Supplementary Information

Khan et al.

Uromodulin T62P variant causes kidney tubular stress and injury modulated by age and polygenic risk

Biobank Datasets:

UK Biobank (UKBB)

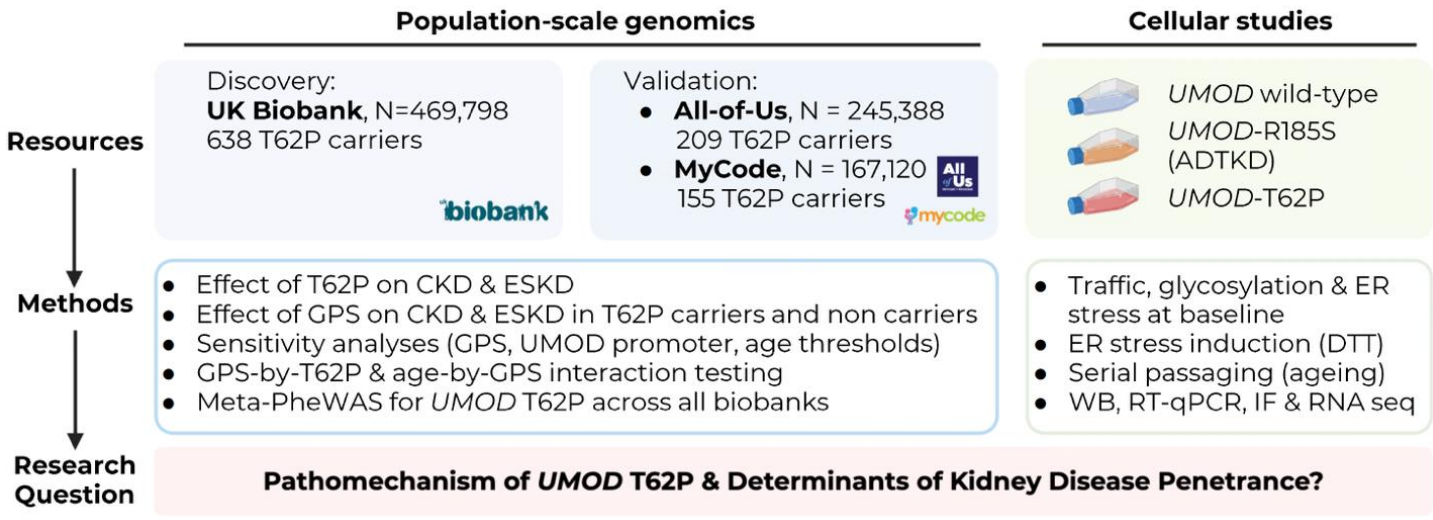
The design and composition of this dataset have been previously described in detail[1]. We accessed and analyzed the most recent data via the UKBB Research Analysis Platform (RAP) on DNAnexus. To extract UMOD variant information, we used exome sequence data for 469,835 participants as previously described[2,3]. The GPS for CKD was calculated based on the imputed microarray data for those individuals who had exome sequencing available. We implemented electronic phenotyping for CKD in this cohort, which resulted in the final number of 10767 cases (CKD stage 3 or above) and 282647 controls included in the downstream statistical analyses.

All-of-US (AoU)

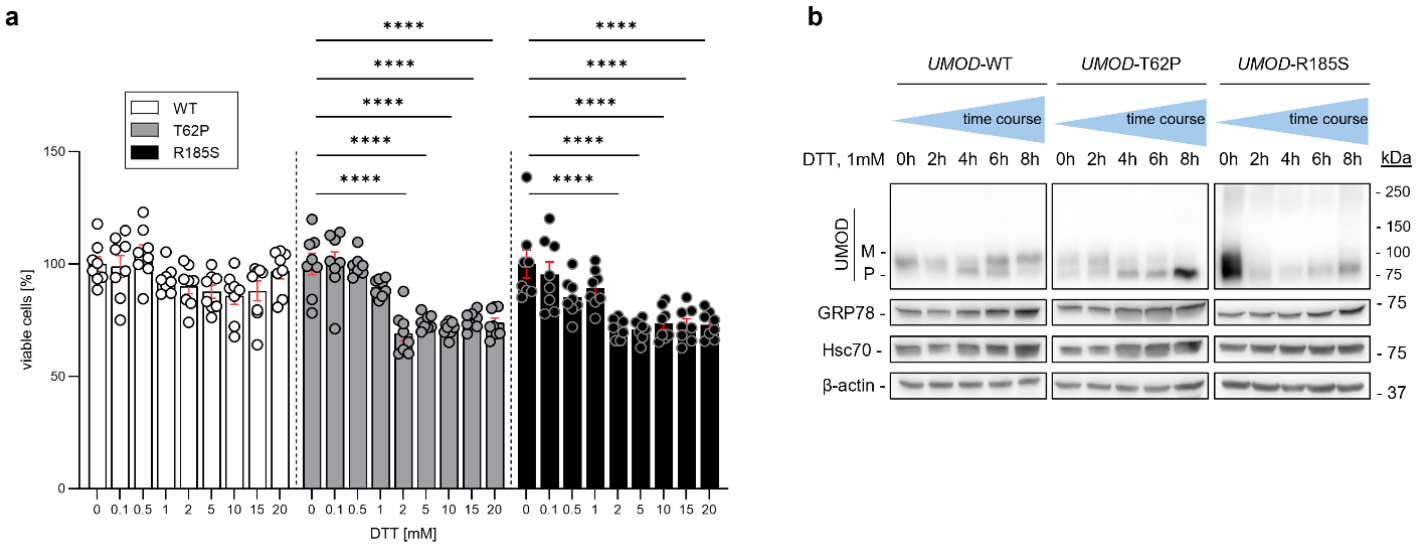
The All-of-Us research program has been described in detail[4]. We used the second data release that included N=312,944 participants. All analyses were performed using AoU Workbench on Google Cloud. To extract UMOD variant information, we used genome sequence data. To calculate the GPS, we used imputed genotype data as described previously[5]. Finally, we implemented our electronic CKD phenotype algorithm to define 26521 cases (CKD stage 3 or above) and 80753 controls for downstream statistical analyses.

MyCode

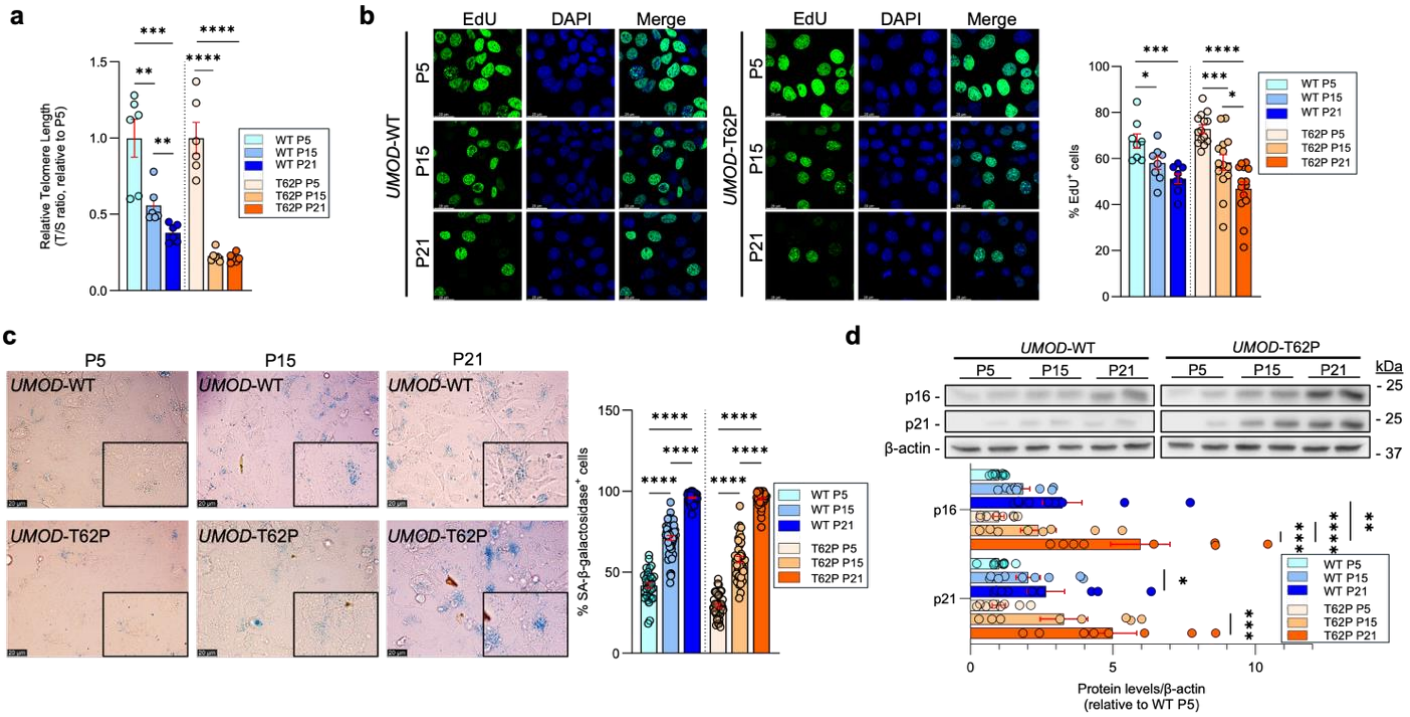
The Geisinger MyCode Community Health Initiative has been described previously in detail[6]. Briefly, MyCode is an integrated research biobank that links electronic health record (EHR) data with genomic data obtained from participants across the Geisinger Health System in Pennsylvania. To extract *UMOD* variant information, we used exome sequencing data from 167,892 participants generated through the DiscovEHR collaboration with Regeneron Genetics Center, and quality-controlled according to previously published protocols[6]. Genotype imputation was performed using microarray data available for the same individuals, allowing calculation of the GPS for CKD in participants with both imputed array data and exome data. We applied an electronic CKD phenotype algorithm to the linked EHR data within MyCode, defining CKD stage 3 or above based on laboratory measurements and diagnostic codes. After application of all inclusion/exclusion criteria, the final dataset consisted of 27299 cases (CKD stage 3 or above) and 65979 controls for downstream statistical analyses. For ESKD, the MyCode cohort used linkage to the USRDS database[7], the national data registry that collects, analyzes, and distributes information about the ESKD population in the U.S, and included 1773 cases and 166119 controls for downstream statistical analyses.



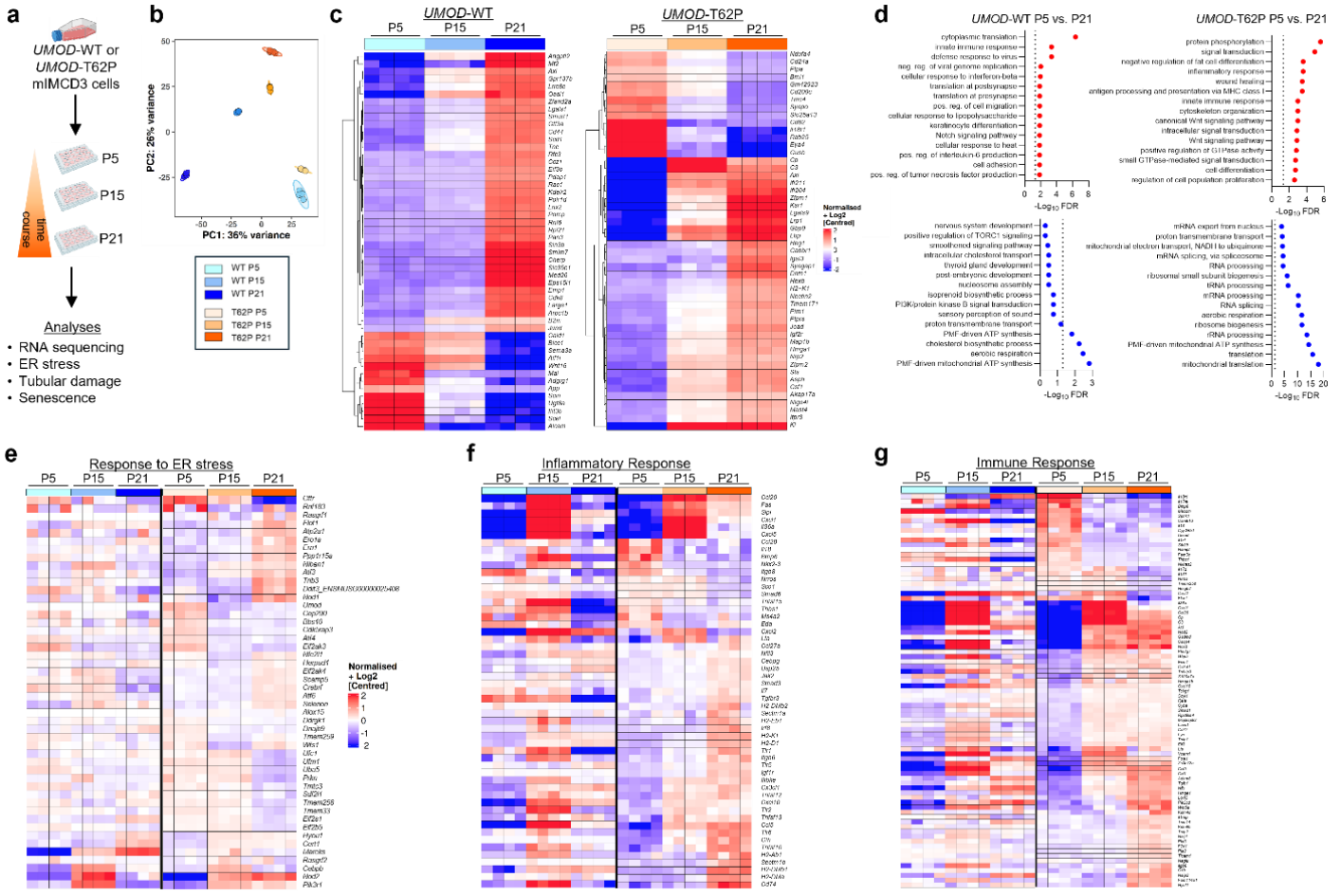
Supplementary Figure 1. Study workflow. Flow of the population and cellular studies investigating the determinants of the penetrance of the intermediate *UMOD* T62P variant.



Supplementary Figure 2. Dose- and time course-dependent effects of DTT treatment on cells expressing *UMOD* variants. (**a**) MTT cell viability assay of *UMOD*- wild-type (WT), T62P or R185S expressing cells, exposed to indicated concentrations of DTT for 8h. Bars represent mean ± SEM. Ordinary one-way ANOVA followed by Dunnett's multiple comparisons test (n = 8). *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001. (**b**) Immunoblot of protein lysates from WT-, T62P- or R185S expressing cells following time-course treatment with 1mM DTT for the indicated duration, probed for UMOD, GRP78 and Hsc70 with β- actin serving as a loading control.



Supplementary Figure 3. Effects of increasing passage number on telomere length, proliferation and senescence. (**a**) Relative telomere length of *UMOD*-WT or T62P expressing mIMCD3 cells harvested at passages 5 (P5), 15 (P15), and 21 (P21) (n=6 biological replicates). Bars represent mean ± SEM. Ordinary one-way ANOVA followed by Tukey's post hoc test. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001. (**b**) Representative immunofluorescence analysis of *UMOD*-WT or T62P-expressing cells at different passages (P5, P15, P21) following 2 h of EdU (green) treatment. Nuclei are counterstained with DAPI. Scale bar: 20 µm. Welch's one-way ANOVA followed by Šídák's T3 multiple comparisons test (n>8 cells per condition). (**c**) Representative bright-field images of *UMOD*-WT or T62P-expressing cells at different passages (P5, P15, P21) stained for senescence-associated beta-galactosidase (SA-β-gal). Scale bar: 20 µm. Welch's one-way ANOVA followed by Šídák's T3 multiple comparisons test (n>32 cells per condition). (**d**) Immunoblot analysis of p16 and p21. β-actin served as a loading control (n=8 biological replicates). Bars represent mean ± SEM. Unpaired two-tailed t test. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001.



Supplementary Figure 4. Effects of increasing passage numbers on telomere length, response to ER stress, inflammatory and immune response in T62P and wild-type UMOD-expressing cells. (**a**) Schematic overview of the workflow. (**b**) Principal component analysis of RNA-sequencing data from WT- or T62P-expressing mIMCD3 cells harvested at passages 5 (P5), 15 (P15), and 21 (P21); PC1 and PC2 explain 36% and 26% of the total variance, respectively. (**c**) Heatmap of the top 50 DEGs in P21 cells compared to P5 cells of the indicated genotype. (**d**) Over-representation analysis showing each 15 most significantly up- (red) and downregulated (blue) Gene Ontology (GO) biological processes in either WT P21 versus P5 (left), or T62P P21 versus T62P P5 cells (right). (**e-g**) Gene expression heatmaps in either WT or T62P cells across passages (P5, P15, P21) for selected gene sets, including (**e**) response to endoplasmic reticulum stress, (**f**) inflammatory response, and (**g**) immune response pathways. Data are shown as $\log_2$-normalized counts per gene.

Supplementary Table 1. The effect of the *UMOD* locus exclusion from the genome-wide polygenic score (GPS). UKBB-based associations between the GPS and CKD risk among carriers of the *UMOD* T62P variant, before and after exclusion of variants located within ±200 kb of the *UMOD* gene. Effect estimates for each predictor (T62P variant; GPS among T62P carriers; and GPS among T62P non-carriers) are adjusted for age, sex, diabetes status, array batch (if applicable), and genetic ancestry.

	GPS in T62P Carriers OR per SD (95%CI), P-value	GPS in Non-Carriers OR per SD (95%CI), P-value
GPS with UMOD locus	4.68 (3.03 – 7.11), P=8.9E-05	1.83 (1.81 – 1.85), P<E-300
GPS without UMOD locus	4.66 (2.15 – 10.1), P=8.9E-05	1.83 (1.79 – 1.87), P<E-300

Supplementary Table 2. Sensitivity analyses of CKD risk estimates stratified by T62P carrier status and GPS risk quantile after changing CKD control definition to eGFR > 60 mL/min/1.73 m² in UKBB. Each odds ratio (OR) was calculated relative to the middle GPS quantile (Q3) among non-carriers. Effect estimates were adjusted for age, sex, diabetes, array batch, and genetic ancestry.

	T62P carriers		T62P non-carriers	
Quintile	**OR (95% CI), P**	**Cases / Controls**	**OR (95% CI), P**	**Cases / Controls**
Quintile 1	0.00 (0.00–>100), 9.26E-01	0 / 122	0.62 (0.57–0.66), 3.19E-37	1,166 / 86,371
Quintile 2	0.77 (0.19–3.16), 7.20E-01	2 / 120	0.78 (0.72–0.83), 7.63E-13	1,459 / 86,078
Quintile 3	3.63 (1.74–7.59), 5.86E-04	8 / 114	**Reference**	
Quintile 4	2.36 (1.02–5.48), 4.51E-02	6 / 116	1.25 (1.17–1.33), 2.12E-12	2,311 / 85,226
Quintile 5	**4.81 (2.51–9.20), 2.10E-06**	**11 / 112**	**1.80 (1.70–1.91), 2.88E-87**	**3,286 / 84,251**

Supplementary Table 3. **Sensitivity analyses using eGFR as a continuous outcome in UKBB.** All estimates are adjusted for age, sex, diabetes, array batch, and genetic ancestry.

Analysis	Group	β (95% CI)	P-value
GPS	All individuals	−0.0332 (−0.0336, −0.0327)	$<1\times10^{-300}$
GPS × Age	All individuals	−0.000223 (−0.000278, −0.000168)	2.46×10^{-15}
GPS × T62P	All individuals	−0.0157(−0.0284, −0.00302)	0.0152
T62P	All individuals	−0.0136 (−0.0251, −0.00219)	0.0196
GPS	T62P carriers	−0.0482 (−0.0619, −0.0344)	1.57×10^{-11}
GPS	T62P non-carriers	−0.0332 (−0.0336, −0.0327)	$<1\times10^{-300}$

Supplementary Table 4. Sensitivity analyses using bias-reduced logistic regression in UKBB. Each odds ratio (OR) was calculated relative to the middle GPS quantile (Q3) among non-carriers. Effect estimates adjusted for age, sex, diabetes, array batch, and genetic ancestry.

PRS quintile	T62P carriers OR (95% CI), P	Cases / Controls	T62P non-carriers OR (95% CI), P	Cases / Controls
Quintile 1	0.00 (0.00–>100), 9.39E-01	0 / 79	0.50 (0.46–0.54), 4.62E-61	933 / 54,428
Quintile 2	0.92 (0.25–3.38), 8.94E-01	2 / 77	0.74 (0.69–0.80), 2.92E-15	1,299 / 54,063
Quintile 3	2.77 (1.18–6.49), 1.88E-02	6 / 74	**Reference**	
Quintile 4	3.13 (1.34–7.32), 8.68E-03	6 / 73	1.40 (1.32–1.50), 6.02E-24	2,266 / 53,096
Quintile 5	**7.10 (3.65–13.8), 8.06E-09**	**13 / 67**	**2.61 (2.46–2.77), 2.97E-210**	**3,898 / 51,464**

Supplementary Table 5. CKD risk estimates stratified by T62P carrier status and quantiles of GPS risk. Each Odds Ratio (OR) was calculated with reference to the middle GPS quantile for non-carriers; all effect estimates were adjusted for age, sex, diabetes, array batch, and genetic ancestry; data from individual cohorts w*ere* pooled using fixed-effects meta-analysis.

		T62P carriers	T62P carriers	T62P non-carriers
Biobank	Quintile	Cases/Controls	OR (95% CI), P	OR (95% CI)
UKBB (Discovery)	Quintile 1	0/79	0.00 (0.00–>100), 9.39E-01	0.49 (0.45–0.53), 3.25E-63
	Quintile 2	2/77	0.74 (0.18–3.10), 6.79E-01	0.74 (0.68–0.80), 2.10E-15
	Quintile 3	6/74	2.60 (1.09–6.20), 3.19E-02	**Reference**
	Quintile 4	6/73	2.92 (1.22–6.97), 1.60E-02	1.40 (1.32–1.50), 5.96E-24
	Quintile 5	**13/67**	**6.96 (3.56–13.6), 1.50E-08**	**2.60 (2.44–2.76), 1.42E-208**
AoU (Validation 1)	Quintile 1	1/11	0.26 (0.03–2.20), 2.15E-01	0.66 (0.62–0.71), 1.09E-30
	Quintile 2	5/7	2.24 (0.62–8.16), 2.21E-01	0.86 (0.80–0.92), 1.13E-05
	Quintile 3	4/9	0.75 (0.19–2.94), 6.76E-01	**Reference**
	Quintile 4	4/8	1.17 (0.29–4.73), 8.21E-01	1.15 (1.07–1.23), 6.41E-05
	Quintile 5	**5/8**	**1.57 (0.45–5.47), 4.77E-01**	**1.64 (1.53–1.75), 1.32E-45**
MyCode (Validation 2)	Quintile 1	3/7	0.63 (0.10–3.96), 6.10E-01	0.55 (0.51–0.59), 0
	Quintile 2	4/8	0.50 (0.11–2.27), 3.50E-01	0.79 (0.75–0.83), 0
	Quintile 3	4/10	0.89 (0.24–3.29), 8.60E-01	**Reference**
	Quintile 4	2/6	0.97 (0.12–7.86), 9.80E-01	1.33 (1.25–1.42), 0
	Quintile 5	**9/2**	**7.33 (1.16–46.3), 3.00E-02**	**2.09 (1.96–2.23), 0**
Validation Meta (AoU + MyCode)	Quintile 1	4/18	0.43 (0.10-1.73), 2.38E-01	0.59 (0.57-0.63), P=2.4E-102
	Quintile 2	9/15	1.19 (0.44-3.17), 7.22E-01	0.81 (0.77-0.84), P=4.8E-24
	Quintile 3	8/19	0.81 (0.31-2.10), 6.78E-10	**Reference**
	Quintile 4	6/14	1.10 (0.34-3.53), 8.63E-01	1.24 (1.17-1.30), P=4.5E-20
	Quintile 5	**14/10**	**2.54 (0.90-7.16), 7.59E-02**	**1.85 (1.76-1.95), P=1.3E-142**
Combined Meta (Discovery + Validation)	Quintile 1	4/97	0.43 (0.11–1.73), 2.37E-01	0.57 (0.55–0.59), 3.0E-160
	Quintile 2	11/92	1.02 (0.45–2.30), 9.56E-01	0.80 (0.77–0.82), 9.3E-37
	Quintile 3	14/93	1.53 (0.81–2.90), 1.93E-01	**Reference**
	Quintile 4	12/87	2.06 (1.02-4.13), 4.28E-02	1.29 (1.24–1.35), 2.3E-40
	Quintile 5	**27/77**	**5.17 (2.94–9.08), 1.04E-08**	**2.11 (2.03–2.19), 1.5E-332**

Supplementary Table 6. The performance of GPS in a re-optimization analysis using 70% of the UKBB after excluding all carriers of the T62P variant. Performance was assessed using logistic regression models with the GPS as the predictor and CKD case-control status as the outcome. Reported metrics include odds ratio (OR) per standard deviation (SD) of each GPS and corresponding P-values. The models were adjusted for age, sex, diabetes status, genotyping batch, and genetic ancestry. The adjusted area under the receiver operating characteristic curve (AUC) reflects performance including covariates, while the crude AUC reflects performance of the GPS alone. Nagelkerke's pseudo-R^2 was used to estimate the variance in CKD status explained by the GPS, independent of covariates. For P+T models, r^2 indicates the LD pruning threshold; for LDpred, ρ (rho) denotes the fraction of causal variants assumed. The best-performing GPS is highlighted in bold red.

Method	Parameter	N variants	OR* per SD of GPS	P-value	AUC (Adjusted*)	AUC (Crude)	Nagelkerke R^2
P+T	P=1E-01	89,880	1.85	P<1.00E-300	0.8382	0.6507	0.0439
P+T	P=1E-02	21,764	1.84	P<1.00E-300	0.8388	0.6511	0.044
P+T	P=1E-03	7,486	1.77	P<1.00E-300	0.8361	0.6415	0.0389
P+T	P=1E-04	3,598	1.76	P<1.00E-300	0.8356	0.6391	0.038
P+T	P=1E-05	2,111	1.74	P<1.00E-300	0.835	0.6374	0.0367
P+T	P=1E-06	1,407	1.74	P<1.00E-300	0.8349	0.6369	0.0363
P+T	P=1E-07	1,028	1.74	P<1.00E-300	0.8345	0.6366	0.0361
P+T	P=1E-08	753	1.72	P<1.00E-300	0.8336	0.6341	0.0345
P+T	P=3.0E-02	41,426	1.85	P<1.00E-300	0.8388	0.6522	0.0446
P+T	P=3.0E-03	11,918	1.81	P<1.00E-300	0.8374	0.6464	0.0415
P+T	P=3.0E-04	4,971	1.77	P<1.00E-300	0.8362	0.6412	0.039
P+T	P=3.0E-05	2,675	1.75	P<1.00E-300	0.8352	0.6378	0.0373
LDpred	ρ=1.0E+00	5,440,627	1.83	P<1.00E-300	0.8376	0.6505	0.043
LDpred	ρ =1.0E-01	5,440,627	1.17	1.51E-31	0.8153	0.5378	0.0027
LDpred	ρ=1.0E-02	5,440,627	1.16	3.46E-30	0.8149	0.5394	0.0026
LDpred	ρ=1.0E-03	5,440,627	1.16	4.64E-29	0.8149	0.5385	0.0025
LDpred	ρ=3.0E-01	5,440,627	1.64	1.11E-307	0.8296	0.6242	0.0286
LDpred	ρ=3.0E-02	5,440,627	1.17	3.74E-32	0.8152	0.5376	0.0027
LDpred	ρ=3.0E-03	5,440,627	1.16	9.38E-30	0.815	0.539	0.0025

*Adjusted for age, sex, diabetes, the first four principal ancestry components, and genotyping batch.

Supplementary Table 7. Sensitivity analyses assessing differences between the subsets of the UKBB used originally for training (70%) and testing (30%) of the GPS. Each odds ratio (OR) was calculated relative to the middle GPS quantile (Q3) among non-carriers. Effect estimates were adjusted for age, sex, diabetes, array batch, and genetic ancestry.

Analysis	30% Testing (3,411 cases; 92,563 controls)	70% Training (6,325 cases; 163,863 controls)
A. Odds ratios (ORs), 95% CI		
GPS (per 1-SD) among T62P carriers	5.42 (1.02–28.9)	4.01 (1.81–8.90)
GPS (per 1-SD) among non-carriers	1.59 (1.54–1.64)	1.98 (1.92–2.04)
GPS Q5 for T62P carriers vs. Q3 of non-carriers	5.26 (1.46–19.0)	8.50 (3.91–18.5)
B. Interaction terms, β (SE)		
GPS × Age	0.015 (SE=0.003)	0.004 (SE=0.002)
GPS × T62P	0.60 (SE=0.70)	0.51 (SE=0.33)

Supplementary Table 8. List of primers used for RT-qPCR analysis.

Gene	Forward primer (5'-3')	Reverse primer (5'-3')	Product size (bp)
hUMOD	CGAGTGTCACCTGGCGTAC	TGTGCTCCAGGAGGGAGATA	155
Atf4	CATGCCAGATGAGCTCTTGA	GGCAACCTGGTCGACTTTTA	145
Ddit3	CCCAGGAAACGAAGAGGAG	CCTCCTGGGCCATAGAACT	273
Cryab	ACTTCCCTGAGCCCCTTCTA	CTTGCCGTGGACCTCAATCA	186
Dnaja4	TGAAGGCATCGGTGGGAAAA	AGTTCTCACAGCGGTCCTTG	176
Dnajb4	GACCCTCCCGTCTCAAACAA	TGATTTTGGTGCCTTCTTTCCAC	195
Hsp5a	CACTGTGGTACCCACCAAGA	GCAGGAGGAATTCCAGTCAG	149
Hsp90aa1	CCCGTGAAATGCTGCAACAA	GTACCGCAACAGCTCTGAAAG	200
Atf6	TCAGCCTGGCCTACATTTCA	TATGGTGAAGGAAGCAGGCA	168
Atf3	TTGTCCAATGGCCAGGGTAT	TGCTTTTGAGGGGCACTTTC	207
Tmem258	CTTCACTGCCTGGTTCTTCG	AAGCAGAGACTTGGTGAGCT	226
Ikbke	GGGTTACATTGGGTCCTGGA	GGCTCCTTTGGGTTTCTGTG	201
Nkb1	CCTTCCGCAAACTCAGCTTT	GGACGATGCAATGGACTGTC	173
Mapk8	AGAAACTGTTCCCCGATGTG	TGATGTATGGGTGCTGGAGA	150
Mapk9	GCCCTGCTCAGTTCCAGTAG	GCATACGAGTCCCCACAGT	198
Mapk14	CCCCAGAGATCATGCTGAA	ACAACGTTCTTCCGGTCAAC	99
Cxcl1	AGTAGAAGGGTGTTGTGCGA	CGTGCGTGTTGACCATACAA	155
Cxcl10	GCCGTCATTTTCTGCCTCAT	GATAGGCTCGCAGGGATGAT	143
Vcam1	ACAGACAGTCCCCTCAATGG	ACAGTGACAGGTCTCCCATG	153
Lgals3	GCCTACCCCAGTGCTCCT	TTGCGTTGGGTTTCACTGTG	151
Adam8	GCGATGTATCAGGCAAAGCA	TCTCTGCTCCCCAACTTCTG	167
Tgfb1	TCGCTTTGTACAACAGCACC	ACTGCTTCCCGAATGTCTGA	156
Smad2	AGAATGGCAGAGACGATGCT	CTATGAGCCTGAGCCTTTGG	159
Smad4	AGCTCCAGCCATCAGTCTGT	GCAGGACTTCATCCAAGAGC	188
Smad7	CCCGAGTGATTGCTTTTCAT	CACACACACACAACCCAACA	195
Wnt5	ACGAGGAGCCATGTTCAGAA	ACGCAGGAGGATAACAACCA	193
Gapdh	TGCACCACCAACTGCTTAGC	GGATGCAGGGATGATGTTCT	176

Supplementary Table 9. List of primary antibodies.

Target antigen	Host species	Dilution	Cat #	Source
hUromodulin	Sheep	WB: 1:500 IF: 1:600	PA546959	Invitrogen
GRP78/BiP	Rabbit	WB 1:1000 IF 1:400	ab21685	Abcam
Hsc70	Rabbit	WB 1:1000	ab51052	Abcam
Lipocalin-2 (LCN2)	Goat	WB 1:500	AF1857	R&D Systems
eIF2α (D7D3)	Rabbit	WB: 1:500	#5324	Cell Signaling
phospho-eIF2α (Ser51) (D9G8)	Rabbit	WB: 1:500	#3398	Cell Signaling
ATF-4 (D4B8)	Rabbit	WB 1:500	#11815	Cell Signaling
PERK (C33E10)	Rabbit	WB: 1:500	#3192	Cell Signaling
Phospho-PERK	Rabbit	WB: 1:500	MA5-15033	Invitrogen
p21 (EPR18021)	Rabbit	WB: 1:1000	ab188224	Abcam
CDKN2A/p16INK4a (EPR20418)	Rabbit	WB: 1:1000	ab211542	Abcam
DRP1 (D6C7)	Rabbit	WB: 1:1000	8570T	Cell Signalling
p-DRP1 (S616)	Rabbit	WB: 1:1000	3455S	Cell Signalling
β-Actin	Mouse	WB 1:10000	A5441	Sigma-Aldrich
Calnexin	Rabbit	IF: 1:400	C4731	Sigma-Aldrich

Supplementary Table 10: List of oligonucleotides used *in vitro*.

Oligonucleotide	Primer sequence (5'-3')	Source
A184C, T62P	GGCGATGGCCTGCCCTGCGTGGACC	Microsynth AG
C598A, R185S	TGGACGAGTACTGGAGCAGCACCGAGTAC	Microsynth AG
mRplp0_fw	GCAGGAAGAGTGTAGGTGGT	Microsynth AG
mRplp0_rev	AGGCTGACTTGGTTGCTTTG	Microsynth AG
Tel1b_fw	CGGTTTGTTTGGGTTTGGGTTTGGGTTTGGGTTTGGGTT	Microsynth AG
Tel2b_rev	GGCTTGCCTTACCCTTACCCTTACCCTTACCCTTACCCT	Microsynth AG

Supplemental References

1. Bycroft, C.*, et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203-+ (2018).
2. Backman, J.D.*, et al.* Exome sequencing and analysis of 454,787 UK Biobank participants. *Nature* **599**, 628-634 (2021).
3. Van Hout, C.V.*, et al.* Exome sequencing and characterization of 49,960 individuals in the UK Biobank. *Nature* **586**, 749-756 (2020).
4. Bick, A.G.*, et al.* Genomic data in the All of Us Research Program. *Nature* **627**(2024).
5. Khan, A.*, et al.* Polygenic risk alters the penetrance of monogenic kidney disease. *Nat Commun* **14**(2023).
6. Carey, D.J.*, et al.* The Geisinger MyCode community health initiative: an electronic health record-linked biobank for precision medicine research. *Genet Med* **18**, 906-913 (2016).
7. System, U.S.R.D. 2024 Researcher's Guide to the USRDS Database. *National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD* (2024).