Original research

A founder *UMOD* variant is a common cause of hereditary nephropathy in the British population

Manoj K Valluru,[1] Noelle KX Chung,[1] Mark Gilchrist [ID],[2] Laura Butland,[3] Jackie Cook,[4] Anna Takou,[5] Abhijit Dixit,[3] Michael N Weedon,[6] Albert C M Ong [ID],[1,7] The Genomics England Research Consortium

ABSTRACT

Background Monogenic disorders are estimated to account for 10%–12% of patients with kidney failure. We report the unexpected finding of an unusual uromodulin *(UMOD)* variant in multiple pedigrees within the British population and demonstrate a shared haplotype indicative of an ancestral variant.

Methods Probands from 12 apparently unrelated pedigrees with a family history of kidney failure within a geographically contiguous UK region were shown to be heterozygous for a pathogenic variant of *UMOD* c.278_289delTCTGCCCCGAAG insCCGCCTCCT.

Results A total of 88 clinically affected individuals were identified, all born in the UK and of white British ethnicity. 20 other individuals with the variant were identified in the UK 100,000 Genomes (100K) Project and 9 from UK Biobank (UKBB). A common extended haplotype was present in 5 of the UKBB individuals who underwent genome sequencing which was only present in <1 in 5000 of UKBB controls. Significantly, rare variants (<1 in 250 general population) identified within 1 Mb of the *UMOD* variant by genome sequencing were detected in all of the 100K individuals, indicative of an extended shared haplotype.

Conclusion Our data confirm a likely founder *UMOD* variant with a wide geographical distribution within the UK. It should be suspected in cases of unexplained familial nephropathy presenting in patients of white British ancestry.

INTRODUCTION

Chronic kidney disease (CKD) is estimated to affect up to 10% of the world's population, making it a leading cause of morbidity and mortality.[1] Although CKD and kidney failure usually results from the interaction of lifestyle or environmental factors with complex polygenic traits, monogenic disorders are estimated to account for 10%–12% of kidney failure.[2] The vast majority of these cases can be attributed to autosomal dominant polycystic kidney disease but pathogenic gene variants especially associated with autosomal dominant tubulointerstitial kidney disease (ADTKD) and Alport syndrome are being increasingly recognised.[2]

ADTKD is a rare genetic cause of progressive CKD and kidney failure. It is genetically heterogenous with pathogenic variants in *MUC1*,[3] *REN*,[4] *HNF1B*[5] and uromodulin *(UMOD)*[6] identified, with *UMOD* comprising the highest proportion of ADTKD cases

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ A rare uromodulin *(UMOD)* variant was first reported in a UK pedigree with familial nephropathy in 2001, followed by isolated reports.

⇒ A previous study concluded that the variant was a recurrent change rather than a founder effect.

WHAT THIS STUDY ADDS

⇒ A further 41 pedigrees with the rare *UMOD* variant identified through regional centres and large national databases.

⇒ Compelling evidence for a pathogenic ancestral variant highly prevalent in the white British population, responsible for kidney failure and chronic kidney disease.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ A high index of clinical suspicion in undiagnosed cases of familial nephropathy of white British ancestry.

⇒ Definition of a large affected group of common ancestry sharing the same variant and thus opportunities to study non-allelic contributions to phenotypic variability.

of up to 50%, with a disease prevalence of 9 per million.[7] ADTKD-*UMOD* is estimated to account for 2% of patients with kidney failure,[8] with susceptibility *UMOD* variants also conferring around 20% increased risk for CKD and 15% for hypertension.[9] Classically, it is characterised by early onset gout, hyperuricaemia, the absence of haematuria or proteinuria and kidney failure usually occurring between 30 and 60 years.

UMOD encodes uromodulin, the most abundant mammalian urinary protein, which is predominantly produced by the thick ascending limb of the loop of Henle in the renal tubule.[10] Uromodulin contains an N-terminal signal peptide, three epidermal growth factor-like (EGF-like) domains, an eight-cysteine domain (D8C), a zona pellucida (ZP) domain and a glycosylphosphatidylinositol anchor segment. Although the physiological roles of uromodulin have not been fully established, studies on *UMOD* knockout mice have suggested protective roles against urinary tract infections (UTIs),[11–13] calcium oxalate formation[14] and in facilitating electrolyte reabsorption in the renal tubules.[15 16] Pathogenic *UMOD* variants lead to

protein misfolding resulting in intracellular accumulation within the endoplasmic reticulum (ER)[17] and a subsequent reduction in urinary secretion.[18] Over 100 distinct *UMOD* variants have been reported to contribute to ADTKD-*UMOD*, with 95% localising in exon 3 and exon 4.[19] The clinical presentation across ADTKD-*UMOD* genotypes appears to be variable with no clear genotype-phenotype correlation.[20]

In this study, we report a rare *UMOD* variant, c.278_289delTCTGCCCCGAAG insCCGCCTCCT detected in a cluster of unrelated families within a contiguous geographical area with unexplained CKD leading to kidney failure. In national UK cohorts (100,000 Genomes (100K) Project and UK Biobank (UKBB)), we detected other cases with the same variant from other geographical regions but who share a common extended haplotype indicative of an ancestral mutation originating in the UK population.

MATERIALS AND METHODS

Study population
All probands were referred to the renal genetic clinics at Sheffield and Nottingham due to a family history of kidney failure of unknown aetiology. Index cases with heterozygous *UMOD* variant c.278_289delTCTGCCCCGAAG insCCGCCTCCT had their family history assessed. Affected relative with CKD and at-risk, apparently healthy family members were subsequently invited to a renal genetics consultation where they were offered genetic testing and had their clinical features, kidney function and kidney morphology (ultrasound) assessed. The age of onset of kidney failure was recorded and patients without kidney failure had their latest estimated glomerular filtration rate (eGFR) estimated using the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) 2009 equation. The definition and staging of CKD followed the Kidney Disease Outcomes Quality Initiative and Kidney Disease Improving Global Outcomes (KDIGO) guidelines. CKD was defined as the presence of an eGFR $<60\,mL/min/1.73\,m^2$ or the presence of markers indicating kidney damage, such as albuminuria for >3 months. Staging of CKD was based on eGFR and severity of albuminuria, and their CKD class specified according to KDIGO guidelines.[21 22] A retrospective review of all patients screened for the *UMOD* variant at both renal genetic clinics was conducted.

Mutational analysis
The recurrent pathogenic *UMOD* variant c.278_289delTCTG-CCCCGAAG insCCGCCTCCT (exon 3) was identified using targeted analysis or panel-based sequencing in familial cases of kidney failure. Five of the index cases were identified on a six gene ADTKD next-generation sequencing gene panel offered through the Sheffield Diagnostic Genetics Service comprising the following genes: *REN, UMOD, HNF1B, SEC61A1, TSC1, TSC2* and another five through a two gene panel (*UMOD, REN*) offered through Oxford University Hospitals. The remaining two cases were diagnosed on direct *UMOD* testing or on a 15 gene tubulointerstitial kidney disease panel (R202, Panel App V.1.3) which includes *UMOD* (Bristol Genetics Laboratory).

Genomics England 100,000 Genomes Project
Inclusion and genotyping of participants in the 100K was managed by Genomics England Limited (GEL). All participants in 100K provided written consent to access their anonymised clinical and genomic data for research purposes (https://re.extge.co.uk/ovd/). Whole-genome sequencing (WGS) was performed on all participants and processed using the GEL rare disease

analysis pipeline as previously described.[23 24] Phenotypes of identified carriers were manually reviewed in the Genomics England Participant Explorer. Initially, affected candidates were filtered based on c.278_289delTCTGCCCCGAAG insCCGCCTCCT *UMOD* variant (rs878855325) in the interactive variant analysis browser (IVA V.2.0, CG38 and RD38). Further detailed analysis of rs878855325 for the selected cases was extracted from the IVA browser and phenotype information from the KIBANA Data discovery (V.3.2) browser. KIBANA and IVA are part of the secured access environment. Cohort statistics were expressed as very rare (frequency <0.1%), rare (frequency <0.5%), average (frequency <5%), common (frequency >5%) or not observed.[25] Conservation was assessed as follows: PhyloP score (positive scores—measure conservation, which is slower evolution than expected, at sites that are predicted to be conserved), PhastCons (scores represent probabilities of negative selection and range between 0 and 1) and Genomic Evolutionary Rate Profiling (score ranges from −12.3 to 6.17, with 6.17 being the most conserved). Selected cases and phenotype history were recorded in an excel file using a pseudo-case ID. Finally, recorded data were subjected to downstream bioinformatics analysis.

UK Biobank
The UKBB comprises approximately 500 000 participants with extensive phenotyping and genetic data linked to clinical care records.[26] The whole-exome sequencing pipeline and quality control has been recently described.[27] We examined the available UKBB cohort exome data for the presence of the c.278_289delTCTGCCCCGAAG insCCGCCTCCT variant in the *UMOD* gene using the NM_003361.3 transcript and obtained demographic data from baseline assessment including age and sex. Clinical phenotype data relevant to ADTKD-*UMOD* including CKD-EPI eGFR, systolic and diastolic blood pressure, albumin:creatinine ratio and serum urate concentration from enrolment in UKBB were obtained along with Hospital Episode Statistics (HES) data for CKD and gout. Ancestry was determined by self-reporting at the assessment centre. Statistical differences in clinical data were determined by independent t-test for continuous data and a Fisher's exact test for categorical data.

Haplotype analysis
We used the directly genotyped SNP chip data from UKBB to phase haplotypes on chromosome 16. The genotyping and quality control of UKBB has been described previously.[26] We only used SNPs with a minor allele frequency >5% with a missingness rate <1%. We used SHAPEIT2[28] for haplotype phasing.

To support the common variant haplotype analyses, we used WGS data available in 150 000 UKBB individuals. Five of the nine individuals with the c.278_289delTCTGCCCCGAAG insCCGCCTCCT variant had WGS data available. The genome sequencing and quality control have been described in detail elsewhere (doi: https://doi.org/10.1101/2021.11.16.468246).

We identified four rare variants (<0.5% frequency in the UKBB) in the 1 Mb window around c.278_289delTCTGC-CCCGAAG insCCGCCTCCT and tested whether they were associated with the presence of the *UMOD* variant using Fisher's exact test. We determined ancestry using principal component analysis as previously described.[29]

We then assessed the frequency of four rare variants associated with c.278_289delTCTGCCCCGAAG insCCGCCTCCT in the UKBB in the 100K participants. The genome sequencing and quality control have been described previously.[24] We used
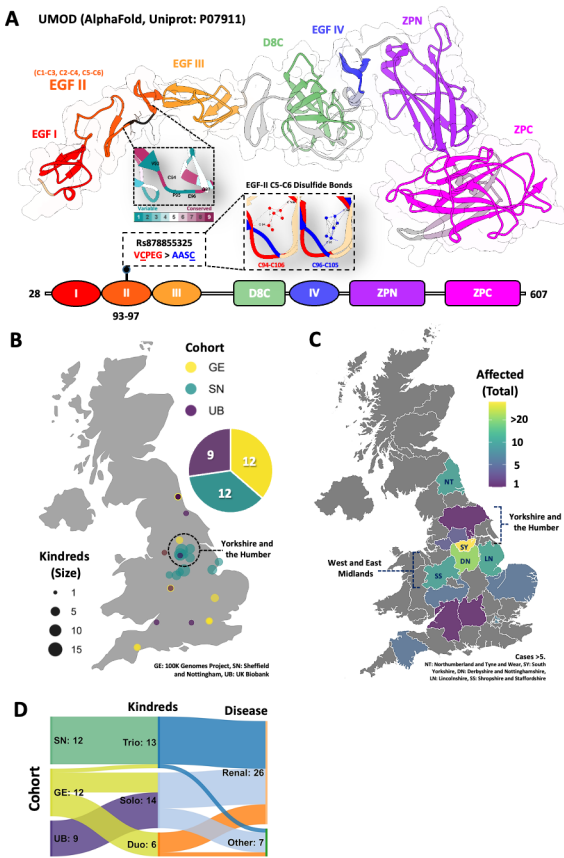
Table 1 Clinical and laboratory features of individuals with the *UMOD* indel variant compared with others in the UKBB cohort

	UMOD indel (n=9)	UKBB (n=450 993)	P value
Age (years)	58.5±8.7	57.3±8.0	0.69
Sex (M/F %)	77.8/22.2	54.3/45.7	0.194
eGFR CKD-EPI (mL/min/1.73 m^2)	69.0±20.3	90.5±13.0	**0.016**
ACR (mg/mmol)	3.1±5.0	1.6±2.5	0.42
SBP (mm Hg)	163.6±20.8	144.2±24.1	**0.023**
DBP (mm Hg)	99.3±13.0	86.4±13.5	**0.017**
Uric acid (µmol/L)	309.7±55.1	309.1±80.4	0.98
Gout	0	9338	1.0

Bold values are statistically significant ie p < 0.05.
ACR, albumin:creatinine ratio; CKD-EPI, Chronic Kidney Disease Epidemiology Collaboration; DBP, diastolic blood pressure; eGFR, estimated glomerular filtration rate; F, female; M, male; SBP, systolic blood pressure; UKBB, UK Biobank; UMOD, uromodulin.

Figure 1 Structural model of wild-type and mutant uromodulin (UMOD) and demographics of the c.278_289delTCTGCCCCGAAG insCCGCCTCCT *UMOD* variant (Indel rs878855325). (A) AlphaFold structure and schematic diagram of uromodulin contains four EGF-like (I–IV), D8C, ZPN and ZPC domains. Indel rs878855325 (VCPEG>AASC) in the EGF II domain (highly conserved C94 and G97 residues) can affect structural integrity by disrupting disulfide bonds in C5-C6 (red: wild type and blue: mutant). (B) UK bubble map showing cohorts of unique pedigrees with Indel rs878855325. (C) UK choropleth map showing total identified cases across different counties in England (cases >5; NT, Northumberland and Tyne and Wear; SY, South Yorkshire; DN, Derbyshire and Nottinghamshire; LN, Lincolnshire; SS, Shropshire and Staffordshire). (D) Sankey diagram of Indel rs878855325 cohorts with paired cohort group and kindreds (trio, duo and solo or singleton). Disease status is indicated (renal and other). GE, 100K Genomes Project; SN, Sheffield and Nottingham; UB, UK Biobank.

Fisher's exact test to test for an increased frequency of these variants in the c.278_289delTCTGCCCCGAAG insCCGCCTCCT heterozygotes compared with the 100K background population.

In silico studies

The three-dimensional (3D) structure of uromodulin (UniProt: P07911) was downloaded from AlphaFold DB (https://alphafold.ebi.ac.uk/). An experimentally validated indel mutant structure is not available, and therefore we generated a mutant structure by introducing the indel mutation in silico, computationally modelled by AlphaFold Colab (https://github.com/deepmind/alphafold).[30] All types of direct interactions: polar and non-polar, favourable and unfavourable, including clashes, were analysed using contacts command in UCSF Chimera V.1.14.[31] In the output, the atom-atom contacts are listed in order of decreasing van der Waals (VDW) overlap: positive where the atomic VDW spheres are intersecting, zero if just touching and negative if separated by space. The superimposed 3D

structural model of the EGF II domain was obtained by superimposing mutated EGF II into the EGF II of the UMOD model, using the MatchMaker tool on UCSF Chimera V.1.14.[31] The evolutionary conservation score of each amino acid of UMOD (EGF II domain) was determined using the ConSurf algorithm, based on the phylogenetic relationships between sequence homologues.[32] The predicted impact of the mutation on protein function was analysed using DeepFRI (Graph Convolutional Network for predicting protein functions).[33]

Maps and plots

The bubble and choropleth map plots were created in R studio (R V.4.0.5). Shapefile: NUTS Level 2 (Counties) Boundaries file was downloaded from geoportal.statistics.gov.uk (valid as of January 2018). A Sankey diagram was created using packages Canvg and d3.js (https://github.com/nowthis/sankeymatic). Pedigree structures were plotted manually in Microsoft PowerPoint.

Statistical analysis

Descriptive measures were presented as mean±SD or median (IQR or range as specified). Renal survival was defined as time to the start of renal replacement therapy and was displayed as Kaplan-Meier survival plots. Patients were censored if they did not receive renal replacement therapy in the study duration and a log-rank test was used for comparison between sex. A χ^2 test was used to determine association between hypertensive status and the development of kidney failure. A p value of <0.05 was deemed as significant. All statistical analysis was done with SPSS V.26.

RESULTS
Clinical findings at presentation
Genetic testing

A total of 91 patients across 12 families with available clinical data were reviewed and their diagnostic outcomes were summarised (online supplemental figure 1). Altogether, there were 88 patients with definite or likely ADTKD-*UMOD*, 38 with a genetic diagnosis and 50 with a presumptive diagnosis (not genotyped) based on their clinical characteristics and family history according to diagnostic criteria laid out by the KDIGO Consensus Report on ADTKD.[34] Of the latter, 94% (47/50) had a recorded diagnosis of kidney failure and 66% (33/50) were deceased.
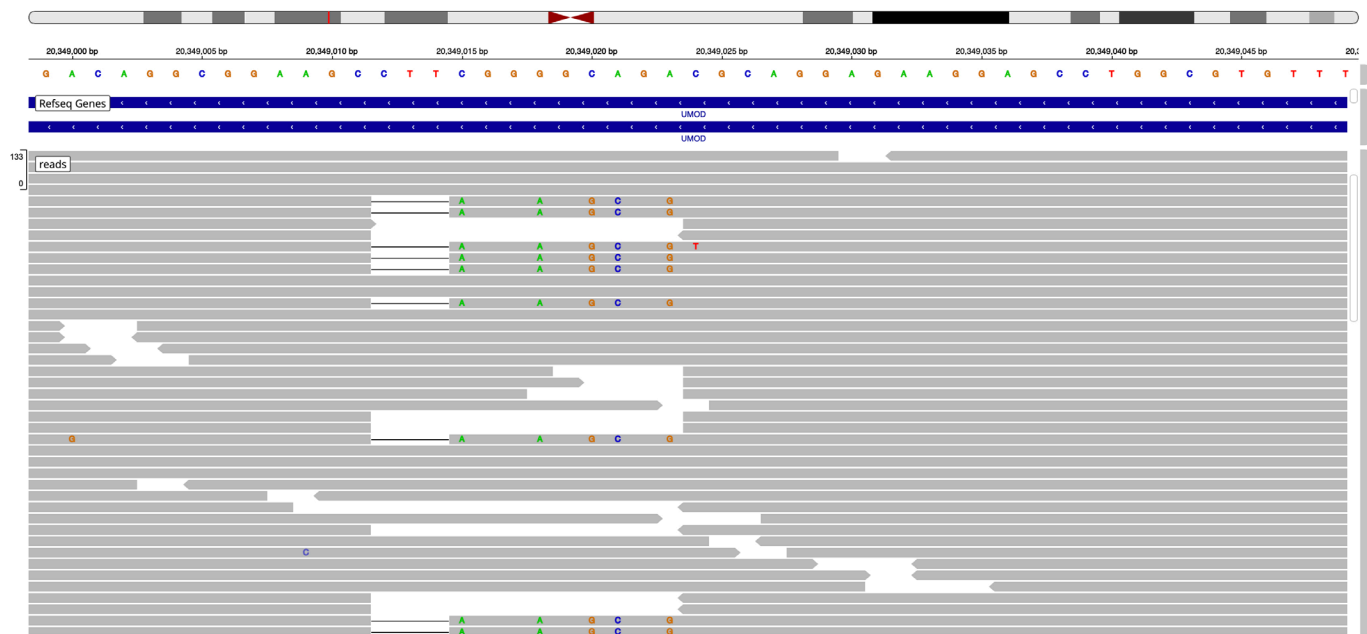
Figure 2 IGV plot showing the c.278_289delTCTGCCCCGAAG insCCGCCTCCT uromodulin *(UMOD)* variant. It has been miscalled as six separate variants.

Kidney function

Online supplemental table 1 shows the breakdown of clinical features in the patients with ADTKD-*UMOD*. Eighty-two per cent (71/88) had a diagnosis of CKD with 69% (61/88) having kidney failure. The median age of kidney failure was 52 years

(range 32–76). The remaining 31% (27/88) who did not have CKD on presentation were younger and had a median age of 32.5 years (range 17–73). The median age of kidney survival was 55 years with no significant sex difference (p=0.31 by log-rank test) based on Kaplan-Meier survival analysis (online
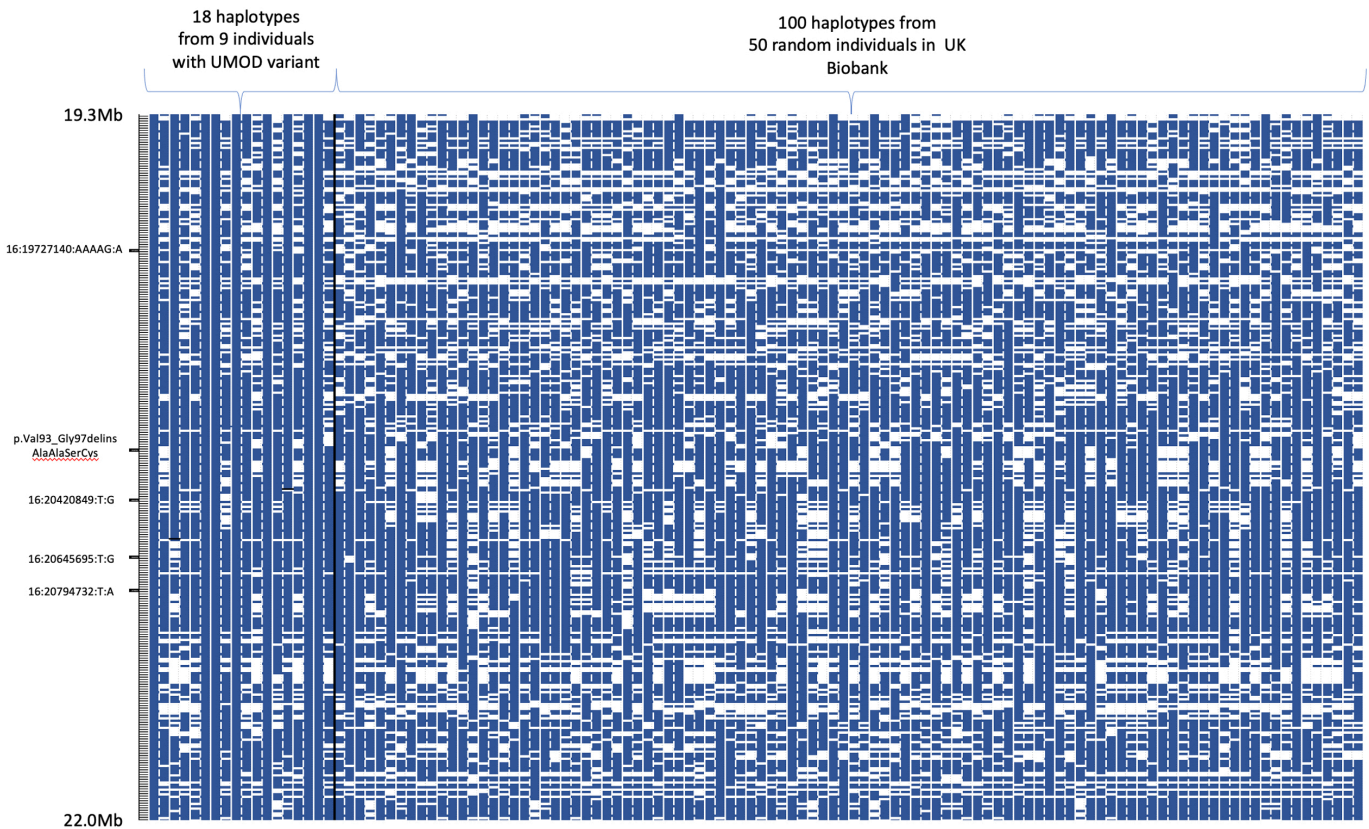
Figure 3 Haplotype analysis of phased common variants demonstrates all nine individuals in the UK Biobank with c.278_289delTCTGCCCCGAAG insCCGCCTCCT carry a shared extended haplotype. Blue indicates where an allele matches the putative shared haplotype for that variant and white where it does not match. The four rare variants identified from whole-genome sequencing are indicated on the left-hand side of the figure. UMOD, uromodulin.

Table 2 Rare variant analysis in UKBB and 100K (GEL) whole-genome sequencing datasets provides strong evidence for shared haplotype in all individuals with the c.278_289delTCTGCCCCGAAG insCCGCCTCCT variant

	UKBB background frequency (n with variant/n without)	UKBB frequency in individuals with UMOD c.278_289delTCTGCCCCGAAG insCCGCCTCCT variant (n with/n total)	P value	GEL background frequency (n with variant/n without)	GEL frequency in individuals with UMOD c.278_289delTCTGCCCCGAAG insCCGCCTCCT variant (n with/n total)	P value
16:19727140:AAAAG:A	329/115140	5/5	5×10^{-12}	410/75211	16/20	$<1\times10^{-16}$
16:20420849:T:G	424/114716	4/5	1×10^{-10}	228/75073	18/20	$<1\times10^{-16}$
16:20645695:T:G	418/114722	3/5	5×10^{-7}	218/75083	18/20	$<1\times10^{-16}$
16:20794732:T:A	109/115031	4/5	4×10^{-12}	76/75225	18/20	$<1\times10^{-16}$

Five of the nine individuals with c.278_289delTCTGCCCCGAAG insCCGCCTCCT variant in the UKBB had whole-genome sequencing data. Specific rare variants (<0.5% UKBB) in a 1 Mb window around the *UMOD* indel variant (column 1) were significantly over-represented in carriers of the *UMOD* variant versus the overall UKBB and 100K cohorts. The location of the variants are shown in figure 3. P values are based on Fisher's exact test and only one individual from each family was included in the analysis.
GEL, Genomics England Limited; 100K, 100,000 Genomes Project; UKBB, UK Biobank; UMOD, uromodulin.

supplemental figure 2). Within each family however, the age of kidney failure varied greatly between individuals with a median age of 52 (range 32–76) years; pedigree 12 had no information on individuals reaching kidney failure (online supplemental figure 3).

Clinical features

Hypertension was present in 65% (30/46) of our cohort with available blood pressure readings. Hypertension showed a non-significant association with the development of kidney failure (p=0.07). Proteinuria was rarely seen and only three patients had proteinuria ≥30 mg/g. Microscopic haematuria (trace or 1+) was detected in 18% (7/40); no case of macroscopic haematuria was reported. Where available, hyperuricaemia was noted in 32% (9/28) at presentation and 61% (17/28) had transient hyperuricaemia recorded on at least one occasion. Nonetheless, gout was not a prominent feature, with only 6.9% (6/87) reporting incident gout at a median age onset of 42.5 (range 29–61) years. Only two patients had renal stones, one was asymptomatic and diagnosed incidentally. Of patients with available information, 15% (8/55) reported an episode of UTI during their lifetime; recurrent UTIs were present in 4% (2/55).

Renal ultrasound

Among patients with available renal ultrasound results (n=28), 61% (17/28) had normal-sized kidneys and the remaining 39% (11/28) had small kidneys. Renal cysts were uncommon: 18% (5/28) had either one or two cysts in each kidney and only one patient had >10 cysts in each kidney.

Table 3 Demographics of reported individuals with the c.278_289delTCTGCCCCGAAG insCCGCCTCCT *UMOD* variant in various cohorts

Study	Pedigrees (n)	Individuals (n)	Reference
Sheffield and Nottingham	12	88	This study
100K (GEL)	12	20	This study
UKBB	9	9	This study
Cambridge	5	35	37
Southampton	4	4	8
Cardiff (Cologne)	2	7	35 36
Wake Forest Registry, USA	11	74	7 19

GEL, Genomics England Limited; 100K, 100,000 Genomes Project; UKBB, UK Biobank; UMOD, uromodulin.

Biopsy findings

Eight patients with the variant had renal biopsies and only one patient had a normal biopsy report (online supplemental table 1). Among biopsies with abnormal findings, interstitial fibrosis and tubular atrophy were the most common findings seen: examples from two patients are shown in online supplemental figure 4. Arterial profiles were mostly thick walled, hyalinosis was rare and seen in only one patient. There was an absence of immune reactants on immunofluorescence. All glomerular basement membranes visualised on electron microscopy were reported to be normal. In one patient (3.5), an enlargement in lysosomes was seen in podocytes.

Overview of the clinical pedigrees presenting to renal genetic clinics

A total of 53 individuals from Sheffield (6 pedigrees) and 35 individuals from Nottingham (6 pedigrees) were identified. The index cases from each of the 12 pedigrees presented at a median age of 50 (range 26–60 years), following a referral from either the renal team or general practice. On initial presentation, all had an eGFR <60 mL/min/1.73 m^2, with the majority having a known diagnosis of hypertension (n=9). Across the Sheffield and Nottingham cohorts, a family history of CKD or kidney failure was present in up to three generations in one pedigree, four generations in nine pedigrees and five generations in two pedigrees. Details of all 12 pedigrees (SN1–12) are summarised in online supplemental figures 5 and 6.

The c.278_289delTCTGCCCCGAAG insCCGCCTCCT variant in *UMOD* leads to a predicted change in EGF II domain structure and altered Ca^{2+} binding

The mature UMOD protein is 616 amino acids in length. The predicted structure of UMOD contains four EGF-like domains (EGF I–IV), a cysteine-rich D8C domain and a bipartite C-terminal ZP domain (ZPN and ZPC) (figure 1A). Domains EGF II and EGF III are predicted to bind Ca^{2+}.

The specific indel variant results in the replacement of five amino acids 'VCPEG', by four novel residues 'AASC'. DeepFRI results were: EGF II wild type—GO:0043169 score=0.84 and GO:0005509 score=0.54 and indel—GO:0043169 score=0.60. Based on in silico predictions using AlphaFold and DeepFRI, this change is predicted to alter protein folding and structural integrity of the EGF II domain and could alter the dynamics of Ca^{2+} binding and protein polymerisation.

The c.278_289delTCTGCCCCGAAG insCCGCCTCCT variant is present in 20 individuals with and without a renal phenotype in the 100,000 Genome Project (GEL)

To identify potential cases from a wider geographical area in England, we accessed the 100K database and identified the c.278_289delTCTGCCCCGAAG insCCGCCTCCT variant (Indel rs878855325, ClinVar 242346) in 20 individuals (online supplemental table 3). Within the RD38 (rare disease) cohort, 17 individuals were identified across 9 families (GE1–9), with 6 families recruited for kidney phenotypes and 3 for non-kidney phenotypes (online supplemental figure 7). A further three individuals were identified within the CG38 (cancer germline) cohort but without additional phenotypic information. The calculated allele frequency was 0.0001582 (0.01582%) for the RD38 population and 0.00009841 (0.009841%) for the CG38 population. We did not identify any positive cases in The Cancer Genome Atlas (PanCancer Atlas Studies, cBioPortal V.4.1.9) and The Genome Aggregation Database (GnomAD V.3.1.2).

Nine individuals are heterozygous for the c.278_289delTCTGCCCCGAAG insCCGCCTCCT variant in UK Biobank

Out of 450993 individuals in the UKBB with exome sequencing data, 9 individuals, 7 males, 2 females were found to possess a single copy of the c.278_289delTCTGCCCCGAAG insCCGGCTCCT variant (table 1). This variant had been miscalled as six separate variants, but visualisation of Integrative Genomics Viewer (IGV) plots of individuals with these variants showed that individuals with these six variants all represent the single p.Val93_Gly97delinsAlaAlaSerCys variant (figure 2). Although we were unable to confirm these cases by Sanger sequencing, we noted that the same pattern was observed in the 100K individuals with this variant, and all have been confirmed by Sanger sequencing. No other *UMOD* variants were identified in these individuals.

All nine individuals were of European descent. Seven were unrelated and two were known to be related to each other. Kidney function was lower in those with the variant: CKD-EPI eGFR 69.0 ± 20.3 vs $90.5\pm13.0\,\text{mL/min/1.73}\,\text{m}^2$ (p=0.016). Systolic blood pressure was higher: 163.6 ± 20.8 vs $144.2\pm24.1\,\text{mm}$ Hg (p=0.023), as was diastolic blood pressure: 99.3 ± 13.0 vs $86.4\pm13.5\,\text{mm}$ Hg (p=0.017). In keeping with previous reports, gout was absent in those with this variant. Serum urate concentration did not differ between those with the variant and the background population 309.7 ± 55.1 vs $309.1\pm80.4\,\mu\text{mol/L}$ (p=0.98). Only one individual had no evidence of CKD or hypertension (age 43.3 years at baseline). Of the remaining eight (all aged >48 years at baseline), all had hypertension with seven having evidence of CKD (based on HES codes or CKD-EPI eGFR $<60\,\text{mL/min/1.73}\,\text{m}^2$ at baseline assessment). The age of those carrying the variant did not differ from the background population: 58.5 ± 8.7 vs 57.3 ± 8.0, p=0.69.

Geographical distribution

In total, 117 individuals with the indel variant were identified from the 3 cohorts. Of interest, individual cases were found to originate from several geographical regions across England (figure 1B). Five regions had more than five cases with the highest concentration of cases in South Yorkshire (figure 1C). For the 100K and UKBB cohorts however, wider family history data were not available leading to a potential underestimation of affected individuals in these pedigrees (figure 1D).

Individuals with c.278_289delTCTGCCCCGAAG insCCGCCTCCT share an extended haplotype

All five cases in UKBB with WGS were from the group comprising white British or Irish ancestry (88%) and all nine individuals in UKBB were of European ancestry by principal component analysis. Only two of the individuals were >third-degree relatives. The nine individuals shared a haplotype around c.278_289delTCTGCCCCGAAG insCCGCCTCCT based on phasing directly genotyped common SNPs from the UKBB SNP chip (figure 3). The shared haplotype extends from 7 Mb to >30 Mb between individuals. This minimal shared haplotype occurs in <1 in 5000 of the remaining UKBB samples.

Five of the individuals in UKBB had undergone WGS. In the 1 Mb window around the c.278_289delTCTGCCCCGAAG insCCGCCTCCT variant, there were four rare variants present in individuals with the c.278_289delTCTGCCCCGAAG insCCGCCTCCT, which were rarely found in the general population (present in <1 in 250 individuals; table 2): all individuals with c.278_289delTCTGCCCCGAAG insCCGCCTCCT carried at least one of these variants. We then assessed whether the 20 individuals confirmed to carry c.278_289delTCTGCCCCGAAG insCCGCCTCCT in the 100K also carried these rare variants. All carried at least one of these rare variants, and most all four (table 2). This demonstrates a shared haplotype across all individuals with c.278_289delTCTGCCCCGAAG insCCGCCTCCT.

Historical cases

Based on published papers, we identified a further 120 individuals with the indel variant drawn from 22 unique pedigrees (table 3). Of relevance, all of these cases originated from the UK and where stated, were of white British ethnicity.[8 35–38]

DISCUSSION

In this study, we provide compelling evidence that the *UMOD* variant c.278_289delTCTGCCCCGAAG insCCGCCTCCT, widely distributed within the UK population, is a founder variant. One hundred seventeen new individuals from 33 pedigrees were ascertained from clinical referrals within a limited geographic region of England, the 100K cohort and from the UKBB population. When published cases are included, a total of 237 individuals from 55 pedigrees have now been reported with this rare variant, all from the UK.

It is worth noting that the c.278_289delTCTGCCCCGAAG insCCGCCTCCT variant was the second most common change (14 pedigrees) reported in the International ADTKD patient registry of 722 individuals, and all cases were of white British ancestry[19] (table 3). These observations are consistent with a founder effect. A previous single-centre study did not identify a common haplotype in four unrelated pedigrees based on limited analysis, the authors concluding that this was likely to be a recurrent change rather than a founder effect.[37] Based on the number of cases since reported, the unusual nature of the variant, their restricted geographical and racial origin and now evidence of a shared haplotype, the evidence now points overwhelmingly to the existence of an ancestral mutation present in the UK population.

In the International ADTKD registry, there was no difference in the median age of kidney failure reported for patients with this variant (48 years) compared with the rest of the ADTKD-*UMOD* population (124 mutations; 47 years).[19] The median age of kidney failure in our patients was however later (52 years) than previously reported. Although previous reports suggested longer kidney survival in females with ADTKD-*UMOD*,[19 39] we

did not detect a significant change in our clinical cohort. We also noted significant interfamilial and intrafamilial variability in the age of onset of kidney failure (range 32–76 years) implying that non-allelic and environmental factors can significantly modify kidney survival, as in other monogenic diseases.

Gout is a common feature in ADTKD-*UMOD*, being present in 50% of patients with a median age of onset of 28 years.[19] The paucity of gout associated with this variant is striking when compared with other *UMOD* variants and has been noted previously.[35 37] The absence of early onset gout as an alerting symptom is likely to have led to delayed clinical diagnosis but the absence of gout did not appear to impact overall kidney survival; hyperuricaemia is therefore unlikely to be a significant pathogenic factor for disease progression in ADTKD-*UMOD*. Hypertension was present in 65% of the clinical cases with available data and was also present in eight of the nine UKBB participants at baseline assessment (aged >48 years). Common *UMOD* variants have been associated with salt-sensitive hypertension, possibly by a functional interaction with the NKCC2 co-transporter.[40 41] Only one patient had significant renal cystic disease (>10 cysts/kidney).

The indel sequence results in the deletion of five amino acids and a replacement with four residues (AlaAlaSerCys), disrupting the structure of the second calcium-binding EGF-like domain.[37] The molecular mechanism underlying disease in ADTKD is more likely to relate to a dominant-negative or gain-of-function induced by the mutant protein rather than to loss-of-function.[9] The intracellular accumulation of the mutant protein likely leads to ER stress and activation of the unfolded protein response pathway.[18 42] Initial biochemical studies showed that the recombinant indel mutant protein showed a glycosylation pattern closer to the wild-type protein and was secreted more efficiently by transfected cells than another mutant (C150S).[37] However, later assays by the same group using atomic force microscopy revealed that all three different *UMOD* variants tested (Indel, C150S, C155R) showed similar premature formation of intracellular fibrillar structures compared with the wild-type protein, despite the differences in glycosylation.[38] These abnormalities are consistent with the similar age of kidney failure noted for this variant compared with others.[19]

CONCLUSION

In summary, we report a pathogenic *UMOD* variant shared by 12 apparently unrelated families with familial kidney failure within a contiguous geographical region of England but which also appears to be distributed more widely within the UK. Our results demonstrate that this is likely to represent a common ancestral variant in unrelated families rather than a recurrent variant. The absence of early onset gout associated with this variant in comparison with classical *UMOD* variants likely led to a delay in diagnosis. Our paper highlights the need for genetic testing in all cases of familial CKD of uncertain aetiology. The *UMOD* variant c.278_289delTCTGCCCCGAAG insCCGCCTCCT should be suspected in all cases of familial nephropathy presenting in patients of white British ancestry.

Author affiliations
[1]Academic Nephrology Unit, Department of Infection, Immunity and Cardiovascular Disease, The University of Sheffield Medical School, Sheffield, UK
[2]College of Medicine and Health, University of Exeter, Exeter, UK
[3]Department of Clinical Genetics, Nottingham University Hospitals NHS Trust, Nottingham, UK
[4]Department of Clinical Genetics, Sheffield Children's Hospital NHS Foundation Trust, Sheffield, UK
[5]Department of Histopathology, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK
[6]Institute of Biomedical and Clinical Science, University of Exeter, Exeter, UK
[7]Sheffield Kidney Institute, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK

Map disclaimer The inclusion of any map (including the depiction of any boundaries therein), or of any geographic or locational reference, does not imply the expression of any opinion whatsoever on the part of BMJ concerning the legal status of any country, territory, jurisdiction or area or of its authorities. Any such expression remains solely that of the relevant source and is not endorsed by BMJ. Maps are provided without any warranty of any kind, either express or implied.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval This study was approved by 100,000 Genome Project and the UK Biobank. Participants provided individual consent for genetic testing or entry into the 100,000 Genome Project and the UK Biobank Project. All data are de-identified.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

ORCID iDs
Mark Gilchrist http://orcid.org/0000-0002-0284-6048
Albert C M Ong http://orcid.org/0000-0002-7211-5400

REFERENCES

1 Foreman KJ, Marquez N, Dolgert A, Fukutaki K, Fullman N, McGaughey M, Pletcher MA, Smith AE, Tang K, Yuan C-W, Brown JC, Friedman J, He J, Heuton KR, Holmberg M, Patel DJ, Reidy P, Carter A, Cercy K, Chapin A, Douwes-Schultz D, Frank T, Goettsch F, Liu PY, Nandakumar V, Reitsma MB, Reuter V, Sadat N, Sorensen RJD, Srinivasan V, Updike RL, York H, Lopez AD, Lozano R, Lim SS, Mokdad AH, Vollset SE, Murray CJL, expectancy Flife. Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: reference and alternative scenarios for 2016-40 for 195 countries and territories. *Lancet* 2018;392:2090 doi:10.1016/S0140-6736(18)31694-5

2 Groopman EE, Marasa M, Cameron-Christie S, Petrovski S, Aggarwal VS, Milo-Rasouly H, Li Y, Zhang J, Nestor J, Krithivasan P, Lam WY, Mitrotti A, Piva S, Kil BH, Chatterjee D, Reingold R, Bradbury D, DiVecchia M, Snyder H, Mu X, Mehl K, Balderes O, Fasel DA, Weng C, Radhakrishnan J, Canetta P, Appel GB, Bomback AS, Ahn W, Uy NS, Alam S, Cohen DJ, Crew RJ, Dube GK, Rao MK, Kamalakaran S, Copeland B, Ren Z, Bridgers J, Malone CD, Mebane CM, Dagaonkar N, Fellström BC, Haefliger C, Mohan S, Sanna-Cherchi S, Kiryluk K, Fleckner J, March R, Platt A, Goldstein DB, Gharavi AG. Diagnostic utility of exome sequencing for kidney disease. *N Engl J Med* 2019;380:142–51.

3 Kirby A, Gnirke A, Jaffe DB, Barešová V, Pochet N, Blumenstiel B, Ye C, Aird D, Stevens C, Robinson JT, Cabili MN, Gat-Viks I, Kelliher E, Daza R, DeFelice M, Hůlková H, Sovová J, Vyleťal P, Antignac C, Guttman M, Handsaker RE, Perrin D, Steelman S, Sigurdsson S, Scheinman SJ, Sougnez C, Cibulskis K, Parkin M, Green T, Rossin E, Zody MC, Xavier RJ, Pollak MR, Alper SL, Lindblad-Toh K, Gabriel S, Hart PS, Regev A, Nusbaum C, Kmoch S, Bleyer AJ, Lander ES, Daly MJ. Mutations causing medullary cystic kidney disease type 1 lie in a large VNTR in MUC1 missed by massively parallel sequencing. *Nat Genet* 2013;45:299–303.

4 Zivná M, Hůlková H, Matignon M, Hodanová K, Vyleťal P, Kalbácová M, Baresová V, Sikora J, Blazková H, Zivny J, Ivánek R, Stránecký V, Sovová J, Claes K, Lerut E, Fryns J-P, Hart PS, Hart TC, Adams JN, Pawtowski A, Clemessy M, Gasc J-M, Gübler M-C, Antignac C, Elleder M, Kapp K, Grimbert P, Bleyer AJ, Kmoch S. Dominant renin gene mutations associated with early-onset hyperuricemia, anemia, and chronic kidney failure. *Am J Hum Genet* 2009;85:204–13.

5 Faguer S, Decramer S, Chassaing N, Bellanné-Chantelot C, Calvas P, Beaufils S, Bessenay L, Lengelé J-P, Dahan K, Ronco P, Devuyst O, Chauveau D, Diagnosis CD. Diagnosis, management, and prognosis of HNF1B nephropathy in adulthood. *Kidney Int* 2011;80:768–76.

6 Hart TC, Gorry MC, Hart PS, Woodard AS, Shihabi Z, Sandhu J, Shirts B, Xu L, Zhu H, Barmada MM, Bleyer AJ. Mutations of the UMOD gene are responsible for medullary cystic kidney disease 2 and familial juvenile hyperuricaemic nephropathy. *J Med Genet* 2002;39:892 doi:10.1136/jmg.39.12.882

7 Olinger E, Hofmann P, Kidd K, Dufour I, Belge H, Schaeffer C, Kipp A, Bonny O, Deltas C, Demoulin N, Fehr T, Fuster DG, Gale DP, Goffin E, Hodaňová K, Huynh-Do U, Kistler A, Morelle J, Papagregoriou G, Pirson Y, Sandford R, Sayer JA, Torra R, Venzin C, Venzin R, Vogt B, Živná M, Greka A, Dahan K, Rampoldi L, Kmoch S, Bleyer AJ, Devuyst O. Clinical and genetic spectra of autosomal dominant tubulointerstitial kidney disease due to mutations in UMOD and MUC1. *Kidney Int* 2020;98:717–31.

8 Gast C, Marinaki A, Arenas-Hernandez M, Campbell S, Seaby EG, Pengelly RJ, Gale DP, Connor TM, Bunyan DJ, Hodaňová K, Živná M, Kmoch S, Ennis S, Venkat-Raman G. Autosomal dominant tubulointerstitial kidney disease-UMOD is the most frequent non polycystic genetic kidney disease. *BMC Nephrol* 2018;19:301.

9 Eddy AA. Scraping fibrosis: UMODulating renal fibrosis. *Nat Med* 2011;17:553–5.

10 Serafini-Cessi F, Malagolini N, Cavallone D. Tamm-Horsfall glycoprotein: biology and clinical relevance. *Am J Kidney Dis* 2003;42:658–76.

11 Garimella PS, Bartz TM, Ix JH, Chonchol M, Shlipak MG, Devarajan P, Bennett MR, Sarnak MJ. Urinary Uromodulin and Risk of Urinary Tract Infections: The Cardiovascular Health Study. *Am J Kidney Dis* 2017;69:744–51.

12 Pak J, Pu Y, Zhang ZT, Hasty DL, Wu XR. Tamm-Horsfall protein binds to type 1 fimbriated Escherichia coli and prevents E. coli from binding to uroplakin Ia and Ib receptors. *J Biol Chem* 2001;276:9924–30.

13 Bates JM, Raffi HM, Prasadan K, Mascarenhas R, Laszik Z, Maeda N, Hultgren SJ, Kumar S. Tamm-Horsfall protein knockout mice are more prone to urinary tract infection: rapid communication. *Kidney Int* 2004;65:791–7.

14 Mo L, Huang H-Y, Zhu X-H, Shapiro E, Hasty DL, Wu X-R. Tamm-Horsfall protein is a critical renal defense factor protecting against calcium oxalate crystal formation. *Kidney Int* 2004;66:1159–66.

15 Mutig K, Kahl T, Saritas T, Godes M, Persson P, Bates J, Raffi H, Rampoldi L, Uchida S, Hille C, Dosche C, Kumar S, Castañeda-Bueno M, Gamba G, Bachmann S. Activation of the bumetanide-sensitive Na+,K+,2Cl- cotransporter (NKCC2) is facilitated by Tamm-Horsfall protein in a chloride-sensitive manner. *J Biol Chem* 2011;286:30200–10.

16 Renigunta A, Renigunta V, Saritas T, Decher N, Mutig K, Waldegger S. Tamm-Horsfall glycoprotein interacts with renal outer medullary potassium channel ROMK2 and regulates its function. *J Biol Chem* 2011;286:2224–35.

17 Williams SE, Reed AAC, Galvanovskis J, Antignac C, Goodship T, Karet FE, Kotanko P, Lhotta K, Morinière V, Williams P, Wong W, Rorsman P, Thakker RV. Uromodulin mutations causing familial juvenile hyperuricaemic nephropathy lead to protein maturation defects and retention in the endoplasmic reticulum. *Hum Mol Genet* 2009;18:2963–74.

18 Rampoldi L, Caridi G, Santon D, Boaretto F, Bernascone I, Lamorte G, Tardanico R, Dagnino M, Colussi G, Scolari F, Ghiggeri GM, Amoroso A, Casari G, MCKD Aof. Allelism of MCKD, FJHN and GCKD caused by impairment of uromodulin export dynamics. *Hum Mol Genet* 2003;12:3369–84.

19 Kidd K, Vyleťal P, Schaeffer C, Olinger E, Živná M, Hodaňová K, Robins V, Johnson E, Taylor A, Martin L, Izzi C, Jorge SC, Calado J, Torres RJ, Lhotta K, Steubl D, Gale DP, Gast C, Gombos E, Ainsworth HC, Chen YM, Almeida JR, de Souza CF, Silveira C, Raposeiro R, Weller N, Conlon PJ, Murray SL, Benson KA, Cavalleri GL, Votruba M, Vrbacká A, Amoroso A, Gianchino D, Caridi G, Ghiggeri GM, Divers J, Scolari F, Devuyst O, Rampoldi L, Kmoch S, Bleyer AJ. Genetic and Clinical Predictors of Age of ESKD in Individuals With Autosomal Dominant Tubulointerstitial Kidney Disease Due to UMOD Mutations. *Kidney Int Rep* 2020;5:1472–85.

20 Bollée G, Dahan K, Flamant M, Morinière V, Pawtowski A, Heidet L, Lacombe D, Devuyst O, Pirson Y, Antignac C, Knebelmann B. Phenotype and outcome in hereditary tubulointerstitial nephritis secondary to UMOD mutations. *Clin J Am Soc Nephrol* 2011;6:2429–38.

21 Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, Feldman HI, Kusek JW, Eggers P, Van Lente F, Greene T, Coresh J, CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration). A new equation to estimate glomerular filtration rate. *Ann Intern Med* 2009;150:604–12.

22 Levey AS, Eckardt K-U, Dorman NM, Christiansen SL, Hoorn EJ, Ingelfinger JR, Inker LA, Levin A, Mehrotra R, Palevsky PM, Perazella MA, Tong A, Allison SJ, Bockenhauer D, Briggs JP, Bromberg JS, Davenport A, Feldman HI, Fouque D, Gansevoort RT, Gill JS, Greene EL, Hemmelgarn BR, Kretzler M, Lambie M, Lane PH, Laycock J, Leventhal SE, Mittelman M, Morrissey P, Ostermann M, Rees L, Ronco P, Schaefer F, St Clair Russell J, Vinck C, Walsh SB, Weiner DE, Cheung M, Jadoul M, Winkelmayer WC. Nomenclature for kidney function and disease: report of a kidney

disease: improving global outcomes (KDIGO) consensus conference. *Kidney Int* 2020;97:1117–29.

23 Turro E, Astle WJ, Megy K, Gräf S, Greene D, Shamardina O, Allen HL, Sanchis-Juan A, Frontini M, Thys C, Stephens J, Mapeta R, Burren OS, Downes K, Haimel M, Tuna S, Deevi SVV, Aitman TJ, Bennett DL, Calleja P, Carss K, Caulfield MJ, Chinnery PF, Dixon PH, Gale DP, James R, Koziell A, Laffan MA, Levine AP, Maher ER, Markus HS, Morales J, Morrell NW, Mumford AD, Ormondroyd E, Rankin S, Rendon A, Richardson S, Roberts I, Roy NBA, Saleem MA, Smith KGC, Stark H, Tan RYY, Themistocleous AC, Thrasher AJ, Watkins H, Webster AR, Wilkins MR, Williamson C, Whitworth J, Humphray S, Bentley DR, Kingston N, Walker N, Bradley JR, Ashford S, Penkett CJ, Freson K, Stirrups KE, Raymond FL, Ouwehand WH, NIHR BioResource for the 100,000 Genomes Project. Whole-genome sequencing of patients with rare diseases in a national health system. *Nature* 2020;583:96–102.

24 100,000 Genomes Project Pilot Investigators, Smedley D, Smith KR, Martin A, Thomas EA, McDonagh EM, Cipriani V, Ellingford JM, Arno G, Tucci A, Vandrovcova J, Chan G, Williams HJ, Ratnaike T, Wei W, Stirrups K, Ibanez K, Moutsianas L, Wielscher M, Need A, Barnes MR, Vestito L, Buchanan J, Wordsworth S, Ashford S, Rehmström K, Li E, Fuller G, Twiss P, Spasic-Boskovic O, Halsall S, Floto RA, Poole K, Wagner A, Mehta SG, Gurnell M, Burrows N, James R, Penkett C, Dewhurst E, Gräf S, Mapeta R, Kasanicki M, Haworth A, Savage H, Babcock M, Reese MG, Bale M, Baple E, Boustred C, Brittain H, de Burca A, Bleda M, Devereau A, Halai D, Haraldsdottir E, Hyder Z, Kasperaviciute D, Patch C, Polychronopoulos D, Matchan A, Sultana R, Ryten M, Tavares ALT, Tregidgo C, Turnbull C, Welland M, Wood S, Snow C, Williams E, Leigh S, Foulger RE, Daugherty LC, Niblock O, Leong IUS, Wright CF, Davies J, Crichton C, Welch J, Woods K, Abulhoul L, Aurora P, Bockenhauer D, Broomfield A, Cleary MA, Lam T, Dattani M, Footitt E, Ganesan V, Grunewald S, Compeyrot-Lacassagne S, Muntoni F, Pilkington C, Quinlivan R, Thapar N, Wallis C, Wedderburn LR, Worth A, Bueser T, Compton C, Deshpande C, Fassihi H, Haque E, Izatt L, Josifova D, Mohammed S, Robert L, Rose S, Ruddy D, Sarkany R, Say G, Shaw AC, Wolejko A, Habib B, Burns G, Hunter S, Grocock RJ, Humphray SJ, Robinson PN, Haendel M, Simpson MA, Banka S, Clayton-Smith J, Douzgou S, Hall G, Thomas HB, O'Keefe RT, Michaelides M, Moore AT, Malka S, Pontikos N, Browning AC, Straub V, Gorman GS, Horvath R, Quinton R, Schaefer AM, Yu-Wai-Man P, Turnbull DM, McFarland R, Taylor RW, O'Connor E, Yip J, Newland K, Morris HR, Polke J, Wood NW, Campbell C, Camps C, Gibson K, Koelling N, Lester T, Németh AH, Palles C, Patel S, Roy NBA, Sen A, Taylor J, Cacheiro P, Jacobsen JO, Seaby EG, Davison V, Chitty L, Douglas A, Naresh K, McMullan D, Ellard S, Temple IK, Mumford AD, Wilson G, Beales P, Bitner-Glindzicz M, Black G, Bradley JR, Brennan P, Burn J, Chinnery PF, Elliott P, Flinter F, Houlden H, Irving M, Newman W, Rahman S, Sayer JA, Taylor JC, Webster AR, Wilkie AOM, Ouwehand WH, Raymond FL, Chisholm J, Hill S, Bentley D, Scott RH, Fowler T, Rendon A, Caulfield M. 100,000 Genomes Pilot on Rare-Disease Diagnosis in Health Care - Preliminary Report. *N Engl J Med* 2021;385:1868–80

25 Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TFC, McCarroll SA, Visscher PM. Finding the missing heritability of complex diseases. *Nature* 2009;461:747–53.

26 Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J, Cortes A, Welsh S, Young A, Effingham M, McVean G, Leslie S, Allen N, Donnelly P, Marchini J. The UK Biobank resource with deep phenotyping and genomic data. *Nature* 2018;562:203–9.

27 Backman JD, Li AH, Marcketta A, Sun D, Mbatchou J, Kessler MD, Benner C, Liu D, Locke AE, Balasubramanian S, Yadav A, Banerjee N, Gillies CE, Damask A, Liu S, Bai X, Hawes A, Maxwell E, Gurski L, Watanabe K, Kosmicki JA, Rajagopal V, Mighty J, Jones M, Mitnaul L, Stahl E, Coppola G, Jorgenson E, Habegger L, Salerno WJ, Shuldiner AR, Lotta LA, Overton JD, Cantor MN, Reid JG, Yancopoulos G, Kang HM, Marchini J, Baras A, Abecasis GR, Ferreira MAR, Regeneron Genetics Center, DiscovEHR. Exome sequencing and analysis of 454,787 UK Biobank participants. *Nature* 2021;599:628–34.

28 Delaneau O, Zagury J-F, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat Methods* 2013;10:5–6.

29 Jones SE, Lane JM, Wood AR, van Hees VT, Tyrrell J, Beaumont RN, Jeffries AR, Dashti HS, Hillsdon M, Ruth KS, Tuke MA, Yaghootkar H, Sharp SA, Jie Y, Thompson WD, Harrison JW, Dawes A, Byrne EM, Tiemeier H, Allebrandt KV, Bowden J, Ray DW, Freathy RM, Murray A, Mazzotti DR, Gehrman PR, Lawlor DA, Frayling TM, Rutter MK, Hinds DA, Saxena R, Weedon MN. Genome-Wide association analyses of chronotype in 697,828 individuals provides insights into circadian rhythms. *Nat Commun* 2019;10:343.

30 Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. Highly accurate protein structure prediction with AlphaFold. *Nature* 2021;596:583–9.

31 Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 2004;25:1605–12.

32 Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, Ben-Tal N. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res* 2016;44:W344–50.

33 Gligorijević V, Renfrew PD, Kosciolek T, Leman JK, Berenberg D, Vatanen T, Chandler C, Taylor BC, Fisk IM, Vlamakis H, Xavier RJ, Knight R, Cho K, Bonneau R. Structure-Based protein function prediction using graph convolutional networks. *Nat Commun* 2021;12:3168.

34 Eckardt K-U, Alper SL, Antignac C, Bleyer AJ, Chauveau D, Dahan K, Deltas C, Hosking A, Kmoch S, Rampoldi L, Wiesener M, Wolf MT, Devuyst O, Kidney Disease: Improving Global Outcomes. Autosomal dominant tubulointerstitial kidney disease: diagnosis, classification, and management--A KDIGO consensus report. *Kidney Int* 2015;88:683 doi:10.1038/ki.2015.28

35 Hateboer N, Gumbs C, Teare MD, Coles GA, Griffiths D, Ravine D, Futreal PA, Rahman N. Confirmation of a gene locus for medullary cystic kidney disease (MCKD2) on chromosome 16p12. *Kidney Int* 2001;60:1239 doi:10.1046/j.1523-1755.2001.00932.x

36 Wolf MTF, Mucha BE, Attanasio M, Zalewski I, Karle SM, Neumann HPH, Rahman N, Bader B, Baldamus CA, Otto E, Witzgall R, Fuchshuber A, Hildebrandt F. Mutations of the uromodulin gene in MCKD type 2 patients cluster in exon 4, which encodes three EGF-like domains. *Kidney Int* 2003;64:1587 doi:10.1046/j.1523-1755.2003.00269.x

37 Smith GD, Robinson C, Stewart AP, Edwards EL, Karet HI, Norden AGW, Sandford RN, Karet Frankl FE. Characterization of a recurrent in-frame UMOD indel mutation causing late-onset autosomal dominant end-stage renal failure. *Clin J Am Soc Nephrol* 2011;6:2766–74. doi:10.2215/CJN.06820711

38 Stewart AP, Sandford RN, Karet Frankl FE, Edwardson JM. Pathogenic uromodulin mutations result in premature intracellular polymerization. *FEBS Lett* 2015;589:89–93. doi:10.1016/j.febslet.2014.11.029

39 Moskowitz JL, Piret SE, Lhotta K, Kitzler TM, Tashman AP, Velez E, Thakker RV, Kotanko P. Association between genotype and phenotype in uromodulin-associated kidney disease. *Clin J Am Soc Nephrol* 2013;8:1349–57. doi:10.2215/CJN.11151012

40 Trudu M, Janas S, Lanzani C, Debaix H, Schaeffer C, Ikehata M, Citterio L, Demaretz S, Trevisani F, Ristagno G, Glaudemans B, Laghmani K, Dell'Antonio G, Loffing J, Rastaldi MP, Manunta P, Devuyst O, Rampoldi L, SKIPOGH team. Common noncoding UMOD gene variants induce salt-sensitive hypertension and kidney damage by increasing uromodulin expression. *Nat Med* 2013;19:1660 doi:10.1038/nm.3384

41 Padmanabhan S, Graham L, Ferreri NR, Graham D, McBride M, Dominiczak AF. Uromodulin, an emerging novel pathway for blood pressure regulation and hypertension. *Hypertension* 2014;64:918–23. doi:10.1161/HYPERTENSIONAHA.114.03132

42 Schaeffer C, Merella S, Pasqualetto E, Lazarevic D, Rampoldi L. Mutant uromodulin expression leads to altered homeostasis of the endoplasmic reticulum and activates the unfolded protein response. *PLoS One* 2017;12:e0175970. doi:10.1371/journal.pone.0175970