Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Integrated serum proteomic and metabolomic profiling prioritises CSF-1 and PD-L1 as candidate circulating biomarkers associated with lupus nephritis: a two-stage cross-sectional study

Zhongyu Wang
Nanjing Normal University
Tiancheng Li
Nanjing Normal University
Haihui Xing
Nanjing Gaochun Hospital of Traditional Chinese Medicine
Wenjing Liu
Nanjing Drum Tower Hospital
Shiting Xie
Nanjing Drum Tower Hospital
Fei Chen
Nanjing Drum Tower Hospital
Yue Tao
Nanjing Drum Tower Hospital
Zhiyang Li
Nanjing Normal University
Ping Yang
pingyang@njglyy.com
Nanjing Normal University

Research Article

Keywords: Lupus nephritis, systemic lupus erythematosus, CSF-1, PD-L1, inflammatory proteomics, metabolomics, ELISA, single-cell transcriptomics

Posted Date: November 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11629375/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Lupus nephritis (LN) contributes substantially to morbidity in systemic lupus erythematosus (SLE), but routine clinical and serological markers incompletely characterise renal involvement. We integrated serum inflammatory proteomics and metabolomics to identify LN-associated protein candidates and assess selected proteins by ELISA.

Methods

This single-centre cross-sectional study included a discovery cohort (n = 73; 22 healthy controls [HCs], 22 non-renal SLE, 29 LN) and an ELISA follow-up cohort (n = 79; 28 HCs, 25 non-renal SLE, 26 LN). Discovery sera underwent Olink Target 96 Inflammation profiling and untargeted metabolomics. Differential abundance and weighted gene co-expression network analyses prioritised candidates for ELISA. Discrimination and clinical associations were assessed by receiver operating characteristic and Spearman correlation analyses; public single-cell datasets were explored for candidate-gene expression.

Results

Four proteins were prioritised: CSF-1, PD-L1, CX3CL1, and SULT1A1. In the discovery cohort, AUCs for LN versus non-renal SLE were 0.857, 0.738, 0.784, and 0.743, respectively. In the follow-up cohort, CSF-1 and PD-L1 were elevated in LN, with AUCs of 0.890 (95% CI, 0.795–0.986) and 0.883 (0.791–0.975); CX3CL1 and SULT1A1 showed weaker discrimination. CSF-1 correlated with 24-hour urinary protein excretion in both cohorts (Spearman's ρ = 0.768 and 0.687). No candidate consistently outperformed C3 or C4. Four metabolomic pathways remained significant after Holm correction; taurine was higher in LN than in both comparison groups. Exploratory single-cell analyses showed candidate-gene transcripts across renal and immune populations.

Conclusions

CSF-1 and PD-L1 are circulating candidates associated with LN, although their renal specificity and incremental clinical value require prospective evaluation.

Background

Systemic lupus erythematosus (SLE) is a heterogeneous autoimmune disease characterized by dysregulated innate and adaptive immune responses that can cause inflammatory injury across multiple organ systems[1–3]. Among its major organ manifestations, lupus nephritis (LN) is a leading contributor to chronic kidney disease, impaired renal function and long-term morbidity, despite advances in

immunosuppressive and targeted therapies[4–6]. Renal involvement is currently assessed using a combination of proteinuria, urinary sediment, serum creatinine, estimated glomerular filtration rate (eGFR), complement levels and anti-double-stranded DNA antibodies. Although essential to clinical evaluation, these measures incompletely capture renal inflammatory processes and may not reliably distinguish active inflammation from established chronic damage. Kidney biopsy remains the reference standard for defining histological class, activity and chronicity, but its invasive nature limits its suitability for repeated monitoring[4, 7, 8]. These limitations underscore the need for accessible circulating biomarkers that can complement conventional clinical and serological assessments. Evaluating whether such biomarkers are specifically associated with renal involvement also requires distinguishing kidney-related molecular alterations from the systemic immune activation characteristic of SLE. Because circulating inflammatory proteins may reflect immune activity across multiple tissues, comparisons between patients with LN and those with SLE without renal involvement are central to assessing their potential relevance to kidney disease; healthy controls, in turn, provide a reference for broader SLE-associated molecular changes.

Proteomic and metabolomic approaches have increasingly been used to characterise the molecular heterogeneity of SLE and identify potential biomarkers of renal involvement. Serum proximity extension assay (PEA)-based profiling has revealed inflammatory protein signatures associated with LN and cumulative organ damage[9]. Studies integrating urinary proteomics with renal single-cell transcriptomics have linked non-invasive molecular measurements to intrarenal inflammatory pathways, particularly interferon-associated responses[10–12]. Subsequent work has linked urinary protein signatures to histological class, activity, chronicity, and treatment response[13], synthesised evidence for urinary markers with independent validation[14], and evaluated urinary biomarkers against biopsy findings or in prospective cohorts[15, 16]. More recently, serum proteomic studies have identified candidate markers in patients with newly diagnosed LN, while ICAM2 and the macrophage-associated protein CD163 have emerged as potential indicators of renal involvement and disease activity[17–19]. Renal tissue proteomics has also demonstrated molecular heterogeneity within LN that is not fully captured by conventional histopathological classification[20]. In parallel, metabolomic studies have identified alterations in lipid metabolism, amino acid utilisation and cellular bioenergetics in SLE, supporting a role for immunometabolic dysregulation in disease pathophysiology[21, 22]. Serum metabolomic profiling has further been investigated for differentiating LN from non-renal SLE and characterising metabolic changes associated with treatment response[23–25] .

Despite these advances, proteomic and metabolomic profiles have often been investigated separately, leaving the relationship between circulating inflammatory proteins and metabolic alterations in LN incompletely characterized. In our previous study, PEA profiling of extracellular vesicle-enriched fractions identified colony-stimulating factor 1 (CSF-1) and programmed death-ligand 1 (PD-L1) among the inflammatory proteins altered in SLE [26]. Whether these proteins are also associated with renal involvement when measured in unfractionated serum, particularly in comparison with SLE without nephritis, remains unclear. The reproducibility of candidate protein findings across analytical platforms and their relationships with clinical measures of renal involvement also require further investigation. In

addition, defining the cellular context of candidate-gene expression may help interpret circulating protein findings, although serum concentrations alone cannot establish a protein's tissue of origin.

In this study, we integrated serum inflammatory proteomic profiling using the Olink Target 96 Inflammation panel with untargeted metabolomics to investigate molecular alterations associated with LN and prioritise candidate circulating proteins. Differential abundance analysis and weighted gene co-expression network analysis (WGCNA) were used to identify proteins associated with clinical phenotypes and integrated molecular profiles. We then evaluated four prioritised candidates using enzyme-linked immunosorbent assay (ELISA) in a separate follow-up cohort. The primary objective was to assess their ability to distinguish LN from SLE without renal involvement, with healthy controls serving as a reference for broader SLE-associated changes. We also examined associations between candidate proteins and clinical and laboratory measures, characterized relevant metabolomic pathways and protein–metabolite relationships, and explored candidate-gene expression in publicly available single-cell transcriptomic datasets[27]. Together, these analyses were designed to identify circulating inflammatory protein candidates associated with LN and provide a multi-omics framework for their further translational evaluation.

Methods

Study design and participants

This single-centre, cross-sectional biomarker study comprised a serum multi-omics discovery cohort and an ELISA follow-up cohort (Fig. 1), both recruited at Nanjing Drum Tower Hospital, Nanjing, China. The discovery cohort included 73 participants enrolled between May 2023 and January 2024: 22 HCs, 22 patients with SLE without clinically documented LN, and 29 patients with LN. The follow-up cohort comprised 79 unique participants (28 HCs, 25 patients with SLE without LN, and 26 patients with LN). Patients with SLE met the 2019 European League Against Rheumatism/American College of Rheumatology classification criteria[28]. Individuals with a history of malignancy, diabetes mellitus, or another concurrent autoimmune disease were excluded. HCs were recruited from the hospital's Health Examination Centre and had no history of autoimmune disease, malignancy, or diabetes mellitus. Discovery samples came from a serum biobank previously used to study extracellular vesicle-enriched fractions[26]. Here, the corresponding unfractionated serum samples underwent inflammatory proteomic and untargeted metabolomic profiling. The study was approved by the Institutional Ethics Committee of Nanjing Drum Tower Hospital (approval no. 2023 – 568). All participants provided written informed consent before enrolment, and the study was conducted in accordance with the Declaration of Helsinki.

Definition of renal involvement and clinical data collection

LN was classified histologically when definitive kidney biopsy findings were available. In patients without definitive biopsy evidence, LN was operationally classified as clinical LN if 24-hour urinary protein excretion was > 0.5 g/day in the context of established SLE, using the guideline threshold that prompts consideration of kidney biopsy[7]. These cases were not considered biopsy-confirmed LN. Pathology reports were available for all 29 discovery-cohort patients and 13 of 26 follow-up-cohort patients with LN; two follow-up reports stated that LN could not be excluded and were not considered definitive histological confirmation. Renal pathological classes are reported as recorded in Tables 1 and 2. Patients were classified as having SLE without LN if they had no documented LN diagnosis and did not meet the proteinuria criterion; available records showed 24-hour urinary protein excretion < 500 mg/day or negative urinary protein testing. Clinical and laboratory data were obtained from medical records and study registries. Disease activity was assessed using the SLE Disease Activity Index (SLEDAI), with renal SLEDAI recorded when available. Renal variables included urinary albumin-to-creatinine ratio (uACR), 24-hour urinary protein excretion, estimated glomerular filtration rate (eGFR), and serum albumin. Characteristics are summarised in Tables 1 and 2, with variable-specific sample sizes reported. Laboratory results dated more than 30 days before or after serum collection were excluded; undated results were retained as recorded. Missing measurements were not imputed. The eGFR was calculated from serum creatinine using the 2021 Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation.

Serum collection and inflammatory proteomic profiling

Blood was collected in BD Vacutainer serum separator tubes (Becton, Dickinson and Company, Franklin Lakes, NJ, USA), allowed to clot at room temperature, and centrifuged at 1,500 × g for 10 min. Serum was aliquoted and stored at − 80°C for up to six months. All samples underwent one freeze–thaw cycle, and collection and processing procedures were consistent across cohorts. Unfractionated serum from all 73 discovery-cohort participants was analysed using the Olink Target 96 Inflammation panel (product no. 95302; Olink Proteomics, Uppsala, Sweden). The proximity extension assay (PEA)-based platform measures 92 inflammation-related proteins using 1 µL of serum per sample. Protein abundance was reported as normalised protein expression (NPX), a relative measure on a $\log_2$ scale. Manufacturer-specified incubation, extension, and detection controls were monitored. NPX values were available for all 92 proteins in every sample; no imputation or retrospective batch correction was required.

Untargeted serum metabolomic profiling

Untargeted serum metabolomics was performed by Metware Biotechnology Co., Ltd. (Wuhan, China) using liquid chromatography–tandem mass spectrometry (LC–MS/MS), following the company's established workflow. Frozen serum was thawed on ice and vortex-mixed for 10 s. Extraction solution (300 µL; acetonitrile:methanol, 1:4, v/v, with internal standards) was added to 50 µL of serum. After vortexing for 3 min and centrifugation at 12,000 rpm and 4°C for 10 min, 200 µL of supernatant was held at − 20°C for 30 min, then centrifuged at 12,000 rpm and 4°C for 3 min. The final 180 µL of supernatant

was collected for analysis. Chromatography used a Waters ACQUITY Premier HSS T3 column (1.8 μm, 2.1 × 100 mm), with 0.1% formic acid in water (A) and acetonitrile (B). The gradient was 5%–20% B over 1 min, 20%–99% B over 2 min, and 99% B for 1.5 min, followed by a return to 5% B in 0.1 min and re-equilibration for 1.4 min. The column temperature was 40°C, flow rate 0.4 mL/min, and injection volume 3 μL. Positive- and negative-mode electrospray ionisation runs were acquired on a Q Exactive mass spectrometer (Thermo Fisher Scientific) using full-scan mode (m/z 70–1,000; resolution 60,000). Ion spray voltages were 3.8 kV and 3.4 kV, respectively; sheath and auxiliary gas settings were 60 and 20. The ion transfer tube and vaporiser temperatures were 320°C and 350°C. Data-dependent MS/MS used collision energies of 30, 40, and 50 V, with dynamic exclusion. Pooled quality-control (QC) samples comprising aliquots of study-sample extracts were analysed alongside study samples. Raw files were converted to mzXML using ProteoWizard; peak detection, alignment, and retention-time correction used an XCMS-based workflow. Feature filtering, missing-value processing, normalisation, and quality assessment followed the service provider's workflow. QC samples were used to assess reproducibility, and features failing QC criteria were excluded. Metabolites were annotated against the company's in-house library and public databases using accurate-mass and tandem mass spectrometry data, supplemented by computational approaches. After peak detection, QC filtering, and annotation, 520 features were retained and $\log_2$-transformed. Mapping the annotated compounds to the Kyoto Encyclopedia of Genes and Genomes (KEGG) yielded 240 unique compounds for pathway analysis.

Integrated multi-omics analysis and candidate protein prioritisation

Principal component analysis (PCA) was performed on centred and scaled proteomic and metabolomic data to assess overall molecular variation and visualise group distributions. Weighted gene co-expression network analysis (WGCNA) was applied to an integrated matrix of 92 inflammatory proteins and 520 metabolomic features[29]. An unsigned network was constructed using a soft-thresholding power (β) of 4. Modules were identified by hierarchical clustering and dynamic tree cutting, with a minimum module size of six features and a deepSplit parameter of 2; no further module merging was performed. Module eigengenes were correlated with binary indicators of SLE and LN status using Pearson correlation. Candidate proteins were selected by intersecting proteins assigned to the module most strongly associated with both clinical indicators with proteins differentially abundant in both comparisons: all patients with SLE versus HCs and patients with LN versus those with SLE without renal involvement. Clinical laboratory variables were not included in network construction. The four proteins selected for ELISA were colony-stimulating factor 1 (CSF-1), programmed death-ligand 1 (PD-L1), CX3CL1, and sulfotransferase 1A1 (SULT1A1).

ELISA measurement of candidate proteins

The four candidates were quantified in the follow-up cohort using commercially available sandwich ELISA kits (ABclonal Biotechnology Co., Ltd., Wuhan, China), according to the manufacturers' instructions. Samples were analysed in duplicate, and concentrations were calculated from the

corresponding standard curves using four-parameter logistic calibration. Olink NPX values and ELISA concentrations were analysed independently on their respective scales; no direct numerical conversion or equivalence between platforms was assumed.

Public transcriptomic and single-cell expression analysis

Public datasets were analysedto examine the four candidate genes in renal and peripheral immune-cell populations (Fig. 6). In GSE180393, glomerular expression of*CSF1*, *CD274*(encoding PD-L1), *CX3CL1*, and*SULT1A1*was compared between nine healthy donors and 15 patients with LN. Expression was displayed in a row-scaled heatmap; group differences were tested using two-sided moderated t-tests inlimma, with Benjamini–Hochberg correction across the four genes. Kidney single-cell analysis used the Broad/CEL-seq2 dataset from the AMP SLE Phase I study (ImmPort SDY997). Original cell annotations were retained, t-distributed stochastic neighbour embedding (t-SNE) coordinates were recalculated, and gene expression was displayed as $\log_2$(1 + counts per 10,000 [CP10K]). The paediatric peripheral blood mononuclear cell (PBMC) dataset GSE135779 was analysed using uniform manifold approximation and projection (UMAP); adult samples were excluded. Donor-level comparisons treated each donor as one biological observation. In kidney, *CSF1*and*CD274*were assessed in leucocytes, *CX3CL1*in epithelial cells, and*SULT1A1*in myeloid cells. Expression was the mean cell-level $\log_2$(1 + CP10K), including cells without detectable expression; groups were compared using two-sided Wilcoxon rank-sum tests. Inpaediatric PBMC classical monocytes, donor-level pseudobulk TMM-normalised$\log_2$ counts per million (CPM) for*SULT1A1*wereanalysed using edgeR quasi-likelihood models adjusted for age, sex, and batch. These analyses were exploratory; single-cell P values were not adjusted for multiple testing and were interpreted as nominal. Kidney and PBMC datasets were analysed separately, and transcript expression was not taken to represent circulating protein abundance or establish the cellular origin of serum proteins.

Statistical analysis

Analyses were performed in R (version 4.5.0). Continuous variables are summarised as mean ± standard deviation (SD) or median and interquartile range (IQR), and categorical variables as counts and percentages. Analyses used available observations without imputation; variable-specific sample sizes are reported in Tables 1 and 2. No prospective sample-size calculation was performed. Protein abundance was compared using Welch's t-tests, with Benjamini–Hochberg (BH) correction across 92 proteins separately for each comparison (all SLE versus HCs, LN versus HCs, and LN versus SLE without renal involvement). Metabolite features were compared across the three groups using Kruskal–Wallis tests with BH correction. Pathway over-representation analysis used metabolites from the eight eligible non-grey WGCNA modules and 240 unique KEGG-mapped compounds as background; Holm correction was applied across pathways. Dunn's post hoc tests compared 18 candidate metabolites across three group contrasts, with Holm correction across 54 comparisons, although only six metabolites are shown in Fig. 4C. Spearman correlations between the six metabolites and four candidate proteins were

adjusted by BH across 24 comparisons. In the follow-up cohort, Wilcoxon rank-sum tests compared candidate-protein concentrations, with Holm correction across 12 comparisons (four proteins and three group contrasts). ROC analyses assessed discrimination between LN and SLE without renal involvement in 51 patients per cohort: 22 without renal involvement and 29 with LN in the discovery cohort, and 25 and 26, respectively, in the ELISA cohort. AUCs and 95% confidence intervals (CIs) were calculated using DeLong's method. Comparisons with C3 and C4 used complete cases with all four markers (51 discovery and 49 follow-up patients); higher CSF-1 and PD-L1 and lower C3 and C4 were prespecified as LN-oriented. Paired DeLong tests compared AUCs, with within-cohort Holm correction. Spearman correlations between four candidate proteins and 13 clinical variables were assessed in patients with LN and SLE without renal involvement, with BH correction across 52 comparisons per cohort. The four bulk transcriptomic gene comparisons were BH-adjusted. Kidney single-cell and paediatric PBMC analyses were exploratory and unadjusted for multiple testing. All tests were two-sided. Adjusted $P < 0.05$ indicated statistical significance for corrected analyses; nominal $P < 0.05$ in exploratory single-cell analyses was considered descriptive and interpreted cautiously.

RESULTS

Clinical characteristics of the discovery and ELISA follow-up cohorts

The discovery and ELISA follow-up cohorts included 73 and 79 participants, respectively; one participant contributed specimens to both cohorts at different sampling dates. Demographic, clinical, and laboratory characteristics are summarised in Tables 1 and 2. In both cohorts, patients with LN had greater disease activity and more pronounced renal abnormalities than patients with non-renal SLE. In the discovery cohort, median SLEDAI-2K was 20.0 (IQR, 18.0–22.0) in LN and 8.5 (IQR, 6.5–12.8) in non-renal SLE. Median uACR was 2,319.8 versus 9.4 mg/g, and median 24-hour urinary protein excretion was 3,756.5 versus 127.0 mg/24 h. Serum albumin, eGFR, and C3 concentrations were lower in patients with LN. In the follow-up cohort, median SLEDAI-2K was 12.0 (IQR, 9.5–15.0) in LN and 2.0 (IQR, 0.0–5.0) in non-renal SLE; median uACR was 1,549.7 mg/g (n = 25) and 14.8 mg/g (n = 15), respectively.

Table 1
Clinical and laboratory characteristics of the multi-omics discovery cohort

Characteristic	HC (n = 22)	SLE without LN (n = 22)	LN (n = 29)
Participants, n	22	22	29
Female sex, n/N (%)	20/22 (90.9%)	20/22 (90.9%)	27/29 (93.1%)
Age, years	28.0 ± 4.7	35.0 ± 16.1	30.0 ± 11.6
SLEDAI	NA	8.5 (6.5–12.8)	20.0 (18.0–22.0)
Urinary albumin-to-creatinine ratio, mg/g	NA	9.4 (7.2–13.4); n = 21	2319.8 (984.2–3686.6); n = 28
24-h urinary protein, mg/24 h	NA	127.0 (99.0–207.0); n = 11	3756.5 (2222.8–6411.5); n = 26
Renal pathological class documented, n/N (%)	NA	NA	29/29 (100.0%)
Class II, n/N (%)	NA	NA	2/29 (6.9%)
Class III, n/N (%)	NA	NA	4/29 (13.8%)
Class IV, n/N (%)	NA	NA	5/29 (17.2%)
Class V, n/N (%)	NA	NA	3/29 (10.3%)
Class III + IV, n/N (%)	NA	NA	1/29 (3.4%)
Class III + V, n/N (%)	NA	NA	6/29 (20.7%)
Class IV + V, n/N (%)	NA	NA	8/29 (27.6%)
C1q, mg/dL	NA	16.5 ± 3.3; n = 16	14.4 ± 3.2; n = 25
Anti-dsDNA positive or weakly positive, n/N (%)	NA	8/20 (40.0%)	13/28 (46.4%)
ANA positive, n/N (%)	NA	18/19 (94.7%)	28/29 (96.6%)
C3, g/L	NA	0.87 ± 0.30	0.68 ± 0.27
C4, g/L	NA	0.15 ± 0.08	0.13 ± 0.11
Erythrocyte sedimentation rate, mm/h	NA	18.0 (13.0–44.0); n = 21	34.0 (15.0–52.5); n = 27
Serum albumin, g/L	NA	38.8 ± 3.4	29.1 ± 5.7
eGFR, mL/min/1.73 m²	130.9 ± 14.9; n = 20	167.7 ± 65.3	109.9 ± 43.6; n = 27
White blood cell count, ×10⁹/L	6.0 ± 1.1	5.1 ± 2.3	5.8 ± 2.5

Characteristic	HC (n = 22)	SLE without LN (n = 22)	LN (n = 29)
Lymphocyte count, ×10⁹/L	2.2 ± 0.6	1.4 ± 0.8	1.2 ± 0.5
Haemoglobin, g/L	139.9 ± 11.0	109.4 ± 21.4	106.7 ± 18.5
Platelet count, ×10⁹/L	227.5 ± 42.2	184.2 ± 102.8	204.0 ± 60.3
Th/Ts ratio	NA	0.9 ± 0.4; n = 19	0.9 ± 0.3; n = 27
25-hydroxyvitamin D, ng/mL	NA	19.0 (12.8–23.9); n = 18	10.6 (8.4–15.9); n = 26

Table 2
Clinical and laboratory characteristics of the ELISA follow-up cohort

Characteristic	HC (n = 28)	SLE without LN (n = 25)	LN (n = 26)
Participants, n	28	25	26
Female sex, n/N (%)	24/28 (85.7%)	22/25 (88.0%)	22/26 (84.6%)
Age, years	37.3 ± 10.6	38.8 ± 12.4	33.5 ± 12.2
SLEDAI	NA	2.0 (0.0–5.0)	12.0 (9.2–15.5)
Renal SLEDAI	NA	0.0 (0.0–4.0)	8.0 (4.0–12.0)
Urinary albumin-to-creatinine ratio, mg/g	NA	13.9 (8.4–27.8); n = 16	1540.3 (552.4–2433.5); n = 24
24-h urinary protein, mg/24 h	NA	128.0 (87.5–253.2); n = 6	2513.5 (1297.8–4712.8); n = 24
Renal pathological class documented, n/N (%)	NA	NA	13/26 (50.0%)
Class III, n/N (%)	NA	NA	3/13 (23.1%)
Class IV, n/N (%)	NA	NA	2/13 (15.4%)
Class V, n/N (%)	NA	NA	2/13 (15.4%)
Class III + V, n/N (%)	NA	NA	1/13 (7.7%)
Class IV + V, n/N (%)	NA	NA	5/13 (38.5%)
C1q, mg/dL	NA	17.4 ± 3.0; n = 22	12.9 ± 4.5; n = 25
Anti-dsDNA positive or weakly positive, n/N (%)	NA	5/13 (38.5%)	6/20 (30.0%)
ANA positive, n/N (%)	NA	17/18 (94.4%)	20/23 (87.0%)
C3, g/L	NA	0.97 ± 0.24; n = 24	0.70 ± 0.29; n = 25
C4, g/L	NA	0.21 ± 0.11; n = 24	0.14 ± 0.11; n = 25
Erythrocyte sedimentation rate, mm/h	NA	16.0 (10.8–23.5); n = 20	48.0 (13.0–81.0); n = 25
Serum albumin, g/L	NA	40.0 ± 4.7	31.7 ± 7.1
eGFR, mL/min/1.73 m²	117.0 ± 11.2	121.2 ± 36.8	85.4 ± 42.4
White blood cell count, ×10⁹/L	5.9 ± 1.0	4.5 ± 2.4	6.0 ± 3.0
Lymphocyte count, ×10⁹/L	2.1 ± 0.5	0.9 ± 0.5	1.3 ± 0.8

Characteristic	HC (n = 28)	SLE without LN (n = 25)	LN (n = 26)
Haemoglobin, g/L	140.2 ± 7.8	118.5 ± 17.6	104.3 ± 26.5
Platelet count, ×10⁹/L	224.8 ± 40.0	194.1 ± 90.4	177.6 ± 88.6

Serum inflammatory proteomic profiles and integrated prioritisation of candidate proteins

Olink Target 96 profiling quantified 92 inflammation-related proteins in all 73 discovery-cohort participants. The first two principal components explained 20.94% and 11.10% of proteomic variance, respectively. HCs and patients with SLE were more clearly separated than patients with LN and non-renal SLE, whose profiles overlapped substantially (Fig. 2A). After Benjamini–Hochberg correction (FDR < 0.05), 49 proteins differed between all patients with SLE and HCs; fewer differences were detected between LN and non-renal SLE, suggesting that many circulating inflammatory changes reflect the broader SLE disease background (Fig. 2B). The heatmap showed protein abundance patterns across the three groups (Fig. 2C). Integrated WGCNA of 92 proteins and 520 metabolomic features identified 20 co-expression modules (Fig. 2D). The turquoise module was strongly positively correlated with SLE status (r = 0.96) and positively correlated with LN status (r = 0.62) (Fig. 2E). Intersecting its protein components with proteins differentially abundant in both the all-SLE versus HC and LN versus non-renal SLE comparisons yielded four candidates: colony-stimulating factor 1 (CSF-1), programmed death-ligand 1 (PD-L1), CX3CL1, and sulfotransferase 1A1 (SULT1A1) (Fig. 2F). These candidates were selected for ELISA evaluation.

Candidate-protein profiles and ELISA follow-up evaluation

In the discovery cohort, Olink measurements showed distinct group profiles for the four prioritised proteins (Fig. 3A). CSF-1, PD-L1, and SULT1A1 NPX values increased across HCs, non-renal SLE, and LN, whereas CX3CL1 was higher in LN than in either comparison group. For the primary comparison of LN versus non-renal SLE, AUCs were 0.857 (95% CI, 0.753–0.962) for CSF-1, 0.784 (0.651–0.917) for CX3CL1, 0.743 (0.597–0.889) for SULT1A1, and 0.738 (0.599–0.878) for PD-L1 (Fig. 3C, middle panel). In the ELISA follow-up cohort, CSF-1 and PD-L1 concentrations were significantly higher in LN than in both non-renal SLE and HCs (Fig. 3B). CX3CL1 was higher in LN than in HCs, but not significantly different between LN and non-renal SLE; SULT1A1 did not differ significantly among the groups. Thus, follow-up results supported the LN-associated differences for CSF-1 and PD-L1, but did not reproduce the discovery findings for CX3CL1 and SULT1A1 to the same extent. For LN versus non-renal SLE, AUCs were 0.890 (95% CI, 0.795–0.986) for CSF-1, 0.883 (0.791–0.975) for PD-L1, 0.636 (0.483–0.790) for CX3CL1, and 0.547 (0.387–0.708) for SULT1A1 (Fig. 3D, middle panel). In secondary analyses distinguishing non-renal SLE from HCs, AUCs for CSF-1, PD-L1, CX3CL1, and SULT1A1 were 0.829, 0.779, 0.467, and 0.899, respectively, in the discovery cohort, and 0.984, 0.893, 0.701, and 0.681 in the follow-up cohort. The corresponding 95% CIs are shown in the left panels of Fig. 3C and D.

Comparison of candidate proteins with conventional complement markers

Discrimination by candidate proteins was compared with C3 and C4 among participants with complete measurements for all four markers (Fig. 3C and D, right panels). The analysis included 51 patients in the discovery cohort (22 with non-renal SLE and 29 with LN) and 49 in the ELISA follow-up cohort (24 and 25, respectively). In the discovery cohort, AUCs were 0.856 (95% CI, 0.750–0.961) for CSF-1, 0.737 (0.597–0.877) for PD-L1, 0.697 (0.547–0.848) for C3, and 0.603 (0.443–0.763) for C4. In the follow-up cohort, they were 0.878 (95% CI, 0.768–0.988), 0.866 (0.757–0.975), 0.742 (0.594–0.890), and 0.652 (0.485–0.819), respectively. Although CSF-1 and PD-L1 had numerically higher AUCs than C3 and C4 in both cohorts, paired comparisons provided limited evidence of superiority. After Holm correction, CSF-1 had a significantly higher AUC than C4 in the discovery cohort; no candidate significantly outperformed either complement marker in the follow-up cohort. Thus, these results do not establish incremental diagnostic value beyond conventional complement measurements.

Serum metabolomic profiles and protein–metabolite associations

Untargeted profiling identified 520 serum metabolomic features in the discovery cohort. The first two principal components explained 15.23% and 5.45% of the variance, respectively. As in the proteomic analysis, HCs were more clearly separated from patients with SLE than LN was from non-renal SLE, whose profiles overlapped considerably (Fig. 4A). KEGG pathway analysis of metabolites assigned to the selected WGCNA modules identified four pathways that remained significant after Holm correction: cholesterol metabolism, fatty-acid biosynthesis, primary bile-acid biosynthesis, and unsaturated fatty-acid biosynthesis (Fig. 4B). Pairwise comparisons of 18 prioritised metabolites across the three groups were Holm-adjusted across 54 comparisons; six metabolites are shown in Fig. 4C. Caprylic acid was lower in both SLE groups than in HCs, and docosahexaenoic acid (DHA) was lower in LN than in HCs. Glycine was lower in LN than in non-renal SLE, whereas taurine was higher in LN than in both comparison groups. Erucic acid and free fatty acid 20:0 showed no significant pairwise differences. Spearman analysis of the four candidate proteins and six selected metabolites showed positive correlations between taurine and all four proteins. Caprylic acid and DHA were inversely correlated with CSF-1, PD-L1, and SULT1A1, while glycine was inversely correlated with CSF-1 (Fig. 4D). These associations indicate coordinated variation but do not establish direct biochemical or regulatory relationships.

Associations of candidate proteins with disease activity and renal clinical measures

Associations with clinical characteristics were assessed in patients with SLE, including those with LN and non-renal SLE. In the discovery cohort, Olink-derived CSF-1 and PD-L1 NPX values were positively correlated with SLEDAI-2K scores (CSF-1: Spearman's $\rho = 0.45$, $P = 0.001$; PD-L1: $\rho = 0.46$, $P < 0.001$). CSF-1 was also positively correlated with 24-hour urinary protein excretion ($\rho = 0.77$, $P < 0.001$; $n = 37$)

and inversely correlated with eGFR (ρ=−0.40, *P* = 0.005; n = 49) (Fig. 5A). In the ELISA follow-up cohort, CSF-1 and PD-L1 concentrations were positively associated with SLEDAI-2K, uACR, 24-hour urinary protein excretion, and erythrocyte sedimentation rate, and inversely associated with eGFR, serum albumin, and C3 (Fig. 5B). Similar patterns were observed in the discovery cohort, where both proteins were positively associated with SLEDAI-2K, uACR, and proteinuria, and inversely associated with serum albumin and eGFR (Fig. 5C). Of the four candidates, CSF-1 showed the most consistent association with 24-hour urinary protein excretion across cohorts. Its association, together with those of CSF-1 and PD-L1 with overall disease activity, links these proteins to clinical measures of renal involvement and systemic SLE activity.

Candidate-gene expression in public transcriptomic and single-cell datasets

Public datasets were analysed to explore the cellular context of the four candidate genes (Fig. 6). In the glomerular dataset GSE180393, *CSF1* and *CX3CL1* expression was higher in samples from 15 patients with LN than in nine healthy controls; group expression patterns are shown in the heatmap (Fig. 6A, B). Single-cell analyses used the adult kidney dataset from AMP SLE Phase I/ImmPort SDY997 and the paediatric PBMC dataset GSE135779. Candidate-gene expression across annotated renal cell populations and peripheral immune-cell clusters is shown in the t-SNE and UMAP maps, respectively (Fig. 6C). Donor-level analyses further examined selected renal leucocyte, epithelial and myeloid populations, and PBMC classical monocytes (Fig. 6D). Expression varied across genes and cell types. The kidney and PBMC datasets represented distinct populations and clinical contexts and were not paired with participants in either serum cohort. These exploratory analyses provide cellular context for the circulating protein findings but do not establish the renal origin of serum CSF-1, PD-L1, *CX3CL1*, or*SULT1A1*.

Discussion

In this study, integrated serum proteomic and metabolomic analyses prioritised four proteins associated with LN: CSF-1, PD-L1, CX3CL1 and SULT1A1. CSF-1 and PD-L1 showed the most consistent differences between LN and non-renal SLE across the Olink discovery and ELISA follow-up cohorts, and CSF-1 was consistently associated with 24-hour urinary protein excretion. These findings identify CSF-1 and PD-L1 as candidates for further evaluation, but do not establish their renal specificity or clinical utility. This distinction is central to interpreting the results: in both omics analyses, separation was more pronounced between SLE and HCs than between LN and non-renal SLE, a pattern consistent with the molecular heterogeneity of SLE and variation in systemic immune activation[30–35]. The circulating signals may therefore reflect a combination of systemic disease activity and renal involvement. Moreover, patients with LN had higher SLEDAI-2K scores, which capture activity across multiple organ systems; the observed protein differences cannot consequently be ascribed to renal involvement alone. Differences in cohort composition, disease activity and assay characteristics may also have contributed to between-

cohort variation, particularly in PD-L1 measurements. The metabolomic and public transcriptomic analyses add biological context to these findings, but do not resolve the mechanisms underlying the protein associations or establish the renal cellular origin of circulating proteins.

CSF-1 was the strongest candidate for renal-focused investigation. Its association with LN was observed on both analytical platforms, and its correlation with 24-hour urinary protein excretion was reproduced in the discovery and follow-up cohorts (Spearman's ρ = 0.768 and 0.687, respectively). As a regulator of monocyte and macrophage survival, differentiation, and function, CSF-1 has a plausible link to renal inflammation. Previous studies have implicated macrophage pathways in lupus renal injury and suggested distinct roles for resident and recruited macrophages[36–38]; macrophage-associated molecules such as CD163 have also been investigated as indicators of LN activity[19]. Our findings extend this work by showing that circulating CSF-1 is associated with clinical measures of renal involvement across two assay platforms. However, systemic myeloid activation, renal inflammation, impaired kidney function, and altered protein handling could each contribute to circulating CSF-1. Moreover, proteinuria was part of the clinical LN definition for some patients without definitive biopsy evidence, so its correlation with CSF-1 is not independent of the classification framework. The cross-sectional design also cannot establish whether elevated CSF-1 precedes renal injury or reflects active or established proteinuric disease. Its relationship with intrarenal inflammation therefore requires evaluation in biopsy-anchored cohorts with tissue-level and longitudinal data.

PD-L1 showed a complementary pattern: its abundance increased across HCs, non-renal SLE, and LN in both cohorts, and it was positively associated with SLEDAI-2K (Spearman's ρ = 0.461 and 0.616, respectively). This may indicate that circulating PD-L1 reflects broader SLE immune activity, with additional variation in LN. As an immune checkpoint molecule involved in T-cell regulation and peripheral tolerance, PD-L1 is relevant to autoimmunity, but its effects depend on cellular context, molecular form, and local immune environment. Previous work reported increased soluble PD-1 and PD-L2, but not soluble PD-L1, in association with SLE activity[39], underscoring that soluble PD-1 pathway analytes are not interchangeable. Differences in populations, assays, and detected molecular forms may contribute to variation between studies, although these factors were not examined here. The reproducible PD-L1 elevation supports its relevance as a circulating inflammation-associated candidate, but its association with overall disease activity leaves its renal specificity uncertain. In complete-case comparisons with C3 and C4, CSF-1 and PD-L1 had numerically higher AUCs in both cohorts, but formal comparisons provided limited evidence of superiority: only CSF-1 exceeded C4 in the discovery cohort after Holm correction, and no candidate outperformed either complement marker in the follow-up cohort. These results do not establish incremental diagnostic value. Larger prospective studies should assess whether either protein improves classification beyond complement, anti-double-stranded DNA antibodies, proteinuria, and kidney function, using multivariable models and clinically relevant performance measures.

Metabolomic analysis identified four pathways that remained significant after multiple-testing correction: cholesterol metabolism, fatty-acid biosynthesis, primary bile-acid biosynthesis, and unsaturated fatty-acid biosynthesis. Taurine was higher in LN than in both comparison groups; caprylic

acid was lower in both SLE groups than in HCs; DHA was lower in LN than in HCs; and glycine was lower in LN than in non-renal SLE. Taurine correlated positively with all four candidate proteins, while caprylic acid and DHA were inversely associated with several candidates. These findings are consistent with coordinated variation in inflammatory and metabolic features, but do not establish direct biochemical or regulatory relationships. Immune-cell activity and immunometabolic reprogramming are closely linked in SLE[40, 41], and lipid-related pathways may reflect changes in membrane composition, inflammatory mediator production, or cellular adaptation. Taurine's higher abundance and positive correlations could reflect altered immune and metabolic states, but cannot distinguish a regulatory role from a compensatory response or changes in systemic metabolism and renal handling. Dietary intake, medication, metabolic status, and kidney function may also influence circulating metabolites. Although studies of mitochondrial damage and neutrophil-derived oxidised mitochondrial DNA provide broader context for interactions between metabolism and interferon activation in lupus [42, 43], these mechanisms were not assessed here. The metabolomic findings are therefore hypothesis-generating and require targeted biochemical and functional validation.

Public bulk and single-cell datasets offered exploratory context for candidate-gene expression in renal and peripheral immune-cell populations. The analyses showed that expression varied by gene, dataset, and cell type, but do not establish that kidney cells are the source of circulating proteins. Transcript abundance does not necessarily correspond to protein abundance, secretion, or activity; moreover, the public datasets were not paired with serum samples from our cohorts and differed in population, tissue, and platform. In particular, the paediatric PBMC dataset does not directly represent peripheral immune-cell expression in the adult serum cohorts. Strengths of the study include its primary comparison of LN with non-renal SLE, integration of proteomics and metabolomics, orthogonal ELISA assessment, and evaluation against clinical measures and complement markers. Several limitations qualify these findings. The study was single-centre and cross-sectional, with modest sample sizes; one participant contributed to both cohorts at different sampling dates, so the follow-up cohort was not fully independent. Biopsy information was incomplete in the follow-up cohort, and some LN classifications relied on the prespecified proteinuria criterion; phenotype misclassification and differences in histological class or activity cannot be excluded. Clinical measurements were incomplete, and treatment exposure was not systematically characterized. In addition, data-driven candidate selection and ROC estimation in the discovery cohort may have led to optimistic performance estimates. ELISA provided an orthogonal assessment but not fully independent external validation, and Olink NPX and ELISA concentrations are not directly comparable. Finally, metabolite annotations were not followed by targeted quantification, correlations may be confounded, and heterogeneous public datasets cannot establish cellular origin.

The next step is a prospective, multicentre, biopsy-anchored study of patients with suspected LN, active non-renal SLE, and appropriate disease controls. It should evaluate CSF-1 and PD-L1 alongside established markers, with systematic data on histological class, activity and chronicity, renal function, proteinuria, treatment, and systemic disease activity. Matched serum, plasma, urine, and biopsy specimens could clarify whether circulating signals correspond to renal-proximal changes; tissue protein

localisation and spatial or single-cell analyses could further investigate cellular sources. Longitudinal sampling is also needed to determine whether changes in CSF-1 or PD-L1 precede, accompany, or follow renal activity. Studies linking post-flare urinary biomarkers with later renal outcomes illustrate the value of testing biomarkers against defined clinical endpoints[44].

Conclusions

In conclusion, CSF-1 and PD-L1 showed the most consistent LN-associated findings across discovery and ELISA analyses, with CSF-1 particularly associated with proteinuria. PD-L1 may also reflect systemic SLE activity. The accompanying metabolomic and transcriptomic findings add context but do not establish mechanism or renal origin. These proteins, especially CSF-1, merit further evaluation, but their independent diagnostic and incremental clinical value remain unestablished.

Abbreviations

AUC
area under the curve
BH
Benjamini–Hochberg
CI
confidence interval
CKD-EPI
Chronic Kidney Disease Epidemiology Collaboration
CPM
counts per million
CP10K
counts per 10,000
CSF-1
colony-stimulating factor 1
CV
coefficient of variation
DHA
docosahexaenoic acid
eGFR
estimated glomerular filtration rate
ELISA
enzyme-linked immunosorbent assay
FDR
false discovery rate
HC
healthy control

IQR
interquartile range
KEGG
Kyoto Encyclopedia of Genes and Genomes
LC–MS/MS
liquid chromatography–tandem mass spectrometry
LN
lupus nephritis
MS/MS
tandem mass spectrometry
NPX
normalised protein expression
ORA
over-representation analysis
PBMC
peripheral blood mononuclear cell
PCA
principal component analysis
PD-L1
programmed death-ligand 1
PEA
proximity extension assay
QC
quality control
ROC
receiver operating characteristic
SD
standard deviation
SLE
systemic lupus erythematosus
SLEDAI-2K
SLE Disease Activity Index 2000
TMM
trimmed mean of M-values
t-SNE
t-distributed stochastic neighbour embedding
uACR
urinary albumin-to-creatinine ratio
UMAP
uniform manifold approximation and projection

WGCNA
weighted gene co-expression network analysis.

Declarations

Ethics approval and consent to participate

The study protocol was approved by the Institutional Ethics Committee of Nanjing Drum Tower Hospital (approval no. 2023 – 568). Written informed consent was obtained from all participants before enrolment. The study was conducted in accordance with the principles of the Declaration of Helsinki.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the Clinical Specialized In-Hospital General Projects of Nanjing Drum Tower Hospital, Affiliated Hospital of Medical School, Nanjing University (grant nos. 2024-LCYJ-MS-11, 2025-LCYJ-PY-04 and 2026-JCYJ-QP-060), and the Nanjing University Institute for Hospital Reform and Development Research Project on Hospital Management (grant no. NDYGN2025008). The funding sources had no role in the design of the study; collection, analysis, or interpretation of the data; preparation of the manuscript; or decision to submit the manuscript for publication.

Author Contribution

ZW conceptualized the study, developed the methodology, performed the formal analysis, and wrote the original draft. TL contributed to the Olink proteomic and metabolomic measurements and formal analysis. HX contributed to the methodology, formal analysis, and manuscript preparation. WL performed software-based analyses, formal analysis, visualization, and validation. SX curated the clinical data and performed the ELISA validation. FC co-supervised the study and critically revised the manuscript. YT and ZL contributed to resources, supervision, project administration, and manuscript review and editing. PY conceptualized the study, acquired funding, supervised and administered the project, and contributed to manuscript review and editing. All authors read and approved the final manuscript and accept responsibility for its content.

Acknowledgements

The authors thank all patients and healthy volunteers who participated in this study. We also acknowledge the clinical and laboratory staff of Nanjing Drum Tower Hospital for their assistance with participant recruitment, sample collection, laboratory testing, and clinical data management.

Availability of data and materials

The de-identified clinical data, Olink NPX measurements, ELISA results, and metabolomic datasets generated and/or analysed during the current study are not publicly available because they contain human participant information and are subject to institutional and ethical restrictions. These data are available from the corresponding author upon reasonable request and subject to institutional approval. The publicly available transcriptomic datasets analysed in this study include GSE180393 and GSE135779, which can be accessed through the Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/). The kidney single-cell datasets from the AMP SLE Phase I study (ImmPort SDY997), including the ampsle1 and ampsle1-metro datasets, are accessible through the Cell Browser public resource (https://immunogenomics.io/cellbrowser/).

References

1. Crow MK. Pathogenesis of systemic lupus erythematosus: risks, mechanisms and therapeutic targets. Ann Rheum Dis. 2023;82:999–1014. https://doi.org/10.1136/ard-2022-223741.
2. Arnaud L, Chasset F, Martin T. Immunopathogenesis of systemic lupus erythematosus: an update. Autoimmun Rev. 2024;23:103648. https://doi.org/10.1016/j.autrev.2024.103648.
3. Siegel CH, Sammaritano LR. Systemic lupus erythematosus: a review. JAMA. 2024;331:1480–91. https://doi.org/10.1001/jama.2024.2315.
4. Anders HJ, Saxena R, Zhao MH, Parodis I, Salmon JE, Mohan C. Lupus nephritis. Nat Rev Dis Primers. 2020;6:7. https://doi.org/10.1038/s41572-019-0141-9.
5. Tsokos GC. Autoimmunity and organ damage in systemic lupus erythematosus. Nat Immunol. 2020;21:605–14. https://doi.org/10.1038/s41590-020-0677-6.
6. Lichtnekert J, Anders HJ. Lupus nephritis-related chronic kidney disease. Nat Rev Rheumatol. 2024;20:699–711. https://doi.org/10.1038/s41584-024-01158-w.
7. Kidney Disease: Improving Global Outcomes (KDIGO) Lupus Nephritis Work Group. KDIGO 2024 clinical practice guideline for the management of lupus nephritis. Kidney Int. 2024;105(1S):S1–69. https://doi.org/10.1016/j.kint.2023.09.002.
8. Sammaritano LR, Askanase A, Bermas BL, Dall'Era M, Duarte-García A, Hiraki LT, et al. 2024 American College of Rheumatology guideline for the screening, treatment, and management of lupus nephritis. Arthritis Care Res (Hoboken). 2025;77(9):1045–65. https://doi.org/10.1002/acr.25528.

9. Petrackova A, Smrzova A, Gajdos P, Schubertova M, Schneiderova P, Kromer P, et al. Serum protein pattern associated with organ damage and lupus nephritis in systemic lupus erythematosus revealed by PEA immunoassay. Clin Proteom. 2017;14:32. https://doi.org/10.1186/s12014-017-9167-8.
10. Fava A, Buyon J, Mohan C, Zhang T, Belmont HM, Izmirly P, et al. Integrated urine proteomics and renal single-cell genomics identify an IFN-γ response gradient in lupus nephritis. JCI Insight. 2020;5:e138345. https://doi.org/10.1172/jci.insight.138345.
11. Fava A, Rao DA, Mohan C, Zhang T, Rosenberg A, Fenaroli P, et al. Urine proteomics and renal single-cell transcriptomics implicate interleukin-16 in lupus nephritis. Arthritis Rheumatol. 2022;74:829–39. https://doi.org/10.1002/art.42023.
12. Li Y, Tang C, Vanarsa K, Thai N, Castillo J, Lea GAB, et al. Proximity extension assay proteomics and renal single-cell transcriptomics uncover novel urinary biomarkers for active lupus nephritis. J Autoimmun. 2024;143:103165. https://doi.org/10.1016/j.jaut.2023.103165.
13. Fava A, Buyon J, Magder L, Hodgin J, Rosenberg A, Demeke DS, et al. Urine proteomic signatures of histological class, activity, chronicity, and treatment response in lupus nephritis. JCI Insight. 2024;9(2):e172569. https://doi.org/10.1172/jci.insight.172569.
14. Vodehnal S, Mohan C. Urinary biomarkers for active lupus nephritis that have survived independent validation across cohorts. Kidney Int. 2024;106(6):1135–45. https://doi.org/10.1016/j.kint.2024.09.007.
15. Hiramoto K, Saito S, Hanaoka H, Kikuchi J, Fukui H, Hashiguchi A, et al. Urinary biomarkers associated with pathogenic pathways reflecting histologic findings in lupus nephritis. Arthritis Rheumatol. 2025;77(3):298–310. https://doi.org/10.1002/art.43017.
16. Lee YJ, Lee EJ, Kim K, Kim M, Yu JY, Kim M, et al. Urinary acetylated protein as a biomarker of lupus nephritis: a prospective cohort study. Arthritis Res Ther. 2025;27(1):220. https://doi.org/10.1186/s13075-025-03684-8.
17. Xiao D, Kong T, Wang T, Huang H, Qin W, Zhao B, et al. Serum protein profiles in lupus nephritis associated with initial-onset systemic lupus erythematosus: Characterization through PEA immunoassay and preliminary development of predictive model. Cytokine. 2026;197:157065. https://doi.org/10.1016/j.cyto.2025.157065.
18. Li Z, Sun Y, Wang Y, Liu F, Pan S, Li S, et al. Proteomics uncovers ICAM2 (CD102) as a novel serum biomarker of proliferative lupus nephritis. Lupus Sci Med. 2025;12:e001446. https://doi.org/10.1136/lupus-2024-001446.
19. Gürel E, Çinar S, Hürdogan Ö, Özlük Y, Kemeç G, Uludağ Ö, et al. CD163: a potential biomarker for lupus nephritis activity. Lupus Sci Med. 2026;13:e002166. https://doi.org/10.1136/lupus-2026-002166.
20. Wang L, Wang W, Hao Y, Chen L, Qi C, Liang D, et al. DIA proteomics of FFPE renal biopsies reveals two molecular subtypes of lupus nephritis and identifies APOL1 as candidate biomarker for stratification. Lupus Sci Med. 2026;13:e002123. https://doi.org/10.1136/lupus-2026-002123.

21. Morel L. Immunometabolism in systemic lupus erythematosus. Nat Rev Rheumatol. 2017;13:280–90. https://doi.org/10.1038/nrrheum.2017.43.
22. Patiño-Martínez E, Kaplan MJ. Immunometabolism in systemic lupus erythematosus. Nat Rev Rheumatol. 2025;21:377–95. https://doi.org/10.1038/s41584-025-01267-0.
23. Guleria A, Phatak S, Dubey D, Kumar S, Zanwar A, Chaurasia S, et al. NMR-based serum metabolomics reveals reprogramming of lipid dysregulation following cyclophosphamide-based induction therapy in lupus nephritis. J Proteome Res. 2018;17:2440–8. https://doi.org/10.1021/acs.jproteome.8b00192.
24. Gavin PG, Allman EL, Jayne D, Mysler E, Amoura Z, Di Poto C, et al. Metabolomic profiling of patients with lupus nephritis reveals unique metabolites that are modulated through type I interferon inhibition by anifrolumab treatment in a phase 2 trial. RMD Open. 2025;11:e005932. https://doi.org/10.1136/rmdopen-2025-005932.
25. Zhang Y, Gan L, Tang J, Liu D, Chen G, Xu B. Metabolic profiling reveals new serum signatures to discriminate lupus nephritis from systemic lupus erythematosus. Front Immunol. 2022;13:967371. https://doi.org/10.3389/fimmu.2022.967371.
26. Zhan S, Wang Z, Xu Y, Zhou S, Ge M, Song Y, et al. Inflammation-related proteomics of extracellular vesicles as novel biomarkers for systemic lupus erythematosus revealed by proximity extension assay. Arthritis Res Ther. 2025;27:211. https://doi.org/10.1186/s13075-025-03681-x.
27. Arazi A, Rao DA, Berthier CC, Davidson A, Liu Y, Hoover PJ, et al. The immune cell landscape in kidneys of patients with lupus nephritis. Nat Immunol. 2019;20:902–14. https://doi.org/10.1038/s41590-019-0398-x.
28. Aringer M, Costenbader K, Daikh D, Brinks R, Mosca M, Ramsey-Goldman R, et al. 2019 European League Against Rheumatism/American College of Rheumatology classification criteria for systemic lupus erythematosus. Arthritis Rheumatol. 2019;71:1400–12. https://doi.org/10.1002/art.40930.
29. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559. https://doi.org/10.1186/1471-2105-9-559.
30. Psarras A, Wittmann M, Vital EM. Emerging concepts of type I interferons in SLE pathogenesis and therapy. Nat Rev Rheumatol. 2022;18:575–90. https://doi.org/10.1038/s41584-022-00826-z.
31. Barrat FJ, Crow MK, Ivashkiv LB. Interferon target-gene expression and epigenomic signatures in health and disease. Nat Immunol. 2019;20:1574–83. https://doi.org/10.1038/s41590-019-0466-2.
32. Banchereau R, Hong S, Cantarel B, Baldwin N, Baisch J, Edens M, et al. Personalized immunomonitoring uncovers molecular networks that stratify lupus patients. Cell. 2016;165:551–65. https://doi.org/10.1016/j.cell.2016.03.008.
33. Haynes WA, Haddon DJ, Diep VK, Khatri A, Bongen E, Yiu G, et al. Integrated, multicohort analysis reveals unified signature of systemic lupus erythematosus. JCI Insight. 2020;5:e122312. https://doi.org/10.1172/jci.insight.122312.
34. Nakano M, Ota M, Takeshima Y, Iwasaki Y, Hatano H, Nagafuchi Y, et al. Distinct transcriptome architectures underlying lupus establishment and exacerbation. Cell. 2022;185:3375–e8921.

https://doi.org/10.1016/j.cell.2022.07.021.

35. Fasano S, Milone A, Nicoletti GF, Isenberg DA, Ciccia F. Precision medicine in systemic lupus erythematosus. Nat Rev Rheumatol. 2023;19:331–42. https://doi.org/10.1038/s41584-023-00948-y.
36. Mohan C, Zhang T, Putterman C. Pathogenic cellular and molecular mediators in lupus nephritis. Nat Rev Nephrol. 2023;19:491–508. https://doi.org/10.1038/s41581-023-00722-z.
37. Richoz N, Tuong ZK, Loudon KW, Patiño-Martínez E, Ferdinand JR, Portet A, et al. Distinct pathogenic roles for resident and monocyte-derived macrophages in lupus nephritis. JCI Insight. 2022;7:e159751. https://doi.org/10.1172/jci.insight.159751.
38. Kwant LE, Vegting Y, Tsang-A-Sjoe MWP, Kwakernaak AJ, Vogt L, Voskuyl AE, et al. Macrophages in lupus nephritis: exploring a potential new therapeutic avenue. Autoimmun Rev. 2022;21:103211. https://doi.org/10.1016/j.autrev.2022.103211.
39. Hirahara S, Katsumata Y, Kawasumi H, Kawaguchi Y, Harigai M. Serum levels of soluble programmed cell death protein 1 and soluble programmed cell death protein ligand 2 are increased in systemic lupus erythematosus and associated with the disease activity. Lupus. 2020;29:686–96. https://doi.org/10.1177/0961203320916517.
40. Romo-Tena J, Kaplan MJ. Immunometabolism in the pathogenesis of systemic lupus erythematosus: an update. Curr Opin Rheumatol. 2020;32:562–71. https://doi.org/10.1097/BOR.0000000000000738.
41. Sharabi A, Tsokos GC. T cell metabolism: new insights in systemic lupus erythematosus pathogenesis and therapy. Nat Rev Rheumatol. 2020;16:100–12. https://doi.org/10.1038/s41584-019-0356-x.
42. Lood C, Blanco LP, Purmalek MM, Carmona-Rivera C, De Ravin SS, Smith CK, et al. Neutrophil extracellular traps enriched in oxidized mitochondrial DNA are interferogenic and contribute to lupus-like disease. Nat Med. 2016;22:146–53. https://doi.org/10.1038/nm.4027.
43. Caielli S, Athale S, Domic B, Murat E, Chandra M, Banchereau R, et al. Oxidized mitochondrial nucleoids released by neutrophils drive type I interferon production in human lupus. J Exp Med. 2016;213:697–713. https://doi.org/10.1084/jem.20151876.
44. Baker R, Whittall Garcia LP, Li Q, Kim M, Bonilla D, Gladman DD, et al. Measurements of urinary biomarkers at 2 years following a lupus nephritis flare are associated with subsequent renal outcomes. Lupus Sci Med. 2026;13:e001724. https://doi.org/10.1136/lupus-2025-001724.

Figures

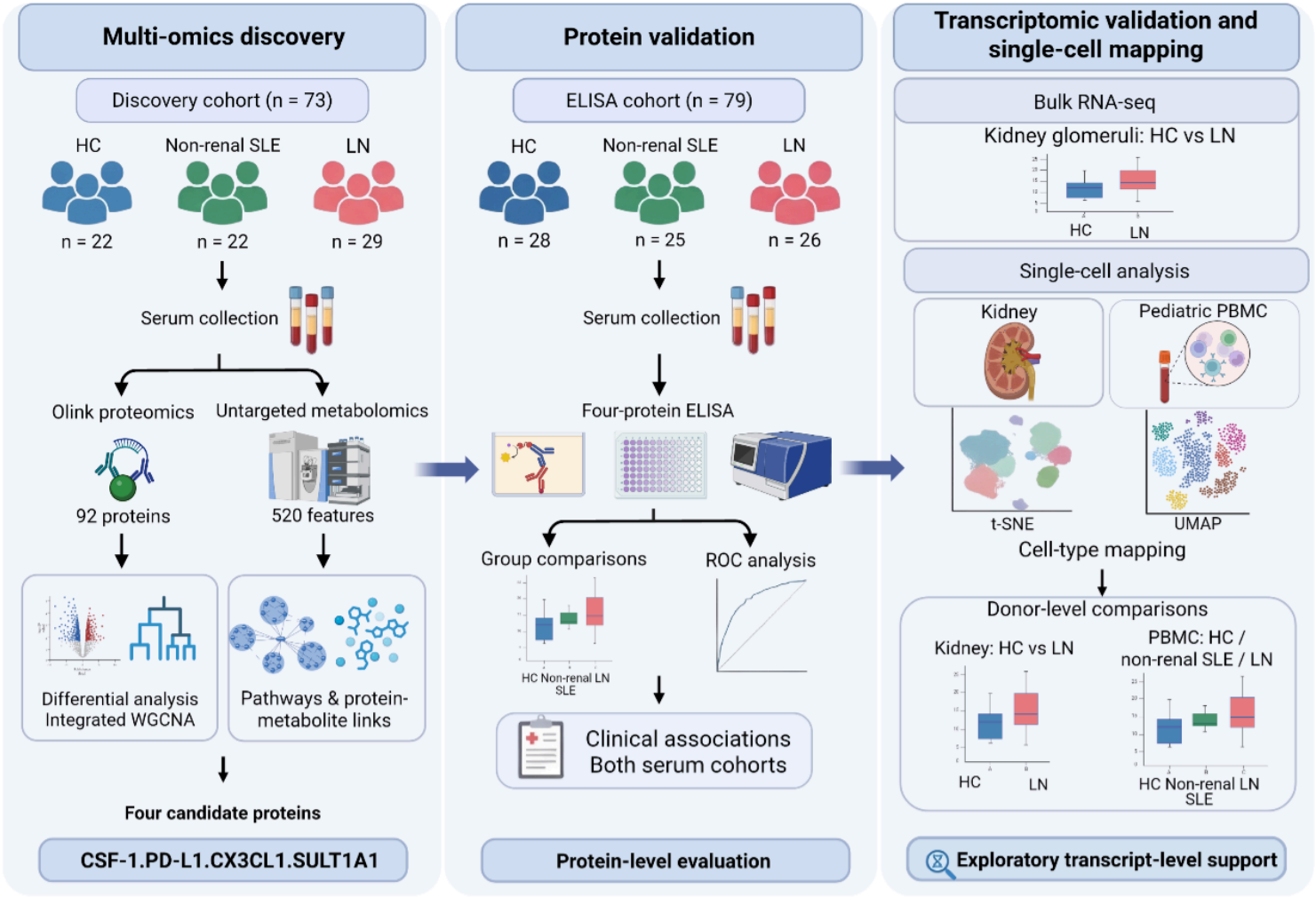


Figure 1

Study design and analytical workflow. The discovery cohort comprised 73 participants: 22 healthy controls (HCs), 22 patients with non-renal systemic lupus erythematosus (SLE) and 29 patients with lupus nephritis (LN). Four candidate proteins identified through integrated proteomic and metabolomic analyses were evaluated by enzyme-linked immunosorbent assay (ELISA) in a follow-up cohort of 79 participants (28 HCs, 25 with non-renal SLE and 26 with LN). One participant contributed specimens to both cohorts at different sampling dates. Public-data analyses included glomerular bulk RNA sequencing and exploratory single-cell analyses of adult kidney and paediatric peripheral blood mononuclear cell (PBMC) datasets, with cell-type mapping and donor-level comparisons. Workflow plots are schematic. Created with BioRender.com.

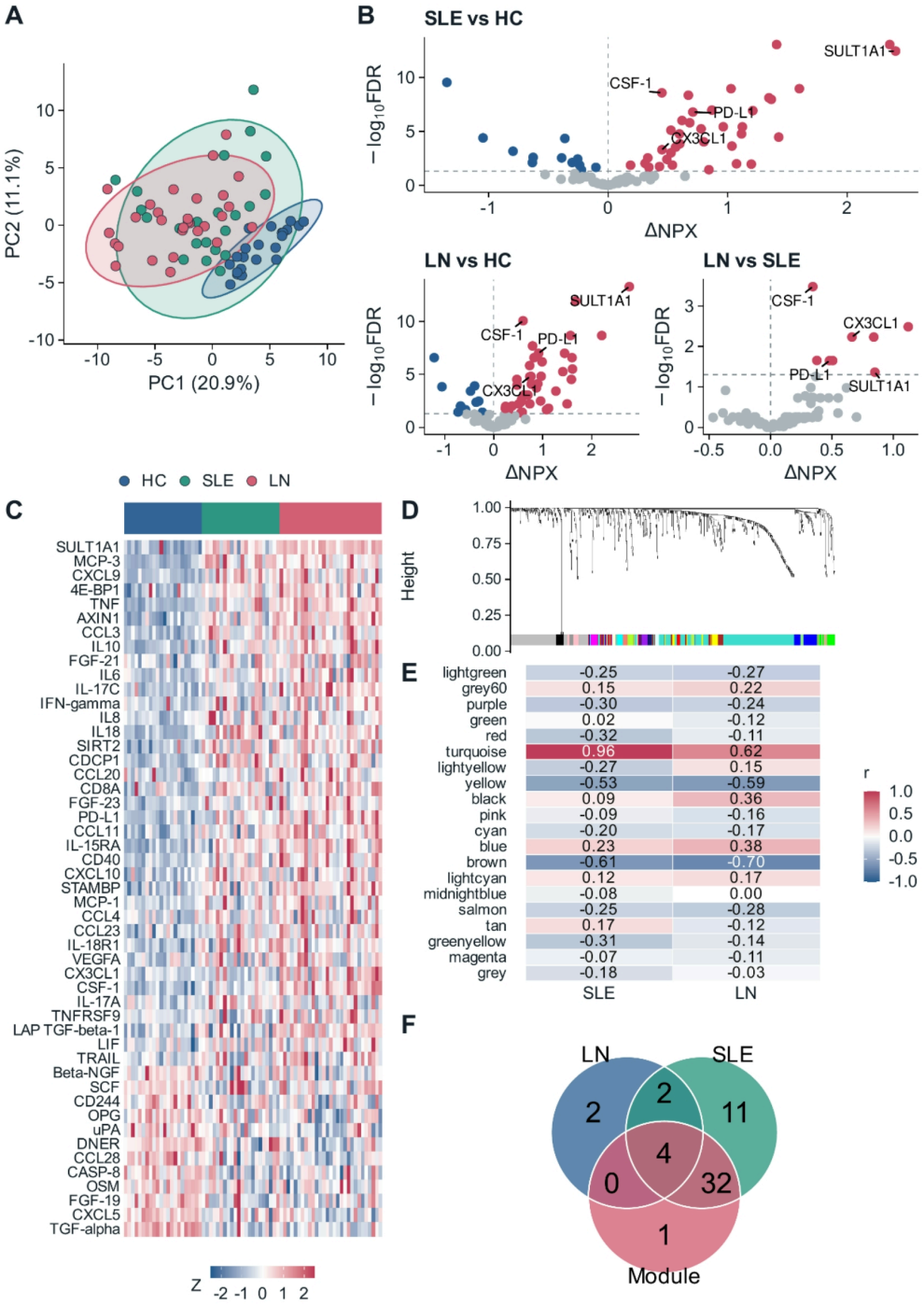


Figure 2

Serum inflammatory proteomic profiling and prioritisation of candidate proteins. (A) Principal component analysis (PCA) of 92 inflammatory proteins measured using the Olink Target 96 Inflammation panel in the discovery cohort (n=73). Ellipses show group distributions. (B) Volcano plots of differential protein abundance for all patients with SLE (non-renal SLE and LN combined) versus HCs, LN versus HCs, and LN versus non-renal SLE. The x-axis shows the difference in normalized protein

expression (ΔNPX), and the y-axis shows $-\log_{10}$ of the Benjamini–Hochberg (BH)-adjusted *P* value. Dashed lines indicate zero difference and the false discovery rate (FDR) threshold of 0.05. Colours denote increased, decreased and non-significant abundance. (C) Heatmap of the 49 proteins that differed between all patients with SLE and HCs after BH correction (FDR<0.05). Values are row-standardized z scores; the original sample order is retained. (D) Hierarchical clustering dendrogram and module assignments from weighted gene co-expression network analysis (WGCNA) of the integrated 612-feature matrix (92 proteins and 520 metabolomic features), displayed using the original hanging convention. (E) Pearson correlations between module eigengenes and binary indicators of SLE and LN status. The turquoise module was positively correlated with SLE status (r=0.96) and LN status (r=0.62). (F) Venn diagram showing the overlap of LN-upregulated proteins, proteins differing between all patients with SLE and HCs, and proteins assigned to the turquoise module. Four proteins were shared across all three sets: CSF-1, PD-L1, CX3CL1 and SULT1A1.

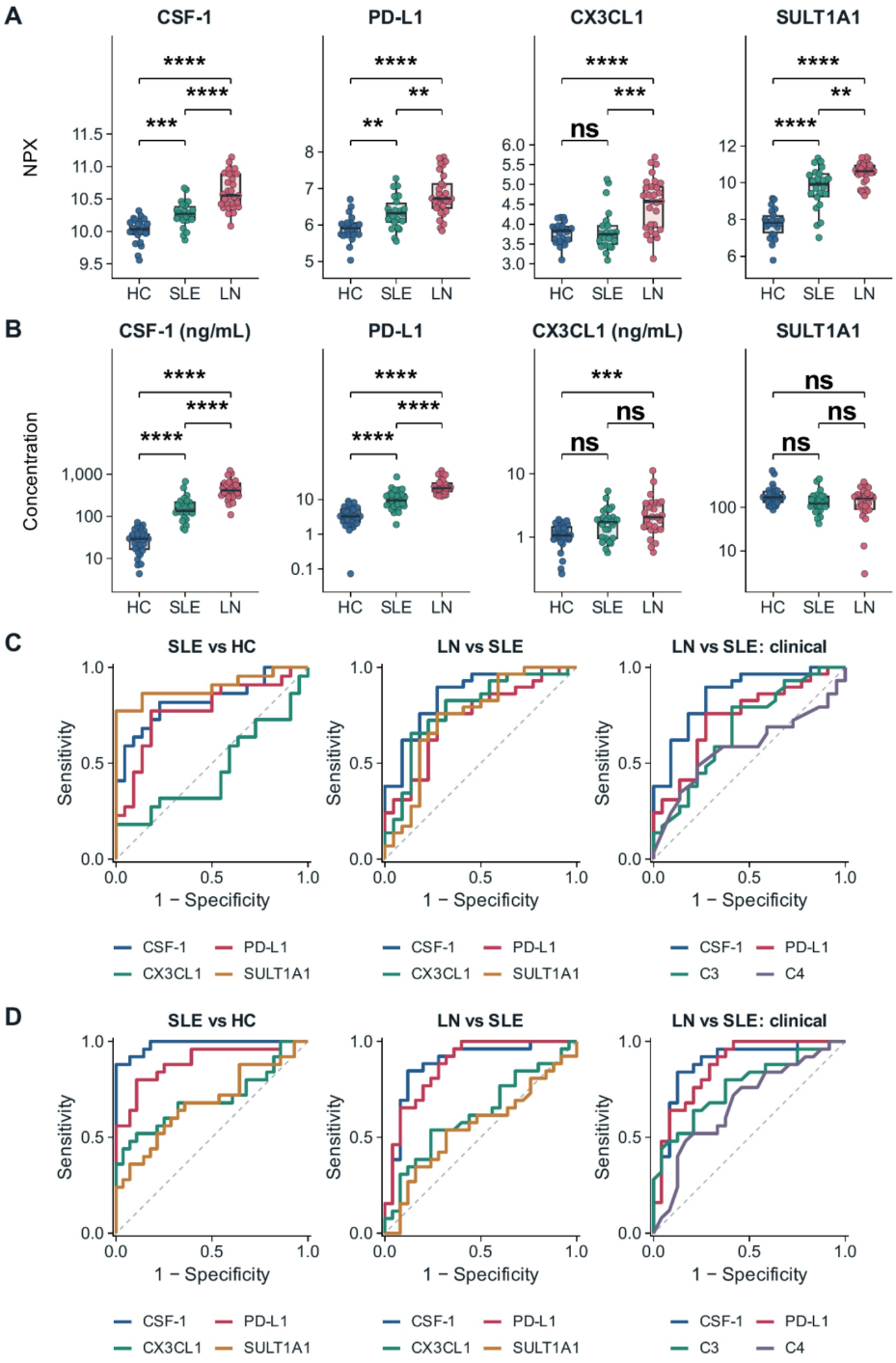


Figure 3

Candidate-protein abundance and discrimination in the discovery and ELISA follow-up cohorts.(A) Olink-derived normalised protein expression (NPX; $\log_2$ scale) of CSF-1, PD-L1, CX3CL1 and SULT1A1 in HCs (n=22), patients with non-renal SLE (n=22) and patients with LN (n=29). (B) ELISA-derived concentrations of the four proteins in HCs (n=28), patients with non-renal SLE (n=25) and patients with LN (n=26). Concentrations are shown on a logarithmic scale. Points represent participants; boxes show

medians and interquartile ranges (IQRs), and whiskers extend to 1.5×IQR. Olink and ELISA comparisons used Welch's t-test and the Wilcoxon rank-sum test, respectively. Asterisks denote Holm-adjusted pairwise P values. (C) Receiver operating characteristic (ROC) curves based on Olink measurements in the discovery cohort. (D) ROC curves based on ELISA measurements in the follow-up cohort. In C and D, panels from left to right show discrimination between non-renal SLE and HCs; discrimination between LN and non-renal SLE using the four candidate proteins; and comparison of CSF-1, PD-L1, complement component 3 (C3) and complement component 4 (C4) among participants with complete measurements of all four markers. Complete-case comparisons included 51 patients in the discovery cohort (22 with non-renal SLE and 29 with LN) and 49 in the follow-up cohort (24 with non-renal SLE and 25 with LN). The candidate-only ELISA ROC analysis for LN versus non-renal SLE included all 51 patients with SLE in the follow-up cohort (25 with non-renal SLE and 26 with LN). For comparisons with complement markers, higher CSF-1 and PD-L1 and lower C3 and C4 values were oriented towards LN; candidate-only ROC analyses retained the original direction convention. Diagonal lines indicate chance-level discrimination. AUC, area under the curve; ns, not significant. Significance thresholds: *adjusted $P<0.05$, **adjusted $P<0.01$, ***adjusted $P<0.001$ and ****adjusted $P<0.0001$.

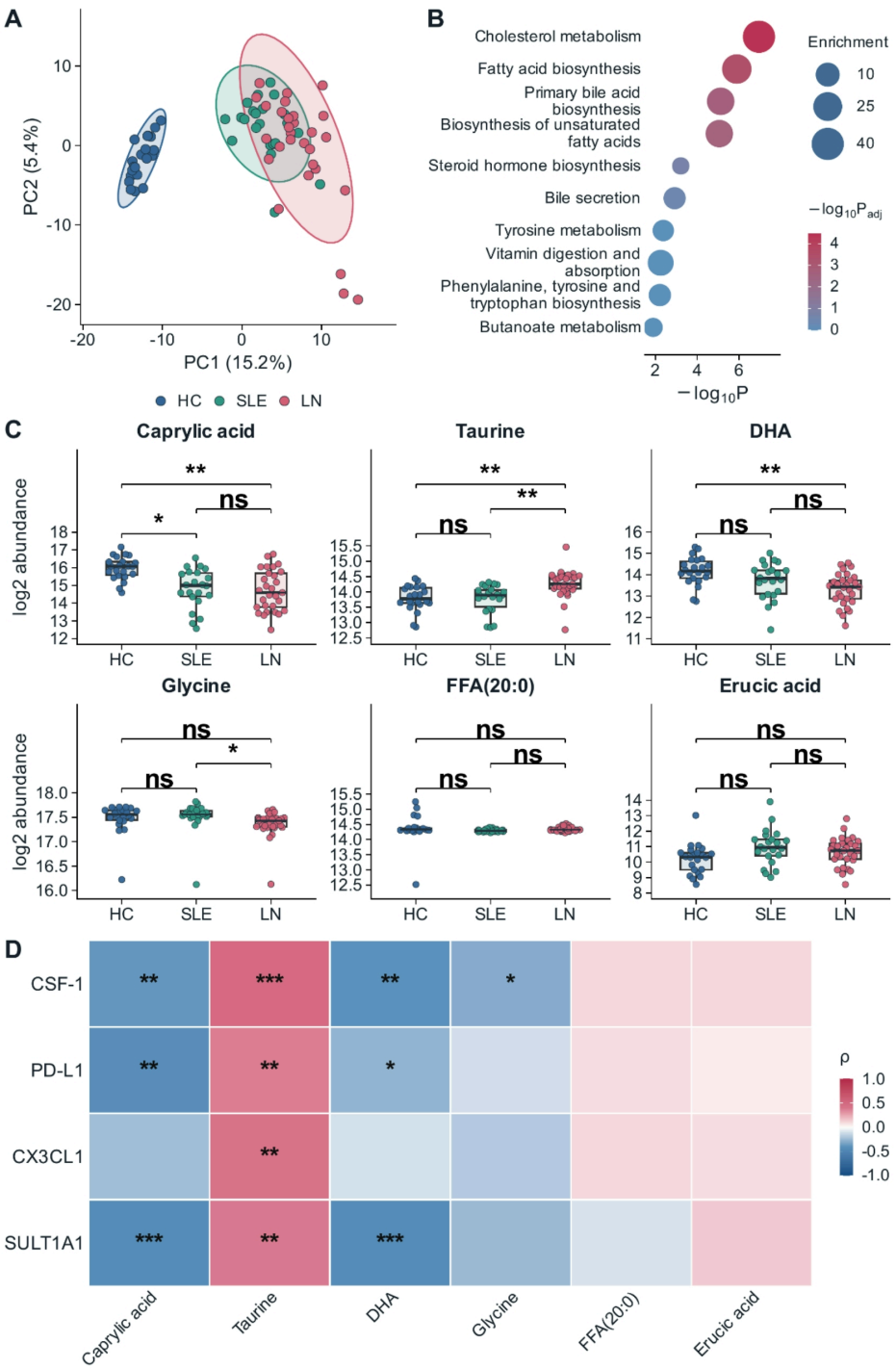


Figure 4

Serum metabolomic profiles and associations with candidate proteins.(A) PCA of 520 metabolomic features in the discovery cohort (n=73). Feature intensities were $\log_2$-transformed, centred and scaled before analysis. (B) Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway over-representation analysis of metabolites assigned to the selected WGCNA modules. The ten displayed pathways are ordered by nominal P value; bubble size represents the enrichment ratio and colour represents $-\log_{10}$ of

the Holm-adjusted P value. Four pathways remained significant after correction. (C) Abundance of six selected metabolites in HCs (n=22), patients with non-renal SLE (n=22) and patients with LN (n=29). Values are $\log_2$-transformed intensities. Points represent participants; boxes show medians and IQRs, and whiskers extend to 1.5×IQR. Pairwise comparisons used Dunn's post hoc test following the Kruskal–Wallis test. Holm correction covered 54 comparisons (three group contrasts for each of 18 candidate metabolites); annotations show the adjusted P values from the original analysis. (D) Spearman correlations between four candidate proteins and six selected metabolites in the discovery cohort (n=73). The heatmap shows the direction and magnitude of the correlation coefficient (ρ). Asterisks indicate BH-adjusted significance across 24 comparisons: *q<0.05, **q<0.01, ***q<0.001 and ****q<0.0001. DHA, docosahexaenoic acid; FFA, free fatty acid.

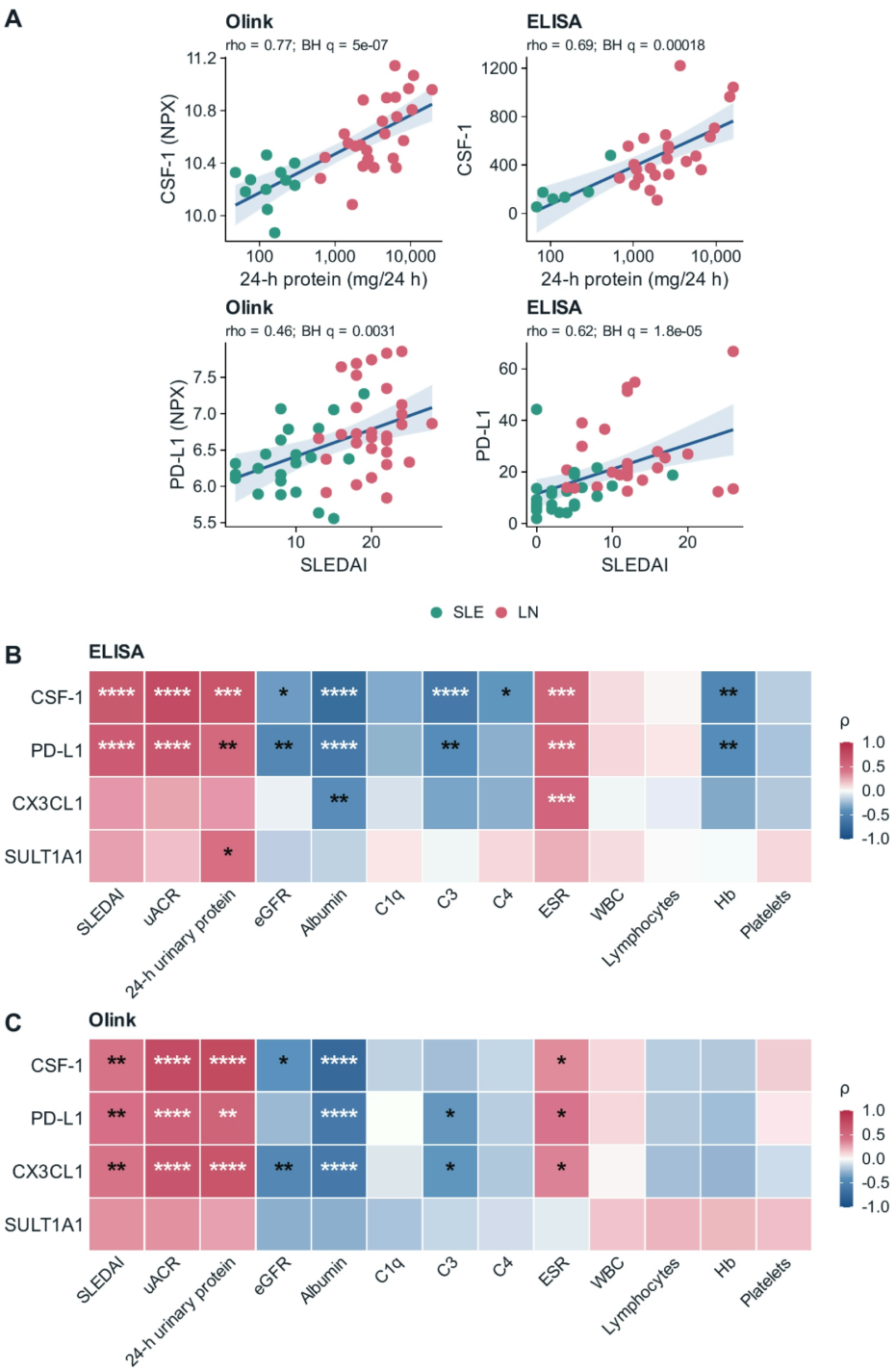


Figure 5

Associations of candidate proteins with disease activity and renal clinical measures.(A) Associations of CSF-1 with 24-hour urinary protein excretion and PD-L1 with SLE Disease Activity Index (SLEDAI) scores in the discovery (Olink; left) and ELISA follow-up (right) cohorts. Analyses included patients with non-renal SLE and LN. Proteinuria is shown on a logarithmic scale. Lines and shaded areas indicate descriptive linear fits and 95% confidence intervals; ρ denotes Spearman's rank correlation coefficient.

Displayed q values are BH-adjusted. (B) Spearman correlations between ELISA-derived concentrations of the four candidate proteins and 13 clinical and laboratory variables in the follow-up cohort. (C) Corresponding correlations based on Olink-derived NPX values in the discovery cohort. Red/coral and blue indicate positive and negative correlations, respectively; colour intensity represents the magnitude of ρ. Measurements dated more than 30 days before or after serum collection were excluded; undated measurements were retained as recorded, and missing values were not imputed. BH correction covered 52 protein–clinical variable comparisons within each cohort. Asterisks indicate BH-adjusted significance: *q<0.05, **q<0.01, ***q<0.001 and ****q<0.0001. M-CSF, macrophage colony-stimulating factor (CSF-1); uACR, urinary albumin-to-creatinine ratio.

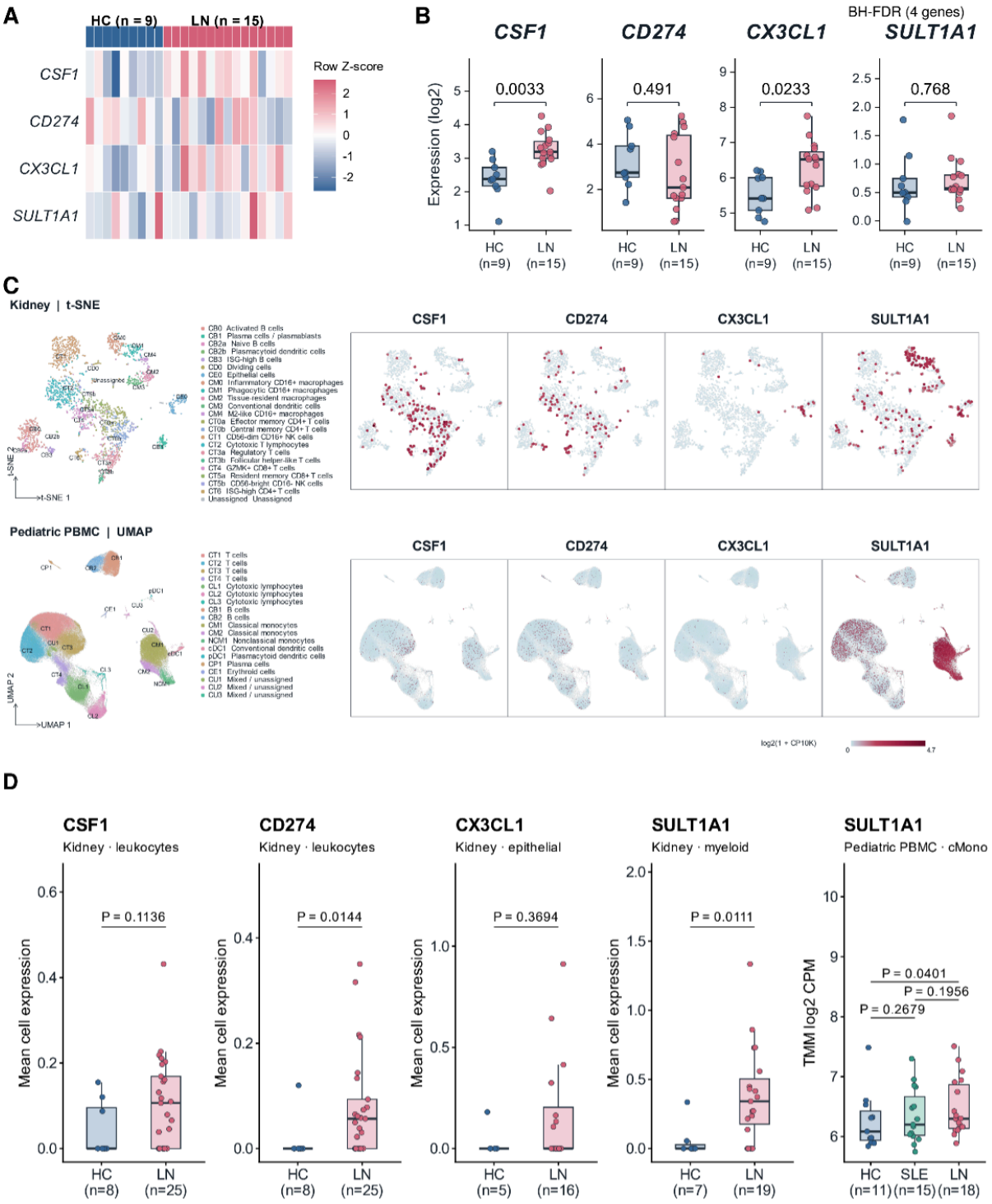


Figure 6

Candidate-gene expression in kidney bulk and single-cell transcriptomic datasets.(A) Row-scaled heatmap of *CSF1*, *CD274*, *CX3CL1* and *SULT1A1*expression in glomerular samples from healthy controls (HCs; n=9) and patients with LN (n=15) in GSE180393. Each column represents one sample. (B) Gene-expression comparisons between HCs and patients with LN. Points represent samples. Two-sided moderated t-tests were performed using limma, with BH correction across the four genes. *CSF1* and

CX3CL1 expression was higher in LN samples (adjusted P=0.0033 and 0.0233, respectively). (C) Recomputed t-distributed stochastic neighbour embedding (t-SNE) maps of adult kidney cells and uniform manifold approximation and projection (UMAP) maps of paediatric PBMCs, showing annotated cell populations and expression of the four candidate genes. Original kidney annotations were retained; display identifiers were used for PBMC clusters. Within each dataset, cell-population and gene-expression maps use the same coordinates. Expression is shown as $\log_2$(1+counts per 10,000 [CP10K]). (D) Donor-level comparisons of *CSF1*and *CD274* in kidney leucocytes (HC n=8; LN n=25), *CX3CL1* in kidney epithelial cells (HC n=5; LN n=16), *SULT1A1* in kidney myeloid cells (HC n=7; LN n=19), and *SULT1A1* in paediatric PBMC classical monocytes (HC n=11; non-renal SLE n=15; LN n=18). Points represent donors. Kidney values are mean cell-level $\log_2$(1+CP10K), including cells without detectable expression; groups were compared using two-sided Wilcoxon rank-sum tests. PBMC expression was analysed using donor-level pseudobulk profiles with trimmed mean of M-values (TMM)-normalized $\log_2$ counts per million (CPM) and edgeR quasi-likelihood models adjusted for age, sex and batch. The figure displays nominal P values for kidney leucocyte *CD274* (0.0144), kidney myeloid *SULT1A1* (0.0111) and the PBMC classical-monocyte comparisons (0.2679, 0.0401 and 0.1956). Boxplots show medians and IQRs; whiskers extend to observations within 1.5×IQR. These analyses and cell-population selections were exploratory; P values were not adjusted for multiple testing. The aggregate leucocyte comparison may reflect cellular composition as well as gene expression, and the *CD274* T-cell comparison was not nominally significant. *SULT1A1* expression was detected in kidney myeloid cells from two of seven healthy donors. Sensitivity analyses using cell numbers and additional populations are provided in the supplementary analyses. The paediatric PBMC and adult kidney datasets represent distinct populations and should not be compared directly. HC, healthy control; LN, lupus nephritis; PBMC, peripheral blood mononuclear cell; SLE, systemic lupus erythematosus.