BRIEF COMMUNICATION

A Multimodal Imaging Pipeline for the Discovery of Molecular Markers of Cellular Neighborhoods

Thai Pham,[1,2,a] Melissa A. Farrow,[1,2,a] Lukasz G. Migas,[3,b] Martin Dufresne,[1,2,b] Madeline E. Colley,[1,2] Jamie L. Allen,[1,2] Mark P. de Caestecker,[4] Raf Van de Plas,[2,3,5] and Jeffrey M. Spraggins [1,2,5,6,7*]

[1]Department of Cell and Developmental Biology, Vanderbilt University School of Medicine, Nashville, TN, USA 37232.
[2]Mass Spectrometry Research Center, Vanderbilt University School of Medicine, Nashville, TN, USA 37232.
[3]Delft Center for Systems and Control, Delft University of Technology, 2628 CD Delft, The Netherlands.
[4]Division of Nephrology and Hypertension, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN USA 37232.
[5]Department of Biochemistry, Vanderbilt University School of Medicine, Nashville, TN, USA 37232.
[6]Department of Pathology, Microbiology and Immunology, Vanderbilt University Medical Center, Nashville, TN USA 37232.
[7]Department of Chemistry, Vanderbilt University, Nashville, TN, USA 37232.

[a]Co-first authors/contributed equally
[b]Co-second authors/contributed equally

[*]Corresponding Author:
Jeffrey M. Spraggins, Ph.D.
465 21st Ave. South
9160 Medical Research Building III
Vanderbilt University
Nashville, TN 37212
jeff.spraggins@vanderbilt.edu

Abstract:

We developed an integrated imaging pipeline for spatio-molecular and immunophenotypic tissue analysis, enabling the discovery of molecular markers across spatial scales—from functional tissue units to cell types and neighborhoods. Our approach facilitates single-tissue section analysis, enabling direct cross-modality mining of imaging mass spectrometry and highly multiplexed immunofluorescence microscopy. Demonstrated on human kidney tissue, we identified lipidomic markers of cellular neighborhoods associated with renal cell types and immune cell infiltration.

Main Text:

Spatial omics assays are rapidly emerging as essential tools in biomedical research, offering unprecedented insights into the molecular and cellular architecture of tissue microenvironments. By integrating complementary molecular imaging technologies, multimodal workflows enable a systems biology approach to understanding tissue organization, cellular interactions, and biochemical processes in both normal aging and disease contexts. Here, we describe an optimized multimodal imaging pipeline combining autofluorescence (AF) microscopy, imaging mass spectrometry (IMS), co-detection by indexing (CODEX) multiplexed immunofluorescence microscopy, and periodic acid-Schiff (PAS)-stained histology from a single tissue section (Fig. 1a). CODEX is a highly multiplexed, antibody-based method that enumerates and characterizes cells in tissue at outstanding spatial resolution[1,2]. IMS is unique in its ability to perform untargeted detection of lipids, metabolites, and small molecules in tissues with exceptional chemical specificity[3]. Our integrated workflow achieves precise spatial co-registration across modalities, directly linking detailed molecular profiles to specific cell types, functional tissue units, and cellular neighborhoods. This facilitates comprehensive analyses of spatially resolved molecular signatures, supporting biomarker discovery and providing novel insights into tissue metabolism, immune cell function, and cellular crosstalk relevant to normal aging and disease pathogenesis.[134]

All modalities were collected from a single tissue section, eliminating spatial ambiguity when relating data between image types (Fig. 1b). Autofluorescence microscopy collected before and after matrix-assisted laser desorption ionization (MALDI) IMS enabled high-precision spatial co-registration of each imaging modality[5,6]. High spatial resolution IMS (5 µm pixel size) was collected before CODEX to prevent analyte delocalization induced by necessary sample preparation steps for immunofluorescence assays. CODEX multiplexed immunofluorescence microscopy using a panel of 22 antibodies was used for spatial cell phenotyping and neighborhood analysis, and PAS-stained images provide histological context. Since all modalities were acquired from the same tissue section, co-registration was inherently more precise than similar experiments performed on serial tissue sections where different cells are sampled and tissue perturbations such as deformations or tears can influence multimodal analyses.

In situ single-cell segmentation and cell type assignments were performed using CODEX data on three human kidney samples (Fig S1-S6) and further analyzed to define cellular neighborhoods (Fig 1c), including calculating the cell type composition of each neighborhood (Fig 1d). Ten discreet cellular neighborhoods comprised of 18 cell types were identified. Of these, FTU level neighborhoods (neighborhood A-collecting duct, B-proximal tubule, C-medullary thick ascending limb, D-cortical thick ascending limb, H-outer glomeruli, I-inner glomeruli, and J-distal tubule) were identified as well as immune (neighborhood G) and vasculature (neighborhoods E and F) specific neighborhoods. While neighborhoods share cell types, each has a distinct population and distribution of cell types that define its spatial niche. For each sample, neighborhoods were overlaid on their associated PAS-stained images (Figs S7-S9), and the cell type composition of each neighborhood was calculated (Fig S10). Neighborhood segmentation masks for the proximal

tubule, immune, and glomerular neighborhoods were overlaid on the PAS-stained image for one donor sample (VAN0031), demonstrating the difference in cell numbers and distributions of each neighborhood (Fig 2a).

Neighborhood segmentation masks were used for mining integrated IMS data. Using interpretable machine learning leveraging Shapley Additive Explanations (SHAP), the biomarker discovery process was translated into a feature ranking problem[4]. First, a classification model was built to assign IMS pixels (Figs S11-S16) to different cellular neighborhoods based on their mass spectra (Figs S17-S22). Then, by applying SHAP, the relative predictive importance of each IMS molecular feature was calculated, with higher-ranking features being more important as molecular markers for each neighborhood (Figs 2b and S23). This approach builds on previous analysis of human kidney tissue lipid profiles linked to the specific functional tissue units (FTU) of the nephron[7]. By using the CODEX-driven neighborhood analysis to mine the lipid data, additional cell types and cellular interactions were detected, and unique information about cell-cell interactions, cross-talk, and niches was observed in terms of their cellular and molecular profiles.

Given the emergence of several FTU-level neighborhoods, we compared their molecular profiles with previous FTU-specific data, yielding results consistent with prior kidney molecular atlas efforts[6,7]. Neighborhood I is predominantly podocyte cells (Figs 1d and 2a), which are known to be enriched in SM(d34:1). Accordingly, the most enriched biomarker for I is SM(d34:1) (*m/z* 703.575) (Fig 2b). Similarly, Neighborhood B, primarily proximal tubule (PT) cells (Figs 1d and 2a), has an enhanced molecular signature of PE(36:4) (*m/z* 738.508), consistent with previous observations[7].

While neighborhoods dominated by FTU-specific cell types show consistency with existing datasets, they also yield new markers. Neighborhood A comprises collecting ducts and the thick ascending limb (TAL), FTUs that reside in the medulla. The carnitine class is highly enriched for Neighborhood A (*m/z* 400.341, 414.357, 428.372, 444.367, and 452.372), with Car(18:0) (*m/z* 428.372) the most robust neighborhood marker. Similarly, when TAL was stratified into medullary (mTAL, Neighborhood C) versus cortical (cTAL, Neighborhood D), carnitines are more highly associated with mTAL (*m/z* 414.357, 426.357, and 444.367). Car(18:1) (*m/z* 426.357) is an exclusive mTAL molecular marker. Carnitines are key molecules for aerobic respiration, which occurs in cortical FTUs where oxygen availability is high. However, the PT neighborhood (B), where aerobic respiration is highest[8], shows an anti-correlative SHAP profile for carnitines. Previous work has shown that PT and mTAL contain the highest density of mitochondria in the kidney, but the mitochondrial bioenergetic functions in these compartments are distinct, with mitochondria isolated from mTAL being more energetically efficient[9]. The specific carnitine profiles of mTAL shown here could provide additional insight into these unique mTAL metabolic processes.

The immune neighborhood (G) is enriched in PC species and was the only neighborhood where all saturated PC (*m/z* 706.538, 720.590, and 734.570) were found to be molecular markers. Saturated PCs are key mediators of lipid raft formation in T-cells, a critical regulator for T-cell receptor (TCR) signaling[10]. Previous work has shown that PC(32:0) was the most abundant PC in the TCR localized lipid raft domain[10]. Consistent with this, PC(16:0_16:0) (*m/z* 734.570) was the most robust marker positively correlated with the T-cell-rich immune neighborhood (Figs 1d and 2b). Aside from saturated PC, PS(18:0_18:1)

was the only other lipid reported to be enriched in the TCR lipid raft region[10]. PS(36:1) (*m/z* 788.544) is also an immune neighborhood marker (Fig 2b). [11,1211,1213]A single lipid marker cannot distinguish the immune neighborhood exclusively. However, both a univariate analysis of the data, as shown in Fig 2c, and the multivariate analysis (Fig 2b) reveal a specific lipid signature for the immune neighborhood. Interpreting the overall profile of the immune neighborhood provides a direct link from the enriched lipids to the function of the cell types present.

Integrated spatial-omics provides a comprehensive view of molecular characteristics across distinct cell populations within tissues. Leveraging this unique IMS-CODEX multimodal molecular imaging pipeline enables precise linkage of untargeted molecular profiles to cellular organization *in situ*. This approach enables discovery of molecular markers associated with functional neighborhoods, highlighting distinct lipidomic profiles between cortical and medullary kidney regions. Additionally, enriched lipids within immune neighborhoods reveal a critical connection between lipid localization and immune function. Combining IMS and CODEX on the same tissue section represents a significant advancement in spatial biology, offering unmatched precision for linking molecular phenotypes to cellular environments and serving as a powerful systems biology toolset for driving discoveries in biomedical research, diagnostics, and personalized medicine.

Acknowledgements:

This work was supported by the National Institutes of Health (NIH) Common Fund and National Institute Of Diabetes And Digestive And Kidney Diseases (NIDDK) under Award Numbers U54DK134302 and U01DK133766 (J.M.S. and R.V.), by the NIH Common Fund and National Eye Institute (NEI) under Award Number U54EY032442 (J.M.S. and R.V.), by the NIH National Institute On Aging (NIA) under Award Number R01AG078803 (J.M.S. and R.V.), and by the National Science Foundation Major Research Instrument Program CBET – 1828299 (J.M.S.). This research was furthermore made possible by the Chan Zuckerberg Initiative DAF, an advised fund of Silicon Valley Community Foundation, under Award Numbers 2021-240339 and 2022-309518 (L.G.M. and R.V.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the funders. The authors also thank the HuBMAP data curation and ingest team, particularly Brendan Honick, Phil Blood, and Jonathan Silverstein for their support.

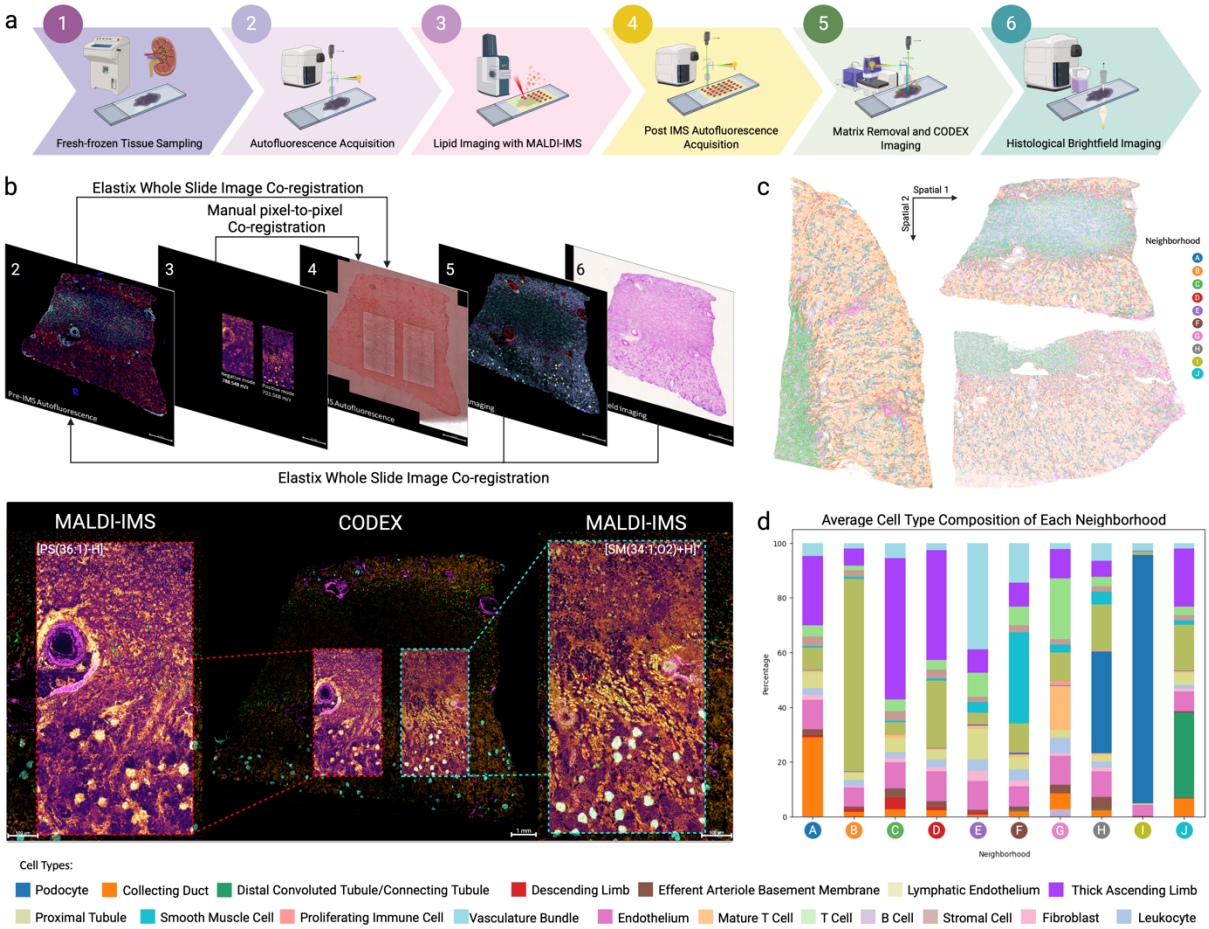

Figure 1: A multimodal workflow for collection and analysis of IMS-CODEX on the same tissue section. a. Overview of analytical pipeline for specimen processing and image acquisition. Autofluorescence was acquired on tissue sections before (pre-IMS AF) and after (post-IMS AF) MALDI IMS to support downstream image co-registration. Lipid imaging by MALDI IMS was done at 5 µm. Following matrix removal, a 22-antibody panel CODEX run was performed and PAS images were collected. **b.** (top panel) Multimodal images from VAN0031 were aligned using a two-step process combining elastix and manual pixel-to pixel co-registration with the post-AF image as the central anchor point. (bottom panel) The CODEX and IMS images were co-registered and assessed by confirming overlay of morphological features between IMS molecular distributions and CODEX cellular structures. Insets are zoomed images of $[PS(36:1)-H]^-$ and $[SM(34:1;O2)+H]^+$, demonstrating precision of the co-registration pipeline. **c.** Ten cell neighborhoods from VAN0005 (left), VAN0031 (top right), and VAN0019 (bottom right), were defined based on their distinct compositions of 18 different cell types derived from protein expression data extracted from the CODEX images. **d.** The average cell type composition for each neighborhood was calculated. Each cell type is color coded as shown in key at the bottom of the figure.

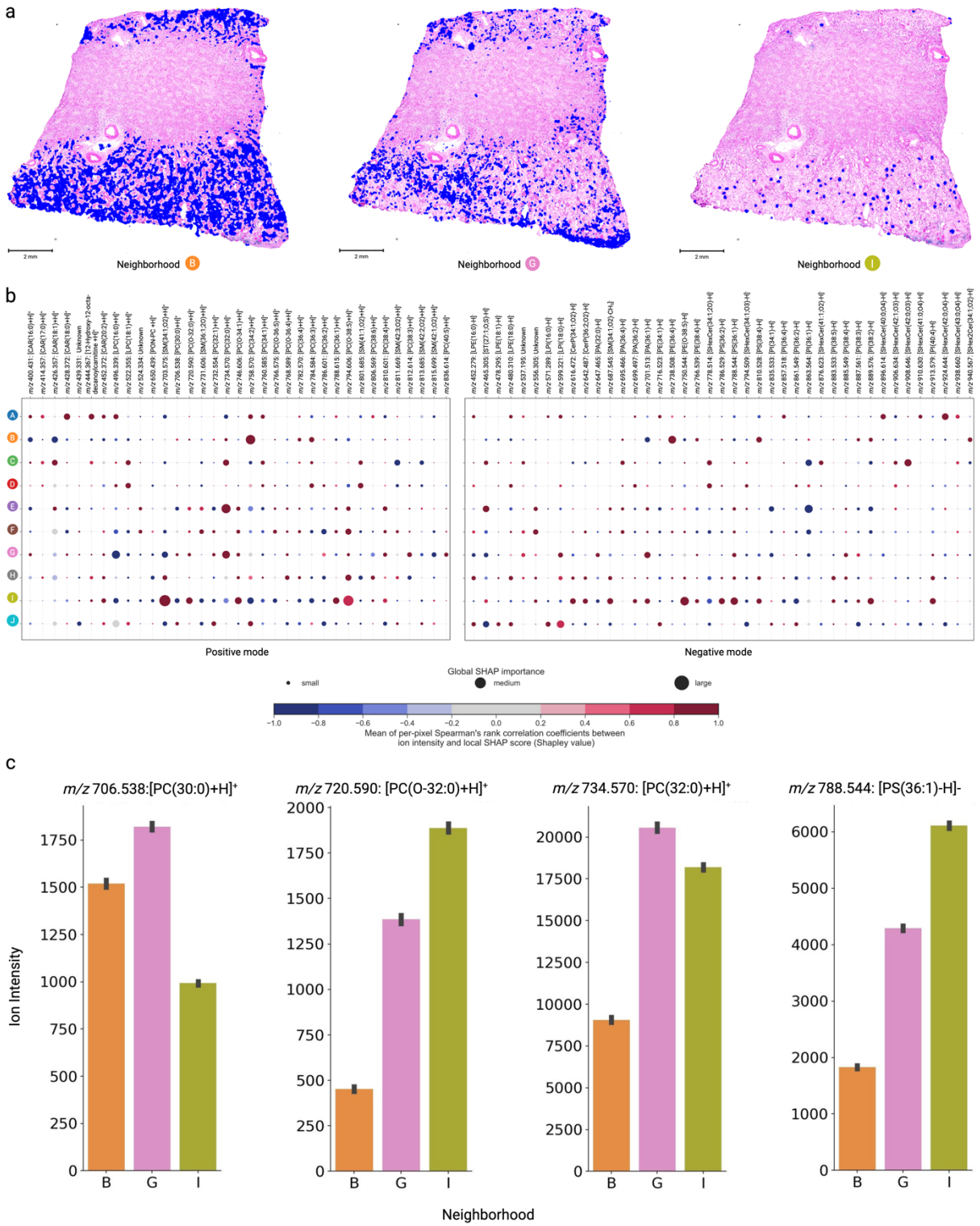

Figure 2: Integration and cross-modality data mining of IMS-CODEX workflow. a. Cell neighborhood masks were overlaid on histological images to illustrate their distinct spatial localizations. Representative examples are shown for Neighborhoods B, G, and I from VAN0031. **b.** SHAP bubble plot quantifying the contribution of each ion in positive (left) and

negative (right) mode to the identification of the 10 neighborhoods. Each column represents an ion with ions shown in ascending *m/z*. Bubble size represents the magnitude of each ion's importance for that neighborhood, while color indicates the direction of correlation ranging from blue (strongly uncorrelated) to red (strongly correlated). **c.** Univariate ion intensities for a selection of ions that showed a strong marker profile for Neighborhood G compared to Neighborhoods B and I.

1. Kennedy-Darling, J. *et al.* Highly multiplexed tissue imaging using repeated oligonucleotide exchange reaction. *Eur J Immunol* **51**, 1262–1277 (2021).
2. Neumann, E. K. *et al.* Highly multiplexed immunofluorescence of the human kidney using co-detection by indexing. *Kidney Int* **101**, 137–143 (2022).
3. Colley, M. E., Esselman, A. B., Scott, C. F. & Spraggins, J. M. High-Specificity Imaging Mass Spectrometry. *Annual Review of Analytical Chemistry* (2024) doi:10.1146/annurev-anchem-083023-024546.
4. Tideman, L. E. M. *et al.* Automated biomarker candidate discovery in imaging mass spectrometry data through spatially localized Shapley additive explanations. *Anal Chim Acta* **1177**, 338522 (2021).
5. Patterson, N. H., Tuck, M., Van de Plas, R. & Caprioli, R. M. Advanced Registration and Analysis of MALDI Imaging Mass Spectrometry Measurements through Autofluorescence Microscopy. *Anal Chem* **90**, 12395–12403 (2018).
6. Esselman, A. B. *et al.* In situ molecular profiles of glomerular cells by integrated imaging mass spectrometry and multiplexed immunofluorescence microscopy. *Kidney Int* **107**, 332–337 (2025).
7. Farrow, M. A. *et al.* A Lipid Atlas of the Human Kidney. Preprint at https://doi.org/10.1101/2022.04.07.487155 (2022).
8. Bhargava, P. & Schnellmann, R. G. Mitochondrial energetics in the kidney. *Nat Rev Nephrol* **13**, 629–646 (2017).
9. Tomar, N. *et al.* Substrate-dependent differential regulation of mitochondrial bioenergetics in the heart and kidney cortex and outer medulla. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1863**, 148518 (2022).
10. Zech, T. *et al.* Accumulation of raft lipids in T-cell plasma membrane domains engaged in TCR signalling. *EMBO J* **28**, 466–476 (2009).
11. Oude Weernink, P. A., López de Jesús, M. & Schmidt, M. Phospholipase D signaling: orchestration by PIP2 and small GTPases. *Naunyn Schmiedebergs Arch Pharmacol* **374**, 399–411 (2007).
12. Porciello, N., Kunkl, M., Viola, A. & Tuosto, L. Phosphatidylinositol 4-Phosphate 5-Kinases in the Regulation of T Cell Activation. *Front Immunol* **7**, 186 (2016).
13. Borges-Araújo, L. *et al.* Acyl-chain saturation regulates the order of phosphatidylinositol 4,5-bisphosphate nanodomains. *Commun Chem* **4**, 164 (2021).