

Immune cell quantification of in situ inflammation partitions human lupus nephritis into mechanistic subtypes

Gabriel Casella,^{1,2} Madeleine S. Torcasso,^{1,2} Junting Ai,¹ Thao P. Cao,^{1,3} Satoshi Hara,^{1,4} Michael S. Andrade,⁵ Deepjyoti Ghosh,¹ Daming Shao,¹ Anthony Chang,⁶ Kichul Ko,¹ Anita S. Chong,⁵ Maryellen L. Giger,² and Marcus R. Clark¹

¹Section of Rheumatology and Gwen and Jules Knapp Center for Immunology and Lupus Research, Department of Medicine and ²Department of Radiology, University of Chicago, Chicago, Illinois, USA.

³Pritzker School of Molecular Engineering, University of Chicago, Chicago, Illinois, USA. ⁴Department of Nephrology and Rheumatology, Kanazawa University Hospital, Kanazawa, Japan. ⁵Department of Surgery and ⁶Department of Pathology, University of Chicago, Chicago, Illinois, USA.

BACKGROUND. In human lupus nephritis (LuN), tubulointerstitial inflammation (TII) is prognostically more important than glomerular inflammation. However, a comprehensive understanding of both TII complexity and heterogeneity is lacking.

METHODS. Herein, we used high-dimensional confocal microscopy, spatial transcriptomics, and specialized computer vision techniques to quantify immune cell populations and localize these within normal and diseased renal cortex structures. With these tools, we compared LuN to renal allograft rejection (RAR) and normal kidney tissues on 54 deidentified biopsies.

RESULTS. In both LuN and RAR, the 33 characterized immune cell populations formed discrete subgroups whose constituents covaried in prevalence across biopsies. In both diseases, these covariant immune cell subgroups organized into the same unique niches. Therefore, inflammation could be resolved into trajectories representing the relative prevalence and density of cardinal immune cell members of each covariant subgroup. Indeed, in any one biopsy, the inflammatory state could be characterized by quantifying constituent immune cell trajectories. Remarkably, LuN heterogeneity could be captured by quantifying a few myeloid immune cell trajectories, while RAR was more complex with additional T cell trajectories.

CONCLUSIONS. Our studies identify rules governing renal inflammation and thus provide an approach for resolving LuN into discrete mechanistic categories.

FUNDING. NIH (U19 AI 082724 [MRC], R01 AI148705 [MRC and ASC]), Chan Zuckerberg Biohub (MRC), and Lupus Research Alliance (MRC).

Introduction

Among the severe manifestations of systemic lupus erythematosus (SLE), lupus nephritis (LuN) is the most common (1–5). Up to 50% of SLE patients develop lupus nephritis, in many cases necessitating treatment with toxic immunosuppressive therapies such as cyclophosphamide or mycophenolate mofetil (6–8). Despite such aggressive treatments, many patients do not respond to therapy, and up to 40% of LuN patients progress to renal failure within 5 years of diagnosis (9–12).

Much work in human LuN has focused on peripheral autoimmunity and its manifestation within the kidney, glomerulonephritis (GN) (13). However, prognosis is more tightly linked to tubulointer-

stitial inflammation (TII) and scarring (11, 14–21). TII, and to some degree GN, are associated with complex in situ immune states. Single cell (sc) RNA-Seq from human lupus biopsies has revealed at least 21 different adaptive and innate immune cell clusters (22), which organize into complex tubulointerstitial structures ranging from small neighborhoods of CD8⁺ T cells through T:B cell aggregates to canonical germinal centers (23–27). TII is associated with selection of B cells expressing unique antibody repertoires and cognate T cell:antigen presenting cell (APC) networks (24, 28, 29). These data indicate that LuN is associated with multiple in situ immune mechanisms hypothesized to drive local inflammation and tissue destruction (30).

Limited studies have related specific in situ immune cell populations to prognosis. For example, infiltrating CD8⁺ T cells have been associated with a poor prognosis (31, 32). In a study of principal adaptive immune cell populations, densities of CD4⁺ T cells were most predictive of resistance to conventional therapy and progression to renal failure (27). Furthermore, intrarenal T cells promote tissue injury in murine LuN and share effector pathways with kidney-infiltrating T cells in human LuN (33). Finally, the efficacy of the calcineurin-inhibitor voclosporin highlights the importance of T cells in some patients with LuN (34).

Authorship note: GC and MST contributed equally to this work.

Conflict of interest: The authors have declared that no conflict of interest exists.

Copyright: © 2025, Casella et al. This is an open access article published under the terms of the Creative Commons Attribution 4.0 International License.

Submitted: February 24, 2025; **Accepted:** August 28, 2025;

Published: September 4, 2025.

Reference information: *J Clin Invest.* 2025;135(21):e192669.

<https://doi.org/10.1172/JCI192669>.

High densities of myeloid cells were associated with progressive disease (35–37). However, these studies used simple markers of myeloid cells, while recent studies have revealed great heterogeneity in the intrarenal myeloid cell compartment (22, 38). It is unclear which in situ myeloid population is most closely linked to prognosis in LuN.

The striking responses of lupus patients to CD19 CAR T cells have renewed interest in the pathogenic role of B cells (39, 40). So far, clinical trials with CD19 CAR T cells have been small, uncontrolled, and with limited mechanistic studies. Furthermore, they appear at odds with observations that in LuN, in situ CD20⁺ B cell densities are associated with a good prognosis (27). These contradictory results could simply reflect different patient populations or the effects of conditioning chemotherapy. However, in aggregate, these observations might suggest unexplored relationships between systemic and in situ autoimmunity and/or in situ functions(s) for B cells not captured by simple total cell densities (41, 42). These functions might include secretion of highly pathogenic antibodies or subpopulations of B cells presenting antigen to large populations of pathogenic T cells (43, 44).

The above studies have each focused on the potential pathogenic role of individual cell populations. There has been no comprehensive study of the relationships between these immune cell populations and how their frequencies and spatial distributions differ between patients with LuN. The latter is likely important, as how adaptive immune cells organize in the LuN kidney provides unique prognostic information (27). Herein, we used high-dimensional confocal microscopy and customized computer vision tools to capture in situ LuN immune cell frequency and spatial distribution heterogeneity and compare these to those of renal allograft rejection (RAR). Our data suggest that in situ LuN and RAR heterogeneity can be described by quantifying a limited number of shared and unique immune cell states. Furthermore, our data suggest that each immune state develops along partially mutually exclusive trajectories, the constituents of which are associated with specific features of renal inflammation and damage.

Results

Lupus nephritis and renal allograft rejection patient biopsies. To probe the in situ immune differences between prototypical autoimmune (e.g., LuN) and alloimmune (e.g., RAR) renal diseases, we acquired 25 LuN and 23 RAR initial diagnostic biopsies. An additional 6 normal kidney control (KC) samples were acquired from nephrectomies for renal cell carcinoma. In LuN, TII is more prognostically important than GN (11, 14–19) while RAR is primarily a tubulointerstitial disease (45). Therefore, our analysis focused on comparing TII between both diseases. Conventional histological scoring by a pathologist of TII, chronicity, interstitial fibrosis, and tubular atrophy did not reveal any significant differences between the 2 disease cohorts (Figure 1A). Supplemental Figure 1A provides LuN class, patient age at diagnosis, and age at biopsy (supplemental material available online with this article; <https://doi.org/10.1172/JCI192669DS1>). Supplemental Figure 1B provides information on RAR patients, including rejection subtype, donor type, age at transplant, and age at time of biopsy. Table 1 lists available clinical and histological data on the LuN cohort.

Quantifying highly multiplex images of full biopsy renal sections. We hypothesized that there would be significant differences in in situ immune cell populations between LuN and RAR. Therefore, we stained each biopsy with antibodies specific for 42 markers that identified known in situ immune cell constituents in LuN and RAR (22, 46), other important cell populations (e.g., $\gamma\delta$ T cells), proximal tubules (CD10), distal tubules (mucin 1, MUC1), as well as markers of inflammation (Myxovirus resistance protein 1, MXA) and scarring (collagen III, COLIII) (Figure 1B).

For our imaging, we used spinning disk confocal microscopy coupled to a codetection by indexing (CODEX) microfluidic head and CODEX chemistry (47). We captured whole slide images with a pixel size of 0.1507 μm . Each cycle of imaging included a nucleus stain (DAPI) imaged at 405 nm and 4 other stains imaged at 488 nm, 561 nm, 637 nm, and 730 nm. In our analytic pipeline, we adapted ASHLAR (48) to stitch and align image tiles into full sections to generate an accurate coordinate space. Background was subtracted and the resulting full biopsy image stacks were min-max normalized to the 99th percentile for each channel.

Analysis of the resulting large and complex datasets required first detection of cells and approximation of cell boundaries (instance segmentation) and then assignment to known cell class (annotation). For nuclear segmentation of DAPI-stained renal tissue, we initially tested several convolutional neural networks (CNNs) including Mask R-CNN (49), Cellpose 2.0 (50), Deepcell (51), and Stardist (52). However, qualitative comparisons to manually annotated images revealed Cellpose to perform best, especially in highly inflamed areas (data not shown). Therefore, Cellpose 2.0 was incorporated into our analytic pipeline (Figure 1C).

To assess cell detection and segmentation performance, we compared the Cellpose output to a manually segmented dataset of 100 regions of interest (ROIs) from 5 LuN biopsies and 5 RAR biopsies (10 ROIs/biopsy). Cellpose had a zero-shot F1 performance score of ~ 0.80 (LuN) and ~ 0.58 (RAR), with an average precision of ~ 0.67 and ~ 0.38 , respectively, at an intersection over union (IoU) of 0.25 (Supplemental Figure 1, C and D). When we adopted human-in-the-loop (HITL) (53) fine-tuning, we achieved a F1 performance score of ~ 0.88 (LuN) and ~ 0.73 (RAR) with an average precision of ~ 0.77 and 0.53, respectively. Approximately 2.19 million cells were detected across all 54 samples.

The Cellpose DAPI nuclear segmentation mask was dilated approximately 1 micron with Voronoi tessellation to capture cytoplasmic staining and approximate whole cell body boundaries. From this cell body mask, we captured mean fluorescence intensity (MFI), which was standardized across all 42 channels. We used uniform manifold approximation and projection (UMAP) for dimension reduction and for plotting a random 10,000 cells from each patient cohort in 2-dimensional space (Figure 1, D and E).

The distribution of cells for each patient cohort was different. All 3 cohorts had similarities and difference in their distributions within the UMAP space. To begin to understand these differences, we projected the distributions of all 42 markers onto the UMAP space (Supplemental Figure 2, A and B). Several lymphocyte markers, including CD3, CD4, CD8, and CD20 strongly colocalize in the upper left of the UMAP plots. In contrast, some myeloid immune cell markers, such as CD14 and BDCA1, colocalize in both the upper left and right quadrants. These data suggest that infiltrating

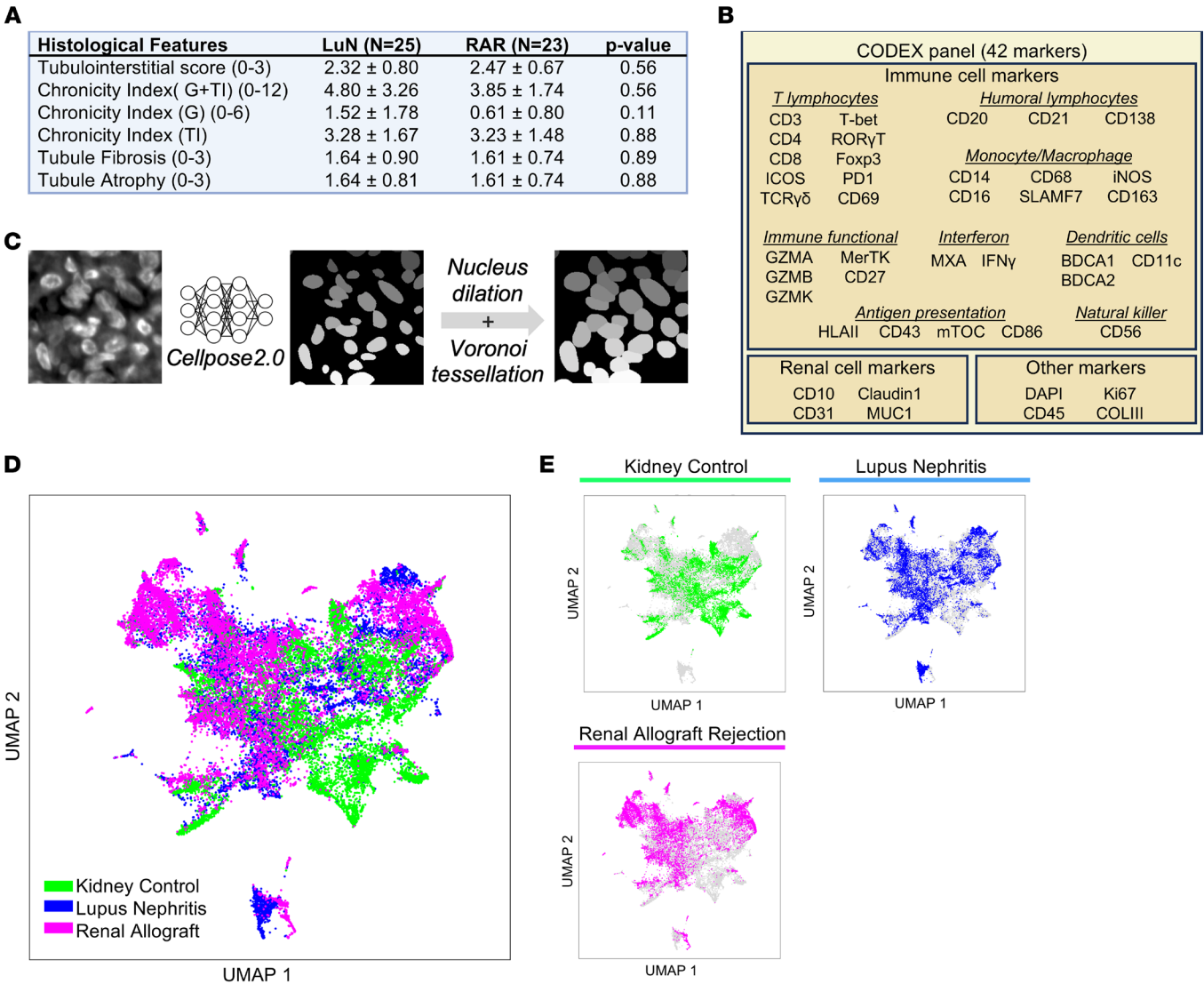


Figure 1. Cell detection in multiplex microscopy imaging of renal biopsy tissue. (A) Histological scoring of patient-paired H&E and PAS biopsies. The mean autoimmune cohort and standard deviation of the histological features are shown. Mann-Whitney-U nonparametric difference in means *P*-values are shown. (B) CODEX antibody marker panel grouped by cell lineage or cellular activity most associated with that marker. (C) Workflow diagram of cell detection and segmentation. Cell nuclei are defined using Cellpose 2.0 on a 512×512 DAPI image after performing Human-in-the-loop model retraining. Cell body was subsequently approximated by performing nuclear dilation with Voronoi tessellation. 512×512 images of DAPI segmentations are then knitted and aligned with the 42 fluorescent channels. (D) UMAP dimensional reduction of cell body MFI from 30,000 cells randomly sampled. 10,000 cells are sampled from each of the 3 cohorts: Normal Kidney Control, Lupus Nephritis, and Renal Allograft Rejection. Cells are colored according to their cohort of origin: KC, LuN, and RAR. (E) UMAP dimensional reduction of cell body MFI from 30,000 cells randomly sampled. Cells are colored according to their cohort of origin.

lymphocytes reside in the upper left quadrant while myeloid populations localize in the upper left and right quadrants.

It was apparent from Supplemental Figure 2 that FOXP3 and other markers had apparent broad distributions. Furthermore, in areas of dense inflammation, fluorescence signals can bleed into proximate cells. Finally, because we were randomly sampling a 2-dimensional cut of a 3-dimensional object, the staining intensity within a given cell class can be variable. For these reasons, identifying different cell populations by K-means clustering in the UMAP space, as is done for scRNA-Seq data, is not adequate (54).

To circumvent these limitations, we emulated the hierarchical approach used to immunophenotype cells by flow cytometry.

Briefly, the mean pixel intensities for 26 cardinal markers were organized into decision trees (DTs) for cell class annotation (47, 55) (Supplemental Figure 3). In Supplemental Figure 3A, the DT for CD45⁺ immune cells is given. Some macrophage populations expressed undetectable levels of CD45 (Supplemental Figure 2A). Therefore, CD45[−] renal structures and macrophage populations were assigned as in Supplemental Figure 3, B and C. T cell subtypes were assigned as in Supplemental Figure 3D. The resulting 33 annotated cell populations are shown in Supplemental Figure 4A. In this way, markers such as FOXP3 were not used for global assignment decisions. Rather, FOXP3 was used to identify regulatory cells within T cell populations.

Table 1. Demographics for patients with LuN

pt	Dx	Age at bx (y)	Sex	Ethnicity	LuN class	Act	Chron	Pulse steroid	Steroids at/after bx (pred eq; mg/d)	HCQ before bx	Immunosuppression before bx
1	LuN	24	F	Black	IV+V	6/24	4/12	N	40	Y	MMF
2	LuN	28	F	Black	V	0/24	N/A	Y	60	Y	0
3	LuN	57	F	Black	III	5/24	1/12	N	20	Y	0
4	LuN	63	F	Black	III	0/24	3/12	N	60	Y	0
5	LuN	19	F	White	IV+V	10/24	2/12	Y	40	Y	MMF
6	LuN	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
7	LuN	28	F	Asian	V	2/24	7/12	N	0	Y	0
8	LuN	44	F	N/A	IV	7/24	5/12	Y	40	Y	MMF
9	LuN	40	F	N/A	VI	0/24	7/12	N/A	N/A	N/A	N/A
10	LuN	17	F	Black	II	0/24	0/12	Y	40	Y	0
11	LuN	17	F	Hispanic	IV	13/24	0/12	Y	20	N	0
12	LuN	16	F	Black	II	0/24	0/12	Y	50	N	0
13	LuN	17	F	Hispanic	III+V	3/24	4/12	N	0	Y	MMF
14	LuN	19	F	Hispanic	IV	12/24	6/12	N	0	Y	0
15	LuN	20	F	Black	III	6/24	2/12	Y	0	Y	0
16	LuN	17	M	N/A	IV	13/24	0/12	N/A	N/A	N/A	N/A
17	LuN	31	F	Hispanic	IV	11/24	2/12	N/A	N/A	N/A	N/A
18	LuN	18	F	N/A	V	1/24	0/12	N/A	N/A	N/A	N/A
19	LuN	54	F	Black/AA	V	8/24	4/12	Y	20	N	0
20	LuN	17	F	Hispanic	IV	10/24	0/12	N/A	N/A	N/A	N/A
21	LuN	34	F	N/A	V	2/24	4/12	N/A	N/A	N/A	N/A
22	LuN	18	M	Hispanic	IV	19/24	2/12	N/A	N/A	N/A	N/A
23	LuN	10	F	Black	IV	9/24	2/12	Y	60	Y	AZA
24	LuN	33	F	Black	IV	8/24	4/12	N	50	Y	0
25	LuN	15	F	Hispanic	IV+V	15/24	3/12	N/A	N/A	N/A	N/A

N/A, not available, pt, patient; Dx, diagnosis; Act, NIK activity index; Chron, chronicity index; bx, biopsy; pred eq, equivalent dose of prednisone; HCQ, hydroxychloroquine; MMF, mycophenolate; AZA, azathioprine. Pulse steroid, IV (intravenous) solumedrol. Dose is unknown although typically 1 gram a day for three days.

Approximately 77% of segmented cells were assigned to a cell class across disease groups. Primarily, renal tubular cells could not be assigned to a class, as identifying markers often stained the plasma membrane, which was beyond the 1-micron DAPI dilation in these large cells. This limitation was circumvented as described below.

Each annotated cell class expressed the expected cardinal markers (Supplemental Figure 4B). Furthermore, across biopsy cohorts, each annotated class manifested similar distributions of cell staining (Supplemental Figure 4C). These data suggest that quantitative comparisons could be made across patient cohorts.

Using the above DTs to annotate cells in the UMAP space, similarities and differences between biopsy cohorts became apparent (Figure 2A). While renal parenchymal structures occupied the center of all 3 cohort UMAPs, there were differences in distributions. Some of these changes could be ascribed differential expression of markers of inflammation (MXA, Claudin 1) and scarring (COL1) in the disease cohorts (Supplemental Figure 4D). Lymphocytes, especially T cells, occupied the upper left quadrant. These populations were scant in normal kidney and highest in RAR. In contrast, myeloid populations were increased in both LuN and RAR. However, there were some differences in the distribution of myeloid cells in the UMAP space, suggesting that each disease might be enriched for specific populations.

A more detailed picture was provided when constituent cell percentages were examined (Supplemental Figure 5A). As expected,

renal parenchymal cells were the most prevalent. As described above, these were under estimations. Among immune cell populations, myeloid cells, especially CD14⁺ macrophages and inflammatory monocytes, were enriched in both LuN and RAR. However, total CD163⁺ macrophages were enriched in LuN. In contrast, most T cell populations, including CD4⁺, CD8⁺, and $\gamma\delta$ T cells, were enriched in RAR. Indeed, comparing cell densities between KC, LuN, and RAR revealed that LuN was enriched for myeloid cell populations, especially CD163⁺ macrophages (Figure 2B). In contrast, RAR was enriched for multiple T cell populations. Notably, CD3⁺ T cells lacking CD4, CD8, or γ coexpression (double negative or DN T cells), were not enriched in either disease compared with KC samples (27, 56). These data suggest that, in our cohort, RAR is characterized by enrichment of T cell and myeloid populations, while LuN is enriched primarily in myeloid cell populations.

In comparison with KC biopsies, there was substantial heterogeneity in disease groups both in terms of lymphocyte versus myeloid cell populations in each biopsy and the constituents within each broad immune cell class (Figure 2C). In LuN, there were 4–5 biopsies with relatively high lymphocyte populations, while, in the rest of the biopsies, myeloid cells predominated. Myeloid cells predominated in about half of RAR biopsies. Furthermore, especially in RAR, the myeloid compartment was dominated by either macrophages or HLA class II positive or negative inflam-

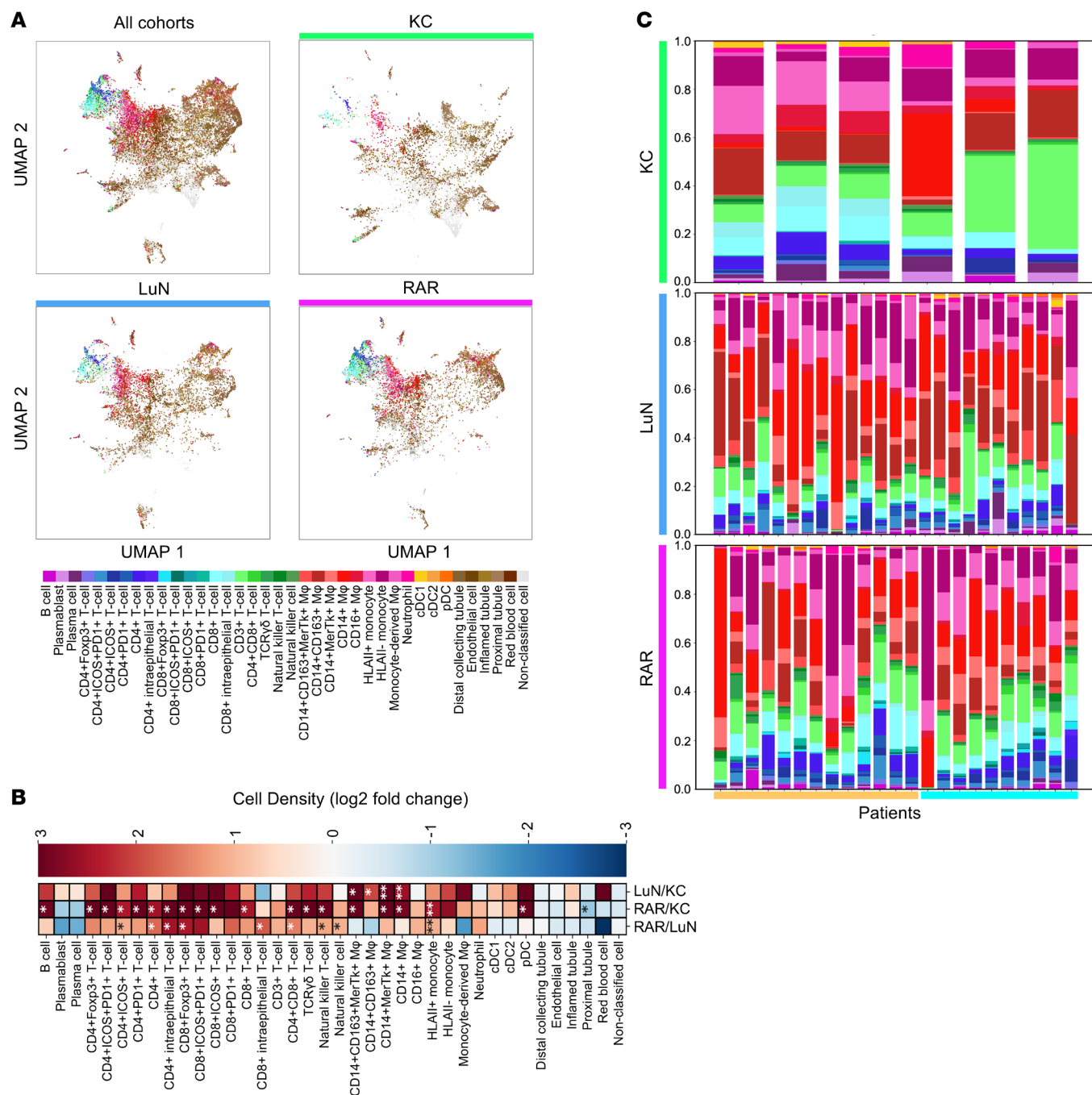


Figure 2. Distribution of immune cell classes in KC, LuN, and RAR. (A) UMAP dimensional reduction of cell body MFI from 30,000 cells randomly sampled. 10,000 cells were sampled from each of the indicated cohorts. (B) Nonparametric Mann-Whitney-U difference of the mean test for population differences in classified cell density between patient cohorts; from the top row: LuN-KC, RAR-KC, and RAR-LuN. Color indicates log₂ fold change. Benjamini-Hochberg *P* value correction was performed. **P* < 0.05, ***P* < 0.01, ****P* < 0.001. (C) Patient-level proportions of the 33 immune cell classes by cohort. Kidney control (top), lupus nephritis (middle), and renal allograft rejection (bottom). Mixed-rejection biopsies are denoted using a beige bar, T-cell mediated rejection biopsies are denoted in light blue. All cell classes (except nonclassified cells) are color coded as displayed in the panel in A. Bars at bottom identify MR (yellow) and TCMR (blue).

mature monocytes (defined by CD14 and CD16 expression) (57). Among the RAR samples were 13 mixed rejection (MR) and 10 T cell mediated rejection (TCMR). However, there were no apparent global differences in the distributions of lymphoid or myeloid cells between these disease classes. As these subgroups were small, no further comparisons between the 2 were performed.

Distinct immune trajectories in LuN and RAR. Plotting Spearman's correlation between immune cell proportions across LuN and RAR biopsies revealed covariance between specific subpopulations. (Figure 3, A and B, and Supplemental Figure 6, A and B). In both diseases, covariance across T cell populations, including CD4⁺, CD8⁺, and $\gamma\delta$ T cells, was the most striking. Interestingly, there was not a

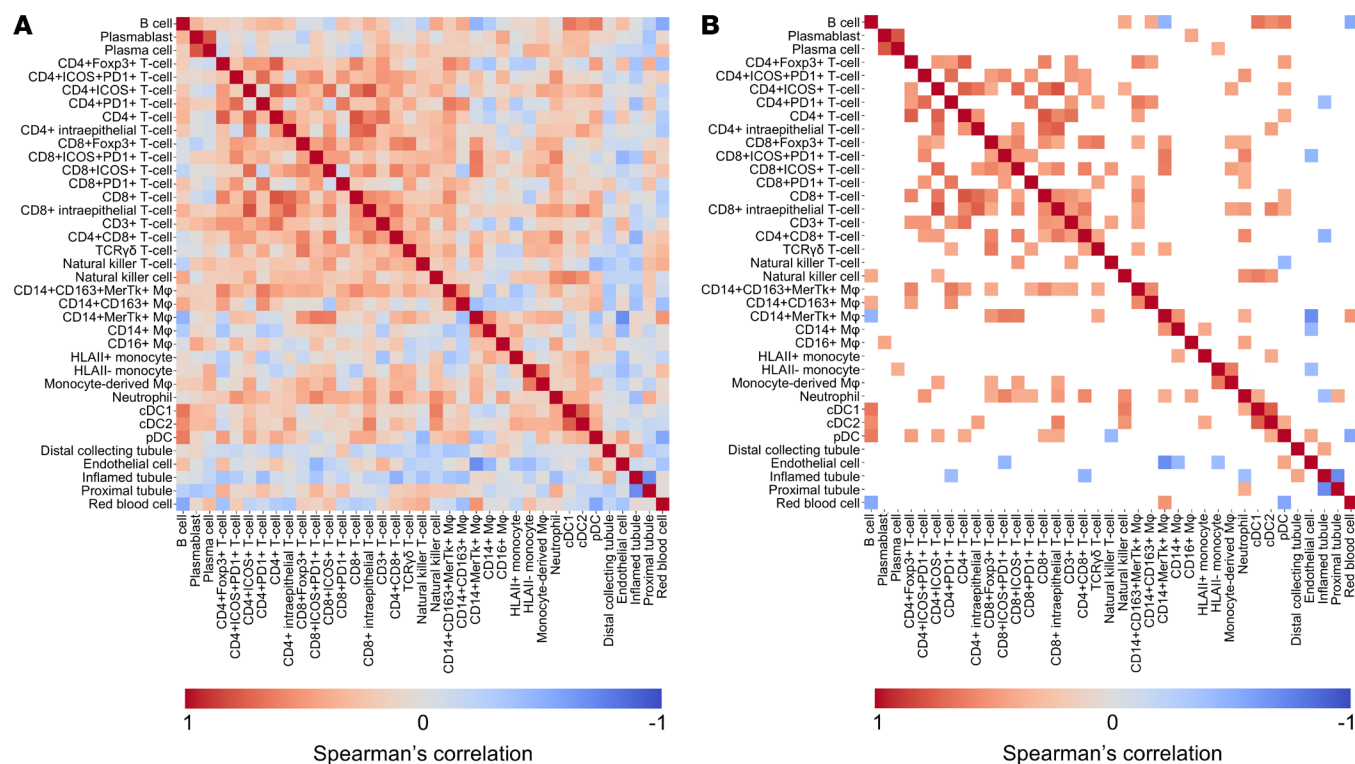


Figure 3. Spearman's correlation of lupus nephritis cell class proportions. (A) Heatmap of the nonparametric spearman's correlations between patient-level immune and nonimmune cell proportions for lupus nephritis patients. (B) Heatmap of statistically significant ($P < 0.05$) nonparametric spearman's correlations between patient-level immune and nonimmune cell proportions for lupus nephritis patients.

strong correlation between B cells and plasma cells or plasmablasts, and in neither disease was there a correlation between plasma cells/blasts and most T cell populations. Focusing on innate cells, there were covariant subpopulations or blocks among the most common cell populations in both LuN and RAR. Specifically, we observed covariance between macrophages expressing CD163, including CD14⁺CD163⁺MerTk⁺ and CD14⁺CD163⁺ populations. There were separate covariant cell blocks for CD163⁺ CD14⁺ macrophages ($\pm$ MerTk) and for CD14⁺CD16⁺ monocytes ($\pm$ HLA class II).

In RAR, some of the above populations had inverse relationships. For example, CD163⁺ macrophages inversely varied with some CD163⁺ macrophage and monocyte populations. In LuN, there was a trend towards these inverse relationships. In LuN, B cell densities have been associated with a good prognosis (27). Therefore, it is interesting that, in LuN, they had an inverse proportional relationship with CD14⁺MerTk⁺ macrophages and a positive correlation with some CD163⁺ macrophage populations, both of which are much more prevalent than B cells. Similar correlations were observed when considering the densities of the different cell populations (Supplemental Figure 6, C–F), suggesting that the observed relationships were robust. These separate, covariant subpopulations, or cell blocks, suggest that a limited number of discrete immune states characterize both LuN and RAR.

Cohort-specific relationships were observed when we plotted the densities of all myeloid cells versus all T cells for each biopsy (Figure 4A). The size of each point reflects total humoral cells. Strikingly, LuN unfolded along a myeloid axis and RAR along a T cell axis with some RAR biopsies also having substantial densities of myeloid cells.

These differences in myeloid and T cell densities were associated with striking visual differences (Figure 4B). We then plotted the densities of CD14⁺MerTk⁺CD163⁺ macrophages versus CD14⁺CD163⁺ macrophages versus total CD8⁺ T cells (Figure 4C). Interestingly, RAR unfolds along the CD14⁺MerTk⁺CD163⁺ axis, which is also rich in CD8⁺ T cells. In contrast, some LuN biopsies also had relatively high CD163⁺ macrophage densities (4 of 25 biopsies) while other biopsies had high densities of CD163⁺ macrophages (5 of 25). Examples from the indicated biopsies are provided in Figure 4D. These data suggest that in our patient cohorts, LuN and RAR biopsies often lie along different immune cell trajectories.

Immune cell localization within renal compartments. The above analyses examined immune cell frequencies across whole biopsies. However, the kidney is structurally complex with glomeruli, tubules, and the tubulointerstitial space. Compared with hematopoietic cells, tubules have a characteristic pattern of DAPI staining. Therefore, to identify tubules, Omnipose (58) was trained on representative 10-fold downsized DAPI kidney images from CODEX image stacks. Glomeruli were segmented manually (Supplemental Figure 7, A and B).

We first assessed the densities of all immune cell populations across the different renal compartments (Figure 5A). From left to right are provided interstitial, tubular, glomerular, peritubular (dilated tubular mask), and periglomerular (dilated glomerular mask) densities for all indicated immune cell populations. In both diseases, and for all immune cell populations, densities were highest in the periglomerular space. In general, there were more plasma cells/blasts in LuN distributed across all compartments. In

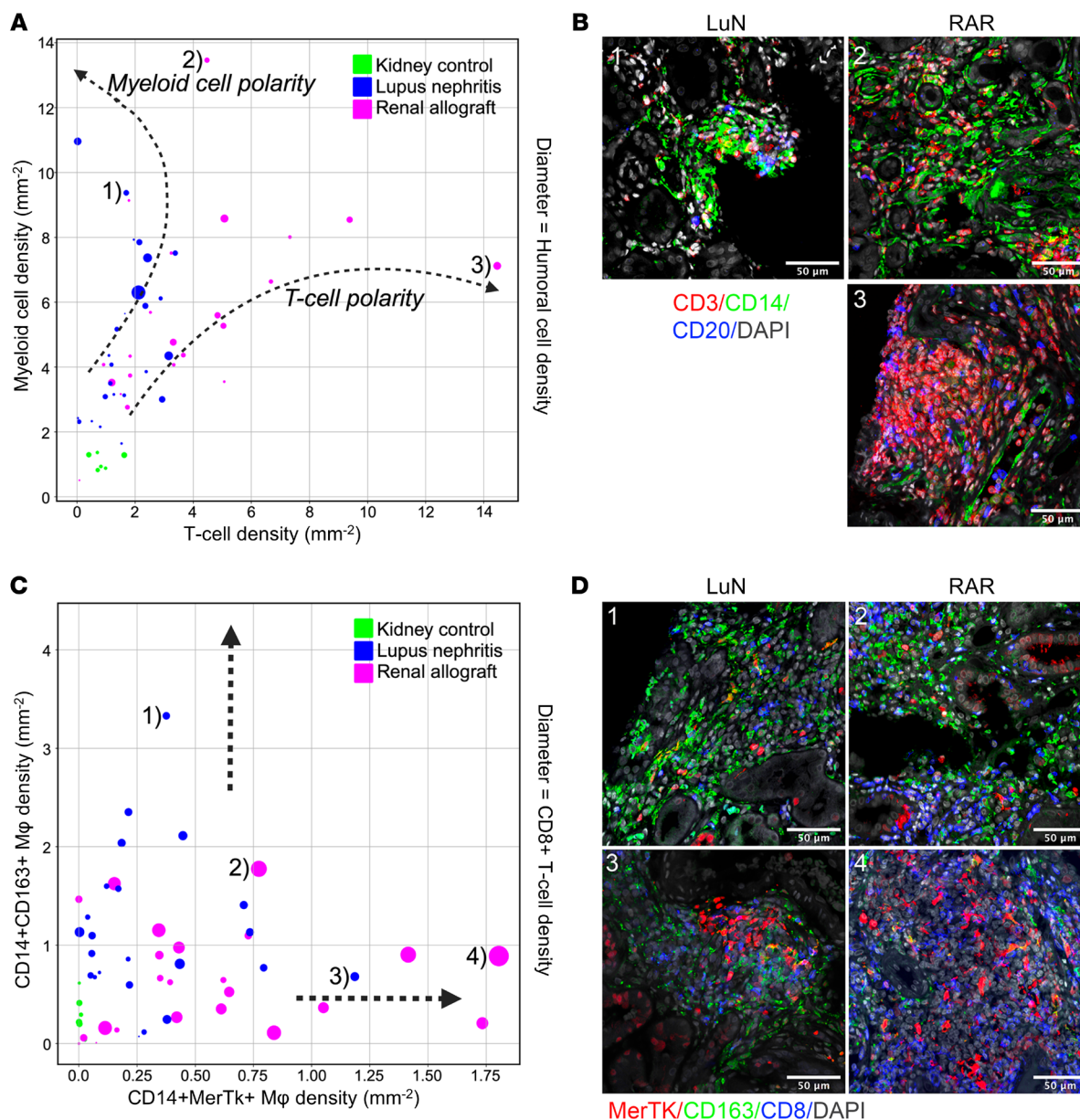


Figure 4. Distinct immune trajectories are associated with distinct pathologic states. (A) Plot of the patient-level T-cell density (x-axis) and myeloid cell density (y-axis) colored by cohort: Kidney control (green), LuN (blue), RAR (magenta). Diameter indicates humoral cell density (B cells plus plasma cells). (B) Representative microscopy images of immune cell polarization in lupus nephritis and renal allograft rejection patient cohorts. Image numbers correspond to biopsies indicated in A. (C) Plot of the patient-level CD14⁺MerTk⁺ macrophage density (x-axis) and CD14⁺CD163⁺ macrophage density (y-axis) colored by cohort: Kidney control (green), LuN (blue), RAR (magenta). Diameter indicates T-cell density. (D) Representative microscopy images of CD14⁺CD163⁺ and CD14⁺MerTk⁺ enriched biopsies. Image numbers correspond to biopsies indicated in C. Scale bars: 50 μ m.

contrast, RAR was generally enriched in T cells distributed across all renal structures except for glomeruli. Furthermore, there was a general enrichment of CD163⁺ macrophages in LuN.

We then focused on 2 potential disease-associated populations, CD8⁺ T cells and CD14⁺MerTk⁺ macrophages (Figure 5B). In both cell populations, the periglomerular enrichment was evident. In addition, in some RAR biopsies, there is an enrichment of both cell populations in the peritubular and tubular space compared with the interstitial space. Indeed, in RAR biopsies there was both peritubular inflammation and tubulitis (Figure 5C). In contrast, interstitial and peritubular densities were similar in LuN and tubulitis was rare

(Figure 5D and data not shown). These data suggest that inflammation in the RAR tubulointerstitium is centered around and in tubules, while, in LuN, it occurs diffusely through the interstitium.

Organization of inflammation into neighborhoods. We next used DBSCAN (59) to determine if in situ inflammation was organized in LuN and RAR. K-means clustering and bootstrapping were first used to estimate the optimal number of states using spatial coordinates for all immune cells and the indicated cell neighborhood size exclusion conditions (Supplemental Figure 8). Based on this analysis, we used 8 as the optimal number of clusters (neighborhoods) for downstream analysis.

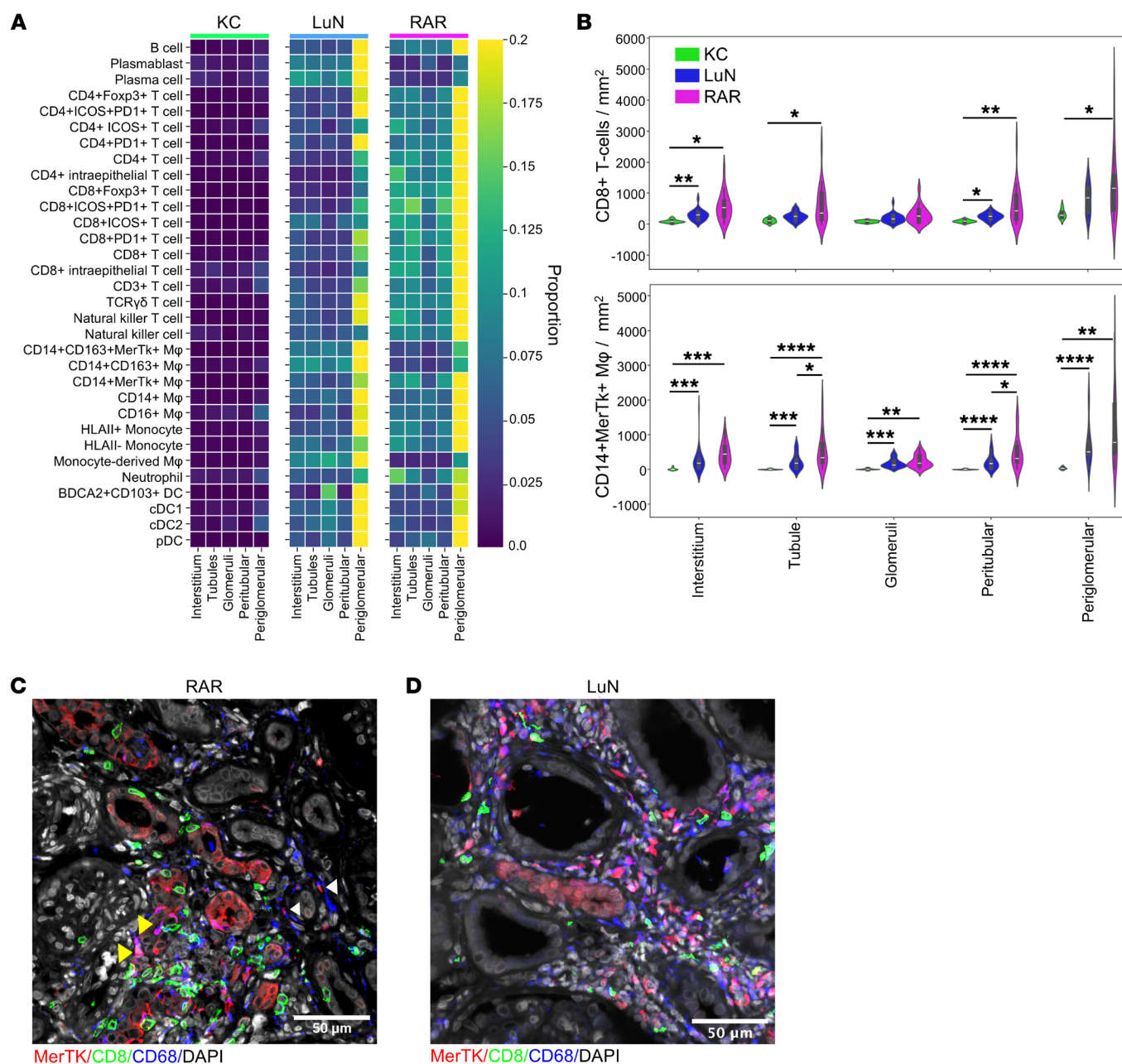


Figure 5. Distribution of inflammation across renal compartments. (A) Heatmap showing the biopsy average density of immune cells of the kidney anatomical compartments, normalized to proportion across cohorts. From the left: interstitium, tubule, glomeruli, peritubular, periglomerular. (Left heatmap) kidney control, (middle heatmap) lupus nephritis, and (right heatmap) renal allograft. (B) Violin plots of the biopsy average number of cells per mm² within the indicated tissue compartments for CD8⁺ T cells and CD14⁺MerTk⁺ macrophages (Mφ). Mann-Whitney U test with Bonferroni correction for multiple tests. **P* < 0.05, ***P* < 0.01, ****P* < 0.001. (C) RAR example demonstrating tubulitis (yellow arrowhead) and peritubular inflammation (white arrowhead). (D) LuN example demonstrating diffuse interstitial inflammation. Scale bars: 50 μm.

We generated a heatmap using features including average total cell count across biopsies and average cell proportion to visualize the unique elements of each cluster (Figure 6A). When cell counts were examined, Clusters 4 and 7 had the most cells, while Clusters 2 and 3 were the most frequent (Figure 6B). Each cell cluster had unique features. Cluster 0 was enriched in HLA class II monocytes and monocyte-derived macrophages, Cluster 1 in CD163⁺MERTK⁺ macrophages, Cluster 2 in MHC class II⁺ monocytes and CD16⁺ macrophages, Cluster 3 in HLA class II⁺ monocytes,

Cluster 4 in T cell populations, Cluster 5 in CD14⁺ macrophages, Cluster 6 in CD163⁺ macrophages, and Cluster 7 in T cell populations including CD4⁺CD8⁺ T cells (27). While most clusters appear more frequent in RAR, the CD163⁺ macrophage Cluster 6 was more common in LuN. Representative examples of clusters 1 and 6 are shown in 1 LuN biopsy, indicating that clusters of CD163⁺ and CD163⁺ macrophages can occur in the same biopsy (Figure 6C). An example in LuN of the inflammatory monocyte and macrophage enriched Cluster 2 is provided in Figure 6D.

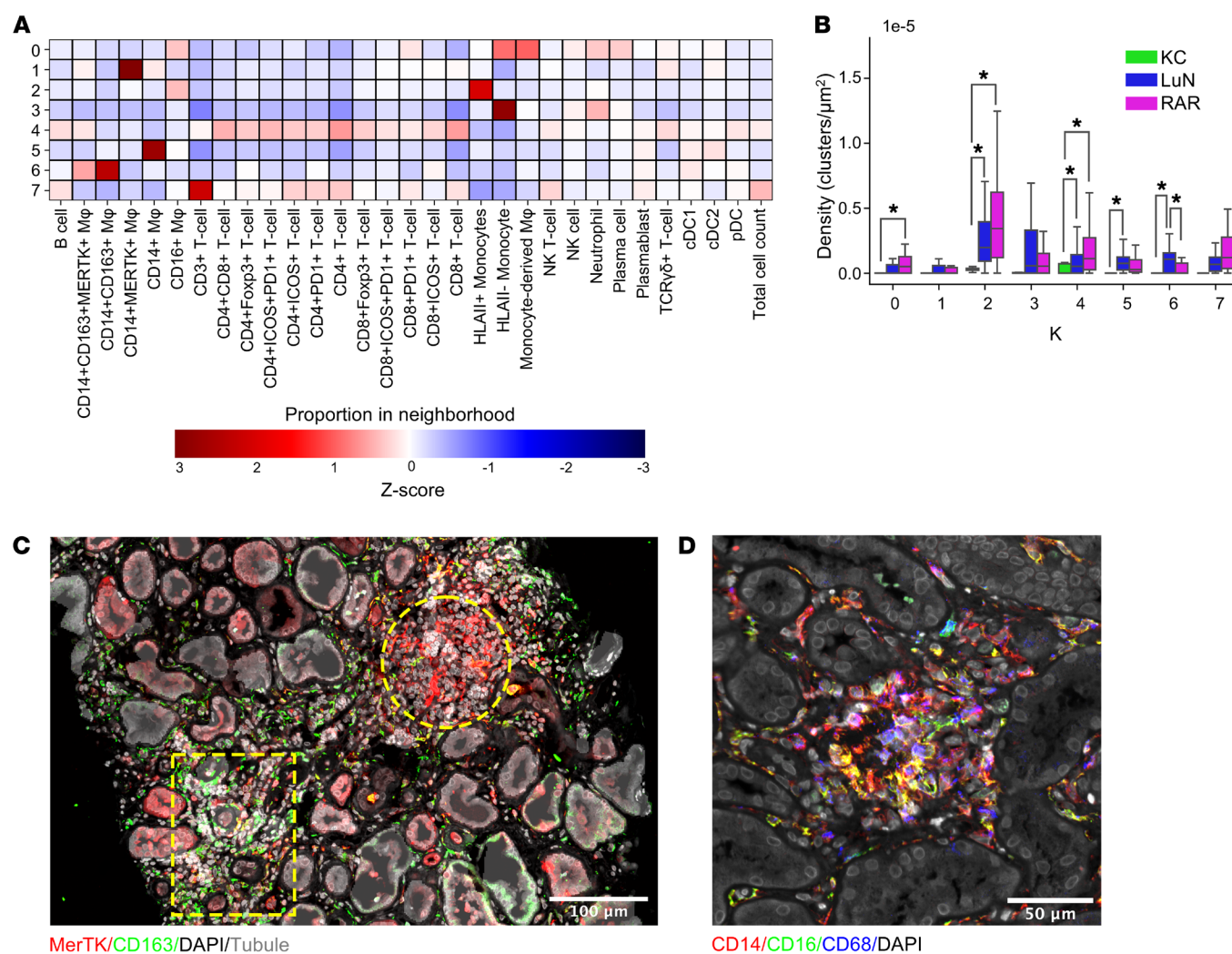


Figure 6. Distribution of cell classes in situ immune clusters. (A) Heatmap of the Z-score from indicated extracted features from DBSCAN cell networks using an optimal K = 8 means clustering of cell clusters larger than 19 cells. Proportions of each cell class were used with the number of cells in each cluster represented in right column. (B) Bar plot showing the total count of the various K=8 DBSCAN cell networks colored by patient cohort: KC (green), LuN (blue), and RAR (magenta). (C) Example of clusters 6 (box) and 1 (circle) within the same LuN biopsy. (D) Example of cluster 2 in a LuN biopsy containing a mix of MHC class II⁺ (data not shown) inflammatory monocytes and macrophages. Scale bars: 100 μm (C), 50 μm (D).

We next sought to confirm our results on a subset of 6 biopsies from our LuN cohort using spatial transcriptomics (MERSCOPE, Vizgen) and the 500 gene human immune-oncology panel. The resulting raw images were resized with a pixel size of 10 nm and DBSCAN performed on *CD3E*, *CD14*, and *FCGR3A* (*CD16*)–expressing channels. Images with DBSCAN labels were further processed to generate neighborhoods for gene density. The Z-scored densities of 20 genes whose protein products were represented in the CODEX panel were used for K-means clustering as above (optimal was 7 clusters, data not shown). Using UMAP to plot the distribution of cell clusters revealed good separation of cluster groups (Supplemental Figure 9A). A Heatmap of the expression distributions of those genes used for clustering revealed similar clusters to those detected by CODEX (Supplemental Figure 9B). Indeed, there were CD163⁺ macrophage clusters (Cluster 3), CD14⁺ enriched monocyte clusters (Cluster 4), CD16⁺ macrophage/monocyte enriched clusters (Cluster

0), and 2 T cell enriched clusters (Clusters 1 and 2). Cluster 6 had similarities to CODEX Cluster 7, while Cluster 5 was not enriched for any specific immune genes. In Supplemental Figure 9C are the top 10 differentially expressed genes in each cluster (55 genes total), demonstrating unique gene expression distributions in each cluster.

We also performed gene set enrichment analysis (GSEA) comparing each cluster to all others (Supplemental Figure 9D). Notably, the CD163⁺ Cluster 3 was enriched in fibroblast proliferation and mitotic gene programs. In contrast, the CD14⁺ monocyte Cluster 4 revealed endothelial cell proliferation, MAPK activation, and receptor tyrosine kinase activation programs. There was a concomitant lack of adaptive immune system expression programs. These results indicate that in both LuN and RAR, renal inflammation is organized into distinct clusters. Furthermore, our spatial transcriptomic data suggest that specific injury mechanisms occur in some immune cell clusters.

Immune cell correlates with tissue inflammation and damage. From our high-dimensional staining panel, we could derive 4 measures of tubulointerstitial inflammation and scarring: (a) total biopsy immune cell density; (b) total biopsy inflamed tubule cell density (MXA and Claudin 1 in tubule mask); (c) COLIII mask area (% of total biopsy area); and (d) MXA mask tissue area (% of total biopsy area). Given the above findings, we plotted these 4 measures versus whole biopsy densities and proportions of total CD8⁺ T cells, total CD4⁺ T cells, CD14⁺MerTK⁺CD163⁺ macrophages, CD14⁺CD163⁺ macrophages, HLA class II⁺ inflammatory monocytes, and HLA class II⁻ inflammatory monocytes, and then performed ordinary least squares linear regression (OLS). Only graphs with positive correlations are provided.

We observed that CD8⁺ T cell densities were correlated with immune cell densities in RAR and LuN (Figure 7A). In contrast, only in RAR were CD8⁺ T cell densities associated with MXA expression (Figure 7B). Consistent with the covariance of CD4⁺ and CD8⁺ T cell densities across biopsies, similar associations were observed for CD4⁺ T cell densities (Figure 7, C and D). CD14⁺MerTK⁺CD163⁺ macrophage cell densities were associated with immune densities in both diseases (Figure 7E). CD14⁺MerTK⁺ macrophages were also associated with inflamed tubule density in RAR with a similar trend in LuN (Figure 7F). There was a strong association between CD163⁺ macrophage cell densities and immune cell densities in LuN but not RAR (Figure 7G). However, CD163⁺ macrophage proportions were associated with COL III scores only in RAR (Figure 7H). Finally, densities of HLA class II⁺ inflammatory monocytes were associated with immune cell densities (Figure 7I) in both diseases. However, they were only associated with MXA scores in RAR (Figure 7J). These data indicate that, in LuN and RAR, immune cell populations having the same surface phenotype can have either similar or different associations with measures of renal inflammation and scarring.

Immune cell trajectories characterize individual biopsies. The above data suggest that inflammation heterogeneity can be resolved into relatively few covariant blocks of the most prevalent immune cells. Immune cell constituents of these different inflammatory states organized into distinct niches and were associated with specific manifestations of renal inflammation and scarring. To begin to graphically quantify inflammation in a way that could be compared across diseases and biopsies, we generated radar diagrams in which each axis was the density of a principal cell population: total CD8⁺ T cells, total CD4⁺ T cells, CD14⁺CD163⁺ macrophages, CD163⁺ macrophages, HLA class II⁺ inflammatory monocytes, and HLA class II⁻ inflammatory monocytes.

We first plotted all 54 individual biopsies on a single radar graph, color coded by clinical cohort (Figure 8A). Compared with both LuN and RAR, the KC samples had far less densities of these 6 immune cell population groups. It is also apparent that some RAR biopsies had higher densities of several immune cell populations compared with LuN biopsies. Indeed, plotting average immune cell densities for each clinical cohort indicate that all populations, except for CD163⁺ macrophages, are higher in RAR than LuN (Figure 8B). These data indicate that, on average, the RAR biopsies are more inflamed, even though the 2 disease cohorts were scored similarly for TII (tubulointerstitial score) by a renal pathologist (Figure 1A).

Plotting LuN and RAR biopsy cohorts separately allowed better visualization of the immune cell densities that characterized each disease cohort. In LuN, inflammation unfolds along 3 major density axes: CD163⁺ macrophages, CD163⁻ macrophages, and HLA Class II⁻ inflammatory monocytes (Figure 8, C and D). Across the disease cohort, RAR appears more complex with significant projections along the CD163⁻ macrophage, HLA Class II⁺ inflammatory monocytes, HLA Class II⁻ inflammatory monocytes, CD4⁺ T cells, and CD8⁺ T cell density axes (Figure 8, E and F). Notable is the lack of substantial CD163⁺ macrophage densities. These data suggest that LuN is characterized by a few myeloid immune cell density axes, while RAR has both myeloid and adaptive immune cell density axes.

We then plotted individual disease biopsies. Of the 25 LuN biopsies, 6 were characterized by prominent CD163⁺ macrophage densities (> 2-fold over CD163⁻ macrophages) and a relative lack of other immune cell populations (Figure 8G and Supplemental Figure 10A). One of these biopsies also had high numbers of HLA class II⁻ inflammatory monocytes (Supplemental Figure 10B). Similarly, 6 biopsies were characterized by primarily CD163⁻ macrophage populations (Figure 8H and Supplemental Figure 10C). Two of these also had substantial HLA class II⁻ monocyte immune cell densities (Figure 8I and Supplemental Figure 10D). The remaining 13 biopsies had both substantial CD163⁺ and CD163⁻ immune cell densities. Seven of these biopsies did not have substantial densities of other populations (Figure 8J and Supplemental Figure 10E). The other 6 had concurrent HLA class II⁻ inflammatory monocyte densities (Figure 8K and Supplemental Figure 10F). Therefore, within our cohort, LuN fell into subsets characterized by CD163⁺ macrophage, CD163⁻ macrophage, and HLA class II⁻ inflammatory monocyte densities.

RAR was more complex. Of the 23 biopsies, 8 had a predominantly single immune cell population (> 2-fold difference over other immune cell populations). In 5 of these, CD163⁻ macrophage densities predominated (Figure 8L and Supplemental Figure 11A). In 2, HLA class II⁻ inflammatory monocytes predominated, and, in 1, CD163⁺ macrophages. Only 2 biopsies were characterized by 2 trajectories, both of which included CD163⁻ macrophages (Supplemental Figure 11B). The rest manifested multiple myeloid and adaptive immune cell densities including 6 that had HLA class II⁻ inflammatory monocytes (Figure 8M and Supplemental Figure 11C) and 7 that did not (Figure 8N and Supplemental Figure 11D). These data suggest that, in our cohort, RAR falls broadly into 2 categories, those that can be characterized by myeloid cells, most commonly CD163⁻ macrophages, and those characterized by both myeloid and adaptive immune cell densities.

Discussion

Using high dimensional imaging and computer vision techniques specifically adapted for the kidney, we provide the first comprehensive assessment of LuN and RAR immune cell constituency, how these cells are organized into neighborhoods, and their relationships to renal cortical structures. These data resolve immune cell heterogeneity into a limited number of cell states. Indeed, in any 1 biopsy, the inflammatory state could be characterized by the relative prevalence and magnitude of cardinal immune cell trajectories. Notably, in many biopsies from both diseases, CD163⁻ macrophages predominated with or without other concurrent immune cell populations. In LuN, concurrent immune cell populations

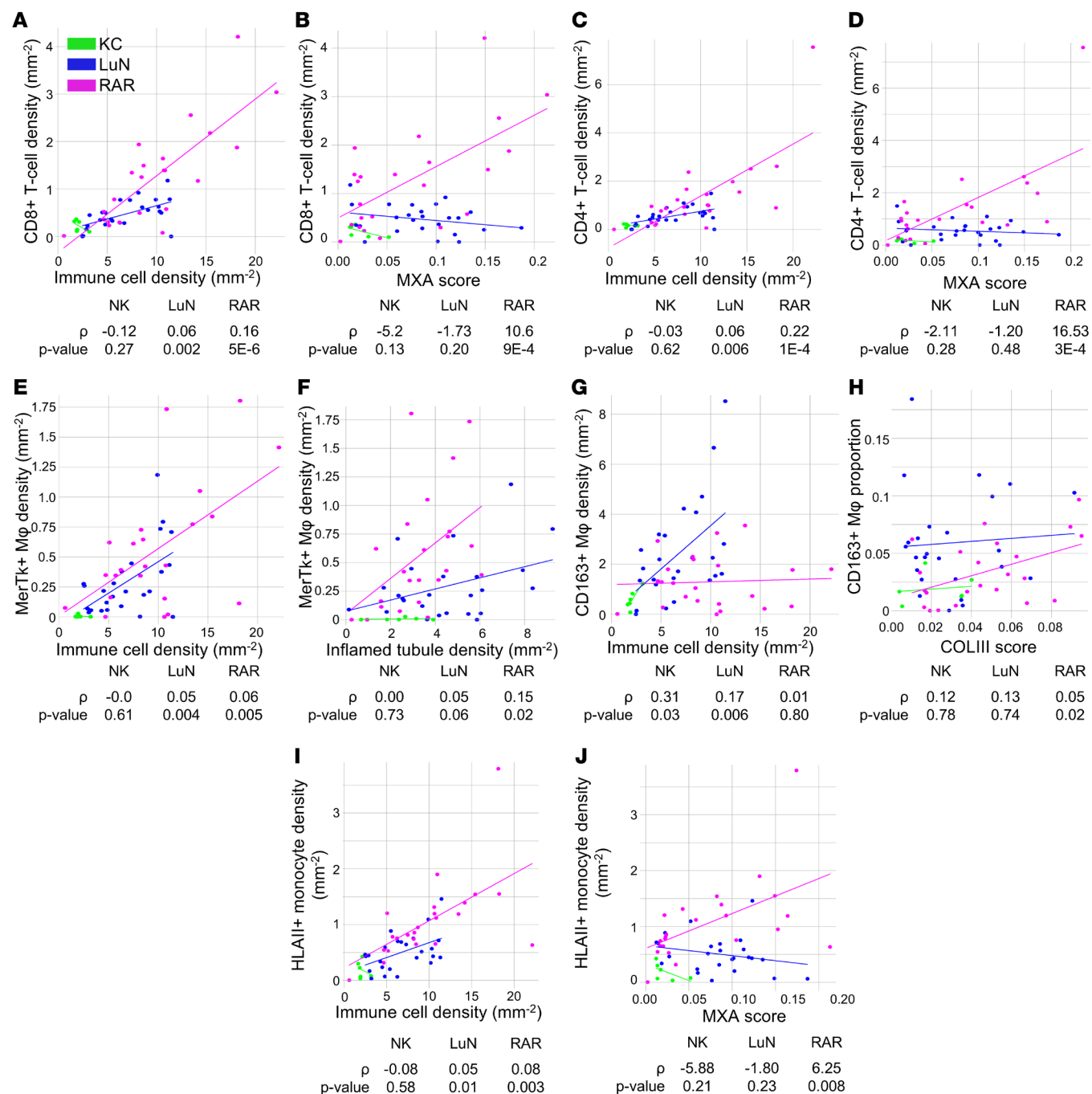


Figure 7. Immune cell trajectories correlate with tissue inflammation and damage. (A and B) CD8⁺ T cell density as a function of immune cell density (A) and score (B). (C and D) CD4⁺ T cell density as a function of immune cell density (C) and myxovirus resistance protein 1 (MXA) score (D). (E and F) MerT-K⁺CD163⁺ macrophage density as a function of immune cell density (E) and inflamed tubule density (F). (G) CD163⁺ macrophage density as a function of immune cell density. (H) CD163⁺ macrophage proportion as a function of Collagen III (COL III) score. (I and J) Inflammatory HLA class II⁺ monocytes immune cell density as a function of immune cell density (I) and MXA score (J).

were of myeloid lineages, while, in RAR, CD163⁺ macrophages cooccurred with both myeloid and adaptive immune cell lineages. While both diseases manifested inflammatory states anchored by CD163⁺ macrophages, only in LuN were there several biopsies in which CD163⁺ macrophages predominated. Our studies suggest an approach to quantifying in situ immunity that allows comparisons between individual biopsies in and across disease states.

Canonical cell markers identified populations of immune cells that covaried with each other and shared common spatial distributions within the tubulointerstitium. CD163⁺, a marker of some M2 macrophages, identified cell populations that formed neighborhoods and distributions different from other macrophage populations regardless of other shared markers such as the phagocytic receptor MerTK (60). Likewise, inflammatory monocytes, especially those expressing MHC class II, did not

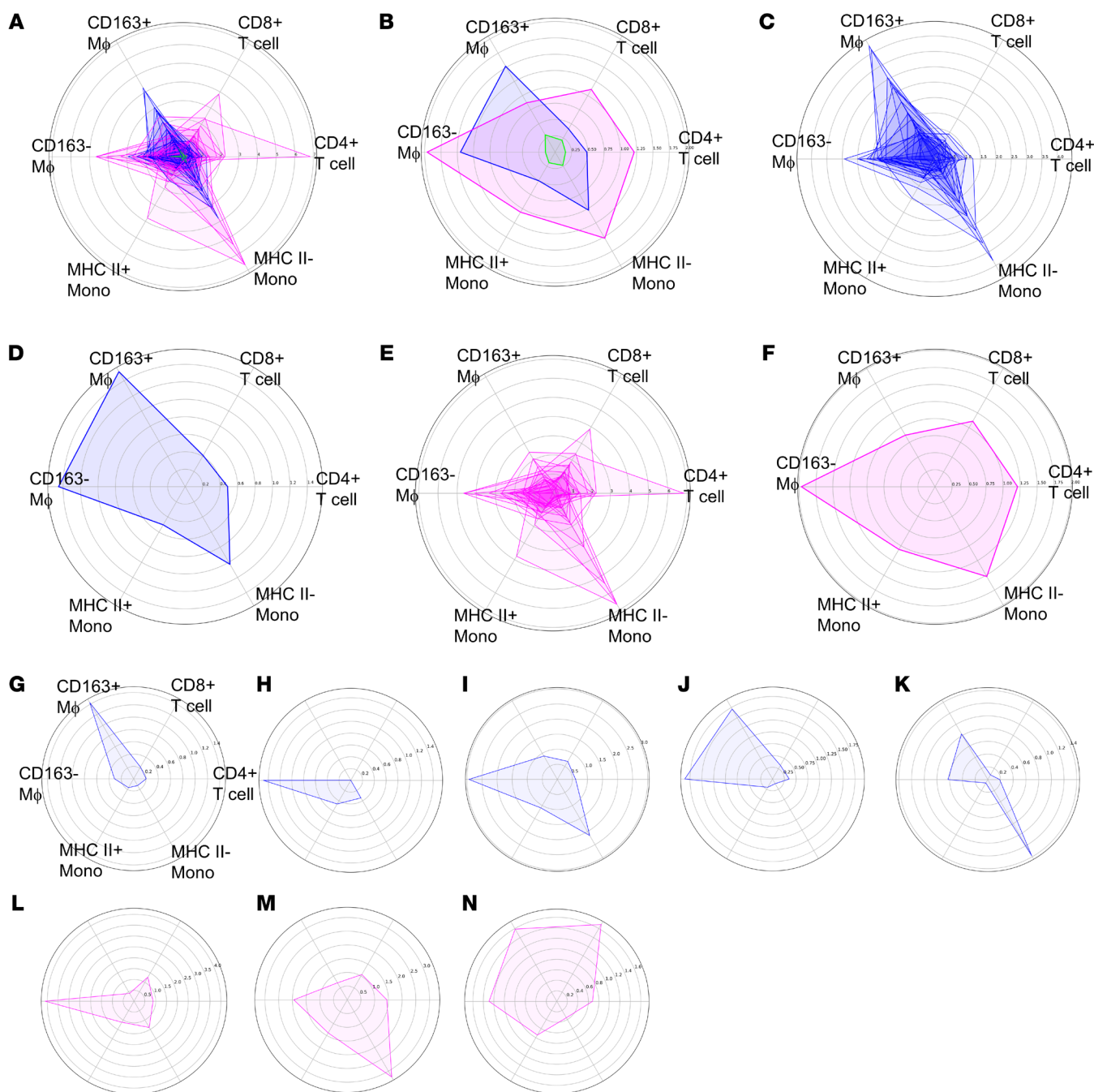


Figure 8. Quantifying in situ immune state using principal immune cell trajectories. (A) Composite web graph of all individual biopsies colored by cohort. Green, normal kidney; blue, LuN; and purple, RAR. (B) Plot of averages for each clinical cohort. (C) Plot of LuN individual biopsies. (D) Average LuN densities for indicated immune cell populations. (E) Plot of RAR individual biopsies. (F) Average RAR densities for indicated immune cell populations. (G) Example of individual LuN biopsy with CD163⁺ macrophage polarity. (H) Example of LuN biopsy with CD163⁻ macrophage polarity. (I) Example of LuN biopsy with CD163⁻ macrophage and HLA II inflammatory monocyte polarities. (J) Example of LuN biopsy with CD163⁺ and CD163⁻ macrophage polarity. (K) Example of LuN biopsy with CD163⁺ and CD163⁻ macrophage polarity in combination with HLA II⁻ inflammatory monocyte polarity. (L) Example of individual RAR biopsy with CD163⁻ macrophage polarity. (M) Example of individual RAR biopsy with multiple immune cell lineages plus HLA II⁻ inflammatory monocytes trajectory. (N) Example of individual RAR biopsy with multiple immune cell lineages without HLA II⁻ inflammatory monocytes trajectory.

covary with any other immune cell blocks/subgroups and formed distinct neighborhoods (57). Our current staining panel did not allow fine categorization of either macrophage or monocyte subsets.

Within the limitations of these small patient cohorts, both CD4⁺ and CD8⁺ T cells largely behaved as a covariant block and

cosegregated into specific neighborhoods. Remarkably, these fundamental relationships were apparent in both LuN and RAR. Rather, these cell subgroups and niches primarily varied in relative prevalence between the 2 diseases. T cell blocks were a dominant feature of RAR, while, in LuN, macrophage subgroups were more

prevalent. These stereotypic relationships between functionally related immune cell populations suggest that, in individual patients, inflammation develops along a limited number of trajectories.

While LuN and RAR can manifest the same immune cell trajectories, they were often associated with different features of tubulointerstitial inflammation and damage. For example, only in RAR were HLA class II⁺ inflammatory monocytes or CD8⁺ T cells associated with MXA expression. The potential mechanism underlying these associations are unclear. However, in neither disease were CD8⁺ T cell densities or proportions associated with fibrosis. Our sample was small and CD8⁺ T cells have been linked with progressive renal disease in RAR and LuN (31, 32). However, recent data from mouse models of LuN suggest that some of these populations have an exhausted or even protective phenotype (61, 62). Indeed, both PD-1 and ICOS expression were wide spread on our infiltrating T cells. Our data indicate that, within our cohorts, macrophages are more associated with renal damage than conventional T cell populations.

There were both similarities and differences in the distribution of immune cell infiltrates in LuN and RAR. In both diseases, there was a strong enrichment in the periglomerular space. In contrast, the distribution of inflammation within the tubulointerstitium was different in LuN and RAR. In LuN, inflammation was evenly distributed within the interstitium without peritubular enrichment. In contrast, in RAR, there was enrichment at the peritubular border with some biopsies also manifesting tubulitis. These data suggest that peritubular and tubular inflammation characterize some RAR biopsies, while LuN is characterized by interstitial inflammation. In LuN and RAR, we did not observe an enrichment of specific immune cells in any tubulointerstitial compartment. This suggests that there are no strong immunological barriers within the tubulointerstitium dictating the evolution of inflammation.

This is a small retrospective study of deidentified patient samples. We could not control for variables such as sex, age, or medication use. Regardless, our investigations suggest approaches to parsing in situ immune state heterogeneity that can be validated in larger, prospective patient cohorts.

Spatial immunology is a new and evolving field in which the imaging and computational tools are rapidly improving (63). Currently, there are still technical limitations. While DAPI nuclear segmentation is still more reliable than whole-cell segmentation, current techniques still under call irregularly shaped nuclei such as those of myeloid cells (64). We dilated nuclear segmentations to capture cytoplasmic staining for annotation. While this is a reliable strategy for assessing lymphocytes, it often fails to capture peripheral staining on large cells such as tubule cells. In part, we circumvented this limitation by segmenting whole tubules. However, CD138 expression by tubules likely led to some plasma cells being included in the tubular mask. Finally, we used a hierarchical decision tree for cell class assignment, which is a well-established approach used by others (47, 55). However, this strategy uses predefined cell classes and can miss novel cell populations. Therefore, our analysis is only an approximation. Nevertheless, within these limitations, we provide insights into in situ immune states and conceptual frameworks for further investigation.

Other approaches have biases and limitations. Single-cell sorting from renal samples has provided a picture of the overall immune landscape that informed our staining panel (22, 46). However, different immune cell populations are likely extracted at different efficiencies from

tissue. Furthermore, some cells, such as plasma cells, survive poorly during extraction and handling (65). These and other factors likely make scRNA-Seq from renal tissue variable and inefficient. In the 1 available study focused on in situ immune cells in LuN, only about 150 CD45⁺ cells were obtained per biopsy (22). In contrast, we identified over 7,000 immune cells per biopsy, with spatial coordinates on each cell. This richness of information per biopsy enabled a dissection of renal inflammation heterogeneity not possible with current scRNA-Seq techniques. Finally, our experiments were done on archived FFPE biopsies. Future studies, with larger numbers of patients from longitudinal registries, can provide clinical context for different in situ immune states.

Our data reveal that, in both LuN and RAR, and across almost all patients, the foundation of in situ inflammation is innate immunity. This is particularly surprising for LuN, which is a manifestation of the canonical systemic autoimmune disease, SLE. Our studies characterize the most prevalent immune cell populations. Adaptive cells are present in situ and there are likely important functional relationships between these and resident innate cell populations. Furthermore, the relationships between systemic adaptive autoimmunity and in situ immunity are largely unexplored in humans (66). Future mechanistic studies in SLE of concurrent blood and tissue biopsy samples, coupled with clinical trials of adaptive immune cell targeted therapies, will begin to unravel the complex interrelationships between innate and adaptive cell programs and those between systemic and in situ autoimmunity.

Methods

Sex as a biological variable. This study was performed on deidentified, archived tissue samples, and therefore knowledge of patient sex is incomplete. Of the LuN patients for which sex is known, > 90% were female (22/24). The sex of the RAR patients is unknown.

Tissue acquisition. We obtained 54 archival blocks of kidney biopsies preserved as formalin-fixed, paraffin-embedded (FFPE) from the University of Chicago Human Tissue Resource Center. Within this, 25 blocks were from the initial diagnostic biopsies of LuN patients, 23 were from RAR patients, and 6 were from the normal renal tissue at the margins of resected renal carcinomas or other pathologies.

FFPE tissue processing and staining. Five-micrometer FFPE tissue sections were cut and mounted on 22 mm × 22 mm glass coverslips. The tissue coverslips were deparaffinized as described here. The paraffin embedding was removed from the tissue sections via a 20-minute incubation at 60°C. Coverslips were transferred into Xylene and sequentially immersed in a fresh Xylene solution 2 times, 5 minutes each; 100% ethanol 2 times 5 minutes each; 95% ethanol 5 minutes; 70% ethanol 5 minutes; 50% ethanol 5 minutes; 30% ethanol 5 minutes and distilled water 5 minutes. Tissue coverslips were then treated with 1× citrate buffer, pH6 (diluted from 100× stock, Abcam ab93678) for 20 minutes in a high pressure cooker. After antigen retrieval, tissues were then stained with Akoya's staining kit for PhenoCycler (Akoya Biosciences, SKU7000008) following their protocol.

Marker panel creation. We designed a 42-marker immunofluorescence panel after conducting a literature review of relevant immune cell populations and those identified on the landmark scRNA-Seq study of LN patient biopsies (Table 2) (22). All antibodies were first validated with immunofluorescence staining on human tonsil and kidney sections. Validated antibodies were then conjugated with DNA barcodes using the conjugation kit (Akoya Biosciences, SKU7000009) and revalidated using single-stain CODEX and multicycle CODEX runs.

Table 2. Primary CODEX antibodies

Target	Clone	Manufacturer	Catalog No.	Dilution	Cycle
BDCA1	OT12F4	Novus Biologicals	NBP2-70345	1:200	1
BDCA2	polyclonal	Novus Biologicals	AF1376	1:200	2
CD10	EPR22867	Abcam	ab256494	1:200	10
CD103(ITGAE)	EPR22590-27	Abcam	AB254201-1001	1:200	9
CD11c	EP1347Y	Abcam	ab216655	1:200	1
CD138	EPR6454	Abcam	ab226108	1:200	4
CD14	SP192	Abcam	AB230903-1001	1:200	3
CD16	D1N9L	Cell Signaling	722045F	1:200	7
CD163	EDHu-1	Novus Biologicals	NB110-40686	1:200	13
CD20	L26	Abcam	ab236434	1:200	8
CD21	EP3093	Abcam	ab271855	1:200	11
CD27	BLR083G	Abcam	ab272072	1:200	10
CD3	SP7	Abcam	ab205228	1:200	12
CD31	EP3095	Abcam	ab226157	1:200	11
CD4	EPR6855	Abcam	ab181724	1:200	6
CD43	SPM503	Novus Biologicals	NBP2-34775	1:200	12
CD45	EP322Y	Abcam	ab214437	1:200	9
CD45	2B11 + PD7/26	Novus Biologicals	NBP2-34287	1:200	9
CD56	OT11G4	OriGene	CF506208	1:200	6
CD68	KP1	eBioscience	#14-0688-82	1:200	13
CD69	EPR21814	Abcam	ab234512	1:200	8
CD8	C8/144B	Invitrogen	MA5-13473	1:200	7
CD86	E2G8P	Cell Signaling		1:100	5
Claudin1	ab238949	Abcam	EPR121871	1:200	10
COL1I	polyclonal	Proteintech	22734-1-AP	1:200	4
FOXP3	PCH101	eBioscience	14-4776-82	1:200	4
GZMA	EPR20161	Abcam	ab251499	1:200	5
GZMB	D6E9W	Cell Signaling	790035F	1:200	9
GZMK	polyclonal	Abcam	EPR24601-164	1:200	8
HLA-DP,-DR,-DQ	CR3/43	Abcam	ab7856	1:200	11
ICOS	D1K2T	Cell Signaling	397405F	1:200	6
IFN-γ	EPR21704	Abcam	ab231301	1:200	7
IL-10	JES3-9D7	Invitrogen	14-7108-81	1:200	2
IL-10	polyclonal	Novus Biologicals	AF-217-NA	1:200	2
iNOS	SP126	Abcam	ab239990	1:100	3
Ki67	B56	BD	BDB556003	1:200	7
MERTK	Y323	Abcam	ab271851	1:200	13
MTOC	TU30	Abcam	ab27074	1:200	12
MUC-1	955	Novus Biologicals	NBP2-44658	1:200	10
MXA	polyclonal	R&D	AF7946	1:200	1
PD1	D4W2J	Cell Signaling	638155F	1:100	2
ROR-γ	polyclonal	ThermoFisher Scientific	PA534164	1:200	3
SLAMF7	E5C4M	Cell Signaling	66110	1:200	1
T-bet	D6N8B	Cell Signaling	#271125F	1:100	5
TCR-delta	H-41	Santa Cruz	sc-100289	1:200	4

Cycle, the numeric position at which the antibody was probed and imaged during the CODEX multiplex imaging process.

Image acquisition and processing. Sections were deparaffinized and stained our panel of 42 antibodies each conjugated to a unique oligonucleotide. Images of full biopsy sections were acquired on an Andor Dragonfly 200 Spinning Disk Confocal Microscope (0.1507 μm pixel size). The staining patterns of our 42-marker panel was acquired through iterative staining with Alexa Fluor 488, Atto550, Cy5/AF647, and AF750 fluo-

rophores conjugated to complementary oligonucleotides and imaging using the PhenoCycler platform. Tissue autofluorescence images were also acquired at each imaging wavelength.

ASHLAR was used to stitch image tiles into a full-section composite and align the resulting channels. Ashlar performance was visually checked across all samples. Areas with insufficient alignment were rejected from downstream analysis. After aligning all image channels, the first blank cycle of imaging was used for background subtraction and normalization of all stained images. First, each channel of the blank cycle was subtracted from the corresponding fluorescence channel in all imaging cycles. Each imaging wavelength has a different dynamic range, so the subtracted images were also divided by the standard deviation of the background image to standardize dynamic range across imaging wavelengths. After standardization relative to imaging wavelength, images were min-max normalized to the 99th percentile. After this preprocessing, instance segmentation of cell nuclei was performed using Cellpose 2.0. Cell body segmentations were approximated by dilating the nucleus segmentations by 7 pixels (1.05 μm). Mean fluorescence intensity (MFI) was calculated from each marker in the panel using the cell body pixel mask as reference.

Decision tree classifier and cell mapping to tissue. A subsample of 26 of the original 42 immunofluorescence markers were used to classify approximately 1.77 million cells. We used a decision-tree classifier for the multiclass annotation of cells that is analogous to flow cytometry-based cell analyses and immunophenotyping. The decision tree considers the known experimental covariance between markers (i.e., CD3⁺ expression would preclude CD4⁺ expression in T-lymphocytes). Cell positivity is determined by applying the multi-Otsu thresholding method using each cell's MFI, across all 1.77 million cells in all 3 cohorts simultaneously. The multi-Otsu threshold that most closely matched manual spot validation was used. This manual spot validation was done by manually circling 5 positive signal cells in the desired cell population (i.e., CD8⁺ expression for cytotoxic lymphocytes) and manually calculating the cell's MFI in ImageJ. We further use our kidney control to check if cell thresholds are not under calling or over calling cells. Once classified, cells are mapped back to computational segmentations of the renal tissue for further downstream spatial analyses.

Distinct immune trajectory. To statistically test for the differential presence of particular cell classes, we performed a nonparametric Mann-Whitney-U test for population differences using cell densities. We performed the following comparisons: LuN-KC, RAR-KC, and LuN-RAR. Benjamini-Hochberg *P*-value correction was performed to control for multiple *P*-value hypothesis testing.

Renal structure segmentation. For the instance segmentation of kidney tubules, Omnipose was trained on 32 tiles of 10-x downsized DAPI kidney images (160 x 160 pixels) randomly selected from 3 kidney biopsies. The total number of training annotations was 355 instances. The change in training image size compared with the training image size in cellular segmentation was to reflect the size differences between a cell and a much bigger tubule structure. The training parameters were as follows: 1,000 epochs, learning rate = 0.1, batch size = 16, number of classes = 2, tyx tuple input = 128×128 . This model was trained from scratch without using any pretrained models. Posttraining validation was done on 16 tiles of downsized DAPI kidney images (160×160 pixels), containing 195 tubule instances.

Whole-slide segmentation of lupus kidney biopsies was performed on DAPI channel after preprocessing and downsizing by a factor of 10. The segmentation parameters are as follows: mask threshold = 2.04, diameter = 30, with affinity segmentation. Postsegmentation of tubules was done to remove false positives (red blood cells clusters, clusters of aggregated lymphocytes in the interstitial spaces) by calculating the mean fluorescence intensity of canonical kidney structural markers per tubule object and subsequently removing the objects which express low tubule markers (CD10, MUC-1) and high expression level of all markers, which is typical of red blood cells. Using normalized mean fluorescence intensity, objects satisfying the following conditions are removed as false positives: $\mu_{CD10} < 0.8$, $\mu_{MUC-1} < 0.8$, $\mu_{CD138} < 1$, $0.25 < (\mu_{MUC-1})/\mu_{CD10} < 6.5$.

DBSCAN. We used density-based spatial clustering of applications with (Noise) DBSCAN to find cellular clusters or neighborhoods in our multiplex microscopy imaging data. An ϵ hyperparameter of 85 pixels (approx. 13 microns) and minimum points hyperparameter of 5 (all points are considered core points). We used bootstrapping to subsample 75% of the DBSCAN data for a total of 3,000 repetitions for the ideal K number of clusters. Using the average bootstrap sum of squared distance plot and the delta sum of square distance plot we found that the empirically best fit K-clusters was around 8. We used 8 as optimal K for downstream analysis; moving forward with phenotyping only those cell neighborhoods with 20 and more cell members. To phenotype these DBSCAN segmented cell neighborhoods, we performed feature extraction by characterizing each neighborhood using total proportion for each cell class. We included the total cell count as another descriptor. Afterwards, to find unique defining features for each of our clusters, we generated a heatmap of the leave-one-out Z-test for every cluster.

MERSCOPE generation of DBSCAN cluster masks. To perform DBSCAN (Python), raw data from MERSCOPE with coordinates of each detected gene was used, and coordinate-to-pixel transformation was performed to downsize the image to a pixel size of 10 μ m. DBSCAN was performed with coordinate information with *CD3E* ($\epsilon = 40$, min_min point = 5), *CD14* ($\epsilon = 20$, min_min point = 5), and *FCGR3A* ($\epsilon = 30$, min_min point = 5). Output masks of DBSCAN were then applied with postprocessing with Skimage including draw (radius_px = 10), binary_closing (footprint of disk radius = 5), remove_small_holes (area_threshold = 500), remove_small_objects (min_size = 100), and gaussian blurring (sigma = 2.0). Manual check and correction were done before quantifying the density of genes in labeled masks (clusters). Density of genes were calculated with the number of transcripts divided by the area of labeled mask.

MERSCOPE clustering of gene densities. Gene densities of all 500 genes for each labeled mask were calculated. Z-scored densities of genes: *CD3E*, *CD4*, *CD8A*, *FOXP3*, *PDCD1*, *ICOS*, *CD163*, *MRC1*, *NOS2*, *CD14*, *FCGR3A*, *MS4A1*, *ITGAX*, and *HLAII* (*HLA-DPA1***HLA-DP-*

*B1***HLA-DRA***HLA-DRB1***HLA-DQA1*) denote canonical immune cell populations were used for K-mean clustering. Top 10 highly expressed genes for each cluster were calculated by wilcoxon rank-sum test comparing to the other 6 clusters using scipy.stats. with alternative= 'greater,' *P* value < 0.05 and mean difference > 0.

MERSCOPE gene set enrichment analysis. Log₂FC data was generated by comparing each cluster with the all the other clusters and was used for gene set enrichment analysis. Analysis was done by using python library gesapy with the prerank function (permutation number=100).

Computational resources. All computational tasks were carried out on the MEL server located in the Radiomics and Machine Learning Facility at the University of Chicago. MEL is equipped with 256 Xeon Gold 6130 CPU cores, 3 TB of DDR4 ECC RAM, 24 TB NVMe SSD storage space, and houses 16 Nvidia Tesla V100 32GB GPU accelerators.

Statistics. Statistical tests used in this paper included Mann-Whitney-U test for population differences with Benjamini-Hochberg *P* value correction for multiple *P* value hypothesis testing and Wilcoxon rank-sum test. A *P* value of less than 0.05 was considered statistically significant.

Study approval. This study was approved by the University of Chicago Institutional Review Board. No consent was obtained for use of the biopsies which were de-identified, left over fragments from clinically indicated biopsies. A routine procedure consent was obtained for the original clinically indicated biopsy.

Data availability. The computer code used for these analyses has been deposited on Github: <https://github.com/bkwalsh/Resolution-of-in-situ-inflammation-in-human-lupus-nephritis-into-principal-immune-cell-trajectories>; commit ID 6b015dd19794bc915bb6a4d8aa3e345601d6f62c. Image data will be provided upon reasonable request. Values for all data points in graphs are reported in the Supporting Data Values file.

Author contributions

GC did most of the analyses and cowrote a draft with MST. MST oversaw the technical aspects of the project including AI analytical development and cowrote the initial manuscript draft with GC. JA and TPC did the imaging for the project. SH, as a renal pathologist, provided clinical context and sample evaluation. MSA developed methods for visualizing raw data. DG did some computer programming, DS did some clinical research. AC provided and clinically scored samples. KK identified patients. ASC oversaw many aspects of the renal allograft rejection control cohort. MLG provided technical guidance to GC and MST. MRC oversaw the whole project and prepared the final manuscript. Both GC and MST made substantial and complementary contributions to the paper.

Funding support

This work is the result of NIH funding, in whole or in part, and is subject to the NIH Public Access Policy. Through acceptance of this federal funding, the NIH has been given a right to make the work publicly available in PubMed Central.

- NIH (U19 AI 082724 [to MRC], R01 AI148705 [to MRC and ASC]).
- Chan Zuckerberg Biohub (to MRC).
- Lupus Research Alliance (to MRC).

Address correspondence to: Marcus R. Clark, University of Chicago, Section of Rheumatology, 5841 S. Maryland Ave., Chicago, Illinois, 60637, USA. Phone: 773.702.0202; Email: mclark@uchicago.edu.

1. Bernatsky S, et al. Mortality in systemic lupus erythematosus. *Arthritis Rheum.* 2006;54(8):2550–2557.
2. McLaughlin JR, et al. Kidney biopsy in systemic lupus erythematosus. III. Survival analysis controlling for clinical and laboratory variables. *Arthritis Rheum.* 1994;37(4):559–567.
3. Cook RJ, et al. Prediction of short term mortality in systemic lupus erythematosus with time dependent measures of disease activity. *J Rheumatol.* 2000;27(8):1892–1895.
4. Mok CC, Tang SSK. Incidence and predictors of renal disease in Chinese patients with systemic lupus erythematosus. *Am J Med.* 2004;117(10):791–795.
5. Cervera R, et al. Morbidity and mortality in systemic lupus erythematosus during a 10-year period: a comparison of early and late manifestations in a cohort of 1,000 patients. *Medicine (Baltimore).* 2003;82(5):299–308.
6. Ginzler EM, et al. Mycophenolate mofetil or intravenous cyclophosphamide for lupus nephritis. *N Engl J Med.* 2005;353(21):2219–2228.
7. Flanc RS, et al. Treatment for lupus nephritis. *Cochrane Database Syst Rev.* 2004;(1):CD002922.
8. Appel G, et al. Mycophenolate mofetil versus cyclophosphamide for induction treatment of lupus nephritis. *J Am Soc Nephrol.* 2009;20(5):1103–1112.
9. Korbet SM, et al. Severe lupus nephritis: racial differences in presentation and outcome. *J Am Soc Nephrol.* 2007;18(1):244–254.
10. Donadio JV Jr., et al. Prognostic determinants in lupus nephritis: a long-term clinicopathologic study. *Lupus.* 1995;4(2):109–115.
11. Neumann K, et al. Lupus in the 1980s: III. Influence of clinical variables, biopsy, and treatment on the outcome in 150 patients with lupus nephritis seen at a single center. *Semin Arthritis Rheum.* 1995;25(1):47–55.
12. Hanly JG, et al. The frequency and outcome of lupus nephritis: results from an international inception cohort study. *Rheumatology (Oxford).* 2016;55(2):252–262.
13. Davidson A, Aranow C. Lupus nephritis: lessons from murine models. *Nat Rev Rheumatol.* 2010;6(1):13–20.
14. Hsieh C, et al. Predicting outcomes of lupus nephritis with tubulointerstitial inflammation and scarring. *Arthritis Care Res (Hoboken).* 2011;63(6):865–874.
15. Austin HA, et al. Prognostic factors in lupus nephritis. Contribution of renal histologic data. *Am J Med.* 1983;75(3):382–391.
16. Appel GB, et al. Long-term follow-up of patients with lupus nephritis. A study based on the classification of the World Health Organization. *Am J Med.* 1987;83(5):877–885.
17. Parichatikanond P, et al. Lupus nephritis: clinicopathological study of 162 cases in Thailand. *J Clin Pathol.* 1986;39(2):160–166.
18. Esdaile JM, et al. The clinical and renal biopsy predictors of long-term outcome in lupus nephritis: a study of 87 patients and review of the literature. *Q J Med.* 1989;72(269):779–833.
19. Williams W, et al. The outcome of lupus nephritis in Jamaican patients. *Am J Med Sci.* 2007;334(6):426–430.
20. Gomes MF, et al. The extent of tubulointerstitial inflammation is an independent predictor of renal survival in lupus nephritis. *J Nephrol.* 2021;34(6):1897–1905.
21. Jeong H, et al. Coexisting tubulointerstitial inflammation and damage is a risk factor for chronic kidney disease in patients with lupus nephritis. *Int J Rheum Dis.* 2024;27(12):e15444.
22. Arazi A, et al. The immune cell landscape in kidneys of patients with lupus nephritis. *Nat Immunol.* 2019;20(7):902–914.
23. Chang A, et al. In situ B cell-mediated immune responses and tubulointerstitial inflammation in human lupus nephritis. *J Immunol.* 2011;186(3):1849–1860.
24. Liarski VM, et al. Cell distance mapping in human nephritis reveals organization of in situ adaptive immune responses. *Sci Transl Med.* 2014;6(230):230ra46.
25. Ko K, et al. Bcl-2 as a therapeutic target in human tubulointerstitial inflammation. *Arthritis Rheumatol.* 2016;68(11):2740–2751.
26. Liarski V, et al. Quantifying in situ adaptive immune cell cognate interactions in humans. *Nat Immunol.* 2019;20(4):503–513.
27. Abraham R, et al. Specific in situ inflammatory states associate with progression to renal failure in lupus nephritis. *J Clin Invest.* 2022;132(13):e155350.
28. Kinloch AJ, et al. Vimentin is a dominant target of in situ humoral immunity in human lupus tubulointerstitial nephritis. *Arthritis Rheumatol.* 2014;66(12):3359–3370.
29. Kinloch AJ, et al. Machine learning to quantify in situ humoral selection in human lupus tubulointerstitial inflammation. *Front Immunol.* 2020;11:593177.
30. Clark MR, et al. The pathogenesis and therapeutic implications of tubulointerstitial inflammation in human lupus nephritis. *Semin Nephrol.* 2015;35(5):455–464.
31. Couzi L, et al. Predominance of CD8⁺ T lymphocytes among periglomerular infiltrating cells and link to the prognosis of class III and class IV lupus nephritis. *Arthritis Rheum.* 2007;56(7):2362–2370.
32. Zhang T, et al. Association between tubulointerstitial CD8⁺ T cells and renal prognosis in lupus nephritis. *Int Immunopharmacol.* 2021;99:107877.
33. Chen PM, et al. Kidney tissue hypoxia dictates T cell-mediated injury in murine lupus nephritis. *Sci Transl Med.* 2020;12(538):eaay1620.
34. Rovin BH, et al. Efficacy and safety of voclosporin versus placebo for lupus nephritis (AURORA 1): a double-blind, randomised, multicentre, placebo-controlled, phase 3 trial. *Lancet.* 2021;397(10289):2070–2080.
35. Hill GS, et al. Proteinuria and tubulointerstitial lesions in lupus nephritis. *Kidney Int.* 2001;60(5):1893–1903.
36. Hill GS, et al. Predictive power of the second renal biopsy in lupus nephritis: significance of macrophages. *Kidney Int.* 2001;59(1):304–316.
37. Olmes G, et al. CD163⁺ M2c-like macrophages predominate in renal biopsies from patients with lupus nephritis. *Arthritis Res Ther.* 2016;18:90.
38. Hoover PJ, et al. Intrarenal myeloid subsets associated with kidney injury are comparable in mice and patients with lupus nephritis [preprint]. <https://doi.org/10.1101/2023.06.24.546409>. Posted on bioRxiv June 25, 2023.
39. Müller F, et al. CD19 CAR T-cell therapy in autoimmune disease - a case series with follow-up. *N Engl J Med.* 2024;390(8):687–700.
40. Mackensen A, et al. Anti-CD19 CAR T cell therapy for refractory systemic lupus erythematosus. *Nat Med.* 2022;28(10):2124–2132.
41. Caielli S, et al. Systemic lupus erythematosus pathogenesis: interferon and beyond. *Annu Rev Immunol.* 2023;41:533–560.
42. Dörner T, Lipsky PE. The essential roles of memory B cells in the pathogenesis of systemic lupus erythematosus. *Nat Rev Rheumatol.* 2024;20(12):770–782.
43. Chan OT, et al. A novel mouse with B cells but lacking serum antibody reveals an antibody-independent role for B cells in murine lupus. *J Exp Med.* 1999;189(10):1639–1648.
44. Shlomchik MJ, et al. The role of B cells in lpr/lpr-induced autoimmunity. *J Exp Med.* 1994;180(4):1295–1306.
45. Mengel M, Adam BA. Emerging phenotypes in kidney transplant rejection. *Curr Opin Organ Transplant.* 2024;29(2):97–103.
46. Lamarche B, et al. Transcriptional and spatial profiling of the kidney allograft unravels a central role for FcγRIII⁺ innate immune cells in rejection. *Nat Commun.* 2023;14(1):4359.
47. Goltsev Y, et al. Deep profiling of mouse splenic architecture with CODEX multiplexed imaging. *Cell.* 2018;174(4):968–981.
48. Muhlich JL, et al. Stitching and registering highly multiplexed whole-slide images of tissues and tumors using ASHLAR. *Bioinformatics.* 2022;38(19):4613–4621.
49. He K, et al. Mask R-CNN. In: Proceedings of the IEEE International Conference on Computer vision. 2961–2969.
50. Stringer C, et al. Cellpose: a generalist algorithm for cellular segmentation. *Nat Methods.* 2021;18(1):100–106.
51. Greenwald NF, et al. Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. *Nat Biotechnol.* 2022;40(4):555–565.
52. Stevens M, et al. StarDist Image Segmentation Improves Circulating Tumor Cell Detection. *Cancers (Basel).* 2022;14(12):2916.
53. Mosqueira-Rey E, et al. Human-in-the-loop machine learning: a state of the art. *Artif Intell Rev.* 2023;56(4):3005–3054.
54. Becht E, et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol.* 2019;37:38–44.
55. Schürch CM, et al. Coordinated cellular neighborhoods orchestrate antitumoral immunity at the colorectal cancer invasive front. *Cell.* 2020;183(3):838.
56. Li H, et al. Systemic lupus erythematosus favors the generation of IL-17 producing double negative T cells. *Nat Commun.* 2020;11(1):2859.
57. Kapellos TS, et al. Human monocyte subsets and phenotypes in major chronic inflammatory diseases. *Front Immunol.* 2019;10:2035.
58. Cutler KJ, et al. Omnipose: a high-precision morphology-independent solution for bacterial cell seg-

- mentation. *Nat Methods*. 2022;19(11):1438–1448.
59. Schubert E, et al. DBSCAN revisited: Why and how you should (still) use DBSCAN. *ACM Trans Database Syst (TODS)*. 2017;42(3):1–21.
60. Rothlin CV, et al. TAM receptor signaling in immune homeostasis. *Annu Rev Immunol*. 2015;33:355–391.
61. Yokose T, et al. Dysfunction of infiltrating cytotoxic CD8⁺ T cells within the graft promotes murine kidney allotransplant tolerance. *J Clin Invest*. 2024;134(16):e179709.
62. Tilstra JS, et al. Kidney-infiltrating T cells in murine lupus nephritis are metabolically and functionally exhausted. *J Clin Invest*. 2018;128(11):4884–4897.
63. Bressan D, et al. The dawn of spatial omics. *Science*. 2023;381(6657):eabq4964.
64. Ma J, et al. The multimodality cell segmentation challenge: toward universal solutions. *Nat Methods*. 2024;21(6):1103–1113.
65. Cassese G, et al. Plasma cell survival is mediated by synergistic effects of cytokines and adhesion-dependent signals. *J Immunol*. 2003;171(4):1684–1690.
66. Thome JJ, et al. Spatial map of human T cell compartmentalization and maintenance over decades of life. *Cell*. 2014;159(4):814–828.