Supplementary Files

UMOD Genotype and Determinants of Urinary Uromodulin in African Populations

Michél Strauss-Kruger[1,2]* PhD, Eric Olinger[3,4,5]* MD, PhD, Patrick Hofmann,[3] MD, Ian J Wilson[6] PhD, Carina Mels[1,2] PhD, Ruan Kruger[1,2] PhD ISHF, Lebo F Gafane-Matemane[1,2] PhD, John A Sayer[4] MB ChB, PhD, Cristian Ricci[7] PhD, Aletta E Schutte[1,2,8,9,10]** PhD FESC FRRSAf ISHF, Olivier Devuyst[3]** MD, PhD

*Joint first authors

**Joint last authors

[1.] Hypertension in Africa Research Team (HART), North-West University, Potchefstroom 2520, North-West Province, South Africa

[2.] MRC Research Unit for Hypertension and Cardiovascular Disease, North-West University, Potchefstroom 2520, North-West Province, South Africa

[3.] Institute of Physiology, University of Zurich, Zurich, Switzerland

[4.] Biosciences Institute, Newcastle University, Newcastle upon Tyne, UK

[5.] Center for Human Genetics, Cliniques Universitaires Saint-Luc, Brussels, Belgium

[6.] Biosciences Institute, Newcastle University, Newcastle upon Tyne, UK

[7.] African Unit for Transdisciplinary Health Research (AUTHeR), North-West University, Potchefstroom 2520, North-West Province, South Africa

[8.] DSI-NRF Centre of Excellence in Human Development and SAMRC/Wits Developmental Pathways for Health Research Unit, University of the Witwatersrand, Johannesburg, South Africa

[9.] The George Institute for Global Health, Sydney NSW 2042, Australia

[10.] School of Population Health, University of New South Wales; Sydney NSW 2052, Australia

Correspondence: Olivier Devuyst, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland. olivier.devuyst@uzh.ch

STUDY DESIGN AND METHODOLOGY

African-PREDICT recruited volunteers from in and around the Potchefstroom area of the North West Province of South Africa, between 2013 and 2017. The inclusion criteria were defined as: participants aged 20-30 years; had a screening clinic BP <140/90 mmHg, were uninfected with HIV, had not been previously diagnosed with any chronic disease, and did not make use of chronic medication (self-reported).

Baseline data collection for the South African leg of PURE took place in 2005 and included N=2010 Black South African volunteers (aged >35 years) from 6000 randomly selected households in a rural and urban setting in the North West Province, South Africa. Data of 1943 participants with baseline uromodulin data were analyzed. We excluded 67 participants that had missing uromodulin data.

Questionnaire: demographic data and dietary intake data

Participant demographics (self-reported ethnicity, sex, age and socioeconomic status [S1]) were obtained using a General Health and Demographic Questionnaire (African-PREDICT) and an Adult Questionnaire (PURE). In African PREDICT, self-defined White participants refer to individuals of European descent, while Black participants from African-PREDICT and PURE refer to individuals of African descent. In this study, Black and White are regarded as ethnic groupings, that also consider historical and cultural backgrounds. Ancestry is used when referring to a shared genetic trait. In African-PREDICT, data on dietary protein intake was obtained by means of three 24-hour dietary recall interviews using a standardized dietary collection kit (example pictures, packages, measurement tools and food models) and the five-step multiple-pass approach. The data were coded according to the South African Medical Research Council Food Composition Tables and the Food Quantities Manual was used to

convert household measures to grams [S2,S3]. Nutrient and food analysis of the dietary data was conducted by the South African Medical Research Council at the Biostatistics Unit.

Biological sampling and biochemical analyses

Early morning spot urine samples were collected from study participants. In African-PREDICT, samples were immediately taken from the Hypertension Research Clinic to the onsite laboratory and prepared according to standardized procedures. In PURE, samples collected from rural areas were rapidly frozen and stored at -18℃ up to the time (less than 5 days) that it was transported to the laboratory bio-freezers. Samples for both studies were stored at -80℃. Spot urine samples were shipped and stored at -80° C in the same biochemical platform at the University of Zurich (Zurich, Switzerland).

From spot urine samples, creatinine, Na^+, K^+, and uric acid were measured using the UniCel® DxC 800 Synchron® Clinical System (Beckman Coulter, Indianapolis, US). The urinary uromodulin concentration, in spot urine, was measured by ELISA as previously described [S4], with a standard curve generated from human uromodulin (stock solution, 100 μg/ml; Millipore). The ELISA has a sensitivity of 2.8 ng/ml, a linearity of 1.0, an inter-assay variability of 3.3%, and an intra-assay variability of 5.5%. The osmolality of the urine was determined using an Advanced Osmometer 2020 (Advanced Instruments, Norwood, MA) based on freezing-point depression. Spot urine nitrite and creatinine were analyzed using gas chromatography–time of flight-mass spectrometry analyses (Leco Pegasus HT GC–TOF-MS system with Agilent 7890A GC front-end).

In African-PREDICT, RAS-Fingerprint[TM] (Attoquant Diagnostics, Vienna, Austria) analyses were performed in serum samples using highly sensitive liquid chromatography-mass spectrometry (LC-MS/MS) multiplex assay to determine angiotensin peptides. Equilibrium angiotensin levels were used to calculate ratios and combined parameters as surrogate

markers (-S) for the activity of circulating renin-S (angiotensin I + angiotensin II) In PURE, renin was measured in plasma samples using a radio immunometric assay (RIA) kit (WIZARD2® Automatic Gamma Counter, PerkinElmer Inc., Waltham, Massachusetts).

Kidney function parameters

The urinary albumin- creatinine ratio (uACR) was calculated from spot urine albumin and creatinine levels/values (Cobas Integra 400plus; Roche, Basel Switzerland). eGFR was determined using the CKD-EPI equation, without the race factor, given that the inclusion of the race variable was shown to overestimate eGFR in Black South African populations [S5,S6].

Genotyping and linkage disequilibrium analysis

Three variants in the *UMOD-PDILT* locus, strongly associated with the urinary levels of uromodulin and with eGFR and CKD in cohorts of mostly European descent, were analyzed: rs4293393 (*UMOD*) and rs12917707 (*UMOD*) and rs12446492 (*PDILT*) Supplementary Figure S1 [S7,S8]. The genotyping for these three SNPs was performed on genomic DNA from blood samples by LGC Genomics (LGC Genomics, Hoddesdon, UK) using the competitive allele-specific PCR (polymerase chain reaction) technique KASP.

In European-ancestry individuals, two independent signals at the *UMOD-PDILT* locus have been associated with the risk of CKD in large meta-GWAS [S9]. The minor alleles (in Europeans) of tag SNPs for these two loci have been associated with a reduction in CKD risk (OR~0.8) [S9]. Consistent with their prevalence in Europeans and their associations in Europeans, we are referring to the major alleles T (rs4293393), G (rs12917707) and T (rs12446492) as "CKD risk" and "UMOD increasing" [S8] alleles and for the minor alleles c (rs4293393), t (rs12917707) and a (rs12446492) as "CKD protective" and "UMOD lowering" alleles.

Differences in LD relationships of candidate variants across populations were analyzed using LDlink [S10], based on phase 3 data from the 1000 Genomes Project [S11] and Ensembl human genome GRCh38 [S12]. The LD plots for the *UMOD-PDILT* locus (GRCh37: 20,344,374-20,416,059) have been generated using Haploview 4.2 [S13] on data from the 1000 Genomes Project Phase 3 [S11]. LD plots for Europeans(n=502) have been generated using these population: Utah residents from North and West Europe (CEU); Toscani in Italia (TSI); Finnish in Finland (FIN), British in England and Scotland (GBR) and Iberian population in Spain (IBS). LD plots for Africans (n=661) included the following populations: African Caribbean in Barbados (ACB), Gambian in Western Divisions (GWD) in the Gambia, Luhya in Webuye, Kenya (LWK), Esan in Nigeria (ESN), Yoruba in Ibadan, Nigeria (YRI), Mende in Sierra Leone (MSL) and Americans with African ancestry in Southwest United States (ASW). The minimum minor allele frequency has been set to 0.1 (rs12917707 was manually included for Africans). The linkage statistics and haplotypes based on the 3 genotyped SNPs in African-PREDICT have been generated for White (n=535) and Black (n=543) subpopulations with R packages "haplo.stats" and "genetics" using maximum likelihood estimates.

Statistical analyses

Categorical variables were expressed as numbers and percentages. Data following a normal distribution are presented as the arithmetic mean ± standard deviation. To better achieve normality, variables with a non-Gaussian distribution were logarithmically transformed and presented as geometric mean (5th and 95th percentiles). The logarithmically transformed variables were used in all further analyses. The distribution of uromodulin indexed to creatinine as well as absolute uromodulin for both cohorts was right skewed but relatively normal after logarithmic transformation. Due to the strong association between urinary uromodulin and creatinine excretion, uromodulin/creatinine ratios were used to explore

independent factors (age, sex, eGFR, uACR, renin, salt intake, protein intake, *UMOD-PDILT* genotype) associated with uromodulin excretion. Patient characteristics of Black and White participants from the African-PREDICT study were compared using *t*-tests, and χ^2 tests for categorical data. The difference in the absolute uromodulin concentrations, as well as uromodulin indexed to creatinine within the genotyped variants rs12917707 (*UMOD*), rs4293393 (*UMOD*), and rs12446492 (*PDILT*), were determined using Welch ANOVA, and graphically displayed.

Pearson correlations were used to identify possible determinants (body composition, kidney function, renin-angiotensin-aldosterone system (RAAS), diet and genotype) of uromodulin/creatinine through a linear bivariate relationship. Covariates with a $p < 0.05$ were considered for inclusion in multiple regression models, while also considering multicollinearity. Standard linear multiple regression analyses (enter method, pairwise deletion, in SPSS) were performed to determine independent predictors of uromodulin/creatinine levels in the African-PREDICT and PURE study cohorts. We additionally explored correlations of uromodulin/creatinine with blood pressure and urinary Na^+, K^+, uric acid and nitrites.

For this study, we reported all analyses using urinary uromodulin normalized for urinary creatinine. In sensitivity analyses, we performed multiple regression analyses with urinary creatinine as the dependent variable to determine whether urinary creatinine may be affected by any of the independent variables. Statistical analyses were performed using SPSS version 27 (IBM; Armonk, New York, USA) and figures were created using Power Point and GraphPad Prism versions 8.0 (GraphPad Software Inc., La Jolla, California, USA). Statistical analyses were performed using pairwise deletion of data, with corresponding N number for participants reported in each Table and Figure presenting data analyses. For the African-PREDICT study, we performed all statistical analyses in the total group and stratified

according to Black and White ethnicity. P-values <0.05 were considered statistically

significant. Significant p values according to the Bonferroni correction are reported in bold.

Supplementary References

S1. Kumar N, Shekhar C, Kumar P, Kundu AS. Kuppuswamy's socioeconomic status scale-updating for 2007. *Indian J Pediatr*. 2007;74(12):1131-1132.

S2. Wolmarans P, Danster N, Dalton A, Rossouw K, Schönfeldt H. Condensed food composition tables for South Africa. *Cape Town: Medical Research Council*. Published online 2010:1-126.

S3. Langenhoven ML. *Medical Research Council Food Quantities Manual*. 2nd ed. Medical Research Council; 1991.

S4. Youhanna S, Weber J, Beaujean V, Glaudemans B, Sobek J, Devuyst O. Determination of uromodulin in human urine: influence of storage and processing. *Nephrol Dial Transplant*. 2014;29(1):136-145. doi:10.1093/ndt/gft345

S5. Inker LA, Eneanya ND, Coresh J, et al. New Creatinine- and Cystatin C-Based Equations to Estimate GFR without Race. *N Engl J Med*. 2021;385(19):1737-1749. doi:10.1056/NEJMoa2102953

S6. Stevens LA, Claybon MA, Schmid CH, et al. Evaluation of the Chronic Kidney Disease Epidemiology Collaboration equation for estimating the glomerular filtration rate in multiple ethnicities. *Kidney Int*. 2011;79(5):555-562. doi:10.1038/ki.2010.462

S7. Troyanov S, Delmas-Frenette C, Bollée G, et al. Clinical, Genetic, and Urinary Factors Associated with Uromodulin Excretion. *Clin J Am Soc Nephrol*. 2016;11(1):62-69. doi:10.2215/CJN.04770415

S8. Olden M, Corre T, Hayward C, et al. Common variants in UMOD associate with urinary uromodulin levels: a meta-analysis. *J Am Soc Nephrol*. 2014;25(8):1869-1882. doi:10.1681/ASN.2013070781

S9. Wuttke M, Li Y, Li M, et al. A catalog of genetic loci associated with kidney function from analyses of a million individuals. *Nat Genet*. 2019;51(6):957-972. doi:10.1038/s41588-019-0407-x

S10. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics*. 2015;31(21):3555-3557. doi:10.1093/bioinformatics/btv402

S11. 1000 Genomes Project Consortium, Auton A, Brooks LD, et al. A global reference for human genetic variation. *Nature*. 2015;526(7571):68-74. doi:10.1038/nature15393

S12. Yates AD, Achuthan P, Akanni W, et al. Ensembl 2020. *Nucleic Acids Res*. 2020;48(D1):D682-D688. doi:10.1093/nar/gkz966

S13. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005;21(2):263-265. doi:10.1093/bioinformatics/bth457

Supplementary Tables

Table S1: Pairwise LD data between *UMOD/PDILT* SNPs in African and European populations from the 1000 Genomes Project and the African-PREDICT cohort.

Population	N	rs4293393 & rs12917707 R^2	rs4293393 & rs12917707 D'	rs4293393 & rs12446492 R^2	rs4293393 & rs12446492 D'	rs12917707 & rs12446492 R^2	rs12917707 & rs12446492 D'
1000 Genomes Project							
Global	**2504**	**0.4344**	**0.9975**	**0.0212**	**0.3403**	**0.0279**	**0.5915**
African	**661**	**0.0972**	**1.00**	**0.0286**	**0.3442**	**<0.001**	**0.0709**
YRI	108	0.0365	1.0	0.0616	0.4953	0.0	0.0092
LWK	99	0.1321	1.0	0.1203	0.6389	0.0005	0.1161
GWD	113	0.0816	1.0	0.0065	0.2231	0.0006	0.2467
MSL	85	0.0246	1.0	0.0004	0.0376	0.0051	1.0
ESN	99	0.0228	1.0	0.0556	0.5	0.0051	1.0
ASW	61	0.1384	1.0	0.0099	0.1644	0.0171	0.5822
ACB	96	0.231	1.0	0.0133	0.2235	0.0001	0.04
European	**503**	**1.00**	**1.00**	**0.1281**	**0.636**	**0.1281**	**0.636**
CEU	99	1.0	1.0	0.1918	0.7715	0.1918	0.7715
TSI	107	1.0	1.0	0.0585	0.5117	0.0585	0.5117
FIN	99	1.0	1.0	0.1256	0.5179	0.1256	0.5179
GBR	91	1.0	1.0	0.1646	0.6847	0.1646	0.6847
IBS	107	1.0	1.0	0.1388	0.7209	0.1388	0.7209
African-PREDICT							
African	535		0.998		0.293		0.0904
European	533		0.999		0.771		0.801

African Caribbean in Barbados **ACB** (*n*=96); Gambian in Western Divisions **GWD** in the Gambia (*n*=113); Luhya in Webuye, Kenya **LWK** (*n*=99); Esan in Nigeria **ESN** (*n*=99); Yoruba in Ibadan, Nigeria **YRI** (*n*=108); Mende in Sierra Leone **MSL** (*n*=85) and Americans with African ancestry in Southwest United States **ASW** (*n*=61).

Utah residents from North and West Europe **CEU** (*n*=99); Toscani in Italia **TSI** (*n*=107); Finnish in Finland **FIN** (*n*=99); British in England and Scotland **GBR** (*n*=91) and Iberian population in Spain **IBS** (*n*=107).

Table S2: Pearson correlations of possible determinants of uromodulin in the African-PREDICT cohort.

	Total group		Black		White	
	Uromodulin/creat. (mg/g)	Uromodulin (μg/mL)	Uromodulin/creat. (mg/g)	Uromodulin (μg/mL)	Uromodulin/creat. (mg/g)	Uromodulin (μg/mL)
Age (years)	r=0.054; p=0.061	r=-0.022; p=0.46	r=0.045; p=0.27	r=-0.010; p=0.81	r=0.062; p=0.13	r=-0.040; p=0.33
BMI (kg/m^2)	r=0.036; p=0.21	r=0.061; p=0.036	r=0.088; p=0.031	r=0.080; p=0.049	r=-0.029; p=0.47	r=0.018; p=0.66
Waist circumference (cm)	r=-0.012; p=0.69	r=0.066; p=0.023	r=0.049; p=0.23	r=0.095; p=0.019	r=-0.082; p=0.047	r=0.001; p=0.97
rs4293393 (TT; TC; CC)	**r=-0.218; p<0.001**	**r=-0.139; p<0.001**	r=-0.055; p=0.20	r=-0.007; p=0.87	**r=-0.399; p<0.001**	**r=-0.255; p<0.001**
rs12917707 (GG; TG; TT)	**r=-0.231; p<0.001**	**r=-0.129; p<0.001**	r=-0.051; p=0.23	r=-0.026; p=0.55	**r=-0.404; p<0.001**	**r=-0.271; p<0.001**
rs12446492 (TT; TA; AA)	**r=-0.196; p<0.001**	**r=-0.105; p=0.001**	r=-0.090; p=0.037	r=0.007; p=0.88	**r=-0.281; p<0.001**	**r=-0.163; p<0.001**
Total protein intake (g)	r=0.035; p=0.23	r=0.040; p=0.17	r=0.041; p=0.33	r=0.018; p=0.67	r=-0.004; p=0.93	r=0.002; p=0.97
Animal protein intake (g)	r=0.031; p=0.29	r=0.039; p=0.18	r=-0.002; p=0.96	r=-0.037; p=0.37	r=0.034; p=0.42	r=0.070; p=0.091
Plant protein intake (g)	r=-0.058; p=0.047	r=-0.058; p=0.048	r=-0.027; p=0.52	r=0.002; p=0.97	r=-0.062; p=0.14	r=-0.070; p=0.092
Estimated NaCl intake (g/day)	r=-0.019; p=0.54	r=-0.012; p=0.71	r=0.060; p=0.19	r=0.027; p=0.55	r=-0.089; p=0.036	r=-0.048; p=0.26
Clinic SBP (mmHg)	**r=-0.133; p<0.001**	r=-0.075; p=0.010	r=-0.068; p=0.095	r=-0.080; p=0.049	**r=-0.193; p<0.001**	r=-0.06; p=0.14
Clinic DBP (mmHg)	r=-0.067; p=0.021	r=-0.072; p=0.012	r=-0.040; p=0.33	r=-0.065; p=0.11	r=-0.081; p=0.048	r=-0.049; p=0.23
Uric acid$_{spot}$ (mg/dL)	r=0.081; p=0.005	**r=0.641; p<0.001**	r=0.083; p=0.040	**r=0.624; p<0.001**	r=0.053; p=0.19	**r=0.653; p<0.001**
eGFR (ml/min/1.73m^2)	r=0.073; p=0.012	r=-0.029; p=0.33	r=0.093; p=0.022	r=-0.055; p=0.18	r=0.080; p=0.053	r=0.030; p=0.47
UACR (mg/mmol)	r=-0.023; p=0.42	**r=-0.349; p<0.001**	r=-0.108; p=0.008	**r=-0.403; p<0.001**	r=0.066; p=0.11	**r=-0.285; p<0.001**
Renin-S (pmol/L)	r=0.074; p=0.011	**r=0.228; p<0.001**	r=0.080; p=0.050	**r=0.198; p<0.001**	r=0.026; p=0.53	**r=0.213; p<0.001**
Angiotensin II (pmol/L)	r=0.072; p=0.013	**r=0.222; p<0.0001**	r=0.081; p=0.047	**r=0.195; p<0.001**	r=0.020; p=0.63	**r=0.203; p<0.001**
Aldosterone (pmol/L)	r=0.076; p=0.009	**r=0.146; p<0.001**	r=0.061; p=0.13	r=0.091; p=0.026	r=0.057; p=0.17	**r=0.132; p=0.001**
Urinary$_{spot}$ Na$^+$ (mmol/l)	r=0.058; p=0.046	**r=0.293; p<0.001**	r=0.091; p=0.026	**r=0.400; p<0.001**	r=0.060; p=0.14	**r=0.259; p<0.001**
Urinary$_{spot}$ K$^+$ (mmol/l)	**r=0.119; p<0.001**	**r=0.530; p<0.001**	**r=0.156; p<0.001**	**r=0.583; p<0.001**	r=0.066; p=0.11	**r=0.481; p<0.001**
Nitrite (μM)	r=-0.064; p=0.027	**r=-0.210; p<0.001**	r=-0.060; p=0.14	**r=-0.242; p<0.001**	r=-0.063; p=0.13	**r=-0.161; p<0.001**
Nitrite (μM/mM Creatinine)	r=-0.028; p=0.337	**r=-0.671; p<0.001**	r=-0.064; p=0.12	**r=-0.689; p<0.001**	r=0.019; p=0.64	**r=-0.652; p<0.001**

BMI, body mass index; CKD-EPI eGFR, Chronic Kidney Disease Epidemiology Collaboration no race formula estimated glomerular filtration rate; DBP, diastolic blood pressure; K$^+$, potassium; Na$^+$, sodium; SBP, systolic blood pressure; UACR, urinary albumin-creatinine ratio. Significant p values according to the Bonferroni correction are reported in bold.

Table S3: Pearson correlations of possible determinant of uromodulin in the PURE cohort.

	Uromodulin/creatinine (mg/g)	Uromodulin (µg/mL)
Age (years)	r=0.035; p=0.12	r=0.009; p=0.70
BMI (kg/m^2)	r=-0.053; p=0.019	r=-0.008; p=0.74
Waist circumference	**r=-0.075; p=0.001**	r=0.015; p=0.50
rs4293393 (TT; TC; CC)	r=-0.052; p=0.034	r=-0.069; p=0.005
rs12917707 (GG; TG; TT)	**r=-0.097; p<0.001**	**r=-0.080; p=0.001**
rs12446492 (TT; TA; AA)	r=0.011; p=0.67	r=0.010; p=0.69
Clinic SBP (mmHg)	r=-0.028; p=0.22	r=-0.063; p=0.006
Clinic DBP (mmHg)	r=-0.038; p=0.099	**r=-0.079; p=0.001**
Uric acid$_{spot}$ (mg/dL)	**r=0.116; p<0.001**	**r=0.448; p<0.001**
eGFR (ml/min/1.73m^2)	r=0.040; p=0.11	r=-0.014; p=0.59
UACR (mg/mmol)	r=0.025; p=0.29	**r=-0.372; p<0.001**
Renin (pg/ml)	r=-0.032; p=0.17	**r=0.082; p<0.001**
Urinary$_{spot}$ Na$^+$ (mmol/l)	**r=0.087; p<0.001**	**r=0.304; p<0.001**
Urinary$_{spot}$ K$^+$ (mmol/l)	r=0.003; p=0.90	**r=0.582; p<0.001**

BMI, body mass index; CKD-EPI eGFR, Chronic Kidney Disease Epidemiology Collaboration no race formula estimated glomerular filtration rate; DBP, diastolic blood pressure; HTN, hypertension; K$^+$, potassium; Na$^+$, sodium; SBP, systolic blood pressure; UACR, urinary albumin-creatinine ratio. Significant p values according to the Bonferroni correction are reported in bold.

Table S4. Multiple regression analyses with urinary uromodulin/creatinine as the main dependent variable in the African-PREDICT cohort.

Total group

	Uromodulin/crea (mg/g) N=1197			Urine creatinine (mg/dL) N=1198		Uromodulin (µg/mL) N=1194		
	Adj. R^2 0.14			Adj. R^2 0.08		Adj. R^2 0.19		
	Std. ß	P	sr^2	Std. ß	P	Std. ß	P	sr^2
Ethnicity (black; white)	0.038	0.33	<0.001	-0.008	0.84	0.049	0.20	0.001
Sex (female; male)	**-0.232**	**<0.001**	**0.047**	**0.141**	**<0.001**	-0.061	0.052	0.003
Age (years)	**0.102**	**0.002**	**0.009**	-0.089	0.009	-0.028	0.38	0.001
BMI (kg/m^2)	-0.012	0.70	<0.001	0.052	0.12	0.023	0.45	<0.001
eGFR (ml/min/1.73m^2)	**0.145**	**<0.001**	**0.018**	**-0.111**	**0.001**	0.043	0.18	0.002
UACR (mg/mmol)	-0.059	0.061	0.003			**-0.347**	**<0.001**	**0.113**
Renin-S (pmol/L)	0.085	0.013	0.006	**0.227**	**<0.001**	**0.178**	**<0.001**	**0.025**
Estimated salt intake (g/day)	0.037	0.25	0.001	-0.044	0.18	-0.002	0.95	<0.001
Plant protein intake (g)	0.013	0.69	<0.001	-0.030	0.39	-0.009	0.78	<0.001
rs4293393 (TT: CT: CC)	-0.075	0.069	0.003	0.004	0.93	-0.028	0.48	<0.001
rs12917707 (GG; TG; TT)	**-0.177**	**<0.001**	**0.017**	-0.031	0.46	**-0.148**	**<0.001**	**0.012**
rs12446492 (TT: AT: AA)	**-0.130**	**<0.001**	**0.014**	0.047	0.17	-0.044	0.17	0.002

Black

	Uromodulin/crea (mg/g) N=604			Urine creatinine (mg/dL) N=604		Uromodulin (µg/mL) N=603		
	Adj. R^2 0.08			Adj. R^2 0.08		Adj. R^2 0.17		
	Std. ß	P	sr^2	Std. ß	P	Std. ß	P	sr^2
Sex (female; male)	**-0.229**	**<0.001**	**0.039**	**0.153**	**0.004**	-0.029	0.58	0.001
Age (years)	**0.106**	**0.036**	**0.010**	-0.091	0.073	-0.003	0.95	<0.001
BMI (kg/m^2)	-0.012	0.82	<0.001	0.090	0.090	0.059	0.24	0.003
eGFR (ml/min/1.73m^2)	**0.173**	**<0.001**	**0.027**	**-0.177**	**<0.001**	0.019	0.68	<0.001
UACR (mg/mmol)	**-0.131**	**0.006**	**0.016**			**-0.389**	**<0.001**	**0.142**
Renin-S (pmol/L)	0.100	0.039	0.009	**0.193**	**<0.001**	**0.155**	**0.001**	**0.022**
Estimated salt intake (g/day)	0.076	0.11	0.006	-0.068	0.15	0.005	0.90	<0.001
Plant protein intake (g)	0.040	0.43	0.001	0.008	0.87	0.016	0.74	<0.001
rs4293393 (TT: CT: CC)	-0.056	0.26	0.003	0.018	0.72	-0.013	0.78	<0.001
rs12917707 (GG; TG; TT)	-0.058	0.24	0.003	0.013	0.79	-0.054	0.25	0.003
rs12446492 (TT: AT: AA)	-0.090	0.058	0.008	0.057	0.22	-0.015	0.74	<0.001

White

	Uromodulin/crea (mg/g) N=593			Urine creatinine (mg/dL) N=594		Uromodulin (µg/mL) N=591		
	Adj. R^2 0.23			Adj. R^2 0.09		Adj. R^2 0.19		
	Std. ß	P	sr^2	Std. ß	P	Std. ß	P	sr^2
Sex (female; male)	**-0.220**	**<0.001**	**0.039**	**0.150**	**0.002**	-0.069	0.13	0.004
Age (years)	**0.096**	**0.027**	**0.008**	**-0.098**	**0.038**	-0.068	0.13	0.004
BMI (kg/m^2)	0.007	0.86	<0.001	0.006	0.89	0.002	0.96	<0.001
eGFR (ml/min/1.73m^2)	**0.120**	**0.005**	**0.012**	-0.059	0.20	0.056	0.20	0.003
UACR (mg/mmol)	0.028	0.51	<0.001			**-0.297**	**<0.001**	**0.081**
Renin-S (pmol/L)	0.049	0.23	0.002	**0.243**	**<0.001**	**0.193**	**<0.001**	**0.034**
Estimated salt intake (g/day)	0.012	0.77	<0.001	-0.012	0.80	0.008	0.86	<0.001
Plant protein intake (g)	-0.018	0.66	<0.001	-0.075	0.097	-0.047	0.27	0.002
rs4293393 (TT: CT: CC)	-0.081	0.48	<0.001	0.052	0.68	0.025	0.83	<0.001
rs12917707 (GG; TG; TT)	**-0.250**	**0.029**	**0.007**	-0.112	0.37	**-0.268**	**0.023**	**0.009**
rs12446492 (TT: AT: AA)	**-0.142**	**0.001**	**0.016**	0.045	0.35	-0.062	0.17	0.003

BMI, body mass index; CKD-EPI eGFR, Chronic Kidney Disease Epidemiology Collaboration no race formula estimated glomerular filtration rate; UACR, urinary albumin-creatinine ratio. Significant p values according to the Bonferroni correction are reported in bold.

Table S5: Multiple regression analyses with urinary uromodulin/creatinine as the main dependent variable in the PURE cohort.

Independent variables	Uromodulin/creatinine (mg/g) N=1943			Urine creatinine (mg/dL) N=1943			Uromodulin (µg/mL) N=1943		
	Adj. R^2 0.04			Adj. R^2 0.05			Adj. R^2 0.15		
	Std. ß	P	sr^2	Std. ß	P	sr^2	Std. ß	P	sr^2
Sex (female; male)	**-0.188**	**<0.001**	**2.40**	**0.219**	**<0.001**	**3.29**	-0.014	0.64	0.01
Age (years)	**0.091**	**0.002**	**0.66**	**-0.080**	**0.006**	**0.52**	0.057	0.043	0.26
BMI (kg/m²)	**-0.108**	**<0.001**	**0.91**	**0.118**	**<0.001**	**1.08**	-0.031	0.27	0.08
eGFR (ml/min/1.73m²)	**0.123**	**<0.001**	**1.01**	**-0.147**	**<0.001**	**1.46**	<0.001	0.99	<0.001
UACR (mg/mmol)	<0.001	0.99	<0.001				**-0.377**	**<0.001**	**13.7**
Renin (pg/ml)	-0.035	0.19	0.12	**0.137**	**<0.001**	**1.83**	**0.069**	**0.007**	**0.46**
rs4293393 (TT: TC: CC)	-0.022	0.44	0.04	-0.039	0.18	0.13	-0.037	0.17	0.11
rs12917707 (GG: TG: TT)	**-0.088**	**0.002**	**0.67**	-0.004	0.89	0.001	-0.055	0.042	0.26
rs12446492 (TT: TA: AA)	0.007	0.79	0.01	0.012	0.64	0.02	0.003	0.89	0.001

All independent variables were baseline measures.
BMI, body mass index; CKD-EPI eGFR, Chronic Kidney Disease Epidemiology Collaboration no race formula estimated glomerular filtration rate; UACR, urinary albumin-creatinine ratio Significant p values according to the Bonferroni correction are reported in bold.

Table S6: Multiple regression with SNPs separately entered into models in the PURE cohort.

Independent variables	Model 1 N=1943 Adj. R^2 0.03			Model 2 N=1943 Adj. R^2 0.04			Model 3 N=1943 Adj. R^2 0.03		
	Std. ß	*P*	sr^2	Std. ß	*P*	sr^2	Std. ß	*P*	sr^2
Sex (female; male)	**-0.189**	**<0.001**	**0.024**	**-0.188**	**<0.001**	**0.024**	**-0.188**	**<0.001**	**0.024**
Age (years)	**0.092**	**0.002**	**0.007**	**0.091**	**0.002**	**0.007**	**0.092**	**0.002**	**0.007**
BMI (kg/m^2)	**-0.109**	**<0.001**	**0.009**	**-0.109**	**<0.001**	**0.009**	**-0.110**	**<0.001**	**0.009**
eGFR (ml/min/1.73m^2)	**0.123**	**<0.001**	**0.010**	**0.123**	**<0.001**	**0.010**	**0.123**	**<0.001**	**0.010**
UACR (mg/mmol)	-0.001	0.96	<0.001	<0.001	0.99	<0.001	-0.003	0.92	<0.001
Renin (pg/ml)	-0.035	0.20	0.001	-0.036	0.19	0.001	-0.034	0.22	0.001
rs4293393 (TT: TC: CC)	-0.054	0.045	0.003	N/A			N/A		
rs12917707 (GG: TG: TT)	N/A			**-0.096**	**<0.001**	**0.009**	N/A		
rs12446492 (TT: TA: AA)	N/A			N/A			0.007	0.78	<0.001

All independent variables were baseline measures.
BMI, body mass index; CKD-EPI eGFR, Chronic Kidney Disease Epidemiology Collaboration no race formula estimated glomerular filtration rate; UACR, urinary albumin-creatinine ratio
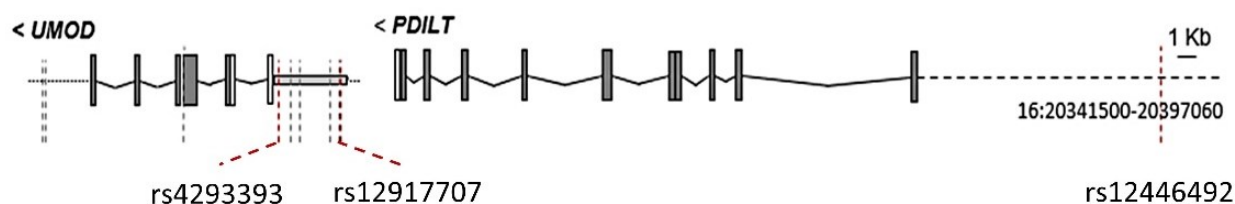Model 1 includes rs4293393 (*UMOD*) SNP
Model 2 includes rs12917707 (*UMOD*) SNP
Model 3 includes rs12446492 (*PDILT*) SNP
Significant p values according to the Bonferroni correction are reported in bold.

Supplementary Figures

rs	Gene	REF/ALT	effect allele	OR or beta for effect allele	trait	population	reference
rs4293393	*UMOD*	A/G	A	1.25	CKD and serum creatinine levels	38765 European (8097 European in replication)	Gudbjartsson DF (20686651)
rs4293393	*UMOD*	A/G	A	0.38 unit incr	Systolic blood pressure	526001 European	Plotnikov D (35762941)
rs12917707	*UMOD/ PDILT*	G/T	T	0.0266 unit incr	Glomerular filtration rate in diabetics (creatinine)	11522 European (replication: 4955 European)	Pattaro C (26831199) 2016
rs12917707	*UMOD/ PDILT*	G/T	T	0.0152 unit incr	Glomerular filtration rate in non diabetics (creatinine)	2826 European	Delgado GE (28242751) 2017
rs12917707	*UMOD/ PDILT*	G/T	T	0.0757 unit decr / 0.0606 unit decr / 0.0762 unit incr / 0.0185 unit decr / 0.0288 unit decr	Creatinine levels / Cystatin C levels / eGFR / Urate levels / Urea levels	342376 European / **6016 African unspecified** / 7339 South Asian	Sinnott-Armstrong N (33462484) 2021
rs12917707	*UMOD/ PDILT*	G/T	T	/	CKD	67093 European (replication: 22982 European)	Kottgen A (20383146) 2010
rs12917707	*UMOD/ PDILT*	G/T	T/G	0.02 ml/min incr / 1.25	Renal function / CKD	19877 European (replication: 18247 NR, 3219 European)	Kottgen A (19430482) 2009
rs12917707	*UMOD/ PDILT*	G/T	T	0.14 unit decr	Kidney function decline traits	45530 European (replication: 18028 European)	Gorski M (25493955) 2014
rs12917707	*UMOD/ PDILT*	G/T	T	0.32 unit decr	Urinary uromodulin levels	10884 European	Olden M (24578125) 2014
rs12446492	*PDILT*	T/A	A	0.15 unit decr	Urinary uromodulin levels	10884 European	Olden M (24578125) 2014

Figure S1: Position of the genotyped variants rs12917707, rs4293393 and rs12446492 along the adjacent *UMOD* and *PDILT* genes on chromosome 16 and reported GWAS associated as found in the GWAS catalogue (https://www.ebi.ac.uk/gwas/home).
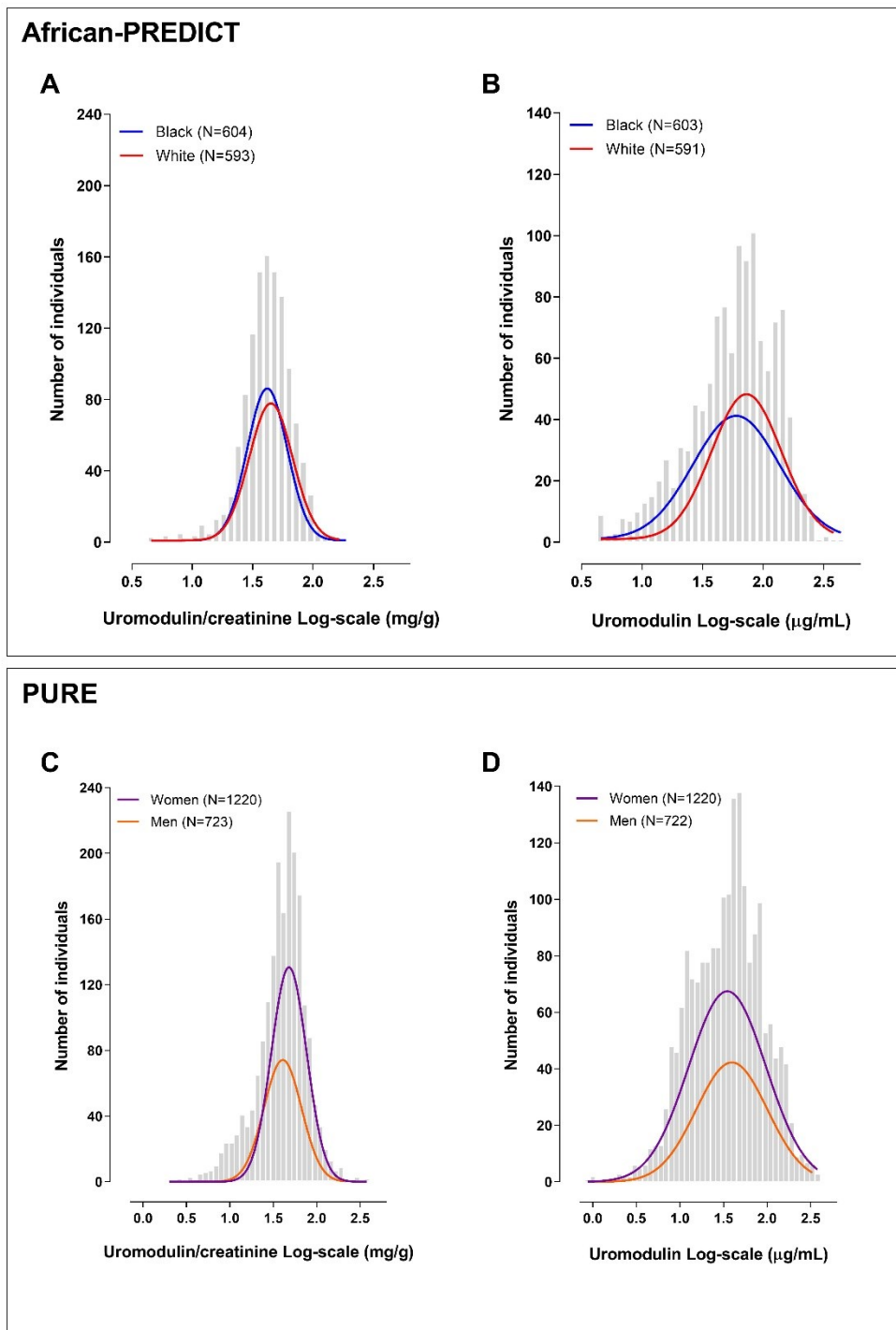
Figure S2. Distribution of urinary UMOD levels Top row: African-PREDICT uromodulin/creatinine (A) and absolute UMOD concentrations (B); bottom row: PURE UMOD/creatinine (C) and absolute UMOD concentrations (D).
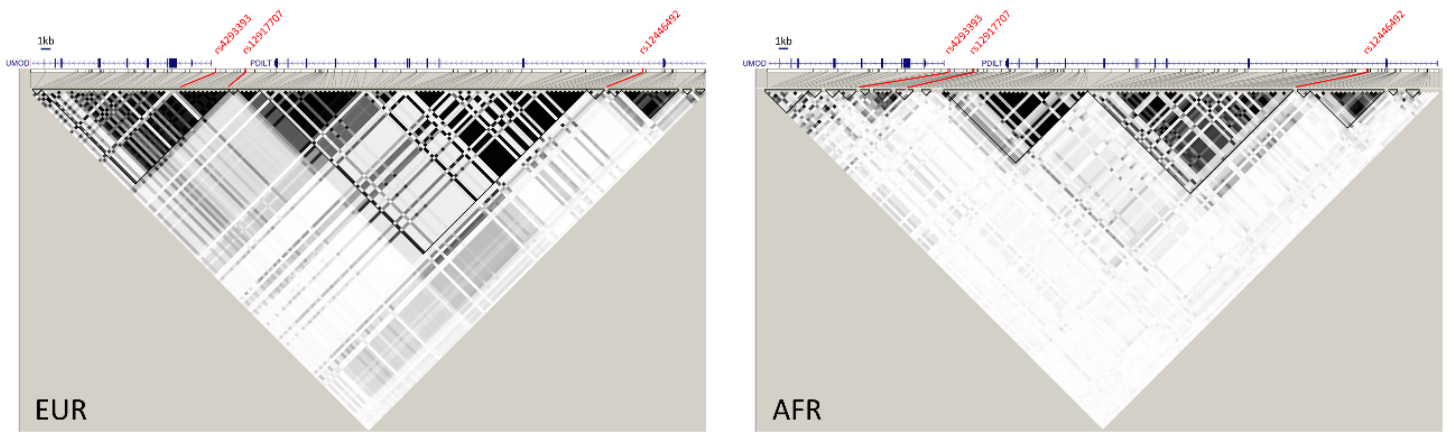
Figure S3: Linkage disequilibrium map of the *UMOD-PDILT* locus based on 1000 Genomes Project summary data for African (n=661) and European (n=502) populations. R^2 values are indicated and the three SNPs of interest are highlighted on the graph. Graph generated using Haploview 4.2.
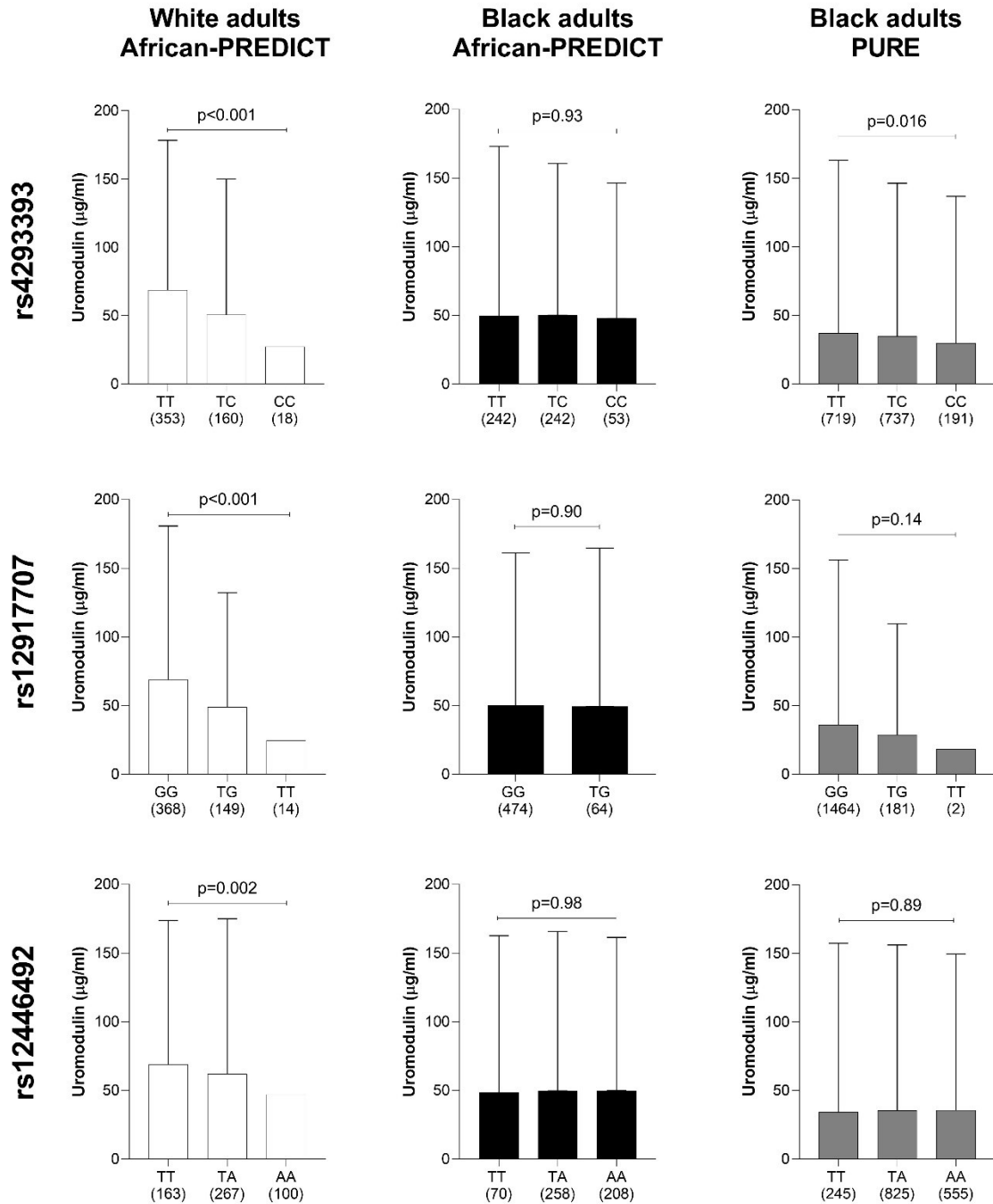
Figure S4: Comparison of absolute urinary uromodulin concentrations levels according to genotypes at rs4293393 (*UMOD*), rs12917707 (*UMOD*) and rs12446492 (*PDILT*) in ● black and ○ white adults from the African-PREDICT study; and ● black adults from the PURE study. Data presented as geometric mean and 95th percentile. p values shown for Welch's ANOVA comparing uromodulin and uromodulin/creatinine across groups.
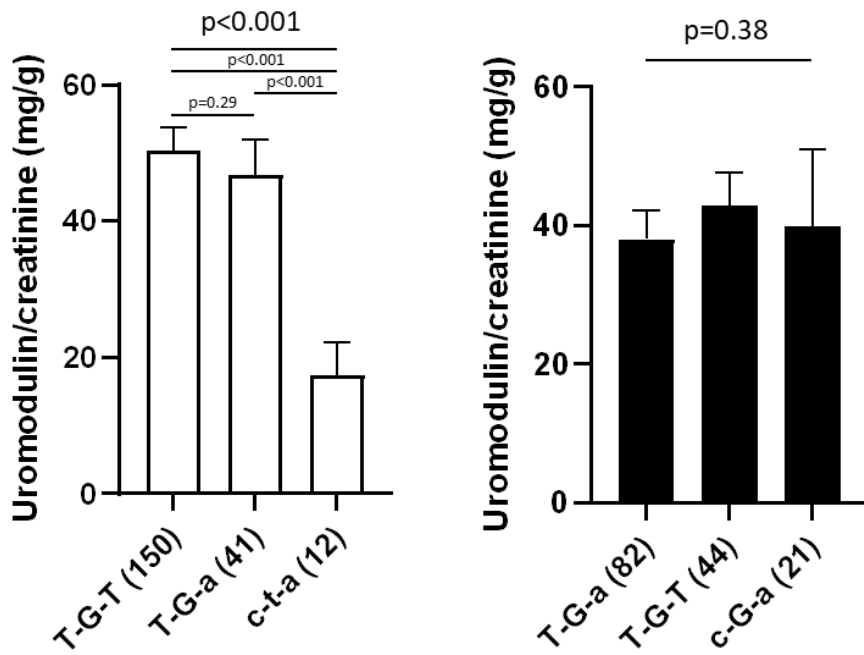
Figure S5: Comparison of urinary uromodulin/creatinine levels in individuals homozygous for haplotypes defined by rs4293393 (*UMOD*), rs12917707 (*UMOD*) and rs12446492 (*PDILT*) in ● black and ○ white adults from the African-PREDICT study. Data presented as geometric mean and 95th percentile. p values shown for Welch's ANOVA comparing uUMOD/creatinine across groups and for Bonferroni multiple comparison test.
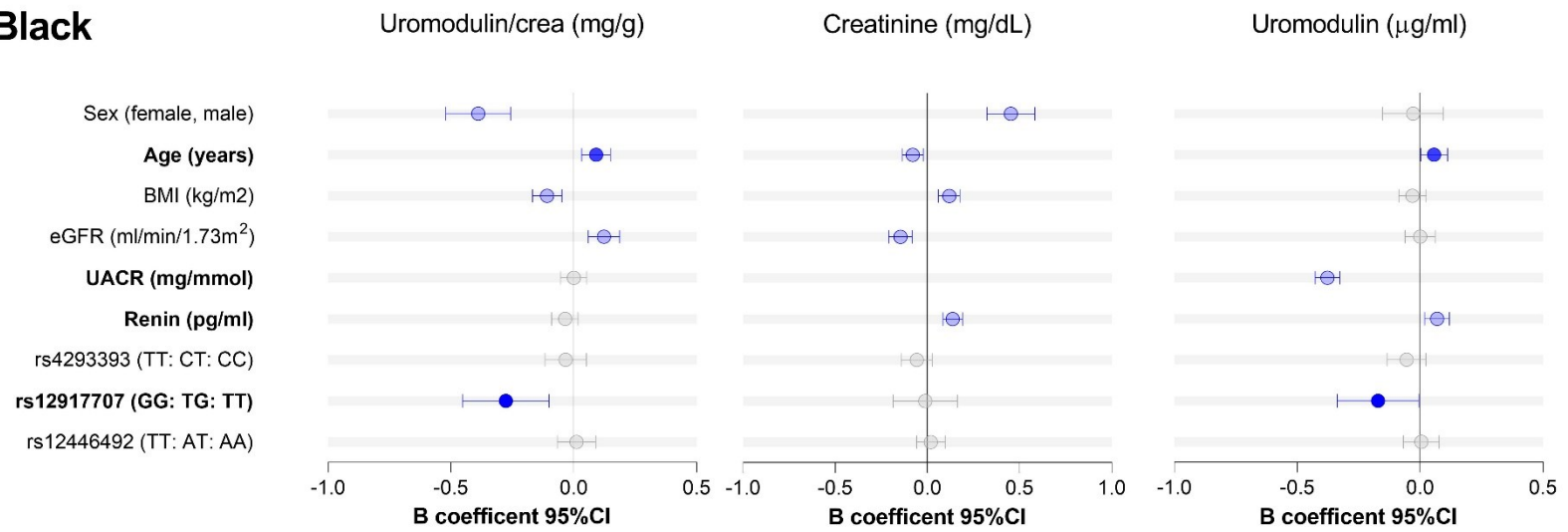
Figure S6. Determinants of urinary uromodulin in the PURE-NWP-SA study. Multiple regression analysis in the PURE study population. Continues variables were standardized by creating z-variables, which were included into multiple regression models for this Forest plot.

STrengthening the REporting of Genetic Association studies (STREGA) reporting recommendations, extended from STROBE Statement

Item	Item no	STROBE Guideline	Extension for Genetic Association Studies (STREGA)	Page no
Title and Abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract.		Page 2
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found.		Page 2
Introduction				
Background rationale	2	Explain the scientific background and rationale for the investigation being reported.		Done
Objectives	3	State specific objectives, including any pre-specified hypotheses	*State if the study is the first report of a genetic association, a replication effort, or both.*	Done
Methods				
Study design	4	Present key elements of study design early in the paper.		Page 5 & SI
Setting	5	Describe the setting, locations and relevant dates, including periods of recruitment, exposure, follow-up and data collection.		Page 5 & SI
Participants	6	(a) **Cohort study –** Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up. **Case–control study –** Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls. **Cross-sectional study –** Give the eligibility criteria, and the sources and methods of selection of participants.	*Give information on the criteria and methods for selection of subsets of participants from a larger study, when relevant.*	Page 5 & SI
		(b) **Cohort study** – For matched studies, give matching criteria and number of exposed and unexposed. **Case–control study –** For matched studies, give matching criteria and the number of controls per case.		
Variables	7	(a) Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.	*(b) Clearly define genetic exposures (genetic variants) using a widely –used nomenclature system. Identify variables likely to be associated with population stratification*	Page 5 & SI

				(confounding by ethnic origin).	
Data sources measurement	8*	(a) For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group.	*(b) Describe laboratory methods, including source and storage of DNA, genotyping methods and platforms (including the allele calling algorithm used, and its version), error rates and call rates. State the laboratory /centre where genotyping was done. Describe comparability of laboratory methods if there is more than one group. Specify whether genotypes were assigned using all of the data from the study simultaneously or in smaller batches.*	Page 5 & SI
Bias	9	(a) Describe any efforts to address potential sources of bias.	*(b) For quantitative outcome variables, specify if any investigation of potential bias resulting from pharmacotherapy was undertaken. If relevant, describe the nature and magnitude of the potential bias, and explain what approach was used to deal with this.*	Done
Study size	10	Explain how the study size was arrived at.		Page 5
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why.	*If applicable, describe how effects of treatment were dealt with.*	n/a
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding.	*State software version used and options (or settings) chosen.*	SI
		(b) Describe any methods used to examine subgroups and interactions.		SI
		(c) Explain how missing data were addressed.		
		(d) **Cohort study –** If applicable, explain how loss to follow-up was addressed. **Case–control study –** If applicable, explain how matching of cases and controls was addressed.		

		Cross-sectional study – If applicable, describe analytical methods taking account of sampling strategy.		
		(e) Describe any sensitivity analyses.		
			(f) State whether Hardy- Weinberg equilibrium was considered and, if so, how.	n/a
			(g) Describe any methods used for inferring genotypes or haplotypes.	SI
			(h) Describe any methods used to assess or address population stratification.	SI
			(i) Describe any methods used to address multiple comparisons or to control risk of false positive findings.	SI
			(j) Describe any methods used to address and correct for relatedness among subjects.	SI
Results				
Participants	13*	(a) Report the numbers of individuals at each stage of the study – e.g. numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up and analysed.	*Report numbers of individuals in whom genotyping was attempted and numbers of individuals in whom genotyping was successful.*	Page 6-7
		(b) Give reasons for non-participation at each stage.		
		(c) Consider use of a flow diagram.		
Descriptive data	14*	(a) Give characteristics of study participants (e.g. demographic, clinical, social) and information on exposures and potential confounders.	*Consider giving information by genotype.*	Cohorts were described in Tables and previously published
		(b) Indicate the number of participants with missing data for each variable of interest.		
		(c) **Cohort study** – Summarize follow-up time, e.g. average and total amount.		
Outcome data	15*	**Cohort study** – Report numbers of outcome events or summary measures over time.	*Report outcomes (phenotypes) for each genotype category over time*	Outcome (uUMOD) reported for each

		Case–control study – Report numbers in each exposure category, or summary measures of exposure.	***Report numbers in each genotype category***	genotype and for clinical factors
		Cross-sectional study – Report numbers of outcome events or summary measures.	***Report outcomes (phenotypes) for each genotype category***	
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g. 95% confidence intervals). Make clear which confounders were adjusted for and why they were included.		Page 10
		(b) Report category boundaries when continuous variables were categorized.		
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period.		
			(d) Report results of any adjustments for multiple comparisons.	Done
Other analyses	17	(a) Report other analyses done – e.g. analyses of subgroups and interactions, and sensitivity analyses.		Done
			(b) If numerous genetic exposures (genetic variants) were examined, summarize results from all analyses undertaken.	Done
			(c) If detailed results are available elsewhere, state how they can be accessed.	n/a
Discussion				
Key results	18	Summarize key results with reference to study objectives.		Done Page 11-12
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias.		Done Page 13
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence.		Done
Generalizability	21	Discuss the generalizability (external validity) of the study results.		Done
Other information				
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based.		Page 14-15

STROBE: STtrengthening the Reporting of Observational Studies in Epidemiology

*Give information separately for cases and controls in case–control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.