



## OPEN ACCESS

EDITED BY  
William C. Cho,  
QEH, Hong Kong SAR, China

REVIEWED BY  
Paul Hoover,  
Brigham and Women's Hospital and  
Harvard Medical School, United States  
Chenglong Xia,  
University of California, Berkeley,  
United States

\*CORRESPONDENCE  
Luke N. Robinson,  
✉ lrobinson@visterrainc.com

†These authors have contributed equally  
to this work and share first authorship

SPECIALTY SECTION  
This article was submitted to Molecular  
Diagnostics and Therapeutics,  
a section of the journal  
Frontiers in Molecular Biosciences

RECEIVED 01 March 2023  
ACCEPTED 21 March 2023  
PUBLISHED 05 April 2023

CITATION  
Pregizer S, Vreven T, Mathur M and  
Robinson LN (2023), Multi-omic single  
cell sequencing: Overview and  
opportunities for kidney disease  
therapeutic development.  
*Front. Mol. Biosci.* 10:1176856.  
doi: 10.3389/fmolb.2023.1176856

COPYRIGHT  
© 2023 Pregizer, Vreven, Mathur and  
Robinson. This is an open-access article  
distributed under the terms of the  
[Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is  
permitted, provided the original author(s)  
and the copyright owner(s) are credited  
and that the original publication in this  
journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Multi-omic single cell sequencing: Overview and opportunities for kidney disease therapeutic development

Steven Pregizer<sup>†</sup>, Thom Vreven<sup>†</sup>, Mohit Mathur and  
Luke N. Robinson\*

Visterra Inc., Waltham, MA, United States

Single cell sequencing technologies have rapidly advanced in the last decade and are increasingly applied to gain unprecedented insights by deconstructing complex biology to its fundamental unit, the individual cell. First developed for measurement of gene expression, single cell sequencing approaches have evolved to allow simultaneous profiling of multiple additional features, including chromatin accessibility within the nucleus and protein expression at the cell surface. These multi-omic approaches can now further be applied to cells *in situ*, capturing the spatial context within which their biology occurs. To extract insights from these complex datasets, new computational tools have facilitated the integration of information across different data types and the use of machine learning approaches. Here, we summarize current experimental and computational methods for generation and integration of single cell multi-omic datasets. We focus on opportunities for multi-omic single cell sequencing to augment therapeutic development for kidney disease, including applications for biomarkers, disease stratification and target identification.

## KEYWORDS

single-cell, multi-omic, scRNA-seq, kidney, integration

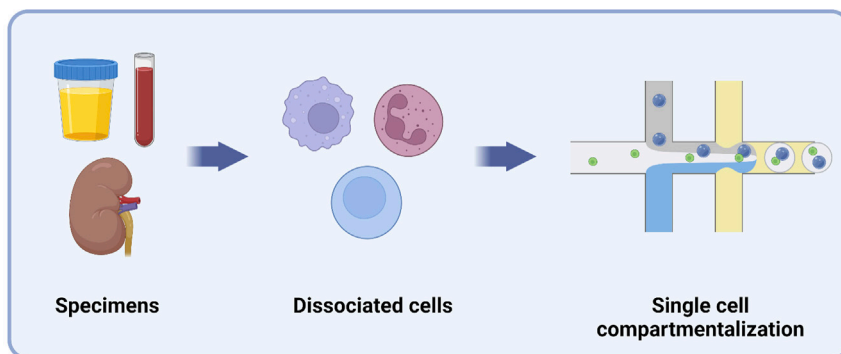
## 1 Single cell multi-omic assays

### 1.1 Single cell assays

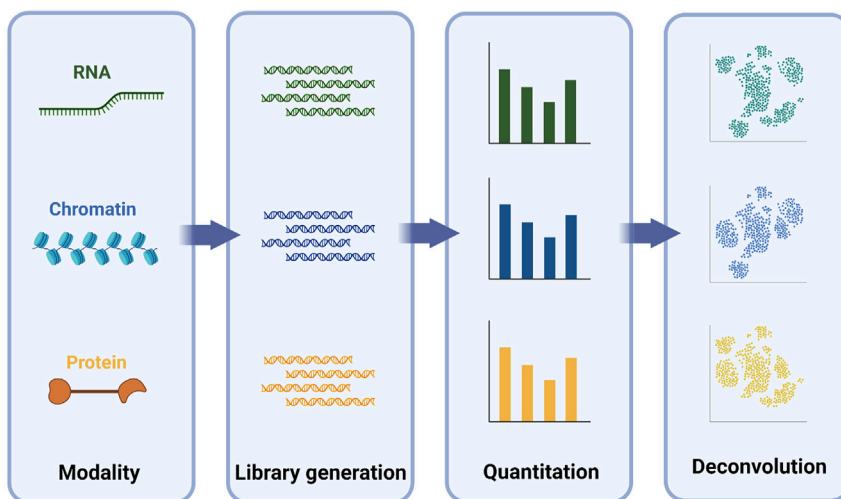
The development of single cell sequencing technology focused initially on profiling the transcriptome (Tang et al., 2009); however, the number of assays that can be adapted to sequencing at the single cell level has undergone rapid growth. Importantly, multiple assays, each measuring a different molecular property, can be performed concurrently on the same cells, enabling a multi-omic deconstruction of individual cells (Figure 1). Collectively, these assays have become a powerful new toolkit for probing myriad aspects of single cells, including transcriptomic, epigenomic, and proteomic signatures. In this review, we will first focus on three different single cell assays.

Single cell transcriptome profiling by RNA-sequencing (scRNA-seq) entails the generation of a barcoded complementary DNA (cDNA) library from the total pool of RNA transcripts present in a given cell at a given time. The chemistries commonly employed for high-throughput single cell cDNA generation typically bias the assay towards the 5' or 3' end of the most abundant polyadenylated transcripts (Klein et al., 2015; Macosko et al., 2015; Cao et al., 2017; Zheng et al., 2017), while lower-throughput chemistries offer higher

### Preparation for single cell sequencing



### Multi-omic single cell assays



### Multi-omic single cell assay integration

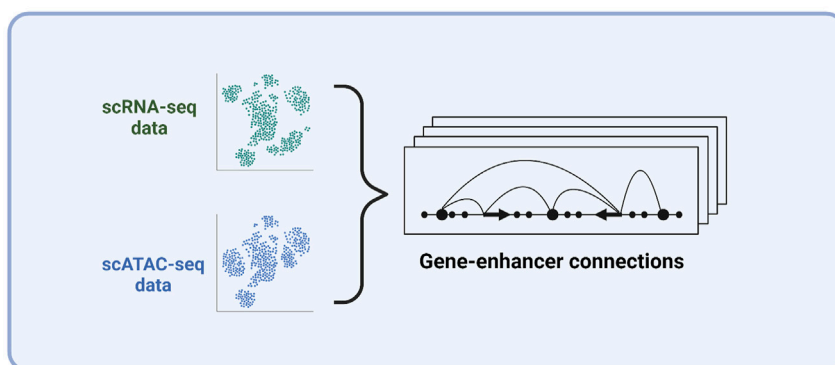


FIGURE 1

Multi-omic single cell sequencing provides an integrated view of cell biology from heterogeneous tissue sources.

sensitivity and full-length coverage, including non-polyadenylated species (Hagemann-Jensen et al., 2020; Hahaut et al., 2022; Salmen et al., 2022). Regardless of the chemistry used, the transcriptome of each cell is effectively sampled *via* this assay, enabling interrogation

of gene expression without the need to pre-select a set of target genes. A key benefit of this target-agnostic approach is the ability to explore fundamental aspects of cell biology without the constraints of *a priori* assumptions. For example, each cell’s “transcriptomic

signature” can be used for cell type classification, typically *via* comparison with reference maps or by clustering and manual inspection of expression of marker genes. With unbiased single-cell resolution, novel cell types may be appreciated that were missed by conventional classification schemes, typically relying on a handful of pre-defined markers (Regev et al., 2017). Likewise, changes in gene expression can be deconstructed to understand biological activities and activated pathways in cells, enabling insights to the functional roles and developmental trajectories of different cells that may elude more biased approaches.

Single cell epigenomic profiling by assay for transposase-accessible chromatin (scATAC) entails fragmentation and tagging of the genome with barcoded adapters for high-throughput sequencing by Tn5 transposase (Buenrostro et al., 2015). Because the enzyme preferentially integrates DNA tags into nucleosome-free regions of “open” chromatin, these regions are enriched in scATAC-seq libraries, while regions of “closed” chromatin are not. The enriched regions typically correspond to active regulatory elements; namely, promoters and enhancers. Thus, scATAC-seq provides genome-wide high-resolution maps of regulatory activity on a cell-by-cell basis. These regulatory maps are essential for pinpointing cell type-specific enhancers that elude identification by bulk methods (Buenrostro et al., 2015; Buenrostro et al., 2018). Transcriptional activity is typically presaged by chromatin changes within a given locus (Ma et al., 2020); thus, single cell regulatory maps derived from scATAC-seq are complementary to transcriptomic profiles derived from scRNA-seq and can likewise be used to deconvolute cell types and states, and in some cases, more effectively (Cusanovich et al., 2018; Shema et al., 2019). Finally, the repertoire of active regulatory elements in a given cell type can be interrogated at the primary sequence level to identify over-represented transcription factor binding motifs (Cusanovich et al., 2018). These molecular “footprints,” in turn, nominate master regulators of transcriptional activity within each cell type. Thus, scATAC-seq offers a bias-free approach for identifying upstream factors governing cell fate and function.

Single cell proteomic profiling by sequencing entails labeling cells with DNA-barcoded antibodies against protein targets, enabling quantitative measurement of their expression. Targets profiled can include those on the cell surface (e.g., CITE-seq, REAP-seq) (Peterson et al., 2017; Shahi et al., 2017; Stoeckius et al., 2017; O’Huallachain et al., 2020; Hwang et al., 2021; Sheng et al., 2022), as well as those within the cytosolic (van Buggenum et al., 2018; Gerlach et al., 2019) and nuclear (Chung et al., 2021) compartments. Sequencing-based single cell proteomic profiling is orthogonal to flow or mass cytometry, which use fluorophore or isotope-coupled antibodies, respectively, to probe antigens on a single-cell level. In contrast to these methods, sequencing-based proteomic profiling allows for as many as 10-fold more proteins to be profiled simultaneously, with studies demonstrating >100 proteins profiled (Su et al., 2020; Nettersheim et al., 2022). The number of proteins profiled, in practice, for sequencing-based single cell proteomic profiling is limited by the availability and inclusion of high-quality antibodies, and thus contrasts somewhat from the assays discussed previously. Nevertheless, sequencing-based single-cell

proteomic profiling, even with a limited number of targets, is quite capable of resolving cell types and states (Shahi et al., 2017). Many cell types, especially immune cells, are defined by the proteins they express on their surface, and thus proteomic profiling by sequencing enables more direct relating of data to classic immunophenotyping, e.g., flow cytometry. Finally, sequencing-based single cell proteomic profiling has the capacity to detect post-translational modifications, allowing signaling events to be directly observed, rather than inferred from expression of pathway constituent and/or target genes (van Buggenum et al., 2018; Gerlach et al., 2019).

## 1.2 Integration of multiple single cell assay formats

The conversion of transcript expression, protein expression, and chromatin accessibility to a common sequence-based readout makes it possible to simultaneously obtain multiple modalities from individual cells in a single experiment. The first reports of multi-omic single cell profiling involved the combined profiling of transcriptomes and proteins. These early multi-omic assays included Cellular Indexing of Transcriptomes and Epitopes by Sequencing (CITE-seq) (Stoeckius et al., 2017), and RNA expression and Protein Sequencing (REAP-seq) (Peterson et al., 2017). Several additional variations on this theme have since emerged (Peterson et al., 2017; Gerlach et al., 2019; Chung et al., 2021; Hwang et al., 2021). Joint profiling of protein and transcript expression in the same cells effectively increases the overall amount of information available for unsupervised clustering/neighborhood mapping, thus providing enhanced ability to resolve cell types and states (Stoeckius et al., 2017). Additionally, it allows functional inferences to be drawn between expression of nuclear proteins and the potential impact on gene expression (Chung et al., 2021).

Joint profiling of transcript expression and chromatin accessibility was described shortly after the introduction of CITE-seq and REAP-seq (Cao et al., 2018), and this was likewise followed by several variations (Chen et al., 2019; Zhu et al., 2019; Ma et al., 2020). As with paired protein and transcript expression, the integration of paired chromatin accessibility and transcript expression from single cells can enhance resolution and identification of cell types and states. Importantly, the ability to directly link the regulatory landscape of a single cell with its transcriptome has powerful implications for deconvoluting gene regulatory networks. Changes in a gene’s transcriptional activity can be directly correlated with changes in accessibility of specific regions in the surrounding locus, allowing assignment of regulatory elements to their cognate target genes to be made with much greater accuracy. This, in turn, leads to more robust hypotheses about how common disease-associated variants mapping to such regulatory elements may be exerting their effects (Ma et al., 2020; Kartha et al., 2022).

Joint single cell profiling of chromatin accessibility and protein expression has been described more recently (Mimitou et al., 2021; Swanson et al., 2021). These methods

were developed in connection with methods that additionally profile transcriptomes from single cells (Mimitou et al., 2021; Swanson et al., 2021). Together with a similar, recently developed trimodal single cell assay (Chen et al., 2022), these methods herald a sea change in our approach to cell biology. These trimodal methods make it possible, for the first time, to directly observe how changes within the nucleus propagate to the cell surface, and *vice versa*. These methods are still in their infancy, and further refinements will be necessary to aid their widespread adoption; however, they offer a glimpse of what promises to be a more holistic era in cell biology.

### 1.3 Opportunities and challenges for single cell assays in kidney disease

Kidney is an inherently challenging organ for single cell analysis, owing to the need to first dissociate cells from the matrix in which they are embedded. This typically entails enzymatic digestion and/or mechanical disruption, processes that can result in cell death and/or stress responses. The latter can be particularly confounding for single cell analysis, as transcriptional responses to stress may obscure native cellular phenotypes (Adam et al., 2017; O'Flanagan et al., 2019). Use of cold-activated proteases can alleviate this somewhat; however, the greater challenge in sequencing single cells from the kidney is that some cell types are more refractory than others to enzymatic dissociation (Wu et al., 2018). When these refractory cell types are also rare, they can be severely under-represented in single cell datasets. Glomerular cells, including podocytes, mesangial cells, and epithelial cells are a classic example of this conundrum (Wu et al., 2019). This can be overcome to some degree by scaling up the number of cells profiled and adapting the method to target a particular cell type (Chung et al., 2020); however, these approaches are not always feasible, especially with clinical samples.

Single nucleus RNA sequencing (snRNA-seq) has emerged as a helpful alternative to single cell sequencing in cases where cell dissociation is necessary and poses a challenge (Grindberg et al., 2013; Habib et al., 2016; Krishnaswami et al., 2016; Lacarr et al., 2016; Lake et al., 2016; Habib et al., 2017; Martin et al., 2023). Nuclei are more resistant to lysis than cells and contain abundant pre-spliced mRNA, in addition to housing the genomic material. Thus, harsher mechanical disruption techniques can be applied to recover nuclei from difficult tissues without compromising the integrity of their contents. As with scRNA-seq, snRNA-seq can be adapted to profile chromatin accessibility in parallel (Cao et al., 2018; Chen et al., 2019; Zhu et al., 2019; Ma et al., 2020); however, profiling surface protein expression is precluded with this approach. At least one study has shown that snRNA-seq is superior to scRNA-seq for recovery of glomerular cells (Wu et al., 2019). An additional advantage of single nucleus sequencing is its applicability to frozen tissues, thus enabling information to be extracted at single-cell resolution from frozen/archived clinical samples that would otherwise not be amenable to single cell sequencing (Lake et al., 2019; Rousselle et al., 2022).

## 2 Computational approaches for integrated analyses of multi-omic single cell datasets

### 2.1 Overview of integration approaches

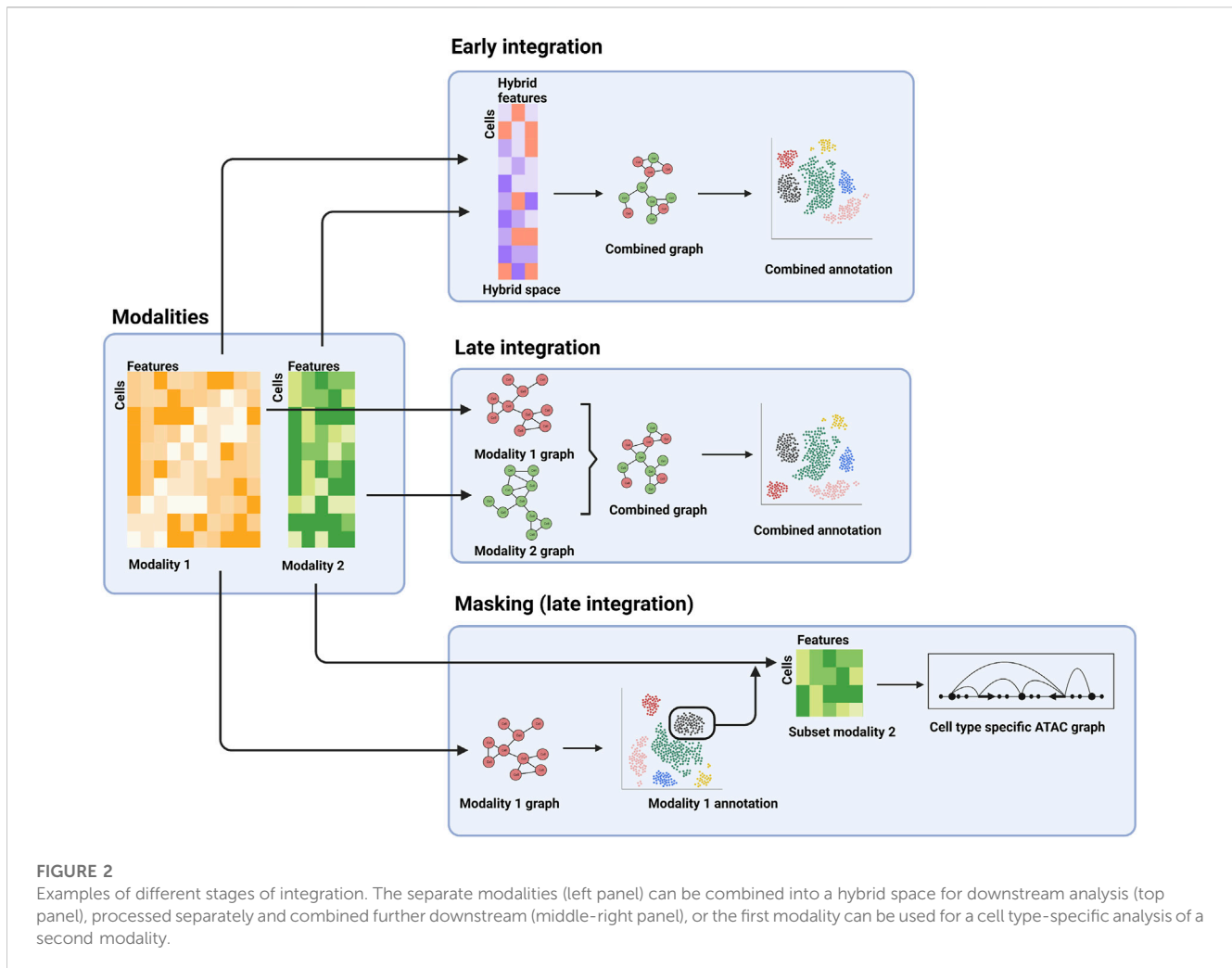
The availability of multiple assays, or modalities, for single cell sequencing can provide a much deeper understanding of the biological processes compared to single modalities. On their own, individual modalities can present complementary evidence or serve as independent validation of biological findings. To unlock a greater potential of multi-omic datasets, the modalities need to be explicitly integrated, which can then substantially expand the insights that can be obtained from the individual modalities.

A range of integration approaches and tools have been developed for multi-omic single cell datasets. Some of these were developed for specific combinations of modalities, while others can handle a broad range of modality combinations. Here we outline the key distinguishing features of these tools noting that the selection of an optimal tool is largely determined by the nature of the data and the biological insights that are sought.

Data from multiple modalities can be from the same cells (paired data) or from different but similar cells (unpaired data) (Argelaguet et al., 2021). Paired data provides a direct mapping between the modalities at the cell-to-cell level, and thus allow the most direct integration. Obtaining paired data, however, may not be straightforward or even possible for all combinations of modalities. Additionally, datasets from altogether different origins may need to be combined, such as scATAC-seq and scRNA-seq of PBMCs from different laboratories on different dates. Consequently, there is a need for integration approaches that can handle both paired and unpaired datasets, and computational methods have been developed for handling either situation.

Another principal distinction between various integration methods is the analysis stage when integration is performed: early stage and late stage (Miao et al., 2021a). For early-stage integration methods, the different modality datasets are integrated at the beginning of the analyses (Figure 2). Effectively this creates a new, hybrid modality dataset that can then be used for downstream analysis, thus enabling application of other analysis tools commonly applied to unimodal data. Whereas single cell data is typically high dimensional (e.g., each gene in a scRNA-seq dataset is represented by one dimension), the hybrid dataset is low dimensional. For example, the TotalVI method for integrating transcriptomic and protein data uses a neural network to place the combined data in a 20-dimensional reduced space (Gayoso et al., 2021). Downstream analysis, such as differential expression analysis, can be performed in this space. The generative component of the neural network can be used to relate the results in the reduced dimensionality space back to the original RNA and protein identities. These early-stage approaches lend themselves well to integration of unpaired assays as the hybrid space can accommodate data from either assay type (Miao et al., 2021b).

In contrast, late-stage integration does not explicitly combine the different modality datasets, but rather combines analysis results obtained for the individual modalities. For example, a typical step in single cell analysis is the construction of a neighbor-graph that



describes the similarities of the cells. The neighbor-graph is then used as input for clustering analysis or reduced dimensionality visualization, such as a uniform manifold approximation and projection (UMAP) or t-distributed stochastic neighbor embedding (t-SNE) map. In a multi-modal context, neighbor-graphs can be constructed for the individual assays, which can then be merged and subsequently used for integrated clustering analysis, visualization, and possibly cell type and state annotation (Figure 2) (Hao et al., 2021).

A third example, which can be regarded as late stage as well, is the identification of subpopulations using one modality, which are then extracted and further analyzed using the second modality (Figure 2). This allows for cell type or state specific analysis of the second modality. As an example, cell types can be determined using a transcriptomic assay (e.g., scRNA-seq), followed by cell type-specific analysis of surface protein abundance (e.g., using CITE-seq) or regulatory landscapes (e.g., using scATAC-seq).

Finally, methods are either designed to handle specific combinations of modalities (see next paragraphs for examples), or for any combination of modalities. Examples of the latter that have been used extensively include the Weighted Nearest Neighbor (WNN) approach implemented in the widely used Seurat package

(Hao et al., 2021), MOFA + which constructs a hybrid modality from paired modalities (Argelaguet et al., 2020), and LIGER which constructs a hybrid modality from unpaired data (Welch et al., 2019). While many challenges exist to improve multi-omic data integration methods, one key challenge is the inherent correction for differences between the modalities, including batch effects. More so than batch correction in unimodal single cell datasets it is a challenge to distinguish technical noise from biological variability. Although not many benchmarking studies are available, a recent study showed that WNN is particularly successful at integrating scRNA-seq and snATAC-seq data even in the presence of complex batch effects (Lee et al., 2023).

## 2.2 Transcriptome and surface proteome

The availability of commercial kits and protocols has increased the accessibility of combined profiling of the transcriptome and surface proteome. Computational approaches for the integration of paired transcriptome and surface proteome data have been increasingly needed for analyzing these datasets.

An example of a late-stage integration approach developed specifically for transcriptomic data and the surface proteome is



CiteFuse (Kim et al., 2020). Cell-to-cell similarity matrices for both modalities are fused, which can then be used for clustering and dimensionality reduction. Results demonstrate that the integrated data enhances cell typing and other downstream analyses relative to single modality data (Kim et al., 2020). In contrast to CiteFuse, TotalVI (Gayoso et al., 2021) and sciPENN (Lakkis et al., 2022) are early-stage integration approaches and are based on neural network machine learning. Although these approaches were designed for integration of paired data, they do allow different cells to have different antibody panels in the protein modality. The neural network is used to impute the missing protein data, in effect relaxing the requirement of the data to be paired. This feature allows for adjusting the antibody panel in the lifetime of a project, while retaining the ability to integrate data obtained from different stages of the project.

### 2.3 Transcriptome and chromatin accessibility

Integration of transcriptomic and ATAC data is a special case as it can take advantage of 'feature correspondence' (Rautenstrauch et al., 2022). The scATAC peaks can be related to specific genes, and thus merged with the transcriptomic data. The result is a hybrid modality that can be used for downstream analysis, but unlike most other hybrid modalities, retains the ability for interpretation in the context of the genes. Many algorithms have been developed specifically for integration of transcriptomic data with scATAC-seq (see (Rautenstrauch et al., 2022) for an overview of the approaches). These various methods employ early- and late-stage frameworks and operate on both paired and unpaired data.

The combination of scRNA-seq and scATAC-seq modalities has proven useful in various ways, as illustrated in a recent study on mouse kidneys (Miao et al., 2021a). First, the combined modalities were able to define more distinct cell types compared to scRNA-seq analysis. Second, as an example of late-stage integration of scRNA-seq and scATAC modalities, the scRNA-seq and/or scATAC-seq modality can be used to annotate the cells, allowing identification of cell-type specific regulatory elements. Gene regulatory networks can then be inferred by relating active genes to active regulatory elements in each cell type (Cao et al., 2018).

### 2.4 Outlook on multi-omic integration methods

With the advent of single-cell methods came efforts to build cell atlases for tissues or even entire species (Regev et al., 2017). More recently, large datasets or reference maps of multi-modal data are also becoming available (Network, 2021). This introduced the need for multi-modal mapping of query datasets to reference maps. Several approaches have been developed specifically for this purpose (Lotfollahi et al., 2022a; Lotfollahi et al., 2022b), while other recent, more general methods are able to map very large datasets that involve different modalities (Hao et al., 2022). Development of such approaches will further expand the use of multi-omic single cell sequencing.

Community efforts such as the Multimodal single cell data integration challenge (Lance et al., 2022) can play important roles in guiding further method development. Notably, the

competition winner, as well as recently developed integration approaches such as bridge learning (Hao et al., 2022), both lessen the distinction of paired and unpaired dataset: Small paired datasets are used to guide integration of the unpaired datasets, with an overall improved performance.

## 3 Overview of single cell studies in kidney diseases

Advances in single-cell sequencing technologies have enabled the analysis of individual cells obtained from tissues, such as kidney biopsies, and biofluids, such as urine and blood, at an unprecedented resolution, revealing cellular signatures of inflammation, cellular injury, and fibrosis in various kidney diseases. These signatures enable a deeper understanding of pathophysiology and can also facilitate development of precision therapeutics for these diseases. Single-cell profiling of kidney biopsies, PBMCs, urine samples and skin lesions from patients with lupus nephritis, acute kidney injury, diabetic nephropathy and focal segmental glomerulosclerosis present potential novel approaches for the diagnosis and monitoring of disease activity (Table 1). These approaches, when performed on urinary cells and PBMCs, in contrast to kidney biopsy, are non-invasive and could be repeated multiple times as needed. Several examples of studies performed using human samples are discussed below.

### 3.1 Lupus nephritis (LN)

LN is a form of glomerulonephritis that constitutes one of the most severe organ manifestations of systemic lupus erythematosus (SLE). Despite increased knowledge of disease pathogenesis and improved treatment options, LN remains a substantial cause of morbidity and death among patients with SLE. Conventional markers of disease activity and response to therapy typically consist of measuring auto-antibody levels, markers of complement activity and laboratory parameters like proteinuria and estimated glomerular filtration rate (eGFR).

Single cell transcriptomics and urine proteomics were used to identify biomarkers that are upregulated in LN, including IL-16, which is found to be highly expressed at sites of kidney injury (Fava et al., 2022). Other studies using scRNA-seq on kidney biopsy and urinary cells have found evidence of upregulation of IFN response genes, fibrotic ECM proteins and chemokine receptors such as CXCR4 and CX3CR1 (Arazi et al., 2019; Der et al., 2019). These findings help our understanding of the pathogenesis of a complex disease like LN and may further help in predicting patients who may better respond to specific therapies.

### 3.2 Acute kidney injury (AKI)

AKI is a major health issue, the outcome of which depends primarily on damage and reparative processes of tubular epithelial cells. According

**TABLE 1** Examples of multi-omic single cell sequencing studies in kidney diseases.

Condition	Technique	Source of samples	Key findings	Reference	
Lupus Nephritis	scRNA-seq and urine proteomics	Kidney biopsy & Urine	<ul style="list-style-type: none"> <li>Identified 237 urinary biomarkers associated with lupus nephritis as compared to controls. Amongst these biomarkers IL-16, CD163, and TGF-<math>\beta</math> mirrored intrarenal nephritis activity</li> <li>IL-16 was found to be highly expressed at key sites of kidney injury and highlight its role in LN pathogenesis and a potentially treatable target and biomarker</li> </ul>	Fava et al. (2022)	
			<ul style="list-style-type: none"> <li>The analysis identified multiple subsets of leukocytes including myeloid cells, T cells, natural killer cells and B cells that demonstrated both pro and anti-inflammatory properties, such as CD16<sup>+</sup> macrophages, CD4<sup>+</sup> TFH cells and FoxP3+Helios + regulatory T cells</li> <li>Two chemokine receptors, CXCR4 and CX3CR1, were broadly expressed, suggesting a key role in cell trafficking</li> <li>Interferon response was observed in dividing CTLs and NK cells</li> </ul>	Arazi et al. (2019)	
	scRNA-seq	Kidney & Skin biopsy	<ul style="list-style-type: none"> <li>Type I IFN-response genes in tubular cells and keratinocytes were much more highly expressed than those of healthy controls and was higher in non-responders, compared to treatment responders</li> <li>Patients non-responsive to treatment demonstrate higher expression of fibrotic ECM proteins compared with responders</li> </ul>	Der et al. (2019)	
			<ul style="list-style-type: none"> <li>Shed podocytes in the urine had loss of canonical podocyte markers, which are required for normal podocyte function, such as NPHS1, NPHS2, and PODXL.</li> <li>Patients with FSGS had higher expression of genes for epithelial-to-mesenchymal transition (EMT) compared to patients with MCD. These markers were relatively higher in treatment non-responders compared to responders</li> </ul>	Latt et al. (2022)	
	FSGS	scRNA-seq	Urine	<ul style="list-style-type: none"> <li>Highest glomerular endothelial cells (GEC) scores were observed in patients with FSGS</li> <li>Molecular endothelial signatures suggested 2 distinct FSGS patient subgroups with <math>\alpha</math>-2 macroglobulin (A2M) as a key downstream mediator of the endothelial cell phenotype and was shown to have prognostic significance</li> </ul>	Menon et al. (2020)
				<ul style="list-style-type: none"> <li>Patients with pre-renal AKI excreted mainly myeloid cells, whereas urine samples from patients with AKI had higher expression of immune and epithelial cells</li> <li>Tubular epithelial cells from urine samples of patients with AKI did not express characteristic segment markers, and instead showed injury-related dedifferentiation and adaptive phenotypes</li> </ul>	Klocke et al. (2022)
AKI	scRNAseq, snRNAseq and spatial transcriptomics	Kidney biopsy	<ul style="list-style-type: none"> <li>This study localized the transcriptomic signature of various immune cells to spatial transcriptomic spots of known renal epithelial cells in ischemia reperfusion injury and cecal ligation puncture models of AKI</li> <li>A subpopulation of injured proximal tubule cells with Activating Transcription Factor 3 (ATF3) and Midkine (Mdk) expression were identified, which may be responsible for neutrophil chemotaxis to the site of injury</li> </ul>	Melo Ferreira et al. (2021)	
			<ul style="list-style-type: none"> <li>This study enabled mapping of all matrix-producing cells at high resolution, revealing distinct subpopulations of pericytes and fibroblasts as the major cellular sources of scar forming myofibroblasts during fibroblasts</li> <li>Using these data, myofibroblast-expressed Nkd2 was identified as a potential therapeutic target</li> </ul>	Kuppe et al. (2021)	
Fibrosis	scRNA-seq, ATAC-seq and spatial transcriptomics	Human & mouse kidney biopsy	<ul style="list-style-type: none"> <li>This study enabled mapping of all matrix-producing cells at high resolution, revealing distinct subpopulations of pericytes and fibroblasts as the major cellular sources of scar forming myofibroblasts during fibroblasts</li> <li>Using these data, myofibroblast-expressed Nkd2 was identified as a potential therapeutic target</li> </ul>	Kuppe et al. (2021)	

(Continued on following page)

TABLE 1 (Continued) Examples of multi-omic single cell sequencing studies in kidney diseases.

Condition	Technique	Source of samples	Key findings	Reference
Diabetic kidney Disease	snRNA-seq	Kidney biopsy	<ul style="list-style-type: none"> <li>• snRNAseq performed on cryopreserved human diabetic kidney samples.</li> <li>• It was seen that diabetic thick ascending limb, late distal convoluted tubule, and principal cells adopted a gene expression signature consistent with increased potassium secretion, including alterations in Na<sup>+</sup>/K<sup>+</sup>-ATPase, WNK1, mineralocorticoid receptor, and NEDD4L expression, along with increases angiogenic signaling</li> </ul>	Wilson et al. (2019)
	scRNA-seq and spatial transcriptomics	Kidney biopsy	<ul style="list-style-type: none"> <li>• There was enrichment of specific cell subpopulations consisting of venous endothelial cells and fibroblasts with elevated expression of CCL21 and IGFBP5.</li> <li>• Spatial analysis revealed that most of the immune cells were localized in areas of renal fibrosis</li> </ul>	Chen et al. (2022)
	snRNA-seq & snATAC-seq	Kidney biopsy	<ul style="list-style-type: none"> <li>• snRNA-seq and snATACseq performed on human DKD and non-diabetic kidney biopsy samples and DKD was associated with an increased proportion of VCAM1+ proximal tubule cells (PT_VCAM1) and infiltrating leukocytes compared to non-diabetics</li> <li>• PT_VCAM1 cell is pro-inflammatory phenotype characterized by enhanced NFkB signaling and failed repair that may underlie transition from acute kidney injury to CKD</li> </ul>	Wilson et al. (2022)

to a recent meta analysis, 1 in 5 adults (21.6%) and 1 in 3 children (33.7%) experienced AKI worldwide (Susantitaphong et al., 2013). Mechanisms underlying AKI remain incompletely understood, effective therapies are lacking and monitoring the course of AKI in clinical routine is limited to measuring urine output and plasma levels of filtration markers. Hence, high-resolution approaches are needed to facilitate a better understanding of the pathogenesis of AKI and to potentially identify therapeutic targets in preventing and treating AKI.

In a recent study (Klocke et al., 2022), scRNA-seq analysis revealed that urinary cells from patients with established AKI had different transcriptional profiles compared to patients with pre-renal AKI, including markers of cellular dedifferentiation. Although single cell sequencing studies have improved our understanding of the transcriptomic signature of different cell types within the kidney, the spatial distribution of injury can be limited to certain regions in the kidney. In another study (Melo Ferreira et al., 2021), investigators were able to localize the transcriptomic signature of various immune cells to spatial transcriptomic spots of known renal epithelial cells in murine models of AKI. The analysis was able to detect a subpopulation of injured proximal tubule cells with Activating Transcription Factor 3 (ATF3) expression which may be responsible for neutrophil chemotaxis to the site of injury.

### 3.3 Focal segmental glomerulosclerosis (FSGS)

FSGS and minimal change disease (MCD) are common causes of nephrotic syndrome and share many common features, such as diffusely effaced podocytes. However, the response to treatment is variable in FSGS, and there is a higher risk of progression to chronic

kidney disease (CKD) in patients with FSGS, compared to patients with MCD, who typically have a benign course. Multiple single cell studies (Menon et al., 2020; Latt et al., 2022) have identified specific markers such as  $\alpha$ -2 macroglobulin (A2M), elevated glomerular endothelial cell score and loss of canonical podocyte markers in patients with FSGS compared to patients with MCD. These findings may help in differentiating FSGS from MCD and in predicting response to treatment (Menon et al., 2020).

### 3.4 Deconvolution of kidney fibrosis

Fibrosis is a characteristic feature in all forms of CKD. Deposition of pathological matrix in the interstitial space and within the walls of glomerular capillaries as well as the cellular processes resulting in this deposition are increasingly recognized as principal factors resulting in progressive kidney damage. It has been challenging to study kidney fibrosis in patients, since kidney biopsies are usually not performed in patients with established CKD. Use of multi-omic techniques (Kuppe et al., 2021) have helped in identifying a population of scar-forming myofibroblasts. These techniques may help in identifying potential therapeutic targets for preventing fibrosis, such as the myofibroblasts identified in the aforementioned study.

### 3.5 Diabetic kidney disease (DKD)

DKD has a high global disease burden and substantially increases the risk of kidney failure and cardiovascular events. Despite treatment, there is substantial residual risk of disease progression with existing therapies. There is an urgent need to better understand the molecular



mechanisms driving DKD to help identify new therapies that slow progression and reduce associated risks (Tuttle et al., 2022). snRNA-seq analysis of cryopreserved diabetic kidney samples showed upregulation of angiogenic and mineralocorticoid markers, consistent with the clinical manifestation of fluid overload and neo angiogenesis seen in patients with DKD (Wilson et al., 2019). Multi-omic techniques incorporating scRNA-seq and spatial transcriptome analyses have been used to generate an atlas of diabetic kidney disease. For example, in one study the investigators identified enrichment of specific cell subpopulations consisting of venous endothelial cells and fibroblasts with elevated expression of CCL21 and IGFBP5 (Chen et al., 2022). Furthermore, spatial analysis revealed that most of the immune cells were localized in areas of renal fibrosis. In another recent study, the investigators performed snRNA-seq, snATACseq and spatial transcriptomics on human DKD kidney biopsy samples and were able to identify increased proportion of VCAM1+ proximal tubule cells (PT\_VCAM1) and infiltrating leukocytes compared to non-diabetics. These changes have adverse implications as they increase the pro-inflammatory milieu and facilitate AKI to CKD transition (Wilson et al., 2022). Thus, single cell sequencing technologies are helping to elucidate the underlying mechanisms of DKD.

## 4 Therapeutic development for kidney disease in single cell era

### 4.1 Target discovery: Genetic variants and gene perturbation

Single cell sequencing provides rich and complex phenotypic portraits of the cellular and molecular circuits involved in disease. These large datasets lead to the identification of an array of cellular and molecular features which are upregulated, downregulated or, in some cases, unique to disease states. Differences observed in disease vs. control samples provide hypotheses of new therapeutic targets, which may have become dysregulated at the cellular or molecular level. Nevertheless, challenges remain in deciphering which observed changes drive pathogenic processes versus those that are passenger effects.

Genetic variants and their loci from genome-wide association studies (GWAS) provide complementary evidence to single cell sequencing datasets for the involvement of genes in the development of diseases. A challenge, however, in identifying specific targets from genomic variant studies is that the large majority of disease-associated loci are in non-coding regions, thus complicating their linkage to specific gene products or cell types that could be candidates for therapies. By integrating single cell sequencing datasets with GWAS candidate genes/loci, greater linkages can be established to identify genes, pathways or cell types with causal relationships to disease. In one analytical framework, genes near disease-associated loci from GWAS studies are compared to various cell subsets identified by scRNA-seq to nominate disease-relevant cell subtypes and pathways. This framework led to the identification of cell clusters enriched for expression of TLR7 (nucleic acid sensing), HIP1 (endocytic participating protein implicated in DC regulation), and LBH (modulates synovial hyperplasia) in LN patients (Arazi et al., 2019). When multiple genes in a shared pathway are implicated

for a disease, a gene signature incorporating a panel of genes, rather than an individual gene, can be used to identify potential disease-relevant cell types and pathways in single cell data. For example, multiple loci related to the type I interferon pathway have been implicated in SLE disease activity by GWAS studies (Rice et al., 2017; Psarras et al., 2022) and gene signature modules of interferon-stimulated genes (ISG) have been shown to correlate with SLE disease activity. When these modules were assessed for enrichment in cell clusters identified in pediatric SLE patients by scRNA-seq, multiple distinct subpopulations were identified (Nehar-Belaid et al., 2020).

Since functionally important gene regulatory regions are mostly nucleosome-free, chromatin accessibility data can be used to aid understanding of candidate risk loci in non-coding regions. Additionally, expressed quantitative loci (eQTL) approaches are commonly used to map variants to causal genes. However, a critical limitation in chromatin accessibility and eQTL analyses for mapping causal genes arises from cell type heterogeneity; bulk methods capture aggregated expression or accessibility across multiple cell types, with cell type diversity and proportionality complicating interpretation [reviewed in (Maria et al., 2022)]. In a pioneering study for kidney disease, Sheng, et al. used single cell multi-omics (scRNA-seq and snATAC-seq), human genetic information, and advanced computational approaches to demonstrate how genetic variants render a functional effect on cell types and specific gene/pathway programs, resulting in the identification of more than 200 genes involved in kidney function and hypertension (Sheng et al., 2021). More recent studies have expanded on this work, providing further mapping of risk alleles to specific cell types, pathways, and genes (Liu et al., 2022; Sandholm et al., 2022), which can augment identification of new targets (Doke et al., 2021a; Doke et al., 2021b). In another study, whole kidney and single cell epigenomic information of hundreds of samples was used with GWAS data to define the genetic association of kidney function in 1.5 million individuals, resulting in identification of 878 loci (126 novel), with prioritized target genes for 87% of the loci (Liu et al., 2022). Collectively, the results have provided meaningful insights to pathologic cell types and disease-causing pathways, which can guide identification of new therapeutic targets.

To further facilitate target identification and drug discovery, single cell profiling has been combined with gene perturbation methods as a forward genetics screening approach to explore phenotypic impacts of gene modulation. Perturb-seq and related approaches integrate pooled CRISPR screening with single cell profiling, enabling systematic determination of the impact of inhibitory and activating perturbations to large numbers of candidate genes (Adamson et al., 2016; Dixit et al., 2016; Jaitin et al., 2016). The phenotypic readouts can identify cell states that are desirable for therapeutic intervention, and can be applied in more complex biological systems, including co-cultures, organoids and in animal models. Perturb-seq methods have been extended to understand the role of non-coding genetic variants associated with disease (Gasperini et al., 2019) and, in combination with CITE-seq, define mechanisms of cancer immune evasion (Frangieh et al., 2021). These forward genetic methods provide complementary

information for a deeper phenotypic understanding of the impact of candidate genes as potential therapeutic targets.

## 4.2 High-resolution diagnostics and biomarkers in kidney disease

A high-resolution picture of all cell types from a disease tissue sample provides a comprehensive and deep interpretation of the underlying biology and should be informative to diagnostics and prognostics. Single cell sequencing can provide a comprehensive capturing of all cell types, as well as information regarding their molecular pathways, thereby providing greater detail of disease endotypes and potential for therapy responsiveness. Indeed, studies have used single-cell profiling of peripheral blood immune cells to identify correlates of anti-PD1 responsiveness in cancer patients (Wu et al., 2020; Luoma et al., 2022).

For diagnostics and biomarkers, a high-resolution molecular picture of the kidney would be of great value, however, a key challenge in molecular diagnostics, including single cell profiling, for kidney disease is the limited ability to obtain kidney biopsy samples for such analyses. As performing a kidney biopsy carries some risk of complications for patients, they are commonly performed at time of diagnosis and at limited additional timepoints, as needed for patient care. Core biopsies contain a relatively low number of cells (single digit thousands for single cell profiling). Moreover, the tissue must be dissociated to single cell suspension, which can introduce artifacts, such as protease-induced changes in gene expression, and lead to biased loss of cell types. To address these key challenges in diagnostics and biomarkers, emerging studies have begun to analyze surrogate tissue sources of cells from urine for single cell profiling, as well as spatial profiling from limited tissue derived from kidney biopsy cores.

Urine sedimentary analysis has been used for decades to inform diagnoses for kidney diseases, however, until the development of single cell profiling approaches, robust and accessible methods to generate an unbiased determination of cell types and their molecular features were lacking. Single cell profiling of urinary cells from DKD patients demonstrated an ability to detect nearly all cell types of the kidney (Abedini et al., 2021). Emerging studies applying single cell sequencing to urinary cells in FSGS (Latt et al., 2022), lupus nephritis (Arazi et al., 2019), and AKI (Cheung et al., 2022; Klocke et al., 2022) have underscored the potential of using this non-invasive cell source for diagnostic and other disease insights. TCR analyses of patients with immune checkpoint-associated nephritis revealed the T cell clonotypes in the kidney are enriched in urine, supporting a direct linkage between T cells in urine and kidney (Singh et al., 2022). In a recent study of acute cell rejection (ACR) of kidney transplantation, investigators applied single cell RNA and TCR sequencing to biopsies from allografts as well as urinary cells (Shi et al., 2023). Interestingly, TCR sequences associated with the expanded CD8 T cell population were also observed in matching urine samples. These results underscore a linkage between immune cellular phenotypes from urine and kidney tissue, and relate this information to treatment response. Collectively, while initial studies showcase the opportunity to be harnessed by applying

single cell sequencing to urinary cells, additional research is needed to better define the relationships between urinary cells and disease biology and progression.

Spatially resolved omic profiling methods, largely rooted in transcriptomics, address multiple challenges with the limited tissue from kidney biopsies. These methods do not require dissociation of tissue, thereby retaining the spatial biological context of cells and eliminating sample biases and loss of cells from dissociation methods. Furthermore, spatial profiling methods often require limited tissue amounts and can be compatible with formalin-fixed, paraffin-embedded (FFPE) tissues, making them more amenable to the limited material of kidney core biopsies from clinical settings that use FFPE as standard of practice. Commercialization of spatial transcriptomic methods has reduced the technical barriers of access to these complex methods, with FFPE-compatible platforms including from 10x Genomics (Visium and Xenium), Nanostring (GeoMx, nCounter, CosMx), Vizgen (MERSCOPE), among others. The technologies are rooted in one of two fundamental methods: *in situ* hybridization (ISH) and next-generation sequencing (NGS)-based methods. ISH methods can provide subcellular localization information but are limited to a selected (and therefore biased) set of several hundred gene probes. In contrast, NGS-based spatial transcriptomics methods combine scRNA-seq technology with spatial barcodes on a specialized slide, enabling amplification of copied transcripts while retaining localization information. The NGS-method provides unbiased, genome-wide transcriptomics but is limited in spatial resolution due to the spot size on the slide. The current 10x Visium methodology incorporates spot sizes of 55  $\mu\text{m}$ , which falls short of single-cell resolution. To improve the resolution, separate conventional (non-spatial) scRNA-seq data can be integrated with the spatial data, and combined with computational deconvolution methods, used to estimate single cell contributions (see reviews (Melo Ferreira et al., 2021; Rao et al., 2021; Zhang et al., 2023)). Studies incorporating spatial transcriptomic profiling methods on human kidney samples are emerging, with examples including AKI to understand immune cell infiltration in histological context (Melo Ferreira et al., 2021), development of an atlas across multiple kidney diseases (Lake et al., 2021), cell-mediated rejection in kidney transplantation (Salem et al., 2022), and small RNA involvement in FSGS (Williams et al., 2022). Multi-omic single cell approaches incorporating spatial profiling have been reported, for example, to understand fibrosis microenvironments in diabetic and hypertensive diseased human kidneys (Abedini et al., 2022). Spatial profiling methods are a powerful emerging technology compatible with kidney biopsy samples that will bring greater understanding of disease processes and treatments.

## 5 Summary and outlook

Multi-omic single cell sequencing has facilitated the creation of high-resolution cellular and molecular maps in the context of kidney disease, providing new insights into disease mechanisms and opportunities for therapeutic intervention. New experimental methods enable measurement of multiple molecular features simultaneously, including gene expression, surface protein expression, TCR/BCR sequences and chromatin accessibility. Computational

tools to integrate these large and diverse datasets have enabled derivation of rich biological insights, from pathogenic mechanisms to new therapeutic targets. And while most studies applying single cell sequencing methods in the context of kidney diseases have utilized single modality datasets (i.e., gene expression), increasingly multi-omic approaches are being pursued.

Single cell profiling is increasingly accessible to researchers, due to developments of more user-friendly experimental workflows (e.g., with commercial kits) and data analysis tools (e.g., graphical interface software). However, tissue availability from kidney core biopsies has been a major limitation for single cell profiling in the context of kidney disease. Emerging alternative approaches compatible with single cell profiling are being developed, including urinary cell analysis and spatial profiling. With studies in which nearly all kidney cell types can be identified in urine by single cell profiling, the potential for routine, non-invasive monitoring of biological changes of individual diseased patients is an exciting one, with ramifications for more optimal treatment selections (e.g., therapeutic classes and doses). These developments portend a future of collection of these high-resolution datasets in a variety of clinical settings, from observational studies to therapeutic interventional trials. Collectively, these advancements are heralding new opportunities for precision medicine in kidney disease, from diagnostics and patient segmentation to prognostics and new targeted therapies, with

the potential to better match the right therapies with the right patients at the right time.

## Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

## Acknowledgments

Figures were created with [Biorender.com](https://biorender.com).

## Conflict of interest

SP, TV, MM, and LR were employees of Visterra Inc.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Abedini, A., Ma, Z., Frederick, J., Dhillon, P., Balzer, M. S., Shrestha, R., et al. (2022). Spatially resolved human kidney multi-omics single cell atlas highlights the key role of the fibrotic microenvironment in kidney disease progression. *bioRxiv* 2022, 513598. doi:10.1101/2022.10.24.513598
- Abedini, A., Zhu, Y. O., Chatterjee, S., Halasz, G., Devalaraja-Narashimha, K., Shrestha, R., et al. (2021). Urinary single-cell profiling captures the cellular diversity of the kidney. *J. Am. Soc. Nephrol.* 32 (3), 614–627. doi:10.1681/ASN.2020050757
- Adam, M., Potter, A. S., and Potter, S. S. (2017). Psychrophilic proteases dramatically reduce single-cell RNA-seq artifacts: A molecular atlas of kidney development. *Development* 144 (19), 3625–3632. doi:10.1242/dev.151142
- Adamson, B., Norman, T. M., Jost, M., Cho, M. Y., Nunez, J. K., Chen, Y., et al. (2016). A multiplexed single-cell CRISPR screening platform enables systematic dissection of the unfolded protein response. *Cell* 167 (7), 1867–1882.e21. doi:10.1016/j.cell.2016.11.048
- Arazi, A., Rao, D. A., Berthier, C. C., Davidson, A., Liu, Y., Hoover, P. J., et al. (2019). The immune cell landscape in kidneys of patients with lupus nephritis. *Nat. Immunol.* 20 (7), 902–914. doi:10.1038/s41590-019-0398-x
- Argelaguet, R., Arnol, D., Bredikhin, D., Deloro, Y., Velten, B., Marioni, J. C., et al. (2020). MOFA+: A statistical framework for comprehensive integration of multi-modal single-cell data. *Genome Biol.* 21 (1), 111. doi:10.1186/s13059-020-02015-1
- Argelaguet, R., Cuomo, A. S. E., Stegle, O., and Marioni, J. C. (2021). Computational principles and challenges in single-cell data integration. *Nat. Biotechnol.* 39 (10), 1202–1215. doi:10.1038/s41587-021-00895-7
- Buenrostro, J. D., Corces, M. R., Lareau, C. A., Wu, B., Schep, A. N., Aryee, M. J., et al. (2018). Integrated single-cell analysis maps the continuous regulatory landscape of human hematopoietic differentiation. *Cell* 173 (6), 1535–1548.e16. doi:10.1016/j.cell.2018.03.074
- Buenrostro, J. D., Wu, B., Litzenburger, U. M., Ruff, D., Gonzales, M. L., Snyder, M. P., et al. (2015). Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* 523 (7561), 486–490. doi:10.1038/nature14590
- Cao, J., Cusanovich, D. A., Ramani, V., Aghamirzaie, D., Pliner, H. A., Hill, A. J., et al. (2018). Joint profiling of chromatin accessibility and gene expression in thousands of single cells. *Science* 361 (6409), 1380–1385. doi:10.1126/science.aau0730
- Cao, J., Packer, J. S., Ramani, V., Cusanovich, D. A., Huynh, C., Daza, R., et al. (2017). Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* 357 (6352), 661–667. doi:10.1126/science.aam8940
- Chen, A. F., Parks, B., Kathiria, A. S., Ober-Reynolds, B., Goronzy, J. J., and Greenleaf, W. J. (2022). NEAT-Seq: Simultaneous profiling of intra-nuclear proteins, chromatin accessibility and gene expression in single cells. *Nat. Methods* 19 (5), 547–553. doi:10.1038/s41592-022-01461-y
- Chen, D., Shao, M., Song, Y., Ren, G., Guo, F., Fan, X., et al. (2022). Combined single-cell RNA-seq and spatial transcriptomics reveal an atlas of human diabetic kidney disease. *SSRN*. [Preprint]. doi:10.2139/ssrn.4117371
- Chen, S., Lake, B. B., and Zhang, K. (2019). High-throughput sequencing of the transcriptome and chromatin accessibility in the same cell. *Nat. Biotechnol.* 37 (12), 1452–1457. doi:10.1038/s41587-019-0290-0
- Cheung, M. D., Erman, E. N., Liu, S., Erdmann, N. B., Ghajar-Rahimi, G., Moore, K. H., et al. (2022). Single-cell RNA sequencing of urinary cells reveals distinct cellular diversity in COVID-19-associated AKI. *Kidney360* 3 (1), 28–36. doi:10.34067/KID.0005522021
- Chung, H., Parkhurst, C. N., Magee, E. M., Phillips, D., Habibi, E., Chen, F., et al. (2021). Joint single-cell measurements of nuclear proteins and RNA *in vivo*. *Nat. Methods* 18 (10), 1204–1212. doi:10.1038/s41592-021-01278-1
- Chung, J. J., Goldstein, L., Chen, Y. J., Lee, J., Webster, J. D., Roose-Girma, M., et al. (2020). Single-cell transcriptome profiling of the kidney glomerulus identifies key cell types and reactions to injury. *J. Am. Soc. Nephrol.* 31 (10), 2341–2354. doi:10.1681/ASN.2020020220
- Cusanovich, D. A., Hill, A. J., Aghamirzaie, D., Daza, R. M., Pliner, H. A., Berletch, J. B., et al. (2018). A single-cell atlas of *in vivo* mammalian chromatin accessibility. *Cell* 174 (5), 1309–1324. doi:10.1016/j.cell.2018.06.052
- Der, E., Suryawanshi, H., Morozov, P., Kustagi, M., Goilav, B., Ranabothu, S., et al. (2019). Tubular cell and keratinocyte single-cell transcriptomics applied to lupus nephritis reveal type I IFN and fibrosis relevant pathways. *Nat. Immunol.* 20 (7), 915. doi:10.1038/s41590-019-0386-1
- Dixit, A., Parnas, O., Li, B., Chen, J., Fulco, C. P., Jerby-Aron, L., et al. (2016). Perturb-Seq: Dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell* 167 (7), 1853–1866. doi:10.1016/j.cell.2016.11.038
- Doke, T., Huang, S., Qiu, C., Liu, H., Guan, Y., Hu, H., et al. (2021a). Transcriptome-wide association analysis identifies DACH1 as a kidney disease risk gene that contributes to fibrosis. *J. Clin. Invest.* 131 (10), e141801. doi:10.1172/JCI141801

- Doke, T., Huang, S., Qiu, C., Sheng, X., Seasock, M., Liu, H., et al. (2021b). Genome-wide association studies identify the role of caspase-9 in kidney disease. *Sci. Adv.* 7 (45), eabi8051. doi:10.1126/sciadv.abi8051
- Fava, A., Rao, D. A., Mohan, C., Zhang, T., Rosenberg, A., Fenaroli, P., et al. (2022). Urine proteomics and renal single-cell transcriptomics implicate interleukin-16 in lupus nephritis. *Arthritis Rheumatol.* 74 (5), 829. doi:10.1002/art.42023
- Frangieh, C. J., Melms, J. C., Thakore, P. I., Geiger-Schuller, K. R., Ho, P., Luoma, A. M., et al. (2021). Multimodal pooled Perturb-CITE-seq screens in patient models define mechanisms of cancer immune evasion. *Nat. Genet.* 53 (3), 332–341. doi:10.1038/s41588-021-00779-1
- Gasparini, M., Hill, A. J., McFaline-Figueroa, J. L., Martin, B., Kim, S., Zhang, M. D., et al. (2019). A genome-wide framework for mapping gene regulation via cellular genetic screens. *Cell* 176 (1–2), 1516–2390 e319. doi:10.1016/j.cell.2019.02.027
- Gayoso, A., Steier, Z., Lopez, R., Regier, J., Nazor, K. L., Streets, A., et al. (2021). Joint probabilistic modeling of single-cell multi-omic data with totalVI. *Nat. Methods* 18 (3), 272–282. doi:10.1038/s41592-020-01050-x
- Gerlach, J. P., van Buggenum, J. A. G., Tanis, S. E. J., Hogeweg, M., Heuts, B. M. H., Muraro, M. J., et al. (2019). Combined quantification of intracellular (phospho-) proteins and transcriptomics from fixed single cells. *Sci. Rep.* 9 (1), 1469. doi:10.1038/s41598-018-37977-7
- Grindberg, R. V., Yee-Greenbaum, J. L., McConnell, M. J., Novotny, M., O'Shaughnessy, A. L., Lambert, G. M., et al. (2013). RNA-sequencing from single nuclei. *Proc. Natl. Acad. Sci. U. S. A.* 110 (49), 19802–19807. doi:10.1073/pnas.1319700110
- Habib, N., Avraham-Davidi, I., Basu, A., Burks, T., Shekhar, K., Hofree, M., et al. (2017). Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat. Methods* 14 (10), 955–958. doi:10.1038/nmeth.4407
- Habib, N., Li, Y., Heidenreich, M., Swiech, L., Avraham-Davidi, I., Trombetta, J. J., et al. (2016). Div-Seq: Single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons. *Science* 353 (6302), 925–928. doi:10.1126/science.aad7038
- Hagemann-Jensen, M., Ziegenhain, C., Chen, P., Ramskold, D., Hendriks, G. J., Larsson, A. J. M., et al. (2020). Single-cell RNA counting at allele and isoform resolution using Smart-seq3. *Nat. Biotechnol.* 38 (6), 708–714. doi:10.1038/s41587-020-0497-0
- Hahaut, V., Pavlinic, D., Carbone, W., Schuierer, S., Balmer, P., Quinodoz, M., et al. (2022). Fast and highly sensitive full-length single-cell RNA sequencing using FLASH-seq. *Nat. Biotechnol.* 40 (10), 1447–1451. doi:10.1038/s41587-022-01312-3
- Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W. M., 3rd, Zheng, S., Butler, A., et al. (2021). Integrated analysis of multimodal single-cell data. *Cell* 184 (13), 3573–3587.e29. doi:10.1016/j.cell.2021.04.048
- Hao, Y., Stuart, T., Kowalski, M., Choudhary, S., Hoffman, P., Hartman, A., et al. (2022). Dictionary learning for integrative, multimodal, and scalable single-cell analysis. *bioRxiv*. 2022, 481684. doi:10.1101/2022.02.24.481684
- Hwang, B., Lee, D. S., Tamaki, W., Sun, Y., Ogorodnikov, A., Hartoularos, G. C., et al. (2021). SCITO-Seq: Single-cell combinatorial indexed cytometry sequencing. *Nat. Methods* 18 (8), 903–911. doi:10.1038/s41592-021-01222-3
- Jaitin, D. A., Weiner, A., Yofe, I., Lara-Astiaso, D., Keren-Shaul, H., David, E., et al. (2016). Dissecting immune circuits by linking CRISPR-pooled screens with single-cell RNA-seq. *Cell* 167 (7), 1883–1896.e15. doi:10.1016/j.cell.2016.11.039
- Kartha, V. K., Duarte, F. M., Hu, Y., Ma, S., Chew, J. G., Lareau, C. A., et al. (2022). Functional inference of gene regulation using single-cell multi-omics. *Cell Genom* 2 (9), 100166. doi:10.1016/j.xgen.2022.100166
- Kim, H. J., Lin, Y., Geddes, T. A., Yang, J. Y. H., and Yang, P. (2020). CiteFuse enables multi-modal analysis of CITE-seq data. *Bioinformatics* 36 (14), 4137–4143. doi:10.1093/bioinformatics/btaa282
- Klein, A. M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., et al. (2015). Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* 161 (5), 1187–1201. doi:10.1016/j.cell.2015.04.044
- Klocke, J., Kim, S. J., Skopnik, C. M., Hinze, C., Boltengagen, A., Metzke, D., et al. (2022). Urinary single-cell sequencing captures kidney injury and repair processes in human acute kidney injury. *Kidney Int.* 102 (6), 1359–1370. doi:10.1016/j.kint.2022.07.032
- Krishnaswami, S. R., Grindberg, R. V., Novotny, M., Venepally, P., Lacar, B., Bhutani, K., et al. (2016). Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. *Nat. Protoc.* 11 (3), 499–524. doi:10.1038/nprot.2016.015
- Kuppe, C., Ibrahim, M. M., Kranz, J., Zhang, X., Ziegler, S., Perales-Paton, J., et al. (2021). Decoding myofibroblast origins in human kidney fibrosis. *Nature* 589 (7841), 281–286. doi:10.1038/s41586-020-2941-1
- Lacar, B., Linker, S. B., Jaeger, B. N., Krishnaswami, S. R., Barron, J. J., Kelder, M. J. E., et al. (2016). Nuclear RNA-seq of single neurons reveals molecular signatures of activation. *Nat. Commun.* 7, 11022. doi:10.1038/ncomms11022
- Lake, B. B., Ai, R., Kaeser, G. E., Salathia, N. S., Yung, Y. C., Liu, R., et al. (2016). Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* 352 (6293), 1586–1590. doi:10.1126/science.aaf1204
- Lake, B. B., Chen, S., Hoshi, M., Plongthongkum, N., Salamon, D., Knoten, A., et al. (2019). A single-nucleus RNA-sequencing pipeline to decipher the molecular anatomy and pathophysiology of human kidneys. *Nat. Commun.* 10 (1), 2832. doi:10.1038/s41467-019-10861-2
- Lake, B. B., Menon, R., Winfree, S., Hu, Q., Ferreira, R. M., Kalhor, K., et al. (2021). An atlas of healthy and injured cell states and niches in the human kidney. *bioRxiv* 2021, 454201. doi:10.1101/2021.07.28.454201
- Lakkis, J., Schroeder, A., Su, K., Lee, M. Y. Y., Bashore, A. C., Reilly, M. P., et al. (2022). A multi-use deep learning method for CITE-seq and single-cell RNA-seq data integration with cell surface protein prediction and imputation. *Nat. Mach. Intell.* 4 (11), 940–952. doi:10.1038/s42256-022-00545-w
- Lance, C., Luecken, M. D., Burkhardt, D. B., Cannoodt, R., Rautenstrauch, P., Laddach, A., et al. (2022). Multimodal single-cell data integration challenge: Results and lessons learned. *bioRxiv* 2022, 487796. doi:10.1101/2022.04.11.487796
- Latt, K. Z., Heymann, J., Jesse, J. H., Rosenberg, A. Z., Berthier, C. C., Arazi, A., Eddy, S., Yoshida, T., Zhao, Y., Chen, V., Nelson, G. W., Cam, M., Kumar, P., Mehta, M., Kelly, M. C., Kretzler, M., et al. (2022). Urine single-cell RNA sequencing in focal segmental glomerulosclerosis reveals inflammatory signatures. *Kidney Int. Rep.* 7 (2), 289–304. doi:10.1016/j.ekir.2021.11.005
- Lee, M. Y. Y., Kaestner, K. H., and Li, M. (2023). Benchmarking algorithms for joint integration of unpaired and paired single-cell RNA-seq and ATAC-seq data. *bioRxiv* 2023, 526609. doi:10.1101/2023.02.01.526609
- Liu, H., Doke, T., Guo, D., Sheng, X., Ma, Z., Park, J., et al. (2022). Epigenomic and transcriptomic analyses define core cell types, genes and targetable mechanisms for kidney disease. *Nat. Genet.* 54 (7), 950–962. doi:10.1038/s41588-022-01097-w
- Lotfollahi, M., Litinetskaya, A., and Theis, F. J. (2022a). Multigrade: Single-cell multi-omic data integration. *bioRxiv* 2022, 484643. doi:10.1101/2022.03.16.484643
- Lotfollahi, M., Naghipourfar, M., Luecken, M. D., Khajavi, M., Buttner, M., Wagenstetter, M., et al. (2022b). Mapping single-cell data to reference atlases by transfer learning. *Nat. Biotechnol.* 40 (1), 121–130. doi:10.1038/s41587-021-01001-7
- Luoma, A. M., Suo, S., Wang, Y., Gunasti, L., Porter, C. B. M., Nabili, N., et al. (2022). Tissue-resident memory and circulating T cells are early responders to pre-surgical cancer immunotherapy. *Cell* 185 (16), 2918–2935.e29. doi:10.1016/j.cell.2022.06.018
- Ma, S., Zhang, B., LaFave, L. M., Earl, A. S., Chiang, S., Hu, Y., et al. (2020). Chromatin potential identified by shared single-cell profiling of RNA and chromatin. *Cell* 183 (4), 1103–1116.e20. doi:10.1016/j.cell.2020.09.056
- Macosko, E. Z., Basu, A., Satija, R., Nemes, J., Shekhar, K., Goldman, M., et al. (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* 161 (5), 1202–1214. doi:10.1016/j.cell.2015.05.002
- Maria, M., Pouyanfar, N., Ord, T., and Kaikkonen, M. U. (2022). The power of single-cell RNA sequencing in eQTL discovery. *Genes (Basel)* 13 (3), 502. doi:10.3390/genes13030502
- Martin, B. K., Qiu, C., Nichols, E., Phung, M., Green-Gladden, R., Srivatsan, S., et al. (2023). Optimized single-nucleus transcriptional profiling by combinatorial indexing. *Nat. Protoc.* 18 (1), 188–207. doi:10.1038/s41596-022-00752-0
- Melo Ferreira, R., Freije, B. J., and Eadon, M. T. (2021). Deconvolution tactics and normalization in renal spatial transcriptomics. *Front. Physiol.* 12, 812947. doi:10.3389/fphys.2021.812947
- Melo Ferreira, R., Sabo, A. R., Winfree, S., Collins, K. S., Janosevic, D., Gulbranson, C. J., et al. (2021). Integration of spatial and single-cell transcriptomics localizes epithelial cell-immune cross-talk in kidney injury. *JCI Insight* 6 (12), e147703. doi:10.1172/jci.insight.147703
- Menon, R., Otto, E. A., Hoover, P., Eddy, S., Mariani, L., Godfrey, B., et al. (2020). Single cell transcriptomics identifies focal segmental glomerulosclerosis remission endothelial biomarker. *JCI Insight* 5 (6), e133267. doi:10.1172/jci.insight.133267
- Miao, Z., Balzer, M. S., Ma, Z., Liu, H., Wu, J., Shrestha, R., et al. (2021a). Single cell regulatory landscape of the mouse kidney highlights cellular differentiation programs and disease targets. *Nat. Commun.* 12 (1), 2277. doi:10.1038/s41467-021-22266-1
- Miao, Z., Humphreys, B. D., McMahon, A. P., and Kim, J. (2021b). Multi-omics integration in the age of million single-cell data. *Nat. Rev. Nephrol.* 17 (11), 710–724. doi:10.1038/s41581-021-00463-x
- Mimitou, E. P., Lareau, C. A., Chen, K. Y., Zorzetto-Fernandes, A. L., Hao, Y., Takeshima, Y., et al. (2021). Scalable, multimodal profiling of chromatin accessibility, gene expression and protein levels in single cells. *Nat. Biotechnol.* 39 (10), 1246–1258. doi:10.1038/s41587-021-00927-2
- Nehar-Belaid, D., Hong, S., Marches, R., Chen, G., Bolisetty, M., Baisch, J., et al. (2020). Mapping systemic lupus erythematosus heterogeneity at the single-cell level. *Nat. Immunol.* 21 (9), 1094–1106. doi:10.1038/s41590-020-0743-0
- Nettersheim, F. S., Armstrong, S. S., Durant, C., Blanco-Dominguez, R., Roy, P., Orecchioni, M., et al. (2022). Titration of 124 antibodies using CITE-Seq on human PBMCs. *Sci. Rep.* 12 (1), 20817. doi:10.1038/s41598-022-24371-7
- Network, B. I. C. C. (2021). A multimodal cell census and atlas of the mammalian primary motor cortex. *Nature* 598 (7879), 86–102. doi:10.1038/s41586-021-03950-0
- O'Flanagan, F., Campbell, K. R., Zhang, A. W., Kabeer, F., Lim, J. L. P., Biele, J., et al. (2019). Dissociation of solid tumor tissues with cold active protease for single-cell RNA-



- seq minimizes conserved collagenase-associated stress responses. *Genome Biol.* 20 (1), 210. doi:10.1186/s13059-019-1830-0
- O'Huallachain, M., Bava, F. A., Shen, M., Dallett, C., Paladugu, S., Samusik, N., et al. (2020). Ultra-high throughput single-cell analysis of proteins and RNAs by split-pool synthesis. *Commun. Biol.* 3 (1), 213. doi:10.1038/s42003-020-0896-2
- Peterson, V. M., Zhang, K. X., Kumar, N., Wong, J., Li, L., Wilson, D. C., et al. (2017). Multiplexed quantification of proteins and transcripts in single cells. *Nat. Biotechnol.* 35 (10), 936–939. doi:10.1038/nbt.3973
- Psarras, A., Wittmann, M., and Vital, E. M. (2022). Emerging concepts of type I interferons in SLE pathogenesis and therapy. *Nat. Rev. Rheumatol.* 18 (10), 575–590. doi:10.1038/s41584-022-00826-z
- Rao, A., Barkley, D., Franca, G. S., and Yanai, I. (2021). Exploring tissue architecture using spatial transcriptomics. *Nature* 596 (7871), 211–220. doi:10.1038/s41586-021-03634-9
- Rautenstrauch, P., Vlot, A. H. C., Saran, S., and Ohler, U. (2022). Intricacies of single-cell multi-omics data integration. *Trends Genet.* 38 (2), 128–139. doi:10.1016/j.tig.2021.08.012
- Regev, A., Teichmann, S. A., Lander, E. S., Amit, I., Benoist, C., Birney, E., et al. (2017). The human cell atlas. *Elife* 6, e27041. doi:10.7554/eLife.27041
- Rice, G. I., Melki, I., Fremont, M. L., Briggs, T. A., Rodero, M. P., Kitabayashi, N., et al. (2017). Assessment of type I interferon signaling in pediatric inflammatory disease. *J. Clin. Immunol.* 37 (2), 123–132. doi:10.1007/s10875-016-0359-1
- Rousselle, T. V., McDaniels, J. M., Shetty, A. C., Bardhi, E., Maluf, D. G., and Mas, V. R. (2022). An optimized protocol for single nuclei isolation from clinical biopsies for RNA-seq. *Sci. Rep.* 12 (1), 9851. doi:10.1038/s41598-022-14099-9
- Salem, F., Perin, L., Sedrakyan, S., Angeletti, A., Ghiggi, G. M., Coccia, M. C., et al. (2022). The spatially resolved transcriptional profile of acute T cell-mediated rejection in a kidney allograft. *Kidney Int.* 101 (1), 131–136. doi:10.1016/j.kint.2021.09.004
- Salmen, F., De Jonghe, J., Kaminski, T. S., Alemany, A., Parada, G. E., Verity-Legg, J., et al. (2022). High-throughput total RNA sequencing in single cells using VASA-seq. *Nat. Biotechnol.* 40 (12), 1780–1793. doi:10.1038/s41587-022-01361-8
- Sandholm, N., Cole, J. B., Nair, V., Sheng, X., Liu, H., Ahlqvist, E., et al. (2022). Genome-wide meta-analysis and omics integration identifies novel genes associated with diabetic kidney disease. *Diabetologia* 65 (9), 1495–1509. doi:10.1007/s00125-022-05735-0
- Shahi, P., Kim, S. C., Haliburton, J. R., Gartner, Z. J., and Abate, A. R. (2017). Abseq: Ultrahigh-throughput single cell protein profiling with droplet microfluidic barcoding. *Sci. Rep.* 7, 44447. doi:10.1038/srep44447
- Shema, E., Bernstein, B. E., and Buenostro, J. D. (2019). Single-cell and single-molecule epigenomics to uncover genome regulation at unprecedented resolution. *Nat. Genet.* 51 (1), 19–25. doi:10.1038/s41588-018-0290-x
- Sheng, J., Hod, E. A., Vlad, G., and Chavez, A. (2022). Quantifying protein abundance on single cells using split-pool sequencing on DNA-barcoded antibodies for diagnostic applications. *Sci. Rep.* 12 (1), 884. doi:10.1038/s41598-022-04842-7
- Sheng, X., Guan, Y., Ma, Z., Wu, J., Liu, H., Qiu, C., et al. (2021). Mapping the genetic architecture of human traits to cell types in the kidney identifies mechanisms of disease and potential treatments. *Nat. Genet.* 53 (9), 1322–1333. doi:10.1038/s41588-021-00909-9
- Shi, T., Burg, A. R., Caldwell, J. T., Roskin, K., Castro-Rojas, C. M., Chukwuma, P. C., et al. (2023). Single cell transcriptomic analysis of renal allograft rejection reveals novel insights into intragraft TCR clonality. *bioRxiv* 2023, 524808. doi:10.1101/2023.02.08.524808
- Singh, S., Clemente, L. C., Parra, E. R., Tchakarov, A., Yang, C., Li, Y., et al. (2022). Urinary T cells are detected in patients with immune checkpoint inhibitor-associated immune nephritis that are clonotypically identical to kidney T cell infiltrates. *Oncimmunology* 11 (1), 2124678. doi:10.1080/2162402X.2022.2124678
- Stoeckius, M., Hafemeister, C., Stephenson, W., Houck-Loomis, B., Chattopadhyay, P. K., Swerdlow, H., et al. (2017). Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* 14 (9), 865–868. doi:10.1038/nmeth.4380
- Su, Y., Chen, D., Yuan, D., Lausted, C., Choi, J., Dai, C. L., et al. (2020). Multi-omics resolves a sharp disease-state shift between mild and moderate COVID-19. *Cell* 183 (6), 1479–1495.e20. doi:10.1016/j.cell.2020.10.037
- Susantitaphong, P., Cruz, D. N., Cerda, J., Abulfaraj, M., Alqahtani, F., Koulouridis, I., et al. (2013). World incidence of AKI: A meta-analysis. *Clin. J. Am. Soc. Nephrol.* 8 (9), 1482–1493. doi:10.2215/CJN.00710113
- Swanson, E., Lord, C., Reading, J., Heubeck, A. T., Genge, P. C., Thomson, Z., et al. (2021). Simultaneous trimodal single-cell measurement of transcripts, epitopes, and chromatin accessibility using TEA-seq. *Elife* 10, e63632. doi:10.7554/eLife.63632
- Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., et al. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods* 6 (5), 377–382. doi:10.1038/nmeth.1315
- Tuttle, K. R., Agarwal, R., Alpers, C. E., Bakris, G. L., Brosius, F. C., Kolkhof, P., et al. (2022). Molecular mechanisms and therapeutic targets for diabetic kidney disease. *Kidney Int.* 102 (2), 248–260. doi:10.1016/j.kint.2022.05.012
- van Buggenum, J. A. G., Gerlach, J. P., Tanis, S. E. J., Hogeweg, M., Jansen, P., Middelwijk, J., et al. (2018). Immuno-detection by sequencing enables large-scale high-dimensional phenotyping in cells. *Nat. Commun.* 9 (1), 2384. doi:10.1038/s41467-018-04761-0
- Welch, J. D., Kozareva, V., Ferreira, A., Vanderburg, C., Martin, C., and Macosko, E. Z. (2019). Single-cell multi-omic integration compares and contrasts features of brain cell identity. *Cell* 177 (7), 1873–1887. doi:10.1016/j.cell.2019.05.006
- Williams, A. M., Jensen, D. M., Pan, X., Liu, P., Liu, J., Huls, S., et al. (2022). Histologically resolved small RNA maps in primary focal segmental glomerulosclerosis indicate progressive changes within glomerular and tubulointerstitial regions. *Kidney Int.* 101 (4), 766–778. doi:10.1016/j.kint.2021.12.030
- Wilson, P. C., Muto, Y., Wu, H., Karihaloo, A., Waikar, S. S., and Humphreys, B. D. (2022). Multimodal single cell sequencing of human diabetic kidney disease implicates chromatin accessibility and genetic background in disease progression. *bioRxiv* 2022, 478204. doi:10.1038/s41467-022-32972-z
- Wilson, P. C., Wu, H., Kirita, Y., Uchimura, K., Ledru, N., Rennke, H. G., et al. (2019). The single-cell transcriptomic landscape of early human diabetic nephropathy. *Proc. Natl. Acad. Sci. U. S. A.* 116 (39), 19619–19625. doi:10.1073/pnas.1908706116
- Wu, H., Kirita, Y., Donnelly, E. L., and Humphreys, B. D. (2019). Advantages of single-nucleus over single-cell RNA sequencing of adult kidney: Rare cell types and novel cell states revealed in fibrosis. *J. Am. Soc. Nephrol.* 30 (1), 23–32. doi:10.1681/ASN.2018090912
- Wu, H., Malone, A. F., Donnelly, E. L., Kirita, Y., Uchimura, K., Ramakrishnan, S. M., et al. (2018). Single-cell transcriptomics of a human kidney allograft biopsy specimen defines a diverse inflammatory response. *J. Am. Soc. Nephrol.* 29 (8), 2069–2080. doi:10.1681/ASN.2018020125
- Wu, T. D., Madireddi, S., de Almeida, P. E., Banchereau, R., Chen, Y. J., Chitre, A. S., et al. (2020). Peripheral T cell expansion predicts tumour infiltration and clinical response. *Nature* 579 (7798), 274–278. doi:10.1038/s41586-020-2056-8
- Zhang, Y., Lin, X., Yao, Z., Sun, D., Lin, X., Wang, X., et al. (2023). Deconvolution algorithms for inference of the cell-type composition of the spatial transcriptome. *Comput. Struct. Biotechnol. J.* 21, 176–184. doi:10.1016/j.csbj.2022.12.001
- Zheng, G. X., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., et al. (2017). Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* 8, 14049. doi:10.1038/ncomms14049
- Zhu, C., Yu, M., Huang, H., Juric, I., Abnoui, A., Hu, R., et al. (2019). An ultra high-throughput method for single-cell joint analysis of open chromatin and transcriptome. *Nat. Struct. Mol. Biol.* 26 (11), 1063–1070. doi:10.1038/s41594-019-0323-x