SARS-CoV-2 receptor networks in diabetic kidney disease, BK-Virus nephropathy and COVID-19 associated acute kidney injury

Rajasree Menon1*, Edgar E. Otto1*, Rachel Sealfon2*, Viji Nair1*, Aaron K. Wong2*, Chandra L. Theesfeld4, Xi Chen2, Yuan Wang4,5, Avinash Boppana5, Peter M. Kasson6, Jennifer A. Schaub1, Celine C. Berthier1, Sean Eddy1, Chrysta C. Lienczewski1, Bradley Godfrey1, Susan L. Dagenais1, Ryann Sohaney4, John Hartman1, Damian Fermin1, Lalita Subramanian1, Helen C. Looker3, Laura H. Mariani1, Abhijit S. Naik1, Robert G. Nelson1, Olga G. Troyanskaya2,4,5#, Matthias Kretzler1#

1Michigan Medicine, Ann Arbor, MI, USA
2Flatiron Institute, Simons Foundation, New York, NY, USA
3Chronic Kidney Disease Section, National Institute of Diabetes and Digestive and Kidney Diseases, Phoenix, AZ, USA
4Princeton University, Lewis-Sigler Institute for Integrative Genomics, Princeton, NJ USA
5Princeton University, Department of Computer Science, Princeton, NJ USA
6Departments of Molecular Physiology and Biomedical Engineering, University of Virginia, Charlottesville, VA, USA

*These authors contributed equally to the manuscript

#Corresponding authors:
Dr. Matthias Kretzler, Warner-Lambert/Parke-Davis Professor of Medicine, Nephrology/Internal Medicine and Computational Medicine and Bioinformatics, University of Michigan, B II, 4544-D 1150 W. Medical Center Dr., Ann Arbor, MI, USA
Email: kretzler@umich.edu

Olga Troyanskaya, Ph.D., Professor Department of Computer Science and Lewis-Sigler Institute for Integrative Genomics, Princeton University, NJ, USA
Deputy Director for Genomics, Flatiron Institute Simons Foundation, New York, NY, USA
Email: ogt@cs.princeton.edu

Running Title: SARS-CoV-2 receptors in diabetic and viral kidney disease
Abstract

COVID-19 morbidity and mortality is significantly increased in patients with diabetes and kidney disease via unknown mechanisms. SARS-CoV-2 uses angiotensin-converting enzyme 2 (ACE2) for entry into human host cells, and ACE2 levels in target cells may influence SARS-CoV-2 susceptibility. We investigated how pre-existing conditions and drug treatments alter receptor expression in kidney tissue. Using single cell RNA profiling (scRNAseq) to assess ACE2 and associated SARS-CoV-2 proteases in healthy living donors (LD) kidneys, diabetic kidney disease (DKD), and in kidney injury during viral infection, ACE2 expression was primarily associated with proximal tubular epithelial cells (PTEC). ACE2 mRNA expression levels were significantly upregulated in DKD versus LD, however, ACE2 levels were not altered by exposures to renin angiotensin aldosterone system (RAAS) inhibitors. ACE2+ expression signatures were defined by differential expression analysis and characterized by Bayesian integrative analysis of a large compendium of public -omics datasets, resulting in the identification of network modules induced in ACE2 positive PTEC in DKD and BK virus nephropathy. These ACE2 upregulated cell programs were linked to viral entry, immune activation, endomembrane reorganization, and RNA processing and overlapped significantly with the cellular responses induced by SARS-CoV-2 infection. Similar cellular programs were activated in ACE2-positive PTEC isolated in a urine sample from a COVID19 patient with acute kidney injury, suggesting a consistent ACE2-coregulated expression program that may interact with SARS-Cov-2 infection processes. The SARS-CoV-2 receptor associated gene signatures could seed further research into therapeutic strategies for COVID-19. Functional networks of gene expression signatures are available for further exploration to researchers at HumanBase (hb.flatironinstitute.org/covid-kidney).
Introduction

COVID-19 disproportionally affects individuals with diabetes, hypertension, and kidney disease \(^1\). Yet the underlying molecular and physiological causes of this association are unknown, and could be as varied as drugs used to treat these conditions \(^4\), disease biology \(^5\)-\(^7\), direct infection of relevant organs by the virus \(^8\), \(^9\) and consequent tissue destruction, and cytokine storm that occurs secondary to infection \(^1\). While upper and lower airway tissues are the primary sites of infection, additional sites are likely to include kidney tissue, and initial data suggest that patients with kidney disease suffer significantly higher mortality than age-matched individuals without these conditions \(^10\)-\(^12\). Understanding the disease-specific molecular processes associated with COVID-19 in patients with kidney disease and diabetes can have a significant impact on public health.

COVID-19 develops from infection with SARS-CoV-2, a betacoronavirus with a single stranded RNA genome. It gains entry into specific cell types through interaction of surface spike protein with a cell surface receptor \(^13\). Studies of severe acute respiratory syndrome (SARS) in the early 2000s identified angiotensin-converting enzyme 2 (ACE2) as the primary cell-entry receptor for the SARS coronavirus (SARS-CoV) in humans \(^14\), \(^15\). Recent studies from the COVID-19 pandemic demonstrate that ACE2 is also the primary cell-entry receptor for SARS-CoV-2 \(^13\), \(^16\), \(^17\) and, similar to SARS-CoV, higher ACE2 expression may lead to higher risk of SARS-CoV-2 infection \(^8\), \(^18\).

ACE2 is a membrane bound zinc metallopeptidase and a master regulator of the renin-angiotensin aldosterone system (RAAS). Normally, ACE2 is expressed as a cell-surface protein to which both SARS-CoV and SARS-CoV-2 can bind \(^4\), \(^19\). Cleavage of spike protein is required for fusion of the virus with host membranes and occurs in a manner similar to SARS-CoV:
TMPRSS2 appears to be the primary means of this activation, at least in lung epithelial cells; however cleavage can also occur in endosomes by acid-activated proteases such as cathepsin L (CTSL) 17, 20, 21.

Direct access to SARS-CoV-2 target tissues combined with the granularity of single cell RNA sequencing (scRNAseq) technology can locate specific cell types with enhanced ACE2 expression to determine whether these cells also possess additional molecular machinery that facilitate initial viral entry and subsequent cellular cytotoxicity. Characterization of these molecular processes in existing tissue from relevant cohorts could greatly accelerate the identification and development of therapeutic options for SARS-CoV-2 infection.

In this study, we explored the expression and associated biological processes of ACE2 and other viral co-receptors in kidney cells from healthy living donors (LD), and from patients with diabetic kidney disease (DKD), BK virus nephropathy (BKVN) and COVID-19 associated AKI (COV-AKI). Using scRNAseq data, we defined the cellular expression of ACE2 in proximal tubular epithelial cells (PTEC), characterized the cellular programs associated with ACE2 and its regulation in DKD, BKVN and COV-AKI, mapped the ACE2 associated changes to emerging data on SARS-CoV-2 induced cellular responses and tested the associations with RAAS medication exposure (Figure 1). Our study focused on DKD since COVID-19 disproportionally affects individuals with diabetes and kidney disease1, 3, 7, 11, 22, 23, and research kidney biopsies were available in a comprehensive DKD cohort for scRNAseq profiling and analysis. The DKD analyses could identify underlying mechanisms of excessive risk associated with COVID-19 in patients with diabetes or chronic kidney disease, with the studies in BKVN and COV-AKI establishing the link between DKD and virus-associated kidney disease.
Methods

Study Population

DKD biopsy cohort

An early DKD cohort was selected as the discovery cohort given the impact of COVID-19 on individuals with diabetes and kidney disease. Between 2013 and 2017, 141 American Indians (39 men, 102 women) from southern Arizona with type 2 diabetes were enrolled in an observational study of DKD that included annual research examinations. Glomerular filtration rate (GFR) was measured at each of these examinations by the urinary clearance of iothalamate. At baseline, each participant also underwent a research kidney biopsy used for quantitative morphometry and for compartment-specific tissue expression studies. For kidney biopsies conducted in the last 44 participants enrolled between 2016 and 2017, processing of the kidney specimens was modified to enable scRNAseq studies using kidney biopsy tissues. These 44 participants (DKD) were included in the present study.

At each annual research examination, medicines taken by the participants were recorded. Blood pressure (BP) was measured with the participant resting in a seated position, and height and weight were obtained with the participant dressed in light clothing without shoes. Body mass index was defined as weight in kilograms divided by the square of height in meters. Urine and serum creatinine were measured by an enzymatic method. Urinary albumin excretion was assessed by nephelometric immunoassay and expressed as the albumin/creatinine ratio (ACR) in mg/g creatinine. GFR and HbA1c were measured by high performance liquid chromatography. Because of the level of obesity in the participants, absolute GFR (ml/min) was reported, uncorrected for body surface area.
Clinical and demographic features and the antihypertensive medicines taken by the DKD cohort at the time of the kidney biopsy are shown in Table 1. Participants had mean diabetes duration of 12.2 ± 7.5 years, HbA1c of 9.2 ± 2.4 %, iGFR of 159 ± 58 ml/min with one patient GFR <60ml/min, and median ACR of 18 [9-53] mg/g (25 persons had a normal ACR <30 mg/g, 14 had moderate albuminuria of 30-300 mg/g and 3 had severe albuminuria >300mg/g of ACR at biopsy).

Allograft Kidney Transplant Surveillance cohort

The University of Michigan has a kidney transplant surveillance program in which biopsies are performed for clinical care at implantation and 3-, 6- and 12-months post transplantation. A subset of patients in this surveillance program participate in the Human Kidney Transplant Transcriptomic Atlas study, which procures an additional biopsy core for research and tissues from these biopsies were used in the present study. Reference healthy tissue from 7 LD biopsies were obtained prior to perfusion and before placement in the recipient.

To evaluate the kidney epithelial response to viral infection 3 biopsies from 2 patients with active BK viremia at the time of their surveillance biopsies were also included in the study. BKVN biopsies were obtained before, at the time of, or after peak BK viremia by which time immunosuppression was reduced.

LD biopsies were selected as healthy reference tissue. Participant age was 49.3 ± 12, range (31-66), five of the 7 donors were female (71.4%), mean spot urine protein to creatinine ratio at the time of evaluation was 0.07 ± 0.03 g/g with range (0.04-0.12), mean iothalamate GFR was 87 ± 5 ml/min/1.73 m^2 with range (81-96). All donors were white, nondiabetic and non-hypertensive.
**BK Virus Nephropathy (BKVN)** biopsies from the transplant atlas cohort study were used, with the first BKVN patient being the recipient of a deceased donor kidney transplant, with a peak viral load >10,000 copies/ml and immunohistochemistry staining positive for SV40 (peak biopsy). After net reduction in immunosuppression the BK viral load dropped to <5000 copies/ml and a subsequent biopsy became negative for SV40 staining (post peak biopsy). The second patient was the recipient of a living donor kidney transplant and had a viral load of >10,000 copies/ml that had still not reached its peak at time of biopsy, but with positive SV40 biopsy staining (pre peak biopsy). Both patients were in the 3rd and 4th decade of life respectively. Clinical characteristics of the transplant atlas cohort participants are provided in Supplementary Table 1.

**COVID-19 AKI Cohort Study (COV-AKI)**

In response to the COVID-19 pandemic at the University of Michigan Medical Center a protocol to capture urine samples from SARS-CoV-2 patients was established to allow cell and molecular biology studies of the kidney manifestations in COVID-19. Patients are offered participation in the study upon admission to the hospital and throughout their hospital course. 50ml of a spot random urine sample is procured specifically for the study, immediately transported to the research laboratory on ice and processed for cell and molecular studies and participants are prospectively followed thereafter.

**COVID-19 associated AKI (COV-AKI)** was studied in urine derived PTEC from a male patient in the 5th decade of his life who had received a combined pancreas and kidney transplant two decades ago. The patient was diagnosed with COVID-19 eleven days before sample collection. Serum creatinine at time of collection was 2.3 mg/dl compared to baseline of 1.6 mg/dl. Urinalysis was significant for proteinuria, but with no RBCs on microscopy. No urinary sediment analysis was available at time of sample procurement. At the time of COVID-19 diagnosis the
patient’s immunosuppression was reduced to dual immunosuppression with tacrolimus and prednisone with mycophenolate mofetil withhold. His kidney function returned to baseline 17 days after admission, 6 days after sample collection (Supplementary Figure 1).

**Kidney biopsy sample processing for single cell**

Single cell transcriptomes were generated with 2-3 mg of the biopsy core samples from 44 CryoStor® (Stemcell Technologies) preserved DKD, 7 LD and 3 BKVN samples. Tissue processing and single cell isolation were performed following our published protocol. Briefly, the tissue samples were thawed at room temperature and dissociated to single cells using Liberase™ TL (research grade, Roche) at 37°C. Single cell samples were immediately transferred to the University of Michigan Advanced Genomics Core facility for further processing.

**Urine cell preparation for scRNAseq in patients with COVID-19**

Urine single cell preparation followed the protocol by Arazi et al. with various modifications. Briefly, COV-AKI urine was filtered through a 30 µm strainer into 50 ml tubes and centrifuged at 200 x g at 4°C for 10 minutes. After removing and storing the supernatant the cell pellet was washed once with 1 ml cold X-VIVO™10 medium (Cat#: 04-743Q, Lonza) and centrifuged in 1.5 ml tubes for 5 minutes at 200 x g at 4°C. The resulting cell pellet was then suspended in 50 µl DMEM/F12 medium supplemented with HEPES and 10% (v/v) FBS and 50,000 cells were immediately loaded onto the 10X Genomics Chromium single cell droplet-based RNAseq platform. Next generation sequencing was carried out on a NovaSeq6000 (Illumina) machine to generate about 200 million reads per sample. All procedures were performed following biosafety precautions as directed by the Centers for Disease Control and Prevention and the University of Michigan guidelines.
scRNAseq data generation and analysis

Individual cell barcoding, reverse RNA transcription, library generation and single cell sequencing using Illumina were all performed using the 10X Genomics protocol\textsuperscript{26}. The output from the sequencer was first processed by CellRanger, the proprietary 10X Chromium single cell gene expression analysis software (https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger). Data analyses were performed on the CellRanger output data files using the Seurat 3 R package (https://cran.r-project.org/web/packages/Seurat/index.html). As a quality control step, cells with less than 500 or greater than 5000 genes per cell were filtered out, with the lower cut off as a threshold for cell viability and the higher cut off to remove cell duplicates. This study used a cutoff of < 50% mitochondrial reads per cell as an additional threshold for cell viability. A combined analysis of the single cell datasets generated from the different sample sources (LD and DKD) was followed by downstream analysis steps including normalization, scaling based on sample mRNA count and mitochondrial RNA content, dimensionality reduction PCA (Principal component analysis) and UMAP (Uniform Manifold Approximation and Projection), sample integration using Harmony algorithm\textsuperscript{28}, standard unsupervised clustering, and the discovery of differentially expressed cell type specific markers\textsuperscript{29}. For cell type annotation we used publicly available resources including published literature, Kidney Interactive Transcriptome (http://humphreyslab.com/SingleCell/), Human Protein Atlas (HPA) (https://www.proteinatlas.org), the Epithelial Systems Biology Laboratory (ESBL) (https://hpcwebapps.cit.nih.gov/ESBL/Database/), scRNAseq data and Immgen (https://www.immgen.org/). The same analysis processes were used for the BKVN and COV-AKI samples except that no Harmony-based integration step was performed. Further functional analyses were focused on PTEC with ACE2 (ACE2+) and without ACE2 (ACE2-) mRNA expression in the different samples studied.
Identification of differentially expressed ACE2+ co-regulated gene signatures

Differentially expressed genes were identified between ACE2+ and ACE2- PTEC in DKD, BKVN, and COV-AKI samples using the FindAllMarkers Seurat function. For DKD and BKVN signatures, all genes with adjusted p-value < 0.05 and found in at least 10% of cells were selected. Due to the comparatively small number of cells in COV-AKI sample, ACE2+ co-regulated genes that passed a nominal uncorrected p-value threshold of 0.05 and were found in at least 10% of cells were selected (Supplementary Tables 2-5).

Overlap of differentially expressed ACE2+ co-regulated gene signatures with SARS-CoV-2 relevant gene sets:

SARS-CoV-2 relevant gene sets were compiled from multiple published sources: 1) Bojkova_proteome is a list of differentially expressed proteins in Caco-2 cell line following SARS-CoV-2 infection (p < 0.05 at any time point)30; 2) Bojkova_translatome is a list of proteins with differential translatome kinetics following SARS-CoV-2 infection (p < 0.05 at any time point)30; 3) Gordon_interactome is a set of host proteins identified as physically interacting with SARS-CoV-2 viral proteins in HEK-293T cells31; 4) Zhou_interactome is a literature-curated list of genes related to diverse coronaviruses17; 5) Blanco-Melo_A459 is a list of differentially expressed genes in response to SARS-CoV-2 infection in human alveolar adenocarcinoma cells32; 6) Blanco-Melo_NHBE is a list of differentially regulated genes in response to SARS-CoV-2 infection in normal human bronchial epithelial cells32. ACE2 was removed from ACE2+ co-expressed gene sets before computing overlaps. P-values were computed using the hypergeometric test with a count of 20,000 genes used as background.
**Functional network analysis**

To determine the biological processes and pathways in the ACE2+ differentially expressed gene sets, functional network clustering of the PTEC gene functional network from GIANT 2.0\textsuperscript{33, 34}. This network was generated through regularized Bayesian integration of 61,400 publicly available expression, physical interaction and other omics experiments to generate a fully connected weighted graph representing functional relationships in biological pathways in PTEC. Community clustering in the network was performed to identify tightly connected sets of genes using HumanBase.io module detection function\textsuperscript{35} (flatironinstitute.org/module). The network was clustered with each set of ACE2+ differentially expressed genes: DKD differentially expressed genes, BKVN differentially expressed genes, the overlap of these two gene sets (DKD\_BK\_overlap, Supplementary Table 6), and ACE2+ differentially expressed genes in PTEC from the COV-AKI sample (Supplementary Table 5). For DKD and BKVN signatures, all genes with adjusted p-value < 0.05 and found in at least 10% of cells were selected. Due to the comparatively small number of cells in COV-AKI sample, ACE2+ co-regulated genes that passed a nominal uncorrected p-value threshold of 0.05 and were found in at least 10% of cells were selected. The Gene Ontology enrichment outputs are Supplementary Tables 7-10 and can be interactively explored at hb.flatironinstitute.org/covid-kidney.

**Statistical analysis of clinical associations**

Spearman correlation were used to evaluate the association of the steady state average PTEC specific gene expression levels of ACE2 and age at time of biopsy for the 44 DKD samples. The Mann-Whitney non parametric test was applied to evaluate the gender and treatment difference of ACE2 expression in the DKD cohort.
**Study approval**

All studies of the DKD population were approved by the Institutional Review Board of the National Institute of Diabetes and Digestive under protocol 13-DK-N151. Transplant biopsies were obtained after review and approval of the Transplant Transcriptomic Atlas study under HUM00150968; DKD kidney biopsies gene expression analysis was performed under HUM00002468; Covid19 urine samples were obtained after review and approval under HUM00004729, all reviewed by the University of Michigan IRBMED.

**Data Access**

LD, BKVN and COV-AKI single cell data sets are searchable at NephroCell.miktmc.org. Owing to ethical considerations and privacy protection concerns, and to avoid identifying individual study participants in vulnerable populations, the Institutional Review Board of the National Institute of Diabetes and Digestive and Kidney Diseases has stipulated that individual-level gene expression and genotype data from the American Indian DKD study cannot be made publicly available.

A dynamic user-friendly interface at HumanBase (hb.flatironinstitute.org/covid-kidney) is available for researchers to explore the functional networks of gene expression signatures.

**Results**

We studied ACE2 expression and viral infection associated processes through scRNASeq and functional network analysis (Figure 1).

**Cohort studied for SARS-CoV-2 receptor expression and regulation:**

Kidney cell expression profiles were obtained from early DKD (n=44), LD (n=7), and BKVN (n=3) kidney biopsies and a urine sample from a COV-AKI patient. Baseline clinical characteristics for
the DKD cohort are provided in Table 1, for BKVN in Supplementary Table 1, for COV-AKI in Supplementary Figure 1, with further details in methods.

**scRNAseq based definition of SARS-CoV-2 receptor expression in kidney tissue**

First, we defined the expression pattern of ACE2 among the cell populations obtained by the scRNAseq analysis in the 44 DKD and 7 LD biopsies. In a combined analysis of the two data sets with a resolution (clustering granularity) of 0.6, 17 cell clusters were defined (Figures 2A and 2B UMAP plots) with 15 clusters populated by cells from both sample sources, indicating presence of cells with comparable identity in both sample sources (Supplementary Figure 3). The 15 cell clusters covered the entire spectrum of kidney cell types found along the nephron and tissue resident immune cells.

Next, to understand pathways activated in viral infection DKD biopsy PTEC DKD were compared to the corresponding cells in BKVN with BKVN scRNAseq signatures obtained before, at and after peak of BK viremia. The combined analysis of the three BKVN samples resulted in 10 clusters from 15135 cells. Unlike LD, we observed prominent immune cell clusters in the BKVN data set (Figure 2C UMAP plot) with cells assigned to PTEC clusters contributing only ~10% to the total cell population.

The violin plots in Figure 2 show a highly restricted, cell-type specific mRNA expression of ACE2 in the PTEC cell cluster consistently across LD, DKD and BKVN samples (Figure 2). Expression level differences of ACE2, TMPRSS2 and CTSL before, at and after peak are shown in supplementary Figure 2.
Defining ACE2+ expression signatures shared with SARS-CoV-2 host factors

To define the functional context of ACE2 expression signatures in kidney disease, we first identified gene expression programs specifically co-activated with the SARS-CoV-2 receptor expression. To this end the expression profiles of ACE2+ PTEC were compared to the ACE2-PTEC cell population in both DKD and BKVN. The resulting gene expression signatures (with p-value adjusted Bonferroni cutoff of 0.05 and expressed in at least 10% of cells) identified sets of genes whose expression is specifically induced in ACE2+ cells in LD (Supplementary Table 2), DKD (Supplementary Table 3) and BKVN (Supplementary Table 4). Critically, we then focused on identifying shared gene expression programs in ACE2+ PTEC in both DKD and BKVN samples (Supplementary Table 6).

We also analyzed the overlap of the ACE2+ signatures with SARS-CoV-2-implicated genes and ACE2 signature genes from multiple publicly available, independent studies and found significant shared components across SARS-CoV-2 datasets (Table 2). Specifically, genes associated with ACE2+ expression in DKD were also significantly enriched among host proteins identified to interact with SARS-CoV-2 proteins, host genes upregulated in SARS-CoV-2-infected cells, and host proteins previously reported to interact with multiple coronaviruses. Such cellular proteins could represent cellular targets of viral infection (permissive factors), cellular responses to viral infection (restriction factors), or a mix. A robust overlap between the ACE2+ DKD and BKVN expression signatures and both SARS-CoV-2-related interactome and differentially expressed gene sets were observed, supporting the activation of a related gene program in SARS-CoV-2 infection.
**Functionally characterizing shared proximal tubule ACE2 co-regulated gene programs**

To functionally characterize these ACE2+ expression signatures, we projected these genes onto networks that represent pathways active in normal PTEC. Intuitively, the HumanBase functional network is constructed by probabilistically integrating a large compendium of thousands of public omics datasets (including expression, protein–protein interaction and motifs) to represent how likely it is that two genes act together in biological processes that function in PTEC\(^{33, 34}\). We clustered genes differentially expressed in ACE2+ vs ACE2- PTEC of the DKD kidneys in this functional network (Figure 3A, Supplementary Table 7). The resulting modules contain key processes of tubular function and failure, including renal tubule development and ion transport, general cellular processes such as cell cycle control and translational regulation, as well as disease signals including inflammatory pathway activation (IL-2 production, JAK/STAT signaling, and Wnt signaling). Importantly, we also identify key processes important to both immune response and viral entry and life cycle, including viral entry, viral gene transcription and translation, endomembrane organization, endosomal transport, and RNA metabolism (Figure 3A, Supplementary Table 7). Immune processes were significantly enriched across multiple modules and encompassed processes including innate immune responses (secretion and response to interferons alpha, beta, and gamma), cytokine production, macrophage activation. The robust activation of immune-related functions links ACE2+ (vs ACE2-) PTEC to these key elements of DKD pathophysiology. Focusing on the program involved in response to viral infection, we find that indeed, many of these modules are also shared with ACE2+ (vs ACE2-) PTEC in BKVN (Supplementary Figure 4, Supplementary Table 8). The congruence between the DKD and BKVN signals at both gene and functional process levels supports the robustness of this ACE2+ molecular signature and suggests that at least some of the programs related to establishment of productive viral infection and to innate immune responses seen in BKVN are also present in the DKD ACE2+ PTEC even without an active viral infection (Figure 3B, Supplementary Table 9). We hypothesize that some of these
expression programs could be permissive to viral infection, along with the upregulation of ACE2 receptor.

**Association of ACE2 mRNA expression in PTEC with clinical features**

We then assessed differential ACE2 expression in PTEC in DKD compared to LD and found a significant overexpression of ACE2 steady state mRNA levels in DKD (adjusted p-value < 11.79 e-23; in cells with normalized read count >= 1.25) (Figure 3A). We focused on cells with ACE2 expression above 1.25 to exclude the ambient ACE2 mRNA expression (~ 1 normalized read count) observed in most cell types as seen in Figure 2. Importantly, proteases thought to be permissive for coronavirus infection including ANPEP, BSG, CTSL, DPP4, ENPEP, FURIN and TMPRSS2 appear to co-express with ACE2 in PTEC (Figure 4A, B).

The expression levels of ACE2 transcripts in PTEC did not show a significant association with baseline DKD participant characteristics (age/gender, Figure 4C and D) or treatment exposures to renin angiotensin aldosterone system (RAAS) inhibitors (Figure 4E-G).

**Transcriptional profile in urine derived proximal tubular epithelial cells in COVID-19**

Finally, the PTEC cellular response in COVID-19 was assessed using urinary scRNAseq analysis constrained to PTEC gene expression clusters. Single cell data analysis of the cells isolated from the urine of a SARS-CoV-2 infected patient yielded 273 cells that passed our quality control threshold. Unsupervised clustering of these cells produced three kidney cell clusters, including PTEC, and two immune cell clusters (Figure 5A and B). 42% of the cells in the PTEC cluster showed ACE2 expression. In addition to ACE2, proteases including ANPEP, BSG, CTSL, DPP4, ENPEP, FURIN and TMPRSS2 were expressed in PTEC from the COV-AKI sample (Figure 5C, Supplementary Table 5). BSG and ANPEP were found to be expressed in all proximal cells identified (Figure 5C). Supplementary Table 11 provides the PTEC markers identified in COV-AKI. Notably, genes nominally upregulated in ACE2+ vs ACE2- COV-AKI
urine derived PTEC (Supplementary Table 10) significantly overlap the ACE2+ DKD and BKVN kidney biopsy disease signature (p-value =7.7x10-7). Examining the ACE2+ co-expressed functional network in the COV-AKI PTEC cell population identified a signature enriched for viral processes (Supplementary Figure 5).

Discussion

With the pandemic spread of COVID-19 and the increased COVID-19 morbidity and mortality in patients with diabetes and kidney disease it is imperative to define the underlying mechanism of the excessive risk in patients with diabetes and rapidly develop protective strategies for risk reduction. Autopsy studies of patients with COVID-19 are consistent with direct viral infection of the kidney and that even within the kidney\textsuperscript{12, 36}, expression of molecules that mediate viral entry may be cell type-specific\textsuperscript{12}. Susceptibility to viral entry may be enhanced by DKD associated changes in SARS-CoV-2 receptor expression in kidney tissue.

Our study set out to establish the cell population within the kidney expressing ACE2 and associated SARS-CoV-2 receptors mRNA in healthy individuals and in patients with DKD and viral associated kidney diseases. ACE2 intra-renal gene expression was localized narrowly to PTEC in surveillance kidney biopsies obtained from LD and in research kidney biopsies from patients with DKD. PTEC are the prime localization of SARS-CoV-2 infection in the kidney\textsuperscript{36}, supporting the link between receptor expression and cellular toxicity in kidneys, as well. In contrast to findings in other cell types, little co-expression of ACE2 with TMPRSS2 co-receptor used for SARS-CoV-2 viral entry was observed, suggesting other cell surface proteases may be associated with viral infection and uptake in the kidney\textsuperscript{12, 36}.

Despite the fact that none of the diabetic patients in this study were infected with SARS-CoV-2 at the time of biopsy, our pathway analysis showed that processes including infection, protein
processing and antigen presentation were enriched in high versus low ACE2 expressing PTEC in DKD. The results from this analysis suggest that ACE2 is co-expressed with a set of genes also having functions in viral entry, establishment of viral replication, and innate immunity.

A key finding of our study is that ACE2+ proximal tubule cells of patients with DKD had significantly higher ACE2 mRNA levels than those of LDs. These higher ACE2 levels may make these cells more permissive to initial viral entry and subsequent increased cellular cytotoxicity in DKD, especially as viral replication and processing machinery also appear to be upregulated in the ACE2+ PTEC. This includes genes involved in endomembrane transport and lysosomal compartments, which we hypothesize could facilitate viral entry and exit, and proteins that interact with viral transcription and translation. It is intriguing that a number of the ACE2+ signature genes interact with predicted or demonstrated SARS-CoV-2 anti-host defense response proteins\textsuperscript{31}, including M (suppresses type1 interferon response and refine cytokine responses)\textsuperscript{37, 38}, N (inhibit autophagy and apoptosis of infected cells)\textsuperscript{39}, Orf3a (modulates IL-1B)\textsuperscript{40}, and Nsp6 (inhibits phagosome expansion and targeting of viral products to lysosome)\textsuperscript{41}. These interactions may simply reflect an element of the functional host response program to cellular stress in DKD. However, if these interactions also involve negative functional regulation of host-defense proteins, ACE2+ proximal tubule cells could be impaired in their ability to mount an appropriate host response, providing one explanation of their observed vulnerability to infection to SARS-CoV-2.

The inflammatory signals observed in the DKD and BKVN PTEC are expected in the context of chronic kidney disease as reported previously by studies in this and other DKD patient populations\textsuperscript{42, 43}. However, these data also could support the hypothesis that this inflammatory state could be maladaptive for a response to viral infection. Recent work on lung epithelial cells infected with SARS-CoV-2 has shown a blunted interferon response despite robust cytokine
production still being present\textsuperscript{32}; a potential hallmark of SARS-CoV-2 pathology. If a comparably asymmetric inflammatory response is already present in DKD kidney tissue, viral infection could amplify the pre-existing inappropriate stress response, increasing cytopathology and enabling further viral propagation. If ACE2 upregulation is indeed correlated with a cellular inflammatory state that permits enhanced viral pathogenesis and replication in multiple organ systems, this linkage could provide an evolutionary driving factor to select ACE2 for viral entry.

The PTEC specific ACE2\textsuperscript{+} networks observed in DKD, BKVN and COV-AKI allow for a wide range of exploratory studies to generate starting points for functional evaluation. Given the urgent need for novel therapeutic options, intersecting the ACE2 regulatory networks with the complex landscape of interactions between drugs and the gene signature defined here can help to identify therapeutic targets for further study. We identified ACE2\textsuperscript{+} co-regulated genes annotated in DrugBank\textsuperscript{44} as targets of medications regularly taken by patients with DKD. For example, insulin (INSR, in M5 in Figure 3B) and the IGF-1 receptor (IGF1R, in M3 in Figure 3B) are overexpressed in ACE2\textsuperscript{+} PTEC as is DPP4 (in M5 in Figure 3B), a protease associated with entry of other coronaviruses and a key factor in glucose metabolism, targeted by DPP4 inhibitors already in clinical use. These interactions suggest that the ACE2\textsuperscript{+} co-expression signature may be influenced by a number of commonly used drugs and can form the starting points for further mechanistic and epidemiological studies.

As data emerged on the role of ACE2 as a receptor for SARS-CoV-2, significant concerns were raised about risks of RAAS inhibitors frequently prescribed in patients with diabetes and chronic kidney diseases. Based on animal models RAAS inhibitors were predicted to increase ACE2 levels raising the potential concern for higher COVID-19 morbidity and mortality\textsuperscript{22, 45, 46}. However, our findings in the DKD cohort do not support the animal model data and do not provide any evidence of transcriptional upregulation of ACE2 mRNAs in PTEC by RAAS.
inhibition. In parallel, a series of case control, database and Electronic Health Record studies did not find an association of RAAS inhibitors with poor outcomes in three independent cohorts of patients with COVID19 and RAAS exposures\textsuperscript{47-49}, consistent with our scRNAseq studies.

ACE2 gene expression studies of kidney tissue to date have been inconclusive concerning regulation in kidney disease \textsuperscript{50-53}. A strength of our study is the definition of the ACE2 expression to the cellular context of PTEC in distinct infectious and non-infectious kidney diseases. As PTEC mass is lost in kidney disease, the cellular resolution of the receptor expression achieved by scRNAseq is critical to detect the increased receptor load per cell versus the average receptor expression level across all cells obtained in analyses of bulk renal biopsies to date. The PTEC dominant ACE2 scRNAseq expression data presented here match the immunohistochemical tissue distribution pattern in the human protein atlas\textsuperscript{54} (https://www.proteinatlas.org/, Supplementary Figure 6), but given the limited sensitivity of scRNAseq, do not rule out low level expression in other renal cell types.

An important element of this study is the confirmation of ACE2 and associated genes and modules in PTEC derived from a COV-AKI sample. The association of ACE2 expression with proteins involved in innate immune response and viral entry and replication provides an intriguing hypothesis for how SARS-CoV-2 can cause increased cytopathology in ACE2+ cells. However, these findings have to be considered exploratory, as they are based on a small population of PTEC which have lost their structural kidney tissue context in a single patient with COV-AKI. Confirmation of these findings in a larger cohort of COVID-19 patients and ideally from kidney tissue is needed.

Our study is also limited by focusing on mRNA levels, which only capture one of several levels of regulation of ACE2 function. For this reason, our analyses focused on the definition of co-
regulation networks to capture the change in transcriptional programs as a functional read out and not protein activity directly. However, as noted above, our mRNA findings for ACE2 are consistent with protein expression specificity from the Human Protein Atlas.

Based on these considerations, further experimental validation of the predicted regulation and therapeutic interventions are critical next steps. Experimental model systems of kidney SARS-CoV-2 toxicity are currently under development, including human kidney organoid systems which do express ACE2 in the correct functional context\(^5\) (nephrocell.miktmc.org,).

In summary, our work identifies the regulation and associated cellular machinery of ACE2 and associated SARS-CoV-2 co-receptors in PTECs in kidney health, metabolic and viral disease. The SARS-CoV-2 receptor associated networks are now available to seed further research into urgently needed therapeutic strategies for COVID-19.
Acknowledgements:

We thank Ms. Lois Jones, RN, Mr. Enrique Diaz, RN, Ms. Bernadine Waseta, and Ms. Camille Waseta for performing the studies in the diabetes cohort and the University of Michigan Advanced Genomics Core for providing expert technical assistance with single cell processing and sequencing.

This work was supported in part by the Intramural Research Program at the National Institute of Diabetes and Digestive and Kidney Diseases (DK069062) to HCL and RGN and (DK083912, DK082841, DK020572, DK092926) to RGK, by the extramural research program of the National Institute of Diabetes and Digestive and Kidney Diseases R24 DK082841 ‘Integrated Systems Biology Approach to Diabetic Microvascular Complications’ and P30 DK081943 ‘University of Michigan O'Brien Kidney Translational Core Center’ to MK, via ‘Kidney Precision Medicine Project’ (KPMP, funded by the following grants from the NIDDK: U2C DK114886, UH3DK114861, UH3DK114866, UH3DK114870, UH3DK114908, UH3DK114915, UH3DK114926, UH3DK114907, UH3DK114920, UH3DK114923, UH3DK114933, and UH3DK114937, with U2C DK114886 and UH3 DK114907) to MK and OGT, via the Chan Zuckerberg Initiative ‘Human Cell Atlas Kidney Seed Network’ to MK and OGT and by JDRF 5-COE-2019-861-S-B ‘JDRF and M-Diabetes Center of Excellence at the University of Michigan’ to MK. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

For the Kidney Precision Medicine Project

American Association of Kidney Patients, Tampa, FL: Richard Knight
Beth Israel Deaconess, Boston, MA: Stewart Lecker, Isaac Stillman
Boston University, Boston, MA: Sushrut Waikar,
Brigham & Women's Hospital, Boston, MA: Gearoid Mcmahon, Astrid Weins, Samuel Short
Broad Institute, Cambridge, MA: Nir Hacohen, Paul Hoover
Case Western Reserve, Cleveland, OH: Mark Aulisio

Cleveland Clinic, Cleveland, OH: Leslie Cooperman, Leal Herlitz, John O'toole, Emilio Poggio, John Sedor, Stacey Jolly


Duke University, Durham, NC: Laura Barisoni

European Molecular Biology Laboratory, Heidelberg, Germany: Theodore Alexandrov

Indiana University, Indianapolis, IN: Tarek Ashkar, Daria Barwinska, Pierre Dagher, Kenneth Dunn, Michael Eadon, Michael Ferkowicz, Katherine Kelly, Timothy Sutton, Seth Winfree

John Hopkins University, Baltimore, MD: Steven Menez, Chirag Parikh, Avi Rosenberg, Pam Villalobos, Rubab Malik, Derek Fine, Mohammed Atta, Jose Manuel Monroy Trujillo, Joslin Diabetes Center, Boston, MA: Alison Slack, Sylvia Rosas, Mark Williams

Mount Sinai, New York, NY: Evren Azeloglu, Cijang (John) He, Ravi Iyengar, Jens Hansen

Ohio State University, Columbus, OH: Samir Parikh, Brad Rovin

Pacific Northwest National Laboratories, Richland, WA: Chris Anderton, Ljiljana Pasa-Tolic, Dusan Velickovic, Jessica Lukowski

Parkland Center for Clinical Innovation, Dallas, TX: George (Holt) Oliver

Patient Partners: Joseph Ardayfio, Jack Bebiak, Keith Brown, Taneisha Campbell, Catherine Campbell, Lynda Hayashi, Nichole Jefferson, Robert Koewler, Glenda Roberts, John Saul, Anna Shpigel, Edith Christine Stutzke, Lorenda Wright, Leslie Miegs, Roy Pinkeney

Princeton University, Princeton, NJ: Rachel Sealfon, Olga Troyanskaya

Providence Medical Research Center, Providence Health Care, Spokane, WA: Katherine Tuttle

Stanford University, Palo Alto, CA: Dejan Dobi, Yury Goltsev

University of California San Diego, La Jolla, CA: Blue Lake, Kun Zhang

University of California San Francisco, San Francisco, CA: Maria Joanes, Zoltan Laszik, Andrew
Schroeder, Minnie Sarwal, Tara Sigdel

*University of Michigan, Ann Arbor, MI*: Ulysses Balis, Victoria Blanc, Oliver He, Jeffrey Hodgin, Matthias Kretzler, Laura Mariani, Rajasree Menon, Edgar Otto, Jennifer Schaub, Becky Steck, Chrysta Lienczewski, Sean Eddy

*University of Pittsburgh, Pittsburgh, PA*: Michele Elder, Daniel Hall, John Kellum, Mary Kruth, Raghav Murugan, Paul Palevsky, Parmjeet Randhawa, Matthew Rosengart, Sunny Sims-Lucas, Mary Stefanick, Stacy Stull, Mitchell Tublin

*University of Washington, Seattle, WA*: Charles Alpers, Ian de Boer, Ashveena Dighe, Jonathan Himmelfarb, Robyn Mcclelland, Sean Mooney, Stuart Shankland, Kayleen Williams, Kristina Blank, Jonas Carson, Frederick Dowd, Zach Drager, Christopher Park

*UT Health San Antonio, Center for Renal Precision Medicine, San Antonio, TX*: Kumar Sharma, Guanshi Zhang, Shweta Bansal, Manjeri Venkatachalam

*UT Southwestern Medical Center, Dallas, TX*: Asra Kermani, Simon Lee, Christopher Lu, Tyler Miller, Orson Moe, Harold Park, Kamalanathan Sambandam, Francisco Sanchez, Jose Torrealba, Toto Robert, Miguel Vazquez, Nancy Wang

*Washington University in St. Louis, St. Louis, MO*: Joe Gaut, Sanjay Jain, Anitha Vijayan

*Yale University, New Haven, CT*: Randy Luciano, Dennis Moledina, Ugwuowo Ugochukwu, Francis Perry Wilson, Sandy Alfano
To identify SARS-CoV-2 receptor networks in diabetic kidney disease (DKD), BK-Virus nephropathy (BKVN) and COVID-19 associated acute kidney injury (COV-AKI), we used scRNAseq profiling and integrative network analyses.

(A) Biopsy samples from DKD, BKVN, and living donors (LD) were processed for scRNAseq profiling and cell-type clusters were identified from each sample. (B) For each dataset, we identified genes that are differentially expressed in ACE2+ compared to ACE2- proximal cells (PTEC). We compared DKD and BKVN ACE2+ co-expression sets to identify a disease-specific signature, and we also compared each signature to published SARS-CoV-2-relevant gene sets. (C) In parallel, we evaluated the association of ACE2 expression levels with clinical characteristics, including exposure to RAAS blockers and ACE inhibitors in DKD. (D) For each ACE2+ co-expression signature, we identified modules in a proximal tubule-specific functional network and found modules enriched with biological processes relevant to viral infection. (E) We compared expression of ACE2 and key proteases between LD and DKD PTEC. (F) Finally, we explored this ACE2+ signature in PTEC cells in BKVN kidney tissue and a COV-AKI urine sample.
Figure 2. Unsupervised clustering of cells from living donor, DKD and BK re-activated kidney biopsies. UMAP plots showing the distributions in unsupervised clustering of (A) 13,055 living donor kidney cells into 15 clusters, (B) 85,872 cells from DKD into 17 clusters, and (C) 15,135 cells from BKVN into 10 clusters. The specific expression of ACE2 in proximal tubular epithelial cells (PTEC) are shown to the right of the corresponding UMAP plots.
Figure 3. Functional summary of ACE2+ expression signature
ACE2+ co-expression signatures were used for community clustering in a PTEC-specific functional network to identify enriched processes and pathways specific to (a) ACE2+ co-expression signatures from DKD biopsy samples and (b) ACE2+ co-expressed genes shared between DKD and BK virus reactivation transplant biopsy samples.
Figure 4. (A and B) ACE2 and related proteases expression in proximal cells

Dot plot showing the expression of proteases in DKD (in red) and LD (in blue) PTEC. Grey color denotes moderate expression in the cell and the size of the dot indicates the percentage of cells expressing the gene. Proteases were grouped into two based on the overall expression pattern in proximal cells (expression in < or >= 30% of the cells). Panel A depicts moderate (<30% of cells) and panel B widely (>=30% of cells) expressed proteases with maximal expression of ACE2 and ANPEP in DKD compared to living donor in contrast to FURIN, TMPRSS2, BSG, CTSL, DPP4 and ENPEP with higher expression in LD.

(C-G) Clinical Associations of ACE expression in proximal tubular epithelial cells in DKD.

Proximal Tubular expression of ACE2 in Female vs. Male (C) correlation with Age (D), and in relationship to exposure to any Renin-Angiotensin-Aldosterone-System Blockers (E); Angiotensin Receptor Blockers (F); Angiotensin Converting Enzyme Inhibitors(G).
Figure 5. Single cell analysis of cells isolated from the urine of patient with covid19 related acute kidney injury. (A) Umap plot showing the 5 clusters identified from 273 cells. (B) Heatmap showing enriched markers identified in the 5 clusters. (C) Expression pattern of SARS-CoV-2 associated proteases, ACE2, ANPEP, BSG, CTSL, ENPEP, FURIN and TMPRSS2. All rights reserved. No reuse allowed without permission.
Supplementary Materials:

Supplementary Figures

Supplementary Figure 1: Serum creatinine (mg/dl) time course across COVID-19 related illness and hospitalization in a kidney pancreas transplant recipient.

Supplementary Figure 2: Violin plot showing the expression of ACE2, TMPRSS2 and CTSL in the three BKVN samples.

Supplementary Figure 3: UMAP plot showing the distribution of the cells in the 17 clusters resulting from the unsupervised clustering of the scRNA-seq data from DKD (red) and LD biopsies (blue).

Supplementary Figure 4: Functional summary of ACE2+ expression signature for BKVN.

Supplementary Figure 5: Functional summary of ACE2+ expression signature from COVID-19 patient.

Supplementary Figure 6: ACE2 immunostaining in human kidney tissue.

Supplementary Tables

Supplementary Table 1: Clinical characteristics of BKVN cohort.

Supplementary Table 2: Differential expression gene list between ACE2+ and ACE2- PTEC in LD.

Supplementary Table 3: Differential expression gene list between ACE2+ and ACE2- PTEC in DKD.

Supplementary Table 4: Differential expression gene list between ACE2+ and ACE2- PTEC in BKVN.

Supplementary Table 5: Differential expression gene list between ACE2+ and ACE2- PTEC in COV-AKI.

Supplementary Table 6: Overlap of differentially expressed genes in DKD and BKVN.
Supplementary Table 7: The Gene Ontology enrichments for network modules of ACE2+ differentially expressed genes in DKD.

Supplementary Table 8: The Gene Ontology enrichments for network modules of ACE2+ differentially expressed genes in BKVN.

Supplementary Table 9: The Gene Ontology enrichments outputs for the network modules of ACE2+ differentially expressed genes overlapping in DKD and BKVN.

Supplementary Table 10: The Gene Ontology enrichments for network modules of ACE2+ differentially expressed genes in COV-AKI.

Supplementary Table 11: PTEC-specific markers in the COV-AKI sample.

References


