Genetic basis of human congenital anomalies of the kidney and urinary tract

Simone Sanna-Cherchi,[1] Rik Westland,[1,2] Gian Marco Ghiggeri,[3] and Ali G. Gharavi[1]

[1]Division of Nephrology, Department of Medicine, Columbia University College of Physicians and Surgeons, New York, New York, USA. [2]Department of Pediatric Nephrology, VU University Medical Center, Amsterdam, Netherlands. [3]Division of Nephrology, Dialysis and Transplantation, Istituto Giannina Gaslini, Genoa, Italy.

The clinical spectrum of congenital anomalies of the kidney and urinary tract (CAKUT) encompasses a common birth defect in humans that has significant impact on long-term patient survival. Overall, data indicate that approximately 20% of patients may have a genetic disorder that is usually not detected based on standard clinical evaluation, implicating many different mutational mechanisms and pathogenic pathways. In particular, 10% to 15% of CAKUT patients harbor an unsuspected genomic disorder that increases risk of neurocognitive impairment and whose early recognition can impact clinical care. The emergence of high-throughput genomic technologies is expected to provide insight into the common and rare genetic determinants of diseases and offer opportunities for early diagnosis with genetic testing.

Introduction

Congenital anomalies of the kidney and urinary tract (CAKUT) represent a spectrum of developmental malformations that include renal agenesis or hypodysplasia (RHD), multicystic dysplastic kidney (MCDK), ureteropelvic junction obstruction (UPJO), duplication of the pelvis, ureter, and/or kidney (DCS), congenital megaureter, ureterovesical junction obstruction (UVJO), vesicoureteral reflux (VUR), and posterior urethral valves (PUVs) (1–4).

CAKUT phenotypes originate from perturbations in kidney and urinary tract development, which is characterized by temporally and spatially coordinated interactions between the metanephric mesenchyme (MM, originating from the nephrogenic cord) and the ureteric bud (UB, originating from the Wolffian or nephric duct). For a detailed description of the molecular pathways of kidney development we refer to work by others (5–9). At embryonic day 10.5 in mice and at the end of the fifth gestational week in humans, the MM and UB send mutually inductive signals resulting in UB outgrowth and branching from the nephric duct into the rapidly differentiating MM (10, 11). The requirement for reciprocal induction (6, 7, 10–13) implies that signaling defects in either compartment can have pleiotropic effects across the entire urinary tract.

As such, malformations can affect single or multiple structures in a symmetric or asymmetric fashion, with significant variability between individuals carrying the same mutation (2, 3, 14–19). CAKUT may also occur in conjunction with other organ defects, indicative of known genetic syndromes. With this degree of variability, anatomical classification is often uninformative with regard to the primary molecular etiology. The advent of new genomic technologies now allows comprehensive examination of germline sequence variation across the genome and determination of molecular etiology and genetic architecture of disease (20). In this review we focus on insights from human genetic studies.

Epidemiology of kidney and urinary tract malformations

CAKUT is identified in more than 1% of overall live births, accounting for up to 23% of overall birth defects (21–24) and 40% to 50% of pediatric end-stage renal disease (ESRD) worldwide (25). Studies have shown that different structural defects have distinct impacts on long-term renal survival and overall mortality, with RHD conferring the greatest risk of adverse events (26–30). A number of extrinsic factors including maternal diabetes, medications, and folate and iron deficiency also increase the risk of CAKUT, highlighting environmental factors that modify expression of disease (31, 32). When ESRD is present at birth, mortality rates reach a striking 93% within the first year of life (33), and children who survive infancy have a 30-fold higher mortality compared with same-age children without ESRD (34). These population data underline the enormous impact of CAKUT on child health. Substantial improvement in early clinical care, such as prenatal detection of CAKUT by fetal ultrasonography and development of surgical and pharmacological approaches, has dramatically improved survival for infants and children with renal failure (35, 36). This, in turn, is resulting in an increased number of adult patients with CAKUT. Thus, the reported 2% to 7% prevalence of CAKUT among adults with ESRD may underestimate the growing impact of these traits among adult populations (25, 26).

Potential genetic models

Because the absence of kidney function is incompatible with postnatal survival without dialysis or transplantation, standard genetic theory predicts that dominantly acting mutations that completely impair kidney development would be strongly subjected to purifying selection and would therefore not reach a high frequency in the general population (37, 38). One can therefore hypothesize that the most severe malformations occurring in offspring of

Conflict of interest: A.G. Gharavi received support from the Renal Research Institute in 2016.
Reference information: *J Clin Invest*. 2018;128(1):4–15.
https://doi.org/10.1172/JCI95300.

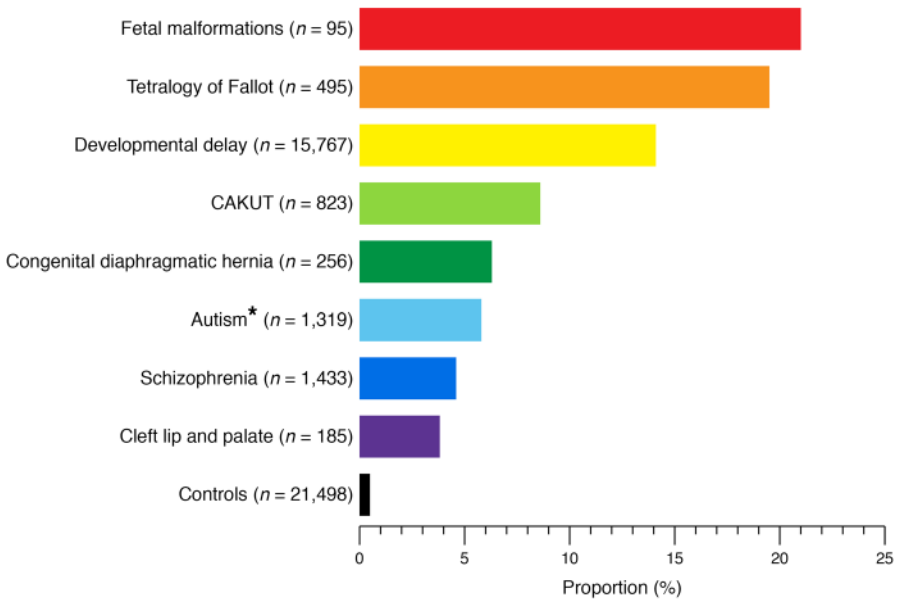

Figure 1. Proportion of patients with known genomic disorders in different human developmental phenotypes and healthy controls. There is a striking enrichment of known genomic disorders in human developmental disease compared with controls (59–62, 64–67, 69–71). The prevalence in controls is based on 21,498 controls generated from previously published studies (70, 71). *The proportion of known genomic disorders in autism spectrum disorder is displayed as the weighted average of two independent studies (64, 65). CAKUT, congenital anomalies of the kidney and urinary tract.

apparently healthy parents are attributable to de novo, dominantly acting mutations in developmentally important genes, or recessive mutations in genes that tolerate haploinsufficiency but not biallelic inactivation. In familial forms of CAKUT, which represent about 10% to 20% of cases, the disease frequently segregates as an autosomal dominant trait with incomplete penetrance (39, 40). In this setting, one can posit incomplete penetrance due to genetic or environmental modifiers, dominant inheritance of hypomorphic mutations impairing but not abrogating nephrogenesis, or disruption of genes/pathways that do not impair the initial steps of nephrogenesis but are important for urinary tract development and function at later stages of life. Other structural disorders such as VUR or DCS are more prevalent (41–44), presumably because they do not greatly impact overall survival; they may remain completely asymptomatic (DCS) or entirely resolve with age (VUR). Such prevalent disorders may also have polygenic determination, produced by the additive burden of common variants with modest effects that are thus less subject to purifying selection.

Recent insights from human genetic studies

With the advent of chromosomal microarrays and next-generation sequencing, over 40 genomic disorders and 50 genes have been implicated in syndromic or nonsyndromic forms of CAKUT (Supplemental Table 1; supplemental material available online with this article; https://doi.org/10.1172/JCI95300DS1). Many implicated genes belong to known developmental pathways, while the mechanism of disease for others has yet to be elucidated. Taken together, these studies implicate virtually every known mutational mechanism in disease pathogenesis and highlight genetic heterogeneity as a prominent biological feature of CAKUT. There is also remarkable diversity in the molecular pathways that have been recently discovered, suggesting that continued genetic analysis of CAKUT patients will uncover novel fundamental pathways in urinary tract development. Below, we will summarize recent examples of genetic discoveries enabled by the advent of newer genomic technologies.

Genomic disorders in sporadic forms of CAKUT. Prior studies demonstrated the association of CAKUT with large cytogenetic defects (45–47). Copy number variations (CNVs), generally defined as any gain or loss of germline DNA ranging from 1 kb to several Mb in size, account for the largest amount of sequence variation in the human genome (48–51). CNVs below 1 to 2 Mb are usually not detectable by conventional cytogenetic techniques but are readily detected by chromosomal microarray, which is now the preferred technique for detection of genomic imbalances for human malformations (52–55). CNVs can be generated via nonallelic homologous recombination, driven by highly repetitive elements flanking the CNVs (such as segmental duplications or low-copy repeats), resulting in recurrent breakpoints in affected individuals. Nonrecurrent CNVs have variable breakpoints in different individuals and are driven by other mechanisms such as nonhomologous end joining or fork stalling and template switching (56–58). Studies using chromosomal microarrays have now delineated over 200 recurrent pathogenic gene-disrupting CNVs, usually encompassing more than one gene, which confer risk for diverse human disorders such as neurodevelopmental syndromes, cardiac defects, craniofacial malformations, and congenital diaphragmatic hernia (59–68). Detection of these CNVs provides a precise molecular diagnosis that can stratify patients, explaining clinical variability between patients with the same clinical diagnosis but also demonstrating shared pathogenesis between some traditionally distinct clinical categories.

With the introduction of chromosomal microarrays, a number of studies have also uncovered an unexpectedly high contribution of genomic disorders to CAKUT (69–75). Remarkably, these studies did not reveal a discrete number of CAKUT-specific CNVs, but identified many known or novel genomic lesions, indicating significant genetic heterogeneity. These CNVs were diagnostic of many well-known human malformation syndromes that were unrecognized based on the clinical workup. For example, a study of 522 children with RHD (with or without syndromic features), recruited from multiple renal and urology clinics in Europe and the USA identified 34 different genomic disorders in 55 individuals (10.5% of cases) (69). The most frequently identified genomic disorder was Chr.17q12 deletion, diagnostic of the renal cyst and

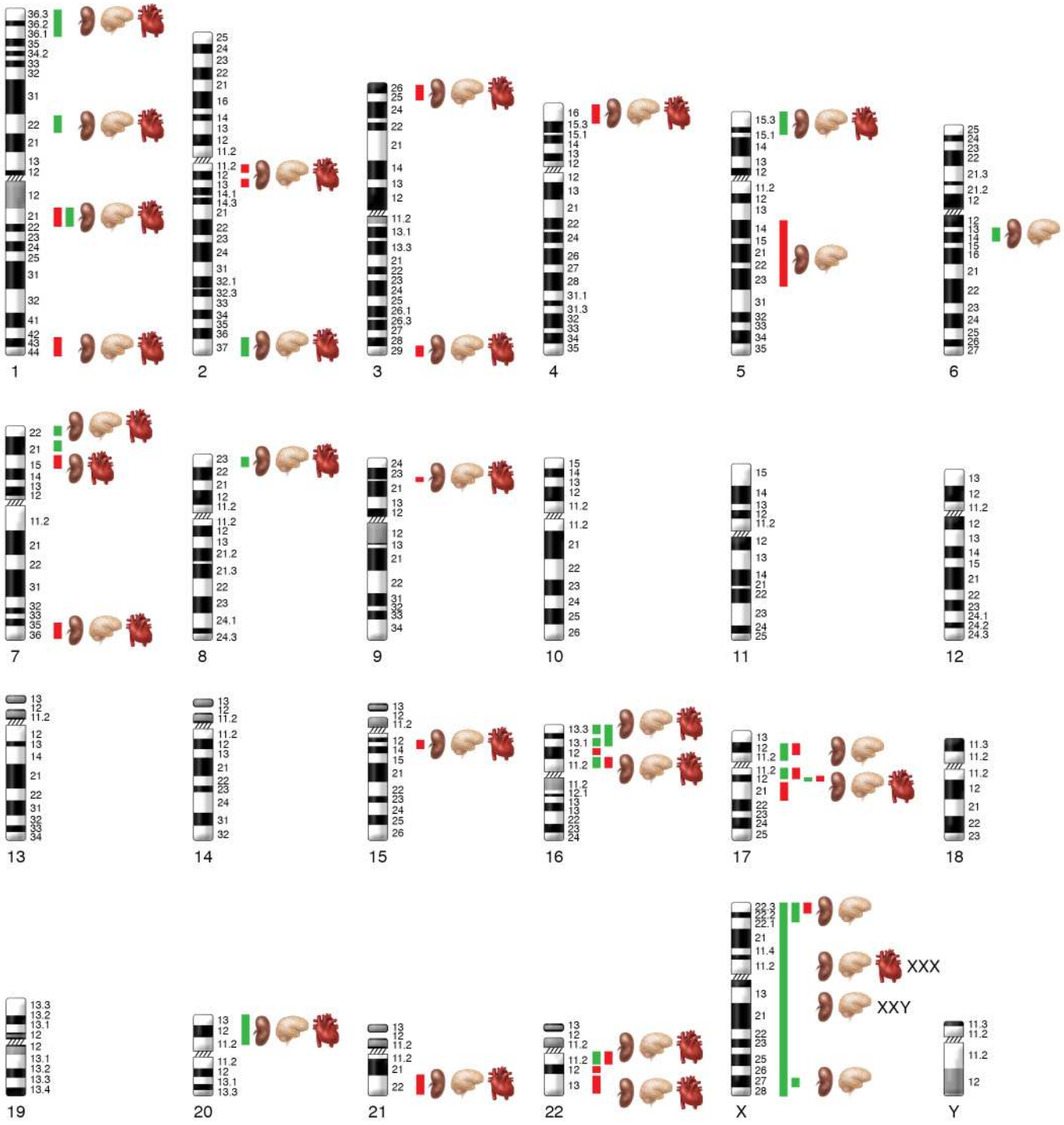

Figure 2. Overview of identified genomic disorders in isolated CAKUT presented on a human chromosomal map. The overlap of known CAKUT genomic disorder loci (kidney symbol) (69–71) with developmental delay (brain symbol) and congenital heart defects (heart symbol) are indicated based on a review of the literature. Red = deletion; green = duplication.

diabetes syndrome (RCAD) (76), followed by Chr.22q11.2 deletion, diagnostic of the DiGeorge/velocardiofacial syndrome (77), and Chr.1q21 deletion (78). The remaining known genetic syndromes were present in few or single individuals, highlighting the extreme genetic heterogeneity of disease. Another 6% of cases harbored rare gene-disrupting CNVs that were likely pathogenic based on gene content, size, and low frequency in controls. In a subsequent study, 13 of 80 (14%) children with a solitary functioning kidney harbored a known or likely pathogenic CNV (71). In another follow-up study of a large prospective observational cohort of children with all-cause CKD (79), 4% of patients harbored genomic imbalances (70), compared with less than 0.5% of 21,575 controls. Among different clinical categories, children with RHD were especially enriched for diagnostic CNVs (odds ratio

[OR] 30.1; $P = 4.9 \times 10^{-16}$ vs. controls), while other CAKUT subcategories, such as VUR and UPJO, had a lower overall prevalence (OR 4.5; $P = 0.03$ vs. controls).

Figure 1 summarizes findings from three large studies totaling 824 CAKUT patients (69–71). Overall, there is a striking 15-fold enrichment for rare genomic disorders as compared with controls, significantly driven by patients with RHD. The frequency of diagnostic CNVs is comparable between CAKUT and several other human developmental disorders, such as developmental delay, neurocognitive disorders, and congenital heart disease, for which microarray analysis is now routinely recommended as first-line diagnostics (59–67). All 38 identified known genomic disorders in CAKUT have been previously associated with developmental delay and/or cardiac malformations (Figure 2). Remarkably, four of these loci account for 44% of the CAKUT cases with a known genomic disorder (Chr.1q21.1, Chr.16p11.2, Chr.17q12, and Chr.22q11.2 deletions/duplications, accounting for 3.5% of total cases vs. 0.1% of total controls, OR 25.3; $P = 6.5 \times 10^{-26}$; Table 1). Comparison of the frequencies of these CNV disorders in CAKUT, developmental delay, and tetralogy of Fallot (59, 80) indicates noteworthy similarities and differences (Figure 3 and Table 1). Deletions are the predominant type of diagnostic CNVs for CAKUT and developmental delay phenotypes, with the Chr.17q12 (RCAD) locus as the most frequently involved site. Patients with cardiac malformations are exquisitely enriched for CNVs at Chr.22q11.2 and carry mostly duplications at Chr.1q21, Chr.16p11.2, and Chr.17q12 loci (Figure 3, A–D). In CAKUT cases where inheritance could be tested, only half of the pathogenic CNVs occurred as de novo events, and the remainders were inherited from apparently healthy or mildly affected parents (69). The reasons for variable phenotypic expression of pathogenic CNVs are incompletely understood, but may in part be attributable to the genetic load imparted and background genetic effects, discussed below under missing heritability.

The next step in understanding the architecture of CAKUT is to identify the genetic drivers that account for the renal malformation phenotypes of these genomic disorders. Although proven challenging (81, 82), knowledge of these drivers increases our understanding of the biological pathways of kidney development and, consequently, can lead to development of precision medicine approaches in these patients.

We highlight the Chr.17q12 deletion (RCAD) syndrome, owing to its relatively high prevalence in CAKUT patients and its association with extrarenal complications (76). The syndrome is also known as maturity-onset diabetes of the young type 5 (MODY5) or hypoplastic glomerulocystic disease. Most of the manifestations of this syndrome are attributable to the disruption of *HNF1B* (83–86), encoding a transcription factor that orchestrates pancreatic, parathyroid, kidney, and urogenital development (87–89). The urinary tract phenotype is highly variable and can present as sparse prenatal renal cysts, small echogenic kidneys, MCDK, and even phenocopy polycystic kidney disease (PKD). Additional tubular manifestations such as hypomagnesemia, hyperuricemia, salt-wasting, and urinary concentrating defects are often present (90). Although most of the renal manifestations are detected in early childhood, this disorder may go unrecognized because diabetes occurs at a mean age of approximately 25 years (86, 91). Other manifestations include defects of the female internal genitalia and hypoparathyroidism; heart and ocular anomalies have occasionally been described (76, 92–95). The Chr.17q12 deletion is also associated with a nearly 14-fold increase in risk of developing autism or schizophrenia later in life (96). This complication may be attributable to disruption of other genes within the interval because neurocognition appears to be preserved in RCAD patients due to intragenic mutations in *HNF1B* (97). Individuals carrying the reciprocal duplication on Chr.17q12 typically present with various degrees of neurodevelopmental delay and behavioral problems, esophageal atresia, cardiac defects, and CAKUT (in particular RHD, cysts, and obstructive uropathy phenotypes) (76, 92).

For other genomic disorders, in-depth bioinformatics analysis can help prioritize genetic drivers for novel CNV phenotypes (71). Recently, we used a more extensive approach to dissect the Chr.22q11.2 locus, which manifests with renal phenotypes in about 20% to 30% of patients (77, 98). Although T-box 1 (*TBX1*) has been implicated in the pathogenesis of some hallmark clinical features of the syndrome (especially the conotruncal cardiac defects), the genetic drivers of the CAKUT phenotype were largely unknown. Using a multidisciplinary approach based on large-scale CNV analysis, whole-exome sequencing (WES), and targeted next-generation sequencing, followed by functional modeling in zebrafish and mouse, we prioritized three genes as likely involved in the pathogenesis of kidney defects in patients with DiGeorge syndrome caused by Chr.22q11.2 microdeletions. These genes, *snap29*, *aifm3*, and *crkl*, recapitulated the renal phenotype in zebrafish. Knockdown or CRISPR/Cas9–mediated inactivation of *crkl* was sufficient to induce pronephric convolution defects. However, *snap29* and *aifm3* appeared to operate in an epistatic fashion, suggesting that deletions at Chr.22q11.2 result in complex genetic determination of CAKUT. We subsequently found protein-altering mutations in *CRKL* in about 1% of CAKUT patients from a large replication cohort. Finally, targeted deletion of *Crkl* exon 2 in mice led to various CAKUT phenotypes (99), a finding that was recently reproduced by other investigators (100). Overall, these data intimate a model where *CRKL* haploinsufficiency is the main determinant of CAKUT in patients with the Chr.22q11.2 microdeletions (99). The incomplete penetrance and variable phenotypic expression of both microdeletions and *CRKL* point mutations in humans and *Crkl* inactivation in mouse suggest a complex model with both local (e.g., *SNAP29*, *AIFM3*) and distantly acting genetic effects, with involvement of environmental and stochastic factors.

Another recent example of genetic driver discovery is *PBX1*, encoding a transcription factor known to regulate ureteric branching in the murine urinary tract (101), which is responsible for the CAKUT phenotype in the 1q23.3-q24.1 deletion syndrome. Using a targeted sequencing approach of candidate genes for CAKUT, Heidet et al. identified de novo loss-of-function mutations (three point mutations and two heterozygous deletions validated by microarray) in *PBX1* in five of 204 RHD cases (2.5%; combined *P* value for de novo occurrence < 0.001) (102). Extrarenal manifestations identified in *PBX1* mutation carriers were deafness, scoliosis, as well as developmental delay.

However, contrary to the above-mentioned scenarios in which dosage imbalances of a single driver may be sufficient to cause CAKUT, we have incomplete understanding of more com-

Table 1. Common genomic disorders in CAKUT, developmental delay, and cardiac malformations

CNV locus	Prevalence in CAKUT cases (%) (*n* = 823)	Renal phenotype	Prevalence in developmental delay cases (%) (*n* = 15,767)	Neurodevelopmental phenotype	Prevalence in tetralogy of Fallot cases (%) (*n* = 495)	Cardiac phenotype	Other associated phenotypic features	Prevalence in in-house controls (%) (*n* = 21,498)
1q21.1	1 in 163 (0.61%)	Renal hypodysplasia (–), vesicoureteral reflux (#)	1 in 217 (0.46%)	Mild to moderate developmental delay (–/#), autism (–/#), schizophrenia (–/#), microcephaly (–), macrocephaly (#), seizures (#), corpus callosum and cerebellar vermis hypoplasia (#)	1 in 123 (0.81%)	Tetralogy of Fallot (#), bicuspid aortic valve with aneurysm, truncus arteriosus (–), transposition of the great vessels (–), patent ductus arteriosus (–), coarctation of the aorta (–)	Facial dysmorphisms (–/#), hand and foot malformations (–), failure to thrive (–)	1 in 2,000 (0.05%)
16p11.2	1 in 163 (0.61%)	Renal hypodysplasia (–/#), obstructive uropathy (–)	1 in 169 (0.59%)	Autism (–/#), seizures (–), macrocephaly (–), microcephaly (#), developmental delay (–/#), behavioral problems/ADHD (–/#), schizophrenia (#/–), tonsillar ectopia (–), Arnold–Chiari malformation (–), gross motor delay (–)	1 in 500 (0.20%)	Tetralogy of Fallot (#), atrial septal defect (–), pulmonary artery stenosis (–)	Vertebral anomalies with congenital scoliosis (–), facial dysmorphisms (–/#), hand and foot malformations (–), obesity (–)	1 in 2,000 (0.05%)
17q12	1 in 58 (1.70%)	Renal hypodysplasia incl. renal cysts (–/#), obstructive uropathy (–)	1 in 500 (0.20%)	Mild to moderate developmental delay (–/#), autism (del), seizures (–/#), corpus callosum hypoplasia (#)	1 in 500 (0.20%)	Tetralogy of Fallot (#), atrial septal defect (#)	Maturity onset of diabetes in the young type 5 (–), facial dysmorphisms (–/#), hypomagnesemia (–), hyperuricemia (–), vertebral anomalies with congenital scoliosis (–), hand and foot malformations (#/–), esophageal atresia (#)	1 in 5,000 (0.02%)
22q11.2	1 in 163 (0.61%)	Renal hypodysplasia (–/#)	1 in 108 (0.93%)	Mild to moderate developmental delay (–/#), autism (–), seizures (–/#), corpus callosum hypoplasia (#), schizophrenia (–), bipolar disorder (–), behavioral problems/ADHD (–), microcephaly (#)	1 in 9 (10.90%)	Tetralogy of Fallot (–/#), truncus arteriosus (–), interrupted aortic arch (–), right aortic arch (–), ventricular septal defect (–), patent ductus arteriosus (–)	Obesity (–), facial dysmorphisms (–/#), vertebral anomalies with congenital scoliosis (–), parathyroid hypoplasia (–), hypocalcemia (–)	1 in 3,000 (0.03%)
Total	1 in 28 (3.52%)	–	1 in 46 (2.18%)	–	1 in 8 (12.12%)	–	–	1 in 714 (0.14%)

The prevalence of known genomic disorders in CAKUT is based on three independent studies (69–71). Phenotypes of CAKUT included in these studies were renal hypodysplasia, obstructive uropathy, and vesicoureteral reflux. The prevalence of known genomic disorders in developmental delay is based on one study (59). Phenotypes in developmental delay included in this study were intellectual disability, developmental delay and/or autism, congenital malformations, hypotonia and feeding difficulties, speech and motor deficits, epilepsy, hearing impairment, craniofacial and skeletal features, and behavioral issues. The prevalence of known genomic disorders in congenital heart defects is based on one study (80). Phenotypes included in this study were tetralogy of Fallot with/without pulmonary atresia. The prevalence of known genomic disorders in controls was generated from in-house available genotyping data (70). The phenotypic features of known genomic disorders were extracted from the aforementioned studies as well as the Online Mendelian Inheritance in Man (OMIM) database (https://www.omim.org). ADHD, attention deficit hyperactivity disorder; CAKUT, congenital anomalies of the kidney and urinary tract; CNV, copy number variation; –, deletion; #, duplication.

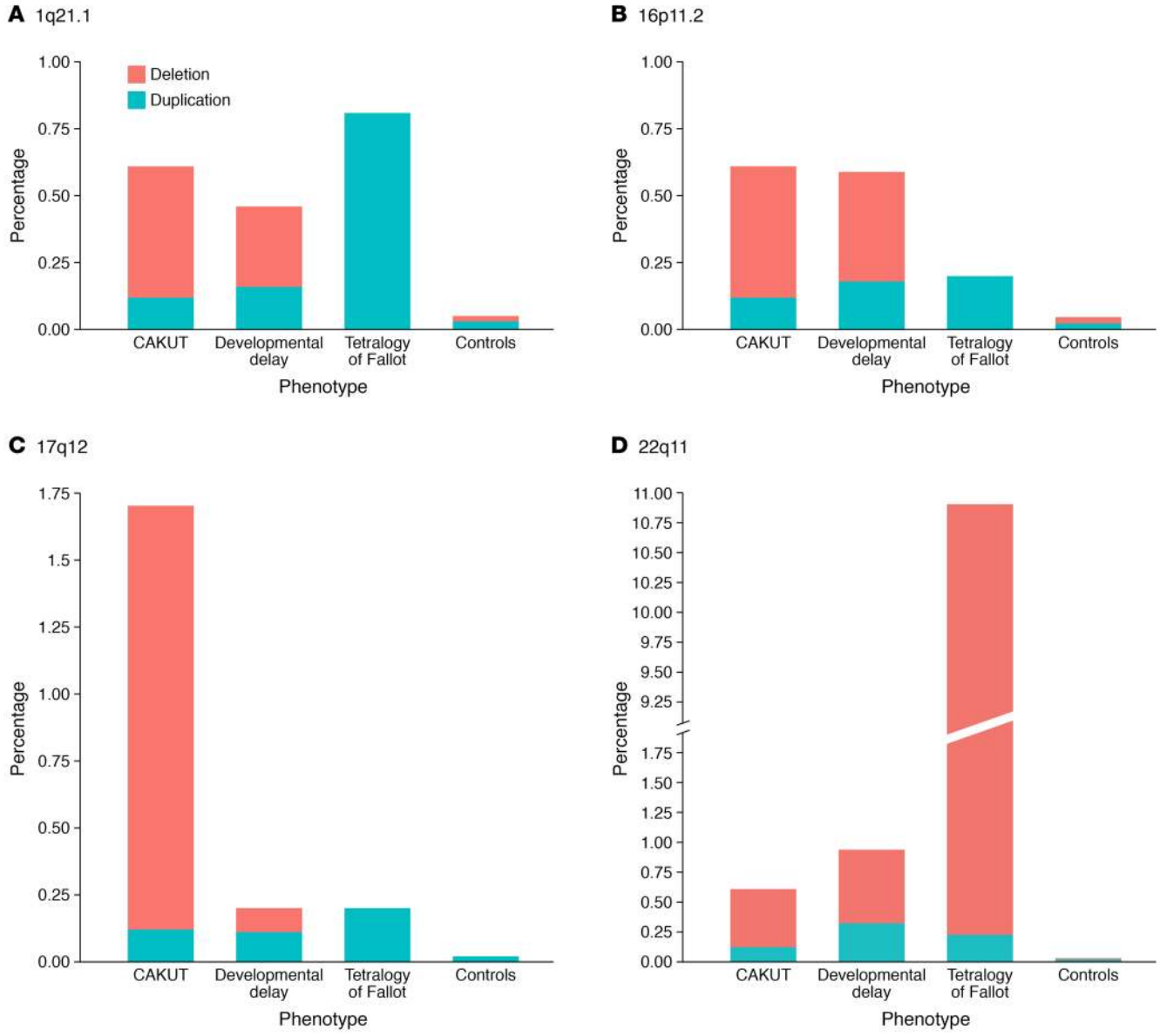

Figure 3. Differences and similarities in the prevalence of the four most commonly implicated CNV loci in CAKUT patients. Bar graphs compare the prevalence of common loci in CAKUT patients (n = 823) to the prevalence of identical loci in patients with developmental delay (n = 15,767), tetralogy of Fallot (n = 495), and in-house genotyping data of healthy controls (n = 21,498) (59, 69–71, 80). Genomic imbalances are enriched for all phenotypes compared with controls. Duplications are shown in green, deletions are shown in red. (**A**) Prevalence of duplications and deletions at chromosomal locus 1q21.1. (**B**) Prevalence of duplications and deletions at chromosomal locus 16p11.2. (**C**) Prevalence of duplications and deletions at chromosomal locus 17q12. As expected, CAKUT patients show a significant enrichment for the renal cysts and diabetes (RCAD) syndrome deletion. (**D**) Prevalence of duplications and deletions at chromosomal locus 22q11.2. Patients with tetralogy of Fallot are typically enriched for the 22q11.2 microdeletion syndrome.

plex scenarios involving many drivers that jointly influence the prevalence and severity of disease (81). These scenarios will be an important subject of future research to define further the genetic architecture of CAKUT.

Autosomal dominant single-gene defects. Initial evidence for an important role of rare point mutations in the pathogenesis of CAKUT is mainly derived from analysis of recognizable syndromic disorders that are attributable to dominant mutations in *PAX2*, *HNF1B*, *SALL1*, *WT1*, *SIX1*, *EYA1*, and others (4, 20, 103–105). These diseases were recognized as specific syndromes based on cooccurrence of specific extrarenal manifestations, such as ocular coloboma, preauricular tags or fistulas, and anorectal malformations, but subsequent to identification of causal genes in syndromic patients, studies demonstrated that mutations in these same genes are also responsible for isolated (nonsyndromic) CAKUT (18). Mutations in *HNF1B* and *PAX2* are relatively frequent (up to 15% of patients with RHD), although the overall fraction of disease that is attributable to these two genes generally does not exceed 10% of patients (18, 85, 90, 106–109). For example, in a recent study the prevalence of diagnostic mutations in *PAX2* and *HNF1B* was much lower than expected (109). The observation that mutations in *HNF1B* are frequently identified in RHD, while they are rare in

isolated lower urinary tract defects such as PUV, may have contributed to this finding (90, 106). The discovery of mutations in genes associated with specific syndromes impacts clinical care because it prompts physicians to reassess patients for undetected extrarenal phenotypes and institute surveillance programs for complications that may develop much later (e.g., diabetes in RCAD syndrome).

The advent of next-generation sequencing has accelerated the pace of gene discovery. We used WES combined with linkage analysis in a family with dominant inheritance and incomplete penetrance of CAKUT phenotypes, predominantly manifested by congenital obstructive uropathy, and identified a mutation in *DSTYK*, encoding an uncharacterized dual-specificity serine/threonine and tyrosine kinase (110). Resequencing of 311 additional CAKUT patients identified five new mutations in seven patients (2.3%), including a loss-of-function variant in the first exon of *DSTYK*. A number of mutation carriers also manifested neurologic phenotypes such as ataxia and epilepsy, suggesting that *DSTYK* mutation can affect neurologic development (110). It is thus interesting that a recent study reported homozygosity for a rare intragenic deletion that encompasses the last two exons and the 3′ UTR of *DSTYK* in three unrelated families of Middle-Eastern ancestry with autosomal recessive spastic paraparesis (111). The proband phenotype in one family also included a horseshoe kidney, suggesting that this particular *DSTYK* mutation can also affect kidney and urinary tract development.

Bekheirnia et al. recently reported WES results in 62 families with CAKUT (112). In addition to detecting pathogenic mutations in *HNF1B*, *PAX2*, and *EYA1* in 5% of affected families, they discovered a deleterious de novo variant in *FOXP1* in a case with RHD, hydrocephalus, and developmental delay. The authors subsequently interrogated their in-house database of individuals subjected to exome sequencing and additionally identified seven cases (out of over 5,000) with de novo mutations in *FOXP1*. These individuals were all characterized by brain anomalies, developmental delay, genital anomalies, and CAKUT. *FOXP1* encodes a forkhead box transcription factor that was previously implicated in mental retardation with language impairment and autistic features, but renal complications had not been previously described (113). This study thus demonstrates the potential of large-scale WES studies for gene discovery and expanding the phenotypic manifestation of syndromic disease. Recent WES studies have also discovered many other genes as monogenic causes for CAKUT, such as *TBX18* and *NRIP1* (114, 115), respectively implicating defects in ureteric mesenchymal cell development and retinoic acid signaling in disease pathogenesis.

Finally, with the availability of large exome control data sets, exome-wide association studies are feasible. This approach, combined with functional in vivo modeling in zebrafish, recently led to the discovery of loss-of-function mutations in *GREB1L* as a new cause of autosomal dominant RHD (116). *GREB1L* (growth regulation by estrogen in breast cancer 1 like) encodes a protein with a poorly understood function, but its role in the pathogenesis CAKUT has recently been validated by others (117). These findings further highlight the diversity of signaling defects that can lead to CAKUT and associated phenotypes.

Autosomal recessive single-gene defects. Many recessive forms of CAKUT have been reported. Although rare, autosomal recessive forms of CAKUT usually involve loss of function, thereby providing immediate insight into pathogenic mechanism (118–120). For example, recessive mutations in genes encoding components of the renin-angiotensin system (RAS) were detected in individuals with a severe form of renal tubular dysgenesis associated with oligohydramnios and perinatal mortality secondary to lung hypoplasia (120). The kidney developmental anomalies precisely phenocopied the effect of angiotensin-converting enzyme inhibitors in fetuses exposed to these drugs, confirming the central role of the RAS in normal kidney development. More recently, targeted next-generation sequencing of 12 recessive murine candidate genes in 590 patients suggested that biallelic missense mutations in *FRAS1*, *FREM1*, *FREM2*, and *GRIP1* may cause isolated CAKUT (121). Interestingly, recessive loss-of-function mutations in these genes cause Fraser syndrome, characterized by genital anomalies, cryptophthalmos, and CAKUT (122, 123). Based on their findings, the authors hypothesized that missense mutations are more likely to represent hypomorphic alleles with milder manifestation such as isolated CAKUT (119, 121). A study on a large cohort of CAKUT and VACTERL (vertebral anomalies, anal atresia, cardiac defects, tracheoesophageal fistula and/or esophageal atresia, renal/radial anomalies, and limb defects) syndrome identified homozygous or compound heterozygous mutations in *TRAP1* in about 0.5% of patients (124). The exact role of TRAP1 in mammalian kidney and urinary tract development is currently unknown.

Missing heritability and gene discovery challenges in CAKUT. Known genetic disorders caused by rare CNVs or point mutations collectively explain at most 20% to 25% of CAKUT cases (particularly in RHD), leaving open the possibility that the remaining patients will be explained by rare monogenic disorders that each explain less than 1% of cases. Nonetheless, a simple single-gene model would not explain the incomplete penetrance and variable expression frequently encountered among individuals harboring the same mutation. It is clear that all mutations are inherited on individual genetic backgrounds that can modify phenotypic effects. For example, recent work on developmental delay and CAKUT shows that about 10% of patients with a known pathogenic structural variant also harbor a second large CNV (69, 125). Consistent with this higher mutational burden, the patients with second-site CNVs had more severe phenotypes and multiple-organ involvement. In a large follow-up study of genomic disorders in patients with developmental delay, some CNVs almost always occurred as de novo events, indicating that they are subjected to negative selection; these manifested with well-defined phenotypic features (e.g., the Smith–Magenis syndrome, caused by deletions at Chr.17p12) (125). On the other hand, some CNVs were more frequently inherited and more likely to be accompanied by a second-site CNV, indicating that they are more tolerated; these disorders had a more variable phenotypic expression, suggesting that their phenotypic expression is modified by background genetic effects (e.g., the Chr.16p11.2 deletion/duplication and the Chr.16p13.11 duplication) (59, 125, 126). Another example of oligogenic effects involves Bardet-Biedl syndrome (BBS, OMIM#209900), a genetically heterogeneous autosomal recessive trait characterized by renal cystic malformations, obesity, polydactyly, retinal degeneration, and other developmental defects (127). Biallelic mutations in 19 distinct genes can cause different BBS subtypes and the phenotypic variability (128, 129) is in

part attributable to the presence of modifier alleles at other BBS loci (129–132). Another example relevant for CAKUT is the case of PKD, in which early-onset, severe phenotypes due to mutations in autosomal dominant PKD genes were attributable to additional mutations in other PKD genes or *HNF1B* (133).

In addition to oligogenic inheritance, GWAS have suggested a polygenic model as the basis of many complex developmental traits, such as cleft palate, bladder exstrophy, and congenital heart disease (134–136). Examples relevant to CAKUT include two GWAS for hypospadias, which revealed a major signal in *DGKK* and multiple additional loci conferring lower risk (137, 138). The top 18 SNPs fell into well-characterized developmental pathways (e.g., *HOXA4*, *IRX5*, *IRX6*, and *EYA1*) and remarkably explained up to 9% of the disease variance, indicating a significant additive contribution of common variants with modest individual effects to the genetic architecture of hypospadias. The polygenic contribution of common variants may be better appreciated with recent application of more complex inheritance models that incorporate multiple signals and estimates of local linkage disequilibrium, explaining a higher proportion of the disease heritability for many disorders (139–143). Complex inheritance models that incorporate the burden of common and rare variants in several neuropsychiatric disorders have indicated that the genetic burden of mutations defines a continuum of risk that can result in a clinical diagnosis at the extremes of the distribution (144–148). As recently proposed (149), one can therefore hypothesize that a polygenic model, in which multiple common causal loci contribute to the variation in the prevalence and severity of CAKUT phenotypes, can potentially explain why the same mutation may manifest very early in some individuals but remain clinically silent in others.

With the increasing feasibility and cost-effectiveness of genome-wide sequencing approaches, a major challenge in studies on isolated sporadic CAKUT is to distinguish disease-causing variants from the large number of variants present within each human genome (150). A typical individual's genome consists of 4.1 to 5 million potential variant sites (151) with 40,000 to 200,000 variants at a frequency below 0.5% in the general population. In addition, 4% to 8% of healthy population controls contain a large rare CNV (>1 Mb), presumably without a known role in human development or health (50, 59, 69). Furthermore, human variation differs greatly between different populations, with African ancestry populations showing the greatest variability and European populations the lowest. Overall, these data clearly underline the significance of population-matched controls for gene discovery studies. Publicly available databases such as ISCA, DECIPHER, Clinvar, Exome Variant Server, 1000 Genomes Project, and ExAC/gNOMAD are extremely valuable for differentiating potential disease-causing variants from standing variation (55, 151–153). The use of reliable assays in cells or organoids, or tractable animal models such as zebrafish will also be invaluable for functional analysis to guide interpretation of these potential variants before one commits to engineering the variants in rodents.

One can also begin to estimate the potential for undetected genetic forms of CAKUT by examining the frequency of putative loss-of-function variants in CAKUT genes in large public WES or genome sequence databases. For example, ExAC, a database of over 60,000 individuals undergoing exome sequencing for non–kidney-related conditions (152), reports loss-of-function mutations in *PAX2* in approximately 1:4,000 individuals, and a *DSTYK* splice site variant in approximately 1:3,000 individuals. Overall, about 1:200 individuals in ExAC harbor a rare loss-of-function variant (allele frequency < 1:1,000) in dominantly inherited genes listed in Supplemental Table 1, suggesting that some patients may carry clinically unrecognized kidney traits, or some variants may be incompletely penetrant or completely benign. Hence, phenotypic characterization of mutation carriers from genetically characterized cohorts can uncover the prevalence of genetic urinary tract anomalies and better define the penetrance and clinical relevance of dominantly inherited variants. At the same time, these findings highlight the challenges of gene discovery and sequence interpretation for CAKUT.

Clinical care implications of genetic testing in CAKUT

With the introduction of clinical grade chromosomal microarray and next-generation sequencing, clinicians now have the opportunity to incorporate genetics into their diagnostic workup. Genetic testing can provide a precise diagnosis that can help individualize clinical care by screening for specific complications (e.g., screen for diabetes or ocular coloboma in patients with *HNF1B* and *PAX2* mutations, respectively; ref. 18) and facilitate medical decision making (e.g., avoiding immunosuppressive therapy in CAKUT clinically misdiagnosed as a glomerular nephropathy because of low-grade proteinuria). The availability of genetic diagnosis can also put an end to diagnostic odysseys for families and provide tailor-made counseling.

Based on the available data and national recommendations for genetic testing for congenital anomalies (54), chromosomal microarrays should be strongly considered as a first-line diagnostic approach for CAKUT cases (55). In particular, data indicate an up to 15% diagnostic rate for cases involving parenchymal kidney defects (e.g., RHD), or cases involving extrarenal malformations (69–71). The identification of genomic disorders enables targeted surveillance and intervention for specific complications, exemplified for the Chr.17q12 syndrome above.

The highly significant overlap between pathogenic CNVs in CAKUT and neuropsychiatric disorders particularly impacts the clinical management of CAKUT patients (Figures 1–3), because a significant fraction of individuals with these genomic imbalances are at increased risk for intellectual disability and other neuropsychiatric complications such as autism and schizophrenia (59, 69–71, 125, 154). This genomic overlap suggests that the same genetic lesion can simultaneously impair nephrogenesis and neurodevelopment, potentially explaining the known association between neurocognitive deficits and CKD in children (155). Consistent with this hypothesis, analysis of the Chronic Kidney Disease in Children (CKiD) study cohort demonstrated that children with pathogenic CNVs had significantly lower neurocognitive scores such as IQ, depression/anxiety, and executive functioning, confirming that genomic diagnosis can provide opportunities for early diagnosis of neurocognitive impairment in children with CKD (156). Moreover, the severity of neurocognitive impairment is attenuated by factors such as higher parental IQ or level of education, indicating modifiability by background hereditary and environmental factors (70,

156–158). These data must be evaluated in the light that CAKUT is usually detected during routine fetal ultrasound investigation, a much earlier time point than neurocognitive defects are diagnosed. Given the plasticity of neurocognitive development in the postnatal period, an early molecular diagnosis can prompt careful clinical evaluation and early detection of subtle neurocognitive impairment and may lead to improved outcome through early and intensive behavioral interventions (159).

When microarray studies are negative, clinicians have the option to follow up with next-generation sequencing tests, such as targeted sequencing panels and whole-exome or whole-genome analysis. The pros and cons of each modality are discussed elsewhere (160–162). In the short term, targeted panels may provide a cost-efficient approach for screening for the most common (but not all) disorders, and negative results may require additional follow-up with another modality. A number of studies have demonstrated that targeted panels will yield a diagnostic rate of 5% to 10% in patients with isolated CAKUT (109, 163). With the rapidly growing list of syndromes and declining cost of sequencing, WES or genome sequencing will likely emerge as the preferred diagnostic modality. The major benefit of these modalities is their genome-wide coverage, which allows analysis of all relevant genes simultaneously and also provides the opportunity for reanalysis as novel genes are identified. They may also reveal incidental genetic findings that are unrelated to the primary indication for testing, providing added health information that may be beneficial (e.g., predisposition to cancer), although this may unintentionally augment the medical decision–making burden on families.

Regardless of sequencing modality, clinical annotation of genomes and classification of variants will be the most challenging tasks for clinicians. As previously mentioned, we have guidelines for interpretation of genetic sequence data and rich population reference data sets and many tools for in-silico prediction of deleteriousness of sequence variants (150–153, 164, 165). Still, molecular pathologists often face many difficulties in differentiating rare variants with predicted deleterious potential from true disease-causing variants. Some clinical databases such ClinVar also contain outdated information about variant pathogenicity, necessitating careful review of original evidence before a clinical report is issued (150, 152, 166). Furthermore, variant interpretation is also complicated for minority populations for which background allele frequency data are scarce, raising the potential for misdiagnosis and generating many variants of unknown significance in clinical reports (167–169). With the availability of sequencing data in larger populations of cases and controls, many variants of unknown significance can be better classified in definite categories such as pathogenic or benign. Conversely, many variants initially classified as pathogenic may be reclassified as benign, posing new dilemmas for clinicians and patients who may have acted on the initial report. Interpretation of alleles that impart large effects on disease risk but are characterized by significant incomplete penetrance will also pose challenges in genetic counseling. Systematic efforts such as ClinGen will be invaluable for curation and annotation of sequence data, but will likely take many years to complete (165). In the meantime, clinicians must balance the need for stringent interpretation of data and the desire to provide answers for their patients. In addition, a large number of ethico-legal and societal issues will need to be considered in the return of primary results and incidental findings. In particular, the ability to obtain genetic diagnoses in the prenatal setting can lead to personalized and preventive treatment strategies, but can also influence parental planning, raising thorny issues that will require thoughtful consideration.

Conclusions

CAKUT is a frequent human developmental defect that imparts significant risk for renal and extrarenal morbidity. Current genetic studies indicate a complex genetic basis for this trait requiring efforts to devise large-scale human genetic studies accompanied by appropriate functional modeling to solve its genetic underpinning. Understanding the genetic architecture of CAKUT, its subcategories, and its complications will be instrumental for developing accurate genetic testing strategies that can guide clinical decision making within a precision medicine framework. A genetic diagnosis prompts physicians to reassess their patient's (renal and extrarenal) phenotype, and is critical in the prevention of future complications of disease. Despite the fact that challenges remain in the interpretation of genetic data, such a precision medicine approach will lead to better treatment and a better clinical outcome for CAKUT patients and their family members.

Acknowledgments

We want to apologize for all the important work of investigators that we were not able to cite in this review due to space restraints. We thank the three anonymous reviewers who commented on this work. This work was supported by the American Heart Association Grant-in-Aid 13GRNT14680075 (to SSC), by the Columbia CTSA Irving Institute/Clinical Trials Office pilot grant UL1 TR000040 (to SSC), by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) 1R21DK098531 (to SSC), the NIDDK 1R01DK103184 (to SSC), by the Italian Ministry of Health "Ricerca Finalizzata" (to SSC and GMG), by a Kolff postdoc abroad grant of the Dutch Kidney Foundation (to RW), by the Fondazione Malattie Renali nel Bambino (to GMG), by the NIDDK 1R01DK080099 (to AGG), and the NIDDK 3U54DK104309 (to AGG).

Address correspondence to: Simone Sanna-Cherchi, Division of Nephrology, Columbia University College of Physicians and Surgeons, 1150 Street Nicholas Avenue, Russ Berrie Pavilion 412b, New York, New York 10032, USA. Phone: 212.851.4925; Email: ss2517@cumc.columbia.edu. Or to: Ali G. Gharavi, Division of Nephrology, Columbia University College of Physicians and Surgeons, 1150 Street Nicholas Avenue, Russ Berrie Pavilion 412, New York, New York 10032, USA. Phone: 212.851.4927; Email: ag2239@columbia.edu.

1. Goodyear PR. Renal dysplasia/hypoplasia. In: Avner ED, Harmon WE, Niaudet P, eds. *Pediatric Nephrology*. 5th ed. Philadelphia, Pennsylvania, USA: Lippincott Williams and Wilkins; 2004:83–91.
2. Ichikawa I, Kuwayama F, Pope JC, Stephens FD, Miyazaki Y. Paradigm shift from classic anatomic theories to contemporary cell biological views of CAKUT. *Kidney Int*. 2002;61(3):889–898.
3. Nakanishi K, Yoshikawa N. Genetic disorders of human congenital anomalies of the kidney and urinary tract (CAKUT). *Pediatr Int*.

2003;45(5):610–616.
4. Sanna-Cherchi S, et al. Genetic approaches to human renal agenesis/hypoplasia and dysplasia. *Pediatr Nephrol*. 2007;22(10):1675–1684.
5. Chen F. Genetic and developmental basis for urinary tract obstruction. *Pediatr Nephrol*. 2009;24(9):1621–1632.
6. Dressler GR. The cellular basis of kidney development. *Annu Rev Cell Dev Biol*. 2006;22:509–529.
7. Schedl A. Renal abnormalities and their developmental origin. *Nat Rev Genet*. 2007;8(10):791–802.
8. Woolf AS, Price KL, Scambler PJ, Winyard PJ. Evolving concepts in human renal dysplasia. *J Am Soc Nephrol*. 2004;15(4):998–1007.
9. Little MH, McMahon AP. Mammalian kidney development: principles, progress, and projections. *Cold Spring Harb Perspect Biol*. 2012;4(5):a008300.
10. Costantini F, Kopan R. Patterning a complex organ: branching morphogenesis and nephron segmentation in kidney development. *Dev Cell*. 2010;18(5):698–712.
11. Short KM, Smyth IM. The contribution of branching morphogenesis to kidney development and disease. *Nat Rev Nephrol*. 2016;12(12):754–767.
12. Grobstein C. Inductive epitheliomesenchymal interaction in cultured organ rudiments of the mouse. *Science*. 1953;118(3054):52–55.
13. Vainio S, Lin Y. Coordinating early kidney development: lessons from gene targeting. *Nat Rev Genet*. 2002;3(7):533–543.
14. Pope JC, Brock JW, Adams MC, Stephens FD, Ichikawa I. How they begin and how they end: classic and new theories for the development and deterioration of congenital anomalies of the kidney and urinary tract, CAKUT. *J Am Soc Nephrol*. 1999;10(9):2018–2028.
15. Vats AN, et al. Steroid-resistant nephrotic syndrome and congenital anomalies of kidneys: evidence of locus on chromosome 13q. *Kidney Int*. 2003;64(1):17–24.
16. Fletcher J, et al. Multicystic dysplastic kidney and variable phenotype in a family with a novel deletion mutation of PAX2. *J Am Soc Nephrol*. 2005;16(9):2754–2761.
17. Nishimoto K, et al. PAX2 gene mutation in a family with isolated renal hypoplasia. *J Am Soc Nephrol*. 2001;12(8):1769–1772.
18. Weber S, et al. Prevalence of mutations in renal developmental genes in children with renal hypodysplasia: results of the ESCAPE study. *J Am Soc Nephrol*. 2006;17(10):2864–2870.
19. Yoshida J, Tsuchiya M, Tatsuma N, Murakami M. Mass screening for early detection of congenital kidney and urinary tract abnormalities in infancy. *Pediatr Int*. 2003;45(2):142–149.
20. Nicolaou N, Renkema KY, Bongers EM, Giles RH, Knoers NV. Genetic, environmental, and epigenetic factors involved in CAKUT. *Nat Rev Nephrol*. 2015;11(12):720–731.
21. Garne E, Dolk H, Loane M, Boyd PA, EUROCAT. EUROCAT website data on prenatal detection rates of congenital anomalies. *J Med Screen*. 2010;17(2):97–98.
22. [No authors listed]. Birth Defects Monitoring Program (BDMP)/Commission on Professional Hospital Activities (CPHA) surveillance data, 1988-1991. *Teratology*. 1993;48(6):658–675.
23. [No authors listed]. Metropolitan Atlanta Congenital Defects Program surveillance data, 1988–1991. *Teratology*. 1993;48(6):695–709.
24. Schulman J, Edmonds LD, McClearn AB, Jensvold N, Shaw GM. Surveillance for and comparison of birth defect prevalences in two geographic areas--United States, 1983-88. *MMWR CDC Surveill Summ*. 1993;42(1):1–7.
25. Wühl E, et al. Timing and outcome of renal replacement therapy in patients with congenital malformations of the kidney and urinary tract. *Clin J Am Soc Nephrol*. 2013;8(1):67–74.
26. Centers for Disease Control Prevention (CDC). State-specific trends in chronic kidney failure — United States, 1990–2001. *MMWR Morb Mortal Wkly Rep*. 2004;53(39):918–920.
27. Ardissino G, et al. Epidemiology of chronic renal failure in children: data from the ItalKid project. *Pediatrics*. 2003;111(4 pt 1):e382–e387.
28. Sanna-Cherchi S, et al. Renal outcome in patients with congenital anomalies of the kidney and urinary tract. *Kidney Int*. 2009;76(5):528–533.
29. Westland R, Schreuder MF, Bökenkamp A, Spreeuwenberg MD, van Wijk JA. Renal injury in children with a solitary functioning kidney — the KIMONO study. *Nephrol Dial Transplant*. 2011;26(5):1533–1541.
30. Corbani V, Ghiggeri GM, Sanna-Cherchi S. 'Congenital solitary functioning kidneys: which ones warrant follow-up into adult life?'. *Nephrol Dial Transplant*. 2011;26(5):1458–1460.
31. Groen In 't Woud S, et al. Maternal risk factors involved in specific congenital anomalies of the kidney and urinary tract: A case-control study. *Birth Defects Res Part A Clin Mol Teratol*. 2016;106(7):596–603.
32. Macumber I, Schwartz S, Leca N. Maternal obesity is associated with congenital anomalies of the kidney and urinary tract in offspring. *Pediatr Nephrol*. 2017;32(4):635–642.
33. Collins AJ, et al. US Renal Data System 2012 Annual Data Report. *Am J Kidney Dis*. 2013; 6(1 suppl 1):A7, e1–A7,476.
34. McDonald SP, Craig JC, Australian New Zealand Paediatric Nephrology Association. Long-term survival of children with end-stage renal disease. *N Engl J Med*. 2004;350(26):2654–2662.
35. Chesnaye NC, et al. Mortality risk disparities in children receiving chronic renal replacement therapy for the treatment of end-stage renal disease across Europe: an ESPN-ERA/EDTA registry analysis. *Lancet*. 2017;389(10084):2128–2137.
36. ESCAPE Trial Group, et al. Strict blood-pressure control and progression of renal failure in children. *N Engl J Med*. 2009;361(17):1639–1650.
37. Cooper GM, et al. Distribution and intensity of constraint in mammalian genomic sequence. *Genome Res*. 2005;15(7):901–913.
38. Samocha KE, et al. A framework for the interpretation of de novo mutation in human disease. *Nat Genet*. 2014;46(9):944–950.
39. Sanna-Cherchi S, et al. Localization of a gene for nonsyndromic renal hypodysplasia to chromosome 1p32-33. *Am J Hum Genet*. 2007;80(3):539–549.
40. Sanna-Cherchi S, et al. Familial vesicoureteral reflux: testing replication of linkage in seven new multigenerational kindreds. *J Am Soc Nephrol*. 2005;16(6):1781–1787.
41. Peters PC, Johnson DE, Jackson JH. The incidence of vesicoureteral reflux in the premature child. *J Urol*. 1967;97(2):259–260.
42. Scott JE, Swallow V, Coulthard MG, Lambert HJ, Lee RE. Screening of newborn babies for familial ureteric reflux. *Lancet*. 1997;350(9075):396–400.
43. Hartman GW, Hodson CJ. The duplex kidney and related abnormalities. *Clin Radiol*. 1969;20(4):387–400.
44. Privett JT, Jeans WD, Roylance J. The incidence and importance of renal duplication. *Clin Radiol*. 1976;27(4):521–530.
45. Groenen PM, Vanderlinden G, Devriendt K, Fryns JP, Van de Ven WJ. Rearrangement of the human CDC5L gene by a t(6;19)(p21;q13.1) in a patient with multicystic renal dysplasia. *Genomics*. 1998;49(2):218–229.
46. Ogata T, et al. Genetic evidence for a novel gene(s) involved in urogenital development on 10q26. *Kidney Int*. 2000;58(6):2281–2290.
47. Vats KR, et al. A locus for renal malformations including vesico-ureteric reflux on chromosome 13q33-34. *J Am Soc Nephrol*. 2006;17(4):1158–1167.
48. Conrad DF, et al. Origins and functional impact of copy number variation in the human genome. *Nature*. 2010;464(7289):704–712.
49. Sebat J, et al. Large-scale copy number polymorphism in the human genome. *Science*. 2004;305(5683):525–528.
50. Redon R, et al. Global variation in copy number in the human genome. *Nature*. 2006;444(7118):444–454.
51. Iafrate AJ, et al. Detection of large-scale variation in the human genome. *Nat Genet*. 2004;36(9):949–951.
52. Wapner RJ, et al. Chromosomal microarray versus karyotyping for prenatal diagnosis. *N Engl J Med*. 2012;367(23):2175–2184.
53. Reddy UM, et al. Karyotype versus microarray testing for genetic abnormalities after stillbirth. *N Engl J Med*. 2012;367(23):2185–2193.
54. South ST, Lee C, Lamb AN, Higgins AW, Kearney HM, Working Group for the American College of Medical Genetics Genomics Laboratory Quality Assurance Committee. ACMG Standards and Guidelines for constitutional cytogenomic microarray analysis, including postnatal and prenatal applications: revision 2013. *Genet Med*. 2013;15(11):901–909.
55. 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.
56. Lupski JR. Genomic rearrangements and sporadic disease. *Nat Genet*. 2007;39(7 suppl):S43–S47.
57. Lee JA, Carvalho CM, Lupski JR. A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. *Cell*. 2007;131(7):1235–1247.
58. Carvalho CM, Lupski JR. Mechanisms underlying structural variant formation in genomic disorders. *Nat Rev Genet*. 2016;17(4):224–238.
59. Cooper GM, et al. A copy number variation morbidity map of developmental delay. *Nat Genet*. 2011;43(9):838–846.
60. 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.
61. Osoegawa K, et al. Identification of novel candidate genes associated with cleft lip and palate using array comparative genomic hybridisation. *J Med Genet*. 2008;45(2):81–86.
62. Serra-Juhé C, et al. Contribution of rare copy number variants to isolated human malformations. *PLoS One*. 2012;7(10):e45530.
63. Brunetti-Pierri N, et al. Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities. *Nat Genet*. 2008;40(12):1466–1471.
64. 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.
65. Sebat J, et al. Strong association of de novo copy number mutations with autism. *Science*. 2007;316(5823):445–449.
66. Stefansson H, et al. Large recurrent microdeletions associated with schizophrenia. *Nature*. 2008;455(7210):232–236.
67. Yu L, et al. De novo copy number variants are associated with congenital diaphragmatic hernia. *J Med Genet*. 2012;49(10):650–659.
68. Holder AM, Klaassens M, Tibboel D, de Klein A, Lee B, Scott DA. Genetic factors in congenital diaphragmatic hernia. *Am J Hum Genet*. 2007;80(5):825–845.
69. 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.
70. Verbitsky M, et al. Genomic imbalances in pediatric patients with chronic kidney disease. *J Clin Invest*. 2015;125(5):2171–2178.
71. Westland R, et al. Copy number variation analysis identifies novel CAKUT candidate genes in children with a solitary functioning kidney. *Kidney Int*. 2015;88(6):1402–1410.
72. Boghossian NS, et al. Rare copy number variants implicated in posterior urethral valves. *Am J Med Genet A*. 2016;170(3):622–633.
73. Faure A, et al. DNA copy number variants: A potentially useful predictor of early onset renal failure in boys with posterior urethral valves. *J Pediatr Urol*. 2016;12(4):227.e1–227.e7.
74. Caruana G, et al. Copy-number variation associated with congenital anomalies of the kidney and urinary tract. *Pediatr Nephrol*. 2015;30(3):487–495.
75. Xi Q, et al. Copy number variations in multicystic dysplastic kidney: update for prenatal diagnosis and genetic counseling. *Prenat Diagn*. 2016;36(5):463–468.
76. Mefford HC, et al. Recurrent reciprocal genomic rearrangements of 17q12 are associated with renal disease, diabetes, and epilepsy. *Am J Hum Genet*. 2007;81(5):1057–1069.
77. Kobrynski LJ, Sullivan KE. Velocardiofacial syndrome, DiGeorge syndrome: the chromosome 22q11.2 deletion syndromes. *Lancet*. 2007;370(9596):1443–1452.
78. Mefford HC, et al. Recurrent rearrangements of chromosome 1q21.1 and variable pediatric phenotypes. *N Engl J Med*. 2008;359(16):1685–1699.
79. Copelovitch L, Warady BA, Furth SL. Insights from the Chronic Kidney Disease in Children (CKiD) study. *Clin J Am Soc Nephrol*. 2011;6(8):2047–2053.
80. 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.
81. Golzio C, Katsanis N. Genetic architecture of reciprocal CNVs. *Curr Opin Genet Dev*. 2013;23(3):240–248.
82. Golzio C, et al. KCTD13 is a major driver of mirrored neuroanatomical phenotypes of the 16p11.2 copy number variant. *Nature*. 2012;485(7398):363–367.
83. Horikawa Y, et al. Mutation in hepatocyte nuclear factor-1 beta gene (TCF2) associated with MODY. *Nat Genet*. 1997;17(4):384–385.
84. Decramer S, et al. Anomalies of the TCF2 gene are the main cause of fetal bilateral hyperechogenic kidneys. *J Am Soc Nephrol*. 2007;18(3):923–933.
85. Ulinski T, et al. Renal phenotypes related to hepatocyte nuclear factor-1beta (TCF2) mutations in a pediatric cohort. *J Am Soc Nephrol*. 2006;17(2):497–503.
86. Bellanné-Chantelot C, et al. Large genomic rearrangements in the hepatocyte nuclear factor-1β (TCF2) gene are the most frequent cause of maturity-onset diabetes of the young type 5. *Diabetes*. 2005;54(11):3126–3132.
87. Poll AV, et al. A vHNF1/TCF2-HNF6 cascade regulates the transcription factor network that controls generation of pancreatic precursor cells. *Diabetes*. 2006;55(1):61–69.
88. Oliver-Krasinski JM, et al. The diabetes gene Pdx1 regulates the transcriptional network of pancreatic endocrine progenitor cells in mice. *J Clin Invest*. 2009;119(7):1888–1898.
89. Hiesberger T, et al. Mutation of hepatocyte nuclear factor-1beta inhibits Pkhd1 gene expression and produces renal cysts in mice. *J Clin Invest*. 2004;113(6):814–825.
90. Heidet L, et al. Spectrum of HNF1B mutations in a large cohort of patients who harbor renal diseases. *Clin J Am Soc Nephrol*. 2010;5(6):1079–1090.
91. Pearson ER, et al. Contrasting diabetes phenotypes associated with hepatocyte nuclear factor-1α and -1β mutations. *Diabetes Care*. 2004;27(5):1102–1107.
92. Nagamani SC, et al. Clinical spectrum associated with recurrent genomic rearrangements in chromosome 17q12. *Eur J Hum Genet*. 2010;18(3):278–284.
93. Bierhals T, Maddukuri SB, Kutsche K, Girisha KM. Expanding the phenotype associated with 17q12 duplication: case report and review of the literature. *Am J Med Genet A*. 2013;161A(2):352–359.
94. Ruderfer DM, et al. Patterns of genic intolerance of rare copy number variation in 59,898 human exomes. *Nat Genet*. 2016;48(10):1107–1111.
95. Adalat S, et al. HNF1B mutations associate with hypomagnesemia and renal magnesium wasting. *J Am Soc Nephrol*. 2009;20(5):1123–1131.
96. Moreno-De-Luca D, et al. Deletion 17q12 is a recurrent copy number variant that confers high risk of autism and schizophrenia. *Am J Hum Genet*. 2010;87(5):618–630.
97. Clissold RL, et al. Chromosome 17q12 microdeletions but not intragenic HNF1B mutations link developmental kidney disease and psychiatric disorder. *Kidney Int*. 2016;90(1):203–211.
98. Portnoï MF. Microduplication 22q11.2: a new chromosomal syndrome. *Eur J Med Genet*. 2009;52(2–3):88–93.
99. Lopez-Rivera E, et al. Genetic drivers of kidney defects in the DiGeorge syndrome. *N Engl J Med*. 2017;376(8):742–754.
100. Haller M, Mo Q, Imamoto A, Lamb DJ. Murine model indicates 22q11.2 signaling adaptor CRKL is a dosage-sensitive regulator of genitourinary development. *Proc Natl Acad Sci U S A*. 2017;114(19):4981–4986.
101. Schnabel CA, Godin RE, Cleary ML. Pbx1 regulates nephrogenesis and ureteric branching in the developing kidney. *Dev Biol*. 2003;254(2):262–276.
102. Heidet L, et al. Targeted exome sequencing identifies PBX1 as involved in monogenic congenital anomalies of the kidney and urinary tract. *J Am Soc Nephrol*. 2017;28(10):2901–2914.
103. Westland R, Schreuder MF, van Goudoever JB, Sanna-Cherchi S, van Wijk JA. Clinical implications of the solitary functioning kidney. *Clin J Am Soc Nephrol*. 2014;9(5):978–986.
104. Hildebrandt F. Genetic kidney diseases. *Lancet*. 2010;375(9722):1287–1295.
105. Vivante A, Kohl S, Hwang DY, Dworschak GC, Hildebrandt F. Single-gene causes of congenital anomalies of the kidney and urinary tract (CAKUT) in humans. *Pediatr Nephrol*. 2014;29(4):695–704.
106. Thomas R, Sanna-Cherchi S, Warady BA, Furth SL, Kaskel FJ, Gharavi AG. HNF1B and PAX2 mutations are a common cause of renal hypodysplasia in the CKiD cohort. *Pediatr Nephrol*. 2011;26(6):897–903.
107. Madariaga L, et al. Severe prenatal renal anomalies associated with mutations in HNF1B or PAX2 genes. *Clin J Am Soc Nephrol*. 2013;8(7):1179–1187.
108. Hoskins BE, et al. Missense mutations in EYA1 and TCF2 are a rare cause of urinary tract malformations. *Nephrol Dial Transplant*. 2008;23(2):777–779.
109. Nicolaou N, et al. Prioritization and burden analysis of rare variants in 208 candidate genes suggest they do not play a major role in CAKUT. *Kidney Int*. 2016;89(2):476–486.
110. Sanna-Cherchi S, et al. Mutations in DSTYK and dominant urinary tract malformations. *N Engl J Med*. 2013;369(7):621–629.
111. Lee JYW, et al. Large intragenic deletion in DSTYK underlies autosomal-recessive complicated spastic paraparesis, SPG23. *Am J Hum Genet*. 2017;100(2):364–370.
112. Bekheirnia MR, et al. Whole-exome sequencing in the molecular diagnosis of individuals with congenital anomalies of the kidney and urinary tract and identification of a new causative gene. *Genet Med*. 2017;19(4):412–420.
113. Hamdan FF, et al. De novo mutations in FOXP1 in cases with intellectual disability, autism, and language impairment. *Am J Hum Genet*. 2010;87(5):671–678.
114. Vivante A, et al. Mutations in TBX18 cause dominant urinary tract malformations via transcriptional dysregulation of ureter development. *Am J Hum Genet*. 2015;97(2):291–301.
115. Vivante A, et al. A dominant mutation in nuclear receptor interacting protein 1 causes urinary tract malformations via dysregulation of retinoic acid signaling. *J Am Soc Nephrol*. 2017;28(8):2364–2376.
116. Sanna-Cherchi S, et al. Exome-wide association study identifies GREB1L mutations in con-

genital kidney malformations. *Am J Hum Genet*. 2017;101(5):789–802.
117. Brophy PD, et al. A gene implicated in activation of retinoic acid receptor targets is a novel renal agenesis gene in humans. *Genetics*. 2017;207(1):215–228.
118. Vivante A, et al. Exome sequencing discerns syndromes in patients from consanguineous families with congenital anomalies of the kidneys and urinary tract. *J Am Soc Nephrol*. 2017;28(1):69–75.
119. Saisawat P, et al. Identification of two novel CAKUT-causing genes by massively parallel exon resequencing of candidate genes in patients with unilateral renal agenesis. *Kidney Int*. 2012;81(2):196–200.
120. Gribouval O, et al. Mutations in genes in the renin-angiotensin system are associated with autosomal recessive renal tubular dysgenesis. *Nat Genet*. 2005;37(9):964–968.
121. Kohl S, et al. Mild recessive mutations in six Fraser syndrome-related genes cause isolated congenital anomalies of the kidney and urinary tract. *J Am Soc Nephrol*. 2014;25(9):1917–1922.
122. van Haelst MM, Scambler PJ, Fraser Syndrome Collaboration Group, Hennekam RC. Fraser syndrome: a clinical study of 59 cases and evaluation of diagnostic criteria. *Am J Med Genet A*. 2007;143A(24):3194–3203.
123. Alazami AM, et al. FREM1 mutations cause bifid nose, renal agenesis, and anorectal malformations syndrome. *Am J Hum Genet*. 2009;85(3):414–418.
124. Saisawat P, et al. Whole-exome resequencing reveals recessive mutations in TRAP1 in individuals with CAKUT and VACTERL association. *Kidney Int*. 2014;85(6):1310–1317.
125. Girirajan S, et al. Phenotypic heterogeneity of genomic disorders and rare copy-number variants. *N Engl J Med*. 2012;367(14):1321–1331.
126. Tropeano M, et al. Male-biased autosomal effect of 16p13.11 copy number variation in neurodevelopmental disorders. *PLoS One*. 2013;8(4):e61365.
127. Beales PL, Elcioglu N, Woolf AS, Parker D, Flinter FA. New criteria for improved diagnosis of Bardet-Biedl syndrome: results of a population survey. *J Med Genet*. 1999;36(6):437–446.
128. Zaghloul NA, Katsanis N. Mechanistic insights into Bardet-Biedl syndrome, a model ciliopathy. *J Clin Invest*. 2009;119(3):428–437.
129. Badano JL, et al. Dissection of epistasis in oligogenic Bardet-Biedl syndrome. *Nature*. 2006;439(7074):326–330.
130. Zaghloul NA, et al. Functional analyses of variants reveal a significant role for dominant negative and common alleles in oligogenic Bardet-Biedl syndrome. *Proc Natl Acad Sci U S A*. 2010;107(23):10602–10607.
131. Katsanis N. The oligogenic properties of Bardet-Biedl syndrome. *Hum Mol Genet*. 2004;13(Spec No 1):R65–R71.
132. Beales PL, et al. Genetic interaction of BBS1 mutations with alleles at other BBS loci can result in non-Mendelian Bardet-Biedl syndrome. *Am J Hum Genet*. 2003;72(5):1187–1199.
133. Bergmann C, et al. Mutations in multiple PKD genes may explain early and severe polycystic kidney disease. *J Am Soc Nephrol*. 2011;22(11):2047–2056.
134. Draaken M, et al. Genome-wide association study and meta-analysis identify ISL1 as genome-wide significant susceptibility gene for bladder exstrophy. *PLoS Genet*. 2015;11(3):e1005024.
135. Yu Y, et al. Genome-wide analyses of non-syndromic cleft lip with palate identify 14 novel loci and genetic heterogeneity. *Nat Commun*. 2017;8:14364.
136. Cordell HJ, et al. Genome-wide association study of multiple congenital heart disease phenotypes identifies a susceptibility locus for atrial septal defect at chromosome 4p16. *Nat Genet*. 2013;45(7):822–824.
137. van der Zanden LF, et al. Common variants in DGKK are strongly associated with risk of hypospadias. *Nat Genet*. 2011;43(1):48–50.
138. Geller F, et al. Genome-wide association analyses identify variants in developmental genes associated with hypospadias. *Nat Genet*. 2014;46(9):957–963.
139. Speed D, Hemani G, Johnson MR, Balding DJ. Improved heritability estimation from genome-wide SNPs. *Am J Hum Genet*. 2012;91(6):1011–1021.
140. Ehret GB, et al. A multi-SNP locus-association method reveals a substantial fraction of the missing heritability. *Am J Hum Genet*. 2012;91(5):863–871.
141. Vilhjálmsson BJ, et al. Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *Am J Hum Genet*. 2015;97(4):576–592.
142. Bulik-Sullivan BK, et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet*. 2015;47(3):291–295.
143. Hormozdiari F, et al. Widespread allelic heterogeneity in complex traits. *Am J Hum Genet*. 2017;100(5):789–802.
144. Purcell SM, et al. A polygenic burden of rare disruptive mutations in schizophrenia. *Nature*. 2014;506(7487):185–190.
145. Robinson EB, et al. Autism spectrum disorder severity reflects the average contribution of de novo and familial influences. *Proc Natl Acad Sci U S A*. 2014;111(42):15161–15165.
146. Robinson EB, et al. Genetic risk for autism spectrum disorders and neuropsychiatric variation in the general population. *Nat Genet*. 2016;48(5):552–555.
147. Weiner DJ, et al. Polygenic transmission disequilibrium confirms that common and rare variation act additively to create risk for autism spectrum disorders. *Nat Genet*. 2017;49(7):978–985.
148. Kosmicki JA, et al. Refining the role of de novo protein-truncating variants in neurodevelopmental disorders by using population reference samples. *Nat Genet*. 2017;49(4):504–510.
149. Boyle EA, Li YI, Pritchard JK. An expanded view of complex traits: from polygenic to omnigenic. *Cell*. 2017;169(7):1177–1186.
150. MacArthur DG, et al. Guidelines for investigating causality of sequence variants in human disease. *Nature*. 2014;508(7497):469–476.
151. 1000 Genomes Project Consortium, et al. A global reference for human genetic variation. *Nature*. 2015;526(7571):68–74.
152. Lek M, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536(7616):285–291.
153. Landrum MJ, et al. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res*. 2016;44(D1):D862–D868.
154. Girirajan S, Campbell CD, Eichler EE. Human copy number variation and complex genetic disease. *Annu Rev Genet*. 2011;45:203–226.
155. Slickers J, Duquette P, Hooper S, Gipson D. Clinical predictors of neurocognitive deficits in children with chronic kidney disease. *Pediatr Nephrol*. 2007;22(4):565–572.
156. Verbitsky M, et al. Genomic disorders and neurocognitive impairment in pediatric CKD. *J Am Soc Nephrol*. 2017;28(8):2303–2309.
157. Hidalgo G, et al. Association of income level with kidney disease severity and progression among children and adolescents with CKD: a report from the Chronic Kidney Disease in Children (CKiD) Study. *Am J Kidney Dis*. 2013;62(6):1087–1094.
158. Kogon AJ, et al. Depressive symptoms in children with chronic kidney disease. *J Pediatr*. 2016;168:164–70.e1.
159. Klintwall L, Eldevik S, Eikeseth S. Narrowing the gap: effects of intervention on developmental trajectories in autism. *Autism*. 2015;19(1):53–63.
160. Yang Y, et al. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N Engl J Med*. 2013;369(16):1502–1511.
161. Biesecker LG, Green RC. Diagnostic clinical genome and exome sequencing. *N Engl J Med*. 2014;370(25):2418–2425.
162. Korf BR, Rehm HL. New approaches to molecular diagnosis. *JAMA*. 2013;309(14):1511–1521.
163. Hwang DY, et al. Mutations in 12 known dominant disease-causing genes clarify many congenital anomalies of the kidney and urinary tract. *Kidney Int*. 2014;85(6):1429–1433.
164. Richards S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405–424.
165. Rehm HL, et al. ClinGen — the Clinical Genome Resource. *N Engl J Med*. 2015;372(23):2235–2242.
166. Dewey FE, et al. Distribution and clinical impact of functional variants in 50,726 whole-exome sequences from the DiscovEHR study. *Science*. 2016;354(6319):aaf6814.
167. Manrai AK, et al. Genetic misdiagnoses and the potential for health disparities. *N Engl J Med*. 2016;375(7):655–665.
168. Popejoy AB, Fullerton SM. Genomics is failing on diversity. *Nature*. 2016;538(7624):161–164.
169. Petrovski S, Goldstein DB. Unequal representation of genetic variation across ancestry groups creates healthcare inequality in the application of precision medicine. *Genome Biol*. 2016;17(1):157.