KI REPORTS
KIReports.org

Complement Proteins Identify Rapidly Progressive Diabetic Kidney Disease

Donghwan Yun[1,2], Sohyun Bae[1], Yuqian Gao[3], Lauren Lopez[4], Dohyun Han[5], Carrie D. Nicora[3], Tae Youn Kim[6], Kyung Chul Moon[7], Dong Ki Kim[1], Thomas L. Fillmore[3], Yon Su Kim[1,2], Avi Z. Rosenberg[8], Weijie Wang[9], Pinaki Sarder[10], Wei-Jun Qian[3], Maryam Afkarian[11] and Seung Seok Han[1]

[1]Department of Internal Medicine, Seoul National University College of Medicine, Seoul, Korea; [2]Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul, Korea; [3]Biological Sciences Division, Pacific Northwest National Laboratory, Richland, Washington, USA; [4]Division of Nephrology, Department of Medicine, University of California, Davis, California, USA; [5]Transdisciplinary Department of Medicine and Advanced Technology, Seoul National University Hospital, Seoul, Korea; [6]School of Nursing, University of California, Davis, Sacramento, California, USA; [7]Department of Pathology, Seoul National University College of Medicine, Seoul, Korea; [8]Department of Pathology, Johns Hopkins University, Maryland, USA; [9]Department of Internal Medicine, University of California, Berkeley, Berkeley, California, USA; [10]Department of Medicine-Quantitative Health, University of Florida College of Medicine, Gainesville, Forida, USA; and [11]Department of Internal Medicine, University of California, Davis, California, USA

Introduction: Mechanisms underlying diabetic kidney disease (DKD) progression remain incompletely understood. This study used untargeted and targeted mass spectrometry-based proteomics in 2 independent cohorts to capture rapidly progressive DKD.

Methods: We conducted untargeted and targeted mass spectrometry on urine samples from Korean patients with type 2 diabetes and biopsy-confirmed diabetic nephropathy (SNUH-DN cohort; $n = 64$) and a DKD subgroup of the Chronic Renal Insufficiency Cohort (CRIC-T2D; $n = 282$), respectively. Urine proteins associated with kidney disease progression (doubling of serum creatinine, $\geq$ 50% decrease in estimated glomerular filtration rates [eGFRs], or progression to end-stage kidney disease[ESKD]) were identified after adjusting for eGFR, proteinuria, and other clinical variables.

Results: In the SNUH-DN patients, urine proteins clustered into 2 groups, with cluster 1 exhibiting a 4.6-fold higher hazard of disease progression (95% confidence interval [CI]: 1.9–11.5) than cluster 0. Proteins in cluster 1 mapped to 10 pathways, 4 of the top 5 being complement-related. A high complement score, derived from urine complement protein abundance, correlated with histopathologic features of DKD and conferred a 2.4-fold greater hazard of disease progression (95% CI: 1.0–5.4) than a low complement score. In CRIC-T2D, targeted mass spectrometry similarly confirmed that complement score stratified patients into rapid and slow DKD progression groups. In both cohorts, complement score exhibited a linear association with disease progression.

Conclusion: The strong association between complement activation and rapid DKD progression highlights the need to explore complement inhibition as a potential therapeutic strategy for DKD.

Kidney Int Rep (2025) ■, ■–■; https://doi.org/10.1016/j.ekir.2025.04.061
KEYWORDS: biomarkers; complement system proteins; diabetic nephropathies; pathology; proteomics

DKD is a major complication of diabetes, significantly contributing to morbidity and mortality worldwide.[1,2] As the leading cause of ESKD globally, DKD accounts for about half of all ESKD cases, thereby posing a significant burden on health care systems worldwide.[3] Despite recent advances in diabetes management, the mechanisms underlying progression (and their respective biomarkers) in human DKD remain incompletely understood. Clinically, this translates into highly variable rates of disease progression, with some patients experiencing rapid progression leading to ESKD.[4,5] This variability presents substantial challenges in clinical practice, because it is currently impossible to identify and intervene in patients who are rapid progressors.

Correspondence: Seung Seok Han, Department of Internal Medicine, Seoul National University College of Medicine, 103 Daehak-ro, Jongno-gu, Seoul, 03080 Republic of Korea. E-mail: hansway7@snu.ac.kr; or Maryam Afkarian, Room 5317, Genome and Biomedical Sciences Facilities, 451 Health Sciences Drive, University of California, Davis, California 95616, USA. E-mail: mafkarian@ucdavis.edu

Received 8 February 2025; revised 14 April 2025; accepted 28 April 2025

Although extensive research is being conducted, the pathophysiological mechanisms responsible for the DKD progression still need to be completely understood. Traditional risk factors, such as hyperglycemia, hypertension, and diabetes duration, are recognized[6,7]; however, they only partially explain the interindividual differences in disease progression. Previous studies have explored various biomarkers, including cytokines and kidney injury–relevant molecules[8-10]; however, a comprehensive understanding that integrates these factors with clinical outcomes remains elusive. This knowledge gap hinders the development of precise predictive models and targeted interventions that are essential for preventing rapid progression in patients with DKD. Therapeutic strategies such as renin-angiotensin-system inhibitors, sodium-glucose cotransporter-2 inhibitors, and mineralocorticoid receptor antagonists are employed to mitigate DKD progression.[11-13] Despite these advances in therapy, it is evident that more targeted therapies are needed to improve DKD outcomes.

In response to these challenges, omics technologies, including genomics, transcriptomics, metabolomics, and proteomics, emerge as powerful tools for unraveling the complex pathobiology of DKD. We used comparative proteomic analyses across 2 independent cohorts (South Korea and the USA) to identify the pathways associated with rapid DKD progression in type 2 diabetes. Recent advances in urinary proteomics have significantly enhanced the ability to predict prognosis and analyze subtypes in DKD by revealing the molecular dynamics of kidney injury.[14,15] This approach allows for the development of personalized treatment plans based on specific molecular profiles.

METHODS

Study Patients

From March 2011 to June 2021, 93 patients with type 2 diabetes and biopsy-proven diabetic nephropathy were recruited at the Seoul National University Hospital (SNUH), referred to as the SNUH-DN cohort. Of this group, 29 were excluded because of eGFR < 30 ml/min per 1.73 m^2 ($n = 25$), combined type 1 and type 2 diabetes ($n = 2$), the presence of concomitant nondiabetic glomerular disease ($n = 1$), or the absence of glomeruli in biopsied kidney tissues ($n = 1$). The remaining 64 patients were included in the final analyses as the SNUH-DN cohort after conducting untargeted proteomics.

In addition, this study incorporated a subset of patients with type 2 diabetes from the CRIC study, termed CRIC-T2D, consisting of American patients with risk factors for cardiovascular disease, progression of chronic kidney disease, and mortality. A total of 3939 patients aged 21 to 74 years with eGFR of 20 to 70 ml/min per 1.73 m^2 were enrolled at 7 clinical centers throughout the USA from June 2003 to December 2008. CRIC oversampled participants with diabetes, who constitute nearly half of the CRIC cohort. This study excluded participants with monogenic renal disease, liver cirrhosis, class III or IV heart failure, HIV, cancer, autoimmune disease, polycystic kidney disease, pregnant women, recipients of organ or bone marrow transplants, and those receiving immunotherapy for primary renal disease or systemic vasculitis within the 6 months preceding recruitment or systemic chemotherapy.

Ethical Considerations

This study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki and was approved by the institutional review boards of all participating institutions. For the SNUH-DN cohort, ethical approval was granted by the institutional review board at SNUH (no. H-2403-066-1519). The CRIC study was approved by the institutional review boards at all participating institutions and was conducted in accordance with the Declaration of Helsinki.[16,17] Urine samples and data for the present study were obtained from the CRIC subset housed at the NIDDK Central Repository (R01 5R01DK104706). The use of deidentified CRIC samples was approved by the institutional review board at the University of California, Davis. The use of deidentified samples and data from the CRIC-T2D subgroup from the NIDDK Central Repository was approved by the Human Subjects Division, University of California, Davis, California. All participants provided written informed consent, having been fully informed of the study purpose, the procedures involved, potential risks, and their rights as study participants, including the right to withdraw at any point without consequence. The confidentiality of patient data was rigorously maintained throughout the study. All personal identifiers for SNUH-DN patients were removed, and data were stored securely to prevent unauthorized access. CRIC-T2D data and samples were deidentified at the NIDDK Central Repository.

Study Outcomes

The primary outcome was the progression of DKD (referred to as kidney disease progression) based on any of the following events: a doubling of baseline serum creatinine, a decline of $> 50\%$ in eGFR, or the onset of ESKD. In the SNUH-DN cohort, the baseline for outcome assessment was defined as the date of the kidney biopsy, whereas in the CRIC-T2D subcohort, the baseline was set at the date of cohort enrollment.

eGFR was calculated using the Chronic Kidney Disease Epidemiology Collaboration equation.[18] ESKD was specifically defined as the need for kidney replacement therapy, including hemodialysis, peritoneal dialysis, or kidney transplantation.

Study Variables

We collected comprehensive patient data encompassing demographic characteristics, such as age; sex; race; body mass index; diabetes duration; and comorbidities of hypertension, ischemic heart disease, cerebrovascular disease, retinopathy, and neuropathy. Laboratory findings included hemoglobin A1c, eGFR, and random urine protein-to-creatinine ratio for the SNUH-DN cohort, whereas the CRIC-T2D subcohort included the 24-hour urine protein amount.

We employed a detailed histopathologic classification system to evaluate kidney biopsies from patients diagnosed with diabetic nephropathy in the SNUH-DN cohort. This classification system follows the criteria described in the reference paper.[19] Each biopsy was required to contain at least 10 glomeruli for accurate assessment, excluding any incomplete glomeruli along the biopsy edges. The glomerular classification was segmented into the following 5 categories: class I, isolated glomerular basement membrane thickening; class IIA, mild mesangial expansion; class IIB, severe mesangial expansion; class III, nodular sclerosis (Kimmelstiel-Wilson lesions); and class IV, advanced diabetic glomerulosclerosis with $>$ 50% global glomerulosclerosis. The assessment of interstitial fibrosis and tubular atrophy, interstitial inflammation, arterial hyalinosis, and arteriosclerosis was quantified using a 4-point or 3-point scale according to the criteria.

Urine Sample Collection

In the SNUH-DN cohort, urine samples were collected at the time of kidney biopsy. Patients were instructed to provide a midstream clean-catch urine sample, which was immediately processed to minimize protein degradation. Upon collecting, urine samples were centrifuged at 3000 rpm for 10 minutes to remove cellular debris and then aliquoted into cryovials. The aliquots were promptly frozen and stored at -80 °C to preserve the integrity of the proteins until preparation for proteomics.

In the CRIC-T2D subcohort, 24-hour urine samples were collected at home by study participants during their CRIC visit number 3. Participants were instructed to keep the samples at 4 °C during collection. Collections that were $<$ 500 ml or collected for $<$ 22 hours or $>$ 26 hours were repeated. Accepted 24-hour collections were thoroughly mixed, and three 10 ml aliquots were obtained. The aliquots were shipped to the CRIC central laboratory on ice packs, where they were further aliquoted and stored at -80 °C. A portion of these samples was shipped to the NIDDK Central Repository when they were sent to a study author (MA). An aliquot of these samples was then shipped to the Pacific Northwest National Laboratory on dry ice and were prepared for targeted proteomics.

Untargeted Proteomics

Sample Preparation

A volume of 1 to 2 ml of urine from each participant was concentrated down to 250 µl using a centrifugal filter with a molecular weight cutoff of 3 kDa (Millipore, Billerica, MA). Protein concentrations were then determined using the Bradford method, employing a commercial kit (Bio-Rad, Hercules, CA). For the analysis, 50 µg of protein from each sample was precipitated using ice-cold acetone and added at 5 times the protein volume. The precipitate was redissolved in 50 µl of SDT buffer (comprising 2% SDS, 0.1 M dithiothreitol, in 0.1 M Tris HCl, pH 8.0), and subjected to heating at 95 °C to denature the proteins. Following denaturation, proteins were enzymatically digested using a filter-aided sample preparation approach, adapted from previously established methods with some modifications.[20] Specifically, proteins were transferred to an Amicon 30 K filter (Millipore, Billerica, MA) and washed multiple times with UA solution (8 M urea in 0.1 M Tris-HCl, pH 8.5) through centrifugation. After 3 washes, cysteines were alkylated using 0.05 M iodoacetamide in the UA solution for 30 minutes at room temperature in the dark. Subsequent buffer exchanges to 40 mM ammonium bicarbonate were performed twice to prepare for enzymatic digestion. Overnight digestion was carried out at 37 °C using trypsin/LysC with a ratio of enzyme to substrate of 1:100. The peptides produced were then collected in new tubes through additional centrifugation, and an extra elution step was performed using 40 mM ammonium bicarbonate mixed with 0.5 M sodium chloride. The eluted peptides were purified and fractionated using custom-made styrene-divinylbenzene reversed-phase sulfonate-StageTips, employing a gradient of acetonitrile (40%, 60%, and 80%) in 1% ammonium hydroxide.[21,22] Following fractionation, peptides were vacuum-dried and stored at -80 °C until further analysis.

Construction of a Peptide Spectral Library

To create a peptide spectral library for using match between runs algorithm of maxquant, a pooling of urine proteins was performed.[23] We combined equal subaliquots of proteins, each containing 5 µg, to a total of 100 µg. This pooled sample was then subjected to a

2-step filter-aided digestion process, consistent with the methods previously outlined.[21,24] Following digestion, peptides were purified using Oasis HLB solid-phase extraction cartridges to ensure their readiness for high-performance liquid chromatography (LC). For the library construction, we used 100 μg of these clean peptides and processed them through an Agilent 1260 bioinert high-performance LC system (Agilent, Santa Clara, CA) equipped with a standard analytical column (4.6 × 250 mm; 5-μm particle size). The peptides were fractionated using a high-pH reversed-phase method, employing a flow rate of 0.8 ml/min across a 60-minute gradient. Solvents included 15 mM ammonium hydroxide in water (solvent A) and 15 mM ammonium hydroxide in 90% acetonitrile (solvent B). We collected 96 separate fractions at 30-second intervals throughout a 48-minute period, which were then concatenated into 24 fractions in stepwise manner. Early-, middle-, and late-eluting peptides were pooled by mixing every 24th original fraction for the proteome (e.g., combining fractions 1, 25, 49, and so on).[25] These concatenated fractions were subsequently dried using a vacuum centrifuge and stored at −80 °C, ready for analysis by LC-tandem mass spectrometry (MS/MS).

LC-MS/MS Analytical Procedure

LC-MS/MS analysis was conducted on a Q-Exactive Plus Quadrupole Orbitrap mass spectrometer (Thermo Fisher Scientific, Waltham, MA), which was interfaced with an Ultimate 3000 RSLC nano system (Dionex, Sunnyvale, CA) featuring a nanoelectrospray ion source. The process incorporated slight modifications from previously established protocols.[20,22] Individual peptide fractions from urine samples underwent separation using a dual-column arrangement consisting of a C18 trap column (300 μm internal diameter × 0.5 cm, with 3 μm particles and 100 Å pore size) and a C18 analytical column (50 μm internal diameter × 50 cm, with 1.9 μm particles and 100 Å pore size). Before injection, dried peptide fractions were dissolved in solvent A (2% acetonitrile and 0.1% formic acid). The nano-LC system loaded the samples and applied a chromatographic gradient over 90 minutes, increasing from 8% to 30% of solvent B (100% acetonitrile with 0.1% formic acid). The setup operated under a spray voltage of 2.0 kV in positive ion mode, and the capillary heater was maintained at 320 °C. Mass spectrometric data were collected in a data-dependent acquisition format, prioritizing the top 15 most abundant precursor ions. The Orbitrap detector analyzed precursor ions across a mass range from 300 to 1650 m/z, achieving a resolution of 70,000 at m/z 200. Fragmentation was performed via higher-energy collisional dissociation at a resolution of 17,500 for m/z 200, using a normalized collision energy setting of 28. The system's ion injection times were set to a maximum of 20 ms for survey scans and 120 ms for MS/MS scans.

Label-Free Quantitative Data Analysis

For the analysis of mass spectrometric data, we employed MaxQuant software (version 1.6.1.0) for label-free quantification.[26] MS/MS spectra were meticulously matched against the Human UniProt protein sequence database (as of December 2014, containing 88,657 entries) via the Andromeda search engine.[27] The analysis parameters were carefully selected, with a precursor ion tolerance set at 6 ppm for total protein analysis and an MS/MS ion tolerance of 20 ppm. We designated carbamido-methylation of cysteine as a fixed modification, whereas N-acetylation of proteins and methionine oxidation was considered variable modifications. The digestion parameter was configured for full tryptic digestion with allowances for up to 2 missed cleavages per peptide, which had to be at least 6 amino acids long. We rigorously maintained a false discovery rate of 1% across peptide, protein, and modification levels to ensure high data integrity. In addition, to enhance quantification accuracy across samples, we used a strategy of matching between runs anchored by the peptide library constructed from pooled urine samples. The Intensity-Based Absolute Quantification algorithm, a component of the MaxQuant suite, facilitated the estimation of protein abundances.[28] In the Intensity-Based Absolute Quantification algorithm, raw intensities were normalized by the count of theoretical peptides, rendering them reflective of the proteins' molar quantities.

Targeted Proteomics

SRM

Assay development was conducted as previously described in detail.[29] Briefly, final surrogate peptides were selected based on the uniqueness of their occurrence in proteins of interest, as well as their presence and chromatographic behavior in shotgun MS/MS data from our previous urine proteomics data repository (Supplementary Table S1). Crude synthetic heavy isotope-labeled versions of the selected peptides were used to generate MS/MS data on an Orbitrap higher-energy collisional dissociation. This data were used to identify the best transitions and optimal collision energy for each transition. Notably, only the nonglycated form of CD59 was identified in both the SNUH-DN and CRIC-T2D cohorts.

Sample Processing and Digestion

Urine samples were processed and underwent LC-MS/MS, as described previously.[29] Briefly, urine

samples were thawed on ice, run through a 10 kDa prewashed (50 mM AMBIC, i.e., NH_4HCO_3, pH 8.0) Amicon ultrafiltration column, and underwent buffer exchange with AMBIC twice. The final urine protein concentration was determined using the BCA protein assay. Urea was added to the urine sample to a final concentration of 8 M, followed by reduction with 500 mM DTT to a final concentration of 10 mM, brief sonication, and incubation at 37 °C for 1 hour while shaking. Alkylation was performed by adding IAA to a final concentration of 40 mM and incubation for 1 hour at 37 °C in the dark while shaking. The sample was then diluted 10-fold with digestion buffer, and trypsin was added to a protein-to-trypsin ratio of 50:1 (w/w), followed by incubation at 37 °C for 3 hours. Digestion was stopped by adding TFA to a final concentration of 0.1%. The tryptic digest was separated using SPE C18 columns, preconditioned using methanol and then SPE conditioning buffer, and washed with SPE washing buffer. Peptides were eluted from the SPE C18 columns with 1 ml SPE elution buffer and, dried under a reduced vacuum using a speed vac and redissolved in MS-grade water. The peptide concentration was determined using the BCA protein assay, and peptide samples were stored at −80 °C until further use. Heavy isotope-labeled peptide standards were mixed in a single 1 ml stock solution, having been reconstituted at 1000 fmol/ml per peptide in 0.1% TFA in water. The digested peptide mix, and the heavy peptide stock were combined while shaking and advanced to LC-selected reaction monitoring (LC-SRM) analysis.

LC-SRM Analysis and Quantification

The urine sample–heavy peptide mixes were analyzed using the final transition list, derived from the orbitrap higher-energy collisional dissociation MS/MS, and optimized in 1 typical urine sample. LC-SRM was performed using a nanoACQUITY UPLC system coupled online to a TSQ Vantage triple quadrupole mass spectrometer. The peak area ratios of endogenous transitions and heavy internal standard transitions represent the molar ratios between the amounts of endogenous and heavy internal standard peptides. The peak area ratios were calculated using Skyline software.[30] The data were imported into a Skyline file, peak boundaries for each peptide were manually verified for each dataset, and detection and optimal transition (for best peak area ratio) were confirmed for each peptide. Protein concentration was calculated using the equation below.

$$\text{Target protein concentration}\left(\frac{\text{ng}}{\text{mL}}\right) = \frac{\text{L}}{\text{H}} \times C_{\text{heavy,LC-SRM}} \times C_{\text{digest,LC-SRM}} \times C_{\text{protein}} \times 400\ \mu\text{l} \times \text{MW} \times \frac{10^{-6}\ \text{ng}}{\text{fg}} \div 10\ \text{ml}$$

where L/H is the peak area ratio of endogenous (L) to heavy isotope-labeled internal standard peptide (H), $C_{\text{heavy, LC-SRM}}$ is the molar concentration of heavy internal standard peptides (fmol/μl) in the final LC-SRM solutions, MW is the molecular weight of targeted protein (Da or g/mol), $C_{\text{digest, LC-SRM}}$ is the mass concentration of total protein digest (μg/μl) in the final LC-SRM solutions, and C_{protein} is mass concentration of total protein (μg/μl) in the 400 μl concentrated retentate from 10 ml of original urine. Among the analyzed proteins, complement 8 beta chain, gamma chain, and complement factor D did not meet the following criteria and were therefore excluded from the analysis; (i) dot product light to heavy > 0.86, (ii) maximum height of light > 2000, and (iii) light/heavy ratio > light/heavy ratio of heavy in Shewanella peptides.

Proteomics Data Processing and Analysis

The proteomic data obtained using label-free quantification of urine proteins were initially processed using the *Seurat* package (version 5.0.3).[31] These values were normalized for urine dilution by indexing (division) by urine creatinine, and then log2-transformed for further normalization using the *NormalizeData* function from the *Seurat* package. Finally, for each identified protein, the normalized data were scaled using the *ScaleData* function from the *Seurat*. Unsupervised clustering techniques were used to group proteins based on their expression patterns, and the *FindNeighbors* and *RunUmap* functions were used to create cluster groups and generate UMAP visualizations, respectively. Gene names related to each protein were identified using *UniProt* accession numbers.[32] Differentially expressed proteins between rapid and slow progressors were identified using the *FindMarkers* function. Proteins with well-documented gene associations were selected for pathway analysis via the *UniProt* database, and differentially represented pathways were identified using *WikiPathways*.[33] When each group contained a sufficient number of genes, Bonferroni-adjusted *P*-values were used for gene assignment; otherwise, nonadjusted *P*-values were applied.

Complement Scores

The complement score was devised to evaluate the prevalence of complement pathway proteins in urinary proteomics data. We identified the proteins involved in the complement pathway by referencing the Kyoto Encyclopedia of Genes and Genomes database.[34] The complement proteins used for scoring are listed in Supplementary Tables S1 and S2 for targeted and

untargeted proteomics, respectively. The average protein expression of denoted genes was calculated based on the assumption of previous studies.[35-37] In addition, we formalized the distinction between activators and inhibitors. For each patient, this score was calculated by taking the mean of the normalized abundance of identified complement proteins, adjusted for their functional roles; inhibitory proteins were assigned a modifier of −1, and activating proteins a modifier of 1. The formula used was:

$$\text{Complement score}_i = \sum_{j=1}^{n}\left(\text{Normalized expression}_{ij} \times \text{Modifier}_j\right) \Big/ n$$

where i denotes the sample number, and j is the individual complement protein. The continuous variable complement score was binarized into low versus high complement score bins by the median value.

Statistical Analysis

Descriptive statistical analyses were conducted on patient information. For continuous variables with a normal distribution, data were presented as the mean ± SD, and for those without a normal distribution, as medians with interquartile ranges (IQR). The Kolmogorov-Smirnov test was used to assess the normality of the data distribution. Categorical variables were shown as percentages. Comparison of categorical variables was performed using either the chi-square test or Fisher exact test, whereas continuous variables were compared using either t test or the Mann-Whitney U test, depending on their distribution. When comparing multiple groups for continuous variables, analysis of variance was used, with the groups treated as ordinal variables. Pairwise comparisons were conducted with Bonferroni correction.

Survival analysis was employed to investigate the association between variables and the risk of kidney disease progression. The hazard ratio and corresponding 95% CIs were calculated using the Cox proportional hazards regression model. Variables included in the model were selected based on clinical relevance, such as age, sex, body mass index, diabetes duration, and baseline kidney function. Adjusted survival curves were then generated to visualize the differences in adjusted cumulative survival between the 2 groups using the *adjustedCurves* package.[38] In addition, penalized spline regression was used to further elucidate the relationship between complement scores and kidney disease progression. All statistical analyses were conducted using the R program (version 4.2.2; https://www.R-project.org/). A 2-sided P-value < 0.05 was considered statistically significant.

Table 1. Baseline characteristics of the SNUH-DN cohort

Variables	Total ($N = 64$)	Cluster 0 ($n = 40$)	Cluster 1 ($n = 24$)	P
Age (yrs)	53.3 ± 11.1	52.1 ± 10.9	55.2 ± 11.5	0.282
Male (%)	71.9	75.0	66.7	0.667
Body mass index (kg/m^2)	25.0 ± 3.3	25.2 ± 3.1	24.7 ± 3.5	0.594
Hypertension (%)	92.2	95.0	87.5	0.355
Diabetes duration (yrs)	10 (4–14)	8 (2–12)	11 (4–20)	0.278
Diabetic retinopathy (%)	54.7	52.5	58.3	0.846
Diabetic neuropathy (%)	29.7	30.0	29.2	1.000
Ischemic heart disease (%)	7.8	12.5	0	0.148
Cerebrovascular disease (%)	4.7	2.5	8.3	0.551
Hemoglobin A1c (%)	7.2 (6.6–8.4)	7.2 (6.6–7.6)	7.4 (6.8–8.9)	0.178
PCR (g/g)	3.1 (1.7–7.0)	2.2 (1.2–3.6)	7.1 (4.9–9.5)	<0.001
eGFR (ml/min per 1.73 m^2)	56 (44–75)	61.2 (48–85)	43 (33–60)	<0.001
Glomerular classification (%)				
I	14.1	17.5	8.3	0.003
IIA	12.5	17.5	4.2	
IIB	32.8	42.5	16.7	
III	12.5	10.0	16.7	
IV	28.1	12.5	54.2	
Interstitial fibrosis and tubular atrophy (%)				
0	4.7	7.5	0.0	<0.001
1	35.9	47.5	16.7	
2	46.9	45.0	50.0	
3	12.5	0.0	33.3	
Interstitial inflammation (%)				
0	10.9	17.5	0.0	0.044
1	84.4	80.0	91.7	
2	4.7	2.5	8.3	
Arteriolar hyalinosis (%)				
0	25.0	35.0	8.3	0.052
1	28.1	22.5	37.5	
2	46.9	42.5	54.2	
Arteriosclerosis (%)				
0	25.0	27.5	20.8	0.674
1	35.9	37.5	33.3	
2	39.1	35.0	45.8	
Follow-up time (yrs)	6.2 (2.4–7.8)	7.1 (4.9–8.4)	2.7 (1.1–5.1)	<0.001

eGFR, estimated glomerular filtration rate; PCR, random urine protein-to-creatinine ratio.

RESULTS

Baseline Characteristics of the SNUH-DN Cohort

This cohort had a mean age of 53.3 ± 11.1 years, with 71.9% male subjects. Hypertension was present in 92.2% of the subjects, and the median duration of diabetes was 10 (IQR: 4–14) years. Diabetic retinopathy and neuropathy were observed in 54.7% and 29.7% of all patients, respectively. The median hemoglobin A1c was 7.2% (IQR: 6.6%–8.4%), whereas the median random urine protein-to-creatinine ratio and eGFR values were 3.1 (IQR: 1.7–7.0) g/g and 55 (IQR: 44–75) ml/min per 1.73 m^2, respectively. Details of histopathologic features are provided in Table 1.

Proteomic Clustering in the SNUH-DN Cohort

Unsupervised clustering of the Intensity-Based Absolute Quantification–normalized urine protein abundances from the SNUH-DN cohort identified 2 distinct clusters, designated as clusters 0 and 1 (Figure 1a). Heatmap of Intensity-Based Absolute Quantification–normalized urine proteins that were abundant and scarce in clusters 0 and 1 are illustrated in Supplementary Figure S1A and B, respectively. These data showed that of the 1877 identified proteins, 71 and 85 were significantly abundant and scarce in cluster 1, compared with cluster 0, after the Bonferroni adjustment, respectively. Many of the abundant proteins mapped to the complement pathway, either complement activators such as complement factor D, component 2, component 8 beta chain, component 6, and component 5 (C5),[39] or complement inhibitors such as complement factor H and serpin family G member 1.[40,41] Conversely, among the scarce proteins, CD55 and CD59 act as complex inhibitors,[42] whereas the protein C receptor functions indirectly as a complement inhibitor.[43]

Kidney disease progression within 1 year after biopsy occurred in 6 out of 24 subjects (25%) with urine proteins in cluster 1, and 2 out of 40 subjects (5%) in cluster 0. Compared with subjects in cluster 0, those in cluster 1 had a 4.62-fold (95% CI: 1.86–11.50) greater hazard of DKD progression in multivariable time-to-event (Cox proportional hazard) regression, adjusted for age, sex, body mass index, hypertension, diabetes

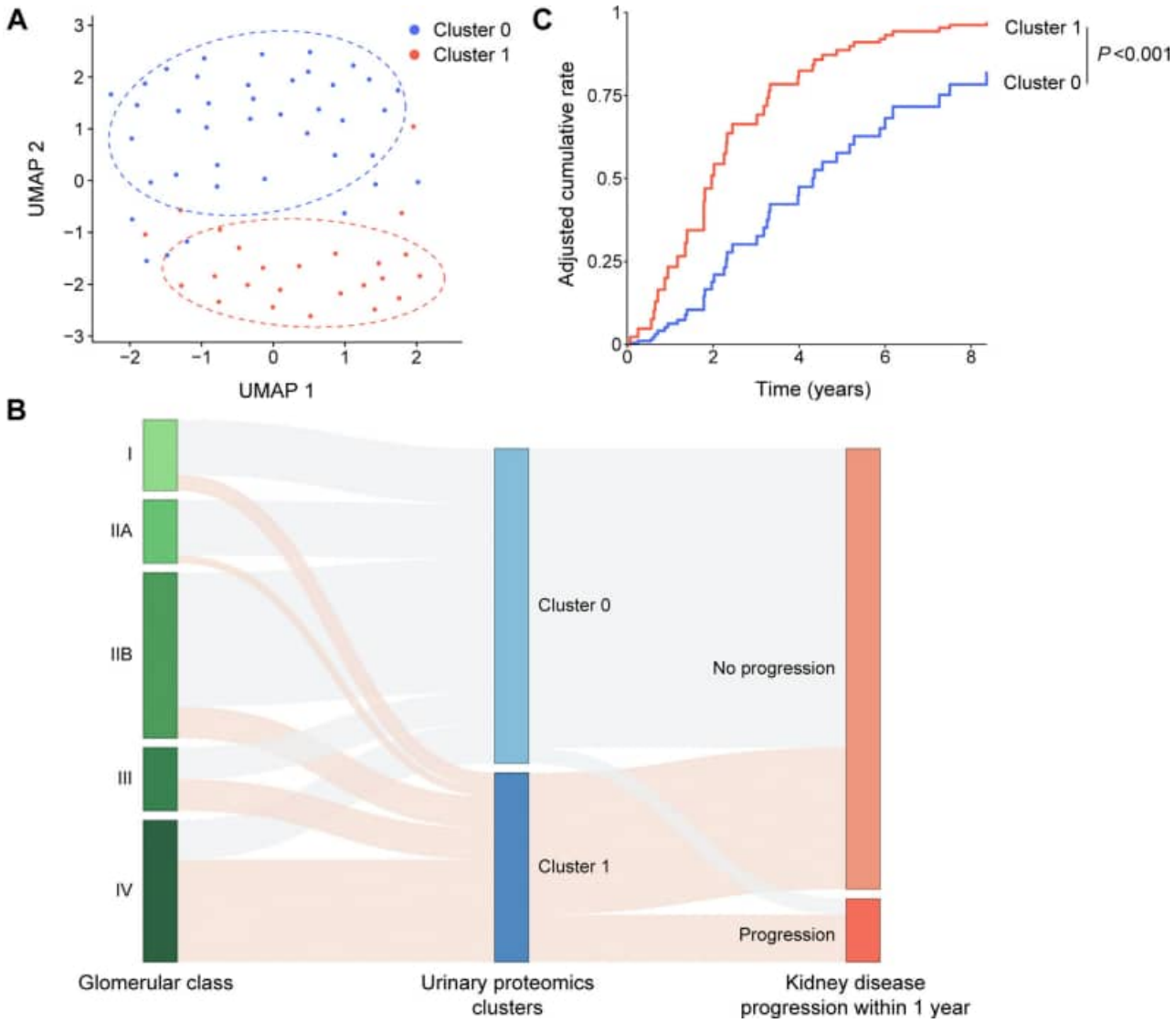

Figure 1. Unsupervised clustering based on untargeted proteomics in the SNUH-DN cohort. (a) Uniform manifold approximation and projection (UMAP) plot with 2 distinct clusters (clusters 0 and 1). (b) Multivariable-adjusted cumulative incidence of kidney disease progression according to clusters, after adjustment for age, sex, body mass index, hypertension, diabetes duration, urine protein-to-creatine ratio, and eGFR. (c) Sankey plot illustrating the relationships among glomerular classes, clusters, and kidney disease progression within 1 year. eGFR, estimated glomerular filtration rate.

Table 2. Association between cluster and kidney disease progression in the SNUH-DN cohort

	Univariable		Multivariable[a]	
Variables	HR (95% CI)	*P*	HR (95% CI)	*P*
Cluster 1 (vs. 0)	4.63 (2.16–9.93)	<0.001	4.62 (1.86–11.50)	0.001
Age (/1 yr)	1.01 (0.98–1.03)	0.759	0.94 (0.90–0.98)	0.004
Female (vs. male)	0.47 (0.20–1.08)	0.074	0.27 (0.10–0.70)	0.007
Body mass index (per 1 kg/m^2)	0.93 (0.83–1.04)	0.224	0.88 (0.78–0.99)	0.039
Hypertension (vs. none)	1.14 (0.35–3.74)	0.827	1.03 (0.29–3.68)	0.964
Diabetes duration (/1 yr)	1.02 (0.98–1.07)	0.380	1.00 (0.95–1.05)	0.969
PCR (per 1 g/g)	1.16 (1.08–1.25)	<0.001	1.16 (1.06–1.28)	0.002
eGFR (per 1 ml/min per 1.73 m^2)	0.98 (0.96–0.99)	0.006	0.99 (0.97–1.00)	0.142

CI, confidence interval; eGFR, estimated glomerular filtration rate; HR, hazard ratio; PCR, random urine protein-to-creatinine ratio.
[a]Adjusted for age, sex, body mass index, hypertension, diabetic duration, proteinuria, and eGFR.

duration, random urine protein-to-creatinine ratio, and eGFR (Table 2). The adjusted cumulative incidence of DKD progression displayed more rapid kidney disease progression in cluster 1 than in cluster 0 (Figure 1b). These results indicate that the unsupervised clustering of urine proteins was able to segregate subjects into 2 clusters, displaying slower and more rapid DKD progression. Notably, the accelerated DKD progression within 1 year and 3 years in cluster 1, compared with cluster 0, was associated with high glomerulosclerosis scores in the paired kidney biopsies from these subjects (Figure 1c and Supplementary Figure S2). Even when the kidney disease progression outcome was divided into subcategories (e.g., a doubling of baseline serum creatinine, a decline of $>$ 50% in eGFR, and the onset of ESKD), an association between cluster 1 and a high glomerulosclerosis score was observed (Supplementary Figure S3).

Different Enrichment in Complement Proteins

To identify pathways associated with kidney disease progression, we performed pathway analysis based on differences in protein expression ratios between the 2 groups. The analysis revealed a notable enrichment of complements and related pathways in cluster 1 (Figure 2a and Supplementary Table S3). The volcano plot of differentially abundant proteins highlighted that complement activators, including complement factor D, component 3 (C3), C5, component 6, and component 8 beta chain, along with some complement inhibitors, including complement factor H, complement factor I, and clusterin, were prominently ranked in cluster 1. Conversely, other complement inhibitors, including CD55 and CD59, were scarce in cluster 1 (Figure 2b). Detailed information about the proteins related to the complement pathway is provided in Supplementary Table 2. A heatmap of normalized complement protein abundance depicted that complement pathway–related proteins were more prominently altered in cluster 1 than in cluster 0 (Figure 2c).

These findings were further supported by comparisons between progressors and nonprogressors at the 1-year timeframe (Supplementary Figure S4A). Notably, even when the criteria for kidney disease progression were extended to 3 years, the complement pathway remained highly ranked (Supplementary Figure S4B), indicating a consistent association between the abundance of complement proteins and kidney disease progression.

Complement Score and Kidney Disease Progression in the SNUH-DN Cohort

To comprehensively assess the effect of complement proteins, we calculated a complement score that considers their expression levels and modifies their contribution by their roles in pathway activation (activators or inhibitors). The overall median complement score for all samples was 0.28 (IQR: 0.08–0.62), with a range from −0.25 to 0.92. Cluster 1 had a median complement score of 0.69 (IQR: 0.51–0.76), whereas cluster 0 had a median score of 0.14 (IQR: 0.01–0.28) ($P < 0.001$) (Figure 3a). Based on the median score of 0.28 from all samples, we divided the patients into high ($>$ 0.28) and low ($\leq$ 0.28) complement score groups (Supplementary Figure 5).

Patients with high complement scores had a 2.41-fold (1.07–5.40) higher risk of kidney disease progression than those with low scores, after adjustment for demographic and clinical variables (Table 3). Multivariable-adjusted cumulative incidence of DKD progression also demonstrated a significant difference between the high and low complement groups (Figure 3b). Notably, in multivariable-adjusted spline regression, the association between the complement score and the hazard of DKD progression was linear (Figure 3c).

To eliminate the potential contribution of the urine creatinine denominator to the outcome, a sensitivity analysis was conducted using the urine abundance of proteins without normalization for urine creatinine. These analyses yielded the same results, with cluster 1 and the high complement score group both significantly associated with rapid DKD progression, compared with cluster 0 and the low complement score group, respectively (Supplementary Figure S6).

Histopathologic Correlation With Complement Scores in SNUH-DN Cohort

In the SNUH-DN cohort, which included paired urine and kidney tissues from the same individuals, urine complement scores were significantly correlated with

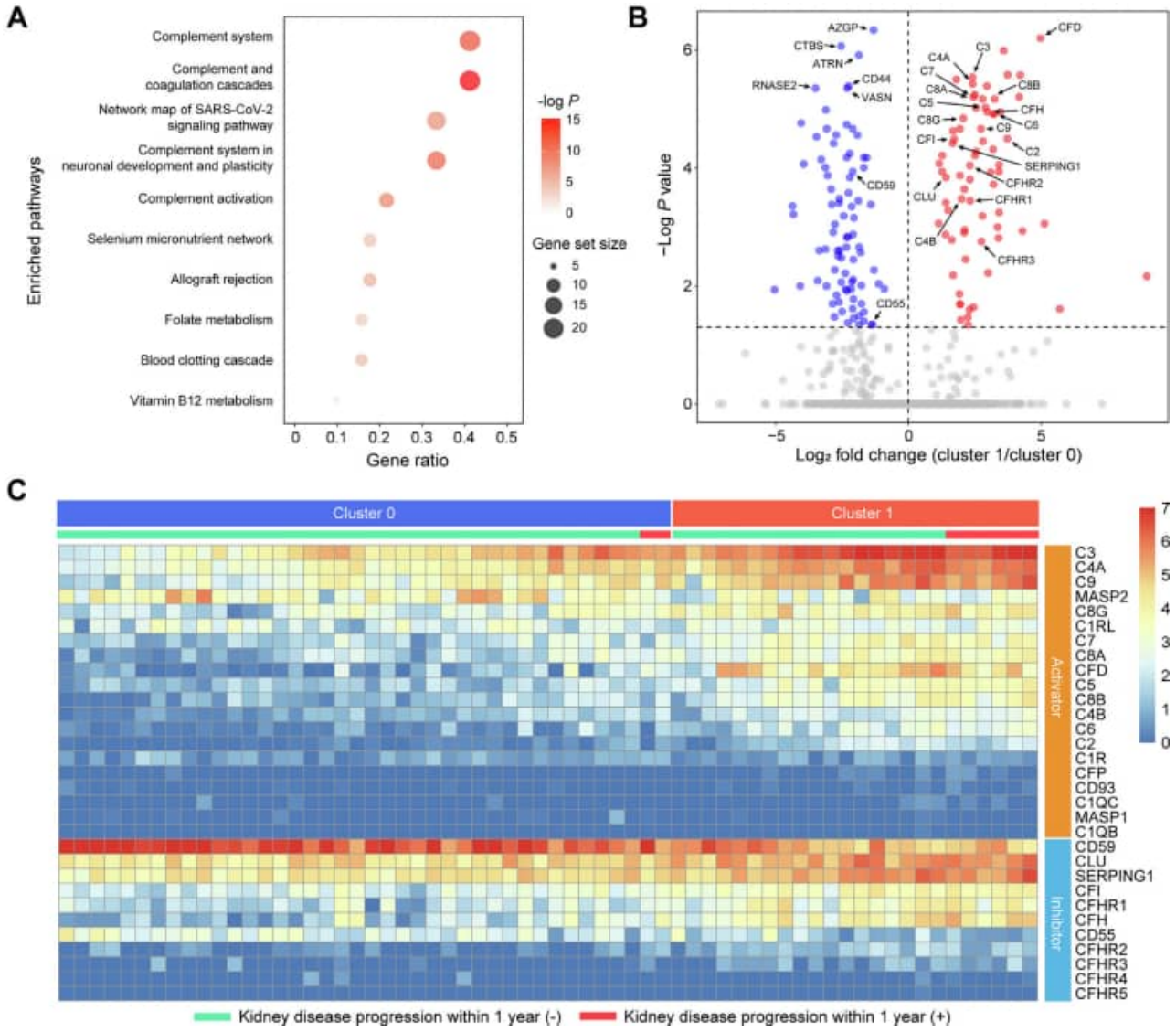

Figure 2. Complement pathway and kidney disease progression in the SNUH-DN cohort. (a) Pathway analysis of the differentially abundant proteins in cluster 1 compared with cluster 0. (b) Volcano plot of the differentially abundant proteins in cluster 1 compared with cluster 0. The dashed line denotes the threshold of significance. Red and blue dots represent proteins increased and decreased in cluster 1, respectively, compared with cluster 0. (c) Heatmap of urinary complement component abundance categorized by clustering, kidney disease progression within 1 year, and complement function.

the following 3 of 5 features used to classify DKD: (i) glomerular lesions, (ii) interstitial fibrosis and tubular atrophy, and (iii) arteriolar hyalinosis. The correlation between complement score and interstitial inflammation did not reach statistical significance. Urine complement score was not correlated with arteriosclerosis in this sample set (Supplementary Figure S7).

Baseline Characteristics of the CRIC-T2D Subcohort

The association between urinary complement proteins and kidney disease progression was examined in an independent cohort: the CRIC-T2D. The CRIC-T2D subcohort (Table 4), comprising 282 patients, was a majority (62%) male group with a mean age of 60.8 ± 7.5 years. Of these patients, 51.8% were non-Hispanic Black and 30.1% were non-Hispanic White; hypertension and cardiovascular disease were identified in 93.6% and 46.5% of the participants, respectively; and the median 24-hour urine protein and eGFR were 0.48 (IQR: 0.10–1.87) g and 42 (IQR: 37–49) ml/min per 1.73 m^2, respectively. The median follow-up duration was 9.7 years (IQR: 6.3–12.0).

Complement Score in the CRIC-T2D Subcohort

In the CRIC-T2D subcohort, targeted urinary proteomics was used to quantify 12 complement proteins, including complement activators such as factor B; C3; component 4; C5; component 6; component 7; component 8 alpha chain; component 9; as well as inhibitors

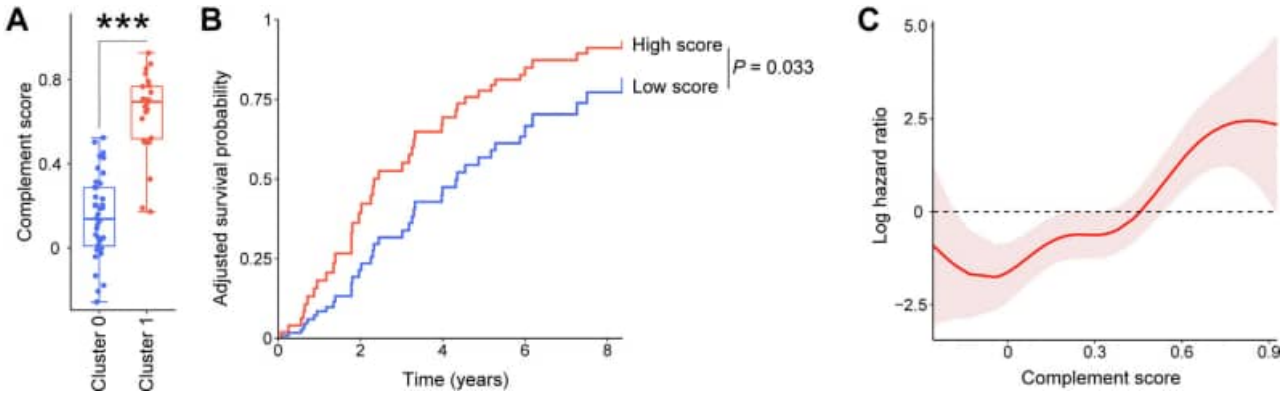

Figure 3. Complement score and kidney disease progression in the SNUH-DN cohort. (a) Bar plot of complement score in clusters 0 and 1. ***$P < 0.001$. (b) Multivariable-adjusted cumulative incidence of kidney disease progression according to complement score groups after adjustment for age, sex, body mass index, hypertension, diabetes duration, urine protein-to-creatinine ratio, and eGFR. (c) Penalized spline regression according to complement score after adjusting for age, sex, body mass index, hypertension, diabetes duration, urine protein-to-creatinine ratio, and eGFR. eGFR, estimated glomerular filtration rate.

such as CD59, complement factor I, clusterin, and complement factor H–related protein 2 (Supplementary Table S2 and Supplementary Figure S8). The complement score was calculated for each participant as described in the complement scores section. The median complement score for all participants was 1.07 (IQR: 0.42–2.10), ranging from −0.76 to 3.04. Based on this median value, the participants were divided into groups with high (> 1.07) and low (≤ 1.07) complement scores (Figure 4a). The median complement scores for the high and low groups were 2.16 (IQR: 1.63–2.45) and 0.41 (IQR: 0.17–0.68), respectively. The group with high scores had a 2.51-fold (95% CI: 1.76–3.57) higher risk of kidney disease progression than the low score group, after adjustment for age, sex, body mass index, hypertension, ethnicity and race, 24-hour urine protein, and eGFR (Figure 4b and Table 5). Multivariable-adjusted spline regression revealed a linear association between the complement score and the hazard of DKD progression (Figure 4c).

Table 3. Association between complement score and kidney disease progression in the SNUH-DN cohort

	Univariable		Multivariable[a]	
Variables	HR (95% CI)	P	HR (95% CI)	P
High complement score (vs. low)	2.67 (1.37–5.18)	0.004	2.41 (1.07–5.40)	0.033
Age (/1 yr)	1.00 (0.98–1.03)	0.759	0.95 (0.91–0.99)	0.011
Female (vs. male)	0.47 (0.20–1.08)	0.074	0.30 (0.11–0.79)	0.015
Body mass index (per 1 kg/m^2)	0.93 (0.83–1.04)	0.224	0.88 (0.78–1.00)	0.050
Hypertension (vs. none)	1.14 (0.35–3.74)	0.827	0.89 (0.25–3.11)	0.852
Diabetes duration (/1 yr)	1.02 (0.98–1.07)	0.380	1.03 (0.98–1.08)	0.311
PCR (per 1 g/g)	1.16 (1.08–1.25)	<0.001	1.14 (1.03–1.26)	0.010
eGFR (per 1 ml/min per 1.73 m^2)	0.98 (0.96–0.99)	0.006	0.99 (0.97–1.00)	0.142

CI, confidence interval; eGFR, estimated glomerular filtration rate; HR, hazard ratio; PCR, random urine protein-to-creatinine ratio.
[a]Adjusted for age, sex, body mass index, hypertension, diabetic duration, proteinuria, and eGFR.

DISCUSSION

We used untargeted and targeted mass spectrometry to investigate the association between complement proteins and DKD progression in the SNUH-DN and CRIC-T2D cohorts. Unsupervised clustering of proteins identified in untargeted urinary proteomics from the SNUH-DN cohort revealed 2 groups with different kidney outcomes. Pathway analysis revealed a marked prominence of complement proteins in the group with rapidly progressive DKD. Complement scores, constructed based on the abundance and function of complement proteins, were associated with histopathologic severity and pace of DKD progression. These results were replicated in the CRIC-T2D subcohort, where high complement scores similarly correlated with rapid progression. Together, these findings underscore the potential of urinary complement proteins

Table 4. Baseline characteristics of the CRIC-T2D subcohort

Variables	Total (N = 282)	No progression (n = 101)	Progression (n = 181)	P
Age (yrs)	60.8 ± 7.5	61.8 ± 7.3	60.3 ± 7.5	0.103
Male (%)	62.4	58.4	64.6	0.365
Body mass index (kg/m^2)	34.7 ± 7.6	34.7 ± 8.2	34.7 ± 7.2	0.997
Race and ethnicity				
Other	18.1	11.9	21.5	0.105
White non-Hispanic	30.1	34.7	27.6	
Black non-Hispanic	51.8	53.5	50.8	
Hypertension (%)	93.6	91.1	95.0	0.297
Cardiovascular disease (%)	46.5	44.6	47.5	0.724
Hemoglobin A1c (%)	7.1 (6.3–8.1)	7.0 (6.2–7.8)	7.1 (6.4–8.2)	0.128
24-hr urine protein (g)	0.48 (0.10–1.87)	0.10 (0.05–0.38)	1.03 (0.30–3.30)	<0.001
eGFR (ml/min per 1.73 m^2)	42 (37–49)	43.7 (38–48)	41 (35–49)	0.031
Follow-up time (yrs)	9.7 (6.3–12.0)	9.7 (5.2–12.0)	9.7 (6.5–12.0)	0.668

eGFR, estimated glomerular filtration rate.

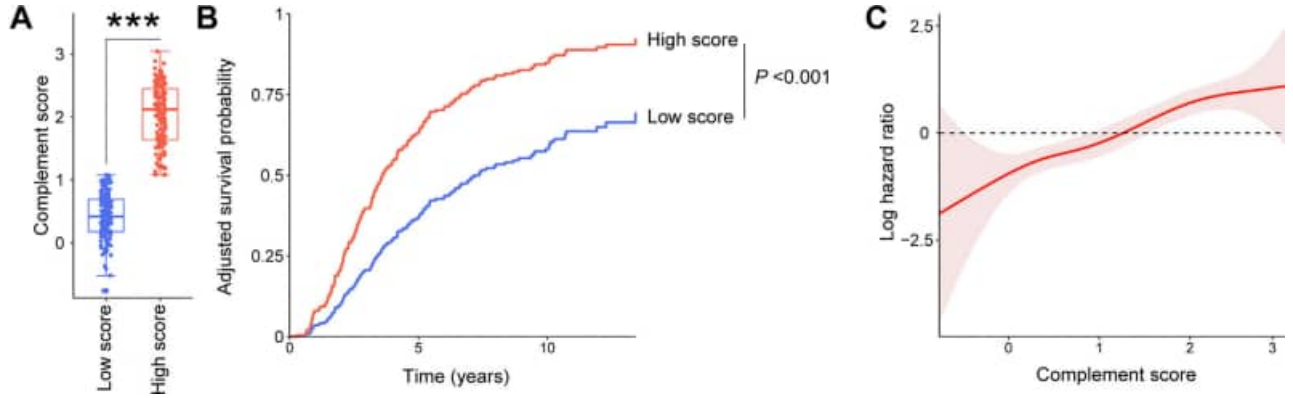

Figure 4. Complement score and kidney disease progression in the CRIC-T2D subcohort. (a) Bar plot of complement score in high and low score groups. ***$P < 0.001$. (b) Multivariable-adjusted cumulative incidence of kidney disease progression according to complement score groups after adjustment for age, sex, body mass index, race, hypertension, 24-hour urine protein, and eGFR. (c) Penalized spline regression according to complement score after adjusting for age, sex, body mass index, race, hypertension, 24-hour urine protein, and eGFR. eGFR, estimated glomerular filtration rate.

as predictive biomarkers for kidney disease progression and the pathway as a target for therapeutic intervention in DKD.

Although agents, including renin-angiotensin-system inhibitors, sodium-glucose cotransporter 2 inhibitors, and mineralocorticoid receptor antagonists, have improved kidney outcomes in patients with DKD, a significant number still experience rapid disease progression.[44,45] The underlying pathophysiology contributing to this residual risk may involve the complement pathway, characterized by an imbalance between activators and inhibitors.[46,47] High urine levels of complement proteins in DKD, comparable to those in autoimmune glomerulonephritis, have been associated with worse kidney outcomes.[48,49] Even in patients with biopsy-confirmed diabetic nephropathy, the association of complement proteins with worse outcomes has been observed.[50,51] Our results support and extend these findings and highlight the importance of further exploring the involvement of the complement pathway in DKD progression.

We add strengths to previous studies in several ways. First, we combined the use of untargeted and targeted proteomics. Employing untargeted proteomics enabled an unbiased examination of the urinary proteome, leading to the initial identification of an increase in complement pathway proteins. Targeted proteomics was then used to validate these observations. Second, to our knowledge, this is the first study to replicate the association between urine complement components and DKD progression in 2 independent cohorts. Third, our approach goes beyond examining individual complement proteins and measures a complement score, which accounts for both activator and inhibitor functions, similar to methodologies in previous studies that used expression levels of molecules involved in specific pathways for scoring.[36,37] This approach allowed both a qualitative and quantitative analysis by categorizing patients into high and low complement score groups and using spline regression to analyze the shape of the association between the complement score and the hazard of disease progression.

We identified several complement activators, such as complement factor D, C3, component 4, C5, component 6, component 8, and mannan-binding lectin serine protease 1 and 2, that were elevated in patients with rapid kidney disease progression. Upregulation of several complement proteins (e.g., C3, C 5 alpha chain, mannan-binding lectin serine protease 1 and 2) has been noted in rats receiving a high-fat diet combined with low-dose streptozotocin.[52] Previous human data on DKD identified component 4 and component 8 as significant risk factors for ESKD and all-cause mortality.[49] Inhibiting these activators may serve as a therapeutic intervention, because a previous study showed

Table 5. Association between complement score and kidney disease progression in the CRIC-T2D subcohort

	Univariable		Multivariable[a]	
Variables	HR (95% CI)	*P*	HR (95% CI)	*P*
High score (vs. low)	3.76 (2.74–5.16)	<0.001	2.51 (1.76–3.57)	<0.001
Age (/1 yr)	0.98 (0.96–1.00)	0.120	1.00 (0.97–1.02)	0.720
Female (vs. male)	1.02 (0.75–1.38)	0.903	1.21 (0.87–1.68)	0.263
Body mass index (per 1 kg/m^2)	1.00 (0.98–1.02)	0.682	1.00 (0.97–1.02)	0.822
Race and ethnicity (vs. other)				
White non-Hispanic	0.55 (0.36–0.84)	0.005	0.79 (0.51–1.23)	0.296
Black non-Hispanic	0.71 (0.48–1.03)	0.070	0.96 (0.65–1.42)	0.839
Hypertension (vs. none)	1.60 (0.82–3.13)	0.171	1.71 (0.87–3.37)	0.121
24-hour urine protein (per 1 g)	1.31 (1.25–1.37)	<0.001	1.23 (1.17–1.30)	<0.001
eGFR (per 1 ml/min per 1.73 m^2)	0.97 (0.96–0.99)	0.003	0.99 (0.97–1.01)	0.263

CI, confidence interval; eGFR, estimated glomerular filtration rate; HR, hazard ratio.
[a]Adjusted for age, sex, body mass index, race and ethnicity, hypertension, 24-hr proteinuria, and eGFR.

that an inhibitor targeting the C 5 alpha chain receptor attenuated kidney fibrosis in streptozotocin-treated mice.[53] The monoclonal antibody targeting C5, eculizumab, is an established clinical intervention in atypical hemolytic uremic syndrome and also tried in glomerulonephritis such as IgA nephropathy and antineutrophil cytoplasmic antibody-mediated vasculitis.[54,55] As a C3 inhibitor, pegcetacoplan, originally effective for paroxysmal nocturnal hemoglobinuria, is currently undergoing clinical trials as a treatment for C3 glomerulopathy and immune complex–mediated membranoproliferative glomerulonephritis.[56,57] The time is approaching to examine complement inhibition as a novel treatment in DKD.

We observed an increase in some complement inhibitors (e.g., complement factor H, clusterin, and complement factor I) and a decrease in others (e.g., CD55 and CD59) in rapid progressors. These findings are consistent with previous studies, which have noted similar patterns of increase in other inhibitors alongside decreased CD55 during DKD progression.[51] The increased complement inhibitors may suggest a compensatory response to an activated complement pathway, perhaps displaying an inadequate regulatory mechanism that is insufficient to counterbalance the activation of the complement pathway.[49] CD55 depletion exacerbates DKD in an animal model.[58] Low urine CD59 was associated with an increased risk of ESKD, cardiovascular complications, and all-cause mortality in patients with DKD,[49,59] and overexpression of CD55 and CD59 attenuated kidney damage in ischemia-reperfusion injury mouse models.[60] Collectively, CD55 and CD59 may exhibit a different pattern and thus play different roles in complement regulation compared with other complement inhibitors, which showed a compensatory increase in DKD. CD59 is an interesting case, where glycation of the native protein blocks its ability to inhibit assembly of the membrane attack complex, and thus complement-mediated tissue damage.[59] Urine CD59 abundance is inversely associated with DKD progression in this study and a previous report.[49] Interestingly, only the nonglycated (i.e., functional) form of CD59 was identified in urine samples from the SNUH-DN cohort, the CRIC-T2D subcohort, and a previous study.[49] The observed inverse association between nonglycated urine CD59 and DKD progression aligns with the notion that complement activation promotes DKD progression. It remains unclear, however, why glycated CD59 was not identified in the urine samples. Several hypotheses can be formulated as follows: the identification of glycated CD59 might depend on the detection of a single peptide, as opposed to multiple peptides from the nonglycated form; or glycated CD59 could be more frequently internalized or be less likely to be shed from kidney tissue into the urine. Determining the exact etiology will require further investigation, such as quantifying tissue dynamics of CD59 expression and serum CD59 levels.

Although the study provides valuable insights, there are certain limitations that warrant further examination. As with other human studies, except randomized controlled trials, it is challenging to confirm causality and elucidate the underlying mechanisms solely based on our findings. We did not examine the trend of changes in complement proteins over time, which limits our understanding of their dynamics. In addition, we did not identify the source of the complement proteins. While the liver is considered a major source of complement proteins, the kidney may also contribute significantly, especially in the context of injury and inflammation.[36,61] Furthermore, the present cohorts primarily consisted of patients with reduced eGFR or advanced stage of DKD, thereby hindering the application of the results to those with earlier stages of DKD.

We find a significant association between the complement pathway and rapid DKD progression in 2 independent populations. Our findings highlight urine complement proteins as potential biomarkers of pathway status and the complement pathway itself as a potential therapeutic target. Comparable findings in 2 ethnically dissimilar cohorts suggest that current observations are likely generalizable. By integrating various complement proteins and validating the results in an independent cohort, we provide robust evidence for the involvement of the complement pathway in rapidly progressive DKD. These findings lay the foundation for its clinical application and advocate for continued research into complement-targeting interventions to improve outcomes in the complex landscape of DKD.

DISCLOSURE

All the authors declared no competing interests.

ACKNOWLEDGMENTS

The work was supported by the grant from the SNUH Research Fund (26-2018-0040 to SSH; 26-2022-0040 to SSH), the Richard A. and Nora Eccles Foundation award (A200111 to MA), Richard A. and Nora Eccles Harrison Endowed Chair in Diabetes Research and the Thomas A. Depner Endowed Chair in Nephrology Research. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the funding institutions. All untargeted proteomics procedures were conducted using the core resources of the Center for Medical Innovation (CMI) of Seoul National University Hospital.

DATA AVAILABILITY STATEMENT

The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request. The untargeted proteomic data have been deposited in the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository (accession no. PXD037505).

AUTHOR CONTRIBUTIONS

SSH and MA designed the study. DY and SB analyzed the collected data. YG, CDN, TLF, and WJQ performed targeted proteomics. LL analyzed the targeted proteomic data. DH performed untargeted proteomics. KCM reviewed kidney biopsy section and classified histopathologic parameters. SSH, DKK, and YSK assisted with the development of the SNUH-DN cohort and sample collections. MA was the recipient of an NIH award with access to CRIC-T2D samples. MA and WW supervised and funded the targeted proteomics study in CRIC-T2D. AZR and PS assisted with the interpretation of untargeted proteomics and assisted with the evaluation of histopathologic parameters. DY, SSH, and MA wrote the manuscript. SSH and MA critically reviewed the manuscript. All the authors approved the final manuscript.

SUPPLEMENTARY MATERIAL

Supplementary File (PDF)

Figure S1. Heatmap of intensity-based absolute quantifications for proteins in the SNUH-DN cohort with Bonferroni adjustment.

Figure S2. Heatmap of urinary complement component abundance categorized by clustering, kidney disease progression within 3 years, and complement function.

Figure S3. Sankey plot illustrating the relationships among glomerular classes, clusters, and clinical outcomes within 1 year.

Figure S4. Pathway analysis of the differentially abundant proteins in kidney disease progressors and no progressors.

Figure S5. Complement scores in the SNUH-DN cohort. Patients were binarized into high and low groups according to the median complement score.

Figure S6. Rates of kidney disease progression using proteins not normalized by urine creatinine in the SNUH-DN cohort, after adjusting for age, sex, body mass index, hypertension, diabetic duration, proteinuria, and estimated glomerular filtration rate.

Figure S7. Correlation between complement scores and histopathologic parameters in the SNUH-DN cohort.

Figure S8. Heatmap of urinary complement component abundance categorized by complement score groups in the CRIC-T2D subcohort.

Table S1. Peptide sequences of targeted proteomics used in LC-MS.

Table S2. List on complement proteins. Intensity of the components was used to calculate the cumulative complement score.

Table S3. Proteins upregulated in cluster 1 mapped to 10 pathways.

STROBE Checklist.

REFERENCES

1. Afkarian M, Sachs MC, Kestenbaum B, et al. Kidney disease and increased mortality risk in type 2 diabetes. *J Am Soc Nephrol*. 2013;24:302–308. https://doi.org/10.1681/ASN.2012070718
2. Hahr AJ, Molitch ME. Management of diabetes mellitus in patients with CKD: core curriculum 2022. *Am J Kidney Dis*. 2022;79:728–736. https://doi.org/10.1053/j.ajkd.2021.05.023
3. Tuttle KR, Bakris GL, Bilous RW, et al. Diabetic kidney disease: a report from an ADA Consensus Conference. *Diabetes Care*. 2014;37:2864–2883. https://doi.org/10.2337/dc14-1296
4. Furuichi K, Shimizu M, Yamanouchi M, et al. Clinicopathological features of fast eGFR decliners among patients with diabetic nephropathy. *BMJ Open Diabetes Res Care*. 2020;8:e001157. https://doi.org/10.1136/bmjdrc-2019-001157
5. Oshima M, Shimizu M, Yamanouchi M, et al. Trajectories of kidney function in diabetes: a clinicopathological update. *Nat Rev Nephrol*. 2021;17:740–750. https://doi.org/10.1038/s41581-021-00462-y
6. MacIsaac RJ, Ekinci EI, Jerums G. Markers of and risk factors for the development and progression of diabetic kidney disease. *Am J Kidney Dis*. 2014;63(2 suppl 2):S39–S62. https://doi.org/10.1053/j.ajkd.2013.10.048
7. Siddiqui K, George TP, Joy SS, Alfadda AA. Risk factors of chronic kidney disease among type 2 diabetic patients with longer duration of diabetes. *Front Endocrinol*. 2022;13:1079725. https://doi.org/10.3389/fendo.2022.1079725
8. Anderson AH, Xie D, Wang X, et al. Novel risk factors for progression of diabetic and nondiabetic CKD: findings from the chronic renal insufficiency cohort (CRIC) study. *Am J Kidney Dis*. 2021;77:56–73.e1. https://doi.org/10.1053/j.ajkd.2020.07.011
9. Colombo M, Looker HC, Farran B, et al. Serum kidney injury molecule 1 and β 2-microglobulin perform as well as larger biomarker panels for prediction of rapid decline in renal function in type 2 diabetes. *Diabetologia*. 2019;62:156–168. https://doi.org/10.1007/s00125-018-4741-9
10. Looker HC, Colombo M, Hess S, et al. Biomarkers of rapid chronic kidney disease progression in type 2 diabetes. *Kidney Int*. 2015;88:888–896. https://doi.org/10.1038/ki.2015.199
11. Rossing P, Caramori ML, Chan JC, et al. KDIGO 2022 clinical practice guideline for diabetes management in chronic kidney disease. *Kidney Int*. 2022;102:S1–S127. https://doi.org/10.1016/j.kint.2022.06.008
12. Barrera-Chimal J, Lima-Posada I, Bakris GL, Jaisser F. Mineralocorticoid receptor antagonists in diabetic kidney disease—mechanistic and therapeutic effects. *Nat Rev Nephrol*. 2022;18:56–70. https://doi.org/10.1038/s41581-021-00490-8

13. Tong L-L, Adler SG. Diabetic kidney disease treatment: new perspectives. *Kidney Res Clin Pract*. 2022;41(suppl 2):S63–S73. https://doi.org/10.23876/j.krcp.21.288

14. Tofte N, Lindhardt M, Adamova K, et al. Early detection of diabetic kidney disease by urinary proteomics and subsequent intervention with spironolactone to delay progression (PRIORITY): a prospective observational study and embedded randomised placebo-controlled trial. *Lancet Diabetes Endocrinol*. 2020;8:301–312. https://doi.org/10.1016/S2213-8587(20)30026-7

15. Zaghlool SB, Halama A, Stephan N, et al. Metabolic and proteomic signatures of type 2 diabetes subtypes in an Arab population. *Nat Commun*. 2022;13:7121. https://doi.org/10.1038/s41467-022-34754-z

16. Gupta J, Mitra N, Kanetsky PA, et al. Association between albuminuria, kidney function, and inflammatory biomarker profile in CKD in CRIC. *Clin J Am Soc Nephrol CJASN*. 2012;7:1938–1946. https://doi.org/10.2215/CJN.03500412

17. Mehta NN, Matthews GJ, Krishnamoorthy P, et al. Higher plasma CXCL12 levels predict incident myocardial infarction and death in chronic kidney disease: findings from the chronic renal insufficiency Cohort study. *Eur Heart J*. 2013;35:2115–2122. https://doi.org/10.1093/eurheartj/eht481

18. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med*. 2009;150:604–612. https://doi.org/10.7326/0003-4819-150-9-200905050-00006

19. Tervaert TWC, Mooyaart AL, Amann K, et al. Pathologic classification of diabetic nephropathy. *J Am Soc Nephrol*. 2010;21:556–563. https://doi.org/10.1681/ASN.2010010010

20. Grzegorski SJ, Hu Z, Liu Y, et al. Disruption of the kringle 1 domain of prothrombin leads to late onset mortality in zebrafish. *Sci Rep*. 2020;10:4049. https://doi.org/10.1038/s41598-020-60840-7

21. Han D, Jin J, Woo J, Min H, Kim Y. Proteomic analysis of mouse astrocytes and their secretome by a combination of FASP and StageTip-based, high pH, reversed-phase fractionation. *Proteomics*. 2014;14:1604–1609. https://doi.org/10.1002/pmic.201300495

22. Rhee SJ, Han D, Lee Y, et al. Comparison of serum protein profiles between major depressive disorder and bipolar disorder. *BMC Psychiatry*. 2020;20:1–11. https://doi.org/10.1186/s12888-020-02540-0

23. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics*. 2014;13:2513–2526. https://doi.org/10.1074/mcp.M113.031591

24. Han D, Jin J, Yu J, Kim K, Kim Y. Integrated approach using multistep enzyme digestion, TiO2 enrichment, and database search for in-depth phosphoproteomic profiling. *Proteomics*. 2015;15:618–623. https://doi.org/10.1002/pmic.201400102

25. Mertins P, Tang LC, Krug K, et al. Reproducible workflow for multiplexed deep-scale proteome and phosphoproteome analysis of tumor tissues by liquid chromatography–mass spectrometry. *Nat Protoc*. 2018;13:1632–1661. https://doi.org/10.1038/s41596-018-0006-9

26. Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc*. 2016;11:2301–2319. https://doi.org/10.1038/nprot.2016.136

27. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res*. 2011;10:1794–1805. https://doi.org/10.1021/pr101065j

28. Schwanhäusser B, Busse D, Li N, et al. Global quantification of mammalian gene expression control. *Nature*. 2011;473:337–342. https://doi.org/10.1038/nature10098

29. Gao Y, Wang H, Nicora CD, et al. LC-SRM-Based targeted quantification of urinary protein biomarkers. *Methods Mol Biol*. 2018;1788:145–156. https://doi.org/10.1007/7651_2017_93

30. MacLean B, Tomazela DM, Shulman N, et al. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics*. 2010;26:966–968. https://doi.org/10.1093/bioinformatics/btq054

31. Hao Y, Stuart T, Kowalski MH, et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol*. 2024;42:293–304. https://doi.org/10.1038/s41587-023-01767-y

32. Consortium U. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res*. 2019;47:D506–D515. https://doi.org/10.1093/nar/gky1049

33. Martens M, Ammar A, Riutta A, et al. WikiPathways: connecting communities. *Nucleic Acids Res*. 2021;49:D613–D621. https://doi.org/10.1093/nar/gkaa1024

34. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res*. 2023;51:D587–D592. https://doi.org/10.1093/nar/gkac963

35. Broquet A, Gourain V, Goronflot T, et al. Sepsis-trained macrophages promote antitumoral tissue-resident T cells. *Nat Immunol*. 2024;25:802–819. https://doi.org/10.1038/s41590-024-01819-8

36. Kim M-G, Yun D, Kang CL, et al. Kidney VISTA prevents IFN-γ/IL-9 axis–mediated tubulointerstitial fibrosis after acute glomerular injury. *J Clin Invest*. 2022;132:e151189. https://doi.org/10.1172/JCI151189

37. Kuppe C, Ibrahim MM, Kranz J, et al. Decoding myofibroblast origins in human kidney fibrosis. *Nature*. 2021;589:281–286. https://doi.org/10.1038/s41586-020-2941-1

38. Denz R, Klaaßen-Mielke R, Timmesfeld N. A comparison of different methods to adjust survival curves for confounders. *Stat Med*. 2023;42:1461–1479. https://doi.org/10.1002/sim.9681

39. Barratt J, Weitz I. Complement factor D as a strategic target for regulating the alternative complement pathway. *Front Immunol*. 2021;12:712572. https://doi.org/10.3389/fimmu.2021.712572

40. Haslund D, Ryø LB, Majidi SS, et al. Dominant-negative SERPING1 variants cause intracellular retention of C1 inhibitor in hereditary angioedema. *J Clin Invest*. 2019;129:388–405. https://doi.org/10.1172/JCI98869

41. Meri S, Haapasalo K. Function and dysfunction of complement factor H during formation of lipid-rich deposits. *Front Immunol*. 2020;11:611830. https://doi.org/10.3389/fimmu.2020.611830

42. De Boer EC, Van Mourik AG, Jongerius I. Therapeutic lessons to be learned from the role of complement regulators as

double-edged sword in health and disease. *Front Immunol*. 2020;11:578069. https://doi.org/10.3389/fimmu.2020.578069

43. Mohan Rao LV, Esmon CT, Pendurthi UR. Endothelial cell protein C receptor: a multiliganded and multifunctional receptor. *Blood*. 2014;124:1553–1562. https://doi.org/10.1182/blood-2014-05-578328
44. Naaman SC, Bakris GL. Slowing diabetic kidney disease progression: where do we stand today? *ADA Clinical Compendia*. 2021;2021:28–32. https://doi.org/10.2337/db20211-28
45. Zoccali C, Floege J. Moderator's view: after SGLT2i and MRA antagonists, where do we go? *Clin Kidney J*. 2024;17:sfae013. https://doi.org/10.1093/ckj/sfae013
46. Flyvbjerg A. The role of the complement system in diabetic nephropathy. *Nat Rev Nephrol*. 2017;13:311–318. https://doi.org/10.1038/nrneph.2017.31
47. Tan SM, Snelson M, Østergaard JA, Coughlan MT. The complement pathway: new insights into immunometabolic signaling in diabetic kidney disease. *Antioxid Redox Signal*. 2022;37:781–801. https://doi.org/10.1089/ars.2021.0125
48. Pelletier K, Bonnefoy A, Chapdelaine H, et al. Clinical value of complement activation biomarkers in overt diabetic nephropathy. *Kidney Int Rep*. 2019;4:797–805. https://doi.org/10.1016/j.ekir.2019.03.004
49. Vaisar T, Durbin-Johnson B, Whitlock K, et al. Urine complement proteins and the risk of kidney disease progression and mortality in type 2 diabetes. *Diabetes Care*. 2018;41:2361–2369. https://doi.org/10.2337/dc18-0699
50. Li X-Q, Chang D-Y, Chen M, Zhao M-H. Complement activation in patients with diabetic nephropathy. *Diabetes Metab*. 2019;45:248–253. https://doi.org/10.1016/j.diabet.2018.04.001
51. Zhao L, Zhang Y, Liu F, et al. Urinary complement proteins and risk of end-stage renal disease: quantitative urinary proteomics in patients with type 2 diabetes and biopsy-proven diabetic nephropathy. *J Endocrinol Invest*. 2021;44:2709–2723. https://doi.org/10.1007/s40618-021-01596-3
52. Huang Y, Xu J, Wu X, et al. High expression of complement components in the kidneys of type 2 diabetic rats with diabetic nephropathy. *Front Endocrinol*. 2019;10:459. https://doi.org/10.3389/fendo.2019.00459
53. Tan SM, Ziemann M, Thallas-Bonke V, et al. Complement C5a induces renal injury in diabetic kidney disease by disrupting mitochondrial metabolic agility. *Diabetes*. 2020;69:83–98. https://doi.org/10.2337/db19-0043
54. Rizk DV, Maillard N, Julian BA, et al. The emerging role of complement proteins as a target for therapy of IgA nephropathy. *Front Immunol*. 2019;10:442135. https://doi.org/10.3389/fimmu.2019.00504
55. Zipfel PF, Wiech T, Rudnick R, Afonso S, Person F, Skerka C. Complement inhibitors in clinical trials for glomerular diseases. *Front Immunol*. 2019;10:464149. https://doi.org/10.3389/fimmu.2019.02166
56. Hillmen P, Szer J, Weitz I, et al. Pegcetacoplan versus eculizumab in paroxysmal nocturnal hemoglobinuria. *N Engl J Med*. 2021;384:1028–1037. https://doi.org/10.1056/NEJMoa2029073
57. Wooden B, Tarragon B, Navarro-Torres M, Bomback AS. Complement inhibitors for kidney disease. *Nephrol Dial Transplant*. 2023;38(suppl 2):ii29–ii39. https://doi.org/10.1093/ndt/gfad079
58. Angeletti A, Cantarelli C, Petrosyan A, et al. Loss of decay-accelerating factor triggers podocyte injury and glomerulosclerosis. *J Exp Med*. 2020;217:e20191699. https://doi.org/10.1084/jem.20191699
59. Ghosh P, Sahoo R, Vaidya A, Chorev M, Halperin JA. Role of complement and complement regulatory proteins in the complications of diabetes. *Endocr Rev*. 2015;36:272–288. https://doi.org/10.1210/er.2014-1099
60. Bongoni AK, Lu B, Salvaris EJ, et al. Overexpression of human CD55 and CD59 or treatment with human CD55 protects against renal ischemia-reperfusion injury in mice. *J Immunol*. 2017;198:4837–4845. https://doi.org/10.4049/jimmunol.1601943
61. Kragesteen BK, Giladi A, David E, et al. The transcriptional and regulatory identity of erythropoietin producing cells. *Nat Med*. 2023;29:1191–1200. https://doi.org/10.1038/s41591-023-02314-7