Published in final edited form as:

Semin Nephrol. 2018 January ; 38(1): 40–51. doi:10.1016/j.semnephrol.2017.09.006.

Precision Medicine for Acute Kidney Injury (AKI): Redefining AKI by Agnostic Kidney Tissue Interrogation and Genetics

Krzysztof Kiryluk¹, Andrew Bomback¹, Yim-Ling Cheng², Pablo G. Camara², Raul Rabadan², Peter Sims^{2,3}, and Jonathan Barasch¹

¹Department of Medicine, Division of Nephrology, College of Physicians & Surgeons, Columbia University, New York, NY, USA

²Department of Systems Biology, College of Physicians & Surgeons, Columbia University, New York, NY, USA

³Department of Biochemistry and Molecular Biophysics, College of Physicians & Surgeons, Columbia University, New York, NY, USA

Abstract

Acute kidney injury (AKI) is currently diagnosed by a temporal trend of a single blood analyte, the serum creatinine. This measurement is neither sensitive nor specific to kidney injury or its protean forms. Newer biomarkers, neutrophil gelatinase-associated lipocalin (NGAL, Lcn2, Siderocalin) or kidney injury molecule-1 (KIM-1, HAVCR1), accelerate the diagnosis of AKI as well as prospectively distinguish rapidly reversible from prolonged causes of serum creatinine elevation. Nonetheless, these biomarkers lack the capacity to further sub-fractionate AKI (e.g. sepsis vs ischemia vs nephrotoxicity from medications, enzymes or metals) or inform us about the primary and the secondary sites of injury. It is also unknown whether all nephrons are injured in AKI, whether all cells in a nephron are affected, and whether injury responses can be stimulus-specific or cell type-specific or both. In this review, we summarize fully agnostic tissue interrogation approaches that may help to redefine AKI in cellular and molecular terms, including single cell and single nuclei RNA sequencing technology. These approaches will empower a shift in the current paradigm of AKI diagnosis, classification and staging, and provide the renal community with a significant advance towards precision medicine in the analysis AKI.

Corresponding Authors: Krzysztof Kiryluk, MD, MS, Department of Medicine, Division of Nephrology, College of Physicians and Surgeons, Columbia University, 1150 St. Nicholas Avenue, Russ Berrie Pavillion #412E, New York, NY 10032, USA. Phone: +1 212 851 5559, kk473@columbia.edu; Jonathan Barasch, MD, PhD, Department of Medicine, Division of Nephrology, College of Physicians and Surgeons, Columbia University, 1150 St. Nicholas Avenue, Russ Berrie Pavillion #413, New York, NY 10032, USA. Phone: +1 917-671 6081, mb4@columbia.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Financial disclosure and conflict of interest statements

Introduction

Classical Nephrology taught that kidney failure should be described on the basis of the etiology and the anatomy of the injury. We believed that identifying the site of injury and the mechanism of dysfunction should inform the correct treatment. For example, a decrease in renal excretion due to the activation of homeostatic mechanisms would be met with fluid therapy, a decrease in renal excretion due to stimuli that directly impact epithelial integrity might require dialysis, while the blockade of the urinary outflow would be met with a urologic intervention. These categories were useful because they predicted the clinical course and suggested appropriate therapeutic interventions. In essence, the founding ideas of clinical Nephrology were very close to what we aspire to in the modern era of personalized medicine – the diagnosis of kidney dysfunction on a patient-specific level in order to implement injury-specific therapies.

Presently, the modern practice of Nephrology uses a single analyte to identify a myriad of injuries to the organ and its related phenomenon. Accordingly, it is believed that any acute rise in this analyte, the serum creatinine (sCr), reflects some degree of injury to the kidney tubule. In fact, published guidelines (RIFLE,¹ AKIN,² or KDIGO³) state that even modest elevations in sCr reflect the initial phase of a continuous pathogenesis of kidney damage. These concepts are encapsulated in the diagnostic acronym “AKI” (Acute Kidney Injury), which is measured by increasing levels of sCr “stages” that are indifferent to the specific pathophysiology or the specific anatomy of the injury.

The definition of kidney dysfunction by the AKI paradigm has allowed for the facile analysis of population based datasets. Further, there is strong evidence that AKI, defined by increases in sCr represents a serious public health problem: the incidence continues to increase among older hospitalized adults,^{1,4,5} and the diagnosis is associated with significant morbidity and mortality,⁶ including increased risk of chronic kidney disease (CKD).⁷ The percentage of patients with an AKI hospitalization in the Medicare fee-for-service population has risen over the past decade, reaching 3.9% in 2013, compared to 1.5% in 2003 (n=51,909 for 2013). Moreover, AKI has gained increasing recognition as a major risk factor for the development of CKD. For example, acute tubular necrosis without recovery is the primary diagnosis for 2–3% of incident end-stage renal disease (ESRD) cases annually. This, however, represents only a small fraction of the renal disease burden resulting from AKI, as studies have demonstrated significantly increased long-term risk of CKD and ESRD following AKI, even after initial recovery of renal function.⁸ Furthermore, this relationship is bi-directional, with CKD patients at substantially greater risk of suffering an episode of AKI.⁹ As a result, AKI superimposed on CKD plays an important role in CKD progression.

While “AKI” definitions have resulted in critical concepts applicable to populations of patients, the terminology implies that there is a single common pathway of injury stratified by sCr “stages”. However, this notion is at variance with the striking lack of correlation between sCr and kidney pathology suggesting that on a cellular level more than a single pathway is involved in acute injury, and in different individuals different pathways may be activated^{10–12}. Indeed, in a study of nearly 3000 Emergency Department patients, the rise in sCr was due to disparate causes, including fully reversible volume depletion in over 15% of

all comers and true intrinsic AKI (tubular damage, prolonged azotemia) in only 5.8% of patients, despite elevation of sCr in both cases¹³. Rapidly reversible elevation of sCr constituted 72% of all enrolled patients with acute changes in sCr and 50% of all cases of elevated sCr. In a prior study, we found rapidly reversible changes in sCr in 14% of all comers and intrinsic AKI in 4.7% meaning that rapidly reversible sCr was the dominant finding (40% of enrolled patients¹⁴). In another study which evaluated 3.8 million patients in the NY Presbyterian Medical system, most of the 25,000 episodes of creatinine elevation (75%) were brief (<72hrs) regardless of the patients' age¹⁵.

Consequently, the initial diagnosis of “injury” broadly applied to individuals who are treated by distinct therapeutic regimens and have entirely different hospital courses is a tentative diagnosis which may be erroneous. In summary, while a single diagnostic category “AKI” has simplified the clinical workflow, including physician service billing (ICD-10), a spot check of serum creatinine does not accommodate the diversity of the condition, intuit the clinical course from the Emergency Department until patient discharge, predict effective therapeutic interventions (e.g. fluid therapy vs dialysis), or help guide us in identifying the specific injured cell.

Rather than a singular readout, recent data demonstrates that gene expression in the kidney and protein signatures in the urine can distinguish different forms of AKI implying a complex system responding to different environmental threats. In fact, some proteins have served as markers of AKI subtypes in different clinical studies. For example, the diagnostic molecules neutrophil gelatinase-associated lipocalin (NGAL), kidney injury molecule-1 (KIM-1), and hundreds of other proteins are activated by tubular damage but not by volume depletion per se, whereas others may have the potential to distinguish extracellular volume depletion from tubular injury (Figures 1 and 2)¹⁵. These data indicate that while the current generation of biomarkers have laid the ground work for individualized diagnosis, further refinement is needed to understand the response of individual cell types in the nephron particularly because the current injury biomarkers are expressed in response to the same signal. Hence, it is now clear that the entire nephron responds to environmental threats, and the next frontier in AKI research must be to find increasingly specific responses of each nephron segment. Here, we discuss the major problems with our current research approaches, highlight the key scientific questions, and propose application of several new agnostic approaches that may finally rectify our current failure to diagnose and understand kidney damage in “AKI”.

The Key Problems in the Field of AKI

Given the above considerations, we have identified several important problems in the clinical and research field of AKI that stand in the way of progress towards precision diagnosis and therapy. The solutions to these problems are required to move the field forward and guide the design of future clinical studies. The first problem relates to the use of serum creatinine (sCr) for diagnosis of AKI¹⁶. Imagine presenting to the Emergency Department, and the physician informs you that you have “AKI”. The underlying diagnosis may be volume depletion “AKI” or sepsis-associated “AKI” or arterial ischemia “AKI” or nephrotoxic “AKI”, yet the physician cannot affirmatively diagnose any of these entities with available

tools at the time of patient encounter^{13,14,17,18}. The principal reason for such failure is that the physician's diagnosis currently depends on a single test, sCr, which is most easily interpretable when its dynamic changes over an unspecified time interval are viewed in retrospect. In addition, interpretation depends on many variables such as muscle mass, nutrition, extracellular volume^{19–23}. Moreover, sCr begins to rise not only in the setting of kidney damage, but in a number of adaptive physiological states, most notably in volume depletion. As a result, our “gold standard” lacks intrinsic characteristics of an appropriate biomarker for kidney injury – it has a high rate of false negatives (e.g. 67.2% of patients in PICU²⁴) and false positives (30% of cases^{25,26}). Not surprisingly, our current ability to treat AKI empirically, based on fluctuations in creatinine, pales in comparison to the targeted approaches of treating AKI when a renal biopsy is utilized to identify a specific diagnosis (cf. a “suspected” case of allergic interstitial nephritis vs. a biopsy-confirmed case of allergic interstitial nephritis). Evidently, we need improved tools to diagnose AKI in order to advance the precision medicine approach.

The second problem relates to the common misinterpretation of biomarkers. The “second generation” of AKI biomarkers, such as NGAL or KIM-1, have partially rectified the inadequacies of sCr, but their interpretation in conjunction with sCr represents a vexing problem that is critical to kidney precision medicine and regulatory agencies such as the Food and Drug Administration (FDA)¹⁶. First, biomarkers and sCr may dissociate at a lower “stages” as defined by RIFLE,¹ AKIN,² or KDIGO³, a problem that has caused much consternation to the regulatory agencies. However, while elevated sCr measures the degree of “organ failure”, biomarkers measure responses at the cellular level. Hence, when damage is not severe enough to cause organ failure, biomarkers are expressed but sCr may be little changed, analogous to a cardiac troponin in the early detection of acute myocardial infarction when myocyte damage is not widespread or strategically positioned^{27,28}. The second critical issue is that two distinctive stimuli, volume depletion and tubular necrosis, induce completely different sets of molecular responses at the same level of sCr²⁹. Hence, a biomarker of tubular injury should not be compared to sCr that has risen as a result of volume depletion^{30,31}. These discrepancies have generated new insights into the diversity of the responses of the nephron to different environmental challenges, and as a result, informing the biomedical community that the only solution to these apparent conundrums is to redefine AKI in molecular terms.

The third problem is due to the unknown anatomy of injury in AKI. Beyond the inadequacies of sCr, the complexity of the human kidney provides additional challenges to revising the AKI paradigm. Remarkably, it is still not known (1) whether all nephrons respond to an injurious stimulus, a question of interest because sCr does not rise until approximately 50% renal mass (the so-called “renal reserve”) is lost, suggesting that at least some nephrons have compensatory but saturable responses; (2) whether all cells in a nephron are affected by a given injury or whether only specific cells are targeted; and (3) whether responses are stimulus-specific or common to all forms of injury. The experience with two biomarkers, NGAL produced in the TALH and the Collecting Duct and KIM-1 in the proximal tubule with different timing after an injurious stimulus favors the idea that cells within different nephron segments can respond to the same stimulus, but by different routes of injury. Hence, we anticipate that co-expressed molecules with greater specificity for

specific forms of damage may be discovered. Consequently, the existing biomarker data supports the hypothesis that kidney damage can be measured by induced expression of segment-specific genes. Yet, our current insights are still insufficient to reach the next step in kidney diagnosis. This is because we must account for greater complexity than a simple segmented structure, since (i) gradients of gene expression are found in both proximal (e.g. megalin) and distal convoluted tubules (e.g. Ca^{2+} transport proteins), while (ii) in other segments some proteins appear “prematurely” (e.g. ENaC in distal convoluted tubules³²) or (iii) a spectrum of cell types overlap in identity (e.g. collecting ducts³³) and (iv) different segments of the nephron are likely to signal one another³⁴ activating secondary events. Consequently, molecular readouts will be complex and form unexpected transcriptomic patterns. Also, given that “AKI” might shift cellular expression patterns^{35–38}, we suggest that a molecular definition of AKI will arise only if we first deconstruct the complexity of the kidney, beginning with maps of the healthy kidney tissue.

The “next generation” tissue interrogation methods enabling a “molecular diagnosis” of AKI may hold promise to address some of these historical problems. From biomarker work published to date, we know that different segments of the nephron respond to injury (Figure 3), but the full anatomy of this response is unknown. Understanding the anatomy of injury will provide a rational framework for the interpretation of molecular responses to a myriad of environmental challenges. Hence, rather than focus on a specific growth factor or cellular pathway, one needs to apply an agnostic approach to discern the patterning of gene expression throughout the entire nephron. However, to create such a map in humans, one must overcome numerous challenges, such as the heterogeneity of injurious insults, inter-individual differences in response to injury, and technical and analytical problems of working with minuscule pieces of highly heterogeneous and anatomically complex human biopsy tissue. We explore these challenges and novel approaches that aim to tackle such challenges in the remainder of this review.

What “next generation” methods are best suited to define the anatomy of injury in AKI?

The classic immunohistological techniques developed over the past 50 years have enabled the characterization of the structural organization of tissues at a molecular level, under both normal and pathological conditions^{39,40}. Although these techniques have had an enormous impact in the field of kidney disease, they have been limited by the fact that only a handful of markers can be simultaneously assessed. More recently, high throughput transcriptional mapping techniques for kidney tissues have been developed in mice using *in situ* RNA technology⁴¹, but similar approaches are needed for human tissues. We and others explored laser capture microscopy²⁹, worked with MARIS⁴², studied *in situ* proteomic techniques, and analyzed cytof assays⁴³ in human tissue, but these techniques can be laborious, are limited by technical issues of cell dissociation, or are only as good as predetermined data and reagents.

The recent development of next-generation single-cell transcriptomics has provided a way to interrogate the expression level of thousands of genes in thousands of individual cells

simultaneously^{44–47}. These methods allow for more comprehensive, unsupervised characterizations of the cellular components from a small tissue sample and their transcriptional profiles under various conditions. However, these methods come with new analytical challenges that has motivated the development of tailored computational approaches^{45,46}. Most challenges come from the large technical variation inherent to single cell sequencing experiments. Differences in the starting quantities of RNA in single cell libraries and imperfections in the capture, reverse transcription, and amplification of the RNA lead to numerous ‘dropouts’ (transcripts that are expressed but not detected) and biases in the expression levels of detected genes.

Accounting for these effects is therefore crucial when assessing biological variability of gene expression across the cells in the sample. In addition, the sparseness of single-cell RNA-seq data (the number of cells in these experiments is typically smaller or comparable to the number of detected genes) requires reducing the dimensionality of the dataset across non-informative directions before performing any statistical inference. Moreover, the scarcity and high dimensionality of single cell transcriptomic data requires the use of tailored mathematical approaches for analysis. Currently available methods are largely based on dimensional reduction followed by clustering and differential expression analysis, but these methods can discern relatively static, highly differentiated cell types rather than the transcriptional logic underlying more dynamic systems, such as tissue injury. In these situations, the continuous, branched, and asynchronous nature of injury limits the application of standard dimensional reduction and clustering methods.

The application of single cell RNA sequencing (sc-RNA-seq) to kidney tissue may be able to highlight the complex patterning of nephrons, but isolating single kidney cells, especially from frozen archival tissues, is currently still problematic. Indeed, most scRNA-seq protocols require fresh material, which is challenging for large-scale, patient biopsy-based studies. Yet, such studies should be the gold standard for AKI research given the above-discussed limitations of creatinine-based definitions. Another challenge arises in tissue with complex architecture, where harsh tissue dissociation to release single cells can lead to transcriptional perturbations and various cell type selection biases. Recent advances include the discovery of enzymes that are active at low temperature which help dissociate the kidney into single cells (Steven Potter, see URL). These protocols obviate the heat stress employed in standard enzymatic dissociation, hence promoting the identification of the native transcriptome of single cells. In addition, the availability of benchtop equipment that sheers tissue fragments in a reproducible fashion has contributed to the advances in single cell analysis.

An additional approach is helpful, particularly for the analysis of archival tissues or in the setting of complex tissue architecture. These obstacles are being overcome by the introduction of methods that assay single nuclei, instead of single cells. This alternative approach takes advantage of the mRNA contained in the nucleus of the cell and consequently avoids the need for cell isolation protocols that are likely to select against low abundance cell types in complex tissues in addition to perturbing gene expression. In contrast to single cell isolation, nuclear isolation protocols are straightforward because nuclei can be isolated with a simple Dounce homogenizer and an Eppendorf desktop

centrifuge. Most importantly, they can accommodate frozen tissues, enabling analysis of archived frozen biopsy material and existing tissue banks that could markedly accelerate the pace of research in this field. Recent comparisons of nuclear and cellular transcriptomes demonstrated that nuclei can substitute for whole cells in most RNA-seq applications, and, for the majority of genes, nuclei yielded expression signatures that were very similar to those obtained from whole-cell controls^{48–52}. For example, Grindberg demonstrated that while single nuclei generated 10–1000 fold less RNA (depending on the gene) than whole cell analysis, similar transcript numbers were detected and there was 98% concordance. Zeng found that 81.3% of genes showed no significant difference between single cell and single nucleus data (FDR = 0.001).

In summary, this technique provides the ability to harvest individual nuclear transcriptomes using fresh, frozen, and archival biopsy cores irrespective of changes in cytoplasmic dynamics. Typically, nuclei are isolated, separated using a variety of techniques, and sequenced using NextGen technology. For example, new microfluidic systems developed by Dr. Sims can distribute individual nuclei in micro-scale wells, capture RNA in solid phase by size-matched polymer beads, and by using a sequence-based barcoding scheme, generate a pooled cDNA library from thousands of individual nuclei at a relatively low cost^{53,54}. We have already successfully applied this approach to healthy human kidney tissue, allowing us to examine transcriptional profiles of individual cells that can be grouped into homogenous clusters of cell types (Figure 4), that can then be mapped back to specific nephron segments. The information generated using this technique will clearly allow for high resolution mapping of cell-specific molecular responses in the injured kidney.

How can we anatomically map specific cell types during evolving tissue injury in humans?

Because of considerable complexity of kidney tissue, anatomical mapping of specific cell types to specific nephron segments, particularly in the setting of evolving tissue injury, represents a major challenge. In the past few years, many algorithms have been proposed for the classification of cells into discrete types using scRNA-seq data with different mathematical approaches. For example, Seurat⁵⁵ implements density clustering in a low-dimensional t-distributed stochastic neighbor embedding⁵⁶ (t-SNE) representation of the data and allows for the integration of in situ hybridization data to reconstruct the single-cell expression data in the context of the three dimensional architecture of the organ. BackSPIN⁵⁷ is a divisive bi-clustering algorithm that clusters cells by selecting exclusive groups of highly-correlated genes within each cluster. PhenoGraph⁵⁸ clusters cells in a k-nearest neighbor representation of the data using the Louvain community detection method⁵⁹. RaceID2^{60,61} seeks the identification of rare cell types by leveraging the use of a multi-step k-medoids clustering approach where outlier cells are identified and clustered apart. Reference component analysis⁶² (RCA) uses a global reference panel of bulk transcriptomes to cluster individual cells into different types according to their transcriptional profile. Several studies have now demonstrated the utility of these algorithms for cell type classification based on RNA-seq data of complex tissue samples^{57,61–63}. These algorithms, however, have limited power to dissect the transcriptional diversity of each

specific cell type. The utility of these methods is therefore reduced in dynamic systems, where cells may form continuous mathematical structures according to their changing transcriptional profiles (for instance, due to ongoing cellular differentiation processes) rather than necessarily forming well-defined clusters. These limitations are particularly apparent in the setting of inflammation or tissue regeneration, both highly relevant to AKI.

The problem of characterizing dynamic cellular systems with single-cell resolution has also emerged in the context of developmental processes. In this context, several approaches have been developed that attempt to order cells according to their transcriptional profile, establishing a developmental pseudo-time^{64–67}. These methods are usually tailored to the analysis of developmental trajectories and cannot be applied to the analysis of more general dynamical systems such as those involved in injury. For example, algorithms for the analysis of developmental trajectories typically assume tree-like differentiation processes with a unique origin (a single stem cell type) and cannot cope with complex relationships such as the de-differentiation of a cellular subpopulation after injury or the presence of independent lineages.

Although more work is required to overcome these limitations, there are additional unbiased algorithms that may be useful for these more general situations. A new method, based on Topological Data Analysis (TDA) has been developed by Drs. Rabadan and Camara to specifically address complex relationships in cell populations. TDA is a nascent branch of mathematics directed towards studies of the continuous structure of high-dimensional data sets. The algorithm scTDA⁶⁸ builds upon topological representations⁶⁹ and statistics to identify and characterize differentially expressed genes without the need of pre-defining cellular populations. These genes are then used to dissect the transcriptional events underlying a dynamic cellular process without assuming any specific structure. Similarly, Destiny⁷⁰ uses diffusion maps⁷¹ to identify combinations of genes associated with the dynamics of a cellular process in an unbiased manner. Hence, these methods may be used to dissect dynamic transcriptional patterns in tissue samples. Similar approaches may provide an agnostic solution to address the complexity of evolving kidney injury, enabling detection of transient cellular states and the complex transcriptional repertoire that accompanies ongoing tissue damage.

How can we account for human diversity in the construction of kidney injury maps?

Any analysis of the molecular response to AKI in humans will be complicated by the heterogeneity of the injurious stimuli and by the diversity of the response to each type of injury. In addition, the kidney is a non-uniform structure meaning that the source of kidney tissue is instrumental in the transcriptomic profile. Therefore, we must first understand the normal anatomy of human kidney tissue as captured by scRNA-seq representation. In other words, we need to construct a “Normal Human Kidney Reference Map”. Such reference map will have to account for a large degree of inter-individual variation in human gene expression profiles, including variations that occur with age, gender, and genetic ancestry as well as the anatomical origin of the tissue (cortex, medulla, papilla), and perhaps even the

time of day that the sample was obtained given known diurnal variations in excretion⁷². Therefore, the same mapping procedure must be reproduced in many healthy kidney samples of diverse backgrounds in order to better quantify each element of natural variability. Some of the tissues for this purpose could be obtained from healthy nephrectomy tissue, autopsy specimens, or kidney donor biopsies. Similar maps are currently being developed for a large number of other human tissues as part of the International Human Cell Atlas (see URL).

Once inter-individual variability due to human genetic and demographic factors are better understood, the next unresolved issue is how to best account for such additional factors in the interpretation of single cell data. Genetic ancestry, specifically, is known to have an effect on gene expression in various tissues and might also mediate a large fraction of inherited disease susceptibility. Unlike recent progress in the genetic discovery for CKD^{73–78}, the genetic effects on the risk of AKI are not well studied. Yet, genetic alleles that are differentially distributed between human populations and represent strong expression quantitative trait loci (eQTLs) are potential confounders in any AKI biomarker studies.

To illustrate ancestry as a potential confounder, we consider a common regulatory polymorphism at the *UMOD* locus (rs4293393) that increases *UMOD* gene expression (represents a strong cis-eQTL) and conveys a higher risk of CKD and hypertension^{79–83}. The risk allele frequency varies by ethnicity (freq.<50% in Africans, but >90% in East Asians). Consequently, any differential gene expression experiment in which cases and controls differ by ethnicity may induce spurious signal for the *UMOD* gene. There are likely thousands of loci with similar genetic effects differentially distributed between ancestral groups. Unfortunately, unlike for many other tissues, comprehensive kidney-specific eQTL catalogues are extremely limited. For example, kidney is not well represented in the Genotype-Tissue Expression Project (GTEx, see URL)⁸⁴, and the only published healthy kidney eQTL analysis to date involved 96 human kidneys, a sample size sufficient only to detect very large cis-eQTL effects⁸⁵. Moreover, many allelic-specific expression patterns are also cell type-specific, but systematic methods for single cell eQTL detection do not yet exist. Similarly, single cell profiling of epigenetic marks that would help define cell-type specific regulatory elements is still at its infancy. This area of methods development will clearly receive a lot of attention in the coming years, and the renal research community must ensure that kidney cell types are studied as extensively as many other, more accessible tissues.

The potential confounding by population stratification in human transcriptomic studies can be controlled for by proper genetic matching of cases and controls, or by regression-based ancestry adjustments, but these techniques require adequate sample size and concurrent DNA profiling by SNP arrays or DNA sequencing. Several well-established methods for computation of human genetic ancestry, including locus-specific admixture, are now available and are standard for human genetic association studies^{86–90}. By incorporating similar methods into tissue interrogation studies, one may be able to improve the ability to partition transcriptomic variability into heritable, environmental, and stochastic components. This, in turn, can improve our ability to differentiate true signals due to the underlying

disease process from spurious signals that may be due to differences in ancestral background. The complex range of genetic, environmental, and physiological factors that can potentially confound transcriptomic studies in AKI are depicted in Figure 5, but similar considerations apply for other tissue interrogation methods, including tissue proteomics and metabolomics.

How can we identify injury-specific AKI subtypes and redefine AKI staging in molecular terms?

Acute kidney injury (AKI) is ideally suited for a precision medicine approach to diagnosis. We envision that, similar to the redefinition of cancer phenotypes⁹¹, AKI can be redefined based on specific molecular patterns that reflect causal mechanisms and kidney anatomy. One largely unanswered question in the field is whether cellular injury responses are stimulus-specific (e.g. sepsis, nephrotoxicity) or shared between diverse set of injurious stimuli. Similarly, it is not known if all nephrons respond to injury, or if the pattern of injury is limited to specific subsets of nephrons. By performing single cell or single nuclei RNA-seq of a large number of cells, one may be able to quantify the fraction of injured cells from a specific nephron segment (the fraction of cells expressing damage related genes) which in turn would reflect the “severity” of AKI affecting that segment of the kidney, hence completely redefining AKI “staging” independent of sCr. There is a clear need for new statistical methods in this area, including identification of appropriate validation sources for these newly identified markers of injury.

How can we translate molecular patterns of injury into diagnostic tests?

As a powerful alternative to simple biomarker screens that utilize case-control differential gene expression design, several network-based methods have been developed for bulk gene expression analysis, such as WGCNA (Weighted Gene Coexpression Network Analysis)⁹² or ARACNe (Algorithm for the Reconstruction of Accurate Cellular Networks)^{93–95}. These methods take advantage of the correlation structure within gene expression experiments and are able to define closely related molecular modules with their central hub genes. The goal would be to identify the central hubs in AKI etiology-specific networks, because this approach will point to mechanism-driven biomarkers that are intrinsic to the evolution of injury. This approach is less likely to be biased due to confounders intrinsic to simple case-control comparative studies in humans. The most inter-connected genes in the regulatory networks should also provide the most robust and reproducible readout of AKI disease activity, identifying mechanism-based biomarker candidates that can be used to design future clinical tests. Importantly, using single cell or nuclei RNA-seq methodology, such central hubs can now be defined in a cell-type specific fashion, greatly increasing the resolution of the approach.

Another way to harness the power of sc-/sn-RNA-seq data is to use it as reference to deconvolute bulk RNA-seq signatures into cell-type specific injury patterns. A variety of machine learning-based reverse engineering approaches already exist that can help to accomplish this task. The main advantage here is that once a sc-/sn-RNA-seq reference map is created, generation of bulk RNA-seq data from additional human kidney tissue samples is

less expensive and methodologically more straightforward. The deconvolution approach would provide a cheaper and more practical means of assessing cell-type specific signals in bulk sequence data, potentially being able to recover some of the diagnostic signals. Both supervised and unsupervised computational methods that attempt to deconvolve diverse, disordered, or highly heterogeneous tissues are presently under active development, mainly in the field of cancer^{96–102}. It may be feasible to apply similar methods to kidney tissues.

Importantly, the molecular patterns of injury derived by the above approaches will inform the development of “third generation” urine tests. For example, the top candidate biomarker transcripts can be filtered for genes encoding secreted proteins by comparisons with the Secreted ProteinDB¹⁰³ and signal sequence predictions (Ensembl^{104,105}). These filtering steps may help to prioritize potential urine biomarkers, which next can be tested in human urine samples from the same individuals or from additional human AKI cohorts. Secreted proteins of great interest would be those that (*i*) originate from damaged sectors of the kidney (*ii*) are found in the urine within hours of the stimulus (*iii*) are generated in proportion with the intensity of the signal (*iv*) are critical to the injury or recovery from injury. Secreted proteins of greatest interest also may be stimulus specific or segment specific. Stimulus specific proteins would report, for example, hemodynamic responses found in heart failure, or septic responses found in Gram-negative or Gram-positive bacterial infections, or SIRS responses due to viral infections. For example, NGAL is not particularly responsive to simple hemodynamic challenges, nor to Gram-positive or viral infections, but dramatically reports the impact of Gram-negative sepsis on the kidney. Segment specific responses would report, for example, the impact of ischemic damage in the proximal tubule, or tubule-interstitial damage, or the effect of urinary obstruction on the kidney. KIM-1 demonstrates the cellular stress of the proximal tubule, while NGAL reports the impact of stimuli that impinge the TALH and the Collecting Duct. The full implementation of these ideas would result in a ‘wiring diagram’ that links the disease stimulus with signal transduction and gene expression in a specific domain of the kidney, generating a unique and measurable outcome. The ‘wiring diagram’ would beg the question as to whether the target cells are responding as a result of a primary effect of the stimulus, or whether they are responding as a result of an intra-renal signal (hormone or metabolite³⁴) that triggers secondary responses. As far as mouse kidneys can replicate the human kidney, the specifics of the ‘wiring diagram’ can be tested with gene knockouts, or alternatively fortuitously identified in humans by naturally occurring polymorphisms that modulate the biomarker readout.

Summary and Conclusions

While standardized measurement and tracking of serum creatinine in the diagnosis of AKI has standardized epidemiologic studies and was adopted by the FDA, the widespread application of creatinine-based AKI diagnosis has provided grossly inadequate quantitative assessment of renal dysfunction. A rise in serum creatinine may not yet be present at the time of patient presentation because this metabolite must accumulate in serum and reach a clinical threshold, and because it is too insensitive to detect AKI when nephron damage does not exceed ‘renal reserve’. Conversely, fundamentally different challenges to the kidney may result in the same increase in creatinine. The consequences of these deficiencies in our

single diagnostic test has numerous negative consequences for our patients, including repeat testing, unnecessary escalation of care, inappropriate use of intravenous fluids, prolonged hospital stays, increased aggregate costs per patient, and unacceptable rates of morbidity and mortality^{106–108}.

The evolving field of precision medicine attempts to individualize diagnostics and therapeutics based on specific molecular causes of disease. For many medically important diseases, advances in massively parallel sequencing technologies have allowed for diagnostic reclassification based on primary genetic defect, enabling precise surveillance and therapy, prognosis and counseling. We believe that this kind of approach offers our community the opportunity to re-introduce, after a prolonged hiatus, both etiologic and anatomical diagnoses for AKI. Coupling the medical context and the genetic predisposition with the detection of the anatomy of injury, defined by histopathology preserved in a state-of-the-art digital kidney tissue atlas, and with specific molecular responses at a single cell level, as defined by single-cell transcriptomic, epigenetic, and proteomic assays factoring in the extent of filtration and tubular dysfunction (serum creatinine) will provide the physician, the bioengineer, the mathematician with a precision map. These data will also reflect the central notion in Nephrology, that the kidney is a sensor of the environment and an effector of homeostasis.

URLs

Cold Active Proteases: <https://www.rebuildingakidney.org/projects/cold-active-proteases/Genotype-Tissue> Expression Project (GTEx): <https://www.gtexportal.org/home> International Human Cell Atlas: www.humancellatlas.org

Acknowledgments

This publication was supported by the following grants from the National Institute of Health (NIH): UG3DK114926 (KK, AB, JB), R01DK105124 (KK), R01MD009223 (AB, KK), K01EB016071 (PS), R01DK073462 (JB), R01DK092684 (JB), U54DK104309 (JB) and the Precision Medicine Pilot from the Irving Institute/Columbia CTSA, UL1TR001873 (JB, KK, PGC, YLC, PS, RR). Additional non-government sources of funding include the Herbert Irving Scholar Award (KK), the March of Dimes Research Grant (JB) and The Coulter Foundation (JB). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or other funding sources. JB is also a co-inventor on the use of NGAL. Columbia has licensed NGAL patents to Biopointo.

Financial support for this work: none.

References

1. Bellomo R, Ronco C, Kellum JA, Mehta RL, Palevsky P, Acute Dialysis Quality Initiative w. Acute renal failure - definition, outcome measures, animal models, fluid therapy and information technology needs: the Second International Consensus Conference of the Acute Dialysis Quality Initiative (ADQI) Group. *Crit Care*. 2004; 8:R204–12. [PubMed: 15312219]
2. Mehta RL, Kellum JA, Shah SV, et al. Acute Kidney Injury Network: report of an initiative to improve outcomes in acute kidney injury. *Crit Care*. 2007; 11:R31. [PubMed: 17331245]
3. Kidney Disease: Improving Global Outcomes (KDIGO) Acute Kidney Injury Work Group. KDIGO Clinical Practice Guideline for Acute Kidney Injury. *Kidney Int Suppl*. 2012; 2:1–138.
4. Anderson S, Eldadah B, Halter JB, et al. Acute kidney injury in older adults. *J Am Soc Nephrol*. 2011; 22:28–38. [PubMed: 21209252]

5. Ishani A, Xue JL, Himmelfarb J, et al. Acute kidney injury increases risk of ESRD among elderly. *J Am Soc Nephrol.* 2009; 20:223–8. [PubMed: 19020007]
6. Bucaloiu ID, Kirchner HL, Norfolk ER, Hartle JE 2nd, Perkins RM. Increased risk of death and de novo chronic kidney disease following reversible acute kidney injury. *Kidney Int.* 2012; 81:477–85. [PubMed: 22157656]
7. Chawla LS, Kimmel PL. Acute kidney injury and chronic kidney disease: an integrated clinical syndrome. *Kidney Int.* 2012; 82:516–24. [PubMed: 22673882]
8. Coca SG, Singanamala S, Parikh CR. Chronic kidney disease after acute kidney injury: a systematic review and meta-analysis. *Kidney Int.* 2012; 81:442–8. [PubMed: 22113526]
9. Garg AX, Parikh CR. Yin and Yang: acute kidney injury and chronic kidney disease. *J Am Soc Nephrol.* 2009; 20:8–10. [PubMed: 19118145]
10. Labban B, Arora N, Restaino S, Markowitz G, Valeri A, Radhakrishnan J. The role of kidney biopsy in heart transplant candidates with kidney disease. *Transplantation.* 2010; 89:887–93. [PubMed: 20220572]
11. Bergler-Klein J, Pirich C, Laufer G, et al. The long-term effect of simultaneous heart and kidney transplantation on native renal function. *Transplantation.* 2001; 71:1597–600. [PubMed: 11435971]
12. Parikh CR, Coca SG. Acute kidney injury: defining prerenal azotemia in clinical practice and research. *Nat Rev Nephrol.* 2010; 6:641–2. [PubMed: 20981121]
13. Nickolas TL, Schmidt-Ott KM, Canetta P, et al. Diagnostic and prognostic stratification in the emergency department using urinary biomarkers of nephron damage: a multicenter prospective cohort study. *J Am Coll Cardiol.* 2012; 59:246–55. [PubMed: 22240130]
14. Nickolas TL, O'Rourke MJ, Yang J, et al. Sensitivity and specificity of a single emergency department measurement of urinary neutrophil gelatinase-associated lipocalin for diagnosing acute kidney injury. *Ann Intern Med.* 2008; 148:810–9. [PubMed: 18519927]
15. Xu K, Rosenstiel P, Paragas N, et al. Unique Transcriptional Programs Identify Subtypes of AKI. *J Am Soc Nephrol.* 2017; 28:1729–40. [PubMed: 28028135]
16. Barasch J, Zager R, JV B. Acute Kidney Injury: A Problem of Definition. *Lancet.* In Press.
17. <http://www.cdc.gov/nchs/fastats/emergency-department.htm>
18. <http://healthpolicy.ucsf.edu/content/icu-outcomes>
19. Moretti C, Frajese GV, Guccione L, et al. Androgens and body composition in the aging male. *J Endocrinol Invest.* 2005; 28:56–64.
20. Kimmel PL, Lew SQ, Bosch JP. Nutrition, ageing and GFR: is age-associated decline inevitable? *Nephrol Dial Transplant.* 1996; 11(Suppl 9):85–8. [PubMed: 9050040]
21. Musso CG, Michelangelo H, Vilas M, et al. Creatinine reabsorption by the aged kidney. *Int Urol Nephrol.* 2009; 41:727–31. [PubMed: 19115077]
22. Sjostrom PA, Odland BG, Wolgast M. Extensive tubular secretion and reabsorption of creatinine in humans. *Scand J Urol Nephrol.* 1988; 22:129–31. [PubMed: 3206217]
23. Sise ME, Forster C, Singer E, et al. Urine neutrophil gelatinase-associated lipocalin identifies unilateral and bilateral urinary tract obstruction. *Nephrol Dial Transplant.* 2011; 26:4132–5. [PubMed: 22049182]
24. Kaddourah A, Basu RK, Bagshaw SM, Goldstein SL. Epidemiology of Acute Kidney Injury in Critically Ill Children and Young Adults. *NEJM.* 2016; doi: 10.1056/NEJMoa1611391
25. Garner AE, Lewington AJ, Barth JH. Detection of patients with acute kidney injury by the clinical laboratory using rises in serum creatinine: comparison of proposed definitions and a laboratory delta check. *Ann Clin Biochem.* 2012; 49:59–62. [PubMed: 22130632]
26. Lin J, Fernandez H, Shashaty MG, et al. False-Positive Rate of AKI Using Consensus Creatinine-Based Criteria. *Clin J Am Soc Nephrol.* 2015; 10:1723–31. [PubMed: 26336912]
27. Sabia P, Afrookteh A, Touchstone DA, Keller MW, Esquivel L, Kaul S. Value of regional wall motion abnormality in the emergency room diagnosis of acute myocardial infarction. A prospective study using two-dimensional echocardiography. *Circulation.* 1991; 84:185–92. [PubMed: 1884510]

28. Luepker RV, Duval S, Jacobs DR Jr, Smith LG, Berger AK. The effect of changing diagnostic algorithms on acute myocardial infarction rates. *Ann Epidemiol.* 2011; 21:824–9. [PubMed: 21982485]
29. Xu K, Rosenstiel P, Paragas N, et al. Unique Transcriptional Programs Identify Subtypes of “Acute Kidney Injury”. *J Am Soc Nephrol.* 2016 In Press.
30. Belcher JM, Sanyal AJ, Peixoto AJ, et al. Kidney biomarkers and differential diagnosis of patients with cirrhosis and acute kidney injury. *Hepatology.* 2014; 60:622–32. [PubMed: 24375576]
31. Nejat M, Pickering JW, Devarajan P, et al. Some biomarkers of acute kidney injury are increased in pre-renal acute injury. *Kidney Int.* 2012; 81:1254–62. [PubMed: 22418979]
32. Biner HL, Arpin-Bott MP, Loffing J, et al. Human cortical distal nephron: distribution of electrolyte and water transport pathways. *J Am Soc Nephrol.* 2002; 13:836–47. [PubMed: 11912242]
33. Werth M, Schmidt-Ott M, Leete T, et al. TFCP2L1 Controls the Cellular Patterning of the Collecting Duct of the Kidney. 2016 Submitted.
34. Azroyan A, Cortez-Retamozo V, Bouley R, et al. Renal intercalated cells sense and mediate inflammation via the P2Y₁₄ receptor. *PLoS One.* 2015; 10:e0121419. [PubMed: 25799465]
35. Paragas N, Qiu A, Zhang Q, et al. The Ngal reporter mouse detects the response of the kidney to injury in real time. *Nat Med.* 2011; 17:216–22. [PubMed: 21240264]
36. Fejes-Toth G, Naray-Fejes-Toth A. Differentiation of renal beta-intercalated cells to alpha-intercalated and principal cells in culture. *Proc Natl Acad Sci U S A.* 1992; 89:5487–91. [PubMed: 1608958]
37. Park EY, Kim WY, Kim YM, et al. Proposed mechanism in the change of cellular composition in the outer medullary collecting duct during potassium homeostasis. *Histol Histopathol.* 2012; 27:1559–77. [PubMed: 23059887]
38. Trepiccione F, Capasso G, Nielsen S, Christensen BM. Evaluation of cellular plasticity in the collecting duct during recovery from lithium-induced nephrogenic diabetes insipidus. *Am J Physiol Renal Physiol.* 2013; 305:F919–29. [PubMed: 23825070]
39. Matos LL, Trufelli DC, de Matos MG, da Silva Pinhal MA. Immunohistochemistry as an important tool in biomarkers detection and clinical practice. *Biomark Insights.* 2010; 5:9–20. [PubMed: 20212918]
40. Yuste R. Fluorescence microscopy today. *Nat Methods.* 2005; 2:902–4. [PubMed: 16299474]
41. Gay L, Miller MR, Ventura PB, et al. Mouse TU tagging: a chemical/genetic intersectional method for purifying cell type-specific nascent RNA. *Genes Dev.* 2013; 27:98–115. [PubMed: 23307870]
42. Hrvatin S, Deng F, O'Donnell CW, Gifford DK, Melton DA. MARIS: method for analyzing RNA following intracellular sorting. *PLoS One.* 2014; 9:e89459. [PubMed: 24594682]
43. Matos TR, Liu H, Ritz J. Research Techniques Made Simple: Experimental Methodology for Single-Cell Mass Cytometry. *J Invest Dermatol.* 2017; 137:e31–e8. [PubMed: 28340685]
44. Saliba AE, Westermann AJ, Gorski SA, Vogel J. Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res.* 2014; 42:8845–60. [PubMed: 25053837]
45. Stegle O, Teichmann SA, Marioni JC. Computational and analytical challenges in single-cell transcriptomics. *Nat Rev Genet.* 2015; 16:133–45. [PubMed: 25628217]
46. Wagner A, Regev A, Yosef N. Revealing the vectors of cellular identity with single-cell genomics. *Nat Biotechnol.* 2016; 34:1145–60. [PubMed: 27824854]
47. Grun D, van Oudenaarden A. Design and Analysis of Single-Cell Sequencing Experiments. *Cell.* 2015; 163:799–810. [PubMed: 26544934]
48. Zeng W, Jiang S, Kong X, et al. Single-nucleus RNA-seq of differentiating human myoblasts reveals the extent of fate heterogeneity. *Nucleic Acids Res.* 2016; 44:e158. [PubMed: 27566152]
49. Grindberg RV, Yee-Greenbaum JL, McConnell MJ, et al. RNA-sequencing from single nuclei. *Proc Natl Acad Sci U S A.* 2013; 110:19802–7. [PubMed: 24248345]
50. Krishnaswami SR, Grindberg RV, Novotny M, et al. Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. *Nat Protoc.* 2016; 11:499–524. [PubMed: 26890679]
51. Lacar B, Linker SB, Jaeger BN, et al. Nuclear RNA-seq of single neurons reveals molecular signatures of activation. *Nat Commun.* 2016; 7:11022. [PubMed: 27090946]

52. Lake BB, Ai R, Kaeser GE, et al. Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science*. 2016; 352:1586–90. [PubMed: 27339989]
53. Yuan J, Sims PA. An Automated Microwell Platform for Large-Scale Single Cell RNA-Seq. *Sci Rep*. 2016; 6:33883. [PubMed: 27670648]
54. Bose S, Wan Z, Carr A, et al. Scalable microfluidics for single-cell RNA printing and sequencing. *Genome Biol*. 2015; 16:120. [PubMed: 26047807]
55. Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol*. 2015; 33:495–502. [PubMed: 25867923]
56. van der Maaten L, Hinton G. Visualizing data using t-SNE. *Journal of Machine Learning Research*. 2008; 9:2579–605.
57. Zeisel A, Munoz-Manchado AB, Codeluppi S, et al. Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science*. 2015; 347:1138–42. [PubMed: 25700174]
58. Levine JH, Simonds EF, Bendall SC, et al. Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell*. 2015; 162:184–97. [PubMed: 26095251]
59. Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*. 2008; 2008:P10008.
60. Grun D, Muraro MJ, Boisset JC, et al. De Novo Prediction of Stem Cell Identity using Single-Cell Transcriptome Data. *Cell Stem Cell*. 2016; 19:266–77. [PubMed: 27345837]
61. Grun D, Lyubimova A, Kester L, et al. Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature*. 2015; 525:251–5. [PubMed: 26287467]
62. Li H, Courtois ET, Sengupta D, et al. Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. *Nat Genet*. 2017; 49:708–18. [PubMed: 28319088]
63. Shekhar K, Lapan SW, Whitney IE, et al. Comprehensive Classification of Retinal Bipolar Neurons by Single-Cell Transcriptomics. *Cell*. 2016; 166:1308–23 e30. [PubMed: 27565351]
64. Setty M, Tadmor MD, Reich-Zeliger S, et al. Wishbone identifies bifurcating developmental trajectories from single-cell data. *Nat Biotechnol*. 2016; 34:637–45. [PubMed: 27136076]
65. Trapnell C, Cacchiarelli D, Grimsby J, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol*. 2014; 32:381–6. [PubMed: 24658644]
66. Haghverdi L, Buttner M, Wolf FA, Buettner F, Theis FJ. Diffusion pseudotime robustly reconstructs lineage branching. *Nat Methods*. 2016; 13:845–8. [PubMed: 27571553]
67. Marco E, Karp RL, Guo G, et al. Bifurcation analysis of single-cell gene expression data reveals epigenetic landscape. *Proc Natl Acad Sci U S A*. 2014; 111:E5643–50. [PubMed: 25512504]
68. Rizvi AH, Camara PG, Kandror EK, et al. Single-cell topological RNA-seq analysis reveals insights into cellular differentiation and development. *Nat Biotechnol*. 2017
69. Cámara PG. Topological methods for genomics: present and future directions. *Current Opinion in Systems Biology*. 2016
70. Angerer P, Haghverdi L, Buttner M, Theis FJ, Marr C, Buettner F. destiny: diffusion maps for large-scale single-cell data in R. *Bioinformatics*. 2016; 32:1241–3. [PubMed: 26668002]
71. Coifman RR, Lafon S, Lee AB, et al. Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. *Proceedings of the National Academy of Sciences of the United States of America*. 2005; 102:7426–31. [PubMed: 15899970]
72. Nikolaeva S, Ansermet C, Centeno G, et al. Nephron-Specific Deletion of Circadian Clock Gene *Bmal1* Alters the Plasma and Renal Metabolome and Impairs Drug Disposition. *J Am Soc Nephrol*. 2016; 27:2997–3004. [PubMed: 27056296]
73. Kiryluk K, Li Y, Sanna-Cherchi S, et al. Geographic differences in genetic susceptibility to IgA nephropathy: GWAS replication study and geospatial risk analysis. *PLoS Genet*. 2012; 8:e1002765. [PubMed: 22737082]
74. Kiryluk K, Li Y, Scolari F, et al. Discovery of new risk loci for IgA nephropathy implicates genes involved in immunity against intestinal pathogens. *Nat Genet*. 2014; 46:1187–96. [PubMed: 25305756]

75. Gharavi AG, Kiryluk K, Choi M, et al. Genome-wide association study identifies susceptibility loci for IgA nephropathy. *Nat Genet.* 2011; 43:321–7. [PubMed: 21399633]
76. Genovese G, Friedman DJ, Ross MD, et al. Association of trypanolytic ApoL1 variants with kidney disease in African Americans. *Science.* 2010; 329:841–5. [PubMed: 20647424]
77. Parsa A, Kao WH, Xie D, et al. APOL1 risk variants, race, and progression of chronic kidney disease. *N Engl J Med.* 2013; 369:2183–96. [PubMed: 24206458]
78. Pattaro C, Teumer A, Gorski M, et al. Genetic associations at 53 loci highlight cell types and biological pathways relevant for kidney function. *Nat Commun.* 2016; 7:10023. [PubMed: 26831199]
79. 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. [PubMed: 19430482]
80. Kottgen A, Pattaro C, Boger CA, et al. New loci associated with kidney function and chronic kidney disease. *Nat Genet.* 2010; 42:376–84. [PubMed: 20383146]
81. Olden M, Corre T, Hayward C, et al. Common variants in UMOD associate with urinary uromodulin levels: a meta-analysis. *J Am Soc Nephrol.* 2014; 25:1869–82. [PubMed: 24578125]
82. 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. [PubMed: 22479191]
83. Padmanabhan S, Melander O, Johnson T, et al. Genome-wide association study of blood pressure extremes identifies variant near UMOD associated with hypertension. *PLoS Genet.* 2010; 6:e1001177. [PubMed: 21082022]
84. Consortium GT. The Genotype-Tissue Expression (GTEx) project. *Nat Genet.* 2013; 45:580–5. [PubMed: 23715323]
85. Ko YA, Yi H, Qiu C, et al. Genetic-Variation-Driven Gene-Expression Changes Highlight Genes with Important Functions for Kidney Disease. *Am J Hum Genet.* 2017; 100:940–53. [PubMed: 28575649]
86. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. *PLoS Genet.* 2006; 2:e190. [PubMed: 17194218]
87. Lee AB, Luca D, Roeder K. A Spectral Graph Approach to Discovering Genetic Ancestry. *Ann Appl Stat.* 2010; 4:179–202. [PubMed: 20689656]
88. Lee AB, Luca D, Klei L, Devlin B, Roeder K. Discovering genetic ancestry using spectral graph theory. *Genet Epidemiol.* 2010; 34:51–9. [PubMed: 19455578]
89. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome research.* 2009; 19:1655–64. [PubMed: 19648217]
90. Price AL, Tandon A, Patterson N, et al. Sensitive detection of chromosomal segments of distinct ancestry in admixed populations. *PLoS Genet.* 2009; 5:e1000519. [PubMed: 19543370]
91. Ciriello G, Gatz ML, Beck AH, et al. Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. *Cell.* 2015; 163:506–19. [PubMed: 26451490]
92. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC bioinformatics.* 2008; 9:559. [PubMed: 19114008]
93. Margolin AA, Wang K, Lim WK, Kustagi M, Nemenman I, Califano A. Reverse engineering cellular networks. *Nat Protoc.* 2006; 1:662–71. [PubMed: 17406294]
94. Margolin AA, Nemenman I, Basso K, et al. ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC bioinformatics.* 2006; 7(Suppl 1):S7.
95. Basso K, Margolin AA, Stolovitzky G, Klein U, Dalla-Favera R, Califano A. Reverse engineering of regulatory networks in human B cells. *Nat Genet.* 2005; 37:382–90. [PubMed: 15778709]
96. Clarke J, Seo P, Clarke B. Statistical expression deconvolution from mixed tissue samples. *Bioinformatics.* 2010; 26:1043–9. [PubMed: 20202973]
97. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods.* 2015; 12:453–7. [PubMed: 25822800]
98. Ahn SG, Lee HM, Lee HW, et al. Prognostic discrimination using a 70-gene signature among patients with estrogen receptor-positive breast cancer and an intermediate 21-gene recurrence score. *International journal of molecular sciences.* 2013; 14:23685–99. [PubMed: 24304542]

99. Gong T, Hartmann N, Kohane IS, et al. Optimal deconvolution of transcriptional profiling data using quadratic programming with application to complex clinical blood samples. *PLoS One*. 2011; 6:e27156. [PubMed: 22110609]
100. Kuhn A, Thu D, Waldvogel HJ, Faull RL, Luthi-Carter R. Population-specific expression analysis (PSEA) reveals molecular changes in diseased brain. *Nat Methods*. 2011; 8:945–7. [PubMed: 21983921]
101. Li Y, Xie X. A mixture model for expression deconvolution from RNA-seq in heterogeneous tissues. *BMC bioinformatics*. 2013; 14(Suppl 5):S11.
102. Qiao W, Quon G, Csaszar E, Yu M, Morris Q, Zandstra PW. PERT: a method for expression deconvolution of human blood samples from varied microenvironmental and developmental conditions. *PLoS computational biology*. 2012; 8:e1002838. [PubMed: 23284283]
103. Chen Y, Zhang Y, Yin Y, et al. SPD—a web-based secreted protein database. *Nucleic Acids Res*. 2005; 33:D169–73. [PubMed: 15608170]
104. Cunningham F, Amode MR, Barrell D, et al. Ensembl 2015. *Nucleic Acids Res*. 2015; 43:D662–9. [PubMed: 25352552]
105. UniProt C. Activities at the Universal Protein Resource (UniProt). *Nucleic Acids Res*. 2014; 42:D191–8. [PubMed: 24253303]
106. Waikar SS, Liu KD, Chertow GM. Diagnosis, epidemiology and outcomes of acute kidney injury. *Clin J Am Soc Nephrol*. 2008; 3:844–61. [PubMed: 18337550]
107. Chawla LS, Amdur RL, Shaw AD, Faselis C, Palant CE, Kimmel PL. Association between AKI and Long-Term Renal and Cardiovascular Outcomes in United States Veterans. *Clin J Am Soc Nephro*. 2014; 9:448–56.
108. Parikh A, Nickolas T. Does NGAL reduce costs? A cost analysis of urine NGAL (UNGAL) and serum creatinine (SCR) for acute kidney injury (AKI) diagnosis. *Crit Care Med*. 2010; 38:A106.

Figure 1. NGAL and KIM-1 urinary biomarker levels

Urine was obtained upon patient contact in the Emergency Department. Post-hoc diagnostic adjudication was based on baseline levels of serum creatinine, the duration of elevated serum creatinine, and the response to fluids over the subsequent 7 days after admission from the Emergency Department. Note that the urinary proteins, collected at the time of the Emergency Department visit distinguished prolonged elevations of serum creatinine (“intrinsic” iAKI) from transient (reversible) elevations of serum creatinine (“prerenal” pAKI) and from stably depressed kidney function (CKD). Unclassified patients (Uncl) could not be adjudicated on the basis of creatinine kinetics. Adapted with permission from Nickolas TL et al. *J Am Coll Cardiol.* 2012;59(3):246-55.

Figure 2. Expression of NGAL in mouse kidney

brief cross clamp of the renal artery induces gene expression in intercalated cells (IC) and in the thick ascending limbs (TALH), whereas volume depleted mice do not express NGAL despite a similar elevation of serum creatinine. Control mice do not express NGAL. Adapted with permission from Xu K et al. *J Am Soc Nephrol.* 2017;17(6):1729-1740.

Figure 3. Nephron segment-specific expression of AKI biomarkers after a brief dose of ischemia
In situ hybridization for NGAL expressed in the thick ascending limbs (TALH) and intercalated (IC) cells, and KIM1 expressed in proximal tubules. Adapted from Paragas N. et al. Nat Med. 2011;17(2):216-22.

Figure 4. Single nuclei RNA-seq analysis of normal kidney tissue

The distribution of segment-specific gene markers in isolated kidney nuclei by snRNA-seq displayed in t- Distributed Stochastic Neighbor Embedding (t-SNE) space.

Figure 5. Next generation transcriptomics in AKI: challenges and methods

The interplay of genetic, environmental and physiologic factors determines baseline variability in normal kidney transcriptome that may confound tissue interrogation efforts. The superimposed tissue injury in AKI may result in distinctive gene expression patterns that are dependent on the type, site, severity, and chronicity of injury. Normal reference map of the human kidney transcriptome derived from a large and diverse human cohort is needed to better differentiate AKI-specific signals from baseline variability. Modern tissue interrogation methods include bulk RNA-seq, laser capture microdissection (LCM) followed by compartment (or segment) specific RNA-seq, single cell RNA-seq and single nuclei RNA-seq. Computational deconvolution of bulk transcriptome into cell-type specific transcriptomic profiles is also under active development. Lastly, a variety of novel multiplex *in situ* methods that aim at spatial mapping of injured cells can be used to confirm the precise anatomical site of injury.