

Title: Kidney transcriptomics signature of prospective rapid diabetic kidney disease progression

Authors: Dianne Acoba MSc^{1,2}, Anna Levin MD³, Anna Witasz PhD³, Kerstin Ebefors PhD⁴, Johan Mölne MD⁵, Peter J. Greasley PhD⁶, Jenny Nyström PhD⁴, Annika Wernerson MD PhD³ and Anna Reznichenko MD PhD¹

Affiliations:

¹Translational Science & Experimental Medicine, Research and Early Development, Cardiovascular, Renal and Metabolism, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden

²Institut Necker Enfants Malades (INEM), Institut National de la Santé et de la Recherche Médicale (INSERM) U1151, Université Paris Cité, Paris, France

³Department of Clinical Science, Intervention and Technology, Division of Renal Medicine, Karolinska Institutet, Stockholm, Sweden

⁴Department of Physiology, Institute of Neuroscience and Physiology, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

⁵Department of Laboratory Medicine, Institute of Biomedicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

⁶Early Clinical Development, Research and Early Development, Cardiovascular, Renal and Metabolism, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden

Correspondence:

Anna Reznichenko

Clinical Renal, Late-stage Development, Cardiovascular, Renal and Metabolism, BioPharmaceuticals R&D, AstraZeneca, Pepparedsleden 1, 431 83 Mölndal, Sweden

+46 72 515 82 70

Anna.Reznichenko@astrazeneca.com

ABSTRACT

Previous cross-sectional transcriptomic studies on diabetic kidney disease (DKD) kidney tissue have shown correlations between gene expression and both disease status and kidney function at time of biopsy, but longitudinal data is scarce. We utilized clinical follow-up data up to five years post-biopsy, linking the transcriptomes of diagnostic kidney biopsies to progression rates and outcomes in 19 DKD patients. Patients were stratified into “rapid progressors” and “non-rapid progressors” based on clinical parameters (eGFR slope, CKD stage advancement, degree of albuminuria, composite of kidney failure or 40% eGFR decline). Differential expression and pathway enrichment analyses were performed to identify dysregulated genes and pathways associated with rapid progression. We identified 265 genes in glomeruli and tubulointerstitium that were significantly modulated in rapid vs non-rapid DKD progression. Rapid progression-associated genes showed enrichment for well-established (extracellular matrix organization, inflammation) as well as novel pathways in the context of DKD (circadian rhythm, cytoskeleton reorganization, *NOTCH* signaling). This study illuminates kidney gene expression patterns that may be predictive of rapid progression in DKD and are distinct from those associated with cross-sectional kidney function.

KEYWORDS

chronic kidney disease, diabetic kidney disease, rapid progression, eGFR slope, gene expression, longitudinal study

INTRODUCTION

Diabetic kidney disease is the leading cause of chronic kidney disease (CKD) and kidney failure,¹ affecting patients with Diabetes Mellitus (DM). The etiology is multifactorial and its clinical presentation is highly heterogeneous with some patients developing DKD at an early stage of DM, while others never do despite multiple risk factors.² This complexity in DKD pathogenesis is also mirrored in its progression pattern where there is a high inter-individual variability in the rate of kidney function decline and levels of albuminuria over time.³ Understanding the pathophysiological drivers and mechanisms of DKD progression is therefore essential to achieve early identification of patients at high risk for rapid kidney function decline as well as to enable identification of novel therapeutic targets to delay or even reverse the disease.

Over the past twenty years, numerous kidney transcriptomics studies have been made, showing significant strides in mapping the human DKD molecular landscape. These studies predominantly compared gene expression profiles of DKD and control kidneys, sometimes employing tissue microdissection, and identified a plethora of disease-implicated pathways. Such were, for example, pathways linked to diminished glomerular tissue repair,⁴ and SRGAP2a as a key gene associated with proteinuria and estimated glomerular filtration rate (eGFR).⁵ The European Renal cDNA Bank (ERCB) study demonstrated association between tubulointerstitial expression of NF- κ B inflammatory pathway genes⁶ and TRAIL⁷ in DKD. In addition, the Wnt signaling pathway was found activated in DKD,⁸ and transcripts involved in inflammation, metabolism, migration, and mitochondrial dysfunction pathways showed correlations with tubulointerstitial damage.⁹ A study by Berthier et al showed that the Janus kinase-STAT pathway is up-regulated in both glomerular and tubulointerstitial compartments,¹⁰ while Woroniecka et al found RhoA, Cdc42, integrin and VEGF signaling to be differentially regulated in the glomeruli and inflammation-related pathways in the tubulointerstitium.¹¹

The introduction of the more unbiased and high-resolution RNA sequencing (RNA-seq) methods has allowed further fine-tuning of the previous findings that were based on microarray technologies. So far, Fan et al used RNA-seq profiling of non-microdissected kidney biopsies from advanced vs early DKD patients and identified retinoic acid pathway to be renoprotective,¹² whereas our group generated RNA-seq data on microdissected DKD kidney biopsies and reported compartment-specific changes, including glomerular upregulation of pathways associated with extracellular matrix reorganization and inflammation, and tubulointerstitial enrichment of apoptotic pathways.¹³

Together, these systematic studies on transcriptome changes in the affected kidney amassed a pool of disease-associated transcripts with hypothesized roles in DKD pathogenesis. However, as all of them used cross-sectional case-control design, it remains unknown whether the associations go beyond baseline, inferring rapid kidney function decline and progression. Thus, longitudinal studies linking baseline kidney biopsy transcriptomics data to prospectively collected clinical data on disease progression would be the next imperative step towards unravelling the molecular drivers of DKD course heterogeneity.

To leverage our previous DKD characterization study,¹³ we collected post-biopsy prospective longitudinal clinical data and stratified DKD patients based on clinically relevant measures and outcomes into rapid- and non-rapid progressors. In this way we were able to uncover rapid DKD progression gene expression signatures in each of the glomerular and tubulointerstitial compartments. Previously reported genes but also unexplored genes in the context of DKD pathophysiology were identified, thus highlighting possible novel DKD progression drivers as well as corroborating prior hypotheses.

METHODS

Study cohort

Patients' baseline characteristics as well as the transcriptomics data generation from kidney biopsies were previously described.¹³ Briefly, 19 adult patients with clinically and histopathologically verified DKD were enrolled at two centers in Sweden between 2010 and 2016. The kidney biopsy tissues, sampled at time of diagnosis, were microdissected into glomerular and tubulointerstitial fractions and profiled with high-throughput RNA sequencing. Written informed consent was obtained from all patients at study enrollment, and the study was approved by Swedish Ethical Review Authorities at both sites (Gothenburg and Stockholm) in accordance with the Declaration of Helsinki 1975, as revised in 2013.

In this study, the patients were followed clinically up to five years post-biopsy and kidney function and events were extracted from patients' files. eGFR was calculated using the CKD-EPI formula but without modifying for race.¹⁴ All patients were on standard of care treatment.

Disease progression definition and cohort stratification

For patients with at least two time-points, "rapid progressor" status was defined according to either one of the following criteria:

- 1) Negative **eGFR slope** greater than -5 ml/min/1.73m²/year;
- 2) Advancement of the **CKD Stage** category compared to baseline¹⁹;
- 3) ≥30% increase in Urine Albumin-Creatinine Ratio (**UACR**) compared to baseline levels²⁰;
- 4) Nephrotic-range albuminuria (**NRA**, defined as UACR > 220mg/mmol) at the last recorded time-point¹⁹;
- 5) **Composite outcome** of kidney failure (KF, defined as the initiation of kidney replacement therapy) or ≥40% eGFR decline.²¹⁻²⁴

Kidney transcriptomics analysis in association with DKD progression

Differential expression analysis (DEA) comparing the kidney transcriptomes of rapid progressors and non-rapid progressors was performed using DESeq2 v1.34.0.²⁵ Research center factor and relevant covariates that are known to affect disease progression (age, sex, BMI, eGFR and UACR at baseline)²⁶ were included in the model design. The p-values were adjusted for multiple testing using the Benjamini-Hochberg method. Genes with an adjusted $P < 0.05$ were considered differentially expressed.

Gene functional annotation was performed using IPA: Ingenuity Pathway Analysis (QIAGEN Inc., <https://www.qiagen.com/us/products/discovery-and-translational-research/next-generation-sequencing/informatics-and-data/interpretation-content-databases/ingenuity-pathway-analysis>) to determine protein class, subcellular location, upstream regulators, and biomarker properties.²⁷ DAVID 2021 was used for Gene Ontology annotation²⁸⁻³¹ and GTEx v8 for kidney cortex enrichment via z-score calculation.³² Weighted Gene Co-expression Network Analysis (WGCNA)³³ was performed and then visualized using Cytoscape.³⁴ Case-control modulation of the progression genes in previous DKD transcriptomics studies (GSE30122, GSE142025, GSE104948, and GSE104954) were also obtained through DEA.^{11-12,35-36}

Over-representation analysis was performed using the Reactome database version 82 and Reactome Pathway Browser v3.7,³⁷ as well as DAVID 2021. Gene set enrichment analysis was performed with the GSEAPreranked function of GSEA v4.3.2 using Gene Ontology gene sets v2022.1.^{30-31,38}

Statistical analyses

All statistical analyses were conducted using R version 4.1.3.¹⁵ Clinical parameters were checked for normality of distributions, intercorrelations and lack of center bias. eGFR slopes were estimated using a mixed model repeated measures analysis, with eGFR as the

dependent variable, time/visit number and baseline eGFR as linear covariates, and patient as a random effect.^{16,17} Transcriptomics data were filtered and only robustly expressed genes (TPM ≥ 1 ¹⁸ in all patients) were included in the downstream analyses: 10 074 genes in glomeruli and 11 964 genes in tubulointerstitium.

RESULTS

Characterization of progressive kidney function loss in the study population

Clinical follow-up data were available for 18 out of 19 patients. Individual and categorized eGFR and UACR trajectories are presented in **Figures 1-5**. During the median follow-up of three years (max five years), 13 patients reached the composite endpoint defined as KF or at least 40% eGFR decline. A total of 17 patients had at least two eGFR time-points and the calculated eGFR slopes ranged from -14 to 3.6 ml/min/year. Ten patients had eGFR declining more than -5 ml/min/year, while 11 patients progressed to a more advanced CKD stage. Out of the 16 patients with at least two measurements of UACR, six patients had at least a 30% increase in UACR and nine patients had NRA during their last clinical follow-up.

Kidney gene expression signature of rapid DKD progression

DEA contrasting rapid progressors versus non-rapid progressors (according to each of the five DKD progression definitions) revealed statistically significant differences in both glomerular and tubulointerstitial transcriptomes, adjusting for known clinical predictors (**Figures 1-5, Table 1**). The full DEA output is shared in **Supplementary Data File S1**. Rapid progression criteria based on the steep eGFR slope and CKD stage advancement yielded the highest numbers of differentially expressed genes (**Table 1**). In total, we identified 265 rapid progression-associated genes—141 genes in the glomeruli and 134 genes in the tubulointerstitium (**Supplementary Data File S2**), with a modest overlap of 10 genes (*MDK, TNFAIP2, LINC00152, C2orf40, ARRB2, MIR4435-1HG, MMP7,*

AL353644.10, COL8A1, MTND2P28) shared between the glomerular and tubulointerstitial fractions.

In silico characterization of the rapid DKD progression signature

To further characterize the rapid progression-associated genes, they were annotated with functional properties and investigated in terms of enriched biological pathways, co-expression, transcriptional regulators, and associations in previous case-control studies (**Supplementary Data File S2**).

The rapid progression-associated genes were diverse in terms of their protein products' classes, which included cytokines, enzymes, G-protein coupled receptors, growth factors, ion channels, kinases, peptidases, phosphatases, transcriptional and translational regulators, transmembrane receptors, and transporters. They displayed predominantly cytoplasmic subcellular localization (103), followed by plasma membrane (51), extracellular space (44), and nucleus (29). Several genes in the rapid progression signature showed strong preferential expression in the healthy kidney cortex as compared to other tissues, such as *AQP6*, *CDH16*, *DHDH*, *LHFPL3-AS2*, *LINC01055*, *LINC01671*, *REN*, *SLC12A3*, and *SPP1*. Based on the IPA BiomarkerFilter analysis, protein products of 124 genes were reported to be measurable in blood, plasma or urine, of which 80 were in the context of urological and nephrological disease, thus supporting their potential biomarker properties.

Upon the compartment-stratified over-representation analysis, the rapid progression-associated genes showed enrichment for specific biological processes (**Figure 6A-D**). Top enriched pathways include integrin cell surface interactions, extracellular matrix assembly and degradation, and interleukin signaling (glomerular subsignature), and pathways for nonsense-mediated decay, ribosomal formation, translation processes, and cellular response to starvation (tubulointerstitial subsignature). Gene set enrichment analysis (GSEA) revealed additional pathways represented in the rapid progression signature such

as mitochondrial dysfunction, oxidative stress, ubiquitination and circadian rhythm regulation.

The complete pathway enrichment analyses results are provided in **Supplementary Data Files S3 and S4**.

The upstream regulator analysis revealed that some of the rapid- progression-associated genes share common transcriptional factors, including *MAFB*, *STAT3*, *SMARCA4*, *HIF1A*, *TP63*, *JUN*, *YAP1*, and *GLI1* in the glomeruli and *SREBF1*, *GLI1*, *TP73*, *YAP1*, *MYC*, *SMARCA4*, and *SP1* in the tubulointerstitium. The comprehensive list of upstream regulators are shared in **Supplementary Data File S5**.

In line with shared transcriptional regulation, the co-expression analysis revealed subnetworks among the rapid progression genes based on gene correlation and module membership (**Figure 6E-F**). Screening for highly connected genes in each module showed that *C16orf58*, *COL1A1*, *CXCL12*, *EMILIN1*, *GLIPR2*, *ITGB2*, *NNMT* and *TNFRSF1B* are hub genes in the glomeruli and *RHOB* in the tubulointerstitium.

Looking over four publicly available DKD data sets, 39 of the 265 rapid progression-associated genes have not been previously reported as differentially expressed between diseased vs healthy patients. The majority of the progression genes (160) were significantly modulated in the same direction in rapid vs non-rapid progressors as in DKD cases vs controls. A small percentage of the progression transcripts either have conflicting modulation directionality (47) or have not been sequenced in previous cross-sectional DKD transcriptomics studies (19). A summary is presented in **Supplementary Data File S6**.

DISCUSSION

To the best of our knowledge, this is the first longitudinal study of DKD kidney biopsy transcriptomics with prospectively collected clinical outcomes data. Such a design has the potential to advance knowledge beyond the amassed cross-sectional evidence and gain novel insights into the local molecular mechanisms driving the disease progression, thus contributing value to the field with respect to deepening biological understanding and ultimately biomarker and drug development.

We identified in total 265 intra-renal transcripts associated with rapid progression of DKD based on different clinical outcome definitions in a model adjusted for baseline kidney function. The majority of the progression-associated genes were up-regulated, which likely reflect increased transcriptional activity, whereas the interpretation of the down-regulated genes may be obscured due to loss of resident cells in the diseased kidney.

Of note, the different clinical phenotypes of rapid progression highlighted diverse gene expression patterns, thus capturing distinct biological nuances of progressive DKD. Stratification using eGFR slope and CKD stage advancement generated the highest numbers of differentially expressed genes. On the other hand, albuminuria-based analyses were hampered by the lack of UACR variability in this high-risk near-uniformly macroalbuminuric population and, therefore, produced few significant hits.

Interestingly, the DKD progression-associated transcriptomics signature displayed a high degree of kidney compartment specificity, which is in line with biological expectations based on the function of different kidney cell types and independent contributions of glomerular and tubular pathologies to kidney outcomes.³⁹ Furthermore, the progression-associated genes are likely not independent players as they displayed co-expression and shared transcriptional regulation, implying participation in common broader biological programs.

This indicates the possibility to select “hub” targets for intervention in order to develop new therapeutic modalities with pleiotropic beneficial effects.

As the heterogeneity of the clinical course of DKD patients remains a challenge, there is a strong need for novel non-invasive biomarkers to enable early prediction of rapid progression to optimally allocate patient care and increase the likelihood of slowing down the kidney function loss in high-risk patients. Almost 50% of the protein products associated with rapid progression in our study are excreted and measurable in biofluids. These potential biomarkers, if validated in large cohort studies and implemented in clinical testing, could improve prediction of DKD progression and would add valuable information besides traditional clinical risk factors.

The most prominent biological themes reflected in the rapid progression signature included many known disease-associated pathways, thus corroborating the domain knowledge. The fact that several rapid progression-associated genes also showed modulation in previous cross-sectional studies add evidence of them having either a potential causal role in driving or perpetuating pathophysiological processes. However, a number of progression-associated genes did not show case-control modulation and thus would have been missed previously. Interestingly, several progression-associated genes are notable to have strong enrichment compared to other organs under normal conditions, thus supporting their kidney relevance and potentially suggesting that they play a role both under healthy and pathophysiological conditions.

For example, the top significant progression-associated transcript in the glomerular subsignature was the kidney-enriched *REN* which encodes renin and is expressed exclusively by the specialized juxtaglomerular cells.⁴⁰ It was strongly down-regulated in the rapid progression cases developing $\geq 30\%$ UACR increase post-baseline compared to those who did not. The renin-angiotensin-aldosterone system (RAAS) is a well-validated

therapeutic target in DKD and CKD, and RAAS inhibitors (RAASi) remain standard of care in DKD and CKD due to their blood pressure and albuminuria lowering effects, thereby preventing eGFR decline.⁴¹ Of note, RAASi treatment is known to impact *REN* mRNA levels through transcriptional regulation, which could make interpretation of patient transcriptomics data challenging.⁴² However, all the participants in this study were on RAASi treatment at the time of biopsy.

The top significant gene in the tubulointerstitium was *SOX4*, SRY box transcription factor 4, up-regulated by almost 200% in the rapid progressors whose CKD stage advanced during follow-up. The *SOX4* protein is a transcriptional factor involved in several different processes including kidney embryogenesis,⁴³ tumorigenesis by regulating cell differentiation,⁴⁴ and apoptosis induction.⁴⁵ It is expressed in fetal human kidneys⁴⁶⁻⁴⁷ and in early tubules of mice.⁴⁸ The role of *SOX4* in the adult kidney remains to be elucidated, but since apoptosis and dedifferentiation of proximal tubular epithelial cells are well-known features of DKD,⁴⁹⁻⁵⁰ *SOX4* can be hypothesized to be involved.

Additional pathways that were highlighted among the rapid progression-associated genes in the glomeruli were cytoskeletal rearrangement and circadian rhythm. *NOTCH3* was up-regulated by almost 200% in the rapid progressor group with a steep negative eGFR slope. *NOTCH* pathway activation has previously been observed in DKD glomeruli with podocyte loss⁵¹ and *NOTCH3*, specifically, is implicated in pro-inflammatory and pro-fibrotic changes in the podocytes.⁵² Cytoskeleton reorganization also goes hand in hand with podocyte injury and proteinuria.⁵ Circadian rhythm dysregulation has been linked to decreased GFR⁵³ and increased risk of CKD.⁵⁴ The glomerular extracellular matrix is subject to local circadian dynamics, and there is mounting evidence suggesting that disruption of the intra-glomerular cellular clock plays a role in disease pathogenesis.⁵⁵ Taken together, the rapid progression-associated genes and pathways in glomeruli appear to be all in interplay.

In the tubulointerstitial signature, mitochondrial and endoplasmic reticulum (ER) processes emerged. Mitochondrial oxidative phosphorylation is the main source of ATP-based energy in the kidney. In DKD, a decrease in the mitochondrial membrane potential in the proximal tubules initiates mitophagy leading to mitochondrial dysfunction and the metabolic switching to glycolysis to compensate for unmet ATP demands.⁵⁶ Consistent with this, we found the mitochondrial oxidative phosphorylation pathways were transcriptionally inhibited in the tubulointerstitium of rapidly progressing DKD cases, likely reflecting impaired metabolism or failure of compensatory mechanisms.

MAP2K3 was strongly up-regulated in the tubulointerstitium of rapid progression cases at risk for developing nephrotic-range albuminuria. *MAPK* signaling can be modulated in response to ER stress⁵⁷ and its activation has been shown to be implicated in tubulointerstitial fibrosis.⁵⁸ Genetically *MAP2K3*-deficient diabetic mice had reduced *MAPK* signaling and observed to be protected from declining kidney function and increasing albuminuria.⁵⁹

Finally, ribosomal translation processes were activated in the tubulointerstitium of rapid progressors based on steep negative eGFR decline and CKD stage advancement. Translation pathways have been reported enriched in the proximal tubules of DKD kidney biopsies compared to controls.⁶⁰ This increase in mRNA translation could be caused by matrix accumulation and higher amounts of growth factors, both of which are prevalent in tubular injury and progressive DKD.⁶¹ An imbalance brought about by an increase in protein synthesis and degradation can progress to ER stress,⁶⁰ which is one of the critical mechanisms in albumin-induced tubular injury.⁶²

Several limitations of this study should be acknowledged. First, the sample size was limited, even though this DKD dataset remains the largest available with kidney biopsy compartment-specific RNA-seq transcriptomics profiling. However, despite the modest

sample size, there was a sufficient degree of inter-individual heterogeneity and disparate disease progression rates, enabling us to stratify the cohort and identify significant progression-associated transcriptional changes. Larger studies in independent cohorts are warranted to validate these findings and further refine DKD progression mechanistic understanding. Secondly, selection bias cannot be ruled out since a clinically indicated kidney biopsy is rarely performed in DKD and thus, the biopsy cohort may not be representative of the general DKD. However, this study highlighted potential non-invasive biomarker candidates that can be further tested in broader DKD populations. Finally, the duration of the follow-up in this study was limited to five years; it is conceivable that longer clinical observation time might reveal additional biological pathways involved in more chronic processes.

To conclude, our study delivered the first longitudinal analysis linking baseline kidney transcriptomics profile with prospective clinical outcome in diabetic kidney disease. These findings shed light onto the molecular mechanisms driving progressive kidney function decline in DKD, the most common etiology of CKD, and hold the potential to contribute to future diagnostic and therapeutic development.

DISCLOSURES

DA, PJG, and AR are AstraZeneca employees. DA is an industrial PhD student at AstraZeneca. All other authors declare no competing interests.

ACKNOWLEDGMENTS

This research was supported by grants provided by the Stockholm County Council (ALF Project), and the European Union Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 860977 titled TrainCKDis.

Initial preliminary results from this work were previously shared as an abstract and poster at the World Congress of Nephrology 2023 (March 30 - April 2, 2023) and as an abstract and oral presentation at the ASN Kidney Week 2023 (November 2-5, 2023).

DATA SHARING STATEMENT

The raw RNA-seq data analyzed in this study is openly available at <http://karokidney.org/rna-seq-dn/>.

AUTHOR CONTRIBUTIONS

All co-authors have contributed to the manuscript. Specifically, A.We., A.Wi, J.N, P.J.G., A.R. designed and directed the project. A.L. collected the data. D.A. analyzed the data and produced the figures. All authors interpreted the results, and provided critical feedback that helped shape the research, analysis and manuscript. D.A., A.L., A.Wi. drafted the manuscript. All authors approved the final version and accept accountability for the overall work by ensuring that questions pertaining to the accuracy or integrity of any portion of the work were appropriately investigated and resolved.

REFERENCES

1. Li H, Lu W, Wang A, et al: Changing epidemiology of chronic kidney disease as a result of type 2 diabetes mellitus from 1990 to 2017: Estimates from Global Burden of Disease 2017. *J Diabetes Investig*, 12: 346-356, 2021 10.1111/jdi.13355
2. Doshi SM, Friedman AN: Diagnosis and Management of Type 2 Diabetic Kidney Disease. *Clin J Am Soc Nephrol*, 12: 1366-1373, 2017 10.2215/CJN.11111016
3. Selby NM, Taal MW: An updated overview of diabetic nephropathy: Diagnosis, prognosis, treatment goals and latest guidelines. *Diabetes Obes Metab*, 22: 3-15, 2020 10.1111/dom.14007
4. Baelde HJ, Eikmans M, Doran PP, et al: Gene expression profiling in glomeruli from human kidneys with diabetic nephropathy. *Am J Kidney Dis*, 43: 636-650, 2004 10.1053/j.ajkd.2003.12.028
5. Pan Y, Jiang S, Hou Q, et al.: Dissection of Glomerular Transcriptional Profile in Patients With Diabetic Nephropathy: SRGAP2a Protects Podocyte Structure and Function. *Diabetes*, 67: 717-730, 2018 10.2337/db17-0755
6. Schmid H, Boucherot A, Yasuda Y, et al.: Modular Activation of Nuclear Factor- κ B Transcriptional Programs in Human Diabetic Nephropathy. *Diabetes*, 55: 2993-3003, 2006 10.2337/db06-0477
7. Lorz C, Benito-Martin A, Boucherot A, et al.: The Death Ligand TRAIL in Diabetic Nephropathy. *J Am Soc Nephrol*, 19: 904-914, 2008 10.1681/ASN.2007050581
8. Cohen CD, Lindenmeyer MT, Eichinger F, et al.: Improved Elucidation of Biological Processes Linked to Diabetic Nephropathy by Single Probe-Based Microarray Data Analysis. *PLoS ONE*, 3: e2937, 2008 10.1371/journal.pone.0002937
9. Nair V, Komorowsky CV, Weil EJ, et al.: A molecular morphometric approach to diabetic kidney disease can link structure to function and outcome. *Kidney Int*, 93: 439-449, 2018 10.1016/j.kint.2017.08.013
10. Berthier CC, Zhang H, Schin M, et al.: Enhanced Expression of Janus Kinase–Signal Transducer and Activator of Transcription Pathway Members in Human Diabetic Nephropathy. *Diabetes*, 58: 469-477, 2009 10.2337/db08-1328
11. Woroniecka KI, Park ASD, Mohtat D, et al.: Transcriptome Analysis of Human Diabetic Kidney Disease. *Diabetes*, 60: 2354-2369, 2011 10.2337/db10-1181
12. Fan Y, Yi Z, D'Agati VD, et al.: Comparison of Kidney Transcriptomic Profiles of Early and Advanced Diabetic Nephropathy Reveals Potential New Mechanisms for Disease Progression. *Diabetes*, 68: 2301-2314, 2019 10.2337/db19-0204
13. Levin A, Reznichenko A, Witasp A, et al.: Novel insights into the disease transcriptome of human diabetic glomeruli and tubulointerstitium. *Nephrol Dial Transplant*, 35: 2059-2072, 2020 10.1093/ndt/gfaa121
14. Levey AS, Stevens LA, Schmid CH, et al.: A New Equation to Estimate Glomerular Filtration Rate. *Ann Intern Med*, 150: 604, 2009 10.7326/0003-4819-150-9-200905050-00006
15. R Core Team: R: The R Project for Statistical Computing. R-project.org. Published 2022. <https://www.r-project.org/>
16. van der Aart-van der Beek AB, Clegg LE, Penland RC, et al.: Effect of once-weekly exenatide on estimated glomerular filtration rate slope depends on baseline renal risk: A *post hoc* analysis of the EXSCEL trial. *Diabetes Obes Metab*, 22: 2493-2498, 2020 10.1111/dom.14175
17. Jhund PS, Solomon SD, Docherty KF, et al.: Efficacy of Dapagliflozin on Renal Function and Outcomes in Patients With Heart Failure With Reduced Ejection Fraction: Results of DAPA-HF. *Circulation*, 143: 298-309, 2021 10.1161/CIRCULATIONAHA.120.050391
18. Thul PJ, Åkesson L, Wiking M, et al.: A subcellular map of the human proteome. *Science*, 356: eaal3321, 2017 10.1126/science.aal3321

19. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group: KDIGO 2012 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease: Chapter 1: Definition and classification of CKD. *Kidney Int Suppl*, 3: 19-62, 2013 10.1038/kisup.2012.64
20. Eijkelkamp WBA, Zhang Z, Remuzzi G, et al.: Albuminuria Is a Target for Renoprotective Therapy Independent from Blood Pressure in Patients with Type 2 Diabetic Nephropathy: Post Hoc Analysis from the Reduction of Endpoints in NIDDM with the Angiotensin II Antagonist Losartan (RENAAL) Trial. *J Am Soc Nephrol*, 18: 1540-1546, 2007 10.1681/ASN.2006050445
21. Neal B, Perkovic V, Mahaffey KW, et al.: Canagliflozin and Cardiovascular and Renal Events in Type 2 Diabetes. *N Engl J Med*, 377: 644-657, 2017 10.1056/NEJMoa1611925
22. Wiviott SD, Raz I, Bonaca MP, et al.: Dapagliflozin and Cardiovascular Outcomes in Type 2 Diabetes. *N Engl J Med*, 380: 347-357, 2019 10.1056/NEJMoa1812389
23. Perkovic V, Koitka-Weber A, Cooper ME, et al.: Choice of endpoint in kidney outcome trials: considerations from the EMPA-REG OUTCOME® trial. *Nephrol Dial Transplant*, 35: 2103-2111, 2020 10.1093/ndt/gfz179
24. Cherney DZI, Charbonnel B, Cosentino F, et al.: Effects of ertugliflozin on kidney composite outcomes, renal function and albuminuria in patients with type 2 diabetes mellitus: an analysis from the randomised VERTIS CV trial. *Diabetologia*, 64: 1256-1267, 2021 10.1007/s00125-021-05407-5
25. Love MI, Huber W, Anders S: Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*, 15: 550, 2014 10.1186/s13059-014-0550-8
26. Radcliffe NJ, Seah J-m, Clarke M, et al.: Clinical predictive factors in diabetic kidney disease progression. *J Diabetes Investig*, 8: 6-18, 2017 10.1111/jdi.12533
27. Krämer A, Green J, Pollard J, et al.: Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics*, 30: 523-530, 2014 10.1093/bioinformatics/btt703
28. Huang DW, Sherman BT, Lempicki RA: Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*, 4: 44-57, 2009 10.1038/nprot.2008.211
29. Sherman BT, Hao M, Qiu J, et al.: DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res*, 50: W216-W221, 2022 10.1093/nar/gkac194
30. Ashburner M, Ball CA, Blake JA, et al.: Gene Ontology: tool for the unification of biology. *Nat Genet*, 25: 25-29, 2000 10.1038/75556
31. The Gene Ontology Consortium: The Gene Ontology resource: enriching a GOLD mine. *Nucleic Acids Res*, 49: D325-D334, 2021 10.1093/nar/gkaa1113
32. Lonsdale J, Thomas J, Salvatore M, et al.: The Genotype-Tissue Expression (GTEx) project. *Nat Genet*, 45: 580-585, 2013 10.1038/ng.2653
33. Langfelder P, Horvath S: WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, 9: 559, 2008 10.1186/1471-2105-9-559
34. Shannon P, Markiel A, Ozier O, et al.: Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res*, 13: 2498-2504, 2003 10.1101/gr.1239303
35. Edgar R: Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res*, 30: 207-210, 2002 10.1093/nar/30.1.207
36. Grayson PC, Eddy S, Taroni JN, et al.: Metabolic pathways and immunometabolism in rare kidney diseases. *Ann Rheum Dis*: annrheumdis-2017-212935, 2018 10.1136/annrheumdis-2017-212935
37. Gillespie M, Jassal B, Stephan R, et al.: The reactome pathway knowledgebase 2022. *Nucleic Acids Res*, 50: D687-D692, 2022 10.1093/nar/gkab1028
38. Subramanian A, Tamayo P, Mootha VK, et al.: Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*, 102: 15545-15550, 2005 10.1073/pnas.0506580102

39. An Y, Xu F, Le W, et al.: Renal histologic changes and the outcome in patients with diabetic nephropathy. *Nephrol Dial Transplant*, 30: 257-266, 2015
10.1093/ndt/gfu250
40. Yamaguchi H, Gomez RA, Sequeira-Lopez MLS: Renin Cells, From Vascular Development to Blood Pressure Sensing. *Hypertension*, 80: 1580-1589, 2023
10.1161/HYPERTENSIONAHA.123.20577
41. Rossing P, Caramori ML, Chan JCN, et al.: Executive summary of the KDIGO 2022 Clinical Practice Guideline for Diabetes Management in Chronic Kidney Disease: an update based on rapidly emerging new evidence. *Kidney Int*, 102: 990-999, 2022
10.1016/j.kint.2022.06.013
42. Samani NJ, Cumin F, Kelly M, et al.: Expression of components of the RAS during prolonged blockade at different levels in primates. *Am J Physiol Endocrinol Metab*, 267: E612-E619, 1994 10.1152/ajpendo.1994.267.4.E612
43. Huang J, Arsenault M, Kann M, et al.: The transcription factor sry-related HMG box-4 (SOX4) is required for normal renal development *in vivo*: SOXC Genes During Renal Development. *Dev Dyn*, 242: 790-799, 2013 10.1002/dvdy.23971
44. Moreno CS: SOX4: The unappreciated oncogene. *Semin Cancer Biol*, 67: 57-64, 2020 10.1016/j.semcancer.2019.08.027
45. Zhang L, Lv L, Zheng N, et al.: Suppression of Sox4 protects against myocardial ischemic injury by reduction of cardiac apoptosis in mice. *J Cell Physiol*, 236: 1094-1104, 2021 10.1002/jcp.29918
46. Cao J, O'Day DR, Pliner HA, et al.: A human cell atlas of fetal gene expression. *Science*, 370: eaba7721, 2020 10.1126/science.aba7721
47. Lindström NO, McMahon JA, Guo J, et al.: Conserved and Divergent Features of Human and Mouse Kidney Organogenesis. *J Am Soc Nephrol*, 29: 785-805, 2018
10.1681/ASN.2017080887
48. Yu J, Valerius MT, Duah M, et al.: Identification of molecular compartments and genetic circuitry in the developing mammalian kidney. *Development*, 139: 1863-1873, 2012 10.1242/dev.074005
49. Habib SL: Diabetes and renal tubular cell apoptosis. *World J Diabetes*, 4: 27, 2013
10.4239/wjd.v4.i2.27
50. Schnaper HW: The Tubulointerstitial Pathophysiology of Progressive Kidney Disease. *Adv Chronic Kidney Dis*, 24: 107-116, 2017 10.1053/j.ackd.2016.11.011
51. Niranjana T, Bielez B, Gruenwald A, et al.: The Notch pathway in podocytes plays a role in the development of glomerular disease. *Nat Med*, 14: 290-298, 2008
10.1038/nm1731
52. El Machhour F, Keuylian Z, Kavvadas P, et al.: Activation of Notch3 in Glomeruli Promotes the Development of Rapidly Progressive Renal Disease. *J Am Soc Nephrol*, 26: 1561-1575, 2015 10.1681/ASN.2013090968
53. Charles LE, Gu JK, Fekedulegn D, et al.: Association Between Shiftwork and Glomerular Filtration Rate in Police Officers. *J Occup Environ Med*, 55: 1323-1328, 2013 10.1097/JOM.0b013e3182a299c9
54. Agarwal R, Light RP: Sleep and Activity in Chronic Kidney Disease: A Longitudinal Study. *Clin J Am Soc Nephrol*, 6: 1258-1265, 2011 10.2215/CJN.10581110
55. Preston R, Meng Q-J, Lennon R: The dynamic kidney matrisome - is the circadian clock in control? *Matrix Biol*, 114: 138-155, 2022 10.1016/j.matbio.2022.05.005
56. Forbes JM, Thorburn DR: Mitochondrial dysfunction in diabetic kidney disease. *Nat Rev Nephrol*, 14: 291-312, 2018 10.1038/nrneph.2018.9
57. Darling NJ, Cook SJ: The role of MAPK signalling pathways in the response to endoplasmic reticulum stress. *Biochim Biophys Acta Mol Cell Res*, 1843: 2150-2163, 2014 10.1016/j.bbamcr.2014.01.009
58. Rhyu DY, Yang Y, Ha H, et al.: Role of Reactive Oxygen Species in TGF- β 1-Induced Mitogen-Activated Protein Kinase Activation and Epithelial-Mesenchymal Transition in Renal Tubular Epithelial Cells. *J Am Soc Nephrol*, 16: 667-675, 2005
10.1681/ASN.2004050425

59. Lim AKH, Nikolic-Paterson DJ, Ma FY, et al.: Role of MKK3–p38 MAPK signalling in the development of type 2 diabetes and renal injury in obese db/db mice. *Diabetologia*, 52: 347-358, 2009 10.1007/s00125-008-1215-5
60. Collins KS, Eadon MT, Cheng Y-H, et al.: Alterations in Protein Translation and Carboxylic Acid Catabolic Processes in Diabetic Kidney Disease. *Cells*, 11: 1166, 2022 10.3390/cells11071166
61. Fan Y, Lee K, Wang N, et al.: The Role of Endoplasmic Reticulum Stress in Diabetic Nephropathy. *Curr Diab Rep*, 17: 17, 2017 10.1007/s11892-017-0842-y
62. Lindenmeyer MT, Rastaldi MP, Ikehata M, et al.: Proteinuria and Hyperglycemia Induce Endoplasmic Reticulum Stress. *J Am Soc Nephrol*, 19: 2225-2236, 2008 10.1681/ASN.2007121313

Table 1. Counts of significantly modulated genes in the kidneys of rapid progressors vs non-rapid progressors.

Rapid disease progression definition	Kidney tissue compartment	Total number of differentially expressed genes	Number of Up-regulated genes	Number of Down-regulated genes
eGFR slope < -5ml/min/year	Glomerular	97	77	20
	Tubulointerstitial	59	39	20
CKD stage advancement	Glomerular	49	38	11
	Tubulointerstitial	66	55	11
≥ 30% UACR increase	Glomerular	4	1	3
	Tubulointerstitial	6	2	4
NRA status	Glomerular	1	1	0
	Tubulointerstitial	9	7	2
Composite of KF or ≥40% eGFR decline	Glomerular	3	0	3
	Tubulointerstitial	2	1	1

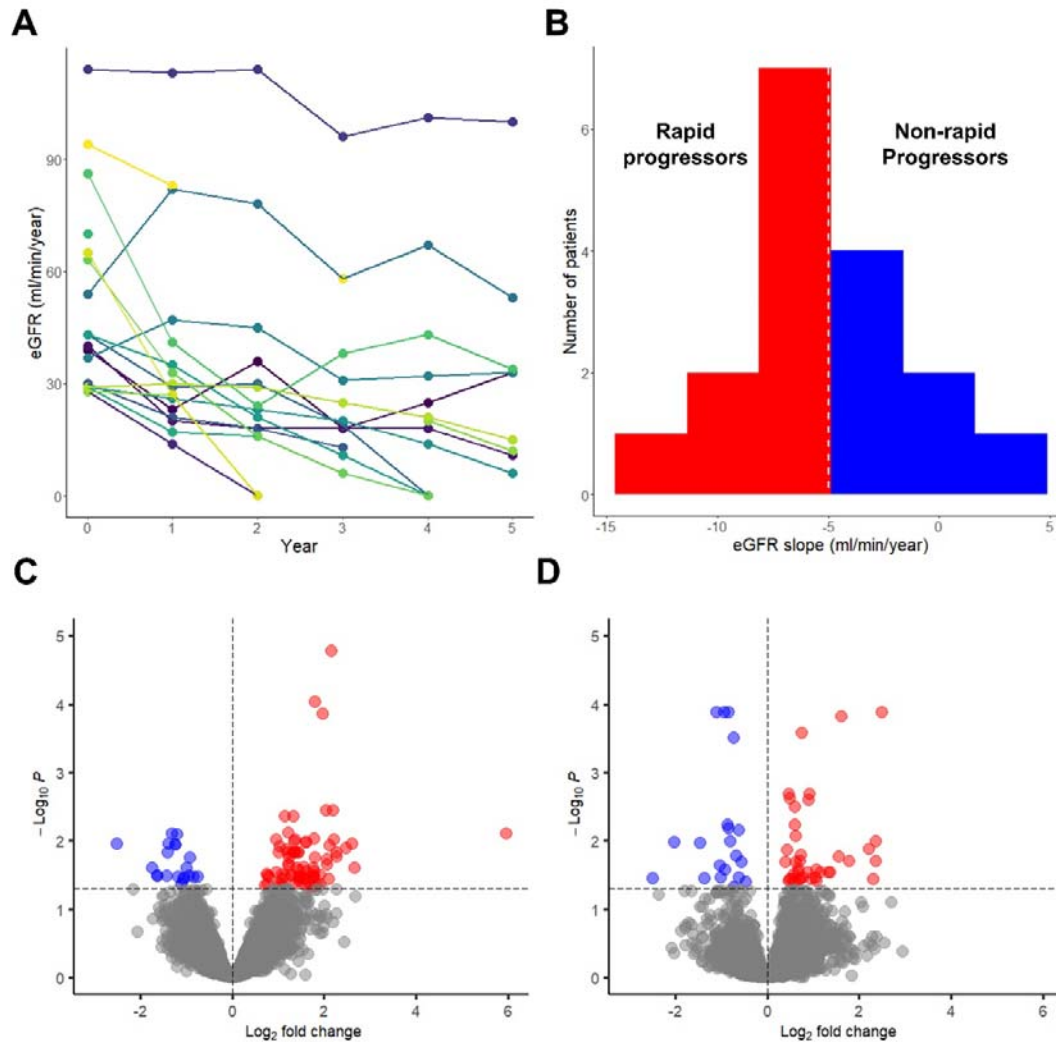


Figure 1. Rapid progression analysis by steep eGFR slope. (A) Individual eGFR trajectories during follow-up. The time-series plot present individual eGFR values from baseline (Year 0) through the duration of follow-up (Year 1-5). eGFR values were set to 0 for patients that reached kidney failure. (B) Histogram of calculated eGFR slopes distributions. Non-rapid progressors are in blue (eGFR slope > -5 ml/min/ year) while rapid progressors are in red (eGFR slope < -5 ml/min/year). (C) DEA of rapid vs non-rapid progressor status in glomeruli. (D) DEA of rapid vs non-rapid progressor status in tubulointerstitium. (C,D) Volcano plots show the log_2 fold changes in gene expression (x-axis) and the negative logarithm of the adjusted p-value (y-axis) between the rapid and non-rapid progressors. The horizontal dashed line indicates the significance threshold at adjusted $P < 0.05$ while the vertical dashed line indicates zero fold change. Each point above the horizontal line corresponds to a significantly differentially expressed gene (red: up-regulated, blue: down-regulated). DEA, differential expression analysis; eGFR, estimated glomerular filtration rate.

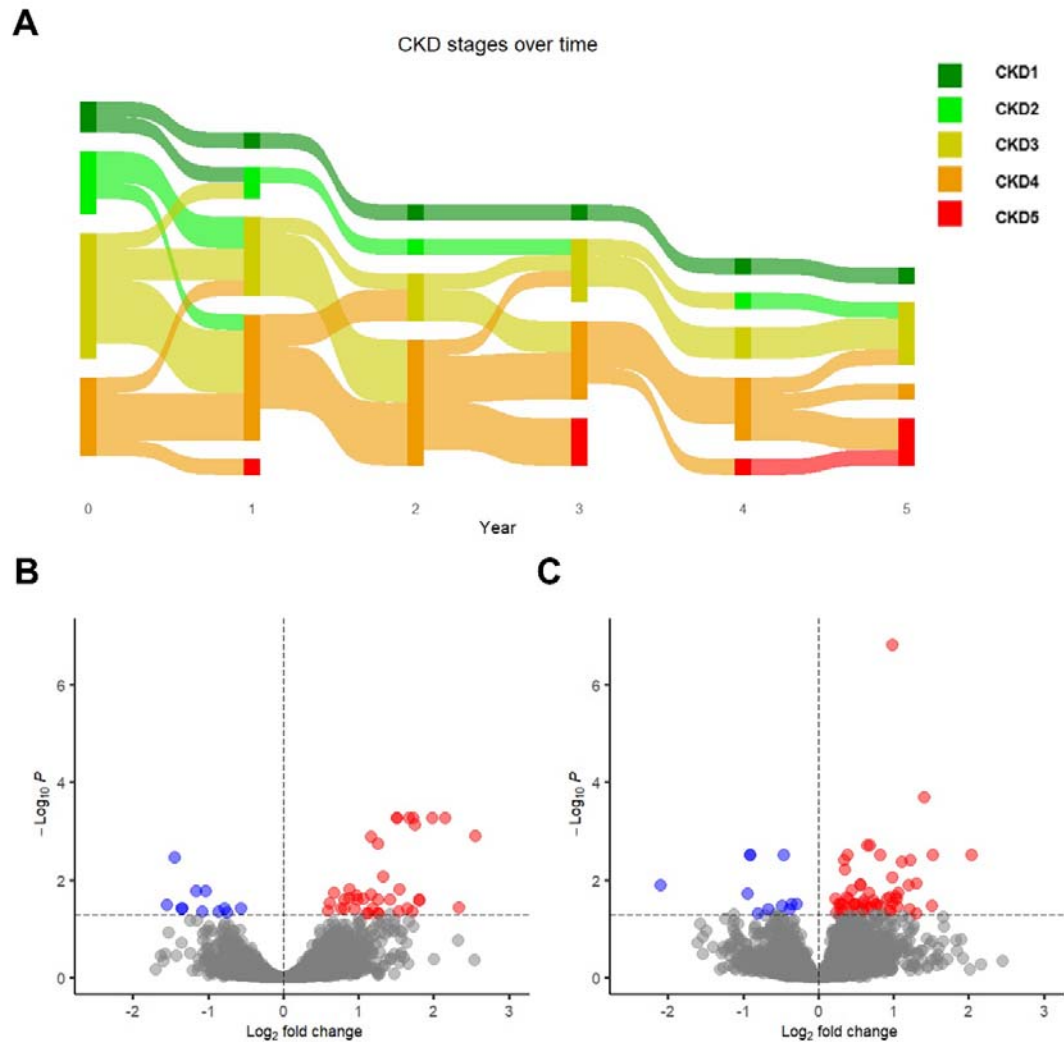


Figure 2. Rapid progression analysis by CKD stage advancement beyond baseline. (A) CKD stage classification changes during follow-up. The Sankey plot demonstrate changes in CKD stage from baseline (Year 0) through the duration of follow-up (Year 1-5). The color ribbons reflect the patient subgroup sizes per category. The unequal number of patients per timepoint exemplifies patients lost to follow-up. (B) DEA rapid vs non-rapid progressor status in glomeruli. (C) DEA rapid vs non-rapid progressor status in tubulointerstitium. (B,C) Volcano plots show the log₂ fold changes in gene expression (x-axis) and the negative logarithm of the adjusted p-value (y-axis) between the rapid and non-rapid progressors. The horizontal dashed line indicates the significance threshold at adjusted P < 0.05 while the vertical dashed line indicates zero fold change. Each point above the horizontal line corresponds to a significantly differentially expressed gene (red: up-regulated, blue: down-regulated). CKD, chronic kidney disease; DEA, differential expression analysis; eGFR, estimated glomerular filtration rate.

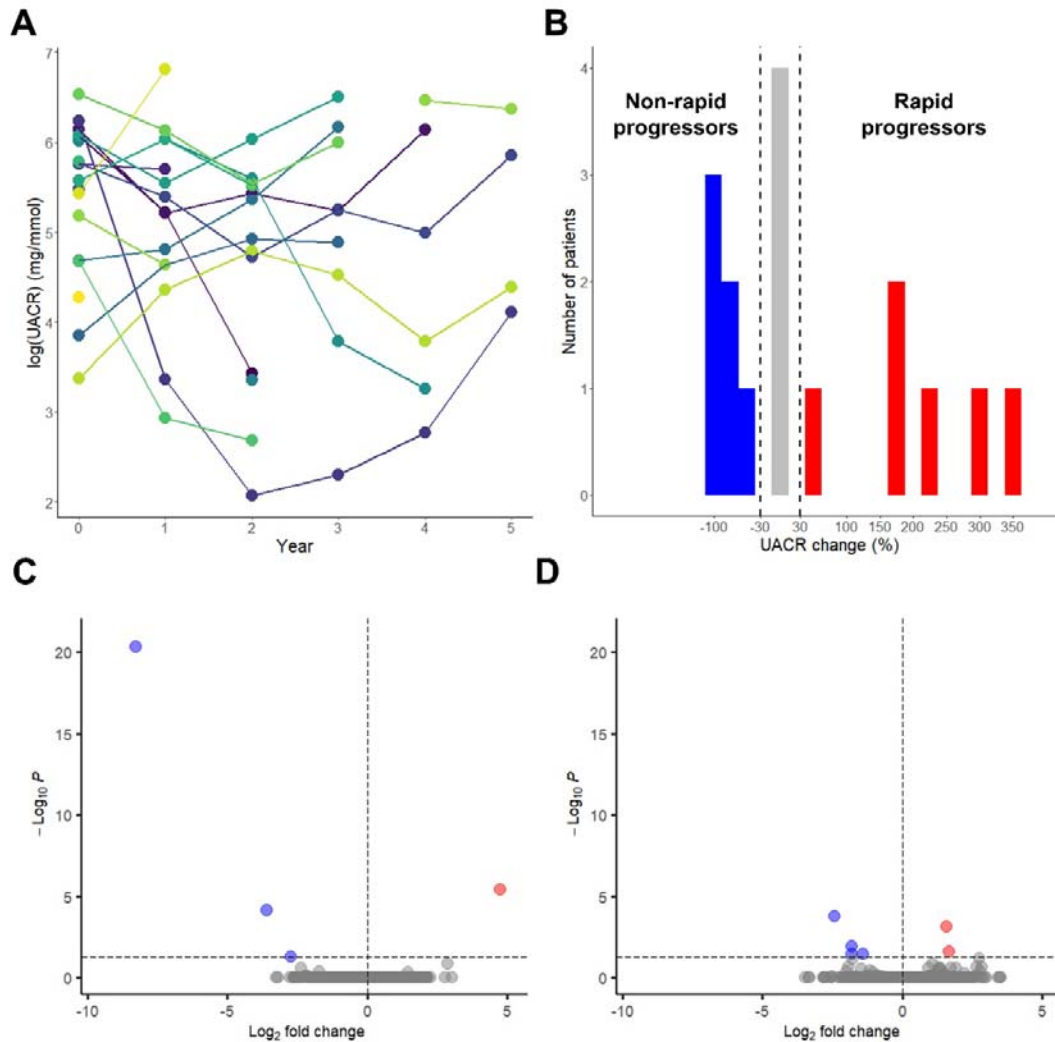


Figure 3. Rapid progression analysis by percentage of UACR change. (A) Individual UACR trajectories during follow-up. The time-series plot present individual $\log(\text{UACR})$ values from baseline (Year 0) through the duration of follow-up (Year 1-5). (B) Histogram of percentage of UACR change distributions. Non-rapid progressors are in blue ($\geq 30\%$ UACR decrease) while rapid progressors are in red ($\geq 30\%$ UACR increase). (C) DEA of rapid vs non-rapid progressor status in glomeruli. (D) DEA of rapid vs non-rapid progressor status in tubulointerstitium. (C,D) Volcano plots show the \log_2 fold changes in gene expression (x-axis) and the negative logarithm of the adjusted p-value (y-axis) between the rapid and non-rapid progressors. The horizontal dashed line indicates the significance threshold at adjusted $P < 0.05$ while the vertical dashed line indicates zero fold change. Each point above the horizontal line corresponds to a significantly differentially expressed gene (red: up-regulated, blue: down-regulated). DEA, differential expression analysis; eGFR, estimated glomerular filtration rate; UACR, urinary albumin-to-creatinine ratio.

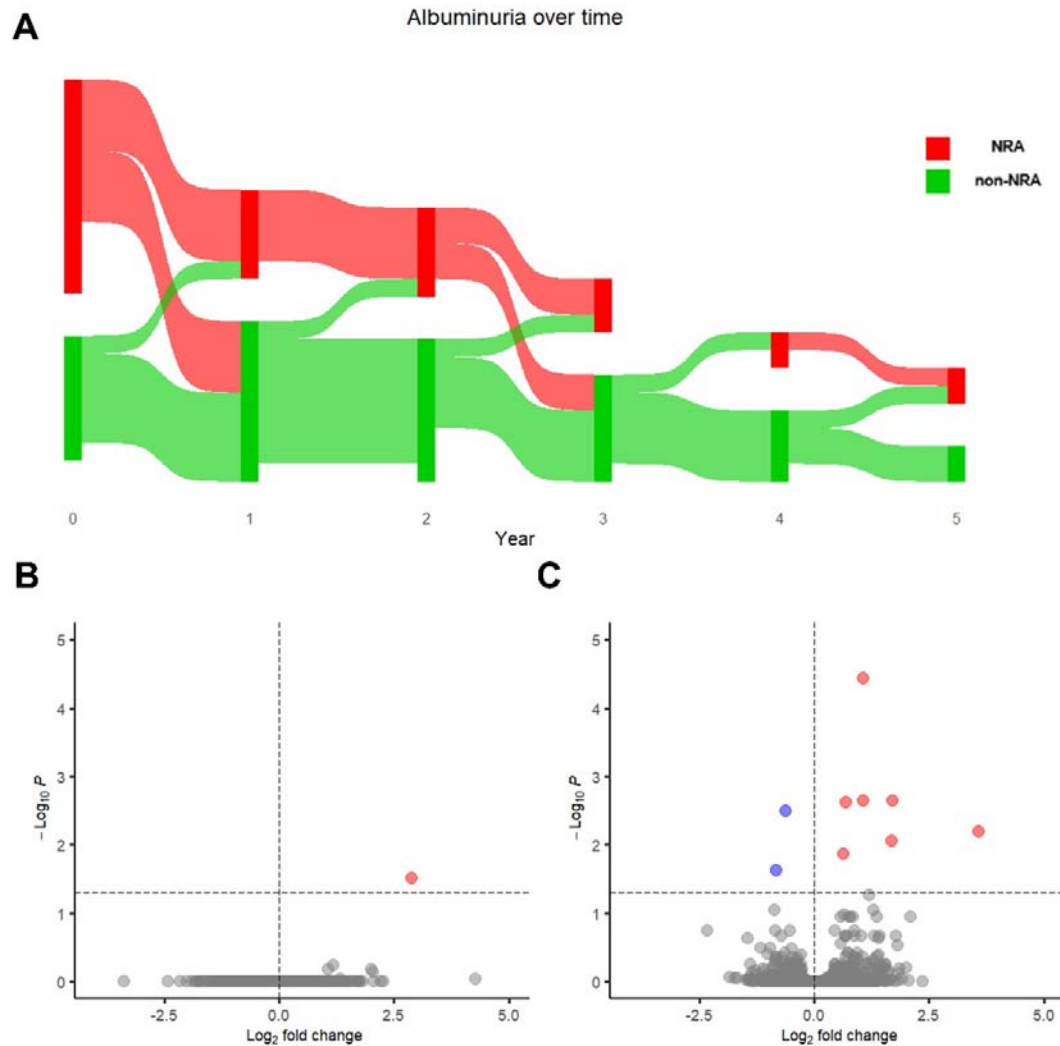


Figure 4. Rapid progression analysis by nephrotic-range albuminuria status. (A) NRA classification changes during follow-up. The Sankey plot demonstrate changes in NRA status from baseline (Year 0) through the duration of follow-up (Year 1-5). The color ribbons reflect the patient subgroup sizes per category. The unequal number of patients per timepoint exemplifies patients lost to follow-up. (B) DEA rapid vs non-rapid progressor status in glomeruli. (C) DEA rapid vs non-rapid progressor status in tubulointerstitium. (B,C) Volcano plots show the log_2 fold changes in gene expression (x-axis) and the negative logarithm of the adjusted p-value (y-axis) between the rapid and non-rapid progressors. The horizontal dashed line indicates the significance threshold at adjusted $P < 0.05$ while the vertical dashed line indicates zero fold change. Each point above the horizontal line corresponds to a significantly differentially expressed gene (red: up-regulated, blue: down-regulated). CKD, chronic kidney disease; DEA, differential expression analysis; eGFR, estimated glomerular filtration rate; NRA, nephrotic-range albuminuria.

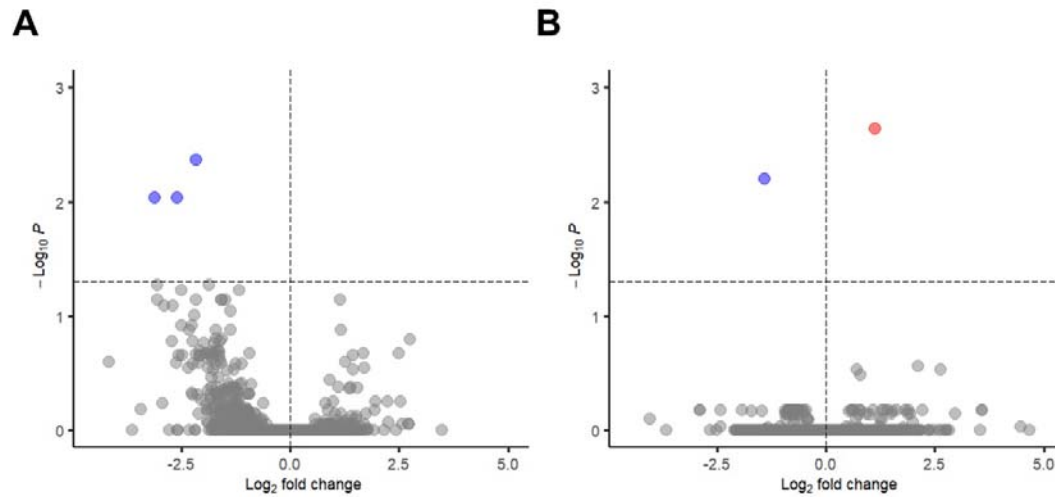


Figure 5. Rapid progression analysis by composite outcome. (A) DEA rapid vs non-rapid progressor status in glomeruli. (B) DEA rapid vs non-rapid progressor status in tubulointerstitium. (A,B) Volcano plots show the \log_2 fold changes in gene expression (x-axis) and the negative logarithm of the adjusted p-value (y-axis) between the rapid and non-rapid progressors. The horizontal dashed line indicates the significance threshold at adjusted $P < 0.05$ while the vertical dashed line indicates zero fold change. Each point above the horizontal line corresponds to a significantly differentially expressed gene (red: up-regulated, blue: down-regulated). DEA, differential expression analysis.

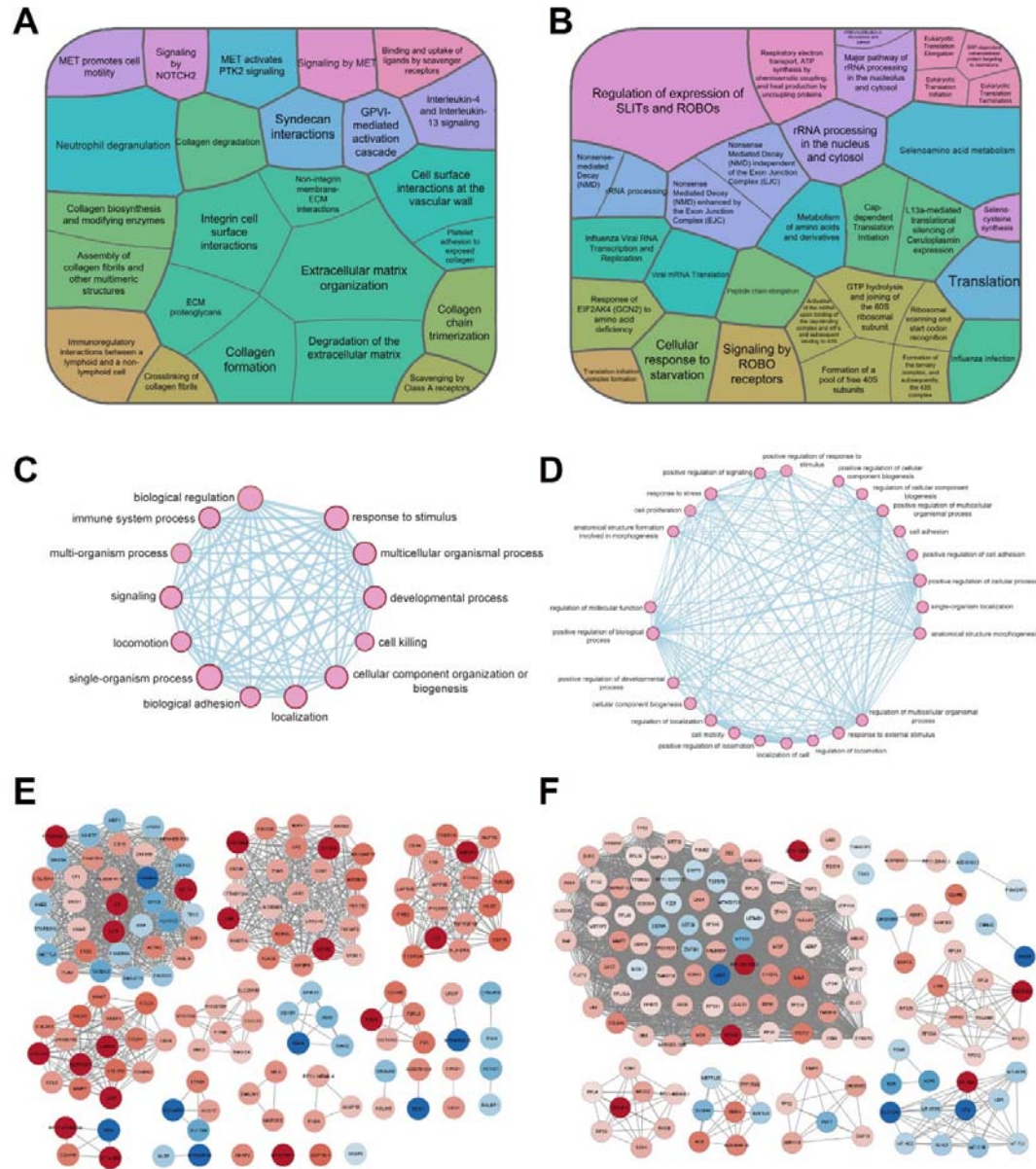


Figure 6. Rapid progression genes and associated pathways. (A,B) Treemaps show the results of the Reactome pathway over-representation analysis of rapid progression-associated genes according to all five disease progression definitions in the (A) glomeruli and (B) tubulointerstitium. The color of the cells denotes biologically related pathways (same Reactome parent pathway). The size of the cells reflects the size of the overlap between the progression-associated genes and the corresponding pathway. Only significant terms (adjusted $P < 0.05$) are shown. (C,D) Network maps show the significant ($FDR < 0.05$) Gene Ontology biological processes in the (C) glomeruli and (D) tubulointerstitium determined through over-representation analysis. (E,F) Co-expression network maps of the rapid progression-associated genes in the (E) glomeruli and (F) tubulointerstitium. Connected genes are co-expressed in the specified module or subnetwork. Genes are colored based on \log_2 fold changes, where red and blue indicate up-regulation and down-regulation, respectively. FDR, false discovery rate.