Blood immunophenotyping identifies distinct kidney histopathology and outcomes in patients with lupus nephritis

Authors: Alice Horisberger[1,2], Alec Griffith[1], Joshua Keegan[1], Arnon Arazi[3], John Pulford[1], Ekaterina Murzin[1], Kaitlyn Howard[1], Brandon Hancock[1], Andrea Fava[4], Takanori Sasaki[1], Tusharkanti Ghosh[5], Jun Inamo[6], Rebecca Beuschel[1], Ye Cao[1], Katie Preisinger[7], Maria Gutierrez-Arcelus[8], Thomas M. Eisenhaure[9], Joel Guthridge[10], Paul J. Hoover[1,9], Maria Dall'Era[11], David Wofsy[11], Diane L. Kamen[12], Kenneth C. Kalunian[13], Richard Furie[3], Michael Belmont[7], Peter Izmirly[7], Robert Clancy[7], David Hildeman[14], E. Steve Woodle[14], William Apruzzese[1], Maureen A. McMahon[15], Jennifer Grossman[15], Jennifer L. Barnas[16], Fernanda Payan-Schober[17], Mariko Ishimori[18], Michael Weisman[18], Matthias Kretzler[19], Celine C. Berthier[19], Jeffrey B. Hodgin[19], Dawit S. Demeke[19], Chaim Putterman[20,21], Accelerating Medicines Partnership: RA/SLE Network*, Michael B. Brenner[1], Jennifer H. Anolik[16], Soumya Raychaudhuri[1], Nir Hacohen[9], Judith A. James[10], Anne Davidson[3], Michelle A. Petri[4], Jill P. Buyon[7], Betty Diamond[3], Fan Zhang[6,#], James A. Lederer[1,#], Deepak A. Rao[1,#]

Corresponding author:
Deepak Rao, MD PhD
60 Fenwood Road
Boston, MA 02115
617-525-1101
darao@bwh.harvard.edu

Affiliations:
1 Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
2 Lausanne University Hospital, University of Lausanne, Switzerland
3 The Feinstein Institutes for Medical Research, Northwell Health, Manhasset, NY, USA
4 Johns Hopkins University School of Medicine, Baltimore, MD, USA
5 School of Public Health, University of Colorado, Anschutz Medical Campus, CO, USA
6 School of Medicine, University of Colorado, Anschutz Medical Campus, CO, USA
7 New York University School of Medicine, New York, NY, USA
8 Boston Children's Hospital, Harvard Medical School, Boston, MA, USA
9 Broad Institute of MIT and Harvard, Cambridge, MA, USA
10 Oklahoma Medical Research Foundation, Oklahoma City, OK, USA
11 University of California San Francisco, San Francisco, CA, USA
12 Medical University of South Carolina, Charleston, SC, USA
13 University of California San Diego School of Medicine, La Jolla, CA, USA
14 University of Cincinnati College of Medicine, Cincinnati, OH, USA
15 University of California, Los Angeles, CA, USA
16 University of Rochester Medical Center, Rochester, NY, USA
17 Texas Tech University Health Sciences Center, El Paso, TX, USA
18 Cedars-Sinai Medical Center, Los Angeles, CA, USA
19 University of Michigan, Ann Arbor, MI, USA
20 Albert Einstein College of Medicine and Montefiore Medical Center, Bronx, NY, USA
21 Azreili Faculty of Medicine, Zefat, Israel
[#] These authors jointly supervised the work

Abstract

Lupus nephritis (LN) is a frequent manifestation of systemic lupus erythematosus, and fewer than half of patients achieve complete renal response with standard immunosuppressants. Identifying non-invasive, blood-based pathologic immune alterations associated with renal injury could aid therapeutic decisions. Here, we used mass cytometry immunophenotyping of peripheral blood mononuclear cells in 145 patients with biopsy-proven LN and 40 healthy controls to evaluate the heterogeneity of immune activation in patients with LN and to identify correlates of renal parameters and treatment response. Unbiased analysis identified 3 immunologically distinct groups of patients with LN that were associated with different patterns of histopathology, renal cell infiltrates, urine proteomic profiles, and treatment response at one year. Patients with enriched circulating granzyme B^+ T cells at baseline showed more severe disease and increased numbers of activated CD8 T cells in the kidney, yet they had the highest likelihood of treatment response. A second group characterized primarily by a high type I interferon signature had a lower likelihood of response to therapy, while a third group appeared immunologically inactive by immunophenotyping at enrollment but with chronic renal injuries. Main immune profiles could be distilled down to 5 simple cytometric parameters that recapitulate several of the associations, highlighting the potential for blood immune profiling to translate to clinically useful non-invasive metrics to assess immune-mediated disease in LN.

Introduction

Lupus nephritis (LN) is a common and severe manifestation of systemic lupus erythematosus (SLE) that affects over 40% of patients[1] and substantially increases SLE-associated morbidity and mortality[2–4]. Advances in the understanding of renal inflammatory processes in patients with severe LN[5–9] have not yet translated into improved renal outcomes[10]. Less than half of the patients treated with standard immunosuppressive regimens achieve complete renal response at one year, and 10% to 30% progress to end-stage renal disease[10–12]. The heterogeneity of immune pathways involved in patients with SLE and LN contributes to the variability in treatment response and poses challenges in developing successful new therapies[13,14].

Current stratification of patients with LN relies on histologic classification of the nature, intensity, and localization of glomerular involvement (ISN/RPS class), as well as the degree of active inflammation and fibrosis in the glomeruli and tubular-interstitium (NIH activity and chronicity indices)[15]. These features currently guide treatment decisions[16,17] and inclusion criteria in clinical trials (class III, IV and/or V). However they have several limitations: they do not inform on the inflammatory pathways involved, they have limited ability to predict long-term renal outcome, and they show discordance with clinical response[7,18]. Stratifying patients based on molecular or cellular features of immune-mediated pathways would improve the predictive value of kidney biopsy[19]. Efforts to characterize inflammatory pathways involved in the kidney tissue in lupus nephritis have identified signatures associated with poor clinical outcome, such as accumulation of $CD4^{neg}$ T cells[20], enriched fibrotic pathways, type I interferon signaling[6], and specific proinflammatory cytokines and chemokines[7,21].

In parallel with analyses of kidney biopsy tissue, identifying non-invasive, easily accessible methods to characterize LN immune heterogeneity and track LN activity could substantially improve disease management. Prior efforts to characterize immunologic diversity among patients

with SLE, not restricted to LN, using blood transcriptomics and proteomics have identified distinct signatures. Evidence of upregulated type I Interferon (IFN) signalling as measured by IFN response gene signatures has been repeatedly identified in 50-80% of patients with SLE and is associated with more severe disease, including renal manifestations[22,23]. Independently, a circulating neutrophil signature has been associated with active renal involvement in patients with SLE[24,25] and appears to improve with renal response to treatment[26,27]. More generally, SLE disease activity has been characterized by specific transcriptomic signatures involving myeloid, plasmablast[24,25,28], B-cell activation, and cell cycle pathways[25]. Similarly, cytometric analyses have highlighted expanded activated B cells, plasmablasts, B-helper T cells and activated myeloid cells in the blood of patients with active SLE, including patients with LN[29–33]. Clustering patients with SLE based on blood immunophenotyping by flow cytometry has shown promise in identifying distinct groups of patients with SLE with different treatment responses[34]. However, the extent to which these cellular signatures can inform renal pathology and treatment response across patients with LN is unclear.

Accordingly, this study was initiated to identify patterns of immune cell signatures in the blood of patients with LN that are associated with renal tissue activity and treatment response. By applying mass cytometry immunophenotyping to a prospective cohort of patients with biopsy proven LN in a multi-ethnic/racial study across the U.S., we have defined alterations in circulating T cells, B cells, myeloid cells and NK cells. Using classifications and unsupervised models, cellular profiles in blood have been identified that discriminate three groups of patients differing in renal activity and outcome. Further using paired single cell RNA-seq of kidney immune cells and proteomic analyses of urine, we found that these blood immunophenotype-defined groups were associated with specific cell infiltrates in kidneys and proteomic signatures in the urine.

Results

Overview of blood immunophenotyping in a longitudinal cohort of patients with active LN

We analyzed peripheral blood mononuclear cells (PBMCs) by mass cytometry using four 48-marker panels in 145 patients with SLE and biopsy-proven LN class III, IV, and/or V, as well as 40 healthy controls enrolled in the AMP SLE phase II study (Supplementary Tables 1, 2). Compared to the controls, patients with LN in this study were younger (p=0.02) and more frequently female, Black (p=0.003), and Hispanic or Latino (p=0.008). Within patients with LN, 68% had positive anti-dsDNA antibody and/or low serum complement levels, and 50% had extrarenal clinical manifestations of SLE at the time of biopsy. All patients had a clinical indication for renal biopsy based on proteinuria (defined as urine protein to creatinine ratio (UPCR) > 0.5 g/g)[16,17]. Most biopsies reported a proliferative classification (III or IV +/- V, 70%) compared to pure membranous (V, 30%). Sixty-nine percent of the patients had a history of previous renal biopsy; 61% had one or more prior episodes of biopsy-proven proliferative or membranous LN. Patients were treated following standard of care according to the judgment of their physician[16,17]. Clinical renal response was defined as previously reported in patients with a UPCR $\geq$ 1 at baseline[35,36]. At 52 weeks, 28% achieved complete response, 24% a partial response and 48% no response. Longitudinal blood samples from 50 patients were analyzed at week-12 and/or week-52 to determine the trajectory of immune cells following treatment.

To minimize technical confounders, PBMC samples were cryopreserved with optimized protocol[37] and then processed for mass cytometry at a single central site; LN and control samples were balanced across 23 batches. Each PBMC sample was stained with panels designed to interrogate T, B, myeloid, or NK cells; a total of 267 samples were retained after quality control (**Supplementary Table 2, Extended Data Fig. 1a,b**). A total of 25-30x10^6 cells/panel were available for analysis after quality control, averaging approximately 111,000 cells/sample/panel (**Extended Data Fig. 1b**). Batch effect correction with Harmony[38] reduced inter-batch signal variation as measured by the local inverse Simpson's index, without impairing cell type

identification or removing signal variation related to LN versus control (**Extended Data Fig. 1c-e**). We then applied a stepwise unsupervised approach to analyze the data generated from each panel, starting from total PBMCs, which avoids potential bias from manual gating. To ensure equal representation of each sample, we downsampled randomly to 10,000 cells/sample. We labeled each cluster as either T, B, myeloid or NK cells, based on cell-type canonical markers, and projected the results in a uniform manifold approximation and projection (UMAP) (**Fig. 1a, Extended Data Fig. 1F**). We then extracted the major cell lineage of interest (T, B, myeloid, and NK) from the dedicated panel and re-clustered the cells to define fine-grained cell-type specific neighborhoods and clusters (**Fig. 1b, Extended Data Fig. 2**).

To validate the robustness of our unbiased analysis, we confirmed our ability to reproduce previously reported alterations in immune cell subset frequencies in patients with SLE. We identified alterations in the frequency of rare cell populations including plasmablasts ($CD20^{-}CD19^{int}CD27^{hi}CD38^{hi}$)[32] and plasmacytoid dendritic cells (pDC, $CD123^{+}HLADR^{+}IRF8^{+}$)[39,40], with abundances similar to prior reports (**Extended Data Fig. 1g-h**). We also confirmed in our cohort an expansion of cell subsets previously associated with LN, including T peripheral helper cells (Tph, cluster T11, $CD4^{+}PD1^{+}ICOS^{+}CXCR5^{-}$)[32] and $CD11c^{+}$ B cell subsets (clusters B5 and B9, $CD11c^{+}T\text{-}bet^{+}CD21^{-}IgD^{+/-}$)[29,30,33], independently from demographic variables (age, sex, ethnicity and race) and immunosuppressive therapy (**Extended Data Fig. 1i**).

A cytometric score captures type I interferon signature heterogeneity in LN

To explore blood immune cell alterations in patients with LN comprehensively, we applied a covarying neighborhood analysis (CNA) in each panel to identify disease-associated cells in a cluster-free manner, while adjusting for sample-level covariates[41]. CNA results were consistent with our initial analysis of Tph and activated B cells, yet additionally detected marked diffuse alterations within each cell type, after adjusting for demographic variables and accounting for false

discovery rate (FDR) < 0.05, suggesting broad differences in the baseline cytometric profiles of patients with LN compared to controls (**Fig. 1c**). CNA results strongly correlated with those from a cluster-based single-cell mixed-effect model (MASC)[42], providing higher granularity of cell-level association, and were not affected by small cluster outliers (**Extended Data Fig. a-b**). Patterns of LN-associated cells were maintained when comparing patients not on immunosuppressive/biologic treatment and on <= 5mg of prednisone at baseline with controls, with striking differences particularly in the T cells and myeloid cells, despite reduced sample size (**Extended Data Fig. 3c**).

We then examined which markers explained the differences observed between LN and controls. Using the results generated from CNA analysis, we were able to identify the proteins associated with the axes of greatest variation characterized in the NAM (cell neighborhood abundances)-derived principal components. In the T panel, the first principal component (NAM-PC1) explained the majority (79%) of the variance observed and was conspicuously correlated with MX1, a type I interferon (IFN)-induced protein[43], compared to the other markers (absolute r=0.82, p<0.001) (**Fig. 1d**). In parallel, in the myeloid panel, monocyte expression of SIGLEC-1, an IFN-inducible protein, was strongly associated with CNA defined LN-enriched cells (**Fig. 1e**). Moreover, in the B cell panel, ISG-15, a third type I IFN-induced marker[24,25,44], was also increased in all cell types from patients with LN and its expression strongly correlated with MX1 (r=0.89, p<0.001) and SIGLEC-1 (r=0.75, p<0.001) at a sample level (**Extended Data Fig. 3d**). MX1 and ISG15 were enriched in all cell types from patients with LN compared to controls, whereas SIGLEC-1^{hi} cells were only detected in myeloid cells, consistent with previous transcriptomic studies[25,45] (**Extended Data Fig. 3e**).

Given the importance of type I IFN signaling in SLE and the emergence of therapeutic IFN receptor blockade[19,46,47], we sought to define a cytometric IFN score based on the expression of these three

proteins (MX1, SIGLEC-1, and ISG-15, see **Method**). We confirmed a significant increase in cytometric IFN score in patients with LN at baseline compared to controls (**Fig. 1f**). In addition, the type I IFN signaling was associated with the presence of serologic manifestations (presence of anti-dsDNA and low complement, **Extended Data Fig. 3f**), consistent with results from gene scores in SLE cohorts[47,48] and with the mechanistic connection between immune complexes inducing type I IFN release, further promoting auto-antibody production[49]. While type I interferon activation pathway is associated with SLE severity, including LN, its correlation with renal histology is less clear[6,23,47,50]. In this LN cohort, the cytometrically-derived type I IFN score did not correlate with proliferative class, renal activity by NIH index, including the interstitial component, or extrarenal clinical manifestations (**Extended Data Fig. 3i-j**). IFN score was also not associated with baseline prednisone dose or the use of immunosuppressants (**Extended Data Fig. 3i**). These results indicate that type I IFN signaling was common across clinically and histologically heterogeneous patients with LN.

Circulating immune cell alterations are associated with histologic patterns of active LN

As the use of immunosuppressive treatment in patients with LN is currently driven by the biopsy-defined proliferative class and the extent of renal activity[16,17,51], we next set out to investigate whether immune alterations in the blood reflected biopsy class and features of active renal inflammation. Both proliferative LN classes (III or IV +/- V) and renal activity were predominantly associated with shifts in naïve B cell phenotypes, after adjusting for demographic features and the history of previous renal biopsy (**Fig. 2a-b**, **Extended Data Fig. 4a-b**). Furthermore, differences in circulating naive B cells were associated with specific histologic features of the activity index, including glomerular endocapillary hypercellularity and the presence and degree of fibrinoid necrosis. The association between naive B cells and fibrinoid necrosis was still evident even when evaluating only the 15 patients with LN who had no immunosuppressant treatment and prednisone $\leq$ 5mg at the time of sampling (**Extended Data Fig. 4c-d**). We also detected a strong association

between cells with a neutrophil phenotype (M3, $CD14^{-}CD16^{+}CD11b^{+}CD15^{hi}$), likely representing low-density neutrophils[52], and fibrinoid necrosis, even when including only non-immunosuppressed (with maximum 5mg of prednisone) patients (**Fig. 2c, Extended Data Fig. 4c-d**). When examining only non-immunosuppressed patients, NIH activity index and fibrinoid necrosis were also associated with a subset of $CD14^{+}$monocytes characterized by low expression of CD86 (M5), a protein downregulated with LPS stimulation and sepsis[53,54], suggesting a inflammatory state. While B and myeloid cells showed an association with glomerular injuries, shifts within the T cells correlated with the interstitial component of the NIH activity index (**Fig. 2d**).

Among B cells, the positive association with different renal inflammatory lesions was primarily driven by cluster B3, a naive subset ($IgD^{+}IgM^{+}CD27^{-}$) with features distinct from transitional cells (B6) and from activated $CD11c^{+}$ subsets (B5, B9), including lack of CD11c, T-bet, CD10, and intermediate CD38 expression (**Fig. 2e-f**). Cluster B3 lacked CD23 expression and had a lower expression of CD21 and CD19 compared to the predominant naïve B cell cluster observed in controls (B0) (**Fig. 2f-g**). Despite the absence of activation or proliferation markers (HLA-DR, CD95, CD86 or Ki67), this cluster B3 was significantly associated with the NIH activity index and low complement levels (e.g., C3 and C4), independently from demographic variables and history of previous biopsies (**Fig. 2h**). In contrast, B3 was not associated with extrarenal manifestations of disease activity (presence of clinical extrarenal SLEDAI, p=0.36). Longitudinal analysis showed a decrease in the proportion of cluster B3 over time in the patients with a complete renal response at 52 weeks, but not in patients with no response (**Fig. 2h**). By manual gating, decreased expression of CD21 within naïve B cells (median CD21 within $CD20^{+}CD27^{-}IgD^{+}$ $CD38^{int/-}CD10^{-}$ $CD11c^{-}$ cells) confirmed: 1) an association with active and proliferative LN, 2) an association with low complement levels, and 3) an increased CD21 expression over time in complete responders (**Extended Data Fig. 5a-b**). Together, these results support a shift in naïve B cells towards an altered phenotype in patients with LN with active renal inflammation and low complement levels.

Cytotoxic T cells, proliferating T-B subsets and type I interferon characterize diverse cellular signatures in patients with LN

These results highlight relationships between blood immune cells and histologic evidence of active renal injury, yet the limited ability of histologic classification to predict renal outcome, including in the AMP phase II cohort[36], and the immune heterogeneity observed in the cohort challenged us to approach the data in an unsupervised manner to identify cellular profiles associated with clinical features of LN. As a first step, we presented the global relationships between immune cell subsets in patients with LN. By applying clustering to the correlation matrix of the identified immune cell clusters (after excluding small cluster outliers, see **Method**) and type I IFN score from LN baseline samples (n=115 stained with four panels), we observed sets of coordinated immune cell abundances (**Fig. 3a**, **Extended Data Fig. 6a**). As expected, cell subsets marked by type I IFN-induced proteins, such as SIGLEC-1^{hi} monocytes (M1) and MX1^{hi} naive CD4 T cells (T2 and T12), were part of a core set correlating with the cytometric type I IFN score (**Fig. 3a**). Interestingly, T12 was phenotypically distinct from T2 by its reduced expression of TCF1, a key transcription factor in T cell stemness that is downregulated upon CD4^{+} effector T cell differentiation[55,56] (**Extended Data Fig. 6b-c**). Supporting the hypothesis that these cells were recently activated in an environment enriched in type I IFN, we used a published bulk RNA-seq dataset[57] and found that anti-CD3/CD28 stimulation of naive CD4 T cells downregulated *TCF7* gene (encoding TCF1) expression, and this was accelerated by IFN-β (**Extended Data Fig. 6d**). Other cells correlating and closely organized with the type I IFN core set included MX1^{hi} TCF1^{lo} naive CD8 T cells (T10), Ki67^{hi} NK cell subsets (NK4, NK6, NK7), transitional B cells (B6), and a small cluster of B cells with features of marginal zone precursors[58] (B13, IgM^{+}CD38^{+}CD27^{-}CD1c^{+}), consistent with the previously reported roles of type I IFN in NK cell activation[59] and in immature B cell expansion[60] (**Fig. 3a**).

We observed two other distinct sets of co-abundant immune cell clusters. One set included Tph (T11), activated memory B/early plasmablasts (B10), plasmablasts (B8, B11), and CD8 GZMB$^+$ Ki67$^+$ (T13) cells, comprising a set of T cells and B cells with a shared proliferative state (Ki67$^+$), consistent with a prior study of patients with newly diagnosed SLE[61] (**Fig. 3a**). A second, distinct set of co-correlated T cells was unified by their expression of granzyme B, including CD8 (T4, T9, T20), CD4, and DN (T16, T17, T18) T cells (**Fig. 3a**). The T13 CD8 GZMB$^+$ Ki67$^+$ cluster strongly correlated with both the proliferative and the granzyme B$^+$ cell sets and also correlated with IFN score. However, non-proliferating granzyme B$^+$ T cells did not correlate with type I IFN signature (**Fig. 3a**), suggesting that an IFN signature and granzyme B expression represent two different axes of immune activation. Together, these results enabled visualization of distinct sets of highly correlated immune cell populations that are unified by specific functional features, providing a framework with which to interpret clinical associations of immune cell clusters in patients with LN.

Blood immunophenotype identifies three subgroups of patients with LN

As a second step to this unbiased approach, we evaluated whether the composition of circulating immune cell populations could distinguish groups of patients with LN. Here we applied a K-means clustering approach including all LN and control samples stained with 4 panels (n = 224), based on the scaled proportions of 55 immune cell subsets. We identified three groups of samples presenting distinct patterns, indicating potential patient subgroups. We confirmed > 80% stability of our clustering results by repeating 1000 K-means clustering on a resampled dataset without replacement (**Extended Data Fig. 7a**). PCA and UMAP visualization confirmed a distinct distribution of these three groups (**Fig. 3b, Extended Data Fig. 7b**). Group G0 included all the controls (n=39) and a minority of patients with LN (n=23), indicating that these patients with LN had a control-like immunophenotype (**Fig. 3c**). The two other groups, G1 and G2, each included 46 patients with LN at baseline.

To understand the main differences observed between these subgroups, we first examined the cell subset loadings on the first two PCs. We noted that both extremes of PC1 were marked by cells correlated with type I IFN score, suggesting that IFN response represented the main axis of variation in blood immunophenotypes (**Fig. 3c**). Type I IFN-correlated subsets, including proliferating NK cells and transitional B cells, were significantly enriched in G1 compared to the other groups (**Extended Data Fig. 7c-d**). Comparing the type I IFN score between the 3 groups of LN, G0 displayed the lowest values compared to the other groups and was not significantly different from controls. Group G1 was marked overall by the highest type I IFN values, whereas group G2 showed and intermediate level on average and the highest variability across patients (**Fig. 3d**).

B and T cells characterized by a proliferative state (B8, B10, B11, T11 and T13) were enriched in both groups G1 and G2 compared to G0. In contrast, group G2 was marked by enrichment in granzyme B^+ cell subsets (**Fig. 3c,e, Extended Data Fig. 7e**). The negative axis of PC2 appeared predominantly driven by T cells with expression of granzyme B (T4, T9, T16, T17, T18, T20 and T13), and the proportion of granzyme B^+ T cells was increased in G2 compared to G0 and G1 (**Fig. 3c, 3f, Extended Data Fig. 7f**). As a complementary approach to the PC loadings, we used CNA to evaluate cells associated with G0, G1, and G2, adjusting for demographic characteristics, history of previous renal biopsy, prednisone dose and mycophenolate mofetil (MMF) use. This analysis confirmed the strong enrichment in granzyme B^+ T cells in group G2. In addition, G2 was also characterized by an expansion of the main neutrophil cluster (M3) in the multivariable CNA model. B cells subsets enriched in G2 included “altered” naive $CD19^{lo}$ $CD21^{lo}$ (B3), and to a lesser extent $CD11c^+$ (B5, B9) clusters (**Fig. 3g**). Thus, this clustering approach distinguished 2 subsets of immunologically active patients with LN: one group (G1) with a prominent IFN signature, and a second group (G2) with a prominent expansion of granzyme B expressing lymphocytes.

Baseline blood-defined LN groups have distinct renal pathology and outcome

We next explored whether these groups of patients with LN differed in relevant renal features. Patients in the group G0 ('control-like') had the lowest renal activity index scores, while the group G2 ('cytotoxic T-enriched') had the highest renal activity scores (**Fig. 4a, Extended Data Table 1**). Consistent with the low activity index, the group G0 was more frequently classified as pure membranous, whereas G1 and G2 were more frequently classified as proliferative (85% in G2, 65% in G1, 35% in G0, $p < 0.001$). Detailed analysis of NIH renal activity index components revealed a significant correlation between G2 and endocapillary hypercellularity, active crescents, and fibrinoid necrosis, further underscoring active and more severe glomerular lesions in group G2 (**Fig. 4b**). In addition, G2 correlated positively with the interstitial activity subscore, although results were at the limit of the FDR threshold of < 0.05 (p=0.01, FDR=0.05). Consistent with a more severe disease in the group G2 compared to G1 at baseline, those in the former group had increased serum creatinine (p=0.04), UPCR (p=0.009), and higher prednisone dose (p<0.001). In contrast, extrarenal manifestations and serologic measures (positive anti-dsDNA and/or low complement) were not significantly different between G1 and G2 (**Extended Data Table 1**). Adjusting for demographic variables, history of previous biopsy and baseline prednisone dose, the group G2 remained significantly associated with increased renal activity indices (beta coefficient [95%CI] = 2.6 [0.8-4.5], p.adj = 0.006) and proliferative class (OR [95%CI] = 4.2 [1.4-14.1], p.adj = 0.015) compared to the other groups (**Fig. 4c**).

Surprisingly, the group G0 had the highest NIH chronicity index and subscores, whereas G1 and G2 were not significantly different (**Fig. 4d, Extended Data Table 1**). We observed that patients in G0 were older and more frequently had a history a previous biopsy (**Extended Data Table 1**). Examining the detailed classes in previous biopsies, we noted a significant increase in past proliferative (III, IV +/- V) disease in the group G0 compared to G2 (65% G0, 59% G1 and 33% G2, overall p=0.01, G0 v G2 p = 0.02). The majority (73%) of the patients with a membranous

class in the group G0 had prior history of proliferative class, a histologic conversion previously reported[62]. Adjusting for demographic variables, history of previous proliferative disease and baseline prednisone dose, G0 remained significantly associated with chronicity compared to the other groups, although to a limited extent (beta coefficient [95% CI] = 1.3 [0.04-2.5], p.adj = 0.044). Together these results suggest that G0 identified a group of patients with little evidence of immune activation in blood, which is associated with minimal current renal activity, yet who may have previously had severe LN that evolved towards chronic renal damage.

We then asked whether these groups of patients had different renal outcomes. In the 88 patients with a defined renal response at week 52 (baseline UPCR $\geq$ 1), we observed a significant difference across groups (p=0.04) and increased renal response in G2, compared to G1 (p=0.03). Adjusting for demographic variables, history of previous biopsy, and baseline prednisone dose, the group G2 had an increased likelihood of complete response at one year (OR [95%CI] = 8.5 [2.2-40.5], p.adj = 0.003). Given the variability in treatment received, we additionally tested the likelihood of response in patients treated with MMF throughout the study (MMF use at week 26 and 52: 52% in G0, 61% in G1 and 50% in G2), and observed a persistent association between G2 and complete response compared to others (OR [95%CI] = 5.6 [1.7-20.8], p = 0.007). These results indicate that G2, characterized by increased proliferative and non-proliferative granzyme B^+ T cells, neutrophils and B cell alterations, identifies a group of patients with severe, active disease but with increased likelihood of response to standard of care, including to an MMF-based regimen.

Baseline blood-defined LN groups differ in kidney immune cell infiltrates and urine profiles

Given the differences in disease activity and outcome between these groups and previous reports associating type I IFN signaling with poorer renal response[6,7], we hypothesized that the groups of patients with LN defined by blood immunophenotyping would show differences in the composition

of the immune infiltrate within the kidney and type I IFN immune response gene expression. We leveraged scRNA-seq data generated on cells from kidney biopsies of 101 patients from the AMP phase II study[37,63,64]. Given the prominent cytotoxic CD8 T cell features of G2 in the blood, as well as the previously recognized presence of CD8 T cell infiltrates in the periglomerular and interstitial space in proliferative LN kidneys[20,65,66], we then asked if CD8 T cell accumulation in the kidney differed in the blood-defined LN patient groups. Indeed, G2 patients showed an enrichment in granzyme B^+ CD8 T cells, as well as a subset of granzyme K^+ cells, in the kidney compared to the other LN groups (**Fig. 5a-c**). In contrast, G2 patients were not associated with populations of $CD8^+$ T cells enriched in *ITGA1, ZNF683* and *XCL1,* and depleted in *KLF2* expression, features suggestive of resident memory T cells[67,68], which are also present in control kidneys[5] (**Fig. 5a**). In addition, we observed differences in subsets of CD4 T cells in the kidney, which were mostly enriched in G0 and reduced in G2 (**Fig. 5a-b**). Using a 21-gene type I IFN list[13,69], we found that G1 patients had the highest IFN type I score in kidney immune cells, while G0 the lowest values, mirroring the pattern seen in the blood (**Fig. 5d**). In total, these results indicate that immunoprofiling based on blood immune cells can identify subsets of patients with LN that differ in their renal immune cell infiltrates, with specific, direct relationships between immunophenotypic features in blood and those in kidney.

To further extend the association between blood immune profiles and kidney injury, we compared the urine proteomic profiles obtained from paired samples (n=115), previously generated using the Raybiotech kiloplex technology and published separately[35]. After adjusting for FDR < 0.10, we observed significant differences between groups (**Fig. 5e**). The group G2 was enriched in granzyme A and with proteins previously associated with increased NIH activity and proliferative LN class, such as Catalase and CD163[35,70] (**Fig. 5e**). Compared to the group G1, G2 had increased secretion of several proteins, including TGFβ1, which had been was reported to be enriched in kidney tissue during proliferative LN flares and was predictive of complete renal

response[7] (**Fig. 5e**). In contrast, G1 was enriched in GITR-L and CD83, proteins expressed by antigen presenting cells, when compared to G2 and was significantly enriched in Siglec-1 when compared to G0 (**Fig. 5h**). Together these results highlighted the association between blood immune alterations and kidney inflammatory process in patients with LN.

Proliferating T-B cells decrease over time in MMF-based treatment responders

Given our results indicating that specific blood immune cell profiles associated with renal activity and outcome, we next aimed to understand how treatment may impact these signatures at baseline, in the context of a clinically diverse LN cohort, and longitudinally. We first used a linear model with penalization at baseline to simultaneously examine the importance of demographic features, renal and extrarenal characteristics, as well as medications, in accounting for the variation in immune cell signatures. We focused this analysis on 5 key blood immune signatures identified in the supervised and unsupervised analyses: type I IFN, proliferative T-B cells, non-proliferative granzyme B^+ T cells, altered naïve B cells, and low-density neutrophils (**Fig. 1-3**). Consistent with preceding analyses, the NIH renal activity index independently positively correlated with an increase in T-B proliferative cells (including T CD8 $GZMB^+$ $Ki67^+$ cells), low-density neutrophils and altered naïve B cells, whereas non-proliferative cytotoxic T cells were positively associated only with the interstitial component of the activity index. Type I IFN was not associated with renal activity (**Fig. 6a**). Prednisone dose at baseline (excluding high intravenous dose of 500-1000mg) was an additional independent factor associated with increased low-density neutrophils, whereas MMF was independently associated with decreased proliferative T-B cells (**Fig. 6a-b**).

We then examined the group changes longitudinally in these signatures. Immunologically inactive patients (group G0) at baseline uniformly remained stable over time (**Fig. 6c**). In contrast, the majority of patients who started in G2 changed their group over time, including some transitioning

to G0, suggesting that the cellular features of G2 are dynamic and can change with clinical course and treatment. Detailed analysis of immune signatures showed that the proportion of altered naïve B cells decreased over time in complete responders but not in partial and non-responders (**Fig. 6d**), while immune features of IFN score, proliferating T-B cells, and cytotoxic T cells did not change consistently in either responders or non-responders (**Fig. 6d**). However, the decreased trend observed in proliferative T-B cells was significant in responders when including only patients with an MMF-based regimen, suggesting a drug effect (**Fig. 6d**). In contrast, the type I IFN signature did not significantly change over time, including in patients treated with MMF, supporting previous reports in general SLE[23,48] that this pathway is not robustly altered by standard of care LN regimens.

Finally, we confirmed that these 5 main cellular features could be captured using a simplified manual gating strategy using a limited number of canonical markers (**Extended Data Fig. 8**). To minimize the number of markers, we measured the expression of CD21 in non-proliferative B cells and for visualization purposes (**Fig 6e,f**) we showed the inversed value, for consistency with previous graphs. Overall, this manual gating strategy, applied to the total dataset of live cells (25-30^10^6), reproduced the key patterns of cellular features associated with renal disease activity and longitudinal changes, including the decrease in proliferative T-B cells and the increase in CD21 expression in non-proliferative B cells over time, supporting that these populations can readily be identified (**Fig. 6e,f**).

Discussion

Here, we have used blood cell immunoprofiling to identify immunologically distinct groups of patients with LN that differ in renal inflammation and outcome in a large cohort of patients who underwent a clinically-indicated renal biopsy. Unsupervised modeling approaches visualized co-abundant cell populations and stratified patients with LN into 3 groups: an immunologically inactive

group (G0), a predominant type I IFN group (G1), and a predominant cytotoxic T cell group (G2). Both immunologically active groups (G1, G2) show increased proportions of proliferating T cells and B cells, yet these 2 groups differ with much higher IFN scores in G1 and increased proportions of cytotoxic T cells, low density neutrophils, and CD21^{low} altered naïve B cells in G2. These results reveal specific features that distinguish the immunologic heterogeneity among patients with LN, highlighting in particular that features of an activated cellular response characterizing the group G2 can be separated from a high IFN response. Further, a separate group of patients with biopsy-demonstrated LN, patients in G0, show little evidence of immune activation, with immunologic profiles that are indistinguishable from control patients.

LN patient groups distinguished by this blood immunophenotyping approach differed in renal injury and renal immune cell infiltration, indicating that blood profiling can reflect relevant intra-renal features. Group G0 patients, who appeared immunologically quiet by immunophenotyping, also showed little renal activity, although had the highest scores for chronic renal damage. The history of prior episodes of proliferative nephritis in these patients suggests that these patients may have proteinuria due to chronic renal damage but little ongoing immune injury. Group G1 patients showed the highest IFN scores in the blood and associated with an increase in circulating proliferative NK cells and increased transitional B cells, and correspondingly showed the highest expression of IFN response in immune cells from kidney. IFN responses may alter the functions of many kidney-infiltrating immune cells as well as kidney parenchymal cells[71–74]. The high IFN signature in blood cells, kidney cells, and in urine of these patients suggests a globally high IFN response, which can be robustly captured by blood cytometry, consistent with prior reports[43,75].

The G2 group, characterized by the expansion of granzyme B^{+} T cells, showed the most active proliferative renal disease. This is consistent with prior reports of expanded effector memory T CD8 cells in the blood of SLE patients with more active disease[76–78]. Our detailed profiling

approach revealed a set of co-correlated granzyme B^+ T cell populations, including subsets of CD8 T effector memory and terminally differentiated effector memory populations, proliferating CD8 T cells, CD4 T cells and $CD8^-CD4^-$ T cells, associated with this group of patients with LN. We further demonstrated that this G2 patient group has a concomitant enrichment in paired kidney samples of $GZMB^+$ and $GZMK^+$ T cell subsets, a major driver of tissue inflammation in autoimmune diseases[79]. A pathogenic role for CD8 T cells in patients with LN has been previously suggested by the enrichment in cytotoxic and proliferating CD8 T cells and production of inflammatory chemokines and interferon-γ[5,9], a cytokine implicated in the pathogenesis on LN disease[80] in kidneys of patients with LN[65,66]. A direct connection between CD8 T cells in blood and kidney is suggested by the description of shared CD8 T cell clones between these sites in patients with LN[66]. Interestingly, we identified different correlations involving specific T cell phenotypes: circulating proliferating cells (including granzyme B^+ T cells) correlated with glomerular pathology, whereas non-proliferating granzyme B^+ T cells correlated with interstitial renal activity.

Notably, the G2 group with a predominant cytotoxic T signature at baseline showed an increased likelihood of response to standard of care at one year compared to the other groups. This is somewhat surprising because expansion of effector CD8 T cells in SLE patients and infiltrating CD8 T cells in LN have been associated with more severe and refractory disease[65,66,78]. In addition, tubulointerstitial activity has been previously associated with poorer outcome[81]. However, by assessing the immune alterations comprehensively, we also found that this group of patients had a distinct urine proteomic profile, suggesting broader immune differences between groups. We hypothesize that these features reflect aspects of immune cell activation that can effectively be suppressed by current immunosuppressive therapies, such as MMF, which was associated with reduction in low density neutrophils and $CD21^{lo}$ naïve B cells populations in this cohort. Additionally, the two other groups of patients identified were either marked by high fibrotic histologic lesions (G0) or high type I interferon responses (G1), both signatures previously

associated with poorer outcome in the AMP phase I study[6]. In addition, we showed that the type I IFN pathway was not reduced by standard of care immunosuppressive therapy, potentially contributing to the low renal response rate in this group. The particularly high expression of type I IFN, contrasting with the variability in the group G2, suggests that considering the intensity of IFN response and combining it with other immune cell signatures might help stratify patients more effectively than the commonly used division of IFN-high versus IFN-low patients. Further studies will help define if patients with a disease characterized by high type I IFN benefit from therapies that effectively disrupt this pathway.

Our study has limitations. Although we examined the blood immunophenotype using 4 panels with 48 proteins each, other relevant signals might be present and could help further determine key cellular features associated with the immune cells enriched in the different groups of patients, including cytokine profiles enriched in the different groups. Ongoing work, as part of the AMP phase II study, on transcriptomic profile of immune and resident kidney cells, as well as blood immune cells will provide more detailed blood-tissue comparison including comprehensive signals potentially further distinguishing these groups of patients with LN. Secondly, controls were not well-matched to our cohort of patients with LN. To account for demographic differences, we systematically analyzed our data in univariate and multivariate models adjusting for age, sex, ethnicity and race. Thirdly, this study was limited by the lack of standardized therapy. Despite these challenges, we were able to identify immune cell changes that were either associated with or independent of the immunosuppressive therapies. Detailed assessment in a randomized trial with standardized therapy will further help understand the immune changes associated with the heterogeneity of response to specific therapies.

In summary, our study highlights the power of blood immunophenotyping to define subsets of patients with LN with distinct patterns of immune activation that reflect ongoing renal inflammation

and injury, which provides critical evidence for future personalized treatment. This work indicates a valuable role for blood cellular biomarkers in evaluating specific dimensions of immune activity of patients with LN to gain insight into the likelihood of response to standard immunosuppressive LN therapies.

Method

Study population

SLE patients 16 years or older were recruited between January 2016 and May 2021 at 14 clinical sites across the United States as part of the Accelerating Medicines Partnership (AMP) RA/SLE Phase II study[37,82] if 1) they fulfilled American College of Rheumatology[83] or Systemic Lupus International Collaborating Clinics classification criteria[84] for SLE, 2) had a clinically-indicated renal biopsy for a UPCR > 0.5[16,17] and 3) had biopsy-confirmed lupus nephritis of class III, IV and/or V, based on the International Society of Nephrology/Renal Pathology Society (ISN/RPS)[15]. Patients with a history of renal transplantation, with a medical condition considered at risk of participating in this study by the investigator, who was pregnant at the time of study entry, or who had recently been exposed to rituximab (within the past six months) were excluded. Healthy controls were recruited from 2 of the clinical sites. All participants provided written informed consent before study enrollment, and human study protocols were approved in accordance with the Declaration of Helsinki by the institutional review boards (IRBs) at each participating sites, which included: Johns Hopkins University, New York University, University of Rochester Medical Center, Oklahoma Medical Research Foundation, University of Cincinnati, Albert Einstein College of Medicine, University of California San Francisco, Northwell Health, Medical University of South Carolina, Texas University El Paso, University of Michigan, University of California San Diego, University of California Los Angeles, Cedars-Sinai Medical Center.

Data and sample collection

Patient characteristics and clinical assessment were collected at each site, including demographic characteristics (age, sex, ethnicity and race), history of previous renal biopsy, SLE disease activity measured by the hybrid Safety of Estrogens in Lupus Erythematosus National Assessment (SELENA)-SLE Disease Activity Index (SLEDAI) (urine protein to creatinine ratio > 0.5 always counted)[85,86], anti-dsDNA antibody status (according to the upper limit of the norm of each site laboratory), C3 and C4 complement values, serum creatinine, estimated glomerular filtration rate, UPCR, and medications received during the study. Patients were treated during the study at the discretion of their physicians based on standard of care. Renal response to treatment was determined clinically at week 52 as complete, partial, or none, adapted from the ACCESS trial definitions[87], and as reported[35], in patients with a baseline UPCR ratio $\geq$ 1. Briefly, complete response (CR) was defined by UPCR <0.5, normal serum creatinine (≤ 1.3 mg/dL) or, if abnormal, <125% of baseline, and prednisone $\leq$10mg/day. Partial response (PR) was defined by >50% reduction in UPCR without meeting UPCR criterion for CR, normal creatinine (≤ 1.3 mg/dL) or, if abnormal, ≤ 125% of baseline, and prednisone dose ≤ 15 mg/day.

Renal biopsy sections were classified and scored according to the ISN/RPS classification as proliferative (class III, IV +/- V) or membranous (class V), and the NIH activity and chronicity indices[15]. Initial assessment was performed by a renal pathologist at each site, followed by a central read by two independent renal pathologists for 101 biopsies. Central reads included a detailed subscoring of each component of the NIH activity and chronicity indices. Scores from central reads were used for further analysis if available, and total scores from the individual site reads were used otherwise.

Blood samples were collected from patients with LN and controls at baseline (week 0), and a subset of patients with LN were recollected at weeks 12 and 52. PBMCs were isolated from blood samples and cryopreserved at each site as previously described[32]. Cryopreserved PBMC were

shipped to the central AMP RA/SLE Biorepository, Oklahoma Medical Research Foundation Biorepository, for storage until sample collection was complete. A total of 275 samples were collected from 152 patients with LN and 40 controls (**Extended Data Fig. 1a**).

Sample staining and data acquisition by mass cytometry

Cryopreserved PBMC samples were sent to the Brigham and Women's Hospital CyTOF Antibody Resource and Core. Samples were then randomly distributed across 23 staining and acquisition days (23 batches) at 20 samples per day (**Extended Data Fig. 1a**). Samples were thawed, stained and mass cytometry data was generated according to the AMP RA/SLE mass cytometry processing workflow, and as previously reported[88]. As described, samples were thawed in a 37 °C water bath for 3 minutes and then mixed with 37 °C thawing media containing: RPMI Medium 1640 (Life Technologies #11875-085) supplemented with 5% heat-inactivated fetal bovine serum (Life Technologies #16000044), 1 mM GlutaMAX (Life Technologies #35050079), antibiotic-antimycotic (Life Technologies #15240062), 2 mM MEM non-essential amino acids (Life Technologies #11140050), 10 mM HEPES (Life Technologies #15630080), 2.5 x 10^{-5} M 2-mercaptoethanol (Sigma-Aldrich #M3148), 20 units/mL sodium heparin (Sigma-Aldrich #H3393), and 25 units/mL benzonase nuclease (Sigma-Aldrich #E1014). 100 µL aliquots of each sample post-thaw were mixed with PBS (Life Technologies #10010023) at a 1:1 ratio to be counted by flow cytometry.

Each sample was then split into 4 to be stained with 4 different panels including lineage and specific markers dedicated to examining B, T, myeloid, and NK cells (**Extended Data Fig. 1a**). Overall, between 0.5 – 1.0 x 10^6 cells were stained per panel for each sample. If the sample before splitting included <2x10^6, a panel priority was used: B panel > T panel > Myeloid panel > NK panel. The list of markers is detailed in **Supplementary Table 3**. All samples were transferred to a polypropylene plate (Corning #3365) to be stained at room temperature for the rest of the

experiment. The samples were spun down and aspirated. Rhodium viability staining reagent (Standard BioTools #201103B) was diluted at 1:1000 and added for five minutes. 16% stock paraformaldehyde (Fisher Scientific #O4042-500) was diluted to 0.4% in PBS and added to the samples for five minutes. After centrifugation and aspiration, Human TruStain FcX Fc receptor blocking reagent (BioLegend #422302) was used at a 1:100 dilution in cell staining buffer (CSB) (PBS with 2.5 g bovine serum albumin [Sigma Aldrich #A3059] and 100 mg of sodium azide [Sigma Aldrich #71289]) for 10 minutes followed by incubation with conjugated surface antibodies (each marker was used at a 1:100 dilution in CSB, unless stated otherwise) for 30 minutes. The Harvard Medical Area CyTOF Antibody Resource and Core (Boston, MA) prepared and validated all antibodies. After centrifugation, samples were resuspended with culture media. 16% stock paraformaldehyde (Fisher Scientific #O4042-500) dissolved in PBS was used at a final concentration of 4% for 10 minutes to fix the samples before permeabilization with the FoxP3/Transcription Factor Staining Buffer Set (ThermoFisher Scientific #00-5523-00). The samples were incubated with SCN-EDTA coupled palladium barcoding reagents for 15 minutes, followed by incubation with Heparin (Sigma-Aldrich #H3149-100KU) diluted 1:10 in PBS. Samples were combined and filtered in a polypropylene tube fitted with a 40μm filter cap. Conjugated intracellular antibodies were added into each tube and incubated for 30 minutes. Cells were then fixed with 4% paraformaldehyde for 10 minutes. DNA was labeled for 20 minutes to identify single-cell events with an 18.75 μM iridium intercalator solution (Standard BioTools #201192B). Samples were subsequently washed and reconstituted in Cell Acquisition Solution (CAS) (Standard BioTools #201240) in the presence of EQ Four Element Calibration beads (Standard BioTools #201078) at a final concentration of $1x10^6$ cells/mL. Samples were acquired on a Helios CyTOF Mass Cytometer (Standard BioTools).

Mass cytometry data processing and quality control

The raw FCS files were normalized to reduce signal deviation between samples during multi-day batch acquisitions, based on the bead standard normalization method established by Finck et al. [89]. The normalized files were then compensated with a panel-specific spillover matrix to subtract cross-contaminating signals[90]. These compensated files were then deconvoluted into individual sample files using a single cell-based debarcoding algorithm established by Zunder et al[91]. Normalized and debarcoded FCS files were then uploaded to OMIQ software from Dotmatics (www.omiq.ai, www.dotmatics.com), as previously reported[88].

Manual gating on Gaussian parameters and DNA was applied to remove debris, outlier events, and doublets. Singlet live cells were then identified by gating on bead-negative and Rhodium-negative cell events. The data was further evaluated for quality control by examining the proportion of viable cells, the presence and frequency of major cell types (B, T, myeloid, and NK) and the distribution of samples using unsupervised clustering with opt-SNE dimensionality reduction and PARC clustering[92], available in the OMIQ software from Dotmatics (www.omiq.ai, www.dotmatics.com). A total of 8 samples were excluded from further analysis due to one or several quality issues: cell viability < 50%, 0 B cells, and samples identified mostly (>90%) in one PARC cluster. Thus, further analysis included 264 samples from 145 patients with LN and 40 controls (**Extended Data Fig. 1a**, **Supplementary Table 2**). In total, over 110 million cells were analyzable across panels, and the median cell counts per sample within each panel after quality control was above 100,000 cells (**Extended Data Fig. 1b**).

Data transformation and batch correction

The data generated from each panel was analyzed separately using the same strategy. After quality control filtering, FCS files including live singlet cells, were uploaded and read in R (R version 4.3.1) using the flowCore package (version 2.14.0) and we then applied arcsinh-transformation with a cofactor of 5 to the protein expression matrix, as described[93]. We leveraged

the single-cell transcriptomic analysis pipeline from the Seurat package (version 4.3.0)[94] for our workflow by transposing our transformed expression matrix and implementing it into a Seurat object with the corresponding metadata. To ensure equal representation of samples (**Extended Data Fig. 1b**) and to save computational time and resources, we randomly downsampled each sample to 10,000 cells (if a sample had < 10,000 cells, all cells were kept), as downsampling does not affect the sensitivity of the clustering analysis[88].

All proteins in each panel were examined and included in further analysis, except for IgG and IgA in the B panel, which were not used for further clustering the cells due an identified strong co-expression of both markers that was determined to be a technical limitation. We used principal component analysis (PCA) for dimensionality reduction, as implemented in the function RunPCA (set without approximation) from the Seurat package and calculated the 20 first principal components (PCs). We then corrected batch effects with the function RunHarmony based on the Harmony[38] application programming interface implemented in the Seurat package (**Extended Data Fig. 1c**). To assess the performance of batch correction we quantified the integration of the batches using the function compute_lisi from the LISI package (version 1.0)[38] (**Extended Data Fig. 1d**).

Graph-based clustering, visualization of protein expression and cell subset annotation

After batch correction, we built a nearest-neighbor graph using the function FindNeighbors using the 20 first harmonized PCs. The clustering was performed using a Louvain-based algorithm implemented in the FindCluster function (Seurat package). For this step we used a resolution of 0.5 for all panels after confirming that all major cell subsets could be identified. We projected cells into two dimensions using the runUMAP function based on the 20 first harmonized PCs, and with the following arguments: metric = cosine, min.dist = 0.08, spread = 1. We obtained between 23 to 29 clusters in each panel. We annotated each cluster based on common lineage markers and

additional markers specific to each panel as either: B cells (including clusters with a plasmablast phenotype), T cells, myeloid/DC cells (including clusters with a plasmacytoid dendritic phenotype, basophil phenotype, and neutrophil phenotype) and NK cells. Clusters that were not attributed clearly to one of these cell types were labeled undetermined and were excluded from further analysis (< 0.1% of total live cells for each panel) (**Extended Data Fig. 1f-h**).

As a second step, we extracted the cell type of interest in the dedicated panel (e.g., T cells in the T panel) and re-clustered the data. We applied a similar pipeline for each subsetted Seurat object by applying again the functions: RunPCA(), RunHarmony(), FindNeighbors(), FindCluster() and RunUMAP(). Additionally, for the cell-type specific nearest-neighbor graph, we used a k=30 based on an elbow plot of the within-cluster sum of squares. For the Louvain-based algorithm, we optimized resolution for each cell type (0.8 for B cells, 0.7 for T cells, 0.3 for Myeloid cells, 0.3 for NK cells), based on the cluster distribution and manual check of expression of critical proteins in each cluster to gain the biological interpretations that made the most sense. Cell-type specific clusters were labeled with a first letter corresponding to the panel they were extracted from, and a number based on the cluster size (e.g., the cluster “B0” corresponds to the largest cluster within B cells extracted from the B panel) and were further annotated based on canonical markers expression and relevant literature.

Overall, for visualization of protein expressions, we showed the arcsinh-transformed data in multiple plot types adapted from the Seurat package: 1) Dotplots to display all proteins in each cluster with dots colored by average of unscaled expression and sized by percentage of cells with non-zero values; 2) Violin plots to show the distribution of a selection of proteins, where the minimal and maximal thresholds were data-driven by the cells expressing the lowest and highest expression; 3) Feature plots to map a selection of protein expression in the UMAP space, with the same min-max threshold definition than in 2).

Association of cell subsets with disease status and clinical characteristics

To examine cell subsets associated with LN disease or with clinical conditions within patients with LN, we applied co-varying neighborhoods analysis (CNA)[41] to our dataset using its r version (rcna package, version 0.0.99) implemented in the Seurat package through the association.Seurat function. CNA defines data-driven neighborhoods (here referring to small regions of cells sharing proteomic similarities) and measures the relative abundance of cells from each sample across all neighborhoods. By further applying PCA to the neighborhood abundance matrix, CNA identifies dominant axes of variation (NAM-PCs) in cell neighborhood abundance across samples and further allows to test for clinical association including potential covariates, in a linear model. We used this approach to test cell-association with individual clinical characteristics in a univariate way and by adjusting for potential confounding factors, as specified in the results and legends. For each model, we examined and reported the global CNA p-value, which was defined through a permutation test, and further evaluated the cell-neighborhoods that passed a threshold of FDR<0.05. To visualize the local associations with the tested variables, we mapped the results in the UMAP space and colored the cell neighborhoods in red or blue, for any significant positive or negative correlation, respectively, passing the FDR threshold.

We tested the consistency of our results by using a separate single-cell cluster-based approach based on a mixed-effect model (MASC)[42]. The model evaluates whether the tested variable (LN relative to controls) improves the null model determining the membership of each cell to its cluster while controlling for inter-individual variability (random effect) and covariates (fixed effects). To run MASC function, we extracted the components from the Seurat object to build a dataframe with one row per cell and the random and fixed variables as columns. MASC results provided the significance and effect size (defined by an odds ratio) of disease status (LN relative to controls) for each cluster. To examine the consistency between the overall results of the 2 approaches, we

tested the correlation between the median cell correlation coefficient, defined by CNA, for each cluster, and the odds ratio defined by MASC (**Extended Data Fig. 3a**).

Protein association with the dominant signal in CNA

To examine the markers driving the main axe of variation when comparing patients with LN to controls, we determined the correlation between NAM-PC1 coordinates (obtained from the CNA analysis) and each protein expression in the panel, at a single-cell level (**Fig. 1d,e**). We extracted the data from the Seurat object and determined the spearman's coefficient of correlation between NAM-PC1 and individual protein expression using the rcorr function from the package Hmisc. Results were validated visually by mapping the expression of the highest marker identified in the UMAP space in comparison to the mapping of CNA results.

Cytometric type I interferon score calculation

To summarize type I interferon induced markers into a score, we identified any proteins in our panels that had been previously and repeatedly used as a gene or a protein associated with type I interferon signaling[13,25,43,75]. We retained MX1 (T panel), SIGLEC-1 (Myeloid panel) and ISG15 (B panel) and examined their expression across cell subsets. We calculated the median expression of MX1 and ISG15 amongst total live cells, as the expression was broadly observed for both markers (**Extended Data Fig. 3e**). In contrast we used the median expression of SIGLEC-1 amongst myeloid cells, as no other cell type expressed it (**Extended Data Fig. 3e**). For each sample stained with these three panels, the values for each marker were standardized to the mean and standard deviation of the controls and then summed to obtain a score.

Identification of circulating immune cell signatures

We examined how immune cell alterations correlate within patients with LN. We extracted the proportions of each cell subsets among cell type for each sample. To prevent signals from small

cluster outliers, we removed cell subset clusters that included more than 50% of samples with a zero value (example shown in **Extended Data Fig. 3b**). We excluded three T cell clusters (T21, T22, and T23), three B cell clusters (B14, B15, and B16) and two myeloid clusters (M10 and M12). We obtained a matrix of 55 cell subset (=cell clusters) proportions with sample representing the rows and columns representing cell subsets.

We studied the correlation between all cell clusters and the type I interferon score, including all baseline samples from patients with LN stained with four panels (n=115). To obtain the Spearman's rho correlations coefficients and p values between all cell subsets and type I interferon score, we used the rcorr function from the Hmisc package (version 5.1-1). We applied FDR correction to the p values to account for multiple testing using the p.adjust function from the stats package (version 4.3.1). We then used the R package corrplot (version 0.92) to explore visually our matrix of correlations by highlighting the correlations that passed FDR and by reordering the correlation by hierarchical clustering using the Ward distance method (**Fig. 3a, Extended Data Fig. 6a**).

Identification of cellular profiles associated with high type I interferon signaling

To examine the cell subsets strongly associated with type I interferon signaling. We first examined the cytometric score and its correlation with all the cell subsets in the correlation plot. To identify the differential expression markers between two T CD4 naïve clusters associated with type I interferon signature, we used the limma package[95], a method used previously to evaluate differential expression in mass cytometry dataset[96], to compute the p values for each marker included in the T panel. After identifying TCF1 as the most differentially expressed marker, we examined it's expression in CD4 naïve T cells after stimulation with CD3/CD28 with or without IFN beta, using a publicly available RNAseq dataset (GSE195541)[57].

Identification of blood-defined immunophenotype groups

We then identified three immunophenotype groups based on K-means clustering on the 55 scaled cell subset proportions for each sample stained with four panels (including controls and LN at all timepoints) using the kmeans function from the stats package. To test for the stability of the clusters we repeating 1000 K-means clustering on a resampled dataset without replacement, using the ConsensusClusterPlus package (version 1.66.0)[97] and further examined the mean cluster consensus across samples for each k and obtained over 80% of stability with a k=3 (**Figure 3a**). We additionnally tested that the differences across groups were also reflected in the PCA and UMAP spaces using the umap function from the uwot package (version 0.1.16) (**Figure 3b, Extended Data Fig. 7b**).

To identify the cells driving the differences between these groups, 1) we examined the cell subset loadings in the first 2 PCs (**Figure 3c**), 2) we run cell-type specific CNA comparing each group with the others including patients with LN at baseline and used a heatmap to combine the results (**Figure 3g**) and 3) we compared the proportions of a selection of cell subsets, either combined or individually (**Figure 3e,f, Extended Data Figure 7d-f**). For visualization of CNA results, we plotted the frequency of cells within each cell subset passing the FDR threshold of <0.05. All CNA models with cells passing FDR had a global p value of < 0.05.

Association of blood-defined LN groups with T cell infiltrating cells and type I IFN score

As part of the AMP phase II study, human kidney biopsies were cryopreserved, thawed, dissociated, sequenced and processed as described[63]. Following alignment, dimensionality reduction was achieved by identifying variable genes and running PCA, followed by batch correction using Harmony[38], graph-based clustering using Seurat (version 4.1.0)[94]. T and NK cell clusters were identified based on the expression of known lineage markers (Al Souz *et al.*, manuscript in preperation). Blood-defined groups were assigned to paired samples and CNA[41] was applied to examine the differences in cell neighborhoods between groups, in addition to

examining the proportions of T cell clusters. To evaluate type I IFN signaling between groups, we used the AddModuleScore function (Seurat package) to define type I IFN score based on a previously described 21-gene list[13], across all immune cells in the kidney.

Urine proteomic from paired patients with LN

The screening of urine proteomic was previously described and reported for patients with LN included in the AMP phase II study[35]. Briefly, the screening was performed using an extended version of the Kiloplex Quantibody (RayBiotech) and the concentration of each analyte was normalized by urine creatinine to account for urine dilution. We compared the abundance of urine proteins in $pg_{protein}/mg_{creatinine}$ between blood-defined groups of patients with LN using a Wilcoxon rank sum test and adjusting for FDR, with a threshold of < 0.10 and < 0.25 as indicated.

Statistics

In addition to the statistical testing mentioned above, we used non-parametric tests for cross-sectional univariate analysis at baseline: Wilcoxon rank sum test to compare continuous or ordinal variables between two groups, Kruskal-Wallis test followed by post hoc Dunn's test to compare continuous or ordinal variables between more than two groups, or Spearman's rho to assess the correlation between two continuous or ordinal variables (using wilcox.test, kruskal.test, dunn_test, cor.test with the method spearman functions in r). To compare categorical variables between groups, we used Fisher exact test and Chi-squared test depending on the size of the tested categorizes (using fisher.test and chisq.test functions in r). For multivariable models, we used linear regression models or generalized linear models for logistic regression (lm and glm function in r) to test for association with a continuous or categorical variable, respectively, including the covariate detailed in the results. To compare LN groups using logistic regression models, we applied a one-vs-rest strategy to examine the specific characteristic of one group compared to the others. To test for clinical and renal factors associated with the variation of selected cell subsets

of interest, we applied a linear model with penalization using an elastic net regression (glmnet function used with the caret package) after 10 random repeats of a 10-fold cross-validation, including the cell proportion as the response variable and the clinical and renal variables as 'predictor' variables. All predictor variables were normalized to a fixed range between 0 and 1. For all above cross-sectional analysis, missing clinical data or samples were excluded from statistical testing. For longitudinal data, we used mixed effect models (lmer function from the lme4 package, version 1.1-34) to test for change over time by including each patient as a random factor to account for paired samples. For longitudinal data, we included all patients with a defined response status at 1 year who had at least two separate samples. All statistical tests were two-sided. For data visualization, we used the packages ggplot, ComplexHeatmap and corrplot. All analysis were preformed on R version 4.3.1.

Data and code availability

Mass cytometry and clinical datasets will be made available on the Synapse platform (synapse.org) upon publication. All analyses can be reproduced using the publicly available versions of the R packages outlined in the method.

Acknowledgments

We would like to thank the patients who participated in this study, and the scientists and clinical sites in the Accelerating Medicines Partnerships in RA/SLE. This work was supported by the Accelerating Medicines Partnership® Rheumatoid Arthritis and Systemic Lupus Erythematosus (AMP® RA/SLE) Network[#]. AMP is a public-private partnership (AbbVie Inc., 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 Funding was provided through grants from the National Institutes of Health (UH2-AR067676, UH2-AR067677, UH2-AR067679, UH2-AR067681, UH2-AR067685, UH2- AR067688, UH2-AR067689, UH2-AR067690, UH2-AR067691, UH2-AR067694, and UM2- AR067678). This work was also conducted with support from UL1TR002541 award through Harvard Catalyst | The Harvard Clinical and Translational Science Center (National Center for Advancing Translational Sciences, National Institutes of Health) and financial contributions from Harvard University and its affiliated academic healthcare centers. Dr Horisberger fellowhip was supported by a grant from the Swiss National Science Foundation Early Postdoc.Mobility P2LAP3_199491 and by the Lausanne University Hospital in Switzerland. This work was also supported by the Arthritis National Research Foundation Award (to F.Z.).

Authors contribution

A.F., K.P., M.D., D.W., D.L.K., K.C.K., R.F., M.B., P.I., R.C., D.H., E.S.W., M.A.M., J.G., J.L.B., F.P.-S.M.I., M.W., M.K., C.P., J.A.J., M.A.P., J.P.B. recruited patients, obtained samples, and curated clinical data. D.A.R, J.A.L., F.Z., A.F., T.M.E., J.G., P.J.H., M.D., D.W., D.L.K., K.C.K., R.F., M.B., P.I., R.C., D.H., E.S.W., W.A., M.A.M., J.G., J.L.B., F.P.-S., M.I., M.W., M.K., C.C.B., J.B.H., D.S.D., C.P., M.B.B., J.H.A., S.R., N.H., J.A.J, A.D., M.A.P., J.P.B., and B.D. contributed to the processing of samples and design of the AMP SLE study. J.B.H. and D.S.D. provided central histopathological read of the kidney biopsies. A.G., J.K., J.P., E.M., K.H., B.H. and, J.A.L. designed and generated mass cytometry data. A.G. and A.H. processed mass cytometry data with the support of T.G. and T.S. A.H. analyzed the data with the support of J.I., F.Z., M.G.A. and D.A.R. T.S., R.B. and Y.C. advised and supported on aspects of the data analysis. T.M.E., N.H. generated the scRNAseq data from kidney tissue and S.R. and A.A. processed and analyzed the scRNAseq data. A.F.

analyzed the urine proteomic data. A.F., P.J.H., P.I., M.B.B., J.H.A., S.R., N.H., J.A.J., A.D., M.A.P., J.P.B., B.D., F.Z., J.A.L., D.A.R. provided disease clinical and immunology input. A.H. and D.A.R. wrote the initial draft. D.A.R., J.A.L, and F.Z. supervised the research. All authors participated in editing the final manuscript.

Declaration of interests

D.A.R. and M.B.B. are co-inventors on patent on Tph cells as a biomarker of autoimmunity.

Figure legends

Figure 1. Blood immunophenotyping by mass cytometry captures the range of type I interferon (IFN) signaling intensity in lupus nephritis (LN). a, Identification of major immune cell types within peripheral blood mononuclear cells from 267 samples (from 145 patients with LN and 40 controls) stained with four different panels. **b**, Detailed cell type specific clustering analysis in each dedicated panel (e.g. T cells in the T panel). **c**, Cell-type specific associations with patients with LN relative to controls at baseline, using covarying neighborhood analysis (CNA) and adjusted for demographic variables (age, sex, ethnicity and race). Cells in the UMAP are expanded (red) or depleted (blue) in patients with LN. **d,** Top 10 correlation coefficients between markers in the T panel or **e**, the Myeloid panel and the major axis of variation of CNA, defined by the first principal component analysis (PC1) on a matrix including all neighborhood abundances (NAM). The correlation coefficient was defined using Spearman's rho testing. **f**, Comparison of a type I IFN cytometric score between LN (n=125) and controls (n=40). Statistical significance was determined using Mann-Whitney U test, ***p<0.001.

Figure 2. Circulating immune cell subsets association with histologic patterns of active lupus nephritis (LN). a, Summary of the results testing the association within each blood cell type (y axis) and different histologic patterns of LN disease (x axis). Statistical significance was determined using covarying neighborhood analysis (CNA) after adjusting for demographic variables (age, sex, ethnicity and race) and history of previous biopsy. The color represent the global CNA p value as indicated and the star indicates if local associations passed FDR < 0.05. **b**, Selection of detailed CNA results between B cell alterations, **c**, myeloid alterations, and **d**, T cell alterations and specified histologic characteristics. **e**, Distribution of NIH activity index-associated cell neighborhood correlations (detailed CNA results) in a selection of B cell clusters. **f**, Violin plots of selected proteins expression in selected B cell clusters. **g**, Scatterplot of cluster B0 and B3 expression of selected proteins. **h**, Proportion of cluster B3 (among B cells) association with LN

histologic class (n=140), activity index (n=124) and complement levels (n=138), as well as longitudinal change in non-responders (n=19) and complete responders (n=23). Statistical significance was determined using linear models adjusting for demographic variables and history of previous biopsy for cross-sectional analysis at baseline, and using mixed effect models for longitudinal analysis to include patients as a random effect.

Figure 3. Unbiased analysis of blood immunophenotype identifies groups of lupus nephritis (LN) patients with distinct immune cell signatures. a, Correlation plot organized by hierarchical clustering showing Spearman's rho correlation coefficient between 55 immune cell clusters as well as type I interferon (IFN) cytometric signature in 115 patients with LN. **b**, Main principal components distribution of 224 samples (39 controls, 115 baseline LN samples, 39 LN samples from week-12 and 31 LN samples from week-52) based on the proportion of 55 immune cell subsets and colored by K-means-defined groups (left panel) of by disease status (right panel). **c**, Immune cell subset loading on the first two principal components. **d-f**, Comparison of key immune cell signatures between each LN groups. Statistical significant was determined using Kruskal-Wallis test followed by Bonferroni correction for multiple testing. **g**, Heatmap of CNA associations of specific immune cell subsets with each LN groups relative to the two other groups. Colors represent the percentage of cells passing FDR < 0.05, as either enriched (red) or depleted (blue). The colors associated with selected cell clusters and groups of cell clusters (green, yellow, purple and red) are consistent through **a-g**.

Figure 4. Blood-defined LN groups are associated with renal pathology and outcome. a, Comparison of available NIH activity index between K-means defined groups (20 G0, 40 G1 and 40 G2) of patients with LN. Statistical significant was determined using Kruskal-Wallis with Dunn's multiple comparisons. **b**, Heatmap of Spearman's rho correlation coefficient between each group against others and each subscores of the NIH activity and chronicity indices, adjusting for multiple

testing (FDR < 0.05 and < 0.1 are indicated). **c**, Multivariable models testing one-vs-rest group association with activity index (linear model) and proliferative class relative to membranous (logistic regression), accounting for demographic variables, history of previous biopsy and prednisone dose. **d**, Comparison of available NIH chronicity index between K-means defined groups (n=20 G0, n=40 G1 and n=40 G2) of patients with LN. Statistical significance was determined using Kruskal-Wallis with Dunn's multiple comparisons. **e**, Multivariable model testing one-vs-rest group association with chronicity index (linear model), accounting for demographic variables, history of previous biopsy and prednisone dose. **f**, Multivariable model testing one-vs-rest group association with response status (non-complete response vs complete response), accounting for demographic variables, history of previous biopsy and prednisone dose. **g**, Univariate model testing one-vs-rest group association with response status in patients treated with MMF throughout the study.

Figure 5. Blood-defined LN groups are associated with specific immune cell infiltration in the kidney and translate to urinary markers of inflammation. a, Distribution of T cells and NK cells and selected gene expression by scRNA seq analysis of kidney tissue from patients with LN. **b**, CNA association of T-NK cells with LN group G2 relative to others (G0 and G1) applied to paired kidney tissue samples. **c**, Comparison of the proportions of CNA-identified granzyme B+ and a subset of granzyme K+ immune cell subsets, and T CD4 memory cells in the kidney tissue between groups. **d**, Comparison of type I interferon (IFN) gene signature (21-gene list previously reported) in all immune cells from the kidney tissue between blood-defined LN groups. Statistical significant was determined using Kruskal-Wallis with Dunn's multiple comparisons in (C and D), *p < 0.05, ** p < 0.01, ***p < 0.001. **e**, Comparison of each urine proteine abundance between specified groups, as displayed by the performance of classification using the AUC (area under the curve). Dots colored in pink passed FDR threshold of 0.10, and labels were assigned to the top associated proteins. Statistical significance was determined using Wilcoxon rank sum test.

Figure 6. Immune cellular signatures heterogeneity across and within patients with LN. a, Importance and direction of the effect of demographic, clinical and renal characteristics on immune cell signatures. The coefficient were defined using a linear model with an elastic net penalization using a 10-fold cross-validation, with the cellular signature as a response variable and the y-axis variables as the predictor variables. **b**, Comparison of immune cell signatures between treated (Yes) or not treated (No) with MMF at baseline. Statistical significance was determined using Wilcoxon rank sum test test, ***p<0.001. **c**, Changes in blood-defined group membership over time. Each band represent a patient, and each patient is colored by the baseline group membership. All patients with LN with samples at three timepoints and samples stained with all four panels were included in this analysis. **d**, Longitudinal changes in immune cell signatures, stratified by response status (NonCR include none and partial responders, CR includes all complete responders). The upper panels include all patients with LN with at least 2 timepoints. The lower panels include patients with LN who were treated with MMF throughout the study. Statistical significance was determined using a mixed effect model including patient as a random effect. *p<0.05, **p<0.01, ***p<0.001. **e**, Correlations between simplified immune cell signatures and NIH activity subscores at baseline. Statistical significance was determined using Spearman's rho correlation followed by FDR correction for multiple testing. For visualization purposes we showed the inversed value of median CD21 in non-proliferative B cells (written as *-1). **f**, Longitudinal changes in simplified immune cell signatures in patients treated with an MMF based therapy. For visualization purposes we showed the inversed value of median CD21 in non-proliferative B cells (written as *-1). Statistical significance was determined using a mixed effect model including patient as a random effect. p = 0.05, *p<0.05, **p<0.01, ***p<0.001.

Extended Data Figure legends

Extended Data Figure 1. Overview of blood immunophenotyping by mass cytometry in a cohort of lupus nephritis (LN). a, Sample processing and quality control pipeline. **b**, Number of

cells analyzable, after filtering for quality control, in each panel and stratified by disease status and timepoints. The dotted grey line represents the threshold of 10,000, used for downsampling. **c**, Representative example of distribution of cells stained with the T panel in the first two principal components, before and after batch correction. **d**, Expression of key markers before and after batch correction in the T panel. **e**, Local inverse Simpson's index (LISI) scores per cell in the T panel measuring the mixture of cells per batch or disease status (SLE versus control). Increased index represents increased mixture of cells. **f**, UMAP representing the expression of key lineage markers by cells stained in the four different panels. **g**, Proportion of the main cell types, amongst total peripheral blood mononuclear cells, identified in each panel. **h**, Identification of plasmablasts and plasmacytoid dendritic cells in the UMAP space and by marker expression, with expected proportions when comparing all baseline LN patients (n=140 in the B panel and n=125 in the myeloid panel) and controls (n=40) (Wilcoxon signed-rank test). **i**, Comparison of proportion of specific cell subsets (% Tph amongst T cells and % of $CD11c^+$ B cells amongst B cells) between 40 controls and total LN patients (**$p<0.01$, ***$p<0.001$ by linear regression adjusting for age, sex, ethnicity and race) or between controls and 15 LN without immunosuppression and prednisone $\leq$ 5mg ($^{\#}p<0.05$, $^{\#\#}p<0.01$, $^{\#\#\#}p<0.001$ by Wilcoxon signed-rank test).

Extended Data Figure 2. Cell-type specific clustering and marker expression per cluster. Cell type specific clusters colored in the UMAP space. All proteins included in each panel and used for clustering cells are shown in the dotplots. Clusters are ordered by size (number of cells) from bottom to top, for each panel.

Extended Data Figure 3. Comprehensive identification of immune alterations in lupus nephritis (LN) patients reveal a proteomic type I interferon signature. **a**, Identification of cell clusters associated with LN (max n=140 in the B panel and min 116 in the NK panel) relative to controls (max n=40 in all panels except for n=39 in the NK panel), using two different approaches

: covarying neighborhood analysis (CNA) and a single-cell mixed-effect model (MASC). Plots represent the correlation between the two models using Spearman's rho. **b**, Example of two cellular cluster outliers identified as outliers in figure a. **c**, Identification of cells associated with a subgroup of LN patients without immunosuppressive therapy and prednisone dose maximum 5 mg at the time of sample collection (max n=15 in the B panel and min n=13 in the NK panel), relative to controls. **d**, Correlation between the median expression per sample of MX1 and ISG amongst total live cells and SIGLEC-1 amongst myeloid cells (Spearman's rho correlation). **e**, Distribution of the level of expression of type I interferon induced proteins in major cell types. **f**, Comparison of a combined type I interferon score (sum of normalized MX1 and ISG15 in live cells and SIGLEC-1 in myeloid cells) in LN patients with serologic parameters. **g**, Lack of association between cytometric type I interferon score and clinical or histologic characteristics, **h**, including with the NIH activity subscores, using Wilcoxon rank-sum test or Spearman's rho correlation.

Extended Data Figure 4. Circulating cell-type specific alterations are associated with histologic patterns of active LN. a, Summary of the results testing the association within each blood cell type (y axis) and different histologic patterns of LN disease (x axis). Statistical significance mentioned in the graph were determined using a univariate covarying neighborhood analysis (CNA). **b**, Selection of detailed CNA results between B cell alterations and specified histologic characteristics. **c**, Summary of CNA results including only LN patients with no immunosuppressive therapy and prednisone $\leq$ 5mg with **d**, representative detailed results in the myeloid and B panels.

Extended Data Figure 5. Naïve B cells shift their phenotype towards a low expressing CD19 and CD21 profile in active and proliferative LN patients. a, Gating strategy to identify naïve B cells by mass cytometry. **b**, Median CD21 expression amongst naïve B cells association with histological characteristics, complement values (C3 and C4) and longitudinal change over time.

Statistical significances were determined either using Wilcoxon sum-rank test (LN class), Spearman's rho correlation (NIH activity index and complement) or a mixed effect model with patient as a random effect (changes over time).

Extended Data Figure 6. Hierarchical clustering of co-correlated cells identifies cell subsets strongly associated with type I interferon signaling. a, Correlation between 55 immune cell subsets organized by hierarchical clustering algorithm. The black boxes represents the first 5 levels of hierarchy. Main co-correlating sets of cells are labeled by the main characteristics of cells including in the groups. **b**, Differential expression of markers included in the T panel between the cell subset T2 and T12. Statistical significance was determined using the limma package. **c**, UMAP of TCF1 expression in T cells. **d**, Changes in TCF7 (TCF1 gene) expression in healthy naïve CD4 T cells stimulated with or without interferon-β, using published bulk RNA-seq dataset.

Extended Data Figure 7. Circulating immune cell subsets characterizing three LN groups. a, Stability of sample membership to clusters depending on the number of clusters by repeating 1000 K-means clustering on a resampled dataset without replacement. **b**, UMAP distribution of all samples included in this study (n=267) based on the proportion of 55 immune cell subsets. **c**, Cell neighborhood associations with the K-means defined groups G1 and G2 relative to the other groups. **d-f** Comparison of selected cell subsets between the three K-means defined groups of LN patients at baseline (23 G0, 46 G1 and 46 G2). Statistical significance was determined by Kruskal-Wallis test followed by Dunn's multiple comparisons.

Extended Data Figure 8. Manual gating strategy to obtain simplified cellular immunophenotype.

References

1. Hoover, P. J. & Costenbader, K. H. Insights into the epidemiology and management of lupus nephritis from the US rheumatologist's perspective. *Kidney Int.* **90**, 487–492 (2016).
2. Tektonidou, M. G., Dasgupta, A. & Ward, M. M. Risk of End-Stage Renal Disease in Patients With Lupus Nephritis, 1971-2015: A Systematic Review and Bayesian Meta-Analysis. *Arthritis Rheumatol* **68**, 1432–1441 (2016).
3. Kharawala, S. *et al.* Health-related quality of life, fatigue and health utilities in lupus nephritis: A systematic literature review. *Lupus* **31**, 1029–1044 (2022).
4. Mok, C. C., Kwok, R. C. L. & Yip, P. S. F. Effect of renal disease on the standardized mortality ratio and life expectancy of patients with systemic lupus erythematosus. *Arthritis Rheum.* **65**, 2154–2160 (2013).
5. Arazi, A. *et al.* The immune cell landscape in kidneys of patients with lupus nephritis. *Nat. Immunol.* **20**, 902–914 (2019).
6. Der, E. *et al.* Tubular cell and keratinocyte single-cell transcriptomics applied to lupus nephritis reveal type I IFN and fibrosis relevant pathways. *Nat. Immunol.* **20**, 915–927 (2019).
7. Parikh, S. V. *et al.* Molecular profiling of kidney compartments from serial biopsies differentiate treatment responders from non-responders in lupus nephritis. *Kidney Int.* **102**, 845–865 (2022).
8. Peterson, K. S. *et al.* Characterization of heterogeneity in the molecular pathogenesis of lupus nephritis from transcriptional profiles of laser-captured glomeruli. *J. Clin. Invest.* **113**, 1722–1733 (2004).

9. Fava, A. *et al.* Integrated urine proteomics and renal single-cell genomics identify an IFN-γ response gradient in lupus nephritis. *JCI Insight* **5**, (2020).
10. Gisca, E., Duarte, L., Farinha, F. & Isenberg, D. A. Assessing outcomes in a lupus nephritis cohort over a 40-year period. *Rheumatology* **60**, 1814–1822 (2021).
11. Furie, R. *et al.* Two-Year, Randomized, Controlled Trial of Belimumab in Lupus Nephritis. *N. Engl. J. Med.* **383**, 1117–1128 (2020).
12. Petri, M., Barr, E. & Magder, L. S. Risk of Renal Failure Within 10 or 20 Years of Systemic Lupus Erythematosus Diagnosis. *J. Rheumatol.* **48**, 222–227 (2021).
13. Jayne, D. *et al.* Phase II randomised trial of type I interferon inhibitor anifrolumab in patients with active lupus nephritis. *Ann. Rheum. Dis.* **81**, 496–506 (2022).
14. Rovin, B. H. *et al.* Efficacy and safety of rituximab in patients with active proliferative lupus nephritis: the Lupus Nephritis Assessment with Rituximab study. *Arthritis Rheum.* **64**, 1215–1226 (2012).
15. Bajema, I. M. *et al.* Revision of the International Society of Nephrology/Renal Pathology Society classification for lupus nephritis: clarification of definitions, and modified National Institutes of Health activity and chronicity indices. *Kidney Int.* **93**, 789–796 (2018).
16. Fanouriakis, A. *et al.* 2019 Update of the Joint European League Against Rheumatism and European Renal Association-European Dialysis and Transplant Association (EULAR/ERA-EDTA) recommendations for the management of lupus nephritis. *Ann. Rheum. Dis.* **79**, 713–723 (2020).
17. Kidney Disease: Improving Global Outcomes (KDIGO) Glomerular Diseases Work Group. KDIGO 2021 Clinical Practice Guideline for the Management of Glomerular Diseases. *Kidney Int.* **100**, S1–S276 (2021).

18. Malvar, A. *et al.* Histologic versus clinical remission in proliferative lupus nephritis. *Nephrol. Dial. Transplant* **32**, 1338–1344 (2017).
19. Morand, E. F. *et al.* Trial of Anifrolumab in Active Systemic Lupus Erythematosus. *N. Engl. J. Med.* **382**, 211–221 (2020).
20. Abraham, R. *et al.* Specific in situ inflammatory states associate with progression to renal failure in lupus nephritis. *J. Clin. Invest.* **132**, (2022).
21. Bauer, J. W. *et al.* Interferon-regulated chemokines as biomarkers of systemic lupus erythematosus disease activity: a validation study. *Arthritis Rheum.* **60**, 3098–3107 (2009).
22. Baechler, E. C. *et al.* Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 2610–2615 (2003).
23. Northcott, M. *et al.* Type 1 interferon status in systemic lupus erythematosus: a longitudinal analysis. *Lupus Sci Med* **9**, (2022).
24. Banchereau, R. *et al.* Personalized Immunomonitoring Uncovers Molecular Networks that Stratify Lupus Patients. *Cell* **165**, 551–565 (2016).
25. Nakano, M. *et al.* Distinct transcriptome architectures underlying lupus establishment and exacerbation. *Cell* **185**, 3375-3389.e21 (2022).
26. Wither, J. E. *et al.* Identification of a neutrophil-related gene expression signature that is enriched in adult systemic lupus erythematosus patients with active nephritis: Clinical/pathologic associations and etiologic mechanisms. *PLoS One* **13**, e0196117 (2018).

27. Jourde-Chiche, N. *et al.* Modular transcriptional repertoire analyses identify a blood neutrophil signature as a candidate biomarker for lupus nephritis. *Rheumatology* **56**, 477–487 (2017).
28. Guthridge, J. M. *et al.* Adults with systemic lupus exhibit distinct molecular phenotypes in a cross-sectional study. *EClinicalMedicine* **20**, 100291 (2020).
29. Tipton, C. M. *et al.* Diversity, cellular origin and autoreactivity of antibody-secreting cell population expansions in acute systemic lupus erythematosus. *Nat. Immunol.* **16**, 755–765 (2015).
30. Jenks, S. A. *et al.* Distinct Effector B Cells Induced by Unregulated Toll-like Receptor 7 Contribute to Pathogenic Responses in Systemic Lupus Erythematosus. *Immunity* **49**, 725-739.e6 (2018).
31. Sanz, I. *et al.* Challenges and Opportunities for Consistent Classification of Human B Cell and Plasma Cell Populations. *Front. Immunol.* **10**, 2458 (2019).
32. Bocharnikov, A. V. *et al.* PD-1hiCXCR5- T peripheral helper cells promote B cell responses in lupus via MAF and IL-21. *JCI Insight* **4**, (2019).
33. Wang, S. *et al.* IL-21 drives expansion and plasma cell differentiation of autoreactive CD11chiT-bet+ B cells in SLE. *Nat. Commun.* **9**, 1758 (2018).
34. Kubo, S. *et al.* Peripheral Immunophenotyping Identifies Three Subgroups Based on T Cell Heterogeneity in Lupus Patients. *Arthritis Rheumatol* **69**, 2029–2037 (2017).
35. Fava, A. *et al.* Urine proteomic signatures of histological class, activity, chronicity, and treatment response in lupus nephritis. *JCI Insight (in press)* (2023).
36. Izmirly P.M. *et al*. Longitudinal Patterns of Response to Standard of Care Therapy for Lupus Nephritis: Data from the Accelerating Medicines Partnership Lupus Network. *ACR Meeting Abstracts* Preprint at Arthritis Rheumatol. 2021; 73 (suppl 9).

37. Hoover, P. *et al.* Accelerating Medicines Partnership: Organizational Structure and Preliminary Data From the Phase 1 Studies of Lupus Nephritis. *Arthritis Care Res.* **72**, 233–242 (2020).

38. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).

39. Psarras, A. *et al.* Functionally impaired plasmacytoid dendritic cells and non-haematopoietic sources of type I interferon characterize human autoimmunity. *Nat. Commun.* **11**, 6149 (2020).

40. Blanco, P., Palucka, A. K., Gill, M., Pascual, V. & Banchereau, J. Induction of dendritic cell differentiation by IFN-alpha in systemic lupus erythematosus. *Science* **294**, 1540–1543 (2001).

41. Reshef, Y. A. *et al.* Co-varying neighborhood analysis identifies cell populations associated with phenotypes of interest from single-cell transcriptomics. *Nat. Biotechnol.* **40**, 355–363 (2022).

42. Fonseka, C. Y. *et al.* Mixed-effects association of single cells identifies an expanded effector $CD4^+$ T cell subset in rheumatoid arthritis. *Sci. Transl. Med.* **10**, eaaq0305 (2018).

43. Huijser, E. *et al.* MxA is a clinically applicable biomarker for type I interferon activation in systemic lupus erythematosus and systemic sclerosis. *Rheumatology* **58**, 1302–1303 (2019).

44. Chiche, L. *et al.* Modular transcriptional repertoire analyses of adults with systemic lupus erythematosus reveal distinct type I and type II interferon signatures. *Arthritis Rheumatol* **66**, 1583–1595 (2014).

45. Nehar-Belaid, D. *et al.* Mapping systemic lupus erythematosus heterogeneity at the single-cell level. *Nat. Immunol.* **21**, 1094–1106 (2020).

46. Bennett, L. *et al.* Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. *J. Exp. Med.* **197**, 711–723 (2003).

47. Feng, X. *et al.* Association of increased interferon-inducible gene expression with disease activity and lupus nephritis in patients with systemic lupus erythematosus. *Arthritis Rheum.* **54**, 2951–2962 (2006).

48. Petri, M. *et al.* Longitudinal expression of type I interferon responsive genes in systemic lupus erythematosus. *Lupus* **18**, 980–989 (2009).

49. Baechler, E. C., Gregersen, P. K. & Behrens, T. W. The emerging role of interferon in human systemic lupus erythematosus. *Curr. Opin. Immunol.* **16**, 801–807 (2004).

50. Der, E. *et al.* Single cell RNA sequencing to dissect the molecular heterogeneity in lupus nephritis. *JCI Insight* **2**, (2017).

51. Parikh, S. V., Almaani, S., Brodsky, S. & Rovin, B. H. Update on Lupus Nephritis: Core Curriculum 2020. *Am. J. Kidney Dis.* **76**, 265–281 (2020).

52. Rahman, S. *et al.* Low-density granulocytes activate T cells and demonstrate a non-suppressive role in systemic lupus erythematosus. *Ann. Rheum. Dis.* **78**, 957–966 (2019).

53. Nolan, A. *et al.* Differential role for CD80 and CD86 in the regulation of the innate immune response in murine polymicrobial sepsis. *PLoS One* **4**, e6600 (2009).

54. Huber, S., Hoffmann, R., Muskens, F. & Voehringer, D. Alternatively activated macrophages inhibit T-cell proliferation by Stat6-dependent expression of PD-L2. *Blood* **116**, 3311–3320 (2010).

55. Escobar, G., Mangani, D. & Anderson, A. C. T cell factor 1: A master regulator of the T cell response in disease. *Sci. Immunol.* **5**, eabb9726 (2020).

56. Johnson, J. L. *et al.* Lineage-Determining Transcription Factor TCF-1 Initiates the Epigenetic Identity of T Cells. *Immunity* **48**, 243-257.e10 (2018).

57. Sumida, T. S. *et al.* Type I interferon transcriptional network regulates expression of coinhibitory receptors in human T cells. *Nat. Immunol.* **23**, 632–642 (2022).

58. Palanichamy, A. *et al.* Novel human transitional B cell populations revealed by B cell depletion therapy. *J. Immunol.* **182**, 5982–5993 (2009).

59. Madera, S. *et al.* Type I IFN promotes NK cell expansion during viral infection by protecting NK cells against fratricide. *J. Exp. Med.* **213**, 225–233 (2016).

60. Bradford, H. F. *et al.* Inactive disease in patients with lupus is linked to autoantibodies to type I interferons that normalize blood IFNα and B cell subsets. *Cell Rep Med* **4**, 100894 (2023).

61. Sasaki, T. *et al.* Longitudinal Immune Cell Profiling in Patients With Early Systemic Lupus Erythematosus. *Arthritis Rheumatol* **74**, 1808–1821 (2022).

62. Fava, A. *et al.* History of proliferative glomerulonephritis predicts end stage kidney disease in pure membranous lupus nephritis. *Rheumatology* **61**, 2483–2493 (2022).

63. Hoover, P. J. *et al.* Intrarenal myeloid subsets associated with kidney injury are comparable in mice and patients with lupus nephritis. *bioRxiv* 2023.06.24.546409 (2023) doi:10.1101/2023.06.24.546409.

64. Fava, A., Raychaudhuri, S. & Rao, D. A. The Power of Systems Biology: Insights on Lupus Nephritis from the Accelerating Medicines Partnership. *Rheum. Dis. Clin. North Am.* **47**, 335–350 (2021).

65. 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.* **56**, 2362–2370 (2007).

66. Winchester, R. *et al.* Immunologic characteristics of intrarenal T cells: trafficking of expanded CD8+ T cell β-chain clonotypes in progressive lupus nephritis. *Arthritis Rheum.* **64**, 1589–1600 (2012).

67. Mackay, L. K. *et al.* Hobit and Blimp1 instruct a universal transcriptional program of tissue residency in lymphocytes. *Science* **352**, 459–463 (2016).

68. Boddupalli, C. S. *et al.* ABC transporters and NR4A1 identify a quiescent subset of tissue-resident memory T cells. *J. Clin. Invest.* **126**, 3905–3916 (2016).

69. Yao, Y. *et al.* Development of Potential Pharmacodynamic and Diagnostic Markers for Anti-IFN-α Monoclonal Antibody Trials in Systemic Lupus Erythematosus. *Hum. Genomics Proteomics* **2009**, (2009).

70. Mejia-Vilet, J. M. *et al.* Urinary Soluble CD163: a Novel Noninvasive Biomarker of Activity for Lupus Nephritis. *J. Am. Soc. Nephrol.* **31**, 1335 (2020).

71. Migliorini, A. *et al.* The antiviral cytokines IFN-α and IFN-β modulate parietal epithelial cells and promote podocyte loss: implications for IFN toxicity, viral glomerulonephritis, and glomerular regeneration. *Am. J. Pathol.* **183**, 431–440 (2013).

72. Gurkan, S. *et al.* Inhibition of type I interferon signalling prevents TLR ligand-mediated proteinuria. *J. Pathol.* **231**, 248–256 (2013).

73. Ferri, D. M. *et al.* Elevated Levels of Interferon-α Act Directly on B Cells to Breach Multiple Tolerance Mechanisms Promoting Autoantibody Production. *Arthritis Rheumatol* **75**, 1542–1555 (2023).

74. Bhargava, R., Li, H. & Tsokos, G. C. Pathogenesis of lupus nephritis: the contribution of immune and kidney resident cells. *Curr. Opin. Rheumatol.* **35**, 107–116 (2023).

75. Zorn-Pauly, L. *et al.* Evaluation of SIGLEC1 in the diagnosis of suspected systemic lupus erythematosus. *Rheumatology* **61**, 3396–3400 (2022).

76. Robinson, G. A. *et al.* Disease-associated and patient-specific immune cell signatures in juvenile-onset systemic lupus erythematosus: patient stratification using a machine-learning approach. *Lancet Rheumatol* **2**, e485–e496 (2020).

77. Blanco, P. *et al.* Increase in activated CD8+ T lymphocytes expressing perforin and granzyme B correlates with disease activity in patients with systemic lupus erythematosus. *Arthritis Rheum.* **52**, 201–211 (2005).

78. McKinney, E. F. *et al.* A CD8+ T cell transcription signature predicts prognosis in autoimmune disease. *Nat. Med.* **16**, 586–91, 1p following 591 (2010).

79. Jonsson, A. H. *et al.* Granzyme K^+ CD8 T cells form a core population in inflamed human tissue. *Sci. Transl. Med.* **14**, eabo0686 (2022).

80. Haas, C., Ryffel, B. & Le Hir, M. IFN-gamma receptor deletion prevents autoantibody production and glomerulonephritis in lupus-prone (NZB x NZW)F1 mice. *J. Immunol.* **160**, 3713–3718 (1998).

81. Leatherwood, C. *et al.* Clinical characteristics and renal prognosis associated with interstitial fibrosis and tubular atrophy (IFTA) and vascular injury in lupus nephritis biopsies. *Semin. Arthritis Rheum.* **49**, 396–404 (2019).

82. Deonaraine, K. K. *et al.* Safety of procuring research tissue during a clinically indicated kidney biopsy from patients with lupus: data from the Accelerating Medicines Partnership RA/SLE Network. *Lupus Sci Med* **8**, (2021).

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

84. Petri, M. *et al.* Derivation and validation of the Systemic Lupus International Collaborating Clinics classification criteria for systemic lupus erythematosus. *Arthritis Rheum.* **64**, 2677–2686 (2012).

85. Petri, M. *et al.* Combined oral contraceptives in women with systemic lupus erythematosus. *N. Engl. J. Med.* **353**, 2550–2558 (2005).

86. Thanou, A. *et al.* Scoring systemic lupus erythematosus (SLE) disease activity with simple, rapid outcome measures. *Lupus Sci Med* **6**, e000365 (2019).

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

88. Inamo, J. *et al.* Deep immunophenotyping reveals circulating activated lymphocytes in individuals at risk for rheumatoid arthritis. *bioRxiv* (2023) doi:10.1101/2023.07.03.547507.

89. Finck, R. *et al.* Normalization of mass cytometry data with bead standards. *Cytometry A* **83**, 483–494 (2013).

90. Chevrier, S. *et al.* Compensation of Signal Spillover in Suspension and Imaging Mass Cytometry. *Cell Syst* **6**, 612-620.e5 (2018).

91. Zunder, E. R. *et al.* Palladium-based mass tag cell barcoding with a doublet-filtering scheme and single-cell deconvolution algorithm. *Nat. Protoc.* **10**, 316–333 (2015).

92. Stassen, S. V. *et al.* PARC: ultrafast and accurate clustering of phenotypic data of millions of single cells. *Bioinformatics* **36**, 2778–2786 (2020).

93. Nowicka, M. *et al.* CyTOF workflow: differential discovery in high-throughput high-dimensional cytometry datasets. *F1000Res.* **6**, 748 (2017).

94. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587.e29 (2021).

95. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

96. Weber, L. M., Nowicka, M., Soneson, C. & Robinson, M. D. diffcyt: Differential discovery in high-dimensional cytometry via high-resolution clustering. *Commun Biol* **2**, 183 (2019).

97. Wilkerson, M. D. & Hayes, D. N. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* **26**, 1572–1573 (2010).

MAIN FIGURES

Fig. 1

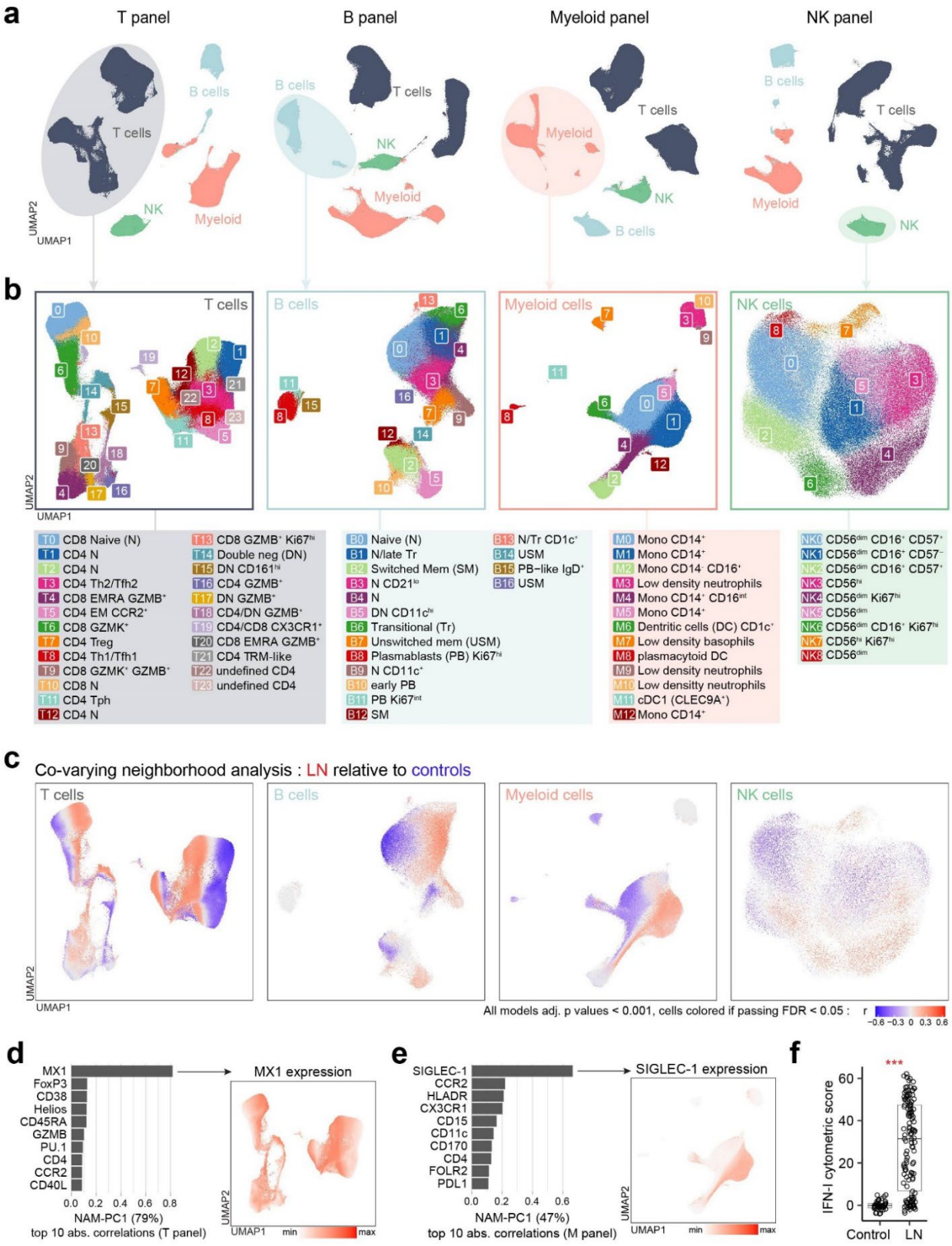

a
T panel
B panel
Myeloid panel
NK panel
T cells
B cells
NK
Myeloid
UMAP2
UMAP1
b
T cells
B cells
Myeloid cells
NK cells
T0 CD8 Naive (N)
T1 CD4 N
T2 CD4 N
T3 CD4 Th2/Tfh2
T4 CD8 EMRA GZMB+
T5 CD4 EM CCR2+
T6 CD8 GZMK+
T7 CD4 Treg
T8 CD4 Th1/Tfh1
T9 CD8 GZMK+ GZMB+
T10 CD8 N
T11 CD4 Tph
T12 CD4 N
T13 CD8 GZMB+ Ki67hi
T14 Double neg (DN)
T15 DN CD161hi
T16 CD4 GZMB+
T17 DN GZMB+
T18 CD4/DN GZMB+
T19 CD4/CD8 CX3CR1+
T20 CD8 EMRA GZMB+
T21 CD4 TRM-like
T22 undefined CD4
T23 undefined CD4
B0 Naive (N)
B1 N/late Tr
B2 Switched Mem (SM)
B3 N CD21lo
B4 N
B5 DN CD11chi
B6 Transitional (Tr)
B7 Unswitched mem (USM)
B8 Plasmablasts (PB) Ki67hi
B9 N CD11c+
B10 early PB
B11 PB Ki67int
B12 SM
B13 N/Tr CD1c+
B14 USM
B15 PB−like IgD+
B16 USM
M0 Mono CD14+
M1 Mono CD14+
M2 Mono CD14- CD16+
M3 Low density neutrophils
M4 Mono CD14+ CD16int
M5 Mono CD14+
M6 Dentritic cells (DC) CD1c+
M7 Low density basophils
M8 plasmacytoid DC
M9 Low density neutrophils
M10 Low densitty neutrophils
M11 cDC1 (CLEC9A+)
M12 Mono CD14+
NK0 CD56dim CD16+ CD57+
NK1 CD56dim CD16+ CD57−
NK2 CD56dim CD16+ CD57+
NK3 CD56hi
NK4 CD56dim Ki67hi
NK5 CD56dim
NK6 CD56dim CD16+ Ki67hi
NK7 CD56hi Ki67hi
NK8 CD56dim
c
Co-varying neighborhood analysis : LN relative to controls
All models adj. p values < 0.001, cells colored if passing FDR < 0.05 :
r
−0.6 −0.3 0 0.3 0.6
d
MX1
FoxP3
CD38
Helios
CD45RA
GZMB
PU.1
CD4
CCR2
CD40L
0.0 0.2 0.4 0.6 0.8
NAM-PC1 (79%)
top 10 abs. correlations (T panel)
MX1 expression
min
max
e
SIGLEC-1
CCR2
HLADR
CX3CR1
CD15
CD11c
CD170
CD4
FOLR2
PDL1
0.0 0.2 0.4 0.6
NAM-PC1 (47%)
top 10 abs. correlations (M panel)
SIGLEC-1 expression
f
***
IFN-I cytometric score
0
20
40
60
Control
LN

Fig.2

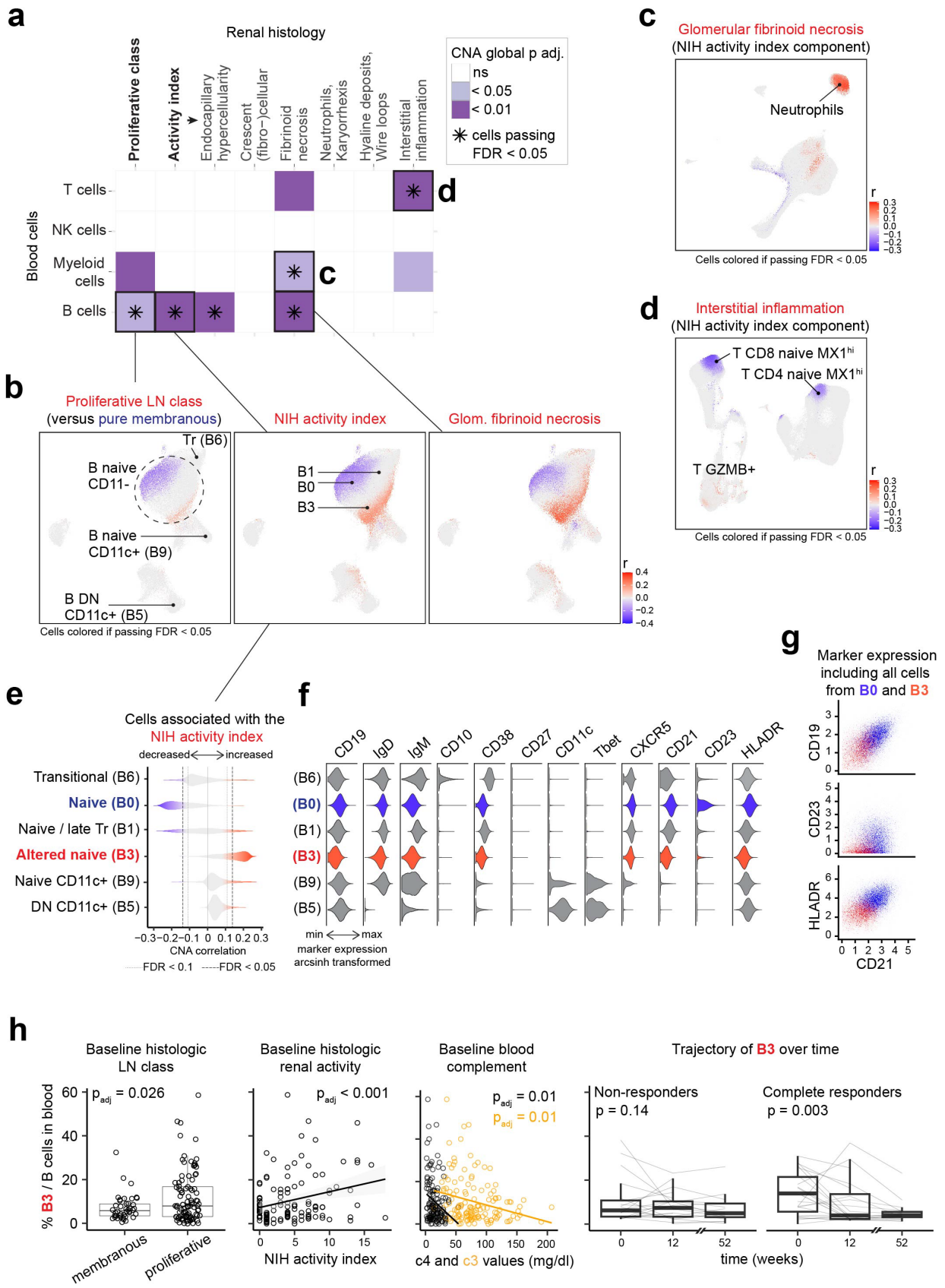
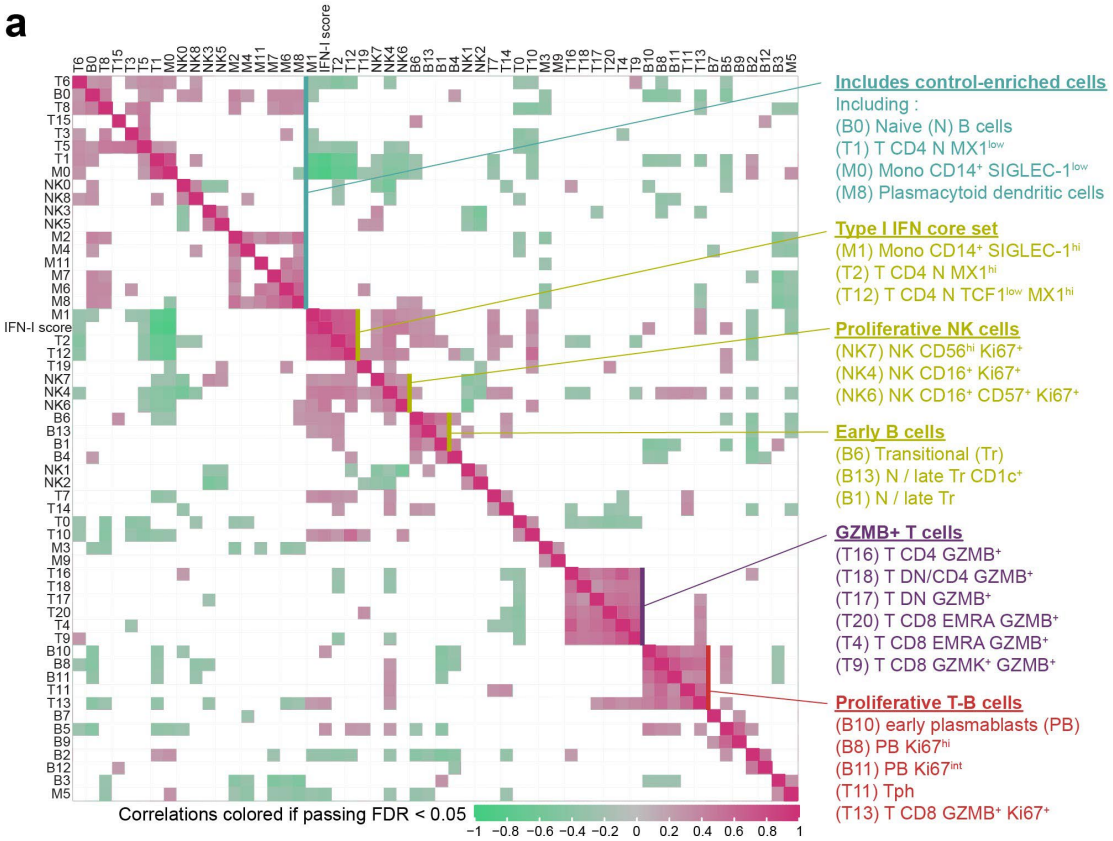
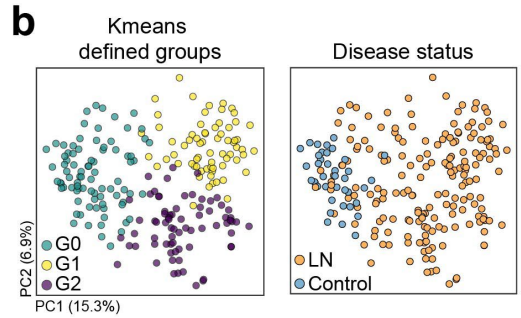
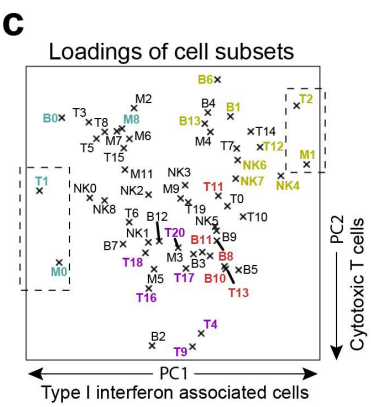
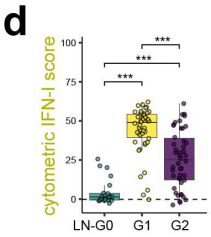

a
Renal histology
Proliferative class
Activity index
Endocapillary hypercellularity
Crescent (fibro−)cellular
Fibrinoid necrosis
Neutrophils, Karyorrhexis
Hyaline deposits, Wire loops
Interstitial inflammation
CNA global p adj.
ns
< 0.05
< 0.01
cells passing FDR < 0.05
Blood cells
T cells
NK cells
Myeloid cells
B cells
c
d
b
Proliferative LN class
(versus pure membranous)
Tr (B6)
B naive CD11-
B naive CD11c+ (B9)
B DN CD11c+ (B5)
Cells colored if passing FDR < 0.05
NIH activity index
B1
B0
B3
Glom. fibrinoid necrosis
r
c
Glomerular fibrinoid necrosis
(NIH activity index component)
Neutrophils
Cells colored if passing FDR < 0.05
d
Interstitial inflammation
(NIH activity index component)
T CD8 naive MX1hi
T CD4 naive MX1hi
T GZMB+
Cells colored if passing FDR < 0.05
e
Cells associated with the NIH activity index
decreased
increased
Transitional (B6)
Naive (B0)
Naive / late Tr (B1)
Altered naive (B3)
Naive CD11c+ (B9)
DN CD11c+ (B5)
CNA correlation
FDR < 0.1
FDR < 0.05
f
CD19
IgD
IgM
CD10
CD38
CD27
CD11c
Tbet
CXCR5
CD21
CD23
HLADR
min
max
marker expression
arcsinh transformed
g
Marker expression including all cells from B0 and B3
CD19
CD23
HLADR
CD21
h
Baseline histologic LN class
% B3 / B cells in blood
p_adj = 0.026
membranous
proliferative
Baseline histologic renal activity
p_adj < 0.001
NIH activity index
Baseline blood complement
p_adj = 0.01
p_adj = 0.01
c4 and c3 values (mg/dl)
Trajectory of B3 over time
Non-responders
p = 0.14
Complete responders
p = 0.003
time (weeks)

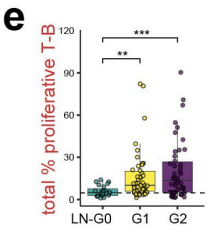
Fig. 3

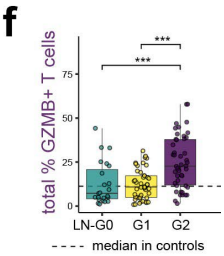

a
IFN-I score
Includes control-enriched cells
Including :
(B0) Naive (N) B cells
(T1) T CD4 N MX1low
(M0) Mono CD14+ SIGLEC-1low
(M8) Plasmacytoid dendritic cells
Type I IFN core set
(M1) Mono CD14+ SIGLEC-1hi
(T2) T CD4 N MX1hi
(T12) T CD4 N TCF1low MX1hi
Proliferative NK cells
(NK7) NK CD56hi Ki67+
(NK4) NK CD16+ Ki67+
(NK6) NK CD16+ CD57+ Ki67+
Early B cells
(B6) Transitional (Tr)
(B13) N / late Tr CD1c+
(B1) N / late Tr
GZMB+ T cells
(T16) T CD4 GZMB+
(T18) T DN/CD4 GZMB+
(T17) T DN GZMB+
(T20) T CD8 EMRA GZMB+
(T4) T CD8 EMRA GZMB+
(T9) T CD8 GZMK+ GZMB+
Proliferative T-B cells
(B10) early plasmablasts (PB)
(B8) PB Ki67hi
(B11) PB Ki67int
(T11) Tph
(T13) T CD8 GZMB+ Ki67+
Correlations colored if passing FDR < 0.05
−1 −0.8 −0.6 −0.4 −0.2 0 0.2 0.4 0.6 0.8 1


b
Kmeans defined groups
Disease status
G0
G1
G2
LN
Control
PC2 (6.9%)
PC1 (15.3%)


c
Loadings of cell subsets
PC2
Cytotoxic T cells
PC1
Type I interferon associated cells


d
cytometric IFN-I score
LN-G0
G1
G2


e
total % proliferative T-B
LN-G0
G1
G2


f
total % GZMB+ T cells
LN-G0
G1
G2
median in controls

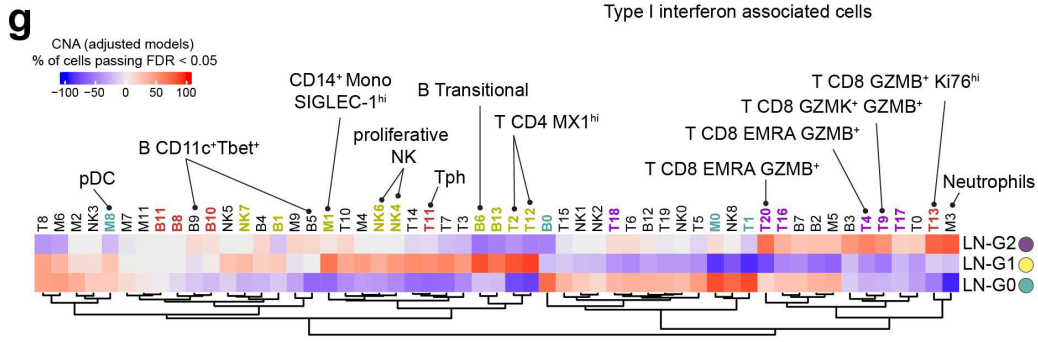

g
CNA (adjusted models)
% of cells passing FDR < 0.05
−100 −50 0 50 100
pDC
B CD11c+Tbet+
CD14+ Mono SIGLEC-1hi
proliferative NK
Tph
B Transitional
T CD4 MX1hi
T CD8 EMRA GZMB+
T CD8 GZMK+ GZMB+
T CD8 GZMB+ Ki76hi
Neutrophils
LN-G2
LN-G1
LN-G0

Fig. 4

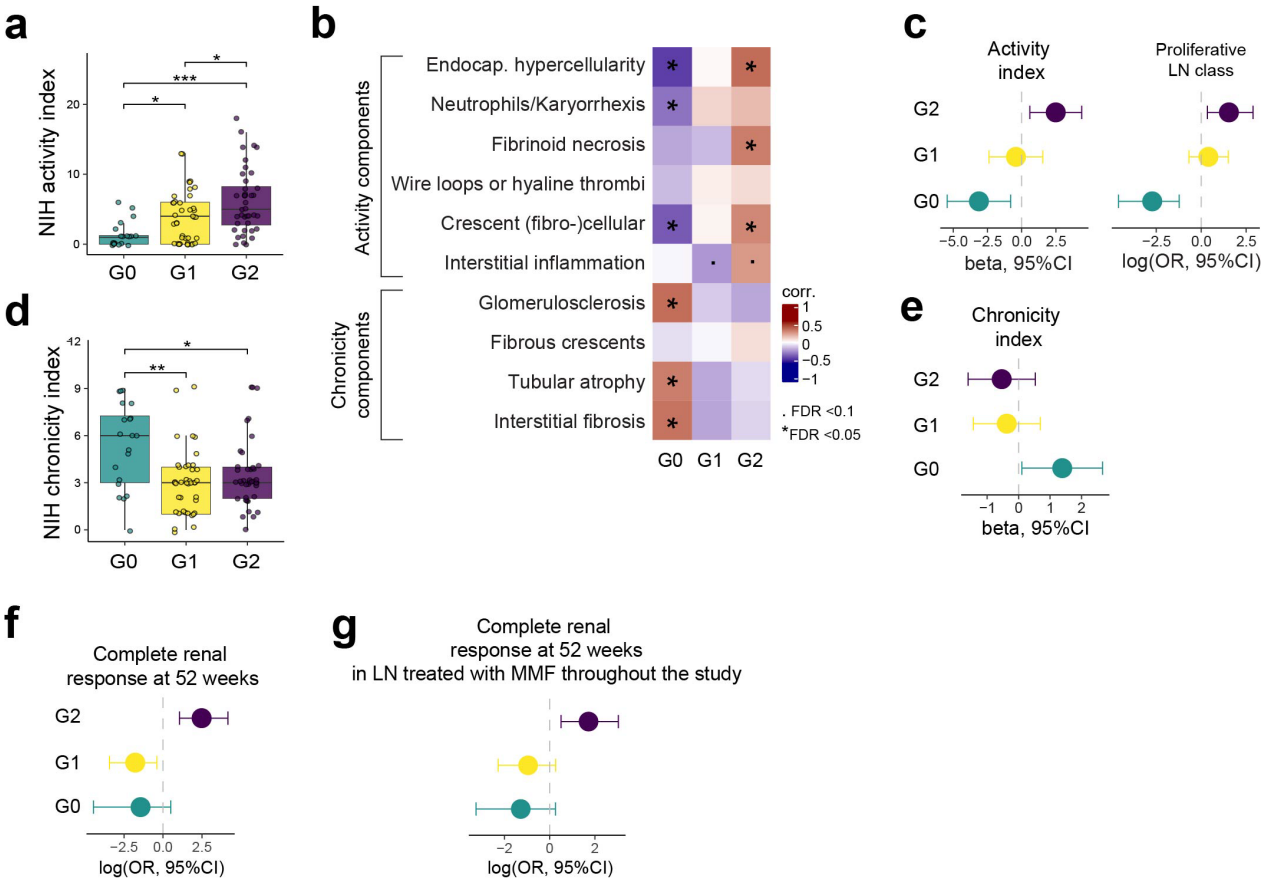
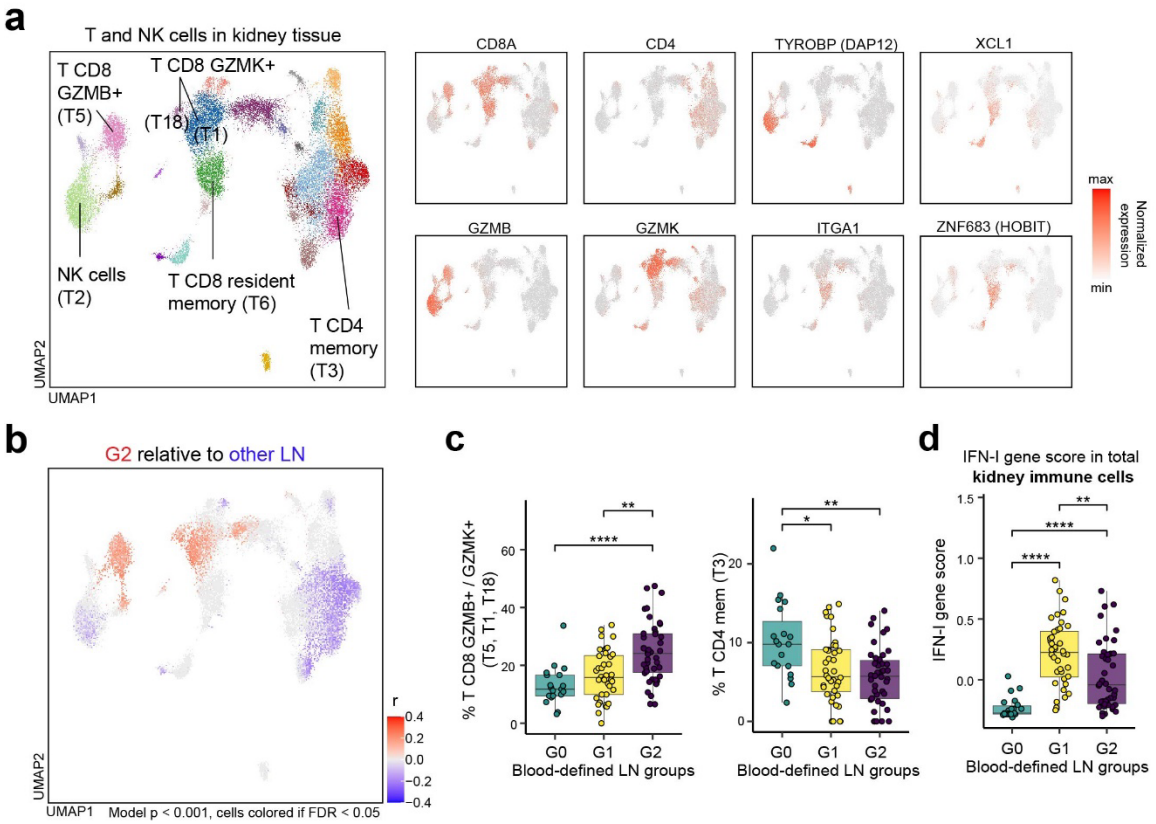

a
NIH activity index
G0
G1
G2
b
Activity components
Chronicity components
Endocap. hypercellularity
Neutrophils/Karyorrhexis
Fibrinoid necrosis
Wire loops or hyaline thrombi
Crescent (fibro-)cellular
Interstitial inflammation
Glomerulosclerosis
Fibrous crescents
Tubular atrophy
Interstitial fibrosis
corr.
. FDR <0.1
*FDR <0.05
c
Activity index
Proliferative LN class
beta, 95%CI
log(OR, 95%CI)
d
NIH chronicity index
e
Chronicity index
f
Complete renal response at 52 weeks
log(OR, 95%CI)
g
Complete renal response at 52 weeks in LN treated with MMF throughout the study

Fig. 5


a
T and NK cells in kidney tissue
T CD8 GZMB+ (T5)
T CD8 GZMK+ (T18) (T1)
NK cells (T2)
T CD8 resident memory (T6)
T CD4 memory (T3)
UMAP2
UMAP1
CD8A
CD4
TYROBP (DAP12)
XCL1
GZMB
GZMK
ITGA1
ZNF683 (HOBIT)
max
min
Normalized expression
b
G2 relative to other LN
r
0.4
0.2
0.0
−0.2
−0.4
Model p < 0.001, cells colored if FDR < 0.05
c
% T CD8 GZMB+ / GZMK+ (T5, T1, T18)
% T CD4 mem (T3)
G0
G1
G2
Blood-defined LN groups
d
IFN-I gene score in total kidney immune cells
IFN-I gene score

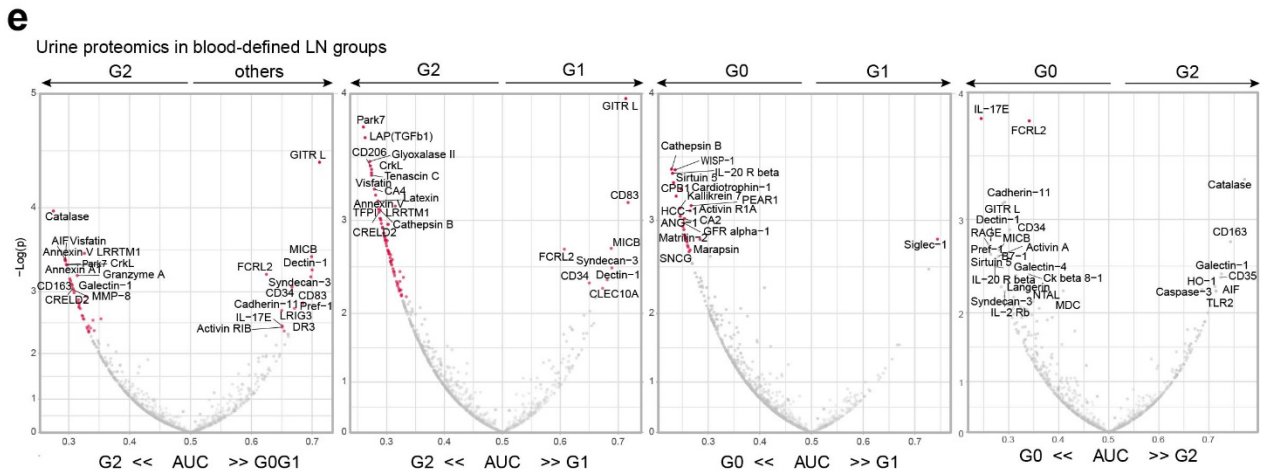
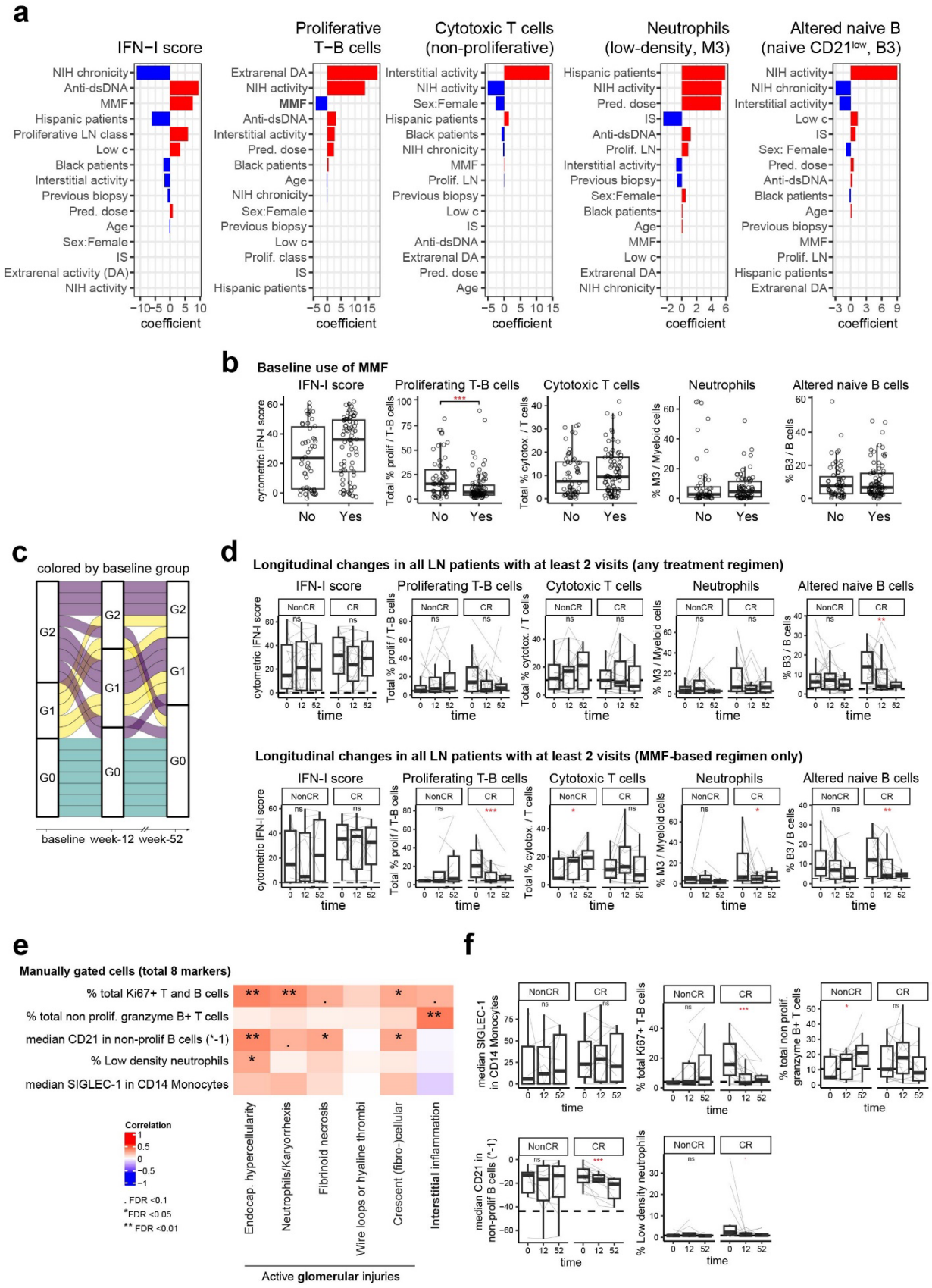

e
Urine proteomics in blood-defined LN groups
G2
others
G1
G0
−Log(p)
G2 << AUC >> G0G1
G2 << AUC >> G1
G0 << AUC >> G1
G0 << AUC >> G2

Fig. 6


a
IFN−I score
Proliferative T−B cells
Cytotoxic T cells (non-proliferative)
Neutrophils (low-density, M3)
Altered naive B (naive CD21low, B3)
coefficient
b
Baseline use of MMF
IFN-I score
Proliferating T-B cells
Cytotoxic T cells
Neutrophils
Altered naive B cells
No
Yes
c
colored by baseline group
baseline
week-12
week-52
d
Longitudinal changes in all LN patients with at least 2 visits (any treatment regimen)
Longitudinal changes in all LN patients with at least 2 visits (MMF-based regimen only)
NonCR
CR
time
e
Manually gated cells (total 8 markers)
% total Ki67+ T and B cells
% total non prolif. granzyme B+ T cells
median CD21 in non-prolif B cells (*-1)
% Low density neutrophils
median SIGLEC-1 in CD14 Monocytes
Correlation
. FDR <0.1
* FDR <0.05
** FDR <0.01
Endocap. hypercellularity
Neutrophils/Karyorrhexis
Fibrinoid necrosis
Wire loops or hyaline thrombi
Crescent (fibro-)cellular
Interstitial inflammation
Active glomerular injuries
f

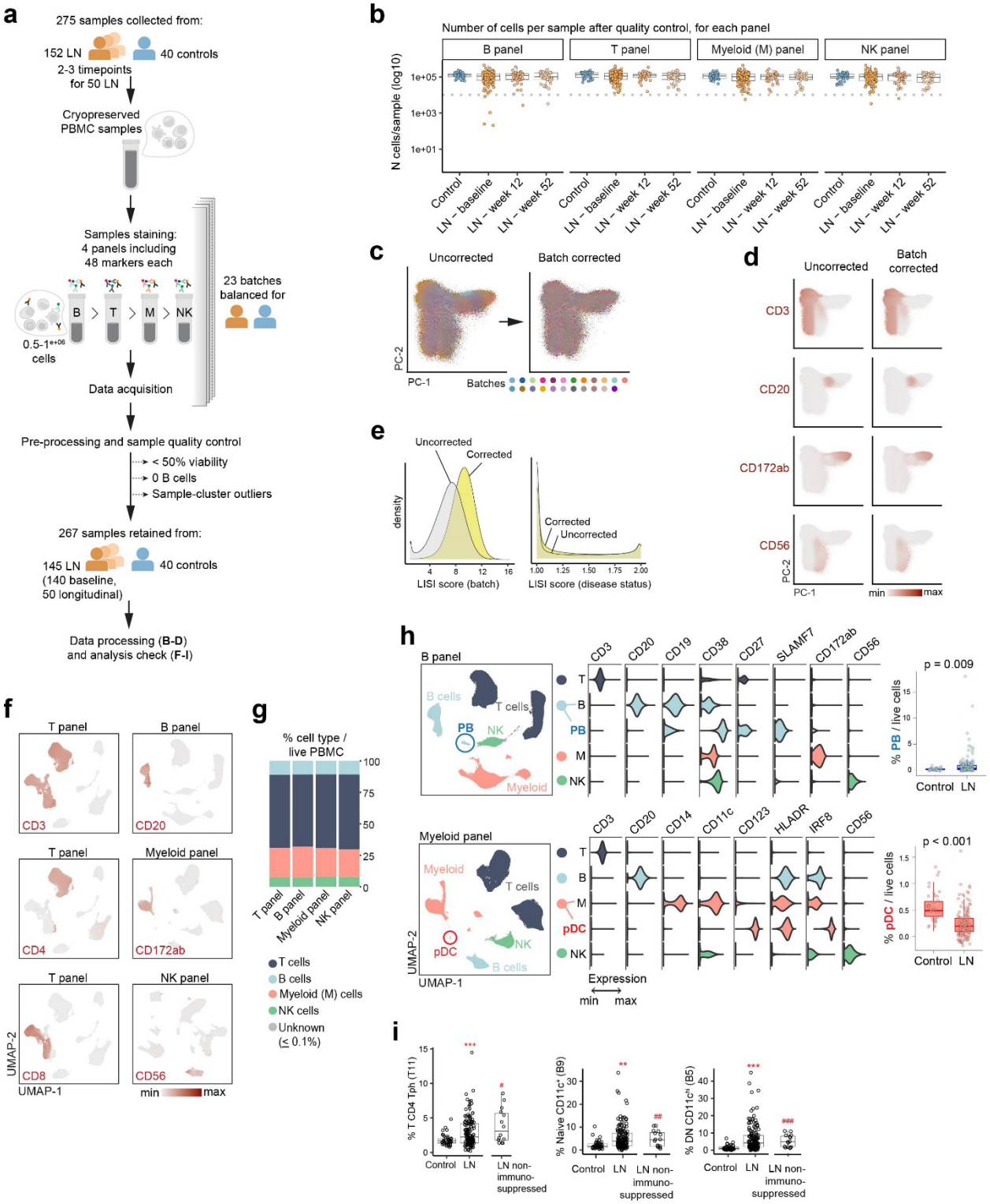
EXTENDED DATA

Extended Data Table 1. Baseline characteristics and renal response at week 52 of LN patients stratified by immunophenotype subgroups (n=115 LN with samples analyzed with 4 panels)

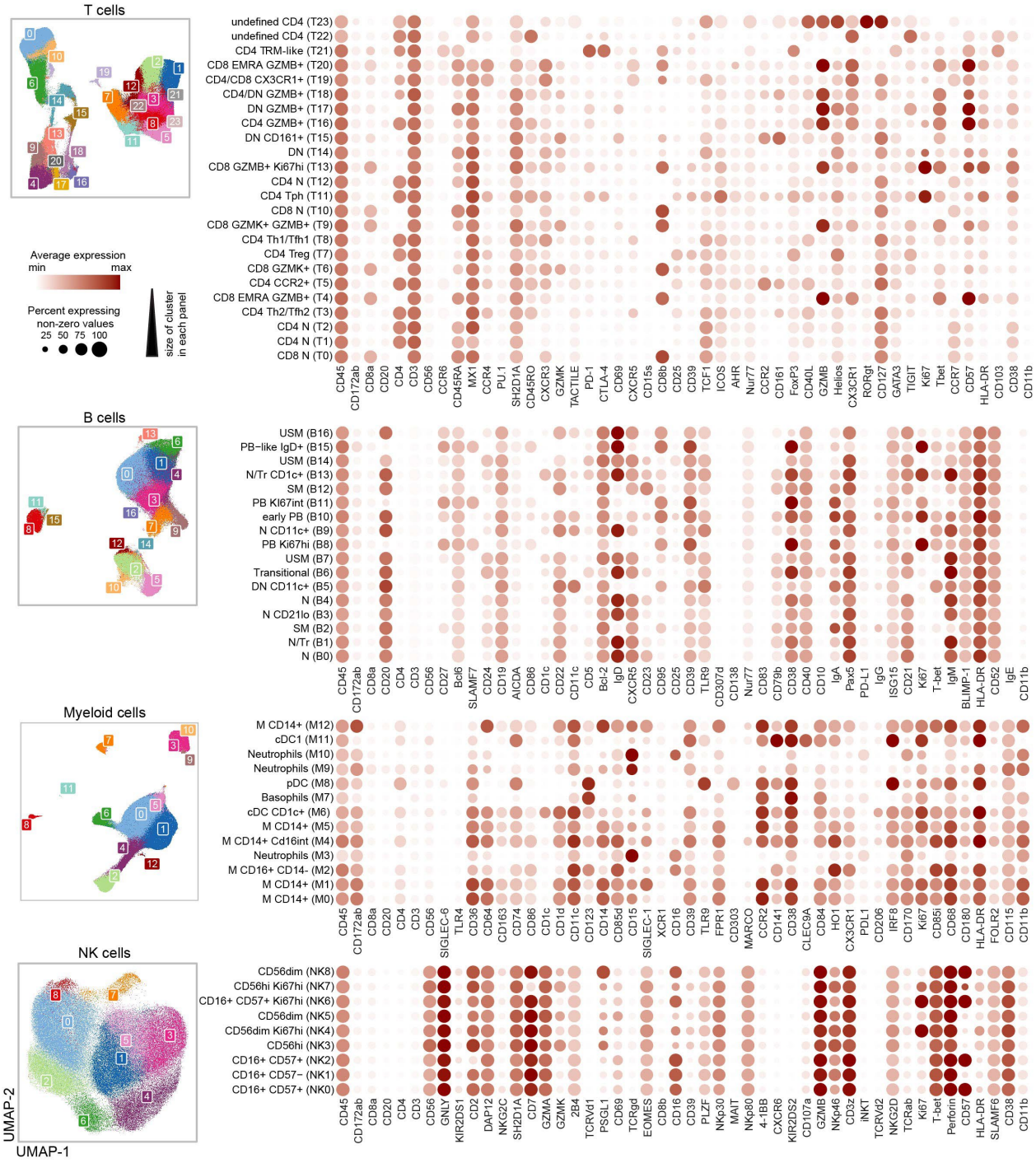
	LN-G0 "control-like"	**LN-G1 "IFN-I high"**	**LN-G2 "cytotoxic T"**	***global p val.* ***	***G1-G2 p val.*** [£]
N	**23**	**46**	**46**		
Age, median (IQR)	41 (31-46)	30 (24-39)	33 (26-44)	0.007	0.06
Female	19 (83)	37 (80)	42 (91)	0.31	
Ethnicity and race					
Hispanic or Latino	11 (48)	7 (15)	19 (41)	0.006	0.01
Black	10/19 (53)	20/43 (47)	23/41 (56)	0.70	
Previous renal biopsy	20 (87)	34 (74)	29 (63)	0.10	
Proliferative class (III or IV +/- V)	9 (39)	31 (67)	39 (85)	<0.001	0.09
Activity index[1], median (IQR)	1 (0-1)	4 (0-6)	5 (3-8)	<0.001	0.03
Chronicity index[1], median (IQR)	6 (3-7)	3 (1-4)	3 (2-4)	0.005	0.26
Serum creatinine mg/ml, median (IQR)	1.0 (0.8-1.5)	0.8 (0.7-1.0)	1.0 (0.8-1.3)	0.04	0.04
UPCR, median (IQR)	3.4 (1.6-4.8)	1.4 (1.0-2.7)	2.2 (1.5-4.3)	0.02	0.009
Presence of cSLEDAI extrarenal[2]	5 (26)	16 (37)	23 (50)	0.05	
Positive anti-dsDNA	5 (22)	38/45 (84)	36 (78)	<0.001	0.63
Low C3 and/or low C4	5 (22)	31/45 (69)	34 (74)	<0.001	0.62
Medication					
Hydroxychloroquine	19 (86)	42 (91)	34 (74)	0.08	
Prednisone, med (IQR)[3]	0 (0-10)	0 (0-7)	20 (5-40)	<0.001	<0.001
Any immunosuppressants[4]	17 (77)	39 (85)	30 (65)	0.08	
Mycophenolate	13 (59)	32 (70)	23 (50)	0.25	
Renal response at week 52[5]					
CR, PR, NR	3(18), 3(18), 11(65)	6(18), 9(27), 19(56)	15(41), 9(24), 13(35)	0.04	0.03

UPCR = urine protein/creatinine ratio. CR, PR, NR = complete, partial and non-responder. Data is presented as N (%) unless specified otherwise. The proportions, median and IQR are calculated on the total number of subject in the group unless specified otherwise. *Chi-square, Fisher exact or Kruskal-Wallis tests when appropriate to test for differences across the three groups. [£]Chi-square, Fisher exact or Mann-Whitney U test to test for differences between G1 and G2. [1]NIH activity and chronicity indices available in n=100. [2]Presence of any clinical extrarenal features of the SLE Disease Activity Index (SLEDAI). [3]Predisone dose or equivalence at baseline in n=110, excluding five patients receiving intravenous methylprednisolone ('pulse') dose ranging from 500-1000mg. [4]Includes any immunosuppressants and biologics (azathioprine, tacrolimus, cyclosporine, methotrexate, abatacept or belimumab). [5]Renal response was determined at week 52 if baseline UPCR was $\geq$ 1 (n=88).

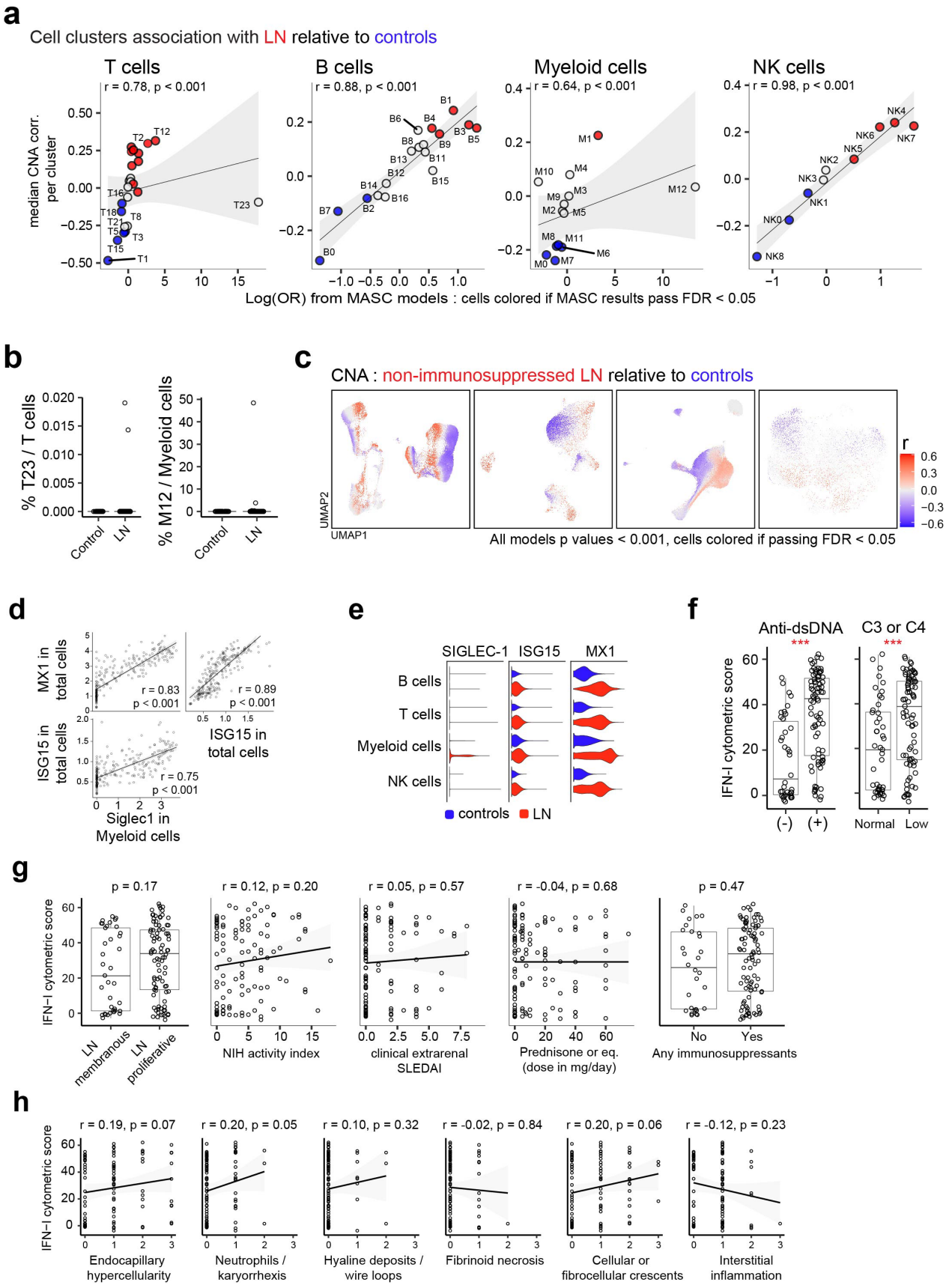
Extended Data Fig. 1



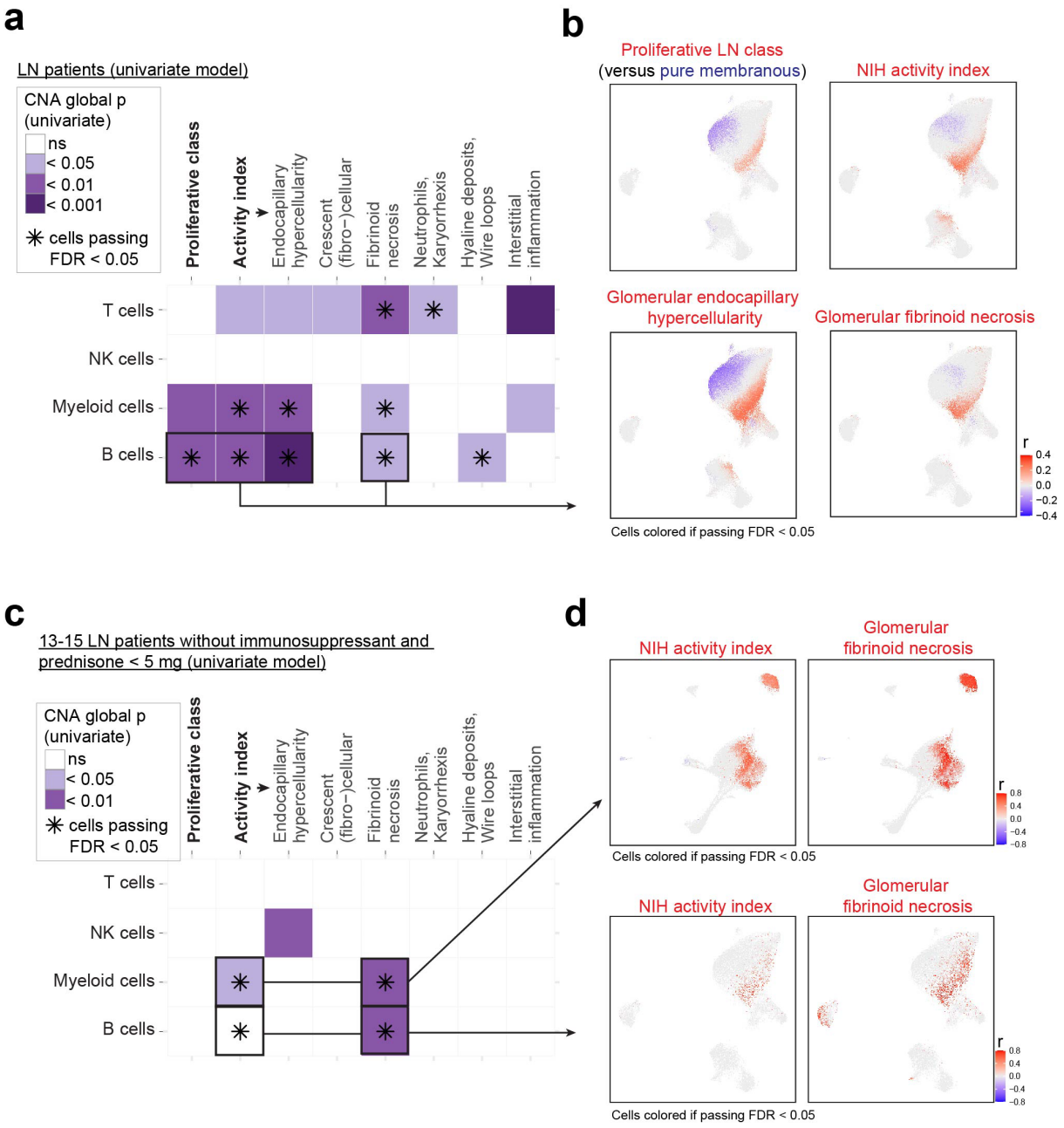
Extended Data Fig. 2



Extended Data Fig. 3

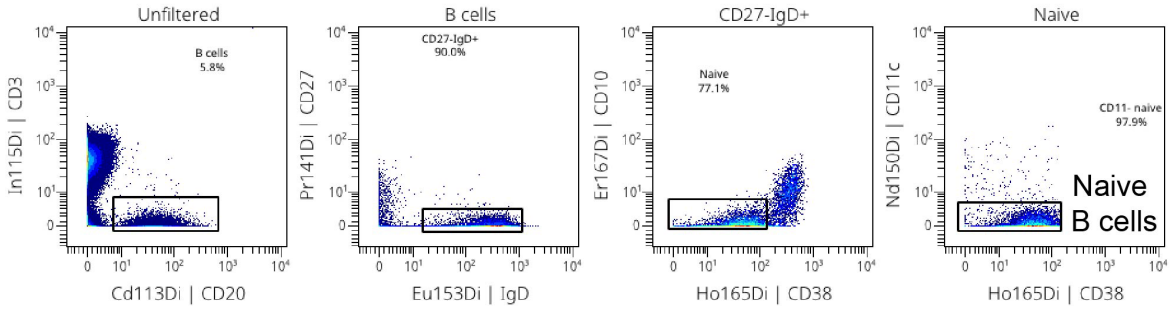


Extended Data Fig. 4

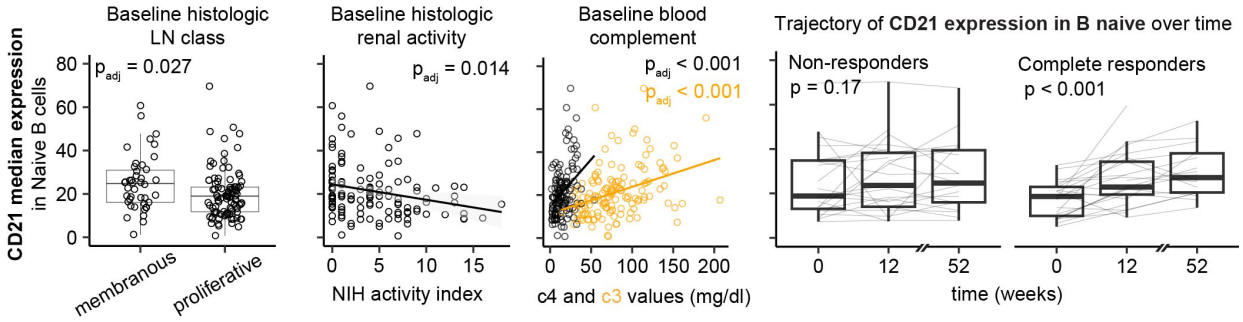


Extended Data Fig. 5

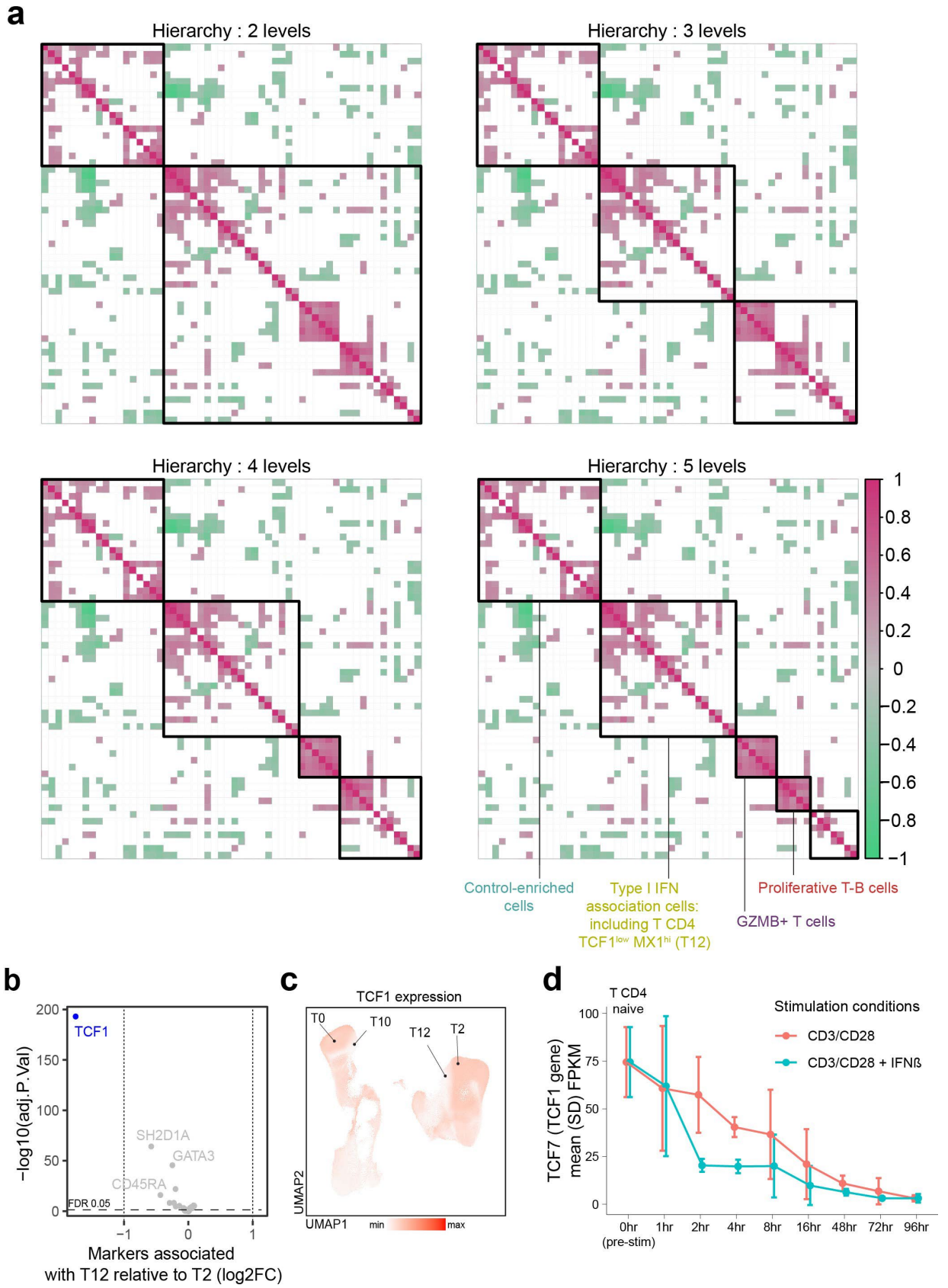
a



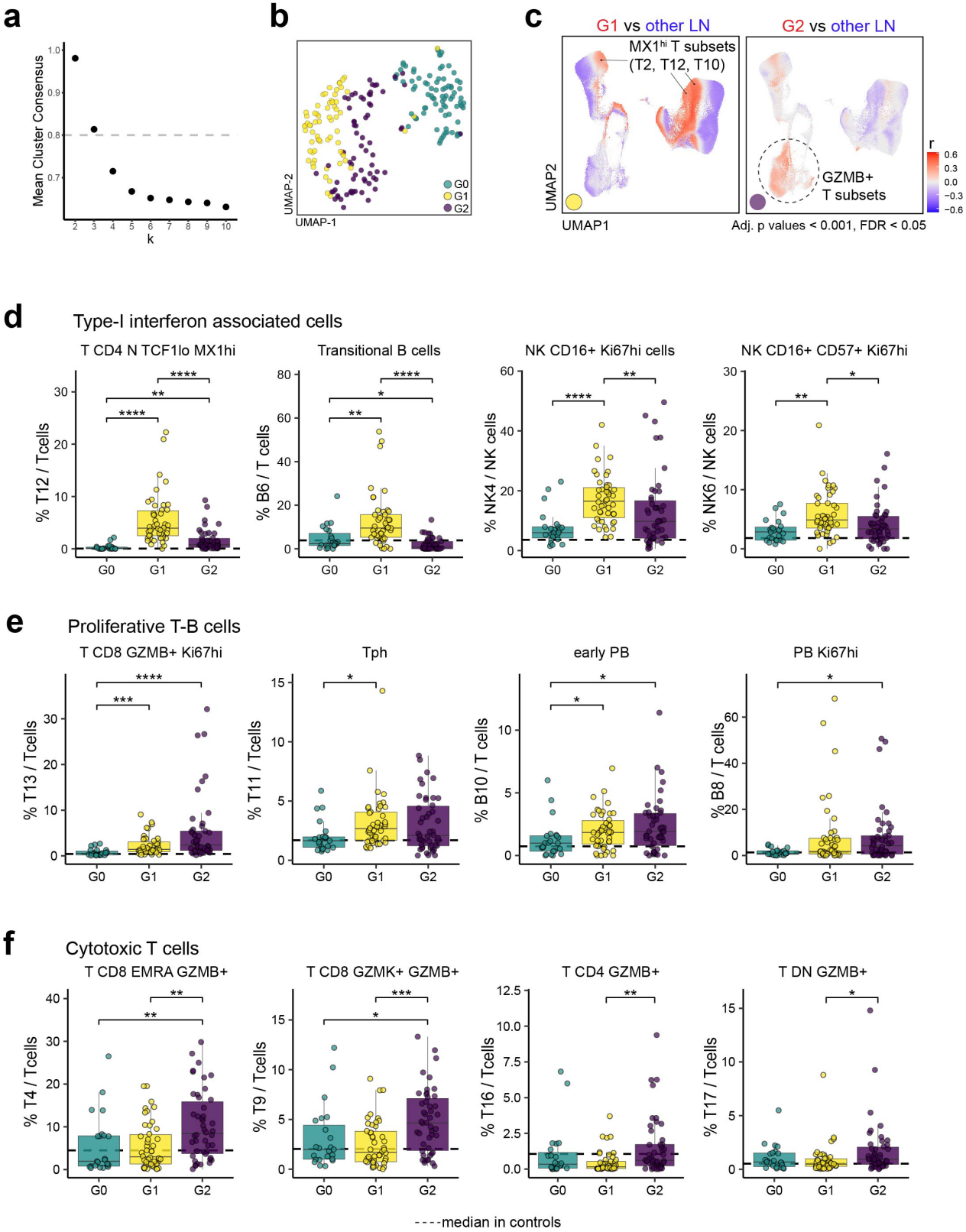
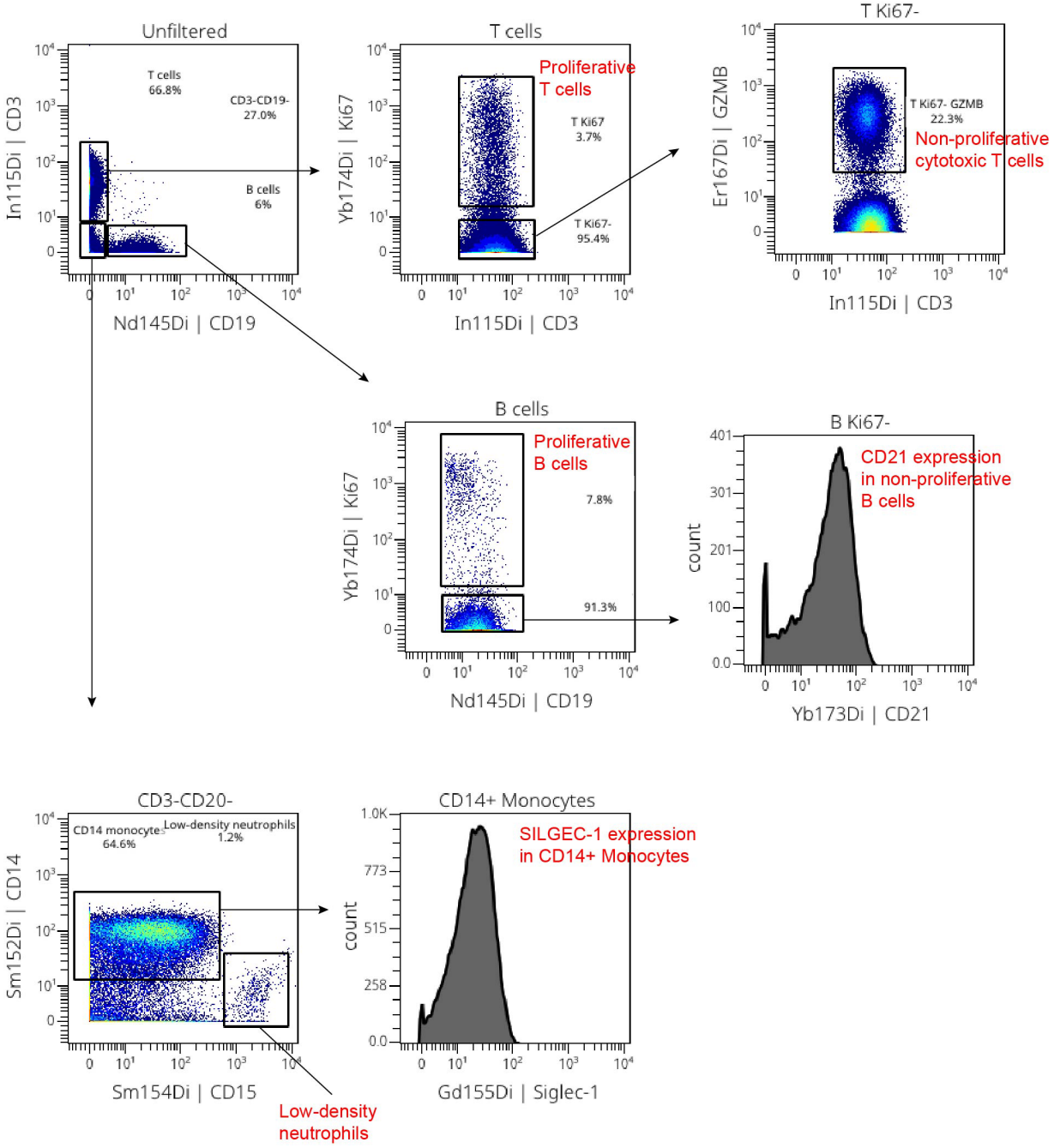
b



Extended Data Fig. 6



Extended Data Fig. 7



Extended Data Fig. 8.