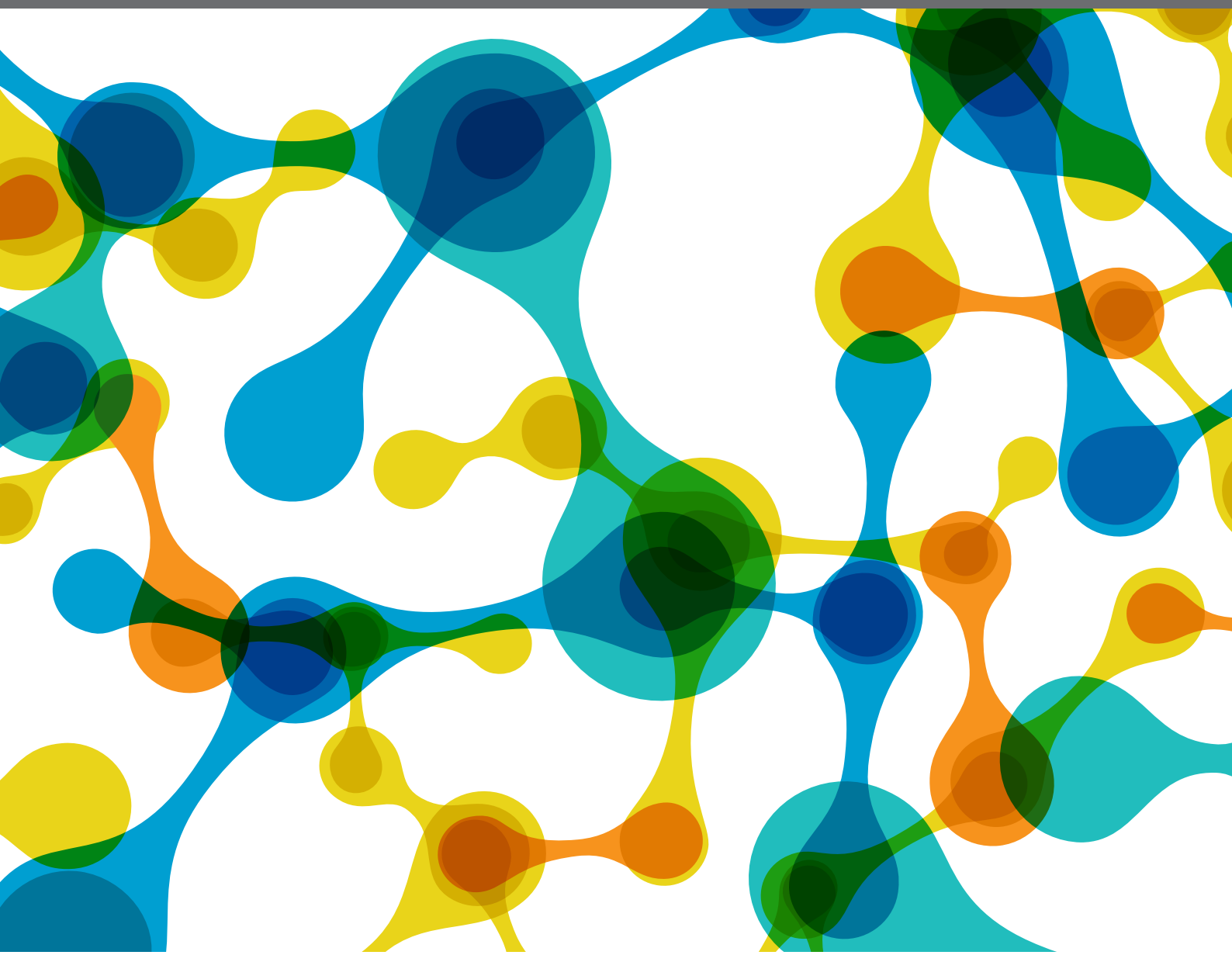
PROCEEDINGS OF THE 2021 INDIANA O'BRIEN CENTER MICROSCOPY WORKSHOP

EDITED BY: Bruce Molitoris, Andrew Hall and Ken William Dunn
PUBLISHED IN: Frontiers in Physiology



frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88976-230-9
DOI 10.3389/978-2-88976-230-9

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

PROCEEDINGS OF THE 2021 INDIANA O'BRIEN CENTER MICROSCOPY WORKSHOP

Topic Editors:
Bruce Molitoris, Indiana University, United States
Andrew Hall, University of Zurich, Switzerland
Ken William Dunn, Indiana University Bloomington, United States

Citation: Molitoris, B., Hall, A., Dunn, K. W., eds. (2022). Proceedings of the 2021 Indiana O'Brien Center Microscopy Workshop. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88976-230-9

Table of Contents

frontiers | Frontiers in Physiology

EDITORIAL
published: 02 May 2022
doi: 10.3389/fphys.2022.891526

Editorial: Proceedings of the 2021 Indiana O'Brien Center Microscopy Workshop

Kenneth W. Dunn[1]*, *Andrew M. Hall*[2] and *Bruce A. Molitoris*[1]

[1]Division of Nephrology, Indiana University School of Medicine, Indianapolis, IN, United States, [2]Institute of Anatomy, University of Zurich, Zurich, Switzerland

Keywords: Microscopy, image analysis, tissue cytometry, spatial transcriptomics, machine learning

Editorial on the Research Topic

Proceedings of the 2021 Indiana O'Brien Center Microscopy Workshop

In June of 2021, the NIH P-30 funded Indiana O'Brien Center for Advanced Renal Microscopic Analysis held its eighth advanced microscopy workshop. Presented every two years since the Center's inception in 2002, the goal of the workshop is to provide renal investigators with hands-on training and encourage the application of exciting new techniques of microscopy and image analysis to the study of kidney disease. Held in Indianapolis, Indiana in the spring the workshop is designed around a model of 4–5 days of morning lectures followed by afternoon hands-on laboratory exercises in confocal microscopy, multiphoton intravital microscopy and digital image analysis. However, due to concerns about the safety of participants during the Covid-19 pandemic, the workshop was held online in 2021. While the switch precluded the hands-on laboratory experiences that were an important part of previous workshops, the virtual format had a silver lining as it facilitated participation by a broad range of microscopy experts from around the world and allowed for more attendees. As a consequence, the workshop featured lectures and demonstrations from four members of the Indiana O'Brien Center and from a stellar field of 20 leaders in the field of microscopy and was viewed by hundreds of attendees from around the world.

The meeting proceedings are presented in this volume of Frontiers in Physiology. Each day had a unique theme underscoring new developments, and their applications, in areas critical for the advancement of microscopy. Specifically designed to catalyze interactions between the four speakers of each day, the meeting offered a unique view of the evolving fields and input into utilization, challenges and future advances expected. Each presenter was asked to contribute a mini-review to this issue and ten agreed.

The first day of the workshop consisted of an overview of the imaging research and development being conducted by the Indiana O'Brien Center. Bruce Molitoris described how intravital multiphoton microscopy has advanced our understanding of kidney physiology and pathophysiology. Tarek Ashkar presented new methods of 3D tissue cytometry that he has developed to elucidate the mechanisms of kidney disease from animal tissues and human biopsies. Michael Eadon introduced novel approaches to spatial transcriptomics, providing unique insights into localized transcriptional responses to kidney injury. Ken Dunn described methods of digital image analysis that the Indiana O'Brien Center has developed to support quantitative studies of living animals and fixed human and animal tissues. Finally, Seth Winfree presented a description and tutorial on the use of the Volumetric Tissue Exploration and Analysis

OPEN ACCESS

Edited by:
Carolyn Mary Ecelbarger, Georgetown University, United States

Reviewed by:
Dennis Brown, Massachusetts General Hospital and Harvard Medical School, United States
Hjalmar Brismar, Royal Institute of Technology, Sweden

***Correspondence:**
Kenneth W. Dunn
kwdunn@iu.edu

Specialty section:
This article was submitted to Renal and Epithelial Physiology, a section of the journal Frontiers in Physiology

Received: *07 March 2022*
Accepted: *28 March 2022*
Published: *02 May 2022*

Citation:
Dunn KW, Hall AM and Molitoris BA (2022) Editorial: Proceedings of the 2021 Indiana O'Brien Center Microscopy Workshop. Front. Physiol. 13:891526. doi: 10.3389/fphys.2022.891526

software, a unique software tool that he developed that provides a complete, interactive workflow solution to tissue cytometry. Perspectives on all five of these topics are included in this volume.

The second day of the workshop was devoted to the new field of large-scale tissue cytometry, an exciting new technique capable of providing a census of every cell and its physiological state in millimeter-scale tissue samples. Bernd Bodenmiller (Zurich, Switzerland) and Lloyd Cantley (Yale, United States) presented examples of how they have applied the technique of quantitative imaging mass cytometry to the study of cancer and kidney disease, respectively. Jeffrey Spraggins (Vanderbilt, United States) described how he has combined imaging mass spectrometry, multiplexed fluorescence microscopy and clinical microscopy to characterize the human kidney for the NIH Human Biomolecular Atlas Program. Michael Gerner (Washington, United States) presented a powerful new integrated approach that he has developed, combining methods of tissue clearing, multiplexed fluorescence microscopy and quantitative image analysis to characterize immune cell microenvironments throughout entire organs. Perspectives from Drs. Cantley and Spraggins are included in this volume.

The third day of the workshop was focused on advances in digital image analysis, a critical aspect of biological microscopy that has become increasingly challenging as the scale and complexity of biological microscopy has grown. Carolina Wahlby (Uppsala, Sweden) and Peter Horvath (Szeged, Hungary) presented an overview of the unique challenges of image analysis in biological microscopy and the solutions developed by their laboratories. David van Valen (Caltech, United States) and Pinaki Sarder (Buffalo, United States) described how they have developed and applied methods of machine learning to improve the quality and power of image analysis in research and pathology. Contributions from Drs. Wahlby and Sarder are included in this volume.

The last day of the workshop featured lectures on new approaches to high-content transcriptomic imaging and efficient large-scale imaging achieved by light-sheet microscopy. Joakim Lundeberg (SciLifeLab, Stockholm, Sweden), Jamie Marshall (Broad Inst., United States) and Yodai Takei (Caltech, United States) described how new techniques of spatial transcriptomics can be used to characterize gene transcription in the spatial context of tissues to better understand human disease. Jonathan Liu (Washington, United States) and Denise Marciano (Texas Southwestern, United States) demonstrated how new methods of light-sheet microscopy facilitate the efficient collection of enormous tissue volumes, extending the power of microscopy as a tool in research and pathology. An overview of Dr. Marshall's presentation is included in this volume.

The unique confluence of imaging experts assembled at the 2021 O'Brien Center Workshop provided participants with an outstanding overview of clever new approaches that have profoundly extended the scale and scope of biological microscopy. They also pointed the way to how we address the challenges to fully realize the potential of this rich data, including software designed for interactive exploration of massive, complex image volumes, methods of automated image analysis supporting reproducible quantitative analysis, and methods of data analysis designed to help researchers explore and detect latent patterns in highly-multiplexed spatial data. Over the course of the workshop, a consistent theme emerged–each of these solutions was developed through a collaboration between experts from disparate, complementary fields, including microscopy, molecular biology, chemistry, computer science, signal processing, machine learning and data analytics. In this sense, the meeting demonstrated the potential of inter-disciplinary science at its best. The editors hope readers find this volume stimulating and look forward to seeing where these collaborations lead in the future.

AUTHOR CONTRIBUTIONS

KD provided the first draft of the editorial. BM and AH reviewed, added additional information and finalized the article.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Physiology

MINI REVIEW
published: 16 December 2021
doi: 10.3389/fphys.2021.812170

Digital Image Analysis Tools Developed by the Indiana O'Brien Center

*Kenneth W. Dunn**

Division of Nephrology, Indiana University School of Medicine, Indianapolis, IN, United States

The scale and complexity of images collected in biological microscopy have grown enormously over the past 30 years. The development and commercialization of multiphoton microscopy has promoted a renaissance of intravital microscopy, providing a window into cell biology *in vivo*. New methods of optical sectioning and tissue clearing now enable biologists to characterize entire organs at subcellular resolution. New methods of multiplexed imaging support simultaneous localization of forty or more probes at a time. Exploiting these exciting new techniques has increasingly required biomedical researchers to master procedures of image analysis that were once the specialized province of imaging experts. A primary goal of the Indiana O'Brien Center has been to develop robust and accessible image analysis tools for biomedical researchers. Here we describe biomedical image analysis software developed by the Indiana O'Brien Center over the past 25 years.

Keywords: image analysis, volume rendering, segmentation, tissue cytometry, intravital microscopy, image registration

OPEN ACCESS

Edited by:
Alexander Staruschenko, University of South Florida, United States

Reviewed by:
Oleg Palygin, Medical University of South Carolina, United States
Georgina Gyarmati, University of Southern California, United States

***Correspondence:**
Kenneth W. Dunn
kwdunn@iu.edu

Specialty section:
This article was submitted to Renal and Epithelial Physiology, a section of the journal Frontiers in Physiology

Received: *09 November 2021*
Accepted: *29 November 2021*
Published: *16 December 2021*

Citation:
Dunn KW (2021) Digital Image Analysis Tools Developed by the Indiana O'Brien Center. Front. Physiol. 12:812170. doi: 10.3389/fphys.2021.812170

INTRODUCTION

Over the past 200 years, biological microscopy has evolved from a largely descriptive technique, documented with pictures and verbal descriptions, into a legitimately quantitative research approach. This evolution was fueled by the widespread deployment of digital detectors in the 1980s and digital computers in the 1990s. As biological microscopy became "digital," biologists increasingly found themselves having to train themselves in methods of digital image analysis in order to visualize and analyze their imaging studies. The past 20 years have witnessed an extraordinary explosion in the development of methods of biological microscopy, extending its scope, scale, complexity and resolution. Realizing the vast potential of these techniques has required that biomedical researchers master increasingly challenging methods of image and data analysis, methods that are generally well outside the realm of their training. Over the course of the Indiana O'Brien Center's existence [see review in Dunn et al. (2021)], we have encountered multiple cases where necessary software tools either do not exist or require an inordinately high level of expertise. A primary goal of the Indiana O'Brien Center has been to develop robust image analysis tools that are accessible to biomedical researchers lacking specialized image analysis experience. Examples of image analysis software developed by the Center are listed in **Table 1**, and described in detail below.

TABLE 1 | Software developed by the Indiana O'Brien Center.

Software	Application	References	Availability
Voxx	3D volume rendering for personal computers	Clendenon et al., 2002. *Am J Physiol Cell Physiol*. 282:C213-218	http://web.medicine.iupui.edu/ICBM/software
IMART	Motion correction for time-series and 3D intravital microscopy images	Dunn et al., 2014. *Intravital*. 3:e28210 Lorenz et al., 2012. *J Microsc*. 245:148-160	http://web.medicine.iupui.edu/ICBM/software
STAFF	Near-continuous measurement of microvascular velocity in 2D networks	Clendenon et al., 2019b. *Microvasc Res*. 123:7-13 Clendenon et al., 2019a. *J Vis Exp*	https://github.com/icbm-iupui/STAFF
VTEA	Interactive exploration of large-scale images and image data for quantitative tissue cytometry	Winfree et al., 2017b. *J Am Soc Nephrol*. 28:2108-2118 Winfree et al., 2017a.*Transl Res*. 189:1-12	https://github.com/icbm-iupui/volumetric-tissue-exploration-analysis
DeepSynth	Segmentation of nuclei in three-dimensional microscopy images	Dunn et al., 2019. *Sci Rep*. 9:18295 Ho et al., 2017. *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*:834-842 Fu et al., 2017. *IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*:704-708	ace@ecn.purdue.edu

INTERACTIVE VISUALIZATION OF THREE-DIMENSIONAL IMAGE VOLUMES – VOXX SOFTWARE

The optical sectioning provided by confocal, and later multiphoton and lightsheet microscopy opened the door to three-dimensional (3D) microscopy. However, when the Indiana O'Brien Center collected its first multiphoton excitation fluorescence image volumes in 2001 visualizing these volumes was challenging. Commonly available software provided static anaglyphs, or sequences of projections, but interactive visualization was limited to scrolling through sequential planes. "Real-time volume rendering" was then an expensive option, requiring costly workstations and surprisingly costly software. However, the rapid growth of video gaming profoundly changed the landscape of computer technology development, moving volume-rendering from a niche scientific market to an enormous consumer market. Jeff Clendenon, a computer engineer in the Indiana O'Brien Center recognized that the graphics capabilities that were once found only on expensive workstations had been reproduced in affordable graphics processors found in personal computers. He proceeded to develop the ground-breaking Voxx scientific volume rendering software, which put real-time volume rendering into the hands of nearly anyone with a personal computer (Clendenon et al., 2002). Voxx (**Figure 1A**) provides 3D renderings of an image volume (maximum projection or alpha-blending), that update in real-time as the user moves the volume around using a mouse, essentially reproducing the experience of rotating an actual 3D object in space. The ability to interactively manipulate the volume is critical to fully exploring a complex image volume. Voxx also supports the ability to export individual images, or to save a volume rendering sequence as a video for presentations. Over the years since Voxx was released, a variety of volume visualization tools have been developed, both free (e.g., ImageJ) and commercial (e.g., Imaris and Amira). However, because of its unique flexibility and capability, Voxx remains a compelling choice, particularly among free software solutions. Voxx continues to be a mainstay tool of the Indiana O'Brien Center and, as of the time of writing, has been cited in over 100 papers and downloaded more than 6000 times[1].

CORRECTING MOTION ARTIFACTS IN INTRAVITAL MICROSCOPY – IMAGE MOTION ARTIFACT REDUCTION TOOL SOFTWARE

Intravital microscopy has been a core technology of the Indiana O'Brien Center since its inception, and a long-standing goal of the Center has been to promote and facilitate intravital microscopy as a powerful tool for understanding the function of the kidney in health and disease. In our first forays into intravital microscopy we immediately discovered that tissue motion, derived primarily from respiration, represented a significant challenge to high resolution *in vivo* imaging. Subsequent studies of liver, pancreas, lymph nodes, and lung demonstrated that tissue motion was a general problem for intravital microscopy of visceral organs. In contrast to the brain, which can be effectively immobilized using stereotaxic devices attached to the skull, visceral organs move relatively freely in the living animal so that imaging at sub-cellular resolution depends upon methods immobilizing tissue to micron precision. We have since developed robust and reproducible methods for mounting the kidney and other internal organs of rat and mice on the stage of an inverted microscope stage in a way that immobilizes the organ without compromising function (Dunn et al., 2018). Even so, there are occasions when tissue motion cannot be controlled, resulting in studies that cannot be quantified or occasionally, even interpreted.

The problem of tissue motion can be addressed at capture, by gating image collection to avoid respiratory motion [see review in Soulet et al. (2020)], an approach that can be augmented for three-dimensional images, by digital reconstruction (Vladymyrov et al., 2020). For time series studies corrupted by relatively few distorted images, the problem of motion artifacts can be addressed by simply discarding distorted images, or portions of images (Soulet et al., 2013). To address the problem

[1]http://web.medicine.iupui.edu/ICBM/software

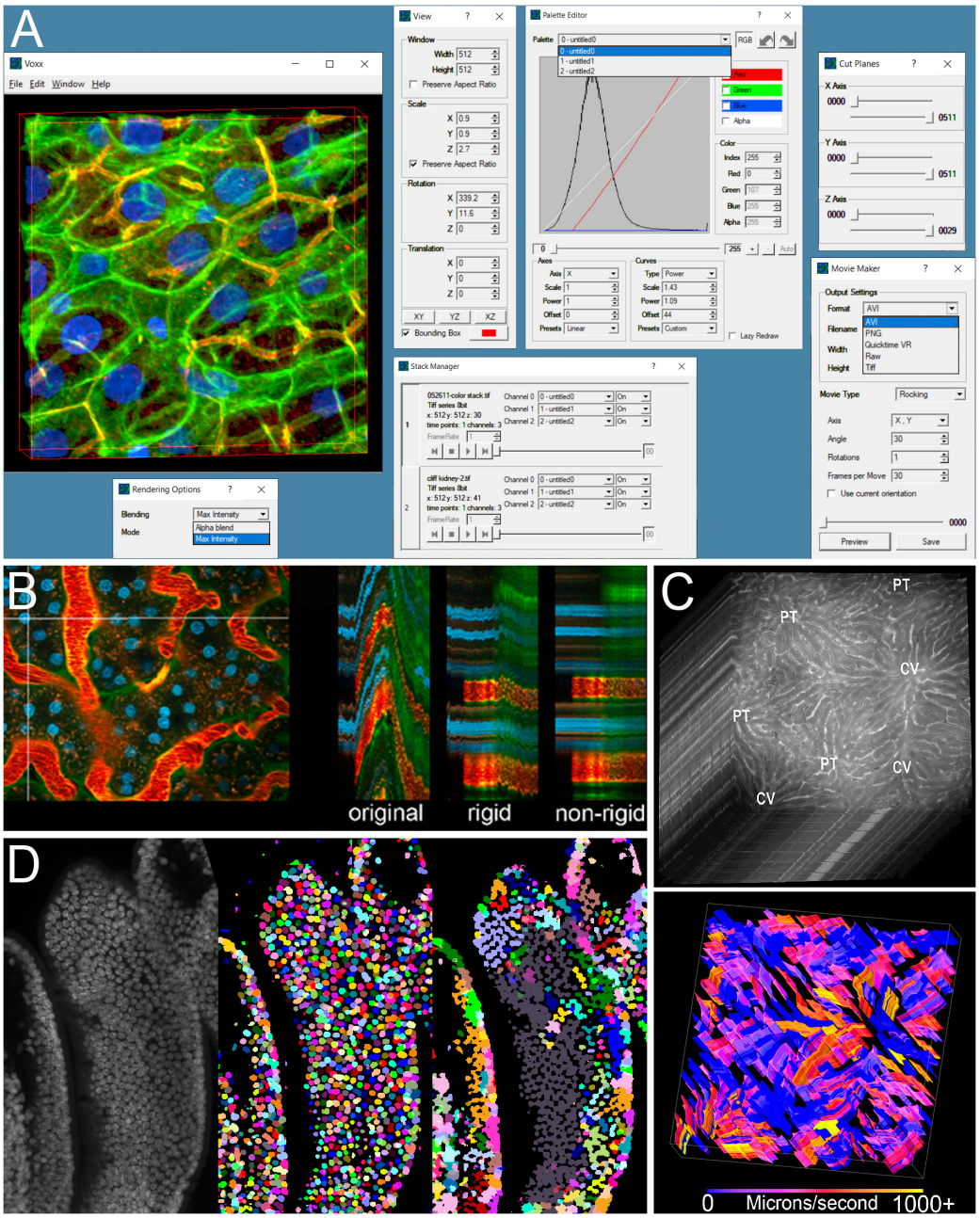


FIGURE 1 | Examples of image processing software developed by the Indiana O'Brien Center. **(A)** Screenshot of Voxx volume visualization software showing the rendered volume, and interactive windows for selecting rendering method, adjusting view and channel palette settings, selecting between multiple volumes, limiting the volume to be displayed, and setting parameters for video outputs. **(B)** Example of IMART image registration. Left – first of a series of images collected over time from the kidney of a living rat. Vertical line indicates region used to generate YT images (two-dimensional images that show the image of a single line, oriented vertically over time, and oriented horizontally). Right – YT images from the original time series, after rigid registration and after rigid and non-rigid registration. **(C)** Example of STAFF microvascular velocity measurements. Top – Series of images collected at the rate of 97.5 frames per second from the liver of a living rat following injection of a fluorescent dextran. Bottom – map of velocities measured over time in which time is presented as a third dimension. **(D)** Comparison of nuclear segmentation results obtained from a 3D volume of mouse intestine (left), using DeepSynth (middle) or CellProfiler (right). Images shown in panels **(B–D)** are modified from previous publications (Dunn et al., 2014, 2019; Clendenon et al., 2019b) and used with permission.

of pervasive image distortion in time series and 3D intravital microscopy, the students from the laboratory of Edward Delp, a Purdue University investigator of the Digital Image Analysis Core of the Indiana O'Brien Center developed novel software to retrospectively correct intravital microcopy images that were compromised by motion artifacts. Based upon an algorithm that

seeks to minimize the differences between images, the Image Motion Artifact Reduction Tool (IMART) software can be used to correct motion artifacts in sequences of images collected over time or in three dimensions (Lorenz et al., 2012; Dunn et al., 2014). Unlike other image registration solutions, IMART can be used to correct for both linear (rigid) translations occurring between sequential frames as well as non-linear (warping) distortions occurring within each frame, distortions that are unique to intravital microscopy (**Figure 1B**). IMART software, which has been downloaded by more than 100 laboratories, has been used to remove motion artifacts from intravital microscopy images collected from the rodent lung and kidney (Presson et al., 2011; Hato et al., 2017), enabling quantitative analysis that would have otherwise been impossible.

CONTINUOUS MEASUREMENT OF MICROVASCULAR FLOW ACROSS ENTIRE OPTICAL SECTIONS – STAFF SOFTWARE

Historically, one of the most common applications of intravital microscopy has been the measurement of microvascular flow. The procedure typically involves measuring the displacement of cells or particles in a series of images collected over time from a capillary segment. Cells can be identified either by fluorescent labeling or as shadows in the lumen of the capillary labeled with a fluorescent fluid probe. The velocity of the cells or particles can then be measured either by manually tracking individual cells or by measuring angles in time-distance kymographs. In either case, the process is laborious enough that velocities are typically measured for only a few vascular segments and only for a very brief interval of time. While accurate measurements can be generated in this way, they are susceptible to the spatial and temporal variability that is characteristic of microvascular flow.

To address this problem, the Indiana O'Brien Center worked Sherry Clendenon of the Indiana Biocomplexity Institute to develop an approach for continuous measurement of microvascular flow across entire microscope fields. Using time series images collected by high-speed intravital microscopy, STAFF (Spatial Temporal Analysis of Fieldwise Flow) automatically generates kymographs for each vascular segment in the field, which are then used to generate a complete map of microvascular velocity in each segment across the entire field (Clendenon et al., 2019b; **Figure 1C**). This approach gives STAFF the unique ability to measure microvascular velocities across entire fields at a temporal resolution on the scale of seconds. Analyses of images collected from the livers of mice demonstrated surprising variability in microvascular flow, with striking differences in flow rates between adjacent sinusoids, and numerous occasions when flow would suddenly stop and later restart. To encourage widespread use, STAFF was developed as a freely available plugin to ImageJ and its use is thoroughly described in a JOVE video (Clendenon et al., 2019a).

IMAGE AND DATA EXPLORATION FOR LARGE SCALE TISSUE CYTOMETRY – VOLUMETRIC TOOL FOR EXPLORATION AND ANALYSIS SOFTWARE

The development of automated microscope systems has enabled researchers to image the distribution of multiple probes at subcellular resolution in centimeter-scale tissue samples. These large and complex image volumes have spurred the development of "tissue cytometry," an image analysis technique capable of providing complete characterizations of the distribution, interactions and physiology of every cell in an organ (Gerner et al., 2012; Coutu et al., 2017; Halse et al., 2018). However, tissue cytometry represents a relatively new domain of image analysis so that quantitative analysis has largely been accomplished using a combination of custom and/or expensive image analysis software.

To address the need for an accessible solution to tissue cytometry, Seth Winfree, a member of the Indiana O'Brien Center developed VTEA (Volumetric Tool for Exploration and Analysis) (Winfree et al., 2017a,b), a unique software tool that provides a complete integrated workflow supporting every step in tissue cytometry, from segmentation, through classification and quantitation to data analysis via a simple, interactive user interface. A fundamental strength of VTEA is that, by integrating image and data analysis into a single software platform, VTEA expedites and encourages the process of discovery, an exciting aspect of large-scale tissue cytometry. Whereas most imaging studies are predicated on tests of hypotheses, tissue cytometry is typically conducted on images whose size and complexity is such that they contain enormous amounts of additional, latent information that may be apparent only upon exploration. VTEA provides a seamless pipeline between image and data analysis so that the user can, for example, quickly identify specific cell populations in the data space, using either supervised or unsupervised strategies, and visualize their distributions and relations to other cells in the image space. Conversely, the user can also identify interesting regions or cell populations in an image and explore the nature of the cells in these regions in the data space, using either scatterplots or tSNE plots.

VTEA has become a critical tool in the quantitative analysis of tissues by members of the Indiana O'Brien Center and beyond, unlocking the promise of tissue cytometry as tool for biomedical research and discovery. There are several powerful software tools currently available to support tissue cytometry, for example, the Cytomapper and Histocat software developed by the Bodenmiller laboratory (Schapiro et al., 2017; Eling et al., 2020), the Xit software developed by the Schroeder lab (Coutu et al., 2018) and the CytoMAP software developed by the Gerner lab (Stoltzfus et al., 2020). However, none incorporate the entire workflow of image processing, quantitation, visualization, and data analysis into a single continuous bidirectional platform that so effectively encourages exploration and analysis refinement. VTEA-based tissue cytometry has made critical contributions to studies of the processes underlying kidney stone formation conducted as part of a NIH-funded program project (Makki et al., 2020;

Winfree et al., 2020) and represents a cornerstone technology of the Indiana University contributions to the Kidney Precision Medicine Project (Winfree et al., 2017a, 2018; El-Achkar et al., 2021; Ferkowicz et al., 2021). Developed as a plug-in to ImageJ, VTEA is freely available online.

ONLINE IMAGE VISUALIZATION AND ANALYSIS – DISTRIBUTED AND NETWORKED ANALYSIS OF VOLUMETRIC IMAGE DATA HIGH PERFORMANCE IMAGE ANALYSIS SYSTEM

Capable of defining the distribution of multiple molecular species at subcellular resolution over regions spanning the full extent of the cortex and medulla, large-scale microscopy image volumes are enormously rich in potential information. However, extracting this information is challenging, not only because of the unique challenges of 3D image analysis, but also because the size and complexity of these image volumes are incompatible with resources available to most researchers. Large-scale image data places an enormous burden on computer and network infrastructure. A four- channel image volume, collected at subcellular resolution to a depth of 100 microns from a 5×6 mm region requires nearly 200 gigabytes of digital storage space. A complete study, which might involve comparison of multiple conditions, each with a reasonable number of replicates, could thus easily involve tens of terabytes of data. Managing data of this magnitude requires extensive and sophisticated computer hardware and network infrastructure beyond that available at most institutions. And the challenges of visualizing and quantifying 3D images, discussed previously, become much larger in image volumes of this scale.

To encourage the application of large-scale tissue cytometry by a broader range of investigators, the Indiana O'Brien Center has a developed an approach to large-scale microscopy that both removes most of these challenges. The O'Brien Center 3D Tissue Imaging Core provides a service whereby samples sent to the Center are imaged using one of the confocal or multiphoton microscopes of the Indiana Center for Biological Microscopy, and the resulting data archived at Indiana University, thus eliminating an investigator's need for extensive storage and network capabilities. The resulting images are also uploaded to a powerful online server system, the DINAVID (Distributed and Networked Analysis of Volumetric Image Data) high performance image analysis system. Hosted by Indiana University and developed by the laboratories of Edward Delp at Purdue and Paul Salama at IUPUI, the DINAVID system is designed to provide remote users throughout the world with an intuitive interface to their image data, supporting interactive visualization, quantitative analysis, and exploration. The DINAVID system is continuously updated with new tools as they are developed by the O'Brien Center Digital Image Analysis Core, including novel methods of 3D segmentation, as described below.

NUCLEAR SEGMENTATION USING A CONVOLUTIONAL NEURAL NETWORK TRAINED IN SYNTHETIC DATA – DEEPSYNTH SEGMENTATION SOFTWARE

Tissue cytometry is formally similar to flow cytometry, except that whereas in flow cytometry the sample is passed through a detector, in tissue cytometry, the detector is passed over the sample. However, unlike flow cytometry, where individual cells are physically separated from one another for quantification, tissue cytometry is complicated by the need for image analysis techniques to distinguish individual cells that are packed into a tissue. The process of distinguishing individual cells, cell "segmentation" is the critical first step in tissue cytometry. In the absence of membrane markers to delineate cell boundaries, individual cells are typically distinguished by their nuclei. Cells are then classified into specific cell types based upon the presence of specific markers in the regions surrounding each nucleus.

Numerous approaches have been developed to segment nuclei in two-dimensional images, supporting automated analysis of thin tissue sections and cells grown in culture. Historically, these approaches have been based upon traditional, morphological image processing operations but increasingly, investigators are demonstrating that deep-learning techniques frequently provide results that are more accurate (Caicedo et al., 2019a,b). Moreover, unlike morphological techniques that typically need to be tuned to the specific characteristics of each image, deep-learning techniques generally provide results that are robust across different images. However, that robustness is typically obtained only when the network is provided with a large amount high-quality "training" data – a library of images that have been manually annotated that are used by the network to "learn" the qualities of nuclei. As manual annotation is a laborious process, training is typically the rate-limiting step in the application of deep-learning techniques to segmentation. The barrier of manual annotation has been addressed in various ways, including side-stepping the annotation process and training networks using publicly available annotated datasets (Caicedo et al., 2019a; Stringer et al., 2021), using crowd-sourcing to annotate images (Moen et al., 2019), or using transfer learning, a process in which a network trained on a large amount of data is refined using a much smaller dataset (Zaki et al., 2020). Interested readers are directed to a recently published review of open-source, deep-learning software for segmentation of biological images (Lucas et al., 2021).

Segmentation of nuclei in three-dimensional tissues is significantly more challenging, in part because of the relatively poor axial resolution of optical microscopy. Accordingly, techniques for segmentation of nuclei in three-dimensional tissues are much less developed, seriously limiting biologists' ability to quantitatively analyze three-dimensional image volumes. As with two-dimensional segmentation, deep-learning represents an exciting approach to segmentation of three-dimensional images. However, the already tedious

task of generating training data is even more onerous in three dimensions. As with two-dimensional data, network performance depends upon annotation of hundreds, if not thousands of nuclei. Extending segmentation to three dimensions means that each nucleus must be manually annotated in multiple focal planes, including those collected from the top and bottom boundaries that are especially difficult to reproducibly delineate.

The Digital Image Analysis Core of the Indiana O'Brien Center addressed this problem by developing DeepSynth, a convolutional neural network trained on synthetic images, essentially eliminating the tedious task of manual annotation (Fu et al., 2017; Ho et al., 2017, 2018; Dunn et al., 2019). As compared with 3D segmentation software based upon traditional morphological segmentation techniques, DeepSynth provides segmentations that are more accurate, particularly for challenging image volumes (**Figure 1D**). A second benefit of the DeepSynth approach is that the quality of segmentations are more consistent throughout large image volumes. At the time of writing, the DeepSynth software, which is freely available from the O'Brien website, has been downloaded by 18 laboratories.

FUTURE DIRECTIONS OF THE DIGITAL IMAGE ANALYSIS CORE OF THE INDIANA O'BRIEN CENTER

The Digital Image Analysis Core is dedicated to the development of image analysis software to further the research of renal investigators. An overarching theme of the core is that the images of biological microscopy are rich in information, and extracting that information depends upon hands-on exploration and analysis by biologists who are not necessarily experts in digital image analysis. Thus image analysis software should be accessible, interactive, and user-friendly. Going forward, the core will continue to improve and refine deep learning-based methods of image segmentation by implementing more accurate models of synthetic data, e.g., by incorporating models of objective lens point-spread functions. The core is also working on methods to address the spatial variability in segmentation quality that we observe in large, three-dimensional image volumes. Insofar as segmentation is fundamental to image quantification, spatial variability in segmentation quantity directly impacts the reliability of tissue cytometry. Since it is impractical to measure segmentation quality at every point in a large image volume, methods are needed to estimate segmentation quality and to convert these estimates into confidence maps that can be used to inform interpretation of cytometry measurements. Finally, the core is also exploring how deep learning can be expanded into additional aspects of image analysis in tissue cytometry, including noise reduction, spectral deconvolution and cell classification.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this review and has approved it for publication.

FUNDING

This work was funded by the Indiana O'Brien Center for Advanced Renal Microscopy (NIH-NIDDK P30DK079312) and by NIH U01 GM111243.

ACKNOWLEDGMENTS

The image processing tools described here reflect the contributions of a large number of talented researchers. In addition to Edward Delp, Paul Salama, Jeff Clendenon, Sherry Clendenon, and Seth Winfree, all of whom were specifically mentioned in the article. I would like to acknowledge crucial contributions from Jason Byars and Carrie Phillips (Voxx), Kevin Lorenz (IMART), Xiao Fu and James Glazier (STAFF), Tarek Ashkar (VTEA), Shuo Han, Alain Chen, Liming Wu (DINAVID), Chichen Fu, David Ho, Soonam Lee, and Shuo Han (DeepSynth).

REFERENCES

Caicedo, J. C., Goodman, A., Karhohs, K. W., Cimini, B. A., Ackerman, J., Haghighi, M., et al. (2019a). Nucleus segmentation across imaging experiments: the 2018 Data Science Bowl. *Nat. Methods* 16, 1247–1253.

Caicedo, J. C., Roth, J., Goodman, A., Becker, T., Karhohs, K. W., Broisin, M., et al. (2019b). Evaluation of deep learning strategies for nucleus segmentation in fluorescence images. *Cytometry A* 95, 952–965. doi: 10.1002/cyto.a.23863

Clendenon, J. L., Phillips, C. L., Sandoval, R. M., Fang, S., and Dunn, K. W. (2002). Voxx: a PC-based, near real-time volume rendering system for biological microscopy. *Am. J. Physiol. Cell Physiol.* 282, C213–C218. doi: 10.1152/ajpcell.2002.282.1.C213

Clendenon, S. G., Fu, X., Von Hoene, R. A., Clendenon, J. L., Sluka, J. P., Winfree, S., et al. (2019a). Spatial temporal analysis of fieldwise flow in microvasculature. *J. Vis. Exp.* 153, 11–18. doi: 10.3791/60493

Clendenon, S. G., Fu, X., Von Hoene, R. A., Clendenon, J. L., Sluka, J. P., Winfree, S., et al. (2019b). A simple automated method for continuous fieldwise measurement of microvascular hemodynamics. *Microvasc. Res.* 123, 7–13. doi: 10.1016/j.mvr.2018.11.010

Coutu, D. L., Kokkaliaris, K. D., Kunz, L., and Schroeder, T. (2017). Three-dimensional map of nonhematopoietic bone and bone-marrow cells and molecules. *Nat. Biotechnol.* 35, 1202–1210. doi: 10.1038/nbt.4006

Coutu, D. L., Kokkaliaris, K. D., Kunz, L., and Schroeder, T. (2018). Multicolor quantitative confocal imaging cytometry. *Nat. Methods* 15, 39–46. doi: 10.1038/nmeth.4503

Dunn, K. W., Fu, C., Ho, D. J., Lee, S., Han, S., Salama, P., et al. (2019). DeepSynth: three-dimensional nuclear segmentation of biological images using neural networks trained with synthetic data. *Sci. Rep.* 9:18295. doi: 10.1038/s41598-019-54244-5

Dunn, K. W., Lorenz, K. S., Salama, P., and Delp, E. J. (2014). IMART software for correction of motion artifacts in images collected in intravital microscopy. *Intravital* 3:e28210. doi: 10.4161/intv.28210

Dunn, K. W., Molitoris, B. A., and Dagher, P. C. (2021). The indiana o'brien center for advanced renal microscopic analysis. *Am. J. Physiol. Renal Physiol.* 320, F671–F682. doi: 10.1152/ajprenal.00007.2021

Dunn, K. W., Sutton, T. A., and Sandoval, R. M. (2018). Live-animal imaging of renal function by multiphoton microscopy. *Curr. Protoc. Cytom.* 83, 12.9.1–12.9.25.

El-Achkar, T. M., Eadon, M. T., Menon, R., Lake, B. B., Sigdel, T. K., Alexandrov, T., et al. (2021). A multimodal and integrated approach to interrogate human kidney biopsies with rigor and reproducibility: guidelines from the kidney precision medicine project. *Physiol. Genomics* 53, 1–11. doi: 10.1152/physiolgenomics.00104.2020

Eling, N., Damond, N., Hoch, T., and Bodenmiller, B. (2020). Cytomapper: an R/bioconductor package for visualisation of highly multiplexed imaging data. *Bioinformatics* 36, 5706–5708. doi: 10.1093/bioinformatics/btaa1061

Ferkowicz, M. J., Winfree, S., Sabo, A. R., Kamocka, M. M., Khochare, S., Barwinska, D., et al. (2021). Large-scale, three-dimensional tissue cytometry of the human kidney: a complete and accessible pipeline. *Lab. Invest.* 101, 661–676. doi: 10.1038/s41374-020-00518-w

Fu, C. C., Ho, D. J., Han, S., Salama, P., Dunn, K. W., and Delp, E. J. (2017). "Nuclei segmentation of fluorescence microscopy images using convolutional neural networks," in *Proceedings of the IEEE 14th International Symposium on Biomedical Imaging*, (Melbourne, VIC: Institute of Electrical and Electronics Engineers), 704–708. doi: 10.1186/s12859-021-04245-x

Gerner, M. Y., Kastenmuller, W., Ifrim, I., Kabat, J., and Germain, R. N. (2012). Histo-cytometry: a method for highly multiplex quantitative tissue imaging analysis applied to dendritic cell subset microanatomy in lymph nodes. *Immunity* 37, 364–376. doi: 10.1016/j.immuni.2012.07.011

Halse, H., Colebatch, A. J., Petrone, P., Henderson, M. A., Mills, J. K., Snow, H., et al. (2018). Multiplex immunohistochemistry accurately defines the immune context of metastatic melanoma. *Sci. Rep.* 8:11158. doi: 10.1038/s41598-018-28944-3

Hato, T., Winfree, S., and Dagher, P. C. (2017). Intravital imaging of the kidney. *Methods* 128, 33–39. doi: 10.1016/j.ymeth.2017.03.024

Ho, D. J., Fu, C. C., Salama, P., Dunn, K. W., and Delp, E. J. (2017). "Nuclei segmentation of fluorescence microscopy images using three dimensional convolutional neural networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, (Honolulu, HI: Institute of Electrical and Electronics Engineers), 834–842.

Ho, D. J., Fu, C. C., Salama, P., Dunn, K. W., and Delp, E. J. (2018). "Nuclei detection and segmentation of fluorescence microscopy images using three dimensional convolutional neural networks," in *Proceedings of the 15th International Symposium on Biomedical Imaging*, (Washington, DC: Institute of Electrical and Electronics Engineers), 418–422.

Lorenz, K. S., Salama, P., Dunn, K. W., and Delp, E. J. (2012). Digital correction of motion artefacts in microscopy image sequences collected from living animals using rigid and nonrigid registration. *J. Microsc.* 245, 148–160. doi: 10.1111/j.1365-2818.2011.03557.x

Lucas, A., Ryder, P., Bin, L., Cimini, B., Eliceiri, K., and Carpenter, A. (2021). Open-source deep-learning software for bioimage segmentation. *Mol. Biol. Cell* 32, 823–829. doi: 10.1091/mbc.E20-10-0660

Makki, M. S., Winfree, S., Lingeman, J. E., Witzmann, F. A., Worcester, E. M., Krambeck, A. E., et al. (2020). A precision medicine approach uncovers a unique signature of neutrophils in patients with brushite kidney stones. *Kidney Int. Rep.* 5, 663–677. doi: 10.1016/j.ekir.2020.02.1025

Moen, E., Bannon, D., Kudo, T., Graf, W., Covert, M., and Van Valen, D. (2019). Deep learning for cellular image analysis. *Nat. Methods* 16, 1233–1246.

Presson, R. G. Jr., Brown, M. B., Fisher, A. J., Sandoval, R. M., Dunn, K. W., Lorenz, K. S., et al. (2011). Two-photon imaging within the murine thorax without respiratory and cardiac motion artifact. *Am. J. Pathol.* 179, 75–82. doi: 10.1016/j.ajpath.2011.03.048

Schapiro, D., Jackson, H., Raghuraman, S., Fischer, J., Zanotelli, V., Schulz, D., et al. (2017). histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data. *Nat. Methods* 14, 873–876. doi: 10.1038/nmeth.4391

Soulet, D., Lamontagne-Proulx, J., Aubé, B., and Davalos, D. (2020). Multiphoton intravital microscopy in small animals: motion artefact challenges and technical solutions. *J. Microsc.* 278, 3–17. doi: 10.1111/jmi.12880

Soulet, D., Paré, A., Coste, J., and Lacroix, S. (2013). Automated filtering of intrinsic movement artifacts during two-photon intravital microscopy. *PLoS One* 8:e53942. doi: 10.1371/journal.pone.0053942

Stoltzfus, C. R., Filipek, F., Gern, B., Olin, B., Leal, J., Wu, Y., et al. (2020). CytoMAP: a spatial analysis toolbox reveals features of myeloid cell organization in lymphoid tissues. *Cell Rep.* 31:107523. doi: 10.1016/j.celrep.2020.107523

Stringer, C., Wang, T., Michaelos, M., and Pachitariu, M. (2021). Cellpose: a generalist algorithm for cellular segmentation. *Nat. Methods* 18, 100–106. doi: 10.1038/s41592-020-01018-x

Vladymyrov, M., Jahromi, N. H., Kaba, E., Engelhardt, B., and Ariga, A. (2020). VivoFollow 2: distortion-free multiphoton intravital imaging. *Front. Phys.* 7:222. doi: 10.3389/fphy.2019.00222

Winfree, S., Dagher, P. C., Dunn, K. W., Eadon, M. T., Ferkowicz, M., Barwinska, D., et al. (2018). Quantitative large-scale three-dimensional imaging of human kidney biopsies: a bridge to precision medicine in kidney disease. *Nephron* 140, 134–139. doi: 10.1159/000490006

Winfree, S., Ferkowicz, M. J., Dagher, P. C., Kelly, K. J., Eadon, M. T., Sutton, T. A., et al. (2017a). Large-scale 3-dimensional quantitative imaging of tissues: state-of-the-art and translational implications. *Transl. Res.* 189, 1–12. doi: 10.1016/j.trsl.2017.07.006

Winfree, S., Khan, S., Micanovic, R., Eadon, M. T., Kelly, K. J., Sutton, T. A., et al. (2017b). Quantitative three-dimensional tissue cytometry to study kidney tissue and resident immune cells. *J. Am. Soc. Nephrol.* 28, 2108–2118. doi: 10.1681/ASN.2016091027

Winfree, S., Weiler, C., Bledsoe, S. B., Gardner, T., Sommer, A. J., Evan, A. P., et al. (2020). Multimodal imaging reveals a unique autofluorescence signature of Randall's plaque. *Urolithiasis* 49, 123–135. doi: 10.1007/s00240-020-01216-4

Zaki, G., Gudla, P. R., Lee, K., Kim, J., Ozbun, L., Shachar, S., et al. (2020). A deep learning pipeline for nucleus segmentation. *Cytometry A* 97, 1248–1264. doi: 10.1002/cyto.a.24257

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

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.

frontiers in Physiology

MINI REVIEW
published: 22 December 2021
doi: 10.3389/fphys.2021.817181

Using Imaging Mass Cytometry to Define Cell Identities and Interactions in Human Tissues

Vijayakumar R. Kakade†, Marlene Weiss† and Lloyd G. Cantley**

Section of Nephrology, Department of Internal Medicine, Yale University School of Medicine, New Haven, CT, United States

In the evolving landscape of highly multiplexed imaging techniques that can be applied to study complex cellular microenvironments, this review characterizes the use of imaging mass cytometry (IMC) to study the human kidney. We provide technical details for antibody validation, cell segmentation, and data analysis specifically tailored to human kidney samples, and elaborate on phenotyping of kidney cell types and novel insights that IMC can provide regarding pathophysiological processes in the injured or diseased kidney. This review will provide the reader with the necessary background to understand both the power and the limitations of IMC and thus support better perception of how IMC analysis can improve our understanding of human disease pathogenesis and can be integrated with other technologies such as single cell sequencing and proteomics to provide spatial context to cellular data.

Keywords: imaging mass cytometry, kidney, highly multiplexed imaging, kidney-MAPPS, cell phenotype, *in situ* imaging

OPEN ACCESS

Edited by:
Bruce Molitoris, Indiana University, United States

Reviewed by:
Christoph Kuppe, RWTH Aachen University, Germany

****Correspondence:***
Vijayakumar R. Kakade
vijayakumar.kakade@yale.edu
Lloyd G. Cantley
lloyd.cantley@yale.edu

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Renal and Epithelial Physiology, a section of the journal Frontiers in Physiology

Received: *17 November 2021*
Accepted: *25 November 2021*
Published: *22 December 2021*

Citation:
Kakade VR, Weiss M and Cantley LG (2021) Using Imaging Mass Cytometry to Define Cell Identities and Interactions in Human Tissues. Front. Physiol. 12:817181. doi: 10.3389/fphys.2021.817181

INTRODUCTION

The predominant methods for cell identification during the pathological analysis of formalin-fixed paraffin-embedded (FFPE) samples from human renal core biopsy tissues include cell morphology and immunohistochemistry or immunofluorescence (IF). The limited amount of tissue obtained from renal biopsy confines the type of analysis performed, preventing extensive analysis at the molecular and cellular level. Thus, most analyses are descriptive, with few efforts to provide quantitative information about the tubular, stromal, and nonresident cell populations in disease states (Kretzler et al., 2002; Zhang and Parikh, 2019). Single-cell RNA sequencing and single nucleus RNA sequencing have markedly increased the depth of information gained from a single biopsy, but lack the spatial information needed to determine cell proximity and cell–cell interactions (Rost et al., 2017; Cippà et al., 2018; Wu et al., 2018, 2019; Lake et al., 2019; Deleersnijder et al., 2021). The large numbers of distinct cell populations and complex cellular arrangement of the human kidney make it particularly difficult to adequately analyze without high-resolution spatial information. To provide that spatial information on such a large number of cells, several platforms for multiplexed imaging have recently been developed, including serial immunofluorescence staining, staining with DNA-barcoded antibodies (CODEX, CO-Detection by indEXing), and staining with metal-tagged antibodies [multiplexed ion beam imaging (MIBI) and imaging mass cytometry (IMC)]. For a detailed comparison of the strengths and weaknesses of these technologies, please see the comprehensive review by Baharlou et al. (2019).

IMAGING MASS CYTOMETRY

Imaging mass cytometry (IMC) is a powerful analytical platform in which a high-resolution laser is combined with a mass cytometer that permits mass spectrometry-based, spatially preserved high-dimensional analysis of intact FFPE and frozen tissues at a resolution of 1 μm^2 per pixel (Giesen et al., 2014; Bodenmiller, 2016; Guo et al., 2020). In IMC, a cocktail of validated antibodies against defined protein epitopes, each of which is covalently bound to a unique rare-earth metal, is hybridized on a tissue section. Pulsed ablation of 1 μm^2 spots with a UV laser is performed, with the vaporized tissue analyzed by a mass spectrometer to identify the combination of heavy metals present in that 1 μm^2 region. The quantity of each metal present at each 1 μm^2 coordinate is used to reconstruct an artificial multilayer image of the initial tissue. Currently IMC can be employed for imaging up to 42 markers on a single section of tissue. The absence of endogenous signal for the heavy metals results in a very low background and improved signal-to-noise ratio, making IMC particularly appealing as compared to fluorescence-based antibody imaging (Zhang et al., 2018).

APPLICATIONS OF IMC

IMC has been used to characterize malignancies, yielding information about the identity, number of, and spatial relationships between immune and resident cells (Giesen et al., 2014; Bodenmiller, 2016; Bertocchi et al., 2021; Li et al., 2021). IMC has also been applied in infectious and autoimmune diseases research and in drug profiling. The 1 μm^2 resolution provides the ability to identify basic subcellular localization of antigens (nucleus vs. cytoplasm), while the large number of antibodies that can be simultaneously analyzed supports both cell identification and cell activation state determination. The technique can therefore be employed to identify cellular markers that provide fundamental tissue architectural layout, as well as secondary cellular responses such as protein modifications, signaling pathway activation, cell injury states, and cell proliferation. IMC allows immune cell markers to be investigated in both healthy and diseased tissues with their distribution pattern and proximity to tissue resident cells spatially mapped (Giesen et al., 2014; Catena et al., 2020; Garcia-Melchor et al., 2021; Patel et al., 2021). Analysis of functional markers generates information on disease states and can potentially be used to identify biomarkers. Importantly, the linearity of the mass spectrometry detection of the heavy metals conjugated to each antibody provides a quantitative assessment of the relative expression levels for the respective antigens, allowing researchers to detect changes not just in cell numbers and localization, but also in cell differentiation and signaling pathway activation. This quantitative information, assigned to each pixel of the generated image, supports a more objective, machine-based evaluation of the tissue that is less subject to observer bias.

IMC has been employed to characterize disease pathophysiology by providing enhanced molecular profiling of biological tissues. With that aim, several research groups have focused on the pathogenesis of type 1 diabetes. Through analysis of 1,581 islets from 12 human donors and eight type 1 diabetic patients, using a panel of antibodies targeted against 35 biomarkers, Damond et al. (2019) found that beta cell destruction is preceded by a beta cell marker loss and by recruitment of cytotoxic and helper T cells. Similarly, using a panel of 33 antibodies and quantifying pancreatic exocrine cells, islet cells, immune cells, and stromal components, Wang et al. (2019) demonstrated a dramatic change in islet architecture, endocrine cell number, and immune cell number in pancreatic sections from type 1 diabetic patients. This study also demonstrated a molecular change indicating altered cell identity and dysregulated cellular protein expression in the histopathology of type 1 diabetes. The described ability to detect and quantify protein-level alterations in cells was reflected in a separate study of neuronal changes in postmortem samples from patients with Parkinson's disease. Using IMC analysis with a panel of antibodies targeting subunits of all five mitochondrial oxidative phosphorylation complexes, Chen et al. (2021) found a widespread decrease in expression of all complexes in Parkinson's neurons as compared to control cases. An IMC study of tissue samples from patients with multiple sclerosis characterized a pool of immune cells including macrophages adjacent to regions of demyelinization and multiple subsets of T and B cells, delineating the cellular makeup of the immune response during exacerbations of disease (Ramaglia et al., 2019). Identifying markers to characterize the pathogenesis of different disease entities allows hypothesis generation for further research. Additionally, IMC contributes to deciphering of so far unknown targets relevant to signaling pathways involved in disease pathogenesis.

IMC has also been used to further characterize hepatitis B virus-associated liver disease and most recently, COVID-19-associated organ manifestations including lung disease (Wang et al., 2020; Zhang et al., 2020; Melms et al., 2021; Rendeiro et al., 2021), brain injury (Schwabenland et al., 2021), and small intestinal infection (Lehmann et al., 2021). Allam et al. (2021) analyzed tonsillitis samples, while at the same time introducing a new algorithm to explore spatial relationships in diverse multiplexed tissue imaging data. In these publications, the corresponding antibody panels reflected potentially relevant signaling pathways in addition to immune cell composition and cross talk.

The ability of IMC to characterize both structural cells and trafficking immune cells has been important for its use in characterizing the tumor microenvironment and treatment effects or pharmacodynamics, respectively. IMC can provide: (i) a means to investigate the cellular heterogeneity of the tumor microenvironment and understanding of cancer progression and resistance to current therapies (Ijsselsteijn et al., 2019; Elaldi et al., 2021); (ii) a basis for supporting efficacy and target engagement validation studies in drug discovery (Bouzekri et al., 2019); (iii) discovery of new biomarkers (Martinez-Morilla et al., 2021); (iv) potential novel drug targets; and (v) improved

Abbreviations: IMC, Imaging mass cytometry; FFPE, Formalin-fixed paraffin-embedded; IF, Immunofluorescence; Kidney-MAPPS, Kidney-multiplexed antibody based profiling with preservation of spatial context; AQP1, Aquaporin-1; AQP2, Aquaporin-2; CK7, Cytokeratin-7; DCT, Distal convoluted tubule; KIM-1, Kidney injury molecule-1; DGF, Delayed graft function.

implementation of already established therapeutics (Carvajal-Hausdorf et al., 2019; Kankeu Fonkoua et al., 2021). These analyses include relevant cell populations and interactions, surrogates of cancer cell states, and immune cell markers. IMC-based analysis of the tumor microenvironment can potentially provide prognostic information predictive of disease outcome.

APPLICATION OF IMC IN THE HUMAN KIDNEY

Currently there are three publications showing IMC analysis of human kidney tissue (Singh et al., 2019; Wang et al., 2020; Avigan et al., 2021) and one publication analyzing murine kidney tissue (Brähler et al., 2018). The first analysis of kidney tissue by IMC was published by our group and characterized the cellular composition of the human kidney, comprising tubular, stromal, glomerular, and non-resident cells such as immune cells (Singh et al., 2019). The underlying hypothesis for the project was the relevance of interaction between kidney-resident and immune cells for homeostasis and disease development. The study (Singh et al., 2019) included the description of a machine learning-based analysis pipeline termed Kidney-MAPPS (Multiplexed Antibody based Profiling with Preservation of Spatial context) for unbiased pixel classification using Ilastik (Berg et al., 2019), nuclear-based cell segmentation using CellProfiler (McQuin et al., 2018), and clustering/neighborhood analysis using HistoCAT (Schapiro et al., 2017). In this approach, four different stacks of pseudocolored tiff images generated from the IMC data are subjected to pixel classification using Ilastik in order to create probability images representing nuclei, tubular, endothelial, and interstitial/glomerular cell types. The machine learning algorithm assigns a probability for the respective cell type to each pixel by differentiating between signal and background in each data set. Analysis in CellProfiler creates the sequential cell segmentation on the probability images and generates a unified mask. Cells are identified based on the presence of a nucleus as the primary object, followed by segmentation of the specific cell type (tubular, endothelial, and interstitial) to delineate the borders of each individual cell. Finally, processing in HistoCAT overlays the unified mask on the raw data from IMC, allowing multiplexed data analysis by unbiased clustering using the PhenoGraph algorithm, a nearest-neighbor-based clustering approach. Individual cell phenotypes are assigned to each cell in an unsupervised manner based on the combinatorial expression of the canonical cell identification markers, supporting classification and quantification of each cell type. **Figure 1** provides a graphic overview of the steps that are used for the Kidney-MAPPS analysis pipeline.

IMC and the Kidney-MAPPS pipeline were successfully applied to 16 pathologist-verified, histopathologically normal FFPE human kidney samples (five from living donors, 11 from carefully selected, banked tumor-remote nephrectomy samples) to quantitatively characterize the cellular makeup of the reference human kidney (Singh et al., 2019). The antibodies that were employed by Singh et al. (2019) listed in **Table 1**, were first validated using a combination of literature search of single-cell expression data, *in vitro* cell expression, morphologic features identified by a pathologist, and cell co-localization of more than one cell-specific marker. Kidney-MAPPS was then employed to successfully generate a two-dimensional quantitative atlas of 22 distinct cell populations that were identified in the normal human kidney. In addition to well-characterized parts of the tubule, interstitium, and glomerulus, the analysis revealed several cell populations that had not yet been described. A population of megalin (low), aquaporin-1 (AQP1)$^+$, aquaporin-2 (AQP2)$^-$, cytokeratin-7 (CK7)$^+$ tubular cells in the medulla were assigned to represent the thin descending limb of the loop of Henle. A cluster of megalin$^+$, AQP1$^+$, vimentin$^+$ cells was suggestive of an injured, fibrotic, or regenerative cell type, while CK7$^+$, AQP2$^+$ cells were felt to represent principal cells. Moreover, the analysis identified cells in connecting tubular segments with morphologically larger diameter than the distal convoluted tubule (DCT) as calbindin$^+$, CK7$^+$, AQP2$^+$ transition cells between the DCT and collecting duct. The Kidney-MAPPS protocol accurately identified, quantified, and localized ~92% of all cells in the human kidney, and quality control measures for validity and reproducibility of unsupervised analysis of data showed no significant differences in the manual vs. Kidney-MAPPS assigned phenotypes, out of over 10,000 cells scored (Singh et al., 2019). Further, the reproducibility analysis of Kidney-MAPPS yielded highly consistent quantitative results when adjacent sections of the same tissue were analyzed and showed no qualitative decrement in staining intensity in FFPE samples banked for >9 years. Tissue analysis on regions corresponding to the original region of interest with a new, identical antibody cocktail 4 months later also led to robust results (Singh et al., 2019).

Our group further expanded the application of IMC and the Kidney-MAPPS analysis pipeline to analyze pre-implant kidney biopsy samples from deceased donor kidneys that were pre-identified as being at high risk for delayed graft function (DGF) following implantation as compared to living donor kidneys that were low risk for subsequent DGF (Avigan et al., 2021). The results showed that the high-risk deceased donor kidneys exhibited a highly significant reduction in tubular cells, and specifically proximal tubule cells, as compared to living donor kidneys. Of note, this decrease in tubular cell numbers did not reflect failure to successfully detect nuclei or failure to assign cell identity in the deceased donor kidney tissue, but rather a quantitative reduction of the number of tubular cells present per area analyzed. Moreover, consistent with our previous study (Singh et al., 2019), $99.5 \pm 4.6\%$ of tubular cells were correctly identified by the Kidney-MAPPS pipeline compared with manual adjudication, indicating that identification of kidney cell populations is not impaired between reference and injured tissues. Intriguingly, the study uncovered a small population of megalin(low), vimentin$^+$ proximal tubular cells exclusively in deceased donor tissue that was surrounded by macrophage-rich infiltrates and co-expressed markers of injury and proliferation, kidney injury molecule 1 (KIM-1) and Ki67, respectively (Avigan et al., 2021). This is consistent with animal data showing that alternatively activated macrophages directly surround the proliferating proximal tubule S3 segment tubular cells during the week after kidney injury (Huen et al., 2015). Moreover,

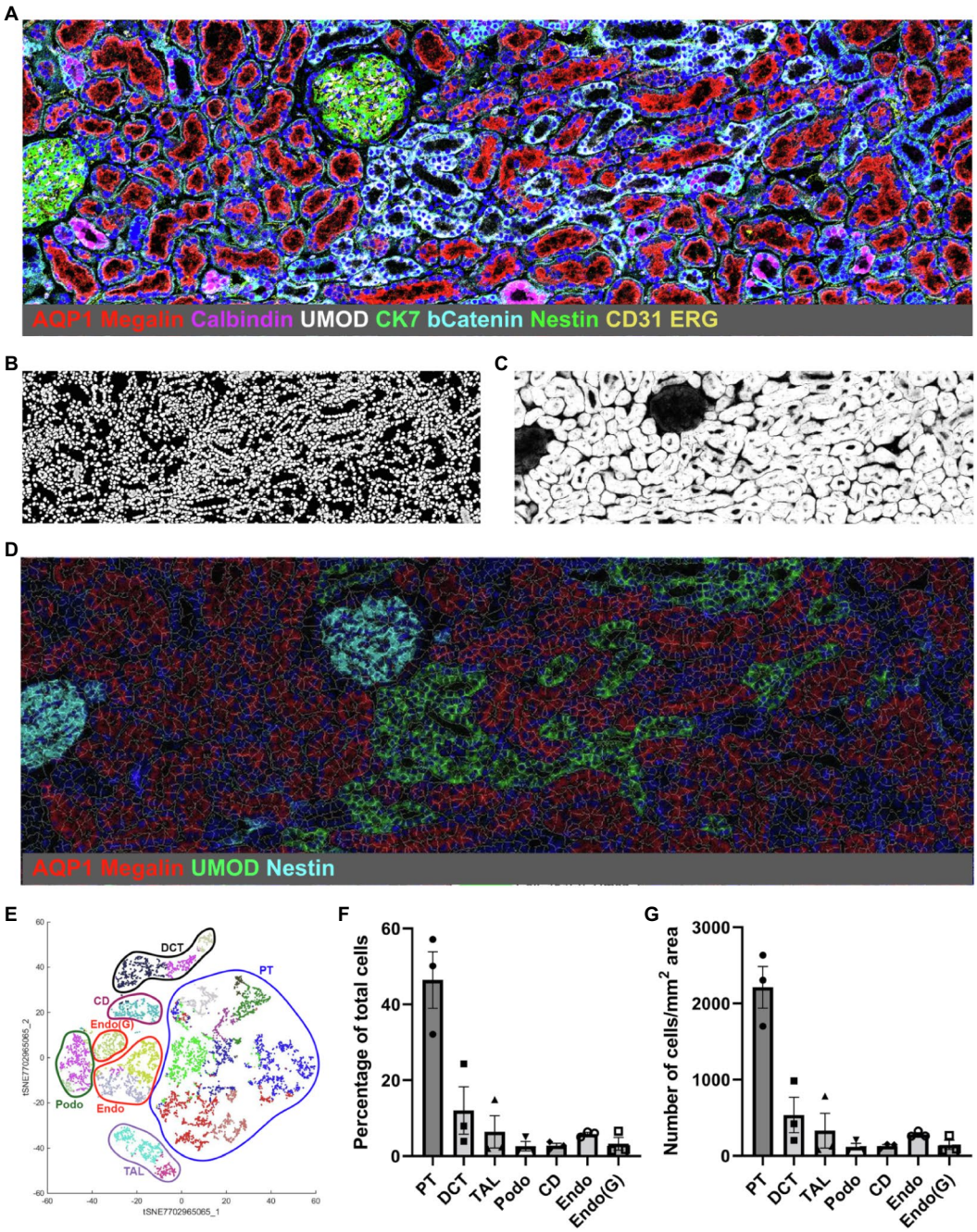


FIGURE 1 | Representative Kidney-MAPPS analysis of healthy human kidney tissue for identification of resident cell populations. **(A)** Pseudocolored image of a region of interest analyzed by IMC. The depicted colors correspond to the markers listed on the image. AQP1, aquaporin-1; UMOD, uromodulin; bCatenin, β-Catenin; CK7, cytokeratin-7; and ERG, ETS-related gene. **(B,C)** Probability images for nuclei **(B)** and tubular cells **(C)**, created using the Ilastik software. **(D)** Segmentation mask generated in CellProfiler overlaid on a pseudocolored image created in HistoCAT. Highlighted markers were selected in order to represent the tissue structure analogous to **(A)** while allowing visibility of the cell segmentation. **(E)** tSNE plot showing different cell types/clusters identified through PhenoGraph analysis of three representative healthy kidney samples in HistoCAT. PT, proximal tubule; DCT, distal convoluted tubule; TAL, thick ascending limb; Podo, podocytes; CD, collecting duct; Endo, endothelial cells (non-glomerular); and Endo(G), glomerular endothelial cells. **(F,G)** Quantification of the cell types identified in **(E)** as a percentage of the total number of cells **(F)** and the number of cells per mm^2 area **(G)**. Bars show the mean ± SD. $n = 3$.

this cell population potentially plays a role in tubular repair because KIM-1 has been shown to be protective in models of ischemia reperfusion injury and renal transplantation (Lee et al., 2018; Al-Bataineh et al., 2021). These results demonstrate that IMC can provide spatially accurate data regarding biologically important cell–cell interactions.

TABLE 1 | List of antibodies used on human kidney tissue.

Antigen	Target	Species	Clone	Metal	Dilution	Supplier	References
DNA intercalator	Nucleus	N/A	N/A	191Ir/193Ir	1:1000	Fluidigm	Singh et al., 2019; Avigan et al., 2021
Histone H3	Nucleus	Rabbit	D1H2	176Yb	1:600	Fluidigm	Singh et al., 2019
β-Catenin	Tubular epithelium	Mouse	D13A1	165Ho	1:250	Fluidigm	Avigan et al., 2021
β-Catenin	Tubular epithelium	Mouse	D10Ab	147Sm	1:500	Fluidigm	Singh et al., 2019
Megalin	Proximal tubule	Mouse	10D5.1	174Yb	1:250	EMD millipore	Avigan et al., 2021
Aquaporin-1	Proximal tubule	Rabbit	EPR11588	173Yb	1:2000	Abcam	Singh et al., 2019; Avigan et al., 2021
Uromodulin	Thick ascending limb	Rat	774056	151Eu	1:1600	R&D Systems	Singh et al., 2019; Avigan et al., 2021
Calbindin	Distal convoluted tubule	Mouse	401025	142Nd	1:400	ThermoFisher	Singh et al., 2019; Avigan et al., 2021
Aquaporin-2	Collecting duct (principal cells)	Rabbit	EPR21080	172Yb	1:300	Abcam	Singh et al., 2019; Avigan et al., 2021
Cytokeratin-7	Connecting tubule, collecting duct	Mouse	RCK105	164Dy	1:150	Fluidigm	Singh et al., 2019; Avigan et al., 2021
CD31	Endothelial cells	Mouse	JC/70A	149Sm	1:100	Abcam	Singh et al., 2019; Avigan et al., 2021
WT1	Podocytes	Rabbit	6F-H2	209Bi	1:25	ThermoFisher	Singh et al., 2019
Nestin	Podocytes	Mouse	196908	146Nd	1:200	Fluidigm	Singh et al., 2019; Avigan et al., 2021
aSMA	Smooth muscle cells	Mouse	1A4	141Pr	1:1600	Fluidigm	Singh et al., 2019; Avigan et al., 2021
Vimentin	Fibroblasts, pericytes, mesangium, podocytes	Mouse	RV202	150Nd	1:500	Abcam	Avigan et al., 2021
Vimentin	Fibroblasts, pericytes, mesangium, podocytes	Rabbit	D21H3	143Nd	1:400	Fluidigm	Singh et al., 2019
Collagen IV	Basement membrane, fibrosis	Mouse	1042	166Er	1:200	ThermoFisher	Singh et al., 2019; Avigan et al., 2021
Renin	Juxtaglomerular cells	Rabbit	EPR20693	171Yb	1:1000	Abcam	Singh et al., 2019; Avigan et al., 2021
CD68	Macrophages	Mouse	KP1	159Tb	1:600	Fluidigm	Singh et al., 2019; Avigan et al., 2021
CD66b	Granulocytes	Mouse	80H3	152Sm	1:200	Fluidigm	Singh et al., 2019; Avigan et al., 2021
CD3	T cells	Mouse	Polyclonal	170Er	1:80	Fluidigm	Singh et al., 2019; Avigan et al., 2021
CD4	Helper T cells	Rabbit	EPR6855	155Gd	1:100	Fluidigm	Singh et al., 2019; Avigan et al., 2021
CD8a	Cytotoxic T cells	Mouse	144B	162Dy	1:100	Fluidigm	Singh et al., 2019; Avigan et al., 2021
CD20	B cells	Mouse	H1	161Dy	1:100	Fluidigm	Singh et al., 2019; Avigan et al., 2021
CD11b	Granulocytes, monocytes/ Macrophages	Rabbit	EPR1344	149Sm	1:50	Fluidigm	Wang et al., 2020
CD11c	Dendritic cells	Rabbit	Polyclonal	154Sm	1:50	Fluidigm	Wang et al., 2020
CD14	Monocytes/ Macrophages	Rabbit	EPR3653	144Nd	1:50	Fluidigm	Wang et al., 2020
CD19	B cells	Rat	6OMP31	142Nd	1:50	Fluidigm	Wang et al., 2020
CD56	NK cells	Mouse	RNL-1	152Sm	1:50	Abcam	Wang et al., 2020
KIM-1	Kidney epithelial injury	Mouse	219211	160Gd	1:75	R&D Systems	Avigan et al., 2021
Ki-67	Proliferation	Mouse	B56	168Er	1:100	Fluidigm	Avigan et al., 2021
TNFalpha	Injury	Rabbit	Polyclonal	160Gd	1:50	Abcam	Wang et al., 2020

IMC has also been applied to COVID-19 patient kidney samples in a study that analyzed immune cell infiltration and proinflammatory cytokine responses in specified organs of three COVID-19 patients. Compared to human healthy control kidneys, IMC analysis showed a higher count of $CD11b^+$ macrophages, $CD11c^+$ dendritic cells, and $CD19^+$ B cells in two of the COVID-19 patient kidney samples (Wang et al., 2020). The antibodies that were used in this study are also listed in **Table 1**. By comparing the immune cell profiles and cytokines in the kidney to other organs, the authors suggested

a tissue-specific immune response to COVID-19 that might lead to differential sensitivities of resident cell populations to the COVID-19 infection and subsequent therapies.

DISCUSSION

IMC is a powerful platform for high-dimensional multiplexed parallel analysis of dozens of proteins localized in intact tissue, providing single-cell resolution while preserving the spatial relationship and morphological features. Its capacity to quantify >40 markers renders IMC a useful technique in the clinical setting where tissue quantities from patient biopsies can be very limited. The ability to use banked FFPE sections without signal attenuation and with the same data outcome as freshly procured tissue also helps with the pathophysiologic study of diseases for which there are few patient samples available. An additional advantage of IMC as compared to IF is that there is no endogenous autofluorescence to negatively impact the signal-to-noise ratio. In terms of sensitivity, IMC can successfully detect a lower limit of approximately 50 copies of an epitope per 1 μm^2 laser pulse (Hartmann and Bendall, 2020).

Limitations include the resolution that presently is at 1 μm^2, meaning that subcellular organelles such as mitochondria, endoplasmic reticulum, and vacuoles cannot be resolved and that areas of cell overlap can make pixel classification and cell segmentation less accurate. The effect this infers on marker expression must be kept in mind when assigning a certain phenotype to a cell. In addition, the acquisition of data is time-consuming, taking 2 h to ablate 1 m^2 area of interest, and approximately 10% of analyzed IMC regions did not yield analyzable data in the study by Singh et al. (2019).

IMC can be easily complemented by other conventional as well as highly multiplexed technologies. For example, Catena et al. (2018) developed a method to perform H&E counterstaining on IMC samples, allowing basic structural analysis of the tissue prior to IMC interrogation. By conjugating RNAscope probes to heavy metals rather than fluorophores, IMC can be used to simultaneously detect RNA and protein expression on the same tissue sample. Using this approach, unique cell identifiers that lack appropriate antibodies can be multiplexed with protein detection in single cells *in situ*, allowing correlation between transcriptional signatures and spatial context (Schulz et al., 2018). This helps gather additional information, i.e., on host–pathogen interactions when a pathogen is only detectable through RNA expression. Focusing on kidney analysis, the validated antibody panel together with the analysis pipeline and quantitative data on the makeup of cells in healthy kidney by Singh et al. (2019) provides a framework and reference cohort for future studies of diseased kidneys. The discovery of new cell types using these combined approaches will help us better understand and characterize the complexity of kidney tissue.

In summary, IMC can be applied to almost any tissue as long as the antibodies employed are appropriately validated and tested, and provides a unique ability to both detect and quantify multiple target antigens on a single archived or fresh tissue sample. We have found this to be particularly useful in the human kidney, which has historically presented many challenges to researchers due to its combination of cellular and architectural complexity. The ability to combine IMC with other tissue interrogation techniques such as single-cell RNA sequencing promises to further increase our understanding of the cellular changes that underlie kidney disease states, and thus identify appropriate interventions in a patient-specific manner.

AUTHOR CONTRIBUTIONS

VK and MW contributed equally to this work and wrote and edited the manuscript. LC oversaw and contributed to writing and editing the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by DK126815 and AG067335 awards to LC and KPMP Opportunity pool award (U2CDK114886) to VK and LC. Additional support came from the German Research Foundation to MW (WE6653/1-1).

ACKNOWLEDGMENTS

We thank Indiana O'Brien Center for Advanced Microscopic Analysis, Indiana Clinical and Translational Sciences Institute, Indianapolis, Indiana, for covering the costs of the publication. We thank our funding sources, including the National Institute of Diabetic and Digestive and Kidney Diseases and KPMP (to LC) and the German Research Foundation (to MW).

REFERENCES

Al-Bataineh, M. M., Kinlough, C. L., Mi, Z., Jackson, E. K., Mutchler, S. M., Emlet, D. R., et al. (2021). KIM-1-mediated anti-inflammatory activity is preserved by MUC1 induction in the proximal tubule during ischemia-reperfusion injury. *Am. J. Physiol. Ren. Physiol.* 321, F135–F148. doi: 10.1152/ajprenal.00127.2021

Allam, M., Hu, T., Cai, S., Laxminarayanan, K., Hughley, R. B., and Coskun, A. F. (2021). Spatially visualized single-cell pathology of highly multiplexed protein profiles in health and disease. *Commun. Biol.* 4:632. doi: 10.1038/s42003-021-02166-2

Avigan, Z. M., Singh, N., Kliegel, J. A., Weiss, M., Moeckel, G. W., and Cantley, L. G. (2021). Tubular cell dropout in preimplantation deceased donor biopsies as a predictor of delayed graft function. *Transplant. Direct* 7:e716. doi: 10.1097/TXD.0000000000001168

Baharlou, H., Canete, N. P., Cunningham, A. L., Harman, A. N., and Patrick, E. (2019). Mass cytometry imaging for the study of human diseases-applications and data analysis strategies. *Front. Immunol.* 10:2657. doi: 10.3389/fimmu.2019.02657

Berg, S., Kutra, D., Kroeger, T., Straehle, C. N., Kausler, B. X., Haubold, C., et al. (2019). Ilastik: interactive machine learning for (bio)image analysis. *Nat. Methods* 16, 1226–1232. doi: 10.1038/s41592-019-0582-9

Bertocchi, A., Carloni, S., Ravenda, P. S., Bertalot, G., Spadoni, I., Lo Cascio, A., et al. (2021). Gut vascular barrier impairment leads to intestinal bacteria dissemination and colorectal cancer metastasis to liver. *Cancer Cell* 39, 708.e11–724.e11. doi: 10.1016/j.ccell.2021.03.004

Bodenmiller, B. (2016). Multiplexed epitope-based tissue imaging for discovery and healthcare applications. *Cell Syst.* 2, 225–238. doi: 10.1016/j.cels.2016.03.008

Bouzekri, A., Esch, A., and Ornatsky, O. (2019). Multidimensional profiling of drug-treated cells by imaging mass cytometry. *FEBS Open Bio* 9, 1652–1669. doi: 10.1002/2211-5463.12692

Brähler, S., Zinselmeyer, B. H., Raju, S., Nitschke, M., Suleiman, H., Saunders, B. T., et al. (2018). Opposing roles of dendritic cell subsets in experimental GN. *J. Am. Soc. Nephrol.* 29, 138–154. doi: 10.1681/ASN.2017030270

Carvajal-Hausdorf, D. E., Patsenker, J., Stanton, K. P., Villarroel-Espindola, F., Esch, A., Montgomery, R. R., et al. (2019). Multiplexed (18-Plex) measurement of signaling targets and cytotoxic T cells in trastuzumab-treated patients using imaging mass cytometry. *Clin. Cancer Res.* 25, 3054–3062. doi: 10.1158/1078-0432.CCR-18-2599

Catena, R., Montuenga, L. M., and Bodenmiller, B. (2018). Ruthenium counterstaining for imaging mass cytometry. *J. Pathol.* 244, 479–484. doi: 10.1002/path.5049

Catena, R., Özcan, A., Kütt, L., Plüss, A., Consortium, I., Schraml, P., et al. (2020). Highly multiplexed molecular and cellular mapping of breast cancer tissue in three dimensions using mass tomography. bioRxiv.

Chen, C., Mcdonald, D., Blain, A., Sachdeva, A., Bone, L., Smith, A. L. M., et al. (2021). Imaging mass cytometry reveals generalised deficiency in OXPHOS complexes in Parkinson's disease. *NPJ Parkinsons Dis.* 7:39. doi: 10.1038/s41531-021-00182-x

Cippà, P. E., Sun, B., Liu, J., Chen, L., Naesens, M., and Mcmahon, A. P. (2018). Transcriptional trajectories of human kidney injury progression. *JCI Insight* 3:e123151. doi: 10.1172/jci.insight.123151

Damond, N., Engler, S., Zanotelli, V. R. T., Schapiro, D., Wasserfall, C. H., Kusmartseva, I., et al. (2019). A map of human type 1 diabetes progression by imaging mass cytometry. *Cell Metab.* 29, 755.e5–768.e5. doi: 10.1016/j.cmet.2018.11.014

Deleersnijder, D., Callemeyn, J., Arijs, I., Naesens, M., Van Craenenbroeck, A. H., Lambrechts, D., et al. (2021). Current methodological challenges of single-cell and single-nucleus RNA-sequencing in glomerular diseases. *J. Am. Soc. Nephrol.* 32, 1838–1852. doi: 10.1681/ASN.2021020157

Elaldi, R., Hemon, P., Petti, L., Cosson, E., Desrues, B., Sudaka, A., et al. (2021). High dimensional imaging mass cytometry panel to visualize the tumor immune microenvironment contexture. *Front. Immunol.* 12:666233. doi: 10.3389/fimmu.2021.666233

Garcia-Melchor, E., Cafaro, G., Macdonald, L., Crowe, L. A. N., Sood, S., Mclean, M., et al. (2021). Novel self-amplificatory loop between T cells and tenocytes as a driver of chronicity in tendon disease. *Ann. Rheum. Dis.* 80, 1075–1085. doi: 10.1136/annrheumdis-2020-219335

Giesen, C., Wang, H. A., Schapiro, D., Zivanovic, N., Jacobs, A., Hattendorf, B., et al. (2014). Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat. Methods* 11, 417–422. doi: 10.1038/nmeth.2869

Guo, N., Van Unen, V., Ijsselsteijn, M. E., Ouboter, L. F., Van Der Meulen, A. E., Chuva, D. E., et al. (2020). A 34-marker panel for imaging mass cytometric analysis of human snap-frozen tissue. *Front. Immunol.* 11:1466. doi: 10.3389/fimmu.2020.01466

Hartmann, F. J., and Bendall, S. C. (2020). Immune monitoring using mass cytometry and related high-dimensional imaging approaches. *Nat. Rev. Rheumatol.* 16, 87–99. doi: 10.1038/s41584-019-0338-z

Huen, S. C., Huynh, L., Marlier, A., Lee, Y., Moeckel, G. W., and Cantley, L. G. (2015). GM-CSF promotes macrophage alternative activation after renal ischemia/reperfusion injury. *J. Am. Soc. Nephrol.* 26, 1334–1345. doi: 10.1681/ASN.2014060612

Ijsselsteijn, M. E., Van Der Breggen, R., Farina Sarasqueta, A., Koning, F., Miranda, D. E., and N., (2019). A 40-marker panel for high dimensional characterization of cancer immune microenvironments by imaging mass cytometry. *Front. Immunol.* 10:2534. doi: 10.3389/fimmu.2019.02534

Kankeu Fonkoua, L. A., Chakrabarti, S., Sonbol, M. B., Kasi, P. M., Starr, J. S., Liu, A. J., et al. (2021). Outcomes on anti-VEGFR-2/paclitaxel treatment after progression on immune checkpoint inhibition in patients with metastatic gastroesophageal adenocarcinoma. *Int. J. Cancer* 149, 378–386. doi: 10.1002/ijc.33559

Kretzler, M., Cohen, C. D., Doran, P., Henger, A., Madden, S., Gröne, E. F., et al. (2002). Repuncturing the renal biopsy: strategies for molecular diagnosis in nephrology. *J. Am. Soc. Nephrol.* 13, 1961–1972. doi: 10.1097/01.ASN.0000020390.29418.70

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.* 10:2832. doi: 10.1038/s41467-019-10861-2

Lee, J. Y., Ismail, O. Z., Zhang, X., Haig, A., Lian, D., and Gunaratnam, L. (2018). Donor kidney injury molecule-1 promotes graft recovery by regulating systemic necroinflammation. *Am. J. Transplant.* 18, 2021–2028. doi: 10.1111/ajt.14745

Lehmann, M., Allers, K., Heldt, C., Meinhardt, J., Schmidt, F., Rodriguez-Sillke, Y., et al. (2021). Human small intestinal infection by SARS-CoV-2 is characterized by a mucosal infiltration with activated CD8(+) T cells. *Mucosal Immunol.* 14, 1381–1392. doi: 10.1038/s41385-021-00437-z

Li, R., Lin, Y., Wang, Y., Wang, S., Yang, Y., Mu, X., et al. (2021). Characterization of the tumor immune microenvironment in lung squamous cell carcinoma using imaging mass cytometry. *Front. Oncol.* 11:620989. doi: 10.3389/fonc.2021.757135

Martinez-Morilla, S., Villarroel-Espindola, F., Wong, P. F., Toki, M. I., Aung, T. N., Pelekanou, V., et al. (2021). Biomarker discovery in patients with immunotherapy-treated melanoma with imaging mass cytometry. *Clin. Cancer Res.* 27, 1987–1996. doi: 10.1158/1078-0432.CCR-20-3340

McQuin, C., Goodman, A., Chernyshev, V., Kamentsky, L., Cimini, B. A., Karhohs, K. W., et al. (2018). CellProfiler 3.0: next-generation image processing for biology. *PLoS Biol.* 16:e2005970. doi: 10.1371/journal.pbio.2005970

Melms, J. C., Biermann, J., Huang, H., Wang, Y., Nair, A., Tagore, S., et al. (2021). A molecular single-cell lung atlas of lethal COVID-19. *Nature* 595, 114–119. doi: 10.1038/s41586-021-03569-1

Patel, J., Maddukuri, S., Li, Y., Bax, C., and Werth, V. P. (2021). Highly multiplexed mass cytometry identifies the immunophenotype in the skin of dermatomyositis. *J. Investig. Dermatol.* 141, 2151–2160. doi: 10.1016/j.jid.2021.02.748

Ramaglia, V., Sheikh-Mohamed, S., Legg, K., Park, C., Rojas, O. L., Zandee, S., et al. (2019). Multiplexed imaging of immune cells in staged multiple sclerosis lesions by mass cytometry. *elife* 8:e48051. doi: 10.7554/eLife.48051

Rendeiro, A. F., Ravichandran, H., Bram, Y., Chandar, V., Kim, J., Meydan, C., et al. (2021). The spatial landscape of lung pathology during COVID-19 progression. *Nature* 593, 564–569. doi: 10.1038/s41586-021-03475-6

Rost, S., Giltnane, J., Bordeaux, J. M., Hitzman, C., Koeppen, H., and Liu, S. D. (2017). Multiplexed ion beam imaging analysis for quantitation of protein expression in cancer tissue sections. *Lab. Investig.* 97, 992–1003. doi: 10.1038/labinvest.2017.50

Schapiro, D., Jackson, H. W., Raghuraman, S., Fischer, J. R., Zanotelli, V. R. T., Schulz, D., et al. (2017). histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data. *Nat. Methods* 14, 873–876. doi: 10.1038/nmeth.4391

Schulz, D., Zanotelli, V. R. T., Fischer, J. R., Schapiro, D., Engler, S., Lun, X. K., et al. (2018). Simultaneous multiplexed imaging of mRNA and proteins with subcellular resolution in breast cancer tissue samples by mass cytometry. *Cell Syst.* 6, 25.e5–36.e5. doi: 10.1016/j.cels.2017.12.001

Schwabenland, M., Salié, H., Tanevski, J., Killmer, S., Lago, M. S., Schlaak, A. E., et al. (2021). Deep spatial profiling of human COVID-19 brains reveals neuroinflammation with distinct microanatomical microglia-T-cell interactions. *Immunity* 54, 1594.e11–1610.e11. doi: 10.1016/j.immuni.2021.06.002

Singh, N., Avigan, Z. M., Kliegel, J. A., Shuch, B. M., Montgomery, R. R., Moeckel, G. W., et al. (2019). Development of a 2-dimensional atlas of the human kidney with imaging mass cytometry. *JCI Insight* 4:e129477. doi: 10.1172/jci.insight.129477

Wang, Y. J., Traum, D., Schug, J., Gao, L., Liu, C., Atkinson, M. A., et al. (2019). Multiplexed *in situ* imaging mass cytometry analysis of the human endocrine pancreas and immune system in type 1 diabetes. *Cell Metab.* 29, 769.e4–783.e4. doi: 10.1016/j.cmet.2019.01.003

Wang, C., Xu, J., Wang, S., Pan, S., Zhang, J., Han, Y., et al. (2020). Imaging mass cytometric analysis of postmortem tissues reveals dysregulated immune cell and cytokine responses in multiple organs of COVID-19 patients. *Front. Microbiol.* 11:600989. doi: 10.3389/fmicb.2020.600989

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, 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, 2069–2080. doi: 10.1681/ASN.2018020125

Zhang, Y., Gao, Y., Qiao, L., Wang, W., and Chen, D. (2020). Inflammatory response cells during acute respiratory distress syndrome in patients with coronavirus disease 2019 (COVID-19). *Ann. Intern. Med.* 173, 402–404. doi: 10.7326/L20-0227

Zhang, W. R., and Parikh, C. R. (2019). Biomarkers of acute and chronic kidney disease. *Annu. Rev. Physiol.* 81, 309–333. doi: 10.1146/annurev-physiol-020518-114605

Zhang, Y., Wang, Y., Cao, W.-W., Ma, K.-T., Ji, W., Han, Z.-W., et al. (2018). Spectral characteristics of autofluorescence in renal tissue and methods for reducing fluorescence background in confocal laser scanning microscopy. *J. Fluoresc.* 28, 561–572. doi: 10.1007/s10895-018-2217-4

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

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.

frontiers in Physiology

MINI REVIEW
published: 06 January 2022
doi: 10.3389/fphys.2021.809346

Principles of Spatial Transcriptomics Analysis: A Practical Walk-Through in Kidney Tissue

Teia Noel[1], Qingbo S. Wang[2,3,4,5], Anna Greka[1,6]* and Jamie L. Marshall[1]**

[1] Kidney Disease Initiative, Broad Institute of MIT and Harvard, Cambridge, MA, United States, [2] Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, United States, [3] Program in Bioinformatics and Integrative Genomics, Harvard Medical School, Boston, MA, United States, [4] Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, United States, [5] Department of Statistical Genetics, Graduate School of Medicine, Osaka University, Osaka, Japan, [6] Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States

Spatial transcriptomic technologies capture genome-wide readouts across biological tissue space. Moreover, recent advances in this technology, including Slide-seqV2, have achieved spatial transcriptomic data collection at a near-single cell resolution. To-date, a repertoire of computational tools has been developed to discern cell type classes given the transcriptomic profiles of tissue coordinates. Upon applying these tools, we can explore the spatial patterns of distinct cell types and characterize how genes are spatially expressed within different cell type contexts. The kidney is one organ whose function relies upon spatially defined structures consisting of distinct cellular makeup. Thus, the application of Slide-seqV2 to kidney tissue has enabled us to elucidate spatially characteristic cellular and genetic profiles at a scale that remains largely unexplored. Here, we review spatial transcriptomic technologies, as well as computational approaches for cell type mapping and spatial cell type and transcriptomic characterizations. We take kidney tissue as an example to demonstrate how the technologies are applied, while considering the nuances of this architecturally complex tissue.

Keywords: slide-seqV2, spatial transcriptomics, kidney spatial transcriptomics, slide-seq, kidney transcriptomics

OPEN ACCESS

Edited by:
Bruce Molitoris, Indiana University, United States

Reviewed by:
Christoph Kuppe, RWTH Aachen University, Germany

****Correspondence:***
Qingbo S. Wang
qingbow@broadinstitute.org
Anna Greka
agreka@broadinstitute.org
Jamie L. Marshall
jmarshal@broadinstitute.org

Specialty section:
This article was submitted to Renal and Epithelial Physiology, a section of the journal Frontiers in Physiology

Received: *04 November 2021*
Accepted: *26 November 2021*
Published: *06 January 2022*

Citation:
Noel T, Wang QS, Greka A and Marshall JL (2022) Principles of Spatial Transcriptomics Analysis: A Practical Walk-Through in Kidney Tissue. Front. Physiol. 12:809346. doi: 10.3389/fphys.2021.809346

APPLICATIONS OF SPATIAL TRANSCRIPTOMICS

Unbiased spatial transcriptomics adds *in situ* spatial context to single cell RNA data providing a powerful tool to characterize the spatial location for whole transcriptome sequencing (10x Genomics, 2007; Chen et al., 2015; Stegle et al., 2015; Ståhl et al., 2016; Rodriques et al., 2019; Liao et al., 2021; Longo et al., 2021; Nature Methods, 2021; Stickels et al., 2021). This is accomplished with two technologies, Spatial Transcriptomics (ST or Visium) and Slide-seqV2 (10x Genomics, 2007; Ståhl et al., 2016; Rodriques et al., 2019; Stickels et al., 2021). Other technologies, such as MERFISH (targeted), GeoMx^{TM} (targeted), DBiT-seq, and Stereo-seq (BGI), allow for higher resolution, in some cases even subcellular, detection of RNA (Chen et al., 2015, 2021; Liu et al., 2020; Zollinger et al., 2020; Su et al., 2021). However, they tend to be more complex to implement. Both ST and Slide-seqV2 use uniquely barcoded beads with Oligo(dT) to capture polyadenylated RNA. ST has a larger feature size (avg 1–10 cells per spot, 50 μm beads with 100 μm spacing between beads (10x Genomics, 2007) than Slide-seqV2 [avg 1–3 cells per spot, 10 μm beads (Rodriques et al., 2019; Stickels et al., 2021)] and thus has relatively limited resolution compared to the near single cell resolution in Slide-seqV2.

On the other hand, ST allows for profiling of a large area of tissue (typically the entire cross section) while a single Slide-seqV2 array only covers a 3 mm diameter of tissue so multiple arrays on serial cryosections are needed to cover the entire tissue cross section. The other advantage of ST is the ability to co-stain the same tissue slice from which the spatial transcriptome is captured with hematoxylin and eosin staining (H&E) for histology or targeted antibodies. Slide-seqV2 and ST capture similar numbers of UMIs across the same spot area (Stickels et al., 2021).

Human and mouse kidneys across health and disease have been profiled with both ST and Slide-seqV2 (Raghubar et al., 2020; Lake et al., 2021; Melo Ferreira et al., 2021). Raghubar et al. (2020) used ST to examine spatial transcriptomic differences between sex and species of human and mouse kidney tissue revealing differences in gene expression correlating with male versus female kidneys and human versus mouse kidneys. Raghubar et al. (2020) used ST to characterize spatial transcriptomics and in particular immune cell clusters in human kidney tissue and mouse models of kidney tissue subjected to ischemia/reperfusion injury and cecal ligation puncture (Melo Ferreira et al., 2021). Raghubar et al. (2020) used both ST and Slide-seqV2 to spatially define altered injury states and a fibrotic niche in human kidney biopsies from healthy, acute kidney injury (AKI), and chronic kidney disease (CKD) participants (Lake et al., 2021).

Our group used Slide-seqV2 to develop a spatial transcriptomic atlas of human and mouse kidney tissue in health and disease (Marshall et al., 2021). We profiled two mouse models of disease, early diabetic kidney disease (DKD, BTBR *ob/ob*) and uromodulin autosomal dominant tubulointerstitial kidney disease (ADTKD, UMOD-KI). The study also contains nine human participants with both cortex and medulla biopsies, one with early DKD and one with injury due to sustained tumor compression. Slide-seqV2 revealed the spatial location of *LYVE1*+ macrophages in human medulla with injury due to sustained tumor compression. In mice with early DKD, we revealed changes in the cellular organization of spatially restricted glomeruli. In UMOD-KI mice, we identified the spatial locations of diseased fibroblasts, macrophages, and *Trem2* + macrophages as well as an upregulation of the unfolded protein response (UPR) pathway in thick ascending limb (TAL) tubules. These results altogether demonstrated the utility of spatial transcriptomics technologies combined with downstream computational analysis to uncover previously unknown human and mouse kidney physiology. Throughout this review, we will discuss such computational approaches for cell type mapping as well as spatial cell type and transcriptomic characterizations focusing on kidney tissue.

PREPROCESSING OF SPATIAL TRANSCRIPTOMICS DATA

The Slide-seqV2 reads are first aligned and mapped to the human or mouse reference transcriptome using tools such as STAR aligner (Dobin et al., 2013). After standard quality control protocols, including filtering based on the number of genes and UMIs per bead, the Slide-seqV2 data is turned into an expression matrix where rows correspond to beads and columns to genes. Additionally, per bead, native spatial coordinates are preserved across 2-D tissue space.

MAPPING THE CELL TYPE FROM scRNA-SEQ DATA

The first step in the analysis of Slide-seqV2 data is to assign cell type identities to each bead. Accurate cell type classification is aided by external scRNA-seq datasets from published materials, where cell types are already annotated (Halpern et al., 2017; Lake et al., 2021; Subramanian et al., 2021). Different computational techniques have been developed to perform such analysis. First, we will focus on a method that was originally developed (Rodriques et al., 2019), NMFreg. Since typical scRNA-seq data is high dimensional (e.g., >10 thousands cells $\times$ >20 thousands genes, where each cell is annotated with specific known cell type), the scRNA-seq expression matrix is projected to a low-dimensional basis of factors by selection of highly variable genes followed by NMF (Lee and Seung, 1999). Choosing the number of dimensions of the low-dimensional space is not trivial; In the Slide-seq paper (Rodriques et al., 2019), the authors evaluated different numbers of dimensions k to semi-manually assign a value to the parameter (they showed that the biological interpretation is roughly consistent across different ks). Every factor is then assigned a unique cell type. To do so, for every cell in the scRNA-seq data, the method computes a loading for all factors, and assigns the cell the factor of maximum weight. Since each cell was previously annotated with a cell type identity, the cell type distribution for the cells assigned to each factor can be calculated. Every factor is then assigned the cell type of maximum count.

Having acquired a set of "metagene" factors for each cell type category, the method next utilizes this set of features to map Slide-seqV2 beads to cell types. Each bead's gene expression profile is approximated as a non-negative sum of the factors using non-negative least square regression (NNLS). Since each factor can be mapped to a unique cell type, the factor loading for each bead can be converted to cell type loading. The cell type with highest load is selected to define the cell type identity for each bead, and the beads that do not have a single dimension with clearly highest load (e.g., beads having uniform load over all the dimensions, or having >1 dimensions with $\gg 0$ load) are filtered out. More recent developments of spatial transcriptomics cell type mapping techniques will be discussed in the next chapter.

CELL TYPE IDENTIFICATION WITH SEURAT TRANSFER LEARNING

A second approach for assigning cell type identities to beads is Seurat's transfer learning method (Stuart et al., 2019). Similar to NMFreg, an scRNA-seq data set is utilized to elucidate cell type labels in a spatial transcriptomics query data set. First, the data of both the reference and the query are projected to

a shared lower-dimensional space. Anchors are then defined to map pairs of similar data points between the reference and the query, based on the amount of shared overlap between their respective neighborhoods. Next, weights are allocated to each anchor in order to emphasize mappings between similar biological states. The weights are based on the following criteria: (1) The distance between the query bead and k nearest reference neighbors and (2) The number of mutual nearest neighbors between the query bead and reference cell. Finally, cell type annotations from scRNA-seq are transferred along these anchors to the spatial transcriptomics reference using a weighted voter classifier (Kuncheva and Rodríguez, 2014). As a result of either the NMFreg or Seurat transfer learning methodologies, not only is each bead informed by a gene expression profile and spatial location, but also its overarching cell type identity. In Marshall et al. (2021), Seurat transfer learning presented higher consistency with physiologically known structures than NMFreg for all cell types in human tissue and in collecting ducts (CD), vascular smooth muscle cells (vSMC), and distal tubules in mouse tissue. An overview of mapped cell types in mouse and human tissue is shown in **Figure 1**.

DECOMPOSING SPATIAL TRANSCRIPTOMIC SPOTS INTO MULTIPLE CELL TYPES

One caveat of many spatial transcriptomics methods, including Slide-seqV2, is that data points are not necessarily representative of single cells. Even in near single-cell resolution spatial transcriptomic technologies such as Slide-seqV2, while data points are about the size of a single cell, their fixed locations may overlap with and capture mRNA from multiple cells. This is particularly apparent when spatial transcriptomics technologies are applied to kidney tissue, wherein the cell type landscape is dense and complex (Chabardès-Garonne et al., 2003; Raghubar et al., 2020). As a result, one collective shortcoming of methods like NMFreg and Seurat transfer learning is that they assume a one-to-one relationship between cell types and data points. More recent approaches have tried to dissect the multifaceted cell type profile of spatial transcriptomics spots, leading to a higher-resolution understanding of cell type distributions across tissue space. One method is RCTD, or robust cell type decomposition (Cable et al., 2021). RCTD leverages a generalized linear model for the total counts per gene on each data point. Specifically, counts are assumed to be Poisson distributed with the rate parameter being the product of the total molecules corresponding to a query data point and a mixture metric representative of the contribution of each gene to all cell types. In particular, one component of the mixture metric is the weighted sum of the average gene expression profiles per cell type (derived from a scRNA-seq reference dataset). One notable asset of RCTD is that it explicitly incorporates into its model platform effects that may mask true biological signals. Ultimately the weights that best fit the observed gene counts in a query data point indicate the proportions of each cell type captured by this data point.

A second example of a method that is able to decompose spatial transcriptomic spots into multiple cell types is SPOTLight (Elosua-Bayes et al., 2021). SPOTLight follows the same trajectory as NMFreg, using both NMF and NNLS at its essence, but incorporates a more biologically driven initialization step and additional NNLS step. First, SPOTLight learns a k-dimensional set of factors (k = number of cell type categories) from a scRNA-seq reference *via* NMF. While the initial factorization may be chosen to be random, here, the authors chose to initialize the matrices so that they encode gene markers for each cell type in the first matrix, and cell type identities of cells in the second matrix. NMF is then run to find the factorization of genes by factors and factors by cells that best fits the scRNA-seq reference. An initial round of NNLS derives the loadings of the basis of factors defined in the previous step for every spatial transcriptomics spot. A second round of NNLS formalizes the relationship between spots and cell types as follows: First, it defines a matrix of factors by cell types, where every column represents the median factor profile per cell type (derived from the first NMF step). Second, the spatial transcriptomic spots are decomposed into a weighted combination of these cell type profiles.

In comparison to RCTD, SPOTlight does not account for platform effects. On the other hand, RCTD has two modes: doublet, which assumes a maximum of two cell types per data point and full, which allows for >2 cell types. The authors forewarn that performance is largely dictated by the mode that the user opts for; in particular, full mode may be subject to overfitting the data. In the context of the kidney, which is complex and tightly structured, doublet mode is not a safe assumption to make. On the other hand, SPOTlight does not make this distinction and searches for the maximum number of cell types, and its proven accuracy is based on this framework.

Additional notable methods have been developed to increase our confidence of both singular and multiple cell type calls per spot. Confidence in cell type classifications can be hindered by the sparse UMI-capture rate of spatial transcriptomic spots in comparison to scRNA-seq. FIST addresses this issue by imputing transcriptomes based on spatial relationships of data points (Li Z. et al., 2021). SPICEMIX utilizes both intrinsic transcriptomic information and the spatial relationships of data points to infer cell type identities (Chidester et al., 2021). Cell2Location utilizes a Bayesian model to predict cell type identities (Kleshchevnikov et al., 2020). Lastly, CellDART tackles the problem of identifying the cell type composition of each spot by leveraging neural networks (Bae et al., 2021).

SPATIAL CURATION OF CELL TYPE IDENTITIES

The cell type identities assigned from the observed gene expression markers through computational techniques is a noisy estimate. For example, the set of mRNAs detected in a 2-D plane might not be representative of the global distribution. As a result, it is advantageous to have prior knowledge of the

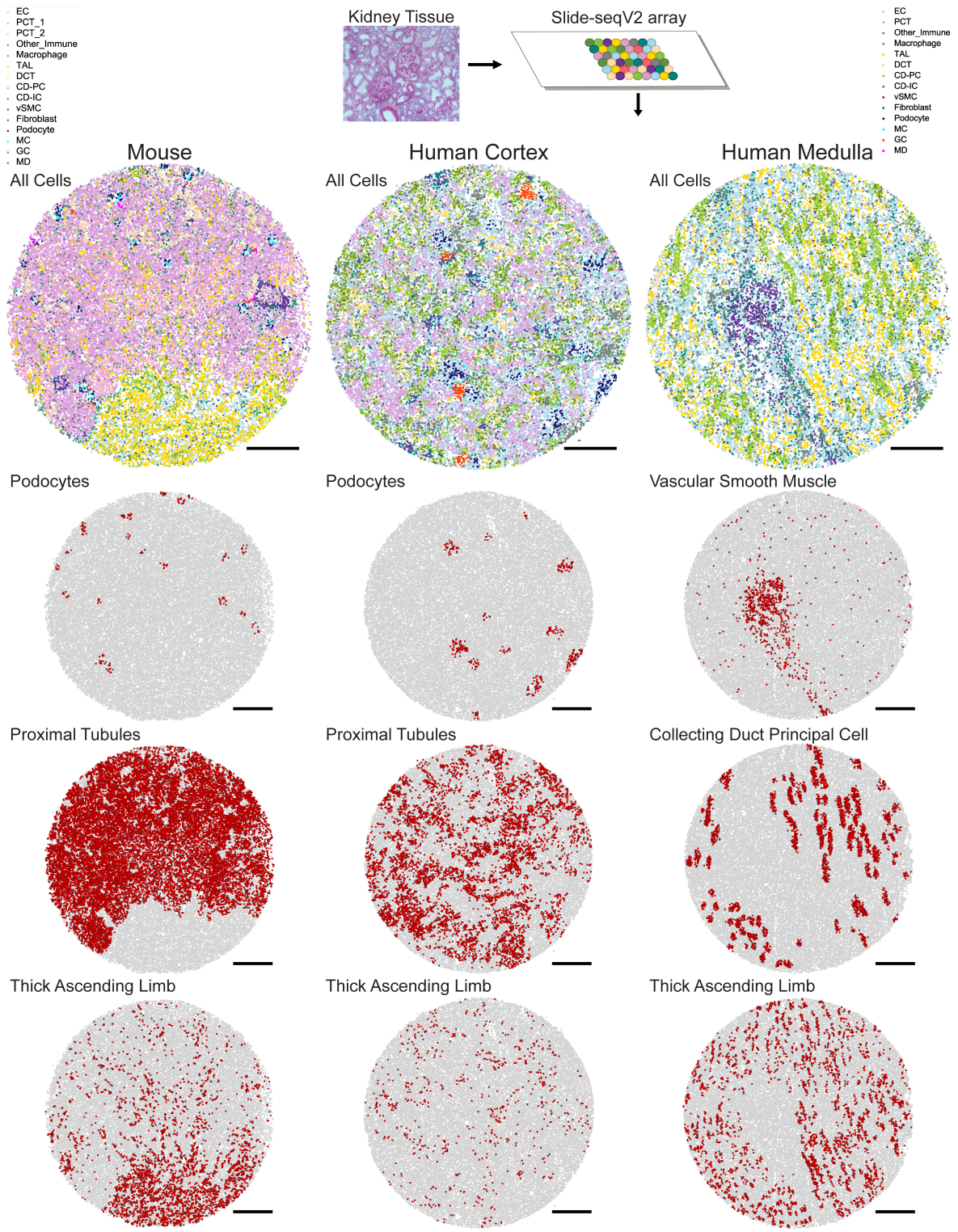


FIGURE 1 | Mapped cell types in human and mouse kidney tissue. In Slide-seqV2, a 10 μm cryosection of kidney tissue is melted onto an array containing 10 μm beads which bind to messenger RNA. Once library preparation is complete, spatially barcoded cDNA corresponding to each bead is assigned a cell identity using Seurat transfer learning. Example mouse, human cortex, and human medulla tissue with all cell types mapped are shown. Individual mappings of podocytes, proximal tubules, and thick ascending limbs are shown for the mouse and human cortex arrays, while vascular smooth muscle, collecting duct principal cell, and thick ascending limb are shown for the human medulla array. Scale bars = 500 μm.

spatial distribution of cell types across cross-sectional tissue space as a method of further filtering out noisy cell type calls. For example, we know that podocytes occur solely within circular structures called glomeruli (Garg, 2018), and distal convoluted tubules (DCT) and principal cells of the collecting duct (CD-PC) form tubular structures (McCormick and Ellison, 2015). In our work, we have implemented automated methods to identify and filter out isolated podocyte or distal tubule calls that do not occur within these denser structures. Specifically, cell type calls that have less than k nearest neighbors with the cell type identity of interest are filtered out (Marshall et al., 2021).

Furthermore, we can use prior knowledge about the known vicinity of multiple cell type classes to hone in on biologically sensical cell type calls. For example, granular cells, macula densa, and glomeruli are known to be adjacent to each other in a structure called the juxtaglomerular apparatus (JGA)(Briggs and Schnermann, 1996). The polygon encapsulating each of these structures can be found and a filtration system can be set up where only those within x units of each other are maintained. Similarly, intercalated cells of the collecting duct (CD-IC) occur within larger tubular structures made up of CD-PC (Rao et al., 2019). Again, we can isolate instances of CD-IC that occur within x units of the edge of any CD-PC structure.

Lastly, certain cell types are thought to occur only within certain regions of tissue. In the kidney, two such regions are the cortex and the medulla. Our prior knowledge (Kriz and Kaissling, 2008) suggests that cell types belonging to glomerular structures (i.e., podocytes, mesangial cells, and endothelial cells) and proximal tubule cells occur only within the cortex. On the other hand, DCT, CD-IC, CD-PC, and the thick ascending limb (TAL) are enriched and form dense structures within the medulla. We can use our knowledge of these distributions of cells across the cortex and medulla to delineate the two regions, and further filter cell types that are called in the unlikely zones.

Not only are cell type calls trusted based on the agreement between their spatial distribution and known morphology of kidney tissue, but we can also verify their identities by looking at the relative expression of known cell type marker genes. Furthermore, cell type proportions can be computed and compared between Slide-seqV2 arrays and Hybridization Chain Reaction (HCR) images (Choi et al., 2018; Marshall et al., 2021).

DOWNSTREAM ANALYSES

Once cell type identities are well-defined across spatial transcriptomic data, the spatial distribution of data points can be explored at two levels: (1) cell types and (2) gene expression profiles within certain cell types. In terms of the former, we can simply ask what the cell type composition looks like in a tissue section at large, or in the medulla or cortex regions. Furthermore, we may be interested in the neighborhoods of certain cell types. That is, what cell types often neighbor a cell type of interest. In our work, we target cell type neighborhoods of *Trem2*+ and *LYVE1*+ macrophages using k nearest neighbors (Marshall et al., 2021). Lastly, we can look at the morphology of cell type structures. Utilizing computational geometry (Gillies et al., 2007[1]; Bellock et al., 2021), we can compute the convex hull or alpha shape of the coordinates of cell types, and compute the area of these polygons. We can then ask if, on average, these metrics shift across groupings of arrays (e.g., healthy and diseased samples).

The full potential of spatial transcriptomics arises when we look into the spatial distribution of cell types, paired with their gene expression profiles. Searching for spatially non-random distributions of gene expression can be done independent of cell type information, but it is hard to distinguish gene hits that are simply markers for cell types, which themselves are characterized by distinct spatial organization. For this reason, it is useful to look for spatially non-random genes within specific cell types. Several methods currently exist that accomplish this task of discovering spatially non-random genes. In the first Slide-seq paper (Rodriques et al., 2019), the authors presented a permutation-based approach, which compares the distribution of randomly selected beads (while accounting for the total number of transcripts for each bead) with the distribution of beads expressing the specific gene of interest. They also defined spatially overlapping/anticorrelating gene pairs, regional gene enrichment and other interpretable measures using similar permutation based methods, to tackle important biological questions such as quantification and visualization of local transcriptional responses to injury. It is notable, however, that the permutation processes they used were time-consuming and included manual downstream filtering.

Another study introduces SPARK (Sun et al., 2020), which utilizes a generative, generalized linear model, argued to be more statistically robust and computationally efficient than preceding methods. It assumes that gene counts can be modeled by an over dispersed Poisson distribution, where the rate parameter is dictated by a non-spatial component (weighted sum of covariates, including batch information, cell-cycle information etc.), and a spatial component, dictated by either a gaussian distribution (representative of a localized gene expression pattern) or periodic distribution. Each component has a different variance, and the method addresses the null hypothesis that the variance of the spatial component is 0. The SPARK developers argue that a generative statistical model is far more efficient than permutation-based approaches.

Another unique advantage of coupling gene expression data with spatial information is inferring cell interactions. Previously, cell interactions were drawn from scRNA-seq, identifying ligand-receptor co-expression across cell type pairs (Cabello-Aguilar et al., 2020; Lu et al., 2021). However, there was no way of knowing if identified signaling was truly occurring between cell types in close vicinity to each other. With the onset of new spatial transcriptomics technologies, new methods have been developed to increase our confidence in identifying signaling between cell types by accounting for spatial information. Giotto looks for coexpression of ligand-receptor pairs, specifically within cell types in close vicinity to each other (Dries et al., 2021b), while MESSI is a method that uses multi-task learning to

[1]Gillies, S., et al. (2007). *Shapely: Manipulation and Analysis of Geometric Objects*. Available online at: https://github.com/shapely/shapely/blob/main/CITATION.txt

predict response genes to intracellular and intercellular signals, considering expression from neighboring cells (Li D. et al., 2021).

CONCLUSION

In this manuscript, we have reviewed the biological and computational basis of spatial transcriptomics analysis, with an example of cell type mapping and downstream applications in kidney tissue. Spatial transcriptomics technologies are evolving at a rapid pace and extending beyond transcriptomics into metabolomics and proteomics (Lundberg and Borner, 2019; Ganesh et al., 2021; Yuan et al., 2021). The integration of unbiased spatial omics technologies will provide a powerful set of tools to characterize disease processes in intact tissue (Dries et al., 2021a). We hope that these technologies will not only develop data rich atlases of healthy and diseased tissues, but also provide a platform for advances in the fundamental understanding of disease mechanisms and highlight new therapeutic targets.

AUTHOR CONTRIBUTIONS

TN, QW, and JM designed, wrote, and edited the manuscript. JM prepared the figure. AG edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

BroadIgnite (JM) CZI Seed Networks grant number CZI2019-02447 (AG).

ACKNOWLEDGMENTS

The authors thank the entire Chen and Greka labs at the Broad Institute of MIT and Harvard for helpful discussions throughout this project.

REFERENCES

10x Genomics (2007). *Spatial Gene Expression – 10x Genomics*. Available online at: https://www.10xgenomics.com/products/spatial-gene-expression (accessed October 31, 2021).

Bae, S., Na, K. J., Koh, J., Lee, D. S., Choi, H., Kim, Y. T., et al. (2021). CellDART: cell type inference by domain adaptation of single-cell and spatial transcriptomic data. *bioRxiv* [Preprint]. doi: 10.1101/2021.04.26.441459

Bellock, K., Godber, N., and Kahn, P. (2021). *bellockk/alphashape: v1.3.1 Release*. Zenodo. doi: 10.5281/ZENODO.4697576

Briggs, J. P., and Schnermann, J. B. (1996). Whys and wherefores of juxtaglomerular apparatus function. *Kidney Int.* 49, 1724–1726. doi: 10.1038/ki.1996.255

Cabello-Aguilar, S., Alame, M., Kon-Sun-Tack, F., Fau, C., Lacroix, M., Colinge, J., et al. (2020). SingleCellSignalR: inference of intercellular networks from single-cell transcriptomics. *Nucleic Acids Res.* 48:e55. doi: 10.1093/nar/gkaa183

Cable, D. M., Murray, E., Zou, L. S., Goeva, A., Macosko, E. Z., Chen, F., et al. (2021). Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat. Biotechnol.* 1–10. doi: 10.1038/s41587-021-00830-w [Epub ahead of Print].

Chabardès-Garonne, D., Mejéan, A., Aude, J. C., Cheval, L., Di Stefano, A., Gaillard, M. C., et al. (2003). A panoramic view of gene expression in the human kidney. *Proc. Natl. Acad. Sci. U.S.A.* 100, 13710–13715.

Chen, A., Liao, S., Cheng, M., Ma, K., Wu, L., Lai, Y., et al. (2021). Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball patterned arrays. *bioRxiv* [Preprint]. doi: 10.1101/2021.01.17.427004

Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S., and Zhuang, X. (2015). RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* 348:aaa6090.

Chidester, B., Zhou, T., and Ma, J. (2021). SPICEMIX: integrative single-cell spatial modeling for inferring cell identity. *bioRxiv* [Preprint]. doi: 10.1101/2020.11.29.383067

Choi, H. M. T., Schwarzkopf, M., Fornace, M. E., Acharya, A., Artavanis, G., Stegmaier, J., et al. (2018). Third-generation hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. *Development* 145:dev165753. doi: 10.1242/dev.165753

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi: 10.1093/bioinformatics/bts635

Dries, R., Chen, J., del Rossi, N., Khan, M. M., Sistig, A., and Yuan, G.-C. (2021a). Advances in spatial transcriptomic data analysis. *Genome Res.* 31, 1706–1718.

Dries, R., Zhu, Q., Dong, R., Eng, C.-H. L., Li, H., Liu, K., et al. (2021b). Giotto: a toolbox for integrative analysis and visualization of spatial expression data. *Genome Biol.* 22:78. doi: 10.1186/s13059-021-02286-2

Elosua-Bayes, M., Nieto, P., Mereu, E., Gut, I., and Heyn, H. (2021). SPOTlight: seeded NMF regression to deconvolute spatial transcriptomics spots with single-cell transcriptomes. *Nucleic Acids Res.* 49:e50. doi: 10.1093/nar/gkab043

Ganesh, S., Hu, T., Woods, E., Allam, M., Cai, S., Henderson, W., et al. (2021). Spatially resolved 3D metabolomic profiling in tissues. *Sci. Adv.* 7:eabd0957.

Garg, P. A. (2018). Review of podocyte biology. *Am. J. Nephrol.* 47(Suppl. 1), 3–13. doi: 10.1159/000481633

Halpern, K. B., Shenhav, R., Matcovitch-Natan, O., Tóth, B., Lemze, D., Golan, M., et al. (2017). Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. *Nature* 542, 352–356.

Kleshchevnikov, V., Shmatko, A., Dann, E., Aivazidis, A., King, H. W., Li, T., et al. (2020). Comprehensive mapping of tissue cell architecture *via* integrated single cell and spatial transcriptomics. *bioRxiv* [Preprint]. doi: 10.1101/2020.11.15.378125

Kriz, W., and Kaissling, B. (2008). Structural organization of the mammalian kidney. *Seldin Giebischs Kidney* 8, 479–563. doi: 10.1016/b978-012088488-9.50023-1

Kuncheva, L. I., and Rodríguez, J. J. (2014). A weighted voting framework for classifiers ensembles. *Knowled. Inform. Syst.* 38, 259–275. doi: 10.1007/s10115-012-0586-6

Lake, B. B., Rajasree, M., Seth, W., Qiwen, H., Melo, S. R., Kian, K., et al. (2021). An atlas of healthy and injured cell states and niches in the human kidney. *bioRxiv* [Preprint]. doi: 10.1101/2021.07.28.454201

Lee, D. D., and Seung, H. S. x (1999). Learning the parts of objects by non-negative matrix factorization. *Nature* 401, 788–791. doi: 10.1038/44565

Li, D., Ding, J., and Bar-Joseph, Z. (2021). Identifying signaling genes in spatial single-cell expression data. *Bioinformatics* 37, 968–975.

Li, Z., Song, T., Yong, J., and Kuang, R. (2021). Imputation of spatially-resolved transcriptomes by graph-regularized tensor completion. *PLoS Comput. Biol.* 17:e1008218. doi: 10.1371/journal.pcbi.1008218

Liao, J., Lu, X., Shao, X., Zhu, L., and Fan, X. (2021). Uncovering an Organ's molecular architecture at single-cell resolution by spatially resolved transcriptomics. *Trends Biotechnol.* 39, 43–58. doi: 10.1016/j.tibtech.2020.05.006

Liu, Y., Yang, M., Deng, Y., Su, G., Enninful, A., Guo, C. C., et al. (2020). High-spatial-resolution multi-omics sequencing *via* deterministic barcoding in tissue. *Cell* 183, 1665–1681.e18. doi: 10.1016/j.cell.2020.10.026

Longo, S. K., Guo, M. G., Ji, A. L., and Khavari, P. A. (2021). Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue

dynamics. *Nat. Rev. Genet.* 22, 627–644. doi: 10.1038/s41576-021-00370-8

Lu, M., Kuo, P. H., Hsu, C. Y., Chiu, Y. H., Liu, Y. W., Lu, M. L., et al. (2021). LR hunting: a random forest based cell-cell interaction discovery method for single-cell gene expression data. *Front. Genet.* 12:708835. doi: 10.3389/fgene.2021.708835

Lundberg, E., and Borner, G. H. H. (2019). Spatial proteomics: a powerful discovery tool for cell biology. *Nat. Rev. Mol. Cell Biol.* 20, 285–302. doi: 10.1038/s41580-018-0094-y

Marshall, J. L., Noel, T., Wang, Q. S., Bazua-Valenti, S., Chen, H., Murray, E., et al. (2021). High resolution slide-seqV2 spatial transcriptomics enables discovery of disease-specific cell neighborhoods and pathways. *bioRxiv* [Preprint]. doi: 10.1101/2021.10.10.463829

McCormick, J. A., and Ellison, D. H. (2015). Distal convoluted tubule. *Compr. Physiol.* 5, 45–98. doi: 10.1002/cphy.c140002

Melo Ferreira, R., Sabo, A. R., Winfree, S., Collins, K. S., Janosevic, D., Gulbronson, C. J., et al. (2021). Integration of spatial and single-cell transcriptomics localizes epithelial cell-immune cross-talk in kidney injury. *JCI Insight* 6:e147703. doi: 10.1172/jci.insight.147703

Nature Methods (2021). Method of the Year 2020: spatially resolved transcriptomics. *Nat. Methods* 18:1. doi: 10.1038/s41592-020-01042-x

Raghubar, A. M., Pham, D. T., Tan, X., Grice, L. F., Crawford, J., Lam, P. Y., et al. (2020). Spatially resolved transcriptome profiles of mammalian kidneys illustrate the molecular complexity of functional nephron segments, cell-to-cell interactions and genetic variants. *bioRxiv* [Preprint]. doi: 10.1101/2020.09.29.317917

Rao, R., Bhalla, V., and Pastor-Soler, N. M. (2019). Intercalated cells of the kidney collecting duct in kidney physiology. *Semin. Nephrol.* 39, 353–367. doi: 10.1016/j.semnephrol.2019.04.005

Rodriques, S. G., Stickels, R. R., Goeva, A., Martin, C. A., Murray, E., Vanderburg, C. R., et al. (2019). Slide-seq: a scalable technology for measuring genome-wide expression at high spatial resolution. *Science* 363, 1463–1467. doi: 10.1126/science.aaw1219

Ståhl, P. L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J. F., Magnusson, J., et al. (2016). Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* 353, 78–82.

Stegle, O., Teichmann, S. A., and Marioni, J. C. (2015). Computational and analytical challenges in single-cell transcriptomics. *Nat. Rev. Genet.* 16, 133–145.

Stickels, R. R., Murray, E., Kumar, P., Li, J., Marshall, J. L., Di Bella, D. J., et al. (2021). Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. *Nat. Biotechnol.* 39, 313–319. doi: 10.1038/s41587-020-0739-1

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M. III, et al. (2019). Comprehensive integration of single-cell data. *Cell* 177, 1888–1902.e21.

Su, G., Qin, X., Enninful, A., Bai, Z., Deng, Y., Liu, Y., et al. (2021). Spatial multi-omics sequencing for fixed tissue *via* DBiT-seq. *STAR Protoc.* 2:100532. doi: 10.1016/j.xpro.2021.100532

Subramanian, A., Vernon, K., Zhou, Y., Marshall, J. L., Alimova, M., Zhang, F., et al. (2021). Obesity-instructed TREM2high macrophages identified by comparative analysis of diabetic mouse and human kidney at single cell resolution. *bioRxiv* [Preprint]. doi: 10.1101/2021.05.30.446342

Sun, S., Zhu, J., and Zhou, X. (2020). Statistical analysis of spatial expression patterns for spatially resolved transcriptomic studies. *Nat. Methods* 17, 193–200. doi: 10.1038/s41592-019-0701-7

Yuan, Z., Zhou, Q., Cai, L., Pan, L., Sun, W., Qumu, S., et al. (2021). SEAM is a spatial single nuclear metabolomics method for dissecting tissue microenvironment. *Nat. Methods* 18, 1223–1232.

Zollinger, D. R., Lingle, S. E., Sorg, K., Beechem, J. M., and Merritt, C. R. (2020). GeoMx^{TM} RNA assay: high multiplex, digital, spatial analysis of RNA in FFPE tissue. *Methods Mol. Biol.* 2148, 331–345. doi: 10.1007/978-1-0716-0623-0_21

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

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.

frontiers in Physiology

MINI REVIEW
published: 11 January 2022
doi: 10.3389/fphys.2021.821217

From What to Why, the Growing Need for a Focus Shift Toward Explainability of AI in Digital Pathology

*Samuel P. Border and Pinaki Sarder**

Department of Pathology and Anatomical Sciences, SUNY Buffalo, Buffalo, NY, United States

While it is impossible to deny the performance gains achieved through the incorporation of deep learning (DL) and other artificial intelligence (AI)-based techniques in pathology, minimal work has been done to answer the crucial question of why these algorithms predict what they predict. Tracing back classification decisions to specific input features allows for the quick identification of model bias as well as providing additional information toward understanding underlying biological mechanisms. In digital pathology, increasing the explainability of AI models would have the largest and most immediate impact for the image classification task. In this review, we detail some considerations that should be made in order to develop models with a focus on explainability.

Keywords: digital pathology, deep learning, artificial intelligence, explainability, interpretability, machine learning, image analysis

OPEN ACCESS

Edited by:
Andrew Hall,
University of Zurich, Switzerland

Reviewed by:
Vasileios Charitatos,
University of Zurich, Switzerland

***Correspondence:**
Pinaki Sarder
pinakisa@buffalo.edu

Specialty section:
This article was submitted to Computational Physiology and Medicine, a section of the journal Frontiers in Physiology

Received: *24 November 2021*
Accepted: *07 December 2021*
Published: *11 January 2022*

Citation:
Border SP and Sarder P (2022) From What to Why, the Growing Need for a Focus Shift Toward Explainability of AI in Digital Pathology. Front. Physiol. 12:821217. doi: 10.3389/fphys.2021.821217

INTRODUCTION

In recent years, the use of artificial intelligence (AI) to classify, segment, and otherwise gain new understanding of medical data has experienced rapid growth. The incorporation of AI in histopathology has great potential, providing pathologists with the ability to quickly render diagnoses for patients in a reproducible, objective, and time-efficient manner. Recent technological advances including the growing popularity of histology slide digitization and accessibility of high-powered computational resources have given rise to a field now referred to as digital pathology (Al-Janabi et al., 2012; Bera et al., 2019; Niazi et al., 2019). While the field of digital pathology has benefited from the advances made in more general domains of AI, it is important to remember the unique considerations that must be made when attempting to understand biological mechanisms. Leveraging domain knowledge held by the medical community is crucial in the development of AI-powered frameworks with a far-reaching impact on patient outcomes.

One of the best areas to study the impact of explainability is for the task of histopathological image classification (Holzinger et al., 2017; Pocevičiūtė et al., 2020). In current practice, pathologists looking at biopsy images synthesize available information based on their decades of education and experience in order to make diagnostic decisions. If a pathologist is asked to explain what specifically influenced their decision, they are able to indicate specific areas of the slide that contain lesions, cellular characteristics, or staining intensity variations that they know are associated with a particular disease. This interpretation by pathologists is the "gold standard" of an explainable

histology system. However, this kind of patient-pathologist consultation is a rare occurrence in current practice despite demonstrated patient interest, particularly in cases of life-changing diagnosis (Gutmann, 2003; Manek, 2012; Lapedis et al., 2019). By incorporating AI-driven pipelines into their workflow, pathologists can greatly increase both the efficiency of diagnoses and their quantitative support. Complex computational models designed to tackle uncertainty through continuous exposure to diverse sets of data and intensive pathologist involvement represent a growing area of personalized medicine. Integrating prior medical knowledge with modern data science is the fundamental goal of Explainable AI, a major focus of this review.

Explainable AI is far from a novel concept in the machine learning (ML) community (Goebel et al., 2018; Tosun et al., 2020a,b). While the presentation of new approaches for *post-hoc* explainers of deep convolutional neural networks (CNNs) is outside of the scope of this review, there are a few simple steps that can increase the interpretability and explainability of an AI-driven study (**Figure 1**). These steps include as: selection of representative units at appropriate scales, extracting quantitative features to discriminate informative units, and aggregating information on the whole slide image (WSI) level in order to generate patient-level conclusions.

The rest of this review will be organized following the above steps, with a focus on presenting the benefits and drawbacks of specific approaches in the current literature.

DEFINING REPRESENTATIVE UNITS

The inherent structure of data for medical ML tasks is hierarchical, consisting of multiple levels of resolution and detail (**Figure 1**). At the highest level, of course, is the patient. Within each patient, we have the results of tests, biopsies, and scans that give pathologists a look into the state of the patient's health. In some cases, the results of genetic tests are also available which provide even finer scale information at the level of the DNA. Integrating information from lower-levels in order to make conclusions on the patient-level can be readily handled within a multi-instance learning (MIL) framework (Dietterich et al., 1997). The original example case of MIL given by Dietterich et al. describes a locked door for which there are several key rings available which might contain the correct key (Dietterich et al., 1997). Assuming the forgetful key master only knows which key rings contain keys that fit that door, we can use MIL to learn characteristics of keys on the positively labeled key rings from which predictions can be drawn for subsequent key rings. In the context of histology slides, WSIs from each patient are treated as the key rings ("bags" in MIL terminology) where the keys ("instances" in MIL terminology) are either individual pixels, patches with much smaller spatial dimensions than the full image, or annotated sub-structures within the image (Campanella et al., 2019; Hao et al., 2019; Diao et al., 2021; Lu et al., 2021). In digital pathology, how these instances are defined can markedly impact how the decisions made by a network can be interpreted in a biological context.

Unlike traditional image classification datasets like ImageNet or MS COCO, histology datasets contain images that are substantially larger in pixel dimensions (Deng et al., 2009; Lin et al., 2014). It is not unusual for WSIs to reach into the gigapixel dimensions, often with only a small fraction of input pixels containing tissue. Gilbertson et al. found that prior to employing JPG2000 compression, a WSI system could output

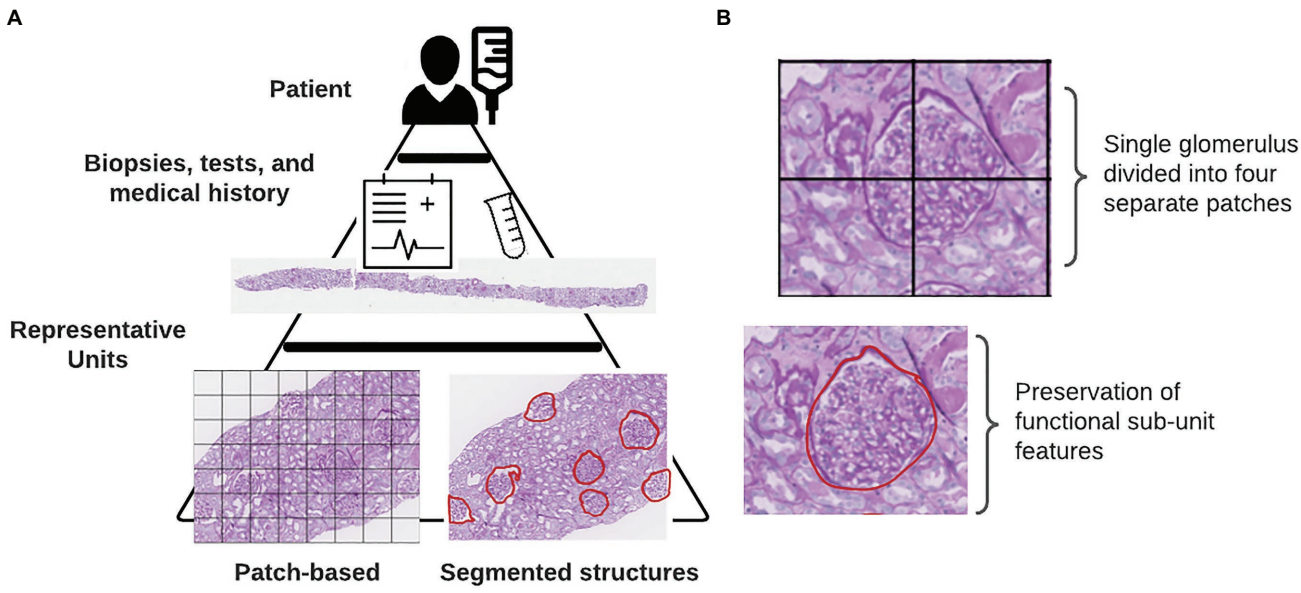


FIGURE 1 | Defining representative units. **(A)** Internal hierarchy of medical data. Each tier represents an increasing complexity or resolution of underlying biological units. At the base of this pyramid is shown the two different methods for defining representative units in a particular study. **(B)** Consequences of using a patch-based instance definition include the partitioning of functional sub-units across multiple patches.

as much as 2.7 GB of imaging data per square centimeter (Foran et al., 1997; Gilbertson et al., 2006). In order to enable DL-based analyses using data from WSIs, it is necessary to load each training image into the computer's memory. Hardware memory constraints prevent the use of entire WSIs from being used as individual training examples. A common approach used by digital pathology researchers to circumvent memory limitations when using WSIs in conjunction with DL algorithms is to break up the image into patches of equal spatial dimensions. These patches typically contain sections of tissue between 40 and 250 kilopixels (0.01–0.625 mm^2 with 0.5 µm/pixel resolution). While several existing works have achieved impressive performance treating these image patches as instances, we contend that the highest amount of explainability is obtained by instead using biologically relevant sub-compartments. Breaking the image into functional sub-units as opposed to arbitrarily assigned blocks has a better chance of conveying the biological relevance of each input object than when it is mixed in with other structures (**Figure 1**). Returning to the analogy of the key rings and the locked door, if our key ring contained thousands of keys, we can imagine that more is learned about what key will unlock the door when we focus on extracting features from each of these keys individually instead of the different parts of multiple keys grouped together.

Previous studies have been carried out with this principle in mind. Diao et al. trained a pair of CNNs to segment specific cell types and tissue regions from which they calculated quantitative features (Diao et al., 2021). This process allowed them to trace back their model predictions to specific cell or tissue types which allowed for simple localization of informative regions (Diao et al., 2021). Similarly, another study by Wang et al. developed a CNN to segment tumor regions from lung adenocarcinoma slides from which a set of 22 morphological features were used in order to predict survival probability (Wang et al., 2018). In both examples, CNNs were trained using pathologist annotations to efficiently generate datasets of specific cell and tissue types. By smartly selecting representative sub-compartments within large WSIs, model explainability is substantially increased in this study.

QUANTITATIVE FEATURE EXTRACTION

After selecting representative sub-units within a WSI, the next step in the pipeline should be to derive a way to compare these sub-units in order to assess the influence of treatment or disease in each of the provided groups. The decision of what type of features to extract from image data can have a substantial impact on the interpretability of the final results.

Standard DL approaches utilize latent features, defined as the pooled output of many convolutional filters, in order to classify images (**Figure 2**). The benefits of using this approach are that researchers are able to generate an arbitrarily large number of fine-grained features which have been shown to be highly discriminative. However, the way in which a computational model looks at image data and how a pathologist looks at image data differ immensely. Pathologists are trained to seek out particular lesions or cellular abnormalities that are known to be prognostic markers. When going through a WSI, pathologists record whether or not specific lesions were observed and at what frequency. Limitations with this kind of information include the requirement of an expert observer in order to properly catalog which results in a much larger amount of time needed per slide compared to fully computational methods. Furthermore, the semi-quantitative or qualitative nature of this kind of information can have a negative impact on inter-rater agreement.

The middle ground between the above two categories of features is referred to as hand-engineered features, which include sets of quantitative measures to describe the size, shape, texture, color, and proximity for given objects in images (**Figure 2**). These features can either directly relate to known morphological changes that are associated with disease, e.g., glomerular area in diabetic nephropathy, or indirectly examining qualitative attributes, such as the loss of mesangial matrix (mesangiolysis) through the calculation of several texture and color features. The specificity of hand-engineered features can also be modified to focus on sub-regions within each image through the use of additional segmentation methods. Color deconvolution, first proposed by Ruifrok et al. for the separation of immunohistochemically stained compartments, allows for the efficient segmentation of areas according to their biological properties (Ruifrok and Johnston, 2001). Studies, such as those by Yu et al. and Zhan et al., make use of an open-source software known as CellProfiler to quickly generate a large number of these hand-engineered features (Kamentsky et al., 2011; Yu et al., 2016; Zhang et al., 2019). CellProfiler provides the user with a wide variety of segmentation tools in a user-friendly interface for repeatable application to large image datasets (Kamentsky et al., 2011). More problem-specific feature extraction pipelines can incorporate existing domain knowledge as they condense the total amount of features to those that are known to be informative for that particular task (Ginley et al., 2019).

EXPANDING PREDICTIONS TO THE PATIENT LEVEL

After creating a quantitative understanding of features captured in each representative unit from a WSI, it now becomes necessary to understand the influence of each one of those units in the broader context of the patient. Modeling the contributions of each sub-unit on the final classification is a popular problem in the field of MIL and it is important that the manner in which instances are combined be interpretable to pathologists. Classical MIL techniques, such as Expectation Maximization-Diverse Density (EM-DD) and Axis-Parallel Rectangles, have demonstrated significant performance in defining bag-level distributions of data given feature values for weakly supervised tasks (Dietterich et al., 1997; Zhang and Goldman, 2001; Foulds and Frank, 2010; Carbonneau et al., 2018). Modern MIL approaches in digital pathology are designed to aggregate high dimensional features that are used by DL algorithms (Cosatto et al., 2013; Campanella et al., 2019; Sudharshan et al., 2019; Yao et al., 2020). Due to the stereological nature of renal biopsies, where

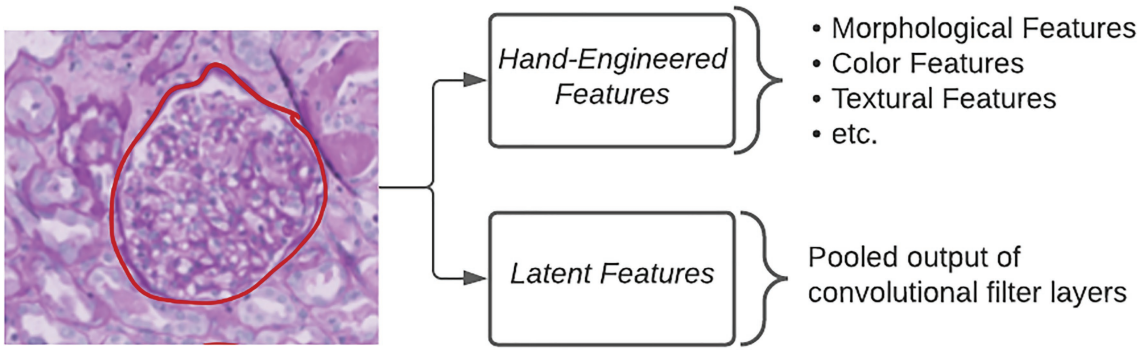


FIGURE 2 | Extracting quantitative features. Different types of quantitative features extracted from images in order to make classifications using ML algorithms.

it is not feasible to sample the entire kidney tissue, it is often the case where a small area of pixels contributes highly to the final WSI diagnosis. To best mimic this natural decision making procedure, some researchers have been incorporating recurrent neural networks (RNNs) and attention-based methods to iteratively learn to select informative regions from which patient-level conclusions are drawn. Campanella et al. employed an RNN that combined the influences of the most "suspicious" patches in order to render a diagnosis (Campanella et al., 2019). The most "suspicious" patches consisted of those with the highest ranked tumor probability by a prior CNN classifier. By using this method, authors were able to trace back their model's diagnostic predictions to a subset of image regions containing patches with the highest probability of belonging to the tumor class. Ginley et al. also demonstrated the efficacy of an RNN to aggregate handcrafted feature values for renal glomeruli presented as a sequence within each renal biopsy (Ginley et al., 2019). For their work, they were more interested in determining the most influential hand-engineered features as opposed to most influential glomeruli, which they determined using a sequential dropout procedure for each feature to measure predictive value. Attention modules were incorporated into a CNN architecture by Ilse et al. in order to differentially weight input patch influences on image class prediction (Ilse et al., 2018). Integrating how patient-level conclusions are deduced from large input images ensures that the result is both accurate and interpretable.

INCORPORATING BIOLOGICAL INTERPRETABILITY

Strict criteria for network interpretability ensure that the model correctly assesses a candidate WSI based on etiologic features that can be interpreted by pathologists. To accomplish this, computational scientists must ensure that the networks they design are not only able to accurately diagnose biopsies, but also allow for the isolation and characterization of informative regions. This characterization process should account for the inherently hierarchical nature of medical data to allow for quick determination of important areas in the slide at multiple levels of magnification. By incorporating these considerations, computational networks can better mimic the "gold standard" of diagnostic explainability.

Methods that seek to determine the focus of Neural Network (NN) models after training are referred to as "*post-hoc*" attention. This includes popular methods, such as saliency maps, deconvolutional networks, Grad-CAM, and DeepLIFT (Zeiler et al., 2010; Simonyan et al., 2013; Selvaraju et al., 2017; Shrikumar et al., 2017; **Figure 3**). While the internal operations vary, the output of each of these methods is a pixel-wise importance value for a specific classification output that is typically displayed as a heatmap overlaid on a tissue region. In addition to being a valuable tool for explaining the decisions made by a CNN, the authors of Grad-CAM also demonstrated how output heatmaps can be used as weak localization cues in a weakly supervised segmentation task (Selvaraju et al., 2017). When paired together with instance definition of functional sub-units, *post-hoc* techniques like Grad-CAM can be powerful tools in translating network predictions to approachable visual displays.

DISCUSSION

Throughout this review, we have assessed different works for their ability to provide users with sufficient levels of interpretability. In the field of digital pathology, interpretability is a critical feature of model design to ensure consistency and quality of patient treatment. Previous work by a mixture of institutions (academic, commercial, and regulatory) has highlighted concerns where AI algorithms have either introduced or mirrored systemic biases in their calculations (Minssen et al., 2020; Mehrabi et al., 2021). Without providing guidance to computational models that is based on prior knowledge, the model is forced to establish its own set of criteria that is not interpretable to a human observer. Through the incorporation of careful instance definition and hand-engineered features, the quality of algorithms using histopathological data can be elevated to the point that they are trusted for a greater range of applications. Model trust, reliability, and robustness require careful domain-specific considerations to be made so that data are appropriately processed to generate explainable results (**Table 1**).

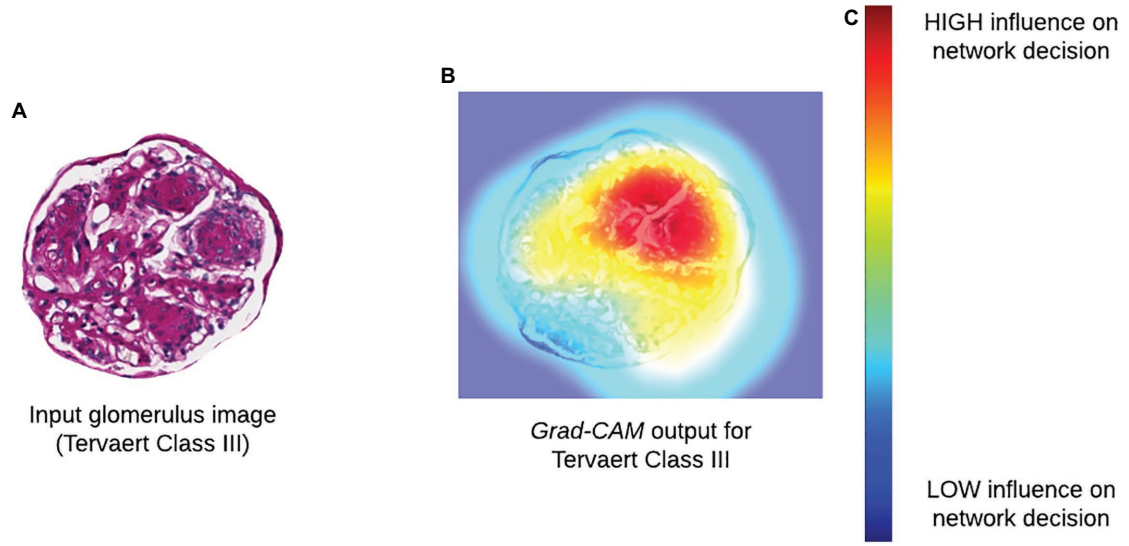


FIGURE 3 | Incorporating biological interpretability. **(A)** Input glomerulus image to a CNN trained to predict severity of progression of diabetic nephropathy according to Tervaert criteria. **(B)** *Grad-CAM* output indicating relative influence of pixels in each region within the original image. **(C)** Colormap for *Grad-CAM* heatmap illustrating degree of influence of a particular region on the decision of a network.

TABLE 1 | Glossary of terms.

Acronym used	Full expansion	Definition
DL	Deep Learning	A sub-field of Machine Learning and Artificial Intelligence where the final predicted value given an input is the result of the aggregation of information from many intermediary layers.
AI	Artificial Intelligence	A simulation of human intelligence by computers in order to solve complex problems.
ML	Machine Learning	Often used interchangeably with Artificial Intelligence, Machine Learning describes a set of algorithms wherein a computer learns to solve problems by analyzing input samples and their corresponding labels.
CNN	Convolutional Neural Network	A type of Machine Learning algorithm commonly used to classify images. Many image filters are compounded to extract information from images using convolution.
WSI	Whole Slide Image	A digitized image of a histology slide captured at full-resolution.
MIL	Multi-Instance Learning	A branch of Machine Learning dealing with data that is organized into groups.
EM-DD	Expectation Maximization-Diverse Density	A Multi-Instance Learning algorithm developed to extract characteristics of groups that best separate individual units into their respective groups.
RNN	Recurrent Neural Network	A type of Machine Learning algorithm commonly used to analyze sequences of data.
NN	Neural Network	A type of Machine Learning algorithm mimicking the flow of information between neurons in the brain.

AUTHOR CONTRIBUTIONS

SB contributed to the writing and figures for this manuscript. PS provided the guidance and editing. All authors contributed to the article and approved the submitted version.

FUNDING

This project was supported by the NIH-NIDDK grant R01 DK114485 (PS), NIH-OD grant R01 DK114485 03S1 (PS), a glue grant (PS) from the NIH-NIDDK Kidney Precision Medicine Project grant U2C DK114886 (Contact: Dr. Jonathan Himmelfarb), a multi-disciplinary small team grant RSG201047.2 (PS) from the State University of New York, a pilot grant (PS) from the University of Buffalo's Clinical and Translational Science Institute (CTSI) grant 3UL1TR00141206 S1 (Contact: Dr. Timothy Murphy), a DiaComp Pilot & Feasibility Project 21AU4180 (PS) with support from NIDDK Diabetic Complications Consortium grants U24 DK076169 and U24 DK115255 (Contact: Dr. Richard A. McIndoe), and NIH-OD grant U54 HL145608 (PS).

REFERENCES

Al-Janabi, S., Huisman, A., and Van Diest, P. J. (2012). Digital pathology: current status and future perspectives. *Histopathology* 61, 1–9. doi: 10.1111/j.1365-2559.2011.03814.x

Bera, K., Schalper, K. A., Rimm, D. L., Velcheti, V., and Madabhushi, A. (2019). Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology. *Nat. Rev. Clin. Oncol.* 16, 703–715. doi: 10.1038/s41571-019-0252-y

Campanella, G., Hanna, M. G., Geneslaw, L., Miraflor, A., Silva, V. W., Busam, K. J., et al. (2019). Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat. Med.* 25, 1301–1309. doi: 10.1038/s41591-019-0508-1

Carbonneau, M.-A., Cheplygina, V., Granger, E., and Gagnon, G. (2018). Multiple instance learning: A survey of problem characteristics and applications. *Pattern Recogn.* 77, 329–353. doi: 10.1016/j.patcog.2017.10.009

Cosatto, E., Laquerre, P.-F., Malon, C., Graf, H.-P., Saito, A., Kiyuna, T., et al. (2013). "Automated gastric cancer diagnosis on H&E-stained sections; ltraining a classifier on a large scale with multiple instance machine learning," in *Medical Imaging 2013: Digital Pathology: International Society for Optics and Photonics, 867605*.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. (2009). "Imagenet: A large-scale hierarchical image database." in *2009 IEEE Conference on Computer Vision and Pattern Recognition*; June 20-25, 2009; IEEE, 248–255.

Diao, J. A., Wang, J. K., Chui, W. F., Mountain, V., Gullapally, S. C., Srinivasan, R., et al. (2021). Human-interpretable image features derived from densely mapped cancer pathology slides predict diverse molecular phenotypes. *Nat. Commun.* 12:1613. doi: 10.1038/s41467-021-21896-9

Dietterich, T. G., Lathrop, R. H., and Lozano-Pérez, T. (1997). Solving the multiple instance problem with axis-parallel rectangles. *Artif. Intell.* 89, 31–71. doi: 10.1016/S0004-3702(96)00034-3

Foran, D. J., Meer, P. P., Papathomas, T., and Marsic, I. (1997). Compression guidelines for diagnostic telepathology. *IEEE Trans. Inf. Technol. Biomed.* 1, 55–60. doi: 10.1109/4233.594046

Foulds, J., and Frank, E. (2010). A review of multi-instance learning assumptions. *Knowl. Eng. Rev.* 25, 1–25. doi: 10.1017/S026988890999035X

Gilbertson, J. R., Ho, J., Anthony, L., Jukic, D. M., Yagi, Y., and Parwani, A. V. (2006). Primary histologic diagnosis using automated whole slide imaging: a validation study. *BMC Clin. Pathol.* 6, 1–19. doi: 10.1186/1472-6890-6-4

Ginley, B., Lutnick, B., Jen, K.-Y., Fogo, A. B., Jain, S., Rosenberg, A., et al. (2019). Computational segmentation and classification of diabetic glomerulosclerosis. *J. Am. Soc. Nephrol.* 30, 1953–1967. doi: 10.1681/ASN.2018121259

Goebel, R., Chander, A., Holzinger, K., Lecue, F., Akata, Z., Stumpf, S., et al. (2018). "Explainable ai: the new 42?" in *International Cross-domain Conference for Machine Learning and Knowledge Extraction*; August 27-30, 2018; Springer, 295–303.

Gutmann, E. J. (2003). Pathologists and patients: can we talk? *Mod. Pathol.* 16, 515–518. doi: 10.1097/01.MP.0000068260.01286.AC

Hao, J., Kosaraju, S. C., Tsaku, N. Z., Song, D. H., and Kang, M. (2019). "PAGE-Net: interpretable and integrative deep learning for survival analysis using histopathological images and genomic data," in *Pacific Symposium on Biocomputing 2020*. eds. R. B. Altman, A. K. Dunler, L. Hunter, T. Murray and T. E. Klein (Singapore: World Scientific), 355–366.

Holzinger, A., Malle, B., Kieseberg, P., Roth, P. M., Müller, H., Reihs, R., et al. (2017). Towards the augmented pathologist: challenges of explainable-ai in digital pathology. arXiv [Epub ahead of preprint].

Ilse, M., Tomczak, J., and Welling, M. (2018). "Attention-based deep multiple instance learning." in *International Conference on Machine Learning*; July 10-15, 2018; PMLR, 2127–2136.

Kamentsky, L., Jones, T. R., Fraser, A., Bray, M.-A., Logan, D. J., Madden, K. L., et al. (2011). Improved structure, function and compatibility for CellProfiler: modular high-throughput image analysis software. *Bioinformatics* 27, 1179–1180. doi: 10.1093/bioinformatics/btr095

Lapedis, C. J., Horowitz, J. K., Brown, L., Tolle, B. E., Smith, L. B., and Owens, S. R. (2019). The patient-pathologist consultation program: A mixed-methods study of interest and motivations in cancer patients. *Arch. Pathol. Lab. Med.* 144, 490–496. doi: 10.5858/arpa.2019-0105-OA

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., et al. (2014). "Microsoft coco: Common objects in context." in *European Conference on Computer Vision*; September 6-12, 2014; Springer, 740–755.

Lu, M. Y., Williamson, D. F. K., Chen, T. Y., Chen, R. J., Barbieri, M., and Mahmood, F. (2021). Data-efficient and weakly supervised computational pathology on whole-slide images. *Nat. Biomed. Eng.* 5, 555–570. doi: 10.1038/s41551-020-00682-w

Manek, S. (2012). The pathology clinic–pathologists should see patients. *Cytopathology* 23, 146–149. doi: 10.1111/j.1365-2303.2012.00985.x

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., and Galstyan, A. (2021). A survey on bias and fairness in machine learning. *CSUR* 54, 1–35. doi: 10.1145/3457607

Minssen, T., Gerke, S., Aboy, M., Price, N., and Cohen, G. (2020). Regulatory responses to medical machine learning. *J. Law Biosci.* 7. doi: 10.1093/jlb/lsaa002

Niazi, M. K. K., Parwani, A. V., and Gurcan, M. N. (2019). Digital pathology and artificial intelligence. *Lancet Oncol.* 20, e253–e261. doi: 10.1016/S1470-2045(19)30154-8

Pocevičiūtė, M., Eilertsen, G., and Lundström, C. (2020). "Artificial intelligence and machine learning for digital Pathology," in *survey of XAI in digital pathology*. eds. A. Holzinger, R. Goebel, M. Mengel and H. Müller (United States: Springer), 56–88.

Ruifrok, A. C., and Johnston, D. A. (2001). Quantification of histochemical staining by color deconvolution. *Anal. Quant. Cytol. Histol.* 23, 291–299.

Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (2017). "Grad-cam: Visual explanations from deep networks via gradient-based localization." in *Proceedings of the IEEE International Conference on Computer Vision*; October 22-29, 2017; 618–626.

Shrikumar, A., Greenside, P., and Kundaje, A. (2017). "Learning important features through propagating activation differences." in *International Conference on Machine Learning*; August 6-11, 2017; PMLR, 3145–3153.

Simonyan, K., Vedaldi, A., and Zisserman, A. (2013). Deep inside convolutional networks: Visualising image classification models and saliency maps. arXiv [Epub ahead of preprint].

Sudharshan, P. J., Petitjean, C., Spanhol, F., Oliveira, L. E., Heutte, L., and Honeine, P. (2019). Multiple instance learning for histopathological breast cancer image classification. *Expert Syst. Appl.* 117, 103–111. doi: 10.1016/j.eswa.2018.09.049

Tosun, A. B., Pullara, F., Becich, M. J., Taylor, D. L., Chennubhotla, S. C., and Fine, J. L. (2020b). "Histomapr™: An explainable ai (xai) platform for computational pathology solutions," in *Artificial Intelligence and Machine Learning for Digital Pathology*. eds. A. Holzinger, R. Goebel, M. Mengel and H. Muller (United States: Springer), 204–227.

Tosun, A. B., Pullara, F., Becich, M. J., Taylor, D., Fine, J. L., and Chennubhotla, S. C. (2020a). Explainable AI (xAI) for anatomic pathology. *Adv. Anat. Pathol.* 27, 241–250. doi: 10.1097/PAP.0000000000000264

Wang, S., Chen, A., Yang, L., Cai, L., Xie, Y., Fujimoto, J., et al. (2018). Comprehensive analysis of lung cancer pathology images to discover tumor shape and boundary features that predict survival outcome. *Sci. Rep.* 8:10393. doi: 10.1038/s41598-018-27707-4

Yao, J., Zhu, X., Jonnagaddala, J., Hawkins, N., and Huang, J. (2020). Whole slide images based cancer survival prediction using attention guided deep multiple instance learning networks. *Med. Image Anal.* 65:101789. doi: 10.1016/j.media.2020.101789

Yu, K.-H., Zhang, C., Berry, G. J., Altman, R. B., Ré, C., Rubin, D. L., et al. (2016). Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. *Nat. Commun.* 7:12474. doi: 10.1038/ncomms12474

Zeiler, M. D., Krishnan, D., Taylor, G. W., and Fergus, R. (2010). "Deconvolutional networks." in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*; June 13-18, 2010; IEEE, 2528–2535.

Zhang, Z., Chen, P., McGough, M., Xing, F., Wang, C., Bui, M., et al. (2019). Pathologist-level interpretable whole-slide cancer diagnosis with deep learning. *Nat. Mach. Intelligence* 1, 236–245. doi: 10.1038/s42256-019-0052-1

Zhang, Q., and Goldman, S. A. (2001). "EM-DD: An improved multiple-instance learning technique", in *Advances in neural Information Processing Systems*, 1073–1080.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

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.

frontiers in Physiology

MINI REVIEW
published: 13 January 2022
doi: 10.3389/fphys.2021.812947

Deconvolution Tactics and Normalization in Renal Spatial Transcriptomics

*Ricardo Melo Ferreira, Benjamin J. Freije and Michael T. Eadon**

Division of Nephrology, Indiana University School of Medicine, Indianapolis, ID, United States

The kidney is composed of heterogeneous groups of epithelial, endothelial, immune, and stromal cells, all in close anatomic proximity. Spatial transcriptomic technologies allow the interrogation of *in situ* expression signatures in health and disease, overlaid upon a histologic image. However, some spatial gene expression platforms have not yet reached single-cell resolution. As such, deconvolution of spatial transcriptomic spots is important to understand the proportion of cell signature arising from these varied cell types in each spot. This article reviews the various deconvolution strategies discussed in the 2021 Indiana O'Brien Center for Microscopy workshop. The unique features of Seurat transfer score methodology, SPOTlight, Robust Cell Type Decomposition, and BayesSpace are reviewed. The application of normalization and batch effect correction across spatial transcriptomic samples is also discussed.

Keywords: spatial transcriptomics, visium gene expression, single nuclear RNA sequencing, nephron, acute kidney injury, biopsy specimen

OPEN ACCESS

Edited by:
Andrew Hall,
University of Zurich, Switzerland

Reviewed by:
Christoph Kuppe,
RWTH Aachen University, Germany

****Correspondence:***
Michael T. Eadon
meadon@iupui.edu

Specialty section:
This article was submitted to Renal and Epithelial Physiology, a section of the journal Frontiers in Physiology

Received: *10 November 2021*
Accepted: *02 December 2021*
Published: *13 January 2022*

Citation:
Melo Ferreira R, Freije BJ and Eadon MT (2022) Deconvolution Tactics and Normalization in Renal Spatial Transcriptomics. Front. Physiol. 12:812947. doi: 10.3389/fphys.2021.812947

INTRODUCTION

Spatial transcriptomics was selected as Nature's Method of the year in 2020 (Marx, 2021). As presented at the 2021 O'Brien Center for Microscopy workshop, Spatial Transcriptomics (ST) represents a powerful tool to reveal *in situ* transcript expression associated with histopathologic features. Countless examples of ST in the development of human tissue atlases are available, identifying key features in breast cancer (Wu et al., 2021), Alzheimer's progression (Navarro et al., 2020), and cardiovascular development (Asp et al., 2019). In the kidney, ST has been applied to understand the regional expression differences in sepsis and ischemia reperfusion injury murine models (Janosevic et al., 2021; Melo Ferreira et al., 2021). A significant limitation of some ST techniques is their resolution. For example, Visium Spatial Gene Expression (VSGE) platform has a spot size of 55 µm and resolution of 100 µm, which invariably encompasses multiple cells within a single spot. Cell atlases of the kidney now include annotation of over 100 different cell types and cell states from a diverse pool of epithelial, stromal, and endothelial cells (Lake et al., 2021). These classes of cell types align very well with the underlying histology of the human kidney (Melo Ferreira et al., 2021). The 55 µm spot size is approximately the size of a cross sectional proximal tubule and will often capture elements of the signature from neighboring peritubular capillaries, dendritic cells, and other stromal cells. To better appreciate the contribution of less represented cell types to a spot's signature, strategies can be employed to deconvolute the proportion of signature arising in a spot using single cell and single nuclear

RNA sequencing (sc/snRNAseq) cluster identities. This brief review outlines the unique features of several deconvolution tactics discussed in the O'Brien center workshop. Normalization and batch effect correction across ST samples are also discussed.

DECONVOLUTION TECHNIQUES

An example of the output from four deconvolution techniques is provided in **Figure 1**. A human deceased donor nephrectomy without evidence of kidney disease was scored to fit the capture area of the Visium slide and a high-resolution image of the Hematoxylin and Eosin (H + E) stained tissue was taken with a Keyence BZ-X810 microscope as mosaics of 10× fields and stitched (1A). The histological image of the nephrectomy had the glomeruli identified and a magnified region is provided. The tissue was permeabilized and mRNA was captured in barcoded spots that allowed downstream informatic localization of each read after sequencing. The resulting expression of *NPHS2* (1B) is concentrated over the outlined glomeruli. Due to its 55 µm diameter, each spot generally covered multiple cell types. Below we present four methods designed to deconvolute the constituents of each spot. As a reference, we use a publicly available human kidney snRNAseq dataset (Lake et al., 2019).

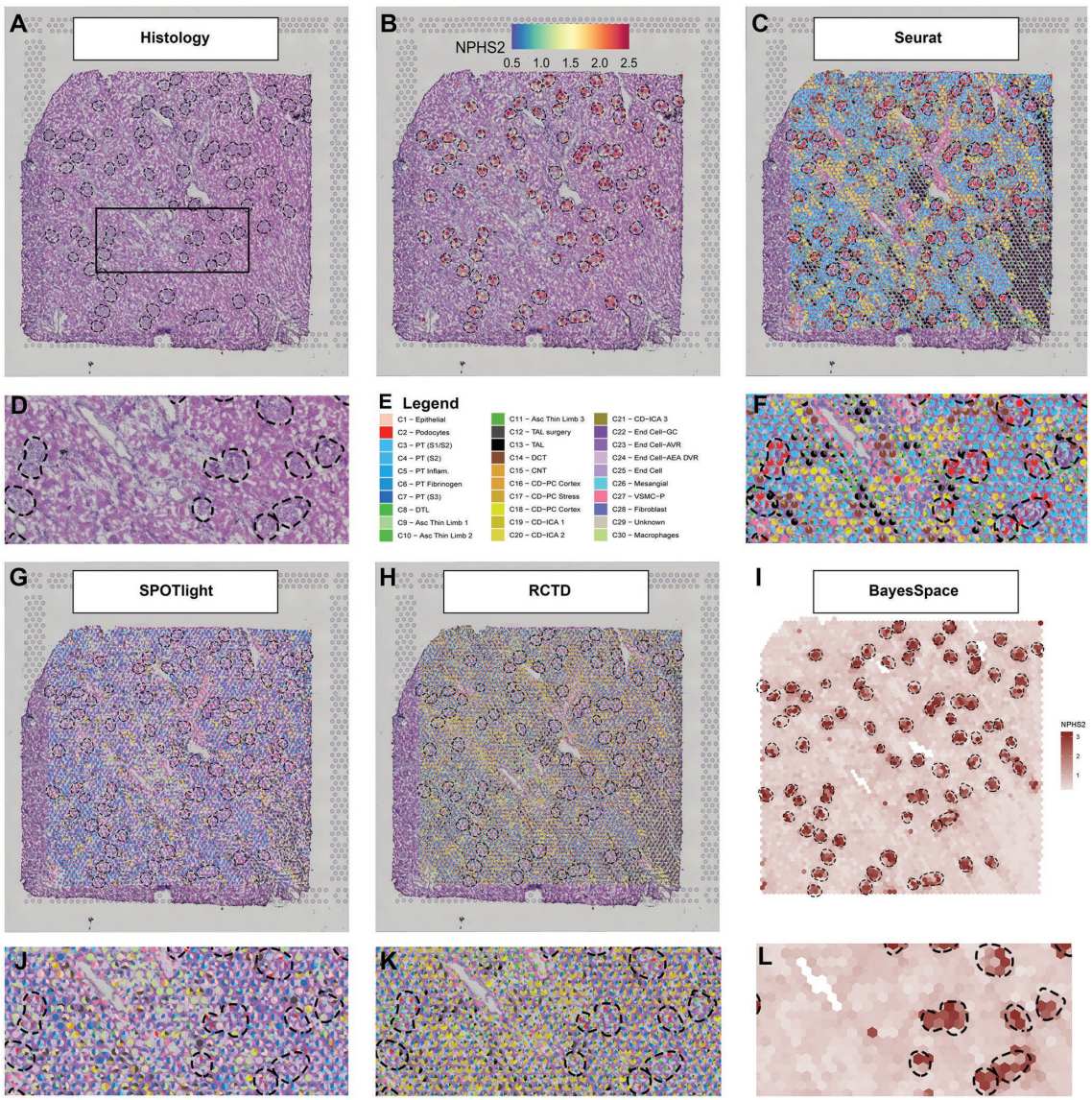


FIGURE 1 | Deconvolution techniques in spatial transcriptomics. **(A)** H + E image of a human nephrectomy sample, **(B)** *NPHS2* (podocin gene) expression localizes over glomeruli. **(C)** Seurat deconvolution in the same nephrectomy field. **(D)** Magnified field of H + E image. **(E)** Cell type identity legend. **(F)** Zoomed image of the Seurat deconvolution. **(G)** SPOTlight deconvolution. **(H)** Robust cell type decomposition deconvolution. **(I)** BayesSpace Deconvolution. **(J)** Zoomed field of SPOTlight deconvolution. **(K)** Zoomed field of RCTD. **(L)** Zoomed field of BayesSpace. Each spot is 55 µm in diameter.

SEURAT

Seurat is a popular tool to process sc/snRNAseq and ST data, with extensive documentation and support from several other analysis packages. In its version 3, Seurat introduced an anchor methodology to integrate multiple datasets (Stuart et al., 2019) that was further adapted to transfer single-cell cluster information to ST. This procedure results in a transfer score and the highest score can be used to label the spots. Alternatively, the relative scores can be displayed in a pie chart including the components of cell signature arising from multiple single-cell clusters (1C). In the example provided, the more prominent scores in the glomeruli spots are derived from the podocyte, glomerular capillary endothelial cell, and mesangial cell clusters. In the magnified region, glomeruli are surrounded by enriched areas of various proximal tubule (PT), distal convoluted tubule (DCT), and collecting duct (CD) cell signatures, as expected. The Seurat pipeline for ST analysis is still under development. For example, version 3.6.3 presents remarkable agreement between the snRNAseq cell type signatures and the underlying kidney histopathology, with the expected proportion of cell signature correlating strongly with the quantitative proportions of cells in the histology (Lake et al., 2021; Melo Ferreira et al., 2021). However, we have noted reduced alignment between the histology and snRNAseq cluster identity in Seurat version 4.

SPOTLIGHT

The SPOTlight deconvolution method uses a negative matrix factorization regression algorithm to define topics as distributions of gene expression across cell types in the reference dataset. Those topics are then used to define the cell type composition of spots and is directly related to cell type expression profiles (Elosua-Bayes et al., 2021). The results are given in proportions, which are easily interpretable. Its source code was adapted to display deconvolution results in three of the four methods discussed in this review. In the example nephrectomy (**Figure 1G**), endothelial cell type signatures, both afferent and efferent arterioles (AEA) and descending vasa recta (DVR), dominated the mapping in the tissue, including spots overlaying glomerular histology. Other expected cell types, such as podocytes, glomerular capillary endothelial cells, and mesangial cells, contributed to the cell signature in a disproportionately smaller degree than the underlying histopathology would suggest. The macrophage signature also contributed to a large proportion of spots in the tubulointerstitium and across the tissue. This methodology may require further adjustment of parameters for the kidney because so many distinct functional structures (glomerulus, PT, DCT, etc.) are located in close proximity to each other. In the example provided, the technique identified a greater proportion of signature from components that are broadly distributed across the whole kidney (like endothelial cells and macrophages) rather than specific localized cell types (like podocytes or DCT cells). However, SPOTlight provides several tools to evaluate and correct the deconvolution method, and with adjustments, the alignment between the histology and snRNAseq cluster identity can be improved.

ROBUST CELL TYPE DECOMPOSITION

Robust cell type decomposition (RCTD; Cable et al., 2021) also defines cell type transcriptomic profiles. This approach considers each spot as a mixture of cells and fits a statistical model to determine each spot composition. Our results (**Figure 1H**) show a large contribution of endothelial cell types [afferent arteriole (AEA), DVR] in the glomerular spots, with podocytes and the glomerular capillary endothelial cells represented to a lesser extent. A very minor contribution is observed from the mesangial cell cluster. Across the nephrectomy, the contribution of proximal tubules to the signature is disproportionately low compared to the histology and the collecting duct signature is minor dominant. RCTD potentially performs better on other ST technologies (like slideSEQ) where more than two cell types are rarely seen underlying a single spot (Stickels et al., 2021). The deconvolution method in the algorithm is designed to report the confidence of doublets or singlets underlying a spot.

BAYESSPACE

The BayesSpace method approaches deconvolution differently than the three previous examples. Instead of deconvoluting the cell types of each spot, it aims to increase the spatial resolution by interpolating the expression between spots (Zhao et al., 2021). This method applies an unsupervised clustering algorithm to the data that requires an *a priori* definition of the number of clusters. It then interpolates expression and defines those clusters in higher resolution. As an example, we present the interpolated expression of *NPHS2* (**Figure 1I**). The expression interpolation could be useful to predict gene expression in smaller structures. The interpolated clusters would be an excellent target to apply cell type decomposition algorithms. However, the Seurat, SPOTlight, and RCTD methods are not currently compatible with BayesSpace because these methods would require either raw counts or method-specific normalized expression to integrate with BayesSpace.

NORMALIZATION AND BATCH CORRECTION

In an effort to create a spatially anchored atlas of the kidney, analysis of multiple ST samples is invariably expected. On the VSGE platform, four samples are run in parallel on a single slide which can lead to batch effects between slides. Furthermore, variations in sample quality can lead to downstream differences in the number of reads mapped to exons in each spot. Differences in permeabilization time, RNA quality, tissue thickness, and tissue sources all contribute to the between sample variability.

In sc/snRNAseq, technical variations are reduced through normalization and batch correction, typically through programs, such as ComBat-seq, Harmony, Liger, and Seurat 3 (Welch et al., 2018; Korsunsky et al., 2019; Stuart et al., 2019; Zhang et al., 2020).

To normalize and batch correct ST data, we provide an example of the regularized negative binomial regression normalization technique, known as SCTransform (Hafemeister and Satija, 2019). To showcase its utility in human samples, nine samples across 3 batches were merged *via* the merge function and normalized or batch-corrected *via* SCTransform. Without normalization and batch correction, the ST samples exhibited inconsistent expression of the house-keeping genes *ACTB* and *GAPDH*, demonstrating a potential need for normalization when comparing across samples (**Figure 2**). Normalization with SCTransform yielded more comparable gene expression of *ACTB* and *GAPDH* across samples. The inclusion of batch as a variable in the SCTransform tool revealed only a minor additional improvement in gene expression alignment compared to normalization without a distinct batch effect variable. This indicates that technical variation in our samples can be modeled by sequencing depth alone. Together, these results suggest SCTransform may be a useful tool for removing intersample technical variation in ST datasets.

CONCLUSION

This brief review presents the result of four common deconvolution techniques and a common normalization procedure applied to the human kidney, as discussed in the 2021 O'Brien Center for Microscopy workshop. The VSGE platform facilitates direct mapping of expression signatures over a H+E stained image. While every organ is different, the kidney has many small, functionally distinct parts of the nephron, all lying in close proximity to each other. Thus, deconvolution of larger spot sizes is essential to mapping the ST signatures. Further, normalization and batch effect correction are important because an atlas must integrate data from multiple sources. The results of the deconvolution methods varied considerably, even when interrogating the same field of tissue. Some methods yielded signatures approximating the underlying histology and others emphasized less abundant cell types. No judgment has been made as to whether the cell type proportions

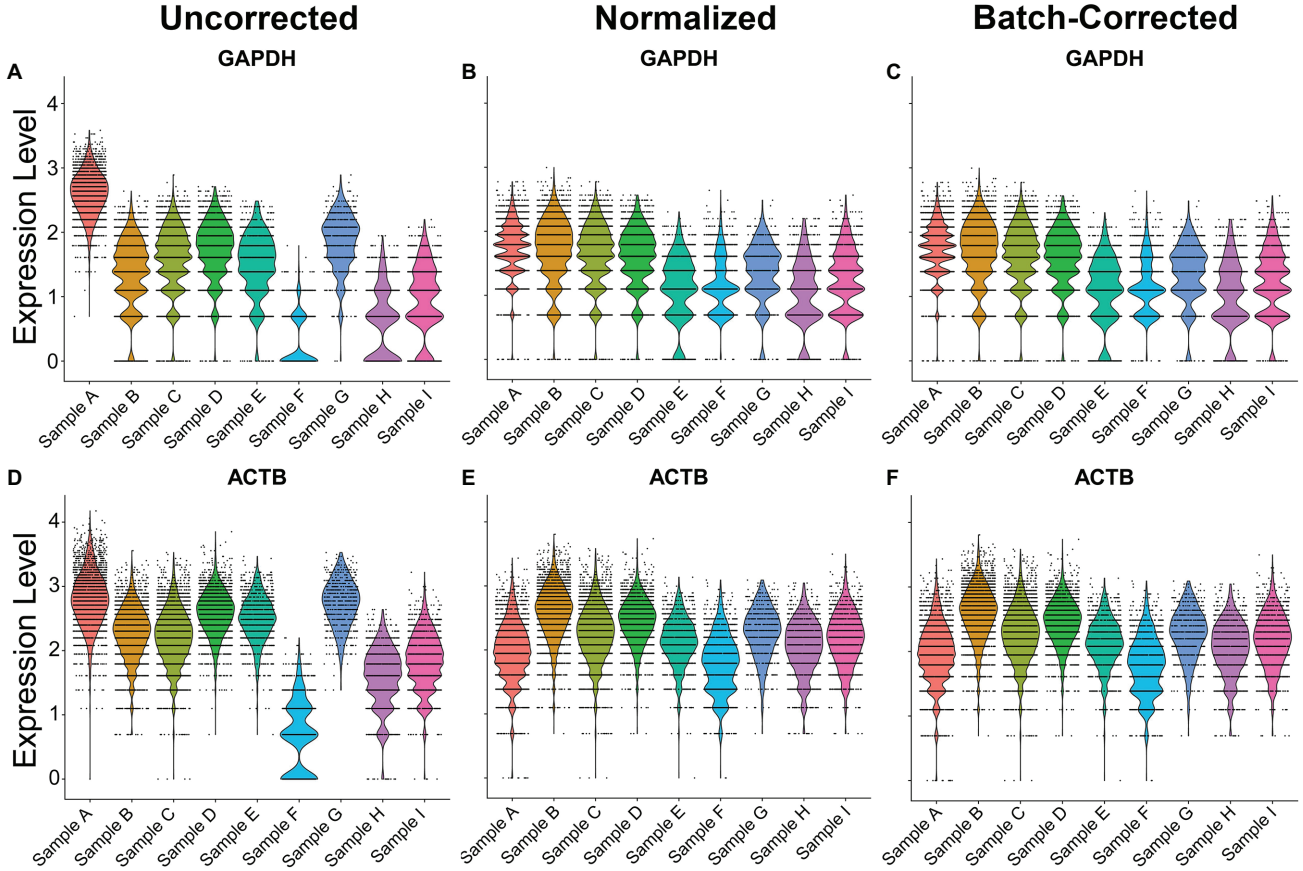


FIGURE 2 | Normalization and batch correction of spatial transcriptomic samples. **(A)** Uncorrected *GAPDH* expression across samples. **(B)** Normalized *GAPDH* expression with SCTransform. **(C)** Normalization and batch correction were performed by adding a batch as a variable in SCTransform. **(D)** Uncorrected expression of *ACTB*. **(E)** Normalized expression of *ACTB*. **(F)** Batch-corrected expression of *ACTB*. Sample A is batch 1, samples B–E are batch 2, and samples F–I are batch 3.

of a spot signature *should* parallel the histologic cell type distribution or whether certain cell types may have an outsized influence on the signature. Differences may arise from how each technique handles cell type heterogeneity or variation in expression. Further, performance can vary based on the fine-tuning of parameters; thus, this review is not intended to compare of each method's value. Instead, it merely provides an example of the diversity of possible results, depending on the approach selected.

AUTHOR CONTRIBUTIONS

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

FUNDING

This work was supported by NIH/NIDDK K08DK107864 (ME); Indiana Grand Challenge Precision Health fund (RM); and the Indiana Center for Biological Microscopy (NIH-NIDDK P30DK079312).

ACKNOWLEDGMENTS

The 10X 3' scRNA sequencing and 10X Visium Spatial 3' transcriptome sequencing were performed in the Center for Medical Genomics at Indiana University School of Medicine. We would like to thank the Indiana Center for Biological Microscopy for imaging assistance.

REFERENCES

Asp, M., Giacomello, S., Larsson, L., Wu, C., Fürth, D., Qian, X., et al. (2019). A spatiotemporal organ-wide gene expression and cell atlas of the developing human heart. *Cell* 179, 1647–1660. doi: 10.1016/j.cell.2019.11.025

Cable, D. M., Murray, E., Zou, L. S., Goeva, A., Macosko, E. Z., Chen, F., et al. (2021). Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat. Biotechnol.* doi: 10.1038/s41587-021-00830-w [Epub Ahead of print].

Elosua-Bayes, M., Nieto, P., Mereu, E., Gut, I., and Heyn, H. (2021). SPOTlight: seeded NMF regression to deconvolute spatial transcriptomics spots with single-cell transcriptomes. *Nucleic Acids Res.* 49:e50. doi: 10.1093/nar/gkab043

Hafemeister, C., and Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* 20, 1–15. doi: 10.1186/s13059-019-1874-1

Janosevic, D., Myslinski, J., McCarthy, T. W., Zollman, A., Syed, F., Xuei, X., et al. (2021). The orchestrated cellular and molecular responses of the kidney to endotoxin define a precise sepsis timeline. *Elife* 10:e62270. doi: 10.7554/eLife.62270

Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., et al. (2019). Fast, sensitive and accurate integration of single-cell data with harmony. *Nat. Methods* 16, 1289–1296. doi: 10.1038/s41592-019-0619-0

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.* 10:2832. 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

Marx, V. (2021). Method of the year: spatially resolved transcriptomics. *Nat. Methods* 18, 9–14. doi: 10.1038/s41592-020-01033-y

Melo Ferreira, R., Sabo, A. R., Winfree, S., Collins, K. S., Janosevic, D., Gulbronson, C. J., et al. (2021). Integration of spatial and single-cell transcriptomics localizes epithelial cell-immune cross-talk in kidney injury. *JCI Insight* 6:e147703. doi: 10.1172/jci.insight.147703

Navarro, J. F., Croteau, D. L., Jurek, A., Andrusivova, Z., Yang, B., Wang, Y., et al. (2020). Spatial Transcriptomics reveals genes associated with Dysregulated mitochondrial functions and stress signaling in Alzheimer disease. *Science* 23:101556. doi: 10.1016/j.isci.2020.101556

Stickels, R. R., Murray, E., Kumar, P., Li, J., Marshall, J. L., Di Bella, D. J., et al. (2021). Highly sensitive spatial transcriptomics at near-cellular resolution with slide-seqV2. *Nat. Biotechnol.* 39, 313–319. doi: 10.1038/s41587-020-0739-1

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M. III, et al. (2019). Comprehensive integration of single-cell data. *Cell* 177, 1888–19021. doi: 10.1016/j.cell.2019.05.031

Welch, J., Kozareva, V., Ferreira, A., Vanderburg, C., Martin, C., and Macosko, E. (2018). Integrative inference of brain cell similarities and differences from single-cell genomics. *bioRxiv* 49, 594–559. doi: 10.1101/459891

Wu, S. Z., Al-Eryani, G., Roden, D. L., Junankar, S., Harvey, K., Andersson, A., et al. (2021). A single-cell and spatially resolved atlas of human breast cancers. *Nat. Genet.* 53, 1334–1347. doi: 10.1038/s41588-021-00911-1

Zhang, Y., Parmigiani, G., and Johnson, W. E. (2020). ComBat-seq: batch effect adjustment for RNA-seq count data. *NAR Genom. Bioinfo.* 2:lqaa078. doi: 10.1093/nargab/lqaa078

Zhao, E., Stone, M. R., Ren, X., Guenthoer, J., Smythe, K. S., Pulliam, T., et al. (2021). Spatial transcriptomics at subspot resolution with BayesSpace. *Nat. Biotechnol.* 39, 1375–1384. doi: 10.1038/s41587-021-00935-2

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

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.

frontiers in Physiology

MINI REVIEW
published: 28 January 2022
doi: 10.3389/fphys.2022.832417

Spatial Statistics for Understanding Tissue Organization

*Andrea Behanova, Anna Klemm and Carolina Wählby**

Department of Information Technology and SciLifeLab, Centre for Image Analysis, Uppsala University, Uppsala, Sweden

Interpreting tissue architecture plays an important role in gaining a better understanding of healthy tissue development and disease. Novel molecular detection and imaging techniques make it possible to locate many different types of objects, such as cells and/or mRNAs, and map their location across the tissue space. In this review, we present several methods that provide quantification and statistical verification of observed patterns in the tissue architecture. We categorize these methods into three main groups: Spatial statistics on a single type of object, two types of objects, and multiple types of objects. We discuss the methods in relation to four hypotheses regarding the methods' capability to distinguish random and non-random distributions of objects across a tissue sample, and present a number of openly available tools where these methods are provided. We also discuss other spatial statistics methods compatible with other types of input data.

Keywords: transcriptomics, spatial statistics, gene expression, tissue analysis, tissue organization, niches

OPEN ACCESS

Edited by:
Zhihui Wang,
Houston Methodist Research Institute, United States

Reviewed by:
Christoph Kuppe,
RWTH Aachen University, Germany

****Correspondence:***
Carolina Wählby
carolina.wahlby@it.uu.se

Specialty section:
This article was submitted to Computational Physiology and Medicine, a section of the journal Frontiers in Physiology

Received: *09 December 2021*
Accepted: *06 January 2022*
Published: *28 January 2022*

Citation:
Behanova A, Klemm A and Wählby C (2022) Spatial Statistics for Understanding Tissue Organization. Front. Physiol. 13:832417. doi: 10.3389/fphys.2022.832417

1. INTRODUCTION

A range of new imaging-based methods make it possible to explore the architecture of tissue samples both at the transcriptomics and proteomics level. Multiplexed *in situ* RNA detection methods (Ke et al., 2013; Shah et al., 2016; Codeluppi et al., 2018; Moffitt et al., 2018; Wang et al., 2018; Eng et al., 2019) map mRNA molecules at sub-cellular resolution, and multiplex immunohistochemical staining (Parra et al., 2019), make it possible to detect and identify a large number of different cell types in the same tissue sample, enabling the discovery of their functional role inside the tissue architecture (Grün and van Oudenaarden, 2015; Svensson et al., 2018). The first step toward further interpretation of the data is detection and decoding, or classification, of each individual object; in this case resulting in maps of the locations of either specific mRNA molecules or cells.

One of the key challenges in fully exploiting this type of spatially resolved data is the availability of appropriate computational methods. The second step in interpretation is to be able to quantify relationships and patterns in an unbiased and reproducible way, and provide confidence measures for observed patterns as compared to a more randomized organization. This is often referred to as spatial statistics.

In this mini-review, we focus on spatial statistics applicable to tissue data independent of image resolution. We start with the assumption that each observed object has a unique position in 2D tissue space, and is assigned a specific type (e.g., cell type or mRNA species). Further, we assume that we also want to take the tissue context, and distribution of other objects, into consideration.

Objects can then be presented either as dots, a graph, a density map, or spatially binned counts in tissue space, as illustrated in **Figures 1A–D**. In the dot representation (**Figure 1A**), a different color would typically be used for each species. In a graph representation (**Figure 1B**), neighboring objects are connected. These connections can be restricted to fulfill criteria, such as a maximum number

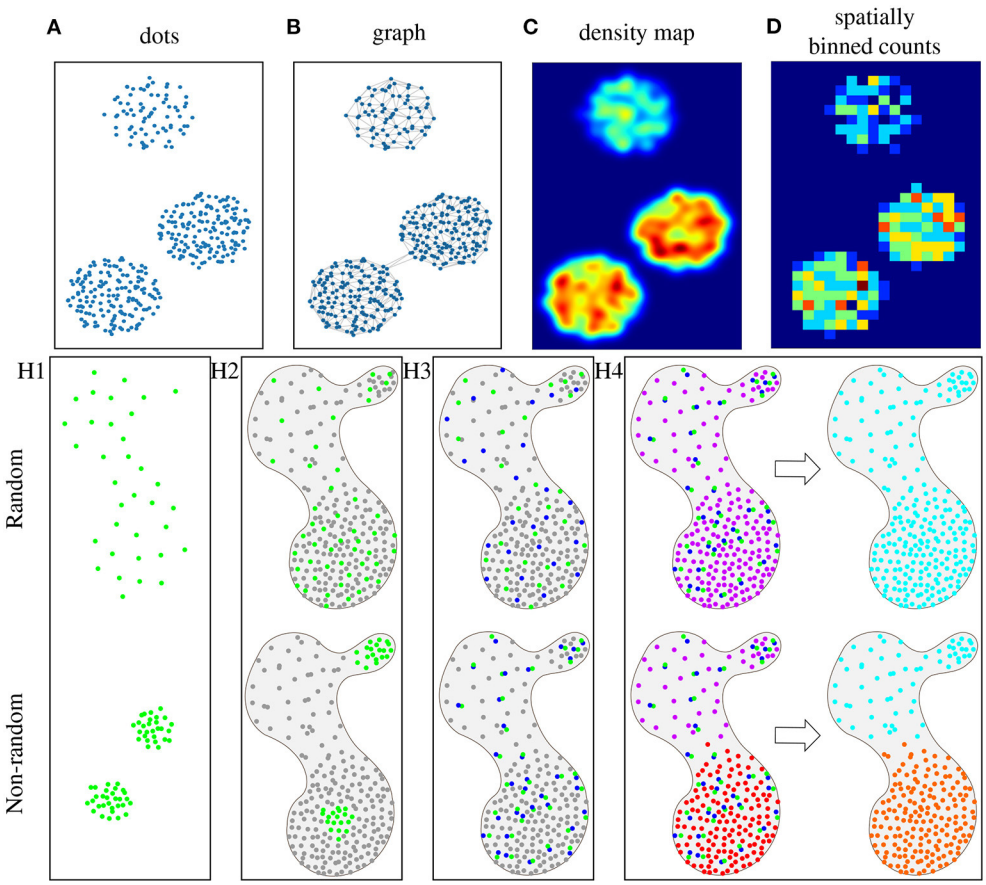


FIGURE 1 | Schematic representations of objects, such as cells or mRNAs, in microscopy images, where each dot represents an object, and the color reflects the object type (where gray is an unspecified type). **(A)** Simple representation, where each dot has a specific location in 2D tissue space. **(B)** The same data represented as a graph, where each dot is a node, and nodes are connected based on a maximum distance criterion. **(C)** Dots can also be represented by a probability density map, where warmer colors represent more dense dots, or **(D)** as counts in fixed spatial bins. Here, bins are squares and warmer colors represent higher object counts per bin. Spatial statistics are used to prove four different hypothesis (with the top row representing the random case): (H1) Visualization of hypothesis H1: Objects of type A (green) are non-randomly distributed. (H2) Visualization of hypothesis H2: Objects of type A (green) are non-randomly distributed as compared to the distribution of other objects (gray) in the same tissue sample. (H3) Visualization of hypothesis H3: Objects of type A (green) and B (blue) are non-randomly distributed in relation to one another within the distribution of other objects (gray) in the same tissue sample. (H4) Visualization of hypothesis H4: There are groups of object types (multiple colors in "niches") that are non-randomly distributed within the tissue sample.

of connections or distance, reflecting a hypothesis on a maximum distance for interaction. The density map representation (**Figure 1C**) translates the object distribution into a probability map, where high values represent high object concentrations, but the exact spatial location of objects is lost. Finally, different types of binning can be applied (**Figure 1D**), providing a lower-resolution map with counts of objects per bin.

We review methods that explore the null hypotheses of randomness for either a single type of objects, pairs of objects, or multiple types of objects. We have created a set of synthetic images describing different scenarios of object distributions within a tissue section, illustrating that the question of randomness is often relative. We first explore a single type of objects, as shown in **Figure 1**H1, and propose the hypothesis **H1**: Objects of type A are non-randomly distributed. In a biological context this could be, e.g., quantifying the distribution of immune cells in the presence or absence of an infection. If we take the tissue context (all objects of other types) into account, as shown in **Figure 1**H2, the hypothesis becomes **H2**: Objects of type A are non-randomly distributed as compared to the distribution of other objects in the same tissue sample. In a biological context this could be, e.g., distribution of a certain cell type in tumor and stroma areas of a tissue. Next, we consider two types of objects, and their potential interaction or repulsion. This is illustrated in **Figure 1**H3, and the hypothesis is **H3**: Objects of type A and B are non-randomly distributed in relation to one another within the distribution of other objects in the same tissue sample. In a biological context the question could, e.g., be whether cancer cells interact with endothelial cells or not. Finally, if there are multiple types of objects, we may want to see if certain groups of objects tend to coincide and form so-called 'niches' of unique combinations of objects in the tissue, as shown in **Figure 1**H4.

In this case, we pose hypothesis **H4**: There are groups of object types ('niches') that are non-randomly distributed within the distribution of other objects in the tissue sample. This could be used for finding mRNAs that are co-expressed, where niches would then correspond to different cell types (Partel and Wählby, 2020).

In the following review, we group different spatial statistics methods according to what types of tissue patterns they investigate, and also summarize and discuss their theoretical ability to answer the four hypotheses we pose above.

2. SPATIAL STATISTICS ON A SINGLE TYPE OF OBJECT

In this section, we describe methods which are capable to test hypothesize H1 (non-random distribution) and H2 (non-random distribution, compared to other objects). The input data can be described as points in space determining the presence of an object. The main idea is to identify and characterize spatially variable objects.

2.1. Ripley's Function

Ripley's function (Ripley, 1976) measures whether objects with discrete positions in space (see **Figure 1A**) follow random, dispersed, or clustered patterns. For each object, the function counts how many other objects of the same type appear within a given distance. Subsequently, the object counts are averaged over the whole dataset and the number is compared with the number of objects one would expect to find based on a completely spatially random pattern (null hypothesis). If the average number of objects found within the given distance is greater than for a random distribution, the dataset is clustered (see green dots in **Figure 1**H1-down). If the number is smaller, the dataset is dispersed. Ripley's K function is generally calculated at multiple distances allowing detection of pattern distributions at multiple scales. For example, at short distances, the objects may be clustered, while at long distances, objects may be dispersed. This method can be used to test hypothesis H1 (non-random distribution).

2.2. Newman's Assortativity

Newman's assortativity (Newman, 2002) evaluates spatial organization using a graph (see **Figure 1B**) as input. The principle is to count existing connections between objects of the same category and compare these counts to the number of connections expected at random object distribution (null hypothesis). This method can be used to test hypothesis H1. **Figure 1**H1-up shows no significant difference in the number of connections compared to a random distribution. However, **Figure 1**H1-down indicates that there would be a significant difference in the number of connections than under the null hypothesis. The difference between Ripley's function and Newman's assortativity is that Ripley's forms an overall cluster analysis providing various evaluations using various distances while Newman's tests the dataset as one object determining clustered patterns. However, the graph structure in Newman's assortativity provides more flexibility since graph connections can be created by different techniques, such as k-nearest neighbors or Delaunay triangulation.

2.3. Centrality Scores

Centrality scores (Everett and Borgatti, 1999) are based on computational analysis to show object patterns in a graph representation (see **Figure 1B**). This provides awareness of complicated relations in large graphs. **Figure 1**H2 can be used as an example where green dots represent one object type (group members, e.g., immune cells) and gray dots represent members of all the other object types (non-group members, e.g., all types of tumor cells). This method can be used to test hypothesis H2. There are four different centrality scores: **Group degree centrality** is interpreted as a ratio of non-group members (gray) that are connected to group members (green). Higher values reveal random distribution. Lower values indicate more grouped objects. This measure helps to identify crucial clusters in a graph. **Group closeness centrality** computes how close the group (green) is to the non-group members (gray). It is defined as the amount of non-group members (gray) divided by the sum of all distances from the group (green) to all non-group members (gray). Higher values reveal random distribution. Lower values indicate more grouped objects. **Group betweenness centrality** calculates the quantity of shortest paths connecting two non-group members (gray) while passing through the group (green). This can be thought of as a measure of cell infiltration. **Average clustering coefficient** measures how likely the group members favor to cluster together.

3. SPATIAL STATISTICS ON TWO TYPES OF OBJECTS

In this section, we describe methods capable of testing hypothesis H3: objects of types A and B are non-randomly distributed in relation to one another within the distribution of other objects in the same tissue sample. The main idea is to identify if different types of objects are closer than what would be expected by chance. It is worth noting that physical closeness is no guarantee for interaction, but a non-random pattern may indicate involvement in similar processes.

3.1. Cluster Co-occurrence Ratio

Cluster co-occurrence ratio (Tosti et al., 2021) describes co-occurrence of two types of objects in the tissue. It measures the probability that an object of type *A* appears in a given distance from an object of type *B* by taking the ratio between occurrences of object type *A* within a distance from object type *B* and occurrences of object type *A* within a distance from object type *B* at random (null hypothesis). It is computed across multiple distances across the tissue area. It measures the probability that an object of type *A* appears in a given distance from conditioned object type *B*. **Figure 1**H3-up shows example of low cluster co-occurrence ratio and **Figure 1**H3-down shows example of high cluster co-occurrence ratio within a short distance.

3.2. Neighborhood Enrichment Test

Neighborhood enrichment test (Schapiro et al., 2017) identifies two non-randomly distributed objects types in relation to one another. The first step is to create a graph (see **Figure 1B**). Then two object types are selected (A and B) and the count of connections between A and B object types (n_{AB}) is compared to random permutations of the objects (null hypothesis). The random configuration is set by keeping the object locations and reshuffling the object identities. Based on these estimates, expected means (μ_{AB}) and standard deviations (σ_{AB}) are calculated for each pair in the randomized dataset. Subsequently, a Z-score is calculated as, $Z_{AB} = \frac{n_{AB} - \mu_{AB}}{\sigma_{AB}}$. The Z-score indicates if an object type pair is over-represented (positive Z-score, see **Figure 1**H3-down) or over-depleted (negative Z-score, see **Figure 1**H3-up) in the connectivity graph. The difference between cluster co-occurrence ratio and neighborhood enrichment test is that cluster co-occurrence ratio evaluates various distances when determining if two objects types are in relation to one another while neighborhood enrichment test examines the dataset as one object determining object relation. However, the graph structure in Neighborhood enrichment test again provides flexibility since graph connections can be created by different techniques.

3.3. Object-Object Correlation Analysis

Object-Object Correlation Analysis (Stoltzfus et al., 2020) investigates the correlation of different object types within neighborhoods over the tissue. A neighborhood is a composition of objects inside a circular area. The neighborhoods' locations are uniformly allocated in a grid pattern throughout the space. The next step is to calculate the Pearson correlation coefficient of two types of objects within the neighborhoods. This method reveals which types of objects are associated with each other or unrelated to each other. **Figure 1D** shows an example of this neighborhood representation. The idea is to create this representation of two object types and then estimate the correlation coefficient across all overlapping neighborhoods.

4. SPATIAL STATISTICS ON MULTIPLE TYPES OF OBJECTS

In this section, we describe methods which are capable to test hypothesis H4 (existence of "niches"). The input data can be described as points in space determining the presence of the object types. The main idea is to identify if there are reoccurring spatial patterns, or 'niches' of objects, in the tissue.

4.1. Spatial Co-expression Patterns

Spatial co-expression patterns (Dries et al., 2021) identify robust patterns of object types that follow correlated spatial expression arrangements throughout the tissue. The first step is to smooth the object expression over the space by averaging in a grid or k-nearest neighbor technique. This results in a one density map for every object type as illustrated in **Figure 1C**. The next step is to calculate the Pearson correlation coefficient of the pair combinations of all object types (e.i., density maps). Subsequently, similarly co-expressed object types are clustered together into modules, and averaging them creates meta-object types to represent the similarly co-expressed object types.

4.2. Spage2vec

Spage2vec (Partel and Wählby, 2020) analyzes the spatial heterogeneity of complex patterns of objects. The input data is a graph (see **Figure 1B**), and it uses a graph representation learning technique based on a graph neural network (GNN). During training, the GNN learns the topological structure of each object's local neighborhood. It does not require labeled training data, but learns to find re-occurring patterns by comparing to a randomization of the data. After training, the observed patterns are summarized in a lower-dimensional embedding space that encapsulates high-dimensional information about each object's neighborhood. The last step is to cluster the multidimensional space using an unsupervised classification method (i.e., Leiden, Traag et al., 2019). Clusters represent combinations of object types that can be identified as specific domain types or 'niches'. **Figure 1**H4-down shows an example, where different neighborhood compositions were identified as different niches. The types of discovered niches can be further identified by correlation between the object composition of the niches and e.g., in the case of *in situ* sequencing data an external dataset of scRNA-seq signatures. The approach has also been applied to detect niches in multiplex fluorescence microscopy data of tissue micro arrays (Solorzano et al., 2021).

4.3. Spot-Based Spatial Cell-Type Analysis by Multidimensional mRNA Density Estimation (SSAM)

SSAM (Park et al., 2021) was defined to identify tissue niches in transcriptomics data. The first step is to create probability maps of the object types. Kernel Density Estimation (KDE) with a Gaussian kernel is applied to every object type resulting in a density map for each object type (see **Figure 1C**). Then all the images are put into a stack creating a multi-channel image where each pixel is a vector describing the local expression profile. Next, group type signatures are computed by clustering using Louvain (Blondel et al., 2008) or DBSCAN (Ester et al., 1996), and outliers (vectors far from their cluster medoid) are removed. The cluster centroids represent the group-type signatures. The third step is to generate a group-type map. Each pixel in the vector field is classified according to the maximum correlation with the group-type signatures. The group-type signatures can be taken from the previous step or an external dataset, such as scRNA-seq. The fourth step is to identify the tissue niches with definite group-type composition. The composition is computed in a circular sliding window over the tissue and clustered by agglomerative hierarchical clustering, merging highly correlating clusters. Finally, each cluster represents a unique tissue niche, an example can be seen in **Figure 1**H4-down where two different niche types were found.

4.4. Vector Approach

Describing local neighborhoods as vectors of counts of object types has been suggested in several publications under multiple names (Stoltzfus et al., 2020; He et al., 2021; Salas et al., 2021).

TABLE 1 | Overview of the methods' functionality.

Object	Method	H1	H2	H3	H4	Toolbox
Single	Ripley's function	yes	no	no	no	Squidpy, PySpacell, CytoMap
	Newman's assortativity	yes	no	no	no	PySpacell
	Centrality scores	no	yes	no	no	Squidpy
Two	Cluster co-occurrence ratio	no	no	yes	no	Squidpy
	Neighborhood enrichment test	no	no	yes	no	Giotto, Matisse, Squidpy, histoCAT
	Object-Object Correlation Analysis	no	no	yes	no	CytoMap
Multiple	Spatial co-expression patterns	no	no	yes	yes	Giotto
	Spage2vec	no	no	no	yes	Spage2vec
	SSAM	no	no	no	yes	SSAM
	Vector approach	no	no	no	yes	CytoMap, ClusterMap, Matisse

Here we refer to it as the vector approach. Its goal is to identify similar neighborhoods across the tissue sample. The first step is to define the neighborhoods. A neighborhood is a composition of object types inside a fixed area. The neighborhoods' locations can be uniformly allocated in a grid pattern throughout the space, constructed around each object from the dataset (Stoltzfus et al., 2020), based on Density peak clustering (He et al., 2021), or be defined by previously segmented tissue structures (Salas et al., 2021). Next, each neighborhood is presented as a vector containing counts of object types normalized, for example, by dividing each object count by the sum of all counts in the neighborhood (local normalization) or by dividing each object count by the sum of all the counts in the sample (global normalization). The normalized vectors are projected to a multidimensional space followed by clustering to identify niches. Examples of supervised clustering methods are common methods such as k-means and hierarchical clustering, or more advanced methods such as Self-Organizing Maps (Kohonen, 1982), Gaussian Distribution Model, or DBSCAN (Ester et al., 1996). Other clustering possibilities are unsupervised approaches such as Leiden (Traag et al., 2019) or Louvain (Blondel et al., 2008).

5. TOOLBOXES

Several toolboxes simplifying spatial statistics are available. Squidpy (Palla et al., 2021) includes four methods from this review: Ripley's function, Centrality scores, Cluster co-occurrence ratio, and the Neighborhood enrichment test. The toolbox PySpacell (Rose et al., 2019) includes methods such as Ripley's function and Newman's assortativity. CytoMap (Stoltzfus et al., 2020) includes Ripley's function, Object-Object correlation analysis and the Vector approach. Giotto focuses mostly on the data consisting of coordinates and quantitative information on multiple measurements per location, but also includes techniques as such as the Neighborhood enrichment test and Spatial co-expression patterns. The recently published Matisse (Salas et al., 2021) includes the Neighborhood enrichment test and the Vector approach. The toolbox histoCAT (Schapiro et al., 2017) includes the Neighborhood enrichment test, and Clustermap (He et al., 2021) includes the Vector approach. **Table 1** summarizes these toolboxes and lists the hypotheses that each of the methods is capable of testing.

6. DISCUSSION

There are many published methods for spatial statistics. However, they differ in the type of input data they can handle. In this review, we focused on methods where the input data can be described as points in 2D tissue space representing the presence of different object types. Another type of input data consists of coordinates and quantitative information on multiple measurements per location, as in e.g., spatial transcriptomics (Larsson et al., 2021). Spatial statistics for exploring this type of data can focus on a single type of objects, with methods such as Binary Spatial extracts (BinSpect, Dries et al., 2021), Getis-Ord General G (Getis and Ord, 2010), Spatial pattern recognition via kernels (SPARK, Sun et al., 2020), spatialDE (Svensson et al., 2018), Trendsceek (Edsgärd et al., 2018), Geary's c (Geary, 1954) or Moran's I (Moran, 1950). In the case of more than a single type of object, there are other methods, such as Spatially informed ligand-receptor pairing (Dries et al., 2021), Object-Object Correlation Analysis (Stoltzfus et al., 2020) and Spatial domain detection (Dries et al., 2021) that can be applied for exploring co-locations, potential interactions and niche discovery.

The methods mentioned above are also applicable on the type of data we present in this paper (input data as points in space determining the presence of the object types) but the data would have to be pre-processed by transferring dots into spatially binned counts for all object types, as exemplified for a single object type in **Figure 1D**. With such a representation, spatial resolution would be lost, but data could be analyzed by methods such as Trendsceek and SPARK.

Many of the methods for analyzing multiple object types include clustering as a final step of the analysis. Different clustering algorithms might lead to different results when applied to the same data, and should be carefully selected. It should also be noted, that proving or disproving a hypothesis regarding spatial statistics will depend on quality and amount of input data. One should also keep in mind that a 2D section may not always be a good representation of a true 3D structure such as an organ.

AUTHOR CONTRIBUTIONS

AB, AK, and CW: conceptualization and investigation. AB and CW: writing. All authors have read and agreed to the published version of the manuscript.

FUNDING

This research was funded by the European Research Council via ERC Consolidator grant CoG 682810 to CW and the SciLifeLab Bioimage Informatics Facility.

REFERENCES

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *J. Stat. Mech.* 2008, P10008. doi: 10.1088/1742-5468/2008/10/P10008

Codeluppi, S., Borm, L. E., Zeisel, A., Manno, G. L., van Lunteren, J. A., Svensson, C. I., et al. (2018). Spatial organization of the somatosensory cortex revealed by osmFISH. *Nat. Methods* 15, 932–935. doi: 10.1038/s41592-018-0175-z

Dries, R., Zhu, Q., Dong, R., Eng, C.-H. L., Li, H., Liu, K., et al. (2021). Giotto: a toolbox for integrative analysis and visualization of spatial expression data. *Genome Biol.* 22, 78. doi: 10.1186/s13059-021-02286-2

Edsgärd, D., Johnsson, P., and Sandberg, R. (2018). Identification of spatial expression trends in single-cell gene expression data. *Nat. Methods* 15, 339–342. doi: 10.1038/nmeth.4634

Eng, C.-H. L., Lawson, M., Zhu, Q., Dries, R., Koulena, N., Takei, Y., et al. (2019). Transcriptome-scale super-resolved imaging in tissues by rna seqfish+. *Nature* 568, 235–239. doi: 10.1038/s41586-019-1049-y

Ester, M., Kriegel, H. -P., Sander, J., and Xu, X. (1996). "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, KDD'96* (Munich: AAAI Press), 226–231.

Everett, M. G., and Borgatti, S. P. (1999). The centrality of groups and classes. *J. Math. Sociol.* 23, 181–201. doi: 10.1080/0022250X.1999.9990219

Geary, R. C. (1954). The contiguity ratio and statistical mapping. *Incorporat. Statist.* 5, 115. doi: 10.2307/2986645

Getis, A., and Ord, J. K. (2010). The analysis of spatial association by use of distance statistics. *Geogr. Anal.* 24, 189–206. doi: 10.1111/j.1538-4632.1992.tb00261.x

Grün, D., and van Oudenaarden, A. (2015). Design and analysis of single-cell sequencing experiments. *Cell* 163, 799–810. doi: 10.1016/j.cell.2015.10.039

He, Y., Tang, X., Huang, J., Zhou, H., Chen, K., Liu, A., et al. (2021). ClusterMap: multi-scale clustering analysis of spatial gene expression. *Nat. Commun.* 12:5909 doi: 10.1038/s41467-021-26044-x

Ke, R., Mignardi, M., Pacureanu, A., Svedlund, J., Botling, J., Wählby, C., et al. (2013). *In situ* sequencing for RNA analysis in preserved tissue and cells. *Nat. Methods* 10, 857–860. doi: 10.1038/nmeth.2563

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biol. Cybern.* 43, 59–69. doi: 10.1007/BF00337288

Larsson, L., Frisén, J., and Lundeberg, J. (2021). Spatially resolved transcriptomics adds a new dimension to genomics. *Nat. Methods* 18, 15–18. doi: 10.1038/s41592-020-01038-7

Moffitt, J. R., Bambah-Mukku, D., Eichhorn, S. W., Vaughn, E., Shekhar, K., Perez, J. D., et al. (2018). Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science* 362:eaau5324. doi: 10.1126/science.aau5324

Moran, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika* 37:17. doi: 10.2307/2332142

Newman, M. E. J. (2002). Assortative mixing in networks. *Phys. Rev. Lett.* 89:208701. doi: 10.1103/PhysRevLett.89.208701

Palla, G., Spitzer, H., Klein, M., Fischer, D., Schaar, A. C., Kuemmerle, L. B., et al. (2021). Squidpy: a scalable framework for spatial single cell analysis. *arxiv [Preprint]*. doi: 10.1101/2021.02.19.431994

Park, J., Choi, W., Tiesmeyer, S., Long, B., Borm, L. E., Garren, E., et al. (2021). Cell segmentation-free inference of cell types from *in situ* transcriptomics data. *Nat. Commun.* 12, 3545. doi: 10.1038/s41467-021-23807-4

Parra, E., Francisco-Cruz, A., and Wistuba, I. (2019). State-of-the-art of profiling immune contexture in the era of multiplexed staining and digital analysis to study paraffin tumor tissues. *Cancers* 11, 247. doi: 10.3390/cancers11020247

Partel, G., and Wählby, C. (2020). Spage2vec: Unsupervised representation of localized spatial gene expression signatures. *FEBS J.* 288, 1859–1870. doi: 10.1111/febs.15572

Ripley, B. D. (1976). The second-order analysis of stationary point processes. *J. Appl. Probab.* 13, 255–266. doi: 10.2307/3212829

Rose, F., Rappez, L., Triana, S. H., Alexandrov, T., and Genovesio, A. (2019). PySpacell: a python package for spatial analysis of cell images. *Cytometry A* 97, 288–295. doi: 10.1002/cyto.a.23955

Salas, S. M., Gyllborg, D., Langseth, C. M., and Nilsson, M. (2021). Matisse: a MATLAB-based analysis toolbox for *in situ* sequencing expression maps. *BMC Bioinformatics* 22:391. doi: 10.1186/s12859-021-04302-5

Schapiro, D., Jackson, H. W., Raghuraman, S., Fischer, J. R., Zanotelli, V. R. T., Schulz, D., et al. (2017). histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data. *Nat. Methods* 14, 873–876. doi: 10.1038/nmeth.4391

Shah, S., Lubeck, E., Zhou, W., and Cai, L. (2016). *In situ* transcription profiling of single cells reveals spatial organization of cells in the mouse hippocampus. *Neuron* 92, 342–357. doi: 10.1016/j.neuron.2016.10.001

Solorzano, L., Wik, L., Bontell, T. O., Wang, Y., Klemm, A. H., Öfverstedt, J., et al. (2021). Machine learning for cell classification and neighborhood analysis in glioma tissue. *Cytometry A*. 99, 1176–1186. doi: 10.1002/cyto.a.24467

Stoltzfus, C. R., Filipek, J., Gern, B. H., Olin, B. E., Leal, J. M., Wu, Y., et al. (2020). CytoMAP: A spatial analysis toolbox reveals features of myeloid cell organization in lymphoid tissues. *Cell Rep.* 31, 107523. doi: 10.1016/j.celrep.2020.107523

Sun, S., Zhu, J., and Zhou, X. (2020). Statistical analysis of spatial expression patterns for spatially resolved transcriptomic studies. *Nat. Methods* 17, 193–200. doi: 10.1038/s41592-019-0701-7

Svensson, V., Teichmann, S. A., and Stegle, O. (2018). SpatialDE: identification of spatially variable genes. *Nat. Methods* 15, 343–346. doi: 10.1038/nmeth.4636

Tosti, L., Hang, Y., Debnath, O., Tiesmeyer, S., Trefzer, T., Steiger, K., et al. (2021). Single-nucleus and *in situ* RNA-sequencing reveal cell topographies in the human pancreas. *Gastroenterology* 160, 1330.e11–1344.e11. doi: 10.1053/j.gastro.2020.11.010

Traag, V. A., Waltman, L., and van Eck, N. J. (2019). From louvain to leiden: guaranteeing well-connected communities. *Sci. Rep.* 9, 5233. doi: 10.1038/s41598-019-41695-z

Wang, X., Allen, W. E., Wright, M. A., Sylwestrak, E. L., Samusik, N., Vesuna, S., et al. (2018). Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science* 361:eaat5691. doi: 10.1126/science.aat5691

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

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.

frontiers in Physiology

MINI REVIEW
published: 11 February 2022
doi: 10.3389/fphys.2022.837773

Uncovering Molecular Heterogeneity in the Kidney With Spatially Targeted Mass Spectrometry

Angela R. S. Kruse[1,2] *and Jeffrey M. Spraggins*[1,2,3,4]*

[1] Department of Biochemistry, Vanderbilt University, Nashville, TN, United States, [2] Mass Spectrometry Research Center, Vanderbilt University, Nashville, TN, United States, [3] Department of Cell and Developmental Biology, Vanderbilt University, Nashville, TN, United States, [4] Department of Chemistry, Vanderbilt University, Nashville, TN, United States

The kidney functions through the coordination of approximately one million multifunctional nephrons in 3-dimensional space. Molecular understanding of the kidney has relied on transcriptomic, proteomic, and metabolomic analyses of kidney homogenate, but these approaches do not resolve cellular identity and spatial context. Mass spectrometry analysis of isolated cells retains cellular identity but not information regarding its cellular neighborhood and extracellular matrix. Spatially targeted mass spectrometry is uniquely suited to molecularly characterize kidney tissue while retaining *in situ* cellular context. This review summarizes advances in methodology and technology for spatially targeted mass spectrometry analysis of kidney tissue. Profiling technologies such as laser capture microdissection (LCM) coupled to liquid chromatography tandem mass spectrometry provide deep molecular coverage of specific tissue regions, while imaging technologies such as matrix assisted laser desorption/ionization imaging mass spectrometry (MALDI IMS) molecularly profile regularly spaced tissue regions with greater spatial resolution. These technologies individually have furthered our understanding of heterogeneity in nephron regions such as glomeruli and proximal tubules, and their combination is expected to profoundly expand our knowledge of the kidney in health and disease.

Keywords: mass spectrometry, kidney, proteomics, metabolomics, lipidomics, multimodal imaging, HuBMAP, KPMP

OPEN ACCESS

Edited by:
Bruce Molitoris, Indiana University, United States

Reviewed by:
Christoph Kuppe, RWTH Aachen University, Germany

***Correspondence:**
Jeffrey M. Spraggins
jeff.spraggins@vanderbilt.edu

Specialty section:
This article was submitted to Renal and Epithelial Physiology, a section of the journal Frontiers in Physiology

Received: *17 December 2021*
Accepted: *04 January 2022*
Published: *11 February 2022*

Citation:
Kruse ARS and Spraggins JM (2022) Uncovering Molecular Heterogeneity in the Kidney With Spatially Targeted Mass Spectrometry. Front. Physiol. 13:837773. doi: 10.3389/fphys.2022.837773

INTRODUCTION

The kidney is a complex and vital organ that filters waste products from the blood, stabilizes electrolyte and water content, and secretes essential hormones (Tryggvason and Wartiovaara, 2005; Ferraro and Fuster, 2021). It functions through nuanced coordination of approximately one million nephrons in 3-dimensional space. Nephrons can be further sub-divided into functional tissue units (FTUs) including vasculature, ducts, tubules, and glomeruli, each with unique molecular functions. FTUs are influenced by proximity to other structures and location within the organ. Individual glomeruli and tubules vary in vascular architecture, molecular environment and drug

Abbreviations: MS, mass spectrometry; FTU, functional tissue unit; PT, proximal tubule; Glom, glomerular; CD, collecting duct; microPOTS, microliter processing in one pot for trace samples; LCM, laser capture microdissection; microLESA, micro, liquid extraction surface analysis; MALDI, matrix assisted laser desorption/ionization; DESI, desorption electrospray ionization; IMS, imaging mass spectrometry.

distributions (Kang et al., 2005; Postnov et al., 2015; Kafarov et al., 2020). This heterogeneity is especially important in the context of renal disease, which can uniquely impact individual FTUs (Weening et al., 2004; Fogo, 2015; Aguayo-Mazzucato et al., 2017). Traditionally, our molecular understanding of renal disease comes from global transcriptomic, proteomic, and metabolomic analyses of kidney lysates. These bulk analyses offer deep and comprehensive molecular coverage and are invaluable for sample profiling and disease biomarker identification (Mayer et al., 2012). However, cell identity and spatial context are lost when tissues are homogenized, and molecular changes at the cellular or FTU level are diluted in bulk tissue analyses. This can result in the lack of detection of inter-individual and disease-associated variation, as well as inability to identify rare cell populations. Molecular characterization of dissociated cells provides cellular information lacking in bulk analyses but does not retain spatial context (Koehler et al., 2020). In addition, enzymatic dissociation of tissues can disrupt the cellular environment and preclude analysis of extracellular matrix molecules that can be relevant in fibrotic kidney disease (Autengruber et al., 2012). Recently, spatially targeted mass spectrometry (MS) technologies have emerged that provide a deeper understanding of the role localized cell types, cellular neighborhoods, and FTUs play in underlying pathomechanisms (Autengruber et al., 2012; Ryan et al., 2019). Each of these MS technologies has unique benefits and drawbacks for the study of human organs. This review highlights the application and potential of spatially targeted MS to illuminate the underlying molecular drivers of kidney health and disease.

SPATIAL MASS SPECTROMETRY TECHNOLOGIES

Spatially targeted MS technologies are characterized as either profiling experiments, where a single spectral signature is collected from a discrete cell type or FTU, or as imaging experiments where MS data are collected from an array of measurement locations (i.e., pixels) to visualize molecular distributions *in situ* (**Figure 1**). Micro-liquid extraction surface analysis (microLESA) is a profiling approach using a robotic fluidic printer to deposit trypsin droplets to specific tissue regions for surface protein digestion (Ryan et al., 2019; Guiberson et al., 2021). Peptides are then recovered using a larger droplet and subjected to liquid chromatography-tandem MS (LC-MS/MS) for protein identification. Laser capture microdissection (LCM) is also commonly employed in profiling experiments and involves dissection of specific sample regions using a cutting laser and subsequent collection into a sample tube using laser propulsion. Collected regions can be analyzed individually or pooled for protein, lipid, or small metabolite profiling (Datta et al., 2015; Knittelfelder et al., 2018; Sigdel et al., 2020). Although the achievable spatial resolution is limited, LCM can also be integrated into quasi-imaging workflows by dissecting tissue in a grid pattern in which each collected square becomes a voxel (Piehowski et al., 2020). Each region can be subjected to proteomic analysis using methods specialized for low sample input such as nanodroplet processing in one pot for trace samples (NanoPOTS), and voxels can be reconstructed to show intensity variation throughout the sample (Zhu et al., 2018; Piehowski et al., 2020).

Imaging mass spectrometry (IMS) is a powerful technology to construct spatial maps of analytes without labeling and in an untargeted manner (Caprioli et al., 1997; Gode and Volmer, 2013; Norris and Caprioli, 2013; Wu et al., 2013; Nilsson et al., 2015; Spengler, 2015). The most common IMS methods use soft ionization such as matrix-assisted laser desorption (MALDI) and desorption electrospray ionization (DESI) (Roach et al., 2010; Eberlin et al., 2011). In MALDI IMS workflows, tissue samples are coated with a matrix that assists with desorption and ionization of endogenous analytes (Franz and Michael, 2000). The tissue surface is then ablated using a laser in a raster pattern, where each laser spot produces a spectrum detecting hundreds to thousands of ions (Caprioli et al., 1997; Norris and Caprioli, 2013; Spraggins et al., 2019; Martín-Saiz et al., 2021). Spectral information from each laser spot (i.e., pixel) is reconstructed to show relative analyte intensity and distribution throughout the sample (Caprioli et al., 1997; Norris and Caprioli, 2013; Spraggins et al., 2019; Martín-Saiz et al., 2021). DESI and nano-DESI workflows use ambient liquid extraction of small tissue regions at regularly spaced measurement regions followed by introduction to a mass spectrometer inlet or primary capillary for electrospray ionization (Roach et al., 2010; Eberlin et al., 2011). MS data from sampled tissue coordinates can similarly be reconstructed into spatial maps in DESI and nano-DESI IMS workflows. Secondary-ion mass spectrometry (SIMS) has achieved the highest spatial resolution to date, but the high energy required for ionization limits the size of molecule that can be analyzed, making this technique more widely applied for analysis of elements and smaller biomolecules ($<$1,000 Da) rather than larger lipids, peptides, and proteins (Wu and Odom, 1996; Heeren et al., 2006; McDonnell et al., 2006).

Each of these strategies requires a trade-off between spatial resolution and sensitivity, where methods approaching cellular or subcellular resolution often detect fewer analytes. Sample preparation and ionization methods also impact the molecular class that can be analyzed. This continuum is especially notable in spatially targeted MS, and researchers must use their judgment to select the technology most suited to their experimental goals (**Figure 1**).

PROTEOMICS

Proteomics offers direct information about downstream effects of transcriptional and translational regulation on cellular function and does not require extrapolation from transcript data (Gry et al., 2009). MS-based proteomics provides an advantage over antibody-based techniques in that it is untargeted, highly multiplexed, and requires no *a priori* knowledge of antibody targets. It also retains information about truncated and post-translationally modified proteoforms that can be impacted by renal disease (Yassine et al., 2016). Spatial proteomics is uniquely advantageous since it can specifically assess protein regulation

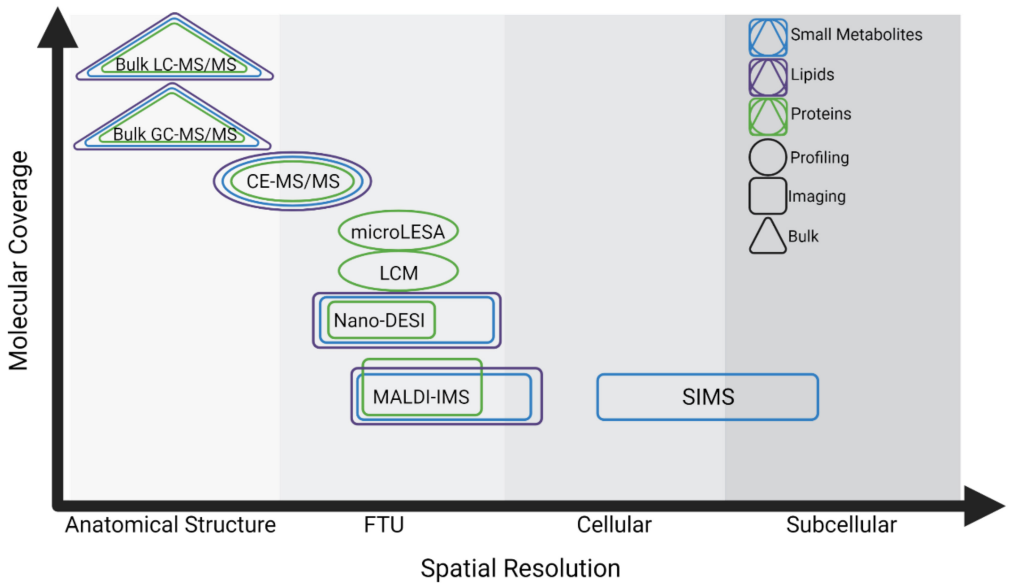


FIGURE 1 | Summary of spatial MS technologies and the trade-offs between spatial resolution and molecular coverage. Technologies are characterized based on their spatial resolution from the level of anatomical structure (>500 μm), Functional Tissue Unit (FTU, ∼50–500 μm), cellular (∼10–50 μm), and subcellular (<10 μm). Triangles indicate technologies for analysis of bulk or homogenate tissues, circles indicate profiling experiments, and rectangles indicate imaging experiments. Methods for analyzing small metabolites, lipids, and proteins are shown in blue, purple, and green, respectively.

in individual kidney FTUs and cell types, and has been used to show that adjacent nephrons vary at the proteomic level (Höhne et al., 2018). Both profiling and imaging MS approaches have been applied to the study of the kidney.

Spatial proteomics profiling experiments rely on the ability to analyze increasingly small amounts of starting material, requiring advances in sample preparation, chromatography, and instrumentation. Protocols employing filters, magnetic beads, or micro-volumes minimize sample loss by performing enzymatic digestion in one tube or droplet (Hughes et al., 2014; Kulak et al., 2014; Moggridge et al., 2018; Xu et al., 2019). NanoPOTS and MicroPOTS have facilitated near-single cell proteomics and are designed for low-input samples (Xu et al., 2019). Polished sample tubes and mass-spectrometry compatible detergents additionally minimize sample loss and the need for detergent removal (Norris et al., 2003, 2005; Grzeskowiak et al., 2016). Ultra-low flow chromatography and fractionation, and capillary electrophoresis can improve protein separation and address the wide range of protein concentrations found in biological samples (Waanders et al., 2008; Aebersold and Mann, 2016; Greguš et al., 2020; Kelly, 2020; Xiang et al., 2020). Pairing these sample preparation and separation techniques with high-resolution MS instrumentation can further facilitate low-input proteomics analysis (Norris et al., 2005). In addition, nanopore sequencing can now be used for single-cell proteomics and will likely be integrated into low-input proteomics workflows (Brinkerhoff et al., 2021).

These advances enabled multiple studies to characterize renal FTU proteomes. The combination of LCM with low-loss sample preparation and chromatography for LC-MS/MS proteomics has been especially successful for analysis of kidney FTUs. Thousands of proteins can be identified from single human or murine glomeruli or 30–40 single microdissected cells (Waanders et al., 2009; Sigdel et al., 2020). One study identified 67 proteins only detected in glomeruli and 25 unique to proximal tubules, with many additional proteins shared by both regions being conserved housekeeping and cytoskeletal proteins (Sigdel et al., 2020). Notably, this study found that proximal tubule proteins comprised a greater fraction of the homogenate proteome than glomerular proteins, and known glomerular markers such as podocin, eva-1 homolog B, and claudin-5 could be identified in dissected glomeruli but not in kidney homogenate (Guiberson et al., 2021). This underscores the value of a spatially targeted approach to study glomeruli (Sigdel et al., 2020). Another LCM-based study investigated proteomic changes associated with proteinuric kidney disease in glomeruli and tubules from murine and human samples. This work implicated a suite of proteins including lysosomal-associated membrane protein 1, cathepsin proteases, albumin, and extracellular matrix proteins in proteinuric kidney disease and proposed further research into cathepsins as potential therapeutic targets (Höhne et al., 2018). Analyses of single glomeruli and glomerular extracellular matrix consistently identify cathepsin proteases and proteins associated with vesicular transport and cellular component organization as differentially abundant in diseased kidney tissue (Hobeika et al., 2017; Höhne et al., 2018; Sigdel et al., 2020). Proteins enriched in proximal tubules were consistently involved in solute transport and small molecule metabolic processing, offering the intriguing possibility of measuring corresponding differences in metabolite abundance and localization (Hobeika et al., 2017; Höhne et al., 2018). Taken together, spatially targeted proteomics of kidney FTUs are invaluable to understanding renal FTU heterogeneity.

Imaging mass spectrometry is a powerful and complementary technology to spatially map proteins and peptides in tissue sections in an untargeted manner and with greater proteomic coverage than antibody-based imaging (Caprioli et al., 1997). Protein imaging requires minimal sample preparation and can be used to visualize proteins under ∼60 kDa depending on MS instrumentation (Chaurand et al., 2006; Norris and Caprioli, 2013). Although its spatial resolution is far superior to profiling-based technologies, a large proportion of the proteome is not available for analysis by this technique. In contrast, peptide imaging provides better proteomic coverage but requires more sample preparation and can suffer from delocalization during on-tissue enzymatic digestion of endogenous proteins (Judd et al., 2019). A major challenge for protein and peptide IMS is ion identification. Most protein IMS experiments rely on exact mass matching within a certain ppm error for identification. Recent computational tools allow for high throughput matching of *m/z* values with candidate identifications based upon intact mass and spatial correlation (Guo et al., 2021). Thus, advances in sample preparation, instrumentation, and computation are improving the feasibility and interpretation of protein and peptide IMS.

Imaging mass spectrometry has been applied to image kidney proteins and peptides, and has great potential as a tool for biomarker discovery and disease characterization (Caprioli et al., 1997; Lalowski et al., 2013; Jones et al., 2014; Casadonte et al., 2015; Srinivasu et al., 2021). Protein IMS was used to identify accumulated cortical transthyretin as a protein biomarker for gentamicin-induced kidney toxicity, and to spatially characterize angiotensin metabolism in murine kidneys (Meistermann et al., 2006; Grobe et al., 2012). Peptide IMS was used to determine that amyloid P component, apolipoprotein E, and vitronectin co-localize with renal amyloid deposits in human biopsy samples (Casadonte et al., 2015). Yet another study found differences in localization of α-enolase peptides in rat kidneys after treatment with nanoparticles commonly found in cosmetic and medical products (Srinivasu et al., 2021). Peptide IMS signal can be enhanced through secondary ionization (MALDI-2) and has been used to show localization of hemoglobin subunit proteins, glutathione-S-transferase, and pyruvate kinase to glomeruli, cortex, and medulla, respectively (McMillen et al., 2021). These IMS studies have benefited from the ability to visualize changes in analyte localization in broad tissue areas, and have leveraged microextraction or homogenate analyses with deeper proteomic coverage to confirm protein identifications (Grobe et al., 2012).

SMALL MOLECULE METABOLOMICS

Mass spectrometry-based metabolomics is essential in basic and clinical renal research (Abbiss et al., 2019). Here, metabolites are defined as small (<1,000 Da) molecules such as amino acids, nucleotides, mono- and disaccharides, and steroids that can be hydrophilic, hydrophobic, or amphipathic (Bijlsma et al., 2006). Liquid or gas chromatography-based metabolomics are routinely used to assess aminoaciduria in clinical samples or tissue homogenates (Rhee, 2018; Abbiss et al., 2019). Early work on the kidney profiled patient samples for disease biomarkers and resulted in the clinical tests now available to physicians (Cisek et al., 2016; Luft, 2021). However, general metabolic markers do not provide information about inter-nephron variation, and there is a gap in understanding sources of metabolic dysfunction on a spatial level and relating these to specific proteins. For example, amino acid transporters have been found to differ in proximal tubules within the same tissue section, implying that solute transport may be performed differently among nephrons and may be contributing uniquely to aminoaciduria and other kidney dysfunctions (Höhne et al., 2018).

Imaging mass spectrometry is uniquely powerful for kidney metabolomics because it is one of few methods that can spatially map metabolites within tissue, since these molecules are not amenable to antibody-based visualization (Prentice et al., 2017). Metabolite IMS has been used to characterize drug distribution in murine kidneys (Römpp et al., 2011), adenosine triphosphate and monophosphate in diabetic murine kidneys (Miyamoto et al., 2016), N-linked glycans in murine kidney (Gustafsson et al., 2015), and is extensively reviewed in Prentice et al. (2017). Metabolites can be routinely imaged with pixel sizes as small as 10 μm and their detection can be enhanced by gas-phase separation approaches such as trapped ion mobility separation (TIMS) (Djambazova et al., 2020; Neumann et al., 2020). Small metabolite IMS in human kidney samples was performed at a spatial resolution of 20 μm and allowed for the detection of >200 unique species using a timsTOF mass spectrometer in qTOF mode only (i.e., without TIMS activated) and >350 species after applying TIMS (Neumann et al., 2020). This study revealed unique distributions of metabolites including argininic acid, acetylcarnitine, and choline in the cortex, medulla, and renal pelvis, respectively (Neumann et al., 2020). Nano-DESI IMS was similarly used to show localization of propionylcarnitine, methylhistidine, sorbitol to the cortex, outer medulla, and inner medulla, respectively (Bergman et al., 2019). Acylcarnitine was shown to accumulate in the cortex of early-diabetic mice (Bergman et al., 2019). These approaches illustrate the excellent spatial resolution achievable by metabolite IMS and provide robust methods to visualize these molecules that cannot routinely be imaged using antibodies or affinity reagents. Future work could integrate metabolomic analyses of isolated FTUs with IMS to leverage the molecular coverage of the former with the spatial resolution of the latter.

LIPIDOMICS

Lipids play crucial and diverse roles in the kidney from establishment of cellular structure and stability to cell-cell interactions (Kinnunen et al., 2012; Balla, 2013). Lipids are metabolized in the kidney via receptor-mediated uptake of plasma lipids in proximal tubules (Moestrup and Nielsen, 2005). Chronic renal disease is associated with abnormal lipid metabolism, elevated apolipoprotein abundance, and elevated plasma lipid levels (Trevisan et al., 2006). Oxidative stress and insulin resistance have been implicated in lipid-mediated renal damage, but the underlying genetic, proteomic, and metabolomic mechanisms are not understood (Trevisan et al., 2006).

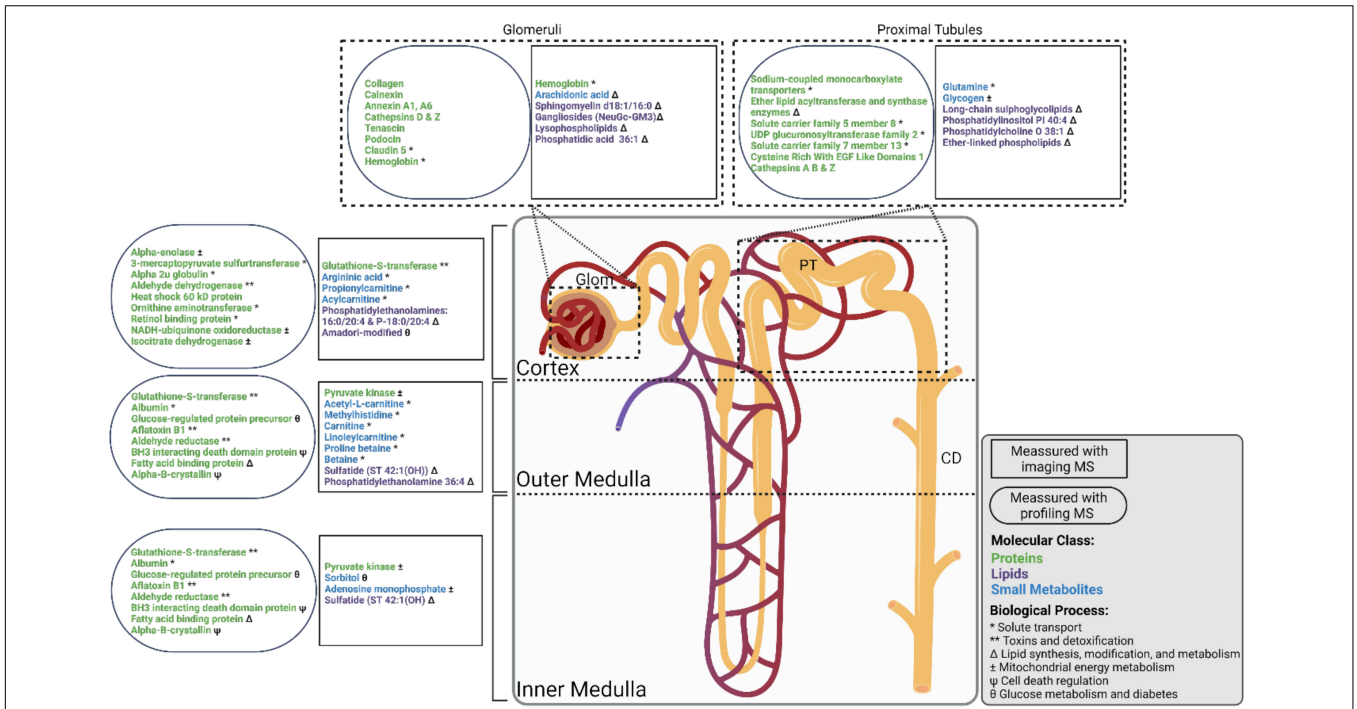


FIGURE 2 | Nephron showing select molecular groups observed in the cortex, outer medulla, inner medulla, glomeruli (Glom), proximal tubules (PT), and collecting ducts (CD). Molecular observations made using imaging mass spectrometry (MS) and profiling MS experiments are shown in rectangles and ovals, respectively. Proteins, small metabolites, and lipids, are shown in green, blue, and purple text, respectively. Biological processes are indicated as follows: * Solute transport, ** Toxins and detoxification, Δ Lipid synthesis, modification, and metabolism, ± Mitochondrial energy metabolism, Ψ Cell death regulation, θ Glucose metabolism and diabetes.

Additionally, altered renal lipid distribution has been associated with nephron dysfunction resulting from pathogen infection (Perry et al., 2019), polycystic kidney disease (Ruh et al., 2013), early diabetes and obesity (Miyamoto et al., 2016; Sugimoto et al., 2016; Bergman et al., 2019), and kidney injury (Rao et al., 2016). MS-based lipidomics globally characterizes how lipid class and the molecular structure influence these processes, and is uniquely informative in the context of renal disease.

Lipid IMS has been widely applied to the kidney and is further reviewed in Miyamoto et al. (2016). Gangliosides, sulfoglycosphingolipids, lysophospholipids, and phosphatidylethanolamines, sphingolipids, and lysolipids were shown to accumulate and spatially relocalize in the kidney due to diabetic nephropathy and severe ischemic injury in murine and porcine samples (Grove et al., 2014; Miyamoto et al., 2016; van Smaalen et al., 2019). Two ether-linked phospholipids were implicated as biomarkers for acute kidney injury in a murine model using sequential window acquisition of all theoretical spectra (SWATH) lipidomics and IMS (Rao et al., 2016). These phospholipids were shown to accumulate in proximal tubules, supporting the combination of profiling and imaging MS to characterize lipid abundance and localization (Rao et al., 2016). Another study showed that the ganglioside NeuGc-GM3, but not other ganglioside species, and several lysophospholipids accumulated in glomeruli of diabetic mice, while long-chain sulfoglycolipids accumulated in renal tubules of diabetic mice (Grove et al., 2014). Amadori-modified phosphatidylethanolamines were also detected in the renal cortex of diabetic mice, providing insight into the metabolic impacts of diabetes (Grove et al., 2014). Each of these approaches showed profound redistribution of lipid species in response to renal disease. To provide further insight into the lipidome and disease, technologies are linking specific lipid species with kidney FTUs based upon histology-informed segmentation of lipid IMS data (Martín-Saiz et al., 2021). These studies illustrate the utility of IMS to detect global lipidomic changes in disease and implicate diverse lipid classes in normal kidney function and pathogenesis.

CONCLUSION AND PERSPECTIVE

Technological advancement fundamentally changes the scale and strategy of scientific research. High-performance mass spectrometry and spatial technologies have moved us into an era of "big data" where the amount of molecular information collected from a single sample would have been previously inconceivable (Beckmann and Lew, 2016). To move beyond simply collecting big data to comprehensive interpretation of complex datasets, we assert that we are beginning an era focused on "multimodal data integration." Scientists will need to cooperatively link and automatically mine large datasets to understand intricate networks of cellular and molecular interactions across vast spatial scales (e.g., anatomical regions to single cells) and wide-ranging molecular classes (e.g., RNA,

proteins, lipids, and metabolites). Multi-institutional consortia are working to address this challenge by constructing molecular atlases of human cells and organs that integrate imaging and omics technologies using spatial anchors through common coordinate frameworks and/or anatomical links (Abbasi, 2017; Hu, 2019; Rozenblatt-Rosen et al., 2020; El-Achkar et al., 2021). These consortia are prioritizing the establishment of broadly accepted standards and quality control for data collection, precise recording of biopsies and tissue sections locations within intact organs, and recording thorough donor metadata data in accessible and stable repositories (Lynch, 2008; Hu, 2019; El-Achkar et al., 2021).

These large-scale kidney research projects are balancing the application of established multi-omic technologies with continued development of cutting-edge spatially targeted MS approaches. Spatially targeted proteomics utilizing profiling strategies has been more widely applied to the kidney than other technologies, and therefore individual glomerular and tubule proteomes are more well characterized (Betsholtz et al., 2007; Höhne et al., 2018; Hoyer et al., 2019; Späth et al., 2019; Koehler et al., 2020; Sigdel et al., 2020; Banki et al., 2021). Metabolomics has been applied to patient samples and kidney homogenates to great effect, and IMS has broadly characterized the spatial distribution of select small metabolites (Abbiss et al., 2019; Neumann et al., 2020; Zhang et al., 2020). Similarly, lipidomics has been used to characterize kidney homogenates, and IMS has generated spatial lipid maps (Rao et al., 2016; Lukowski et al., 2020; Zhang et al., 2020; Martín-Saiz et al., 2021). The next challenge will be to integrate these analytical modalities into workflows combining multiple spatially targeted MS technologies and to develop tools necessary to perform these analyses at scale across statistically relevant numbers of samples. The integration of these modalities in a systems-biology approach can provide us with a more comprehensive understanding of kidney biology (**Figure 2**; Mayer et al., 2012; Cisek et al., 2016; Rhee, 2018; Zhang et al., 2018; Neumann et al., 2021a).

High performance computing and development of necessary machine learning algorithms are playing an important role in technology integration. In addition to combining spatially targeted MS approaches, we anticipate that it will become more common for these data to be combined with other advanced molecular imaging technologies such as microscopy and spatial transcriptomics. Examples of this have already demonstrated integration of spatially targeted MS data with autofluorescence microscopy and multiplexed immunohistochemistry approaches such as imaging mass cytometry (IMC) and co-detection by indexing (CODEX) to molecularly characterize and discover markers for kidney FTUs and cell types (Patterson et al., 2018; Singh et al., 2019; Martín-Saiz et al., 2021; Neumann et al., 2021a,b). To enable these multimodal approaches, computational tools are emerging that automatically annotate, integrate, and mine molecular imaging data from orthogonal technologies for unbiased data interpretation and identification of candidate biomarkers (Van de Plas et al., 2015; Palmer et al., 2017; Vollnhals et al., 2017; Balluff et al., 2019; Race et al., 2020; Martín-Saiz et al., 2021; Tideman et al., 2021).

In summary, spatially targeted MS is a powerful set of technologies for the discovery of molecular profiles of critical FTUs and cell types in the kidney. As the field matures, multimodal data integration will certainly become more common requiring interdisciplinary, and often multi-institutional collaborations bringing together researchers with a wide array of expertise including cell biologists, pathologists, analytical chemists, computer scientists, and mathematical engineers. The application of this diverse set of expertise and technological capabilities is expected to dramatically enhance our understanding of the cellular and molecular makeup of the kidney to personalize medical care and improve health outcomes.

AUTHOR CONTRIBUTIONS

AK and JS completed the original manuscript draft and contributed to critical review and editing of the manuscript. Both authors contributed to the article and approved the submitted version.

FUNDING

The authors gratefully acknowledge support from the NIH Common Fund and National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) (U54DK120058 awarded to JS) and the NIH Common Fund and National Eye Institute (U54EY032442 awarded to JS). AK is supported by a National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) training grant (T32-DK-101003) and the NIH Common Fund (3OT2 OD026675-01S5 awarded to AK). Figures were made using BioRender under the publication license agreement numbers SB23BQJQE4 and GQ23BQJJT2.

REFERENCES

Abbasi, J. (2017). An international human cell atlas consortium takes shape. *JAMA* 318, 685–686. doi: 10.1001/jama.2017.5640

Abbiss, H., Maker, G. L., and Trengove, R. D. (2019). Metabolomics approaches for the diagnosis and understanding of kidney diseases. *Metabolites* 9:34. doi: 10.3390/metabo9020034

Aebersold, R., and Mann, M. (2016). Mass-spectrometric exploration of proteome structure and function. *Nature* 537, 347–355. doi: 10.1038/nature19949

Aguayo-Mazzucato, C., van Haaren, M., Mruk, M., Lee, T. B. Jr., Crawford, C., Hollister-Lock, J., et al. (2017). β Cell aging markers have heterogeneous distribution and are induced by insulin resistance. *Cell Metab.* 25, 898–910.e5. doi: 10.1016/j.cmet.2017.03.015

Autengruber, A., Gereke, M., Hansen, G., Hennig, C., and Bruder, D. (2012). Impact of enzymatic tissue disintegration on the level of surface molecule expression and immune cell function. *Eur. J. Microbiol. Immunol.* 2, 112–120. doi: 10.1556/EuJMI.2.2012.2.3

Balla, T. (2013). Phosphoinositides: tiny lipids with giant impact on cell regulation. *Physiol. Rev.* 93, 1019–1137. doi: 10.1152/physrev.00028.2012

Balluff, B., Buck, A., Martin-Lorenzo, M., Dewez, F., Langer, R., McDonnell, L. A., et al. (2019). Integrative clustering in mass spectrometry imaging for enhanced patient stratification. *Proteomics Clin. Appl.* 13:e1800137. doi: 10.1002/prca.201800137

Banki, E., Fisi, V., Moser, S., Wengi, A., Carrel, M., Loffing-Cueni, D., et al. (2021). Specific disruption of calcineurin-signaling in the distal convoluted tubule impacts the transcriptome and proteome, and causes hypomagnesemia and metabolic acidosis. *Kidney Int.* 100, 850–869. doi: 10.1016/j.kint.2021.06.030

Beckmann, J. S., and Lew, D. (2016). Reconciling evidence-based medicine and precision medicine in the era of big data: challenges and opportunities. *Genome Med.* 8:134. doi: 10.1186/s13073-016-0388-7

Bergman, H. M., Lindfors, L., Palm, F., Kihlberg, J., and Lanekoff, I. (2019). Metabolite aberrations in early diabetes detected in rat kidney using mass spectrometry imaging. *Anal. Bioanal. Chem.* 411, 2809–2816. doi: 10.1007/s00216-019-01721-5

Betsholtz, C., He, L., Takemoto, M., Norlin, J., Sun, Y., Patrakka, J., et al. (2007). The glomerular transcriptome and proteome. *Nephron Exp. Nephrol.* 106, e32–e36. doi: 10.1159/000101790

Bijlsma, S., Bobeldijk, I., Verheij, E. R., Ramaker, R., Kochhar, S., Macdonald, I. A., et al. (2006). Large-scale human metabolomics studies: a strategy for data (pre-) processing and validation. *Anal. Chem.* 78, 567–574. doi: 10.1021/ac051495j

Brinkerhoff, H., Kang, A. S. W., Liu, J., Aksimentiev, A., and Dekker, C. (2021). Multiple rereads of single proteins at single-amino acid resolution using nanopores. *Science* 374:eabl4381. doi: 10.1126/science.abl4381

Caprioli, R. M., Farmer, T. B., and Gile, J. (1997). Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal. Chem.* 69, 4751–4760. doi: 10.1021/ac970888i

Casadonte, R., Kriegsmann, M., Deininger, S. O., Amann, K., Paape, R., Belau, E., et al. (2015). Imaging mass spectrometry analysis of renal amyloidosis biopsies reveals protein co-localization with amyloid deposits. *Anal. Bioanal. Chem.* 407, 5323–5331. doi: 10.1007/s00216-015-8689-z

Chaurand, P., Norris, J. L., Cornett, D. S., Mobley, J. A., and Caprioli, R. M. (2006). New developments in profiling and imaging of proteins from tissue sections by MALDI mass spectrometry. *J. Proteome Res.* 5, 2889–2900. doi: 10.1021/pr060346u

Cisek, K., Krochmal, M., Klein, J., and Mischak, H. (2016). The application of multi-omics and systems biology to identify therapeutic targets in chronic kidney disease. *Nephrol. Dial Transplant* 31, 2003–2011. doi: 10.1093/ndt/gfv364

Datta, S., Malhotra, L., Dickerson, R., Chaffee, S., Sen, C. K., and Roy, S. (2015). Laser capture microdissection: big data from small samples. *Histol. Histopathol.* 30, 1255–1269. doi: 10.14670/HH-11-622

Djambazova, K. V., Klein, D. R., Migas, L. G., Neumann, E. K., Rivera, E. S., Van de Plas, R., et al. (2020). Resolving the complexity of spatial lipidomics using MALDI TIMS imaging mass spectrometry. *Anal. Chem.* 92, 13290–13297. doi: 10.1021/acs.analchem.0c02520

Eberlin, L. S., Ferreira, C. R., Dill, A. L., Ifa, D. R., and Cooks, R. G. (2011). Desorption electrospray ionization mass spectrometry for lipid characterization and biological tissue imaging. *Biochim. Biophys. Acta* 1811, 946–960. doi: 10.1016/j.bbalip.2011.05.006

El-Achkar, T. M., Eadon, M. T., Menon, R., Lake, B. B., Sigdel, T. K., Alexandrov, T., et al. (2021). A multimodal and integrated approach to interrogate human kidney biopsies with rigor and reproducibility: guidelines from the kidney precision medicine project. *Physiol. Genomics* 53, 1–11. doi: 10.1152/physiolgenomics.00104.2020

Ferraro, P. M., and Fuster, D. G. (2021). Clinical physiology of the kidney, electrolytes and lithiasis. the "old" meets the "new". *J. Nephrol.* 34, 29–30. doi: 10.1007/s40620-020-00850-w

Fogo, A. B. (2015). Causes and pathogenesis of focal segmental glomerulosclerosis. *Nat. Rev. Nephrol.* 11, 76–87. doi: 10.1038/nrneph.2014.216

Franz, H., and Michael, K. (2000). Matrix-assisted laser desorption/ionisation, an experience. *Int. J. Mass Spectrom.* 200, 71–77.

Gode, D., and Volmer, D. A. (2013). Lipid imaging by mass spectrometry - a review. *Analyst* 138, 1289–1315. doi: 10.1039/c2an36337b

Greguš, M., Kostas, J. C., Ray, S., Abbatiello, S. E., and Ivanov, A. R. (2020). Improved sensitivity of ultralow flow LC-MS-based proteomic profiling of limited samples using monolithic capillary columns and FAIMS technology. *Anal. Chem.* 92, 14702–14712. doi: 10.1021/acs.analchem.0c03262

Grobe, N., Elased, K. M., Cool, D. R., and Morris, M. (2012). Mass spectrometry for the molecular imaging of angiotensin metabolism in kidney. *Am. J. Physiol. Endocrinol. Metab.* 302, E1016–E1024. doi: 10.1152/ajpendo.00515.2011

Grove, K. J., Voziyan, P. A., Spraggins, J. M., Wang, S., Paueksakon, P., Harris, R. C., et al. (2014). Diabetic nephropathy induces alterations in the glomerular and tubule lipid profiles. *J. Lipid Res.* 55, 1375–1385. doi: 10.1194/jlr.M049189

Gry, M., Rimini, R., Strömberg, S., Asplund, A., Pontén, F., Uhlén, M., et al. (2009). Correlations between RNA and protein expression profiles in 23 human cell lines. *BMC Genomics* 10:365. doi: 10.1186/1471-2164-10-365

Grzeskowiak, R., Hamels, S., and Gancarek, E. (2016). *Comparative Analysis of Protein Recovery Rates in Eppendorf LoBind® and Other "Low Binding" Tubes. AG Application Note.2016: No, 382*. Eppendorf: Hamburg.

Guiberson, E. R., Weiss, A., Ryan, D. J., Monteith, A. J., Sharman, K., Gutierrez, D. B., et al. (2021). Spatially targeted proteomics of the host-pathogen interface during staphylococcal abscess formation. *ACS Infect. Dis.* 7, 101–113. doi: 10.1021/acsinfecdis.0c00647

Guo, G., Papanicolaou, M., Demarais, N. J., Wang, Z., Schey, K. L., Timpson, P., et al. (2021). Automated annotation and visualisation of high-resolution spatial proteomic mass spectrometry imaging data using HIT-MAP. *Nat. Commun.* 12:3241. doi: 10.1038/s41467-021-23461-w

Gustafsson, O. J., Briggs, M. T., Condina, M. R., Winderbaum, L. J., Pelzing, M., McColl, S. R., et al. (2015). MALDI imaging mass spectrometry of N-linked glycans on formalin-fixed paraffin-embedded murine kidney. *Anal. Bioanal. Chem.* 407, 2127–2139. doi: 10.1007/s00216-014-8293-7

Heeren, R. M. A., McDonnell, L. A., Amstalden, E., Luxembourg, S. L., Altelaar, A. F. M., and Piersma, S. R. (2006). Why don't biologists use SIMS?: a critical evaluation of imaging MS. *Appl. Surf. Sci.* 252, 6827–6835. doi: 10.1016/j.apsusc.2006.02.134

Hobeika, L., Barati, M. T., Caster, D. J., McLeish, K. R., and Merchant, M. L. (2017). Characterization of glomerular extracellular matrix by proteomic analysis of laser-captured microdissected glomeruli. *Kidney Int.* 91, 501–511. doi: 10.1016/j.kint.2016.09.044

Höhne, M., Frese, C. K., Grahammer, F., Dafinger, C., Ciarimboli, G., Butt, L., et al. (2018). Single-nephron proteomes connect morphology and function in proteinuric kidney disease. *Kidney Int.* 93, 1308–1319. doi: 10.1016/j.kint.2017.12.012

Hoyer, K. J. R., Dittrich, S., Bartram, M. P., and Rinschen, M. M. (2019). Quantification of molecular heterogeneity in kidney tissue by targeted proteomics. *J. Proteomics* 193, 85–92. doi: 10.1016/j.jprot.2018.03.001

Hu, B. C. (2019). The human body at cellular resolution: the NIH Human biomolecular atlas program. *Nature* 574, 187–192. doi: 10.1038/s41586-019-1629-x

Hughes, C. S., Foehr, S., Garfield, D. A., Furlong, E. E., Steinmetz, L. M., and Krijgsveld, J. (2014). Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol. Syst. Biol.* 10:757. doi: 10.15252/msb.20145625

Jones, E. E., Powers, T. W., Neely, B. A., Cazares, L. H., Troyer, D. A., Parker, A. S., et al. (2014). MALDI imaging mass spectrometry profiling of proteins and lipids in clear cell renal cell carcinoma. *Proteomics* 14, 924–935. doi: 10.1002/pmic.201300434

Judd, A. M., Gutierrez, D. B., Moore, J. L., Patterson, N. H., Yang, J., Romer, C. E., et al. (2019). A recommended and verified procedure for in situ tryptic digestion of formalin-fixed paraffin-embedded tissues for analysis by matrix-assisted laser desorption/ionization imaging mass spectrometry. *J. Mass Spectrom.* 54, 716–727. doi: 10.1002/jms.4384

Kafarov, E., Zenin, O., Fyodorov, S., Bataev Kh, M., and Vezirkhanov, A. (2020). "Variant anatomy of renal vein and its intra-organ branches," in *Proceedings of International Conference "Health and Wellbeing in Modern Society"(ICHW 2020)*, (Dordrecht: Atlantis Press).

Kang, J., Dai, X. S., Yu, T. B., Wen, B., and Yang, Z. W. (2005). Glycogen accumulation in renal tubules, a key morphological change in the diabetic rat kidney. *Acta Diabetol.* 42, 110–116. doi: 10.1007/s00592-005-0188-9

Kelly, R. T. (2020). Single-cell proteomics: progress and prospects. *Mol. Cell Proteomics* 19, 1739–1748. doi: 10.1074/mcp.R120.002234

Kinnunen, P. K., Kaarniranta, K., and Mahalka, A. K. (2012). Protein-oxidized phospholipid interactions in cellular signaling for cell death: from biophysics to clinical correlations. *Biochim. Biophys. Acta* 1818, 2446–2455. doi: 10.1016/j.bbamem.2012.04.008

Knittelfelder, O., Traikov, S., Vvedenskaya, O., Schuhmann, A., Segeletz, S., Shevchenko, A., et al. (2018). Shotgun lipidomics combined with laser capture microdissection: a tool to analyze histological zones in cryosections of tissues. *Anal. Chem.* 90, 9868–9878. doi: 10.1021/acs.analchem.8b02004

Koehler, S., Kuczkowski, A., Kuehne, L., Jüngst, C., Hoehne, M., Grahammer, F., et al. (2020). Proteome analysis of isolated podocytes reveals stress responses in glomerular sclerosis. *J. Am. Soc. Nephrol.* 31, 544–559. doi: 10.1681/ASN.2019030312

Kulak, N. A., Pichler, G., Paron, I., Nagaraj, N., and Mann, M. (2014). Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat. Methods* 11, 319–324. doi: 10.1038/nmeth.2834

Lalowski, M., Magni, F., Mainini, V., Monogioudi, E., Gotsopoulos, A., Soliymani, R., et al. (2013). Imaging mass spectrometry: a new tool for kidney disease investigations. *Nephrol. Dial Transplant* 28, 1648–1656. doi: 10.1093/ndt/gft008

Luft, F. C. (2021). Biomarkers and predicting acute kidney injury. *Acta Physiol.* 231:e13479.

Lukowski, J. K., Pamreddy, A., Velickovic, D., Zhang, G., Pasa-Tolic, L., Alexandrov, T., et al. (2020). Storage conditions of human kidney tissue sections affect spatial lipidomics analysis reproducibility. *J. Am. Soc. Mass Spectrom.* 31, 2538–2546. doi: 10.1021/jasms.0c00256

Lynch, C. (2008). Big data: how do your data grow? *Nature* 455, 28–29. doi: 10.1038/455028a

Martín-Saiz, L., Mosteiro, L., Solano-Iturri, J. D., Rueda, Y., Martín-Allende, J., Imaz, I., et al. (2021). High-resolution human kidney molecular histology by imaging mass spectrometry of lipids. *Anal. Chem.* 93, 9364–9372. doi: 10.1021/acs.analchem.1c00649

Mayer, P., Mayer, B., and Mayer, G. (2012). Systems biology: building a useful model from multiple markers and profiles. *Nephrol. Dial Transplant* 27, 3995–4002. doi: 10.1093/ndt/gfs489

McDonnell, L. A., Heeren, R. M., de Lange, R. P., and Fletcher, I. W. (2006). Higher sensitivity secondary ion mass spectrometry of biological molecules for high resolution, chemically specific imaging. *J. Am. Soc. Mass Spectrom.* 17, 1195–1202. doi: 10.1016/j.jasms.2006.05.003

McMillen, J. C., Gutierrez, D. B., Judd, A. M., Spraggins, J. M., and Caprioli, R. M. (2021). Enhancement of tryptic peptide signals from tissue sections using MALDI IMS postionization (MALDI-2). *J. Am. Soc. Mass Spectrom.* 32, 2583–2591. doi: 10.1021/jasms.1c00213

Meistermann, H., Norris, J. L., Aerni, H. R., Cornett, D. S., Friedlein, A., Erskine, A. R., et al. (2006). Biomarker discovery by imaging mass spectrometry: transthyretin is a biomarker for gentamicin-induced nephrotoxicity in rat. *Mol. Cell Proteomics* 5, 1876–1886. doi: 10.1074/mcp.M500399-MCP200

Miyamoto, S., Hsu, C. C., Hamm, G., Darshi, M., Diamond-Stanic, M., Declèves, A. E., et al. (2016). Mass spectrometry imaging reveals elevated glomerular ATP/AMP in diabetes/obesity and identifies sphingomyelin as a possible mediator. *EBioMedicine* 7, 121–134. doi: 10.1016/j.ebiom.2016.03.033

Moestrup, S. K., and Nielsen, L. B. (2005). The role of the kidney in lipid metabolism. *Curr. Opin. Lipidol.* 16, 301–306. doi: 10.1097/01.mol.0000169350.45944.d4

Moggridge, S., Sorensen, P. H., Morin, G. B., and Hughes, C. S. (2018). Extending the compatibility of the SP3 paramagnetic bead processing approach for proteomics. *J. Proteome Res.* 17, 1730–1740. doi: 10.1021/acs.jproteome.7b00913

Neumann, E. K., Migas, L. G., Allen, J. L., Caprioli, R. M., Van de Plas, R., and Spraggins, J. M. (2020). Spatial metabolomics of the human kidney using MALDI trapped ion mobility imaging mass spectrometry. *Anal. Chem.* 92, 13084–13091. doi: 10.1021/acs.analchem.0c02051

Neumann, E. K., Patterson, N. H., Allen, J. L., Migas, L. G., Yang, H., Brewer, M., et al. (2021a). Protocol for multimodal analysis of human kidney tissue by imaging mass spectrometry and CODEX multiplexed immunofluorescence. *STAR Protoc.* 2:100747. doi: 10.1016/j.xpro.2021.100747

Neumann, E. K., Patterson, N. H., Rivera, E. S., Allen, J. L., Brewer, M., deCaestecker, M. P., et al. (2021b). Highly multiplexed immunofluorescence of the human kidney using co-detection by indexing. *Kidney Int.* 101, 137–143. doi: 10.1016/j.kint.2021.08.033

Nilsson, A., Goodwin, R. J., Shariatgorji, M., Vallianatou, T., Webborn, P. J., and Andrén, P. E. (2015). Mass spectrometry imaging in drug development. *Anal. Chem.* 87, 1437–1455.

Norris, J. L., and Caprioli, R. M. (2013). Analysis of tissue specimens by matrix-assisted laser desorption/ionization imaging mass spectrometry in biological and clinical research. *Chem. Rev.* 113, 2309–2342. doi: 10.1021/cr3004295

Norris, J. L., Hangauer, M. J., Porter, N. A., and Caprioli, R. M. (2005). Nonacid cleavable detergents applied to MALDI mass spectrometry profiling of whole cells. *J. Mass Spectrom.* 40, 1319–1326. doi: 10.1002/jms.914

Norris, J. L., Porter, N. A., and Caprioli, R. M. (2003). Mass spectrometry of intracellular and membrane proteins using cleavable detergents. *Anal. Chem.* 75, 6642–6647. doi: 10.1021/ac034802z

Palmer, A., Phapale, P., Chernyavsky, I., Lavigne, R., Fay, D., Tarasov, A., et al. (2017). FDR-controlled metabolite annotation for high-resolution imaging mass spectrometry. *Nat. Methods* 14, 57–60. doi: 10.1038/nmeth.4072

Patterson, N. H., Tuck, M., Lewis, A., Kaushansky, A., Norris, J. L., Van de Plas, R., et al. (2018). Next generation histology-directed imaging mass spectrometry driven by autofluorescence microscopy. *Analyt. Chem.* 90, 12404–12413. doi: 10.1021/acs.analchem.8b02885

Perry, W. J., Spraggins, J. M., Sheldon, J. R., Grunenwald, C. M., Heinrichs, D. E., Cassat, J. E., et al. (2019). *Staphylococcus aureus* exhibits heterogeneous siderophore production within the vertebrate host. *Proc. Natl. Acad. Sci. U.S.A.* 116, 21980–21982. doi: 10.1073/pnas.1913991116

Piehowski, P. D., Zhu, Y., Bramer, L. M., Stratton, K. G., Zhao, R., Orton, D. J., et al. (2020). Automated mass spectrometry imaging of over 2000 proteins from tissue sections at 100-μm spatial resolution. *Nat. Commun.* 11:8. doi: 10.1038/s41467-019-13858-z

Postnov, D. D., Holstein-Rathlou, N. H., and Sosnovtseva, O. (2015). Laser speckle imaging of intra organ drug distribution. *Biomed. Opt. Express* 6, 5055–5062. doi: 10.1364/BOE.6.005055

Prentice, B. M., Caprioli, R. M., and Vuiblet, V. (2017). Label-free molecular imaging of the kidney. *Kidney Int.* 92, 580–598. doi: 10.1016/j.kint.2017.03.052

Race, A. M., Rae, A., Vorng, J. L., Havelund, R., Dexter, A., Kumar, N., et al. (2020). Correlative hyperspectral imaging using a dimensionality-reduction-based image fusion method. *Anal. Chem.* 92, 10979–10988. doi: 10.1021/acs.analchem.9b05055

Rao, S., Walters, K. B., Wilson, L., Chen, B., Bolisetty, S., Graves, D., et al. (2016). Early lipid changes in acute kidney injury using SWATH lipidomics coupled with MALDI tissue imaging. *Am. J. Physiol. Renal. Physiol.* 310, F1136–F1147. doi: 10.1152/ajprenal.00100.2016

Rhee, E. P. (2018). A systems-level view of renal metabolomics. *Semin. Nephrol.* 38, 142–150. doi: 10.1016/j.semnephrol.2018.01.005

Roach, P. J., Laskin, J., and Laskin, A. (2010). Nanospray desorption electrospray ionization: an ambient method for liquid-extraction surface sampling in mass spectrometry. *Analyst* 135, 2233–2236. doi: 10.1039/c0an00312c

Römpp, A., Guenther, S., Takats, Z., and Spengler, B. (2011). Mass spectrometry imaging with high resolution in mass and space (HR(2) MSI) for reliable investigation of drug compound distributions on the cellular level. *Anal. Bioanal. Chem.* 401, 65–73. doi: 10.1007/s00216-011-4990-7

Rozenblatt-Rosen, O., Regev, A., Oberdoerffer, P., Nawy, T., Hupalowska, A., Rood, J. E., et al. (2020). The human tumor atlas network: charting tumor transitions across space and time at single-cell resolution. *Cell* 181, 236–249. doi: 10.1016/j.cell.2020.03.053

Ruh, H., Salonikios, T., Fuchser, J., Schwartz, M., Sticht, C., Hochheim, C., et al. (2013). MALDI imaging MS reveals candidate lipid markers of polycystic kidney disease. *J. Lipid Res.* 54, 2785–2794. doi: 10.1194/jlr.M040014

Ryan, D. J., Patterson, N. H., Putnam, N. E., Wilde, A. D., Weiss, A., Perry, W. J., et al. (2019). MicroLESA: integrating autofluorescence microscopy, in situ micro-digestions, and liquid extraction surface analysis for high spatial resolution targeted proteomic studies. *Anal. Chem.* 91, 7578–7585. doi: 10.1021/acs.analchem.8b05889

Sigdel, T. K., Piehowski, P. D., Roy, S., Liberto, J., Hansen, J. R., Swensen, A. C., et al. (2020). Near-single-cell proteomics profiling of the proximal tubular and glomerulus of the normal human kidney. *Front. Med.* 7:499. doi: 10.3389/fmed.2020.00499

Singh, N., Avigan, Z. M., Kliegel, J. A., Shuch, B. M., Montgomery, R. R., Moeckel, G. W., et al. (2019). Development of a 2-dimensional atlas of the human kidney with imaging mass cytometry. *JCI Insight* 4:e129477. doi: 10.1172/jci.insight.129477

Späth, M. R., Bartram, M. P., Palacio-Escat, N., Hoyer, K. J. R., Debes, C., Demir, F., et al. (2019). The proteome microenvironment determines the protective effect of preconditioning in cisplatin-induced acute kidney injury. *Kidney Int.* 95, 333–349. doi: 10.1016/j.kint.2018.08.037

Spengler, B. (2015). Mass spectrometry imaging of biomolecular information. *Anal. Chem.* 87, 64–82. doi: 10.1021/ac504543v

Spraggins, J. M., Djambazova, K. V., Rivera, E. S., Migas, L. G., Neumann, E. K., Fuetterer, A., et al. (2019). High-performance molecular imaging with MALDI Trapped Ion-Mobility Time-of-Flight (timsTOF) mass spectrometry. *Analyt. Chem.* 91, 14552–14560. doi: 10.1021/acs.analchem.9b03612

Srinivasu, B. Y., Arumugaperumal, A., Mitra, A., Muralidharan, M., Das, R., and Mandal, A. K. (2021). Effect of nanoparticle exposure in a living system: probed by quantification of Fetuin-B in plasma proteome and kidney tissue imaging using MALDI imaging mass spectrometry in a rat model. *J. Nanopart. Res.* 23, 125.

Sugimoto, M., Wakabayashi, M., Shimizu, Y., Yoshioka, T., Higashino, K., Numata, Y., et al. (2016). Imaging mass spectrometry reveals Acyl-chain- and region-specific sphingolipid metabolism in the kidneys of sphingomyelin synthase 2-deficient mice. *PLoS One* 11:e0152191. doi: 10.1371/journal.pone.0152191

Tideman, L. E. M., Migas, L. G., Djambazova, K. V., Patterson, N. H., Caprioli, R. M., Spraggins, J. M., et al. (2021). Automated biomarker candidate discovery in imaging mass spectrometry data through spatially localized Shapley additive explanations. *Anal. Chim. Acta* 1177:338522. doi: 10.1016/j.aca.2021.338522

Trevisan, R., Dodesini, A. R., and Lepore, G. (2006). Lipids and renal disease. *J. Am. Soc. Nephrol.* 17(4 Suppl. 2), S145–S147.

Tryggvason, K., and Wartiovaara, J. (2005). How does the kidney filter plasma? *Physiology* 20, 96–101. doi: 10.1152/physiol.00045.2004

Van de Plas, R., Yang, J., Spraggins, J., and Caprioli, R. M. (2015). Image fusion of mass spectrometry and microscopy: a multimodality paradigm for molecular tissue mapping. *Nat. Methods* 12, 366–372. doi: 10.1038/nmeth.3296

van Smaalen, T. C., Ellis, S. R., Mascini, N. E., Siegel, T. P., Cillero-Pastor, B., Hillen, L. M., et al. (2019). Rapid identification of ischemic injury in renal tissue by mass-spectrometry imaging. *Anal. Chem.* 91, 3575–3581. doi: 10.1021/acs.analchem.8b05521

Vollnhals, F., Audinot, J. N., Wirtz, T., Mercier-Bonin, M., Fourquaux, I., Schroeppel, B., et al. (2017). Correlative microscopy combining secondary ion mass spectrometry and electron microscopy: comparison of intensity-hue-saturation and laplacian pyramid methods for image fusion. *Anal. Chem.* 89, 10702–10710. doi: 10.1021/acs.analchem.7b01256

Waanders, L. F., Almeida, R., Prosser, S., Cox, J., Eikel, D., Allen, M. H., et al. (2008). A novel chromatographic method allows on-line reanalysis of the proteome. *Mol. Cell Proteomics* 7, 1452–1459. doi: 10.1074/mcp.M800141-MCP200

Waanders, L. F., Chwalek, K., Monetti, M., Kumar, C., Lammert, E., and Mann, M. (2009). Quantitative proteomic analysis of single pancreatic islets. *Proc. Natl. Acad. Sci. U.S.A.* 106, 18902–18907. doi: 10.1073/pnas.0908351106

Weening, J. J., D'Agati, V. D., Schwartz, M. M., Seshan, S. V., Alpers, C. E., Appel, G. B., et al. (2004). The classification of glomerulonephritis in systemic lupus erythematosus revisited. *Kidney Int.* 65, 521–530.

Wu, C., Dill, A. L., Eberlin, L. S., Cooks, R. G., and Ifa, D. R. (2013). Mass spectrometry imaging under ambient conditions. *Mass Spectrom. Rev.* 32, 218–243. doi: 10.1002/mas.21360

Wu, K. J., and Odom, R. W. (1996). Matrix-enhanced secondary ion mass spectrometry:? a method for molecular analysis of solid surfaces. *Anal. Chem.* 68, 873–882. doi: 10.1021/ac950717i

Xiang, P., Zhu, Y., Yang, Y., Zhao, Z., Williams, S. M., Moore, R. J., et al. (2020). Picoflow liquid chromatography-mass spectrometry for ultrasensitive bottom-up proteomics using 2-μm-i.d. open tubular columns. *Anal. Chem.* 92, 4711–4715. doi: 10.1021/acs.analchem.9b05639

Xu, K., Liang, Y., Piehowski, P. D., Dou, M., Schwarz, K. C., Zhao, R., et al. (2019). Benchtop-compatible sample processing workflow for proteome profiling of <100 mammalian cells. *Anal. Bioanal. Chem.* 411, 4587–4596. doi: 10.1007/s00216-018-1493-9

Yassine, H. N., Trenchevska, O., Dong, Z., Bashawri, Y., Koska, J., Reaven, P. D., et al. (2016). The association of plasma cystatin C proteoforms with diabetic chronic kidney disease. *Proteome Sci.* 14:7. doi: 10.1186/s12953-016-0096-7

Zhang, G., Darshi, M., and Sharma, K. (2018). The warburg effect in diabetic kidney disease. *Semin. Nephrol.* 38, 111–120. doi: 10.1016/j.semnephrol.2018.01.002

Zhang, G., Zhang, J., DeHoog, R. J., Pennathur, S., Anderton, C. R., Venkatachalam, M. A., et al. (2020). DESI-MSI and METASPACE indicates lipid abnormalities and altered mitochondrial membrane components in diabetic renal proximal tubules. *Metabolomics* 16:11. doi: 10.1007/s11306-020-1637-8

Zhu, Y., Piehowski, P. D., Zhao, R., Chen, J., Shen, Y., Moore, R. J., et al. (2018). Nanodroplet processing platform for deep and quantitative proteome profiling of 10-100 mammalian cells. *Nat. Commun.* 9:882. doi: 10.1038/s41467-018-03367-w

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Physiology

MINI REVIEW
published: 04 March 2022
doi: 10.3389/fphys.2022.832457

Tissue Cytometry With Machine Learning in Kidney: From Small Specimens to Big Data

Tarek M. El-Achkar[1], Seth Winfree[2], Niloy Talukder[3], Daria Barwinska[1], Michael J. Ferkowicz[1] and Mohammad Al Hasan[3]**

[1] Division of Nephrology, Department of Medicine, Indiana University, Indianapolis, IN, United States, [2] Department of Pathology and Microbiology, University of Nebraska Omaha, Omaha, NE, United States, [3] Department of Computer and Information Science, Indiana University–Purdue University Indianapolis, Indianapolis, IN, United States

Advances in cellular and molecular interrogation of kidney tissue have ushered a new era of understanding the pathogenesis of kidney disease and potentially identifying molecular targets for therapeutic intervention. Classifying cells *in situ* and identifying subtypes and states induced by injury is a foundational task in this context. High resolution Imaging-based approaches such as large-scale fluorescence 3D imaging offer significant advantages because they allow preservation of tissue architecture and provide a definition of the spatial context of each cell. We recently described the Volumetric Tissue Exploration and Analysis cytometry tool which enables an interactive analysis, quantitation and semiautomated classification of labeled cells in 3D image volumes. We also established and demonstrated an imaging-based classification using deep learning of cells in intact tissue using 3D nuclear staining with 4′,6-diamidino-2-phenylindole (DAPI). In this mini-review, we will discuss recent advancements in analyzing 3D imaging of kidney tissue, and how combining machine learning with cytometry is a powerful approach to leverage the depth of content provided by high resolution imaging into a highly informative analytical output. Therefore, imaging a small tissue specimen will yield big scale data that will enable cell classification in a spatial context and provide novel insights on pathological changes induced by kidney disease.

Keywords: 3D imaging, cytometry analysis, kidney injury, artificial intelligence, deep learning

OPEN ACCESS

Edited by:
Bruce Molitoris, Indiana University, United States

Reviewed by:
Andrew Hall, University of Zurich, Switzerland

****Correspondence:***
Tarek M. El-Achkar
telachka@iu.edu
Mohammad Al Hasan
alhasan@iupui.edu

Specialty section:
This article was submitted to Medical Physics and Imaging, a section of the journal Frontiers in Physiology

Received: *09 December 2021*
Accepted: *28 January 2022*
Published: *04 March 2022*

Citation:
El-Achkar TM, Winfree S, Talukder N, Barwinska D, Ferkowicz MJ and Al Hasan M (2022) Tissue Cytometry With Machine Learning in Kidney: From Small Specimens to Big Data. Front. Physiol. 13:832457. doi: 10.3389/fphys.2022.832457

INTRODUCTION

Understanding the biology and function of an organ requires detailed assessment of various cells and structures in the intact tissue environment (Asp et al., 2019; Stewart et al., 2019; Barwinska et al., 2021). This is particularly needed for the kidney, an organ with complex architecture where zonation of specialized cells and structures is directly linked with physiological function (Hato et al., 2013; El-Achkar and Dagher, 2015; Berry et al., 2017; Barwinska et al., 2021; Ferkowicz et al., 2021). Furthermore, disease states are associated with alteration in tissue architecture and

changes in cell distribution, activity, and/or state (Wilson et al., 2019; Lake et al., 2021; Muto et al., 2021). Technological advancements such as single cell RNA sequencing that provide high content information at the cell and molecular levels have enhanced our ability to further classify cells into subtypes, and study alterations in cell states, which could be linked to disease pathogenesis and outcomes (Park et al., 2018; Lake et al., 2019, 2021; Wilson et al., 2019; Menon et al., 2020). Innovative approaches in high content and high-volume imaging of kidney tissue are also rapidly evolving (Winfree et al., 2017a, 2018, 2021; Singh et al., 2019; Black et al., 2021a,b; Ferkowicz et al., 2021; Lipp et al., 2021; Liu et al., 2021; Melo Ferreira et al., 2021; Neumann et al., 2021), and these advancements are urgently needed to: (1) provide a platform of discovery based on imaging data, thereby delivering a unique context within an intact tissue environment and (2) anchor and interpret *in situ* emerging findings from technologies that lose the spatial context (Winfree et al., 2021). In the last decade, we saw an evolution of imaging kidney tissue from a qualitative toward a highly quantitative science (Winfree et al., 2017b, 2021; Singh et al., 2019; Martins et al., 2020; Black et al., 2021a,b; Melo Ferreira et al., 2021; Neumann et al., 2021). This progress has been enhanced by the advancements in various modalities of microscopy that could perform high-resolution large-scale imaging. The ability to image multiple labels simultaneously (multiplexing) has significantly increased the depth of content acquired (Singh et al., 2019; Woloshuk et al., 2020; Ferkowicz et al., 2021; Melo Ferreira et al., 2021; Neumann et al., 2021). Furthermore, imaging in all 3 dimensions using optical sectioning has allowed faithful preservation of tissue architecture and spatial context (Puelles et al., 2016; Klingberg et al., 2017; Winfree et al., 2017b; Ferkowicz et al., 2021; Lake et al., 2021; Liu et al., 2021). These advancements were catalyzed by the availability of novel software tools that allow streamlined image processing and quantitative analysis (Dao et al., 2016; Winfree et al., 2017a; Czech et al., 2019; Stoltzfus et al., 2020). These significant developments were discussed during the 2021 Indiana University O'Brien Center for Advanced Microscopy Analysis workshop (Dunn et al., 2021).

In this mini-review we will focus on advancement in large scale 3D imaging of kidney tissue and analysis using tissue cytometry with the Volumetric Tissue Exploration and Analysis (VTEA) software tool (**Figure 1**; Winfree et al., 2017b; Ferkowicz et al., 2021). We will also discuss how incorporating novel machine learning approaches and algorithms with tissue cytometry has enhanced the ability to expand and transform the analysis of image volumes toward discovery (Winfree et al., 2021). Particularly, developing deep neural networks that allow classification of cells independent of specific labels will not only increase the power and usefulness of cytometry in classifying cells based on imaging data (Woloshuk et al., 2020), but will also enable unbiased and non-exhaustive discovery of cell subtypes *in situ*. These novel subtypes can then be visualized and mapped back in the image volumes, which will allow biological interpretation. Therefore, this could become a unique opportunity whereby the learning could become interpretable. Furthermore, when large scale 3D imaging is coupled with advanced computational tools that allow processing of large image volumes, hundreds thousand cells or more could be analyzed from a single tissue specimen, thereby allowing the generation of big data from these imaging experiments.

TISSUE CYTOMETRY

Tissue cytometry refers to the process of surveying all cells within an image volume of a tissue, and transforming cells into "analysis-ready" objects with associated variables based on labels (such as fluorescence marker intensities) or spatial parameters. Frequently, the nuclei are used as fiduciaries for the cells because: (1) nuclear staining can be easily incorporated into most experimental designs, and (2) nuclei can be consistently segmented using several standard approaches (Winfree et al., 2017b, 2021; Dunn et al., 2019). The segmented nuclei representing individual cells could then be used in an analytical pipeline that allows quantitative analysis based on the various parameters associated with each cell. The simplest form of analysis is a plot displaying 2 dimensions in the x and y axis, where specific gates could be drawn based on a threshold such as fluorescent label intensity (**Figure 1**). Two key components of tissue cytometry are obtaining quantitative measurements of the cell populations of interest *and* direct visualization by mapping back the cells of interest into the image volume. The latter allows on-the-spot validation of the "choice" of cells (whether by direct gating or other methods) and biological interpretation (particularly when specific distribution patterns start to emerge). Multiple software tools (open-source or commercial) have been developed to perform image analysis, and can be used for tissue cytometry (Gerner et al., 2012; Winfree et al., 2017a; Stoltzfus et al., 2020, 2021; Stirling et al., 2021). We have described the VTEA tool (Winfree et al., 2017b), which was applied specifically to perform tissue cytometry on 3D image volumes of kidney tissue (**Figure 1**). Potential advantages of VTEA include: open-source as a plugin to ImageJ, a single platform that allows image processing, segmentation and cytometry analysis, extensibility and easy incorporation of novel computational approaches, leveraging existing ImageJ tools for image analysis, interactive interplay between the image volume and the analytical process used. We have used tissue cytometry with VTEA in various settings such as to study the abundance and distribution of epithelial and immune cells in the mouse and human kidney (Winfree et al., 2017b; Ferkowicz et al., 2021; Lake et al., 2021), understand the association of epithelial and immune cells to injury in the setting of human acute and chronic kidney disease and stone disease (Lake et al., 2021), quantify and localize the activation of c-JUN in the mouse kidney (Lafavers et al., 2019), study changes in lymphatics in various models of kidney injury (Black et al., 2021a,b). Large scale 3D imaging and tissue cytometry with VTEA is a key tissue interrogation technology used by the Kidney Precision Medicine Project (KPMP) consortium to extract cellular and molecular information from kidney biopsies of patients with kidney disease (De Boer et al., 2021; El-Achkar et al., 2021; Lake et al., 2021). Therefore, the application of tissue cytometry in analyzing kidney

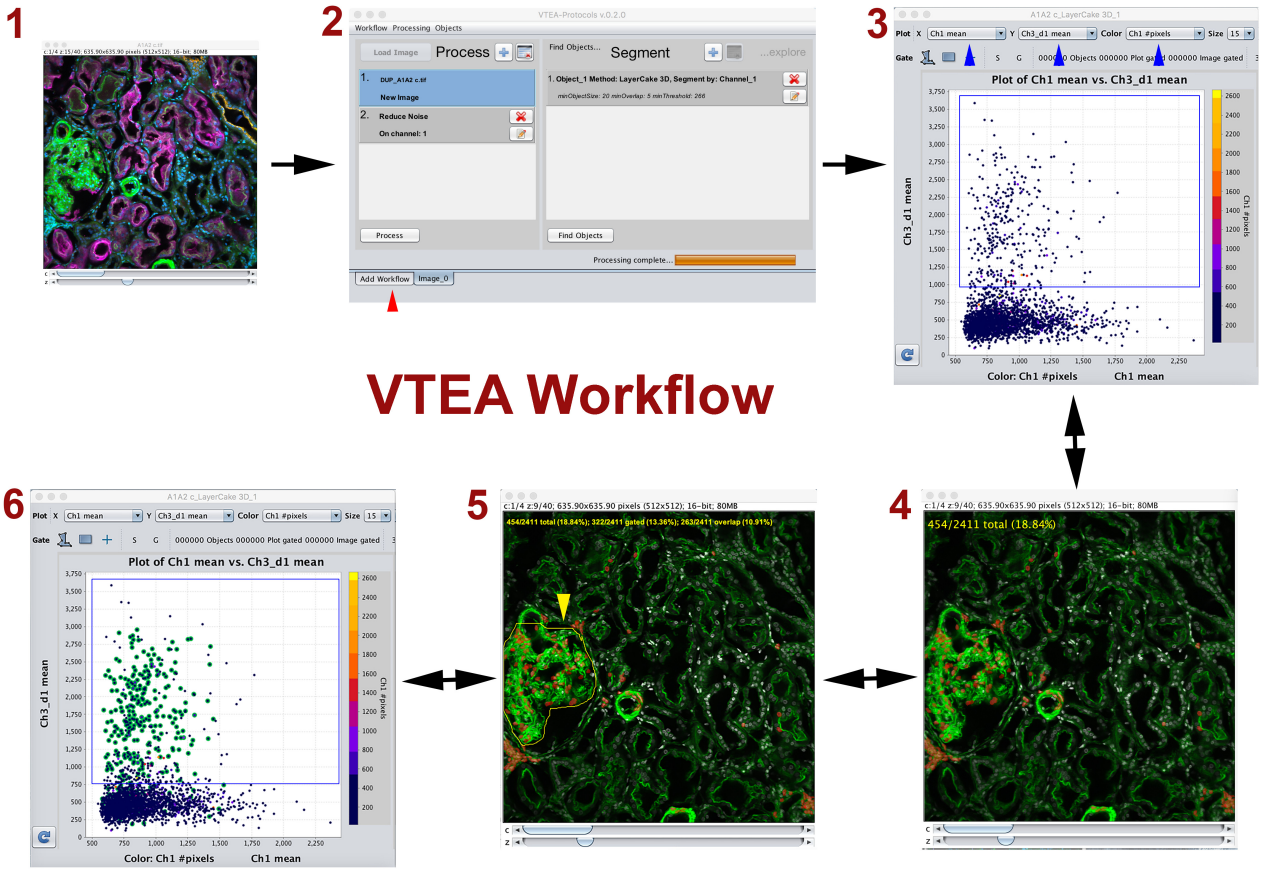


FIGURE 1 | Volumetric Tissue Exploration and Analysis (VTEA) basic workflow. VTEA is a user-friendly platform that allows interactive exploration of image volume (1) where image processing, segmentation, analysis and exploration could occur in a single workspace (2). In the analytical plot, each dot represents a cell with various features. The simplest analysis is in the form of a 2D scatter plot displaying features on the *x* and *y* axis (3). Gates can be drawn to chose and quantify a specific population of cells that can be directly visualized in the image volume (4). Conversely, regions of interest can be drawn in the image (5) to locate cells of interest in the scatter plot (6). This process allows for an explorative interplay between the image and the analytical space. Red arrowhead shows different tabs available in the workspace. Figure adapted and used with permission from Winfree et al. (2017b).

tissue is expanding, and has proven to be complementary to other technologies that do not preserve the tissue architecture.

TISSUE CYTOMETRY AND MACHINE LEARNING

Since multiple parameters can be extracted for each single cell using high resolution multiplexed imaging, advancing the analytical approach to take into account the effect of all these parameters in thousands of cells becomes a big data problem. It is then reasonable to incorporate machine learning algorithms to help cluster, classify and visualize cell subtypes into the analytical space. Indeed, the extensibility of VTEA to incorporate available libraries of machine learning algorithms is a significant development that enables a semi-automated unsupervised classification of cells (Winfree et al., 2022). We demonstrated that this approach could be useful in classifying cells from reference kidney tissue (Woloshuk et al., 2020). In addition, the ability to understand cell-cell and cell-structure interactions could be enhanced by performing neighborhood analysis, as implemented, for example in CytoMAP or VTEA (Stoltzfus et al., 2020; Lake et al., 2021; Winfree et al., 2022). We recently used VTEA to perform a cell centric neighborhood analysis on >1.2 million cells from various kidney biopsies of patients with kidney disease (Lake et al., 2021). This approach uncovered spatial associations that were validated by other transcriptomics-based technologies. One of the key advantages of performing such cell-centric neighborhood analysis (percent of cells within a distance from each cell) is the ability to merge analysis from various specimens into one analytical space, since such analysis is by default normalized (Lake et al., 2021).

Multiplexing various probes into one imaging experiment offers significant advantages for cell classification based on particular labels. For example, using highly multiplexed detection such as imaging mass cytometry or co-detection by indexing allows the classification of multiple cell subtypes (Singh et al., 2019; Melo Ferreira et al., 2021; Neumann et al., 2021). However, multiplexing has limitations, particularly in its application in 3D and its practicality when kidney tissue is of limited availability

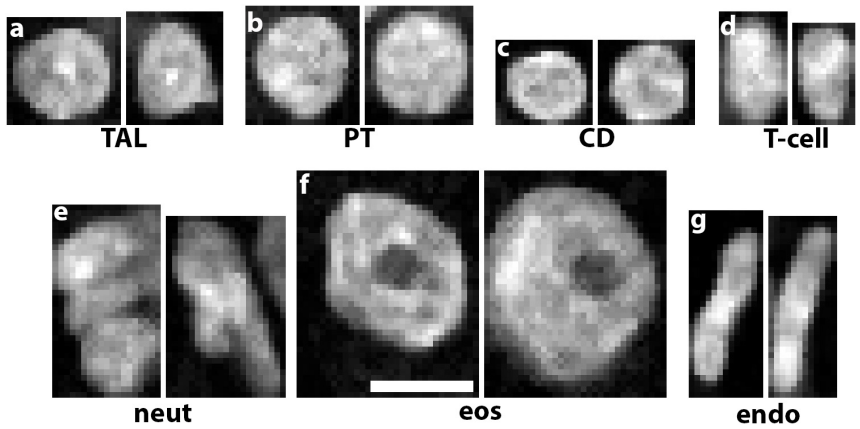


FIGURE 2 | Unique nuclear staining signatures of various kidney cell types. DAPI staining alone reveals distinct signatures of chromatin condensation states and nuclear morphology of **(a)** thick ascending limbs (TAL), **(b)** proximal tubules (PT), **(c)** collecting ducts (CD), **(d)** T-cells, **(e)** neutrophils, **(f)** eosinophils, and **(g)** endothelial cells. Scale bar = 5 μm.

(Winfree et al., 2021). In addition, using pre-specified labels limits the potential of agnostic discovery of novel cell types and subtypes based on imaging data. To circumvent these limitations, we recently devised a deep learning approach to classify cells based only on nuclear staining (Woloshuk et al., 2020). The premise is based on the fact that nuclear staining has unique features for each cell type (**Figure 2**) and its changes could represent alterations in cell states (Gustafsdottir et al., 2013; Eulenberg et al., 2017). Therefore, these studies are confined within a biologically interpretable context (Woloshuk et al., 2020). This work presented us also with a unique opportunity to test several unexplored questions such as: is 2D enough or do we need the information in 3D image volumes of nuclei? Can we use classical machine learning classifiers that extract features or do a deep neural network work better? Does the context of the nucleus (i.e., neighboring nuclei) improve cell classification accuracy? Our results showed that we could successfully classify cells from human reference kidneys into eight different classes based on machine learning approaches, but the highest accuracy was achieved with a 3D deep neural network trained on 3D image volume of nuclei with context (Woloshuk et al., 2020). Our efforts are currently to extend this approach to kidney disease, and use the 3D leaning network to uncover cell subtypes induced by injury. This could be done by using various approaches. For example, the features extracted by the 3D network from the nuclear staining could be used to reclassify and visualize cells using tissue cytometry. Importantly, novel machine learning tools could be applied on these features to achieve non-exhaustive learning and agnostically discover new cell subtypes that can be further vetted using tissue cytometry. This will be discussed next.

LEVERAGING MACHINE LEARNING FOR AGNOSTIC DISCOVERY

Agnostic discovery is the exploration process for the identification and localization of novel kidney cell subtypes induced by injury. In an agnostic discovery scenario, obtaining labeled cell examples for the injury cell subtypes is a hard task for many reasons: first, the nature of injury to the morphology of kidney cells due to disease is unknown so accurate labeling is difficult; second, we may not yet have a suitable marker for such cells, which makes us unable to correctly label them using cytometry; finally, due to lack of knowledge regarding the injury it is even hard for us to know the definite count of number of possible injury subtypes. While lack of labeled data makes agnostic discovery a difficult task, recent advances in supervised classification can help us in this regard (**Figure 3**).

In supervised classification, identifying novel classes (example: novel injury states) for which no examples are available in the training data (ground truth datasets used in training machine learning classifier) has received wide-spread attention from the machine learning community in recent years. There are different approaches for solving such machine learning tasks. Most prominent among these is called zero-shot learning (ZSL), which is well studied by the deep learning community (Romera-Paredes and Torr, 2015; Zhang and Saligrama, 2015). ZSL is also becoming a promising direction in the medical domain. In recent works, ZSL has been used in diagnosis and classification of disease in chest radiographs (Hayat et al., 2021; Paul et al., 2021). Bayesian non-exhaustive classification is another prominent direction (Görür and Edward Rasmussen, 2010; Ben-Yosef and Weinshall, 2018).

For zero-shot learning, the number of novel classes along with side-information (also known as semantic information) about all the classes needs to be provided upfront. During training, the learning algorithm utilizes the side information to compensate for the lack of labeled data for the unknown class. Spatial neighborhood data around a cell can be used as side-information. For instance, we expect that the concentration of inflammatory immune cells (such as neutrophils or T-cells) around injured kidney cells would be higher, and hence, such side information will be relevant for classifying injured kidney cell subclass instances. Another potential candidate for side

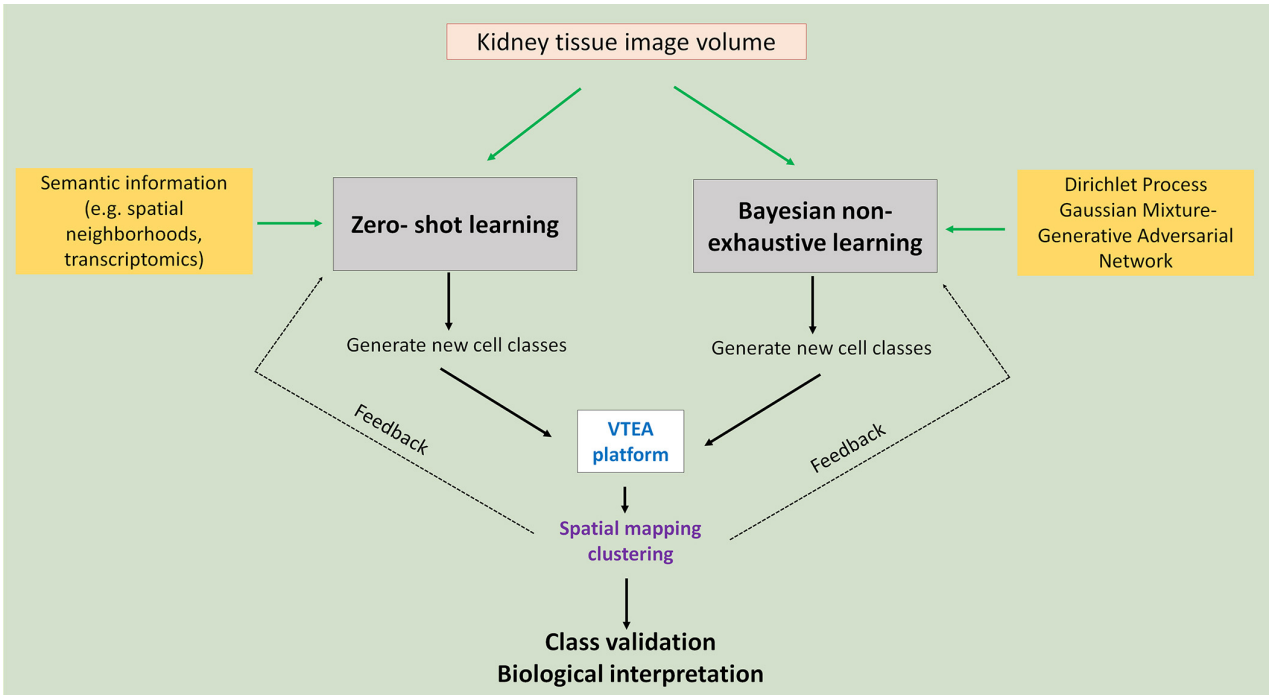


FIGURE 3 | Agnostic discovery using machine learning and tissue cytometry. Proposed approach to use imaging data of cell nuclei in machine learning workflows that allow non-exhaustive classification of new classes that could be visualized and further analyzed using tissue cytometry. Green arrows point to the two proposed approaches: zero-shot and Bayesian non-exhaustive learning.

information is to use data from alternate modality, say, single cell RNA sequencing and/or spatial transcriptomics data. Using such data will not only help us identify novel kidney cells, but also will provide more information regarding the pathways that control the injury progression in the kidney cells over time. A challenge of zero-shot learning is that it requires that the number of novel classes is known during training time, which is often not feasible for agnostic discovery. In that case, Bayesian non-exhaustive classification can be used. Using Bayesian technique, it learns some parametric probability distribution for the known classes. During inference, it identifies instances which are far away from the distribution of the known classes and create a new class along with its probability distribution. Generally, Dirichlet process Gaussian Mixture Model is used for non-exhaustive classification (Görür and Edward Rasmussen, 2010; Zhuang and Al Hasan, 2021). The challenge in Bayesian non-exhaustive classification is that their performance becomes very poor if the assumed data distribution does not follow the actual data distribution.

CHALLENGES FOR IMAGE-BASED CLASSIFICATION IN HUMAN KIDNEY BIOPSIES

The novel imaging-based approaches discussed to characterize cell types and subtypes in human kidney tissue specimens are very promising. However, it is also important to discuss some of the challenges and limitations that need to be addressed to make these methodologies more robust and accessible. First, variation in tissue processing practices and fixation may alter the quality of the tissue and the downstream imaging data. The effects of changes in tissue processing on the ability to classify cells using tissue cytometry and machine learning are unknown. Fortunately, collaborative studies (such as the KPMP consortium) that are focused on interrogating kidney tissue biopsy specimens are rigorously standardizing tissue acquisition and processing, which would allow to set standards and perform comparison with data acquired from archived tissues originating from other sources (De Boer et al., 2021; El-Achkar et al., 2021). Second, it is possible that some of the changes in cell states that are induced by disease may not be accompanied by significant alteration in nuclear activity or morphology. Therefore, expanding classification strategies to include another common marker that tracks changes in cell morphology and activity in the cytoplasm such as F-actin will likely increase the sensitivity and dynamic range of capturing subtle changes in cell states. Finally, performing imaging and data analysis is frequently limited to centers with appropriate expertise and resources, which may limit accessibility to the broader research community. Furthermore, the computational breadth needed for data access, storage and transfer may also be a restrictive factor. Therefore, increasing the accessibility of these approaches by using and disseminating open-source software, public imaging data repositories and accessible cloud-based imaging visualization and analysis tools will provide reasonable first steps to make these innovative tools more reachable by the broader community.

CONCLUSION AND FUTURE OUTLOOK

We highlighted in this mini-review advances in tissue cytometry of kidney tissue, emphasizing novel analytical approaches that transform imaging-based data into highly quantifiable big data outputs that can also be used for discovery while incorporating the richness of the spatial context. These advances are crucial to understand kidney disease, which frequently displays regional heterogeneity at the cellular and molecular levels. Leveraging novel machine learning approaches will allow unbiased discoveries such as novel cell types and subtypes which are spatially anchored and linked to other features that allow biological interpretation. In the future, we anticipate that with relatively few labels, the combination of tissue cytometry with machine learning will enable a form of enhanced "*virtual multiplexing,*" which could classify most cell types *in situ* within kidney tissue and allow the agnostic discovery of novel cell types based on imaging. For the kidney, imaging and analyzing tissue will certainly become a very important issue!

AUTHOR CONTRIBUTIONS

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

FUNDING

This work was supported by funding from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) P30 DK079312 (TE-A) and NIDDK Diabetic Complications Consortium (RRID: SCR_001415, www.diacomp.org), grants DK076169 and DK115255 (TE-A and MA).

REFERENCES

Asp, M., Giacomello, S., Larsson, L., Wu, C., Furth, D., Qian, X., et al. (2019). A Spatiotemporal Organ-Wide Gene Expression and Cell Atlas of the Developing Human Heart. *Cell* 179, 647-1660.e19. doi: 10.1016/j.cell.2019.11.025

Barwinska, D., El-Achkar, T. M., Melo Ferreira, R., Syed, F., Cheng, Y. H., and Winfree, S. (2021). Molecular characterization of the human kidney interstitium in health and disease. *Sci Adv* 7:eabd3359. doi: 10.1126/sciadv.abd3359

Ben-Yosef, M., and Weinshall, D. (2018). Gaussian mixture generative adversarial networks for diverse datasets, and the unsupervised clustering of images. *arXiv preprint arXiv* 1808.10356. [preprint]

Berry, M. R., Mathews, R. J., Ferdinand, J. R., Jing, C., Loudon, K. W., Wlodek, E., et al. (2017). Renal Sodium Gradient Orchestrates a Dynamic Antibacterial Defense Zone. *Cell* 170, 860-874.e19.

Black, L. M., Farrell, E. R., Barwinska, D., Osis, G., Zmijewska, A. A., Traylor, A. M., et al. (2021a). VEGFR3 tyrosine kinase inhibition aggravates cisplatin nephrotoxicity. *Am. J. Physiol. Renal. Physiol.* 321, F675–F688. doi: 10.1152/ajprenal.00186.2021

Black, L. M., Winfree, S., Khochare, S. D., Kamocka, M. M., Traylor, A. M., Esman, S. K., et al. (2021b). Quantitative 3-dimensional imaging and tissue cytometry reveals lymphatic expansion in acute kidney injury. *Lab Invest.* 101, 1186–1196. doi: 10.1038/s41374-021-00609-2

Czech, E., Aksoy, B. A., Aksoy, P., and Hammerbacher, J. (2019). Cytokit: a single-cell analysis toolkit for high dimensional fluorescent microscopy imaging. *BMC Bioinformatics* 20:448. doi: 10.1186/s12859-019-3055-3

Dao, D., Fraser, A. N., Hung, J., Ljosa, V., Singh, S., and Carpenter, A. E. (2016). CellProfiler Analyst: interactive data exploration, analysis and classification of large biological image sets. *Bioinformatics* 32, 3210–3212. doi: 10.1093/bioinformatics/btw390

De Boer, I. H., Alpers, C. E., Azeloglu, E. U., Balis, U. G. J., Barasch, J. M., and Barisoni, L. (2021). Rationale and design of the Kidney Precision Medicine Project. *Kidney Int.* 99, 498–510. doi: 10.1016/j.kint.2020.08.039

Dunn, K. W., Fu, C., Ho, D. J., Lee, S., Han, S., Salama, P., et al. (2019). DeepSynth: Three-dimensional nuclear segmentation of biological images using neural networks trained with synthetic data. *Sci. Rep.* 9:18295.

Dunn, K. W., Molitoris, B. A., and Dagher, P. C. (2021). The Indiana O'Brien Center for Advanced Renal Microscopic Analysis. *Am. J. Physiol. Renal Physiol.* 320, F671–F682. doi: 10.1152/ajprenal.00007.2021

El-Achkar, T. M., and Dagher, P. C. (2015). Tubular cross talk in acute kidney injury: a story of sense and sensibility. *Am. J. Physiol. Renal. Physiol.* 308, F1317–F1323. doi: 10.1152/ajprenal.00030.2015

El-Achkar, T. M., Eadon, M. T., Menon, R., Lake, B. B., Sigdel, T. K., Alexandrov, T., et al. (2021). A multimodal and integrated approach to interrogate human kidney biopsies with rigor and reproducibility: guidelines from the Kidney Precision Medicine Project. *Physiol. Genomics* 53, 1–11. doi: 10.1152/physiolgenomics.00104.2020

Eulenberg, P., Kohler, N., Blasi, T., Filby, A., Carpenter, A. E., Rees, P., et al. (2017). Reconstructing cell cycle and disease progression using deep learning. *Nat. Commun.* 8:463. doi: 10.1038/s41467-017-00623-3

Ferkowicz, M. J., Winfree, S., Sabo, A. R., Kamocka, M. M., Khochare, S., and Barwinska, D. (2021). Large-scale, three-dimensional tissue cytometry of the human kidney: a complete and accessible pipeline. *Lab Invest.* 101, 661–676. doi: 10.1038/s41374-020-00518-w

Gerner, M. Y., Kastenmuller, W., Ifrim, I., Kabat, J., and Germain, R. N. (2012). Histo-cytometry: a method for highly multiplex quantitative tissue imaging analysis applied to dendritic cell subset microanatomy in lymph nodes. *Immunity* 37, 364–376. doi: 10.1016/j.immuni.2012.07.011

Görür, D., and Edward Rasmussen, C. (2010). Dirichlet Process Gaussian Mixture Models: Choice of the Base Distribution. *J. Comput. Sci. Technol.* 25, 653–664.

Gustafsdottir, S. M., Ljosa, V., Sokolnicki, K. L., Anthony Wilson, J., Walpita, D., Kemp, M. M., et al. (2013). Multiplex cytological profiling assay to measure diverse cellular states. *PLoS One* 8:e80999. doi: 10.1371/journal.pone.0080999

Hato, T., El-Achkar, T. M., and Dagher, P. C. (2013). Sisters in arms: myeloid and tubular epithelial cells shape renal innate immunity. *Am. J. Physiol. Renal Physiol.* 304, F1243–F1251. doi: 10.1152/ajprenal.00101.2013

Hayat, N., Lashen, H., and Shamout, F. E. (2021). Multi-Label Generalized Zero Shot Learning for the Classification of Disease in Chest Radiographs. *arXiv preprint arXiv* 1808.10356. [PREPRINT]

Klingberg, A., Hasenberg, A., Ludwig-Portugall, I., Medyukhina, A., Mann, L., Brenzel, A., et al. (2017). Fully Automated Evaluation of Total Glomerular Number and Capillary Tuft Size in Nephritic Kidneys Using Lightsheet Microscopy. *J. Am. Soc. Nephrol.* 28, 452–459. doi: 10.1681/ASN.2016020232

Lafavers, K. A., Macedo, E., Garimella, P. S., Lima, C., Khan, S., Myslinski, J., et al. (2019). Circulating uromodulin inhibits systemic oxidative stress by inactivating the TRPM2 channel. *Sci. Transl. Med.* 11:eaaw3639. doi: 10.1126/scitranslmed.aaw3639

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.* 10:2832. 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*.[PREPRINT] 2021.2007.2028.454201.

Lipp, S. N., Jacobson, K. R., Hains, D. S., Schwarderer, A. L., and Calve, S. (2021). 3D Mapping Reveals a Complex and Transient Interstitial Matrix During Murine Kidney Development. *J. Am. Soc. Nephrol.* 32, 1649–1665. doi: 10.1681/ASN.2020081204

Liu, H., Hiremath, C., Patterson, Q., Vora, S., Shang, Z., Jamieson, A. R., et al. (2021). Heterozygous Mutation of Vegfr3 Reduces Renal Lymphatics without

Renal Dysfunction. *J. Am. Soc. Nephrol.* 32, 3099–3113. doi: 10.1681/ASN.2021010061

Martins, J. R., Haenni, D., Bugarski, M., Figurek, A., and Hall, A. M. (2020). Quantitative intravital Ca(2+) imaging maps single cell behavior to kidney tubular structure. *Am. J. Physiol. Renal Physiol.* 319, F245–F255. doi: 10.1152/ajprenal.00052.2020

Melo Ferreira, R., Sabo, A. R., Winfree, S., Collins, K. S., Janosevic, D., Gulbronson, C. J., et al. (2021). Integration of spatial and single-cell transcriptomics localizes epithelial cell-immune cross-talk in kidney injury. *JCI Insight* 6:e147703. doi: 10.1172/jci.insight.147703

Menon, R., Otto, E. A., Hoover, P., Eddy, S., Mariani, L., and Godfrey, B. (2020). Single cell transcriptomics identifies focal segmental glomerulosclerosis remission endothelial biomarker. *JCI Insight* 5:e133267. doi: 10.1172/jci.insight.133267

Muto, Y., Wilson, P. C., Ledru, N., Wu, H., Dimke, H., Waikar, S. S., et al. (2021). Single cell transcriptional and chromatin accessibility profiling redefine cellular heterogeneity in the adult human kidney. *Nat. Commun.* 12:2190. doi: 10.1038/s41467-021-22368-w

Neumann, E. K., Patterson, N. H., Rivera, E. S., Allen, J. L., Brewer, M., Decaestecker, M. P., et al. (2021). Highly multiplexed immunofluorescence of the human kidney using co-detection by indexing. *Kidney Int*. 101, 137–143. doi: 10.1016/j.kint.2021.08.033

Park, J., Shrestha, R., Qiu, C., Kondo, A., Huang, S., Werth, M., et al. (2018). Single-cell transcriptomics of the mouse kidney reveals potential cellular targets of kidney disease. *Science* 360, 758–763. doi: 10.1126/science.aar2131

Paul, A., Shen, T. C., Lee, S., Balachandar, N., Peng, Y., Lu, Z., et al. (2021). Generalized Zero-Shot Chest X-Ray Diagnosis Through Trait-Guided Multi-View Semantic Embedding With Self-Training. *IEEE Trans. Med. Imaging* 40, 2642–2655. doi: 10.1109/TMI.2021.3054817

Puelles, V. G., Van Der Wolde, J. W., Schulze, K. E., Short, K. M., Wong, M. N., Bensley, J. G., et al. (2016). Validation of a Three-Dimensional Method for Counting and Sizing Podocytes in Whole Glomeruli. *J. Am. Soc. Nephrol.* 27, 3093–3104. doi: 10.1681/ASN.2015121340

Romera-Paredes, B., and Torr, P. (2015). "An embarrassingly simple approach to zero-shot learning", in *Proceedings of the 32nd International Conference on Machine Learning*, July 2015, Lille, 2152–2161.

Singh, N., Avigan, Z. M., Kliegel, J. A., Shuch, B. M., Montgomery, R. R., Moeckel, G. W., et al. (2019). Development of a 2-dimensional atlas of the human kidney with imaging mass cytometry. *JCI Insight* 4:e129477. doi: 10.1172/jci.insight.129477

Stewart, B. J., Ferdinand, J. R., Young, M. D., Mitchell, T. J., Loudon, K. W., and Riding, A. M. (2019). Spatiotemporal immune zonation of the human kidney. *Science* 365, 1461–1466. doi: 10.1126/science.aat5031

Stirling, D. R., Carpenter, A. E., and Cimini, B. A. (2021). CellProfiler Analyst 3.0: Accessible data exploration and machine learning for image analysis. *Bioinformatics* 3:btab634. doi: 10.1093/bioinformatics/btab634

Stoltzfus, C. R., Filipek, J., Gern, B. H., Olin, B. E., Leal, J. M., Wu, Y., et al. (2020). CytoMAP: A Spatial Analysis Toolbox Reveals Features of Myeloid Cell Organization in Lymphoid Tissues. *Cell Rep* 31:107523. doi: 10.1016/j.celrep.2020.107523

Stoltzfus, C. R., Sivakumar, R., Kunz, L., Olin Pope, B. E., Menietti, E., Speziale, D., et al. (2021). Multi-Parameter Quantitative Imaging of Tumor Microenvironments Reveals Perivascular Immune Niches Associated With Anti-Tumor Immunity. *Front Immunol* 12:726492. doi: 10.3389/fimmu.2021.726492

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, 19619–19625. doi: 10.1073/pnas.1908706116

Winfree, S., Al Hasan, M., and El-Achkar, T. M. (2021) . Profiling immune cells in the kidney using tissue cytometry and machine learning. *Kidney* 360, doi: 10.34067/KID.0006802020 [Epub ahead of print].

Winfree, S., Dagher, P. C., Dunn, K. W., Eadon, M. T., Ferkowicz, M., Barwinska, D., et al. (2018). Quantitative Large-Scale Three-Dimensional Imaging of Human Kidney Biopsies: A Bridge to Precision Medicine in Kidney Disease. *Nephron* 140, 134–139. doi: 10.1159/000490006

Winfree, S., Ferkowicz, M. J., Dagher, P. C., Kelly, K. J., Eadon, M. T., Sutton, T. A., et al. (2017a). Large-scale 3-dimensional quantitative imaging of tissues: state-of-the-art and translational implications. *Transl. Res.* 189, 1–12. doi: 10.1016/j.trsl.2017.07.006

Winfree, S., Khan, S., Micanovic, R., Eadon, M. T., Kelly, K. J., Sutton, T. A., et al. (2017b). Quantitative Three-Dimensional Tissue Cytometry to Study Kidney Tissue and Resident Immune Cells. *J. Am. Soc. Nephrol.* 28, 2108–2118. doi: 10.1681/ASN.2016091027

Winfree, S., Mcnutt, A. T., Khochare, S., Borgard, T. J., Barwinska, D., Sabo, A. R., et al. (2022). Integrated cytometry with machine learning applied to high-content imaging of human kidney tissue for in-situ cell classification and neighborhood analysis. *bioRxiv*.[PREPRINT] 2021.2012.2027.474025.

Woloshuk, A., Khochare, S., Almulhim, A. F., Mcnutt, A. T., Dean, D., Barwinska, D., et al. (2020). *In Situ* Classification of Cell Types in Human Kidney Tissue Using 3D Nuclear Staining. *Cytometry A* 99, 707–721. doi: 10.1002/cyto.a.24274

Zhang, Z., and Saligrama, V. (2015). *Zero-shot learning via semantic similarity embedding", in: Proceedings of the IEEE international conference on computer vision*). Manhattan, NY: IEEE international, 4166–4174.

Zhuang, J., and Al Hasan, M. (2021). "Non-exhaustive Learning Using Gaussian Mixture Generative Adversarial Networks," in *Machine Learning and Knowledge Discovery in Databases. Research Track*, (Midtown Manhattan, NY: Springer International Publishing), 3–18. doi: 10.1007/978-3-030-86520-7_1

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

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.

frontiers in Physiology

MINI REVIEW
published: 10 March 2022
doi: 10.3389/fphys.2022.833333

User-Accessible Machine Learning Approaches for Cell Segmentation and Analysis in Tissue

*Seth Winfree**

Department of Pathology and Microbiology, University of Nebraska Medical Center, Omaha, NE, United States

Advanced image analysis with machine and deep learning has improved cell segmentation and classification for novel insights into biological mechanisms. These approaches have been used for the analysis of cells *in situ,* within tissue, and confirmed existing and uncovered new models of cellular microenvironments in human disease. This has been achieved by the development of both imaging modality specific and multimodal solutions for cellular segmentation, thus addressing the fundamental requirement for high quality and reproducible cell segmentation in images from immunofluorescence, immunohistochemistry and histological stains. The expansive landscape of cell types-from a variety of species, organs and cellular states-has required a concerted effort to build libraries of annotated cells for training data and novel solutions for leveraging annotations across imaging modalities and in some cases led to questioning the requirement for single cell demarcation all together. Unfortunately, bleeding-edge approaches are often confined to a few experts with the necessary domain knowledge. However, freely available, and open-source tools and libraries of trained machine learning models have been made accessible to researchers in the biomedical sciences as software pipelines, plugins for open-source and free desktop and web-based software solutions. The future holds exciting possibilities with expanding machine learning models for segmentation via the brute-force addition of new training data or the implementation of novel network architectures, the use of machine and deep learning in cell and neighborhood classification for uncovering cellular microenvironments, and the development of new strategies for the use of machine and deep learning in biomedical research.

Keywords: machine learning, deep learning—artificial neural network, segmentation, classification, neighborhoods, microenviroment, bio-imaging tools

OPEN ACCESS

Edited by:
Bruce Molitoris,
Indiana University, United States

Reviewed by:
Noriko F. Hiroi,
Keio University Shonan Fujisawa Campus, Japan

***Correspondence:**
Seth Winfree
swinfree@unmc.edu

Specialty section:
This article was submitted to Computational Physiology and Medicine, a section of the journal Frontiers in Physiology

Received: *13 December 2021*
Accepted: *12 January 2022*
Published: *10 March 2022*

Citation:
Winfree S (2022) User-Accessible Machine Learning Approaches for Cell Segmentation and Analysis in Tissue. Front. Physiol. 13:833333. doi: 10.3389/fphys.2022.833333

INTRODUCTION

Image use in the biomedical sciences varies from demonstrative and representative to data for quantitative interrogation. Quantitative analyses of tissue and cells, the basic building blocks in biology, requires the accurate segmentation of cells or surrogates of cells and methods for classifying cells and quantitative analysis of cell type, cell states and function. Cellular segmentation has been an intense focus in biomedical image analysis for decades and has evolved from largely *ad hoc* approaches to generalizable solutions (Meijering, 2012). The classification strategies for cell type (e.g., immune cell, epithelium, stromal, etc.) and state (e.g., injured, repairing, dividing) have

Abbreviations: ML, machine learning; DL, deep learning; GUI, graphical user interface; API, application programming interface.

developed rapidly (Meijering, 2012, 2020; Meijering et al., 2016). How different cell types organize into microenvironments or neighborhoods is import for our understanding of pathogenesis and biology. The identification and classification of these neighborhood or microenvironments is of significant interest to the bioimaging community (Allam et al., 2020; Stoltzfus et al., 2020; Solorzano et al., 2021). This mini review will cover the current state of quantitative analysis of tissues and cells in imaging data, with a discussion of segmentation, classification, and neighborhood analysis, specifically highlighting the application of machine learning, including recent advancements, challenges, and the tools available to the biomedical researchers.

SEGMENTATION

A cornucopia of segmentation approaches has been developed for specific experimental situations, tissue types or cell populations including clusters of cells, specific cell types, etc. (Meijering, 2012; Meijering et al., 2016). Often these approaches are built as pipelines in image processing software, enabling the sharing of segmentation methods (Berthold et al., 2007; de Chaumont et al., 2012; Schindelin et al., 2012; Bankhead et al., 2017; McQuin et al., 2018; Berg et al., 2019). A common approach is to first differentiate foreground, the cell, from background in a semantic segmentation step. Secondly, objects of interest in the image are isolated, or instance segmentation, by identifying and then splitting touching cells. Meijering outlined five fundamental methods for segmentation: intensity thresholding (Otsu, 1979), feature detection, morphologically based, deformable model fitting and region accumulation or splitting (Meijering, 2012). These methods are often combined sequentially. For instance, cell segmentation might include *semantic* segmentation of a foreground of all nuclei with pixel intensity, followed by a second *instance* segmentation for identifying an individual nucleus using a region accumulation approach like watershed (Beucher and Lantuejoul, 1979). A common limitation is the *ad hoc* nature of segmentation approaches: the applicability of a segmentation method may be limited by constraints in the datasets including differences in staining or imaging modality (fluorescence vs. histology staining), artifacts in image capture (out-of-focus light or uneven field illumination) or morphological differences (spherical epithelial vs. more cylindrical muscle nuclei). These constraints, and others, have limited the development of generalizable segmentation algorithms.

Cell segmentation with machine learning is well established- a popular approach is to perform semantic pixel segmentation with a Random Forest Classifier (Hall et al., 2009; McQuin et al., 2018; Berg et al., 2019). Segmentation with a Random Forest Classifier, as with all machine learning approaches, requires training data. In cell segmentation this is data that has been annotated to indicate which pixels in images are foreground, nuclei, vs. background. ilastik provides an intuitive and iterative solution for generating training data with a GUI that allows a user to: (1) highlight pixels to indicate nuclei vs. background-training data, (2) test classification and segmentation, (3) repeat and add or subtract highlighted pixels, to improve the classification and segmentation. This process is powerful but can become labor-intensive in different tissues where there may be a variety of nuclei (e.g., shape, texture, size, clustering, etc.) in smooth muscle, epithelium, endothelium, and immune cells in varying densities and distributions. Unfortunately, while high quality cell culture nuclei training datasets and tissue image datasets exist, 2D training data of nuclei in tissue is limited or fractured across multiple repositories (Ljosa et al., 2012; Williams et al., 2017; Ellenberg et al., 2018; Kume and Nishida, 2021). Furthermore, while 3D electron microscopy data is readily available, 3D fluorescence image or training datasets of nuclei is limited (Ljosa et al., 2012; Iudin et al., 2016; El-Achkar et al., 2021; Lake et al., 2021). The availability of training data is one of the most significant barriers to the application of machine learning to cell image segmentation (Ching et al., 2018). Fortunately, the number of venues to share imaging datasets should not limit the dissemination of training datasets as they are generated (**Table 1**, Datasets and Repositories).

Recently, three novel approaches were developed to address the dearth of segmentation training data for the variety of cell-types and imaging modalities. The first, and most direct approach has been the concerted effort of a number of groups including the Van Valen and Lundberg laboratories to establish "human-in-the-loop" pipelines and infrastructure of software and personnel, including collaborative crowd sourcing, to generate ground truth from imaging datasets (Sullivan et al., 2018; Moen et al., 2019; Bannon et al., 2021). A limitation of this approach is the requirement for on-going support for personnel; on-going support is critical to long term success. To ease the generation of high-quality training data with a "human-in-the-loop" approach, methods have also been established around segmentation refinement (Sullivan et al., 2018; Lutnick et al., 2019; Moen et al., 2019; Govind et al., 2021; Lee et al., 2021). An alternative to these brute-force approach has been to generate synthetic training data by combining "blob" models of cells with real images using generative adversarial networks (Dunn et al., 2019; Wu et al., 2021). Further, to leverage training data across imaging modalities NucleAIzer[1] relies on style transfer with a generative-adversarial-network to generate synthetic data using prior training data from other modalities (fluorescence, histological stains, or immunohistochemistry). Thus, this approach can expand training data by mapping to a common modality, giving a nearly general solution to segmentation across 2D imaging modalities (Hollandi et al., 2020).

The on-going search for generalizable segmentation is an area of active research in deep learning and is critical to establishing rigorous and reproducible segmentation approaches. To this end, a pipeline that requires little to no tuning on multiple datasets and modalities was demonstrated recently (Waibel et al., 2021). In the interim, the field will continue to make progress with generalizable segmentation, existing approaches, networks, etc. can provide the foundation for novel segmentation solutions. For instance, deep learning approaches to address 2D and 3D cell segmentation are often based on existing networks (Haberl et al., 2018; Schmidt et al., 2018; Falk et al., 2019; Weigert et al., 2020; Minaee et al., 2021; Stringer et al., 2021), using training data augmentation (Moshkov et al., 2020), or transfer learning

[1]https://www.nucleaizer.org/

TABLE 1 | End user accessibility of tools supporting machine and/or deep learning for bioimage analysis.

User	Application	Name	Support or demonstrated		Description	Software type	URL	References
			Classical learning	Deep learning (DL)				
Novice, Practitioner, Developer	Datasets and repositories	Image data resource (IDR)	Not determined	Yes, ex. Idr0042	Tissue and cell images with cell based training datasets	Repository	https://idr.openmicroscopy.org	Williams et al., 2017
		Broad bioimage benchmark collection	Yes	Yes	Cell images training datasets	Repository	https://bbbc.broadinstitute.org	Ljosa et al., 2012
		Cell image library	Not determined	CDeep3M	Mulitmodal cell images, linked to *CDeep3M* for testing	Repository	http://cellimagelibrary.org/pages/datasets	NA
		BioImageDbs	Yes	Yes	R package and repository for images	Bioconductor package	https://kumes.github.io/BioImageDbs/	Kume and Nishida, 2021
		EMPIAR	Yes, ex. EMPIAR-10069	Yes, ex. EMPIAR-10592	Electron microscopy images	Repository	https://www.ebi.ac.uk/empiar/	Iudin et al., 2016
		SciLifeLab	Not determined	Yes	Scientific data, images and figure	Repository	https://www.scilifelab.se/data/repository/	NA
		BioImage Archive	Yes	Yes	Archive of IDR and EMPIAR	Repository	https://www.ebi.ac.uk/bioimage-archive/	Ellenberg et al., 2018
	Segmentation	DeepCell Kiosk		Establishing a cellwise dataset	Tool for segmentation in the cloud	Web interface	deepcell.org	Moen et al., 2019
		Cellpose		Segmentation	Tool for segmentation in the cloud and python GUI	Web interface, application	https://github.com/MouseLand/cellpose	Stringer et al., 2021
		NucleAIzer		Transfer learning	Tool for segmentation in the cloud	Web interface	www.nucleaizer.org	Hollandi et al., 2020
		CDeep3M		Electron microscopy segmentation	Multiple trained networks for distinct structures in EM images	Web interface, model zoo	https://cdeep3m.crbs.ucsd.edu/	Haberl et al., 2018
		QuPath	Feature design for segmentation	Inference with StarDist	ML segmentation with GUI	Application, plugin	qupath.github.io	Bankhead et al., 2017
		DeepImageJ		Inference in ImageJ with BioImage.IO	Tool for inference on the desktop	Plugin	deepimagej.github.io/deepimagej/	Gómez-de-Mariscal et al., 2021
		Ilastik	Feature design for segmentation	Interfaces with BioImage.IO	Segmentation with GUI	Application, plugin	www.ilastik.org	Berg et al., 2019
		CellProfiler and CellAnalyst	Feature design for classification	Unet Segmentation	Pipeline Based image processing tool with ML and DL support	Application	cellprofiler.org	Dao et al., 2016; McQuin et al., 2018
		StarDist		Segmentation	Python and Java (ImageJ/FIJI) tool for segmentation	Plugin	https://github.com/stardist/stardist	Weigert et al., 2020
		HistomicsML2		Model for training and tools for inference	Framework for training and inference on imaging data	Web interface	https://histomicsml2.readthedocs.io/	Lee et al., 2021
		CSBDeep		Image restoration, segmentation	FIJI plugins and python for image restoration and segmentation	Python, plugin	https://csbdeep.bioimagecomputing.com/	Schmidt et al., 2018; Weigert et al., 2020
	Cellwise Classification, Neighborhoods	CytoMAP	Feature design for neighborhoods		Cell classification and neighborhood analysis with GUI	Application	gitlab.com/gernerlab/cytomap/-/wikis/home	Stoltzfus et al., 2020
		Volumetric tissue exploration and analysis	Feature design for classification and segmentation		Cell segmentation, classification and neighborhood analysis with GUI	Plugin	https://vtea.wiki	Winfree et al., 2017
	Segmentation	modelzoo.co		Models for many datatypes	Open source and pretrained networks	Web repository	modelzoo.co	NA
		InstantDL		Segmentation and classification	Broadly applicable segmentation and classification framework	Python, CoLab	https://github.com/marrlab/InstantDL/	Waibel et al., 2021
		BioImage.IO		Models for specifically for bioimaging	DL networks for the bioimaging community	Web repository	bioimage.io	NA
		ZeroCostDL4Mic		Training and inference with BioImage.IO	Tool for training and inference in the cloud	Cloud based, CoLab	github.com/HenriquesLab/ZeroCostDL4Mic	von Chamier et al., 2021
		OpSeF		DL network training and inference	Python framework in Jupyter notebooks	Python	github.com/trasse/OpSeF-IV	Rasse et al., 2020
		Weka	Extensive library of classifiers and tools		ML frame work for Java, Plugin for ImageJ	API, application, plugin	www.cs.waikato.ac.nz/ml/weka/	Hall et al., 2009

(Zhuang et al., 2021). Thus, until there is a generalizable solution, new deep learning segmentation approaches can be developed quickly by building on existing work with focused training datasets specific to tissue, cell-type and imaging modality.

CLASSIFICATION

Using specific protein or structural markers is a common way to determine cell-types in cytometry approaches like flow and image cytometry. Image cytometry is complicated by defining which pixels are associated with which cells. While a nuclear stain can be used to identify the nucleus, membrane, and cytoplasmic markers may be inconsistent across cell-types, cell-states, and tissues. A common solution is to measure markers in pixels proximal to segmented nuclei. These pixels can be defined by using a limited cell-associated region-of-interest that wraps around an existing nuclear segmentation or by performing a tessellation with a Voronoi segmentation (Winfree et al., 2017; Goltsev et al., 2018; McQuin et al., 2018).

The mean-fluorescence intensity (or other intensity measure, mode, upper-quartile mean, etc.) of markers in cell-associated segmented regions is frequently used for classification. A common supervised approach is to perform a series of sequential selections or gates based on marker intensities like flow cytometry. This "gating strategy" can easily identify specific cell-types with a predefined cell-type hierarchy. Cell classification can be semi-automated with unsupervised or semi-supervised machine learning using classifiers and clustering approaches. Popular approaches include Bayesian and Random Forest classifiers and clustering with k-means or graph based community clustering like the Louvain algorithm (Hall et al., 2009; Dao et al., 2016; McQuin et al., 2018; Phillip et al., 2021; Solorzano et al., 2021). Importantly, analyzing highly multiplexed image datasets, more than twenty markers, with a supervised "gating" approach may prove intractable necessitating machine learning approaches (Levine et al., 2015; Goltsev et al., 2018; Neumann et al., 2021).

Deep learning has been broadly applied to classification of images (Gupta et al., 2019). One of the strengths of a deep learning classification approach, as with segmentation, is that it is possible to start with a pretrained network-potentially reducing training set sizes. For instance, in 2D image classification, a convolutional neural network (CNN) like ResNet-50 initially trained on natural images (e.g., animals, vehicles, plants, etc.) can be retrained with a new label structure and training data that might include, for instance, cell nuclei (Woloshuk et al., 2021). Some deep learning models can further simplify workflows, like regional-CNNs, performing both segmentation and classification (Caicedo et al., 2019).

One image dataset that presents an interesting challenge and unique opportunity in both segmentation and classification is multiplexed fluorescence *in situ* hybridization (FISH). These approaches can, through combinatorial labeling of fluorophores, generate images of nearly all putative transcripts (Coskun and Cai, 2016). Although a semantic and instance segmentation approach can be used to identify and classify cells using associated FISH probes (Littman et al., 2021), a recent pixelwise-segmentation free approach has been proposed. This approach organizes the detected FISH-probes into spatial clusters using graphs from which signatures of cells and cell-types are determined (Shah et al., 2016; Andersson et al., 2020; Partel and Wählby, 2021).

MICROENVIRONMENTS AS NEIGHBORHOODS

The classification of microenvironments in tissues informs our understanding of the role of specific cells and structures in an underlying biology. This has led to the development of neighborhood analysis strategies that involve the segmentation of groups of cells or structures which are then classified with machine learning using neighborhood features such as cell-type census and location (Stoltzfus et al., 2020; Solorzano et al., 2021; Winfree et al., 2021). This process mirrors the segmentation and classification of single cells by protein and RNA markers where the types of cells or the distributions of cell types in neighborhoods are the markers used to classify the neighborhoods. The segmentation strategies for defining neighborhoods usually rely on either regular sampling of a tissue or cell centric approaches (e.g., distance from a cell or the *k*-nearest neighborhoods) (Jackson et al., 2020; Stoltzfus et al., 2020; Lake et al., 2021; Winfree et al., 2021). The impact of neighborhood size and defining it variably and locally (e.g., microenvironments may be different near arterioles vs. microvasculature) are under explored avenues in the analysis of cellular microenvironments in bioimaging datasets. Importantly, further development of neighborhoods analyses is critical as it has demonstrated mechanistic insight in human disease when used with highly multiplexed chemical and fluorescence imaging (Jackson et al., 2020; Schürch et al., 2020; Stoltzfus et al., 2021).

TECHNOLOGY AND TOOL ACCESSIBILITY

Minimizing the exclusivity of segmentation and classification advancements with the development of user accessible tools, is critical to the democratization of image analysis. In the above discussions of both segmentation and classification, most researchers and developers paid careful attention to providing tools for use by biomedical scientists. Example tools include web interfaces, stand-alone applications, or plugins for open-source image processing software (**Table 1**). These tools provide access to users that are novices in image analysis, day-to-day practitioners and super-users or developers across the three fundamental tasks of cell segmentation, cell classification and neighborhood analysis (**Table 1**). Furthermore, deep learning networks are available through online repositories such as github.com and modelzoo.co. An exciting development is the recent set of publications that have defined one-stop-shops for deep learning models and accessible tools for using and training existing deep learning networks (Iudin et al., 2016; Berg et al., 2019; Rasse et al., 2020; Gómez-de-Mariscal et al., 2021; von Chamier et al., 2021, p. 4). This includes the integration of segmentation tools

with online repositories of trained deep learning networks that can be easily downloaded and tested on cells and modalities of interest. With this added accessibility, there is a risk of misuse and possible abuse. However, the ease of reproducibility may outweigh this risk.

CONCLUSION

The bioimaging community has recognized for decades that image data is more than a picture. Mining imaging data collected in the biomedical sciences has blossomed in the past 20 years, pushed by advancements in multiplexed tissue labeling, image capture technologies, computational capacity, and machine learning. It will be exciting to see the next developments in image analysis with machine learning approaches. Perhaps we will witness: (1) a fully generalizable multidimensional cell segmentation approach; (2) novel approaches to cell-segmentation independent of pixelwise classification (as with some FISH data), or (3) new models of neighborhoods to characterize cellular microenvironments and niches. Furthermore, with web-based repositories to share datasets and tools that are suitable for all levels of expertise, these and other developments will be accessible to both experts, practitioners, *and* researchers new to imaging and image analysis. The broad accessibility of image data and tools could facilitate the adoption of common and rigorous processes for meaningful biological insight from image datasets across fields of study for so much more than a picture.

AUTHOR CONTRIBUTIONS

SW conceived and outlined the manuscript.

REFERENCES

Allam, M., Cai, S., and Coskun, A. F. (2020). Multiplex bioimaging of single-cell spatial profiles for precision cancer diagnostics and therapeutics. *NPJ Precis. Oncol*. 4:11. doi: 10.1038/s41698-020-0114-1

Andersson, A., Partel, G., Solorzano, L., and Wahlby, C. (2020). "Transcriptome-Supervised Classification of Tissue Morphology Using Deep Learning," in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, (Piscataway: IEEE), 1630–1633. doi: 10.1109/ISBI45749.2020.9098361

Bankhead, P., Loughrey, M. B., Fernández, J. A., Dombrowski, Y., and McArt, D. G. (2017). QuPath: open source software for digital pathology image analysis. *Sci. Rep*. 7:16878. doi: 10.1038/s41598-017-17204-5

Bannon, D., Moen, E., Schwartz, M., Borba, E., Kudo, T., Greenwald, N., et al. (2021). DeepCell Kiosk: scaling deep learning–enabled cellular image analysis with Kubernetes. *Nat. Methods* 18, 43–45. doi: 10.1038/s41592-020-01023-0

Berg, S., Kutra, D., Kroeger, T., Straehle, C. N., Kausler, B. X., Haubold, C., et al. (2019). ilastik: interactive machine learning for (bio)image analysis. *Nat. Methods* 16, 1226–1232. doi: 10.1038/s41592-019-0582-9

Berthold, M. R., Cebron, N., Dill, F., Gabriel, T. R., Kötter, T., Meinl, T., et al. (2007). "KNIME: The Konstanz Information Miner," in *Studies in Classification, Data Analysis, and Knowledge Organization*, eds C. Preisach, H. Burkhardt, L. Schmidt-Thieme, and R. Decker (Berlin: Springer).

Beucher, S., and Lantuejoul, C. (1979). "Use of Watersheds in Contour Detection," in *International Workshop on Image Processing, Real-Time Edge and Motion Detection*, (Rennes: Centre de Morphologie Mathématique). doi: 10.1016/s0893-6080(99)00105-7

Caicedo, J. C., Goodman, A., Karhohs, K. W., Cimini, B. A., Ackerman, J., Haghighi, M., et al. (2019). Nucleus segmentation across imaging experiments: the 2018 Data Science Bowl. *Nat. Methods* 16, 1247–1253. doi: 10.1038/s41592-019-0612-7

Ching, T., Himmelstein, D. S., Beaulieu-Jones, B. K., Kalinin, A. A., Do, B. T., Way, G. P., et al. (2018). Opportunities and obstacles for deep learning in biology and medicine. *J. R. Soc. Interface* 15:20170387. doi: 10.1098/rsif.2017.0387

Coskun, A. F., and Cai, L. (2016). Dense transcript profiling in single cells by image correlation decoding. *Nat. Methods* 13, 657–660. doi: 10.1038/nmeth.3895

Dao, D., Fraser, A. N., Hung, J., Ljosa, V., Singh, S., and Carpenter, A. E. (2016). CellProfiler Analyst: interactive data exploration, analysis and classification of large biological image sets. *Bioinformatics* 32, 3210–3212. doi: 10.1093/bioinformatics/btw390

de Chaumont, F., Dallongeville, S., Chenouard, N., Hervé, N., Pop, S., Provoost, T., et al. (2012). Icy: an open bioimage informatics platform for extended reproducible research. *Nat. Methods* 9, 690–696. doi: 10.1038/nmeth.2075

Dunn, K. W., Fu, C., Ho, D. J., Lee, S., Han, S., and Salama, P. (2019). DeepSynth: three-dimensional nuclear segmentation of biological images using neural networks trained with synthetic data. *Sci. Rep*. 9:18295. doi: 10.1038/s41598-019-54244-5

El-Achkar, T. M., Eadon, M. T., Menon, R., Lake, B. B., Sigdel, T. K., Alexandrov, T., et al. (2021). A multimodal and integrated approach to interrogate human kidney biopsies with rigor and reproducibility: guidelines from the Kidney Precision Medicine Project. *Physiol. Genomics* 53, 1–11. doi: 10.1152/physiolgenomics.00104.2020

Ellenberg, J., Swedlow, J. R., Barlow, M., Cook, C. E., Sarkans, U., Patwardhan, A., et al. (2018). A call for public archives for biological image data. *Nat. Methods* 15, 849–854. doi: 10.1038/s41592-018-0195-8

Falk, T., Mai, D., Bensch, R., Çiçek, Ö, Abdulkadir, A., and Marrakchi, Y. (2019). U-Net: deep learning for cell counting, detection, and morphometry. *Nat. Methods* 16, 67–70. doi: 10.1038/s41592-018-0261-2

Goltsev, Y., Samusik, N., Kennedy-Darling, J., Bhate, S., Hale, M., Vazquez, G., et al. (2018). Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging. *Cell* 174, 968–981.e15. doi: 10.1016/j.cell.2018.07.010

Gómez-de-Mariscal, E., García-López-de-Haro, C., Ouyang, W., Donati, L., Lundberg, E., Unser, M., et al. (2021). DeepImageJ: a user-friendly environment to run deep learning models in ImageJ. *Nat. Methods* 18, 1192–1195. doi: 10.1038/s41592-021-01262-9

Govind, D., Becker, J. U., Miecznikowski, J., Rosenberg, A. Z., Dang, J., Tharaux, P. L., et al. (2021). PodoSighter: a Cloud-Based Tool for Label-Free Podocyte Detection in Kidney Whole-Slide Images. *J. Am. Soc. Nephrol.* 32, 2795–2813. doi: 10.1681/ASN.2021050630

Gupta, A., Harrison, P. J., Wieslander, H., Pielawski, N., Kartasalo, K., Partel, G., et al. (2019). Deep Learning in Image Cytometry: a Review. *Cytometry A*. 95, 366–380. doi: 10.1002/cyto.a.23701

Haberl, M. G., Churas, C., Tindall, L., Boassa, D., Phan, S., Bushong, E. A., et al. (2018). CDeep3M—Plug-and-Play cloud-based deep learning for image segmentation. *Nat. Methods* 15, 677–680. doi: 10.1038/s41592-018-0106-z

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H. (2009). The WEKA data mining software: an update. *ACM SIGKDD Explor Newsl*. 11, 10–18. doi: 10.1145/1656274.1656278

Hollandi, R., Szkalisity, A., Toth, T., Tasnadi, E., Molnar, C., Mathe, B., et al. (2020). nucleAIzer: a Parameter-free Deep Learning Framework for Nucleus Segmentation Using Image Style Transfer. *Cell Syst*. 10, 453–458.e6. doi: 10.1016/j.cels.2020.04.003

Iudin, A., Korir, P. K., Salavert-Torres, J., Kleywegt, G. J., and Patwardhan, A. (2016). EMPIAR: a public archive for raw electron microscopy image data. *Nat. Methods* 13, 387–388. doi: 10.1038/nmeth.3806

Jackson, H. W., Fischer, J. R., Zanotelli, V. R. T., Ali, H. R., Mechera, R., Soysal, S. D., et al. (2020). The single-cell pathology landscape of breast cancer. *Nature* 578, 615–620. doi: 10.1038/s41586-019-1876-x

Kume, S., and Nishida, K. (2021). BioImageDbs: Bio- and biomedical imaging dataset for machine learning and deep learning (for ExperimentHub). R package version 1.2.2. Available online at: https://bioconductor.org/packages/release/data/experiment/html/BioImageDbs.html (accessed January 21, 2022).

Lake, B. B., Menon, R., Winfree, S., Qiwen, H., Ricardo Melo, F., Kian, K., et al. (2021). An Atlas of Healthy and Injured Cell States and

Niches in the Human Kidney. *bioRxiv* [Preprint]. doi: 10.1101/2021.07.28.454201

Lee, S., Amgad, M., Mobadersany, P., McCormick, M., Pollack, B. P., Elfandy, H., et al. (2021). Interactive Classification of Whole-Slide Imaging Data for Cancer Researchers. *Cancer Res*. 81, 1171–1177. doi: 10.1158/0008-5472.CAN-20-0668

Levine, J. H., Simonds, E. F., Bendall, S. C., Davis, K. L., Amir el, A. D., and Tadmor, M. D. (2015). Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell* 162, 184–197.

Littman, R., Hemminger, Z., Foreman, R., Arneson, D., Zhang, G., Gómez-Pinilla, F., et al. (2021). Joint cell segmentation and cell type annotation for spatial transcriptomics. *Mol. Syst. Biol.* 17:e10108. doi: 10.15252/msb.202010108

Ljosa, V., Sokolnicki, K. L., and Carpenter, A. E. (2012). Annotated high-throughput microscopy image sets for validation. *Nat. Methods* 9, 637–637. doi: 10.1038/nmeth.2083

Lutnick, B., Ginley, B., Govind, D., McGarry, S. D., LaViolette, P. S., Yacoub, R., et al. (2019). An integrated iterative annotation technique for easing neural network training in medical image analysis. *Nat. Mach. Intell.* 1, 112–119. doi: 10.1038/s42256-019-0018-3

McQuin, C., Goodman, A., Chernyshev, V., Kamentsky, L., Cimini, B. A., Karhohs, K. W., et al. (2018). CellProfiler 3.0: next-generation image processing for biology. *PLoS Biol*. 16:e2005970. doi: 10.1371/journal.pbio.2005970

Meijering, E. (2012). Cell Segmentation: 50 Years Down the Road [Life Sciences]. *IEEE Signal Process. Mag.* 29, 140–145. doi: 10.1109/MSP.2012.2204190

Meijering, E. (2020). A bird's-eye view of deep learning in bioimage analysis. *Comput. Struct. Biotechnol. J.* 18, 2312–2325. doi: 10.1016/j.csbj.2020.08.003

Meijering, E., Carpenter, A. E., Peng, H., Hamprecht, F. A., and Olivo-Marin, J. C. (2016). Imagining the future of bioimage analysis. *Nat. Biotechnol*. 34, 1250–1255. doi: 10.1038/nbt.3722

Minaee, S., Boykov, Y. Y., Porikli, F., Plaza, A. J., Kehtarnavaz, N., and Terzopoulos, D. (2021). Image Segmentation Using Deep Learning: a Survey. *IEEE Trans. Pattern Anal. Mach. Intell.* 1–1. [Online ahead of print] doi: 10.1109/TPAMI.2021.3059968

Moen, E., Bannon, D., Kudo, T., Graf, W., Covert, M., and Van Valen, D. (2019). Deep learning for cellular image analysis. *Nat. Methods* 16, 1233–1246. doi: 10.1038/s41592-019-0403-1

Moshkov, N., Mathe, B., Kertesz-Farkas, A., Hollandi, R., and Horvath, P. (2020). Test-time augmentation for deep learning-based cell segmentation on microscopy images. *Sci. Rep*. 10:5068. doi: 10.1038/s41598-020-61808-3

Neumann, E. K., Patterson, N. H., Allen, J. L., Migas, L. G., Yang, H., Brewer, M., et al. (2021). Protocol for multimodal analysis of human kidney tissue by imaging mass spectrometry and CODEX multiplexed immunofluorescence. *STAR Protoc*. 2:100747. doi: 10.1016/j.xpro.2021.100747

Otsu, N. A. (1979). Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man Cybern*. 9, 62–66. doi: 10.1109/TSMC.1979.4310076

Partel, G., and Wählby, C. (2021). Spage2vec: unsupervised representation of localized spatial gene expression signatures. *FEBS J*. 288, 1859–1870. doi: 10.1111/febs.15572

Phillip, J. M., Han, K. S., Chen, W. C., Wirtz, D., and Wu, P. H. (2021). A robust unsupervised machine-learning method to quantify the morphological heterogeneity of cells and nuclei. *Nat. Protoc*. 16, 754–774. doi: 10.1038/s41596-020-00432-x

Rasse, T. M., Hollandi, R., and Horvath, P. (2020). OpSeF: open Source Python Framework for Collaborative Instance Segmentation of Bioimages. *Front. Bioeng. Biotechnol*. 8:558880. doi: 10.3389/fbioe.2020.558880

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682. doi: 10.1038/nmeth.2019

Schmidt, U., Weigert, M., Broaddus, C., and Myers, G. (2018). "Cell Detection with Star-Convex Polygons," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. Lecture Notes in Computer Science*, eds A. F. Frangi, J. A. Schnabel, C. Davatzikos, C. Alberola-López, and G. Fichtinger (Berlin: Springer International Publishing), 265–273. doi: 10.1007/978-3-030-00934-2_30

Schürch, C. M., Bhate, S. S., Barlow, G. L., Phillips, D. J., Noti, L., Zlobec, I., et al. (2020). Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front. *Cell* 182, 1341–1359.e19. doi: 10.1016/j.cell.2020.07.005

Shah, S., Lubeck, E., Zhou, W., and Cai, L. (2016). In Situ Transcription Profiling of Single Cells Reveals Spatial Organization of Cells in the Mouse Hippocampus. *Neuron* 92, 342–357. doi: 10.1016/j.neuron.2016.10.001

Solorzano, L., Wik, L., Olsson Bontell, T., Wang, Y., Klemm, A. H., Öfverstedt, J., et al. (2021). Machine learning for cell classification and neighborhood analysis in glioma tissue. *Cytometry A* 99, 1176–1186. doi: 10.1002/cyto.a.24467

Stoltzfus, C. R., Filipek, J., Gern, B. H., Olin, B. E., Leal, J. M., Wu, Y., et al. (2020). CytoMAP: a Spatial Analysis Toolbox Reveals Features of Myeloid Cell Organization in Lymphoid Tissues. *Cell Rep*. 31:107523. doi: 10.1016/j.celrep.2020.107523

Stoltzfus, C. R., Sivakumar, R., Kunz, L., Olin Pope, B. E., Menietti, E., Speziale, D., et al. (2021). Multi-Parameter Quantitative Imaging of Tumor Microenvironments Reveals Perivascular Immune Niches Associated With Anti-Tumor Immunity. *Front. Immunol*. 12:726492. doi: 10.3389/fimmu.2021.726492

Stringer, C., Wang, T., Michaelos, M., and Pachitariu, M. (2021). Cellpose: a generalist algorithm for cellular segmentation. *Nat. Methods* 18, 100–106. doi: 10.1038/s41592-020-01018-x

Sullivan, D. P., Winsnes, C. F., Åkesson, L., Hjelmare, M., Wiking, M., Schutten, R., et al. (2018). Deep learning is combined with massive-scale citizen science to improve large-scale image classification. *Nat. Biotechnol*. 36, 820–828. doi: 10.1038/nbt.4225

von Chamier, L., Laine, R. F., Jukkala, J., Spahn, C., Krentzel, D., Nehme, E., et al. (2021). Democratising deep learning for microscopy with ZeroCostDL4Mic. *Nat. Commun*. 12:2276. doi: 10.1038/s41467-021-22518-0

Waibel, D. J. E., Shetab Boushehri, S., and Marr, C. (2021). InstantDL - An easy-to-use deep learning pipeline for image segmentation and classification. *BMC Bioinformatics* 22:103. doi: 10.1186/s12859-021-04037-3

Weigert, M., Schmidt, U., Haase, R., Sugawara, K., and Myers, G. (2020). "Star-convex Polyhedra for 3D Object Detection and Segmentation in Microscopy," in *The IEEE Winter Conference on Applications of Computer Vision (WACV)*, (Snowmass: IEEE). doi: 10.1109/WACV45572.2020.9093435

Williams, E., Moore, J., Li, S. W., Rustici, G., Tarkowska, A., Chessel, A., et al. (2017). Image Data Resource: a bioimage data integration and publication platform. *Nat. Methods* 14, 775–781. doi: 10.1038/nmeth.4326

Winfree, S., Al Hasan, M., and El-Achkar, T. M. (2021). Profiling immune cells in the kidney using tissue cytometry and machine learning. *Kidney360* 10.34067/KID.0006802020. [Online ahead of Print] doi: 10.34067/KID.0006802020

Winfree, S., Khan, S., Micanovic, R., Eadon, M. T., Kelly, K. J., Sutton, T. A., et al. (2017). Quantitative Three-Dimensional Tissue Cytometry to Study Kidney Tissue and Resident Immune Cells. *J. Am. Soc. Nephrol.* 28, 2108–2118. doi: 10.1681/ASN.2016091027

Woloshuk, A., Khochare, S., Almulhim, A. F., McNutt, A. T., Dean, D., Barwinska, D., et al. (2021). In Situ Classification of Cell Types in Human Kidney Tissue Using 3D Nuclear Staining. *Cytometry A*. 99, 707–721. doi: 10.1002/cyto.a.24274

Wu, L., Han, S., Chen, A., Salama, P., Dunn, K. W., and Delp, E. J. (2021). RCNN-SliceNet: a Slice and Cluster Approach for Nuclei Centroid Detection in Three-Dimensional Fluorescence Microscopy Images. *ArXiv* [Preprint]. Available online at: http://arxiv.org/abs/2106.15753 (Accessed October 15, 2021).

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., et al. (2021). A Comprehensive Survey on Transfer Learning. *Proc. IEEE* 109, 43–76. doi: 10.1109/JPROC.2020.3004555

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Physiology

REVIEW
published: 24 March 2022
doi: 10.3389/fphys.2022.827280

Intravital Multiphoton Microscopy as a Tool for Studying Renal Physiology, Pathophysiology and Therapeutics

Bruce A. Molitoris, Ruben M. Sandoval and Mark C. Wagner*

Indiana Center for Biological Microscopy, Indiana University School of Medicine, Indianapolis, IN, United States

Intravital multiphoton microscopy has empowered investigators to study dynamic cell and subcellular processes *in vivo* within normal and disease organs. Advances in hardware, software, optics, transgenics and fluorescent probe design and development have enabled new quantitative approaches to create a disruptive technology pioneering advances in understanding of normal biology, disease pathophysiology and therapies. Offering superior spatial and temporal resolution with high sensitivity, investigators can follow multiple processes simultaneously and observe complex interactions between different cell types, intracellular organelles, proteins and track molecules for cellular uptake, intracellular trafficking, and metabolism in a cell specific fashion. The technique has been utilized in the kidney to quantify multiple dynamic processes including capillary flow, permeability, glomerular function, proximal tubule processes and determine the effects of diseases and therapeutic mechanisms. Limitations include the depth of tissue penetration with loss of sensitivity and resolution due to scattered emitted light. Tissue clearing technology has virtually eliminated penetration issues for fixed tissue studies. Use of multiphoton microscopy in preclinical animal models offers distinct advantages resulting in new insights into physiologic processes and the pathophysiology and treatment of diseases.

Keywords: proximal tubule, glomerular filtration, endocytosis, renal hemodynamics, fluorescent biomarkers

OPEN ACCESS

Edited by:
Grazia Tamma,
University of Bari Aldo Moro, Italy

Reviewed by:
David M. Small,
Cornell University, United States
Ann Seynhaeve,
Erasmus University Medical Center, Netherlands

****Correspondence:***
Bruce A. Molitoris
bmolitor@iu.edu

Specialty section:
This article was submitted to Renal and Epithelial Physiology, a section of the journal Frontiers in Physiology

Received: *01 December 2021*
Accepted: *09 February 2022*
Published: *24 March 2022*

Citation:
Molitoris BA, Sandoval RM and Wagner MC (2022) Intravital Multiphoton Microscopy as a Tool for Studying Renal Physiology, Pathophysiology and Therapeutics. Front. Physiol. 13:827280. doi: 10.3389/fphys.2022.827280

INTRODUCTION

Intravital multiphoton microscopy (MPM) of the kidney has been conducted for 20 years (Dunn et al., 2002, 2021). During this time advances in optics, lasers, computer software and hardware have led to more powerful systems having improvements in sensitivity and speed leading to a wide variety of new techniques exploring questions *in vivo* that were before unapproachable. Intravital multi-photon microscopy allows for visualization and quantification of dynamic cellular processes in normal functioning and diseased cells *in vivo*. A wealth of fluorescent biomarkers utilizing, blue, green, red, and far-red emitting fluorophores now allow four channels to be viewed simultaneously in three dimensions (3D) over time resulting in four-dimensional data. This has markedly increased the ability to observe and relate events involving multiple cell types and or intracellular organelles. Several laboratories have pioneered approaches and taken advantage of the many of these technological advances to study kidney physiology and pathophysiology (Dunn et al., 2002). Our laboratory has been aided along the way by numerous scientific collaborations and a NIH supported O'Brien Center for the past 20 years (Dunn et al., 2021). **Table 1** lists some of the processes that can be quantified. In particular, the ability to study, within the same nephron, the interdependent roles of the glomerulus and proximal tubule (PT) simultaneously

has been an exciting development for our laboratory. This mini-review will highlight a number of the advantages, techniques developed and utilized to quantify various aspects of renal physiology, pathophysiology and drug therapies, and will end indicating some of the existing limitations and challenges to the field.

We will start by giving an overview of the imaging set up used, and some basics of what can be visualized and then proceed to individual structural components of the nephron (**Figure 1A**). The set up developed and utilized by our laboratory is shown in **Figure 1B**. We prefer the inverted microscope as experience has taught us that we can limit motion more thoroughly using this approach. Maintaining body temperature, volume status and appropriate anesthesia are essential to a successful study. We usually have an IV infusion ongoing and also measure the arterial blood pressure with a transducer to insure adequate hydration and physiologic parameters. Maintaining adequate anesthesia, but not too much, minimizes movement which is essential. We prefer inhaled anesthetics as they allow for fine tuning of the state of anesthesia. **Figure 1C** shows a low power view of the outer cortex of a Munich Wistar Frömter rat revealing a surface glomerulus surrounded by numerous tubules. Tubule types such as proximal tubules are identified by their endogenous autofluorescence and apical brush border membrane. Collecting ducts and distal tubules are indistinguishable from each other as they lack endogenous autofluorescence or any other visible landmark and appear as large empty patches similar in size to proximal tubules. Surrounding the tubules is the interstitial space containing dendritic cells and other cell types, especially during and following injury, and a network of peritubular capillaries and erythrocytes and white blood cells appearing as dark objects as they do not take up the fluorescent molecule contained in the plasma. Large molecular weight fluorescent molecules, that remain stable in the vasculature, are used to demarcate vessels, evaluate permeability and localize the interstitial space (**Figure 1C**). **Figure 1** also shows a high magnification micrograph of a shallow 5 μm, 3D reconstruction of proximal tubule from the same series in **Figure 1C**. A 10 kDa filtered blue dextran is seen in early endosomes in the sub-apical region. The techniques and probes to be described have allowed us to understand normal renal physiology, the effect and pathophysiology of disease processes and the mechanisms of effective therapies.

RENAL BLOOD FLOW DYNAMICS

Intravital MPM reveals a heterogeneous landscape of normal red blood cell flow within the peritubular vasculature and glomerular capillary loops. It has allowed important insights into ischemic and septic injury to the microvasculature. Large molecular weight fluorescent molecules create shadows of blood cells in the vasculature and the velocity of these cells can be inferred from the angles of these streaks in 2D images, or more accurately from the angle of the streaks in kymographs derived from line scans. This allows for assessment of red blood cell (RBC) flow rates and the degree of white blood cell rolling and attachment following ischemic injury and during sepsis (Dunn et al., 2002, 2021; Molitoris and Sandoval, 2005, 2011; Sharfuddin et al., 2009; Sandoval and Molitoris, 2017; Sandoval et al., 2019). In disease states red blood cells can stack together to form rouleaux reducing their oxygen delivery capacity and resulting in partial or complete peritubular capillary obstruction. These structures are easily identified, often lodged behind adherent white blood cells in the microvasculature, and can exit the kidney in the venous outflow (Sharfuddin et al., 2009; Molitoris and Sandoval, 2011; Sharfuddin and Molitoris, 2011). They may lodge in other microvascular beds in distant organs but the importance of this has not been determined. This has been used extensively to visualize and quantify the changes in peritubular capillary blood flow rates and microvascular dropout following ischemic injury (Basile, 2019).

Labeling White Blood Cells (WBC) nuclei using Hoechst 33342, and using distinctive nuclear morphology, gives a qualitative idea of the number and types of WBC flowing freely within the renal vasculature or found within the interstitium. In disease or injury models activated WBCs adhere to or roll along endothelial cells reducing flow (Sharfuddin et al., 2009; Molitoris and Sandoval, 2011; Sharfuddin and Molitoris, 2011). This can be visualized in the peritubular capillaries of S1 and S2 PT segments. Unfortunately, due to limited depth penetration, the S3 segment of PT cannot be visualized using intravital MPM. The S3 PT nephron segment is known to suffer the greatest capillary injury and microvascular dropout in ischemic models (Sharfuddin and Molitoris, 2011; Basile, 2019).

Ratiometric imaging of two non-overlapping fluorescent vascular dyes has been used in pre-clinical studies to determine glomerular filtration rates in rats under physiologic and following ischemic acute kidney injury (Yu et al., 2005; Wang et al., 2010). The glomerular sieving coefficient (GSC) of a fluorescent compound is the ratio of fluorescence in Bowman's Space divided by fluorescence in the glomerular capillary plasma. A small 5 kDa dextran, with a GSC of 1.0 is rapidly and completely filtered across glomerular capillaries, and a large dextran, 150 kDa, has a very low GSC and is retained and stable in the vasculature. This approach has been adapted to clinical studies to provide both quantitative GFR and plasma volume determinations (Rizk et al., 2018; Molitoris et al., 2019).

GLOMERULAR IMAGING

In MWF rats surface glomeruli are easily identified allowing the dynamic aspects of glomerular capillary vessel diameter, RBC flow rates, single nephron GFR, and permeability to be quantified (Dunn et al., 2002, 2021; Sandoval et al., 2019). Of these different parameters the measurement of glomerular permeability of macromolecules is likely the most clinically important and has created controversy in the literature. Previous methods to quantify glomerular permeability were based on micropuncture or urinary fractional clearance studies. These techniques compare tubular lumen filtrate and urinary concentrations to plasma concentrations, respectively. Unfortunately, there is no allowance for PT mediated loss of material from the lumen due to tubular

reabsorption via fluid phase or receptor mediated endocytosis prior to the collection location (Russo et al., 2007a,b, 2009; Sandoval et al., 2012; Sandoval and Molitoris, 2013; Wagner et al., 2016a). Since the early S1 segment is primarily responsible for albumin reabsorption, micropuncture studies often miss the most endocytic S1 portion thus underestimating the amount of filtered albumin due to the needle placement away from Bowman's Space. Thus, our MPM studies have shown the glomerular sieving coefficients (GSCa) higher than most micropuncture studies, in the 0.012–0.015 range (Sandoval et al., 2012; Sandoval and Molitoris, 2013), while micropuncture studies have for the most part found values in the 0.0005 range for MWF rats. A recent micropuncture study did show a much higher GSCa (Hu et al., 2016). Russo et al. (2007b, 2009) used Munich Wistar Simonsen rats and found their GSCa was in the range of 0.025–0.030. The reason for the GSCa differences in MWF and MWS has not been investigated, but we have validated the high GSCa in MWS rats. Interestingly, early streptozocin diabetic Munich

TABLE 1 | Potential uses of multiphoton microscopes in kidney processes.

Dynamic cell specific events:	References
Cellular labeling and uptake	
Cell type specific:	
Epithelial	Tanner et al., 2005; Ashworth et al., 2007; Sandoval and Molitoris, 2017; Dunn et al., 2021
Endothelial	Dunn et al., 2002; Sutton et al., 2003; Molitoris and Sandoval, 2011; Desposito et al., 2021; Gyarmati et al., 2021a
Glomerular labeling	Hackl et al., 2013; Schiessl et al., 2020; Gyarmati et al., 2021a
Uptake site:	
Apical	Dunn et al., 2002, 2021; Sandoval et al., 2004, 2012; Sandoval and Molitoris, 2017
Basolateral	Horbelt et al., 2007
Mechanism:	
Endocytosis	Dunn et al., 2002; Sandoval et al., 2004, 2019; Kalakeche et al., 2011; Wagner et al., 2016a; Sandoval and Molitoris, 2017; Schuh et al., 2018
Carrier/transporter mediated	Horbelt et al., 2007
Cell number	Hackl et al., 2013; Schiessl et al., 2020; Gyarmati et al., 2021b
Pattern distribution	Hackl et al., 2013; Schiessl et al., 2020; Shroff et al., 2021
Cellular distribution	
Site specific intracellular organelle accumulation	Weinberg and Molitoris, 2009; Hall et al., 2013
Cytosol accumulation	Molitoris and Sandoval, 2006
Cell function	
Endocytosis-quantitative analysis	Sandoval and Molitoris, 2008, 2017; Schuh et al., 2018; Sandoval et al., 2019
Intracellular trafficking	Sandoval et al., 2004, 2012; Molitoris and Sandoval, 2006; Sandoval and Molitoris, 2017
Transcytosis/exocytosis	Sandoval et al., 2012
Renin secretion	Schiessl et al., 2020
Dynamic structural/functional effects within the kidney:	
Glomerular:	
Size/volume	Sandoval et al., 2019
Permeability	Russo et al., 2007b; Nakano et al., 2012; Sandoval et al., 2012, 2014, 2019; Sandoval and Molitoris, 2013, 2014, 2017; Schiessl and Castrop, 2013; Dickson et al., 2014; Schiessl et al., 2015; Wagner et al., 2016a,b; Kidokoro et al., 2019; Gyarmati et al., 2021a
Fibrosis/Sclerosis	Ranjit et al., 2016
snGFR	Kang et al., 2006; Kidokoro et al., 2019
Afferent arteriole	Kang et al., 2006; Hackl et al., 2013; Gyarmati et al., 2021b
Macula densa	Shroff et al., 2021
Microvasculature:	
Blood flow rate	Molitoris and Sandoval, 2005; Sharfuddin et al., 2009; Sandoval and Molitoris, 2017; Sandoval et al., 2019
Endothelial permeability	Molitoris and Sandoval, 2011; Sandoval and Molitoris, 2017; Sandoval et al., 2019
WBC adherence/rolling/tissue invasion	Sandoval and Molitoris, 2017; Sandoval et al., 2019
Vasoconstriction	Kidokoro et al., 2019
Epithelial cell:	
Cell injury in necrosis, apoptosis	Dunn et al., 2002; Kelly et al., 2003; Ashworth et al., 2007; Kalakeche et al., 2011; Sandoval and Molitoris, 2017
Surface membrane/blebbing	Tanner et al., 2005; Ashworth et al., 2007
Tubular flow	Ferrell et al., 2015

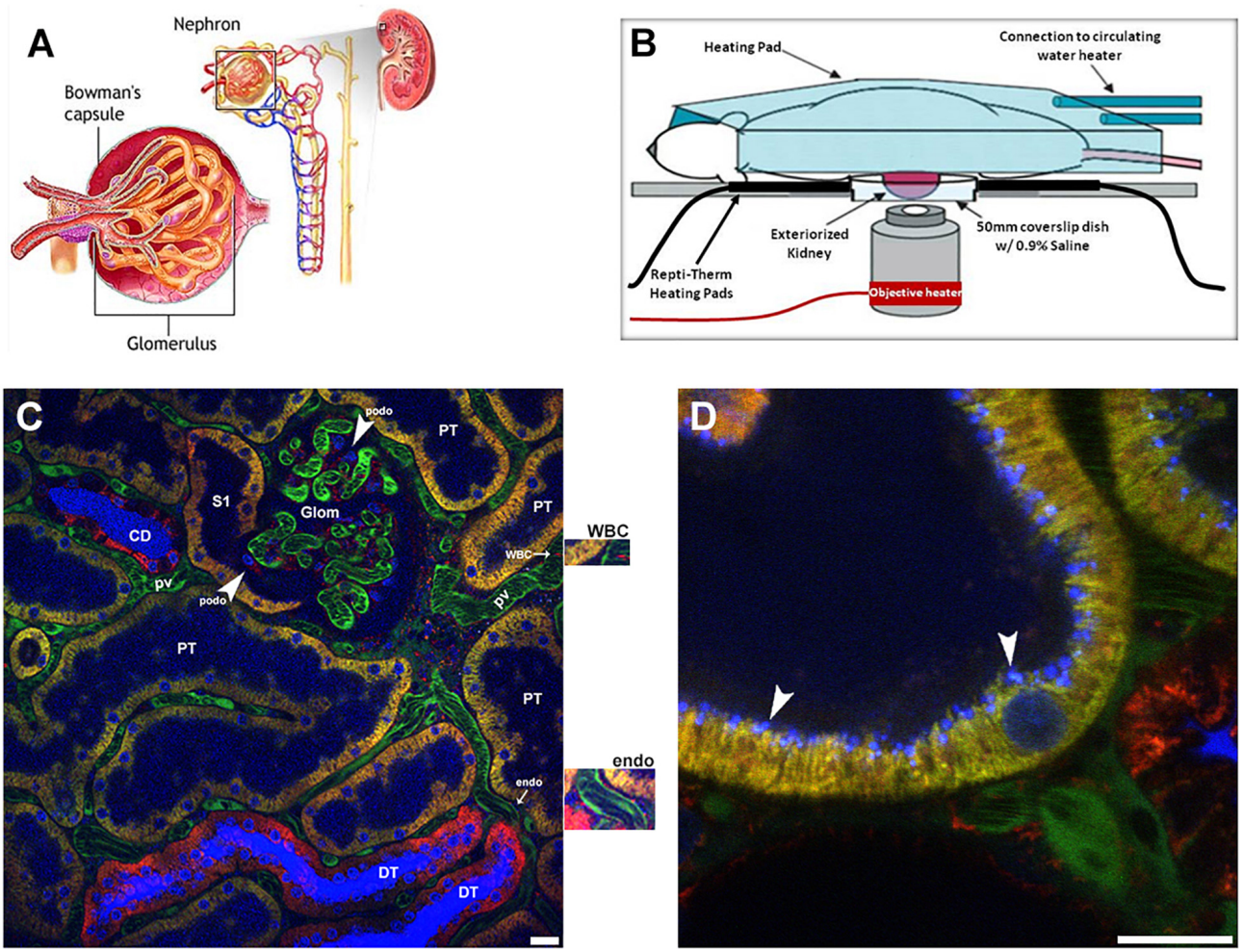


FIGURE 1 | Visual resolving power of intravital multi-photon microscopy: a schematic of the renal architecture to subcellular resolution in proximal tubules. **(A)** Shows a classic diagram of the kidney in cross-section (upper right), with an inset of the complete nephron (center), and glomerulus (left). **(B)** Shows a schematic of the set-up used for 2-photon imaging. Placing the left exteriorized kidney onto a coverslip bottom dish on an inverted microscope is the most efficient way to minimize motion artifact from breathing. The various heating elements shown are used to maintain the proper kidney and body temperature, which is monitored and regulated. **(C)** Shows a single plane cross section with a glomerulus (glom) in the upper center and surrounding peritubular vasculature made visible using a large 150 kDa FITC dextran (green) retained in the plasma. Faint blue fluorescence in proximal tubule lumens (PT) and intense blue fluorescence in the lumen of collecting ducts (CD) and distal tubules (DT) comes from a small rapidly filtered 10 kDa Cascade Blue dextran bolus administered earlier. In DT and CD segments water removal concentrates the dextran intensifying the color. Three different mitochondrial dyes are seen in this image. Rhodamine 123 (yellow) predominantly labels the mitochondria of proximal tubule cells (PT); the open S1 PT segment is seen juxtapose to the glomerulus on the left. Tetramethylrhodamine methyl ester (TMRM, red) predominantly labels mitochondria in the distal tubules (DT) and collecting ducts (CD). Note the heterogenous labeling seen in the collecting ducts between the intercalated and principal cells. Finally, rhodamine B hexyl ester (also red) labels circulating white blood cells (WBC), podocytes (podo) in the glomerulus, endothelial cells (endo) see surrounding the peritubular vasculature, and a variety of cells in the interstitial space. Enlargements for the white blood cell and endothelial cell staining with rhodamine B hexyl ester are immediately adjacent to the right of the image. Hoechst 33342 labels the nuclei of all cell types (blue-cyan) in this micrograph (Bar = 20 μm). **(D)** Shows a high magnification micrograph of a shallow 5 μm, 3D reconstruction of proximal tubule from the same series in **(C)**. Endocytosis of the 10 kDa Cascade Blue dextran by the proximal tubule is seen accumulating at the sub-apical space in discrete blue vesicles (arrowheads). Note the various sizes of endosomes the microscope can resolve (Bar = 20 μm).

Wistar Simonsen rats had no change in their GSCa, but showed reduced PT reabsorption of filtered albumin and increased urinary albumin (Russo et al., 2009).

Unfortunately, glomerular studies in mice are challenging as glomeruli are rarely within 100 microns of the surface after 4 weeks of age in all mice strains (Schiessl et al., 2013). To circumvent this challenge prolonged ureteral obstruction, or the use of non-steroidal anti-inflammatory agents, have been used to force glomeruli to the surface secondary to cortical destruction. However, this approach is known to cause excessive inflammation, fibrosis, and loss of proximal tubule (PT) structure and function leaving interpretation of the results problematic especially when studying a disease model (Chevalier et al., 2009; Yang et al., 2010).

To get accurate GSCa using MPM one must have the necessary sensitivity to correct for measurement and subtraction of background values (Yang et al., 2010; Sandoval et al., 2014; Sandoval and Molitoris, 2014). Setting the background too high lessens sensitivity and reduces the GSCa. This requires using the full dynamic range of the system's detectors and particularly the correct offset or black level. Setting the black level for Bowman's Space to zero to eliminate all background signal, markedly reduces detector sensitivity and distorts the results of the low intensity signals (Sandoval et al., 2014; Sandoval and Molitoris, 2014). The old adage, your results are only as good as your sensitivity holds in this situation.

PROXIMAL TUBULE ENDOCYTOSIS AND TRANSCYTOSIS

Proximal tubules function to reabsorb filtered fluid, electrolytes and macromolecules to prevent loss via urinary excretion. They also function to "sense" the internal environment and have immunologic surveillance capabilities (Hato et al., 2013). Intravital MPM has played an important role in understanding the processes involved, intracellular trafficking and catabolism of the reabsorbed material (Molitoris and Sandoval, 2005; Horbelt et al., 2007; Sandoval and Molitoris, 2008, 2017; Sandoval et al., 2019). This has been particularly important for macromolecules including therapeutic agents (Sandoval et al., 2004; Molitoris and Sandoval, 2006; Molitoris et al., 2009; Kalakeche et al., 2011). Kinetic studies with p53 fluorescent siRNA showed PT endocytosis, cytosolic delivery and a short intracellular half-life corresponding to the rate and duration of the synthesis of p53 (Molitoris et al., 2009).

Macromolecule reabsorption across the apical membrane occurs via receptor mediated and fluid phase endocytosis, **Figure 1D**. The Hall laboratory, using a tissue clearing technique to allow for deeper MPM penetration in fixed tissue showed the S1 segment uses receptor mediated endocytosis (RME) primarily whereas the S2 and S3 segments primarily use the fluid phase endocytosis (FPE; Schuh et al., 2018). Thereafter endocytic trafficking sorts material into two main pathways, lysosomal for catabolism and transcytosis for reclamation. Intravital MPM has helped expand the investigative focus, beyond glomerular dysfunction, to elucidate the role tubular injury plays in proteinuric and albuminuric diseases previously thought to be associated solely with damage to the filtration barrier (Sandoval et al., 2012; Sandoval and Molitoris, 2013; Wagner et al., 2016a). In quantifying uptake it is important not to saturate the intensity of the endosomal pool (particularly lysosomes) as this will underestimate the amount of the material therein (Sandoval and Molitoris, 2008, 2014; Sandoval et al., 2014).

Careful consideration to the background fluorescence must also be accounted for when quantifying uptake of any compound into the lysosomal/endosomal pool. This value must be subtracted from the raw images to determine true and meaningful intensity values (Sandoval et al., 2019). We typically take three background 3D volumes of proximal tubules at different laser transmissivities prior to imaging, and calculate average intensity values at each laser power to mathematically compensate for saturating intensities. This generates intensity correction factors between the different laser powers used to normalize background subtracted images taken at different laser powers. Thresholding is used to help correct for autofluorescence in lysosomes, and is fluorophore channel specific (Sandoval and Molitoris, 2013; Sandoval et al., 2019).

Transcytosis has not been extensively studied in PT cells *in vivo* because it is difficult to characterize the amount of transcytosis based on basolateral transport into the interstitial space (Sandoval et al., 2012; Dickson et al., 2014). We observed albumin transcytosis via both finger like projections and vesicles from endosomal accumulations reaching basolateral membranes of proximal tubules (Sandoval et al., 2012). This is in agreement with FcRn mediated immunoglobulin transcytosis in cultured cells (Ward et al., 2005). Transcytosis of albumin was confirmed using molecular techniques, but the amount of albumin undergoing transcytosis remains unknown (Tenten et al., 2013). A potential role for PT sorting of glycated, carbamylated and other potentially toxic albumins, mediated by FcRn binding, for catabolism via lysosomal trafficking has been proposed as a mechanism to rid the body of altered albumins while preserving physiologic albumin for transcytosis (Dickson et al., 2014; Wagner et al., 2016b; Yadav et al., 2021). Proximal tubule transcytosis of folic acid and other vitamins is known to occur (Sandoval et al., 2004). Transcytosis from the basolateral membrane has also been demonstrated for PT cells using other techniques (Hu et al., 2016).

MITOCHONDRIAL STRUCTURE AND FUNCTION AND ASSOCIATED PROCESSES

Mitochondrial structure and function can both be studied using intravital MPM in multiple cell types simultaneously. Multiple cell membrane permeable dyes can be used to determine the mitochondrial potential and follow its loss during injury (Weinberg and Molitoris, 2009; Hall et al., 2013). These studies identified the relative resistance in cellular mitochondrial potential to ischemia among the different tubular epithelial cells and structural changes in PT mitochondria. Three different mitochondrial dyes are used to label various cortical cells (**Figure 1C**). Rhodamine 123 predominantly labels proximal tubule cells, Tetramethylrhodamine methyl ester (TMRM), labels the collecting ducts and distal tubule cells. However, increased loading concentrations can cause accumulation in other tubule types. A second red dye, Rhodamine B hexyl ester, is used to stain mitochondria of endothelial cells, podocytes, circulating white blood cells, and cells within the interstitial space. All three dyes can be used simultaneously at lower loading concentrations, even the two red dyes because of the disparate cell types they label (Hall et al., 2013). The differences in cells labeling by each dye may relate to the organic ion transport processes in each cell type.

Apoptosis is another intracellular process that can easily be followed and quantified using MPM and can be differentiated from necrosis using Hoechst 33342 and the vital dye propidium

iodide (Dunn et al., 2002, 2021; Kelly et al., 2003). Bright condensed staining along the edge of the nucleus, as well as bright fragmented structures, are hallmark changes that occur during apoptosis. Staining of nuclei with propidium iodide is indicative of a necrotic cell with a compromised cell membrane as this dye is membrane impermeant.

ALTERNATIVE METHODS OF PROBE DELIVERY

While intravascular delivery remains the mainstay for delivering fluorescent biomarkers to the kidney, a major advantage of MPM is the ability to pair other techniques with it and observe the subsequent process *in vivo* using a biomarker that cannot be delivered via the vascular route. We have used micropuncture techniques to deliver plasmids to fluorescently label cellular actin structures (Tanner et al., 2005; Ashworth et al., 2007), deliver fluorescently labeled bacteria to the lumen of PT cells to follow growth, invasion and cellular responses of PT (Mansson et al., 2007; Melican et al., 2011; Choong et al., 2012; Schulz et al., 2018), to endothelium and WBC (Sutton et al., 2003; Molitoris and Sandoval, 2011) and hydrodynamic delivery of genes to cells throughout the kidney (Kolb et al., 2018). These techniques are done just prior to imaging the animal or on the microscope stage during imaging. The ability to follow the result in a small area of cells eliminates the need to deliver the probe to the entire kidney.

CHALLENGES TO STUDYING THE KIDNEY

Imaging the kidney intravitally has a number of challenges that must be understood and minimized. The kidney has reduced optical penetration, compared to the many tissues, due to increased blood flow, cellular heterogeneity, and inherent autofluorescence. This results in scattering and absorption of the emitted light. This limits the depth of penetration allowing for high sensitivity and resolution to less than 100 microns whereas studies in brain can penetrate over a millimeter (Sandoval et al., 2012; Sandoval and Molitoris, 2013). **Figure 2** shows the effect of imaging depth on sensitivity and resolution from 20 to 70 μm utilizing two different wavelengths, 880 and 890 nm, even when using the Linear Z-Compensation feature on the Leica Dive Multi-Photon system. Orthogonal views, **Figures 2A,B** and single plane images **Figures 2C–F** show the drop off in both sensitivity and resolution regardless of the wavelength used. As we have shown before, this drop off is greater for fluorophores emitting in the green spectrum due to enhanced quenching by hemoglobin (Sandoval et al., 2012; Sandoval and Molitoris, 2013).

Two approaches have recently been advanced to allow for deeper penetration and visualization. Schuh et al. (2016) using specialized longer wavelength excitation lasers and far-red probes, demonstrated greater depth advantages when conducting intravital 2 and 3-photon studies of the kidney. Adaptive optics may also be able to improve the depth of penetration by compensating for system and sample aberrations in the excitation beam wavefront. This will improve the focus resulting in higher intensities and improved spatial confinement at depth (Ji, 2017). However, adaptive optics has not been applied to imaging the kidney.

It is also more difficult to stabilize the kidney leading to increased motion artifacts. These challenges and approaches to minimize them have been carefully described previously (Dunn et al., 2002, 2021; Sandoval and Molitoris, 2017; Sandoval et al., 2019).

The use of fluorescent probes or biomarkers to delineate aspects of glomerular filtration, peritubular capillary function and tubular function in health and disease is critical but not without challenges. For instance, commercially available fluorescent dextrans all too often have a wide molecular weight dispersion limiting their accurate characterization of processes such as glomerular permeability (Sandoval et al., 2012; Sandoval and Molitoris, 2013). We have solved this problem by first obtaining a highly uniform dextran, with low MW dispersion, and second by performing the fluorophore conjugation directly (Sandoval et al., 2012; Sandoval and Molitoris, 2013, 2017).

Measurement of the fluorescence intensity of labeled compounds is the basis for many quantitative studies including glomerular permeability, PT reabsorption, co-localization and many others. Quantitative intensity-based data analysis requires strict attention to how instrument parameters and sensitivity are managed to completely utilize the full dynamic range of the system (Sandoval and Molitoris, 2013; Sandoval et al., 2014, 2019). For example, if settings are not correct the ratiometric intensities of the same compound in two different compartments can vary by orders of magnitude. Full dynamic range utilization requires system detectors with correct offset, or black level settings, showing only a few pixels in the image randomly flash as having values of zero (Sandoval et al., 2014; Sandoval and Molitoris, 2014). When acquiring background images setting all pixels to zero, in an effort to remove background during acquisition, decreases sensitivity thus reducing the ability to detect low intensity values (Nakano et al., 2012; Schiessl and Castrop, 2013; Sandoval and Molitoris, 2014; Schiessl et al., 2015).

When studying a protein it is essential to make sure the conjugation of the fluorophore does not alter its physiologic binding and or function. For albumin we have found a 1:1 ratio of protein to fluorophore and use of a multi-carbon spacer on the fluorophore, maintains physiologic binding affinity (Wagner et al., 2016a). Increasing the conjugation ratio often leads to reduced function and altered kidney metabolism and vascular clearance (Wagner et al., 2016a). Therefore, it is essential, but often overlooked, to ask, and test if possible, whether the labeled protein has the same biological properties as the native protein before undertaking imaging studies (Wagner et al., 2016a; Sandoval et al., 2019).

The answer to which animal model to use is primarily dictated by the question being asked. Mice have several advantages including a wealth of transgenic strains and many strains with fluorescently labeled cells such as the Tie-2 mice labeled endothelial cells. The relative ease of generating unique

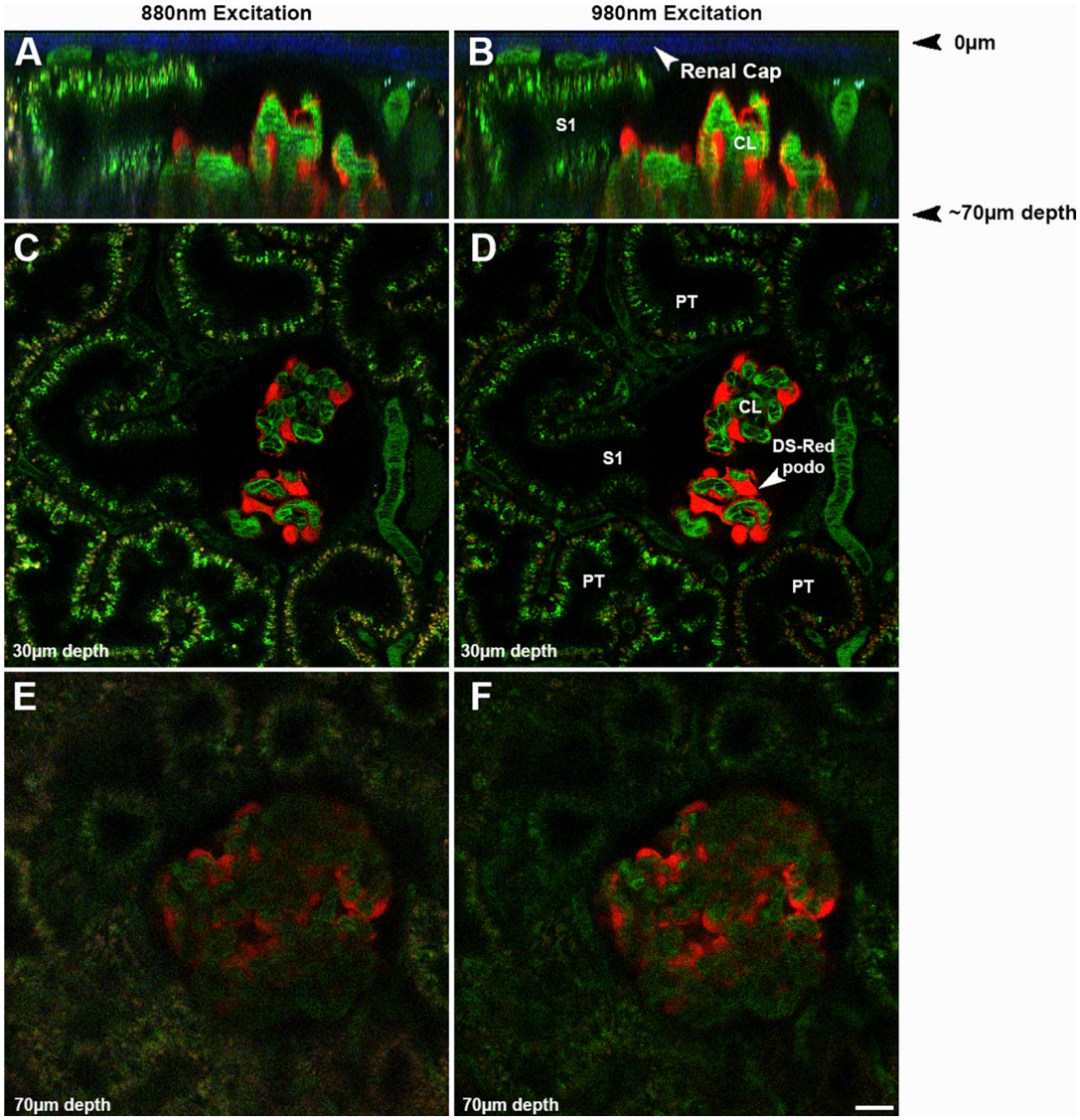


FIGURE 2 | Effect of imaging depth on image intensity and resolution: The same glomerulus in a strain of Munich Wistar Frömter rats, expressing the fluorescent protein DS Red selectively in podocytes, was imaged from 20 to 70 microns from the surface using two different wavelengths, 880 **(A,C,E)** and 980 nm **(B,D,F)**. Oregon Green 488 labeled rat serum albumin (OG488-RSA) was injected I.V. and can be seen in the vasculature, proximal tubules, and glomerulus. To assure illumination remained relatively constant from the upper to the lower optical sections in the image volume, the Linear Z-Compensation feature on the Leica Dive Multi-Photon system was utilized. **(A,B)** Show orthogonal the Linear Z-Compensation feature on the Leica Dive Multi-Photon system X-Z projections of a 73 μm micron volume, with the glomerular surface at the top of the image; note the 2nd harmonic excitation of collagen in the renal capsule (arrows, blue). These orthogonal projections in show the degradation in resolution and intensity in the deeper portions of the tissue, due to light scattering of the emitted light. The S1 cross section from both excitation wavelengths can resolve small endosomes and tubular structures rich in OG488-RSA at the upper cross section at 30 μm. The lower cross section of the same S1 has a hazier appearance, with only a few individual endosomes identifiable. Single plane images shown in **(C–F)** show the loss in resolution and intensity of the individual endosomes in **(E,F)** (taken at 70 μm), as compared to **(C,D)** (taken at 30 μm). The loss in resolution at the lower depths extends to small structures like endosomes and also includes losing the ability to discern circulating red blood cells in peritubular vessels and glomerular capillary loops. The inability to clearly discern the boundaries of glomerular capillary loops or peritubular blood vessels and the general haze makes intensity based (such as GSCs) or morphology based (such as RBC flow) analysis nearly impossible and greatly increases error and data variability (Bar = 20 μm).

transgenic mice has also been an important advantage. The Peti-Peterdi laboratory has used this approach to follow endothelial and glomerular epithelial regeneration using serial intravital multi-photon microscopy (Hackl et al., 2013; Schiessl et al., 2020; Desposito et al., 2021). These studies have shed light on the dynamic alterations, spatial distribution and fate of single renal cells or cell populations and their migration patterns in the same tissue region over several days in response to various stimuli. As delineated above, glomerular studies in mice are challenging as glomeruli are rarely within 100 microns of the

surface after 4 weeks of age in all mice strains (Schiessl et al., 2013). Ureteral obstruction for 6–12 weeks, or non-steroidal anti-inflammatory agents, have been used to induce surface glomeruli, but this comes at the cost of tubular destruction, cortical atrophy and fibrosis (Chevalier et al., 2009; Yang et al., 2010). This same team of investigators have shown, in their recent studies, deep glomeruli in mice can be imaged and both afferent and efferent arteriolar RBC flow can be quantified (Gyarmati et al., 2021a,b; Shroff et al., 2021). To do this they again used longer wavelength light to visualize normal cortical depth glomeruli in mice. However, resolution does suffer and not all processes can be quantified at this depth. A recent study also shows mice glomeruli may increase on the surface during progressive disease in a mouse model of Alport's Syndrome (Gyarmati et al., 2021b).

Another investigative team has used cortical resection to expose subsurface glomeruli in mice. While necessary for glomerular visualization, this approach resulted in a very high GSC for albumin, the ratio of glomerular filtrate to capillary albumin fluorescence, of 0.2–0.3, likely resulting form injury induced by the resection (Kidokoro et al., 2019). We have chosen to primarily study Munich Wistar Frömter rat strains (Simonsen and Frömter) that have easily imaged surface glomeruli. The Peti-Peterdi laboratory has developed a nice technique to quantify single nephron GFR and renal blood flow in these rats (Kang et al., 2006). The Frömter strain has up to three times more than the Simonsen's strain. Surface glomerular capillaries are seen within Bowman's Capsule, lack any autofluorescence, and are surrounded by proximal tubules (**Figures 2C,D**). The rat glomerulus consists of lobules and is about 100 microns in diameter allowing full 3D studies to be conducted. Unfortunately, the afferent arteriole usually lies at the bottom of the glomerulus making studies of it difficult due to decreased resolution and sensitivity at that cortical depth.

The S1 segment of the proximal tubule can be easily identified having a direct opening into the glomerulus making identifying and studying this unique and very endocytic segment easy (Dunn et al., 2002; Molitoris and Sandoval, 2005, 2011; Yu et al., 2005; Russo et al., 2007a; Sharfuddin et al., 2009; Wang et al., 2010; Sharfuddin and Molitoris, 2011; Sandoval and Molitoris, 2013, 2017; Dickson et al., 2014; Rizk et al., 2018; Basile, 2019; Molitoris et al., 2019; Sandoval et al., 2019). Being able to identify and study the S1 segment of the proximal tubule is important as this segment has the greatest capacity for endocytosis of macromolecules. This includes filtered proteins, vitamins, drugs, and endogenous and exogenous nephrotoxins. Differentiating S1 from S2 PT can be done in mice based on endogenous autofluorescence but not in rats (Kalakeche et al., 2011). We have found that anionic and neutral dextrans are endocytosed differently between S1 and S2 thus providing another way to distinguish these PTs in rats.

CONCLUSION

In summary, intravital MPM can serve as an invaluable tool to enhance the research objectives of many laboratories studying the physiology, pathophysiology and therapy of the kidney, or any organ that is accessible to exposure, placement and stabilization for intravital MPM microscopy. Multiple aspects can be studied individually and up to four fluorescent dyes can be visualized and spectrally separated. Since these dyes localize differently within tissue compartments, a greater number of cellular of processes can be simultaneously studied than the number of detector channels.

AUTHOR CONTRIBUTIONS

BM designed the study. BM, RS, and MW wrote the manuscript. RS created the figure. All authors contributed to the article and approved the submitted version.

FUNDING

The authors acknowledge grant support to BM from the National Institutes of Health (NIH) (DK091623 and DK079312).

REFERENCES

Ashworth, S. L., Sandoval, R. M., Tanner, G. A., and Molitoris, B. A. (2007). Two-photon microscopy: visualization of kidney dynamics. *Kidney Int.* 72, 416–421. doi: 10.1038/sj.ki.5002315

Basile, D. P. (2019). The case for capillary rarefaction in the AKI to CKD progression: insights from multiple injury models. *Am. J. Physiol. Renal Physiol.* 317, F1253–F1254. doi: 10.1152/ajprenal.00468.2019

Chevalier, R. L., Forbes, M. S., and Thornhill, B. A. (2009). Ureteral obstruction as a model of renal interstitial fibrosis and obstructive nephropathy. *Kidney Int.* 75, 1145–1152. doi: 10.1038/ki.2009.86

Choong, F. X., Sandoval, R. M., Molitoris, B. A., and Richter-Dahlfors, A. (2012). Multiphoton microscopy applied for real-time intravital imaging of bacterial infections in vivo. *Methods Enzymol.* 506, 35–61. doi: 10.1016/B978-0-12-391856-7.00027-5

Desposito, D., Schiessl, I. M., Gyarmati, G., Riquier-Brison, A., Izuhara, A. K., Kadoya, H., et al. (2021). Serial intravital imaging captures dynamic and functional endothelial remodeling with single-cell resolution. *JCI Insight* 6:e123392. doi: 10.1172/jci.insight.123392

Dickson, L. E., Wagner, M. C., Sandoval, R. M., and Molitoris, B. A. (2014). The proximal tubule and albuminuria: really! *J. Am. Soc. Nephrol.* 25, 443–453. doi: 10.1681/asn.2013090950

Dunn, K. W., Molitoris, B. A., and Dagher, P. C. (2021). The Indiana O'Brien Center for advanced renal microscopic analysis. *Am. J. Physiol. Renal Physiol.* 320, F671–F682. doi: 10.1152/ajprenal.00007.2021

Dunn, K. W., Sandoval, R. M., Kelly, K. J., Dagher, P. C., Tanner, G. A., Atkinson, S. J., et al. (2002). Functional studies of the kidney of living animals using multicolor two-photon microscopy. *Am. J. Physiol. Cell Physiol.* 283, C905–C916. doi: 10.1152/ajpcell.00159.2002

Ferrell, N., Sandoval, R. M., Bian, A., Campos-Bilderback, S. B., Molitoris, B. A., and Fissell, W. H. (2015). Shear stress is normalized in glomerular capillaries following 5/6 nephrectomy. *Am. J. Physiol. Renal Physiol.* 308, F588–F593. doi: 10.1152/ajprenal.00290.2014

Gyarmati, G., Jacob, C. O., and Peti-Peterdi, J. (2021a). New endothelial mechanisms in glomerular (Patho)biology and proteinuria development captured by Intravital Multiphoton imaging. *Front. Med. (Lausanne)* 13:765356. doi: 10.3389/fmed.2021.765356

Gyarmati, G., Shroff, U. N., Izuhara, A., Hou, X., Da Sacco, S., Sedrakyan, S., et al. (2021b). Intravital imaging reveals glomerular capillary distension and endothelial and immune cell activation early in Alport syndrome. *JCI Insight* 7:e152676. doi: 10.1172/jci.insight.152676

Hackl, M. J., Burford, J. L., Villanueva, K., Lam, L., Susztak, K., Schermer, B., et al. (2013). Tracking the fate of glomerular epithelial cells in vivo using serial multiphoton imaging in new mouse models with fluorescent lineage tags. *Nat. Med.* 19, 1661–1666. doi: 10.1038/nm.3405

Hall, A. M., Rhodes, G. J., Sandoval, R. M., Corridon, P. R., and Molitoris, B. A. (2013). In vivo multiphoton imaging of mitochondrial structure and function during acute kidney injury. *Kidney Int.* 83, 72–83. doi: 10.1038/ki.2012.328

Hato, T., El-Achkar, T. M., and Dagher, P. C. (2013). Sisters in arms: myeloid and tubular epithelial cells shape renal innate immunity. *Am. J. Physiol. Renal Physiol.* 304, F1243–F1251. doi: 10.1152/ajprenal.00101.2013

Horbelt, M., Wotzlaw, C., Sutton, T. A., Molitoris, B. A., Philipp, T., Kribben, A., et al. (2007). Organic cation transport in the rat kidney in vivo visualized by time-resolved two-photon microscopy. *Kidney Int.* 72, 422–429. doi: 10.1038/sj.ki.5002317

Hu, M. C., Shi, M., Zhang, J., Addo, T., Cho, H. J., Barker, S. L., et al. (2016). Renal production, uptake, and handling of circulating alphaKlotho. *J. Am. Soc. Nephrol.* 27, 79–90. doi: 10.1681/ASN.2014101030

Ji, N. (2017). Adaptive optical fluorescence microscopy. *Nat. Methods* 14, 374–380. doi: 10.1038/nmeth.4218

Kalakeche, R., Hato, T., Rhodes, G., Dunn, K. W., El-Achkar, T. M., Plotkin, Z., et al. (2011). Endotoxin uptake by S1 proximal tubular segment causes oxidative stress in the downstream S2 segment. *J. Am. Soc. Nephrol.* 22, 1505–1516. doi: 10.1681/ASN.2011020203

Kang, J. J., Toma, I., Sipos, A., McCulloch, F., and Peti-Peterdi, J. (2006). Quantitative imaging of basic functions in renal (patho)physiology. *Am. J. Physiol. Renal Physiol.* 291, F495–F502. doi: 10.1152/ajprenal.00521.2005

Kelly, K. J., Sandoval, R. M., Dunn, K. W., Molitoris, B. A., and Dagher, P. C. (2003). A novel method to determine specificity and sensitivity of the TUNEL reaction in the quantitation of apoptosis. *Am. J. Physiol. Cell Physiol.* 284, C1309–C1318. doi: 10.1152/ajpcell.00353.2002

Kidokoro, K., Cherney, D. Z. I., Bozovic, A., Nagasu, H., Satoh, M., Kanda, E., et al. (2019). Evaluation of glomerular hemodynamic function by empagliflozin in diabetic mice using in vivo imaging. *Circulation* 140, 303–315. doi: 10.1161/CIRCULATIONAHA.118.037418

Kolb, A. L., Corridon, P. R., Zhang, S., Xu, W., Witzmann, F. A., Collett, J. A., et al. (2018). Exogenous gene transmission of Isocitrate Dehydrogenase 2 Mimics Ischemic preconditioning protection. *J. Am. Soc. Nephrol.* 29, 1154–1164. doi: 10.1681/ASN.2017060675

Mansson, L. E., Melican, K., Molitoris, B. A., and Richter-Dahlfors, A. (2007). Progression of bacterial infections studied in real time–novel perspectives provided by multiphoton microscopy. *Cell Microbiol.* 9, 2334–2343. doi: 10.1111/j.1462-5822.2007.01019.x

Melican, K., Sandoval, R. M., Kader, A., Josefsson, L., Tanner, G. A., Molitoris, B. A., et al. (2011). Uropathogenic *Escherichia coli* P and Type 1 fimbriae act in synergy in a living host to facilitate renal colonization leading to nephron obstruction. *PLoS Pathog.* 7:e1001298. doi: 10.1371/journal.ppat.1001298

Molitoris, B. A., and Sandoval, R. M. (2005). Intravital multiphoton microscopy of dynamic renal processes. *Am. J. Physiol. Renal. Physiol.* 288, F1084–F1089. doi: 10.1152/ajprenal.00473.2004

Molitoris, B. A., and Sandoval, R. M. (2006). Pharmacophotonics: utilizing multiphoton microscopy to quantify drug delivery and intracellular trafficking in the kidney. *Adv. Drug Deliv. Rev.* 58, 809–823. doi: 10.1016/j.addr.2006.07.017

Molitoris, B. A., and Sandoval, R. M. (2011). Kidney endothelial dysfunction: ischemia, localized infections and sepsis. *Contrib. Nephrol.* 174, 108–118. doi: 10.1159/000329248

Molitoris, B. A., Dagher, P. C., Sandoval, R. M., Campos, S. B., Ashush, H., Fridman, E., et al. (2009). siRNA targeted to p53 attenuates ischemic and cisplatin-induced acute kidney injury. *J. Am. Soc. Nephrol.* 20, 1754–1764. doi: 10.1681/ASN.2008111204

Molitoris, B. A., George, A. G., Murray, P. T., Meier, D., Reilly, E. S., Barreto, E., et al. (2019). A novel fluorescent clinical method to rapidly quantify plasma volume. *Cardiorenal. Med.* 9, 168–179. doi: 10.1159/000496480

Nakano, D., Kobori, H., Burford, J. L., Gevorgyan, H., Seidel, S., Hitomi, H., et al. (2012). Multiphoton imaging of the glomerular permeability of angiotensinogen. *J. Am. Soc. Nephrol.* 23, 1847–1856. doi: 10.1681/ASN.2012010078

Ranjit, S., Dobrinskikh, E., Montford, J., Dvornikov, A., Lehman, A., Orlicky, D. J., et al. (2016). Label-free fluorescence lifetime and second harmonic generation imaging microscopy improves quantification of experimental renal fibrosis. *Kidney Int.* 90, 1123–1128. doi: 10.1016/j.kint.2016.06.030

Rizk, D. V., Meier, D., Sandoval, R. M., Chacana, T., Reilly, E. S., Seegmiller, J. C., et al. (2018). A Novel method for rapid bedside measurement of GFR. *J. Am. Soc. Nephrol.* 29, 1609–1613. doi: 10.1681/ASN.2018020160

Russo, L. M., Sandoval, R. M., Brown, D., Molitoris, B. A., and Comper, W. D. (2007a). Controversies in nephrology: response to 'renal albumin handling, facts, and artifacts'. *Kidney Int.* 72, 1195–1197. doi: 10.1038/sj.ki.5002528

Russo, L. M., Sandoval, R. M., McKee, M., Osicka, T. M., Collins, A. B., Brown, D., et al. (2007b). The normal kidney filters nephrotic levels of albumin retrieved by proximal tubule cells: retrieval is disrupted in nephrotic states. *Kidney Int.* 71, 504–513. doi: 10.1038/sj.ki.5002041

Russo, L. M., Sandoval, R. M., Campos, S. B., Molitoris, B. A., Comper, W. D., and Brown, D. (2009). Impaired tubular uptake explains albuminuria in early diabetic nephropathy. *J. Am. Soc. Nephrol.* 20, 489–494. doi: 10.1681/ASN.2008050503

Sandoval, R. M., and Molitoris, B. A. (2008). Quantifying endocytosis in vivo using intravital two-photon microscopy. *Methods Mol. Biol.* 440, 389–402. doi: 10.1007/978-1-59745-178-9_28

Sandoval, R. M., and Molitoris, B. A. (2013). Quantifying glomerular permeability of fluorescent macromolecules using 2-photon microscopy in Munich Wistar rats. *J. Vis. Exp.* 74:50052. doi: 10.3791/50052

Sandoval, R. M., and Molitoris, B. A. (2014). Letter to the editor: "Quantifying albumin permeability with multiphoton microscopy: why the difference?" *Am. J. Physiol. Renal Physiol.* 306, F1098–F1100.

Sandoval, R. M., and Molitoris, B. A. (2017). Intravital multiphoton microscopy as a tool for studying renal physiology and pathophysiology. *Methods* 128, 20–32. doi: 10.1016/j.ymeth.2017.07.014

Sandoval, R. M., Kennedy, M. D., Low, P. S., and Molitoris, B. A. (2004). Uptake and trafficking of fluorescent conjugates of folic acid in intact kidney determined using intravital two-photon microscopy. *Am. J. Physiol. Cell Physiol.* 287, C517–C526.

Sandoval, R. M., Molitoris, B. A., and Palygin, O. (2019). Fluorescent imaging and microscopy for dynamic processes in rats. *Methods Mol. Biol.* 2018, 151–175. doi: 10.1007/978-1-4939-9581-3_7

Sandoval, R. M., Wagner, M. C., Patel, M., Campos-Bilderback, S. B., Rhodes, G. J., Wang, E., et al. (2012). Multiple factors influence glomerular albumin permeability in rats. *J. Am. Soc. Nephrol.* 23, 447–457. doi: 10.1681/ASN.2011070666

Sandoval, R. M., Wang, E., and Molitoris, B. A. (2014). Finding the bottom and using it: offsets and sensitivity in the detection of low intensity values in vivo with 2-photon microscopy. *Intravital* 2:e23674. doi: 10.4161/intv.23674

Schiessl, I. M., and Castrop, H. (2013). Angiotensin II AT2 receptor activation attenuates AT1 receptor-induced increases in the glomerular filtration of albumin: a multiphoton microscopy study. *Am. J. Physiol. Renal Physiol.* 305, F1189–F1200. doi: 10.1152/ajprenal.00377.2013

Schiessl, I. M., Bardehle, S., and Castrop, H. (2013). Superficial nephrons in BALB/c and C57BL/6 mice facilitate in vivo multiphoton microscopy of the kidney. *PLoS One* 8:e52499. doi: 10.1371/journal.pone.0052499

Schiessl, I. M., Fremter, K., Burford, J. L., Castrop, H., and Peti-Peterdi, J. (2020). Long-term cell fate tracking of individual renal cells using serial Intravital microscopy. *Methods Mol. Biol.* 2150, 25–44.

Schiessl, I. M., Kattler, V., and Castrop, H. (2015). In vivo visualization of the antialbuminuric effects of the angiotensin-converting enzyme inhibitor enalapril. *J. Pharmacol. Exp. Ther.* 353, 299–306. doi: 10.1124/jpet.114.222125

Schuh, C. D., Haenni, D., Craigie, E., Ziegler, U., Weber, B., Devuyst, O., et al. (2016). Long wavelength multiphoton excitation is advantageous for intravital kidney imaging. *Kidney Int.* 89, 712–719. doi: 10.1038/ki.2015.323

Schuh, C. D., Polesel, M., Platonova, E., Haenni, D., Gassama, A., Tokonami, N., et al. (2018). Combined structural and functional imaging of the kidney reveals major axial differences in proximal tubule endocytosis. *J. Am. Soc. Nephrol.* 29, 2696–2712. doi: 10.1681/ASN.2018050522

Schulz, A., Chuquimia, O. D., Antypas, H., Steiner, S. E., Sandoval, R. M., Tanner, G. A., et al. (2018). Protective vascular coagulation in response to bacterial

infection of the kidney is regulated by bacterial lipid A and host CD147. *Pathog. Dis.* 76:fty087. doi: 10.1093/femspd/fty087

Sharfuddin, A. A., and Molitoris, B. A. (2011). Pathophysiology of ischemic acute kidney injury. *Nat. Rev. Nephrol.* 7, 189–200. doi: 10.1038/nrneph.2011.16

Sharfuddin, A. A., Sandoval, R. M., Berg, D. T., McDougal, G. E., Campos, S. B., Phillips, C. L., et al. (2009). Soluble thrombomodulin protects ischemic kidneys. *J. Am. Soc. Nephrol.* 20, 524–534. doi: 10.1681/ASN.2008060593

Shroff, U. N., Gyarmati, G., Izuhara, A., Deepak, S., and Peti-Peterdi, J. (2021). A new view of macula densa cell protein synthesis. *Am. J. Physiol. Renal Physiol.* 321, F689–F704. doi: 10.1152/ajprenal.00222.2021

Sutton, T. A., Mang, H. E., Campos, S. B., Sandoval, R. M., Yoder, M. C., and Molitoris, B. A. (2003). Injury of the renal microvascular endothelium alters barrier function after ischemia. *Am. J. Physiol. Renal Physiol.* 285, F191–F198. doi: 10.1152/ajprenal.00042.2003

Tanner, G. A., Sandoval, R. M., Molitoris, B. A., Bamburg, J. R., and Ashworth, S. L. (2005). Micropuncture gene delivery and intravital two-photon visualization of protein expression in rat kidney. *Am. J. Physiol. Renal Physiol.* 289, F638–F643. doi: 10.1152/ajprenal.00059.2005

Tenten, V., Menzel, S., Kunter, U., Sicking, E. M., van Roeyen, C. R., Sanden, S. K., et al. (2013). Albumin is recycled from the primary urine by tubular transcytosis. *J. Am. Soc. Nephrol.* 24, 1966–1980. doi: 10.1681/ASN.2013010018

Wagner, M. C., Campos-Bilderback, S. B., Chowdhury, M., Flores, B., Lai, X., Myslinski, J., et al. (2016a). Proximal tubules have the capacity to regulate uptake of albumin. *J. Am. Soc. Nephrol.* 27, 482–494. doi: 10.1681/ASN.2014111107

Wagner, M. C., Myslinski, J., Pratap, S., Flores, B., Rhodes, G., Campos-Bilderback, S. B., et al. (2016b). Mechanism of increased clearance of glycated albumin by proximal tubule cells. *Am. J. Physiol. Renal Physiol.* 310, F1089–F1102. doi: 10.1152/ajprenal.00605.2015

Wang, E., Sandoval, R. M., Campos, S. B., and Molitoris, B. A. (2010). Rapid diagnosis and quantification of acute kidney injury using fluorescent ratiometric determination of glomerular filtration rate in the rat. *Am. J. Physiol. Renal Physiol.* 299, F1048–F1055. doi: 10.1152/ajprenal.00691.2009

Ward, E. S., Martinez, C., Vaccaro, C., Zhou, J., Tang, Q., and Ober, R. J. (2005). From sorting endosomes to exocytosis: association of Rab4 and Rab11 GTPases with the Fc receptor, FcRn, during recycling. *Mol. Biol. Cell.* 16, 2028–2038. doi: 10.1091/mbc.e04-08-0735

Weinberg, J. M., and Molitoris, B. A. (2009). Illuminating mitochondrial function and dysfunction using multiphoton technology. *J. Am. Soc. Nephrol.* 20, 1164–1166. doi: 10.1681/ASN.2009040419

Yadav, S. P. S., Sandoval, R. M., Zhao, J., Huang, Y., Wang, E., Kumar, S., et al. (2021). Mechanism of how carbamylation reduces albumin binding to FcRn contributing to increased vascular clearance. *Am. J. Physiol. Renal Physiol.* 320, F114–F129. doi: 10.1152/ajprenal.00428.2020

Yang, H. C., Zuo, Y., and Fogo, A. B. (2010). Models of chronic kidney disease. *Drug Discov. Today Dis. Models* 7, 13–19.

Yu, W., Sandoval, R. M., and Molitoris, B. A. (2005). Quantitative intravital microscopy using a Generalized Polarity concept for kidney studies. *Am. J. Physiol. Cell Physiol.* 289, C1197–C1208. doi: 10.1152/ajpcell.00197.2005

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

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.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership