Deep phenotyping and precision medicine approaches to understand and treat Autosomal Dominant Tubulointerstitial Kidney Disease due to *UMOD* mutations

Dr Holly Rachel Mabillard

Supervisors:
Professor John A Sayer
Professor Heather J Cordell

This thesis is submitted for the degree of Doctor of Philosophy

Translational and Clinical Research Institute
August 2025

Author's Declaration

This thesis is submitted for the degree of Doctor of Philosophy at Newcastle University. I, Holly Rachel Mabillard, declare that the work described here is my own, unless where clearly acknowledged and stated otherwise. I certify that I have not submitted any of the material in this thesis for a degree qualification at this or any other university.

Dr Holly Rachel Mabillard

Abstract

Autosomal Dominant Tubulointerstitial Kidney Disease due to *UMOD* variants (ADTKD-*UMOD*) is a rare monogenic kidney disorder, historically viewed as a Mendelian condition with uniform clinical expression. However, growing evidence reveals significant variability in disease severity, progression, and phenotype. This thesis presents the most comprehensive characterisation of ADTKD-*UMOD* undertaken in the UK, integrating clinical, genetic, and biomarker data across multiple cohorts to investigate the complex modifiers of disease expression and progression.

Using data from the UK RaDaR registry, local cohorts, and UK Biobank, this work defines the national spectrum of *UMOD* variants and shows that the recurrent *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant likely represents a UK-specific founder variant. Despite this genetic homogeneity, significant intra-genotypic heterogeneity was observed in gout onset, salt-wasting symptoms, age of kidney failure, and kidney imaging. Genome-wide association studies (GWAS) and gene burden analyses identified genetic modifiers of disease progression, including loci involved in uric acid handling (e.g. *ABCG2*), mitochondrial function (e.g. *COX6B2*), and transcriptional regulation (e.g. *KMT5C*). Chromatin interaction mapping suggested a regulatory link between *ABCG2* and *PKD2,* indicating shared pathways between cystic and tubulointerstitial kidney diseases. Polygenic risk scores for chronic kidney disease were independently associated with kidney survival, emphasising the contribution of background genetic architecture in a monogenic context.

Biomarker analyses confirmed that serum and urinary uromodulin concentrations correlate with disease severity beyond the primary *UMOD* variant and kidney function and may serve as prognostic tools. Individuals with the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant showed partial preservation of uromodulin secretion, suggesting variant-specific trafficking and a unique metabolic sub-phenotype with lower serum urate and absence of gout.

Together, these findings reposition ADTKD-*UMOD* as a model of complex monogenic disease and lay the groundwork for genetically stratified trials and therapeutic targeting, with broader relevance to chronic kidney disease.

Acknowledgements

This thesis would not have been possible without the unwavering support, encouragement, and inspiration of many people to whom I am deeply grateful.

First and foremost, to my husband, the most supportive and encouraging partner I could ever ask for, and an incredible father to our children. I owe you more than I can express and you have been the foundation beneath this journey. You have not only carried the load during the hardest moments, but have willingly placed my ambitions above your own, making sacrifices in your own career so that I could pursue mine.

To my two wonderful children, Luke and Sienna, thank you for being my greatest motivation and for reminding me what truly matters.

To my parents, Alison and John, thank you for the countless opportunities and role modelling you've given me and for instilling in me the values of education, purpose and curiosity. To my parents in law, Gill and Ian, thank you for your support, kindness, and encouragement throughout this journey.

To my supervisor, Professor John Sayer, thank you for seeing my potential long before I saw it myself. Your mentorship has helped me grow in confidence, ambition, and self-belief. I am immeasurably grateful for the opportunities you have created for me and for the generosity with which you share your knowledge, time, and belief in others.

To my supervisor, Professor Heather Cordell, thank you for being such a constant source of support, wisdom, kindness and inspiration throughout this process. Your encouragement and reassurance have been invaluable, both professionally and personally.

To my colleague, Dr Eric Olinger, thank you for so generously devoting your time to teach me the foundations I needed to step away from clinical training and pursue a PhD. Your guidance opened the door to this opportunity and gave me the tools to make the most of it.

Finally, to the patients and families who live with ADTKD-*UMOD*, thank you for inspiring and supporting me with this work and for reminding me why it matters.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

Despite major advances in medical science, chronic kidney disease (CKD) remains a silent epidemic; under-recognised, underdiagnosed, and under-prioritised (Borg *et al.*, 2023). Its progression is often insidious, its early symptoms easily missed, and its clinical burden devastating (Chertow *et al.*, 2024; Francis *et al.*, 2024; Global, regional, and national burden of chronic kidney disease, 2020). Unlike cancer or cardiovascular disease, CKD rarely commands headlines, yet it affects more than one in ten people globally and disproportionately impacts the most vulnerable (Borg *et al.*, 2023). At its core, lies a final common pathway, interstitial fibrosis, that transcends diverse aetiologies and is responsible for nearly all cases of progressive kidney function decline (Huang, Fu and Ma, 2023). Within this spectrum, genetic kidney diseases offer a unique lens through which to examine CKD pathogenesis in its most pure form (Peltonen *et al.*, 2006).

This thesis centres on one such condition, Autosomal Dominant Tubulointerstitial Kidney Disease due to *UMOD* variants (ADTKD-*UMOD*), a rare, inherited disorder that exemplifies the challenges of precision nephrology and the opportunities it holds. As a monogenic disease in which tubulointerstitial fibrosis is the primary and defining pathology, ADTKD-*UMOD* provides an opportunistic human model of kidney fibrosis (Devuyst *et al.*, 2014, 2019). Through the study of this condition, this work seeks not only to unravel the complex modifiers of disease progression in a monogenic context, but also to illuminate broader mechanisms relevant to fibrosis and biomarker discovery in CKD.

1.1 Chronic Kidney Disease

Chronic kidney disease is a major public health burden with limited treatments, poor outcomes, and high cost. Furthermore, CKD has one of the fastest rising incidence rates of all diseases and a current estimated global prevalence of 11-13% with an even higher hypothesised true prevalence (Hill *et al.*, 2016). The Global Burden of Disease (GBD) study (2015) estimated that kidney disease was directly responsible for 1.2 million deaths globally which had increased by 32% in the previous 10 years (GBD Chronic Kidney Disease Collaboration, 2020). Strikingly, 44% of individuals over the age of 60 years were found to be living with undiagnosed CKD on screening as part of the Oxford Health Study (Hirst *et al.*, 2020). There are many risk factors for CKD (Kazancioğlu, 2013) and having kidney disease is

associated with increased cardiovascular morbidity, mortality and premature death (Tonelli *et al.*, 2006).

CKD is associated with psychosocial problems, poorer income and educational attainment, reduced wellbeing and quality of life in comparison to those without it. Huge health inequalities exist both socially and globally (Stanifer *et al.*, 2016; Luyckx, Cherney and Bello, 2020) for CKD which greatly affect outcomes and result regarding inadequate and unequitable access to healthcare (Levin *et al.*, 2017; Sever *et al.*, 2021).

The socio-economic burden of CKD is huge. The proportion of patients with CKD who experienced hospitalisation was up to 698% higher than those with normal kidney function and hospital stays are significantly longer for this group (Freeman, 2018). Direct costs of healthcare for a patient with CKD stage 1-2 compared to one without was 49% higher. These costs rise as CKD progresses. The National Health Service (NHS) spends more than one billion pounds treating kidney-associated disease annually (Freeman, 2018).

It is therefore no surprise that kidney disease has been described as the most neglected chronic disease. Focus must be diverted further from reaction and towards prevention if we want to reduce the multifaceted burden of CKD (Vanholder *et al.*, 2017). The United Nations now have 'tackling kidney disease' added to their list of sustainable development goals evidencing increased recognition of the need to act (Luyckx, Tonelli and Stanifer, 2018).

Known and hypothesised causes of CKD involve metabolic conditions such as diabetes, environmental exposures including nephrotoxic medications and heavy metals, socioeconomic and structural determinants of health, and genetic factors (Webster *et al.*, 2017; Luyckx, Tonelli and Stanifer, 2018; KDIGO Conference Participants, 2022). Over one hundred genetic loci have now been associated with CKD through trans-ancestry meta-analyses of genome-wide association studies which may inform polygenic risk scores to aid earlier detection of the disease (Wuttke *et al.*, 2019). Despite this, molecular mechanisms associated with these loci remain undetermined. Efforts should therefore be channelled into deciphering genetic and molecular mechanisms involved in all forms of CKD to enhance early detection, treatment options and prevention (Zoccali *et al.*, 2017).

1.2 Genetic Kidney Disease

At least 10% of adults and almost all children progressing to end-stage kidney disease (ESKD) suffer from a monogenic disease (Devuyst *et al.*, 2014). Monogenic causes of CKD have been detected in up to 42% of some cohorts (Al-Hamed *et al.*, 2022).

Monogenic (or Mendelian) diseases are often rare diseases and are often highly heterogeneous. The definition of a rare disease depends on which country you find yourself in. An example of the spectrum of definitions include a disease that affects 'less than 200 000 people nationally' (United States), 'less than 1 in 2000 people' (Europe) (Schieppati *et al.*, 2008) or 'less than 1 in 500 000 people' (China) (Liu *et al.*, 2010).

There is an identified genetic origin in about 80% of rare diseases and about 30 million Europeans have been affected by one of the 6000-8000 rare diseases (Schieppati *et al.*, 2008) with incidence varying greatly between ethnic groups.

However, not all inherited kidney diseases are rare diseases. Autosomal Dominant Polycystic Kidney Disease (ADPKD) has a prevalence of 1 in 1000 people and because of this, large pressure groups of patients have drawn much attention to inherited kidney diseases which has resulted in research funding, evolution of treatments and even genetic kidney disease clinics.

Rare kidney diseases represent the fifth most common cause of ESKD after diabetes, hypertension, glomerulonephritis and pyelonephritis (Devuyst *et al.*, 2014). 150 different disorders have been identified as rare kidney diseases and have a prevalence of 60-80 cases per 100 000 in Europe and the United States (Soliman, 2012; Nguengang Wakap *et al.*, 2020).

Due to the availability of dialysis and kidney transplantation, those who have a rare kidney disease can usually live for many years despite disease progression; however, quality of life and general health is often poor. This is often due to the consequences of disease in childhood affecting physical, psychological and social development. In addition, the multisystem nature of many of these diseases can pose a huge burden to health and wellbeing.

Patients with rare diseases, or aptly named 'orphan diseases', are done a huge disservice in healthcare. Those that receive a diagnosis for their rare disease wait, on average, four years

for this, and 22% have visited more than five clinicians before receiving their diagnosis (Nutt, S., Limb, L., 2010; Muir, E., 2016). In a study of 3315 patients with CKD who had whole exome sequencing performed, diagnostic variants were detected in 9.3%. Of those who had pathogenic variants detected, 17.1% had CKD of 'unknown cause', highlighting widespread diagnostic struggles (Groopman *et al.*, 2019). Typical challenges from a physician's lens include variable phenotypes, unknown genetic causes, insufficient ontology, scattered patient cohorts with fragmented clinical data, absence of biomarkers, lack of standardised diagnostic procedures, carrier status implications, insufficient disease models, poor clinician knowledge on disease presentation and mechanisms in addition to poor genetic literacy (Hayashi and Umeda, 2008).

High throughput Next Generation Sequencing (NGS) technologies have allowed us to define over 160 rare kidney diseases (Devuyst *et al.*, 2014). This is a rapid advancement from the first use of genetics in 1985 to map ADPKD (Reeders *et al.*, 1985). Soon, disease-specific NGS panels, affordable exome sequencing and single-cell NGS will be widely available and should improve diagnostic yield, but genetic literacy of clinicians needs to catch up in order for this to be fully integrated into healthcare. Gene editing of appropriate model organisms, precision diagnostic tools and personalised cell systems are offering improved ways to study disease. Disease registries, networks, databases and biorepositories are becoming ever more prevalent to better overcome cohort fragmentation. This, along with unified international collaborative efforts, means there is hope of a better life for patients with rare and inherited diseases.

1.3 Autosomal Dominant Tubulointerstitial Kidney Disease (ADTKD)

Autosomal Dominant Tubulointerstitial Kidney Disease (ADTKD) is a form of genetic kidney disease defined by progressive kidney dysfunction, bland urinalysis, and interstitial fibrosis with tubular atrophy (IF/TA) seen on kidney histology. It encompasses a genetically heterogeneous group of disorders, each defined by variants in different genes but sharing a common histopathological and clinical phenotype.

The most common subtypes are ADTKD-*UMOD* and ADTKD-*MUC1*, caused by variants in the *UMOD* and *MUC1* genes, respectively. *UMOD* encodes uromodulin (also known as Tamm-Horsfall protein), and variants typically lead to intracellular retention and toxic accumulation within the thick ascending limb of the nephron. *MUC1* variants result in the production of an

abnormal mucin protein fragment (MUC1-fs) that accumulates in kidney cells but is challenging to detect with conventional sequencing technologies.

Rarer subtypes include:

- **ADTKD-*REN***: caused by variants in the *REN* gene encoding renin, often presenting with early-onset anaemia, hypotension, and hyperuricemia in childhood (Zivná *et al.*, 2009).
- **ADTKD-*HNF1B***: associated with variants in *HNF1B*, a transcription factor important in kidney and pancreatic development. This form is often syndromic and may present with renal cysts, diabetes, hypomagnesemia, or genital tract anomalies (Eckardt *et al.*, 2015; Bleyer *et al.*, 2022).
- **ADTKD-*SEC61A1***: due to variants in *SEC61A1*, which is involved in protein translocation in the endoplasmic reticulum; this subtype is extremely rare and poorly understood (Bolar *et al.*, 2016).
- **ADTKD-*DNAJB11***: considered an atypical form that mimics features of both tubulointerstitial and cystic kidney diseases. *DNAJB11* encodes a co-chaperone involved in protein folding, and variants may present with milder cystic changes and slowly progressive CKD (Huynh *et al.*, 2020).
- **ADTKD-*APOA4***: an emerging and very rare subtype caused by variants in *APOA4*, involved in lipid metabolism, recently associated with interstitial kidney disease and where medullary amyloid is detected on kidney biopsy (Kmochová *et al.*, 2024).

In addition, **mitochondrial variants** have also been identified as rare causes of ADTKD-like presentations. These may co-occur with multisystem involvement or be isolated to the kidney (Lorenz *et al.*, 2020).

ADTKD is now thought to be one of the most common genetic kidney diseases after Autosomal Dominant Polycystic Kidney Disease (ADPKD) and Alport Syndrome (AS) **(figure 1),** accounting for an estimated 5% of all monogenic causes of CKD (Gast *et al.*, 2018; Groopman *et al.*, 2019). Of course, the NGS era has resulted in increasing recognition and improved definition of this genetically heterogenous disease, however true prevalence probably remains underestimated due to limited recognition by clinicians and the challenges associated with obtaining a genetic diagnosis in ADTKD-*MUC1*.

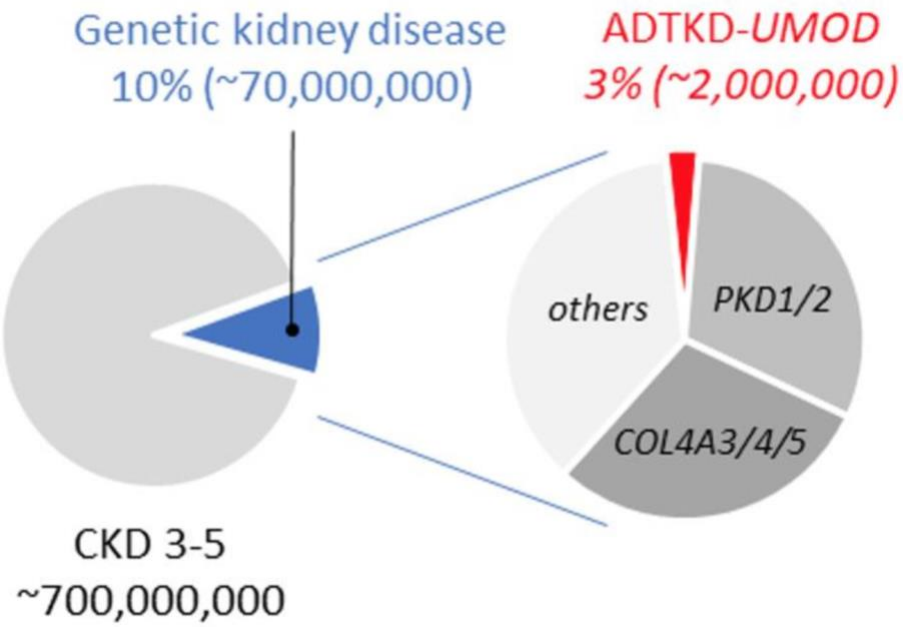

Figure 1: Global prevalence of genetic kidney disease and estimated prevalence of ADTKD-*UMOD*

Since its discovery in 2006, global efforts to amass patient cohorts and determine underlying disease pathways in ADTKD have resulted in critical insights into this disease and how it might aid understanding of common fibrotic mechanisms in CKD.

As with other rare diseases, ADTKD has suffered from its previously confusing nomenclature. Previous umbrella terms of 'familial juvenile hyperuricaemic nephropathy' and 'medullary cystic kidney disease' (amongst other variations) were used to describe the unifying features of (usually isolated) kidney interstitial fibrosis and tubular atrophy and an autosomal dominant inheritance pattern (Ekici *et al.*, 2014; Eckardt *et al.*, 2015). Improved molecular genetics have resulted in re-definition of the six aforementioned ADTKD subtypes where ADTKD is affixed with its causal gene (e.g. ADTKD-*UMOD*) or, if not yet determined, is labelled ADTKD-*NOS* (not otherwise specified) (Eckardt *et al.*, 2015). This refined definition has resulted in increased clinical recognition of ADTKD as a fundamental subtype of monogenic kidney disease. Furthermore, this has increased international effort and collaboration which has been critical to development and new understanding of genetic and phenotypic spectra, correlations and diagnostic approaches.

Nine years after discovery, ADTKD received official recognition by Kidney Disease: Improving Global Outcomes (KDIGO) in 2015 where diagnostic and management strategies were first outlined on the global stage (Eckardt *et al.*, 2015). A positive family history of progressive CKD,

bland urine sediment or absence of proteinuria and normal/small sized kidneys should raise clinical suspicion of ADTKD**.** Genetic testing should then be the first-line diagnostic approach so a definitive diagnosis can be made by identification of a variant in one of the known causative genes (Eckardt *et al.*, 2015). Kidney biopsy should be avoided in favour of genetic testing but typical histology findings include IF/TA, lamellation and thickening of the tubular basement membrane and tubular dilatations (Eckardt *et al.*, 2015; Ayasreh *et al.*, 2018; Onoe *et al.*, 2021). As additional causative genetic loci still require identification, an absence of variant does not exclude a diagnosis of ADTKD (Devuyst *et al.*, 2019). Rate of kidney function decline varies greatly both within and between families often despite the same primary variant. This is thought to be due to the influence of environmental, clinical and genetic modifiers but reasons for this remain largely unknown (Bleyer *et al.*, 2014; Kidd *et al.*, 2020; Olinger *et al.*, 2020).

As the risk of transmission to the disease offspring is 50%, genetic counselling should be offered to all who are diagnosed. Other than renal transplantation for ESKD, there are currently no disease-specific therapies available and any recommendations around management are based on limited evidence (Eckardt *et al.*, 2015; Devuyst *et al.*, 2019). Clinicians should follow the general recommendations for CKD management (Eckardt *et al.*, 2015) (with the exception of ACE-inhibitors in ADTK-*REN* which should be avoided) and disease does not recur after kidney transplantation of which outcomes are no worse than for other forms of CKD. See **figure 2** for a typical family pedigree and patient journey with ADTKD-*UMOD*.

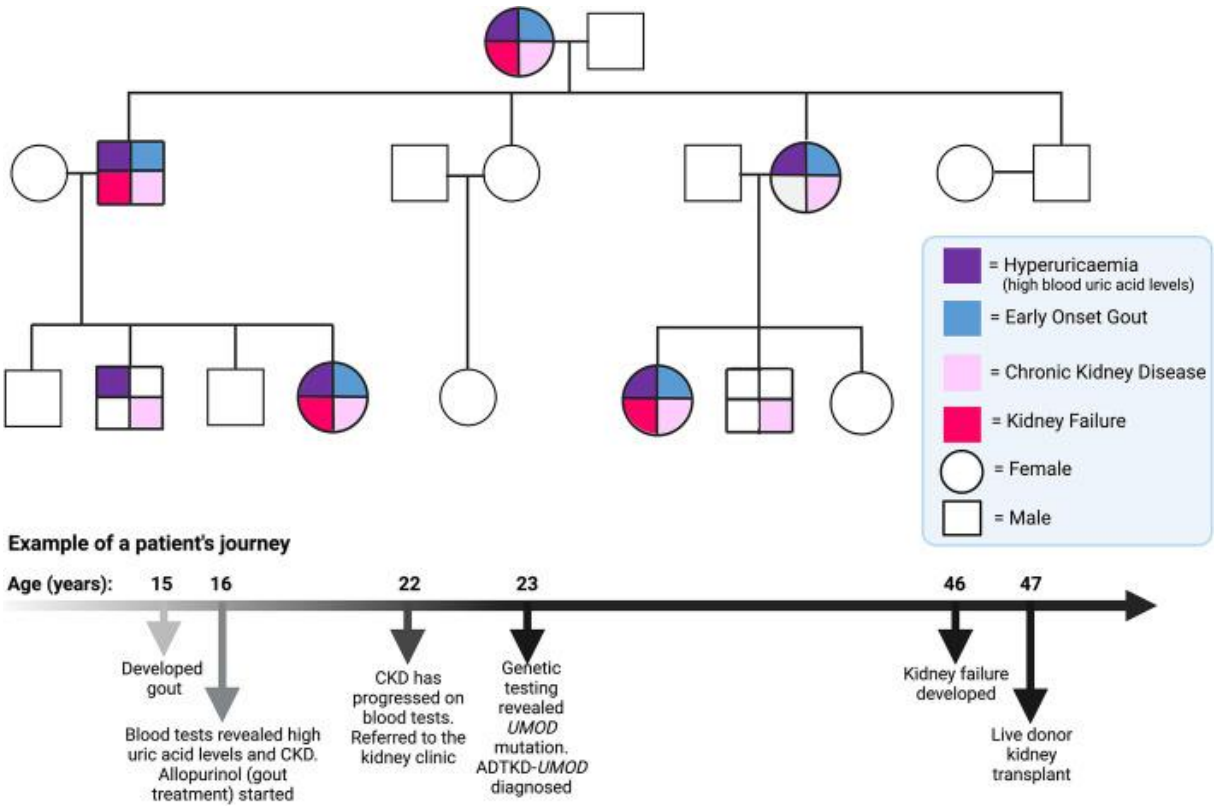

Figure 2: A typical pedigree of a family with ADTKD-*UMOD**

*[*Taken from (Mabillard, Olinger and Sayer, 2022) which demonstrates a typical pedigree of a family with ADTKD-UMOD and a typical patient's journey from birth to diagnosis to kidney failure and beyond.]*

A few therapies exist to aid the management of some ADTKD subtypes. Fludrocortisone and erythropoietin therapy ameliorate hyperkalaemia, acidosis, hypotension and severe anaemia in ADTKD-*REN* and fludrocortisone has been associated with increases in eGFR in some cases (Zivná *et al.*, 2009; Bleyer *et al.*, 2010). Optimistically, a recent high-content screen to identify candidate therapies in ADTKD-*MUC1* has revealed a promising therapeutic candidate which is being investigated further and may become the first mechanistic therapy in ADTKD (Dvela-Levitt *et al.*, 2019).

As described in the previous section, ADTKD suffers with many of the challenges associated with rare diseases. However, recent novel insights through international collaboration and cohort amalgamation are improving momentum in the avenues to treatment for this patient group (Huynh *et al.*, 2020; Izzi *et al.*, 2020; Kidd *et al.*, 2020; Olinger *et al.*, 2020; Živná *et al.*, 2020). There is, however, urgent need for disease-specific biomarkers to allow us to better

monitor disease activity when CKD progression is slow if we want faster clinical trials and interest from pharmaceutical industry.

1.4 ADTKD due to UMOD Variants (ADTKD-UMOD)

The most abundant protein found in mammalian urine during physiological conditions is uromodulin (Tamm-Horsfall protein), an 85 kDa glycoprotein which is encoded by the *UMOD* gene (NM_003361.4). Uromodulin (NP_003352.2) is exclusively expressed in epithelial cells of the thick ascending limb of the Loop of Henle in the nephron (Schaeffer, Devuyst and Rampoldi, 2021).

Uromodulin is a fascinating protein with multiple significant biological functions **(figure 3).** It regulates electrolyte transport in the thick ascending limb of the Loop of Henle and early distal convoluted tubules**.** It protects against urinary tract infections and calcium containing kidney stone formation. It also plays a role in both local and systemic immunomodulation (Schaeffer, Devuyst and Rampoldi, 2021).

After engaging all 48 conserved cysteine residues in the formation of 24 intramolecular disulphide bonds, uromodulin leaves the endoplasmic reticulum (ER) and enters the cell's secretory pathway. Uromodulin is heavily glycosylated in the Golgi and polymerizes in the urine after hepsin-mediated cleavage at the cell membrane owing to the bipartite zona pellucida domain of the protein **(figure 4)**. Other structural elements of uromodulin include four epidermal growth factor (EGF-like) domains and a cysteine-rich domain (D8C) (Schaeffer, Devuyst and Rampoldi, 2021).

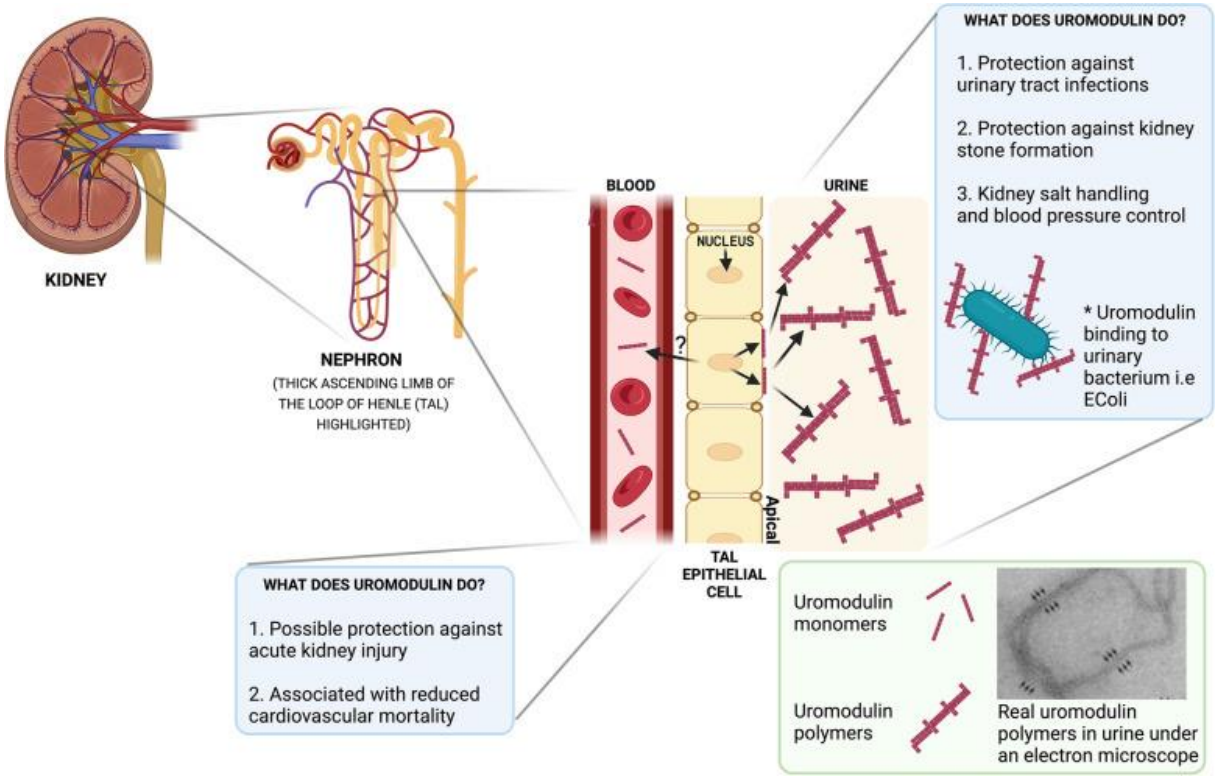

Figure 3: Uromodulin functions in urine and the systemic circulation*

*[*Taken from (Mabillard, Olinger and Sayer, 2022) which demonstrates where uromodulin is produced in the Thick Ascending Limb of the Loop of Henle (TAL) in addition to some of its functions in urine and the systemic circulation. We now also understand that uromodulin is also present in its monomeric form in the urine* (Nanamatsu *et al.*, 2024) *but the reasons for its presence here is yet to be determined.]*

Variants in *UMOD* are the leading cause of ADTKD resulting in the ADTKD-*UMOD* subtype. *UMOD* variants cluster in exons 3 and 4 (Olinger *et al.*, 2020). *UMOD*-deficient mice do not develop ADTKD and the *UMOD* gene is not constrained for predicted loss-of-function variants demonstrating a toxic gain-of-function mechanism in ADTKD-*UMOD* (Devuyst *et al.*, 2019). Mutant uromodulin is retained in the ER due to impaired trafficking. Some of the mutant protein does escape quality control in the ER and forms aggregates at the plasma membrane but mostly, reduced urinary levels of uromodulin are seen (Bernascone *et al.*, 2006; Schaeffer *et al.*, 2012; Johnson *et al.*, 2017; Piret *et al.*, 2017). ER accumulation results in ER stress (also shown in human kidneys), activation of the Unfolded Protein Response (UPR) and different branches of this have been implicated in the disease in *in vitro* and *in vivo* models (Schaeffer, C., et al., 2017). Conflict as to whether apoptosis is increased has arisen where this has been seen *in vitro* but not *in vivo* (Johnson *et al.*, 2017; Piret *et al.*, 2017). Inflammation occurs followed by mitochondrial dysfunction, impaired energy metabolism and fibrosis (Kemter *et*

al., 2017; Trudu *et al.*, 2017). Recently, it has been demonstrated in a mouse model of ADTKD-*UMOD* that mesencephalic astrocyte-derived neurotrophic factor (MANF) enhances autophagy/mitophagy, clears mutant UMOD, promotes mitochondrial biogenesis via p-AMPK activation, and protects kidney function, highlighting MANF as a potential biotherapeutic for proteinopathies such as this (Kim *et al.*, 2023).

TAL dysfunction and reduced expression of the sodium-potassium-chloride (NKCC2) cotransporter has been evidenced in mouse and human tissue (Bernascone *et al.*, 2006; Labriola *et al.*, 2015; Ma *et al.*, 2017). Further to this, the mechanism of hyperuricemia has always been intriguing. It has been widely hypothesised that upregulation of the proximal tubule Na+-coupled urate transporters are upregulated to compensate for volume contraction which is thought to drive hyperuricaemia but this is yet to be investigated (Ma *et al.*, 2017).

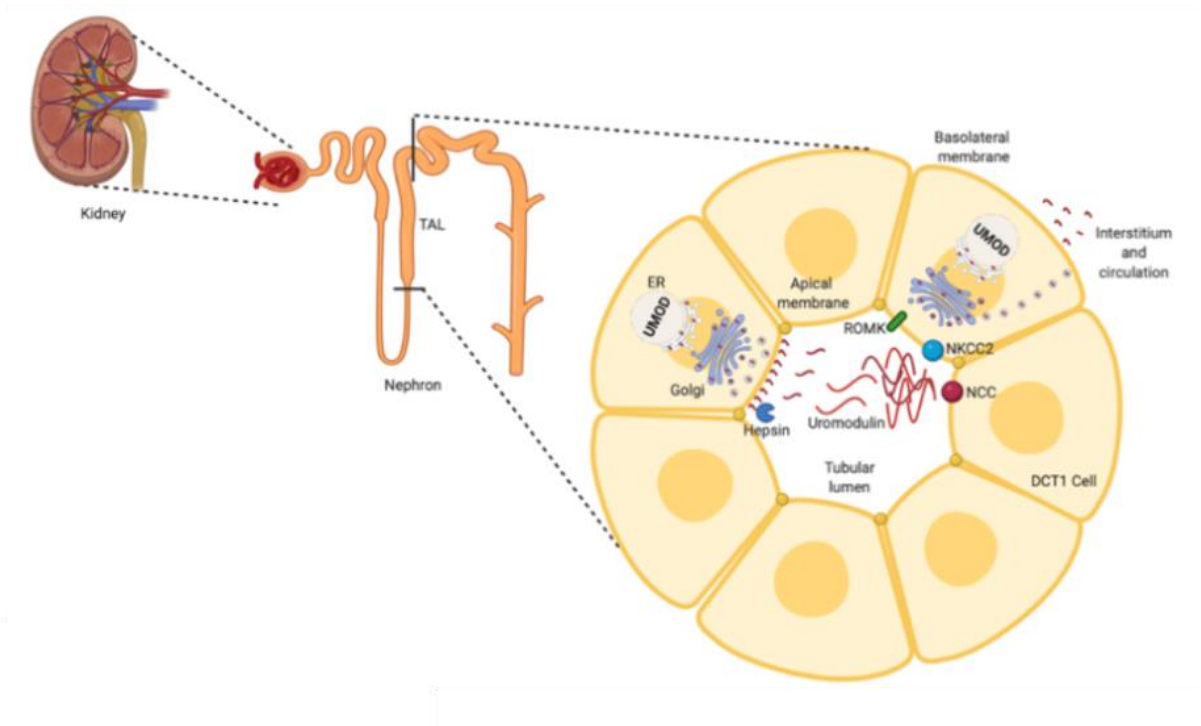

Figure 4: Uromodulin production and secretion in the thick ascending limb of the loop of Henle*

*[*A schematic illustration adapted from (Ponte et al., 2021) of uromodulin production and secretion in the thick ascending limb (TAL) of the nephron, with focus on its intracellular trafficking and apical release into the tubular lumen. Uromodulin (encoded by the UMOD gene) is synthesised in the endoplasmic reticulum (ER) of TAL cells, where it undergoes proper folding and post-translational modifications. It is then transported through the Golgi apparatus, glycosylated, and then transported to the apical membrane, where it is cleaved by the serine*

protease, hepsin, and released into the tubular lumen. Here, it polymerises and forms a major component of normal urine. A smaller proportion of uromodulin is secreted basolaterally into the interstitium and circulation, where it is measurable in serum and may have systemic effects. In Autosomal Dominant Tubulointerstitial Kidney Disease caused by UMOD variant*s (ADTKD-UMOD)***,** *pathogenic variants typically result in misfolded uromodulin protein that fails to exit the ER. This leads to intracellular retention, ER stress, and activation of the unfolded protein response (UPR). Chronic ER stress contributes to tubular cell dysfunction, apoptosis, and progressive interstitial fibrosis. This defective trafficking of uromodulin reduces its secretion into both urine and serum, impairing its physiological roles.]*

ADTKD-*UMOD* is thought to have a prevalence of 1% in those with CKD stage 3-5 and 2% in those with ESKD (Gast *et al.*, 2018). In a cohort of >3000 patients mostly with ESKD who had undergone whole exome sequencing, a *UMOD* variant was detected in 0.3% of the cohort and in 3% of all patients with a diagnosis of monogenic kidney disease (Groopman *et al.*, 2019).

ADTKD-*UMOD* presents with progressive CKD, absence of protein or blood on urinalysis, blood pressure is typically normal or elevated and ultrasound reveals normal or small-sized kidneys. Biopsy findings of FSGS with partial podocyte effacement and sub-nephrotic proteinuria can be encountered, usually as a consequence of late presentation with secondary FSGS once kidney damage has become more advanced (Chun *et al.*, 2020). Although an autosomal dominant pattern of inheritance is typically found, de novo presentations can occur (Olinger *et al.*, 2020).

ADTKD-*UMOD* has a high prevalence of hypouricosuric hyperuricaemia and gout (preceding CKD onset) which is the only extra-renal manifestation in this subtype of ADTKD. Gout was present in 75% of men and 50% of women with a median age of onset of 21 years in a European cohort (Bollée *et al.*, 2011). Hyperuricaemia (corrected for glomerular filtration rate) was reported in 71% of the same cohort. This differentiates the *UMOD* subtype from ADTKD-*MUC1* where gout is much less prevalent (Kemter *et al.*, 2014; Olinger *et al.*, 2020; Gong *et al.*, 2021). Kidney ultrasound reveals cysts in one third of patients and these are almost always cortical (Bollée *et al.*, 2011).

Hyperuricaemia and gout are well-recognised features of ADTKD-*UMOD*, but their mechanistic basis remains incompletely understood. Unlike most mammals, humans and other great apes lack the enzyme uricase, leading to inherently higher serum urate levels and

a predisposition to urate-related disorders (Chang, 2014). In the general population, serum uric acid levels are tightly regulated by renal and intestinal transport pathways, and heritability estimates range between 38% and 63% (Yang *et al.*, 2005; Stiburkova and Ichida, 2025). In ADTKD-*UMOD*, it is hypothesised that TAL dysfunction and associated volume contraction drive upregulation of proximal tubular urate reabsorption, but this has not been formally tested. While hyperuricaemia typically precedes CKD onset in this condition, its role as a biomarker or modulator of disease progression remains unclear. The unique features of urate handling in ADTKD-*UMOD* suggests the need for mechanistic studies that can disentangle urate's causal role from its correlation with tubular dysfunction and fibrosis.

Cohorts have defined a median age of onset of ESKD as 47 (range 18-87) however, as for other forms of ADTKD, there is significant inter and intra-familial variability of disease progression despite the same causative *UMOD* variant (Kidd *et al.*, 2020). There is some evidence of genotype-phenotype correlation in ADTKD-*UMOD,* mainly in relation to variant type and associated protein domain affected altering age of onset of ESKD (Kemter *et al.*, 2014; Olinger *et al.*, 2020; Gong *et al.*, 2021). Surprisingly, variants which involve cysteine residues often seem to be associated with better outcomes (Olinger *et al.*, 2020; Gong *et al.*, 2021). Male sex is a significant predictor of worse kidney outcomes. However, multivariate analyses have shown that the best predictors of kidney survival in the disease are sex and mean age of ESKD onset within the family (Kidd *et al.*, 2020).

Although ADTKD-*UMOD* is typically inherited in an autosomal dominant pattern with heterozygous variants, rare homozygous cases have been described. These offer unique insights into gene dosage effects. One such case reported a homozygous *UMOD* p.C120Y variant in a consanguineous family, resulting in markedly reduced urinary uromodulin excretion and an altered fragment profile, yet without the severe phenotype that might be expected (Rezende-Lima *et al.*, 2004; Edwards *et al.*, 2017). Compared with heterozygous family members, the homozygous proband had earlier-onset gout and a more rapid decline in renal function, consistent with dose-dependent toxic gain-of-function effects. Nonetheless, she did not exhibit overt features associated with uromodulin deficiency, such as recurrent UTIs, kidney stones, or hypotension. These findings suggest that even very low levels of aberrant, glycosylated uromodulin may retain partial physiological function. The existence of both pathogenic and hypomorphic *UMOD* variants further supports a spectrum of functional impact. Hypomorphic alleles, variants with reduced but not abolished function such as *UMOD*

p.Cys255Tyr and *UMOD* p.Thr62Pro, can lead to attenuated ER stress and lower levels of intracellular retention, possibly explaining their milder clinical trajectory. Taken together, these observations support the concept that the clinical heterogeneity in ADTKD-*UMOD* reflects both the nature of the variant and gene dosage effects, with heterozygous and homozygous variants occupying different positions along a phenotypic continuum (Rezende-Lima *et al.*, 2004; Edwards *et al.*, 2017; Olinger *et al.*, 2022).

Aligned with guidance for all forms of ADTKD acknowledged above, definitive diagnosis of ADTKD-*UMOD* is via genetic testing. This can be via Sanger sequencing or NGS-based approaches such as gene panels or whole exome sequencing. A genetic diagnosis is also useful as is usually a prerequisite for future clinical trial inclusion and to allow prompt variant-based assessment of family members for diagnosis, pre-conception counselling and eligibility for live-related kidney donation (Olinger *et al.*, 2020).

Where biopsy is performed, it can be informative for subtype characterisation where uromodulin immunostaining can be performed. 85% of ADTKD-*UMOD* patients will have visible accumulation of uromodulin aggregates under periodic acid-Schiff staining under light microscopy. This is not usually widely available, however, and requires both adequate controls and an experienced operator (Onoe *et al.*, 2021).

The future may see the measurement of urinary uromodulin levels integrated into diagnostic algorithms as a non-invasive tool to differentiate ADTKD-*UMOD* from other ADTKD subtypes (Olinger *et al.*, 2020).

As for other ADTKD subtypes, no specific therapy is available apart from kidney transplantation, and general CKD recommendations apply (Eckardt *et al.*, 2015; Cormican *et al.*, 2020) . If RAAS blockade is necessary, losartan is the best choice due to its beneficial uricosuric effects when gout is present. Urate lowering therapies such as allopurinol or febuxostat and a low purine diet help to prevent gout attacks but do not influence CKD progression (Eckardt *et al.*, 2015). As salt wasting can be an early feature of the disease, diuretics should be used with caution (Labriola *et al.*, 2015).

Attempts have been made to develop therapies in ADTKD-*UMOD*. Due to its beneficial role in inflammatory pathways, *in vivo* tumour necrosis factor alpha blockade has slowed disease progression in a mouse model of ADTKD-*UMOD* (Johnson *et al.*, 2017). The chemical

chaperone sodium-4-phenylbutyrate showed promising *in vitro* effects but these have not been reproduced successfully *in vivo*. Finally, testing of the small molecule BRD4780 which has shown successful outcomes on ADTKD-*MUC1* is also being tested in ADTKD-*UMOD* (Kemter *et al.*, 2014). *In vitro* data has shown effects on mutant uromodulin accumulation which is promising news but preclinical data is awaited (Dvela-Levitt *et al.*, 2019).

The genetic and environmental modifiers of ADTKD-*UMOD* progression are unknown due to heterogeneous and scattered patient cohorts. Genetic modifier alleles determining uromodulin expression in the TAL or modulating pathological pathways operating in ADTKD-*UMOD* have not been studied. Early gout with reduced fractional excretion of uric acid is a hallmark of ADTKD-*UMOD* (Olinger *et al.*, 2020). However, the role of hyperuricemia as a biomarker for disease severity and/or as a factor modulating decline in kidney function has not been systematically investigated. A retrospective cohort study found that age of gout onset correlated with disease progression (Kidd *et al.*, 2020) but the underlying mechanisms driving hyperuricemia and its potential value as prognostic marker are unclear. Mouse models with different *UMOD* variants recapitulate the hallmarks of the disease including endoplasmic reticulum (ER) retention of mutant uromodulin and ER stress (Piret *et al.*, 2017). The molecular mechanisms leading from ER accumulation to interstitial fibrosis are however ill-defined. Mechanistic studies are hampered by the lack of appropriate cell systems reflecting the *in vivo* situation. Faithful *in vitro* disease modelling should integrate biallelic *UMOD* expression and the effects of genetic modifiers and thus ideally calls for a human cell system with endogenous *UMOD* expression and patient-specific genetic background. ADTKD is thus paradigmatic for difficulties typically encountered in rare diseases: genotype-phenotype correlations and disease modifiers are unknown, relevant human cell systems are missing, and limited pathophysiological insights translate into a lack of specific therapies.

1.5 The UK Experience of ADTKD-UMOD

The UK has pioneered in rare and genetic disease research due to the establishment of population biobanks, nationalised patient registries and integrated whole genome sequencing available within the National Health Service (NHS) (Turro *et al.*, 2020). Examples of these include The National Registry of Rare Kidney Disease (RaDaR) which is a national web-based registry containing longitudinal, prospective datasets of biochemical, imaging and pathological data with commonly accepted disease ontology available to researchers for data collection, secondary audit, study recruitment and patient engagement in research ('National

Registry of Rare Kidney Diseases (RaDaR)', 2022). UK Biobank (a large-scale biomedical population database and research resource containing thorough health and genetic information from 500 000 UK participants) ('UK Biobank', 2022) and Genomics England 100 000 Genomes project (over 35 000 individuals with rare diseases have had their whole genome sequenced to date and results are directly reported back to ensure prompt clinical influence) are other examples (Turro *et al.*, 2020).

At present, 179 UK families with ADTKD have been recruited to RaDaR and the 100 000 Genomes Project of which 61 families have received a genetic diagnosis of ADTKD-*UMOD* (Mabillard, Sayer and Olinger, 2023). These registries have allowed us to identify a unique genetic landscape of ADTKD-*UMOD* in the UK as almost half of families carry a specific in-frame 'indel' p.Val93_Gly97delinsAlaAlaSerCys variant in exon 3 which has been shown to result in a milder clinical phenotype (slower kidney function decline and a very low prevalence of gout) (Smith *et al.*, 2011; Olinger *et al.*, 2020). A previous study of four large families with this variant suggested that this 'indel' was a recurrent change rather than a founder effect (Smith *et al.*, 2011). However, a recent, larger and more geographically dispersed cohort study has shown the 'indel' to be a pathogenic ancestral variant highly prevalent in the British population due to a shared common extended haplotype amongst individuals with the variant (Valluru *et al.*, 2023).

A less genetically heterogeneous ADTKD-*UMOD* population is the ideal situation to study the influence of clinical, environmental and genetic modifiers of disease progression.

1.6 Broader insights from the study of ADTKD-UMOD

The study of rare kidney diseases provides insights into more common disorders and the *UMOD* gene is a well-documented example of this. Common regulatory Single Nucleotide Polymorphisms (SNPs) in the *UMOD* promoter have shown it to be the most strongly associated locus with the risk of CKD and hypertension in genome-wide association studies (Rampoldi *et al.*, 2011; Schaeffer, Devuyst and Rampoldi, 2021).

Since it has been discovered that a fraction of uromodulin is released through the basolateral membrane of tubular cells into the interstitium and, subsequently, the circulation, much interest has been taken into its systemic role and potential for clinical utility (El-Achkar *et al.*, 2013). Urinary and serum concentrations of uromodulin have emerged as novel surrogates for

renal tubular function, integrity and nephron mass (Dawnay and Cattell, 1981; Thornley, Dawnay and Cattell, 1985; Risch *et al.*, 2014; Pruijm *et al.*, 2016; Steubl *et al.*, 2016; Garimella and Sarnak, 2017; Scherberich *et al.*, 2018). Reduced urine and serum uromodulin have been associated with deterioration of kidney function (Leiherer *et al.*, 2018), as well as sensitive markers of kidney allograft dysfunction prior to deterioration of serum creatinine or development of albuminuria/proteinuria (Avis *et al.*, 1985; McLaughlin *et al.*, 1993). A lower baseline serum uromodulin concentration has also been associated with kidney allograft loss at follow up (Steubl *et al.*, 2017; Bostom *et al.*, 2018). The increase in serum uromodulin after acute kidney injury has also been suggested as a prognostic biomarker for kidney recovery. Serum uromodulin is also an independent predictor of mortality and risk of cardiovascular events however much more is to be elucidated as to the systemic role of this fascinating glycoprotein (Ponte and Devuyst, 2020).

Intense research efforts globally have been targeted at mechanisms driving interstitial fibrosis and at rescuing proteostasis defects. As a monogenic disease model of kidney interstitial fibrosis, a better understanding of the clinical and genetic architecture associated with ADTKD, and pathways involved in the variability of disease progression should improve our understanding of more common target mechanisms involved in fibrotic kidney disease which is the endpoint of all CKD. Furthermore, the study of ADTKD has caused us to question the role of the distal tubule, ER, proteostasis and metabolic processes involved in renal fibrosis development (Mabillard, Sayer and Olinger, 2023). If we can uncover new actionable targets in ADTKD, these might be relevant for more common diseases in which fibrosis is the main driver of progression.

1.7 Phenotypic Heterogeneity in Kidney Disease

Phenotypic heterogeneity in rare Mendelian diseases is a serious concern due to consequential misdiagnosis and/or delayed diagnosis (Maroilley and Tarailo-Graovac, 2019). Next generation sequencing techniques that survey the whole genome or exome give us the power to take a hypothesis-free approach to genetic testing and has enabled us to realise the vast variation and deviation of phenotypes from text-book descriptors of disease (Stokman *et al.*, 2016; Ars and Torra, 2017). Possible explanations for such phenotypic variability in kidney disease include additional clinical factors such as diabetes, obesity and hypertension. Environmental factors can influence disease from embryonic development to old age and

include diet, drugs, climate, illness and stress with many driving the aforementioned clinical factors (Nicolaou *et al.*, 2015; Ars and Torra, 2017).

Genetic heterogeneity influences many kidney diseases including ADTKD (Jones *et al.*, 1999; Eckardt *et al.*, 2015). For example, Polycystic Kidney Disease patients with a *PKD1* variant reach ESKD around 20 years earlier than those with a *PKD2* variant (Hateboer *et al.*, 1999). Furthermore, different variants in a particular gene can result in different phenotypes (allelic heterogeneity) and genotype-phenotype correlations are typically weaker for autosomal dominant disorders than for autosomal recessive ones (Devuyst and Thakker, 2010; Gunay-Aygun *et al.*, 2010; Hopp *et al.*, 2012; Cornec-Le Gall *et al.*, 2013). Other reasons for such variability in phenotype include epigenetic regulation (dynamic alterations such as DNA methylation and histone modifications can alter gene expression), mosaicism (two or more populations of cells with different genotypes exist in the affected individual or organ), X inactivation (if the wild-type allele is inactivated in the majority of cells, the disease would be more severe than expected in a female) and splicing variants (when the proportion of aberrant splicing is higher, disease is typically more severe) (Ars and Torra, 2017).

Most inherited kidney diseases are considered to be fully penetrant but incomplete penetrance can be seen in some diseases (Torra *et al.*, 2004), for example, some patients with *MUC1* variants have had normal kidney function in their 70s (Eckardt *et al.*, 2015; Ars and Torra, 2017). This phenomenon is thought to be due to either the causative gene itself or the impact of modifier genes.

Oligogenic inheritance (where more than one gene exerts an effect of comparable magnitude on a phenotype) has been suggested in some inherited kidney diseases such as atypical HUS (Lemaire *et al.*, 2013; Sánchez Chinchilla *et al.*, 2014; Mele *et al.*, 2015), Alport Syndrome (Mencarelli *et al.*, 2015), Bardet-Biedl Syndrome (Katsanis, 2004) and Nephronophthisis (Hoefele *et al.*, 2007). Similarly, the impact of modifier genes (either in the causative gene or in other genes impacting common disease pathways) have been shown to contribute to the phenotype in many diseases. Modifier genes are often suspected when intra-familial variability is seen in adult-onset disease and these have been detected in ADPKD and Alport Syndrome where, in some cases, age of onset of ESKD can differ by over 30 years (Persu *et al.*, 2004; Vujic *et al.*, 2010; Bullich *et al.*, 2015).

Genetic association studies have been fundamental at identifying genetic causes of phenotypic variability, even in rare diseases (Drumm *et al.*, 2005; Gu *et al.*, 2009; Rahit and Tarailo-Graovac, 2020; Parodi *et al.*, 2022). Identification of genetic modifiers have uncovered entirely new and unsuspected patho-mechanisms in various diseases revealing new opportunities for therapeutic targets.

1.8 Genetic Association Studies

Genetic association studies have revolutionised the field of disease genetics since the first genome-wide association study (GWAS) was performed for age-related macular degeneration in 2005 (Klein *et al.*, 2005; Visscher *et al.*, 2012, 2017). GWAS, in particular, have enabled us to identify thousands of disease susceptibility genes and corresponding underlying biological pathways (Tam *et al.*, 2019). This has led to clinical advances such as identification both of biomarkers (Owen *et al.*, 2010) and drug targets in addition to disease risk prediction and personalised medicine approaches where treatments are optimised according to genotype (Nelson *et al.*, 2015; Visscher *et al.*, 2017). Other benefits of GWAS include insight into ethnic variation of complex traits (Liu *et al.*, 2015), ability to identify low-frequency and rare variants (Peloso *et al.*, 2014; Fuchsberger *et al.*, 2016; Turcot *et al.*, 2018) in addition to other classes of genetic variation associated with disease risk such as structural variants (Pinto *et al.*, 2010; Malhotra *et al.*, 2011), mosaic events (Jacobs *et al.*, 2012; Bonnefond *et al.*, 2013) and even novel monogenic and oligogenic diseases (Hirschhorn, 2009). Beyond disease identification, GWAS can be used for many applications from ancestry determination (Jakkula *et al.*, 2008) and paternity testing (Kerr *et al.*, 2013) to forensic analysis (Homer *et al.*, 2008). Following important technological and methodological advances, GWAS are relatively straightforward to perform and are highly reproducible when a standardised stringent significance threshold ($P < 5 \times 10^{-8}$) is adopted (Pe'er *et al.*, 2008; Tam *et al.*, 2019). GWAS are paradigmatic for the benefits of collaborative efforts as the increasing availability of publicly available GWAS data with increased data sharing is enabling further novel associations.

GWAS are not without limitations and criticism. Multiple testing burden (Pulit, de With and de Bakker, 2017), missing heritability (Eichler *et al.*, 2010), epistasis (Wei, Hemani and Haley, 2014), spurious associations (McClellan and King, 2010) and cryptic population stratification (McClellan and King, 2010) are factors known to compromise or be compromised by methodological rigor. It is unlikely that GWAS will ever explain 100% of the heritability of complex traits due to gene-gene and gene-environment interactions and the challenges

associated with detecting rare variants or variants with small effect sizes (Altshuler, Daly and Lander, 2008). Identification of a locus does not necessarily mean that causal variants can be discerned due to linkage disequilibrium encompassing multiple variants, association localised to variants in non-coding regions, discovery of multiple hits or, in contrast, no associations reaching the required significance thresholds (Altshuler, Daly and Lander, 2008; Tam *et al.*, 2019). However, many advances have been made to overcome these obstacles such as dense and custom SNP arrays, imputation, meta-analyses from different ethnic backgrounds, advances in statistical fine-mapping and functional genomics and regulatory element databases (Boyle *et al.*, 2012; Andersson *et al.*, 2014; Roadmap Epigenomics Consortium *et al.*, 2015; Mägi *et al.*, 2017; Schaid, Chen and Larson, 2018; Tam *et al.*, 2019).

Many of the limitations of GWAS can now be overcome to some extent. With careful and rigorous methodological design, many GWAS in recent years have successfully uncovered genetic associations and significant underlying biology with transformative clinical implications (Visscher *et al.*, 2017). The growing number of published studies is testament to the ongoing success of this approach in medical research.

1.9 Polygenic Risk Scores

While GWAS have identified common genetic variants that contribute to both complex traits and diseases, their combined effects can be captured through Polygenic Risk Scores (PRS) which can estimate an individual's genetic predisposition (Hingorani *et al.*, 2023). More recently, research has demonstrated that PRS can modify the expressivity and penetrance of monogenic diseases, offering new insights into variable disease outcomes among carriers of pathogenic variants (Khan *et al.*, 2023).

More broadly, much hype has been placed on PRS, namely in refining disease risk prediction and enhancing existing clinical models for disease stratification and screening. This includes speculation that learning about disease risk could encourage an individuals' behaviour change towards improved lifestyle adoption. However, studies suggest that this knowledge does not strongly influence behaviour and, indeed, that other expressed benefits of PRS are both premature and overstated (Hingorani *et al.*, 2023; Sud *et al.*, 2023).

A study which analysed data from the Polygenic Score Catalog found that PRS generally perform poorly across different applications (population screening, individual risk prediction and risk stratification) due to low detection rates and high false positives, limiting their clinical

utility. There were even only modest differences between high and low quintiles in disease stratification applications aiming to identify high and low risk-subgroups. Conclusions from this work suggest that PRS utility should be carefully evaluated against existing screening and prevention strategies (Hingorani *et al.*, 2023; Sud *et al.*, 2023). However, PRS may still have niche applications such as modifying risk predictions in monogenic disease or identifying potential drug targets, but they are unlikely to revolutionise disease prevention or screening as their performance, measured by discrimination and calibration, often falls short of levels necessary to guide these strategies (Hingorani *et al.*, 2023; Sud *et al.*, 2023).

Considering monogenic disease, where rare, high-impact variants cause strong disruption to specific biological pathways, PRS should and have started to be explored particularly in the realm of explaining incomplete penetrance (Khan *et al.*, 2023). Currently understood reasons for incomplete penetrance are environmental factors, stochastic (random) factors and inherited genetic modifiers (Kingdom and Wright, 2022). Genome-wide PRS can explain the cumulative effect of many small-effect variants across the genome and, indeed, have been shown to explain variable penetrance in other monogenic disorders including Familial Hypercholesterolaemia (FH), Hereditary breast and ovarian cancer syndromes due to *BRCA1/2* variants and in Lynch Syndrome, a hereditary colorectal cancer (Fahed *et al.*, 2020).

PRS may modify disease risk in individuals with monogenic variants, making these individuals more or less likely to develop the disease despite carrying a high-risk genetic variant. Khan *et al.* (Khan *et al.*, 2023). applied a Polygenic Risk Score for CKD to Autosomal Dominant Polycystic Kidney Disease (ADPKD) and Collagen Type IV Alpha-Associated Nephropathy (COL4-AN) variant carriers and found a 54-fold increased CKD risk versus a 3-fold increased risk in the highest versus lowest PRS tertile in the ADPKD-variant carriers. Similarly, the authors found a 2.5-fold versus no significant CKD risk elevation in the highest versus lowest tertile respectively in COL4-AN-variant carriers suggesting that polygenic background influences monogenic disease penetrance (Khan *et al.*, 2023). This interaction between monogenic and polygenic risk has not been well studied in kidney diseases beyond this and could be an important area for future research, particularly in ADTKD-*UMOD* where the *UMOD* gene displays a whole spectrum of rare Mendelian disease-causing variants to common variants affecting kidney function, blood pressure and uric acid levels across the general population (Olinger *et al.*, 2022).

Furthermore, for rare diseases, where clinical trials are so desperately needed, PRS could also serve to improve patient stratification to enrich study populations and allow smaller sample sizes to increase statistical power, aid the efficacy of trial treatment tailoring and allow for an improved personalised approach to therapy. This could add momentum to testing of treatments in diseases like ADTKD-*UMOD* and increase the speed at which patients can gain access to therapies with better-informed risk-benefit decision making. Finally, it must be emphasised that PRS, like GWAS, should be tested across ancestries to ensure application is both equitable and as unbiased as possible (Khan *et al.*, 2022).

1.10 Biomarkers of Disease Progression

Biomarkers are critical for the diagnosis, prognosis and therapeutic development for rare diseases like ADTKD-*UMOD*. Lack of disease-specific treatments, vast disease heterogeneity, small and dispersed patient cohorts and limitations of conventional clinical biomarkers such as serum creatinine, eGFR and invasive kidney biopsy reinforce how novel biomarkers are desperately needed (Canki, Kho and Hoenderop, 2024). ADTKD-*UMOD* is paradigmatic for the challenges of rare disease research (Tang, Yuan and Chen, 2022).

The lack of symptoms and distinguishable clinical features such as proteinuria or early hypertension can make early diagnosis very challenging. Furthermore, traditional measure to detect disease progression such as eGFR decline often occur years beyond the onset of kidney damage, further delaying opportunities for clinical interventions. Genetic biomarkers such as primary *UMOD* variant, provide a definitive diagnosis, but do not accurately dictate or predict severity of disease progression. Urinary biomarkers such as uromodulin, B2-microglobulin, kidney injury molecule-1 (KIM-1) and neutrophil gelatinase-associated lipocalin (NGAL) have shown potential at detecting early tubular injury but have not been studied in large disease-specific cohorts or as a biomarker of disease progression specifically (Ix and Shlipak, 2021; Canki, Kho and Hoenderop, 2024).

The heterogeneity of disease progression in ADTKD-*UMOD* necessitates reliable prognostic biomarkers to identify and stratify those at higher risk of rapid kidney function decline. Serum and urinary biomarkers reflective of cellular uromodulin secretion, tubular stress and interstitial fibrosis could allow disease severity stratification, better informing clinical decision making (Ix and Shlipak, 2021). Predictive biomarkers can also facilitate precision medicine

approaches which can allow for more appropriate treatment tailoring, risk-benefit clinical decision making and avoidance of unnecessary therapeutic complications (R. G. Fassett, 2011).

Biomarkers are essential to drug development in this disease, however lack of non-invasive and real-time biomarkers of disease activity have hampered interest from the pharmaceutical industry. Serum uric acid has been explored as a potential biomarker of disease progression in CKD and ADTKD-*UMOD*, particularly in the context of trials involving urate-lowering therapies (Sampson, Singer and Walters, 2017; Olinger *et al.*, 2020). Surrogate biomarkers that reflect tubular function and progression of fibrosis could accelerate clinical trials by providing early indicators of therapeutic efficacy and reduce reliance on long-term outcomes and high-risk invasive kidney biopsies (Ix and Shlipak, 2021). Biomarker-driven trials can aid patient stratification, ensuring inclusion of individuals most likely to benefit from novel therapies and potentially increase statistical power (R. G. Fassett, 2011). More precise methods for monitoring drug response could also expedite regulatory approval of therapeutics.

Development and validation of biomarkers in ADTKD-*UMOD* needs to overcome the limitations of existing clinical and biomarker-driven strategies and focus on advancing early diagnosis, predicting rate of disease progression and facilitate therapeutic decision making (Fassett, 2011). In the clinical setting, this integrated approach of combining genetic, biochemical and molecular biomarkers could optimise diagnosis, prognosis, treatment strategies, shared-decision making and outcomes for individuals living with ADTKD-*UMOD*.

1.11 Gaps in Current Knowledge and Unmet Needs

1) The variability in ADTKD-*UMOD* disease progression has not been studied in homogeneous longitudinal patient populations. The contributions of environmental, clinical, and genetic factors to this variability are unknown.
2) Biomarkers of disease progression and precision diagnostic tools are lacking.
3) The molecular mechanisms underlying hyperuricemia and the pathophysiological pathways driving kidney disease progression are unknown.

1.12 Hypothesis

Genetic and clinical factors are responsible for disease progression and significant inter- and intra-familial phenotypic variability in ADTKD-*UMOD*. Applying unbiased approaches to

identify factors contributing to disease variability will indicate pathophysiological pathways involved in disease progression.

1.13 Summary of Aims and Objectives

The overarching aim of this work is to determine clinical and genetic factors that are responsible for disease progression in ADTKD-*UMOD*. I also intended to determine relationships between potential disease biomarkers such as uric acid, urine uromodulin and serum uromodulin and disease progression in patients with ADTKD-*UMOD*. These aims will be achieved through 3 strands of work:

1) **Deep longitudinal phenotyping of ADTKD-*UMOD* in a large cohort of UK patients**

ADTKD is paradigmatic for difficulties typically encountered in rare diseases: genotype-phenotype correlations and disease modifiers are unknown and limited pathophysiological insights translate into lack of specific therapies. I built a deep phenotypic dataset within the National Registry of Rare Kidney Disease (RaDaR) ADTKD cohort and expanded on this dataset with the help of international collaborations which I established to better determine the phenotypic spectrum of this disease. The unique genetic landscape in the UK with a high number of individuals with a recurrent 'indel' variant (*UMOD* p. Val93_gly97delinsAlaAlaSerCys) offered a unique opportunity to identify genotype-independent disease modifiers.

2) **Define clinical and genetic modifiers of kidney disease progression in ADTKD-*UMOD***

The genetic and environmental modifiers of ADTKD-*UMOD* progression had previously been unknown due to heterogeneous and scattered patient cohorts, and genetic modifier alleles determining uromodulin expression in the thick ascending limb (TAL) or modulating pathophysiological pathways operating in ADTKD-*UMOD* had not been studied. Using appropriate statistical models (Cox regression models), I determined whether factors such as primary *UMOD* variant, uric acid levels, presence of diabetes, hypertension and other factors are independently associated with kidney survival. I utilized compiled data from the UK ADTKD-*UMOD* Registry and collaborator-sought data to substantiate disease progression between individuals and performed an unbiased genome-wide association study to determine genetic factors that might be responsible for this variability.

3) **Define factors that could serve as biomarkers of kidney disease progression in ADTKD-*UMOD***

The severity of *in vitro* trafficking defects of different primary *UMOD* variants was shown to correlate with age of end stage kidney disease suggesting the use of urine and/or serum uromodulin as a prognostic marker in ADTKD-*UMOD* (Kidd *et al.*, 2020). Using established ELISA, I systematically assessed urine and serum uromodulin levels in patients and tested correlations between these levels and severity of kidney disease. Early gout with reduced fractional excretion of uric acid is a hallmark of ADTKD-*UMOD* (Olinger *et al.*, 2020). However, the role of hyperuricemia as a biomarker of disease severity and/or as a factor modulating decline in kidney function had not been systematically investigated. A retrospective cohort study found that age of gout onset correlated with disease progression (Kidd *et al.*, 2020) but the underlying mechanisms driving hyperuricemia and its potential value as a prognostic marker remained unclear. I assessed parameters relevant for uric acid handling, including fractional excretion of uric acid and dynamics of hyperuricemia and gout in ADTKD-*UMOD* patients. I also determined the relationship between serum uric acid levels and glomerular filtration rate in the general population using data from the UK Biobank to better define the intriguing defects in uric acid handling observed in ADTKD-*UMOD.*

1.14 Conclusion

Better characterising both the clinical and genetic architecture of ADTKD-*UMOD,* using human data, will enable new and better targeted approaches to diagnosis and treatment as new genotype-phenotype correlations are discovered and detailed phenotypic understanding is advanced. Disease biomarkers can not only aid early detection and hasten clinical trial approval but can facilitate a personalised approach to disease prognosis and intervention. Studying mechanisms that modify disease progression can provide valuable insights into disease pathophysiology and uncover important biological pathways and new actionable targets and might be relevant for more common diseases of the kidney in which fibrosis is the main driver of progression.

This thesis is structured to first establish a comprehensive clinical and genetic understanding of ADTKD-*UMOD* before systematically exploring its modifiers and potential biomarkers of disease progression. Chapter 2 characterises the clinical landscape of ADTKD-*UMOD* in the UK

through in-depth phenotyping of national registry data, supplemented by local and biobank cohorts. This chapter also addresses variant spectrum, genotype-phenotype correlations, and under-recognised clinical features, establishing the foundation for subsequent analyses. Chapter 3 builds on this by investigating genetic modifiers of disease progression through genome-wide association studies (GWAS), gene burden testing, and chromatin interaction mapping, offering novel insights into uric acid transporter impact, mitochondrial pathways, and the role of polygenic background. Chapter 4 narrows the focus to the intriguing relationship between uric acid handling and kidney survival in ADTKD-*UMOD*, leveraging both patient and population-level data to dissect this metabolic phenotype. Chapter 5 then evaluates serum and urinary uromodulin as mechanistically anchored biomarkers of disease severity and progression, providing a proof-of-concept for their potential integration into clinical practice. Finally, Chapter 6 synthesises findings across all strands of work, reframing ADTKD-*UMOD* as a complex monogenic disease shaped by specific modifying variants, common genetic background, and likely regulatory and mitochondrial influences.

This structure reflects a deliberate progression: from disease characterisation to modifier discovery, to biomarker validation. Each step grounded in the clinical realities of patients and informed by molecular data. By adopting this integrated, multi-layered approach, this thesis addresses key knowledge gaps in ADTKD-*UMOD*, while also contributing broadly to our understanding of fibrosis, biomarker development, and genetic architecture in chronic kidney disease.

In the next chapter, the focus shifts to defining the clinical and genetic architecture of ADTKD-*UMOD* within the UK. Using data from the National Registry of Rare Kidney Diseases (RaDaR), local cohort studies, and the UK Biobank, this chapter explores the distribution and spectrum of *UMOD* variants, including the high-frequency UK-specific *UMOD* p.Val93_Gly97delinsAlaAlaSerCys 'indel' variant. It examines genotype-phenotype correlations, intra- and inter-familial variability, and highlights under-recognised clinical features such as sex-specific differences, pregnancy complications, and diagnostic delay. This comprehensive characterisation provides the necessary foundation for the subsequent investigation of genetic modifiers and biomarkers, enabling deeper insight into the factors that drive disease heterogeneity and progression.

Chapter 2. The Clinical and Genetic Landscape of ADTKD-*UMOD* in the UK

2.1 Introduction

ADTKD-*UMOD* has been increasingly characterised through national and international cohorts, each contributing important insights into disease presentation, progression, heterogeneity and genetic landscape.

International Cohorts

To date, the largest multicentre cohort description includes 722 individuals from 249 families across 19 countries. 52.4% were genetically confirmed and clinical features in 347 included hyperuricaemia (67%), gout (35%) and end stage kidney disease (ESKD) (49.7%) (Olinger *et al.*, 2020). Median age of gout onset was 28 years (range 9-69) and median age of ESKD was 54 years (range 17-83). Fractional excretion of uric acid was <5% in 78% of tested individuals. Prior to this, a French and Spanish cohort had been characterised with similar characteristics. Bollee *et al* reported in their French cohort of 143 individuals across 33 families with *UMOD* variants that 94% of individuals had hyperuricaemia and 69% had gout, with fractional excretion of uric acid <5% present in 91% of cases (Bollée *et al.*, 2011). Median age of CKD onset was 25 years (range 3-60) with a median ESKD age of 56 (range 36-77). Similarly, in the Spanish cohort, Ayasreh *et al,* reported on 9 families with ADTKD-*UMOD* (Ayasreh *et al.*, 2018). Of these 58% individuals had gout, median age of CKD onset was 26 (range 12-51) and the median age of ESKD was 54 years (range 36-70). Fractional excretion of uric acid was <5% in 86% of tested individuals. In an Irish cohort, the gout prevalence in those with ADTKD-*UMOD* was 68% and mean age of ESKD was 46.9 years. Beyond this Eurocentric landscape, Chinese and Japanese cohorts have now been described. Gong *et al* included 11 probands with ADTKD-*UMOD* in a Chinese cohort (Gong *et al.*, 2021). All affected individuals had hyperuricaemia, over 70% developed gout and 50% of patients had progressed to ESKD. In the Japanese cohort, Tanaka *et al* describe a median diagnosis age of 30.5 years, hyperuricaemia presence in 94% and a median age of end stage kidney failure of 56 years (Tanaka *et al.*, 2025).

Genotype-Phenotype Correlations

More than 135 *UMOD* variants have been identified to date (Olinger *et al.*, 2020; Schaeffer, Devuyst and Rampoldi, 2021). Most are missense variants involving conserved cysteine residues, critical for disulphide bond formation and proper protein folding. About 95% of

variants map to exons 3 and 4, corresponding to the N-terminal portion of the protein and code for the four EGF-like domains and the D8C domain. Approximately, 60% of variants delete, insert or replace a cysteine residue, crucial for complex tertiary protein folding. Interestingly, there is a distinct geographical clustering of two specific variants, *UMOD* p.Val93_Gly97delinsAlaAlaSerCys, frequent to the UK and *UMOD* p.H177_R185del in the USA. Disease penetrance is understood to be ~100% however there are a few cases of unaffected heterozygotes reported (Smith *et al.*, 2011; Kidd *et al.*, 2020; Olinger *et al.*, 2020; Gong *et al.*, 2021).

Kidd *et al* demonstrated that the relationship between *UMOD* variant and clinical severity are often modest and variable (Kidd *et al.*, 2020). The authors of this large study found that the specific class of variant was not independently predictive of ESKD however the *in vitro* trafficking defect (ratio of ER-retained versus mature uromodulin) of the variant did significantly correlate with kidney survival. Individuals with more severe trafficking defects had an earlier onset of ESKD (mean age 47 for highest severity versus 59 years for the lowest). Further to this, Gong *et al* reported that variants which substituted a cysteine (affecting disulfide bonding and folding) were associated with a better kidney survival in comparison to non-cysteine variants, although domain location was not significantly predictive (Gong *et al.*, 2021). Bollée *et al* found only modest genotype-phenotype correlations and emphasised high intrafamilial variability (Bollée *et al.*, 2011). Similarly, Moskowitz *et al* reported earlier ESKD onset in patients with variants in the EGF-like domains (median 45–52 years) compared to those in the cysteine rich domain (median 60–65 years), although this effect was modest and may have been confounded by family size and data quality (Moskowitz *et al.*, 2013).

Returning to the in frame *UMOD* p.Val93_Gly97delinsAlaAlaSerCys insertion-deletion variant; this change was associated with a milder *in vitro* phenotype, virtual absence of gout and slower decline in kidney function when first described but a second study showed no statistical difference in median age (52 years; range 32-76) of kidney failure to broader population studies. 90% of these individuals had hyperuricaemia, 63% had gout and all had a fractional excretion of uric acid <5% that were tested (Smith *et al.*, 2011; Valluru *et al.*, 2023). The aforementioned *in vitro* phenotype refers to experimental observations made in cultured cells transfected with expression constructs encoding the mutant *UMOD* protein. In this case, cells expressing the p.Val93_Gly97delinsAlaAlaSerCys variant showed relatively mild impairment in

uromodulin trafficking, measured as a higher ratio of mature, glycosylated uromodulin successfully exiting the endoplasmic reticulum compared to the amount retained intracellularly. This suggests reduced ER stress and better cellular processing of the mutant protein relative to more deleterious *UMOD* variants. Such *in vitro* trafficking scores have been shown to correlate with age of ESKD onset in ADTKD-*UMOD* and are increasingly used to stratify variants along a continuum of pathogenic severity.

Olinger *et al* characterised an intermediate-effect *UMOD* variant, p.Thr62Pro, which confers a significantly increased risk of CKD and ESKD, but with reduced severity compared to classical ADTKD-*UMOD* variants (Olinger *et al.*, 2022). Found in approximately 1 in 1000 individuals of European descent, this variant demonstrates incomplete penetrance, an intermediate protein *in vitro* trafficking defect, modest ER stress induction and a 20-year delay in ESKD onset relative to other ADTKD-*UMOD* variants. Despite this, the variant can segregate with disease in ADTKD-like families consistent with a Mendelian inheritance pattern in some cases. The variant is enriched in familial CKD clusters in addition to population-based kidney failure cohorts shedding more light on the genotype-phenotype spectrum of *UMOD*-associated kidney diseases as well as the genetic architecture of CKD.

Overall, whilst certain variant features such as uromodulin trafficking defect severity, cysteine involvement or variant position can influence phenotype, recent findings suggest that some variants such as *UMOD* p.Thr62Pro exhibit intermediate effects with partial penetrance and delayed progression. Nonetheless, genotype-phenotype correlations in the Mendelian context remain modest with considerable variability both within and between families (Kidd *et al.*, 2020; Olinger *et al.*, 2022).

Modifiers of Disease Progression

Kidd *et al* demonstrated that male sex is significantly associated with an earlier onset of end-stage kidney disease (ESKD) (Kidd *et al.*, 2020). While the presence of gout was not independently predictive of renal outcomes, an earlier age of gout onset was significantly associated with more rapid progression to kidney failure. Parental age at ESKD, particularly maternal age, also predicted the age of ESKD in the offspring, with a stronger effect observed in daughters. In the same study, the minor allele of the *UMOD* promoter variant rs4293393, associated with lower uromodulin expression when *in cis* with the disease-causing variant

suggesting a potential protective effect, however, its uneven distribution precluded a formal Mendelian randomisation analysis. Collectively, these findings indicate that sex, age of gout onset, parental disease course and potentially *UMOD* promoter polymorphisms may all contribute to phenotypic heterogeneity in ADTKD-*UMOD.*

Clinical Sub-Phenotypes

The ADTKD-*UMOD* literature consistently describes a disease phenotype centred around progressive tubulointerstitial fibrosis, progressive kidney function decline and bland urinary sediment. Several characteristic clinical and pathological features beyond CKD, gout, and hyperuricaemia are highlighted across cohort studies and mechanistic reports (Eckardt *et al.*, 2015; Devuyst *et al.*, 2019).

Renal cysts are variably reported in ADTKD-*UMOD* and are not considered a defining feature. In the French cohort described by Bollée *et al,* renal cysts were identified in approximately 25% of patients, though their presence was not specific or pathognomonic (Bollée *et al.*, 2011). Olinger *et al* noted that imaging typically reveals kidneys of normal or slightly reduced size, with cysts being uncommon or mild (Olinger *et al.*, 2020). Gong *et al* did not find a significant difference in cyst prevalence between genetically confirmed ADTKD-*UMOD* patients and those with suspected disease, reinforcing the notion that cysts are neither sensitive nor specific markers of the disease (Gong *et al.*, 2021).

A core pathophysiological feature is dysfunction of the thick ascending limb (TAL), where uromodulin is exclusively produced. Mutant uromodulin disrupts trafficking and expression of NKCC2 and ROMK, key channels for sodium and potassium reabsorption in the TAL. This results in impaired sodium reabsorption and a mild salt-wasting phenotype. This mechanism is supported by functional studies in animal models; however, this phenotype is yet to be clearly described in human disease cohorts (Bernascone *et al.*, 2010; Trudu *et al.*, 2013).

Overall, these additional features, particularly TAL dysfunction, and subtle imaging findings, reinforce ADTKD-*UMOD* as a distinct tubulointerstitial disease with early tubular injury and subtle or non-specific imaging clues which should be characterised further.

2.2 Methods

2.2.1 RaDaR ADTKD Cohort Acquisition and Analysis

A data requisition application was made to the UK RaDaR registry ('National Registry of Rare Kidney Diseases (RaDaR)', 2022). Following approval by the RaDaR Data Access Committee, anonymised patient data relevant to the study objective was securely transferred in accordance with data governance protocols. All data handling complied with relevant ethical and data protection guidelines. Data was only used for those in the cohort who had a genetically confirmed diagnosis of ADTKD-*UMOD* and where their primary *UMOD* variant was available.

2.2.2 Local ADTKD Cohort Acquisition and Analysis

A locally acquired cohort of 70 patients with genetically confirmed ADTKD-*UMOD* was amalgamated. The Electronic Patient Record (EPR) was interrogated for more detailed information on phenotype in this cohort. All data was anonymised, and data handling complied with relevant ethical and data protection guidelines.

2.2.3 Determination of Non-Canonical UMOD Variant Prevalence in UK Biobank

This was performed by a colleague, Dr Robert Geraghty. UK Biobank is a large population-based cohort consisting of clinical, biochemical, imaging and genetic data (Allen *et al.*, 2024). Using PLINK (Chang *et al.*, 2015), we filtered the interim release (450k) whole exome sequence chromosome 16 bed file to include likely pathogenic and pathogenic variants within exons 3 and 4 of *UMOD*. Pathogenicity was derived from entries in dbSNP (Sherry *et al.*, 2001). Code for phenotype construction and genotype extraction is detailed at: https://github.com/rg2u17/umod.

2.2.4 Genetic Spectrum Analysis

Primary *UMOD* variant data was extracted from RaDaR ADTKD data to determine the spectrum of *UMOD* variants. These were then plotted in ProteinPaint (Zhou *et al.*, 2016). The proportions of patients with each variant were determined and the ACMG Classification for each variant, where determined, were checked using ClinVar (Richards *et al.*, 2015; Landrum *et al.*, 2016).

2.2.5 Genotype-Phenotype Correlations

Clinical and Genetic data, where available, from RaDaR and the local ADTKD-*UMOD* cohort (named the 'UK Cohort') were matched. Data from the UK cohort was then compared to previously published data from the Wake Forrest and International Cohorts. Appropriate statistical tests (Chi-square and Fisher's exact) were performed for cohort comparison. Where scatter plots were performed, a Mann-Whitney test was applied and where kidney survival was analysed, a Cox Proportional Hazards model was applied. Salt-wasting features were based on commonly used clinical and published criteria and diagnostic features. Kidney imaging and kidney biopsy data was taken from existing local reports written by radiologists or histopathologists.

2.2.6 Determination of Clinical Modifiers of Disease Progression in ADTKD-UMOD

A locally acquired cohort of 70 patients with genetically confirmed ADTKD-*UMOD* was analysed as this provided richness of clinical data taken from the Electronic Patient Record (EPR). The outcome variable was defined as kidney survival and a Cox Proportional Hazards model was applied.

2.3 Results

To guide the reader through the structured presentation of results, this chapter is organised to reflect a logical progression from national cohort ascertainment to increasingly granular phenotype and genotype characterisation. **Section 2.3.1** begins by outlining the national landscape of ADTKD family recruitment across the UK, establishing the foundation for subsequent analyses. **Section 2.3.2** then presents clear evidence of phenotypic heterogeneity beyond the primary causative *UMOD* variant, illustrated through both inter- and intra-familial comparisons. **Section 2.3.3** details the full genetic spectrum of *UMOD* variants identified in the UK, including classification and novel observations. **Section 2.3.4** explores genotype-phenotype correlations, providing extensive survival, imaging, and phenotypic data stratified by variant type, and includes detailed subgroup analyses such as salt-wasting and homozygote phenotypes. Finally, **section 2.3.5** investigates clinical modifiers that may influence disease progression in ADTKD-*UMOD*. Each section presents relevant figures and tables in sequence, which are intended to be interpreted descriptively in the first instance. Integrated interpretation, clinical implications, and comparisons with existing literature are reserved for the discussion **(section 2.4)** that follows.

2.3.1 The UK RaDaR landscape of ADTKD families

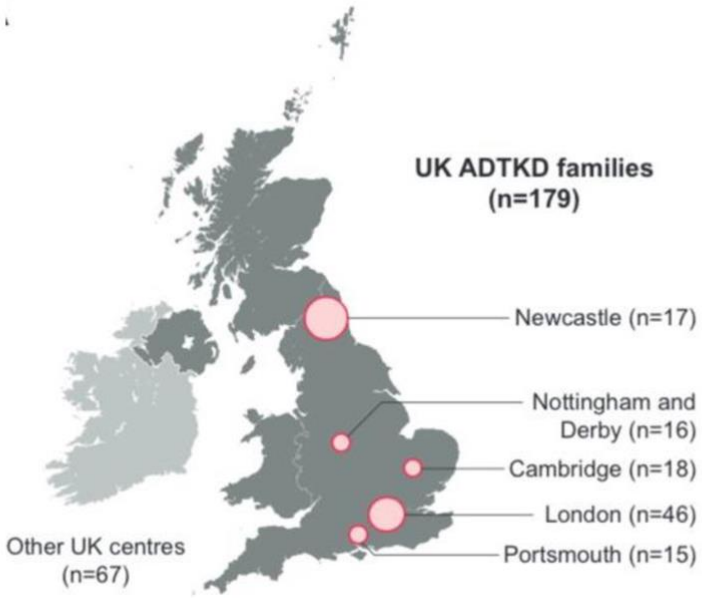

Figure 5: Major UK centres (with family numbers) from which ADTKD families have been recruited into the RaDaR and Genomics England 100 000 Genomes Project*

*[*A geographical representation of the United Kingdom taken from (Mabillard, Sayer and Olinger, 2023). The number of families with ADTKD in the RaDaR registry (n=179) is acknowledged (registry snapshot obtained July 2021). The major UK centers who have contributed the most ADTKD families to the UK RaDaR Registry are labelled.]*

2.3.2 Evidence of Disease Heterogeneity beyond the Primary UMOD Variant

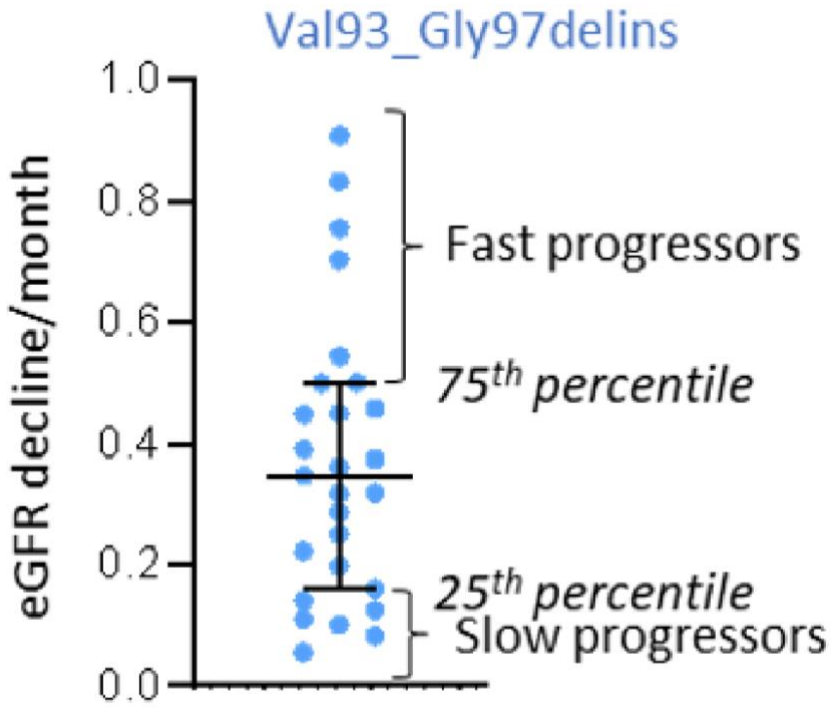

Figure 6: Variability in Disease Progression in ADTKD-*UMOD* Patients with the same Primary *UMOD* p.Val93_Gly97delinsAlaAlaSerCys Variant*

*[*A scatter plot encompassing 28 individuals with ADTKD-UMOD secondary to the UMOD p.Val93_Gly97delinsAlaAlaSerCys variant. These data were derived from aggregate eGFR values available in the national RaDaR registry, rather than full clinical records. To calculate the slope of eGFR decline per month, serial eGFR values were plotted against time in months from the first measurement. A linear trendline was applied to the scatter plot, and the slope of the resulting regression equation (eGFR = m × time + c) was taken to represent the monthly rate of eGFR change (mL/min/1.73m² per month), where* m *reflects the average rate of decline and* c *is the estimated baseline eGFR. The plot suggests that there is variability in the rate of eGFR decline between individuals despite the same primary disease-causing UMOD variant.]*

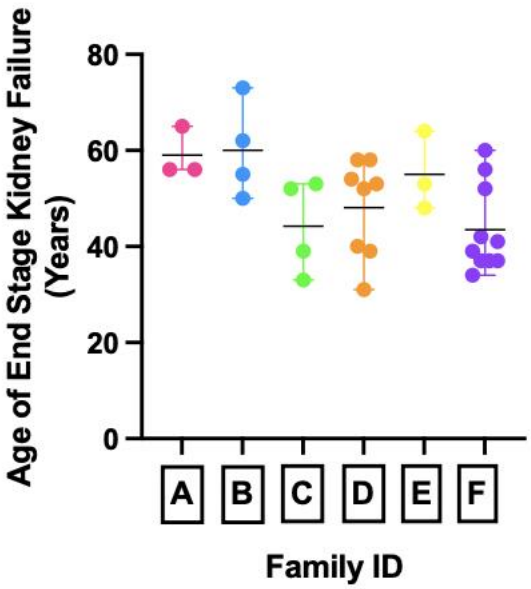

Figure 7: Intra-familial variability in Age of Onset of Kidney Failure in Six Families with ADTKD-*UMOD**

*[*A scatter plot encompassing 32 individuals from the six largest families with ADTKD-UMOD from both the RaDaR registry and from data obtained from collaborators in other centres for those who have already reached End Stage Kidney Disease (ESKD). Each family is represented by its own colour and letter (A to F), and each dot represents an individual within that family who has reached ESKD. Mean (years) per family (A to F): 59, 60, 44.25, 48.13, 55, 43.5. Range (years) per family (A to F): 9, 23, 20, 27, 16, 26. The plot suggests that there is significant variability of age of ESKD spanning almost 30 years in some families despite the same primary disease-causing UMOD variant.]*

2.3.3 Genetic Spectrum of ADTKD-UMOD in the UK

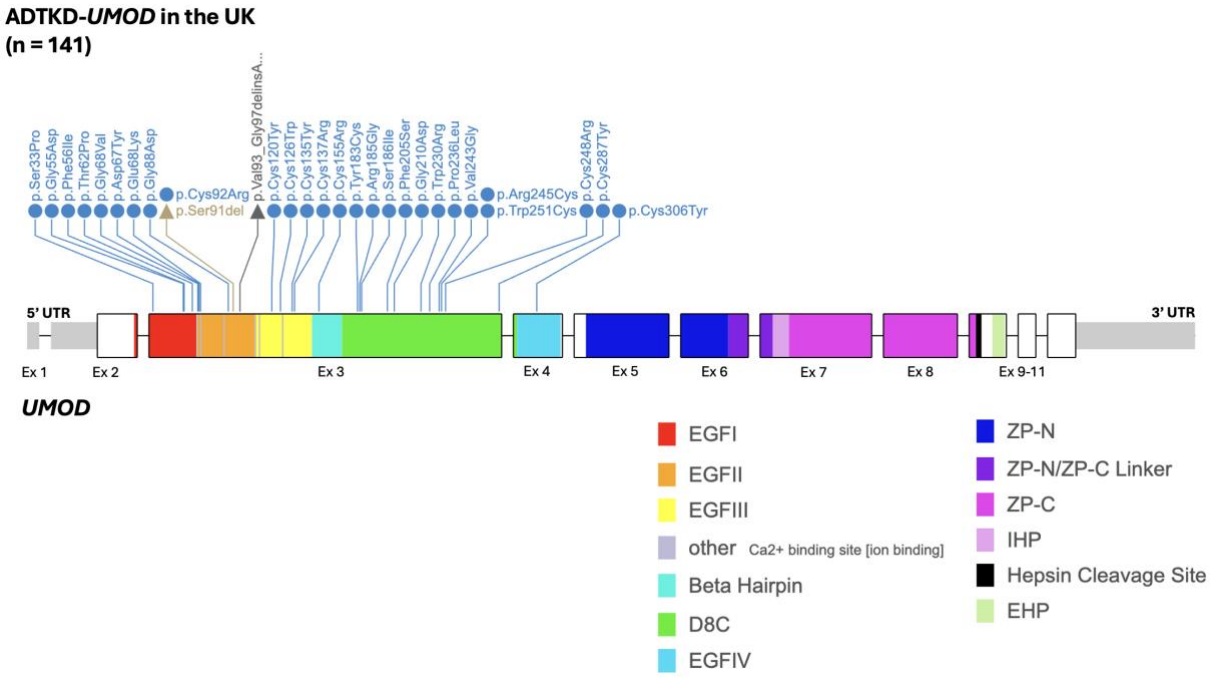

Figure 8: The Spectrum of *UMOD* variants in the UK in those with ADTKD-*UMOD**

*[*A schematic representation of the UMOD protein structure with exons 1-11 aligned to their corresponding protein domains. Of the 141 patients, 29 UMOD variants were found in the UK ADTKD-UMOD Cohort and are shown above with annotated protein consequences. The blue circles indicate missense variants while the triangles represent deletions and the recurring UMOD p.Val93_Gly97delinsAlaAlaSerCys variant present in the UK population. The coloured blocks indicate key structural and functional domains in the protein. As can be observed, all UK UMOD variants occur on exons 3 and 4 of the gene where clustering occurs in the highly conserved cysteine-rich regions critical to the protein's intricate folding.]*

Primary *UMOD* Variant (NM_003361.4)	Number of patients	Allele Frequency (gnomAD European non-Finnish)	ACMG Classification
p.Val93_Gly97delinsAlaAlaSerCys	81	0.00001198	Pathogenic
p.Ser91del	10	0.000	Likely pathogenic
p.Cys120Tyr	6	0.000	Uncertain significance
p.Trp230Arg	4	8.538e-7	Likely pathogenic
p.Gly55Asp	4	8.474e-7	-
p.Phe56Ile	3	0	-
p.Cys137Arg	3	0	-
p.Tyr183Cys	2	0.000003491	Likely pathogenic/Uncertain significance
p.Thr62Pro	2	0.0006856	Uncertain significance/Likely benign/Benign
p.Ser186Ile	2	0	-
p.Gly68Val	2	0	-
p.Cys306Tyr	2	0	-
p.Cys287Tyr	2	0	-
p.Cys155Arg	2	0	-
p.Arg185Gly	2	0	-
p.Val243Gly	1	0	-
p.Trp251Cys	1	0	-
p.Tyr33Pro	1	0	Uncertain significance
p.Pro236Leu	1	8.518e-7	Pathogenic/Likely pathogenic
p.Phe205Ser	1	0	-
p.Gly88Asp	1	0	Likely pathogenic/Uncertain significance
p.Gly210Asp	1	0	-
p.Glu68Lys	1	0	Uncertain significance
p.Cys92Arg	1	0	Pathogenic
p.Cys248Arg	1	0	Likely pathogenic
p.Cys135Tyr	1	0	-
p.Cys126Trp	1	0	-
p.Asp67Tyr	1	0	-
p.Arg245Cys	1	0	-

Table 1: The UK Spectrum of *UMOD* variants with ACMG Classifications*

*[*The spectrum of UMOD (NM_003361.4) variants in the UK ADTKD-UMOD Cohort along with the number of individuals with each of the 29 variants and the variant American College of Medical Genetics and Genomics Classification where assigned.]*

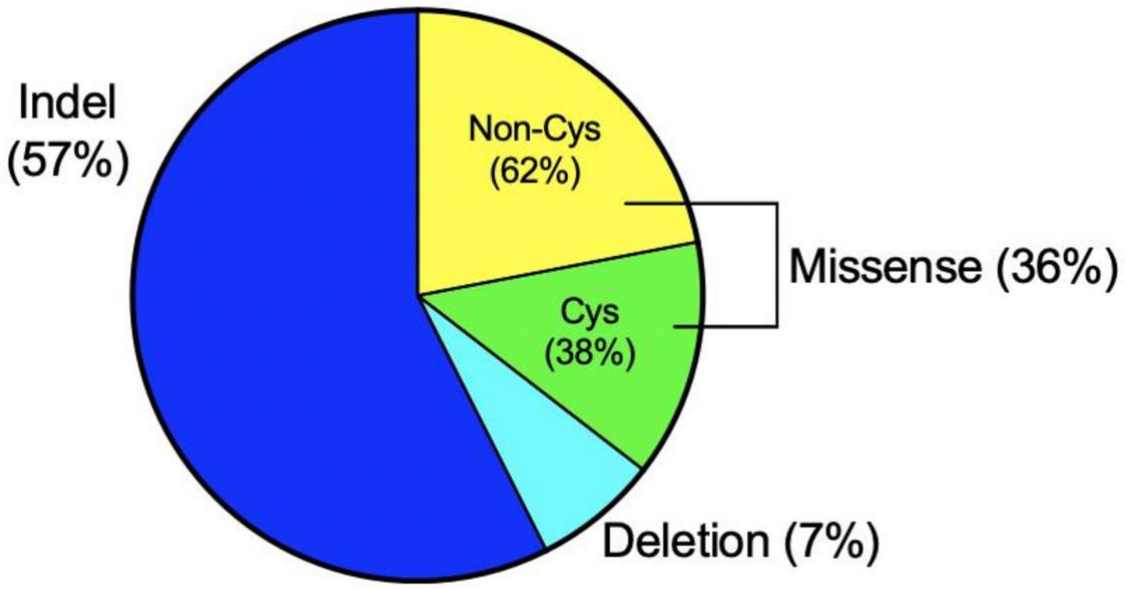

Figure 9: The UK spectrum of Primary *UMOD* variant type*

*[*A pie chart representing the UK ADTKD-UMOD Cohort spectrum of UMOD variants, categorised into the following groups: Deletion (these are the small in-frame single amino acid deletion: UMOD p.Ser91del), Missense and Insertion-Deletion 'indel' with associated percentage proportions. The missense variants are further broken down into whether there is involvement of one of the 48 highly conserved cysteine residues necessary for complex tertiary protein folding.]*

Unreported *UMOD* Variants (NM_003361.4)	Number	Sex	Current Age	Highest serum uric Acid (umol/L)	First attack of gout (years)	Age of ESKD	If no ESKD, current eGFR (CKD-EPI; ml/min/1.73m^2)
p.Gly55Asp	4	F	78	362	77	NA	35
		M	73	398	NA	70	NA
		M	83	363	71	74	NA
		M	64	469	NA	NA	35
p.Phe56Ile	3	F	53	?	NA	50	NA
		F	74	?	NA	43	NA
		F	28	237	NA	NA	80
p.Cys137Arg	3	F	61	340	NA	NA	79
		F	73	?	NA	73	NA
		F	68	363	NA	55	NA
p.Ser186Ile	2	F	52	?	?	?	? (chronic TIN on kidney biopsy)
		F	?	?	?	?	? (glomerulosclerosis on kidney biopsy)
p.Gly68Val	2	F	91	?	?	NA	33
		F	57	390	?	45	NA
p.Cys287Tyr	2	F	59	700	?	27	NA
		F	26	430	?	NA	18
p.Val243Gly	1	M	38	778	19	NA	38
p.Phe205Ser	1	M	61	?	34	?	?
p.Gly210Asp	1	F	68	?	NA	46	NA
p.Cys126Trp	1	M	58	431	NA	NA	50
p.Asp67Tyr	1	M	53	364	NA	NA	26
p.Arg245Cys	1	M	43	608	36	NA	53

Table 2: Unreported UK *UMOD* Variant Phenotype Details*

*[*The 12 UMOD variants that have not been reported in the literature or have been found in other cohorts. This was confirmed using Varsome, ClinVar, GnomAD and LitVar. The table highlights the patient demographics along with kidney function and serum uric acid data to aid genotype-phenotype correlations and the deleterious nature of each unreported variant.]*

2.3.4 Genotype-Phenotype Correlations

	UK Cohort (n=260)	Published Wake Forrest Cohort	Published International Cohort	
ADTKD-*UMOD* Genotyped	**134 (52%)**	**44 (61%)**	**322 (98%)**	
Number of Individuals who reached ESKD	64/136 (47%) 124/260 (48%) Unknown	303 (53%)	123 (37%)	Chi-square test p<0.0001
Age of ESKD (median)	50.00	48.65	48.88	
Race	204 (78%) White 38 (15%) Unreported 12 (5%) Asian 2 (1%) Black	627 (96%) White 16 (2%) Asian/pacific islander 6 (1%) Unreported 3 (0.5%) Black 2 (0.3%) Indian	271 (82%) White 58 (18%) Unreported	
Ethnicity	195 (75%) White British 38 (15%) Unreported 10 (4%) Pakistani 5 (2%) White European 3 (1%) Chinese 3 (1%) White Irish 2 (1%) Caribbean 1 (<1%) Indian	613 (94%) Not Hispanic/Latino 22 (3%) Unreported 13 (2%) Other 6 (1%) Hispanic/Latino	251 (76%) Not Hispanic/Latino 54 (16%) Unreported 24 (7%) Hispanic/Latino	
Female	139 (53%) 19 Unknown (7%)	325 (50%)	168 (51%)	Chi-square test 0.1176
Male	102 (39%) 19 Unknown (7%)	325 (50%)	162 (49%)	Chi-square test 0.1176
Gout	24/104 (23%) *156 (60%) Uncertain	43 (17%) *198 (78%) Uncertain	245 (37%) *207 (32%) Uncertain	Chi-square test p<0.0001

Proportion of Females with Gout	11/24 (46%)	12 (11%) *90 (83%) Uncertain	108 (33%) *97 (29%) Uncertain	Chi-square test p<0.0001
Proportion of Males with Gout	13/24 (54%)	31 (21%) *108 (7%) Uncertain	137 (42%) *110 (34%) Uncertain	Chi-square test p<0.0001
Age of Gout Onset	39.69 +/- 15.99 (n=16)	27.2 +/- 9.5 (n = 20)	30.2 +/- 11.4 (n = 217)	
Age of Gout Onset (Females)	38 +/- 15.57 (n=8)	25.4 +/- 8.0 (n = 7)	31.7 +/- 12.7 (n = 99)	
Age of Gout Onset (Males)	41.38 +/- 17.30 (n=8)	28.2 +/- 10.4 (n = 13)	29.0 +/- 9.9 (n = 118)	

Table 3: Summary of UK ADTKD-*UMOD* Cohort Demographics and Phenotypes*

*[*The characteristics of UK ADTKD-UMOD Cohort in an international context. Demographics and phenotype information have been compared to published information on the Wake Forrest (USA) and International ADTKD-UMOD Cohorts (Kidd et al., 2020).]*

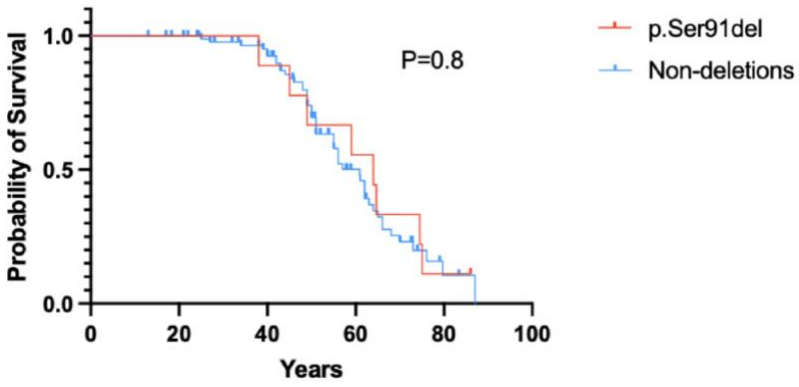

Figure 10: Kidney Survival for UK *UMOD* p.Ser91del Variants*

*[*Kidney survival differences between individuals in the UK ADTKD-UMOD cohort with the UMOD p.Ser91del variant versus the rest of the cohort. When a Cox proportional hazards model was applied, no statistically significant difference in kidney survival was observed.]*

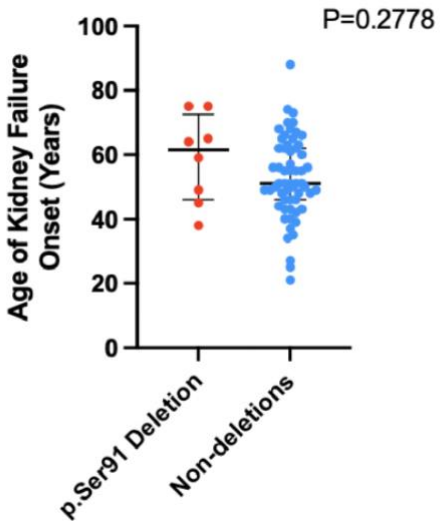

Figure 11: Age of Onset of Kidney Failure for UK *UMOD* p.Ser91del Variants*

*[*The distribution of age of onset of kidney failure in years for those in the UK ADTKD-UMOD cohort with the UMOD p.Ser91del variant versus the rest of the cohort. A Mann-Whitney test was applied which illustrated a non-significant difference in the median age of onset of kidney failure between the two groups.]*

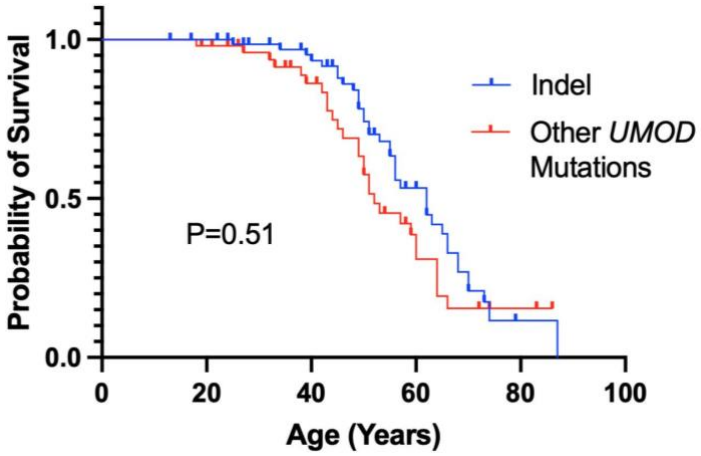

Figure 12: Kidney Survival for UK *UMOD* p.Val93_Gly97delinsAlaAlaSerCys Variant*

*[*Kidney survival differences between individuals in the UK ADTKD-UMOD cohort with the UMOD p.Val93_Gly97delinsAlaAlaSerCys variant versus the rest of the cohort. When a Cox proportional hazards model was applied, no statistically significant difference in kidney survival was observed.]*

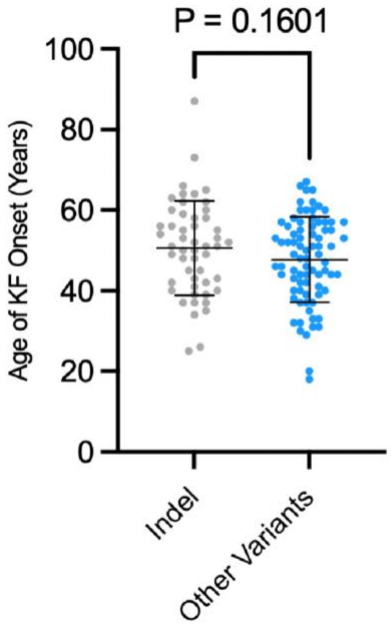

Figure 13: Age of Onset of Kidney Failure for UK *UMOD* p.Val93_Gly97delinsAlaAlaSerCys Variant*

*[*The distribution of age of onset of kidney failure in years for those in the UK ADTKD-UMOD cohort with the UMOD p.Val93_Gly97delinsAlaAlaSerCys variant versus the rest of the cohort. A Mann-Whitney test was applied which illustrated a non-significant difference in the median age of onset of kidney failure between the two groups.]*

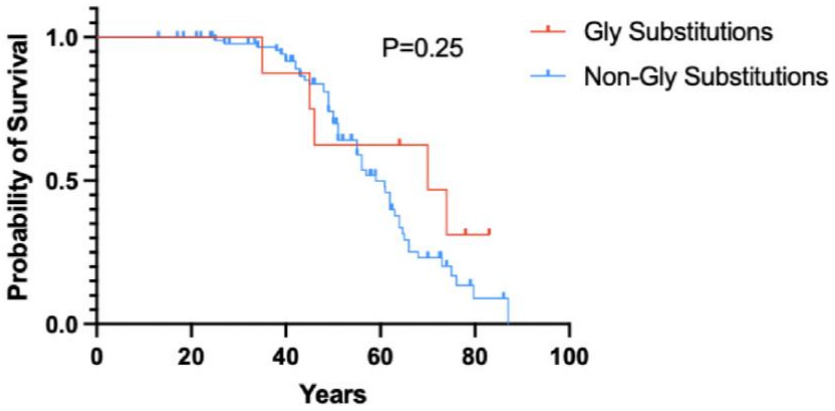

Figure 14: Kidney Survival for UK *UMOD* Glycine Substitutions*

*[*Kidney survival differences between individuals in the UK ADTKD-UMOD cohort with UMOD variants resulting in glycine substitutions versus the rest of the cohort. When a Cox proportional hazards model was applied, no statistically significant difference in kidney survival was observed.]*

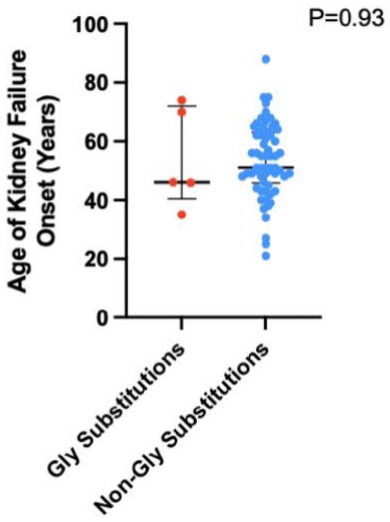

Figure 15: Age of Onset of Kidney Failure for UK *UMOD* Glycine Substitutions*

*[*The distribution of age of onset of kidney failure in years for those in the UK ADTKD-UMOD cohort with UMOD variants resulting in glycine substitutions versus the rest of the cohort. A Mann-Whitney test was applied which illustrated a non-significant difference in the median age of onset of kidney failure between the two groups.]*

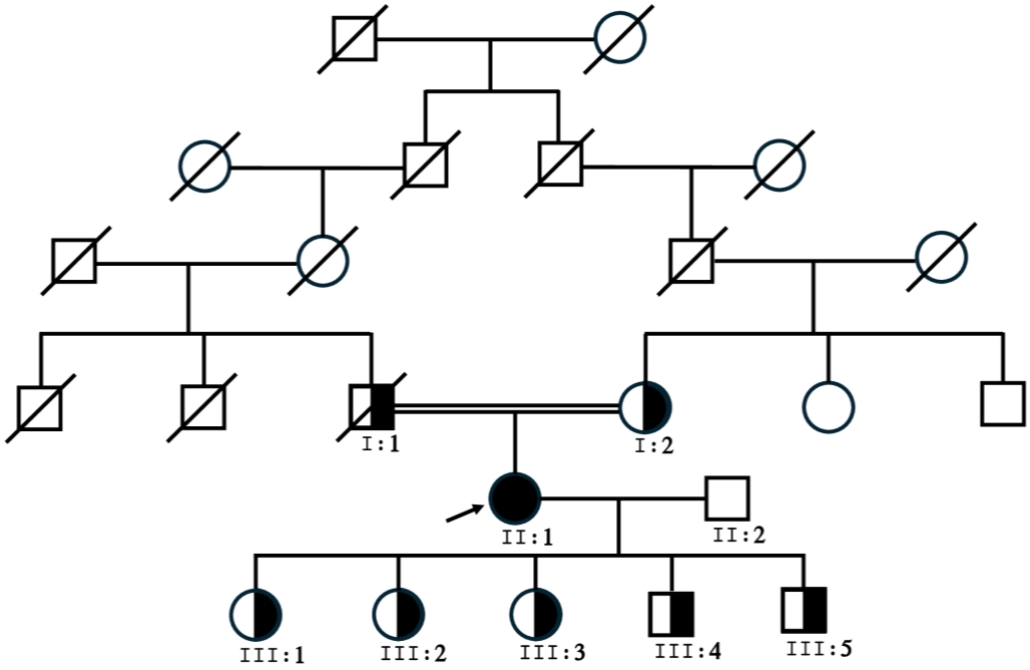

Figure 16: ADTKD-*UMOD* Homozygote Family Pedigree*

*[*A family pedigree for the only homozygote in the UK ADTKD-UMOD Cohort. The pedigree shows this individual who is homozygous for the UMOD p.Cys120Tyr variant, including her five obligate heterozygote children and heterozygote parents.]*

Patient	Sex	Genotype p.C120Y	Current Age	Highest serum uric acid (umol/L)	Fractional excretion of uric acid	First attack of gout (years)	Current serum creatinine (umol/L)	Current eGFR (CKD-EPI; ml/min/1.73m^2)
Homozygote II:1	F	Y/Y	52	614	0.026	31	164	31
Heterozygote I:1	M	C/Y	Deceased aged 60	Unknown	Unknown	Unknown	Unknown	Unknown
Heterozygote I:2	F	C/Y	83	Unknown	Unknown	64	Unknown	Unknown
Heterozygote III:1	F	C/Y	33	364	0.052	NA	90	72
Heterozygote III:2	F	C/Y	32	318	0.087	NA	72	>90
Heterozygote III:3	F	C/Y	30	322	0.113	NA	81	84
Heterozygote III:4	M	C/Y	28	432	0.045	NA	121	69
Heterozygote III:5	M	C/Y	23	477	0.041	NA	116	76

Table 4: ADTKD-*UMOD* Homozygote Family Pedigree Phenotype Details*

*[*Demographic and phenotype details for the individual who is homozygous for the UMOD p.Cys120Tyr variant and her heterozygous family members. This variant was previously suggested in the literature as being a hypomorphic variant, however, this more recent phenotypic information suggest a more deleterious effect (Edwards et al., 2017).*

Phenotype	Local cohort (n=70)	*UMOD* p.(Val93_Gly97delinsAlaAlaSerCys) variant (n=54)	Other *UMOD* variants (n=16)	Fisher's Exact Test
Sex	**Female = 53 (76%)** **Male = 17 (24%)**	**Female = 41 (76%)** **Male = 13 (24%)**	**Female 12 (75%)** **Male = 4 (25%)**	P>0.9999
Ethnicity	White = 64 (91%) British-Pakistani = 6 (9%)	White = 54 (100%)	White = 10 (63%) British-Pakistani – 6 (37%)	**P<0.0001**
FH of kidney disease	Yes = 69 (99%) No = 1 (1%)	Yes = 53 (98%) No = 1 (2%)	Yes = 16 (100%) No = 0 (0%)	P>0.9999
Gout	Yes = 9 (13%) No = 58 (83%) Unknown = 3 (4%)	Yes = 6 (11%) No = 46 (85%) Unknown = 2 (4%)	Yes = 3 (19%) No = 12 (75%) Unknown = 1 (6%)	P=0.4075
Gout Onset <30 years	Yes = 4 (6%) No = 60 (87%) Unknown = 5 (7%) N/A = 1	Yes = 1 (19%) No = 48 (91%) Unknown = 4 (8%) NA = 1	Yes = 3 (19%) No = 12 (75%) Unknown = 1 (6%)	**P=0.0372**
Gout Onset >30 years	Yes = 3 (5%) No = 58 (95%) N/A = 9	Yes = 3 (7%) No = 43 (93%) NA = 8	Yes = 0 (0%) No = 15 (100%) NA = 1	P=0.5686
Serum uric acid >500 (μmol/L)	Yes = 7 (10%) No = 51 (73%) Unknown = 12 (17%)	Yes = 5. (9%) No = 40 (74%) Unknown = 9 (17%)	Yes = 2 (12.5%) No = 11 (69%) Unknown = 3 (19%)	P=0.6478
Serum uric Acid 360-500 (μmol/L)	Yes = 20 (29%) No = 38 (54%) Unknown = 12 (17%)	Yes = 16 (30%) No = 29 (54%) Unknown = 9 (17%)	Yes = 4 (25%) No = 9 (56%) Unknown = 3 (19%)	P>0.9999
Serum uric Acid always normal	Yes = 31 (44%) No = 27 (39%) Unknown = 12 (17%)	Yes = 24 (44%) No = 21 (39%) Unknown = 9 (17%)	Yes = 7 (44%) No = 6 (38%) Unknown = 3 (19%)	P>0.9999
Haematuria	Yes = 16 (23%) No = 40 (57%) Unknown = 14 (20%)	Yes = 4 (6%) No = 39 (72%) Unknown = 11 (20%)	Yes = 1 (6%) No = 12 (75%) Unknown = 3 (19%)	P>0.9999

Proteinuria	Yes = 5 (7%) No = 51 (73) Unknown = 14 (20%)	Yes = 14 (26%) No = 29 (54%) Unknown = 11 (20%)	Yes = 2 (12.5%) No = 11 (69%) Unknown = 3 (19%)	P=0.3078
Haematoproteinuria	Yes = 3 (4%) No = 53 (76%) Unknown = 14 (20%)	Yes = 3 (6%) No = 40 (74%) Unknown = 11 (20%)	Yes = 0 (0%) No = 13 (81%) Unknown = 3 (19%)	P>0.9999
Diabetes (clinical dx or HbA1c >47mmol/mol)	Yes = 4 (6%) No = 63 (90%) Unknown = 3 (4%)	Yes = 4 (7%) No = 48 (89%) Unknown = 2 (4%)	Yes = 0 (0%) No = 15 (94%) Unknown = 1 (6%)	P=0.5675
Hypertension (clinical dx) or 2 BPs >130/80 mmHg	Yes = 28 (40%) No = 32 (46%) Unknown = 10 (14%)	Yes = 24 (44%) No = 22 (41%) Unknown = 8 (15%)	Yes = 4 (25%) No =10 (63%) Unknown = 2 (12.5%)	P=0.1403
BP low or salt-wasting symptoms (from clinic letters)	Yes = 6 (9%) No = 53 (76%) Unknown = 11 (16%)	Yes = 4 (7%) No = 41 (76%) Unknown = 9 (17%)	Yes = 2 (12.5%) No = 12 (75%) Unknown = 2 (12.5%)	P=0.6204
Abnormal renal USS	Yes = 25 (36%) No = 31 (44%) Unknown = 14 (20%)	Yes = 20 (37%) No = 25 (46%) Unknown = 9 (17%)	Yes = 5 (31%) No = 6 (38%) Unknown = 5 (31%)	P>0.9999
Cysts on US	Yes = 15 (21%) No = 41 (59%) Unknown = 14 (20%)	Yes = 11 (20%) No = 34 (63%) Unknown = 9 (17%)	Yes = 4 (25%) No = 7 (44%) Unknown = 5 (31%)	P=0.4610
Structural defect on US (no cysts)	Yes = 14 (20%) No = 42 (60%) Unknown = 14 (20%)	Yes = 12 (22%) No = 33 (61%) Unknown = 9 (17%)	Yes = 2 (12.5%) No = 9 (56%) Unknown = 5 (31%)	P=0.7115
BMI >26 and <30 (kg/m^2)	Yes = 10 (14%) No = 34 (49%) Unknown = 26 (37%)	Yes = 9 (17%) No = 25 (46%) Unknown = 20 (37%)	Yes = 1 (6%) No = 9 (56%) Unknown = 6 (38%)	P=0.4105
BMI >30 (kg/m^2)	Yes = 14 (20%) No = 30 (43%) Unknown = 26 (37%)	Yes = 11 (20%) No = 23 (43%) Unknown = 20 (37%)	Yes = 3 (19%) No = 7 (44%) Unknown = 6 (38%)	P>0.9999
Smoker/previous smoker	Yes = 17 (24%) No = 42 (60%)	Yes = 15 (28%) No = 31 (57%)	Yes = 2 (12.5%) No = 11 (69%)	P=0.3101

	Unknown = 11 (16%)	Unknown = 8 (15%)	Unknown = 3 (19%)	
Recurrent UTIs	Yes = 8 (11%) No = 60 (86%) Unknown = 2 (3%)	Yes = 6 (11%) No = 46 (85%) Unknown = 2 (4%)	Yes = 2 (12.5%) No = 14 (88%)	P>0.9999
Kidney Stone Episode(s)	Yes = 2 (3%) No = 67 (96%) Unknown = 1 (1%)	Yes = 2 (4%) No = 51 (94%) Unknown = 1 (2%)	Yes = 0 (0%) No = 16 (100%)	P>0.9999

Table 5: Detailed Phenotype Details for Local ADTKD-*UMOD* Cohort*

*[*Detailed phenotype differences between those in the local ADTKD-UMOD cohort who carry the UMOD p.Val93_Gly97delinsAlaAlaSerCys 'indel' variant in comparison to those who carry a different UMOD variant.*

Total = 27	Females (n = 20)			Males (n = 7)			Fisher's Exact Test
	All	**Indels (n = 13)**	**Non-indels (n = 7)**	**All**	**Indels (n = 4)**	**Non-indels (n = 3)**	
Family History of Gout	**10 (50%)**	**4 (31%)**	**6 (86%)**	**4 (57%)**	**2 (50%)**	**2 (67%)**	P=0.0573 (Indels vs Non-indels)
Pregnancy complications **Those who have experienced pregnancy*	8/*14 (57%)	5	3	NA	NA	NA	P>0.9999
Symptoms suggestive of salt-wasting							
Bedwetting	5 (25%)	4 (31%)	1 (14%)	1 (14%)	0 (0%)	1 (33%)	0.3285 (> 1 symptoms vs no symptoms in Indels vs Non-indels)
Fainting/postural symptoms	4 (20%)	1 (8%)	3 (43%)	1 (14%)	0 (0%)	1 (33%)	
Low blood pressure (<125/75)	7 (35%)	4 (31%)	3 (43%)	1 (14%)	0 (0%)	1 (33%)	
Perceived polyuria	11 (55%)	5 (38%)	6 (86%)	2 (29%)	2 (50%)	0 (0%)	
Nocturia	10 (50%)	5 (38%)	5 (71%)	1 (14%)	0 (0%)	1 (33%)	
Urinary frequency	9 (45%)	4 (31%)	5 (71%)	2 (29%)	1 (25%)	1 (33%)	
Excessive thirst	7 (35%)	5 (38%)	2 (29%)	0 (0%)	0 (0%)	0 (0%)	

Table 6: Further Phenotype Details for Local ADTKD-*UMOD* Cohort*

*[*Further phenotype details for those in the local ADTKD-UMOD Cohort some of which haven't been described in detail in the literature to date. This includes pregnancy complications and details of salt-wasting symptoms. 'Indels' represent those with the UMOD p.Val93_Gly97delinsAlaAlaSerCys variant. Pregnancy complications included: recurrent miscarriages, Deep Vein Thrombosis, pre-eclampsia, polyhydramnios, pre-term labour, neonatal brain oedema and death at 4 days, recurrent ectopics, hypotension, pyelonephritis, Gestational Diabetes.]*

	Salt-wasting symptoms (out of 7)	eGFR (mls/min/1.73m^2)	Serum Osmolality (mOsm/kg)	Urine Osmolality (mOsm/kg)	Urine Sodium (mmol/L)	Fractional Excretion of Urate
1	0	24	296	478	48	0.059
2	3	45	297	279	37	0.055
3	4	29	309	489	94	0.157
4	4	76	287	756	112	0.047
5	2	90	296	625	89	0.045
6	0	90	287	697	62	0.075
7	2	19	307	327	48	0.084
8	2	50	292	855	204	0.043
9	0	90	283	756	81	0.073
10	0	90	290	843	131	0.059
11	4	34	312	350	66	0.026
12	6	74	285	768	112	0.052
13	0	76	285	342	28	0.041
14	3	72	291	1077	155	0.087
15	0	81	284	249	33	0.045
16	3	50	291	547	88	0.075
17	1	86	288	917	112	0.043
18	2	90	296	827	170	0.064
19	1	79	296	266	38	0.051
20	0	90	297	431	128	0.098
21	5	90	292	?	?	?
22	3	90	297	?	?	?
23	6	80	287	583	92	0.085
24	3	19	304	371	41	?
25	1	90	293	841	151	?
26	2	90	287	897	176	0.025
27	0	90	295	370	57	0.081
Total (%)	**Score >4: 6 (22%)**	**-**	**Serum Osmolality <275: 0 (0%)**	**Urine Osmolality >300: 22 (88%)**	**Urine Osmolality >50: 18 (72%)**	**FEurate <0.05: 8 (35.7%)**

Table 7: Salt-wasting Phenotypes in the Local ADTKD-*UMOD* Cohort (A)*

*[*A biochemical characterisation of salt-wasting and its association with fractional excretion of uric acid. Salt-wasting symptoms, kidney function, biochemical features suggesting salt-wasting and fractional excretion of uric acid in a subset of the local ADTKD-UMOD Cohort. Biochemical salt-waster defined by: Urine Na >50 mmol/L, Urine Osm >300 mOsm/kg, Serum Osm <275 mOsm/kg*.

Variable 1	Variable 2	Fisher's Exact Test	Interpretation
Over 4 salt-wasting symptoms	Urine Osmolality >300mOsm/kg	**P=0.9999**	No significant association
Over 4 salt-wasting symptoms	Urine Sodium >50mmol/L	**P=0.2801**	No significant association
Over 4 salt-wasting symptoms	FEurate <0.05	**P=0.9999**	No significant association
Urine Osmolality >300mOsm/kg	FEurate <0.05	**P=0.9999**	No significant association
Urine Sodium >50mmol/L	FEurate <0.05	**P=0.9999**	No significant association
Over 4 salt-wasting symptoms	Female sex	**P=0.1372**	No significant association

Table 8: Salt-wasting Phenotypes in the Local ADTKD-*UMOD* Cohort (B)*

*[*A biochemical characterisation of salt-wasting and its association with fractional excretion of uric acid. Use of Fisher's exact test to determine significance of association salt-wasting symptoms, biochemistry, sex and fractional excretion of uric acid in the same individuals from the local ADTKD-UMOD Cohort.]*

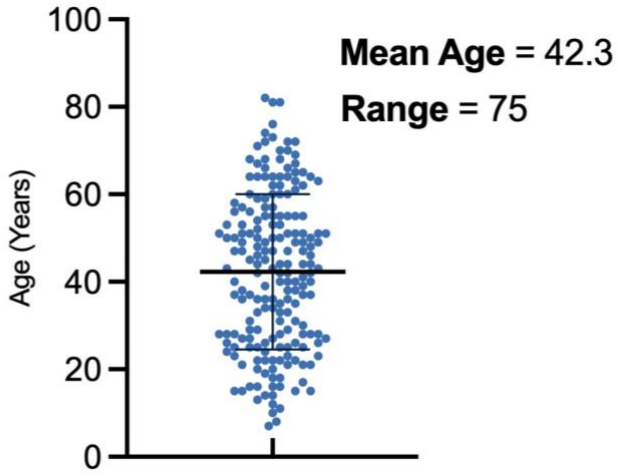

Figure 17: Age of Diagnosis in the UK ADTKD-*UMOD* Cohort*

*[*A scatter plot of the age of diagnosis (years) for individuals within the UK ADTKD-UMOD cohort, where available, which includes the mean age as 42.3 years and a large range of 75 years spanning from <15 to >80. Each dot represents one individual. Data points are horizontally spread using GraphPad Prism to improve visualisation of overlapping values.]*

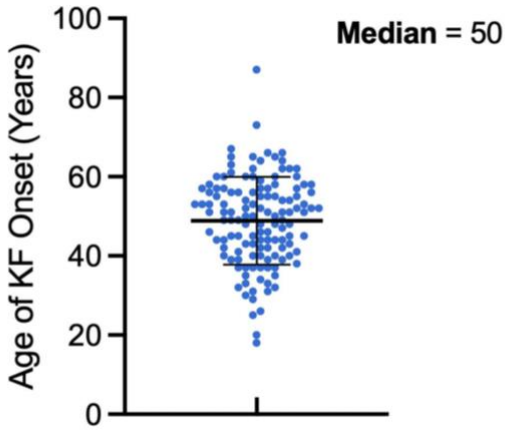

Figure 18: Age of Kidney Failure Onset in the UK ADTKD-*UMOD* Cohort*

*[*A scatter plot of the age of kidney failure onset (years) for individuals within the UK ADTKD-UMOD cohort, where available, which includes a median age of 50 and a large range from <20 to >80 years. Kidney failure was defined as an estimated Glomerular Filtration Rate of <10mls/min/1.73m^2 or in receipt of kidney replacement therapy. Each dot represents one individual. Data points are horizontally spread using GraphPad Prism to improve visualisation of overlapping values.]*

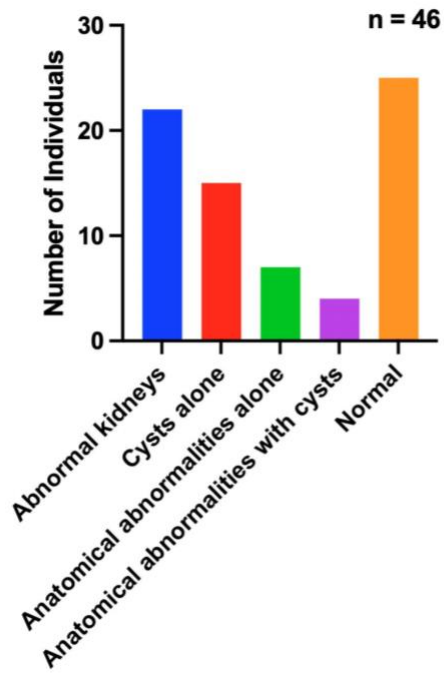

Figure 19: Native Kidney Ultrasound Findings in the Local ADTKD-*UMOD* Cohort*

*[*A bar chart to reflect the proportion of individuals within the UK ADTKD-UMOD cohort who had kidney cysts and/or structural abnormalities on ultrasound. Almost half of patients had an 'abnormal ultrasound', one third of patients had kidney cysts and less than 10% of patients had anatomical abnormalities alone. Findings were extracted from formal ultrasound reports*

documented by radiologists as part of routine clinical care and as featured on the Electronic Patient Record (EPR).]

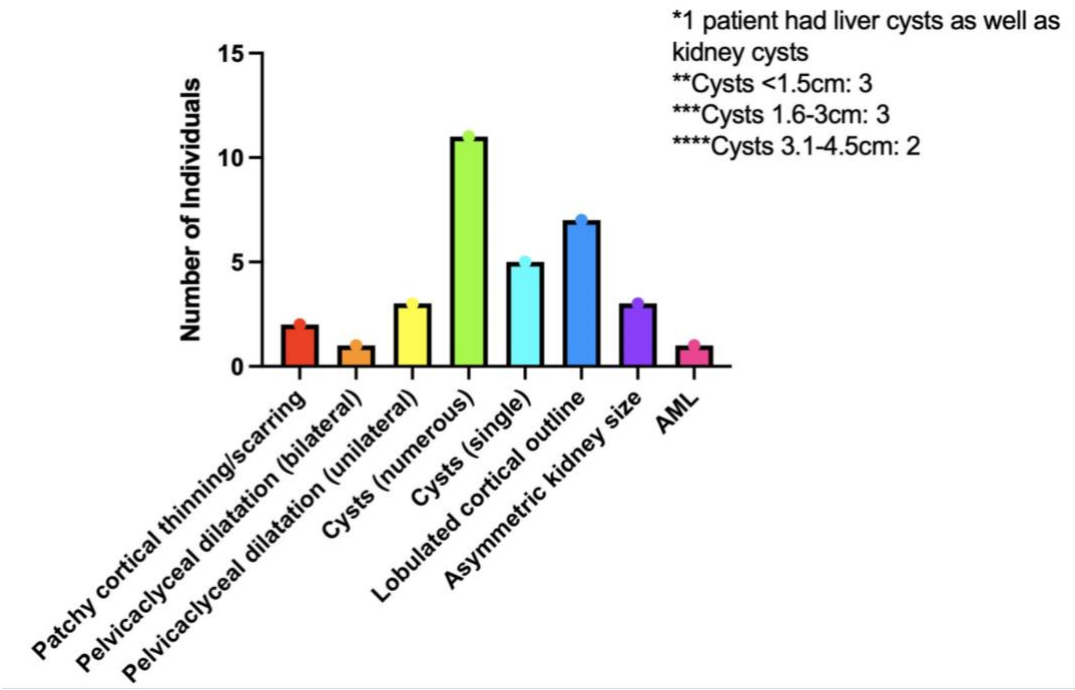

Figure 20: Native Kidney Imaging Characteristics in the Local ADTKD-*UMOD* Cohort*

*[*A bar chart of more detailed ultrasound characteristics of the local ADTKD-UMOD patients for those patients who had an 'abnormal ultrasound'. Findings were extracted from formal ultrasound reports documented by radiologists as part of routine clinical care and as featured on the Electronic Patient Record (EPR).]*

Abnormal USS	Cysts on USS	Structural Defect on USS	Abnormal USS
***UMOD* p.(Val93_Gly97delinsAlaAlaSerCys) variant**	19/43 (44%)	11/43 (26%)	11/43 (26%)
Other *UMOD* variants	5/13 (38%)	4/13 (31%)	2/13 (15%)
Fisher's exact test	**P=0.76**	**P=0.73**	**P=0.71**

Table 9: Kidney Imaging in Local ADTKD-UMOD Cohort: those with *UMOD* p.(Val93_Gly97delinsAlaAlaSerCys) variant compared to those with other *UMOD* variants*

*[*Kidney ultrasound imaging in those with the UMOD p.Val93_Gly97delinsAlaAlaSerCys 'indel' variant in comparison to those with other UMOD variants along with a Fisher's exact test to*

determine any potential statistically meaningful differences of ultrasound findings between these ADTKD-UMOD subtypes.]

Patient (Age at Biopsy)	Glomeruli	Tubules	Interstitium	Blood vessels	Immunofluorescence	Electron Microscopy	Summary
1 (64 years)	2/5 globally sclerosed. 1/5 periglomerular fibrosis and ischaemic wrinkling.	Moderate atrophy. Red cell casts seen.	Moderate mononuclear cell infiltrate. Moderate fibrosis.	Mild arteriolar arterialisation. Mild intimal fibrosis and intimal thickening	No specific deposition	No specific abnormality	Non-specific chronic tubulointerstitial damage. Glomerular-origin haematuria.
2 (47 years)	11 normal glomeruli.	Mild tubular atrophy	Mild interstitial fibrosis. Subcapsular focal infiltrate of acute and chronic inflammatory cells.	Mild arterial intimal fibroelastosis	No specific deposition	No specific abnormality	Mild tubular atrophy and interstitial fibrosis
3 (48 years)	14/36 glomeruli globally sclerosed. Diffuse increase in mesangial cellularity and matrix. 2/49 show focal sclerosis/hyalinosis. Some show periglomerular fibrosis.	Moderate atrophy. Occasional red cell casts seen.	Moderate mononuclear cell infiltrate. Moderate fibrosis.	Severe arteriolar arterialisation. Severe intimal fibrosis and intimal thickening	Moderate segmental granular deposition of IgM and C3.	Not performed	Appearances suggest FSGS. Vascular changes are consistent with hypertension.

4 (36 years)	No comment	Mild tubular atrophy	Mild interstitial fibrosis	No comment	No comment	No comment	Mild chronic tubulointerstitial damage
5 (27 years)	No glomeruli available	Mild tubular atrophy	Mild mononuclear cell infiltrate. Mild interstitial fibrosis. Large collection of acellular material with a fibrillary appearance in the medulla consistent with an inflammatory reaction.	Unremarkable	Weak granular segmental mesangial deposition of IgM and C1q.	No specific abnormality	Mild chronic tubulointerstitial damage, The large collection of acellular fibrillary material in the medulla is highly suspicious of ADTKD.

Table 10: Kidney Biopsy Findings in Local ADTKD-*UMOD* Cohort*

*[*Kidney biopsy findings for any local ADTKD-UMOD patients who have received a kidney biopsy and a breakdown of features based on different features throughout the nephron including on light microscopy, immunofluorescence and electron microscopy.]*

2.3.5 Clinical Modifiers of Disease Progression in ADTKD-UMOD

Predictor Variable/ Covariates (pre-failure)	Outcome Variable	Method	N	Number of Events	Regression Coefficients ('coef')	Hazard ratio ('exp(coef)')	Se(coef)	Wald-statistic value ('Z')	Likelihood ratio test	P-value
Proteinuria	Kidney Survival	CoxPH	56	25	0.5193	1.6808	0.4106	1.265	1.58	0.2084
Gout	Kidney Survival	CoxPH	67	33	0.6909	1.9956	0.4668	1.48	1.92	0.1656
Family History of Kidney Disease	Kidney Survival	CoxPH	70	36	-2.91241	0.05434	1.15739	-2.516	3.66	0.05568
Gout Onset <Age 30 years	Kidney Survival	CoxPH Removed those under age 30 if they had not developed gout yet	53	30	0.7038	2.0215	1.0566	0.666	0.37	0.5434
Gout Onset >Age 30 years	Kidney Survival	CoxPH Removed those under age 30 and those who developed gout <age 30	47	29	0.3083	1.3611	0.6230	0.495	0.23	0.6335

Serum Uric Acid Always Normal	Kidney Survival	CoxPH	58	26	0.4945	1.6397	0.4133	1.196	1.43	0.2318
Haematuria	Kidney Survival	CoxPH	56	25	-0.2776	0.7576	0.7446	-0.373	0.15	0.6987
Haematoproteinuria	Kidney Survival	CoxPH	56	25	-0.1572	0.8545	0.7464	-0.211	0.05	0.8297
Male Sex	Kidney Survival	CoxPH	70	36	0.6147	1.8491	0.4301	1.429	1.81	0.1783
Female Sex	Kidney Survival	CoxPH	70	36	-0.6147	0.5408	0.4301	-1.429	1.81	0.1783
Diabetes	Kidney Survival	CoxPH	67	33	-0.4170	0.6590	0.5445	-0.766	0.65	0.4218
Hypertension	Kidney Survival	CoxPH	60	26	-0.9441	0.3890	0.5775	-1.635	2.39	0.1221
Salt Wasting Symptoms	Kidney Survival	CoxPH	59	25	0.7827	2.1873	0.7707	1.016	0.86	0.3537
Abnormal Kidney USS	Kidney Survival	CoxPH	56	31	-0.2359	0.7899	0.3721	-0.634	0.4	0.5275
Kidney Cysts On USS* **	Kidney Survival	CoxPH	56	31	-0.8777	0.4157	0.4282	-2.05	4.66	0.0308*
Non-cystic Kidney Structural Defect On USS	Kidney Survival	CoxPH	56	31	0.5848	1.7946	0.3873	1.51	2.16	0.1419
BMI 26-29	Kidney Survival	CoxPH Removed those with BMI $\geq$ 30	30	19	0.07288	1.07560	0.47844	0.152	0.02	0.8791

BMI ≥ 30	Kidney Survival	CoxPH Removed those with BMI 26-29	34	17	-0.01106	0.98901	0.52644	-0.021	0	0.9832
Smoking History	Kidney Survival	CoxPH	59	26	-0.3621	0.6962	0.4096	-0.884	0.8	0.3715
Recurrent Urinary Tract Infections	Kidney Survival	CoxPH	68	34	0.04299	1.04392	0.48758	0.088	0.01	0.9301
Kidney Stones	Kidney Survival	CoxPH	69	35	-0.7480	0.4733	1.0191	-0.734	0.69	0.4078

Table 11: Clinical Modifiers in the Local ADTKD-*UMOD* Cohort*

*[*A summary of predictor variables collected in the local ADTKD-UMOD cohort and, upon application of a Cox Proportional Hazards model, their association with the trait of kidney survival in this cohort. Statistically significant results are highlighted in grey.*

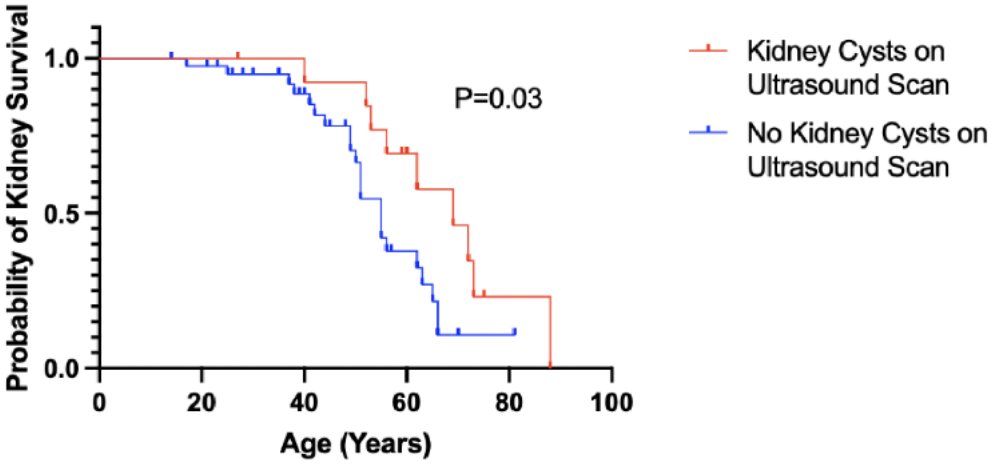

Figure 21: Kidney Survival in those with Kidney Cysts in the Local ADTKD-*UMOD* Cohort*

*[*Kidney survival differences between individuals in the local ADTKD-UMOD cohort who had kidney cysts on ultrasound in comparison to those who did not. When a Cox proportional hazards model was applied, a statistically significant difference in kidney survival was observed in favour of those who had kidney cysts present on ultrasound.]*

2.4 Discussion

This study provides the first detailed characterisation of the UK ADTKD-*UMOD* cohort thanks to integration of both UK RaDaR data ('National Registry of Rare Kidney Diseases (RaDaR)', 2022) **(figure 5)** and deep phenotypic data from a local cohort of ADTKD-*UMOD* patients. This work has allowed the development of new insights into the genetic and clinical landscape of the disease.

2.4.1 The UK ADTKD-UMOD variant spectrum reflects international patterns but is uniquely enriched for a single UMOD p.Val93_Gly97delinsAlaAlaSerCys variant suggesting a potential founder effect

This UK cohort included 141 individuals with Autosomal Dominant Tubulointerstitial Kidney Disease due to *UMOD* variants **(table 1)**. As can be observed in **figure 8**, variants predominantly cluster within exon 3 of the *UMOD* gene reinforcing this region as a variant 'hot spot' with deleterious effects likely due to the functional and structural importance of this domain. This region in exon 3 encodes the Epithelial Growth Factor-like (EGF-like) domains and the D8C domain of the uromodulin protein, critical for proper protein folding and intracellular trafficking (Olinger *et al.*, 2020). The majority of variants are missense, often

involving cysteine residues which echoes the genotypic spectrum in other international cohorts and reinforces the importance of this region for careful disulfide bond formation to maintain its complex tertiary structure **(figure 9).** The proportion and clinical trajectory of in frame deletions **(figures 10 & 11)** echo what is observed in other cohorts internationally (Bollée *et al.*, 2011; Ayasreh *et al.*, 2018; Olinger *et al.*, 2020; Gong *et al.*, 2021; Tanaka *et al.*, 2025). Given the spectrum closely mirrors international datasets, the UK cohort is representative and further supports its utility in broader genotype-phenotype correlation and cross-cohort studies **(table 3)**. In contrast, however, 57% of the cohort have the same pathogenic *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant thought to represent a founder variant in the British population (Valluru *et al.*, 2023). This represents an additional opportunity to study the disease without the confounding effects of primary *UMOD* variant as well as a clinical reminder to consider this as a recurrent monogenic cause of CKD of high prevalence in the British population **(figures 12 & 13).** Phenotypic impact of uromodulin glycine substitutions was explored as they have been shown to influence kidney survival in Alport syndrome (Pagniez *et al.*, 2025). These were found to have a non-significant impact on age of kidney failure or kidney survival in this ADTKD-*UMOD* cohort suggesting that this doesn't have an overt magnified effect on uromodulin folding or that exon location of these is perhaps more important and should be explored next **(figures 14 & 15).**

2.4.2 The UK ADTKD-UMOD cohort reveals lower genotype reporting and significantly later gout onset than the current published international landscape

Table 3 and figure 18 echo that the UK cohort broadly reflects international trends in end stage kidney failure onset and sex distribution of the disease (Bollée *et al.*, 2011; Ayasreh *et al.*, 2018; Olinger *et al.*, 2020; Gong *et al.*, 2021; Tanaka *et al.*, 2025). The UK cohort does, however, demonstrate lower genotype reporting and later gout onset. Later gout onset is likely to reflect the high proportion of those with the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant in the UK cohort and its lack of association with gout as explored in detail in **Chapter 4**. Overall, these differences highlight the need for improved genetic data integration into the RaDaR registry, improved genetic diagnosis and reporting practises across all UK centres and additional emphasis on disease detection with lack of early onset gout as a key clinical feature **(figure 17)**. To overcome this, greater emphasis should be placed on family history, bland urinary abnormalities, mild chronic kidney disease, and routine genetic testing in unexplained familial kidney disease, even in the absence of gout.

2.4.3 Unreported UMOD variants in this UK cohort show inter- and intra-familial clinical heterogeneity

A subset of individuals in the UK ADTKD-*UMOD* cohort were found to carry previously unreported *UMOD* variants **(table 2)**. Whilst classical ADTKD-*UMOD* is characterised by early-onset gout and marked hyperuricaemia, many of the patients with these unreported variants did not have these hallmark features (Devuyst *et al.*, 2019). Although not a diagnostic necessity, documented gout was often absent or late onset in several cases. Kidney outcomes were similarly variable with some individuals progressing to kidney failure in early adulthood, while others maintain preserved kidney function into older age **(figures 6 & 7).** Furthermore, certain unreported variants affected conserved cysteine residues (e.g. p.Cys137Arg, p.Cys126Trp, p.Arg245Cys) which are commonly implicated in protein misfolding and ER retention (Serafini-Cessi *et al.*, 1993; Moskowitz *et al.*, 2013). Others clustered within the N-terminal region (e.g. p.Gly55Asp and p.Phe56Ile) potentially indicating emerging functional relevance (Olinger *et al.*, 2022) **(figure 8)**. Two individuals with the p.Ser186Ile variant had kidney biopsy findings of interstitial fibrosis, tubular atrophy and glomerulosclerosis supporting kidney involvement even in the absence of classical extra-renal features **(table 2)**. Overall, these findings reiterate substantial clinical heterogeneity among individuals with unreported *UMOD* variants, some which are consistent with pathogenicity and other which require further functional characterisation.

2.4.4 UMOD deletion variants, Gly-substitutions and the UMOD p.Val93_Gly97delinsAlaAlaSerCys variant were not associated with a significantly different kidney survival in comparison to the wider ADTKD-UMOD cohort

As can be seen in the Kaplan-Meier and scatter plots in **figures 10-15**, deletion variants, glycine substitutions (specifically *UMOD* p.Gly55Asp, p.Gly68Val, p.Gly88Asp, p.Gly210Asp) and the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant were not associated with a significantly different kidney survival or age of kidney failure in comparison to the broader ADTKD-*UMOD* cohort, although numbers for the first two variants are both small. Deletion variants are often assumed to be more deleterious because if their potential to induce frame shifts or disturb in-region folding domains, but *UMOD* p.Ser91del is an in-frame deletion making it less likely to be more deleterious (Moskowitz *et al.*, 2013). Furthermore, glycine substitutions in Alport syndrome in collagenous domains are typically associated with more severe disease

phenotypes due to the exaggerated disruption of critical structural motifs because of glycine's particular role due to its small size and the strict triple-helical requirement of the organised protein. These particular amino acid substitutions in the *UMOD* gene may not disrupt folding or function severely enough to accelerate disease progression, or they may fall in less structurally constrained regions. Additionally, modifiers and cohort characteristics may dilute detectable clinical differences. Interestingly, as presented in **table 2**, those with the *UMOD* p.Gly55Asp substitution have a much milder kidney phenotype, which may be due to its location outside highly constrained structural domains, limited impact on protein folding, and the potential for partial protein trafficking (Serafini-Cessi *et al.*, 1993; Moskowitz *et al.*, 2013). These factors, along with possible protective genetic modifiers, may explain the relatively attenuated phenotype observed in this subgroup which needs to be explored further functionally. It was also previously described that those with the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant had milder disease. However, in ADTKD-*UMOD*, these variant classes did not confer a worse, or better, prognosis suggesting that disease progression is not solely dictated by variant type and is likely influenced by other genetic, clinical and environmental disease modifiers (Kidd *et al.*, 2020). Taken together, these findings support the pathogenicity of such variants and further describe the phenotypic spectrum across different *UMOD* variant types.

2.4.5 Updated phenotypic data from the previously reported UMOD p.C120Y homozygous family reinforce gene dosage effects and suggest a more deleterious impact of this variant than originally appreciated

The proband of the family reported in Edwards *et al,* 2017, now aged 52, remains the only known homozygote in the UK ADTKD-*UMOD* cohort **(figure 16)**. She demonstrates progressive CKD (eGFR 31 ml/min/1.73 m²) and early-onset gout (first attack at age 31), with markedly elevated serum uric acid (614 µmol/L) and a very low fractional excretion of uric acid (0.026) **(table 4)**. These findings are consistent with the predicted toxic gain-of-function and intracellular retention of mutant uromodulin previously described. Compared to her five heterozygous offspring, aged 23–33, all of whom remain normouricaemic or only mildly hyperuricaemic, with better kidney function (eGFR 69 - 90 ml/min/1.73 m²), no history of gout, and higher fractional excretion of urate, this supports a gene dosage-dependent phenotype. The proband's heterozygous mother, now aged 83, also had onset of gout at 64, consistent with a delayed and milder phenotype. These real-world clinical trajectories provide further

evidence that homozygosity for p.C120Y results in earlier and more severe disease, albeit not dramatically so when compared to some heterozygous cases with classic *UMOD* variants. Furthermore, despite undetectable urinary uromodulin and biochemical evidence of severe tubular dysfunction, the homozygous patient does not exhibit nephrocalcinosis, recurrent UTIs, or kidney stones. This supports the hypothesis that residual mutant uromodulin, though misfolded, may retain partial physiological function. Overall, this extended dataset reiterates the pathogenicity of p.C120Y (originally thought to represent a hypomorphic allele) and adds support to a model in which gene dosage modifies disease severity in ADTKD-*UMOD* (Edwards *et al.*, 2017).

2.4.6 Uric acid handling phenotypes were the strongest clinical distinction observed between those with the UMOD p.Val93_Gly97delinsAlaAlaSerCys variant and those with other UMOD variants

Table 5 details clinical and demographic features of a local cohort of 70 individuals with ADTKD-*UMOD*, stratifying between those individuals with the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant and those with other *UMOD* variants. The sex distribution was similar across groups, with approximately 75–76% female representation. However, those without the *UMOD* p.(Val93_Gly97delinsAlaAlaSerCys) variant displayed greater ethnic diversity, including 37% of British-Pakistani ancestry, compared to 100% White ethnicity among carriers of the *UMOD* p.(Val93_Gly97delinsAlaAlaSerCys) variant adding weight to the possibility it is a founder variant. A positive family history of kidney disease was near-universal in both groups, consistent with the autosomal dominant pattern of inheritance. Gout was reported in a minority (only 13%) of the full cohort, with slightly higher prevalence among those without the *UMOD* p.(Val93_Gly97delinsAlaAlaSerCys) variant, who also accounted for most cases of gout onset before the age of 30 reinforcing the likelihood that gout is not a prominent feature in those with the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant. Further to this point, 44% of individuals had normal serum uric acid on initial measurement, and only 10% had levels exceeding 500 µmol/L, suggesting variable uric acid phenotypes across genotypes. Proteinuria and haematuria were generally uncommon as expected. Hypertension was common across the cohort (40–44%). Other features such as recurrent urinary tract infections and kidney stones were rare. These findings highlight the phenotypic heterogeneity of ADTKD-*UMOD*, with *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant and non-*UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant carriers demonstrating

overlapping but distinct clinical profiles. Those without the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant seem to have earlier-onset gout and milder hypertensive phenotypes possibly due to a more exaggerated salt-wasting effect from the variant, but this is speculative only. A family history of gout was more commonly reported among those without the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant in both females (86% vs. 31%) and males (67% vs. 50%), reinforcing potential variant-specific patterns of hyperuricemia.

2.4.7 A high burden of pregnancy complications existed in females with ADTKD-UMOD

Pregnancy complications have had limited study in ADTKD-*UMOD* and the disease is thought not to confer an increased risk of clinical complications (Bleyer *et al.*, 2023). Among the 14 females represented in the subgroup in **table 6** who had experienced pregnancy, 8 (57%) reported complications, suggesting that reproductive outcomes may be an underappreciated aspect of the disease and should be urgently explored further. Reported complications included a wide range of significant obstetric and perinatal events such as recurrent miscarriage, pre-eclampsia, polyhydramnios, gestational diabetes, pre-term labour, hypotension, pyelonephritis, and neonatal loss due to brain oedema. One individual experienced multiple ectopic pregnancies, and another reported deep vein thrombosis (DVT) during pregnancy. These findings raise the possibility of increased susceptibility to vascular and renal stress during gestation in females with ADTKD-*UMOD*, which may be exacerbated by pre-existing subclinical tubular dysfunction or impaired salt and water handling. However, these complications may relate to oestrogen-driven upregulation of the *UMOD* gene. As uromodulin is now recognised as an oestrogen-responsive protein, rising oestrogen levels during pregnancy could exacerbate the production of misfolded mutant uromodulin, increasing tubular stress and contributing to renal or systemic complications during gestation (Intapad, 2023). Moreover, systemic uromodulin plays a role in immune modulation and kidney homeostasis, and its dysregulation in ADTKD-*UMOD* may further compromise maternal adaptation to pregnancy (Säemann *et al.*, 2005; Nanamatsu *et al.*, 2024). Given the frequency and severity of pregnancy outcomes here, this needs to now be studied in a larger cohort for validation purposes. If an increase in adverse pregnancy outcomes is confirmed, proactive preconception counselling, multidisciplinary antenatal care, and close monitoring of blood pressure and renal function during pregnancy should be provided in this population. These

findings highlight an important, often overlooked aspect of ADTKD-*UMOD* in women and support the need for further research into reproductive health in this population.

2.4.8 Salt-wasting symptoms vary by sex and UMOD variant type in ADTKD-UMOD

Salt-wasting symptoms were frequently reported in this patient population, particularly among females and individuals without the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant, although the sex and *UMOD* variant group difference was not statistically significant **(tables 6-8)**. The most common symptoms included perceived polyuria (55%), nocturia (50%), and urinary frequency (45%). Those without the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant reported higher frequencies of most salt-wasting symptoms compared to those with the variant, especially among females but this was not statistically significant for either variant or sex when a Fisher's exact test was applied. Fainting or postural symptoms, low blood pressure, and excessive thirst were also more frequent in this group. In contrast, males reported fewer salt-wasting symptoms overall, which may reflect milder expression or underreporting. These findings suggest emerging genotype-phenotype and sex-specific differences in symptom burden in ADTKD-*UMOD*, with implications for personalised symptom management, fluid and salt intake guidance. In a subset of 27 individuals with ADTKD-*UMOD* **(table 7)***,* biochemical parameters were assessed alongside self-reported salt-wasting symptoms to explore potential correlations. While 22% of participants reported five or more salt-wasting symptoms, there was no statistically significant association between symptom burden and objective biochemical markers, including urine osmolality, urine sodium, and fractional excretion of urate (FEurate). Most individuals demonstrated preserved urinary concentrating ability, with 88% showing urine osmolality >300 mOsm/kg, and none meeting the threshold for hypoosmolality (serum osmolality <275 mOsm/kg). Urine sodium was >50 mmol/L in 72%, yet this did not correlate with the number of symptoms reported. Similarly, low FEurate (<0.05), a known feature of ADTKD-*UMOD*, was observed in only 36% of individuals and did not associate with symptom burden. These findings suggest a mismatch between patient-reported symptoms and 'spot' biochemical testing, raising the possibility that biochemical measurements were too confounded by other factors in this study and cannot be captured in single 'spot' measurements rather than the possibility that tubular dysfunction is intermittent or an inaccuracy of patient-reported symptoms (Hew-Butler *et al.*, 2018). In conclusion, this analysis emphasises the challenge of defining biochemical salt-wasting in ADTKD-*UMOD* using static parameters and that biochemical salt-wasting may be better informed using dietary-

intake diaries in tandem with 24-hour urine collections in those with preserved kidney function.

2.4.9 A delayed diagnosis is common despite a predictable age of kidney failure for those with ADTKD-UMOD in the UK

As observed in **figure 17**, the timing of diagnosis and kidney failure onset in the UK ADTKD-*UMOD* cohort reveals important insights into disease recognition and progression. The mean age at diagnosis was 42.3 years, with a striking range of 75 years, reflecting substantial variability in when individuals are diagnosed with ADTKD-*UMOD*. This wide range probably reflects a combination of factors, including differences in access to or confidence with genetic testing, clinical awareness among healthcare providers, and variable presentation within families. In contrast, the age of kidney failure onset was more tightly clustered, with a median of 50 years, like other international cohorts yet still with a significant range **(figure 18)** (Bollée *et al.*, 2011; Ayasreh *et al.*, 2018; Olinger *et al.*, 2020; Gong *et al.*, 2021; Tanaka *et al.*, 2025). Most individuals progressed to kidney failure in mid-adulthood, typically between the late 30s and early 60s. The disparity between age at diagnosis and age at kidney failure highlights a diagnostic delay in many cases, potentially spanning years or even decades, during which earlier intervention, counselling, and family screening could have occurred. Given median age of diagnosis was only 7.3 years less than median age of kidney failure, a significant missed opportunity for the aforementioned interventions in addition to education and support is evident. These findings reiterate the urgent need to raise clinical awareness of ADTKD-*UMOD*, to enable timely treatment, screening, counselling, education and support.

2.4.10 Renal ultrasound findings are variable and not clearly genotype specific in ADTKD-UMOD

Limited characterisation of architectural kidney phenotypes have been defined in the literature of ADTKD-*UMOD* to date (Wolf *et al.*, 2009; Ayasreh *et al.*, 2018). Kidney ultrasound findings in the UK ADTKD-*UMOD* cohort demonstrate heterogeneity, both in the presence and type of abnormalities observed **(figures 19 & 20)**. Among 46 individuals with available imaging data, 50% had abnormal findings, while the remaining half exhibited normal renal ultrasounds, highlighting the limited sensitivity of imaging alone in detecting ADTKD-*UMOD*. Cysts were identified in approximately 30% of cases and were more often multiple rather than solitary, but more prevalent than proportions from the Spanish cohort (Ayasreh *et al.*, 2018). Cyst size

varied from <1.5 cm to >3 cm in a minority of cases, with one individual also presenting with concurrent liver cysts, suggesting that although ADTKD-*UMOD* is not classically a cystic disease, cysts detected by ultrasound may be more common than appreciated. In addition to cysts, structural abnormalities such as lobulated cortical outlines, asymmetric kidney size, and pelvicalyceal dilatation were reported possibly reflecting chronic tubulointerstitial injury and variable anatomical remodelling. When stratified by genotype, no significant differences in imaging findings were observed between individuals with the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant versus individuals with different *UMOD* variants **(table 9)**. Statistical analysis using Fisher's exact test confirmed no significant genotype-imaging associations, but numbers were, indeed, small. These findings suggest that although renal ultrasound may contribute to the clinical suspicion of ADTKD-*UMOD*, it lacks discriminatory power for *UMOD* variant type and is frequently normal. Therefore, ultrasound imaging should be viewed as a supportive tool rather than a defining diagnostic feature, reinforcing the need for genetic testing in at-risk individuals. Furthermore, for those patients with CKD of unclear cause and the presence of some kidney cysts, sequencing of the *UMOD* gene should be undertaken along with other cystic kidney disease genes. Secondly, more detailed imaging modalities should be studied such as MRI and contrast-enhanced CT to better characterise subtle or atypical renal abnormalities that may be under-recognised on standard ultrasound. This approach could improve phenotypic classification, support earlier recognition of ADTKD-*UMOD*, and refine our understanding of genotype-imaging correlations.

2.4.11 Additional non-specific biopsy features to classical findings reiterate the need for genetic diagnosis in ADTKD-UMOD

Kidney biopsies from individuals with ADTKD-*UMOD* consistently showed features of chronic interstitial fibrosis and tubular atrophy and, in some cases, mononuclear infiltrates **(table 10)**. Glomerular changes were generally mild, though one patient showed features suggestive of Focal Segmental Glomerulosclerosis, a recognised presentation of the disease particularly in its advanced stages (Chun *et al.*, 2020). Red cell casts supported glomerular-origin haematuria in some cases, and vascular changes (arteriolar hyalinosis, intimal fibrosis) were common, reflecting chronic ischaemia and could certainly lead clinicians to overlook a possible diagnosis of ADTKD as it is often misdiagnosed as 'hypertensive nephrosclerosis', 'IgA nephropathy', or 'benign familial haematuria', suggesting the importance of immunofluorescence, electron-microscopy and genetic testing in these specific cases. Immunofluorescence was mostly non-

specific, and electron microscopy was unremarkable. In one case, dense fibrillary material in the medulla was present and was considered to be characteristic of chronic inflammatory change by the diagnosing histopathologist. Overall, biopsies showed non-specific but characteristic features of ADTKD-*UMOD*, reinforcing the importance of the clinical and genetic context for diagnosis as additional biopsy findings could potentially lead clinicians astray (Onoe *et al.*, 2021).

2.4.12 Clinical features are of limited prognostic utility in ADTKD-UMOD, but this conclusion requires validation within a larger cohort

The Cox proportional hazards analysis of clinical predictors of kidney survival in individuals with ADTKD-*UMOD* revealed that the most commonly assessed features in the clinical environment, such as gout, proteinuria, haematuria, hypertension, salt-wasting symptoms, and sex, were not significantly associated with kidney survival **(table 11)**. This suggests that while these variables may be part of the broader phenotypic spectrum, they are not reliable independent predictors of disease progression in this cohort based on this analysis. Interestingly, the only variable to reach statistical significance was the presence of kidney cysts on ultrasound, which was associated with a reduced risk of progression to kidney failure (HR = 0.42, P = 0.0308) **(figure 21)**. While this finding appears counterintuitive, given that cysts are typically viewed as markers of structural renal damage, it may reflect diagnostic or surveillance bias, whereby individuals with milder disease and incidental cysts are more likely to undergo imaging, or may be captured, screened or genetically diagnosed earlier in their disease course. Alternatively, this could suggest a distinct cystic phenotype with slower progression in a subset of patients, but this is less likely, though further validation in larger cohorts is needed. Taken together, this analysis highlights the limitations of relying solely on conventional clinical features to stratify risk. While Cox proportional hazards modelling is widely used, it has some limitations here (Abd ElHafeez *et al.*, 2021). It reduces longitudinal kidney function data into a binary time-to-event outcome, limiting sensitivity to more gradual or early differences in disease progression. The proportional hazards assumption may not hold in this condition as relative risk of progression may not remain constant over time in 'the real-world', and small, related cohorts such as this one pose challenges due to non-independence of observations. In addition, substantial censoring can reduce model robustness. Alternative approaches, such as eGFR slope modelling using linear mixed-effects models, may better capture individual

variability and allow more nuanced assessment of genotype-phenotype correlations and this can be done when more eGFR data becomes available in time.

2.5 Conclusion

These findings demonstrate the considerable clinical and genetic heterogeneity of ADTKD-*UMOD* in the UK as seen internationally and, most importantly, expose key shortcomings in existing diagnostic and prognostic strategies. Despite a relatively consistent median age of kidney failure, many individuals remain undiagnosed until late in the disease course, reflecting both under-recognition and the challenges posed by delayed onset of hallmark features such as gout. The unique enrichment of the p.Val93_Gly97delinsAlaAlaSerCys variant suggests a potential founder effect and reinforces the need for UK and region-specific genetic awareness. Traditional clinical features including gout, salt-wasting, urine sediment, ultrasound imaging, and biopsy findings showed poor correlation with kidney survival and limited genotype specificity. In summary, these data support a need to shift toward earlier, genotype-led diagnosis and suggest that molecular tools, such as variant classification and new prognostication methods, may offer greater value for risk stratification and personalised care than conventional phenotypic markers alone. Fundamentally, the exploration of genetic modifiers, polygenic risk and biomarkers to aid prognostication is imperative and will now be explored in the next chapters.

The findings from Chapter 2 highlight the considerable heterogeneity in clinical outcomes among individuals with ADTKD-*UMOD*, even in the context of shared pathogenic variants. This variability, particularly evident within families carrying the recurrent *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant, strongly suggests the presence of additional genetic or regulatory modifiers that influence disease trajectory. In Chapter 3, this hypothesis is explored through a series of genome-wide and gene-level analyses designed to identify variants associated with kidney survival in ADTKD-*UMOD*. These approaches aim to uncover novel biological pathways underlying disease progression and to refine our understanding of how polygenic and regulatory architecture may shape outcomes in what is traditionally considered a Mendelian disease.

Chapter 3. Genetic Modifiers of Disease Progression in ADTKD-*UMOD*

3.1 Introduction

As described in **Section 1.7**, monogenic inheritance alone does not explain the significant variability in disease progression and age of onset of end stage kidney failure within and between families with ADTKD-*UMOD*. Factors such as male gender, younger age of gout onset, maternal age of onset of ESKD, primary *UMOD* variant and genotype of the *UMOD* promotor variant rs4293393 have been shown to be associated with variability of disease progression in ADTKD-*UMOD* in one published cohort but contribution of genetic modifiers remain mostly unestablished and unexplored (Kidd *et al.*, 2020).

Regarding genetic factors, in a study of 147 ADTKD-*UMOD* families, the minor G allele in the *UMOD* promotor at rs4293393, which reduces uromodulin expression, was shown to have a protective effect whereby disease progression was slower. However, this allele was significantly under-represented (11%) in affected families included in the study, was in cis with the wild-type allele in 17% of cases, and, because of its non-random distribution, a Mendelian randomisation study could not be conducted (Kidd *et al.*, 2020). Given the high proportion of individuals with a *UMOD* promotor variant that increases the quantity of expressed uromodulin, this is thought, when coinciding with the pathogenic disease-causing variant,, to worsen the disease phenotype and may contribute to disease heterogeneity (Trudu *et al.*, 2013; Ghirotto *et al.*, 2016). *UMOD* variants occurring in the epidermal growth factor domains 2 and 3 were found to have an earlier onset of kidney failure (range 45-52 years) in comparison to *UMOD* variants in the cysteine-rich domains (range 60-65 years) (Moskowitz *et al.*, 2013). However, a different study found only a modest non-significant effect of primary *UMOD* variants on kidney survival (Bollée *et al.*, 2011). Further work has shown that an *in vitro* score rather than the class of amino acid substitution influences kidney survival. Of course, as with any rare disease, all of these studies have limited power to detect statistical differences between groups along with missing clinical and genetic data, suggesting a need for an amalgamation of cohorts and subsequent genetic data internationally to determine the role of genetic modifiers in ADTKD-*UMOD*.

Monogenic kidney diseases demonstrate substantial phenotypic heterogeneity. For instance, in those with ADPKD-*PKD1*, inherited modifiers account for an estimated 18-59% of the

phenotypic variability (Fain *et al.*, 2005). Evidence for modifier genes is extensive in both humans and organisms and can be subtle or profound (e.g. complete suppression of the disease in those who with a monogenic disease-causing variant) (Nadeau, 2001). GWAS is traditionally used to determine loci associated with disease and is typically conducted to include thousands of participants. As explained in **Sections 1.8 and 1.9**, more recently, GWAS has been used to detect modifiers of disease progression and studies have been very successful. The concern of ensuring a properly powered study to detect modifiers depends on the size of effect potential modifiers have on the phenotype, the heritability of the trait, the minor allele frequency, and the rigor of study design. Recently, a GWAS performed in 253 patients with dystrophinopathy detected a strong modifier (*THBS1*) demonstrating that genome-wide significance can be achieved with small numbers (Weiss *et al.*, 2018). The discovery of a rare *APOE* Christchurch variant which delays the onset of Alzheimer's disease in carriers with a highly penetrant *PSEN* variant, originally reported from a single remarkable case, highlights the profound potential of investigating genetic modifiers to unravel biological mechanisms (Quiroz *et al.*, 2024). Furthermore, the recent identification of a protective loss-of-function *APOL1* variant provides compelling evidence supporting *APOL1* inhibition as a therapeutic strategy for chronic kidney disease and illustrates the profound benefit genetic modifier discovery can provide to identify novel therapeutic targets in diseases of high global-burden (Genovese, Friedman, *et al.*, 2010, Gupta *et al.*, 2023, Hung *et al.*, 2023).

Genome-wide association studies have repeatedly identified common variants at the *UMOD-PDILT* locus as among the strongest genetic signals associated with estimated glomerular filtration rate (eGFR) and chronic kidney disease (CKD) in European and East Asian populations (Köttgen *et al.*, 2009). These associations are primarily mediated via increased expression of uromodulin. However, a recent South African study has demonstrated that these same *UMOD-PDILT* SNPs exert markedly attenuated or absent effects on uromodulin levels in individuals of African ancestry, with reduced linkage disequilibrium and distinct haplotype architecture at the locus. Importantly, the 'protective' c-t-a haplotype common in Europeans was almost absent in Black individuals, highlighting the limitations of transposing GWAS findings across ancestries and highlighting the importance of ancestry-specific fine-mapping and functional validation in genetic modifier studies (Strauss-Kruger *et al.*, 2024).

GWAS provides an unbiased, hypothesis-free approach to detect genetic modifiers of disease progression in monogenic diseases such as ADTKD-*UMOD* which have evidence of incomplete penetrance and variable expressivity. Furthermore, genome-wide polygenic scores have emerged as a powerful tool to quantify the contribution of polygenic effects, if indeed contributions to the variability in monogenic disease penetrance is mitigated or exacerbated by the broader range of mechanisms associated with CKD rather than variants that disrupt crucial disease pathways. Given the spectrum of *UMOD* variants and their differing magnitude of effects, from rare Mendelian variants to common, small effect variants that contribute to CKD risk, determining genetic factors at play in ADTKD-*UMOD* may be relevant to all forms of CKD in which fibrosis is a driver of disease progression.

3.2 Methods

3.2.1 ADTKD-UMOD Cohort DNA Acquisition

A data requisition application was made to the UK RaDaR registry to identify hospitals which had high numbers of patients registered with ADTKD-*UMOD*. RaDaR is approved by the Southwest - Central Bristol Research Ethics Committee (REC reference: 19/SW/0173), providing national ethical clearance for patient recruitment, data collection, and optional DNA sampling across participating renal units. Clinicians with high numbers of patients with this condition were contacted for DNA and study contributions. DNA for individuals who had relevant clinical information, and a genetic diagnosis were pooled after Material Transfer Agreements and where local ethical approval was in place. DNA had already been extracted and stored for patients in order to obtain a genetic diagnosis, so an aliquot of this existing store was taken. Collaborators in the USA and Ireland who had large local ADTKD-*UMOD* cohorts were also contacted to contribute. Overall, DNA for 275 individuals with ADTKD-*UMOD* were collected for SNP genotyping.

3.2.2 SNP Array Genotyping

All pooled DNA samples were measured for concentration and purity using a Nanodrop spectrophotometer, ensuring a DNA quantity of 200ng in 4µL was loaded per sample where possible. Samples were aliquoted into Abgene 96-well deep-well plates, sealed with Microseal 'F' PCR Plate Seals, and wrapped in parafilm to prevent seal disruption. Prior to shipment, sample metadata, including a 96-well layout, DNA integrity assessments using an Agarose gel

and a sample sheet for Laboratory Information Management System (LIMS) integration was prepared, checked and submitted to the genotyping company. Plates were shipped on dry ice to the Genomics Core Facility at Erasmus MC for genotyping using the Illumina Infinium Global Screening Array v3.0.

3.2.3 ADTKD-UMOD Cohort Phenotyping

Phenotype data consisted of current age and eGFR (CKD-EPI) if the individual had not yet developed end stage kidney failure or age at which the individual developed end stage kidney failure. End stage kidney failure was defined as, the point at which kidney replacement therapy (regular dialysis or pre-emptive kidney transplantation) was initiated or an eGFR of 10mL/min/1.73m^2. Biological sex was also obtained to fulfil quality control processes outlined in **Section 3.2.4** and primary *UMOD* variant was obtained to allow for covariate adjustment outlined in **Section 3.2.5**. This data was obtained from either the RaDaR registry using the individual's unique RaDaR numerical identifier or via collaborators using unique numerical identifiers and encrypted file transfer protocols, ensuring compliance with data protection regulations. Phenotype and covariate data were manually entered into a curated phenotype text file and checked for accuracy.

3.2.4 Quality Control Analysis

A quality control (QC) pipeline was developed and performed on genotype data using PLINK to ensure data integrity before conducting genetic association studies (Chang *et al.*, 2015). The data consisted of 275 individuals and 725 497 variants at the start of the quality control process.

Initial SNP Quality Control

Rare variants were excluded using a minor allele frequency (MAF) threshold of 0.01, removing 214 446 variants. Genotype call rates were assessed, and SNPs with greater than 3% missing data were excluded resulting in the removal of 2892 variants. Hardy-Weinberg equilibrium (HWE) tests were conducted and SNPs with a HWE $p < 1 \times 10^{-15}$ were removed (22 variants removed).

Individual Quality Control

Sex was updated using phenotype records and a sex check was performed using chromosome-wide heterozygosity. Four individuals had discrepancies between reported and imputed sex and were excluded. Missing data rates and heterozygosity were assessed and individuals (n=4) with more than 3% missingness or more than 3 standard deviations from the mean for heterozygosity were excluded. Identity by descent (IBD) analysis was performed using a pruned SNP set to identify duplicate individuals of which there was one duplicate (one of these individuals was removed). Principal Component Analysis (PCA) was conducted using HapMap reference populations revealing an almost completely European population (Tanaka, 2005). One individual was removed due to divergent ancestry. In total, 9 individuals were removed based on this individual-based QC.

Marker Quality Control

SNPs with a missing call rate of over 5% were excluded (0 variants excluded). Hardy-Weinberg equilibrium was reassessed, and SNPs with HWE $p < 1 \times 10^{-6}$ were removed (57 variants excluded). Minor allele frequency thresholds (0.01) were reapplied removing a further 2247 variants. Following QC, the final dataset consisted of 266 individuals and 498 288 variants.

3.2.5 Determination of Covariates

Effects of the major primary *UMOD* variants within the cohort were originally determined using Cox proportional Hazards regression analysis. As some of these variants were found to have a statistically significant impact on kidney survival within the cohort, primary *UMOD* variant was included as a covariate in the Genome Wide Association Analysis. Ten principal components of relatedness, generated using PLINK within the cohort, were also used as covariates. The primary *UMOD* variant segregated into 4 additional covariates. These were CV11 (*UMOD* p.Val93_Gly97delinsAlaAlaSerCys), CV12 (*UMOD* p.His177-Arg185del), CV13 (*UMOD* p.Cys106Phe) and CV14 (*UMOD* p.Arg178Pro).

3.2.6 Genome-wide Association Analysis

A genome-wide association study was performed in two cohorts, the 266-person cohort of all individuals with ADTKD-*UMOD* and the 92-person cohort of individuals with the same primary *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant in order to eliminate the effect of the primary *UMOD* variant on kidney survival.

A survival-based association R package (SPACox) was used to assess genetic variants associated with the phenotype of kidney survival in these two cohorts separately (Bi *et al.*, 2020). Phenotypic data and 10 principal components of relatedness as covariates were extracted from the phenotype file and matched to genotype data stored in PLINK binary format (with the additional consideration of 4 primary *UMOD* variant covariates in the full 266-person cohort). A null Cox proportional hazards model was generated using the SPACox_Null_model function from the SPACox package, with individual IDs aligned between phenotype and genotype files.

SNP association testing was conducted using the SPACox.plink function, generating results in an output file. To assess genomic inflation, chi-square test statistics were calculated and adjusted using the genomic control factor. Adjusted chi-square values and genomic control corrected p-values were derived (Devlin and Roeder, 1999).

Quantile-quantile (QQ) plots were created to compare expected vs observed chi-squared statistics before and after genomic control adjustment. A Manhattan plot was generated using the mahattan() function from the qqman R package to highlight genome-wide and suggestive loci associating with the trait.

Given that no loci were of genome-wide significance ($p < 5 \times 10^{-8}$), a bespoke p-value threshold was determined to capture potentially relevant SNPs approaching genome-wide significance. These SNPs are listed, detailed and characterized in **Appendix A**. These SNPs were subsetted and stored for further downstream analysis outlined in **Section 3.2.8**.

3.2.7 Imputation

Pre-imputation data preparation (for the 266-person full ADTKD-*UMOD* cohort) was performed on the Newcastle University HPC system and an interactive node was allocated. PLINK 2.0 was loaded to process genotype data. Allele frequencies for the cohort were calculated which is crucial for imputation. The annotation file for the Haplotype Reference Consortium r1.1 reference panel was downloaded from the Sanger Institute based on GRCh37 (hg19) coordinates (Loh *et al.*, 2016). Additionally, the HRC-1000G-check-bim tool was downloaded and extracted to validate genotype data. A Perl script was executed to check and clean the PLINK .bim file to ensure consistence with the reference panel. Each chromosome

was converted to Variant Call Format (VCF) using PLINK 2.0 which incorporated the updated allele definitions and was repeated for chromosomes 1-23 to ensure compatibility with the Michigan Imputation Server (MIS). The VCF files were compressed for storage and transfer to the MIS for imputation. After imputation on the MIS, files were downloaded to the HPC. Each chromosome was extracted and quality control was performed based on a minor allele frequency (MAF) cutoff of ≥ 0.01 and imputation quality score (Rsq) of ≥ 0.5. PLINK 1.9 was used to extract high confidence SNPs and duplicated SNPs were identified and removed using R and PLINK. The final cleaned and imputed data was merged across all chromosomes using PLINK ready for downstream analysis.

For the 92-person *UMOD* p.Val93_Gly97delinsAlaAlaSerCys cohort, individuals without this primary *UMOD* variant were identified within the full cohort, listed in a text file and excluded using PLINK. Variants with a MAF <0.01 were removed to retain common variants only for downstream analysis.

3.2.8 SNP Prioritisation for Further Investigation

All four genome-wide association analysis output files were uploaded to the Functional Mapping and Annotation of genome-wide association studies (FUMA GWAS) server and analysed separately (Watanabe *et al.*, 2017; 'FUMA GWAS', 2024). GWAS results were functionally annotated using the FUMA SNP2GENE package. Standard FUMA settings were selected to process summary statistics. Pre-selected SNPs that were approaching genome-wide significance from the GWAS were specified for further analysis. LocusZoom plots were generated to visualise regional association signals for each SNP (Boughton *et al.*, 2021). SNPs without LocusZoom support were eliminated from further consideration and SNPs with LocusZoom support were retained for further prioritisation. Each SNP was assigned a prioritisation score based on multiple functional and genetic annotation criteria **(figure 22)**. SNPs were then ranked based on their prioritisation scores. The highest scoring SNPs were considered 'key candidate SNPs' to be considered for further biological prioritisation.

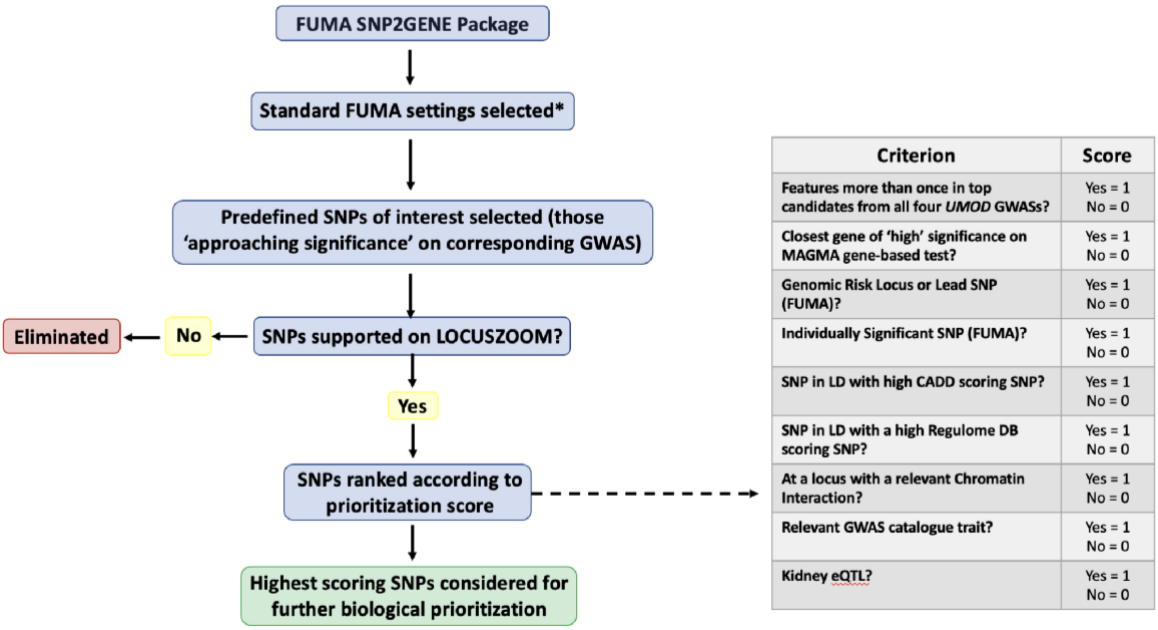

Figure 22: SNP Prioritisation Method

3.2.9 Gene-burden Analysis

All four GWAS output files were uploaded and analysed as described in **Section 3.2.6**. MAGMA gene-burden analysis was performed within Functional Mapping and Annotation of genome-wide association studies (FUMA GWAS) to identify genes associated with the trait of kidney survival (Guo *et al.*, 2016; 'FUMA GWAS', 2024). SNPs were mapped to genes based on a +/- 10kb window surrounding each gene boundary. MAGMA performed a multi-SNP regression test to assess the combined effect of SNPs within each gene on the trait and a Bonferroni correction to correct for multiple testing. In total, 19 997 genes were included and a genome-wide significance threshold of $-\log_{10}(p) > 5.6$; $p < 2.5 \times 10^{-6}$ was applied. Genes approaching genome-wide significance are listed, detailed and characterized in **Appendix B**

3.2.10 Gene Prioritisation for Further Investigation

The R software was used to extract top-ranking genes from the MAGMA output file for further downstream analysis (Dessau and Pipper, 2008). Top ranking genes were eliminated if the gene product protein or mRNA was not expressed in the kidney using Human Protein Atlas as a reference (Thul and Lindskog, 2018; 'Human Protein Atlas', 2024). Each gene was then assigned a prioritisation score based on multiple functional and genetic annotation criteria **(figure 23)**. Genes were then ranked based on prioritisation scores. The highest scoring genes were selected for further biological prioritisation and were considered 'key candidate genes'.

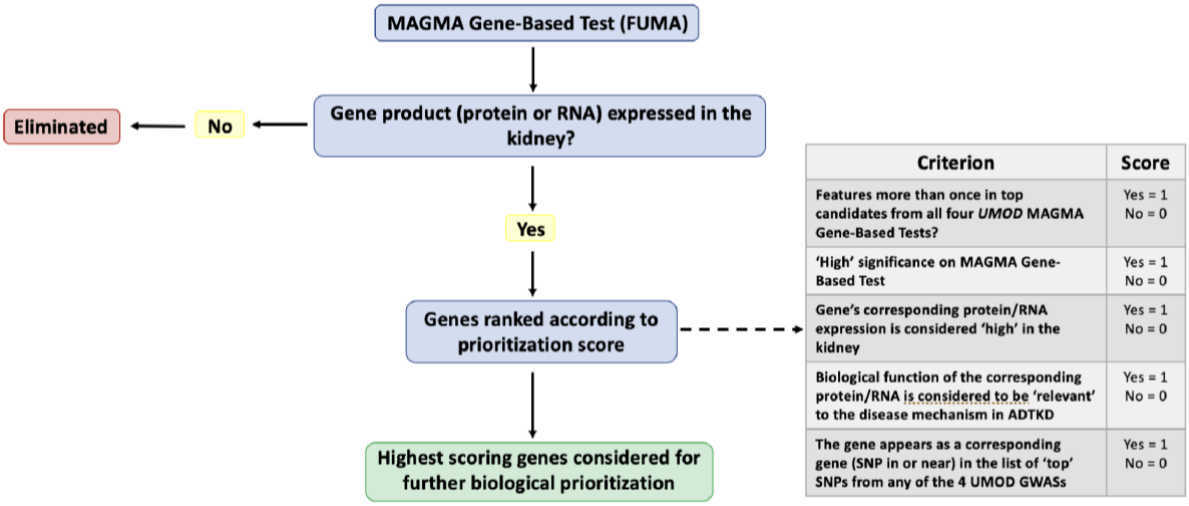

Figure 23: Gene Prioritisation Method

3.2.11 Conditional Analysis

GCTA-COJO was used initially to perform conditional analysis with the intention of identifying independent SNP associations with the trait of kidney survival whist conditioning on lead variants (Yang *et al.*, 2012). Summary statistics from the full 266-person imputed GWAS cohort were entered in R. The standard error was calculated from the square root of the Variance. Then Beta (the effect size) was calculated from the product of the Z score and standard error. A constant sample size of 266 was assigned for each SNP. The data was then formatted to meet GCTA-COJO requirements and saved in a tab-delimited file. Genotype data was prepared by extracting all autosomes from the original GWAS binary file using PLINK. Conditional analysis was then conducted using the two datasets using GCTA-COJO, conditioning on four lead SNPs individually. A p-value threshold of 5×10^{-5} was set for SNP inclusion.

Due to uncertainty in the reliability of results when using GCTA-COJO explained later in **Section 3.4.6**, an alternative method using PLINK for conditional analysis was subsequently adopted. Each SNP of interest was listed in a text file and allele dosage of each of the 4 SNPs of interest was determined for each individual within the full 266-person imputed GWAS dataset using the "--recode A" command in PLINK. The GWAS phenotype file including all previously used covariates was read into R and the allele dosages for each of the four SNPs of interest in each individual were merged by IID as four additional covariates and saved as a tab-delimited file. A subset of chromosome 19 was extracted from the genotype file using PLINK and the SPACox null model was re-run including the individual allele dosages as covariates for each of the four

SNPs of interest. The p-values for the other three SNPs of interest were compared before and after conditioning on the fourth SNP to determine the extent each association is dependent on the lead SNP tested. This process was repeated with each of the four SNPs of interest acting as the 'lead variant' in turn.

3.2.12 Chromatin-landscape Interaction Mapping

Chromatin interaction mapping for top candidate SNPs and genes was carried out within FUMA (Watanabe *et al.*, 2017; 'FUMA GWAS', 2024). Significant chromatin interactions were defined by a False Discovery Rate (FDR) <0.05, and Hi-C and promoter-capture Hi-C datasets from multiple tissues, including kidney, were used (Belton *et al.*, 2012). FUMA generated Circos plots to visualise these chromatin interactions.

To validate these findings for particular SNPs of interest, further analysis was performed in UCSC using data from the 1000 Genomes Project (phase 3, GBR population). Pairwise correlations between the SNPs of interest were analysed by looking at Linkage Disequilibrium patterns in UCSC (Nassar *et al.*, 2023; 'UCSC Genome Browser', 2024). Publicly available RNA-seq datasets (GTEx Portal and Human Protein Atlas) were analysed to examine gene expression patterns across tissues (GTEx Consortium, 2013; 'GTEx Portal', 2024; 'Human Protein Atlas', 2024). Epigenomic ChIP-seq data for H3K4me3 and H3K27ac histone modifications were analysed to determine active enhancers and promotors in proximity to SNPs of interest using the Kidney Precision Medicine Project ('Kidney Precision Medicine Project', 2024). Chromatin accessibility was determined by analysing ATAC-seq data using the Kidney Precision Medicine Project (Ong *et al.*, 2020). To predict enhancer-target gene interactions, the Activity-by-Contact (ABC) model was used and eQTL analysis was performed in The Open Targets Genetics Portal to identify expression quantitative trait loci (eQTL) for each SNP of interest (Ghoussaini *et al.*, 2021; Hecker *et al.*, 2023; 'Open Targets Genetics', 2024).

3.2.13 Polygenic Risk Score Application

A polygenic risk score for CKD was downloaded from the authors' webpage (Khan *et al.*, 2022). The downloaded PRS SNPs were matched against the cohort's imputed genotype data by comparing variant IDs and allele codes. Where variant IDs did not match, SNPs were matched according to chromosome and base pair position. In total, 6583 SNPs of the 471 316 SNPs in

the PRS were missing from the genotype dataset and these were excluded. Amongst the remaining matching variants, 35298 required an updated variant ID and 42130 required the effect allele to be forced due to minor allele assumption discrepancies. All updated SNP IDs and allele codes were recorded in a text file and PLINK was used to update genotype files based on this updated information. 463569 valid predictors were included in the final PRS application. The --score command was used in PLINK to apply the PRS variant effect sizes to the genotype data for every individual, resulting in an individual score for all 266 cohort participants. A Cox regression model was applied to test association between PRS and kidney survival in the cohort, including adjustment for the 14 covariates (10 principal components of relatedness and 4 primary *UMOD* variant categories). Polygenic scores for the 92-person imputed cohort with the same *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant were tested for association with kidney survival using the same Cox regression model, including covariate adjustment for 10 principal components of relatedness.

3.2.14 Polygenic Risk Score Variation

To assess the specific regional contribution of the *UMOD* gene and its promotor to the PRS in the full 266-person imputed cohort, PRS SNPs were matched with the study's genotype dataset by comparing variant IDs and allele codes between the two files (464 733 SNPs). These were extracted and allele frequencies were calculated using the --freq command in PLINK. From the genotype file for the cohort, SNPs located in the *UMOD* gene and its promotor were extracted and copied to a text file. A merged dataset was created including PRS SNPs, allele frequencies and *UMOD*-specific SNPs. For both *UMOD* and non-*UMOD* SNPs, variance calculations were performed using the expected mean SNP score, expected squared SNP score and variance of the SNP score for (a) the full set of SNPs contributing to the score and (b) SNPs within the *UMOD* gene and its promotor.

3.2.15 Haplotype Estimation of the UMOD p.Val93_Gly97delinsAlaAlaSerCys variant

This work was kindly performed by Dr Fabienne Jabot-Hanin (Paris Cite University / Imagine Institute, France) through research collaboration. Files for the 92 patients with the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant were transferred securely in compliance with Newcastle University's data sharing policy and agreement. The genotyped SNPs on chromosome 16 were phased with SHAPEIT2 software (Delaneau *et al.*, 2013) and used to define a haplotype shared by all patients. This common haplotype and the flanking SNPs were

then used to estimate the age of the most recent common ancestor using ESTIAGE software (Genin *et al.*, 2004) which implements a likelihood-based method. Allele frequencies were taken from the gnomAD database and genetic distances (cM) were obtained from the 1000 Genomes Phase 3 data; positions absent from this map were interpolated. A variant rate from 10^{-4} to 10^{-6} at each marker were considered in the model.

3.3 Results

This chapter presents a comprehensive series of genetic analyses undertaken to identify potential modifiers of disease progression in ADTKD-*UMOD*, with kidney survival as the primary phenotype of interest. The results are presented sequentially, beginning in **section 3.3.1** with a detailed overview of the GWAS cohort, including genotype distributions, geographical representation, and kidney survival stratified by variant and sex. **Sections 3.3.2** to **3.3.5** describe the genome-wide association study (GWAS) findings across the full cohort and within the subgroup carrying the common *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant, both with and without genotype imputation. **Section 3.3.6** highlights specific top-ranked SNPs for further exploration, with survival plots and regional association visualisations. In **section 3.3.7**, signals within the *UMOD* promoter and *PDILT* loci are interrogated, while **sections 3.3.8** to **3.3.11** present results from gene-burden analyses under the same cohort and subgroup structure as the GWAS. **Section 3.3.12** then prioritises top genes from these burden tests. **Section 3.3.13** includes conditional analyses at the top chromosome 19 locus to further define signal independence. Chromatin interaction mapping is presented in **section 3.3.14** to explore potential regulatory interactions for identified loci. In **sections 3.3.15** and **3.3.16**, polygenic risk scores (PRS) for chronic kidney disease are applied to the cohort to assess their predictive value for kidney survival. Finally, **section 3.3.17** investigates haplotype structure surrounding the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant to explore a potential founder effect. As in the preceding chapter, each section presents relevant figures and tables in a descriptive manner. Interpretation, context, and implications of these findings are reserved for the discussion section that follows **(section 3.4)**.

3.3.1 GWAS Cohort Genotypes

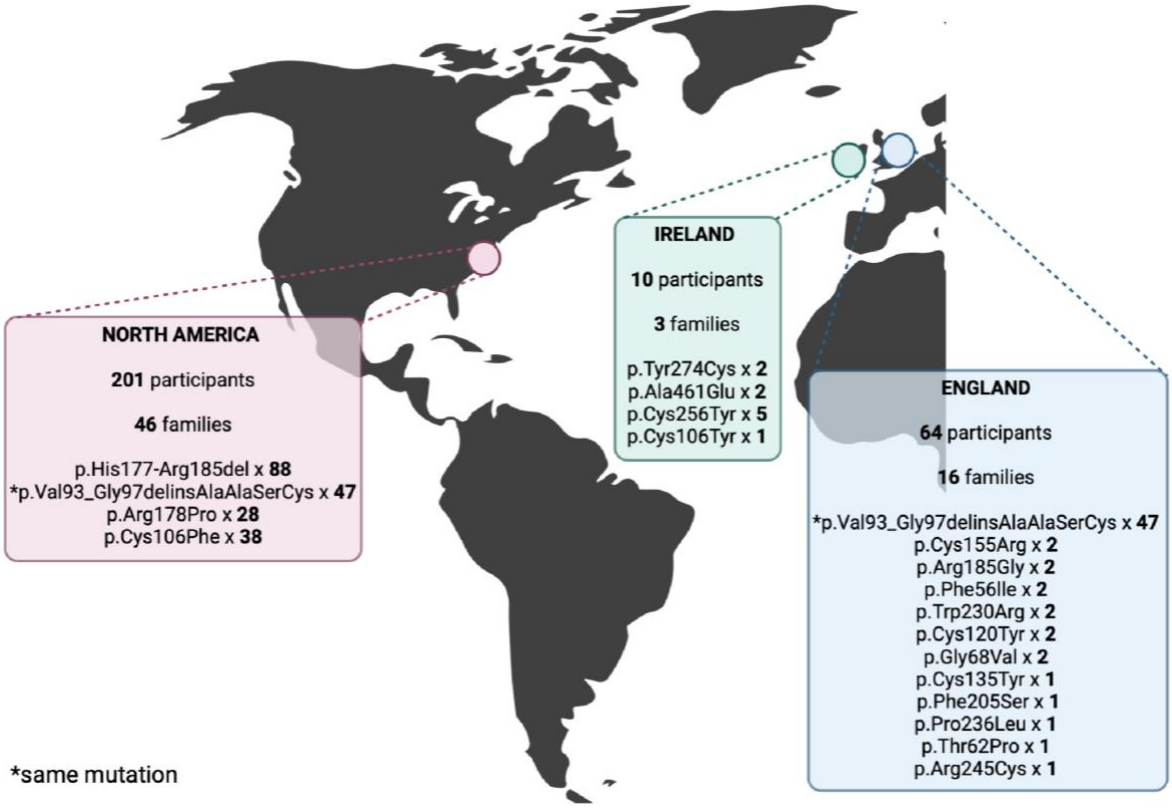

Figure 24: Geographical Source of ADTKD-*UMOD* GWAS Participants*

*[*A visual representation of the Genome Wide Association Study (GWAS) cohort's geographical acquisition. The majority of GWAS participant DNA has come from North American collaborators (via their RedCap registry). The UK participant DNA has mostly come from Newcastle upon Tyne, followed by a small number of participants in Cambridge. Ten further participant DNA samples were provided by collaborators in Dublin, Ireland. Overall, this represents a geographical pattern of probable European decent. Although the UMOD p.Val93_Gly97delinsAlaAlaSerCys variant is believed to be unique to the UK, 47 participants from the North American registry also had the UMOD p.Val93_Gly97delinsAlaAlaSerCys variant.]*

Primary UMOD Variant (n=19)	Number of Participants
p.Val93_Gly97delinsAlaAlaSerCys	94
p.His177-Arg185del	88
p.Cys106Phe	38
p.Arg178Pro	28
p.Cys256Tyr	5
p.Tyr274Cys	2
p.Trp230Arg	2
p.Gly68Val	2
p.Ala461Glu	2
p.Arg185Gly	2
p.Cys155Arg	2
p.Phe56Ile	2
p.Cys120Tyr	(1 homozygote + 1 heterozygote)
p.Thr62Pro	1
p.Pro236Leu	1
p.Phe205Ser	1
p.Cys135Tyr	1
p.Cys106Tyr	1
p.Arg245Cys	1
Total	**275**

Table 12: Primary *UMOD* Variant in GWAS Cohort Participants*

*[*The proportion of GWAS participants with each primary UMOD variant. Most of the GWAS cohort had one of four primary UMOD variants and all were heterozygotes except for one homozygote participant with a consanguineous family history. The UMOD p.Val93_Gly97delinsAlaAlaSerCys variant represented the most abundant variant in this cohort.]*

The following figures show kidney survival curves for each of the four most abundant primary *UMOD* variants featuring in the GWAS cohort. This was analysed with the intention of determining the impact of primary *UMOD* variant on the phenotypic trait to be assessed in the GWAS and if the primary *UMOD* variant category needed to be considered as an important contributor and thus adjusted for as a covariate.

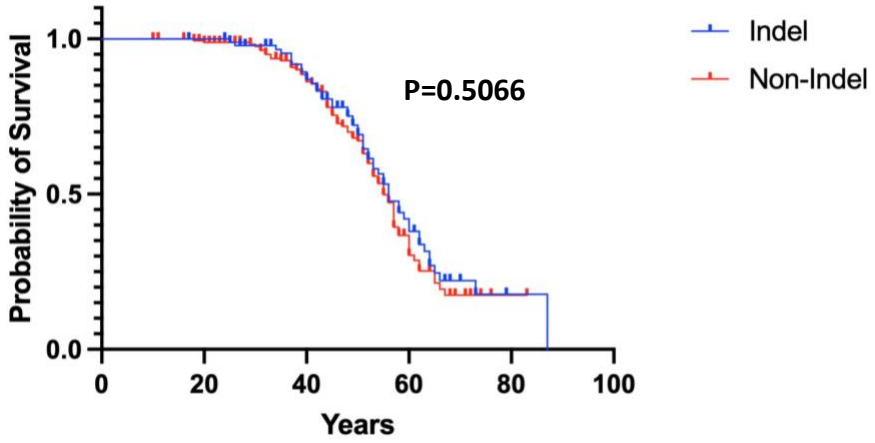

Figure 25: Kidney Survival in GWAS Participant with ADTKD-*UMOD* due to the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys Variant*

*[*Kidney survival differences between GWAS participants with the UMOD p.Val93_Gly97delinsAlaAlaSerCys variant vs. those with the other remaining variants described in **table 12**. When a Cox proportional hazards model was applied, there was no statistically significant difference in kidney survival between those with the UMOD p.Val93_Gly97delinsAlaAlaSerCys variant vs. those with the other remaining variants.]*

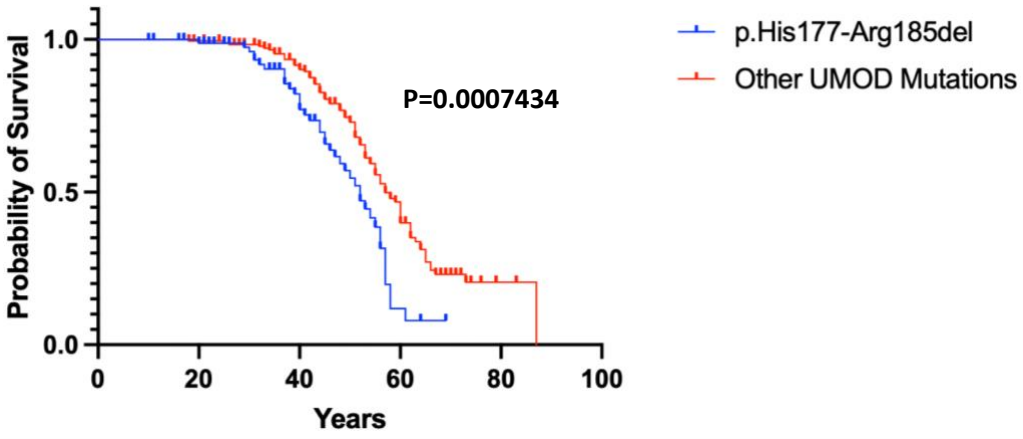

Figure 26: Kidney Survival in GWAS Participant with ADTKD-*UMOD* due to *UMOD* p.His177-Arg185del*

*[*Kidney survival differences between GWAS participants with the UMOD p.His177-Arg185del variant vs. those with the other remaining variants described in **table 12**. When a Cox proportional hazards model was applied, there was a statistically significant difference in kidney survival between those with the UMOD p.His177-Arg185del variant vs. those with the other remaining variants. The group with the UMOD p.His177-Arg185del had a statistically worse kidney survival than the rest of the cohort.]*

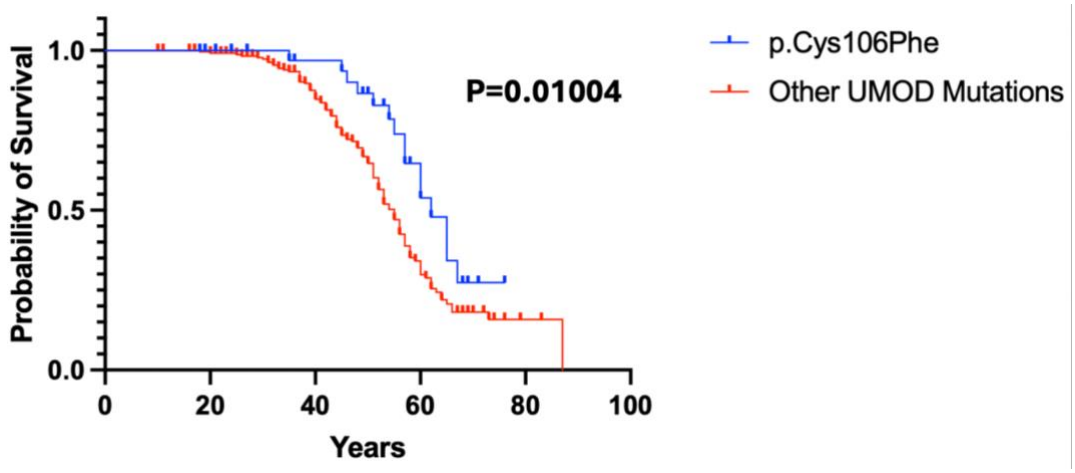

Figure 27: Kidney Survival in GWAS Participant with ADTKD-*UMOD* due to *UMOD* p.Cys106Phe*

*[*Kidney survival differences between GWAS participants with the UMOD p.Cys106Phe variant vs. those with the other remaining variants described in **table 12**. When a Cox proportional hazards model was applied, there was a statistically significant difference in kidney survival between those with the UMOD p.Cys106Phe variant vs. those with the other remaining variants. The group with the UMOD p.Cys106Phe variant had a statistically better kidney survival than the rest of the cohort.]*

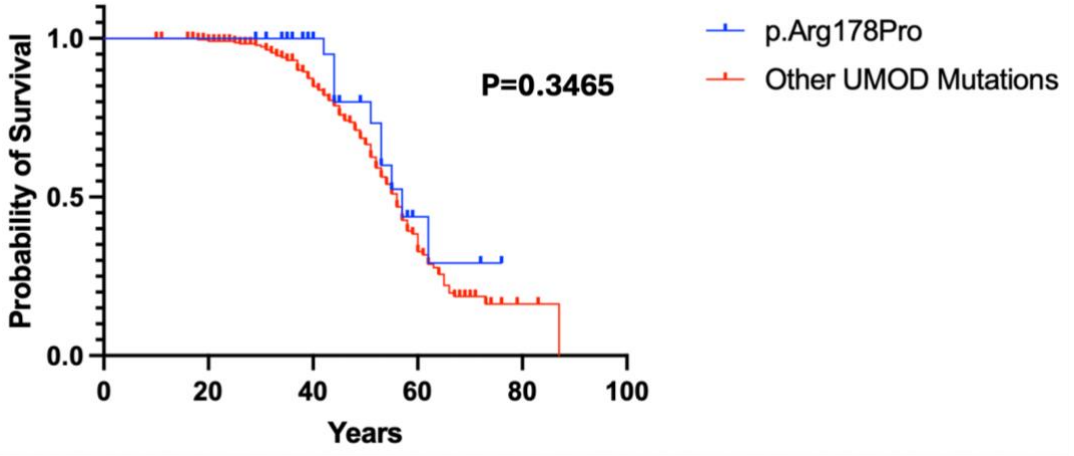

Figure 28: Kidney Survival in GWAS Participant with ADTKD-*UMOD* due to *UMOD* p.Arg178Pro*

*[*Kidney survival differences between GWAS participants with the UMOD p.Arg178Pro variant vs. those with the other remaining variants described in **table 12**. When a Cox proportional hazards model was applied, there was no statistically significant difference in kidney survival*

between those with the UMOD p.Arg178Pro variant vs. those with the other remaining variants.]

Next, the literature in other cohorts suggest that sex could be a modifier of kidney survival in ADTKD-*UMOD*, with males experiencing worse kidney survival outcomes. To assess, if this was the case in this cohort, a Cox proportional hazards model was applied to the full cohort and in a smaller sub-cohort including only participants with the same *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant to eliminate the impact of the primary *UMOD* variant.

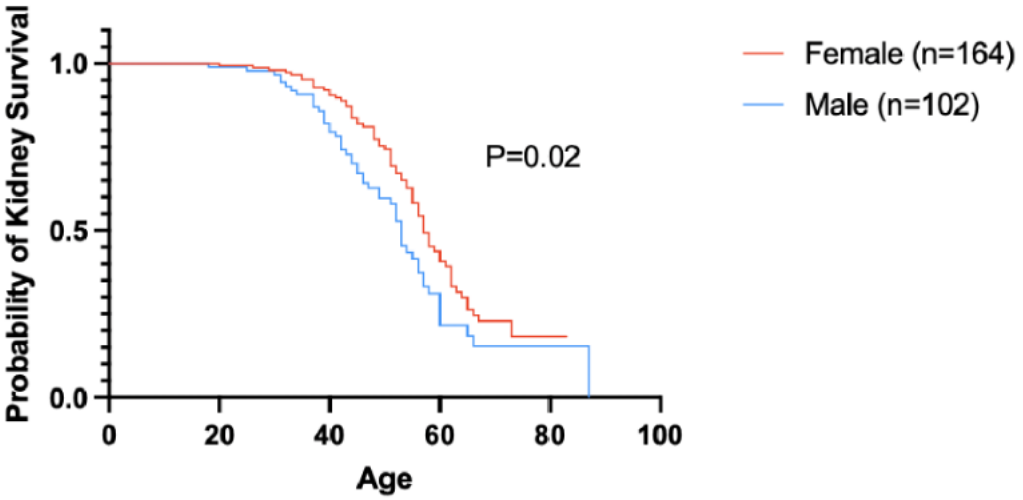

Figure 29: Kidney Survival by Biological Sex in GWAS Participants with ADTKD-*UMOD**

*[*Kidney survival differences between GWAS participants by biological sex. When a Cox proportional hazards model was applied, there was a statistically significant difference in kidney survival between males and females with males having a statistically worse kidney survival.]*

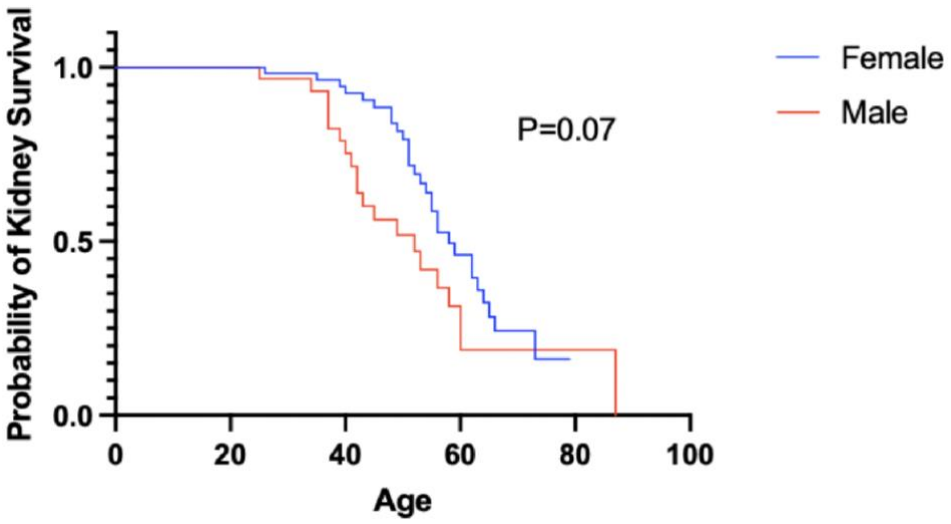

Figure 30: Kidney Survival by Biological Sex in GWAS Participants with the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant*

*[*Kidney survival differences by biological sex between GWAS participants with the same UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant*. When a Cox proportional hazards model was applied, there was survival curve separation but not a statistically significant difference in kidney survival between males and females.]*

3.3.2 Genome Wide Association Results for Kidney Survival in ADTKD-UMOD

After covariates for adjustment were determined, four genome-wide association studies (corresponding to two cohorts, the 266-person cohort of all individuals with ADTKD-*UMOD* and the 92-person cohort of individuals with the same primary *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant, using either genotyped or imputed SNPs) were performed to assess genetic factors associated with kidney survival in ADTKD-*UMOD*. The four Manhattan plots shown in **figures 31-34** are a visual representation of the results generated from the four GWASs.

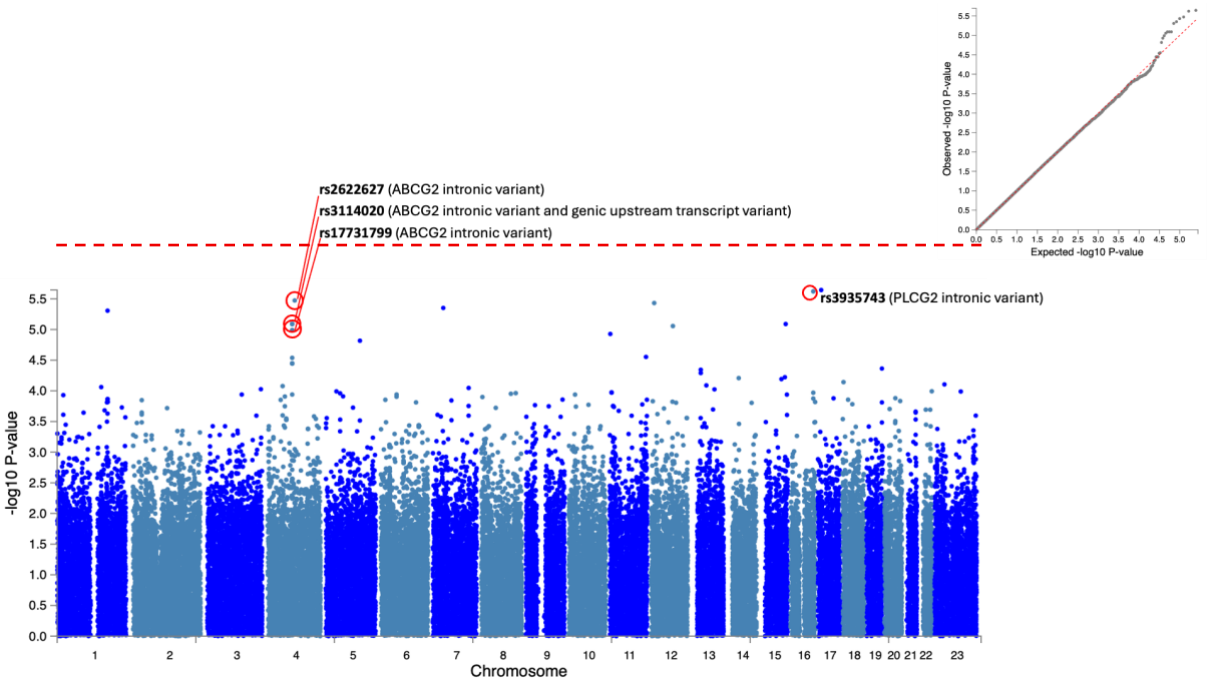

Figure 31: Full 266-person ADTKD-*UMOD* Genome Wide Association Study (Phenotype of Kidney Survival)*

*[*A Manhattan Plot and Quantile-Quantile (Q-Q) Plot (top right corner) of the full 266-person Genome Wide Association Study of participants with ADTKD-UMOD assessing the trait of kidney survival. The Q-Q Plot represents the observed -log$_{10}$(p-values) (y-axis) versus expected -log$_{10}$(p-values) (x-axis) from the GWAS under the null hypothesis. The black diagonal line indicated the expected distribution of no association under the null hypothesis, therefore deviation from this line suggests potential true associations or population stratification. The Manhattan Plot displays log$_{10}$(p-value) of each single nucleotide polymorphism (SNP) tested for association with the trait of kidney survival across the genome. Each point represents a SNP, plotted according to chromosome position on the x-axis and its corresponding p-value on the y-axis. SNPs circled in red and labelled, represent 'candidate SNPs of interest' according to the methodology outlined in **Section 3.2.8**. These results show that there are no regions or SNPs of genome wide significance ($p<5 \times 10^{-8}$) however there is a locus approaching genome-wide significance whereby the three most statistically significant SNPs feature in the ABCG2 gene.]*

3.3.3 Genome Wide Association Results for Kidney Survival in ADTKD-UMOD due to UMOD p.Val93_Gly97delinsAlaAlaSerCys Variant

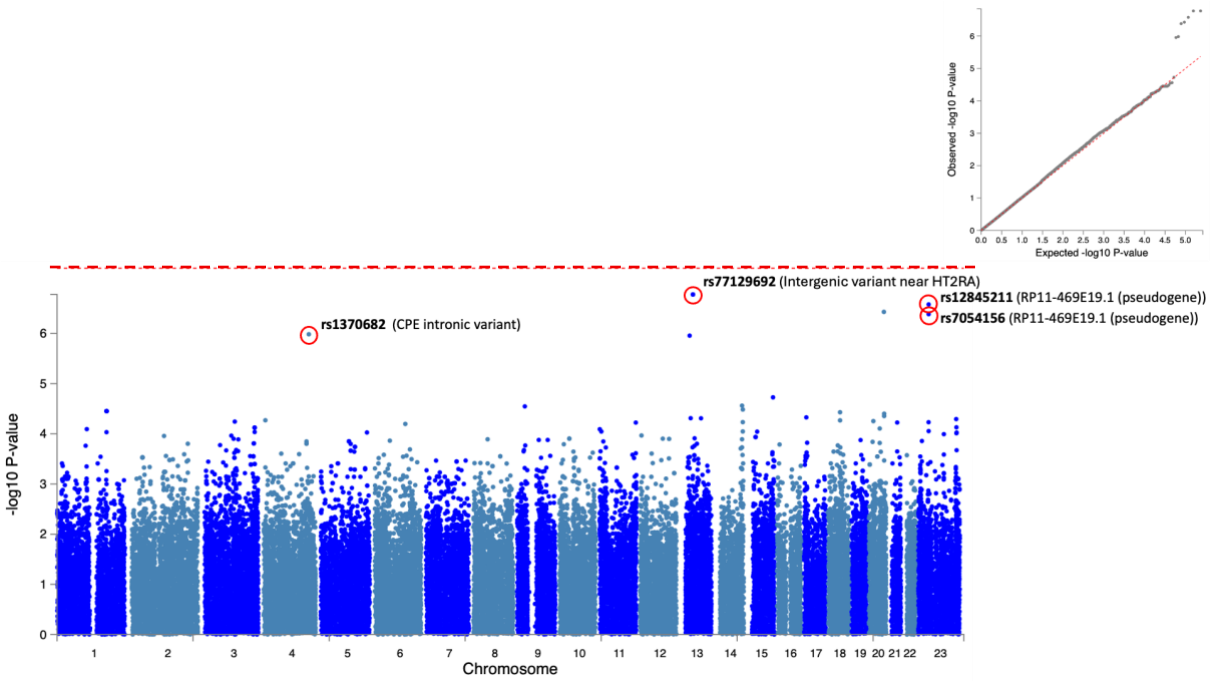

Figure 32: 92-person ADTKD-*UMOD* due to *UMOD* p.Val93_Gly97delinsAlaAlaSerCys Genome Wide Association Study (Phenotype of Kidney Survival)*

*[*A Manhattan Plot and Quantile-Quantile (Q-Q) Plot (top right corner) of the 92-person Genome Wide Association Study of participants with ADTKD-UMOD with the same UMOD p.Val93_Gly97delinsAlaAlaSerCys variant assessing the trait of kidney survival. The Q-Q Plot represents the observed -log$_{10}$(p-values) (y-axis) versus expected -log$_{10}$(p-values) (x-axis) from the GWAS under the null hypothesis. The black diagonal line indicated the expected distribution of no association under the null hypothesis, therefore deviation from this line suggests potential true associations or population stratification. The Manhattan Plot displays log$_{10}$(p-value) of each single nucleotide polymorphism (SNP) tested for association with the trait of kidney survival across the genome. Each point represents a SNP, plotted according to chromosome position on the x-axis and its corresponding p-value on the y-axis. The horizontal red line indicated the genome-wide significant threshold (P <5 x 10^{-8}). SNPs circled in red and labelled, represent 'candidate SNPs of interest' according to the methodology outlined in **Section 3.2.8**. These results show that there are no regions or SNPs of genome wide significance (p<5 x 10^{-8}) however there are SNPs approaching this degree of association.]*

3.3.4 Genome Wide Association Results for Kidney Survival in ADTKD-UMOD using Imputation

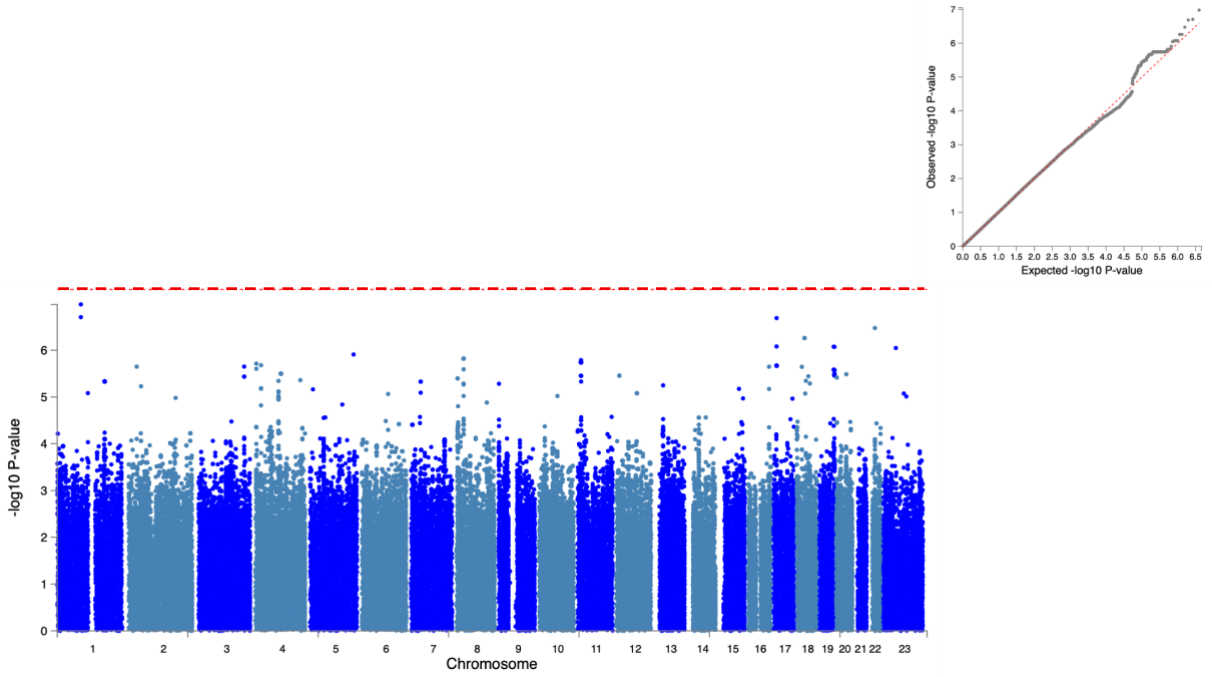

Figure 33: Full 266-person ADTKD-*UMOD* Imputed Genome Wide Association Study (Phenotype of Kidney Survival)*

*[*A Manhattan Plot and Quantile-Quantile (Q-Q) Plot (top right corner) of the full 266-person Imputed Genome Wide Association Study of participants with ADTKD-UMOD assessing the trait of kidney survival. The Q-Q Plot represents the observed -log_{10}(p-values) (y-axis) versus expected -log_{10}(p-values) (x-axis) from the GWAS under the null hypothesis. The black diagonal line indicated the expected distribution of no association under the null hypothesis, therefore deviation from this line suggests potential true associations or population stratification. The Manhattan Plot displays log_{10}(p-value) of each single nucleotide polymorphism (SNP) tested for association with the trait of kidney survival across the genome. Each point represents a SNP, plotted according to chromosome position on the x-axis and its corresponding p-value on the y-axis. The horizontal red line indicated the genome-wide significant threshold ($P < 5 \times 10^{-8}$). These results show that there are no regions or SNPs of genome wide significance ($p < 5 \times 10^{-8}$), although some SNPs are approaching this but are not considered to be 'candidate SNPs of interest' according to the methodology outlined in **Section 3.2.8**.]*

3.3.5 Genome Wide Association Results for Kidney Survival in ADTKD-UMOD due to UMOD p.Val93_Gly97delinsAlaAlaSerCys Variant using Imputation

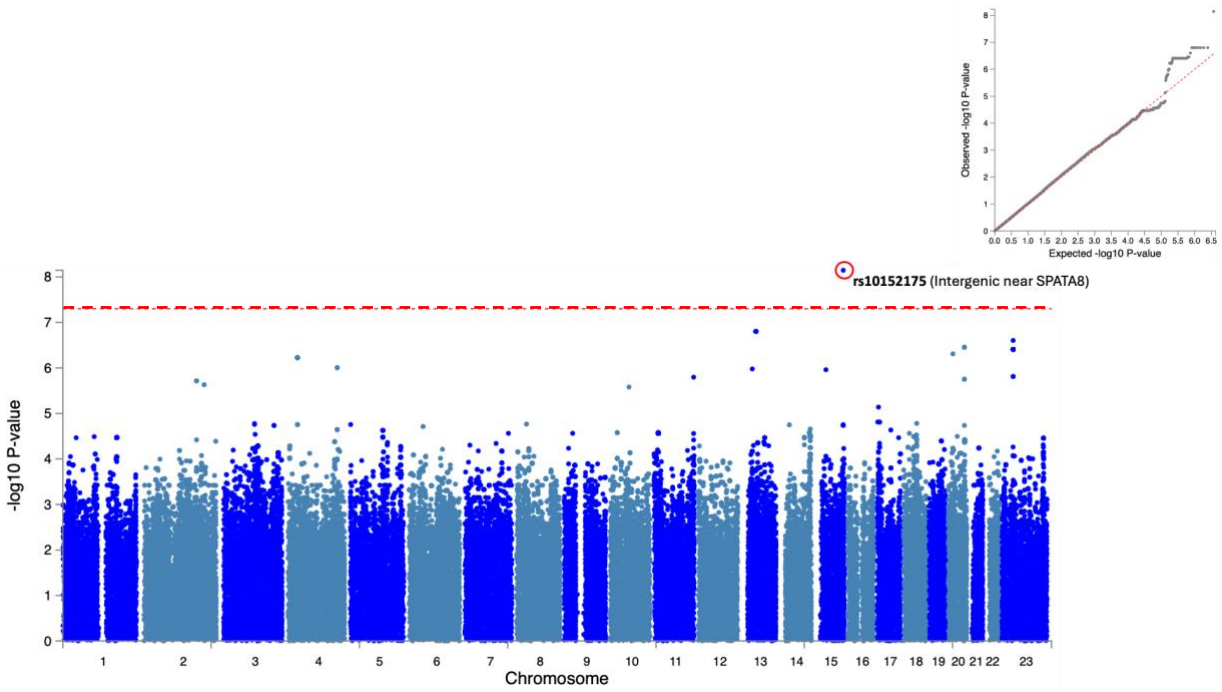

Figure 34: 92-person ADTKD-*UMOD* due to *UMOD* p.Val93_Gly97delinsAlaAlaSerCys Imputed Genome Wide Association Study (Phenotype of Kidney Survival)*

*[*A Manhattan Plot and Quantile-Quantile (Q-Q) Plot (top right corner) of the 92-person Imputed Genome Wide Association Study of participants with ADTKD-UMOD with the same UMOD p.Val93_Gly97delinsAlaAlaSerCys variant assessing the trait of kidney survival. The Q-Q Plot represents the observed -log$_{10}$(p-values) (y-axis) versus expected -log$_{10}$(p-values) (x-axis) from the GWAS under the null hypothesis. The black diagonal line indicated the expected distribution of no association under the null hypothesis, therefore deviation from this line suggests potential true associations or population stratification. The Manhattan Plot displays log$_{10}$(p-value) of each single nucleotide polymorphism (SNP) tested for association with the trait of kidney survival across the genome. Each point represents a SNP, plotted according to chromosome position on the x-axis and its corresponding p-value on the y-axis. The horizontal red line indicated the genome-wide significant threshold (P <5 x 10^{-8}). The SNP circled in red and labelled, represents a 'candidate SNPs of interest' according to the methodology outlined in **Section 3.2.8**. These results show that there is one SNP above the genome wide significance (p<5 x 10^{-8}) threshold and there are also some SNPs approaching this degree of association, however none of these were considered to be a 'candidate SNP of interest' according to the methodology outlined in **Section 3.2.8**.]*

3.3.6 Top SNPs to Explore Further

To illustrate all nine top 'candidate SNPs to explore further' from all four GWASs, **Table 13** highlights the variant name, genomic position, type and various corresponding functional annotation details according to the methodology outlined in **Section 3.2.8. Table 14** highlights both genomic and functional annotation features for the same nine SNPs with the aim of further determining biological relevance, population frequency and potential clinical impact.

TOP GWAS SNPS TO EXPLORE FURTHER ***See below table for subheading meanings****													
Marker ID	**Position (GRCh37)**	**Variant Type**	**Nearest gene**	**LZ**	**GWAS**	**MAGMA**	**FUMA1**	**FUMA2**	**CADD**	**RDB**	**CBI**	**GWAS Catalog**	**Kidney eQTL?**
rs3935743	16:81909980	Intron variant	*PLCG2*	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No
rs2622627	4:89065353	Intron variant	*ABCG2*	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	No
rs3114020	4:89083666	Intron variant, Genic upstream transcript variant	*ABCG2*	Yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
rs17731799	4:89068455	Intron variant	*ABCG2*	Yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
rs77129692	13:47485064	Intergenic	*HTR2A*	Yes	Yes	No	Yes	Yes	Yes	No	Yes	No	No
rs12845211	X:40889466	Intergenic	*RP11-469E19.1*	?	Yes	No	Yes	Yes	Yes	No	Yes	No	No

			(pseudo gene)										
rs7054156	X:40903663	Intergenic	*RP11-469E19.1* (pseudo gene)	?	Yes	No	No	Yes	Yes	No	Yes	No	No
rs1370682	4:166296379	Intron variant	*CPE*	Yes	Yes	No	Yes	Yes	Yes	No	No	No	No
rs10152175	15:97350625	Intergenic	*SPATA8*	Yes	No	No	Yes	Yes	Yes	Yes	Yes	No	No

Table 13: Top Ranked SNPs from Genome Wide Association Studies for Further Exploration (A)*

[(LZ) Hit supported on LOCUSZOOM, (GWAS) Multiple hits across GWASs done, (MAGMA) Gene featured in top MAGMA Gene-burden studies, (FUMA1) Genomic risk locus or lead SNP on FUMA, (FUMA2) Individually significant SNP on FUMA, (CADD) SNP in LD with, SNP in top CADD score list (FUMA), (RDB) SNP in LD with SNP in top RDB score list (FUMA), (CBI) Relevant chromatin bridging interaction, (GWASCatalog) Relevant GWASCatalog trait.]*

TOP GWAS SNPS TO EXPLORE FURTHER							
Marker ID	**Conservation Score Across 100 Vertebral Species (PhyloP on UCSC)**	**Alleles (dbSNP)**	**1000 GenomeMAF**	**Documented Clinical Significance (dbSNP)**	**Kidney Expression of gene? (Human Protein Atlas)**	**Transcript (Human Protein Atlas)**	**Protein Localization (Human Protein Atlas)**
rs3935743	-0.148134	T>A,G	T=0.1522	1 citation, not reported on ClinVar	Yes	Low in glomeruli and low in tubules 15.3 nTPM (RNA)	Cytosol and Plasma Membrane
rs2622627	-0.0583543	C>A,T	A=0.4854	0 citations, not reported on ClinVar	Yes	Medium in CD and DCT 3.4 nTPM (RNA)	Nucleus and Plasma Membrane
rs3114020	0.730622	T>C,G	C=0.4249	4 citations, not reported on ClinVar	Yes	Medium in CD and DCT 3.4 nTPM (RNA)	Nucleus and Plasma Membrane
rs17731799	0.909394	G>A,C,T	T=0.4509	1 citation, not reported on ClinVar	Yes	Medium in CD and DCT 3.4 nTPM (RNA)	Nucleus and Plasma Membrane
rs77129692	-0.332638	C>A,T	T=0.0405	0 citations, not reported on ClinVar	Yes	Protein not detected 0.1 nTPM (RNA)	Plasma Membrane
rs12845211	0.630331	A>G,T	A=0.3468	0 citations, not reported on ClinVar	*Pseudogene*	*Pseudogene*	*Pseudogene*

rs7054156	-1.22305	A>C,G,T	A=0.3650	0 citations, not reported on ClinVar	*Pseudogene*	*Pseudogene*	*Pseudogene*
rs1370682	0.136937	G>A,C,T	A=0.2262	0 citations, not reported on ClinVar	Yes	Protein not detected 66.3 nTPM (RNA)	Extracellular, Golgi and Plasma Membrane
rs10152175	0.322299	G>A,T	T=0.0441	0 citations, not reported on ClinVar	*RNA Gene*	*RNA Gene*	*RNA Gene*

Table 14: Top Ranked SNPs from Genome Wide Association Studies for Further Exploration (B)*

[(Marker ID) SNP identifier (e.g. rsID) from dbSNP, (Conservation Score Across 100 Vertebral Species (PhyloP on UCSC)) PhyloP score reflecting evolutionary conservation across 100 vertebrate species; higher scores indicate greater conservation, (Alleles (dbSNP)) Reported reference and alternate alleles for the variant in dbSNP, (1000 Genome MAF) Minor allele frequency of the variant in the 1000 Genomes Project, (Documented Clinical Significance (dbSNP)) Clinical annotations of the variant (e.g., pathogenic, benign) from dbSNP or linked resources, (Kidney Expression of Gene? (Human Protein Atlas)) Whether the gene is expressed in kidney tissue based on Human Protein Atlas data, (Transcript (Human Protein Atlas)) Transcript-level expression evidence in kidney from the Human Protein Atlas, (Protein Localization (Human Protein Atlas)) Cellular localization of the encoded protein in kidney tissue, based on immunohistochemistry data.]*

SNP rs2622627 was chosen as a particular SNP of interest given its biological relevance with the disease. This SNP in *ABCG2*, a uric acid transporter, is of interest due to ADTKD-*UMOD* being associated with juvenile-onset hyperuricemia and/or early onset gout. **Figure 35** represents further exploration of the SNPs genotype effects on kidney survival.

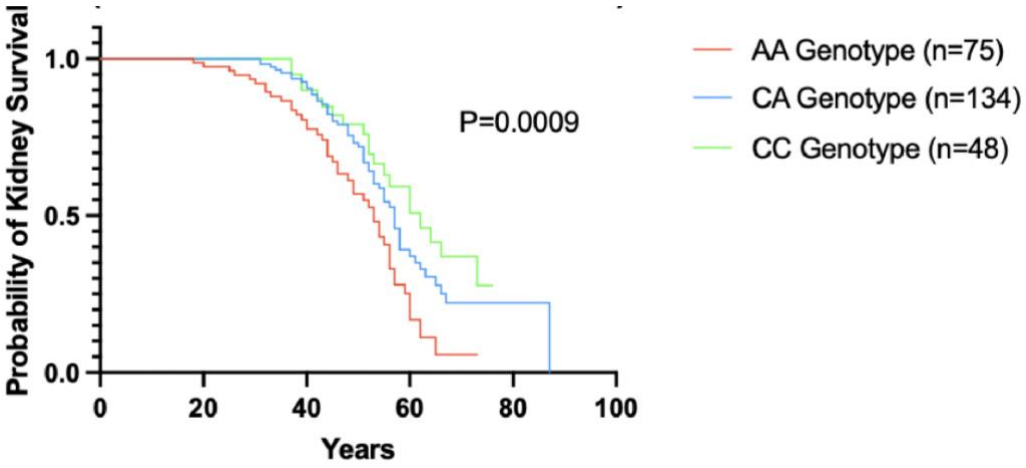

Figure 35: Kidney Survival for rs2622627 Genotypes in 266 Participants with ADTKD-*UMOD**

*[*Kidney survival curves for each of the three genotypes at SNP rs2622627 within the full 266-person ADTKD-UMOD GWAS cohort. When a Cox proportional hazards model was applied, there was a statistically significant difference in kidney survival between each of the three genotypes. The AA genotype was associated with the worst kidney survival in the cohort and the CC genotype was associated with the best. Kidney Survival Curves for all nine top SNPs of interest are featured in the **Appendix**.]*

To provide further insights into the strength of association, linkage disequilibrium patterns and regional genomic context, **figure 36** demonstrates the locus of interest in the ABCG2 gene from the aforementioned GWAS studies.

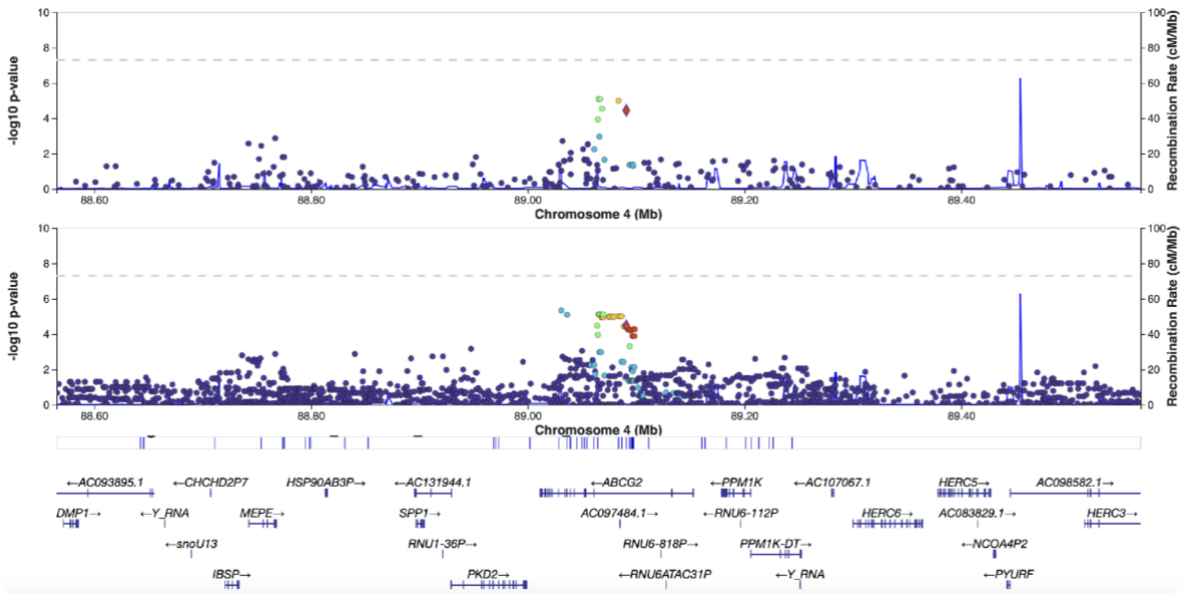

Figure 36: LocusZoom Plots of Main GWAS Signal of Interest in *ABCG2**

*[*Two LocusZoom plots (top from the full 266-person ADTKD-UMOD GWAS and bottom from the imputed full 266-person ADTKD-UMOD GWAS). The X-axis represents the physical location of the SNPs along the chromosome according to their base pair position. The purple diamond represents one of the lead SNPs of interest (rs2622627, rs2622626, rs3114020, rs17731799, rs11732936 and rs10011796). The Y-axis shows the strength of association between each SNP and the trait of kidney survival with the dotted horizontal grey line marking genome-wide significance $-\log_{10}(5 \times 10^{-8})$. SNPs are coloured based on their LD (r^2 value) with the lead SNP. Red (high LD (r^2 >0.8)), orange (moderate LD (r^2 >0.4 and <0.8)) and green/light blue (low LD (r^2 <0.4)). Genes in the region are annotated at the bottom of the figure. As can be seen, SNPs specifically within the ABCG2 are approaching the genome-wide significant threshold in both GWASs and a number are in high to medium LD with the lead SNPs suggesting a regional association with the trait of kidney survival in ADTKD-UMOD. Adjacent to ABCG2 lies PKD2, which encodes polycystin-2, a calcium-permeable ion channel found in human kidney cells. This proximity could suggest that they are part of the same LD block and inherited together, that they could be involved in overlapping biological pathways, have shared regulatory mechanisms or that the ABCG2 locus could be capturing PKD2 variants though LD.]*

3.3.7 *Involvement of the UMOD Promotor and PDILT Locus in kidney survival in a cohort of ADTKD-UMOD patients*

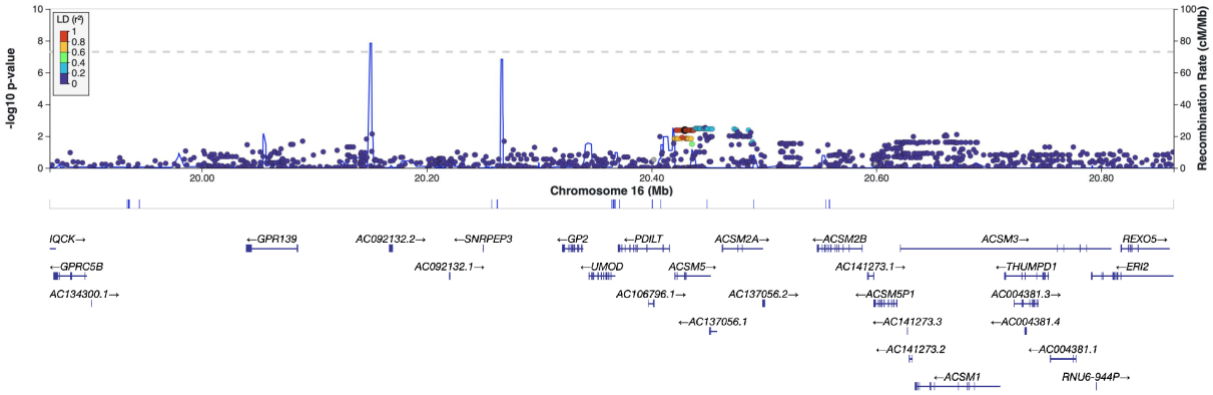

Figure 37: LocusZoom plot of the full 266-person ADTKD-*UMOD* GWAS Cohort focussed on *UMOD* and its promotor region*

*[*A LocusZoom plot of the full 266-person ADTKD-UMOD GWAS Cohort. The X-axis represents the physical location of the SNPs along the chromosome according to their base pair position. The Y-axis shows the strength of association between each SNP and the trait of kidney survival with the dotted horizontal grey line marking genome-wide significance $-\log_{10}(5 \times 10^{-8})$. SNPs are coloured based on their LD (r^2 value) with the lead SNP which was a SNP chosen in the UMOD promotor region to locate this region of interest. Red (high LD (r^2 >0.8)), orange (moderate LD (r^2 >0.4 and <0.8)) and green/light blue (low LD (r^2 <0.4)). Genes in the region are annotated at the bottom of the figure. As can be observed, no SNPs show elevated significance values in relation to the trait of kidney survival in the UMOD gene, PDILT locus or surrounding promotor regions.]*

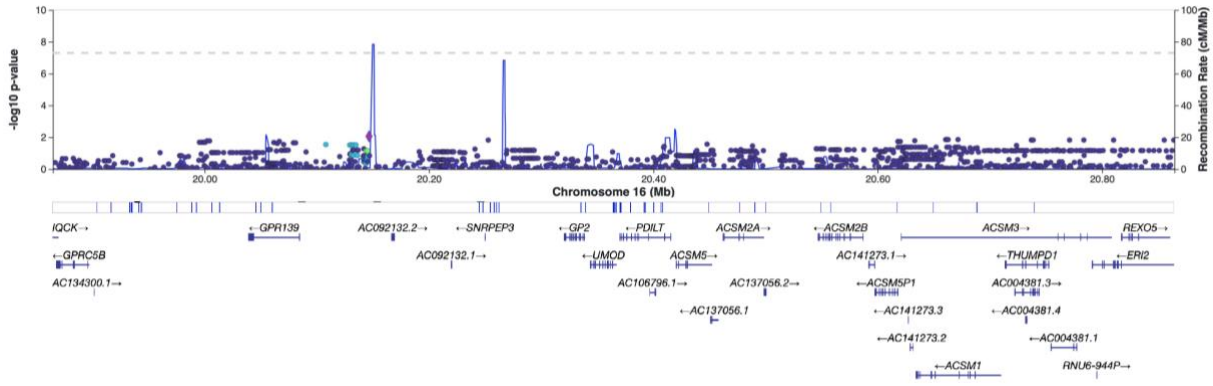

Figure 38: LocusZoom plot of the 92-person ADTKD-*UMOD* GWAS Cohort with the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant focussed on *UMOD* and its promotor region*

*[*A LocusZoom plot of the 92-person ADTKD-UMOD GWAS Cohort of individuals with the UMOD p.Val93_Gly97delinsAlaAlaSerCys variant. The X-axis represents the physical location of the SNPs along the chromosome according to their base pair position. The Y-axis shows the strength of association between each SNP and the trait of kidney survival with the dotted horizontal grey line marking genome-wide significance -log$_{10}$(5 x 10^{-8}). SNPs are coloured based on their LD (r^2 value) with the lead SNP which was a SNP chosen in the UMOD promotor region to locate this region of interest. Red (high LD (r^2 >0.8)), orange (moderate LD (r^2 >0.4 and <0.8)) and green/light blue (low LD (r^2 <0.4)). Genes in the region are annotated at the bottom of the figure. As can be observed, no SNPs show elevated significance values in relation to the trait of kidney survival in the UMOD gene, PDILT locus or surrounding promotor regions.]*

3.3.8 Gene-burden Analysis Results for Kidney Survival in ADTKD-UMOD

Gene-burden studies were then performed by aggregating GWAS signals across genes to improve power to detect associations and aid detection of candidate genes associating with kidney survival in ADTKD-*UMOD*.

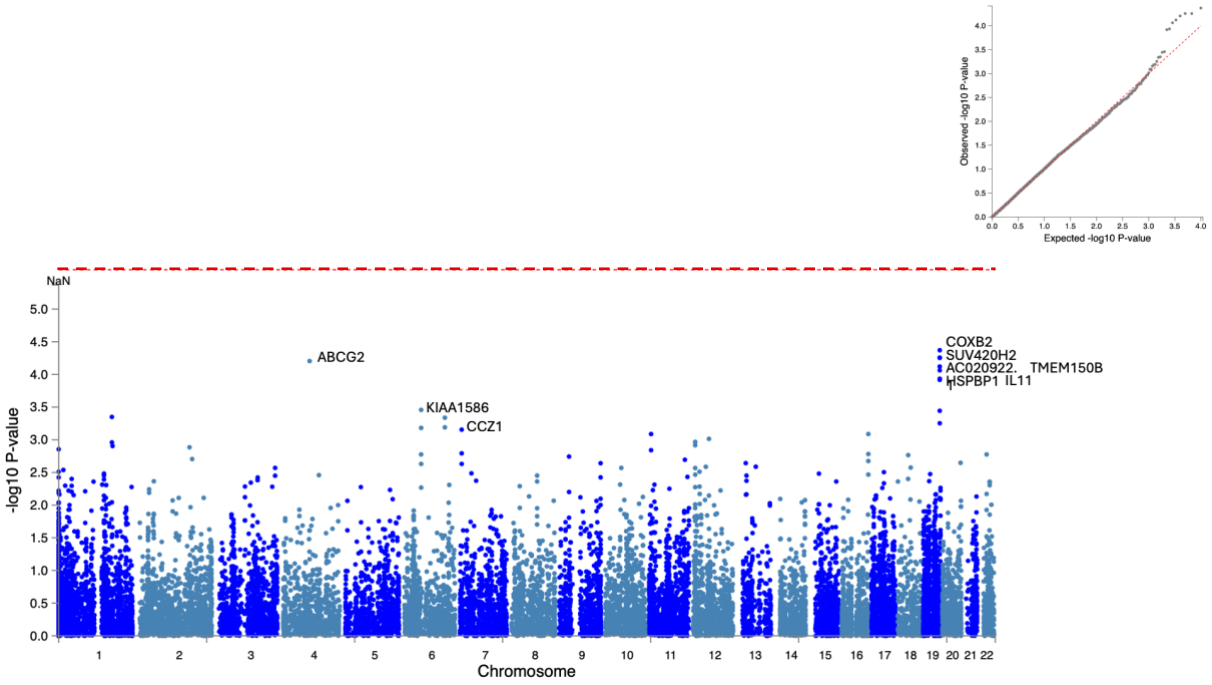

Figure 39: Full 266-person ADTKD-UMOD Gene-burden Study (Phenotype of Kidney Survival)*

*[*A Manhattan Plot and Quantile-Quantile (Q-Q) Plot (top right corner) of the full 266-person Gene-burden Study of participants with ADTKD-UMOD assessing the trait of kidney survival. The Q-Q Plot represents the observed -log$_{10}$(p-values) (y-axis) versus expected -log$_{10}$(p-values) (x-axis) for genes from the Gene-burden Study under the null hypothesis. The black diagonal*

line indicated the expected distribution of no association under the null hypothesis, therefore deviation from this line suggests potential true associations or population stratification. The Manhattan Plot displays log_{10}(p-value) of each gene tested for association with the trait of kidney survival across the genome. Each point represents a gene, plotted according to chromosome position on the x-axis and its corresponding p-value on the y-axis. Genes labelled with their name, represent 'candidate genes of interest' according to the methodology outlined in ***Section 3.2.10****. These results show that there are no regions or genes of genome wide significance (-log_{10} (p) 5.6 ; $p < 2.5 \times 10^{-6}$) however there is a locus approaching genome-wide significance on chromosome 19 and ABCG2 stands out as another gene with standalone importance but approaching rather than reaching genome-wide significance.]*

3.3.9 Gene-burden Analysis Results for Kidney Survival in ADTKD-UMOD due to UMOD p.Val93_Gly97delinsAlaAlaSerCys Variant

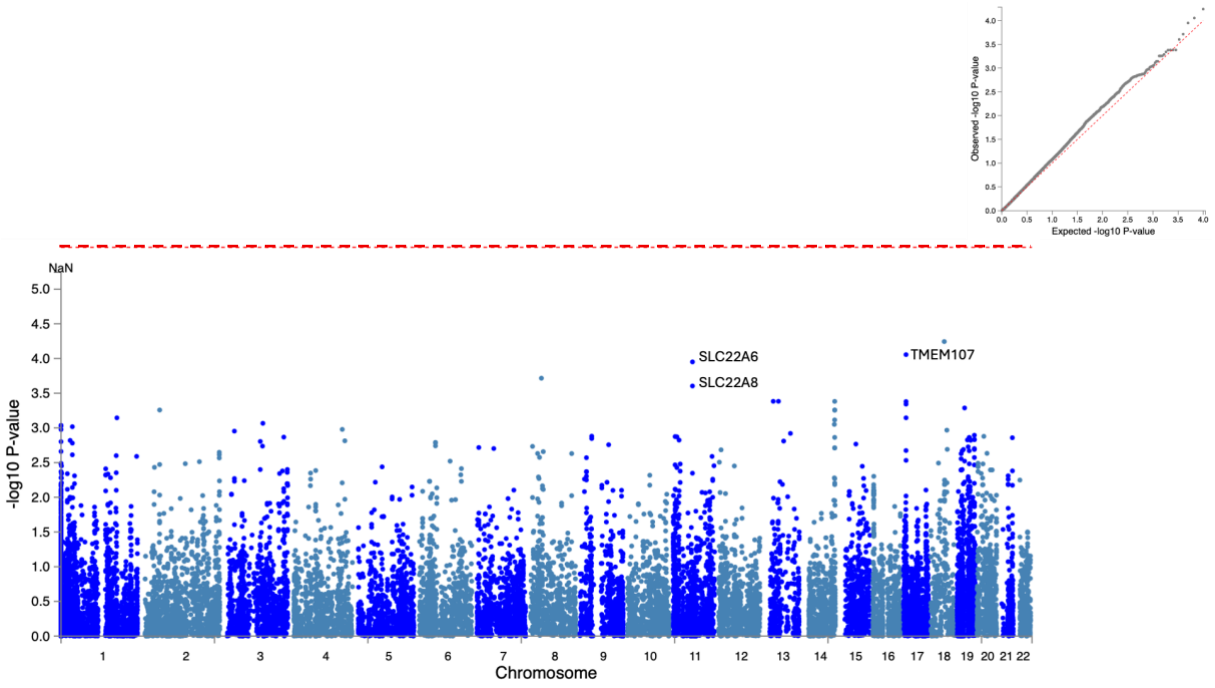

Figure 40: 92-person ADTKD-UMOD due to UMOD p.Val93_Gly97delinsAlaAlaSerCys Gene-burden Study (Phenotype of Kidney Survival)*

*[*A Manhattan Plot and Quantile-Quantile (Q-Q) Plot (top right corner) of the 92-person Gene-burden Study of participants with ADTKD-UMOD with the same UMOD p.Val93_Gly97delinsAlaAlaSerCys variant assessing the trait of kidney survival. The Q-Q Plot represents the observed -log_{10}(p-values) (y-axis) versus expected -log_{10}(p-values) (x-axis) for genes from the Gene-burden study under the null hypothesis. The black diagonal line indicated the expected distribution of no association under the null hypothesis, therefore deviation from*

this line suggests potential true associations or population stratification. The Manhattan Plot displays log_{10}(p-value) of each gene tested for association with the trait of kidney survival across the genome. Each point represents a gene, plotted according to chromosome position on the x-axis and its corresponding p-value on the y-axis. The horizontal red line indicated the genome-wide significant threshold (-log_{10} (p) 5.6 ; $p < 2.5 \times 10^{-6}$). Genes labelled with their name, represent 'candidate genes of interest' according to the methodology outlined in ***Section 3.2.10****. These results show that there are no regions or genes of genome wide significance (-log_{10} (p) 5.6 ; $p < 2.5 \times 10^{-6}$) however the three genes labelled and approaching genome-wide significance have been put forward for further investigation. These include two additional renal uric acid transporters (SLC22A6 and SLC22A8).]*

3.3.10 Gene-burden Analysis Results for Kidney Survival in ADTKD-UMOD using Imputation

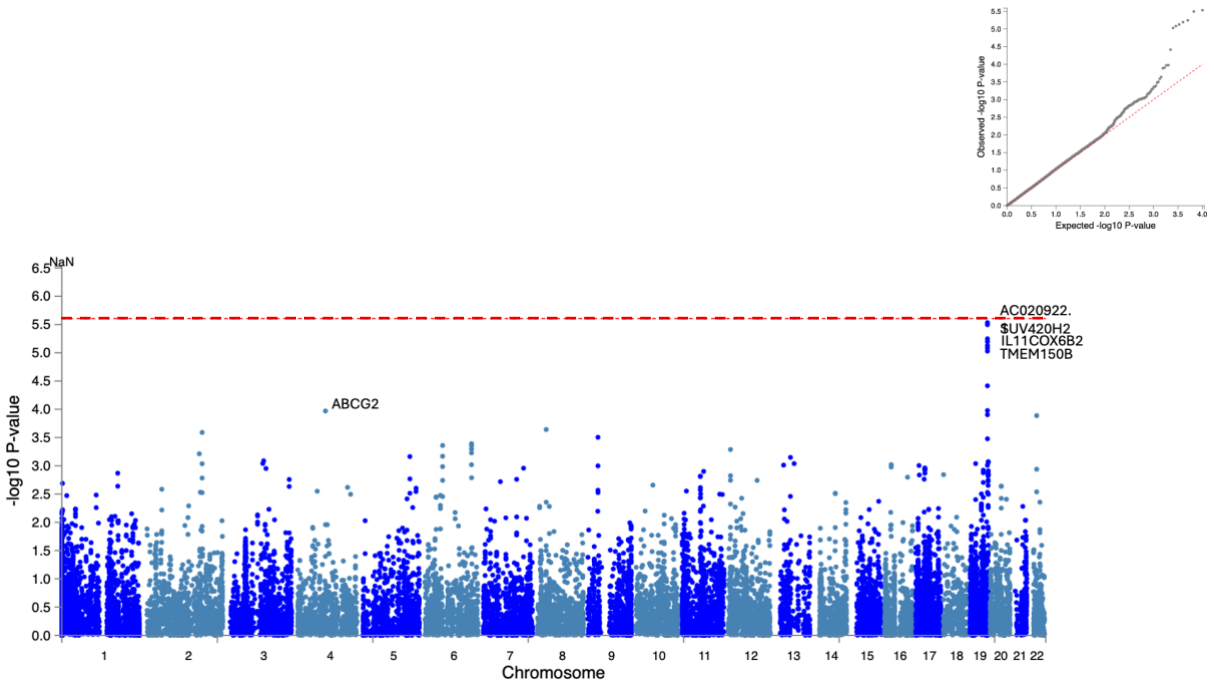

Figure 41: Full 266-person ADTKD-*UMOD* Imputed Gene-burden Study (Phenotype of Kidney Survival)*

*[*A Manhattan Plot and Quantile-Quantile (Q-Q) Plot (top right corner) of the full 266-person Imputed Gene-burden Study of participants with ADTKD-UMOD assessing the trait of kidney survival. The Q-Q Plot represents the observed -log_{10}(p-values) (y-axis) versus expected -log_{10}(p-values) (x-axis) for genes from the Gene-burden Study under the null hypothesis. The black diagonal line indicated the expected distribution of no association under the null hypothesis, therefore deviation from this line suggests potential true associations or*

population stratification. The Manhattan Plot displays log_{10}(p-value) of each gene tested for association with the trait of kidney survival across the genome. Each point represents a gene, plotted according to chromosome position on the x-axis and its corresponding p-value on the y-axis. The horizontal red line indicated the genome-wide significant threshold (-log_{10} (p) 5.6 ; $p < 2.5 \times 10^{-6}$). These results show that the previously specified locus on chromosome 19 is reaching genome-wide significance and ABCG2 is also labelled as a standalone gene approaching the genome-wide significant threshold. Genes labelled signify 'candidate genes of interest' according to the methodology outlined in ***Section 3.2.10****.]*

3.3.11 Gene-burden Analysis Results for Kidney Survival in ADTKD-UMOD due to UMOD p.Val93_Gly97delinsAlaAlaSerCys Variant UMOD using Imputation

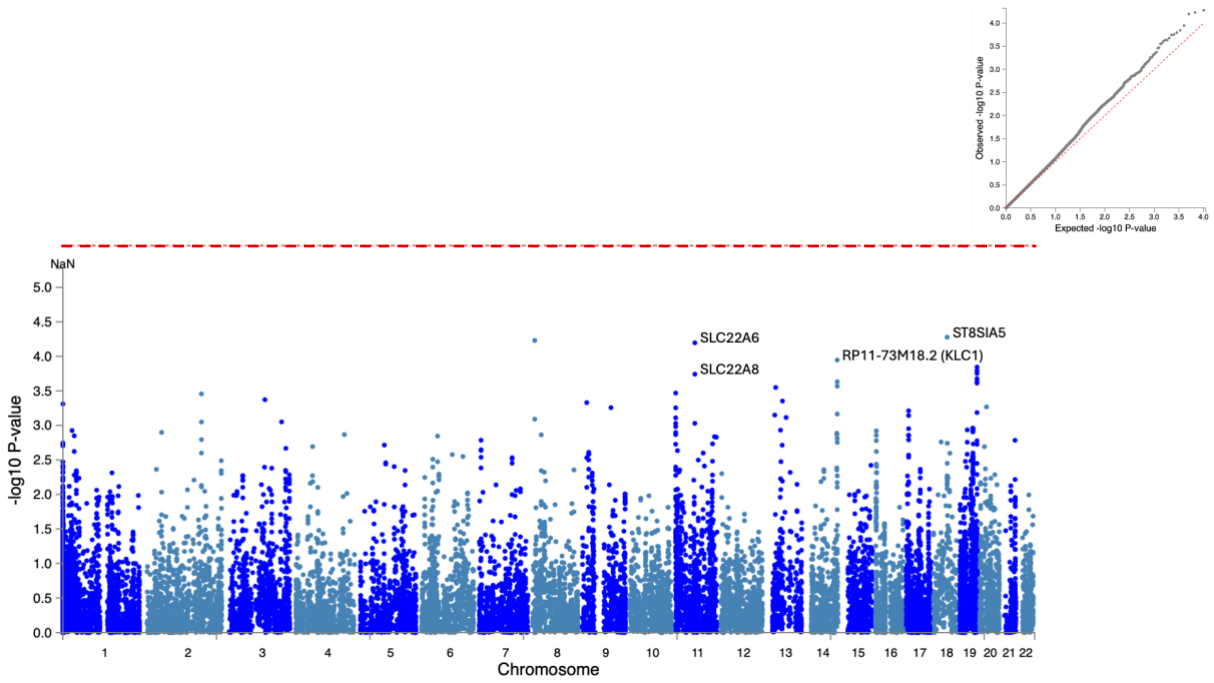

Figure 42: 92-person ADTKD-*UMOD* due to *UMOD* p.Val93_Gly97delinsAlaAlaSerCys Imputed Gene-burden Study (Phenotype of Kidney Survival)*

*[*A Manhattan Plot and Quantile-Quantile (Q-Q) Plot (top right corner) of the 92-person Imputed Gene-burden Study of participants with ADTKD-UMOD with the same UMOD p.Val93_Gly97delinsAlaAlaSerCys variant assessing the trait of kidney survival. The Q-Q Plot represents the observed -log_{10}(p-values) (y-axis) versus expected -log_{10}(p-values) (x-axis) from the Gene-burden study under the null hypothesis. The black diagonal line indicated the expected distribution of no association under the null hypothesis, therefore deviation from this line suggests potential true associations or population stratification. The Manhattan Plot displays log_{10}(p-value) of each gene tested for association with the trait of kidney survival*

across the genome. Each point represents a gene, plotted according to chromosome position on the x-axis and its corresponding p-value on the y-axis. The horizontal red line indicated the genome-wide significant threshold ($-\log_{10}$ (p) 5.6 ; $p < 2.5 \times 10^{-6}$). Genes labelled with their name, represent 'candidate genes of interest' according to the methodology outlined in ***Section 3.2.10*** *and as can be seen, no genes reach genome-wide significance, but the same chromosome 19 locus of interest appears again as approaching this threshold along with other genes of interest labelled. These include the previously acknowledged two renal uric acid transporters (SLC22A6 and SLC22A8).]*

3.3.12 Top Genes to Explore Further

TOP MAGMA GENES TO EXPLORE FURTHER							
Gene	**CHR (GRCh38)**	**Appears more than once in MAGMA Studies**	**MAGMA Significance High**	**Protein/RNA Expressed in kidney**	**Protein/RNA Expression HIGH in kidney**	**Biological function relevant to ADTKD-*UMOD***	**Gene in GWAS list of top SNPs**
ABCG2	Chromosome 4: 88,090,150-88,231,628	Yes	Yes	Yes	Yes	Yes	Yes
COX6B2	Chromosome 19: 55,349,306-55,354,719	Yes	Yes	Yes	No	Yes	No
SUV420H2 (KMT5C)	Chromosome 19: 55,339,853-55,348,121	Yes	Yes	Yes	No	Yes	No
TMEM150B	Chromosome 19: 55,312,801-55,334,048	Yes	Yes	Yes	No	Yes	No
SLC22A6	Chromosome 11: 62,936,385-62,984,967	Yes	No	Yes	Yes	Yes	No
SLC22A8	Chromosome 11: 62,989,154-63,015,841	Yes	No	Yes	Yes	Yes	No
IL11	Chromosome 19: 55,364,382-55,370,463	Yes	No	Yes	No	Yes	No
TMEM107	Chromosome 17: 8,172,457-8,176,399	No	Yes	Yes	Yes	Yes	No
KLC1	Chromosome 14: 103,561,896-103,714,249	Yes	No	Yes	Yes	Yes	No
ST8SIA5	Chromosome 18: 46,667,821-46,759,257	No	Yes	Yes	No	Yes	Yes

KIAA1586	Chromosome 6: 57,046,532-57,055,239	No	No	Yes	Yes	Yes	No
HSPBP1	Chromosome 19: 55,262,223-55,280,381	No	No	Yes	Yes	Yes	No
CCZ1	Chromosome 7: 5,898,725-5,926,550	No	No	Yes	Yes	Yes	No

Table 15: Top Ranked Gene-burden Test Genes for Further Exploration*

*[*To illustrate all thirteen 'candidate genes to explore further' from all four Gene-burden Studies, this table highlights the gene name, genomic positions and various corresponding functional annotation details according to the methodology outlined in **Section 3.2.10** with the aim of further determining the biological relevance and potential clinical impact.]*

SNP rs12459907 was chosen to examine as it has the highest p-value in the full imputed 266-person GWAS within the gene of interest (KMT5C). **Figure 43** represents further exploration of the SNPs genotype effects on kidney survival.

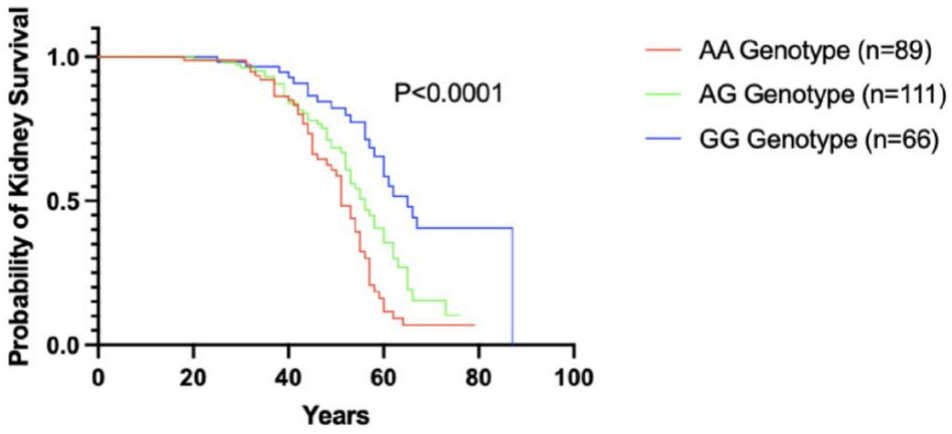

Figure 43: Kidney Survival for *rs12459907* Genotypes (in *KMT5C*) in 266 Participants with ADTKD-*UMOD

*[*Kidney survival curves for each of the three genotypes at SNP rs12459907 within the full 266-person ADTKD-UMOD GWAS cohort. When a Cox proportional hazards model was applied, there was a statistically significant difference in kidney survival between each of the three genotypes. The AA genotype was associated with the worst kidney survival in the cohort and the GG genotype was associated with the best.]*

To provide further insights into the strength of association, linkage disequilibrium patters and regional genomic context, **Figure 44** demonstrates the locus of interest on chromosome 19 from the aforementioned Gene-burden studies.

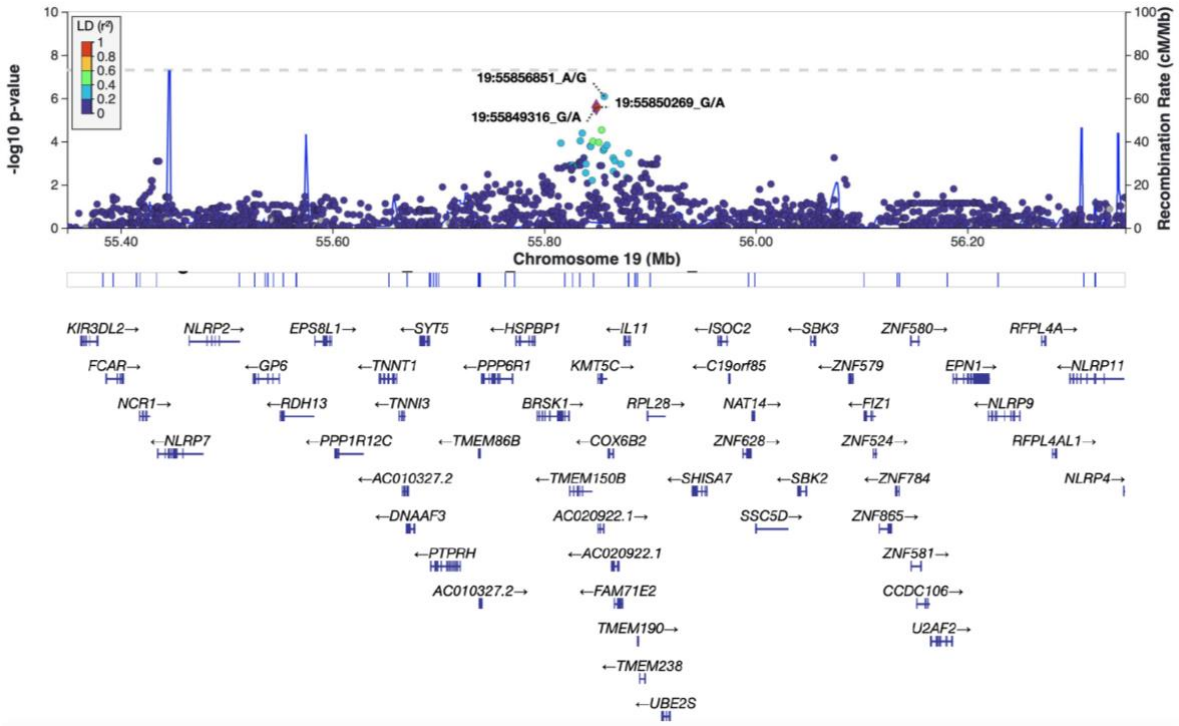

Figure 44: LocusZoom Plots of Main Gene-Burden Study Signal of Interest*

*[*A LocusZoom plot (GRCh37) from the imputed full 266-person ADTKD-UMOD GWAS. The X-axis represents the physical location of the SNPs along the chromosome according to their base pair position. The purple diamond represents one of the lead SNPs of interest (rs1870073). The Y-axis shows the strength of association between each SNP and the trait of kidney survival with the dotted horizontal grey line marking genome-wide significance -log$_{10}$(5 x 10^{-8}). SNPs are coloured based on their LD (r^2 value) with the lead SNP. Red (high LD (r^2 >0.8)), orange (moderate LD (r^2 >0.4 and <0.8)) and green/light blue (low LD (r^2 <0.4)). Genes in the region are annotated at the bottom of the figure. As can be seen, SNPs specifically within the KMT5C, COX6B2, AC020922.1 and other surrounding genes are approaching the genome-wide significant threshold, and a number are in medium to low LD with the lead SNPs suggesting a regional association with the trait of kidney survival in ADTKD-UMOD.]*

3.3.13 Conditional Analysis of Chromosome 19 Locus

GRCh37	p.value.spa Pre-adjustment for 19:55856851	p.value.spa Post-adjustment for 19:55856851	p.value.spa Pre-adjustment for 19:55849316	p.value.spa Post-adjustment for 19:55849316	p.value.spa Pre-adjustment for 19:55850269	p.value.spa Post-adjustment for 19:55850269	p.value.spa Pre-adjustment for 19:55854389	p.value.spa Post-adjustment for 19:55854389
19:55856851 (rs12459907)	-	-	2.01e-06	0.13	2.01e-06	0.13	2.01e-06	0.21
19:55849316 (rs79012440)	5.76e-06	0.33	-	-	5.76e-06	0.99	5.76e-06	0.29
19:55850269 (rs1870073)	5.76e-06	0.33	5.76e-06	0.99	-	-	5.76e-06	0.29
19:55854389 (rs12610863)	5.57e-05	0.81	5.57e-05	0.51	5.57e-05	0.51	-	-

Table 16: Conditional Analysis Table of Results of Top SNPs at Chromosome 19 Locus*

*[*The results of conditional analysis using PLINK. The results reveal that the, at the four SNPs of interest, the association signals observed are probably driven by a shared underlying genetic effect on the trait of kidney survival. When each SNP is conditioned on, the p-values of the remaining three SNPs significantly increase suggesting a loss of their independent significance. This pattern suggests that the SNPs are tagging each other due to the presence of linkage disequilibrium with the causal variant. The lead SNP, 19:55856851, appears to be the signal's primary driver as the other SNPs' associations are effectively nullified after conditioning on this lead SNP. Overall, the findings suggest a shared genetic effect of these four SNPs rather than multiple independent regional signals.]*

3.3.14 Chromatin Interaction Mapping in Loci of Interest

Using the methods as described in **Section 3.2.12**, the three ABCG2 SNPs of interest (*rs2622627, rs17731799* and *rs3114040*) are in linkage disequilibrium with *rs2622627* linked to *rs17731799*, which in turn is linked to *rs3114020*. ABCG2 is expressed in the proximal convoluted tubule in the kidney and exists in three main isoforms, each with active promotors. However, gene expression of *ABCG2* is primary expressed in the GI tract and in haemopoietic cells such as macrophages and erythroblasts. In the kidney, there is enhancer accessibility adjacent to promotor 1 and promotor 2, near *rs3114020*, in PCT cells, with promotor 1 confirmed as active. Enhancers near *rs2622627* and *rs17731799*, previously associated with activity in kidney and mesenchymal cells, are also shown to be active in epithelial cells, regulating *ABCG2* itself and *PKD2*, and in the case of *rs2622627*, *HERC5*. In contract, the enhancer adjacent to *rs3114020* appears to regulate *ABCG2* and is specifically active in macrophages. All three SNPs are eQTLs for *ABCG2* and *SPP1* in blood cells, but none are eQTLs for *PKD2*.

To further assess the two regions of greatest interest, the chromosome 19 signal and the *ABCG2* gene specifically, Circos plots were generated to visually highlight chromatin bridging interactions at large scale.

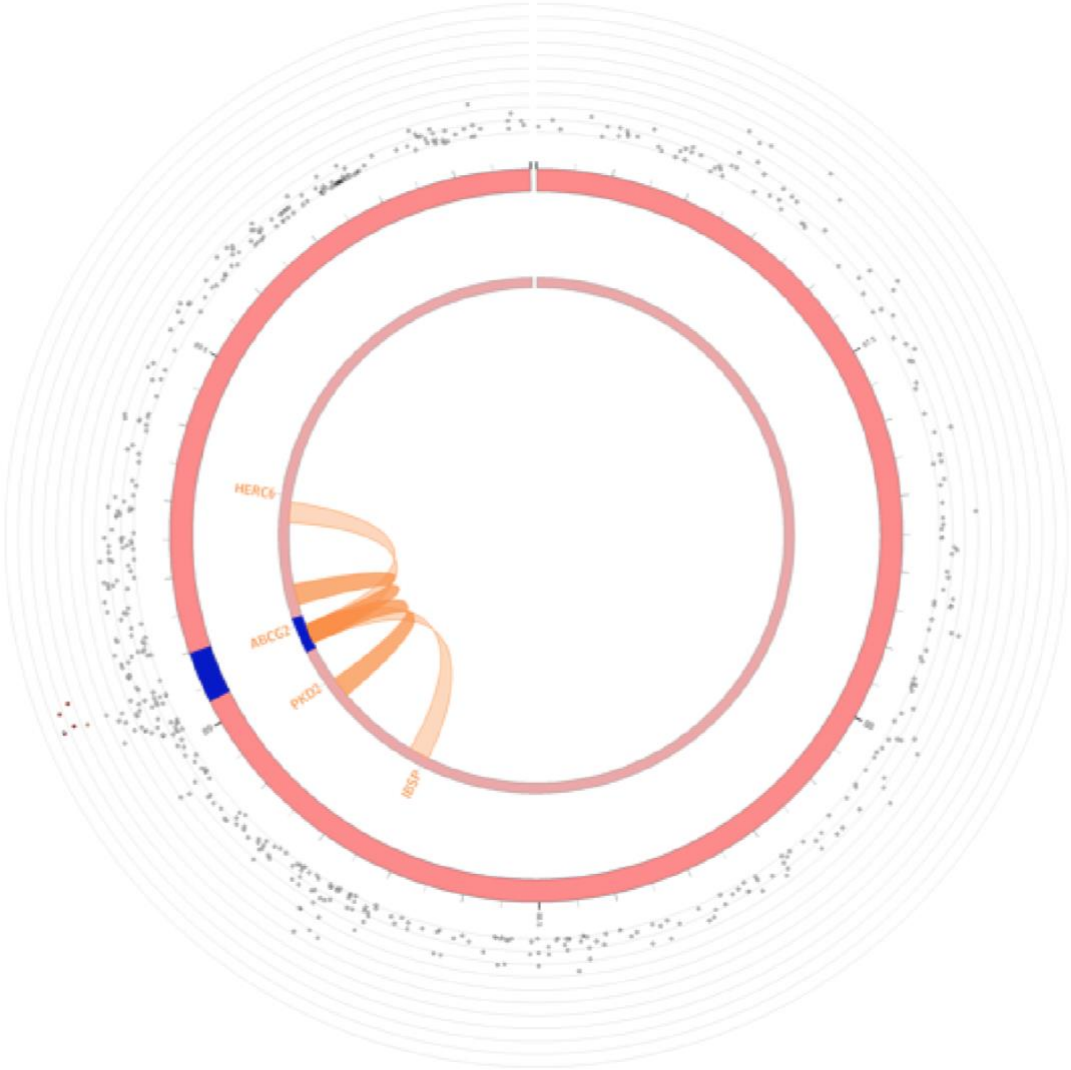

Figure 45: Circos Plot demonstrating Chromatin Bridging Interactions on Chromosome 4*

*[*A circular visualisation of chromatin bridging interactions on the locus of interest at ABCG2. The outermost ring represents the associations of SNPs with the trait of kidney survival from the full 266-person ADTKD-UMOD GWAS. Higher dots represent stronger associations with the trait. The red ring represents the chromosome of interest (chromosome 4) and the physical positions of SNPs according to base pair position. Four genes are plotted according to their genomic coordinates (ABCG2, PKD2, HERC4 and ISPF). The orange arcs represent chromatin bridging interactions. The stronger the interaction, the darker the orange. The region highlighted in dark blue marks the locus of interest in ABCG2. Overall, this plot suggests a regulatory chromatin bridging interaction between ABCG2 and PKD2.]*

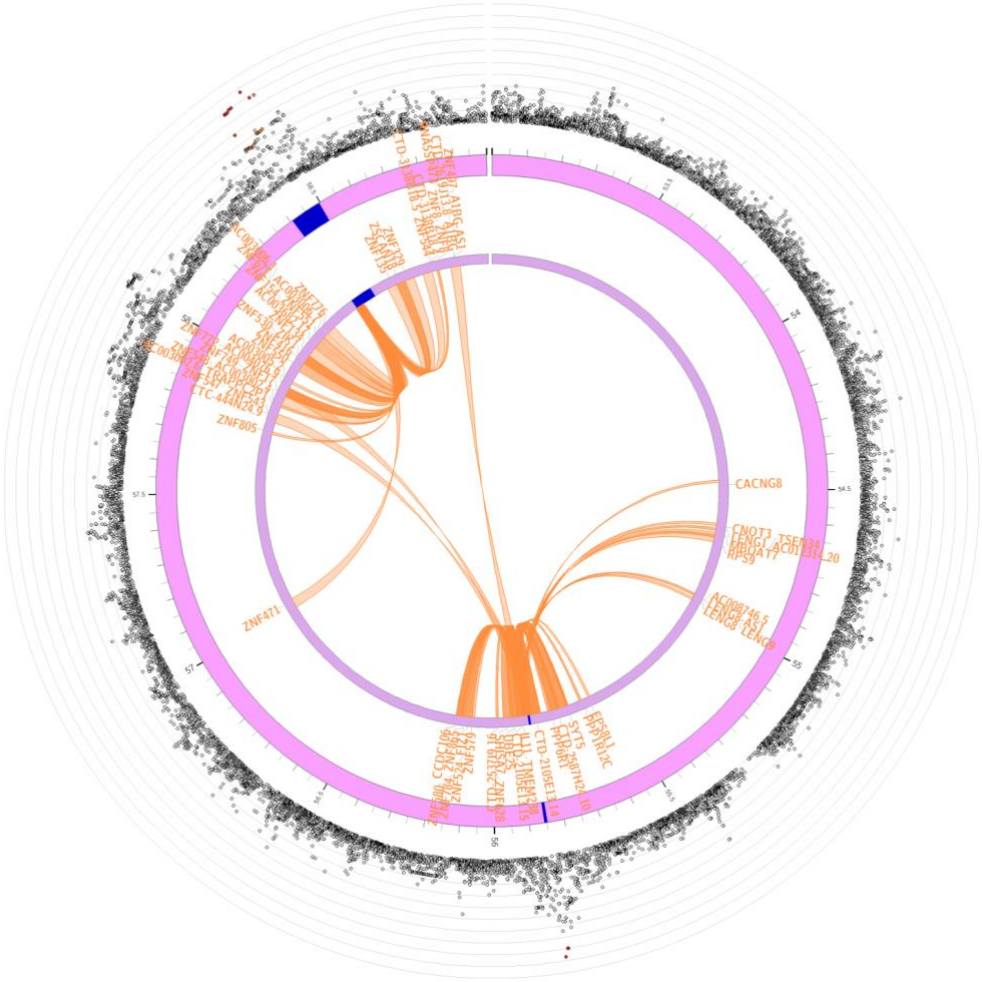

Figure 46: Circos Plot demonstrating Chromatin Bridging Interactions on Chromosome 19*

*[*A circular visualisation of chromatin bridging interactions on the locus of interest on chromosome 19. The outermost ring represents the associations of SNPs with the trait of kidney survival from the full imputed 266-person ADTKD-UMOD GWAS. Higher dots represent stronger associations with the trait. The pink ring represents the chromosome of interest (chromosome 19) and the physical positions of SNPs according to base pair position. Multiple genes are plotted according to their genomic coordinates. The orange arcs represent chromatin bridging interactions. The stronger the interaction, the darker the orange. The region highlighted in dark blue marks the locus of interest in KMT5C. Overall, this plot suggests a region particularly rich in regulatory function with multiple regulatory chromatin bridging interactions from both the locus of interest as well as a second locus of interest on the same chromosome.]*

3.3.15 Polygenic Risk Score for CKD as a Predictor of Kidney Survival in ADTKD-UMOD

Association of Polygenic Risk Score Applied to ADTKD-*UMOD* Cohort with Kidney Survival					
Cohort	**Likelihood ratio test**	**P value on specified degrees of freedom**	**N**	**Number of Events**	**Score P value**
Full 266-Person ADTKD-*UMOD* Cohort	47.53	3.029e-05 (on 15 degrees of freedom)	266	124	**0.065**
92-Person 'indel' Cohort with *UMOD* p.Val93_Gly97delinsAlaAlaSerCys	23.55	0.01476 (on 11 degrees of freedom)	92	49	**0.012038**

Table 17: Association of Polygenic Risk Score Applied to ADTKD-*UMOD* Cohort with Kidney Survival*

*[*The results of the application of a Polygenic Risk Score for CKD (Khan et al., 2022) to both the full 266-person ADTKD-UMOD cohort and the 92-person cohort of individuals with the same primary UMOD p.Val93_Gly97delinsAlaAlaSerCys variant and its association with kidney survival. The PRS for CKD does not statistically associate with kidney survival in the full 266-person ADTKD-UMOD cohort but the p-value of 0.065 is not far off the conventional threshold for statistical significance. This indicates that there is currently not strong enough evidence that the PRS is associated with kidney survival. However, when applied to the 92-person UMOD p.Val93_Gly97delinsAlaAlaSerCys cohort, the PRS has a statistically significant association with kidney survival.]*

To confirm if the PRS for CKD can clinically differentiate between fast and slow progressors of kidney function decline in ADTKD-*UMOD*, the following Kaplan-Meier curves and scatter plots were conceived using data from both cohorts.

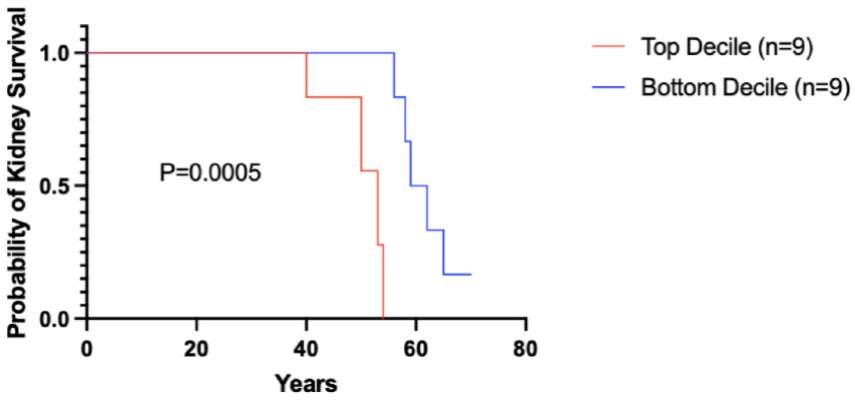

Figure 47: Kidney Survival for Extreme Tails of Polygenic Risk Score in those with the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant*

*[*Kidney survival curves for the two extreme tails of PRS score (9 highest and 9 lowest scoring individuals) in the 92-person UMOD p.Val93_Gly97delinsAlaAlaSerCys cohort. When a Cox proportional hazards model was applied, there was a statistically significant difference in kidney survival between each extreme tail reinforcing association of the PRS with kidney survival in this cohort.]*

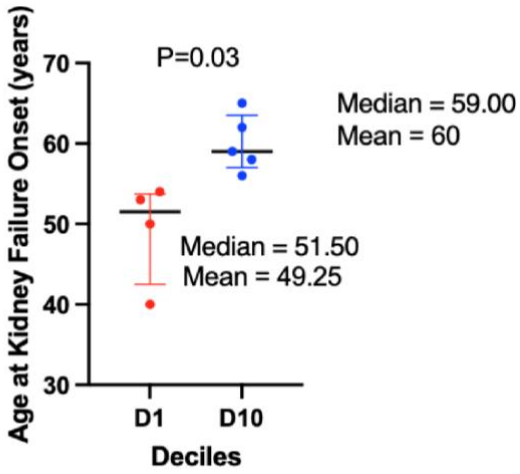

Figure 48: Age of Kidney Failure Onset for Extreme Tails of Polygenic Risk Score in those with the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant*

*[*A scatter plot for the two extreme tails (top and bottom score deciles) of PRS score for those who have already reached end stage kidney failure in the 92-person UMOD p.Val93_Gly97delinsAlaAlaSerCys cohort. An unpaired t test with Welch's correlation was applied. There is a statistically significant difference between the median values for age of end stage kidney failure for both score groups reinforcing association of the PRS with age of kidney failure onset in this cohort.]*

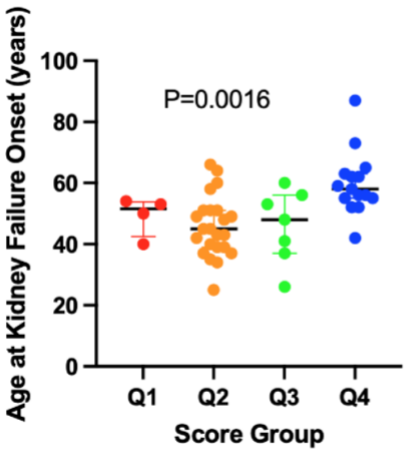

Figure 49: Age of Kidney Failure Onset for Polygenic Risk Score Quartiles in those with the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant*

*[*A scatter plot for the age of onset of kidney failure in the 92-person UMOD p.Val93_Gly97delinsAlaAlaSerCys cohort for each PRS score group when broken into quartiles by score threshold. An unpaired t test with Welch's correlation was applied. There is a statistically significant difference between the median values for age of end stage kidney failure for both score groups reinforcing association of the PRS with age of kidney failure onset in this cohort.]*

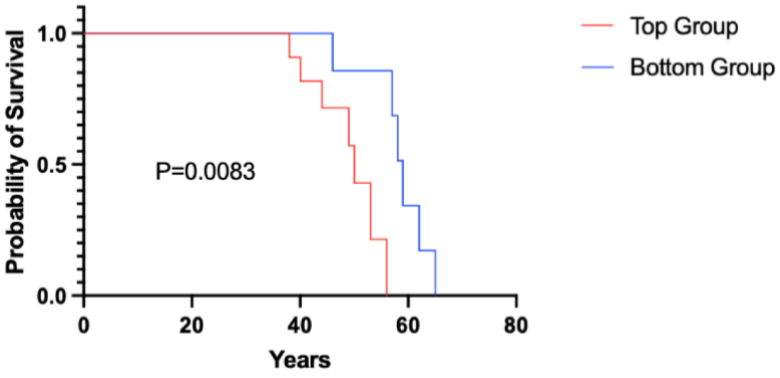

Figure 50: Kidney Survival for Extreme Tails of Polygenic Risk Score in Full ADTKD-*UMOD* Cohort*

*[*Kidney survival curves for the two extreme tails of PRS score in the full 266-person ADTKD-UMOD cohort. When a Cox proportional hazards model was applied, there was a statistically significant difference in kidney survival between each extreme tail suggesting association of the PRS with kidney survival in this cohort even though this did not statistically associate in this cohort using the model outlined in Table 17. Top group n=13, bottom group n=10.]*

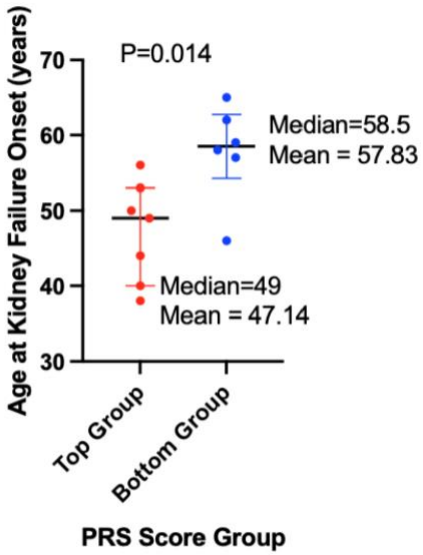

Figure 51: Age of Kidney Failure Onset for Extreme Tails of Polygenic Risk Score in Full ADTKD-*UMOD* Cohort*

*[*A scatter plot for the age of onset of kidney failure in the full 266-person ADTKD-UMOD cohort for the highest and lowest PRS score groups based on score thresholds. An unpaired t test with Welch's correlation was applied. There is a statistically significant difference between the median values for age of end stage kidney failure for both score groups even though the PRS did not statistically associate with kidney survival in this cohort using the model outlined Table 17.]*

3.3.16 Polygenic Risk Score Variation Contributions

After calculations described in **Section 3.2.14**, 0.17% of the overall variation of the PRS is accounted for by the variants in *UMOD* and surrounding regions including its promotor.

3.3.17 Haplotype Estimation of the UMOD p.Val93_Gly97delinsAlaAlaSerCys variant

Ten unrelated patients from the US cohort and twenty unrelated patients from the UK cohort, all with the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant, appear to have a common ancestor who lived 3 generations before with a confidence interval from 2 to 4 generations. The longest common haplotype is 477123 base pairs in length (or 0.95 cM in length) and is found in the 3 unrelated patients (2007, 1052, 1158). The shared haplotype to all patients is only around 20kb long.

3.4 Discussion

Genome Wide Association Studies (GWAS) identify genetic variants across the genome that associate with a trait of interest by analysing genetic differences between individuals. GWAS can offer an unbiased, hypothesis-free and powerful approach to determine genetic modifiers of disease progression by analysing genetic variation in individuals with the same disease (Visscher *et al.*, 2017). The challenge of small sample sizes in rare disease research is always a limiting factor but taking advantage of methods such as imputation, meta-analysis, gene-burden studies and pathway analysis can aid detection of genetic loci implicated in disease pathogenesis. Identifying genetic variants associated with disease progression can provide insights into underlying biological mechanisms and uncover potential therapeutic targets and allow us to risk stratify patients with the disease. GWAS does, however, come with several challenges in this setting. Small sample sizes can limit detection of rare variant contributions, and, in this highly related dataset, population stratification and relatedness can introduce confounding. It should be noted that GWAS identifies associations rather than causal mechanisms, requiring functional studies to interpret biological significance of findings (Tam *et al.*, 2019).

After interrogating the UK RaDaR Registry ('National Registry of Rare Kidney Diseases (RaDaR)', 2022) and reaching out to potential national and international collaborators, I was able to acquire DNA from 275 people with genetically confirmed ADTKD-*UMOD* **(figure 24 & table 12)**. This was a larger number in the registry than outlined in the introduction due to ongoing recruitment to RaDaR and sourcing of study participants thanks to international collaborations. After genotyping, this allowed me to conduct genome-wide association studies and both imputation and gene-burden studies allowed further statistical power to be gained (Das *et al.*, 2016; Guo *et al.*, 2016). Functional annotation of SNPs approaching genome-wide significance allowed prioritisation of key candidate SNPs **(tables 13 & 14)** and chromatin interaction mapping revealed further functional relevance of key variants (Watanabe *et al.*, 2017). Application of a Polygenic Risk Score for CKD enabled contributary assessment of multiple aggregated polygenic factors in disease progression (Khan *et al.*, 2022, 2023).

3.4.1 The ADTKD-UMOD GWAS cohort was limited by size but advantaged by genetic homogeneity and precious consolidation of rare disease genome-wide data

As with any rare disease, ADTKD-*UMOD* faces the challenges of dispersed patient cohorts, disease heterogeneity and low patient numbers for research participation (Devuyst *et al.*, 2014). Genetic diagnosis and recruitment to RaDaR in ADTKD-*UMOD* patients was mostly limited to centres of research interest and RaDaR requires strategies for recruitment improvement to enhance cohort sizes. Due to the 'bland' clinical disease characteristics and poor genetic literacy amongst nephrologists the disease is undoubtedly underdiagnosed both clinically and genetically. Collaboration has been beneficial in this work to expand cohort size, but even with this, ensuring accompanying clinical data remained a further crucial limiting factor. The nature of diagnosis, through symptomatic advanced CKD, allowed for severe phenotypes to be included, and milder phenotypes were typically detected through cascade screening allowing for a heterogeneous cohort.

Due to the 'bland' clinical characteristics and lower prevalence compared to genetic kidney diseases such as Alport syndrome and ADPKD, this condition is underrepresented in registries such as Genomics England or UK Biobank (Devuyst *et al.*, 2014; Wong *et al.*, 2024). Consequentially, this is the largest cohort encompassing genome-wide genetic data globally for this condition.

This cohort also benefits from many patients with the same primary disease-causing variant (*UMOD* p. Val93_gly97delinsAlaAlaSerCys) which allows for a cleaner dataset without confounding of the primary *UMOD* variant, however this smaller sample size was less statistically powered in a GWAS and there was more confounding in-terms of the cohort being more highly related (Sul, Martin and Eskin, 2018). Finally, as acknowledged in the previous chapter, patients with the *UMOD* p.Val93_gly97delinsAlaAlaSerCys had significantly less gout suggesting that this sub-population may represent a slightly different disease phenotype that cannot be completely extrapolated to those with other disease causing *UMOD* variants.

3.4.2 Interpretation of GWAS findings was limited by statistical power and genetic loci were subsequently prioritised on biological plausibility

The small cohort size for GWAS required genotyping with a high SNP density, imputation, clear phenotyping with rigorous end points and a kidney survival approach to GWA analysis to ensure the greatest statistical power (Bi *et al.*, 2020). Genotyping was limited by cost, so a broad-coverage SNP array was used to allow high-density genotyping of clinically relevant variants with improved imputation accuracy. However, limitations included poorer detection of rare and structural variants with such an array. Phenotyping was limited by the balance between inclusion of as many participants as possible with available clinical data. The clinical endpoints of eGFR and age were clear and well defined however many patients lacked more than one serum creatinine or eGFR value because they were young or had a milder disease stage. Using slope of eGFR decline necessitates a binary approach and requires more than one serum creatinine or eGFR value per participant and, for some, from many years ago (pre-kidney replacement therapy) for which this data was not available. Using a presumptive eGFR of, say 100mls/min/1.73^2 at birth, was not chosen due to the wide variability of nephron endowment amongst individuals and not thought to be accurate (Charlton *et al.*, 2021). Furthermore, performing a binary approach (e.g. fast versus slow kidney function progression) to phenotyping meant that larger numbers would be needed to gain statistical power, so a kidney survival approach was used instead (Bi *et al.*, 2020).

Survival models can enhance power to detect genetic effects on the trait by incorporating time-to-event data making data utilisation more efficient. Individuals who have not yet reached the end point of kidney failure can be included using 'censoring' and both 'endpoint' and 'time' are included in the statistical model. Not all participants will reach the end point of kidney failure, leading to right-censoring and limiting the number of informative cases. A further limitation of this approach is the underrepresentation of young individuals with rapidly progressive disease as they have not yet had sufficient follow up to develop eGFR decline and may be censored as 'non-progressors' due to their early disease stage. This could lead to bias in effect estimates and reduce power to detect genetic variants influencing early disease trajectories (Abd ElHafeez *et al.*, 2021). An earlier end point such as CKD stage 4 was considered to try to reclaim some of this underrepresentation however, with many

participants being years post-kidney replacement therapy, this early data in their disease course was not available.

The quality control pipeline developed effectively ensures data quality because stringent filters were used to remove low-quality individual and variant DNA. Of course, the removal of rare variants can lead to loss of potentially biologically relevant findings, but due to the frequently found clinically evident difference in kidney survival even within families **(Figures 6 & 7)** I was expecting most detectable contributing variants to the trait to be common (Turner *et al.*, 2011).

Proposed clinical modifiers of disease progression include biological sex **(figure 29)**, maternal age of end stage kidney failure and primary *UMOD* variant (Kidd *et al.*, 2020). Ten principal components of relatedness were introduced as covariates to reduce the confounding effects of this highly related cohort as well as effects of the primary *UMOD* variant due to differential effects on kidney survival **(figures 25-30).** Maternal age of ESKD was not included as a covariate simply due to lack of this data but could reflect mitochondrial, environmental and epigenetic factors on disease progression in ADTKD-*UMOD*. Biological sex is not a necessary GWAS covariate as it does not confound genetic associations except for SNPs on the sex chromosomes. While adjusting for sex can sometimes reduce variability and increase power, it can reduce power in certain cases. Many studies have conducted a GWAS without sex adjustment and perform a secondary analysis to include sex, often finding minimal differences in results. If the goal is to assess sex-specific genetic effects, this would require either testing for SNP x sex interactions or analysing males and females separately (Khramtsova *et al.*, 2023).

Although no loci approached genome-wide significance in any of the GWAS **(figures 31-34)**, selecting SNPs approaching this threshold and filtering them using regional LocusZoom plots ensured potentially relevant variants were supported by strong association signals whilst acknowledging the overall sample size was small and potentially underpowered to detect effects (Boughton *et al.*, 2021). The prioritisation score based on functional criteria and kidney expression **(figure 22)** allowed for objective ranking and cross-GWAS comparison, increasing confidence in selected candidates considering that the disease is limited to the kidney. However, limitations are that excluding SNPs without LocusZoom support may eliminate biologically relevant variants and bias towards well annotated genes potentially overlooking

novel or rare variants. Furthermore, reliance on pre-defined scoring criteria may miss functionally important SNPs not captured in databases. Whilst this approach refines SNP selection based on biological candidacy for loci where genome-wide significance is not achieved, experimental validation remains essential to confirm the role of these SNPs in disease progression (French and Edwards, 2020).

3.4.3 The UMOD promoter is not a GWAS signal, but may still modify disease in cis

In the GWAS of kidney survival among 266 individuals with genetically confirmed ADTKD-*UMOD*, no SNPs within the *UMOD* gene, its promoter region, or the adjacent *PDILT* locus reached genome-wide significance **(figures 37 & 38)**. This includes the well-characterised promoter variant rs4293393, previously shown to reduce mutant *UMOD* expression and associate with improved kidney outcomes in population-based studies (Kidd *et al.*, 2020). The lack of association in this context is likely attributable to a key limitation of GWAS and its inability to determine allelic phase, however, when the *UMOD* and its promotor was assessed for its contribution to a polygenic risk score for CKD in the cohort to predict kidney survival, its contribution was incredibly small **(section 3.3.16)**. In ADTKD-*UMOD*, only promoter variants located *in cis* with the pathogenic *UMOD* variant are hypothesised to attenuate disease severity through reduced mutant protein expression (Kidd *et al.*, 2020). Standard GWAS approaches are limited in their ability to detect cis-regulatory effects because they do not determine the phase, for example, whether variants are located on the same chromosome (*in cis*) or on opposite chromosomes (*in trans*). In the context of ADTKD-*UMOD*, where the promoter variant rs4293393 may only modify disease if it resides on the same allele as the pathogenic *UMOD* variant, cis-phase information is essential. Without knowing whether the promoter variant is *in cis* or *trans*, the protective effect of reduced mutant uromodulin expression cannot be accurately captured. Cis-phase aware analyses, using long-read sequencing, family-based phasing, or statistical phasing, are therefore required to assess the true impact of regulatory variants in monogenic diseases like ADTKD-*UMOD*. Kidd *et al* showed that at rs4293393, the minor allele was significantly underrepresented in the cohort compared to population frequency (11.6% vs. 19%, $P < 0.01$) which resulted in Hardy-Weinberg disequilibrium and preventing Mendelian randomisation analysis. These findings demonstrate the limitations of GWAS for uncovering certain modifiers in rare autosomal dominant conditions and highlight the importance of cis-phase aware analyses to elucidate the full impact of regulatory variants.

3.4.4 Gene burden testing uncovered an important genetic locus on chromosome 19 and potential roles of renal uric acid transporter genes

Gene-burden testing offers advantages beyond GWAS. Unlike GWAS, which identifies more common variants of smaller effects, gene burden testing aggregates multiple rare variants within a gene which both increases statistical power and reduces the multiple testing burden. This approach captures the cumulative effects of functionally relevant variants making it more biologically meaningful, particularly in protein-coding regions (Guo *et al.*, 2016). Some diseases may be influenced by oligogenic contributions of multiple rare variants rather than a single common variant which GWAS might overlook. Gene-burden testing can detect these and provide clinical relevance by highlighting genes with potential clinical relevance. Due to the aggregation of functionally relevant variants, clearer insights are provided for precision medicine, drug targeting and clinical risk stratification compared to GWAS where loci may not have clear mechanistic links. Of course, disadvantages include the inclusion of coding variants only, which means regulatory elements can be overlooked. The assumption that all variants in a gene contribute in the same direction to the trait when effects may be discordant can result in true associations being diluted (Guo *et al.*, 2016). **Table 15** highlights the top prioritised 'candidate genes' from gene-burden studies.

The locus on chromosome 19 reaches the genome-wide significance threshold in the 266-person imputed cohort gene-burden study for association with the trait of kidney survival **(figure 41)**. This signal is also evident, although less statistically significant in the full 266-person non-imputed cohort and the imputed cohort of 92 patients with the same *UMOD* p. Val93_gly97delinsAlaAlaSerCys variant **(figure 42)**. Interestingly, three uric acid transporter encoding genes stood out as approaching genome-wide significance. *ABCG2* appeared as gene of interest on both imputed and non-imputed full 266-person cohorts with ADTKD-*UMOD* (imputed cohort p= 1.0721e-04) **(figures 39 & 41)**. *SLC22A6* (imputed cohort p=6.3718e-05) and *SLC22A8* (imputed cohort p=1.8117e-04) appeared as genes of interest in both imputed and non-imputed 92-person cohorts with the same *UMOD* p. Val93_gly97delinsAlaAlaSerCys variant **(figures 40 & 42)**. This was clinically highly relevant as young onset hyperuricaemia and gout are hallmarks of the disease. The lack of *ABCG2* as a hit in the 92-person *UMOD* p. Val93_gly97delinsAlaAlaSerCys cohort is interesting as gout is much less prevalent in this

disease sub-group and this different phenotype may reflect differing disease mechanisms between genetic sub-types of ADTKD-*UMOD*. This is explored in more detail in **Chapter 4**.

3.4.5 GWAS and Gene Burden Studies suggest disease progression may be modified by genes involved in uric acid biology, mitochondrial homeostasis and epigenetic regulation

Focusing on the locus of interest on chromosome 19, three genes of interest appear from this locus as candidates from the gene burden studies: *KMT5C, COX6B2 and AC020922.1* **(table 15, figures 39-43)**. When the LocusZoom **(figure 44)** plot of this GWAS locus is analysed, the most statistically significant SNPs occur in *KMT5C*. Conditional analysis was subsequently required to distinguish true variants from this locus associated with the trait of kidney survival in this context **(table 16)** (Yang *et al.*, 2012).

KMT5C (SUV420H2) encodes the histone methyltransferase, lysine methyltransferase 5C, which trimethylates nucleosomal histone H4 on lysine-20 and thus regulates transcription and maintenance of genome integrity. H4 'Lys-20' trimethylation represents a specific tag for epigenic transcriptional repression and mostly functions in pericentric regions of heterochromatin (Wu *et al.*, 2013). Lysine methyltransferase 5Cmis targeted to histone H3 via interaction with RB1 family proteins and, when DNA damage occurs, facilitates TP53BP1 foci formation and non-homologous end-joining (NHEJ)-directed DNA repair by acting as the catalyst for di- and trimethylation of 'Lys-20' of histone H4 (Bromberg *et al.*, 2017). Given the kidneys, particularly tubular epithelial cells, are known to show rich transcriptional activity as they rapidly adapt to metabolic stress and finely tune tubular transporter gene expression, these actions require a dynamic epigenetic landscape (Tanemoto, Nangaku and Mimura, 2022). KMT5C-mediated chromatin modifications may underpin this rich regulatory landscape (as illustrated in **figure 46)** necessary for nephron function. Disruption of this pathway could result in impaired DNA repair, altered gene expression and contribute to tubular injury or fibrosis, offering a potential epigenic link to CKD progression (Tanemoto, Nangaku and Mimura, 2022).

COX6B2 encodes Cytochrome C Oxidase Subunit 6B2 which is the final enzyme in the mitochondrial electron transport chain which drives oxidative phosphorylation. *COX6B2* has been associated with Bardet-Biedl Syndrome 18 (BBS18), an autosomal recessive ciliopathy

characterised by retinitis pigmentosa, obesity, kidney failure and cognitive disability where chronic tubulointerstitial nephropathy is also seen on kidney biopsy. In addition to this *COX6B2* has also been associated with mitochondrial disease. Mitochondrial dysfunction has been implicated repeatedly in our early understanding of the patho-mechanisms of ADTKD-*UMOD*. TAL cells have demonstrated a reduction in mitochondria and mitochondrial proteins along with disrupted activation of mitochondrial transcription factor NFR1. Reduced oxidative phosphorylation and AMPK-pathway activation further suggests impaired energy homeostasis and mitochondrial dysfunction in the disease (Kemter *et al.*, 2017). Further work has shown that mesencephalic astrocyte-derived neurotrophic factor (MANF), an ER secreted protein, stimulated autophagy/mitophagy, clear mutant uromodulin and promotes mitochondrial biogenesis in ADTKD-*UMOD* (Kim *et al.*, 2023). The potential role of *COX6B2* adds weight to the potentially critical role of mitochondrial quality and heath in mitigating tubular injury in ADTKD-*UMOD* and advances our focus on mitochondrial pathways as potential therapeutic targets for disease modification.

SNPs in *ABCG2* appeared as candidate SNPs of interest from both the full-266 person ADTKD-*UMOD* GWAS **(figures 31, 35 & 36, table 13)** and the gene-burden studies **(figures 39 & 41)** with the same cohorts reinforcing association of this locus with the trait of kidney survival. *ABCG2* encodes the ATP-binding cassette subfamily G member 2 transporter protein which play a crucial role in both drug metabolism and uric acid excretion, including in the human kidney (Ohashi, Toyoda, *et al.*, 2023). Furthermore, *SLC22A6* and *SLC22A8* which were candidate genes of interest from both *UMOD* p. Val93_gly97delinsAlaAlaSerCys cohort gene-burden studies **(figures 40 & 42)**, encode Organic Ion Transporters 1 (OAT1) and 3 (OAT3) respectively which are both key players in renal tubular regulation of uric acid (Nigam, 2018). As previously mentioned, this intriguing relevance of uric acid regulation to the phenotypic hallmark of hyperuricaemia in ADTKD-*UMOD* warrants further exploration and is explored and discussed in detail in the next chapter **(Chapter 4).**

3.4.6 Conditional analysis was limited by cohort relatedness and did not discriminate a single contributing SNP on the chromosome 19 locus

To determine if the leading SNPs on the chromosome 19 locus were independent genetic signals, GCTA-COJO was used to perform conditional analysis of the top four SNPs of interest: *rs12459907, rs79012440, rs1870073* and *rs12610863* **(table 16)** (Yang *et al.*, 2012). A

challenge with performing this method was that the 'Betas' and 'Standard Errors' provided to GCTA-COJO are not necessarily the "correct" quantities expected by the software (namely the logOR and standard error from logistic regression analysis of a case/control study), as they were derived from a completely different type of analysis, namely SPACox. Furthermore, the 'n' column output by GCTA_COJO did not reflect the full cohort sample size of 266 individuals, possibly because GCTA-COJO used the estimated effective sample size rather than the actual cohort sample size due to the relatedness of the cohort. Subsequently, an allele dosage method was used in PLINK to directly condition SNPs on the allele dosage, as a covariate, of each of the four SNPs of interest in turn (Chang *et al.*, 2015). This allowed a direct comparison of p-values and effect sizes of the remaining 3 SNPs of interest when the 4th was conditioned on. Since neither of these two methods fully accounts for relatedness within the cohort, a linear mixed models (LMM) approach would be more appropriate as it adjusts for shared genetic background using a kinship matrix, though LMM is not widely available for time-to-event analyses such as this (Wang, Aragam and Xing, 2022).

Theoretically, related individuals share more alleles than would be expected by chance, leading to inflation of test statistics if population structure and relatedness are not properly adjusted for. Interestingly, all four GWAS analyses showed a genomic inflation factor of <1, suggesting mild deflation despite the cohort being highly related. This likely reflects the small sample size and the effectiveness of principal component (PC) adjustment in the initial SPACox models (Sul, Martin and Eskin, 2018). While GCTA-COJO does not directly model relatedness or population structure, it operates on summary statistics that were generated from PC-adjusted models, and thus indirectly incorporates these adjustments. However, a potential limitation of COJO is that it may not fully account for the model assumptions underlying those summary statistics, especially for more complex or survival-based models. For this reason, the second PLINK allele dosage method, which directly adjusts for relatedness/population structure using PCs at the individual level, was preferred due to its clearer and more transparent treatment of these confounders.

After conditioning, using both methods, none of the four SNPs of interest remained significant after conditioning suggesting that none of the four SNPs had an independent effect on the trait. This could indicate that the chromosome 19 locus is one primary signal rather than

multiple independent associations of these four SNPs. It could, alternatively, reflect limitations in the sample size and power to detect any secondary signals.

3.4.7 Chromatin interaction mapping uncovered a biologically interesting regulatory pathway between ABCG2 and PKD2

Chromatin interaction mapping enables the identification of both long- and short-range regulatory loops that extend beyond proximity-based annotation, enhancing the functional interpretation of genome-wide association findings (He and Bonasio, 2017). This approach has been instrumental in elucidating regulatory architectures at loci associated with uric acid transport and kidney disease (Leask *et al.*, 2020). For instance, chromatin conformation analysis revealed a biologically compelling interaction between *ABCG2* and *PKD2* **(figure 45)**, suggesting a shared transcriptional network relevant to renal tubular physiology.

The three intronic SNPs identified in *ABCG2*; rs2622627, rs17731799, and rs3114020 **(table 13)**, lie within regions enriched for enhancer activity in kidney, epithelial, and haematopoietic cells. SpliceAI did not predict splicing effects, suggesting that the functional relevance of these SNPs is regulatory rather than coding (de Sainte Agathe *et al.*, 2023). Their localisation to active enhancer elements and the chromatin loop formation between *ABCG2* and *PKD2* supports a model in which this region may act as a cis-regulatory hub governing epithelial gene programmes related to tubular function, urate handling, and possibly cystogenesis (Reiterová and Tesař, 2022).

Although *ABCG2* is predominantly expressed in the gastrointestinal tract and immune cells such as macrophages, it is also localised to the apical membrane of proximal tubule cells, where it contributes to solute and xenobiotic excretion (Robey *et al.*, 2009). By contrast, *UMOD* is not directly expressed in the proximal tubule, raising questions about how *ABCG2*-mediated effects might influence disease progression in ADTKD-*UMOD*. One possible explanation lies in the concept of TAL–S3 cross-talk, wherein basolateral release of uromodulin (THP) from the thick ascending limb exerts paracrine anti-inflammatory effects on neighbouring S3 segments of the proximal tubule (El-Achkar and Dagher, 2006; El-Achkar and Wu, 2012; Micanovic *et al.*, 2020). This mechanism, supported by microscopy and animal models of ischaemic injury, highlights a plausible anatomical and signalling pathway through which *ABCG2* and *UMOD* could interact indirectly (El-Achkar *et al.*, 2013; Micanovic *et al.*, 2018).

In this context, experimental evidence from (Cristóbal-García *et al.*, 2015) is noteworthy, demonstrating that chronic hyperuricaemia in rats induces renal cortical oxidative stress, mitochondrial dysfunction, and tubular damage; effects prevented by antioxidant treatment. While, short-term, hyperuricaemia induced hypertension without affecting mitochondrial function, long-term exposure resulted in reduced ATP production and enhanced oxidative injury, emphasising the relevance of chronic urate burden to tubular energetics.

These findings are consistent with human cohort studies showing that elevated serum urate, even within the high-normal range, may accelerate kidney function decline. In a Japanese cohort of healthy men (Kuma *et al.*, 2021), serum urate levels above 7.0 mg/dL were associated with a greater fall in eGFR over five years, and urate reduction attenuated this decline.

Additionally, the identification of *ABCG2* eQTLs in blood **(section 3.3.14)** raises the possibility of an immunomodulatory role, whereby monocyte/macrophage lineage cells influenced by *ABCG2* may contribute to renal interstitial inflammation or repair processes. Such immune-mediated mechanisms could represent novel pathways in ADTKD-*UMOD* pathogenesis, though functional validation is required. The potential for immune involvement is supported by prior studies linking ABCG2 dysfunction to altered macrophage trafficking and cytokine production, and by the work of (Ohashi, Kuriyama, *et al.*, 2023; Ohashi *et al.*, 2023) who found that individuals with impaired ABCG2 function experienced faster CKD progression even in the absence of gout.

The observed enhancer interaction between *ABCG2* and *PKD2* is particularly notable given the established role of *PKD2* in autosomal dominant polycystic kidney disease (ADPKD) (Reiterová and Tesař, 2022). *PKD2* encodes polycystin-2, a cation channel essential for tubular architecture and flow sensing, and its interaction with *ABCG2* could suggest convergence of regulatory programmes underlying tubular maintenance, inflammation, or cyst development (Wang, Kang and Xie, 2024). While cysts are not universally present in ADTKD-*UMOD*, their occurrence in approximately one-third of affected individuals raises the possibility of shared pathways with ADPKD. Although genotype–phenotype analysis of *ABCG2* variants and cyst prevalence in this study was underpowered (n=9), future studies with expanded imaging datasets are warranted.

Further support for the pathological significance of *ABCG2* variants comes from a study of familial paediatric-onset hyperuricaemia (Toyoda *et al.*, 2021) which identified two rare dysfunctional alleles (p.M131I and p.R236X) that abolished or severely impaired urate transport. These variants segregated with early-onset hyperuricaemia and gout across three generations and highlight how *ABCG2* loss of function can drive urate accumulation from a young age, even in the absence of secondary causes.

Finally, the pharmacological landscape of *ABCG2* modulation offers translational potential. Based on the work by (Robey *et al.*, 2009), several *ABCG2* inhibitors and substrates have been described that could modulate its activity and downstream effects:

Candidate drugs that inhibit or interact with *ABCG2*:

- Tyrosine kinase inhibitors (TKIs): such as imatinib, nilotinib, and gefitinib
- mTOR inhibitors: including everolimus and sirolimus
- Flavonoids: such as quercetin and genistein
- Xanthine oxidase inhibitors: allopurinol may have indirect effects by altering uric acid levels transported via *ABCG2*
- Other *ABCG2* inhibitors under study: Ko143, fumitremorgin C (experimental)

While several of these agents are already approved for other indications, further study is needed to assess their safety and efficacy in the context of ADTKD-*UMOD*. Of particular interest are drugs that may modulate *ABCG2*-mediated urate transport or, separately, immune cell trafficking, offering novel avenues for disease modification, however the next step must be to validate this signal in a larger genotyped cohort followed by mechanistic studies (Robey *et al.*, 2009).

3.4.8 A polygenic risk score for CKD is associated with kidney survival in those with ADTKD-UMOD

The integration of a genome-wide polygenic score (PGS) predictive for CKD progression with the ADTKD-*UMOD* cohort genetic data suggested that the PGS was significantly associated with kidney survival in those with the *UMOD* p. Val93_gly97delinsAlaAlaSerCys **(table 17, figures 47-51)** (Khan *et al.*, 2022). These findings underscore the complex genetic architecture of monogenic kidney disease and suggest that monogenic variants alone do not fully

determine disease severity. The interplay between rare high-impact variants and polygenic background explains some of the variability in penetrance within families and those with the same disease-causing variant. The implications of this on risk stratification and clinical management are considerable as polygenic scores could be integrated with monogenic screening and improve accuracy of disease severity prediction (Khan *et al.*, 2023). Application of this score prospectively in a secondary cohort of individuals with ADTKD-*UMOD* would be an excellent validation for its predictive ability for disease progression. It would also be useful to see if the PGS is predictive for disease progression in a larger cohort of participants with mixed ADTKD-*UMOD* variants.

3.4.9 The recurrent UMOD p. Val93_gly97delinsAlaAlaSerCys variant represents a probable founder variant dominant in the British population

This analysis provides robust evidence that the recurrent *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant represents a fairly recent founder variant within the British population **(section 3.3.17)**. While Smith *et al.*, 2011, initially proposed a variant hotspot model based on limited haplotype overlap, Valluru *et al.*, 2023, later suggested a UK-specific founder effect across 22 families, though their conclusions were constrained by the absence of formal phasing and variant age estimation. By applying genome-wide phasing (SHAPEIT2) (Delaneau *et al.*, 2013) and likelihood-based estimation of the most recent common ancestor (ESTIAGE) (Genin *et al.*, 2004), the current study formally characterises a shared haplotype spanning 20 kb, with the longest extending 477123 base pairs in length. The founder event is estimated to have occurred just three generations ago (95% CI: 2–4 generations). In contrast to classical founder variants such as the *BRCA1* c.5207T>C in Orkney or *BRCA2* c.517-2A>G in Shetland, which are associated with shared haplotypes spanning 3–9 Mb and arose 10–20 generations ago (Kerr *et al.*, 2024), the relatively short haplotype observed here is striking given the recent timing. This may reflect a localised recombination hotspot near the variant or complex ancestral structure such as inheritance through multiple, partially overlapping genealogical lineages or unrecognised pedigree loops. Such patterns are common in endogamous populations and can accelerate haplotype decay even in the context of a recent founder event. Comparable phenomena have been observed in other rare disease contexts, including the Finnish disease heritage (Peltonen, Jalanko and Varilo, 1999), the French-Canadian *BRCA1* C4446T and *BRCA2* 8765delAG variants (Tonin *et al.*, 1998), and the Icelandic *BRCA2* variant (Thorlacius *et al.*, 1996), where founder effects have led to elevated

regional carrier frequencies and extended shared haplotypes. However, unlike these classical examples, typically marked by longer shared haplotypes due to older origin or reduced recombination, the *UMOD* founder variant described here displays a paradoxically short core haplotype despite its recency. Importantly, shared haplotype structure was also observed in a subset of US individuals, demonstrating how rare disease genomic cohorts can resolve the genetic architecture of monogenic diseases. This builds upon and refines earlier observations (Devuyst *et al.*, 2019; Olinger *et al.*, 2020) highlighting the utility of international comparative genomics in disentangling founder versus recurrent variant models. These findings underscore the power of integrating international datasets for validating founder events, particularly when haplotype length alone might misleadingly suggest older origins and provide a compelling model for how national rare disease networks can uncover clinically significant population-specific variants of recent origin.

3.5 Conclusion

The combination of GWAS, gene-burden studies, polygenic risk score application and gene regulation analysis have enabled the determination of genetic factors associated with disease progression in individuals with ADTKD-*UMOD*. A validation cohort and larger numbers of patients would be ideal to confirm these findings. Indeed, the variability in disease progression in ADTKD-*UMOD* cannot be fully explained by genetic factors as there are many contenders at play that are known to influence kidney disease risk, penetrance and severity. These are multifaceted and include environmental exposure, foetal prematurity, nephron endowment, prenatal exposures, birth weight and mitochondrial health (Starr and Hingorani, 2018; Good *et al.*, 2023; Ghelichi-Ghojogh *et al.*, 2024; Oulerich and Sferruzzi-Perri, 2024; Pavlović *et al.*, 2025) . The degree of contribution of genetic factors to this is difficult to truly determine but this work does further insights into kidney biology, disease patho-mechanisms and potential therapeutic targets for ADTKD-*UMOD*.

The findings in Chapter 3 suggest that disease progression in ADTKD-*UMOD* is influenced by a complex interplay between polygenic background, and loci involved in uric acid handling, mitochondrial function, and epigenetic regulation. Particularly, *ABCG2* emerged as a putative genetic modifier of kidney survival. These observations raise important questions about the metabolic signatures of ADTKD-*UMOD* and whether distinct sub-phenotypes, such as hyperuricaemia, might reflect underlying pathophysiological mechanisms. Chapter 4 builds on

this by exploring uric acid metabolism in detail, examining the relationship between serum urate levels, genetic variants in urate transporter genes, and kidney survival. By comparing patient-level and population-level data, this chapter aims to clarify the role of uric acid as both a diagnostic marker and a potential contributor to disease heterogeneity in ADTKD-*UMOD.*

Chapter 4. Uric Acid and Disease Implications in ADTKD-*UMOD*

4.1 Introduction

Uric acid (2,6,8-trihydroxypurine), the end product of purine metabolism in humans, has an intriguing dualistic action; it is a powerful central nervous system antioxidant, yet has pro-oxidant properties in the blood (Nieto *et al.*, 2000; Wang, Wen and Kong, 2020). It is well stablished that hyperuricaemia can trigger crystal formation, leading to gout and has been implicated in a range of conditions such as CKD, cardiovascular disease, autoimmune disease and the metabolic syndrome (Edwards, 2009; Keskin *et al.*, 2021; Sun *et al.*, 2021). Hyperuricaemia is also a cause of kidney stones and acute kidney injury from tumour lysis syndrome (Johnson *et al.*, 2018; Adomako and Moe, 2020). In contrast, hypouricaemia has been associated with an increased risk of neurodegenerative disorders such as Parkinson's and Alzheimer's disease suggesting the requirement of uric acid for physiological health (Ascherio *et al.*, 2009; Ye *et al.*, 2016). Unlike most mammals, humans and other great apes lack the enzyme uricase, which normally converts uric acid into the more soluble allantoin, leading to inherently higher serum urate levels in these species, an evolutionary loss thought to confer antioxidant advantages but at the cost of increased susceptibility to urate-related diseases (Chang, 2014).

Globally, gout affects 1-4% of the population (incidence range 0.1-0.3%) and both statistics increase with age reaching 11-13% prevalence (0.4% incidence) among those over 80 years old (Han *et al.*, 2024). Gout is extremely rare in children and adolescents and usually occur in the context of inherited renal and metabolic disorders. Serum uric acid levels are influenced by genetic predisposition, gene-environment interactions, environmental exposures, age and sex and heritability estimates are between 38% and 63% (Yang *et al.*, 2005; Stiburkova and Ichida, 2025).

During infancy, serum uric acid levels are low (131–149 µmol/L) which accompany a high fractional excretion of uric acid (>90%). The fractional excretion of uric acid decreases to 8% by age 1 and serum levels rise to (208–268 µmol/L) and fraction excretion further decreases in boys after age 12 but remains unchanged in girls (Stiburkova and Bleyer, 2012). Endogenous uric acid production is influenced by purine intake, de novo synthesis, degradation and recycling. Impaired excretion is the primary driver of hyperuricaemia in the general population

and as little as a 1% reduction in excretion can results in almost a 5-fold increase in serum uric acid concentration. Deficiencies in purine metabolic enzymes have a very small impact on serum uric acid concentration due to minimal effects on total uric acid production, however even minor effects on renal or intestinal uric acid transporter function can lead to a clinically significant increase or decrease in serum uric acid (Stiburkova and Ichida, 2025). This is reinforced by a meta-analysis of over 16 000 individuals which demonstrated that 23.9% of hyperuricaemia risk of accounted for by genetic variation in contrast to <0.3% attributed to dietary factors (Major *et al.*, 2018).

Hyperuricaemia can be classified into two types: 'Underexcretion type' typically due to reduced renal clearance, often due to genetic variants in SLC2A9 and ABCG2 and, in rare instances, d-lactate dehydrogenase abnormalities may contribute to increased reabsorption via URAT1; 'Overproduction type' is driven by increased purine turnover as seen in HPRT deficiency (Lesch-Nyhan Spectrum), PRPS1 (Phosphoribosyl pyrophosphate synthetase 1) superactivity or increased consumption of ATP. Those with Glycogen Storage diseases have a combination of increased uric acid production and reduced renal excretion secondary to lactic and ketoacidosis. Hyperuricaemia may also be drug-induced (e.g. thiazides, valproate, ciclosporin and phenobarbital), secondary to acute illness (e.g. gastroenteritis, anaemia or malignancy). In the case of gastroenteritis, intestinal inflammation and dehydration can impair ABCG2-mediated uric acid excretion (Stiburkova and Ichida, 2025). Most other species, metabolise uric acid to allantoin, urea or ammonia via Urate oxidase so hyperuricaemia does not reflect what we see in humans and higher primates (Johnson, Lanaspa and Gaucher, 2011). Asymptomatic hyperuricaemia is commonly seen in both paediatric and adult populations. In a paediatric 10-year retrospective study of 33 900 children, hyperuricaemia had a prevalence of 12.6% (Stiburkova, Lukesova and Zeman, 2025).

Gout is the most common inflammatory arthritis and has a prevalence of nearly 4% in the US (5.9% men and 2% women) but some populations including Māori, Pacific Islanders and Indigenous Taiwanese have over twice the average prevalence (Zhu, Pandya and Choi, 2011; Sumpter *et al.*, 2023). Gout onset in the general population occurs between the fourth and sixth decades with a mean age of 61.9 in the UK (Rothenbacher *et al.*, 2011). As oestrogen promotes urate excretion, post-menopausal women have an increased risk of gout as oestrogen declines. As elevated fructose causes ATP depletion and increased uric acid

generation, gout incidence has been rising due to the global concerning rise in fructose-rich ultra-processed diets (Stiburkova and Ichida, 2025). Although very rare in paediatric populations, gout can develop following persistent hyperuricaemia. The UK General Practice Research Database reported an incidence of 12 cases per 255 950 males and 1 case per 246 364 females under the age of 25 (Mikuls *et al.*, 2005). In Japan, a study of 700 000 individuals under age 18, found a prevalence of 0.033% for asymptomatic hyperuricaemia and 0.007% for gout (Kim *et al.*, 2017).

Uric acid is tightly regulated by both renal (eliminating 2/3 uric acid) and intestinal transporters (eliminating 1/3 uric acid) that mediate both reabsorption and excretion. Adults produce about 1000mg of uric acid per day under stable conditions. In the kidney, roughly 90% of filtered uric acid is reabsorbed in the proximal convoluted tubule (PCT), primarily through URAT1 (*SLC22A12*), GLUT9 (*SLC2A9*) an OAT4 (*SLC22A11*). Uric acid is secreted in PCT via transporters such as OAT1 (*SLC22A6*), OAT3 (*SLC22A8*), NPT1 (*SLC17A1*), NPT4 (*SLC17A3*) and ABCG2 **(figure 52)** (Stiburkova and Ichida, 2025).

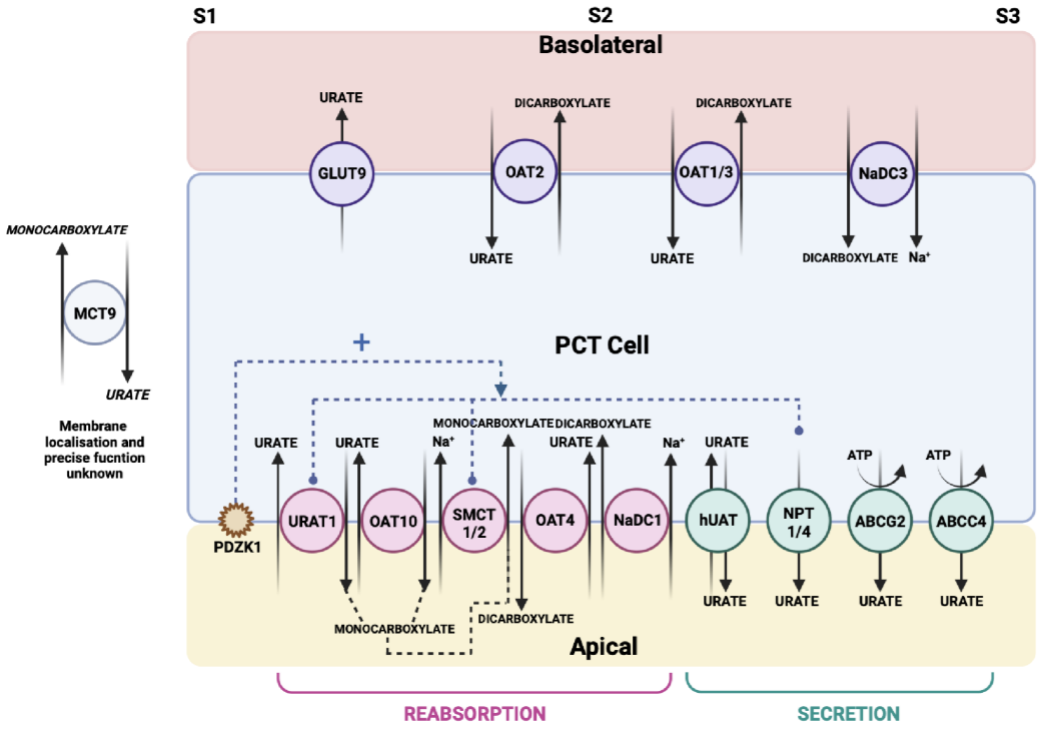

Figure 52: Uric Acid Transport in the Kidney Proximal Convoluted Tubule*

*[*The current understanding of uric acid handling in the proximal tubule within the kidney. The circles represent the key uric acid transporters on the apical and basolateral membranes. Uric acid reabsorption is primarily mediated by URAT1, OAT10 and GLUT9 whilst uric acid secretion*

mostly occurs via OAT1/3, OAT2, NPT1/4, ABCG2 and ABCC4. The sodium-coupled transporters NaDC1 and SMCT1/2 work with the Organic Anion Transporters to regulate balance of uric acid across the renal epithelial cells. Mechanisms remain completely understood but include additional regulatory proteins such as PDZK1 and MCT9 (Bobulescu and Moe, 2012).]

It has been well demonstrated that genetic variants in these transporters significantly influence and result in heritable differences in serum uric acid levels as well as susceptibility to gout, CKD and cardiovascular disease. Loss-of-function variants in *SLC22A12* (encoding URAT1) and *SLC2A9* (encoding GLUT9) can cause familial renal hypouricaemia, or conversely, polymorphisms in *ABCG2* have been linked to early-onset gout thus clinical effects depend on whether the variant impairs or enhances uric acid reabsorption (Stiburkova and Ichida, 2025). This also reinforces the potential for genetic screening to establish gout risk so specific therapies and be appropriately tailored. Furthermore, gaining understanding of the mechanistic genetic and molecular architecture of these transporters has resulted in development of uricosuric drugs as well as repurposing existing drugs for their therapeutic potential on hyperuricaemia such as losartan that inhibits URAT1. Current strategies to either reduce uric acid production or enhance its excretion include Xanthine oxidase inhibitors (e.g. allopurinol and febuxostat) which reduce uric acid production, Uricosurics (e.g. benzbromarone and lesinurad) which promote excretion and newer dual inhibitors of Xanthine oxidase and URAT1 (e.g. dotinurad and verinurad) which are under investigation. Uricases such as pegloticase are used in refractory cases to enzymatically degrade uric acid to Allantoin (Sun *et al.*, 2021). Better understanding of hyperuricaemia and gout in ADTKD may further insights into precision-based treatment of hyperuricaemia and gout in the disease in addition to hastening discovery of potentially new molecular pathways in the disease biology.

Non-renal pathways, principally intestinal excretion of uric acid via ABCG2 become increasingly physiologically important as kidney function declines in CKD as this route provides a compensatory mechanism for uric acid elimination and has thus become a focus of more recent research (Stiburkova and Ichida, 2025).

Hyperuricaemia and gout are well-recognised and described features in those with ADTKD-*UMOD*. The prevalence, age of onset and family history of gout in ADTKD-*UMOD* has been examined in various cohorts internationally. The largest to date (726 individuals with ADTKD)

suggested an overall gout prevalence of 79% in those with ADTKD-*UMOD* with a median age of gout onset of 27 years. Gout is more prevalent in males compared to females in all cohorts analysed, and the Wake Forrest cohort demonstrated that gout typically developed between the ages of 15-40 (Olinger *et al.*, 2020). The presence of gout was not associated with age of ESKD however if gout occurred at a younger age, this was associated with a younger age of ESKD (Kidd *et al.*, 2020).

The mechanistic nature of hyperuricaemia and gout in ADTKD-*UMOD* is limited to hypotheses at present based on clinical observations and early experimental models. A proposed explanation is based on the impaired urinary concentration ability from TAL Na-K-2Cl cotransporter dysfunction in the disease leading to polyuria and polydipsia. The subsequent fluid loss and volume contraction results in compensatory sodium reabsorption in the proximal tubule with concurrent increased Na-dependent urate reabsorption and subsequent hyperuricaemia (Scolari *et al.*, 2004; Liu *et al.*, 2018). This is supported by a reduced Na-K-2Cl cotransporter activity in aged *UMOD* KO mice. Interestingly though, gout is considerably less prevalent in ADTKD-*MUC1* where patho-mechanisms also originate in the distal tubule leaving further questions about the mechanistic origin of hyperuricaemia and gout in ADTKD-*UMOD* (Olinger *et al.*, 2020). Gout is also observed occasionally in other salt-wasting conditions but usually as kidney function progresses and is not a unifying feature of diseases where volume contraction is common such as Gitleman and Bartter syndrome and diabetes insipidus.

Understanding the role of uric acid in CKD and whether uric acid lowering therapies may slow CKD progression remains a long-standing inconclusive debate. Experimental studies in cellular and animal models have implicated hyperuricaemia in the induction of oxidative stress, endothelial dysfunction and renin-angiotensin system activation, thereby contributing to both hypertension and kidney injury. Additionally, hyperuricaemia may influence insulin resistance and fat accumulation, linking it to the metabolic syndrome, however this is thought to be mediated by the intracellular effects of uric acid rather than crystal deposition (Johnson *et al.*, 2018). Despite strong experimental and epidemiological evidence, Mendelian Randomisation studies have not found a causal relationship between elevate serum uric acid and chronic diseases. This may be due to environmental interactions, genetic complexity and the complex differences between intracellular and extracellular effects of uric acid (Johnson *et al.*, 2018). To add to this, clinical trials exploring urate-lowering therapy (mostly allopurinol and

febuxostat) have shown mixed results, some report improvements in kidney function, blood pressure and metabolic parameters, while others find no benefit (Johnson *et al.*, 2018). Most trials have been small, short term and have used surrogate endpoints, yet a more recent large trial (PERL) also showed no benefit of uric acid lowering therapy in modifying long-term outcomes (Doria *et al.*, 2020).

As ADTKD-*UMOD* is considered a monogenic model of kidney fibrosis, the endpoint of all CKD, uncovering treatment targets for both the disease and hyperuricaemia, might further understanding of these intertwined pathways. This disease offers a unique model to dissect these mechanisms linking tubular dysfunction, hyperuricaemia and kidney disease progression. The well-defined genetic origin, young-onset hyperuricaemia and variable manifestation of gout in ADTKD-*UMOD* makes the disease an ideal candidate for targeted studies aimed at urate-lowering or tubule-specific modulation. Understanding involvement of uric acid in disease biology may clarify if uric acid or its genetic relationships are a true driver of disease progression. Advancing the mechanistic understanding of this genetically homogeneous population could help resolve the long-standing debate around the causal role of uric acid in kidney disease and the potential utility of urate-lowering therapies in CKD.

4.2 Methods

4.2.1 Serum Uric Acid and eGFR Data Acquisition

Initial serum uric acid concentration and eGFR data was acquired via the data requisition application to the RaDaR registry as described in **Section 3.2.1**. For those participants who were not included in the RaDaR registry, collaborators were contacted to provide this information.

4.2.2 UK Biobank Control Data Acquisition

Access to UK Biobank via the UK Biobank Access Management System was already granted through colleagues in the research group (‘UK Biobank’, 2022). A research proposal was submitted to gather serum uric acid, eGFR and biological sex data in line with UK Biobank’s ethics and access guidelines. Following review and approval, the data was obtained. All data handling and analyses were conducted in accordance with UK Biobank’s approved protocols and data security policies.

4.2.3 Serum Uric Acid Distribution Plotting

The relationship between serum uric acid concentration and eGFR (CKD-EPI equation) was stratified by biological sex using UK Biobank data and was implemented using the transplantr package. Uric acid and eGFR were modelled using quantile regression (10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th} percentiles) with a quadratic function for eGFR. Sex-stratified models were fitted separately for males and females and results were visualised using ggplot2. Quantile lines were overlaid and those with the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant were coloured in red.

4.2.4 Genetic Association Study Adjusting for Hyperuricaemia

A genome-wide association study was performed in the full 266-person imputed GWAS cohort using the same methods as outlined in **Section 3.2.6**, however, this time, the analysis was adjusted for an extra covariate; the presence of early onset gout (before age 40) and/or hyperuricaemia on first serum uric acid measurement. Those who had not yet reached age 40, those where their first uric acid level was elevated and corresponding eGFR was <30mls/min/1.73m^2 and those who were taking uric acid lowering therapy were coded as 'unknown' in the covariate column to eliminate selection bias and confounding. The adjusted p-values (adjusted for genomic control) for the six most statistically significant SNPs in ABCG2 for association with kidney survival were compared before and after adjustment for hyperuricemia and early-onset gout.

Gene-burden analysis was performed in the full 266-person imputed GWAS cohort using the same methods outlined in **Section 3.2.9**, however, the input file outlined in the previous paragraph where the GWAS was adjusted for hyperuricaemia and early gout was used. The adjusted p-values (adjusted for genomic control) for the ABCG2 gene were compared before and after adjustment for hyperuricemia and early-onset gout.

4.2.5 Custom Gene-set Analysis

A custom gene-set analysis was conducted using MAGMA v1.10 (Guo *et al.*, 2016). The literature was searched to determine 12 kidney-expressed uric acid transporters and their corresponding encoding genes. A custom gene set file was created manually to include these 12 genes. SNP positions were extracted from the imputed GWAS .bim file to generate a SNP

location file. The MAGMA-provided build 37 gene location file was downloaded and unzipped, and gene annotation was performed within MAGMA. Gene-level association analysis was conducted using the full 266-person imputed GWAS summary statistics and genomic-inflation factor adjusted p-value. Custom gene set analysis was executed by linking gene-level results to the gene-set file. The analysis accounted for confounders including gene size, gene density and minor allele count using MAGMA's default settings. This was repeated for two control sets, performed in the same way: one of 12 genes encoding various non-uric acid renal proximal tubule transporters and one of renal magnesium transporter encoding genes.

4.2.6 Kidney Survival in Kidney Uric Acid Transporter Genotypes

To determine if there was a genotype effect on kidney survival in the renal uric acid transporter genes, the literature was searched to determine 12 kidney-expressed uric acid transporters and their corresponding encoding genes. The genome-build 37 genomic coordinates for each gene were determined using Ensembl (McLaren *et al.*, 2016) and SNPs within these coordinates were extracted from GWAS summary statistics for the full 266-person imputed GWAS and ranked according to genomic-inflation factor adjusted p-value. The genotypes for the most statistically significant SNP associated with kidney survival in each of the uric acid transporter genes were extracted using PLINK. A Cox regression model was applied to test association of each genotype with kidney survival and survival curves were drawn using GraphPad Prism.

4.3 Results

This chapter examines serum uric acid and uric acid transporter-encoding genes as clinical and genetic modifiers of disease expression in ADTKD-*UMOD*, progressing from descriptive cohort characterisation to targeted genomic interrogation. **Section 4.3.1** establishes the distribution of baseline serum uric acid in the ADTKD-*UMOD* cohort overall, by sex, and by *UMOD* variant subgroup, and compares these distributions with population data from UK Biobank; it also explores kidney survival across extremes of serum uric acid. **Section 4.3.2** evaluates whether genomic signals associated with kidney survival are mediated through hyperuricaemia and early-onset gout, re-running imputed GWAS and MAGMA gene-burden analyses with these traits as covariates and focussing in particular on *ABCG2*-associated loci. **Section 4.3.3** extends this by testing pre-specified renal uric acid transporter gene sets (with biologically grounded control gene sets) to determine enrichment of association signals. Building on these findings,

section 4.3.4 assesses kidney survival according to genotypes at the most significant SNPs within the top contributing uric acid transporter genes in the full cohort. **Section 4.3.5** then interrogates the role of *ABCG2* variation in greater depth, contrasting effects on kidney survival between individuals carrying the common *UMOD* p.(Val93_Gly97delinsAlaAlaSerCys) variant and those with other *UMOD* variants. As in prior chapters, figures and tables are presented descriptively and sequentially; integrated interpretation, mechanistic context, and clinical implications are developed in the discussion **(section 4.4)** that follows.

4.3.1 Serum Uric Acid in ADTKD-UMOD

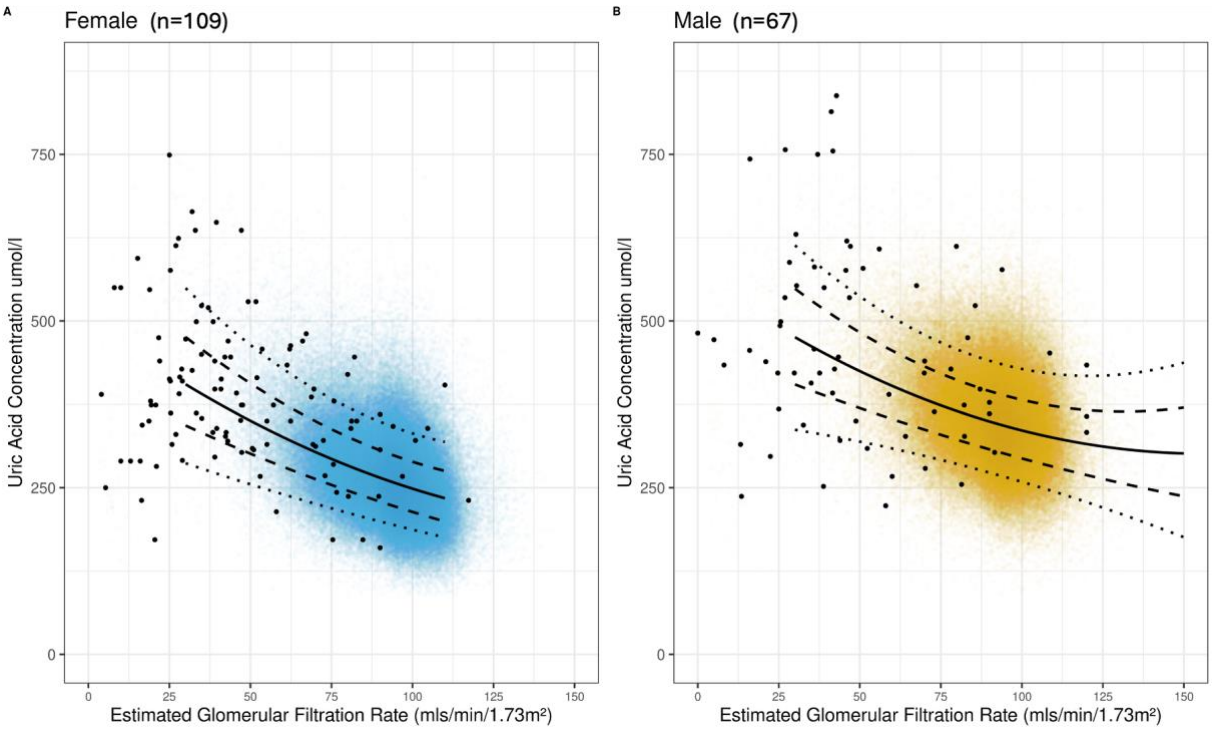

Figure 53: Distribution of Initial Serum Uric Acid Values for ADTKD-*UMOD* Cohort in comparison to Values from UK Biobank Participants*

*[*Sex specific quantile regression plots of the distribution of serum uric acid according to eGFR (female participants on the left and male participants on the right). The blue and yellow dots represent this data in UK Biobank participants in females and males respectively. The black dots represent sex specific participants with ADTKD-UMOD. Overall, males seem to have higher serum uric acid values than females and those with ADTKD tend to have higher serum uric acid levels than those reflective of the general population for both sexes.]*

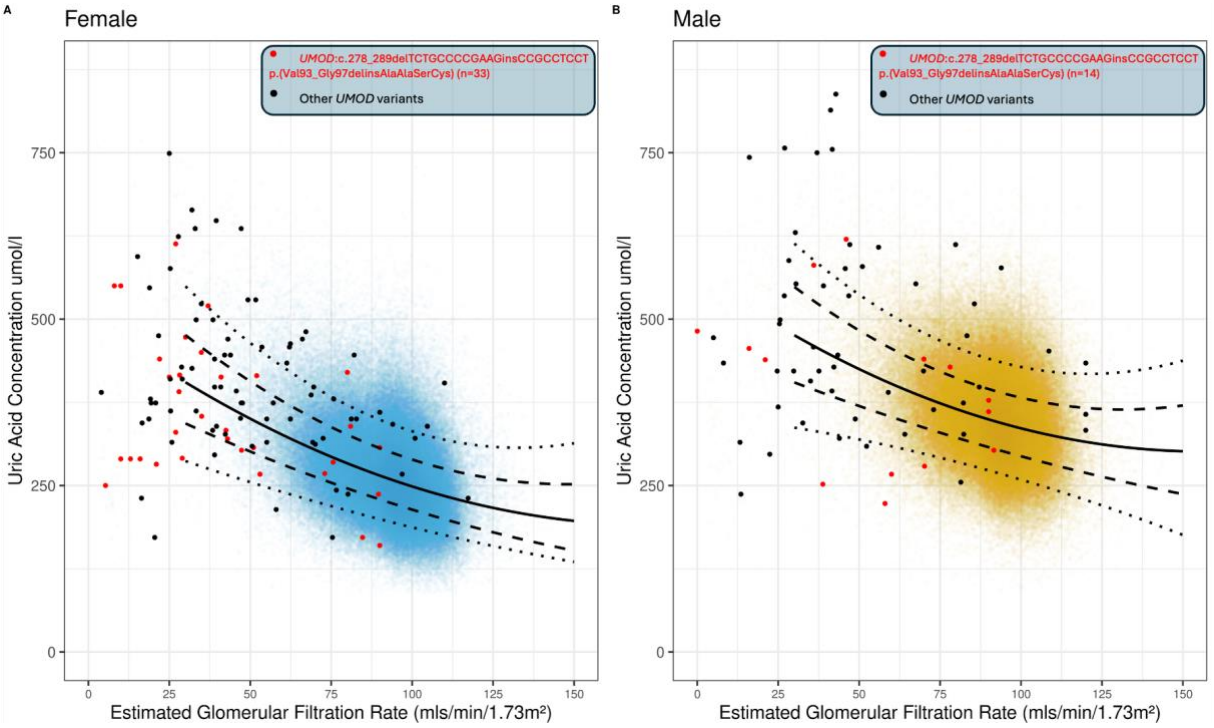

Figure 54: Distribution of Initial Serum Uric Acid Values for ADTKD-*UMOD* p.(Val93_Gly97delinsAlaAlaSerCys Cohort in comparison to Values from UK Biobank Participants*

*[*Sex specific quantile regression plots of the distribution of serum uric acid according to eGFR (female participants on the left and male participants on the right) but patients with the UMOD p.(Val93_Gly97delinsAlaAlaSerCys) variant are labelled as red dots. The blue and yellow dots represent this data in UK Biobank participants in females and males respectively. The black dots represent sex specific participants with ADTKD-UMOD without the UMOD p.(Val93_Gly97delinsAlaAlaSerCys) variant. Overall, those with the UMOD p.(Val93_Gly97delinsAlaAlaSerCys) variant have a distribution of serum uric acid that is more reflective of the general population.]*

Tables 18-20 outline the percentages of participants with ADTKD-*UMOD* that have serum uric acid values that fall above the specified percentiles based on UK Biobank (UKBB) values that are reflective of the general population.

Males with ADTKD-*UMOD*		
UKBB Percentile	**Above (N)**	**Percentage Above (%)**
10	57	**85.07463**
25	51	**76.11940**
50	37	**55.22388**
75	26	**38.80597**
90	19	**28.35821**

Table 18: Serum Uric Acid Concentration Distribution for males with ADTKD-*UMOD*

Females with ADTKD-*UMOD*		
UKBB Percentile	**Above (N)**	**Percentage Above (%)**
10	**98**	**89.90826**
25	87	**79.81651**
50	66	**60.55046**
75	41	**37.61468**
90	23	**21.10092**

Table 19: Serum Uric Acid Concentration Distribution for females with ADTKD-*UMOD*

	***UMOD* p.(Val93_Gly97delinsAlaAlaSerCys) variant**	**Other *UMOD* variants**	**Total**
>50th Percentile	19 (40.4%)	84 (65.1%)	103
<50th Percentile	28 (59.6%)	45 (34.9%)	73
Total	47	129	176
Fisher's Exact Test	P= 0.0053		

Table 20: Serum Uric Acid Concentration Distribution for those with *UMOD* p.(Val93_Gly97delinsAlaAlaSerCys) variant compared to other *UMOD* variants

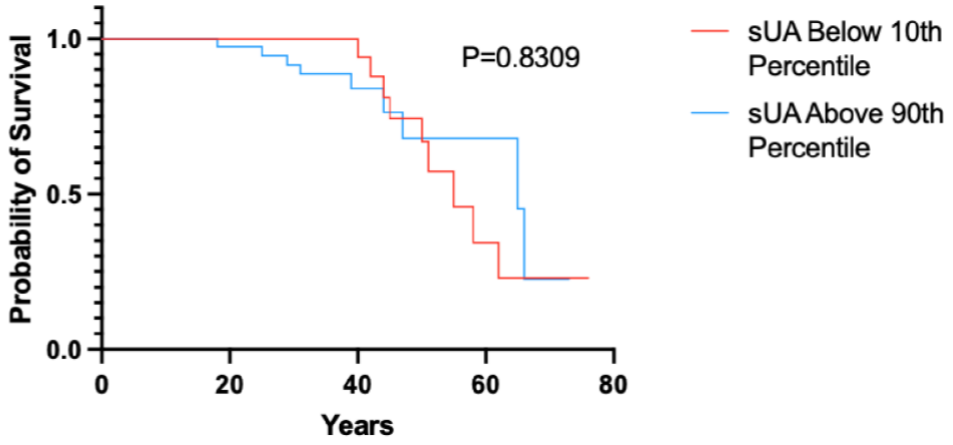

Figure 55: Kidney Survival in UK Biobank Extreme Tails of Serum Uric Acid Concentration in those with ADTKD-*UMOD**

*[*The probability of kidney survival for the two extreme tails of serum uric acid measurements (based on first measured uric acid value) in those with ADTKD-UMOD. There was no statistical difference in kidney survival between the two groups. (sUA below 10th percentile n = 21, sUA above the 90th percentile n = 42)]*

4.3.2 Adjustment of Genomic Association Studies for Hyperuricaemia

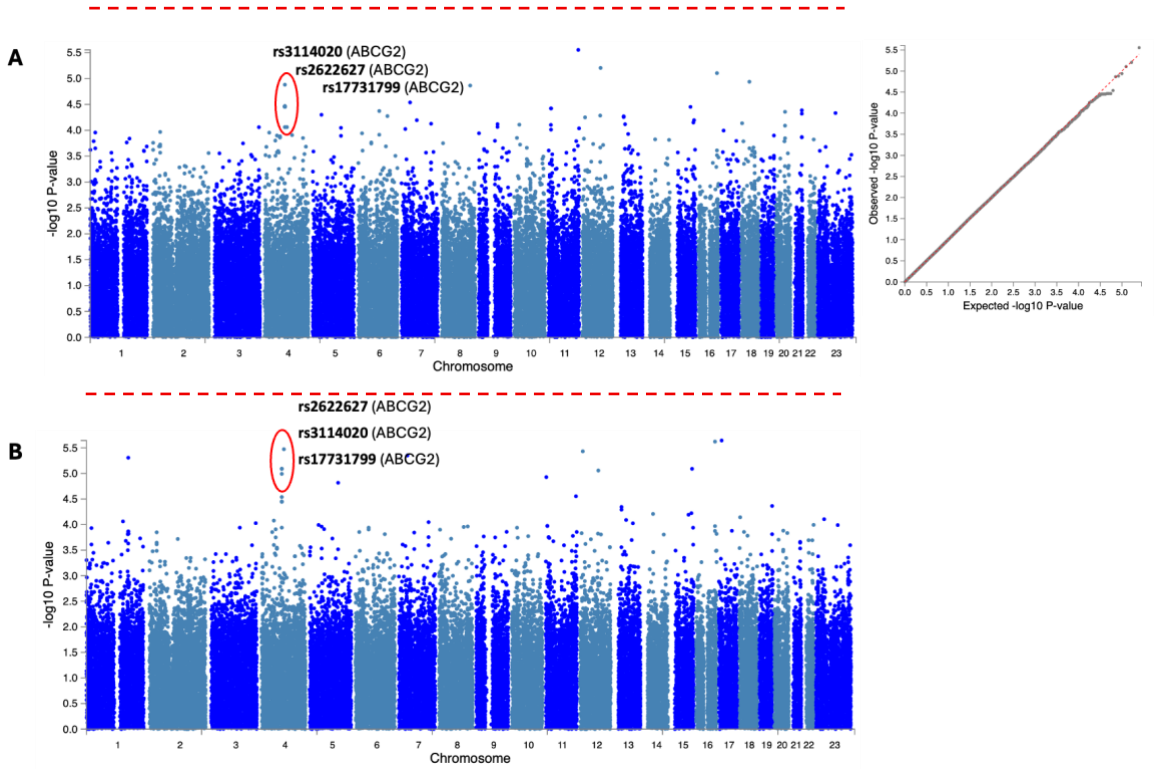

Figure 56: Adjustment of 266-Person ADTKD-*UMOD* Genome Wide Association Study (Imputed) for Hyperuricaemia and Early Onset Gout*

*[*Manhattan plots and a Quantile-Quantile (Q-Q) Plot (top right corner) of the full 266-person Genome Wide Association Study of participants with ADTKD-UMOD assessing the trait of kidney survival before (top) and after (bottom) adjustment for hyperuricaemia and early-onset gout as specified in Section 4.2.4. The Q-Q Plot represents the observed -log$_{10}$(p-values) (y-axis) versus expected -log$_{10}$(p-values) (x-axis) from the GWAS under the null hypothesis. The black diagonal line indicated the expected distribution of no association under the null hypothesis, therefore deviation from this line suggests potential true associations or population stratification. The Manhattan Plot displays log$_{10}$(p-value) of each single nucleotide polymorphism (SNP) tested for association with the trait of kidney survival across the genome. Each point represents a SNP, plotted according to chromosome position on the x-axis and its corresponding p-value on the y-axis. SNPs circled in red and labelled, represent 'candidate SNPs of interest' according to the methodology outlined in Section 3.2.8. These results show that, after adjustment for hyperuricaemia and early-onset gout, there is no loss of the statistical significance of candidate SNPs (rs3114020, rs262627 and rs17731799) in the ABCG2 gene.]*

	Marker ID	MAF	Pval_ Adjchisq (ORIGINAL GWAS)	Pval_ Adjchisq (GWAS ADJUSTED FOR HYPERURICAEMIA/GOUT)
1	**rs3114020**	0.47556391	1.019699e-05	1.315878e-05
2	**rs2622627**	0.43233083	8.109745e-06	3.422581e-05
3	**rs2622626**	0.43233083	8.109745e-06	3.422581e-05
4	**rs17731799**	0.47718631	2.895609e-05	3.423609e-05
5	**rs11732936**	0.48308271	3.571847e-05	3.533143e-05
6	**rs10011796**	0.48308271	3.571847e-05	3.533143e-05

Table 21: Genome-wide Significance Values for ABCG2 SNPs Before and After Adjustment of 266-Person ADTKD-*UMOD* Genome Wide Association Study (Imputed) for Hyperuricaemia and Early Onset Gout*

*[*The top six candidate SNPs of significance within the ABCG2 gene and their genomic-inflation factor adjusted p-values before and after adjustment for hyperuricemia and early-onset gout for the trait of kidney survival in the full 266-person ADTKD-UMOD cohort.]*

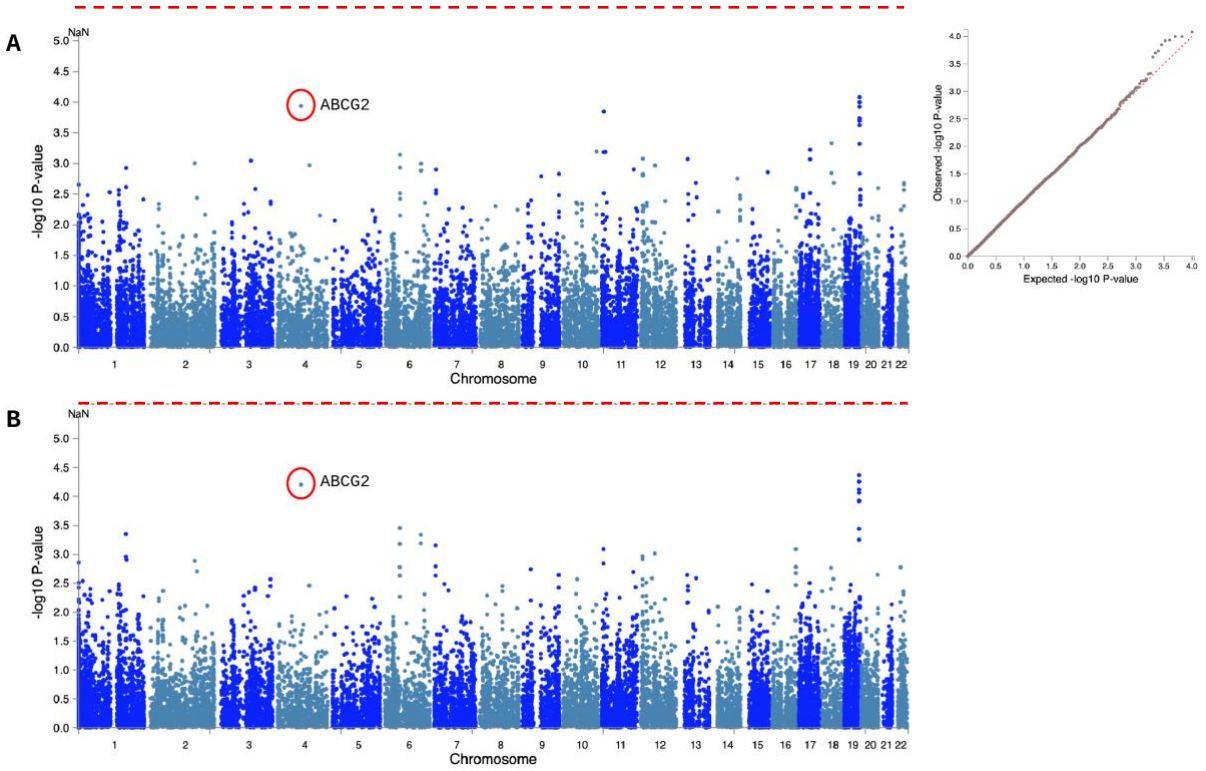

Figure 57: Adjustment of 266-Person ADTKD-*UMOD* MAGMA Gene-Burden Study for Hyperuricaemia and Early Onset Gout*

*[*A Manhattan Plot and Quantile-Quantile (Q-Q) Plot (top right corner) of the full 266-person Gene-burden Study of participants with ADTKD-UMOD assessing the trait of kidney survival before (top) and after (bottom) adjustment for hyperuricaemia and early-onset gout as specified in **section 4.2.4**. The Q-Q Plot represents the observed -log$_{10}$(p-values) (y-axis) versus expected -log$_{10}$(p-values) (x-axis) for genes from the Gene-burden Study under the null hypothesis. The black diagonal line indicated the expected distribution of no association under the null hypothesis, therefore deviation from this line suggests potential true associations or population stratification. The Manhattan Plot displays log$_{10}$(p-value) of each gene tested for association with the trait of kidney survival across the genome. Each point represents a gene, plotted according to chromosome position on the x-axis and its corresponding p-value on the y-axis. ABCG2 is circled in red and there seems to be no apparent loss of statistical significance of this gene for the trait of kidney survival after adjustment for hyperuricemia and early onset gout.]*

	START	STOP	NSNPS	NPARAM	ZSTAT	P-VALUE
ORIGINAL MAGMA STUDY	88961416	89202474	75	24	3.8366	6.2381e-05
2ND MAGMA STUDY (ADJUSTED FOR HYPERURICAEMIA/GOUT)	88961416	89202474	75	24	3.6817	1.1584e-04

Table 22: Genome-wide Significance Values for *ABCG2* Gene Before and After Adjustment of 266-Person ADTKD-*UMOD* Genome Wide Association Study (Imputed) for Hyperuricaemia and Early Onset Gout*

*[*The ABCG2 gene and the genomic-inflation factor adjusted p-value before and after adjustment for hyperuricemia and early-onset gout for the trait of kidney survival in the full 266-person ADTKD-UMOD cohort. (NSNPS) number of SNPs included in the gene-based test, (NPARAM) number of parameters used in the regression model, (ZSTAT) Z-statistic from the gene-based association test, (P-value) statistical significance of the gene's association with the phenotype.]*

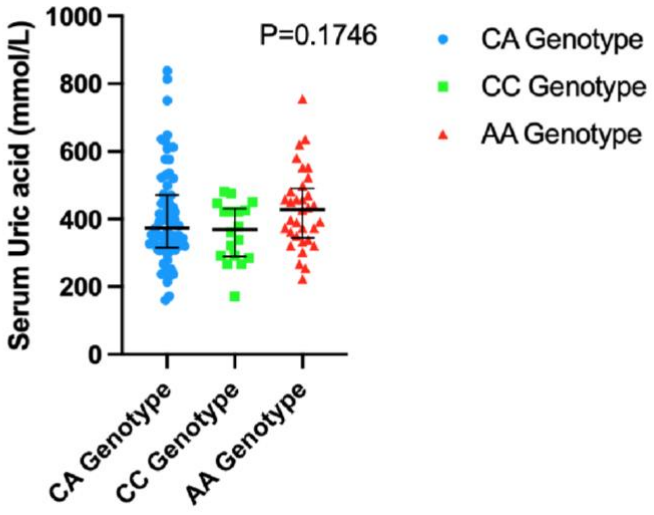

Figure 58: First Serum Uric Acid Values for the *rs2622627* (*ABCG2*) Genotype in full 266-person ADTKD-*UMOD* Cohort*

*[*A scatter plot of serum uric acid values (first serum uric acid measurement) for participant genotype at rs2622627 in those with ADTKD-UMOD. Those participants with an eGFR of <30mls/min/1.73m^2 to reduce confounding of eGFR and those taking uric acid lowering therapy were also removed. A non-parametric Kruskal-Wallis test was performed and showed a non-significant difference in median values between each genotype.]*

	Marker ID	Alleles	MAF in GWAS Cohort	Highest MAF in other populations (1000G)	Ancestral allele	Location within ABCG2	SNP Phenotype Associations	SNP Citations	Functional predictions (CADD Score)
1	**rs3114020**	T/C/G	0.47556391	0.5 (C) Puerto Rican	T	Intron variant	Gout (C allele)	20 citations relating to drug response or gout	C:4.194 G:2.841
2	**rs2622627**	C/A/T	0.43233083	0.5 (A) Columbian	C	Intron variant	-	I citation relating to drug response	A:1.440 T:1.624
3	**rs2622626**	C/A	0.43233083	0.5 (C) Gujarati Indians in Houston, Texas, USA	A	Intron variant	-	2 citations relating to drug response or gout	A:1.654
4	**rs17731799**	G/A/C /T	0.47718631	0.5 (G) South Asian Superpopula tion	G	Intron variant	-	10 citations relating to drug response, gout, colorectal cancer risk and kidney-disease related traits	A:4.817 C:3.379 T:3.285
5	**rs11732936**	A/G/T	0.48308271	0.5 (G) Sri Lankan Tamil	A	Intron variant	4-hydroxychlorothalonil concentration (G allele)	2 citations relating to drug response and metabolites	G:0.577 T:0.550
6	**rs10011796**	T/C/G	0.48308271	0.5 (C) Latino/Admi xed American	C	Intron variant	Gout	14 citations relating to gout, drug response and allopurinol response	C:1.558, G:1.451

Table 23: Annotation of top six *ABCG2* SNPs from the full 266-person ADTKD-*UMOD* GWAS*

*[*The top six candidate SNPs of significance within the ABCG2 gene and their population, phenotypic and functional characteristics.]*

4.3.3 Custom Gene-set Analyses

GENE	CHR	START	STOP	NSNPS	NPARAM	N	ZSTAT	P	ZFITTED_BASE	ZRESID_BASE
SLC22A13	3	38307298	38320161	18	2	266	0.73966	0.24079	0	0.73966
SLC2A9	4	9827848	10041872	1022	21	266	1.2695	0.11466	0	1.2695
ABCG2	4	89011416	89152569	436	28	266	3.7111	9.6487e-05	0	3.7111
SLC17A1	6	25783125	25832287	152	12	266	-0.88431	0.80526	0	-0.88431
SLC17A3	6	25845328	25874471	82	11	266	-1.1914	0.88004	0	-1.1914
SLC5A12	11	26688566	26743574	186	26	266	1.8709	0.028087	0	1.8709
SLC22A6	11	62744069	62752495	18	5	266	1.1629	0.12049	0	1.1629
SLC22A8	11	62760296	62783317	34	13	266	0.43165	0.32501	0	0.43165
SLC22A11	11	64323098	64339002	19	2	266	-0.21693	0.57162	2.7756e-17	-0.21693
SLC22A12	11	64358113	64369825	35	4	266	-0.83113	0.79167	0	-0.83113
SLC5A8	12	101549994	101604016	144	12	266	0.37608	0.34854	0	0.37608
LGALS9	17	25958174	25976586	39	4	266	0.61968	0.26842	0	0.61968
GENE SET	**P = 0.0219368**									

Table 24: MAGMA Custom Gene-Set Analysis for Renal Uric Acid Transporters*

*[*The custom gene-set analysis results for twelve renal uric acid transporter encoding genes. There was a statistically significant association of this gene-set with the trait of kidney survival in the full 266-person imputed ADTKD-UMOD cohort**.** (GENE) name of the tested gene, (CHR) Chromosome on which the gene is located (NSNP**S**) number of SNPs included in the gene-based test, (NPARAM) number of parameters used in the regression model, (ZSTA**T**) Z-statistic from the gene-based association test, (P-value) statistical significance of the gene's association with the phenotype, (ZFITTED_BASE) fitted Z-score based on the baseline model, accounting for gene size and SNP density, (ZRESID_BASE) residual Z-score after removing the effect of baseline predictors.]*

Gene	Top SNP	pval_adjchisq	Location of SNP
ABCG2	4:89030920:C:G	4.526857e-06	Intron variant
SLC5A12 **(SMCT2)**	11:26690854:A:G	6.634499e-05	Genic downstream transcript variant, 3 prime UTR variant
SLC2A9 **(GLUT9)**	4:10019984:T:C	0.003224219	Intron variant, genic upstream transcript variant
SLC17A3 **(NPT4)**	6:25847039:C:T	0.009376723	Intron variant
SLC5A8 **(SMCT1)**	12:101571879:T:C	0.01206438	Intron variant
SLC17A1 **(NPT1)**	6:25809887:G:A	0.03401188	Intron variant
SLC22A6	11:62745799:T:C	0.03451490	Intron variant
SLC22A8	11:62764789:C:T	0.05896264	Intron variant
LGALS9 **(hUAT)**	17:25976524:A:G	0.08608371	3 prime UTR variant, non-coding transcript variant, genic transcript variant
SLC22A13 **(OAT10)**	3:38311953:C:T	0.1370549	Intron variant
SCL22A12 **(URAT1)**	11:64368428:C:T	0.1383749	Intron variant
SLC22A11 **(OAT4)**	11:64330324:A:C	0.1769599	Intron variant

Table 25: Uric Acid Transporter Gene Top SNPs*

*[*The most statistically significant SNPs for each of the 12 renal uric acid transporter encoding genes, their genomic-inflation factor adjusted p-value (pval_adjchisq) from the full 266-person imputed ADTKD-UMOD GWAS and the genomic location of each variant within its respective gene.]*

Tables 26 & 27 show the control custom gene-set analysis results for renal proximal tubule transporter encoding genes and renal magnesium transporter encoding genes. There was no statistically significant association of either gene-set with the trait of kidney survival in the full 266-person imputed ADTKD-*UMOD* cohort.

Renal Proximal Tubule Transporter Gene Set								
GENE	**CHR**	**START**	**STOP**	**NSNPS**	**NPARAM**	**N**	**ZSTAT**	**P**
SLC9A3	5	473298	524549	269	27	266	1.138	0.11999
SLC5A1	22	32439019	32509016	188	10	266	-0.68067	0.75978
SLC5A2	16	31494439	31502091	6	3	266	0.8964	0.18252
SLC6A19	5	1201710	1225232	116	27	266	-2.6132	0.99469
SLC6A20	3	45796941	45838039	85	24	266	-1.4902	0.92408
SLC6A14	X	115567747	115592625	53	9	266	-0.75803	0.76716
SLC34A1	5	176811432	176825849	19	3	266	-0.37734	0.64098
SLC34A3	9	140125209	140131006	18	7	266	-1.092	0.86288
SLC2A2	3	170714137	170744768	77	7	266	-0.21762	0.58834
AQP1	7	30951415	30965131	40	11	266	-0.17088	0.56163
SLC4A4	4	72053003	72437804	659	36	266	-2.1884	0.98627
SLC22A2	6	160637794	160679963	98	5	266	-0.030938	0.51544
GENE SET	**P = 0.97529**							

Table 26: MAGMA Custom Gene-Set Analysis: Control Gene Sets (Renal Proximal Tubule Genes)*

[(GENE) name of the tested gene, (CHR) Chromosome on which the gene is located (NSNPS) number of SNPs included in the gene-based test, (NPARAM) number of parameters used in the regression model, (ZSTAT) Z-statistic from the gene-based association test, (P-value) statistical significance of the gene's association with the phenotype.]*

Renal Magnesium Transporter Gene Set								
GENE	**CHR**	**START**	**STOP**	**NSNPS**	**NPARAM**	**N**	**ZSTAT**	**P**
TRPM6	9	77337411	77503010	296	29	266	0.83235	0.20261
TRPM7	15	50849351	50979012	395	16	266	-1.0212	0.84641
SLC41A1	1	205758221	205782887	61	14	266	1.1104	0.13341
SLC41A3	3	125725200	125820398	318	12	266	-0.58547	0.72089
MRS2	6	24403141	24425816	106	11	266	1.0956	0.13663
GENE SET	**P = 0.15343**							

Table 27: MAGMA Custom Gene-Set Analysis: Control Gene Sets (Renal Magnesium Transporter Genes)*

[(GENE) name of the tested gene, (CHR) Chromosome on which the gene is located (NSNPS) number of SNPs included in the gene-based test, (NPARAM) number of parameters used in the regression model, (ZSTAT) Z-statistic from the gene-based association test, (P-value) statistical significance of the gene's association with the phenotype.]*

4.3.4 Kidney Survival in Kidney Uric Acid Transporter Genotypes

Figures 59-61 show probability of kidney survival at the different genotypes of participants with ADTKD-*UMOD* at the most statistically significant SNP for the three most significant genes (*ABCG2, SMCT2* and *GLUT9*) from custom gene set analysis of renal uric acid transporter encoding genes. All show kidney survival curve separation at different genotypes with statistically significant differences (Cox proportional hazards model) between each SNP genotype within the full 266-person full imputed ADTKD-*UMOD* cohort.

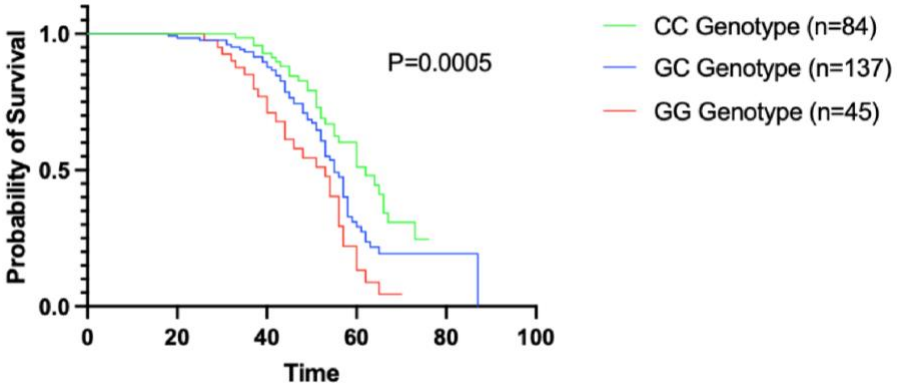

Figure 59: Kidney Survival for the Most Significant SNP (*rs2622621)* in the Top Three Contributing Renal Uric Acid Transporter Genes in 266-Person ADTKD-*UMOD* GWAS Cohort*

*[*Kidney survival curves for each of the three genotypes at SNP **rs2622621 (in ABCG2)** within the full 266-person ADTKD-UMOD GWAS cohort. When a Cox proportional hazards model was applied, there was a statistically significant difference in kidney survival between each of the three genotypes. The GG genotype was associated with the worst kidney survival in the cohort and the CC genotype was associated with the best.]*

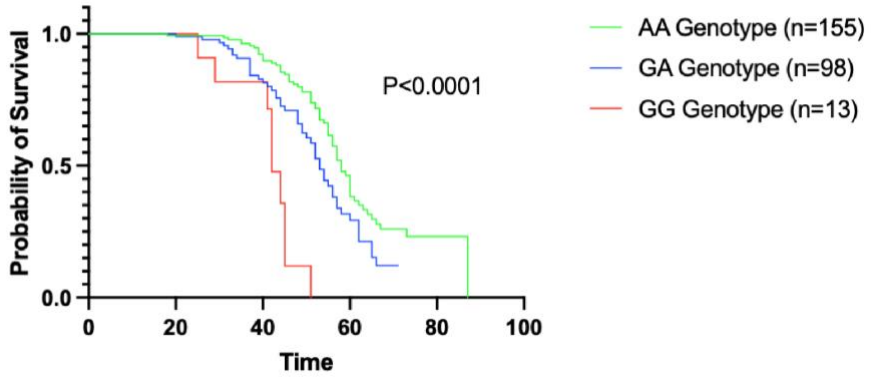

Figure 60: Kidney Survival for the Most Significant SNP (*rs61580147*) in the Top Three Contributing Renal Uric Acid Transporter Genes in 266-Person ADTKD-*UMOD* GWAS Cohort*

*[*Kidney survival curves for each of the three genotypes at SNP **rs61580147 (in SLC5A12 which encodes SMCT2)** within the full 266-person ADTKD-UMOD GWAS cohort. When a Cox proportional hazards model was applied, there was a statistically significant difference in kidney survival between each of the three genotypes. The GG genotype was associated with the worst kidney survival in the cohort and the AA genotype was associated with the best.]*

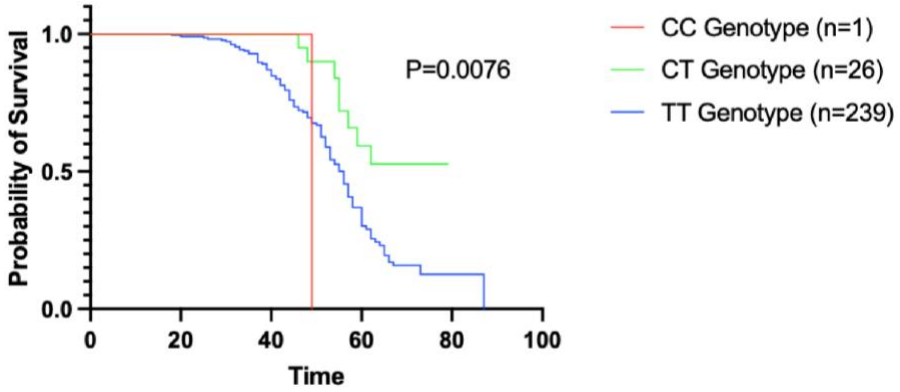

Figure 61: Kidney Survival for the Most Significant SNP (*rs3775943*) in the Top Three Contributing Renal Uric Acid Transporter Genes in 266-Person ADTKD-*UMOD* GWAS Cohort*

*[*Kidney survival curves for each of the three genotypes at SNP **rs3775943 (in SLC2A9 which encodes GLUT9)** within the full 266-person ADTKD-UMOD GWAS cohort. When a Cox proportional hazards model was applied, there was a statistically significant difference in kidney survival between each of the three genotypes. The CC genotype was associated with the worst kidney survival in the cohort and the CT genotype was associated with the best.]*

4.3.5 Effects of ABCG2 on Kidney Survival in the UMOD p.(Val93_Gly97delinsAlaAlaSerCys) Cohort

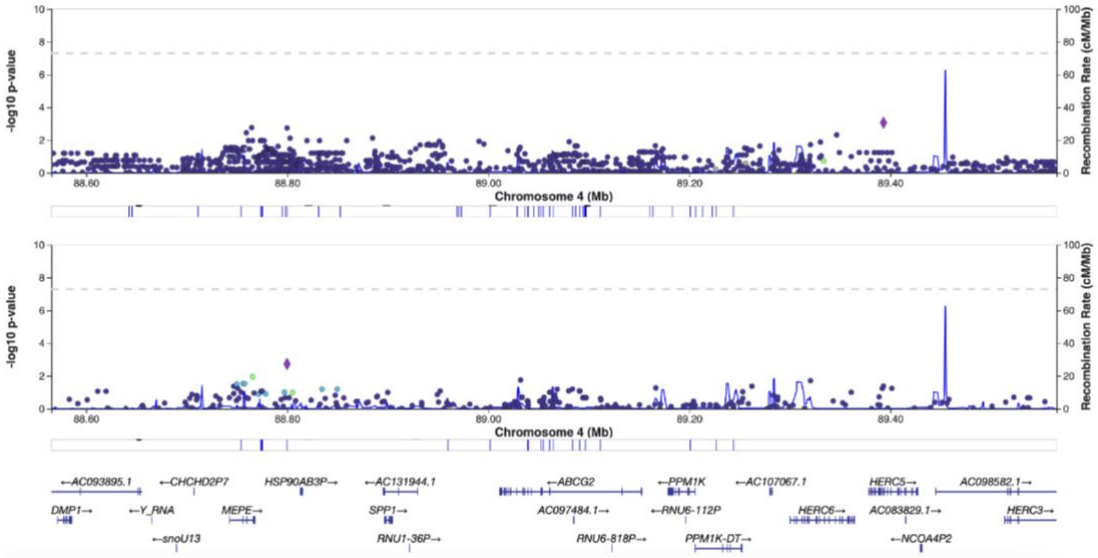

Figure 62: LocusZoom Plot of *ABCG2* in the 92-person *UMOD* p.(Val93_Gly97delinsAlaAlaSerCys) Cohort*

*[*A LocusZoom Plot of the region in and round ABCG2 from the 92-person imputed GWAS for those with the UMOD p.Val93_gly97delinsAlaAlaSerCys variant and the non-imputed version below. No obvious signal is observed in this region for association with the trait of kidney survival in this group.]*

Figures 63 & 64 demonstrate the probability of kidney survival at the different genotypes of the most statistically significant SNP in *ABCG2* from the imputed full 266-person ADTKD-*UMOD* GWAS with/without the UMOD p.Val93_gly97delinsAlaAlaSerCys variant. A statistically significant difference (Cox proportional hazards model) can be observed for the different genotypes in rs2622627 in those without the *UMOD* p.Val93_gly97delinsAlaAlaSerCys variant (**Figure 64**) and this is not observed at the same SNP in those with the *UMOD* p.Val93_gly97delinsAlaAlaSerCys variant (**Figure 63**).

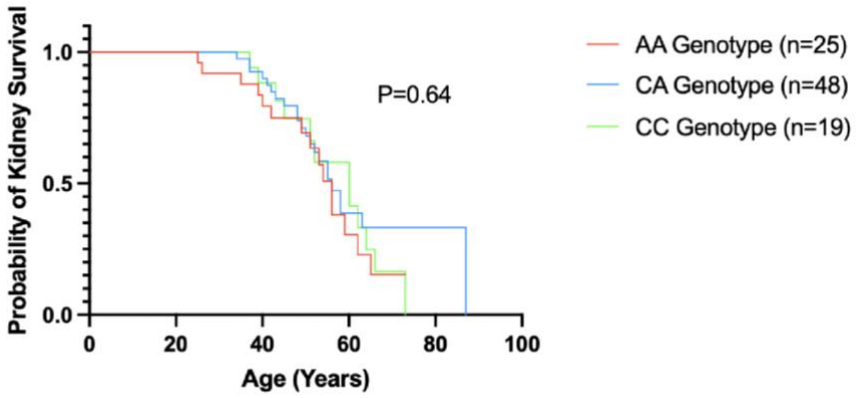

Figure 63: Kidney Survival for the *rs2622627* Genotypes (in *ABCG2*) in *the UMOD* p.Val93_gly97delinsAlaAlaSerCys variant cohort*

*[*Kidney survival curves for each of the three genotypes at SNP **rs2622627 (in ABCG2)** in 92 people with the UMOD* p.Val93_gly97delinsAlaAlaSerCys variant*. When a Cox proportional hazards model was applied, there was no statistically significant difference in kidney survival between each of the three genotypes.]*

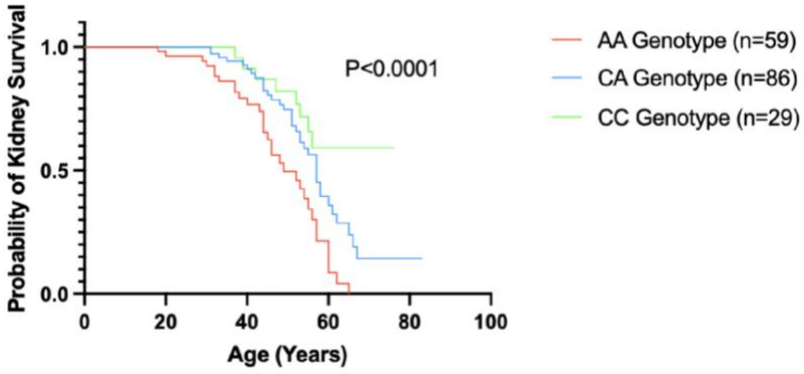

Figure 64: Kidney Survival for the *rs2622627* Genotypes (in *ABCG2*) in those with ADTKD-UMOD without the *UMOD* p.Val93_gly97delinsAlaAlaSerCys variant*

*[*Kidney survival curves for each of the three genotypes at SNP **rs2622627 (in ABCG2)** in 175 people without the UMOD* p.Val93_gly97delinsAlaAlaSerCys variant who have other pathogenic *UMOD* variants*. When a Cox proportional hazards model was applied, there was a statistically significant difference in kidney survival between each of the three genotypes. The AA genotype was associated with the worst kidney survival in the cohort and the CC genotype was associated with the best.]*

4.4 Discussion

The role of uric acid regulation in ADTKD-*UMOD* has always been an intriguing hallmark of the disease but little is understood about this clinical feature and its association with disease pathophysiology. Whilst gain-of-function variants leading to protein misfolding, ER retention and ER-stress is a well justified patho-mechanism in ADTKD-*UMOD* (Bernascone *et al.*, 2010; Johnson *et al.*, 2017; Trudu *et al.*, 2017), only hypotheses have been generated to explain early-onset gout and hyperuricemia prior to CKD onset. Given the absence of uric acid transport in the TAL, it has been postulated to date that hyperuricaemia results from increased proximal tubule reabsorption of uric acid in response to chronic volume contraction from the salt-wasting mechanisms in the disease (Scolari *et al.*, 2004; Devuyst *et al.*, 2019). Whilst this may remain a plausible consideration, further study of hyperuricaemia is required, both to determine the pattern of clinical manifestation, utility as a biomarker and exploration of genetic and functional regulation of uric acid in the disease.

Through the support of collaborators and the UK RaDaR Registry ('National Registry of Rare Kidney Diseases (RaDaR)', 2022), I was able to collect clinical uric acid measurements from the GWAS participants with ADTKD-*UMOD*. This allowed me to explore this as a modifier of disease progression, its relationship with renal uric acid transporter encoding genes of interest from the GWASs I performed and its potential role as a disease biomarker, thanks to the utility of UK Biobank ('UK Biobank', 2022).

4.4.1 Serum uric acid is unlikely to serve as a sensitive or specific stand-alone diagnostic biomarker and is not prognostic for kidney function decline in ADTKD-UMOD

The availability of large population-based databases such as UK Biobank allows for a valuable context in which to evaluate the distribution of serum uric acid concentration in ADTKD-*UMOD* (Allen *et al.*, 2024). **Figures 53 & 54** provide a nuanced interpretation of serum uric acid as a biomarker in ADTKD-*UMOD* as, there is a general trend toward more elevated serum uric acid and an upward distributional shift relative those from the general population, but this is not specific to all patients with ADTKD-*UMOD* as can be seen. Only 21% of females and 28% of males with ADTKD-*UMOD* have serum uric acid levels above the 90th percentile, limiting its broad diagnostic utility **(tables 18-20).** To determine if the degree of

hyperuricaemia was predictive of disease progression, **figure 55** illustrates the comparison in kidney survival between the two extreme tails of serum uric acid concentration in those with ADTKD-*UMOD* which did not show any statistical difference, further supporting that serum uric acid is unable to serve as a biomarker for disease progression and should not be used for kidney function prognostication based on this data.

4.4.2 Hyperuricaemia does not appear to be a prominent feature beyond observed population levels in individuals with the UMOD p. Val93_gly97delinsAlaAlaSerCys variant and this may represent a unique metabolic sub-phenotype

Further to the discussion in **Section 4.4.1**, **tables 18-20** demonstrate a statistically significant difference in the distribution of serum uric acid levels between patients with the *UMOD* p.Val93_gly97delinsAlaAlaSerCys variant and those with other disease causing *UMOD* variants. High serum uric acid levels are much less common in patients with the *UMOD* p.Val93_gly97delinsAlaAlaSerCys variant and, as seen in the final table, only 40.4% of those with this variant have serum uric acid levels above the 50th percentile of the UK Biobank derived measurements. This suggests that serum uric acid levels for those with the *UMOD* p.Val93_gly97delinsAlaAlaSerCys variant are no different than normal population levels. Furthermore, this could suggest that this variant may influence a specific metabolic phenotype where uric acid handling is like the general population as well as a unique disease sub-phenotype. Furthermore, this may be further supported by the lack of modifying effects of *ABCG2* in patients with this variant from the GWAS studies I performed as shown in **Section 4.3.5**. This milder metabolic phenotype could be due a variety of factors, for instance, a different functional impact from this variant such as less severe dysfunction in TAL sodium reabsorption (no individual symptom of salt-wasting was significantly more or less common in those with the *UMOD* p.Val93_gly97delinsAlaAlaSerCys variant based on **Table 6** but numbers were small and thus probably inconclusive), less ER retention of the protein, uromodulin secretion differences (further supported by data in **Section 5.3.1**) or alterations in regulatory genetic effects. This variant was associated with a milder, but not the mildest *in vitro* score, a surrogate for ER retention in previous work (Kidd *et al.*, 2020) and functional studies would be beneficial to investigate this further. This observation may also explain why *ABCG2*, a known urate transporter and genetic determinant of hyperuricaemia, did not emerge as a significant modifier in individuals with the *UMOD*

p.Val93_Gly97delinsAlaAlaSerCys variant **(figures 62-64)**. In the apparent absence of a hyperuricaemic substrate, the modifying effect of *ABCG2* may be functionally or statistically silent in this genetically defined sub-group.

4.4.3 SNPs in ABCG2 are modifying kidney survival in ADTKD-UMOD independent of their effects on serum uric acid

Adjustment of the GWAS and gene-burden studies as described in **Section 4.2.4**, for the presence of hyperuricaemia and early-onset gout allowed determination of whether the effects of *ABCG2* SNPs **(table 23)** on kidney survival were independent of or, alternatively, potentially mediated through, serum uric acid **(figure 56 & 57, tables 21 & 22)**. The alternative method of Mendelian Randomisation (MR) would typically assume that the exposure (serum uric acid) is independent of confounders and has a causal effect on kidney function progression where this is, in fact, not necessarily the case (de Leeuw *et al.*, 2022). First, because serum uric acid, as detailed in **Section 4.4.2,** is not a predominant feature in those with the *UMOD* p.Val93_gly97delinsAlaAlaSerCys variant who were included in the GWAS and is also heavily influenced by confounders such as dietary intake regardless of eGFR (Kuwabara, Kanbay and Hisatome, 2023). Second, it is known that CKD progression is primarily driven by tubular injury and interstitial fibrosis in ADTKD-*UMOD* so hyperuricaemia may be a disease marker rather than a driver and therefore less biologically meaningful in this context (Johnson *et al.*, 2023). Taken altogether, MR might be more susceptible to confounding so genetic association with adjustment was performed instead. When considering adjustment, those taking urate lowering therapies and those with an accompanying eGFR of <30mls/min/1.73m^2 were not considered due to the confounding effects of both (Kuwabara, Kanbay and Hisatome, 2023). As can be seen in **Section 4.3.2,** the modifying effects of *ABCG2* on kidney survival in the disease did not disappear after adjustment for hyperuricemia and early-onset gout suggesting that the modifying effects from SNPs in this gene are independent of their impact on serum uric acid. This is reinforced by **figure 55** where serum uric acid concentration did not have a statistically meaningful impact on kidney survival and the different *ABCG2* SNP genotypes did not have a significantly meaningful impact on serum uric acid concentration either **(figure 58)**.

4.4.4 Renal uric acid transporter encoding genes modify disease progression in ADTKD-UMOD

Gene-set analysis can offer advantages to complement existing GWA Studies. Focussing on groups of genes involved in mechanistic pathways or which have shared functions, gene-set analysis offers greater insights into the biological context of disease. As multiple signals across multiple genes are aggregated, this increases the statistical power and reduces the multiple testing burden which is advantageous in the rare-disease context where numbers of study participants are scarce (Fridley and Biernacka, 2011; Taruscio and Gahl, 2024).

As demonstrated in **Section 4.3.3**, a custom gene set analysis of these twelve renal uric acid transporter-encoding genes was associated with kidney survival in ADTKD-*UMOD* **(tables 24-27)**. This effect was primarily driven by *ABCG2* but *SLC5A12* also contributed in a statistically meaningful way to the kidney survival phenotype. *SLC22A6, SLC22A8, SLC5A8, SLC2A9, LGALS9* and *SLC22A13* had positive Z-statistics but with non-significant p-values suggesting a possible but non-definitive association with kidney survival. **Section 4.3.4** shows that the probability of kidney survival is statistically different based on SNP genotype for the top 3 genes of interest (*ABCG2, SLC5A12* and *SLC2A9*) from custom gene-set analysis **(figures 59-61)** but validation of this work would be beneficial to determine contributions of each of the renal uric acid transporter-encoding genes to disease progression before functional studies can be taken forward.

4.5 Conclusion

Plotting serum uric acid values in patients with ADTKD-*UMOD* against this metric in UK Biobank participants as a surrogate for the general population has enabled disease-specific patterns to be established and assessment of biomarker potential. Validating this with a separate cohort would provide further confirmation that serum uric acid concentration is not suitable as a standalone diagnostic biomarker or a prognostic biomarker for disease progression. Adjusting genetic association studies for hyperuricaemia and early-onset gout allowed exploration of whether the effect of *ABCG2* is independent of serum uric acid. This was important because of the hypothesis that hyperuricaemia itself may be directly or indirectly causing kidney damage through various mechanisms (Johnson *et al.*, 2023). Looking at kidney survival differences in those taking uric acid lowering therapies in this cohort could

also add weight to this understanding. Given the spectrum of *UMOD* variants and their association with CKD and considering ADTKD-*UMOD* is a monogenic model of kidney fibrosis, the terminal point of all forms of CKD, understanding the role of hyperuricaemia in this disease could be very relevant especially as hyperuricaemia develops as all forms of CKD progress (Johnson *et al.*, 2018; Lee *et al.*, 2021). Further work identified as part of these findings include performing a GWAS in the ADTKD-*UMOD* cohort without those with the *UMOD* p.Val93_gly97delinsAlaAlaSerCys variant as these patients may be confounding the exploration of renal uric acid transporter genes and their role in this disease potentially acting as a unique metabolic disease subtype. It would also be of value to perform a GWAS of early-onset gout development in this cohort to further understand the patho-mechanisms of this trait.

Chapter 4 demonstrated that while hyperuricaemia has traditionally been seen as a hallmark feature of ADTKD-*UMOD*, this is only the case in certain genetic subgroups and its relationship with disease progression is complex and not solely determined by serum urate levels. The identification of *ABCG2* as a modifier of kidney survival, independent of its urate-lowering effects, reinforces the notion that uric acid may be more reflective of underlying tubular dysfunction than a direct driver of fibrosis. These findings shift attention toward more proximal markers of tubular health. In Chapter 5, the focus turns to uromodulin itself, the product of the *UMOD* gene, as a mechanistically anchored biomarker of disease. By investigating serum and urinary uromodulin concentrations in individuals with ADTKD-*UMOD*, this chapter evaluates their potential to serve as non-invasive indicators of disease severity and progression and explores how variant-specific trafficking defects influence biomarker profiles.

Chapter 5. Investigation of Serum and Urine Uromodulin in ADTKD-*UMOD*

5.1 Introduction

Biomarkers in ADTKD-*UMOD* are desperately needed to aid early disease detection, prognosis and facilitate clinical trial development and therapeutic decision making as outlined in **Section 1.10**. Recent advances in understanding of uromodulin biology have provided important insights into its potential as a biomarker in the context of ADTKD-*UMOD* and the need to determine if it has a role in predicting or reflecting disease progression.

Uromodulin (Tamm-Horsfall protein), the most abundant protein in human urine is a kidney-specific glycoprotein secreted bidirectionally into both urine and circulation (El-Achkar and Wu, 2012; Devuyst, Olinger and Rampoldi, 2017; Schaeffer, Devuyst and Rampoldi, 2021). Uromodulin's discovery predates modern medicine however its precise role in both health and disease remains enigmatic. Research has progressively revealed that uromodulin, in fact, exists in multiple forms in the body and is a key player in urinary and systemic homeostasis. The majority of urinary uromodulin undergoes polymerisation into organised filaments contributing to urinary tract defence and the prevention of kidney stone formation (Hallson *et al.*, 1997; Chen *et al.*, 2001; Serafini-Cessi, Monti and Cavallone, 2005; Weiss *et al.*, 2020; Micanovic *et al.*, 2022). In contrast, non-polymerising uromodulin is found in both urine and circulation and harbours distinct roles in immune modulation, oxidative stress regulation and vascular homeostasis (LaFavers *et al.*, 2019; Alesutan *et al.*, 2021; LaFavers, Hage, *et al.*, 2022). These various functions depend on the structural form of the protein in addition to its site of action (El-Achkar and Wu, 2012; Devuyst, Olinger and Rampoldi, 2017).

Only recently has it been understood that uromodulin exists in at least two distinct forms: a polymerising form secreted apically into the urine and a non-polymerising form secreted either apically or basolaterally into the circulation (Stsiapanava *et al.*, 2020; Weiss *et al.*, 2020; Micanovic *et al.*, 2022; LaFavers *et al.*, 2023). Th polymerising form undergoes hepsin-mediated cleavage to remove its external hydrophobic patch (EHP), allowing for filament assembly which contributes to urinary homeostasis (Schaeffer *et al.*, 2009; Brunati *et al.*, 2015; Bokhove *et al.*, 2016). In contrast, the non-polymerising form retains the EHP motif and is not filamentous. This form is dominant in the circulation and has been found to have systemic antioxidant effects, regulate immune cell function and inhibit inflammatory cytokine signalling

(LaFavers *et al.*, 2019; Alesutan *et al.*, 2021; LaFavers, Hage, *et al.*, 2022). Basolateral secretion of uromodulin appears to increase during recovery from ischaemia-reperfusion injury and sepsis, suggesting that circulating uromodulin may play a protective role (El-Achkar *et al.*, 2013; LaFavers, Hage, *et al.*, 2022). Furthermore, genetic studies suggest that variations in glycosylation, including *B4GALNT2* variants, influence circulating uromodulin levels (Li *et al.*, 2022).

Studies involving analysis of *UMOD*-knockout mice further emphasise the importance of uromodulin in disease protection. These mice exhibit increased susceptibility to urinary tract infections, tubular injury and dysregulated immune responses including increased neutrophil infiltration, altered macrophage phenotypes and disrupted IL-23/IL-17 axis signalling (Bates *et al.*, 2004; Mo *et al.*, 2004; Raffi *et al.*, 2005, 2009). Moreover, the balance between circulating and urinary forms of uromodulin may be relevant in disease states such as AKI, CKD, ADTKD and their progression (El-Achkar *et al.*, 2011; Micanovic *et al.*, 2015, 2018). Whilst urinary uromodulin is a marker of nephron mass and tubular integrity (Youhanna *et al.*, 2014; Pruijm *et al.*, 2016; Pivin *et al.*, 2018; Scherberich *et al.*, 2018; LaFavers, Micanovic, *et al.*, 2022), serum uromodulin may more directly reflect intracellular processing, trafficking and basolateral secretion (Dawnay and Cattell, 1981; Risch *et al.*, 2014; Fedak *et al.*, 2016; Scherberich *et al.*, 2018; Enko *et al.*, 2021).

Focussing on its expression, uromodulin is restricted to the thick ascending limb of the loop of Henle and, to a lesser degree, the distal convoluted tubule (Sikri *et al.*, 1981; Tokonami *et al.*, 2018). Here, uromodulin is predominantly localised to the apical membrane for urinary secretion, but a small portion is targeted to the basolateral membrane for release into the interstitium and circulation where it is found at a concentration of around one thousand times lower than found in urine (Sikri *et al.*, 1981; Bachmann, Koeppen-Hagemann and Kriz, 1985; El-Achkar *et al.*, 2013). The *UMOD* gene (found on chromosome 16p12.3-16p13.11) is transcriptionally regulated by hepatic nuclear factor 1beta (*HNF1B*) and is responsive to salt intake, sex hormones and glucocorticoids (Pook *et al.*, 1993; Ying and Sanders, 1998; Gresh *et al.*, 2004; Torffvit, Melander and Hultén, 2004; Trudu *et al.*, 2013; Srivastava *et al.*, 2014; Ponte *et al.*, 2021; LaFavers *et al.*, 2023; Nanamatsu *et al.*, 2023). Common SNPs in the promotor region of *UMOD*, such as rs4293393 and rs12917707, are associated with increased expression and secretion of uromodulin (Köttgen, Hwang, *et al.*, 2010; Trudu *et al.*, 2013; Olden *et al.*,

2014; Troyanov *et al.*, 2016). Interestingly, these variants are also paradoxically associated with an increased risk of hypertension and chronic kidney disease, raising important questions about the context-dependent effects of elevated uromodulin levels (Chambers *et al.*, 2010; Gudbjartsson *et al.*, 2010; Köttgen, Pattaro, *et al.*, 2010; Padmanabhan *et al.*, 2010; Pattaro *et al.*, 2012; Morris *et al.*, 2019; Wuttke *et al.*, 2019).

In the setting of ADTKD-*UMOD*, mutant uromodulin misfolds and accumulated intracellularly in the endoplasmic reticulum (ER), impairing uromodulin trafficking and secretion. Consequently, uromodulin deficiency in urine and serum occurs along with downstream effects including ER stress, activation of the unfolded protein response, cellular apoptosis, inflammation, mitochondrial dysfunction and interstitial fibrosis (Vylet'al *et al.*, 2006; Bernascone *et al.*, 2010; Kemter *et al.*, 2013, 2017; Johnson *et al.*, 2017; Ma *et al.*, 2017; Piret *et al.*, 2017; Trudu *et al.*, 2017; Kim *et al.*, 2023; Schiano *et al.*, 2023). It should be specified that the pathological effects are primarily driven by intracellular aggregation of polymerisation-incompetent mutant protein rather than the polymerised urinary form (Schaeffer *et al.*, 2012; Micanovic *et al.*, 2022). This disruption of normal uromodulin trafficking can additionally affect wild-type uromodulin which further compounds the protein deficiency (Schiano *et al.*, 2023). Mutant uromodulin accumulation results in inflammatory milieu which contributes to the slowly progressive kidney dysfunction with minimal glomerular involvement (Gast *et al.*, 2018; Devuyst *et al.*, 2019; Groopman *et al.*, 2019).

Olinger *et al.* demonstrate that urinary uromodulin is a key biomarker distinguishing ADTKD-*UMOD* from ADTKD-*MUC1*. Individuals with ADTKD-*UMOD* showed a markedly reduced urine uromodulin concentration attributed to intracellular retention of mutant uromodulin in TAL cells, resulting in ER stress. In contrast, individuals with ADTKD-*MUC1* displayed a normal urine uromodulin concentration. Those with ADTKD-*UMOD* had a markedly reduced urine uromodulin when compared to healthy controls with matched eGFR and was characteristic of this disease subtype but not sufficiently diagnostic as a standalone biomarker. When combined with a clinical scoring system, urine uromodulin concentration offered a sensitive and specific non-invasive tool to guide genetic testing and distinguish ADTKD subtypes and improved diagnostic performance of ADTKD-*UMOD* over urine uromodulin utility alone (94.1% sensitivity, 74.3% specificity and 84.2% positive predictive value) (Olinger *et al.*, 2020).

Despite uromodulin emerging as a promising biomarker of tubular health, with relevance for both monogenic and complex kidney diseases, its interpretation is complicated by inter-individual variability influenced by both genetic and non-genetic factors. A large recent study in South Africa revealed that urinary uromodulin levels are not uniformly governed by *UMOD-PDILT* genotype across ancestries. Whilst risk alleles robustly predicted urine uromodulin levels in white individuals, they showed minimal effect in Black individuals, in whom factors such as renin activity and albuminuria were more strongly associated. These findings emphasise the need to account for ancestry and broader physiological context when using uromodulin as a biomarker (Strauss-Kruger *et al.*, 2024).

In light of all the complexities specified, both serum and urine uromodulin hold promise as biomarkers that reflect distinct aspects of TAL function, secretory pathway integrity and compensatory systemic responses. The work in this chapter aims to build on this mechanistic understanding and explore how these biomarkers behave in individuals with genetically confirmed ADTKD-*UMOD* and whether they could serve as accessible, clinically relevant indicators of disease progression.

5.2 Methods

5.2.1 Identification of Patients with ADTKD-UMOD for Uromodulin Measurement

An application for ethical approval was submitted to the Health Research Authority (HRA) and approved for collection of patient blood and urine samples for the purpose of measuring serum and urine uromodulin concentration after review by a Research Ethics Committee (REC). Patients were then identified from the local ADTKD-*UMOD* cohort. For inclusion to the study, patients required a genetic diagnosis of ADTKD-*UMOD*, an eGFR (CKD-EPI) of >15 and no prior receipt of a kidney transplant. For control samples, inclusion criteria were no diagnosis of ADTKD-*UMOD*, an eGFR (CKD-EPI) of >15 and no prior receipt of a kidney transplant. Study participants had to provide written informed consent after reading a Participant Information Sheet (PIS) and a matched serum Creatinine sample was collected in the same venesection for normalisation.

5.2.2 Collection of Phenotype Data

Phenotype data included primary *UMOD* variant, genetic diagnosis of patients used as controls and delta of serum Creatinine over time used to define rate of renal function decline. The first recorded serum creatinine measurement was used to subtract from the value taken as part of the study and this was divided by the number of months between these two samples.

5.2.3 Serum ELISA

Human uromodulin concentration in serum was quantified using the BioVendor Human uromodulin ELISA kit (Cat. No. RD191163200R), following the manufacturer's protocol with minor procedural adaptations. Samples and reagents were brought to room temperature prior to use. The lyophilised master standard was reconstituted to a stock concentration of 32ng/ml and serially diluted in dilution buffer to achieve a seven-point standard curve of 0.5-32ng/ml. High- and low-quality control samples were reconstituted according to the certificate of analysis. Patient samples were diluted 1:2, 1:5 or 1:10 based on expected uromodulin concentration and control samples were diluted 1:50 to ensure readings fell within the assay's dynamic range. Standards, controls and diluted samples were pipetted in duplicate into antibody-coated wells and incubated for one hour at room temperature on an orbital shaker. Following sequential incubations with biotin-labelled detection antibody and streptavidin-HRP conjugate, signal was developed using TMB substrate and stopped after 8 minutes. Absorbance was read at 450nm with a 630nm reference filter, and uromodulin concentrations were interpolated from the standard curve, adjusting for sample dilution factors. After initial assay results, it was established that samples for those with ADTKD-*UMOD* should be plated with 1:5 dilution and controls; 1:50 dilution and this was done for all samples in the study. For normalisation across varying degrees of kidney function, uromodulin concentration (ng/ml) was divided by serum creatinine (μmol/L).

5.2.4 Urine ELISA

The second morning urine sample of the day was collected to reduce the confounding of osmolality on urine uromodulin measurement. Samples were immediately frozen at -80 degrees Celsius and then shipped, for external measurement, to the Devuyst lab in Zurich, Switzerland (https://zfin.org/ZDB-LAB-190207-1#summary) on dry ice. ELISA was conducted

via collaborators in the Devuyst lab according to the following published method (Youhanna *et al.*, 2014).

A sandwich ELISA was performed using a sheep anti-human uromodulin polyclonal antibody (Meridian Life Science, K90071C) for plate coating, followed by detection with a mouse monoclonal anti-human uromodulin antibody (Cedarland, CL 1032A) and a horseradish peroxidase-conjugated goat anti-mouse IgG secondary antibody (Bio-Rad 172.1011). Colour development was achieved suing O-phenylenediamine (Sigma-Aldrich) and the optical density was read at 492nm. A human uromodulin standard (Millipore, 100ug/mL) was used to generate a standard curve, and all samples were run in multiple dilutions to ensure reliability. Uromodulin levels were expressed in ug/mL and normalised to urinary creatinine concentrations, which were measured using the Synchron System Creatinine Assay (Beckman Coulter), in accordance with the manufacturer's protocol.

5.2.5 Statistical Analysis

Patients were categorised into separate progressor groups based on their delta serum creatinine over time. Numerical thresholds were applied to define these groups; these cut-offs were selected pragmatically in the absence of established clinical criteria to enable stratification for exploratory analysis. Mann-Whitney or Kruskal-Wallis tests were used to analyse differences between median values.

5.3 Results

These results are structured to first explore the diagnostic utility of uromodulin concentration across disease states and genotypes, followed by its relationship to disease progression. **Section 5.3.1** focuses on serum uromodulin as a potential diagnostic biomarker, comparing concentrations across different kidney diseases and among individuals with various *UMOD* variants. **Section 5.3.2** then assesses the association between serum uromodulin levels and disease progression in ADTKD-*UMOD*, including both the full cohort and those carrying the common *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant, stratified by binary and categorical progressor subtypes. **Section 5.3.3** shifts focus to urine uromodulin, comparing it to serum levels and evaluating differences across *UMOD* variant subgroups. **Section 5.3.4** investigates whether urine uromodulin concentrations similarly reflect disease progression in ADTKD-*UMOD*, again stratified by progressor subtype. As in previous chapters, figures and tables are

presented descriptively and in order of appearance. Interpretation of findings, mechanistic insights, and clinical implications are developed in the discussion **(section 5.4)** that follows.

5.3.1 Serum Uromodulin as a Biomarker of ADTKD-UMOD

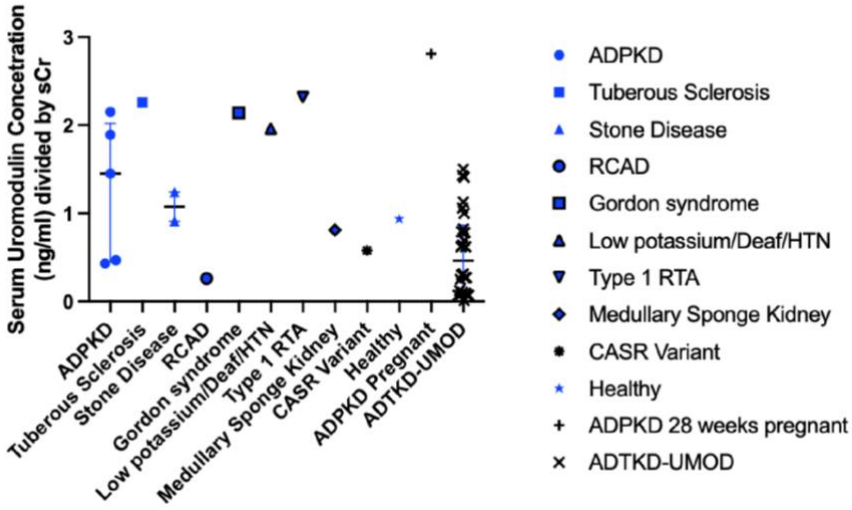

Figure 65: Serum Uromodulin Concentration across Different Kidney Diseases*

*[*A scatter plot which shows the measured uromodulin concentration (ng/ml) in the serum (adjusted for serum Creatinine) from 42 different individuals: 26 with ADTKD-UMOD and 16 with a range of genetic kidney or tubular diseases including one unaffected relative from an ADTKD-UMOD family. Autosomal Dominant Polycystic Kidney Disease (ADPKD); Renal Cysts and Diabetes Syndrome (RCAD); Type 1 Renal Tubular Acidosis (Type 1 RTA); Calcium Sensing Receptor (CASR).]*

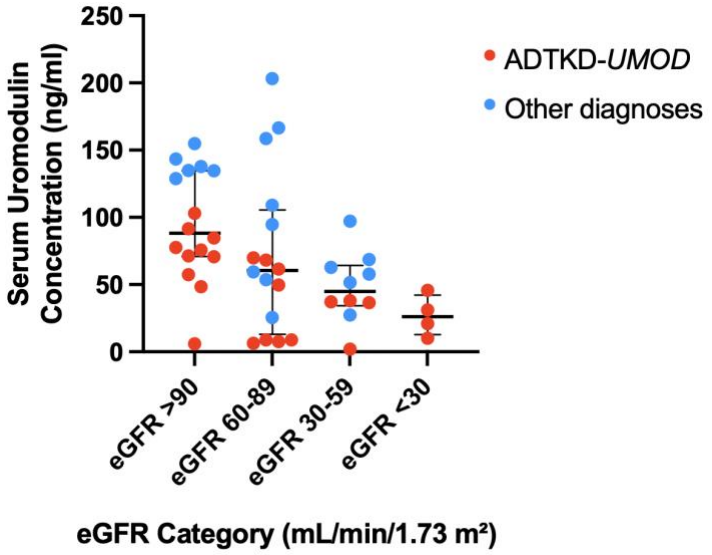

Figure 66: Serum Uromodulin as a Potential Diagnostic Marker in ADTKD-*UMOD**

*[*Serum uromodulin concentration (ng/mL) is shown across categories of estimated glomerular filtration rate (eGFR; mL/min/1.73 m²) for individuals with ADTKD-UMOD (red) and other kidney diagnoses (blue). Each point represents one individual, with horizontal lines indicating the median and interquartile range for each group. Across all eGFR strata, individuals with ADTKD-UMOD exhibit consistently lower serum uromodulin concentrations compared with those with other kidney diseases, supporting its potential as a disease-specific biomarker independent of kidney function.]*

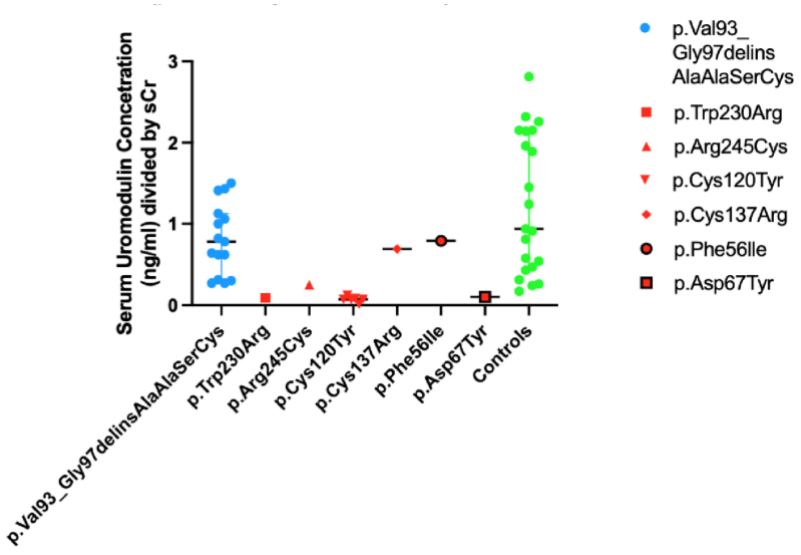

Figure 67: Serum Uromodulin across Different Primary *UMOD* Variants in ADTKD-*UMOD**

*[*A scatter plot which shows the measured uromodulin concentration (ng/ml) in the serum (adjusted for serum Creatinine) for those with ADTKD-UMOD according to primary UMOD*

variant (shown in red). Those with other kidney diagnoses are represented in green. Median values are shown and those with the UMOD p.Val93_gly97delinsAlaAlaSerCys variant are represented in blue.]

5.3.2 Serum Uromodulin as a Biomarker of Disease Progression in ADTKD-UMOD

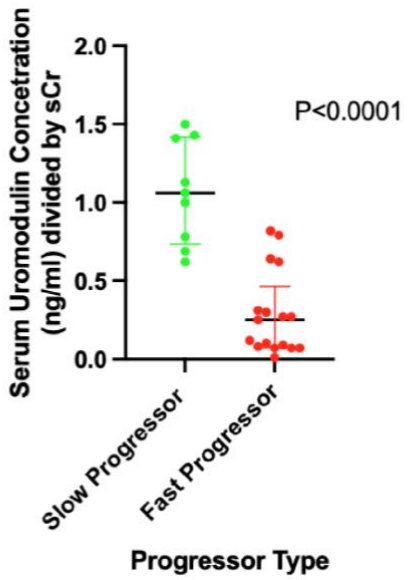

Figure 68: Serum Uromodulin Concentration in Binary Progressor Subtypes in ADTKD-*UMOD**

*[*A scatter plot which shows the measured uromodulin concentration (ng/ml) in the serum (adjusted for serum Creatinine) for those with ADTKD-UMOD (n=26) broken into two binary disease progressor groups (slow and fast) defined by the change between the most recent and the earliest recorded serum creatinine divided by the time between the two values in months (delta of serum Creatinine/time between values(months)). A Mann-Whitney U test demonstrates a statistically significant difference in serum uromodulin concentrations between the two groups, with individuals from the 'fast progressor' group exhibiting significantly lower median levels.]*

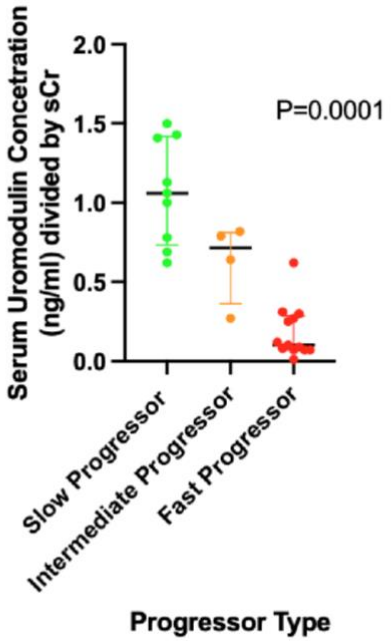

Figure 69: Serum Uromodulin Concentration in Different Progressor Subtypes in ADTKD-*UMOD**

*[*A scatter plot which shows the measured uromodulin concentration (ng/ml) in the serum (adjusted for serum Creatinine) for those with ADTKD-UMOD (n=26) broken into three disease progressor groups (slow, intermediate and fast) defined by the change between the most recent and the earliest recorded serum creatinine divided by the time between the two values in months (delta of serum Creatinine/time between values(months)). A Kruskal-Wallis test demonstrates a statistically significant difference in serum uromodulin concentrations between the three groups, with individuals from the 'fast progressor' group exhibiting significantly lower median levels.]*

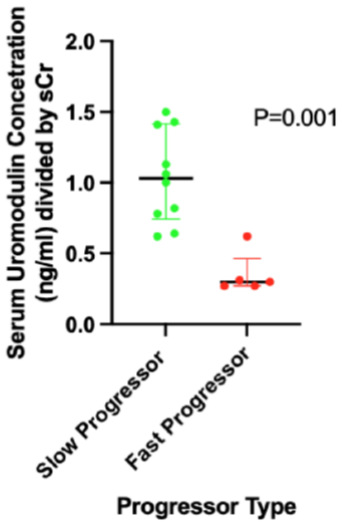

Figure 70: Serum Uromodulin Concentration in Binary Progressor Subtypes in ADTKD-*UMOD* due to the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys Variant*

*[*A scatter plot which shows the measured uromodulin concentration (ng/ml) in the serum (adjusted for serum Creatinine) for those with ADTKD-UMOD due to the UMOD p.Val93_gly97delinsAlaAlaSerCys variant (n=15) broken into two binary disease progressor groups (slow and fast) defined by the change between the most recent and the earliest recorded serum creatinine divided by the time between the two values in months (delta of serum Creatinine/time between values(months)). A Mann-Whitney U test demonstrates a statistically significant difference in serum uromodulin concentrations between the two groups, with individuals from the 'fast progressor' group exhibiting significantly lower median levels.]*

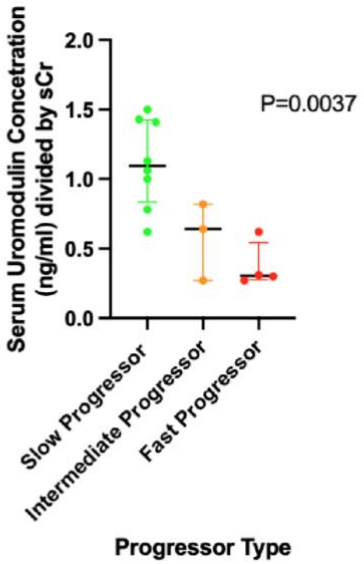

Figure 71: Serum Uromodulin Concentration in Different Progressor Subtypes in ADTKD-*UMOD* due to the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys Variant*

*[*A scatter plot which shows the measured uromodulin concentration (ng/ml) in the serum (adjusted for serum Creatinine) for those with ADTKD-UMOD due to the UMOD p.Val93_gly97delinsAlaAlaSerCys variant (n=15) broken into three disease progressor groups (slow, intermediate and fast) defined by the change between the most recent and the earliest recorded serum creatinine divided by the time between the two values in months (delta of serum Creatinine/time between values(months)). A Kruskal-Wallis test demonstrates a statistically significant difference in serum uromodulin concentrations between the three groups, with individuals from the 'fast progressor' group exhibiting significantly lower median levels.]*

5.3.3 *Urine Uromodulin as a Biomarker in ADTKD-UMOD*

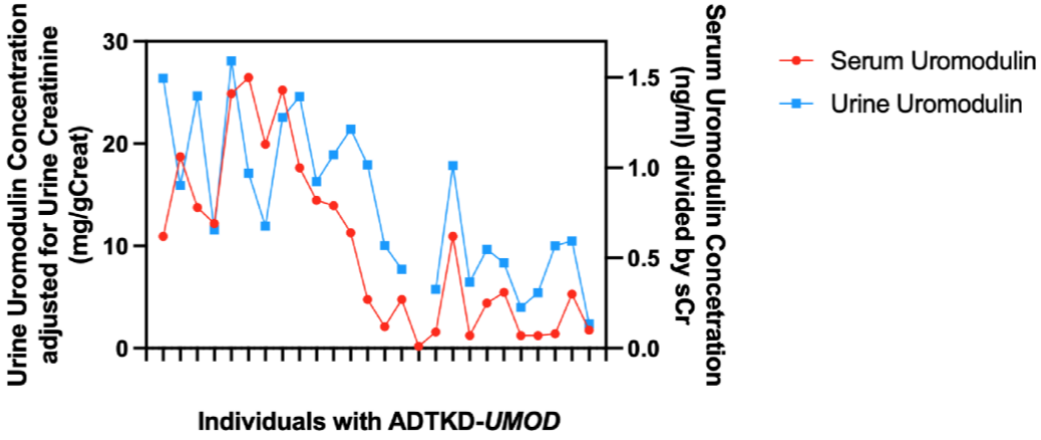

Figure 72: A Comparison of Serum and Urine Concentration in those with ADTKD-*UMOD**

*[*A scatter plot which shows the measured uromodulin concentration (ng/ml) in the serum (adjusted for serum Creatinine) shown in red and in the Urine (mg/gCreat) shown in blue for those with ADTKD-UMOD per individual (n=26).]*

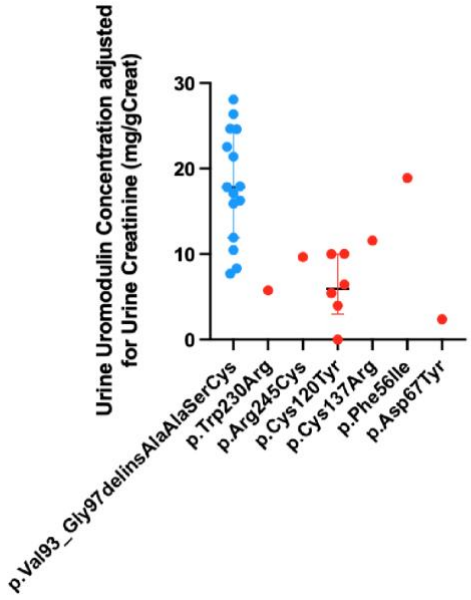

Figure 73: Urine Uromodulin across Different Primary *UMOD* Variants in ADTKD-*UMOD**

*[*A scatter plot which shows the measured uromodulin concentration (mg/gCreat) in the urine (adjusted for urine Creatinine) for those with ADTKD-UMOD according to primary UMOD variant (shown in red). Median values are shown and those with the UMOD p.Val93_gly97delinsAlaAlaSerCys variant are represented in blue.]*

5.3.4 Urine Uromodulin as a Biomarker of Disease Progression in ADTKD-UMOD

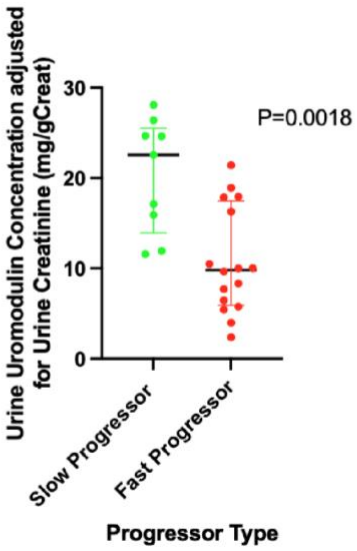

Figure 74: Urine Uromodulin Concentration in Binary Progressor Subtypes in those with ADTKD-*UMOD**

*[*A scatter plot which shows the measured uromodulin concentration (mg/gCreat) in the urine (adjusted for urine Creatinine) for those with ADTKD-UMOD (n=26) broken into two binary disease progressor groups (slow and fast) defined by the change between the most recent and the earliest recorded serum creatinine divided by the time between the two values in months (delta of serum Creatinine/time between values(months)). A Mann-Whitney U test demonstrates a statistically significant difference in urine uromodulin concentrations between the two groups, with individuals from the 'fast progressor' group exhibiting significantly lower median levels.]*

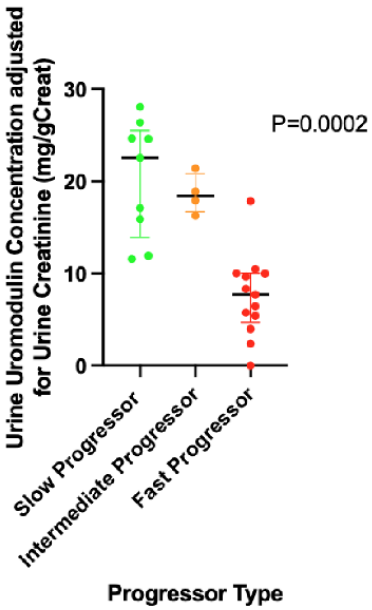

Figure 75: Urine Uromodulin Concentration in Different Progressor Subtypes in those with ADTKD-*UMOD**

*[*A scatter plot which shows the measured uromodulin concentration (mg/gCreat) in the urine (adjusted for urine Creatinine) for those with ADTKD-UMOD (n=26) for those with ADTKD-UMOD (n=26) broken into three disease progressor groups (slow, intermediate and fast) defined by the change between the most recent and the earliest recorded serum creatinine divided by the time between the two values in months (delta of serum Creatinine/time between values(months)). A Kruskal-Wallis test demonstrates a statistically significant difference in urine uromodulin concentrations between the three groups, with individuals from the 'fast progressor' group exhibiting significantly lower median levels.]*

5.4 Discussion

The pathogenicity of many *UMOD* variants are well established, however the extent to which they differentially impact uromodulin secretion and correlate with disease progression remains to be completely understood (Scolari *et al.*, 2004; Vyletal, Bleyer and Kmoch, 2010; Kidd *et al.*, 2020). Smith *et al.* first demonstrated that the recombinant 'indel' (*UMOD* p.Val93_gly97delinsAlaAlaSerCys variant) mutant protein displayed a glycosylation pattern more closely resembling the wild-type and was secreted more efficiently by transfected cells than the C150S mutant (Smith *et al.*, 2011). However, subsequent atomic force microscopy studies by the same group showed that all three *UMOD* variants examined ('indel', C150S, C155R) formed premature intracellular fibrillar structures compared to the wild-type protein,

despite their differences in glycosylation patterns (Stewart *et al.*, 2015). As discussed in the chapter introduction, uromodulin has emerged as a potential biomarker of tubular health and has some diagnostic potential in ADTKD-*UMOD*, yet its utility in stratifying disease severity or progression risk in ADTKD-*UMOD* has not been explored (Troyanov *et al.*, 2016; Olinger *et al.*, 2020; Thielemans *et al.*, 2023). Serum and urine uromodulin concentrations are thought to serve as surrogates for basolateral and apical secretion respectively, reflecting the two primary routes by which uromodulin is released from the thick ascending limb (TAL) in the nephron (Scherberich *et al.*, 2018; Nanamatsu *et al.*, 2024). In a local cohort of genetically confirmed ADTKD-*UMOD* patients, I measured serum and urine uromodulin concentration urine an Enzyme-Linked Immunosorbent Assay (ELISA) with the help of collaborators in the Devuyst lab, Zurich. I integrated this data with *UMOD* genotype information and clinical progression phenotypes to determine if uromodulin secretion patterns could reflect underlying *UMOD* variant biology or disease trajectory. This approach is intended to advance current understanding of the functional consequences of *UMOD* variants in disease and the potential value of uromodulin as a disease biomarker in ADTKD-*UMOD*.

5.4.1 Low serum uromodulin concentration identifies ADTKD-UMOD amongst kidney disease cohorts but not as a standalone biomarker

Referring to **figures 65-66**, serum uromodulin concentration varies within and between those with ADTKD-*UMOD* and in those with different kidney diagnoses. In general, serum uromodulin does appear to be reduced in those individuals with ADTKD-*UMOD* consistent with the known pathogenic effect of *UMOD* variants on uromodulin production and secretion. In contrast, those with ADPKD and Tuberous Sclerosis demonstrate higher or preserved levels suggesting that these conditions do not impair uromodulin secretion in the same way. Notably, an individual who was 28 weeks pregnant with ADPKD had a much higher value suggesting pregnancy may influence uromodulin levels, though only based on a single data point. This is consistent with early work suggesting that uromodulin excretion is altered in pregnancy and seems to be driven by human maternal BMI and ethnicity but is also predictive of pregnancy-associated hypertension in rodent models (Easton *et al.*, 2000; Mary *et al.*, 2024). Other kidney disorders such as Gordon syndrome, Kidney stone disease and Medullary Sponge Kidney show variable uromodulin concentrations, generally falling between healthy and the low levels seen in ADTKD-*UMOD*. Interestingly, the individual with Renal Cysts and Diabetes Syndrome (RCAD) due to an *HNF1B* variant, had a very low uromodulin

concentration. Given that *HNF1B* is a known transcriptional regulator of *UMOD*, this could explain the reduced expression level observed (Gresh *et al.*, 2004). This would be important to validate in a cohort of patients with RCAD. Overall, adjusted serum uromodulin concentration appears to partially distinguish ADTKD-*UMOD* from other kidney diseases as can be seen by a statistically significant lower median value of those with ADTKD-*UMOD* when compared to controls **(figure 65)**, however, this is probably not specific enough as a standalone biomarker given the low values in some other individuals. Thus, serum uromodulin has potential utility as a diagnostic biomarker only if in conjunction with other clinical indictors of disease.

5.4.2 Adjusted serum uromodulin differentiates functional impact of UMOD variants

Firstly, adjusting for serum Creatinine allowed for better comparison of intrinsic difference in uromodulin secretion across genotypes as it accounted for differences in kidney function. The data in **figure 67** suggests that most *UMOD* variants severely impair the secretory process when serum uromodulin concentration is considered a reflection of basolateral secretion from the TAL (El-Achkar *et al.*, 2013). Variants such as p.Trp230Arg, p.Arg245Cys and p.Asp67Tyr are associated with very low adjusted serum uromodulin suggesting a significant disruptive effect on protein trafficking and/or secretion. In contrast, those with the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant have low levels in comparison to controls but slightly higher levels than those with other *UMOD* variants. This could reflect less severe disruption of uromodulin handling or alternative mechanisms which could be affecting its secretion.

5.4.3 Partial preservation of uromodulin secretion in those with UMOD p.Val93_Gly97delins suggests distinct effects on uromodulin trafficking and secretion

The relatively higher serum uromodulin level in those with the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant, compared to those with different *UMOD* variants might indicate a less severe disruption in uromodulin handling **(figure 67)**. This could be because the mutant protein my retain some capacity for correct folding allowing for escape from Endoplasmic Reticulum (ER) retention and thus some progression through the secretory pathway. Therefore, in contrast to certain variants, especially those involving cysteine residues, known to impair folding in this cysteine-rich protein and accumulate in the ER (Serafini-Cessi *et al.*, 1993; Kidd *et al.*, 2020), the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant may allow for more efficient protein trafficking and folding. Supporting this, Kidd *et al*

did not find that loss or gain of a cysteine residue (affecting protein polarity) was associated with kidney survival, it was, rather, the individual's *in vitro* trafficking score. It could be, however, that my findings suggest different mechanisms that could affect basolateral secretion such as altered post-translational processing or compensatory cellular responses. In the previous chapter, it was found that those with the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant have serum uric acid levels akin to general population levels in contrast to other pathogenic *UMOD* variants. Thus, the preservation of some serum uromodulin secretion in tandem with the absence of pathological serum urate could reflect a less severe defect in tubular transport mechanisms if we hypothesise that serum uric acid may be a biomarker of tubular health (Giordano *et al.*, 2015; Alex *et al.*, 2024). These findings reinforce the concept that *UMOD* variants can exert graded effects on protein trafficking, tubular function and systemic consequences such as hyperuricaemia. Furthermore, serum uromodulin could serve as a functional redout of variant-specific pathogenicity in ADTKD-*UMOD*.

5.4.4 Serum uromodulin stratifies disease progression risk in ADTKD-UMOD and is independent of primary UMOD variant

Serum uromodulin, when adjusted for serum creatinine, appears to correlate with clinical rate of kidney disease progression in individuals with ADTKD-*UMOD* **(figures 68-71)**. Individuals classed as slow progressors exhibited higher uromodulin values compared to intermediate and fast progressors. This suggests that reduced uromodulin secretion, likely due to more severe intracellular retention and trafficking defects may be linked to a more aggressive disease course. Interestingly, this relationship persists even when comparing individuals with the same primary *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant. This suggests that serum uromodulin could serve as a potential biomarker of disease progression risk even beyond the effect of the primary *UMOD* variant and complements the *in vitro* trafficking score suggested to predict kidney survival in ADTKD-*UMOD* yet is considerably easier to measure, more accessible and scalable as a method for clinical and research utility (Kidd *et al.*, 2020). This could suggest that there are secretion-related as well as variant-specific mechanisms in determining the clinical trajectory of ADTKD-*UMOD* and further supports explanations for the wide variability of disease progression within families with the same *UMOD* variant.

5.4.5 Serum and urinary uromodulin show concordant reductions across individuals with ADTKD-UMOD

It is clear that in most individuals displayed in **figure 72**, serum and urine uromodulin levels track together, suggesting a broadly consistent reduction in both apical and basolateral secretion across individuals. This could imply that the degree of uromodulin trafficking impairment affects both secretory pathways simultaneously. This coordinated reduction may reflect the extent of intracellular retention of misfolded uromodulin and the overall degree of secretory dysfunction and both, together, could be a robust readout of variant-specific pathogenicity and disease severity beyond the primary disease-causing variant.

5.4.6 Adjusted urine uromodulin concentration is influenced by primary UMOD variant and correlates with disease progression

Like serum uromodulin, urine uromodulin concentration varies across individuals with ADTKD-*UMOD* and appears to reflect both underlying *UMOD* variant and clinical progression status over and above the primary gene variant as seen in **figures 73-75**. As with serum concentration, adjusted urine uromodulin is associated with disease progression for all those measured and for those with the same *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant. Those with a higher adjusted urine uromodulin concentration were classified as slow progressors and those with the lowest values were classified as fast progressors based on change in eGFR over time. Taken together, adjusted urine uromodulin concentration is influenced by specific *UMOD* variant as well as clinical disease progression risk suggesting its potential utility as a biomarker for both variant impact and disease progression stratification.

Despite its promise, several challenges remain in the clinical implementation of uromodulin as a biomarker. Current dipstick-based platforms used in routine urinalysis are not designed to detect or quantify uromodulin due to its polymeric structure, extensive glycosylation, and the need for specific detection methods such as ELISA (Kavuru *et al.*, 2020; Thielemans *et al.*, 2023). The development of a reliable, point-of-care dipstick capable of measuring uromodulin remains a technical hurdle but could transform early detection of tubulointerstitial pathology if overcome.

In parallel, serum uromodulin is emerging as a robust and potentially more clinically accessible biomarker. Unlike urinary levels, which are influenced by hydration status and voiding

behaviour, serum uromodulin reflects systemic secretion from viable thick ascending limb (TAL) cells into the interstitium and bloodstream, offering a stable surrogate for nephron mass (Scherberich *et al.*, 2018; Thielemans *et al.*, 2023; Nanamatsu *et al.*, 2024). In broader CKD populations, reduced serum uromodulin has been associated with lower eGFR, cardiovascular risk, and adverse renal outcomes, suggesting that it should enter clinical practice as a biomarker of tubular health and disease severity.

More recently, attention has turned to the *qualitative* properties of uromodulin, particularly its oxidative status and other post-translational modifications such as glycosylation profile, which can now be assessed using mass spectrometry coupled to high-performance liquid chromatography (Patabandige, Go and Desaire, 2021; Chaiyarit and Thongboonkerd, 2022). The oxidised forms of uromodulin may offer a window into tubular cell stress and dysfunction, particularly relevant in ADTKD-*UMOD* where intracellular trafficking and ER stress are central to disease pathogenesis (Nanamatsu *et al.*, 2024). In this context, mass spectrometry techniques not only allow quantification but also enable discrimination between functional and dysfunctional protein isoforms, laying the foundation for a future in which uromodulin is employed as both a quantitative and functional biomarker of tubular health (Micanovic *et al.*, 2022).

Collectively, both urinary and serum uromodulin hold strong promise as mechanistically anchored biomarkers. Their integration into future diagnostic and prognostic frameworks for monogenic and complex kidney diseases will depend on assay standardisation, regulatory validation, and incorporation into clinical trials assessing therapy response.

5.5 Conclusion

Overall, this work demonstrates that both serum and urine uromodulin concentrations, adjusted for serum and urine creatinine respectively, differ across individuals with ADTKD-*UMOD* in a manner which reflects both the underlying primary *UMOD* variant and the rate of disease progression independent of the primary variant. Individuals with fast progressing disease and those with *UMOD* missense variants exhibited markedly reduced serum and urine uromodulin levels, consistent with more severe defects in protein folding and/or secretion. In contrast, those with the in-frame insertion-deletion *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant, demonstrated less depleted levels in serum and urine than the other *UMOD* variants. The parallel trend in reduction of uromodulin in serum and urine in each individual with

ADTKD-*UMOD* suggest a coordinated impairment in basolateral and apical secretion. These findings provide fresh insights into the functional heterogeneity of *UMOD* variants and highlight the potential of both serum and urine uromodulin as easily measured biomarkers of disease progression risk in ADTKD-*UMOD* independent of primary *UMOD* variant and glomerular filtration rate. This work contributes to the growing understanding of genotype-phenotype correlations in ADTKD and supports the integration of biomarker data into future diagnostic and prognostic frameworks which are also desperately needed for clinical trial development.

The findings in Chapter 5 provide compelling evidence that serum and urinary uromodulin concentrations reflect both the underlying molecular pathology and clinical trajectory of ADTKD-*UMOD*. Low levels of uromodulin, influenced in part by variant-specific trafficking defects, were associated with more advanced disease, supporting its role as a mechanistically informative and prognostically relevant biomarker. These insights underscore the value of integrating molecular, genetic, and phenotypic data to enhance clinical stratification in rare kidney diseases. In Chapter 6, these strands are brought together in a broader discussion of the clinical, genetic, and translational implications of this work. This final chapter reflects on how ADTKD-*UMOD* serves as a model for complex monogenic disease, explores opportunities for therapeutic development, and considers how lessons learned here may inform future approaches to diagnosis, classification, and personalised care in chronic kidney disease more widely.

Chapter 6. General Discussion

Autosomal Dominant Tubulointerstitial Kidney Disease due to *UMOD* variants (ADTKD-*UMOD*) is an inherited renal disorder defined by progressive chronic kidney disease, bland urinalysis and tubulointerstitial fibrosis (Devuyst *et al.*, 2019; Mabillard, Sayer and Olinger, 2023). While the causative role of *UMOD* variants is well established, the considerable inter- and intra-familial variability in disease progression remains poorly understood (Kidd *et al.*, 2020). At the outset of this research, I hypothesised that both genetic and clinical modifiers contribute to this phenotypic heterogeneity, and that systematic characterisation and stratification of these factors could generate novel insights into disease pathophysiology, biomarker development, precision prognostication and potential therapeutic targets.

The overarching aim of this work was to define the clinical and genetic landscape of ADTKD-*UMOD* in the UK, evaluate genome-wide modifiers of disease progression and assess the utility of potential biomarkers, specifically uric acid and uromodulin in predicting disease progression. This general discussion integrates the research findings across these strands of work and situates them within the broader literature, highlighting novel contributions, implications for clinical management of ADTKD-*UMOD* and directions for future research.

6.1 A Distinct Genetic Landscape: Founder Effects and Variant Spectrum in the UK

This thesis presents the first genetic characterisation of ADTKD-*UMOD* across the UK and reveals a variant spectrum broadly consistent with that observed internationally (Bollée *et al.*, 2011; Ayasreh *et al.*, 2018; Olinger *et al.*, 2020; Gong *et al.*, 2021; Tanaka *et al.*, 2025). However, the more detailed characterisation of a recurrent in-frame insertion-deletion variant, *UMOD* p.Val93_Gly97delinsAlaAlaSerCys, enriched within the UK population, suggests a potential founder effect. This variant alone accounted for a substantial proportion of genetically confirmed cases in the cohort, a pattern not observed in comparable European or North American series (Devuyst *et al.*, 2019; Olinger *et al.*, 2020). The geographic and familial clustering of this variant within specific UK regions, particularly Northeast England, is highly suggestive of a shared ancestral haplotype and illustrates the utility of national datasets in uncovering population-specific founder events (Valluru *et al.*, 2023). This phenomenon is well established in population genetics, wherein a rare variant introduced by a common

ancestor becomes enriched through genetic drift in an isolated or demographically constrained population (Peltonen, Jalanko and Varilo, 1999; Jakkula *et al.*, 2008).

Smith *et al.* initially reported that closely linked markers were not haploidentical in four families with this variant, supporting a variant hotspot model rather than a single ancestral origin (Smith *et al.*, 2011). Valluru *et al.* (2022) subsequently provided evidence for a UK-specific founder variant by demonstrating extended haplotype sharing across 22 families; however, the absence of formal haplotype phasing, variant age estimation, and the small number of sequenced individuals limited the strength of their conclusions (Valluru *et al.*, 2023). The current analysis substantially improves upon these prior studies through the application of SHAPEIT2 for genome-wide phasing and ESTIAGE for likelihood-based estimation of the age of the most recent common ancestor (MRCA) (Genin *et al.*, 2004; Delaneau *et al.*, 2013; Delaneau, Zagury and Marchini, 2013). By analysing phased genotypes across chromosome 16 and incorporating recombination rates from the 1000 Genomes Project alongside allele frequencies from gnomAD, this approach enables the formal characterisation of both shared and maximal haplotypes and provides a time-based estimate of the founder event (1000 Genomes Project Consortium *et al.*, 2015; Gudmundsson *et al.*, 2022). The shared haplotype surrounding the variant spans approximately 18–20 kb, with the longest observed haplotype extending ~477 kb in a subset of individuals. The most recent common ancestor was estimated to have lived just 3 generations ago (95% CI: 2–4) demonstrating the recency of this founder event. Importantly, this analysis was performed across two ancestral populations, UK and US individuals, further reinforcing the robustness of the founder effect. Despite differences in geographic origin, a shared haplotype structure was observed in both populations, supporting a single ancestral origin rather than recurrent variant hotspots. This cross-population consistency highlights the power of integrating international datasets in identifying and characterising clinically relevant founder variants. This work represents the most rigorous evidence to date for a UK founder variant in ADTKD-*UMOD*, reinforcing the case for ancestry-informed genetic testing and variant prioritisation. The use of formal haplotype reconstruction and time-based modelling is a significant methodological advance over prior studies, providing not only structural insight but also critical information on the age and dissemination of the variant. Importantly, these findings were achieved using a relatively large, nationally ascertained cohort with representative coverage of affected UK regions, enhancing the generalisability of the conclusions.

The concept of founder effects has gained substantial traction in human medical genetics. Founder variants have been implicated in a range of Mendelian disorders within genetically isolated populations or those experiencing demographic bottlenecks. Examples include glycogen storage disease type IIIa and congenital sucrase-isomaltase deficiency in the Inuit (Evans, 2015), infantile hypophosphatasia in Mennonites, and pathogenic *BRCA1* and *BRCA2* variants in Ashkenazi Jews, Icelanders, and populations in Orkney (Thorlacius *et al.*, 1996; Kerr *et al.*, 2024). In these settings, repeated colonisation events and geographic isolation have facilitated the persistence and enrichment of specific alleles, with measurable impacts on disease prevalence and penetrance. Importantly, the identification of founder variants is directly translatable into clinical practice, offering a route to streamlined genetic testing, improved diagnostic yield, and targeted family cascade screening (Ankala *et al.*, 2015).

Recognising a recurrent variant with founder characteristics in ADTKD-*UMOD*, now confirmed across both UK and US populations, opens the possibility of applying similar strategies. In regions such as Northeast England where this variant is clustered, targeted testing could efficiently identify presymptomatic carriers, reduce diagnostic delay, and facilitate early monitoring and clinical management. Moreover, the establishment of a genetically defined, trial-ready cohort is critical in a rare disease context to accelerate recruitment into therapeutic trials. This approach mirrors successful implementations in other diseases, such as ancestry-targeted *BRCA* testing in Ashkenazi Jewish and Northern Isles populations, where carrier frequency is markedly elevated and targeted testing is both cost-effective and clinically impactful (Kerr *et al.*, 2024).

Interestingly, contrary to prior reports suggesting a milder phenotype (Smith *et al.*, 2011), no significant differences in kidney survival were observed between individuals with the *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant and those with other *UMOD* variants including Gly-substitutions and in-frame deletion variants. This observation reiterates previous findings that primary variant class, such as resultant loss or gain of a cysteine residue, does not predict prognosis (Kidd *et al.*, 2020), reinforcing that genotype alone is insufficient to explain disease course. Similarly, Bollée *et al.*, 2011; Moskowitz *et al.*, 2013 identified substantial intrafamilial variability and only modest effects on kidney survival, further supporting the hypothesis that clinical heterogeneity in ADTKD-*UMOD* is shaped by additional modifiers beyond the primary variant itself. The marked clinical heterogeneity observed within and between families carrying this founder variant emphasises the need to investigate additional modifying factors.

This also echoes findings in other founder-associated conditions, where variable expressivity despite identical genotypes implicates the role of genetic modifiers, polygenic risk, epigenetic regulation, and environmental influences (Evans, 2015; Kingdom and Wright, 2022).

Finally, while the founder effect concept was originally developed in the context of island biogeography and speciation (Kingdom and Wright, 2022), its relevance in human health is now well recognised. As shown in work by Kerr *et al.*, 2024, recognising founder variants in underrepresented or underserved populations not only improves access to testing and clinical services but also supports the equitable delivery of precision medicine. In the case of ADTKD-*UMOD*, the identification of a recent UK-specific founder variant represents an important advance in our understanding of the disease's genetic architecture and offers a model for integrating population genetics with translational medicine. This is particularly important in Northeast England, a region disproportionately affected by health inequities and comparatively underfunded in terms of healthcare spending (Corris *et al.*, 2020), where ancestry-informed testing strategies could provide a vital route to improving outcomes in rare genetic kidney disease.

6.2 Clinical Variability, Diagnostic Delay, and Missed Opportunities

This thesis presents the first national characterisation of ADTKD-*UMOD* in the UK, drawing on a genetically confirmed cohort recruited via the RaDaR Registry ('National Registry of Rare Kidney Diseases (RaDaR)', 2022). By integrating clinical, genetic, and biomarker-based data, it extends and refines the international landscape of ADTKD research. Markedly, this UK cohort is uniquely enriched for the recurrent *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant, previously identified as a potential founder variant (Valluru *et al.*, 2023). Consistent with findings from Smith *et al.*, 2011, individuals with this variant had an apparently milder disease phenotype with regard to serum uric acid and gout incidence but, in contrast, did not show a significantly different kidney survival compared to other genotypic groups, further highlighting the limited prognostic utility of variant subtype alone.

The UK cohort revealed several distinctive clinical features that extend beyond what is commonly reported in the ADTKD-*UMOD* literature (Bollée *et al.*, 2011; Ayasreh *et al.*, 2018; Olinger *et al.*, 2020; Gong *et al.*, 2021; Tanaka *et al.*, 2025). One of the most notable findings was the significantly later onset of gout in comparison to international cohorts. In many

individuals, gout occurred well after the onset of renal impairment diminishing its utility as an early diagnostic sign. This finding contrasts with studies that have emphasised early-onset gout as a hallmark of ADTKD-*UMOD* (Bollée *et al.*, 2011; Olinger *et al.*, 2020) and suggests that reliance on hyperuricaemia or gout as diagnostic cues may be problematic in specific populations or variant backgrounds.

A further clinical observation was the high burden of pregnancy complications among females with ADTKD-*UMOD*, an underreported feature in existing literature especially given this cohort all experienced pregnancies when CKD was early in its disease course, at either CKD stage G1-2. These complications included recurrent miscarriages, medical complications such as gestational hypertension and adverse perinatal outcomes. While the international cohort described by Bleyer et al. (2022) suggested relatively benign pregnancy outcomes for individuals with ADTKD, reporting low rates of hypertension (12%), preterm delivery (11%), and NICU admission (12%), our findings, although numbers were small, present a more concerning narrative. In our UK cohort, adverse pregnancy outcomes were considerably more common despite very early CKD stages. This is consistent with broader epidemiological evidence from the BJOG systematic review (Jeyaraman *et al.*, 2024), which demonstrated that women with CKD stages 3–5 face significantly elevated risks of pre-eclampsia (OR 55.18), preterm birth (OR 20.24) and NICU admission (OR 19.32). Importantly, these risks far exceed those reported in the general pregnant population, where complications affect approximately 15% of pregnancies (Changalidis *et al.*, 2022).

This discrepancy between Bleyer's findings and our own may be due to differences in cohort composition, CKD severity, or methodology (Bleyer *et al.*, 2023). The UK data benefit from contemporaneous clinical record linkage and clinician-led retrospective recall versus Bleyer's voluntary self-report questionnaires which could introduce varying degrees of recall and selection bias. Our findings align more closely with the wider CKD-in-pregnancy literature and challenges the notion that ADTKD-*UMOD* confers a low obstetric risk. Instead, they reinforce the importance of recognising ADTKD-*UMOD* as a condition that, like other kidney diseases, demands enhanced surveillance, pre-pregnancy counselling, and multidisciplinary antenatal care.

Mechanistically, emerging evidence may explain why pregnancy complications are more frequent in this condition. Nanamatsu *et al.*, 2023 identified *UMOD* as an oestrogen-responsive gene with circulating uromodulin levels significantly higher in women than men. Their study demonstrated that 17β-estradiol upregulates *UMOD* transcription in thick ascending limb cells probably via non-canonical oestrogen response elements. During pregnancy, rising oestrogen levels could further amplify *UMOD* transcription in women carrying pathogenic variants increasing mutant protein production and enhancing tubular stress. Theoretically, this may compound underlying renal injury and contribute to worsening function or obstetric complications. These hormonal mechanisms are consistent with animal and cell studies showing increased *UMOD* mRNA and protein expression with oestradiol exposure. In this context, the observed complications in our cohort may reflect a combination of physiological stress, altered immunomodulation, and enhanced pathogenic gene expression, warranting heightened clinical vigilance. Pregnancy complications must now be more rigorously explored in a larger ADTKD-*UMOD* cohort and by methodologically limiting selection bias.

Next, delayed diagnosis of ADTKD-*UMOD* was common despite predictable inheritance and consistent median age of kidney failure. Many individuals reached advanced CKD without a molecular diagnosis, many only identified through research efforts or retrospective genetic testing. This diagnostic inertia reflects broader patterns described in the 2024 Rare Barometer survey, which reported a mean time to diagnosis of 4.7 years across 6,507 patients with rare diseases in Europe (Faye *et al.*, 2024). Several determinants highlighted in that study were mirrored in this cohort, particularly: early symptom onset, misdiagnosis, and multiple healthcare contacts. Most individuals in the 2024 Rare Barometer survey presented with symptoms before age 30, a group shown to be at significantly higher risk of diagnostic delay (OR = 3.11 for childhood, 4.79 for adolescence). From personal clinical observation, multiple prior consultations and frequent misdiagnosis e.g. reflux nephropathy or hypertensive nephrosclerosis in those with ADTKD-*UMOD* also aligned with the survey's findings (OR = 5.15 and OR = 2.48, respectively).

Uniquely within Nephrology, the presence of an established renal replacement therapy pathway may reduce the perceived urgency for diagnostic precision, particularly in the absence of approved disease-modifying therapies. This contrasts with other specialties where

aetiology directly informs management and, in some cases, survival. As a result, both patients and clinicians may be less motivated to pursue a molecular diagnosis. This disincentive is compounded by health system-level barriers. The Rare Barometer survey identified Northern and Western Europe as regions with higher risk of diagnostic delay (OR = 2.15 and 1.96, respectively) likely reflecting more complex referral systems and the UK's 'three-tier healthcare system' in the National Health Service. The UK, situated within both regions, exemplifies this trend despite having a strong specialist care infrastructure and rare disease registry initiatives like RaDaR ('National Registry of Rare Kidney Diseases (RaDaR)', 2022; Faye *et al.*, 2024).

These findings are corroborated by recent data from the national RaDaR cohort, which reported that individuals with rare kidney diseases, including ADTKD-*UMOD*, have significantly higher five-year incidence of kidney failure (28%) compared to the general CKD population (1%) yet better overall survival (standardised mortality ratio 0.42) (Wong *et al.*, 2024). This reinforces the importance of early and accurate diagnosis in rare diseases. While overall survival may be preserved, delayed diagnosis precludes targeted surveillance, genetic and reproductive counselling as well as access to potential interventions within the critical therapeutic window (Wong *et al.*, 2024). Indeed, the median age at kidney failure for ADTKD-*UMOD* in this UK cohort was 50 years and the mean age at diagnosis only 7.7 years prior, highlights the narrow window of opportunity for intervention in addition to missed opportunities for future clinical trial eligibility if this delayed diagnostic trend continues.

Nevertheless, uptake of genetic testing for ADTKD-*UMOD* remains inconsistent across UK renal centres, likely due to a combination of low awareness of tubulointerstitial diseases, uneven access to genomics services, and limited confidence in genetics among clinicians. Educational deficits in rare and genetic kidney disease persist within adult nephrology contributing to under-recognition and late referral. As shown in the Rare Barometer survey, such delays carry substantial downstream consequences. Not only do they preclude timely cascade testing and reproductive planning, but they also delay access to supportive services and increase unmet psychological and financial needs (Faye *et al.*, 2024). In this context, the missed opportunity for clinical trial inclusion, particularly with novel therapies on the horizon, is another critical consequence.

Improving early diagnosis of ADTKD-*UMOD* in the UK requires system-wide change. While the centralised NHS referral model offers an opportunity to streamline diagnostics, current pathways are disjointed (Nun *et al.*, 2025). Embedding rare kidney disease flags into electronic health records, expanding access to panel-based testing, integration of genomic practitioners in nephrology units and leveraging platforms like RaDaR for early case detection will be key. Equally, embedding genetics education within nephrology training and creating diagnostic decision-support tools can empower general nephrologists to suspect and investigate monogenic disease earlier as well as the emphasis on 'don't let the cause of a person's CKD go unexplained' reinforcing the term CKDx as an important clinical consideration (Halbritter *et al.*, 2025).

Ultimately, addressing delayed diagnosis in ADTKD-*UMOD* will require both top-down system reform and bottom-up cultural change in how kidney clinicians approach unexplained CKD (CKDx) (Halbritter *et al.*, 2025). As genomic medicine continues to evolve, the UK is well-placed to lead a new paradigm of precision nephrology; one that prioritises early recognition, diagnosis, and management of monogenic kidney diseases.

6.3 Beyond Gout: Revisiting Salt-Wasting and Uric Acid in ADTKD-UMOD

Historically, ADTKD-*UMOD* has been associated with defective salt handling in the thick ascending limb (TAL) and impaired renal urate excretion, manifesting as low fractional excretion of uric acid and early-onset gout (Olinger *et al.*, 2020). In this work, salt-wasting symptoms were reported variably across the cohort, with differences by sex and *UMOD* variant, and, although not statistically significant, numbers were very small. Whilst the underlying mechanisms remain unclear, this phenotypic variability suggests a potential role for hormonal or transporter-mediated modulation of TAL function.

Experimental models provide key mechanistic insights. Trudu *et al.*, 2013 demonstrated that common noncoding *UMOD* promoter variants increase uromodulin expression, enhance NKCC2 phosphorylation, and raise blood pressure. In transgenic mice overexpressing wild-type uromodulin, this resulted in salt-sensitive hypertension and tubular injury, despite preserved glomerular function. These findings support a gain-of-function model whereby membrane-anchored uromodulin amplifies NKCC2 activity, leading to sodium retention and tubulointerstitial damage. Conversely, in ADTKD-*UMOD*, mutant uromodulin is retained in the

endoplasmic reticulum (ER), impairing apical trafficking and surface expression of NKCC2 and ROMK resulting in chronic but subtle salt-wasting (Bollée *et al.*, 2011; Devuyst, Olinger and Rampoldi, 2017).

This mechanistic spectrum is extended by findings from Brunati *et al.*, 2015, who showed that impaired hepsin-mediated cleavage of uromodulin causes its intracellular accumulation and dysregulation of NKCC2 in the TAL. Hepsin-deficient mice developed a Bartter-like phenotype under salt-loading conditions, with tubular injury, NaCl wasting, and loss of blood pressure regulation. These models collectively demonstrate that both overexpressed and misprocessed uromodulin disrupt TAL sodium homeostasis and render this nephron segment vulnerable to proteostatic stress.

Clinical parallels are found in inherited salt-wasting tubulopathies such as Bartter and Gitelman syndromes. Longitudinal studies show that in the setting of chronic hypokalaemia and salt-wasting, kidney function is often affected, suggesting that these features may drive tubulointerstitial fibrosis and tubular atrophy (Balavoine *et al.*, 2011; Tseng *et al.*, 2012; Liu *et al.*, 2023). In contrast, ADTKD-*UMOD* presents a more complex pathophysiological picture. Although it shares phenotypic overlap with these syndromes in terms of salt-wasting, the co-existence of mutant uromodulin-induced ER stress and early tubular injury suggests a synergistic mechanism. Ultimately, salt-wasting signs or symptoms did not statistically associate with kidney survival, but numbers were small, there are many confounders in the measurement of 'salt-wasting' and symptoms are, of course, subjective.

These disease models highlight the importance of tightly regulated sodium transport in maintaining tubular integrity. In ADTKD-*UMOD*, salt-wasting may not be overt, but chronic hypovolaemia and compensatory renin-angiotensin-aldosterone activation might contribute to progressive kidney fibrosis. Similar mechanisms are described in Bartter syndrome, where sustained RAAS activation leads to interstitial injury (Zhu *et al.*, 2019). Together, these data support a role for even mild salt-wasting as a pathophysiologically relevant, potentially modifiable contributor to ADTKD-*UMOD* progression especially in the context of increased energy demands and mitochondrial vulnerability that, based on the data from this thesis, we cannot yet rule out. Targeting sodium transport or maladaptive RAAS responses may offer therapeutic opportunities (Trudu *et al.*, 2013; Devuyst *et al.*, 2019) but, first, more work needs

to be done to better characterise salt-wasting in larger ADTKD-*UMOD* cohorts, perform genotype-phenotype correlations and further mechanistic studies.

Unexpectedly, serum urate levels were not uniformly elevated in our cohort, particularly among carriers of the UK *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant, despite biochemical evidence of impaired renal urate handling. Moreover, serum urate was not predictive of kidney disease progression. These findings challenge the traditional view of hyperuricaemia as a prognostic hallmark of ADTKD-*UMOD* and instead point toward the existence of metabolic sub-phenotypes within the disease spectrum. The identification of such sub-phenotypes has significant implications for precision medicine, as it reveals pathophysiological heterogeneity even among individuals with the same *UMOD* variant. Variable downstream effects on urate handling, salt reabsorption, and tubular metabolism may reflect the influence of additional genetic or environmental modifiers.

Crucially, the lack of a consistent hyperuricaemic phenotype and the absence of a predictive relationship with disease progression argue for a more nuanced approach to diagnosis, risk stratification, and therapeutic targeting. Stratifying patients into distinct metabolic subgroups may inform the selective use of urate-lowering therapies, enable the development of more meaningful biomarkers, and refine the design of clinical trials by identifying those most likely to benefit from specific interventions. Ultimately, delineating these subtypes enhances our understanding of ADTKD-*UMOD* pathogenesis and supports the development of personalised therapeutic strategies.

This reinterpretation is supported by emerging genetic evidence, which implicates variants in *ABCG2* (a uric acid transporter encoding gene) as modifiers of kidney outcomes, independent of serum urate levels and that those with hyperuricaemia show variable fractional excretion of uric acid results (Kannangara *et al.*, 2016; Stiburkova *et al.*, 2019; Ohashi, Kuriyama, *et al.*, 2023). These variants may act through pathways involving mitochondrial dysfunction, interstitial fibrosis or tubular injury, rather than through urate accumulation *per se* (Kang *et al.*, 2002; Balakumar *et al.*, 2020; Braga, Foresto-Neto and Camara, 2020; Sun *et al.*, 2021). Indeed, the absence of hyperuricaemia in carriers of the UK *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant, despite impaired urate handling, adds further complexity to the debate over uric acid's role in chronic kidney disease. While experimental

models have linked hyperuricaemia to oxidative stress, endothelial dysfunction, RAAS activation, and the metabolic syndrome (Johnson *et al.*, 2018), our findings suggest that elevated serum urate may not be neither necessary nor sufficient to drive progressive kidney injury in ADTKD-*UMOD.* More broadly, this is further supported by Mendelian Randomisation studies, which have failed to establish a causal relationship between serum urate and CKD progression, despite robust epidemiological associations (Johnson *et al.*, 2018). This data points toward a more complex interaction potentially involving intracellular urate effects, variant-specific pathophysiology, and compensatory metabolic adaptations that may mitigate systemic hyperuricaemia in certain genetic backgrounds. Furthermore, it may be that hyperuricaemia is not the primary driver of tubulointerstitial injury, but rather a biomarker of underlying tubular dysfunction or impaired urate transport. In this context, ABCG2 dysfunction may exert pathogenic effects via urate-independent mechanisms, such as mitochondrial stress, impaired autophagy, inflammatory signalling or other mechanisms discussed in this thesis. Nevertheless, the frequent presence of hyperuricaemia in ADTKD-*UMOD* could reflect the confounding and aggregated effects of impaired urate excretion, due to both uromodulin- and ABCG2-mediated transport defects, mitochondrial stress and systemic adaptations to reduced nephron mass. Thus, while elevated urate may not be causative, it remains a phenotypic hallmark that co-segregates with other modifiers of kidney injury in this disease.

Of particular interest is the role of *ABCG2* as a modifier of kidney function in ADTKD-*UMOD*, independent of its urate transport activity. Traditionally associated with extra-renal urate excretion, *ABCG2* appears to influence disease severity through alternative mechanisms, possibly involving intracellular trafficking, mitochondrial dynamics, or modulation of tubular injury responses. This challenges the urate-centric view of pathogenesis and highlights that, even in monogenic disorders like ADTKD-*UMOD*, disease progression is shaped by a broader landscape of metabolic and molecular modifiers.

These insights reframe ADTKD-*UMOD* not only as a model of inherited kidney fibrosis but also as a platform for exploring the pathogenic versus associative roles of uric acid in CKD. They highlight the importance of focusing on upstream cellular stress responses, rather than serum urate alone, and suggest that modifier genes like *ABCG2* may reveal novel therapeutic targets beyond conventional urate-lowering strategies.

Hyperuricaemia is also a recognised feature of several mitochondrial disorders and may result from converging mechanisms. Impaired mitochondrial ATP production leads to increased purine turnover and accumulation of uric acid as a breakdown product (Huang *et al.*, 2021). Moreover, mitochondrial cytopathies are often accompanied by lactic acidosis, with lactate competition with urate for excretion in the proximal tubule (Edwards *et al.*, 2019). Tubular mitochondrial injury itself may also impair urate handling directly. These mechanisms help explain the presence of hyperuricaemia in disorders such as MELAS and MERRF (Du *et al.*, 2024). In ADTKD-*UMOD*, mutant uromodulin is known to induce ER stress and disrupt mitochondrial function, suggesting that mitochondrial dysfunction may contribute to altered urate metabolism in a variant-specific manner. This may explain why some *UMOD* variants, such as the UK *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant, do not result in hyperuricaemia, whereas others are associated with elevated urate levels. The identification of *ABCG2* as a modifier, independent of urate levels, further implicates mitochondrial and cellular stress pathways in shaping disease progression rather than the direct effects of uric acid.

While ABCG2 is well known for its role in urate efflux (Dehghan *et al.*, 2008), recent data suggest its broader physiological functions include the transport of porphyrins and protection against intracellular accumulation of toxic metabolites such as protoporphyrin IX. Experimental studies in ABCG2-deficient hepatocytes have shown disrupted mitochondrial morphology, impaired oxidative phosphorylation, and increased reactive oxygen species (ROS) generation, all of which are mediated through intracellular porphyrin accumulation and dysregulation of mitochondrial fission via DRP-1 activation (Lin *et al.*, 2013). Analogous processes may occur in renal tubular cells, where *ABCG2* is also expressed, and where mitochondrial integrity is critical for energy-dependent solute transport. In the context of ADTKD-*UMOD*, which already features mitochondrial vulnerability and ER stress, the presence of *ABCG2* risk alleles could compound injury through additive or synergistic metabolic strain (Tseng *et al.*, 2018). This model helps explain why some individuals with *ABCG2* variants experience more rapid kidney function decline even in the absence of overt hyperuricaemia. It also offers a potential reconciliation of apparently discordant findings: while serum urate is frequently elevated in ADTKD-*UMOD* and associated with impaired urate handling, it may represent a parallel phenotypic marker rather than a direct driver of tubular damage. These

insights support a shift in perspective, elevated urate may not be causative, but may instead co-segregate with underlying defects in mitochondrial function or tubular resilience, mediated in part by *ABCG2* dysfunction (Matsuo *et al.*, 2016).

The finding that uric acid transporter genes modify kidney survival in ADTKD-*UMOD* has several key implications. First, it reinforces the concept that disease progression is not solely driven by the primary *UMOD* variant but may also be influenced by polygenic modifiers, including genes regulating urate transport. Second, it suggests that uric acid transporters may play active roles in renal cellular homeostasis, influencing stress responses, metabolic adaptation, and tubular resilience. Third, it opens the door to novel therapeutic strategies targeting urate transporter expression, trafficking, or function, particularly those expressed in the kidney or intestine, rather than focusing solely on lowering serum uric acid. Finally, these findings support a precision medicine framework in which urate transporter genotype may help identify high-risk individuals and guide trial design or therapeutic decision-making.

Given these insights, future research should focus on dissecting the role of uric acid transporters in modifying ADTKD-*UMOD* progression. Functional studies in cellular and animal models are essential. Renal tubular cells expressing mutant *UMOD* could be engineered to knock down or overexpress *ABCG2*, *SLC5A12* and *SLC2A9* enabling investigation of their roles in ER stress, mitochondrial function, and cellular viability. iPSC-derived kidney organoids from genotyped patients offer a 3D platform for mechanistic exploration (Forbes *et al.*, 2018). *In vivo,* crossing *UMOD* mutant mice with transporter-deficient models could clarify transporter-specific effects on fibrosis, urate handling, and kidney outcomes.

In parallel, large-scale genotype-phenotype studies should evaluate how transporter variants influence kidney survival, serum and urinary uric acid, salt-wasting severity, and serum and urine uromodulin. Integrating transcriptomic, proteomic, and metabolomic data may reveal modifier-specific signatures of disease activity. Pharmacological studies using transporter modulators, such as benzbromarone or probenecid, could determine whether targeting urate-independent functions of these proteins yields therapeutic benefit.

Polygenic risk scores incorporating urate transporter loci may ultimately support personalised stratification of patients within this monogenic disorder. Together, these strategies will

enhance our mechanistic understanding of ADTKD-*UMOD* and may uncover new pathways for intervention in a disease currently lacking targeted treatments.

6.4 Imaging in ADTKD-UMOD: Utility and Limitations

In contrast to other genetic kidney diseases such as autosomal dominant polycystic kidney disease (ADPKD), where imaging plays a central role in prognostication, conventional diagnostic modalities were of limited prognostic utility in this ADTKD-*UMOD* cohort although numbers were small (Alam *et al.*, 2015). In ADPKD, total kidney volume (TKV) measured by MRI is a validated predictor of progression risk, with the Mayo Imaging Classification stratifying patients based on age-adjusted TKV to identify rapid progressors (Irazabal *et al.*, 2015). TKV is also accepted as a surrogate endpoint in clinical trials and has guided therapeutic development, such as in the TEMPO 3:4 trial of Tolvaptan (Torres *et al.*, 2012).

More recently, advanced MRI techniques such as diffusion-weighted imaging (DWI), T1/T2 mapping, and magnetisation transfer imaging are being explored in ADPKD to quantify interstitial fibrosis and tubular integrity, offering potential for earlier and more nuanced detection of progression risk (Caroli *et al.*, 2023). These emerging modalities have yet to be evaluated in ADTKD-*UMOD* but could offer significant promise given the central role of tubulointerstitial fibrosis in this condition. Applying these imaging tools may provide novel, non-invasive biomarkers to complement genetic and molecular findings and facilitate risk stratification in patients.

By contrast, renal ultrasound findings in ADTKD-*UMOD* were heterogeneous, with cysts observed inconsistently and without clear genotype specificity or faithful correlation with progression. These findings highlight the limitations of conventional imaging in ADTKD-*UMOD* and reinforce the importance of genetic testing and molecular phenotyping. In parallel, studying more sensitive and targeted imaging approaches, as successfully applied in ADPKD, may enhance diagnostic precision and prognostication in ADTKD-*UMOD*.

6.5 Genetic and Epigenetic Modifiers of Progression: From GWAS to Regulatory Architecture

A major focus of this thesis was to investigate the contribution of genetic modifiers to the variability of disease progression among individuals with ADTKD-*UMOD*. The use of genome-

wide association studies (GWAS) in rare monogenic diseases has historically been constrained by limited cohort sizes and heterogeneous variant backgrounds. However, this study capitalised on the unique homogeneity of the UK ADTKD-*UMOD* cohort, particularly the high prevalence of the UMOD p.Val93_Gly97delinsAlaAlaSerCys, to reduce the potential confounding of the primary *UMOD* variant and enhance analytical power. Despite the relatively small sample size, several important observations emerged. Most notably, while the *UMOD* promoter variant (rs4293393) previously associated with kidney function in population studies (Trudu *et al.*, 2013) did not emerge as a significant modifier in this cohort, loci on chromosome 19 and the uric acid transport gene *ABCG2*, showed potential modifying effects. Intriguingly, *ABCG2* variants were associated with differences in kidney survival independent of serum urate levels, suggesting urate-independent roles in modulating tubular stress, mitochondrial function, or tubular epithelial resilience. This hypothesis is supported by *in vivo* models of hyperuricaemia in rats, which demonstrated that chronic exposure to elevated urate levels induces renal oxidative stress, mitochondrial dysfunction, and decreased ATP content in the renal cortex (Cristóbal-García *et al.*, 2015). Importantly, these effects were attenuated by antioxidant treatment, suggesting that ABCG2-associated urate handling may modulate tubular energy stress and vulnerability to fibrosis via oxidative pathways. Functional studies using CRISPR-edited tubular cells or transporter overexpression/knockdown models could further define whether *ABCG2* modifies disease via urate-independent roles in epithelial stress or mitochondrial homeostasis.

Furthermore, *ABCG2* plays a broader physiological role beyond renal urate handling. It is a high-capacity ATP-dependent urate exporter expressed in multiple tissues, including the kidney, intestine, blood-brain barrier, and placenta. In settings of kidney dysfunction or impaired renal urate excretion, *ABCG2*-mediated intestinal urate elimination becomes critically important. In apparently healthy individuals, even mild elevations in serum urate above 7.0 mg/dL have been shown to associate with accelerated eGFR decline over time (Kuma *et al.*, 2021), and longitudinal data suggest that maintaining urate levels below this threshold can slow the progression of kidney dysfunction. This raises the possibility that ABCG2 function may contribute to chronic kidney risk not only via overt hyperuricaemia, but also via subtle effects on tubular integrity and repair capacity. Studies in anuric patients and nephrectomised rodent models have demonstrated that intestinal *ABCG2* compensates for reduced urinary excretion and can account for up to 60% of daily urate clearance when kidney

function is lost. This highlights the importance of extra-renal urate handling. Moreover, *ABCG2* variants, especially the common functional polymorphisms p.Q141K and p.Q126X, have been robustly associated with early-onset hyperuricaemia and gout. These variants compromise *ABCG2*'s capacity for urate excretion and are enriched in individuals with childhood- and adolescent-onset gout outside the ADTKD setting. In addition to common variants, rare dysfunctional alleles of *ABCG2*, such as p.M131I and p.R236X, have been implicated in familial paediatric-onset hyperuricaemia and early-onset gout (Toyoda *et al.*, 2021). These variants exhibit severely reduced or absent urate transport function and segregate with hyperuricaemia across generations, underscoring the critical role of ABCG2 in early urate handling and its potential impact when renal reserve is limited. Furthermore, these variants can also affect the handling of urate-lowering therapies, including allopurinol, by modulating drug transport and metabolism. As such, *ABCG2* represents not only a potential genetic modifier of kidney progression, but also a pharmacogenetic determinant with relevance for therapeutic selection and response. The dual role of *ABCG2*, as a mediator of both urate homeostasis and drug excretion, positions it as a strategic target for precision medicine approaches in tubulointerstitial kidney disease. Moreover, recent cohort studies suggest that ABCG2 dysfunction may predict more rapid CKD progression independent of overt gout or hyperuricaemia (Ohashi, Kuriyama, *et al.*, 2023), further supporting its role as a prognostic biomarker and potential therapeutic target in chronic tubulointerstitial disease.

Chromatin interaction mapping further revealed a potential regulatory link between *ABCG2* and *PKD2*, a key genetic player in cystic kidney disease. This points toward shared pathogenic mechanisms between ADTKD-*UMOD* and other tubulointerstitial or cystic nephropathies. These findings reinforce the concept that ADTKD-*UMOD* is not solely a disorder of protein misfolding but is part of broader kidney pathophysiological networks. To further validate the regulatory relationship between *ABCG2* and *PKD2*, a CRISPRi/a model to disturb enhancer elements and assess effects on *PKD2* transcription and tubular phenotype could be attempted. CRISPRi/a are modified CRISPR systems that use a dead Cas9 protein to control gene expression. CRISPRi silences genes, while CRISPRa activates them, without cutting the DNA (Li *et al.*, 2020). In addition to epithelial effects, ABCG2 is also expressed in immune cell subsets such as macrophages. Emerging evidence suggests that *ABCG2* polymorphisms may influence inflammatory signalling or monocyte-derived cell trafficking (Cleophas *et al.*, 2017), providing a possible immunological dimension to its modifier role.

A further candidate modifier emerging from this work is *COX6B2*, a nuclear-encoded subunit of cytochrome c oxidase that contributes to mitochondrial oxidative phosphorylation (Nie *et al.*, 2020; Pham *et al.*, 2024). Recent studies have implicated *COX6B2* in the regulation of the AMPK pathway, a key energy-sensing mechanism that integrates metabolic stress signals and promotes mitochondrial biogenesis and autophagy. Downregulation of AMPK signalling is increasingly recognised as a contributor to metabolic dysfunction-associated kidney disease (MDAKD), particularly in the context of obesity, insulin resistance, and impaired fatty acid oxidation (Juszczak *et al.*, 2020). In the comprehensive review by Bansal and Chonchol, 2025, AMPK dysregulation was highlighted as a central mechanism linking metabolic stress with proximal tubular injury, hypoxia, and progressive fibrosis (Bansal and Chonchol, 2025). The identification of a *COX6B2* variant in individuals with more severe ADTKD-*UMOD* phenotypes raises the possibility that impaired *COX6B2* function may limit mitochondrial resilience under stress, reducing AMPK responsiveness and amplifying the deleterious effects of misfolded uromodulin and salt-wasting. This hypothesis is consistent with broader evidence from metabolic kidney diseases in which mitochondrial dysfunction, lipid accumulation, and defective mitophagy converge to drive tubular injury and interstitial fibrosis (Pavlović *et al.*, 2025). Future mechanistic work is needed to determine whether *COX6B2* variants impair renal mitochondrial homeostasis directly, or through modulation of energy-sensing networks such as AMPK, thereby acting as a metabolic amplifier of tubulointerstitial damage in genetically susceptible individuals. Recent experimental data further support a central role of AMPK-regulated mitochondrial quality control in ADTKD-*UMOD* pathogenesis (Kim *et al.*, 2023). In a knock-in mouse model expressing a common *UMOD* deletion variant, MANF enhanced mitophagy, suppressed maladaptive STING signalling, and reduced fibrosis via p-AMPK activation, restoring tubular energy homeostasis. This provides biological plausibility for our GWAS finding implicating COX6B2, suggesting that variants impairing mitochondrial resilience or AMPK responsiveness could exacerbate the toxic effects of misfolded uromodulin and metabolic stress, amplifying disease progression.

With this work, several caveats should be considered. The genetic homogeneity of this founder variant cohort, while analytically advantageous, may have obscured modifiers that act through pathways less active in this metabolically milder subgroup. The relative absence of hyperuricaemia and gout may have limited detection of modifiers linked to urate or sodium

handling, such as transporter genes whose influence may be more pronounced in individuals with greater biochemical disruption. Additionally, the lack of a GWAS signal at the *UMOD* promoter may reflect limitations in resolving cis-acting effects in this disease, for which more work needs to be done to explore this. Moreover, the key locus on chromosome 19 lies within a region of dense epigenetic regulation, known to harbour multiple enhancers and transcriptional control elements. Conditional analysis was unable to isolate a single causal variant, raising the possibility of true polygenicity or alternatively, of epigenetic mechanisms, such as DNA methylation or chromatin conformation, modulating gene expression in a context-specific manner. Combining statistical, epigenomic, functional-genetic, and single-cell approaches can pinpoint which regulatory elements and epigenetic modifications drive gene dysregulation at the chromosome-19 locus.

Furthermore, this work reinforces the importance of integrating genetic and phenotypic diversity in kidney research. The *UMOD-PDILT* locus exemplifies how a robust GWAS signal in European populations, driven by increased uromodulin expression, may not hold equivalent predictive or functional relevance in other ancestries. As demonstrated by Strauss-Kruger *et al.* (2024), the relationship between genotype and urinary uromodulin levels is substantially attenuated in African populations, reflecting both distinct linkage disequilibrium structure and differing environmental or physiological modulators. These insights accentuate the limitations of Eurocentric discovery frameworks in nephrology and highlight the need for inclusive genomic research to ensure the equitable development of biomarkers and therapeutic strategies (Strauss-Kruger *et al.*, 2024).

These observations at the chromosome 19 locus in ADTKD-*UMOD* align with findings in other monogenic and complex disorders, where loci enriched in regulatory elements, such as enhancers, promoters, and chromatin interaction hubs, have been shown to modify disease progression. For example, in chronic kidney disease, the well-characterised *APOL1* locus contains non-coding risk variants within cytokine-responsive enhancers, which modify disease risk and severity, particularly under inflammatory stress (Genovese, Tonna, *et al.*, 2010). Similarly, in Systemic Lupus Erythematosus, the *IRF5* locus harbours regulatory polymorphisms that affect promoter activity and epigenetic state, influencing interferon-driven immune dysregulation (Graham *et al.*, 2007; Gateva *et al.*, 2009). In cardiovascular disease, non-coding variants in an upstream enhancer of *SORT1* affect hepatic gene expression

through long-range chromatin looping, modulating LDL levels and coronary artery disease risk (Musunuru *et al.*, 2010). Even within nephrology, the *UMOD* promoter variants associated with hypertension and CKD in the general population represent a paradigmatic case of regulatory elements modifying tubular transporter activity and disease progression (Trudu *et al.*, 2013). These examples reinforce that epigenetically complex loci can act as critical modifiers of disease expression and support the hypothesis that chromosome 19 harbours regulatory variants, either alone or in combination, that influence kidney survival in ADTKD-*UMOD*. Therefore, prioritising functional annotation and epigenomic profiling of this locus may yield insights of both mechanistic and translational value.

These insights from ADTKD-*UMOD* reflect a broader trend across monogenic kidney diseases, where modifier genes increasingly explain variation in penetrance and progression. In autosomal dominant polycystic kidney disease (ADPKD), some variants have been linked to disease severity (Peters and Breuning, 2001), while a polygenic risk score for CKD significantly modulated outcomes among individuals with *PKD1/PKD2* variants (Khan *et al.*, 2023). In Alport syndrome, *LAMA5* and *MYH9* variants have been implicated as modifiers (Deltas *et al.*, 2023), and skewed X-inactivation explains phenotypic variability in female *COL4A5* carriers (Rheault *et al.*, 2010). In nephronophthisis, oligogenic inheritance with secondary variants in *TTC21B* and *CEP290* is well established (Davis *et al.*, 2011; Schueler *et al.*, 2016). Even in cystinosis and Gitelman syndrome, variants affecting oxidative stress pathways and RAAS-related genes respectively contribute to phenotypic variation (Besouw, Kleta and Bockenhauer, 2020; Nozu *et al.*, 2020). These findings position ADTKD-*UMOD* within an emerging paradigm: monogenic kidney diseases often function as genetically complex traits, modulated by rare, common, and epigenetic variants.

Looking forward, replication of these findings in larger and more diverse cohorts is essential. Ongoing international collaborations through ADTKD-NET (Europe) and RedCap registries (US) have facilitated a more expansive dataset and will enable validation of existing findings, testing for ancestry-specific modifiers and polygenic risk interaction (Harris *et al.*, 2019). In parallel, deeper mechanistic exploration of the chromosome 19 locus is warranted as previously discussed. Importantly, understanding these epigenetic mechanisms opens translational opportunities. Epigenetic modulation has already proven clinically effective in other diseases, DNA methyltransferase inhibitors (e.g. azacitidine, decitabine) and histone

deacetylase inhibitors (e.g. vorinostat) are approved for malignancies such as myelodysplastic syndromes and cutaneous T-cell lymphoma (Esteller, 2008; West and Johnstone, 2014). Inflammatory diseases also show promising responses to epigenetic therapies targeting immune gene expression (Ballestar, 2011; Angiolilli *et al.*, 2017). If dysregulated enhancer activity or transcription factor binding is confirmed in ADTKD-*UMOD*, CRISPR-dCas9-based epigenetic editing tools or pharmacological modifiers of chromatin state may offer future therapeutic avenues (Fadul, Arshad and Mehmood, 2023).

Recently published findings also facilitate the importance of transcript-level regulation as a potential modifier of disease expression in ADTKD-*UMOD*. Nanamatsu *et al.*, 2025 identified a novel alternatively spliced isoform of *UMOD* (AS-UMOD), generated through exon 10 skipping, which lacks the GPI anchor and localises to mitochondria rather than being secreted. This isoform was shown to enhance NAD^+ and ATP production via interactions with mitochondrial SLC25 transporters, promoting tubular energy recovery in models of mild or reversible AKI. Notably, AS-UMOD was absent in settings of severe AKI or CKD, suggesting that its induction represents a stress-adaptive and context-specific epithelial response. While highly relevant to tubular resilience and energy metabolism, such splicing events would not be captured in conventional GWAS, which detect common germline variants influencing stable phenotypes. Alternative splicing is governed by complex post-transcriptional regulation, often influenced by splicing factors, tissue stress, and epigenetic state rather than inherited sequence variants. Moreover, AS-UMOD expression is transient and conditional, arising only under specific injury contexts that are poorly modelled in cross-sectional population studies. Its discovery highlights the need to integrate transcriptomic and splicing analyses into studies of monogenic kidney disease, as they may reveal previously unappreciated mechanisms of adaptation and vulnerability. In ADTKD-*UMOD*, dysregulation of protective splicing responses, due to variant-induced ER stress or altered splicing factor activity, could plausibly compromise mitochondrial adaptation, accelerating disease progression. Therapeutically, the successful use of splice-switching oligonucleotides to induce AS-UMOD *in vivo* raises the exciting possibility of RNA-targeted therapies to enhance tubular resilience in hereditary kidney disorders (Havens and Hastings, 2016). In support of the broader concept that post-transcriptional regulation of UMOD may modulate tubular vulnerability, recent work has also implicated microRNAs in controlling UMOD expression. Specifically, miR-103a-3p was shown to directly target the UMOD 3'UTR and suppress its

expression in renal tubular cells. Silencing of miR-103a-3p in a rat model of hyperoxaluria led to increased UMOD and TRPV5 levels, enhanced calcium reabsorption, and protection against oxalate-induced tubular injury (Cui *et al.*, 2022). These findings reiterate the importance of microRNA-mediated regulation of UMOD in maintaining tubular homeostasis under stress and raise the possibility that microRNA perturbations may act as secondary modifiers of disease expression in ADTKD-*UMOD*.

Finally, integrating epigenetic, transcriptomic, and proteomic data into biomarker frameworks may refine risk prediction and clinical trial design. For example, urinary exosomes enriched with transcriptional regulators could serve as early, non-invasive markers of disease activity. Epigenetic signatures could also be added into polygenic models to better reflect gene-environment interactions underlying progression (Zhou *et al.*, 2008; Ung *et al.*, 2014; Wu *et al.*, 2022).

In summary, ADTKD-*UMOD* offers a model for understanding how monogenic disease expression is shaped by complex regulatory networks. A deeper mechanistic resolution of modifier pathways, particularly those involving epigenetic regulation and solute transporter function, will be essential to bridge the gap between genetic insight and disease-modifying therapies in ADTKD-*UMOD* and related disorders.

6.6 Uromodulin as a Mechanistic and Prognostic Biomarker

Biomarkers have become increasingly valuable in our understanding and study of genetic kidney diseases enabling earlier detection of disease activity and more accurate risk stratification than traditional clinical features or imaging alone. In ADPKD, biomarkers such as copeptin and urinary MCP-1 have been used to predict cyst growth and therapeutic response, while in Alport syndrome, urinary collagen fragments and podocyte injury markers help detect early glomerular damage (Janssens *et al.*, 2021; Rhode *et al.*, 2023). Building on these models, this research evaluated serum and urinary uromodulin as candidate biomarkers in ADTKD-*UMOD*, a disease in which tubular injury precedes clinically evident functional decline. Uromodulin, the most abundant protein in human urine, is exclusively produced in the thick ascending limb and plays key roles in salt handling, immune defence, and tubular homeostasis. Its dysregulation due to *UMOD* variants makes it a pathophysiologically relevant biomarker candidate (Devuyst *et al.*, 2019).

This work demonstrated that both serum and urinary uromodulin concentrations were significantly reduced in individuals with ADTKD-*UMOD*, consistent with intracellular retention and defective trafficking of mutant uromodulin (Schaeffer *et al.*, 2012; Kidd *et al.*, 2020). Importantly, both markers were associated with kidney disease progression even in a subset with the same *UMOD* p.Val93_Gly97delinsAlaAlaSerCys variant. This suggests that serum levels may more accurately reflect residual functional tubular mass and biosynthetic capacity *in vivo*. An additional novel finding was the partial preservation of uromodulin secretion in individuals with the p.Val93_Gly97delins variant, pointing to genotype-specific differences in protein trafficking and possibly explaining the milder uric acid phenotype seen in this subgroup. These results support the use of serum and urine uromodulin as a biomarker for risk stratification and therapeutic monitoring in ADTKD-*UMOD* and align with broader efforts in genetic kidney disease to integrate mechanistic biomarkers into clinical practice and trial design. By demonstrating independent associations between serum uromodulin and disease progression, this work advances prior research focused primarily on urinary uromodulin and provides a robust, mechanistically grounded biomarker that may help overcome the limitations of imaging and histology in this disease.

Understanding the varying degrees of protein trafficking impairment in ADTKD-*UMOD* is essential to elucidate the pathophysiological heterogeneity underlying this disease. While all pathogenic *UMOD* variants disrupt uromodulin maturation and apical secretion, the extent of ER retention differs between variants as shown by Kidd *et al.*, 2020. This variation in trafficking efficiency likely determines the severity of tubular stress, influencing the degree of interstitial damage and, ultimately, disease progression. For example, the partial preservation of uromodulin secretion observed in individuals carrying the p.Val93_Gly97delins variant may explain the relatively mild biochemical phenotype and attenuated hyperuricaemia in this subgroup (Kidd *et al.*, 2020). These genotype-specific differences in trafficking offer a mechanistic explanation for the diverse clinical trajectories seen across the ADTKD-*UMOD* spectrum. Moreover, as serum and urinary uromodulin emerge as candidate biomarkers, recognising differences in trafficking capacity is critical for accurate biomarker interpretation and risk stratification. Importantly, this information may also have therapeutic implications: individuals with residual trafficking may respond differently to emerging treatments targeting proteostasis, autophagy, or ER stress pathways. Stratifying patients by trafficking phenotype

could therefore refine trial design and enable more precise therapeutic targeting. Beyond ADTKD-*UMOD*, these insights contribute to the broader understanding of protein misfolding diseases, reinforcing the central role of intracellular processing defects in tubular injury and fibrosis.

Despite these promising findings, further work is needed to validate and extend the use of serum and urinary uromodulin as biomarkers in ADTKD-*UMOD*. Longitudinal studies in larger, multi-centre cohorts are required to confirm their predictive value over time and across different *UMOD* variants. Standardisation of assay methodologies, particularly for serum uromodulin, is essential to ensure reproducibility and comparability across studies and clinical settings. Future research should also explore how uromodulin levels change in response to emerging therapeutic interventions, to evaluate their potential role as pharmacodynamic biomarkers. Moreover, integrating uromodulin measurements with other candidate biomarkers (e.g. mitochondrial stress markers, urinary cytokines) and imaging or genetic modifiers may yield more comprehensive risk stratification models. Finally, mechanistic studies are needed to further delineate the pathways linking uromodulin secretion patterns to disease activity, and to understand whether residual secretion reflects true preservation of tubular integrity or compensatory processing in variant-specific contexts. Together, these efforts will be critical for establishing uromodulin as a clinically useful tool in the personalised management of ADTKD-*UMOD,* potentially as well as for other forms of CKD.

6.7 Polygenic Risk in a Monogenic Disease: A New Layer of Precision

While ADTKD-*UMOD* is a dominantly inherited monogenic disease, the findings presented in this work strongly support the presence of a polygenic component influencing disease severity. The successful application of a chronic kidney disease (CKD) polygenic risk score (PRS) to predict kidney survival in this cohort is particularly notable (Khan *et al.*, 2022). This represents the first demonstration that common variants, when aggregated into a risk score, can modify clinical outcomes in ADTKD-*UMOD*, thereby challenging the classical notion of fully penetrant Mendelian inheritance.

The polygenic signal likely captures the cumulative influence of background pathways, such as inflammation, fibrosis susceptibility, tubular regeneration, or vascular integrity, that are independent of *UMOD* dysfunction yet fundamentally shape the course of CKD. These results

reposition ADTKD-*UMOD* not simply as a Mendelian disorder with binary outcomes, but rather as a complex, polygenic-modulated disease, whereby both rare and common genetic variants interact to determine phenotype expression.

This model is increasingly recognised across other monogenic conditions. In familial hypercholesterolaemia, polygenic scores for LDL cholesterol modify the degree of hyperlipidaemia and cardiovascular risk, even in carriers of *LDLR* or *APOB* variants (Khera *et al.*, 2018). In hereditary breast and ovarian cancer, background polygenic risk significantly alters the penetrance and age of onset among *BRCA1/2* variant carriers (Mavaddat *et al.*, 2019). Similarly, in cystic fibrosis, polygenic variation in inflammatory and epithelial repair pathways contributes to variable lung disease severity among individuals with identical *CFTR* genotypes (Corvol *et al.*, 2015).

Most relevant to nephrology is the recent study by Khan *et al.*, 2023, which demonstrated that polygenic background substantially modifies CKD risk in individuals with pathogenic variants in *PKD1*, *PKD2* (ADPKD), and *COL4A* genes (COL4A-AN). Using data from the UK Biobank and All of Us, the study showed that monogenic variant carriers in the highest PRS tertile had dramatically higher CKD risk than those in the lowest tertile, for example, a 54-fold vs. 3-fold increased risk among ADPKD carriers. These findings strongly reinforce the notion that polygenic architecture modulates the penetrance and severity of monogenic kidney diseases and validate the application of PRS in rare disease settings.

In this context, the results from the current study place ADTKD-*UMOD* within this emerging paradigm. Identifying individuals with both high monogenic burden (e.g. pathogenic *UMOD* variant) and high polygenic risk for CKD could enable more accurate prediction of disease progression, facilitate risk-adapted monitoring, and refine eligibility criteria for clinical trials. Furthermore, understanding the interaction between polygenic risk and variant-specific features, such as differential uromodulin trafficking, could support more nuanced patient stratification and, in time, therapeutic targeting. Ultimately, incorporating polygenic frameworks into the study of monogenic diseases like ADTKD-*UMOD* reflects a broader shift toward precision medicine, where outcomes are viewed along a spectrum shaped by the full complement of inherited variation.

In the clinical setting, polygenic scores could be used to refine risk prediction among carriers of *UMOD* variants, supporting more personalised surveillance strategies based on predicted rate of progression. This may be particularly valuable in childhood or early adulthood, when eGFR remains preserved but long-term risk varies widely. Polygenic data could also inform shared decision-making around emerging interventions, such as early use of nephroprotective agents, by identifying individuals at highest risk of decline. In families undergoing genetic counselling, background polygenic risk may help clarify the potential for variable expressivity or incomplete penetrance. In individuals without a clear monogenic diagnosis, polygenic scores could also serve as an adjunct to assess CKD risk early, even before clinical symptoms emerge.

In the clinical trial context, PRS may enable smarter patient selection, for example, enriching trial populations with those at higher risk of progression, thereby improving power to detect treatment effects, especially in the rare disease setting where recruitment is more challenging. Polygenic background could also be used to stratify randomisation arms, enabling adjustment for background genetic risk. Trials targeting early intervention or disease modification may particularly benefit from PRS-based identification of individuals in the preclinical high-risk window, when intervention is most likely to alter long-term outcomes. Moreover, as pharmaco-genomic and gene-environment interactions become more deeply studied, polygenic risk may influence not only progression risk but treatment responsiveness, offering a new frontier for precision therapeutics in inherited kidney disease.

Despite these promising findings, further work is needed to establish the clinical utility of polygenic risk scoring in ADTKD-*UMOD*. In particular, the PRS developed in this study requires prospective validation in independent and ideally multi-ethnic ADTKD-*UMOD* cohorts, using longitudinal data to determine whether it can reliably predict disease progression over time. Future studies should also assess how polygenic risk integrates with other predictive features, such as *UMOD* variant class, serum and urine uromodulin levels, and clinical phenotype, to improve overall risk stratification. Additionally, it will be important to explore the performance and calibration of existing PRSs across different ancestral backgrounds, ensuring equitable application in diverse populations. Ultimately, validating and refining the PRS through such studies will be essential to enable its translation into both clinical and trial settings.

6.8 Implications for Diagnosis, Classification, and Clinical Management

These findings have immediate implications for the classification and management of ADTKD-*UMOD*. Firstly, they argue against the use of serum urate or gout as standalone diagnostic criteria, particularly in variant subgroups where these features are absent or attenuated. Secondly, they highlight the limited prognostic utility of conventional clinical and histopathological markers and instead support a model that integrates genetic diagnosis with molecular biomarkers (such as serum and urine uromodulin) and, where available, polygenic risk profiling. This multi-dimensional approach may enable earlier and more precise diagnosis, particularly in individuals with atypical or non-classic presentations.

Improving diagnosis in the UK will require greater awareness among clinicians, streamlined access to genetic testing, and wider implementation of national rare disease diagnostic pathways. Current diagnostic rates for ADTKD-*UMOD* remain suboptimal, often delayed by years due to non-specific features and reliance on clinical suspicion alone. Further embedding ADTKD gene panels within mainstream nephrology practice, particularly for patients with familial CKD, tubulointerstitial histology, or unexplained hyperuricaemia, would support earlier identification. Integration with the UK Rare Kidney Disease Registry (RaDaR) and systematic screening of at-risk family members could also enhance detection and improve cascade testing.

In parallel, there is a growing need to develop a clinical trial-ready cohort for ADTKD-*UMOD*. This requires harmonised phenotyping, consensus definitions of progression, and reliable biomarkers to stratify patients by risk. The current findings, particularly around serum uromodulin and polygenic risk, significantly contribute to building this framework. Incorporating these markers into longitudinal follow-up, alongside digital health record linkage and real-world data capture, will enable rapid identification of trial-eligible patients. In turn, this will support the evaluation of emerging interventions targeting proteostasis, autophagy, or inflammation.

More broadly, ADTKD-*UMOD* can serve as a monogenic model of kidney fibrosis, the final common pathway of all CKD. As a disease driven by intracellular protein accumulation, ER stress, and tubular injury, it offers an amenable framework for studying early fibrogenesis in

isolation from confounding glomerular or systemic factors. This positions ADTKD-*UMOD* as a valuable platform for biomarker discovery, drug testing, and mechanistic exploration relevant to common forms of CKD. Insights gained from studying modifier genes, mitochondrial dysfunction, and tubular resilience in this setting may ultimately inform therapies applicable across the CKD spectrum that could have very broad benefit especially as the global burden of CKD is rapidly expanding (Luyckx, Tonelli and Stanifer, 2018).

In terms of disease classification, this work suggests that ADTKD-*UMOD* may be more accurately conceptualised as a spectrum disorder with both Mendelian and complex features. Such a framework mirrors recent discussions in genomic nephrology, where the boundaries between monogenic and polygenic conditions are increasingly recognised as blurred (Groopman *et al.*, 2019). For ADTKD-*UMOD*, this reframing could help resolve longstanding diagnostic uncertainties, guide variant interpretation, and facilitate better patient stratification for emerging therapies. It also accentuates the necessity of viewing monogenic diseases not in isolation, but as dynamic entities influenced by polygenic background, molecular heterogeneity, and environmental context, principles that are increasingly central to the future of precision nephrology.

6.9 Limitations and Future Directions for Translational Impact

While this thesis represents the most comprehensive characterisation of ADTKD-*UMOD* in the UK to date, spanning genetic, clinical, biomarker, and polygenic analyses, several limitations should be acknowledged.

First, although the cohort was large by the standards of rare disease research, statistical power remained limited for some analyses, particularly the genome-wide association study (GWAS). The genetic homogeneity introduced by the UK founder variant enhanced internal consistency but may have reduced the generalisability of findings to more genetically diverse populations or to those carrying rarer *UMOD* variants which are more prevalent in other populations globally. In particular, the reduced prevalence of gout and hyperuricaemia in the founder group may have masked associations that would be detectable in cohorts with higher biochemical penetrance.

Second, while serum and urinary uromodulin were measured cross-sectionally, prospective validation over time was not available. Longitudinal biomarker trajectories, and their ability to dynamically reflect disease progression or therapeutic response, remain unexplored and should be prioritised in future studies. Similarly, while the observed association between serum and urine uromodulin and kidney survival was independent of eGFR or serum Creatinine, mechanistic studies are needed to confirm whether serum levels truly reflect tubular biosynthetic capacity or whether they are confounded by other systemic processes.
Third, while the polygenic risk score (PRS) demonstrated prognostic utility in this cohort, the score was derived from general population data and not specifically optimised for rare tubulointerstitial disease. Furthermore, the applicability of this PRS in non-European ancestries remains uncertain due to known limitations in PRS across populations (Sud *et al.*, 2023). Future work should include ancestry-aware PRS development and calibration, along with independent validation in international ADTKD-*UMOD* cohorts.

Fourth, epigenetic regulation, despite being implicated by chromatin interaction and regional enrichment data, was not directly assessed. Given the location of the modifier signal on chromosome 19 within a region of known epigenetic complexity, future studies should incorporate methylation profiling, chromatin accessibility assays, or single-cell transcriptomics to investigate potential non-coding or environmentally responsive regulatory mechanisms. This would be particularly relevant for understanding variant-specific expression patterns and could also uncover novel therapeutic targets.

Fifth, while pregnancy complications were reported more frequently than previously described, this data was retrospectively captured and may be subject to reporting bias or confounding by indication. A dedicated, prospective study of obstetric outcomes in ADTKD-*UMOD*, with standardised outcome definitions and genotype-specific analysis, is warranted to better understand risk and inform pre-pregnancy counselling.

To address these limitations and build on the findings of this work, several future directions are now underway. I have assembled a validation cohort of over 400 additional ADTKD-*UMOD* patients, expanding the original UK-US dataset and enabling replication of key genetic, biomarker, and polygenic findings. This cohort is being further enriched through collaboration with the European ADTKD-NET consortium, a pan-European research network that facilitates

data harmonisation and international recruitment. In parallel, established ADTKD-*UMOD* disease registries such as the US RedCap registry, coordinated by Wake Forest University, offer valuable infrastructure for cross-cohort validation and global clinical trial readiness. Utilising these international datasets will be essential for evaluating the generalisability of biomarker and polygenic risk models, and for accelerating translational efforts across diverse populations.

The urgent need for a trial-ready cohort is underscored by the emergence of potential disease-modifying therapies. Notably, the small molecule BRD4780, initially developed for ADTKD-*MUC1*, has shown promise in pre-clinical models of ADTKD-*UMOD*, where it facilitates the clearance of mutant protein via interaction with the ER–Golgi trafficking receptor TMED9 (Dvela-Levitt *et al.*, 2019). This cross-applicability highlights the potential for shared therapeutic pathways across ADTKD subtypes and reinforces the importance of assembling deeply phenotyped and genetically stratified cohorts in anticipation of clinical trials. Establishing the infrastructure to rapidly identify eligible participants, based on genotype, biomarkers, and progression risk, will be critical for testing such agents as they enter early-phase studies.

Additional future priorities include:

- Prospective validation of serum and urinary uromodulin as biomarkers across multi-centre cohorts and therapeutic settings

- Multi-omic longitudinal studies to define molecular trajectories and identify novel therapeutic targets

- Mechanistic dissection of urate transporter gene function and their urate-independent roles in tubular injury

- Epigenetic profiling of key regulatory regions implicated in genetic modifier analysis

- Evaluation of ancestry-specific implementation of founder variant screening and PRS tools in clinical practice

- Clinical trial infrastructure development, utilising RaDaR and ADTKD-NET for harmonised phenotyping and cohort readiness

Together, these directions provide a clear roadmap for translating this body of research into improved diagnostics, predictive modelling, and ultimately, targeted therapies for individuals with ADTKD-*UMOD* and other forms of inherited CKD.

6.10 Concluding Remarks

This thesis presents the most detailed characterisation of ADTKD-*UMOD* undertaken in the UK to date, integrating genetic, clinical, biomarker, and polygenic data to unravel the complex modifiers of disease expression and progression. It establishes that while *UMOD* variants are the primary cause of disease, they do not act in isolation. Instead, disease severity is shaped by a dynamic interplay between variant-specific properties, background genetic architecture and probable epigenetic modifiers. Together, these findings challenge traditional views of ADTKD-*UMOD* as a purely Mendelian disorder and reposition it as a model for complex monogenic disease, one that is well-suited to the emerging paradigm of precision medicine.

Key advances made through this work include the improved characterisation and validation of a UK-specific founder variant (*UMOD* p.Val93_Gly97delinsAlaAlaSerCys), the demonstration of substantial intra-genotypic variability, and the validation of serum and urine uromodulin as mechanistically anchored biomarkers of disease progression. Genome-wide analyses revealed that urate transporter genes, particularly *ABCG2*, can modify kidney survival in a urate-independent manner, implicating these pathways in disease modulation. Chromatin interaction mapping uncovered a potential regulatory link between *ABCG2* and *PKD2*, suggesting convergence between cystic and tubulointerstitial nephropathies. Additionally, *COX6B2*, a mitochondrial gene involved in oxidative phosphorylation and the AMPK signalling axis, emerged as a candidate modifier, highlighting the role of cellular energy metabolism in disease variability. Importantly, these discoveries position chromosome 19 as a locus of regulatory richness, where epigenetic mechanisms may prove a strong influence on disease trajectory. The incorporation of a chronic kidney disease polygenic risk score further highlights the relevance of common genetic background in shaping outcomes, even within a monogenic disease.

Together, these findings support a polygenic and regulatory model of ADTKD-*UMOD* progression, in which rare pathogenic variants, common susceptibility alleles, and non-coding regulatory elements interact to determine clinical phenotype. Importantly, this work also lays the foundation for future clinical translation. By assembling a genetically stratified cohort with rich phenotypic and molecular annotation, this thesis contributes directly to the formation of a trial-ready population, essential for evaluating targeted therapies such as BRD4780, which has shown promise in clearing mutant uromodulin in pre-clinical studies (Dvela-Levitt *et al.*, 2019). This work also reveals the potential of epigenetic modulation as a therapeutic avenue. As demonstrated in oncology and immunology, targeting chromatin dynamics, enhancer activity, or transcriptional repression offers a means to restore cellular homeostasis. If enhancer dysregulation or context-specific chromatin accessibility is validated as central to disease progression in ADTKD-*UMOD*, future interventions may harness CRISPR-based epigenetic editing or pharmacological modulators of chromatin state to alter disease course.

Furthermore, this work highlights the emerging importance of transcript-level regulation, potentially mediated by chromatin-modifying genes such as *KMT5C*, as an additional contributor to disease variability in ADTKD-*UMOD*. Although not directly investigated in this thesis, recent experimental data point to the role of alternative splicing in tubular stress adaptation. For example, a newly described isoform of uromodulin (AS-UMOD), which localises to mitochondria and enhances cellular energy production, was shown to support epithelial recovery in acute injury models. Although not yet studied in the context of ADTKD-*UMOD*, such findings suggest that dysregulation of protective splicing programmes, whether due to *UMOD* variants, ER stress, or altered splicing factor activity, could compromise mitochondrial resilience and contribute to disease progression. These transcript-level mechanisms, though invisible to conventional GWAS, may represent novel therapeutic targets. RNA-based approaches, such as splice-switching oligonucleotides, offer a promising avenue for modulating tubular stress responses and improving outcomes in inherited kidney diseases.

Finally, these findings have broad relevance to CKD as a whole. ADTKD-*UMOD*, with its defined genetic cause, measurable biomarkers, and early-stage tubular injury, serves as an ideal model for studying progressive renal fibrosis in isolation from glomerular disease or systemic

comorbidity. The insights gained here, particularly those related to cellular stress, mitochondrial vulnerability, and epithelial resilience, may inform therapeutic strategies across both rare and common disorders of the kidney, important in the global context whereby over 10% of the current global population is affected with CKD (Kovesdy, 2022).

In summary, this work advances the field of ADTKD-*UMOD* in fundamental ways: it clarifies the contribution of genetic and epigenetic modifiers, validates new biomarkers, reframes disease classification, and establishes the infrastructure for therapeutic innovation. It exemplifies how rare disease research, when deeply and systematically undertaken, can yield generalisable insights into kidney pathophysiology and shape the future for personalised nephrology.

References

1000 Genomes Project Consortium *et al.* (2015) 'A global reference for human genetic variation', *Nature*, 526(7571), pp. 68–74. Available at: https://doi.org/10.1038/nature15393.

Abd ElHafeez, S. *et al.* (2021) 'Methods to Analyze Time-to-Event Data: The Cox Regression Analysis', *Oxidative Medicine and Cellular Longevity*, 2021, p. 1302811. Available at: https://doi.org/10.1155/2021/1302811.

Alam, A. *et al.* (2015) 'Total Kidney Volume in Autosomal Dominant Polycystic Kidney Disease: A Biomarker of Disease Progression and Therapeutic Efficacy', *American Journal of Kidney Diseases: The Official Journal of the National Kidney Foundation*, 66(4), pp. 564–576. Available at: https://doi.org/10.1053/j.ajkd.2015.01.030.

Alesutan, I. *et al.* (2021) 'Circulating uromodulin inhibits vascular calcification by interfering with pro-inflammatory cytokine signalling', *Cardiovascular Research*, 117(3), pp. 930–941. Available at: https://doi.org/10.1093/cvr/cvaa081.

Alex, R. *et al.* (2024) 'Comparative Levels of Urinary Biomarkers of Renal Injury and Inflammation Among Patients With Diabetic Nephropathy With or Without Hyperuricemia', *Journal of Clinical Rheumatology: Practical Reports on Rheumatic & Musculoskeletal Diseases* [Preprint]. Available at: https://doi.org/10.1097/RHU.0000000000002068.

Al-Hamed, M.H. *et al.* (2022) 'Exome sequencing unravels genetic variants associated with chronic kidney disease in Saudi Arabian patients', *Human Mutation*, 43(12), pp. e24–e37. Available at: https://doi.org/10.1002/humu.24480.

Allen, N.E. *et al.* (2024) 'Prospective study design and data analysis in UK Biobank', *Science Translational Medicine*, 16(729), p. eadf4428. Available at: https://doi.org/10.1126/scitranslmed.adf4428.

Altshuler, D., Daly, M.J. and Lander, E.S. (2008) 'Genetic mapping in human disease', *Science (New York, N.Y.)*, 322(5903), pp. 881–888. Available at: https://doi.org/10.1126/science.1156409.

Angiolilli, C. *et al.* (2017) 'Histone deacetylase 3 regulates the inflammatory gene expression programme of rheumatoid arthritis fibroblast-like synoviocytes', *Annals of the Rheumatic*

Diseases, 76(1), pp. 277–285. Available at: https://doi.org/10.1136/annrheumdis-2015-209064.

Ankala, A. *et al.* (2015) 'Clinical applications and implications of common and founder mutations in Indian subpopulations', *Human Mutation*, 36(1), pp. 1–10. Available at: https://doi.org/10.1002/humu.22704.

Ars, E. and Torra, R. (2017) 'Rare diseases, rare presentations: recognizing atypical inherited kidney disease phenotypes in the age of genomics', *Clinical Kidney Journal*, 10(5), pp. 586–593. Available at: https://doi.org/10.1093/ckj/sfx051.

Avis, P.J. *et al.* (1985) 'Serum Tamm-Horsfall protein (THP) concentration in cyclosporin treated renal transplant recipients', *Lancet (London, England)*, 2(8447), p. 154. Available at: https://doi.org/10.1016/s0140-6736(85)90255-7.

Ayasreh, N. *et al.* (2018) 'Autosomal Dominant Tubulointerstitial Kidney Disease: Clinical Presentation of Patients With ADTKD-UMOD and ADTKD-MUC1', *American Journal of Kidney Diseases: The Official Journal of the National Kidney Foundation*, 72(3), pp. 411–418. Available at: https://doi.org/10.1053/j.ajkd.2018.03.019.

Bachmann, S., Koeppen-Hagemann, I. and Kriz, W. (1985) 'Ultrastructural localization of Tamm-Horsfall glycoprotein (THP) in rat kidney as revealed by protein A-gold immunocytochemistry', *Histochemistry*, 83(6), pp. 531–538. Available at: https://doi.org/10.1007/BF00492456.

Balakumar, P. *et al.* (2020) 'Mechanistic insights into hyperuricemia-associated renal abnormalities with special emphasis on epithelial-to-mesenchymal transition: Pathologic implications and putative pharmacologic targets', *Pharmacological Research*, 161, p. 105209. Available at: https://doi.org/10.1016/j.phrs.2020.105209.

Balavoine, A.S. *et al.* (2011) 'Phenotype-genotype correlation and follow-up in adult patients with hypokalaemia of renal origin suggesting Gitelman syndrome', *European Journal of Endocrinology*, 165(4), pp. 665–673. Available at: https://doi.org/10.1530/EJE-11-0224.

Ballestar, E. (2011) 'An introduction to epigenetics', *Advances in Experimental Medicine and Biology*, 711, pp. 1–11. Available at: https://doi.org/10.1007/978-1-4419-8216-2_1.

Bansal, A. and Chonchol, M. (2025) ‘Metabolic dysfunction-associated kidney disease: pathogenesis and clinical manifestations’, *Kidney International*, pp. S0085-2538(25)00351–5. Available at: https://doi.org/10.1016/j.kint.2025.01.044.

Bates, J.M. *et al.* (2004) ‘Tamm-Horsfall protein knockout mice are more prone to urinary tract infection: rapid communication’, *Kidney International*, 65(3), pp. 791–797. Available at: https://doi.org/10.1111/j.1523-1755.2004.00452.x.

Belton, J.-M. *et al.* (2012) ‘Hi-C: a comprehensive technique to capture the conformation of genomes’, *Methods (San Diego, Calif.)*, 58(3), pp. 268–276. Available at: https://doi.org/10.1016/j.ymeth.2012.05.001.

Bernascone, I. *et al.* (2006) ‘Defective intracellular trafficking of uromodulin mutant isoforms’, *Traffic (Copenhagen, Denmark)*, 7(11), pp. 1567–1579. Available at: https://doi.org/10.1111/j.1600-0854.2006.00481.x.

Bernascone, I. *et al.* (2010) ‘A transgenic mouse model for uromodulin-associated kidney diseases shows specific tubulo-interstitial damage, urinary concentrating defect and renal failure’, *Human Molecular Genetics*, 19(15), pp. 2998–3010. Available at: https://doi.org/10.1093/hmg/ddq205.

Besouw, M.T.P., Kleta, R. and Bockenhauer, D. (2020) ‘Bartter and Gitelman syndromes: Questions of class’, *Pediatric Nephrology (Berlin, Germany)*, 35(10), pp. 1815–1824. Available at: https://doi.org/10.1007/s00467-019-04371-y.

Bi, W. *et al.* (2020) ‘A Fast and Accurate Method for Genome-Wide Time-to-Event Data Analysis and Its Application to UK Biobank’, *American Journal of Human Genetics*, 107(2), pp. 222–233. Available at: https://doi.org/10.1016/j.ajhg.2020.06.003.

Bleyer, A.J. *et al.* (2010) ‘Clinical and molecular characterization of a family with a dominant renin gene mutation and response to treatment with fludrocortisone’, *Clinical Nephrology*, 74(6), pp. 411–422. Available at: https://doi.org/10.5414/cnp74411.

Bleyer, A.J. *et al.* (2014) ‘Variable clinical presentation of an MUC1 mutation causing medullary cystic kidney disease type 1’, *Clinical journal of the American Society of Nephrology: CJASN*, 9(3), pp. 527–535. Available at: https://doi.org/10.2215/CJN.06380613.

Bleyer, A.J. *et al.* (2022) 'Autosomal dominant tubulointerstitial kidney disease: more than just HNF1β', *Pediatric Nephrology (Berlin, Germany)*, 37(5), pp. 933–946. Available at: https://doi.org/10.1007/s00467-021-05118-4.

Bleyer, A.J. *et al.* (2023) 'Maternal health and pregnancy outcomes in autosomal dominant tubulointerstitial kidney disease', *Obstetric Medicine*, 16(3), pp. 162–169. Available at: https://doi.org/10.1177/1753495X221133150.

Bobulescu, I.A. and Moe, O.W. (2012) 'Renal transport of uric acid: evolving concepts and uncertainties', *Advances in Chronic Kidney Disease*, 19(6), pp. 358–371. Available at: https://doi.org/10.1053/j.ackd.2012.07.009.

Bokhove, M. *et al.* (2016) 'A structured interdomain linker directs self-polymerization of human uromodulin', *Proceedings of the National Academy of Sciences of the United States of America*, 113(6), pp. 1552–1557. Available at: https://doi.org/10.1073/pnas.1519803113.

Bolar, N.A. *et al.* (2016) 'Heterozygous Loss-of-Function SEC61A1 Mutations Cause Autosomal-Dominant Tubulo-Interstitial and Glomerulocystic Kidney Disease with Anemia', *American Journal of Human Genetics*, 99(1), pp. 174–187. Available at: https://doi.org/10.1016/j.ajhg.2016.05.028.

Bollée, G. *et al.* (2011) 'Phenotype and outcome in hereditary tubulointerstitial nephritis secondary to UMOD mutations', *Clinical journal of the American Society of Nephrology: CJASN*, 6(10), pp. 2429–2438. Available at: https://doi.org/10.2215/CJN.01220211.

Borg, R. *et al.* (2023) 'The Growing Challenge of Chronic Kidney Disease: An Overview of Current Knowledge', *International Journal of Nephrology*, 2023, p. 9609266. Available at: https://doi.org/10.1155/2023/9609266.

Boughton, A.P. *et al.* (2021) 'LocusZoom.js: interactive and embeddable visualization of genetic association study results', *Bioinformatics (Oxford, England)*, 37(18), pp. 3017–3018. Available at: https://doi.org/10.1093/bioinformatics/btab186.

Braga, T.T., Foresto-Neto, O. and Camara, N.O.S. (2020) 'The role of uric acid in inflammasome-mediated kidney injury', *Current Opinion in Nephrology and Hypertension*, 29(4), pp. 423–431. Available at: https://doi.org/10.1097/MNH.0000000000000619.

Bromberg, K.D. *et al.* (2017) 'The SUV4-20 inhibitor A-196 verifies a role for epigenetics in genomic integrity', *Nature Chemical Biology*, 13(3), pp. 317–324. Available at: https://doi.org/10.1038/nchembio.2282.

Brunati, M. *et al.* (2015) 'The serine protease hepsin mediates urinary secretion and polymerisation of Zona Pellucida domain protein uromodulin', *eLife*, 4, p. e08887. Available at: https://doi.org/10.7554/eLife.08887.

Canki, E., Kho, E. and Hoenderop, J.G.J. (2024) 'Urinary biomarkers in kidney disease', *Clinica Chimica Acta; International Journal of Clinical Chemistry*, 555, p. 117798. Available at: https://doi.org/10.1016/j.cca.2024.117798.

Caroli, A. *et al.* (2023) 'Diffusion magnetic resonance imaging for kidney cyst volume quantification and non-cystic tissue characterisation in ADPKD', *European Radiology*, 33(9), pp. 6009–6019. Available at: https://doi.org/10.1007/s00330-023-09601-4.

Chaiyarit, S. and Thongboonkerd, V. (2022) 'Oxidized forms of uromodulin promote calcium oxalate crystallization and growth, but not aggregation', *International Journal of Biological Macromolecules*, 214, pp. 542–553. Available at: https://doi.org/10.1016/j.ijbiomac.2022.06.132.

Chambers, J.C. *et al.* (2010) 'Genetic loci influencing kidney function and chronic kidney disease', *Nature Genetics*, 42(5), pp. 373–375. Available at: https://doi.org/10.1038/ng.566.

Chang, B.S.W. (2014) 'Ancient insights into uric acid metabolism in primates', *Proceedings of the National Academy of Sciences of the United States of America*, 111(10), pp. 3657–3658. Available at: https://doi.org/10.1073/pnas.1401037111.

Chang, C.C. *et al.* (2015) 'Second-generation PLINK: rising to the challenge of larger and richer datasets', *GigaScience*, 4, p. 7. Available at: https://doi.org/10.1186/s13742-015-0047-8.

Changalidis, A.I. *et al.* (2022) 'Aggregation of Genome-Wide Association Data from FinnGen and UK Biobank Replicates Multiple Risk Loci for Pregnancy Complications', *Genes*, 13(12), p. 2255. Available at: https://doi.org/10.3390/genes13122255.

Charlton, J.R. *et al.* (2021) 'Nephron number and its determinants: a 2020 update', *Pediatric Nephrology (Berlin, Germany)*, 36(4), pp. 797–807. Available at: https://doi.org/10.1007/s00467-020-04534-2.

Chen, W.C. *et al.* (2001) 'Effects of Tamm-Horsfall protein and albumin on calcium oxalate crystallization and importance of sialic acids', *Molecular Urology*, 5(1), pp. 1–5. Available at: https://doi.org/10.1089/109153601750124186.

Chertow, G.M. *et al.* (2024) 'Projecting the clinical burden of chronic kidney disease at the patient level (Inside CKD): a microsimulation modelling study', *EClinicalMedicine*, 72, p. 102614. Available at: https://doi.org/10.1016/j.eclinm.2024.102614.

Chun, J. *et al.* (2020) 'Autosomal Dominant Tubulointerstitial Kidney Disease-Uromodulin Misclassified as Focal Segmental Glomerulosclerosis or Hereditary Glomerular Disease', *Kidney International Reports*, 5(4), pp. 519–529. Available at: https://doi.org/10.1016/j.ekir.2019.12.016.

Cleophas, M.C. *et al.* (2017) 'ABCG2 polymorphisms in gout: insights into disease susceptibility and treatment approaches', *Pharmacogenomics and Personalized Medicine*, 10, pp. 129–142. Available at: https://doi.org/10.2147/PGPM.S105854.

Corris, V. *et al.* (2020) 'Health inequalities are worsening in the North East of England', *British Medical Bulletin*, 134(1), pp. 63–72. Available at: https://doi.org/10.1093/bmb/ldaa008.

Corvol, H. *et al.* (2015) 'Genome-wide association meta-analysis identifies five modifier loci of lung disease severity in cystic fibrosis', *Nature Communications*, 6, p. 8382. Available at: https://doi.org/10.1038/ncomms9382.

Cristóbal-García, M. *et al.* (2015) 'Renal oxidative stress induced by long-term hyperuricemia alters mitochondrial function and maintains systemic hypertension', *Oxidative Medicine and Cellular Longevity*, 2015, p. 535686. Available at: https://doi.org/10.1155/2015/535686.

Cui, Z. *et al.* (2022) 'miR-103a-3p Silencing Ameliorates Calcium Oxalate Deposition in Rat Kidney by Activating the UMOD/TRPV5 Axis', *Disease Markers*, 2022, p. 2602717. Available at: https://doi.org/10.1155/2022/2602717.

Das, S. *et al.* (2016) 'Next-generation genotype imputation service and methods', *Nature Genetics*, 48(10), pp. 1284–1287. Available at: https://doi.org/10.1038/ng.3656.

Davis, E.E. *et al.* (2011) 'TTC21B contributes both causal and modifying alleles across the ciliopathy spectrum', *Nature Genetics*, 43(3), pp. 189–196. Available at: https://doi.org/10.1038/ng.756.

Dawnay, A.B. and Cattell, W.R. (1981) 'Serum Tamm-Horsfall glycoprotein levels in health and in renal disease', *Clinical Nephrology*, 15(1), pp. 5–8.

Dehghan, A. *et al.* (2008) 'Association of three genetic loci with uric acid concentration and risk of gout: a genome-wide association study', *Lancet (London, England)*, 372(9654), pp. 1953–1961. Available at: https://doi.org/10.1016/S0140-6736(08)61343-4.

Delaneau, O. *et al.* (2013) 'Haplotype estimation using sequencing reads', *American Journal of Human Genetics*, 93(4), pp. 687–696. Available at: https://doi.org/10.1016/j.ajhg.2013.09.002.

Delaneau, O., Zagury, J.-F. and Marchini, J. (2013) 'Improved whole-chromosome phasing for disease and population genetic studies', *Nature Methods*, 10(1), pp. 5–6. Available at: https://doi.org/10.1038/nmeth.2307.

Deltas, C. *et al.* (2023) 'Genetic Modifiers of Mendelian Monogenic Collagen IV Nephropathies in Humans and Mice', *Genes*, 14(9), p. 1686. Available at: https://doi.org/10.3390/genes14091686.

Dessau, R.B. and Pipper, C.B. (2008) '["R"--project for statistical computing]', *Ugeskrift for Laeger*, 170(5), pp. 328–330.

Devlin, B. and Roeder, K. (1999) 'Genomic control for association studies', *Biometrics*, 55(4), pp. 997–1004. Available at: https://doi.org/10.1111/j.0006-341x.1999.00997.x.

Devuyst, O. *et al.* (2014) 'Rare inherited kidney diseases: challenges, opportunities, and perspectives', *Lancet (London, England)*, 383(9931), pp. 1844–1859. Available at: https://doi.org/10.1016/S0140-6736(14)60659-0.

Devuyst, O. *et al.* (2019) 'Autosomal dominant tubulointerstitial kidney disease', *Nature Reviews. Disease Primers*, 5(1), p. 60. Available at: https://doi.org/10.1038/s41572-019-0109-9.

Devuyst, O., Olinger, E. and Rampoldi, L. (2017) 'Uromodulin: from physiology to rare and complex kidney disorders', *Nature Reviews. Nephrology*, 13(9), pp. 525–544. Available at: https://doi.org/10.1038/nrneph.2017.101.

Doria, A. *et al.* (2020) 'Serum Urate Lowering with Allopurinol and Kidney Function in Type 1 Diabetes', *The New England Journal of Medicine*, 382(26), pp. 2493–2503. Available at: https://doi.org/10.1056/NEJMoa1916624.

Du, L. *et al.* (2024) 'Hyperuricemia and its related diseases: mechanisms and advances in therapy', *Signal Transduction and Targeted Therapy*, 9(1), p. 212. Available at: https://doi.org/10.1038/s41392-024-01916-y.

Dvela-Levitt, M. *et al.* (2019) 'Small Molecule Targets TMED9 and Promotes Lysosomal Degradation to Reverse Proteinopathy', *Cell*, 178(3), pp. 521-535.e23. Available at: https://doi.org/10.1016/j.cell.2019.07.002.

Easton, R.L. *et al.* (2000) 'Pregnancy-associated changes in the glycosylation of tamm-horsfall glycoprotein. Expression of sialyl Lewis(x) sequences on core 2 type O-glycans derived from uromodulin', *The Journal of Biological Chemistry*, 275(29), pp. 21928–21938. Available at: https://doi.org/10.1074/jbc.M001534200.

Eckardt, K.-U. *et al.* (2015) 'Autosomal dominant tubulointerstitial kidney disease: diagnosis, classification, and management--A KDIGO consensus report', *Kidney International*, 88(4), pp. 676–683. Available at: https://doi.org/10.1038/ki.2015.28.

Edwards, A. *et al.* (2019) 'A model of uric acid transport in the rat proximal tubule', *American Journal of Physiology. Renal Physiology*, 316(5), pp. F934–F947. Available at: https://doi.org/10.1152/ajprenal.00603.2018.

Edwards, N. *et al.* (2017) 'A novel homozygous UMOD mutation reveals gene dosage effects on uromodulin processing and urinary excretion', *Nephrology, Dialysis, Transplantation: Official Publication of the European Dialysis and Transplant Association - European Renal Association*, 32(12), pp. 1994–1999. Available at: https://doi.org/10.1093/ndt/gfx066.

Eichler, E.E. *et al.* (2010) 'Missing heritability and strategies for finding the underlying causes of complex disease', *Nature Reviews. Genetics*, 11(6), pp. 446–450. Available at: https://doi.org/10.1038/nrg2809.

Ekici, A.B. *et al.* (2014) 'Renal fibrosis is the common feature of autosomal dominant tubulointerstitial kidney diseases caused by mutations in mucin 1 or uromodulin', *Kidney International*, 86(3), pp. 589–599. Available at: https://doi.org/10.1038/ki.2014.72.

El-Achkar, T.M. *et al.* (2011) 'Tamm-Horsfall protein-deficient thick ascending limbs promote injury to neighboring S3 segments in an MIP-2-dependent mechanism', *American Journal of Physiology. Renal Physiology*, 300(4), pp. F999-1007. Available at: https://doi.org/10.1152/ajprenal.00621.2010.

El-Achkar, T.M. *et al.* (2013) 'Tamm-Horsfall protein translocates to the basolateral domain of thick ascending limbs, interstitium, and circulation during recovery from acute kidney injury', *American Journal of Physiology. Renal Physiology*, 304(8), pp. F1066-1075. Available at: https://doi.org/10.1152/ajprenal.00543.2012.

El-Achkar, T.M. and Dagher, P.C. (2006) 'Renal Toll-like receptors: recent advances and implications for disease', *Nature Clinical Practice. Nephrology*, 2(10), pp. 568–581. Available at: https://doi.org/10.1038/ncpneph0300.

El-Achkar, T.M. and Wu, X.-R. (2012) 'Uromodulin in kidney injury: an instigator, bystander, or protector?', *American Journal of Kidney Diseases: The Official Journal of the National Kidney Foundation*, 59(3), pp. 452–461. Available at: https://doi.org/10.1053/j.ajkd.2011.10.054.

Enko, D. *et al.* (2021) 'Individual uromodulin serum concentration is independent of glomerular filtration rate in healthy kidney donors', *Clinical Chemistry and Laboratory Medicine*, 59(3), pp. 563–570. Available at: https://doi.org/10.1515/cclm-2020-0894.

Esteller, M. (2008) 'Epigenetics in cancer', *The New England Journal of Medicine*, 358(11), pp. 1148–1159. Available at: https://doi.org/10.1056/NEJMra072067.

Evans, J.A. (2015) 'Old meets new: identifying founder mutations in genetic disease', *CMAJ: Canadian Medical Association journal = journal de l'Association medicale canadienne*, 187(2), pp. 93–94. Available at: https://doi.org/10.1503/cmaj.141509.

Fadul, S.M., Arshad, A. and Mehmood, R. (2023) 'CRISPR-based epigenome editing: mechanisms and applications', *Epigenomics*, 15(21), pp. 1137–1155. Available at: https://doi.org/10.2217/epi-2023-0281.

Fahed, A.C. *et al.* (2020) 'Polygenic background modifies penetrance of monogenic variants for tier 1 genomic conditions', *Nature Communications*, 11(1), p. 3635. Available at: https://doi.org/10.1038/s41467-020-17374-3.

Fain, P.R. *et al.* (2005) 'Modifier genes play a significant role in the phenotypic expression of PKD1', *Kidney International*, 67(4), pp. 1256–1267. Available at: https://doi.org/10.1111/j.1523-1755.2005.00203.x.

Faye, F. *et al.* (2024) 'Time to diagnosis and determinants of diagnostic delays of people living with a rare disease: results of a Rare Barometer retrospective patient survey', *European journal of human genetics: EJHG*, 32(9), pp. 1116–1126. Available at: https://doi.org/10.1038/s41431-024-01604-z.

Fedak, D. *et al.* (2016) 'Serum uromodulin concentrations correlate with glomerular filtration rate in patients with chronic kidney disease', *Polskie Archiwum Medycyny Wewnetrznej*, 126(12), pp. 995–1004. Available at: https://doi.org/10.20452/pamw.3712.

Forbes, T.A. *et al.* (2018) 'Patient-iPSC-Derived Kidney Organoids Show Functional Validation of a Ciliopathic Renal Phenotype and Reveal Underlying Pathogenetic Mechanisms', *American Journal of Human Genetics*, 102(5), pp. 816–831. Available at: https://doi.org/10.1016/j.ajhg.2018.03.014.

Francis, A. *et al.* (2024) 'Chronic kidney disease and the global public health agenda: an international consensus', *Nature Reviews. Nephrology*, 20(7), pp. 473–485. Available at: https://doi.org/10.1038/s41581-024-00820-6.

Freeman, C (2018) 'SP301 THE ECONOMIC BURDEN OF CHRONIC KIDNEY DISEASE: FINDINGS FROM A SYSTEMATIC LITERATURE REVIEW', *Nephrology Dialysis Transplantation*, 33(Issue suppl_1), p. i445.

French, J.D. and Edwards, S.L. (2020) 'The Role of Noncoding Variants in Heritable Disease', *Trends in genetics: TIG*, 36(11), pp. 880–891. Available at: https://doi.org/10.1016/j.tig.2020.07.004.

Fridley, B.L. and Biernacka, J.M. (2011) 'Gene set analysis of SNP data: benefits, challenges, and future directions', *European journal of human genetics: EJHG*, 19(8), pp. 837–843. Available at: https://doi.org/10.1038/ejhg.2011.57.

'FUMA GWAS' (2024). Available at: https://fuma.ctglab.nl.

Gast, C. *et al.* (2018) 'Autosomal dominant tubulointerstitial kidney disease-UMOD is the most frequent non polycystic genetic kidney disease', *BMC nephrology*, 19(1), p. 301. Available at: https://doi.org/10.1186/s12882-018-1107-y.

Gateva, V. *et al.* (2009) 'A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus', *Nature Genetics*, 41(11), pp. 1228–1233. Available at: https://doi.org/10.1038/ng.468.

GBD Chronic Kidney Disease Collaboration (2020) 'Global, regional, and national burden of chronic kidney disease, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017', *Lancet (London, England)*, 395(10225), pp. 709–733. Available at: https://doi.org/10.1016/S0140-6736(20)30045-3.

Genin, E. *et al.* (2004) 'Estimating the age of rare disease mutations: the example of Triple-A syndrome', *Journal of Medical Genetics*, 41(6), pp. 445–449. Available at: https://doi.org/10.1136/jmg.2003.017962.

Genovese, G., Tonna, S.J., *et al.* (2010) 'A risk allele for focal segmental glomerulosclerosis in African Americans is located within a region containing APOL1 and MYH9', *Kidney International*, 78(7), pp. 698–704. Available at: https://doi.org/10.1038/ki.2010.251.

Genovese, G., Friedman, D.J., *et al.* (2010) 'Association of trypanolytic ApoL1 variants with kidney disease in African Americans', *Science (New York, N.Y.)*, 329(5993), pp. 841–845. Available at: https://doi.org/10.1126/science.1193032.

Ghelichi-Ghojogh, M. *et al.* (2024) 'Environmental factors and chronic kidney disease: a case-control study', *Scientific Reports*, 14(1), p. 26511. Available at: https://doi.org/10.1038/s41598-024-72685-5.

Ghoussaini, M. *et al.* (2021) 'Open Targets Genetics: systematic identification of trait-associated genes using large-scale genetics and functional genomics', *Nucleic Acids Research*, 49(D1), pp. D1311–D1320. Available at: https://doi.org/10.1093/nar/gkaa840.

Giordano, C. *et al.* (2015) 'Uric Acid as a Marker of Kidney Disease: Review of the Current Literature', *Disease Markers*, 2015, p. 382918. Available at: https://doi.org/10.1155/2015/382918.

'Global, regional, and national burden of chronic kidney disease, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet, 2020. 395(10225): p. 709-733.' (no date).

Gong, K. *et al.* (2021) 'Autosomal dominant tubulointerstitial kidney disease genotype and phenotype correlation in a Chinese cohort', *Scientific Reports*, 11(1), p. 3615. Available at: https://doi.org/10.1038/s41598-020-79331-w.

Good, P.I. *et al.* (2023) 'Low nephron endowment increases susceptibility to renal stress and chronic kidney disease', *JCI insight*, 8(3), p. e161316. Available at: https://doi.org/10.1172/jci.insight.161316.

Graham, R.R. *et al.* (2007) 'Three functional variants of IFN regulatory factor 5 (IRF5) define risk and protective haplotypes for human lupus', *Proceedings of the National Academy of Sciences of the United States of America*, 104(16), pp. 6758–6763. Available at: https://doi.org/10.1073/pnas.0701266104.

Gresh, L. *et al.* (2004) 'A transcriptional network in polycystic kidney disease', *The EMBO journal*, 23(7), pp. 1657–1668. Available at: https://doi.org/10.1038/sj.emboj.7600160.

Groopman, E.E. *et al.* (2019) 'Diagnostic Utility of Exome Sequencing for Kidney Disease', *The New England Journal of Medicine*, 380(2), pp. 142–151. Available at: https://doi.org/10.1056/NEJMoa1806891.

GTEx Consortium (2013) 'The Genotype-Tissue Expression (GTEx) project', *Nature Genetics*, 45(6), pp. 580–585. Available at: https://doi.org/10.1038/ng.2653.

'GTEx Portal' (2024). Available at: https://www.gtexportal.org/home/.

Gudbjartsson, D.F. *et al.* (2010) 'Association of variants at UMOD with chronic kidney disease and kidney stones-role of age and comorbid diseases', *PLoS genetics*, 6(7), p. e1001039. Available at: https://doi.org/10.1371/journal.pgen.1001039.

Gudmundsson, S. *et al.* (2022) 'Variant interpretation using population databases: Lessons from gnomAD', *Human Mutation*, 43(8), pp. 1012–1030. Available at: https://doi.org/10.1002/humu.24309.

Gunay-Aygun, M. *et al.* (2010) 'PKHD1 sequence variations in 78 children and adults with autosomal recessive polycystic kidney disease and congenital hepatic fibrosis', *Molecular Genetics and Metabolism*, 99(2), pp. 160–173. Available at: https://doi.org/10.1016/j.ymgme.2009.10.010.

Guo, M.H. *et al.* (2016) 'Determinants of Power in Gene-Based Burden Testing for Monogenic Disorders', *American Journal of Human Genetics*, 99(3), pp. 527–539. Available at: https://doi.org/10.1016/j.ajhg.2016.06.031.

Gupta, Y. *et al.* (2023) 'Strong protective effect of the APOL1 p.N264K variant against G2-associated focal segmental glomerulosclerosis and kidney disease', *Nature Communications*, 14(1), p. 7836. Available at: https://doi.org/10.1038/s41467-023-43020-9.

Halbritter, J. *et al.* (2025) 'Chronic Kidney Disease of unexplained cause (CKDx): a consensus statement by the Genes & Kidney Working Group of the ERA', *Nephrology, Dialysis, Transplantation: Official Publication of the European Dialysis and Transplant Association - European Renal Association*, p. gfaf092. Available at: https://doi.org/10.1093/ndt/gfaf092.

Hallson, P.C. *et al.* (1997) 'Effects of Tamm-Horsfall protein with normal and reduced sialic acid content upon the crystallization of calcium phosphate and calcium oxalate in human urine', *British Journal of Urology*, 80(4), pp. 533–538. Available at: https://doi.org/10.1046/j.1464-410x.1997.00366.x.

Han, T. *et al.* (2024) 'Epidemiology of gout - Global burden of disease research from 1990 to 2019 and future trend predictions', *Therapeutic Advances in Endocrinology and Metabolism*, 15, p. 20420188241227295. Available at: https://doi.org/10.1177/20420188241227295.

Harris, P.A. *et al.* (2019) 'The REDCap consortium: Building an international community of software platform partners', *Journal of Biomedical Informatics*, 95, p. 103208. Available at: https://doi.org/10.1016/j.jbi.2019.103208.

Hateboer, N. *et al.* (1999) 'Comparison of phenotypes of polycystic kidney disease types 1 and 2. European PKD1-PKD2 Study Group', *Lancet (London, England)*, 353(9147), pp. 103–107. Available at: https://doi.org/10.1016/s0140-6736(98)03495-3.

Havens, M.A. and Hastings, M.L. (2016) 'Splice-switching antisense oligonucleotides as therapeutic drugs', *Nucleic Acids Research*, 44(14), pp. 6549–6563. Available at: https://doi.org/10.1093/nar/gkw533.

Hayashi, S. and Umeda, T. (2008) '35 years of Japanese policy on rare diseases', *Lancet (London, England)*, 372(9642), pp. 889–890. Available at: https://doi.org/10.1016/S0140-6736(08)61393-8.

He, C. and Bonasio, R. (2017) 'A cut above', *eLife*, 6, p. e25000. Available at: https://doi.org/10.7554/eLife.25000.

Hecker, D. *et al.* (2023) 'The adapted Activity-By-Contact model for enhancer-gene assignment and its application to single-cell data', *Bioinformatics (Oxford, England)*, 39(2), p. btad062. Available at: https://doi.org/10.1093/bioinformatics/btad062.

Hew-Butler, T.D. *et al.* (2018) 'Dehydration is how you define it: comparison of 318 blood and urine athlete spot checks', *BMJ open sport & exercise medicine*, 4(1), p. e000297. Available at: https://doi.org/10.1136/bmjsem-2017-000297.

Hill, N.R. *et al.* (2016) 'Global Prevalence of Chronic Kidney Disease - A Systematic Review and Meta-Analysis', *PloS One*, 11(7), p. e0158765. Available at: https://doi.org/10.1371/journal.pone.0158765.

Hingorani, A.D. *et al.* (2023) 'Performance of polygenic risk scores in screening, prediction, and risk stratification: secondary analysis of data in the Polygenic Score Catalog', *BMJ medicine*, 2(1), p. e000554. Available at: https://doi.org/10.1136/bmjmed-2023-000554.

Hirschhorn, J.N. (2009) ‘Genomewide association studies--illuminating biologic pathways’, *The New England Journal of Medicine*, 360(17), pp. 1699–1701. Available at: https://doi.org/10.1056/NEJMp0808934.

Hirst, J.A. *et al.* (2020) ‘Prevalence of chronic kidney disease in the community using data from OxRen: a UK population-based cohort study’, *The British Journal of General Practice: The Journal of the Royal College of General Practitioners*, 70(693), pp. e285–e293. Available at: https://doi.org/10.3399/bjgp20X708245.

Hoefele, J. *et al.* (2007) ‘Evidence of oligogenic inheritance in nephronophthisis’, *Journal of the American Society of Nephrology: JASN*, 18(10), pp. 2789–2795. Available at: https://doi.org/10.1681/ASN.2007020243.

Homer, N. *et al.* (2008) ‘Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays’, *PLoS genetics*, 4(8), p. e1000167. Available at: https://doi.org/10.1371/journal.pgen.1000167.

Huang, R., Fu, P. and Ma, L. (2023) ‘Kidney fibrosis: from mechanisms to therapeutic medicines’, *Signal Transduction and Targeted Therapy*, 8(1), p. 129. Available at: https://doi.org/10.1038/s41392-023-01379-7.

Huang, Z. *et al.* (2021) ‘From purines to purinergic signalling: molecular functions and human diseases’, *Signal Transduction and Targeted Therapy*, 6(1), p. 162. Available at: https://doi.org/10.1038/s41392-021-00553-z.

‘Human Protein Atlas’ (2024). Available at: https://www.proteinatlas.org.

Hung, A.M. *et al.* (2023) ‘Genetic Inhibition of APOL1 Pore-Forming Function Prevents APOL1-Mediated Kidney Disease’, *Journal of the American Society of Nephrology: JASN*, 34(11), pp. 1889–1899. Available at: https://doi.org/10.1681/ASN.0000000000000219.

Huynh, V.T. *et al.* (2020) ‘Clinical spectrum, prognosis and estimated prevalence of DNAJB11-kidney disease’, *Kidney International*, 98(2), pp. 476–487. Available at: https://doi.org/10.1016/j.kint.2020.02.022.

Intapad, S. (2023) ‘Uromodulin and Estrogen’, *Kidney360*, 4(9), pp. e1201–e1202. Available at: https://doi.org/10.34067/KID.0000000000000259.

Irazabal, M.V. *et al.* (2015) 'Imaging classification of autosomal dominant polycystic kidney disease: a simple model for selecting patients for clinical trials', *Journal of the American Society of Nephrology: JASN*, 26(1), pp. 160–172. Available at: https://doi.org/10.1681/ASN.2013101138.

Ix, J.H. and Shlipak, M.G. (2021) 'The Promise of Tubule Biomarkers in Kidney Disease: A Review', *American Journal of Kidney Diseases: The Official Journal of the National Kidney Foundation*, 78(5), pp. 719–727. Available at: https://doi.org/10.1053/j.ajkd.2021.03.026.

Izzi, C. *et al.* (2020) 'Variable Expressivity of HNF1B Nephropathy, From Renal Cysts and Diabetes to Medullary Sponge Kidney Through Tubulo-interstitial Kidney Disease', *Kidney International Reports*, 5(12), pp. 2341–2350. Available at: https://doi.org/10.1016/j.ekir.2020.09.042.

Jacobs, K.B. *et al.* (2012) 'Detectable clonal mosaicism and its relationship to aging and cancer', *Nature Genetics*, 44(6), pp. 651–658. Available at: https://doi.org/10.1038/ng.2270.

Jakkula, E. *et al.* (2008) 'The genome-wide patterns of variation expose significant substructure in a founder population', *American Journal of Human Genetics*, 83(6), pp. 787–794. Available at: https://doi.org/10.1016/j.ajhg.2008.11.005.

Janssens, P. *et al.* (2021) 'Enhanced MCP-1 Release in Early Autosomal Dominant Polycystic Kidney Disease', *Kidney International Reports*, 6(6), pp. 1687–1698. Available at: https://doi.org/10.1016/j.ekir.2021.03.893.

Jeyaraman, D. *et al.* (2024) 'Adverse pregnancy outcomes in pregnant women with chronic kidney disease: A systematic review and meta-analysis', *BJOG: an international journal of obstetrics and gynaecology*, 131(10), pp. 1331–1340. Available at: https://doi.org/10.1111/1471-0528.17807.

Johnson, B.G. *et al.* (2017) 'Uromodulin p.Cys147Trp mutation drives kidney disease by activating ER stress and apoptosis', *The Journal of Clinical Investigation*, 127(11), pp. 3954–3969. Available at: https://doi.org/10.1172/JCI93817.

Johnson, R.J. *et al.* (2018) 'Hyperuricemia, Acute and Chronic Kidney Disease, Hypertension, and Cardiovascular Disease: Report of a Scientific Workshop Organized by the National Kidney Foundation', *American Journal of Kidney Diseases: The Official Journal of the National*

Kidney Foundation, 71(6), pp. 851–865. Available at: https://doi.org/10.1053/j.ajkd.2017.12.009.

Johnson, R.J. *et al.* (2023) 'Uric Acid and Chronic Kidney Disease: Still More to Do', *Kidney International Reports*, 8(2), pp. 229–239. Available at: https://doi.org/10.1016/j.ekir.2022.11.016.

Johnson, R.J., Lanaspa, M.A. and Gaucher, E.A. (2011) 'Uric acid: a danger signal from the RNA world that may have a role in the epidemic of obesity, metabolic syndrome, and cardiorenal disease: evolutionary considerations', *Seminars in Nephrology*, 31(5), pp. 394–399. Available at: https://doi.org/10.1016/j.semnephrol.2011.08.002.

Juszczak, F. *et al.* (2020) 'Critical Role for AMPK in Metabolic Disease-Induced Chronic Kidney Disease', *International Journal of Molecular Sciences*, 21(21), p. 7994. Available at: https://doi.org/10.3390/ijms21217994.

Kang, D.-H. *et al.* (2002) 'A role for uric acid in the progression of renal disease', *Journal of the American Society of Nephrology: JASN*, 13(12), pp. 2888–2897. Available at: https://doi.org/10.1097/01.asn.0000034910.58454.fd.

Kannangara, D.R.W. *et al.* (2016) 'Hyperuricaemia: contributions of urate transporter ABCG2 and the fractional renal clearance of urate', *Annals of the Rheumatic Diseases*, 75(7), pp. 1363–1366. Available at: https://doi.org/10.1136/annrheumdis-2015-208111.

Katsanis, N. (2004) 'The oligogenic properties of Bardet-Biedl syndrome', *Human Molecular Genetics*, 13 Spec No 1, pp. R65-71. Available at: https://doi.org/10.1093/hmg/ddh092.

Kavuru, V. *et al.* (2020) 'Dipstick analysis of urine chemistry: benefits and limitations of dry chemistry-based assays', *Postgraduate Medicine*, 132(3), pp. 225–233. Available at: https://doi.org/10.1080/00325481.2019.1679540.

Kazancioğlu, R. (2013) 'Risk factors for chronic kidney disease: an update', *Kidney International Supplements*, 3(4), pp. 368–371. Available at: https://doi.org/10.1038/kisup.2013.79.

KDIGO Conference Participants (2022) 'Genetics in chronic kidney disease: conclusions from a Kidney Disease: Improving Global Outcomes (KDIGO) Controversies Conference', *Kidney*

International, 101(6), pp. 1126–1141. Available at: https://doi.org/10.1016/j.kint.2022.03.019.

Kemter, E. *et al.* (2013) 'Type of uromodulin mutation and allelic status influence onset and severity of uromodulin-associated kidney disease in mice', *Human Molecular Genetics*, 22(20), pp. 4148–4163. Available at: https://doi.org/10.1093/hmg/ddt263.

Kemter, E. *et al.* (2014) 'No amelioration of uromodulin maturation and trafficking defect by sodium 4-phenylbutyrate in vivo: studies in mouse models of uromodulin-associated kidney disease', *The Journal of Biological Chemistry*, 289(15), pp. 10715–10726. Available at: https://doi.org/10.1074/jbc.M113.537035.

Kemter, E. *et al.* (2017) 'Mitochondrial Dysregulation Secondary to Endoplasmic Reticulum Stress in Autosomal Dominant Tubulointerstitial Kidney Disease - UMOD (ADTKD-UMOD)', *Scientific Reports*, 7, p. 42970. Available at: https://doi.org/10.1038/srep42970.

Kerr, S.M. *et al.* (2013) 'Pedigree and genotyping quality analyses of over 10,000 DNA samples from the Generation Scotland: Scottish Family Health Study', *BMC medical genetics*, 14, p. 38. Available at: https://doi.org/10.1186/1471-2350-14-38.

Kerr, S.M. *et al.* (2024) 'Two founder variants account for over 90% of pathogenic BRCA alleles in the Orkney and Shetland Isles in Scotland', *European journal of human genetics: EJHG*, 32(12), pp. 1624–1631. Available at: https://doi.org/10.1038/s41431-024-01704-w.

Khan, A. *et al.* (2022) 'Genome-wide polygenic score to predict chronic kidney disease across ancestries', *Nature Medicine*, 28(7), pp. 1412–1420. Available at: https://doi.org/10.1038/s41591-022-01869-1.

Khan, A. *et al.* (2023) 'Polygenic risk alters the penetrance of monogenic kidney disease', *Nature Communications*, 14(1), p. 8318. Available at: https://doi.org/10.1038/s41467-023-43878-9.

Khera, A.V. *et al.* (2018) 'Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations', *Nature Genetics*, 50(9), pp. 1219–1224. Available at: https://doi.org/10.1038/s41588-018-0183-z.

Khramtsova, E.A. *et al.* (2023) 'Quality control and analytic best practices for testing genetic models of sex differences in large populations', *Cell*, 186(10), pp. 2044–2061. Available at: https://doi.org/10.1016/j.cell.2023.04.014.

Kidd, K. *et al.* (2020a) 'Genetic and Clinical Predictors of Age of ESKD in Individuals With Autosomal Dominant Tubulointerstitial Kidney Disease Due to UMOD Mutations', *Kidney International Reports*, 5(9), pp. 1472–1485. Available at: https://doi.org/10.1016/j.ekir.2020.06.029.

Kidd, K. *et al.* (2020b) 'Genetic and Clinical Predictors of Age of ESKD in Individuals With Autosomal Dominant Tubulointerstitial Kidney Disease Due to UMOD Mutations', *Kidney International Reports*, 5(9), pp. 1472–1485. Available at: https://doi.org/10.1016/j.ekir.2020.06.029.

'Kidney Precision Medicine Project' (2024). Available at: https://www.kpmp.org.

Kim, J.-W. *et al.* (2017) 'Prevalence and incidence of gout in Korea: data from the national health claims database 2007-2015', *Rheumatology International*, 37(9), pp. 1499–1506. Available at: https://doi.org/10.1007/s00296-017-3768-4.

Kim, Y. *et al.* (2023) 'MANF stimulates autophagy and restores mitochondrial homeostasis to treat autosomal dominant tubulointerstitial kidney disease in mice', *Nature Communications*, 14(1), p. 6493. Available at: https://doi.org/10.1038/s41467-023-42154-0.

Kingdom, R. and Wright, C.F. (2022) 'Incomplete Penetrance and Variable Expressivity: From Clinical Studies to Population Cohorts', *Frontiers in Genetics*, 13, p. 920390. Available at: https://doi.org/10.3389/fgene.2022.920390.

Kmochová, T. *et al.* (2024) 'Autosomal dominant ApoA4 mutations present as tubulointerstitial kidney disease with medullary amyloidosis', *Kidney International*, 105(4), pp. 799–811. Available at: https://doi.org/10.1016/j.kint.2023.11.021.

Köttgen, A. *et al.* (2009) 'Multiple loci associated with indices of renal function and chronic kidney disease', *Nature Genetics*, 41(6), pp. 712–717. Available at: https://doi.org/10.1038/ng.377.

Köttgen, A., Pattaro, C., *et al.* (2010) 'New loci associated with kidney function and chronic kidney disease', *Nature Genetics*, 42(5), pp. 376–384. Available at: https://doi.org/10.1038/ng.568.

Köttgen, A., Hwang, S.-J., *et al.* (2010) 'Uromodulin levels associate with a common UMOD variant and risk for incident CKD', *Journal of the American Society of Nephrology: JASN*, 21(2), pp. 337–344. Available at: https://doi.org/10.1681/ASN.2009070725.

Kovesdy, C.P. (2022) 'Epidemiology of chronic kidney disease: an update 2022', *Kidney International Supplements*, 12(1), pp. 7–11. Available at: https://doi.org/10.1016/j.kisu.2021.11.003.

Kuma, A. *et al.* (2021) 'Alteration of normal level of serum urate may contribute to decrease in estimated glomerular filtration rate decline in healthy Japanese men', *Renal Failure*, 43(1), pp. 1408–1415. Available at: https://doi.org/10.1080/0886022X.2021.1988969.

Kuwabara, M., Kanbay, M. and Hisatome, I. (2023) 'Tips and pitfalls in uric acid clinical research', *Hypertension Research: Official Journal of the Japanese Society of Hypertension*, 46(3), pp. 771–773. Available at: https://doi.org/10.1038/s41440-022-01148-z.

Labriola, L. *et al.* (2015) 'Paradoxical response to furosemide in uromodulin-associated kidney disease', *Nephrology, Dialysis, Transplantation: Official Publication of the European Dialysis and Transplant Association - European Renal Association*, 30(2), pp. 330–335. Available at: https://doi.org/10.1093/ndt/gfu389.

LaFavers, K.A. *et al.* (2019) 'Circulating uromodulin inhibits systemic oxidative stress by inactivating the TRPM2 channel', *Science Translational Medicine*, 11(512), p. eaaw3639. Available at: https://doi.org/10.1126/scitranslmed.aaw3639.

LaFavers, K.A., Micanovic, R., *et al.* (2022) 'Evolving Concepts in Uromodulin Biology, Physiology, and Its Role in Disease: a Tale of Two Forms', *Hypertension (Dallas, Tex.: 1979)*, 79(11), pp. 2409–2418. Available at: https://doi.org/10.1161/HYPERTENSIONAHA.122.18567.

LaFavers, K.A., Hage, C.A., *et al.* (2022) 'The kidney protects against sepsis by producing systemic uromodulin', *American Journal of Physiology. Renal Physiology*, 323(2), pp. F212–F226. Available at: https://doi.org/10.1152/ajprenal.00146.2022.

LaFavers, K.A. *et al.* (2023) 'Water Loading and Uromodulin Secretion in Healthy Individuals and Idiopathic Calcium Stone Formers', *Clinical journal of the American Society of Nephrology: CJASN*, 18(8), pp. 1059–1067. Available at: https://doi.org/10.2215/CJN.0000000000000202.

Landrum, M.J. *et al.* (2016) 'ClinVar: public archive of interpretations of clinically relevant variants', *Nucleic Acids Research*, 44(D1), pp. D862-868. Available at: https://doi.org/10.1093/nar/gkv1222.

Leask, M.P. *et al.* (2020) 'The Shared Genetic Basis of Hyperuricemia, Gout, and Kidney Function', *Seminars in Nephrology*, 40(6), pp. 586–599. Available at: https://doi.org/10.1016/j.semnephrol.2020.12.002.

Lee, T.H. *et al.* (2021) 'Hyperuricemia and Progression of Chronic Kidney Disease: A Review from Physiology and Pathogenesis to the Role of Urate-Lowering Therapy', *Diagnostics (Basel, Switzerland)*, 11(9), p. 1674. Available at: https://doi.org/10.3390/diagnostics11091674.

de Leeuw, C. *et al.* (2022) 'Understanding the assumptions underlying Mendelian randomization', *European journal of human genetics: EJHG*, 30(6), pp. 653–660. Available at: https://doi.org/10.1038/s41431-022-01038-5.

Leiherer, A. *et al.* (2018) 'The value of uromodulin as a new serum marker to predict decline in renal function', *Journal of Hypertension*, 36(1), pp. 110–118. Available at: https://doi.org/10.1097/HJH.0000000000001527.

Lemaire, M. *et al.* (2013) 'Recessive mutations in DGKE cause atypical hemolytic-uremic syndrome', *Nature Genetics*, 45(5), pp. 531–536. Available at: https://doi.org/10.1038/ng.2590.

Levin, A. *et al.* (2017) 'Global kidney health 2017 and beyond: a roadmap for closing gaps in care, research, and policy', *Lancet (London, England)*, 390(10105), pp. 1888–1917. Available at: https://doi.org/10.1016/S0140-6736(17)30788-2.

Li, K. *et al.* (2020) 'Interrogation of enhancer function by enhancer-targeting CRISPR epigenetic editing', *Nature Communications*, 11(1), p. 485. Available at: https://doi.org/10.1038/s41467-020-14362-5.

Li, Y. *et al.* (2022) 'Genome-wide studies reveal factors associated with circulating uromodulin and its relationships to complex diseases', *JCI insight*, 7(10), p. e157035. Available at: https://doi.org/10.1172/jci.insight.157035.

Lin, Y.-H. *et al.* (2013) 'Protoporphyrin IX accumulation disrupts mitochondrial dynamics and function in ABCG2-deficient hepatocytes', *FEBS letters*, 587(19), pp. 3202–3209. Available at: https://doi.org/10.1016/j.febslet.2013.08.011.

Liu, B. *et al.* (2010) 'A cross-national comparative study of orphan drug policies in the United States, the European Union, and Japan: towards a made-in-China orphan drug policy', *Journal of Public Health Policy*, 31(4), pp. 407–420; discussion 420-421. Available at: https://doi.org/10.1057/jphp.2010.30.

Liu, J. *et al.* (2023) 'Bartter syndrome type III with glomerular dysplasia and chronic kidney disease: A case report', *Frontiers in Pediatrics*, 11, p. 1169486. Available at: https://doi.org/10.3389/fped.2023.1169486.

Liu, J.Z. *et al.* (2015) 'Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations', *Nature Genetics*, 47(9), pp. 979–986. Available at: https://doi.org/10.1038/ng.3359.

Loh, P.-R. *et al.* (2016) 'Reference-based phasing using the Haplotype Reference Consortium panel', *Nature Genetics*, 48(11), pp. 1443–1448. Available at: https://doi.org/10.1038/ng.3679.

Lorenz, R. *et al.* (2020) 'Homoplasmy of the Mitochondrial DNA Mutation m.616T>C Leads to Mitochondrial Tubulointerstitial Kidney Disease and Encephalopathia', *Nephron*, 144(3), pp. 156–160. Available at: https://doi.org/10.1159/000504412.

Luyckx, V.A., Cherney, D.Z.I. and Bello, A.K. (2020) 'Preventing CKD in Developed Countries', *Kidney International Reports*, 5(3), pp. 263–277. Available at: https://doi.org/10.1016/j.ekir.2019.12.003.

Luyckx, V.A., Tonelli, M. and Stanifer, J.W. (2018) 'The global burden of kidney disease and the sustainable development goals', *Bulletin of the World Health Organization*, 96(6), pp. 414-422D. Available at: https://doi.org/10.2471/BLT.17.206441.

Ma, L. *et al.* (2017) 'Point mutation in D8C domain of Tamm-Horsfall protein/uromodulin in transgenic mice causes progressive renal damage and hyperuricemia', *PloS One*, 12(11), p. e0186769. Available at: https://doi.org/10.1371/journal.pone.0186769.

Mabillard, H., Olinger, E. and Sayer, J.A. (2022) 'UMOD and you! Explaining a rare disease diagnosis', *Journal of Rare Diseases (Berlin, Germany)*, 1(1), p. 4. Available at: https://doi.org/10.1007/s44162-022-00005-4.

Mabillard, H., Sayer, J.A. and Olinger, E. (2023) 'Clinical and genetic spectra of autosomal dominant tubulointerstitial kidney disease', *Nephrology, Dialysis, Transplantation: Official Publication of the European Dialysis and Transplant Association - European Renal Association*, 38(2), pp. 271–282. Available at: https://doi.org/10.1093/ndt/gfab268.

Major, T.J. *et al.* (2018) 'Evaluation of the diet wide contribution to serum urate levels: meta-analysis of population based cohorts', *BMJ (Clinical research ed.)*, 363, p. k3951. Available at: https://doi.org/10.1136/bmj.k3951.

Malhotra, D. *et al.* (2011) 'High frequencies of de novo CNVs in bipolar disorder and schizophrenia', *Neuron*, 72(6), pp. 951–963. Available at: https://doi.org/10.1016/j.neuron.2011.11.007.

Maroilley, T. and Tarailo-Graovac, M. (2019) 'Uncovering Missing Heritability in Rare Diseases', *Genes*, 10(4), p. 275. Available at: https://doi.org/10.3390/genes10040275.

Mary, S. *et al.* (2024) 'Pregnancy-associated changes in urinary uromodulin excretion in chronic hypertension', *Journal of Nephrology*, 37(3), pp. 597–610. Available at: https://doi.org/10.1007/s40620-023-01830-6.

Matsuo, H. *et al.* (2016) 'Genome-wide association study of clinically defined gout identifies multiple risk loci and its association with clinical subtypes', *Annals of the Rheumatic Diseases*, 75(4), pp. 652–659. Available at: https://doi.org/10.1136/annrheumdis-2014-206191.

Mavaddat, N. *et al.* (2019) 'Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes', *American Journal of Human Genetics*, 104(1), pp. 21–34. Available at: https://doi.org/10.1016/j.ajhg.2018.11.002.

McClellan, J. and King, M.-C. (2010) 'Genetic heterogeneity in human disease', *Cell*, 141(2), pp. 210–217. Available at: https://doi.org/10.1016/j.cell.2010.03.032.

McLaren, W. *et al.* (2016) 'The Ensembl Variant Effect Predictor', *Genome Biology*, 17(1), p. 122. Available at: https://doi.org/10.1186/s13059-016-0974-4.

Mencarelli, M.A. *et al.* (2015) 'Evidence of digenic inheritance in Alport syndrome', *Journal of Medical Genetics*, 52(3), pp. 163–174. Available at: https://doi.org/10.1136/jmedgenet-2014-102822.

Micanovic, R. *et al.* (2015) 'Tamm-Horsfall Protein Regulates Granulopoiesis and Systemic Neutrophil Homeostasis', *Journal of the American Society of Nephrology: JASN*, 26(9), pp. 2172–2182. Available at: https://doi.org/10.1681/ASN.2014070664.

Micanovic, R. *et al.* (2018) 'Tamm-Horsfall Protein Regulates Mononuclear Phagocytes in the Kidney', *Journal of the American Society of Nephrology: JASN*, 29(3), pp. 841–856. Available at: https://doi.org/10.1681/ASN.2017040409.

Micanovic, R. *et al.* (2020) 'Uromodulin (Tamm-Horsfall protein): guardian of urinary and systemic homeostasis', *Nephrology, Dialysis, Transplantation: Official Publication of the European Dialysis and Transplant Association - European Renal Association*, 35(1), pp. 33–43. Available at: https://doi.org/10.1093/ndt/gfy394.

Micanovic, R. *et al.* (2022) 'The kidney releases a nonpolymerizing form of uromodulin in the urine and circulation that retains the external hydrophobic patch domain', *American Journal of Physiology. Renal Physiology*, 322(4), pp. F403–F418. Available at: https://doi.org/10.1152/ajprenal.00322.2021.

Mikuls, T.R. *et al.* (2005) 'Gout epidemiology: results from the UK General Practice Research Database, 1990-1999', *Annals of the Rheumatic Diseases*, 64(2), pp. 267–272. Available at: https://doi.org/10.1136/ard.2004.024091.

Mo, L. *et al.* (2004) 'Ablation of the Tamm-Horsfall protein gene increases susceptibility of mice to bladder colonization by type 1-fimbriated Escherichia coli', *American Journal of Physiology. Renal Physiology*, 286(4), pp. F795-802. Available at: https://doi.org/10.1152/ajprenal.00357.2003.

Morris, A.P. *et al.* (2019) 'Trans-ethnic kidney function association study reveals putative causal genes and effects on kidney-specific disease aetiologies', *Nature Communications*, 10(1), p. 29. Available at: https://doi.org/10.1038/s41467-018-07867-7.

Moskowitz, J.L. *et al.* (2013) 'Association between genotype and phenotype in uromodulin-associated kidney disease', *Clinical journal of the American Society of Nephrology: CJASN*, 8(8), pp. 1349–1357. Available at: https://doi.org/10.2215/CJN.11151012.

Muir, E. (2016) 'The Rare Reality –an insight into the patient and family experience of rare disease.' Available at: https://www.raredisease.org.uk/media/1588/the-rare-reality-an-insight-into-the-patient-and-family-experience-of-rare-disease.pdf.

Musunuru, K. *et al.* (2010) 'From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus', *Nature*, 466(7307), pp. 714–719. Available at: https://doi.org/10.1038/nature09266.

Nadeau, J.H. (2001) 'Modifier genes in mice and humans', *Nature Reviews. Genetics*, 2(3), pp. 165–174. Available at: https://doi.org/10.1038/35056009.

Nanamatsu, A. *et al.* (2023) 'Healthy Women Have Higher Systemic Uromodulin Levels: Identification of Uromodulin as an Estrogen Responsive Gene', *Kidney360*, 4(9), pp. e1302–e1307. Available at: https://doi.org/10.34067/KID.0000000000000197.

Nanamatsu, A. *et al.* (2024) 'Advances in uromodulin biology and potential clinical applications', *Nature Reviews. Nephrology*, 20(12), pp. 806–821. Available at: https://doi.org/10.1038/s41581-024-00881-7.

Nanamatsu, A. *et al.* (2025) 'Alternative splicing of uromodulin enhances mitochondrial metabolism for adaptation to stress in kidney epithelial cells', *The Journal of Clinical Investigation*, 135(12), p. e183343. Available at: https://doi.org/10.1172/JCI183343.

Nassar, L.R. *et al.* (2023) 'The UCSC Genome Browser database: 2023 update', *Nucleic Acids Research*, 51(D1), pp. D1188–D1195. Available at: https://doi.org/10.1093/nar/gkac1072.

'National Registry of Rare Kidney Diseases (RaDaR)' (2022). Available at: Available from: https://renal.org/rare-renal/radar.

Nguengang Wakap, S. *et al.* (2020) 'Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database', *European journal of human genetics: EJHG*, 28(2), pp. 165–173. Available at: https://doi.org/10.1038/s41431-019-0508-0.

Nie, K. *et al.* (2020) 'COX6B2 drives metabolic reprogramming toward oxidative phosphorylation to promote metastasis in pancreatic ductal cancer cells', *Oncogenesis*, 9(5), p. 51. Available at: https://doi.org/10.1038/s41389-020-0231-2.

Nieto, F.J. *et al.* (2000) 'Uric acid and serum antioxidant capacity: a reaction to atherosclerosis?', *Atherosclerosis*, 148(1), pp. 131–139. Available at: https://doi.org/10.1016/s0021-9150(99)00214-2.

Nigam, S.K. (2018) 'The SLC22 Transporter Family: A Paradigm for the Impact of Drug Transporters on Metabolic Pathways, Signaling, and Disease', *Annual Review of Pharmacology and Toxicology*, 58, pp. 663–687. Available at: https://doi.org/10.1146/annurev-pharmtox-010617-052713.

Nozu, K. *et al.* (2020) 'Inherited salt-losing tubulopathy: An old condition but a new category of tubulopathy', *Pediatrics International: Official Journal of the Japan Pediatric Society*, 62(4), pp. 428–437. Available at: https://doi.org/10.1111/ped.14089.

Nun, A. *et al.* (2025) 'Real-life implementation and evaluation of the e-referral system SIPILINK', *International Journal of Medical Informatics*, 194, p. 105605. Available at: https://doi.org/10.1016/j.ijmedinf.2024.105605.

Nutt, S., Limb, L. (2010) 'Survey of patients' and families' experiences of rare diseases reinforces calls for a rare disease strategy', *Social Care and Neurodisability*, 2(4), pp. 195–199.

Ohashi, Y., Toyoda, M., *et al.* (2023) 'Evaluation of ABCG2-mediated extra-renal urate excretion in hemodialysis patients', *Scientific Reports*, 13(1), p. 93. Available at: https://doi.org/10.1038/s41598-022-26519-x.

Ohashi, Y., Kuriyama, S., *et al.* (2023) 'Urate Transporter ABCG2 Function and Asymptomatic Hyperuricemia: A Retrospective Cohort Study of CKD Progression', *American Journal of Kidney Diseases: The Official Journal of the National Kidney Foundation*, 81(2), pp. 134-144.e1. Available at: https://doi.org/10.1053/j.ajkd.2022.05.010.

Olden, M. *et al.* (2014) 'Common variants in UMOD associate with urinary uromodulin levels: a meta-analysis', *Journal of the American Society of Nephrology: JASN*, 25(8), pp. 1869–1882. Available at: https://doi.org/10.1681/ASN.2013070781.

Olinger, E. *et al.* (2020) 'Clinical and genetic spectra of autosomal dominant tubulointerstitial kidney disease due to mutations in UMOD and MUC1', *Kidney International*, 98(3), pp. 717–731. Available at: https://doi.org/10.1016/j.kint.2020.04.038.

Olinger, E. *et al.* (2022) 'An intermediate-effect size variant in UMOD confers risk for chronic kidney disease', *Proceedings of the National Academy of Sciences of the United States of America*, 119(33), p. e2114734119. Available at: https://doi.org/10.1073/pnas.2114734119.

Ong, E. *et al.* (2020) 'Modelling kidney disease using ontology: insights from the Kidney Precision Medicine Project', *Nature Reviews. Nephrology*, 16(11), pp. 686–696. Available at: https://doi.org/10.1038/s41581-020-00335-w.

Onoe, T. *et al.* (2021) 'Significance of kidney biopsy in autosomal dominant tubulointerstitial kidney disease-UMOD: is kidney biopsy truly nonspecific?', *BMC nephrology*, 22(1), p. 1. Available at: https://doi.org/10.1186/s12882-020-02169-x.

'Open Targets Genetics' (2024). Available at: https://genetics.opentargets.org.

Oulerich, Z. and Sferruzzi-Perri, A.N. (2024) 'Early-life exposures and long-term health: adverse gestational environments and the programming of offspring renal and vascular disease', *American Journal of Physiology. Renal Physiology*, 327(1), pp. F21–F36. Available at: https://doi.org/10.1152/ajprenal.00383.2023.

Owen, K.R. *et al.* (2010) 'Assessment of high-sensitivity C-reactive protein levels as diagnostic discriminator of maturity-onset diabetes of the young due to HNF1A mutations', *Diabetes Care*, 33(9), pp. 1919–1924. Available at: https://doi.org/10.2337/dc10-0288.

Padmanabhan, S. *et al.* (2010) 'Genome-wide association study of blood pressure extremes identifies variant near UMOD associated with hypertension', *PLoS genetics*, 6(10), p. e1001177. Available at: https://doi.org/10.1371/journal.pgen.1001177.

Pagniez, M.-S. *et al.* (2025) 'Exon location of glycine substitutions impacts kidney survival in autosomal dominant Alport Syndrome', *Nephrology, Dialysis, Transplantation: Official*

Publication of the European Dialysis and Transplant Association - European Renal Association, p. gfaf011. Available at: https://doi.org/10.1093/ndt/gfaf011.

Patabandige, M.W., Go, E.P. and Desaire, H. (2021) 'Clinically Viable Assay for Monitoring Uromodulin Glycosylation', *Journal of the American Society for Mass Spectrometry*, 32(2), pp. 436–443. Available at: https://doi.org/10.1021/jasms.0c00317.

Pattaro, C. *et al.* (2012) 'Genome-wide association and functional follow-up reveals new loci for kidney function', *PLoS genetics*, 8(3), p. e1002584. Available at: https://doi.org/10.1371/journal.pgen.1002584.

Pavlović, N. *et al.* (2025) 'Mitochondrial Dysfunction: The Silent Catalyst of Kidney Disease Progression', *Cells*, 14(11), p. 794. Available at: https://doi.org/10.3390/cells14110794.

Peloso, G.M. *et al.* (2014) 'Association of low-frequency and rare coding-sequence variants with blood lipids and coronary heart disease in 56,000 whites and blacks', *American Journal of Human Genetics*, 94(2), pp. 223–232. Available at: https://doi.org/10.1016/j.ajhg.2014.01.009.

Peltonen, L. *et al.* (2006) 'Lessons from studying monogenic disease for common disease', *Human Molecular Genetics*, 15 Spec No 1, pp. R67-74. Available at: https://doi.org/10.1093/hmg/ddl060.

Peltonen, L., Jalanko, A. and Varilo, T. (1999) 'Molecular genetics of the Finnish disease heritage', *Human Molecular Genetics*, 8(10), pp. 1913–1923. Available at: https://doi.org/10.1093/hmg/8.10.1913.

Persu, A. *et al.* (2004) 'Comparison between siblings and twins supports a role for modifier genes in ADPKD', *Kidney International*, 66(6), pp. 2132–2136. Available at: https://doi.org/10.1111/j.1523-1755.2004.66003.x.

Peters, D.J. and Breuning, M.H. (2001) 'Autosomal dominant polycystic kidney disease: modification of disease progression', *Lancet (London, England)*, 358(9291), pp. 1439–1444. Available at: https://doi.org/10.1016/S0140-6736(01)06531-X.

Pham, L. *et al.* (2024) 'Regulation of mitochondrial oxidative phosphorylation through tight control of cytochrome c oxidase in health and disease - Implications for

ischemia/reperfusion injury, inflammatory diseases, diabetes, and cancer', *Redox Biology*, 78, p. 103426. Available at: https://doi.org/10.1016/j.redox.2024.103426.

Piret, S.E. *et al.* (2017) 'A mouse model for inherited renal fibrosis associated with endoplasmic reticulum stress', *Disease Models & Mechanisms*, 10(6), pp. 773–786. Available at: https://doi.org/10.1242/dmm.029488.

Pivin, E. *et al.* (2018) 'Uromodulin and Nephron Mass', *Clinical journal of the American Society of Nephrology: CJASN*, 13(10), pp. 1556–1557. Available at: https://doi.org/10.2215/CJN.03600318.

Ponte, B. *et al.* (2021) 'Uromodulin, Salt, and 24-Hour Blood Pressure in the General Population', *Clinical journal of the American Society of Nephrology: CJASN*, 16(5), pp. 787–789. Available at: https://doi.org/10.2215/CJN.11230720.

Ponte, B. and Devuyst, O. (2020) 'Circulating Uromodulin and Risk of Cardiovascular Events and Kidney Failure', *Clinical journal of the American Society of Nephrology: CJASN*, 15(5), pp. 589–591. Available at: https://doi.org/10.2215/CJN.03580320.

Pook, M.A. *et al.* (1993) 'Localization of the Tamm-Horsfall glycoprotein (uromodulin) gene to chromosome 16p12.3-16p13.11', *Annals of Human Genetics*, 57(4), pp. 285–290. Available at: https://doi.org/10.1111/j.1469-1809.1993.tb00902.x.

Pruijm, M. *et al.* (2016) 'Associations of Urinary Uromodulin with Clinical Characteristics and Markers of Tubular Function in the General Population', *Clinical journal of the American Society of Nephrology: CJASN*, 11(1), pp. 70–80. Available at: https://doi.org/10.2215/CJN.04230415.

Pulit, S.L., de With, S.A.J. and de Bakker, P.I.W. (2017) 'Resetting the bar: Statistical significance in whole-genome sequencing-based association studies of global populations', *Genetic Epidemiology*, 41(2), pp. 145–151. Available at: https://doi.org/10.1002/gepi.22032.

Quiroz, Y.T. *et al.* (2024) 'APOE3 Christchurch Heterozygosity and Autosomal Dominant Alzheimer's Disease', *The New England Journal of Medicine*, 390(23), pp. 2156–2164. Available at: https://doi.org/10.1056/NEJMoa2308583.

R. G. Fassett (2011) ‘Biomarkers in chronic kidney disease: a review’, *Kidney International*, 80(8), pp. 806–821.

Raffi, H.S. *et al.* (2005) ‘Tamm-Horsfall protein acts as a general host-defense factor against bacterial cystitis’, *American Journal of Nephrology*, 25(6), pp. 570–578. Available at: https://doi.org/10.1159/000088990.

Raffi, H.S. *et al.* (2009) ‘Tamm-horsfall protein protects against urinary tract infection by proteus mirabilis’, *The Journal of Urology*, 181(5), pp. 2332–2338. Available at: https://doi.org/10.1016/j.juro.2009.01.014.

Rahit, K.M.T.H. and Tarailo-Graovac, M. (2020) ‘Genetic Modifiers and Rare Mendelian Disease’, *Genes*, 11(3), p. 239. Available at: https://doi.org/10.3390/genes11030239.

Reeders, S.T. *et al.* (1985) ‘A highly polymorphic DNA marker linked to adult polycystic kidney disease on chromosome 16’, *Nature*, 317(6037), pp. 542–544. Available at: https://doi.org/10.1038/317542a0.

Reiterová, J. and Tesař, V. (2022) ‘Autosomal Dominant Polycystic Kidney Disease: From Pathophysiology of Cystogenesis to Advances in the Treatment’, *International Journal of Molecular Sciences*, 23(6), p. 3317. Available at: https://doi.org/10.3390/ijms23063317.

Rezende-Lima, W. *et al.* (2004) ‘Homozygosity for uromodulin disorders: FJHN and MCKD-type 2’, *Kidney International*, 66(2), pp. 558–563. Available at: https://doi.org/10.1111/j.1523-1755.2004.00774.x.

Rheault, M.N. *et al.* (2010) ‘X-inactivation modifies disease severity in female carriers of murine X-linked Alport syndrome’, *Nephrology, Dialysis, Transplantation: Official Publication of the European Dialysis and Transplant Association - European Renal Association*, 25(3), pp. 764–769. Available at: https://doi.org/10.1093/ndt/gfp551.

Rhode, H. *et al.* (2023) ‘Urinary Protein-Biomarkers Reliably Indicate Very Early Kidney Damage in Children With Alport Syndrome Independently of Albuminuria and Inflammation’, *Kidney International Reports*, 8(12), pp. 2778–2793. Available at: https://doi.org/10.1016/j.ekir.2023.09.028.

Richards, S. *et al.* (2015) 'Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology', *Genetics in Medicine: Official Journal of the American College of Medical Genetics*, 17(5), pp. 405–424. Available at: https://doi.org/10.1038/gim.2015.30.

Risch, L. *et al.* (2014) 'The serum uromodulin level is associated with kidney function', *Clinical Chemistry and Laboratory Medicine*, 52(12), pp. 1755–1761. Available at: https://doi.org/10.1515/cclm-2014-0505.

Robey, R.W. *et al.* (2009) 'ABCG2: a perspective', *Advanced Drug Delivery Reviews*, 61(1), pp. 3–13. Available at: https://doi.org/10.1016/j.addr.2008.11.003.

Rothenbacher, D. *et al.* (2011) 'Frequency and risk factors of gout flares in a large population-based cohort of incident gout', *Rheumatology (Oxford, England)*, 50(5), pp. 973–981. Available at: https://doi.org/10.1093/rheumatology/keq363.

Säemann, M.D. *et al.* (2005) 'Tamm-Horsfall glycoprotein links innate immune cell activation with adaptive immunity via a Toll-like receptor-4-dependent mechanism', *The Journal of Clinical Investigation*, 115(2), pp. 468–475. Available at: https://doi.org/10.1172/JCI22720.

de Sainte Agathe, J.-M. *et al.* (2023) 'SpliceAI-visual: a free online tool to improve SpliceAI splicing variant interpretation', *Human Genomics*, 17(1), p. 7. Available at: https://doi.org/10.1186/s40246-023-00451-1.

Sampson, A.L., Singer, R.F. and Walters, G.D. (2017) 'Uric acid lowering therapies for preventing or delaying the progression of chronic kidney disease', *The Cochrane Database of Systematic Reviews*, 10(10), p. CD009460. Available at: https://doi.org/10.1002/14651858.CD009460.pub2.

Schaeffer, C. *et al.* (2009) 'Analysis of uromodulin polymerization provides new insights into the mechanisms regulating ZP domain-mediated protein assembly', *Molecular Biology of the Cell*, 20(2), pp. 589–599. Available at: https://doi.org/10.1091/mbc.e08-08-0876.

Schaeffer, C. *et al.* (2012) 'Urinary secretion and extracellular aggregation of mutant uromodulin isoforms', *Kidney International*, 81(8), pp. 769–778. Available at: https://doi.org/10.1038/ki.2011.456.

Schaeffer, C., Devuyst, O. and Rampoldi, L. (2021) 'Uromodulin: Roles in Health and Disease', *Annual Review of Physiology*, 83, pp. 477–501. Available at: https://doi.org/10.1146/annurev-physiol-031620-092817.

Schaeffer, C., et al. (2017) 'Mutant uromodulin expression leads to altered homeostasis of the endoplasmic reticulum and activates the unfolded protein response.', *PLoS ONE [Electronic Resource]*, 12(4), p. e0175970.

Scherberich, J.E. *et al.* (2018) 'Serum uromodulin-a marker of kidney function and renal parenchymal integrity', *Nephrology, Dialysis, Transplantation: Official Publication of the European Dialysis and Transplant Association - European Renal Association*, 33(2), pp. 284–295. Available at: https://doi.org/10.1093/ndt/gfw422.

Schiano, G. *et al.* (2023) 'Allelic effects on uromodulin aggregates drive autosomal dominant tubulointerstitial kidney disease', *EMBO molecular medicine*, 15(12), p. e18242. Available at: https://doi.org/10.15252/emmm.202318242.

Schieppati, A. *et al.* (2008) 'Why rare diseases are an important medical and social issue', *Lancet (London, England)*, 371(9629), pp. 2039–2041. Available at: https://doi.org/10.1016/S0140-6736(08)60872-7.

Schueler, M. *et al.* (2016) 'Large-scale targeted sequencing comparison highlights extreme genetic heterogeneity in nephronophthisis-related ciliopathies', *Journal of Medical Genetics*, 53(3), pp. 208–214. Available at: https://doi.org/10.1136/jmedgenet-2015-103304.

Scolari, F. *et al.* (2004) 'Uromodulin storage diseases: clinical aspects and mechanisms', *American Journal of Kidney Diseases: The Official Journal of the National Kidney Foundation*, 44(6), pp. 987–999. Available at: https://doi.org/10.1053/j.ajkd.2004.08.021.

Serafini-Cessi, F. *et al.* (1993) 'Biosynthesis and oligosaccharide processing of human Tamm-Horsfall glycoprotein permanently expressed in HeLa cells', *Biochemical and Biophysical Research Communications*, 194(2), pp. 784–790. Available at: https://doi.org/10.1006/bbrc.1993.1890.

Serafini-Cessi, F., Monti, A. and Cavallone, D. (2005) 'N-Glycans carried by Tamm-Horsfall glycoprotein have a crucial role in the defense against urinary tract diseases', *Glycoconjugate Journal*, 22(7–9), pp. 383–394. Available at: https://doi.org/10.1007/s10719-005-2142-z.

Sever, M.Ş. *et al.* (2021) 'A roadmap for optimizing chronic kidney disease patient care and patient-oriented research in the Eastern European nephrology community', *Clinical Kidney Journal*, 14(1), pp. 23–35. Available at: https://doi.org/10.1093/ckj/sfaa218.

Sherry, S.T. *et al.* (2001) 'dbSNP: the NCBI database of genetic variation', *Nucleic Acids Research*, 29(1), pp. 308–311. Available at: https://doi.org/10.1093/nar/29.1.308.

Sikri, K.L. *et al.* (1981) 'Localization of Tamm-Horsfall glycoprotein in the human kidney using immuno-fluorescence and immuno-electron microscopical techniques', *Journal of Anatomy*, 132(Pt 4), pp. 597–605.

Smith, G.D. *et al.* (2011a) 'Characterization of a recurrent in-frame UMOD indel mutation causing late-onset autosomal dominant end-stage renal failure', *Clinical journal of the American Society of Nephrology: CJASN*, 6(12), pp. 2766–2774. Available at: https://doi.org/10.2215/CJN.06820711.

Smith, G.D. *et al.* (2011b) 'Characterization of a recurrent in-frame UMOD indel mutation causing late-onset autosomal dominant end-stage renal failure', *Clinical journal of the American Society of Nephrology: CJASN*, 6(12), pp. 2766–2774. Available at: https://doi.org/10.2215/CJN.06820711.

Soliman, N.A. (2012) 'Orphan kidney diseases', *Nephron. Clinical Practice*, 120(4), pp. c194-199. Available at: https://doi.org/10.1159/000339785.

Srivastava, R. *et al.* (2014) 'An intricate network of conserved DNA upstream motifs and associated transcription factors regulate the expression of uromodulin gene', *The Journal of Urology*, 192(3), pp. 981–989. Available at: https://doi.org/10.1016/j.juro.2014.02.095.

Stanifer, J.W. *et al.* (2016) 'Chronic kidney disease in low- and middle-income countries', *Nephrology, Dialysis, Transplantation: Official Publication of the European Dialysis and Transplant Association - European Renal Association*, 31(6), pp. 868–874. Available at: https://doi.org/10.1093/ndt/gfv466.

Starr, M.C. and Hingorani, S.R. (2018) 'Prematurity and future kidney health: the growing risk of chronic kidney disease', *Current Opinion in Pediatrics*, 30(2), pp. 228–235. Available at: https://doi.org/10.1097/MOP.0000000000000607.

Steubl, D. *et al.* (2017) 'Serum uromodulin predicts graft failure in renal transplant recipients', *Biomarkers: Biochemical Indicators of Exposure, Response, and Susceptibility to Chemicals*, 22(2), pp. 171–177. Available at: https://doi.org/10.1080/1354750X.2016.1252957.

Stewart, A.P. *et al.* (2015) 'Pathogenic uromodulin mutations result in premature intracellular polymerization', *FEBS letters*, 589(1), pp. 89–93. Available at: https://doi.org/10.1016/j.febslet.2014.11.029.

Stiburkova, B. *et al.* (2019) 'The impact of dysfunctional variants of ABCG2 on hyperuricemia and gout in pediatric-onset patients', *Arthritis Research & Therapy*, 21(1), p. 77. Available at: https://doi.org/10.1186/s13075-019-1860-8.

Stiburkova, B. and Bleyer, A.J. (2012) 'Changes in serum urate and urate excretion with age', *Advances in Chronic Kidney Disease*, 19(6), pp. 372–376. Available at: https://doi.org/10.1053/j.ackd.2012.07.010.

Stiburkova, B. and Ichida, K. (2025) 'Genetic Background of Selected Hyperuricemia Causing Gout with Pediatric Onset', *Joint Bone Spine*, p. 105884. Available at: https://doi.org/10.1016/j.jbspin.2025.105884.

Stiburkova, B., Lukesova, M. and Zeman, J. (2025) 'Pediatrics hyperuricemia in clinical practice: A retrospective analysis in 1753 children and adolescents with hyperuricemia', *Joint Bone Spine*, 92(1), p. 105796. Available at: https://doi.org/10.1016/j.jbspin.2024.105796.

Stokman, M.F. *et al.* (2016) 'The expanding phenotypic spectra of kidney diseases: insights from genetic studies', *Nature Reviews. Nephrology*, 12(8), pp. 472–483. Available at: https://doi.org/10.1038/nrneph.2016.87.

Strauss-Kruger, M. *et al.* (2024) 'UMOD Genotype and Determinants of Urinary Uromodulin in African Populations', *Kidney International Reports*, 9(12), pp. 3477–3489. Available at: https://doi.org/10.1016/j.ekir.2024.09.015.

Stsiapanava, A. *et al.* (2020) 'Cryo-EM structure of native human uromodulin, a zona pellucida module polymer', *The EMBO journal*, 39(24), p. e106807. Available at: https://doi.org/10.15252/embj.2020106807.

Sud, A. *et al.* (2023) 'Realistic expectations are key to realising the benefits of polygenic scores', *BMJ (Clinical research ed.)*, 380, p. e073149. Available at: https://doi.org/10.1136/bmj-2022-073149.

Sul, J.H., Martin, L.S. and Eskin, E. (2018) 'Population structure in genetic studies: Confounding factors and mixed models', *PLoS genetics*, 14(12), p. e1007309. Available at: https://doi.org/10.1371/journal.pgen.1007309.

Sun, H.-L. *et al.* (2021) 'Function of Uric Acid Transporters and Their Inhibitors in Hyperuricaemia', *Frontiers in Pharmacology*, 12, p. 667753. Available at: https://doi.org/10.3389/fphar.2021.667753.

Tam, V. *et al.* (2019) 'Benefits and limitations of genome-wide association studies', *Nature Reviews. Genetics*, 20(8), pp. 467–484. Available at: https://doi.org/10.1038/s41576-019-0127-1.

Tanaka, T. (2005) '[International HapMap project]', *Nihon Rinsho. Japanese Journal of Clinical Medicine*, 63 Suppl 12, pp. 29–34.

Tanaka, Y. *et al.* (2025) 'Phenotype and genotype of autosomal dominant tubulointerstitial kidney disease in a Japanese cohort', *Clinical and Experimental Nephrology*, 29(6), pp. 788–796. Available at: https://doi.org/10.1007/s10157-025-02629-4.

Tanemoto, F., Nangaku, M. and Mimura, I. (2022) 'Epigenetic memory contributing to the pathogenesis of AKI-to-CKD transition', *Frontiers in Molecular Biosciences*, 9, p. 1003227. Available at: https://doi.org/10.3389/fmolb.2022.1003227.

Tang, S., Yuan, K. and Chen, L. (2022) 'Molecular biomarkers, network biomarkers, and dynamic network biomarkers for diagnosis and prediction of rare diseases', *Fundamental Research*, 2(6), pp. 894–902. Available at: https://doi.org/10.1016/j.fmre.2022.07.011.

Taruscio, D. and Gahl, W.A. (2024) 'Rare diseases: challenges and opportunities for research and public health', *Nature Reviews. Disease Primers*, 10(1), p. 13. Available at: https://doi.org/10.1038/s41572-024-00505-1.

Thielemans, R. *et al.* (2023) 'Unveiling the Hidden Power of Uromodulin: A Promising Potential Biomarker for Kidney Diseases', *Diagnostics (Basel, Switzerland)*, 13(19), p. 3077. Available at: https://doi.org/10.3390/diagnostics13193077.

Thorlacius, S. *et al.* (1996) 'A single BRCA2 mutation in male and female breast cancer families from Iceland with varied cancer phenotypes', *Nature Genetics*, 13(1), pp. 117–119. Available at: https://doi.org/10.1038/ng0596-117.

Thul, P.J. and Lindskog, C. (2018) 'The human protein atlas: A spatial map of the human proteome', *Protein Science: A Publication of the Protein Society*, 27(1), pp. 233–244. Available at: https://doi.org/10.1002/pro.3307.

Tokonami, N. *et al.* (2018) 'Uromodulin is expressed in the distal convoluted tubule, where it is critical for regulation of the sodium chloride cotransporter NCC', *Kidney International*, 94(4), pp. 701–715. Available at: https://doi.org/10.1016/j.kint.2018.04.021.

Tonelli, M. *et al.* (2006) 'Chronic kidney disease and mortality risk: a systematic review', *Journal of the American Society of Nephrology: JASN*, 17(7), pp. 2034–2047. Available at: https://doi.org/10.1681/ASN.2005101085.

Tonin, P.N. *et al.* (1998) 'Founder BRCA1 and BRCA2 mutations in French Canadian breast and ovarian cancer families', *American Journal of Human Genetics*, 63(5), pp. 1341–1351. Available at: https://doi.org/10.1086/302099.

Torffvit, O., Melander, O. and Hultén, U.L. (2004) 'Urinary excretion rate of Tamm-Horsfall protein is related to salt intake in humans', *Nephron. Physiology*, 97(1), pp. p31-36. Available at: https://doi.org/10.1159/000077600.

Torra, R. *et al.* (2004) 'Collagen type IV (alpha3-alpha4) nephropathy: from isolated haematuria to renal failure', *Nephrology, Dialysis, Transplantation: Official Publication of the European Dialysis and Transplant Association - European Renal Association*, 19(10), pp. 2429–2432. Available at: https://doi.org/10.1093/ndt/gfh435.

Torres, V.E. *et al.* (2012) 'Tolvaptan in patients with autosomal dominant polycystic kidney disease', *The New England Journal of Medicine*, 367(25), pp. 2407–2418. Available at: https://doi.org/10.1056/NEJMoa1205511.

Toyoda, Y. *et al.* (2021) 'Identification of Two Dysfunctional Variants in the ABCG2 Urate Transporter Associated with Pediatric-Onset of Familial Hyperuricemia and Early-Onset Gout', *International Journal of Molecular Sciences*, 22(4), p. 1935. Available at: https://doi.org/10.3390/ijms22041935.

Troyanov, S. *et al.* (2016) 'Clinical, Genetic, and Urinary Factors Associated with Uromodulin Excretion', *Clinical journal of the American Society of Nephrology: CJASN*, 11(1), pp. 62–69. Available at: https://doi.org/10.2215/CJN.04770415.

Trudu, M. *et al.* (2013) 'Common noncoding UMOD gene variants induce salt-sensitive hypertension and kidney damage by increasing uromodulin expression', *Nature Medicine*, 19(12), pp. 1655–1660. Available at: https://doi.org/10.1038/nm.3384.

Trudu, M. *et al.* (2017) 'Early involvement of cellular stress and inflammatory signals in the pathogenesis of tubulointerstitial kidney disease due to UMOD mutations', *Scientific Reports*, 7(1), p. 7383. Available at: https://doi.org/10.1038/s41598-017-07804-6.

Tseng, C.-C. *et al.* (2018) 'Next-generation sequencing profiling of mitochondrial genomes in gout', *Arthritis Research & Therapy*, 20(1), p. 137. Available at: https://doi.org/10.1186/s13075-018-1637-5.

Tseng, M.-H. *et al.* (2012) 'Genotype, phenotype, and follow-up in Taiwanese patients with salt-losing tubulopathy associated with SLC12A3 mutation', *The Journal of Clinical Endocrinology and Metabolism*, 97(8), pp. E1478-1482. Available at: https://doi.org/10.1210/jc.2012-1707.

Turner, S. *et al.* (2011) 'Quality control procedures for genome-wide association studies', *Current Protocols in Human Genetics*, Chapter 1, p. Unit1.19. Available at: https://doi.org/10.1002/0471142905.hg0119s68.

Turro, E. *et al.* (2020) 'Whole-genome sequencing of patients with rare diseases in a national health system', *Nature*, 583(7814), pp. 96–102. Available at: https://doi.org/10.1038/s41586-020-2434-2.

'UCSC Genome Browser' (2024). Available at: https://genome.ucsc.edu.

'UK Biobank' (2022). Available at: https://www.ukbiobank.ac.uk/.

Ung, T.H. *et al.* (2014) 'Exosome proteomics reveals transcriptional regulator proteins with potential to mediate downstream pathways', *Cancer Science*, 105(11), pp. 1384–1392. Available at: https://doi.org/10.1111/cas.12534.

Valluru, M.K. *et al.* (2023) 'A founder UMOD variant is a common cause of hereditary nephropathy in the British population', *Journal of Medical Genetics*, 60(4), pp. 397–405. Available at: https://doi.org/10.1136/jmg-2022-108704.

Vanholder, R. *et al.* (2017) 'Reducing the costs of chronic kidney disease while delivering quality health care: a call to action', *Nature Reviews. Nephrology*, 13(7), pp. 393–409. Available at: https://doi.org/10.1038/nrneph.2017.63.

Visscher, P.M. *et al.* (2012) 'Five years of GWAS discovery', *American Journal of Human Genetics*, 90(1), pp. 7–24. Available at: https://doi.org/10.1016/j.ajhg.2011.11.029.

Visscher, P.M. *et al.* (2017) '10 Years of GWAS Discovery: Biology, Function, and Translation', *American Journal of Human Genetics*, 101(1), pp. 5–22. Available at: https://doi.org/10.1016/j.ajhg.2017.06.005.

Vylet'al, P. *et al.* (2006) 'Alterations of uromodulin biology: a common denominator of the genetically heterogeneous FJHN/MCKD syndrome', *Kidney International*, 70(6), pp. 1155–1169. Available at: https://doi.org/10.1038/sj.ki.5001728.

Vyletal, P., Bleyer, A.J. and Kmoch, S. (2010) 'Uromodulin biology and pathophysiology--an update', *Kidney & Blood Pressure Research*, 33(6), pp. 456–475. Available at: https://doi.org/10.1159/000321013.

Wang, H., Aragam, B. and Xing, E.P. (2022) 'Trade-offs of Linear Mixed Models in Genome-Wide Association Studies', *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology*, 29(3), pp. 233–242. Available at: https://doi.org/10.1089/cmb.2021.0157.

Wang, S., Kang, Y. and Xie, H. (2024) 'PKD2: An Important Membrane Protein in Organ Development', *Cells*, 13(20), p. 1722. Available at: https://doi.org/10.3390/cells13201722.

Watanabe, K. *et al.* (2017) 'Functional mapping and annotation of genetic associations with FUMA', *Nature Communications*, 8(1), p. 1826. Available at: https://doi.org/10.1038/s41467-017-01261-5.

Webster, A.C. *et al.* (2017) 'Chronic Kidney Disease', *Lancet (London, England)*, 389(10075), pp. 1238–1252. Available at: https://doi.org/10.1016/S0140-6736(16)32064-5.

Wei, W.-H., Hemani, G. and Haley, C.S. (2014) 'Detecting epistasis in human complex traits', *Nature Reviews. Genetics*, 15(11), pp. 722–733. Available at: https://doi.org/10.1038/nrg3747.

Weiss, G.L. *et al.* (2020) 'Architecture and function of human uromodulin filaments in urinary tract infections', *Science (New York, N.Y.)*, 369(6506), pp. 1005–1010. Available at: https://doi.org/10.1126/science.aaz9866.

Weiss, R.B. *et al.* (2018) 'Long-range genomic regulators of THBS1 and LTBP4 modify disease severity in duchenne muscular dystrophy', *Annals of Neurology*, 84(2), pp. 234–245. Available at: https://doi.org/10.1002/ana.25283.

West, A.C. and Johnstone, R.W. (2014) 'New and emerging HDAC inhibitors for cancer treatment', *The Journal of Clinical Investigation*, 124(1), pp. 30–39. Available at: https://doi.org/10.1172/JCI69738.

Wolf, M.T.F. *et al.* (2009) 'Mutation analysis of the Uromodulin gene in 96 individuals with urinary tract anomalies (CAKUT)', *Pediatric Nephrology (Berlin, Germany)*, 24(1), pp. 55–60. Available at: https://doi.org/10.1007/s00467-008-1016-6.

Wong, K. *et al.* (2024) 'Effects of rare kidney diseases on kidney failure: a longitudinal analysis of the UK National Registry of Rare Kidney Diseases (RaDaR) cohort', *Lancet (London, England)*, 403(10433), pp. 1279–1289. Available at: https://doi.org/10.1016/S0140-6736(23)02843-X.

Wu, A. *et al.* (2022) 'Using human urinary extracellular vesicles to study physiological and pathophysiological states and regulation of the sodium chloride cotransporter', *Frontiers in Endocrinology*, 13, p. 981317. Available at: https://doi.org/10.3389/fendo.2022.981317.

Wu, H. *et al.* (2013) 'Crystal structures of the human histone H4K20 methyltransferases SUV420H1 and SUV420H2', *FEBS letters*, 587(23), pp. 3859–3868. Available at: https://doi.org/10.1016/j.febslet.2013.10.020.

Wuttke, M. *et al.* (2019) 'A catalog of genetic loci associated with kidney function from analyses of a million individuals', *Nature Genetics*, 51(6), pp. 957–972. Available at: https://doi.org/10.1038/s41588-019-0407-x.

Yang, J. *et al.* (2012) 'Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits', *Nature Genetics*, 44(4), pp. 369–375, S1-3. Available at: https://doi.org/10.1038/ng.2213.

Yang, Q. *et al.* (2005) 'Genome-wide search for genes affecting serum uric acid levels: the Framingham Heart Study', *Metabolism: Clinical and Experimental*, 54(11), pp. 1435–1441. Available at: https://doi.org/10.1016/j.metabol.2005.05.007.

Ye, B.S. *et al.* (2016) 'Does serum uric acid act as a modulator of cerebrospinal fluid Alzheimer's disease biomarker related cognitive decline?', *European Journal of Neurology*, 23(5), pp. 948–957. Available at: https://doi.org/10.1111/ene.12969.

Ying, W.Z. and Sanders, P.W. (1998) 'Dietary salt regulates expression of Tamm-Horsfall glycoprotein in rats', *Kidney International*, 54(4), pp. 1150–1156. Available at: https://doi.org/10.1046/j.1523-1755.1998.00117.x.

Youhanna, S. *et al.* (2014) 'Determination of uromodulin in human urine: influence of storage and processing', *Nephrology, Dialysis, Transplantation: Official Publication of the European Dialysis and Transplant Association - European Renal Association*, 29(1), pp. 136–145. Available at: https://doi.org/10.1093/ndt/gft345.

Zhou, H. *et al.* (2008) 'Urinary exosomal transcription factors, a new class of biomarkers for renal disease', *Kidney International*, 74(5), pp. 613–621. Available at: https://doi.org/10.1038/ki.2008.206.

Zhou, X. *et al.* (2016) 'Exploring genomic alteration in pediatric cancer using ProteinPaint', *Nature Genetics*, 48(1), pp. 4–6. Available at: https://doi.org/10.1038/ng.3466.

Zhu, B. *et al.* (2019) 'A novel CLCNKB mutation in a Chinese girl with classic Bartter syndrome: a case report', *BMC medical genetics*, 20(1), p. 137. Available at: https://doi.org/10.1186/s12881-019-0869-9.

Zhu, Y., Pandya, B.J. and Choi, H.K. (2011) 'Prevalence of gout and hyperuricemia in the US general population: the National Health and Nutrition Examination Survey 2007-2008', *Arthritis and Rheumatism*, 63(10), pp. 3136–3141. Available at: https://doi.org/10.1002/art.30520.

Zivná, M. *et al.* (2009) 'Dominant renin gene mutations associated with early-onset hyperuricemia, anemia, and chronic kidney failure', *American Journal of Human Genetics*, 85(2), pp. 204–213. Available at: https://doi.org/10.1016/j.ajhg.2009.07.010.

Živná, M. *et al.* (2020) 'An international cohort study of autosomal dominant tubulointerstitial kidney disease due to REN mutations identifies distinct clinical subtypes', *Kidney International*, 98(6), pp. 1589–1604. Available at: https://doi.org/10.1016/j.kint.2020.06.041.

Zoccali, C. *et al.* (2017) 'The systemic nature of CKD', *Nature Reviews. Nephrology*, 13(6), pp. 344–358. Available at: https://doi.org/10.1038/nrneph.2017.52.

Appendices

Appendix A: Most Significant SNPs from Genome-wide Association Studies

FULL 266 PERSON ADTKD-*UMOD* GWAS (Adjusted Significance Order)											
	Marker ID	**MAF**	**P-value**	**Chi-sq**	**Adj chi-sq**	**Pval_ adjchisq**	**CHR**	**cM**	**BP Position**	**Allele 1**	**Allele 2**
1	**rs11651009**	0.16917293	4.807934e-06	20.91232	22.35261	2.269043e-06	17	33.21048	11620993	A	G
2	**rs3935743**	0.08270677	5.044726e-06	20.82023	22.25418	2.388378e-06	16	105.9021	81909980	T	G
3	**rs1462758**	0.29511278	6.942719e-06	20.20895	21.60081	3.357104e-06	4	103.441	98200637	A	G
4	**rs11054458**	0.04887218	7.597481e-06	20.03658	21.41656	3.695651e-06	12	26.63298	11947535	C	T
5	**rs1450186**	0.02462121	9.038353e-06	19.70460	21.06172	4.447250e-06	7	58.9194	37722873	C	T
6	**rs61814582**	0.02631579	9.950082e-06	19.52098	20.86545	4.927017e-06	1	182.4422	181780959	A	G
7	**rs2622627**	0.43233083	1.588037e-05	18.62880	19.91182	8.109745e-06	4	95.82188	89065353	C	A
8	**rs2622626**	0.43233083	1.588037e-05	18.62880	19.91182	8.109745e-06	4	95.82292	89066715	G	T
9	**rs66548062**	0.06766917	1.591704e-05	18.62440	19.90712	8.129707e-06	15	108.7928	94767385	T	C
10	**rs73148239**	0.02819549	1.709797e-05	18.48797	19.76130	8.774144e-06	12	93.06915	79407375	G	A
11	**rs3114020**	0.47556391	1.968699e-05	18.21934	19.47416	1.019699e-05	4	95.83585	89083666	T	C
12	**rs56259150**	0.04580153	2.259698e-05	17.95685	19.19359	1.181091e-05	11	1.529965	2047489	T	C

13	**rs72785963**	0.02067669	2.867553e-05	17.50366	18.70919	1.522470e-05	5	127.9061	123525292	G	T
14	**rs61909183**	0.03383459	5.073205e-05	16.42056	17.55149	2.796314e-05	11	143.7847	130818354	T	C
15	**rs17731799**	0.47718631	5.242086e-05	16.35849	17.48515	2.895609e-05	4	95.82425	89068455	G	T
IMPUTED FULL 266 PERSON ADTKD-*UMOD* GWAS (Adjusted Significance Order)											
1	**rs143840619**	0.03007519	2.881995e-07	26.32721	28.28452	1.047302e-07	1	-	90535832	A	G
2	**rs80352850**	0.03007519	5.173444e-07	25.19805	27.07141	1.960762e-07	1	-	90537444	C	A
3	**rs8074402**	0.21992481	5.395491e-07	25.11701	26.98434	2.051102e-07	17	-	11626047	G	T
4	**rs5761621**	0.32894737	8.531873e-07	24.23387	26.03555	3.351888e-07	22	-	26999358	G	A
5	**rs28562846**	0.01879699	1.349610e-06	23.35118	25.08723	5.479454e-07	18	-	29778049	C	G
6	**rs16962801**	0.01879699	1.349610e-06	23.35118	25.08723	5.479454e-07	18	-	29778090	T	G
7	**rs62060825**	0.15037594	1.988855e-06	22.60578	24.28642	8.302247e-07	17	-	11562294	A	G
8	**rs12459907**	0.45676692	2.006895e-06	22.58843	24.26778	8.382970e-07	19	-	55856851	A	G
9	**rs499779**	0.38721805	2.029414e-06	22.56699	24.24475	8.483812e-07	19	-	58475451	C	T
10	**rs1307415325**	0.05451128	2.137521e-06	22.46730	24.13764	8.969007e-07	X	-	51018388	G	A
11	**rs10058279**	0.19924812	2.878469e-06	21.89591	23.52377	1.233799e-06	5	-	166864788	G	C
12	**rs7001584**	0.34774436	3.474839e-06	21.53470	23.13571	1.509611e-06	8	-	27548165	T	A

13	**rs6558012**	0.34774436	3.474839e-06	21.53470	23.13571	1.509611e-06	8	-	27548456	T	A
14	**rs4545046**	0.34774436	3.474839e-06	21.53470	23.13571	1.509611e-06	8	-	27556526	A	C
15	**rs72869201**	0.19360902	3.728948e-06	21.39936	22.99031	1.628200e-06	11	-	13405015	C	T
16	**rs72870911**	0.18609023	4.110538e-06	21.21260	22.78966	1.807356e-06	11	-	13408922	A	G
17	**rs17452383**	0.18609023	4.110538e-06	21.21260	22.78966	1.807356e-06	11	-	13410371	G	A
18	**rs56384052**	0.18609023	4.110538e-06	21.21260	22.78966	1.807356e-06	11	-	13414046	G	A
19	**rs142074501**	0.18609023	4.110538e-06	21.21260	22.78966	1.807356e-06	11	-	13422059	T	C
20	**rs3761862**	0.18609023	4.110538e-06	21.21260	22.78966	1.807356e-06	11	-	13424506	G	C
21	**rs72870943**	0.18609023	4.110538e-06	21.21260	22.78966	1.807356e-06	11	-	13430092	C	T
22	**rs72870946**	0.18609023	4.110538e-06	21.21260	22.78966	1.807356e-06	11	-	13431441	C	G
23	**rs143305491**	0.18609023	4.110538e-06	21.21260	22.78966	1.807356e-06	11	-	13433612	G	A
24	**rs149617495**	0.18609023	4.110538e-06	21.21260	22.78966	1.807356e-06	11	-	13437761	T	C
25	**rs145159588**	0.18609023	4.110538e-06	21.21260	22.78966	1.807356e-06	11	-	13448711	T	C
26	**rs72857026**	0.18609023	4.110538e-06	21.21260	22.78966	1.807356e-06	11	-	13450861	T	C
27	**rs72857027**	0.18609023	4.110538e-06	21.21260	22.78966	1.807356e-06	11	-	13451999	C	T
28	**rs72857028**	0.18609023	4.110538e-06	21.21260	22.78966	1.807356e-06	11	-	13454137	T	G

29	**rs72857030**	0.18609023	4.110538e-06	21.21260	22.78966	1.807356e-06	11	-	13455010	A	T
30	**rs72857043**	0.18609023	4.110538e-06	21.21260	22.78966	1.807356e-06	11	-	13461457	A	C
31	**rs72857050**	0.18609023	4.110538e-06	21.21260	22.78966	1.807356e-06	11	-	13465573	T	C
32	**rs1077956**	0.18609023	4.110538e-06	21.21260	22.78966	1.807356e-06	11	-	13469165	T	C
33	**rs72857069**	0.18609023	4.110538e-06	21.21260	22.78966	1.807356e-06	11	-	13474291	T	A
34	**rs72857071**	0.18609023	4.110538e-06	21.21260	22.78966	1.807356e-06	11	-	13474711	C	T
35	**rs74543117**	0.18609023	4.110538e-06	21.21260	22.78966	1.807356e-06	11	-	13475595	C	T
36	**rs2061438**	0.18609023	4.110538e-06	21.21260	22.78966	1.807356e-06	11	-	13476665	C	G
37	**rs4356206**	0.18609023	4.110538e-06	21.21260	22.78966	1.807356e-06	11	-	13481096	A	G
38	**rs72857077**	0.18609023	4.110538e-06	21.21260	22.78966	1.807356e-06	11	-	13482744	C	A
39	**rs6605364**	0.33082707	4.374010e-06	21.09354	22.66175	1.931757e-06	4	-	3925152	A	G
40	**rs146302733**	0.01503759	4.711703e-06	20.95105	22.50866	2.091981e-06	4	-	21998438	T	C
41	**rs11650039**	0.16917293	4.807934e-06	20.91232	22.46705	2.137792e-06	17	-	11602011	C	A
42	**rs62061778**	0.16917293	4.807934e-06	20.91232	22.46705	2.137792e-06	17	-	11602552	T	C
43	**rs72810873**	0.16917293	4.807934e-06	20.91232	22.46705	2.137792e-06	17	-	11610875	G	A
44	**rs8079335**	0.16917293	4.807934e-06	20.91232	22.46705	2.137792e-06	17	-	11611213	C	T

45	**rs11655363**	0.16917293	4.807934e-06	20.91232	22.46705	2.137792e-06	17	-	11611795	G	C
46	**rs11651009**	0.16917293	4.807934e-06	20.91232	22.46705	2.137792e-06	17	-	11620993	A	G
47	**rs17602516**	0.16917293	4.807934e-06	20.91232	22.46705	2.137792e-06	17	-	11622978	G	T
48	**rs28456990**	0.45300752	5.016280e-06	20.83106	22.37976	2.237197e-06	3	-	175103020	A	T
49	**rs114270590**	0.01503759	5.037222e-06	20.82308	22.37118	2.247205e-06	2	-	30529903	C	A
50	**rs3935743**	0.08270677	5.044726e-06	20.82023	22.36812	2.250792e-06	16	-	81909980	G	T
51	**rs113059329**	0.03947368	5.070612e-06	20.81043	22.35759	2.263169e-06	18	-	19563074	C	T
52	**rs7437393**	0.31203008	5.536869e-06	20.64198	22.17662	2.486843e-06	4	-	3928134	T	C
53	**rs13257161**	0.41541353	5.688764e-06	20.59016	22.12095	2.560006e-06	8	-	27542439	C	T
54	**rs79012440**	0.39661654	5.762232e-06	20.56560	22.09456	2.595444e-06	19	-	55849316	A	G
55	**rs1870073**	0.39661654	5.762232e-06	20.56560	22.09456	2.595444e-06	19	-	55850269	A	G
56	**rs1544929**	0.37593985	5.899772e-06	20.52044	22.04605	2.661873e-06	19	-	58436041	G	A
57	**rs34647569**	0.37593985	5.899772e-06	20.52044	22.04605	2.661873e-06	19	-	58442162	T	C
58	**rs257676**	0.36654135	6.886606e-06	20.22448	21.72808	3.141594e-06	19	-	58485439	T	C
59	**rs2865786**	0.29511278	6.942719e-06	20.20895	21.71140	3.169026e-06	4	-	98199403	T	C
60	**rs1350974**	0.29511278	6.942719e-06	20.20895	21.71140	3.169026e-06	4	-	98199626	T	G

61	**rs1350973**	0.29511278	6.942719e-06	20.20895	21.71140	3.169026e-06	4	-	98199636	A	C
62	**rs1462759**	0.29511278	6.942719e-06	20.20895	21.71140	3.169026e-06	4	-	98200136	C	T
63	**rs1462758**	0.29511278	6.942719e-06	20.20895	21.71140	3.169026e-06	4	-	98200637	T	C
64	**rs7688316**	0.29511278	6.942719e-06	20.20895	21.71140	3.169026e-06	4	-	98201537	C	T
65	**rs1599660**	0.29511278	6.942719e-06	20.20895	21.71140	3.169026e-06	4	-	98202008	A	T
66	**rs73066266**	0.36654135	6.997989e-06	20.19379	21.69510	3.196061e-06	19	-	58496101	T	C
67	**rs34243962**	0.05827068	7.144058e-06	20.15427	21.65265	3.267584e-06	20	-	39392361	C	T
68	**rs7255565**	0.37781955	7.501492e-06	20.06090	21.55233	3.443036e-06	19	-	58422281	C	T
69	**rs7253038**	0.37781955	7.501492e-06	20.06090	21.55233	3.443036e-06	19	-	58425118	G	T
70	**rs12972704**	0.37781955	7.501492e-06	20.06090	21.55233	3.443036e-06	19	-	58426883	G	T
71	**rs7254431**	0.37781955	7.501492e-06	20.06090	21.55233	3.443036e-06	19	-	58429340	T	C
72	**rs11054458**	0.04887218	7.597481e-06	20.03658	21.52621	3.490256e-06	12	-	11947535	C	T
73	**rs11054460**	0.04887218	7.597481e-06	20.03658	21.52621	3.490256e-06	12	-	11954171	C	T
74	**rs115613990**	0.19172932	7.656779e-06	20.02171	21.51024	3.519448e-06	11	-	13394335	A	G
75	**rs72869173**	0.19172932	7.656779e-06	20.02171	21.51024	3.519448e-06	11	-	13396748	A	G
76	**rs58295558**	0.19172932	7.656779e-06	20.02171	21.51024	3.519448e-06	11	-	13400278	G	C

77	**rs114080582**	0.02443609	7.817501e-06	19.98199	21.46756	3.598651e-06	18	-	44315867	C	T
78	**rs4894486**	0.45676692	7.941622e-06	19.95187	21.43520	3.659896e-06	3	-	175100948	A	G
79	**rs9883274**	0.45676692	7.941622e-06	19.95187	21.43520	3.659896e-06	3	-	175101948	T	C
92 PERSON ADTKD-*UMOD* p.(Val93_Gly97delinsAlaAlaSerCys) variant GWAS (Adjusted Significance Order)											
1	**rs77129692**	0.04347826	9.878013e-07	23.95176	27.37919	1.672213e-07	13	50.65763	47485064	T	C
2	**rs73175544**	0.04347826	9.878013e-07	23.95176	27.37919	1.672213e-07	13	50.65798	47485407	C	T
3	**rs12845211**	0.42391304	1.474425e-06	23.18106	26.49820	2.637831e-07	23	63.69046	40889466	G	A
4	**rs62224727**	0.04891304	1.993384e-06	22.60141	25.83561	3.717662e-07	20	90.15533	54697933	A	G
5	**rs7054156**	0.42391304	2.187610e-06	22.42281	25.63145	4.132495e-07	23	63.71739	40903663	G	A
6	**rs1370682**	0.27173913	4.914788e-06	20.87021	23.85668	1.037811e-06	4	160.4881	166296379	A	G
7	**rs9564536**	0.29891304	5.202245e-06	20.76135	23.73224	1.107124e-06	13	33.28389	35258606	A	C
IMPUTED 92 PERSON ADTKD-*UMOD* p.(Val93_Gly97delinsAlaAlaSerCys) variant GWAS (Adjusted Significance Order)											
1	**rs10152175**	0.10869565	6.595647e-08	29.18000	33.49543	7.143163e-09	15	-	97350625	T	C
2	**rs73175542**	0.04347826	9.878013e-07	23.95176	27.49398	1.575842e-07	13	-	47478865	A	G
3	**rs73175544**	0.04347826	9.878013e-07	23.95176	27.49398	1.575842e-07	13	-	47485407	C	T
4	**rs73175546**	0.04347826	9.878013e-07	23.95176	27.49398	1.575842e-07	13	-	47487779	T	C
5	**rs115416253**	0.04347826	9.878013e-07	23.95176	27.49398	1.575842e-07	13	-	47491911	A	T

6	**rs55812936**	0.04347826	9.878013e-07	23.95176	27.49398	1.575842e-07	13	-	47495063	T	C
7	**rs73175549**	0.04347826	9.878013e-07	23.95176	27.49398	1.575842e-07	13	-	47496679	A	G
8	**rs112452802**	0.04347826	9.878013e-07	23.95176	27.49398	1.575842e-07	13	-	47502783	T	C
9	**rs12845211**	0.42391304	1.474425e-06	23.18106	26.60930	2.490416e-07	X	-	40889466	G	A
10	**rs62224727**	0.04891304	1.993384e-06	22.60141	25.94393	3.514790e-07	20	-	54697933	A	G
11	**rs118059905**	0.04891304	1.993384e-06	22.60141	25.94393	3.514790e-07	20	-	54712293	T	A
12	**rs4827251**	0.42391304	2.187610e-06	22.42281	25.73891	3.908661e-07	X	-	40900317	G	A
13	**rs5918100**	0.42391304	2.187610e-06	22.42281	25.73891	3.908661e-07	X	-	40900922	T	C
14	**rs4827252**	0.42391304	2.187610e-06	22.42281	25.73891	3.908661e-07	X	-	40902022	G	A
15	**rs7054156**	0.42391304	2.187610e-06	22.42281	25.73891	3.908661e-07	X	-	40903663	G	A
16	**rs17322309**	0.42391304	2.187610e-06	22.42281	25.73891	3.908661e-07	X	-	40904595	G	A
17	**rs5918103**	0.42391304	2.187610e-06	22.42281	25.73891	3.908661e-07	X	-	40908740	G	A
18	**rs5917405**	0.42391304	2.187610e-06	22.42281	25.73891	3.908661e-07	X	-	40908885	T	C
19	**rs36042689**	0.42391304	2.187610e-06	22.42281	25.73891	3.908661e-07	X	-	40909349	C	T
20	**rs12689198**	0.42391304	2.187610e-06	22.42281	25.73891	3.908661e-07	X	-	40909637	C	T
21	**rs1263919**	0.42391304	2.187610e-06	22.42281	25.73891	3.908661e-07	X	-	40911207	C	T

22	**rs1263917**	0.42391304	2.187610e-06	22.42281	25.73891	3.908661e-07	X	-	40913393	T	C
23	**rs1263915**	0.42391304	2.187610e-06	22.42281	25.73891	3.908661e-07	X	-	40914444	G	A
24	**rs1263912**	0.42391304	2.187610e-06	22.42281	25.73891	3.908661e-07	X	-	40915698	A	G
25	**rs1150526**	0.42391304	2.187610e-06	22.42281	25.73891	3.908661e-07	X	-	40918086	A	G
26	**rs1150520**	0.42391304	2.187610e-06	22.42281	25.73891	3.908661e-07	X	-	40919699	C	A
27	**rs1150533**	0.42391304	2.187610e-06	22.42281	25.73891	3.908661e-07	X	-	40920165	G	A
28	**rs1150513**	0.42391304	2.187610e-06	22.42281	25.73891	3.908661e-07	X	-	40921754	C	T
29	**rs1150509**	0.42391304	2.187610e-06	22.42281	25.73891	3.908661e-07	X	-	40922534	T	C
30	**rs1150508**	0.42391304	2.187610e-06	22.42281	25.73891	3.908661e-07	X	-	40922745	G	A
31	**rs1263911**	0.42391304	2.187610e-06	22.42281	25.73891	3.908661e-07	X	-	40924859	G	A
32	**rs1263909**	0.42391304	2.187610e-06	22.42281	25.73891	3.908661e-07	X	-	40925114	T	A
33	**rs1263908**	0.42391304	2.187610e-06	22.42281	25.73891	3.908661e-07	X	-	40925187	A	G
34	**rs6034344**	0.08152174	2.667606e-06	22.04192	25.30169	4.902778e-07	20	-	15971823	G	C
35	**rs10025399**	0.27173913	4.914788e-06	20.87021	23.95670	9.852657e-07	4	-	166295607	T	A
36	**rs1370682**	0.27173913	4.914788e-06	20.87021	23.95670	9.852657e-07	4	-	166296379	A	G
37	**rs969411**	0.22282609	3.158703e-06	21.71766	24.92948	5.946608e-07	4	-	32278133	T	C

38	**rs12650672**	0.22282609	3.158703e-06	21.71766	24.92948	5.946608e-07	4	-	32279524	T	G
39	**rs9993569**	0.22282609	3.158703e-06	21.71766	24.92948	5.946608e-07	4	-	32282212	A	G
40	**rs1117427**	0.22282609	3.158703e-06	21.71766	24.92948	5.946608e-07	4	-	32286601	G	A
41	**rs1563182**	0.22282609	3.158703e-06	21.71766	24.92948	5.946608e-07	4	-	32276958	G	A
42	**rs10005923**	0.22282609	3.158703e-06	21.71766	24.92948	5.946608e-07	4	-	32277495	T	C
43	**rs35493938**	0.22282609	3.158703e-06	21.71766	24.92948	5.946608e-07	4	-	32277873	T	C
44	**rs9564536**	0.29891304	5.202245e-06	20.76135	23.83174	1.051345e-06	13	-	35258606	A	C
45	**rs78466902**	0.03804348	5.390851e-06	20.69315	23.75346	1.094985e-06	15	-	38938033	A	G
46	**rs6610503**	0.42934783	7.273864e-06	20.11983	23.09535	1.541632e-06	X	-	40890831	G	A
47	**rs5918096**	0.42934783	7.273864e-06	20.11983	23.09535	1.541632e-06	X	-	40895356	A	G
48	**rs11223239**	0.32608696	7.510588e-06	20.05858	23.02504	1.599051e-06	11	-	132706250	A	G
49	**rs17429356**	0.09782609	8.209498e-06	19.88845	22.82975	1.770048e-06	20	-	54726631	T	C
50	**rs1884616**	0.09782609	8.209498e-06	19.88845	22.82975	1.770048e-06	20	-	54615829	C	A
51	**rs847129**	0.11413043	8.850235e-06	19.74480	22.66485	1.928636e-06	2	-	176932555	A	G
52	**rs847127**	0.11413043	8.850235e-06	19.74480	22.66485	1.928636e-06	2	-	176936443	T	C
53	**rs62195469**	0.22282609	1.048579e-05	19.42082	22.29296	2.340633e-06	2	-	202713495	T	C

54	**rs77918077**	0.05978261	1.160759e-05	19.22674	22.07018	2.628612e-06	10	-	63723517	A	G

Table 28: Most Significant SNPs from Genome Wide Association Studies*

*[*The most statistically significant SNPs from each of the genome wide association studies performed along with non-adjusted and adjusted p-values and SNP position details]*

Marker ID	LZ	Nearest gene	GWAS	MAGMA	FUMA 1	FUMA 2	CADD	RDB	CBI	GWAS Catalog	Kidney eQTL?	Prioritisation Score (out of 9)
FULL 266 PERSON ADTKD-*UMOD* GWAS												
rs3935743	Yes	*PLCG2*	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No	**7**
rs2622627	Yes	*ABCG2*	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	No	**7**
rs3114020	Yes	*ABCG2*	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	**6**
rs17731799	Yes	*ABCG2*	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	**5**
rs61814582	Yes	*RP11-33M22.2*	No	No	Yes	Yes	Yes	Yes	No	No	No	**4**
rs2622626	Yes	*ABCG2*	Yes	Yes	No	No	No	No	Yes	No	No	**3**
rs56259150	Yes	*H19*	No	No	Yes	Yes	No	No	Yes	No	No	**3**
rs61909183	Yes	*SNX19*	No	No	Yes	Yes	No	No	Yes	No	No	**3**
rs66548062	Yes	*MCTP2*	No	No	Yes	Yes	No	No	No	No	No	**2**
rs73148239	Yes	*SYT1 LOC105369863*	No	No	Yes	Yes	No	No	No	No	No	**2**

rs72785963	Yes	*LINC01170*	No	No	Yes	Yes	No	No	No	No	No	**2**
rs1462758	Yes	*SGPT2*	Yes	No	No	No	No	No	No	No	No	**1**
rs11054458	No	*ETV6*	Yes	No	Yes	Yes	No	No	Yes	No	No	**-**
rs1450186	No	*GPR141*	No	No	Yes	Yes	No	No	Yes	No	No	**-**
rs11651009	No	*DNAH9*	Yes	No	Yes	Yes	Yes	Yes	Yes	No	No	**-**
IMPUTED FULL 266 PERSON ADTKD-*UMOD* GWAS												
rs143840619	Yes	*ZNF326*	Yes	No	Yes	Yes	Yes	No	No	No	No	**4**
rs5761621	Yes	*CRYBA4 CRYBB1*	No	Yes	Yes	Yes	No	Yes	No	No	No	**4**
rs8074402	Yes	*DNAH9*	Yes	No	Yes	Yes	No	No	No	No	No	**3**
rs12459907	Yes	*KMT5C*	Yes	Yes	No	No	No	No	No	No	No	**2**
rs80352850	Yes	*ZNF326*	Yes	No	No	No	No	No	No	No	No	**1**
rs28562846	Yes	*MEP1B*	Yes	No	No	No	No	No	No	No	No	**1**
rs16962801	Yes	*MEP1B*	Yes	No	No	No	No	No	No	No	No	**1**
rs62060825	Yes	*DNAH9*	Yes	No	No	No	No	No	No	No	No	**1**
rs499779	Yes	*C19orf18*	Yes	No	No	No	No	No	No	No	No	**1**
rs7001584	Yes	*SCARA3 LOC124901921*	Yes	No	No	No	No	No	No	No	No	**1**
rs6558012	Yes	*SCARA3*	Yes	No	No	No	No	No	No	No	No	**1**

		LOC124901921										
rs4545046	Yes	*SCARA3 LOC124901920*	Yes	No	No	No	No	No	No	No	No	**1**
rs72869201	Yes	*ARNTL*	Yes	No	No	No	No	No	No	No	No	**1**
rs72870911	Yes	*ARNTL*	Yes	No	No	No	No	No	No	No	No	**1**
rs17452383	Yes	*BTBD10*	Yes	No	No	No	No	No	No	No	No	**1**
rs56384052	Yes	*BTBD10*	Yes	No	No	No	No	No	No	No	No	**1**
rs142074501	Yes	*BTBD10*	Yes	No	No	No	No	No	No	No	No	**1**
rs3761862	Yes	*BTBD10*	Yes	No	No	No	No	No	No	No	No	**1**
rs72870943	Yes	*BTBD10*	Yes	No	No	No	No	No	No	No	No	**1**
rs72870946	Yes	*BTBD10*	Yes	No	No	No	No	No	No	No	No	**1**
rs143305491	Yes	*BTBD10*	Yes	No	No	No	No	No	No	No	No	**1**
rs149617495	Yes	*BTBD10*	Yes	No	No	No	No	No	No	No	No	**1**
rs145159588	Yes	*BTBD10*	Yes	No	No	No	No	No	No	No	No	**1**
rs72857026	Yes	*BTBD10*	Yes	No	No	No	No	No	No	No	No	**1**
rs72857027	Yes	*BTBD10*	Yes	No	No	No	No	No	No	No	No	**1**

rs72857028	Yes	*BTBD10*	Yes	No	No	No	No	No	No	No	No	**1**
rs72857030	Yes	*BTBD10*	Yes	No	No	No	No	No	No	No	No	**1**
rs72857043	Yes	*BTBD10*	Yes	No	No	No	No	No	No	No	No	**1**
rs72857050	Yes	*BTBD10*	Yes	No	No	No	No	No	No	No	No	**1**
rs1077956	Yes	*BTBD10*	Yes	No	No	No	No	No	No	No	No	**1**
rs72857069	Yes	*BTBD10*	Yes	No	No	No	No	No	No	No	No	**1**
rs72857071	Yes	*BTBD10*	Yes	No	No	No	No	No	No	No	No	**1**
rs74543117	Yes	*BTBD10*	Yes	No	No	No	No	No	No	No	No	**1**
rs2061438	Yes	*BTBD10*	Yes	No	No	No	No	No	No	No	No	**1**
rs4356206	Yes	*BTBD10*	Yes	No	No	No	No	No	No	No	No	**1**
rs72857077	Yes	*BTBD10*	Yes	No	No	No	No	No	No	No	No	**1**
rs6605364	Yes	*LOC124900652*	Yes	No	No	No	No	No	No	No	No	**1**
rs11650039	Yes	*DNAH9*	Yes	No	No	No	No	No	No	No	No	**1**
rs62061778	Yes	*DNAH9*	Yes	No	No	No	No	No	No	No	No	**1**
rs72810873	Yes	*DNAH9*	Yes	No	No	No	No	No	No	No	No	**1**

rs8079335	Yes	*DNAH9*	Yes	No	No	No	No	No	No	No	No	**1**
rs11655363	Yes	*DNAH9*	Yes	No	No	No	No	No	No	No	No	**1**
rs11651009	Yes	*DNAH9*	Yes	No	No	No	No	No	No	No	No	**1**
rs17602516	Yes	*DNAH9*	Yes	No	No	No	No	No	No	No	No	**1**
rs28456990	Yes	*NAALADL2*	Yes	No	No	No	No	No	No	No	No	**1**
rs3935743	Yes	*PLCG2*	Yes	No	No	No	No	No	No	No	No	**1**
rs7437393	Yes	*LOC124900652*	Yes	No	No	No	No	No	No	No	No	**1**
rs13257161	Yes	*SCARA3*	Yes	No	No	No	No	No	No	No	No	**1**
rs1870073	Yes	*KMT5C*	Yes	Yes	No	No	No	No	No	No	No	**1**
rs1544929	Yes	*ZNF418*	Yes	No	No	No	No	No	No	No	No	**1**
rs34647569	Yes	*ZNF418*	Yes	No	No	No	No	No	No	No	No	**1**
rs257676	Yes	*C19orf18*	Yes	No	No	No	No	No	No	No	No	**1**
rs2865786	Yes	*STPG2*	Yes	No	No	No	No	No	No	No	No	**1**
rs1350974	Yes	*STPG2*	Yes	No	No	No	No	No	No	No	No	**1**
rs1350973	Yes	*STPG2*	Yes	No	No	No	No	No	No	No	No	**1**
rs1462759	Yes	*STPG2*	Yes	No	No	No	No	No	No	No	No	**1**
rs1462758	Yes	*STPG2*	Yes	No	No	No	No	No	No	No	No	**1**
rs7688316	Yes	*STPG2*	Yes	No	No	No	No	No	No	No	No	**1**
rs1599660	Yes	*STPG2*	Yes	No	No	No	No	No	No	No	No	**1**
rs73066266	Yes	*ZNF606*	Yes	No	No	No	No	No	No	No	No	**1**

rs34243962	Yes	*MAFB*	Yes	No	No	No	No	No	No	No	No	**1**
rs7255565	Yes	*ZNF417*	Yes	No	No	No	No	No	No	No	No	**1**
rs7253038	Yes	*ZNF417*	Yes	No	No	No	No	No	No	No	No	**1**
rs12972704	Yes	*ZNF417*	Yes	No	No	No	No	No	No	No	No	**1**
rs11054458	Yes	*ETV6*	Yes	No	No	No	No	No	No	No	No	**1**
rs11054460	Yes	*ETV6*	Yes	No	No	No	No	No	No	No	No	**1**
rs115613990	Yes	*ARNTL*	Yes	No	No	No	No	No	No	No	No	**1**
rs72869173	Yes	*ARNTL*	Yes	No	No	No	No	No	No	No	No	**1**
rs58295558	Yes	*ARNTL*	Yes	No	No	No	No	No	No	No	No	**1**
rs4894486	Yes	*NAALADL2*	Yes	No	No	No	No	No	No	No	No	**1**
rs9883274	Yes	*NAALADL2*	Yes	No	No	No	No	No	No	No	No	**1**
rs114080582	Yes	*ST8SIA5*	No	No	No	No	No	No	No	No	No	**0**
rs7254431	Yes	*ZNF417*	No	No	No	No	No	No	No	No	No	**0**
rs79012440	Yes	*KMT5C*	No	No	No	No	No	No	No	No	No	**0**
rs113059329	Yes	*LINC01900*	No	No	No	No	No	No	No	No	No	**0**
rs114270590	Yes	*LBH*	No	No	No	No	No	No	No	No	No	**0**
rs146302733	Yes	*KCNIP4*	No	No	No	No	No	No	No	No	No	**0**

rs1307415325	?	*LOC105373204*	?	No	No	No	No	No	No	No	No	-
rs10058279	No	*TENM2*	No	No	No	No	No	No	No	No	No	-
92 PERSON ADTKD-*UMOD* p.(Val93_Gly97delinsAlaAlaSerCys) variant GWAS												
rs77129692	Yes	*HTR2A*	Yes	No	Yes	Yes	Yes	No	Yes	No	No	**5**
rs12845211	?	*RP11-469E19.1*	Yes	No	Yes	Yes	Yes	No	Yes	No	No	**5**
rs7054156	?	*RP11-469E19.1*	Yes	No	No	Yes	Yes	No	Yes	No	No	**4**
rs1370682	Yes	*CPE*	Yes	No	Yes	Yes	Yes	No	No	No	No	**4**
rs73175544	Yes	*HTR2A*	Yes	No	No	Yes	No	No	Yes	No	No	**3**
rs9564536	Yes	*LOC105370159*	Yes	No	Yes	Yes	No	No	No	No	No	**3**
rs62224727	No	*RNA5SP487*	Yes	No	Yes	Yes	No	No	No	No	No	-
IMPUTED 92 PERSON ADTKD-*UMOD* p.(Val93_Gly97delinsAlaAlaSerCys) variant GWAS												
rs10152175	Yes	*SPATA8*	No	No	Yes	Yes	Yes	Yes	Yes	No	No	**5**
rs73175542	Yes	*HTR2A*	Yes	No	No	No	No	No	No	No	No	**1**
rs73175544	Yes	*HTR2A*	Yes	No	No	No	No	No	No	No	No	**1**
rs73175546	Yes	*HTR2A*	Yes	No	No	No	No	No	No	No	No	**1**
rs115416253	Yes	*HTR2A*	Yes	No	No	No	No	No	No	No	No	**1**
rs55812936	Yes	*HTR2A*	Yes	No	No	No	No	No	No	No	No	**1**
rs73175549	Yes	*HTR2A*	Yes	No	No	No	No	No	No	No	No	**1**
rs112452802	Yes	*HTR2A*	Yes	No	No	No	No	No	No	No	No	**1**
rs12845211	?	*RP11-469E19.1*	Yes	No	No	No	No	No	No	No	No	**1**
rs62224727	Yes	*RNA5SP487*	Yes	No	No	No	No	No	No	No	No	**1**

rs118059905	Yes	*RNA5SP487*	Yes	No	No	No	No	No	No	No	No	**1**
rs4827251	?	*USP9X*	Yes	No	No	No	No	No	No	No	No	**1**
rs5918100	?	*USP9X*	Yes	No	No	No	No	No	No	No	No	**1**
rs4827252	?	*USP9X*	Yes	No	No	No	No	No	No	No	No	**1**
rs7054156	?	*RP11-469E19.1*	Yes	No	No	No	No	No	No	No	No	**1**
rs17322309	?	*USP9X*	Yes	No	No	No	No	No	No	No	No	**1**
rs5918103	?	*USP9X*	Yes	No	No	No	No	No	No	No	No	**1**
rs5917405	?	*USP9X*	Yes	No	No	No	No	No	No	No	No	**1**
rs36042689	?	*USP9X*	Yes	No	No	No	No	No	No	No	No	**1**
rs12689198	?	*USP9X*	Yes	No	No	No	No	No	No	No	No	**1**
rs1263919	?	*USP9X*	Yes	No	No	No	No	No	No	No	No	**1**
rs1263917	?	*USP9X*	Yes	No	No	No	No	No	No	No	No	**1**
rs1263915	?	*USP9X*	Yes	No	No	No	No	No	No	No	No	**1**
rs1263912	?	*USP9X*	Yes	No	No	No	No	No	No	No	No	**1**
rs1150526	?	*USP9X*	Yes	No	No	No	No	No	No	No	No	**1**
rs1150520	?	*USP9X*	Yes	No	No	No	No	No	No	No	No	**1**
rs1150533	?	*USP9X*	Yes	No	No	No	No	No	No	No	No	**1**
rs1150513	?	*USP9X*	Yes	No	No	No	No	No	No	No	No	**1**
rs1150509	?	*USP9X*	Yes	No	No	No	No	No	No	No	No	**1**
rs1150508	?	*USP9X*	Yes	No	No	No	No	No	No	No	No	**1**
rs1263911	?	*USP9X*	Yes	No	No	No	No	No	No	No	No	**1**
rs1263909	?	*USP9X*	Yes	No	No	No	No	No	No	No	No	**1**
rs1263908	?	*USP9X*	Yes	No	No	No	No	No	No	No	No	**1**
rs10025399	Yes	*CPE*	Yes	No	No	No	No	No	No	No	No	**1**
rs1370682	Yes	*CPE*	Yes	No	No	No	No	No	No	No	No	**1**
rs969411	Yes	*RP11-240A16.1*	Yes	No	No	No	No	No	No	No	No	**1**

rs12650672	Yes	*RP11-240A16.1*	Yes	No	No	No	No	No	No	No	No	**1**
rs9993569	Yes	*RP11-240A16.1*	Yes	No	No	No	No	No	No	No	No	**1**
rs1117427	Yes	*RP11-240A16.1*	Yes	No	No	No	No	No	No	No	No	**1**
rs1563182	Yes	*RP11-240A16.1*	Yes	No	No	No	No	No	No	No	No	**1**
rs10005923	Yes	*RP11-240A16.1*	Yes	No	No	No	No	No	No	No	No	**1**
rs35493938	Yes	*RP11-240A16.1*	Yes	No	No	No	No	No	No	No	No	**1**
rs9564536	Yes	*LOC105370159*	Yes	No	No	No	No	No	No	No	No	**1**
rs6610503	?	*USP9X*	Yes	No	No	No	No	No	No	No	No	**1**
rs5918096	?	*USP9X*	Yes	No	No	No	No	No	No	No	No	**1**
rs17429356	Yes	*LOC105372680*	Yes	No	No	No	No	No	No	No	No	**1**
rs1884616	Yes	*LOC105372680*	Yes	No	No	No	No	No	No	No	No	**1**
rs847129	Yes	*EVX2*	Yes	No	No	No	No	No	No	No	No	**1**
rs847127	Yes	*EVX2*	Yes	No	No	No	No	No	No	No	No	**1**
rs6034344	Yes	*MACROD2*	No	No	No	No	No	No	No	No	No	**0**
rs78466902	Yes	*LINC02694*	No	No	No	No	No	No	No	No	No	**0**
rs11223239	Yes	*OPCML*	No	No	No	No	No	No	No	No	No	**0**
rs62195469	Yes	*CDK15*	No	No	No	No	No	No	No	No	No	**0**

rs77918077	Yes	*ARID5B*	No	No	No	No	No	No	No	No	No	**0**

Table 29: Prioritisation Score and Ranking of SNPs of Interest from Genome Wide Association Studies*

*[*The prioritisation annotation of the most statistically significant SNPs from each of the genome wide association studies performed. Each variant is evaluated across multiple lines of evidence, including statistical significance, gene proximity, bioinformatic prediction tools (e.g., CADD, MAGMA, FUMA), regulatory potential, and kidney-specific expression quantitative trait loci (eQTLs), to generate an overall prioritisation score out of 9. (LZ) Hit supported on LOCUSZOOM, (GWAS) Multiple hits across GWASs done, (MAGMA) Gene featured in top MAGMA Gene-burden studies, (FUMA1) Genomic risk locus or lead SNP on FUMA, (FUMA2) Individually significant SNP on FUMA, (CADD) SNP in LD with, SNP in top CADD score list (FUMA), (RDB) SNP in LD with SNP in top RDB score list (FUMA), (CB!) Relevant chromatin bridging interaction, (GWASCatalog) Relevant GWASCatalog trait.]*

TOP GWAS SNPS TO EXPLORE FURTHER								
Marker ID	**Nearest Gene**	**Function of Associated Gene**	**GWAS Catalog Trait**	**Kidney Expression of Protein/RNA (HPA)**	**Cellular Expression (HPA)**	**PubMed Info for SNP**	**Animal Model**	**Human Phenotype (Gene)**
rs3935743	*PLCG2*	The protein encoded by this gene is a transmembrane signalling enzyme that catalyses the conversion of 1-phosphatidyl-1D-myo-inositol 4,5-bisphosphate to	-	Low in glomeruli and low in tubules 15.3 nTPM (RNA)	Cytosol and Plasma Membrane	1 citation – one of 237 SNPs for Autism Spectrum Disorder classifier in the CEU cohort	3 x mouse models	Colitis Ulcerative colitis Cataract Vitiligo PLAID (immune deficiency) (AD) APLAID (autoinflammation) (AD) Ibrutinib resistance in CLL

		1D-myo-inositol 1,4,5-trisphosphate (IP3) and diacylglycerol (DAG) using calcium as a cofactor. IP3 and DAG are second messenger molecules important for transmitting signals from growth factor receptors and immune system receptors across the cell membrane. Variants in this gene have been found in autoinflammation, antibody deficiency, and immune dysregulation syndrome and familial cold autoinflammatory syndrome 3.						Inflammatory response in Alzheimer's Disease Familial Cold Autoinflammatory Syndrome (AD) Parkinson's Disease Allergy Osteoarthritis Rheumatoid arthritis Solid Cancers

rs2622627	*ABCG2*	The membrane-associated protein encoded by this gene is included in the superfamily of ATP-binding cassette (ABC) transporters. ABC proteins transport various molecules across extra- and intra-cellular membranes. Acts as a urate exporter functioning in both renal and extra-renal urate excretion. In kidney, it also functions as a physiological exporter of the uremic toxin indoxyl sulfate (By similarity). Also involved in the excretion of steroids like estrone 3-	Serum uric acid levels (Yang, B., 2014) – 4e20 Serum uric acid levels (Lee, J., 2018) – 2e16	Medium in CD and DCT 3.4 nTPM (RNA)	Nucleus and Plasma Membrane	0 citations	4 x mouse models	Gout Hyperuricaemia Renal and extra-renal uric acid excretion Erythropoiesis (blood group) Porphyrin homeostasis Drug resistance Calcium oxalate stone Reduced GFR Secretion of vitamins into breast milk Placental limitation of drug passage to foetus Early stem cell renewal Steroid excretion *Promotes renal dysfunction in CKD, increase systemic inflammatory responses and decrease cellular autophagic responses to stress (PMID 28461764)*

		sulfate/E1S, 3beta-sulfooxy-androst-5-en-17-one/DHEAS, and other sulfate conjugates.						
rs3114020	*ABCG2*	The membrane-associated protein encoded by this gene is included in the superfamily of ATP-binding cassette (ABC) transporters. ABC proteins transport various molecules across extra- and intra-cellular membranes. Acts as a urate exporter functioning in both renal and extrarenal urate excretion. In kidney, it also functions as a physiological exporter of the uremic toxin	Gout (Nakayama, A., 2016) – 9e35 Urate Levels (Tin, A., 2019) – 2e18 Serum uric acid in response to allopurinol in Gout (Wen, C.C., 2015) – 2e8	Medium in CD and DCT 3.4 nTPM (RNA)	Nucleus and Plasma Membrane	SNP associated with Lamotrigine concentration in epilepsy SNP associated with gout risk in a Chinese Han population SNP predicts clinical outcome of NSCLC in a Chinese	4 x mouse models	Gout (AD) Hyperuricaemia Renal and extra-renal urate excretion Erythropoiesis (blood group) Porphyrin homeostasis Drug resistance Calcium oxalate kidney stones Reduced GFR Secretion of vitamins into breast milk Placental limitation of drug passage to foetus Early stem cell renewal Steroid excretion *Promotes renal dysfunction in CKD, increase systemic inflammatory responses and decrease cellular autophagic responses to stress (PMID 28461764)*

		indoxyl sulfate (By similarity). Also involved in the excretion of steroids like estrone 3-sulfate/E1S, 3beta-sulfooxy-androst-5-en-17-one/DHEAS, and other sulfate conjugates.	Gout (Sumpter, N.A., 2022) – 1e19 Serum uric acid in response to allopurinol in Gout (Wen, C.C., 2015) – 1e7			population Genetic association with gout		
rs17731799	*ABCG2*	The membrane-associated protein encoded by this gene is included in the superfamily of ATP-binding cassette (ABC) transporters. ABC proteins transport various molecules across	GFR (Kanai, M., 2018) – 9e9	Medium in CD and DCT 3.4 nTPM (RNA)	Nucleus and Plasma Membrane	SNP associated with gout risk in a Chinese Han population	4 x mouse models	Gout (AD) Hyperuricaemia Renal and extra-renal urate excretion Erythropoiesis (blood group) Porphyrin homeostasis Drug resistance Calcium oxalate kidney stones Reduced GFR

		extra- and intra-cellular membranes. Acts as a urate exporter functioning in both renal and extrarenal urate excretion. In kidney, it also functions as a physiological exporter of the uremic toxin indoxyl sulfate (By similarity). Also involved in the excretion of steroids like estrone 3-sulfate/E1S, 3beta-sulfooxy-androst-5-en-17-one/DHEAS, and other sulfate conjugates.						Secretion of vitamins into breast milk Placental limitation of drug passage to foetus Early stem cell renewal Steroid excretion *Promotes renal dysfunction in CKD, increase systemic inflammatory responses and decrease cellular autophagic responses to stress (PMID 28461764)*
rs77129692	*HTR2A*	This gene encodes one of the receptors for serotonin, a neurotransmitter with many roles.	-	Protein not detected 0.1 nTPM (RNA)	Plasma Membrane	0 citations	3 x mouse models	Susceptibility to Schizophrenia Susceptibility to OCD Response to Citalopram in major depressive disorder Urate measurement

		Variants in this gene are associated with susceptibility to schizophrenia and obsessive-compulsive disorder and are also associated with response to the antidepressant citalopram in patients with major depressive disorder (MDD). MDD patients who also have a variant in intron 2 of this gene show a significantly reduced response to citalopram as this antidepressant downregulates expression of this gene. Multiple transcript variants encoding different isoforms						

		have been found for this gene.						
rs12845211	*RP11-469E19.1*		-	*Pseudogene*	*Pseudogene*	0 citations	*Pseudogene*	*Pseudogene*
rs7054156	*RP11-469E19.1*		-	*Pseudogene*	*Pseudogene*	0 citations	*Pseudogene*	*Pseudogene*
rs1370682	*CPE*	This gene encodes a member of the M14 family of metallocarboxypeptidases. The encoded preproprotein is proteolytically processed to generate the mature peptidase. This peripheral membrane protein cleaves C-terminal amino acid residues and is involved in the biosynthesis of peptide hormones and neurotransmitters, including insulin. This	-	Protein not detected 66.3 nTPM (RNA)	Extracellular, Golgi and Plasma Membrane	0 citations	2 x mouse models	None

		protein may also function independently of its peptidase activity, as a neurotrophic factor that promotes neuronal survival, and as a sorting receptor that binds to regulated secretory pathway proteins, including prohormones. Variants in this gene are implicated in type 2 diabetes.						
rs10152175	*SPATA8*	SPATA8 (Spermatogenesis Associated 8) is an RNA Gene and is affiliated with the lncRNA class.	-	*RNA Gene*	*RNA Gene*	0 citations	None	None

Table 30: Further Information on Top Ranked SNPs from Genome Wide Association Studies prioritized for Further Exploration*

*[*Detailed functional and biological context for the top-ranked SNPs identified in the genome wide association studies performed, which were prioritised for further investigation. It integrates information on gene function, associated traits from the GWAS Catalog, kidney and cellular expression data from Human Protein Atlas (HPA), literature evidence, and the availability of relevant animal models and human phenotypes sourced from GWAS Catalog, OMIM and Ensembl data.]*

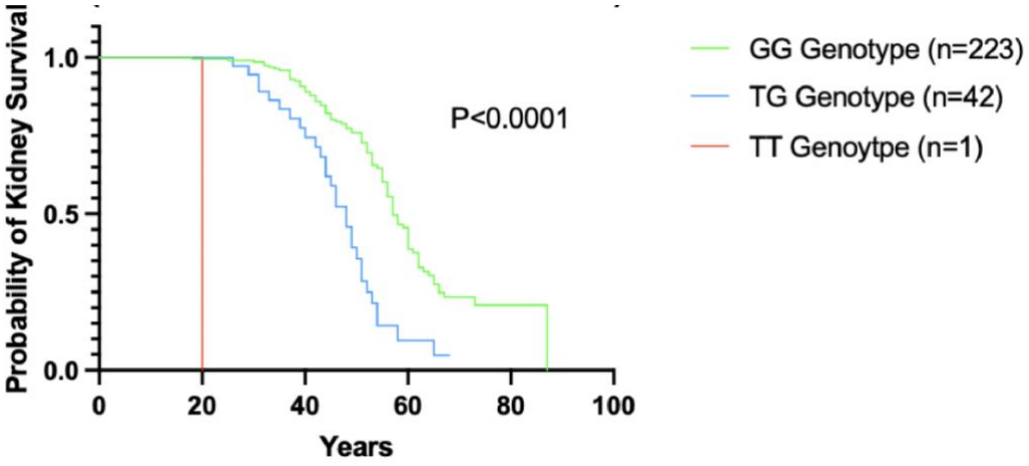

Figure 76: Kidney Survival Curves for Top Prioritised SNPs from Genome Wide Association Studies for Further Exploration (*rs3935743)**

*[*Kidney survival curves for each of the three genotypes at SNP rs3935743 within the full 266-person ADTKD-UMOD GWAS cohort. When a Cox proportional hazards model was applied, there was a statistically significant difference in kidney survival between each of the three genotypes. The TT genotype was associated with the worst kidney survival in the cohort and the GG genotype was associated with the best.]*

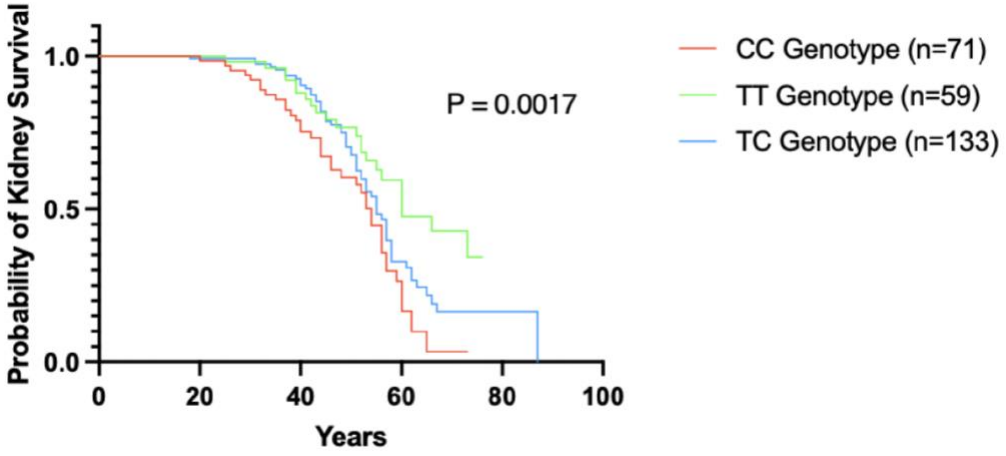

Figure 77: Kidney Survival Curves for Top Prioritised SNPs from Genome Wide Association Studies for Further Exploration (*rs3114020*)*

*[*Kidney survival curves for each of the three genotypes at SNP rs3114020 within the full 266-person ADTKD-UMOD GWAS cohort. When a Cox proportional hazards model was applied, there was a statistically significant difference in kidney survival between each of the three genotypes. The CC genotype was associated with the worst kidney survival in the cohort and the TT genotype was associated with the best.]*

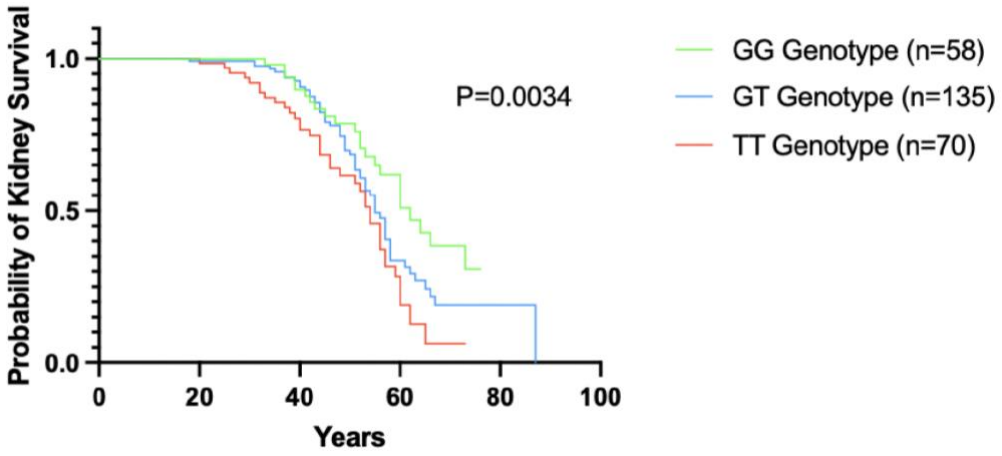

Figure 78: Kidney Survival Curves for Top Prioritised SNPs from Genome Wide Association Studies for Further Exploration (*rs17731799*)*

*[*Kidney survival curves for each of the three genotypes at SNP rs17731799 within the full 266-person ADTKD-UMOD GWAS cohort. When a Cox proportional hazards model was applied, there was a statistically significant difference in kidney survival between each of the three genotypes. The TT genotype was associated with the worst kidney survival in the cohort and the GG genotype was associated with the best.]*

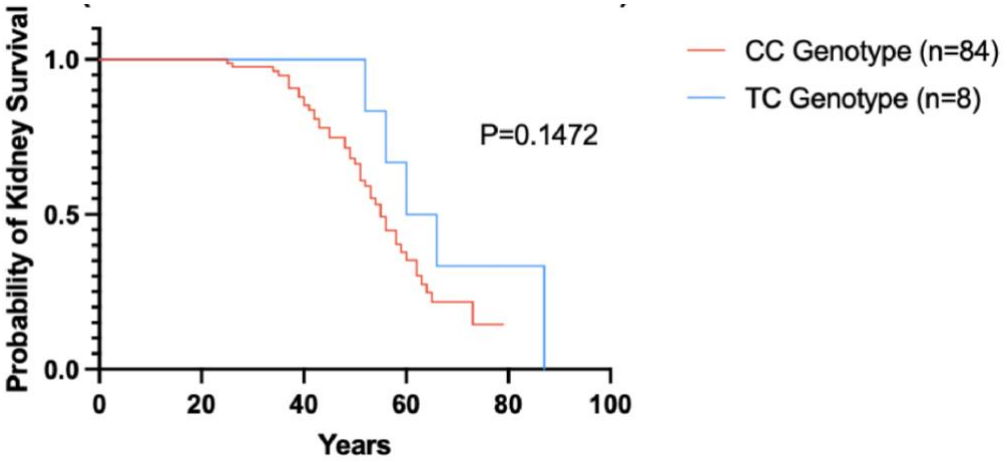

Figure 79: Kidney Survival Curves for Top Prioritised SNPs from Genome Wide Association Studies for Further Exploration (*rs77129692*)*

*[*Kidney survival curves for each of the three genotypes at SNP rs77129692 within the full 266-person ADTKD-UMOD GWAS cohort. When a Cox proportional hazards model was applied, there was no statistically significant difference in kidney survival between each of the three genotypes. The CC genotype was associated with the worst kidney survival in the cohort and the TC genotype was associated with the best.]*

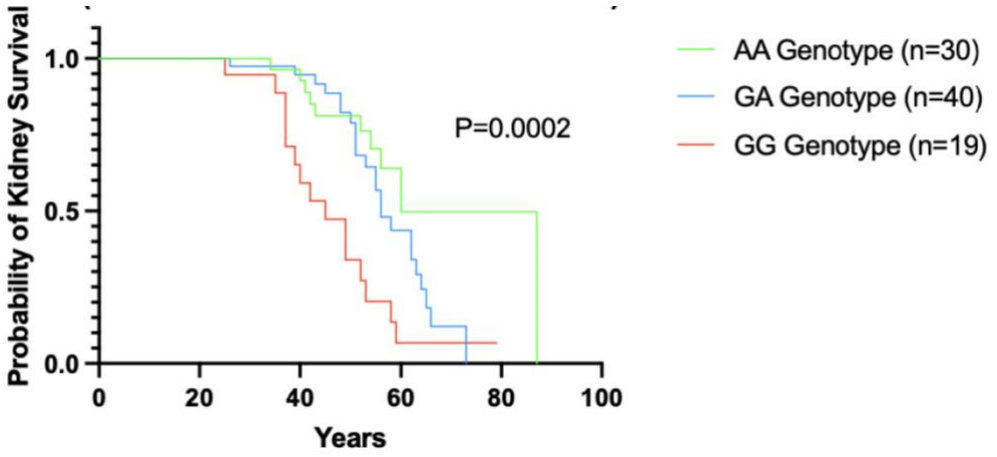

Figure 80: Kidney Survival Curves for Top Prioritised SNPs from Genome Wide Association Studies for Further Exploration (*rs12845211*)*

*[*Kidney survival curves for each of the three genotypes at SNP rs12845211 within the full 266-person ADTKD-UMOD GWAS cohort. When a Cox proportional hazards model was applied, there was a statistically significant difference in kidney survival between each of the three genotypes. The GG genotype was associated with the worst kidney survival in the cohort and the AA genotype was associated with the best.]*

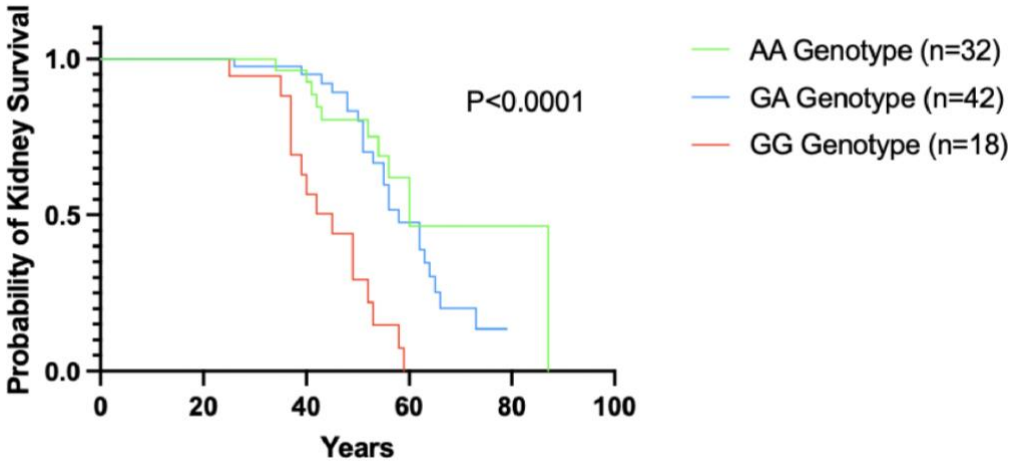

Figure 81: Kidney Survival Curves for Top Prioritised SNPs from Genome Wide Association Studies for Further Exploration (*rs7054156*)*

*[*Kidney survival curves for each of the three genotypes at SNP rs7054156 within the full 266-person ADTKD-UMOD GWAS cohort. When a Cox proportional hazards model was applied, there was a statistically significant difference in kidney survival between each of the three genotypes. The GG genotype was associated with the worst kidney survival in the cohort and the AA genotype was associated with the best.]*

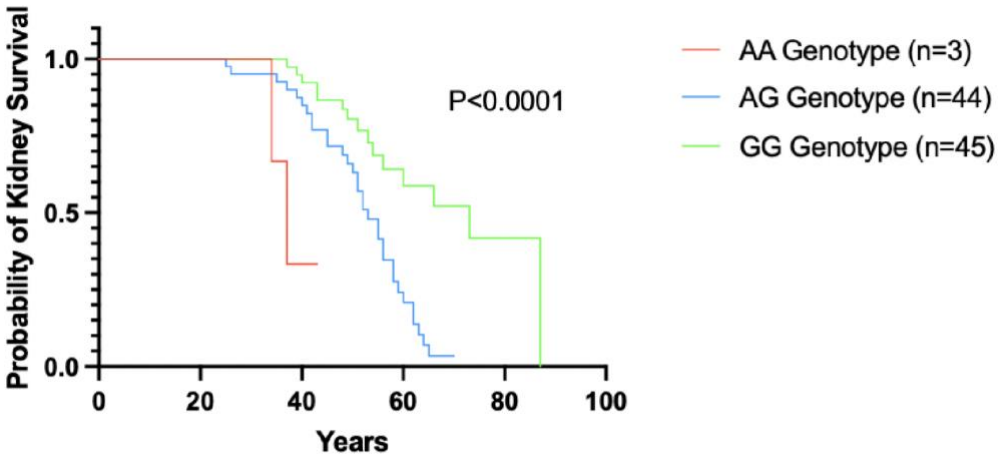

Figure 82: Kidney Survival Curves for Top Prioritised SNPs from Genome Wide Association Studies for Further Exploration (*rs1370682)

*[*Kidney survival curves for each of the three genotypes at SNP rs1370682 within the full 266-person ADTKD-UMOD GWAS cohort. When a Cox proportional hazards model was applied, there was a statistically significant difference in kidney survival between each of the three genotypes. The AA genotype was associated with the worst kidney survival in the cohort and the GG genotype was associated with the best.]*

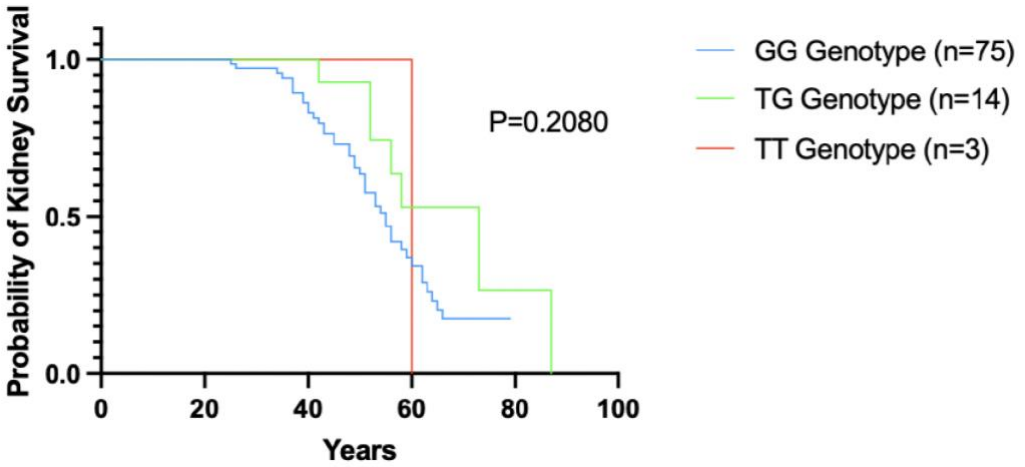

Figure 83: Kidney Survival Curves for Top Prioritised SNPs from Genome Wide Association Studies for Further Exploration (*rs10152175*)*

*[*Kidney survival curves for each of the three genotypes at SNP rs10152175 within the full 266-person ADTKD-UMOD GWAS cohort. When a Cox proportional hazards model was applied, there was no statistically significant difference in kidney survival between each of the three genotypes. The TT genotype was associated with the worst kidney survival in the cohort and the TG genotype was associated with the best.]*

Appendix B: Most Significant Genes from Gene-burden Studies

FULL 266 PERSON ADTKD-*UMOD* COHORT					
GENE	**CHR**	**NSNPS**	**NPARAM**	**ZSTAT**	**P-value**
COX6B2	19	23	11	3.9291	4.2640e-05
SUV420H2	19	24	10	3.8650	5.5550e-05
AC020922.1	19	24	10	3.8650	5.5550e-05
ABCG2	4	75	24	3.8366	6.2381e-05
TMEM150B	19	24	10	3.7885	7.5770e-05
CTD-2105E13.6	19	21	10	3.7552	8.6609e-05
FAM71E2	19	23	11	3.6806	1.1636e-04
BRSK1	19	26	10	3.6704	1.2110e-04
KIAA1586	6	7	4	3.3904	3.4899e-04
HSPBP1	19	30	11	3.3816	3.6030e-04
FAM163A	1	31	14	3.3217	4.4735e-04
SLC35D3	6	26	5	3.3142	4.5957e-04
IL11	19	21	11	3.2592	5.5863e-04
PEX7	6	30	6	3.2179	6.4571e-04
ZNF451	6	11	5	3.2123	6.5847e-04
CCZ1	7	14	9	3.1946	7.0011e-04
OR56B3P	11	17	10	3.1499	8.1652e-04
IL17C	16	19	10	3.1499	8.1667e-04
IMPUTED FULL 266 PERSON ADTKD-*UMOD* COHORT					
AC020922.1	19	332	43	4.5312	2.9328e-06
SUV420H2	19	353	48	4.5134	3.1893e-06
IL11	19	351	49	4.3899	5.6711e-06
COX6B2	19	351	47	4.3649	6.3594e-06
CTD-2105E13.6	19	355	48	4.3306	7.4356e-06
FAM71E2	19	357	48	4.3076	8.2530e-06
TMEM150B	19	341	44	4.2798	9.3532e-06
BRSK1	19	332	38	3.9536	3.8496e-05
TMEM190	19	321	46	3.7057	1.0542e-04
ABCG2	4	622	55	3.7014	1.0721e-04
TMEM238	19	326	48	3.6621	1.2509e-04
CRYBA4	22	382	42	3.6531	1.2956e-04
92 PERSON ADTKD-*UMOD* p.(Val93_Gly97delinsAlaAlaSerCys) VARIANT COHORT					
SIGLEC15	18	8	5	3.8586	5.7020e-05
TMEM107	17	12	6	3.7515	8.7901e-05
SLC22A8	11	32	13	3.6906	1.1188e-04
UNC5D	8	51	30	3.5506	1.9218e-04
SLC22A6	11	37	15	3.4819	2.4892e-04
IMPUTED 92 PERSON ADTKD-*UMOD* p.(Val93_Gly97delinsAlaAlaSerCys) VARIANT COHORT					
ST8SIA5	18	573	68	3.8773	5.2805e-05
ANGPT2	8	847	67	3.8508	5.8869e-05
SLC22A6	11	257	28	3.8314	6.3718e-05

RP11-73M18.2	14	479	25	3.6877	1.1313e-04
AC020922.1	19	327	42	3.6274	1.4313e-04
SUV420H2	19	346	45	3.5977	1.6054e-04
IL11	19	343	47	3.5725	1.7676e-04
SLC22A8	11	261	32	3.5661	1.8117e-04
CTD-2105E13.6	19	347	46	3.5262	2.1079e-04
KLC1	14	507	22	3.4996	2.3301e-04
FAM71E2	19	349	46	3.4991	2.3339e-04
COX6B2	19	344	44	3.4871	2.4414e-04
APOPT1	14	325	18	3.4594	2.7064e-04
UBL3	13	496	30	3.4479	2.8244e-04

Table 31: Most Significant Genes from Gene-burden Studies*

*[*The most statistically significant genes from each of the gene burden studies performed along with p-values and gene position details and number of SNPs included on each gene. (GENE) name of the tested gene, (CHR) Chromosome on which the gene is located (NSNPS) number of SNPs included in the gene-based test, (NPARAM) number of parameters used in the regression model, (ZSTAT) Z-statistic from the gene-based association test, (P-value) statistical significance of the gene's association with the phenotype.]*

FULL 266 PERSON ADTKD-*UMOD* COHORT							
Gene	**Appears more than once in MAGMA Studies**	**MAGMA Significance High**	**Protein/RNA Expressed in kidney**	**Protein/RNA Expression HIGH in kidney**	**Biological function relevant to ADTKD-*UMOD***	**Gene in GWAS list of top SNPs**	**Score**
ABCG2	Yes	Yes	Yes	Yes	Yes	Yes	**6**
COX6B2	Yes	Yes	Yes	No	Yes	No	**4**
SUV420H2	Yes	Yes	Yes	No	Yes	No	**4**
TMEM150B	Yes	Yes	Yes	No	Yes	No	**4**
AC020922.1	Yes	Yes	Yes	No	No	No	**3**
KIAA1586	No	No	Yes	Yes	Yes	No	**3**
HSPBP1	No	No	Yes	Yes	Yes	No	**3**
IL11	Yes	No	Yes	No	Yes	No	**3**
CCZ1	No	No	Yes	Yes	Yes	No	**3**
BRSK1	Yes	No	Yes	No	No	No	**2**
PEX7	No	No	Yes	Yes	Yes	No	**2**
ZNF451	No	No	Yes	No	Yes	No	**2**
FAM163A	No	No	Yes	No	No	No	**1**
FAM71E2	Yes	No	No	No	Yes	No	**-**
SLC35D3	No	No	No	No	Yes	No	**-**
CTD-2105E13.6	Yes	Yes	No	No	?	No	**-**
IL17C	No	No	No	No	Yes	No	**-**
OR56B3P	No	No	No	No	No	No	**-**
IMPUTED FULL 266 PERSON ADTKD-*UMOD* COHORT							
ABCG2	Yes	No	Yes	Yes	Yes	Yes	**5**
COX6B2	Yes	Yes	Yes	No	Yes	No	**4**
SUV420H2	Yes	Yes	Yes	No	Yes	No	**4**
TMEM150B	Yes	Yes	Yes	No	Yes	No	**4**
IL11	Yes	Yes	Yes	No	Yes	No	**4**

AC020922.1	Yes	Yes	Yes	No	No	No	**3**
TMEM190	No	No	Yes	No	Yes	No	**2**
BRSK1	Yes	No	Yes	No	No	No	**2**
TMEM238	No	No	Yes	No	Yes	No	**2**
CRYBA4	No	No	No	No	No	Yes	**-**
CTD-2105E13.6	Yes	Yes	No	No	?	No	**-**
FAM71E2	Yes	Yes	No	No	Yes	No	**-**
92 PERSON ADTKD-*UMOD* p.(Val93_Gly97delinsAlaAlaSerCys) VARIANT COHORT							
SLC22A6	Yes	No	Yes	Yes	Yes	No	**4**
SLC22A8	Yes	No	Yes	Yes	Yes	No	**4**
TMEM107	No	Yes	Yes	Yes	Yes	No	**4**
SIGLEC15	No	Yes	Yes	No	Yes	No	**3**
UNC5D	No	No	Yes	No	Yes	No	**2**
IMPUTED 92 PERSON ADTKD-*UMOD* p.(Val93_Gly97delinsAlaAlaSerCys) VARIANT COHORT							
SLC22A6	Yes	Yes	Yes	Yes	Yes	No	**5**
SLC22A8	Yes	No	Yes	Yes	Yes	No	**4**
ST8SIA5	No	Yes	Yes	No	Yes	Yes	**4**
RP11-73M18.2	Yes	No	Yes	Yes	Yes	No	**4**
KLC1	Yes	No	Yes	Yes	Yes	No	**4**
SUV420H2	Yes	No	Yes	No	Yes	No	**3**
IL11	Yes	No	Yes	No	Yes	No	**3**
COX6B2	Yes	No	Yes	No	Yes	No	**3**
APOPT1	No	No	Yes	Yes	Yes	No	**3**
UBL3	No	No	Yes	Yes	Yes	No	**3**
ANGPT2	No	Yes	Yes	No	Yes	No	**3**
AC020922.1	Yes	No	Yes	No	No	No	**2**
FAM71E2	Yes	No	No	No	Yes	No	**-**
CTD-2105E13.6	Yes	No	No	No	?	No	**-**

Table 32: Prioritisation Score and Ranking of Most Significant Genes from Gene-burden Studies*

*[*The prioritisation of genes identified through gene-burden testing in each of the gene-burden studies performed. Each gene is evaluated based on MAGMA significance, recurrence across analyses, kidney-specific expression (including expression level) taken from Human Protein Atlas data, functional relevance to ADTKD-UMOD, and overlap with top-ranked GWAS SNPs, generating a composite prioritisation score.]*

FULL 266 PERSON ADTKD-*UMOD* COHORT								
GENE	**CHR**	**Occurs in other MAGMA GWAS (x times)**	**Full Gene Name**	**Function**	**Kidney Expression of Protein/RNA (HPA)**	**Cellular Expression (HPA)**	**Top 5 Interacting Proteins**	**Associated Human Phenotypes**
ABCG2	4	2	ATP Binding Cassette Subfamily G Member 2 (JR blood group)	The membrane-associated protein encoded by this gene is included in the superfamily of ATP-binding cassette (ABC) transporters. ABC proteins transport various molecules across extra- and intra-cellular membranes. Acts as a urate exporter functioning in both renal and extrarenal urate excretion. In kidney, it also functions as a physiological exporter of the uremic toxin indoxyl sulfate (By similarity). Also involved in the excretion of steroids like estrone 3-sulfate/E1S, 3beta-sulfooxy-androst-5-	Medium in CD and DCT 3.4 nTPM (RNA)	Nucleus and Plasma Membrane	ATP5F1B ATP5F1A ATP5F1C ATP5F1D PCDH7	Gout Hyperuricaemia Renal and extra-renal uric acid excretion Erythropoiesis (blood group) Porphyrin homeostasis Drug resistance Calcium oxalate stone Reduced GFR Secretion of vitamins into breast milk Placental limitation of drug passage to foetus

				en-17-one/DHEAS, and other sulfate conjugates.				Early stem cell renewal Steroid excretion *Promotes renal dysfunction in CKD, increase systemic inflammatory responses and decrease cellular autophagic responses to stress (PMID 28461764)*
COX6B2	19	3	Cytochrome C Oxidase Subunit 6B2	The last enzyme in the mitochondrial electron transport chain which drives oxidative phosphorylation. Related to AMPK enzyme complex pathway.	No protein data 0.7 nTPM (RNA)	Mitochondrial membrane in crista	COX5B CYC1 UQCRFS1 UQCR10 COX6C	Body height
SUV420H2 (KMT5C)	19	3	Lysine Methyltransferase 5C	Histone methyltransferase that specifically methylates monomethylated 'Lys-20' (H4K20me1) and dimethylated 'Lys-20' (H4K20me2) of histone H4 to produce respectively dimethylated 'Lys-20' (H4K20me2) and trimethylated 'Lys-20'	No protein data 5.1 nTPM (RNA)	Nucleus	CBX3 CBX5 CBX1 H4C6 PSTPIP1	Platelet volume Reticulocyte count Body height Monocyte count Mean corpuscular volume VTE

				(H4K20me3) and thus regulates transcription and maintenance of genome integrity.				
TMEM150B	19	2	Transmembrane Protein 150B	Modulator of macroautophagy that causes accumulation of autophagosomes under basal conditions and enhances autophagic flux. Represses cell death and promotes long-term clonogenic survival of cells grown in the absence of glucose in a macroautophagy-independent manner. May have some role in extracellular matrix engulfment or growth factor receptor recycling, both of which can modulate cell survival.	No protein data 3.1 nTPM (RNA)	Plasma membrane	-	Age at menopause Ovulation Oestradiol measurement GFR
SLC22A6	11	2	Solute Carrier Family 22 Member 6	The protein encoded by this gene is involved in the sodium-dependent transport and excretion of organic anions, some of which are potentially toxic.	High protein expression in the proximal tubules 189.5 nTPM (RNA)	Plasma membrane	ERGIC3 SLC26A2 APPBP2 NEDD4L NEDD4	Uric acid measurement Blood protein measurement BMI-adjusted wait-hip ratio
SLC22A8	11	2	Solute Carrier Family 22 Member 8	This gene encodes a protein involved in the sodium-independent	Medium protein expression	Plasma membrane	NEDD4L NEDD4 EPHA5	Calcium measurement

				transport and excretion of organic anions, some of which are potentially toxic. The encoded protein is an integral membrane protein and appears to be localized to the basolateral membrane of the kidney.	in the tubules 322.8 nTPM (RNA)	and Extracellular	MTX3 SAMM50	Lipid measurement
IL11	19	3	Interleukin 11	The protein encoded by this gene is a member of the gp130 family of cytokines. These cytokines drive the assembly of multisubunit receptor complexes, all of which contain at least one molecule of the transmembrane signalling receptor IL6ST (gp130). This cytokine is shown to stimulate the T-cell-dependent development of immunoglobulin-producing B cells. It is also found to support the proliferation of hematopoietic stem cells and megakaryocyte progenitor cells.	No protein data 4.6 nTPM (RNA)	Extracellular	IL11RA S100P MAGEA11 IL6ST FOS	TMP190 measurement Protein measurement Age at menopause Body height Lean body mass Ovulation Osteoarthritis Age at menarche Periodontitis Blood lipid measurement Corpus callosum volume
TMEM107	17	1	Transmembrane protein 107	This gene encodes a transmembrane protein and component of the primary cilia transition zone. The encoded protein regulates ciliogenesis and ciliary protein composition.	Low protein expression in the tubules 14.3 nTPM (RNA)	Plasma membrane	ATP11C TMEM237 TMEME179B TMEM231 TMEM216	Meckel syndrome 13 Joubert syndrome 29 Orofaciodigital Syndrome

				Human fibroblasts expressing a mutant allele of this gene exhibit reduced numbers of cilia, altered cilia length, and impaired sonic hedgehog signalling.				Sex hormone measurement Body height Protein measurement Metabolic syndrome Potassium measurement Osteoarthritis
KLC1	14	2	Kinesin Light Chain 1	It associates with kinesin heavy chain through an N-terminal domain, and six tetratricopeptide repeat (TPR) motifs are thought to be involved in binding of cargos such as vesicles, mitochondria, and the Golgi complex. Thus, kinesin light chains function as adapter molecules and not motors per se.	Protein not detected 34.6 nTPM (RNA)	Cytosol and Cytoskeleton	KIF5B FIF5A KIF5C KLC4 KLC2	Uric acid measurement Gout GFR Blood pressure Blood protein measurement BMI
ST8SIA5	18	1	ST8 Alpha-N-Acetyl-Neuraminide Alpha-2, 8-Sialyltransferase 5	The protein encoded by this gene is a type II membrane protein that may be present in the Golgi apparatus. Among its related pathways are Synthesis of substrates in N-glycan biosythesis and Metabolism of proteins.	Protein not detected 1.0 nTPM (RNA)	Golgi apparatus	SELENON GALNT14 ADAM15 APPBP2 MST1R	Bone density Protein measurement Educational attainment
KIAA1586	6	1	KIAA1586	Enables SUMO ligase activity. Involved in protein sumoylation.	Medium expression	Nucleus	SPACA1 CLEC2D FBXO7	Coronary artery disease

					in glomeruli and tubules 3.5 nTPM (RNA)		AURKB KIF20A	Educational attainment
HSPBP1	19	1	HSPA (Heat Shock 70KDa) Binding Protein 1	Enables ubiquitin protein ligase binding activity. Involved in positive regulation of proteasomal ubiquitin-dependent protein catabolic process and positive regulation of protein ubiquitination. Predicted to be active in endoplasmic reticulum.	Medium expression in glomeruli and tubules 37.7 nTPM (RNA)	Endoplasmic Reticulum and Extracellular	HSPA1B FAU HSPA1A HSPA8 HSPA2	Protein level Age at menopause Ovulation Mean corpuscular volume Cholesterol measurement
CCZ1	7	1	CCZ1 Homolog, Vacuolar Protein Trafficking And Biogenesis Associated	Enables guanyl-nucleotide exchange factor activity. Predicted to be involved in vesicle-mediated transport.	Protein expression low in glomeruli and high and tubules 26.5 nTPM (RNA)	Cytosol and Endosome	VPS18 VPS41 VPS39 VPS16 VPS11	Platelet count Mean corpuscular volume Body weight Agre at menopause Triglyceride measurement Educational attainment

Table 33: Further Information on Prioritised Top Genes from Gene-burden Studies*

*[*Extended annotation for top-prioritised genes identified through the prioritised genes from gene-burden studies. It includes chromosomal location, recurrence across MAGMA studies, full gene names, known biological functions, kidney-specific and cellular expression patterns sourced from Human Protein Atlas data, key protein–protein interactions, and any associated human phenotypes sourced from GWAS Catalog, OMIM and Ensembl data, offering deeper biological insight into gene-level GWAS signals.*