Nephrogenomics

https://doi.org/10.1038/s41581-025-00952-3

GWAS scorecard prioritizes kidney genes using coding and regulatory variants

Matthias Wuttke & Cristian Pattaro

Based on genome-wide association study data from 2.2 million individuals, a functional prioritization scorecard integrates classical omics with allele-specific gene expression, chromatin accessibility and gene regulatory circuits in kidney tissues and specific cell types. This approach prioritized 601 kidney function-related genes, including genes associated with nephropathy in mice and humans.

REFERS TO Liu, H. et al. Kidney multiome-based genetic scorecard reveals convergent coding and regulatory variants. *Science* **387**, eadp4753 (2025).

Over the past 20 years, genome-wide association studies (GWAS) have proven to be powerful agnostic tools for identifying genetic variants linked to chronic kidney disease (CKD)[1]. As GWAS sample sizes have increased, the number of associated genetic loci has grown to several hundreds, which has shifted the focus from mere discovery to the characterization of the implicated genes and the mechanisms that link them to kidney function. Recently, Liu et al.[2] pooled GWAS data from 2.2 million individuals and created a kidney disease genetic scorecard that integrates classical omics with allele-specific data, such as allele-specific gene expression (ASE) and allele-specific chromatin accessibility, and gene regulatory circuits.

CKD occurs across a broad allele frequency spectrum, and both common and rare genetic variants contribute to its development[3]. Common variants are typically studied using GWAS, whereas rare variations are investigated through sequencing technologies. As sample sizes increase, the power for imputation and the detection of less frequent alleles improves[4], although rare mutations remain difficult to detect in GWAS[5]. Fine mapping, particularly with multi-ancestry datasets, enhances precision. Multi-omics approaches integrate genomic, epigenomic and transcriptomic data to provide a comprehensive view of biological processes and clarify how genetic variants translate into changes in biological function.

Liu et al. identified 1,026 genetic loci associated with estimated glomerular filtration rate based on serum creatinine (eGFRcrea), 87% of which were validated as being related to kidney function rather than reflecting creatinine metabolism, on the basis of consistent association with cystatin C-based eGFR or blood urea nitrogen. Of note, although the pace of newly uncovered signals in people of European ancestry seems to be slowing down (owing to the current large sample size), GWAS insights from people of other ancestries are still in their infancy and most continents are still severely underrepresented. As different populations have distinct genetic architectures, allele frequencies and linkage disequilibrium patterns, integration of diverse ancestry groups can reveal novel associations and improve fine mapping of causal variants. The improvement of signal detection obtained by integrating GWAS from three different ancestries is a call for the further expansion of GWAS efforts to underrepresented groups. By harnessing the power of large sample sizes combined with advanced Bayesian fine mapping, the researchers demonstrated the ability to distinguish independent signals even within small genomic loci. This ability is exemplified by transportin 3 (*TNPO3*), a novel gene implicated in kidney function that is adjacent to interferon-regulatory factor 5 (*IRF5*), which has been previously associated with eGFRcrea[6] and CKD incidence[7].

> "Integration of diverse ancestry groups can reveal novel associations and improve fine mapping of causal variants"

To better understand the functional implications of eGFRcrea-associated variants, the researchers analysed ASE in glomerular and tubular kidney tissues. ASE refers to differential gene expression between the maternal and paternal alleles within the same individual, which means that two copies of the same gene might not result in the same level of gene transcription owing to differences in genomic sequences or epigenetic effects. Their analysis of kidney tissues uncovered 10,398 genes with allele-specific expression, 17% of which were previously undetected by traditional expression quantitative trait locus (QTL) studies. Bulk assay for transposase-accessible chromatin (ATAC)-based allele-specific accessibility (bASA) identified ~500,000 variants associated with chromatin accessibility. Eighteen per cent of bASA loci did not overlap with DNA methylation or ASE, which suggests that other regulatory mechanisms might be influencing expression in these loci. Single-nuclei ATAC sequencing and allele-specific accessibility (snASA) mapped 25,766 variants, of which 86% showed cell-type specificity. Notably, 669 snASA variants were significantly associated with eGFR, of which 63% disrupted transcription factor binding site motifs. For example, *HNF4A* and *HNF4G* motifs showed cell type-specific disruption in the proximal tubule. *HNF4A* regulates genes that are essential for solute transport and metabolism directly, and thereby maintains epithelial integrity and facilitates proper reabsorption processes in the proximal tubule[8]. Combining single-nucleus RNA and ATAC sequencing, the researchers created a joint atlas of gene expression and chromatin accessibility, which links 7,137 eGFRcrea-associated variants to the expression of 1,351 target genes.

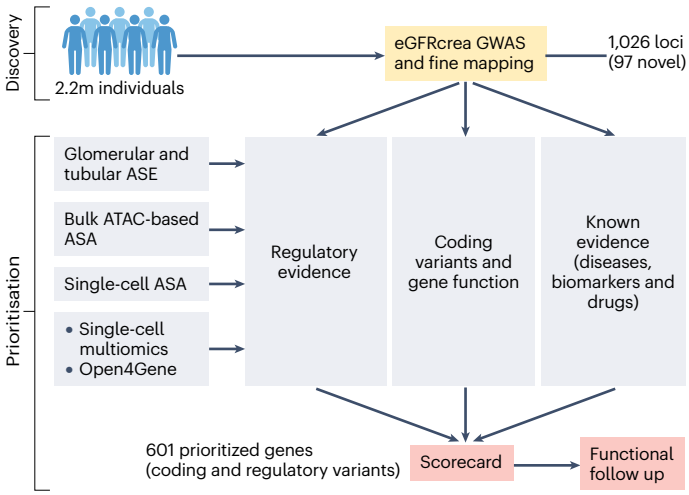

Fig. 1 | Kidney Disease Genetic Scorecard workflow. Fine-mapping genome-wide association study (GWAS) data for estimated filtration rate based on creatinine (eGFRcrea) from 2.2 million (2.2m) individuals identified 1,026 associated loci. Associated variants were annotated on the basis of regulatory evidence, coding variant function and existing knowledge on diseases, biomarkers and drugs to prioritize genes with a potential causal role in kidney function. ASA, allele-specific chromatin accessibility; ASE, allele-specific gene expression; ATAC, assay for transposase-accessible chromatin.

Liu et al. also introduced the 'Kidney Disease Genetic Scorecard', a tool that attributes scores to regulatory variants, common coding variants and target genes. The score reflects the sum of all instances in which the genomic feature of a variant or gene (for example, its expression) coincides with an eGFRcrea genomic signal. The approach integrates existing datasets from expression QTLs, methylation QTLs, exome-wide association studies and complementary trait GWAS. The scorecard was instrumental for the prioritization of 601 genes influenced by both common coding and regulatory variants (Fig. 1): 161 genes showed convergence of coding and regulatory variants, which robustly supports genetic influence on kidney function, and 124 genes are targets of FDA-approved drugs.

This study is one of the largest and most comprehensive investigations of kidney function genomics to date, and has a strong focus on transcriptomics and open chromatin. Integrating single-cell multi-omics and ASE provided key insights into cell-type-specific regulatory elements and transcription factors, which advances our understanding of kidney function and disease mechanisms. Beyond genomics and epigenomics, the main novelty lies in their refined transcriptomic analysis, which demonstrated the convergence of common coding and regulatory variants in key kidney disease genes. The Kidney Disease Genetic Scorecard allowed for systematic prioritization of genetic variants, and effectively synthesized multiple sources of evidence into a clear framework. This structured approach enhances the ability to identify disease-related genes and their mechanisms, which paves the way for novel therapeutic insights and drug repurposing. An example of a gene with convergent evidence from multiple sources is ubiquitin specific peptidase 24 (*USP24*), which is implicated through shared evidence across open chromatin, CpG methylation and gene expression (which supports previous evidence of a role in kidney function[9]). *USP24* is a deubiquitinating enzyme that regulates protein turnover, cellular stress responses and inflammatory pathways, which are crucial processes for tubular injury repair and immune-mediated kidney damage.

> **"Refined transcriptomic analysis ... demonstrated the convergence of common coding and regulatory variants"**

The increasing number of genes previously known for specific kidney conditions emerging through such large-scale GWAS integrated with multiple omics layers is notable, and confirms their relevance in kidney disease. Also of interest was the observation that >60% of prioritized regulatory variants were intronic and nearly 30% of them coincided with transcription factor binding sites, which highlights the role of the noncoding genome. The results of this study are even more exceptional when considering that the focus was on common autosomal variants. Bridging the gap with whole-exome and whole-genome sequencing studies that capture both common and rare variants within a single framework is the next step. For example, Liu et al. were the first to identify common kidney disease variants in *PKD2*, yet previous whole-exome sequencing studies had already implicated rare variants[4]. This discrepancy underscores the importance of integrating different sequencing methods to achieve comprehensive variant discovery.

The researchers are to be commended for the outstanding effort to process over 700 microdissected human kidneys for transcriptional and open chromatin characterization. Understanding how genetic variants translate into protein-level changes by using proteomics might further enhance our ability to pinpoint therapeutic targets based on transcriptomic regulation insights, and facilitate clinical translation.

Matthias Wuttke [1] ✉ **& Cristian Pattaro** [2]
[1]Institute of Genetic Epidemiology, Medical Center – University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany. [2]Eurac Research, Institute for Biomedicine, Bolzano, Italy.
✉e-mail: matthias.wuttke@uniklinik-freiburg.de

Published online: 20 March 2025

References

1. Köttgen, A. & Pattaro, C. The CKDGen Consortium: ten years of insights into the genetic basis of kidney function. *Kidney Int.* **97**, 236–242 (2020).
2. Liu, H. et al. Kidney multiome-based genetic scorecard reveals convergent coding and regulatory variants. *Science* **387**, eadp4753 (2025).
3. Wuttke, M. & Köttgen, A. Insights into kidney diseases from genome-wide association studies. *Nat. Rev. Nephrol.* **12**, 549–562 (2016).
4. Wuttke, M. et al. Imputation-powered whole-exome analysis identifies genes associated with kidney function and disease in the UK Biobank. *Nat. Commun.* **14**, 1287 (2023).
5. Auer, P. L. & Lettre, G. Rare variant association studies: considerations, challenges and opportunities. *Genome Med.* **7**, 16 (2015).
6. Wuttke, M. et al. A catalog of genetic loci associated with kidney function from analyses of a million individuals. *Nat. Genet.* **51**, 957–972 (2019).
7. Fumeron, F. et al. Genetic variants of interferon-response factor 5 are associated with the incidence of chronic kidney disease: the D.E.S.I.R. study. *Genes Immun.* **24**, 303–308 (2023).
8. Marable, S. S., Chung, E. & Park, J.-S. Hnf4a is required for the development of Cdh6-expressing progenitors into proximal tubules in the mouse kidney. *J. Am. Soc. Nephrol.* **31**, 2543–2558 (2020).
9. Hattori, K., Tsubota, S., Hagita, J., Doke, T. & Maruyama, S. Transcriptome-wide association study identified USP24 as a kidney disease risk gene: FR-PO986. *J. Am. Soc. Nephrol.* **34**, 680 (2023).

Competing interests

The authors declare no competing interests.