JCI The Journal of Clinical Investigation

Genomic imbalances in pediatric patients with chronic kidney disease

Miguel Verbitsky, … , Craig S. Wong, Ali G. Gharavi

J Clin Invest. 2015;125(5):2171-2178. https://doi.org/10.1172/JCI80877.

Clinical Medicine | Genetics | Nephrology

BACKGROUND. There is frequent uncertainty in the identification of specific etiologies of chronic kidney disease (CKD) in children. Recent studies indicate that chromosomal microarrays can identify rare genomic imbalances that can clarify the etiology of neurodevelopmental and cardiac disorders in children; however, the contribution of unsuspected genomic imbalance to the incidence of pediatric CKD is unknown.

METHODS. We performed chromosomal microarrays to detect genomic imbalances in children enrolled in the Chronic Kidney Disease in Children (CKiD) prospective cohort study, a longitudinal prospective multiethnic observational study of North American children with mild to moderate CKD. Patients with clinically detectable syndromic disease were excluded from evaluation. We compared 419 unrelated children enrolled in CKiD to multiethnic cohorts of 21,575 children and adults that had undergone microarray genotyping for studies unrelated to CKD.

RESULTS. We identified diagnostic copy number disorders in 31 children with CKD (7.4% of the cohort). We detected 10 known pathogenic genomic disorders, including the 17q12 deletion *HNF1 homeobox B* (*HNF1B*) and triple X syndromes in 19 of 419 unrelated CKiD cases as compared with 98 of 21,575 control individuals (OR 10.8, $P = 6.1 \times 10^{-20}$). In an additional […]

Genomic imbalances in pediatric patients with chronic kidney disease

Miguel Verbitsky,[1] Simone Sanna-Cherchi,[1] David A. Fasel,[1] Brynn Levy,[2] Krzysztof Kiryluk,[1] Matthias Wuttke,[3] Alison G. Abraham,[4] Frederick Kaskel,[5] Anna Köttgen,[3,4] Bradley A. Warady,[6] Susan L. Furth,[7] Craig S. Wong,[8] and Ali G. Gharavi[1]

[1]Department of Medicine, Division of Nephrology, and [2]Department of Pathology and Cell Biology, Columbia University, College of Physicians and Surgeons, New York, New York, USA. [3]Department of Nephrology, Medical Center – University of Freiburg, Freiburg, Germany. [4]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA. [5]Division of Pediatric Nephrology, Albert Einstein College of Medicine, New York, New York, USA. [6]Division of Pediatric Nephrology, Children's Mercy Hospital, Kansas City, Missouri, USA. [7]Departments of Pediatrics and Epidemiology, Perelmen School of Medicine at the University of Pennsylvania, Division of Nephrology, Children's Hospital of Philadelphia (CHOP), Philadelphia, Pennsylvania, USA. [8]Division of Pediatric Nephrology, University of New Mexico Children's Hospital, Albuquerque, New Mexico, USA.

BACKGROUND. There is frequent uncertainty in the identification of specific etiologies of chronic kidney disease (CKD) in children. Recent studies indicate that chromosomal microarrays can identify rare genomic imbalances that can clarify the etiology of neurodevelopmental and cardiac disorders in children; however, the contribution of unsuspected genomic imbalance to the incidence of pediatric CKD is unknown.

METHODS. We performed chromosomal microarrays to detect genomic imbalances in children enrolled in the Chronic Kidney Disease in Children (CKiD) prospective cohort study, a longitudinal prospective multiethnic observational study of North American children with mild to moderate CKD. Patients with clinically detectable syndromic disease were excluded from evaluation. We compared 419 unrelated children enrolled in CKiD to multiethnic cohorts of 21,575 children and adults that had undergone microarray genotyping for studies unrelated to CKD.

RESULTS. We identified diagnostic copy number disorders in 31 children with CKD (7.4% of the cohort). We detected 10 known pathogenic genomic disorders, including the 17q12 deletion *HNF1 homeobox B* (*HNF1B*) and triple X syndromes in 19 of 419 unrelated CKiD cases as compared with 98 of 21,575 control individuals (OR 10.8, $P = 6.1 \times 10^{-20}$). In an additional 12 CKiD cases, we identified 12 likely pathogenic genomic imbalances that would be considered reportable in a clinical setting. These genomic imbalances were evenly distributed among patients diagnosed with congenital and noncongenital forms of CKD. In the vast majority of these cases, the genomic lesion was unsuspected based on the clinical assessment and either reclassified the disease or provided information that might have triggered additional clinical care, such as evaluation for metabolic or neuropsychiatric disease.

CONCLUSION. A substantial proportion of children with CKD have an unsuspected genomic imbalance, suggesting genomic disorders as a risk factor for common forms of pediatric nephropathy. Detection of pathogenic imbalances has practical implications for personalized diagnosis and health monitoring in this population.

TRIAL REGISTRATION. ClinicalTrials.gov NCT00327860.

FUNDING. This work was supported by the NIH, the National Institutes of Diabetes and Digestive and Kidney Diseases (NIDDK), the National Institute of Child Health and Human Development, and the National Heart, Lung, and Blood Institute.

Introduction

Chronic kidney disease (CKD) is a major public health problem that affects up to 13% of the United States population (1–3). The impact of CKD in children is particularly profound, with increased morbidity from anemia, hypertension, and cardiovascular complications as well as neurodevelopmental and behavioral deficits. Children with end-stage renal disease have a 1-year mortality rate of 35 per 1,000 patient years (1), but it is not known what proportion of mortality and comorbidities are sequelae of progressive renal dysfunction or attributable to additional independent factors (4, 5). As a complicating factor, the etiology of CKD in children is sometimes not known or is classified nonspecifically within likely heterogeneous pathologic descriptions such as focal segmental glomerulosclerosis. A precise diagnosis of the underlying etiologies may facilitate screening, prevention, or management of comorbid conditions and complications.

Copy number variants (CNV) constitute a substantial source of genetic variation in humans (6). Chromosomal microarray analysis is the preferred modality for the diagnosis of CNV disorders and has

▶ Related Commentary: p. 1799

Conflict of interest: The authors have declared that no conflict of interest exists.
Submitted: January 7, 2015; **Accepted:** March 12, 2015.
Reference information: *J Clin Invest*. 2015;125(5):2171–2178. doi:10.1172/JCI80877.

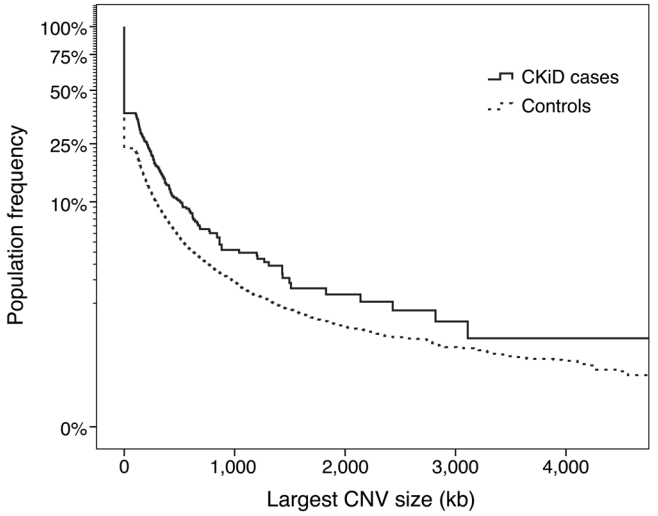

Figure 1. CNV burden based on the largest CNV per genome. There is an excess burden of rare, large, gene-disrupting CNVs in CKiD cases as compared with controls (log-rank test, $P = 9.9 \times 10^{-8}$). Only rare (frequency ≤ 0.1% of controls) and exon-intersecting CNVs are analyzed. The *y* axis (exponential scale) represents the percentage of individuals, with their largest CNV above each size threshold (in kb) on the *x* axis.

better resolution compared with traditional karyotyping for detection of genomic imbalances (7, 8). Using microarray analysis, many studies have implicated CNV disorders in different congenital and neurodevelopmental defects (7–14). These studies have shown that the same genomic imbalances can have pleiotropic effects on neurologic, cardiac, and skeletal development (9–14), suggesting that they affect common developmental pathways. Further supporting these data, we recently showed that nearly 10.5% of children and young adults with kidney malformations harbor pathogenic genomic imbalances that were unsuspected by clinical exam (15) and overlapped significantly with CNV disorders implicated in neurodevelopmental disorders (9, 11, 12). However, the contribution of these genomic imbalances to broader etiologies of CKD has not been systematically studied. Here, we performed an analysis of high-density microarrays in children enrolled in the Chronic Kidney Disease in Children (CKiD) prospective cohort study (2) to determine the prevalence of pathogenic genomic imbalances in different categories of pediatric CKD.

Results

Prevalence of genomic disorders among controls. Because genomic disorders have a very low frequency in the population, we assembled a large data set of 21,575 pediatric and adult controls from publicly available genetic studies unrelated to kidney disease to accurately estimate the frequencies of rare CNVs. As in previous studies of pediatric and neurodevelopmental disorders (9, 15, 16), we deliberately included adults as controls because many CNVs have an age-related penetrance and adults who were screened and consented for genetic studies are less likely to carry undetected genomic imbalances predisposing to pediatric CKD. Examination of the 21,575 controls revealed large, rare, genic, autosomal CNVs in 5,056 individuals (23.4%) in this population. We next annotated these CNVs for 131 known genomic disorders listed in Supplemental Table 1 (supplemental material available online with this article; doi:10.1172/JCI80877DS1). We detected 28 distinct syndromes in 98 individuals (0.45% of the controls, Supplemental Figure 1). No control individual carried more than one known genomic disorder. The overall population frequency of genomic disorders was consistent with previous reports (9, 15–17), and there were no differences in prevalence among different ethnicities. Similar to prior studies, the most frequent abnormalities were 17p12 (*PMP22*) deletion/duplication (12 and 8 carriers, respectively), 1q21.1 deletion/reciprocal duplication (susceptibility locus for thrombocytopenia–absent radius syndrome, 7 and 8 carriers, respectively), and 1q21.1 distal recurrent microdeletion/microduplication (4 and 5 carriers, respectively). These data establish normative data for the prevalence of CNVs, indicating that genomic disorders have a very low frequency in the control populations.

Prevalence of rare genomic imbalances in children with CKD. Compared with the control population, we observed an excess burden of large, rare gene-disrupting CNVs among the 419 unrelated CKID cases (Figure 1). Altogether, 158 (37.7%) of the CKiD children had at least one large, rare gene-disrupting autosomal CNV compared with 5,056 (23.4%) of controls (OR = 2.0, $P = 2.2 \times 10^{-10}$). These data suggest that potentially up to 14.3% of the pediatric CKD cases (60 individuals) might be attributable to a CNV of 100 kb or larger. We next performed detailed annotation of CNVs as delineated below.

Known genomic disorders. In our annotation for 131 known genomic disorders, we identified diagnostic CNVs in 19 of 419 unrelated CKiD cases (4.53%) as compared with 98 of 21,575 (0.45%) control individuals, (OR 10.78; $P = 6.08 \times 10^{-20}$). Consis-

Table 1. Prevalence of known and likely pathogenic genomic imbalances in CKiD cases as compared with controls

	n	Individuals with a known genomic disorder (%)	Odds ratio (*P* value)	Individuals with a likely pathogenic CNV (%)
All CKD	419	19 (4.5%)	10.8 (6.08×10^{-20})	12 (2.9%)
CAKUT	221	10 (4.5%)	11.0 (2.96×10^{-12})	7 (3.2%)
CAKUT due to RHD	67	7 (10.5%)	30.1 (4.91×10^{-16})	1 (1.5%)
Non-CAKUT	198	9 (4.6%)	10.3 (8.74×10^{-11})	5 (2.5%)
Controls	21,575	98 (0.5%)	1	

Odds ratios with associated *P* values are derived from logistic regression analysis of known pathogenic imbalances in CKiD cases in comparison with controls (reference group); the analysis is adjusted for race/ethnicity and sex. CAKUT, congenital abnormality of the kidney and urinary tract. CAKUT includes ureteropelvic junction obstruction, reflux nephropathy, and RHD. RHD cases are also shown separately. Other CKiD cases were included in the non-CAKUT category.

Table 2. Genomic disorders in CKiD children

Chromosomal region	Start (Mb)	End (Mb)	Size (Mb)	CNV type	Syndrome	Clinical diagnosis	Patient ID	Age at diagnosis (yr)	eGFR	Cystatin C (Siemens, mg/l)
1q21.1	144.80	146.30	1.50	Dup	1q21.1 Recurrent microduplication[A]	Focal segmental glomerulosclerosis	54	0.0	29.9	2.8
1q21.1	144.80	146.30	1.50	Del	1q21.1 Recurrent microdeletion[A]	Hemolytic uremic syndrome	153	2.5	47.0	2.0
1q21.1	144.94	146.30	1.36	Dup	1q21.1 Recurrent microduplication[A]	Reflux nephropathy	322	0.0	51.1	1.4
2q13	110.21	110.34	0.13	Del	2q13 Homozygous deletion (nephronophthisis 1)[A]	Chronic glomerulonephritis	90	9.5	40.7	1.8
2q13	110.21	110.34	0.13	Del	2q13 Homozygous deletion (nephronophthisis 1)[A]	RHD	236	9.2	28.8	2.8
4p16.3-p15.33	0.06	14.53	14.47	Del	Wolf-Hirschhorn[A]	RHD	282	0.1	37.4	2.3
15q24.1-q24.2	70.75	73.86	3.11	Del	15q24 BP0-BP1 Deletion (includes BBS4, PMI1)[A]	Membranous nephropathy	300	7.5	28.1	2.5
16p11.2	29.49	30.11	0.62	Del	16p11.2 Deletion[A]	Obstructive uropathy	311	0.5	38.9	2.1
16p11.2	29.49	30.11	0.62	Del	16p11.2 Deletion[A]	RHD	377	0.0	53.4	1.4
17p12	14.04	15.42	1.38	Del	Hereditary liability to pressure palsies	RHD	368	13.5	65.0	1.0
17q12	31.89	33.32	1.43	Del	RCAD deletion[A]	RHD	360	0.0	41.7	2.0
17q12	31.89	33.32	1.43	Del	RCAD deletion[A]	RHD	390	0.0	92.0	0.8
17q12	31.89	33.32	1.43	Del	RCAD deletion[A]	Glomerular other	146	6.5	42.3	1.6
17p13.3	3.45	3.51	0.06	Del	CTNS homozygous deletion[A]	Cystinosis	136	0.5	25.1	2.0
17p13.3	3.45	3.51	0.06	Del	CTNS homozygous deletion[A]	Cystinosis	157	0.8	32.6	1.9
17p13.3	3.45	3.51	0.06	Del	CTNS homozygous deletion[A]	Cystinosis	144	0.8	30.7	2.3
X	0.00	154.91	154.9	Dup	XXX[A]	Recessive polycystic kidney disease	314	0.0	31.6	2.0
X	0.00	154.91	154.9	Dup	XXX[A]	Reflux nephropathy	115	11.5	29.3	2.8
X	0.00	154.91	154.9	Dup	XXX[A]	RHD	429	1.5	48.8	1.6

[A]Genomic disorders with known associations with nephropathy or genitourinary tract defects. Start and end positions (in Mb) are based on UCSC genome build hg18. CNV type, deletion (Del), or duplication (Dup), clinical diagnosis and deidentified patient number are indicated for each instance of a known genomic disorder found in CKiD cases. eGFR is in ml/min/1.73 m^2.

tent with our prior study (15), the subset of patients clinically diagnosed with renal hypodysplasia (RHD) was particularly enriched for known genomic disorders (7 of 67 cases [10.45%]; OR 30.09, $P = 4.91 \times 10^{-16}$, vs. controls, Tables 1 and 2). However, the excess of known genomic disorders was detected among other clinical causes of CKD, including patients previously diagnosed with glomerular disorders (Table 2). The excess of genomic disorders was significant after adjusting for sex and race/ethnicity (Table 1) or after comparison with only pediatric controls (Supplemental Table 2). The most frequent genomic imbalances were 17q12 deletion (HNFI homeobox B [*HNF1B*], diagnostic of renal cysts and diabetes [RCAD] syndrome, 3 cases vs. 0 controls, $P = 7 \times 10^{-6}$), chromosome (Chr) X duplication (XXX syndrome, 3 cases vs. 1 control, $P = 3 \times 10^{-5}$), 1q21.1 deletion/duplication (3 cases vs. 9 controls, $P = 0.001$), and cystinosin (CTNS) homozygous deletion (cystinosis, 3 cases vs. 0 controls, $P = 7 \times 10^{-6}$).

Likely pathogenic CNVs and candidate novel syndromes. After exclusion of the 19 cases with known genomic disorders, we still detected an excess of large, rare gene-disrupting CNVs ($P = 1.38 \times 10^{-7}$, Table 2, Supplemental Table 3, and Supplemental Figure 2). For example, we detected 35 CKiD cases with CNVs greater than 500 kb, which is nearly twice the frequency seen in controls (Supplemental Table 3). Annotation of these CNVs according to strict criteria adapted from prior recommendations for interpretation of microarray data revealed another 12 unrelated individuals (2.9% of the cohort) who carried a likely pathogenic imbalance, including one patient with a complex rearrangement on Chr 12 and 7 patients with CNVs larger than 1 Mb (Table 3). These lesions fulfilled very strict criteria for pathogenicity and would be considered reportable in a clinical setting (7). These CNVs disrupted genes implicated in kidney development (e.g., *FOXF1*, *FOX1*, *HOXD10/HOXD13*, *CDH19*, *CDH7*, and *ERRB4*), identifying novel candidate genes for human kidney disease. This subset of patients was not enriched for any particular clinical diagnosis, again demonstrating enrichment of CNVs across multiple forms of CKD.

Individuals with multiple rare gene-disrupting CNVs. Recent studies have reported a high prevalence of second-site CNVs in patients with known genomic disorders (16, 18–20). Among the 31 patients with known or likely pathogenic CNVs, 7 (23%) carried a second-site rare CNV, which is comparable to studies of patients with developmental delay (16, 18–20). We also identified 7 additional individuals carrying

Table 3. Likely pathogenic CNVs in CKiD children

Chromosomal region	CNV type	Start (Mb)	End (Mb)	Size (Mb)	Clinical diagnosis	ID	Evidence of likely pathogenicity
2p11.2	Del	87.48	89.91	2.43	Reflux nephropathy	127	Includes region of novel likely pathogenic deletion associated with CAKUT (15) (>1 Mb); includes *EIF2AK3* Wolcott-Rallison syndrome gene, which can have renal manifestations
2q21.1	Dup	131.20	131.71	0.50	Obstructive uropathy	32	Includes *ARHGEF4* (Asef), an OMIM gene associated with tubular injury (31); overlaps (>70%) with deletion in Decipher patient 2311
2q31.1	Dup	176.51	177.04	0.53	Obstructive uropathy	38	Includes *HOXD13* (brachidactyly; VACTERL association); *HOXD10* (foot deformity of Charcot-Marie-Tooth disease; vertical talus); includes duplicated region in Decipher patient 2055; includes and overlaps ISCA database pathogenic deletion regions
2q34	Del	211.55	212.17	0.63	Nonglomerular other	350	Includes gene *ERBB4*: implicated in kidney development (32) and polycystic kidney disease (in cpk mice; ref. 33) and associated with diabetic nephropathy in a GWAS (34)
4p15.2	Dup	25.09	25.86	0.77	Obstructive uropathy	28	Includes OMIM gene *SLC34A2* (testicular microlithiasis); overlaps duplicated region in Decipher patient 248972; included in ISCA database pathogenic CNV region.
5q35.1	Del	168.43	170.57	2.14	RHD	348	>2 Mb; includes *FOXI1*, involved in collecting duct development in mice (35), included in ISCA database pathogenic CNV
9q34.3	Dup	138.05	139.88	1.83	Focal segmental glomerulosclerosis	58	>1 Mb; includes 9q subtelomeric deletion region and partially overlaps with 9q34 duplication region.
10q11.22-q11.23	Del	46.42	51.50	5.09	Chronic glomerulonephritis	152	>5 Mb; overlaps (>70%) with ISCA database pathogenic deletions and CNVs in Decipher patients; includes OMIM Cockayne syndrome type B gene; and ChAT
12p13.33-p13.31	Dup	0.08	7.28	7.21	Focal segmental glomerulosclerosis	271	>7 Mb; includes *WNK1* and other OMIM genes; overlaps (>70%) region of ISCA database likely pathogenic duplications
12q24.33	Dup	129.08	131.57	2.49	Focal segmental glomerulosclerosis	271	>2 Mb; overlaps regions of pathogenic deletions and duplications in ISCA database and duplications in Decipher patients; part of a larger complex rearrangement
12q24.33	Del	131.58	132.29	0.71	Focal segmental glomerulosclerosis	271	Overlaps (>70%) region of ISCA database pathogenic deletion; part of a larger complex rearrangement
15q21.3	Dup	53.25	53.94	0.69	Nonglomerular other	210	Includes region of ISCA database pathogenic duplication, OMIM genes
16q24.1	Dup	84.03	85.30	1.27	Reflux nephropathy	169	Includes *FOXF1* (possibly associated with VACTERL; ref. 36), *FOXC2* (lymphedema-distichiasis syndrome with renal disease and diabetes mellitus); disrupts Wilms' tumor 3 gene and overlaps with larger pathogenic CNVs regions in the ISCA database
18q21.33-q22.1	Del	59.64	62.45	2.82	Obstructive uropathy	94	Overlaps with (1 > 70%, many larger) ISCA database pathogenic deletion regions, included in 18q del syndrome region; includes *CDH19* and *CDH7*

Start and end positions (in Mb) are based on UCSC genome build hg18. CNV type, deletion or duplication, and clinical diagnosis are indicated. Evidence of likely pathogenicity is also indicated.

at least 2 variants of unknown significance (VOUS) (Table 4). In total, 14 of 419 CKiD unrelated cases (3.3%), but only 194 of 19,685 controls (1%), harbored 2 or more rare, large CNVs, representing a significant enrichment (OR = 3.47, $P = 1.25 \times 10^{-03}$, Table 4). The 14 individuals with 2 or more rare gene-disrupting CNVs carried a broad range of clinical diagnoses, further supporting the association of large, gene-disrupting CNVs with all-cause pediatric CKD.

Clinical correlations and potential impact of CNVs for care plan. We found no differences in baseline clinical and demographic variables between carriers of pathogenic CNVs and noncarriers (sex and ethnic distribution, age of diagnosis, height, body mass index, Tanner stage, or blood pressure; Supplemental Table 4). However, carriers of pathogenic CNVs had a nominally reduced estimated glomerular filtration rate (eGFR), elevated cystatin C levels, and increased proteinuria at study enrollment, even after adjustment for variables such as age, sex, age at diagnosis, and duration of CKD (e.g., median Schwartz eGFR 37.5 vs. 46.4 ml/min, $P = 0.03$, uncorrected for multiple testing, Supplemental Table 4). The absence of major baseline differences is consistent with the exclusion criteria of clinically discernible syndromic disease or known major chromosomal disorders for the CKiD study, while the nominal difference in eGFR and proteinuria suggests that CNVs may affect kidney function.

In the 31 patients with a known or likely pathogenic genomic imbalance, we determined whether the genomic diagnosis was consistent with the clinical diagnosis and provided additional information that potentially could alter management. We found that 3 of 8 patients with a clinical diagnosis of cystinosis were homozygous for a known recurrent 63-kb *CTNS*-disrupting deletion (21). In these cases, the CNV analysis can be considered to be confirmatory of the clinical diagnosis (positive control) and could also provide precise mutation information for family screening. We also identified 3 additional cystinosis patients who were heterozygous for the same *CTNS* deletion, suggesting that these individuals carry a point mutation in the second *CTNS* allele (these individuals were not included in the count for genomic disorders).

In the remaining 28 patients with a known or likely pathogenic imbalance, the genomic diagnosis was unsuspected based on the clinical assessment and either resulted in reclassification of the disease or provided additional information that would have

Table 4. CKiD children with 2 or more pathogenic CNVs, likely pathogenic CNVs or VOUS

Chromosomal region	CNV type	Start (Mb)	End (Mb)	Size (Mb)	Clinical diagnosis	CNV combination	Patient
2q31.1	Dup	176.51	177.04	0.53	Obstructive uropathy	P+V	38
2q32.1	Dup	182.87	183.25	0.38			
1q21.1	Dup	144.80	146.30	1.50	Focal segmental glomerulosclerosis	P+V+V	54
4q34.1	Del	172.70	173.18	0.48			
4q34.3, 4q35.1	Del	182.22	182.79	0.56			
9q34.3	Dup	138.05	139.88	1.83	Focal segmental glomerulosclerosis	P+V	58
20q12	Dup	39.52	40.32	0.80			
19q13.12	Del	41.77	42.18	0.41	Reflux nephropathy	V+P	115
X	Dup	0.00	154.91	154.9			
7p21.3	Del	7.43	7.70	0.27	Nonglomerular other	V+V	143
13q21.32	Dup	65.39	66.17	0.78			
5q35.3	Dup	180.36	180.63	0.27	Chronic glomerulonephritis	P+V	152
10q11.22, 10q11.23	Del	46.42	51.50	5.09			
1p36.33	Dup	1.11	1.53	0.42	Hemolytic uremic syndrome	V+V	163
8q24.3	Dup	144.86	145.25	0.39			
1p13.3	Dup	109.25	109.62	0.37	Nonglomerular other	V+V	165
9q34.3	Dup	139.04	139.32	0.28			
1p36.33	Dup	1.12	1.47	0.35	Hemolytic uremic syndrome	V+V	191
2q31.1	Dup	174.20	174.52	0.32			
1p36.33	Dup	1.12	1.44	0.33	Reflux nephropathy	V+V	207
20q13.33	Dup	60.46	60.88	0.42			
12p13.33-p13.31	Dup	0.08	7.28	7.21	Focal segmental glomerulosclerosis	P+P+P+V	271
12q24.33	Dup	129.08	131.57	2.49			
12q24.33	Del	131.58	132.29	0.71			
Xp22.31	Dup	8.37	9.15	0.78			
18q21.31	Dup	53.49	53.86	0.38	RHD	V+V	359
19q13.31	Hom. Del.	48.01	48.43	0.42			
1q21.1,	Dup.	144.94	146.30	1.36	Reflux nephropathy	P+V	322
4q32.3,	Dup.	165.98	166.28	0.31			
4q26, 4q25	Dup	113.73	114.35	0.62	RHD	V+V	387
4q26	Dup	119.73	120.06	0.33			

Start and end positions (in Mb) are based on UCSC genome build hg18. CNV type, deletion, or duplication, clinical diagnosis and deidentified patient number are indicated for each instance. CNV combination is indicated (e.g., V+V = patients with 2 VOUS; V+P = patient with 1 VOUS and 1 known/likely pathogenic CNV). Hom., homozygous.

warranted genetic counseling, targeted workup, or surveillance based on published recommendations (Table 2 and Supplemental Table 5). For example, the CNV analysis reclassified disease to RCAD syndrome and nephronophthisis for 3 patients carrying imprecise clinical diagnoses of glomerulopathy or a nonspecific nonglomerular disorder (Table 2). In other patients diagnosed with isolated congenital defects, the CNV analysis implicated a genomic disorder that, if recognized, would have warranted a targeted workup or surveillance (e.g., screening for diabetes, hypomagnesemia, and hyperuricemia in patients with RCAD; other detailed targeted workups listed in Supplemental Table 5). The remaining 16 cases involved single instances of CNV-nephropathy associations, which can represent coincidental findings, phenotype expansion of known syndromes, or potentially novel genetic syndromes. Strikingly, the majority of pathogenic genomic imbalances detected in this population have a known association with developmental delay, intellectual disability, and/or seizure disorders (16), which would warrant screening for neuropsychiatric illness (Supplemental Table 5).

Discussion

In this study, we demonstrate that children with CKD harbor a 10-fold excess of large genomic imbalances that were not suspected based on standard clinical evaluation. This high genomic load was supported by an analysis of global CNV burden by detailed characterization of CNVs, ultimately identifying 31 patients with a known or likely pathogenic imbalance. These lesions fulfilled very strict criteria for pathogenicity and would be considered reportable in the clinical setting. The 7.4% detection rate was comparable to the yield for microarrays obtained for prenatal diagnosis of major developmental disorders (8). Of note, the known genomic imbalances were identified based on their genomic coordinates and their well-known involvement in developmental disorders. Thus, the criteria for detection of these known genomic disorders were independent of their frequency in controls. In fact, our study likely underestimated the prevalence of pathogenic imbalances in pediatric CKD because the CKiD study excluded major chromosomal disorders and clinically discernible syndromic diseases. These data suggest that children with CKD merit a thorough clinical evalua-

tion for subtle signs of syndromic disorders, and a chromosomal microarray should be considered in the diagnostic workup, particularly for those individuals with RHD.

In the CKiD cohort, the diagnoses were provided by the subject's pediatric nephrology center. We show that in many cases, microarray analysis can provide an alternative diagnosis or a personal genomic diagnosis that defines risk for specific extrarenal disorders, warranting a targeted follow-up based on current recommendations (Supplemental Table 5). For example, the 17q12 deletion (RCAD) warrants regular monitoring for diabetes, hypomagnesemia, and hyperuricemia as well as a screen for uterine abnormalities; the XXX syndrome warrants periodic EEG monitoring and screening for hypogonadism; and the 1q21.1 deletion/duplication syndrome raises considerations for the performance of an ophthalmologic exam, a cardiac evaluation, and neuroimaging studies. The detection of genomic imbalances also has immediate implications for genetic counseling and family planning. Most importantly, the majority of genomic imbalances we detected are associated with the risk of developmental delay, learning disability, and other neuropsychiatric disorders that benefit from early detection and intervention (Supplemental Table 5). These findings should alert clinicians that poor neuropsychiatric performance or behavioral disorders in children with CKD may not always be attributable to the burden of chronic illness, but may reflect an unsuspected primary genetic disorder that directly impairs both kidney and neurocognitive function.

It is known that many CNV disorders have pleiotropic effects on organ function, simultaneously predisposing to cardiac malformations, metabolic disorders, autism, schizophrenia, intellectual disability, and seizure disorders (9, 10, 22–24). We previously detected an excess of CNV disorders in patients with congenital kidney malformations (15), but this study now implicates CNV disorders in the pathogenesis of clinically diverse forms of CKD. For example, we had 7.7% and 7.1% detection rates among patients diagnosed with congenital and noncongenital forms of CKD, respectively (Table 1). How do genomic imbalances contribute to CKD of diverse etiology? One can hypothesize that some CNVs have a direct causal role by inactivating known kidney disease genes (e.g., *HNF1B*), while others may have a more generalized effect on kidney growth, nephron number, and renal reserve, serving as progression factors for primary kidney diseases, such as glomerulonephritis or thrombotic disorders. Finally, CNVs can also contribute to kidney failure indirectly by augmenting the burden of comorbid conditions; for example, concomitant neuropsychiatric illness can result in suboptimal adherence to medical regimens prescribed to reduce the progression of CKD. These hypotheses can be pursued by analysis of animal models and larger human cohorts to permit genetic dissection of CNVs and elucidation of their effects on the development of disease and renal progression.

In recent years, chromosomal microarray analysis has emerged as a major diagnostic tool for evaluation of congenital anomalies or neuropsychiatric disorders (7, 8, 25). Our data further support the utility of this technology for diagnosis of pediatric CKD. Based on prior studies that have indicated that about half of pathogenic imbalances in congenital defects and developmental delay are inherited (9, 10, 22–24), parental testing would be expected to be similarly informative for genetic counseling and family planning in pediatric CKD. It is very likely that the diagnostic yield would be even higher if chromosomal microarrays are applied to pediatric CKD populations with concomitant neurocognitive deficits, structural heart disease, or end-stage organ failure (these were all exclusion criteria for the CKiD study). In the near future, whole-genome sequencing, which allows simultaneous detection of pathogenic CNVs and single-nucleotide variants, will emerge as the preferred modality for genomic diagnostics and will likely augment the overall diagnostic yield. Nephrology may be ideally positioned to spearhead the adoption of genomic diagnostics in clinical practice. Perinatal imaging and lab tests are routinely performed to assess kidney development and function, and detection of CKD could become an indication to search for genomic imbalances to achieve personalized medical interventions at an early stage.

Methods

Cases and controls. The CKiD study is an NIH-sponsored longitudinal observational study of children with CKD from 48 participating clinical sites across North America. Details of the CKiD study design, consent procedures, and adjudication of causes for CKD are described elsewhere (2, 3). Briefly, specific aims of CKiD include the following: (a) identifying and quantifying novel and traditional risk factors for progression of CKD; (b) characterizing CKD progression effects on neurodevelopment, cognitive abilities, and behavior; (c) describing the prevalence of cardiovascular disease and associated risk factors; and (d) examining the effects of declining kidney function on growth in children with CKD. Participants are followed longitudinally until they are 21 years of age, transplanted, or transferred to an adult center. A total of 439 individuals consented to the genetic substudies. The primary clinical diagnoses are listed in Table 1 and Supplemental Table 6A. CKiD cases included 289 mixed European (68.2%), 71 black or African American (16.7%), 10 Asian (2.4%), and 54 (12.7%) children of other or unknown ancestry (CKiD cases, Supplemental Table 6B). Raw data from this study are available in the NCBI's dbGaP database (phs000650.v1.p1; http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000650.v1.p1) and in the NIDDK Central Repository (https://www.niddkrepository.org/studies/ckid/). The control population consisted of a multiethnic cohort of 21,575 individuals (9% pediatric) obtained from 11 different studies with genome-wide genotyping on high-density Illumina platforms as part of case-control or longitudinal studies of complex traits unrelated to nephropathy or healthy controls for studies related to nephropathy: 7 studies of Parkinson disease, Alzheimer disease, blood clotting, melanoma, adiposity, addiction, and hypertension (dbGaP accession numbers phs000196.v2.p1, phs000168.v1.p1, phs000304.v1.p1, phs000187.v1.p1, phs000092.v1.p1); 2 control data sets from studies of IgA nephropathy (phs000431.v2.p1); and 1 reference pediatric control data set from the CHOP (phs000199.v1.p1); see Supplemental Table 7A. The controls included 18,147 mixed European (84.1%), 2,357 black or African American (10.9%), 897 Asian (4.2%), and 174 (0.8%) individuals of other or unknown ancestry (Supplemental Table 7B).

Data set and CNV calls. DNA was obtained from lymphoblastoid cell lines derived from peripheral blood cells at the NIDDK biorepository. Parental DNA was not collected in the CKiD study. CKiD samples were genotyped on Illumina Omni2.5 microarrays, while control samples were genotyped on Illumina arrays of Hap550v1 or

higher, which share a similar backbone and have a significant number of overlapping probes (Supplemental Table 7B). To avoid potential bias in variant calls, primary microarray data files were used for ab initio processing and variant calls, using the same standardized method in both cases and controls. Raw intensity data were processed in GenomeStudio v2011 (Illumina). PennCNV (26) was used to determine CNV calls. CNVs were mapped to the human reference genome hg18 and annotated with UCSC RefGene and RefExon using the CNVision program (27). Only CNVs with confidence scores of 30 or more were considered in the analyses based on experimental validation from our prior study (15). To minimize the potential effect of cell line artifacts in the analyses, individuals with more than 19 (99th percentile) large (≥ 100 kb), rare (frequency ≤ 0.02%) CNVs were excluded from both CKiD cases and controls. A total of 424 (419 unrelated) CKiD participants passed quality control (QC) filters, performed with GenomeStudio v.2011 (Illumina), PLINK, and PennCNV software (26, 28). All CNVs reported in Tables 1 through 4 were examined visually in Illumina Genome Viewer 1.9.0 to rule out possible artifacts. CNV frequencies were calculated on the basis of the entire control data set of 21,575 individuals. Similar to our prior study (27), 2 CNVs were considered to be identical if they had the same copy number value and had 70% or greater reciprocal overlap; otherwise, they were considered to be distinct.

Known genomic disorders and likely pathogenic CNVs. We defined known genomic disorders when we detected 70% or greater CNV overlap with the coordinates of 1 of 131 known syndromes whose coordinates were taken from the Decipher database (https://decipher.sanger.ac.uk/) and 3 recent studies (Supplemental Table 1 and refs. 8, 9, 15). Each one of the 131 known genomic disorders had frequencies of 1% or less in controls. The criteria for likely pathogenic and reportable CNVs were adapted from prior recommendations for interpretation of microarray data (7, 8, 25, 29): (a) CNV size of 500 kb or greater, with frequency of 0.02% or less in controls and absence in the CHOP cohort (since this represents a reference healthy pediatric population; ref. 30) and (b) partial overlap with a known genomic disorder or overlap with likely pathogenic CNVs reported in the International Standards for Cytogenomic Arrays (ISCA) database (http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000205.v2.p1) or in our previous study on RHD and/or gene content relevant to kidney development or pathology.

Analysis of CNV burden and 2-hit CNV model. Only the 419 unrelated individuals were considered for proportions and burden analyses. CNV burden analysis was performed by counting large (size ≥ 100 kb), rare (frequency < 0.1% in any single cohort and in the whole control data set), autosomal CNVs that contained or intersected at least 1 exonic sequence. CNVs larger than 10 Mb were excluded from the burden analysis. The proportions of cases and controls with the largest CNVs within given size ranges were compared. We also examined the population frequency of the largest CNV per genome. This analysis was repeated after exclusion of individuals carrying a known genomic disorder. To evaluate a 2-hit CNV model, we selected unrelated patients carrying 2 or more CNVs absent in the CHOP pediatric cohort that were either a known syndrome or a gene-disrupting CNV of 250 or greater kb with a frequency of less than 0.02% in controls.

Statistics. To examine global CNV burden, population frequencies of the largest CNV per genome were analyzed using a log-rank test (SPSS IBM v.21). The proportions of cases and controls with the largest CNVs within given size ranges and with 2-CNV hits were compared using 2-sided Fisher's exact tests (R v2.14). The distribution of genomic disorders between cases and controls was compared by logistic regression, correcting for race/ethnicity and sex (R v2.14). To control for age, we also conducted a separate comparison of known genomic disorders in CKiD cases versus pediatric controls. *P* values of less than 0.005, corrected for 10 independent analyses, were considered significant for the above analyses. To compare clinical variables between cases with and without genomic disorders, nonparametric Wilcoxon test was performed on residuals after adjusting for covariates by logistic regression (R v2.14). Nominal *P* values are reported in Supplemental Table 4, but none of the associations with clinical variables were significant after multiple hypothesis testing.

Study approval. The study was approved by the Institutional Review Board at Columbia University, the CKiD participating sites, and the CKiD steering committee. Signed written informed consent by parents or guardians plus the participant children's assent were obtained for CKiD genetic substudies, according to specific IRB requirements for each site.

Acknowledgments

We thank the patients for participating in this study. This study was supported by grants RO1DK082394 (to C.S. Wong)and 1R01DK080099 and 1U54DK104309 (to A.G. Gharavi). S. Sanna-Cherchi is supported by grants R21DK098531, R01DK103184, American Heart Association (AHA) 13GRNT14680075, and the New York State Empire Clinical Research Investigator Program (ECRIP). K. Kiryluk is supported by K23DK090207, R03DK099564, an American Society of Nephrology Carl W. Gottschlak Research Scholar Grant, and the New York State ECRIP. M. Wuttke and A. Köttgen were funded by the German Research Foundation (DFG KO 3598/2-1 to A. Köttgen). The genotyping data utilized for this study were collected with support by the NIDDK (RO1DK082394). Data in this manuscript were collected by the CKiD prospective cohort study, with clinical coordinating centers (principal investigators) at Children's Mercy Hospital and the University of Missouri — Kansas City (Bradley Warady) and CHOP (Susan Furth), the central laboratory (principal investigator) at the Department of Pediatrics, University of Rochester Medical Center (George Schwartz), and the data coordinating center (principal investigator) at the Johns Hopkins Bloomberg School of Public Health (Alvaro Muñoz). CKiD is funded by the NIDDK, with additional funding from the National Institute of Child Health and Human Development and the National Heart, Lung, and Blood Institute (U01 DK066143, U01 DK066174, U01 DK082194, U01 DK066116). The CKiD website is located at http://www.statepi.jhsph.edu. We thank the investigators who made their data available via dbGAP for public use. The full list of funding sources for dbGAP studies is in the Supplemental Acknowledgments.

Address correspondence to: Craig S. Wong, Division of Pediatric Nephrology, University of New Mexico Children's Hospital, MSC10-5590 1, University of New Mexico, Phone: 505.272.3887; E-mail: cwong@salud.unm.edu. Or to: Ali Gharavi, Columbia University, College of Physicians and Surgeons, Department of Medicine, Division of Nephrology, 1150 St. Nicholas Ave., Russ Berrie Pavilion Room 413, New York, New York 10032, USA. Phone: 212.851.5556; E-mail: ag2239@columbia.edu.

1. US Renal Data System. *USRDS 2013 Annual Data Report: Atlas of End-Stage Renal Disease in the United States*. Bethesda, Maryland, USA: National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases; 2013.
2. Furth SL, et al. Design and methods of the Chronic Kidney Disease in Children (CKiD) prospective cohort study. *Clin J Am Soc Nephrol*. 2006;1(5):1006–1015.
3. Wong CJ, Moxey-Mims M, Jerry-Fluker J, Warady BA, Furth SL. CKiD (CKD in children) prospective cohort study: a review of current findings. *Am J Kidney Dis*. 2012;60(6):1002–1011.
4. Gerson AC, et al. Neurocognitive outcomes in children with chronic kidney disease: Current findings and contemporary endeavors. *Ment Retard Dev Disabil Res Rev*. 2006;12(3):208–215.
5. Hooper SR, et al. Neurocognitive functioning of children and adolescents with mild-to-moderate chronic kidney disease. *Clin J Am Soc Nephrol*. 2011;6(8):1824–1830.
6. Khaja R, et al. Genome assembly comparison identifies structural variants in the human genome. *Nat Genet*. 2006;38(12):1413–1418.
7. Miller DT, et al. Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet*. 2010;86(5):749–764.
8. Wapner RJ, et al. Chromosomal microarray versus karyotyping for prenatal diagnosis. *N Engl J Med*. 2012;367(23):2175–2184.
9. Cooper GM, et al. A copy number variation morbidity map of developmental delay. *Nat Genet*. 2011;43(9):838–846.
10. Greenway SC, et al. De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot. *Nat Genet*. 2009;41(8):931–935.
11. Liu X, et al. Increased rate of sporadic and recurrent rare genic copy number variants in Parkinson's disease among Ashkenazi Jews. *Mol Genet Genomic Med*. 2013;1(3):142–154.
12. Marshall CR, et al. Structural variation of chromosomes in autism spectrum disorder. *Am J Hum Genet*. 2008;82(2):477–488.
13. Silversides CK, et al. Rare copy number variations in adults with tetralogy of Fallot implicate novel risk gene pathways. *PLoS Genet*. 2012;8(8):e1002843.
14. van Duyvenvoorde HA, et al. Copy number variants in patients with short stature. *Eur J Hum Genet*. 2014;22(5):602–609.
15. Sanna-Cherchi S, et al. Copy-number disorders are a common cause of congenital kidney malformations. *Am J Hum Genet*. 2012;91(6):987–997.
16. Girirajan S, et al. Phenotypic heterogeneity of genomic disorders and rare copy-number variants. *N Engl J Med*. 2012;367(14):1321–1331.
17. Itsara A, et al. Population analysis of large copy number variants and hotspots of human genetic disease. *Am J Hum Genet*. 2009;84(2):148–161.
18. Pinto D, et al. Functional impact of global rare copy number variation in autism spectrum disorders. *Nature*. 2010;466(7304):368–372.
19. Leblond CS, et al. Genetic and functional analyses of SHANK2 mutations suggest a multiple hit model of autism spectrum disorders. *PLoS Genet*. 2012;8(2):e1002521.
20. Girirajan S, et al. A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay. *Nat Genet*. 2010;42(3):203–209.
21. Shotelersuk V, et al. CTNS mutations in an American-based population of cystinosis patients. *Am J Hum Genet*. 1998;63(5):1352–1362.
22. Pescosolido MF, Gamsiz ED, Nagpal S, Morrow EM. Distribution of disease-associated copy number variants across distinct disorders of cognitive development. *J Am Acad Child Adolesc Psychiatry*. 2013;52(4):414–430.
23. Guilmatre A, et al. Recurrent rearrangements in synaptic and neurodevelopmental genes and shared biologic pathways in schizophrenia, autism, and mental retardation. *Arch Gen Psychiatry*. 2009;66(9):947–956.
24. Glessner JT, et al. Increased frequency of de novo copy number variants in congenital heart disease by integrative analysis of single nucleotide polymorphism array and exome sequence data. *Circ Res*. 2014;115(10):884–896.
25. Reddy UM, et al. Karyotype versus microarray testing for genetic abnormalities after stillbirth. *N Engl J Med*. 2012;367(23):2185–2193.
26. Wang K, et al. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res*. 2007;17(11):1665–1674.
27. Sanders SJ, et al. Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron*. 2011;70(5):863–885.
28. Purcell S, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559–575.
29. Kearney HM, et al. American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants. *Genet Med*. 2011;13(7):680–685.
30. Shaikh TH, et al. High-resolution mapping and analysis of copy number variations in the human genome: a data resource for clinical and research applications. *Genome Res*. 2009;19(9):1682–1690.
31. Cheng HT, Juang IP, Chen LC, Lin LY, Chao CH. Association of Asef and Cdc42 expression to tubular injury in diseased human kidney. *J Investig Med*. 2013;61(7):1097–1103.
32. Veikkolainen V, et al. ErbB4 modulates tubular cell polarity and lumen diameter during kidney development. *J Am Soc Nephrol*. 2012;23(1):112–122.
33. Zeng F, Miyazawa T, Kloepfer LA, Harris RC. Deletion of ErbB4 accelerates polycystic kidney disease progression in cpk mice. *Kidney Int*. 2014;86(3):538–547.
34. Sandholm N, et al. New susceptibility loci associated with kidney disease in type 1 diabetes. *PLoS Genet*. 2012;8(9):e1002921.
35. Al-Awqati Q, Schwartz GJ. A fork in the road of cell differentiation in the kidney tubule. *J Clin Invest*. 2004;113(11):1528–1530.
36. Agochukwu NB, et al. Analysis of FOXF1 and the FOX gene cluster in patients with VACTERL association. *Eur J Med Genet*. 2011;54(3):323–328.