

FUSION: a web-based application for in-depth exploration of multi-omics data with brightfield histology

Received: 9 October 2024

Accepted: 7 August 2025

Published online: 25 September 2025


 Check for updates

Samuel P. Border¹, Ricardo Melo Ferreira ², Nicholas Lucarelli¹, Suhas Katari Chaluva Kumar ¹, Anindya S. Paul¹, David Manthey ³, Laura Barisoni ^{4,5}, Yulia A. Levites Strelakova⁶, Jessica Ray⁷, Ying-Hua Cheng², Avi Z. Rosenberg⁸, John E. Tomaszewski⁹, Sayat Mimar ¹, Jeffrey B. Hodgins ¹⁰, John W. Hickey¹¹, Bei Wei¹², Fiona Ginty ¹³, Arivarasan Karunamurthy¹⁴, Juexin Wang ¹⁵, Mauminah Raina¹⁵, Gloria S. Pryhuber ¹⁶, Jeffrey Purkerson¹⁶, HuBMAP consortium*, Tarek M. El-Achkar ^{2,17} , Sanjay Jain ¹⁸ , Michael T. Eadon^{2,17}  & Pinaki Sarder ¹ 

Spatial technologies examining the cell and tissue microenvironment at near single-cell resolution are revealing important molecular insights. However, few tools enable integrated, interactive analysis of spatial-omics with tissue morphology in the same functional tissue unit. Here, we present FUSION (Functional Unit State Identification in Whole Slide Images), a web-based platform for visualizing and analyzing spatial-omics data with high-resolution histology. FUSION provides workflows for assessing cell compositions, quantitative morphometrics, and comparative tissue analyses. We demonstrate applicability across spatial assays, including 10x Visium, Visium HD, 10x Xenium, Cell DIVE, and PhenoCycler, applied to healthy and diseased tissues from kidney, small intestine, lung, and skin in the Human BioMolecular Atlas Program. FUSION is cloud-based, open-source, and accessible at <https://fusion.hubmapconsortium.org/>, hosting over 50 paired datasets and tutorials. In a series of use cases, we show its capacity to distinguish renal glomeruli injury states, quantify morphometric changes, and characterize fibrosis with immune infiltration.

Spatially-resolved, molecular-omics methods are poised to revolutionize medicine both in clinical care and investigative research^{1–3}. These techniques provide investigators with the ability to analyze biological processes down to subcellular resolution within large regions of interest (ROI)^{4,5}. When further aligned with histology images, spatial molecular data enrich observations of tissue characteristics and lesions, thereby imparting further insights into localized tissue injury responses. However, the high dimensionality of -omics data represents a barrier towards alignment with histologic findings and paired interpretation of both modes of data. When a pathologist interprets a

biopsy specimen, information from large whole slide images (WSIs), often consisting of over one billion pixels at 40X magnification, is condensed into a few discrete metrics which inform treatment and diagnostic decisions. Spatial-omics, by comparison, can contain thousands of measurements across the entire tissue area. A similar depth of histologic annotation, wherein every cell or functional tissue unit (FTU) of an organ is manually characterized at the pixel level, would exceed a pathologist's clinical capacity. Therefore, paired interpretation of spatial-omics and histology data require the ability to query each modality of data, mining heterogeneous information in

A full list of affiliations appears at the end of the paper. *A list of authors and their affiliations appears at the end of the paper.  e-mail: telachka@iu.edu; sanjayjain@wustl.edu; meadon@iu.edu; pinaki.sarder@medicine.ufl.edu

diverse regions of gigapixel size. Facilitating this complex integration necessitates a tool, allowing users the freedom to simultaneously interact and align multimodal information in a scalable, user-friendly interface, where data dimensionality is tractable and comparable. Such a tool will empower users with automated segmentation of tissue structures using deep learning (DL), quantification of tissue characteristics (“pathomics”), dimensionality reduction and statistical analysis of spatial data, and merging multiple modalities of data in a biologically meaningful fashion. To enhance accessibility and interoperability with other data analysis pipelines, the proposed tool should enable integrated histology and derived data output to share results with the broader research community.

We developed a cloud-based visualization and analysis tool, which combines morphological interpretation of FTUs at high resolution with genome-wide spatial-omics data. This tool, called Functional Unit State Identification for WSIs (*FUSION*), enables dynamic interaction between users and their data as well as running algorithms with high-computational costs (e.g., GPUs for automated segmentation) utilizing cloud resources (Fig. 1). A key distinction that we make between *FUSION* and other existing tools in this space is in the implementation of interactive analysis. This type of analysis incorporates user inputs in a “human-in-the-loop”⁶ model to iteratively refine the generation of derived data. For example, identification of distinct structures in biopsied tissue typically involves the recruitment of trained pathologists or subject-matter-experts to provide annotations. These annotations are then incorporated into downstream tasks to either train DL-based segmentation algorithms or to serve as a basis for quantitative morphometric extraction. Further along in the course of analysis, experts may weigh in to further narrow down which characteristics of

which structures are most informative and important to consider in the context of that study. In many ways, this process can be considered as a supervised form of feature selection. Instead of columns of data tables, regions of images and clusters of data points are considered for further interrogation. Facilitating interactive analysis requires the development of tools which can handle large scales of computation with substantial amounts of user guidance. The resulting system is then not as lightweight as tools which are designed just to facilitate visualization of existing data. Tools such as *Vitessce*, *Minerva*, *Loupe Browser*, and others are examples of tools optimized for visualization of large amounts of data but are not designed to enable users to generate additional data based on observations^{7–9}. On the opposite end of the spectrum, several tools are also available which are designed primarily to enable users to run large analysis jobs on their data and export the results of that analysis. Tools such as *WebMeV*, *spatialGE*, *Xena*, and *Crosto* present web-based solutions for large data analysis, storage, and organization^{10–13}. However, these analytical steps are not designed to dynamically respond to user interactions with data and are typically restricted to routine tasks such as quality control, normalization, dimensional reduction, and clustering. While both sets of visualization tools and analytical tools are generalizable to different types of input data, there remains a gap which has the potential to have a strong impact in the paired analysis of spatial-omics and histology datasets.

FUSION was designed in collaboration with the Human BioMolecular Atlas Program (HuBMAP) with the goal of aiding users in quantitatively linking spatial-omics datasets to histopathology and determining the distribution of cells and cell states within healthy and diseased tissue¹⁴. Due to its structural diversity and complex functions,

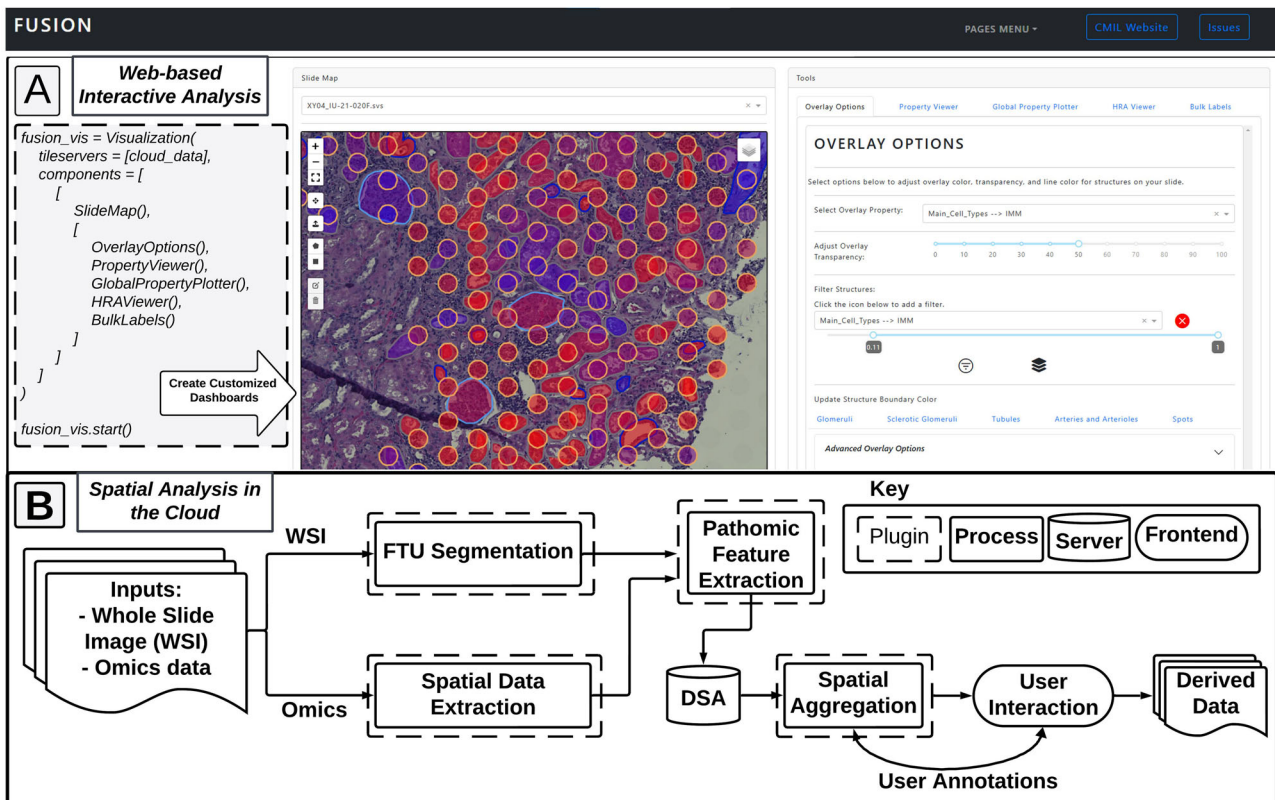


Fig. 1 | Schematic of *FUSION*. *FUSION* offers cloud-based unimodal and multimodal analysis of histology images and associated spatial-omics datasets. **A** Code block demonstrating how interactive components are added to a custom dashboard incorporating cloud-stored data. Example visualization shows the usage of the *OverlayOptions* component to overlay a colormap on automatically segmented

functional tissue units (FTUs) in the kidney alongside 10x Visium spots and apply a filter based on spatially aggregated properties. **B** Preprocessing flow diagram for separate processing of histological FTUs and spatial-omics data prior to spatial aggregation and user interaction, including interactive data analysis.

we used the kidney as a prototype organ to demonstrate how *FUSION* can generate easy-to-interpret visualizations of FTUs with their cellular composition. However, this process can also be applied to spatial-omics datasets from all major organs.

Results

Overview

We begin by discussing a broad system-level overview of *FUSION* (<https://fusion.hubmapconsortium.org>), and its various functionalities for diverse users, then describe features that are not available in existing cloud tools for large-scale image analytics of digital pathology data. We next use paired, multi-modal histology tissue images and associated *10x Visium* spatial transcriptomics (ST) data from kidney tissues of reference (healthy), chronic kidney disease (CKD), and acute kidney injury (AKI) subjects as a vehicle to demonstrate the potential application of *FUSION* in biological research and molecular pathology. Samples that are indicated as healthy reference are those that do not have prior clinical indication of other diseases with which they are compared. We provide an overview of how the procedure used in these cases can be expanded to include diverse spatial-omics data including *PhenoCycler*, *Cell DIVE*, *10x Xenium*, and *Visium HD* across multiple organs. Additional examples for each of these technologies are provided in the Supplementary Information.

FUSION and fusion-tools

FUSION, and the associated software package *fusion-tools*, allow for the custom generation of web-hosted interactive dashboards, integrating both cloud-hosted and local spatial-omics datasets (Fig. 1A). Interactive components in *FUSION* can be arranged in different layouts with other components in columns and rows or even across multiple pages depending on the user's preference. These components are built using *Dash*, a web-based framework for Python applications, specifically *DashBlueprints* (a part of the *dash-extensions* library) which enables the embedding of components included in *fusion-tools* in other application layouts as well as the creation and embedding of custom components into a *FUSION* layout. The primary instance of *FUSION* (<https://fusion.hubmapconsortium.org>), is organized into several different pages to facilitate uploading and processing of user's datasets, building datasets based on sample-level metadata, and interactive analysis. Processes with high computational cost are offloaded in the form of Dockerized, Slicer CLI-web plugins to a separate server where a running instance of Digital Slide Archive (DSA)¹⁵ facilitates data storage and access through a RESTful-API (Fig. 1B).

FUSION currently hosts example datasets consisting of *10x Visium*, *Cell DIVE* and *PhenoCycler*, *10x Xenium*, and *Visium HD* data for a variety of different organs. In each of these samples, hematoxylin and eosin (H&E) stained histological sections are present for the same section, allowing for multi-modal data alignment between spatial-omics and FTU morphometrics. One of the initial steps in our processing pipeline is to extract boundaries of FTUs using one of several DL-based algorithms (Methods). After automated segmentation, morphometrics properties are calculated for each structure prior to performing spatial aggregation of intersecting spatial-omics data. For image-based spatial-omics data (as in *Cell DIVE* and *PhenoCycler*), spatial aggregation consists of quantifying a summary of the intensity of each channel within the boundaries of a segmented structure. For more diffuse methods such as *10x Visium* and *10x Xenium*, we begin by identifying which spatial-omics units (spots for *Visium* and cells for *Xenium*) intersect with each segmented structure and then aggregate the -omics properties from those units using a weighted average. For the *10x Visium* and *Xenium* kidney biopsies, transcriptomic counts are translated into proportions of cell subtypes and states through a process known as cell deconvolution incorporating a large atlas of snRNA-seq data created by the Kidney Precision Medicine Project¹⁶.

On the Visualization page in *FUSION*, users can visualize pathomic or cellular feature abundance as color-mapped overlays on top of the tissue, enabling efficient comparisons across segmented FTUs or manually selected regions of tissue (Supplementary Fig. 1-3). Plots summarizing the distribution of user-selected properties across FTUs in the current slide, in the current viewport, and across multiple user-selected slides can be found in the *PropertyPlotter*, *PropertyViewer*, and *GlobalPropertyPlotter* components, respectively. In the *BulkLabels* component, users can employ a combination of spatial queries and property filters to apply labels to structures on the current slide, aiding in tasks such as cell annotation and pathological lesion assignment.

FUSION incorporates information from broader anatomical structure and cell type ontologies. This is facilitated in the form of interactive components such as the *HRAViewer*¹⁷⁻¹⁹, incorporating organ anatomical structure, cell type, and biomarker hierarchies, as well as in non-interactive utility functions which translate *Ensembl* gene IDs into more human readable names. The consolidation of these functionalities into a single platform facilitates more efficient data interpretation processes.

FUSION was designed with *FAIR* (findable, accessible, interoperable, reusable) principles in mind²⁰. Our approach focuses on modular engineering and systems-level design for computation, scalability, and expansion to enable future integration of spatial data from emerging technologies. Codes and documentation for *FUSION* can be accessed at the GitHub repository: <https://github.com/SarderLab/FUSION>. Additionally, the Python package *fusion-tools* is available through the Python Package Index (PyPI) with extensive documentation available at <https://github.com/SarderLab/fusion-tools>. Details on individual processing plugins may be found in the Methods section with links to the associated Slicer CLI-web Plugin.

Multi-modal comparison between disease and reference tissues

In the following sections, we present illustrative use cases of *FUSION* that integrate histology images and *10x Visium* ST data of kidney tissues from reference CKD, diabetic, and AKI patients. These data are analyzed separately using sections from 12 FFPE and 23 frozen section slides. For brevity, we only report analysis of FTU-level features in normal and sclerotic glomeruli, which are defined in the Supplementary Information Data description, including demographic and clinical data associated with the data, are made available in Supplementary Tables 1 and 2 as needed.

Structure-level analysis of digital pathology image data

At a granular level, *FUSION* enables the analysis of distributions of derived features within structures, and the examination of the spatial relationships between those structures. This type of query identifies structural phenotypes which are associated with certain cell types or histomorphological features and can provide insights into pathogenetic mechanisms. Users can select custom combinations of FTU features to analyze, generate figures, and perform statistical analysis according to biological hypotheses. For a single feature, violin plots are generated showing the distribution of the selected feature across user-selected labels present in the associated metadata. For two features, a scatter plot is generated to assess linearity between the two selected features. For more than two features, *FUSION* dimensionally reduces the assembled feature vectors into a Uniform Manifold Approximation and Projection (UMAP) which preserves relative sample variation²¹. Users can then use their mouse to select individual data points or "lasso" groups of points to view the image and cell composition associated with that point, or group of points, on the current plot. This type of visualization allows users to associate the quantitative distribution of features with other qualitative findings in individual FTUs. In addition to hypothesis testing, these features can also result in discovery as the clustering is data driven and agnostic and therefore

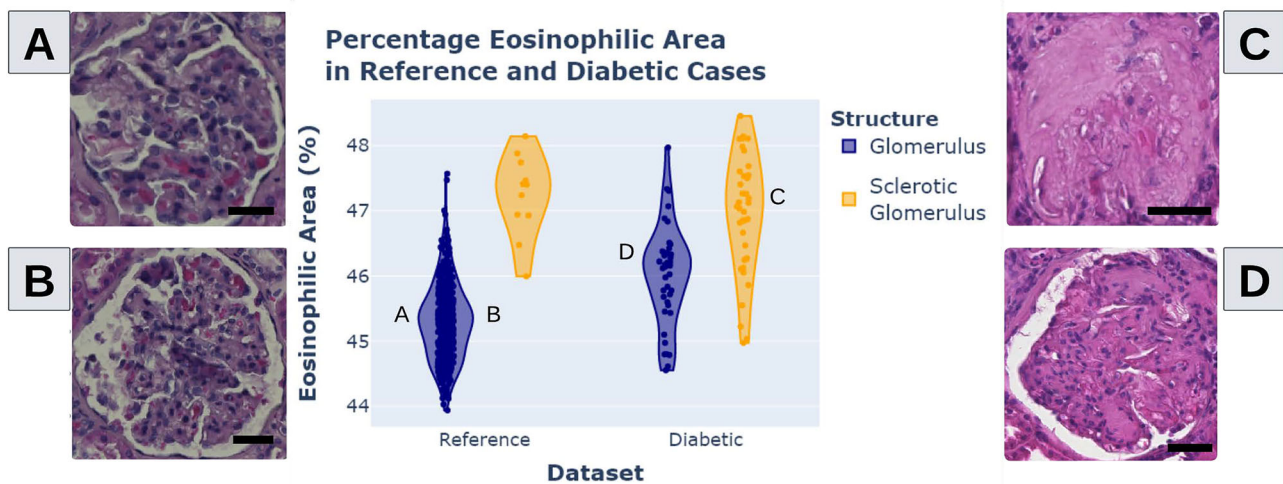


Fig. 2 | Local comparison of digital histology images using a single feature. As an example, glomerular hypertrophy is compared in diabetes and reference tissue. Percentage eosinophilic area between normal and sclerotic glomeruli from diabetic and reference kidney samples. **A–D** Example glomeruli selected from the feature distribution as indicated. Dots represent individual glomeruli. Sclerosis increases

eosinophilic area by reducing open capillary lumens and Bowman's space. Diabetic kidney disease, in the absence of sclerosis, increases eosinophilic area through hypertrophy and expansion of the mesangium. Scale bar indicates 25 μm . Two-sided unpaired Student's *t* test. Data reflect $n = 3$ independent replicates per group. Source data are provided as a Source Data file.

can reveal unanticipated relationships using morphology and molecular datasets.

To demonstrate *FUSION*'s ability to succinctly and accurately report on histopathology and molecular features, we examined glomerular hypertrophy in reference and diabetic kidney disease (DKD) samples. Glomerular hypertrophy, or the inordinate growth in tissue, is a manifestation of DKD resulting from mechanical and chemical stressors²². To assess for hypertrophy, we measured the area of eosinophilic tissue within each glomerulus (i.e. the histological structure). In glomeruli, these eosinophilic regions comprise cell cytoplasm, mesangial matrix, and basement membranes. As expected, diabetic and sclerotic glomeruli depicted higher eosinophilic area (Fig. 2), coinciding with both the physical enlargement of the glomerulus in response to adaptive hyperfiltration, mesangial expansion, and the loss of both Bowman's space and open capillary lumens in the setting of glomerulosclerosis^{22,23}.

To demonstrate structure-level analysis using multiple morphometric features, we next examine morphological features related to the relative thickness of each sub-compartment. These features are measured using the distance transformation, or the pixel-wise distance from one pixel in a sub-compartment to a pixel of another sub-compartment. Assembled features are initially scaled so that each feature has zero mean and unit standard deviation prior to UMAP dimensionality reduction. The resulting scatter plot is then dynamically relabeled in *FUSION* so that the color of each point corresponds to either disease or structure phenotypes (Fig. 3). These plots may be used to assess data heterogeneity within respective phenotypes and to make inferences on key features within a particular group.

One of the initial observations derived from the UMAP plot in Fig. 3A, B is that the sclerotic glomeruli predominantly group on the left-hand side along with non-sclerotic glomeruli from diabetic cases. To quantify this observed clustering behavior, silhouette scores are automatically calculated in *FUSION* and presented alongside other statistical measures and feature summaries in the *Plot Report* component. For the UMAP containing all glomeruli, silhouette scores of 0.22 and 0.06 are found for the diabetic and reference datasets, respectively, indicating a relatively stronger intra-class grouping of diabetic glomeruli compared to those from reference patients. This

observation hints at significant heterogeneity in reference glomeruli, while diabetic glomeruli are concluded to be more homogeneous in this dataset.

FUSION enables the assessment of both inter- and intra-subject variability. For example, we observed within the UMAP that the glomeruli from one reference tissue section grouped together more closely to each other than other members of the reference group (review using Fig. 3C). After selecting these points using the lasso tool, the presence of several red (possibly erythrocytes/RBCs) inclusions are identified within glomerular capillary lumens and are likely artifact related to tissue preparation. It is common that histopathologic specimens will vary in their preparation, staining character, and image quality. *FUSION* can help identify artifacts so users can assess whether there is biological relevance or whether artifacts should be excluded or corrected in their analyses. See Supplementary Fig. 4. for slide labels for each glomerulus in Fig. 3. See Supplementary Fig. 5 for another similar example.

By selecting other points in the UMAP plot, we can begin to make qualitative conclusions on the relative ordering of glomeruli in the high-dimensional scatter (Fig. 3C). Specifically, points sampled from right to left on the lower edge of the UMAP plot seem to depict gradual changes from normal glomeruli to fully sclerotic glomeruli. Furthermore, glomeruli towards the upper side of the UMAP plot seem to have a larger overall area. The tool allows discovery of different sub-groups of glomeruli including large sclerotic, small sclerotic, small dense, large dense, and generally normal glomeruli. Importantly, because *FUSION* integrates the molecular features (spatial transcriptomics), one can survey molecular expression in these different morphological variants and infer the molecular changes that may underlie these transitions.

The included examples show *FUSION*'s ability to conduct spatial data QC, study data heterogeneity, and find patterns using an interactive approach applicable to disease and reference kidney biopsy sections (Figs. 2–3).

Local Comparison of Spatial Molecular Data

After using *FUSION*'s analysis of image morphometrics, our next focus is on the analysis of spatial molecular data and the demonstration of multi-modal data integration.

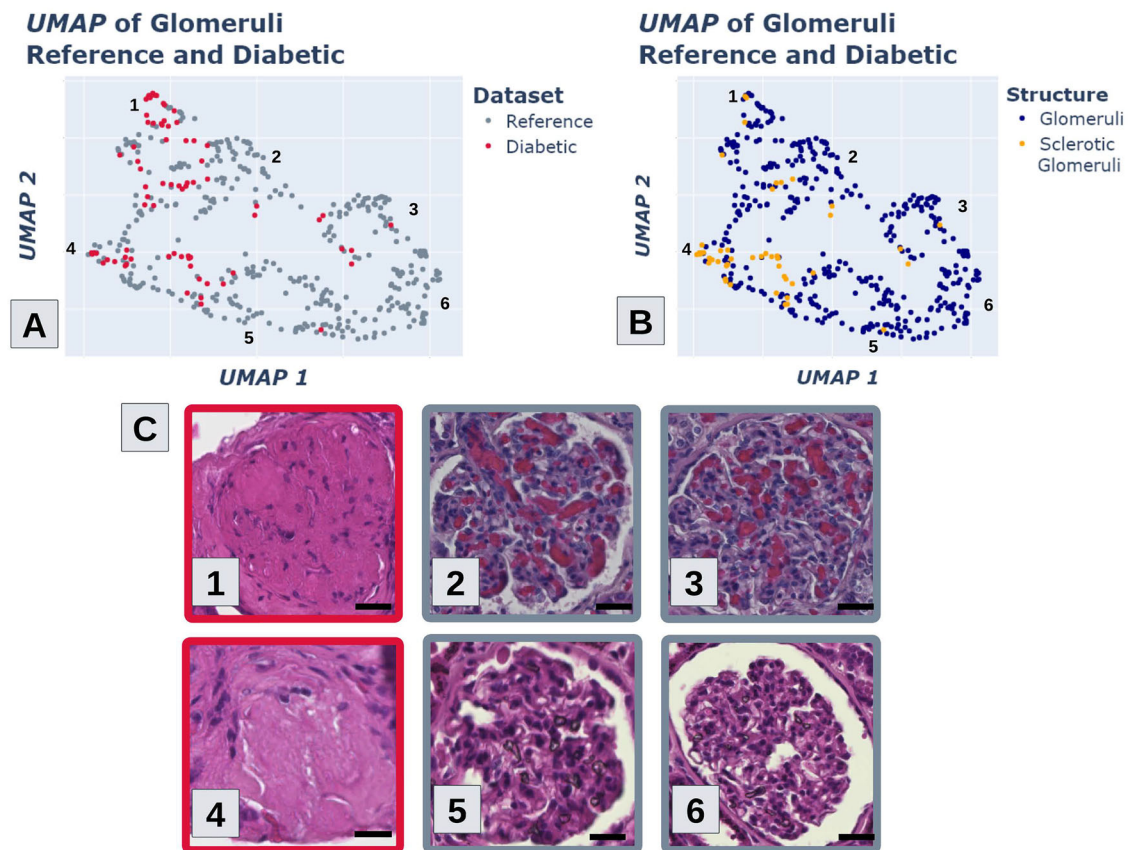


Fig. 3 | Local comparison of digital histology images using multiple features. Uniform Manifold Approximation and Projection (UMAP) projection of glomerular sub-compartment distance transform features indicating (A) whether each glomerulus is from a diabetic or reference slide, (B) whether each glomerulus is normal or sclerotic, and (C) images taken from numbered locations in the plots above.

Several distinct sub-groups of glomeruli are observed, including large sclerotic, small sclerotic, small dense, large dense, and generally normal glomeruli. Scale bar indicates 25 μm . UMAP clustering shown; silhouette scores calculated per dataset. Images are representative and reproducible across independent samples. Data reflect $n = 3$ independent replicates. Source data are provided as a Source Data file.

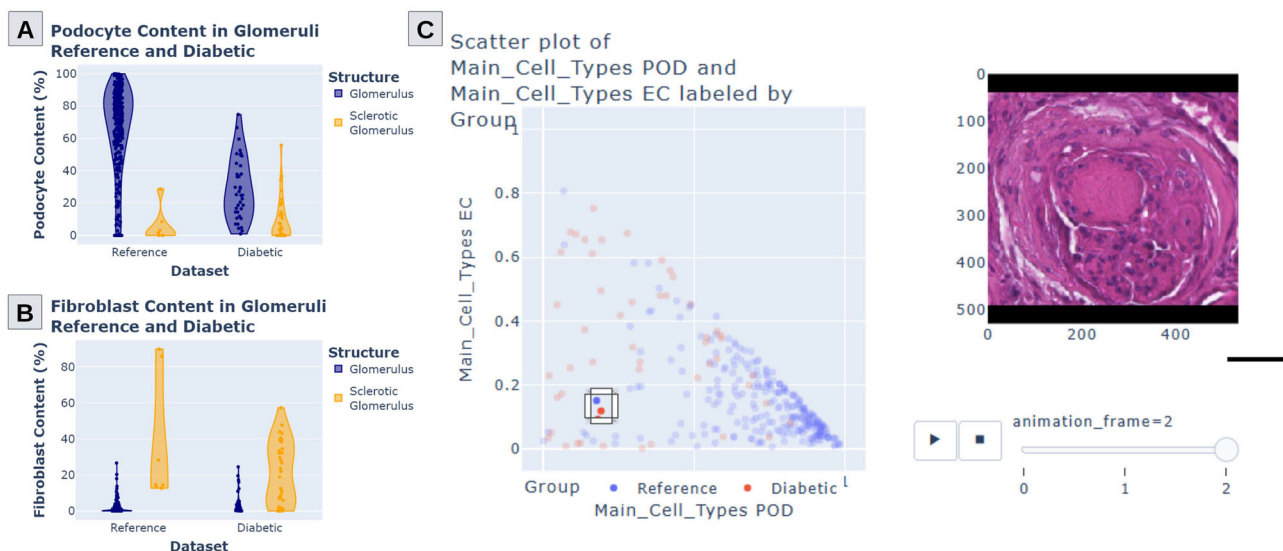


Fig. 4 | Local comparison of spatial molecular data. Spatial distribution of major cell types in glomerular functional tissue units in diabetic and reference datasets. **A** Podocyte proportion in normal and sclerotic glomeruli from reference and diabetic cases. **B** Fibroblast proportion in glomeruli and sclerotic glomeruli from reference and diabetic cases. **C** FUSION example plotting podocyte proportion against endothelial cell proportion, with a representative diabetic glomerulus

showing both low podocyte and endothelial cell content. Each dot represents a glomerulus. Scale bar indicates 25 μm . Two-sided unpaired Student's t test. Data reflect $n = 3$ biological replicates and per group as applicable. Representative glomeruli are reproducible across independent tissue sections. Source data are provided as a Source Data file.

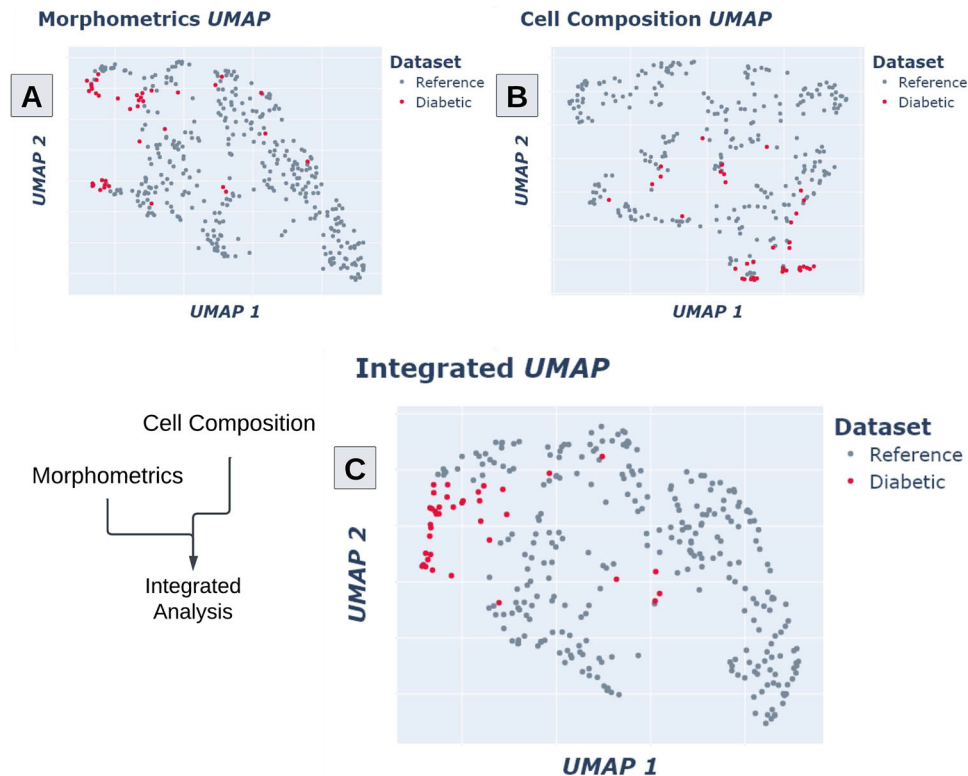


Fig. 5 | Multimodal data integration in FUSION combining image features and cell abundance data derived from spatial-omics. Glomerular features from reference and diabetic samples were used to generate Uniform Manifold Approximation and Projection (UMAP) plots. **A** UMAP generated using only morphometric

features. **B** UMAP generated using only cell composition features. **C** UMAP generated using a combined set of morphometric and cell composition features. Each dot represents a glomerulus. Source data are provided as a Source Data file.

Among the primary cell types of the glomerulus are podocytes and endothelial cells. As described above, normal glomeruli exhibit a significantly higher proportion of podocyte transcript expression when compared to diabetic glomeruli ($p = 7.8376E-18$; two-sided unpaired Student's t test; Fig. 4A). Using *FUSION*'s plotting tool, lasso selection tool, and relative cell composition pie charts, a biomedical scientist user can confirm known changes in cell type composition of diabetic glomeruli, wherein the proportion of podocyte to endothelial cell signature shifts (Fig. 4). This shift reflects two phenomena: a reduction in podocyte transcript expression related to injury and effacement and an increase in glomerular capillary endothelial cell (EC-GC) transcript expression, potentially resultant from neovascularization in diabetic kidney disease²⁴. Furthermore, a significant difference was found in fibroblast cell fraction between sclerotic and normal glomeruli from both diabetic and reference tissue sections (Fig. 4B).

Local Comparison Integrating Pathomic Image Features and Spatial Molecular Data

Integrative analysis is achieved in *FUSION* by clustering combined morphometric and molecular features. The UMAP in Fig. 5A was generated using only morphometric features from the distance transform and morphological categories. Diabetic samples are loosely grouped toward the left-hand side of the plot, with some glomeruli intermixed with the reference group. Two separate trends are apparent along the spectrum of reference glomeruli (right) to diabetic glomeruli (left), indicating two sub-groups of diabetic glomeruli with distinct morphological appearances.

When evaluating glomerulus cell composition features (including non-tubule cell types such as podocytes, aggregate endothelial cells, mesangial cells, immune cells, and fibroblasts) (Fig. 5B), we observe

heterogeneity in the reference glomeruli, resulting in several clusters (upper portion of the plot). Diabetic glomeruli were intermixed, with a subpopulation of glomeruli grouped (bottom right).

Integrating these two modes of data resulted in the strongest separable clustering of diabetic and reference glomeruli. Diabetic glomeruli were grouped towards the upper left-hand portion of the resulting UMAP plot while reference glomeruli were spread between the middle to lower-right portions with some intermixing of the two groups in between (Fig. 5C). By combining morphometrics and cell composition features, silhouette scores for glomeruli in the diabetic group increased from 0.35 with just morphometrics and 0.37 with just cell composition to 0.53. *FUSION* integrates a common analytical technique found in the *Seurat R* package, *Cluster Markers*, to determine key features which separate different groups into dimensionally reduced clusters. This backend implementation offsets the computational costs of this process away from the visualization server, greatly increasing scalability of this process. By selecting the "Find Cluster Markers" button, users can send all the data that is used to generate the current plot to the backend server to determine representative features for each group and their adjusted p -values. In the diabetic and reference glomeruli, the significant discriminative features for glomeruli from the diabetic group include both mean and maximum distance transform values for eosinophilic regions (normalized by object area), increased contents of fibroblasts and immune cells, and nuclear morphology. Specifically, some measures of nuclear aspect ratio indicated a departure from circular nuclei to a more elliptical shape. These features consistently demonstrate a change from a normal glomerulus phenotype to that of a typical diabetic glomerulus, often associated with fibrosis, inflammation, and changes in cellular composition. Glomeruli from the reference group exhibited a higher

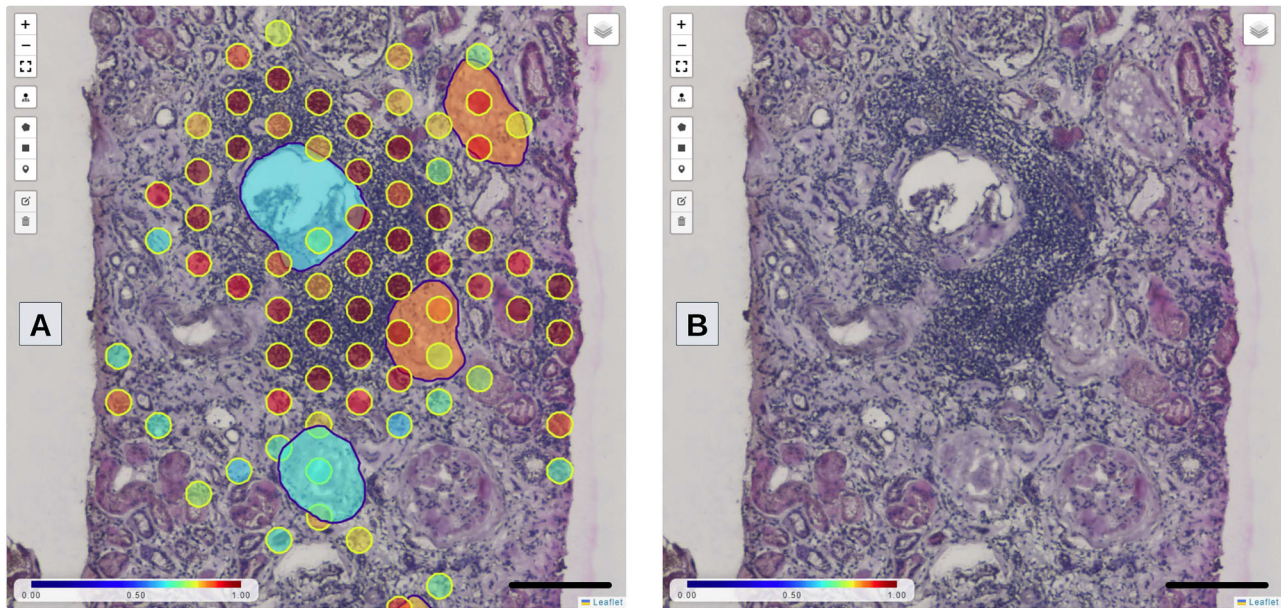


Fig. 6 | Spatial visualization of immune cell infiltration. **A** Screenshot of a chronic kidney disease section in *FUSION* with the heatmap indicating the proportion of immune cells. **B** Same view as in (A) with annotations turned off to better

show the underlying histology. Scale bar indicates 200 μm . Representative spatial visualization shown; image is reproducible across multiple annotated regions from independent tissue samples.

proportion of podocyte transcript expression and reduced fibroblast transcript expression. For reproducibility, we repeated the clustering of Fig. 5 using slide identity as labels of the respective glomeruli and found the expected cluster patterns (Supplementary Fig. 6–8).

Spatial visualization of immune cell infiltration

Thin FFPE sections generally have improved morphology over thick frozen sections. Despite this, *FUSION* has the sensitivity to distinguish some morphologic features in frozen tissue. In the next use case, 23 kidney samples prepared using a frozen tissue section protocol are derived from individuals with AKI ($n_a = 6$), CKD ($n_c = 11$), and reference ($n_r = 6$). The presence and characteristics of inflammation are known to vary greatly between these conditions^{25–28}. AKI, being an acute pathology, often features acute inflammation in the form of infiltrating lymphocytes, but fewer chronic fibrotic changes^{25,26}. CKD, by contrast, shows evidence of continual injury over a longer time, manifested as fibrosis throughout the tubulointerstitium in the underlying tissue^{29,30}. The reference tissues may feature some age-related glomerulosclerosis or tubulointerstitial fibrosis but typically have minimal inflammation^{31,32}.

In Fig. 6A, an example depicts the *10x Visium* spots colored according to immune cell content and glomeruli with greater than 30% immune cell composition. Strong localization of immune cells is seen around a central glomerulus. Fully sclerotic glomeruli in the lower portion of the image exhibit comparatively less immune cell infiltration. This use-case illustrates that *FUSION* is able to identify FTUs and regions of inflammation even with the sub-optimal morphology of a frozen section.

Integrating Pathomics and Cell Composition to Study the Impact of Inflammation

Morphological properties can be selected in the clustering tab of *FUSION* along with non-tubule cell types to generate UMAP visualizations (Fig. 7). Reference glomeruli clustered distinctly from those of diseased samples (Fig. 7A). Cluster markers for these groups revealed statistically significant increased content of immune cells in both AKI and CKD groups compared to the reference glomeruli, while the

fibroblast content helped to distinguish CKD glomeruli from the AKI glomeruli. *FUSION* enables users to determine the key features driving clustering (Fig. 7B). In the *Reference* group, the primary source of heterogeneity between glomeruli stemmed from the different proportions of both podocytes and endothelial cells. Both cell types are crucial to normal glomerular function. As expected, the podocyte content of glomeruli in the AKI group (mean = 39.46%) was not substantially lower than that of the reference group (mean = 49.1%), although the difference was found to be statistically significant ($p = 1.5017\text{E}-02$; Tukey's Honestly Significant Difference test; Supplementary Fig. 9). This finding and possible explanation is elaborated in Supplementary Fig. 9 and in related text. By contrast, more immune cell infiltration was found in AKI glomeruli (mean = 16.1%) compared to reference glomeruli (mean = 1.9%). An increased proportion of fibroblast transcripts were found in CKD glomeruli (mean = 10.9%) compared to both AKI (mean = 6.9%) and reference (mean = 3.2%) which aligns with the increased fibrosis observed in these samples (Fig. 7C, D)^{25,26}. See in Supplementary Fig. 10 the same UMAPs in Fig. 7A, B but using slide identity phenotypes as labels of the respective glomeruli.

Application of *FUSION* to other spatial-omics technologies and organs

We next present use cases of *FUSION* for other spatial-omics technologies and organ systems to demonstrate the generalizability of the platform. Supplementary Fig. 11 depicts data from *10x Visium* for a healthy uterus sample from the HuBMAP data portal. We also discuss applications for data from *10x Xenium*, *Cell DIVE*, and *PhenoCycler* with their respective overlaid histology below, and from recently released *Visium HD* technology in Supplementary Fig. 12.

10x xenium

As opposed to *10x Visium*, *Xenium* enables single-cell specificity for spatial transcriptomics, increasing spatial resolution³³. The total number of measurable transcripts in this dataset was ~300 (compared to ~17k in *Visium*), however, similar methods were applied to deconvolve the observed transcripts into predicted cell types and states. These resulting cell subtypes were aggregated to intersecting

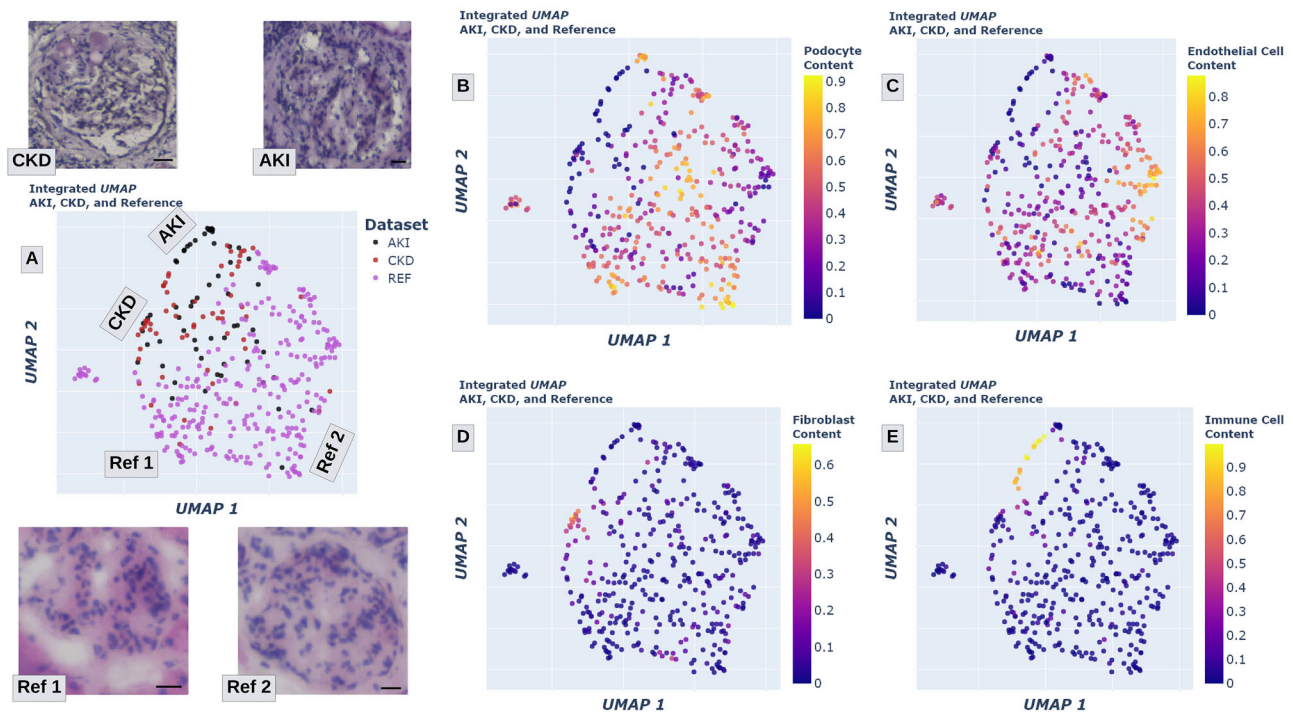


Fig. 7 | Integration of pathomics and cell composition to study the impact of inflammation. **A** Uniform Manifold Approximation and Projection (UMAP) plot generated using combined non-tubule cell type composition and morphologic features of glomeruli from acute kidney injury, chronic kidney disease, and reference groups. Selected glomeruli shown from indicated locations. Each dot

represents a glomerulus. **B–E** Same UMAP as in **(A)** with each dot colored by the content of **(B)** podocyte, **(C)** endothelial cell, **(D)** fibroblast, and **(E)** immune cell transcripts. Scale bar indicates 25 μm . Source data are provided as a Source Data file.

histological structures, showing *FUSION* can achieve summary cell compositions from single-cell input data.

For this use case, we examined two kidney sections. One section was from a patient with diabetic kidney disease and a second was from a healthy patient without diabetes. As in previous examples, one of the main hallmarks of diabetic kidney disease progression is podocyte depletion, followed by the formation of chronic lesions (i.e., nodular mesangial sclerosis) in key renal structures such as the glomeruli, tubules, and the interstitium. In the healthy glomerulus (Fig. 8A, B), the primary cell types consisted of podocytes, endothelial cells, and parietal epithelial cells while diabetic glomeruli (Fig. 8C, D) featured an increased quantity of fibroblasts and immune cells. Sclerotic glomeruli of the diabetic patient also depict similar cell composition. Furthermore, immune cell infiltration was observed in the diabetic case, spatially localized around some non-sclerotic glomeruli (Supplementary Fig. 13).

FUSION integrates images and transcriptomics by spatially variable genes (SVGs)

Spatially variable genes (SVGs) are biologically significant due to their differential expression across distinct regions or cell types within a tissue³⁴. Linking histologic regions of interest to enriched pathological annotations of SVGs is useful in helping biologists and pathologists understand the progression of disease.

SVG analysis was performed by integrating the Big Small Patch (BSP) package with *FUSION* on eight *10x Xenium* kidney samples³⁵. Enriched SVG function ontologies were identified using the SVGs in two conditions: CKD where 154 enriched genes contributed to ontology mapping and healthy reference where 132 genes contributed to ontologies (Supplementary Table 3-5). Leukocyte migration was a significantly enriched gene ontology term (GO_0050900) (q value 1.09E-09) active in the CKD gene set. The activity of this functional ontology was calculated by gene expression and mapped to histologic

images to identify regions of epithelial-immune crosstalk. Figure 8E shows a section of a CKD sample displaying regions with high expression of leukocyte migration using *FUSION*. *Region I* is an area of interstitial fibrosis with ontology mapping of leukocyte migration in immune cells. In contrast, *Region 2* encircles an area of interstitial fibrosis with infiltrating immune cells and upregulation of leukocyte migration within the epithelial nuclei of tubules. To showcase the capability of *FUSION* to enhance histology analysis with SVG function ontology information from sequencing, the CKD sample herein was uploaded to *FUSION*'s interactive viewer along with the pathway expression overlay (Supplementary Fig. 14). Figure 8F displays a section of the CKD sample enriched with the leukocyte migration (GO_0050900) ontology mapped. *FUSION* visualization highlights the cells of interest and displays a clear view of the sample histology. The built-in selection tool allows highlighting of regions of interest. *Region I* points to two sections, both indicating upregulation of leukocyte migration within the epithelial nuclei of the highlighted tubules. *Region II* highlights an area of interstitial fibrosis that contains infiltrating immune cells. With additional overlays of tubules, arteries, and glomeruli boundaries, *FUSION* simplifies and accelerates the process of identifying areas of high SVG domain expression and linking the underlying information to the histology.

Multiplexed immunofluorescence (MxIF) and PhenoCycler

PhenoCycler is a multiplexed immunofluorescence (MxIF) imaging method that allows for single-cell resolution of spatial protein expression detected by DNA-conjugated antibodies across the tissue. Although the number of available markers is limited (up to 100 proteins), and additional markers require careful optimization, PhenoCycler and other MxIF methods (MIBI—multiplexed ion beam imaging, Cell DIVE, CycIF—cyclic immunofluorescence, etc.) enable researchers to achieve detailed views of underlying cell types and interactions between neighboring cell types. Furthermore, since *PhenoCycler* is not

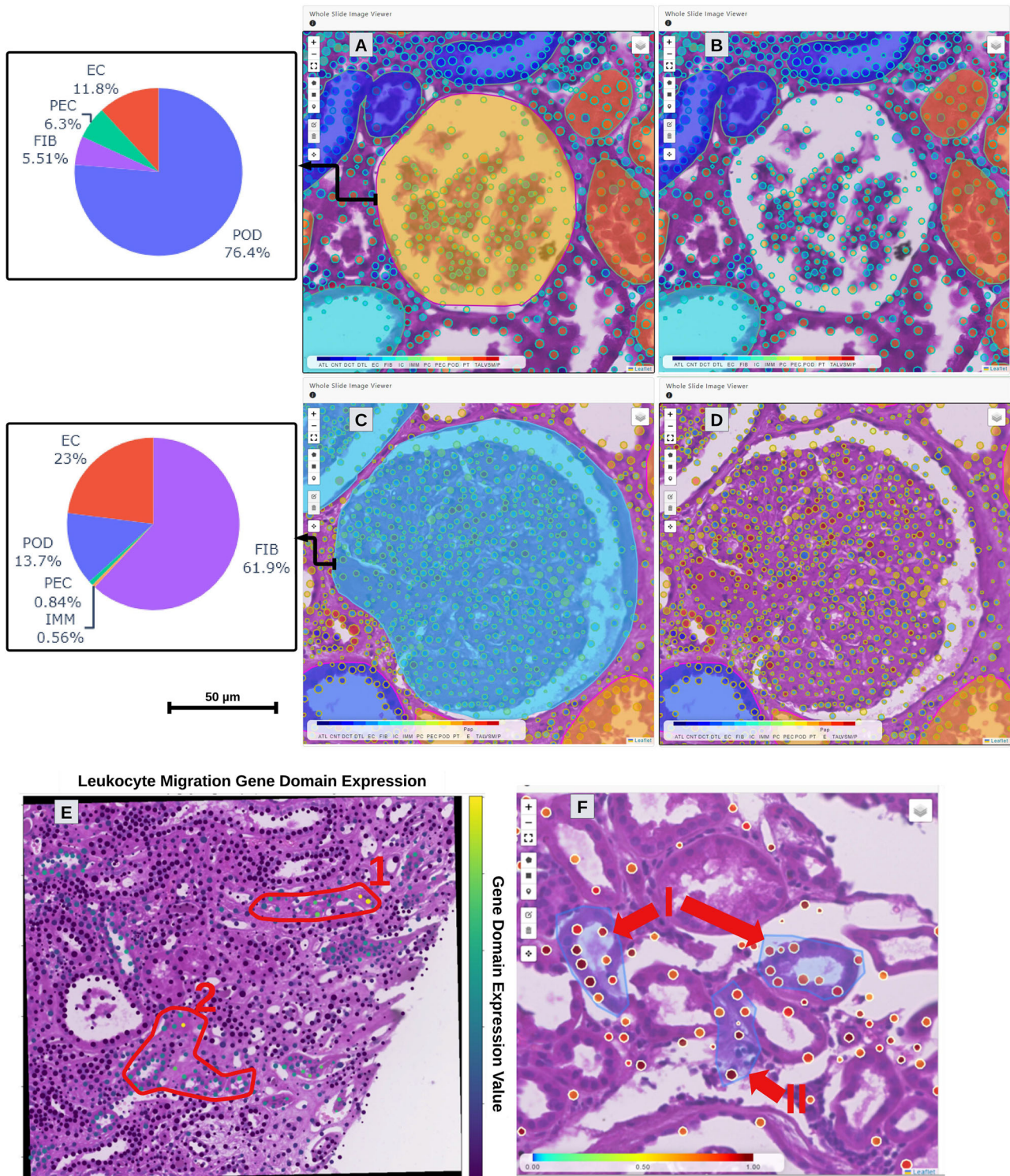


Fig. 8 | Application of FUSION for 10x Xenium data. A, B Glomerulus from a healthy patient with automatically segmented cells and functional tissue units overlaid. **B** Same view as in (A) with the glomerular overlay turned off to highlight segmented cells. A high proportion of detected cells (76.4%) in this glomerulus were identified as podocytes. **C, D** Glomerulus from a diabetic patient with segmented cells and functional tissue units overlaid. **D** Same view as in (C) with the glomerular overlay turned off to highlight segmented cells. Compared to the reference glomerulus, this glomerulus contains a higher proportion of fibroblasts and endothelial cells. Scale bar indicates 50 μ m. Color bar indicates the most abundant cell type in each structure. **E** Spatially variable gene analysis highlighting

leukocyte migration (Gene Ontology term GO_0050900) pathway expression aligned with the histology image in a chronic kidney disease sample. Region 1 shows immune infiltration overlying interstitial matrix expansion. Region 2 denotes immune infiltration with upregulation of the pathway in epithelial nuclei. **F** Visualization of the same leukocyte migration pathway in a chronic kidney disease sample in FUSION with a minimum threshold of 0.35. Region 1 shows upregulation in epithelial nuclei, while Region 2 indicates ontology mapping to immune cells in the interstitium. FUSION ontology mapping helps identify regions of epithelial-immune crosstalk.

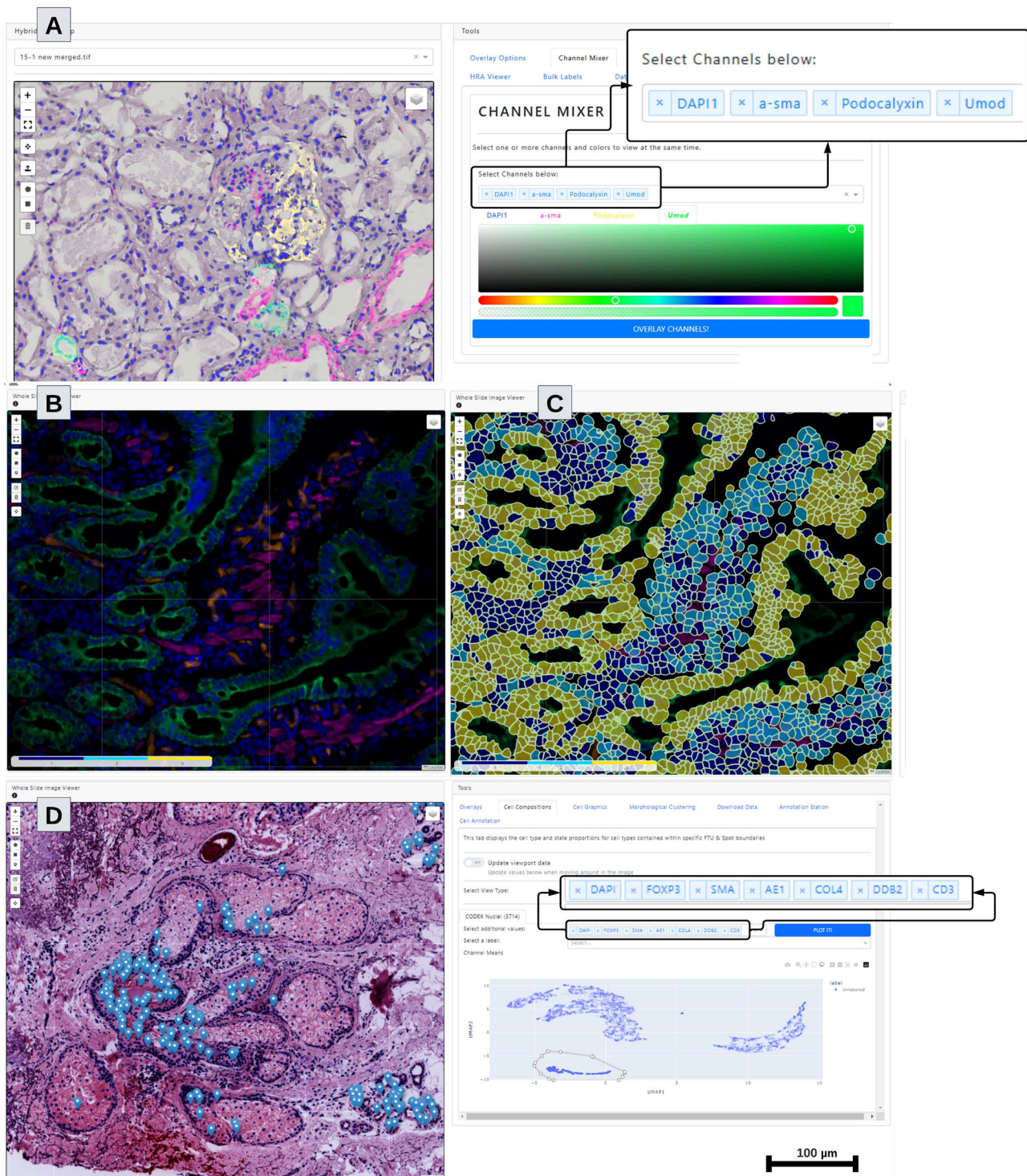


Fig. 9 | Application of FUSION for multiplexed immunofluorescence data.

A Kidney sample with PhenoCycler channels overlaid in multiple colors on a registered hematoxylin and eosin (H&E) image. **B** Large intestine sample showing selected PhenoCycler channels viewed together. **C** Overlaid cell segmentations with colors corresponding to unsupervised cluster assignments for the same tissue

section as in **(B)**. **D** Skin CELL DIVE sample with dynamic clustering of selected channel means for each cell in the current view. By selecting groups of points in the generated Uniform Manifold Approximation and Projection (UMAP) plot, users can visualize where specific cells appear in the corresponding histology or multiplexed immunofluorescence image. Scale bar indicates 100 μm.

a destructive technique, histological staining (H&E) may be performed on the same section (Fig. 9A) post-PhenoCycler³⁶. The addition of H&E-stained tissue allows for direct association between biomarker expression and tissue morphology. One challenge of adding H&E staining to a PhenoCycler experiment is that it will often require image

registration, as tissues may be shifted, shrunk, or damaged by additional processing. As it is important to achieve accurate cell-to-cell alignment for proper interpretation of these paired datasets, we developed a specific registration procedure that combines classical image analysis techniques and established image registration

algorithms. More information on the cloud implementation of this registration pipeline is provided in the Methods section. Pixel-level fluorescent marker expression data is stored as a 2D channel for each marker in a *MxIF* sample. Individually, these channels are grayscale representations of the fluorescence intensity at each pixel ranging from black (no expression) to white (high expression or saturation). Combined visualization of multiple channels is therefore aided by incorporating different colors to each channel. In *FUSION*, users can dynamically select which *MxIF* channels they would like to see at the same time as well as the color they would like that channel to be (Fig. 9A, B). This color flexibility allows users to view each channel with personalized degrees of contrast depending on an individual's color vision range.

Cell type annotation is one of the preeminent challenges in high-resolution tissue image analysis. Through the use of spatial-omics techniques with single-cell resolution, such as *PhenoCycler*, *Cell DIVE*, and *10x Xenium*, it is now possible to achieve well-informed cell annotations across an individual piece of tissue. The reproducibility of these cell type annotations relies on using established ontological frameworks like the HRA¹⁷. *FUSION* enables preliminary annotation of cell types based on the intensity of different *MxIF* channels within segmented nuclei (Fig. 9D). These annotations can then be exported directly from the *FUSION* interface for downstream analysis. See Supplementary Fig. 15–18 for more use-case examples of using *FUSION* for *PhenoCycler* and *Cell DIVE* data overlaid on histology for multiple organ systems.

Discussion

We present *FUSION*, an integrated visualization, analytic, and interactive platform for spatially resolved molecular data that combines histomorphological evaluation with spatial-omics. *FUSION* empowers a wide variety of users to incorporate spatial-omics into their workflows by providing them with user-friendly tools to reveal underlying pathobiological mechanisms. This process serves to increase the accessibility of quantitative computational analyses to a larger user base, which normally requires specific skills.

Continuously linking derived data with the source histology image allows for broad associations of quantitative features with established pathological lesions. For a clinical pathologist or researcher, this type of interaction enables more direct comparison of segmented structures from one tissue section with those in other sections. When presented with information from many tissue sections, the extensive data visualization methods implemented in *FUSION* provide supportive contextual information which can help clarify ambiguity associated with disease severity grading. This is especially true if longitudinal data are available for those other tissue sections in the dataset. An additional benefit of continuously enabling image and quantitative feature association is that it allows users to validate whether a given image morphometric or textural feature is robust in distinguishing different types of structures or diseases. Outlier detection is also possible by identifying data points that correspond to images that contain tissue artifacts, such as folds, tears, or increased presence of RBCs, which can interfere with accurate feature calculation. By enabling users to interact with their data at both the level of a single tissue section as well as for a dataset of many tissue sections, the efficiency and rigor of biological discovery is enhanced. One strength of *FUSION* is its extensibility to other modes of spatial-omics. In this work, we present several examples using *10x Visium* spatial data which consists of many ROIs covering the tissue. To effectively translate this data to provide users with structural cell composition, we implemented specialized analytical methods for spatial aggregation of molecular data. These aggregation techniques can be expanded for technologies that provide single-cell resolution such as *PhenoCycler*, *Cell DIVE*, *10x Xenium*, and *Visium HD* providing direct insights into the cellular composition of complex FTUs^{37–39}. Moreover, our tool is generalizable for any organ systems with spatial and histology data (Fig. 9

and Supplementary Fig. 11), and such data can be accessed through the HuBMAP portal⁴⁰.

It is noteworthy to discuss one limitation of *FUSION*. Due to the requirement for large, annotated, single-nucleus RNA-sequencing (snRNAseq) datasets, cell deconvolution is reliable for kidney samples¹. We have implemented one reference-free method for cell type deconvolution (Methods) which should accommodate examination of data from other organs without expert annotations. The use-cases presented in this manuscript serve to highlight the various functionalities of *FUSION* to integrate multi-modal data and are not powered to definitively identify biologic insights.

Through integrations with DSA, *FUSION* enables users to run complex and computationally intense machine learning algorithms on their own data. The *Slicer CLI* plugin framework provides computational researchers with a method to share their pipelines with the community in accordance with FAIR principles²⁰. Future iterations of *FUSION* will primarily focus on developing plugins for feature extraction, segmentation, and analysis of spatial-omics data.

Methods

Ethics statement

This study complies with all relevant ethical regulations and was approved by the Institutional Review Board at the University of Florida. All human tissue samples analyzed were obtained under IRB-approved protocols from contributing institutions. Samples were fully de-identified prior to access and analysis. Informed consent was obtained by the contributing institutions, and no compensation was provided to participants.

Statistics and reproducibility

Statistical results presented in this work are descriptive in nature and are not intended to reflect an underlying experimental design. No statistical method was used to predetermine sample size. No data was excluded from the analyses. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

No sampling was performed from any distribution, and labels for individual structures (e.g., spots, glomeruli, sclerotic glomeruli), as well as labels for slides, were not blinded to investigators. Each experiment involving biological tissue was repeated independently at least three times with consistent results, unless otherwise noted. All statistical tests used, including test type, sidedness, and *p*-values, are specified in the figure legends. Representative examples are clearly indicated in the text and captions. No inferential claims are made beyond the demonstration of platform functionality.

Datasets

Input data structure. *10x Visium* data used in the development of *FUSION* consisted of two separate types of files as input into the pre-processing steps. The initial set of files was the high-resolution scanned Whole Slide Images. Whole Slide Images varied in file extension depending on the scanner brand used or due to downstream post-processing (OME-TIFF). Reading these different types of image files can be accomplished through a number of open-source libraries including *OpenSlide*, *Tiffslide*, and *large-image* which is developed by *Kitware*^{41,42}. These libraries are capable of opening and extracting image regions for a large variety of file formats. *FUSION*, through DSA, uses the *large-image* Python library to read uploaded WSIs, and generate tile sources which are then exposed through an endpoint in the DSA WebAPI⁴².

The second set of files contains transcriptomics read counts and spot center coordinate information. These files can either consist of a single RDS file that is the output of the *Load10x_Spatial()* command in *Seurat* (v4.0), the filtered features matrix, tissue positions, and scale factors files that are output by *10x's spaceranger* tool, or an AnnData-

Table 1 | Breakdown of main cell types, subtypes, and their states

Main Cell Types	Sub-Types	States
Ascending Thin Limb (ATL)	ATL, dATL	Reference, degenerative
Connecting Tubule (CNT)	CNT, CNT-PC, dCNT, cycCNT	Reference, degenerative, cycling
Distal Convolutd Tubule (DCT)	DCT, DCT1, DCT2, dDCT, cycDCT	Reference, degenerative, cycling
Descending Thin Limb (DTL)	DTL, DTL1, DTL2, DTL3, dDTL3	Reference, degenerative
Endothelial Cell (EC)	EC, EC-GC, EC-AEA, EC-DVR, EC-PTC, dEC-PTC, EC-AVR, dEC, cycEC, EC-LYM	Reference, degenerative, cycling
Fibroblast (FIB)	FIB, MYOF, cycMYOF, M-FIB, dM-FIB, aFIB, dFIB	Reference, cycling, degenerative, adaptive
Immune (IMM)	IMM, B, PL, T, NKT, MAST, MAC-M2, cycMNP, MDC, cDC, pDC, ncMON, N	Reference, cycling
Intercalated Cell (IC)	IC, C-IC-A, CCD-IC-A, CNT-IC-A, dC-IC-A, OMCD-IC-A, M-IC-A, tPC-IC, IC-B	Reference, degenerative, transitioning
Neural-like Cells (NEU)	NEU, SC/NEU	Reference
Papillary Tip Epithelial (PapE)	PapE	Reference
Principal Cell (PC)	PC, C-PC, CCD-PC, OMCD-PC, M-PC, dOMCD-PC, dM-PC, IMCD, dIMCD	Reference, degenerative
Parietal Epithelial Cell (PEC)	PEC	Reference
Podocyte (POD)	POD, dPOD	Reference, degenerative
Proximal Tubule (PT)	PT, PT-S1, PT-S2, PT-S3, aPT, cycPT, dPT, dPT/DTL	Reference, adaptive, cycling, degenerative
Thick Ascending Limb (TAL)	TAL, aTAL1, aTAL2, M-TAL, dM-TAL, C-TAL, dC-TAL, MD	Reference, adaptive, degenerative
Vascular Smooth Muscle/ Pericyte (VSM/P)	VSM/P, MC, REN, VSMC, VSMC/P, dVSMC	Reference, degenerative

Kidney cell types and states determined using single cell kidney atlas produced by Lake et al.¹

formatted (h5ad) file that is the output of HuBMAP processing pipelines⁴³. Cell subtype proportions (Table 1) for each spot are generated using one of the methods for cell deconvolution described in the *Analytics* section^{17,44}.

Spot-level cell type composition data is stored as metadata within the structure annotation text file along with geometry. In *FUSION*, these are formatted following *GeoJSON* convention, where each spot is a *Feature* within a *FeatureCollection* for each slide. Following automated segmentation of select FTUs, cell type composition data is projected from the spots to intersecting FTUs using a weighted sum, where the weight for each spot is equal to the percentage of FTU-area occupied by that spot. Intersection tests, intersection area, and structural area are calculated using the *Shapely* package in Python.

Analytics

Cell deconvolution. Cell deconvolution for *10x Visium* spatial transcriptomic counts can be carried out in three different ways. If a user is uploading data from a kidney sample, label transfer is carried out using the reference atlas containing sequencing single-nucleus information from over 200,000 cells created as a part of the KPMP Atlas V1¹. Alternatively, users may select *STdeconvolve* as a method for cell deconvolution using an unsupervised approach without reference data⁴⁵. Cell proportions are represented as “*Topics*” which may be manually annotated later in a downstream process. Finally, users may also choose to use the reference object provided by Azimuth for cell deconvolution⁴⁶. This includes a wide variety of organs in the current implementation, although this method has not been optimized for usage on *10x Visium* samples.

Automated segmentation of FTUs. Segmentation of FTUs is performed using one of a few plugin implementations of DL deployed using the *Slicer CLI* Web schema. Each of these models includes PAS-stained kidney sections in their training data. More details on each of the models can be found in Supplementary Table 6. Our prior work *HistoCloud* outlines plugin creation using DSA^{42,47}.

Morphometric feature extraction. Morphometrics are calculated for each segmented FTU and added to per-structure metadata alongside

cell type composition data. These morphometrics can then be accessed either on an individual slide basis in the *PropertyViewer* component, taking into account spatial localization of morphometric properties, or compared across several slides in a Visualization session in the *GlobalPropertyPlotter* component. Morphometrics calculated in this step quantify the size, shape, color and texture of sub-compartments (lumina, eosin+ area, and nuclei) within each segmented FTU. These sub-compartments are segmented using user-defined thresholds during the upload process in an interactive procedure. A full description of these features is provided in Supplementary Table 7.

Cell segmentation (PhenoCycler). Cell segmentation in *FUSION* was performed using *DeepCell*⁴⁸. A key contribution to this existing model was facilitating the use of the *DeepCell NucleiSegmentation* model on large WSIs. For a given batch of equal sized patches, our plugin implementation of *DeepCell* generates nuclei predictions, post-processes results to limit the number of multi-nucleus clumps, and then extracts channel-wise means for use in cell clustering and feature visualization in *FUSION*.

Multi-modal image registration. Image registration was performed in an automated fashion by applying classical image analysis methods to both the input histology (H&E stained) and *PhenoCycler* images. By performing color deconvolution on the histology image, a grayscale image was generated indicating the presence of cell nuclei⁴⁹. This was then compared with the corresponding DAPI-channel of the *PhenoCycler* image. The Python package, *pystackreg*, was used to approximate the affine transform necessary to minimize the difference between these two masks⁵⁰. The *large-image-converter* library was used to apply the derived affine transformation and to then merge the two modalities into a single image for paired visualization⁵¹.

Cell phenotyping for xenium data. Due to the reduced number of available markers for *Xenium* sections compared to *Visium* (~300 compared to ~17k), a dimensionally reduced version of the kidney atlas generated in Lake et al. was employed¹. We suppressed degenerative

cell type mapping due to their over-representation in *Xenium* because of the lower quantity of canonical cell markers on the *Xenium* platform. Otherwise, the label transfer approach annotated cells in the spatial context using snRNA-seq labels in an identical manner to that of the *Visium* samples. Cell segmentation was conducted using *Xenium Explorer (v2.0)* using 10x proprietary methodology⁵². However, downstream FTU segmentation, morphometric calculation, and spatial aggregation were performed analogously to those of *Visium*.

Frontend

FUSION is built using *Dash*, a web-based visualization framework for Python applications^{53–55}. Components within *FUSION* are derived from either the core components library or several open-source community components including *Dash-Leaflet* (for large image visualization), *Dash-Bootstrap-Components* (for grid-style layouts, icons, and organization), *Dash-Extensions* (for integrating *JavaScript* functions), among others.

The visualization page (<https://fusion.hubmapconsortium.org/Visualization>) contains the majority of interactive components in *FUSION*. These are divided amongst WSI visualization and overlays, interactive plotting, cell graphics and hierarchy viewing, and data downloading. The WSI Viewer card contains the WSI, FTU/spot annotations and their associated overlays, and custom ROIs. Integrated *JavaScript* functions are used to control boundary color, fill color, transparency, and filtration of overlaid FTU and spot annotations according to user inputs. Annotations and image tiles are hosted in DSA and accessed through the annotations and item tiles endpoints. By logging in through *FUSION*, users can access their personal uploaded images. Although annotation correction is not implemented for FTUs in *FUSION*, users may correct predicted FTU boundaries using *HistomicsUI*, which is accessed through DSA at the host address. Combined morphometrics and cell composition plotting is available in the *GlobalPropertyPlotter* tab. In this tab, users can select features, labels, and filters to generate an interactive plot of data in their current dataset. For single features, violin plots of the raw data are generated. Selecting two or more features results in a two-dimensional scatter plot containing either raw data (for two features) or UMAP coordinates for dimensionally reduced data. Users can select one or multiple points to view the FTU or spot with that feature value. By selecting multiple points, an animation is generated that allows the user to scroll through several images or FTUs at a given location at a time (maximum 100 points, for efficiency) as well as view their mean cell composition. This visualization allows users to make broader connections for qualitative observations to the quantitative morphometric or -omics values present in each FTU. Users can also download the cell composition data for FTUs which are in the current view in either Comma Separated Values (CSV) or Excel (XLSX) format. These files will list each FTU in the current view and have a column corresponding to a particular cell.

From the dataset builder page (<https://fusion.hubmapconsortium.org/Dataset-Builder>), users can manually select which datasets they would like to examine in their current visualization session. We provide 56 pre-processed samples from both FFPE and frozen preparations (Supplementary Table 1) from both healthy, reference and disease groups. Once a user is satisfied with their collection of slides, they can then click the 'Visualization' button from the page menu in the navigation bar to be redirected to the visualization page where their selected slides can be viewed, and their clustering features accessed.

Processing workflows are accessed through the data upload page. Here, users can select what kind of data they would like to upload, upload the required files, and input parameters to use in processing plugins. All of which is accomplished using simple user-interface components with descriptive instructions throughout. Supplementary Fig. 19–21 elaborate on the layout of upload process, pipelines, and various available options for processing including different cell deconvolution methods.

Backend

FUSION works in concert with a running instance of DSA⁴². DSA is a containerized cloud platform which provides users with a secure, flexible, and scalable interface for the analysis of WSIs. Included in DSA is a rich RESTful API (Representational State Transfer Application Programming Interface) and Girder for data management operations. DSA (specifically the Web API) is used by *FUSION* to store pre-processed and uploaded data, run Slicer CLI Web plugins (specially formatted Docker images), access Whole Slide Image tiles for large image visualization at multiple levels of resolution, and manage user registration. Using the *Girder Client* Python library as well as the *requests* Python library, requests can be made to the DSA server from an external host. This means that the deployment of *FUSION* and its corresponding DSA instance can have separate resource allocations/specifications which fit what is expected of frontend and backend servers. For example, WSI segmentation jobs are accelerated using graphics processing units (GPUs), and cell deconvolution requires a significantly higher amount of random-access memory (RAM) compared to rendering visualization components on the frontend.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Source data are provided with this paper. Public datasets used in this study are available through the data portals of participating consortia, including the Human Biomolecular Atlas Program (HuBMAP; <https://portal.hubmapconsortium.org>) and the Kidney Precision Medicine Project (KPMP; <https://www.kpmp.org/available-data>). For KPMP data access, contact info@kpmp.org. For HuBMAP data access, use the "Contact Support" feature on their portal. Demonstration datasets for testing the data upload interface are deposited at the Open Science Framework (OSF) under <https://doi.org/10.17605/OSF.IO/U2ZY5>. Additional data including example 10x Visium and Pheno-Cycler images are available at <https://trailblazer.box.com/s/qzo3n9tam6pmo6igwsjtxi8k1lxjbj9>. Source data are provided with this paper.

Code availability

The source code for *FUSION* is available through GitHub at <https://github.com/SarderLab/FUSION> and is archived at Zenodo under <https://doi.org/10.5281/zenodo.15880605.54>. The *fusion-tools* Python package is available at <https://github.com/SarderLab/fusion-tools> and is archived at Zenodo under <https://doi.org/10.5281/zenodo.16002725.55>. Additional spatial -omics plugins, including those used for image registration and DeepCell-based segmentation, are hosted at <https://spatial-omics-plugins.readthedocs.io/en/latest/> and Docker Hub https://hub.docker.com/r/dsarchive/histomicstx_extras.

References

1. Lake, B. B. et al. An atlas of healthy and injured cell states and niches in the human kidney. *Nature* **619**, 585–594 (2023).
2. Bressan, D., Battistoni, G. & Hannon, G. J. The dawn of spatial omics. *Science* **381**, eabq4964 (2023).
3. Wu, Y., Cheng, Y., Wang, X., Fan, J. & Gao, Q. Spatial omics: Navigating to the golden era of cancer research. *Clin. Transl. Med.* **12**, e696 (2022).
4. Marx, V. Method of the Year: spatially resolved transcriptomics. *Nat. Methods* **18**, 9–14 (2021).
5. Williams, C. G., Lee, H. J., Asatsuma, T., Vento-Tormo, R. & Haque, A. An introduction to spatial transcriptomics for biomedical research. *Genome Med.* **14**, 1–18 (2022).

6. Lutnick, B. et al. An integrated iterative annotation technique for easing neural network training in medical image analysis. *Nat. Mach. Intell.* **1**, 112–119 (2019).
7. Keller, M. S. et al. Vitesse: integrative visualization of multimodal and spatially resolved single-cell data. *Nat. Methods* **22**, 63–67 (2025).
8. Hoffer, J. et al. Minerva: a light-weight, narrative image browser for multiplexed tissue images. *J. Open Source Softw.* **5**, 2579 (2020).
9. Genomics, x. Loupe Browser, <https://www.10xgenomics.com/support/software/loupe-browser/latest> (2025).
10. Wang, Y. E., Kutnetsov, L., Partensky, A., Farid, J. & Quackenbush, J. WebMeV: a cloud platform for analyzing and visualizing cancer genomic data. *Cancer Res.* **77**, e11–e14 (2017).
11. Ospina, O. E. et al. spatialGE: quantification and visualization of the tumor microenvironment heterogeneity using spatial transcriptomics. *Bioinformatics* **38**, 2645–2647 (2022).
12. Goldman, M. J. et al. Visualizing and interpreting cancer genomics data via the Xena platform. *Nat. Biotechnol.* **38**, 675–678 (2020).
13. Wang, G. et al. CROST: a comprehensive repository of spatial transcriptomics. *Nucleic Acids Res.* **52**, D882–D890 (2024).
14. Jain, S. et al. Advances and prospects for the Human BioMolecular Atlas Program (HuBMAP). *Nat. Cell Biol.* **25**, 1089–1100 (2023).
15. HistomicsTk, <https://digitalslidearchive.github.io/HistomicsTK/index.html>.
16. Melo Ferreira, R., Freije, B. J. & Eadon, M. T. Deconvolution tactics and normalization in renal spatial transcriptomics. *Front. Physiol.* **12**, 812947 (2022).
17. Börner, K. et al. Tissue registration and exploration user interfaces in support of a human reference atlas. *Commun. Biol.* **5**, 1369 (2022).
18. Börner, K. et al. Anatomical structures, cell types, and biomarkers tables plus 3d reference organs in support of a Human Reference Atlas. *bioRxiv* <https://doi.org/10.1101/2021.05.31.446440> (2021).
19. Bueckle, A. et al. The HRA organ gallery affords immersive superpowers for building and exploring the human reference Atlas with virtual reality. *bioRxiv* <https://doi.org/10.1101/2023.02.13.528002> (2023).
20. Jacobsen, A. et al. 2 10-29 (MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info, 2020).
21. McInnes, L., Healy, J. & Melville, J. UMAP: uniform manifold approximation and projection for dimension reduction. Preprint at <https://arxiv.org/abs/1802.03426> (2018).
22. Tervaert, T. W. C. et al. Pathologic classification of diabetic nephropathy. *J. Am. Soc. Nephrol.* **21**, 556–563 (2010).
23. Fogo, A. B. Glomerular hypertension, abnormal glomerular growth, and progression of renal diseases. *Kidney Int.* **57**, S15–S21 (2000).
24. Ferkowicz, M. J. et al. Molecular signatures of glomerular neovascularization in a patient with diabetic kidney disease. *Clin. J. Am. Soc. Nephrol.* **19**, 266–275 (2024).
25. Kinsey, G. R., Li, L. & Okusa, M. D. Inflammation in acute kidney injury. *Nephron Exp. Nephrol.* **109**, e102–e107 (2008).
26. Zuk, A. & Bonventre, J. V. Acute kidney injury. *Annu. Rev. Med.* **67**, 293–307 (2016).
27. Akchurin, O. M. & Kaskel, F. Update on inflammation in chronic kidney disease. *Blood Purif.* **39**, 84–92 (2015).
28. Mihai, S. et al. Inflammation-related mechanisms in chronic kidney disease prediction, progression, and outcome. *J. Immunol. Res.* **2018** (2018).
29. Eddy, A. A. Progression in chronic kidney disease. *Adv. Chronic Kidney Dis.* **12**, 353–365 (2005).
30. Panizo, S. et al. Fibrosis in chronic kidney disease: pathogenesis and consequences. *Int J. Mol. Sci.* **22**, 408 (2021).
31. Kaplan, C., Pasternack, B., Shah, H. & Gallo, G. Age-related incidence of sclerotic glomeruli in human kidneys. *Am. J. Pathol.* **80**, 227 (1975).
32. Kremers, W. K. et al. Distinguishing age-related from disease-related glomerulosclerosis on kidney biopsy: the Aging Kidney Anatomy study. *Nephrol. Dial. Transpl.* **30**, 2034–2039 (2015).
33. Ke, R. et al. In situ sequencing for RNA analysis in preserved tissue and cells. *Nat. Methods* **10**, 857–860 (2013).
34. Chen, C., Kim, H. J. & Yang, P. Evaluating spatially variable gene detection methods for spatial transcriptomics data. *Genome Biol.* **25**, 18 (2024).
35. Wang, J. et al. Dimension-agnostic and granularity-based spatially variable gene identification using BSP. *Nat. Commun.* **14**, 7367 (2023).
36. Eadon, M. T. et al. Kidney histopathology and prediction of kidney failure: a retrospective cohort study. *Am. J. Kidney Dis.* **76**, 350–360 (2020).
37. Black, S. et al. CODEX multiplexed tissue imaging with DNA-conjugated antibodies. *Nat. Protoc.* **16**, 3802–3835 (2021).
38. Salas, S. M. et al. Optimizing xenium in situ data utility by quality assessment and best practice analysis workflows. *BioRxiv* 2023-2002 (2023).
39. Rosenbloom, A. et al. A complete pipeline for high-plex spatial proteomic profiling and analysis of neural cell phenotypes on the CosMx™ Spatial Molecular Imager and AtoMx™ Spatial Informatics Platform. *Alzheimer's Dement.* **19**, e076665 (2023).
40. Project, H. B. A. <https://portal.hubmapconsortium.org>.
41. Goode, A., Gilbert, B., Harkes, J., Jukic, D. & Satyanarayanan, M. OpenSlide: a vendor-neutral software foundation for digital pathology. *J. Pathol. Inform.* **4**, 27 (2013).
42. Gutman, D. A. et al. The digital slide archive: a software platform for management, integration, and analysis of histology for cancer research. *Cancer Res.* **77**, e75–e78 (2017).
43. Seurat (V 3.2.3), <https://satijalab.org/seurat/>.
44. de Boer, I. H. et al. Rationale and design of the kidney precision medicine project. *Kidney Int.* **99**, 498–510 (2021).
45. Miller, B. F., Huang, F., Atta, L., Sahoo, A. & Fan, J. Reference-free cell type deconvolution of multi-cellular pixel-resolution spatially resolved transcriptomics data. *Nat. Commun.* **13**, 2339 (2022).
46. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e3529 (2021).
47. Lutnick, B. et al. A user-friendly tool for cloud-based whole slide image segmentation with examples from renal histopathology. *Commun. Med.* **2**, 105 (2022).
48. Greenwald, N. F. et al. Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. *Nat. Biotechnol.* **40**, 555–565 (2022).
49. Ruifrok, A. C. & Johnston, D. A. Quantification of histochemical staining by color deconvolution. *Anal. Quant. Cytol. Histol.* **23**, 291–299 (2001).
50. Thevenaz, P., Ruttimann, U. E. & Unser, M. A pyramid approach to subpixel registration based on intensity. *IEEE Trans. Image Process.* **7**, 27–41 (1998).
51. Kitware, I. *large-image* (2024) https://github.com/girder/large_image.
52. Genomics, x. *Xenium Explorer* (2024). <https://www.10xgenomics.com/support/software/xenium-explorer/latest>.
53. Plotly, D. <https://dash.plotly.com>.
54. Border, S. & Sarder, P. FUSION: A web-based application for in-depth exploration of multi-omics data with brightfield histology. *SarderLab/FUSION*. **0** <https://doi.org/10.5281/zenodo.15880605> (2025).
55. Border, S. & Sarder, P. FUSION: A web-based application for in-depth exploration of multi-omics data with brightfield histology. *SarderLab/fusion-tools* **6**, 19 (2025).
56. Biorender. <https://www.biorender.com> (2024).
57. *Kidney Precision Medicine Project*, <https://kpmp.org/>.

Acknowledgements

The authors thank the Kidney Precision Medicine Project for providing some of the kidney disease data. Lillian Atchison and Samantha Hoffman helped with manuscript editing. Some of the icons in the cell diagrams are informed and inspired by KPMP.org, Biorender.com, and the Human BioMolecular Atlas Program (HuBMAP)^{17,56,57}. T.M.E., S.J., M.T.E., and R.M.F. were supported by U54 DK134301. G.P. is supported by HuBMAP U54 HL165443 and LungMAP U01 HL148861. PS's work was funded by NIH funding from NIDDK—R21 DK128668 & R01 DK114485, and from OD OT2 ODO33753. FG's work was supported by U54 AR081775. M.T.E. and J.W. were supported by R01DK138504.

Author contributions

S.P.B. conceived the idea of *FUSION*, wrote the code and implemented it, and wrote the manuscript. NL assisted in data curation and pre-processing pipelines for spatial-omics data. S.K.C.K., A.S.P., D.M., S.M. contributed to system infrastructure and implementation of *FUSION*. YALS and J.R. assisted in usability-focused design of different components in *FUSION* for various user groups. A.Z.R., J.B.H., and L.B. provided insights from the perspective of research and clinical renal pathologists. J.E.T. contributed to the development of data visualization approaches for handcrafted features. Critiqued and edited the manuscript. TME contributed to concept development, data generation and manuscript editing. S.J. conceptualized the idea of linking histomorphology with multi-omics, provided data, contributed to the idea of different features in *FUSION* and edited the manuscript. M.T.E., YH.C., and R.M.F. generated the *Visium* and *Xenium* data, provided biological interpretation, contributed to feature development within *FUSION*, and edited the manuscript. J.W. and M.R. provided analysis of spatially variable genes in *Xenium* data. F.G. and A.K. contributed skin data. G.S.P. and J.P. contributed lung data. J.W.H. and B.W. contributed small intestine data. P.S. conceived the idea of *FUSION*, conducted study design, coordinated with the study team, mentored S.P.B., and wrote and edited the manuscript. The following authors are members of the HuBMAP consortium: S.P.B., M.T.E., N.L., Y.A.L.S., S.K.C.K., G.S.P., F.G., R.M.F., J.W.H., S.M., Y.H.C., A.S.P., B.W., S.J., A.K., J.P., T.M.E., and P.S.

Competing interests

J.E.T. serves in the scientific advisory board of Neurovascular Diagnostics, and the Board of Directors of American Registry of Pathology. P.S. serves on the advisory board of DigPath Inc. *FUSION* software is protected by ©Copyright 2023-25 University of Florida Research

Foundation, Inc. All Rights Reserved. All other authors declare no competing interests: S.P.B., M.T.E., N.L., Y.A.L.S., S.K.C.K., G.S.P., F.G., R.M.F., J.W., J.W.H., S.M., Y.H.C., A.Z.R., D.M., M.R., L.B., J.R., B.W., S.J., A.K., J.P., J.B.H., and T.M.E.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-63050-9>.

Correspondence and requests for materials should be addressed to Tarek M. El-Achkar, Sanjay Jain, Michael T. Eadon or Pinaki Sarder.

Peer review information *Nature Communications* thanks John Imig and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025

¹Department of Medicine – Section of Quantitative Health, University of Florida, Gainesville, FL, USA. ²Department of Medicine, Indiana University School of Medicine, Indianapolis, IN, USA. ³Kitware, Inc, Clifton Park, NY, USA. ⁴Department of Pathology, Division of AI and Computational Pathology, Duke University, Durham, NC, USA. ⁵Department of Medicine, Division of Nephrology, Duke University, Durham, NC, USA. ⁶College of Public Health and Health Professions, University of Florida, Gainesville, FL, USA. ⁷Department of Health Outcomes & Biomedical Informatics, University of Florida, Gainesville, FL, USA. ⁸Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, MD, USA. ⁹Department of Pathology & Anatomical Sciences, Jacobs School of Medicine & Biomedical Sciences, University at Buffalo, State University of New York, Buffalo, NY, USA. ¹⁰Department of Pathology, University of Michigan, Ann Arbor, MI, USA. ¹¹Department of Biomedical Engineering, Duke University, Durham, NC, USA. ¹²Department of Genetics, Stanford School of Medicine, Stanford, CA, USA. ¹³GE HealthCare Technology and Innovation Center, Chicago, IL, USA. ¹⁴Department of Dermatology, University of Pittsburgh, Pittsburgh, PA, USA. ¹⁵Department of Biomedical Engineering and Informatics, Indiana University Indianapolis, Indianapolis, IN, USA. ¹⁶Department of Pediatrics, Division of Neonatology, University of Rochester, Rochester, NY, USA. ¹⁷Indianapolis VA Medical Center, Indianapolis, IN, USA. ¹⁸Department of Medicine, Division of Nephrology, Washington University School of Medicine, St. Louis, MO, USA. ✉ e-mail: telachka@iu.edu; sanjayjain@wustl.edu; meadon@iu.edu; pinaki.sarder@medicine.ufl.edu

HuBMAP consortium

Samuel P. Border¹, Ricardo Melo Ferreira², Nicholas Lucarelli¹, Suhas Katari Chaluva Kumar¹, Anindya S. Paul¹, Yulia A. Levites Strelakova⁶, Ying-Hua Cheng², Sayat Mimar¹, John W. Hickey¹¹, Bei Wei¹², Fiona Ginty¹³, Arivarasan Karunamurthy¹⁴, Gloria S. Pryhuber¹⁶, Jeffrey Purkerson¹⁶, Tarek M. El-Achkar¹⁸, Sanjay Jain¹⁸, Michael T. Eadon^{2,17} & Pinaki Sarder¹