Genetic loci associated with renal function measures and chronic kidney disease in children: the Pediatric Investigation for Genetic Factors Linked with Renal Progression Consortium

Matthias Wuttke[1], Craig S. Wong[2,*], Elke Wühl[3], Daniel Epting[1], Li Luo[4], Anselm Hoppmann[1], Anke Doyon[3], Yong Li[1], CKDGen Consortium, Betül Sözeri[5], Daniela Thurn[6], Martin Helmstädter[1], Tobias B. Huber[1], Tom D. Blydt-Hansen[7], Albrecht Kramer-Zucker[1], Otto Mehls[3], Anette Melk[6], Uwe Querfeld[8], Susan L. Furth[9], Bradley A. Warady[10], Franz Schaefer[3,*] and Anna Köttgen[1,11,*]

[1]Renal Division, Department of Internal Medicine, Medical Center, University of Freiburg, Freiburg im Breisgau, Germany, [2]Division of Pediatric Nephrology, University of New Mexico Children's Hospital, Albuquerque, NM, USA, [3]Division of Pediatric Nephrology, University Medical Center Heidelberg, Heidelberg, Germany, [4]Division of Epidemiology, Department of Internal Medicine, University of New Mexico, Albuquerque, NM, USA, [5]Faculty of Medicine, Ege University, Izmir, Turkey, [6]Department of Pediatric Nephrology, Hannover Medical School, Hannover, Germany, [7]University of Manitoba, Winnipeg, Manitoba, Canada, [8]Charite Universitätsmedizin Berlin, Berlin, Germany, [9]Departments of Pediatrics and Epidemiology, Perelman School of Medicine at the University of Pennsylvania, Childrens Hospital of Philadelphia, Philadelphia, PA, USA, [10]Division of Pediatric Nephrology, Children's Mercy Hospital, Kansas City, MO, USA and [11]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

Correspondence and offprint requests to: Craig S. Wong; E-mail: cwong@salud.unm.edu
*These authors shared joint oversight of the project.

ORIGINAL ARTICLE

ABSTRACT

Background. Chronic kidney disease (CKD) in children is characterized by rapid progression and a high incidence of end-stage renal disease and therefore constitutes an important health problem. While unbiased genetic screens have identified common risk variants influencing renal function and CKD in adults, the presence and identity of such variants in pediatric CKD are unknown.

Methods. The international Pediatric Investigation for Genetic Factors Linked with Renal Progression (PediGFR) Consortium comprises three pediatric CKD cohorts: Chronic Kidney Disease in Children (CKiD), Effect of Strict Blood Pressure Control and ACE Inhibition on the Progression of CRF in Pediatric Patients (ESCAPE) and Cardiovascular Comorbidity in Children with CKD (4C). Clean genotype data from >10 million genotyped or imputed single-nucleotide polymorphisms (SNPs) were available for 1136 patients with measurements of serum creatinine at study enrolment. Genome-wide association studies were conducted to relate the SNPs to creatinine-based estimated glomerular filtration rate (eGFR$_{crea}$) and proteinuria (urinary albumin- or protein-to-creatinine ratio ≥300 and

≥500 mg/g, respectively). In addition, European-ancestry Ped-iGFR patients (cases) were compared with 1347 European-ancestry children without kidney disease (controls) to identify genetic variants associated with the presence of CKD.

Results. SNPs with suggestive association P-values $<1\times10^{-5}$ were identified in 10 regions for eGFR$_{crea}$, four regions for proteinuria and six regions for CKD including some plausible biological candidates. No SNP was associated at genome-wide significance ($P < 5\times10^{-8}$). Investigation of the candidate genes for proteinuria in adults from the general population provided support for a region on chromosome 15 near *RSL24D1/UNC13C/RAB27A*. Conversely, targeted investigation of genes harboring GFR-associated variants in adults from the general population did not reveal significantly associated SNPs in children with CKD.

Conclusions. Our findings suggest that larger collaborative efforts will be needed to draw reliable conclusions about the presence and identity of common variants associated with eGFR, proteinuria and CKD in pediatric populations.

Keywords: pediatric chronic kidney disease, genetic epidemiology, genome-wide association study, glomerular filtration rate, proteinuria

INTRODUCTION

Chronic kidney disease (CKD) is an important global health problem [1]. Although uncommon in children [2–4], CKD in this population is associated with increased risk for poor growth, cardiovascular disease and mortality [5–7]. Despite the lower prevalence of CKD in children when compared with adults, a higher proportion of pediatric CKD patients progress to end-stage renal disease. The average life expectancy of children with renal replacement therapy (RRT) is 63 years for those receiving a transplant and only 38 years for those remaining on dialysis [8].

Thus, gaining insights into pathophysiological mechanisms that underlie the development and progression of pediatric CKD is of great importance. Previous studies in adults support a genetic component to CKD and kidney function that is also observed in individuals from the general population without clear monogenic causes of CKD [9–11]. Genome-wide association studies (GWAS) were used successfully to identify common genetic variants associated with measures of renal function and CKD in population-based studies [12–17] as well as with specific kidney diseases such as IgA nephropathy [18, 19] and membranous nephropathy [20].

Previous GWAS of CKD and renal function measures were conducted among middle-aged or older adults. It is unclear whether common genetic variants also contribute to CKD risk in children. While monogenic causes account for a substantial fraction of pediatric CKD, it is conceivable that common genetic variants, each typically conferring small increases in risk, contribute to the development and/or progression of CKD. To date, no pediatric GWAS of renal function parameters have been conducted because of the lack of sufficiently large study populations. The Pediatric Investigation for Genetic Factors Linked with Renal Progression (PediGFR) Consortium has now assembled several pediatric CKD cohorts. Here we report on GWAS utilizing baseline parameters of renal function, glomerular filtration rate (GFR) and proteinuria and CKD from data of 1136 PediGFR participants.

MATERIALS AND METHODS

The PediGFR Consortium is an international collaborative effort to identify genetic risk factors for kidney disease. The Chronic Kidney Disease in Children (CKiD) prospective cohort study [21] enrolled 540 children aged 1–16 years with an estimated GFR (eGFR) of 30–75 mL/min/1.73 m^2, 440 of which were genotyped. The Effect of Strict Blood Pressure Control and ACE Inhibition on the Progression of CRF in Pediatric Patients (ESCAPE) trial recruited 468 children. Of these, 315 children aged from 3–18 years and with an eGFR of 15–80 mL/min/1.73 m^2 at the baseline visit were genotyped [22]. The Cardiovascular Comorbidity in Children with CKD (4C) prospective study enrolled 705 children aged 6–17 years with an initial eGFR of 10–45 mL/min/1.73 m^2 [23], 691 of whom were genotyped.

Publicly available data from the Study of the Genetic Causes of Complex Pediatric Disorders [24] from the Children's Hospital of Philadelphia (CHOP) were obtained [database of Genotypes and Phenotypes (dbGAP) accession number phs000490.v1.p1] and used as an external control population (interquartile age range 6–17 years). Patients with kidney-related diseases (n = 41) were excluded based on primary diagnosis International Classification of Diseases, Ninth Revision codes (580–594, 598, 403, 404, 250.4 and 753).

GFR was estimated using equations by Schwartz et al. [25]: eGFR$_{crea}$ = 42.3 × (height[m]/serum creatinine[mg/dl])$^{0.78}$) and eGFR$_{cycr}$ = 41.6 × (height[m]/serum creatinine[mg/dl])$^{0.443}$ × (1.8/serum cystatin C [mg/dl])$^{0.479}$). Iohexol GFR was measured for the CKiD cohort [21]. Proteinuria was defined as a urinary albumin-to-creatinine ratio (UACR) ≥300 mg/g or urinary protein-to-creatinine ratio (UPCR) ≥500 mg/g. We chose to analyze proteinuria as a binary phenotype because of its clinical relevance and its high prevalence in CKD populations, and to minimize an effect of medication intake that may result in smaller alterations in the UACR/UPCR on a continuous scale.

For the case–control analyses, all PediGFR patients were treated as CKD cases, while CHOP children without kidney disease were treated as controls.

Using the Illumina Infinium 2.5M-8 microarray, 1450 PediGFR participants were successfully genotyped at ∼2.4 million single-nucleotide polymorphism (SNP) markers. Data cleaning was performed according to standard protocols [26] separately for ESCAPE/4C and for CKiD and resulted in exclusion of n = 84 samples (Supplementary data, Figure S1).

Genomic ancestry was assessed by principal component analysis using Eigensoft SmartPCA v4.2 [27]. The CEU, CHB/JPT and YRI groups (HapMap phase 3, release 2) were used as reference populations to visually derive ancestry-specific clusters (Supplementary data, Figure S2). Participants were assigned to five groups of sufficient size for analysis [CKiD European Ancestry (n = 262), ESCAPE Central European (n = 223), ESCAPE Turkish (n = 62), 4C Central European (n = 270) and 4C Turkish (n = 341)].

Genotyped markers were removed when their call rate was <95%, yielding an average of 1.7 million remaining markers per group. Subsequently, genotype imputation was performed according to standard protocols [28] separately for each group using SHAPEIT v2.r644 [29] and IMPUTE2 v2.3.0 [30]. Phased haplotypes from the 1000 Genomes project [31] (Phase 1, release v3 date 2012-04-19, ALL subset) were used as the reference panel. After imputation, ∼10.5 million SNPs of high imputation quality (information measure ≥0.8) were retained in each group.

A flow chart outlining data preparation for the analyses using the external control population is presented in the Supplementary data (Figure S3). Control samples had been genotyped on the Illumina Human 610-Quad v1 (n = 1480) and the Illumina Human Hap550 v3 arrays (n = 1151). Genotype datasets of only European-ancestry PediGFR cases and CHOP controls were merged. Based on the top 20 calculated principal components, 2005 individuals of similar genetic ancestry were retained for analyses (658 cases, 1347 controls; Supplementary data, Figure S4). Because of the different

genotyping arrays and data sources, stringent data cleaning was performed prior to imputation, removing SNPs with minor allele frequency (MAF) <1%, Hardy–Weinberg equilibrium (HWE) P-value <0.01 (in cases and controls separately) and call rate <95%. Imputation of the combined case–control dataset with the same settings as outlined above resulted in 8.3 million high-quality SNPs (information measure ≥0.8).

Genome-wide association tests were performed using SNPtest v2.5 [32]. Within each of the five PediGFR groups, a linear regression model was used to relate continuous eGFR (ln transformed) to each SNP using an additive model and genotype dosages. The analysis for the binary outcome proteinuria among the PediGFR children was performed similarly using logistic regression. We adjusted for sex, age and any of the first 10 principal components associated with the analyzed phenotype (P < 0.05).

A fixed-effects inverse-variance weighted meta-analysis combined estimates across the five groups using the software GWAMA [33]. Only SNPs with MAF >1% and present in >50 individuals in each group were meta-analyzed. Data were combined over all five groups ($n = 1136$) and separately for individuals of central European ancestry ($n = 744$) and Turkish ancestry ($n = 397$). Genomic control was applied at the individual study level and post meta-analysis [34]. A P-value threshold $<1\times10^{-5}$ was used to indicate suggestive signals and $<5\times10^{-8}$ to indicate genome-wide significance.

For the comparison of PediGFR cases with external controls, power calculations using Quanto 1.2.4 showed that, for a discovery subset of 658 cases and 2 controls per case, there was at least 80% power to detect a CKD odds ratio of 1.87 (1.57) or higher at genome-wide significance at a MAF of 10% (30%), assuming a population prevalence of pediatric CKD of 0.01%. A logistic regression model was used to relate CKD or proteinuria case status to each SNP using an additive model with covariates as described above.

Experimental follow-up was attempted for novel genome-wide significant signals and genes in loci suggestively associated with the clinically important phenotype proteinuria. For genes in the linkage disequilibrium (LD) block of the index SNP, relevant orthologs in *Danio rerio*, including all paralogs, were identified. All animal work was carried out according to relevant national and international guidelines. Zebrafish of the strain AB/TL wild-type were maintained, and the embryos were staged as previously described [35].

Zebrafish orthologs were identified for *clmp*, *gramd1b*, *rab27a* and *rsl24d1*. *SAMD3* had no ortholog; the extended LD block containing *FAM151B* contained too many genes to make experimental follow-up feasible. Amplification was carried out from zebrafish cDNA with specific primers, cloned into TOPO (Invitrogen) and linearized with corresponding restriction enzymes. Whole-mount *in situ* hybridization (WISH) analysis using digoxigenin-labeled probes was performed as described [36] using 4-Nitro blue tetrazolium chloride (blue) (Roche) as substrate.

In order to evaluate the presence of risk variants in genes previously associated with the complex traits CKD, eGFR, creatinine and/or cystatin C levels in adults, a list of 34 genes was assembled using the NHGRI GWAS Catalog [37]

(Supplementary data, Table S1). For each of these genes plus a 5 kb flanking region, the SNP of MAF >5% with the lowest P-value was retrieved from the PediGFR meta-analysis data for the corresponding phenotype. To correct for multiple testing, a Bonferroni correction for the number of independent SNPs per gene (determined using a pruning procedure of the 1000 Genomes Project EUR data; $r^2 <0.2$, window size 50 bp, offset 5 bp) and the number of investigated genes was used.

We further investigated the genes implicated by the PedGFR data in the Chronic Renal Insufficiency Cohort (CRIC) study [38], a study of adult patients with CKD. Data were acquired from dbGAP (accession number phs000524.v1.p1). PediGFR candidate genes were queried in CRIC baseline eGFR association data in the subsample of patients of European ancestry without diabetes ($n = 926$).

In addition, we queried PediGFR proteinuria-associated regions for their association with the UACR in data from up to 54 450 mostly population-based adults studied in the CKDGen Consortium ([14] and personal communication). Significance thresholds were derived based on the number of independent SNPs in the region as described above. Lastly, we investigated the expression of these genes in publicly available data from freshly isolated murine podocytes and other glomerular cells [39]. Differences in gene expression between podocytes and other glomerular cells were corrected for transcriptome-wide multiple testing and additionally for the number of regions investigated ($n = 4$).

RESULTS

Demographic parameters were similar across the three PedGFR cohorts (Table 1). Differences in baseline GFR, serum creatinine and cystatin C as well as the proportion of proteinuric children were consistent with differences in study-specific inclusion criteria.

GWAS of eGFR in PediGFR participants

Association analyses of eGFR$_{crea}$ showed no evidence for systematic inflation of results in the individual groups (1.00 < λ < 1.02) or after meta-analysis (λ = 1.01; Figure 1A). While 10 genomic regions contained at least one SNP associated

Table 1. Characteristics of the PediGFR study cohorts

	CKiD	ESCAPE	4C	All
Subjects (n)	262	285	611	1158
Sex (female) (%)	36	44	35	37
Age (years)	11 ± 4	12 ± 4	12 ± 3	12 ± 4
Height (cm)	136 ± 27	143 ± 23	140 ± 20	140 ± 23
Weight (kg)	37 ± 19	39 ± 18	38 ± 16	38 ± 17
Creatinine (mg/dl)	1.4 ± 0.7	2.1 ± 1.3	2.6 ± 1.4	2.2 ± 1.3
Cystatin C (mg/l)	1.8 ± 0.6	2.5 ± 0.9	2.9 ± 0.9	2.5 ± 0.9
eGFR (creatinine) (ml/min)	46 ± 15	37 ± 14	30 ± 11	36 ± 15
eGFR (cystatin C/creatinine) (ml/min)	46 ± 15	35 ± 12	29 ± 11	34 ± 15
iGFR (iohexol measured) (ml/min)	45 ± 16	–	–	–
Proteinuria/albuminuria[a] (%)	40	55	52	53

[a]UACR >300 mg/g or UPCR >0.5 g/g.

with eGFR$_{crea}$ at P < 10^{-5}, none of these loci reached genome-wide significance. Table 2 shows the SNP with the lowest P-value (index SNP) at each of the loci. Regional association plots of the associated regions are shown in the Supplementary data.

To assess whether the association with creatinine-based eGFR reflected an association with kidney function rather than with creatinine metabolism, we assessed the association of the index SNPs with an alternative estimate of GFR, eGFR$_{cycr}$. Both effect sizes and directions of association were consistent (Supplementary data, Table S2). Second, measured iohexol GFR was available in a subset of CKiD study participants. Again, the effects of the index SNPs were consistent with those found in the main analysis (Figure 2), indicating

the use of eGFR$_{crea}$ as the primary phenotype was appropriate. Allele frequencies of the index SNPs were similar across the five groups.

In the secondary meta-analyses of the three groups of central European ancestry and the two groups of Turkish ancestry separately, there were also no genome-wide significant findings for eGFR$_{crea}$ (Supplementary data, Table S3).

GWAS of proteinuria in PediGFR

In association analyses comparing PediGFR children with proteinuria with PediGFR children without proteinuria, we observed no evidence for systematic inflation of study-specific or combined results (Figure 1B). Four genomic regions contained index SNPs of MAF ≥5% and P < 10^{-5}, mapping in or near

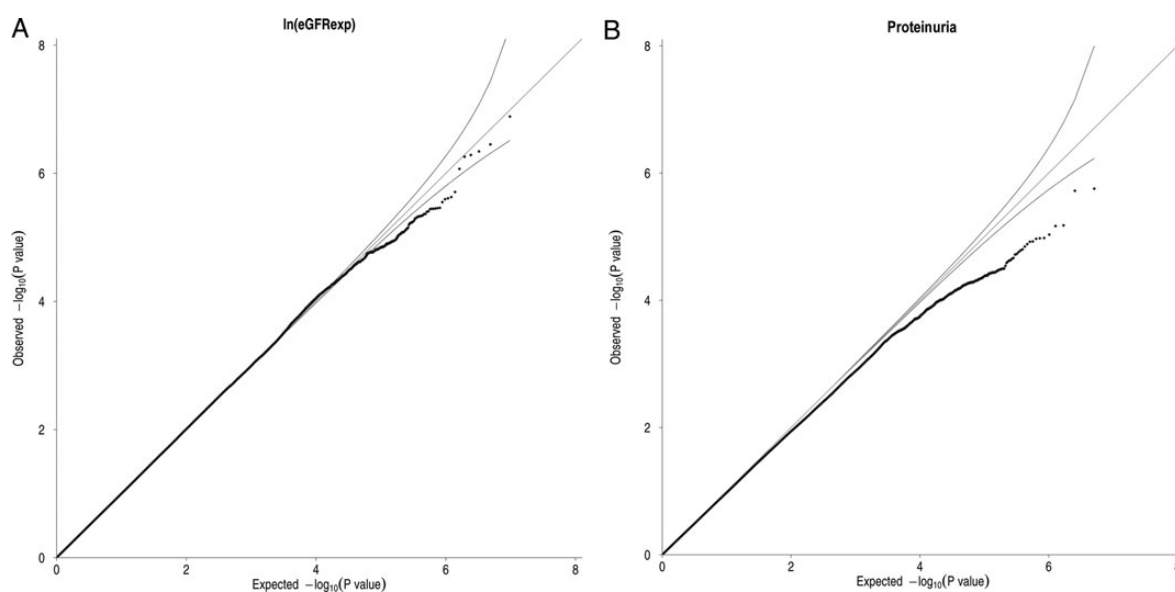**FIGURE 1:** (**A**) Quantile–quantile (QQ) plots for the PediGFR eGFR$_{crea}$ quantitative GWAS and (**B**) the PediGFR proteinuria GWAS. X-axis, expected $-\log_{10}$ P-value; Y-axis, observed $-\log_{10}$ P-value. Thin lines above and below the identity line indicate the 95% confidence interval.

Table 2. SNPs associated with eGFR or proteinuria with P < 10^{-5} in the PediGFR Consortium

SNP identifier	Chr.	Pos.	Coded allele	Coded allele frequency	Effect/OR	SE/95% CI	P-value	I^2	Gene symbols
eGFR (creatinine) (*n* = 1136)					Effect	SE			
rs7541937	1	35341982	G	0.47	−0.07	0.01	3.5E-06	0.36	*DLGAP3*
rs71413769	2	139833140	G	0.10	0.11	0.02	4.7E-06	0.01	NXPH2
rs10928150	2	143254723	A	0.46	−0.07	0.01	7.3E-06	0.00	LRP1B, KYNU
rs138904061	3	105174202	C	0.16	−0.09	0.02	6.1E-06	0.10	*ALCAM*
rs10054413	5	88605629	G	0.14	−0.10	0.02	7.0E-06	0.23	MEF2C
rs62441014	7	5533319	C	0.41	0.07	0.02	3.9E-06	0.24	*FBXL18*
rs11066679	12	114123402	C	0.23	−0.08	0.02	2.8E-06	0.00	RBM19, LHX5
rs12898179	14	67964171	C	0.79	−0.08	0.02	8.8E-06	0.00	*TMEM229B*
rs344364	16	1973516	A	0.22	0.08	0.02	5.0E-06	0.00	HS3ST6, MSRB1
rs6036759	20	2491209	A	0.41	0.07	0.01	2.4E-06	0.00	ZNF343, TMC2
Proteinuria (*n* = 1065)					OR	95% CI			
rs145768835	5	79834287	C	0.18	0.55	0.42–0.71	6.7E-06	0.00	*FAM151B*
rs61201007	6	130515494	T	0.37	1.54	1.27–1.87	9.3E-06	0.02	*SAMD3*
rs55847827	11	123189003	G	0.18	1.80	1.39–2.32	6.8E-06	0.00	*MIR4493, CLMP*
rs76158983	15	55309006	A	0.05	0.35	0.22–0.55	5.2E-06	0.00	RSL24D1, UNC13C

Variants are filtered for MAF >5% and I^2 <50% and must be present in all five analysis samples. Underlined gene names indicated intronic localization of the SNPs.

Chr., chromosome; Pos., position; OR, odds ratio; SE, standard error; 95% CI, 95% confidence interval. Positions refer to build 37.

FAM151B, *SAMD3*, *MIR4493/CLMP* and *RSL24D1/UNC13C/RAB27A* (Table 2), but none of these associations reached genome-wide significance.

GWAS of CKD and proteinuria comparing PediGFR cases with external pediatric controls

In order to test for the presence of common genetic variants associated with the presence of CKD or proteinuria, we compared PediGFR cases with external pediatric controls without kidney disease. There was no systematic inflation of results comparing PediGFR participants (CKD cases) versus CHOP controls ($\lambda = 1.03$) and for the comparison of PediGFR patients with proteinuria versus CHOP controls ($\lambda = 1.00$). Table 3 shows that while there were no genome-wide significant
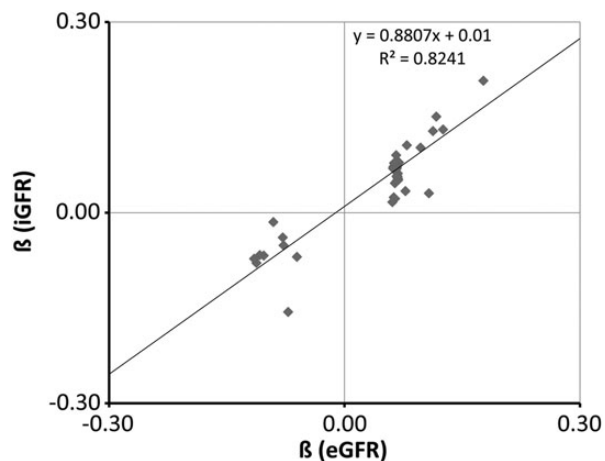
FIGURE 2: Scatter plot comparing the effects of the index SNPs associated with eGFR$_{crea}$ with their effects on iohexol-measured GFR (iGFR) in a subsample of CKiD patients with measured GFR. Black: regression line.

associations for both analyses, six regions contained index SNPs associated with CKD at $P < 10^{-5}$ and 10 regions contained such SNPs for proteinuria.

Experimental evaluation of results

In lieu of suitable and sufficiently powered pediatric replication cohorts, we sought experimental evidence and examined kidney-specific expression during zebrafish development. Genes in regions suggestively associated with proteinuria were selected because of the clinical relevance of proteinuria in CKD progression. None of the selected genes showed specific glomerular expression during zebrafish development (Supplementary data, Figure S5), although general expression was observed for genes *clmp*, *gramd1b*, *rab27a* and *rsl24dl*. Among other organ systems, *rab27a* showed expression in the pronephric tubule.

Candidate gene evaluation

Finally, we used evidence from studies of CKD or complex renal traits in adults to examine overlap with the loci implicated here. None of the 34 known risk genes for CKD or complex renal traits in adults contained SNPs significantly associated with the corresponding traits in the PediGFR or the PediGFR-CHOP datasets after correcting for multiple testing. Some nominally significant associations, such as rs11097415 in the *SHROOM3* region and CKD, were observed (P = 7.3×10^{-4}). In addition, we queried the 10 genes associated with eGFR in PediGFR (Table 2) in eGFR association results of the CRIC baseline data, a study of adult patients with CKD. Again, the search did not yield any significantly associated SNPs.

One of the four candidate regions for proteinuria, *RSL24D1/UNC13C/RAB27A*, contained SNPs significantly associated with UACR in the CKDGen dataset. The index SNP in the CKDGen Consortium was rs1528472 (P = 5×10^{-7}). This variant falls into

Table 3. SNPs associated with CKD with P < 10^{-5} (European-ancestry PediGFR patients as cases and CHOP children without kidney disease as controls)

SNP identifier	Chr.	Pos.	Coded allele	Coded allele frequency	OR	95% CI	P-value	Gene symbols
CKD (*n* = 2005, *n* = 658 cases), $\lambda = 1.03$								
rs13072153	3	123178132	G	0.38	0.70	0.60–0.81	2.6E-06	*ADCY5, PTPLB*
chr5:54032778:I	5	54032778	CT	0.53	0.70	0.60–0.81	2.0E-06	*(SNX18, ESM1)*
rs7129190	11	113818752	A	0.53	0.71	0.62–0.83	4.9E-06	*HTR3B, HTR3A*
rs9535020	13	48923808	G	0.13	1.69	1.35–2.10	4.2E-06	*RB1*
rs6497441	16	9829848	T	0.29	1.44	1.23–1.69	5.2E-06	*GRIN2A*
rs8117834	20	8090640	T	0.14	1.68	1.34–2.10	5.6E-06	*PLCB1, TMX4*
Proteinuria (*n* = 1641, *n* = 294 cases), $\lambda = 1.00$								
rs792064	2	5518678	T	0.42	1.64	1.33–2.01	2.0E-06	*SOX11*
rs11720887	3	20940843	A	0.13	1.87	1.44–2.43	4.0E-06	*(VENTXP7)*
rs11011653	10	20163993	G	0.34	0.59	0.47–0.74	1.4E-06	*PLXDC2*
rs4963124	11	762791	C	0.71	1.71	1.34–2.18	8.6E-06	*TALDO1*
chr11:87335171:D	11	87335171	G	0.33	0.60	0.48–0.76	9.0E-06	*TMEM135, RAB38*
rs78057470	15	84149984	C	0.32	1.61	1.31–1.98	6.6E-06	*SH3GL3*
rs35219335	17	10077836	T	0.14	0.49	0.35–0.67	4.1E-06	*GAS7*
rs9895232	17	4519139	C	0.35	1.62	1.32–1.99	2.7E-06	*SMTNL2, ALOX15*
rs12601558	17	80179147	C	0.22	0.56	0.43–0.73	5.5E-06	*CCDC57, SLC16A3*
rs56052703	19	13432940	C	0.34	0.60	0.48–0.75	3.4E-06	*CACNA1A*

Variants are filtered to remove SNPs with HWE P-value <0.01 and MAF <10%. Underlined gene names indicate intronic localization of the SNP. OR refers to the effect of each additional copy of the coded allele.

Chr., chromosome; Pos., position; OR, odds ratio; 95% CI, 95% confidence interval.

the same LD block as the PediGFR index variant rs76158983, which was not present in the HapMap-imputation-based CKDGen data that also did not contain a good proxy ($r^2 > 0.8$) for rs76158983.

One of the genes, *FAM151B*, showed significantly higher expression in podocytes compared with other glomerular cells in the murine glomerular expression dataset (1.4-fold change, corrected P = 3×10^{-4}).

Supplementary data Table S5 summarizes results attempting to prioritize genes in the implicated regions through several bioinformatics resources that leverage information on gene expression and functional variant annotation as well as a literature search.

DISCUSSION

This study is the first GWAS meta-analysis based on data from 1136 children with CKD recruited in large international collaborations. Our analyses revealed suggestive associations of several common genetic variants with eGFR and proteinuria as well as with CKD status at study baseline. However, none of the associations reached genome-wide significance.

Multiple GWASs among adults from the general population identified common risk variants associated with measures of kidney function and CKD [12–17, 40]. The variants identified in these studies confer modest effects, as evidenced by the large sample sizes required for their identification. For example, GWASs among ~20 000 adults identified SNPs in four eGFR-associated regions [12], whereas almost 70 000 individuals were required to identify 25 such regions [13].

Consequently, genetic studies in 1136 children would only be able to identify common risk variants if they conferred larger risks than those observed in population-based adults. It is conceivable that a given susceptibility variant may exert a larger effect in children because additional factors, e.g. the presence of micro- and macro-vascular diseases and of lifestyle-related factors such as smoking and obesity, can be assumed to have a smaller impact. In support, a previous study of growth and height reported a larger effect of the same common variant in children compared with adults [41]. Our results do not support the presence of common variants of large effect associated with GFR or proteinuria in this sample of children, since our study had >85% power to detect, for a variant with MAF of 20%, a change in average GFR ≥16% or an odds ratio for CKD ≥1.7 per risk allele at genome-wide significance.

There are at least three potential explanations for the negative findings at genome-wide significance: first, common genetic susceptibility variants may be present in children but only confer small changes in risk, similar to adults. This is supported by published GWASs of other phenotypes in pediatric populations that reported mostly small effect sizes of significant and replicated results [24, 42, 43]. The discovery of such variants in children with CKD would require a larger study sample than PediGFR, and additional samples for replication. This represents a considerable challenge in light of the low incidence of pediatric CKD and the fact that PediGFR already represents the currently largest international effort collecting patients across continents for more than a decade.

A second explanation could be that the genetic architecture of CKD differs between children and adults, and that common genetic variants of moderate effect do not play an equally important role in children as they do in adults. The substantial fraction of children with CKD due to a single-gene mutation [44–46] may limit the role of common genetic variants that modify rather than determine disease. As GWASs, by design, do not allow for the detection of single-gene mutations, and because for many children in PediGFR, monogenic causes of CKD were not comprehensively evaluated, this explanation represents a hypothesis that can be addressed in the future once whole genome sequences are available.

Third, the general phenotypes studied in PediGFR arise as a result of different underlying diseases. It is plausible that this etiologic heterogeneity complicates the detection of cause-specific common variants of larger effect. The combined study of renal function parameters and CKD in all individuals was performed to increase statistical power for the discovery of variants that impact pathophysiological mechanisms irrespective of disease etiology. Evidence from studies in adults with CKD suggests that common variants of large effects indeed exist for specific renal diseases, for example, membranous nephropathy [20] or *APOL1*-related kidney disease [47]. Again, given the low incidence of pediatric CKD, subgroup analyses will be extremely challenging.

Since there were no regions associated at genome-wide significance and suggestive regions could not be replicated because of the lack of adequate studies, we refrain from detailed discussions of the genes mapping into suggestive regions. One locus that received some support through additional evidence is the one at *RSL24D1/UNC13C/RAB27A* on chromosome 15 that showed an association with proteinuria in CKD. This region was significantly associated with a closely related phenotype, UACR, in population-based data from the CKDGen Consortium.

The strengths of our study are the large international collection of patients and the careful and comprehensive phenotypic characterization, including the availability of several renal function markers as well as measured GFR in a subset of the participants. Data cleaning and analysis were carried out using state-of-the-art methods, which together with the public posting of the genotype data facilitates reproducibility and allows for future pooling with data from additional cohorts. The consortium assembled data from understudied ethnicities such as the Turkish population, where the Chronic Renal Disease in Turkey (CREDIT) study estimated that the prevalence of pediatric CKD stages 3–5 was 2600 per million age-related population as opposed to 68.9 in Western Europe [48].

A number of limitations deserve acknowledgement. The study did not have sufficient statistical power to detect genetic risk variants that confer only moderate changes in risk. As such, it is not possible to draw definite conclusions about the presence of such variants, the genetic architecture of CKD in children when compared with adults or their effect on specific CKD etiologies. Another important limitation is the lack of suitable external replication cohorts. Finally, because of the study inclusion criteria, the full spectrum of eGFR could not be examined. As the studies were designed with the objective to study CKD progression, GWAS of kidney function decline and incident

end-stage renal disease should be performed once sufficient follow-up data have been collected.

In this first GWAS meta-analysis of pediatric CKD, we did not identify common genetic susceptibility variants significantly associated with GFR, proteinuria or pediatric CKD status, although several suggestive candidates emerged. Even larger efforts will be needed to draw reliable conclusions about the presence of such variants in pediatric populations.

SUPPLEMENTARY DATA

Supplementary data are available online at http://ndt.oxford-journals.org.

CONFLICT OF INTEREST STATEMENT

The results presented in this paper have not been published previously in whole or part, except in abstract format. The authors declare no conflicts of interest.

(See related article by Böger and Hayward. The PediGFR consortium: studying the genetics of complex kidney traits in children. *Nephrol Dial Transplant* 2016; 31: 173–175.)

REFERENCES

1. Eckardt K-U, Coresh J, Devuyst O et al. Evolving importance of kidney disease: from subspecialty to global health burden. Lancet 2013; 382: 158–169
2. Peco-Antic A, Bogdanovic R, Paripovic D et al. Epidemiology of chronic kidney disease in children in Serbia. Nephrol Dial Transplant 2012; 27: 1978–1984
3. Ardissino G, Dacco V, Testa S et al. Epidemiology of chronic renal failure in children: data from the ItalKid project. Pediatrics 2003; 111: e382–e387
4. Harambat J, van Stralen KJ, Kim JJ et al. Epidemiology of chronic kidney disease in children. Pediatr Nephrol 2012; 27: 363–373
5. Al-Uzri A, Matheson M, Gipson DS et al. The impact of short stature on health-related quality of life in children with chronic kidney disease. J Pediatr 2013; 163: 736–741 e1
6. Abraham AG, Schwartz GJ, Furth S et al. Longitudinal formulas to estimate GFR in children with CKD. Clin J Am Soc Nephrol 2009; 4: 1724–1730
7. Mitsnefes MM. Cardiovascular disease in children with chronic kidney disease. J Am Soc Nephrol 2012; 23: 578–585
8. Warady BA, Chadha V. Chronic kidney disease in children: the global perspective. Pediatr Nephrol 2007; 22: 1999–2009
9. Lei HH, Perneger TV, Klag MJ et al. Familial aggregation of renal disease in a population-based case-control study. J Am Soc Nephrol 1998; 9: 1270–1276
10. Satko SG, Freedman BI. The familial clustering of renal disease and related phenotypes. Med Clin North Am 2005; 89: 447–456
11. Fox CS, Yang Q, Cupples LA et al. Genomewide linkage analysis to serum creatinine, GFR, and creatinine clearance in a community-based population: the Framingham Heart Study. J Am Soc Nephrol 2004; 15: 2457–2461
12. Kottgen A, Glazer NL, Dehghan A et al. Multiple loci associated with indices of renal function and chronic kidney disease. Nat Genet 2009; 41: 712–717
13. Kottgen A, Pattaro C, Böger CA et al. New loci associated with kidney function and chronic kidney disease. Nat Genet 2010; 42: 376–384
14. Böger CA, Chen MH, Tin A et al. CUBN is a gene locus for albuminuria. J Am Soc Nephrol 2011; 22: 555–570
15. Pattaro C, Kottgen A, Teumer A et al. Genome-wide association and functional follow-up reveals new loci for kidney function. PLoS Genet 2012; 8: e1002584
16. Chambers JC, Zhang W, Lord GM et al. Genetic loci influencing kidney function and chronic kidney disease. Nat Genet 2010; 42: 373–375
17. Okada Y, Sim X, Go MJ et al. Meta-analysis identifies multiple loci associated with kidney function-related traits in East Asian populations. Nat Genet 2012; 44: 904–909
18. Kiryluk K, Li Y, Scolari F et al. Discovery of new risk loci for IgA nephropathy implicates genes involved in immunity against intestinal pathogens. Nat Genet 2014; 46: 1187–1196
19. Gharavi AG, Kiryluk K, Choi M et al. Genome-wide association study identifies susceptibility loci for IgA nephropathy. Nat Genet 2011; 43: 321–327
20. Stanescu HC, Arcos-Burgos M, Medlar A et al. Risk HLA-DQA1 and PLA (2)R1 alleles in idiopathic membranous nephropathy. N Engl J Med 2011; 364: 616–626
21. Furth SL, Cole SR, Moxey-Mims M et al. Design and methods of the Chronic Kidney Disease in Children (CKiD) prospective cohort study. Clin J Am Soc Nephrol 2006; 1: 1006–1015
22. Wuhl E, Trivelli A, Picca S et al. Strict blood-pressure control and progression of renal failure in children. N Engl J Med 2009; 361: 1639–1650
23. Querfeld U, Anarat A, Bayazit AK et al. The Cardiovascular Comorbidity in Children with Chronic Kidney Disease (4C) study: objectives, design, and methodology. Clin J Am Soc Nephrol 2010; 5: 1642–1648
24. Sleiman PM, Flory J, Imielinski M et al. Variants of DENND1B associated with asthma in children. N Engl J Med 2010; 362: 36–44
25. Schwartz GJ, Schneider MF, Maier PS et al. Improved equations estimating GFR in children with chronic kidney disease using an immunonephelometric determination of cystatin C. Kidney Int 2012; 82: 445–453
26. Anderson CA, Pettersson FH, Clarke GM et al. Data quality control in genetic case-control association studies. Nat Protoc 2010; 5: 1564–1573
27. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. PLoS Genet 2006; 2: e190
28. Marchini J, Howie B. Genotype imputation for genome-wide association studies. Nat Rev Genet 2010; 11: 499–511
29. Delaneau O, Zagury JF, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. Nat Methods 2013; 10: 5–6
30. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet 2009; 5: e1000529

31. Abecasis GR, Auton A, Brooks LD et al. An integrated map of genetic variation from 1,092 human genomes. Nature 2012; 491: 56–65
32. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007; 447: 661–678
33. Magi R, Morris AP. GWAMA: software for genome-wide association meta-analysis. BMC Bioinformatics. 2010; 11: 288
34. Devlin B, Roeder K, Wasserman L. Genomic control, a new approach to genetic-based association studies. Theor Popul Biol 2001; 60: 155–166
35. Kimmel CB, Ballard WW, Kimmel SR et al. Stages of embryonic development of the zebrafish. Dev Dyn 1995; 203: 253–310
36. Epting D, Slanchev K, Boehlke C et al. The Rac1 regulator ELMO controls basal body migration and docking in multiciliated cells through interaction with Ezrin. Development 2015; 142: 174–184
37. Welter D, MacArthur J, Morales J et al. The NHGRI GWAS catalog, a curated resource of SNP-trait associations. Nucleic Acids Res 2014; 42: D1001–D1006
38. Feldman HI, Appel LJ, Chertow GM et al. The Chronic Renal Insufficiency Cohort (CRIC) study: design and methods. J Am Soc Nephrol 2003; 14: S148–S153
39. Boerries M, Grahammer F, Eiselein S et al. Molecular fingerprinting of the podocyte reveals novel gene and protein regulatory networks. Kidney Int 2013; 83: 1052–1064
40. Böger CA, Heid IM. Chronic kidney disease: novel insights from genome-wide association studies. Kidney Blood Press Res 2011; 34: 225–234
41. van der Valk RJ, Kreiner-Moller E, Kooijman MN et al. A novel common variant in DCST2 is associated with length in early life and height in adulthood. Hum Mol Genet 2015; 24: 1155–1168
42. Bonnelykke K, Sleiman P, Nielsen K et al. A genome-wide association study identifies CDHR3 as a susceptibility locus for early childhood asthma with severe exacerbations. Nat Genet 2014; 46: 51–55
43. Rothenberg ME, Spergel JM, Sherrill JD et al. Common variants at 5q22 associate with pediatric eosinophilic esophagitis. Nat Genet 2010; 42: 289–291
44. Sadowski CE, Lovric S, Ashraf S et al. A single-gene cause in 29.5% of cases of steroid-resistant nephrotic syndrome. J Am Soc Nephrol 2014; 1: 1–4
45. Hwang DY, Dworschak GC, Kohl S et al. Mutations in 12 known dominant disease-causing genes clarify many congenital anomalies of the kidney and urinary tract. Kidney Int 2014; 85: 1429–1433
46. Devuyst O, Knoers NV, Remuzzi G et al. Rare inherited kidney diseases: challenges, opportunities, and perspectives. Lancet 2014; 383: 1844–1859
47. Genovese G, Friedman DJ, Ross MD et al. Association of trypanolytic ApoL1 variants with kidney disease in African Americans. Science 2010; 329: 841–845
48. Soylemezoglu O, Duzova A, Yalcinkaya F et al. Chronic renal disease in children aged 5–18 years: a population-based survey in Turkey, the CREDIT-C study. Nephrol Dial Transplant 2012; 27(Suppl 3): iii146–iii151

ORIGINAL ARTICLE

Seasonal influenza vaccination is associated with reduced morbidity and mortality in peritoneal dialysis patients

I-Kuan Wang[1,2,3], Cheng-Li Lin[4,5], Po-Chang Lin[6], Shih-Ni Chang[7], Che-Yi Chou[3], Tzung-Hai Yen[8,9], Chiz-Tzung Chang[3], Chiu-Ching Huang[3] and Fung-Chang Sung[1,4,10]

[1]Graduate Institute of Clinical Medical Science, China Medical University, Taichung, Taiwan, [2]Department of Internal Medicine, College of Medicine, China Medical University, Taichung, Taiwan, [3]Division of Kidney Disease, China Medical University Hospital, Taichung, Taiwan, [4]Management Office for Health Data, China Medical University Hospital, Taichung, Taiwan, [5]College of Medicine, China Medical University, Taichung, Taiwan, [6]Division of Infection, China Medical University Hospital, Taichung, Taiwan, [7]Department of Medical Research, Taichung Veterans General Hospital, Taichung, Taiwan, [8]Division of Nephrology, Chang Gung Memorial Hospital, Taipei, Taiwan, [9]Chang Gung University College of Medicine, Taoyuan, Taiwan and [10]Department of Health Services Administration, China Medical University, Taichung, Taiwan

Correspondence and offprint requests to: Fung-Chang Sung; E-mail: fcsung1008@yahoo.com

ABSTRACT

Background. Studies on the effectiveness of seasonal influenza vaccination in peritoneal dialysis (PD) patients are limited. The aim of the present study is to evaluate the effectiveness of seasonal influenza vaccination in reducing morbidity and mortality in incident end-stage renal disease patients on PD.

Methods. From Taiwan's National Health Insurance Research Database, we identified 2089 incident PD patients with seasonal influenza vaccination and 2089 propensity score matched incident PD patients without the vaccination during