Human C-terminal *CUBN* variants associate with chronic proteinuria and normal renal function

Supplemental Material

Supplementary Text: Study populations and NGS panels

Supplemental Figure 1: Proportion and phenotypes of *CUBN* variant carriers in genetic kidney disease cohort I and chronic PU cohort.
Supplemental Figure 2: Proportions of mutation types found in the three cohorts of this study.

Supplemental Table 1: Clinical and biological data of patients with or without biallelic pathogenic *CUBN* variants
Supplemental Table 2: List of patients with biallelic filtered *CUBN* variants.
Supplemental Table 3: Additional information on available renal biopsies.
Supplemental Table 4: List of *CUBN* variants identified in the cohort and additional cases.
Supplemental Table 5: Description of structural models for the identified missense variants.
Supplemental Table 6: Patients with identified hemizygous mutations in Dent's disease genes *CLCN5* and *OCRL1*.
Supplemental Table 7: Individual with one filtered variant and one GWAS variant in trans.
Supplemental Table 8: Association of *CUBN* SNPs with albuminuria and kidney function (eGFR) in Europeans (with and without diabetes).
Supplemental Table 9: Association of *CUBN* SNPs with albuminuria and kidney function (eGFR) in Europeans stratified for diabetes status.
Supplemental Table 10: Additive genetic model of *CUBN* SNPs for albuminuria.
Supplemental Table 11: Recessive genetic model of *CUBN* SNPs for albuminuria.
Supplemental Table 12: Genotypic distribution of the *CUBN* SNPs and power estimate for albuminuria.

Supplementary Text:

Study populations and NGS panels

Between 2008 and 2019, DNA samples were collected for diagnostic purposes in our reference center for hereditary renal disease at the Department of Genetics of Hôpital Necker Enfants-malades. Patients with suspected genetic renal disease were initially recruited through adult and pediatric nephrology departments throughout France (plus some additional countries) and, after obtaining informed consent, blood samples were addressed to our center (1, 2). Familial data were collected by the clinicians in charge of the patients, including pedigrees, data on consanguinity and data on parents and siblings. Clinical data was obtained from the patient files and from standardized questionnaires filled in by the referring clinicians or by asking back to the clinicians. Before 2017 genetic diagnostics were performed with Sanger sequencing of individual genes or with the help of smaller panels (e.g. only AS genes), whereas after 2017 all DNA samples were subjected to the so-called „Renome“ panel. Based on this, two cohorts were selected out of the larger cohort for this study:

Genetic kidney disease I: Between 2017 and 2019, 759 individuals were subjected to NGS sequencing using the Renome panel and retrospectively analyzed. The Renome panel includes 309 genes (also *CUBN*), whose mutations are responsible for wide range of renal diseases (see Methods for disease and for the control group).

Chronic PU cohort: 107 individuals were selected because they exhibited isolated chronic proteinuria and because previous genetic testing in the “pre-Renome” era (using Sanger sequencing or smaller panels (AS or SRNS)) had not yielded a genetic diagnosis. In particular, the *CUBN* gene had not been sequenced before. Isolated chronic proteinuria was mostly below the nephrotic range (mostly 0.5-3 g/d, serum albumin > 30 g/l or no edema) based on 24 hour-urine, urinary protein-to-creatinine ratio, urinary albumin or protein dipstick measurements. In some cases, the presence of proteinuria was noted in the questionnaire without any values. Whenever biopsy data was available, non-genetic forms of glomerulopathies such as membranous nephropathy, C3 glomerulonephritis and lupus nephrits were excluded. Patients were re-sequenced with a custom-made next-generation sequencing panel targeting seventeen genes known to have functions in the proximal tubule, incl. *CUBN*.

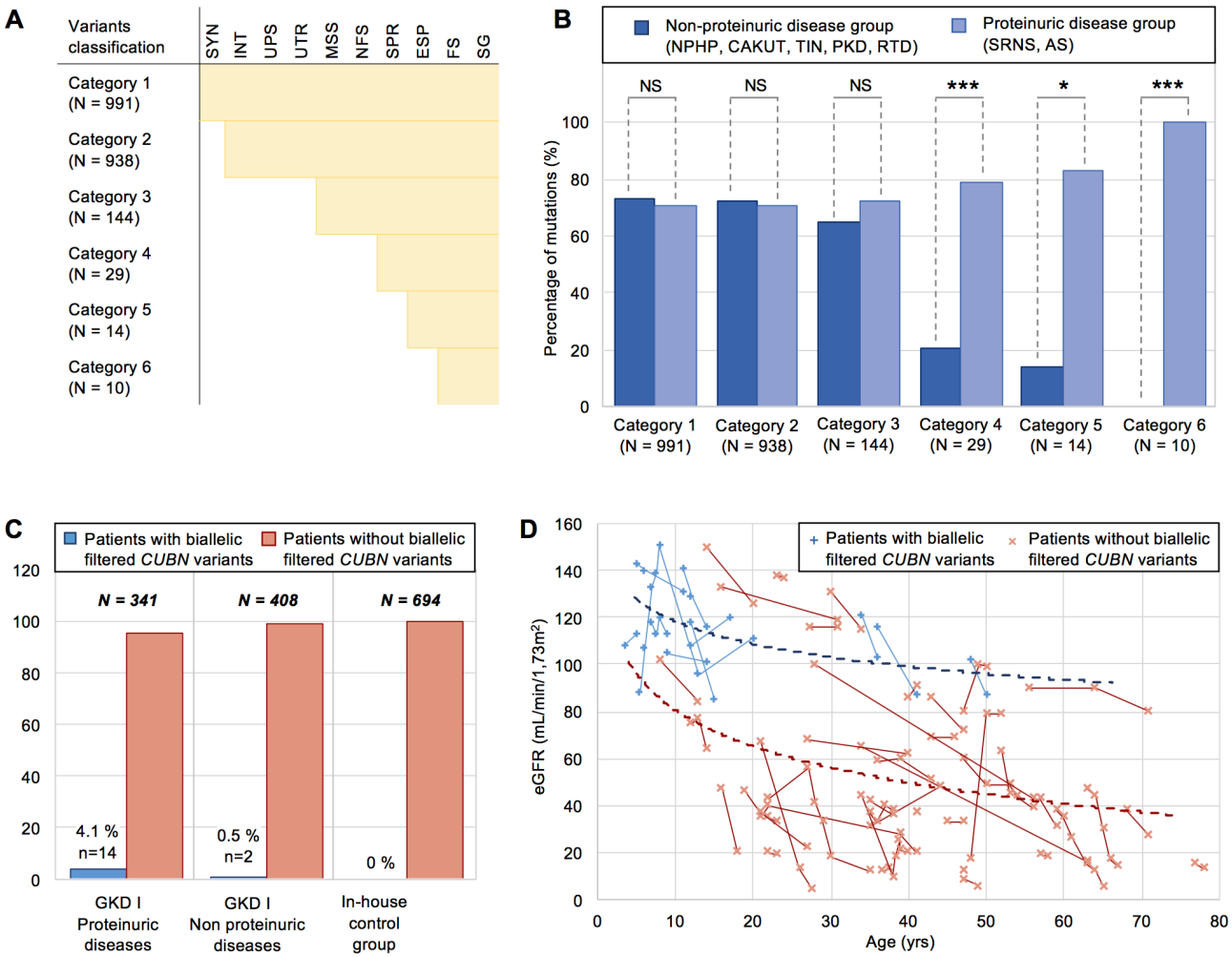
Our third cohort, *Genetic kidney disease cohort II,* is a German and Austrian cohort consisting of 1350 patients primarily with suspected SRNS, FSGS and AS. The DNA samples were collected between 2009 and 2019. Like in Genetic kidney cohort I, patients were sequenced in the diagnostic

setting (3). The applied NGS panels contained up to 324 genes, mostly with high expression in the glomerulus ("GLOM" panel) (4). As *CUBN* was also included in the panels, patients were retrospectively analyzed with regard to the *CUBN* gene. All the clinical information was provided and collected by clinicians prescribing the genetic testing.

Recruitment centers	French multi-center		German/Austrian multi-center
Renal genetics unit	Department of Genetics Hôpital Necker		Human Genetics Bioscientia
Cohorts	Genetic kidney cohort I	Chronic PU cohort	Genetic kidney cohort II
Renal diseases	SRNS, AS, NPH, CAKUT, TIN, PKD, RTD	SRNS, AS	SRNS, FSGS, AS
Sample size (n)	759	107	1350
NGS panel (incl. *CUBN*)	309 renal diseae genes (Renome)	17 proximal tubular genes	up to 324 renal genes (GLOM)

References:

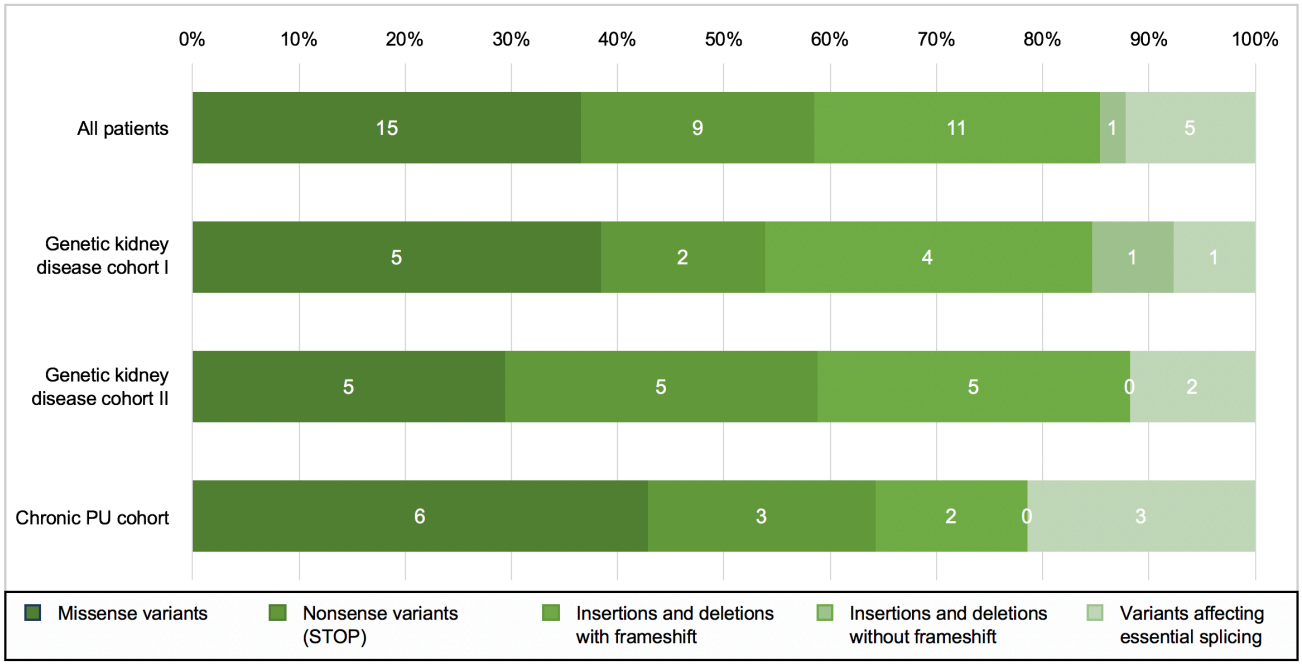
1. Gribouval O, Boyer O, Hummel A, Dantal J, Martinez F, Sberro-Soussan R, Etienne I, Chauveau D, Delahousse M, Lionet A, et al. Identification of genetic causes for sporadic steroid-resistant nephrotic syndrome in adults. *Kidney Int.* 2018;94(5):1013-22.
2. Heidet L, Moriniere V, Henry C, De Tomasi L, Reilly ML, Humbert C, Alibeu O, Fourrage C, Bole-Feysot C, Nitschke P, et al. Targeted Exome Sequencing Identifies PBX1 as Involved in Monogenic Congenital Anomalies of the Kidney and Urinary Tract. *J Am Soc Nephrol.* 2017;28(10):2901-14.
3. Lu H, Galeano MCR, Ott E, Kaeslin G, Kausalya PJ, Kramer C, Ortiz-Bruchle N, Hilger N, Metzis V, Hiersche M, et al. Mutations in DZIP1L, which encodes a ciliary-transition-zone protein, cause autosomal recessive polycystic kidney disease. *Nat Genet.* 2017;49(7):1025-34.
4. Rinschen MM, Godel M, Grahammer F, Zschiedrich S, Helmstadter M, Kretz O, Zarei M, Braun DA, Dittrich S, Pahmeyer C, et al. A Multi-layered Quantitative In Vivo Expression Atlas of the Podocyte Unravels Kidney Disease Candidate Genes. *Cell Rep.* 2018;23(8):2495-508.



Supplemental Figure 1. Proportion and renal function of *CUBN* variant carriers in genetic kidney disease cohort I and chronic PU cohort.

(A) Definition of variant categories by sequential removal of variants per type (SYN, synonymous; INT, intronic; UPS, upstream; UTR, untranslated region; MSS, missense; NFS, no-frameshift insertion or deletion; SPR, splicing region; ESP, essential splicing; FS, frameshift insertion or deletion; SG, stop gain mutation). (B) Barplot of the percentage of *CUBN* variants per category found in the proteinuric group (Alport syndrome (AS) and steroid-resistant nephrotic syndrome (SRNS) vs. the non-proteinuric group (nephronophthisis (NPHP), congenital anomalies of the kidney and urinary tract (CAKUT), tubulointerstitial nephritis (TIN), polycystic kidney disease (PKD) and renal tubular dysgenesis (RTD)). The percentage corresponds to the ratio of the number of variants found in the considered group on the total number of variants of this category found in the Genetic kidney disease cohort I (Fisher's exact test: *p < 0.01; *p < 0.001; ***p < 0,0001; NS, not significant). (C) Proportion of patients with biallelic *CUBN* variants passing our filtering criteria (see Methods) in the proteinuric, non-proteinuric group of genetic kidney disease cohort I (GKD I) and a

control group of in-house 694 exomes. The two cases in the non-proteinuric group are the following: one fetus with renal agenesis belonging to the HYP group, and one male case with chronic proteinuria, consistent with *CUBN* deficiency. It should be noted, however, that he and several other family members also seemed to suffer from an autosomal-dominant renal insufficiency, which is why this case was classified as TIN (tubulointerstitial nephritis) in the first place. (D) eGFR trend for individuals of genetic kidney disease cohort I (GKD I) and chronic PU cohort, for which more than one serum creatinine value was available (complementing Figure 4D). eGFR was calculated using Schwartz formula for children based on their height and serum creatinine, and using CKD-EPI formula for adults, based on their age and serum creatinine. Values above the 95% confidence interval were excluded for this graphic representation.



Supplemental Figure 2. Proportions of mutation types found in the three cohorts of this study.

Barplot representing the proportions of different types of *CUBN* variants identified in the Genetic kidney disease cohorts I and II and Chronic PU cohort.

Supplemental Table 1. Clinical and biological data of patients with or without biallelic pathogenic *CUBN* variants (Part 2)							
	Genetic kidney disease cohort I	***Genetic kidney disease cohort II***	***Chronic PU cohort***				
Clinical and biological data	Patients with biallelic filtered *CUBN* variants (95% IC or %)	Patients with biallelic filtered *CUBN* variants (95% IC or %)	All patients (95% IC or %)	Patients with biallelic filtered *CUBN* variants (95% IC or %)	Patients without biallelic filtered *CUBN* variants (95% IC or %)	*P value*	
N	**14**	**13**	**107**	**12**	**95**		
Edema at onset	0 [n=14]	1 (7.6) [n=13]	2 (2.2) [n=89]	0 [n=10]	2 (2.5) [n=79]	NS (1.000)	*(b)*
High blood pressure	0 [n=14]	1 (7.6) [n=13]	25 (27.5) [n=91]	0 [n=9]	25 (30.86) [n=81]	NS (0.0577)	*(b)*
Microscopic hematuria	4 (69.23) [n=13]	0 [n=13]	13 (23.6) [n=55]	1 (11.1) [n=9]	12 (26.1) [n=46]	NS (0.6691)	*(b)*
Gross hematuria	0 [n=13]	5 (38.5) [n=13]	2 (3.6) [n=55]	1 (11.1) [n=9]	1 (2.2) [n=46]	NS (0.3030)	*(b)*
Glycosuria	0 [n=9]	0 [n=10]	*NA*	0 [n=4]	*NA*	*NA*	*(b)*
Hypercalciuria	0 [n=8]	0 [n=10]	*NA*	0 [n=2]	*NA*	*NA*	*(b)*
Urinary α1-microglobulin (detected but normal)	3 (75.0) [n=4]	8 (88.9) [n=9]	*NA*	2 [n=2]	*NA*	*NA*	*(b)*
Urinary β2-microglobulin (detected but normal)	0 [n=7]	0 [n=4]	*NA*	2 (66.7) [n=3]	*NA*	*NA*	*(b)*
Steroids with no response	2 (100) [n=2]	3 (100) [n=3]	11 (91.67) [n=12]	1 (100) [n=1]	10 (90.91) [n=11]	NS (1.000)	*(a)*
ACE inhibitors with no response	7 (100) [n=7]	2 (100) [n=2]	15 (88.2) [n=17]	6 (100) [n=6]	9 (81.8) [n=11]	NS (0.5147)	*(a)*
Transplant	0 [n=14]	0 [n=13]	15 (14.42) [n=104]	0 [n=12]	15 (16.30) [n=92]	NS (0.2069)	*(a)*

(a) Two-sample t-test for significance between *CUBN* mutated and non-mutated patients from the chronic PU cohort
(b) Fisher exact test for significance between *CUBN* mutated and non-mutated patients from the chronic PU cohort
[n] indicates the total number of patients with the given information; NA, not available; yr, years; FSGS, focal segmental glomerulosclerosis; ACE, Angiotensin-conversion enzyme; ESRD, end-stage renal disease.
* Proteinuria in g/24h were measured in most patients. For children and adults where only urinary protein to creatinine ratios in spot urine were reported, conversion into g/24h was performed (56). Patients not included here are patients where only albuminuria or protein dipstick was available or patients without further information on the amount of proteinuria.

Supplemental Table 2. List of patients with biallelic filtered *CUBN* variants.							
					Segregation		
	Family No.	**Patient No.**	**Gene**	**Exon (zygosity)**	**Father**	**Mother**	**Variants**
Genetic Kidney Disease Cohort I	1	N3791	*CUBN*	46 (het)	WT	het	c.7165dup – p.Asp2399Glyfs*19
				66 (het)	het	WT	c.10559G>A – p.Gly3520Asp
	2	A4256	*CUBN*	45 (het)	*NA*	*NA*	c.6928_6934delGAGGTTA – p.Glu2310Cysfs*3
				61 (het)	*NA*	*NA*	c.9725_9726insACCT – p.Ala3243Profs*9
	3	A3639	*CUBN*	51 (hom)	*NA*	*NA*	c.7968_7969ins17 – p.Leu2656_Pro2657delinsPheValIleProTyrIleThr
	4	N4126	*CUBN*	51 (hom)	*NA*	*NA*	c.7968_7969ins17 – p.Leu2656_Pro2657delinsPheValIleProTyrIleThr
	5	A4306	*CUBN*	2 (het)	*NA*	*NA*	c.164C>T – p.Thr55Met
				40 (het)	*NA*	*NA*	c.6088C>T – p.Arg2030*
	6	**N3998**	*CUBN*	62 (het)	het	WT	c.9956delA – p.Lys3319Argfs*2
				65 (het)	WT	het	c.10474G>T – p.Asp3492Tyr
	6	**N3999**	*CUBN*	62 (het)	het	WT	c.9956delA – p.Lys3319Argfs*2
				65 (het)	WT	het	c.10474G>T – p.Asp3492Tyr
	7	**N4355**	*CUBN*	57 (hom)	het	het	c.9053A>C – p.Tyr3018Ser
	7	**N4356**	*CUBN*	57 (hom)	het	het	c.9053A>C – p.Tyr3018Ser
	7	**N4357**	*CUBN*	57 (hom)	het	het	c.9053A>C – p.Tyr3018Ser
	8	N945	*CUBN*	57 (hom)	het	hom*	c.9053A>C – p.Tyr3018Ser
	9	A4500	*CUBN*	51(het)	*NA*	*NA*	c.7968_7969ins17 – p.Leu2656_Pro2657delinsPheValIleProTyrIleThr
				54 (het)	*NA*	*NA*	c.8493T>A – p.Cys2831*
	10	N4841	*CUBN*	-2_ex41 (hom)	*NA*	*NA*	c.6125–2A>G
	11	A4669	*CUBN*	54 (hom)	*NA*	*NA*	c.8465C>T – p.Pro2822Leu
Genetic	12	CB10094	*CUBN*	38 (het)	het	WT	c.5560G>T – p.Asp1854Tyr

Kidney Disease Cohort II				52 (het)	WT	het	c.8165del – p.Pro2722Hisfs*39
	13	**CB1179**	*CUBN*	38 (hom)	het	het	c.5600del – p.Phe1867Serfs*16
	13	**CB1178**	*CUBN*	38 (hom)	het	het	c.5600del – p.Phe1867Serfs*16
	14	CB8449	*CUBN*	57 (het)	WT	het	c.9053A>C – p.Tyr3018Ser
				57 (het)	het	WT	c.9079G>A – p.Gly3027Arg
	15	CB8549	*CUBN*	37 (het)	WT	het	c.5511dup – p.Gly1838Trpfs*17
				-2_41 (het)	het	WT	c.6125-2A>G
	16	CB1115	*CUBN*	62 (hom)	*NA*	*NA*	c.9922T>C – p.Trp3308Arg
	17	CB2847	*CUBN*	37 (het)	het	WT	c.5428C>T – p.Arg1810*
				50 (het)	WT	het	c.7797C>G – p.Cys2599Trp
	18	CB3447	*CUBN*	33 (het)	*NA*	*NA*	c.4921del – p.Tyr1641Ilefs*17
				62 (het)	*NA*	*NA*	c.9949C>T – p.Gln3317*
	19	CB7512	*CUBN*	-2_ex38 (het)	*NA*	*NA*	c.5549-2A>C
				57 (het)	*NA*	*NA*	c.9053A>C – p.Tyr3018Ser
	20	CB8631	*CUBN*	24 (het)	*NA*	*NA*	c.3473G>A – p.Trp1158*
				57 (het)	*NA*	*NA*	c.9053A>C – p.Tyr3018Ser
	21	CB1211	*CUBN*	37 (het)	*NA*	*NA*	c.5428C>T – p.Arg1810*
				57 (het)	*NA*	*NA*	c.6088C>T – p.Arg2030*
	22	CB1524	*CUBN*	34 (het)	*NA*	*NA*	c.4973del – p.Asn1658Ilefs*23
				57 (het)	*NA*	*NA*	c.9079G>A – p.Gly3027Arg
	23	CB2804	*CUBN*	2_ex6 (het)	*NA*	*NA*	c.489+2_489+4dup
				30 (het)	*NA*	*NA*	c.4459C>T – p.Arg1487*
Chronic PU Cohort	**24**	**N811**	*CUBN*	45 (het)	*NA*	*NA*	c.6926_6928delinsAA – p.Gly2309fs*6
				55 (het)	*NA*	*NA*	c.8707C>T – p.Gln2903*
	24	**N812**	*CUBN*	45 (het)	*NA*	*NA*	c.6926_6928delinsAA – p.Gly2309fs*6
				55 (het)	*NA*	*NA*	c.8707C>T – p.Gln2903*
	25	N846	*CUBN*	27 (het)	*NA*	*NA*	c.3907A>C – p.Asn1303His
				63 (het)	*NA*	*NA*	c.10097C>G – p.Ser3366Cys
	26	N3997	*CUBN*	57 (hom)	*NA*	*NA*	c.9053A>C – p.Tyr3018Ser

	27	**N2187**	*CUBN*	65 (hom)	*NA*	*NA*	c.10474G>T – p.Asp3492Tyr
	27	**N2188**	*CUBN*	39 (het)	*NA*	*NA*	c.5783G>T – p.Gly1928Val
				67 (het)	*NA*	*NA*	c.10825G>C – p.Asp3609His
	28	N3109	*CUBN*	44 (het)	*NA*	*NA*	c.6782T>G – p.Leu2261Arg
				-1_ex50 (het)	*NA*	*NA*	c.7706-1G>T
	29	N3123	*CUBN*	-2_ex41 (het)	*NA*	*NA*	c.6125-2A>G
				2_ex44 (het)	*NA*	*NA*	c.6821+2T>C
	30	N3622	*CUBN*	2_ex44 (het)	*NA*	*NA*	c.6821+2T>C
				54 (het)	*NA*	*NA*	c.8498G>A – p.Trp2833*
	31	N3953	*CUBN*	39 (het)	*NA*	*NA*	c.5913_5916del – p.Thr1972Leufs*10
				2_ex44 (het)	*NA*	*NA*	c.6821+2T>C
	32	N4012	*CUBN*	57 (het)	*NA*	*NA*	c.9053A>C – p.Tyr3018Ser
				67 (het)	*NA*	*NA*	c.10852C>T – p.Arg3618*

NA, not available; het, heterozygous; hom, homozygous.
In bold are related cases (all are siblings except for N2187 and N2188 who are third cousins).
* Mother of N945: Proteinuria ~ 1 g/d; First manifestation during childhood; Normal renal function, No ESRD; 43.1 years old.

Supplemental Table 3. Additional information on available renal biopsies.				
Cohort	**Patient**	**Biopsy report LM or IF**	**Biopsy report EM**	**Conclusion**
Genetic Kidney Disease Cohort I	N4126	No lesion, no deposits	*NA*	No lesion
	A3639	No lesion, no deposits	*NA*	No lesion
	N845	No lesion, no deposits	*NA*	No lesion
	N4841	No lesion, no deposits	normal	No lesion
	N4355	Isolated hypertrophic podocytes, some IgM and C1q deposits	normal	No specific lesion: MCD
	N4357	Normal glomeruli, some C3 deposits	*NA*	No specific lesion: MCD
	A4500	2 out of 20 glomeruli: small size, thickened capsule and vascular pole lesion, no deposits	*NA*	No specific lesion: MCD
	A4256	Mild mesangial hypertrophy, some IgM and IgA deposits	*NA*	No specific lesion: MCD
	A4306	No lesion, no deposits	*NA*	No lesion
Chronic PU Cohort	N3123	No lesion, no deposits	*NA*	No lesion
	N3997	No lesion, no deposits	*NA*	No lesion
	N3953	No lesion, no deposits	*NA*	No lesion
	N3622	Discrete accentuation of mesangium	*NA*	No specific lesion: MCD
	N811	Discrete mesangial fibrosis	*NA*	No specific lesion: MCD
	N812	2 out of 27 glomeruli with synechia	*NA*	No specific lesion: MCD
Genetic Kidney Disease Cohort I	CB1179	7 out of 18 glomeruli with synechiae	synechiae	FSGS at "early stage"
	CB2847	No lesion	*NA*	No specific lesion: MCD
	CB7512	1 out of 22 glomeruli with tip lesion	synechia	No specific lesion: MCD
	CB1211	No glomerular lesion, unspecific tubular lesions	*NA*	No specific lesion: MCD

GKD: Genetic kidney disease cohort; CPU: Chronic PU cohort; EM: electron microscopy; FPE:foot processes effacement; FSGS: focal segmental glomerulosclerosis; GBM: glomerular IF: IF: Immunofluorescence; LM: light microscopy; MCD: minimal change disease; FPE: foot process effacement; GBM: Glomerular basement membrane.
For Table 1, only CB1179 was classified as „Biopsy showing FSGS“.

Supplemental Table 4. List of *CUBN* variants identified in the cohort and additional cases											
	Variant rsID	**cDNA**	**Protein**	**Exon**	**Protein domain**	**Allele frequency (gnomAD)**	**Frequency (hom)**	**Mutation taster**	**PolyPhen 2**	**SIFT**	**Reference (HGMD)**
Missense	rs774556167	c.164C>T	p.Thr55Met	2	N-ter	0.000016	0	DC (0.525)	PSD 0.557	D 0.01	—
	rs200203056	c.3907A>C	p.Asn1303His	27	CUB8	0.000032	0	P (1)	B 0.219	D 0.01	—
		c.5560G>T	p.Asp1854Tyr	38	CUB13	*Absent*	—	P (0.982)	PD 0.946	D 0.01	—
	rs201513648	c.5783G>T	p.Gly1928Val	39	CUB13	0.000035	0	DC (1)	PD 0.998	D 0.00	—
	rs147617753	c.5840C>A	p.Ser1947Tyr	39	CUB13	0.000244	0	P (0.98)	PD 0.995	D 0.00	(26)
	rs779959064	c.6782T>G	p.Leu2261Arg	44	CUB16	0.000004	0	DC (1)	PD 0.963	D 0.00	—
		c.7797C>G	p.Ser2599Trp	50	CUB19	*Absent*	—	DC (1)	PD 0.997	D 0.00	—
	rs776663892	c.8465C>T	p.Pro2822Leu	54	CUB21	0.000032	0	DC (0.98)	PD 0.972	D 0.00	—
	rs370778353	c.9053A>C	p.Tyr3018Ser	57	CUB22	0.000120	0	DC (0.991)	PD 0.976	D 0.00	—
	rs150202444	c.9079G>A	p.Gly3027Arg	57	CUB22	0.000198	0	DC (1)	PD 1.00	D 0.00	—
	rs752843169	c.9922T>C	p.Trp3308Arg	62	CUB25	0.000139	0	DC (1)	PD 1.00	D 0.00	—
	rs201157846	c.10097C>G	p.Ser3366Cys	63	CUB25	0.000032	0	DC (0.759)	PD 0.999	D 0.00	—
	rs764917718	c.10474G>T	p.Asp3492Tyr	65	CUB26	0.000064	0	DC (0.999)	PD 0.993	D 0.02	—
		c.10559G>A	p.Gly3520Asp	66	CUB27	*Absent*	—	DC (1)	PD 0.992	D 0.00	—
	rs775742147	c.10825G>C	p.Asp3609His	67	CUB27	0.000008	0	P (1)	PSD 0.601	D 0.02	—
Nonsense		c.3473G>A	p.Trp1158*	24	CUB6	*Absent*	—	—	—	—	—

	rs145661597	c.4459C>T	p.Arg1487*	30	CUB9	0.000049	0	—	—	—	—
	rs143944436	c.5428C>T	p.Arg1810*	37	CUB12	0.000142	0	—	—	—	—
	rs374417889	c.6088C>T	p.Arg2030*	40	CUB14	0.000039	0	—	—	—	—
		c.8493T>A	p.Cys2831*	54	CUB21	*Absent*	—	—	—	—	—
	rs778949688	c.8498G>A	p.Trp2833*	54	CUB21	0.000008	0	—	—	—	—
		c.8707C>T	p.Gln2903*	55	CUB21	*Absent*	—	—	—	—	—
		c.9949C>T	p.Gln3317*	62	CUB25	*Absent*	—	—	—	—	—
	rs750303687	c.10852C>T	p.Arg3618*	67	CUB27	0.000036	0	—	—	—	—
Insertion and deletions		c.4921del	p.Tyr1641Ilefs*17	33	CUB11	*Absent*	—	—	—	—	—
		c.4973del	p.Asn1658Ilefs*23	34	CUB11	*Absent*	—	—	—	—	—
	rs1168074679	c.5511dup	p.Gly1838Trpfs*17	37	CUB12	0.000004	0	—	—	—	—
	rs747417629	c.5600del	p.Phe1867Serfs*16	38	CUB13	0.000004	0	—	—	—	—
	rs765301342	c.5913_5916del	p.Thr1972Leufs*10	39	CUB13	0.000049	0	—	—	—	—
		c.6926_6928delinsAA	p.Gly2309fs*6	45	CUB16	*Absent*	—	—	—	—	—
	rs757649673	c.6928_6934delGAGGTTA	p.Glu2310Cysfs*3	45	CUB16	0.000169	0	—	—	—	—
		c.7195dup	p.Asp2399Glyfs*19	46	CUB17	*Absent*	—	—	—	—	—
		c.7968_7969ins17	p.Leu2656_Pro2657delinsPheValIleProTyrIleThr	51	CUB19	*Absent*	—	—	—	—	—
	rs1315612892	c.8165del	p.Pro2722Hisfs*39	52	CUB20	0.000004	0	—	—	—	—
		c.9725_9726insACCT	p.Ala3243Profs*9	61	CUB25	*Absent*	—	—	—	—	—
		c.9956delA	p.Lys3319Argfs*2	61	CUB25	*Absent*	—	—	—	—	—
Essential		c.489+2_489+4dup	—	2_Ex6	EGF-like 2	*Absent*	—	—	—	—	—

splicing		c.5549-2A>C	—	-2_Ex38	CUB13	*Absent*	—	—	—	—	—
	rs75386064	c.6125-2A>G	—	-2_Ex41	CUB14	0.000748	0	—	—	—	—
	rs150901286	c.6821+2T>C	—	2_Ex44	CUB16	0.000198	0	—	—	—	—
		c.7706-1G>T	—	-1_Ex50	CUB19	*Absent*	—	—	—	—	—
GWAS	rs141640975	c.5069C>T	p.Ala1690Val	34	CUB11	0.00173	0	DC (1)	PD 0.994	T 0.13	(20, 22,25)
	rs144360241	c.6469A>G	p.Asn2157Asp	43	CUB15	0.00513	0.0001	DC (0.999)	PD 0.990	D 0.00	(22)
	rs45551835	c.8741C>T	p.Ala2914Val	55	CUB21	0.01216	0.0001	DC (0.993)	PD 0.984	D 0.00	(22)
	rs1801239	c.8950A>G	p.Ile2984Val	57	CUB22	0.08754	0.00464	P (0.182)	B 0.017	D 0.00	(21, 26)
NA, not available; het, heterozygous; hom, homozygous; hemi, hemizygous; PD, Probably damaging; PSD, possibly damaging; B, benign; D, deleterious; T, tolerated; DC, Disease causing; P, polymorphism; gnomAD, genome aggregation database.											

Supplemental Table 5. Description of structural models for the identified missense variants.			
Missense mutation	**Domain**	**Location**	**Predicted consequence**
Thr55Met	N-term	-	Thr55 is located in the β-helix of the Cubilin N-terminal, which interacts with AMN. Thr55 packs in the interior of the β-helix and mutation to the larger and hydrophobic methionine will most likely disturb the folding of the β-helix. The Met44Lys mutation located in the same area of cubilin has been characterized and cause ER retention of both cubilin and AMN. A similar consequence is expected of Thr55Met.
Asn1303His	CUB8	L3	Asn1303 is located directly in the CUB8-IF interface and mutation to His could affect IF-B12 binding to cubilin.
Ala1690Val	CUB11	L6	Ala1690 is positioned in interior part of loop 6. Mutation to the larger valine could alter the conformation of the loop. If CUB11 loop 6 is involved in domain interfaces the mutation could affect the general structure of cubilin.
Asn1854Arg	CUB13	β2	Asn1854 is located at the N-terminal of CUB13 immediately after the linker between CUB12 and CUB13. Mutation to arginine could disturb the domain interface between CUB12 And CUB13 and prevent correct folding of the receptor.
Gly1928Val	CUB13	L7	Gly1928 is positioned at the tip of the hairpin in loop 7 in CUB11. Mutation to valine could alter the main chain conformation and prevent formation of the hairpin. This could in turn destabilize the entire domain.
Ser1947Tyr	CUB13	L9	Ser1947 is positioned in the Ca^{2+}-binding loop 9 of CUB13. Mutation to the larger aromatic tyrosine residue will most likely affect the conformation of loop 9 and prevent Ca^{2+}-binding. If the Ca^{2+}-site of CUB13 mediates ligand binding the mutation will likely cause a decrease in ligand affinity.

Asn2157Asp	CUB15	L6	Asn2157 is located in the loop 6 of CUB15. Mutation to aspartate places an unnatural negative charge in the area, which could potentially affect cubilin domain interfaces.
Leu2261Arg	CUB16	β5	Leu2261 is positioned in the hydrophobic core of CUB16 and mutation to arginine will most likely prevent correct folding of the entire domain.
Ser2599Trp	CUB19	β10	Ser2599 is located in β10 of CUB19, where it packs in the interior of the domain. Mutation to the much larger tryptophan most likely prevents correct folding of the entire domain.
Pro2822Leu	CUB21	L3	Pro2822 is surface exposed in L3 at the tip of CUB21. Mutation to Leucine is not expected to affect the folding of CUB21, but could possibly disturb inter-domain interactions or ligand binding.
Ala2914Val	CUB21	β10	Ala2914 is positioned in the hydrophobic core of CUB21. Mutation to the slightly larger hydrophobic residue valine might introduce steric clashes, which could affect folding of CUB21.
Ile2984Val	CUB22	β6	Ile2984 is positioned in the hydrophobic core of CUB22. Mutation to the slightly smaller could introduce minor changes in the hydrophobic core of CUB22, which could affect folding of the domain.
Tyr3018Ser	CUB22	β9	Tyr3018 is positioned in β9. The residue is solvent exposed and could potentially be directly involved in ligand binding. Mutation to serine could disrupt ligand interaction.
Gly3027Arg	CUB22	L9	Gly3027 is positioned in L9 of CUB22., where it is in close contact with residues from the adjacent L5. Mutation to the much larger arginine will most likely affect the conformation of both L5 and L9, which could be directly involved in ligand binding.
Trp3308Arg	CUB25	β4	Trp3308 is located in β4 of CUB25, where it packs in the interior of the domain. Mutation to longer and hydrophilic arginine most likely prevents correct folding of the entire domain.

Ser3366Cys	CUB25	L8	Ser3366 is positioned solvent exposed in loop 8. Introduction of a cysteine residue in this position could cause unspecific disulfides preventing correct folding of the receptor.
Asp3492Tyr	CUB26	L9	Asp3492 is part of the putative Ca^{2+}-binding site of CUB26 and potentially directly coordinating Ca^{2+}. Mutation to tyrosine will most likely prevent Ca^{2+}-binding. If the Ca^{2+}-site of CUB26 mediates ligand binding the mutation will likely cause a decrease in ligand affinity.
Gly3520Asp	CUB27	β3	Gly3520 is positioned in β3 and buried in the interior of CUB27. Mutation to larger and charged aspartate could introduce steric clashed that in turn prevent proper folding of CUB27.
Asp3609His	CUB27	L9	Asp3609 of CUB27 corresponds to Asp3492 in CUB26. While the triad of acidic residues (Glu-Asp-Asp) constituting Ca2+-binding sites in CUB-domains is not strictly conserved in CUB27, the triad of CUB27 formed by the residues Asp, Asn and Asp could nevertheless be involved in Ca2+ binding site. In such a case, mutation to histidine would disturb the interaction to Ca2+.
β, β-sheets of CUB domains, L, loops of CUB domains.			

Supplemental Table 6. Patients with identified hemizygous mutations in Dent's disease genes *CLCN5* and *OCRL1*.

Patient No.	Gene	Exon (zygosity)	Mutation	Mutation Taster	PolyPhen 2	SIFT	MAF (gnomAD)	References (HGMD)
N3806	*CLCN5*	10 (hemi)	c.941C>T – p.Ser314Leu	DC (1)	PD 0.986	D 0.01	0	yes
N4111	*OCRL1*	11 (hemi)	c.952C>T – p.Arg318Cys	DC (1)	PD 1.000	D 0.00	0	yes

hemi, hemizygous; DC, disease causing; PD, probably damaging; D, deleterious.

Supplemental Table 7. Individual with one filtered variant and one GWAS variant in trans.																	
Patient No.	**Gene**	**Exon (zygo-sity)**	**Segregation**		**Mutation**	**Poly Phen 2**	**SIFT**	**Muta-tion Taster**	**gnomAD frequen-cy (het)**	**Consan-guinity**	**Age at first manifes-tation**	**UPCR (g/ mmol)**	**UACR (g/ mmol)**	**UAPR (%)**	**eGFR (ml/ min/m^2)**	**Biopsy**	**Glyco-suria**
			Father	**Mother**													
N4142	*CUBN*	39 (het)	het	WT	c.5840C>A – p.Ser1947 Tyr	PD 0.995	D 0.00	P (0.98)	0.00027	yes	3.0	0.14	0.098	70.0	147	Minimal glomerular lesions	Neg-ative
		43 (het)	WT	het	c.6469A>G – p.Asn2157 Asp	PD 0.990	D 0.00	DC (0.999)	0.0001								
	het, heterozygous; WT, wild type; PD, Probably damaging; D, deleterious; P, polymorphism; DC, Disease causing; UPCR, Urinary protein to creatinine ratio; UACR, Urinary albumin to creatinine ratio; UAPR, Urinary albumin to protein ratio.																

Supplemental Table 8. Association of *CUBN* SNPs with albuminuria and kidney function (eGFR) in Europeans (with and without diabetes).												
Gene	**rsID**	**Protein**	**EA/OA**	**EAF**	***Albuminuria***				***eGFR***			
					Effect on log(UACR)	**SE**	**P**	**N**	**Effect on log (eGFR)**	**SE**	**P**	**N**
CUBN	rs45551835	A2914V	A/G	0.019	0.14	0.037	1.9×10^{-4}	13226	0.018	0.008	0.03	13550
	rs141640975	A1690V	A/G	0.0085	0.26	0.061	1.2×10^{-5}	13124	0.026	0.013	0.04	13383
	rs1801239	I2984V	G/A	0.015	0.055	0.017	0.0017	13225	0.008	0.003	0.02	13382
Albuminuria (UAER/UACR) and eGFR have been log transformed in each linear regression analysis adjusting for sex, age, first 10 principal components, followed by an inverse variance fixed effects meta analysis of 5 cohorts. EA/OA: Effect allele, Other allele; EAF: Effect allele frequency, *f* is the frequency, Effect on log(UACR) and on log(eGFR) is the effect size or "beta" (positive value means positive directionality and therefore an increased value of albuminuria and eGFR); SE is the standard error; P is the P-value; and N is the sample size. The unit for UACR is mg/g and for eGFR is ml min^{-1} per 1.73 m^2.												

Supplemental Table 9. Association of *CUBN* SNPs with albuminuria and kidney function (eGFR) in Europeans stratified for diabetes status.										
			Albuminuria							
			DIABETES				**NON DIABETES**			
Gene	**rsID**	**Protein**	**Effect on log(UACR)**	**SE**	**P**	**N**	**Effect on log(UACR)**	**SE**	**P**	**N**
CUBN	rs141640975	A1690V	0.69	0.16	2.2×10^{-5}	3873	0.19	0.065	0.002	9251
	rs45551835	A2914V	0.17	0.11	0.11	3896	0.13	0.04	5.8×10^{-4}	9330
	rs1801239	I2984V	0.093	0.04	0.05	3896	0.04	0.018	0.011	9329
			eGFR							
			DIABETES				**NON DIABETES**			
Gene			**Effect on log (eGFR)**	**SE**	**P**	**N**	**Effect on log (eGFR)**	**SE**	**P**	**N**
CUBN	rs141640975	A1690V	0.017	0.013	0.04	3967	0.037	0.015	0.01	9407
	rs45551835	A2914V	0.045	0.008	0.03	3881	0.009	0.009	0.34	9499
	rs1801239	I2984V	0.012	0.003	0.02	3990	0.006	0.004	0.16	9392
	eGFR: Estimated glomerular filtration rate; albuminuria (UAER/UACR) and eGFR have been log transformed in each linear regression analysis adjusting for sex, age, first 10 principal components. Position in base pairs corresponding to GRCh37 assembly; EA/OA: Effect allele, Other allele; EAF: Effect allele frequency, *f* is the frequency, Effect on log(UACR) and on log(eGFR) is the effect size or "beta" (positive value means positive directionality and therefore an increased value of albuminuria and eGFR); SE is the standard error; P is the P-value; and N is the sample size. The unit for UACR is mg/g and for eGFR is ml min^{-1} per 1.73 m^2.									

Supplemental Table 10. Additive genetic model of *CUBN* SNPs for albuminuria.					
SNP	EA/OA	EAF	**Effect on log(UACR)**	SE	P_association
rs1801239 (I2984V)	G/A	0.105	0.069	0.019	5.2×10^{-4}
rs45551835 (A2914V)	A/G	0.019	0.16	0.043	1.9×10^{-4}
rs141640975 (A1690V)	A/G	0.009	0.33	0.066	5.2×10^{-7}

All SNP trait associations are based on linear regression additive model on pooled genotype data from the Danish cohorts (n= upto 13,295) and correcting for sex, age, population substructure (PCs1-4), and study cohorts.
EAF: Effect allele frequency

Supplemental Table 11. Recessive genetic model of *CUBN* SNPs for albuminuria.					
SNP	EA/OA	EAF	**Effect on log(UACR)**	SE	P_association
rs1801239 (I2984V)	G/A	0.105	0.28	0.084	6.4×10^{-4}
rs45551835 (A2914V)	A/G	0.019	-	-	-
rs141640975 (A1690V)	A/G	0.009	-	-	-

All main SNP trait associations (P_association) are based on linear regression recessive genetic models on pooled genotype data from the Danish cohorts (n= upto 13,295) and correcting for sex, age, population substructure (PCs1-4), and study cohorts. EAF: Effect allele frequency

Supplemental Table 12. Genotypic distribution of the *CUBN* SNPs and power estimate for albuminuria.					
SNP	**Genotype distribution**			**Power* (%)**	**Power** (%)**
rs1801293[a]	AA	AG	GG	100	~40
	10639	2515	141		
rs45551835[b]	GG	GA	AA	89	1
	12782	506	8		
rs141640975[c]	GG	GA	AA	94	1
	13076	218	2		

***Statistical power based on Additive model. **Statistical power based on recessive model**
[a]Power estimate based on n ~13,500 individuals, relative risk (or estimate) = 1.2, trait prevalence = 0.15, effect allele frequency = 0.105, P_significance =0.015;
[b]Power estimate based on n ~13,500 individuals, relative risk (or estimate) = 1.3, trait prevalence = 0.15, effect allele frequency = 0.019, P_significance =0.015;
[c]Power estimate based on n ~13,500 individuals, relative risk (or estimate) = 1.5, trait prevalence = 0.15, effect allele frequency = 0.009, P_significance =0.015

Statement about reporting guidelines:

Our retrospective study summarizes genetic data obtained by sequencing of patients in the clinical diagnostic setting based on informed patient consent. As our study is not a clinical trial, it requires neither a trial registration number nor standardized trial flow diagrams and checklists.