Genomic imbalances in pediatric patients with chronic kidney disease

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


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 × 10^{-20}). In an additional […]

Find the latest version:
http://jci.me/80877/pdf
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,5 Frederick Kaskel,5 Anna Köttgen,3,4 Bradley A. Warady,6 Susan L. Furth,7 Craig S. Wong,4 and Ali G. Gharavi1

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

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 (4), 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.

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.
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 × 10^{-20}). 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 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 × 10^{-20}). Consis-

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

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Individuals with a known genomic disorder (%)</th>
<th>Odds ratio (P value)</th>
<th>Individuals with a likely pathogenic CNV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All CKD</td>
<td>419</td>
<td>19 (4.5%)</td>
<td>10.8 (6.08 × 10^{-20})</td>
<td>12 (2.9%)</td>
</tr>
<tr>
<td>CAKUT</td>
<td>221</td>
<td>10 (4.5%)</td>
<td>11.0 (2.96 × 10^{-10})</td>
<td>7 (3.2%)</td>
</tr>
<tr>
<td>CAKUT due to RHD</td>
<td>67</td>
<td>7 (10.5%)</td>
<td>30.1 (4.91 × 10^{-4})</td>
<td>1 (1.5%)</td>
</tr>
<tr>
<td>Non-CAKUT</td>
<td>198</td>
<td>9 (4.6%)</td>
<td>10.3 (8.74 × 10^{-4})</td>
<td>5 (2.5%)</td>
</tr>
<tr>
<td>Controls</td>
<td>21,575</td>
<td>98 (0.5%)</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

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.
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 = 7 × 10–6). In addition to the excess of known genomic disorders, we still observed an excess of large, rare gene-disrupting CNVs (P = 4.91 × 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) and after comparison with only pediatric controls (Supplemental Table 3, 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., FOXC1, CDH7, CDH19, CDH8, HOXD10/13, 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 2. Genomic disorders in CKiD children

<table>
<thead>
<tr>
<th>Chromosomal region</th>
<th>CNV type</th>
<th>Syndrome</th>
<th>Clinical diagnosis</th>
<th>Patient ID</th>
<th>Age at diagnosis (yr)</th>
<th>eGFR (ml/min/1.73 m²)</th>
<th>Cystatin C (mg/l)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1q21.1</td>
<td>Dup</td>
<td>1q21.1 Recurrent microduplication</td>
<td>Focal segmental glomerulosclerosis</td>
<td>54</td>
<td>0.0</td>
<td>29.9</td>
<td>2.8</td>
</tr>
<tr>
<td>1q21.1</td>
<td>Del</td>
<td>1q21.1 Recurrent microdeletion</td>
<td>Hemolytic uremic syndrome</td>
<td>153</td>
<td>2.5</td>
<td>47.0</td>
<td>2.0</td>
</tr>
<tr>
<td>17q12</td>
<td>Del</td>
<td>RCAD deletion</td>
<td>RHD</td>
<td>360</td>
<td>0.0</td>
<td>41.7</td>
<td>2.0</td>
</tr>
<tr>
<td>17p12</td>
<td>Del</td>
<td>RHD</td>
<td>Obstructive uropathy</td>
<td>311</td>
<td>0.5</td>
<td>38.9</td>
<td>2.1</td>
</tr>
<tr>
<td>17p12</td>
<td>Del</td>
<td>RHD</td>
<td>Cystinosis</td>
<td>377</td>
<td>0.0</td>
<td>53.4</td>
<td>1.4</td>
</tr>
<tr>
<td>17q12</td>
<td>Del</td>
<td>RHD</td>
<td>Membranous nephropathy</td>
<td>300</td>
<td>7.5</td>
<td>28.1</td>
<td>2.5</td>
</tr>
<tr>
<td>17q12</td>
<td>Del</td>
<td>RHD</td>
<td>Glomerular other</td>
<td>146</td>
<td>6.5</td>
<td>42.3</td>
<td>1.6</td>
</tr>
<tr>
<td>17p13.3</td>
<td>Del</td>
<td>RHD</td>
<td>Cystinosis</td>
<td>136</td>
<td>0.5</td>
<td>25.1</td>
<td>2.0</td>
</tr>
<tr>
<td>17p13.3</td>
<td>Del</td>
<td>RHD</td>
<td>Cystinosis</td>
<td>157</td>
<td>0.8</td>
<td>32.6</td>
<td>1.9</td>
</tr>
<tr>
<td>17p13.3</td>
<td>Del</td>
<td>RHD</td>
<td>Recessive polycystic kidney disease</td>
<td>314</td>
<td>0.0</td>
<td>31.6</td>
<td>2.0</td>
</tr>
</tbody>
</table>

^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 m².
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 × 10^{-6}, 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 clinical age 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 3. Likely pathogenic CNVs in CKiD children

<table>
<thead>
<tr>
<th>Chromosomal region</th>
<th>CNV Type</th>
<th>Start (Mb)</th>
<th>End (Mb)</th>
<th>Size (Mb)</th>
<th>Clinical Diagnosis</th>
<th>ID</th>
<th>Evidence of likely pathogenicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>2p11.2</td>
<td>Del</td>
<td>87.48</td>
<td>89.91</td>
<td>2.43</td>
<td>Reflux nephropathy</td>
<td>127</td>
<td>Includes region of novel likely pathogenic deletion associated with CAKUT (15) (&gt;1Mb); includes EF2A2 Wolcott-Rallison syndrome gene, which can have renal manifestations</td>
</tr>
<tr>
<td>2q21.1</td>
<td>Dup</td>
<td>131.20</td>
<td>131.71</td>
<td>0.50</td>
<td>Obstructive uropathy</td>
<td>32</td>
<td>Includes ARHGEF4 (Asef), an OMIM gene associated with tubular injury (31); overlaps (&gt;70%) with deletion in Decipher patient 2311</td>
</tr>
<tr>
<td>2q31.1</td>
<td>Dup</td>
<td>176.51</td>
<td>177.04</td>
<td>0.53</td>
<td>Obstructive uropathy</td>
<td>38</td>
<td>Includes HOXD13 (brachidactyly; VACTERL association); HOXD10 (foot deformity of Charcot-Marie-Tooth disease; vertical talus); includes duplicated region in Decipher patient 2255; includes and overlaps ISCA database pathogenic deletion regions</td>
</tr>
<tr>
<td>2q34</td>
<td>Del</td>
<td>211.55</td>
<td>212.17</td>
<td>0.63</td>
<td>Nonglomerular other</td>
<td>350</td>
<td>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)</td>
</tr>
<tr>
<td>4p15.2</td>
<td>Dup</td>
<td>25.09</td>
<td>25.86</td>
<td>0.77</td>
<td>Obstructive uropathy</td>
<td>28</td>
<td>Includes OMIM gene SLC34A2 (testicular microcystis); overlaps duplicated region in Decipher patient 248972; included in ISCA database pathogenic CNV region.</td>
</tr>
<tr>
<td>5q35.1</td>
<td>Del</td>
<td>168.43</td>
<td>170.57</td>
<td>2.14</td>
<td>RHD</td>
<td>348</td>
<td>&gt;2 Mb; includes FOXF1, involved in collecting duct development in mice (35), included in ISCA database pathogenic CNV</td>
</tr>
<tr>
<td>9q34.3</td>
<td>Del</td>
<td>138.05</td>
<td>139.88</td>
<td>1.83</td>
<td>Focal segmental glomerulosclerosis</td>
<td>58</td>
<td>&gt;1Mb; includes 9q subtelomeric deletion region and partially overlaps with 9q34 duplication region.</td>
</tr>
<tr>
<td>10q11.22-11.23</td>
<td>Del</td>
<td>46.42</td>
<td>51.50</td>
<td>5.09</td>
<td>Chronic glomerulonephritis</td>
<td>152</td>
<td>&gt;5 Mb; overlaps (&gt;70%) with ISCA database pathogenic deletions and CNVs in Decipher patients; includes OMIM Cockayne syndrome type B gene; and ChAT</td>
</tr>
<tr>
<td>12p13.33-p13.31</td>
<td>Dup</td>
<td>0.08</td>
<td>7.28</td>
<td>7.21</td>
<td>Focal segmental glomerulosclerosis</td>
<td>271</td>
<td>&gt;7 Mb; includes WNK1 and other OMIM genes; overlaps (&gt;70%) region of ISCA database likely pathogenic duplications</td>
</tr>
<tr>
<td>12q24.33</td>
<td>Dup</td>
<td>129.08</td>
<td>131.57</td>
<td>2.49</td>
<td>Focal segmental glomerulosclerosis</td>
<td>271</td>
<td>&gt;2 Mb; overlaps regions of pathogenic deletions and duplications in ISCA database and duplications in Decipher patients; part of a larger complex rearrangement</td>
</tr>
<tr>
<td>12q24.33</td>
<td>Del</td>
<td>131.58</td>
<td>132.29</td>
<td>0.71</td>
<td>Focal segmental glomerulosclerosis</td>
<td>271</td>
<td>Overlaps (&gt;70%) region of ISCA database pathogenic deletion; part of a larger complex rearrangement</td>
</tr>
<tr>
<td>15q21.3</td>
<td>Dup</td>
<td>53.25</td>
<td>53.94</td>
<td>0.69</td>
<td>Nonglomerular other</td>
<td>210</td>
<td>Includes region of ISCA database pathogenic duplication, OMIM genes</td>
</tr>
<tr>
<td>16q24.1</td>
<td>Dup</td>
<td>84.03</td>
<td>85.30</td>
<td>1.27</td>
<td>Reflux nephropathy</td>
<td>169</td>
<td>Includes FOXF1 (possibly associated with VACTERL; ref. 36), FOX2 (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</td>
</tr>
<tr>
<td>18q21.33-q22.1</td>
<td>Del</td>
<td>59.64</td>
<td>62.45</td>
<td>2.82</td>
<td>Obstructive uropathy</td>
<td>94</td>
<td>Overlaps with (1 &gt; 70%, many larger) ISCA database pathogenic deletion regions, included in 18q del syndrome region; includes CDH9 and CDH7</td>
</tr>
</tbody>
</table>

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.
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 evaluation warranted genetic counseling, targeted workup, or surveillance based on published recommendations (Table 2 and Suppmental Table 5). For example, the CNV analysis reclassified disease to RCAD syndrome and nephropathies 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).

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

<table>
<thead>
<tr>
<th>Chromosomal region</th>
<th>CNV type</th>
<th>Start (Mb)</th>
<th>End (Mb)</th>
<th>Size (Mb)</th>
<th>Clinical diagnosis</th>
<th>CNV combination</th>
<th>Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>2q33.1</td>
<td>Dup</td>
<td>176.51</td>
<td>177.04</td>
<td>0.53</td>
<td>Obstructive uropathy</td>
<td>P+V</td>
<td>38</td>
</tr>
<tr>
<td>2q23.1</td>
<td>Dup</td>
<td>182.87</td>
<td>183.25</td>
<td>0.38</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1q21.1</td>
<td>Dup</td>
<td>144.80</td>
<td>146.30</td>
<td>1.50</td>
<td>Focal segmental glomerulosclerosis</td>
<td>P+V+V</td>
<td>54</td>
</tr>
<tr>
<td>4q34.1</td>
<td>Del</td>
<td>172.70</td>
<td>173.18</td>
<td>0.48</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4q34.3, 4q35.1</td>
<td>Del</td>
<td>182.22</td>
<td>182.79</td>
<td>0.56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9q34.3</td>
<td>Dup</td>
<td>138.05</td>
<td>139.88</td>
<td>1.83</td>
<td>Focal segmental glomerulosclerosis</td>
<td>P+V</td>
<td>58</td>
</tr>
<tr>
<td>20q12</td>
<td>Del</td>
<td>39.52</td>
<td>40.32</td>
<td>0.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19q13.12</td>
<td>Del</td>
<td>41.77</td>
<td>42.18</td>
<td>0.41</td>
<td>Reflux nephropathy</td>
<td>V+P</td>
<td>115</td>
</tr>
<tr>
<td>7p21.3</td>
<td>Del</td>
<td>7.43</td>
<td>7.70</td>
<td>0.27</td>
<td>Nonglomerular other</td>
<td>V+V</td>
<td>143</td>
</tr>
<tr>
<td>19q21.32</td>
<td>Del</td>
<td>65.39</td>
<td>66.17</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5q35.3</td>
<td>Dup</td>
<td>180.36</td>
<td>180.63</td>
<td>0.27</td>
<td>Chronic glomerulonephritis</td>
<td>P+V</td>
<td>152</td>
</tr>
<tr>
<td>10q11.22, 10q11.23</td>
<td>Del</td>
<td>46.42</td>
<td>51.50</td>
<td>5.09</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8q43.3</td>
<td>Del</td>
<td>111.1</td>
<td>115.33</td>
<td>4.23</td>
<td>Hemolytic urane nephritis</td>
<td>V+V</td>
<td></td>
</tr>
<tr>
<td>1p36.3</td>
<td>Del</td>
<td>144.86</td>
<td>145.25</td>
<td>0.39</td>
<td>Nonglomerular other</td>
<td>V+V</td>
<td>165</td>
</tr>
<tr>
<td>9q34.3</td>
<td>Del</td>
<td>139.04</td>
<td>139.32</td>
<td>0.28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1p36.33</td>
<td>Del</td>
<td>112.0</td>
<td>112.34</td>
<td>0.35</td>
<td>Hemolytic urane nephritis</td>
<td>V+V</td>
<td>191</td>
</tr>
<tr>
<td>2p33.1</td>
<td>Del</td>
<td>174.20</td>
<td>174.52</td>
<td>0.32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1p36.33</td>
<td>Del</td>
<td>112.0</td>
<td>114.44</td>
<td>0.33</td>
<td>Reflux nephropathy</td>
<td>V+V</td>
<td>207</td>
</tr>
<tr>
<td>20q13.33</td>
<td>Del</td>
<td>60.46</td>
<td>60.88</td>
<td>0.42</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12p13.33-p13.31</td>
<td>Del</td>
<td>5.08</td>
<td>7.28</td>
<td>7.21</td>
<td>Focal segmental glomerulosclerosis</td>
<td>P+P+P+P</td>
<td>271</td>
</tr>
<tr>
<td>12q23.3</td>
<td>Del</td>
<td>129.08</td>
<td>131.57</td>
<td>2.49</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12q24.3</td>
<td>Del</td>
<td>131.55</td>
<td>132.29</td>
<td>0.71</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1q22.31</td>
<td>Del</td>
<td>8.37</td>
<td>9.15</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18q11.21</td>
<td>Hom. Del.</td>
<td>53.49</td>
<td>53.86</td>
<td>0.38</td>
<td>RHD</td>
<td>V+V</td>
<td>359</td>
</tr>
<tr>
<td>19q13.11</td>
<td>Del</td>
<td>48.01</td>
<td>48.43</td>
<td>0.42</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1p22.11</td>
<td>Del</td>
<td>144.94</td>
<td>146.10</td>
<td>1.36</td>
<td>Reflux nephropathy</td>
<td>P+V</td>
<td>322</td>
</tr>
<tr>
<td>4q23.3</td>
<td>Del</td>
<td>165.98</td>
<td>166.28</td>
<td>0.31</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4q34.1</td>
<td>Del</td>
<td>113.73</td>
<td>114.35</td>
<td>0.62</td>
<td>RHD</td>
<td>V+V</td>
<td>387</td>
</tr>
<tr>
<td>4q26</td>
<td>Del</td>
<td>119.73</td>
<td>120.06</td>
<td>0.33</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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.
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 extra-renal 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 neuropsychogenic 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 num-
ber of overlapping probes (Supplemental Table 7B). To avoid poten-
tial 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 pro-
cessed 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, per-
formed with GenomeStudio v.2011 (Illumina), PLINK, and PennCNV
software (26, 28). All CNVs reported in Tables 1 through 4 were exam-
ined 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 num-
ber 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 report-
able CNVs were adapted from prior recommendations for interpreta-
tion 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 popula-
tion; ref. 30) and (b) partial overlap with a known genomic disorder
or overlap with likely pathogenic CNVs reported in the International
nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000205.
v2.pl) 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 unre-
related individuals were considered for proportions and burden analy-
ses. CNV burden analysis was performed by counting large (size >100
kb), rare (frequency ≤0.1%) in any single cohort and in the whole con-
trol 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 larg-
est 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 genom-
ic 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 frequen-
cies 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-CNVS were com-
pared 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 consid-
ered 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 covarai-
tes by logistic regression (R v2.14). Nominal P values are reported in
Supplemental Table 4, but none of the associations with clinical vari-
ables 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 Pro-
gram (ECRIP). K. Kyriluk is supported by K23DK090207,
R03DK099564, an American Society of Nephrology Carl W. Gott-
slach Research Scholar Grant, and the New York State ECRIP.
M. Wuttke and A. Köttgen were funded by the German Research
Foundation (DFG KO 3958/2-1 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 investiga-
tor) at the Department of Pediatrics, University of Rochester Medi-
cal 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 avail-
able 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 Pedia-
tric 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.


