<1e-04), except in B/Plasma cells, where low control cell counts were unable to generate an informative CNA result. We did not adjust for age and sex in our analysis as doing so had minimal impact (**Fig. S20**). Within T/NK cells, the greatest shifts were in tissue-specific populations (**Fig. 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) (**Fig. 4B**).

We identified shifts in the renal parenchymal populations. Activated/injury-associated PECs (GLOM1, *SDK1*^{high}*HDAC9*^{high}) were unsurprisingly expanded in LN, while homeostatic PECs (GLOM0, BCAM^{high}AKAP12^{high}) contracted (**Fig. 4C**)(*48*, *49*). Podocytes (GLOM2) were relatively expanded in LN, potentially due to depletion of healthy PECs. In Endothelial/Stromal cells, fibroblasts (STR8), EC-AVR (EC1), and EC-AEA (EC3) expanded, while VSMC/pericytes (STR9) contracted (**Fig. 4D**). In the loop of Henle, neighborhoods within TAL (LOH0) and C-TAL (LOH2) demonstrate heterogeneity, with finer subpopulations having positive and negative correlation values, indicating statistically significant expansion and contraction in subpopulations of cells. The median negative correlation indicates that, considered in aggregate, these populations are contracting. Other cell types such as DTL/ATL (LOH1) and Macula Densa (LOH4) expanded in LN samples (**Fig. 4E**). In the Distal Nephron, *SPP1*^{high} CNT (DN4) and CD-PC (DN1) expanded (**Fig. 4F**). Among Proximal Tubule cells, *PKHD1*^{high}*SLC3A1*^{high} S2/S3 (PT6) was expanded, while *SLC7A7*^{high}*ALB*^{high} S1/S2 (PT1) was contracted (**Fig. 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(*50*). To confirm that this is not driven by case-control cell number imbalances (**Table S2A**), we subsetted the number of case samples to match controls and further downsampled cells within cases to match cell-type distributions (**Fig. S21A**). Repeating CNA in this reduced data set finds that the associations remain concordant and globally significant (**Fig. S21B, C**).

Differential gene expression in LN and control samples revealed widespread upregulation of genes across cell types (**Fig. S19C**). Unsurprisingly, interferon-stimulated genes (ISGs) were enriched (*IFI44L*, *IF16*, *MX1*) across cell types, as were genes related to immune activation (*STAT1*, *SPP1*)(*51*, *52*). Several genes were broadly downregulated in LN, including innate and innate-like lymphoid cell regulators (*53*) (*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) (**Fig. S19C**). Gene set enrichment analysis (GSEA) confirmed upregulation of the well-described interferon response pathway (**Fig. S19D**)(*23*, *54*).

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). Although SLE manifestations differ between sexes (*55*), we did not find any globally significant associations with sex **(Fig. 5A)**. This may be due to the imbalance in our predominantly-female cohort, as is typical of SLE **(Table S1).** 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 (**Fig. 5A**). We speculated that many alterations due to increased chronicity reflect worsening tissue damage to the kidney.

Defining Cell State Correlations with Chronicity

The chronicity index reflects irreversible damage to the kidney and is associated with reduced treatment response(*16*, *56*). Differential expression analysis revealed gene programs associated with chronicity index across cell types (**Fig. 5B, C**). 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(*57*, *58*), was highly upregulated in multiple renal cell types. While the most significant changes occurred in parenchymal cell types, immune populations also showed upregulation of immune response genes (**Table S5**). 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 tracked with increased permanent tissue damage

(**Fig. S22C**). GSEA revealed enrichment for ECM remodeling, epithelial-mesenchymal transition, and translation pathways (**Fig. 5C**). ECM remodeling was enriched in endothelial/stromal (nominally) and proximal tubule cells, consistent with prior reports of profibrotic genes in these cell types(*59*), as well as in distal nephron and loop of Henle cells. Interferon response pathways were downregulated in T/NK, B, Myeloid, endothelial/stromal, and glomerular cells (**Fig. 5C**) 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(*23*). LN cases had higher ISG scores than controls in both tissue and immune populations (**Fig. S23A**). Among LN cases, the ISG score declined with increasing chronicity but remained stable across activity index levels (**Fig. S23B**).

We noted that chronicity was correlated with collection variables, including non-first biopsy status and biopsy site (**Fig. S2D, E**). Since chronicity increases with ongoing disease, the chronicity index was lower in the first biopsy (p=1.14e-06) (**Fig. S23F**). Patient recruitment site was correlated with many clinical variables, including first biopsy status (**Fig. S23D**); 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(*16*), chronicity was lower in complete responders than in non-responders (p=0.002) and partial responders (p=0.04) (**Fig. S23F**).

CNA analysis of broad cell types revealed associations with increasing chronicity (FDR<0.05, p<10.00e-04); these associations included expansion of T/NK, myeloid, B/plasma, and glomerular cells, with heterogeneous shifts in proximal tubule and loop of Henle cells (**Fig. S23C**). CNA within each cell type revealed fine cell state shifts **(Table S6)**. 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 (**Fig. 5D**). This expansion was further corroborated at the sample level in NK1 (1.4% at chronicity=0, with +0.46% per unit chronicity; P = 1E-4), T11 (0.32% at chronicity=0, + 0.13% per unit chronicity; P = 8E-4), and T15 (2.2% at chronicity=0, +0.46% per unit chronicity; P = 6.7E-5) **(Table S7)**. In our data set

2.44 points of chronicity correspond to 1 standard deviation. 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) (**Fig. 5E**). Conversely, tissue-specific macrophages (M5, M14) and cDCs (DC18) were expanded with increasing chronicity. These populations were also expanded at the sample-level **(Table S7)**. 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 stromal compartment (**Fig. S24A**). In glomeruli, podocytes were contracted (GLOM2) (FDR<0.05, p<9.99e-05) while injury-associated PEC subsets (GLOM1, GLOM3) were expanded (**Fig. S24B**). 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 (**Fig. 5F**). Cell states in the loop of Henle and distal nephron showed both expansions and contractions with chronicity (**Fig. S24C, D**, FDR<0.05, p<10.0e-04 both). Adjusting for treatment response had minimal impact (**Table S8, Fig. S24E**).

Although we were underpowered to detect B cell associations with CNA, we attempted to utilize clonotypes to try and uncover clones expanded in higher chronicity samples. However, clonotyping was limited by low expression of immunoglobulin transcripts **(Fig. S25A)**, a known issue of 3' sequencing's ability to determine clonotypic information(*60*). Light chain gene expression was especially sparse in B cells, and of the cells that did have non-zero light chain gene expression, no light chain genes were expanded with higher chronicity **(Fig. S25B).**

While B cells are more clinically relevant, plasma cells had more light chain gene expression **(Fig. S25A)**. However, we did not observe any clones across multiple samples that expanded with

chronicity **(Fig. S25C)**. IGKV3-11 and IGKV3D-11 gene usage was expanded in a small subset of cells, but all of these cells originate from a single high chronicity patient sample **(Fig. S25D).**

Proximal Tubule Injury Drives Chronicity-Associated Cell State Shifts

In CNA analysis, the proximal tubule was the most significantly associated population with chronicity, and had the largest number of significant (FDR<5%) local neighborhood shifts **(Table S9)**. Of the proximal tubule cell states, a key epithelial injury state, PT0 maladaptive/failed repair (R=0.542, P=1.01E-10) was most correlated with at the sample-level with increasing chronicity (**Table S7**) but only weakly associated with treatment response and not associated with ISN class (**Fig. S26A, B**). PT0 had a frequency of 14% at chronicity=0, and expanded 3.93% per unit chronicity (p<4.6E-12).

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<10e-04), distal nephron (p<10.00e-04), endothelial/stromal (p<10.00e-04), glomerular (p<10.00e-04), T/NK (p=2.6e-03), and myeloid compartments (p=2.00e-04) (**Fig. S26C**). The PT0-associated shifts were concordant with those associated with increasing chronicity (**Fig. S26D-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 (**Fig. S26C**). 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 the size of the PT0 population, is a proxy for the cellular effects of higher chronicity index and may better reflect tissue damage than chronicity index in LN itself. Within PT0, we find no differentially expressed genes that distinguish between response groups **(Fig. S27A-C).**

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(*61*), 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 (**Fig. S23D**). 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 (**Fig. 6A**). These associations were not confounded by age or sex **(Fig. S28).** Differential gene expression further revealed the largest number of genes associated with activity index in these same compartments (**Table S10**).

Although activity index correlates with ISN class and prednisone use (**Fig. S29A**) and has been associated with treatment response in previous studies(*62*), adjusting for these variables did not obviate associations (**Fig. S29B-C**), suggesting that activity independently drives these cell state shifts.
CNA on myeloid cells showed expansion of macrophages (**Fig. 6B**), 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) (**Fig. 6B**). Populations shared with blood, including *TPSB2*+ MAST cells (M4) and *CD16*+*CXC3CR1*+ Monocytes (M0), contracted with increasing activity. While M5 was also expanded with increasing chronicity, all other activity-expanded populations did not expand with increasing chronicity. For instance, *CCL2*+*CCL3*+ Monocytes (M3) and *CLEC10A*^{high} 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.

We observed that marker genes for Scar Associated Macrophages (SAMs) defined populations expanding with increasing activity. *CD9* and *GPNMB* were highly correlated ($R>0.7$) with CNA activity associations in myeloid cells, differentially expressed with activity index, and highly expressed in M5 (**Fig. S29H**). 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(*49*). *GPNMB* (R=0.713) is a known biomarker of SLE previously found in the urine and exosomes in a patient with acute LN(*63*). Both genes are known markers for SAMs(*64*, *65*) 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) (**Fig. S29D**). While a small subset of SAM macrophages expand with increasing chronicity (**Fig. S29E)**, almost all SAMs are strongly expanding with activity (**Fig. 6C**). SAM proportions of the myeloid compartment at the sample level similarly showed expansion (6.4% at 0 activity; +2.00% per unit activity; P = 5.2E-14) **(Fig. S30A),** which is consistent with the CNA results. Three out of four SAM clusters show significant sample-level expansion, specifically M5 (1.2% at 0 activity; +0.77% per unit activity; P = 1.3E-8), M7 (3.4% at 0 activity; +0.87% per unit activity; P = 1.2E-9), and M11 (1.3% at 0 activity; +0.26% per unit activity; P = 4.4E-7) **(New Table S11).** As with activity, SAM proportions of the myeloid compartment at the time of biopsy do not appear to be indicative of treatment response in our cohort **(Fig. S30B)**. 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 (**Fig. S29F**).

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. Differential gene expression between SAM and other myeloid populations reveals upregulation of pro-inflammatory markers such as the receptor Integrin Beta-1 (ITGB1) (**Fig. S30C**), which belongs to a family of integrin receptors known to bind to and activate latent TGFB(*66*). Consistent with this, integrin receptors and other upregulated genes are enriched for the EMT pathway, which is the most significantly enriched in SAMs (**Fig. S30D**). EMT has been shown in previous work as a necessary prerequisite in murine organ fibrosis(*67*).

To better understand where SAMs localize, we examined a spatial data set of 11 samples (see section **Comparing Population Frequencies in Single Cell and Spatial Transcriptomic Data**)(*25*). To identify SAMs in this dataset, we trained a logistic regression classifier to distinguish SAM versus non-SAM myeloid cells using 258 genes that are present in both our data set and the spatial transcriptomic data set, and are highly expressed in the Danaher et al. dataset. **(Methods).** Testing the classifier on an unseen subset of across 10 folds myeloid cells in our data demonstrated an average AUC of 0.933 **(Fig. S31A).** Applying this classifier to spatial data, we find that macrophages and SAMs localize to glomerular niches **(Fig. 6F, Fig. S31B, C)**. Across all spatial samples, SAMs comprise a higher percentage of total cells in SLE cases over control samples **(Fig. S31D)**. Within case samples, SAMs are present in the glomeruli at a higher rate than non-SAM myeloid cells **(Fig. S31E).**

Conversely, SAMs obtained by scRNAseq and snRNAseq resemble macrophages previously identified in spatial data to be localized to the glomeruli. From spatial transcriptomic data, we defined a Glomerular Infiltration Score (GIS) consisting of the 196 differentially expressed genes previously reported upregulated or downregulated in the glomeruli-localized macrophages over interstitial macrophages (FDR < 0.05)(*25*)**.** Applying this score on our dataset we find that SAMs have a higher infiltration score over other non-SAMs **(Fig. 6G)**. Samples with higher median infiltration score have higher activity index **(Fig. S31F)**. This suggests that our SAMs are not only associated with glomerular damage as measured by the activity index but are also infiltrating the glomeruli at a higher proportion than other non-activity associated macrophage populations.

Among glomerular cells, cell states associated with increasing activity index were similar to those associated with chronicity (**Fig. S29G,** R=0.487), including strong expansion of *SDK1*^{high}*HDAC9*^{high} PEC (GLOM1) (**Fig. 6B, C**). This expansion was also observed at the sample level (15.0% at 0 activity; +1.90% per unit activity; P = 6.6E-7) **(New Table S11).** 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*, *64*, *68*, *69*). Notably, GLOM1 selectively expresses *TNC* (Tenascin C), an extracellular matrix glycoprotein that is transiently induced during tissue injury and repair(*70*). We have previously shown that urinary Tenascin C levels correlate with histological activity—and, to a lesser extent, chronicity—in patients with LN(*71*), 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*) (**Fig. S29H**). *MFGE8*, previously implicated in SLE disease activity and risk(*72*, *73*), was also enriched. All are markers of the activity-associated cluster GLOM1. Differential expression analysis within PECs **(Fig. S32A)** suggests that GLOM1 is enriched for the epithelial to mesenchymal transition pathway **(Fig. S32B)**. To test whether activity-associated expansions are due to proliferation, we developed a proliferation score from 93 cell cycle genes and found that GLOM1 had a higher proliferation score than other PEC populations (**Fig. S32C**).

Given the localization of SAMs to the glomeruli, we wondered if they interacted with glomerular cells. We employed receptor-ligand analysis between glomerular cells and tissue immune cells. We observed that the glomerular-myeloid pairing has the most interactions among the cell types **(Fig. S32F).** Statistically significant glomerular and myeloid receptor-ligand pairings implicate SPP1 and TNC. SPP1, a known SAM marker interacts with cell adhesion molecule CD44 and 4 integrin receptor heterodimers **(Fig. S32E)**. Two of these heterodimers show strong interaction with SPP1 in GLOM1 **(Fig. S32F, G)**. ITGB1, which forms a heterodimer with ITGAV, is upregulated in GLOM1 compared with other PECs **(Fig. S32A)** and is also part of the EMT pathway. Another integrin heterodimer, ITGB3 + ITGAV, is strongly enriched for SPP1 and GLOM1 interactions; ITGB3 is not significantly expressed in GLOM1 over other PECs ($p = 0.99$). Although not previously characterized in SLE, secreted SPP1 promotes cell migration as part of the EMT pathway (*74*, *75*)**.**

GLOM1 upregulated Tenascin C (TNC) which has two main binding partners: SDC4 and the ITGAV+ITGB3 heterodimer. TNC expressed by GLOM1 is found to interact with the SDC4 receptor

found on SAMs and proliferating myeloid cells **(Fig. S32I).** Furthermore, GLOM1-expressed TNC interacts with the aforementioned pan-glomerular ITGAV+ITGB3 receptor **(Fig. S32J)**. As previously mentioned, our group has found TNC as a potential urine biomarker correlated with activity in SLE. While the integrin-TNC link has not been explicitly studied in SLE, it is well-known to trigger the EMT pathway in cancer(*76*, *77*); the EMT pathway has been shown to contribute to renal fibrosis(*78*). Given that the chronicity index is divided into indices representing glomerular or tubulointerstitial damage and 21 out of 24 points in the activity index are based on measures of glomerular inflammation, we examined whether applying CNA to the individual subcomponents could reflect different aspects of LN pathology.

Activity and Chronicity Index Subcomponents

Activity and chronicity indices are comprised of subindices representing glomerular or tubulointerstitial damage. For example, 21 out of 24 points in the activity index are based on measures of glomerular inflammation. We examined whether applying CNA to the individual subcomponents could reflect different aspects of LN pathology.

In chronicity, however, apart from fibrous crescents, most indices correlate with each other, including measures of glomerular or tubulointerstitial damage **(Fig. S33A).** CNA on fibrous crescents reveals new neighborhoods in distal nephron and myeloid cells **(Fig. S33B, C)** that were not previously identified as associated with chronicity **(Table S11).** While DN4 primarily expands with higher chronicity, higher fibrous crescents are marked exclusively by expanded distal convolute tubule (DN0) populations **(Fig. 33D, E).** In the myeloid compartment, fibrous crescent CNA has a lower correlation with chronicity CNA (R = 0.442) than it does with activity CNA (R = 0.715) **(Fig. S33F)**. While M5 remains the most significant with fibrous crescent as it does with chronicity, dendritic cells are no longer significantly expanded **(Fig. S33G)**. Other macrophage populations M11 and M7 remain significantly expanded. Previous pathological studies suggest that fibrous crescents may indicate early lesions that

eventually lead to glomerulosclerosis (*79*). Furthermore, our group has previously reported that the urine proteomic signature in high fibrous crescent index is more correlated with activity rather than chronicity (*71*). So we felt that quantification of fibrous crescent may capture a mixture of chronicity and activity.

In activity, three components correlated with overall activity at the sample level index with R < 0.6, fibrinoid necrosis, wire loops, and interstitial inflammation **(Fig. S34A)**. Notably, interstitial inflammation is the lone component that does not reflect glomerular damage. However, these associations were only significant in myeloid cells **(Table S12)**. Within myeloid cells, neighborhood-level associations for these indices were highly correlated with activity **(Fig. S34B)** and did not identify many novel neighborhoods not previously identified as passing FDR significance in activity CNA **(Fig. S34C-E).**

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 (ISN class V) vs proliferative nephritis, since membranous nephritis is clinically defined as having low activity.

Class V was significantly associated with Myeloid (p<9.99e-05) and glomerular (p=0.0029) populations (**Fig. S36A, 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 included $SPP1^{low}FABP5^{high}$ (M8), $GPNMB^{high}LYVE1^{low}$ (M11), $SELENOP^{inter}LYVE1^{inter}$ Resident (M10), $SPP1^{high}FABP5^{high}$ (M7), and $MERTK^{high}FABP5^{high}$ (M9) macrophages (**Fig. S35A**). In the Glomeruli, $SDK1^{high}HDAC9^{high}$ PEC (GLOM1) were also contracted (**Fig. S35B**). Results were unchanged after adjustment for prednisone use and Black race, both of which correlated with ISN Class V (**Fig. S23D**, **Fig. S35C**).

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 (**Fig. 35D, 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) (**Fig. S36A-C**). LN cases had expansion of ISG-high monocytes (bl-M6), 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 (*80*). Dendritic cells (bl-DC10, bl-DC11, bl-DC12) were contracted, consistent with previous studies of SLE blood(*81*, *82*). ABCs (bl-B2) were expanded, aligning with prior human patient and murine lupus models(*40*, *83*). 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} Central Memory (bl-T5), CD4+ 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 (**Fig. S36D**).

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) (**Table S13, 14**). 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 (**Figure S36E**). Notably, after correcting for age, sex, there were no significant associations with activity. The limited activity index associations in blood does not rule out the possibility of immunophenotype associations that may be present with different patient subgroups; for example, in a recent study on the same cohort, statistical differences were seen in B cell states across patient subgroups (*84*).

Similarly, differential expression analyses revealed only weak gene associations with chronicity index (**Fig. S36F,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 the blood of SLE patients(*85*), 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 in LN, 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(*23*), 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 only a modest decline with increasing chronicity in some cell types, and no association with activity.
We observed subtle shifts, such as expansion of *GZMK+CD8+* T cells in blood with increasing chronicity, that paralleled changes in tissue. However, overall, blood expression and cell states had limited associations 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(*48*). In contrast, non-invasive biomarkers, including macrophage-derived proteins in urine, have been associated with proliferative disease and decrease in treatment response(*71*).

We were particularly interested in the activity index, which reflects ongoing renal inflammation and is thought to be modifiable with immunomodulatory therapies. Even brief reductions in activity index has been associated with reduced organ damage(*86*). 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)(*63*) 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(*87–89*). We leverage a classifier built on SAM transcriptional profiles to find that SAM macrophages identified in a spatial dataset infiltrate the glomeruli at a higher proportion than non-SAMs. In addition, receptor-ligand analysis indicates that SAMs interact specifically with activity-associated glomerular populations through the EMT pathway.

SAM subsets are thought to be heterogeneous, with some populations being pro-fibrotic and others being reparative. (*90–92*). The persistence of the M5 subset in association with both disease activity and chronicity may suggest a particularly pathogenic role. These populations appear to be similar

to SAMs reported in other diseases and tissue contexts(*93*). We note that SAMs may not be specific to LN and may play a role in other renal diseases. Further comparative studies between other renal diseases may address this in the future.

This data does not definitively establish whether these states drive disease activity, mediate tissue repair, or promote fibrosis but rather highlights distinct, disease-associated states with therapeutic potential. Furthermore, while our trajectory analysis suggest that SAMs may infiltrate from blood into 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 (*94*, *95*). Mouse models may be useful in defining causal functions of this and other myeloid populations (*96*). Current SLE therapies primarily target B and T cells (*5*, *97*). While B cell targeting has been effective, our study was underpowered to detect activity associations with B cell states. We recognize that single cell analysis requires tissue disaggregation and droplet-based assays, which limits the amount of cells profiled. In addition, there may be biases in terms of cell type proportions. We find that, compared to spatial datasets, our single cell data has a lower proportion of glomerular and endothelial/stromal cells. We find that immune populations are not significantly different between spatial and single-cell modalities, and our lack of B cells may reflect limited infiltration rather than technical collection limitations. The low frequency of B cells in LN is consistent with the idea that B cells are acting elsewhere in other tissues to produce damaging immune complexes. The data presented here supports the potential for targeting myeloid cells in LN in conjunction with B cells.

While we are, to our knowledge, the largest single-cell LN tissue dataset, spatial methods collect many more cells per biopsy and may result in even larger atlases. Single cell data, however, still offers much greater depth of information for each individual cell with over >30-fold transcripts per cell enabling fine cell state definitions. The next phase of this project, AMP AIM, is currently working towards building a large data set of spatial transcriptomics for LN to complement the information in this dataset(*98*).

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 (*99*). 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(*16*). 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. Intriguingly, B cells were still relatively rare in the tissue for our study, despite the large number of cells and samples assayed. This was surprising considering the importance of B cells as an SLE drug target. Given the similar proportions of B cells in our atlas and an external spatial dataset, it is likely that this is not a technology issue. B cells are pathogenically active in other tissues, such as lymph nodes, and are less active in LN kidneys. It is possible that there are important rare cell states that track with disease activity in the kidney, but our study would lack the power to identify this given the few cells assayed.

However, other pathogenically important cell types, such as glomerular cells, are still relatively depleted over spatial data where tissue architecture is not as disturbed. This is a known limitation of single-cell data, but glomerular cells have been rescued in other studies through a Percoll gradient(*100*, *101*). Additionally, renal immune cell states were defined transcriptionally, without surface protein markers, which can improve immune cell state annotation accuracy(*102*). 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.

Methods and Materials

Study Design

We have described the details of enrollment in the AMP SLE cohort elsewhere(*16*, *103*, *104*). 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(*105*) or the Systemic Lupus Erythematosus International Cooperating Clinics(*106*) 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(*13*, *107*). 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. One sample was later discovered to have chronic lymphocytic leukemia and was excluded from downstream analyses.

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.

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(*108*). 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(*109*). 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(*35*, *110*).

Defining PBMC Fine Grain Cell States

We processed data within each PBMC each obtained cell type. 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 scRNAseq

and snRNAseq data, and then scaled the data. We then applied Harmony[24] to principal components weighted by scRNAseq to snRNAseq ratio to correct for processing batch, site, sample, and differences in scRNAseq and snRNAseq 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(*111*), and Perez et. al(*21*) **(Fig. S14B).**. These markers were matched with clusters via identification of cluster specific gene expression sourced from the presto package(*108*).

Clustering within T and NK Cells obtained 20 clusters spanning CD56^{bright} and CD56^{dim} NK Cells (bl-NK0, bl-NK1, bl-NK2), GZMK+CD8+ (bl-T3, bl-T4, bl-T14, bl-T15), and gamma/delta T cells (bl-T11, bl-T12). TCAT multilabel annotation(*112*) leverages RNA expression data to distinguish between CD4+ central memory (bl-T5) and naive (bl-T10) populations **(Fig. S37A)**. Key marker genes designated between effector memory populations (bl-T9, bl-T13), *GZMK*+*TEMRA* (bl-T7), *CD8*+*GZMB*+proliferating CTLs (bl-T18, bl-T19), Treg (bl-T16), and an interferon-stimulated gene (ISG) high cell state (bl-T17) **(Fig. S37C, D).** CITE-Seq obtained for the PBMC cells further validates these annotations, marking bl-T10 CD4+ naive populations as CD45RA+, CD62L+, and CD45RO- and bl-T5. CD4+ central memory populations as predominantly CD45RO+, CD62L+, and CD45RA- **(Fig. S37B, C)**.

13 monocyte and dendritic cell clusters were delineated as *CD14*+*CD16*- conventional monocytes (bl-M0, bl-M1, bl-M2, bl-M5), *CD14*+ *CD16*+ intermediate monocytes (bl-M4, bl-M8), and *CD14*^{dim} *CD16*++ non-conventional monocytes (bl-M3) (**Fig. S14E, F**). We also found a cluster of ISG^{high} monocytes (bl-M6), CD1C+ DC2 dendritic cells (bl-DC10), *CLEC9A*+ *XCR1*+ DC1 dendritic cells (bl-DC12), *TCF4*+ *CLEC4C*+ pDC dendritic cells (bl-DC11), and platelet cells (bl-P9). B Cells were clustered into 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). (**Fig. S14G, H**). 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).

Case-Control Statistical Associations

To understand cell states enriched and depleted in cases, we utilized Covarying Neighborhood Analysis (CNA) (*30*). 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(*113*). 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.

Statistical Analysis

Correlations between clinical variables were calculated using Pearson correlation for all 155 cases. For correlations with collection sites, we performed multivariate regression correlating clinical variables with all sites. When correlating gene expression with CNA results, only genes that were marked differential using a Wilcoxon rank-sum test via presto with logFC > 0.5 and Benjamin-Hochberg adjusted $p<0.01$ were used. Correlations were then calculated using Pearson correlation with Bonferroni-corrected p-values. Statistical methods for other analyses are described in-depth in their appropriate sections.

References and Notes:

1. D. J. Wallace, Progress toward better treatment of lupus. *N. Engl. J. Med.* **387**, 939–940 (2022).

2. P. M. Izmirly, H. Parton, L. Wang, W. J. McCune, S. S. Lim, C. Drenkard, E. D. Ferucci, M. Dall'Era, C. Gordon, C. G. Helmick, E. C. Somers, 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. P. M. Izmirly, I. Wan, S. Sahl, J. P. Buyon, H. M. Belmont, J. E. Salmon, A. Askanase, J. M. Bathon, L. Geraldino-Pardilla, Y. Ali, E. M. Ginzler, C. Putterman, C. Gordon, C. G. Helmick, H. Parton, 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. M. Dall'Era, M. G. Cisternas, K. Snipes, L. J. Herrinton, C. Gordon, C. G. Helmick, 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. B. H. Rovin, Y. K. O. Teng, E. M. Ginzler, C. Arriens, D. J. Caster, J. Romero-Diaz, K. Gibson, J. Kaplan, L. Lisk, S. Navarra, S. V. Parikh, S. Randhawa, N. Solomons, R. B. Huizinga, 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. R. Furie, B. H. Rovin, F. Houssiau, A. Malvar, Y. K. O. Teng, G. Contreras, Z. Amoura, X. Yu, C.-C. Mok, M. B. Santiago, A. Saxena, Y. Green, B. Ji, C. Kleoudis, S. W. Burriss, C. Barnett, D. A. Roth, Two-year, randomized, controlled trial of belimumab in lupus nephritis. *N. Engl. J. Med.* **383**, 1117–1128 (2020).

7. K. L. Gibson, D. S. Gipson, S. A. Massengill, M. A. Dooley, W. A. Primack, M. A. Ferris, S. L. Hogan, 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. L. Plantinga, S. S. Lim, R. Patzer, W. McClellan, M. Kramer, M. Klein, S. Pastan, C. Gordon, C. Helmick, C. Drenkard, 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. C. C. Mok, R. C. L. Kwok, P. S. F. Yip, 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. D. Y. H. Yap, C. S. O. Tang, M. K. M. Ma, M. F. Lam, T. M. Chan, Survival analysis and causes of mortality in patients with lupus nephritis. *Nephrol. Dial. Transplant* **27**, 3248–3254 (2012).

11. H.-J. Anders, R. Saxena, M.-H. Zhao, I. Parodis, J. E. Salmon, C. Mohan, Lupus nephritis. *Nat. Rev. Dis. Primers* **6**, 7 (2020).

12. A. O. Alsuwaida, A. A. Bakhit, F. A. Alsuwaida, J. J. Wadera, H. M. Kfoury, S. Husain, The long-term outcomes and histological transformation in class II lupus nephritis. *Saudi Med. J.* **39**, 990–993 (2018).

13. I. M. Bajema, S. Wilhelmus, C. E. Alpers, J. A. Bruijn, R. B. Colvin, H. T. Cook, V. D. D'Agati, F. Ferrario, M. Haas, J. C. Jennette, K. Joh, C. C. Nast, L.-H. Noël, E. C. Rijnink, I. S. D. Roberts, S. V. Seshan, S. Sethi, A. B. Fogo, 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).

14. 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).

15. M. F. Ugarte-Gil, E. Acevedo-Vásquez, G. S. Alarcón, C. A. Pastor-Asurza, J. L. Alfaro-Lozano, J. M. Cucho-Venegas, M. I. Segami, D. Wojdyla, E. R. Soriano, C. Drenkard, J. C. Brenol, A. C. de Oliveira e Silva Montandon, L. T. L. Costallat, L. Massardo, J. F. Molina-Restrepo, M. Guibert-Toledano, L. H. Silveira, M. C. Amigo, L. A. Barile-Fabris, R. Chacón-Díaz, M. H. Esteva-Spinetti, G. J. Pons-Estel, G. McGwin Jr, B. A. Pons-Estel, GLADEL, 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).

16. P. M. Izmirly, M. Y. Kim, P. M. Carlucci, K. Preisinger, B. Z. Cohen, K. Deonaraine, D. Zaminski, M. Dall'Era, K. Kalunian, A. Fava, H. M. Belmont, M. Wu, C. Putterman, J. Anolik, J. L. Barnas, B. Diamond, A. Davidson, D. Wofsy, D. Kamen, J. A. James, J. M. Guthridge, W. Apruzzese, D. A. Rao, M. H. Weisman, Accelerating Medicines Partnership in RA/SLE Network, M. Petri, J. Buyon, R. Furie, 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).

17. V. Domingues, B. A. Levinson, N. Bornkamp, J. D. Goldberg, J. Buyon, H. M. Belmont, Serum albumin at 1 year predicts long-term renal outcome in lupus nephritis. *Lupus Sci. Med.* **5**, e000271 (2018).

18. M. Dall'Era, M. G. Cisternas, D. E. Smilek, L. Straub, F. A. Houssiau, R. Cervera, B. H. Rovin, M. Mackay, Predictors of long-term renal outcome in lupus nephritis trials: lessons learned from the Euro-Lupus Nephritis cohort. *Arthritis Rheumatol.* **67**, 1305–1313 (2015).

19. A. Malvar, V. Alberton, B. Lococo, M. Ferrari, P. Delgado, H. N. Nagaraja, B. H. Rovin, Kidney biopsy-based management of maintenance immunosuppression is safe and may ameliorate flare rate in lupus nephritis. *Kidney Int.* **97**, 156–162 (2020).

20. R. Banchereau, S. Hong, B. Cantarel, N. Baldwin, J. Baisch, M. Edens, A.-M. Cepika, P. Acs, J. Turner, E. Anguiano, P. Vinod, S. Kahn, G. Obermoser, D. Blankenship, E. Wakeland, L. Nassi, A. Gotte, M. Punaro, Y.-J. Liu, J. Banchereau, J. Rossello-Urgell, T. Wright, V. Pascual, Personalized immunomonitoring uncovers molecular networks that stratify lupus patients. *Cell* **165**, 551–565 (2016).

21. R. K. Perez, M. G. Gordon, M. Subramaniam, M. C. Kim, G. C. Hartoularos, S. Targ, Y. Sun, A. Ogorodnikov, R. Bueno, A. Lu, M. Thompson, N. Rappoport, A. Dahl, C. M. Lanata, M. Matloubian, L. Maliskova, S. S. Kwek, T. Li, M. Slyper, J. Waldman, D. Dionne, O. Rozenblatt-Rosen, L. Fong, M. Dall'Era, B. Balliu, A. Regev, J. Yazdany, L. A. Criswell, N. Zaitlen, C. J. Ye, Single-cell RNA-seq reveals cell type-specific molecular and genetic associations to lupus. *Science* **376**, eabf1970 (2022).

22. M. Nakano, M. Kono, K. Asahara, T. Katsuyama, S. Kubo, E. Katsuyama, Y. Fujita, H. Inokuchi, T.

Nishino, T. Arakawa, T. Kawashima, S. Noma, R. Bagherzadeh, Y. Matsumoto, J. Inamo, H. Takahashi, B. Natsumoto, X. Zhang, S.-C. Bae, A. Suzuki, H. Hatano, C. Terao, Y. Tanaka, Y. Matsumoto, K. Yamamoto, K. Ishigaki, Cellular and molecular fine mapping pinpoints new immunopathology of lupus, *bioRxiv* (2025)p. 2025.12.08.693110.

23. A. Arazi, D. A. Rao, C. C. Berthier, A. Davidson, Y. Liu, P. J. Hoover, A. Chicoine, T. M. Eisenhaure, A. H. Jonsson, S. Li, D. J. Lieb, F. Zhang, K. Slowikowski, E. P. Browne, A. Noma, D. Sutherby, S. Steelman, D. E. Smilek, P. Tosta, W. Apruzzese, E. Massarotti, M. Dall'Era, M. Park, D. L. Kamen, R. A. Furie, F. Payan-Schober, W. F. Pendergraft 3rd, E. A. McInnis, J. P. Buyon, M. A. Petri, C. Putterman, K. C. Kalunian, E. S. Woodle, J. A. Lederer, D. A. Hildeman, C. Nusbaum, S. Raychaudhuri, M. Kretzler, J. H. Anolik, M. B. Brenner, D. Wofsy, N. Hacohen, B. Diamond, Accelerating Medicines Partnership in SLE network, The immune cell landscape in kidneys of patients with lupus nephritis. *Nat. Immunol.* **20**, 902–914 (2019).

24. A. Malvar, V. Alberton, B. Lococo, M. Lourenco, J. Martinez, L. Burna, C. Besso, J. Navarro, H. N. Nagaraja, A. Khatiwada, B. Wolf, B. Rovin, Remission of lupus nephritis: the trajectory of histological response in successfully treated patients. *Lupus Sci. Med.* **10**, e000932 (2023).

25. P. Danaher, N. Hasle, E. D. Nguyen, J. E. Roberts, N. Rosenwasser, C. Rickert, E. W. Y. Hsieh, K. Hayward, D. M. Okamura, C. E. Alpers, R. C. Reed, S. K. Baxter, S. W. Jackson, Childhood-onset lupus nephritis is characterized by complex interactions between kidney stroma and infiltrating immune cells. *Sci. Transl. Med.* **16**, eadl1666 (2024).

26. Z. Sultana, R. Khatri, B. Yousefi, N. Shaikh, S. L. Jauch-Speer, D. P. Schaub, J. Engesser, M. Hellmig, V. Piegsa, A. Hube, V. Sivayoganathan, A. Borchers, A. Peters, A. Kaffke, S. Zielinski, H.-J. Paust, T. Goldbeck-Strieder, U. O. Wenzel, V. G. Puelles, E. Hoxha, T. Wiech, C. Meyer-Schwesinger, T. B. Huber, U. Panzer, S. Bonn, C. F. Krebs, Spatiotemporal interaction of immune and renal cells controls glomerular crescent formation in autoimmune kidney disease. *Nat. Immunol.* **26**, 1977–1988 (2025).

27. P. Hoover, E. Der, C. C. Berthier, A. Arazi, J. A. Lederer, J. A. James, J. Buyon, M. Petri, H. M. Belmont, P. Izmirly, D. Wofsy, N. Hacohen, B. Diamond, C. Putterman, A. Davidson, Accelerating Medicines Partnership: Organizational structure and preliminary data from the phase 1 studies of lupus nephritis. *Arthritis Care Res. (Hoboken)* **72**, 233–242 (2020).

28. I. Korsunsky, K. Wei, M. Pohin, E. Y. Kim, F. Barone, T. Major, E. Taylor, R. Ravindran, S. Kemble, G. F. M. Watts, A. H. Jonsson, Y. Jeong, H. Athar, D. Windell, J. B. Kang, M. Friedrich, J. Turner, S. Nayar, B. A. Fisher, K. Raza, J. L. Marshall, A. P. Croft, T. Tamura, L. M. Sholl, M. Vivero, I. O. Rosas, S. J. Bowman, M. Coles, A. P. Frei, K. Lassen, A. Filer, F. Powrie, C. D. Buckley, M. B. Brenner, S. Raychaudhuri, Cross-tissue, single-cell stromal atlas identifies shared pathological fibroblast phenotypes in four chronic inflammatory diseases. *Med (N. Y.)* **3**, 481–518.e14 (2022).

29. I. Korsunsky, N. Millard, J. Fan, K. Slowikowski, F. Zhang, K. Wei, Y. Baglaenko, M. Brenner, P.-R. Loh, S. Raychaudhuri, Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).

30. Y. A. Reshef, L. Rumker, J. B. Kang, A. Nathan, I. Korsunsky, S. Asgari, M. B. Murray, D. B. Moody, S. Raychaudhuri, Co-varying neighborhood analysis identifies cell populations associated with phenotypes of interest from single-cell transcriptomics. *Nat. Biotechnol.* **40**, 355–363 (2022).

31. J.-M. Oh, M. An, D.-S. Son, J. Choi, Y. B. Cho, C. E. Yoo, W.-Y. Park, 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).

32. H. Wu, Y. Kirita, E. L. Donnelly, B. D. Humphreys, 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).

33. F. Zhang, K. Wei, K. Slowikowski, C. Y. Fonseka, D. A. Rao, S. Kelly, S. M. Goodman, D. Tabechian, L. B. Hughes, K. Salomon-Escoto, G. F. M. Watts, A. H. Jonsson, J. Rangel-Moreno, N. Meednu, C. Rozo, W. Apruzzese, T. M. Eisenhaure, D. J. Lieb, D. L. Boyle, A. M. Mandelin 2nd, Accelerating Medicines Partnership Rheumatoid Arthritis and Systemic Lupus Erythematosus (AMP RA/SLE) Consortium, B. F. Boyce, E. DiCarlo, E. M. Gravallese, P. K. Gregersen, L. Moreland, G. S. Firestein, N. Hacohen, C. Nusbaum, J. A. Lederer, H. Perlman, C. Pitzalis, A. Filer, V. M. Holers, V. P. Bykerk, L. T. Donlin, J. H. Anolik, M. B. Brenner, S. Raychaudhuri, Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry. *Nat. Immunol.* **20**, 928–942 (2019).

34. Y. Kirita, H. Wu, K. Uchimura, P. C. Wilson, B. D. Humphreys, Cell profiling of mouse acute kidney injury reveals conserved cellular responses to injury. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 15874–15883 (2020).

35. L. M. S. Gerhardt, J. Liu, K. Koppitch, P. E. Cippà, A. P. McMahon, 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).

36. B. B. Lake, R. Menon, S. Winfree, Q. Hu, R. Melo Ferreira, K. Kalhor, D. Barwinska, E. A. Otto, M. Ferkowicz, D. Diep, N. Plongthongkum, A. Knoten, S. Urata, L. H. Mariani, A. S. Naik, S. Eddy, B. Zhang, Y. Wu, D. Salamon, J. C. Williams, X. Wang, K. S. Balderrama, P. J. Hoover, E. Murray, J. L. Marshall, T. Noel, A. Vijayan, A. Hartman, F. Chen, S. S. Waikar, S. E. Rosas, F. P. Wilson, P. M. Palevsky, K. Kiryluk, J. R. Sedor, R. D. Toto, C. R. Parikh, E. H. Kim, R. Satija, A. Greka, E. Z. Macosko, P. V. Kharchenko, J. P. Gaut, J. B. Hodgin, KPMP Consortium, M. T. Eadon, P. C. Dagher, T. M. El-Achkar, K. Zhang, M. Kretzler, S. Jain, An atlas of healthy and injured cell states and niches in the human kidney. *Nature* **619**, 585–594 (2023).

37. M. S. Balzer, T. Doke, Y.-W. Yang, D. L. Aldridge, H. Hu, H. Mai, D. Mukhi, Z. Ma, R. Shrestha, M. B. Palmer, C. A. Hunter, K. Susztak, Single-cell analysis highlights differences in druggable pathways underlying adaptive or fibrotic kidney regeneration. *Nat. Commun.* **13**, 4018 (2022).

38. R. Gray, Z. Al-Ani, Risk management in clinical practice. Part 8. Temporomandibular disorders. *Br. Dent. J.* **209**, 433–449 (2010).

39. K. Jeong, J. Je, T. Dusabimana, H. Kim, S. W. Park, Early growth response 1 contributes to renal IR injury by inducing proximal tubular cell apoptosis. *Int. J. Mol. Sci.* **24** (2023).

40. Y. Qin, M.-L. Cai, H.-Z. Jin, W. Huang, C. Zhu, A. Bozec, J. Huang, Z. Chen, 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).

41. P. Angerer, L. Haghverdi, M. Büttner, F. J. Theis, C. Marr, F. Buettner, destiny: diffusion maps for large-scale single-cell data in R. *Bioinformatics* **32**, 1241–1243 (2016).

42. K. Street, D. Risso, R. B. Fletcher, D. Das, J. Ngai, N. Yosef, E. Purdom, S. Dudoit, Slingshot: cell

lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477 (2018).

43. D. Sichien, B. N. Lambrecht, M. Guilliams, C. L. Scott, Development of conventional dendritic cells: from common bone marrow progenitors to multiple subsets in peripheral tissues. *Mucosal Immunol.* **10**, 831–844 (2017).

44. F. P. Siegal, N. Kadowaki, M. Shodell, P. A. Fitzgerald-Bocarsly, K. Shah, S. Ho, S. Antonenko, Y. J. Liu, The nature of the principal type 1 interferon-producing cells in human blood. *Science* **284**, 1835–1837 (1999).

45. A. H. Jonsson, F. Zhang, G. Dunlap, E. Gomez-Rivas, G. F. M. Watts, H. J. Faust, K. V. Rupani, J. R. Mears, N. Meednu, R. Wang, G. Keras, J. S. Coblyn, E. M. Massarotti, D. J. Todd, J. H. Anolik, A. McDavid, Accelerating Medicines Partnership RA/SLE Network, K. Wei, D. A. Rao, S. Raychaudhuri, M. B. Brenner, Granzyme K+ CD8 T cells form a core population in inflamed human tissue. *Sci. Transl. Med.* **14**, eabo0686 (2022).

46. D. A. Rao, M. F. Gurish, J. L. Marshall, K. Slowikowski, C. Y. Fonseka, Y. Liu, L. T. Donlin, L. A. Henderson, K. Wei, F. Mizoguchi, N. C. Teslovich, M. E. Weinblatt, E. M. Massarotti, J. S. Coblyn, S. M. Helfgott, Y. C. Lee, D. J. Todd, V. P. Bykerk, S. M. Goodman, A. B. Pernis, L. B. Ivashkiv, E. W. Karlson, P. A. Nigrovic, A. Filer, C. D. Buckley, J. A. Lederer, S. Raychaudhuri, M. B. Brenner, Pathologically expanded peripheral T helper cell subset drives B cells in rheumatoid arthritis. *Nature* **542**, 110–114 (2017).

47. A. Gupta, K. Weinand, A. Nathan, S. Sakaue, M. J. Zhang, Accelerating Medicines Partnership RA/SLE Program and Network, L. Donlin, K. Wei, A. L. Price, T. Amariuta, S. Raychaudhuri, Dynamic regulatory elements in single-cell multimodal data implicate key immune cell states enriched for autoimmune disease heritability. *Nat. Genet.* **55**, 2200–2210 (2023).

48. C.-C. Liu, A. H. Kao, S. Manzi, J. M. Ahearn, Biomarkers in systemic lupus erythematosus: challenges and prospects for the future. *Ther. Adv. Musculoskelet. Dis.* **5**, 210–233 (2013).

49. H. Lazareth, C. Henique, O. Lenoir, V. G. Puelles, M. Flamant, G. Bollée, C. Fligny, M. Camus, L. Guyonnet, C. Millien, F. Gaillard, A. Chipont, B. Robin, S. Fabrega, N. Dhaun, E. Camerer, O. Kretz, F. Grahammer, F. Braun, T. B. Huber, D. Nochy, C. Mandet, P. Bruneval, L. Mesnard, E. Thervet, A. Karras, F. Le Naour, E. Rubinstein, C. Boucheix, A. Alexandrou, M. J. Moeller, C. Bouzigues, P.-L. Tharaux, The tetraspanin CD9 controls migration and proliferation of parietal epithelial cells and glomerular disease progression. *Nat. Commun.* **10**, 3303 (2019).

50. R. G. Evans, D. W. Smith, C.-J. Lee, J. P. Ngo, B. S. Gardiner, What makes the kidney susceptible to hypoxia? *Anat. Rec. (Hoboken)* **303**, 2544–2552 (2020).

51. W. Gao, D. Liu, H. Sun, Z. Shao, P. Shi, T. Li, S. Yin, T. Zhu, SPP1 is a prognostic related biomarker and correlated with tumor-infiltrating immune cells in ovarian cancer. *BMC Cancer* **22**, 1367 (2022).

52. R. Bill, P. Wirapati, M. Messemaker, W. Roh, B. Zitti, F. Duval, M. Kiss, J. C. Park, T. M. Saal, J. Hoelzl, D. Tarussio, F. Benedetti, S. Tissot, L. Kandalaft, M. Varrone, G. Ciriello, T. A. McKee, Y. Monnier, M. Mermod, E. M. Blaum, I. Gushterova, A. L. K. Gonye, N. Hacohen, G. Getz, T. R. Mempel, A. M. Klein, R. Weissleder, W. C. Faquin, P. M. Sadow, D. Lin, S. I. Pai, M. Sade-Feldman, M. J. Pittet, CXCL9:SPP1 macrophage polarity identifies a network of cellular programs that control human cancers. *Science* **381**, 515–524 (2023).

53. A.-P. Mao, I. E. Ishizuka, D. N. Kasal, M. Mandal, A. Bendelac, A shared Runx1-bound Zbtb16

enhancer directs innate and innate-like lymphoid lineage development. *Nat. Commun.* **8**, 863 (2017).

54. J. Banchereau, V. Pascual, Type I interferon in systemic lupus erythematosus and other autoimmune diseases. *Immunity* **25**, 383–392 (2006).

55. K. Albrecht, W. Troll, J. Callhoff, A. Strangfeld, S. Ohrndorf, J. Mucke, Sex- and gender-related differences in systemic lupus erythematosus: a scoping review. *Rheumatol. Int.* **45**, 160 (2025).

56. S. Singh, H. C. Sreenidhi, 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).

57. Y. Yang, F. Li, Y. Li, X. Li, Z. Zhao, N. Zhang, H. Li, 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).

58. W. Xu, L. Hou, Knockdown of nicotinamide N-methyltransferase ameliorates renal fibrosis caused by ischemia-reperfusion injury and remodels sphingosine metabolism. *Clin. Exp. Nephrol.* **28**, 1241–1253 (2024).

59. L. S. Gewin, Renal fibrosis: Primacy of the proximal tubule. *Matrix Biol.* **68-69**, 248–262 (2018).

60. M. Singh, G. Al-Eryani, S. Carswell, J. M. Ferguson, J. Blackburn, K. Barton, D. Roden, F. Luciani, T. Giang Phan, S. Junankar, K. Jackson, C. C. Goodnow, M. A. Smith, A. Swarbrick, High-throughput targeted long-read single cell sequencing reveals the clonal and transcriptional landscape of lymphocytes. *Nat. Commun.* **10**, 3120 (2019).

61. R. F. van Vollenhoven, M. Mosca, G. Bertsias, D. Isenberg, A. Kuhn, K. Lerstrøm, M. Aringer, H. Bootsma, D. Boumpas, I. N. Bruce, R. Cervera, A. Clarke, N. Costedoat-Chalumeau, L. Czirják, R. Derksen, T. Dörner, C. Gordon, W. Graninger, F. Houssiau, M. Inanc, S. Jacobsen, D. Jayne, A. Jedryka-Goral, A. Levitsky, R. Levy, X. Mariette, E. Morand, S. Navarra, I. Neumann, A. Rahman, J. Rovensky, J. Smolen, C. Vasconcelos, A. Voskuyl, A. Voss, H. Zakharova, A. Zoma, M. Schneider, Treat-to-target in systemic lupus erythematosus: recommendations from an international task force. *Ann. Rheum. Dis.* **73**, 958–967 (2014).

62. A. Alsuwaida, S. Husain, M. Alghonaim, N. AlOudah, J. Alwakeel, A. Ullah, H. Kfoury, Strategy for second kidney biopsy in patients with lupus nephritis. *Nephrol. Dial. Transplant* **27**, 1472–1478 (2012).

63. P.-S. Tsou, A. H. Sawalha, Glycoprotein nonmetastatic melanoma protein B: A key mediator and an emerging therapeutic target in autoimmune diseases. *FASEB J.* **34**, 8810–8823 (2020).

64. P. Ramachandran, R. Dobie, J. R. Wilson-Kanamori, E. F. Dora, B. E. P. Henderson, N. T. Luu, J. R. Portman, K. P. Matchett, M. Brice, J. A. Marwick, R. S. Taylor, M. Efremova, R. Vento-Tormo, N. O. Carragher, T. J. Kendall, J. A. Fallowfield, E. M. Harrison, D. J. Mole, S. J. Wigmore, P. N. Newsome, C. J. Weston, J. P. Iredale, F. Tacke, J. W. Pollard, C. P. Ponting, J. C. Marioni, S. A. Teichmann, N. C. Henderson, Resolving the fibrotic niche of human liver cirrhosis at single-cell level. *Nature* **575**, 512–518 (2019).

65. T. Fabre, A. M. S. Barron, S. M. Christensen, S. Asano, K. Bound, M. P. Lech, M. H. Wadsworth 2nd, X. Chen, C. Wang, J. Wang, J. McMahon, F. Schlerman, A. White, K. M. Kravarik, A. J. Fisher, L. A. Borthwick, K. M. Hart, N. C. Henderson, T. A. Wynn, K. Dower, Identification of a broadly

fibrogenic macrophage subset induced by type 3 inflammation. *Sci. Immunol.* **8**, eadd8945 (2023).

66. J. S. Munger, D. Sheppard, Cross talk among TGF-β signaling pathways, integrins, and the extracellular matrix. *Cold Spring Harb. Perspect. Biol.* **3**, a005017 (2011).

67. J. Yang, P. Antin, G. Berx, C. Blanpain, T. Brabletz, M. Bronner, K. Campbell, A. Cano, J. Casanova, G. Christofori, S. Dedhar, R. Derynck, H. L. Ford, J. Fuxe, A. García de Herreros, G. J. Goodall, A.-K. Hadjantonakis, R. Y. J. Huang, C. Kalcheim, R. Kalluri, Y. Kang, Y. Khew-Goodall, H. Levine, J. Liu, G. D. Longmore, S. A. Mani, J. Massagué, R. Mayor, D. McClay, K. E. Mostov, D. F. Newgreen, M. A. Nieto, A. Puisieux, R. Runyan, P. Savagner, B. Stanger, M. P. Stemmler, Y. Takahashi, M. Takeichi, E. Theveneau, J. P. Thiery, E. W. Thompson, R. A. Weinberg, E. D. Williams, J. Xing, B. P. Zhou, G. Sheng, EMT International Association (TEMTIA), Guidelines and definitions for research on epithelial-mesenchymal transition. *Nat. Rev. Mol. Cell Biol.* **21**, 341–352 (2020).

68. E. Cantero-Navarro, S. Rayego-Mateos, M. Orejudo, L. Tejedor-Santamaria, A. Tejera-Muñoz, A. B. Sanz, L. Marquez-Exposito, V. Marchant, L. Santos-Sanchez, J. Egido, A. Ortiz, T. Bellon, R. R. Rodrigues-Diez, M. Ruiz-Ortega, Role of macrophages and related cytokines in kidney disease. *Front. Med. (Lausanne)* **8**, 688060 (2021).

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

70. R. Chiquet-Ehrismann, M. Chiquet, Tenascins: regulation and putative functions during pathological stress. *J. Pathol.* **200**, 488–499 (2003).

71. A. Fava, J. Buyon, L. Magder, J. Hodgin, A. Rosenberg, D. S. Demeke, D. A. Rao, A. Arazi, A. I. Celia, C. Putterman, J. H. Anolik, J. Barnas, M. Dall'Era, D. Wofsy, R. Furie, D. Kamen, K. Kalunian, J. A. James, J. Guthridge, M. G. Atta, J. Monroy Trujillo, D. Fine, R. Clancy, H. M. Belmont, P. Izmirly, W. Apruzzese, D. Goldman, C. C. Berthier, P. Hoover, N. Hacohen, S. Raychaudhuri, A. Davidson, B. Diamond, Accelerating Medicines Partnership in RA/SLE network, M. Petri, Urine proteomic signatures of histological class, activity, chronicity, and treatment response in lupus nephritis. *JCI Insight* **9** (2024).

72. M. Ushikubo, S. Saito, J. Kikuchi, M. Takeshita, K. Yoshimoto, H. Yasuoka, K. Yamaoka, N. Seki, K. Suzuki, H. Oshima, T. Takeuchi, 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).

73. W.-Y. Baek, J.-M. Woo, H.-A. Kim, J.-Y. Jung, C.-H. Suh, 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).

74. D. W. Erikson, R. C. Burghardt, K. J. Bayless, G. A. Johnson, Secreted phosphoprotein 1 (SPP1, osteopontin) binds to integrin alpha v beta 6 on porcine trophectoderm cells and integrin alpha v beta 3 on uterine luminal epithelial cells, and promotes trophectoderm cell adhesion and migration. *Biol. Reprod.* **81**, 814–825 (2009).

75. J. Kim, D. W. Erikson, R. C. Burghardt, T. E. Spencer, G. Wu, K. J. Bayless, G. A. Johnson, F. W. Bazer, Secreted phosphoprotein 1 binds integrins to initiate multiple cell signaling pathways, including FRAP1/mTOR, to support attachment and force-generated migration of trophectoderm cells. *Matrix Biol.* **29**, 369–382 (2010).

76. D. Katoh, K. Nagaharu, N. Shimojo, N. Hanamura, M. Yamashita, Y. Kozuka, K. Imanaka-Yoshida, T. Yoshida, Binding of αvβ1 and αvβ6 integrins to tenascin-C induces epithelial-mesenchymal transition-like change of breast cancer cells. *Oncogenesis* **2**, e65 (2013).

77. T. Yoshida, T. Akatsuka, K. Imanaka-Yoshida, Tenascin-C and integrins in cancer. *Cell Adh. Migr.* **9**, 96–104 (2015).

78. S. Lovisa, V. S. LeBleu, B. Tampe, H. Sugimoto, K. Vadnagara, J. L. Carstens, C.-C. Wu, Y. Hagos, B. C. Burckhardt, T. Pentcheva-Hoang, H. Nischal, J. P. Allison, M. Zeisberg, R. Kalluri, Epithelial-to-mesenchymal transition induces cell cycle arrest and parenchymal damage in renal fibrosis. *Nat. Med.* **21**, 998–1009 (2015).

79. A. Malvar, V. Alberton, B. Lococo, S. J. Martinez, I. Chidid, R. Heguilen, B. H. Rovin, The evolution of crescents in lupus nephritis. *Kidney Int. Rep.* **10**, 1579–1581 (2025).

80. E. C. Baechler, F. M. Batliwalla, G. Karypis, P. M. Gaffney, W. A. Ortmann, K. J. Espe, K. B. Shark, W. J. Grande, K. M. Hughes, V. Kapur, P. K. Gregersen, T. W. Behrens, 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).

81. K. Migita, T. Miyashita, Y. Maeda, H. Kimura, M. Nakamura, H. Yatsuhashi, H. Ishibashi, K. Eguchi, Reduced blood BDCA-2+ (lymphoid) and CD11c+ (myeloid) dendritic cells in systemic lupus erythematosus. *Clin. Exp. Immunol.* **142**, 84–91 (2005).

82. S. Blomberg, M.-L. Eloranta, M. Magnusson, G. V. Alm, L. Rönnblom, 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).

83. W. Zhang, H. Zhang, S. Liu, F. Xia, Z. Kang, Y. Zhang, Y. Liu, H. Xiao, L. Chen, C. Huang, N. Shen, H. Xu, F. Li, 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).

84. A. Horisberger, A. Griffith, J. Keegan, A. Arazi, J. Pulford, E. Murzin, K. Howard, B. Hancock, A. Fava, T. Sasaki, T. Ghosh, J. Inamo, R. Beuschel, Y. Cao, K. Preisinger, M. Gutierrez-Arcelus, T. M. Eisenhaure, J. Guthridge, P. J. Hoover, M. Dall'Era, D. Wofsy, D. L. Kamen, K. C. Kalunian, R. Furie, M. Belmont, P. Izmirly, R. Clancy, D. Hildeman, E. S. Woodle, W. Apruzzese, M. A. McMahon, J. Grossman, J. L. Barnas, F. Payan-Schober, M. Ishimori, M. Weisman, M. Kretzler, C. C. Berthier, J. B. Hodgin, D. S. Demeke, C. Putterman, M. B. Brenner, J. H. Anolik, S. Raychaudhuri, N. Hacohen, J. A. James, A. Davidson, M. A. Petri, J. P. Buyon, B. Diamond, F. Zhang, J. A. Lederer, D. A. Rao, Blood immunophenotyping identifies distinct kidney histopathology and outcomes in patients with lupus nephritis. *J. Clin. Invest.*, doi: 10.1172/JCI181034 (2025).

85. I. C. Mouat, E. Goldberg, M. S. Horwitz, Age-associated B cells in autoimmune diseases. *Cell. Mol. Life Sci.* **79**, 402 (2022).

86. J. Hunnicutt, M. E. Georgiou, A. Richards, H. Quasny, L. Magder, D. Goldman, M. A. Petri, 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).

87. A. Subramanian, K. A. Vernon, Y. Zhou, J. L. Marshall, M. Alimova, C. Arevalo, F. Zhang, M. Slyper, J. Waldman, M. S. Montesinos, D. Dionne, L. T. Nguyen, M. S. Cuoco, D. Dubinsky, J. Purnell, K. Keller, S. H. Sturner, E. Grinkevich, A. Ghoshal, A. Kotek, G. Trivioli, N. Richoz, M. B. Humphrey, I. G. Darby, S. J. Miller, Y. Xu, A. Weins, A. Chloe-Villani, S. L. Chang, M. Kretzler, O. Rosenblatt-Rosen, J. L. Shaw, K. A. Zimmerman, M. R. Clatworthy, A. Regev, A. Greka, Protective role for kidney TREM2high macrophages in obesity- and diabetes-induced kidney injury. *Cell Rep.* **43**, 114253 (2024).

88. R. Xu, N. Vujić, V. Bianco, I. Reinisch, D. Kratky, J. Krstic, A. Prokesch, Lipid-associated macrophages between aggravation and alleviation of metabolic diseases. *Trends Endocrinol. Metab.* **35**, 981–995 (2024).

89. A. Reggio, C. Fuoco, R. Deodati, A. Palma, SPP1 macrophages across diseases: A call for reclassification? *FASEB J.* **39**, e70448 (2025).

90. M. Cruz Tleugabulova, S. P. Melo, A. Wong, A. Arlantico, M. Liu, J. D. Webster, J. Lau, A. Lechner, B. Corak, J. J. Hodgins, V. S. Garlapati, M. De Simone, B. Korin, S. Avraham, J. Lund, S. Jeet, A. Reiss, H. Bender, C. D. Austin, S. Darmanis, Z. Modrusan, H. Brightbill, S. Durinck, M. S. Diamond, C. Schneider, A. S. Shaw, M. Nitschké, Induction of a distinct macrophage population and protection from lung injury and fibrosis by Notch2 blockade. *Nat. Commun.* **15**, 9575 (2024).

91. E. M. King, Y. Zhao, C. M. Moore, B. Steinhart, K. C. Anderson, B. Vestal, P. K. Moore, S. A. McManus, C. M. Evans, K. J. Mould, E. F. Redente, A. L. McCubbrey, W. J. Janssen, Gpnmb and Spp1 mark a conserved macrophage injury response masking fibrosis-specific programming in the lung. *JCI Insight* **9** (2024).

92. S.-I. Biniaris-Georgallis, T. Aschman, K. Stergioula, F. Schreiber, V. Jafari, A. Taranko, T. Karmalkar, A. Kasapi, T. Lenac Rovis, V. Jelencic, D. A. Bejarano, L. Fabry, M. Papacharalampous, I. Mattiola, M. Molgora, J. Hou, K. W. Hublitz, F. Heinrich, G. M. Guerra, P. Durek, G. Patone, E. L. Lindberg, H. Maatz, O. Hölsken, G. Krönke, A. Mortha, R. E. Voll, A. J. Clarke, A. E. Hauser, M. Colonna, K. Thurley, A. Schlitzer, C. Schneider, E. G. Stamatiades, M.-F. Mashreghi, S. Jonjic, N. Hübner, A. Diefenbach, M. Kanda, A. Triantafyllopoulou, Amplification of autoimmune organ damage by NKp46-activated ILC1s. *Nature* **634**, 952–960 (2024).

93. A. Palma, The landscape of SPP1 + macrophages across tissues and diseases: A comprehensive review. *Immunology* **176**, 179–196 (2025).

94. F. F. De Ponti, A. Bujko, Z. Liu, P. J. Collins, S. Schuermans, C. Maueroder, S. Amstelveen, T. Thoné, L. Martens, J. G. McKendrick, P. A. Louwe, A. Sànchez Cruz, W. Saelens, K. P. Matchett, K. J. Waller, C. Zwicker, A. Buglar-Lamb, B. Vanneste, F. Parmentier, M. Binte Abdul Latib, A. Remmerie, L. Kertesz, A. Kremer, J. Verbeke, D. H. Ipsen, D. R. Pfister, Z. Liu, M. Guilliams, N. C. Henderson, K. Ravichandran, P. E. Marques, C. L. Scott, Spatially restricted and ontogenically distinct hepatic macrophages are required for tissue repair. *Immunity* **58**, 362–380.e10 (2025).

95. A. Silvin, S. Uderhardt, C. Piot, S. Da Mesquita, K. Yang, L. Geirsdottir, K. Mulder, D. Eyal, Z. Liu, C. Bridlance, M. S. Thion, X. M. Zhang, W. T. Kong, M. Deloger, V. Fontes, A. Weiner, R. Ee, R. Dress, J. W. Hang, A. Balachander, S. Chakarov, B. Malleret, G. Dunsmore, O. Cexus, J. Chen, S. Garel, C. A. Dutertre, I. Amit, J. Kipnis, F. Ginhoux, Dual ontogeny of disease-associated microglia and disease inflammatory macrophages in aging and neurodegeneration. *Immunity* **55**, 1448–1465.e6 (2022).

96. P. J. Hoover, C. Raparia, D. J. Lieb, Y. Tzur, J. Kang, A. Arazi, R. Leavitt, R. Mishra, S. I. Shah, D.

Simmons, S. Li, M. Peters, T. Eisenhaure, T. J. Few-Cooper, S. S. Gurajala, A. Sonny, J. B. Hodgin, C. C. Berthier, J. M. Guthridge, A. Fava, R. M. Clancy, C. Putterman, P. M. Izmirly, H. M. Belmont, K. Kalunian, D. Kamen, D. Wofsy, J. P. Buyon, J. A. James, M. Petri, B. Diamond, S. Raychaudhuri, S. S. Shen-Orr, Accelerating Medicines Partnership in RA/SLE network, N. Hacohen, A. Davidson, A human-mouse atlas of intrarenal myeloid cells identifies conserved disease-associated macrophages in lupus nephritis. *J. Exp. Med.* **222**, e20241873 (2025).

97. R. Furie, M. Petri, O. Zamani, R. Cervera, D. J. Wallace, D. Tegzová, J. Sanchez-Guerrero, A. Schwarting, J. T. Merrill, W. W. Chatham, W. Stohl, E. M. Ginzler, D. R. Hough, Z. J. Zhong, W. Freimuth, R. F. van Vollenhoven, BLISS-76 Study Group, 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).

98. RePORT 〉 RePORTER. https://reporter.nih.gov/search/n17gcxvd0EiSaYry0_ayxg/project-details/10452169.

99. D. Janosevic, T. De Luca, Kidney Precision Medicine Project, M. T. Eadon, The Kidney Precision Medicine Project and single-cell biology of the injured proximal tubule. *Am. J. Pathol.* **195**, 7–22 (2025).

100. P. C. Breda, T. Wiech, C. Meyer-Schwesinger, F. Grahammer, T. Huber, U. Panzer, G. Tiegs, K. Neumann, Renal proximal tubular epithelial cells exert immunomodulatory function by driving inflammatory CD4+ T cell responses. *Am. J. Physiol. Renal Physiol.* **317**, F77–F89 (2019).

101. S. Liu, Y. Zhao, S. Lu, T. Zhang, M. T. Lindenmeyer, V. Nair, S. E. Gies, G. Wu, R. G. Nelson, J. Czogalla, H. Aypek, S. Zielinski, Z. Liao, M. Schaper, D. Fermin, C. D. Cohen, D. Delic, C. F. Krebs, F. Grahammer, T. Wiech, M. Kretzler, C. Meyer-Schwesinger, S. Bonn, T. B. Huber, Single-cell transcriptomics reveals a mechanosensitive injury signaling pathway in early diabetic nephropathy. *Genome Med.* **15**, 2 (2023).

102. Y. Hao, S. Hao, E. Andersen-Nissen, W. M. Mauck 3rd, S. Zheng, A. Butler, M. J. Lee, A. J. Wilk, C. Darby, M. Zager, P. Hoffman, M. Stoeckius, E. Papalexi, E. P. Mimitou, J. Jain, A. Srivastava, T. Stuart, L. M. Fleming, B. Yeung, A. J. Rogers, J. M. McElrath, C. A. Blish, R. Gottardo, P. Smibert, R. Satija, Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).

103. P. M. Carlucci, J. Li, A. Fava, K. K. Deonaraine, D. Wofsy, J. A. James, C. Putterman, B. Diamond, A. Davidson, D. M. Fine, J. Monroy-Trujillo, M. G. Atta, W. DeJager, J. M. Guthridge, K. Haag, D. A. Rao, M. B. Brenner, J. A. Lederer, W. Apruzzese, H. M. Belmont, P. M. Izmirly, D. Zaminski, M. Wu, S. Connery, F. Payan-Schober, R. Furie, M. Dall'Era, K. Cho, D. Kamen, K. Kalunian, J. Anolik, J. Barnas, M. Ishimori, M. H. Weisman, Accelerating Medicines Partnership (AMP) RA/SLE Network, J. P. Buyon, M. Petri, High incidence of proliferative and membranous nephritis in SLE patients with low proteinuria in the Accelerating Medicines Partnership. *Rheumatology (Oxford)* **61**, 4335–4343 (2022).

104. K. K. Deonaraine, P. M. Carlucci, A. Fava, J. Li, D. Wofsy, J. A. James, C. Putterman, B. Diamond, A. Davidson, D. M. Fine, J. Monroy-Trujillo, M. G. Atta, K. Haag, D. A. Rao, W. Apruzzese, H. M. Belmont, P. M. Izmirly, M. Wu, S. Connery, F. Payan-Schober, R. A. Furie, C. C. Berthier, M. Dall'Era, K. Cho, D. L. Kamen, K. Kalunian, J. Anolik, M. Ishimori, M. H. Weisman, Accelerating Medicines Partnership RA/SLE network, M. A. Petri, J. P. Buyon, 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).

105. M. C. Hochberg, Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum.* **40**, 1725 (1997).

106. M. Petri, A.-M. Orbai, G. S. Alarcón, C. Gordon, J. T. Merrill, P. R. Fortin, I. N. Bruce, D. Isenberg, D. J. Wallace, O. Nived, G. Sturfelt, R. Ramsey-Goldman, S.-C. Bae, J. G. Hanly, J. Sánchez-Guerrero, A. Clarke, C. Aranow, S. Manzi, M. Urowitz, D. Gladman, K. Kalunian, M. Costner, V. P. Werth, A. Zoma, S. Bernatsky, G. Ruiz-Irastorza, M. A. Khamashta, S. Jacobsen, J. P. Buyon, P. Maddison, M. A. Dooley, R. F. van Vollenhoven, E. Ginzler, T. Stoll, C. Peschken, J. L. Jorizzo, J. P. Callen, S. S. Lim, B. J. Fessler, M. Inanc, D. L. Kamen, A. Rahman, K. Steinsson, A. G. Franks Jr, L. Sigler, S. Hameed, H. Fang, N. Pham, R. Brey, M. H. Weisman, G. McGwin Jr, L. S. Magder, Derivation and validation of the Systemic Lupus International Collaborating Clinics classification criteria for systemic lupus erythematosus. *Arthritis Rheum.* **64**, 2677–2686 (2012).

107. J. J. Weening, V. D. D'Agati, M. M. Schwartz, S. V. Seshan, C. E. Alpers, G. B. Appel, J. E. Balow, J. A. Bruijn, T. Cook, F. Ferrario, A. B. Fogo, E. M. Ginzler, L. Hebert, G. Hill, P. Hill, J. C. Jennette, N. C. Kong, P. Lesavre, M. Lockshin, L.-M. Looi, H. Makino, L. A. Moura, M. Nagata, International Society of Nephrology Working Group on the Classification of Lupus Nephritis, Renal Pathology Society Working Group on the Classification of Lupus Nephritis, The classification of glomerulonephritis in systemic lupus erythematosus revisited. *Kidney Int.* **65**, 521–530 (2004).

108. I. Korsunsky, A. Nathan, N. Millard, S. Raychaudhuri, Presto scales Wilcoxon and auROC analyses to millions of observations, *Bioinformatics* (2019). https://www.biorxiv.org/content/10.1101/653253v1.abstract.

109. A. Nathan, J. I. Beynor, Y. Baglaenko, S. Suliman, K. Ishigaki, S. Asgari, C.-C. Huang, Y. Luo, Z. Zhang, K. Lopez, C. S. Lindestam Arlehamn, J. D. Ernst, J. Jimenez, R. I. Calderón, L. Lecca, I. Van Rhijn, D. B. Moody, M. B. Murray, S. Raychaudhuri, Multimodally profiling memory T cells from a tuberculosis cohort identifies cell state associations with demographics, environment and disease. *Nat. Immunol.* **22**, 781–793 (2021).

110. A. A. Alli, D. Desai, A. Elshika, M. Conrad, B. Proneth, W. Clapp, C. Atkinson, M. Segal, L. A. Searcy, N. D. Denslow, S. Bolisetty, B. Mehrad, L. Morel, Y. Scindia, Kidney tubular epithelial cell ferroptosis links glomerular injury to tubulointerstitial pathology in lupus nephritis. *Clin. Immunol.* **248**, 109213 (2023).

111. A.-C. Villani, R. Satija, G. Reynolds, S. Sarkizova, K. Shekhar, J. Fletcher, M. Griesbeck, A. Butler, S. Zheng, S. Lazo, L. Jardine, D. Dixon, E. Stephenson, E. Nilsson, I. Grundberg, D. McDonald, A. Filby, W. Li, P. L. De Jager, O. Rozenblatt-Rosen, A. A. Lane, M. Haniffa, A. Regev, N. Hacohen, Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356** (2017).

112. D. Kotliar, M. Curtis, R. Agnew, K. Weinand, A. Nathan, Y. Baglaenko, K. Slowikowski, Y. Zhao, P. C. Sabeti, D. A. Rao, S. Raychaudhuri, Reproducible single-cell annotation of programs underlying T cell subsets, activation states and functions. *Nat. Methods* **22**, 1964–1980 (2025).

113. J. W. Squair, M. Gautier, C. Kathe, M. A. Anderson, N. D. James, T. H. Hutson, R. Hudelle, T. Qaiser, K. J. E. Matson, Q. Barraud, A. J. Levine, G. La Manno, M. A. Skinnider, G. Courtine, Confronting false discoveries in single-cell differential expression. *Nat. Commun.* **12**, 5692 (2021).

114. ACCESS Trial Group, Treatment of lupus nephritis with abatacept: the Abatacept and Cyclophosphamide Combination Efficacy and Safety Study. *Arthritis Rheumatol.* **66**, 3096–3104 (2014).

115. C. Chatelle, Y. G. Bodien, C. Carlowicz, S. Wannez, V. Charland-Verville, O. Gosseries, S. Laureys, R. T. Seel, J. T. Giacino, Detection and interpretation of impossible and improbable Coma Recovery Scale-Revised scores. *Arch. Phys. Med. Rehabil.* **97**, 1295–1300.e4 (2016).

116. J. G. Doench, N. Fusi, M. Sullender, M. Hegde, E. W. Vaimberg, K. F. Donovan, I. Smith, Z. Tothova, C. Wilen, R. Orchard, H. W. Virgin, J. Listgarten, D. E. Root, Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* **34**, 184–191 (2016).

117. K. R. Sanson, R. E. Hanna, M. Hegde, K. F. Donovan, C. Strand, M. E. Sullender, E. W. Vaimberg, A. Goodale, D. E. Root, F. Piccioni, J. G. Doench, Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. *Nat. Commun.* **9**, 5416 (2018).

118. A. A. Hardigan, B. S. Roberts, D. E. Moore, R. C. Ramaker, A. L. Jones, R. M. Myers, CRISPR/Cas9-targeted removal of unwanted sequences from small-RNA sequencing libraries. *Nucleic Acids Res.* **47**, e84 (2019).

119. B. B. Lake, S. Chen, M. Hoshi, N. Plongthongkum, D. Salamon, A. Knoten, A. Vijayan, R. Venkatesh, E. H. Kim, D. Gao, J. Gaut, K. Zhang, S. Jain, A single-nucleus RNA-sequencing pipeline to decipher the molecular anatomy and pathophysiology of human kidneys. *Nat. Commun.* **10**, 2832 (2019).

120. H. M. Kang, M. Subramaniam, S. Targ, M. Nguyen, L. Maliskova, E. McCarthy, E. Wan, S. Wong, L. Byrnes, C. M. Lanata, R. E. Gate, S. Mostafavi, A. Marson, N. Zaitlen, L. A. Criswell, C. J. Ye, Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nat. Biotechnol.* **36**, 89–94 (2018).

121. F. Zhang, J. R. Mears, L. Shakib, J. I. Beynor, S. Shanaj, I. Korsunsky, A. Nathan, Accelerating Medicines Partnership Rheumatoid Arthritis and Systemic Lupus Erythematosus (AMP RA/SLE) Consortium, L. T. Donlin, S. Raychaudhuri, 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).

122. E. Becht, L. McInnes, J. Healy, C.-A. Dutertre, I. W. H. Kwok, L. G. Ng, F. Ginhoux, E. W. Newell, Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* **37**, 38–44 (2018).

123. S. L. Wolock, R. Lopez, A. M. Klein, Scrublet: Computational identification of cell Doublets in Single-cell transcriptomic data. *Cell Syst.* **8**, 281–291.e9 (2019).

124. F. Zhang, A. H. Jonsson, A. Nathan, N. Millard, M. Curtis, Q. Xiao, M. Gutierrez-Arcelus, W. Apruzzese, G. F. M. Watts, D. Weisenfeld, S. Nayar, J. Rangel-Moreno, N. Meednu, K. E. Marks, I. Mantel, J. B. Kang, L. Rumker, J. Mears, K. Slowikowski, K. Weinand, D. E. Orange, L. Geraldino-Pardilla, K. D. Deane, D. Tabechian, A. Ceponis, G. S. Firestein, M. Maybury, I. Sahbudin, A. Ben-Artzi, A. M. Mandelin 2nd, A. Nerviani, M. J. Lewis, F. Rivellese, C. Pitzalis, L. B. Hughes, D. Horowitz, E. DiCarlo, E. M. Gravallese, B. F. Boyce, Accelerating Medicines Partnership: RA/SLE Network, L. W. Moreland, S. M. Goodman, H. Perlman, V. M. Holers, K. P. Liao, A. Filer, V. P. Bykerk, K. Wei, D. A. Rao, L. T. Donlin, J. H. Anolik, M. B. Brenner, S. Raychaudhuri, Deconstruction of rheumatoid arthritis synovium defines inflammatory subtypes. *Nature* **623**, 616–624 (2023).

125. I. Tirosh, B. Izar, S. M. Prakadan, M. H. Wadsworth 2nd, D. Treacy, J. J. Trombetta, A. Rotem, C.

Rodman, C. Lian, G. Murphy, M. Fallahi-Sichani, K. Dutton-Regester, J.-R. Lin, O. Cohen, P. Shah, D. Lu, A. S. Genshaft, T. K. Hughes, C. G. K. Ziegler, S. W. Kazer, A. Gaillard, K. E. Kolb, A.-C. Villani, C. M. Johannessen, A. Y. Andreev, E. M. Van Allen, M. Bertagnolli, P. K. Sorger, R. J. Sullivan, K. T. Flaherty, D. T. Frederick, J. Jané-Valbuena, C. H. Yoon, O. Rozenblatt-Rosen, A. K. Shalek, A. Regev, L. A. Garraway, Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).

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

127. S. Jin, C. F. Guerrero-Juarez, L. Zhang, I. Chang, R. Ramos, C.-H. Kuan, P. Myung, M. V. Plikus, Q. Nie, Inference and analysis of cell-cell communication using CellChat. *Nat. Commun.* **12**, 1088 (2021).

Acknowledgements

This work was generously supported by the Accelerating Medicines Partnership® Rheumatoid Arthritis and Systemic Lupus Erythematosus (AMP® RA/SLE) Network. AMP is a public-private partnership (AbbVie, Arthritis Foundation, Bristol-Myers Squibb Company, Foundation for the National Institutes of Health, GlaxoSmithKline, Janssen Research and Development, LLC, Lupus Foundation of America, Lupus Research Alliance, Merck & Co., Inc., National Institute of Allergy and Infectious Diseases, National Institute of Arthritis and Musculoskeletal and Skin Diseases, Pfizer, Inc., Rheumatology Research Foundation, Sanofi and Takeda Pharmaceuticals International, Inc.) created to develop new ways of identifying and validating promising biological targets for diagnostics and drug development. Accelerating Medicines Partnership and AMP are registered service marks of the US Department of Health and Human Services.

Funding:

National Institutes of Health UH2-AR067676 (AMP RA/SLE)

National Institutes of Health UH2-AR067677 (AMP RA/SLE)

National Institutes of Health UH2-AR067679 (AMP RA/SLE)

National Institutes of Health UH2-AR067681 (AMP RA/SLE)

National Institutes of Health UH2-AR067685 (AMP RA/SLE)

National Institutes of Health UH2- AR067688 (AMP RA/SLE)

National Institutes of Health UH2-AR067689 (AMP RA/SLE)

National Institutes of Health UH2-AR067690 (AMP RA/SLE)

National Institutes of Health UH2-AR067691 (AMP RA/SLE)

National Institutes of Health UH2-AR067694 (AMP RA/SLE)

National Institutes of Health UM2- AR067678 (AMP RA/SLE)

Author Contributions

Conceptualization: DAR, AF, TME, J. Gunthridge, PJH, MD, DW, DLK, KCK, RF, MB, PI, RC, DH, ESW, WA, MAM, J.Grossman, CC, HP, JLB, FP-S, CCB, JBH, CP, JHA, SR, NH, JAJ, AD, MAP, JPB, BD

Histology: JBH

scRNA and snRNA Data: TME, NH, MP, RR

Single-cell pipeline: SG, JM, QX, AA, SS, SR

Statistical Analysis: SG, NWS, MC, YZ, SS, SR, AF

Histological Analysis: BR, AF

Website: NWS

Supervision: SR, BD, JLB, JPB, MAP, AD, DAR

Writing – Original Draft: SG, NWS, MC, SR, BD

Writing – Review and Editing: All authors contributed

AMP RA/SLE Network members contributed to this work by managing patient recruitment, curating clinical data, obtaining and processing synovial tissue samples, managing biorepositories, conducting histologic or computational analysis, providing software code, providing website support and/or providing input on data analysis and interpretation.

Competing interests

SR is a founder for Mestag, Inc, on advisory boards for Pfizer, Janssen and Sonoma, and a consultant for Abbvie, Biogen, Nimbus and Magnet. JAJ has served as an advisor for GSK and Novartis.

Code availability

The code for analyzing this dataset can be found on our GitHub:

https://github.com/immunogenomics/AMP_SLE_Kidney.

Data availability

In accordance with AMP RA/SLE policies the processed count matrices will be made available to the public via the ARK Portal (https://arkportal.synapse.org/) once this study has been accepted for publication in a peer-reviewed journal.

The results published here are in whole or in part based on data obtained from the ARK Portal (http://arkportal.synapse.org). The Accelerating Medicines Partnership® RA/SLE Network data used for this publication are available under:

1. scRNAseq (https://arkportal.synapse.org/Explore/Datasets/DetailsPage?id=54406421)
2. snRNAseq (https://arkportal.synapse.org/Explore/Datasets/DetailsPage?id=54407831)

The ARK Portal hosts data generated by a network of research teams working collaboratively to deepen the understanding of Arthritis and Autoimmune and Related Diseases. It was established by the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) and includes data from the Accelerating Medicines Partnership® (AMP®) RA/SLE program.

The specific data used in this publication is available as a controlled-access dataset. Researchers seeking to use these data must submit:

1. A detailed intended data use statement
2. A completed and signed data use certificate

These access requirements ensure responsible and ethical data sharing within the research community. Instructions for access are available at https://help.arkportal.org/help/data-use-certificate#DataUse&Acknowledgement-Acknowledgement.

An interactive portal with marker gene visualization is available on our website: https://immunogenomics.io/ampsle2/app/?ds=b_pbmcs.

Figures / Legends

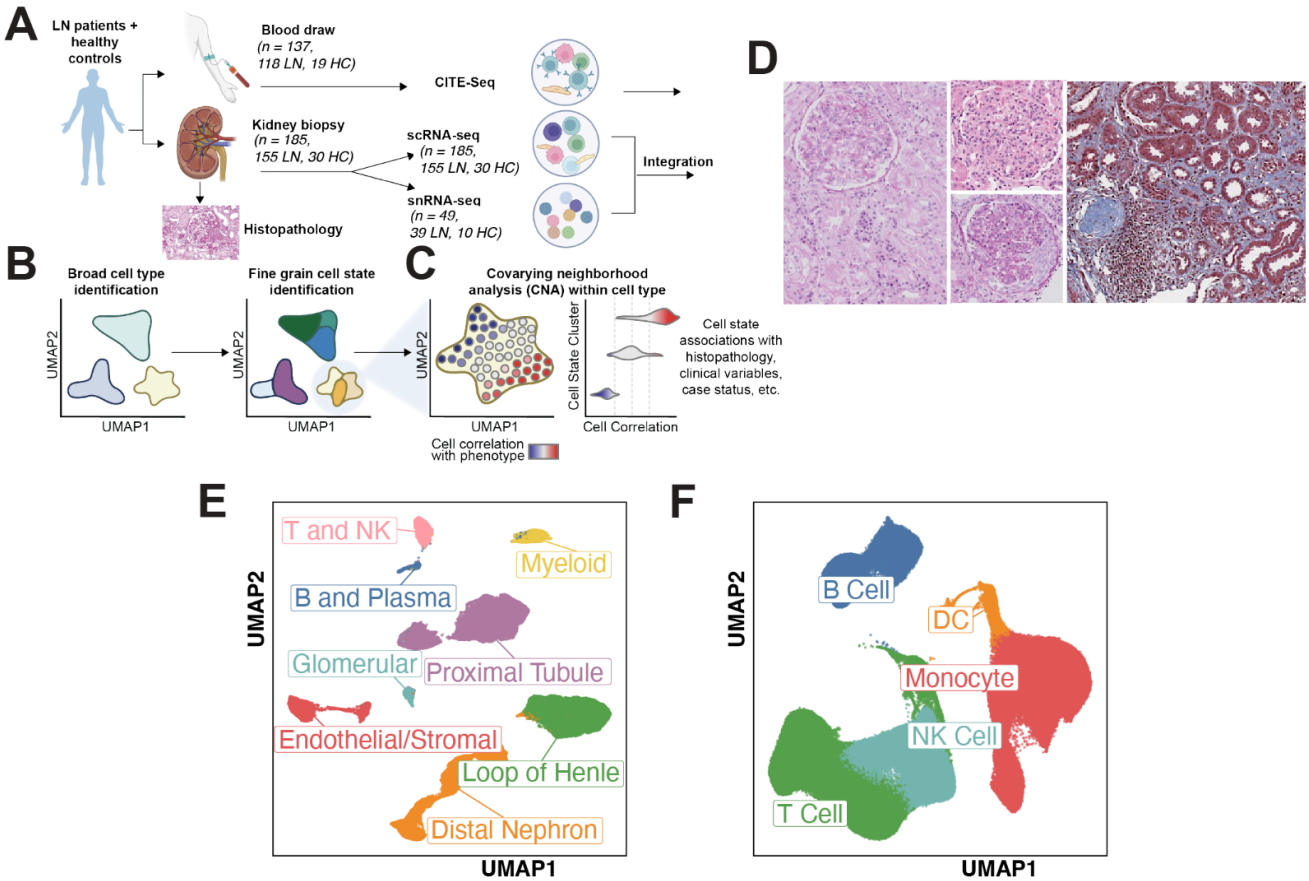

Fig. 1. Dataset overview. A-C, Schematic overview of the data collection and analysis pipeline. (**A**) Blood and kidney samples were collected from LN patients and healthy controls (HC). Histopathology and clinical scoring was performed on renal biopsies. CITE-seq was performed on blood samples, and 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. **(D)** 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. **(E-F),** UMAPs of all single cell or single nuclear annotations in tissue (**E**) and blood (**F**), colored by broad cell type annotation.

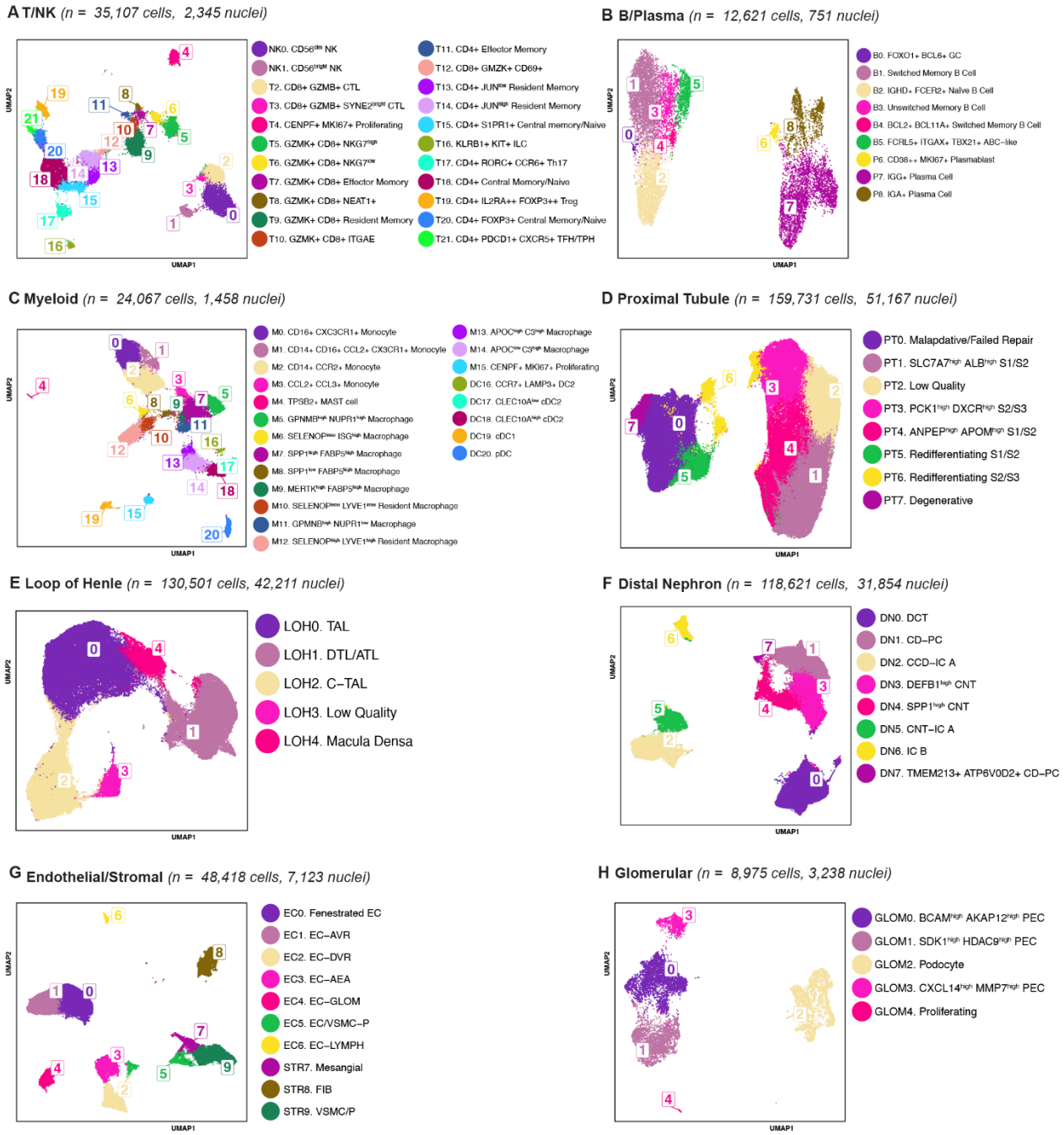

Fig. 2. Cell state annotations in tissue. (A) UMAP of T/NK cell states in tissue, colored by fine grain cell state annotations. Annotation boxes are located at the centroid (mean) of each cluster. **(B-H),** Same as **(A)** for B/plasma, myeloid, proximal tubule, loop of Henle, distal nephron, endothelial/stromal, and glomerular cells, respectively.

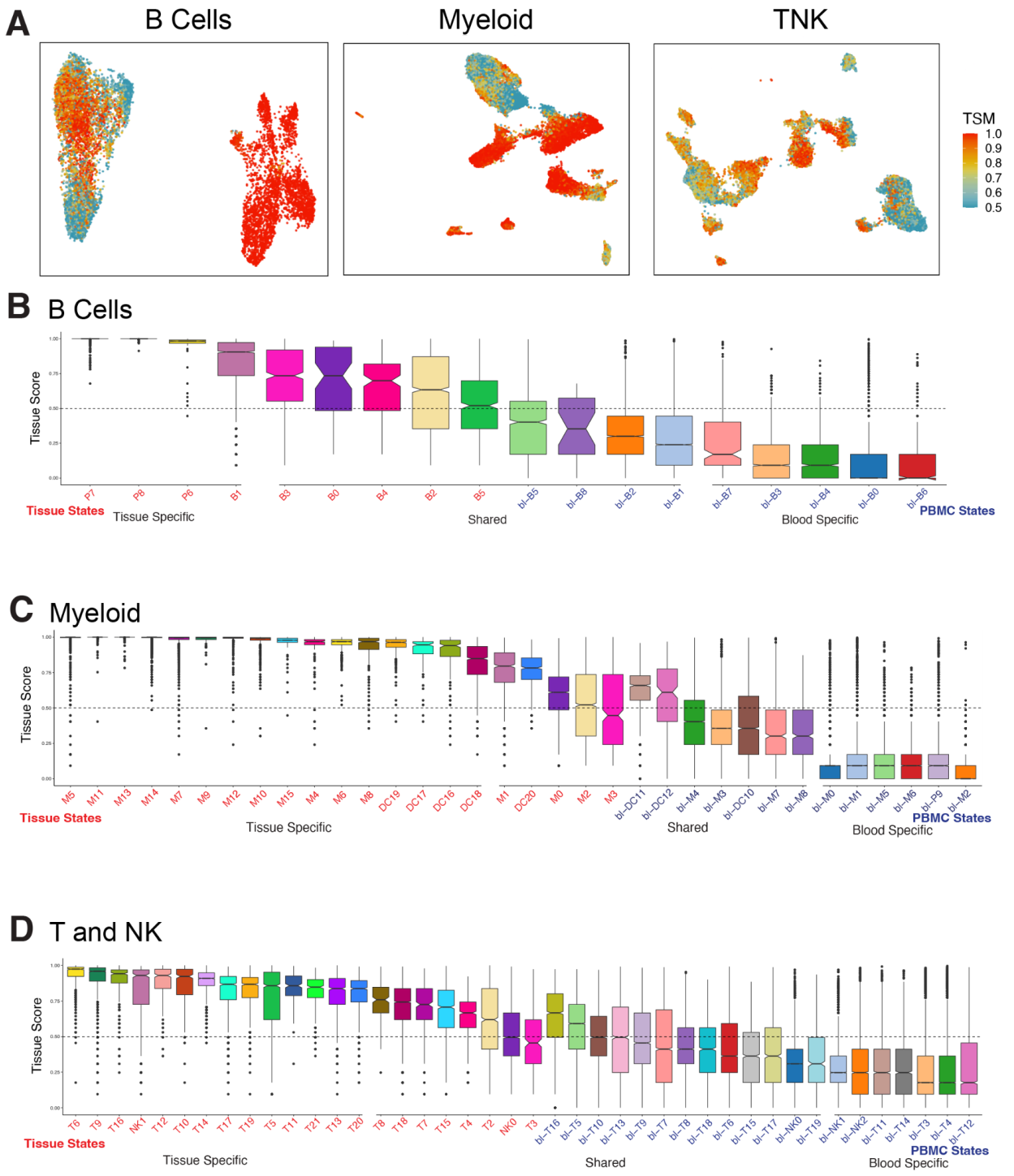

Fig. 3. Tissue Specificity Metric. (A) UMAPs of TSM score for immune cells from kidney tissue for B cells (left), myeloid cells (center) and T and NK cells (right). (**B-D)** Tissue specificity metric of tissue and blood immune cells stratified by cluster for B cells **(B),** Myeloid cells **(C)**, and T and NK cells **(D)**.

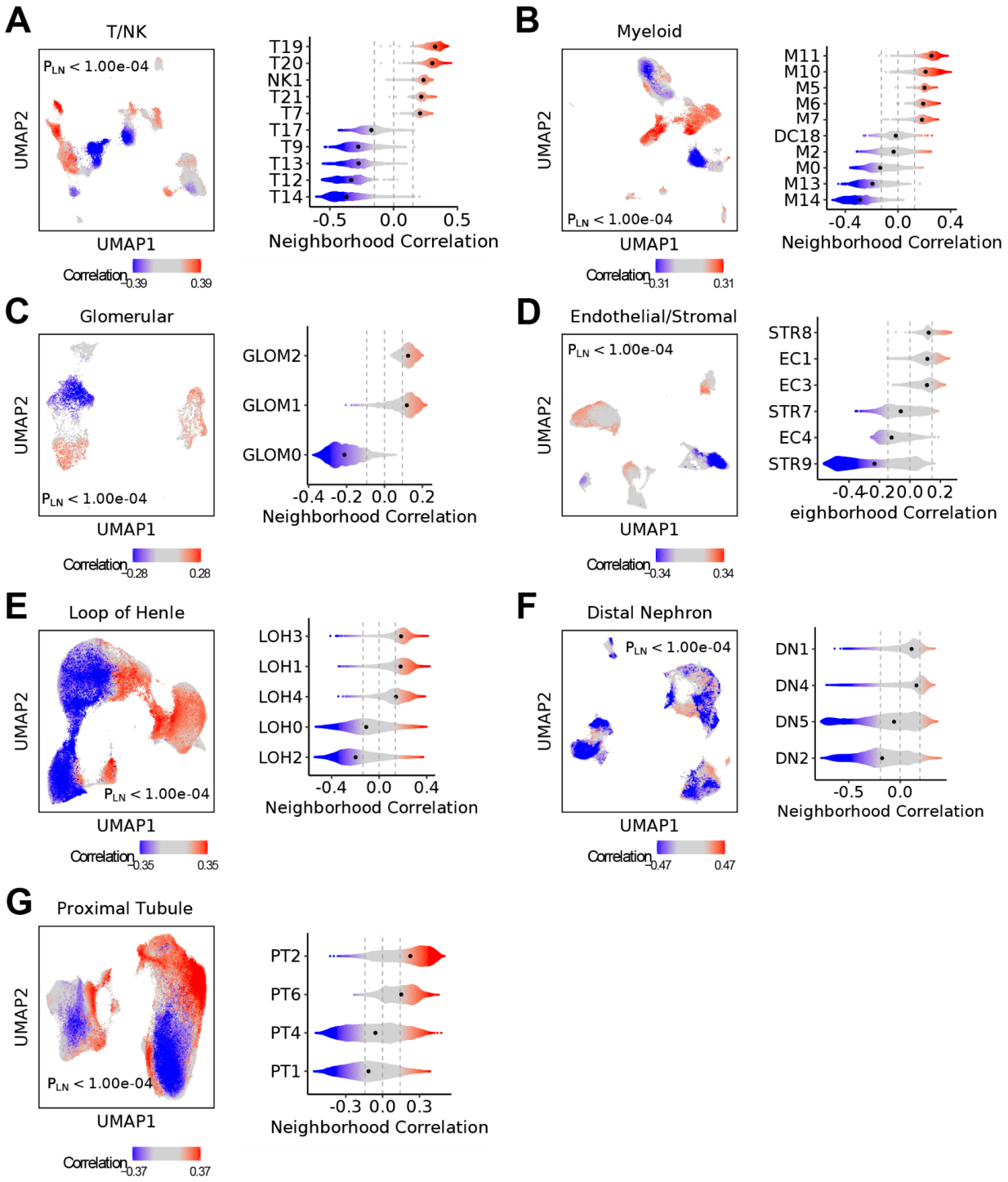

Fig. 4. Case-control associations within cell types. A-G CNA results for case-control association with no covariate corrections, in scRNA-seq data. Left - UMAP displaying significant per-cell associations with LN, with FDR cutoff of 0.05. Blue and red denote contracted and expanded neighborhoods that 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),** endothelial/stromal cells, (**E),** loop of Henle cells, (**F),** distal nephron, **(G),** proximal tubule cells.

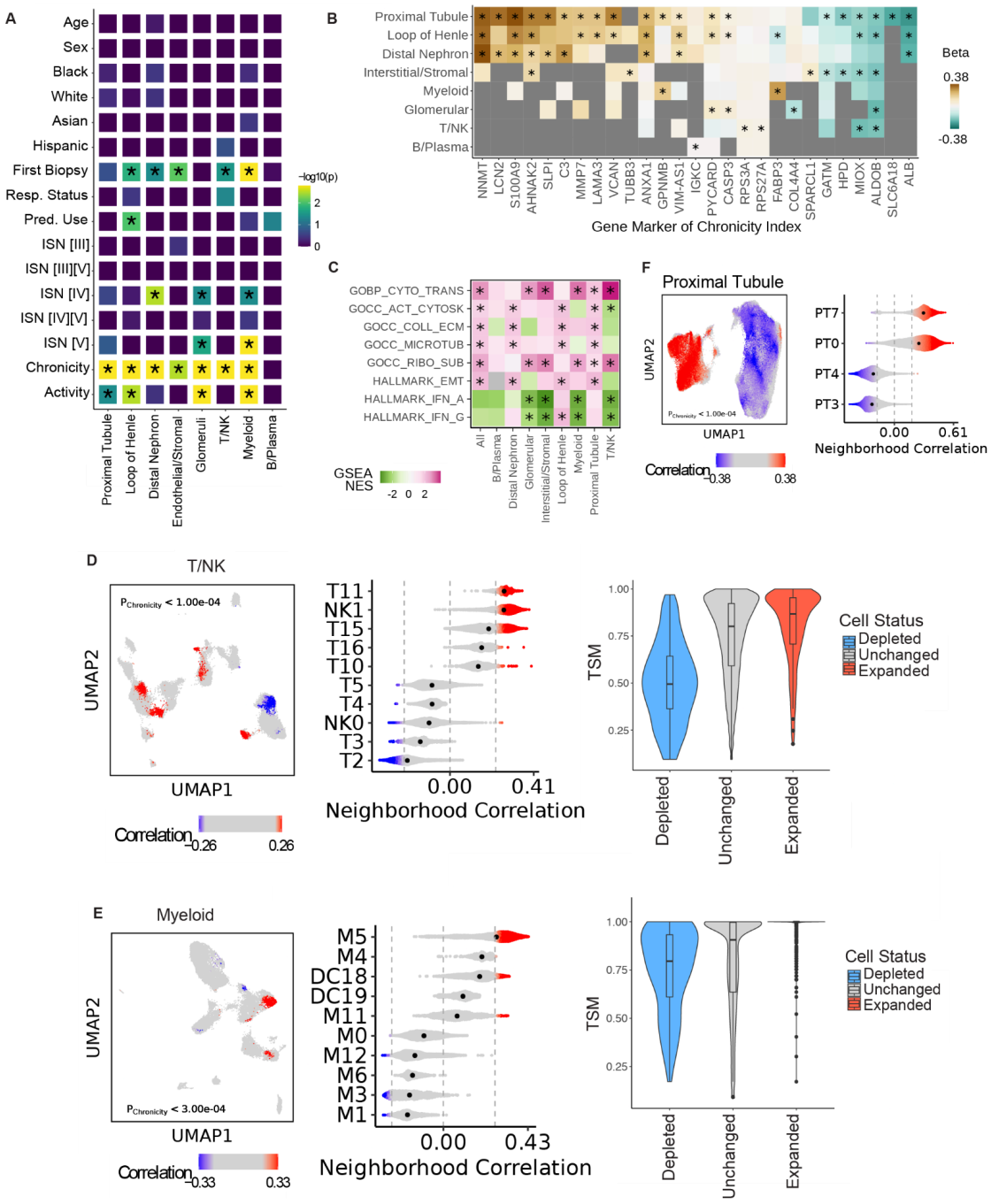

Fig. 5. Cell states associated with chronicity index. (A), Univariate CNA of each cell type testing for relevant clinical and demographic variables. * indicates a global p-value < 0.05. (**B),** For each cell type, differential expression for genes associated with chronicity index. * indicates a Bonferroni-adjusted 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 0.1. 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 the top 5 and bottom 5 clusters by median cell state correlation containing cells passing FDR significance for chronicity index association. Dashed vertical lines represent the correlation threshold with FDR < 0.1. Right - Within T/NK 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 Myeloid cells. **(F),** Same as **(D)** for PT cells, without TSM score derivations.

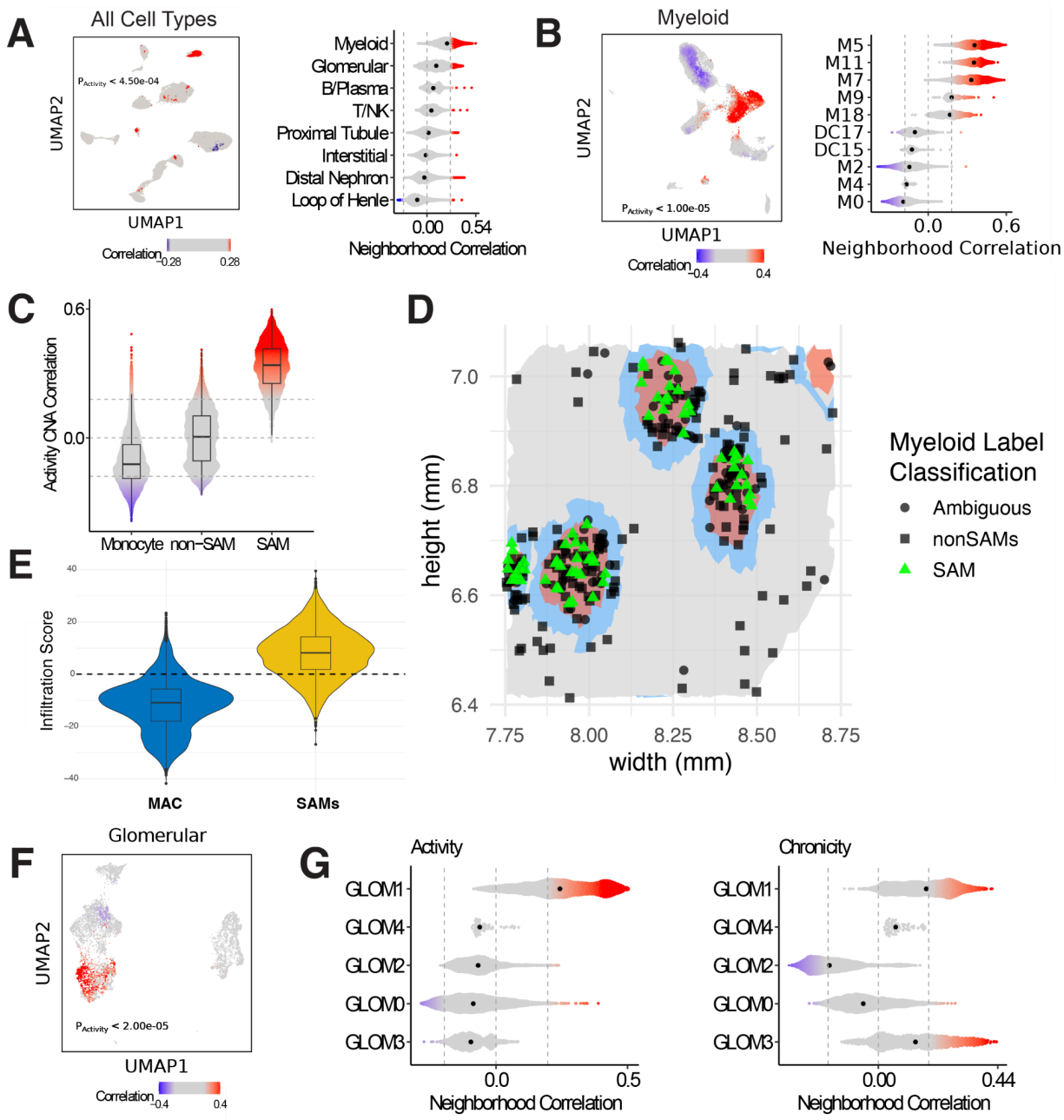

Fig. 6. Cell states associated with activity index. (A), Across all cell types, CNA results for activity index association while adjusting for first biopsy, site, and chronicity index, in scRNA-seq data. Left - UMAP displaying significant per-cell associations with activity index, with FDR cutoff of 0.1. Dashed vertical lines represent the correlation threshold with FDR < 0.1. 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)** for myeloid cells. Left - Same as **(A)** for myeloid cells. Right - Within myeloid cells, violin plots of clusters containing cells passing FDR significance for activity index association (adjusting for first biopsy, site, and chronicity). **(C),** Violin plot of activity CAN correlation grouped by broad myeloid category. **(D),** Select case FOV spatial plot of non-myeloid cells colored by spatial niche. Gray indicates tubulointerstitium, blue indicates the periglomerular region and red indicates the glomerular region. Myeloid cells distinguished by shape and color. **(E)** Violin plot of infiltration score applied on scRNAseq and snRNAseq myeloid cells. **(F)** Same as **(A)** Left for glomerular cells. **(G),** Violin plots for CNA correlations in glomerular cell types for activity (left) and chronicity (right)

Supplementary Materials

Supplementary Methods/Materials

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 (*114*). 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 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 Collection and Disassociation

Case samples were collected via needle biopsy at the time of their initial visit as part of the AMP program. Control samples were taken from living donors following kidney removal and perfusion. Both case and control kidney biopsies were processed and cryopreserved in batches of 3 following a previously described protocol(*23*) 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(*115*) method for compatibility with 10x libraries to reduce mitochondrial content. We selected a panel of 61 guide RNA sequences (**Table S15**) to tile regions of the mitochondrial genome heavily represented in 3' 10x libraries from kidney samples using on-target and off-target predictions from the Broad Institute Genetic Perturbation Platform algorithm(*116*, *117*). We used two non-targeted mitochondrial genes: MT-ND5 and MT-ND6 to estimate the mitochondrial content per cell. We find that MT-ND5 and MT-ND6 correlate to total mitochondrial read content in our pilot data (R=0.982 and R=0.861, respectively) **(Fig. S38A, B).** The summation MT-ND5 + MT-ND6 also correlates to overall mitochondrial read content **(Fig S38C).** A 1% increase in MT-ND5 or MT-ND6 corresponded with a 12.7% and 13.7% increase in total mitochondrial content. A 1% increase in the summation MT-ND5 + MT-ND6 corresponds to a 7.3% increase. We removed cells with more than 3% read contributions of 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 of total mitochondrial reads. Sensitivity analysis of our mitochondrial threshold in scRNAseq suggests that a 1% change in MT-ND5 and MT-ND6 resulted in a 10% change in cell count **(Fig. S38D).**

T7 reverse transcriptase template was generated from oligonucleotides (IDT) as previously described(*115*). 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(*118*), 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(*119*), 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.

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(*120*) 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.

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. 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 (**Fig. S9B**).

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 (*121*) and initially characterized in Korsunsky et. al (*28*) to balance the contribution from single cell and single nuclear cells when computing principal components. We then applied harmony 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(*122*). 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(*123*) 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(*36*), the AMP Phase II Rheumatoid Arthritis publication(*124*), and the AMP Phase I LN publication. 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 endothelial/stromal compartment, and *CFH* and *PODXL* from cells from the glomerulus.

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:

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

Groups with a median score of above 0.7 were assigned as tissue-specific, while a median score of below 0.3 was designated as blood-specific. All other populations are demarcated as being shared.

Pseudotime Analysis

Pseudotime lineages were computed by embedding PBMC and tissue cells into a joint embedding using Destiny (*41*) on harmonized Principal Components (PCs). Trajectories were subsequently computed using Slingshot (*42*). Start points were chosen from PBMC clusters that were categorized as blood-specific by TSM.

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(*23*) and the commonly used method for single-cell scoring with a gene set used in Seurat's AddModuleScore and Scanpy's score_genes functions (*125*, *126*). Briefly, within each cell type for scRNAseq 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, endothelial/stromal) cell types.

Receptor-Ligand Analysis

Cell–cell communication analysis was performed using the CellChat R package(*127*). We performed ligand receptor analysis subsetting only to immune and glomerular cells from scRNAseq experiments. We use the full human ligand-receptor interaction database provided to CellChat. Overexpressed genes and ligand–receptor interactions were identified using CellChat's native identifyOverExpressedGenes and identifyOverExpressedInteractions functions. Communication probabilities between cell groups were inferred using the computeCommunProb function with the tri-mean method. Inferred interactions were aggregated into signaling networks using aggregateNet.

SAM Spatial Logistic Regression Classifier

We build our logistic regression classifier using the glmnet package in R. We designate M5, M7, M9, and M11 as SAMs and all other myeloid cells as non-SAMs. Training data was selected from a random 90% of samples with the remaining 10% withheld for testing. Featurized genes were selected from the overlap between the AMP and Danaher et al. dataset and further subsetted to include genes that were expressed in at least 25% of genes in spatial transcriptomic cells. Cross fold validation was performed across all the potential folds and the first fold was arbitrarily selected. To optimize regularization, we performed additional cross-validation using cv.glmnet to optimize lambda. The trained model was applied to all myeloid cells, and cells were classified as "SAMs" if predicted probability >0.8, "non-SAMs" if <0.2, and "ambiguous" if probabilities fell between these thresholds.

Supplementary Figures

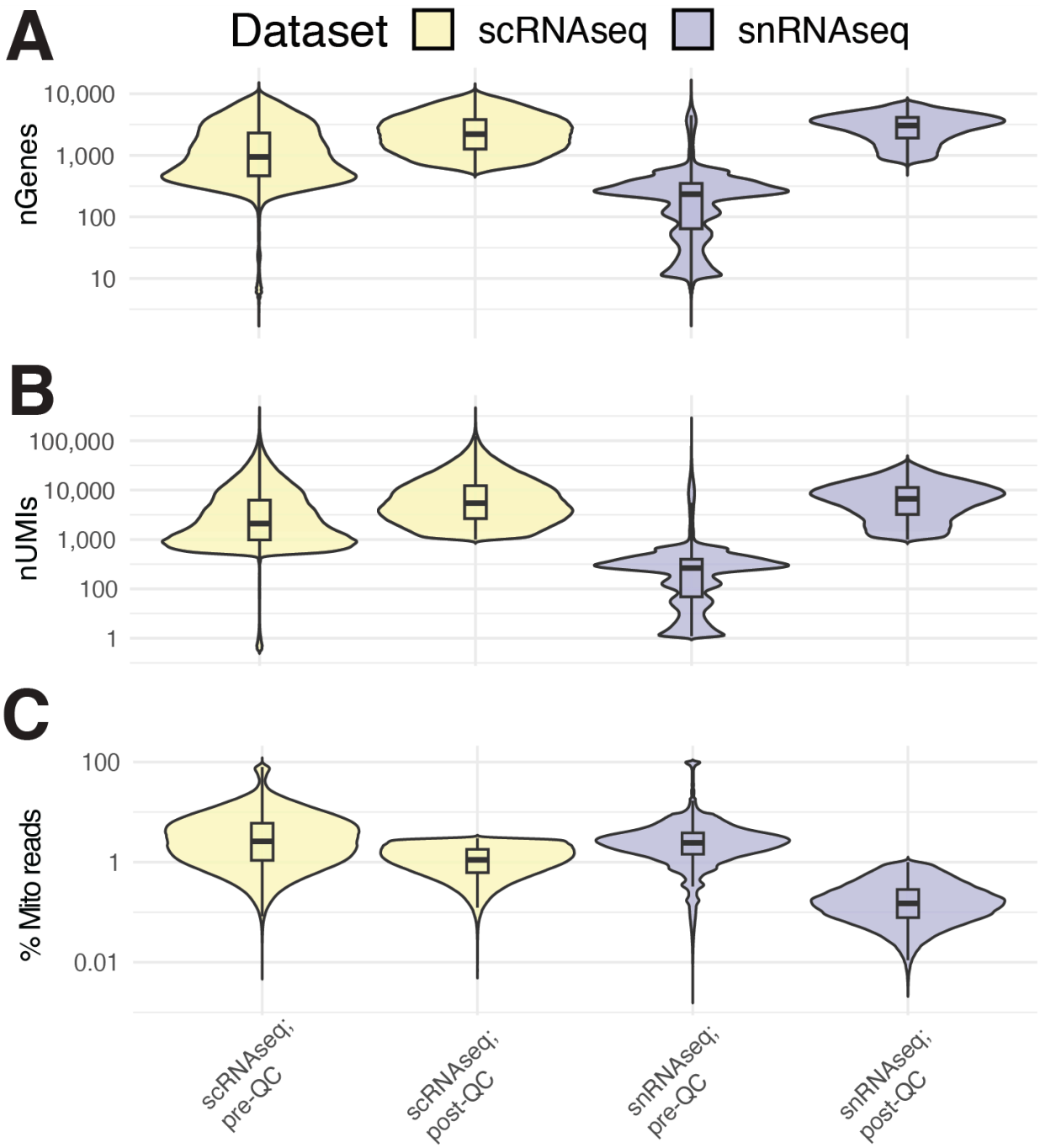

Fig. S1. Histogram of scRNAseq and snRNAseq prior to quality control. Histograms of uniquely expressed genes (top), UMIs (center), and % mitochondrial reads (bottom) stratified for scRNAseq (A) and snRNAseq (B) modalities. Vertical lines correspond to post quality control cutoffs. Lines are drawn at > 500 genes, > 1,000 UMIs, and >3% MT reads in scRNAseq and between 500 and 7500 uniquely expressed genes, between 1,000 and 40,000 UMIs, and less than 1% mitochondrial reads in snRNAseq.

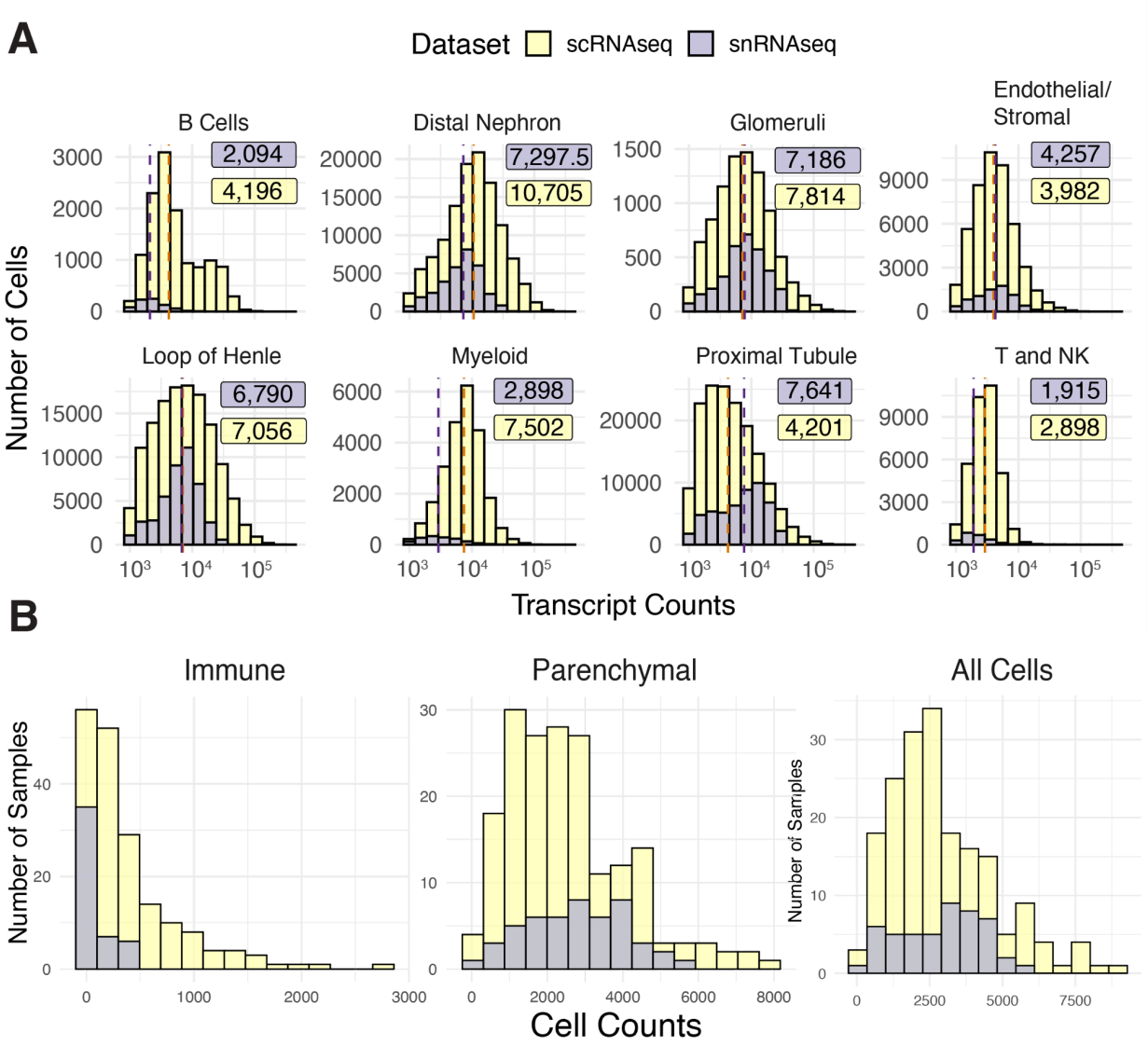

Fig. S2. Transcript Count Histograms. (A) Histogram of transcript counts per modality, faceted by cell type. Vertical lines are drawn at each modality's median in scRNAseq (deep orange) and snRNAseq (deep purple). Medians for scRNAseq (yellow box) and snRNAseq (purple box) are included for each cell type. (B) Histogram of cell counts per sample in each modality, faceted by immune cells (left), parenchymal cells (center), or both immune and parenchymal cells (right).

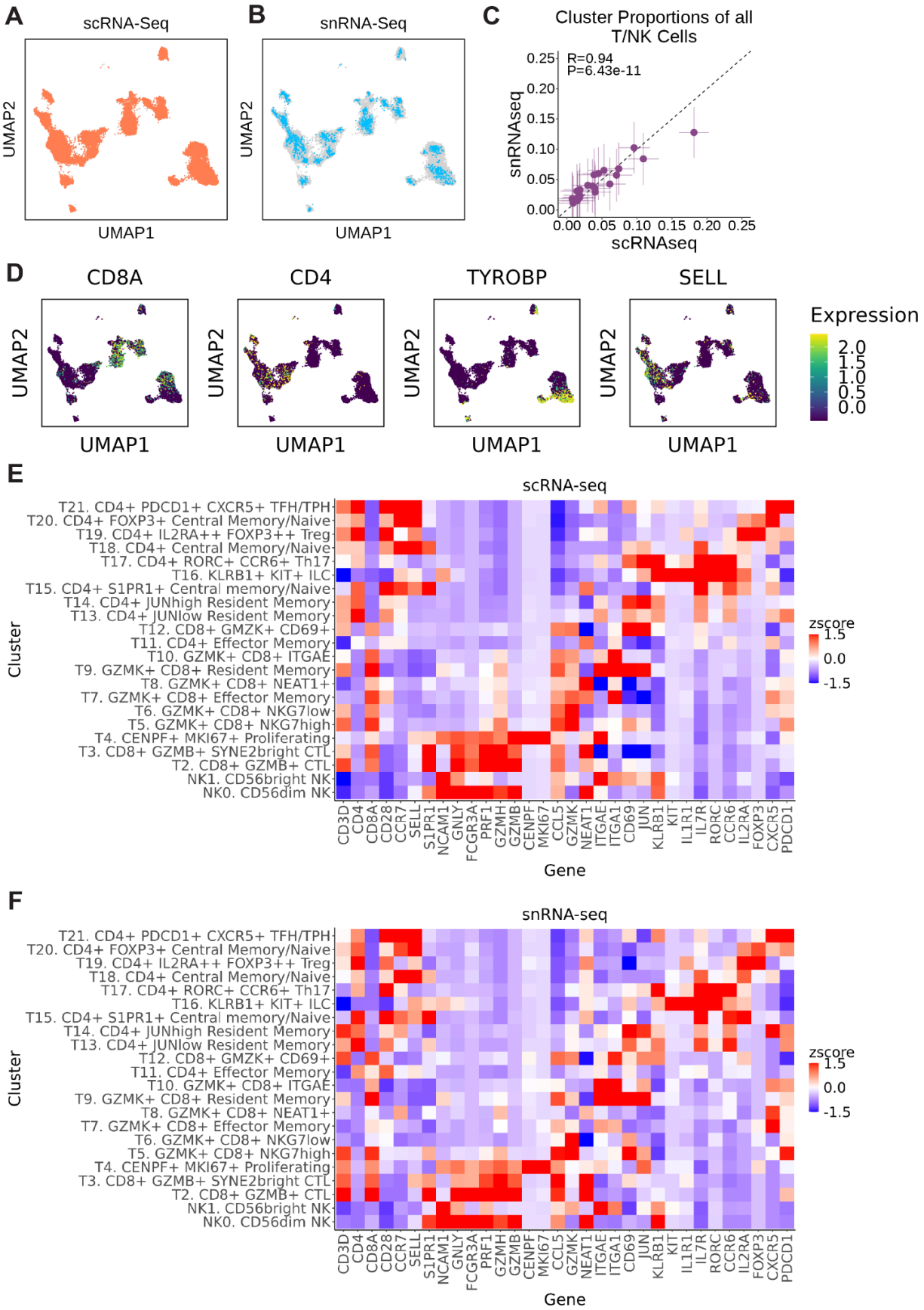
A
scRNA-Seq
UMAP2
UMAP1
B
snRNA-Seq
UMAP2
UMAP1
C
Cluster Proportions of all T/NK Cells
R=0.94
P=6.43e-11
snRNAseq
scRNAseq
0.00
0.05
0.10
0.15
0.20
0.25
D
CD8A
CD4
TYROBP
SELL
UMAP2
UMAP1
Expression
2.0
1.5
1.0
0.5
0.0
E
scRNA-seq
Cluster
T21. CD4+ PDCD1+ CXCR5+ TFH/TPH
T20. CD4+ FOXP3+ Central Memory/Naive
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

Fig. S3. 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)**.

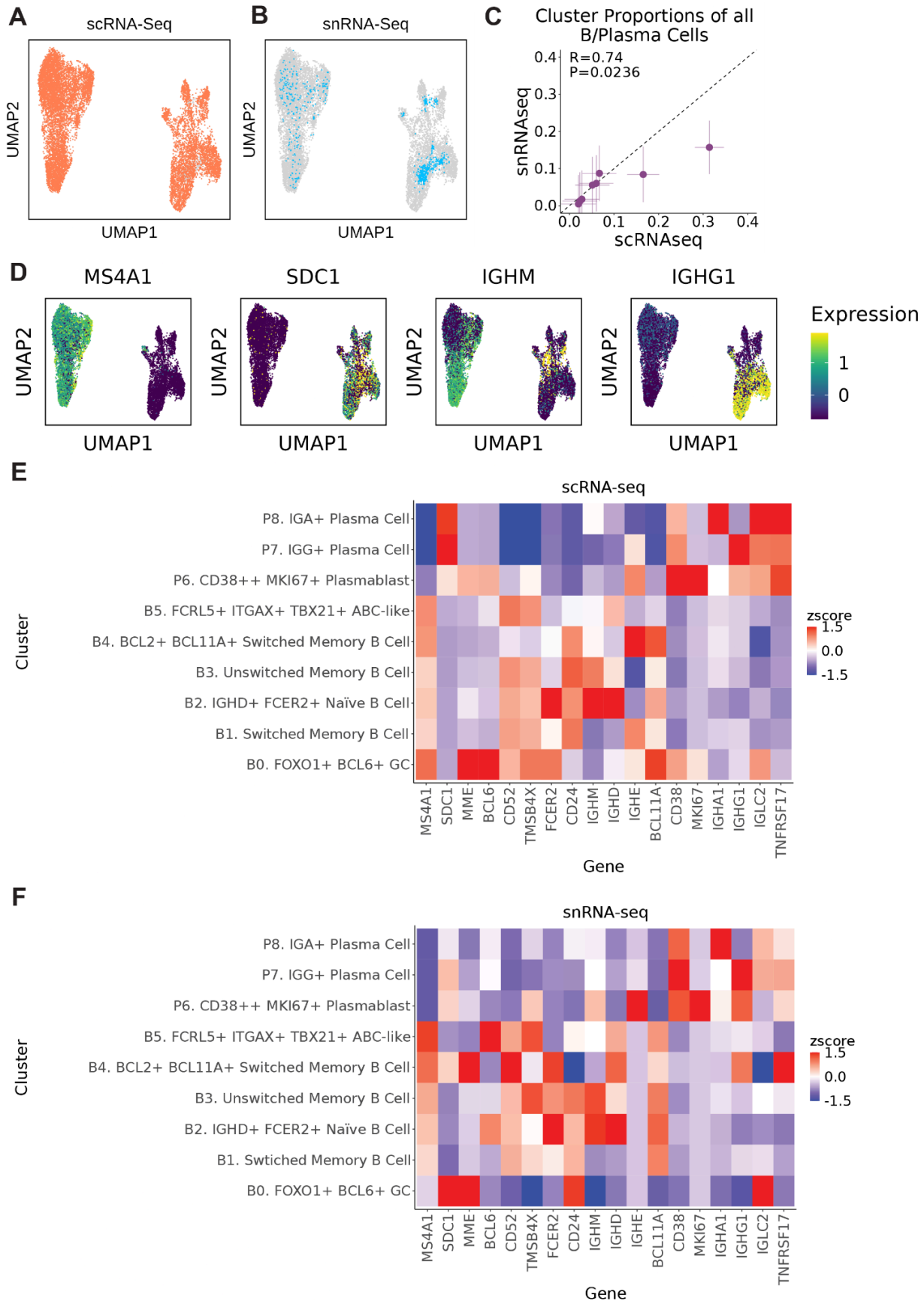
A
scRNA-Seq
UMAP2
UMAP1
B
snRNA-Seq
UMAP2
UMAP1
C
Cluster Proportions of all B/Plasma Cells
R=0.74
P=0.0236
snRNAseq
scRNAseq
0.0
0.1
0.2
0.3
0.4
D
MS4A1
SDC1
IGHM
IGHG1
UMAP2
UMAP1
Expression
1
0
E
scRNA-seq
Cluster
P8. IGA+ Plasma Cell
P7. IGG+ Plasma Cell
P6. CD38++ MKI67+ Plasmablast
B5. FCRL5+ ITGAX+ TBX21+ ABC-like
B4. BCL2+ BCL11A+ Switched Memory B Cell
B3. Unswitched Memory B Cell
B2. IGHD+ FCER2+ Naïve B Cell
B1. Switched Memory 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+ Switched Memory B Cell
B3. Unswitched Memory B Cell
B2. IGHD+ FCER2+ Naïve B Cell
B1. Swtiched Memory 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

Fig. S4. 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)**.

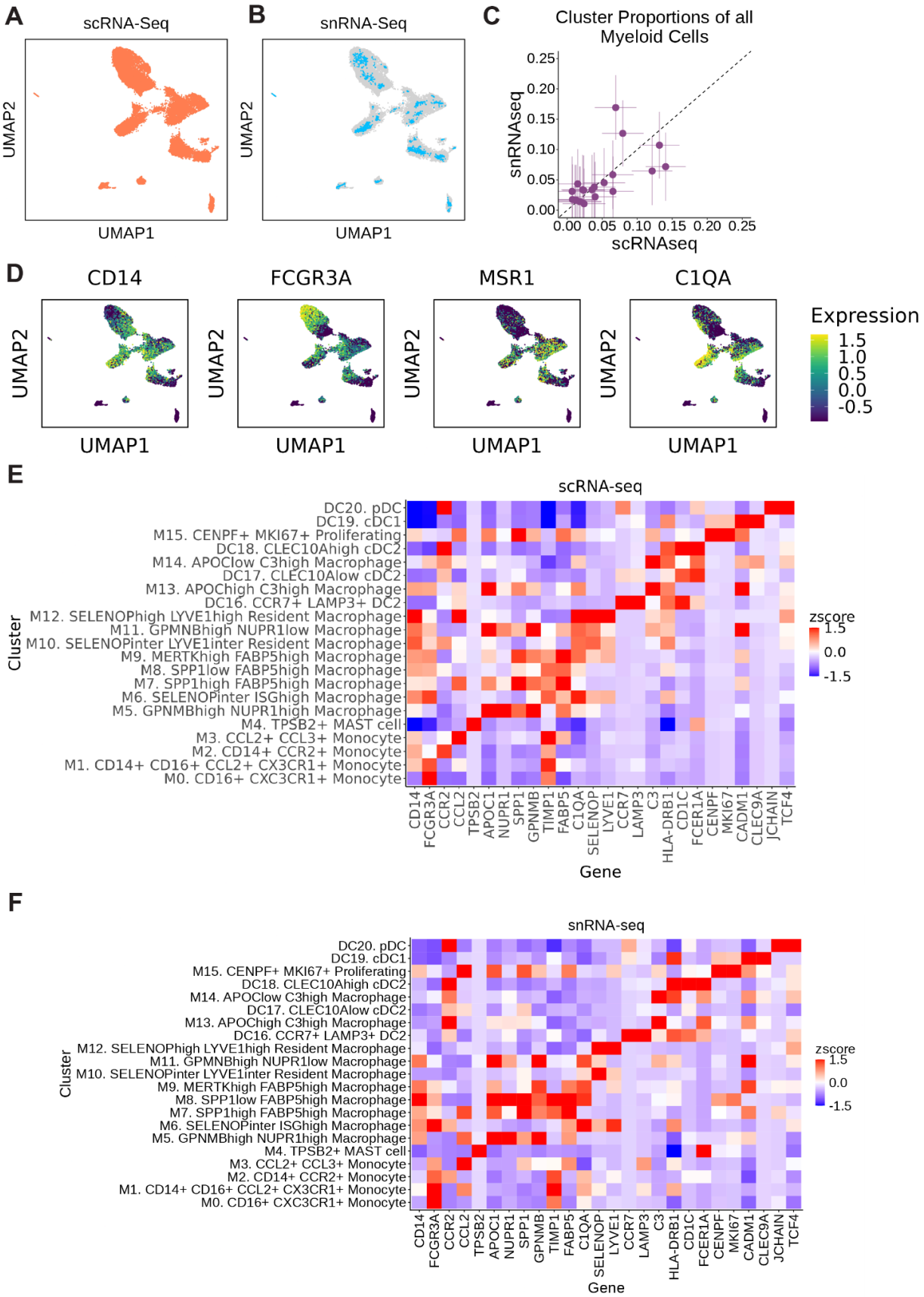
A
scRNA-Seq
UMAP2
UMAP1
B
snRNA-Seq
UMAP2
UMAP1
C
Cluster Proportions of all Myeloid Cells
snRNAseq
scRNAseq
0.00
0.05
0.10
0.15
0.20
0.25
D
CD14
FCGR3A
MSR1
C1QA
UMAP2
UMAP1
Expression
1.5
1.0
0.5
0.0
-0.5
E
scRNA-seq
Cluster
DC20. pDC
DC19. cDC1
M15. CENPF+ MKI67+ Proliferating
DC18. CLEC10Ahigh cDC2
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
snRNA-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

Fig. S5. 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)**.

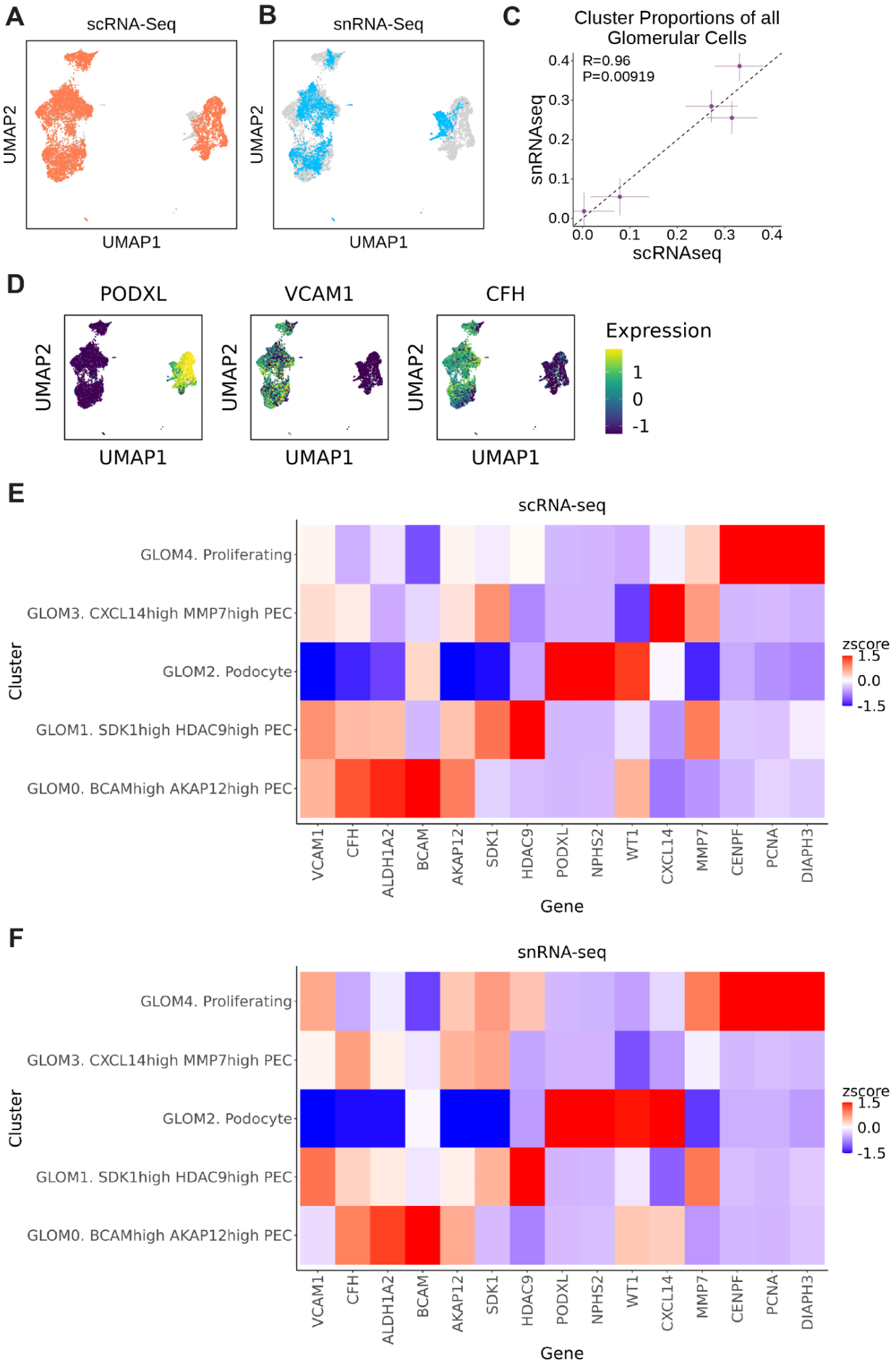
A
scRNA-Seq
UMAP2
UMAP1
B
snRNA-Seq
UMAP2
UMAP1
C
Cluster Proportions of all Glomerular Cells
R=0.96
P=0.00919
snRNAseq
scRNAseq
0.0
0.1
0.2
0.3
0.4
D
PODXL
VCAM1
CFH
UMAP2
UMAP1
Expression
1
0
-1
E
scRNA-seq
Cluster
GLOM4. Proliferating
GLOM3. CXCL14high MMP7high PEC
GLOM2. Podocyte
GLOM1. SDK1high 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

Fig. S6. 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)**.

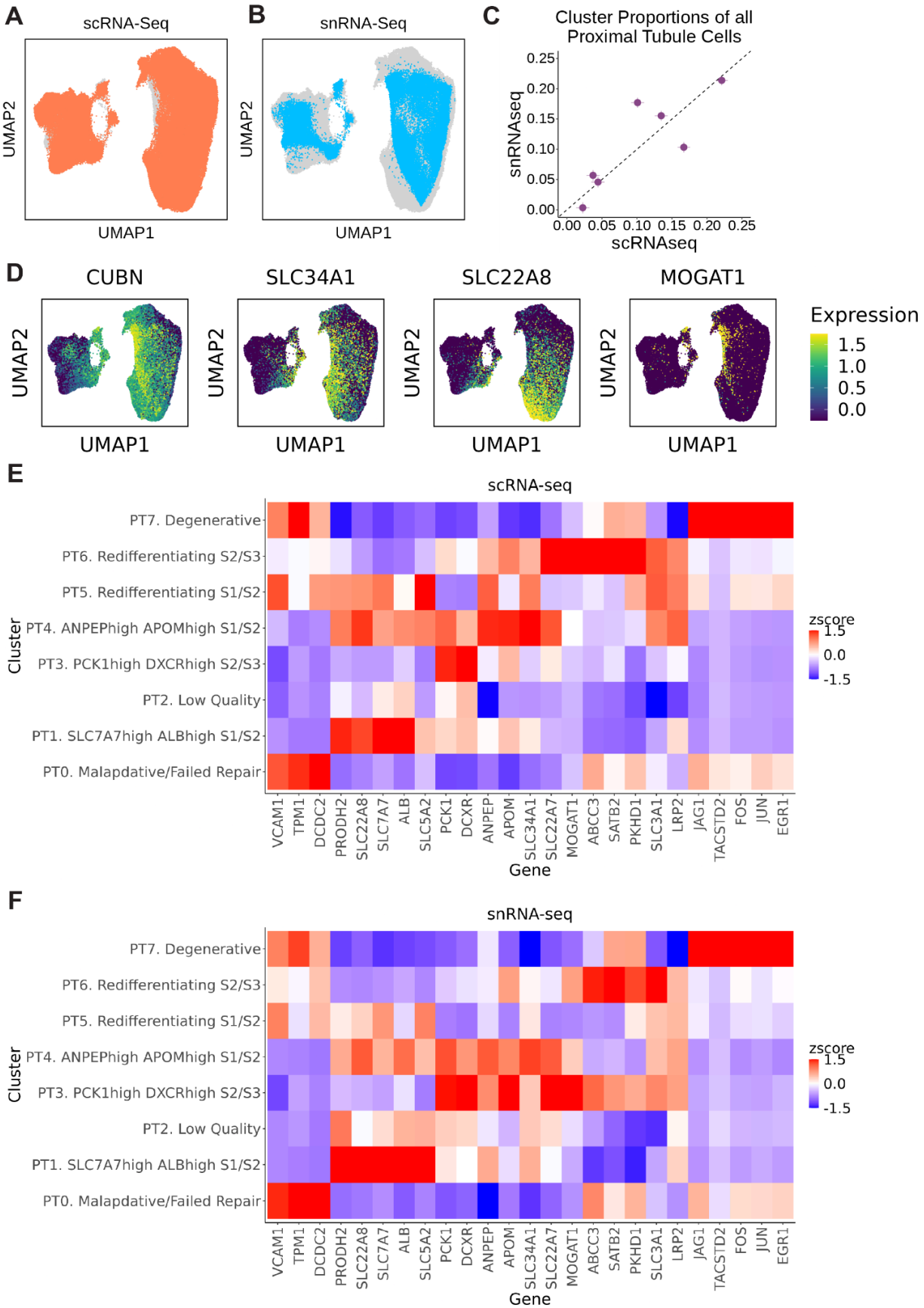
A
scRNA-Seq
UMAP2
UMAP1
B
snRNA-Seq
UMAP2
UMAP1
C
Cluster Proportions of all Proximal Tubule Cells
snRNAseq
scRNAseq
0.00
0.05
0.10
0.15
0.20
0.25
D
CUBN
SLC34A1
SLC22A8
MOGAT1
UMAP2
UMAP1
Expression
1.5
1.0
0.5
0.0
E
scRNA-seq
Cluster
PT7. Degenerative
PT6. Redifferentiating S2/S3
PT5. Redifferentiating 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

Fig. S7. 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)**.

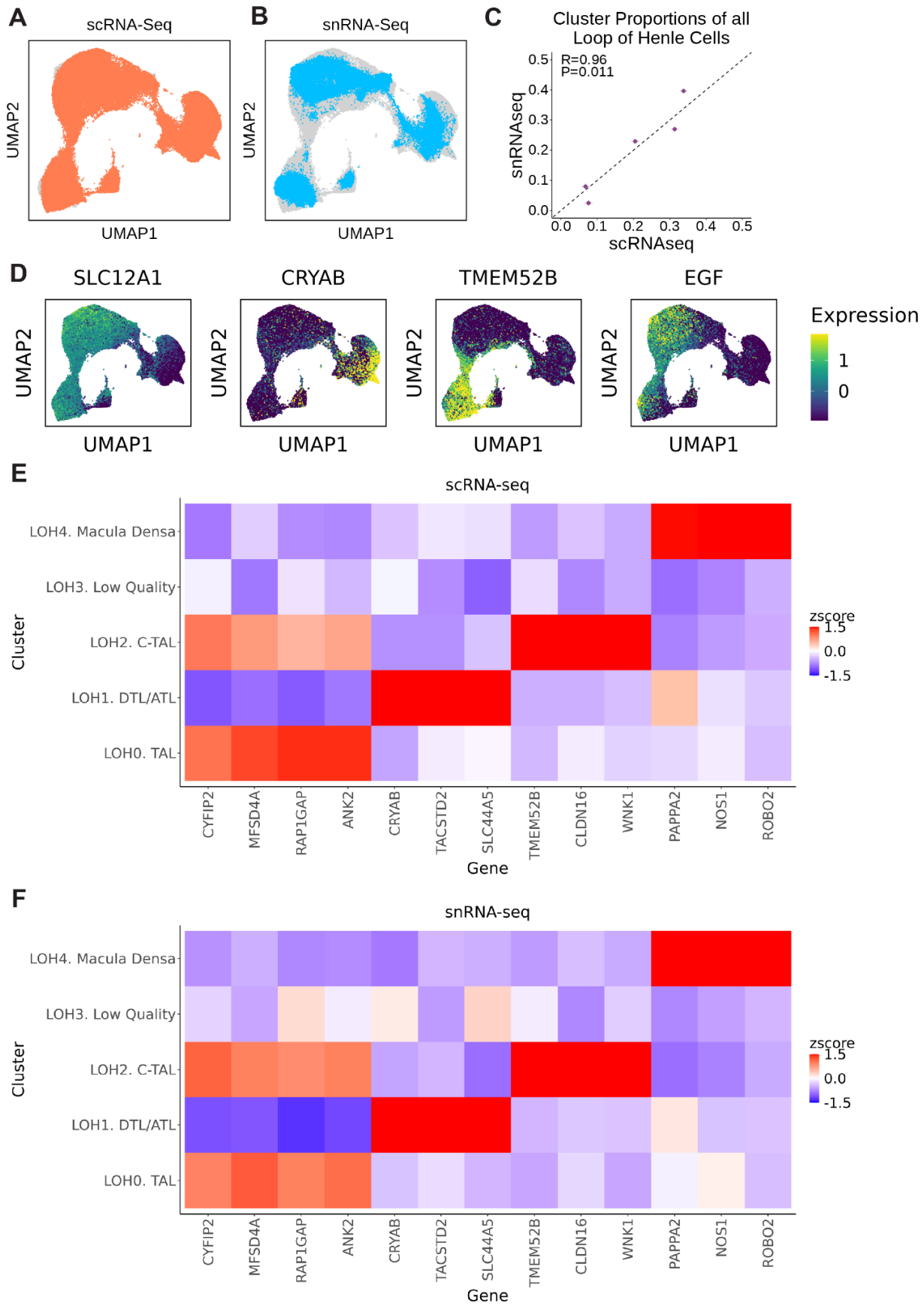
A
scRNA-Seq
UMAP2
UMAP1
B
snRNA-Seq
UMAP2
UMAP1
C
Cluster Proportions of all Loop of Henle Cells
R=0.96
P=0.011
snRNAseq
scRNAseq
0.0
0.1
0.2
0.3
0.4
0.5
D
SLC12A1
CRYAB
TMEM52B
EGF
UMAP2
UMAP1
Expression
1
0
E
scRNA-seq
LOH4. Macula Densa
LOH3. Low Quality
LOH2. C-TAL
LOH1. DTL/ATL
LOH0. TAL
Cluster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

Fig. S8. 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)**.

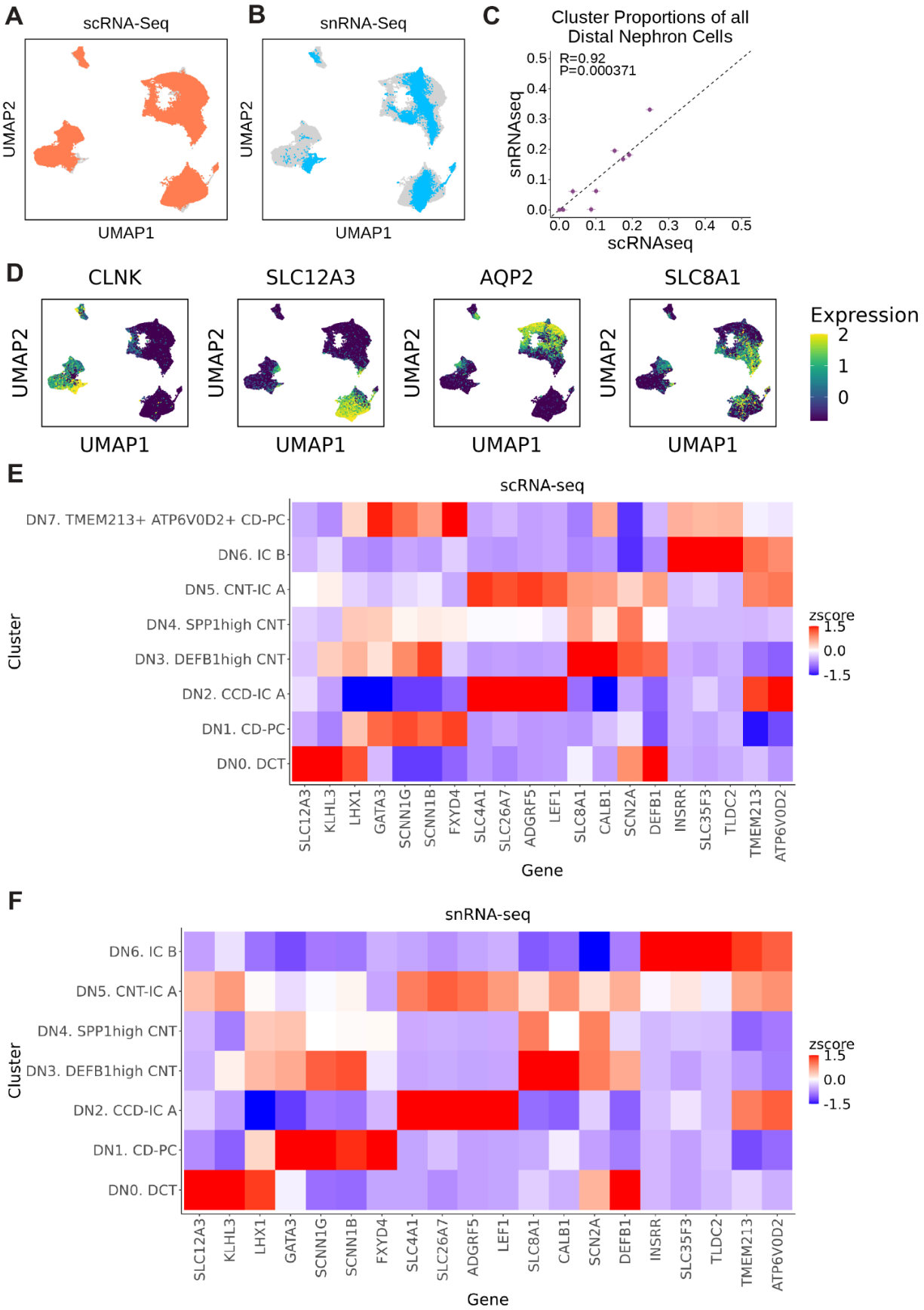
A
scRNA-Seq
UMAP2
UMAP1
B
snRNA-Seq
UMAP2
UMAP1
C
Cluster Proportions of all Distal Nephron Cells
R=0.92
P=0.000371
snRNAseq
scRNAseq
0.0
0.1
0.2
0.3
0.4
0.5
D
CLNK
SLC12A3
AQP2
SLC8A1
UMAP2
UMAP1
Expression
2
1
0
E
scRNA-seq
Cluster
DN7. TMEM213+ ATP6V0D2+ CD-PC
DN6. IC B
DN5. CNT-IC A
DN4. SPP1high 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

Fig. S9. 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)**.

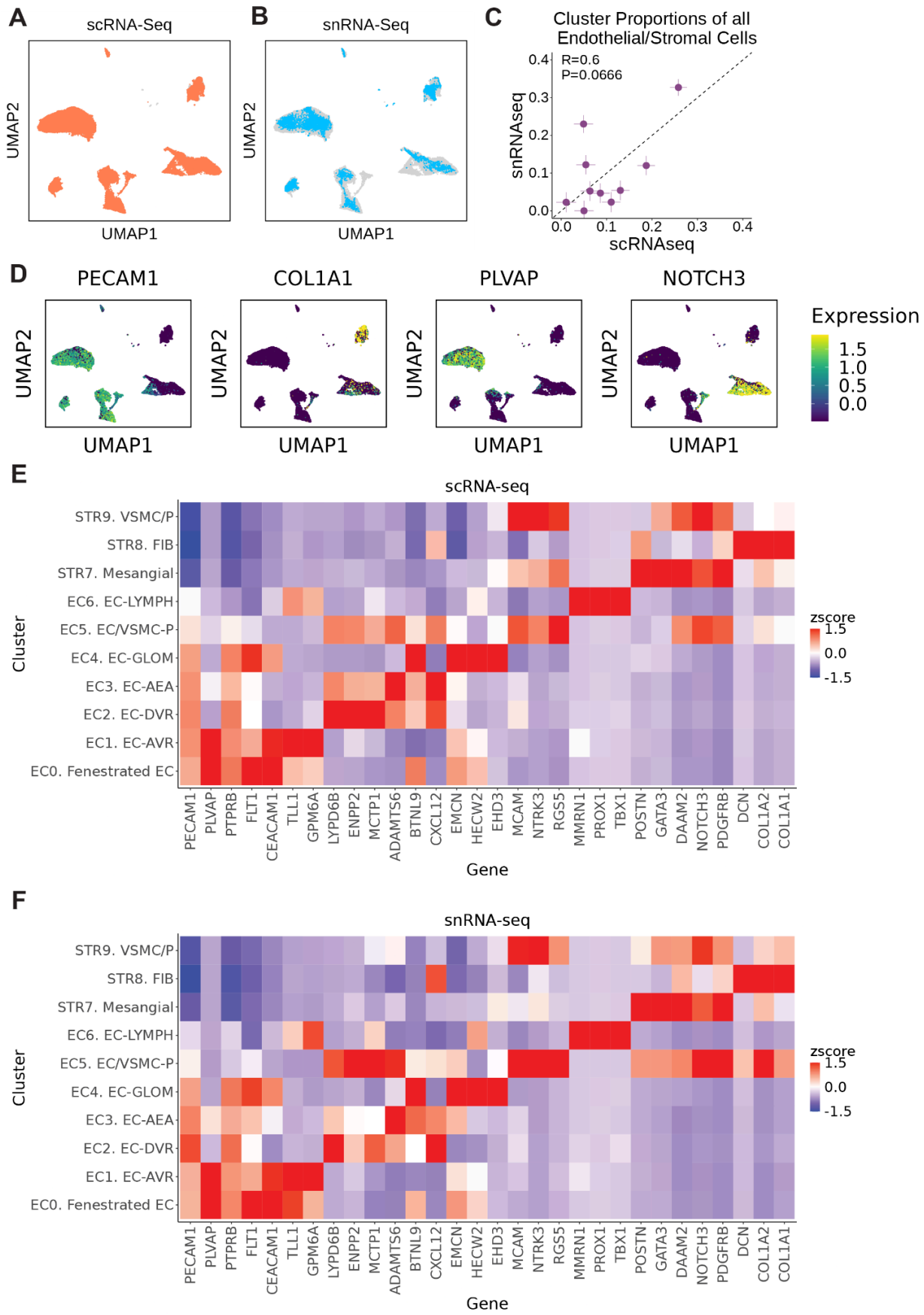
A
scRNA-Seq
UMAP2
UMAP1
B
snRNA-Seq
UMAP2
UMAP1
C
Cluster Proportions of all Endothelial/Stromal Cells
R=0.6
P=0.0666
snRNAseq
scRNAseq
0.0
0.1
0.2
0.3
0.4
D
PECAM1
COL1A1
PLVAP
NOTCH3
UMAP2
UMAP1
Expression
1.5
1.0
0.5
0.0
E
scRNA-seq
Cluster
STR9. VSMC/P
STR8. FIB
STR7. Mesangial
EC6. EC-LYMPH
EC5. 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

Fig. S10. Endothelial/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 endothelial/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 endothelial/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)**.

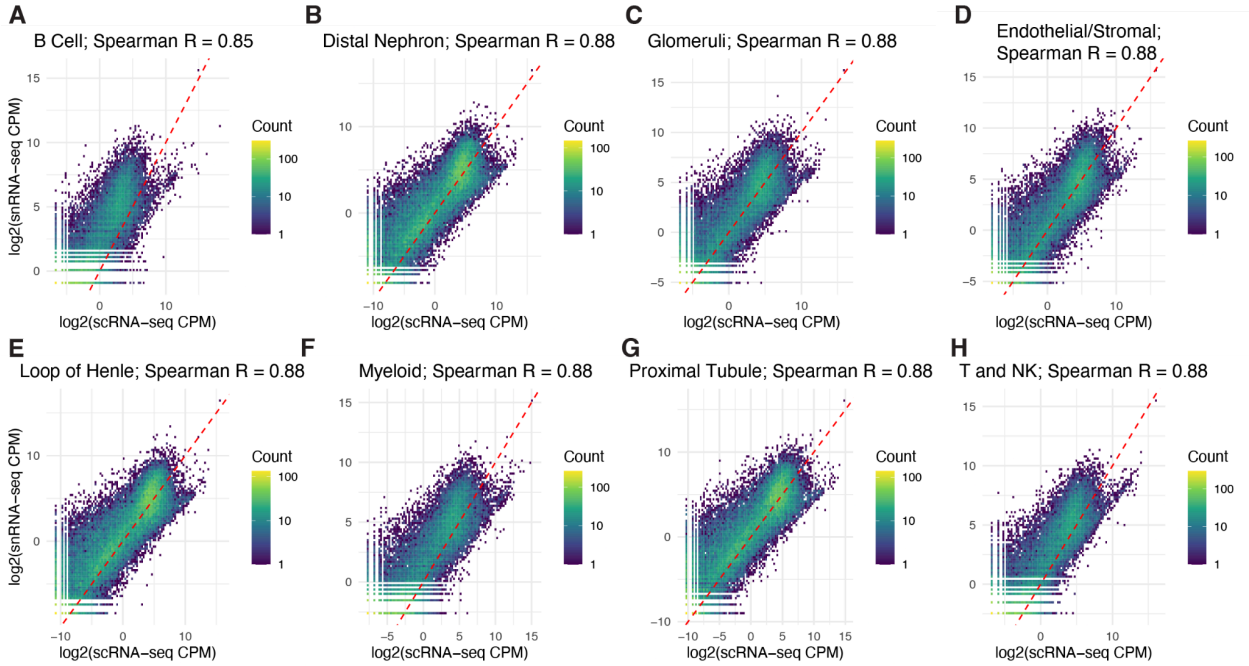

Fig. S11. Global transcript correlations within cell type. Density plots of log2 counts per million per gene stratified by B cells **(A),** distal nephron **(B),** glomeruli **(C),** endothelial/stromal cells **(D),** loop of Henle **(E),** myeloid **(F),** proximal tubule **(G),** and T and NK cells **(H).** Spearman correlations are included.

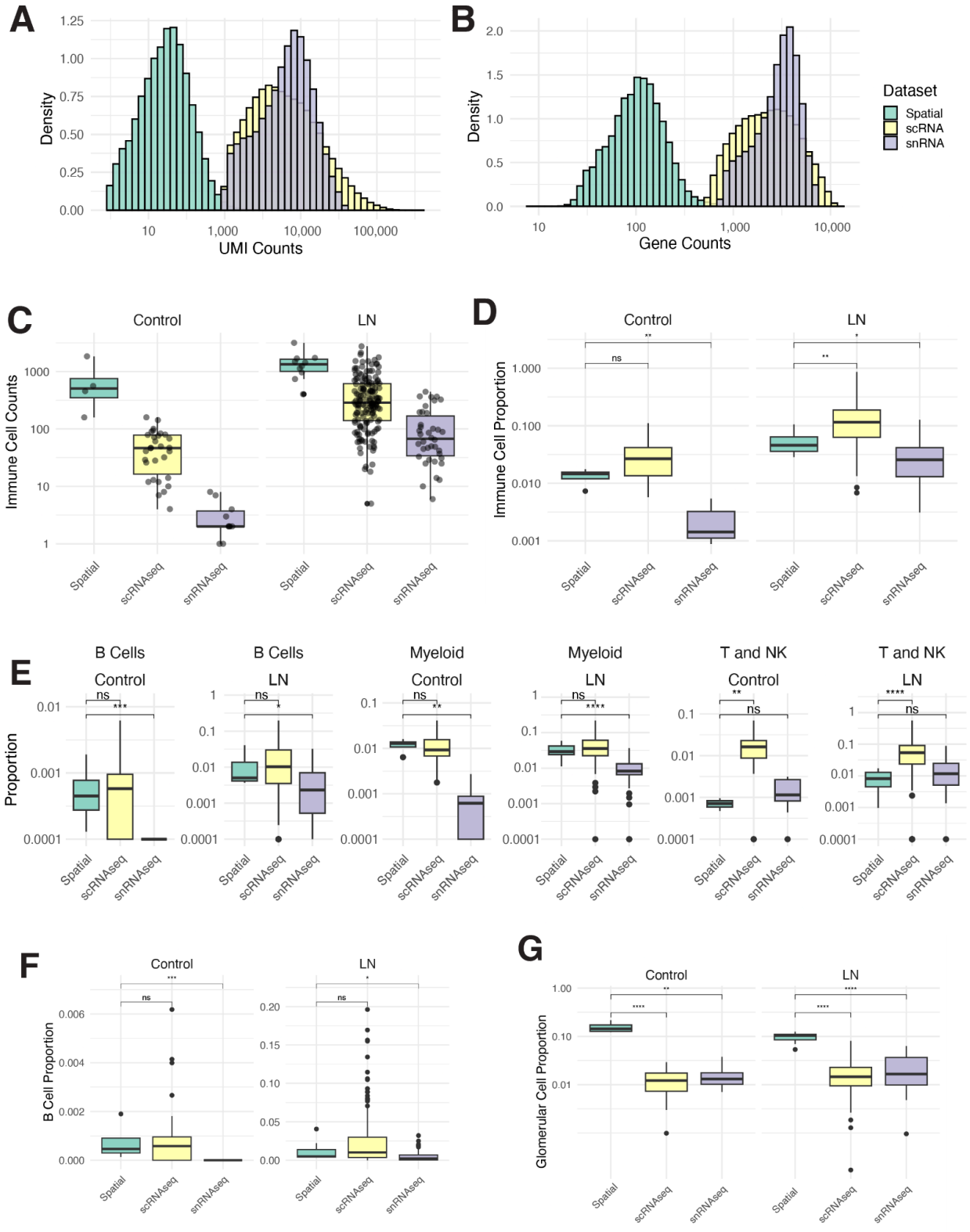

Fig. S12. Sparsity of Immune Cells and Glomerular Cells. (A) Histogram of total gene counts per cell,

stratified by modality. Histograms are normalized such that total area per group is equal. **(B)** Same as **(A)**, but for total genes expressed. **(C)** Total immune cells collected per biopsy, stratified by collection technology. **(D)** Proportion of immune cells per biopsy, stratified by collection technology. **(E)** Proportion of immune cells per biopsy, faceted by B cells (left), myeloid cells (middle) and T and NK cells (right). **(F)** Proportion of B cells of all cells per collection method (G) same as (F), but for glomerular cells as a proportion of all collected cells per sample.

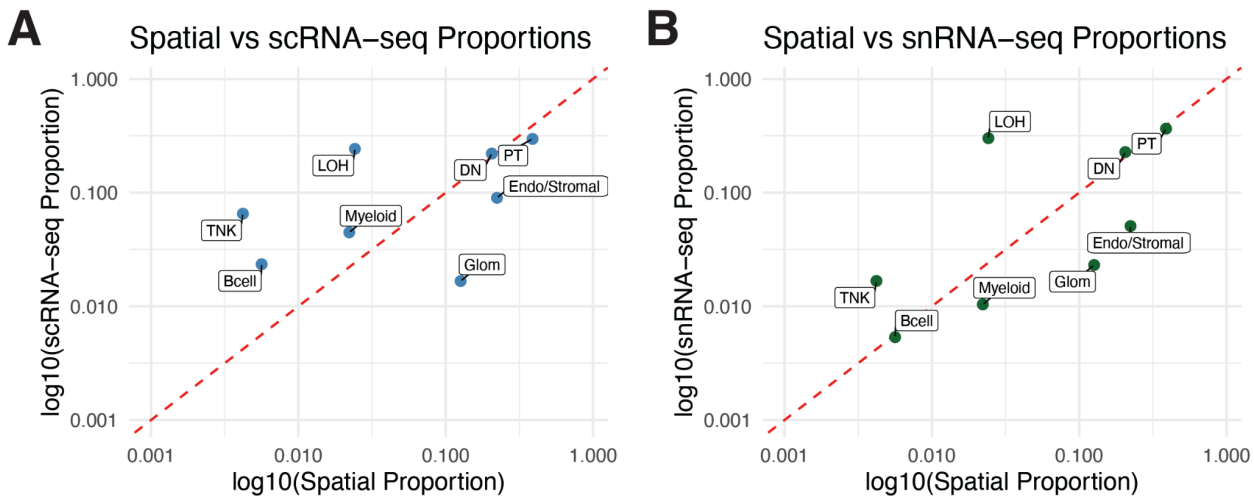

Fig. S13. Cell Type Proportions between spatial and single-cell modalities. **(A)** Proportion of broad cell types in Spatial data versus scRNAseq data. **(B)** Proportion of broad cell types in Spatial data versus snRNAseq data.

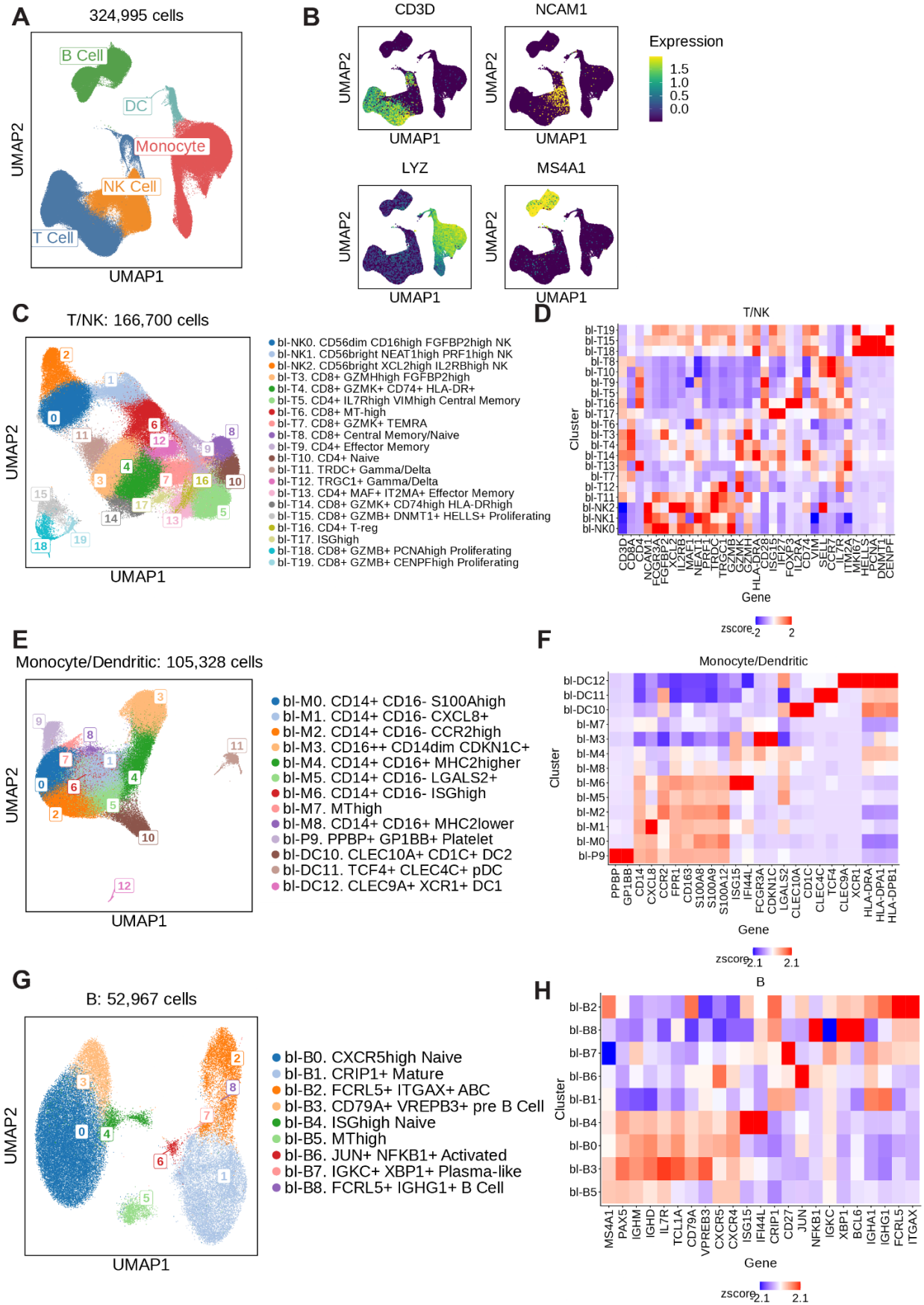
A
324,995 cells
B Cell
DC
Monocyte
NK Cell
T Cell
UMAP2
UMAP1
B
CD3D
NCAM1
LYZ
MS4A1
Expression
1.5
1.0
0.5
0.0
C
T/NK: 166,700 cells
bl-NK0. CD56dim CD16high FGFBP2high NK
bl-NK1. CD56bright NEAT1high PRF1high NK
bl-NK2. CD56bright XCL2high IL2RBhigh NK
bl-T3. CD8+ GZMHhigh FGFBP2high
bl-T4. CD8+ GZMK+ CD74+ HLA-DR+
bl-T5. CD4+ IL7Rhigh VIMhigh Central Memory
bl-T6. CD8+ MT-high
bl-T7. CD8+ GZMK+ TEMRA
bl-T8. CD8+ Central Memory/Naive
bl-T9. CD4+ Effector Memory
bl-T10. CD4+ 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
D
T/NK
Cluster
Gene
zscore
-2
2
E
Monocyte/Dendritic: 105,328 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-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
Monocyte/Dendritic
Cluster
Gene
zscore
-2.1
2.1
G
B: 52,967 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
-2.1
2.1

Fig. S14. PBMC dataset overview. (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.

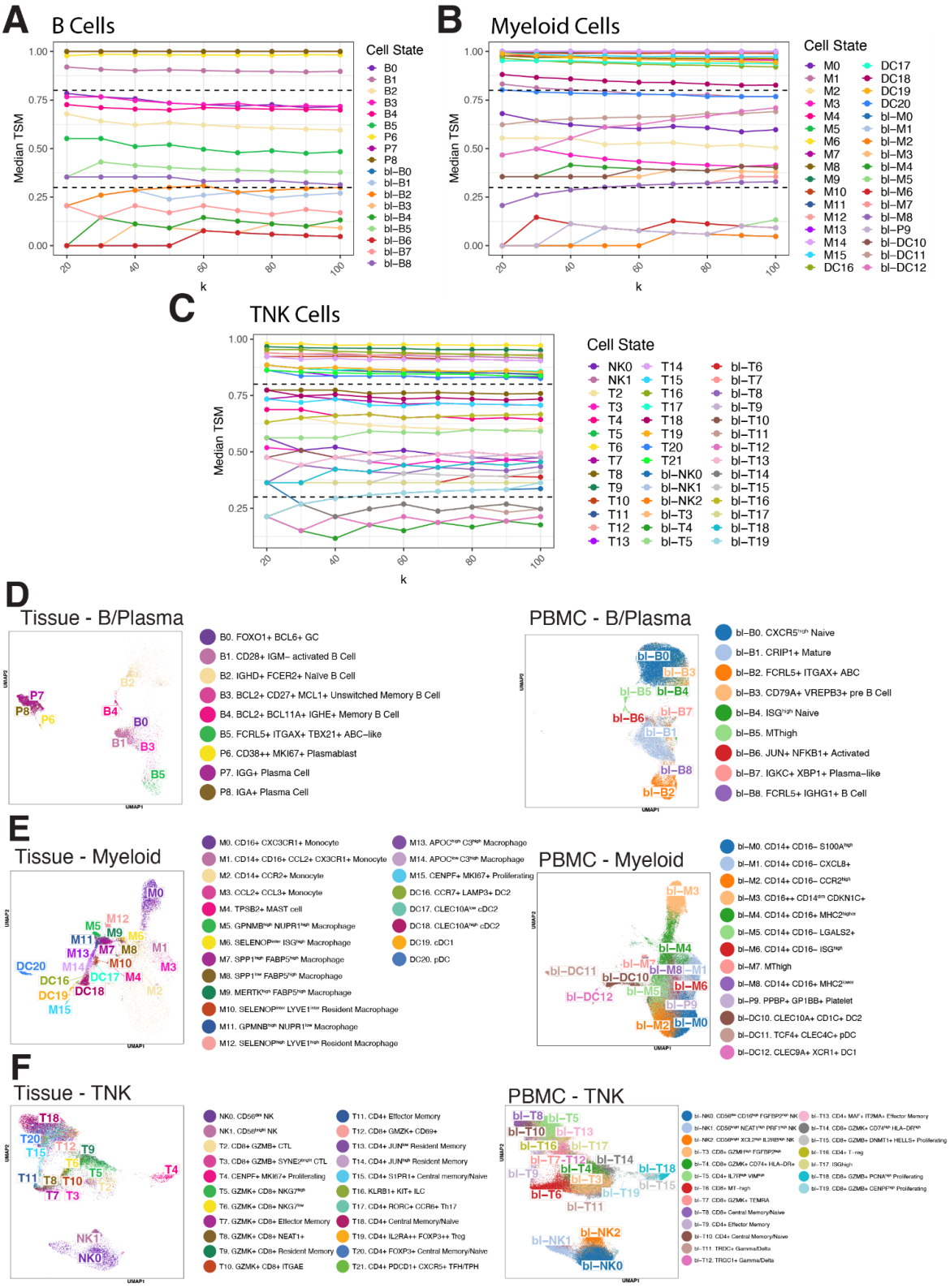

Fig. S15. Robustness of TSM. Median TSM scores per cell state for B **(A),** Myeloid **(B),** and TNK **(C)**

cell types. Dashed lines denote tissue-specific (> 0.8) and blood-specific (< 0.3) cutoffs. **(D-F)** Joint embedding of tissue (left) and blood (right) immune cells colored by the original clustering for B and plasma cells **(D),** myeloid cells **(E),** and T and NK cells **(F).**

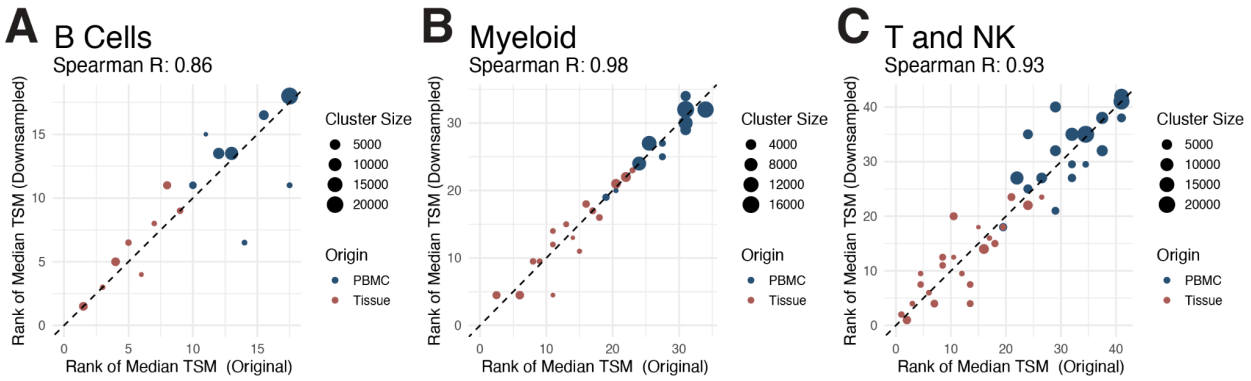

Fig. S16. Median cluster TSM rankings before and after downsampling. Correlational plots of per-cluster median TSM rankings for B cells **(A),** Myeloid cells **(B),** and T and NK cells **(C).** Dot sizes are proportional to relative cluster size within that cell type and colored by origin.

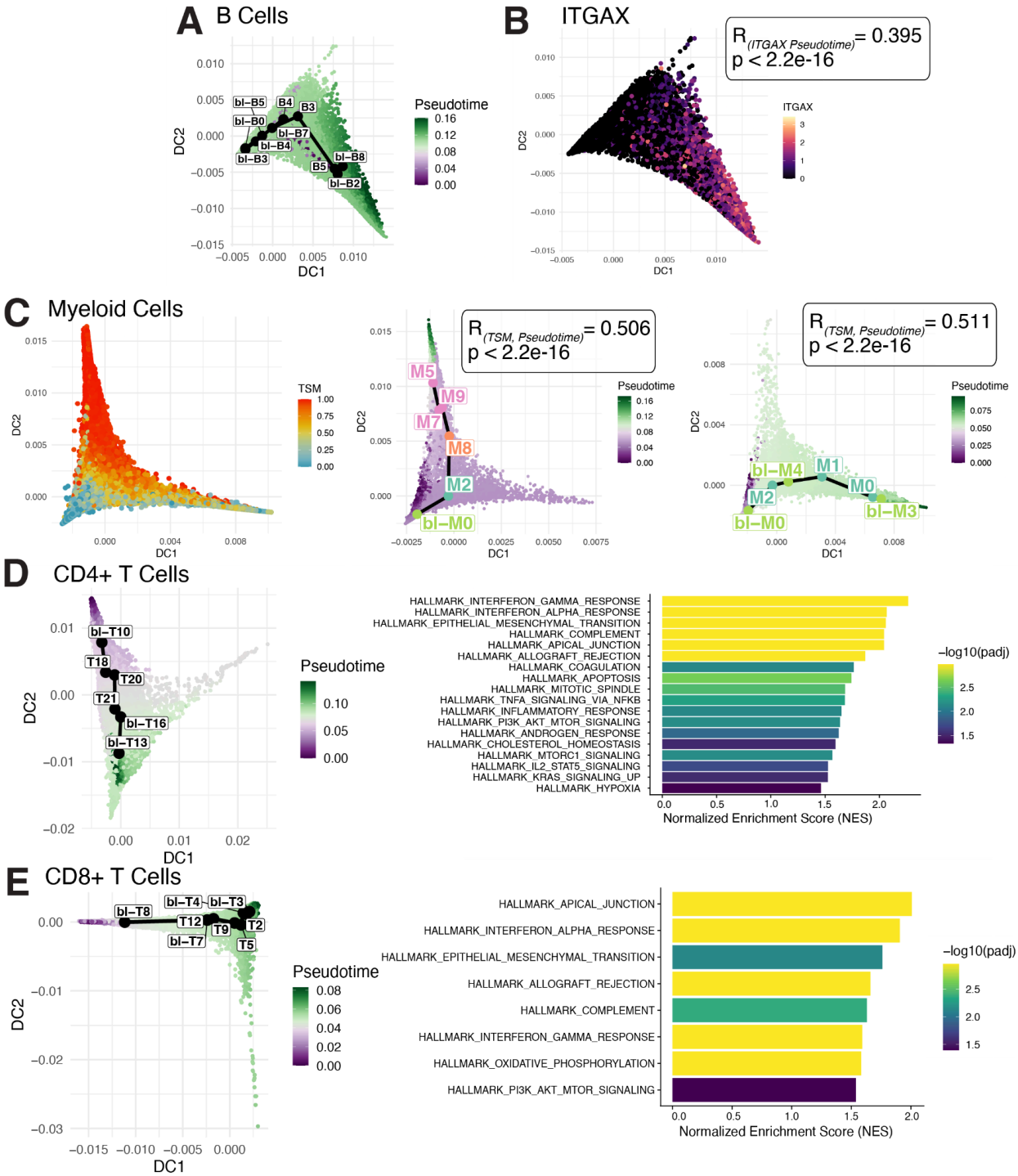
A
B Cells
Pseudotime
DC1
DC2
B
ITGAX
R (ITGAX Pseudotime) = 0.395
p < 2.2e-16
ITGAX
C
Myeloid Cells
TSM
R (TSM, Pseudotime) = 0.506
p < 2.2e-16
R (TSM, Pseudotime) = 0.511
p < 2.2e-16
D
CD4+ T Cells
HALLMARK_INTERFERON_GAMMA_RESPONSE
HALLMARK_INTERFERON_ALPHA_RESPONSE
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION
HALLMARK_COMPLEMENT
HALLMARK_APICAL_JUNCTION
HALLMARK_ALLOGRAFT_REJECTION
HALLMARK_COAGULATION
HALLMARK_APOPTOSIS
HALLMARK_MITOTIC_SPINDLE
HALLMARK_TNFA_SIGNALING_VIA_NFKB
HALLMARK_INFLAMMATORY_RESPONSE
HALLMARK_PI3K_AKT_MTOR_SIGNALING
HALLMARK_ANDROGEN_RESPONSE
HALLMARK_CHOLESTEROL_HOMEOSTASIS
HALLMARK_MTORC1_SIGNALING
HALLMARK_IL2_STAT5_SIGNALING
HALLMARK_KRAS_SIGNALING_UP
HALLMARK_HYPOXIA
Normalized Enrichment Score (NES)
−log10(padj)
E
CD8+ T Cells
HALLMARK_APICAL_JUNCTION
HALLMARK_INTERFERON_ALPHA_RESPONSE
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION
HALLMARK_ALLOGRAFT_REJECTION
HALLMARK_COMPLEMENT
HALLMARK_INTERFERON_GAMMA_RESPONSE
HALLMARK_OXIDATIVE_PHOSPHORYLATION
HALLMARK_PI3K_AKT_MTOR_SIGNALING
Normalized Enrichment Score (NES)
−log10(padj)

Fig. S17. Pseudotime analysis of joint immune cell embeddings. (A) Pseudotime trajectory for B cells on destiny embedding. **(B)** ABC-marker ITGAX expression in B cells. Includes R coefficient between ITGAX expression and pseudotime computed by slingshot for the trajectory. **(C)** Left - myeloid cell TSM score colored on destiny embedding. Center - pseudotime trajectory for myeloid cells on destiny embedding. Includes R coefficient between TSM and pseudotime computed by slingshot for trajectory. Right - Same as (center), but for a monocyte-only trajectory. **(D)** Left - pseudotime trajectory for CD4+ T cells. Right - GSEA barplot on genes correlated with pseudotime trajectory. **(E)** Same as **(D)** but for CD8+ T cells

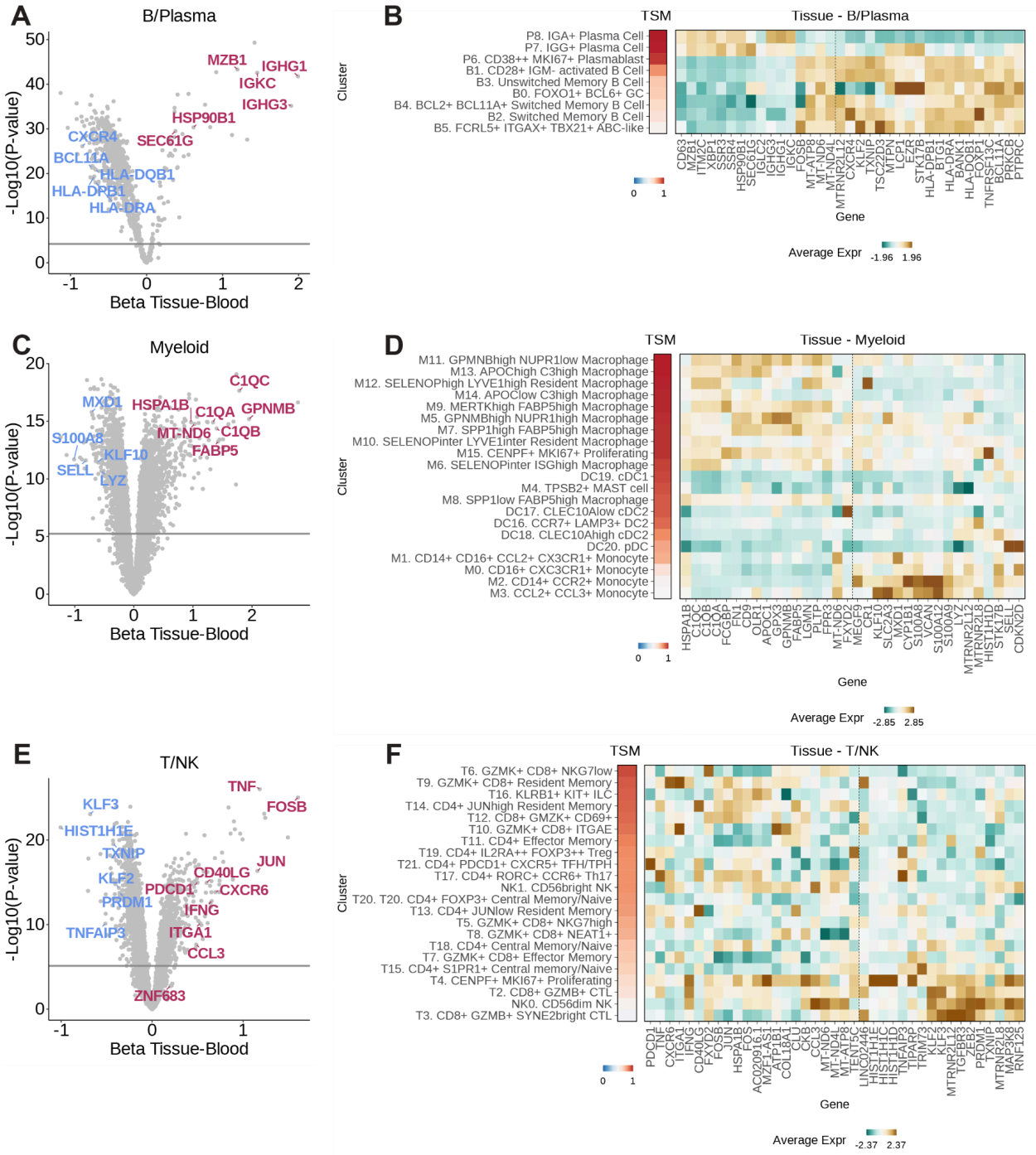

Fig. S18. Differential gene expression between immune compartment in blood and tissue. (A) In B/plasma cells, differential expression results between tissue and blood samples. Betas and p-values are derived from pseudobulk linear modeling. Genes upregulated with respect to tissue are displayed in red, 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.

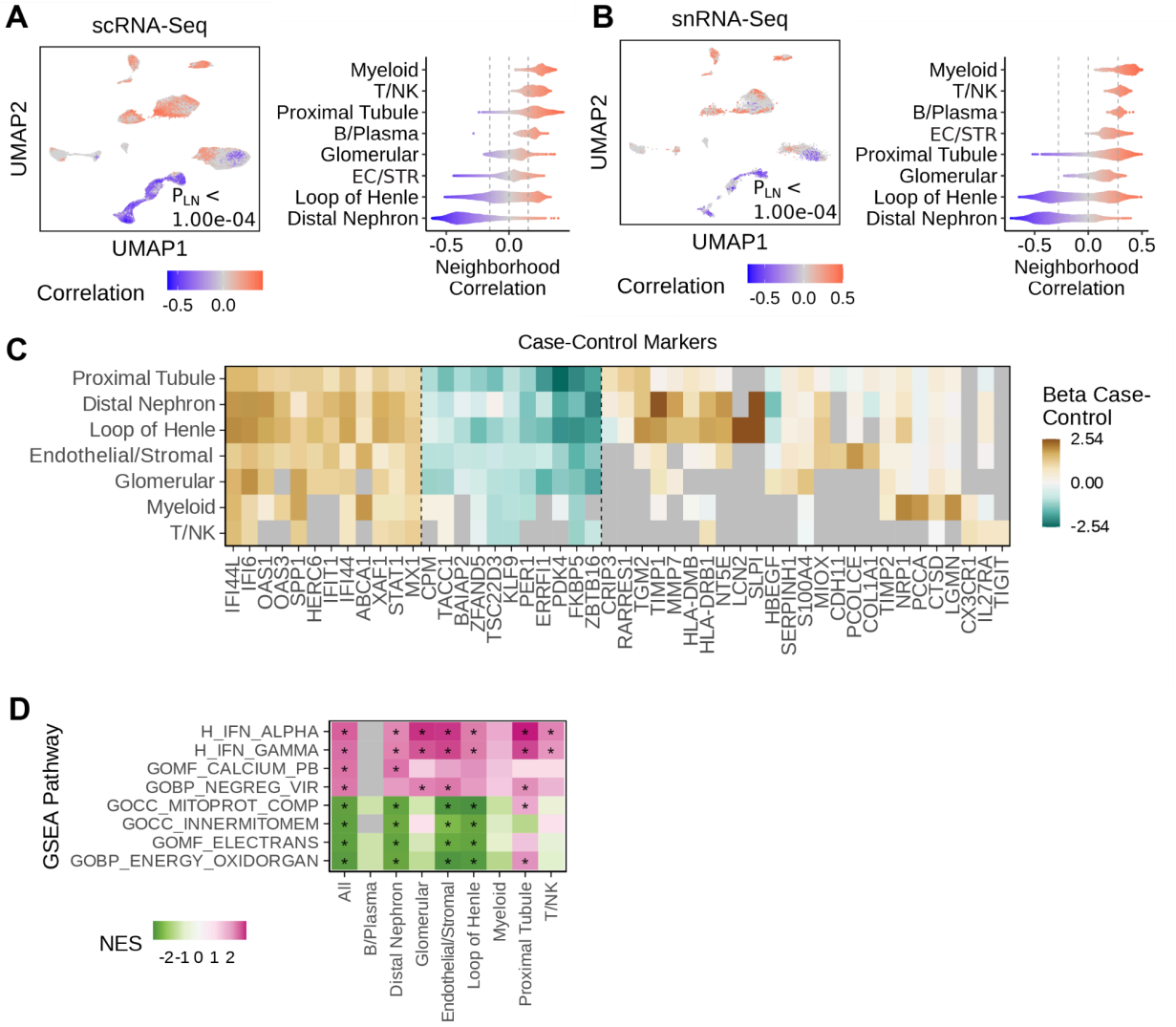

Fig. S19. Whole-tissue case-control CNA and differential expression. (A) Across all cell types, CNA results for case-control association with no covariate corrections, in scRNA-seq data. Left - UMAP displaying significant per-cell associations with LN, with FDR cutoff of 0.05. Non-significant associations 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. Dashed vertical lines represent the correlation threshold with FDR < 0.05. **(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. **(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.

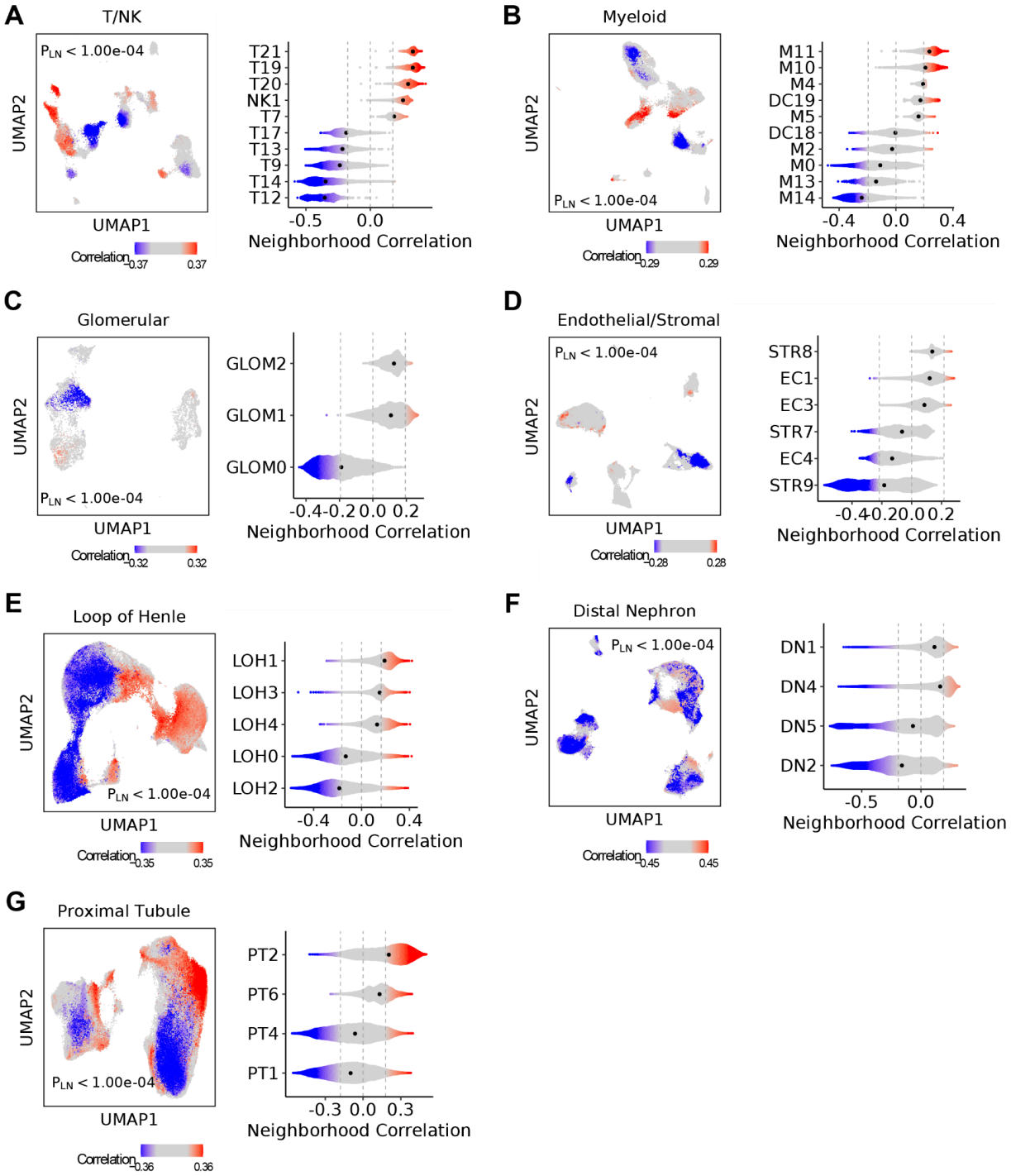
A
T/NK
P_{LN} < 1.00e-04
UMAP2
UMAP1
Correlation
-0.37
0.37
T21
T19
T20
NK1
T7
T17
T13
T9
T14
T12
-0.5
0.0
Neighborhood Correlation
B
Myeloid
P_{LN} < 1.00e-04
UMAP2
UMAP1
Correlation
-0.29
0.29
M11
M10
M4
DC19
M5
DC18
M2
M0
M13
M14
-0.4
0.0
0.4
Neighborhood Correlation
C
Glomerular
P_{LN} < 1.00e-04
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
Proximal Tubule
P_{LN} < 1.00e-04
UMAP2
UMAP1
Correlation
-0.36
0.36
PT2
PT6
PT4
PT1
-0.3
0.0
0.3
Neighborhood Correlation

Fig. S20. Cell state case-control CNA. (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. Dashed vertical lines represent the correlation threshold with FDR < 0.05. **(B)** Same as **(A)** for Myeloid cells. **(C)** Same as (**A)** for B/Plasma cells. **(D)** Same as **(A)** for endothelial/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.

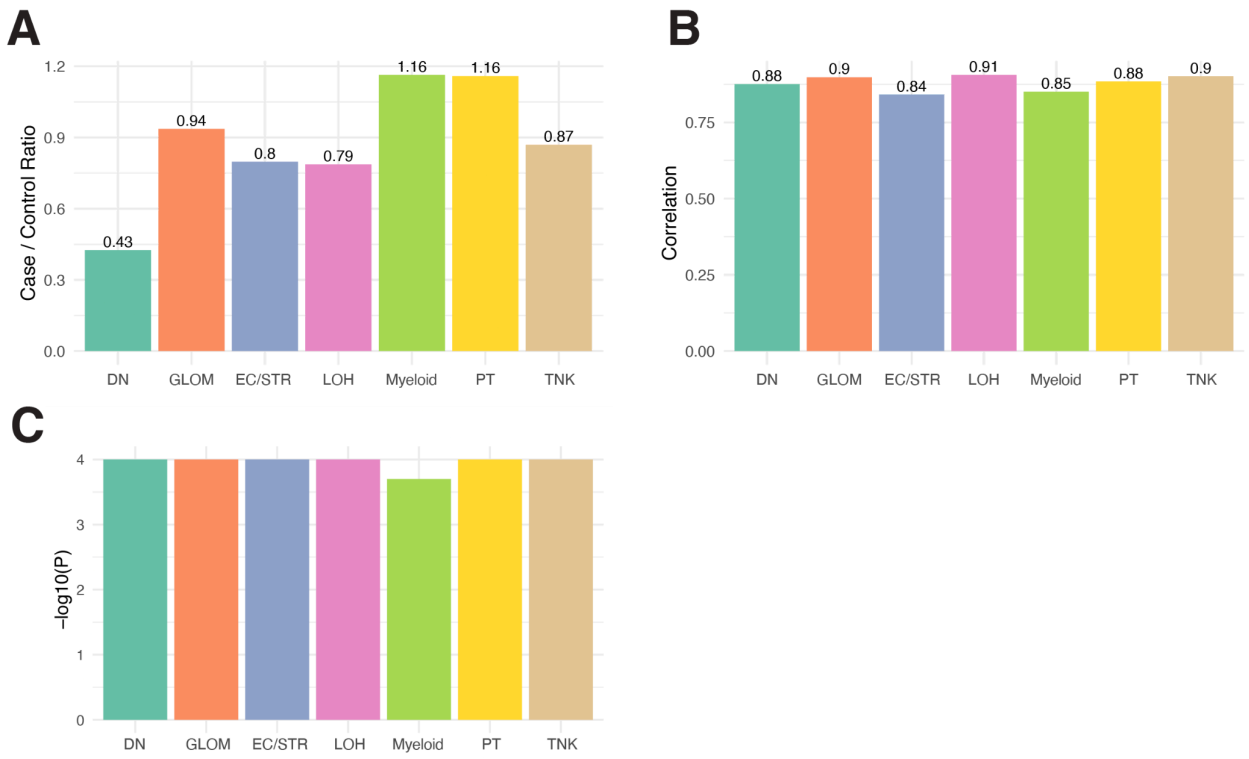

Fig. S21. Subsampling of case-control CNA associations. (A) Ratio of total case/control cell counts stratified by cell type. **(B)** Neighborhood-level correlations of shared cells between case-control CNA and case-control CNA with case samples subsetted to match control sample count and cell count distributions. These results suggest that similar cell populations are expanded or contracted in both data sets. **(C)** Global P value of subsetted case-control associations within cell type.

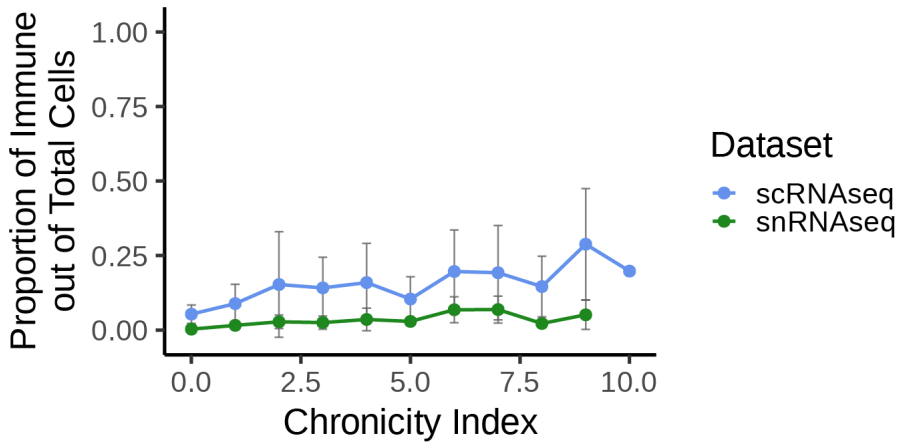

Fig. S22. scRNAseq vs snRNAseq proportions. Per-sample proportions of immune cells (T/NK, B/plasma, myeloid) out of all cells in tissue, subset by data modality. Points represent the mean proportion across samples for a given chronicity index value. Error bars represent the standard deviation across samples at that chronicity index value.

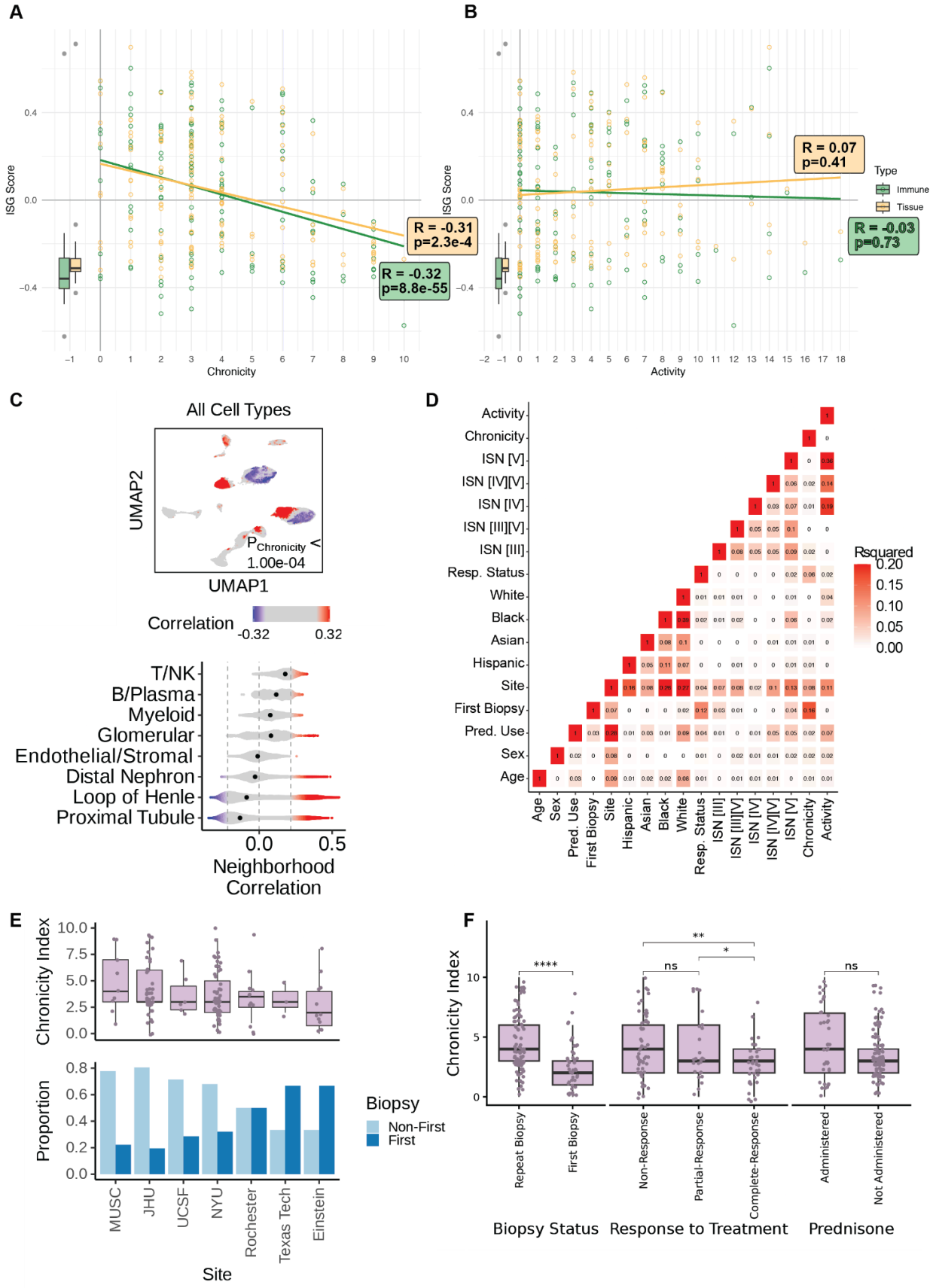
A
ISG Score
Chronicity
R = -0.31
p=2.3e-4
R = -0.32
p=8.8e-55
B
ISG Score
Activity
R = 0.07
p=0.41
R = -0.03
p=0.73
Type
Immune
Tissue
C
All Cell Types
UMAP2
UMAP1
P Chronicity < 1.00e-04
Correlation
-0.32
0.32
T/NK
B/Plasma
Myeloid
Glomerular
Endothelial/Stromal
Distal Nephron
Loop of Henle
Proximal Tubule
Neighborhood Correlation
D
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
Rsquared
E
Chronicity Index
Proportion
Biopsy
Non-First
First
MUSC
JHU
UCSF
NYU
Rochester
Texas Tech
Einstein
Site
F
Chronicity Index
Repeat Biopsy
First Biopsy
Non-Response
Partial-Response
Complete-Response
Administered
Not Administered
Biopsy Status
Response to Treatment
Prednisone

Fig. S23. ISG-score and sample-level covariate analysis. 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. Dashed vertical lines represent the correlation threshold with FDR < 0.05. **(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 $p>0.05$.

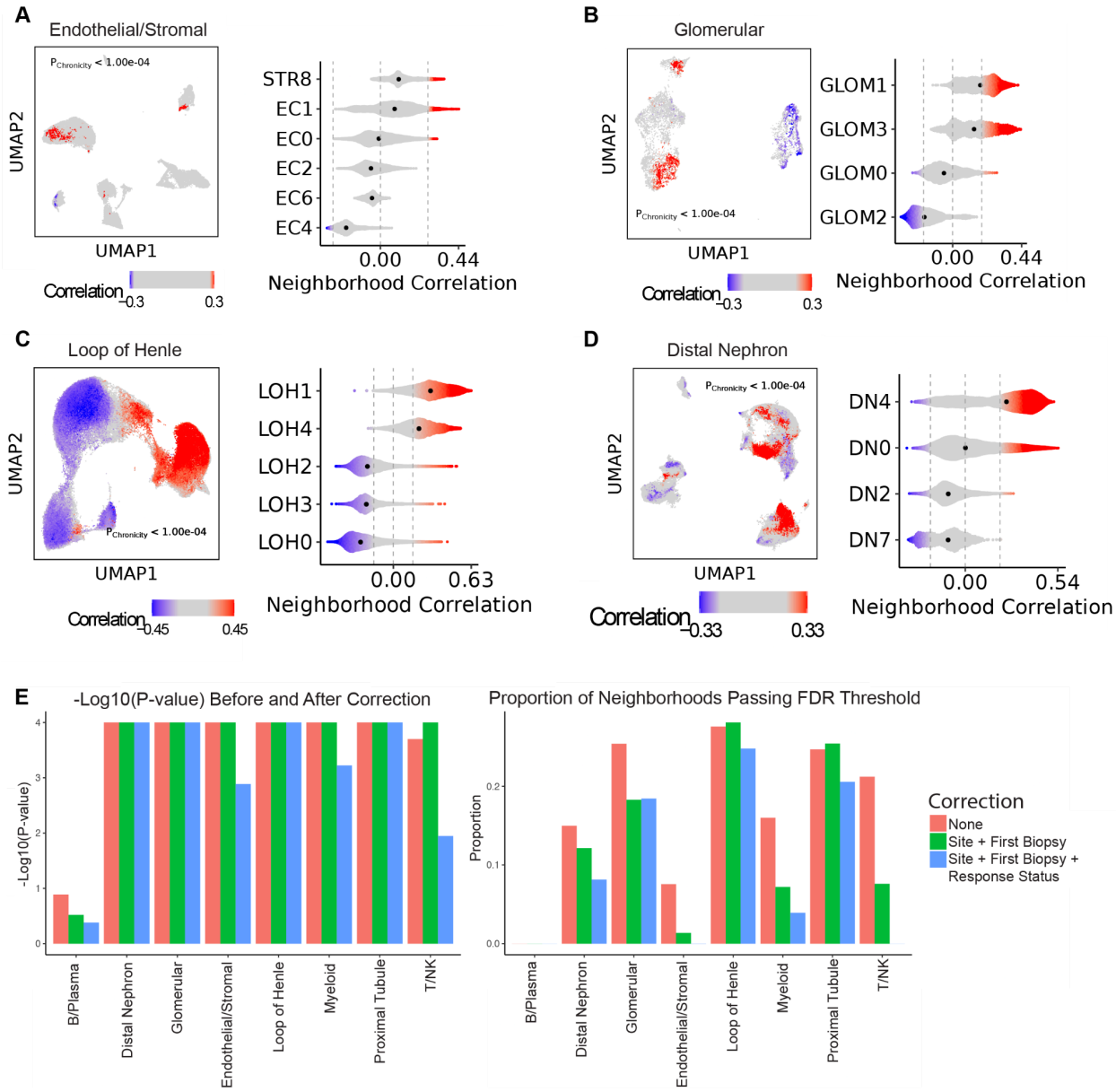

Fig. S24. Cell state chronicity CNA analysis. (A-D) CNA testing correcting for major collection sites and whether the sample is from the initial biopsy following diagnosis for the endothelial/stromal compartment **(A)**, glomerulus **(B)**, loop of Henle **(C)**, or the distal nephron **(D)**. **(E)** Global p values (left) and the percent of neighborhoods that pass FDR<0.10 (right) for CNA testing of chronicity accounting for either no or additional relevant clinical covariates.

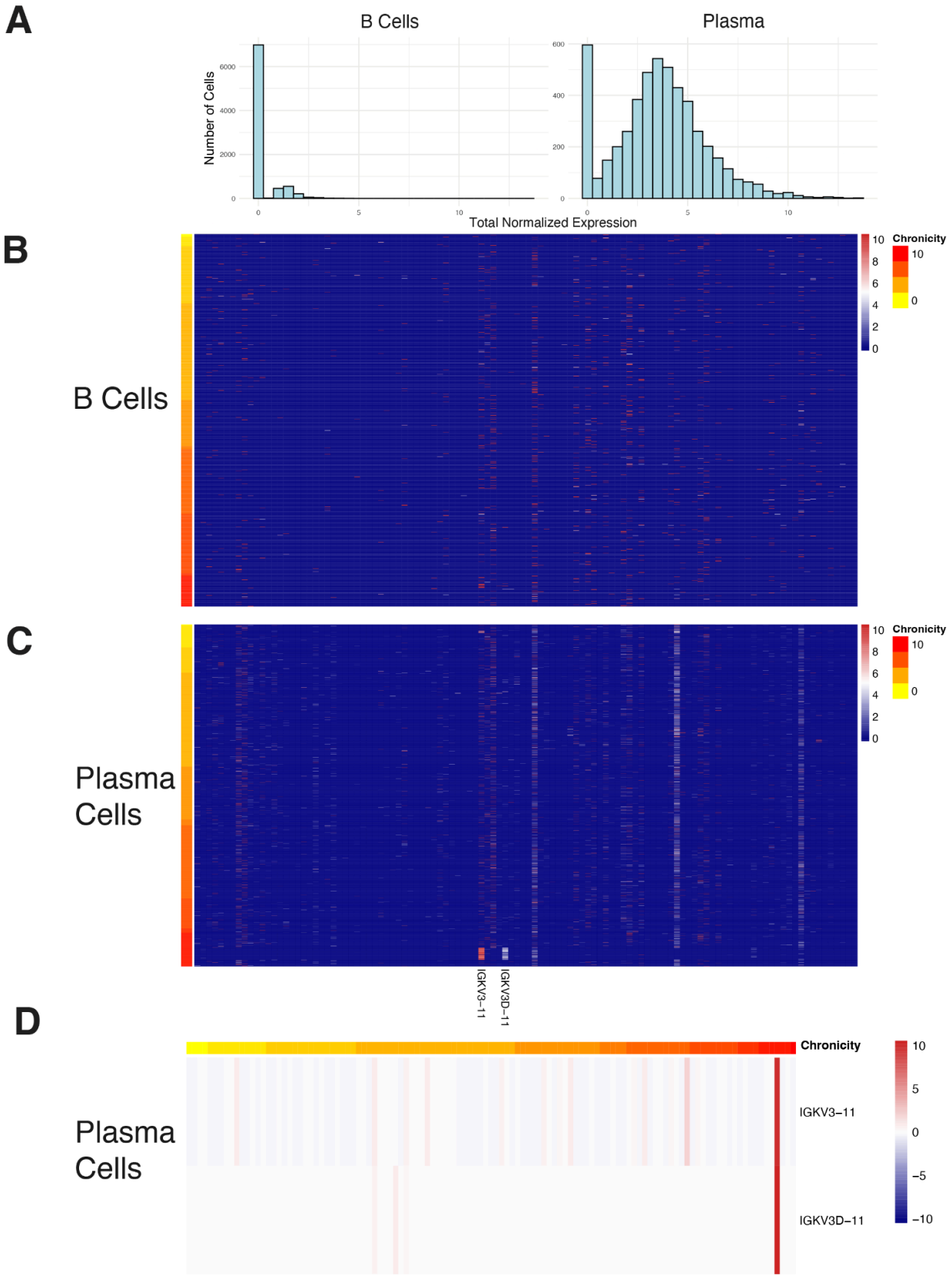
A
B Cells
Plasma
Number of Cells
Total Normalized Expression
B
B Cells
Chronicity
C
Plasma Cells
IGKV3−11
IGKV3D−11
D
Plasma Cells
Chronicity
IGKV3−11
IGKV3D−11

Fig. S25. Clonotyping Analysis of Tissue B cells. (A), Histogram of cell counts across various levels of total normalized expression of light chain genes stratified by B cells (left) and plasma cells (right). **(B)** Heatmap of light chain expression for each light chain (column) per B cell (row). B cells are ordered from cells originating from low-chronicity samples to high-chronicity samples. **(C)** Same as **(B)**, but for plasma cells. **(D)** Light chain genes IGKV3-11 and IGKV3D-11 pseudobulked across plasma cells per sample.

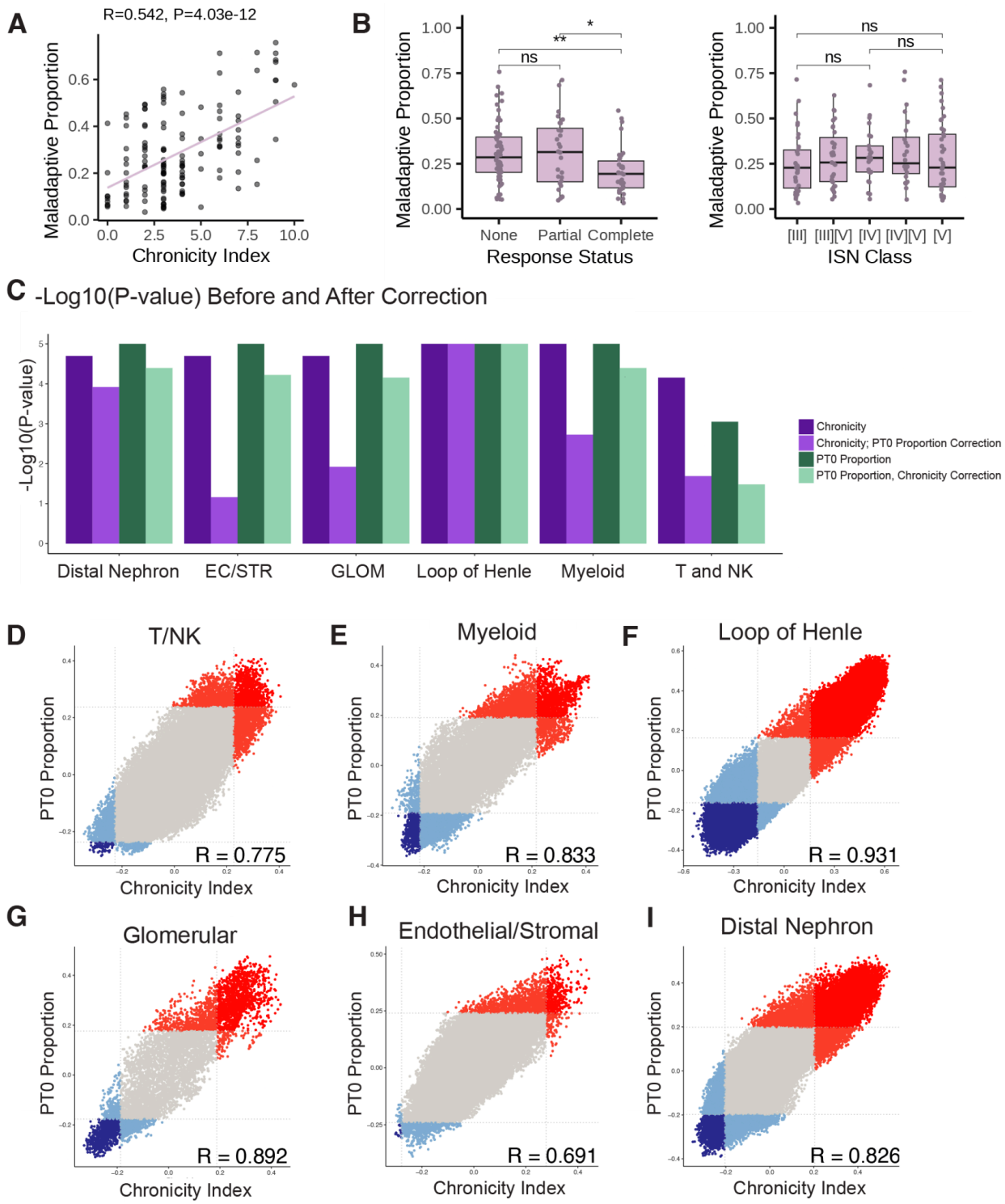
A
R=0.542, P=4.03e-12
Maladaptive Proportion
Chronicity Index
B
Maladaptive Proportion
ns
**
*
None
Partial
Complete
Response Status
ns
ns
ns
[III]
[III][V]
[IV]
[IV][V]
[V]
ISN Class
C
-Log10(P-value) Before and After Correction
-Log10(P-value)
Distal Nephron
EC/STR
GLOM
Loop of Henle
Myeloid
T and NK
Chronicity
Chronicity; PT0 Proportion Correction
PT0 Proportion
PT0 Proportion, Chronicity Correction
D
T/NK
R = 0.775
E
Myeloid
R = 0.833
F
Loop of Henle
R = 0.931
G
Glomerular
R = 0.892
H
Endothelial/Stromal
R = 0.691
I
Distal Nephron
R = 0.826
PT0 Proportion
Chronicity Index

Fig. S26. PT0-chronicity analysis. (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)**, endothelial/stromal **(H)**, or distal nephron **(I)** compartments.

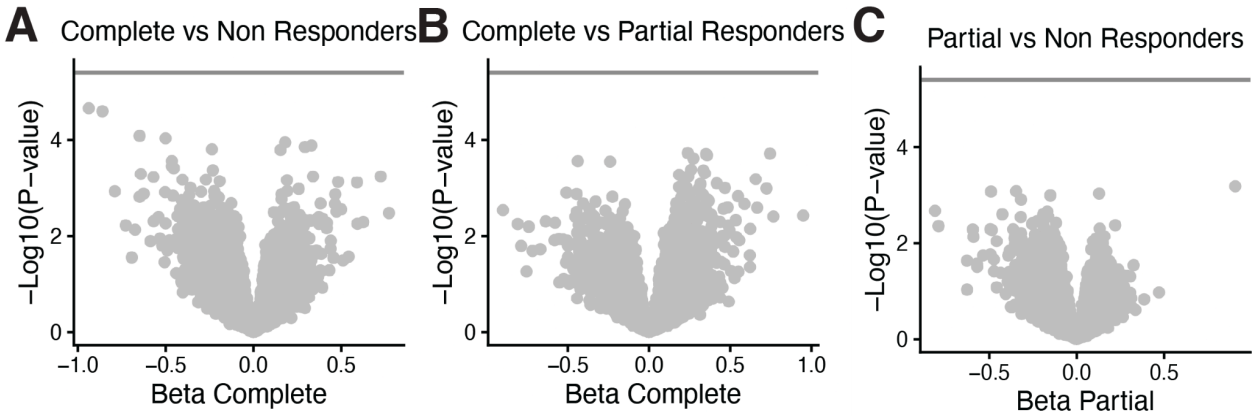

Fig. S27. Differential gene expression of PT0 between treatment response groups. Volcano plots of differentially expressed genes within PT0 between treatment groups for complete responders vs non responders **(A)**, complete vs partial responders **(B)**, and partial vs non responders **(C).**

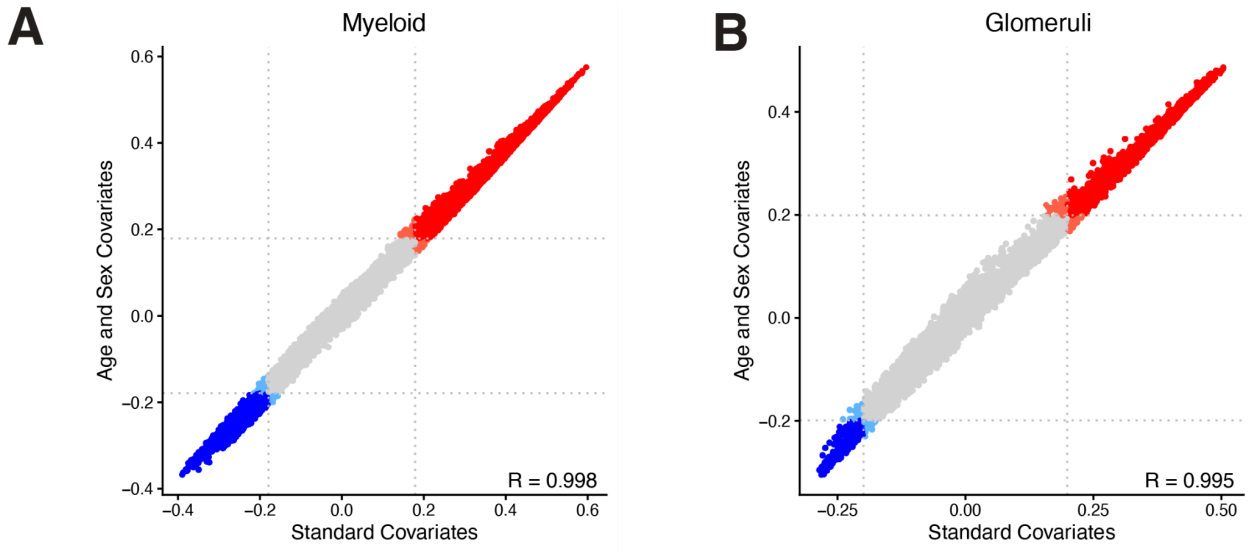

Fig. S28. Concordances between standard covariates and additional age and sex covariates in activity index CNA. **(A, B)** Correlations of activity index CNA with standard covariates compared to additional age and sex covariates in myeloid **(A)** and glomerular **(B)** cell types. Horizontal lines are drawn at each individual analyses FDR < 0.10 threshold.

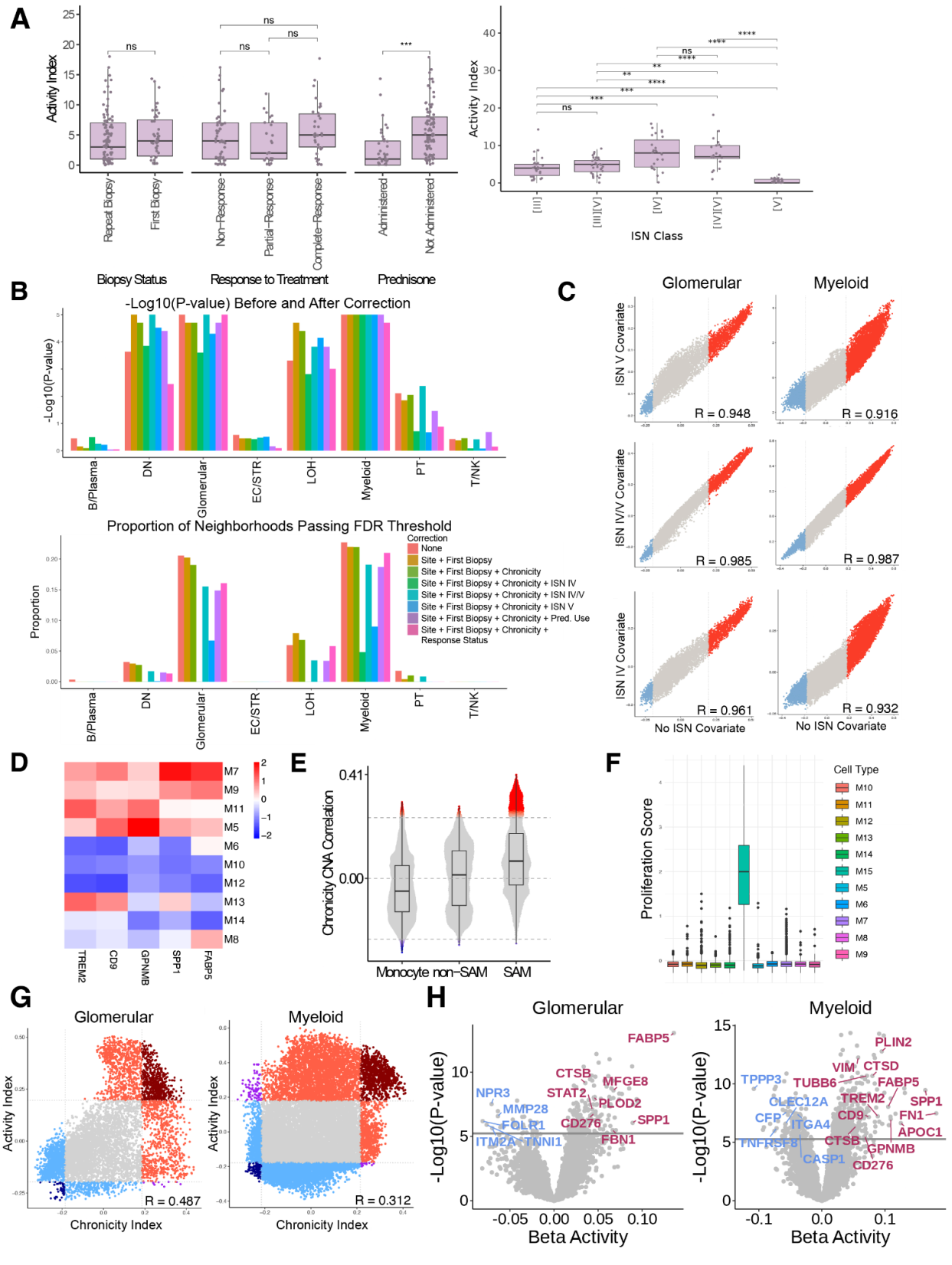
A
Activity Index
Repeat Biopsy
First Biopsy
Biopsy Status
Non-Response
Partial-Response
Complete-Response
Response to Treatment
Administered
Not Administered
Prednisone
ISN Class
B
-Log10(P-value) Before and After Correction
-Log10(P-value)
B/Plasma
DN
Glomerular
EC/STR
LOH
Myeloid
PT
T/NK
Proportion 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
R = 0.948
R = 0.916
R = 0.985
R = 0.987
R = 0.961
R = 0.932
No ISN Covariate
D
M7
M9
M11
M5
M6
M10
M12
M13
M14
M8
TREM2
CD9
GPNMB
SPP1
FABP5
E
Chronicity CNA Correlation
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

Fig. S29. Activity CNA analysis. (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<0.1 (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. 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 $p<0.05$.

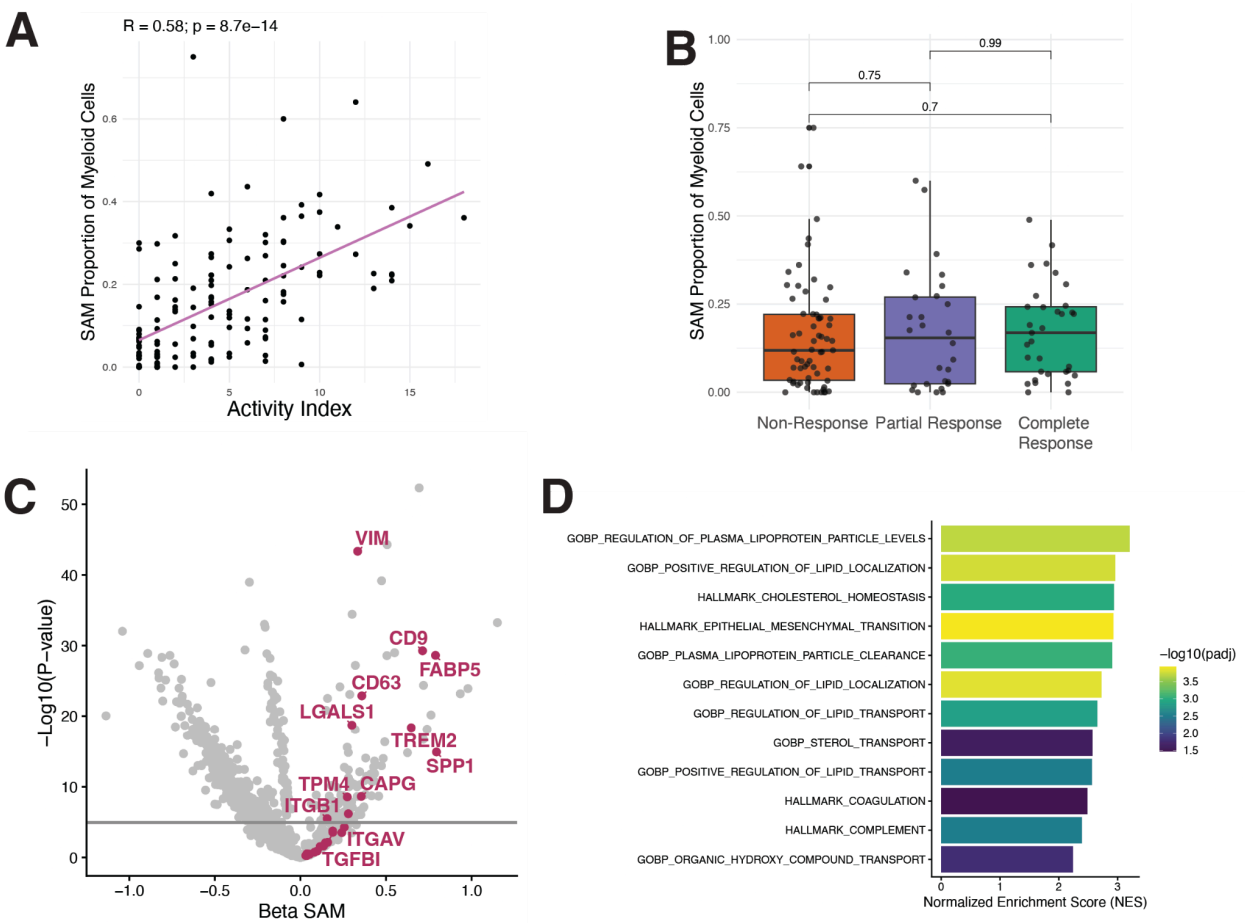

Fig. S30. Differential gene expression of SAMs over other myeloid cells. **(A)** Correlation plot of activity and per-sample SAM proportion of myeloid cells. **(B)** Boxplot of per-sample SAM proportion stratified by treatment response. **(C)** Volcano plot of SAM vs nonSAM myeloid cell differential gene expression. Horizontal line drawn at -log10(p) = 0.05/4,455 genes tested. **(D)** GSEA bar plot indicating Normalized Enrichment Score (NES) and -log10(Bonferonni-corrected p value).

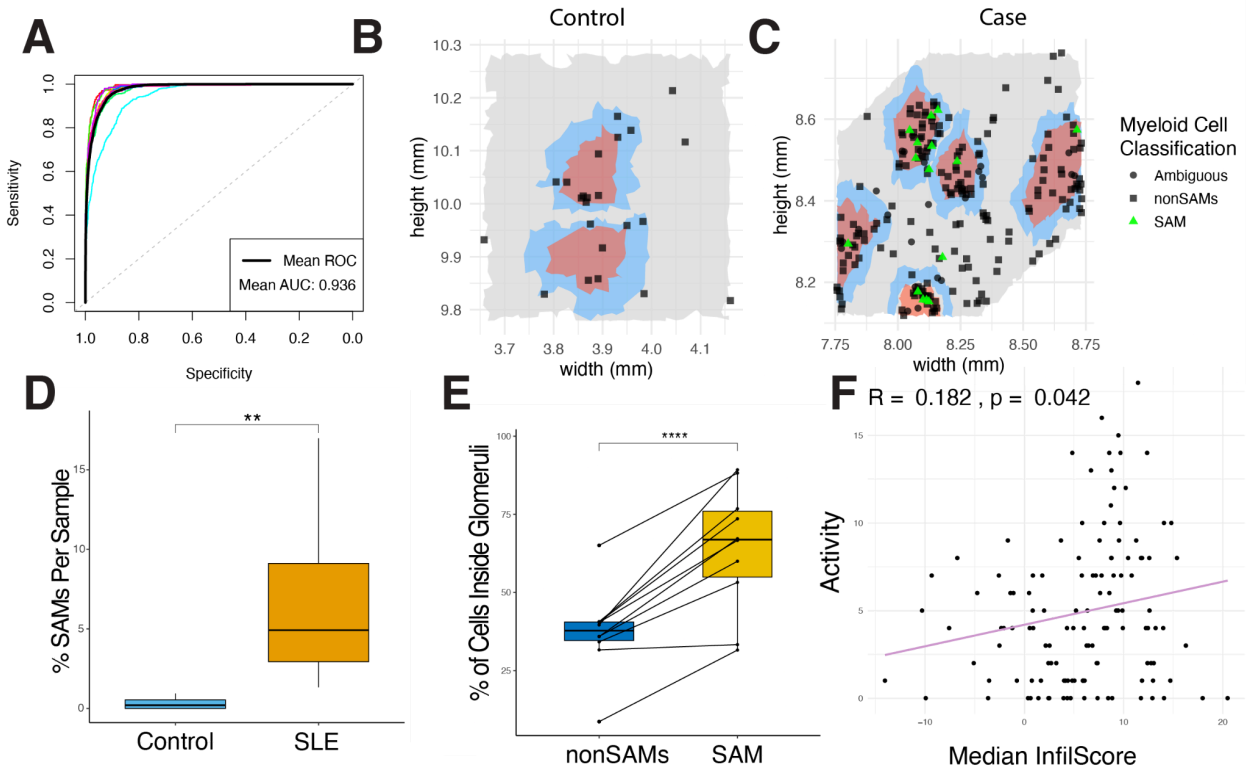

Fig. S31. Integration of Myeloid Classifier with Spatial Data. (A) Average AUC of myeloid cell classifier. **(B, C)** Non-myeloid cells colored by spatial niche and myeloid cells distinguished by shape and colored by SAM marker FABP5 for a select control **(B)** and case **(C)** FOV. **(D)** Boxplot of percent SAMs per sample in control and SLE cases in spatial data. **(E)** Boxplot of percent of myeloid cells in case samples inside glomerular niches. **(F)** Per-sample scatterplot of median sample infiltration score versus clinical activity index.

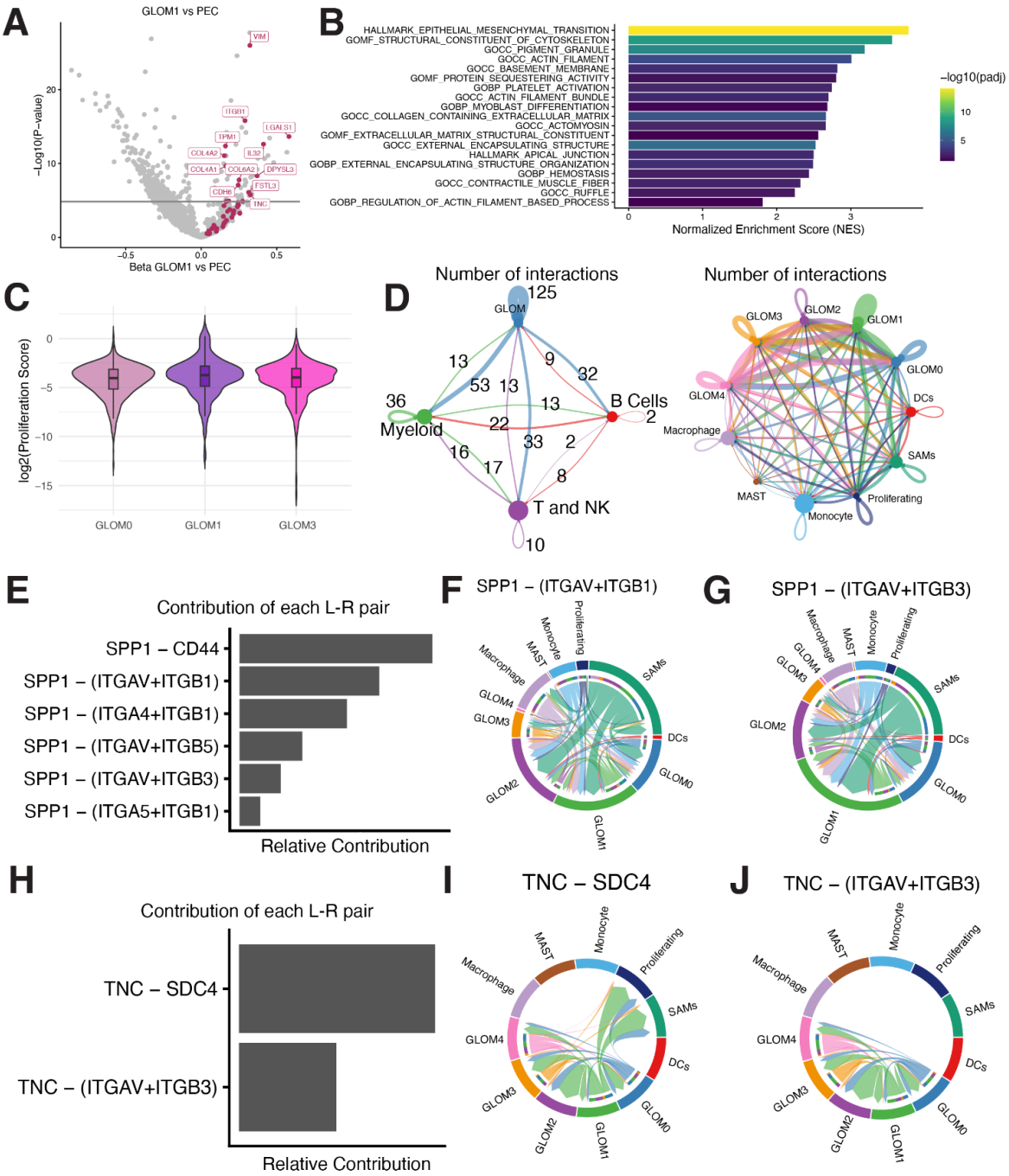
A
GLOM1 vs PEC
−Log10(P−value)
Beta GLOM1 vs PEC
VIM
ITGB1
LGALS1
TPM1
IL32
COL4A2
COL4A1
COL6A2
DPYSL3
FSTL3
CDH6
TNC
B
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION
GOMF_STRUCTURAL_CONSTITUENT_OF_CYTOSKELETON
GOCC_PIGMENT_GRANULE
GOCC_ACTIN_FILAMENT
GOCC_BASEMENT_MEMBRANE
GOMF_PROTEIN_SEQUESTERING_ACTIVITY
GOBP_PLATELET_ACTIVATION
GOCC_ACTIN_FILAMENT_BUNDLE
GOBP_MYOBLAST_DIFFERENTIATION
GOCC_COLLAGEN_CONTAINING_EXTRACELLULAR_MATRIX
GOCC_ACTOMYOSIN
GOMF_EXTRACELLULAR_MATRIX_STRUCTURAL_CONSTITUENT
GOCC_EXTERNAL_ENCAPSULATING_STRUCTURE
HALLMARK_APICAL_JUNCTION
GOBP_EXTERNAL_ENCAPSULATING_STRUCTURE_ORGANIZATION
GOBP_HEMOSTASIS
GOCC_CONTRACTILE_MUSCLE_FIBER
GOCC_RUFFLE
GOBP_REGULATION_OF_ACTIN_FILAMENT_BASED_PROCESS
Normalized Enrichment Score (NES)
−log10(padj)
C
log2(Proliferation Score)
GLOM0
GLOM1
GLOM3
D
Number of interactions
GLOM
Myeloid
B Cells
T and NK
Number of interactions
GLOM3
GLOM2
GLOM1
GLOM0
GLOM4
DCs
Macrophage
SAMs
MAST
Proliferating
Monocyte
E
Contribution of each L-R pair
SPP1 – CD44
SPP1 – (ITGAV+ITGB1)
SPP1 – (ITGA4+ITGB1)
SPP1 – (ITGAV+ITGB5)
SPP1 – (ITGAV+ITGB3)
SPP1 – (ITGA5+ITGB1)
Relative Contribution
F
SPP1 – (ITGAV+ITGB1)
G
SPP1 – (ITGAV+ITGB3)
H
Contribution of each L-R pair
TNC – SDC4
TNC – (ITGAV+ITGB3)
Relative Contribution
I
TNC – SDC4
J
TNC – (ITGAV+ITGB3)

Fig. S32. GLOM1 Gene Signatures. (A) Volcano plot of differentially expressed genes between GLOM1 and other PEC cells. Epithelial-mesenchymal transition genes are highlighted in red. A horizontal line is drawn at $p < 0.05$ after Bonferroni correction for genes tested. **(B)** Top significantly upregulated GSEA enrichment of differentially expressed genes in GLOM1 compared to other PEC populations. **(C)** Violin plot of log2 of infiltration score stratified by cell type. **(D)** Receptor-ligand analysis between glomeruli and broad immune tissue cell types (left) and subcategorized macrophage populations (left). **(E)** Relative contributions of ligand-receptor pairs of the SPP1 pathway. **(F-G)** Chord plot of select ligand-receptor pairings in the SPP1 pathway grouped by cell state for SPP1 and ITGAV+ITVB1 heterodimer (left) and ITGAV+ITGB3 heterodimer (right). **(H)** Relative contributions of ligand-receptor pairs of the Tenasin pathway. **(I-J)** Chord plot of select ligand-receptor pairings in the SPP1 pathway grouped by cell state for SPP1 and SDC4 (left) and ITGAV+ITGB3 heterodimer (right).

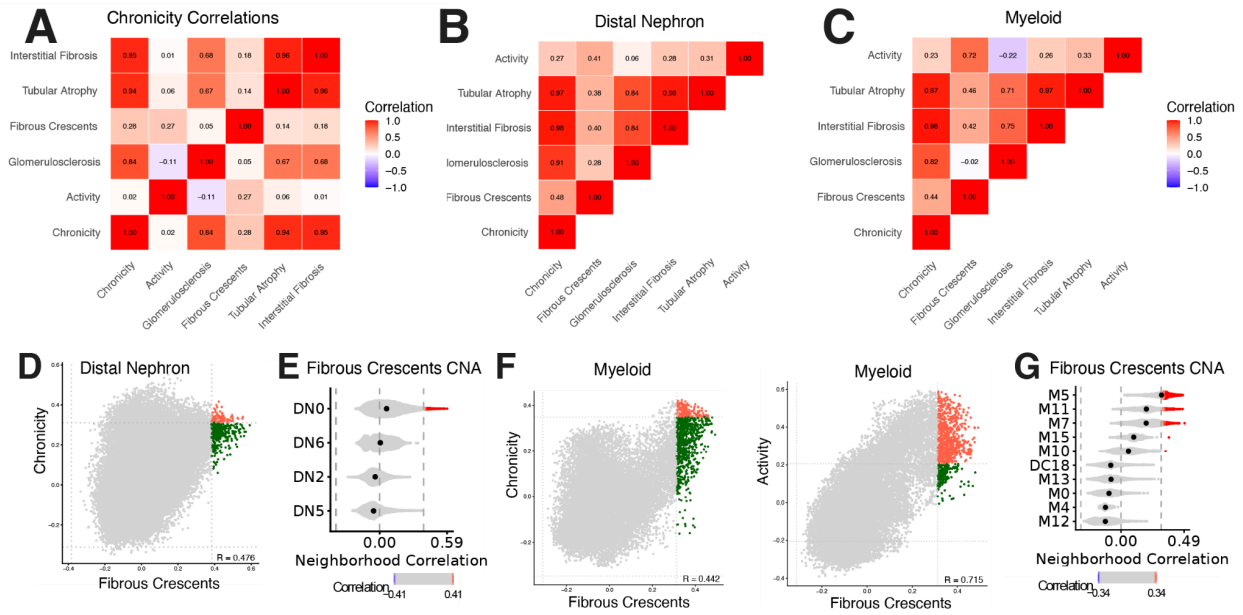

Fig. S33. CNA with Chronicity Index Components. (A) Sample level correlations for chronicity components. **(B-C)** Neighborhood-level correlations for each chronicity component in distal nephron **(B)** and myeloid **(C)** cells. **(D)** Neighborhood-level correlation scatterplot for chronicity and fibrous crescent CNA. Grey lines are added for each test's respective FDR threshold. Neighborhoods concordantly expanded in both CNA and fibrous crescent CNA are colored red, and novel neighborhoods significantly expanded in fibrous crescents but not in chronicity are colored in green. **(E)** Violin plot of neighborhood correlations of fibrous crescent CNA in distal nephron cells. Dashed gray line added at FDR < 10%. **(F)** Same as **(D)** for myeloid cells. **(G)** Same as **(E)** for myeloid cells. Neighborhoods that are significant in the subindex but not overall activity are highlighted in green.

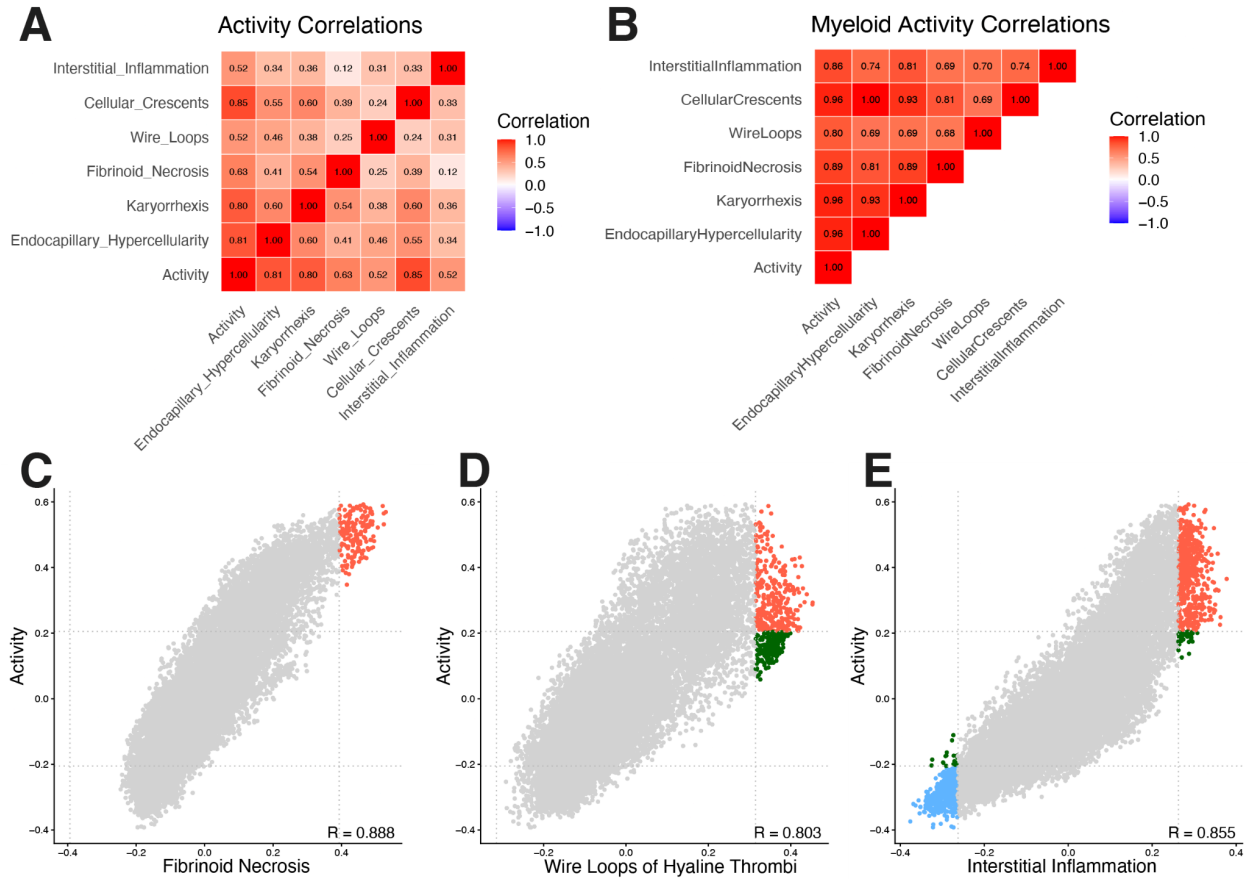

Fig. S34. CNA with Activity Index Components. **(A)** Sample level correlations for activity components. **(B)** Neighborhood-level correlations for each activity component in myeloid cells. **(C-E)** Neighborhood-level correlation scatterplot lines between activity CNA and fibrinoid necrosis **(C),** wire **l**oops or hyaline thrombi **(D)**, and interstitial inflammation **(E)**. Dashed gray lines are added for each test's respective FDR threshold. New neighborhoods that are significant in subindex CNA but not activity CNA are highlighted in green. Neighborhoods that are concordantly significantly expanded or depleted in both are colored red and blue, respectively.

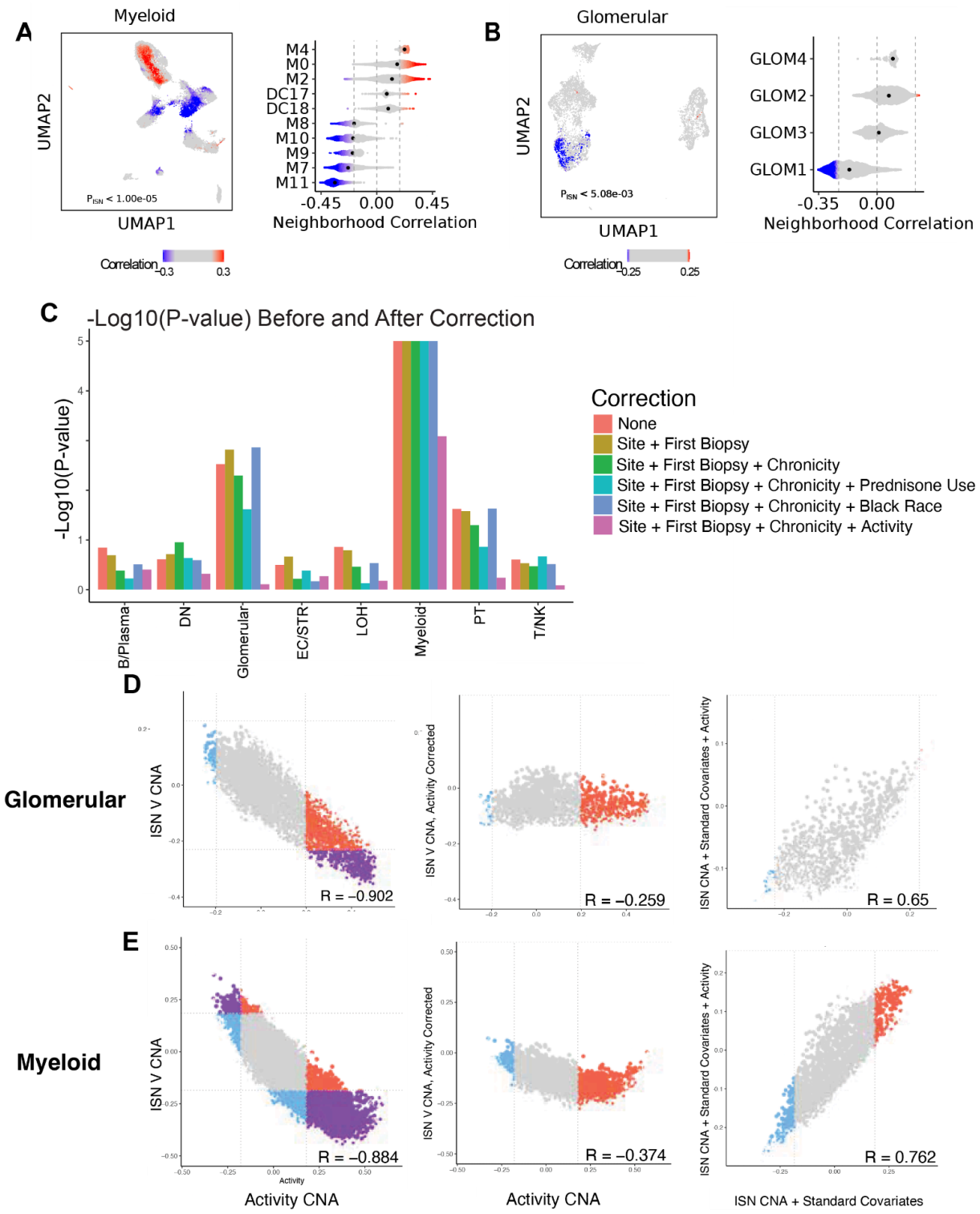
A
Myeloid
UMAP2
UMAP1
P_ISN < 1.00e-05
Correlation
-0.3
0.3
M4
M0
M2
DC17
DC18
M8
M10
M9
M7
M11
-0.45
0.00
0.45
Neighborhood Correlation
B
Glomerular
UMAP2
UMAP1
P_ISN < 5.08e-03
Correlation
-0.25
0.25
GLOM4
GLOM2
GLOM3
GLOM1
-0.35
0.00
Neighborhood Correlation
C
-Log10(P-value) Before and After Correction
-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

Fig. S35. Activity CNA and ISN V CNA analysis. (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. Left - UMAP displaying significant per-cell associations with ISN Class V, with FDR cutoff of 0.1. 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 ISN class V association. Dashed vertical lines represent the correlation threshold with FDR < 0.1. **(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.

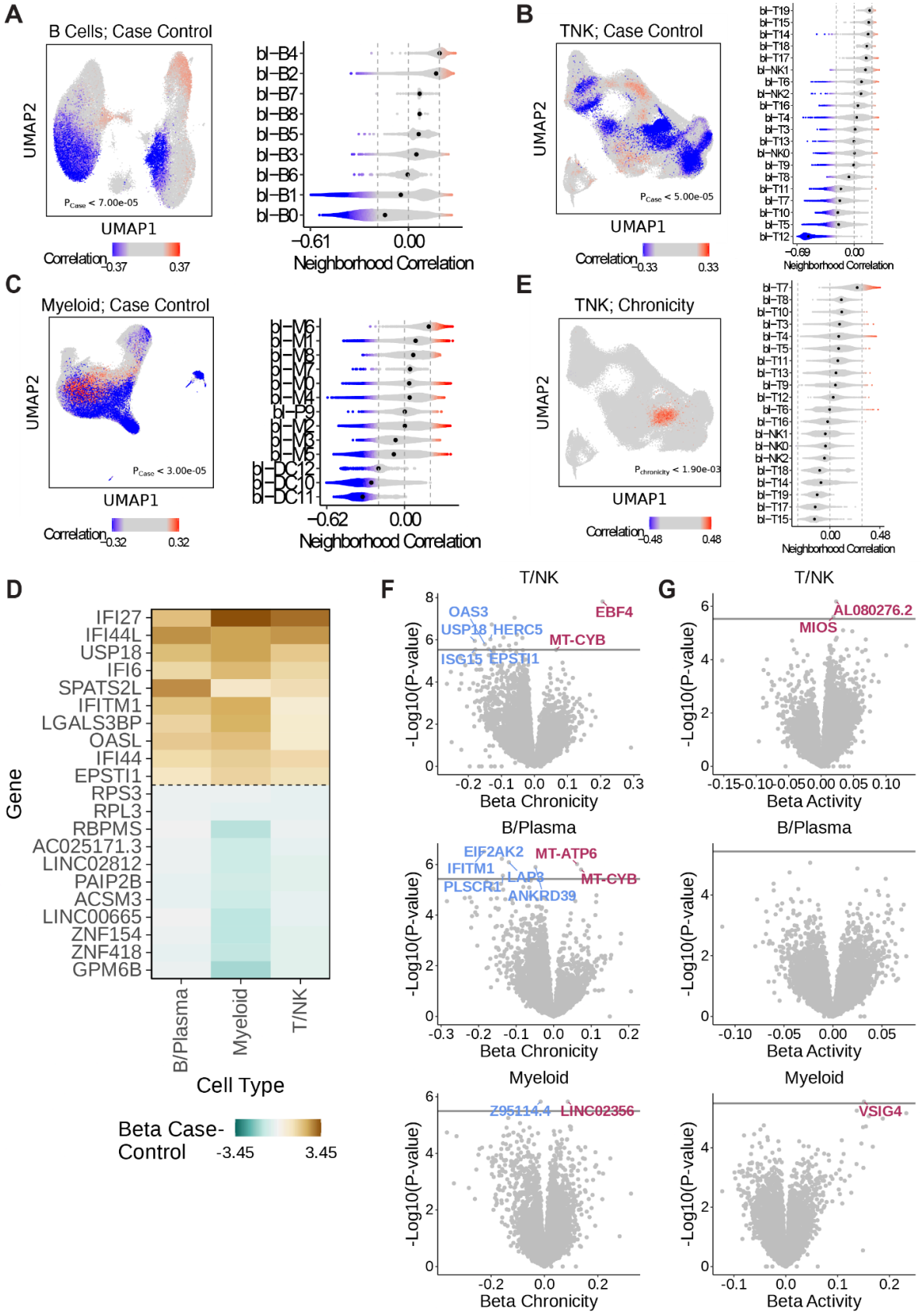
A
B Cells; Case Control
UMAP2
UMAP1
P_Case < 7.00e-05
Correlation
-0.37
0.37
bl−B4
bl−B2
bl−B7
bl−B8
bl−B5
bl−B3
bl−B6
bl−B1
bl−B0
−0.61
0.00
Neighborhood Correlation
B
TNK; Case Control
P_Case < 5.00e-05
-0.33
0.33
-0.69
C
Myeloid; Case Control
P_Case < 3.00e-05
-0.32
0.32
-0.62
E
TNK; Chronicity
P_chronicity < 1.90e-03
-0.48
0.48
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
Myeloid
Z95114.4
LINC02356
G
T/NK
AL080276.2
MIOS
Beta Activity
B/Plasma
Myeloid
VSIG4

Fig. S36. PBMC CNA analysis. (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. Left - UMAP displaying significant per-cell associations. Non-significant associations are colored in grey. P-value is the global P-value for cell state associations. Right - violin plots of clusters containing cells passing FDR thresholding, indicated by the vertical lines. **(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.

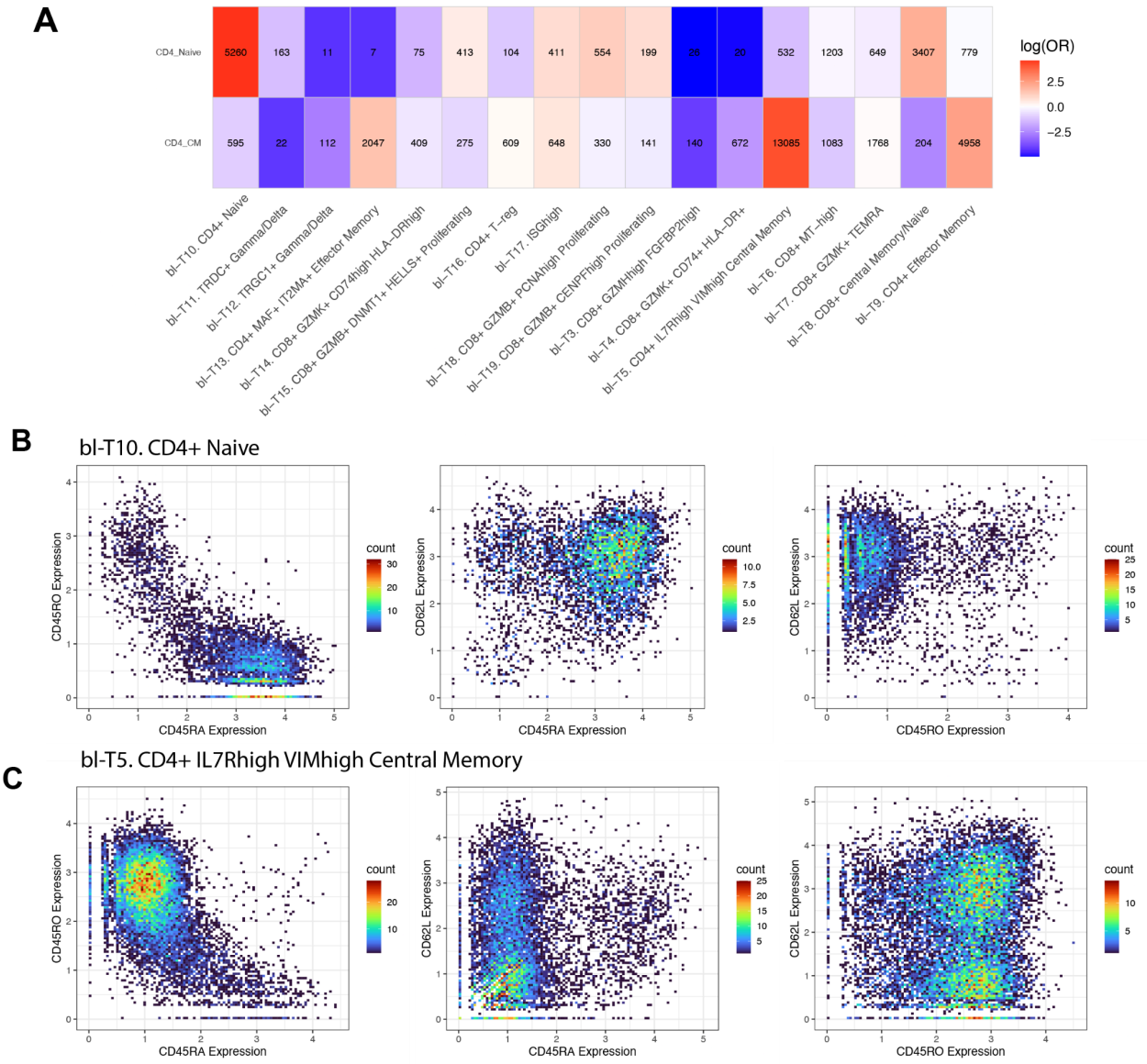

Fig. S37. CITE-Seq of PBMC T cells. (A) log odds ratio heatmap of TCAT multinomial classification on PBMC T cells. Boxes are labeled with the number of cells assigned by TCAT to that label. **(B-E)** Normalized expression of CD45RA **(B, D)** and CD45RO **(C, E)** protein on all the T cells **(B, C)** and the bl-T10 cluster **(D, E)**. **(F)** Scatterplot of CD45RA protein and CD45RO protein expression within bl-T10.

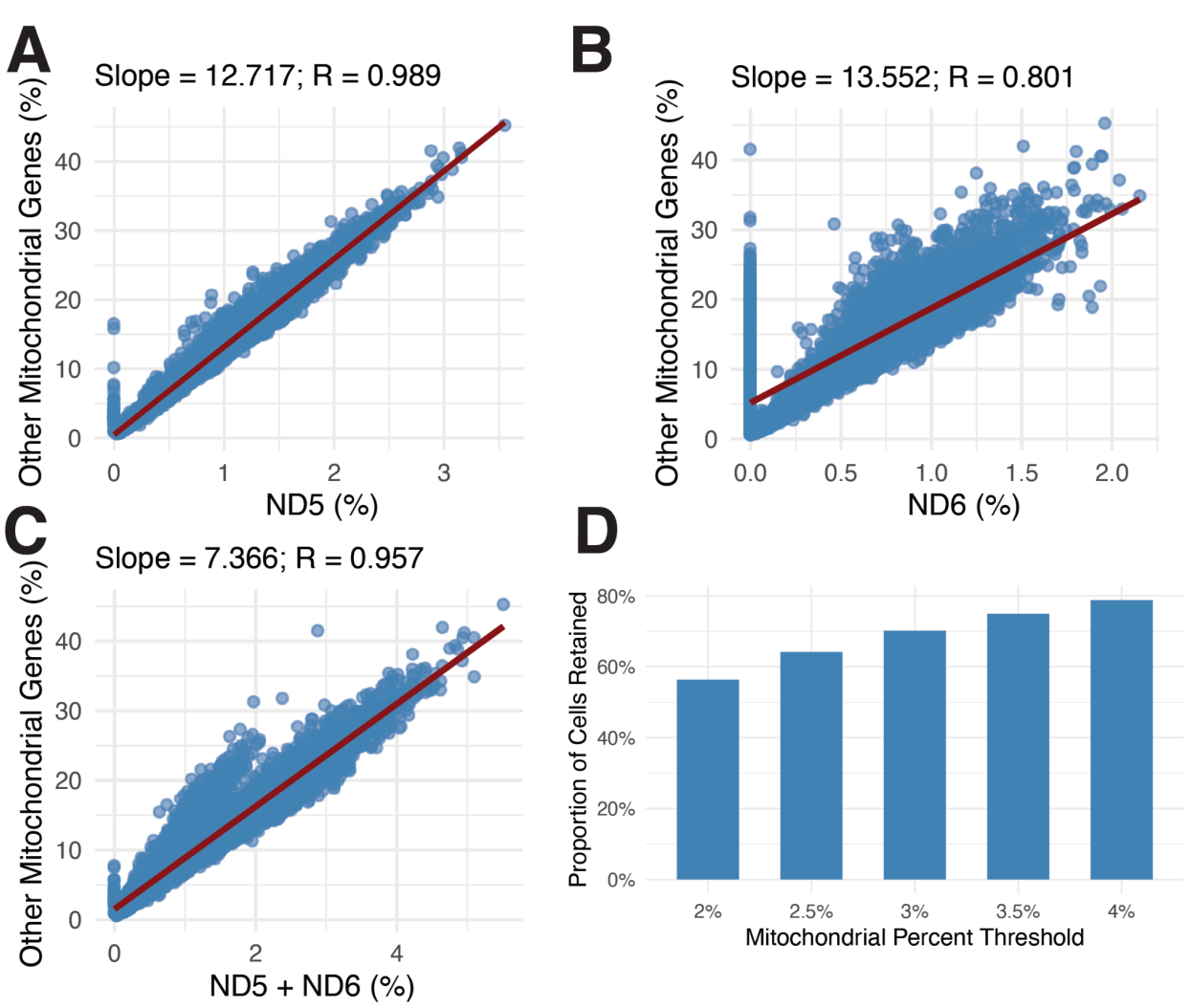

Fig. S38. DASH Mitochondrial quality control. (A-B) Correlational plots of percentage of mitochondrial genes for MT-ND5 **(A)**, MT-ND6 **(B)**, and MT-ND5 + MT-ND6 against total mitochondrial content **(C)**. **(D)** Barplots of proportion of cells retained in AMP dataset at various MT-ND5 and MT-ND6 cutoffs.

Supplementary Tables

Table S1. Cohort demographics

Table S2. Cell and sample counts for data modalities

Table S3. Cell counts before and after quality control

Table S4. Tissue cell state statistics

Table S5. Chronicity differential gene expression results

Table S6. CNA cell state correlations with chronicity index

Table S7. Sample-level correlations with chronicity index

Table S8. Chronicity CNA p-values in tissue

Table S9. Broad cell type Chronicity CNA FDR thresholding

Table S10. Activity differential gene expression results

Table S11. Sample-level correlations with activity index

Table S12. Chronicity subcomponent CNA p-values in tissue

Table S13. Activity subcomponent CNA p-values in tissue

Table S14. Donor counts for blood CNA analysis

Table S15. P values for blood CNA analysis

Table S16. DASH Sequences