GENETICS OF KIDNEY DISEASES

EDITED BY: Harvest F. Gu and Martin H. De Borst
PUBLISHED IN: Frontiers in Genetics

frontiers Research Topics

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

GENETICS OF KIDNEY DISEASES

Topic Editors:
Harvest F. Gu, China Pharmaceutical University, China
Martin H. De Borst, University Medical Center Groningen, Netherlands

Table of Contents

Editorial: Genetics of Kidney Diseases

Martin H. de Borst[1]* and Harvest F. Gu[2]

[1] *Division of Nephrology, Department of Internal Medicine, University Medical Center Groningen, University of Groningen, Groningen, Netherlands,* [2] *Center for Pathophysiology, School of Basic Medicine and Clinical Pharmacy, China Pharmaceutical University, Nanjing, China*

Editorial on the Research Topic

Genetics of Kidney Diseases

Worldwide, more than 850 million people have chronic kidney disease (CKD) (Jager et al., 2019). CKD can be caused by a variety of individual diseases including primary kidney diseases and systemic diseases such as diabetes. Diabetic kidney disease (DKD) worldwide develops in ∼40% of patients with diabetes and is the leading cause of CKD (Alicic et al., 2017). CKD is accompanied by an excessively elevated risk of premature mortality, particularly due to its predisposition to accelerated cardiovascular disease. In more than 10% of CKD patients the primary cause of disease is unknown (Groopman et al., 2018). It is crucial to increase the yield of the diagnostic work-up in these patients, since specific treatments exist for several kidney diseases.

Recent developments in genetics including next-generation sequencing (NGS) have strongly enhanced the diagnostic potential for patients with CKD. Moreover, novel genetic tools have contributed to growing insight in the etiology of both monogenic and multifactorial types of CKD. This is strongly illustrated by the rare disease atypical hemolytic uremic syndrome (aHUS), for which now numerous causal mutations in complement-related genes have been identified (Jokiranta, 2017). In this Research Topic, Schönauer et al. have demonstrated the case of a patient with aHUS, which was caused by a novel splice site variant in the complement factor H gene. Another paper in this Research Topic (Li P. et al.) also reports on a splice site mutation, in this case in the alpha-galactosidase A (*GLA*) gene. Several mutations in this gene have been previously found that lead to Fabry disease, a rare X-linked recessive hereditary systemic disorder of glycosphingolipid metabolism caused by totally or partially decreased activity of *GLA* (Simonetta et al., 2018). The aforementioned studies illustrate the power of genetic tools to reach a clinically relevant diagnosis driving the CKD phenotype (although the disorder may also affect other organ system as for example in Fabry disease), which may have therapeutic consequences as specific treatments exist both for aHUS and Fabry.

At the same time, clinical practice in nephrology is facing new challenges including optimal patient selection, implementation, counseling, and therapeutic consequences of the outcomes of NGS-based diagnostics. De Haan et al. discuss the advantages and limitations of NGS-based tools, and specifically focus on how these tools could improve diagnostic yield in patients with CKD of unknown or unclear etiology. In fact, further prospective cohort studies are needed to define the optimal positioning of NGS-based genetic testing in the diagnostic workup of CKD with unknown etiology.

For multifactorial diseases, which affect the majority of CKD patients, the impact of the "genomic revolution" has so far been relatively limited. At the same time these advances, among

others in terms of technology and infrastructure, have the potential to revolutionize precision medicine in the field of nephrology. This is illustrated by International Genetics & Translational Research in Transplantation Network (iGeneTRAiN): a multi-site consortium that encompasses >45 genetic studies with genome-wide genotyping from over 51,000 transplant samples, including genome-wide data from >30 kidney transplant cohorts ($n = 28,015$) (Fishman et al.). The potential of large-scale collaborations such as iGeneTrain to contribute to the understanding of disease etiologies has already been shown in several studies in CKD and kidney transplantation (Snoek et al., 2018; Reindl-Schwaighofer et al., 2019). Furthermore, genome-wide association studies (GWAS) performed in diverse populations can be useful to define the robustness of previously identified CKD risk loci such as the APOL1 across different ethnicities. This multi-ethnic study by Lin et al. also identified a novel risk locus for CKD near the NMT2 gene. Yet another approach is to study very specific high-risk populations. Thomson et al. performed a GWAS of limited size in a cohort of Australian Aboriginal Tiwi islanders, a population prone to develop CKD that had not been extensively studied before. The authors identified a variant near the CRIM1 gene that was associated with albuminuria, and remained significant after adjusting for multiple testing (Thomson et al.).

Another important application of genetic tools is in patient stratification, and in the identification of modifier genes that enhance susceptibility to morbidities such as CKD and DKD. This particularly applies to patients with type 1 and type 2 diabetes, the most common cause of end-stage kidney disease (ESRD) worldwide (Gu). Valoti et al. identified a variant in the complement factor H (CFH) gene that confer patients with type 2 diabetes at increased risk of microalbuminuria

and cardiovascular complications. Moreover, patients carrying the variant were less likely to benefit from ACEi therapy. Although CFH mutations are well-known to predispose to aHUS (Jokiranta, 2017), it was so far not known that CFH variants could enhance susceptibility to adverse kidney outcomes in patients with diabetes. Two variants in the beta-actin (ACTB) gene were identified in another study, which were associated with a higher risk of DKD in a large cohort of patients with type 2 diabetes (Li M. et al.). Thereby, ACTB, as a housekeeping gene, is suggested preferably not to be used as internal control for gene expression studies at the mRNA and protein levels in diabetes and DKD.

Despite efforts summarized above to find missing pieces of the genetic puzzle in CKD and DKD, it remains challenging to explain the complete heritability with currently available methods and datasets. Although studies beyond "conventional GWAS" have focused on telomeres, copy number variants, mitochondrial DNA and sex chromosomes, there remains considerable unexplained heritability in CKD and DKD (Cañadas-Garre et al.).

Furthermore, in addition to the multitude of information coming to us from genetic studies, we are facing new challenges as these data need to be interpreted in the context of other data dimensions including proteomics, single-cell RNA-sequencing, metabolomics, and the microbiome (Parsa et al., 2013; Hanna et al., 2017). Therefore, in order to profoundly impact clinical practice, comprehensive approaches are needed, particularly through the integration of multiple-omics data.

AUTHOR CONTRIBUTIONS

MB and HG co-edited the Research Topic, wrote, edited, and approved the final version of the Editorial.

REFERENCES

Alicic, R. Z., Rooney, M. T., and Tuttle, K. R. (2017). Diabetic kidney disease: challenges, progress, and possibilities. Clin. J. Am. Soc. Nephrol. 12, 2032–2045. doi: 10.2215/CJN.11491116

Groopman, E. E., Rasouly, H. M., and Gharavi, A. G. (2018). Genomic medicine for kidney disease. Nat. Rev. Nephrol. 14, 83–104. doi: 10.1038/nrneph.2017.167

Hanna, M. H., Dalla Gassa, A., Mayer, G., Zaza, G., Brophy, P. D., Gesualdo, L., et al. (2017). The nephrologist of tomorrow: towards a kidney-omic future. Pediatr. Nephrol. 32, 393–404. doi: 10.1007/s00467-016-3357-x

Jager, K. J., Kovesdy, C., Langham, R., Rosenberg, M., Jha, V., and Zoccali, C. (2019). A single number for advocacy and communication—worldwide more than 850 million individuals have kidney diseases. Nephrol. Dial. Transplant. 34, 1803–1805. doi: 10.1093/ndt/gfz174

Jokiranta, T. S. (2017). HUS and atypical HUS. Blood 129, 2847–2856. doi: 10.1182/blood-2016-11-709865

Parsa, A., Kao, W. H. L., Xie, D., Astor, B. C., Li, M., Hsu, C., et al. (2013). APOL1 risk variants, race, and progression of chronic kidney disease. N. Engl. J. Med. 369, 2183–2196. doi: 10.1056/NEJMoa1310345

Reindl-Schwaighofer, R., Heinzel, A., Kainz, A., van Setten, J., Jelencsics, K., Hu, K., et al. (2019). Contribution of non-HLA incompatibility between donor and recipient to kidney allograft survival: genome-wide analysis in a prospective cohort. Lancet 393, 910–917. doi: 10.1016/S0140-6736(18)32473-5

Simonetta, I., Tuttolomondo, A., Di Chiara, T., Miceli, S., Vogiatzis, D., Corpora, F., et al. (2018). Genetics and gene therapy of anderson-fabry disease. Curr. Gene Ther. 18, 96–106. doi: 10.2174/1566523218666180404161315

Snoek, R., Van Setten, J., Keating, B. J., Israni, A. K., Jacobson, P. A., Oetting, W. S., et al. (2018). NPHP1 (Nephrocystin-1) gene deletions cause adult-onset ESRD. J. Am. Soc. Nephrol. 29, 1772–1779. doi: 10.1681/ASN.2017111200

A Novel α-Galactosidase A Splicing Mutation Predisposes to Fabry Disease

Ping Li[1]*†, Lijuan Zhang[1]†, Na Zhao[1], Qiuhong Xiong[1], Yong-An Zhou[2], Changxin Wu[1]* and Han Xiao[1]*

[1] Institutes of Biomedical Sciences, Shanxi University, Taiyuan, China, [2] Bluttransfusion, The Second Hospital, Shanxi Medical University, Taiyuan, China

Fabry disease (FD) is a rare X-linked α-galactosidase A (*GLA*) deficiency, resulting in progressive lysosomal accumulation of globotriaosylceramide (Gb3) in a variety of cell types. Here, we report a novel splicing mutation (c.801 + 1G > A) that results in alternative splicing in *GLA* of a FD patient with variable phenotypic presentations of renal involvement. Sequencing of the RT-PCR products from the patient's blood sample reveals a 36-nucleotide (nt) insertion exists at the junction between exons 5 and 6 of the *GLA* cDNA. Splicing assay indicates that the mutated minigene produces an alternatively spliced transcript which causes a frameshift resulting in an early termination of protein expression. Immunofluorescence shows puncta in cytoplasm for mutated *GLA* whereas uniform staining small dots evenly distributed inside cytoplasm for wild type *GLA* in transfected HeLa cells. The increased senescence and decreased GLA enzyme activity suggest that the abnormalities might be due to the altered localization which further might result from the lack of the C-terminal end of GLA. Our study reveals the pathogenesis of splicing mutation c.801 + 1G > A to FD and provides scientific foundation for accurate diagnosis and precise medical intervention for FD.

Keywords: Fabry disease, GLA, splicing mutation, c.801 + 1G > A, novel mutation

INTRODUCTION

Fabry disease (OMIM #301500, FD) is a rare X-linked recessive hereditary systemic disorder of glycosphingolipid metabolism, caused by total or partial decreased activity of alpha-galactosidase A (a-Gal or GLA, EC 3.2.1.22; UniProt P06280) (Brady et al., 1967; Kint, 1970) and results in lysosomal accumulations of globotriaosylceramide (Gb3), and other neutral glycosphingolipids in various cells and tissues including skin, eye, kidney, heart, brain, and peripheral nervous system (Zarate and Hopkin, 2008).

Classical FD is a complex multisystemic disorder with prominent features like neuropathic pain, exercise intolerance, gastrointestinal abnormalities, hyperhidrosis, corneal changes, angiokeratomas, progressive renal and cardiac deterioration, and a reduced life expectancy (Zarate and Hopkin, 2008). The disease may also present in milder forms involving primarily the heart or the kidneys (von Scheidt et al., 1991; Nakao et al., 2003). The milder forms of the disease have a later onset and are usually associated with some residual levels of GLA enzyme activities. The ubiquitously expressed *GLA* gene (located at position Xq22.1, OMIM 300644, RefSeq X14448, HGNC 4296; NCBI reference sequence NM_000169.2) contains 7 exons encoding the 429 amino acid GLA polypeptide, including an N-terminal 31-residue signal peptide.

The vast majority of human genes are discontinuous and contain more than one exon. Following transcription, genes are expressed as pre-mRNAs. Pre-mRNA splicing is a nuclear process, during which intronic sequences are removed from eukaryotic pre-mRNA transcripts and exons are joined together to produce a functional mRNA molecule. In order for the spliceosome to carry out the splicing reaction, it must first recognize canonical splice sites present at exon-intron junctions, including the 5′ and 3′ splice sites at the 5′ and 3′ termini of the intron, respectively, and the branchpoint sequence (BPS) locates a short distance upstream of the 3′ splice site (Hastings and Krainer, 2001). Most introns have a 5′ splice site beginning with GT and a 3′ splice site ending with AG, and some introns have distinct splice-site consensus sequences and exhibit either AT-AC termini or GT-AG termini (Patel and Steitz, 2003).

So far, hundreds of mutations in *GLA* that causing FD was identified (Human Gene Mutation Database[1] and Fabry mutants list[2]). Some GLA missense variants (p.P60L, p.E66Q, p.R118C, p.A143T and p.I198T) have been described as causative for FD when first discovered in subsequent clinical, functional and population studies (van der Tol et al., 2014; Ferreira et al., 2015; Smid et al., 2015; Lenders et al., 2016). However, other GLA missense variants have been only reported in clinical case reports which lack functional study. Besides the pathogenetic variants, several intronic variants and one missense variant (p.D313Y) which cause false positive in the enzyme assay through a pseudodeficiency have been described. Nevertheless, this missense variant (p.D313Y) is identified as non-pathogenic (Froissart et al., 2003; Yasuda et al., 2003; Ferri et al., 2012; Ferreira et al., 2015).

Many *GLA* splicing mutations have also been described in case reports, but only the deep intronic mutation c.639 + 919 G > A was well studied (Ishii et al., 2002; Chien et al., 2016; Palhais et al., 2016; Chang et al., 2017; Chiang et al., 2017). In the present study, we identified a novel splicing mutation in a FD patient in which the first nucleotide of *GLA* intron 5 is changed from G to A. This mutation, c.801 + 1G > A, alters the 5′ splice site recognition sequence that is crucial for splicing. The aim of this study was to characterize the molecular effect and mechanism of the *GLA* GT-AG intron mutation that causes FD.

MATERIALS AND METHODS

Patients
This study was approved by the local Ethics Committees and written informed consent was obtained from all patients participating in the study. Four patients from four unrelated Chinese families were recruited from Fabry Disease patient organization in China, and the patients were geographically localized in Shanxi, Anhui, Jilin, and Shanghai, respectively. Patients' medical records were reviewed and evaluated, and clinical and physical examinations were performed. Percutaneous renal biopsies were done by nephrologists in the hospital.

Based on all medical records, clinical presentation, data given by examinations and pathologic findings, the patients were diagnosed as FD.

Sequencing Analysis
Genomic DNA was extracted from peripheral blood samples using the DNeasy Blood & Tissue Kit (Qiagen, Cat NO. 69506) according to the manufacturer's instructions. All coding regions and exon–intron splice junctions of the *GLA* gene were analyzed using PCR amplification in combination with Sanger sequencing using the primers described previously (Shabbeer et al., 2005). PCR products were purified using the SanPrep Column DNA Gel Extraction Kit (Sangon Biotech, Cat. No. B518131).

Cell Culture
HEK293T and HeLa cells were maintained in DMEM supplemented with 10% (v/v) FBS, 100 U/ml penicillin, and 100 mg/ml streptomycin at 37°C and 5% CO_2. Cells were transfected by the Polyetherimide (PEI) (PolyScience, Cat. No. 23966-2) according to the manufacturer's instructions. Transfected cells were incubated for 24–48 h post-transfection.

RT-PCR and qRT-PCR Analysis
To evaluate the transcript variants of *GLA* by RT-PCR and qRT-PCR in the blood cells from patient and three healthy volunteers, total RNA was extracted with the TRIzol following the instructions of the supplier (Invitrogen, Cat. No. 15596-018). First-strand cDNA synthesis was performed using the M-MLV reverse transcriptase RNase H Minus-kit from Promega. The primer pair for qRT-PCR of *GLA* were used: Fw: 5′-GTTGGAATGACCCAGATATGTTA-3′ and Rv: 5′-CTGATTGATGGCAATTACGTCC-3′. For normalization, the expression of GAPDH (Forward: CGGAGTCAACGGATTTGGTCGTAT; Reverse: AGCCTTCTCCATGGTGGTGAAGAC) was used.

The Minigene Constructs
Minigene constructs encompassing exon 5, intron 5 and exon 6 were amplified using the primer pair of *GLA* E5-in 5-E6 Fw: 5′-GCGCTCGAG CCCAATTAT ACAGAAATCCGACAG-3′ and *GLA* E5-in 5-E6 Rv 5′-GCGG AATTCCTGTCTAAGCTGGTACCCTTG-3′. The amplified minigene products were cloned into pcDNA 3.1(−) and pEGFP-C3 cloning vector at the Xho I and EcoR I sites. The complete sequence of the minigene constructs was verified by sequencing. Transient transfection of minigene constructs in HEK293T cells were performed with polyetherimide (PEI) (PolyScience, Cat. No. 23966-2), and the minigene RT-PCR was amplified with a set of primers: 5′-TGCTGACATTGATGATTCCTGG-3′ and 5′-GTTACTTGCTGATTCCAGCTG-3′.

Western Blotting
Cells were grown to about 90% confluency in 6 well plates, washed twice with ice cold PBS and resuspended in modified radio-immunoprecipitation (RIPA) lysis buffer (50 mM Tris/HCl, pH 7.5, 150 mM NaCl, 1% NP-40, 0.5%

---

[1]http://www.hgmd.org
[2]http://fabry-database.org/mutants/

Na-desoxycholate, 1 mM dithiothreitol, 1 mM benzamidine, 1 mM PMSF). Cell suspensions were passed through a 0.45 μm needle 10 times and incubated for 15 min on ice, followed by sonication. The lysates were cleared by centrifugation at 10,000 rpm for 10 min at 4°C. The samples were resuspended in 5× SDS sample buffer and heated at 98°C for 5 to 10 min. For western blot analysis, proteins were resolved in 12 or 15% SDS polyacrylamide gels and transferred to NC membrane (GE) using the wet blot transfer over 3 h (I = 300 mA). The blots were probed with primary antibodies, then washed with TBST (10 mM Tris/HCl, pH 8.0, 150 mM NaCl, and 0.05% Tween-20) three times and incubated with horseradish peroxidase-conjugated secondary antibody (Thermo). The blots were washed again with TBST and signals were visualized using chemiluminescence (ECL) system (GE, Cat. NO. RPN2232). The following antibodies were used: GFP mouse monoclonal antibody (Proteintech, Cat. No. 66002-1-Ig), GAPDH mouse monoclonal antibody (Proteintech, Cat. No. 60004-1-Ig).

Analysis of Evolutionary Conservation of Amino Acid Residues and Structure Prediction of the Mutant Protein

Evolutionary conservation of amino acid residue alteration was analyzed by comparing across different species. The homology modeling programs Swiss-Model[3] was used to develop an appropriate model to mimic the effects of the mutated region. The structures were displayed by PDB-Viewer software.

GLA Enzyme Activity Assay

HEK293T cells were cultured and transfected with plasmids containing either wild type or mutant *GLA*, respectively. At 48 h post-transfection, cells were collected after centrifugation at 1500 rpm using bench top centrifuge. Every 5 million cells were resuspended with 1 mL extraction buffer followed by sonication. The lysates were cleared by centrifugation at 15 000 G for 10 min at 4°C. The GLA enzyme activity of the supernatants were measured according to the manufacturer's instructions (Solarbio, Cat. NO. BC2575).

In details, the samples were set in a 96-well plate, and the assay reagents were added and mix thoroughly according to the instructions given by the supplier. The absorbance A at 400 nm was measured and the ΔA = A measurement-A control was calculated. For α-GAL activity calculation, a standard curve was established based on the absorbance (x) and concentration (y, nmol/ml) of the standard sample, and ΔA was taken into the standard curve to calculate the amount of product (nmol/ml) produced by the samples. For definition of enzyme activity unit, 1 nmol of p-nitrophenol per hour per 10,000 cells is defined as an enzyme activity unit, and α-GAL activity (nmol/h/104 cell) = $(y \times V1) \div (500 \times V \div V2) \div T = 0.028 \times y$ (V1: total volume of the reaction system, 0.07 mL; V: used sample volume in the reaction system, 0.01 mL; V2: volume of added extract solution, 1 mL; 500: total number of cells, 5 million; T: reaction time, 0.5 h).

---

[3]http://swissmodel.expasy.org

Immunofluorescence Analysis

Cells grown on coverslips were fixed in 4% paraformaldehyde in PBS for 10 min followed by permeabilization with 0.5% Triton X-100 for 5 min for DAPI (Sigma) staining and mounted in gelvatol. Slides were imaged on a DeltaVision Image Restoration Microscope with a ×100 objective (DeltaVision Elite, GE) (Li and Noegel, 2015).

Senescence-Associated β-Galactosidase Assays

HEK293T cells were seeded in 6 well plates at 0.8 million per well the day before transfection, and cells were transfected with plasmids containing either wild type or mutant *GLA*, respectively. At 18 h post-transfection, the transfected cells were washed with PBS and fixed with 3% PFA (5 min, RT) (Li et al., 2014). Cells were washed twice with PBS and incubated at 37°C with freshly prepared senescence-associated-Gal (SA-Gal) staining solution (Solarbio, Cat. No. BC2580) for β-galactosidase assays. After incubation for 4–8 h, staining was checked and visualized under bright field microscopy at 200× magnification.

Statistical Analysis

All data are presented as the mean ± SD from at least three separate experiments. The *p*-values were determined by two-tailed Student's *t*-test. $P < 0.05$ was considered as being significant.

RESULTS

Family Pedigree/Patient Information

All patients included in this study with clinical manifestations of FD were recruited from four unrelated Chinese families. The control blood samples were collected from three healthy volunteers. The blood samples were obtained and Sanger sequencing of genomic DNA isolated from the samples were performed for *GLA* gene. The blood samples of two affected males (III-3 and IV-2) and one female (III-6) from family-1 were obtained. Sanger sequencing revealed that these two males contains the hemizygous *GLA* mutation and the female harbors the heterozygous *GLA* mutation (c.119C > A, p. Pro40His, P40H) (**Figure 1**). Two affected females (II-1 and II-3) from family-2 were tested and the sequencing results showed that these two females harbor the heterozygous *GLA* mutation (c.101A > G, p. Asn34Ser, N34S) (**Figure 1**). The blood samples of one affected male (III-2) and his mother (II-3) from family-3 were analyzed. Sequencing results indicated that patient III-2 inherited the described *GLA* mutation (c.680G > C, p. Arg227Pro, R227P) from his mother (II-3) (**Figure 1**). In family-4, a novel *GLA* mutation c.801 + 1G > A (p.L268IfsX3) was detected (**Figure 1**). This patient was diagnosed as Fabry nephropathy using electron microscope and histopathology of renal biopsy (data not shown). In this four-generation Chinese family, three individuals were affected but we just got the blood sample from patient III-4 and confirmed the novel splicing mutation c.801 + 1G > A by Sanger sequencing.
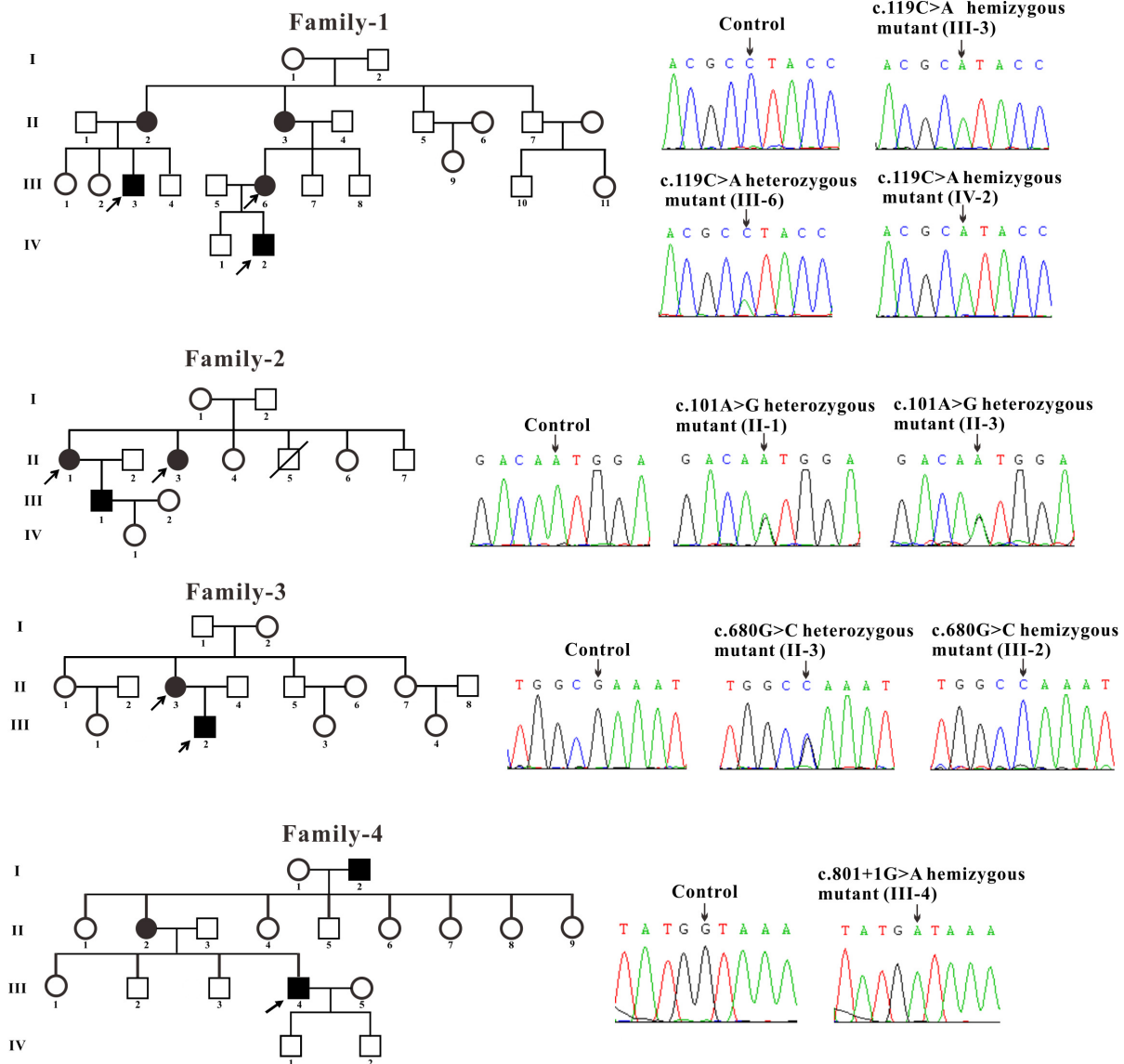
FIGURE 1 | Pedigree and sequencing of patients from four unrelated Chinese families. The patient involved in this study is pointed by an arrow. Sanger sequencing analysis performed on the genomic DNA from indicated patients. The gene variation is shown by black arrow.

Three families were carrying the previously described *GLA* mutations: c.119C > A (p. Pro40His, P40H), c.101A > G (p. Asn34Ser, N34S) and c.680G > C (p. Arg227Pro, R227P) (Eng et al., 1993; Meng et al., 2010; Zizzo et al., 2016). A novel *GLA* mutation c.801 + 1G > A (p.L268IfsX3) was detected from the fourth family and subjected to further study.

Splicing Defect in *GLA* c.801 + 1G > A FD Patient

In the present study, we focused on the novel c.801 + 1G > A mutation which is located at the boundary between exon 5 and intron 5 and affects the first nucleotide of intron 5. The flanking intronic regions are always considered to be related

with alternative splicing (Patel and Steitz, 2003). To further characterize the abnormal splicing, RT-PCR for the mentioned region that includes exons 5 and 6 was performed. Two bands were visualized on gel electrophoresis: ∼160 bp fragment, the expected wild type and ∼200 bp, an extra larger fragment (**Figure 2A**), suggesting that the abnormal DNA fragment was generated by a rare splicing event within the *GLA* gene. Sequencing of the RT-PCR product revealed a 36-nucleotide (nt) insertion at the junction between exons 5 and 6 of the *GLA* cDNA (**Figure 2B**), consistent with the observed size of the RT-PCR products from the patient's blood (**Figure 2A**). This insertion corresponded to an intronic sequence which is identified as sequence of 5′-end of intron 5. This in-frame insertion caused a premature termination codon TGA at the 12th nucleotide

FIGURE 2 | Alternative splicing of *GLA* (c.801+1G > A). **(A)** mRNA from the patient blood was extracted and amplified by RT-PCR for alternative splicing of *GLA* (c.801+1G > A). RT-PCR of the region that includes exons 5 and 6 was performed and the PCR products were separated by gel electrophoresis. **(B)** Sanger sequencing analysis for the alternative splicing products. The sequence of the upper band reveals a 36-nucleotide (nt) insertion at the junction between exons 5 and 6 of the *GLA* cDNA compare to wild type. **(C)** Schematic representation of *GLA*. Uppercase letters indicate the exonic sequences, whereas lowercase letters indicate the intronic sequences. The encoded amino acids are depicted in single-letter code. The changed nucleotide in mutant *GLA* cDNA is labeled by underscore. **(D)** qRT-PCR analysis performed on total RNA obtained from blood samples of patient III4 and three healthy volunteers individuals. Levels were normalized to the amount of GAPDH. Data represent the mean ± SE of three independent measurements performed in triplicate (**p < 0.01).

downstream of the *GLA* exon 5 (**Figure 2C**), giving the predicted product of the mutant *GLA* mRNA was a truncated protein of 270 amino acid residues.

To identify the mRNA expression level of *GLA* in patient samples, qRT-PCR was performed. Total RNA was isolated from the blood samples of patient III-4 and three healthy volunteers, and qRT-PCR was normalized to that of GAPDH. The results clearly showed that the level of *GLA* mRNA was reduced to one third in the samples from patient compared to control (**Figure 2D**). Furthermore, the mRNA and protein expression were also decreased in HEK293T cells transfected with mutant *GLA* compared with that transfected with equal amount wild type *GLA* or vector control (**Supplementary Figure S1**). These results suggest that the mRNA of mutant *GLA* is not stable in the cells thus lead to a lower expression of the protein, and imply that an insufficient normal GLA in the patient.

Corroboration of *GLA* c.801+1G > A as the Main Determinant of Alternative Splicing in Minigene Splicing Assay

To investigate whether the alternative splicing was due to the c.801 + 1G > A mutation, we constructed a minigene containing entire exon 5, exon 6, and intron 5 sequence with G or A at c.801 + 1 of *GLA* (**Figure 3A**). After transfection in HEK293T cells, total mRNAs were isolated and RT-PCR was performed. RT-PCR products were separated by electrophoresis analysis and isoforms were identified by Sanger sequencing. In both untransfected and minigene transfected HEK293T cells, two bands with the same size were shown which was identified as exon 5 + exon 6, but an extra band was detected in mutant minigene transfected cells, same size as the band given by mRNA isolated from the patient. As in the alternative transcript analysis, the wild type construct revealed one normal size band of 159 bp which is corresponding to exon 5 + exon 6. The c.801 + 1G > A substitution resulted in one extra band of 195 bp showing up, which is bigger and even stronger than exon 5 + exon 6. Sanger sequencing showed the extra band contains 36 nucleotides between exon 5 and exon 6 (**Figure 3B**).

To further investigate the protein expression for the minigenes, western blot analysis was performed. Proteins were extracted from the transfected HEK293T cells and resolved in 15% SDS polyacrylamide gels. After transferred to NC membrane, the blots were probed with GFP monoclonal antibodies. In pEGFP-C3 vector transfected HEK293T cells, a strong band was observed with the molecular weight about 27 kD which represents the GFP protein alone (**Figure 3C**). In pEGFP-*GLA*-minigene wild type transfected cells, a band with molecular weight at ∼38 kD was observed which was identified as the protein products of exon 5 + exon 6 fusion with GFP (**Figure 3C**). However, in pEGFP-*GLA*-minigene mutant transfected cells, one extra band at ∼31 kD molecular weight was shown compared to wild type *GLA* minigene transfected cells, suggesting that the c.801 + 1G > A substitution caused a frameshift resulting in early termination of protein expression which is consistent with the results of alternative transcript analysis (**Figures 2B,C**, **3C**). Both

mRNA and protein expressions confirmed the single nucleotide substitution (c.801 + 1G > A) is the main determinant of alternative splicing.

Abnormalities of Localization and Senescence in Mutant GLA Transfected Cells

The human GLA structure is a homodimer with each monomer containing a $(\beta/\alpha)_8$ domain containing the active site and a C-terminal domain containing eight antiparallel β strands on two sheets in a β sandwich (Garman and Garboczi, 2004). After removal of the 31-residue signal sequence, the first domain extends from residues 32 to 330 and contains the active site formed by the C-terminal ends of the β strands at the center of a barrel, a typical location for the active site in $(\beta/\alpha)_8$ domains. The second domain comprised of residues 331–429, packs against the first with an extensive interface, burying 2500 Å^2 of surface area within one monomer (Garman and Garboczi, 2004).

Through sequence alignment, we found the c.801 + 1G > A mutation caused a translational frameshift, and the premature stop codon appeared at codon 272 (p.L268IfsX3), which caused a 161-amino-acid-residue change and partial loss of C-terminal domain (**Figures 4A,B**). Evolutionary conservation analysis of amino acid residues showed that these impaired amino acid residues in the truncated protein were most highly evolutionary conserved among GLA proteins from different species, indicating the mutation was likely causative mutation predisposing to FD (**Figure 4B**). This mutation results in a truncated protein lacking the C-terminal end containing part of the first domain (residues 269–330) and the whole second domain (residues 331–429) (**Figure 4C**).

To identify whether the mutant GLA has altered localization inside the transfected cells, we introduced the mutation into wild type GFP-tagged full length GLA and expressed the corresponding proteins in HeLa cells. In GFP alone transfected cells, GFP signal was detected in both nucleus and cytoplasm. In GFP-GLA-WT transfected cells, the GLA fusions were uniformly expressed in the cytoplasm. Surprisingly, puncta structures were observed in the cytoplasm of GFP-GLA-MT transfected cells (**Figure 4D**). The altered localization and structure might result from lack of the GLA C-terminal end.

To understand whether the GLA mutation affects the phenotype of transfected cells, we examined senescence associated β-galactosidase (SA-β-Gal) in GFP alone, GFP-GLA-WT and GFP-GLA-MT transfected cells. In GFP alone and GFP-GLA-WT transfected cells, less than 10% of the cells were β-galactosidase positive, however, in GFP-GLA-MT transfected cells, almost all the cells were β-galactosidase positive (**Figure 4E**). Our results indicate that *GLA* c.801 + 1G > A (p.L268IfsX3) mutation results in the mis-localization of GLA protein and increased senescence of transfected cells.

c.801 + 1G > A Mutation Resulted in Reduced Enzyme Activity

The c.801 + 1G > A mutation causes an in-frame insertion with 36-nucleotide (nt) which contains a premature termination TGA

FIGURE 3 | Minigene constructs for splicing pattern investigation. **(A)** Schematic representation of the *GLA* minigene construct, extending from exon 5 to exon 6. The nucleotide sequences of the wild type (WT) and mutant (MT) minigenes are shown below. The changed nucleotide in intro 5 of both wild type and mutant *GLA* minigene are labeled by underscore. **(B)** Minigene splicing assay. *GLA* minigenes harboring the wild type or mutant *GLA* pseudoexon were transiently transfected into HEK293T cells. After RNA isolation the splicing products were analyzed by RT-PCR. The lower bands represent correctly spliced exons, whereas the higher band represent the *GLA* pseudoexon inserted 36 nt from intro 5 between minigene exons 5 and 6. **(C)** Western blot analysis for the expression of the wild type and mutant pEGFP-*GLA*-minigenes. Whole cell lysates were separated by SDS–PAGE (12% acrylamide). GFP monoclonal antibodies were used.

at the 12th nucleotide downstream from exon 5 (**Figure 2B**). This stop codon results in a truncated protein lacking 163 amino acids at the C terminal but with 3 amino acids insertion from the sequence of intron 5 (**Figure 2C**). To examine whether the truncated protein exhibits any residual enzyme activity, an expression construct pEGFP-GLA full length wild type and pEGFP-GLA mutant were prepared and expressed in HEK293T cells. The fluorescence microscope visualization and western blot analysis were used, and the similar transfection efficiency of those transfectants was observed, GAPDH was used as loading control (**Figures 5A,B**). The

full length and mutant protein of 47 and 30 kD were detected by western blot analysis, respectively (**Figure 5B**). The GLA enzyme activity of GFP alone, GFP-GLA-WT and GFP-GLA-MT transfected cells were detected using the kit from Solarbio (Cat NO. BC2575). Before comparing the enzyme activities for GFP-GLA-WT and GFP-GLA-MT transfected cells, the activity of GFP alone transfected cells was subtracted as endogenous enzyme activity according to manufacturer's guidelines. The enzyme activity of HEK293T cells transfected with mutant GLA construct was significantly lower with a relative GLA enzyme activity down to 20% compared to
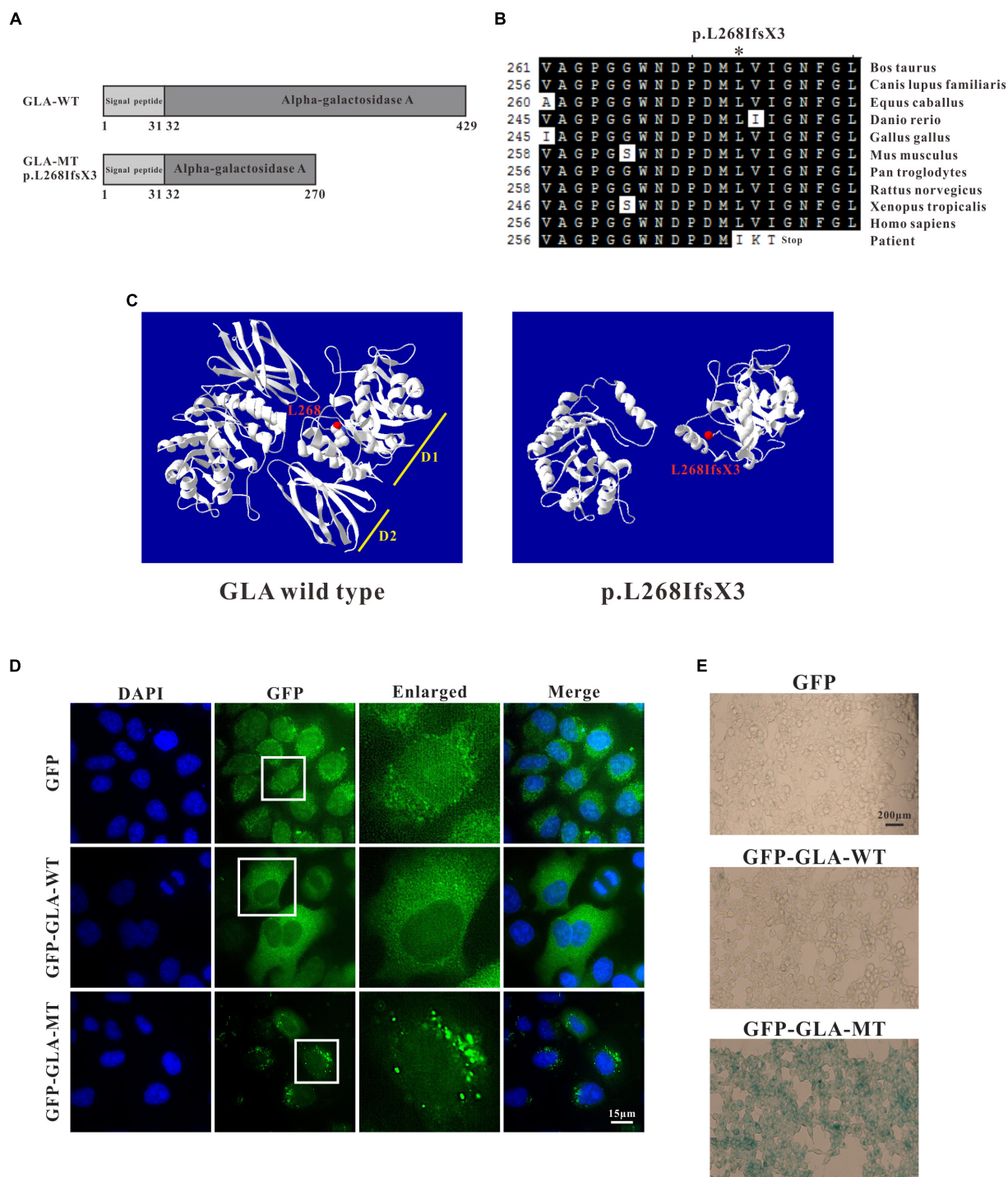
FIGURE 4 | Analysis of GLA mutation. **(A)** Schematics of the secondary structure and functional domains of the GLA protein. **(B)** Evolutionary conservation of amino acid residues altered by c.801 + 1G > A (p.L268IfsX3) across different species. NCBI accession numbers are: Bos Taurus: NP_001179665; Canis lupus familiaris: XP_538109; Equus caballus: XP_001492699; Danio rerio: NP_001006103; Gallus gallus: XP_420183; Mus musculus: NP_038491; Pan troglodytes: XP_003954083; Rattus norvegicus: NP_001102290; Xenopus tropicalis: NP_001120606; Homo sapiens: NP_000160. **(C)** The mutant protein p.L268IfsX3 were predicted to result in the loss of the C-terminal domain by Swiss-Model online software compared to the wild type. Ribbon representation of the human GLA and map of the studied variant localization obtained by homology modeling analysis. The WT and MT monomers are shown in white. The altered amino acid is shown as red ball. Each monomer is composed of two domains as pointed out: Domain 1 (D1) with the catalytic site, and Domain 2 (D2). **(D)** HEK293T were transfected with GFP alone, GFP-GLA-WT, and GFP-GLA-MT plasmids and the localization of wild type and mutant GLA were studied by immunofluorescence. Bar: 15 μm. **(E)** Visualization of senescence associated β-galactosidase show differences between wild type and mutant GLA transfected cells. Examination for staining was done after incubation for 4 to 8 h under bright field microscopy at 200× magnification. Bar: 200 μm.
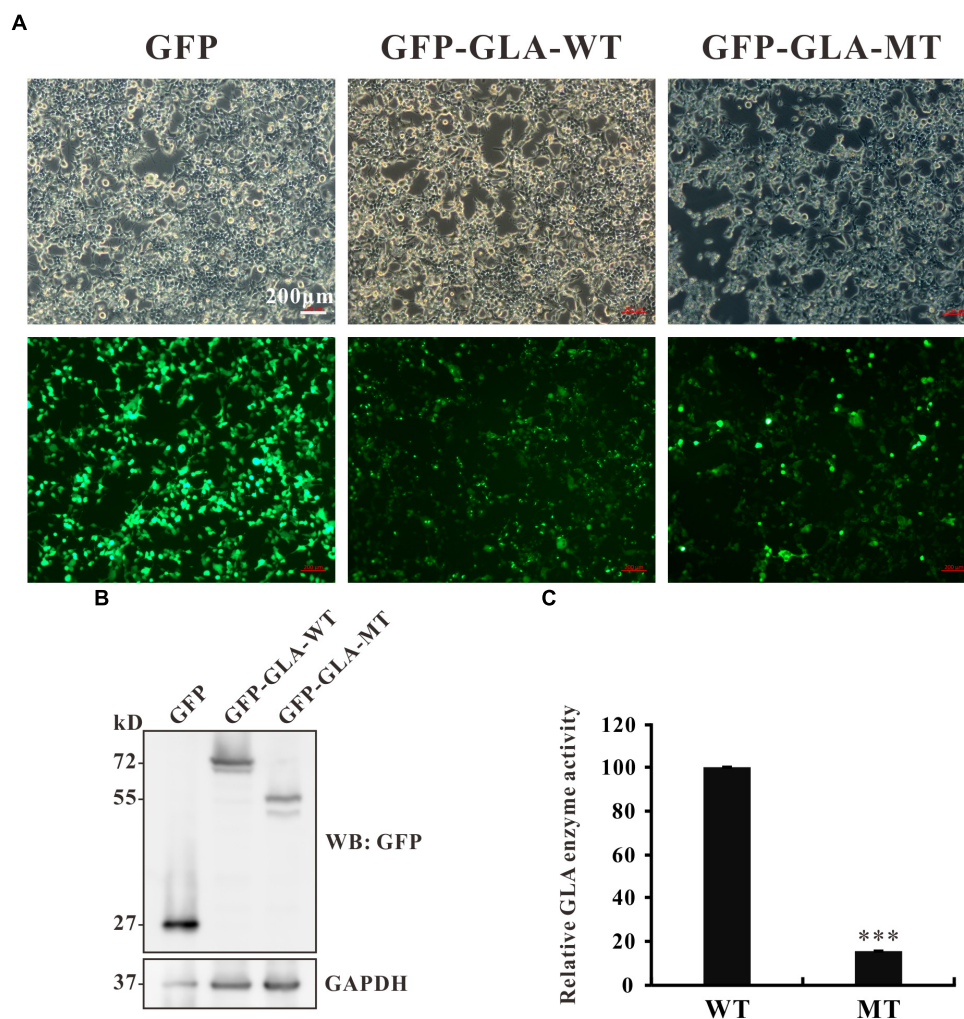
FIGURE 5 | Enzyme activity of the mutant GLA. **(A)** Images taking from bright and fluorescence fields were used to show the similar transfection efficiency of GFP, GFP-GLA-WT, or GFP-GLA-MT plasmids in HEK293T cells. **(B)** Western blot analysis for HEK293T lysates transfected with GFP, GFP-GLA-WT, or GFP-GLA-MT plasmids. GAPDH was used as loading control. **(C)** The result of enzyme activity assay from HEK293T cells transfected with wild type or mutant GLA plasmids. Data were presented as the mean ± SD from three independent experiments; ***$p < 0.001$.

that of GLA wild type transfected cells (**Figure 5C**). This result indicated that the C-terminal truncated protein had no enzyme activity.

DISCUSSION

Fabry disease (FD) is a rare X-linked recessive hereditary systemic disorder with nearly complete penetrance in male patients with mutations in the *GLA* gene. Diagnosis in male patients is made by an enzymatic assay measuring GLA activity. The detection of residual GLA enzyme activity is a time consuming and complicate diagnosis, either in females or in males. It has been reported that higher residual enzyme activities can lead to milder phenotype (Schaefer et al., 2005). Identifying of a novel *GLA* variant can leave confirmation of FD uncertain, particularly in patients harboring missense variants and exhibiting mild manifestations or in female probands (Froissart et al., 2003; Yasuda et al., 2003; Ferreira et al., 2015). However, the relationship between clinical manifestations, biochemical abnormalities, genetic mutations has not been clearly established.

In this study, we recruited the patients from four unrelated Chinese families and found three described *GLA* missense mutations: c.119C > A (p. Pro40His, P40H), c.101A > G (p. Asn34Ser, N34S), and c.680G > C (p. Arg227Pro, R227P) (Eng et al., 1993; Meng et al., 2010; Zizzo et al., 2016) from the first three families and a novel splicing mutation c.801 + 1G > A (p.L268IfsX3) from the fourth family which is subjected to further study. This novel splicing mutation c.801 + 1G > A (p.L268IfsX3) was detected between exon 5 and exon 6 from the blood sample of patient III-4 who exhibits classical renal Fabry features. This mutation is located

at the boundary of exon 5 and intron 5. The flanking intronic regions are always considered to be related to alternative splicing (Patel and Steitz, 2003). Sequencing of the RT-PCR products revealed a 36-nucleotide (nt) insertion at the junction between exons 5 and 6 of the *GLA* cDNA which corresponded to the intronic sequence of intron 5 (**Figure 2B**). The first intron sequences ever characterized revealed highly conserved dinucleotides GT and AG at the 5′ and 3′ termini, respectively. The nucleotide G has very high frequency of occurrence at the 3′ termini of exon which means the boundary sequence GGT are very important for the recognition by spliceosome (Patel and Steitz, 2003). In our study, we found nucleotide G and GT are at position 36 and 37–38 of intron 5, respectively. And this GGT appears for the first time after the splicing site. Once the GT is altered at 5′ termini of intron, the spliceosome will go to next GGT which could explain how the 36-nucleotide remains.

This in-frame insertion caused a premature termination TGA at the 12th nucleotide downstream from exon 5 which results in a truncated GLA protein of 270 amino acid residues (**Figure 2C**). Mutations that generate premature termination codons (PTCs) can reduce the stability of mRNA via nonsense-mediated decay (NMD) (Maquat, 1995). In our case, the level of the patient *GLA* mRNA containing a PTC was remarkably reduced to one third compared to that of healthy volunteers (**Figure 2D**), showing that the *GLA* mRNAs were might subject to NMD.

To further confirm whether the alternative splicing was caused by the c.801 + 1G > A mutation, minigene was constructed to mimic the splicing process *in vitro*. Consistent with our result of RT-PCR products from patient blood, the mutated minigene produced an alternative splicing product which contains a 36-nucleotide insertion from intron 5 as well (**Figure 3B**). Furthermore, the protein expressed in HEK293T cells by the mutated minigene showed two bands (38 and 31 kD) compared to the one band (38 kD) from wild type minigene (**Figure 3C**). These results suggest that the single nucleotide substitution (c.801 + 1G > A) is the only causative factor of alternative splicing.

The structure of GLA is a homodimer with each monomer containing a $(\beta/\alpha)_8$ domain with the active site and an antiparallel β domain (Garman and Garboczi, 2004). The c.801 + 1G > A (p.L268IfsX3) mutation results in a truncated protein lacks the C-terminal end (**Figures 4A,C**). Evolutionary conservation analysis of amino acid residues showed that these lost amino acid residues were most highly evolutionary conserved among GLA proteins from different species, indicating the mutation was likely pathological (**Figure 4B**). Immunofluorescence study revealed that the overexpressed GFP- GLA-WT were uniformly distributed in the cytoplasm. Strikingly, overexpressed GFP-GLA-MT formed puncta in the cytoplasm of transfected HEK293T cells. Each GLA monomer is composed of two domains: domain 1 contains the catalytic site, and domain 2 packing against domain 1 with an extensive interface (Garman and Garboczi, 2004). The formation of mutant GLA puncta might due to the instability caused by lack of the C-terminal domain (**Figure 4D**). Furthermore, the overexpressed GFP-GLA-MT increased the senescence of the transfected cells that might be caused from the abnormal puncta structures.

The FD mutations are partitioned into two classes: those that locally perturb the active site of the enzyme and those that adversely affect the folding of the protein. Residues from seven loops in domain 1 form the active site: β1-a1, β2-a2, β3-a3, β4-a4, β5-a5, β6-a6, and β7-a7. The active site is formed by the side-chains of residues W47, D92, D93, Y134, C142, K168, D170, E203, L206, Y207, R227, D231, D266, and M267, with C172 making a disulfide bond to C142 (Garman and Garboczi, 2004). The c.801 + 1G > A (p.L268IfsX3) mutation described in this article is next to the catalytic site M267, therefore, the truncated protein probably loss the function of catalytic. To confirm our hypothesis, GLA enzyme activity was measured. The cells transfected with mutant *GLA* construct showed significantly lower GLA enzyme activity compared with wild type *GLA* transfected cells (**Figure 5C**). The reduced activity might due to the impaired binding with its substrate or the loss of its enzyme activity. To address this question, further studies are required.

In summary, we report a novel intronic mutation c.801 + 1G > A causes a remarkable increase in the alternatively spliced GLA transcript and, consequently, results in the renal phenotype of FD. Our study confirms that c.801+1G > A is a Fabry-causative mutation and data in this study enrich Fabry mutation database and provide a FD causative mutation for accurate molecular diagnosis as well as scientific information.

AUTHOR CONTRIBUTIONS

PL designed the study and wrote the manuscript. LZ performed the practical work and was assisted by NZ. QX and Y-AZ analyzed the patients' data. CW and HX conceived the study and edited the manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019. 00060/full#supplementary-material

REFERENCES

Brady, R. O., Gal, A. E., Bradley, R. M., Martensson, E., Warshaw, A. L., and Laster, L. (1967). Enzymatic defect in Fabry's disease. Ceramidetrihexosidase deficiency. *New Engl. J. Med.* 276, 1163–1167. doi: 10.1056/NEJM196705252762101

Chang, W. H., Niu, D. M., Lu, C. Y., Lin, S. Y., Liu, T. C., and Chang, J. G. (2017). Modulation the alternative splicing of GLA (IVS4+919G > A) in Fabry disease. *PLoS One* 12:e0175929. doi: 10.1371/journal.pone.0175929

Chiang, H. L., Wang, N. H., Song, I. W., Chang, C. P., Wen, M. S., Chien, Y. H., et al. (2017). Genetic epidemiological study doesn't support GLA IVS4+919G > a variant is a significant mutation in Fabry disease. *Mol. Genet. Metab.* 121, 22–27. doi: 10.1016/j.ymgme.2017.03.005

Chien, Y., Chien, C. S., Chiang, H. C., Huang, W. L., Chou, S. J., Chang, W. C., et al. (2016). Interleukin-18 deteriorates Fabry cardiomyopathy and contributes to the development of left ventricular hypertrophy in Fabry patients with GLA IVS4+919 G > a mutation. *Oncotarget* 7, 87161–87179. doi: 10.18632/oncotarget.13552

Eng, C. M., Resnick-Silverman, L. A., Niehaus, D. J., Astrin, K. H., and Desnick, R. J. (1993). Nature and frequency of mutations in the alpha-galactosidase a gene that cause Fabry disease. *Am. J. Hum. Genet.* 53, 1186–1197.

Ferreira, S., Ortiz, A., Germain, D. P., Viana-Baptista, M., Caldeira-Gomes, A., Camprecios, M., et al. (2015). The alpha-galactosidase A p.Arg118Cys variant does not cause a Fabry disease phenotype: data from individual patients and family studies. *Mol. Genet. Metab.* 114, 248–258. doi: 10.1016/j.ymgme.2014.11.004

Ferri, L., Guido, C., la Marca, G., Malvagia, S., Cavicchi, C., Fiumara, A., et al. (2012). Fabry disease: polymorphic haplotypes and a novel missense mutation in the GLA gene. *Clin. Genet.* 81, 224–233. doi: 10.1111/j.1399-0004.2011.01689.x

Froissart, R., Guffon, N., Vanier, M. T., Desnick, R. J., and Maire, I. (2003). Fabry disease: D313Y is an alpha-galactosidase a sequence variant that causes pseudodeficient activity in plasma. *Mol. Genet. Metab.* 80, 307–314. doi: 10.1016/S1096-7192(03)00136-7

Garman, S. C., and Garboczi, D. N. (2004). The molecular defect leading to Fabry disease: structure of human alpha-galactosidase. *J. Mol. Biol.* 337, 319–335. doi: 10.1016/j.jmb.2004.01.035

Hastings, M. L., and Krainer, A. R. (2001). Pre-mRNA splicing in the new millennium. *Curr. Opin. Cell Biol.* 13, 302–309. doi: 10.1016/S0955-0674(00)00212-X

Ishii, S., Nakao, S., Minamikawa-Tachino, R., Desnick, R. J., and Fan, J. Q. (2002). Alternative splicing in the alpha-galactosidase A gene: increased exon inclusion results in the Fabry cardiac phenotype. *Am. J. Hum. Genet.* 70, 994–1002. doi: 10.1086/339431

Kint, J. A. (1970). Fabry's disease: alpha-galactosidase deficiency. *Science* 167, 1268–1269. doi: 10.1126/science.167.3922.1268

Lenders, M., Weidemann, F., Kurschat, C., Canaan-Kuhl, S., Duning, T., Stypmann, J., et al. (2016). Alpha-galactosidase a p.A143T, a non-Fabry disease-causing variant. *Orphanet J. Rare Dis.* 11:54. doi: 10.1186/s13023-016-0441-z

Li, P., Meinke, P., Huong le, T. T., Wehnert, M., and Noegel, A. A. (2014). Contribution of SUN1 mutations to the pathomechanism in muscular dystrophies. *Hum. Mutat.* 35, 452–461. doi: 10.1002/humu.22504

Li, P., and Noegel, A. A. (2015). Inner nuclear envelope protein SUN1 plays a prominent role in mammalian mRNA export. *Nucleic Acids Res.* 43, 9874–9888. doi: 10.1093/nar/gkv1058

Maquat, L. E. (1995). When cells stop making sense: effects of nonsense codons on RNA metabolism in vertebrate cells. *RNA* 1, 453–465.

Meng, Y., Zhang, W. M., Shi, H. P., Wei, M., and Huang, S. Z. (2010). [Clinical manifestations and mutation study in 16 Chinese patients with Fabry disease]. *Zhonghua Yi Xue Za Zhi* 90, 551–554.

Nakao, S., Kodama, C., Takenaka, T., Tanaka, A., Yasumoto, Y., Yoshida, A., et al. (2003). Fabry disease: detection of undiagnosed hemodialysis patients and identification of a "renal variant" phenotype. *Kidney Int.* 64, 801–807. doi: 10.1046/j.1523-1755.2003.00160.x

Palhais, B., Dembic, M., Sabaratnam, R., Nielsen, K. S., Doktor, T. K., Bruun, G. H., et al. (2016). The prevalent deep intronic c. 639+919 G > A GLA mutation causes pseudoexon activation and Fabry disease by abolishing the binding of hnRNPA1 and hnRNP A2/B1 to a splicing silencer. *Mol. Genet. Metab.* 119, 258–269. doi: 10.1016/j.ymgme.2016.08.007

Patel, A. A., and Steitz, J. A. (2003). Splicing double: insights from the second spliceosome. *Nat. Rev. Mol. Cell Biol.* 4, 960–970. doi: 10.1038/nrm1259

Schaefer, E., Mehta, A., and Gal, A. (2005). Genotype and phenotype in Fabry disease: analysis of the Fabry outcome survey. *Acta Paediatr.* 94, 87–92; discussion 79. doi: 10.1080/08035320510031045

Shabbeer, J., Robinson, M., and Desnick, R. J. (2005). Detection of alpha-galactosidase a mutations causing Fabry disease by denaturing high performance liquid chromatography. *Hum. Mutat.* 25, 299–305. doi: 10.1002/humu.20144

Smid, B. E., Hollak, C. E., Poorthuis, B. J., van den Bergh Weerman, M. A., Florquin, S., Kok, W. E., et al. (2015). Diagnostic dilemmas in Fabry disease: a case series study on GLA mutations of unknown clinical significance. *Clin. Genet.* 88, 161–166. doi: 10.1111/cge.12449

van der Tol, L., Smid, B. E., Poorthuis, B. J., Biegstraaten, M., Deprez, R. H., Linthorst, G. E., et al. (2014). A systematic review on screening for Fabry disease: prevalence of individuals with genetic variants of unknown significance. *J. Med. Genet.* 51, 1–9. doi: 10.1136/jmedgenet-2013-101857

von Scheidt, W., Eng, C. M., Fitzmaurice, T. F., Erdmann, E., Hubner, G., Olsen, E. G., et al. (1991). An atypical variant of Fabry's disease with manifestations confined to the myocardium. *New Engl. J. Med.* 324, 395–399. doi: 10.1056/NEJM199102073240607

Yasuda, M., Shabbeer, J., Benson, S. D., Maire, I., Burnett, R. M., and Desnick, R. J. (2003). Fabry disease: characterization of alpha-galactosidase a double mutations and the D313Y plasma enzyme pseudodeficiency allele. *Hum. Mutat.* 22, 486–492. doi: 10.1002/humu.10275

Zarate, Y. A., and Hopkin, R. J. (2008). Fabry's disease. *Lancet* 372, 1427–1435. doi: 10.1016/S0140-6736(08)61589-5

Zizzo, C., Monte, I., Pisani, A., Fatuzzo, P., Riccio, E., Rodolico, M. S., et al. (2016). Molecular and clinical studies in five index cases with novel mutations in the GLA gene. *Gene* 578, 100–104. doi: 10.1016/j.gene.2015.12.024

New Genetic Loci Associated With Chronic Kidney Disease in an Indigenous Australian Population

Russell J. Thomson[1]*, Brendan McMorran[2], Wendy Hoy[3], Matthew Jose[4,5], Lucy Whittock[6], Tim Thornton[7], Gaétan Burgio[2], John Duncan Mathews[8] and Simon Foote[2]

[1] Centre for Research in Mathematics, School of Computing, Engineering and Mathematics, Western Sydney University, Sydney, NSW, Australia, [2] John Curtin School of Medical Research, Australian National University, Canberra, ACT, Australia, [3] Centre for Chronic Disease, Faculty of Health, The University of Queensland, Brisbane, QLD, Australia, [4] Menzies Institute of Medical Research, College of Health and Medicine, University of Tasmania, Hobart, TAS, Australia, [5] School of Medicine, College of Health and Medicine, University of Tasmania, Hobart, TAS, Australia, [6] Institute for Marine and Antarctic Studies, College of Sciences and Engineering, University of Tasmania, Hobart, TAS, Australia, [7] Department of Biostatistics, School of Public Health, University of Washington, Seattle, WA, United States, [8] Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, VIC, Australia

The common occurrence of renal disease in Australian Aboriginal populations such as Tiwi Islanders may be determined by environmental and genetic factors. To explore genetic contributions, we performed a genome-wide association study (GWAS) of urinary albumin creatinine ratio (ACR) in a sample of 249 Tiwi individuals with genotype data from a 370K *Affymetrix* single nucleotide polymorphism (SNP) array. A principal component analysis (PCA) of the 249 individual Tiwi cohort and samples from 11 populations included in phase III of the HapMap Project indicated that Tiwi Islanders are a relatively distinct and unique population with no close genetic relationships to the other ethnic groups. After adjusting for age and sex, the proportion of ACR variance explained by the 370K SNPs was estimated to be 37% (using the software GCTA.31; likelihood ratio = 8.06, p-value = 0.002). The GWAS identified eight SNPs that were nominally significantly associated with ACR ($p < 0.0005$). A replication study of these SNPs was performed in an independent cohort of 497 individuals on the eight SNPs. Four of these SNPs were significantly associated with ACR in the replication sample ($p < 0.05$), rs4016189 located near the *CRIM1* gene ($p = 0.000751$), rs443816 located in the gene encoding *UGT2B11* ($p = 0.022$), rs6461901 located near the *NFE2L3* gene, and rs1535656 located in the *RAB14* gene. The SNP rs4016189 was still significant after adjusting for multiple testing. A structural equation model (SEM) demonstrated that the rs4016189 SNP was not associated with other phenotypes such as estimated glomerular filtration rate (eGFR), diabetes, and blood pressure.

Keywords: chronic kidney disease, genome-wide association study, Australian Aboriginal and Torres Strait Islanders, indigenous genetics, gene–environment interaction, urinary albumin creatinine ratio

INTRODUCTION

There is an epidemic of renal disease in the Australian Aboriginal population. In the Northern Territory (NT), diabetes deaths in Aboriginal people are increased up to 15-fold, cardiovascular deaths by three- to sixfold (Hoy, 1996), while rates of treated end-stage renal disease (ESRD) are more than 20 times that of non-Indigenous Australians overall, and more than 60-fold greater in some regions (Spencer et al., 1998). The burden of morbidity is great, costs of hospitalization and ESRD treatment are huge (You et al., 2002) and premature deaths of young and middle-aged adults have incalculable impact on families and communities. Studies in the NT Aboriginal communities of Nauiyu, Wadeye, Borroloola (the Outreach communities), the Tiwi Islands, and Angurugu, showed that the rates of hypertension, albuminuria/proteinuria, and diabetes were high, although they differed by community (Hoy et al., 2006b). Relative to the (mostly non-Indigenous) participants in the nationwide AusDiab study, rates of proteinuria in the first three of these Aboriginal communities were increased 2.3- to 8.7-fold, hypertension was increased 2.7- to 3.6-fold and diabetes was increased 6- to 9.6-fold (Hoy et al., 2006b). Incidence increased markedly with age. In addition, there was pronounced overlap of these conditions that also increased with increasing age.

The causes of this disease epidemic are poorly understood but are very likely linked to changes associated with compressed transition to a quasi-western lifestyle in the last 50 years. However, smoking and obesity do not appear to fully explain the disease (Wang and Hoy, 2003), and central fat deposition, inflammation and infections, psychosocial stress, birthweight, and a genetic propensity to develop renal disease in the presence of other factors have been suggested as causes. For example, streptococcal impetigo, and other markers of Group A Streptococcal infection are predictive of albuminuria in Aboriginal children and adults (Van Buynder et al., 1992; Goodfellow et al., 1999). The same association has been seen in this population (Hoy et al., 2012).

Renal disease in the Australian Aboriginal population has classic pathological findings, characterized by the presence of fewer, larger glomeruli with or without increased sclerosis (Hoy et al., 2006a). Kidneys from Aboriginal people with no known kidney disease have 30% fewer glomeruli (Douglas-Denton et al., 2006; Hoy et al., 2006a) and these glomeruli tend to be larger, as determined by autopsy. It is possible that the smaller number of glomeruli is compensated by glomerulomegaly. Interestingly, the most common lesion seen in biopsies from individuals with renal disease from remote communities is glomerulomegaly and not diabetic, or other change (Hoy et al., 2014). There is a hypothesis that the smaller kidneys could be a consequence of low birth weight, a variation on Barker's hypothesis (Manalich et al., 2000; Hoy et al., 2016). However, it is also possible that there are other factors at play.

Kidneys with fewer, large glomeruli are termed oligomeganephronic kidneys. Elsewhere in the world, this condition is most often associated with low birth weight and virtually always occurs in infants with ESRD occurring by early childhood (Wilhelm-Bals et al., 2014). A review of the literature (Alves et al., 2012) found only six cases of adult-onset oligomeganephronia. Histology on biopsies showed a pathology reminiscent of that seen in the remote Aboriginal communities with variable presence of glomerosclerosis (Alves et al., 2012). Mouse studies also suggest an inverse relationship between glomerular quantity and the presence of albuminuria (Long et al., 2013).

The Tiwi Islanders living on Bathurst and Melville Islands north of Darwin have the one of the highest measured prevalence of renal disease among Australian Aboriginals. The problem is so severe that the 3000-strong islander community requires its own dialysis unit and renal failure (start of dialysis or dying without dialysis) is the most common cause of death. It has been found that kidney disease is more common in Tiwi Islanders than three other Aboriginal communities screened and nearly 10 times that of non-Aboriginal people (OR 9.7 [8.2,11.7]) (Hoy et al., 2006b). This suggests that the Tiwi Islanders have a far greater risk of developing renal disease than other northern Australian Aboriginal communities, despite having a profile for either hypertension or diabetes similar to other communities. The Tiwi Islanders have been isolated for 15,000 years up until the last 50–100 years (Tiwi Land Council, 2018). Consequently, we hypothesize that genetic variants associated with kidney disease (in the presence of other factors brought on by quasi-western lifestyle) have become more common among the Tiwi Islanders than mainland Aboriginals. These risk alleles could have become more common with random drift, when the population was genetically isolated and the environmental risks were not present.

There are over 200 genetic changes that cause or are associated with either renal disease or renal developmental abnormalities (Vivante and Hildebrandt, 2016). While many of these are likely to be irrelevant to the Tiwi renal disease, there are several genetic changes that are associated with proteinuria alone or oligonephronia or oligomeganephronia. These include a deletion of the short arm of chromosome 4 with a residual monosomy known as Wolf–Hirschhorn syndrome where oligomeganephronia is a rare accompaniment (Park and Chi, 1993; Anderson et al., 1997). Mutations in PAX2 in both mice and humans give rise to oligomeganephronia with the occasional co-occurrence of a coloboma. And finally, hepatocyte nuclear factor-1b (HNF 1b) mutations give rise to a syndrome that includes meganephronia (Bingham et al., 2001). Current findings suggest that the excess burden of kidney disease in African Americans is due to risk alleles in the *APOL1* gene; however, these same risk alleles are not found in the Tiwi Islander community (Hoy et al., 2017).

There is evidence that renal disease in Aboriginal populations has an important genetic component. There is a significant association between a common polymorphism at codon 72 of the *p53* gene and urine albumin creatinine ratio (ACR) among a remote coastal Aboriginal Australian community in the East Arnhem region of NT (McDonald et al., 2002).

The heritability of the natural log of ACR is estimated to be 64% in an Aboriginal population, while for systolic and diastolic blood pressure, it was only 26 and 11%, respectively. The heritability of eGFR was not found to be significantly different from zero (Duffy et al., 2016). This study also displayed evidence of association of variants in *ACE* and *TP53* with ACR.

Here we describe a classical genome-wide association study (GWAS) of Tiwi Islanders in which we identified several loci associated with markers of renal disease. Two of these loci, single nucleotide polymorphisms (SNPs) rs443816 located in the gene encoding *UGT2B11* and rs4016189 located in the gene encoding *CRIM1*, were also significantly associated with the renal disease phenotype in an independently sampled cohort.

MATERIALS AND METHODS

Sample Collection

In the 1990s, in close consultation with the Tiwi Land Council, the protocol was approved by the institutional ethics committee and written informed consent was obtained from all participants. Phenotype data were collected for 1492 samples, with sufficient DNA obtained for 249 samples.

In 2013–2014, data were collected on a second cohort of 497 Tiwi Islander study participants. This study was carried out in accordance with the recommendations of the Human Ethics Committee (Tasmania) Network ethics reference number H0012832, and in conjunction with all institutional ethics committees of the research team and the Human Research Ethics Committee of the NT Department of Health and Menzies School of Health Research. Written informed consent was obtained from all participants. Participants donated blood and urine samples, resulting in good quality DNA samples for 492 individuals. Clinical data collected include: date of birth, gender, smoking and alcohol history, blood pressure or kidney disease medicine use, height, weight, waist circumference, random blood glucose, glycated hemoglobin, serum albumin (measured by radioimmunoassay), serumcreatinine (measured by jaffe reaction), estimated glomerular filtration rate (eGFR), urinary protein, and urinary ACR. eGFR was estimated from serum creatinine levels using the MDRD Study equation (Levey et al., 1999).

In both cohorts, samples were collected on all participants who presented at the clinic and consented to be a part of the study.

DNA Preparation and Analysis

The blood samples from which the DNA was derived for the *Affymetrix* genotyping study were collected from Tiwi Islanders during 1995–1996. Lymphocytes were obtained from buffy coat preparations, and stored under liquid nitrogen. The genomic DNA was extracted from the lymphocyte samples in 2007 using a Dneasy Blood and Tissue Kit (Qiagen). The extracted DNA was subsequently amplified using a GenomiPhi V2 DNA Amplification Kit V2 (GE Healthcare) to provide sufficient quantity and quality genomic DNA for the genotyping. For each sample, two separate 20 µL amplification reactions were conducted, and then pooled and purified using the AMPure PCR Purification System (Agencourt). The quantity and quality of the amplified and purified DNA were determined using a Nanodrop DNA Analyser (Thermo Fisher Scientific). Out of 257 lymphocyte samples subjected to DNA extraction and subsequent amplification, 251 met the minimum requirements

for the genotyping analysis, namely: $>2\,\mu g$ DNA, 100 ng/µL, and OD260/280 ratio between 1.7–1.9.

Whole genome SNP genotyping was performed on the amplified DNA using *Affymetrix* 5.0 SNP chips. The accuracy of the genotype calls was assessed by using the following analyses. First, comparison of 25 duplicate samples indicated a concordance rate of 98.4% (range: 96.1–99.2%). Second, we used the genotype data to ascertain first- and second-degree relationships among the samples (using *PLINK* software; Purcell et al., 2007). These inferences were correctly predictive for 87% of the reported relationships. Discrepancies here are likely due to errors in self-reporting as many Tiwi Islanders are brought up by extended families and the concept of primary carer is different to other cultures.

The DNA used for the genotyping replication studies was obtained from blood samples collected from Tiwi Islanders in a separate study on the population conducted during 2015. DNA was extracted from whole blood samples using the Nucleospin Blood XL silica matrix system (Machery-Nagel) at the Australian Genome Research Facility (AGRF) in Australia. Genotyping of these samples for the SNPs identified in the GWAS stage of the study was performed using TaqMan[TM] genotyping primers (Applied Biosystems). The SNP analysis of the *CRIM1* region was performed using the Sequenom platform (by the AGRF).

Blood samples from 12 anonymous Tiwi Islanders were sent to *Illumina* for whole genome sequencing in 2013, using the *HumanOmni2.5*-8v1 chip. The mean fold coverage for these samples ranged from 33.6 to 42.1. Given the small sample size, and the lack of phenotype information, the sequence data could not be used for association or imputation. The sequence data could be used to identify polymorphic SNPs for fine mapping.

Population Stratification and Linkage Disequilibrium

A principal component analysis (PCA) was conducted using approximately 50,000 SNPs that were polymorphic in the Tiwi samples [minor allele frequency (MAF) > 0.2] and that were not in linkage disequilibrium (as determined using 73 unrelated Tiwi samples and *PLINK*). The PCA was performed using *EIGENSOFT* (Patterson et al., 2006) on samples from 11 populations in release 3 of phase III of the International Haplotype Map Project (hapmap3) samples. The Tiwi samples were not included in generating the principal components, to prevent the PCs from being due to differences in platforms or genotyping errors. The hapmap3 and the Tiwi samples were then projected on to a plane (**Figure 1**) using the first two principal components to generate a PCA plot. Eight Tiwi samples were self-reported as being of "mixed-race" and these were indicated on the PCA plot. The genetic distance between the midpoint of the Tiwi samples and the 11 hapmap3 samples was calculated using Euclidean distance including all PCs.

The median length of haplotype blocks across the genome was estimated for each of the Tiwi samples and the 11 hapmap3 samples; the median haplotype block size was then calculated
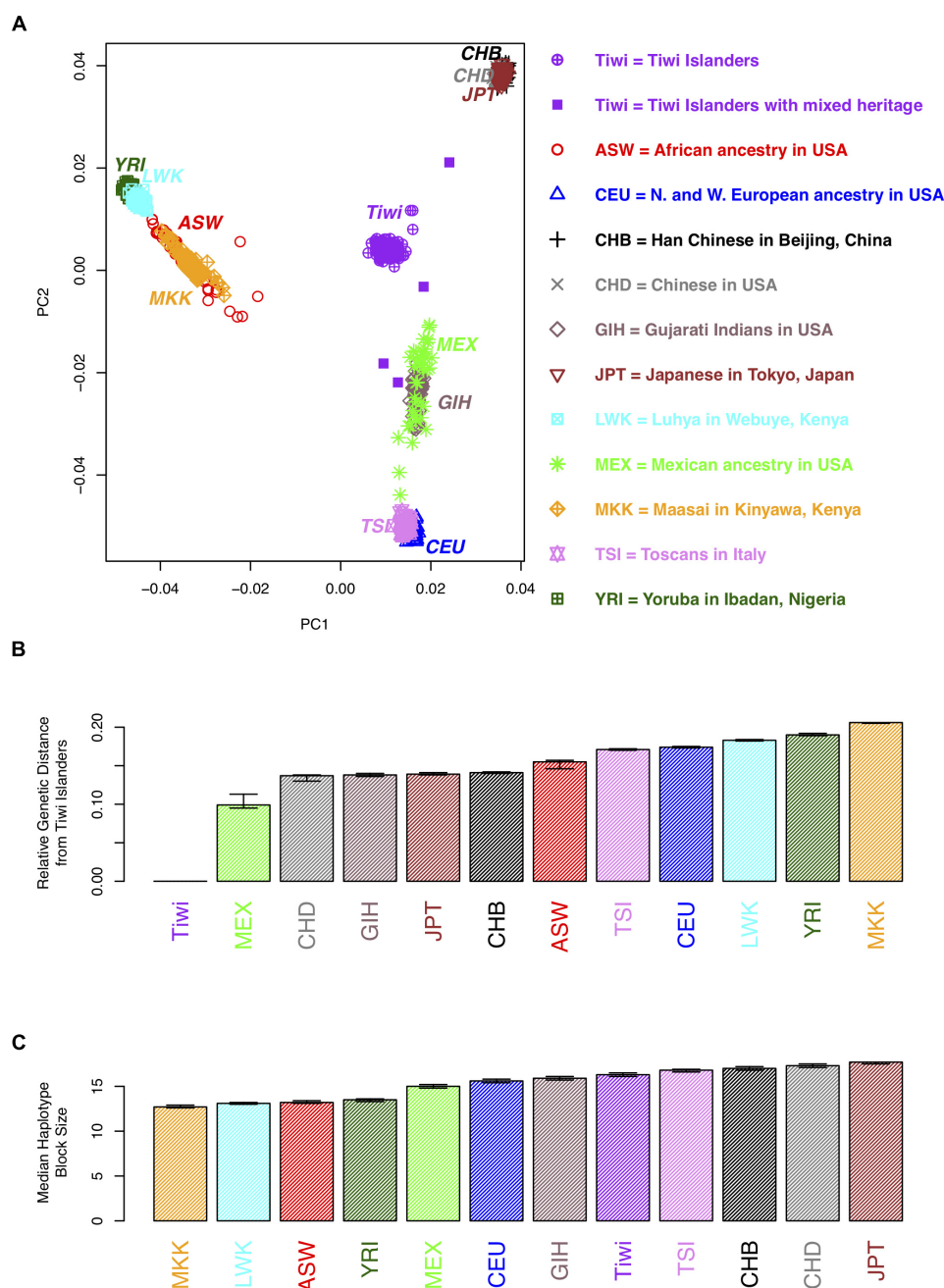
FIGURE 1 | Population genetic analyses of 73 unrelated individuals from the 249 Tiwi cohort from the 1990s. **(A)** A plot of the first two principal components (PCs) for an analysis of the hapmap3 samples from 11 populations and the Tiwi cohort. Each point represents an individual. More distant points are more genetically distinct. Tiwi study participants that self-reported as mixed heritage are indicated by opaque purple squares. **(B)** The genetic distance between the midpoints of each of the hapmap3 samples and the Tiwi samples was calculated using all principal components. **(C)** The median haplotype block size of the hapmap3 samples from 11 populations and the 1990s Tiwi cohort. Haplotypes were defined using software *Haploview* based on the *four gamete rule*; 95% confidence intervals, estimated using bootstrap sampling, are displayed for the estimates in **B** and **C**.

across each population. The *four gamete rule* was used to define a haplotype block, dictating that all pairs of SNPs within a haplotype block exhibit less than four possible haplotypes or gametes, showing no evidence of a historical recombination within the haplotype block (Thorisson et al., 2005).

Association Analysis

Genetic regions of homozygosity (ROH) were inferred using the "Runs of Homozygosity" option in the software *PLINK*. Homozygosity association was carried out by searching for ROH that were shared by study participants with high ACR. The data

were examined to see if the high level of kidney disease could be explained through consanguinity by seeing if there was any correlation between log(ACR) and the total number and length of ROH for each individual.

For the association analysis, the natural log of ACR was used, and age and sex were included as covariates. In a previous study from the same population, heritability of ACR was estimated to be 64%, while the heritability of eGFR was not significantly greater than zero (Duffy et al., 2016). Detectable reduction in eGFR, assessed through creatinine-based measures, is a late event in the expression of renal disease in this population which is uniformly predicted by years of albuminuria, and once established progresses quickly to kidney failure. Thus, it manifests in only small numbers of people at any given study which excludes people on dialysis, and is not such a good endophenotype for kidney disease (Hoy et al., 2001).

The variance of ACR that can be explained by the *Affy 5.0* genotypes was examined using the *GCTA* software (Yang et al., 2011). The association analysis was conducted with a linear mixed model, also using the software *GCTA*. An additive genetic model was assumed and age was adjusted for by treating it as a covariate. The linear mixed effects regression model accounts for the non-independence among family members by modeling the variance structure of the relationships between individuals as a random effect. The analysis adjusted for relatedness by inferring it from the genome wide data. The software *GCTA* is just one of many methods for adjusting for relatedness in genome wide studies (Thomson and McWhirter, 2017). It has the advantage that we were able to obtain more power, by not including the chromosome that the locus of interest was on, in the relatedness inference (option –mlma-loco in GCTA).

To examine whether our top hit (rs4016189), which was in a gene desert, could affect transcriptional regulation of the nearest gene (*CRIM1*), we interrogated an eQTL dataset published previously (Xia et al., 2012). This dataset contains *p*-values for the association between gene expression (of all known genes) and a genome wide set of SNPs, in 11 hapmap3 samples. From this dataset, the *p*-values were extracted for association between the SNPs within the genomic region of our top hit, and regulation of *CRIM1* in the Mexican hapmap population.

The Follow-Up Dataset

The top eight SNPs from the GWAS, were re-genotyped in the second cohort at the AGRF. These SNPs were chosen if the *p*-value of the corresponding SNP and the next closest SNP were both less than 0.0005 [corresponding to a $-\log_{10}(p\text{-value})$ greater than 3.3]. Whole genome data were not collected on all 497 Tiwi Islanders from the second cohort, so it was not possible to adjust for relatedness. A simple linear regression model was used to measure the significance of association for the top eight SNPs and the fine mapping analyses. The covariate, age was adjusted for, and an additive genetic model was assumed.

Structural Equation Model

A structural equation model (SEM) was used to examine the genotype most associated with ACR, and its effect on other phenotypes. SEMs provide an approach to propose a causal hypothesis in a statistical framework that can then yield causal interpretations conditional on the model assumptions and data. The *R* package; *piecewiseSEM* was used to estimate the coefficients that describe the strength of the causal pathways in the SEM, as it allowed the use of generalized linear models (Lefcheck, 2016). Dependent variables in the SEM are ACR, eGFR, systolic blood pressure (sysBP), and diabetes (HbA1c > 6.5). ACR was logged and diabetes is modeled using a logistic regression model to match the assumptions of regression. The SEM was modeled assuming a causal relationship from diabetes to blood pressure and to ACR/eGFR and that ACR and eGFR are correlated. Using the standardized coefficients, the effect sizes of the genotype (rs4016189), age, sex, and body mass index (BMI) on the dependent variables were compared.

RESULTS

Phenotypes

The summary statistics of kidney disease, blood pressure, diabetes, age, and BMI are shown in **Table 1** for the two cohorts.

The variance in ACR, explained by genotype as measured by the *Affymetrix* 370K SNP chips was 37% (estimated using the software GCTA), after adjusting for age and sex. This was significantly greater than zero (likelihood ratio = 8.06, *p*-value = 0.002), showing that the genetic variation in the SNPs genotyped has a significant effect on the risk of kidney disease, and the SNPs explained 37% of the variation in ACR among the Tiwi Islanders. We would expect that the variance explained by

TABLE 1 | Mean/median/proportion of the variables of interest in the two cohorts of Tiwi study participants.

	1990s Cohort	2013–4 Cohort
Number of participants with genotype data	249	497
Age (mean, range) (years)	33.3 (12–73)	39.7 (17–75)
Sex (number of males/females)	133 M, 116 F	253 M, 239 F
ACR (median, IQR, range) (mg/g creat)	20 (7–249) (1–6058)	23 (7–159) (1–17,555)
eGFR (mean, range) (mls/min/1.73 m^2)	105.9 (25–243)	84.4 (5–90)
Systolic BP (mean, range) (mmHg)	121.2 (84–170)	114.5 (71–187)
Diastolic BP (mean, range) (mmHg)	73.2 (39–122)	73.8 (48–124)
Diabetes (Y/N in 1990s cohort*, hba1c > 6.5 in 2013–4 cohort)	8.5%	18.2%
BMI (median and IQR)	22.3 (19–27)	23.8 (21–28)
Relationships	135 (0.5%) are first degree relationships, 307 (1.2%) are second degree relationships.	Unknown

*In the 1990s cohort, a study participant was deemed to have diabetes if a clinical history was in the records, or if they were on hypoglycemic medications, or WHO diagnostic levels of glycemia fasting or were 2 h after oral glucose challenge.

the whole genome sequence data to be closer to the heritability estimate of 64% (Duffy et al., 2016), as there are likely to be rare variants that also contribute to variation in ACR.

Genotypes

The *Affymetrix* 5.0 SNP chip is made up of approximately 500,000 SNPs that are known to have a MAF > 0.05 in a Caucasian population. Based on the 249 individuals that were part of the 1990s sample, for the Tiwi population, 40% of these SNPs have a MAF < 0.05 and 35,000 of these SNPs are monomorphic. This high rate of monomorphism may be because *Affymetrix* 5.0 SNPs were selected to be polymorphic in a Caucasian population, while other SNPs could be polymorphic in the Tiwi population.

We used the genome wide SNP data to conduct a PCA to compare the Tiwi samples with several other ethnic groups and to distinguish whether there was any evidence of population stratification within the Tiwi samples. The PCA results (**Figures 1A,B** and **Supplementary Table S1**) distinguish several different ethnic groups into individual clusters, including the African Masssai, Luhya and Yoruba people, and Mexicans. The Han and United States-based Chinese and Japanese groups clustered together in a separate domain, as did the Tuscans and CEPH European ancestry groups. Strikingly, the Tiwi clustered together in another completely separate domain of the PCA plot. Eight Tiwi samples were self-reported as being of "mixed-race." Four of these were noticeably separated from the other Tiwi samples. However, there was no evidence of population stratification within the Tiwi samples.

The median haplotype block size of the Tiwi samples was 16.3 kb (95% C.I.: 16.1–16.5 kb), based on the *four gamete rule*. When comparing this estimate to the median haplotype block size to the other populations, LD extends further in the Tiwi samples than the African and some other populations (**Figure 1C** and **Supplementary Table S1**). However, LD does not extend further than the Chinese, Japanese, or Tuscan populations, suggesting that the Tiwi population have a distant common ancestor, despite their small effective population size.

Association

Genetic ROH were inferred using the "Runs of Homozygosity" option in the software *PLINK*. We found that there were no runs of inferred homozygosity that were shared by more than two individuals. The total length of inferred ROH was not significantly longer for individuals with higher ACR (*kendall's r* = 0.018, *p*-value = 0.67), suggesting that higher levels homozygosity do not explain the higher levels of kidney disease in the Tiwi population. There were very few ROH inferred that were less than 2 Mb in length. This may be due to the high genotyping error rate or low coverage. Consequently, it may be possible that the higher levels of kidney disease in the Tiwi population could be due, in part, to shorter runs of homozygosity that are less than 2 Mb in length.

A GWAS was carried out using GCTA analysis, adjusting for age and sex, with log(ACR) as the phenotype. This is a mixed effects analysis that accounts for relatedness. **Figure 2A** presents all the association results as $-\log_{10}(p\text{-values})$ across the genome as a Manhattan plot. A quantile-quantile plot

(**Figure 2B**) displays the expected vs observed *p*-values as points on a scatter plot, ordered from largest to smallest. The majority of the points on the scatterplot lie on the one-to-one line, suggesting that the *p*-values are not over-inflated (as could happen with unaccounted relatedness or population stratification). The most significant SNPs (largest $-\log_{10}$ *p*-values) were not as significant as expected, suggesting that the study may be underpowered, necessitating replication.

The top eight SNPs were chosen for follow up in an independent sample (**Table 2**). These SNPs were chosen when the *p*-value of the corresponding SNP and the next closest SNP were both less than 0.0005 [corresponding to a $-\log_{10}(p\text{-value})$ greater than 3.3]. Among the top hits identified in the GWAS were SNPs located in or near genes with potential roles affecting kidney function and development. The SNP with the lowest significance, rs4016189 (*p*-value = 9.76×10^{-5}), was located in a large intergenic region and approximately 614 kb upstream of the *CRIM1* gene on chromosome 2p2.22; the next closest protein-coding gene to this SNP was more than 1 Mb distant (in the opposite direction to *CRIM1*). The association between this variant and ACR levels also remained highly significant in the replication study (*p*-value = 0.000751). The associated allele was more common in both Tiwi cohorts, than other populations (**Table 3**). *CRIM1* encodes Cysteine Rich Transmembrane BMP Regulator 1, a transmembrane protein containing six cysteine-rich repeat domains and an insulin-like growth factor-binding domain. It is known to have roles in renal development (Wilkinson et al., 2012; Phua et al., 2013) and mice with a *Crim1* knockdown develop a glomerulomegaly (Georgas et al., 2000) and therefore was of high interest to our study. To test the possibility that the SNP could be involved in transcriptional regulation of *CRIM1*, we interrogated the genomic region containing this SNP in an eQTL dataset published previously (Xia et al., 2012). We observed a significant association between several variants, including rs4016189, and *CRIM1* expression in the hapmap3 MEX population (**Supplementary Figure S1**). We also conducted a fine scale genetic analysis of the region near rs4016189 and the *CRIM1* locus. To do this, we identified variants spanning approximately every 2 kb within the region (using whole genome sequences from 12 Tiwi individuals), developed Sequenom assays for each variant and genotyped these in the 497 samples in the 2013–14 cohort. The results of the association tests conducted for each of these SNPs (97 in total) with ACR levels are presented in **Figure 3**. As well as confirming the rs4016189 association using this independent assay, we observed multiple SNPs in this region, including those located nearest *CRIM1* (e.g., rs2968579; 455 kb away from the gene) were also significantly associated. This provided additional evidence that a region upstream and near to *CRIM1* was associated ACR levels. **Figure 3** also displays the LD between these 97 SNPs. Three distinct haplotype blocks can be seen across the region, although there is some evidence that the SNPs around rs4016189 are in LD with the SNPs around rs2968579.

Three other SNPs, rs4438816, rs6823947, and rs12511454 associated with ACR in the GWAS were located within the same gene, *UGT2B11* on chromosome 4q13.2. The gene encodes the enzyme UDP Glucuronosyltransferase

FIGURE 2 | (A) A Manhattan plot displaying the −log10(p-values) for the association between the phenotype ACR and the SNPs from the *Affymetrix* SNP Chip. **(B)** A QQ-plot, showing the observed p-values on the y-axis and the expected p-values on the x-axis under the null hypothesis of no association.
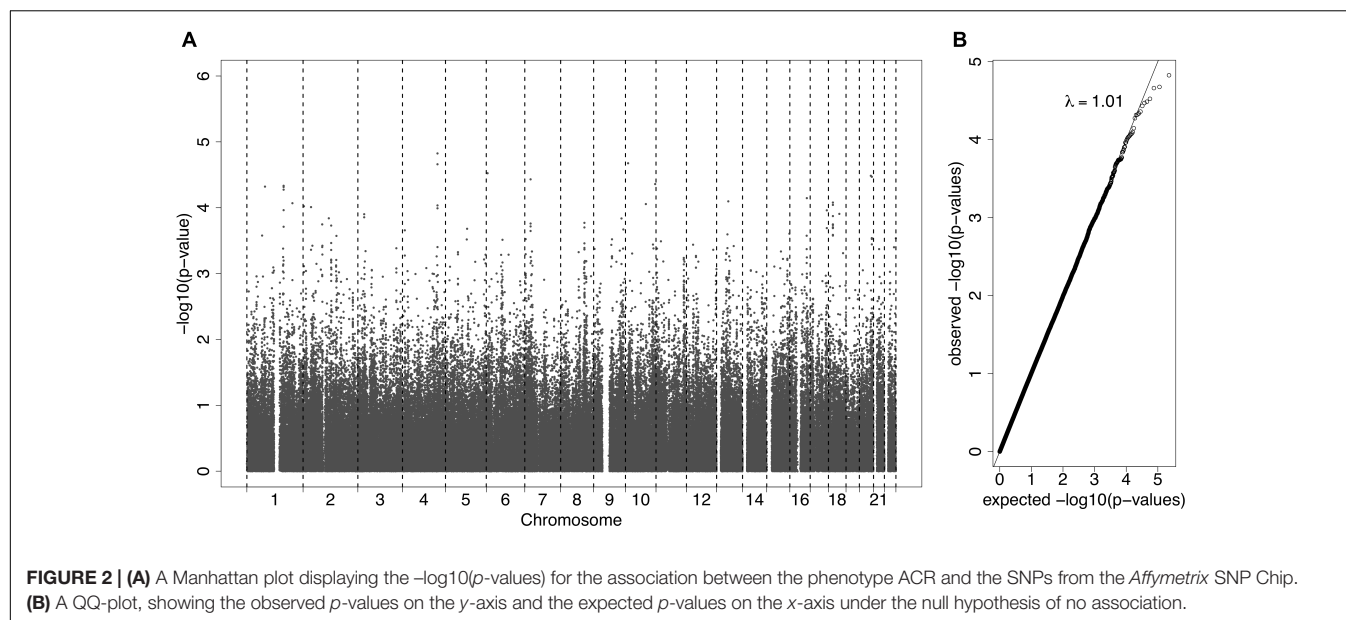
TABLE 2 | The top eight hits in the GWAS study on the original cohort, along with the association results in the replication cohort.

Chr	SNP	Position (bp) GRCh38.p12	Assoc. allele	Nearest gene	*p*-value in original sample	*p*-value of nearest SNP in original sample	*p*-value in replication sample
2	rs4016189	35741807	C	CRIM1 (614 kbp away)	9.76E-05	0.000433	0.000751
4	rs4438816	69165940	T	UGT2B11	0.000739	5.24E-05	0.022
4	rs6823947	69545677	A	UGT2B11	5.24E-05	0.000294	0.079
4	rs12511454	69566288	G	UGT2B11	0.000294	5.24E-05	0.078
7	rs6461901	25990123	A	NFE2L3 (162 kb away)	3.70E-05	0.000173	0.035
9	rs1535656	121188709	C	RAB14	0.000144	0.000145	0.037
17	rs2279057	76313344	C	PRPSAP1	0.000109	0.000203	0.069
20	rs2254239	56994842	C	lncRNA, LOC105372683 (8 kb away)	3.39E-05	0.000288	0.962

TABLE 3 | Allele frequencies of the associated alleles for SNPs that were significant in the replication study (p < 0.05).

Chr	SNP	Assoc allele	Gene	Allele freq. in Tiwi Cohort 1	Allele freq. in Tiwi Cohort 2	Allele freq in MEX	Allele freq. in YRI	Allele freq. in CEU
2	rs4016189	C	CRIM1 (614 kbp away)	0.83	0.82	0.36	0.55	0.52
4	rs4438816	T	UGT2B11	0.40	0.46	0.25	0.26	0.18
7	rs6461901	A	NFE2L3 (162 kb away)	0.21	0.25	0.08	0.75	0.14
9	rs1535656	C	RAB14	0.27	0.24	0.26	0.01	0.10

Allele frequencies for the other populations were obtained from https://www.ncbi.nlm.nih.gov/variation/tools/1000genomes/. MEX = Mexican Ancestry in the United States, YRI = Yoruba in Ibadan, Nigeria, CEU = North and Western European Ancestry in United States.

(UGT) Family 2 Member B11 (EC 2.4.1.17) – none of the associated SNPs changed the coding sequence of this gene. The former SNP was also significantly associated with ACR in the replication study (p = 0.022), although associations for the other two did not quite reach significance (p = 0.079 and 0.078). The associated allele for rs4438816 was more common in both Tiwi cohorts, than other populations (**Table 3**). Members of the UGT enzyme family catalyze an intracellular process known as glucuronidation, in which potentially toxic endogenous compounds as well as many

drugs and xenobiotics are conjugated and subsequently eliminated from the body. Previous studies have shown that some xenobiotics cause kidney disease, such as Balkan endemic nephropathy (BEN) caused by exposure to aristolochic acid present in flour produced from wheat from the endemic areas (Stefanovic and Polenakovic, 2009).

Two additional SNPs, rs6461901 and rs1535656, were also significantly associated with ACR levels in the replication study (p < 0.05). The associated alleles in these two SNPs did not consistently have higher allele frequencies in the Tiwi samples,
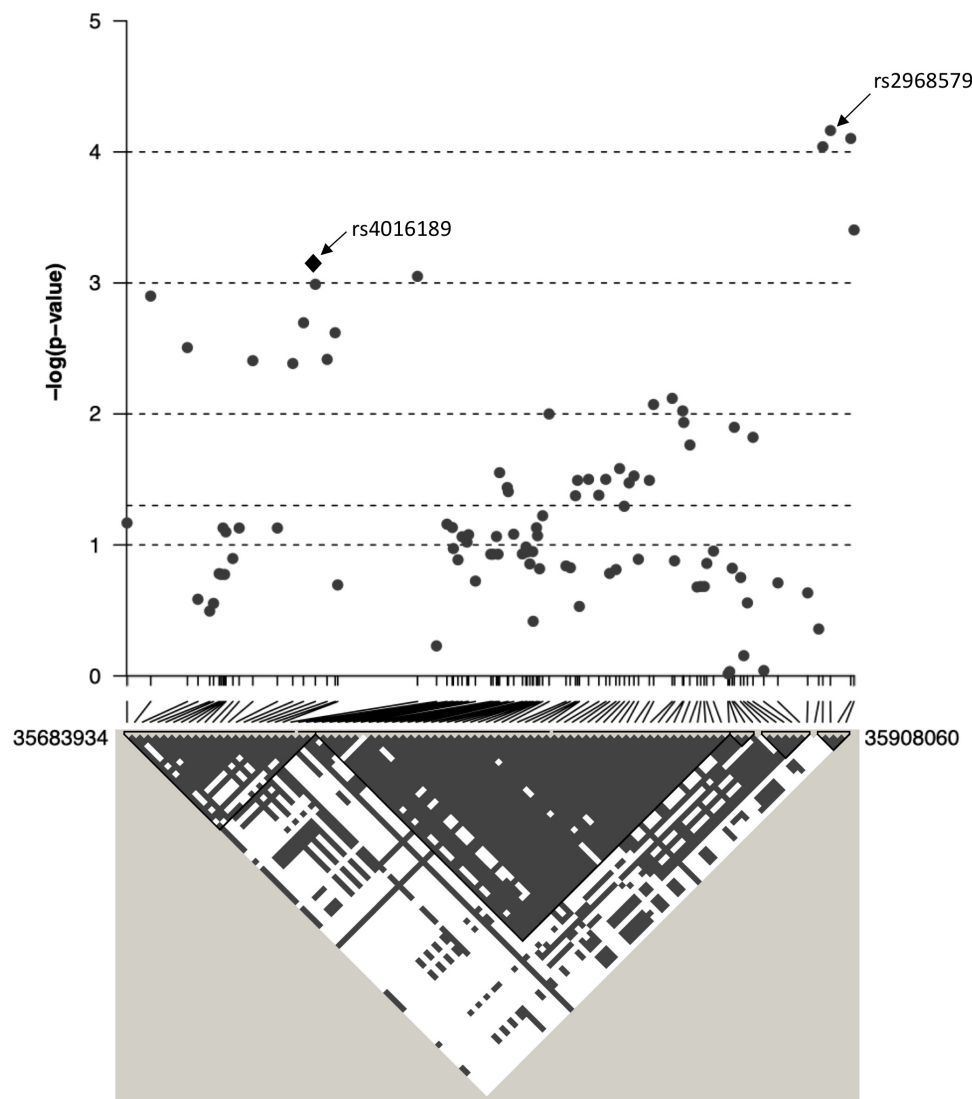
FIGURE 3 | The association between ACR and 97 SNPs in the region surrounding rs4016189 is displayed as –log10 *p*-values. CRIM1 gene lies 448 kb upstream of the region displayed on this graph. Positions are based on ensembl GRCh38.p12. Below the graph, the linkage disequilibrium between each of the 97 SNPs is displayed as gray/white squares, using the *four gamete rule*. SNPs that do not show all four possible haplotypes/gametes (i.e., in LD) are colored gray (as opposed to white).

compared to other populations (**Table 3**). The former SNP is located in an intergenic region and approximately 162 kb upstream of *NFE2L3* on chromosome 2q31.2. *NFE2L3* encodes a transcription factor, nuclear factor erythroid 2-related factor 2 (NRF2), which has been widely studied for its cytoprotective functions, including in kidney disease (Nezu et al., 2017). The other SNP was located in an intronic region of *RAB14* on chromosome 9p33.2. *RAB14* encodes a GTPase expressed in many cell types, including the kidney, and is involved the regulation of ER protein trafficking (Junutula et al., 2004). No obvious roles in kidney function or development have been ascribed to this gene. None of the other SNP identified in the GWAS were significantly associated with ACR levels in the replication study.

Associations With Other Phenotypes

A SEM was fitted to compare the effect sizes of the SNP rs4016189, that is near CRIM1, with known risk factors for kidney disease, blood pressure, and diabetes (**Figure 4** and **Supplementary Table S2**). The SEM consisted of four regression equations, explaining the variation in four dependent variables; ACR, eGFR, sysBP, and diabetes (HbA1c > 6.5). ACR and eGFR were negatively correlated (red curved arrow) as expected, since high ACR and low eGFR correspond to kidney disease.

The SEM displayed that the SNP, rs4016189 was significantly independently associated with ACR, but not with sysBP, diabetes, or eGFR. The effect size of this SNP on ACR was smaller than age, BMI, and sysBP, but larger than gender.
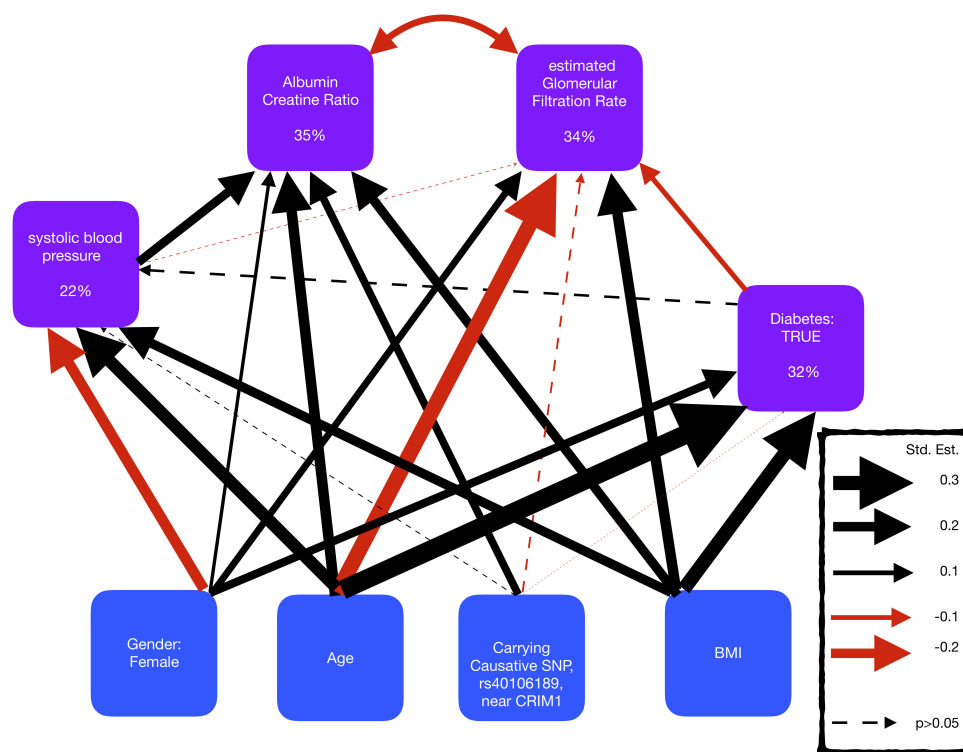
FIGURE 4 | Representation of a structural equation model, exploring the comorbidity of kidney disease, blood pressure, and diabetes. The boxes represent the measures and the arrows show the relationships between the variables (black/red arrows for positive/negative relationships, respectively). The thickness of the arrows has been scaled based on the magnitude of standardized regression coefficients for putative causative relationships (straight arrows) and correlation coefficients for correlated variables (curved arrows). The line type of the arrow denotes significance (solid = $>p$-value < 0.05; dotted => p-value > 0.05). The percentage variance explained by the risk factors is also shown in the boxes. The models also adjusted for cohort, but coefficients are not shown here to make the graphs simpler.

The largest standardized effect size in the SEM was the positive effect of age on diabetes, followed by the negative effect of age on eGFR. The risk factors described the largest proportion of the variation in ACR, and the smallest variation in sysBP, suggesting that there are other risk factors involved in sysBP.

DISCUSSION

Australian Aboriginals experience a high burden of kidney disease that is at least partly independent of other comorbidities that affect kidney function. The problem is particularly acute among the Tiwi Islanders, where rates of ESKD are 30 times that of the general Australian population. The Tiwi have lived in relative isolation for probably thousands of years, and anecdotally, did not suffer from renal disease prior to adopting a quasi-Western lifestyle sometime during the mid-late 20th century. It is possible that dietary and/or other environmental factors introduced since this time have contributed to the rapid emergence of the disease. However, our previous studies indicated a significant heritable component (Duffy et al., 2016), raising the possibility that individuals are genetically predisposed to the disease, but only in the presence of these hypothetical environmental risk factors, i.e., gene–environment interactions. In this study, we investigated the genetic contribution to renal

disease in Tiwi Islanders by conducting a GWAS, in which associations were tested between SNP genetic variants and single measure ACR levels. A number of nominally significantly associated SNPs were identified. These SNPs did not reach genome wide significance, probably due to the small sample size. The top eight SNPs were re-tested for association in a separately collected cohort from the same population. Four of these SNPs were significantly associated with ACR in the replication sample ($p < 0.05$), and through examination of their known biology and functions, there are possible mechanisms through which renal disease in Tiwi Islanders may manifest.

We identified a genomic region upstream and near to the *CRIM1* locus. The prototypical GWAS SNP in this region, rs4016189, was also significantly associated in the replication study, as were a set of additional SNPs located between this SNP and *CRIM1*. The indicative region of association spans at least 250 kb; however, our analysis did not extend sufficiently to include the *CRIM1* locus itself. The associated region may be involved in regulating *CRIM1* transcription, as suggested by the positive eQTL signal from the rs4016189 SNP. It is also possible that there are other more distantly located variants that are in linkage disequilibrium that influence gene function, although we did not observe any non-synonymous changes in *CRIM1* in the 12 genome-sequenced samples. A further in-depth study of the interval and the relationship between sequence variation

and *CRIM1* gene expression will be required to understand the functional link between the associated variant and disease phenotype. The function of *CRIM1* is not well understood, but it has several important roles in development, including in the kidney. Basic studies in the mouse demonstrate that the murine homolog, *Crim1*, is expressed in the developing kidney (Georgas et al., 2000); it is also localized in descending loop of Henle cells in the adult kidney (Park et al., 2018). In humans, immunohistochemical studies of have localized CRIM1 to the podocyte slit diaphragm of the adult human kidney (Nystrom et al., 2009). *Crim1* knockout mice die *in utero* from multiple organ defects. However, a subset of mice homozygous for a *Crim1* hypomorphic mutation, called KTS264 (Phua et al., 2013) survive to adulthood and go on to develop a glomerulomegaly that is reminiscent of that seen in biopsy studies of Australian Aboriginal renal patients. It is therefore tempting to speculate that the renal disease associated variants modify *CRIM1* expression in such a way as to affect the developing kidney and lead, possibly in conjunction with added environmental factor(s), to the renal disease susceptibility phenotype in the Tiwi. That the associated variant may operate directly through effects on kidney function is further supported by our structural equation modeling, which showed no relationships between the variant and kidney disease risk factors such as diabetes and sysBP. However, age, BMI, and blood pressure were all shown to have a greater effect on urinary ACR than the prototypical SNP. Therefore, the underlying causes of disease are clearly complex.

The GWAS also implicated a locus encoding a UGT, *UGT2B11*, in the renal disease susceptibility. Three SNPs in this gene were independently associated with ACR, although only one remained significant in the replication study. The functions of UTGs, and their known involvement in disease, indicate their possible involvement in renal disease pathophysiology. UGTs catalyze the conjugation of glucuronide moieties to many drugs and xenobiotic compounds, an important step in the Phase II detoxification pathway of drug metabolism. Exposure to certain xenobiotics has been shown to cause renal disease. For example, BEN which results from exposure to aristolochic acid present in flour produced in this region (Arsenović et al., 2005) and in animals exposed to mycotoxins (fungal toxins) in contaminated feedstock (Austwick, 1984). Interestingly, BEN is only observed in some households despite the ubiquitous exposure to this xenobiotic, and both a genetic predisposition and environmental (xenobiotic) influence have therefore been implicated. A potential role for genetic factors influencing xenobiotic metabolism is exemplified in a common polymorphism causing UGT2B17 deficiency. UGT2B17 is known to conjugate testosterone and other androgens, and thus enable their excretion in the urine. The genetic deficiency causes dysregulation of testosterone levels including reduced urinary excretion (Jakobsson et al., 2006). In androgen anabolic steroid (AAS) use, urinary detection of the drug is compromised in UGT2B17 deficient individuals and affected abusers can develop renal function problems, presumably because of build-up of AAS (Deshmukh et al., 2010). Therefore, we hypothesize that the renal disease associated UGT2B11 variant identified in the Tiwi may

compromise an ability to conjugate and thus excrete a nephrotoxic xenobiotic or produce a nephrotoxic xenobiotic by conjugation of pro-substance. It is also conceivable that the introduction of this compound coincided with their recent adoption of a Westernized life-style and changed living conditions.

A third associated SNP, rs6461901, lies near a gene whose product, NRF2, is a well-characterized transcriptional regulator that directs the expression of cellular systems employed to protect against oxidant and electrophile-induced stress. Oxidative damage is central to organ damage following ischemic injury, for example, and in the kidney upregulation of NRF2 function provides protection against the development of chronic kidney disease resulting from such injury. This has been shown in many studied in mice with genetic modifications affecting NRF2 function and in human clinical trials using small molecule activators of NRF2 (reviewed in Nezu et al., 2017). We therefore speculate that altered NRF2 expression or function could affect the severity of renal disease in the Tiwi, although we have no knowledge of how the associated variant may affect the gene or its product.

The Tiwi people have been proactively engaged with research in their community for 30 years or more. In the 1980s, research was seen as exploitative by many Indigenous leaders; at that time when there was no national leadership to provide ethical guidance for researchers and support for communities (Kowal and Anderson, 2012). The Tiwi people worked with the Menzies School of Health Research from its earliest years; Menzies had established an ethics committee with Indigenous members to oversight its research. Subsequently, the Tiwi Land Council, through its Health Board, signed an historic research agreement with Menzies to formalize Tiwi control of the research priorities, research information, and samples. This is believed to have been the first ever such legal agreement between an Indigenous community and a research organization.

The Menzies School of Health Research had also helped to focus the national debate by organizing the 1986 conference in Alice Springs, attended by Indigenous leaders, researchers, and NHMRC representatives; this helped to launch a process that eventually led to NHMRC research guidelines to support community decision-making and control of Indigenous health research.

The Tiwi people's engagement with research has had a particular emphasis on chronic disease and kidney disease. They have articulated and shaped the research directions, raised local financial support and raised external funds, specifically the Stanley Tipiloura Fund (to honor the first Tiwi MLA) to support the first phases of this research, worked with government agencies, including Australia's National Health and Medical Research Council, with Kidney Health Australia (and through them, Rio Tinto), and with Servier, Australia. Tiwi people have worked as staff and leaders in all the research projects conducted within their community, and Tiwi people have participated in many surveys, treatment programs, and controlled clinical trials. Tiwi people pioneered governance of research processes with formation of the Tiwi Health Board, founded their own Scientific

Research and Advisory Group and Committee. Tiwi advocated for and house the first satellite dialysis unit in the NT. Tiwi people encouraged application of genetics research, to illuminate their origins, migrations, customs, relationships, and health issues, and they are proud to be contributors on the world stage.

We also investigated the relative degree of genetic relatedness between the Tiwi and other populations using PCA and a set of polymorphic SNPs identified in the Tiwi GWAS study. We observed that the Tiwi samples clustered in a single group, that was well-separated and distant from all of the other populations tested. As an example of the relative genetic differences observed, the approximate genetic distance separating the closest population to the Tiwi, Mexicans, was similar to that separating Mexicans from Europeans. Collectively, the analysis is strongly suggestive of the Tiwi's long existence as a genetically distinct and isolated population. These data have been discussed with members of the Tiwi Land Council and has elicited a large degree of interest, as it is pertinent to the story of the origins of the Tiwi.

ETHICS STATEMENT

In the 1990s, in close consultation with the Tiwi Land Council, the protocol was approved by the institutional ethics committee and written informed consent was obtained from all participants. In 2013–2014, data were collected on a second cohort of 497 Tiwi Islander study participants. This study was carried out in accordance with the recommendations of the Human Ethics Committee (Tasmania) Network ethics reference number H0012832, and in conjunction with all institutional ethics committees of the research team and the Human Research Ethics Committee of the Northern Territory Department of Health and Menzies School of Health Research. Written informed consent was obtained from all participants.

AUTHOR CONTRIBUTIONS

RT carried out the statistical analyses and wrote the manuscript. TT contributed to the statistical analyses. WH, JM, and SF conceived the study. MJ contributed clinical knowledge. BM and GB carried out biochemical analyses and genotyping. LW carried out the DNA amplification. All authors read and approved the final manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene. 2019.00330/full#supplementary-material

REFERENCES

Alves, R. J., Oppermann, K., Schein, L. E., and Pegas, K. L. (2012). A case of late-onset oligomeganephronia. *J. Bras. Nefrol.* 34, 392–394. doi: 10.5935/0101-2800.20120030

Anderson, C. E., Wallerstein, R., Zamerowski, S. T., Witzleben, C., Hoyer, J. R., Gibas, L., et al. (1997). Ring chromosome 4 mosaicism coincidence of oligomeganephronia and signs of Seckel syndrome. *Am. J. Med. Genet.* 72, 281–285. doi: 10.1002/(SICI)1096-8628(19971031)72:3<281::AID-AJMG5>3.0.CO;2-U

Arsenović, A., Bukvić, D., Trbojević, S., Marić, I., and Djukanović, L. (2005). Detection of renal dysfunctions in family members of patients with Balkan endemic nephropathy. *Am. J. Nephrol.* 25, 50–54. doi: 10.1159/000084105

Austwick, P. K. (1984). Fungal nephrotoxins. *Vet. Res. Commun.* 7, 145–154. doi: 10.1007/BF02228610

Bingham, C., Bulman, M. P., Ellard, S., Allen, L. I., Lipkin, G. W., Hoff, W. G., et al. (2001). Mutations in the hepatocyte nuclear factor-1beta gene are associated with familial hypoplastic glomerulocystic kidney disease. *Am. J. Hum. Genet.* 68, 219–224. doi: 10.1086/316945

Deshmukh, N., Petroczi, A., Barker, J., Szekely, A. D., Hussain, I., Naughton, D. P., et al. (2010). Potentially harmful advantage to athletes: a putative connection between UGT2B17 gene deletion polymorphism and renal disorders with prolonged use of anabolic androgenic steroids. *Subst. Abuse Treat Prev. Policy* 5:7. doi: 10.1186/1747-597X-5-7

Douglas-Denton, R. N., McNamara, B. J., Hoy, W. E., Hughson, M. D., and Bertram, J. F. (2006). 'Does nephron number matter in the development of kidney disease? *Ethn. Dis.* 16(2 Suppl. 2), S2–S40.

Duffy, D. L., McDonald, S. P., Hayhurst, B., Panagiotopoulos, S., Smith, T. J., Wang, X. L., et al. (2016). Familial aggregation of albuminuria and arterial hypertension in an Aboriginal Australian community and the contribution of variants in ACE and TP53. *BMC Nephrol.* 17:183. doi: 10.1186/s12882-016-0396-2

Georgas, K., Bowles, J., Yamada, T., Koopman, P., and Little, M. H. (2000). Characterisation of Crim1 expression in the developing mouse urogenital tract reveals a sexually dimorphic gonadal expression pattern. *Dev. Dyn.* 219, 582–587. doi: 10.1002/1097-0177(2000)9999:9999<::AID-DVDY1072>3.0.CO;2-I

Goodfellow, A. M., Hoy, W. E., Sriprakash, K. S., Daly, M. J., Reeve, M. P., Mathews, J. D., et al. (1999). Proteinuria is associated with persistence of antibody to streptococcal M protein in Aboriginal Australians. *Epidemiol. Infect.* 122, 67–75. doi: 10.1017/S0950268898001812

Hoy, W. E. (1996). 'Renal disease in Australian aboriginals'. *Med. J. Aust.* 165, 126–127. doi: 10.5694/j.1326-5377.1996.tb124882.x

Hoy, W. E., Bertram, J. F., and Hughson, M. D. (2014). Nephron hypertrophy and glomerulosclerosis in normal donor kidneys. *Clin. J. Am. Soc. Nephrol.* 9, 1832–1834. doi: 10.2215/CJN.08680814

Hoy, W. E., Hughson, M. D., Singh, G. R., Douglas-Denton, R., and Bertram, J. F. (2006a). Reduced nephron number and glomerulomegaly in Australian Aborigines: a group at high risk for renal disease and hypertension. *Kidney Int.* 70, 104–110.

Hoy, W. E., Kondalsamy-Chennakesavan, S., McDonald, S., and Wang, Z. (2006b). Renal disease, the metabolic syndrome, and cardiovascular disease. *Ethn. Dis.* 16, S2–S46.

Hoy, W. E., Kopp, J. B., Mott, S. A., and Winkler, C. A. (2017). Absence of APOL1 risk alleles in a remote living Australian Aboriginal group with high rates of CKD, hypertension, diabetes, and cardiovascular disease. *Kidney Int.* 91, 990. doi: 10.1016/j.kint.2016.11.031

Hoy, W. E., Wang, Z., VanBuynder, P., Baker, P. R., and Mathews, J. D. (2001). The natural history of renal disease in Australian Aborigines. Part 1. Changes in albuminuria and glomerular filtration rate over time. *Kidney Int.* 60, 243–248. doi: 10.1046/j.1523-1755.2001.00792.x

Hoy, W. E., White, A. V., Dowling, A., Sharma, S. K., Bloomfield, H., Tipiloura, B. T., et al. (2012). Post-streptococcal glomerulonephritis is a strong risk factor for chronic kidney disease in later life. *Kidney Int.* 81, 1026–1032. doi: 10.1038/ki.2011.478

Hoy, W. E., White, A. V., Tipiloura, B., Singh, G. R., Sharma, S., Bloomfield, H., et al. (2016). The influence of birthweight, past poststreptococcal glomerulonephritis and current body mass index on levels of albuminuria in young adults: the multideterminant model of renal disease in a remote Australian Aboriginal population with high rates of renal disease and renal failure. *Nephrol. Dial. Transplant* 31, 971–977. doi: 10.1093/ndt/gfu241

Jakobsson, J., Ekstrom, L., Inotsume, N., Garle, M., Lorentzon, M., Ohlsson, C., et al. (2006). 'Large differences in testosterone excretion in Korean and Swedish men are strongly associated with a UDP-glucuronosyl transferase 2B17 polymorphism'. *J. Clin. Endocrinol. Metab.* 91, 687–693. doi: 10.1210/jc.2005-1643

Junutula, J. R., De Maziere, A. M., Peden, A. A., Ervin, K. E., Advani, R. J., van Dijk, S. M., et al. (2004). 'Rab14 is involved in membrane trafficking between the Golgi complex and endosomes'. *Mol. Biol. Cell* 15, 2218–2229. doi: 10.1091/mbc.e03-10-0777

Kowal, E., and Anderson, I. (2012). *Genetic Research in Aboriginal and Torres Strait Islander Communities: Continuing the Conversation.* Carlton, VIC: Lowitja Institute.

Lefcheck, J. S. (2016). 'piecewiseSEM: piecewise structural equation modelling in r for ecology, evolution, and systematics'. *Methods Ecol. Evol.* 7, 573–579. doi: 10.1111/2041-210X.12512

Levey, A. S., Bosch, J. P., Lewis, J. B., Greene, T., Rogers, N., Roth, D., et al. (1999). A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Modification of diet in renal disease study group. *Ann. Intern. Med.* 130, 461–470. doi: 10.7326/0003-4819-130-6-199903160-00002

Long, D. A., Joannou, M. K., Price, K. L., Baradez, C. D., Huang, J. L., Papakrivopoulou, E., et al. (2013). Albuminuria is associated with too few glomeruli and too much testosterone. *Kidney Int.* 83, 1118–1129. doi: 10.1038/ki.2013.45

Manalich, R, Reyes, L., Herrera, M., Melendi, C., and Fundora, I. (2000). Relationship between weight at birth and the number and size of renal glomeruli in humans: a histomorphometric study. *Kidney Int.* 58, 770–773. doi: 10.1046/j.1523-1755.2000.00225.x

McDonald, S. P., Hoy, W. E., Maguire, G. P., Duarte, N. L., Wilcken, D. E., and Wang, X. L. (2002). The p53Pro72Arg polymorphism is associated with albuminuria among aboriginal Australians. *J. Am. Soc. Nephrol.* 13, 677–683.

Nezu, M., Suzuki, N., and Yamamoto, M. (2017). Targeting the KEAP1-NRF2 system to prevent kidney disease progression. *Am. J. Nephrol.* 45, 473–483. doi: 10.1159/000475890

Nystrom, J., Hultenby, K., Ek, S., Sjolund, J., Axelson, H., Jirstrom, K., et al. (2009). 'CRIM1 is localized to the podocyte filtration slit diaphragm of the adult human kidney'. *Nephrol. Dial. Transplant* 24, 2038–2044. doi: 10.1093/ndt/gfn743

Park, J., Shrestha, R., Qiu, C., Kondo, A., Huang, S., Werth, M., et al. (2018). 'Single-cell transcriptomics of the mouse kidney reveals potential cellular targets of kidney disease'. *Science* 360, 758–763. doi: 10.1126/science.aar2131

Park, S. H., and Chi, J. G. (1993). Oligomeganephronia associated with 4p deletion type chromosomal anomaly. *Pediatr. Pathol.* 13, 731–740. doi: 10.3109/15513819309048260

Patterson, N., Price, A. L., and Reich, D. (2006). 'Population structure and eigenanalysis. *PLoS Genet.* 2:e190. doi: 10.1371/journal.pgen.0020190

Phua, Y. L., Martel, N., Pennisi, D. J., Little, M. H., and Wilkinson, L. (2013). 'Distinct sites of renal fibrosis in Crim1 mutant mice arise from multiple cellular origins'. *J. Pathol.* 229, 685–696. doi: 10.1002/path.4155

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795

Spencer, J. L., Silva, D. T., Snelling, P., and Hoy, W. E. (1998). An epidemic of renal failure among Australian Aboriginals. *Med. J. Aust.* 168, 537–541. doi: 10.5694/j.1326-5377.1998.tb139080.x

Stefanovic, V., and Polenakovic, M. (2009). Fifty years of research in Balkan endemic nephropathy: where are we now? *Nephron Clin. Pract.* 112, c51–c56. doi: 10.1159/000213081

Thomson, R., and McWhirter, R. (2017). "Adjusting for familial relatedness in the analysis of GWAS data," in *Bioinformatics*, ed. J. M. Keith (New York, NY: Humana Press).

Thorisson, G. A., Smith, A. V., Krishnan, L., and Stein, L. D. (2005). The international hapmap project web site. *Genome Res.* 15, 1592–1593. doi: 10.1101/gr.4413105

Tiwi Land Council (2018). *Timeline: Dreamtime to 1978.* Available at: http://tiwilandcouncil.com/ (accessed April 5, 2019).

Van Buynder, P. G., Gaggin, J. A., Martin, D., Pugsley, D., and Mathews, J. D. (1992). 'Streptococcal infection and renal disease markers in Australian aboriginal children'. *Med. J. Aust.* 156, 537–540.

Vivante, A., and Hildebrandt, F. (2016). Exploring the genetic basis of early-onset chronic kidney disease. *Nat. Rev. Nephrol.* 12, 133–146. doi: 10.1038/nrneph.2015.205

Wang, Z., and Hoy, W. E. (2003). Hypertension, dyslipidemia, body mass index, diabetes and smoking status in Aboriginal Australians in a remote community. *Ethn. Dis.* 13, 324–330.

Wilhelm-Bals, A., Chehade, H., Girardin, E., Gonzalez, E., and Parvex, P. (2014). Premature kidneys and oligonephronia: long term repercussion. *Rev. Med. Suisse* 10, 438–441.

Wilkinson, L., Kurniawan, N. D., Phua, Y. L., Nguyen, M. J., Li, J., Galloway, G. J., et al. (2012). 'Association between congenital defects in papillary outgrowth and functional obstruction in Crim1 mutant mice'. *J. Pathol.* 227, 499–510. doi: 10.1002/path.4036

Xia, K., Shabalin, A. A., Huang, S., Madar, V., Zhou, Y. H., Wang, W., et al. (2012). 'seeQTL: a searchable database for human eQTLs'. *Bioinformatics* 28, 451–452. doi: 10.1093/bioinformatics/btr678

Yang, J., Lee, S. H., Goddard, M. E., and Visscher, P. M. (2011). GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* 88, 76–82. doi: 10.1016/j.ajhg.2010.11.011

You, J., Hoy, W., Zhao, Y., Beaver, C., and Eagar, K. (2002). 'End-stage renal disease in the Northern territory: current and future treatment costs'. *Med. J. Aust.* 176, 461–465.

frontiers
in Genetics

Deleterious Impact of a Novel *CFH* Splice Site Variant in Atypical Hemolytic Uremic Syndrome

Ria Schönauer[1], Anna Seidel[1], Maik Grohmann[2], Tom H. Lindner[1], Carsten Bergmann[2] and Jan Halbritter[1]*

[1] Division of Nephrology, University Hospital Leipzig, Leipzig, Germany, [2] Center for Human Genetics, Bioscientia, Ingelheim, Germany

Atypical hemolytic uremic syndrome (aHUS) is a heterogeneous disorder characterized by microangiopathic hemolytic anemia (MAHA), thrombocytopenia, and acute kidney injury (AKI). In about 50% of cases, pathogenic variants in genes involved in the innate immune response including complement factors complement factor H (CFH), CFI, CFB, C3, and membrane co-factor protein (MCP/CD46) put patients at risk for uncontrolled activation of the alternative complement pathway. As aHUS is characterized by incomplete penetrance and presence of additional triggers for disease manifestation, genetic variant interpretation is challenging and streamlined functional variant evaluation is urgently needed. Here, we report the case of a 27-year-old female without previous medical and family history who presented with confusion, petechial bleeding, and anuric AKI. Kidney biopsy revealed glomerular thrombotic microangiopathy (TMA). Targeted next generation sequencing identified a paternally transmitted novel heterozygous splice site variant in the *CFH* gene [c.3134-2A>G; p.Asp1045_Thr1053del] which resulted in a partial in-frame deletion of exon 20 transcript as determined by cDNA analysis. On the protein level, the concomitant loss of 9 amino acids in the short consensus repeat (SCR) domains 17 and 18 of CFH includes a highly conserved cysteine residue, which is assumed to be essential for proper structural folding and protein function. Treatment with steroids, plasmapheresis, and the complement inhibitor eculizumab led to complete hematological and clinical remission after several months and stable renal function up to 6 years later. In conclusion, genetic investigation for pathogenic variants and evaluation of their functional impact, in particular in the case of splice site variants, is clinically relevant and enables not only better molecular understanding but helps to guide therapy with complement inhibitors.

Keywords: complement factor H, atypical hemolytic uremic syndrome, splice site variant, short consensus repeat 18, eculizumab, CFH, aHUS

INTRODUCTION

Atypical hemolytic uremic syndrome (aHUS; MIM# 235400/612922/612923/612924/615008/6129 25/612926) is a disease complex characterized by the uncontrolled over-activation of the alternative pathway of the complement system (AP). Clinical symptoms include microangiopathic hemolytic anemia (MAHA), thrombocytopenia, and acute kidney injury (AKI) (Maga et al., 2010). In contrast

to the "typical" form of HUS initiated by infection with Shiga-toxigenic *Escherichia coli* (STEC) and characterized by diarrheal illness (D+), atypical forms (D−) have a poorer prognosis and progress to end-stage renal disease (ESRD) in 30–60% of all cases (Loirat et al., 2008; Westra et al., 2010). Pathogenic variants in AP-regulating genes including *CFH*, *CFI*, or *MCP/CD46* are found in more than 50% of aHUS-patients and variants in *CFH* represent the most frequent genetic finding (Fremeaux-Bacchi et al., 2013).

The protein complement factor H (CFH) encoded by *CFH* plays a key role in the control of complement activation in fluid phase and on cell membranes thereby protecting self-surfaces from immune attacks. It exerts its effects by preventing the formation of the central complement protein C3b from C3 and C5b from C5 by accelerating the decay of C3 and C5 convertases (C3bBb and C3bBbC3b), respectively. Furthermore, it acts as a

co-factor of complement factor I (CFI, encoded by *CFI*) in the proteolytic inactivation of C3b (**Figure 1**; Merinero et al., 2018).

Complement factor H is a 155 kDa plasma protein that consists of 20 homologous short consensus repeat (SCR) domains of 60 amino acids length, each containing 4 conserved cystein residues that form two structure-defining disulfide-bridges. The region spanning SCR1-4 has been proposed to act as regulatory domain, mediating C3b binding in fluid phase, CFI interaction and convertase decay accelerating activity (Goicoechea de Jorge et al., 2013; Perkins et al., 2014; Clark and Bishop, 2015; de Vriese et al., 2015; Sepúlveda et al., 2016). Whereas SCR6, SCR7, and SCR12-14 seem to be additionally involved in glycosaminoglycan (GAG) or C3b binding, SCR19 and SCR20 predominantly serve as surface-binding domain, comprising the ability to interact with both GAGs and C3b at self-membranes. Therefore, this C-terminal area is assumed to play a key role in

FIGURE 1 | Influence of CFH on complement activation. CFH inhibits the conversion of C3 to C3b by accelerating the decay of C3 convertases and enhancing the inactivation of C3b and prevents conversion of C5 to C5b by accelerating the decay of C5 convertases. CFB: complement factor B; CFD: complement factor D; MAC: membrane attack complex.

FIGURE 2 | Phenotypical characterization and treatment. **(A)** Kidney biopsy showing signs of pre- and intraglomerular TMA (black arrows). Periodic Acid-Schiff (PAS) staining; 10× magnification. **(B)** Diagram illustrating course of serum (S)-creatinine and platelets during plasma exchange/dialysis and treatment with Eculizumab.

distinguishing between host and pathogenic cells. It has been proposed that mutations within the N-terminal part of the protein often result in uncontrolled complement activation in fluid phase leading to glomerulonephritis, whereas C-terminal mutations are more frequently associated with aHUS leading to defects in the recognition and activity at the endogenous endothelia resulting in thrombotic microangiopathy (TMA) (de Vriese et al., 2015). However, the majority of pathogenic *CFH* variants show incomplete penetrance and constitute predisposing factors lowering the threshold of aHUS/TMA disease manifestation. Upon presence of additional triggers, such as infection, pregnancy, or immunosuppressive drugs, patients with predisposing variants are prone to develop aHUS/TMA. Thus, pathogenic *CFH* variants are also associated with a significant risk for disease recurrence after renal transplantation. Under these circumstances, complement-targeting therapies

(e.g., eculizumab) in addition to the immunosuppressive regimen are able to provide better outcome after renal transplantation in patients with aHUS/TMA (Legendre et al., 2013; Münch et al., 2017). Therefore, genetic testing for pathogenic variants within the AP regulatory genes is indispensable to establish responsible prognosis and treatment strategies, particularly in the context of renal transplantation.

CASE PRESENTATION

A 27-year-old female without prior medical and family history presented with nausea, confusion, petechial bleeding, and anuric AKI necessitating admission on intensive care unit and immediate initiation of renal replacement therapy by continuous veno-venous hemofiltration (CVVH). Laboratory examination

FIGURE 3 | Genetic analysis and *in silico* evaluation of *CFH* c.3134-2A>G; p.Asp1045_Thr1053del. **(A)** Exon structure of the *CFH* gene indicating the position of the identified variant (red arrow and letters) at the splice acceptor site of the intron (gray letters) 19/exon (black letters) 20 boundary. The variant activates an exonic splice acceptor and leads to a partial loss of exon 20 from the RNA transcript (red italic letters). **(B)** Domain structure of CFH (dark gray: regulatory domain; brown: proteoglycan binding; blue: surface binding domain); the variant results in an in-frame deletion of 9 amino acids of SCR17 and SCR18 including Cys[1048] (red italic letters) leading to the loss of a structure-determining disulfide bond (red-dotted line). **(C)** Model of the C-terminal domains SCR18-20 of CFH (PDB: 3SW0) modified from Morgan et al. (2012), indicating the deleted sequence part (red color) (Morgan et al., 2012). SCR, short consensus repeat; C3b, C3 binding site; PA, polyanion binding site.

revealed thrombocytopenia and Coombs-negative hemolytic anemia (hemoglobin 3.7 mmol/L, platelets 118×10^9/L, haptoglobin < 0.1 g/L, fragmented erythrocytes 1%, lactate dehydrogenase > 13 mmol/L) but normal ADAMTS13 levels and activity (>50%) and absence of ADAMTS13 autoantibodies. Complement analysis yielded reduced levels for C3 (0.5 – 0.7 g/l; reference range: 0.8 – 1.6) and normal levels for C4. Percutaneous kidney biopsy evidenced signs of acute and non-acute preglomerular and intraglomerular TMA (**Figure 2A**). The patient was initially treated with intravenous glucocorticoids and 6 weeks of plasma exchange. Only after addition of the C5-inhbitor eculizumab (4 weeks of induction followed by 6 months of maintenance) the patient's condition slowly resolved with complete hematological and clinical remission, accompanied by gradual recovery of kidney function over several months, allowing to terminate hemodialysis (**Figure 2B**). Long-term follow up over 6 years showed no relapse and stable renal function at CKD stage 3 (CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration) 40–50 ml/min/1.73 m^2) without continued maintenance therapy.

Targeted next-generation sequencing using a gene panel consisting of 14 aHUS-associated genes (including *ADAMTS13, C3, CFB, CFD, CFH, CFHR1, CFHR2, CFHR3, CFHR5, CFI, DGKE, MCP, MMACHC,* and *THBD*) identified a novel heterozygous canonical splice site variant in the *CFH* gene (c.3134-2A>G; NM_000186.3), which was absent from SNP-databases (gnomAD). In addition, the following aHUS-risk alleles were detected: *CFH*-H3 (heterozygous), *MCP*-H2 (homozygous), and *CFHR1**B (homozygous). Copy number variations (CNV) of *CFH, CFHR1-3* and *CFHR5* were excluded by multiplex ligation-dependent probe amplification (MLPA). Segregation analysis yielded paternal transmission despite negative family history. The variant is located within the splice acceptor at the boundary between intron 19 and exon 20 (**Figure 3A** and **Supplementary Figure 1**). Sanger sequencing

of patient cDNA indicated, that as a consequence of the variant c.3134-2A>G; r.3134_3160del, activation of an intra-exonic splice acceptor site results in an in-frame deletion of the first 27 base pairs of exon 20 transcript (**Supplementary Figure 1**). On the protein level, the variant leads to a loss of the very C-terminal amino acid of SCR17 and the eight N-terminal amino acids of SCR18 p.Asp1045_Thr1053del (**Figure 3B**). Importantly, the missing sequence includes Cys[1048], which forms one of the two highly conserved intra-domain disulfide bonds with Cys[1091] that is thought to be essential for correct domain folding (**Figure 3C**). In conclusion, this leads to a variant classification as *pathogenic* according to the American College of Medical Genetics (ACMG) (Richards et al., 2015).

Written informed consent was obtained from the participant for the publication of this report.

DISCUSSION

We here report a novel aHUS-associated *CFH* splice-site variant [c.3134-2A>G; p.Asp1045_Thr1053del] and predict its potential impact on the protein function by an in-frame deletion of 9 amino acids mainly affecting SCR18 of CFH. Although the activating disease trigger could not be clearly established in this case, we were able to detect the underlying genetic predisposition which allowed us to make the diagnosis of complement-mediated aHUS/TMA. In contrast to non-complement mediated forms (e.g., *DGKE, MMACHC*), patients with *CFH*-related aHUS are amenable to treatment with the C5-inhibitor eculizumab.

Complement factor H represents the key regulating protein of the alternative pathway of complement activation, required for fine-tuning of complement activity and differentiation between self-surfaces and pathogenic molecules. Consequently, the majority of pathogenic variants within the *CFH* gene

FIGURE 4 | Statistical evaluation of aHUS-associated *CFH* variants according to HGMD 2018.4. **(A)** Frequency of aHUS-associated variant subtypes. **(B)** Mapping of total (light colors) and aHUS-associated (dark colors) missense/non-sense variants onto distinct SCR domains of CFH. SCR 18 is highlighted in red. SCR, short consensus repeat.
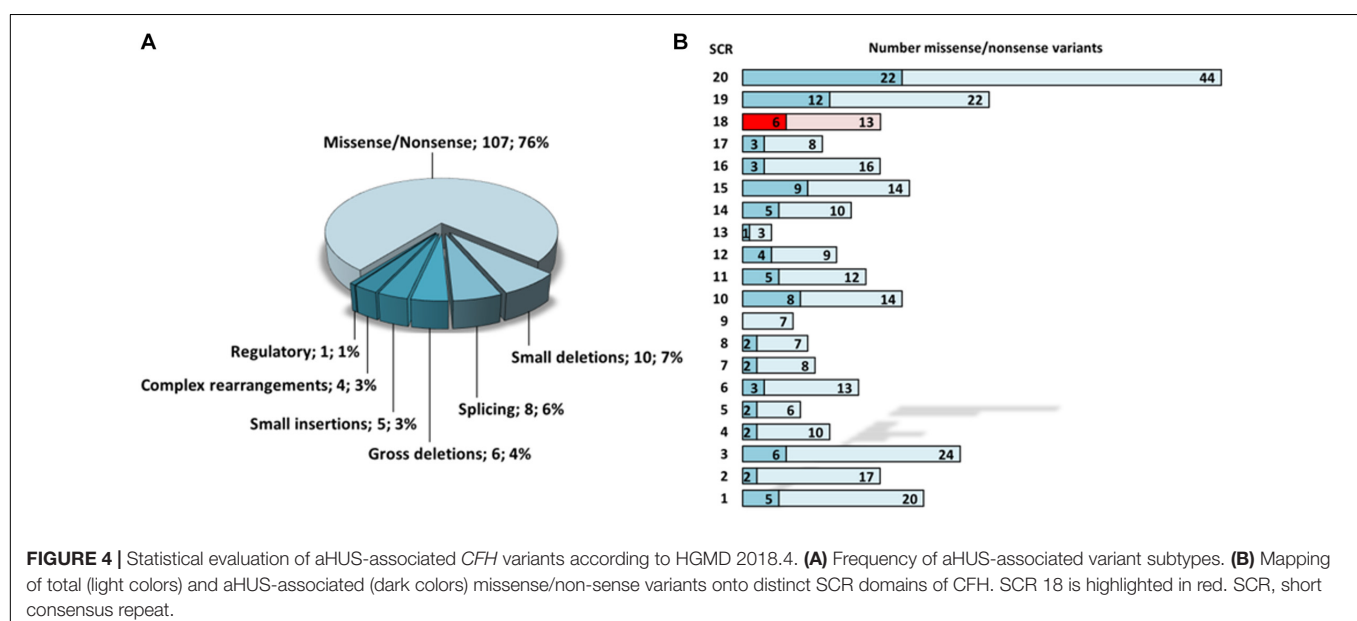
lead to phenotypes related to disturbed immune processes including aHUS/TMA, C3-glomerulopathy (including dense deposit disease and C3 glomerulonephritis) or age-related macular degeneration (AMD) (Boon et al., 2008; de Córdoba and de Jorge, 2008; Morgan et al., 2012).

To date, a total of 346 disease-associated variants were reported for *CFH*, 141 of which refer to the phenotypical term "aHUS" (HGMD 2018.4). Of those, the vast majority (76%) represents missense or non-sense variants, whereas indels, complex rearrangements (such as *CFH/CFHR1* hybrid genes), and splice site variants account for less than 10% each (**Figure 4A**). Regarding the domain distribution of the reported missense and non-sense variants, exemplifying variants with known consequences on the protein level, the region spanning SCR19 (total = 22, aHUS-associated = 12) and SCR20 (total = 44, aHUS-associated = 22) clearly represents a mutational hotspot (**Figure 4B**). Thus, there is emerging evidence for the speculation that variants responsible for the aHUS phenotype might predominantly affect the C-terminal surface recognition domain (de Vriese et al., 2015; Bu et al., 2018).

To effectively allow cell protection, the CFH-molecule has to adopt a U-shaped structure that enables the interaction of its N- and C-terminal domains (SCR 1-4 and SCR19-20) with cell surface-associated C3b and GAG containing recognition sites (Morgan et al., 2012). Previously, it has been shown that SCR18 is connected to SCR19 by a flexible linker, whereas SCR19 and SCR20 are rigidly associated, allowing a kink-like conformational rearrangement of the CFH C-terminal domains with presumably functional implications. To date, six out of 13 known missense and non-sense variants and one of the variants categorized as "small deletion" [c.3269delT; p.(Met1090Serfs*3)] within SCR18 were reported to be associated with aHUS. In addition, we now identified a pathogenic splice-site variant leading to the deletion of a 9 amino acid sequence mainly located within SCR18. Unfortunately, the majority of newly detected splice-site variants are not routinely subjected to further genetic or functional investigation. Thus, deeper understanding of the disease-causing effects arising from distinct transcript rearrangements remains most often elusive. For *CFH*, 13 splice-site variants have been described so far (total = IVS 2–6, 8, 16, 18, 19, 21), eight of which were associated with aHUS. Due to the domain-oriented exonic structure of the *CFH* gene, splice-site variants can affect the previous or the following SCR domain or even lead to a frame-shift resulting in a premature translational stop. The splice-site of intron 19 affected in the patient described here, was already found to be mutated in two previous reports. The variant c.3134-5T>A was found in a genetic screening of a patient cohort known to suffer from renal failure, malignant hypertension, and hypertension-associated TMA (Larsen et al., 2018). Interestingly, this exchange is located only 3bp upstream within the same splice acceptor site and *in silico* prediction tools (MutationTaster, PolyPhen2 and SIFT) resulted in mixed ratings. Additionally, the variant c.3133+1G>A that is located within the splice donor site of intron 19 was reported in a patient diagnosed with aHUS, however, only *in silico* analyses (Human Splicing Finder) were conducted suggesting abrogation of the splice donor.

By means of functional analysis, we were able to determine that the defective splice site in our case results in the activation of an intra-exonic splice acceptor, leading to the deletion of an amino acid stretch within SCR18 that contains one of the four conserved cysteine residues of the consensus sequence known to be required for the establishment of the two domain structure-defining disulfide bonds. Loss of these cysteines is highly likely to result in incorrect folding and can be assumed to severely affect protein function. 41 out of 346 total known (12%) and 23 out of 141 aHUS-associated (16%) variants (missense/non-sense) affect a cysteine residue further underlining their obvious functional importance. An accumulation of abrogated cysteine residues was also reported when considering the CFH/ membrane cofactor protein (MCP) consensus sequence (Rodriguez et al., 2014). A defective structural organization of SCR18 probably prevents the required conformational flexibility that is needed for a correct orientation of SCR19 and SCR20 which could negatively affect recognition of interaction partners and host surfaces.

CONCLUSION

In summary, we detected and analyzed a novel pathogenic *CFH* splice site variant in a patient with aHUS/TMA, which probably leads to an incorrect protein fold impairing inhibitory control of AP activity. Incomplete penetrance demonstrated by the clinically asymptomatic father underlines the influence of an additional disease trigger for aHUS manifestation. Genetic assessment and functional variant analysis, particularly in the case of detected splice site variants, helps guiding patients' treatment and assessing the probability of disease recurrence.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the institutional review board of the University of Leipzig with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by institutional review board of the University of Leipzig.

AUTHOR CONTRIBUTIONS

RS generated and analyzed the data (genetic) and wrote the manuscript. AS generated the data (clinical) and edited the manuscript. MG generated the data (genetic) and edited the manuscript. TL contributed the clinical data. CB generated the data (genetic) and edited the manuscript. JH initiated and supervised functional evaluation, contributed clinical data, and edited the manuscript.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.00465/full#supplementary-material

FIGURE S1 | Experimental analysis of the splice-site variant c.3134-2A>G. **(A)** Agarose gel electrophoresis showing PCR fragments of *CFH* cDNA derived from primary dermal fibroblast RNA (WT, wild type control; PAT, patient); the additional splice product is depicted with an white arrow. **(B)** Sequence of the *CFH* cDNA showing the consequences of the splice site variant; exons 18 – 20 are indicated as upper case letters (wild type sequence: alternating black and blue, missing bases: red) and intron 19 as lower case letters (gray: wildtype sequence, red: base exchange c.3134-2A>G); primer binding sites are shown in gray and sequencing results in yellow and light blue. **(C)** Chromatogram of the additional PCR-fragment (reverse sequence); assigned positions of exon 19 and exon 20 are depicted in light blue and yellow, respectively, and the missing part of exon 20 is shown in red.

REFERENCES

Boon, C. J., Klevering, B. J., Hoyng, C. B., Zonneveld-Vrieling, M. N., Nabuurs, S. B., Blokland, E., et al. (2008). Basal laminar drusen caused by compound heterozygous variants in the CFH gene. *Am. J. Hum. Genet.* 82, 516–523. doi: 10.1016/j.ajhg.2007.11.007

Bu, F., Zhang, Y., Wang, K., Borsa, N. G., Jones, M. B., Taylor, A. O., et al. (2018). Genetic Analysis of (400) patients refines understanding and implicates a new gene in atypical hemolytic uremic syndrome. *J. Am. Soc. Nephrol.* 29, 2809–2819. doi: 10.1681/ASN.2018070759

Clark, S. J., and Bishop, P. N. (2015). Role of Factor H and related proteins in regulating complement activation in the macula, and relevance to age-related macular degeneration. *J. Clin. Med.* 4, 18–31. doi: 10.3390/jcm4010018

de Córdoba, S. R., and de Jorge, E. G. (2008). Translational mini-review series on complement factor H: genetics and disease associations of human complement factor H. *Clin. Exp. Immunol.* 151, 1–13. doi: 10.1111/j.1365-2249.2007. 03552.x

de Vriese, A. S., Sethi, S., van Praet, J., Nath, K. A., and Fervenza, F. C. (2015). Kidney disease caused by dysregulation of the complement alternative pathway: an etiologic approach. *J. Am. Soc. Nephrol.* 26, 2917–2929. doi: 10.1681/ASN. 2015020184

Fremeaux-Bacchi, V., Fakhouri, F., Garnier, A., Bienaimé, F., Dragon-Durey, M. A., Ngo, S., et al. (2013). Genetics and outcome of atypical hemolytic uremic syndrome: a nationwide French series comparing children and adults. *Clin. J. Am. Soc. Nephrol.* 8, 554–562. doi: 10.2215/CJN.04760512

Goicoechea de Jorge, E., Caesar, J. J., Malik, T. H., Patel, M., Colledge, M., Johnson, S., et al. (2013). Dimerization of complement factor H-related proteins modulates complement activation in vivo. *Proc. Natl. Acad. Sci. U.S.A.* 110, 4685–4690. doi: 10.1073/pnas.1219260110

Larsen, C. P., Wilson, J. D., Best-Rocha, A., Beggs, M. L., and Hennigar, R. A. (2018). Genetic testing of complement and coagulation pathways in patients with severe hypertension and renal microangiopathy. *Modern Pathol.* 31, 488–494. doi: 10.1038/modpathol.2017.154

Legendre, C. M., Licht, C., Muus, P., Greenbaum, L. A., Babu, S., Bedrosian, C., et al. (2013). Terminal complement inhibitor eculizumab in atypical hemolytic-uremic syndrome. *N. Engl. J. Med.* 368, 2169–2181. doi: 10.1056/ NEJMoa1208981

Loirat, C., Noris, M., and Fremeaux-Bacchi, V. (2008). Complement and the atypical hemolytic uremic syndrome in children. *Pediatr. Nephrol.* 23, 1957–1972. doi: 10.1007/s00467-008-0872-4

Maga, T. K., Nishimura, C. J., Weaver, A. E., Frees, K. L., and Smith, R. J. (2010). Mutations in alternative pathway complement proteins in American patients with atypical hemolytic uremic syndrome. *Hum. Mutat.* 31, E1445–E1460. doi: 10.1002/humu.21256

Merinero, H. M., García, S. P., García-Fernández, J., Arjona, E., Tortajada, A., and Rodríguez de Córdoba, S. (2018). Complete functional characterization of disease-associated genetic variants in the complement factor H gene. *Kidney Int.* 93, 470–481. doi: 10.1016/j.kint.2017.07.015

Morgan, H. P., Mertens, H. D. T., Guariento, M., Schmidt, C. Q., Soares, D. C., Svergun, D. I., et al. (2012). Structural analysis of the C-terminal region (modules 18-20) of complement regulator factor H (FH). *PloS One* 7:e32187. doi: 10.1371/journal.pone.0032187

Münch, J., Bachmann, A., Grohmann, M., Mayer, C., Kirschfink, M., Lindner, T. H., et al. (2017). Effective immunosuppressive management with belatacept and eculizumab in post-transplant aHUS due to a homozygous deletion of CFHR1/CFHR3 and the presence of CFH antibodies. *Clin. Kidney J.* 10, 742–746. doi: 10.1093/ckj/sfx053

Perkins, S. J., Fung, K. W., and Khan, S. (2014). Molecular Interactions between complement Factor H and its heparin and heparan sulfate ligands. *Front. Immunol.* 5:126. doi: 10.3389/fimmu.2014. 00126

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., 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. *Genet. Med.* 17, 405–424. doi: 10.1038/gim.2015.30

Rodriguez, E., Rallapalli, P. M., Osborne, A. J., and Perkins, S. J. (2014). New functional and structural insights from updated mutational databases for complement factor H, Factor I, membrane cofactor protein and C3. *Biosci. Rep.* 34:e00146. doi: 10.1042/BSR20140117

Sepúlveda, R. A., Tagle, R., and Jara, A. (2016). Novel Complement Factor H mutation, a case report of atypical hemolytic uremic syndrome. *CRIM* 4, 263–265. doi: 10.5430/crim.v4n2p13

Westra, D., Volokhina, E., van der Heijden, E., Vos, A., Huigen, M., Jansen, J., et al. (2010). Genetic disorders in complement (regulating) genes in patients with atypical haemolytic uraemic syndrome (aHUS). *Nephrol. Dial. Transplant.* 25, 2195–2202. doi: 10.1093/ndt/gfq010

Genetics of Chronic Kidney Disease Stages Across Ancestries: The PAGE Study

Bridget M. Lin[1], Girish N. Nadkarni[2,3], Ran Tao[4,5], Mariaelisa Graff[6], Myriam Fornage[7], Steven Buyske[8], Tara C. Matise[8], Heather M. Highland[6], Lynne R. Wilkens[9], Christopher S. Carlson[10,11], S. Lani Park[12], V. Wendy Setiawan[12], Jose Luis Ambite[13], Gerardo Heiss[6], Eric Boerwinkle[7], Dan-Yu Lin[1], Andrew P. Morris[14,15], Ruth J. F. Loos[2,16], Charles Kooperberg[10], Kari E. North[6], Christina L. Wassel[17†] and Nora Franceschini[6*†]

[1] Department of Biostatistics, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States, [2] Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, United States, [3] Division of Nephrology, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, United States, [4] Department of Biostatistics, Vanderbilt University Medical Center, Nashville, TN, United States, [5] Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, TN, United States, [6] Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States, [7] The University of Texas Health Science Center at Houston, Houston, TX, United States, [8] Department of Genetics, Rutgers University, Piscataway, NJ, United States, [9] Epidemiology Program, University of Hawaii Cancer Center, Honolulu, HI, United States, [10] Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, United States, [11] Department of Epidemiology, University of Washington, Seattle, WA, United States, [12] Department of Preventive Medicine, University of Southern California, Los Angeles, CA, United States, [13] Information Sciences Institute, University of Southern California, Marina del Rey, CA, United States, [14] Department of Biostatistics, University of Liverpool, Liverpool, United Kingdom, [15] Wellcome Centre for Human Genetics, University of Oxford, Oxford, United Kingdom, [16] Mindich Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, NY, United States, [17] Applied Sciences, Premier, Inc., Charlotte, NC, United States

Background: Chronic kidney disease (CKD) is common and disproportionally burdens United States ethnic minorities. Its genetic determinants may differ by disease severity and clinical stages. To uncover genetic factors associated CKD severity among high-risk ethnic groups, we performed genome-wide association studies (GWAS) in diverse populations within the Population Architecture using Genomics and Epidemiology (PAGE) study.

Methods: We assembled multi-ethnic genome-wide imputed data on CKD non-overlapping cases [4,150 mild to moderate CKD, 1,105 end-stage kidney disease (ESKD)] and non-CKD controls for up to 41,041 PAGE participants (African Americans, Hispanics/Latinos, East Asian, Native Hawaiian, and American Indians). We implemented a generalized estimating equation approach for GWAS using ancestry combined data while adjusting for age, sex, principal components, study, and ethnicity.

Results: The GWAS identified a novel genome-wide associated locus for mild to moderate CKD nearby *NMT2* (rs10906850, $p = 3.7 \times 10^{-8}$) that replicated in the United Kingdom Biobank white British ($p = 0.008$). Several variants at the *APOL1* locus were associated with ESKD including the *APOL1* G1 rs73885319 ($p = 1.2 \times 10^{-9}$). There was no overlap among associated loci for CKD and ESKD traits, even at

the previously reported *APOL1* locus ($p = 0.76$ for CKD). Several additional loci were associated with CKD or ESKD at *p*-values below the genome-wide threshold. These loci were often driven by variants more common in non-European ancestry.

Conclusion: Our genetic study identified a novel association at *NMT2* for CKD and showed for the first time strong associations of the *APOL1* variants with ESKD across multi-ethnic populations. Our findings suggest differences in genetic effects across CKD severity and provide information for study design of genetic studies of CKD in diverse populations.

Keywords: genetics, chronic kidney disease stages, genome-wide association studies, *APOL1*, end stage kidney disease, diverse populations, single nucleotide polymorphisms

INTRODUCTION

Chronic kidney disease (CKD) affects 15% of United States. adults and is a leading cause of death globally (Global Burden of Disease 2016 Causes of Death Collaborators, 2017). CKD is classified based on its causes, kidney function (estimated glomerular filtration rate, eGFR), and markers of kidney damage (Levin and Stevens, 2014). The risk of adverse outcomes and disability greatly increases in advanced CKD (Go et al., 2004; Saran et al., 2018). There is a high burden of CKD in non-European ancestry groups, including African Americans and Hispanics/Latinos (Collins et al., 2011). Genetic susceptibility explains in part ethnic differences in the burden of CKD, as illustrated by the African-ancestry *APOL1* G1 and G2 genotypes that contribute to increased CKD risk in individuals with African ancestry (Genovese et al., 2010; Kramer et al., 2017). Approximately 13% of African Americans carry two *APOL1* risk genotypes G1 (composed of two missense variants), G2 (a 6-base pair in-frame deletion) or are compound heterozygous of G1 and G2 genotypes. *APOL1* encodes an HDL cholesterol-binding protein but mechanisms related to CKD risk are unknown.

Few genome-wide association studies (GWAS) have been published for CKD as the primary outcome. These include studies of CKD progression such as the Chronic Renal Insufficiency Cohort Study (Parsa et al., 2017), and cause-specific CKD such as GWAS consortia that compared individuals with diabetic nephropathy with non-CKD diabetes controls (Iyengar et al., 2015; van Zuydam et al., 2018), in addition to studies of glomerular diseases (for example, IgA nephropathy, membranous nephropathy) (GWAS Catalog, 2019). CKD is a heterogeneous condition and its genetic determinants may vary by CKD severity, with more advanced CKD possibly reflecting stronger genetic risk. The genetic determinants of CKD severity have not been previously studied, particularly among individuals of diverse ancestries that vary in their genetic susceptibility.

Our recent research in diverse populations within the Continental Origins and Genetic Epidemiology Network (COGENT) Kidney Consortium identified 93 novel loci for eGFR, displaying homogenous effects across four major ancestries (Morris et al., 2019). Using Mendelian Randomization, we have shown that identified single nucleotide variants (SNVs) were causally related to a clinical diagnosis of CKD from

International Classification of Disease (ICD) diagnosis billing codes in the United Kingdom Biobank. These SNVs, identified from the *trans*-ethnic analyses and showing homogenous effects across ancestries, more likely capture CKD genetic risk across diverse populations.

To identify novel risk loci associated with CKD severity (stages), we assembled multi-ethnic data on cases (4,150 mild to moderate CKD and 1,105 ESKD) and non-CKD controls for samples up to 41,041 participants within the Population Architecture using Genomics and Epidemiology (PAGE) study (Bien et al., 2016). We also examined the association of eGFR-identified GWAS variants with CKD stages using genetic variants reported by the COGENT-Kidney consortium.

MATERIALS AND METHODS

PAGE Study Description

The PAGE consortium includes eligible minority participants from four studies. The Women's Health Initiative (WHI) is a long-term, prospective, multi-center, and multi-ethnic cohort study investigating post-menopausal women's health recruited from 1993 to 1998 at 40 centers across the United States (Anderson et al., 2003). WHI participants of European descent were excluded from analyses. The Hispanic Community Health Study/Study of Latinos (HCHS/SOL) is a multi-center study of Hispanic/Latinos with the goal of determining the role of acculturation in the prevalence and development of diseases relevant to Hispanic/Latino health. Starting in 2006, household sampling was used to recruit self-identified Hispanic/Latinos from four sites in San Diego, CA, Chicago, IL, Bronx, NY, and Miami, FL (Sorlie et al., 2010). The Multiethnic Cohort (MEC) is a population-based prospective cohort study recruiting men and women aged 45–75 from Hawaii and Los Angeles, California, in 1993–1996, that examines lifestyle risk factors and genetic susceptibility for cancer across five racial/ethnic groups (Kolonel et al., 2000). Only the African American, Japanese American, and Native Hawaiian participants for MEC were included in analyses. The BioMe BioBank is an Electronic Medical Record-linked biobank that integrates research data and clinical care information for consented patients at the Mount Sinai Medical Center, which serves diverse local

communities of upper Manhattan with broad health disparities. Recruitment began in 2007 and continues at 30 clinical care sites throughout New York City. BioMe participants were African American, Hispanic/Latino, primarily of Caribbean origin (36%), Caucasian (30%), and Others who did not identify with any of the available options (9%) (Nadkarni et al., 2014). All PAGE participants have provided informed consent. Up to 41,041 participants with kidney phenotype information were included in analyses.

Genotypes and Imputation

The genotyping and quality control (QC) in PAGE has been previously described (Bien et al., 2016). Briefly, 53,426 samples were genotyped centrally at the Center for Inherited Disease Research (CIDR) in the Johns Hopkins University using the Multi-Ethnic Genotyping Array (MEGA), Consortium version, consisting of 1,705,969 single nucleotide variants (SNV). Genotypes were called using the GenomeStudio version 2001.1, Genotyping Module 1.9.4, and GenTrain version 1.0. Extensive QC was performed to the combined genotyping data, which included checks for gender discrepancies, Mendelian inconsistencies, unexpected duplication, unexpected non-duplication, poor performance, or DNA mixture. Samples with identity issues, restricted consent, and duplicates were also removed (final sample 51,520 subjects). SNVs were filtered if they had a missing call rate $\geq$ 2%, more than 6 discordant calls in 988 study duplicates, > 1 Mendelian errors in 282 trios and 1,439 duos, a Hardy–Weinberg p-value $< 10^{-4}$, sex difference in allele frequency ≥ 0.2 and sex difference in heterozygosity > 0.3 for autosome chromosomes. After SNV QC, a total of 1,438,399 SNVs were available for analyses.

Imputation was done centrally at the University of Washington in combined samples. The study samples were phased with SHAPEIT2 (Delaneau et al., 2013) and imputed with IMPUTE2 (Howie et al., 2009) to the 1000 Genomes Project Phase 3 data release. Reference panel variants were restricted to a minor allele count (MAC) ≥ 2 across all 1000. Kinship coefficients were estimated using PC-Relate (Conomos et al., 2016). Principal components (PCs) were estimated in unrelated individuals within the global study population using SNVRelate package (Zheng et al., 2012). The first 10 PCs explained most of the genetic variation in the PAGE study population.

CKD Phenotypes

For studies with available serum creatinine (HCHS/SOL, WHI), we calculated eGFR using the CKD-EPI equation and baseline cohort data (Inker et al., 2012). Mild to moderate CKD (referred as CKD) was defined by an eGFR between 15 and 60 ml/min/1.73 m^2 (HCHS/SOL, WHI) or by an ICD-9 or 10 code in medical claims (585.1-585.5, 585.9, N18.1-N18.5, N18.9) (MEC, BioMe) (Nadkarni et al., 2014). Advanced CKD (referred as ESKD) was defined by an eGFR < 15 ml/min/1.73 m^2 (HCHS/SOL), an ICD-9 or 10 code of 585.6 or N18.6 related to ESKD (MEC) or ESKD obtained through linkage to the United States Renal Data System (BioMe). CKD and ESKD cases were mutually exclusive. Controls were individuals with an eGFR > 60 ml/min/1.73 m^2 or without ICD codes related to CKD or ESKD. In sensitivity analyses, we used two additional definitions for mild to moderate CKD: one based on ICD codes (MEC and BioMe data) and one based on eGFR from cohort studies (HCHS/SOL, WHI).

Statistical Analyses

We performed genome-wide association analyses of the combined data using the software SUGEN, which implements a generalized estimating equation approach and empirically estimates within-family correlations without modeling the correlation structures of complex pedigrees (Lin et al., 2014). SUGEN adopts a modified version of the sandwich variance estimator, which replaces the empirical covariance matrix of the score vectors by the Fisher information matrix for unrelated subjects. We used logistic regression to analyze categorical phenotypes, and included age, sex, ten PCs, study, center (if available), and ethnicity as covariates. We filtered variants with an effective number <50 based on minor allele frequency of cases (MAF), number of participants (N) and imputation score from impute2 (info) using the following calculation [2*MAF*(1-MAF)*N*info] where N = total sample for a given phenotype. P-values were generated by score tests. Significant threshold for GWAS was $p < 5.0 \times 10^{-8}$ and suggestive threshold was a $p < 1.0 \times 10^{-7}$.

For SNVs with $p < 10^{-7}$, we used the clumping procedure INDEP in Easystrata to identify independent signals at each locus which included on 1 Mb genomic interval flanking the lead SNVs. We also examined if the published loci for eGFR were associated with CKD or ESKD. We prioritized eGFR variants identified in the multi-ethnic COGENT-Kidney consortium and also used variants available in the Genome catalog for CKD and ESKD. For the *APOL1* locus, we performed analysis conditioning on the most significant SNV in the region. In sensitivity analysis, we assessed the significance of SNVs identified in the GWAS for CKD using CKD definition based on ICD codes or eGFR thresholds.

Associations in the United Kingdom Biobank

We assessed the association of our identified variants in the United Kingdom Biobank for SNVs available in European ancestry listed in **Tables 2**, **3**. We extracted p-values from United Kingdom Biobank using GeneATLAS (2019) for ICD-10 diagnosis codes N18 (chronic renal failure, 4,905 cases, and 447,359 controls), N19 (unspecified renal failure including uremia, kidney failure, azotemia, 1,516 cases and 450,748 controls) and renal/kidney failure (759 cases and 451,505 controls) (Gene ATLAS). Publicly available replication samples for CKD/ESKD in non-European ancestry were not available.

RESULTS

Participant Characteristics

Overall, 41,041 individuals contributed data for CKD analyses (10.1% cases) and 31,694 individuals to ESKD analyses (3.4%

TABLE 1 | Descriptive characteristics of mild to moderate CKD and advance CKD stages.

Characteristics	CKD (mild to moderate)	Control	ESKD	Control
Total number	4,150	41,041	1,105	31,694
Mean age (SD), years	63.1 (9.5)	52.5 (13.1)	57.9 (11.4)	50.6 (13.8)
Female, %	61.2	68.2	45.9	59.4
Diabetes, %	53.3	24.7	56.1	24.1
Hypertension, %	92.5	49.6	98.7	48.0
African American, N (%)	1, 532 (37)	13, 761 (34)	584 (53)	8, 429 (27)
Hispanic/Latino, N (%)	1, 336 (32)	19, 317 (47)	282 (26)	15, 852 (50)
East Asian, N (%)	575 (14)	3, 540 (9	112 (10)	3, 536 (11)
Native Hawaiian, N (%)	608 (15)	2, 892 (7)	96 (9)	2, 892 (9)
American Indian, N (%)	42 (1)	595 (2)	1 (0)	49 (0)
Other ethnicity, N (%)	57 (1)	936 (2)	30 (3)	936 (3)

SD, standard deviation; N, number.

TABLE 2 | Main findings for association with mild to moderate CKD at $p < 10^{-7}$.

Chr:position (hg19)	SNV	Alleles	Allele frequency PAGE	Allele frequency 1000 genome project			Odds ratio	95% confidence interval	p	Nearby gene	Function or location
			Multi-ethnic	AFR	AMR	EUR					
1:226523031	rs76064236	G/T	0.005	NA	0.001	0.002	2.28	1.69, 3.16	9.5×10^{-7}	PARP1	intergenic
3:128388015	rs146639727	C/G	0.004	0.013	NA	NA	3.05	2.04, 4.56	5.1×10^{-8}	RPN1	intergenic
8:131963415	rs138873021	T/TAAG	0.629	0.260	0.752	0.818	1.17	1.11, 1.25	1.6×10^{-7}	ADCY8	intron
9:12141349	rs186208070	A/T	0.009	0.020	0.003	0.001	2.07	1.59, 2.70	6.7×10^{-8}	TYRP1	intergenic
10:15225054	rs10906850	C/T	0.225	0.109	0.242	0.325	1.18	1.11, 1.25	3.7×10^{-8}	NMT2	intergenic
11:125840145	rs183951714	T/C	0.011	0.000	0.000	0.000	1.66	1.36, 2.02	7.0×10^{-7}	CDON	intron
12:105784903	rs117329947	C/A	0.014	0.001	0.027	0.027	1.79	1.42, 2.25	9.3×10^{-7}	C12orf75	intergenic
16:63000290	rs11645800	G/A	0.184	0.404	0.085	0.109	1.20	1.12, 1.28	1.4×10^{-7}	CDH8	intergenic
16:86761390	rs147084429	T/C	0.019	0.050	0.001	NA	0.63	0.53, 0.76	4.5×10^{-7}	FOXL1	intergenic
17:48825516	rs144210385	A/G	0.006	0.001	0.009	0.021	2.85	1.94, 4.17	7.6×10^{-8}	LUC7L3	intron
21:41667378	rs2837554	G/A	0.701	NA	0.001	0.002	1.16	1.10, 1.24	3.1×10^{-7}	DSCAM	intron
21:43048485	rs187652497	C/T	0.007	0.017	0.001	NA	2.23	1.66, 3.00	1.1×10^{-7}	LINC00111	intergenic

Chr, chromosome; SE, standard error; SNV, single nucleotide variant; NA, not available; AFR, African; EUR, European; AMR, Ad Mixed American. Models adjusted for age, sex, study, race/ethnicity and PCs.

cases) with non-overlapping cases. Cases were older and had a lower proportion of women compared to controls, and more comorbidities (**Table 1**). The study-specific contribution for cases and controls is shown in **Supplementary Table 1**.

Multi-Ethnic GWAS Findings

The main findings from the GWAS of CKD and ESKD in combined multi-ethnic samples are shown in **Tables 2**, **3**. Manhattan plots are shown in **Figure 1** and quantile-quantile (QQ) are shown in **Supplementary Figure 1**. The genomic control lambdas were 1.025 for CKD and 1.026 for ESKD. These analyses identified two genome-wide associated loci: a chromosome 10 locus nearby *NMT2* associated with CKD (rs10906850, allele frequency 0.23, $p = 3.7 \times 10^{-8}$) (**Table 2 and Figure 2A**) and the chromosome 22 *APOL1* locus associated with ESKD (four common SNVs, including the two highly correlated *APOL1* G1 missense variants rs73885319 and rs60910145) (**Figure 2B** and **Supplementary Table 2**). *APOL1* G2 indel was not available in our data. Conditional analysis on the

most significant SNV supported an independent association at the *APOL1* locus.

Several loci with low frequency SNVs had suggestive evidence for association including SNVs located nearby *RPN1* ($p = 5.1 \times 10^{-8}$), *TYRP1* ($p = 6.7 \times 10^{-8}$), and *LUC7L3* ($p = 7.6 \times 10^{-8}$) associated with CKD, and *MIR4790* ($p = 8.6 \times 10^{-8}$) associated with ESKD. Except for *LUC7L3*, the most significant SNV at these loci was rare or not present in European ancestry. For example, the *TYRP1* variant allele frequency in 1000 Genomes Project is 0.02 in African ancestry and 0.001 in European ancestry, and *RPN1* and *MIR4790* SNVs are not available in 1000 Genomes Project European ancestry samples. Additional SNVs associated with ESKD at $p < 10^{-7}$ included a low frequency intronic variant in *FTO* (rs7189997, $p = 2.8 - 10^{-7}$) and a common variant nearby *IRX3* (rs8050506, $p = 9.94 \times 10^{-7}$). Both SNVs were also more common in African ancestry reference panels (**Table 3**).

Nine of the SNVs with significant or suggestive association with CKD or ESKD (listed in **Tables 2**, **3**) were available

TABLE 3 | Main association findings with ESKD at $p < 10^{-7}$.

Chr:position (hg19)	SNV	Alleles	Allele Freq PAGE	Allele Frequency 1000 Genome Project			Odds Ratio	95% Confidence Interval	p	Nearby gene	Function or location
			Multi-ethnic	AFR	AMR	EUR					
1:78634021	rs77138376	A/G	0.009	NA	NA	NA	2.63	1.83, 3,78	1.9×10^{-7}	GIPC2	intergenic
1: 146810798	rs12032578	C/T	0.088	0.160	0.033	0.007	1.44	1.25, 1.66	3.1×10^{-7}	LINC00624	intergenic
3: 5547449	rs112407915	A/T	0.020	0.047	0.003	NA	1.96	1.53, 2.51	8.6×10^{-8}	MIR4790	intergenic
4: 118841271	rs114425659	A/G	0.039	0.023	0.063	0.080	1.87	1.45, 2.39	9.1×10^{-7}	NDST3	intergenic
7:11617925	rs74614630	C/T	0.012	0.030	0.003	NA	2.25	1.63, 3.11	9.2×10^{-7}	THSD7A	intron
9: 79608946	rs115747230	C/T	0.021	0.061	0.004	0.001	1.88	1.47, 2.39	3.6×10^{-7}	FOXB2	intergenic
11:24231483	rs116510623	G/A	0.034	0.107	0.009	0.001	1.65	1.35,2.01	6.4×10^{-7}	LUZP2	intergenic
13:32883196	rs60236946	T/C	0.021	0.071	0.006	NA	1.86	1.45, 2.38	9.2×10^{-7}	ZAR1L	intron
14:56825071	rs115007604	A/G	0.008	0.023	0.006	NA	2.67	1.81, 3.95	8.7×10^{-7}	PELI2	intergenic
16: 53886224	rs7189997	G/A	0.009	0.027	0.003	NA	2.44	1.73, 3.42	2.8×10^{-7}	FTO	intron
16: 54413730	rs8050506	A/G	0.211	0.324	0.108	0.119	1.29	1.17, 1.43	7.5×10^{-7}	IRX3	5'-UTR
17:14189909	rs191540116	A/G	0.011	0.026	0.01	0.000	2.32	1.66, 3.25	8.6×10^{-7}	HS3ST3B1	intergenic
18: 54764414	rs12963285	C/T	0.166	0.269	0.170	0.038	0.74	0.66, 0.84	9.7×10^{-7}	LINC-ROR	5'-UTR
22:29461285	rs138572244	A/G	0.008	0.022	0.001	NA	2.90	1.94, 4.35	2.4×10^{-7}	C22orf31	intergenic
22: 36661906	rs73885319	G/A	0.077	0.260	0.009	NA	1.51	1.32, 1.72	1.9×10^{-9}	APOL1	missense

Chr, chromosome; SE, standard error; SNV, single nucleotide polymorphism; NA, not available; AFR, African; EUR, European; AMR, Ad Mixed American. Models adjusted for age, sex, study, race/ethnicity and PCs.

for replication in the United Kingdom Biobank white British samples. The SNV rs10906850 nearby *NMT2* was significantly associated with ICD code for renal/kidney failure ($p = 0.008$) and rs11645800 nearby *CDH8* was nominally associated with renal/kidney failure (**Table 4**). The common indel rs138873021 was not available in the United Kingdom Biobank.

Cross-Association of Significant SNVs for CKD and ESKD and Sensitivity Analyses

At the *NMT2* locus associated for CKD, the most significant SNV was not associated with ESKD ($p = 0.82$). At the *APOL1* locus associated with ESKD, rs73885319 (and other variants) were not associated with CKD ($p = 0.76$). To explore heterogeneity in the definition for CKD that could explain our findings, we examined the association of the genome-wide associated significant SNVs in samples stratified by CKD definition based on ICD code ($n = 4,698$ cases, $n = 18,764$ controls) or eGFR thresholds ($n = 3,179$ cases, $n = 18,550$ controls). The *NMT2* SNV rs10906850 was associated with CKD using both definitions ($p = 2.4 \times 10^{-6}$ for ICD codes and $p = 8.2 \times 10^{-3}$ for eGFR thresholds) and there was consistency in direction of effects. Conversely, *APOL1* rs73885319 was not associated with CKD using either definition ($p > 0.05$).

Association of Previously Reported eGFR SNVs From the COGENT-Kidney Consortium With CKD Stages in PAGE

Given PAGE studies included diverse (non-European) participants, we selected 93 eGFR SNVs identified in the *trans*-ethnic COGENT-Kidney Consortium to assess their

association with CKD and ESKD in PAGE. Seventeen loci were associated with CKD and six loci were associated with ESKD at nominal *p*-values ($p < 0.05$). These SNVs had concordant effect estimates between the COGENT-Kidney eGFR lowering allele that showed increased odds of CKD or ESKD (**Supplementary Table 3**). The *PDILT/UMOD* was the only locus that was associated with both CKD and ESKD.

DISCUSSION

The main finding of this study is the identification of a new locus for mild to moderate CKD nearby *NMT2* for a SNV common across ancestries. The study also shows for the first time genome-wide associations of the *APOL1* SNVs with ESKD across multi-ethnic populations. Several other leading low frequency variants showed suggestive association with CKD traits. Most of the low frequency associated SNVs were more common in reference datasets of African ancestry and rare or absent in individuals of European ancestry, particularly for findings related to ESKD. These findings were driven by our discovery samples, which is composed of a large number of non-European ancestry including African Americans (34%) and Hispanics/Latinos (46%). Only nine SNVs were available in the United Kingdom Biobank for a sample of European ancestry. We replicated the association at the *NMT2* locus, which also showed consistent direction of effects for alleles between PAGE and the United Kingdom Biobank samples. Of the remaining eight SNVs brought for replication, four had consistent direction of effect between CKD discovery and the United Kingdom Biobank ICD code N18 replication, and one showed consistent direction of effect between ESKD discovery and United Kingdom Biobank ICD code N19 and renal/kidney
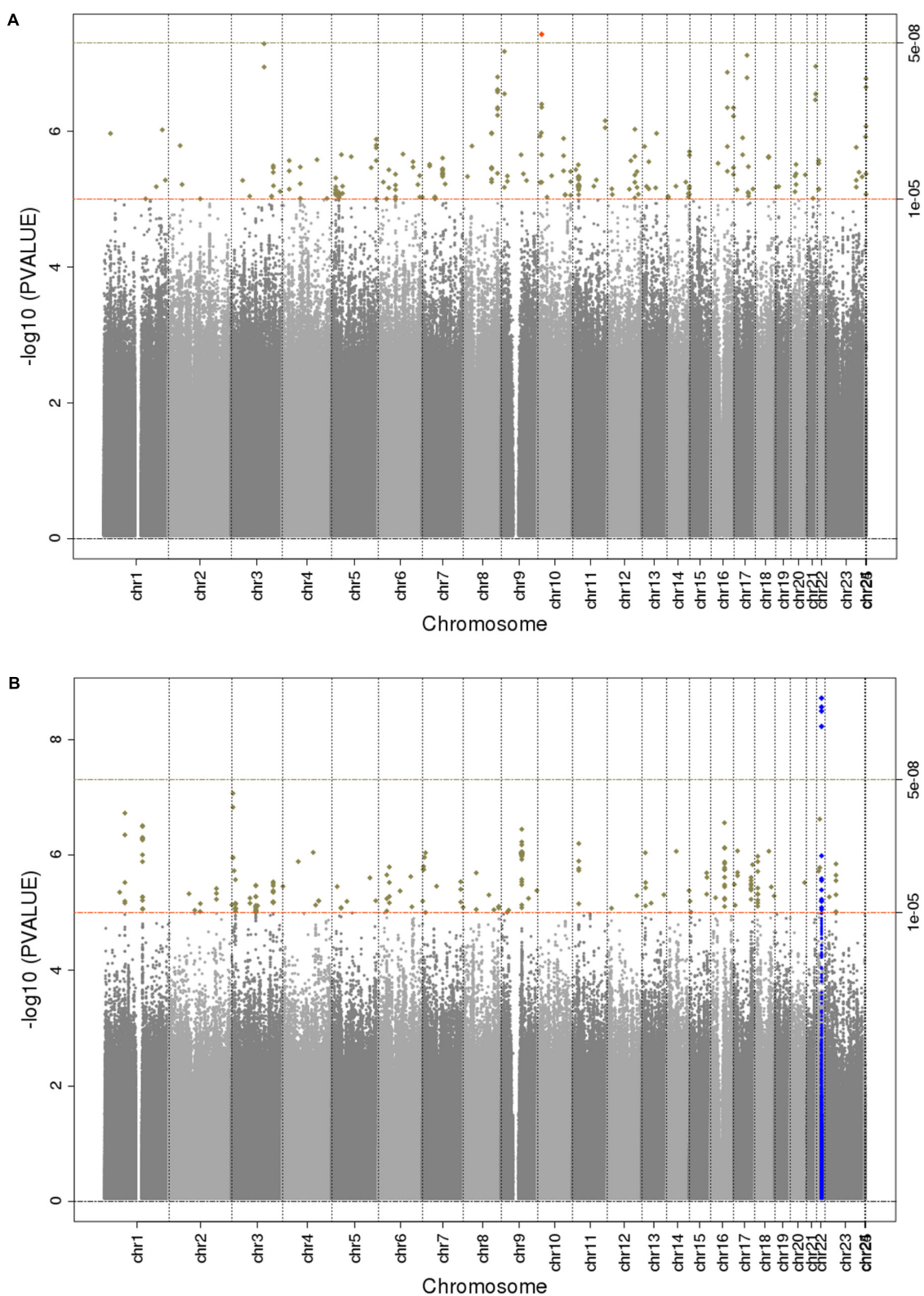
FIGURE 1 | Manhattan plots for *trans*-ethnic GWAS of CKD **(A)** and ESKD **(B)**. Significant novel (red) and known (blue) loci are highlighted.
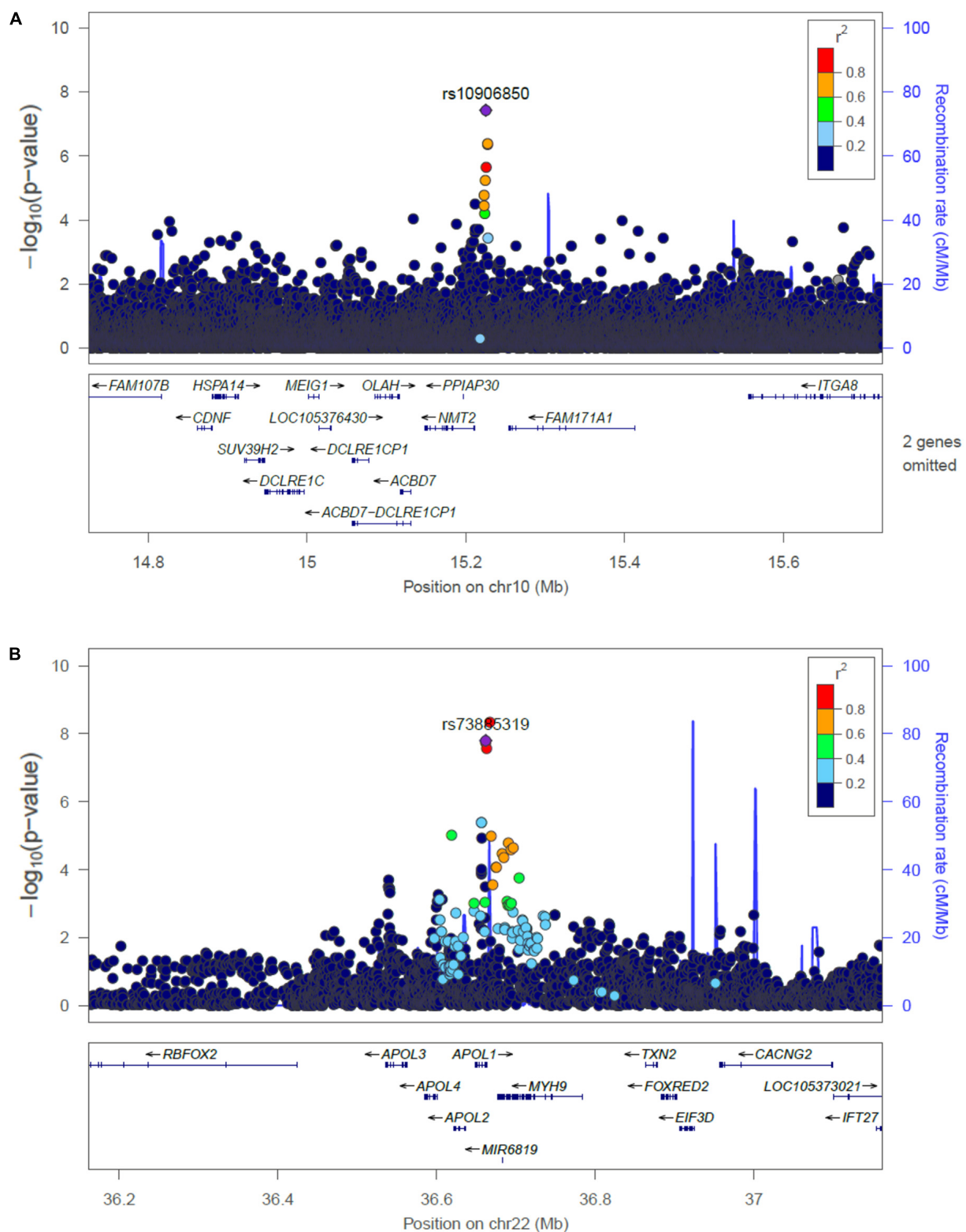
FIGURE 2 | Regional plots of the association at the *NMT2* **(A)** with CKD, and the *APOL1* locus associated with ESKD **(B)**. Linkage disequilibrium was estimated from the PAGE multi-ethnic data.

TABLE 4 | Association of SNVs identified for CKD or ESKD in white British in the United Kingdom Biobank for three CKD diagnosis.

SNV	Coded allele	Minor Allele frequency United Kingdom Biobank	N18 chronic renal failure*		N19 unspecified renal failure*		Renal/kidney failure*		PAGE discovery samples
			p	OR	p	OR	p	OR	OR
CKD SNVs									
rs10906850	C	0.37	0.16	1.03	0.83	1.01	0.0089	1.15	1.18
rs117329947	C	0.03	0.75	1.02	0.60	0.95	0.34	1.02	1.79
rs11645800	G	0.35	0.63	1.01	0.33	1.04	0.028	1.13	1.20
rs144210385	A	0.02	0.94	1.01	0.31	0.86	0.69	0.92	2.85
rs2837554	G	0.08	0.57	1.02	0.79	0.98	0.12	0.86	1.16
ESKD SNVs									
rs12032578	C	0.003	0.96	1.01	0.17	0.61	0.34	0.62	1.44
rs114425659	A	0.05	0.32	0.96	0.87	0.99	0.12	0.83	1.87
rs8050506	A	0.11	0.35	1.03	0.84	1.01	0.64	0.96	1.29
rs12963285	C	0.04	0.34	0.95	0.02	0.80	0.35	0.88	0.74

Indels rs61555057 and rs138873021 are not available in the United Kingdom Biobank. SNV, single nucleotide variant; OR, odds ratio for association of coded allele.
ICD-10 codes in the United Kingdom Biobank.

failure replication, although p-values were not significant. We were unable to replicate other variants due to their low frequency or unavailability in samples of European ancestry, and lack of comparable publicly available summary results for CKD traits in non-European populations.

Single nucleotide variant rs10906850 is an intergenic variant located nearby *NMT2*, a gene that encodes a protein involved in regulating the function and localization of signaling proteins. The SNV is an expression quantitative trait (eQTL) for *NMT2* in tibial artery ($p = 4.7 \times 10^{-9}$), adipose tissue ($p = 5.7 \times 10^{-8}$) and skin ($p = 1.6 \times 10^{-7}$). This locus has not been previously associated with kidney traits. Additional locus identified at $p < 10^{-7}$ for CKD includes *ADCY8*, which has been previously described in the CRIC study for association with eGFR decline among non-diabetic African Americans, although not at genome-wide significant level (Parsa et al., 2017). Our SNV in the region (rs138873021) is in linkage disequilibrium with the SNV identified in the CRIC study (rs4492355, $p = 1.3 \times 10^{-7}$ in the CRIC study, D' = 0.90, $r^2 = 0.01$ in 1000 Genomes Project African ancestry), although rs4492355 is not associated with CKD in the study ($p = 0.53$). Two loci associated with ESKD have been previously associated with obesity traits (*FTO* and *IRX3*), but our SNVs are low frequency and more common in African ancestry. Our findings related to the association of eGFR SNVs identified in the COGENT-Kidney consortium with CKD stages supports heterogeneity in genetic effects across CKD stages. We found a larger number of COGENT-Kidney eGFR lowering SNVs associated with increased CKD than eGFR lowering SNVs associated with increased ESKD. Only one locus was associated with both CKD and ESKD in PAGE: rs77924615 at the *PDILT/UMOD* locus, which showed nominal associations with these traits.

Chronic kidney disease is a heterogeneous disease in its etiology and clinical manifestation, with varying rates of progression to advanced stages. There is still little understanding on the mechanisms related to these varying patterns of disease severity even within the same disease etiology, for example, diabetic nephropathy. An interesting finding of our study is that there may be differences in genetic susceptibility based on the severity of the disease manifestation. We found little overlap of the most significant loci associated with the CKD phenotype (that includes mild to moderate CKD stages) and ESKD (which reflects advanced CKD). For example, the known *APOL1* G1 genotype was not associated with CKD. The *NMT2* SNV was significantly associated with CKD but not with ESKD. Although our study lacks information on the *APOL1* G2 SNV, we did not identify additional associations at the chromosome 22 locus in conditional analyses. The *APOL1* risk genotypes, including G1, were identified in admixture mapping for ESKD attributed to hypertension, FSGS or HIV (Genovese et al., 2010), and associations have been replicated in population studies although not at the genome-wide significant level (Foster et al., 2013; Kramer et al., 2017). Our study provides evidence for *APOL1* association with ESKD among diverse populations and for ESKD not selected for a specific disease etiology.

There are several possible explanations for the different genetic findings by CKD stages. Advanced CKD (ESKD) may have a stronger genetic component related to CKD progression, whereas mild to moderate CKD may capture more genetic factors related to CKD initiation. Mild to moderate CKD likely has less genetic influences due to inclusion of older individuals (aging process) and due to environmental factors. Alternatively, there is greater heterogeneity in our definition of mild to moderate CKD, opening the possibility of misclassified cases among individuals with an eGFR around the threshold of 60 ml/min/1.73 m^2. However, our sensitivity analyses in CKD subgroups defined by ICD codes or eGFR thresholds did not show differences in the association for our significant *NMT2* locus. Overall, these findings provide important information for the study design

of genetic studies for CKD, which should consider phenotype heterogeneity and severity of disease particularly when CKD is defined using ICD billing codes and when the definition includes a mixed case of CKD identified through biomarkers or clinical disease.

In conclusion, our multi-ethnic study identified a novel locus for mild to moderate CKD and replicated a known locus for ESKD. Our results highlight the need for more studies in diverse populations to identify genetic risk factors in populations at higher risk for CKD. It also underscores the current limitations of genetic research in these populations, including the lack of suitable replication samples for non-European ancestry variants.

ETHICS STATEMENT

All human research was approved by the relevant institutional review boards and conducted according to the Declaration of Helsinki. All participants provided written informed consent.

AUTHOR CONTRIBUTIONS

NF, GN, and CW conceived and designed the experiments. NF and CW coordinated the project. JA, RT, MG, HMH, SB, TM, and D-YL performed the quality control of genotypes and phenotype data, or support for statistical methods. BL, NF, and CW performed the statistical analyses. BL, NF, LW, and RT drafted the manuscript. All authors critically revised the manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene. 2019.00494/full#supplementary-material

REFERENCES

Anderson, G. L., Manson, J., Wallace, R., Lund, B., Hall, D., and Davis, S. (2003). Implementation of the women's health initiative study design. *Ann. Epidemiol.* 13, S5–S17.

Bien, S. A., Wojcik, G. L., Zubair, N., Gignoux, C. R., Martin, A. R., and Kocarnik, J. M. (2016). Study, strategies for enriching variant coverage in candidate disease loci on a multiethnic genotyping array. *PLoS One* 11:e0167758. doi: 10.1371/journal.pone.0167758

Collins, A. J., Foley, R. N., Gilbertson, D. T., and Chen, S. C. (2011). United States renal data system public health surveillance of chronic kidney disease and end-stage renal disease. *Kidney Int. Suppl.* 5, 2–7. doi: 10.1038/kisup.2015.2

Conomos, M. P., Reiner, A. P., Weir, B. S., and Thornton, T. A. (2016). Model-free estimation of recent genetic relatedness. *Am. J. Hum. Genet.* 98, 127–148. doi: 10.1016/j.ajhg.2015.11.022

Delaneau, O., Zagury, J. F., and Marchini, J. (2013). Improved whole-chromosome phasing for disease and population genetic studies. *Nat. Methods* 10, 5–6. doi: 10.1038/nmeth.2307

Foster, M. C., Coresh, J., Fornage, M., Astor, B. C., Grams, M., and Franceschini, N. (2013). APOL1 variants associate with increased risk of CKD among African Americans. *J. Am. Nephrol.* 24, 1484–1491. doi: 10.1681/ASN.2013010113

GeneATLAS Gene ATLAS is a Large Database of Associations Between Hundreds of Traits and Millions of Variants using the UK Biobank cohort. Available at: http://geneatlas.roslin.ed.ac.uk/ (accessed March 14, 2019).

Genovese, G., Friedman, D. J., Ross, M. D., Lecordier, L., Uzureau, P., and Freedman, B. I. (2010). Association of trypanolytic ApoL1 variants with kidney disease in African Americans. *Science* 329, 841–845. doi: 10.1126/science.1193032 doi: 10.1126/science.1193032

Global Burden of Disease 2016 Causes of Death Collaborators (2017). Global, regional, and national age-sex specific mortality for 264 causes of death, 1980-2016: a systematic analysis for the global burden of disease study 2016. *Lancet* 390, 1151–1210.

Go, A. S., Chertow, G. M., Fan, D., McCulloch, C. E., and Hsu, C. Y. (2004). Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. *N. Engl. J. Med.* 351, 1296–1305. doi: 10.1056/nejmoa041031

GWAS Catalog *The NHGRI-EBI Catalog of Published Genome-Wide Association studies.* Available at: https://www.ebi.ac.uk/gwas/ (accessed March 14, 2019).

Howie, B. N., Donnelly, P., and Marchini, J. (2009). A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* 5:e1000529. doi: 10.1371/journal.pgen.1000529 doi: 10.1371/journal.pgen.1000529

Inker, L. A., Schmid, C. H., Tighiouart, H., Eckfeldt, J. H., Feldman, H. I., and Greene, T. (2012). Estimating glomerular filtration rate from serum creatinine and cystatin C. *N. Engl. J. Med.* 367, 20–29. doi: 10.1056/NEJMoa1114248

Iyengar, S. K., Sedor, J. R., Freedman, B. I., Kao, W. H., Kretzler, M., and Keller, B. J. (2015). Genome-wide association and trans-ethnic meta-analysis for advanced diabetic kidney disease: family investigation of nephropathy and diabetes (FIND). *PLoS Genet.* 11:e1005352. doi: 10.1371/journal.pgen.1005352

Kolonel, L. N., Henderson, B. E., Hankin, J. H., Nomura, A. M., Wilkens, L. R., and Pike, M. C. (2000). A multiethnic cohort in Hawaii and Los Angeles: baseline characteristics. *Am. J. Epidemiol.* 151, 346–357. doi: 10.1093/oxfordjournals.aje.a010213

Kramer, H. J., Stilp, A. M., Laurie, C. C., Reiner, A. P., Lash, J., and Daviglus, M. L. (2017). African ancestry-specific alleles and kidney disease risk in hispanics/latinos. *J. Am. Soc. Nephrol.* 28, 915–922. doi: 10.1681/ASN.2016030357

Levin, A., and Stevens, P. E. (2014). Summary of KDIGO 2012 CKD guideline: behind the scenes, need for guidance, and a framework for moving forward. *Kidney Int.* 85, 49–61. doi: 10.1038/ki.2013.444

Lin, D. Y., Tao, R., Kalsbeek, W. D., Zeng, D., Gonzalez, F. II, and Fernandez-Rhodes, L. (2014). Genetic association analysis under complex survey sampling: the hispanic community health study/study of latinos. *Am. J. Hum. Genet.* 95, 675–688. doi: 10.1016/j.ajhg.2014.11.005

Morris, A. P., Le, T. H., Wu, H., Akbarov, A., van der Most, P. J., and Hemani, G. (2019). Trans-ethnic kidney function association study reveals putative causal genes and effects on kidney-specific disease aetiologies. *Nat. Commun.* 10:29. doi: 10.1038/s41467-018-07867-7

Nadkarni, G. N., Gottesman, O., Linneman, J. G., Chase, H., Berg, R. L., and Farouk, S. (2014). Development and validation of an electronic phenotyping algorithm for chronic kidney disease. *AMIA Annu. Symp. Proc.* 2014, 907–916.

Parsa, A., Kanetsky, P. A., Xiao, R., Gupta, J., Mitra, N., and Limou, S. (2017). Genome-wide association of CKD progression: the chronic renal insufficiency cohort study. *J. Am. Soc. Nephrol.* 28, 923–934. doi: 10.1681/ASN.2015101152

Saran, R., Robinson, B., Abbott, K. C., Agodoa, L. Y. C., Bhave, N., and Bragg-Gresham, J. (2018). US renal data system 2017 annual data report: epidemiology of kidney disease in the United States. *Am. J. Kidney Dis.* 71:A7.

Sorlie, P. D., Aviles-Santa, L. M., Wassertheil-Smoller, S., Kaplan, R. C., Daviglus, M. L., and Giachello, A. L. (2010). Design and implementation of the hispanic community health study/study of latinos. *Ann. Epidemiol.* 20, 629–641. doi: 10.1016/j.annepidem.2010.03.015

van Zuydam, N. R., Ahlqvist, E., Sandholm, N., Deshmukh, H., and Rayner, N. W. (2018). A genome-wide association study of diabetic kidney disease in subjects with type 2 diabetes. *Diabetes Metab. Res. Rev.* 67, 1414–1427.

Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., and Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* 28, 3326–3328. doi: 10.1093/bioinformatics/bts606

frontiers
in Genetics

Check for
updates

Genetic Susceptibility to Chronic Kidney Disease – Some More Pieces for the Heritability Puzzle

Marisa Cañadas-Garre[1]*, Kerry Anderson[1†], Ruaidhri Cappa[1†], Ryan Skelly[1†], Laura Jane Smyth[1†], Amy Jayne McKnight[1†] and Alexander Peter Maxwell[1,2]

[1] Epidemiology and Public Health Research Group, Centre for Public Health, Queen's University of Belfast, Belfast, United Kingdom, [2] Regional Nephrology Unit, Belfast City Hospital, Belfast, United Kingdom

Chronic kidney disease (CKD) is a major global health problem with an increasing prevalence partly driven by aging population structure. Both genomic and environmental factors contribute to this complex heterogeneous disease. CKD heritability is estimated to be high (30–75%). Genome-wide association studies (GWAS) and GWAS meta-analyses have identified several genetic loci associated with CKD, including variants in *UMOD*, *SHROOM3*, solute carriers, and E3 ubiquitin ligases. However, these genetic markers do not account for all the susceptibility to CKD, and the causal pathways remain incompletely understood; other factors must be contributing to the missing heritability. Less investigated biological factors such as telomere length; mitochondrial proteins, encoded by nuclear genes or specific mitochondrial DNA (mtDNA) encoded genes; structural variants, such as copy number variants (CNVs), insertions, deletions, inversions and translocations are poorly covered and may explain some of the missing heritability. The sex chromosomes, often excluded from GWAS studies, may also help explain gender imbalances in CKD. In this review, we outline recent findings on molecular biomarkers for CKD (telomeres, CNVs, mtDNA variants, sex chromosomes) that typically have received less attention than gene polymorphisms. Shorter telomere length has been associated with renal dysfunction and CKD progression, however, most publications report small numbers of subjects with conflicting findings. CNVs have been linked to congenital anomalies of the kidney and urinary tract, posterior urethral valves, nephronophthisis and immunoglobulin A nephropathy. Information on mtDNA biomarkers for CKD comes primarily from case reports, therefore the data are scarce and diverse. The most consistent finding is the A3243G mutation in the *MT-TL1* gene, mainly associated with focal segmental glomerulosclerosis. Only one GWAS has found associations between X-chromosome and renal function (rs12845465 and rs5987107). No loci in the Y-chromosome have reached genome-wide significance. In conclusion, despite the efforts to find the genetic basis of CKD, it remains challenging to explain all of the heritability with currently available methods and datasets. Although additional

biomarkers have been investigated in less common suspects such as telomeres, CNVs, mtDNA and sex chromosomes, hidden heritability in CKD remains elusive, and more comprehensive approaches, particularly through the integration of multiple –"omics" data, are needed.

INTRODUCTION

Chronic kidney disease is a major global health problem with an increasing prevalence (Levey et al., 2007; Bash et al., 2009; Centers for Disease Control and Prevention, 2015). By 2040, it is estimated that CKD will have become the fifth leading cause of death (Foreman et al., 2018). This increasing CKD burden is driven in part by aging population structure (CKD is ∼8x more common in adults > 70 years old compared to persons < 40 years of age) (Bash et al., 2009; Centers for Disease Control and Prevention, 2015). Diabetes and hypertension are common risk factors for kidney damage (Kazancioğlu, 2013) and are therefore major contributors to the increased CKD prevalence (Bash et al., 2009).

There is a marked gender imbalance in CKD with a higher incidence (11.0 vs. 9.6 per 1,000 person-years) and higher prevalence (16.0% vs. 12.4%) in women (Bash et al., 2009; Centers for Disease Control and Prevention, 2015). Nevertheless, women have a lower risk of CKD progression and men are more likely to develop ESRD (Ricardo et al., 2018).

Chronic kidney disease is a complex heterogeneous disease, with contributions from both genomic and environmental factors. CKD heritability has been estimated to be high (30–75%) (Satko and Freedman, 2005; O'Seaghdha and Fox, 2011; Regele et al., 2015). CKD can be identified by well-established clinical biomarkers such as SCr levels, eGFR, albuminuria, or UACR (Cañadas-Garre et al., 2018a,b). Unfortunately, these clinical biomarkers are limited in their utility to predict individual risk of CKD or likelihood for later progression to ESRD. Major efforts have been made to understand the heritability in CKD but the causal pathways remain incompletely understood. Four major approaches have been proposed to uncover the missing heritability; exploration of rare variants, increased samples sizes, study of molecular factors not involving variants in the DNA sequence and consideration of whether family studies overestimated heritability risk (Bourrat et al., 2017). In CKD, meta-analyses of GWAS have provided a useful and

relatively inexpensive strategy to increase the statistical power by combining data summaries from different individual GWAS, helping to attenuate the issue of small sample size and identifying many genetic loci associated with CKD and/or kidney function traits (Köttgen et al., 2009, 2010; Chambers et al., 2010; Böger et al., 2011; Parsa et al., 2013; Pattaro et al., 2016; Gorski et al., 2017). Rare variants in *UMOD*, *SHROOM3*, solute carriers, and E3 ubiquitin ligases have also been associated with CKD, eGFR or SCr (Köttgen et al., 2012; Sveinbjornsson et al., 2014; Prokop et al., 2018). However, these genetic markers do not account for all the susceptibility to CKD, therefore other factors must be contributing to the missing heritability. Part of the missing heritability may correspond to genetic interactions (epistasis), rather than to missing variants (Zuk et al., 2012). Telomere length is a biological factor that has been associated with CKD prevalence and/or CKD progression in a small number of studies (Ameh et al., 2017). Structural variants, such as CNVs, insertions, deletions, inversions and translocations are, in general, poorly covered in commercial arrays and may explain part of the missing heritability (Manolio et al., 2009). Mitochondrial proteins, encoded by nuclear genes, and specific mtDNA encoded genes have also been associated with CKD (Skelly et al., 2019). The sex chromosomes, often excluded from GWAS studies, may help explain gender imbalances in CKD.

In this review, we outline some recent findings on molecular biomarkers for CKD (telomeres, CNVs, mtDNA variants, X and Y chromosomes) that typically have received less attention than single nucleotide polymorphisms (SNPs) present on, or imputed from, GWAS arrays. These less commonly studied biomarkers may be part of the "missing heritability" for CKD.

Telomeres and CKD

Telomeres are specialized nucleoprotein complexes that help protect the ends of linear chromosomes (Sfeir, 2012). There are inter-individual and intra-individual differences in the length of telomeres. Shorter telomere length has been associated with multi-system diseases, early life stressors, increasing chronological age and all-cause mortality (Dlouha et al., 2014; De Meyer et al., 2018; Desai et al., 2018; Mangaonkar and Patnaik, 2018; Wang et al., 2018; Willis et al., 2018) (**Figure 1**). The majority of studies have analyzed relative telomere length in peripheral blood leukocytes, but telomere length differs between tissues within a single individual, with greater heterogeneity in telomere length evident in older people (Butler et al., 1998; Dlouha et al., 2014). Telomere length has a reported heritability of 28–82%, however, not all genetic factors (Broer et al., 2013; Codd et al., 2013) or environmental influences on telomere

FIGURE 1 | Telomeres and kidney disease. Copyright disclosure: mouse: https://commons.wikimedia.org/wiki/File:Vector_diagram_of_laboratory_mouse_%28black _and_white%29.svg; https://creativecommons.org/licenses/by-sa/4.0/deed.en; attribution, "By Gwilz [CC BY-SA 4.0 (https://creativecommons.org/licenses/by-sa/ 4.0)], from Wikimedia Commons." Kidney: https://commons.wikimedia.org/wiki/File:Kidney_Cross_Section.png; By Artwork by Holly Fischer [CC BY 3.0 (https://creativecommons.org/licenses/by/3.0)], via Wikimedia Commons.

length are known (Cubiles et al., 2018; Dugdale and Richardson, 2018; Gao et al., 2018; Lu et al., 2018). Meta-analysis of telomere length may help confirm discovery associations across multiple collections, however, this is challenging with different wet-lab techniques (such as time at sample collection, storage and processing of biological material, absolute compared to relative telomere length evaluation, platform employed) and *in silico* analyses (such as normalization, controls, covariates, association, and correction tools) having significant effects on the reported measurements. There is also limited traditional epidemiological evidence exploring the mechanistic basis or causality of reported associations.

Nonetheless, there is evidence that telomere length is associated with disease states, particularly age-related diseases, beyond the most commonly studied cancers (Rizvi et al., 2014; Jafri et al., 2016). Conflicting reports have been published for the association of telomere length with renal disease, however, most publications, albeit in relatively small sample sizes with modest significance values, report that shorter telomere length is associated with renal dysfunction. Shorter telomeres have been reported as associated with progression of CKD (defined as a doubling of baseline SCr and/or ESRD), in the MMKD ($n = 59$ patients had confirmed CKD progression) and CRISIS ($n = 105$ patients had confirmed CKD progression) studies, with the effect size strengthened by smoking and the presence of diabetes (Raschenberger et al., 2015). Telomere shortening has been associated with IgAN in 177 patients, but not in 30 patients with DKD or 30 patients with FSGS compared to 83 controls (Lu et al., 2014). A study examining DNA from peripheral blood

and urine in 15 patients with IgAN showed shorter telomere length correlated with declining renal function (Szeto et al., 2005). Multiple studies have been performed for DKD, with the majority linking shorter telomere length to the development and progression of kidney disease in people with both type 1 (Astrup et al., 2010, 273 patients; Fyhrquist et al., 2010, 176 patients, 21 progressed) and type 2 diabetes (Tentolouris et al., 2007, 168 patients; Verzola et al., 2008, 17 patients; Testa et al., 2011, 501 patients; Gurung et al., 2018, 691 patients). Shorter telomere length is associated with diabetic complications (Testa et al., 2011) and all-cause mortality (Astrup et al., 2010). The Heart and Soul Study is a longitudinal cohort of individuals with stable coronary heart disease; shorter telomere length at baseline and more rapid telomere shortening over 5 years were associated with reduced kidney function, but these changes were not significant when accounting for age (Bansal et al., 2012). It is noteworthy that the largest study published considered less than 1,000 individuals (Testa et al., 2011), which provides limited power to draw robust conclusions in this era of mega-consortia studying the genetics of CKD.

Premature telomere shortening is associated with duration of dialysis treatment in terms of months to years (Boxall et al., 2006). A cross-sectional study of 175 hemodialysis patients reported shorter telomere length in men with CKD, despite women having an older average age in this cohort; association of shorter telomeres was also observed with increasing age and male sex (Carrero et al., 2008). Shorter telomeres were associated with CKD in 203 Japanese hemodialysis patients compared to 203 age and sex-matched controls without CKD, with shorter telomeres

also associated with new onset cardiovascular events (Hirashio et al., 2014). A less reactive immune system is associated with healthy aging in the general population and ESRD enhances premature immunological aging with shorter telomeres observed in 137 patients with ESRD compared to 144 individuals without kidney disease (Betjes et al., 2011).

Histologically normal and abnormal human kidney tissue samples from 24 individuals highlighted age-related shorter telomere length with telomeres typically shorter in the cortex than in the medulla (Melk et al., 2000). Premature senescence is an important feature of renal fibrosis that accelerates when cells are exposed to stressful environments such as more ROS and higher glucose (Verzola et al., 2008; Carracedo et al., 2013; Cao et al., 2018). Increasing age and sex related telomere shortening is observed in kidneys, with shorter telomeres observed in male rats (Cherif et al., 2003). Multiple animal models of kidney disease show telomere shortening associated with renal dysfunction, however, a careful experimental design is required for accurate telomere measurement (Hastings et al., 2004). Exploring renal ischemia/reperfusion injury in wild-type and telomerase deficient mice also suggests that shorter telomeres impair recovery from acute kidney injury (Westhoff et al., 2010; Song et al., 2011; Cheng et al., 2015). Severe renal failure induces telomere shortening (Wong et al., 2009) with rapid telomere loss observed during kidney transplantation in a rat model of chronic rejection (Joosten et al., 2003). Tucker and colleagues demonstrated that high-intensity interval training was beneficial protecting against telomere erosion in a rat model of CKD (Tucker et al., 2015).

Large-scale studies using carefully collected biological samples with harmonized phenotypes and analysis protocols will help determine the true association of telomere length for CKD. Potential therapies exist to minimize premature telomere shortening (Townsley et al., 2016; Rodrigues et al., 2017), but further work is needed to define the mechanistic links between telomere length and kidney function.

Copy Number Variation and Larger Chromosomal Re-arrangements Association With CKD

Copy number variants are genetic structural variants which involve DNA regions being deleted or duplicated. This can occur throughout the genome affecting stretches of DNA ranging from kilo- to mega-base pairs in length and can result in abnormal gene amplification (Thapar and Cooper, 2013; Sampson, 2016). CNVs can be both inherited and arise de novo, and are increasingly being recognized as a significant source of genetic variation relating to both population diversity and disease, including renal diseases (Sampson, 2016), neuropsychiatric diseases (Lew et al., 2018), and cancer (Liang et al., 2016).

There is often uncertainty about the genetic basis of CKD in pediatric patients, but recent studies have indicated that chromosomal microarrays have the potential to partly address this. Verbitsky et al. (2015) assessed 419 children enrolled in the CKD in children (CKiD) study alongside 21,575 children and adults who had undergone microarray genotyping for non-CKD studies. CNV disorders were identified in 31 children

with CKD and 10 known pathogenic genomic disorders were detected including HNF1B deletion at 17q12. A further 12 pathogenic genomic imbalances were identified using this technique, distributed evenly among patients diagnosed with congenital and non-congenital forms of CKD. Overall, large gene-altering CNVs were more common in the CKiD population compared with the controls (38 vs. 23%), but the specific genetic alterations identified in several of the individuals would require personalized recommendations in future healthcare.

Copy number variants have been linked to CAKUT (Sanna-Cherchi et al., 2012; Caruana et al., 2014; Bekheirnia et al., 2017; Siomou et al., 2017). In a study by Caruana et al. (2014), DNA from 178 Australian children who presented with any abnormality associated with CAKUT was screened using SNP arrays. In total, CNVs were identified in 18 children, of which 11 children presented with genomic disorders of unknown significance. Of these 11 participants, four were reported as having duplications of 1q23.1, 4p16.1, 7q33, and 8q13.2q13.3 regions, containing genes NEPH1, SLC2A9, AKR1B1, and EYA1, respectively. Each of these genes have previously been associated with renal abnormalities.

In an investigation undertaken by Siomou et al. (2017), seven children with CAKUT were assessed from three unrelated families using array comparative genomics hybridization. Of these participants, one reportedly had ureterovesical junction obstruction and a 1.4 Mb deletion at 17q12, containing two genes, HNF1B, which has been previously associated with CAKUT, and ACACA (Thomas et al., 2011; Caruana et al., 2014).

A recent study published by Bekheirnia et al. (2017), suggested whole exome sequencing (WES) as a viable method to detect CNVs in individuals with CAKUT. These investigators performed WES in 112 individuals from 62 families, to identify SNVs and CNVs in 35 genes previously related to CAKUT. They identified a de novo triplication in one family at 22q11, and overall, 6.5% of the individuals assessed in this investigation were shown to have pathogenic CNVs.

Posterior urethral valves are one of the most common causes of CKD in children. Faure et al. (2016) assessed the phenotypic effects of and relationship between renal outcomes and CNVs in 45 boys with PUV. In total, 13 CNVs were identified in 12 boys, two of which, at positions 3p25.1p25.2 and 17p12, were pathogenic in nature. Additionally, those CNVs identified which were > 100 kb in size, were significantly associated with earlier onset of renal failure in children with PUV.

Nephronophthisis (NPH) is a Mendelian genetic disease, which often leads to ESRD by around 13 years of age. Snoek et al. (2018) sought to investigate the prevalence of NPH in adult-onset ESRD, through assessment of the CNVs in the NPHP1 gene (>90 kb) because a homozygous full gene deletion is a prominent cause of NPH. These investigators assessed 5,606 adult renal transplant recipients, 26 of whom showed evidence of the homozygous NPHP1 deletion, compared to none of the 3,311 controls. Despite this, only 12% of the patients with the homozygous NPHP1 gene deletion were clinically diagnosed with NPH.

Copy number variants have also been investigated in association with IgAN, which is the most common cause of

primary glomerulonephritis (Ai et al., 2016). The multi-allelic CNV in the defensin alpha 1 and alpha 3 gene locus (*DEFA1A3*) was assessed in two independent IgAN cohorts of Chinese Han individuals (Ai et al., 2016). This locus can present as tandem repeats of a 19kb DNA stretch, containing one copy of either *DEFA1* or *DEFA3*, and several bi-allelic polymorphisms. The protein products of *DEFA1A3*, human neutrophil peptides 1–3, are abundant neutrophil granule proteins and function in the regulation of both the complement system and pro-inflammatory cytokine production. Each of these have been previously linked with IgAN.

Evaluation of the presence of CNVs yields potentially useful clinical information, especially for pediatric individuals with CKD.

Copy number variants in the human genome are likely to contribute to healthy development, but have additionally been linked to several human diseases (Sampson, 2016; Liang et al., 2016; Lew et al., 2018). The molecular mechanisms that trigger the formation of CNVs are not fully understood, but recurrent CNVs with common breakpoints reportedly arise through unequal meiotic or non-allelic homologous recombination (Arlt et al., 2012). Recent evidence has suggested that *de novo* and non-recurrent CNVs may develop following either replicative errors, chromosome shattering or chromothripsis (Kloosterman et al., 2011; Arlt et al., 2012; Nazaryan-Petersen et al., 2018).

Replication stress occurring during DNA replication has been linked to the collapse of the DNA replication fork and creation of a single-ended double strand break (Arlt et al., 2012). It has been considered that this could result in a high frequency of *de novo* CNVs. Both the fork collapse and strand break could result in the activation of damage checkpoint and repair pathways to correctly reactivate replication, thus preventing the creation of structural variants. However, CNVs are understood to be created if this reactivation occurs in an incorrect location using a template switch, or when an incorrect repair occurs, which joins two distant DNA breaks and causes a large deletion (Arlt et al., 2012). Any present mutations which inhibit the ability of the cell to accurately respond to a collapsed fork, are thought to ultimately increase the formation of CNVs (Arlt et al., 2012).

Single Nucleotide Polymorphisms and Chronic Kidney Disease

In the last decade, GWAS have become essential for investigating the genetic contribution to CKD, with over 50 germline genetic loci identified as biomarkers of kidney disease risk or associated with SCr, cystatin-C and/or microalbuminuria (Cañadas-Garre et al., 2018a). The *UMOD* gene, coding for uromodulin, the most abundant urinary protein (Devuyst et al., 2017), is the gene with most of the consistently replicated genetic associations (Cañadas-Garre et al., 2018a). Several common *UMOD* variants (rs12917707, rs4293393, rs11864909, rs13329952) are associated with both CKD and eGFR (Köttgen et al., 2009, 2010; Gudbjartsson et al., 2010; Pattaro et al., 2012, 2016). More recently, the higher frequency in ESRD of another common *UMOD* variant (rs13333226), has been confirmed in 638 Chinese patients with ESRD and 366 controls (Chen et al., 2016). Several

common variants in the myosin heavy chain type II isoform A (*MYH9*) gene have been associated with non-diabetic ESRD in African Americans (Kao et al., 2008; Kopp et al., 2008; Chambers et al., 2010). Common variants in *APOL1* are also associated with non-diabetic ESRD (Genovese et al., 2010; Tzur et al., 2010; Foster et al., 2013). Common variants in *ELMO1* gene have been associated with DKD and its progression to ESRD in several populations, although in this case with less consistency (Shimazaki et al., 2005; Leak et al., 2009; Pezzolesi et al., 2009a,b; Narres et al., 2016). A more recent meta-analysis of GWAS, including data from 133,413 individuals and subsequently replicated in 42,166 individuals, identified 24 new loci associated with eGFR (*BCAS1, AP5B1, A1CF, PTPRO, UNCX, NFKB1, TP53INP2, KCNQ1, CACNA1S, WNT7A, TSPAN9, IGFBP5, KBTBD2, RNF32, SYPL2, SDCCAG8, ETV5, DPEP1, LRP2, SIPA1L3, INHBC, ZNF204, SKIL,* and *NFATC1*) (Pattaro et al., 2016). The *trans*-ethnic meta-analysis showed that 12 loci had fully consistent effect direction on eGFR across European, Asian and African individuals (*SDCCAG8, LRP2, IGFBP5, SKIL, UNCX, KBTBD2, A1CF, KCNQ1, AP5B1, PTPRO, TP53INP2,* and *BCAS1*). Regarding other measures of kidney function, a variant rs1801239 in the *CUBN* gene was proposed as a predictor of UACR and microalbuminuria in a meta-analysis of 63,153 individuals of European ancestry (Böger et al., 2011), and another variant in the same gene, rs10795433, has been associated with UACR in 5,825 individuals of European ancestry with diabetes compared to 46,061 without diabetes (Teumer et al., 2016). A recent discovery GWAS of UACR in 382,500 unrelated European participants of the UK Biobank, a population-based cohort, reported 33 common variants, 20 of them sharing a consistent direction of effect with the study by Teumer et al. (2016), including *CUBN, HOTTIP, LOC101927609, NR3C2, ARL15, SHROOM3, MAPKBP1, ICA1L, SNX17, LRMDA, SBF2, SPATA5L1, FUT1/IZUMO1* genes and additional variants in chromosomes 1, 2, 7, 14, and 15: rs10157710, rs12032996, rs1276720, rs17158386, rs2023844, rs2472297, rs4410790, rs6535594, rs702634, rs7654754, rs8035855, rs10207567, rs1047891, rs4665972, rs13394343, rs67339103, rs17368443, rs4288924, rs1145074, and rs838142 (Haas et al., 2018). This GWAS also identified 11 common novel associations in *CUBN*, PRKCI, *EFNA3-EFNA1, MIR548AR-LOC646736, COL4A4, SPHKAP-PID1, INC01262-FRG1, RIB1-LINC00861,* and *BAHCC1* genes. UACR had previously been associated with another common variant in *SHROOM3* (rs17319721) in a meta-analysis of 31,580 and 27,746 Caucasian patients, although it did not reach GWA significance ($p_{\text{discovery}} = 1.9 \times 10^{-6}$) (Böger et al., 2011).

Although GWAS have successfully identified SNPs associations for the different traits associated with CKD, most of them are common DNA variants of small effect size. The proportion of phenotypic variance of eGFR explained by the 24 novel loci and the 29 previously identified by Pattaro et al. was 3.22%, therefore of limited help in CKD prediction (Pattaro et al., 2016).

An alternative to the concept of SNPs as single biomarkers is the use of PGRS, which provide individual estimates of the risk of presenting a determinate trait calculated from the

combination of specific risks associated to SNPs. However, PGRS may provide only a partial solution in complex diseases. A recent analysis of 32 highly relevant traits related to five disease areas in 13,436 subjects of the Lifelines Cohort reported only 10.7% of the common-SNP heritability of these traits was explained by the different weighted PGRS, compiled from genome-wide significantly associated index SNPs based on previous GWAS (Nolte et al., 2017). The percentage of variance explained by the PGRS for SCr, composed of three SNPs of high imputation quality ($R^2 > 0.5$) was 0.2% for both weighted and unweighted PGRS (Nolte et al., 2017). Addition of one low-quality SNP increased the variance up to 0.21% (weighted PGRS). For the eGFR PGRS, composed of 33 SNPs (high-quality), the percentage of variance was 1.6% for unweighted PGRS and 1.8% for the weighted PGRS. Addition of 19 low-quality SNPs increased the variance up to 2.01% (weighted PGRS). There were no high-quality SNPs associated with UACR, so it was not possible to construct this PGRS. The inclusion of one low-quality SNP explained 0.12% of the variance with both weighted and unweighted PGRS (Nolte et al., 2017). The PGRS for urate, composed of 20 SNPs (high-quality), explained from 2.0 to 4.2%, depending if either an unweighted or weighted PGRS was considered. Addition of eight low-quality SNPs increased the variance up to 4.52% (weighted PGRS) (Nolte et al., 2017).

Next generation sequencing has an increasing role for both research and diagnosis of kidney disease. Recently, a NGS panel for a spectrum of genetic nephropathies, covering 301 genes, was designed and validated in a CLIA-approved laboratory (Larsen et al., 2016). The assay showed excellent performance characteristics and was able to provide a specific molecular pathogenesis-based diagnosis in 46% of biopsies studied. An NGS panel covering all coding and regulatory regions of *UMOD* identified 119 genetic variants in 23 ESRD patients (compared to 22 controls without renal disease). Ninety of those variants were SNPs, 60 of them with minor allele frequency greater than 5%. Linkage disequilibrium allowed 20 SNPs to capture 100% of the alleles with a mean R^2 of 0.97, providing a set of independent SNPs suitable for association analysis in larger cohorts (Bailie et al., 2017).

Whole-exome sequencing provided a diagnosis in 22 out of 92 adults with CKD of unknown cause, familial nephropathy or hypertension (22/92; 24%) (Lata et al., 2018). The confirmation of the clinical diagnosis by WES allowed the appropriate genetic counseling and screening for the family members of some affected patients and helped in clarifying or entirely reclassified the disease in other cases (Lata et al., 2018). WES also identified *PARN* haploinsufficiency as a new genetic cause of CKD in this study (Lata et al., 2018). The *PARN* gene encodes a poly(A)-specific ribonuclease which mediates the post-transcriptional maturation of the telomerase RNA component (*TERC)* and causes telomere disorders (Moon et al., 2015). Exome sequencing has recently identified 11 loci ($p < 1 \times 10^{-4}$) in eight genes (*PLEKHN1, NADK, RAD51AP2, RREB1, PEX6, GRM8, PRX, APOL1*) associated with T2DM-ESRD in 2476 cases and 2057 non-nephropathy control individuals of African American origin (Guan et al., 2018). However, exome data from 7974 self-identified healthy adults has recently demonstrated

an implausibly high rate of candidate pathogenic variants for kidney and genitourinary diseases (1.4%), much higher than the prevalence of genetic renal/genitourinary disorders, even after stringent filtering criteria (removal of indels and minor allele frequency cutoffs of < 0.01% and <0.1% for dominant and recessive disorders, respectively) (Rasouly et al., 2018). This overestimation of potential pathogenic variants may increase the burden of uncertain diagnoses and medical referrals rather than alleviate it, therefore minimizing the utility of exome sequencing in clinical practice (Rasouly et al., 2018).

Mitochondria and Their Association With Chronic Kidney Disease

Mitochondria are organelles which generate ATP through OXPHOS and thus represent the primary energy source for normal function of the cell and body (Cooper, 2000; Lodish et al., 2012; Chaban et al., 2014). The majority of mitochondrial proteins are encoded by nuclear genes (Timmis et al., 2004; Dolezal et al., 2006). However, mitochondria also have their own circular genome 16,569 base pairs long that contains 37 genes which encode 13 proteins of the electron transport chain essential for OXPHOS (Meiklejohn et al., 2013) along with two rRNAs and 22 tRNAs (Taanman, 1999; Cooper, 2000; Gray et al., 2008). Mitochondrial dysfunction in kidney tissue may severely impact renal health and has previously been implicated in CKD development (Rahman and Hall, 2013; Wallace, 2013; Zhan et al., 2013; Che et al., 2014; Douglas et al., 2014; Swan et al., 2015; Galvan et al., 2017).

If mitochondrial metabolism is adversely affected by genetic variants it can result in kidney disease, sometimes as part of a wider clinical disorder (Rahman and Hall, 2013). Somatic mtDNA mutations may be associated with aging, resulting in decline of mitochondrial function in older individuals (Wallace, 2013). Increased levels of mtDNA mutations have previously been associated with several disorders including various forms of kidney disease (**Figure 2**) (Wallace, 2013).

Mitochondrial dysfunction can occur via a number of pathways, for example persistent hyperglycemia (associated with diabetes) results in increased tubular oxygen consumption, and in turn leads to hypoxia of the kidney tissue (Hansell et al., 2013). Mitochondrial dysfunction can be associated with increased electron leakage from the respiratory chain during OXPHOS, which results in ROS being generated which can cause kidney injury (Granata et al., 2015) including direct damage to DNA (Marnett, 2000). Genetic variation in mtDNA (**Figure 2**) or nuclear genes (**Figure 3**) which influence mitochondrial function may impair respiratory chain complex activities leading to an increase in production of ROS resulting in a negative feedback loop, increasing mitochondrial dysfunction, OXPHOS defects and ROS generation along with a reduction in ATP production which leads to increased oxidative stress which may lead to uncontrolled autophagy, mitophagy, and further ROS production (Fernandez-Marcos and Auwerx, 2011; Kim et al., 2012; Zaza et al., 2013). Mitochondrial dysfunction, ROS generation and the resulting dysregulation of autophagic mechanisms may also lead to an upregulation of the intrinsic pathway of apoptosis which in turn leads to inflammation and fibrosis in the renal tubules and
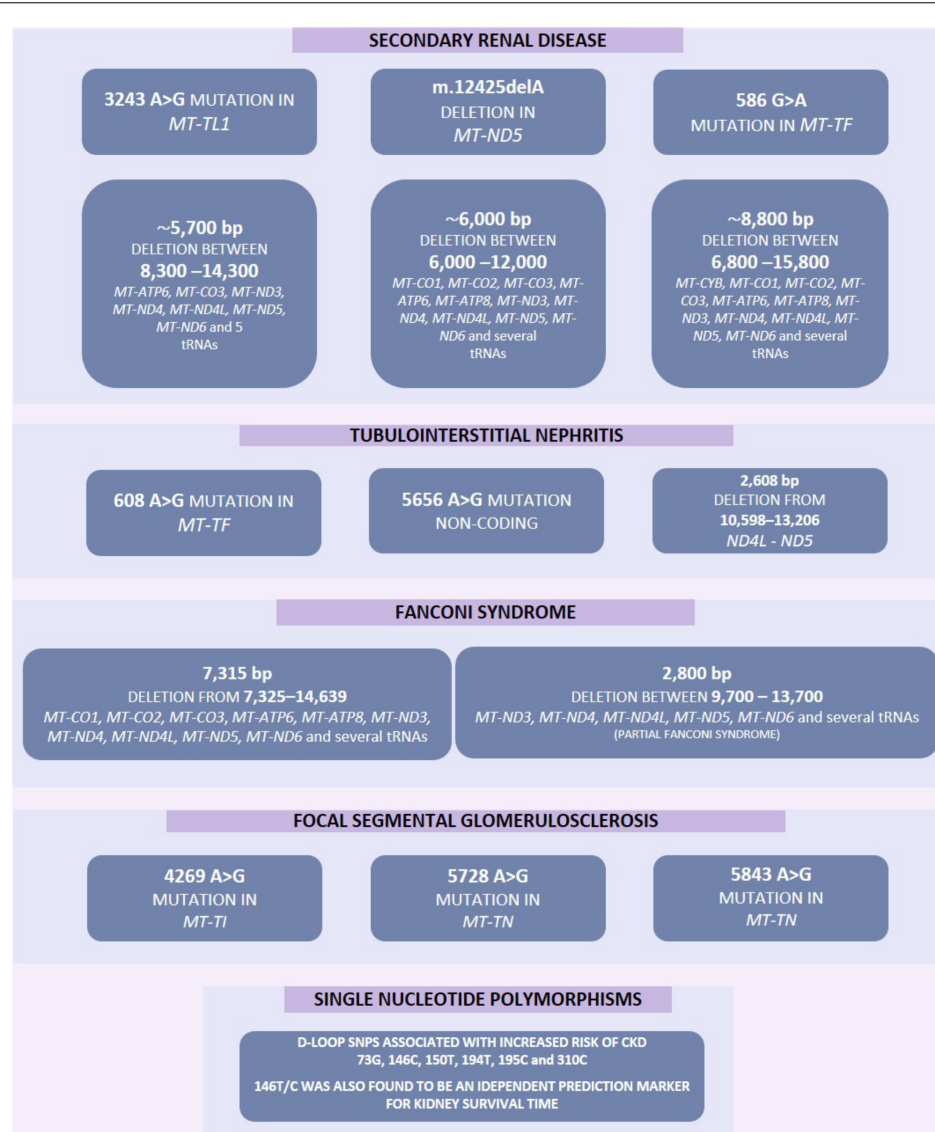
FIGURE 2 | Increased mutation rate in mtDNA have previously been associated with several diseases including various forms of kidney disease. These mutations include point mutations, deletions and single nucleotide polymorphisms. Some of these mutations may result in several pathological phenotypes and these have been highlighted by a solid line between such genes.

glomeruli (Tanaka et al., 2005; Song et al., 2010; Ye et al., 2010; Coughlan et al., 2016).

Despite the mitochondrial genome being widely ignored in relation to CKD, a number of studies have identified mitochondrial genomic loci associated with specific forms of renal disease (**Table 1**). SNPs within *MT-HV2*, *MT-CO1*, and *MT-CO2c* have been associated with IgAN (Douglas et al., 2014); the A3243G point mutation in the leucine^{UUR} tRNA gene (*MT-TL1*) was identified in patients with FSGS (Jansen et al., 1997; Kurogouchi et al., 1998; Nakamura et al., 1999; Doleris et al., 2000; Hotta et al., 2001; Hirano et al., 2002; Guéry et al., 2003), other forms of renal disease (Guéry et al., 2003) and in a male with a history of MELAS syndrome including kidney cancer, who rapidly developed renal failure after removal of the cancerous

kidney (Piccoli et al., 2012). In general, mtDNA biomarkers have not been considered as potential biomarkers in association studies, therefore most findings concerning the mitochondrial genome in relation to CKD come from case reports. The *MT-TW* tRNA (m.5538 G > A) mutation was identified as causing FSGS in a male (Lim et al., 2017). The (m.547 A > T) and tRNA^{Phe} (m.616 T > C) mutations were found in patients suffering from inherited tubulointerstitial kidney disease, who did not display typical symptoms of mitochondrial disease (Connor et al., 2017). A novel mutation in mtDNA (09155 A > G) was described in a Caucasian female with a history of renal disease, and symptoms of Maternally inherited deafness and diabetes (MIDD) (Adema et al., 2016). Mutations in nuclear genes associated with mitochondrial function have also been associated with renal

FIGURE 3 | Genetic variation or altered expression of nuclear genes which encode mitochondrial proteins may impair respiratory chain complex activities leading to an increase in production of reactive oxygen species (ROS). This initiates a negative feedback loop, further reducing mitochondrial function, and ATP production along with an increase in OXPHOS defects and ROS generation leading to increased oxidative stress which may lead to uncontrolled autophagy, mitophagy and further ROS production. Mitochondrial dysfunction, ROS generation and the resulting dysregulation of autophagic mechanisms may also activate intrinsic apoptotic mechanisms resulting in inflammation and fibrosis in the renal tubules, glomerulus and podocytes eventually leading to kidney disease. Red arrows indicate underexpression and green arrows overexpression.

disease. The P99L mutation in the *BCS1L* gene was found in a female infant suffering from Neonatal Toni–Debré–Fanconi Syndrome, including renal tubulopathy (Ezgu et al., 2013). R45C and R56X mutations in the *BCS1L* gene were described in two siblings suffering from congenital lactic acidosis, including renal tubulopathy (De Meirleir et al., 2003), and nine different mutations in *FBXL4* were identified in nine individuals suffering from mitochondrial encephalomyopathy including renal tubular acidosis (Gai et al., 2013).

Despite the limited published literature, the known significance of the mitochondrial genome with relation to renal function and the multiple case reports relating to individuals suffering from renal dysfunction associated with mutations in mitochondrial or mitochondrial-associated genes, suggest that there exists considerable potential for genetic mutations, resulting in mitochondrial dysfunction, to contribute toward CKD.

X and Y Chromosomes

In CKD research, despite the efforts of extensive GWAS and other genomic analyses in this area, a "blind spot" still

exists in the form of X- and Y-chromosome analysis. Fifty-three of the 3,643 publications found in the online GWAS catalog (hosted by the National Human Genome Research Institute-European Bioinformatics Institute) examined CKD and/or kidney-associated traits (MacArthur et al., 2017). Over 450 genome-wide associations ($p < 5 \times 10^{-8}$) with renal disease and/or related traits were found at 140 loci across the genome (**Table 2**).

As depicted in **Table 2**, the number of associations per chromosome is the lowest for chromosome Y (no associations) and the fourth lowest number of associations for chromosome X (four associations). This is not surprising for chromosome Y. Historically thought of as a "genetic wasteland" (Skaletsky et al., 2003), association analyses usually exclude the Y-chromosome. Indeed, in the 53 studies examining renal disease/traits, only one included the Y-chromosome in the association analysis (Nanayakkara et al., 2014). Given that the Y-chromosome is the smallest and contains the fewest number of genes per number of base pairs (Zerbino et al., 2018), the lack of significant associations in this study is not unexpected.

However, for a chromosome of its size and gene content, the small number of associations found between X-chromosome SNPs and renal disease/traits raises questions as to why there are so few reported. Indeed, the only chromosomes with fewer reported associations are chromosomes 14 and 21, both of which are smaller and contain fewer genes than chromosome X. The lack of reported associations with sex chromosome SNPs could be due to a true lack of association or under-representation of sex chromosomes in GWAS.

Of the 53 GWAS in renal disease/traits, 10 are unclear as to whether X- and Y-chromosome SNPs were included in association analysis. Over half (62%) of the studies did not report sex chromosome association results, with many actively excluding the X- and Y-chromosomes from the association (Chambers et al., 2010; McDonough et al., 2011) or meta-analysis stages (Köttgen et al., 2009). Of the 10 studies (18%) that explicitly state that the X-chromosome analysis was included, only one study found associations between X-chromosome SNPs and renal traits (SCr and eGFR) that reached genome-wide significance (Kanai et al., 2018). Two SNPS, rs12845465, and rs5987107, were both associated with SCr and eGFR ($p < 5 \times 10^{-8}$). In only one study does Y-chromosome analysis appear to be included, where no SNPs reached genome-wide significance (Nanayakkara et al., 2014).

Therefore, with less than 20% of studies reporting X-chromosome results and Y-chromosome exclusion almost ubiquitous, it is not surprising that very few sex chromosome SNPs have shown association in studies of renal disease/traits. A possible explanation for sex chromosome exclusion is that traditional imputation methods call for the use of autosomes only (Marchini et al., 2007). Even now that methods of X-chromosome imputation have been introduced (Marchini and Howie, 2010; König et al., 2014), greater expertise is required and the X-chromosome is imputed separately from the autosomes, and these issues may lead some researchers to simply exclude it. The lack of

TABLE 1 | Studies in mitochondrial genome and in nuclear genes associated with mitochondrial function (FSGS, focal segmental glomerulosclerosis; MELAS, mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes; MIDD, maternally inherited diabetes and deafness).

Gene	Mutation	Disease	Methodology	Function	References
?	09155 A > G	MIDD, renal disease	Whole exome sequencing	?	Adema et al., 2016
MT-TL1	3243 A > G	MELAS syndrome, renal cancer, renal failure	Restriction fragment length polymorphism (ApaI)	Encodes tRNALeu(UUR), forms amino acids containing Leucine	Piccoli et al., 2012
		"Kidney disease"	Allele-specific PCR	Encodes tRNALeu(UUR), forms amino acids containing Leucine	Guéry et al., 2003
		FSGS	Restriction fragment length polymorphism (ApaI)	Encodes tRNALeu(UUR), forms amino acids containing Leucine	Kurogouchi et al., 1998
		Kidney disease, diabetes, hearing loss	Restriction fragment length polymorphism (ApaI)	Encodes tRNALeu(UUR), forms amino acids containing Leucine	Jansen et al., 1997
		Kidney disease, diabetes, hearing loss, cataract	Restriction fragment length polymorphism (ApaI)	Encodes tRNALeu(UUR), forms amino acids containing Leucine	Nakamura et al., 1999
		FSGS, diabetes, hearing loss (3/4 cases), Cerebellar syndrome (1/4 cases)	Allele-specific amplification	Encodes tRNALeu(UUR), forms amino acids containing Leucine	Doleris et al., 2000
		FSGS, diabetes (2/4 cases), hearing loss (1/4 cases)	Restriction fragment length polymorphism (ApaI)	Encodes tRNALeu(UUR), forms amino acids containing Leucine	Hotta et al., 2001
		Renal disease, hearing loss	?	Encodes tRNALeu(UUR), forms amino acids containing Leucine	Hirano et al., 2002
Control region	m.547 A > T	Inherited tubulointerstitial kidney disease	Sanger sequencing	?	Connor et al., 2017
MT-TF	m.616 T > C	Inherited tubulointerstitial kidney disease	Sanger sequencing	Encodes tRNA Phenylalanine, forms amino acids containing phenylalanine	Connor et al., 2017
MT-TW	m.5538 G > A	FSGS	Whole exome sequencing	Encodes tRNA tryptophan, forms amino acids containing tryptophan	Lim et al., 2017
BCS1L	P99L	Neonatal Toni–Debré–Fanconi Syndrome, renal tubulopathy	?	Assembly of mitochondrial respiratory chain complex III	Ezgu et al., 2013
BCS1L	R45C, R56X	Congenital lactic acidosis, renal tubulopathy	Direct sequencing	Assembly of mitochondrial respiratory chain complex III	De Meirleir et al., 2003
FBXL4	(c.1703G > C [p.Gly568Ala], c.1444C > T [p.Arg482Trp], c.1694A > G (p.Asp565Gly), c.1652T > A (p.Ile551Asn), c.1067del (p.Gly356Alafs*15), c.1790A > C (p.Gln597Pro), (c.[614T > C; 106A > T], p.[Ile205Thr; Arg36*]), (c.1229C > T [p.Ser410Phe])	Mitochondrial encephalomyopathy, renal tubular acidosis	Whole exome sequencing	Phosphorylation-dependent ubiquitination in mitochondria	Gai et al., 2013

reported analysis of X-chromosome SNPs in renal disease then leads to its exclusion from meta-analysis, as X-chromosome results are not common between all included studies. Poor genotyping of X-chromosome SNPs may also account for a reduced number of significant associations. Evidence has suggested that removal of X-chromosome SNPs during quality control is significantly more likely, due to a higher rate of chromosomal anomalies or missing call rate than autosomal SNPs (Wise et al., 2013). However, despite the successful imputation of the X-chromosome, chromosome Y lags behind. Despite recent efforts (Zhang et al., 2013), haplogroup-based Y-chromosome imputation is still not widely used, with authors opting to instead use only directly genotyped SNPs (Charchar et al., 2012).

The lack of sex chromosome inclusion in CKD GWAS may be one reason that the relationship between sex and CKD incidence/progression is so unclear. By regularly excluding these chromosomes from renal GWAS, we may miss SNPs that infer either increased CKD risk or protection to one gender in particular.

Traditionally, a greater risk of CKD incidence and progression to ESRD was associated with males. While current evidence still supports an increased rate of progression in men to ESRD (Yang et al., 2014), the risk inferred by gender on incidence of CKD is unclear. A study which used several definitions of incidence found that when using eGFR-based definitions of CKD (<60 ml/min/1.73^2), incident CKD was significantly higher in women than men ($p = 0.02$), but when using a minimum increase in SCr to detect CKD, men had a significantly higher incidence ($p = 0.001$) (Bash et al., 2009). Gender adjustment occurs in eGFR calculation, which may explain this difference. A study conducted to develop a CKD risk score also found that female sex was associated with prevalent CKD ($p = 0.02$) (Bang et al., 2007), as did a Turkish population study ($p < 0.001$) (Süleymanlar et al., 2011). Additionally, a comprehensive review revealed that 38 studies found CKD was more prevalent in women, while 13 found it was more prevalent in men (Hill et al., 2016). Therefore, while women seem to make up a larger proportion of the individuals affected by CKD, affected men seem to progress at a much faster rate, highlighting the difference in the way that CKD affects men and women.

Clinical evidence and recent literature support a link between the sex chromosomes and impaired renal function. Arising as a result of a mutation in *COL4A5* on chromosome X, Alport syndrome is caused by impaired production or function of collagen in various basement membranes throughout the body, including in the glomerulus (Kashtan, 2017). The condition is characterized by hearing loss, ocular abnormalities and

TABLE 2 | Comparison of associations reaching genome-wide significance (5×10^{-8}) per chromosome in renal disease or related traits (bp, base pairs; Chr, chromosome; GWAS, genome-wide association studies).

Chr	Size Chr (bp)	Coding Genes per Chr	Number of Associations with Renal Traits in GWAS Catalog	
			Total	Renal-Associated Loci
1	248956422	2050	21	15
2	242193529	1301	47	12
3	198295559	1079	18	8
4	190214555	753	25	6
5	181538259	884	19	7
6	170805979	1045	60	11
7	159345973	992	26	6
8	145138636	1021	16	5
9	138394717	778	8	3
10	133797422	731	14	6
11	135086622	1316	33	10
12	133275309	1036	26	10
13	114364328	321	7	3
14	107043718	821	1	1
15	101991189	616	33	8
16	90338345	862	29	5
17	83257441	1188	24	4
18	80373285	269	13	5
19	58617616	1474	9	4
20	64444167	543	9	5
21	46709983	232	2	1
22	50818468	492	8	3
X	156040895	846	4	2
Y	57227415	63	0	0
			Total: 452	Total: 140

progression to ESRD, where up to 30% of women reach ESRD by age 60 (Savige et al., 2016) and the majority of affected males will require transplant or dialysis by their late twenties (Temme et al., 2012).

DISCUSSION AND CONCLUSION

Extensive efforts have been made to harness existing GWAS data and improve the sample size statistical power via GWAS meta-analyses to uncover true associations between genetic variants and CKD. Nevertheless, it remains challenging to explain all of the heritability of CKD with currently available methods and datasets.

The definition of CKD phenotype (based on SCr, eGFR and/or urinary albumin measurements) varies between published studies which impacts on the strength of genetic associations observed. CKD is phenotypically heterogeneous and CKD risk may be amplified by co-morbidities such as obesity. Many genetic studies have a cross-sectional case-control design with the determination of CKD based on a single measurements of kidney function. This limits the ability to explore dynamic gene-environment interactions over time, e.g., the impact of diet, gut microbiome, smoking, physical activity, stress, medication use or long-term glycemic control on genetic risk of developing CKD (Simon et al., 2016; Sandoval-Motta et al., 2017).

Prospective follow up of longitudinal cohorts at risk of developing CKD, such as the UK Biobank population may help to unravel some of the complex interplay of genetic background and environmental stressors contributing to kidney damage (Kim et al., 2017). Stratification by co-morbidity, e.g., elevated BMI in T2DM patients, may help identify additional risk variants with a stronger genetic predisposition to CKD (Perry et al., 2012).

The molecular biomarkers for CKD that have received less attention (telomeres, CNVs, mtDNA variants, X and Y chromosomes) are pieces of the missing heritability puzzle. Shorter telomere length is associated with renal dysfunction and CKD progression, even though reported results are conflicting. CNVs have been linked to CAKUT (1q23.1, 22q11, 4p16.1, 7q33, 8q13.2q13.3, and 17q12 regions), PUV (3p25.1p25.2 and 17p12), nephronophthisis (*NPHP1* gene) and IgAN (*DEFA1A3* locus). Information on mtDNA biomarkers is mostly from case

reports, but the A3243G mutation in the *MT-TL1* gene has been associated with FSGS. One GWAS has found associations between X-chromosome SNPs and renal function (rs12845465 and rs5987107). No SNPs in the Y-chromosome have reached genome-wide significance.

Unraveling the missing heritability of CKD will need coherent integration of the different sources contributing to total heritability, and not just inclusion of missing gene variants. Using multiple –"omics" data by combining elements of the phenome, genome, epigenome, transcriptome, metabolome, proteome, and microbiome and translating these data into a useful individual CKD risk assessment remains a major challenge. These research goals efforts will likely help to increase our understanding of the mechanisms of kidney function and disease, and improve disease prediction.

AUTHOR CONTRIBUTIONS

MC-G, KA, RC, RS, LJS, AJM, and APM contributed to the conception or design of the work, acquisition, analysis and interpretation of data for the work, drafting the work and revising it critically for important intellectual content, provided the approval for publication of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

FUNDING

REFERENCES

Adema, A. Y., Janssen, M. C. H., and van der Heijden, J. W. (2016). A novel mutation in mitochondrial DNA in a patient with diabetes, deafness and proteinuria. *Neth. J. Med.* 74, 455–457.

Ai, Z., Li, M., Liu, W., Foo, J. N., Mansouri, O., Yin, P., et al. (2016). Low (-defensin gene copy number increases the risk for IgA nephropathy and renal dysfunction. *Sci. Transl. Med.* 8, 345–388. doi: 10.1126/scitranslmed.aaf2106

Ameh, O. I., Okpechi, I. G., Dandara, C., and Kengne, A. P. (2017). Association between telomere length, chronic kidney disease, and renal traits: a systematic review. *OMICS* 21, 143–155. doi: 10.1089/omi.2016.0180

Arlt, M. F., Wilson, T., and Glover, T. W. (2012). Replication stress and mechanisms of CNV formation. *Curr. Opin. Genet. Dev.* 22, 204–210. doi: 10.1016/j.gde.2012.01.009

Astrup, A. S., Tarnow, L., Jorsal, A., Lajer, M., Nzietchueng, R., Benetos, A., et al. (2010). Telomere length predicts all-cause mortality in patients with type 1 diabetes. *Diabetologia* 53, 45–48. doi: 10.1007/s00125-009-1542-1

Bailie, C., Kilner, J., Maxwell, A. P., and McKnight, A. J. (2017). Development of next generation sequencing panel for UMOD and association with kidney disease. *PLoS One* 12:e0178321. doi: 10.1371/journal.pone.0178321

Bang, H., Vupputuri, S., Shoham, D. A., Klemmer, P. J., Falk, R. J., and Mazumdar, M. (2007). SCreening for occult REnal disease (SCORED). *Arch. Intern. Med.* 167, 374–381. doi: 10.1001/archinte.167.4.374

Bansal, N., Whooley, M. A., Regan, M. E., McCulloch, C. E., Ix, J. H., Epel, E., et al. (2012). Association between kidney function and telomere length: the heart and soul study. *Am. J. Nephrol.* 36, 405–411. doi: 10.1159/000343495

Bash, L. D., Coresh, J., Köttgen, A., Parekh, R. S., Fulop, T., Wang, Y., et al. (2009). Defining incident chronic kidney disease in the research setting: the ARIC study. *Am. J. Epidemiol.* 170, 414–424. doi: 10.1093/aje/kwp151

Bekheirnia, M. R., Bekheirnia, N., Bainbridge, M. N., Gu, S., Coban Akdemir, Z. H., Gambin, T., et al. (2017). Whole-exome sequencing in the molecular diagnosis of individuals with congenital anomalies of the kidney and urinary tract and identification of a new causative gene. *Genet. Med.* 19, 412–420. doi: 10.1038/gim.2016.131

Betjes, M. G., Langerak, A. W., van der Spek, A., de Wit, E. A., and Litjens, N. H. (2011). Premature aging of circulating T cells in patients with end-stage renal disease. *Kidney Int.* 80, 208–217. doi: 10.1038/ki.2011.110

Böger, C. A., Chen, M. H., Tin, A., Olden, M., Köttgen, A., de Boer, I. H., et al. (2011). CUBN is a gene locus for albuminuria. *J. Am. Soc. Nephrol.* 22, 555–570. doi: 10.1681/ASN.2010060598

Bourrat, P., Lu, Q., and Jablonka, E. (2017). Why the missing heritability might not be in the DNA. *Bioessays* 39:1700067. doi: 10.1002/bies.201700067

Boxall, M. C., Goodship, T. H. J., Brown, A. L., Ward, M. C., and von Zglinicki, T. (2006). Telomere shortening and haemodialysis. *Blood Purif.* 24, 185–189. doi: 10.1159/000090517

Broer, L., Codd, V., Nyholt, D. R., Deelen, J., Mangino, M., Willemsen, G., et al. (2013). Meta-analysis of telomere length in 19,713 subjects reveals high heritability, stronger maternal inheritance and a paternal age effect. *Eur. J. Hum. Genet.* 21, 1163–1168. doi: 10.1038/ejhg.2012.303

Butler, M. G., Tilburt, J., DeVries, A., Muralidhar, B., Aue, G., Hedges, L., et al. (1998). Comparison of chromosome telomere integrity in multiple tissues from subjects at different ages. *Cancer Genet. Cytogenet.* 105, 138–144. doi: 10.1016/s0165-4608(98)00029-6

Cañadas-Garre, M., Anderson, K., McGoldrick, J., Maxwell, A. P., and McKnight, A. J. (2018a). Genomic approaches in the search for molecular biomarkers in chronic kidney disease. *J. Transl. Med.* 16:292. doi: 10.1186/s12967-018-1664-7

Cañadas-Garre, M., Anderson, K., McGoldrick, J., Maxwell, A. P., and McKnight, A. J. (2018b). Proteomic and metabolomic approaches in the search for biomarkers in chronic kidney disease. *J. Proteom.* 193, 93–122. doi: 10.1016/j.jprot.2018.09.020

Cao, D. W., Jiang, C. M., Wan, C., Zhang, M., Zhang, Q. Y., Zhao, M., et al. (2018). Upregulation of MiR-126 delays the senescence of human glomerular mesangial cells induced by high glucose via telomere-p53-p21-Rb signaling pathway. *Curr. Med. Sci.* 38, 758–764. doi: 10.1007/s11596-018-1942-x

Carracedo, J., Buendía, P., Merino, A., Soriano, S., Esquivias, E., Martín-Malo, A., et al. (2013). Cellular senescence determines endothelial cell damage induced by uremia. *Exp. Gerontol.* 48, 766–773. doi: 10.1016/j.exger.2013.04.004

Carrero, J. J., Stenvinkel, P., Fellström, B., Qureshi, A. R., Lamb, K., Heimbürger, O., et al. (2008). Telomere attrition is associated with inflammation, low fetuin-A levels and high mortality in prevalent haemodialysis patients. *J. Intern. Med.* 263, 302–312. doi: 10.1111/j.1365-2796.2007.01890.x

Caruana, G., Wong, M. N., Walker, A., Heloury, Y., Webb, N., Johnstone, L., et al. (2014). Copy-number variation associated with congenital anomalies of the kidney and urinary tract. *Pediatr. Nephrol.* 30, 487–495. doi: 10.1007/s00467-014-2962-9

Centers for Disease Control and Prevention (2015). *Chronic Kidney Disease Initiative*. Atlanta: Centers for Disease Control and Prevention. Available at: https://www.cdc.gov/kidneydisease/index.html (accessed December 03, 2018).

Chaban, Y., Boekema, E. J., and Dudkina, N. V. (2014). Structures of mitochondrial oxidative phosphorylation supercomplexes and mechanisms for their stabilisation. *Biochim. Biophys. Acta* 1837, 418–426. doi: 10.1016/j.bbabio.2013.10.004

Chambers, J. C., Zhang, W., Lord, G. M., van der Harst, P., Lawlor, D. A., Sehmi, J. S., et al. (2010). Genetic loci influencing kidney function and chronic kidney disease in man. *Nat. Genet.* 42, 373–375. doi: 10.1038/NG.566

Charchar, F. J., Bloomer, L. D. S., Barnes, T. A., Cowley, M. J., Nelson, C. P., Wang, Y., et al. (2012). Inheritance of coronary artery disease in men: an analysis of the role of the Y chromosome. *Lancet* 379, 915–922. doi: 10.1016/S0140-6736(11)61453-0

Che, R., Yuan, Y., Huang, S., and Zhang, A. (2014). Mitochondrial dysfunction in the pathophysiology of renal diseases. *Am. J. Physiol. Renal Physiol.* 306, F367–F378. doi: 10.1152/ajprenal.00571.2013

Chen, T., Wang, Q., Li, G., and Wang, L. (2016). A single nucleotide polymorphism in the UMOD promoter is associated with end stage renal disease. *BMC Med. Genet.* 17:95. doi: 10.1186/s12881-016-0358-3

Cheng, H., Fan, X., Lawson, W. E., Paueksakon, P., and Harris, R. C. (2015). Telomerase deficiency delays renal recovery in mice after ischemia–reperfusion injury by impairing autophagy. *Kidney Int.* 88, 85–94. doi: 10.1038/ki.2015.69

Cherif, H., Tarry, J. L., Ozanne, S. E., and Hales, C. N. (2003). Ageing and telomeres: a study into organ- and gender-specific telomere shortening. *Nucleic Acids Res.* 31, 1576–1583. doi: 10.1093/nar/gkg208

Codd, V., Nelson, C. P., Albrecht, E., Mangino, M., Deelen, J., Buxton, J. L., et al. (2013). Identification of seven loci affecting mean telomere length and their association with disease. *Nat. Genet.* 45, 422–427.e2. doi: 10.1038/ng.2528

Connor, T. M., Hoer, S., Mallett, A., Gale, D. P., Gomez-Duran, A., and Posse, V. (2017). Mutations in mitochondrial DNA causing tubulointerstitial kidney disease. *PLoS Genet.* 13:e1006620. doi: 10.1371/journal.pgen.1006620

Cooper, G. M. (2000). *The Cell: A Molecular Approach*, 2nd Edn. Sunderland, MA: Sinauer Associates.

Coughlan, M. T., Higgins, G. C., Nguyen, T. V., Penfold, S. A., Thallas-Bonke, V., Tan, S. M., et al. (2016). Deficiency in apoptosis-inducing factor recapitulates chronic kidney disease via aberrant mitochondrial homeostasis. *Diabetes* 65, 1085–1098. doi: 10.2337/db15-0864

Cubiles, M. D., Barroso, S., Vaquero-Sedas, M. I., Enguix, A., Aguilera, A., and Vega-Palas, M. A. (2018). Epigenetic features of human telomeres. *Nucleic Acids Res.* 46, 2347–2355. doi: 10.1093/nar/gky006

De Meirleir, L., Seneca, S., Damis, E., Sepulchre, B., Hoorens, A., Gerlo, E., et al. (2003). Clinical and diagnostic characteristics of complex III deficiency due to mutations in the BCS1L gene. *Am. J. Med. Genet. Part A* 121, 126–131. doi: 10.1002/ajmg.a.20171

De Meyer, T., Nawrot, T., Bekaert, S., De Buyzere, M. L., Rietzschel, E. R., and Andrés, V. (2018). Telomere length as cardiovascular aging biomarker: JACC review topic of the week. *J. Am. Coll. Cardiol.* 72, 805–813. doi: 10.1016/j.jacc.2018.06.014

Desai, K., Berkman, N., Cohen-Manheim, I., Sinnreich, R., Aviv, A., and Kark, J. D. (2018). Rapid shortening of leukocyte telomeres is associated with poorer pulmonary function among healthy adults. *Respir. Med.* 145, 73–79. doi: 10.1016/j.rmed.2018.10.026

Devuyst, O., Olinger, E., and Rampoldi, L. (2017). Uromodulin: from physiology to rare and complex kidney disorders. *Nat. Rev. Nephrol.* 13, 525–544. doi: 10.1038/nrneph.2017.101

Dlouha, D., Maluskova, J., Kralova Lesna, I., Lanska, V., and Hubacek, J. A. (2014). Comparison of the relative telomere length measured in leukocytes and eleven different human tissues. *Physiol. Res.* 63(Suppl. 3), S343–S350.

Doleris, L. M., Hill, G. S., Chedin, P., Nochy, D., Bellanne-Chantelot, C., Hanslik, T., et al. (2000). Focal segmental glomerulosclerosis associated with mitochondrial cytopathy. *Kidney Int.* 58, 1851–1858. doi: 10.1111/j.1523-1755.2000.00356.x

Dolezal, P., Likic, V., Tachezy, J., and Lithgow, T. (2006). Evolution of the molecular machines for protein import into mitochondria. *Science* 313, 314–318. doi: 10.1126/science.1127895

Douglas, A. P., Vance, D. R., Kenny, E. M., Morris, D. W., Maxwell, A. P., and McKnight, A. J. (2014). Next-generation sequencing of the mitochondrial genome and association with IgA nephropathy in a renal transplant population. *Sci. Rep.* 4:7379. doi: 10.1038/srep07379

Dugdale, H. L., and Richardson, D. S. (2018). Heritability of telomere variation: it is all about the environment! *Philos. Trans. R. Soc. B Biol. Sci.* 373:20160450. doi: 10.1098/rstb.2016.0450

Ezgu, F., Senaca, S., Gunduz, M., Tumer, L., Hasanoglu, A., Tiras, U., et al. (2013). Severe renal tubulopathy in a newborn due to BCS1L gene mutation: effects of different treatment modalities on the clinical course. *Gene* 528, 364–366. doi: 10.1016/j.gene.2013.07.007

Faure, A., Bouty, A., Caruana, G., Williams, L., Burgess, T., Wong, M. N., et al. (2016). DNA copy number variants: a potentially useful predictor of early onset renal failure in boys with posterior urethral valves. *J. Pediatr. Urol.* 12, 227.e1–227.e7. doi: 10.1016/j.jpurol.2016.02.020

Fernandez-Marcos, P. J., and Auwerx, J. (2011). Regulation of PGC-1α, a nodal regulator of mitochondrial biogenesis. *Am. J. Clin. Nutr.* 93, 884S–890S. doi: 10.3945/ajcn.110.001917

Foreman, K. J., Marquez, N., Dolgert, A., Fukutaki, K., Fullman, N., McGaughey, M., et al. (2018). Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: reference and alternative scenarios for 2016-40 for 195 countries and territories. *Lancet* 392, 2052–2090. doi: 10.1016/S0140-6736(18)31694-5

Foster, M. C., Coresh, J., Fornage, M., Astor, B. C., Grams, M., and Franceschini, N. (2013). APOL1 variants associate with increased risk of CKD among African Americans. *J. Am. Soc. Nephrol.* 24, 1484–1491. doi: 10.1681/ASN.2013010113

Fyhrquist, F., Tiitu, A., Saijonmaa, O., Forsblom, C., Groop, P. H., and FinnDiane Study Group (2010). Telomere length and progression of diabetic nephropathy in patients with type 1 diabetes. *J. Intern. Med.* 267, 278–286. doi: 10.1111/j.1365-2796.2009.02139.x

Gai, X., Ghezzi, D., Johnson, M. A., Biagosch, C. A., Shamseldin, H. E., and Haack, T. B. (2013). Mutations in FBXL4, encoding a mitochondrial protein, cause early-onset mitochondrial encephalomyopathy. *Am. J. Hum. Genet.* 93, 482–495. doi: 10.1016/j.ajhg.2013.07.016

Galvan, D. L., Green, N. H., and Danesh, F. R. (2017). The hallmarks of mitochondrial dysfunction in chronic kidney disease. *Kidney Int.* 92, 1051–1057. doi: 10.1016/j.kint.2017.05.034

Gao, X., Zhang, Y., Mons, U., and Brenner, H. (2018). Leukocyte telomere length and epigenetic-based mortality risk score: associations with all-cause mortality among older adults. *Epigenetics* 13, 846–857. doi: 10.1080/15592294.2018.1514453

Genovese, G., Friedman, D. J., Ross, M. D., Lecordier, L., Uzureau, P., Freedman, B. I., et al. (2010). Association of trypanolytic ApoL1 variants with kidney disease in African Americans. *Science* 329, 841–845. doi: 10.1126/science.1193032

Gorski, M., van der Most, P. J., Teumer, A., Chu, A. Y., Li, M., Mijatovic, V., et al. (2017). 1000 genomes-based meta-analysis identifies 10 novel loci for kidney function. *Sci. Rep.* 7:45040. doi: 10.1038/srep45040

Granata, S., Dalla Gassa, A., Tomei, P., Lupo, A., and Zaza, G. (2015). Mitochondria: a new therapeutic target in chronic kidney disease. *Nutr. Metab.* 12:49. doi: 10.1186/s12986-015-0044-z

Gray, M. W., Burger, G., and Lang, B. F. (2008). Mitochondrial evolution. *Science* 1476:a011403. doi: 10.1126/science.283.5407.1476

Guan, M., Keaton, J. M., Dimitrov, L., Hicks, P. J., Xu, J., and Palmer, N. D. (2018). An exome-wide association study for type 2 diabetes-attributed end-stage kidney disease in African Americans. *Kidney Int. Rep.* 3, 867–878. doi: 10.1016/j.ekir.2018.03.002

Gudbjartsson, D. F., Holm, H., Indridason, O. S., Thorleifsson, G., Edvardsson, V., Sulem, P., et al. (2010). Association of variants at UMOD with chronic kidney disease and kidney stones-role of age and comorbid diseases. *PLoS Genet.* 6:e1001039. doi: 10.1371/journal.pgen.1001039

Guéry, B., Choukroun, G., Noël, L. H., Clavel, P., Rötig, A., Lebon, S., et al. (2003). The spectrum of systemic involvement in adults presenting with renal lesion and mitochondrial tRNA(Leu) gene mutation. *J. Am. Soc. Nephrol.* 14, 2099–2108. doi: 10.1097/01.ASN.0000080180.51098.02

Gurung, R. L., M, Y., Liu, S., Liu, J. J., and Lim, S. C. (2018). Short leukocyte telomere length predicts albuminuria progression in individuals with type 2 diabetes. *Kidney Int. Rep.* 3, 592–601. doi: 10.1016/j.ekir.2017.12.005

Haas, M. E., Aragam, K. G., Emdin, C. A., Bick, A. G., International Consortium for Blood Pressure, Hemani, G., et al. (2018). Genetic association of albuminuria with cardiometabolic disease and blood pressure. *Am. J. Hum. Genet.* 103, 461–473. doi: 10.1016/j.ajhg.2018.08.004

Hansell, P., Welch, W. J., Blantz, R. C., and Palm, F. (2013). Determinants of kidney oxygen consumption and their relationship to tissue oxygen tension in diabetes and hypertension. *Clin. Exp. Pharmacol. Physiol.* 40, 123–137. doi: 10.1111/1440-1681.12034

Hastings, R., Li, N. C., Lacy, P. S., Patel, H., Herbert, K. E., Stanley, A. G., et al. (2004). Rapid telomere attrition in cardiac tissue of the ageing wistar rat. *Exp. Gerontol.* 39, 855–857. doi: 10.1016/j.exger.2004.02.003

Hill, N. R., Fatoba, S. T., Oke, J. L., Hirst, J. A., O'Callaghan, C. A., Lasserson, D. S., et al. (2016). Global prevalence of chronic kidney disease - a systematic review and meta-analysis. *PLoS One* 11:e0158765. doi: 10.1371/journal.pone.0158765

Hirano, M., Konishi, K., Arata, N., Iyori, M., Saruta, T., Kuramochi, S., et al. (2002). Renal complications in a patient with A-to-G mutation of mitochondrial DNA at the 3243 position of leucine tRNA. *Int. Med.* 41, 113–118. doi: 10.2169/internalmedicine.41.113

Hirashio, S., Nakashima, A., Doi, S., Anno, K., Aoki, E., and Shimamoto, A. (2014). Telomeric G-tail length and hospitalization for cardiovascular events in hemodialysis patients. *Clin. J. Am. Soc. Nephrol.* 9, 2117–2122. doi: 10.2215/CJN.10010913

Hotta, O., Inoue, C. N., Miyabayashi, S., Furuta, T., Takeuchi, A., and Taguma, Y. (2001). Clinical and pathologic features of focal segmental glomerulosclerosis with mitochondrial tRNALeu(UUR) gene mutation. *Kidney Int.* 59, 1236–1243. doi: 10.1046/j.1523-1755.2001.0590041236.x

Jafri, M. A., Ansari, S. A., Alqahtani, M. H., and Shay, J. W. (2016). Roles of telomeres and telomerase in cancer, and advances in telomerase-targeted therapies. *Genome Med.* 8:69. doi: 10.1186/s13073-016-0324-x

Jansen, J. J., Maassen, J. A., van der Woude, F. J., Lemmink, H. A., van den Ouweland, J. M., and t'Hart, L. M. (1997). Mutation in mitochondrial tRNA(Leu(UUR)) gene associated with progressive kidney disease. *J. Am. Soc. Nephrol.* 8, 1118–1124.

Joosten, S. A., van Ham, V., Nolan, C. E., Borrias, M. C., Jardine, A. G., Shiels, P. G., et al. (2003). Telomere shortening and cellular senescence in a model of chronic renal allograft rejection. *Am. J. Pathol.* 162, 1305–1312. doi: 10.1016/S0002-9440(10)63926-0

Kanai, M., Akiyama, M., Takahashi, A., Matoba, N., Momozawa, Y., Ikeda, M., et al. (2018). Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. *Nat. Genet.* 50, 390–400. doi: 10.1038/s41588-018-0047-6

Kao, W. H., Klag, M. J., Meoni, L. A., Reich, D., Berthier-Schaad, Y., Li, M., et al. (2008). MYH9 is associated with nondiabetic end-stage renal disease in African Americans. *Nat. Genet.* 40, 1185–1192. doi: 10.1038/ng.232

Kashtan, C. (2017). Alport syndrome: facts and opinions. *F1000Res.* 6:50. doi: 10.12688/f1000research.9636.1

Kazancioğlu, R. (2013). Risk factors for chronic kidney disease: an update. *Kidney Int. Suppl.* 3, 368–371. doi: 10.1038/kisup.2013.79

Kim, H., Grueneberg, A., Vazquez, A. I., Hsu, S., and de Los Campos, G. (2017). Will big data close the missing heritability gap? *Genetics* 207, 1135–1145. doi: 10.1534/genetics.117.300271

Kim, K., Kim, S.-J., Cho, N.-C., Jung, H.-J., Ahn, K., Ryu, I. W., et al. (2012). Reactive oxygen species-dependent transcriptional regulation of peroxisome proliferator-activated receptor γ coactivator 1α in a human hepatocarcinoma cell line. *Genes Genomics* 34, 709–713. doi: 10.1007/s13258-012-0086-9

Kloosterman, W. P., Guryev, V., van Roosmalen, M., Duran, K. J., de Bruijn, E., Bakker, S. C., et al. (2011). Chromothripsis as a mechanism driving complex de novo structural rearrangements in the germline†. *Hum. Mol. Genet.* 20, 1916–1924. doi: 10.1093/hmg/ddr073

König, I. R., Loley, C., Erdmann, J., and Ziegler, A. (2014). How to include chromosome X in your genome-wide association study. *Genet. Epidemiol.* 38, 97–103. doi: 10.1002/gepi.21782

Kopp, J. B., Smith, M. W., Nelson, G. W., Johnson, R. C., Freedman, B. I., Bowden, D. W., et al. (2008). MYH9 is a major-effect risk gene for focal segmental glomerulosclerosis. *Nat. Genet.* 40, 1175–1184. doi: 10.1038/ng.226

Köttgen, A., Glazer, N. L., Dehghan, A., Hwang, S. J., Katz, R., Li, M., et al. (2009). Multiple loci associated with indices of renal function and chronic kidney disease. *Nat. Genet.* 41, 712–717. doi: 10.1038/ng.377

Köttgen, A., Pattaro, C., Böger, C. A., Fuchsberger, C., Olden, M., Glazer, N. L., et al. (2010). New loci associated with kidney function and chronic kidney disease. *Nat. Genet.* 42, 376–384. doi: 10.1038/ng.568

Köttgen, A., Yang, Q., Shimmin, L. C., Tin, A., Schaeffer, C., Coresh, J., et al. (2012). Association of estimated glomerular filtration rate and urinary uromodulin concentrations with rare variants identified by UMOD gene region sequencing. *PLoS One* 7:e38311. doi: 10.1371/journal.pone.0038311

Kurogouchi, F., Oguchi, T., Mawatari, E., Yamaura, S., Hora, K., Takei, M., et al. (1998). A case of mitochondrial cytopathy with a typical point mutation for MELAS, presenting with severe focal-segmental glomerulosclerosis as main clinical manifestation. *Am. J. Nephrol.* 18, 551–556. doi: 10.1159/000013406

Larsen, C. P., Durfee, T., Wilson, J. D., and Beggs, M. L. (2016). A custom targeted next-generation sequencing gene panel for the diagnosis of genetic nephropathies. *Am. J. Kidney Dis.* 67, 992–993. doi: 10.1053/j.ajkd.2015.11.023

Lata, S., Marasa, M., Li, Y., Fasel, D. A., Groopman, E., Jobanputra, V., et al. (2018). Whole-exome sequencing in adults with chronic kidney disease: a pilot study. *Ann. Intern. Med.* 168, 100–109. doi: 10.7326/M17-1319

Leak, T. S., Perlegas, P. S., Smith, S. G., Keene, K. L., Hicks, P. J., Langefeld, C. D., et al. (2009). Variants in intron 13 of the ELMO1 gene are associated with diabetic nephropathy in African Americans. *Ann. Hum. Genet.* 73, 152–159. doi: 10.1111/j.1469-1809.2008.00498.x

Levey, A. S., Atkins, R., Coresh, J., Cohen, E. P., Collins, A. J., Eckardt, K.-U., et al. (2007). Chronic kidney disease as a global public health problem: approaches and initiatives – a position statement from kidney disease improving global outcomes. *Kidney Int.* 72, 247–259. doi: 10.1038/SJ.KI.5002343

Lew, A. R., Kellermayer, T. R., Sule, B. P., and Szigeti, K. (2018). Copy number variations in adult-onset neuropsychiatric diseases. *Curr. Genomics* 19, 420–430. doi: 10.2174/1389202919666180330153842

Liang, L., Fang, J. Y., and Xu, J. (2016). Gastric cancer and gene copy number variation: emerging cancer drivers for targeted therapy. *Oncogene* 35, 1475–1482. doi: 10.1038/onc.2015.209

Lim, K., Steele, D., Fenves, A., Thadhani, R., Heher, E., and Karaa, A. (2017). Focal segmental glomerulosclerosis associated with mitochondrial disease. *Clin. Nephrol. Case Stud.* 5, 20–25. doi: 10.5414/CNCS109083

Lodish, H., Berk, A., Zipursky, P., Lawrence, S. M., Baltimore, D., and Darnell, J. (eds). (2012). "Electron transport and oxidative phosphorylation," in *Molecular Cell Biology*, 4th Edn, Vol. 474, (New York, NY: W.H. Freeman), 1187. doi: 10.1016/S1470-8175(01)00023-6

Lu, A. T., Xue, L., Salfati, E. L., Chen, B. H., Ferrucci, L., Levy, D., et al. (2018). GWAS of epigenetic aging rates in blood reveals a critical role for TERT. *Nat. Commun.* 9:387. doi: 10.1038/s41467-017-02697-5

Lu, Y. Y., Yang, X., Chen, W. Q., Ju, Z. Y., Shou, Z. F., Jin, J., et al. (2014). Proteins induced by telomere dysfunction are associated with human IgA nephropathy. *J. Zhejiang Univ. Sci. B* 15, 566–574. doi: 10.1631/jzus.B1300115

MacArthur, J., Bowler, E., Cerezo, M., Gil, L., Hall, P., Hastings, E., et al. (2017). The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic. Acids Res.* 45, D896–D901. doi: 10.1093/nar/gkw1133

Mangaonkar, A. A., and Patnaik, M. M. (2018). Short telomere syndromes in clinical practice: bridging bench and bedside. *Mayo Clin. Proc.* 93, 904–916. doi: 10.1016/j.mayocp.2018.03.020

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., et al. (2009). Finding the missing heritability of complex diseases. *Nature* 461, 747–753. doi: 10.1038/nature08494

Marchini, J., and Howie, B. (2010). Genotype imputation for genome-wide association studies. *Nat. Rev. Genet.* 11, 499–511. doi: 10.1038/nrg2796

Marchini, J., Howie, B., Myers, S., McVean, G., and Donnelly, P. (2007). A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.* 39, 906–913. doi: 10.1038/ng2088

Marnett, L. J. (2000). Oxyradicals and DNA damage. *Carcinogenesis* 21, 361–370. doi: 10.1093/carcin/21.3.361

McDonough, C. W., Nicholette, D. P., Hicks, P. J., Roh, B. H., An, S. S., Cooke, J. N., et al. (2011). A genome-wide association study for diabetic nephropathy genes in African Americans. *Kidney Int.* 79, 563–572. doi: 10.1038/ki.2010.467

Meiklejohn, C. D., Marissa, A. H., Mohammad, A. S., Dawn, N. A., David, M. R., and Kristi, L. M. (2013). An incompatibility between a mitochondrial tRNA and its nuclear-encoded tRNA synthetase compromises development and fitness in *Drosophila*. *PLoS Genet.* 9:e1003238. doi: 10.1371/journal.pgen.1003238

Melk, A., Ramassar, V., Helms, L. M., Moore, R., Rayner, D., Solez, K., et al. (2000). Telomere shortening in kidneys with age. *J. Am. Soc. Nephrol.* 11, 444–453.

Moon, D. H., Segal, M., Boyraz, B., Guinan, E., Hofmann, I., Cahan, P., et al. (2015). Poly(A)-specific ribonuclease (PARN) mediates 3'-end maturation of the telomerase RNA component. *Nat. Genet.* 47, 1482–1488. doi: 10.1038/ng.3423

Nakamura, S., Yoshinari, M., Doi, Y., Yoshizumi, H., Katafuchi, R., Yokomizo, Y., et al. (1999). Renal complications in patients with diabetes mellitus associated with an A to G mutation of mitochondrial DNA at the 3243 position of leucine tRNA. *Diabetes Res. Clin. Pract.* 44, 183–189. doi: 10.1016/s0168-8227(99)00051-0

Nanayakkara, S., Senevirathna, S. T., Abeysekera, T., Chandrajith, R., Ratnatunga, N., Gunarathne, E. D., et al. (2014). An Integrative Study of the Genetic, social and environmental determinants of chronic kidney disease characterized by tubulointerstitial damages in the North Central Region of Sri Lanka. *J. Occup. Health* 56, 28–38. doi: 10.1539/joh.13-0172-oa

Narres, M., Claessen, H., Droste, S., Kvitkina, T., Koch, M., Kuss, O., et al. (2016). The incidence of end-stage renal disease in the diabetic (Compared to the Non-Diabetic) population: a systematic review. *PLoS One* 11:e0147329. doi: 10.1371/journal.pone.0147329

Nazaryan-Petersen, L., Eisfeldt, J., Pettersson, M., Lundin, J., Nilsson, D., Wincent, J., et al. (2018). Replicative and non-replicative mechanisms in the formation of clustered CNVs are indicated by whole genome characterization. *PLoS Genet.* 14:e1007780. doi: 10.1371/journal.pgen.1007780

Nolte, I. M., van der Most, P. J., Alizadeh, B. Z., de Bakker, P. I., Boezen, H. M., Bruinenberg, M., et al. (2017). Missing heritability: is the gap closing? An analysis of 32 complex traits in the lifelines cohort study. *Eur. J. Hum. Genet.* 25, 877–885. doi: 10.1038/ejhg.2017.50

O'Seaghdha, C. M., and Fox, C. S. (2011). Genome-wide association studies of chronic kidney disease: what have we learned? *Nat. Rev. Nephrol.* 8, 89–99. doi: 10.1038/nrneph.2011.189

Parsa, A., Fuchsberger, C., Köttgen, A., O'Seaghdha, C. M., Pattaro, C., de Andrade, M., et al. (2013). Common variants in mendelian kidney disease genes and their association with renal function. *J. Am. Soc. Nephrol.* 24, 2105–2117. doi: 10.1681/ASN.2012100983

Pattaro, C., Köttgen, A., Teumer, A., Garnaas, M., Böger, C. A., Fuchsberger, C., et al. (2012). Genome-wide association and functional follow-up reveals new loci for kidney function. *PLoS Genet.* 8:e1002584. doi: 10.1371/journal.pgen.1002584

Pattaro, C., Teumer, A., Gorski, M., Chu, A. Y., Li, M., Mijatovic, V., et al. (2016). Genetic associations at 53 loci highlight cell types and biological pathways relevant for kidney function. *Nat. Commun.* 7:10023. doi: 10.1038/ncomms10023

Perry, J. R. B., Voight, B. F., Yengo, L., Amin, N., Josée, D., Martha, G., et al. (2012). Stratifying type 2 diabetes cases by BMI identifies genetic risk variants in LAMA1 and enrichment for risk variants in lean compared to obese cases. *PLoS Genet.* 8:e1002741. doi: 10.1371/journal.pgen.1002741

Pezzolesi, M. G., Katavetin, P., Kure, M., Poznik, G. D., Skupien, J., and Mychaleckyj, J. C. (2009a). Confirmation of genetic associations at ELMO1 in the GoKinD collection supports its role as a susceptibility gene in diabetic nephropathy. *Diabetes* 58, 2698–2702. doi: 10.2337/db09-0641

Pezzolesi, M. G., Poznik, G. D., Mychaleckyj, J. C., Paterson, A. D., Barati, M. T., Klein, J. B., et al. (2009b). Genome-wide association scan for diabetic nephropathy susceptibility genes in type 1 diabetes. *Diabetes* 58, 1403–1410. doi: 10.2337/db08-1514

Piccoli, G. B., Laura, D. B., Paola, C., Federica, N. V., Martina, F., Federica, F., et al. (2012). Chronic kidney disease, severe arterial and arteriolar sclerosis and kidney neoplasia: on the spectrum of kidney involvement in MELAS syndrome. *BMC Nephrol.* 13:9. doi: 10.1186/1471-2369-13-9

Prokop, J. W., Yeo, N. C., Ottmann, C., Chhetri, S. B., Florus, K. L., Ross, E. J., et al. (2018). Characterization of coding/noncoding variants for SHROOM3 in patients with CKD. *J. Am. Soc. Nephrol.* 29, 1525–1535. doi: 10.1681/ASN.2017080856

Rahman, S., and Hall, A. M. (2013). Mitochondrial disease–an important cause of end-stage renal failure. *Pediatr. Nephrol.* 28, 357–361. doi: 10.1007/s00467-012-2362-y

Raschenberger, J., Kollerits, B., Ritchie, J., Lane, B., Kalra, P. A., and Ritz, E. (2015). Association of relative telomere length with progression of chronic kidney disease in two cohorts: effect modification by smoking and diabetes. *Sci. Rep.* 5:11887. doi: 10.1038/srep11887

Rasouly, H. M., Groopman, E. E., Heyman-Kantor, R., Fasel, D. A., Mitrotti, A., Westland, R., et al. (2018). The burden of candidate pathogenic variants for kidney and genitourinary disorders emerging from exome sequencing. *Ann. Intern. Med.* doi: 10.7326/M18-1241 [Epub ahead of print].

Regele, F., Jelencsics, K., Shiffman, D., Paré, G., McQueen, M. J., and Mann, J. F. (2015). Genome-wide studies to identify risk factors for kidney disease with a focus on patients with diabetes. *Nephrol. Dial. Transplant.* 30(Suppl. 4), iv26–iv34. doi: 10.1093/ndt/gfv087

Ricardo, A. C., Yang, W., Sha, D., Appel, L. J., Chen, J., Krousel-Wood, M., et al. (2018). Sex-related disparities in CKD progression. *J. Am. Soc. Nephrol.* 30, 137–146. doi: 10.1681/ASN.2018030296

Rizvi, S., Raza, S. T., and Mahdi, F. (2014). Telomere length variations in aging and age-related diseases. *Curr. Aging Sci.* 7, 161–167. doi: 10.2174/1874609808666150122153151

Rodrigues, C. E., Capcha, J. M., de Bragança, A. C., Sanches, T. R., Gouveia, P. Q., de Oliveira, P. A., et al. (2017). Human umbilical cord-derived mesenchymal stromal cells protect against premature renal senescence resulting from oxidative stress in rats with acute kidney injury. *Stem Cell Res. Ther.* 8:19. doi: 10.1186/s13287-017-0475-8

Sampson, M. G. (2016). Actualizing the benefits of genomic discovery in pediatric nephrology. *J. Pediatr. Genet.* 5, 69–75. doi: 10.1055/s-0035-1557113

Sandoval-Motta, S., Aldana, M., Martínez-Romero, E., and Frank, A. (2017). The human microbiome and the missing heritability problem. *Front. Genet.* 8:80. doi: 10.3389/fgene.2017.00080

Sanna-Cherchi, S., Kiryluk, K., Burgess, K. E., Bodria, M., Sampson, M. G., Hadley, D., et al. (2012). Copy-number disorders are a common cause of congenital kidney malformations. *Am. J. Hum. Genet.* 91, 987–997. doi: 10.1016/j.ajhg.2012.10.007

Satko, S. G., and Freedman, B. I. (2005). The familial clustering of renal disease and related phenotypes. *Med. Clin. North Am.* 89, 447–456. doi: 10.1016/J.MCNA.2004.11.011

Savige, J., Colville, D., Rheault, M., Gear, S., Lennon, R., and Lagas, S. (2016). Special feature alport syndrome in women and girls. *Clin. J. Am. Soc. Nephrol.* 11, 1713–1720. doi: 10.2215/CJN.00580116

Sfeir, A. (2012). Telomeres at a glance. *J. Cell Sci.* 125(Pt 18), 4173–4178. doi: 10.1242/jcs.106831

Shimazaki, A., Kawamura, Y., Kanazawa, A., Sekine, A., Saito, S., Tsunoda, T., et al. (2005). Genetic variations in the gene encoding ELMO1 are associated with susceptibility to diabetic nephropathy. *Diabetes* 54, 1171–1178. doi: 10.2337/diabetes.54.4.1171

Simon, P. H., Sylvestre, M. P., Tremblay, J., and Hamet, P. (2016). Key considerations and methods in the study of gene–environment interactions. *Am. J. Hypertens.* 29, 891–899. doi: 10.1093/ajh/hpw021

Siomou, E., Mitsioni, A. G., Giapros, V., Bouba, I., Noutsopoulos, D., and Georgiou, I. (2017). Copy-number variation analysis in familial nonsyndromic congenital anomalies of the kidney and urinary tract: evidence for the causative role of a transposable element-associated genomic rearrangement. *Mol. Med. Rep.* 15, 3631–3636. doi: 10.3892/mmr.2017.6462

Skaletsky, H., Kuroda-Kawaguchi, T., Minx, P. J., Cordum, H. S., Hillier, L., Brown, L. G., et al. (2003). The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* 423, 825–837. doi: 10.1038/nature01722

Skelly, R., Maxwell, A. P., and McKnight, A. J. (2019). Mitochondria and chronic kidney disease: a molecular update. *SPG Biomed.* (in press). doi: 10.32392/biomed.36

Snoek, R., van Setten, J., Keating, B. J., Israni, A. K., Jacobson, P. A., Oetting, W. S., et al. (2018). NPHP1 (Nephrocystin-1) gene deletions cause adult-onset ESRD. *J. Am. Soc. Nephrol.* 29, 1772–1779. doi: 10.1681/ASN.2017111200

Song, J., Czerniak, S., Wang, T., Ying, W., Carlone, D. L., and Breault, D. T. (2011). Characterization and fate of telomerase-expressing epithelia during kidney repair. *J. Am. Soc. Nephrol.* 22, 2256–2265. doi: 10.1681/ASN.2011050447

Song, Y. R., You, S. J., Lee, Y. M., Chin, H. J., Chae, D. W., and Oh, Y. K. (2010). Activation of hypoxia-inducible factor attenuates renal injury in rat remnant kidney. *Nephrol. Dial. Transplant.* 25, 77–85. doi: 10.1093/ndt/gfp454

Süleymanlar, G., Utaş, C., Arinsoy, T., Ateş, K., Altun, B., Altiparmak, M. R., et al. (2011). A population-based survey of chronic REnal disease in turkey–the CREDIT study. *Nephrol. Dial. Transplant.* 26, 1862–1871. doi: 10.1093/ndt/gfq656

Sveinbjornsson, G., Mikaelsdottir, E., Palsson, R., Indridason, O. S., Holm, H., Jonasdottir, A., et al. (2014). Rare mutations associating with serum creatinine and chronic kidney disease. *Hum. Mol. Genet.* 23, 6935–6943. doi: 10.1093/hmg/ddu399

Swan, E. J., Salem, R. M., Sandholm, N., Tarnow, L., Rossing, P., and Lajer, M. (2015). Genetic risk factors affecting mitochondrial function are associated with kidney disease in people with type 1 diabetes. *Diabet. Med.* 32, 1104–1109. doi: 10.1111/dme.12763

Szeto, C. C., Poon, P. Y., Lai, F. M., Chow, K. M., Szeto, C. Y., and Li, P. K. (2005). Chromosomal telomere shortening of kidney cells in IgA nephropathy by the measurement of DNA in urinary sediment. *Clin. Nephrol.* 64, 337–342. doi: 10.5414/cnp64337

Taanman, J. W. (1999). The mitochondrial genome: structure, transcription, translation and replication. *Biochim. Biophys. Acta* 1410, 103–123. doi: 10.1016/S0005-2728(98)00161-3

Tanaka, T., Kojima, I., Ohse, T., Ingelfinger, J. R., Adler, S., and Fujita, T. (2005). Cobalt promotes angiogenesis via hypoxia-inducible factor and protects tubulointerstitium in the remnant kidney model. *Lab. Invest.* 85, 1292–1307. doi: 10.1038/labinvest.3700328

Temme, J., Kramer, A., Jager, K. J., Lange, K., Peters, F., Müller, G. A., et al. (2012). Outcomes of male patients with alport syndrome undergoing renal replacement therapy. *Clin. J. Am. Soc. Nephrol.* 7, 1969–1976. doi: 10.2215/CJN.02190312

Tentolouris, N., Nzietchueng, R., Cattan, V., Poitevin, G., Lacolley, P., and Papazafiropoulou, A. (2007). White blood cells telomere length is shorter in males with type 2 diabetes and microalbuminuria. *Diabetes Care* 30, 2909–2915. doi: 10.2337/dc07-0633

Testa, R., Olivieri, F., Sirolla, C., Spazzafumo, L., Rippo, M. R., Marra, M., et al. (2011). Leukocyte telomere length is associated with complications of type 2 diabetes mellitus. *Diabet. Med.* 28, 1388–1394. doi: 10.1111/j.1464-5491.2011.03370.x

Teumer, A., Tin, A., Sorice, R., Gorski, M., Yeo, N. C., Chu, A. Y., et al. (2016). Genome-wide association studies identify genetic loci associated with albuminuria in diabetes. *Diabetes* 65, 803–817. doi: 10.2337/db15-1313

Thapar, A., and Cooper, M. (2013). Copy number variation: what is it and what has it told us about child psychiatric disorders? *J. Am. Acad. Child Adolesc. Psychiatry* 52, 772–774. doi: 10.1016/j.jaac.2013.05.013

Thomas, R., Sanna-Cherchi, S., Warady, B. A., Furth, S. L., Kaskel, F. J., and Gharavi, A. G. (2011). HNF1B and PAX2 mutations are a common cause of renal hypodysplasia in the CKiD cohort. *Pediatr. Nephrol.* 26, 897–903. doi: 10.1007/s00467-011-1826-9

Timmis, J. N., Ayliffe, M. A., Huang, C. Y., and Martin, W. (2004). Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat. Rev. Genet.* 5, 123–135. doi: 10.1038/nrg1271

Townsley, D. M., Dumitriu, B., and Young, N. S. (2016). Danazol treatment for telomere diseases. *N. Engl. J. Med.* 374, 1922–1931. doi: 10.1056/NEJMoa1515319

Tucker, P. S., Scanlan, A. T., Vella, R. K., and Dalbo, V. J. (2015). Genomic integrity is favourably affected by high-intensity interval training in an animal model of early-stage chronic kidney disease. *Sports Med. Open* 2:28. doi: 10.1186/s40798-016-0055-y

Tzur, S., Rosset, S., Shemer, R., Yudkovsky, G., Selig, S., Tarekegn, A., et al. (2010). Missense mutations in the APOL1 gene are highly associated with end stage kidney disease risk previously attributed to the MYH9 gene. *Hum. Genet.* 128, 345–350. doi: 10.1007/s00439-010-0861-0

Verbitsky, M., Sanna-Cherchi, S., Fasel, D. A., Levy, B., Kiryluk, K., Wuttke, M., et al. (2015). Genomic imbalances in pediatric patients with chronic kidney disease. *J. Clin. Invest.* 125, 2171–2178. doi: 10.1172/JCI80877

Verzola, D., Gandolfo, M. T., Gaetani, G., Ferraris, A., Mangerini, R., Ferrario, F., et al. (2008). Accelerated senescence in the kidneys of patients with type 2 diabetic nephropathy. *Am. J. Physiol. Renal Physiol.* 295, F1563–F1573. doi: 10.1152/ajprenal.90302.2008

Wallace, D. C. (2013). A mitochondrial bioenergetic etiology of disease. *J. Clin. Investig.* 123, 1405–1412. doi: 10.1172/JCI61398

Wang, Q., Zhan, Y., Pedersen, N. L., Fang, F., and Hägg, S. (2018). Telomere length and all-cause mortality: a meta-analysis. *Ageing Res. Rev.* 48, 11–20. doi: 10.1016/j.arr.2018.09.002

Westhoff, J. H., Schildhorn, C., Jacobi, C., Hömme, M., Hartner, A., Braun, H., et al. (2010). Telomere shortening reduces regenerative capacity after acute kidney injury. *J. Am. Soc. Nephrol.* 21, 327–336. doi: 10.1681/ASN.2009010072

Willis, M., Reid, S. N., Calvo, E., Staudinger, U. M., and Factor-Litvak, P. (2018). A scoping systematic review of social stressors and various measures of telomere length across the life course. *Ageing Res. Rev.* 47, 89–104. doi: 10.1016/j.arr.2018.07.006

Wise, A. L., Gyi, L., and Manolio, T. A. (2013). eXclusion: toward integrating the X chromosome in genome-wide association analyses. *Am. J. Hum. Genet.* 92, 643–647. doi: 10.1016/j.ajhg.2013.03.017

Wong, L. S., Windt, W. A., Roks, A. J., van Dokkum, R. P., Schoemaker, R. G., and de Zeeuw, D. (2009). Renal failure induces telomere shortening in the rat heart. *Neth. Heart J.* 17, 190–194. doi: 10.1007/bf03086245

Yang, W., Xie, D., Anderson, A. H., Joffe, M. M., Greene, T., Teal, V., et al. (2014). Association of kidney disease outcomes with risk factors for CKD: findings from the chronic renal insufficiency cohort (CRIC) study. *Am. J. Kidney Dis.* 63, 236–243. doi: 10.1053/j.ajkd.2013.08.028

Ye, J., Li, J., Yu, Y., Wei, Q., Deng, W., and Yu, L. (2010). L-carnitine attenuates oxidant injury in HK-2 cells via ROS-mitochondria pathway. *Regul. Pept.* 161, 58–66. doi: 10.1016/j.regpep.2009.12.024

Zaza, G., Granata, S., Masola, V., Rugiu, C., Fantin, F., and Gesualdo, L. (2013). Downregulation of nuclear-encoded genes of oxidative metabolism in dialyzed chronic kidney disease patients. *PLoS One* 8:e77847. doi: 10.1371/journal.pone.0077847

Zerbino, D. R., Achuthan, P., Akanni, W., Amode, M. R., Barrell, D., Bhai, J., et al. (2018). Ensembl 2018. *Nucleic Acids Res.* 46, D754–D761. doi: 10.1093/nar/gkx1098

Zhan, M., Brooks, C., Liu, F., Sun, L., and Dong, Z. (2013). Mitochondrial dynamics: regulatory mechanisms and emerging role in renal pathophysiology. *Kidney Int.* 83, 568–581. doi: 10.1038/ki.2012.441

Zhang, F., Chen, R., Liu, D., Yao, X., Li, G., and Jin, Y. (2013). YHap: a population model for probabilistic assignment of y haplogroups from re-sequencing data. *BMC Bioinformatics* 14:331. doi: 10.1186/1471-2105-14-331

Zuk, O., Hechter, E., Sunyaev, S. R., and Lander, E. S. (2012). The mystery of missing heritability: genetic interactions create phantom heritability. *Proc. Natl. Acad. Sci. U.S.A.* 109, 1193–1198. doi: 10.1073/pnas.1119675109

Genetic and Epigenetic Studies in Diabetic Kidney Disease

Harvest F. Gu*

Center for Pathophysiology, School of Basic Medicine and Clinical Pharmacy, China Pharmaceutical University, Nanjing, China

Chronic kidney disease is a worldwide health crisis, while diabetic kidney disease (DKD) has become the leading cause of end-stage renal disease (ESRD). DKD is a microvascular complication and occurs in 30–40% of diabetes patients. Epidemiological investigations and clinical observations on the familial clustering and heritability in DKD have highlighted an underlying genetic susceptibility. Furthermore, DKD is a progressive and long-term diabetic complication, in which epigenetic effects and environmental factors interact with an individual's genetic background. In recent years, researchers have undertaken genetic and epigenetic studies of DKD in order to better understand its molecular mechanisms. In this review, clinical material, research approaches and experimental designs that have been used for genetic and epigenetic studies of DKD are described. Current information from genetic and epigenetic studies of DKD and ESRD in patients with diabetes, including the approaches of genome-wide association study (GWAS) or epigenome-wide association study (EWAS) and candidate gene association analyses, are summarized. Further investigation of molecular defects in DKD with new approaches such as next generation sequencing analysis and phenome-wide association study (PheWAS) is also discussed.

Keywords: diabetic kidney disease, diabetes, end-stage renal disease, genetics, epigenetics, phenotypes

INTRODUCTION

Diabetes is a major public health problem that is approaching epidemic proportions globally. According to the latest report from the IDF, the prevalence of diabetes will increase from 425 million persons in 2017 to 629 million by 2045 (IDF 2017[1]). Diabetic kidney disease (DKD, previously termed diabetic nephropathy, DN) is a microvascular complication and progresses gradually over many years in approximately 30–40% of individuals with T1D and T2D mellitus (Harjutsalo and Groop, 2014; Thomas et al., 2015; Barrett et al., 2017). DKD is now the main cause of chronic kidney disease (CKD) worldwide and the leading cause of end-stage-renal disease (ESRD) requiring renal replacement therapy (dialysis or transplantation). The presence of CKD is the single strongest predictor of mortality for persons with diabetes (Dousdampanis et al., 2016; Papadopoulou-Marketou et al., 2017). Pathological findings in DKD include glomerular

---

Abbreviations: ACR, albumin-to-creatinine ratio; ADA, American Diabetes Association; BMI, body mass index; CNV, copy number variant; DKD, diabetic kidney disease; ESRD, end-stage renal disease; EWAS, epigenome-wide association study; GFR, glomerular filtration rate; GWAS, genome-wide association study; IDF, International Diabetes Federation; IHME, Institute for Health Metrics and Evaluation; LD, Linkage disequilibrium; PheWAS, phenome-wide association study; SNP, single nucleotide polymorphism; T1D, type 1 diabetes; T2D, type 2 diabetes; UAE, urinary albumin excretion.
[1]http://www.diabetesatlas.org/

hypertrophy, mesangial matrix expansion, reduced podocyte number, glomerulosclerosis, tubular atrophy and tubulointerstitial fibrosis. Clinical criteria used to diagnose the subjects with DKD are urine ACR higher than 300 mg/g, while microalbuminuria is diagnosed when ACR is between 30–300 mg/g (Bouhairie and McGill, 2016). Accumulating evidence has indicated that podocyte loss and epithelial dysfunction play important roles in DKD pathogenesis with further progression associated with inflammation but the exact molecular mechanisms responsible for DKD are not fully known (Badal and Danesh, 2014; Reidy et al., 2014; Gnudi et al., 2016).

Both clinical and epidemiological studies have demonstrated that there is familial aggregation of DKD in different ethnic groups, indicating that genetic factors contribute to development of the disease. Furthermore, genetic risk factors in DKD interact with the environmental factors (for example, lifestyle, diet and medication) (Freedman et al., 2007a; Murea et al., 2012; Thomas et al., 2012; Kato and Natarajan, 2014). **Figure 1** is a schematic diagram representing the relationship between genetic, epigenetic and environmental factors that are involved in the development and progression of DKD. Genetic studies of DKD are mainly focused on association analyses between genomic DNA variation (for example, single nucleotide polymorphisms, SNPs, copy number variants, CNVs, and microsatellites) and clinical phenotypes of the disease (Freedman et al., 2007a; Gu and Brismar, 2012; Thomas et al., 2012; Florez, 2016). Epigenetics studies of DKD examine potentially heritable changes in gene expression that occur without variation in the original DNA nucleotide sequence (Villeneuve and Natarajan, 2010; Kato and Natarajan, 2014; Thomas, 2016; Keating et al., 2018). Therefore, epigenetic studies of DKD may provide information to help understand how environmental factors modify the expression of genes that are involved in DKD progression. Combined genetic, epigenetic and phenotypic studies together may generate information to understand new pathogenic pathways and to search for new biomarkers for early diagnosis and prediction as part of prevention programs in DKD. The results may also be useful in finding novel targets for the treatment of DKD.

SNPs are the most common form of genomic DNA variation. The updated dbSNP database of more than 500 million reference SNPs (rs) with allele frequency data[2] has provided fundamental information for genetic studies of complex diseases including, DKD. The genetic studies in DKD have implicated previously unsuspected biological pathways and subsequently improved our knowledge for understanding of the genetic basis of the disease. For most common traits studied in DKD, however, the identified genes and their SNPs only explain a fraction of associated risk, suggesting that human genomic DNA variations are only a part of underlying susceptibility to DKD. This has led to evolving interest in epigenetics to help explain some of the missing heritability of DKD. Epigenetic mechanisms mainly consist of DNA methylation, chromosome histone modification and non-coding RNA (ncRNA) regulation (Kato and Natarajan, 2014; Allis and Jenuwein, 2016). Epigenetic related ncRNAs include miRNA, siRNA, piRNA, and lncRNA (Holoch and Moazed, 2015).

There are more than 30,000 identified CpG islands in the human genome. Detailed information for these CpG islands can be found in the public database[3]. The CpG islands are defined as stretches of DNA > 200 bp long with a GC percentage greater than 50% and an observed-to-expected CpG ratio of more than 60%. The CpG islands are often found at promoters and contain the 5′ end of the transcript, while DNA methylation occurs at 5′-cytosines of "CpG" dinucleotides[4] (Cross and Bird, 1995). In DKD, the effects of DNA methylation have been studied in terms of transgenerational inheritance of the disease to explore environmental and other non-genetic factors that may influence epigenetic modifications in the genes involved in DKD (Deaton and Bird, 2011; Jones, 2012). Identification of differentially methylated CpG sites in promoters or other functional regions of genes and the analysis of the DNA methylation changes that are associated with DKD have become the most common approaches used in epigenetic studies of the disease (Villeneuve and Natarajan, 2010; Kato and Natarajan, 2014; Thomas, 2016). Furthermore, ncRNAs, particularly long ncRNAs are known to be involved in epigenetic processes. ncRNAs certainly play an important role in chromatin formation, histone modification, DNA methylation and consequently gene transcription silencing.

Genetic and epigenetic studies of DKD, initially using candidate gene approaches and more recently at genome-wide scale (known as GWAS and EWAS), have been undertaken to identify many genes conferring susceptibility or resistance to DKD. In this review, clinical phenotypes, research approaches and experimental designs that have been used for genetic and epigenetic studies of DKD are described. These research approaches and experimental designs can also be used for study of CKD. Current information from genetic and epigenetic studies of DKD is summarized. Further investigation of molecular defects in DKD with new generation sequencing analyses and phenome-wide association studies (PheWAS) are discussed.

BIOLOGICAL MATERIAL, RESEARCH APPROACHES AND STUDY DESIGNS USED IN GENETIC AND EPIGENETIC INVESTIGATIONS OF DIABETIC KIDNEY DISEASE

Two major research approaches either at genome-wide scale or focused on candidate gene(s) have been widely used for comparative studies between cases (patients with DKD) and controls (diabetes patients without DKD). Case-control studies by recruiting large numbers of subjects can increase the statistical power of reported associations. The aim is to discover the genes presented differentially in genomic structure or genetic expression. Genome-wide or epigenome-wide association studies (GWAS or EWAS) are hypothesis—generating approaches (Rakyan et al., 2011; Do et al., 2017; Lappalainen and Greally, 2017). These

---

[2] https://www.ncbi.nlm.nih.gov/feed/rss.cgi?ChanKey=dbsnpnews

[3] https://genome.ucsc.edu/cgi-bin/hgTables
[4] https://en.wikipedia.org/wiki/CpG_site

FIGURE 1 | This is a schematic diagram representing the relationship between genetic, epigenetic and phenotypic studies in diabetic kidney disease (DKD). Genetic association studies are fundamentally important for identification of susceptibility or resistance genes (G). Epigenetic studies analyzing genomic DNA methylation changes, chromosome histone modification and ncRNA regulation are useful for dissecting the interaction of the genes with environmental factors. The combined data from genetic, epigenetic and phenotypic (Phe) studies may provide the opportunity for us to understand new pathways underlying the pathogenesis of DKD and to discover new biomarkers for early diagnosis and to find targets for prevention and treatment programs of this disease. The different sizes of the 'G" and "Phe" represent the variation of genetic and phenotypic effects.

study designs have benefited from rapid development of human genome research, including the creation of publicly available databases of SNPs, haplotypes and CpG islands and the rapid technical improvements in analyzing genomic variation using high-throughput techniques and high-density SNP or CpG arrays. Another approach is to focus on candidate genes and study a more limited number of genes potentially involved in the pathogenesis of DKD based upon our known knowledge or hypothesis. In genetic and epigenetic studies of DKD, DNA samples used are commonly extracted from peripheral blood samples because they are clinically accessible. Dick et al. (2014) have comparatively analyzed DNA methylation changes related to BMI by using both approaches of whole-blood DNA methylation profiling and adipose tissue specific methylation measurement. Data suggests that analysis of blood DNA methylation is worthwhile because the results can reflect the DNA methylation changes in relevant tissues for a particular phenotype. Nevertheless, there is still limited information concerning the correlation between whole blood DNA methylation profiles and kidney tissue specific DNA methylation changes in part due to the heterogeneity of cell types within the kidney. To improve the tissue specific DNA methylation analysis of kidney diseases, including DKD, it is necessary to construct biobanks of renal biopsies. Karolinska Institutet has established a biobank in KaroKidney with more than 750 renal biopsies[5]. The advantages and limitations of these two approaches, as well as the clinical materials and experimental

[5]http://karokidney.org

design used in genetic and epigenetic studies of DKD are summarized in **Table 1**.

RECENT DATA FROM GENETIC STUDIES IN DIABETIC KIDNEY DISEASE

Considerable amounts of data from genetic studies in DKD have accumulated. A list of the genes that are reported to be associated with susceptibility or resistance to DKD are summarized in **Table 2**. The genes are listed in alphabetical order. Surprisingly, there are more than 150 genes. Most of them have been identified by genetic association studies employing candidate gene approaches over the past 20 years. Furthermore, a number of GWAS in DKD have been published in the last 10 years. By using GWAS approaches, approximately 33 genes have been found to be associated with the DKD, i.e., *ABCG2, AFF3, AGER, APOL1, AUH, CARS, CERS2, CDCA7/SP3, CHN2, CNDP1, ELMO1, ERBB4, FRMD3, GCKR, GLRA3, KNG1, LIMK2, MMP9, NMUR2, MSRB3/HMGA2, MYH9, PVT1, RAET1L, RGMA/MCTP2, RPS12, SASH1, SCAF8/CNKSR3, SHROOM3, SLC12A3, SORBS1, TMPO, UMOD,* and *ZMIZ1* (Hanson et al., 2007; Sandholm et al., 2012, 2014; Maeda et al., 2013; Thameem et al., 2013; Bailey et al., 2014; Palmer et al., 2014; Guan et al., 2016; Teumer et al., 2016; Lim et al., 2017; Roden, 2017; Charmet et al., 2018; van Zuydam et al., 2018). However, most of these genes (~80%) reportedly associated with DKD still need to be confirmed by further replication studies and detailed analysis of their functional role in DKD in experimental models. Polymorphisms in these candidate

TABLE 1 | Clinical material, research approaches and experimental designs used in genetic and epigenetic studies of diabetic kidney disease.

	Study	Advantage	Disadvantage
Clinical material	Blood or saliva	Clinical accessible	Possible bias from mixed cell types
	Kidney tissues	Gene specific methylation and expression can be analyzed	Difficult to access
	Renal cell lines	Intervention and mechanism study	*In vitro* experiment
Research approach	Candidate gene DNA variation or methylation analysis	Study of candidate genes with potential biological functions	Less information on the studied genes
	Global genomic DNA variation or methylation analyses	General information of DNA polymorphisms and methylation in genome wide scale	Analysis of repeated sequence alteration and methylation changes Lack of gene specific information
	Genome or epigenome-wide association studies	Numerous SNP, CNV or CpG sites methylation information in genome wide scale	Higher cost Strict validation is needed
Experimental design	Case-control study	Many cohorts exist	Difficult to control genetic and environmental confounders
	Twin study	Control for genetics	Few large cohorts
	Family study	Study of potential inheritance	Few large cohorts
	Longitudinal study	Determine causality	Time consuming

CNV, copy-number variation; CpG sites, the regions of DNA where a cytosine nucleotide is followed by a guanine nucleotide in the linear sequence of bases along its 5′ → 3′ direction; SNP, single-nucleotide polymorphism.

genes association with DKD studies are listed in **Table 2A**, while their potential biological relevance and genetic effects in DKD are briefly described. Of them, 34 genes are originally predicted by GWAS and the statistical association with DKD summarized in **Table 2B**.

The *CNDP1* (carnosine dipeptidase 1) gene is located in chromosome 18q22.3 and contains 5-leucine (CTG) trinucleotide repeat length polymorphism (D18S880) in the coding region (Wanic et al., 2008). This trinucleotide repeat polymorphism is found to have gender specificity and to confer the susceptibility for DKD and ESRD in T2D (Albrecht et al., 2017b). Furthermore, serum carnosinase (CN-1) activity is negatively correlated with time on hemodialysis (Peters et al., 2016). In addition, several SNPs in this gene are also associated with DKD and ESRD (Janssen et al., 2005; Freedman et al., 2007b; McDonough et al., 2009; Alkhalaf et al., 2010; Mooyaart et al., 2010; Ahluwalia et al., 2011b; Chakkera et al., 2011; Kurashige et al., 2013). Interestingly, an experimental study in BTBR ob/ob mice has demonstrated that treatment with carnosine as the target of CNDP1 improves glucose metabolism and albuminuria, suggesting that carnosine may be a novel therapeutic strategy to treat patients with DKD (Albrecht et al., 2017a).

The *ELMO1* (engulfment and cell motility 1) gene is located on chromosome p14.1 and encodes a member of the engulfment and cell motility protein family. The protein interacts with dedicator of cytokinesis proteins and subsequently promotes phagocytosis and cell migration. Increased expression of *ELMO1* and dedicator of cytokinesis 1 may promote glioma cell invasion (Patel et al., 2010). Furthermore, several SNPs in this gene are found to be associated with DKD in both T1D and T2D (Shimazaki et al., 2005, 2006; Craig et al., 2009; Leak et al., 2009; Pezzolesi et al., 2009a; Hanson et al., 2010; Wu et al., 2013; Alberto Ramirez-Garcia et al., 2015; Bodhini et al., 2016; Hathaway et al., 2016; Mehrbzadeh et al., 2016; Sharma et al., 2016). The variants associated with DKD, however, are different

in the several populations studied, suggesting the presence of allelic heterogeneity probably resulting from the diverse ancestral genetic backgrounds of the different racial groups.

The *FRMD3* (FERM domain containing 3) gene is located in chromosome 9q21.32. The *FRMD3* gene is expressed in adult brain, fetal skeletal muscle, thymus, ovaries, and podocytes (Ni et al., 2003). Pezzolesi et al. (2009b) have demonstrated that *FRMD3* expression in kidneys of a DKD mouse model is decreased as compared with non-diabetic mice. Genetic polymorphisms in the *FRMD3* gene are associated with DKD and ESRD in T1D and T2D (Freedman et al., 2011; Al-Waheeb et al., 2016). Furthermore, the members of the bone morphogenetic protein (BMP) interact with *FRMD3*, which implies that *FRMD3* may influence the risk of DKD through regulation of the BMP pathway (Martini et al., 2013; Palmer and Freedman, 2013).

The *MMP9* (matrix metallopeptidase 9) gene is located in chromosome 20q13.12. The MMP family members are involved in the breakdown of extracellular matrix (ECM) in physiological processes, such as tissue remodeling, reproduction and embryonic development, while *MMP9* is the ninth member in the family. *MMP9* may play an essential role in local proteolysis of the extracellular matrix and in leukocyte migration. Moreover, MMPs, including *MMP9*, are zinc-dependent endopeptidases and the major proteases in ECM degradation. There are common variants such as rs3918242 (-1562C/T) and microsatellites (CA)n in the promoter region and several SNPs rs481480, rs2032487, rs4281481, rs3752462 and rs3918242 are found to be associated with the susceptibility to DKD (Hirakawa et al., 2003; Nair et al., 2008; Ahluwalia et al., 2009; Freedman et al., 2011; Cooke et al., 2012; Zhang et al., 2015; Feng et al., 2016).

Both *UMOD* (uromodulin) and *SLC12A3* (solute carrier family 12 member 3) genes are located in the same chromosome but in short and long arms, respectively, i.e., 16p12.3 and 16q13. SLC12A3 is also known as thiazide-sensitive sodium-chloride cotransporter in kidney distal convoluted tubules,

TABLE 2A | Current data from genetic association studies in diabetic kidney disease by using candidate gene approach.

Gene symbol	Genomic DNA polymorphisms	Disease
ABCG2	rs2231142	T2D-uric acid
ACACB	rs2268388	T2D-DKD
ACE	rs4646994 (289bp Alu I/D), rs4343, rs1799752, rs1800764, rs12449782	T1D-DKD, T2D-DKD, T2D-ESRD
ADPOQ	rs266729, rs17300539, rs2241766, rs1063537, rs2241767, rs2082940	T1D-DKD, T2D-DKD
ADRB2	Arg16Gly, Gln27Glu	T2D-eGFR
AFF3	rs7583877	T1D-ESRD
AGER	rs2070600, rs2071288	T2D-DKD
AGT	rs5050, rs4762, Met235Thr	T2D-DKD
AGTR1	rs5186, +1166A/C, −106C/T, rs12695897	T1D-DKD, T2D-ESRD
AGTR2	+1675G/A, +1818A/T	T1D-DKD
AKR1B1	rs759853	T2D-DKD, T2D-ESRD
ALOX12	rs14309	T2D-DKD+CVD
APOE	e4 allele, e2/e3 alleles	T2D-DKD
APOL1	rs136161, rs713753, rs767855, Ser342Gly, Ile384Met	T2D-ESRD
AUH	rs773506	T2D-ESRD
BID	rs181390	T1D-ESRD
CALD1	rs3807337	T1D-DKD
CARS	rs452041, rs739401	T1D-DKD, T2D-DKD
CASR	rs3804594	T2D-DKD
CAT	rs1001179	T2D-ESRD
CERS2	rs267734, rs267738	T1D-DKD, T2D-DKD
CDH13	rs11646213, rs3865188	T1D-ESRD
CFH	rs379489	T2D-ESRD
CHN2	rs39059	T1D-DKD
CNDP1	(CTG)5, rs4892249, rs6566815, rs2346061, rs1295330, rs6566810, rs11151964, rs17817077	T2D-dialysis, T2D-DKD, T1D-ESRD, T2D-ESRD
CNDP2	rs7577, rs4892247	T2D-ESRD
CYP11B2	−344T/C	T2D-DKD
COQ5	rs1167726, rs614226, rs1167725	T1D-ESRD
COX6A1	rs12310837	T1D-ESRD
COX10	rs7213412	T1D-ESRD
CUBN	rs1801239	T1D-albuminuria, T2D-ESRD
CYBA	rs4673, rs9932581	T1D-ESRD, T2D-DKD
eNOS	−786C/T, +786T/C, +894G/T, Glu298Asp	T1D-DKD, T2D-DKD
ELMO1	rs741301, rs1345365, rs11769038, rs10951509, rs1882080, rs6462776, rs6462777	T1D-DKD, T1D-ESRD, T2D-DKD
ENPP1	rs1044498, rs7754586, rs1974201	T1D-DKD, T2D-DKD, T2D-ESRD
EPHX2	rs751141	T2D-DKD
EPO	rs1617640	T1D-ESRD, T2D-DKD
ERBB4	rs7588550	T1D-DKD
ESR1	rs12197043, rs11964281, rs1569788, rs9340969	T2D-DKD
FNDC5	rs16835198	T2D-DKD
FRMD3	rs1888747, rs10868025, rs942280, rs942278, rs942263, rs1535753, rs2378658, rs13288659	T1D-ESRD, T2D-DKD
GAS6	Intron 8, c.834+7G/A	T2D-DKD
GATC	rs2235222, rs7137953	T1D-ESRD
GCK	rs730947	T2D-ESRD
GCKR	rs1260326	T2D-eGFR
GFPT2	Ile147Val	T2D-DKD
GLRA3	rs1564939	T1D-AER
GPX1	rs3448	T1D-DKD

(Continued)

TABLE 2A | Continued

Gene symbol	Genomic DNA polymorphisms	Disease
GREM1	rs1129456	T1D-DKD
GSTP1	rs1695 (Ile105Val)	T2D-DKD, T2D-ESRD
H19-IGF2 cluster	rs2839698, rs10732516, rs201858505	T2D-DKD
HIF1α	rs11549465 (Pro582Ser)	T1D-DKD, T2D-DKD
HO1	−413T/A	T2D-DKD
HSP70	rs2763979, rs2227956	T2D-DKD
ICAM1	rs5498	T1D-DKD, T2D-DKD
IGFBP1	rs1065780, rs3828998, rs3793344, rs4619	T2D-DKD
IGF2BP2	rs4402960	T2D-DKD
IL1α	−889C/T	T2D-DKD
IL1β	rs16944, −511C/T	T2D-DKD
IL6	−634G/C, −174G/C, rs1800796, rs1524107, rs1800795, rs1800796	T2D-DKD
IL10	−819T/C, −592A/C, −1082A/G	T2D-DKD
IL18	rs360719	T2D-DKD
INSR	rs2059806	T2D-DKD
IRAK4	rs4251532	T2D-DKD
KCNQ1	I/D in intron 12, rs2237897	T2D-eGFR, T2D-DKD
KLRA1	rs2168749	T1D-ESRD
KNG1	+7965C/T	T1D-DKD
LIMK2	rs2106294	T2D-ESRD
LTA	Thr60Asn	T1D-DKD
LRP2	rs17848169	T2D-ESRD
MAPRE1P2	rs1670754	T1D-ESRD
MCF2L2	Leu359Ile	T1D-DKD
MGP	−138T/C	T2D-DKD
MME	rs3796268, rs3773885	T1D-DKD
MMP12	rs1277718, rs652438, Asn357Ser	T1D-DKD
MMP9	(CA)n in promoter, rs481480, rs2032487, rs4281481, rs3752462, rs3918242	T2D-ESRD, T2D-DKD
NMUR2	rs982715, rs4958531, rs4958532, rs4958535	T1D-DKD
MSC	rs9298190	T1D-ESRD
MT2A	rs28366003	T2D-DKD
MTHFR	rs1801133	T1D-DKD, T2D-DKD
MTOR	rs7212142	T2D-DKD
MyD88	rs6853	T2D-DKD
MYH9	rs5750250	T2D-ESRD
NCALD	rs1131863, +999T/A, +1298A/C, +1307A/G	T2D-DKD
near IRS2	rs1411766	T1D-DKD, T1D-ESRD, T2D-DKD
NOS2	rs1137933	T2D-DKD
NOS3	rs3918188, Glu298Asp, Gly894Thr	T1D-DKD, T2D-DKD
NQO1	rs1800566	T2D-DKD
NPHS1	rs35238405	T2D-ESRD
NPY	Leu7Pro	T1D-DKD
PACRG	rs2147653, rs1408705	T1D-ESRD
PAI1	4G/5G	T2D-DKD
PARK2	rs4897081	T2D-DKD
PARP1	C410T, G1672A, Val762Ala	T2D-DKD
PFKFB2	rs17258746, rs11120137	T2D-DKD
PLEKHH2	rs1368086, rs725238, rs11886047	T1D-DKD
PLXDC2	rs1571942, rs12219125	T1D-DKD
PON1	Leu55Met, Gln192Arg	T1D-DKD, T2D-ACR
PON2	rs12704795	T2D-DKD
PPARG	rs1805192, rs1801282	T1D-DKD, T2D-DKD

(Continued)

TABLE 2A | Continued

Gene symbol	Genomic DNA polymorphisms	Disease
PPARG2	Pro12Ala	T2D-eGFR, T2D-DKD
PPARGC1A	Gly482Ser	T2D-DKD
PRKAA2	rs2746342, rs10789038	T2D-DKD
PROX1	rs340841	T2D-DKD
PSMD9	rs1043307, rs14259, +460A/G, +437T/C, Glu197Gly	T2D-DKD
PRKCB1	−1504C/T, −546C/T, −348A/G, −278C/T, −238C/G	T1D-DKD, T2D-eGFR
PTX3	rs2305619, rs2120243	T2D-DKD
PVT1	rs2648875, rs2720709	T2D-ESRD
RAGE	−429T/C, −374T/A, +2184A/G	T1D-ESRD, T2D-DKD
RAET1L	rs1543547	T1D-DKD
RBP4	rs3758538, rs10882278, rs7094671, rs12766992	T2D-eGFR
REN	rs41317140	T2D-DKD
RREB1	rs9379084, rs41302867	T2D-ESRD
TOP1MT	rs7387720, 724037	T1D-ESRD
TXNRD2	rs17745445, rs17745433, rs5992495, rs5992493	T1D-ESRD
RPS12	rs7769051	T2D-ESRD
RTN1	rs1952034, rs12431381, rs12434215	T2D-ESRD
SASH1	rs6930576	T2D-ESRD
SCAF8/CNKSR3	rs12523833	T2D-DKD
SEMA6D/SLC24A5	rs12917114	T1D-ESRD
SERPINB7	rs1720843	T2D-DKD
SERPINE1	4G/5G polymorphism	T2D-DKD
SHROOM3	rs1739721	T2D-eGFR
SIK1	rs2838302	T1D-ESRD
SIRT1	rs4746720	T2D-DKD
SLC2A1	rs3820589, HaeIII polymorphism	T1D-DKD, T2D-DKD
SLC2A2	+16459C/T	T1D-DKD
SLC2A9	rs11722228, rs3775948	T2D-uric acid
SLC12A3	rs11643718	T2D-DKD, T2D-ESRD
SOD1	rs2234694	T1D-DKD
SOD2	Ala9Val, Val16Ala	T1D-DKD
SORBS1	rs1326934	T1D-DKD
SOX2	rs11915160	T1D-DKD
SPTLC2	rs176903	T1D-ESRD
SUMO4	rs237025	T2D-DKD
SUV39H2	rs17353856	T1D-DKD
TCF7L2	rs7903146	T2D-DKD
TGFβ1	rs1800470	T1D-DKD, T2D-DKD
THP	rs12444268	T1D-DKD
TMPO	rs4762495	T1D-ESRD
TNFα	rs1800629, rs1800470, rs1800469, rs1800630, rs1799964	T2D-DKD, T2D-ESRD
TRAF6	rs16928973	T2D-DKD
TRIB3	rs2295490	T2D-DKD
UMOD	rs12917707, rs13333226	T2D-DKD
VDR	RaqI variant	T2D-DKD
VEGF	−1499C/T, rs2010963	T1D-DKD, T2D-DKD
VEGFA	rs3025021	T1D-DKD
WNT4/ZBTB40	rs12137135	T1D-ESRD
ZMIZ1	rs1749824	T1D-ESRD
miRNA-146a	rs2910164	T1D-DKD, T2D-DKD
miRNA-125	rs12976445	T2D-DKD

which is important for electrolyte homeostasis. Mutations in this gene are characterized by hypokalemic alkalosis combined with hypomagnesemia, low urinary calcium, but increased renin activity. Tanaka et al. (2003) performed a GWAS in Japanese T2D subjects and reported that the *SLC12A3* Arg913Gln polymorphism was associated with reduced risk of DKD. Nishiyama et al. (2005) then conducted another 10-year longitudinal study in the same population. The results confirmed that the 913Gln allele of *SLC12A3* Arg913Gln polymorphism conferred a protective effect in DKD (Nishiyama et al., 2005). More recently, Abu Seman et al. (2014) performed a further genetic study of *SLC12A3* polymorphisms in a Malaysian population, including the meta-analysis of the association between the *SLC12A3* Arg913Gln polymorphism and DKD from all the previous studies. *SLC12A3* Arg913Gln polymorphism was

found to be associated with T2D ($P = 0.028$, OR = 0.772, 95% CI = 0.612–0.973) and DKD ($P = 0.038$, OR = 0.547, 95% CI = 0.308–0.973) in the Malaysian cohort. The meta-analysis confirmed the protective effects of the *SLC12A3* 913Gln allele in DKD (Z-value = -1.992, $P = 0.046$, OR = 0.792). In addition, the authors investigated the role of *slc12a3* expression in the progress of DKD with db/db mice and in kidney development with zebrafish embryos. With knockdown of zebrafish ortholog, slc12a3 led to structural abnormality of kidney pronephric distal duct at 1-cell stage. Slc12a3 mRNA and protein expression levels were upregulated in kidneys of db/db mice from 6, 12, and 26 weeks at the age. The authors thus concluded that *SLC12A3* is a susceptibility gene in DKD, while allele 913Gln but not allele Arg913 has a preventive effect in the disease (Abu Seman et al., 2014). This association of the *SLC12A3*

TABLE 2B | Current data from genetic association studies in diabetic kidney disease by using genome wide association approach.

Gene symbol	Genomic DNA polymorphisms	P-value	Disease	References
ABCG8	rs4148217	$P = 0.003$	T2D-ESRD	Nicolas et al., 2015
AFF3	rs7583877, rs7562121	$P = 1.2 \times 10(-8)$ and $< 1 \times 10(-6)$	T1D-ESRD	Sandholm et al., 2012, 2017
AGER	rs2070600, rs2071288	$P < 0.001$	T2D-DKD	Lim et al., 2017
AGTR1	rs12695897	$P = 0.032$	T2D-ESRD	Palmer et al., 2014
APOL1	rs136161, rs713753, rs767855	$P = 0.006–0.037$	T2D-ESRD	Palmer et al., 2014
AUH	rs7735506	$P = 2.57 \times 10(-4)$	T2D-ESRD	McDonough et al., 2011
BID	rs181390	$P = 0.006$	T1D-ESRD	Craig et al., 2009
CARS	rs452041, rs739401	$P = 3.1 \times 10(-6)$	T1D-DKD, T2D-DKD	Pezzolesi et al., 2009b
CERS2	rs267734, rs267738	$P = 0.0013$ and 0.0015	T1D-DKD, T2D-DKD	Shiffman et al., 2014
CDCA7-SP3	rs4972593	$P = 5 \times 10(-8)$	T1D-ESRD in women	Sandholm et al., 2013
CHN2	rs17157914	$P = 0.029$	T2D-ESRD	Palmer et al., 2014
CNDP1	rs4892249, rs6566815	$P = 0.0043$ and 0.0076	T2D-ESRD	Palmer et al., 2014
CNTNAP2	rs1989248	$P < 1 \times 10(-6)$	T1D-ESRD	Sandholm et al., 2017
ELMO1	rs741301 rs1345365, rs11769038, rs10951509, rs1882080, rs6462776, rs6462777	$P = 0.004$	T2D-DKD	Wu et al., 2013
ERBB4	rs7588550	$P = 2.1 \times 10(-7)$	T1D-DKD	Sandholm et al., 2012
FRMD3	rs942278, rs1888747, rs10868025, rs942280, rs942263, rs1535753, rs2378658, rs13288659	$P = 5.0 \times 10(-7)$	T1D-ESRD, T2D-ESRD	Pezzolesi et al., 2009a; Freedman et al., 2011
GABRR1	rs9942471	$P = 4.5 \times 10(-8)$	T2D-DKD	van Zuydam NR
GCKR	rs1260326	$P = 3.23 \times 10(-3)$	T2D-eGFR	Deshmukh et al., 2013
GLRA3	rs1564939	$P = 0.0013$	T1D-AER	Sandholm et al., 2018
KLKB	rs4253311	$P = 5.5 \times 10(-8)$	Plasma renin activity	Lieb et al., 2015
KNG1	rs5030062	$P = 0.001$	Plasma renin activity	Lieb et al., 2015
LIMK2	rs2106294, rs4820043	$P = 7.49E-04$ and 0.001	T2D-ESRD	McDonough et al., 2011
MMP9	rs481480, rs2032487, rs4281481	$P = 0.038$, 0.045 and 0.048 $P = 0.053$, 0.054 and 0.055	T2D-ESRD T2D-DKD	Freedman et al., 2009; Cooke et al., 2012
MYH9	rs5750250, rs92280	$P = 4.3 \times E(-4)$ $P = 3 \times 10(-7)$	T2D-ESRD	Freedman et al., 2011; McDonough et al., 2011
PTPN13	rs61277444	$P < 1 \times 10(-6)$	T1D-DKD	Sandholm et al., 2017
PVT1	rs2648875, rs2720709	$P = 1.8–2.1 \times (-7)$	T2D-ESRD	Hanson et al., 2007
RAET1L	rs1543547	$P = 1 \times 10(-5)$	T1D-DKD	McKnight et al., 2009
RGMA-MCTP2	rs12437854	$P = 2 \times 10(-9)$	T1D-ESRD	Sandholm et al., 2012
RPS12	rs9493454	$P = 8.79 \times 10(-4)$	T2D-ESRD	McDonough et al., 2011
SHROOM3	rs1739721	$P = 3.18 \times 10(-3)$	T2D-eGFR	Deshmukh et al., 2013

(Continued)

Gene symbol	Genomic DNA polymorphisms	P-value	Disease	References
SLC12A3	rs11643718	$P = 0.021$	T2D-DKD, T2D-ESRD	Tanaka et al., 2003
TMPO	rs4762495	$P = 0.0006$	T1D-ESRD	Craig et al., 2009
UMOD	rs12917707	$P = 8.84 \times 10(-4)$	T2D-eGFR	Deshmukh et al., 2013
ZMIZ1	rs1749824	$P = 8.1 \times 10(-5)$	T1D-ESRD	Craig et al., 2009

Data were extracted from more than 300 references in PubMed and most studies were carryout with genetic association study of candidate gene(s). CNVs, Copy Number Variants; DKD, Diabetic Kidney Disease; eGFR, estimated Glomerular Filtration Rate; T1D, Type 1 Diabetes Mellitus; T2D, Type 2 Diabetes Mellitus; ABCG, ATP Binding Cassette Subfamily G; ACACB, Acetyl-CoA Carboxylase Beta; ACE, Angiotensin I Converting Enzyme; ADPOQ, Adiponectin; ADRB2, Adrenoceptor Beta 2; AFF3, AF4/FMR2 Family Member 3; AGER, Advanced Glycosylation End-Product Specific Receptor; AGT, Angiotensinogen; AGTR, Angiotensin II Receptor; AKR1B1, Aldo-Keto Reductase Family 1 Member B; ALOX12, Arachidonate 12-Lipoxygenase, 12S Type; ApoE, Apolipoprotein E; APOL1, Apolipoprotein L1; AUH, AU RNA Binding Methylglutaconyl-CoA Hydratase; BID, BH3 Interacting Domain Death Agonist; CALD1, Caldesmon 1; CaSR, Calcium-Sensing Receptor; CARS, Cysteinyl-TRNA Synthetase; CAT, Catalase; CERS2, Ceramide Synthase 2; CDCA7, Cell Division Cycle Associated 7; CDH13, Cadherin 13; CHN2, Chimerin 2; CNDP, Carnosine Dipeptidase; COQ5, Coenzyme Q5, Methyltransferase; COX6A1, Cytochrome C Oxidase Subunit 6A1; COX10, COX10, Heme A:Farnesyltransferase Cytochrome C Oxidase Assembly Factor; CUBN, Cubilin; CYBA, Cytochrome B-245 Alpha Chain; CYP11B2, Cytochrome P450 Family 11 Subfamily B Member 2; ELMO1, Engulfment And Cell Motility 1; eNOS, Nitric Oxide Synthase; ENPP1, Ectonucleotide Pyrophosphatase/Phosphodiesterase 1; EPO, Erythropoietin; EPHX2, Epoxide Hydrolase 2; ERBB4, Erb-B2 Receptor Tyrosine Kinase 4; ESR1, Estrogen Receptor 1; FRMD3, FERM Domain Containing 3; FNDC5, Fibronectin Type III Domain Containing 5; GAS6, Growth Arrest Specific 6; GATC, Glutamyl-TRNA Amidotransferase Subunit C; GCK, Glucokinase; GCKR, Glucokinase Regulator; GFPT2, Glutamine-Fructose-6-Phosphate Transaminase 2; GLRA3, Glycine Receptor Alpha 3; GPX1, Glutathione Peroxidase 1; GREM1, Gremlin 1, DAN Family BMP Antagonist; GSTP1, Glutathione S-Transferase Pi 1; HIF1α, Hypoxia Inducible Factor 1 Subunit Alpha; H19, H19, Imprinted Maternally Expressed Transcript; HMGA2, High Mobility Group AT-Hook 2; HO1, Heme Oxygenase 1; HSP70, Heat Shock Protein 70; ICAM1, Intercellular Adhesion Molecule 1; IGF2, Insulin Like Growth Factor 2; IGFBP1, Insulin Like Growth Factor Binding Protein 1; IL, Interleukin; IRAK4, Interleukin 1 Receptor Associated Kinase 4; INSR, Insulin Receptor; IRS2, Insulin Receptor Substrate 2; KCNQ1, Potassium Voltage-Gated Channel Subfamily Q Member 1; KLRA1, Killer Cell Lectin Like Receptor A1; KNG1, Kininogen 1; LTA, Lymphotoxin Alpha; LIMK2, LIM Domain Kinase 2; MAPRE1P2, MAPRE1 Pseudogene 2; MCF2L2, MCF.2 Cell Line Derived Transforming Sequence-Like 2; MGP, Matrix Gla Protein; MME, Membrane Metalloendopeptidase; MMP, Matrix Metallopeptidase; MSC, Musculin; MTHFR, Methylenetetrahydrofolate Reductase; MT2A, Metallothionein 2A; MSRB3, Methionine Sulfoxide Reductase B3; MTOR, Mechanistic Target of Rapamycin Kinase; MyD88, Myeloid Differentiation Primary Response 88; MYH9, Myosin Heavy Chain 9; NCALD, Neurocalcin Delta; NOS, Nitric Oxide Synthase; NQO1, NAD(P)H Quinone Dehydrogenase 1; NPHS1, NPHS1, Nephrin; NPY, Neuropeptide Y; PACRG, Parkin Coregulated; PAI1, Plasminogen Activator Inhibitor 1; PARK2, Parkin RBR E3 Ubiquitin Protein Ligase; PFKFB2, 6-Phosphofructo-2-Kinase/Fructose-2,6-Biphosphatase 2; PLXDC2, Plexin Domain Containing 2; PLEKHH2, Pleckstrin Homology, MyTH4 and FERM Domain Containing H2; PON, Paraoxonase; PPARG, Peroxisome Proliferators-Activated Receptor Gamma; PPARGC1A, Peroxisome Proliferators-Activated Receptor Gamma Co-activator 1 alpha; PRKAA2, Protein Kinase AMP-Activated Catalytic Subunit Alpha 2; PROX1, Prospero Homeobox 1; PSMD9, Proteasome 26S Subunit, Non-ATPase 9; PRKCB1, Protein Kinase C Beta; PTX3, Pentraxin 3; PVT1, Pvt1 Oncogene; RAGE, Advanced Glycosylation End-Product Specific Receptor; RAET1L, Retinoic Acid Early Transcript 1L; RBP4, Retinol Binding Protein 4; REN, Renin; RGMA, Repulsive Guidance Molecule BMP Co-Receptor A; RREB1, Ras Responsive Element Binding Protein 1; TOP1MT, DNA Topoisomerase I Mitochondrial; RPS12, Ribosomal Protein S12; RTN1, Reticulon 1; SASH1, SAM And SH3 Domain Containing 1; SCAF8, SR-Related CTD Associated Factor 8; SEMA6D, Semaphorin 6D; SERPINB, Serpin Family; SHROOM3, Shroom Family Member 3; SIK1, Salt Inducible Kinase 1; SIRT1, Sirtuin 1; SLC2A, Solute Carrier Family 2; SLC12A3, Solute Carrier Family 12 Member 3; SOD, Superoxide Dismutase; SOX2, SRY-Box 2; SORBS1, Sorbin and SH3 Domain Containing 1; SP3, Sp3 Transcription Factor; SUMO4, Small Ubiquitin-Like Modifier 4; SUV39H2, Suppressor Of Variegation 3-9 Homolog 2; TCF7L2, Transcription Factor 7 Like 2; TGFβ1, Transforming Growth Factor Beta 1; TMPO, Thymopoietin; TNFα, Tumor Necrosis Factor alpha; THP, Tamm-Horsfall protein; TRAF6, TNF Receptor Associated Factor 6; TRIB3, Tribbles Pseudokinase 3; UMOD, Uromodulin; VEGF, Vascular Endothelial Growth Factor; VEGFA, Vascular Endothelial Growth Factor A; VDR, Vitamin D Receptor; WNT4, Wnt Family Member 4; ZBTB40, Zinc Finger and BTB Domain Containing 40; ZMIZ1, Zinc Finger MIZ-Type Containing 1.

Arg913Gln polymorphism with DKD has been very recently replicated in a Chinese population (Zhang et al., 2018). The UMOD gene encoded glycoprotein is synthesized exclusively in renal tubular cells and released into urine. Furthermore, UMOD may prevent urinary tract infection and inhibit formation of liquid containing supersaturated salts and subsequent formation of salt crystals. SNPs rs4293393 and rs1297707 in the UMOD gene are found to be associated with the susceptibility to DKD in T2D (Ahluwalia et al., 2011a; Prudente et al., 2017; van Zuydam et al., 2018).

The Human Genome Project has revealed that there are more than twenty thousand protein coding genes, and probably more than one million of RNA genes[6]. Genetic association studies of RNA gene polymorphisms with DKD are very limited. Up to date, only two SNPs, i.e., rs2910164 and rs12976445 in the genes for miRNA-146a and miRNA-125 have been found to be associated with DKD in T1D and T2D (Li et al., 2014; Kaidonis et al., 2016). Further investigation of RNA genetic variation conferring susceptibility to DKD needs to be undertaken.

---

[6]https://www.genecards.org/

CURRENT INFORMATION FROM EPIGENETIC STUDIES IN DIABETIC KIDNEY DISEASE

Similar to genetic association studies, epigenome-wide (EWAS) and candidate gene DNA methylation analyses have been used for epigenetic studies of DKD. Current information from epigenetic studies in DKD are represented in **Table 3**. An EWAS suggested that several genes, including SLC22A12, TRPM6, AQP9, HP, AGTX, and HYAL2, may have epigenetic effects in DKD (VanderJagt et al., 2015). Interestingly, SLC22A12 encodes for urate anion transporter 1 (URAT1), which is a kidney-specific urate transporter that transports urate across the apical membrane of the proximal tubule in kidneys. Loss-of-function SLC22A12 mutations are associated with renal hypouricaemia and affected persons can develop exercise-induced acute kidney injury and are at increased risk of developing urate stones (Lee et al., 2008). TRPM6 is a member of transient receptor potential superfamily of cation channels. This gene is widely expressed in the body, including kidneys along the nephron. The TRPM6

TABLE 3 | Current information from epigenetic studies in diabetic kidney disease.

Analysis	Gene symbol/ Target	Material and methods	Results	References
DNA methylation	*AKR1B1, TIMP-2*	T2DM-DKD	Hypomethylation of the genes are associated with albuminuria	Aldemir et al., 2017
	AKR1B1, IGF1, SLC12A3	T2DM-DKD and ESRD	Those genes implicated in DKD based upon the inter-individual epigenetic differences	Sapienza et al., 2011
	CTGF	T2DM-DKD Glomerular and mesangial cells	Hypomethylation through the decreased Dnmt3a binding in the gene promoter	Zhang et al., 2014
	IGFBP1	T1DM-DKD	Hypermethylation	Gu et al., 2014
	IL13RA1, IL15, EDG3, INHA	Hemodialyzed patients with DKD	Hypermethylation	Korabecna et al., 2013
	MTHFR	Diabetic complications, including DKD	Hypermethylation	Dos Santos Nunes et al., 2018
	MTHFR	T2DM-DKD	Demethylation	Yang et al., 2016
	MIOX	Human and mouse	Hypomethylation	Sharma et al., 2017
	PIK3C2B	Glomeruli in DKD	Up-regulated with methylation in glomeruli	Wang et al., 2018
	POLR2G, DDB1, ZNF230		Down-regulated with methylation in glomeruli	
	SLC30A8	T2DM-DKD	Hypermethylation	Seman et al., 2015
	SLC22A12, TRPM6, AQP9, HP, AGXT, HYAL2	Pre-diabetes and T2DM-DN	Hypermethylation found in 174 of 694 CpG sites	VanderJagt et al., 2015
	TAMM41, PMPCB, TSFM, AUH	T1DM-DKD	DNA methylation changes in these genes and influence with mitochondrial function	Swan et al., 2015
	UNC13B	T1DM-DKD	An intronic polymorphism rs13293564 in the gene is associated with DKD DNA methylation levels in 19 CpG sites are changed	Bell et al., 2010
	KLF4	Glomerular podocytes in human and mouse	DNA methylation levels in the promoters of genes encoding mesenchymal markers are increased	Hayashi et al., 2014
	aPC	Podocytes	aPC epigenetically controls p66(Shc) expression	Bock et al., 2013
	egfr	Cultured proximal tubule (normal rat kidney) cells	Inhibition of histone deacetylase in eGFR	Gilbert et al., 2011
	pxr	db/db mice and proximal tubular cells	Demethylation of DNA	Watanabe et al., 2018
	dnmt1	db/db mice	Hypomethylation	Zhang et al., 2017
	agt, abcc4, cyp4a10, glut5	db/m mouse	Hypomethylation	Marumo et al., 2015
	kif20b, cldn18, slco1a1		Hypomethylation	
	sglt2, pck1, g6pc, hnf4a	db/db mice	Demethylated in the proximal tubules	Marumo et al., 2015
	tgfb1, tet2	db/db mice	Decreased DNA methylation	Yang et al., 2018
Histone modification	*MTHFR*	T2D with DN	MTHFR regulates histone modification rs1801133 C677T in the gene is associated with DN	Zhou et al., 2015
	TGFB1	Glomerular and mesangial cells	TGF-β1 increases expression of the H3K4 methyltransferase SET7/9	Sun et al., 2010
	12/15-LO	Glomerular and mesangial cells	Up-regulation of histone lysine modifications	Yuan et al., 2016

(Continued)

TABLE 3 | Continued

Analysis	Gene symbol/ Target	Material and methods	Results	References
	h3k9/14ac, at1r	Glomerular and mesangial cells db/db mice	Losartan attenuated increased H3K9/14Ac at RAGE, PAI-1 and MCP-1 promoters, while the chromatin state at these genes are mediated in part by AT1R	Reddy et al., 2014
	h3k9, h3k23	db/db and C57BL/6 mice	Acetylation	Sayyed et al., 2010
	h3k4 in serine 10		Demethylation and phosphorylation	
	h3k9/14ac	db/+ mice	Losartan reversed permissive epigenetic changes in renal glomeruli	Reddy et al., 2014
	set7/9	db/db mice	Induced histone modification and mcp-1 expression	Chen et al., 2014
	xbp1	db/db mice	XBP1s-mediated of histone SET7/9 and consequently decreased MCP-1 expression	Chen et al., 2014
	opn/h3k27me3	Sur1-E1506K mice	Histone modification with opn	Cai et al., 2016
	txnip, h3k9ac, h3k4me3, h3k4me1, h3k27me3	Sur1-E1506K mice	Histone acetylation changes	De Marinis et al., 2016
	egfr	Cultured proximal tubule (normal rat kidney) cells	Inhibition of histone deacetylase in eGFR	Gilbert et al., 2011
	grp78/histone h4	Diabetic rats	Acetylation changes	Sun et al., 2016
	mfn2	Diabetic rats	Histone acetylation at collagen IV promoter	Mi et al., 2016
	h3 and hsp-27, map kinase p28	Sprague-Dawley rats	Dephosphorylation and acetylation of h3	Tikoo et al., 2008
Non-coding RNA dysregulation	miR-9-3, miR34a, miR-137	DKD and diabetic retinopathy	DNA methylation changes	Dos Santos Nunes et al., 2018
	miR-199b-5p, klotho	T2DM-DKD and STZ mice	Increased serum klotho levels are mediated by miR-199b-5p	Kang and Xu, 2016
	microRNA Let-7a-3	T2DM with DKD	DNA methylation levels in the promoter are increased by targeting UHRF1	Peng et al., 2015
	microRNA 1207-5P	Glomerular and mesangial cells	This PVT1-derived microRNA is upregulated by glucose and TGF-β1	Alvarez et al., 2013
	creb1, miR-10a	HFD/STZ mice	This microRNA regulate epigenetic modification by targeting creb1	Shan et al., 2016

DKD, Diabetic Kidney Disease; T1D, Type 1 Diabetes; T2DM, Type 2 Diabetes. The genes predicted by epigenome-wide association analysis are shown in bold, while genes from rodent studies are shown in lower case. AKR1B1, Aldo-Keto Reductase family 1, member B1; aPC, activated Protein C; AQP9, Aquaporin; AT1R, Angiotensin II Receptor type 1; AUH, AU RNA binding protein/enoyl-CoA hydratase; EGFR, epidermal growth factor receptor; CTGF, Connective Tissue Growth Factor; DDB1, Damage Specific DNA Binding Protein 1; EDG3, Endothelial Differentiation G-protein coupled receptor 3; DNMT1, DNA methyltransferase 1; HFD, High Fat Diet; IGF1, Insulin like Growth Factor 1; IGFBP1, Insulin-like Growth Factor Binding Protein 1; IL13RA1, interleukin 13 receptor subunit alpha 1; IL15, Interleukin 15; INHA, Inhibin alpha; KLF4, Kkruppel-like factor 4; MTHFR, Methylenetetrahydrofolate Reductase; MIOX, Myo-Inositol Oxygenase; PIK3C2B, Phosphatidylinositol-4-Phosphate 3-Kinase Catalytic Subunit Type 2 Beta; PMPCB, Peptidase, Mitochondrial Processing beta subunit; POLR2G, RNA Polymerase II Subunit G; SLC12A3, Solute Carrier family 12 member 3; SLC22A12, Solute Carrier family 22 member 12; SLC30A8, Solute Carrier family 30 member 8; TAMM41, TAM41 Mitochondrial translocator assembly and maintenance homolog; tet2, tet methylcytosine dioxygenase 2; TIMP2, TIMP metallopeptidase inhibitor 2; TRPM6, Transient Receptor Potential cation channel subfamily M member 6; TSFM, Ts translation elongation Factor, Mitochondrial; UHRF1, Ubiquitin like with PHD and Ring Finger domains 1; UNC13B, Unc-13 homolog B member 3; XBP1, X-Box Binding Protein 1; ZNF230, Zinc Finger Protein 230; 12/15-LO, 12/15-lipoxygenase; TGFB1, Transforming Growth Factor Beta 1.

channels are mainly located in the renal distal convoluted tubule, the site of active transcellular calcium and magnesium transport in the kidney (Felsenfeld et al., 2015). As described previously, several studies have implicated *UMOD* genetic polymorphisms in the susceptibility to DKD (Ahluwalia et al., 2011a; Prudente et al., 2017; van Zuydam et al., 2018). A recent study has demonstrated that UMOD regulates renal magnesium homeostasis through TRPM6 (Nie et al., 2018). Furthermore, analyses of the candidate genes such as *IGFBP1* and *MTHFR* have also provided evidence that

DNA methylation changes in these genes may be involved in the pathogenesis of DKD (Gu et al., 2013, 2014; Yang et al., 2016). Combining and analyzing data from genetic and epigenetic studies together may help understand some of the pathophysiology in DKD.

ncRNAs regulate gene expression at the post-transcriptional level and are involved in chromatin histone modification. Most of studies concerning histone modification and ncRNA dysregulation have been performed in diabetic animal models, while a few studies have been undertaken in subjects with

DKD (**Table 3**). Reddy et al. (2014) have analyzed histone modification profiles in genes associated with DKD pathology and the modified regulation of these genes following treatment with the angiotensin II type 1 receptor (AT1R) blocker losartan. The data indicate that losartan attenuates key parameters of DKD and modifies gene expression, and reverses some epigenetic changes in db/db mice. Losartan also attenuates increased H3K9/14Ac at RAGE, PAI-1, and MCP-1 promoters in mesangial cells cultured under diabetic conditions (Reddy et al., 2014). In a recent study of subjects of T2D and diabetic complications (including DKD) (Dos Santos Nunes et al., 2018) the methylation profiles of miR gene were compared and related to the presence of diabetic complications. Results indicated that miRs can modulate the expression of a variety of genes and methylation changes of *miR-9-3*, *miR-34a*, and *miR-137* were found to be associated with diabetic complications (Dos Santos Nunes et al., 2018). These two studies provide evidence suggesting that therapies targeting epigenetic regulators might be beneficial in the treatment of DKD.

SUMMARY AND PERSPECTIVES

Researchers have made major efforts to undertake well powered genetic and epigenetic studies in DKD to help understand its pathogenesis. The data, however, need to be confirmed by several strategies, for instance, replication studies could be performed with better selection of subjects with similar genetic background to limit influences from migration; intermarriage; cultural preferences; coupled with further investigation of DNA variation and methylation changes in RNA regulation genes and biological experiments to determine functional impact of these variants. Furthermore, new technologies for DNA and ncRNA sequencing analysis such as third generation sequencing and a PheWAS approach have recently been developed.

New Generation Sequencing

DNA sequencing analysis is used for determining the accurate order of nucleotides along chromosomes and genomes. Second-generation sequencing, commonly known as next-generation sequencing (NGS), has presently become popular in DNA sequencing analysis because NGS can enable a massively-paralleled approach capable of producing large numbers of reads at high coverages along the genome and therefore dramatically reduce the cost of DNA sequencing analysis (Treangen and Salzberg, 2011; Gu et al., 2018; Mone et al., 2018). Today, third-generation sequencing (often called as long-read sequencing) is a new generation sequencing method, which works by reading the nucleotide sequences at single molecule level in contrast to the first and second generations of DNA sequencing (van Dijk et al., 2018). Moreover, it is necessary to develop the molecular instruments for whole genome sequencing to make this new generation sequencing commercially available. The advanced sequencing technologies will improve genetic and epigenetic studies in DKD in the near future.

ncRNA Genetic and Epigenetic Studies

In the human genome, RNA genes are much more abundant than protein coding genes, while ncRNAs mainly include miRNAs and lncRNAs. Both forms of ncRNAs have been found to be involved in chromatin histone modifications, and subsequently can have epigenetic effects on the target genes. Therefore, identification of RNA genetic variation and investigation of biological alteration of these RNA genes should be included in research plans. Kato has very recently pointed out a hypothesis that transforming growth factor-β (TGF1β) may play an important role in early stage development of DKD, while some miRNAs and lncRNAs regulate the key molecules in the TGF1β pathway. These ncRNAs may be served as biomarkers for predicting the potential targets for prevention and treatment in DKD (Kato, 2018). Furthermore, Smyth et al. (2018) have compared Sanger sequencing and NGS to validate the five top ranked miRNAs that are predicted to be associated with DKD by EWAS. This study suggests that targeted NGS may offer a more cost-effective and sensitive approach and implied that the methylated miR-329-2, in which region SNP rs10132943 is located, and miR-429 where SNPs rs7521584 and rs112695918 exist, are associated with DKD (Smyth et al., 2018). Although these two studies are preliminary, they may be good examples to help direct further DKD research.

Phenome-Wide Association Study (PheWAS)

PheWAS is a new approach to analyze many phenotypes in comparison with a single genetic variant. This approach was originally described using electronic medical record (EMR) data from EMR-linked with a DNA biobank and also can be combined with GWAS and EWAS. Therefore, PheWAS has become a powerful tool to investigate the impact of genetic variation on drug response among many individuals and may expand our knowledge of new drug targets and effects (Pendergrass and Ritchie, 2015; Denny et al., 2016; Roden, 2017). Clearly, combined with GWAS and EWAS, PheWAS will provide us with the possibility to discover the associations with drug effects, including therapeutic response and side effect profiles in DKD (Hebbring, 2014).

Taken together, application of these advanced studies in DKD will be very useful not only for evaluating current data from genetic and epigenetic studies but also for generating new knowledge for dissecting the complexity of this disease.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

REFERENCES

Abu Seman, N., He, B., Ojala, J. R., Wan Mohamud, W. N., Östenson, C. G., Brismar, K., et al. (2014). Genetic and biological effects of sodium-chloride cotransporter (SLC12A3) in diabetic nephropathy. *Am. J. Nephrol.* 40, 408–416. doi: 10.1159/000368916

Ahluwalia, T. S., Khullar, M., Ahuja, M., Kohli, H. S., Bhansali, A., Mohan, V., et al. (2009). Common variants of inflammatory cytokine genes are associated with risk of nephropathy in type 2 diabetes among Asian Indians. *PLoS One* 4:e5168. doi: 10.1371/journal.pone.0005168

Ahluwalia, T. S., Lindholm, E., Groop, L., and Melander, O. (2011a). Uromodulin gene variant is associated with type 2 diabetic nephropathy. *J. Hypertens.* 29, 1731–1734. doi: 10.1097/HJH.0b013e328349de25

Ahluwalia, T. S., Lindholm, E., and Groop, L. C. (2011b). Common variants in CNDP1 and CNDP2, and risk of nephropathy in type 2 diabetes. *Diabetologia* 54, 2295–2302. doi: 10.1007/s00125-011-2178-5

Alberto Ramirez-Garcia, S., Charles-Niño, C., Mazariegos-Rubí, M., Rosalba Topete-González, L., Topete-González, R., Javier Flores-Alvarado, L., et al. (2015). Association of the ELMO1 gene (snp rs1345365) with development of type 2 diabetes mellitus in the Mexican mestizo population. *Invest. Clin.* 56, 341–355.

Albrecht, T., Schilperoort, M., Zhang, S., Braun, J. D., Qiu, J., Rodriguez, A., et al. (2017a). Carnosine attenuates the development of both type 2 diabetes and diabetic nephropathy in BTBR ob/ob mice. *Sci. Rep.* 7:44492. doi: 10.1038/srep44492

Albrecht, T., Zhang, S., Braun, J. D., Xia, L., Rodriquez, A., and Qiu, J. (2017b). The CNDP1 (CTG)(5) polymorphism is associated with biopsy-proven diabetic nephropathy, time on hemodialysis, and diabetes duration. *J. Diabetes Res.* 2017:9506730. doi: 10.1155/2017/9506730

Aldemir, O., Turgut, F., and Gokce, C. (2017). The association between methylation levels of targeted genes and albuminuria in patients with early diabetic kidney disease. *Ren. Fail.* 39, 597–601. doi: 10.1080/0886022X.2017.1358180

Alkhalaf, A., Bakker, S. J., Bilo, H. J., Gans, R. O., Navis, G. J., Postmus, D., et al. (2010). A polymorphism in the gene encoding carnosinase (CNDP1) as a predictor of mortality and progression from nephropathy to end-stage renal disease in type 1 diabetes mellitus. *Diabetologia* 53, 2562–2568. doi: 10.1007/s00125-010-1863-0

Allis, C. D., and Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. *Nat. Rev. Genet.* 17, 487–500. doi: 10.1038/nrg.2016.59

Alvarez, M. L., Khosroheidari, M., Eddy, E., and Kiefer, J. (2013). Role of microRNA 1207-5P and its host gene, the long non-coding RNA Pvt1, as mediators of extracellular matrix accumulation in the kidney: implications for diabetic nephropathy. *PLoS One* 8:e77468. doi: 10.1371/journal.pone.0077468

Al-Waheeb, S., Alwohhaib, M., Abdelghani, A., Al-Sharrah, S., Al-Shafey, E., Al-Sahow, A., et al. (2016). Evaluation of associations between single nucleotide polymorphisms in the FRMD3 and CARS genes and diabetic nephropathy in a Kuwaiti population. *Genet. Mol. Res.* 15:gmr7619. doi: 10.4238/gmr.15017619

Badal, S. S., and Danesh, F. R. (2014). New insights into molecular mechanisms of diabetic kidney disease. *Am. J. Kidney Dis.* 63(2 Suppl. 2), S63–S83. doi: 10.1053/j.ajkd.2013.10.047

Bailey, J. N. C., Palmer, N. D., Ng, M. C. Y., Bonomo, J. A., Hicks, P. J., Hester, J. M., et al. (2014). Analysis of coding variants identified from exome sequencing resources for association with diabetic and non-diabetic nephropathy in African Americans. *Hum. Genet.* 133, 769–779. doi: 10.1007/s00439-013-1415-z

Barrett, E. J., Liu, Z., Khamaisi, M., King, G. L., Klein, R., Klein, B. E. K., et al. (2017). Diabetic microvascular disease: an endocrine society scientific statement. *J. Clin. Endocrinol. Metab.* 102, 4343–4410. doi: 10.1210/jc.2017-01922

Bell, C. G., Teschendorff, A. E., Rakyan, V. K., Maxwell, A. P., Beck, S., and Savage, D. A. (2010). Genome-wide DNA methylation analysis for diabetic nephropathy in type 1 diabetes mellitus. *BMC Med. Genomics* 3:33. doi: 10.1186/1755-8794-3-33

Bock, F., Shahzad, K., Wang, H., Stoyanov, S., Wolter, J., Dong, W., et al. (2013). Activated protein C ameliorates diabetic nephropathy by epigenetically inhibiting the redox enzyme p66Shc. *Proc. Natl. Acad. Sci. U.S.A.* 110, 648–653. doi: 10.1073/pnas.1218667110

Bodhini, D., Chidambaram, M., Liju, S., Revathi, B., Laasya, D., Sathish, N., et al. (2016). Association of rs11643718 SLC12A3 and rs741301 ELMO1 variants

with diabetic nephropathy in south Indian population. *Ann. Hum. Genet.* 80, 336–341. doi: 10.1111/ahg.12174

Bouhairie, V. E., and McGill, J. B. (2016). Diabetic kidney disease. *Mol. Med.* 113, 390–394.

Cai, M., Bompada, P., Atac, D., Laakso, M., Groop, L., and De Marinis, Y. (2016). Epigenetic regulation of glucose-stimulated osteopontin (OPN) expression in diabetic kidney. *Biochem. Biophys. Res. Commun.* 469, 108–113. doi: 10.1016/j.bbrc.2015.11.079

Chakkera, H. A., Hanson, R. L., Kobes, S., Millis, M. P., Nelson, R. G., Knowler, W. C., et al. (2011). Association of variants in the carnosine peptidase 1 gene (CNDP1) with diabetic nephropathy in American Indians. *Mol. Genet. Metab.* 103, 185–190. doi: 10.1016/j.ymgme.2011.02.010

Charmet, R., Duffy, S., Keshavarzi, S., Gyorgy, B., Marre, M., Rossing, P., et al. (2018). Novel risk genes identified in a genome-wide association study for coronary artery disease in patients with type 1 diabetes. *Cardiovasc. Diabetol.* 17:61. doi: 10.1186/s12933-018-0705-0

Chen, J., Guo, Y., Zeng, W., Huang, L., Pang, Q., Nie, L., et al. (2014). ER stress triggers MCP-1 expression through SET7/9-induced histone methylation in the kidneys of db/db mice. *Am. J. Physiol. Renal Physiol.* 306, F916–F925. doi: 10.1152/ajprenal.00697.2012

Cooke, J. N., Bostrom, M. A., Hicks, P. J., Ng, M. C., Hellwege, J. N., Comeau, M. E., et al. (2012). Polymorphisms in MYH9 are associated with diabetic nephropathy in European Americans. *Nephrol. Dial. Transplant.* 27, 1505–1511. doi: 10.1093/ndt/gfr522

Craig, D. W., Millis, M. P., and DiStefano, J. K. (2009). Genome-wide SNP genotyping study using pooled DNA to identify candidate markers mediating susceptibility to end-stage renal disease attributed to Type 1 diabetes. *Diabet. Med.* 26, 1090–1098. doi: 10.1111/j.1464-5491.2009.02846.x

Cross, S. H., and Bird, A. P. (1995). CpG islands and genes. *Curr. Opin. Genet. Dev.* 5, 309–314.

De Marinis, Y., Cai, M., Bompada, P., Atac, D., Kotova, O., Johansson, M. E., et al. (2016). Epigenetic regulation of the thioredoxin-interacting protein (TXNIP) gene by hyperglycemia in kidney. *Kidney Int.* 89, 342–353. doi: 10.1016/j.kint.2015.12.018

Deaton, A. M., and Bird, A. (2011). CpG islands and the regulation of transcription. *Genes Dev.* 25, 1010–1022. doi: 10.1101/gad.2037511

Denny, J. C., Bastarache, L., and Roden, D. M. (2016). Phenome-wide association studies as a tool to advance precision medicine. *Annu. Rev. Genomics Hum. Genet.* 17, 353–373. doi: 10.1146/annurev-genom-090314-024956

Deshmukh, H. A., Palmer, C. N., Morris, A. D., and Colhoun, H. M. (2013). Investigation of known estimated glomerular filtration rate loci in patients with type 2 diabetes. *Diabet. Med.* 30, 1230–1235. doi: 10.1111/dme.12211

Dick, K. J., Nelson, C. P., Tsaprouni, L., Sandling, J. K., Aïssi, D., Wahl, S., et al. (2014). DNA methylation and body-mass index: a genome-wide analysis. *Lancet* 383, 1990–1998. doi: 10.1016/S0140-6736(13)62674-4

Do, C., Shearer, A., Suzuki, M., Terry, M. B., Gelernter, J., Greally, J. M., et al. (2017). Genetic-epigenetic interactions in cis: a major focus in the post-GWAS era. *Genome Biol.* 18:120. doi: 10.1186/s13059-017-1250-y

Dos Santos Nunes, M. K., Silva, A. S., Wanderley de Queiroga Evangelista, I., Modesto Filho, J., Alves Pegado Gomes, C. N., Ferreira do Nascimento, R. A., et al. (2018). Analysis of the DNA methylation profiles of miR-9-3, miR-34a, and miR-137 promoters in patients with diabetic retinopathy and nephropathy. *J. Diabetes Complications* 32, 593–601. doi: 10.1016/j.jdiacomp.2018.03.013

Dousdampanis, P., Trigka, K., and Mouzaki, A. (2016). Tregs and kidney: from diabetic nephropathy to renal transplantation. *World J. Transplant.* 6, 556–563. doi: 10.5500/wjt.v6.i3.556

Felsenfeld, A. J., Levine, B. S., and Rodriguez, M. (2015). Pathophysiology of calcium, phosphorus, and magnesium dysregulation in chronic kidney disease. *Semin. Dial.* 28, 564–577. doi: 10.1111/sdi.12411

Feng, S., Ye, G., Bai, S., Wei, H., Liao, X., and Li, L. (2016). Matrix metalloproteinase-9 -1562C/T gene polymorphism is associated with diabetic nephropathy. *Biomed Res. Int.* 2016:1627143. doi: 10.1155/2016/1627143

Florez, J. C. (2016). Genetics of diabetic kidney disease. *Semin. Nephrol.* 36, 474–480.

Freedman, B. I., Bostrom, M., Daeihagh, P., and Bowden, D. W. (2007a). Genetic factors in diabetic nephropathy. *Clin. J. Am. Soc. Nephrol.* 2, 1306–1316.

Freedman, B. I., Hicks, P. J., Bostrom, M. A., Comeau, M. E., Divers, J., Bleyer, A. J., et al. (2009). Non-muscle myosin heavy chain 9 gene *MYH9* associations in

African Americans with clinically diagnosed type 2 diabetes mellitus-associated ESRD. *Nephrol. Dial. Transplant.* 24, 3366–3371. doi: 10.1093/ndt/gfp316

Freedman, B. I., Hicks, P. J., Sale, M. M., Pierson, E. D., Langefeld, C. D., Rich, S. S., et al. (2007b). A leucine repeat in the carnosinase gene CNDP1 is associated with diabetic end-stage renal disease in European Americans. *Nephrol. Dial. Transplant.* 22, 1131–1135. doi: 10.1093/ndt/gfl717

Freedman, B. I., Langefeld, C. D., Lu, L., Divers, J., Comeau, M. E., Kopp, J. B., et al. (2011). Differential effects of MYH9 and APOL1 risk variants on FRMD3 association with diabetic ESRD in African Americans. *PLoS Genet.* 7:e1002150. doi: 10.1371/journal.pgen.1002150

Gilbert, R. E., Huang, Q., Thai, K., Advani, S. L., Lee, K., Yuen, D. A., et al. (2011). Histone deacetylase inhibition attenuates diabetes-associated kidney growth: potential role for epigenetic modification of the epidermal growth factor receptor. *Kidney Int.* 79, 1312–1321. doi: 10.1038/ki.2011.39

Gnudi, L., Coward, R. J. M., and Long, D. A. (2016). Diabetic nephropathy: perspective on novel molecular mechanisms. *Trends Endocrinol. Metab.* 27, 820–830. doi: 10.1016/j.tem.2016.07.002

Gu, H. F., and Brismar, K. (2012). Genetic association studies in diabetic nephropathy. *Curr. Diabetes Rev.* 8, 336–344. doi: 10.2174/157339912802083522

Gu, T., Falhammar, H., Gu, H. F., and Brismar, K. (2014). Epigenetic analyses of the insulin-like growth factor binding protein 1 gene in type 1 diabetes and diabetic nephropathy. *Clin. Epigenetics* 6:10. doi: 10.1186/1868-7083-6-10

Gu, T., Gu, H. F., Hilding, A., Sjöholm, L. K., Ostenson, C. G., Ekström, T. J., et al. (2013). Increased DNA methylation levels of the insulin-like growth factor binding protein 1 gene are associated with type 2 diabetes in Swedish men. *Clin. Epigenetics* 5:21. doi: 10.1186/1868-7083-5-21

Gu, W., Miller, S., and Chiu, C. Y. (2018). Clinical metagenomic next-generation sequencing for pathogen detection. *Annu. Rev. Pathol.* [Epub ahead of print].

Guan, M., Ma, J., Keaton, J. M., Dimitrov, L., Mudgal, P., Stromberg, M., et al. (2016). Association of kidney structure-related gene variants with type 2 diabetes-attributed end-stage kidney disease in African Americans. *Hum. Genet.* 135, 1251–1262. doi: 10.1007/s00439-016-1714-2

Hanson, R. L., Craig, D. W., Millis, M. P., Yeatts, K. A., Kobes, S., Pearson, J. V., et al. (2007). Identification of PVT1 as a candidate gene for end-stage renal disease in type 2 diabetes using a pooling-based genome-wide single nucleotide polymorphism association study. *Diabetes* 56, 975–983. doi: 10.2337/db06-1072

Hanson, R. L., Millis, M. P., Young, N. J., Kobes, S., Nelson, R. G., Knowler, W. C., et al. (2010). ELMO1 variants and susceptibility to diabetic nephropathy in American Indians. *Mol. Genet. Metab.* 101, 383–390. doi: 10.1016/j.ymgme.2010.08.014

Harjutsalo, V., and Groop, P. H. (2014). Epidemiology and risk factors for diabetic kidney disease. *Adv. Chronic Kidney Dis.* 21, 260–266. doi: 10.1053/j.ackd.2014.03.009

Hathaway, C. K., Chang, A. S., Grant, R., Kim, H. S., Madden, V. J., Bagnell, C. R. Jr., et al. (2016). High Elmo1 expression aggravates and low Elmo1 expression prevents diabetic nephropathy. *Proc. Natl. Acad. Sci. U.S.A.* 113, 2218–2222. doi: 10.1073/pnas.1600511113

Hayashi, K., Sasamura, H., Nakamura, M., Azegami, T., Oguchi, H., Sakamaki, Y., et al. (2014). KLF4-dependent epigenetic remodeling modulates podocyte phenotypes and attenuates proteinuria. *J. Clin. Invest.* 124, 2523–2537. doi: 10.1172/JCI69557

Hebbring, S. J. (2014). The challenges, advantages and future of phenome-wide association studies. *Immunology* 141, 157–165. doi: 10.1111/imm.12195

Hirakawa, S., Lange, E. M., Colicigno, C. J., Freedman, B. I., Rich, S. S., and Bowden, D. W. (2003). Evaluation of genetic variation and association in the matrix metalloproteinase 9 (MMP9) gene in ESRD patients. *Am. J. Kidney Dis.* 42, 133–142. doi: 10.1016/s0272-6386(03)00416-5

Holoch, D., and Moazed, D. (2015). RNA-mediated epigenetic regulation of gene expression. *Nat. Rev. Genet.* 16, 71–84. doi: 10.1038/nrg3863

Janssen, B., Hohenadel, D., Brinkkoetter, P., Peters, V., Rind, N., Fischer, C., et al. (2005). Carnosine as a protective factor in diabetic nephropathy: association with a leucine repeat of the carnosinase gene CNDP1. *Diabetes* 54, 2320–2327. doi: 10.2337/diabetes.54.8.2320

Jones, P. A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* 13, 484–492. doi: 10.1038/nrg3230

Kaidonis, G., Gillies, M. C., Abhary, S., Liu, E., Essex, R. W., Chang, J. H., et al. (2016). A single-nucleotide polymorphism in the MicroRNA-146a gene is associated with diabetic nephropathy and sight-threatening diabetic retinopathy in Caucasian patients. *Acta Diabetol.* 53, 643–650. doi: 10.1007/s00592-016-0850-4

Kang, W. L., and Xu, G. S. (2016). Atrasentan increased the expression of klotho by mediating miR-199b-5p and prevented renal tubular injury in diabetic nephropathy. *Sci. Rep.* 6:19979. doi: 10.1038/srep19979

Kato, M. (2018). Noncoding RNAs as therapeutic targets in early stage diabetic kidney disease. *Kidney Res. Clin. Pract.* 37, 197–209. doi: 10.23876/j.krcp.2018.37.3.197

Kato, M., and Natarajan, R. (2014). Diabetic nephropathy—emerging epigenetic mechanisms. *Nat. Rev. Nephrol.* 10, 517–530. doi: 10.1038/nrneph.2014.116

Keating, S. T., van Diepen, J. A., Riksen, N. P., and El-Osta, A. (2018). Epigenetics in diabetic nephropathy, immunity and metabolism. *Diabetologia* 61, 6–20. doi: 10.1007/s00125-017-4490-1

Korabecna, M., Pazourkova, E., Horinek, A., Mokrejsova, M., and Tesar, V. (2013). Methylation status of immune response genes promoters in cell-free DNA differs in hemodialyzed patients with diabetic nephropathy according to the intensity of anemia therapy. *Blood Purif.* 36, 280–286. doi: 10.1159/000356094

Kurashige, M., Imamura, M., Araki, S., Suzuki, D., Babazono, T., Uzu, T. (2013). The influence of a single nucleotide polymorphism within CNDP1 on susceptibility to diabetic nephropathy in Japanese women with type 2 diabetes. *PLoS One* 8:e54064. doi: 10.1371/journal.pone.0054064

Lappalainen, T., and Greally, J. M. (2017). Associating cellular epigenetic models with human phenotypes. *Nat. Rev. Genet.* 18, 441–451. doi: 10.1038/nrg.2017.32

Leak, T. S., Perlegas, P. S., Smith, S. G., Keene, K. L., Hicks, P. J., Langefeld, C. D., et al. (2009). Variants in intron 13 of the ELMO1 gene are associated with diabetic nephropathy in African Americans. *Ann. Hum. Genet.* 73, 152–159. doi: 10.1111/j.1469-1809.2008.00494.x

Lee, J. H., Choi, H. J., Lee, B. H., Kang, H. K., Chin, H. J., Yoon, H. J., et al. (2008). Prevalence of hypouricaemia and SLC22A12 mutations in healthy Korean subjects. *Nephrology* 13, 661–666. doi: 10.1111/j.1440-1797.2008.01029.x

Li, S. Y., Huang, P. H., Yang, A. H., Tarng, D. C., Yang, W. C., Lin, C. C., et al. (2014). Matrix metalloproteinase-9 deficiency attenuates diabetic nephropathy by modulation of podocyte functions and dedifferentiation. *Kidney Int.* 86, 358–369. doi: 10.1038/ki.2014.67

Lieb, W., Chen, M. H., Teumer, A., de Boer, R. A., Lin, H., and Fox, E. R. (2015). EchoGen consortium. Genome-wide meta-analyses of plasma renin activity and concentration reveal association with the kininogen 1 and prekallikrein genes. *Circ. Cardiovasc. Genet.* 8, 131–140. doi: 10.1093/hmg/ddr092

Lim, S. C., Dorajoo, R., Zhang, X., Wang, L., Ang, S. F., Tan, C. S. H., et al. (2017). Genetic variants in the receptor for advanced glycation end products (RAGE) gene were associated with circulating soluble RAGE level but not with renal function among Asians with type 2 diabetes: a genome-wide association study. *Nephrol. Dial. Transplant.* 32, 1697–1704. doi: 10.1093/ndt/gfw263

Maeda, S., Imamura, M., Kurashige, M., Araki, S., Suzuki, D., Babazono, T., et al. (2013). Replication study for the association of 3 SNP loci identified in a genome-wide association study for diabetic nephropathy in European type 1 diabetes with diabetic nephropathy in Japanese patients with type 2 diabetes. *Clin. Exp. Nephrol.* 17, 866–871. doi: 10.1007/s10157-013-0797-5

Martini, S., Nair, V., Patel, S. R., Eichinger, F., Nelson, R. G., Weil, E. J., et al. (2013). From single nucleotide polymorphism to transcriptional mechanism: a model for FRMD3 in diabetic nephropathy. *Diabetes* 62, 2605–2612. doi: 10.2337/db12-1416

Marumo, T., Yagi, S., Kawarazaki, W., Nishimoto, M., Ayuzawa, N., Watanabe, A., et al. (2015). Diabetes induces aberrant DNA methylation in the proximal tubules of the kidney. *J. Am. Soc. Nephrol.* 26, 2388–2397. doi: 10.1681/ASN.2014070665

McDonough, C. W., Hicks, P. J., Lu, L., Langefeld, C. D., Freedman, B. I., and Bowden, D. W. (2009). The influence of carnosinase gene polymorphisms on diabetic nephropathy risk in African-Americans. *Hum. Genet.* 126, 265–275. doi: 10.1007/s00439-009-0667-0

McDonough, C. W., Palmer, N. D., Hicks, P. J., Roh, B. H., An, S. S., Cooke, J. N., et al. (2011). A genome-wide association study for diabetic nephropathy genes in African Americans. *Kidney Int.* 79, 563–572. doi: 10.1038/ki.2010.467

McKnight, A. J., Currie, D., Patterson, C. C., Maxwell, A. P., Fogarty, D. G., and Warren, 3/UK GoKinD Study Group. (2009). Targeted genome-wide investigation identifies novel SNPs associated with diabetic nephropathy. *Hugo J.* 3, 77–82. doi: 10.1007/s11568-010-9133-2

Mehrabzadeh, M., Pasalar, P., Karimi, M., Abdollahi, M., Daneshpour, M., Asadolahpour, E., et al. (2016). Association between ELMO1 gene polymorphisms and diabetic nephropathy in an Iranian population. *J. Diabetes Metab. Disord.* 15:43.

Mi, X., Tang, W., Chen, X., Liu, F., and Tang, X. (2016). Mitofusin 2 attenuates the histone acetylation at collagen IV promoter in diabetic nephropathy. *J. Mol. Endocrinol.* 57, 233–249. doi: 10.1530/JME-16-0031

Mone, F., Quinlan-Jones, E., and Kilby, M. D. (2018). Clinical utility of exome sequencing in the prenatal diagnosis of congenital anomalies: a review. *Eur. J. Obstet. Gynecol. Reprod. Biol.* 231, 19–24. doi: 10.1016/j.ejogrb.2018.10.016

Mooyaart, A. L., Zutinic, A., Bakker, S. J., Grootendorst, D. C., Kleefstra, N., van Valkengoed, I. G., et al. (2010). Association between CNDP1 genotype and diabetic nephropathy is sex specific. *Diabetes* 59, 1555–1559. doi: 10.2337/db09-1377

Murea, M., Ma, L., and Freedman, B. I. (2012). Genetic and environmental factors associated with type 2 diabetes and diabetic vascular complications. *Rev. Diabet. Stud.* 9, 6–22. doi: 10.1900/RDS.2012.9.6

Nair, S., Phillips, A. O., Norton, N., Spurlock, G., Williams, H. J., Craig, K. J., et al. (2008). Further evidence for the association of MMP9 with nephropathy in type 2 diabetes and application of DNA pooling technology to candidate gene screening. *J. Nephrol.* 21, 400–405.

Ni, X., Ji, C., Cao, G., Cheng, H., Guo, L., Gu, S., et al. (2003). Molecular cloning and characterization of the protein 4.1O gene, a novel member of the protein 4.1 family with focal expression in ovary. *J. Hum. Genet.* 48, 101–106. doi: 10.1007/s100380300015

Nicolas, A., Fatima, S., Lamri, A., Bellili-Muñoz, N., Halimi, J. M., Saulnier, P. J., et al. (2015). ABCG8 polymorphisms and renal disease in type 2 diabetic patients. *Metabolism* 64, 713–719. doi: 10.1016/j.metabol.2015.03.005

Nie, M., Bal, M. S., Liu, J., Yang, Z., Rivera, C., Wu, X. R., et al. (2018). Uromodulin regulates renal magnesium homeostasis through the ion channel transient receptor potential melastatin 6 (TRPM6). *J. Biol. Chem.* 293, 16488–16502. doi: 10.1074/jbc.RA118.003950

Nishiyama, K., Tanaka, Y., Nakajima, K., Mokubo, A., Atsumi, Y., Matsuoka, K., et al. (2005). Polymorphism of the solute carrier family 12 (sodium/chloride transporters) member 3, SLC12A3, gene at exon 23 (+78G/A: Arg913Gln) is associated with elevation of urinary albumin excretion in Japanese patients with type 2 diabetes: a 10-year longitudinal study. *Diabetologia* 48, 1335–1338. doi: 10.1007/s00125-005-1785-4

Palmer, N. D., and Freedman, B. I. (2013). Diabetic nephropathy: FRMD3 in diabetic nephropathy–guilt by association. *Nat. Rev. Nephrol.* 9, 313–314. doi: 10.1038/nrneph.2013.81

Palmer, N. D., Ng, M. C., Hicks, P. J., Mudgal, P., Langefeld, C. D., Freedman, B. I., et al. (2014). Evaluation of candidate nephropathy susceptibility genes in a genome-wide association study of African American diabetic kidney disease. *PLoS One* 9:e88273. doi: 10.1371/journal.pone.0088273

Papadopoulou-Marketou, N., Chrousos, G. P., and Kanaka-Gantenbein, C. (2017). Diabetic nephropathy in type 1 diabetes: a review of early natural history, pathogenesis, and diagnosis. *Diabetes Metab. Res. Rev.* 33:e2841. doi: 10.1002/dmrr.2841

Patel, M., Margaron, Y., Fradet, N., Yang, Q., Wilkes, B., Bouvier, M., et al. (2010). An evolutionarily conserved autoinhibitory molecular switch in ELMO proteins regulates Rac signaling. *Curr. Biol.* 20, 2021–2027. doi: 10.1016/j.cub.2010.10.028

Pendergrass, S. A., and Ritchie, M. D. (2015). Phenome-wide association studies: leveraging comprehensive phenotypic and genotypic data for discovery. *Curr. Genet. Med. Rep.* 3, 92–100. doi: 10.1007/s40142-015-0067-9

Peng, R., Liu, H., Peng, H., Zhou, J., Zha, H., Chen, X., et al. (2015). Promoter hypermethylation of let-7a-3 is relevant to its down-expression in diabetic nephropathy by targeting UHRF1. *Gene* 570, 57–63. doi: 10.1016/j.gene.2015.05.073

Peters, V., Kebbewar, M., Janssen, B., Hoffmann, G. F., Möller, K., and Wygoda, S. (2016). CNDP1 genotype and renal survival in pediatric nephropathies. *J. Pediatr. Endocrinol. Metab.* 29, 827–833. doi: 10.1515/jpem-2015-0262

Pezzolesi, M. G., Katavetin, P., Kure, M., Poznik, G. D., Skupien, J., Mychaleckyj, J. C., et al. (2009a). Confirmation of genetic associations at ELMO1 in the GoKinD collection supports its role as a susceptibility gene in diabetic nephropathy. *Diabetes* 58, 2698–2702. doi: 10.2337/db09-0641

Pezzolesi, M. G., Poznik, G. D., Mychaleckyj, J. C., Paterson, A. D., Barati, M. T., and Klein, J. B. (2009b). Genome-wide association scan for diabetic nephropathy susceptibility genes in type 1 diabetes. *Diabetes* 58, 1403–1410.

Prudente, S., Di Paola, R., Copetti, M., Lucchesi, D., Lamacchia, O., and Pezzilli, S. (2017). The rs12917707 polymorphism at the UMOD locus and glomerular filtration rate in individuals with type 2 diabetes: evidence of heterogeneity across two different European populations. *Nephrol. Dial. Transplant.* 32, 1718–1722. doi: 10.1093/ndt/gfw262

Rakyan, V. K., Down, T. A., Balding, D. J., and Beck, S. (2011). Epigenome-wide association studies for common human diseases. *Nat. Rev. Genet.* 12, 529–541. doi: 10.1038/nrg3000

Reddy, M. A., Sumanth, P., Lanting, L., Yuan, H., Wang, M., Mar, D., et al. (2014). Losartan reverses permissive epigenetic changes in renal glomeruli of diabetic db/db mice. *Kidney Int.* 85, 362–373. doi: 10.1038/ki.2013.387

Reidy, K., Kang, H. M., Hostetter, T., and Susztak, K. (2014). Molecular mechanisms of diabetic kidney disease. *J. Clin. Invest.* 124, 2333–2340. doi: 10.1172/JCI72271

Roden, D. M. (2017). Phenome-wide association studies: a new method for functional genomics in humans. *J. Physiol.* 595, 4109–4115. doi: 10.1113/JP273122

Sandholm, N., Forsblom, C., Mäkinen, V. P., McKnight, A. J., Osterholm, A. M., and He, B. (2014). Genome-wide association study of urinary albumin excretion rate in patients with type 1 diabetes. *Diabetologia* 57, 1143–1153. doi: 10.1007/s00125-014-3202-3

Sandholm, N., Haukka, J. K., Toppila, I., Valo, E., Harjutsalo, V., Forsblom, C., et al. (2018). Confirmation of GLRA3 as a susceptibility locus for albuminuria in Finnish patients with type 1 diabetes. *Sci. Rep.* 8:12408. doi: 10.1038/s41598-018-29211-1

Sandholm, N., McKnight, A. J., Salem, R. M., Brennan, E. P., Forsblom, C., and Harjutsalo, V. (2013). FinnDiane Study Group and the GENIE Consortium. Chromosome 2q31.1 associates with ESRD in women with type 1 diabetes. *J. Am. Soc. Nephrol.* 24, 1537–1543. doi: 10.1681/ASN.2012111122

Sandholm, N., Salem, R. M., McKnight, A. J., Brennan, E. P., Forsblom, C., and Isakova, T. (2012). New susceptibility loci associated with kidney disease in type 1 diabetes. *PLoS Genet.* 8:e1002921. doi: 10.1371/journal.pgen.1002921

Sandholm, N., Van Zuydam, N., Ahlqvist, E., Juliusdottir, T., Deshmukh, H. A., and Rayner, N. W. (2017). The genetic landscape of renal complications in type 1 diabetes. *J. Am. Soc. Nephrol.* 28, 557–574. doi: 10.1681/ASN.2016020231

Sapienza, C., Lee, J., Powell, J., Erinle, O., Yafai, F., Reichert, J., et al. (2011). DNA methylation profiling identifies epigenetic differences between diabetes patients with ESRD and diabetes patients without nephropathy. *Epigenetics* 6, 20–28. doi: 10.4161/epi.6.1.13362

Sayyed, S. G., Gaikwad, A. B., Lichtnekert, J., Kulkarni, O., Eulberg, D., Klussmann, S., et al. (2010). Progressive glomerulosclerosis in type 2 diabetes is associated with renal histone H3K9 and H3K23 acetylation, H3K4 dimethylation and phosphorylation at serine 10. *Nephrol. Dial. Transplant.* 25, 1811–1817. doi: 10.1093/ndt/gfp730

Seman, N. A., Mohamud, W. N., Östenson, C. G., Brismar, K., and Gu, H. F. (2015). Increased DNA methylation of the SLC30A8 gene promoter is associated with type 2 diabetes in a Malay population. *Clin. Epigenetics* 7:30. doi: 10.1186/s13148-015-0049-5

Shan, Q., Zheng, G., Zhu, A., Cao, L., Lu, J., Wu, D., et al. (2016). Epigenetic modification of miR-10a regulates renal damage by targeting CREB1 in type 2 diabetes mellitus. *Toxicol. Appl. Pharmacol.* 306, 134–143. doi: 10.1016/j.taap.2016.06.010

Sharma, I., Dutta, R. K., Singh, N. K., and Kanwar, Y. S. (2017). High glucose-induced hypomethylation promotes binding of Sp-1 to myo-inositol oxygenase: implication in the pathobiology of diabetic tubulopathy. *Am. J. Pathol.* 187, 724–739. doi: 10.1016/j.ajpath.2016.12.011

Sharma, K. R., Heckler, K., Stoll, S. J., Hillebrands, J. L., Kynast, K., Herpel, E., et al. (2016). ELMO1 protects renal structure and ultrafiltration in kidney

development and under diabetic conditions. *Sci. Rep.* 6:37172. doi: 10.1038/srep37172

Shiffman, D., Pare, G., Oberbauer, R., Louie, J. Z., Rowland, C. M., Devlin, J. J., et al. (2014). A gene variant in CERS2 is associated with rate of increase in albuminuria in patients with diabetes from ONTARGET and TRANSCEND. *PLoS One* 9:e106631. doi: 10.1371/journal.pone.0106631

Shimazaki, A., Kawamura, Y., Kanazawa, A., Sekine, A., Saito, S., Tsunoda, T., et al. (2005). Genetic variations in the gene encoding ELMO1 are associated with susceptibility to diabetic nephropathy. *Diabetes* 54, 1171–1178. doi: 10.2337/diabetes.54.4.1171

Shimazaki, A., Tanaka, Y., Shinosaki, T., Ikeda, M., Watada, H., Hirose, T., et al. (2006). ELMO1 increases expression of extracellular matrix proteins and inhibits cell adhesion to ECMs. *Kidney Int.* 70, 1769–1776. doi: 10.1038/sj.ki.5001939

Smyth, L. J., Maxwell, A. P., Benson, K. A., Kilner, J., McKay, G. J., and McKnight, A. J. (2018). Validation of differentially methylated microRNAs identified from an epigenome-wide association study; Sanger and next generation sequencing approaches. *BMC Res. Notes* 11:767. doi: 10.1186/s13104-018-3872-x

Sun, G., Reddy, M. A., Yuan, H., Lanting, L., Kato, M., and Natarajan, R. (2010). Epigenetic histone methylation modulates fibrotic gene expression. *J. Am. Soc. Nephrol.* 21, 2069–2080. doi: 10.1681/ASN.2010060633

Sun, X. Y., Qin, H. J., Zhang, Z., Xu, Y., Yang, X. C., Zhao, D. M., et al. (2016). Valproate attenuates diabetic nephropathy through inhibition of endoplasmic reticulum stress-induced apoptosis. *Mol. Med. Rep.* 13, 661–668. doi: 10.3892/mmr.2015.4580.

Swan, E. J., Maxwell, A. P., and McKnight, A. J. (2015). Distinct methylation patterns in genes that affect mitochondrial function are associated with kidney disease in blood-derived DNA from individuals with Type 1 diabetes. *Diabet. Med.* 32, 1110–1115. doi: 10.1111/dme.12775

Tanaka, N., Babazono, T., Saito, S., Sekine, A., Tsunoda, T., Haneda, M., et al. (2003). Association of solute carrier family 12 (sodium/chloride) member 3 with diabetic nephropathy, identified by genome-wide analyses of single nucleotide polymorphisms. *Diabetes* 52, 2848–2853. doi: 10.2337/diabetes.52.11.2848

Teumer, A., Tin, A., Sorice, R., Gorski, M., Yeo, N. C., and Chu, A. Y. (2016). Genome-wide association studies identify genetic loci associated with albuminuria in diabetes. *Diabetes* 65, 803–817. doi: 10.2337/db15-1313

Thameem, F., Igo, R. P. Jr., Freedman, B. I., Langefeld, C., Hanson, R. L., and Schelling, J. R. (2013). Family investigation of nephropathy and diabetes research group. A genome-wide search for linkage of estimated glomerular filtration rate (eGFR) in the family investigation of nephropathy and diabetes (FIND). *PLoS One* 8:e81888. doi: 10.1371/journal.pone.0081888

Thomas, M. C. (2016). Epigenetic mechanisms in diabetic kidney disease. *Curr. Diab. Rep.* 16:31. doi: 10.1007/s11892-016-0723-9

Thomas, M. C., Brownlee, M., Susztak, K., Sharma, K., Jandeleit-Dahm, K. A., Zoungas, S., et al. (2015). Diabetic kidney disease. *Nat. Rev. Dis. Primers* 1:15018. doi: 10.1038/nrdp.2015.18

Thomas, M. C., Groop, P.-H., and Tryggvason, K. (2012). Towards understanding the inherited susceptibility for nephropathy in diabetes. *Curr. Opin. Nephrol. Hypertens.* 21, 195–202. doi: 10.1097/MNH.0b013e328350313e

Tikoo, K., Meena, R. L., Kabra, D. G., and Gaikwad, A. B. (2008). Change in post-translational modifications of histone H3, heat-shock protein-27 and MAP kinase p38 expression by curcumin in streptozotocin-induced type I diabetic nephropathy. *Br. J. Pharmacol.* 153, 1225–1231. doi: 10.1038/sj.bjp.0707666

Treangen, T. J., and Salzberg, S. L. (2011). Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nat. Rev. Genet.* 13, 36–46. doi: 10.1038/nrg3117

van Dijk, E. L., Jaszczyszyn, Y., Naquin, D., and Thermes, C. (2018). The third revolution in sequencing technology. *Trends Genet.* 34, 666–681. doi: 10.1016/j.tig.2018.05.008

van Zuydam, N. R., Ahlqvist, E., Sandholm, N., Deshmukh, H., Rayner, N. W., and Abdalla, M. S. (2018). A genome-wide association study of diabetic kidney disease in subjects with type 2 diabetes. *Diabetes* 67, 1414–1427. doi: 10.2337/db17-0914

VanderJagt, T. A., Neugebauer, M. H., Morgan, M., Bowden, D. W., and Shah, V. O. (2015). Epigenetic profiles of pre-diabetes transitioning to type 2

diabetes and nephropathy. *World J. Diabetes* 6, 1113–1121. doi: 10.4239/wjd.v6.i9.1113

Villeneuve, L. M., and Natarajan, R. (2010). The role of epigenetics in the pathology of diabetic complications. *Am. J. Physiol. Renal Physiol.* 299, F14–F25. doi: 10.1152/ajprenal.00200.2010

Wang, T., Zhang, Y., Wang, N., Liu, Q., Wang, Z., Liu, B., et al. (2018). Synergistical action of the β2 adrenoceptor and fatty acid binding protein 2 polymorphisms on the loss of glomerular filtration rate in Chinese patients with type 2 diabetic nephropathy. *Int. Urol. Nephrol.* 50, 715–723. doi: 10.1007/s11255-018-1812-2

Wanic, K., Placha, G., Dunn, J., Smiles, A., Warram, J. H., and Krolewski, A. S. (2008). Exclusion of polymorphisms in carnosinase genes (CNDP1 and CNDP2) as a cause of diabetic nephropathy in type 1 diabetes: results of large case-control and follow-up studies. *Diabetes* 57, 2547–2551. doi: 10.2337/db07-1303

Watanabe, A., Marumo, T., Kawarazaki, W., Nishimoto, M., Ayuzawa, N., Ueda, K., et al. (2018). Aberrant DNA methylation of pregnane X receptor underlies metabolic gene alterations in the diabetic kidney. *Am. J. Physiol. Renal Physiol.* 314, F551–F560. doi: 10.1152/ajprenal.00390.2017

Wu, H. Y., Wang, Y., Chen, M., Zhang, X., Wang, D., Pan, Y., et al. (2013). Association of ELMO1 gene polymorphisms with diabetic nephropathy in Chinese population. *J. Endocrinol. Invest.* 36, 298–302. doi: 10.3275/8525

Yang, L., Zhang, Q., Wu, Q., Wei, Y., Yu, J., Mu, J., et al. (2018). Effect of TET2 on the pathogenesis of diabetic nephropathy through activation of transforming growth factor β1 expression via DNA demethylation. *Life Sci.* 207, 127–137. doi: 10.1016/j.lfs.2018.04.044

Yang, X. H., Cao, R. F., Yu, Y., Sui, M., Zhang, T., Xu, J. Y., et al. (2016). A study on the correlation between MTHFR promoter methylation and diabetic nephropathy. *Am. J. Transl. Res.* 8, 4960–4967.

Yuan, H., Reddy, M. A., Deshpande, S., Jia, Y., Park, J. T., Lanting, L. L., et al. (2016). Epigenetic histone modifications involved in profibrotic gene regulation by 12/15-lipoxygenase and its oxidized lipid products in diabetic nephropathy. *Antioxid. Redox Signal.* 24, 361–375. doi: 10.1089/ars.2015.6372

Zhang, H., Cai, X., Yi, B., Huang, J., Wang, J., and Sun, J. (2014). Correlation of CTGF gene promoter methylation with CTGF expression in type 2 diabetes mellitus with or without nephropathy. *Mol. Med. Rep.* 9, 2138–2144. doi: 10.3892/mmr.2014.2067

Zhang, L., Zhang, Q., Liu, S., Chen, Y., Li, R., Lin, T., et al. (2017). DNA methyltransferase 1 may be a therapy target for attenuating diabetic nephropathy and podocyte injury. *Kidney Int.* 92, 140–153. doi: 10.1016/j.kint.2017.01.010

Zhang, R., Zhuang, L., Li, M., Zhang, J., Zhao, W., Ge, X., et al. (2018). Arg913Gln of SLC12A3 gene promotes development and progression of end-stage renal disease in Chinese type 2 diabetes mellitus. *Mol. Cell. Biochem.* 437, 203–210. doi: 10.1007/s11010-017-3120-z

Zhang, Z., Wu, X., Cai, T., Gao, W., Zhou, X., Zhao, J., et al. (2015). Matrix metalloproteinase 9 gene promoter (rs3918242) mutation reduces the risk of diabetic microvascular complications. *Int. J. Environ. Res. Public Health* 12, 8023–8033. doi: 10.3390/ijerph120708023

Zhou, T. B., Drummen, G. P., Jiang, Z. P., and Li, H. Y. (2015). Methylenetetrahydrofolate reductase (MTHFR) C677T gene polymorphism and diabetic nephropathy susceptibility in patients with type 2 diabetes mellitus. *Ren. Fail.* 37, 1247–1259. doi: 10.3109/0886022X.2015.1064743

ACTB Variants Confer the Genetic Susceptibility to Diabetic Kidney Disease in a Han Chinese Population

Mengxia Li[1†], Ming Wu[2†], Yu Qin[2], Jinyi Zhou[2], Jian Su[3], Enchun Pan[3], Qin Zhang[3], Ning Zhang[4], Hongyan Sheng[4], Jiayi Dong[1], Ye Tong[1] and Chong Shen[1*]

[1] Department of Epidemiology, School of Public Health, Nanjing Medical University, Nanjing, China, [2] Department of Non-communicable Chronic Disease Control, Jiangsu Provincial Center for Disease Control and Prevention, Nanjing, China, [3] Department of Chronic Disease Prevention and Control, Huai'an City Center for Disease Control and Prevention, Huai'an, China, [4] Changshu County Center for Disease Control and Prevention, Suzhou, China

Beta-actin (ACTB) loss-of-function mutations result in a pleiotropic developmental disorder of kidney. The present study aims to explore whether the common variants at the *ACTB* gene contribute to diabetic kidney disease (DKD) susceptibility in patients with type 2 diabetes mellitus (T2DM). From the baseline population of 20,340 diabetic patients, 1,510 DKD cases and 1,510 age-matched T2DM controls were selected. All subjects were Han Chinese. Three tagging single nucleotide polymorphisms (SNPs), rs852423, rs852426, and rs2966449, at the *ACTB* gene were genotyped. Logistic regression was performed to estimate the association with DKD. SNPs, rs852426 and rs2966449, were significantly associated with DKD [additive model; odds ratio (OR), 1.217 and 1.151; $P = 0.001$ and 0.018, respectively]. The association of rs852426 with DKD still remained statistically significant after Bonferroni correction and particularly significant in the population older than 70 years rather than the 70 years or younger ($P = 0.047$ for heterogeneity test). Furthermore, the association of rs852426 with DKD was observed in populations of male and females without smoking, drinking, and with duration for T2DM 10–20 years. The association of rs2966449 with DKD was also found in the populations older than 70 years, male, not smoking, not drinking, and with duration for T2DM over 20 years. The estimated glomerular filtration rate (eGFR) levels of the individuals with TT or CC genotypes of rs2966449 were significantly lower than that of TC genotype in DKD cases ($P = 0.021$). The present study provides evidence that the *ACTB* variants, i.e., rs852426 and rs2966449, may confer the genetic susceptibility to DKD in a Han Chinese population.

Keywords: *ACTB* gene, single nucleotide polymorphisms, diabetic kidney disease, estimated glomerular filtration rate, type 2 diabetes mellitus

INTRODUCTION

According to the International Diabetes Federation (IDF) survey in 2017, approximately 451 million people aged 18 to 99 years worldwide are suffering from diabetes mellitus, and the number was expected to increase to 693 million by year 2045 (Cho et al., 2018). China Kadoorie Biobank (CKB) data from 170,287 participants showed that the estimated standardized prevalence of total diagnosed and undiagnosed diabetes in urban and rural China reached 10.9% in 2013 and that of

pre-diabetic patients was 35.7% (Wang et al., 2017). In reality, diabetes mellitus and its complications seriously deteriorate patients' quality of life (Trikkalinou et al., 2017).

Diabetic kidney disease (DKD), a well-established microvascular complications of diabetes mellitus, is also the most frequent primary cause of end-stage renal disease (ESRD) (Kdoqi, 2007). In 2014, the American Diabetes Association (ADA) and the National Kidney Foundation (NKF) defined DKD as a chronic kidney disease (CKD) caused by diabetes mellitus, which was characterized by glomerular basement membrane thickening, mesangial expansion, nodular sclerosis, and finally, advanced diabetic glomerulosclerosis (Tervaert et al., 2010). DKD accounted in more than 40% of ESRD (Remuzzi et al., 2002; Zelnick et al., 2017).

The all-age DKD mortality rates in China increased by 33.3% from 1990 to 2016 (Liu et al., 2018). Accordingly, the Global Burden of Disease (GBD) reported that deaths from DKD all over the world also significantly grew up with an increase percentage of 40.73% from 2007 to 2017 (Collaborators GBDCoD, 2018). Therefore, the growing prevalence and mortality due to DKD constitute a huge health and social economic burden and have become a serious public health problem in China and all over the world (Cooper, 2012).

DKD is a complex disease with multiple etiologies and pathogenesis, including genetic determinants, glucose/lipid metabolism disorders, glomerular hemodynamic changes, and abnormal expression of cytokines (Kong et al., 2013; Mauer et al., 2015; Kitada et al., 2016). Identification of genetic markers linked to DKD could provide deep insights into the underlying biological processes of the renal dysfunction. Previous studies suggested that variants in TGF-β/Smad classic pathway and Wnt/β-catenin signaling pathway were associated with DKD (Regele et al., 2015; Ying and Wu, 2017).

The *ACTB* gene encodes an abundant cytoskeletal housekeeping protein β-actin, which is ubiquitous in eukaryotic cells (Pollard and Cooper, 2009). Recently, a study suggests that a critically reduced amount of β-actin alters cell function and gene expression to the detriment of brain, heart, and kidney development (Cuvertino et al., 2017). *ACTB* loss-of-function mutations result in a pleiotropic developmental disorder, including unilateral renal agenesis, pelvic kidney, and kidney cysts, whereas there are few reports about the effect of the common variants at *ACTB* on the renal function. Although five genomewide association studies (GWAS) of DKD identified several susceptible loci (Iyengar et al., 2015; Lim et al., 2017; Gurung et al., 2018; van Zuydam et al., 2018; Ahluwalia et al., 2019; Guan et al., 2019), no near loci linked to *ACTB* were observed as well as the data from GWAS of the National Human Genome Research Institute (https://research.nhgri.nih.gov/).

This study aimed to evaluate the association between three tagging single-nucleotide polymorphisms (tagSNPs) in the *ACTB* gene and DKD in a Han Chinese cohort of T2DM patients. The findings would provide a better understanding of the genetic susceptibility of *ACTB* in DKD.

PATIENTS AND METHODS

Study Participants

Subjects in this analysis were selected from the project of "Comprehensive Research on the Prevention and Control of the Diabetes (CRPCD)," which has been previously described in detail (Miao et al., 2017; Shen et al., 2018). According to the criteria proposed by the ADA in 2010 (American Diabetes Association, 2010), 19,992 subjects with type 2 diabetes mellitus (T2DM) were defined from all 20,340 participants. In the current study, we constructed a case-control study in T2DM subjects with equal DKD cases and controls. Referring to the recommendations released by NKF in 2007 (Kdoqi, 2007), a total of 3,443 T2DM patients who met one of the following three criteria would be selected as DKD cases: 1) eGFR less than 60 ml·min^{-1}·1.73 m^{-2} (n = 2,193); 2) self-reported kidney disease (n = 810); 3) concomitant microvascular disease (mainly diabetic retinopathy) accompanied by CKD stage 2 (eGFR) less than 90 ml·min^{-1}·1.73 m^{-2}) (n = 1,663). Meanwhile, subjects with the following characteristics were excluded: 1) time interval between T2DM and DKD less than 2 years (n = 94); 2) non-DKD in mentioned self-reported kidney disease (n = 185); 3) diabetes mellitus duration less than 5 years (n = 1,281). Thus, 1,883 DKD cases could be included. We used propensity score matching (PSM) to match a non-DKD control, by age (2 years as an interval), for every case. Finally, 1,510 pairs were involved in this study. Informed consent was obtained from all participants and approved by the Research Ethics Committee of Jiangsu Provincial Center for Disease Control and Prevention and Nanjing Medical University.

Questionnaire Survey and Anthropometric Measurements

All staff were uniformly trained and qualified in a standardized manner. Individuals accepted the survey by a questionnaire addressing their demographic characteristics, smoking, drinking, and self-reported chronic disease history. Anthropometric measurements including height, weight, waistline, and hipline were measured with light clothing. Body mass index (BMI) was then calculated as weight (kg)/height squared (m^2).

Definition of Drinking and Smoking

Drinking was defined as at least one alcoholic drink per month. Subjects who have ever smoked 100 cigarettes in the past were considered as smokers.

SNP Selection

The *ACTB* gene maps to the chromosome 7p22.1 (Gene ID: 60; Locus NC_000007.14) and spans 3.4 kbp and consists of six exons. We searched the SNPs from the upstream 5 kb to the downstream 2 kb and selected tagSNPs through the database of the Chinese Han population in Beijing (CHB) and China of the International Hap MAP Project (HapMap Data Rel 24/phase II

Nov08, on NCBI B36 assembly, dbSNPb126). All tagSNPs were included with a minor allele frequency (MAF) ≥ 0.05 and the criterion of linkage disequilibrium (LD) r^2 ≥0.8. Furthermore, a functional candidate strategy was also applied to select potential functional SNPs on the bioinformatics effect prediction website (SNPINFO, https://snpinfo.niehs.nih.gov/). Finally, three tagSNPs of *ACTB* gene, rs852423 (A > G), rs852426 (T > C), and rs2966449 (T > C), were selected and genotyped in this study. Detailed biological information was summarized in **Table S1**.

Blood Samples Collecting and Genotyping

After an overnight fasting, blood samples were collected in vacuum anticoagulation tubes with ethylenediamine tetraacetic acid dipotassium salt (EDTA-K2)-containing for genotyping. DNA was extracted from the frozen whole blood by a protein precipitation method (Eaglink Cat EGEN2024, NANJING YININGFUSHENG Biotech. Co., Ltd. Nanjing, China) and stored at −20°C. The concentration and the purity of each DNA sample were determined using the NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA). Amplifications of three SNPs were performed using a polymerase chain reaction (PCR)-TaqMan MGB probe array in the GeneAmp® PCR system 9700 thermal cycler (Applied Biosystems, USA). The results were determined on the platform of 7900HT real-time PCR System (Applied Biosystems, Foster City, CA) with a Sequence Detection System (SDS) 2.3 software. To test the genetic effect of SNP, the code of wild types (WT) was 0 (reference), heterozygous types (HT) was 1, and the mutant types (MT) was 2 in additive model. Similar principle was employed in the dominant and recessive models. The successful call rates of SNPs genotyping were 100%. The genotype results were validated by Sanger sequencing, and the consistence was 100% (**Supplementary Materials**).

Statistical Analysis

EpiData 3.0 software (The EpiData Association, Odense, Denmark) was performed for duplicate entry, and consistency check was also available. Quantitative variables were expressed as the mean ± standard deviation (SD), and the categorical variables were presented by their counts and proportions. A Fisher's exact test was used to estimate whether the genotype frequencies in controls met the Hardy–Weinberg equilibrium (HWE). Chi-square (χ^2) test was performed to display the frequencies and distributions of alleles. Binary logistic regression was applied to calculate the odds ratio (OR) and corresponding 95% confidence interval (CI) with adjustment for covariates, and evaluate the genetic effects of different models. All above statistical analyses were conducted on SPSS version 17.0 (SPSS, Inc, Chicago, IL). Stata SE version 12.0 (StataCorp LP, College Station, TX) was used to estimate heterogeneity between groups. Statistical significance was set at 0.05 ($P \leq 0.05$).

RESULTS

Demographic and Clinical Characteristics of DKD Cases and Controls

Table 1 shows the detailed demographic characteristics of included 1,510 couple individuals. The age (mean ± SD) of 69.87 ± 8.61 years in DKD case group was definitely comparable with that of 69.50 ± 8.74 years in control group (P = 0.190). BMI and the proportions of smokers and drinkers in case and control groups were also evenly matched. Nevertheless, the percentage of female in case group was slightly higher than that in the comparison group. Subjects in DKD group had higher levels of systolic pressure (SBP), diastolic pressure (DBP), triglyceride (TG), total cholesterol (TC), low-density lipoprotein cholesterol (LDL-C),

TABLE 1 | Demographic and clinical characteristics of DKD cases and controls.

Characteristics	Group	Case n = 1,510	Control n = 1,510	t/χ^2	P
Age (year)		69.87 ± 8.61	69.50 ± 8.74	1.186	0.190
Gender (%)	Male	543 (35.96%)	606 (40.13%)	7.576	0.018
	Female	967 (64.04%)	904 (59.87%)		
SB*P* (mm Hg)		153.78 ± 22.49	150.98 ± 19.98	3.603	<0.001
DB*P* (mm Hg)		79.47 ± 11.9	79.13 ± 10.4	0.820	<0.001
BMI (kg/m²)		25.02 ± 3.45	24.93 ± 3.52	0.661	0.258
TC (mmol/L)		5.41 ± 1.64	5.22 ± 1.12	3.788	<0.001
TG (mmol/L)		2.13 ± 1.64	1.86 ± 1.54	4.570	0.002
HDL-C (mmol/L)		1.50 ± 0.5	1.51 ± 0.4	0.299	<0.001
LDL-C (mmol/L)		3.26 ± 1.24	3.15 ± 0.92	2.953	<0.001
GLU (mmol/L)		9.68 ± 4.25	9.24 ± 3.05	3.222	<0.001
eGFR (ml·min⁻¹·1.73 m⁻²)		58.69 ± 18.52	83.41 ± 13.44	41.986	<0.001
Smoking (%)	Yes	396 (26.23%)	404 (26.75%)	0.109	0.741
	No	1114 (73.77%)	1106 (73.25%)		
Drinking (%)	Yes	257 (17.02%)	306 (20.26%)	5.247	0.073
	No	1247 (82.58%)	1198 (79.34%)		
	Unknown	6 (0.40%)	6 (0.40%)		

SBP, systolic blood pressure; DBP, diastolic blood pressure; BMI, body mass index; TC, total cholesterol; TG, triglyceride; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; GLU, glucose; eGFR, estimated glomerular filtration rate.

glucose (GLU), and shorter diabetic duration while lower mean high-density lipoprotein cholesterol (HDL-C) and eGFR than controls (all *P* values <0.05).

Association Analysis of the *ACTB* Polymorphisms with DKD

In this study, the genotype distributions of three tagSNPs exactly followed HWE in the control population (all *P* values > 0.05). We observed that there was difference of rs852423 genotypes between DKD cases and controls, after the adjustment for age, gender, smoking, drinking, BMI, diabetic duration, TG, TC, HDL-C, and LDL-C and the borderline *P* values were 0.046, 0.048, and 0.287 for additive, dominant, and recessive models, respectively (**Table 2**). rs852426 T > C variation presented a remarkable higher risk of DKD in the whole study population, the adjusted OR [95% confidence interval (CI)] of additive model was 1.217 (1.079–1.373) with a *P* value of 0.001. After Bonferonni correction (*P*×3), the association still remains statistically significant (*P* = 0.003). The association of rs2966449 (T > C) with DKD was also significant, and the adjusted OR (95% CI) of additive model was 1.151 (1.025–1.293) with a *P* value of 0.018. When rs852426 and rs2966449 entered the logistic regression model at the same time, the association of rs852426 still remained statistically significant (*P* = 0.016), whereas rs2966449 did not (*P* = 0.202).

Stratified Analyses by Age, Gender, Smoking, Drinking, and DM Duration

The ORs (95% CI) of rs852426 (TC + CC vs. TT) for DKD in 70 years or younger and older than 70 years populations were 1.058 (0.857–1.306) and 1.443 (1.168–1.782) with *P* values of 0.601 and 0.001, respectively. Remarkably, there was significant heterogeneity of the association between the two groups with *P* of 0.047 (**Table 3**).

Also, a significant association of rs852426 and DKD was found in populations of male and female, without smoking and without drinking and with DM duration of 10–20 years (**Tables S2–5**). However, no significant heterogeneity was observed between different strata of gender, smoking, drinking, and DM duration.

In addition, the variation of rs2966449 showed significant association with DKD in the populations of older than 70 years, male, without smoking, and without drinking and with DM duration over 20 years (**Tables S2–5**). No significant heterogeneity was observed between different strata of gender, smoking, drinking, and DM duration either.

Stratified Analyses by DKD Severity

Further, we divided the DKD cases by its severity with an eGFR of 60 ml·min^{-1}·1.73 m^{-2}, and the same control group was used to explore the effect of DKD severity on the associations of SNPs and DKD susceptibility (**Table S6**). The results indicated that rs852426 was significantly associated with DKD, the ORs (95% CI) and corresponding *P* values were 1.231 (1.047–1.448) (*P* = 0.012) and 1.181 (1.027–1.357) (*P* = 0.019) in additive model.

Analyses for SNPs of the *ACTB* Polymorphisms and eGFR

One-way ANOVA revealed that the eGFR levels of the individuals with TT and CC genotypes of rs2966449 were significantly lower than that of TC genotype in DKD cases (*P* = 0.021). Even adjusting for gender, smoking, drinking, BMI, and diabetic duration, the difference still remained (*P* = 0.013), whereas no significant difference of eGFR was found among the genotypes of rs852423 and rs852426 (**Table 4**).

TABLE 2 | Association analysis of SNPs of the *ACTB* gene with DKD.

SNP	Group	WT/HT/MT	Genotype OR (95% CI) P value[a]			Allele gene		
			Additive model	Dominant model	Recessive model	Major/minor	OR (95%CI)/*P*[b]	*P*[c]
		AA/AG/GG				A/G		
rs852423	Case	716/648/146	1.12 (1.002–1.253) *P* = 0.046	1.157 (1.001–1.336) *P* = 0.048	1.147 (0.891–1.478) *P* = 0.287	2080/940	1.115 (0.999–1.245) *P* = 0.053	0.960
	Control	765/619/126				2149/871		
		TT/TC/CC				T/C		
rs852426	Case	890/518/102	1.217 (1.079–1.373) *P* = 0.001	1.226 (1.057–1.422) *P* = 0.007	1.522 (1.108–2.089) *P* = 0.009	2298/722	1.219 (1.079–1.376) *P* = 0.001	0.300
	Control	961/479/70				2401/619		
		TT/TC/CC				T/C		
rs2966449	Case	810/589/111	1.151 (1.025–1.293) *P* = 0.018	1.201 (1.039–1.389) *P* = 0.013	1.151 (0.865–1.533) *P* = 0.334	2289/811	1.107 (0.987–1.243) *P* = 0.083	0.307
	Control	874/540/96				2288/732		

WT, wild type; HT, heterozygote; MT, mutant type.
[a]*Adjusted for age, gender, smoking, drinking, BMI and diabetic duration, TG, TC, HDL-C and LDL-C.*
[b]*P value of χ² text for comparison of allele frequencies between case and control groups.*
[c]*P value of Fisher's exact two-sided chi-squared test for Hardy–Weinberg equilibrium.*

TABLE 3 | Stratified analysis of the *ACTB* gene with DKD by age.

SNP	Age	Group	WT/HT/MT	Genotype OR (95% CI)[a]					
				Additive model		Dominant model		Recessive model	
			AA/AG/GG		Heterogeneity test		Heterogeneity test		Heterogeneity test
rs852423	≤70 years	Case	338/305/70	1.124 (0.959–1.318)	$P = 0.994$	1.147 (0.933–1.41)	$P = 0.912$	1.197 (0.837–1.712)	$P = 0.854$
		Control	387/324/66	$P = 0.150$		$P = 0.193$		$P = 0.324$	
	>70 years	Case	378/343/76	1.125 (0.960–1.317)		1.166 (0.951–1.429)		1.140 (0.794–1.637)	
		Control	378/295/60	$P = 0.145$		$P = 0.139$		$P = 0.477$	
			TT/TC/CC		Heterogeneity test		Heterogeneity test		Heterogeneity test
rs852426	≤70 years	Case	426/240/47	1.130 (0.949–1.344)	$P = 0.208$	1.058 (0.857–1.306)	$P = 0.047$	1.815 (1.132–2.908)	$P = 0.375$
		Control	472/274/31	$P = 0.170$		$P = 0.601$		$P = 0.013$	
	>70 years	Case	464/278/55	1.323 (1.116–1.568)		1.443 (1.168–1.782)		1.332 (0.866–2.048)	
		Control	489/205/39	$P = 0.001$		$P = 0.001$		$P = 0.192$	
			TT/TC/CC		Heterogeneity test		Heterogeneity test		Heterogeneity test
rs2966449	≤70 years	Case	386/270/57	1.143 (0.969–1.35)	$P = 0.834$	1.121 (0.911–1.380)	$P = 0.336$	1.446 (0.960–2.178)	$P = 0.190$
		Control	4336/296/45	$P = 0.113$		$P = 0.280$		$P = 0.078$	
	>70 years	Case	424/319/54	1.172 (0.994–1.381)		1.296 (1.055–1.592)		0.959 (0.640–1.436)	
		Control	438/244/51	$P = 0.059$		$P = 0.013$		$P = 0.838$	

[a]*Adjusted for gender, smoking, drinking, BMI and diabetic duration, TG, TC, HDL-C and LDL-C.*

TABLE 4 | Comparison of eGFR among the genotypes of *ACTB*.

SNP	DKD Group	eGFR (mean ± SD)			*P*
rs852423		AA	AG	GG	
	Case	57.97 ± 18.26	59.67 ± 18.74	57.83 ± 18.74	0.203
	Control	83.02 ± 13.58	84.17 ± 13.37	82.11 ± 12.83	0.152
rs852426		TT	TC	CC	
	Case	58.46 ± 18.68	59.06 ± 18.26	58.74 ± 18.57	0.840
	Control	83.01 ± 13.82	84.34 ± 12.78	82.67 ± 12.51	0.189
rs2966449		TT	TC	CC	
	Case	57.88 ± 18.46	60.26 ± 18.53	56.23 ± 18.47	0.021[a]
	Control	83.02 ± 13.35	84.39 ± 13.62	81.55 ± 13.04	0.065

[a]*The average eGFR levels of the TT and CC genotypes of rs2966449 was significantly lower than that of TC genotype in DKD cases (P < 0.05 by SNK test).*

DISCUSSION

Although the GWAS have identified several DN associated loci for European and African populations (Pezzolesi and Krolewski, 2013; Raina et al., 2015), only few of them have been validated, and many remained unvalidated (Regele et al., 2015). Previous GWAS identified 16 SNPs accounted for only 1.4% of the variability of eGFR (Kottgen et al., 2009). Besides, the functional effect of genetic loci often differs across different ethnicities and populations. More promising candidate genes for DKD remain to be explored.

Because *ACTB* loss-of-function mutations result in renal agenesis and impairment, the study of the common variants at *ACTB* and CKD as well as eGFR would be deserved. In our present study, rs852426 T to C variation, located at the upstream of *ACTB*, is significantly associated with DKD, and the association is particularly significant in the older than 70 years populations with even higher OR (95% CI) of 1.425 (1.155–1.759). The results suggest that aging might promote the impact of rs852426 T to C variation on the development of DKD. We also observed that individuals with TT or CC genotypes of rs2966449 presented lower eGFR level than those with TC genotype, and the result is consistent with the functional characteristics that *ACTB* haploinsufficiency leads to reduced cell proliferation, altered expression of cell-cycle genes, and decreased amounts of nuclear (Cuvertino et al., 2017). Previous studies found that DKD occurs incrementally during T2DM in those with long diabetic duration or advanced age (Satirapoj and Adler, 2014). Thus, this study would provide an available genetic marker of precisive prevention from DKD for T2DM patients.

Obviously, the average age in this study, which is near 70 years, has efficient power to find the cut-point of age heterogeneity of the association. The association of rs852426 and DKD also emerged in the subgroup without smoking and drinking and with DM duration of 10–20 years. So, further replication study would warrant large sample to identify the target population of genetic susceptibility to the variation of rs852426. The weak association of rs2966449 and DKD observed in this study is necessary to be replicated by further study.

Previous study demonstrated that β-cytoplasmic actin structures were disorganized, and its expression was downregulated in the epithelial-to-mesenchymal transition (EMT) process in the involvement of fibroblastic phenotype (Shagieva et al., 2012). Ishizawa et al. (2014) reported that Rho kinase was stimulated by its upstream factors in cultured renal tubular cells, triggering EMT to develop to renal fibrosis. In addition, the ACTB gene is a downstream gene in the Rho/ROCK pathway. Besides, animal experiment has reported that neonatal rats subjected to partial unilateral ureteral obstruction (PUUO) could produce more actin filaments to control and maintain kidney shape and internal structure (Zhao et al., 2016). Along with the evidence mentioned above, we therefore speculated with caution that the ACTB gene may take a role in the EMT-induced DKD. Bioinformatic analysis indicated that the variations of three tagging SNPs selected in this study as well as other SNPs with MAF >0.05 in ACTB would not cause the structural changes of amino acids; we speculate that the variations of rs852426 and rs2966449 might affect the transcriptional level of ACTB by an epigenetic modification way. Further biological functional research would be warranted to explore the potential effect of the variants at ACTB on gene transcription, splicing, or mRNA stability as well as the molecular mechanism of DKD.

Our study for the first time explored the associations of the ACTB genetic variations with DKD and age- and gender-matched comparison, which was adopted to balance the common confounding factors. Besides, a standard that subjects with T2DM duration more than 5 years were qualified for the study was applied to weaken the discrimination from diabetic duration, which was also guaranteed by statistical adjustments. Moreover, the current subjects were representative because they were selected from about 20,340 T2DM patients from community-based chronic disease management system, distributed in the southern and northern of Jiangsu Province.

There indeed consisted several limitations that should not be ignored. First, we selected candidate SNPs at the ACTB with the criterion of MAF ≥ 0.05 and could have missed the rare variants with MAF <0.05 that may have substantial effects on the occurrence and progress of DKD. Second, case-control design cannot avoid the causal confusion caused by time sequence. Besides, the diabetic duration in this study referred to the time gap from first-diagnosed T2DM to the investigation day, which may overestimate the onset time of DKD. Further follow-up study is warranted.

CONCLUSION

Data from the present study suggest that the ACTB genetic variants confer the susceptibility to DKD. SNP rs852426 at this gene is associated with the increased risk of DKD while the association is particularly marked in T2DM patients over 70 years old. SNP rs2966449, however, shows a mild association with DKD and eGFR.

ETHICS STATEMENT

Informed consent was obtained from all participants and approved by the Research Ethics Committee of Jiangsu Provincial Center for Disease Control and Prevention and Nanjing Medical University.

AUTHOR CONTRIBUTIONS

CS and MW designed the program. MW, YQ, JZ, JS, EP, QZ, NZ, and HS collected the data. ML, JD, and YT extracted DNA and performed genotyping work. ML analyzed the data and wrote the manuscript. CS edited and proofed the manuscript.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.00663/full#supplementary-material

REFERENCES

Ahluwalia, T. S., Schulz, C. A., Waage, J., Skaaby, T., Sandholm, N., van Zuydam, N., et al. (2019). A novel rare CUBN variant and three additional genes identified in Europeans with and without diabetes: results from an exome-wide association study of albuminuria. Diabetologia 62, 292–305. doi: 10.1007/s00125-018-4783-z

American Diabetes Association (2010). Diagnosis and classification of diabetes mellitus. Diabetes Care 33 Suppl 1, S62–S69. doi: 10.2337/dc10-S062

Cho, N. H., Shaw, J. E., Karuranga, S., Huang, Y., da Rocha Fernandes, J. D., Ohlrogge, A. W., et al. (2018). IDF Diabetes Atlas: global estimates of diabetes prevalence for 2017 and projections for 2045. Diabetes Res. Clin. Pract. 138, 271–281. doi: 10.1016/j.diabres.2018.02.023

Collaborators GBDCoD (2018). Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 392, 1736–1788. doi: 10.1016/S0140-6736(18)32203-7

Cooper, M. E. (2012). Diabetes: treating diabetic nephropathy-still an unresolved issue. Nat. Rev. Endocrinol. 8, 515–516. doi: 10.1038/nrendo.2012.125

Cuvertino, S., Stuart, H. M., Chandler, K. E., Roberts, N. A., Armstrong, R., Bernardini, L., et al. (2017). ACTB loss-of-function mutations result in a pleiotropic developmental disorder. Am. J. Hum. Genet. 101, 1021–1033. doi: 10.1016/j.ajhg.2017.11.006

Guan, M., Keaton, J. M., Dimitrov, L., Hicks, P. J., Xu, J., Palmer, N. D., et al. (2019). Genome-wide association study identifies novel loci for type 2

diabetes-attributed end-stage kidney disease in African Americans. *Hum. Genomics* 13, 21. doi: 10.1186/s40246-019-0205-7

Gurung, R. L., Dorajoo, R., Liu, S., Yiamunaa, M., Liu, J.-J., Wang, L., et al. (2018). Genetic markers for urine haptoglobin is associated with decline in renal function in type 2 diabetes in East Asians. *Sci. Rep.* 8, 5109. doi: 10.1038/s41598-018-23407-1

Ishizawa, S., Takahashi-Fujigasaki, J., Kanazawa, Y., Matoba, K., Kawanami, D., Yokota, T., et al. (2014). Sphingosine-1-phosphate induces differentiation of cultured renal tubular epithelial cells under Rho kinase activation *via the* S1P2 receptor. *Clin. Exp. Nephrol.* 18, 844–852. doi: 10.1007/s10157-014-0933-x

Iyengar, S. K., Sedor, J. R., Freedman, B. I., Kao, W. H., Kretzler, M., Keller, B. J., et al. (2015). Genome-wide association and trans-ethnic meta-analysis for advanced diabetic kidney disease: family investigation of nephropathy and diabetes (FIND). *PLoS Genet.* 11, e1005352. doi: 10.1371/journal.pgen.1005352

Kdoqi (2007). KDOQI clinical practice guidelines and clinical practice recommendations for diabetes and chronic kidney disease. *Am. J. Kidney Dis.* 49, S12–S154. doi: 10.1053/j.ajkd.2006.12.005

Kitada, M., Ogura, Y., Suzuki, T., Sen, S., Lee, S. M., Kanasaki, K., et al. (2016). A very-low-protein diet ameliorates advanced diabetic nephropathy through autophagy induction by suppression of the mTORC1 pathway in Wistar fatty rats, an animal model of type 2 diabetes and obesity. *Diabetologia* 59, 1307–1317. doi: 10.1007/s00125-016-3925-4

Kong, L. L., Wu, H., Cui, W. P., Zhou, W. H., Luo, P., Sun, J., et al. (2013). Advances in murine models of diabetic nephropathy. *J. Diabetes Res.* 2013, 797548. doi: 10.1155/2013/797548

Kottgen, A., Glazer, N. L., Dehghan, A., Hwang, S. J., Katz, R., Li, M., et al. (2009). Multiple loci associated with indices of renal function and chronic kidney disease. *Nat. Genet.* 41, 712–717. doi: 10.1038/ng.377

Lim, S. C., Dorajoo, R., Zhang, X., Wang, L., Ang, S. F., Tan, C. S. H., et al. (2017). Genetic variants in the receptor for advanced glycation end products (RAGE) gene were associated with circulating soluble RAGE level but not with renal function among Asians with type 2 diabetes: a genome-wide association study. *Nephrol. Dial. Transplant.* 32, 1697–1704. doi: 10.1093/ndt/gfw263

Liu, M., Liu, S. W., Wang, L. J., Bai, Y. M., Zeng, X. Y., Guo, H. B., et al. (2018). Burden of diabetes, hyperglycaemia in China from to 2016: findings from the 1990 to 2016, global burden of disease study. *Diabetes Metab.* 45 (3), 286–293. doi: 10.1016/j.diabet.2018.08.008

Mauer, M., Caramori, M. L., Fioretto, P., and Najafian, B. (2015). Glomerular structural-functional relationship models of diabetic nephropathy are robust in type 1 diabetic patients. *Nephrol. Dial. Transplant.* 30, 918–923. doi: 10.1093/ndt/gfu279

Miao, D. D., Pan, E. C., Zhang, Q., Sun, Z. M., Qin, Y., and Wu, M. (2017). Development and validation of a model for predicting diabetic nephropathy in Chinese People. *Biomed. Environ. Sci.* 30, 106–112. doi: 10.3967/bes2017.014

Pezzolesi, M. G., and Krolewski, A. S. (2013). The genetic risk of kidney disease in type 2 diabetes. *Med. Clin. North Am.* 97, 91–107. doi: 10.1016/j.mcna.2012.10.005

Pollard, T. D., and Cooper, J. A. (2009). Actin, a central player in cell shape and movement. *Science* 326, 1208–1212. doi: 10.1126/science.1175862

Raina, P., Sikka, R., Kaur, R., Sokhi, J., Matharoo, K., Singh, V., et al. (2015). Association of transforming growth factor beta-1 (TGF-beta1) genetic variation with type 2 diabetes and end stage renal disease in two large population samples from North India. *OMICS* 19, 306–317. doi: 10.1089/omi.2015.0005

Regele, F., Jelencsics, K., Shiffman, D., Pare, G., McQueen, M. J., Mann, J. F., et al. (2015). Genome-wide studies to identify risk factors for kidney disease with a focus on patients with diabetes. *Nephrol. Dial. Transplant.* 30 Suppl 4, iv26–iv34. doi: 10.1093/ndt/gfv087

Remuzzi, G., Schieppati, A., and Ruggenenti, P. (2002). Clinical practice. Nephropathy in patients with type 2 diabetes. *N. Engl. J. Med.* 346, 1145–1151. doi: 10.1056/NEJMcp011773

Satirapoj, B., and Adler, S. G. (2014). Comprehensive approach to diabetic nephropathy. *Kidney Res. Clin. Pract.* 33, 121–131. doi: 10.1016/j.krcp.2014.08.001

Shagieva, G. S., Domnina, L. V., Chipysheva, T. A., Ermilova, V. D., Chaponnier, C., and Dugina, V. B. (2012). Actin isoforms and reorganization of adhesion junctions in epithelial-to-mesenchymal transition of cervical carcinoma cells. *Biochemistry (Mosc.).* 77, 1266–1276. doi: 10.1134/S0006297912110053

Shen, C., Wen, J., Pan, X., Su, J., Du, W., Pan, E., et al. (2018). Age at menarche and age at natural menopause as predictors of glycemic control in type 2 diabetic patients. *J. Diabetes Complicat.* 32, 623–629. doi: 10.1016/j.jdiacomp.2018.04.012

Tervaert, T. W., Mooyaart, A. L., Amann, K., Cohen, A. H., Cook, H. T., Drachenberg, C. B., et al. (2010). Pathologic classification of diabetic nephropathy. *J. Am. Soc. Nephrol.* 21, 556–563. doi: 10.1681/ASN.2010010010

Trikkalinou, A., Papazafiropoulou, A. K., and Melidonis, A. (2017). Type 2 diabetes and quality of life. *World J. Diabetes* 8, 120–129. doi: 10.4239/wjd.v8.i4.120

van Zuydam, N. R., Ahlqvist, E., Sandholm, N., Deshmukh, H., Rayner, N. W., Abdalla, M., et al. (2018). A genome-wide association study of diabetic kidney disease in subjects with type 2 diabetes. *Diabetes* 67, 1414–1427. doi: 10.2337/db17-0914

Wang, L., Gao, P., Zhang, M., Huang, Z., Zhang, D., Deng, Q., et al. (2017). Prevalence and ethnic pattern of diabetes and prediabetes in China in 2013. *JAMA* 317, 2515–2523. doi: 10.1001/jama.2017.7596

Ying, Q., and Wu, G. (2017). Molecular mechanisms involved in podocyte EMT and concomitant diabetic kidney diseases: an update. *Ren. Fail.* 39, 474–483. doi: 10.1080/0886022X.2017.1313164

Zelnick, L. R., Weiss, N. S., Kestenbaum, B. R., Robinson-Cohen, C., Heagerty, P. J., Tuttle, K., et al. (2017). Diabetes and CKD in the United States Population, 2009–2014. *Clin. J. Am. Soc. Nephrol.* 12, 1984–1990. doi: 10.2215/CJN.03700417

Zhao, Q., Xue, Y., Yang, Y., Niu, Z., Wang, C., Hou, Y., et al. (2016). Screening and identification of the differentially expressed proteins in neonatal rat kidney after partial unilateral ureteral obstruction. *Mol. Med. Rep.* 14, 681–688. doi: 10.3892/mmr.2016.5338

Impact of a Complement Factor H Gene Variant on Renal Dysfunction, Cardiovascular Events, and Response to ACE Inhibitor Therapy in Type 2 Diabetes

Elisabetta Valoti[1†], Marina Noris[1*†], Annalisa Perna[1], Erica Rurali[1], Giulia Gherardi[1], Matteo Breno[1], Aneliya Parvanova Ilieva[1], Ilian Petrov Iliev[1], Antonio Bossi[2], Roberto Trevisan[3], Alessandro Roberto Dodesini[3], Silvia Ferrari[1], Nadia Stucchi[1], Ariela Benigni[1], Giuseppe Remuzzi[1,4,5†] and Piero Ruggenenti[1,4†] on behalf of the BENEDICT Study Group

[1] Aldo e Cele Daccò Clinical Research Center for Rare Diseases, Istituto di Ricerche Farmacologiche Mario Negri—IRCCS, Ranica, Italy, [2] Units of Diabetology of Treviglio Hospital, Treviglio, Italy, [3] Unit of Diabetology, Azienda Socio-Sanitaria Territoriale Papa Giovanni XXIII, Bergamo, Italy, [4] Unit of Nephrology, Azienda Socio-Sanitaria Territoriale Papa Giovanni XXIII, Bergamo, Italy, [5] Department of Biomedical and Clinical Sciences, University of Milan, Milan, Italy

Complement activation has been increasingly implicated in the pathogenesis of type 2 diabetes and its chronic complications. It is unknown whether complement factor H (CFH) genetic variants, which have been previously associated with complement-mediated organ damage likely due to inefficient complement modulation, influence the risk of renal and cardiovascular events and response to therapy with angiotensin-converting enzyme inhibitors (ACEi) in type 2 diabetic patients. Here, we have analyzed the c.2808G>T, (p.Glu936Asp) CFH polymorphism, which tags the H3 CFH haplotype associated to low plasma factor H levels and predisposing to atypical hemolytic uremic syndrome, in 1,158 type 2 diabetics prospectively followed in the Bergamo nephrologic complications of type 2 diabetes randomized, controlled clinical trial (BENEDICT) that evaluated the effect of the ACEi trandolapril on new onset microalbuminuria. At multivariable Cox analysis, the p.Glu936Asp polymorphism (Asp/Asp homozygotes, recessive model) was associated with increased risk of microalbuminuria [adjusted hazard ratio (HR) 3.25 (95% CI 1.46–7.24), $P = 0.0038$] and cardiovascular events [adjusted HR 2.68 (95% CI 1.23–5.87), $P = 0.013$]. The p.Glu936Asp genotype significantly interacted with ACEi therapy in predicting microalbuminuria. ACEi therapy was not nephroprotective in Asp/Asp homozygotes [adjusted HR 1.54 (0.18–13.07), $P = 0.691$ vs. non-ACEi-treated Asp/Asp patients], whereas it significantly reduced microalbuminuria events in Glu/Asp or Glu/Glu patients [adjusted HR 0.38 (0.24–0.60), $P < 0.0001$ vs. non-ACEi-treated Glu/Asp or Glu/Glu patients]. Among ACEi-treated patients, the risk of developing cardiovascular events was higher in Asp/Asp homozygotes than in Glu/Asp or Glu/Glu patients [adjusted HR 3.26 (1.29–8.28), $P = 0.013$]. Our results indicate that type 2 diabetic patients Asp/Asp homozygotes in the p.Glu936Asp

CFH polymorphism are at increased risk of microalbuminuria and cardiovascular complications and may be less likely to benefit from ACEi therapy. Further studies are required to confirm our findings.

Keywords: diabetes, complement, factor H, ACE inhibitors, microalbuminuria, cardiovascular risk, diabetes complications

INTRODUCTION

Complement activation products have been implicated in renal and cardiovascular complications of type 2 diabetes by promoting endothelial dysfunction and inflammation (Onat et al., 2011; Ghosh et al., 2015; Merle et al., 2015). Deposits of C3 activation fragments and of the terminal complement complex C5b-9 are observed in glomeruli and arteries of diabetic rats (Fischetti et al., 2011). The close association of these deposits with proteinuria, mesangial expansion, vascular hypertrophy, extracellular matrix deposition, and increased expression of adhesion molecules and growth factors strongly suggests that complement activation is involved in the onset and progression of renal and vascular complications of experimental diabetes (Fujita et al., 2013). Consistently, glomerular functional and structural changes are blunted by treatment with complement inhibitors, and vascular changes are almost fully prevented in C6-deficient diabetic rats that cannot form C5b-9 (Fujita et al., 1999; Fischetti et al., 2011).

In diabetic patients, elevated glomerular and tubular expression of C3 and Factor B—the two components of the alternative pathway C3 convertase C3bBb that cleaves C3 to C3a and C3b—has been associated with overt diabetic nephropathy (Woroniecka et al., 2011). In addition, plasma levels of Bb and C3a were higher in diabetic patients with nephropathy than in those without renal involvement (Li et al., 2018). Finally, plasma levels of C3 activation products correlated with the risk of severe atherosclerosis and ischemic heart disease in type 2 diabetics (Figueredo et al., 1993; Woroniecka et al., 2011; Fujita et al., 2013; Hertle et al., 2014; Li et al., 2018). The previously discussed findings converge to indicate that complement activation via the alternative pathway may contribute to the onset and progression of renal and vascular complications in human diabetes. This possibility is confirmed by the recent observation that among patients with diabetic nephropathy, the presence of deposits of C3 activation products in kidney biopsy was associated with lower renal function and more severe tubular and glomerular damage (Sun et al., 2018).

Should complement activation play a key role in the pathophysiology of chronic diabetic complications, then the availability of modulators of complement activity for clinical use would have major implications. Approximately one third of type 2 diabetics continue to develop micro- and macro-vascular disease, despite optimized metabolic and blood pressure control and early treatment with renin–angiotensin system (RAS) inhibitors. In these patients, renal and vascular dysfunctions, often heralded by a transition from normo- to microalbuminuria, substantially increase the risk of major renal and cardiovascular events, including overt nephropathy and progression to end-stage kidney disease, and coronary artery disease, myocardial infarction, and stroke (Holtkamp et al., 2011; Ruggenenti et al., 2011).

Complement factor H (CFH) plays a central role in the modulation of the complement alternative pathway by facilitating C3b degradation by the plasma serine protease factor I and enhancing C3 convertase dissociation (Jozsi and Zipfel, 2008). Factor H acts both in the fluid phase and on the endothelial cell surface, where it binds to heparan sulfate molecules and to C3 activation products deposited on cell membranes (Heinen et al., 2007). This is instrumental for protecting the host from excess complement activation and complement-mediated renal and vascular injury upon cell/tissue exposure to agents that may activate the alternative pathway. Consistently, reduced CFH bioavailability or activity, due to gene mutations or autoantibodies, may result in uncontrolled complement activation and consequent severe vascular damage, as observed in patients with atypical hemolytic uremic syndrome (aHUS), a rare thrombotic microangiopathy that targets the microvasculature of the kidney and other organs (Noris and Remuzzi, 2009). In addition, the CFH H3 common haplotype that includes polymorphisms in the promoter (rs3753394, c-331C>T) and the coding region of CFH (rs3753396, c.A2016G, p.Q672Q; and rs 1065489, c.2808G>T, p.Glu936Asp) confers an increased risk of aHUS and favors a poorer renal prognosis (Bernabeu-Herrero et al., 2015). Notably, among healthy controls, subjects homozygous for the H3 haplotype were found to have lower plasma CFH levels compared with individuals with zero H3 copies (Bernabeu-Herrero et al., 2015; Pouw et al., 2018). Furthermore, the previously mentioned CFH polymorphisms have been associated with susceptibility to other kidney diseases (Bonomo et al., 2014), eye diseases (Miki et al., 2014; Garcia et al., 2015; Wang et al., 2016), and infections (Davila et al., 2010; Zhang et al., 2013).

We postulated that activation of the complement system, possibly mediated by genetically determined reduced CFH bioavailability, could play a central role in the onset and progression of renal and vascular complications of diabetes. To explore whether and to what extent the CFH H3 haplotype may affect the risk of renal involvement in diabetes, we performed a post hoc analysis of the c.2808G>T (p.Glu936Asp) polymorphism, which is strongly associated with the H3 haplotype and determines an amino-acidic change close to the cysteine 931 involved in CFH folding (Reid and Day, 1989), in a large cohort of normoalbuminuric type 2 diabetics who were prospectively monitored through serial measurements of urinary albumin excretion (UAE) in the context of the

Bergamo nephrologic complications of type 2 diabetes randomized, controlled clinical trial (BENEDICT) (Ruggenenti et al., 2004). Since the trial found that the ACE inhibitor (ACEi) trandolapril—alone or in combination with verapamil—reduced the risk of progression to microalbuminuria and the number of cardiovascular events (Ruggenenti et al., 2004; Ruggenenti et al., 2012), here we sought to explore whether and to what extent the c.2808G>T (p.Glu936Asp) *CFH* polymorphism could affect the interactions of diabetes and ACE inhibition with the previously discussed outcomes.

RESEARCH DESIGN AND METHODS

This is a post hoc analysis of the BENEDICT trial, a multicenter, double-blind, placebo-controlled, randomized clinical trial, designed to assess whether the ACEi trandolapril and the non-dihydropyridine calcium-channel blocker verapamil, alone or in combination, would prevent microalbuminuria in 1,204 subjects with hypertension, type 2 diabetes, and normal UAE. Detailed information about the trial is provided elsewhere (The BENEDICT Group, 2003; Ruggenenti et al., 2004).

Whereas the effect of verapamil was similar to that of the placebo, either trandolapril alone or in combination with verapamil reduced to a similar extent the risk of progression to microalbuminuria (The BENEDICT Group, 2003; Ruggenenti et al., 2004), a specific effect possibly related to RAS inhibition, which was independent of blood pressure and metabolic control. Thus, for the purpose of this study, patients were pooled in two cohorts according to their original allocation to ACEi or non-ACEi therapy regardless of concomitant therapy with verapamil or placebo: the ACEi group included patients treated with trandolapril or trandolapril plus verapamil, and the non-ACEi group included patients treated with verapamil or with placebo. Gene-by-treatment interactions were tested according to ACEi therapy (yes or no).

Objectives

To investigate whether the p.Glu936Asp CFH variant could affect the risk of microalbuminuria and the protective effect of ACEi against this event, we genotyped the c.2808G>T (rs1065489) single nucleotide polymorphism (SNP) in 1,158 of the 1,204 type 2 normoalbuminuric diabetic patients included in the BENEDICT phase A study who consented to genetic analyses. Patients were actively followed until June 2004, when results of the final analysis became available (Ruggenenti et al., 2004). One hundred forty-seven nondiabetic volunteers from the same geographical region (Lombardy) served as healthy controls.

Primary—We first compared the allele frequency of the p.Glu936Asp CFH polymorphism between diabetic patients who developed microalbuminuria and those who did not develop microalbuminuria during follow-up. Thereafter, we aimed to evaluate the association between the p.Glu936Asp genotype according to different models and progression from normo- to microalbuminuria. We then evaluated the interactive role of the p.Glu936Asp CFH polymorphism and ACEi therapy in predicting new-onset microalbuminuria.

Secondary—Then, we evaluated the role of the previously mentioned interaction in predicting first onset of one of the components of a composite endpoint of fatal (including sudden death) or nonfatal major cardiovascular events, including events related to coronary (acute myocardial infarction, unstable angina pectoris, or coronary revascularization by bypass grafting or percutaneous transluminal angioplasty), cerebrovascular (stroke, transient ischemic attack, pre-cerebral artery revascularization) or peripheral artery (amputation, revascularization) disease, and hospitalization because of congestive heart failure.

Competing events—Patients who developed microalbuminuria were not censored, and their follow-up was continued until study end in order to also capture cardiovascular events occurring after progression to microalbuminuria. In addition, the risk that major cardiovascular events precluded the occurrence of microalbuminuria can be considered negligible. Indeed, during the BENEDICT phase A core trial, only five participants died from cardiovascular events. Moreover, the majority of patients experiencing nonfatal cardiovascular events remained in the core trial and had their albuminuria evaluated after the event.

Definitions

The new onset of microalbuminuria was defined as UAE ≥20 and <200 μg/min in at least two of three consecutive overnight urine collections at two consecutive visits 2 months apart in previously normoalbuminuric (UAE < 20 μg/min in at least two out of three consecutive overnight samples) subjects (The BENEDICT Group, 2003; Ruggenenti et al., 2004). All cardiovascular events were adjudicated by two cardiologists (Brigitte Kalsh and Piero Ruggenenti) blinded to treatment and genetic analysis.

Genotyping

Genomic DNA was extracted from peripheral blood leukocytes by Nucleon BACC2 kit (Amersham). Genotyping for c.2808G>T SNP of *CFH* was performed with Sanger direct sequencing using a primer pair designed to amplify exon 19 of CFH (Reid and Day, 1989) and the 3730-XL sequencer (Applied Biosystems).

Ethics and Data Handling

The study was approved by the local ethics committee, and all study participants provided written informed consent according to the Helsinki Declaration guidelines. Data were handled with respect for patient confidentiality and anonymity.

Statistical Analyses

The outcomes of interest were time to microalbuminuria or to major cardiovascular events. For patients who did not reach the endpoint, we censored time at the last follow-up visit with available data for albuminuria and at the last follow-up visit for major cardiovascular events.

All time-to-event endpoints were analyzed using Cox proportional hazard regression models, and results were expressed as hazard ratio (HR) and 95% confidence interval (CI). The Kaplan–Meier method was used to plot the probability of achieving the endpoints, according to the c.2808G>T (p.Glu636Asp) polymorphism and the

ACEi treatment. Multivariable models for the endpoints included c.2808G>T (p.Glu636Asp) genotype, ACEi treatment and all the baseline covariates that, at the univariable Cox analysis, significantly (*P* < 0.05) associated with the outcome without exceeding the limit of one independent variable included in the model for every at least 10 outcome events available for the analyses. The previously discussed variables were also tested in Cox models with genotype × ACEi treatment interaction terms. In the multivariable approach, blood glucose was not considered because of its high colinearity with glycosylated hemoglobin (HbA1c) level that is a better biomarker of metabolic control as compared with blood glucose.

The covariates excluded in the univariable approach were considered in sensitivity analyses by testing each of the previously mentioned covariates in the multivariable models with genotype × ACEi treatment interaction. The same approach was used in exploratory statistical analyses to consider the predictive value of mean systolic blood pressure (SBP), diastolic blood pressure (DBP), mean arterial pressure, and Hba1C in the previously mentioned multivariable model, calculated on the basis of mean values during follow-up instead of baseline values.

Tests of the proportional hazard assumption were based on Schoenfeld residuals. To test possible differences between the four groups derived from genotype × ACEi treatment interaction, we used linear mixed effect models for DBP, SBP, and HbA1c. Not normally distributed covariates were log transformed before analysis. Normality for continuous variables was assessed by means of the Shapiro–Wilk and Kolmogorov–Smirnov tests.

Baseline characteristics were presented as numbers and percentages, means and standard deviations (SD), or medians and interquartile ranges (IQR), as appropriate. Comparisons between groups were made using unpaired t-test, Wilcoxon rank sum test, Chi-squared test, or paired t-test, as appropriate. Comparison between groups of UAE changes from baseline to the final visit was carried out by analysis of covariance. All P values were two-sided. Bonferroni correction was applied for multiple testing. Analyses were carried out using SAS (version 9.2) and Stata (version 13).

RESULTS

The genotype distribution of the c.2808G>T (p.Glu936Asp) C*FH* SNP in the 1,158 type 2 diabetics from the BENEDICT trial was comparable with the distribution in 145 nondiabetic, healthy volunteers with similar ancestry and geographic origin and did not deviate from the Hardy–Weinberg equilibrium (**Table 1**).

TABLE 1 | Genotypic distribution of *CFH* SNP rs1065489 (c.2808 G>T, p.Glu936Asp) in type 2 diabetic patients of BENEDICT study and in healthy subjects.

c.2808 G>T	n	G/G	G/T	T/T	Hardy–Weinberg equilibrium
T2D patients	1,158	801	321	36	$X^2 = 0.174$; $P = 0.917$
Healthy subjects	145	97	46	2	$X^2 = 0.971$; $P = 0.615$
			$X^2 = 2.152$; $P = 0.341$		

Complement Factor H Single Nucleotide Polymorphism Variants According to Progression to Microalbuminuria

Over a median (IQR) follow-up of 42 (12–51) months, 98 of the 1,158 participants (8.5%) progressed to new-onset microalbuminuria (**Figure 1**). Allele frequencies of the *CFH* c.2808G>T SNP differed significantly between patients with or without events (**Table 2**).

We found a suggestive evidence for potential association of the minor T variant (p.936Asp) with microalbuminuria, according to genotype distribution and the additive and the recessive models (**Table 2**). Considering the recessive model, 7 out of 36 Asp/Asp (T/T) homozygotes (19.4%) vs. 91 out of 1,122 Glu/Glu+Glu/Asp (TG/GG) patients (8.1%) developed new-onset microalbuminuria [Cox univariable analysis: HR (95% CI): 2.46 (1.14–5.32), *P* = 0.021, **Table 3**, Kaplan–Meier curve is shown in **Figure 2A**]. Cox univariable analysis with the additive model was not significant after Bonferroni correction [HR (95% CI): 1.4 (1.02–1.98), *P* = 0.04, **Table 3**, Kaplan–Meier curve is shown in **Supplementary Figure 1A**].
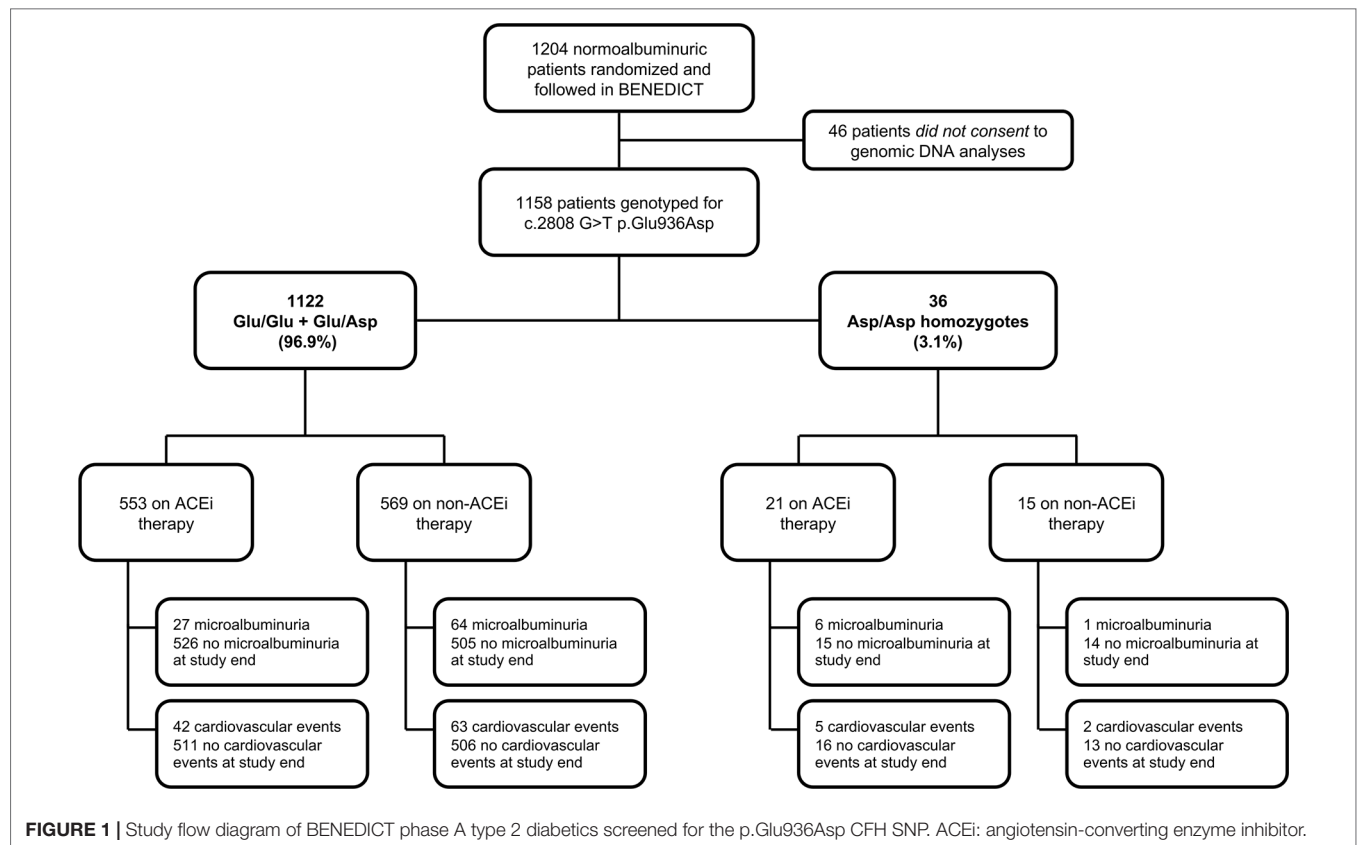
Baseline characteristics, the proportion of patients allocated to ACEi or non-ACEi therapy (**Table 4**), and the distribution of concomitant medications (**Supplementary Table 1**) were similar between the Asp/Asp and Glu/Glu+Glu/Asp genotype groups with the exception of a higher baseline HbA1c in Asp/Asp homozygotes [Asp/Asp: median 5.8% (IQR: 5.2–7.4), Glu/Glu+Glu/Asp: median 5.6 (IQR: 4.8–6.5), *P* < 0.05, **Table 4**].

Interactions Between p.Glu936Asp Genotype, Microalbuminuria, and Response to Angiotensin-Converting Enzyme Inhibitor Treatment

At univariable Cox analyses, the following covariates, male gender, smoking, and higher baseline UAE, HbA1c, and blood glucose were significantly associated with an increased risk of microalbuminuria, and ACEi therapy was associated with protection against this event (**Table 3**). Baseline UAE and HbA1c, the p.Glu936Asp genotype (recessive model), and ACEi therapy retained an independent association with new-onset microalbuminuria at multiregression analysis (**Table 5**), which included all the baseline covariates that significantly predicted the outcome in the univariable approach (**Table 3**). With the additive model, no significant association was found between the p.Glu936Asp genotype and new-onset microalbuminuria at multiregression analysis (**Supplementary Table 3**).

Among Asp/Asp homozygotes, there were more participants with insulin monotherapy in the non-ACEi group compared with the ACEi group (**Supplementary Table 1**). Among Glu/Glu+Glu/Asp patients, there were more patients receiving diuretics and sympatholytic agents, in the non-ACEi group, compared with the ACEi group (**Supplementary Table 1**).

In a Cox model including the p.Glu936Asp genotype (recessive model), ACEi therapy, and their interaction, the p.Glu936Asp genotype significantly interacted with ACEi therapy in predicting microalbuminuria [HR = 0.096, 95% CI (0.011–0.837), *P* = 0.034, **Table 6**].

FIGURE 1 | Study flow diagram of BENEDICT phase A type 2 diabetics screened for the p.Glu936Asp CFH SNP. ACEi: angiotensin-converting enzyme inhibitor.

Consistently, progression to microalbuminuria was observed in 6 of 21 Asp/Asp homozygotes on ACEi (28.6%) compared with 1 of 15 (6.7%) on non-ACEi [HR 4.03, 95% CI (0.49–33.50)] and in 27 of 553 (4.9%) Glu/Glu+Glu/Asp patients on ACEi compared with 64 of 569 (11.3%) on non-ACEi [HR 0.39 (0.25–0.61), $P <$ 0.0001] (**Figure 1** and **Figure 3A**, **Supplementary Table 2**). Similar results were obtained after adjustment for baseline covariates that, at univariable analyses, were significantly associated with the event [**Table 7**, Asp/Asp homozygotes, ACEi vs. non-ACEi, adjusted HR 1.54, 95% CI (0.18–13.07), Glu/Glu+Glu/Asp ACEi vs. non-ACEi, adjusted HR 0.38, 95% CI (0.24–0.60) $P <$ 0.0001]. Among ACEi-treated patients, Asp/Asp homozygotes had more microalbuminuria events than patients with Glu/Glu+Glu/Asp genotypes (adjusted HR 4.72, 95% CI [1.93–11.52], $P =$ 0.001, **Table 7**), while among the non-ACEi patients, microalbuminuria events were comparable in the two genotype groups [adjusted HR 1.16, 95% CI (0.16–8.63), **Table 7**]. The adjusted HR for progression to microalbuminuria events increased progressively from 1 in Glu/Glu+Glu/Asp on ACEi (reference group) to 2.63 (95% CI 1.67–4.14, $P <$ 0.0001) in Glu/Glu+Glu/Asp on non-ACEi, to 3.06 (95% CI 0.41–23.04) in Asp/Asp homozygotes on non-ACEi, and to 4.72 (95% CI 1.93–11.52, $P =$ 0.001) in Asp/Asp on ACEi (**Table 7** and **Figure 4**).

Notably, among patients on non-ACEi, the risk of microalbuminuria events was comparable between Asp/Asp homozygotes and Glu/Glu+Glu/Asp patients (**Table 7**, **Figure 4**). Consistent findings were observed when the changes in albuminuria levels from baseline to the end of the study (or to new onset of microalbuminuria) were considered as a

continuous variable (**Figure 5**). Thus, in ACEi-treated patients, UAE increased by 7% (median value) vs. baseline in Asp/Asp homozygotes, while it decreased by 8% in Glu/Glu+Glu/Asp patients. In non-ACEi patients, UAE increased by 6.5% and 3% in Asp/Asp homozygotes and in Glu/Glu+Glu/Asp patients, respectively (**Figure 5**). After adjustment for UAE at baseline, ACEi treatment did not impact on final UAE values in Asp/Asp homozygotes ($P =$ 0.493 by analysis of covariance), whereas it significantly reduced UAE ($P <$ 0.0015) in Glu/Glu+Glu/Asp patients compared with Glu/Glu+Glu/Asp on non-ACEi.

Relationship Between Albuminuria and Cardiovascular Outcomes

Baseline UAE values were significantly associated with risk of cardiovascular events at both univariable (**Table 3**) and multivariable [HR 1.74 (95% CI 1.26–2.40), $P =$ 0.0007] analyses (**Table 5**). There was also a significant association between new-onset microalbuminuria during the core study and the risk of cardiovascular events throughout the whole study period [**Table 3**, HR 1.85, 95% CI (1.11–3.11), $P =$ 0.0191].

Interactions Between p.Glu936Asp Genotype and Cardiovascular Events

During a median (IQR) follow-up of 60 (52–67) months, at least one major cardiovascular event was observed in 112 (9.6%) of the 1,158 patients (**Figure 1**). According to the recessive

TABLE 2 | Association analysis of genotypes and alleles of CFH c.2808G>T SNP with new-onset microalbuminuria in type 2 diabetic patients from BENEDICT phase A.

SNP	M/m	Renal events	Allele frequencies			Genotype frequencies				Additive model	Recessive model			Dominant model		
			M	m	P	MM	Mm	mm	P*	P*	MM/Mm	mm	P*	MM	Mm/mm	P*
rs1065489 c.2808G>T	G/T	Yes	0.77	0.23	**0.025**	0.612	0.316	0.072	0.028	0.02	0.928	0.072	0.036	0.612	0.388	0.096
		No	0.836	0.164		0.699	0.274	0.027			0.973	0.027		0.699	0.301	

*A Bonferroni-corrected threshold of P = 0.0125 was considered as significant, while 0.0125 < P < 0.05 was considered suggestive of evidence for a potential association (in italics). P values in bold are statistically significant.

model, 7 of 36 Asp/Asp homozygotes (19.4%) vs. 105 of 1,122 Glu/Glu+Glu/Asp patients (9.4%) developed major cardiovascular events [Cox univariable analysis: HR 2.43, 95% CI (1.13–5.22), P = 0.023, **Table 3**, Kaplan–Meier curve is shown in **Figure 2B**]. At univariable and multivariable analyses, cardiovascular events were predicted by age, body mass index, hypertension duration, baseline UAE, HbA1c, and low-density lipoprotein cholesterol and by the p.Glu936Asp genotype (recessive model, **Tables 3** and **5**). With the additive model, no significant association was found between the p.Glu936Asp genotype and cardiovascular events either at univariable or multivariable analyses (**Table 3**, **Supplementary Figure 1B**, and **Supplementary Table 3**).

Among Asp/Asp homozygotes, major cardiovascular events were observed in 5 of 21 of those on ACEi (23.8%) and in 2 of 15 (13.3%) of those on non-ACEi [HR 1.75, 95% CI (0.34–9.03), **Supplementary Table 2** and **Figure 3B**], whereas among Glu/Glu+Glu/Asp patients, 42 of 553 (7.6%) of those on ACEi and 63 of 569 (11.1%) of those on non-ACEi [HR 0.67, 95% CI (0.46–0.99), P = 0.046, **Supplementary Table 2**, **Figure 1**, and **Figure 3B**] developed major cardiovascular events.

Adjustment for baseline covariates indicated that ACEi had no protective effect on cardiovascular risk in Asp/Asp patients [adjusted HR 1.11, 95% CI (0.21–5.82), **Table 7**], whereas it tended to reduce the risk of events [adjusted HR 0.73, 95% CI (0.49–1.08)] in the Glu/Glu+Glu/Asp group (**Table 7**), an effect that, however, failed to achieve statistical significance.

Among ACEi-treated participants, the risk of developing cardiovascular events was significantly higher in Asp/Asp homozygotes than in Glu/Glu+Glu/Asp subjects [**Table 7**, adjusted HR 3.26, 95% CI (1.29–8.28), P = 0.013], while in the non-ACEi group, no significant difference was found in the risk of cardiovascular events between Asp/Asp and Glu/Glu+Glu/Asp patients (**Table 7**).

The adjusted HR for developing an event increased progressively from Glu/Glu+Glu/Asp patients on ACEi (reference group), to Glu/Glu+Glu/Asp patients on non-ACEi [HR 1.38, 95% CI (0.93–2.05)], to Asp/Asp homozygotes on non-ACEi [HR 2.93, 95% CI (0.70–12.31)], and to Asp/Asp homozygotes on ACEi [HR 3.26, 95% CI (1.29–8.28), P = 0.013] (**Figure 4**).

Adjustments for Blood Pressure and Metabolic Control

The relationships between the p.Glu936Asp genotype and microalbuminuria or cardiovascular events did not change appreciably when analyses were adjusted for SBP or DBP or HbA1C levels, either at baseline or mean levels during the study period (**Tables 8** and **9**).

DISCUSSION

In a large cohort of normoalbuminuric type 2 diabetics prospectively followed in the context of a randomized clinical trial (Ruggenenti et al., 2004), we found evidence indicating that carriers of the Asp/Asp genotype of the p.Glu936Asp CFH polymorphism were at a higher risk of progression to microalbuminuria than carriers of one or two wild-type Glu alleles. Surprisingly, the

TABLE 3 | Univariable Cox analyses for microalbuminuria and cardiovascular events.

	New-onset microalbuminuria			Cardiovascular events		
	Hazard ratio	95% CI	*P* value	Hazard ratio	95% CI	*P* value
p.Glu936Asp (*Rec model*)°	2.465	1.143–5.319	**0.0214**	2.425	1.128–5.215	**0.0233**
p.Glu936Asp (*Addit model*)°	1.417	1.016–1.977	*0.0401*	1.227	0.882–1.706	0.2251
Age (years)	1.014	0.989–1.040	0.2801	1.052	1.027–1.078	**<0.0001**
Gender (male)	2.339	1.494–3.663	**0.0002**	1.571	1.067–2.312	**0.0221**
BMI	1.023	0.983–1.064	0.2699	0.944	0.903–0.986	**0.0097**
Diabetes duration (years)	1.007	0.977–1.038	0.6393	1.013	0.985–1.041	0.3726
Hypertension duration (years)	0.996	0.967–1.027	0.8180	1.024	1.001–1.048	**0.0451**
Smoking habit	2.045	1.367–3.061	**0.0005**	1.139	0.785–1.652	0.4935
HbA1c*	6.788	2.936–15.696	**<0.0001**	2.828	1.271–6.293	**0.0108**
Glucose (mg/dl)	1.006	1.002–1.010	**0.0013**	1.004	1.000–1.007	**0.0488**
SBP (mmHg)	1.001	0.987–1.015	0.9004	1.000	0.987–1.013	0.9568
DBP (mmHg)	0.998	0.962–1.015	0.3873	0.979	0.955–1.004	0.0965
MAP (mmHg)	0.994	0.971–1.019	0.994	0.988	0.966–1.011	0.3135
Serum creatinine (mg/dl)	1.736	0.500–6.023	0.3851	4.286	1.390–13.214	**0.0113**
Triglycerides (mg/dl)*	0.978	0.654–1.463	0.9147	1.186	0.815–1.726	0.3740
Total cholesterol (mg/dl)	0.996	0.990–1.002	0.1530	1.004	0.999–1.009	0.1580
HDL cholesterol (mg/dl)*	0.865	0.406–1.845	0.7073	0.984	0.967–1.000	0.0558
LDL cholesterol (mg/dl)	0.995	0.989–1.001	0.0988	1.005	1.000–1.010	**0.0363**
UAE (µg/min)*	9.870	6.451–15.101	**<0.0001**	1.879	1.396–2.529	**<0.0001**
ACEi therapy	0.461	0.303–0.702	**0.0003**	0.722	0.496–1.050	0.0884
New onset micro	–	–	–	1.854	1.106–3.106	**0.0191**

°The Bonferroni-corrected threshold was P = 0.025. In bold are shown P value significant after Bonferroni correction; P values between 0.025 and 0.05 are shown in italics.
*Log transformed.
Rec, recessive model; Addit, additive model; MAP, mean arterial pressure; LDL, low-density lipoprotein; BMI, body mass index.

excess risk in Asp/Asp homozygotes tended to increase with ACEi therapy that exerted its expected renal protective effect only in Glu/Glu+Glu/Asp diabetic patients. Consequently, ACEi-treated Asp/Asp homozygotes were the patients at highest risk of new-onset microalbuminuria, whereas Glu/Glu+Glu/Asp patients on ACEi were those at the lowest risk. Consistently, ACEi therapy failed to prevent the progressive increase in albuminuria from baseline to the end of the study in Asp/Asp homozygotes, and its anti-albuminuric effect was restricted to carriers of one or two wild-type Glu alleles. These results may be taken to indicate that among normoalbuminuric type 2 diabetics, those with the CFH Asp/Asp genotype are at higher risk of renal involvement, and in this subset, the interaction between genotype, treatment, and outcome results in lack of response to the anti-proteinuric action of ACEi therapy.

Similar findings were observed when cardiovascular events were considered as outcomes. Again, the Asp/Asp genotype appeared to be associated with excess risk of cardiovascular events and poor responsiveness to ACEi. Thus, we find evidence indicating an incremental risk of cardiovascular events from ACEi-treated Glu/Glu+Glu/Asp diabetics to ACEi-treated Asp/Asp homozygotes, who were the patients at the highest risk of events.

Findings that risk of microalbuminuria and cardiovascular events, as well as the protective effect of ACEi against these events, were similarly affected by the underlying *CFH* genotype, are in harmony with consolidated evidence that renal and cardiovascular outcomes in subjects at risk are strongly associated (de Zeeuw et al., 2006). Consistently, both higher albuminuria at baseline and progression to microalbuminuria on follow-up significantly predicted cardiovascular events. Finding that outcome data did not appreciably change when analyses were adjusted for baseline and mean follow-up HbA1c and SBP

and DBP values makes it possible to reasonably exclude major confounding effects of these concomitant risk factors.

In addition to modulating the alternative complement pathway in the fluid phase, CFH may bind specific sites on the renal microvascular endothelium and the glomerular capillary wall, to serve as a fixed complement regulator (Abrera-Abeleda et al., 2006). Thus, in experimental animals and humans, genetically determined CFH dysfunction results in uncontrolled alternative complement pathway activation, leading to complement-mediated renal endothelial damage (Pickering et al., 2007; Noris and Remuzzi, 2009). Consistently, the common *CFH*-H3 haplotype, including the T variant of the c.2808G>T (p.Glu936Asp) SNP, predisposes to aHUS, a rare disease characterized by complement-mediated glomerular endothelial injury (Caprioli et al., 2003; Pickering et al., 2007).

CFH is composed of 20 short consensus repeats (SCR) (Zipfel et al., 2006). The N-terminal SCRs 1–4 display complement regulatory activity, while the C-terminal SCRs 19–20 contain the recognition domain for cell surface proteoglycans and surface-bound C3 activation fragments (Zipfel et al., 2006). Once CFH C-terminus interacts with surface-bound C3b, CFH bends back on itself so that its N-terminus recognizes C3b and exerts its regulatory activity. The central SCRs 5–18, which connect N-terminal and C-terminal C3b-binding sites, are crucial for determining the flexibility required by CFH to achieve the folded bent-back structure and simultaneously occupy both sites on C3b (Morgan et al., 2011). Interestingly, the c.2808G>T SNP determines the glutammic (Glu) to aspartic (Asp) 936 amino-acidic change in SCR16 that is located 5 amino acids away from cysteine 931, which is crucial for S-S bridge formation and proper CFH folding (Reid and Day, 1989). Finding mutations affecting few amino
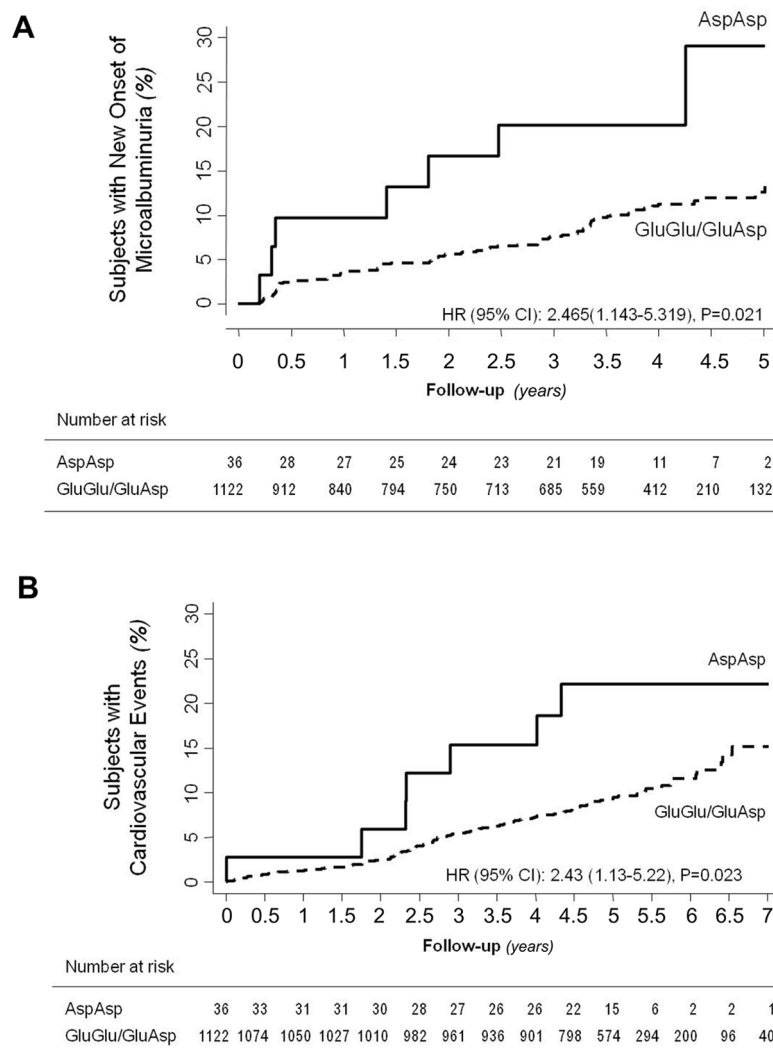
FIGURE 2 | Impact of p.Glu936Asp CFH polymorphism on new-onset microalbuminuria and cardiovascular events. Kaplan–Meier curves show the fraction of Asp/Asp homozygotes or Glu/Glu+Glu/Asp diabetics who progressed to microalbuminuria (panel **A**) or developed cardiovascular events (panel **B**) throughout the study period. *P* values and HR (95% CI) of unadjusted Cox analyses are shown.

acids away from the p.Glu936Asp polymorphism, in patients with genetically determined disease of the endothelium, (Bresin et al., 2013) provides additional evidence that a functional site located in SCR16 may be crucial to allowing CFH to acquire the suitable conformation to prevent complement activation on host cells. Moreover, the T variant of the c.2808G>T (p.Glu936Asp) SNP tags the CFH H3 haplotype, which was associated with lower CFH levels in subjects carrying two H3 copies as compared with subjects with zero copies (Bernabeu-Herrero et al., 2015; Pouw et al., 2018).

Altogether, the previously discussed observations would suggest that in Asp/Asp homozygous diabetic patients, conformational modifications and/or lower plasma concentration of CFH impair its protective effect on endothelial cells against the attack of complement components flowing in the blood, leading to endothelial dysfunction (Heinen et al., 2007). Endothelial dysfunction in renal microvasculature is a key factor in the early phase of diabetic nephropathy and may contribute to initial hyperfiltration and

decreased permselectivity due to neo-angiogenesis and alterations in endothelial glycocalix (Nakagawa et al., 2011). It might also induce tubulo-interstitial injury in diabetes through plasma leaking from injured peritubular capillaries, with a secondary inflammatory tubulo-interstitial response (Temm and Dominguez, 2007). Complement activation products associated with diabetes (Ostergaard et al., 2005; Woroniecka et al., 2011; Fujita et al., 2013) could play a role in increased renal microvascular permeability, as suggested by the ability of C5a to cause endothelial cell retraction, gap formation, and increased fluorescein isothiocyanate–dextran passage through the cell monolayer (Schraufstatter et al., 2002; Khan et al., 2013). The circulating sC5b-9 complex also promotes vascular leakage of proteins and fluids (Bossi et al., 2004).

Thus, we suggest that the Asp936 CFH variant, in the context of the complement hyperactivation status associated with the diabetic milieu, regulates complement alternative pathway on the surface of endothelial cells less efficiently than the wild-type

TABLE 4 | Clinical characteristics at randomization of patients with type 2 diabetes of BENEDICT phase A according to p.Glu936Asp CFH polymorphism.

	Asp/Asp homozygotes	Glu/Glu+Glu/Asp
n	36	1,122
Age (years)	63.4 (9.2)	62.2 (8.1)
Sex (male)	22 (61.1%)	585 (52.1%)
BMI	28.3 (26.1–32.5)	28.4 (25.8–31.6)
Diabetes duration (years)	7 (3–11.5)	6 (3–11)
Hypertension duration (years)	0.5 (0–7)	2 (0–8)
Smokers (current/former)	20 (55.6%)	463 (41.3%)
HbA1c (%)	5.8 (5.2–7.4)	5.6 (4.8–6.5)*
Glucose (mg/dl)	162.5 (40.1)	161.4 (47.1)
SBP (mmHg)	149.0 (14.5)	150.9 (14.1)
DBP (mmHg)	85.9 (7.6)	87.6 (7.6)
MAP (mmHg)	106.9 (8.7)	108.7 (8.3)
Serum creatinine (mg/dl)	0.92 (0.17)	0.91 (0.16)
Triglycerides (mg/dl)	137 (92–202)	125 (92–179)
Total cholesterol (mg/dl)	209.2 (22.7)	209.7 (37.0)
HDL cholesterol (mg/dl)	43 (36–57)	45 (38–54)
LDL cholesterol (mg/dl)	163.6 (25.2)	162.8 (36.1)
UAE (μg/min)	5.0 (3.6–9.2)	5.3 (3.6–9.3)
ACEi at randomization (yes)	21 (58.3%)	553 (49.3%)

Continuous variables are expressed as mean and SD and compared using unpaired t test or expressed as median and IQR and compared using Mann–Whitney test. Categorical variables are expressed in percentage and compared using Chi-square test or Fisher's exact test as appropriate.

*P < 0.05 vs Asp/Asp homozygotes.

HbA1c, glycosylated hemoglobin; SBP, systolic blood pressure; DBP, diastolic blood pressure; UAE, urinary albumin excretion; MAP, mean arterial pressure; LDL, low-density lipoprotein; BMI, body mass index.

TABLE 5 | Multivariable Cox analysis for microalbuminuria and cardiovascular endpoints.

	Microalbuminuria number of events = 98		Cardiovascular events Number of events = 112	
	Hazard ratio (95% CI)	P value	Hazard ratio (95% CI)	P value
p.Glu936Asp (*Rec model*)	3.254 (1.462–7.241)	**0.0038**	2.683 (1.227–5.868)	**0.0134**
ACEi therapy	0.415 (0.270–0.639)	**<0.0001**	0.716 (0.485–1.055)	0.0908
Gender (male)	1.532 (0.913–2.569)	0.1061	1.610 (0.982–2.638)	0.0589
Smoking habits	1.401 (0.881–2.226)	0.1541	–	–
HbA1c[†]	5.170 (2.207–12.113)	**0.0002**	2.892 (1.231–6.797)	**0.0149**
UAE[†]	9.415 (6.042–14.791)	**<0.0001**	1.742 (1.262–2.404)	**0.0007**
Age	–	–	1.047 (1.020–1.075)	**0.0006**
Hypertension duration	–	–	1.026 (1.003–1.050)	**0.0250**
BMI	–	–	0.940 (0.895–0.986)	**0.0117**
Serum Creatinine	–	–	2.220 (0.595–8.287)	0.2352
LDL cholesterol	–	–	1.007 (1.002–1.012)	**0.0112**

[†]log transformed. Rec, recessive model; LDL, low-density lipoprotein; BMI, body mass index. P values in bold are statistically significant.

Glu936, which might result in renal endothelial dysfunction, albuminuria, and structural kidney damage.

Moreover, complement activation products might contribute directly to diabetic cardiovascular complications (Bjerre et al., 2008; Hertle et al., 2014) by promoting endothelial dysfunction and thrombotic events. In particular, C3a and C5a, and C5b-9, may induce endothelial cell detachment with exposure of subendothelium and secondary platelet aggregation and switch endothelial cells to a procoagulant phenotype due to expression of tissue factor (Tedesco et al., 1997; Schraufstatter et al., 2002). The excess of cardiac complications reported in patients with genetically determined endothelial disease associated with *CFH* mutations and chronic alternative pathway dysregulation (Noris et al., 2010) provides evidence that CFH dysfunction and consequent alternative pathway activation may play a major role in the pathogenesis of cardiovascular events in diabetics.

Why ACEi therapy had no protective effect against microalbuminuria or cardiovascular events in Asp/Asp homozygotes is a matter of speculation. It is well known that angiotensin II regulates the secretion of renin through a homeostatic mechanism-defined "feedback loop." In animal models, the blockade of angiotensin II by ACEi causes an immediate and large increase in plasma renin concentration (Chen et al., 2010). A similar renin increase is documented in humans during treatment with ACEi (Johnston et al., 1979; Seifarth et al., 2002). Recently, Békássy and colleagues by *in vitro* studies provided evidence that renin triggers complement activation by cleaving C3 and generating functional C3b and C3a (Bekassy et al., 2018). Administration of a renin inhibitor in three patients with dense deposit disease (DDD), a rare kidney disease characterized by complement hyperactivation, decreased plasma C3a and C5a levels and complement deposition in the renal biopsy (Bekassy et al., 2018).

Based on these observations, we would speculate that in Asp/Asp homozygous diabetic patients treated with ACEi, an increase in renin plasma levels combined with defective CFH-mediated complement regulation could predispose to complement activation locally in the kidney, leading to the onset and progression of renal complications.

ACEi, by blocking the main plasma enzyme degrading bradykinin, may result in abnormally elevated levels of this potent vascular permeability factor, which occasionally associates with acute plasma extravasation and angioedema (Schmidt et al., 2010; Tomita et al., 2012). Finding that the permeabilizing activity of sC5b-9 *in vitro* and *in vivo* is inhibited by a bradykinin 2 receptor antagonist suggests a cross talk between the complement and the kinin system in inducing vascular permeability (Bossi et al., 2004).

A plausible explanation of our findings could be that in Asp/Asp diabetics, ineffective modulation of the complement alternative pathway combined with excess renin and bradykinin levels associated with ACEi treatment might enhance complement activation and microvascular permeability and offset the anti-proteinuric and cardioprotective effects of RAS blockade.

Limitations and Strengths

This was a post hoc analysis of a study originally designed for other purposes. The number of cardiovascular events was relatively small, which limited the statistical power to evaluate

TABLE 6 | Multivariable Cox analysis with genotype–ACEi treatment interaction for microalbuminuria and cardiovascular endpoints (*without other covariates, reference: ACEi yes*).

	New onset microalbuminuria Number of events = 98		Cardiovascular events Number of events = 112	
	Hazard ratio (95% CI)	*P* value	Hazard ratio (95% CI)	*P* value
p.Glu936Asp°	6.210 (2.563–15.044)	**<0.0001**	3.655 (1.445–9.242))	**0.006**
ACEi therapy	2.584 (1.648–4.052)	**<0.0001**	1.489 (1.007–2.200)	**0.046**
p.Glu936Asp*ACEi therapy	0.096 (0.011–0.837)	**0.034**	0.384 (0.071–2.073)	0.266

°Recessive model. P values in bold are statistically significant.

FIGURE 3 | Impact of p.Glu936Asp CFH polymorphism and ACEi therapy on new-onset microalbuminuria and cardiovascular events. Kaplan–Meier curves show the fraction of Asp/Asp homozygous or Glu/Glu+Glu/Asp diabetic patients with or without ACEi therapy who progressed to microalbuminuria (panel **A**) or developed cardiovascular events (panel **B**) throughout the study period. P values and HR (95% CI) of unadjusted Cox analyses are shown.
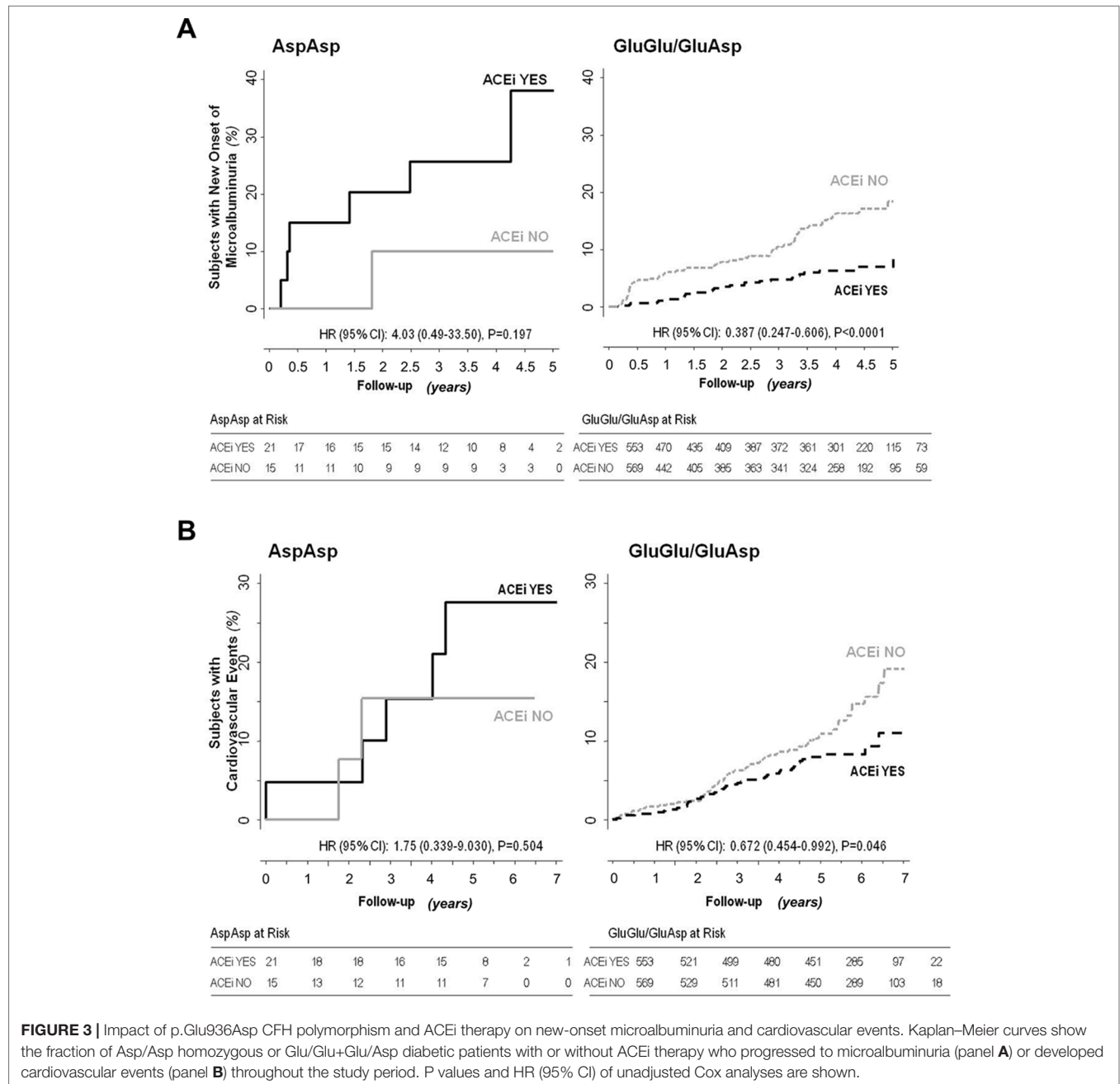
TABLE 7 | Panel A: HRs of the comparisons between ACEi-treated and non-ACEi-treated patients within the two genotype groups. Panel B: HRs of the comparisons between Asp/Asp homozygotes and Glu/Glu+Glu/Asp patients in ACEi or non-ACEi arms (genotype ACEi use interaction, with other covariates).

A	ACEi vs. non-ACEi	
	New-onset microalbuminuria	Cardiovascular events
Asp/Asp homozygotes	HR = 1.543, P = 0.691 95% CI (0.182–13.072)	HR = 1.112, P = 0.900 95% CI (0.212–5.821)
Glu/Glu+Glu/Asp	HR = 0.381, **P < 0.0001** 95% CI (0.241–0.601)	HR = 0.726, P = 0.113 95% CI (0.488–1.079)

B	Asp/Asp homozygotes vs. Glu/Glu+Glu/Asp	
	New onset microalbuminuria	Cardiovascular events
non-ACEi	HR = 1.164, P = 0.882 95% CI (0.157–8.627)	HR = 2.128, P = 0.298 95% CI (0.513–8.831)
ACEi	HR = 4.717, **P = 0.001** 95% CI (1.931–11.519)	HR = 3.262, **P = 0.013** 95% CI (1.285–8.282)

P values in bold are statistically significant.

FIGURE 5 | Changes in urinary albumin excretion (UAE) rate, according to p.Glu936Asp polymorphism and ACEi therapy. Percent changes of UAE at last visit vs. baseline. On the left, boxplots on the entire data range. On the right, boxplots were zoomed to show median (the central line) and IQR (lower and upper hinges).

possible interactions of baseline covariates and treatment with cardiovascular outcomes. Replication studies are needed to confirm the association between the p.Glu936Asp CFH polymorphism and the risk of renal and cardiovascular complications. This study was focused on the analysis of a single SNP selected on the basis of the hypothesis that the p.Glu936Asp CFH polymorphism that tags the H3 haplotype and has been found in association with a rare kidney disease with microvascular involvement may play a predisposing role in microvascular complications of diabetes. On the other hand, we recognize that a genome-wide analysis would address the multiple testing approach and would allow to assess population stratification, possible cryptic relatedness, and to evaluate pathways.

Major strengths were standardized study conduction, centralized measurement of all considered laboratory variables, including microalbuminuria—which was assessed by gold
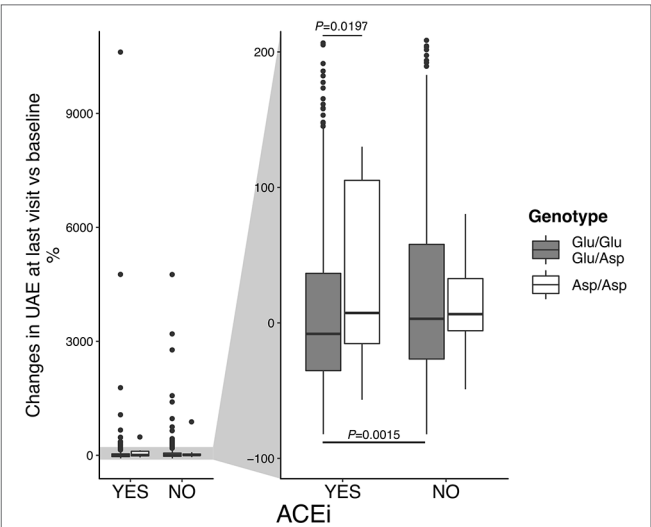
standard procedures in triplicate overnight urine collections (Ruggenenti et al., 2004)—and the centralized adjudication of cardiovascular events by two cardiologists who were blinded to treatment allocation. Study findings are widely generalizable because more than 96% of the BENEDICT population was genotyped and data were observed in type 2 diabetics with normoalbuminuria and hypertension, a typology of patients that accounts for at least 90% of the whole diabetic population.

Finally, at variance with most previous similar genetic studies that included also patients with micro- or macroalbuminuria who were largely driving the study findings, or did not measure albuminuria levels, our study investigated the impact on renal outcome of the p.Glu936Asp CFH polymorphism in a pure
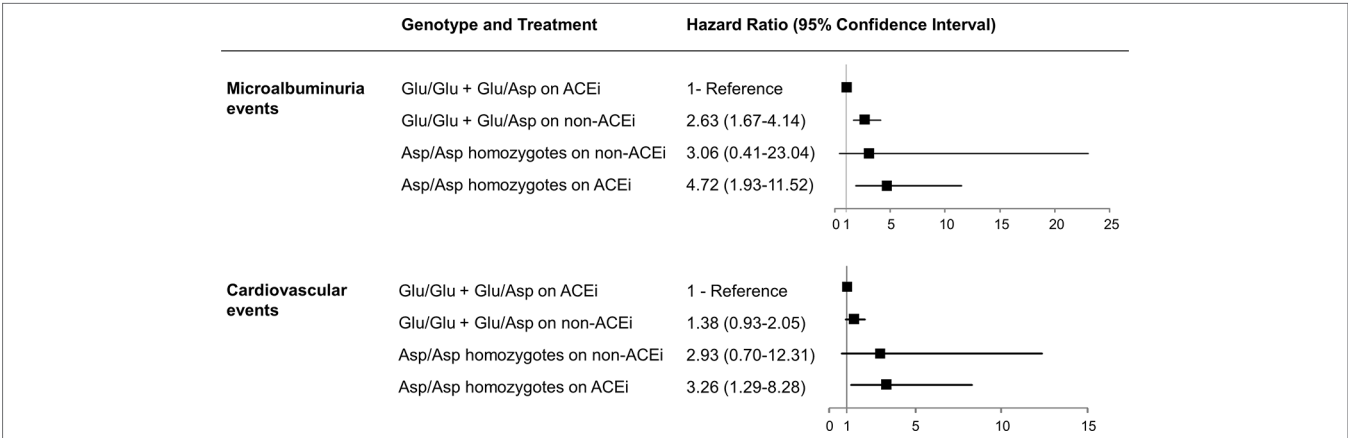

FIGURE 4 | Hazard ratios for considered events according to p.Glu936Asp genotype and ACEi treatment. Adjusted hazard ratios (95% confidence intervals) for microalbuminuria and cardiovascular events according to p.Glu936Asp genotype and ACEi therapy compared with Glu/Glu+Glu/Asp patients on ACEi therapy taken as the reference group are shown.

TABLE 8 | Multivariable Cox analysis for microalbuminuria based on 98 events.

	HbA1c fup + MAP fup		HbA1c fup + DBP fup		HbA1c fup + SBP fup	
	Hazard ratio (95% CI)	P value	Hazard ratio (95% CI)	P value	Hazard ratio (95% CI)	P value
p.Glu936Asp*	3.518 (1.568–7.893)	**0.0023**	3.702 (1.654–8.285)	**0.0014**	3.532 (1.572–7.932)	**0.0022**
ACEi therapy	0.483 (0.311–0.750)	**0.0012**	0.457 (0.295–0.707)	**0.0004**	0.491 (0.316–0.761)	**0.0015**
Gender (male)	1.584 (0.921–2.722)	0.0961	1.628 (0.943–2.811)	0.0804	1.689 (0.982–2.905)	0.0583
Smoking habits	1.439 (0.894–2.316)	0.1336	1.401 (0.872–2.250)	0.1635	1.433 (0.888–2.313)	0.0583
UAE baseline[†]	9.654 (6.085–15.315)	**<0.0001**	9.947 (6.291–15.729)	**<0.0001**	9.433 (5.922–15.027)	**<0.0001**
HbA1c fup	1.398 (1.192–1.639)	**<0.0001**	1.409 (1.204–1.649)	**<0.0001**	1.404 (1.197–1.648)	**<0.0001**
MAP fup	1.040 (1.004–1.077)	**0.0287**	–	–	–	–
DBP fup	–	–	1.016 (0.976–1.058)	0.4365	–	–
SBP fup	–	–	–	–	1.030 (1.010–1.050)	**0.0030**

*fup, mean value during follow up; MAP, mean arterial pressure. *Recessive model. [†]Log transformed. P values in bold are statistically significant.*

TABLE 9 | Multivariable Cox analysis for cardiovascular endpoints based on 112 CV events.

	Hba1c fup + MAP fup		Hba1c fup + DBP fup		Hba1c fup + SBP fup	
	Hazard ratio (95% CI)	P value	Hazard ratio (95% CI)	P value	Hazard ratio (95% CI)	P value
p.Glu936Asp*	2.406 (1.030–5.618)	**0.0425**	2.412 (1.036-5.618)	**0.0412**	2.441 (1.046-5.696)	**0.039**
ACEi therapy	0.887 (0.580–1.356)	0.5789	0.863 (0.564–1.320)	0.4973	0.853 (0.562–1.297)	0.458
Gender (male)	1.621 (0.932–2.819)	0.0869	1.637 (0.938–2.857)	0.0826	1.700 (0.983–2.943)	0.058
HbA1c follow up	1.090 (0.912–1.301)	0.3436	1.091 (0.914–1.303)	0.335	1.093 (0.916–1.304)	0.325
UAE baseline[†]	1.578 (1.109–2.247)	**0.0113**	1.601 (1.126–2.276)	**0.0087**	1.549 (1.086–2.208)	**0.0156**
Age	1.048 (1.109–1.078)	**0.0012**	1.055 (1.025–1.086)	**0.0002**	1.039 (1.008–1.070)	**0.0121**
Hypertension duration	1.007 (0.979–1.036)	0.6378	1.011 (0.0983–1.039)	0.454	1.007 (0.980–1.036)	0.615
BMI	0.924 (0.876–0.975)	**0.0039**	0.924 (0.875–0.975)	**0.004**	0.928 (0.880–0.978)	**0.0054**
Serum creatinine	1.760 (0.400–7.740)	0.454	1.680 (0.378–7.466)	0.4953	1.884 (0.431–8.234)	0.4
LDL cholesterol	1.008 (1.003–1.014)	**0.0032**	1.009 (1.003–1.014)	**0.0027**	1.009 (1.003–1.048)	**0.0026**
MAP follow-up	1.045 (1.010–1.081)	**0.0116**	–	–	–	–
DBP follow-up	–	–	1.040 (1.002–1.081)	**0.0415**	–	–
SBP follow-up	–	–	–	–	1.027 (1.005–1.048)	**0.0142**

*MAP, mean arterial; BMI, body mass index; LDL, low-density lipoprotein. pressure. *recessive model. [†]Log transformed. P values in bold are statistically significant.*

population of patients with no evidence of renal involvement at baseline (Parving et al., 2008; Bacci et al., 2011; Mooyaart et al., 2011; Ahlqvist et al., 2015).

Conclusions

Screening for the p.Glu936Asp polymorphism may help identify subjects at increased risk of microalbuminuria and cardiovascular events and those who might not benefit from ACEi therapy. Notably, observational evidence that ACEi therapy was associated with an excess risk of events in patients with the Asp/Asp CFH genotype suggests that renal and cardiac effects of ACEi in diabetes may be influenced by genetic variations and highlights the urgent need for genome-wide studies to primarily address this critical issue in larger number of patients.

The present findings should not be taken as a reason for not offering ACEi therapy to type 2 diabetic patients who in general may benefit of nephro- and cardio-protective effects of this class of drugs. Instead, our data indicate that genotyping for the c.2808G>T (p.Glu936Asp) polymorphism could help identifying a subpopulation of patients, accounting for approximately 3% of type 2 diabetics, who are less likely to benefit from these drugs to avoid unnecessary exposure to potentially serious, treatment-related side effects.

DATA AVAILABILITY

The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

ETHICS STATEMENT

The study was approved by the Ethics Committee of Azienda Sanitaria Locale, Bergamo, Italy. All study participants provided written informed consent according to the Helsinki Declaration guidelines. Data were handled with respect for patient confidentiality and anonymity.

AUTHOR CONTRIBUTIONS

EV, MN, GR, and PR designed research, interpreted data, and wrote the paper; EV and ER performed the research and analyzed the data; AP and MB performed statistical analyses; GG organized blood and informed consent collection; API, IPI, AB, RT, and AD participated to BENEDICT study and provided detailed clinical information of patients; SF and NS

performed laboratory measurements; AB analyzed the data and critically revised the manuscript. GR was the principal investigator of BENEDICT study. PR was the study coordinator of the BENEDICT study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.00681/full#supplementary-material

REFERENCES

Abrera-Abeleda, M. A., Nishimura, C., Smith, J. L., Sethi, S., McRae, J. L., Murphy, B. F., et al. (2006). Variations in the complement regulatory genes factor H (CFH) and factor H related 5 (CFHR5) are associated with membranoproliferative glomerulonephritis type II (dense deposit disease). *J. Med. Genet.* 43 (7), 582–589. doi: 10.1136/jmg.2005.038315

Ahlqvist, E., van Zuydam, N. R., Groop, L. C., and McCarthy, M. I. (2015). The genetics of diabetic complications. *Nat. Rev. Nephrol.* 11 (5), 277–287. doi: 10.1038/nrneph.2015.37

Bacci, S., Rizza, S., Prudente, S., Spoto, B., Powers, C., Facciorusso, A., et al. (2011). The ENPP1 Q121 variant predicts major cardiovascular events in high-risk individuals: evidence for interaction with obesity in diabetic patients. *Diabetes* 60 (3), 1000–1007. doi: 10.2337/db10-1300

Bekassy, Z. D., Kristoffersson, A. C., Rebetz, J., Tati, R., Olin, A. I., and Karpman, D. (2018). Aliskiren inhibits renin-mediated complement activation. *Kidney Int.* 94 (4), 689–700. doi: 10.1016/j.kint.2018.04.004

Bernabeu-Herrero, M. E., Jimenez-Alcazar, M., Anter, J., Pinto, S., Sanchez Chinchilla, D., Garrido, S., et al. (2015). Complement factor H, FHR-3 and FHR-1 variants associate in an extended haplotype conferring increased risk of atypical hemolytic uremic syndrome. *Mol. Immunol.* 67 (2 Pt B), 276–286. doi: 10.1016/j.molimm.2015.06.021

Bjerre, M., Hansen, T. K., and Flyvbjerg, A. (2008). Complement activation and cardiovascular disease. *Horm. Metab. Res.* 40 (9), 626–634. doi: 10.1055/s-0028-1083786

Bonomo, J. A., Palmer, N. D., Hicks, P. J., Lea, J. P., Okusa, M. D., Langefeld, C. D., et al. (2014). Complement factor H gene associations with end-stage kidney disease in African Americans. *Nephrol. Dial. Transplant.* 29 (7), 1409–1414. doi: 10.1093/ndt/gfu036

Bossi, F., Fischetti, F., Pellis, V., Bulla, R., Ferrero, E., Mollnes, T. E., et al. (2004). Platelet-activating factor and kinin-dependent vascular leakage as a novel functional activity of the soluble terminal complement complex. *J. Immunol.* 173 (11), 6921–6927. doi: 10.4049/jimmunol.173.11.6921

Bresin, E., Rurali, E., Caprioli, J., Sanchez-Corral, P., Fremeaux-Bacchi, V., Rodriguez de Cordoba, S., et al. (2013). Combined complement gene mutations in atypical hemolytic uremic syndrome influence clinical phenotype. *J. Am. Soc. Nephrol.* 24 (3), 475–486. doi: 10.1681/ASN.2012090884

Caprioli, J., Castelletti, F., Bucchioni, S., Bettinaglio, P., Bresin, E., Pianetti, G., et al. (2003). Complement factor H mutations and gene polymorphisms in haemolytic uraemic syndrome: the C-257T, the A2089G and the G2881T polymorphisms are strongly associated with the disease. *Hum. Mol. Genet.* 12 (24), 3385–3395. doi: 10.1093/hmg/ddg363

Chen, L., Kim, S. M., Eisner, C., Oppermann, M., Huang, Y., Mizel, D., et al. (2010). Stimulation of renin secretion by angiotensin II blockade is Gsalpha-dependent. *J. Am. Soc. Nephrol.* 21 (6), 986–992. doi: 10.1681/ASN.2009030307

Davila, S., Wright, V. J., Khor, C. C., Sim, K. S., Binder, A., Breunis, W. B., et al. (2010). Genome-wide association study identifies variants in the CFH region associated with host susceptibility to meningococcal disease. *Nat. Genet.* 42 (9), 772–776. doi: 10.1038/ng.640

de Zeeuw, D., Parving, H. H., and Henning, R. H. (2006). Microalbuminuria as an early marker for cardiovascular disease. *J. Am. Soc. Nephrol.* 17 (8), 2100–2105. doi: 10.1681/ASN.2006050517

Figueredo, A., Ibarra, J. L., Bagazgoitia, J., Rodriguez, A., Molino, A. M., Fernandez-Cruz, A., et al. (1993). Plasma C3d levels and ischemic heart disease in type II diabetes. *Diabetes Care* 16 (2), 445–449. doi: 10.2337/diacare.16.2.445

Fischetti, F., Candido, R., Toffoli, B., Durigutto, P., Bernardi, S., Carretta, R., et al. (2011). Innate immunity, through late complement components activation, contributes to the development of early vascular inflammation and morphologic alterations in experimental diabetes. *Atherosclerosis* 216 (1), 83–89. doi: 10.1016/j.atherosclerosis.2011.01.050

Fujita, T., Ohi, H., Komatsu, K., Endo, M., Ohsawa, I., and Kanmatsuse, K. (1999). Complement activation accelerates glomerular injury in diabetic rats. *Nephron* 81 (2), 208–214. doi: 10.1159/000045278

Fujita, T., Hemmi, S., Kajiwara, M., Yabuki, M., Fuke, Y., Satomura, A., et al. (2013). Complement-mediated chronic inflammation is associated with diabetic microvascular complication. *Diabetes Metab. Res. Rev.* 29 (3), 220–226. doi: 10.1002/dmrr.2380

Garcia, M., Alvarez, L., Nogacka, A. M., Gonzalez-Iglesias, H., Escribano, J., Fernandez-Vega, B., et al. (2015). CFH polymorphisms in a Northern Spanish population with neovascular and dry forms of age-related macular degeneration. *Acta Ophthalmol.* 93 (8), e658–e666. doi: 10.1111/aos.12790

Ghosh, P., Sahoo, R., Vaidya, A., Chorev, M., and Halperin, J. A. (2015). Role of complement and complement regulatory proteins in the complications of diabetes. *Endocr. Rev.* 36 (3), 272–288. doi: 10.1210/er.2014-1099

Heinen, S., Jozsi, M., Hartmann, A., Noris, M., Remuzzi, G., Skerka, C., et al. (2007). Hemolytic uremic syndrome: a factor H mutation (E1172Stop) causes defective complement control at the surface of endothelial cells. *J. Am. Soc. Nephrol.* 18 (2), 506–514. doi: 10.1681/ASN.2006091069

Hertle, E., Stehouwer, C. D., and van Greevenbroek, M. M. (2014). The complement system in human cardiometabolic disease. *Mol. Immunol.* 61 (2), 135–148. doi: 10.1016/j.molimm.2014.06.031

Hertle, E., van Greevenbroek, M. M., Arts, I. C., van der Kallen, C. J., Geijselaers, S. L., Feskens, E. J., et al. (2014). Distinct associations of complement C3a and its precursor C3 with atherosclerosis and cardiovascular disease. The CODAM study. *Thromb. Haemost.* 111 (6), 1102–1111. doi: 10.1160/TH13-10-0831

Holtkamp, F. A., de Zeeuw, D., de Graeff, P. A., Laverman, G. D., Berl, T., Remuzzi, G., et al. (2011). Albuminuria and blood pressure, independent targets for cardioprotective therapy in patients with diabetes and nephropathy: a post hoc analysis of the combined RENAAL and IDNT trials. *Eur. Heart J.* 32 (12), 1493–1499. doi: 10.1093/eurheartj/ehr017

Johnston, C. I., Millar, J. A., McGrath, B. P., and Matthews, P. G. (1979). Long-term effects of captopril (SQ14 225) on blood-pressure and hormone levels in essential hypertension. *Lancet* 2 (8141), 493–496. doi: 10.1016/S0140-6736(79)91552-6

Jozsi, M., and Zipfel, P. F. (2008). Factor H family proteins and human diseases. *Trends Immunol.* 29 (8), 380–387. doi: 10.1016/j.it.2008.04.008

Khan, M. A., Maasch, C., Vater, A., Klussmann, S., Morser, J., Leung, L. L., et al. (2013). Targeting complement component 5a promotes vascular integrity and limits airway remodeling. *Proc. Natl. Acad. Sci. U. S. A.* 110 (15), 6061–6066. doi: 10.1073/pnas.1217991110

Li, X. Q., Chang, D. Y., Chen, M., and Zhao, M. H. (2018). Complement activation in patients with diabetic nephropathy. *Diabetes Metab.* 45 (3), 248–253. doi: 10.1016/j.diabet.2018.04.001

Merle, N. S., Church, S. E., Fremeaux-Bacchi, V., and Roumenina, L. T. (2015). Complement system Part I - molecular mechanisms of activation and regulation. *Front. Immunol.* 6, 262. doi: 10.3389/fimmu.2015.00262

Miki, A., Kondo, N., Yanagisawa, S., Bessho, H., Honda, S., and Negi, A. (2014). Common variants in the complement factor H gene confer genetic susceptibility to central serous chorioretinopathy. *Ophthalmology* 121 (5), 1067–1072. doi: 10.1016/j.ophtha.2013.11.020

Mooyaart, A. L., Valk, E. J., van Es, L. A., Bruijn, J. A., de Heer, E., Freedman, B. I., et al. (2011). Genetic associations in diabetic nephropathy: a meta-analysis. *Diabetologia* 54 (3), 544–553. doi: 10.1007/s00125-010-1996-1

Morgan, H. P., Schmidt, C. Q., Guariento, M., Blaum, B. S., Gillespie, D., Herbert, A. P., et al. (2011). Structural basis for engagement by complement factor H of C3b on a self surface. *Nat. Struct. Mol. Biol.* 18 (4), 463–470. doi: 10.1038/nsmb.2018

Nakagawa, T., Tanabe, K., Croker, B. P., Johnson, R. J., Grant, M. B., Kosugi, T., et al. (2011). Endothelial dysfunction as a potential contributor in diabetic nephropathy. *Nat. Rev. Nephrol.* 7 (1), 36–44. doi: 10.1038/nrneph.2010.152

Noris, M., and Remuzzi, G. (2009). Atypical hemolytic-uremic syndrome. *N. Engl. J. Med.* 361 (17), 1676–1687. doi: 10.1056/NEJMra0902814

Noris, M., Caprioli, J., Bresin, E., Mossali, C., Pianetti, G., Gamba, S., et al. (2010). Relative role of genetic complement abnormalities in sporadic and familial aHUS and their impact on clinical phenotype. *Clin. J. Am. Soc. Nephrol.* 5 (10), 1844–1859. doi: 10.2215/CJN.02210310

Onat, A., Can, G., Rezvani, R., and Cianflone, K. (2011). Complement C3 and cleavage products in cardiometabolic risk. *Clin. Chim. Acta* 412 (13-14), 1171–1179. doi: 10.1016/j.cca.2011.03.005

Ostergaard, J., Hansen, T. K., Thiel, S., and Flyvbjerg, A. (2005). Complement activation and diabetic vascular complications. *Clin. Chim. Acta* 361 (1–2), 10–19. doi: 10.1016/j.cccn.2005.04.028

Parving, H. H., de Zeeuw, D., Cooper, M. E., Remuzzi, G., Liu, N., Lunceford, J., et al. (2008). ACE gene polymorphism and losartan treatment in type 2 diabetic patients with nephropathy. *J. Am. Soc. Nephrol.* 19 (4), 771–779. doi: 10.1681/ASN.2007050582

Pickering, M. C., de Jorge, E. G., Martinez-Barricarte, R., Recalde, S., Garcia-Layana, A., Rose, K. L., et al. (2007). Spontaneous hemolytic uremic syndrome triggered by complement factor H lacking surface recognition domains. *J. Exp. Med.* 204 (6), 1249–1256. doi: 10.1084/jem.20070301

Pouw, R. B., Gomez Delgado, I., Lopez Lera, A., Rodriguez de Cordoba, S., Wouters, D., Kuijpers, T. W., et al. (2018). High Complement Factor H-Related (FHR)-3 Levels Are Associated With the Atypical Hemolytic-Uremic Syndrome-Risk Allele CFHR3*B. *Front. Immunol.* 9, 848. doi: 10.3389/fimmu.2018.00848

Reid, K. B., and Day, A. J. (1989). Structure-function relationships of the complement components. *Immunol. Today* 10 (6), 177–180. doi: 10.1016/0167-5699(89)90317-4

Ruggenenti, P., Fassi, A., Ilieva, A. P., Bruno, S., Iliev, I. P., Brusegan, V., et al. (2004). Preventing microalbuminuria in type 2 diabetes. *N. Engl. J. Med.* 351 (19), 1941–1951. doi: 10.1056/NEJMoa042167

Ruggenenti, P., Fassi, A., Ilieva, A., Iliev, I. P., Chiurchiu, C., Rubis, N., et al. (2011). Effects of verapamil added-on trandolapril therapy in hypertensive type 2 diabetes patients with microalbuminuria: the BENEDICT-B randomized trial. *J. Hypertens.* 29 (2), 207–216. doi: 10.1097/HJH.0b013e32834069bd

Ruggenenti, P., Porrini, E., Motterlini, N., Perna, A., Ilieva, A. P., Iliev, I. P., et al. (2012). Measurable urinary albumin predicts cardiovascular risk among normoalbuminuric patients with type 2 diabetes. *J. Am. Soc. Nephrol.* 23 (10), 1717–1724. doi: 10.1681/ASN.2012030252

Schmidt, P. W., Hirschl, M. M., and Trautinger, F. (2010). Treatment of angiotensin-converting enzyme inhibitor-related angioedema with the bradykinin B2

receptor antagonist icatibant. *J. Am. Acad. Dermatol.* 63 (5), 913–914. doi: 10.1016/j.jaad.2010.03.023

Schraufstatter, I. U., Trieu, K., Sikora, L., Sriramarao, P., and DiScipio, R. (2002). Complement c3a and c5a induce different signal transduction cascades in endothelial cells. *J. Immunol.* 169 (4), 2102–2110. doi: 10.4049/jimmunol.169.4.2102

Seifarth, C., Trenkel, S., Schobel, H., Hahn, E. G., and Hensen, J. (2002). Influence of antihypertensive medication on aldosterone and renin concentration in the differential diagnosis of essential hypertension and primary aldosteronism. *Clin. Endocrinol. (Oxf.)* 57 (4), 457–465. doi: 10.1046/j.1365-2265.2002.01613.x

Sun, Z. J., Li, X. Q., Chang, D. Y., Wang, S. X., Liu, G., Chen, M., et al. (2018). Complement deposition on renal histopathology of patients with diabetic nephropathy. *Diabetes Metab.* S1262-3636(18)30170-8. doi: 10.1016/j.diabet.2018.08.011

Tedesco, F., Pausa, M., Nardon, E., Introna, M., Mantovani, A., and Dobrina, A. (1997). The cytolytically inactive terminal complement complex activates endothelial cells to express adhesion molecules and tissue factor procoagulant activity. *J. Exp. Med.* 185 (9), 1619–1627. doi: 10.1084/jem.185.9.1619

Temm, C., and Dominguez, J. H. (2007). Microcirculation: nexus of comorbidities in diabetes. *Am. J. Physiol. Renal Physiol.* 293 (2), F486–F493. doi: 10.1152/ajprenal.00503.2006

The BENEDICT Group (2003). The BErgamo NEphrologic DIabetes Complications Trial (BENEDICT): design and baseline characteristics. *Control Clin. Trials* 24 (4), 442–461. doi: 10.1016/S0197-2456(03)00028-X

Tomita, H., Sanford, R. B., Smithies, O., and Kakoki, M. (2012). The kallikrein-kinin system in diabetic nephropathy. *Kidney Int.* 81 (8), 733–744. doi: 10.1038/ki.2011.499

Wang, Q. F., Huang, X. F., Zheng, Z. L., Dai, M. L., Cai, W. J., Yang, M. M., et al. (2016). Association of CD59 and CFH polymorphisms with acute anterior uveitis in Chinese population. *Eye (Lond.)* 30 (11), 1452–1457. doi: 10.1038/eye.2016.146

Woroniecka, K. I., Park, A. S., Mohtat, D., Thomas, D. B., Pullman, J. M., and Susztak, K. (2011). Transcriptome analysis of human diabetic kidney disease. *Diabetes* 60 (9), 2354–2369. doi: 10.2337/db10-1181

Zhang, D. F., Huang, X. Q., Wang, D., Li, Y. Y., and Yao, Y. G. (2013). Genetic variants of complement genes ficolin-2, mannose-binding lectin and complement factor H are associated with leprosy in Han Chinese from Southwest China. *Hum. Genet.* 132 (6), 629–640. doi: 10.1007/s00439-013-1273-8

Zipfel, P. F., Heinen, S., Jozsi, M., and Skerka, C. (2006). Complement and diseases: defective alternative pathway control results in kidney and eye diseases. *Mol. Immunol.* 43 (1–2), 97–106. doi: 10.1016/j.molimm.2005.06.015

Genome-Wide Study Updates in the International Genetics and Translational Research in Transplantation Network (iGeneTRAiN)

Claire E. Fishman[1], Maede Mohebnasab[1], Jessica van Setten[2], Francesca Zanoni[3], Chen Wang[3], Silvia Deaglio[4,5], Antonio Amoroso[4,5], Lauren Callans[1], Teun van Gelder[6], Sangho Lee[7], Krzysztof Kiryluk[3], Matthew B. Lanktree[8] and Brendan J. Keating[1]*, on behalf of the iGeneTRAiN consortium

[1] Division of Transplantation Department of Surgery, University of Pennsylvania, Philadelphia, PA, United States, [2] Department of Cardiology, University Medical Center Utrecht, University of Utrecht, Utrecht, Netherlands, [3] Department of Medicine, Division of Nephrology, Vagelos College of Physicians & Surgeons, Columbia University, New York, NY, United States, [4] Immunogenetics and Biology of Transplantation, Città della Salute e della Scienza, University Hospital of Turin, Turin, Italy, [5] Medical Genetics, Department of Medical Sciences, University Turin, Turin, Italy, [6] Department of Hospital Pharmacy, University Medical Center Rotterdam, Rotterdam, Netherlands, [7] Department of Nephrology, Khung Hee University, Seoul, South Korea, [8] Division of Nephrology, St. Joseph's Healthcare Hamilton, McMaster University, Hamilton, ON, Canada

The prevalence of end-stage renal disease (ESRD) and the number of kidney transplants performed continues to rise every year, straining the procurement of deceased and living kidney allografts and health systems. Genome-wide genotyping and sequencing of diseased populations have uncovered genetic contributors in substantial proportions of ESRD patients. A number of these discoveries are beginning to be utilized in risk stratification and clinical management of patients. Specifically, genetics can provide insight into the primary cause of chronic kidney disease (CKD), the risk of progression to ESRD, and post-transplant outcomes, including various forms of allograft rejection. The International Genetics & Translational Research in Transplantation Network (iGeneTRAiN), is a multi-site consortium that encompasses >45 genetic studies with genome-wide genotyping from over 51,000 transplant samples, including genome-wide data from >30 kidney transplant cohorts (n = 28,015). iGeneTRAiN is statistically powered to capture both rare and common genetic contributions to ESRD and post-transplant outcomes. The primary cause of ESRD is often difficult to ascertain, especially where formal biopsy diagnosis is not performed, and is unavailable in ~2% to >20% of kidney transplant recipients in iGeneTRAiN studies. We overview our current copy number variant (CNV) screening approaches from genome-wide genotyping datasets in iGeneTRAiN, in attempts to discover and validate genetic contributors to CKD and ESRD. Greater aggregation and analyses of well phenotyped patients with genome-wide datasets will undoubtedly yield insights into the underlying pathophysiological mechanisms of CKD, leading the way to improved diagnostic precision in nephrology.

Keywords: genomics, kidney disease, GWAS, whole exome sequencing analyses, whole genome sequencing

INTRODUCTION

The global prevalence of end-stage renal disease (ESRD) continues to climb. In 2016, 19,301 kidney transplants were performed in the United States, and approximately five times as many were performed worldwide[1,2]. Due to improvements in surgical techniques, immunosuppression protocols, and clinical management of post-transplant complications, the five-year graft survival rates for kidneys obtained from deceased and living donors reached highs of 75.3% and 85.3%, respectively (Cohen et al.,2006; Serur et al., 2011; Vignolini et al., 2019). However, the prevalence of ESRD cases in the US has continued to rise by ~20,000 cases per year over the past three decades, creating an increased need for kidney allografts[1]. This increase is believed to be due primarily to worsening diets and other modifiable factors associated with Western lifestyle but also to an increase in the longevity of pre-transplant ESRD cases.

It is well established that genetic factors contribute to the development and progression of specific types of chronic kidney disease (CKD), yet many previous studies have been limited in scope due to small sample sizes and genotyping strategies (Azarpira et al., 2014; Misra et al., 2014; Phelan et al 2014; Parsa et al., 2017; Stapleton et al., 2018). Studies of families with severe phenotypes of diseases, such as Alport's Syndrome and Fabry Disease, have significantly contributed to the understanding of the genetic characteristics of these conditions (Gillion et al., 2018; Kashiwagi et al., 2018; McCloskey et al., 2018). However, milder forms of these diseases and their role in the development of ESRD have yet to be explored in great depth.

Genome-Wide Genotyping Arrays

Array based genome-wide genotyping from diverse patient populations facilitates very precise ancestry determination using methods such as principal component analysis (Cai et al., 2013; Li et al., 2015). Genome-wide association studies (GWAS) among patients with CKD have detected both rare and common genetic variants significantly associated with estimated glomerular filtration rate (eGFR) decline and microalbuminuria, some of the strongest predictors of CKD outcomes, despite >80% of GWAS participants having eGFRs in the normal range (Boger et al., 2011a; Boger et al., 2011b;Reznichenko et al., 2012; Gorski et al., 2015; Parsa et al., 2017; Limou et al., 2018).

The findings of genome-wide studies may also provide new therapeutic targets to slow the progression of CKD to ESRD, which may delay or impact the need for transplantation in some patient populations (Wuttke & Kottgen, 2016; Kalatharan et al., 2018). For example, nephropathic cystinosis, a rare autosomal recessive disease, is caused by a 57-kb deletion in the *CTNS* gene in ~75% of patients of European ancestry and progresses to ESRD if left untreated (Brodin-Sartorius et al., 2012). However, treatment with oral cysteamine by five years of age has been found to significantly decrease the prevalence and delay the onset

of ESRD (Brodin-Sartorius et al., 2012). Additionally, at least 38 genes have been associated with the development of genetic focal segmental glomerulosclerosis (FSGS), some of which have been shown to be responsive to glucocorticoid treatment (Rosenberg & Kopp, 2017). GWAS findings can also provide insight into the biology of ESRD, helping to remove diagnostic heterogeneity. The two *APOL1* risk alleles (G1 and G2) found in high frequency in sub-Saharan African populations and strongly associated with FSGS and HIV nephropathy were found to activate protein kinase R, thus inducing glomerular injury and proteinuria (Kopp et al., 2011; Limou et al., 2014; Okamoto et al., 2018). Overall, results from genome-wide screening can enable physicians to provide accurate genetic diagnoses for the primary cause of ESRD, enabling timely and effective therapeutic managemenvwt and aiding in the evaluation of family members as living donors (Snoek et al., 2018).

Whole-Exome and Whole-Genome Sequencing

In the last decade, whole-exome sequencing (WES) and whole-genome sequencing (WGS) approaches have been used very successfully to discover and diagnose genetic disorders in a clinical context (Mallawaarachchi et al., 2016; Lata et al., 2018; Warejko et al., 2018; Groopman et al., 2019). WES typically yields sufficient depth of sequencing coverage across ~95% of nucleotides in coding regions captured and has been used to diagnose rare high penetrant, Mendelian disorders, discover common variants, and identify causal mutations in cancer (Huang et al., 2018; Zhang et al., 2018). WES has recently been implemented as a first-line diagnostic tool in clinical medicine. In a study on fetuses with congenital anomalies of the kidney and urinary tract (CAKUT), pathogenic variants were discovered in 13% of cases (Lei et al., 2017). WES has also been applied to adult-onset CKD and ESRD, in which ~10% of cases are caused by Mendelian mutations (Wuhl et al., 2014; Lata et al., 2018; Groopman et al., 2019). In a cohort of >3,000 patients with advanced CKD and ESRD ascertained for a clinical trial, WES identified diagnostic variants in 9.3% of patients encompassing 66 monogenic disorders (Groopman et al., 2019). Of the 343 detected variants, 141 (41%) had not been previously reported as pathogenic. Additionally, diagnostic variants were identified in 17.1% of individuals with nephropathy of unknown origin, altering medical management by initiating multidisciplinary care, prompting referral to clinical trials, and guiding donor selection for transplantation (Groopman et al., 2019). However, it should be noted that many CKD studies using WES have struggled to obtain adequate control populations. iGeneTRAiN has a large pool of healthy donors (in kidney and in other organs), which represents a strong advantage for our study designs.

WGS is the most comprehensive approach for the detection of inherited variants due to more complete genome-wide coverage, although there are additional challenges compared to WES. WGS can capture single nucleotide genetic variants, small Insertions and Deletions (Indels), and Copy-Number Variants (Cnvs) throughout the human genome. Although it has a higher cost per sample and can be more difficult to analyze than wes, greater diagnostic yields are evident in

---

[1]https://www.usrds.org

[2]http://www.transplant-observatory.org

patients with negative or inconclusive wes results (Alfares et al., 2018; Lionel et al., 2018). WGS has been shown to identify a diagnostic genetic variant in ~10–50% of individuals with a suspected genetic disorder, depending on the clinical study population(S-) being screened (van Der Ven et al., 2018; Groopman et al., 2019; Mann et al., 2019).

International Genetics and Translational Research in Transplantation Network

Despite technological advances that enable research to be carried out on a genome-wide scale, many studies have been hindered by small sample sizes in single transplant sites, as well as the vast number of complex donor and recipient clinical covariates and disease-related phenotypes observed in transplantation. The International Genetics & Translational Research in Transplantation Network (iGeneTRAiN) is a multi-site consortium that encompasses >45 genetic studies with ~51,210 solid-organ transplant subjects (International Genetics and Translational Research in Transplantation Network (iGeneTRAiN), 2015). The iGeneTRAiN consortium aims to discover and validate solid organ transplant related genetic factors and post-transplant complications, including primary disease, disease recurrence, drug- and cardio-metabolic related adverse events, and different forms of allograft rejection (International Genetics and Translational Research in Transplantation Network (iGeneTRAiN), 2015). Of the iGeneTRAiN samples, 54% (n = 28,015) are from kidney transplant cohorts and include 17,742 (63.3%) recipients and 10,273 (36.7%) donors. The genotyped

donor DNA provides control samples for all iGeneTRAiN studies, a large advantage over previously published genetic studies.

The iGeneTRAiN consortium designed and developed a genome-wide genotyping array, the "TxArray," which was enriched with content relating to known or putative transplant-specific genetic associations (Li et al., 2015). The TxArray version 1 contains ~782,000 genetic markers, with tailored transplant-specific content to capture variants across *HLA*, *KIR*, loss-of-function, pharmacogenomic, and cardio-metabolic loci. The array also contains extensive overlap with the UK Biobank Axiom® Array and the Axiom Biobank Genotyping Array, enabling future joint studies or meta-analyses using conventional, hypothesis-free GWAS approaches (Li et al., 2015).

The first wave of iGeneTRAiN kidney cohorts had a wide geographic representation with participants from various sites in the United States, Canada, Australia, The Netherlands, United Kingdom, and Ireland, including both adult and pediatric sites (**Figure 1**). Over the past few years, many genetic discoveries have been made within the iGeneTRAiN cohorts related to kidney, heart, liver, and lung transplants (Oetting et al., 2016; Greenland et al., 2017; Shaked et al., 2017; Hernandez-Fuentes et al., 2018; Oetting et al., 2018; Snoek et al., 2018; Oetting et al., 2019; Reindl-Schwaighofer et al., 2019). The Wellcome Trust Case Control Consortium (WTCCC) carried out the first large-scale GWAS with both kidney transplant donor and recipient DNA with the goal of identifying genetic variants, in addition to the *HLA* regions, that significantly contribute to long- and/ or short-term renal allograft survival (Hernandez-Fuentes
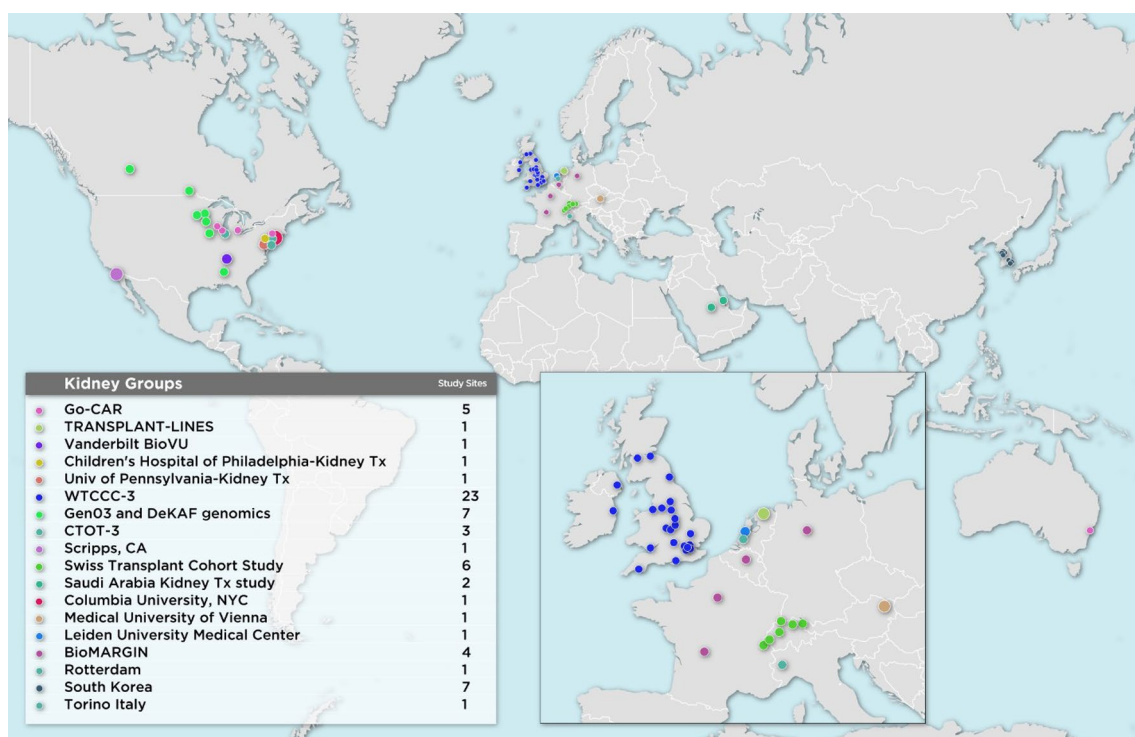

FIGURE 1 | Geographical Map of Current iGeneTRAiN Kidney Sites and Sub-sites.

et al., 2018). No non-*HLA* signals were observed at genome-wide significance in this initial study, illustrating the need for harmonization of larger, well-phenotyped kidney transplant cohorts. In addition to the previously discovered common loss-of-function variant *CYP3A5*3* allele (rs776746), the Deterioration of Kidney Allograft Function (DeKAF) Trial identified two *CYP3A5* variants, rs10264272 and rs41303343, and one *CYP3A4* variant, rs35599367, that explain additional portions of variance observed in dose-adjusted tacrolimus (TAC) through blood concentrations for African American (AA) and European ancestry (EA) kidney transplant recipients, respectively (Dai et al., 2006; Jacobson et al., 2011; Oetting et al., 2016; Oetting et al., 2018). These findings illustrate the utility of genome-wide studies when determining immunosuppression therapy regimens post-transplant, potentially contributing to improvements in renal allograft survival. Another iGeneTRAiN study showed that GWAS performed in nontransplant settings can predict post-transplant complications. Polygenic risk scores calculated from non-melanoma skin cancer (NMSC) GWAS in the general population predicted risk of and time to post-transplant NMSC and added additional predictive value beyond that explained by clinical variables (Stapleton et al., 2019).

Ongoing iGeneTRAiN Kidney Genome-Wide Studies

Recently, additional kidney transplant cohorts from Austria, Belgium, Germany, France, Italy, The Netherlands, Saudi Arabia, South Korea, Switzerland, and additional United States sites have joined iGeneTRAiN. This greatly increases ancestral diversity of recipients and donors, as well as statistical power to detect transplant related genetic variants that impact primary disease and transplant outcomes (**Figure 1**). Our large sample sizes are enabling us to investigate both donor and recipient characteristics that effect ESRD cause, treatment, and transplant-related outcomes.

Where available, we obtained formal clinical diagnoses of primary cause of ESRD, organized into disease categories of diabetic, arteriopathic, glomerular, acute kidney injury, infective and obstructive nephropathy, congenital, familial, toxic nephropathy, and malignancies, for all iGeneTRAiN kidney cohort subjects (**Table 1**). With these datasets, we are working to increase our understanding of the genetic underpinnings of ESRD and primary disease through single nucleotide polymorphism (SNP) based GWAS, copy number variant (CNV) screening, donor-recipient properties, allogenicity, and transplant outcomes.

Copy Number Variant Screening in iGeneTRAiN Cohorts

Genome-wide genotyping arrays are well established as an effective means for identification of known and novel CNVs (Sallustio et al., 2015; Ai et al., 2016; Verbitsky et al., 2019). CNV screening within iGeneTRAiN subjects is of major interest for both assessing the genetic architecture of primary disease and for allogenicity studies. iGeneTRAiN has developed an extensive loss-of-function (LoF) pipeline which includes haplotype phasing of over 10 million directly genotyped and imputed variants. We

are particularly interested in two copy LoF (by single-nucleotide variants and/or CNVs) and integration of one or two copy LoF variants for donor-recipient interaction analyses, for association with time-to-rejection and graft loss events (International Genetics and Translational Research in Transplantation Network (iGeneTRAiN), 2015).

CNV screening in *a priori* regions for primary disease has been performed in iGeneTRAiN cohorts. For example, we performed CNV screening in patients with nephronophthisis (NPH), the most common genetic cause of ESRD in children and often caused by homozygous *NPHP1* full gene deletions (Levy and Feingold, 2000; Hildebrandt, 2010; Wolf & Hildebrandt, 2011). In iGeneTRAiN, we previously examined this region in a subset of iGeneTRAiN studies for adult-onset ESRD (n = 5,606 patients). Of the subjects analyzed, 26 patients showed homozygous *NPHP1* CNV deletions. Interestingly only 12% of these patients were previously diagnosed as having NPH and many presented with ESRD later in adulthood (Snoek et al., 2018). Thus, using the two copy gene loss of *NPHP1* from GWG arrays to ascertain NPH status and examine NPH-related information in iGeneTRAiN studies, including accuracy of case-ascertainment and age-of-onset, shows a strong proof-of-principle for use in other high penetrant autosomal recessive/dominant cases, and the need for further sequencing for rare single-nucleotide variants in adult-onset ESRD patients. Furthermore, in a recent genome-wide analysis of CNVs in almost 3,000 cases of CAKUT, 45 distinct, known genomic disorders at 37 independent genomic loci were identified in 4% of CAKUT cases, and novel genomic disorders were found in an additional ~2% of cases (Verbitsky et al., 2019). Genome-wide genotyping and imputation using large whole-genome sequencing (WGS) datasets, such as the 1000 genomes project (1KGP), typically cannot identify variants in the most common ancestral populations to a minor allele frequency (MAF) of <0.005, yet it is often possible to identify rare CNVs using monomorphic or SNP based probes across loci.

Our previous analyses of the Axiom TxArray genome-wide genotyping data was primarily limited to approximately 2,000 *a priori* CNV regions of interest that had specific probes designed onto the TxArray. Initial analyses used an adaption of the BRLMM-P CNV algorithm adapted from algorithms previously used to cluster genotypes across many samples (Yeung et al., 2008). However, BRLMM-P could only identify up to three clusters and thus was only able to detect 0, 1, and 2 copy deletions. The newer Axiom Analysis Suite 4.0 software allows streamlined, targeted, and *de-novo* whole-genome CNV region analysis[3]. A major advantage of the newer software is the ability to detect duplications as well as 0, 1, and 2 copy deletions (**Figure 2**).

DISCUSSION

Genome-wide genotyping studies have become very affordable and streamlined. However, large sample sizes, on the order of 10,000–100,000, are needed in order to detect both rare variants
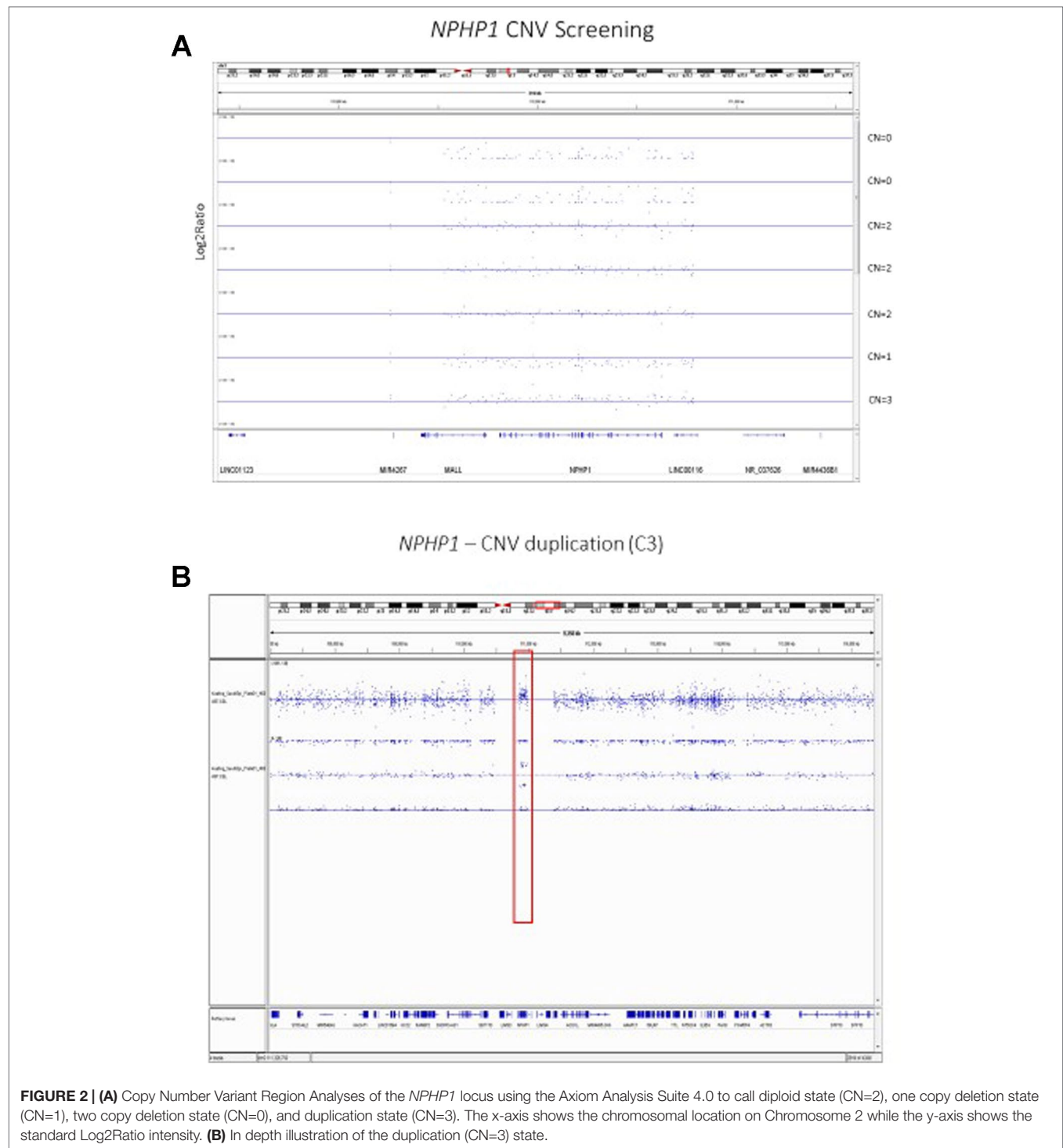
---

[3] https://www.thermofisher.com

TABLE 1 | Primary cause of ESRD in iGeneTRAiN kidney cohorts.

Primary Cause of ESRD	University of Pennsylvania	Transplant-LINES	BioMARGIN‡	Swiss Transplant Cohort Study†	WTCCC-3†	Go-CAR	Columbia Transplant Biobank	Torino Transplant Cohort‡	Rotterdam‡	DeKAF*	Gen03*	Scripps	Leiden*	Vienna	Vanderbilt	Seoul†,‡
Diabetic nephropathy	310 (26.8%)	51 (4.0%)	115 (8.5%)	116 (7.4%)	186 (7.3%)	219 (37.2%)	176 (15.6%)	22 (3.3%)	– (17.0%)	553 (28.5%)	178 (22.8%)	–	22 (25.0%)	113 (12.8%)	–	240 (20.9%)
Arteriopathic renal disease	202 (17.5%)	104 (8.2%)	67 (4.9%)	183 (11.6%)	111 (4.4%)	109 (18.5%)	123 (10.9%)	36 (5.4%)	– (13.0%)	132 (6.8%)	43 (5.5%)	–	–	88 (10.0%)	–	244 (21.3%)
Glomerular disease	340 (29.4%)	401 (31.6%)	329 (24.2%)	433 (27.5%)	482 (19.0%)	111 (18.9%)	431 (38.3%)	198 (29.7%)	– (20.0%)	446 (23.0%)	226 (29.0%)	–	14 (15.9%)	296 (33.6%)	–	251 (21.9%)
AKI	3 (0.3%)	1 (0.1%)	3 (0.2%)	0 (0.0%)	8 (0.3%)	0 (0.0%)	5 (0.4%)	2 (0.3%)	– (4.0%)	–	–	–	–	4 (0.5%)	–	0 (0.0%)
Infective and obstructive nephropathy	31 (2.7%)	107 (8.4%)	79 (5.8%)	85 (5.4%)	203 (8.0%)	22 (3.7%)	17 (1.5%)	87 (13.1%)	– (6.0%)	–	–	–	–	56 (6.4%)	–	6 (0.5%)
Congenital diseases	19 (1.6%)	68 (5.4%)	57 (4.2%)	40 (2.5%)	26 (1.0%)	14 (2.4%)	58 (5.2%)	31 (4.7%)	– (3.0%)	–	–	–	–	18 (2.1%)	–	0 (0.0%)
Familial hereditary renal diseases	143 (12.4%)	235 (18.5%)	245 (18.0%)	327 (20.8%)	361 (14.2%)	67 (11.4%)	145 (12.9%)	110 (16.5%)	– (14.0%)	308 (15.9%)	123 (15.8%)	–	52 (59.1%)	147 (16.7%)	–	25 (2.2%)
Toxic nephropathy	2 (1.7%)	15 (1.2%)	19 (1.4%)	46 (2.9%)	16 (0.6%)	18 (3.1%)	31 (2.8%)	9 (1.4%)	– (1.0%)	–	–	–	–	25 (2.8%)	–	4 (0.4%)
Neoplasms	5 (0.4%)	6 (0.5%)	21 (1.5%)	0 (0.0%)	20 (0.8%)	2 (0.3%)	6 (0.5%)	5 (0.8%)	– (3.0%)	–	–	–	–	14 (1.6%)	–	0 (0.0%)
Other disease	50 (4.3%)	59 (4.6%)	64 (4.7%)	344 (21.9%)	165 (6.5%)	12 (2.0%)	9 (0.8%)	6 (0.9%)	– (10.0%)	432 (22.3%)	166 (21.3%)	–	–	17 (1.9%)	–	0 (0.0%)
ESRD of unknown etiology	29 (2.5%)	168 (13.2%)	161 (11.8%)	0 (0.0%)	0 (0.0%)	14 (2.4%)	125 (11.1%)	160 (24.0%)	– (9.0%)	–	–	–	–	100 (11.4%)	–	185 (16.1%)
Missing information	4 (0.4%)	56 (4.4%)	200 (14.7%)	0 (0.0%)	966 (38.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	– (0.0%)	67 (3.5%)	44 (5.6%)	–	–	2 (0.2%)	–	193 (16.8%)
Total	**1156**	**1271**	**1360**	**1574**	**2544**	**588**	**1126**	**666**	**–**	**1938**	**780**	**–**	**88**	**880**	**–**	**1148**

*Data includes Caucasian only ancestry. ‡Recently added iGeneTRAiN kidney cohort.

†National level iGeneTRAiN kidney cohort. See **Supplementary Table 1** for full breakdown of primary cause of ESRD.

FIGURE 2 | (A) Copy Number Variant Region Analyses of the *NPHP1* locus using the Axiom Analysis Suite 4.0 to call diploid state (CN=2), one copy deletion state (CN=1), two copy deletion state (CN=0), and duplication state (CN=3). The x-axis shows the chromosomal location on Chromosome 2 while the y-axis shows the standard Log2Ratio intensity. **(B)** In depth illustration of the duplication (CN=3) state.

with large contributions and common variants with minor contributions to a specific phenotype(s) (Korte and Farlow, 2013). While it is very important to bolster statistical power to detect genetic underpinnings of transplant-related phenotypes by aggregating similar cohorts, great caution must be exercised when combining genotyping and phenotyping datasets, especially as transplant study covariates are very complex and can vary greatly

by era and geographical region. iGeneTRAiN does have a unified quality control/quality assurance GWAS pipeline, including adjustment for population-based stratification (International Genetics and Translational Research in Transplantation Network (iGeneTRAiN), 2015). Association study analyses do adjust for all known/available study covariates, including patient demographics and clinical characteristics, and we adjust for each

transplant site alone to look for confounders. Genome-wide genotyping arrays are generally poor at detecting rarer frequency pathogenic variants, with the exception of medium to large CNVs. Significant advances in genomic technologies and the decreasing cost of WES/WGS efforts over the past several years have made it increasingly feasible to carry out better designed genome-wide studies in a clinical environment (Gumpinger et al., 2018). However, there are still significant advantages to having genome-wide genotyping array datasets, as rigorous quality control and quality assurance measures are generally performed on the original DNA, and gender, ancestry, and HLA (amino acid imputation) concordance checks can be performed before progressing to WES or WGS pipelines for deeper genetic characterization. GWAS are able to provide insight into genetic risk scores and pathogenic CNVs, as genome-wide variants are covered in conventional genome-wide genotyping arrays (Sampson and Juppner, 2013; Li et al., 2015 Marigorta et al., 2018; Snoek et al., 2018; Stapleton et al., 2019). For example, a meta-analysis across 36 articles identified three genetic variants that are significantly associated with new onset diabetes after transplantation (NODAT), all of which are also known risk factor variants for Type 2 diabetes. The integration and analysis of large and complex multi-omic datasets has been demonstrated in a number of recent high impact publications, which in general increase, by approximately 10-fold, the statistical power to detect and illustrate functional variants (Chen et al., 2012; Piening et al., 2018; Schüssler-Fiorenza Rose et al., 2019; Zhou et al., 2019). iGeneTRAiN genomic data can be integrated with results from proteome-, metabolome-, and transcriptome-wide transplant studies to further characterize clinical risks and allow for personalized treatments, as a number of iGeneTRAiN studies have multi-omic datasets/samples (International Genetics and Translational Research in Transplantation Network (iGeneTRAiN), 2015).

The advent of single-cell RNA sequencing (scRNASeq) has yielded major insights into the biology of CKD. Expression quantitative trait loci (eQTL) atlases have been generated for glomerular and tubular compartments from human kidney cells. Integrating results from genome-wide studies of CKD with eQTL from scRNAseq as well as known regulatory region maps has been shown to identify novel CKD genes (Qiu et al., 2018). The Human Cell Atlas project is a major international initiative which aims to create comprehensive reference maps of all human cells to gain fundamental insight into the understanding of human health and will undoubtedly aid in the diagnoses and surveillance of a range of diseases (Regev et al., 2017).

Future of iGeneTRAiN Kidney Cohorts Analyses

As the population of kidney transplant recipients and donors continues to grow in the iGeneTRAiN consortium and as post-transplant outcomes accrue, we will be able to further increase our knowledge of the genetic underpinnings of ESRD, primary disease, and post-transplant outcomes, such as acute rejection and graft loss. These sequencing approaches may provide additional insight into donor-recipient (D-R) interactions that influence graft outcomes. Although it is well established that allelic matches across HLA loci impact clinical outcomes post-transplant, there is a paucity of genome-wide research conducted to identify donor-recipient interactions independent of HLA (Thorsby, 2009; Chan-On and Sarwal, 2016 ;Stapleton et al., 2018). One recent iGeneTRAiN kidney D-R study showed decreased allograft survival of recipients with increased D-R kidney transmembrane non-synonymous SNPs (nsSNPs). We further demonstrated that we could detect alloantibodies against customized amino-acid peptides designed with a number of these kidney transmembrane nsSNPs using sera from these patients (Reindl-Schwaighofer, et al., 2019). Finally, data from all solid-organ transplant studies in the iGeneTRAiN consortium will be utilized in cross-organ studies in order to gain additional insight into the genetics of acute rejection, allograft/patient survival, and pharmacogenomic outcomes.

ETHICS STATEMENT

All data used in this publication was collected in accordance with local IRB stipulations.

AUTHOR CONTRIBUTIONS

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.01084/full#supplementary-material

REFERENCES

Ai, Z., Li, M., Liu, W., Foo, J. N., Mansouri, O., Yin, P., et al. (2016). Low alpha-defensin gene copy number increases the risk for IgA nephropathy and renal dysfunction. *Sci. Transl. Med.* 8 (345), 345ra388. doi: 10.1126/scitranslmed.aaf2106

Alfares, A., Aloraini, T., Subaie, L. A., Alissa, A., Qudsi, A. A., Alahmad, A., et al. (2018). Whole-genome sequencing offers additional but limited clinical utility compared with reanalysis of whole-exome sequencing. *Genet. Med.* 20 (11), 1328–1333. doi: 10.1038/gim.2018.41

Azarpira, N., Kazemi, K., and Darai, M. (2014). Influence of p53 (rs1625895) polymorphism in kidney transplant recipients. *Saudi J. Kidney Dis. Transpl* 25 (6), 1160–1165. doi: 10.4103/1319-2442.144248

Boger, C. A., Chen, M. H., Tin, A., Olden, M., Kottgen, A., de Boer, I. H., et al. (2011a). CUBN is a gene locus for albuminuria. *J. Am. Soc. Nephrol.* 22 (3), 555–570. doi: 10.1681/ASN.2010060598

Boger, C. A., Gorski, M., Li, M., Hoffmann, M. M., Huang, C., Yang, Q., et al. (2011b). Association of eGFR-Related Loci Identified by GWAS with Incident CKD and ESRD. *PloS Genet.* 7 (9), e1002292. doi: 10.1371/journal.pgen.1002292

Brodin-Sartorius, A., Tete, M. J., Niaudet, P., Antignac, C., Guest, G., Ottolenghi, C., et al. (2012). Cysteamine therapy delays the progression of nephropathic cystinosis in late adolescents and adults. *Kidney Int.* 81 (2), 179–189. doi: 10.1038/ki.2011.277

Cai, M., Dai, H., Qiu, Y., Zhao, Y., Zhang, R., Chu, M., et al. (2013). SNP set association analysis for genome-wide association studies. *PloS One* 8 (5), e62495. doi: 10.1371/journal.pone.0062495

Calabrese, D. R., Florez, R., Dewey, K., Hui, C., Torgerson, D., Chong, T., et al. (2018). Genotypes associated with tacrolimus pharmacokinetics impact clinical outcomes in lung transplant recipients. *Clin. Transplant.* 32 (8), e13332. doi: 10.1111/ctr.13332

Chan-On, C., and Sarwal, M. (2016). Non-HLA Antibodies in clinical transplantation. *Clin. Transpl* 32, 45–61.

Chen, R., Mias, G. I., Li-Pook-Than, J., Jiang, L., Lam, H. Y., Chen, R., et al. (2012). Personal omics profiling reveals dynamic molecular and medical phenotypes. *Cell* 148 (6), 1293–1307. doi: 10.1016/j.cell.2012.02.009

Cohen, D. J., St Martin, L., Christensen, L. L., Bloom, R. D., and Sung, R. S. (2006). Kidney and pancreas transplantation in the United States, 1995-2004. *Am. J. Transplant.* 6 (5 Pt 2), 1153–1169. doi: 10.1111/j.1600-6143.2006.01272.x

Dai, Y., Hebert, M. F., Isoherranen, N., Davis, C. L., Marsh, C., Shen, D. D., et al. (2006). Effect of CYP3A5 polymorphism on tacrolimus metabolic clearance in vitro. *Drug Metab. Dispos* 34 (5), 836–847. doi: 10.1124/dmd.105.008680

Gillion, V., Dahan, K., Cosyns, J. P., Hilbert, P., Jadoul, M., Goffin, E., et al. (2018). Genotype and outcome after kidney transplantation in alport syndrome. *Kidney Int. Rep.* 3 (3), 652–660. doi: 10.1016/j.ekir.2018.01.008

Gorski, M., Tin, A., Garnaas, M., McMahon, G. M., Chu, A. Y., Tayo, B. O., et al. (2015). Genome-wide association study of kidney function decline in individuals of European descent. *Kidney Int.* 87 (5), 1017–1029. doi: 10.1038/ki.2014.361

Greenland, J. R., Sun, H., Calabrese, D., Chong, T., Singer, J. P., Kukreja, J., et al. (2017). HLA mismatching favoring host-versus-graft NK cell activity via KIR3DL1 is associated with improved outcomes following lung transplantation. *Am. J. Transplant.* 17 (8), 2192–2199. doi: 10.1111/ajt.14295

Groopman, E. E., Marasa, M., Cameron-Christie, S., Petrovski, S., Aggarwal, V. S., Milo-Rasouly, H., et al. (2019). Diagnostic utility of exome sequencing for kidney disease. *N Engl. J. Med.* 380 (2), 142–151. doi: 10.1056/NEJMoa1806891

Gumpinger, A. C., Roqueiro, D., Grimm, D. G., and Borgwardt, K. M. (2018). Methods and tools in genome-wide association studies. *Methods Mol. Biol.* 1819, 93–136. doi: 10.1007/978-1-4939-8618-7_5

Hernandez-Fuentes, M. P., Franklin, C., Rebollo-Mesa, I., Mollon, J., Delaney, F., Perucha, E., et al. (2018). Long- and short-term outcomes in renal allografts with deceased donors: a large recipient and donor genome-wide association study. *Am. J. Transplant.* 18 (6), 1370–1379. doi: 10.1111/ajt.14594

Hildebrandt, F. (2010). Genetic kidney diseases. *Lancet* 375 (9722), 1287–1295. doi: 10.1016/S0140-6736(10)60236-X

Huang, K. L., Mashl, R. J., Wu, Y., Ritter, D. I., Wang, J., Oh, C., et al. (2018). Pathogenic germline variants in 10,389 adult cancers. *Cell* 173 (2), 355–370 e314. doi: 10.1016/j.cell.2018.03.039

International Genetics and Translational Research in Transplantation Network (iGeneTRAiN). (2015). Design and Implementation of the International Genetics and Translational Research in Transplantation Network. *Transplantation* 99 (11), 2401–2412. doi: 10.1097/TP.0000000000000913

Jacobson, P. A., Oetting, W. S., Brearley, A. M., Leduc, R., Guan, W., Schladt, D., et al. (2011). Novel polymorphisms associated with tacrolimus trough concentrations: results from a multicenter kidney transplant consortium. *Transplantation* 91 (3), 300–308. doi: 10.1097/TP.0b013e318200e991

Kalatharan, V., Lemaire, M., and Lanktree, M. B. (2018). Opportunities and challenges for genetic studies of end-stage renal disease in Canada. *Can. J. Kidney Health Dis.* 5, 2054358118789368. doi: 10.1177/2054358118789368

Kashiwagi, Y., Suzuki, S., Agata, K., Morishima, Y., Inagaki, N., Numabe, H., et al. (2018). A family case of X-linked Alport syndrome patients with a novel variant in COL4A5. *CEN Case Rep.* 8 (2), 75–78. doi: 10.1007/s13730-018-0368-4

Kopp, J. B., Nelson, G. W., Sampath, K., Johnson, R. C., Genovese, G., An, P., et al. (2011). APOL1 genetic variants in focal segmental glomerulosclerosis and HIV-associated nephropathy. *J. Am. Soc. Nephrol.* 22 (11), 2129–2137. doi: 10.1681/ASN.2011040388

Korte, A., and Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods* 9, 29. doi: 10.1186/1746-4811-9-29

Lata, S., Marasa, M., Li, Y., Fasel, D. A., Groopman, E., Jobanputra, V., et al. (2018). Whole-Exome Sequencing in Adults With Chronic Kidney Disease: A Pilot Study. *Ann. Intern Med.* 168 (2), 100–109. doi: 10.7326/M17-1319

Lei, T. Y., Fu, F., Li, R., Wang, D., Wang, R. Y., Jing, X. Y., et al. (2017). Whole-exome sequencing for prenatal diagnosis of fetuses with congenital anomalies of the kidney and urinary tract. *Nephrol. Dial Transplant.* 32 (10), 1665–1675. doi: 10.1093/ndt/gfx031

Levy, M., and Feingold, J. (2000). Estimating prevalence in single-gene kidney diseases progressing to renal failure. *Kidney Int.* 58 (3), 925–943. doi: 10.1046/j.1523-1755.2000.00250.x

Li, Y. R., van Setten, J., Verma, S. S., Lu, Y., Holmes, M. V., Gao, H., et al. (2015). Concept and design of a genome-wide association genotyping array tailored for transplantation-specific studies. *Genome Med.* 7, 90. doi: 10.1186/s13073-015-0211-x

Limou, S., Nelson, G. W., Kopp, J. B., and Winkler, C. A. (2014). APOL1 kidney risk alleles: population genetics and disease associations. *Adv. Chronic Kidney Dis.* 21 (5), 426–433. doi: 10.1053/j.ackd.2014.06.005

Limou, S., Vince, N., and Parsa, A. (2018). Lessons from CKD-related genetic association studies-moving forward. *Clin. J. Am. Soc. Nephrol.* 13 (1), 140–152. doi: 10.2215/CJN.09030817

Lionel, A. C., Costain, G., Monfared, N., Walker, S., Reuter, M. S., Hosseini, S. M., et al. (2018). Improved diagnostic yield compared with targeted gene sequencing panels suggests a role for whole-genome sequencing as a first-tier genetic test. *Genet. Med.* 20 (4), 435–443. doi: 10.1038/gim.2017.119

Mallawaarachchi, A. C., Hort, Y., Cowley, M. J., McCabe, M. J., Minoche, A., Dinger, M. E., et al. (2016). Whole-genome sequencing overcomes pseudogene homology to diagnose autosomal dominant polycystic kidney disease. *Eur. J. Hum. Genet.* 24 (11), 1584–1590. doi: 10.1038/ejhg.2016.48

Mann, N., Braun, D. A., Amann, K., Tan, W., Shril, S., Connaughton, D. M., et al. (2019). whole-exome sequencing enables a precision medicine approach for kidney transplant recipients. *J. Am. Soc. Nephrol.* 30 (2), 201–215. doi: 10.1681/ASN.2018060575

Marigorta, U. M., Rodriguez, J. A., Gibson, G., and Navarro, A. (2018). Replicability and prediction: lessons and challenges from GWAS. *Trends Genet.* 34 (7), 504–517. doi: 10.1016/j.tig.2018.03.005

McCloskey, S., Brennan, P., and Sayer, J. A. (2018). Variable phenotypic presentations of renal involvement in Fabry disease: a case series. *Res* 7, 356. doi: 10.12688/f1000research.13708.1 F1000.

Misra, M. K., Kapoor, R., Pandey, S. K., Sharma, R. K., and Agrawal, S. (2014). Association of CTLA-4 gene polymorphism with end-stage renal disease and renal allograft outcome. *J. Interferon Cytokine Res.* 34 (3), 148–161. doi: 10.1089/jir.2013.0069

Oetting, W. S., Schladt, D. P., Guan, W., Miller, M. B., Remmel, R. P., Dorr, C., et al. (2016). Genomewide association study of tacrolimus concentrations in african american kidney transplant recipients identifies multiple CYP3A5 alleles. *Am. J. Transplant.* 16 (2), 574–582. doi: 10.1111/ajt.13495

Oetting, W. S., Wu, B., Schladt, D. P., Guan, W., Remmel, R. P., Mannon, R. B., et al. (2018). Genome-wide association study identifies the common variants

in CYP3A4 and CYP3A5 responsible for variation in tacrolimus trough concentration in Caucasian kidney transplant recipients. *Pharmacogenomics J.* 18 (3), 501–505. doi: 10.1038/tpj.2017.49

Oetting, W. S., Wu, B., Schladt, D. P., Guan, W., van Setten, J., Keating, B. J., et al. (2019). Genetic variants associated with immunosuppressant pharmacokinetics and adverse effects in the DeKAF genomics genome wide association studies. *Transplantation* 103 (6), 1131–1139. doi: 10.1097/TP.0000000000002625

Okamoto, K., Rausch, J. W., Wakashin, H., Fu, Y., Chung, J. Y., Dummer, P. D., et al. (2018). APOL1 risk allele RNA contributes to renal toxicity by activating protein kinase R. *Commun. Biol.* 1, 188. doi: 10.1038/s42003-018-0188-2

Parsa, A., Kanetsky, P. A., Xiao, R., Gupta, J., Mitra, N., Limou, S., et al. (2017). Genome-wide association of CKD progression: the chronic renal insufficiency cohort study. *J. Am. Soc. Nephrol.* 28 (3), 923–934. doi: 10.1681/ASN.2015101152

Phelan, P. J., Conlon, P. J., and Sparks, M. A. (2014). Genetic determinants of renal transplant outcome: where do we stand? *J. Nephrol.* 27 (3), 247–256. doi: 10.1007/s40620-014-0053-4

Piening, B. D., Zhou, W., Contrepois, K., Röst, H., Gu Urban, G. J., Mishra, T., et al. (2018). Integrative personal omics profiles during periods of weight gain and loss. *Cell Syst.* 6 (2), 157–170. doi: 10.1016/j.cels.2017.12.013

Qiu, C., Huang, S., Park, J., Park, Y., Ko, Y. A., Seasock, M. J., et al. (2018). Renal compartment-specific genetic variation analyses identify new pathways in chronic kidney disease. *Nat. Med.* 24 (11), 1721–1731. doi: 10.1038/s41591-018-0194-4

Regev, A., Teichmann, S. A., Lander, E. S., Amit, I., Benoist, C., Birney, E., et al. (2017). The human cell atlas. *Elife* 6, 1–30. doi: 10.7554/eLife.27041

Reindl-Schwaighofer, R., Heinzel, A., Kainz, A., van Setten, J., Jelencsics, K., Hu, K., et al. (2019). Contribution of non-HLA incompatibility between donor and recipient to kidney allograft survival: genome-wide analysis in a prospective cohort. *Lancet* 393 (10174), 910–917. doi: 10.1016/S0140-6736(18)32473-5

Reznichenko, A., Boger, C. A., Snieder, H., van den Born, J., de Borst, M. H., Damman, J., et al. (2012). UMOD as a susceptibility gene for end-stage renal disease. *BMC Med. Genet.* 13, 78. doi: 10.1186/1471-2350-13-78

Rosenberg, A. Z., and Kopp, J. B. (2017). Focal segmental glomerulosclerosis. *Clin. J. Am. Soc. Nephrol.* 12 (3), 502–517. doi: 10.2215/CJN.05960616

Sallustio, F., Cox, S. N., Serino, G., Curci, C., Pesce, F., De Palma, G., et al. (2015). Genome-wide scan identifies a copy number variable region at 3p21.1 that influences the TLR9 expression levels in IgA nephropathy patients. *Eur. J. Hum. Genet.* 23 (7), 940–948. doi: 10.1038/ejhg.2014.208

Sampson, M. G., and Juppner, H. (2013). Genes, exomes, genomes, copy number: what is their future in pediatric renal disease. *Curr. Pediatr. Rep.* 1 (1), 52–59. doi: 10.1007/s40124-012-0001-5

Schüssler-Fiorenza Rose, S. M., Contrepois, K., Moneghetti, K. J., Zhou, K., Mishra, T., Mataraso, S., et al. (2019). A longitudinal big data approach for precision health. *Nat. Med.* 25 (5), 792–804. doi: 10.1038/s41591-019-0414-6

Serur, D., Saal, S., Wang, J., Sullivan, J., Bologa, R., Hartono, C., et al. (2011). Deceased-donor kidney transplantation: improvement in long-term survival. *Nephrol. Dial Transplant.* 26 (1), 317–324. doi: 10.1093/ndt/gfq415

Shaked, A., Chang, B. L., Barnes, M. R., Sayre, P., Li, Y. R., Asare, S., et al. (2017). An ectopically expressed serum miRNA signature is prognostic, diagnostic, and biologically related to liver allograft rejection. *Hepatology* 65 (1), 269–280. doi: 10.1002/hep.28786

Snoek, R., van Setten, J., Keating, B. J., Israni, A. K., Jacobson, P. A., Oetting, W. S., et al. (2018). NPHP1 (Nephrocystin-1) gene deletions cause adult-onset ESRD. *J. Am. Soc. Nephrol.* 29 (6), 1772–1779. doi: 10.1681/ASN.2017111200

Stapleton, C. P., Birdwell, K. A., McKnight, A. J., Maxwell, A. P., Mark, P. B., Sanders, M. L., et al. (2019). Polygenic risk score as a determinant of risk of non-melanoma skin cancer in a European-descent renal transplant cohort. *Am. J. Transplant.* 19 (3), 801–810. doi: 10.1111/ajt.15057

Stapleton, C. P., Conlon, P. J., and Phelan, P. J. (2018). Using omics to explore complications of kidney transplantation. *Transpl Int.* 31 (3), 251–262. doi: 10.1111/tri.13067

Thorsby, E. (2009). A short history of HLA. *Tissue Antigens* 74 (2), 101–116. doi: 10.1111/j.1399-0039.2009.01291.x

van der Ven, A. T., Connaughton, D. M., Ityel, H., Mann, N., Nakayama, M., Chen, J., et al. (2018). Whole-exome sequencing identifies causative mutations in families with congenital anomalies of the kidney and urinary tract. *J. Am. Soc. Nephrol.* 29 (9), 2348–2361. doi: 10.1681/ASN.2017121265

Verbitsky, M., Westland, R., Perez, A., Kiryluk, K., Liu, Q., Krithivasan, P., et al. (2019). The copy number variation landscape of congenital anomalies of the kidney and urinary tract. *Nat. Genet.* 51 (1), 117–127. doi: 10.1038/s41588-018-0281-y

Vignolini, G., Campi, R., Sessa, F., Greco, I., Larti, A., Giancane, S., et al. (2019). Development of a robot-assisted kidney transplantation programme from deceased donors in a referral academic centre: technical nuances and preliminary results. *BJU Int.* 123 (3), 474–484. doi: 10.1111/bju.14588

Warejko, J. K., Tan, W., Daga, A., Schapiro, D., Lawson, J. A., Shril, S., et al. (2018). Whole exome sequencing of patients with steroid-resistant nephrotic syndrome. *Clin. J. Am. Soc. Nephrol.* 13 (1), 53–62. doi: 10.2215/CJN.04120417

Wolf, M. T., and Hildebrandt, F. (2011). Nephronophthisis. *Pediatr. Nephrol.* 26 (2), 181–194. doi: 10.1007/s00467-010-1585-z

Wuhl, E., van Stralen, K. J., Wanner, C., Ariceta, G., Heaf, J. G., Bjerre, A. K., et al. (2014). Renal replacement therapy for rare diseases affecting the kidney: an analysis of the ERA-EDTA Registry. *Nephrol. Dial Transplant.* 29 Suppl 4, iv1–iv8. doi: 10.1093/ndt/gfu030

Wuttke, M., and Kottgen, A. (2016). Insights into kidney diseases from genome-wide association studies. *Nat. Rev. Nephrol.* 12 (9), 549–562. doi: 10.1038/nrneph.2016.107

Yeung, J. M., Sham, P. C., Chan, A. S., and Cherny, S. S. (2008). OpenADAM: an open source genome-wide association data management system for Affymetrix SNP arrays. *BMC Genomics* 9, 636. doi: 10.1186/1471-2164-9-636

Zhang, J., Qiu, W., Liu, H., Qian, C., Liu, D., Wang, H., et al. (2018). Genomic alterations in gastric cancers discovered via whole-exome sequencing. *BMC Cancer* 18 (1), 1270. doi: 10.1186/s12885-018-5097-8

Zhou, W., Sailani, M. R., Contrepois, K., Zhou, Y., Ahadi, S., Leopold, S. R., et al. (2019). Longitudinal multi-omics of host-microbe dynamics in prediabetes. *Nature* 569 (7758), 663–671. doi: 10.1038/s41586-019-1236-x

Diagnostic Yield of Next-Generation Sequencing in Patients With Chronic Kidney Disease of Unknown Etiology

Amber de Haan[1], Mark Eijgelsheim[1], Liffert Vogt[2], Nine V. A. M. Knoers[3] and Martin H. de Borst[1]*

[1] Department of Internal Medicine, Division of Nephrology, University Medical Center Groningen, University of Groningen, Groningen, Netherlands, [2] Section Nephrology, Amsterdam Cardiovascular Sciences, Department of Internal Medicine, Amsterdam University Medical Centre, University of Amsterdam, Amsterdam, Netherlands, [3] Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen, Netherlands

Advances in next-generation sequencing (NGS) techniques, including whole exome sequencing, have facilitated cost-effective sequencing of large regions of the genome, enabling the implementation of NGS in clinical practice. Chronic kidney disease (CKD) is a major contributor to global burden of disease and is associated with an increased risk of morbidity and mortality. CKD can be caused by a wide variety of primary renal disorders. In about one in five CKD patients, no primary renal disease diagnosis can be established. Moreover, recent studies indicate that the clinical diagnosis may be incorrect in a substantial number of patients. Both the absence of a diagnosis or an incorrect diagnosis can have therapeutic implications. Genetic testing might increase the diagnostic accuracy in patients with CKD, especially in patients with unknown etiology. The diagnostic utility of NGS has been shown mainly in pediatric CKD cohorts, while emerging data suggest that genetic testing can also be a valuable diagnostic tool in adults with CKD. In addition to its implications for unexplained CKD, NGS can contribute to the diagnostic process in kidney diseases with an atypical presentation, where it may lead to reclassification of the primary renal disease diagnosis. So far, only a few studies have reported on the diagnostic yield of NGS-based techniques in patients with unexplained CKD. Here, we will discuss the potential diagnostic role of gene panels and whole exome sequencing in pediatric and adult patients with unexplained and atypical CKD.

Keywords: end-stage renal disease, exome sequencing, diagnostic utility, genetic testing, kidney disease

INTRODUCTION

Chronic kidney disease (CKD) affects over 850 million individuals worldwide, and is associated with an increased risk of morbidity and mortality (Go et al., 2004; Gansevoort et al., 2013; Hill et al., 2016; Jager et al., 2019) and a high burden in terms of quality of life and costs (Smith et al., 2004; Baumeister et al., 2010; Eriksson et al., 2016; Nguyen et al., 2018). Recent predictions by the Institute of Health Metrics and Evaluations indicated that by 2040 CKD will be the fifth cause of life years lost on a global scale (Foreman et al., 2018).

CKD is defined as decreased renal function [an estimated glomerular filtration rate (eGFR) lower than 60 ml/min/1.73 m²] or the presence of kidney damage, for 3 months or longer, irrespective of the underlying cause (Levey et al., 2011). According to the Kidney Disease

Improving Global Outcomes (KDIGO) guidelines, CKD can be classified into six stages based on eGFR (**Table 1**) and into three categories based on albuminuria (**Table 2**) (KDIGO, 2013). The first stages of CKD are often clinically silent and most patients start developing symptoms in stages 4 and 5, hampering the early diagnosis of CKD.

A positive family history is reported by 24–34% of patients with CKD, and familial clustering is a common phenomenon in patients with end-stage renal disease (ESRD) (Akrawi et al., 2014; Skrunes et al., 2014; Connaughton et al., 2015). This indicates that, during the diagnostic workup of a CKD patient, the possibility of a hereditary cause cannot be neglected. This is supported by the fact that monogenic mutations are frequently found in early-onset CKD and that inherited kidney diseases are a common cause of ESRD in both adults and children (Devuyst et al., 2014).

Approximately 17% of patients with ESRD do not have a primary renal disease (PRD) diagnosis, and are therefore labeled as CKD with unknown origin. Furthermore, in many patients the primary diagnosis is inaccurate (Hoekstra et al., 2017). Making a correct diagnosis in these patients may have therapeutic implications. For example, a genetic mutation predisposing to atypical hemolytic uremic syndrome (aHUS) may provide crucial therapeutic possibilities with the availability of specific monoclonal antibodies targeting the complement system (Zuber et al., 2012). Moreover, a genetic diagnosis may be of pivotal importance for family counseling and in the setting of kidney transplantation, particularly when living-related donation is involved (Aymé et al., 2017).

A recent study in a cohort of CKD patients, of whom 65% had ESRD, demonstrated that a molecular diagnosis can be derived with next-generation sequencing (NGS) in approximately one in ten adults (Groopman et al., 2018). Since unexplained diseases have a higher risk of a genetic origin, it is likely that the proportion of patients in whom a molecular diagnosis can be identified is even larger when considering patients with unexplained CKD, or those with an atypical presentation. However, genetic testing is currently only performed in a minority of these patients.

NGS-based techniques are thus likely to improve the diagnostic accuracy in patients with CKD of unknown origin. It could also improve patient care, since NGS-based testing has the potential to identify the cause of CKD in an early stage of the disease, which allows timely intervention to delay or prevent development of ESRD. However, given the novelty of the technique, the value of NGS-based multi-gene panels and whole exome sequencing for these patients in clinical practice remains to be demonstrated.

In this review, we will discuss the utility of NGS-based testing in CKD patients with childhood-onset and adult-onset disease. Subsequently, we will focus on the potential role of NGS in the diagnostic workup of adult patients with an unknown PRD.

PREVALENCE AND CAUSES OF CHRONIC KIDNEY DISEASE

The global prevalence of CKD is 13.4%, with a higher prevalence in women compared to men and in patients over 70 years of age (Mills et al., 2015; Hill et al., 2016). Data on trends in the prevalence of CKD over the past decades are conflicting (Coresh et al., 2007; Aitken et al., 2014; Murphy et al., 2016), and the prevalence of CKD varies widely across countries (Bruck et al., 2016; de Grauw et al., 2018). For example, in Europe the prevalence estimates for CKD stage 1–5 vary from 3.3% in Norway, to 10.4% in the Netherlands, and 17.3% in north-eastern Germany (Bruck et al., 2016; de Grauw et al., 2018). CKD prevalence may even vary within one country (Zhang et al., 2012). There are many factors which can partially explain the reported differences in prevalence between and within countries. Healthcare organization and policies in prevention and screening programs for CKD and its risk factors may vary substantially between countries. Other factors include, but are not limited to, communicable diseases, variations in laboratory methods, environmental factors like diet and toxins, ethnicity, genetic susceptibility, and differences in the prevalence of CKD risk factors like diabetes, hypertension and obesity (Jha et al., 2013; Agyemang et al., 2016; Stanifer et al., 2016; Glassock et al., 2017; Webster et al., 2017).

While hypertension and diabetes are the leading CKD causes in Western countries, other common causes include glomerulonephritis and inherited forms of CKD (Jha et al., 2013; Mallett et al., 2014). Interestingly, hypertension is not only a cause of CKD, but it can also occur as a result of CKD. Therefore, it can sometimes be difficult to determine which occurred first (Hamrahian and Falkner, 2017). It could thus be possible that some CKD patients with the PRD diagnosis "hypertensive nephropathy" have a different disease etiology, with hypertension being rather a consequence than the primary cause of the disease. In clinical practice, isolated essential hypertension rarely leads to ESRD, in contrast with, for example, malignant hypertension.

TABLE 1 | Estimated glomerular filtration rate stages for chronic kidney disease.

GFR stage	eGFR	Description
1	>90	Normal or high
2	60–89	Mildly decreased
3a	45–59	Mildly to moderately decreased
3b	30–44	Moderately to severely decreased
4	15–29	Severely decreased
5	<15	Kidney failure

GFR, glomerular filtration rate; eGFR, estimated glomerular filtration rate. Based on KDIGO guidelines (KDIGO, 2013).

TABLE 2 | Albuminuria categories for chronic kidney disease.

Albuminuria category	ARC range		Description
	mg/g	mmol/g	
1	<30	<3	Normally to mildly increased
2	30–300	3–30	Moderately increased
3	>300	>30	Severely increased

ACR, albumin-to-creatinine ratio. Based on KDIGO guidelines (KDIGO, 2013).

A hereditary form of CKD can be found in approximately 10% of the adult CKD population, comprising both patients with ESRD and patients with pre-dialysis kidney disease (Devuyst et al., 2014; Mallett et al., 2014). Compared to adults, children with CKD are more likely to have a hereditary form of CKD. Current estimates are that at least 15–20% of early-onset CKD, that is before 25 years of age, is caused by a genetic form. It is also estimated that nearly all children who progress to ESRD have an inherited form of CKD (Devuyst et al., 2014; Vivante and Hildebrandt, 2016). The most common inherited kidney diseases seen in the pediatric and adult populations are likely to be different from each other. For instance, autosomal dominant polycystic kidney disease (ADPKD) usually presents in adulthood, while nephronophthisis is more prevalent in children (Hildebrandt, 2010).

The number of ESRD patients with an unknown PRD diagnosis is increasing. In the Netherlands, the prevalence of ESRD of unknown etiology is 17% (Hoekstra et al., 2017). Similar percentages are found elsewhere in Europe (Kramer et al., 2018). Patients with an unknown PRD would benefit from an earlier diagnosis, which could permit early intervention to reduce the risk of cardiovascular complications and slow down or prevent the progression of CKD.

NEXT GENERATION SEQUENCING

Genetic testing can have an important role in accurately diagnosing PRD in cases that have a higher risk of an inherited form of CKD; for example patients with familial CKD, early-onset CKD, an unusual disease course, or patients with an unknown cause of CKD (Vivante and Hildebrandt, 2016). One of the benefits of genetic testing is that it can identify the cause of disease regardless of disease stage, contrary to diagnostic renal biopsies that often fail to reveal a diagnosis in very early or late stages of the disease (Renkema et al., 2014; Stokman et al., 2016). Moreover, genetic testing is non-invasive and can shorten the diagnostic odyssey. Identification of a genetic diagnosis can play an important role in family planning, options for kidney transplantation, targeted surveillance of extra-renal features and can guide treatment and genetic counseling (Adams and Eng, 2018; Nestor et al., 2018; Armstrong and Thomas, 2019). Counseling and genetic testing can also be of importance for relatives of the patient with genetic CKD. Identification of a genetic risk variant in otherwise healthy relatives can make them eligible for monitoring and early intervention to prevent progression of CKD and its complications, such as cardiovascular disease. In addition, it can exclude the relative for living kidney donation as to reduce the risk of future CKD.

Diagnostic genetic testing has benefited from advances in NGS, which has enabled cost-effective sequencing of large regions of the genome (Bick and Dimmock, 2011; Petersen et al., 2017). NGS, or massively parallel sequencing, allows simultaneous sequencing of multiple genes associated with a particular phenotype (gene panels), all 20,000 protein-coding genes (whole exome-sequencing), or the entire genome (whole genome

sequencing) (Adams and Eng, 2018; Groopman et al., 2018). Each approach has its own advantages and limitations. Gene panels often have a higher sequencing coverage and depth than whole exome sequencing (WES) or whole genome sequencing (WGS), resulting in a greater diagnostic yield (Xue et al., 2015). However, the coverage and depth of NGS-based techniques varies depending on the capture systems used (Chilamakuri et al., 2014; Shigemizu et al., 2015; García-García et al., 2016). Coverage and depth of a system will also change over time as new updates and techniques become available, such as capture free WGS. It is thus important to select a suitable system.

Gene panel testing reduces the risk of incidental findings, being designed to only include a pre-selected subset of genes. Disadvantages of gene panels are that they need to be regularly updated to include newly discovered genes and the inherent requirement of a correctly interpreted clinical context, in order to avoid an inappropriate gene panel to be performed (Prakash and Gharavi, 2015; Groopman et al., 2018).

In contrast with gene panels, WES sequences all the protein-coding genes of the genome and is thus not limited to a preselected set of genes. This allows a more flexible analysis compared to gene panels and the opportunity to identify new causative genes. Furthermore, WES data can be stored for future reanalysis as new genes are discovered and variants are reclassified (Groopman et al., 2018; Jayasinghe et al., 2018). Although WES sequences the entire exome, it is possible to only target a specified subset of genes with an *in silico* panel (targeted WES). This approach gives similar results as gene panels, but has the advantage that the original WES data can be "opened" for further analysis if new genes are discovered or if a causative variant cannot be identified in the initial analysis (Preston et al., 2017; Jayasinghe et al., 2018). Because of these advantages, most of the current diagnostic gene panels are based on targeted WES.

An important issue when performing WES is the possibility of detecting incidental findings, which are causative variants not related to the primary purpose for genetic testing. For example, when a genetic variant that predisposes to cancer is identified in a patient undergoing WES for hereditary kidney disease, this finding could not only have consequences for the patient (for instance intensified cancer screening or preventive surgery), but also for the relatives of the patient. The identification of a genetic risk variant for cancer can also influence treatment. Kidney transplant recipients have an increased risk of developing cancer, which can be partly attributed to immunosuppressive therapy. To minimize the risk of developing cancer, reduction of immunosuppression or other treatment options should be considered in transplant recipients with a genetic predisposition for cancer. There is, however, no consensus about when to report an incidental finding.

Another drawback of WES is that not all genomic regions are equally covered and regions with high guanine-cytosine (GC) rich content, copy number variants, and high sequence homology with pseudogenes may be missed (Xue et al., 2015). For example, WES is of limited use in diagnosing ADPKD, which is caused by mutations in *PKD1* and *PKD2*. The *PKD1* gene has a high degree of sequence homology with six pseudogenes, which complicates variant identification (Ali et al., 2019).

Some of the limitations of WES can be addressed by WGS. For instance, WGS can identify copy number variants and has a more complete per base coverage compared to WES (Belkadi et al., 2015; Wu et al., 2016). WGS has the same advantages as WES, but has the additional benefit of being able to sequencing nearly all types of genetic variations in both the coding and non-coding regions of the genome (Taylor et al., 2015; Wu et al., 2016; Lionel et al., 2018). For instance, WGS was able to identify a genetic variant in 86% of patients with ADPKD. This suggest that WGS can discriminate between the original *PKD1* gene and the pseudogenes (Mallawaarachchi et al., 2016).

Other studies have reported that the diagnostic yield of WGS is higher than WES in a variety of disorders and that WGS can identify a causative variant in 20–40% of the patients in whom no genetic cause could be identified with WES (Gilissen et al., 2014; Ellingford et al., 2016). Nevertheless, WGS is not commonly used in clinical practice. This could be due to the costs and time associated with WGS, the requirements for data analysis and data storage, and the complex interpretation of unknown variants, especially intronic and other non-coding variants. However, it is likely that the capability to interpret variants in noncoding regions of the genome will improve over time. This, together with an expected decline in sequencing costs, will increase the advantages of WGS in the future (Lionel et al., 2018).

It is important to recognize that all forms of NGS-based testing have common limitations. For example, all NGS-based techniques were unable to detect causative variants in *MUC1* in six unrelated families with autosomal dominant tubulointerstitial kidney disease. It was likely missed due to the highly repetitive GC-rich sequence and the variant was only identified by long-range polymerase chain reaction and molecular cloning (Kirby et al., 2013). NGS-based testing is furthermore of limited value in most patients with acquired diseases and the translation of genetic findings to clinical practice may be challenging (Stokman et al., 2016). In addition, the social, ethical, and legal concerns of genetic testing cannot be neglected (Guay-Woodford and Knoers, 2009; Clarke, 2014).

Some limitations of NGS are caused by the short-read lengths that are used to maintain high coverage and depth. A new technique, long-read sequencing, may overcome some NGS-specific limitations. Long-read single molecule platforms offer the ability to directly evaluate many difficult or even previously unsequenceable regions of the genome, such as repetitive elements, non-targeted structural variant breakpoints at base-pair resolution, pseudogene discrimination, and epigenetics (Mantere et al., 2019). In the future, long-read sequencing based WGS could thus increase the diagnostic yield in patients with genetic CKD and give more patients a (correct) PRD diagnosis. Because of the advantages of genetic testing in general, and specifically the advent of NGS in the past decade, genetic testing is becoming increasingly relevant in clinical practice. However, the interpretation of variants remains a challenge with both WES and WGS.

Most studies on genetic testing in the field of nephrology are restricted to a research setting (e.g., cohort studies), and so far little is known about its diagnostic utility in clinical practice.

NEXT-GENERATION SEQUENCING IN CHILDREN WITH CHRONIC KIDNEY DISEASE

The most common causes of renal disease in children are congenital anomalies of the kidney and urinary tract (CAKUT), glomerulonephritis, steroid-resistant nephrotic syndrome (SRNS) and renal ciliopathies (Harambat et al., 2012; Becherucci et al., 2016). These renal diseases can all have a genetic etiology. Reports show that in approximately 20% of adolescents with early-onset CKD a monogenic cause of disease can be identified (Devuyst et al., 2014; Vivante and Hildebrandt, 2016). The high contribution of genetic causes in pediatric CKD is underlined by the fact that almost all children who progress to ESRD have an inherited form of CKD.

Extensive genetic research has been done in the pediatric CKD population, although data obtained in the clinical practice setting remain scarce. For example, a recent study investigated the diagnostic yield of targeted WES in a pediatric kidney transplant recipient cohort. The authors included children irrespective of clinical PRD diagnosis, and identified a causative variant in 34 of 104 (32.7%) of patients (Mann et al., 2019). They found a higher diagnostic yield in patients with a positive family history for CKD, patients with extra-renal manifestation or with a history of consanguinity (Mann et al., 2019).

Other studies examined the diagnostic yield of NGS-based testing in pediatrics cohorts with a specific phenotype. For example, with NGS a genetic cause was identified in 55–80% of patients with Alport syndrome and in 21–25% of patients with nephronophthisis-related ciliopathies (Fallerini et al., 2014; Moriniere et al., 2014; Schueler et al., 2016). In patients with CAKUT, gene panels identified a causative variant in 2.5–6% (Hwang et al., 2014; Kohl et al., 2014), whereas targeted WES identified diagnostic variants in 5–14% of patients with CAKUT (Bekheirnia et al., 2017; van der Ven et al., 2018).

The diagnostic utility of targeted WES has also been addressed in SRNS, where a diagnostic variant was found in 11.1–29.5% of patients (Sadowski et al., 2015; Bierzynska et al., 2017; Tan et al., 2018; Warejko et al., 2018).

Gene panels identified a causative variant in 16.8–20.8% of patients with childhood onset nephrolithiasis, which is significantly higher than the adult-onset setting where a genetic cause was identified in 11.4% (Halbritter et al., 2015; Braun et al., 2016). In comparison, targeted WES in patients with nephrolithiasis with onset before the age of 25 years identified a diagnostic variant in 29.4%. According to the authors, the higher diagnostic yield can be attributed to a younger age of onset, the inclusion or more familial cases, and presence of consanguinity (Daga et al., 2018).

The nephrolithiasis studies show that the clinical characteristics of patients can influence the diagnostic yield. The study design (sample size, number of genes sequenced, and whether patients are recruited from tertiary hospitals or specialized centers) can likewise affect the diagnostic yield. In addition, most studies have been conducted in patients predominantly of European ancestry and the result may therefore not always apply to other ethnic

groups. The results from the mentioned studies can therefore not always be generalized to the broader CKD population.

The data from the abovementioned studies are mostly obtained in a research setting. However, the potential diagnostic utility of targeted WES has also been shown in a clinical practice setting. Offering targeted WES to families of children with familial CKD led to an overall diagnostic yield as high as 46% (Mallett et al., 2017). These findings highlight the diagnostic potential of NGS-based testing in pediatric patients with early-onset CKD. It also underlines the need for further research in a diagnostic setting, to further define the position of NGS-based diagnostics in clinical practice.

NEXT-GENERATION SEQUENCING IN ADULTS WITH CHRONIC KIDNEY DISEASE

Genome-wide association studies (GWAS) and GWAS meta-analyses have identified several common variants in genetic loci associated with CKD, including variants in *UMOD, SHROOM3,* solute carriers, and E3 ubiquitin ligases, as reviewed elsewhere (Cañadas-Garre et al., 2019). Despite the common genetic risk variants and the relatively high prevalence of genetic causes of CKD in adults, the role of NGS diagnostics in clinical practice has been limited compared to the pediatric population. A possible explanation is that genetic testing has traditionally been considered to be more successful in children compared to adults (Mallett et al., 2017). However, targeted WES in familial CKD in a diagnostic setting showed similar diagnostic rates among children (31 of 67 patients, 46.3%) and adults (27 of 68 patients, 39.7%) (Mallett et al., 2017). There were, however, significant differences in diagnostic yield between adults and children for certain phenotypes such as aHUS (17% in adults *vs.* 67% in children) (Mallett et al., 2017).

These findings were confirmed by a recent study from Ireland. Targeted WES showed comparable diagnostic rates among pediatric onset CKD (20 of 50 patients, 40%) and adult onset CKD (35 of 85 patients, 41%) (Connaughton et al., 2019). The similar overall diagnostic rate between adults and children seems in contrast with previous observations that showed an inverse correlation between proband age and the chance of identifying a genetic cause for a specific inherited type of kidney disease (Sadowski et al., 2015).

Subsequent studies have further highlighted the diagnostic potential of NGS in an adult CKD population. For example, targeted WES in a broad CKD population identified diagnostic variants in 307 of 3,315 (9.3%) patients (Groopman et al., 2018). This cohort consisted of 91.6% adults (over 21 years of age) and 64.7% of the patients had ESRD. In the 307 patients with a genetic diagnosis a subsequent analysis was done with a phenotype-specific gene panel. This resulted, at most, in a genetic diagnosis in 136 of 307 (44.3%) patients (Groopman et al., 2018).

Several studies have investigated the diagnostic rate of NGS-based testing in adults with specific renal phenotypes. For example, genetic testing identified causative variants in 14% of adult-onset SRNS and gene panels found diagnostic variants in 11.4% of adult-onset nephrolithiasis. In both studies a lower genetic yield was associated with a higher proband age (Santín et al., 2011; Halbritter et al., 2015).

The diagnostic yield of NGS-based testing in patients with no clear PRD or with a family history of CKD is likely higher than in the overall CKD population. Indeed, a pilot study in adults with familial or unexplained CKD demonstrated that targeted WES identified a causative variant in 22 of 92 (24%) patients. These encompassed 13 different genetic disorders and influenced medical management in 21 of 22 patients (Lata et al., 2018).

As mentioned before, the results from these studies cannot always be generalized to the broader CKD population due to ascertainment bias. For instance, the authors of the pilot study mention that the generalizability of the results to the general CKD population is limited by a small sample size and the fact that the cohort is enriched for familial or suspected genetic CKD (Lata et al., 2018).

These studies confirm the importance of NGS-based testing in adults with CKD. Increased awareness for the possibility of genetic testing in adults could help to provide more patients an early and specific diagnosis, aiming to prevent progression of CKD and its complications. However, the diagnostic yield of genetic testing may be lower in patients who are older at disease onset. More research is needed to determine the effect of age at onset on the yield of genetic testing in the context of CKD; such studies will direct guidelines towards cost-effective implementation of NGS-based testing. Moreover, the diagnostic yield in CKD is currently limited by the fact that in many cases the underlying cause is polygenic in nature; the development of polygenic risk scores may further improve diagnostic yield in the future.

NEXT-GENERATION SEQUENCING TO RECLASSIFY THE PRIMARY RENAL DIAGNOSIS

Most inherited kidney diseases have a high degree of genetic heterogeneity and are associated with a wide range of phenotypes (Kalatharan et al., 2018). This can make it challenging to distinguish certain kidney diseases based on clinical presentation, and may lead to misdiagnosis.

The emergence of NGS has enabled a more precise differentiation between overlapping phenotypes based on a genetic diagnosis. In some cases, the genetic diagnosis can lead to reclassification of the PRD diagnosis in individual patients. For example, it may be difficult to discriminate between Alport syndrome, mesangial proliferative glomerulonephritis (MPGN), and primary focal segmental glomerulosclerosis (FSGS) based on clinical presentation and renal biopsies (Yao et al., 2012). However, the distinction between these diseases is of great importance with respect to treatment and clinical outcomes. Genetic studies have shown that patients with Alport syndrome have been misdiagnoses as MPGN (Adam et al., 2014). In addition, multiple studies have shown that patients with FSGS were reclassified as Alport syndrome based on mutations found

in *COL4A3, COL4A4,* and *COL4A5* (Adam et al., 2014; Gast et al., 2016; Yao et al., 2019). Genetic testing can distinguish between different etiologies in heterogeneous diseases and prevent an incorrect diagnosis.

Another example of reclassification has been illustrated in a report describing the use of WES to obtain a genetic diagnosis (Choi et al., 2009). In a cohort of patients with suspected Bartter syndrome, WES identified no candidate variants in Bartter-associated genes but instead found a variation in *SLC26A3*. Variants in this gene are associated with chloride diarrhea, and clinical follow-up confirmed the diagnosis of chloride diarrhea in these patients (Choi et al., 2009).

Moreover, in a recently published analysis from the iGeneTrain consortium, it was found that 0.5% of patients with adult-onset ESRD had a full gene deletion of the *NPHP1* gene. This percentage increased to almost 1% when only the patients of 18–50 years at ESRD onset were considered (n = 24). Full gene deletions of *NPHP1* are associated with nephronophthisis, the most common genetic cause for ESRD in children. Of the 26 patients with a full *NPHP1* gene deletion, only 3 had a correct clinical PRD diagnosis of nephronophthisis (Snoek et al., 2018). This is not surprising, since nephronophthisis had previously been considered to result in ESRD exclusively at child age.

Other studies have also reported on the reclassification of PRD diagnosis based on genetic findings, albeit in cohorts with limited sample size. For example, Lata et al. identified a genetic diagnosis in 22 of 92 adults with CKD, and in two of these patients the kidney disease was reclassified. The reclassification of diagnosis, from FSGS to autosomal dominant Alport syndrome and to autosomal recessive Alport syndrome, had direct consequence for clinical management (Lata et al., 2018). Furthermore, in a study by Connaughton et al. targeted WES identified a molecular diagnosis in 42 of 114 families with CKD, and in 9 of these 42 families the initial CKD diagnosis was changed based on genetic findings (Connaughton et al., 2019).

Similarly, in a pediatric cohort consisting of 79 consanguineous or familial cases with a suspected diagnosis of nephronophthisis, targeted WES identified causal variants in 63% of cases. While in 64% of these families the suspected diagnosis was confirmed, in a noteworthy 36% of these families a different molecular diagnosis was found. The CKD diagnoses in these families were reclassified as Alport syndrome (8%), CAKUT (6%), renal tubulopathies (16%), autosomal-recessive polycystic kidney disease (4%), and auto-immune nephropathy (2%) (Braun et al., 2016). These finding show that (targeted) WES can provide a correct diagnosis in genetically and phenotypically heterogeneous diseases. The above findings demonstrate that it is very likely that a subset of CKD patients has a misdiagnosis and that with the increased availability of genetic testing more primary diagnoses will be reclassified.

NEXT-GENERATION SEQUENCING IN UNEXPLAINED CHRONIC KIDNEY DISEASE

WES facilitated the molecular diagnosis and PRD reclassification in both pediatric and adult cohorts and informed prognosis,

therapy, and counseling in these patients. Moreover, WES has the potential to resolve cases with unknown etiology (Need et al., 2012; Hauer et al., 2018) and could thus be a valuable analytical tool in the diagnostic process of patients with an unknown PRD or in patients who have an atypical/nonspecific clinical presentation.

The diagnostic value of WES in CKD patients with unknown etiology is illustrated by a pilot study by Lata et al, including 16 patients who had CKD with unknown etiology (Lata et al., 2018). A causative variant was identified in nine of these patients, indicating a diagnostic yield of 56%. The established diagnoses consisted of a variety of nephropathies, including CHARGE syndrome, X-linked and autosomal forms of Alport syndrome, *HNF1B*-associated disease, Dent's disease, and autosomal dominant tubulointerstitial kidney disease (**Table 3**). In addition to providing a kidney disease diagnosis for these patients, the genetic diagnosis influenced medical management. Clinical implications included a change in therapy (for example avoidance of immunosuppressive therapy), screening of at-risk family members, screening for extra-renal features, and optimal selection of living related kidney donors (Lata et al., 2018).

Another study in which the diagnostic value of WES in patients with undiagnosed CKD is demonstrated is the recently published study of Groopman et al. Targeted WES provided a molecular diagnosis in 48 of 281 (17.1%) patients with CKD of unknown origin. The diagnostic yield in this group was the second highest in this study, the yield was only higher in patients with a clinical diagnosis of congenital or cystic renal disease (23.9%) (**Table 3**) (Groopman et al., 2018). In addition, the authors found that nephropathy of unknown origin, a family history of CKD, and clinical diagnosis of congenital or cystic renal disease were independent predictors of yielding a genetic diagnosis (Groopman et al., 2018).

A recent study by Connaughton et al. also reported on the diagnostic value of targeted WES in patients with unexplained CKD (Connaughton et al., 2019). A causative variant was identified in 16 of 34 (47%) families with CKD of unknown origin. The newfound diagnoses consisted, among others, of X-linked and autosomal forms of Alport syndrome, Dent's disease, Fanconi anemia, Wolfram-like syndrome, nephronophthisis, and nephrocalcinosis/nephrolithiasis. The diagnostic yield in the unexplained CKD family group represented 38% of the identified molecular diagnosis in this study (16 of 42 families) (Connaughton et al., 2019) (**Table 3**).

The diagnostic value of gene panels in patients with undetermined ESRD is demonstrated by Ottlewksi *et al*, who tested 50 patients on the kidney transplantation waiting list with undetermined CKD (Ottlewski et al., 2019). The renal gene panel, consisting of 209 genes associated with ESRD, identified a causative variant in 6 of 50 (12%) these patients. The genetic diagnosis consisted of X-linked and autosomal dominant forms of Alport syndrome and familial FSGS. The identification of a genetic diagnosis in six patients significantly reduced the proportion of patients with unexplained ESRD on the kidney transplant waiting list (Ottlewski et al., 2019) (**Table 3**).

TABLE 3 | Overview of studies on the diagnostic yield of next-generation sequencing-based testing in patients with unexplained chronic kidney disease.

	N sequenced, unexplained/ total cohort	Age at presentation	Patients characteristics	Positive family history, n (%)	New (genetic) diagnoses	Number of genes in NGS panel	Yield in unexplained CKD, %
(Lata et al., 2018)	16/92	37 (SD 11) years	Recruited from tertiary care medical center, predominantly Caucasian	9 (56%)	CHARGE syndrome, DENT's disease, AS, *HNF1B* associated disease, ADTKD, PARN mutations (n = 3)	287	56
(Groopman et al., 2018)	281/3315	At inclusion 91.6% of patients is older than 21 years*	Recruited from the AURORA study (Fellström et al., 2009) and a tertiary care medical center, more than 60% has ESRD, 35.6% non-European ancestry**	Available for 2187 patients: 619 (28.3%)†	NA	625	17.1
(Connaughton et al., 2019)	34/114 families	36% childhood onset, 62% adult onset*	Recruited from nephrology services in Ireland, predominantly Caucasian, included mostly patients with positive family history or extra-renal features, 65% has ESRD**	33 (80%)	Cystic kidney disease or NPHP, syndromic CAKUT, AS, tubulointerstitial kidney disease, hypertensive nephropathy, nephrocalcinosis/ nephrolithiasis	478	47
(Ottlewski et al., 2019)	50/135	Median age at ESRD 43.3 (range 15.4–66.5)	Recruited from waitlist for kidney transplantation at a tertiary care medical center, predominantly Caucasian	NA	AS, familial FSGS	209	12

ADTKD, autosomal dominant tubulointerstitial kidney disease; AS, Alport syndrome; CHARGE, coloboma, heart defects, atresia choanae, growth retardation, genital abnormalities and ear abnormalities; CKD, chronic kidney disease; ESRD, end-stage renal disease; FSGS, focal segmental glomerulosclerosis; NA, not available; NPHP, nephronophthisis; SD, standard deviation.

*Age at presentation from the total cohort, not specific for unexplained CKD subset.

**Patient characteristics from the total cohort, not specified for unexplained CKD subset.

†Number of patients with a positive family history for CKD in the total cohort, not specified for unexplained CKD subset.

The dissimilarities in diagnostic yield between these studies likely result from differences in sample size, different inclusion criteria, NGS approach, and different selection of genes. The study of Lata et al. included only patients with familial or suspected genetic CKD (Lata et al., 2018). This enrichment for genetic diagnosis could explain, together with the limited sample size, the high yield. Groopman et al. included CKD patients irrespective of diagnosis or family history (Groopman et al., 2018), while Connaughton et al. included mostly CKD patients with a positive family history or extra-renal features and only a small proportion of CKD patients who had neither familial CKD nor extra-renal features (Connaughton et al., 2019). The study by Ottlewski et al. also has a limited sample size, but it is not clear what the proportion of patients with familial CKD or extra-renal features is. However, five of six patients in which a genetic diagnosis was identified had a positive family history (Ottlewski et al., 2019). In addition, it is the only study using gene panels instead of targeted WES. The diagnostic yield is thus dependent on the characteristics of the included patients and the study design. This can make it difficult to make comparisons between different studies, especially when some patient characteristic are unknown.

The findings above highlight the diagnostic value of NGS based testing for patients with unexplained CKD, especially in patients with early-onset or familial CKD. Further research is needed to explore the diagnostic utility of WES for CKD of unknown etiology in larger cohorts and in a clinical setting. Such studies should guide the position of NGS diagnostics in the work-up of patients with unknown PRD, by revealing subpopulations with the highest diagnostic yield that may be preferentially offered this diagnostic testing.

CONCLUSION

NGS, including gene panels and WES, shows promising results as a diagnostic tool in both pediatric and adult CKD cohorts. In a considerable proportion of patients, particularly in those with familial CKD (~40%), NGS has the potential to resolve CKD cases with an unknown etiology. NGS also enables the reclassification of PRD diagnoses. Counseling on the potential implications of a positive test result is essential, particularly for WES, given the possibility of incidental findings. Although further research is needed to determine which subpopulations will have the highest

diagnostic yield, we foresee an expanding role for NGS-based diagnostics in clinical nephrology in the coming years.

AUTHOR CONTRIBUTIONS

AH wrote the first draft of the manuscript. ME, LV, NK, and MB gave feedback and contributed to manuscript revision. All authors read and approved the submitted version.

REFERENCES

Adam, J., Connor, T. M. F., Wood, K., Lewis, D., Naik, R., Gale, D. P., et al. (2014). Genetic testing can resolve diagnostic confusion in Alport syndrome. *Clin. Kidney J.* 7, 197–200. doi: 10.1093/ckj/sft144

Adams, D. R., and Eng, C. M. (2018). Next-generation sequencing to diagnose suspected genetic disorders. *N. Engl. J. Med.* 379, 1353–1362. doi: 10.1056/NEJMra1711801

Agyemang, C., Snijder, M. B., Adjei, D. N., Van Den Born, B. J. H., Modesti, P. A., Peters, R. J. G., et al. (2016). Ethnic disparities in CKD in the Netherlands: The healthy life in an urban setting (HELIUS) Study. *Am. J. Kidney Dis.* 67, 391–399. doi: 10.1053/j.ajkd.2015.07.023

Aitken, G. R., Roderick, P. J., Fraser, S., Mindell, J. S., O'Donoghue, D., Day, J., et al. (2014). Change in prevalence of chronic kidney disease in England over time: comparison of nationally representative cross-sectional surveys from 2003 to 2010. *BMJ Open* 4, e005480. doi: 10.1136/bmjopen-2014-005480

Akrawi, D. S., Li, X., Sundquist, J., Sundquist, K., and Zöller, B. (2014). Familial risks of kidney failure in Sweden: a nationwide family study. *PloS One* 9, e113353. doi: 10.1371/journal.pone.0113353

Ali, H., Al-Mulla, F., Hussain, N., Naim, M., Asbeutah, A. M., AlSahow, A., et al. (2019). PKD1 Duplicated regions limit clinical utility of whole exome sequencing for genetic diagnosis of autosomal dominant polycystic kidney disease. *Sci. Rep.* 9, 4141. doi: 10.1038/s41598-019-40761-w

Armstrong, M., and Thomas, C. (2019). Diagnosis of monogenic chronic kidney diseases. *Curr. Opin. Nephrol. Hypertens.* 28, 183–194. doi: 10.1097/MNH.0000000000000486

Aymé, S., Bockenhauer, D., Day, S., Devuyst, O., Guay-Woodford, L. M., Ingelfinger, J. R., et al. (2017). Common elements in rare kidney diseases: conclusions from a kidney disease: improving global outcomes (KDIGO) controversies conference. *Kidney Int.* 92, 796–808. doi: 10.1016/j.kint.2017.06.018

Baumeister, S. E., Böger, C. A., Krämer, B. K., Döring, A., Eheberg, D., Fischer, B., et al. (2010). Effect of chronic kidney disease and comorbid conditions on health care costs: a 10-year observational study in a general population. *Am. J. Nephrol.* 31, 222–229. doi: 10.1159/000272937

Becherucci, F., Roperto, R. M., Materassi, M., and Romagnani, P. (2016). Chronic kidney disease in children. *Clin. Kidney J.* 9, 583–591. doi: 10.1093/ckj/sfw047

Bekheirnia, M. R., Bekheirnia, N., Bainbridge, M. N., Gu, S., Akdemir, Z. H. C., Gambin, T., et al. (2017). Whole-exome sequencing in the molecular diagnosis of individuals with congenital anomalies of the kidney and urinary tract and identification of a new causative gene. *Genet. Med.* 19, 412–420. doi: 10.1038/gim.2016.131

Belkadi, A., Bolze, A., Itan, Y., Cobat, A., Vincent, Q. B., Antipenko, A., et al. (2015). Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. *Proc. Natl. Acad. Sci. U.S.A.* 112, 5473–5478. doi: 10.1073/pnas.1418631112

Bick, D., and Dimmock, D. (2011). Whole exome and whole genome sequencing. *Curr. Opin. Pediatr.* 23, 594–600. doi: 10.1097/MOP.0b013e32834b20ec

Bierzynska, A., McCarthy, H. J., Soderquest, K., Sen, E. S., Colby, E., Ding, W. Y., et al. (2017). Genomic and clinical profiling of a national nephrotic syndrome cohort advocates a precision medicine approach to disease management. *Kidney Int.* 91, 937–947. doi: 10.1016/j.kint.2016.10.013

Braun, D. A., Lawson, J. A., Gee, H. Y., Halbritter, J., Shril, S., Tan, W., et al. (2016). Prevalence of monogenic causes in pediatric patients with nephrolithiasis or nephrocalcinosis. *Clin. J. Am. Soc Nephrol.* 11, 664–672. doi: 10.2215/CJN.07540715

Braun, D. A., Schueler, M., Halbritter, J., Gee, H. Y., Porath, J. D., Lawson, J. A., et al. (2016). Whole exome sequencing identifies causative mutations in the majority of consanguineous or familial cases with childhood-onset increased renal echogenicity. *Kidney Int.* 89, 468–475. doi: 10.1038/ki.2015.317

Bruck, K., Stel, V. S., Gambaro, G., Hallan, S., Volzke, H., Arnlov, J., et al. (2016). CKD prevalence varies across the european general population. *J. Am. Soc Nephrol.* 27, 2135–2147. doi: 10.1681/ASN.2015050542

Cañadas-Garre, M., Anderson, K., Cappa, R., Skelly, R., Smyth, L. J., McKnight, A. J., et al. (2019). Genetic susceptibility to chronic kidney disease - some more pieces for the heritability puzzle. *Front. Genet.* 10, 453. doi: 10.3389/fgene.2019.00453

Chilamakuri, C. S. R., Lorenz, S., Madoui, M.-A., Vodák, D., Sun, J., Hovig, E., et al. (2014). Performance comparison four exome capture systems for deep sequencing. *BMC Genomics* 15, 449. doi: 10.1186/1471-2164-15-449

Choi, M., Scholl, U. I., Ji, W., Liu, T., Tikhonova, I. R., Zumbo, P., et al. (2009). Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc. Natl. Acad. Sci.* 106, 19096–19101. doi: 10.1073/pnas.0910672106

Clarke, A. J. (2014). Managing the ethical challenges of nextgeneration sequencing in genomic medicine. *Br. Med. Bull.* 111, 17–30. doi: 10.1093/bmb/ldu017

Connaughton, D. M., Bukhari, S., Conlon, P., Cassidy, E., O'Toole, M., Mohamad, M., et al. (2015). The irish kidney gene project - prevalence of family history in patients with kidney disease in Ireland. *Nephron* 130, 293–301. doi: 10.1159/000436983

Connaughton, D. M., Kennedy, C., Shril, S., Mann, N., Murray, S. L., Williams, P. A., et al. (2019). Monogenic causes of chronic kidney disease in adults. *Kidney Int.* S0085-2538 (18), 30839–30831. doi: 10.1016/j.kint.2018.10.031

Coresh, J., Selvin, E., Stevens, L. A., Manzi, J., Kusek, J. W., Eggers, P., et al. (2007). Prevalence of chronic kidney disease in the United States. *J. Am. Med. Assoc.* 298, 2038–2047. doi: 10.1001/jama.298.17.2038

Daga, A., Majmundar, A. J., Braun, D. A., Gee, H. Y., Lawson, J. A., Shril, S., et al. (2018). Whole exome sequencing frequently detects a monogenic cause in early onset nephrolithiasis and nephrocalcinosis. *Kidney Int.* 93, 204–213. doi: 10.1016/j.kint.2017.06.025

de Grauw, W. J. C., Scherpbier de Haan, N. D., and van Dipten, C. (2018). Classification of chronic kidney disease 10 years on: what have we learnt in the Netherlands. *Fam. Pract.* 35, 744–745. doi: 10.1093/fampra/cmy036

Devuyst, O., Knoers, N. V. A. M., Remuzzi, G., and Schaefer, F. (2014). Rare inherited kidney diseases: challenges, opportunities, and perspectives. *Lancet* 383, 1844–1859. doi: 10.1016/S0140-6736(14)60659-0

Ellingford, J. M., Barton, S., Bhaskar, S., Williams, S. G., Sergouniotis, P. I., O'Sullivan, J., et al. (2016). Whole genome sequencing increases molecular diagnostic yield compared with current diagnostic testing for inherited retinal disease. *Ophthalmology* 123, 1143–1150. doi: 10.1016/j.ophtha.2016.01.009

Eriksson, J. K., Neovius, M., Jacobson, S. H., Elinder, C. G., and Hylander, B. (2016). Healthcare costs in chronic kidney disease and renal replacement therapy: a population-based cohort study in Sweden. *BMJ Open* 6, e012062. doi: 10.1136/bmjopen-2016-012062

Fallerini, C., Dosa, L., Tita, R., Del Prete, D., Feriozzi, S., Gai, G., et al. (2014). Unbiased next generation sequencing analysis confirms the existence of autosomal dominant Alport syndrome in a relevant fraction of cases. *Clin. Genet.* 86, 252–257. doi: 10.1111/cge.12258

Fellström, B. C., Jardine, A. G., Schmieder, R. E., Holdaas, H., Bannister, K., Beutler, J., et al. (2009). Rosuvastatin and cardiovascular events in patients undergoing hemodialysis. *N Engl. J. Med.* 362, 1450. doi: 10.1056/NEJMoa0810177

Foreman, K. J., Marquez, N., Dolgert, A., Fukutaki, K., Fullman, N., McGaughey, M., et al. (2018). Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: reference and alternative scenarios for 2016–40 for 195 countries and territories. *Lancet* 392, 2052–2090. doi: 10.1016/S0140-6736(18)31694-5

Gansevoort, R. T., Correa-Rotter, R., Hemmelgarn, B. R., Jafar, T. H., Heerspink, H. J. L., Mann, J. F., et al. (2013). Chronic kidney disease and cardiovascular risk: epidemiology, mechanisms, and prevention. *Lancet* 382, 339–352. doi: 10.1016/S0140-6736(13)60595-4

García-García, G., Baux, D., Faugère, V., Moclyn, M., Koenig, M., Claustres, M., et al. (2016). Assessment of the latest NGS enrichment capture methods in clinical context. *Sci. Rep.* 6, 20948. doi: 10.1038/srep20948

Gast, C., Pengelly, R. J., Lyon, M., Bunyan, D. J., Seaby, E. G., Graham, N., et al. (2016). Collagen (COL4A) mutations are the most frequent mutations underlying adult focal segmental glomerulosclerosis. *Nephrol. Dial Transplant.* 31, 961–970. doi: 10.1093/ndt/gfv325

Gilissen, C., Hehir-Kwa, J. Y., Thung, D. T., van de Vorst, M., van Bon, B. W. M., Willemsen, M. H., et al. (2014). Genome sequencing identifies major causes of severe intellectual disability. *Nature* 511, 344–347. doi: 10.1038/nature13394

Glassock, R. J., Warnock, D. G., and Delanaye, P. (2017). The global burden of chronic kidney disease: Estimates, variability and pitfalls. *Nat. Rev. Nephrol.* 13, 104–114. doi: 10.1038/nrneph.2016.163

Go, A. S., Chertow, G. M., Fan, D., McCulloch, C. E., and Hsu, C. (2004). Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. *N. Engl. J. Med.* 351, 1296–1305. doi: 10.1056/NEJMoa041031

Groopman, E., Marasa, M., Cameron-Christie, S., Petrovski, S., Aggarwal, V., Milo-Rasouly, H., et al. (2018). Diagnostic utility of exome sequencing for kidney disease. *N. Engl. J. Med.* 380, 142–151. doi: 10.1056/nejmoa1806891

Groopman, E. E., Rasouly, H. M., and Gharavi, A. G. (2018). Genomic medicine for kidney disease. *Nat. Rev. Nephrol.* 14, 84–104. doi: 10.1038/nrneph.2017.167

Guay-Woodford, L. M., and Knoers, N. V. A. M. (2009). Genetic testing: considerations for pediatric nephrologists. *Semin. Nephrol.* 29, 338–348. doi: 10.1016/j.semnephrol.2009.03.010

Halbritter, J., Baum, M., Hynes, A., Rice, S., Thwaites, D., Gucev, Z., et al. (2015). Fourteen monogenic genes account for 15% of nephrolithiasis/nephrocalcinosis. *J. Am. Soc Nephrol.* 26, 543–551. doi: 10.1681/ASN.2014040388

Hamrahian, S., and Falkner, B. (2017). Hypertension in chronic kidney disease. *Adv. Exp. Med. Biol.* 956, 307–325. doi: 10.1007/5584_2016_84

Harambat, J., van Stralen, K. J., Kim, J. J., and Tizard, E. J. (2012). Epidemiology of chronic kidney disease in children. *Pediatr. Nephrol.* 27, 363–373. doi: 10.1007/s00467-011-1979-6

Hauer, N. N., Popp, B., Schoeller, E., Schuhmann, S., Heath, K. E., Hisado-Oliva, A., et al. (2018). Clinical relevance of systematic phenotyping and exome sequencing in patients with short stature. *Genet. Med.* 20, 630–638. doi: 10.1038/gim.2017.159

Hildebrandt, F. (2010). Genetic kidney diseases. *Lancet* 375, 1287–1295. doi: 10.1016/S0140-6736(10)60236-X

Hill, N. R., Fatoba, S. T., Oke, J. L., Hirst, J. A., O'Callaghan, C. A., Lasserson, D. S., et al. (2016). Global prevalence of chronic kidney disease -a systematic review and meta-analysis. *PloS One* 11, e0158765. doi: 10.1371/journal.pone.0158765

Hoekstra, T., Hemmelder, M. H., and van Ittersum, F. J. (2017). *RENINE Annual Report 2015.* Nederlandse Federatie voor Nefrologie.

Hwang, D. Y., Dworschak, G. C., Kohl, S., Saisawat, P., Vivante, A., Hilger, A. C., et al. (2014). Mutations in 12 known dominant disease-causing genes clarify many congenital anomalies of the kidney and urinary tract. *Kidney Int.* 85, 1429–1433. doi: 10.1038/ki.2013.508

Jager, K. J., Kovesdy, C., Langham, R., Rosenberg, M., Jha, V., and Zoccali, C. (2019). A single number for advocacy and communication-worldwide more than 850 million individuals have kidney diseases. *Nephrol Dial Transplant* 34, 1803–1805. doi: 10.1093/ndt/gfz174

Jayasinghe, K., Quinlan, C., Stark, Z., Patel, C., Mallawaarachchi, A., Wardrop, L., et al. (2018). Renal genetics in Australia: kidney medicine in the genomic age. *Nephrology* 24, 279–286. doi: 10.1111/nep.13494

Jha, V., Garcia-Garcia, G., Iseki, K., Li, Z., Naicker, S., Plattner, B., et al. (2013). Chronic kidney disease: global dimension and perspectives. *Lancet* 382, 260–272. doi: 10.1016/S0140-6736(13)60687-X

Kalatharan, V., Lemaire, M., and Lanktree, M. B. (2018). Opportunities and challenges for genetic studies of end-stage renal disease in canada. *Can. J. Kidney Heal Dis.* 5, 2054358118789368. doi: 10.1177/2054358118789368

KDIGO (2013). KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney Int. Suppl.* 3, 1–150. doi: 10.1038/kisup.2012.76

Kirby, A., Gnirke, A., Jaffe, D. B., Barešová, V., Pochet, N., Blumenstiel, B., et al. (2013). Mutations causing medullary cystic kidney disease type 1 lie in a large VNTR in MUC1 missed by massively parallel sequencing. *Nat. Genet.* 45, 299–303. doi: 10.1038/ng.2543

Kohl, S., Hwang, D.-Y., Dworschak, G. C., Hilger, A. C., Saisawat, P., Vivante, A., et al. (2014). Mild recessive mutations in six fraser syndrome-related genes cause isolated congenital anomalies of the kidney and urinary tract. *J. Am. Soc Nephrol.* 25, 1917–1922. doi: 10.1681/ASN.2013101103

Kramer, A., Pippias, M., Noordzij, M., Stel, V. S., Afentakis, N., Ambühl, P. M., et al. (2018). The european renal association – european dialysis and transplant association (ERA-EDTA) registry annual report 2015: a summary. *Clin. Kidney J.* 11, 108–122. doi: 10.1093/ckj/sfx149

Lata, S., Marasa, M., Li, Y., Fasel, D. A., Groopman, E., Jobanputra, V., et al. (2018). Whole-exome sequencing in adults with chronic kidney disease: a pilot study. *Ann. Intern. Med.* 168, 100–109. doi: 10.7326/M17-1319

Levey, A. S., De Jong, P. E., Coresh, J., Nahas, M. E., Astor, B. C., Matsushita, K., et al. (2011). The definition, classification, and prognosis of chronic kidney disease: A KDIGO Controversies Conference report. *Kidney Int.* 80, 17–28. doi: 10.1038/ki.2010.483

Lionel, A. C., Costain, G., Monfared, N., Walker, S., Reuter, M. S., Hosseini, S. M., et al. (2018). Improved diagnostic yield compared with targeted gene sequencing panels suggests a role for whole-genome sequencing as a first-tier genetic test. *Genet Med.* 20, 435–443, doi: 10.1038/gim.2017.119

Mallawaarachchi, A. C., Hort, Y., Cowley, M. J., McCabe, M. J., Minoche, A., Dinger, M. E., et al. (2016). Whole-genome sequencing overcomes pseudogene homology to diagnose autosomal dominant polycystic kidney disease. *Eur. J. Hum. Genet.* 24, 1584–1590. doi: 10.1038/ejhg.2016.48

Mallett, A., Patel, C., Salisbury, A., Wang, Z., Healy, H., and Hoy, W. (2014). The prevalence and epidemiology of genetic renal disease amongst adults with chronic kidney disease in Australia. *Orphanet J. Rare Dis.* 9, 98. doi: 10.1186/1750-1172-9-98

Mallett, A. J., McCarthy, H. J., Ho, G., Holman, K., Farnsworth, E., Patel, C., et al. (2017). Massively parallel sequencing and targeted exomes in familial kidney disease can diagnose underlying genetic disorders. *Kidney Int.* 92, 1493–1506. doi: 10.1016/j.kint.2017.06.013

Mann, N., Braun, D., Amann, K., Tan, W., Shril, S., Connaughton, D., et al. (2019). Whole-exome sequencing enables a precision medicine approach for kidney transplant recipients. *J. Am. Soc Nephrol.* 30, 201–215. doi: 10.1681/ASN.2018060575

Mantere, T., Kersten, S., and Hoischen, A. (2019). Long-read sequencing emerging in medical genetics. *Front. Genet.* 10, 426. doi: 10.3389/fgene.2019.00426

Mills, K. T., Xu, Y., Zhang, W., Bundy, J. D., Chen, C. S., Kelly, T. N., et al. (2015). A systematic analysis of worldwide population-based data on the global burden of chronic kidney disease in 2010. *Kidney Int.* 88, 950–957. doi: 10.1038/ki.2015.230

Moriniere, V., Dahan, K., Hilbert, P., Lison, M., Lebbah, S., Topa, A., et al. (2014). Improving mutation screening in familial hematuric nephropathies through next generation sequencing. *J. Am. Soc Nephrol.* 25, 2740–2751. doi: 10.1681/ASN.2013080912

Murphy, D., McCulloch, C. E., Lin, F., Banerjee, T., Bragg-Gresham, J. L., Eberhardt, M. S., et al. (2016). Trends in prevalence of chronic kidney disease in the United States. *Ann. Intern. Med.* 165, 473–481. doi: 10.7326/M16-0273

Need, A. C., Shashi, V., Hitomi, Y., Schoch, K., Shianna, K. V., McDonald, M. T., et al. (2012). Clinical application of exome sequencing in undiagnosed genetic conditions. *J. Med. Genet.* 49, 353–361. doi: 10.1136/jmedgenet-2012-100819

Nestor, J. G., Groopman, E. E., and Gharavi, A. G. (2018). Towards precision nephrology: the opportunities and challenges of genomic medicine. *J. Nephrol.* 31, 46–60. doi: 10.1007/s40620-017-0448-0

Nguyen, N. T. Q., Cockwell, P., Maxwell, A. P., Griffin, M., O'Brien, T., and O'Neill, C. (2018). Chronic kidney disease, health-related quality of life and their associated economic burden among a nationally representative sample of community dwelling adults in England. *PloS One* 13, e0207960. doi: 10.1371/journal.pone.0207960

Ottlewski, I., Münch, J., Wagner, T., Schönauer, R., Bachmann, A., Weimann, A., et al. (2019). Value of renal gene panel diagnostics in adults waiting for kidney transplantation due to undetermined end-stage renal disease. *Kidney Int.* 96, 222–230. doi: 10.1016/j.kint.2019.01.038

Petersen, B. S., Fredrich, B., Hoeppner, M. P., Ellinghaus, D., and Franke, A. (2017). Opportunities and challenges of whole-genome and -exome sequencing. *BMC Genet.* 18, 14. doi: 10.1186/s12863-017-0479-5

Prakash, S., and Gharavi, A. G. (2015). Diagnosing kidney disease in the genetic era. *Curr. Opin. Nephrol. Hypertens.* 24, 380–387. doi: 10.1097/MNH.0000000000000139

Preston, R., Stuart, H. M., and Lennon, R. (2017). Genetic testing in steroid-resistant nephrotic syndrome: why, who, when and how? *Pediatr. Nephrol.* 34, 195–210. doi: 10.1007/s00467-017-3838-6

Renkema, K. Y., Stokman, M. F., Giles, R. H., and Knoers, N. V. A. M. (2014). Next-generation sequencing for research and diagnostics in kidney disease. *Nat. Rev. Nephrol.* 10, 433–444. doi: 10.1038/nrneph.2014.95

Sadowski, C. E., Lovric, S., Ashraf, S., Pabst, W. L., Gee, H. Y., Kohl, S., et al. (2015). A single-gene cause in 29.5% Of cases of steroid-resistant nephrotic syndrome. *J. Am. Soc Nephrol.* 26, 1279–1289. doi: 10.1681/ASN.2014050489

Santín, S., Bullich, G., Tazón-Vega, B., García-Maset, R., Giménez, I., Silva, I., et al. (2011). Clinical utility of genetic testing in children and adults with steroid-resistant nephrotic syndrome. *Clin. J. Am. Soc Nephrol.* 6, 1139–1148. doi: 10.2215/cjn.05260610

Schueler, M., Halbritter, J., Phelps, I. G., Braun, D. A., Otto, E. A., Porath, J. D., et al. (2016). Large-scale targeted sequencing comparison highlights extreme genetic heterogeneity in nephronophthisis-related ciliopathies. *J. Med. Genet.* 53, 208–214. doi: 10.1136/jmedgenet-2015-103304

Shigemizu, D., Momozawa, Y., Abe, T., Morizono, T., Boroevich, K. A., Takata, S., et al. (2015). Performance comparison of four commercial human whole-exome capture platforms. *Sci. Rep.* 5, 12742. doi: 10.1038/srep12742

Skrunes, R., Svarstad, E., Reisæter, A. V., and Vikse, B. E. (2014). Familial clustering of esrd in the norwegian population. *Clin. J. Am. Soc Nephrol.* 9, 1692–1700. doi: 10.2215/CJN.01680214

Smith, D. H., Gullion, C. M., Nichols, G., Keith, D. S., and Brown, J. B. (2004). Cost of medical care for chronic kidney disease and comorbidity among enrollees in a large HMO population. *J. Am. Soc Nephrol.* 15, 1300–1306. doi: 10.1097/01.ASN.0000125670.64996.BB

Snoek, R., van Setten, J., Keating, B. J., Israni, A. K., Jacobson, P. A., Oetting, W. S., et al. (2018). NPHP1 (Nephrocystin-1) Gene Deletions Cause Adult-Onset ESRD. *J. Am. Soc. Nephrol.* 29, 1772–1779. doi: 10.1681/ASN.2017111200

Stanifer, J. W., Muiru, A., Jafar, T. H., and Patel, U. D. (2016). Chronic kidney disease in low- and middle-income countries. *Nephrol. Dial Transplant.* 31, 868–874. doi: 10.1093/ndt/gfv466

Stokman, M. F., Renkema, K. Y., Giles, R. H., Schaefer, F., Knoers, N. V. A. M., and Van Eerde, A. M. (2016). The expanding phenotypic spectra of kidney diseases: Insights from genetic studies. *Nat. Rev. Nephrol.* 12, 472–483. doi: 10.1038/nrneph.2016.87

Tan, W., Lovric, S., Ashraf, S., Rao, J., Schapiro, D., Airik, M., et al. (2018). Analysis of 24 genes reveals a monogenic cause in 11.1% of cases with steroid-resistant nephrotic syndrome at a single center. *Pediatr. Nephrol.* 33, 305–314. doi: 10.1007/s00467-017-3801-6

Taylor, J. C., Martin, H. C., Lise, S., Broxholme, J., Cazier, J.-B., Rimmer, A., et al. (2015). Factors influencing success of clinical genome sequencing across a broad spectrum of disorders. *Nat. Genet.* 47, 717–726. doi: 10.1038/ng.3304

van der Ven, A. T., Connaughton, D. M., Ityel, H., Mann, N., Nakayama, M., Chen, J., et al. (2018). Whole-exome sequencing identifies causative mutations in families with congenital anomalies of the kidney and urinary tract. *J. Am. Soc Nephrol.* 29, 2348–2361. doi: 10.1681/ASN.2017121265

Vivante, A., and Hildebrandt, F. (2016). Exploring the genetic basis of early-onset chronic kidney disease. *Nat. Rev. Nephrol.* 12, 133–146. doi: 10.1038/nrneph.2015.205

Warejko, J. K., Tan, W., Daga, A., Schapiro, D., Lawson, J. A., Shril, S., et al. (2018). Whole exome sequencing of patients with steroid-resistant nephrotic syndrome. *Clin. J. Am. Soc Nephrol.* 13, 52–62. doi: 10.2215/CJN.04120417

Webster, A. C., Nagler, E. V., Morton, R. L., and Masson, P. (2017). Chronic Kidney disease. *Lancet* 389, 1238–1252. doi: 10.1016/S0140-6736(16)32064-5

Wu, J., Wu, M., Chen, T., and Jiang, R. (2016). Whole genome sequencing and its applications in medical genetics. *Quant. Biol.* 4, 115–128. doi: 10.1007/s40484-016-0067-0

Xue, Y., Ankala, A., Wilcox, W. R., and Hegde, M. R. (2015). Solving the molecular diagnostic testing conundrum for Mendelian disorders in the era of next-generation sequencing: single-gene, gene panel, or exome/genome sequencing. *Genet. Med.* 17, 444–451. doi: 10.1038/gim.2014.122

Yao, X. D., Chen, X., Huang, G. Y., Yu, Y. T., Xu, S. T., Hu, Y. L., et al. (2012). Challenge in pathologic diagnosis of Alport syndrome: Evidence from correction of previous misdiagnosis. *Orphanet J. Rare Dis.* 7, 100. doi: 10.1186/1750-1172-7-100

Yao, T., Udwan, K., John, R., Rana, A., Haghighi, A., Xu, L., et al. (2019). Integration of genetic testing and pathology for the diagnosis of adults with FSGS. *Clin. J. Am. Soc. Nephrol.* 14, 213–223. doi: 10.2215/CJN.08750718

Zhang, L., Wang, F., Wang, L., Wang, W., Liu, B., Liu, J., et al. (2012). Prevalence of chronic kidney disease in China: a cross-sectional survey. *Lancet* 379, 815–822. doi: 10.1016/S0140-6736(12)60033-6

Zuber, J., Fakhouri, F., Roumenina, L. T., Loirat, C., and Frémeaux-Bacchi, V. (2012). Use of eculizumab for atypical haemolytic uraemic syndrome and C3 glomerulopathies. *Nat. Rev. Nephrol.* 8, 643–657. doi: 10.1038/nrneph.2012.214

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read
for greatest visibility
and readership

FAST PUBLICATION
Around 90 days
from submission
to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative,
and constructive
peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers
acknowledged by name
on published articles

REPRODUCIBILITY OF RESEARCH
Support open data
and methods to enhance
research reproducibility

DIGITAL PUBLISHING
Articles designed
for optimal readership
across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics
track visibility across
digital media

EXTENSIVE PROMOTION
Marketing
and promotion
of impactful research

LOOP RESEARCH NETWORK
Our network
increases your
article's readership