Electronic Supplemental Material

Supplemental Materials

Urine proteomics LC-MS/MS details.

Follow-up study cohort blood and urine cathepsin D ELISA measurements.

Mouse proximal tubular epithelial cell culture and Akita mouse cathepsin D enzyme activity studies: reagents, cell culture, animal models, cathepsin D activity, total protein extractions from the mice kidney cortex, double label fluorescent immunohistochemistry, assessment of IL-1, IL-6, and KIM-1 levels, statistical analysis.

References.

Tables

Table S1. Urine peptides measured in the JDRF Biomarkers Consortium across discovery and validation sets.

Table S2. Endogenous peptides monitored in the targeted proteomics assay.

Table S3. Distribution of JDRF cohorts by laboratory processing Subsets.

Table S4. Baseline characteristics of participants from cohorts used for follow-up cathepsin D serum and urine measurements.

Table S5. Associations of 38 urine peptides with rapid eGFR decline in the JDRF Biomarkers Consortium discovery set.

Table S6. Associations of 20 urine proteins with rapid eGFR decline in the JDRF Biomarkers Consortium discovery set.

Table S7. Associations of urine cathepsin D peptides and protein with rapid eGFR decline in the JDRF Biomarkers Consortium validation set.

Table S8. Associations of cathepsin D protein with rapid eGFR decline in the JDRF Biomarkers Consortium discovery set, sequentially adjusted for full covariate model variables.

Table S9. Associations of cathepsin D protein with rapid eGFR decline in the JDRF Biomarkers Consortium validation set, sequentially adjusted for full covariate model variables.

Table S10. Associations of cathepsin D peptides and protein with rapid eGFR decline in the discovery set, stratified by baseline UACR and eGFR.

Table S11. Associations of cathepsin D peptides and protein with rapid eGFR decline in the validation set, stratified by baseline UACR and eGFR.

Table S12. Associations between serum cathepsin D and histopathological features in the Boston Kidney Biopsy Cohort.

Table S13. Associations between urine creatinine-normalized cathepsin D and serum cathepsin D in follow-up analyses.

Table S14. Baseline characteristics of Pima Indian study cohort subjects with type 2 diabetes.

Table S15. Characteristics of Kidney Precision Medicine Project (KPMP) subjects with diabetes and chronic kidney disease used for cathepsin D single cell RNAseq kidney tissue measurements.

Figures

Figure S1. Cathepsin D activity in the media of MCT cells transfected with CatD-siRNA or Ctr-siRNA after 24 hours of treatment.

Figure S2. Cathepsin D follow-up cohort measurements.

Figure S3. Violin plots depicting cathepsin D (CTSD) mRNA expression levels across kidney cell types derived using single-cell RNA sequencing from Kidney Precision Medicine Project (KPMP) biopsy samples.

Figure S4. Cathepsin D, Akita mouse model studies

Figure S5. MCT cell inflammatory/injury marker studies.

Supplemental Materials

Urine proteomics LC-MS/MS details

Each aliquot of urine (0.2 mL) was supplemented with 1 µg of α-amylase from *Aspergillus oryzea* (10 µL of 0.1 µg/µL in phosphate buffered saline). Proteins were then precipitated by adding 0.8 mL ice-cold acidified acetone:methanol (50%:50% v:v; acidified acetone is made with 120 mL acetone and 24 µL of 5 M HCl in water), briefly vortexing, and incubating for exactly 24 h at -20°C. Proteins were pelleted for 15 min at 4°C and 15,000x*g* and then dried, first at room temperature and atmospheric pressure, and then under vacuum with a centrifugal evaporator (15 min at room temperature). The pellets were reconstituted in 100 µL freshly prepared 0.5% sodium deoxycholate in 50 mM ammonium bicarbonate in a Thermomixer (1000 rpm, 2 h at room temperature). The proteins were reduced in 5 mM dithiothreitol (1 µL of 0.5 M dithiothreitol) for 30 min at 60°C, alkylated in 15 mM iodoacetamide (3 µL of freshly prepared 0.5 M iodoacetamide) for 30 min at room temperature in the dark, and then digested with 1 µg trypsin (5 µL of 0.2 µg/µL Worthington Biomedical trypsin, Lakewood, NJ) for 2 h at 37°C with shaking (Thermomixer, 1400 rpm). Deoxycholate was precipitated by the addition of HCl to a final concentration of 200 mM (4.4 µL of 5 M HCl) and incubation at 4°C for 15 min and was pelleted by centrifuging at 4°C and 15,000x*g* for 20 min. The supernatant was removed, centrifuged again, and then desalted using solid phase extraction (HLB Prime µElute plate, Waters, Milford, MA) on a positive pressure manifold (Biotage). To desalt, samples were loaded onto the extraction plate, desalted 3 times with 100 µL 0.1% trifluoroacetic acid, followed by 100 µl 0.5% formic acid. Peptides were eluted 2 times with 100 µL of 80% acetonitrile/0.1% formic acid and dried in a centrifugal vacuum concentrator. The peptides were reconstituted in 30 µL 5% acetonitrile in 0.1% formic acid with vortexing (2 h at room temperature), spiked with internal standard peptides (30 µL internal standards in 0.1% formic acid), vortexed briefly, centrifuged for 10 min at room temperature (15,000x*g*), and then transferred to an injection vial (20 µL). Peptides of interest (see table **) were quantified using LC-MS/MS on a Thermo Orbitrap mass spectrometer connected to an Easy-nLC liquid chromatography system (Themo Scientific, Waltham, MA). Peptides were monitored by selecting their precursor ions in the isolation quadrupole (selection window 1.5 Da) and collecting full scan MS/MS after HCD fragmentation (NCE 27) in the Orbitrap analyzer using lowest resolution (17,500) setting to increase scan rate. Scheduled acquisition with five-minute acquisition windows were set up for each peptide precursor using Skyline software (MacCoss Laboratory, University of

Washington) allowing a maximum of 30 concurrent PRM experiments at any given time. Acquired data was processed in Skyline and automated integration was manually checked for each peptide chromatogram. Identity of the chromatographic peaks was ascertained by matching the PRM MS/MS spectrum to the spectra from a shotgun source dataset (dot product > 0.8 and mass precision < 5 ppm).

Follow-up study cohort details and blood and urine cathepsin D ELISA measurements

Urine and serum cathepsin D measurements were performed in 96 adults with diabetes from the Boston Kidney Biopsy Cohort (n=56), the Healthy Kidney Study (HKS, n = 19), and a pilot Vitamin D trial in subjects with type 2 diabetes and microalbuminuria (n=21). The HKS is a cohort study based in Seattle beginning in 2004 following adults over a range of kidney functions with the goal of determining risk factors for kidney disease progression.[1] The pilot Vitamin D trial randomized adults with type 2 diabetes, eGFR $\geq$ 60ml/min/1.73m^2, and UACR 30-300mg/g to 12 months of vitamin D_3 2000 IU/day versus placebo.

Urine and serum cathepsin D measurements were performed with a commercially available Abcam enzyme-linked immunosorbent assay (ab119586; Cambridge, Massachusetts, USA). Urine samples were diluted 20-fold and serum samples were diluted 80-fold. For urine measurements, the mean intra-and inter-assay coefficients of variation (CV) were 17.2% and 15.0%, respectively. For serum measurements, the mean intra- and inter-assay CVs were 16.4% and 16.4%, respectively. Correlations between serum and urine cathepsin D were assessed using the Pearson method. The association between serum and urine cathepsin D was additionally evaluated using linear regression models adjusting for sex, age, eGFR, and log(UACR).

Mouse proximal tubular epithelial cell culture and Akita mouse cathepsin D enzyme activity studies

Reagents. Unless otherwise stated, all chemicals were of reagent-grade quality and were purchased from Sigma Chemical Co. (St. Louis, MO). Cathepsin D substrate was purchased from ENZO (Farmingdale, NY). Rabbit polyclonal anti-cathepsin D, mouse monoclonal anti-LAMP1 antibody and mouse monoclonal anti-aquaporin 1 antibody were purchased from Abcam (ab75852, ab25630 and ab9566, respectively). Alexa Fluor 488 and 594 antibodies were purchased from ThermoFisher Scientific (Waltham, MA). VECTASHIELD® Antifade Mounting Medium with DAPI (Vector Laboratories, Burlingame, CA). siRNA and transfection

reagents DharmaFECT 1 were purchased from GE Healthcare Dharmacon (Lafayette, CO). IL1, IL6 and Kim1 ELISA kits were purchased from BioVision (Milpitas, CA). Pierce BCA protein assay kit from ThermoFisher Scientific (Waltham, MA). Other reagents for immunohistochemistry have been purchased from Agilent (Santa Clara, CA).

Cell culture. For the experiment we used MCT (murine kidney proximal tubular epithelial) cells which in culture express in vivo characteristics of proximal tubular epithelial cells.[2] MCT cells were maintained in Dulbecco's modified Eagle's medium containing 5 mM glucose (DMEM), 100 units/ml penicillin, 100 µg/ml streptomycin and 2 mm glutamine, plus 10% fetal bovine serum in a 5% CO2atmosphere at 37°C. Cells were used between passages 6 and 15. Before addition of agents and transfections, the medium was changed overnight to DMEM without FBS and antibiotics. To induce cathepsin D activity, cells were treated with 25 mM glucose, or AGE-modified BSA (100–200 µg/mL) or control BSA (100 µg/mL) with 5 mM glucose for 24 hours. AGE-BSA concentrations were chosen based on the previously published study.[3] For the following experiments 100 ug/mL of AGE-BSA were used as a minimal concentration which significantly induced CatD activity.

For the RNA interference experiments, a SMART-pool consisting of four short or small interfering RNA (siRNA) duplexes specific for mouse CatD was obtained from Dharmacon. The SMARTpool of siRNAs was introduced into the cells by transfection using DharmaFECT 1 according to the manufacture protocols. Scrambled (nontargeting) siRNAs served as controls to validate the specificity of the siRNAs. Scrambled control and siRNA for CatD were used at a concentration of 125 pmol/mL.

Animal models. Samples for the immunofluorescent staining and CatD activity were from a previously published study.[4] To generate the diabetic kidney disease mouse model, we bred C57BL/6-Ins2^{Akita} male mice to female DBA/2 mice to produce F1(DBA/2 X C57BL/6)-Ins2^{Akita} diabetic mice (F1 Akita mice). As a control we used offspring of a C57BL/6 male mated to female DBA/2 mice (F1[DBA/2 X C57BL/6] control mice). All mice were purchased from Jackson Laboratories. 16-week-old mice were used. Mice were fasted for 3 h prior to spontaneous urine or blood collection for Cat D activity analysis. Kidneys were fixed

in 4% formalin and embedded in paraffin for the next immunofluorescence microscopy. Part of kidney were frozen in the liquid nitrogen and stored at -80 ^{0}C and were used for the CatD activity.

Cathepsin D activity was measured by monitoring hydrolysis of the CatD-specific fluorogenic substrate, MCA-Gly-Lys-Pro-Ile-Leu-Phe-Phe-Arg-Leu-Lys(Dnp)-D-Arg-NH_2 in the in the 0.1 M acetate buffer with 10 mM DTT, pH=4.5.[5] Fluorescence of free aminomethyl coumarin was determined as a kinetic interval assay, with readings taken every 5 minutes for 3 hours at 30 ^{0}C by excitation at 328 nm and emission at 460 nm using Tecan Spark 10M fluorescent plate reader. Representative results were taken after one hour of reaction. Pepstatin A were used as controls to confirm the detection of specific activities. Absorbance in samples containing pepstatin A was subtracted from that in samples lacking pepstatin A.

Double-label fluorescent immunohistochemistry was assessed as described, with minor modifications.[6] The sections of kidney were deparaffinized in xylene, hydrated in decreasing concentrations of ethanol, and washed in water. For double fluorescent immunohistochemistry of the kidney cortex, anti-cathepsin D antibody was applied in combination with anti-aquaporin 1 antibody, both diluted 1:100 in 1% BSA in Tris-buffered solution (TBS) and applied overnight at 4 ^{0}C in the humidity chamber. After incubation with primary antibodies, all the sections were washed in PBS containing 0.01% Triton X-100. A cocktail of secondary Alexa Fluor 488 and 594 antibodies (both in 1:150 dilution) were applied for 2 h at room temperature. Then the sections were washed, coverslipped, and mounted with Prolong Gold Antifade Reagent with DAPI. At the least, 10 fields of each section were examined to select 1 representative image. To detect colocalization of Cathepsin D and LAMP1 (lysosomal marker), MCT cells were fixed in ethanol:methanol solution (50:50 v/v) for 15 min at -20 ^{0}C. Anti-cathepsin D antibody was applied in combination with anti-LAMP1 antibody, diluted 1:100 and 1:50, respectively, in 1% BSA in TBS and applied overnight at 4 0C in the humidity chamber. Next steps were performed similarly like mentioned above. Representative images were photographed using Axio Imager A1 microscope (Carl Zeiss, Melville, NY). The Adobe Photoshop CS6 program (Adobe Systems Inc., San Jose, CA) was used to merge the images. To verify specificity of all immunostainings, primary antibodies were omitted in negative controls.

Total protein extractions from the mice kidney cortex were obtained from liquid nitrogen powdered tissue and following buffer: 0.25 M sucrose, 1 mM MgCl2, 2 mM EGTA, 25 mM HEPES, pH 7.5 and protease inhibitor cocktail. CatD activity were measured in the cytosolic fraction after centrifugation at 40,000 rpm for 30 min at 4 ^{0}C. A Pierce BCA protein assay kit was used to quantify the protein concentration.

Assessment of IL1, IL6 and Kim1 levels were performed by using according ELISA kits. All procedures were performed according to the manufacture protocols. Briefly samples were incubated with capture antibody for 2 hours at room temperature, washed with followed by 2 hours of incubation with detection antibodies. After washing, streptavidin-HRP were applied for the next 20 min of incubation at room temperature. Color development reaction were performed with substrate solution for the 20 min of incubation at room temperature with next applying of sulfuric acid stop solution. Determination of optical density were performed by using Tecan Spark 10M microplate reader at 450 nm with correction at 570 nm. Obtained results were corrected to the total protein level in the cell medium. A Pierce BCA protein assay kit was used to quantify the protein concentration.

Statistical analyses. The results are expressed as mean ± SEM. Statistical analysis was based on Student's t-test for comparison of two groups and one-way analysis of variance for multiple group comparisons with the ANOVA. Significance was defined at $P \leq 0.05$.

References

1. Jing J, Isoherranen N, Robinson-Cohen C, Petrie I, Kestenbaum BR, Yeung CK. Chronic Kidney Disease Alters Vitamin A Homeostasis via Effects on Hepatic RBP4 Protein Expression and Metabolic Enzymes. *Clin Transl Sci*. 2016;9(4):207-215. doi:10.1111/cts.12402

2. Haverty TP, Kelly CJ, Hines WH, Amenta PS, Watanabe M, Harper RA, Kefalides NA, and Neilson EG. Characterization of a renal tubular epithelial cell line which secretes the autologous target antigen of autoimmune experimental interstitial nephritis. J Cell Biol 107: 1359 –1368, 1988.
3. Liu WJ, Gan Y, Huang WF, Wu HL, Zhang XQ, Zheng HJ, Liu HF. Lysosome restoration to activate podocyte autophagy: a new therapeutic strategy for diabetic kidney disease. Cell Death Dis. 10:806. 2019.
4. You Y-H, Quach T, Saito R, Pham J, Sharma K. Metabolomics reveals a key role for fumarate in mediating the effects of NADPH oxidase 4 in diabetic kidney disease. *Journal of the American Society of Nephrology*. 2016;27(2):466–481.
5. Bestion E, Jilkova ZM, Mège JL, et al. GNS561 acts as a potent anti-fibrotic and pro-fibrolytic agent in liver fibrosis through TGF-β1 inhibition. *Ther Adv Chronic Dis.* 2020;11:2040622320942042. doi: 10.1177/2040622320942042. eCollection 2020.
6. Drel VR, Lupachyk S, Shevalye H, et al. New therapeutic and biomarker discovery for peripheral diabetic neuropathy: PARP Inhibitor, Nitrotyrosine, and Tumor Necrosis Factor-α. *Endocrinology*. 2010;151(6): 2547–2555.

Supplemental Tables

Table S1. Urine peptides measured in the JDRF Biomarkers Consortium across discovery and validation sets. Data obtained from UniProt database.

Protein Name	UniProt ID	Peptide Sequence	Full Protein Name
AQP2	P41181	SLAPAVVTGK	Aquaporin-2
CATD	P07339	VGFAEAAR	Cathepsin D
CATD	P07339	LSPEDYTLK	Cathepsin D
CNDP2	Q96KP4	EGGSIPVTLTFQEATGK	Cytosolic non-specific dipeptidase 2
CNDP2	Q96KP4	TGQEIPVNVR	Cytosolic non-specific dipeptidase 2
CO1A1	P02452	GEAGPQGPR	Collagen alpha-1(I) chain
CO1A1	P02452	GSEGPQGVR	Collagen alpha-1(I) chain
CO3A1	P02461	GPVGPSGPPGK	Collagen alpha-1(III) chain
CO3A1	P02461	DGNPGSDGLPGR	Collagen alpha-1(III) chain
CTGF	P29279	LPSPDCPFPR	Connective tissue growth factor
CTGF	P29279	DGAPCIFGGTVYR	Connective tissue growth factor
EGF	P01133	LGTAWCSCR	Epidermal growth factor
EGF	P01133	AYIWVDLER	Epidermal growth factor
EGF	P01133	EDDTWEPEQK	Epidermal growth factor
ESAM	Q96AP7	DSGPYSCSVNVQDK	Endothelial cell-selective adhesion molecule
ESAM	Q96AP7	SKPAVQYQWDR	Endothelial cell-selective adhesion molecule
IBP2	P18065	LEGEACGVYTPR	Insulin-like growth factor binding protein 2
IBP2	P18065	EPGCGCCSVCAR	Insulin-like growth factor binding protein 2
IBP3	P17936	ALAQCAPPPAVCAELVR	Insulin-like growth factor binding protein 3
IBP7	Q16270	AGAAAGGPGVSGVCVCK	Insulin-like growth factor binding protein 7
IBP7	Q16270	GHYGVQR	Insulin-like growth factor binding protein 7
ICAM1	P05362	DGTFPLPIGESVTVTR	Intercellular adhesion molecule 1
ICAM1	P05362	DLEGTYLCR	Intercellular adhesion molecule 1
MMP7	P09237	DLPHITVDR	Matrix metalloproteinase 7

MMP7	P09237	FYLYDSETK	Matrix metalloproteinase 7
PODXL	O00592	ATFNPAQDK	Podocalyxin
PODXL	O00592	LISLICR	Podocalyxin
RENR	O75787	SFDTSLIR	Renin receptor
SDC4	P31431	ETEVIDPQDLLEGR	Syndecan-4
SDC4	P31431	AGSGSQVPTEPK	Syndecan-4
SEPP1	P49908	LPTDSELAPR	Selenoprotein P
SEPP1	P49908	QPPAWSIR	Selenoprotein P
UPAR	Q03405	VEECALGQDLCR	Urokinase receptor
UPAR	Q03405	LWEEGEELELVEK	Urokinase receptor
UROK	P00749	KPSSPPEELK	Urokinase-type plasminogen activator
UROK	P00749	SDALQLGLGK	Urokinase-type plasminogen activator
VCAM1	P19320	ELQVYISPK	Vascular cell adhesion protein 1
VCAM1	P19320	SLEVTFTPVIEDIGK	Vascular cell adhesion protein 1

Table S2. Endogenous peptides monitored in the targeted proteomics assay.

Protein	Uniprot ID	Peptide	Position		Precursor (Endogenous Peptide)	Precursor (Internal Standard Peptide)	Variability of Peak Area Ratio (PAR), %CV	Variability of Protein Peak Area Ratio (PAR), %CV	Fragments
			Start	End	m/z				y-ion (charge)
Podocalyxn (PODXL)	O00592	ATFNPAQDK	363	371	496.2458	500.2529	14%	16%	y5(1+), y5(2+), y7(1+)
		LISLICR	353	359	437.7626	442.7667	22%		y3(1+), y4(1+)
Urokinase-type plasminogen activator (UROK)	P00749	KPSSPPEELK	135	144	556.3033	560.3104	23%	20%	y6(1+), y8(1+)
		SDALQLGLGK	88	97	501.2849	505.2920	20%		y6(1+), y8(1+)
Epidermal Growth Factor (EGF)	P01133	AEDDTWEPEQK	283	293	674.2886	678.2957	18%	18%	y4(1+), y6(1+), y7(1+), y9(1+)
		IYWVDLER	65	72	547.2875	552.2916	25%		y5(1+), y6(1+), y7(1+)
		LGTAWCSCR	737	745	555.7446	560.7487	19%		y4(1+), y5(1+), y6(1+), y7(1+)
Collagen alpha-1(I) chain (CO1A1)	P02452	GEAGPQGPR	330	338	434.7172	439.7214	40%	40%	y5(1+), y6(1+), y7(1+)
		GSEGPQGVR	339	347	443.7225	448.7266	44%		y5(1+), y6(1+), y7(1+)
Collagen alpha-1(III) chain (CO3A1)	P02461	DGNPGSDGLPGR	990	1001	571.2653	576.2694	60%	57%	y9(2+)
		GPVGPSGPPGK	1116	1126	475.2587	479.2658	35%		y7(1+), y8(1+)
Intercellular adhesion molecule 1 (ICAM1)	P05362	DGTFPLPIGESVTVTR	406	421	844.9463	849.9505	32%	31%	y10(1+), y10(2+), y12(1+)
		DLEGTYLCR	422	430	563.7635	568.7676	32%		y4(1+), y6(1+), y7(1+)
Cathepsin D (CATD)	P07339	LSPEDYTLK	328	336	533.2768	537.2839	14%	15%	y5(1+), y6(1+), y7(2+)
		VGFAEAAR	383	390	410.7192	415.7234	19%		y4(1+), y5(1+), y6(1+)
Matrilysin (MMP7)	P09237	DLPHITVDR	101	109	533.2880	538.2921	13%	14%	y7(2+)
		FYLYDSETK	24	32	583.2742	587.2813	17%		y6(1+), y8(1+)
Insulin-like growth factor-binding protein 2 (IBP2)	P18065	EPGCGCCSVCAR	45	56	706.7623	711.7665	38%	26%	y6(1+), y7(1+), y8(1+)
		LEGEACGVYTPR	57	68	676.3192	681.3233	22%		y6(1+), y7(1+), y11(1+)
Vascular cell adhesion protein 1 (VCAM1)	P19320	ELQVYISPK	191	199	538.8030	542.8101	12%	24%	y5(1+), y6(1+)
		SLEVTFTPVIEDIGK	152	166	824.4456	828.4527	31%		y8(1+), y9(1+), y11(1+)
Connective tissue growth factor (CTGF)	P29279	DGAPCIFGGTVYR	72	84	706.8350	711.8391	26%	26%	y7(1+), y8(1+), y9(1+)
		LPSPDCPFPR	116	125	593.2897	598.2938	27%		y5(1+), y7(2+), y8(1+)
Syndecan-4 (SDC-4)	P31431	AGSGSQVPTEPK	74	85	579.2935	583.3006	8%	21%	y5(1+), y6(1+), y7(1+), y8(1+), y9(1+)
		ETEVIDPQDLLEGR	4	17	807.4045	812.4086	30%		y8(1+), y9(1+), y10(1+)
Selenoprotein P (SEPP1)	P49908	LPTDSELAPR	269	278	549.7931	554.7973	22%	39%	y6(1+), y7(1+), y9(2+)
		QPPAWSIR	10	17	477.7614	482.7656	39%		y4(1+), y5(1+), y6(2+)
Urokinase plasminogen activator surface receptor (UPAR)	Q03405	LWEEGEELELVEK	30	42	801.8985	805.9056	37%	33%	y9(1+), y10(1+), y11(1+)
		VEECALGQDLCR	13	24	725.3267	730.3308	40%		y6(1+), y7(1+), y8(1+), y9(1+), y11(1+)
Insulin-like growth factor-binding protein 7 (IBP7)	Q16270	AGAAAGGPGVSGVCVCK	71	87	759.3636	763.3707	18%	18%	y11(1+), y12(1+)
		GHYGVQR	172	178	408.7092	413.7133	47%		y4(1+), y5(1+), y6(1+)
Endothelial cell-selective adhesion molecule (ESAM)	Q96AP7	DSGPYSCSVNVQDK	90	103	778.3383	782.3454	25%	22%	y7(1+), y8(1+)
		SKPAVQYQWDR	149	159	689.3491	694.3533	18%		y6(1+), y9(1+)
Cytosolic non-specific dipeptidase (CNDP2)	Q96KP4	EGGSIPVTLTFQEATGK	412	428	867.9491	871.9562	23%	21%	y10(1+), y12(1+), y12(2+)
		TGQEIPVNVR	148	157	556.8066	561.8107	16%		y5(1+), y6(1+), y7(1+)
Aquaporin-2 (AQP2)	P41181	SLAPAVVTGK	187	196	471.7846	475.7917	23%		y5(1+), y6(1+), y7(1+), y7(2+), y9(1+)
Insulin-like growth factor-binding protein 3 (IBP3)	P17936	ALAQCAPPPAVCAELVR	19	35	911.9688	916.9729	37%		y11(1+), y11(2+), y12(1+), y13(1+)
Renin receptor (RENR)	O75787	SFDTSLIR	251	258	469.7507	474.7549	17%		y4(1+), y6(1+)

Table S3. Distribution of JDRF cohorts by laboratory processing Subsets.

Cohort	Discovery set				Validation set	
	Subset 1		Subset 2		Subset 3	
	Case	Control	Case	Control	Case	Control
FinnDiane (n = 561)	70	67	70	66	150	138
Steno (n = 352)	23	57	18	81	39	134
EDC (n=138)	14	20	16	18	33	37
CACTI (n=219)	19	36	19	36	37	72
Total (n=1270)	126	180	123	201	259	381

Table S4. Baseline characteristics of participants from cohorts used for follow-up cathepsin D serum and urine measurements.

	Healthy Kidney Study n = 19	Pilot Vitamin D Trial n = 21	Boston Kidney Biopsy Cohort n = 56	Overall n = 96
Demographics				
Female sex	10 (53)	7 (33)	31 (55)	48 (50)
Age (years)	62.6 (11.6)	60.3 (11.6)	57.9 (14.8)	59.3 (13.5)
Race & ethnicity				
White	18 (95)	15 (71)	27 (48)	60 (62.5)
Black	0 (0)	2 (10)	21 (38)	23 (24.0)
Hispanic	0 (0)	1 (5)	8 (14)	9 (9.4)
American Indian / Native Alaskan	1 (5)	0 (0)	0 (0)	1 (1.0)
Native Hawaiian/ Pacific Islander	0 (0)	1 (5)	0 (0)	1 (1.0)
Asian	0 (0)	2 (10)	2 (4)	4 (4.2)
Unknown/other	0 (0)	1 (5)	6 (11)	7 (7.3)
BMI	31.1 (6.6)	36.8 (9.6)	31 (5.9)	33.0 (8.1)
Medical history and clinical characteristics				
Hypertension history or on antihypertensive medications	11 (57.9)	19 (90.4)	55 (98.2)	85 (88.5)
Laboratory data at baseline				
eGFR (ml/min/1.72m^2)	86 (17)	84 (17)	32 (19)	54 (32)
UACR (mg/g), median (IQR)	9 (6-14)	35 (23-93)	2310 (912-4603)	441 (24-2893)

Categorical variables are listed as N (%), continuous variables are listed as mean (SD) unless otherwise stated.

Number (%) of missing values for each variable in the overall study population: BMI 38 (40), eGFR 1 (1), UACR 1 (1).

eGFR = estimated glomerular filtration rate; UACR = urine albumin creatinine ratio

Table S5. Associations of 38 urine peptides with rapid eGFR decline in the JDRF Biomarkers Consortium discovery set. Results from minimal, intermediate, and full covariate models are shown. The minimal model is adjusted for age, sex, and urine creatinine. The intermediate model is adjusted for these variables plus diabetes duration, antihypertensive medication use, systolic blood pressure, hemoglobin A1c. The full model is adjusted for variables in the preceding models plus urine albumin creatinine ratio and eGFR. Values for peptides with FDR < 5% are bolded and shaded.

	Minimally adjusted model		Intermediately adjusted model		Fully adjusted model	
	OR (95% CI)	**FDR**	**OR (95% CI)**	**FDR**	**OR (95% CI)**	**FDR**
Cathepsin D (VGFA…)	**1.66 (1.37, 2.03)**	**1.8e-05**	**1.54 (1.26, 1.89)**	**1.2e-03**	**1.44 (1.17, 1.78)**	**3.0e-02**
Cathepsin D (LSPE…)	**1.55 (1.27, 1.88)**	**2.6e-04**	**1.44 (1.18, 1.76)**	**8.2e-03**	1.34 (1.09, 1.65)	1.1e-01
Matrix metalloproteinase 7 (DLPH…)	**1.44 (1.19,1.74)**	**1.4e-03**	**1.34 (1.11, 1.63)**	**2.9e-02**	1.24 (1.02, 1.52)	4.2e-01
Insulin-like growth factor binding protein 2 (LEGE…)	**1.42 (1.16, 1.73)**	**4.2e-03**	1.33 (1.08, 1.63)	5.8e-02	1.23 (0.99, 1.52)	4.9e-01
Selenoprotein P (LPTD…)	**1.47 (1.21, 1.78)**	**1.1e-03**	**1.36 (1.11, 1.65)**	**2.9e-02**	1.20 (0.97, 1.48)	4.9e-01
Collagen alpha-1(I) chain (GEAG…)	1.01 (0.85, 1.20)	6.9e-01	1.09 (0.90, 1.32)	6.0e-01	1.19 (0.96, 1.47)	4.9e-01
Cytosolic non-specific dipeptidase 2 (TGQE…)	1.17 (0.97, 1.42)	1.9e-01	1.16 (0.95, 1.40)	4.0e-01	1.18 (0.97, 1.45)	4.9e-01
Cytosolic non-specific dipeptidase 2 (EGGS…)	1.21 (1.00, 1.46)	1.5e-01	1.17 (0.97, 1.42)	3.6e-01	1.18 (0.97, 1.43)	4.9e-01
Vascular cell adhesion protein 1 (SLEV…)	**1.41 (1.16, 1.71)**	**3.9e-03**	1.28 (1.05, 1.56)	8.5e-02	1.17 (0.96, 1.44)	4.9e-01
Connective tissue growth factor (LPSP…)	1.27 (1.04, 1.56)	8.7e-02	1.18 (0.96 1.45)	3.6e-01	1.17 (0.94, 1.45)	4.9e-01
Matrix metalloproteinase 7 (FYLY…)	**1.35 (1.11, 1.64)**	**1.5e-02**	1.27 (1.04, 1.55)	9.5e-02	1.16 (0.95, 1.42)	4.9e-01
Vascular cell adhesion protein 1 (ELQV…)	**1.38 (1.14, 1.66)**	**4.2e-03**	1.25 (1.04, 1.52)	9.5e-02	1.15 (0.94, 1.41)	4.9e-01
Collagen alpha-1(I) chain (GSEG…)	0.95 (0.80, 1.14)	6.9e-01	1.01 (0.84, 1.22)	9.3e-01	1.13 (0.92, 1.38)	6.4e-01
Selenoprotein P (QPPA…)	1.06 (0.87, 1.28)	6.8e-01	1.04 (0.85, 1.26)	8.8e-01	1.10 (0.89, 1.36)	7.5e-01
Insulin-like growth factor binding protein 3	1.23 (1.01, 1.50)	1.4e-01	1.13 (0.92, 1.38)	4.9e-01	1.10 (0.89, 1.36)	7.5e-01
Connective tissue growth factor (DGAP…)	1.07 (0.88, 1.29)	6.3e-01	1.03 (0.85, 1.25)	8.8e-01	1.09 (0.89, 1.33)	7.5e-01
Urokinase receptor (VEEC…)	1.21 (0.99, 1.47)	1.5e-01	1.12 (0.92, 1.37)	4.9e-01	1.06 (0.86, 1.30)	9.3e-01
Collagen alpha-1(III) chain (GPVG…)	0.86 (0.72, 1.04)	2.0e-01	0.92 (0.76, 1.12)	6.4e-01	1.04 (0.85, 1.27)	9.3e-01

Aquaporin-2	1.11 (0.92, 1.36)	4.3e-01	1.07 (0.88, 1.31)	7.0e-01	1.04 (0.84, 1.28)	9.3e-01
Urokinase receptor (LWEE…)	1.20 (1.00, 1.44)	1.5e-01	1.11 (0.92, 1.34)	5.1e-01	1.03 (0.85, 1.25)	9.3e-01
Renin receptor	1.02 (0.85, 1.22)	9.1e-01	1.01 (0.84, 1.21)	9.3e-01	1.02 (0.84, 1.23)	9.3e-01
Insulin-like growth factor binding protein 7 (GHYG…)	0.96 (0.80, 1.15)	7.3e-01	1.00 (0.82, 1.20)	9.6e-01	1.02 (0.83, 1.25)	9.3e-01
Syndecan-4 (AGSG…)	0.82 (0.69, 0.99)	1.4e-01	0.89 (0.74, 1.09)	4.9e-01	0.99 (0.80, 1.22)	9.3e-01
Intercellular adhesion molecule 1 (DLEG…)	1.20 (0.99, 1.45)	1.6e-01	1.11 (0.91, 1.34)	5.1e-01	0.99 (0.82, 1.20)	9.3e-01
Insulin-like growth factor binding protein 7 (AGAA…)	0.97 (0.81, 1.17)	8.5e-01	0.95 (0.79, 1.16)	8.1e-01	0.99 (0.81, 1.21)	9.3e-01
Epidermal growth factor (IYWV…)	0.91 (0.76, 1.11)	5.0e-01	0.93 (0.76, 1.13)	6.9e-01	0.99 (0.80, 1.21)	9.3e-01
Urokinase-type plasminogen activator (SDAL…)	0.85 (0.71, 1.01)	1.6e-01	0.89 (0.74, 1.07)	4.9e-01	0.98 (0.80, 1.20)	9.3e-01
Syndecan-4 (ETEV…)	0.82 (0.69, 0.99)	1.4e-01	0.94 (0.77, 1.14)	7.2e-01	0.98 (0.80, 1.20)	9.3e-01
Epidermal growth factor (AEDD…)	0.89 (0.73, 1.09)	4.0e-01	0.91 (0.74, 1.12)	6.0e-01	0.97 (0.78, 1.21)	9.3e-01
Podocalyxin (ATFN…)	1.09 (0.90, 1.32)	5.1e-01	1.01 (0.83, 1.24)	9.3e-01	0.96 (0.78, 1.18)	9.3e-01
Insulin-like growth factor binding protein 2 (EPGC…)	1.07 (0.88, 1.30)	6.3e-01	1.01 (0.83, 1.24)	9.3e-01	0.96 (0.78, 1.18)	9.3e-01
Endothelial cell-selective adhesion molecule (DSGP…)	1.00 (0.83, 1.22)	9.9e-01	0.95 (0.79, 1.16)	8.1e-01	0.96 (0.78, 1.18)	9.3e-01
Urokinase-type plasminogen activator (KPSS…)	0.85 (0.71, 1.01)	1.6e-01	0.89 (0.74, 1.08)	4.9e-01	0.95 (0.78, 1.15)	9.3e-01
Collagen alpha-1 (III) chain (DGNP…)	0.85 (0.71, 1.03)	1.9e-01	0.86 (0.71, 1.04)	3.6e-01	0.93 (0.76, 1.13)	8.2e-01
Intercellular adhesion molecule I (DGTF…)	1.00 (0.84, 1.19)	9.9e-01	0.97 (0.81, 1.17)	8.8e-01	0.91 (0.75, 1.11)	7.5-01
Endothelial cell-selective adhesion molecule (SKPA…)	0.87 (0.73, 1.03)	2.0e-01	0.87 (0.72, 1.05)	4.1e-01	0.90 (0.74, 1.10)	7.5e-01
Podocalyxin (LISL…)	0.90 (0.75, 1.07)	3.8e-01	0.89 (0.74, 1.06)	4.9e-01	0.87 (0.72, 1.06)	4.9e-01
Epidermal growth factor (LGTA…)	0.85 (0.70, 1.02)	1.6e-01	0.84 (0.70, 1.02)	3.2e-01	0.85 (0.69, 1.05)	4.9e-01

Table S6. Associations of 20 urine proteins with rapid eGFR decline in the JDRF Biomarkers Consortium discovery set. Results from minimal, intermediate, and full covariate models are shown. The minimal model is adjusted for age, sex, and urine creatinine. The intermediate model is adjusted for these variables plus diabetes duration, antihypertensive medication use, systolic blood pressure, hemoglobin A1c. The full model is adjusted for variables in the preceding models plus urine albumin creatinine ratio and eGFR. Values for proteins with FDR < 5% are bolded and shaded.

	Minimally adjusted model		**Intermediately adjusted model**		**Fully adjusted model**	
	OR (95% CI)	**FDR**	**OR (95% CI)**	**FDR**	**OR (95% CI)**	**FDR**
Cathepsin D	**1.63 (1.34, 1.99)**	**3.6-05**	**1.51 (1.23, 1.85)**	**1.9e-03**	**1.41 (1.14, 1.74)**	**3.7e-02**
Matrix metalloproteinase 7	**1.45 (1.18, 1.77)**	**2.9e-03**	**1.35 (1.10, 1.66)**	**4.8e-02**	1.23 (0.99, 1.52)	4.9e-01
Selenoprotein P	**1.38 (1.0, 1.74)**	**2.8e-02**	1.28 (1.01, 1.62)	2.0e-01	1.20 (0.94, 1.54)	4.9e-01
Cytosolic non-specific dipeptidase 2	1.20 (99, 1.46)	1.4e-01	1.18 (0.97, 1.43)	4.2e-01	1.19 (0.97, 1.46)	4.9e-01
Vascular cell adhesion protein 1	**1.41 (1.17, 1.71)**	**2.9e-03**	1.28 (1.05, 1.56)	8.9e-02	1.17 (0.95, 1.45)	4.9e-01
Collagen alpha-1(I) chain	0.98 (0.82, 1.18)	8.4e-01	1.06 (0.87, 1.28)	6.9ee-01	1.17 (0.95, 1.46)	4.9e-01
Connective tissue growth factor	1.18 (0.96, 1.46)	1.8e-01	1.11 (0.90, 1.38)	4.6e-01	1.15 (0.92, 1.43)	6.5e-01
Insulin-like growth factor binding protein 3	1.23 (1.01, 1.50)	1.1e-01	1.13 (0.92, 1.38)	4.3e-01	1.10 (0.89, 1.36)	7.7e-01
Insulin-like growth factor binding protein 2	1.25 (1.02, 1.54)	1.1e-01	1.17 (0.95, 1.45)	4.3e-01	1.09 (0.88, 1.36)	7.7e-01
Urokinase-type plasminogen activator	1.23 (1.01, 1.51)	1.1e-01	1.13 (0.92, 1.40)	4.3e-01	1.05 (0.85, 1.30)	9.1e-01
Aquaporin 2	1.11 (0.92, 1.36)	4.0e-01	1.07 (0.88, 1.31)	6.2e-01	1.04 (0.84, 1.28)	9.1e-01
Renin receptor	1.02 (0.85, 1.22)	8.4e-01	1.01 (0.84, 1.21)	8.9e-01	1.02 (0.84, 1.23)	9.1e-01
Insulin-like growth factor binding protein 7	0.95 (0.76, 1.18)	7.6e-01	0.96 (0.77, 1.22)	8.0e-01	1.01 (0.79, 1.28)	9.7e-01
Collagen alpha-1(III) chain	0.83 (0.67, 1.02)	1.4e-01	0.87 (0.70, 1.07)	4.3e-01	0.98 (0.78, 1.22)	9.1e-01
Syndecan-4	0.83 (0.68, 102)	1.5e-01	0.90 (0.72, 1.11)	4.6e-01	0.98 (0.78, 1.23)	9.1e-01
Urokinase receptor	0.82 (0.67, 0.99)	1.1e-01	0.87 (0.71, 1.07)	4.3e-01	0.96 (0.77, 1.19)	9.1e-01
Intercellular adhesion molecule 1	1.11 (0.91, 1.36)	4.1e-01	1.04 (0.85, 1.28)	7.5e-01	0.94 (0.76, 1.16)	8.4e-01
Epidermal growth factor	0.85 (0.69, 1.06)	2.2e-01	0.86 (0.69, 1.08)	4.3e-01	0.92 (0.72, 1.16)	7.7e-01
Endothelial cell-selective adhesion molecule	0.91 (0.74, 1.11)	4.3e-01	0.88 (0.71, 1.09)	4.3e-01	0.91 (0.73, 1.14)	7.7e-01
Podocalyxin	0.98 (0.80, 1.19)	8.4e-01	0.93 (0.76, 1.14)	6.2e-01	0.90 (0.73, 1.12)	7.7e-01

Table S7. Associations of urine cathepsin D peptides and protein with rapid eGFR decline in the JDRF Biomarkers Consortium validation set. Results from minimal, intermediate, and full covariate models are shown. The minimal model is adjusted for age, sex, urine creatinine, and sample processing site. The intermediate model is adjusted for these variables plus diabetes duration, antihypertensive medication use, systolic blood pressure, hemoglobin A1c. The full model is adjusted for variables in the preceding models plus urine albumin creatinine ratio and eGFR.

	Minimally adjusted model		**Intermediately adjusted model**		**Fully adjusted model**	
	OR (95% CI)	**p-value**	**OR (95% CI)**	**p-value**	**OR (95% CI)**	**p-value**
Cathepsin D (VGFA…)	1.51 (1.23, 1.86)	0.0001	1.44 (1.16, 1.79)	0.001	1.21 (0.96, 1.53)	0.11
Cathepsin D (LSPE…)	1.40 (1.14, 1.72)	0.001	1.34 (1.08, 1.66)	0.007	1.12 (0.89, 1.40)	0.34
Cathepsin D	1.47 (1.19, 1.81)	0.0003	1.40 (1.13, 1.74)	0.003	1.17 (0.92, 1.47)	0.20

Table S8. Associations of cathepsin D protein with rapid eGFR decline in the JDRF Biomarkers Consortium discovery set, sequentially adjusted for full covariate model variables.

Variable	Model 0 OR (95% CI)	Model 1 OR (95% CI)	Model 2 OR (95% CI)	Model 3 OR (95% CI)	Model 4 OR (95% CI)	Model 5 OR (95% CI)	Model 6 OR (95% CI)	Model 7 OR (95% CI)	Model 8 OR (95% CI)	Model 9 OR (95% CI)
Cathepsin D (per 1 SD)	1.31 (1.11, 1.55)	1.33 (1.12, 1.58)	1.32 (1.11, 1.57)	1.63 (1.34, 1.99)	1.62 (1.33, 1.98)	1.55 (1.27, 1.89)	1.55 (1.27, 1.90)	**1.51 (1.23, 1.85)**	**1.36 (1.10, 1.67)**	1.41 (1.14, 1.74)
Age (per 1 SD)		0.62 (0.53, 0.74)	0.62 (0.52, 0.74)	0.60 (0.50, 0.72)	0.55 (0.44, 0.69)	0.48 (0.37, 0.61)	0.48 (0.37, 0.61)	0.48 (0.38, 0.62)	0.57 (0.44, 0.73)	0.64 (0.49, 0.84)
Sex (reference: male)			0.91 (0.65, 1.27)	0.74 (0.52, 1.05)	0.74 (0.52, 1.05)	0.78 (0.55, 1.12)	0.78 (0.55, 1.12)	0.80 (0.56, 1.15)	0.80 (0.55, 1.15)	0.86 (0.59, 1.25)
Log-2 Urine creatinine (per 1 SD)				0.58(0.47, 0.72)	0.59 (0.48, 0.73)	0.62 (0.50, 0.77)	0.62 (0.50, 0.77)	0.64 (0.52, 0.79)	0.70 (0.57, 0.87)	0.70 (0.57, 0.87)
Diabetes duration (per 1 SD)					1.15 (0.92 1.145)	1.17 (0.93, 1.47)	1.17 (0.93, 1.47)	1.19 (0.95, 1.50)	1.09 (0.87, 1.38)	1.16 (0.91, 1.47)
SBP (per 1 SD)						1.36 (1.12, 1.65)	1.36 (1.12, 1.65)	1.33 (1.09, 1.62)	1.17 (0.95, 1.45)	1.23 (1.00, 1.52)
Antihypertensive medication (reference: no)							1.01 (0.69, 1.47)	0.98 (0.67, 1.43)	0.70 (0.46, 1.06)	0.76 (0.49, 1.16)
HbA1c (per SD)								1.17 (0.98, 1.40)	1.10 (0.92, 1.33)	1.08 (1.44, 2.27)
Log-2 UACR (per 1 SD)									1.66 (1.34, 2.07)	1.81 (1.44, 2.27)
eGFR (per 1 SD)										1.47 (1.17, 1.85)

Table S9. Associations of cathepsin D protein with rapid eGFR decline in the JDRF Biomarkers Consortium validation set, sequentially adjusted for full covariate model variables.

Variable	Model 0 OR (95% CI)	Model 1 OR (95% CI)	Model 2 OR (95% CI)	Model 3 OR (95% CI)	Model 4 OR (95% CI)	Model 5 OR (95% CI)	Model 6 OR (95% CI)	Model 7 OR (95% CI)	Model 8 OR (95% CI)	Model 9 OR (95% CI)	Model 10 OR (95% CI)
Cathepsin D (per 1 SD)	1.13 (0.96, 1.33)	1.12 (0.96, 1.32)	1.14 (0.97, 1.34)	1.25 (1.04, 1.49)	1.23 (1.03, 1.47)	1.34 (1.12, 1.62)	1.33 (1.11, 1.61)	**1.34 (1.11, 1.62)**	**1.19 (0.98, 1.45)**	1.16 (0.95, 1.42)	1.17 (0.92, 1.47)
Age (per 1 SD)		0.68 (0.58, 0.81)	0.70 (0.59, 0.82)	0.66 (0.56, 0.79)	0.57 (0.45, 0.71)	0.45 (0.35, 0.58)	0.45 (0.35, 0.58)	0.47 (0.36, 0.60)	0.54 (0.41, 0.70)	0.65 (0.49, 0.86)	0.65 (0.49, 0.87)
Sex (reference: male)			1.29 (0.93, 1.78)	1.14 (0.81, 1.60)	1.11 (0.79, 1.57)	1.30 (0.91, 1.85)	1.29 (0.91, 1.84)	1.38 (0.96, 1.98)	1.41 (0.97, 2.04)	1.53 (1.05, 2.25)	1.53 (1.05, 2.25)
Log-2 Urine creatinine (per 1 SD)				0.77 (0.64, 0.92)	0.78 (0.65, 0.94)	0.80 (0.67, 0.97)	0.81 (0.67, 0.98)	0.85 (0.70, 1.03)	0.94 (0.78, 1.15)	0.96 (0.79, 1.18)	0.96 (0.79, 1.18)
Diabetes duration (per 1 SD)					1.28 (1.02, 1.60)	1.27 (1.00, 1.59)	1.25 (0.99, 1.58)	1.33 (1.05, 1.69)	1.23 (0.96, 1.56)	1.27 (0.99, 1.63)	1.27 (0.99, 1.64)
SBP (per 1 SD)						1.69 (1.38, 2.06)	1.64 (1.34, 2.02)	1.55 (1.26, 1.91)	1.35 (1.09, 1.68)	1.37 (1.09, 1.70)	1.36 (1.09, 1.71)
Antihypertensive medication (reference: no)							1.20 (0.83, 1.73)	1.09 (0.75, 1.58)	0.70 (0.46, 1.06)	0.77 (0.50, 1.18)	0.77 (0.50, 1.18)
HbA1c (per 1 SD)								1.55 (1.29, 1.88)	1.42 (1.17, 1.73)	1.38 (1.13, 1.69)	1.38 (1.13, 1.69)
Log-2 UACR (per 1 SD)									1.84 (1.48, 2.30)	2.13 (1.68, 2.71)	2.13 (1.67, 2.71)
eGFR (per 1 SD)										1.62 (1.29, 2.02)	1.62 (1.29, 2.02)
Sample processing batch (ref: no)											1.01 (0.58, 1.75)

Table S10. Associations of cathepsin D peptides and protein with rapid eGFR decline in the discovery set, stratified by baseline UACR and eGFR. Results from full covariate models are shown, adjusted for age, sex, urine creatinine, diabetes duration, antihypertensive medication use, systolic blood pressure, hemoglobin A1c, urine albumin creatinine ratio, and eGFR. Values for proteins with p <0.05 are bolded.

	Cathepsin D (LSPE…)		Cathepsin D (VGFA…)		Cathepsin D	
	OR (95% CI)	**p-value**	**OR (95% CI)**	**p-value**	**OR (95% CI)**	**p-value**
UACR group (mg/g)						
Normo: ≤30	**1.34 (1.01, 1.77)**	**0.043**	**1.38 (1.04, 1.85)**	**0.029**	**1.37 (1.03, 1.83)**	**0.033**
Micro: >30,<300	**2.10 (1.25, 3.55)**	**0.0059**	**2.25 (1.33, 3.81)**	**0.003**	**2.27 (1.33, 3.87)**	**0.0032**
Macro: ≥300	0.83 (0.33, 2.11)	0.70	1.04 (0.42, 2.61)	0.93	0.93 (0.36, 2.40)	0.88
eGFR (ml/min/1.73m^2)						
120 < eGFR ≤ 150	0.75 (0.33, 1.67)	0.48	1.00 (0.46, 2.18)	0.99	0.86 (0.38, 1.92)	0.71
90 < eGFR ≤ 120	**1.61 (1.14, 2.28)**	**0.007**	**1.68 (1.18, 2.40)**	**0.0043**	**1.68 (1.17, 2.39)**	**0.0048**
60 eGFR ≤ 90	**1.55 (1.05, 2.27)**	**0.028**	**1.64 (1.12, 2.40)**	**0.012**	**1.62 (1.10, 2.39)**	**0.016**

Table S11. Associations of cathepsin D peptides and protein with rapid eGFR decline in the validation set, stratified by baseline UACR and eGFR. Results from full covariate models are shown, adjusted for age, sex, urine creatinine, diabetes duration, antihypertensive medication use, systolic blood pressure, hemoglobin A1c, urine albumin creatinine ratio, eGFR, and sample processing site.

	Cathepsin D (LSPE…)		Cathepsin D (VGFA…)		Cathepsin D	
	OR (95% CI)	**p-value**	**OR (95% CI)**	**p-value**	**OR (95% CI)**	**p-value**
UACR group (mg/g)						
Normo: ≤30	1.00 (0.73, 1.37)	0.99	1.08 (0.78, 1.48)	0.64	1.04 (0.75, 1.43)	0.82
Micro: >30,<300	1.58 (0.96, 2.57)	0.072	1.66 (0.99, 2.79)	0.059	1.64 (0.98, 2.73)	0.061
Macro: ≥300	3.70 (0.74, 18.49)	0.12	3.27 (0.75, 14.26)	0.12	3.54 (0.75, 16.72)	0.12
eGFR (ml/min/1.73m²)						
120 < eGFR ≤ 150	0.90 (0.45, 1.77)	0.75	1.10 (0.55, 2.18)	0.79	0.99 (0.50, 1.96)	0.97
90 < eGFR ≤ 120	1.03 (0.71, 1.50)	0.87	1.10 (0.75, 1.60)	0.63	1.07 (0.73, 1.56)	0.74
60 eGFR ≤ 90	1.07 (0.71, 1.62)	0.74	1.20 (0.76, 1.87)	0.44	1.13 (0.73, 1.75)	0.58

Table S12. Associations between serum cathepsin D and histopathological features in the Boston Kidney Biopsy Cohort.

Histopathologic lesion	Severity	N (% total)	Mean serum cathepsin D in ng/mL (SD)	Adjusted difference in cathepsin D concentration in ng/mL (95% CI)*	Trend test p-value*
Tubulointerstitial compartment					
Interstitial fibrosis and tubular atrophy	≤ 50%	22	389.3 (229.7)	Ref.	0.11
	>50%	31	401.6 (171.6)	24.8 (-85.3, 134.8)	
Tubulointerstitial inflammation, with fibrosed area	Mild, 11-25%	33	401.2 (221.3)	Ref.	0.78
	Moderate, 26-50% or Severe >50%	20	388.9 (149.5)	-16.3 (-124.6, 91.9)	
Tubulointerstitial inflammation, with preserved area	None, < 10%	46	396.8 (206.2)	Ref.	0.78
	Mild, 11-25%, Moderate, 26- 50% or Severe, > 50%	4	408.0 (122.5)	27.5 (-183.8, 238.8)	
Acute tubular injury	None	27	364.0 (171.1)	Ref.	0.35
	Mild or Moderate	26	430.3 (216.8)	63.1 (-40.4, 166.6)	
Glomerular compartment					
Global glomerulosclerosis	None < 10% or Mild, 11-25%	15	394.3 (269.8)	Ref.	0.85
	Moderate, 26-50%	22	426.5 (182.3)	53.0 (-80.6, 186.5)	
	Severe, >50%	18	367.8 (147.2)	-18.4 (-162.8, 126.0)	
Segmental glomerulosclerosis	None, <10%	43	393.3 199.9)	Ref.	0.08
	Mild, 11-25%, Moderate, 26-50%, or Severe, > 50%	10	437.3 (207.5)	58.1 (-80.7, 197.0)	
Mesangial matrix expansion	Mild or Moderate	19	395.2 (215.4)	Ref.	0.39
	Severe	36	405.8 (191.4)	34.4 (-77.2, 145.9)	
Mesangial hypercellularity	None or Mild	14	456.1 (265.3)	Ref.	0.15
	Moderate	27	418.5 (175.0)	-32.8 (-160.2, 94.5)	
	Severe	10	309.0 (132.1)	-139.5 (-299.2, 20.3)	
Vascular compartment					
Arterial sclerosis	Mild or Moderate	23	409.2 (221.8)	Ref.	0.38
	Severe	30	386.8 (176.7)	-43.2 (-150.5, 64.1)	
Arteriolar hyalinosclerosis	Mild or Moderate	15	420.1 (272.4)	Ref.	0.72
	Severe	38	387.2 (159.7)	-39.3 (-156.1, 77.4)	

*Adjusted for age and sex. Calculated using all available histopathologic lesion severity categories instead of combined categories.

Histopathologic lesion severities were collapsed to create balanced distributions. Trend test p-values were calculated for original, non-collapsed histopathologic lesion severity categories.

Number (%) of missing values for each variable in the overall study population: interstitial fibrosis and tubular atrophy 3 (5), tubulointerstitial inflammation, with fibrosed area 3 (5), tubulointerstitial inflammation, with preserved area 6 (11), acute tubular injury 3 (5), global glomerulosclerosis 1 (2), segmental glomerulosclerosis 3 (5), mesangial matrix expansion 1 (2), mesangial hypercellularity 5 (9), arterial sclerosis 3 (5), arteriolar hyalinosclerosis 3 (5).

Table S13. Associations between urine creatinine-normalized cathepsin D and serum cathepsin D in follow-up analyses.

	Model 1		**Model 2**		**Model 3**	
	Difference in urine creatinine-normalized cathepsin D in ng/mg (95% CI)	p-value	Difference in urine creatinine-normalized cathepsin D in ng/mg (95% CI)	p-value	Difference in urine creatinine-normalized cathepsin D in ng/mg (95% CI)	p-value
Serum cathepsin D (per 10ng/mL higher)	1.1 (-0.1, 2.2)	0.07	**1.2 (0.1, 2.3)**	**0.03**	-0.1 (-1.1, 1.0)	0.92
Sex (ref: male)			**53.5 (12.6-94.4)**	**0.01**	**41.8 (5.6, 78.1)**	**0.03**
Age (per 1 year older)			-1.1 (-2.6, 0.4)	0.15	-0.4 (-1.8, 0.9)	0.55
eGFR (per 10 ml/min/m2 higher)					-1.1 (-9.5, 7.3)	0.80
UACR (per doubling)					**15.0 (7.8, 22.2)**	**< 0.001**

Table S14. Baseline characteristics of Pima Indian study cohort subjects with type 2 diabetes.

Baseline Characteristics	T2D Pima Indian cohort (N = 74)
Male, N%	22 (30)
Age, years	46 (10)
Diabetes duration, years	15.50 (6.05)
HbA1C, %	9.3 (2.01)
GFR, ml/min	145 (53)
UACR, mg/g Cr	35 [11-147]

Values reported as N(%), Mean(SD), %N or Median [IQR].

Abbreviations: SD – standard deviation, IQR – interquartile range, GFR – measured glomerular filtration rate, UACR – albumin to creatinine ratio

Table S15. Characteristics of Kidney Precision Medicine Project (KPMP) subjects with diabetes and chronic kidney disease used for cathepsin D single cell RNAseq kidney tissue measurements.

Variable	N (%)
Female sex	7 (70)
Age, years	
30-39	2 (20)
60-69	5 (50)
70-79	3 (30)
Race	
White	7 (70)
Black or African American	2 (20)
Baseline eGFR, ml/min/1.73m^2	
≥90	2 (20)
60-89	2 (20)
30-49	5 (50)
15-29	1 (10)
UPCR, mg/g	
< 150	2 (20)
150-499	1 (10)
500-999	2 (20)
≥ 1000mg	3 (20)
UACR, mg/g	
< 30	2 (20)
30-299	1 (10)
≥ 300	5 (50)
A1c, %	
< 6.5	1 (10)
6.5-7.5	4 (40)
> 8.5	2 (20)
Hypertension present	9 (90)
Hypertension duration, years	9 (90)
< 5	3 (30)
5-9	1 (10)
> 10	5 (50)
Renin-angiotensin-aldosterone system blockade use	7 (70)

Number (%) of missing values for each variable: Race 1 (10), proteinuria 2 (20), albuminuria 2 (20), A1c 3 (30), HTN duration 1 (10).
UPCR = urine protein-creatinine ratio; UACR = urine albumin-creatinine ratio

Supplemental Figures

Figure S1. Cathepsin D activity in the media of MCT cells transfected with CatD-siRNA or Ctr-siRNA after 24 hours of treatment. (A) Cathepsin D activity level normalized to actine is reduced in cells transfected with cathepsin D siRNA (CatD-siRNA) but not control siRNA (Ctr-siRNA) or non-transfected cells. (B) Cathepsin D enzyme activity is reduced in MCT cells transfected with CatD-siRNA compared to cells transfected with Ctr-siRNA when exposed to either AGE-BSA or Ctr-BSA. Cathepsin D activity is noted in relative fluorescent units (RFU). (C) Western blot illustrating reduced cathepsin D with CatD-siRNA. ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$

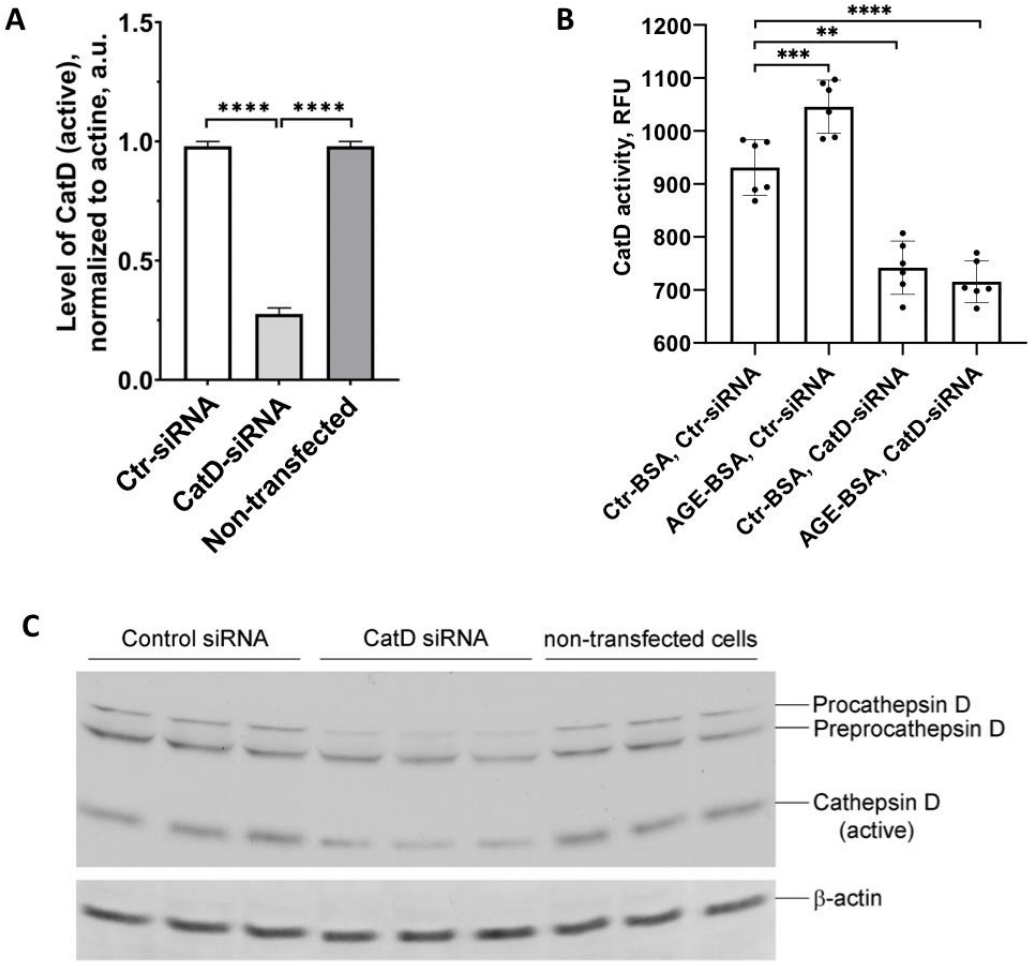
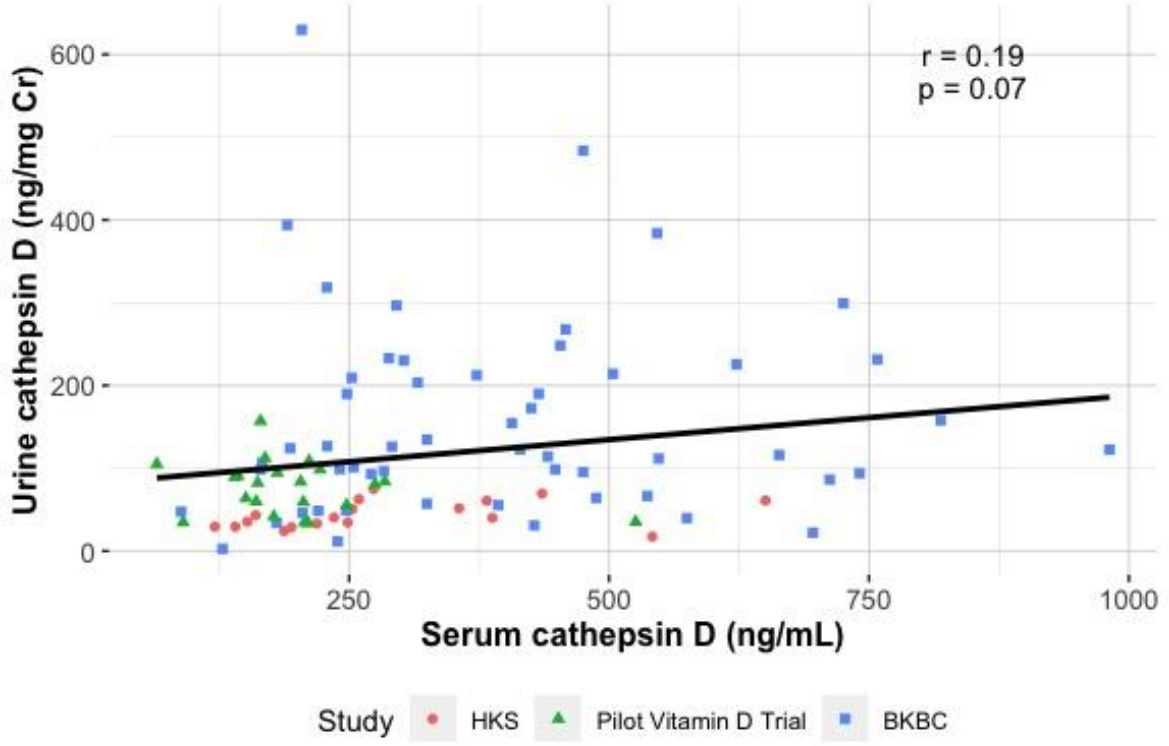


Figure S2. Cathepsin D follow-up cohort measurements. (A) Laboratory characteristics of follow-up cohort participants. Entries are mean (SD) for continuous variables unless otherwise indicated. Urine and serum cathepsin D are higher with greater albuminuria and lower eGFR. The interclass correlation coefficient for pilot Vitamin D trial urine cathepsin D measurements at baseline, 3, and 6 months was 0.52. (B) Urine and serum cathepsin D correlate only weakly in people with diabetes and a broad range of kidney damage.

A.

	HKS n = 19	**Pilot Vitamin D Trial n = 21**	**BKBC n = 56**
UACR (mg/g), median (Q_1, Q_3)	9 (6-14)	35 (23-93)	2310 (913-4604)
eGFR (ml/min/1.73m^2)	86 (17)	84 (17)	32 (19)
Urine Cathepsin D (ng/mg Cr)	44 (17)	76 (32)	155 (121)
Serum Cathepsin D (ng/mL)	285 (141)	200 (91)	399 (198)

B.



UACR = urine albumin-creatinine ratio; eGFR = estimated glomerular filtration rate; HKS = Health Kidney Study; BKBC = Boston Kidney Biopsy Cohort

Figure S3. Violin plots depicting cathepsin D (*CTSD*) mRNA expression levels across kidney cell types derived using single-cell RNA sequencing from Kidney Precision Medicine Project (KPMP) biopsy samples. Each point represents a single cell in the respective cluster/group. DKD: diabetic kidney disease, LD: living donor; ATL: thin ascending loop of Henle; Bcells: B cells; CNT: connecting tubule; DCT: distal convoluted tubule; DTL: Descending thin limb cell; EC: glomerular endothelial cells; FIB: Fibroblast; IC: collecting duct intercalated cells; MC: mesangial cells; MYL: Myeloid cell; PEC: https://www.miktmc.org; PC: collecting duct principal cells; PT: proximal tubule cells; POD: podocytes. TAL: thick ascending loop of Henle; Tcells: T cells; vSMC: Vascular smooth muscle cell.

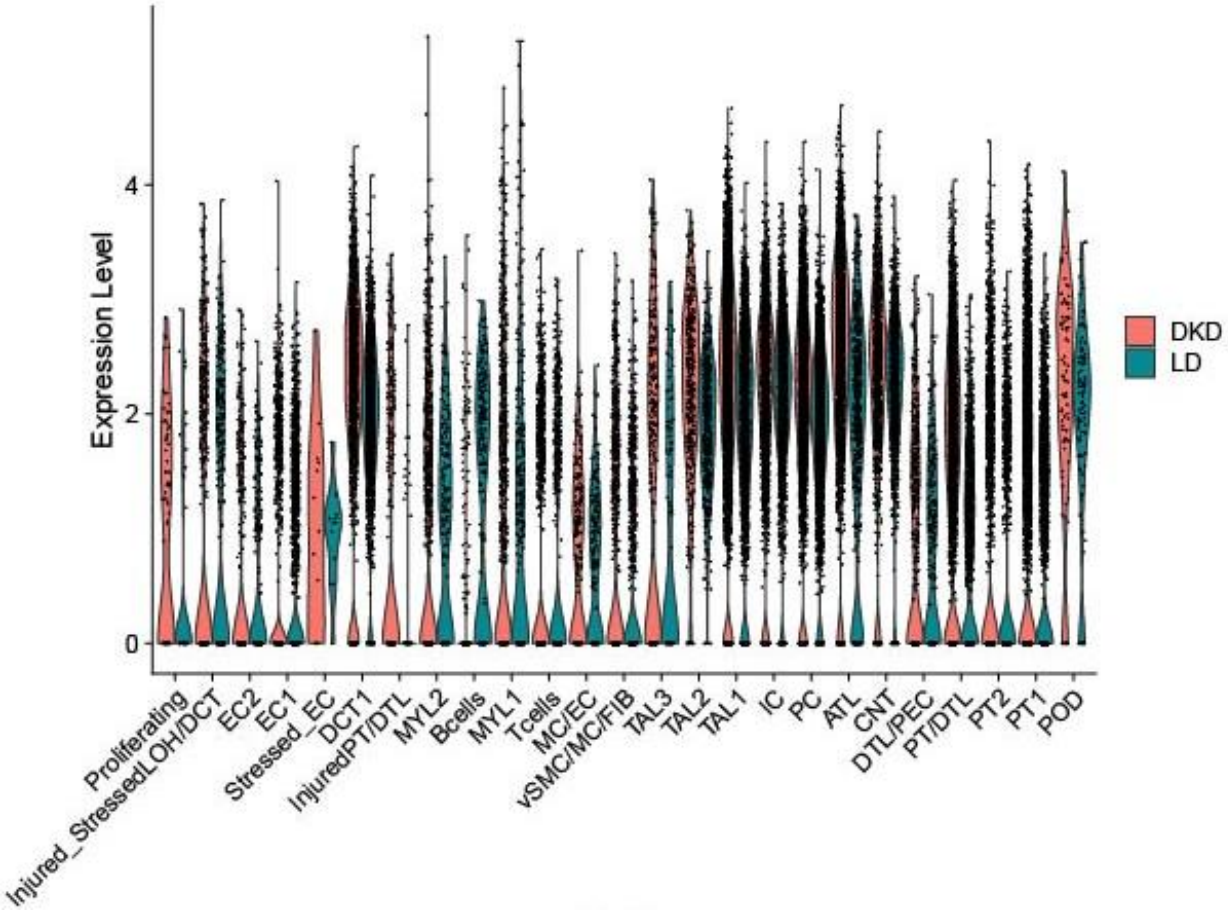


Figure S4. Cathepsin D, Akita mouse model studies. Cathepsin D enzyme activity in the urine of 8 and 16-weeks (A) old mice but not in the plasma (B) of 16-weeks old mice. (C) Representative microphotographs of immunofluorescent staining of aquaporin 1 (green) and cathepsin D (red) in the kidney of non-diabetic control (upper panel) and Akita mice (lower panel). White arrows indicate cathepsin D localization in the glomeruli and yellow arrows indicate cathepsin D localization in tubules. Scale bar = 30 µm. Cathepsin D activity is noted in relative fluorescent units (RFU). Mean ± sem. **** $p < 0.0001$ vs control.

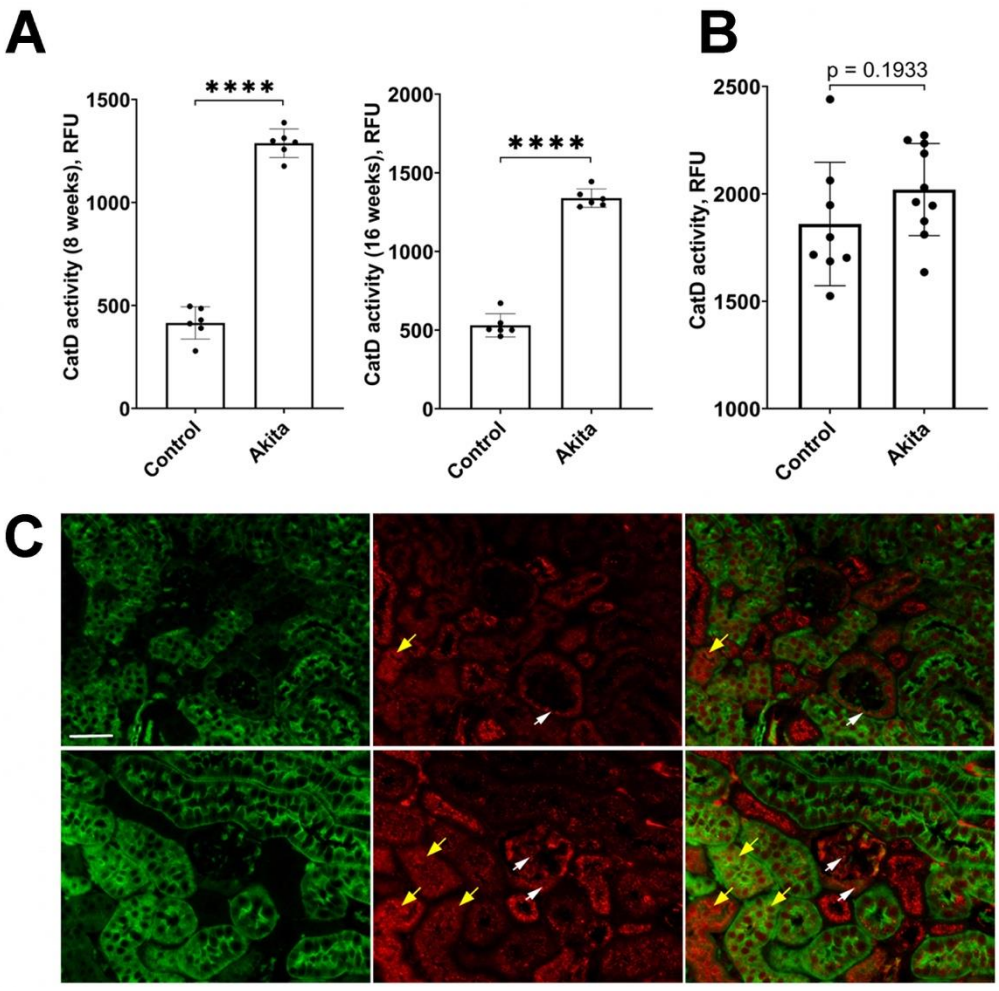


Figure S5. MCT cell inflammatory/injury marker studies. Effects of 5.5 mM glucose, 25 mM glucose, 100 ug/mL control BSA, 200 ug/mL control BSA, 100 ug/mL AGE-BSA, and 200ug AGE-BSA on MCT cell media concentrations of IL-1 (A), IL-6 (B), and KIM-1 (C). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$