Supplemental Material Table of Contents

Supplemental Spreadsheet 1. List of 1,305 proteins measured using SOMAscan.

Supplemental Spreadsheet 2. Proteins associated with eGFR in linear regression model, with or without adjustment for age, gender, race and BMI.

Supplemental Spreadsheet 3. Proteins associated with renal interstitial fibrosis in linear regression model, with or without adjustment for age, gender, race and BMI.

Supplemental Spreadsheet 4. DAVID gene ontology analysis for renal interstitial fibrosis-associated proteins, and proteins in the brown module of WGCNA.

Supplemental Spreadsheet 5. Proteins associated with MMP7 in linear regression model.

Supplemental Spreadsheet 6. Paired expression of 1,225 mRNA and protein levels in human kidney samples.

Supplemental Spreadsheet 7. Cell numbers of 13 clusters between LD and DKD in the KPMP dataset.

Supplemental Spreadsheet 8. DEGs of all 13 clusters between LD and DKD in the KPMP dataset.

Supplemental Table 1. Patient characteristics of validation database.

Supplemental Figure 1. Correlation of clinical parameters with MMP7.

Supplemental Figure 2. MMP7 expression by single-cell RNA-seq, and validation of MMP7 expression by bulk RNA-seq in kidneys.

**Supplemental Table 1. Patient characteristics of validation database.**

Validation database 1

Characteristic	Control (*n*=78)	DM (*n*=94)	HTN (*n*=102)	CKD (*n*=89)	DKD (*n*=70)
Age years (SD)	52.9 (13.6)	62.7 (9.7)	60.1 (11.7)	63.4 (12.6)	68.5 (9.8)
Male %	65.4	66.0	65.7	65.2	50.0
Race					
White, *n* (%)	47 (60.2)	66 (70.2)	70 (68.7)	64 (71.9)	45 (64.2)
African American, *n* (%)	13 (16.7)	15 (16.0)	18 (17.6)	17 (19.1)	14 (20.0)
Asian, *n* (%)	2 (2.6)	2 (2.1)	1 (1.0)	0 (0)	2 (2.9)
Hispanic, *n* (%)	2 (2.6)	1 (1.1)	4 (3.9)	1 (1.1)	2 (2.9)
Other and unknown, *n* (%)	14 (17.9)	10 (10.6)	9 (8.8)	7 (7.9)	7 (10.0)
DM, *n* (%)	0 (0)	94 (100)	0 (0)	0 (0)	70 (100)
Hypertension, *n* (%)	0 (0)	80 (85.0)	102 (100)	70 (78.6)	65 (92.9)
BMI kg/m^2 (SD)	29.1 (7.1)	33.6 (9.5)	29.3 (7.1)	29.4 (5.8)	32.7 (8.3)
SBP mmHg (SD)	125 (16)	139 (21)	140 (19)	140 (25)	139 (21)
DBP mmHg (SD)	75 (11)	76 (12)	82 (12)	79 (14)	73 (12)
eGFR ml/min per 1.73 m^2 (SD)	85 (16)	82 (15)	83 (16)	41 (19)	38 (16)
Serum glucose mg/dl (SD)	112 (30)	160 (55)	105 (30)	104 (25)	188 (101)
Histology					
Glomerulosclerosis % (SD)	5.8 (9.4)	11.2 (14.1)	7.8 (10.2)	21.8 (25.8)	26.9 (25.4)
Interstitial Fibrosis % (SD)	6.6 (11.0)	9.8 (10.6)	6.7 (8.9)	24.4 (29.8)	27.1 (30.0)

Validation database 2

Characteristic	Control (*n*=152)	DKD (*n*=34)
Age years (SD)	58.3 (12.9)	67.1 (11.8)
Male %	67.8	70.6
Race		
White, *n* (%)	132 (89.2)	32 (94.1)
African American, *n* (%)	14 (9.5)	2 (5.9)
Others, *n* (%)	2 (1.3)	0 (0)
DM, *n* (%)	0 (0)	34 (100)
Hypertension, *n* (%)	92 (62)	27 (79)
BMI kg/m^2 (SD)	30.8 (7.6)	32.5 (6.7)
SBP mmHg (SD)	136 (18)	142 (19)
DBP mmHg (SD)	78 (12)	72 (10)
eGFR ml/min per 1.73 m^2 (SD)	78 (11)	40 (15)

Data are presented as mean and standard deviation with
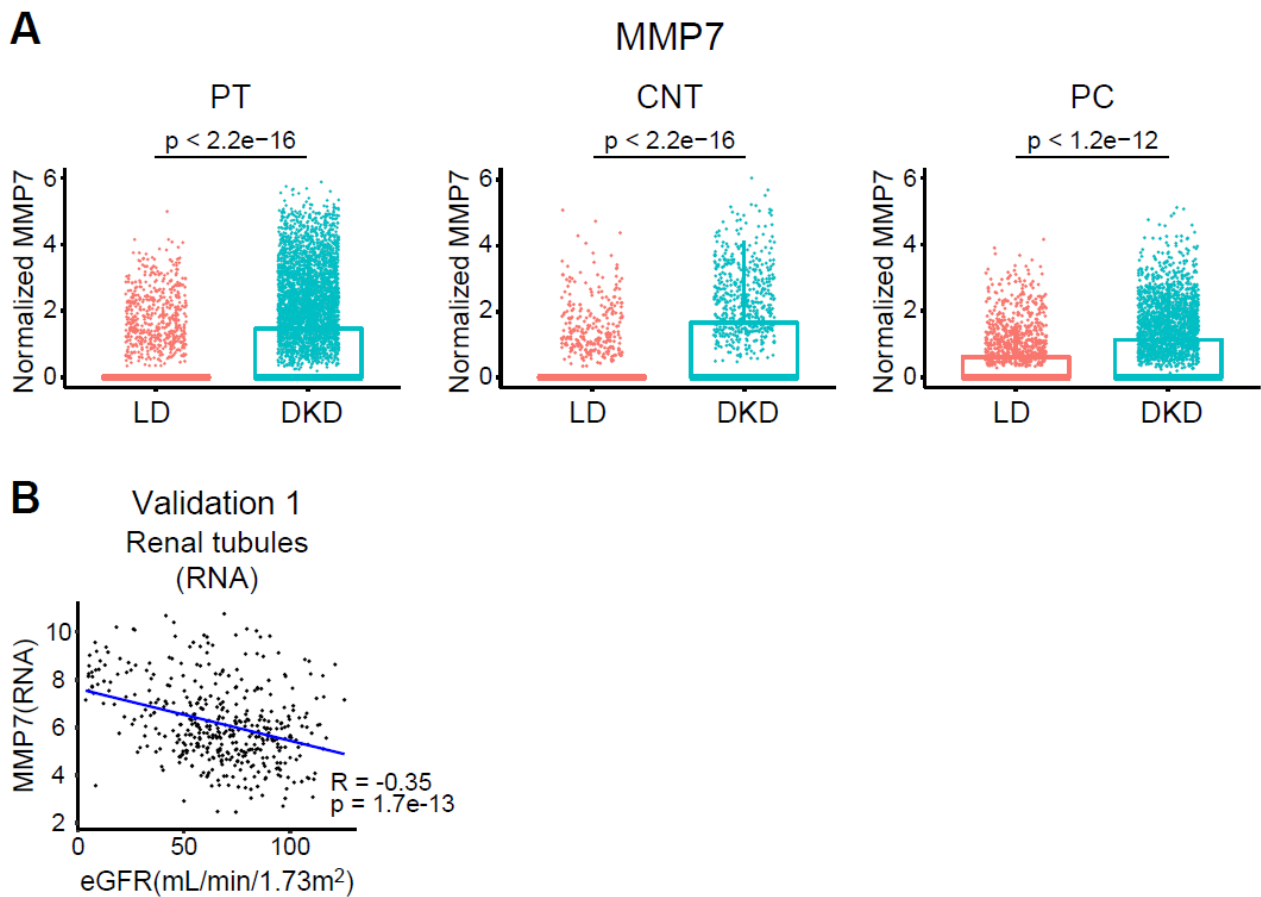the median values or percentage (%).
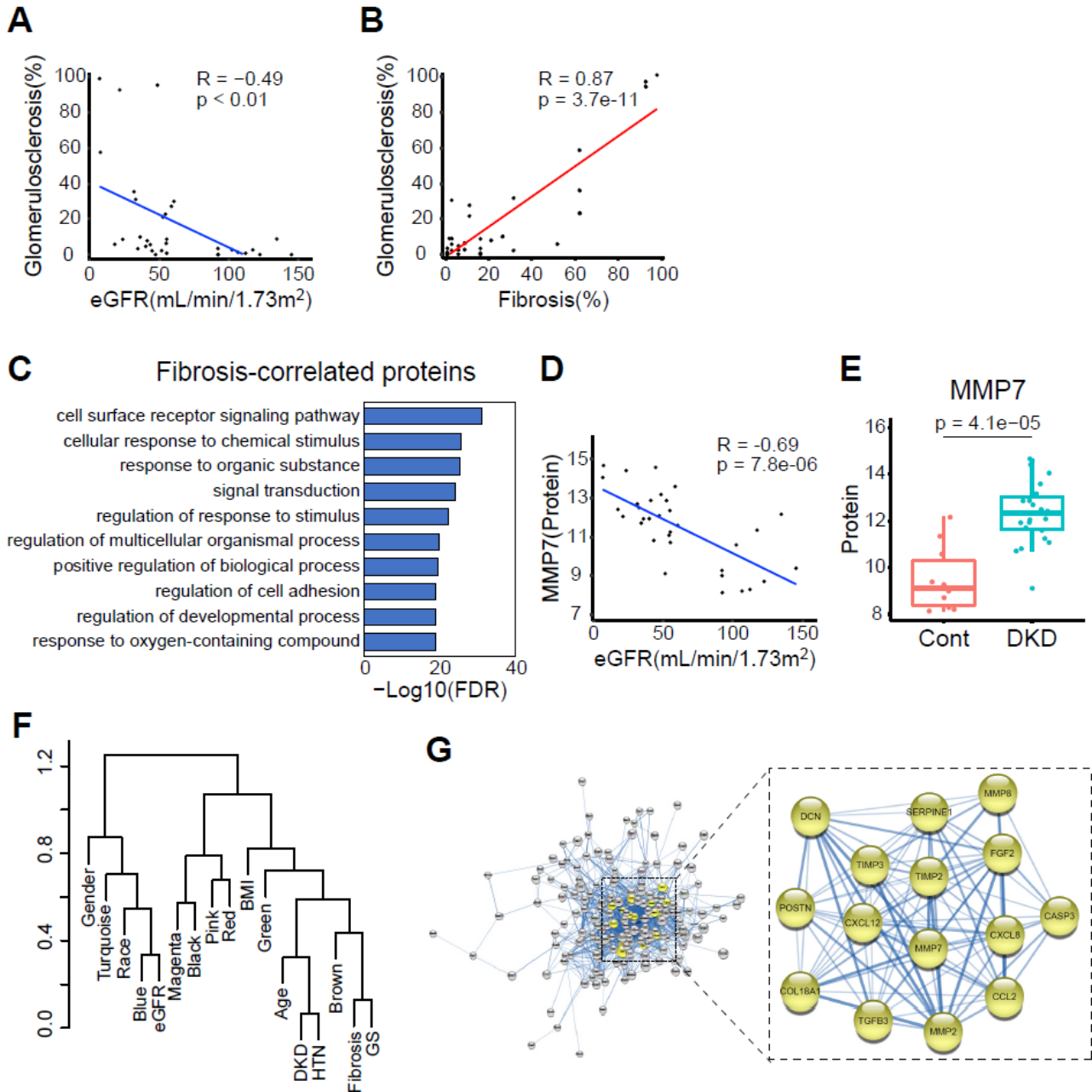
Supplemental Figure 1. Correlation of clinical parameters with MMP7, pathway analysis regarding interstitial fibrosis-correlated proteins.

(A) Correlation between eGFR and glomerulosclerosis, with Pearson R of −0.49. (B) Correlation between interstitial fibrosis and glomerulosclerosis, with Pearson R of 0.87. (C) Pathway analysis using GO showing the top pathways enriched in the protein-sets of interstitial fibrosis-correlated proteins. (D) Correlation between renal MMP7 expression measured by SOMAmer array and eGFR, with Pearson R of −0.69. (E) Box plots showing renal MMP7 expression measured by SOMAmer

array in DKD and control subjects. P values were calculated with the Wilcoxon rank-sum test (for two group comparison). (F) WGCNA dendrogram of 8 module eigenproteins and 9 clinical traits. (G) Protein-protein interaction (PPI) networks in the protein-sets of the WGCNA brown module. Distance between nodes is determined by strength of the correlation. MCODE plugin of Cytoscape identified the highly connected regions in the PPI network shown enlarged, which represent molecular clusters (see Methods section).

Con, control; DKD, diabetic kidney disease; HTN, hypertension; BMI, body mass index; GS, glomerular sclerosis.

A MMP7

PT
p < 2.2e−16

CNT
p < 2.2e−16

PC
p < 1.2e−12

B Validation 1
Renal tubules
(RNA)

R = -0.35
p = 1.7e-13

Supplemental Figure 2. Box plots showing *MMP7* expression by single-cell RNA-seq, and validation of *MMP7* expression by bulk RNA-seq in the kidney.

(A) Box plots showing the *MMP7* expression in PT, CNT, and PC clusters between DKD subjects and LD. P values were calculated with the Wilcoxon rank-sum test (for two group comparison). (B) Correlation between *MMP7* expression by bulk RNA-seq in 433 microdissected human kidney tubules and eGFR, with Pearson R of −0.35.

PT, proximal tubule; CNT, connecting tubule; PC, principal cell; DKD, diabetic kidney diseas; LD, living donor.