Focus: Precision medicine in nephrology

Diagnoses of uncertain significance: kidney genetics in the 21st century

Daniel P. Gale[1], Andrew Mallett[2,3,4], Chirag Patel[2,5], Tam P. Sneddon[6], Heidi L. Rehm[7,8,9], Matthew G. Sampson[8,9,10*] and Detlef Bockenhauer[1,11*]

1. Department of Renal Medicine, University College London, London, UK
2. KidGen Collaborative, Australian Genomics Health Alliance, Australia
3. Department of Renal Medicine, Royal Brisbane and Women's Hospital, Brisbane, Australia
4. Institute for Molecular Bioscience and Faculty of Medicine, The University of Queensland, Brisbane, Australia
5. Genetic Health Queensland, Royal Brisbane and Women's Hospital, Brisbane, Australia
6. Department of Pathology and Laboratory Medicine, University of North Carolina, , Chapel Hill, NC, USA
7. Center for Genomic Medicine, Massachusetts General Hospital, Cambridge Street, Boston, USA
8. Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, USA
9. Harvard University School of Medicine, Boston, USA
10. Division of Nephrology, Boston Children's Hospital
11. Renal Unit, Great Ormond Street Hospital for Children, NHS Foundation Trust, London, UK

*E-mail: Matthew.Sampson@childrens.harvard.edu

*E-mail: d.bockenhauer@ucl.ac.uk

Standfirst

The increasing availability of sequencing has accelerated the discovery of genetic causes of kidney disease, with clear benefits for patients. However, insufficient or contradictory evidence exists for numerous variants that were previously reported to be pathogenic, calling into question some proposed gene–disease associations. Rigorous re-appraisal of evidence is needed to assure diagnostic accuracy.

Main commentary

Genetics has had a dramatic impact on modern medicine, including nephrology. Over the past 2–3 decades, hundreds of genes associated with kidney disease have been discovered, providing precise diagnoses and revealing fundamental insights into renal physiology and pathology. The rapid pace of kidney genomic discovery and its clinical implementation is exciting and holds great promise. However, concerns exist that the ability to generate genomic data from patients is outpacing our ability to interpret the data properly. If we are to effectively integrate genomic medicine into nephrology, accurate assessment of genes and variants is needed. Here, we aim to highlight the scale and scope of the challenge in kidney genomics and discuss strategies to improve diagnostic accuracy.

[H1] *Relevance of diagnostic accuracy*

A genetic diagnosis can end protracted diagnostic odysseys, enable targeted treatments, inform the risk of disease recurrence after transplantation and aid the assessment of potential donors. Conversely, an erroneous interpretation of a genetic test may lead to unsuitable treatment and can misinform risk of disease recurrence, risk to family members (including potential donors) and decisions with regard to family planning or pregnancy termination. In short, diagnostic accuracy of a genetic test is of utmost importance and proper evaluation of the results and methods is critical for patient care.

[H1] *Scope of diagnostic inaccuracy*

The problem of diagnostic inaccuracy is not restricted to nephrology. Over the past ten years, a number of public databases such as gnomAD (gnomad.broadinstitute.org) that aggregate large amounts of sequencing data have become publicly available. These resources provide insights into variant frequency within reference populations, enabling robust assessment of published disease-associated variants. Such assessments have revealed that many published

variants are unlikely to cause disease due to their excessive frequency in healthy controls and a lack of independent, robust evidence for pathogenicity. One review estimated that up to a quarter of published variants might be erroneously considered pathogenic.[1]

[H1] *Variant classification*

Professional societies, such as the American College of Medical Genetics (ACMG) and the Association for Molecular Pathology (AMP) have proposed the use of graded evidence to classify variants as benign, likely benign, of uncertain significance, likely pathogenic or pathogenic and computational tools are now available to support this approach.[2] Importantly, the absence of a variant from population databases alone constitutes insufficient evidence of pathogenicity. Moreover, judgement needs to be applied to some criteria, such as functional studies. Usually, these should be graded as supportive, since only well-validated clinical assays can provide strong evidence of causality.[3]

[H1] *Variants of uncertain significance*

Multiple evidence types, with at least one being strong or moderate, are needed to classify a variant as (likely) pathogenic. Most novel missense variants will be variants of uncertain significance (VUS), which is therefore a common outcome of genetic studies. Critical review of the phenotype is necessary to determine whether a particular variant fits the clinical findings. If the prior probability of a given monogenic disorder is low VUS are very unlikely to be disease-causing (Supplementary Figure 1). Evidence to prove variant pathogenicity is often insufficient, even when the gene matches the phenotype.[4]

Genes of uncertain significance

In some instances, application of graded evidence can call into question the entire association of a gene with a disease, so that it becomes a gene of uncertain significance. For example, a 2008 study identified 3 different heterozygous variants in *SLC9A3R1* (which encodes NHERF1) in 4 of 92 European individuals with nephrolithiasis or osteopenia,[5] with clinical and functional evidence consistent with a causative role for the variants in renal phosphate loss, leading to nephrolithiasis or osteopenia. However, analysis of population databases shows that the cumulative frequency of those 3 variants in Europeans is 5.2%. Not only is this frequency essentially the same as that in the study group (4.4%), but if these 3 variants were indeed disease causative, the implication is that *SLC9A3R1*-related disease would be the most common genetic kidney disorder, exceeding ADPKD by approximately 50-fold.

Despite these observations, which call into question a pathogenic role of *SLC9A3R1* variants, these are still described by OMIM (#604990) and by several published papers as a genetic cause of kidney disease, highlighting the need to critically review published evidence to aid variant interpretation

Small study sizes also pose a problem when assessing the causality of identified variants. Whilst the number of reported disease genes has increased dramatically in recent years, each novel disease gene identified tends to be causative in ever-smaller subsets of patients. Consequently, reporting numbers can be as low as a single patient, increasing the likelihood of an incidental association.[6]

[H1] *Defining genetic architecture*

The likelihood of a genetic test correctly revealing the molecular diagnosis is critically dependent on the prevalence of detectable monogenic disorders in the test population — that is, the prior probability of individuals within the test population having an identifiable, disease-causing variant. In situations in which similar clinical presentations can result from monogenic or non-monogenic disorders (for example, nephrotic syndrome or C3 glomerulopathy), it is vitally important that accurate data are available to allow reliable estimation of the prior probability. Two methodological issues commonly contribute to overestimation of the prior probability. The first is that studies often enrol patients who have been highly selected based on their likelihood of having a monogenic disease (for example, based on family history, parental consanguinity and/or syndromic features). The second involves the misclassification of variants as pathogenic in cohort studies on which the estimate is based. Such misclassification is a particular issue in older genetic studies, conducted before the ACMG–AMP criteria were formulated and/or prior to the availability of large-scale population sequencing data. Yet despite the availability of these resources, misclassified variants remain a problem.

Another example of potential misunderstanding of genetic architecture concerns recessive disorders that can be caused by a number of genes, such as Bartter or nephrotic syndromes. Increasingly, digenic or oligogenic inheritance of variants has been proposed, whereby the disease is supposedly caused by heterozygous variants in different causative genes. Such reports seem to be emerging for a couple of reasons. First, owing to the ability to sequence many genes simultaneously and second because of the temptation to assign causality to single variants across multiple recessive disease genes implicated in a particular disease. However, solid evidence from controlled studies that recessive kidney disorders can be caused by such

combination of heterozygous variants is lacking. A key problem with many such studies is that they consider cases only and do not assess the frequency of individually rare, heterozygous variants ascertained comparably in ethnically matched controls. Unfortunately, public sequencing databases aggregate variants on a population level, thereby preventing comparisons of the burden of rare variants at the individual level.

[H1] *Drivers of inaccuracy*

The understandable desire of researchers and clinicians to provide a definitive molecular diagnosis for patients potentially increases the risk of a well-recognised statistical reasoning error termed the 'prosecutor's fallacy'.[7] In the context of genetic testing, this error would arise from an individual researcher observing a suspicious variant in a patient with a rare disease and deciding intuitively that it constitutes sufficient evidence for causation, even if the formal ACMG–AMP criteria for pathogenicity are not met. In addition, experts for a particular disease or gene will be obvious and perhaps enthusiastic reviewers for a paper suggesting an association involving their gene or disease of interest, but their expertise in the application of general gene-disease evidence review may be limited. Further contributors might include the problems of publication bias and a commercial-academic reward system that favours high diagnostic yields and the discovery of novel disease-causing genes.

[H1] *A way forward*

A number of steps can be implemented to improve the accuracy of attribution of pathogenicity to genes and variants. Any study that identifies a pathogenic variant must ensure the following. Firstly, that newly proposed gene–disease associations are appropriately annotated with the level of evidence supporting the claim,[8] and that the assessment criteria applied to variants in well-established genes are explicitly stated. Second, sufficient details about patients sequenced are provided to enable assessment of clinical diagnosis and whether the cohort is enriched for patients with monogenic disease. Lastly, that appropriate controls with matched ancestry and sequencing pipelines are included to allow unbiased comparisons of variant burden. We ask editors and reviewers of relevant publications to also ensure these steps have been followed. The ability of such an approach to reveal insights into the underlying genetic architecture of primary membranoproliferative glomerulonephritis — a disorder that was frequently considered to be a monogenic disease — has now been highlighted.[9]

In addition, increased collaborative efforts are needed. An example is the Clinical Genome Resource (ClinGen; clinicalgenome.org), which was founded to provide a central resource that defines the clinical relevance of genes and variants for use in precision medicine and research.[10] We are forming a series of ClinGen Expert Panels for kidney disorders to enable a reappraisal of published genes and variants. Such an effort is critically needed to establish a foundation for the genomic diagnosis of patients with suspected genetic kidney disorders.

References

1 MacArthur, D. G. *et al. Nature* **508**, 469-476, (2014).
2 Richards, S. *et al. Genet. Med.* **17**, 405-424 (2015).
3 Brnich, S. E. *et al. Genome Med* **12**, 3 (2019).
4 Sampson, M. G. *et al. J. Am Soc. Nephrol.* **27**, 1970-1983 (2016).
5 Karim, Z. *et al. N. Engl. J. Med.* **359**, 1128-1135, (2008).
6 Ashton, E. & Bockenhauer, D. *Kidney Int.* **97**, 455-457, (2020).
7 Thompson, W. C. & Schumann, E. L. *Law Human Behav* **11**, 167-187 (1987).
8 Strande, N. T. *et al. Am J Hum Genet* **100**, 895-906 (2017).
9 Levine, A. P. *et al. J. Am Soc. Nephrol. JASN*, doi:10.1681/ASN.2019040433 (2020).
10 Rehm, H. L. *et al. N. Engl. J. Med.* **372**, 2235-2242 (2015).

Competing interests

The authors declare no competing interests.

Related links:

gnomAD: gnomad.broadinstitute.org

Clinical Genome Resource: clinicalgenome.org