Genes, Exomes, Genomes, Copy Number: What is Their Future in Pediatric Renal Disease

Matthew G. Sampson · Harald Jüppner

Published online: 22 December 2012
© Springer Science + Business Media New York 2012

Abstract The influence of genetic variation on the pathogenesis of pediatric kidney disease extends from the earliest stages of kidney development in utero to conditions arising throughout a child's life. Major advances in genomic technologies, computing power, and bioinformatics analyses have resulted in the accelerated discovery of novel genes and risk loci associated with both inherited and sporadic forms of pediatric kidney disease. In this review, we will highlight studies over the past year that used diverse approaches to discover novel genes and loci associated with pediatric renal disease. We will also discuss reports that investigate the association with disease of previously discovered risk variants in novel populations, different phenotypes, or in model systems. Finally, we will discuss how we believe genomic inquiry will evolve in pediatric kidney disease in the future. Together, these studies illustrate that almost every child with a kidney condition could participate in some form of genomic investigation.

Keywords Pediatric nephrology · Genes · Next-generation sequencing · Exome · Genomics · Genome · Transcriptome · CNV · GWAS ·

Single nucleotide variant · Nephrotic syndrome · Chronic kidney disease · Family studies · Phenotyping

Introduction

The influence of genetic variation on the pathogenesis of pediatric kidney disease extends from the earliest stages of kidney development in utero to conditions that arise throughout a child's life. Most successful gene discovery efforts have come from studying rare, large kindreds with multiply affected members [1–4], groups of children who share a similar renal phenotype [5], or children with a chromosomal abnormality that have multiple anomalies, including those of the kidney [6]. Using techniques such as linkage analysis, positional cloning, and karyotype analysis, the single gene or chromosomal region for innumerable kidney conditions has been defined, ranging from autosomal recessive polycystic kidney disease [7] to branchio-oto-renal syndrome [5] to Wolf-Hirschhorn syndrome [8].

Major advances in genomic technologies and increases in computing power have resulted in the ability to generate, process, and store genotype and/or sequence data on a genome-wide level. There has been a parallel development of sophisticated analytical approaches to incorporate genomics data into the study of human disease, led by biostatisticians, bioinformaticians, and human geneticists. Importantly, these advancements have been accompanied by continually decreasing costs of sequencing, which have allowed more investigators to study a greater number of individuals.

These technologies have been successfully incorporated into existing family-based studies of Mendelian disease. In addition, the technologies and accompanying methodologies are allowing investigators to ask new sets of questions

M. G. Sampson (✉)
Division of Pediatric Nephrology, Department of Pediatrics and Communicable Disease, University of Michigan School of Medicine, West Medical Center Drive, A510D MSRB1, Ann Arbor, MI 48109, USA
e-mail: mgsamps@med.umich.edu

H. Jüppner
Pediatric Nephrology and Endocrine Unit, Massachusetts General Hospital and Harvard Medical School, 50 Blossom Street, Thier 10, Boston, MA 02114, USA

using large cohorts of unrelated individuals. These analyses include the role of low-frequency genetic variants or the contribution of copy number variants (CNV) in disease pathogenesis.

In this review, we will first define key terminology as it relates to describing types of genetic variants and methodologies for genomic inquiry. We will then highlight studies over the past year that have used these approaches to discover novel genes and loci associated with pediatric renal disease. We will also review studies that utilize risk variants, previously identified through genome-wide association studies (GWAS), to study the affects of these alleles in new populations or identify the functional role of these variants in model systems. Finally, we will discuss how we believe genomic inquiry will evolve in pediatric kidney disease research moving forward.

Definitions for Genomic Inquiry

Prior to discussing recent studies that employ innovative genomics approaches to define genes and loci associated with pediatric kidney disease, it is important to develop a working understanding of the terms and methodologies employed in genomic discovery investigations (Table 1).

Genomic Variants

Across the genome, nucleotide positions can be found that vary from the reference sequence in a certain percentage of the population. These can be referred to as “single nucleotide variants (SNVs).” Rare, disease-causing SNVs, absent in healthy subjects, were historically referred to as “mutations,” such as those in nephrin in congenital

nephrotic syndrome [9] and alanine/glyoxylate amino-transferase in primary hyperoxaluria type 1 [10].

Beginning in the early 2000s, SNVs and their frequencies were systematically defined genome-wide in different populations. Millions of SNVs with a minor allele frequency greater than 5 % were discovered. These were called single nucleotide polymorphisms (SNPs) and were employed to perform GWAS. With the deeper sequencing of more subjects over the past decade, increasingly rare SNVs have been annotated. A common nomenclature that has arisen refers to SNVs of a frequency between 0.5 and 5 % as “low-frequency variants” and less than 0.5 % as “rare variants.” A generalizable observation is that there is an inverse relationship between disease-associated variant frequency and magnitude of association with disease risk (lower frequency = increasing effect size). SNV may follow recessive, dominant, or additive modes of disease association.

Protein-Encoding Genes

Protein encoding genes are composed of exons, with intervening introns, and also have non-coding features associated with them, such as promoters or enhancers, which regulate their transcription. Of the 3,572 disease phenotypes listed in the Online Mendelian Inheritance of Man (OMIM) database, the vast majority are due to either exonic variants or deletions or duplication of genes [<http://omim.org/>]. The traditional method for gene discovery has been by identification of a candidate causal gene followed by Sanger sequencing of the exons and splice sites of candidate genes. While incredibly successful and accurate, it is time-consuming, costly, and interrogates only one gene at a time.

Table 1 Genomic discovery methodologies

Method	Key characteristics
Sanger sequencing	Assess for variants in a single candidate gene or screen large cohorts for variant in very few known disease associated genes
Whole exome sequencing	Detection of variants of all frequencies in all exons of the genome. Hypothesis: disease causing variant is a coding variant in an unknown gene
Targeted exon sequencing	Detection of coding variants of all frequencies in a subset of known or candidate disease genes. Can range from 10s to 1000s of genes
Whole genome sequencing	Detection of sequence and structural variants of entire genome, in both coding and regulatory regions. Current limitations = expensive and analytically challenging
Genome-wide association study	Identify and compare allele frequencies in case–control study design using common variants across genic and nongenic regions of genome. Hypothesis-independent. Risk variants/haplotypes discovered most often not causal. Instead, marks a genomic region for follow up fine-mapping to discover causative variant
CNV detection	Found with SNP arrays or aCGH. Discovers regions of the genome that have common or rare duplications or deletions. Length of CNV from 1 Kb–many Mb. In genic and nongenic regions

LD linkage disequilibrium, *aCGH* array comparative genomic hybridization, *CNV* copy number variant, *Kb* kilobase, *Mb* megabase, *SNP* single nucleotide polymorphism

Genomes and Exomes

Besides protein-encoding genes, other functional elements of DNA include promoters, enhancers, transcription factor binding sites, epigenetic marks, and genes that are transcribed into micro-RNAs and long non-coding RNAs [11•]. DNA also has structural variants, with duplications and deletions, ranging from less than 10 base pair insertion–deletions (indels) to entire chromosomes [12]. Finally, DNA has a three-dimensional structure, involving histones, chromatin, and physical interactions of DNA segments that are far away in terms of linear distance across the nucleotide sequence [12, 13]. Taken together, all protein-coding genes, as well as these other characteristics of DNA, comprise the genome, and the study of these elements is referred to as “genomics.”

Next Generation Sequencing Technologies

There are many children with kidney disease for whom a clinical phenotype of family history suggests that there is a genetic component. Many times there are no obvious candidate genes on which to perform Sanger Sequencing, and either small family size or limited numbers of affected patients preclude traditional linkage or positional cloning approaches. In this situation, whole exome sequencing (WES), a method in which every exon of every protein-coding gene in an individual undergoes simultaneous next-generation sequencing (NGS), can discover rare, causal variants in a novel gene [14]. In other situations in which the clinical phenotype of an unsolved disease suggests a certain mechanism of disease (e.g., a ciliopathy), targeted exon capture can be used to select large numbers (10–100 s) of hypothesis-based candidates for sequencing, rather than all 20,000 genes, to discover a causal variant [15]. This method can also be used to determine the prevalence of variants in a group of known disease genes in individuals with a genetically heterogeneous disease.

Whole genome sequencing (WGS), in which every nucleotide is sequenced, provides additional information about variants in non-coding, yet conserved regions of the genome. It can also determine copy number variants (CNV) in intergenic regions that may contribute to disease. However, its cost and the complexity of analysis of a dataset that is $\sim 50\times$ larger than that of WES has thus far precluded its widespread use.

Genotyping Arrays

Another major tool of genomic discovery is the genome-wide SNP genotyping array or “SNP array.” These can include millions of SNPs across the genome, or can be customized to SNPs that have a particular functionality or association with a certain type of disease. SNP arrays are

cost efficient, allowing investigators to study genomic variation in large cohorts of affected individuals across the genome in an undirected manner. In this way, investigators can identify risk loci and follow up these leads through fine mapping of the implicated regions.

SNP arrays can also detect CNV [16], segments of the genome that have either greater or fewer copies as compared to the reference sequence. This can be detected at a resolution down to 1 kb [17•]. CNV can be inherited or occur *de novo*, and can be benign or deleterious [18, 19]. CNV have been associated with disease through mechanisms such as direct changes of the dosage of a particular gene, or altering regulatory sites of the genome. CNV can also be detected by array comparative genome hybridization (aCGH) [20].

We will now highlight studies that have emerged over the past 12 months that illustrate how investigators in the human genetics of kidney disease have taken advantage of each of these technologies to (1) improve their ability to discover new genes causing rare, Mendelian diseases, (2) search for more common variants associated with complex diseases, or (3) study known genomic risk loci in different populations, different phenotypes, or in model systems.

Rare, Causal Variants

Pseudohypoaldosteronism Type II (PHAII), a rare Mendelian disease characterized by hypertension, hyperkalemia, and metabolic acidosis, has been previously found to be caused by mutations in a class of genes called WNK kinases [21]. However, many affected subjects remain who lack variants in these genes. Boyden et al. [22•] recruited 52 patients with PHAII and their immediate families, of which only 7 had *WNK* mutations, to discover new genes causing this disease. They performed whole exome sequencing (WES) on 11 unrelated affected individuals to discover rare variants in novel genes, and found 124 genes with three or more protein altering variants. At the same time, they performed SNP genotyping of the index cases and their families and determined that variants in 28/124 genes segregated with the disease state in two or more families. Using their knowledge of the physiology of PHAII and the distribution of rare variants in these candidates, they focused on kelch-like 3 (*KLHL3*), which had both heterozygous and homozygous variants in 5 of the initial 11 kindreds undergoing WES. Targeted sequencing of this gene in the affected members of all 52 kindreds found *KLHL3* mutations in 24 of the investigated kindreds. A main binding partner of *KHL3* is cullin 3 (*CUL3*), and rare heterozygous variants in this gene were found in 2/11 individuals initially undergoing WES. Thus targeted resequencing of *CUL3* was performed in all unsolved cases, and mutations were found in 17 individuals.

KLHL3 and *CUL3* are present in the distal nephron and act together to ubiquitinate targets. They are posited to cause the PHAII phenotype by ubiquitination of substrates that activate the Na–Cl cotransporter, and affect K⁺, and H⁺ handling. In this study, patients with *CUL3* mutations were found to be younger at presentation, with more severe hyperkalemia and acidosis, as well as growth impairment. Subjects with *KLHL3* mutations had a disease severity in between those with causal variants in *CUL3* and *WNK*.

Two independent groups used similar methodologies to discover rare variants in myosin 1E (*MYO1E*) as a cause of focal segmental glomerulosclerosis FSGS [23, 24•]. Both groups identified a consanguineous kindred that had three siblings with FSGS. By genotyping the family through high-density SNP arrays and performing both linkage analysis and homozygosity mapping (an approach to find autosomal recessive disease genes), the investigators were able to identify a region on chromosome 15 presumed to harbor the disease-causing variant. In one study, WES had already been performed on these subjects, so analysis was restricted to the 263 positional candidates at this locus. In the other study, targeted exon capture followed by NGS was performed on the 112 genes at the locus. Variant analysis and bioinformatic query identified homozygous variants in *MYO1E* as the most likely causal gene for FSGS in this family. Rare variants in this gene were subsequently replicated in an unrelated family and functional studies using cell cultures and model systems confirmed *MYO1E* as a newly discovered cause of FSGS (24). They showed that *MYO1E*, a non-muscle myosin, localizes to the podocyte and affects actin cytoskeleton dynamics, and when knocked out, causes NS in mice (24).

Congenital anomalies of the kidney and urinary tract (CAKUT) have been associated with rare variants in numerous genes, discovered in both children with this condition as well as animal models of CAKUT. This represents an ideal condition for targeted exon sequencing to test in many children at once whether variants in these known disease, or biologically plausible, genes contribute to the CAKUT phenotype. Saisawat et al. [25•] pooled the DNA from 40 children with non-syndromic CAKUT, 29 with unilateral renal agenesis (URA) and 11 with other forms of CAKUT. Simultaneous PCR of 313 exons of 30 candidate URA genes was performed followed by NGS and variant filtering and analysis. After numerous filtering steps, heterozygous variants were confirmed in 7/40 subjects, consisting of two known human CAKUT genes, *RET* and *BMP4*, as well as two novel ones, *FRAS1* and *FREM2*. Homozygous rare variants in *FRAS1* and *FREM2* have been implicated in URA in the context of Fraser Syndrome, but have not been previously implicated in non-syndromic CAKUT. The results of this study may lead to consideration of these genes as a more prevalent cause of non-syndromic CAKUT.

Copy Number Variants CNV

The pathogenic role of CNV in pediatric renal disease has been most apparent in CAKUT, with a previous report of four large pathogenic CNV in 3/30 children who had CAKUT plus an additional extrarenal anomaly [19]. These large CNV contained hundreds of genes and regulatory regions that require further investigation to determine the relationship with CAKUT. A recently published report builds upon this observation and expands its scope.

Sanna-Cherchi et al. [26•] used SNP arrays to examine the burden of rare CNV, >500 kb, in 192 patients with renal agenesis and/or hypodysplasia (RHD). Known and novel disorders of genomic copy number were discovered in this cohort and then replicated in 330 RHD cases from two independent cohorts. As compared to CNV of controls, subjects with RHD were found to have larger CNV and enrichment of duplications versus deletions. There were 34 known genomic disorders detected in 55/522 cases, including renal cysts and diabetes syndrome, Potocki-Lupski syndrome, and Wolf-Hirschhorn syndrome. In addition, another 40/522 cases had large CNVs impacting genes that were absent or extremely rare in 13,838 controls, thus identifying 38 novel or rare genomic disorders for RHD. Altogether, this indicates that about 18 % of subjects with CAKUT have a large, rare CNV.

Genome-Wide Association Studies GWAS

The discovery of rare, causal variants associated with pediatric kidney diseases has been invaluable. However, because of their rarity, these specific variants will not impact the risk of disease across a population. To identify common variants associated with increased risk of disease in populations, a GWAS can be employed. Once discovered, the higher frequency of these risk variants allow them to be studied in other populations. Specific questions can be asked about how this risk SNP impacts other populations, disease phenotypes, or outcomes.

Encouragingly, in the past 4 years, GWAS has discovered risk alleles associated with glomerular diseases such as FSGS [27, 28], membranous nephropathy MN [29], IgA nephropathy (IGAN) [30], and acquired nephrotic syndrome [31], as well as with phenotypes relating to kidney function, such as estimated glomerular filtration rate (eGFR), chronic kidney disease (CKD) [32] and albuminuria [33]. The GWAS of FSGS and MN are notable in that variants of large effect size were identified using small, but very well, histologically characterized, cohorts of affected individuals. Identification of specific risk SNPs in all of these phenotypes have allowed them to be focused upon in numerous follow up investigations. We will discuss some

of these pertinent studies that relate to the risk alleles associated with FSGS and eGFR/CKD.

Initial GWAS of African-Americans (AA) with FSGS and hypertensive ESRD implicated a chromosome 22 locus with greatly increased risk of these conditions [27, 34]. *MYH9* was the initial gene implicated as harboring the causal risk variant for these phenotypes in AA. Four years later, it is known that while *MYH9* is associated with modestly increased odds of disease, it is in fact, variants in apolipoprotein L1 (*APOLI*) at this locus that confer the majority of the increased risk of FSGS in this population [28]. There is an increased odds ratio (OR) of 17 for FSGS in individuals with two copies of the risk allele [35]. In the past year, *APOLI* risk alleles have been studied in adults to characterize how their impact on a variety of kidney phenotypes or clinical decisions.

In a disease-based cohort study, FSGS associated with two *APOLI* risk alleles was associated with earlier onset of disease and faster progression of disease [36•]. The same study showed that having two risk alleles also puts an AA at a 4 % lifetime risk for FSGS and explains 18 % of FSGS in AA. The presence of two copies of the *APOLI* risk variants in AA adults were associated with microalbuminuria and eGFR <60 ml/min/m² [37], increased risk of hypertensive end-stage nephropathy and rate of kidney functional decline [38], and lower age of initiation of dialysis in those with non-diabetic ESRD [39, 40]. Finally, it was shown that in kidney transplants, allograft survival was shorter if the AA donor had two *APOLI* risk alleles [41], but was not affected by the *APOLI* status of the recipient [42].

In regards to loci that impact eGFR and CKD, Kottgen et al. performed two GWAS in 2009 and 2010 on more than 85,000 adults of European ancestry and discovered common variants in more than twenty loci associated with eGFR, prevalent CKD, or creatinine production/secretion [32, 43]. As opposed to the two *APOLI* risk alleles in FSGS that had large OR, all of these risk variants for CKD/eGFR together explain less than 2 % of the variance seen in eGFR in individuals. As in *APOLI*, follow-up investigations of these validated eGFR/CKD SNPs has focused on identifying the effects of these variants in other populations and in other phenotypes. There has also been interesting work done to use existing biological knowledge and functional studies to both prioritize risk loci for further investigation and to characterize the role of these SNPs and their associated genes in disease pathogenesis.

eGFR/CKD risk SNPs were applied to other phenotypes. SNPs associated with a lower eGFR were not found to be associated with urinary albumin to creatinine ratio or albuminuria [44•]. However, a SNP tagging the gene cubilin was found via GWAS to be associated with albuminuria [31], and it was also subsequently found to be

associated with ESRD in native kidneys and graft failure in transplanted kidneys [45]. The majority of eGFR-associated SNPs were shown to be associated with or show a strong trend towards association with incident CKD, but were not strongly associated with ESRD [46].

In terms of clinical translation of GWAS findings in CKD, investigators created a genetic risk score composed of 16 SNPs associated with stage 3 CKD [47•]. They tested its performance to determine its ability to predict incident CKD stage 3 in a previously collected prospective cohort. They found that this test's performance was no better than utilizing known clinical risk factors.

Finally, multiple studies are integrating eGFR/CKD GWAS data with biological functional data, such as gene expression data or gene ontology analysis. One successful study used this approach to help prioritize novel risk SNPs for kidney function [48•]. Another study determined the genes linked to the top SNPs implicated in a GWAS of CKD, and then used morpholino knockdown of these genes in zebrafish to illustrate the importance of these gene on the kidney phenotype [49, 50].

Future Directions

In both the near- and long-term, the future of genomic inquiry in pediatric nephrology is exciting (Table 2). Aided by the ability to successfully perform gene discovery studies using small kindreds or single affected individuals, the genetic lesions underlying other Mendelian forms of pediatric renal disease will be discovered using WES and WGS. In addition, the role of rare variants in highly conserved, non-coding or regulatory regions of the genome may be implicated in some inherited forms of renal disease. The deleterious effects of non-coding variants is becoming increasingly evident in other recently published work [51, 52].

WES-based or WGS-based studies of unrelated children with the same clinical phenotype will identify rare SNV associated with disease. Additionally, meaningful genetic risk variants may be found in association studies utilizing relatively small cohorts of well-phenotyped children with non-Mendelian forms of kidney diseases. The discovery of significant risk variants with large effect sizes in a GWAS of FSGS using 190 subjects [27] and of MN in a 75 subject cohort [29] supports this concept. Genotyping children for the risk SNPs discovered from adult GWAS, such as those associated with eGFR and CKD [32, 43] and IGAN, may help us to understand the genetic influence of these diseases in pediatrics patients [30].

We will capitalize on large, deeply phenotyped cohorts of children with non-Mendelian, complex disease such as NS in the Nephrotic Syndrome Study Network (NEPTUNE) [53] or CKD in the Chronic Kidney Disease in

Table 2 Near- and long-term future directions for genomic research of pediatric kidney disease

Proliferation of custom exon sequencing “panels” consisting of all known disease genes in cohorts of children with genetically heterogeneous diseases
Increased number of children being resequenced for known disease-genes due to dropping costs, resulting in increased understanding of genotype–phenotype correlations
Investigation of the role that combinations of rare and common variants play in the phenotypic heterogeneity/variable expressivity of pediatric kidney disease
Use of NGS on single affected children, single sib-pairs, or small kindreds to discover novel disease causing genes for known and unnamed conditions
Genotyping of risk alleles discovered from adult studies in cohorts of children with similar phenotypes
Identification of novel disease-associated risk alleles through genotyping or WGS of cohorts of children with the same kidney disease; taking advantage of family genetic structure when possible
Increased understanding of pathogenic role of genomic variation in noncoding regions of genome
Integration of genomic data with epigenetic, transcriptomic, proteomic, and metabolomic data from the same subject in a disease- and tissue-specific manner
Derivation of induced pluripotent stem cells from blood cells or fibroblasts of children with deleterious variants to allow creation of specific kidney cells containing these variants
Early and accurate prenatal diagnoses through sequencing of fetal DNA free-floating within blood of pregnant mother
NGS next-generation sequencing, WGS whole genome sequencing, SNP single nucleotide polymorphism, ENCODE encyclopedia of DNA elements

Children (CKiD) Study [54]. Genotyping arrays could identify novel disease-associated loci using either GWAS or specialized pathway-based SNV analyses. Genome-wide sequencing or targeted NGS efforts can be used as well to discover novel rare variants. Finally, these cohorts would allow the study of the effects of known risk variants previously identified in adult studies.

Finally, we will continue to integrate variant data with previous biological knowledge and newly generated gene expression information to expand our understanding of the genomic influence on the pathogenesis and progression of pediatric kidney disease. Recent results from the ENCODE project indicated that 80 % of the human genome has a functional role [55•]. The series of studies that used eGFR/CKD GWAS data demonstrated that defining whether a non-coding SNP from a GWAS has a functional effect can both narrow down a list of candidate SNPs as well as elucidate the biological effect of the particular variant on the disease under investigation [43, 48, 49].

This integrative, functional genomics approach will also be employed by NEPTUNE [53] to understand the role of regulatory variation in the pathogenesis of NS. Subjects with NS will undergo WGS, targeted exon sequencing, and exome chip genotyping, as well as generation of gene expression data derived from glomerular and tubulointerstitial tissue obtained via renal biopsy. With these DNA- and RNA-based datasets, subjects with known or novel genetic risk variants will be identified, and the effects of these variants on gene expression in their own diseased kidney tissue can then be functionally characterized. This information will then be integrated with clinical data to

connect coding and non-coding genetic variation to clinical phenotypes and outcomes. In the future, this approach could be applied to many pediatric renal diseases in cohorts of varying sizes.

Conclusion

In this era, it is apparent that almost every child with a kidney condition can be enrolled in some form of genomic investigation. Family-based studies can now be performed in kindreds of all sizes to discover rare, causal variants for conditions such as disorders of bone and mineral metabolism or nephronophthisis. A child with a particular phenotype, such as renal dysplasia, that has not been evaluated genetically, can be enrolled in a cohort with similarly affected, unrelated children. If a child has a phenotype (such as steroid-resistant nephrotic syndrome or atypical hemolytic uremic syndrome) with known disease-causing genes, they can all be screened simultaneously to achieve a genetic diagnosis. For those children who remain undiagnosed, further discovery efforts can be done in a genome-wide manner or by analyzing hundreds of candidate genes simultaneously. Children with sporadic or syndromic forms of CAKUT may benefit from undergoing an analysis for CNV. Finally, for children affected by those conditions that do not follow simple Mendelian inheritance patterns, such as NS, CKD, or nephrolithiasis, enrollment into a cohort with similar individuals could allow unique approaches that incorporate common and rare sequence and structural variant and/or gene expression to discover risk alleles.

Moving forward, efforts should be made to increase awareness that many of children with a renal phenotype may have a genetic component to their conditions. It is important to help clinicians determine whether a clinical- or research-based genetic study is indicated. And from a research perspective, the process of connecting patients and families with a particular condition to the appropriate investigator for the optimal genetic test should be made easier. By identifying the genomic contribution to pediatric renal disease in increasing numbers of children, diagnostic tests can be improved to include those genes known to be involved in the disease, genotype–phenotype correlations could be discovered or refined to inform clinical decisions making, and key biological pathways and networks can be identified and targeted for therapeutic development and intervention. It is with these advances that the discoveries from genomic research can be increasingly translated into genomic medicine for children with kidney disease.

Acknowledgments MS is supported in part by NIH grants from the George M. O'Brien Kidney Center Grant, P30-DK-08194 and NEPTUNE, the Nephrotic Syndrome Study Network, U54-DK-083912. MS is also supported by the National Center for Research Resources, Grant UL1RR024986, which is now at the National Center for Advancing Translational Sciences, Grant 2UL1TR000433.

Disclosure No potential conflicts of interest relevant to this article were reported.

References

1. Brown EJ, Schlondorff JS, Becker DJ, et al. Mutations in the formin gene INF2 cause focal segmental glomerulosclerosis. *Nat Genet.* 2010;42:72–6.
2. Bastepe M, Frohlich LF, Hendy GN, et al. Autosomal dominant pseudohypoparathyroidism type Ib is associated with a heterozygous microdeletion that likely disrupts a putative imprinting control element of GNAS. *J Clin Invest.* 2003;112:1255–63.
3. Bergwitz C, Roslin NM, Tieder M, et al. SLC34A3 mutations in patients with hereditary hypophosphatemic rickets with hypercalciuria predict a key role for the sodium-phosphate cotransporter NaPi-IIc in maintaining phosphate homeostasis. *Am J Hum Genet.* 2006;78:179–92.
4. Pook MA, Wrong O, Wooding C, et al. Dent's disease, a renal Fanconi syndrome with nephrocalcinosis and kidney stones, is associated with a microdeletion involving DXS255 and maps to Xp11.22. *Hum Mol Genet.* 1993;2:2129–34.
5. Kochhar A, Fischer SM, Kimberling WJ, et al. Branchio-oto-renal syndrome. *Am J Med Genet A.* 2007;143A:1671–8.
6. Perez Jurado LA, Peoples R, Kaplan P, et al. Molecular definition of the chromosome 7 deletion in Williams syndrome and parent-of-origin effects on growth. *Am J Hum Genet.* 1996;59:781–92.
7. Ward CJ, Hogan MC, Rossetti S, et al. The gene mutated in autosomal recessive polycystic kidney disease encodes a large, receptor-like protein. *Nat Genet.* 2002;30:259–69.
8. McKeown C, Read AP, Dodge A, et al. Wolf-Hirschhorn locus is distal to D4S10 on short arm of chromosome 4. *J Med Genet.* 1987;24:410–2.
9. Kestila M, Lenkkeri U, Mannikko M, et al. Positionally cloned gene for a novel glomerular protein—nephrin—is mutated in congenital nephrotic syndrome. *Mol Cell.* 1998;1:575–82.
10. Purdue PE, Takada Y, Danpure CJ. Identification of mutations associated with peroxisome-to-mitochondrion mistargeting of alanine:glyoxylate aminotransferase in primary hyperoxaluria type 1. *J Cell Biol.* 1990;111:2341–51.
11. • Bernstein BE, Birney E, Dunham I, et al. An integrated encyclopedia of DNA elements in the human genome. *Nature.* 2012;489:57–74. *Provides an overview of the insights regarding the organization and regulation of the genome gained through the Encyclopedia of DNA Elements project. Demonstrates that 80 % of our DNA has a biochemical function. Non-coding regions can be related to human disease.*
12. Thurman RE, Rynes E, Humbert R, et al. The accessible chromatin landscape of the human genome. *Nature.* 2012;489:75–82.
13. Sanyal A, Lajoie BR, Jain G, et al. The long-range interaction landscape of gene promoters. *Nature.* 2012;489:109–13.
14. Ng SB, Buckingham KJ, Lee C, et al. Exome sequencing identifies the cause of a Mendelian disorder. *Nat Genet.* 2010;42:30–5.
15. Otto EA, Hurd TW, Airik R, et al. Candidate exome capture identifies mutation of SDCCAG8 as the cause of a retinal-renal ciliopathy. *Nat Genet.* 2010;42:840–50.
16. Cooper GM, Mefford HC. Detection of copy number variation using SNP genotyping. *Methods Mol Biol.* 2011;767:243–52.
17. • Southard AE, Edlmann LJ, Gelb BD. Role of copy number variants in structural birth defects. *Pediatrics.* 2012;129:755–63. *Reviews the literature over the past year to identify studies that implicate CNV to structural birth defects. Described specific defects and loci associated with CNV and disease pathogenesis.*
18. Greenway SC, Pereira AC, Lin JC, et al. De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot. *Nat Genet.* 2009;41:931–5.
19. Weber S, Landwehr C, Renkert M, et al. Mapping candidate regions and genes for congenital anomalies of the kidneys and urinary tract CAKUT by array-based comparative genomic hybridization. *Nephrol Dial Transplant.* 2011;26:136–43.
20. Zhang F, Gu W, Hurles ME, et al. Copy number variation in human health, disease, and evolution. *Annu Rev Genomics Hum Genet.* 2009;10:451–81.
21. Furgeson SB, Linas S. Mechanisms of type I and type II pseudohypoaldosteronism. *J Am Soc Nephrol.* 2010;21:1842–5.
22. • Boyden LM, Choi M, Choate KA, et al. Mutations in kelch-like 3 and cullin 3 cause hypertension and electrolyte abnormalities. *Nature.* 2012;482:98–102. *Integrates whole exome sequencing with SNP-based linkage studies to discover novel single gene causes of pseudohypoaldosteronism type IIa. Describes variant filtering strategies, functional studies, and genotype-phenotype correlations.*
23. Sanna-Cherchi S, Burgess KE, Nees SN, et al. Exome sequencing identified MYO1E and NEIL1 as candidate genes for human autosomal recessive steroid-resistant nephrotic syndrome. *Kidney Int.* 2011;80:389–96.
24. • Mele C, Iatropoulos P, Donadelli R, et al. MYO1E mutations and childhood familial focal segmental glomerulosclerosis. *N Engl J Med.* 2011;365:295–306. *Defines novel FSGS genes using exome sequencing targeted at a region defined by family information. Demonstrates that disease causing genes can be discovered in single, small kindreds.*

25. • Saisawat P, Tasic V, Vega-Warner V, 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;81:196–200. *Successfully identifies genes associated with CAKUT using a targeted exon sequencing method. Model of using prior biologic/genetic knowledge of a disease to guide sequencing efforts.*
26. • Sanna-Cherchi S, Kiryluk K, Burgess KE, et al. Copy number disorders are a common cause of congenital kidney malformations. *Am J Hum Genet.* 2012. doi:10.1016/j.ajhg.2012.10.007. *Implicates CNV as a significant cause of renal hypodysplasia. Illustrates ability to define genomic risk factors from modest sized cohorts of unrelated children using SNP-genotyping.*
27. Kopp JB, Smith MW, Nelson GW, et al. MYH9 is a major-effect risk gene for focal segmental glomerulosclerosis. *Nat Genet.* 2008;40:1175–84.
28. Genovese G, Friedman DJ, Ross MD, et al. Association of trypanolytic ApoL1 variants with kidney disease in African Americans. *Science.* 2010;329:841–5.
29. Stanescu HC, Arcos-Burgos M, Medlar A, et al. Risk HLA-DQA1 and PLA2R1 alleles in idiopathic membranous nephropathy. *N Engl J Med.* 2011;364:616–26.
30. Gharavi AG, Kiryluk K, Choi M, et al. Genome-wide association study identifies susceptibility loci for IgA nephropathy. *Nat Genet.* 2011;43:321–7.
31. Okamoto K, Tokunaga K, Doi K, et al. Common variation in GPC5 is associated with acquired nephrotic syndrome. *Nat Genet.* 2011;43:459–63.
32. Kottgen A, Pattaro C, Boger CA, et al. New loci associated with kidney function and chronic kidney disease. *Nat Genet.* 2010;42:376–84.
33. Boger CA, Chen MH, Tin A, et al. CUBN is a gene locus for albuminuria. *J Am Soc Nephrol.* 2011;22:555–70.
34. Kao WH, Klag MJ, Meoni LA, et al. MYH9 is associated with nondiabetic end-stage renal disease in African Americans. *Nat Genet.* 2008;40:1185–92.
35. Pollak MR, Genovese G, Friedman DJ 2012 APOL1 and kidney disease. *Curr Opin Nephrol Hypertens.* 2008;21:179–82.
36. • Kopp JB, Nelson GW, Sampath K, et al. APOL1 genetic variants in focal segmental glomerulosclerosis and HIV-associated nephropathy. *J Am Soc Nephrol* 2011;22:2129–37. *Recent comprehensive study of clinical effects and risk associated with APOL1 risk alleles in adults. 17× greater odds of FSGS with two risk alleles, earlier onset, and faster progression to ESRD.*
37. Friedman DJ, Kozlitina J, Genovese G, et al. Population-based risk assessment of APOL1 on renal disease. *J Am Soc Nephrol.* 2011;22:2098–105.
38. Lipkowitz MS, Freedman BI, Langefeld CD, et al. Apolipoprotein L1 gene variants associate with hypertension-attributed nephropathy and the rate of kidney function decline in African Americans. *Kidney Int.* 2012. doi:10.1038/ki.2012.263.
39. Tzur S, Rosset S, Skorecki K, et al. APOL1 allelic variants are associated with lower age of dialysis initiation and thereby increased dialysis vintage in African and Hispanic Americans with non-diabetic end-stage kidney disease. *Nephrol Dial Transplant.* 2012;27:1498–505.
40. Kanji Z, Powe CE, Wenger JB, et al. Genetic variation in APOL1 associates with younger age at hemodialysis initiation. *J Am Soc Nephrol.* 2011;22:2091–7.
41. Reeves-Daniel AM, DePalma JA, Bleyer AJ, et al. The APOL1 gene and allograft survival after kidney transplantation. *Am J Transplant.* 2011;11:1025–30.
42. Lee BT, Kumar V, Williams TA, et al. The APOL1 genotype of African American kidney transplant recipients does not impact 5-year allograft survival. *Am J Transplant.* 2012;12:1924–8.
43. Kottgen A, Glazer NL, Dehghan A, et al. Multiple loci associated with indices of renal function and chronic kidney disease. *Nat Genet.* 2009;41:712–7.
44. • Ellis JW, Chen MH, Foster MC, et al. Validated SNPs for eGFR and their associations with albuminuria. *Hum Mol Genet.* 2012;21:3293–8. *Studies association of eGFR SNPs found from GWAS in independent population to examine their effect on incident albuminuria. Example of studying known risk loci in other related phenotype to better understand pathophysiology.*
45. Reznichenko A, Snieder H, van den Born J, et al. CUBN as a novel locus for end-stage renal disease: insights from renal transplantation. *PLoS ONE.* 2012;7:e36512.
46. Boger CA, Gorski M, Li M, et al. Association of eGFR-related loci identified by GWAS with incident CKD and ESRD. *PLoS Genet.* 2011;7:e1002292.
47. • O'Seaghdha CM, Yang Q, Wu H, et al. Performance of a genetic risk score for CKD stage 3 in the general population. *Am J Kidney Dis.* 2012;59:19–24. *Uses SNPs with established association of increased risk of CKD to create a risk prediction score for incident CKD and applied this to independent cohort. A model for future efforts to integrate genetic risk information with clinical data.*
48. • Chasman DI, Fuchsberger C, Pattaro C, et al. Integration of genome-wide association studies with biological knowledge identifies six novel genes related to kidney function. *Hum Mol Genet.* 2012;21(24):5329–43. *Integrates GWAS data with prior biological knowledge and bioinformatics methods to identify novel genes associated with kidney function. Demonstrates power of multilevel data synthesis that combines genetic associations and biologic knowledge to improve understanding of CKD.*
49. Pattaro C, Kottgen A, Teumer A, et al. Genome-wide association and functional follow-up reveals new loci for kidney function. *PLoS Genet.* 2012;8:e1002584.
50. Liu CT, Garnaas MK, Tin A, et al. Genetic association for renal traits among participants of African ancestry reveals new loci for renal function. *PLoS Genet.* 2011;7:e1002264.
51. Huang L, Jolly LA, Willis-Owen S, et al. A noncoding, regulatory mutation implicates HCFC1 in nonsyndromic intellectual disability. *Am J Hum Genet.* 2012;91(4):694–702.
52. Vernot B, Stergachis AB, Maurano MT, et al. Personal and population genomics of human regulatory variation. *Genome Res.* 2012;22:1689–97.
53. Gadebeku CA, Gipson DS, Holzman L, et al. Design of the nephrotic syndrome study network (NEPTUNE): a multi-disciplinary approach to understanding primary glomerular nephropathy. *Kid Int.* 2012.
54. Furth SL, Cole SR, Moxey-Mims M, et al. Design and methods of the Chronic Kidney Disease in children (CKiD) prospective cohort study. *Clin J Am Soc Nephrol.* 2006;1(5):1006–15.
55. • Maher B. ENCODE: the human encyclopaedia. *Nature.* 2012;489:46–48. *A brief but highly informative editorial that describes the development, key findings, and potential impact of the ENCODE project.*