OmicsNavigator: an LLM-driven multi-agent system for autonomous zero-shot biological analysis in spatial omics

Li Yiyao
The University of Hong Kong
liyiyao@connect.hku.hk

Nirvi Vakharia
Purdue University
nvakhari@purdue.edu

Weixin Liang
Stanford University
wxliang@cs.stanford.edu

Aaron T. Mayer
Enable Medicine
aaron@enablemedicine.com

Ruibang Luo
The University of Hong Kong
rbluo@cs.hku.hk

Alexandro E. Trevino
Enable Medicine
alex@enablemedicine.com

Zhenqin Wu
The University of Hong Kong
zqwu@cs.hku.hk

Abstract

Spatial omics provides unprecedented high-resolution insights into molecular tissue compositions but poses significant analytical challenges due to massive data volumes, complex hierarchical spatial structures, and domain-specific interpretive demands. To address these limitations, we introduce OmicsNavigator, an LLM-driven multi-agent system that autonomously distills expert-level biological insights from raw spatial omics data without domain-specific fine-tuning. OmicsNavigator encodes spatial data into concise natural language summaries, enabling zero-shot annotation of structural components, quantitative analysis of pathological relevance, and semantic search of regions of interest using free-form text queries. We evaluated OmicsNavigator on multiple spatial omics studies of kidney cohorts with different phenotypes and biomarker panels, where OmicsNavigator achieved outstanding performances in structural annotations, pathology assessments, and semantic search across studies. OmicsNavigator offers a scalable, interpretable, and modality-agnostic solution for spatial omics analysis.

1. Introduction

Spatial omics, including spatial transcriptomics and proteomics, is reshaping biological research by enabling high-resolution spatially-resolved mapping of the molecular composition of entire tissues at previously unattainable scales[12, 37]. Platforms such as 10x Visium[25], MERFISH[4], Xenium[19], CODEX[5], and MIBI-TOF[6]

now routinely generate terabytes of data from sequencing or imaging experiments with multiplexed gene or protein expression measurements, thereby providing unprecedented opportunities for in situ investigations of developmental processes, immune responses, and disease progression [10, 40].

Yet the very richness that makes spatial omics attractive also presents substantial analytical challenges: (i) the sheer data volume renders manual inspection infeasible [33]; (ii) the high dimensionality of spatial omics data, encompassing numerous protein biomarker or gene transcripts, lacks an intuitive representation [34]; (iii) structures that are inherently hierarchical, ranging from subcellular protein assemblies and multicellular structures to centimeter-scale anatomical regions, require different interpretation and analysis tools that are customized for the spatial scale and resolution [35]; and (iv) translating statistical patterns into mechanistic insight still relies heavily on domain knowledge from human experts [28].

Concurrently, large language models (LLMs) have redefined the state-of-the-art in natural language understanding and generation across scientific disciplines [14]. Several variants focusing on biomedical knowledge have demonstrated expert-level performance on question-answering and evidence retrieval [18, 22–24]. Multimodal vision–language models integrating radiology or pathology images with free-text reasoning have shown capabilities in automatic report generation, captioning, and interactive querying [1, 9, 17, 32, 39]. In the realm of single-cell and spatial omics, foundation models directly trained on massive amounts of experimental measurements with self-supervised learning strategies [21, 29, 38] have demon-
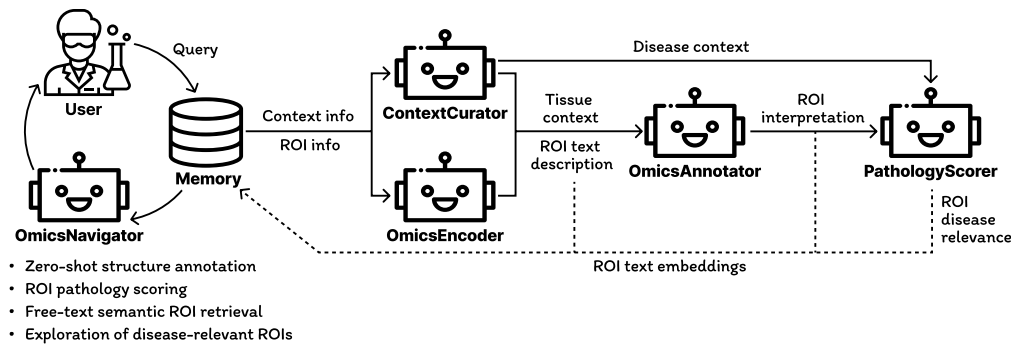
Figure 1. Overview and capabilities of OmicsNavigator.

strated improvements across diverse downstream tasks, through these methods use customized architectures and require downstream fine-tuning, which limits their accessibility and generalizability. In parallel, initial applications of LLM-based tools in single-cell and spatial omics have demonstrated strong performances in tasks such as cell-type annotation [8, 13, 29], agentic multi-step analysis pipelines [30], and zero-shot reasoning for single-cell readouts [20]. These successes hint at a broader opportunity: if spatial omics could be rendered in a language-native form, one might harness the emergent reasoning capabilities of foundation models—without any domain-specific fine-tuning—to accelerate data exploration and biological discovery.

However, this opportunity remains largely untapped. Most current LLM-based bioinformatics tools employ LLMs to support routine tasks—such as literature summarization, metadata normalization, or basic question answering—yet lack the capacity to directly work with raw omics data in its specific biological or pathological contexts. This limitation stems in part from the complex and diverse data formats spatial omics measurements take. Raw sequencing and imaging outputs, as well as dense numerical arrays of cellular tables after preprocessing, are inherently incompatible with the text-only language models without tailed modality-specific embedding tools. Existing attempts reduce the complexity of omics data by representing it as sequences of gene symbols to enable LLM querying [13] or task-specific fine-tuning [20]. Yet, these abstractions have not yet been able to capture the spatial topology and biological contexts essential for robust inference.

To bridge this gap, we present OmicsNavigator, an LLM-driven multi-agent system that autonomously ingests spatial omics data and distills expert-level biological insights without requiring domain-specific fine-tuning (**Figure 1**). OmicsNavigator consists of multiple interacting agents that share information through a common memory module. These individual agents carry out tasks including curation of study-specific contexts, transforming spatial omics data

to free-form text, biological reasoning for annotations and pathological relevance analysis. We demonstrate the capabilities of OmicsNavigator on a range of tasks including zero-shot annotations, quantitative analysis of disease relevance, semantic search, and data-driven discovery of spatial signatures.

Our key contributions are as follows:
- We introduce the first general-purpose omics-to-text encoding mechanism that maps high-dimensional spatial omics data into representations compatible with the reasoning capabilities of LLMs;
- We develop a LLM-driven multi-agent system that produces qualitative and quantitative insights from spatial omics inputs in a zero-shot manner;
- We propose a language-native retrieval framework that accurately identifies canonical structures and disease-relevant hot spots from multiple large-scale datasets, reaching outstanding retrieval accuracies, and offering a lightweight, interpretable and study-agnostic assistant for spatial biology.

2. Methods

2.1. OmicsNavigator

OmicsNavigator is an LLM-driven multi-agent system designed to perform biological reasoning on arbitrary regions of interest (ROIs) from spatial omics studies. It constructs background knowledge and ROI-specific information as natural language representations, followed by reasoning over and interpreting the biological contents and significance of ROIs. An overview of the internal workflow on an example ROI sampled from a kidney region is shown in **Figure 2**.

2.1.1. Curation of study-specific contexts

While LLMs have demonstrated remarkable success in addressing general biomedical questions[23, 24], our experiments show that they often struggle with domain-specific, research-level questions in the absence of adequate back-
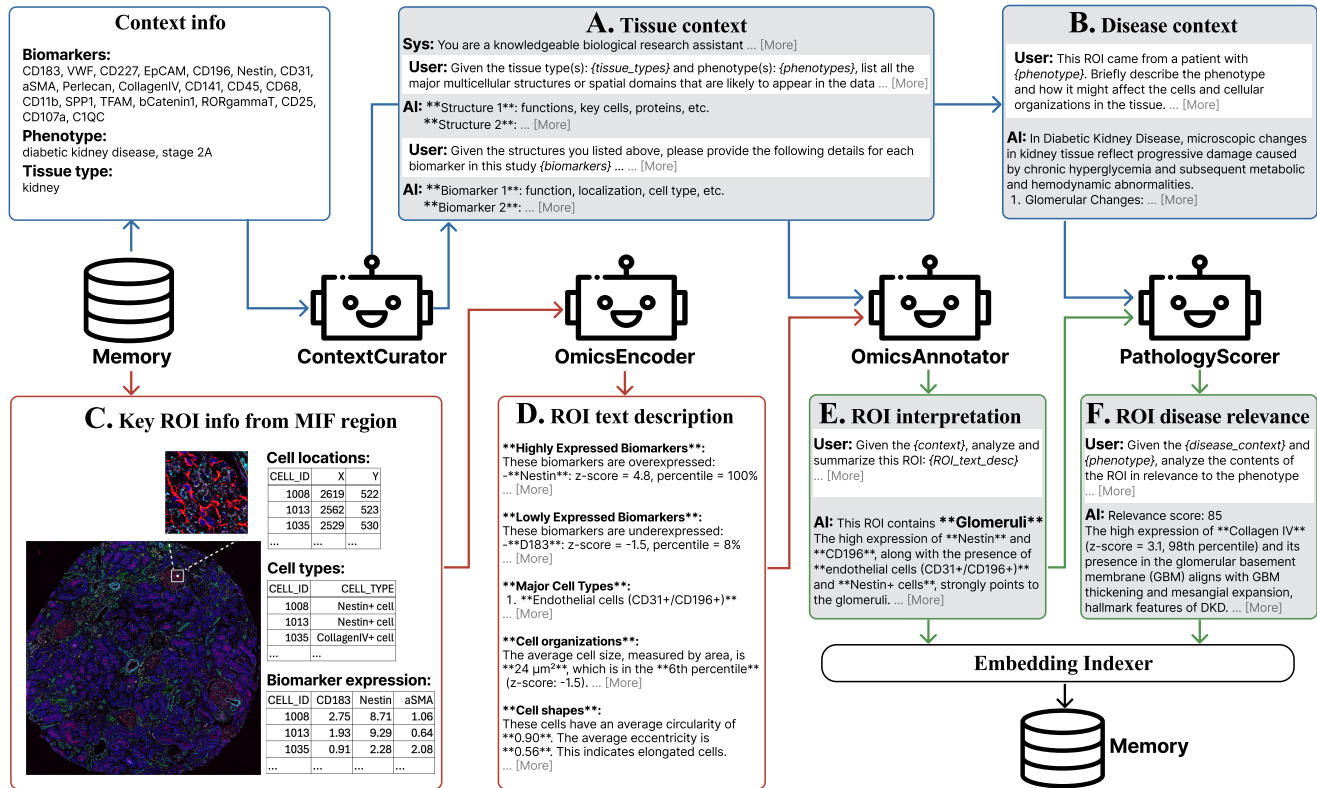
Figure 2. Internal workflow of OmicsNavigator: For a tissue region profiled with spatial omics experiments, the ContextCurator first curates biological contexts regarding the tissue (**A**) and disease (**B**) according to metadata including tissue type, phenotype, biomarker panel, etc. For an arbitrary ROI in the region (**C**), its information—cell locations, types, morphology, and biomarker expression, among others—is extracted and assembled into a natural language description (**D**) by OmicsEncoder. Guided by the contexts, the OmicsAnnotator and PathologyScorer annotate the ROI (**E**) and interpret its disease relevance (**F**). All contexts, ROI descriptions and interpretations are indexed and stored in the memory module for downstream retrieval and analysis tasks.

ground information. To address this limitation and ensure applicability across domains, OmicsNavigator employs an LLM-powered agent—referred to as "ContextCurator"—to curate domain-specific contexts of the specific tissue and disease using metadata of the patient and the tissue sample (**Figure 2A-B**).

This process involves interactive conversation with the agent, leveraging metadata including study background, tissue type, disease phenotype of the patient, and all biomarkers measured in the omics experiment. These inputs are used to generate tissue- and disease-specific contexts that guide downstream reasoning. In addition, information regarding the biomarkers can be optionally enhanced by referencing external knowledge bases, such as Human Protein Atlas[26]), which has proven highly effective in our experiments. Importantly, these contexts are generated once, stored in the memory module of OmicsNavigator, and reused across all ROIs from the same omics experiment.

2.1.2. Omics-to-text encoding

To address the modality gap between spatial omics data and the textual input required by LLMs, we introduce a transformation module called "OmicsEncoder", which converts an arbitrary ROI into a natural language representation. This module is inspired by the analytical workflows of human experts, wherein key information of the ROI, such as cellular composition and differentially-expressed biomarkers, are extracted and assembled.

This process begins by segmenting and characterizing all cells within the tissue sample. For cells within the queried ROI, features such as biomarker expression, morphology, and/or cell type are extracted and aggregated to define the ROI-specific features. These features are then compared against cellular features randomly sampled from the full region for additional normalization. Key distinguishing characteristics—including top differentially expressed biomarkers, distinctive morphological patterns, and predominant cell types—are extracted and summarized in natural language (**Figure 2D**). The resulting textual description

includes sections on highly expressed biomarkers, major cell types, spatial localization patterns and other informative cues that closely resemble the elements typically used by human experts in biological reasoning.

2.1.3. LLM-aided ROI annotation

OmicsNavigator then employs an LLM-powered agent, "OmicsAnnotator", to interpret the contents of the queried ROI based on its natural language description and the associated contexts (**Figure 2**E). In our experiments, the agent was explicitly prompted to generate annotations for the major multicellular structures or spatial domains in the region—a task extensively studied in the spatial omics analysis. This annotation task can be viewed as a specialized form of visual question answering (VQA), where the "visual" modality is replaced by textual encodings of spatial omics data. Extending this framework to support more diverse forms of content-based interpretation will be explored in future work.

2.1.4. Inference of disease relevance

A natural extension of the LLM-aided ROI interpretation is to infer its relevance to patient-specific clinical characteristics, such as disease state. To this end, we introduce another LLM-powered agent, "PathologyScorer", which infers disease relevance using the queried ROI's textual description, structural interpretation generated by "OmicsAnnotator", and the corresponding tissue- and disease-specific contexts (**Figure 2F**). Notably, in addition to generating a natural language explanation, the agent is explicitly instructed to output a quantitative relevance score, enabling cross-ROI comparisons of disease relevance.

2.2. Coupling existing unsupervised annotation tools with OmicsNavigator

In addition to annotating arbitrarily selected ROIs, OmicsNavigator seamlessly integrate with existing computational tools[15, 27, 36] that perform unsupervised annotations on spatial omics data (**Figure 5**). Using cell-level unsupervised annotations produced by an arbitrary external tool, we can identify and extract ROIs enriched with cells from specific clusters. OmicsNavigator can then be applied to interpret these cluster-specific ROIs. By further enumerating and aggregating individual interpretations of cluster-specific ROIs, a consensus annotation can be derived for each cluster. This approach ensures comprehensive coverage of all major tissue components discovered by the external tool, while also enabling cross-region integration or comparison of clusters.

2.3. Semantic search

Beyond annotation, OmicsNavigator also opens the path to semantic search by linking spatial omics with natural language descriptions and interpretations (**Figure 6A**). To con-

struct the searchable text corpus, we enumerate all ROIs within the region, with each ROI encoded based on its feature. Subsequently, these ROIs are grouped using a multistage clustering pipeline, with in total 40 to 50 representatives selected from all clusters, termed "key ROIs". This approach is inspired by prior work on content-based image retrieval techniques[3, 11], where clustering and pruning techniques are used to reduce the search space of candidate patches. Each key ROI is analyzed by OmicsNavigator, and the resulting textual interpretations are encoded and stored in the memory module to support subsequent semantic retrieval.

During a semantic search, an input query text is encoded and compared against entries in the memory. Pairwise distances and exact text matches are computed to identify the top hits among the key ROIs. These hits serve as anchors to compute the median distances between all ROI features and the anchor features. Finally, the most relevant ROIs according to median distance to anchors are retrieved and visualized for further analysis.

3. Dataset

3.1. Diabetic kidney disease (DKD) dataset

We used a previously published CODEX[5] dataset of kidney samples from diabetic patients [16], hereinafter referred to as the DKD dataset. A total of 17 regions comprising 137,000 cells were curated for OmicsNavigator experiments. These regions span multiple stages of diabetic kidney disease (early to stage 3) and exhibit a broad spectrum of disease-associated histological changes.

Each region in the DKD dataset was preprocessed using a standard pipeline involving cell segmentation, biomarker integration, and unsupervised cell typing via Leiden clustering, resulting in 11 identified cell types, including tubular, endothelial, immune, and other cells. Pathologist annotations of five major anatomical structures—proximal tubules, distal tubules, glomeruli, blood vessels, and interstitium—were used to evaluate OmicsNavigator's zero-shot annotation accuracy in subsequent experiments.

3.2. Transplant rejection (TR) dataset

We adapted a different kidney dataset [2] comprising samples profiled using CODEX from 5 normal and 7 transplant rejection kidney regions, hereinafter denoted as the TR dataset. In total around 350,000 cells were identified and classified after the same preprocessing pipeline, yielding 20 different cell types including immune, tubular, fibroblast, and smooth muscle cells. Importantly, the TR dataset was profiled using a different biomarker panel, with only 7 of the 52 markers (13%) overlapping with the DKD dataset. As demonstrated in the results below, OmicsNavigator is agnostic to biomarker panels and can perform joint
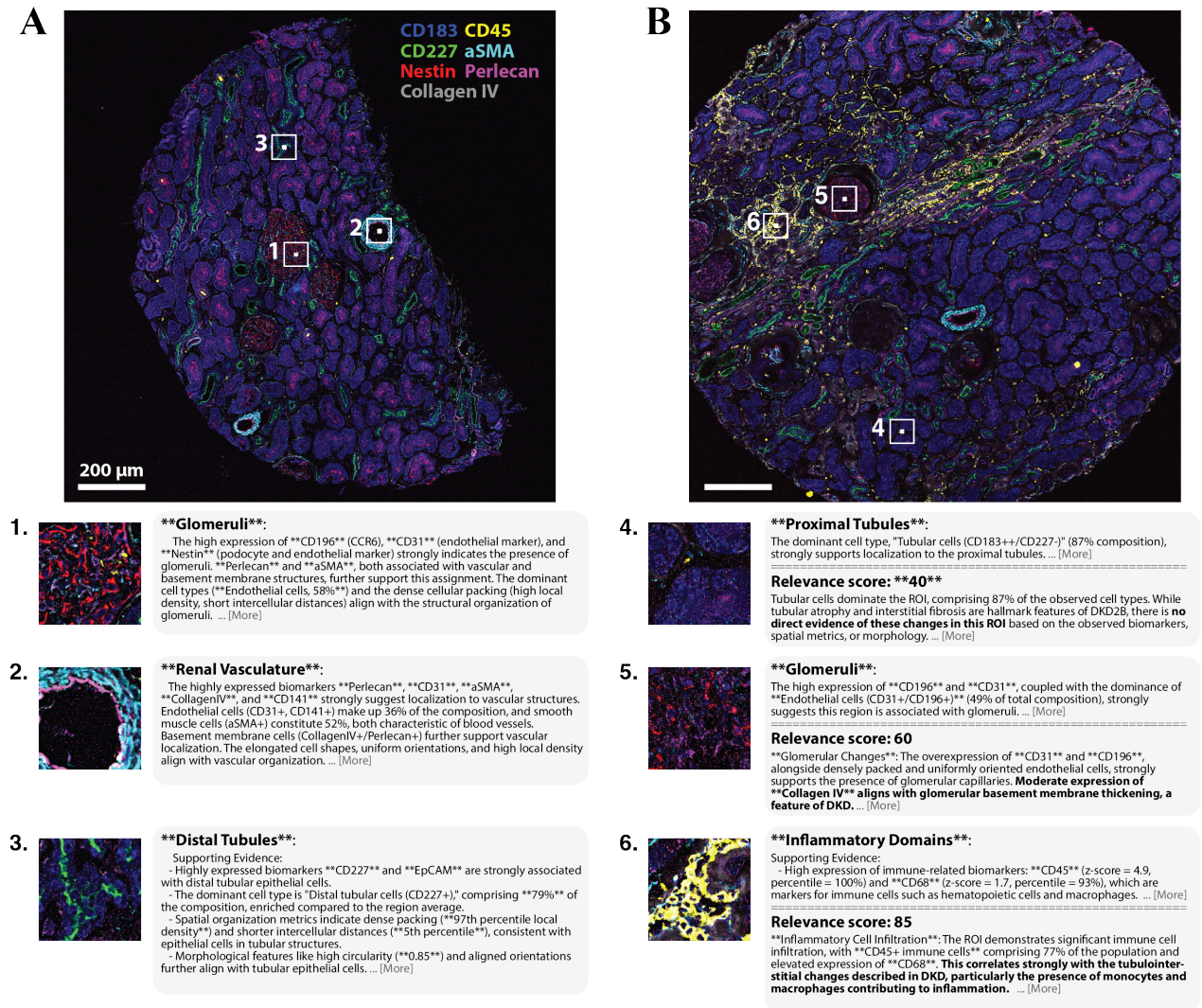
Figure 3. Two kidney regions from the DKD dataset with ROIs annotated by OmicsNavigator. (**A**) Tissue section from a patient with early-stage DKD. Three ROIs (1–3), marked by white boxes, are annotated for structural components, with key sections in the textual interpretations from OmicsNavigator shown. All annotations aligned with pathologist evaluations. (**B**) Tissue section from a patient with late-stage (IIB) DKD. Three ROIs (4–6) are annotated for structural components and disease relevance by OmicsNavigator. Varying levels of disease-associated changes were observed, accurately captured by OmicsNavigator and quantified using relevance scores.

analysis of both studies without any adjustments.

3.3. Kidney allograft rejection (KAR) dataset

Another dataset consists of 8 T cell-mediated and 8 antibody-mediated rejection kidney cases, referred to as the KAR dataset[7], was used in parallel to test the generalizability of OmicsNavigator. This dataset, similarly, used a different biomarker panel and contained around 558,000 cells of 13 distinct cell types, with a particular focus on immune populations within kidney structures.

All three datasets were collected from different institutions and used customized biomarker panels. Cell segmentation and annotations were performed independently, therefore yielding non-overlapping sets of cell types. OmicsNavigator operates on regions from each study in an independent manner, thereby minimizing the influence of potential batch effects and panel discrepancy. The interpretations and semantic search results are directly comparable across studies, thanks to their language-native format.
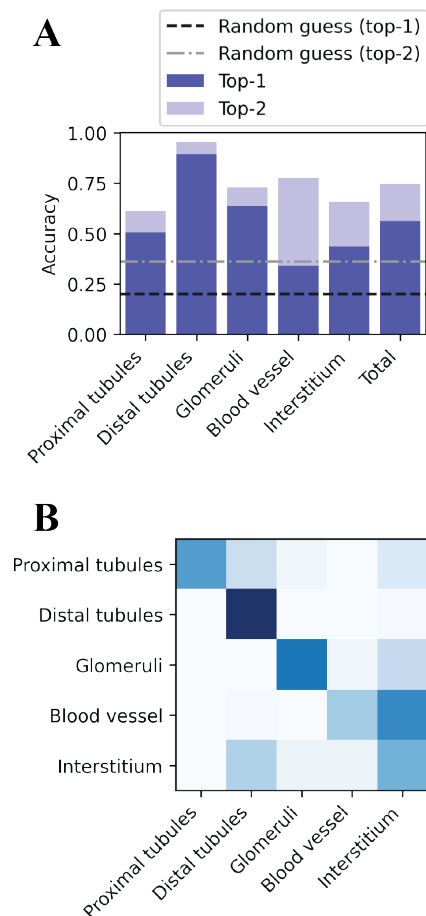
A

B

Figure 4. Accuracy of annotations from OmicsNavigator: We randomly sampled 500 ROIs around pathologist-annotated cells and prompted OmicsNavigator to generate up to two guesses for annotating components. OmicsNavigator does not have access to the ground truth classes. (**A**) Per-class and average accuracy for top-1 and top-2 guesses, with baseline accuracy of randomly guessing ground truth classes shown as dashed lines. (**B**) Confusion matrix comparing ground truth classes to top-1 guesses from OmicsNavigator. Notable confusion between the blood vessel and interstitium classes arises due to overlapping signature biomarkers and ambiguities in class definitions.

4. Results

4.1. ROI interpretation with OmicsNavigator

To illustrate how OmicsNavigator performs ROI-level interpretation, we present two representative examples of different disease stages from the DKD dataset. Region shown in **Figure 3A** is sampled from a patient with early-stage DKD, while Region in **Figure 3B** is sampled from a late-stage (IIB) disease patient. For each region, several ROIs were selected and annotated using OmicsNavigator in a zero-shot manner, which generates textual interpretations on the ROIs' structural components and assesses their disease rel-

evance. The outputs closely match expert pathologist evaluations, correctly identifying structures such as glomeruli and distal convoluted tubules.

Furthermore, OmicsNavigator captures disease-associated alterations in ROIs. Analysis on the three ROIs sampled from the late-stage disease region show varying levels of disease relevance. OmicsNavigator highlights an ROI heavily infiltrated by immune cells (No. 6 in **Figure 3B**), evidenced by the strong CD45 expression, as well as the structural annotation of "inflammatory domains". In addition, OmicsNavigator also identifies a glomerulus-associated ROI in the region (No. 5 in **Figure 3B**) and suggests its disease relevance indicated by glomerular basement membrane. These observations demonstrate the outstanding capabilities of OmicsNavigator in identifying and highlighting disease-associated signals in an interpretable and context-aware fashion.

We further quantitatively evaluated OmicsNavigator's annotation accuracy by randomly sampling 500 ROIs surrounding pathologist-annotated cells. For each ROI, OmicsNavigator was prompted to generate up to two candidate annotations without access to ground truth labels. The results show that OmicsNavigator substantially outperforms random baselines, achieving outstanding top-1 and top-2 accuracies of 56% and 75%, respectively, across various tissue structures (**Figure 4**A). Analysis of the confusion matrix between ground truth classes and the top-1 guess by OmicsNavigator (**Figure 4**B) suggests ambiguity between the class of blood vessel and interstitium, which largely stems from overlapping signature biomarkers (i.e., aSMA, CollagenIV). In such cases, allowing multiple guesses with justifications—a unique advantage offered by LLM-based approach—provides substantial help.

Overall, OmicsNavigator demonstrates exceptional capability in ROI-level interpretation by accurately identifying structural components and assessing disease relevance.

4.2. Unsupervised cluster annotation with OmicsNavigator

OmicsNavigator can be seamlessly integrated with existing computational biology tools that perform unsupervised annotations of spatial omics data [15, 27, 36]. As illustrated in **Figure 5A-B**, we analyzed two representative tissue regions independently using an existing unsupervised annotation tools developed for spatial omics (i.e., SCGP[36]). Such class of tools leverage a variety of machine learning and deep learning approaches to identify multicellular spatial domains based on cellular-level profiles. However, most of them produce unsupervised outputs that require manual inspection and annotation for further interpretation. In studies involving multiple regions, clusters identified across runs (e.g., on different regions, batches, or case/control groups) will need to be matched and integrated for larger-scale anal-
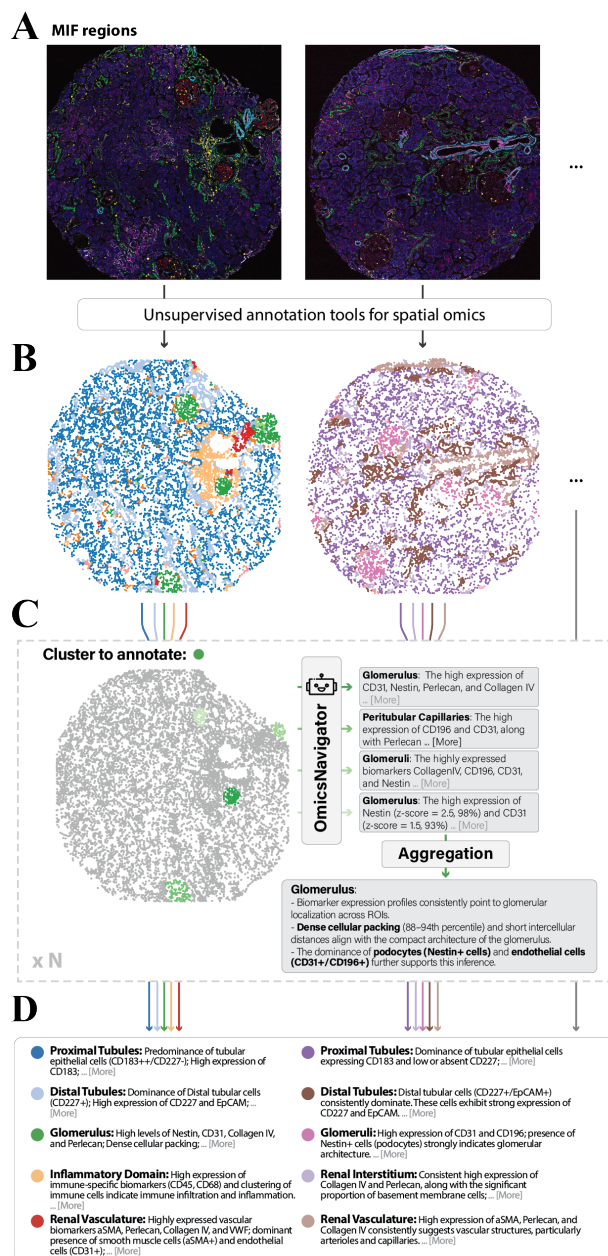
Figure 5. Coupling unsupervised annotation tools with OmicsNavigator. Two randomly selected regions from the DKD dataset (**A**) were analyzed using an existing unsupervised annotation tool to identify tissue structures / spatial domains, marked by different colors (**B**). OmicsNavigator can be used to further annotate these unsupervised clusters (**C**) by sampling subsets of cells and aggregating their interpretations. Clusters across different regions were annotated and matched (**D**), revealing common structures (e.g., glomeruli in green and pink) and region-specific domain (e.g., inflammatory domain in orange).

yses. OmicsNavigator can streamline these post hoc annotation steps through detailed biological reasoning and inter-

pretations, which typically rely on domain expertise from human experts.

In practice, we employed a technique similar to self-consistency decoding [31]. For each unsupervised cluster, representative subsets of cells were selected, and ROIs enriched with these cells—referred to as cluster-specific ROIs—were defined. OmicsNavigator was used to interpret and annotate multiple cluster-specific ROIs from the same cluster, and the resulting interpretations were further aggregated to form a consensus annotation. An example is shown in **Figure 5**C, where four subsets of cells are annotated independently and aggregated into a consensus annotation. Notably, the consensus annotation remained accurate even though the four cluster-specific ROI annotations were inconsistent. This procedure improves the overall robustness and accuracy of OmicsNavigator by leveraging diverse samples of cluster-specific ROIs, mitigating the impact of noise and errors in individual annotations. In our experiments, OmicsNavigator produced only two mistakes in its consensus annotations across 92 clusters identified in the DKD dataset.

Furthermore, the language-native outputs of OmicsNavigator ensure that post hoc annotations are naturally consistent across regions, allowing clusters from different regions to be matched. As shown in **Figure 5**D, OmicsNavigator identifies both common tissue structures (e.g., glomeruli in green and pink) and region-specific domain (e.g., inflammatory domain in orange) across the two regions. In practice, we were able to identify common clusters across all 17 regions of the DKD dataset, enabling comparative analyses on the same structure to reveal disease-associated alterations.

In summary, OmicsNavigator complements existing unsupervised annotation tools by providing interpretable and consistent post hoc annotations of clusters, facilitating downstream integration and cross-region analyses.

4.3. Semantic search with OmicsNavigator

OmicsNavigator enables text-based semantic search over spatial omics regions by linking ROIs to natural language descriptions. As illustrated in **Figure 6**A, the process begins by identifying representative ROIs of the subject region through a multi-stage unsupervised clustering pipeline that leverages ROI features and coordinates. These representatives, termed key ROIs, are then analyzed for their biological contents and disease relevance using OmicsNavigator. Textual results of the interpretations are converted and encoded into searchable entries. Once entries from all key ROIs are indexed and stored, OmicsNavigator can retrieve semantically similar key ROIs, or hits, from the memory When provided with a free-form text query (e.g., "Glomeruli", "Immune-infiltrated interstitium"). These hits serve as anchors to identify relevant ROIs across the entire tissue region, based on feature distances.
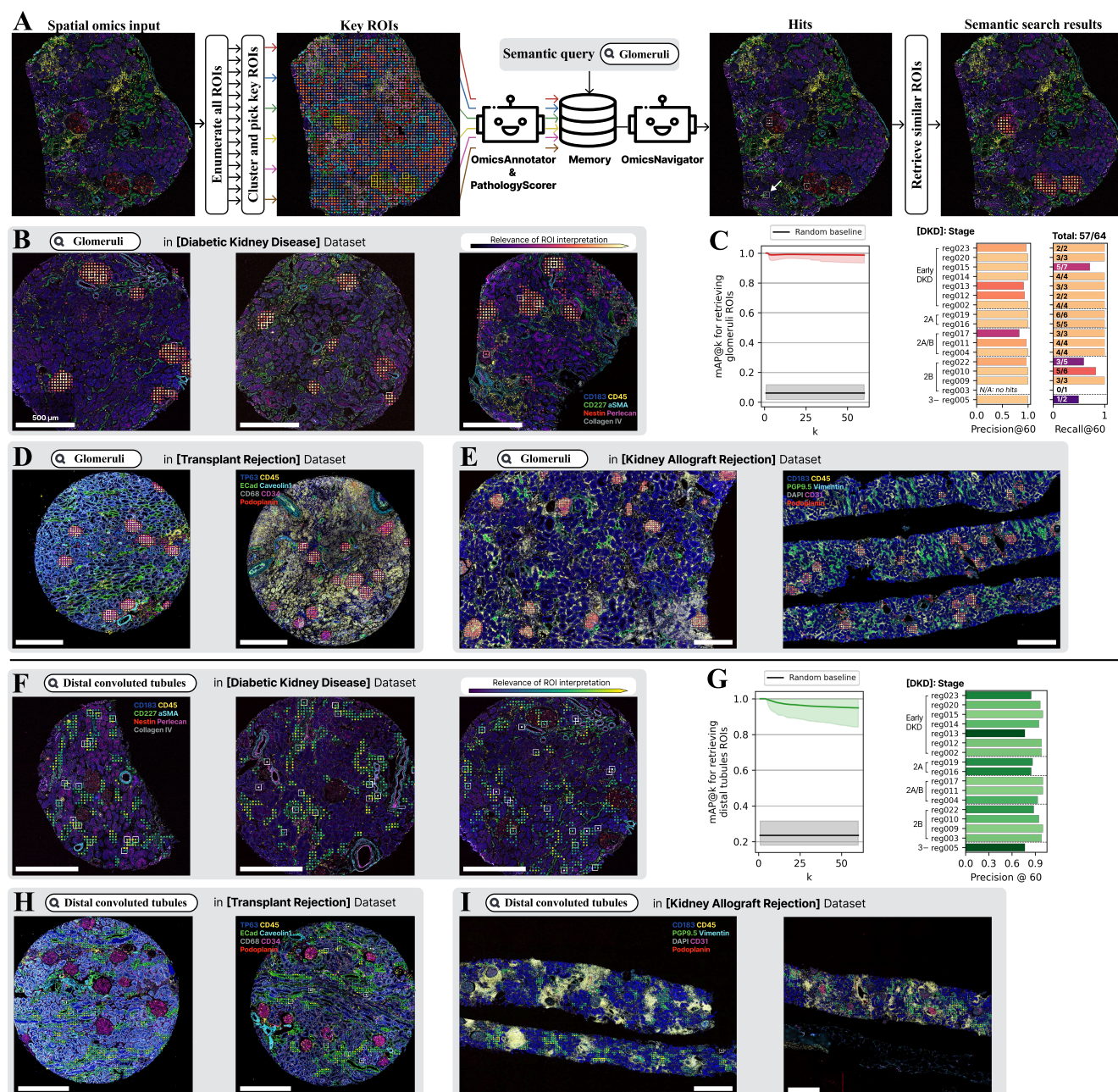
Figure 6. OmicsNavigator enables semantic search using natural language queries. (**A**) In a semantic search pipeline, key ROIs are defined through multi-stage unsupervised clustering and are subsequently interpreted using OmicsAnnotator/PathologyScorer. For a given text query, top hits in key ROIs are identified based on semantic similarities and subsequently used as anchors to retrieve all relevant ROIs. (**B**) Results of searching for the kidney structure "Glomeruli" in example regions from the DKD dataset. (**C**) Quantitative metrics of retrieval performance across the whole DKD dataset, achieving an overall precision of 0.97 and a recall of 0.89 when counting top 60 ROIs per region. Shaded areas in the mAP@k plot represent the 90% confidence interval. (**D-E**) Results of searching for the same keyword "Glomeruli" in the TR and KAR datasets, both achieving outstanding visual alignment. Note that colors represent different biomarkers in these three datasets. (**F**) Results of searching for the kidney structure "Distal convoluted tubules" in the DKD dataset. (**G**) Quantitative metrics of retrieval performance, with an overall precision of 0.93 when counting top 60 ROIs. Shaded areas in the mAP@k plot represent the 90% confidence interval. (**H-I**) Results of searching for "Distal convoluted tubules" in the TR and KAR datasets, both achieving outstanding visual alignment.

The design of key ROIs as an intermediate step offers two major advantages. First, by focusing on representative ROIs, this strategy drastically reduces the cost and time required to run OmicsNavigator while maintaining comprehensive search coverage. Second, this approach is error-tolerant: final search results are calculated based on median feature distances to anchors (i.e., hits). Minor mistakes from OmicsNavigator when analyzing key ROIs (e.g., the white box marked by arrow in **Figure 6A**) have minimal impact on the search results.

We assessed our approach across all three studies comprising samples of different disease phenotypes, tissue sources, batches, biomarker panels, and cell type compositions, where OmicsNavigator achieved robust performance in all scenarios. For instance, glomerular structures were consistently retrieved with high precision and recall across the DKD, TR, and KAR datasets (**Figure 6B,D,E**), despite considerable inter-dataset heterogeneity. Quantitatively, OmicsNavigator achieved strong retrieval metrics (**Figure 6C**): when counting top 60 ROIs, it attained a precision of 0.97 and a recall of 0.89, calculated as the amount of individual glomerulus identified in the search results. Similarly, OmicsNavigator demonstrated outstanding performances in retrieving the kidney structure of distal convoluted tubules (**Figure 6F,H,I**), with high quantitative metrics (**Figure 6G**).

These results highlight the potential of OmicsNavigator's language-native format to bridge the gap between spatial omics data and intuitive, text-based exploration, enabling precise and efficient retrieval of arbitrary regions.

4.4. Exploration of disease-relevant spatial patterns with OmicsNavigator

In addition to the identification and annotation of structures, investigation of disease-associated spatial patterns is another pivotal component of spatial omics data analysis, which could provide critical insights into the mechanisms underlying disease onset, progression, and prognosis. While substantial amount of biological knowledge has been accumulated over the years, computational analysis of spatial omics data that autonomously makes use of this wealth of knowledge remains an under-explored direction. By leveraging LLM-driven agents equipped with disease-specific contexts, OmicsNavigator bridges this gap, enabling disease-aware exploration of spatial patterns with augmentation from existing biological knowledge.

In a targeted setting, semantic search can be used to retrieve ROIs of known pathological significance. For instance, using the keyword "Immune aggregates", OmicsNavigator successfully retrieved ROIs enriched for immune cells and immune-related biomarkers including CD45, CD68, and CD3e **Figure 7A**. The distributions of these ROIs closely reflected region-level phenotypes asso-

ciated with inflammation levels and disease stages. Visualizations of contrasting regions from each study further demonstrate OmicsNavigator's ability to capture meaningful inter-sample differences. Very importantly, OmicsNavigator does not show signs of hallucinating findings in early-stage or mild disease cases, as indicated by the low number of hits and retrieved ROIs in corresponding region groups across all studies. This ensures that the outputs of OmicsNavigator remain faithful to the underlying data and prevent over-interpretation. Broadly, this approach holds strong potential in scenarios where analysts could identify structures of known pathological significance, such as localizing tertiary lymphoid structures and necrotic cores in tumor microenvironments.

Beyond predefined queries, OmicsNavigator can also be used to explore disease-relevant spatial patterns in a data-driven manner **Figure 7B**. By ranking key ROIs based on their disease relevance scores—automatically assigned by PathologyScorer, an LLM agent equipped with disease-specific contexts—high-scoring regions can be selected and analyzed, enabling unbiased discovery of pathologically-relevant spatial signatures. Quantitatively, we observed similar trends, where OmicsNavigator identifies most highly disease-relevant ROIs in severe disease cases, avoiding over-interpretation in mild cases. Through further clustering and summarizing these highly disease-relevant ROIs, consensus patterns can be identified, each representing a distinct mode of tissue alteration. In the examples shown in **Figure 7B**, five major disease-associated alterations are identified and localized in the regions without any human guidance.

Collectively, OmicsNavigator opens new avenues for contextualizing and analyzing spatial omics data by integrating existing biological knowledge. It enables both hypothesis-driven and unbiased discovery approaches to study spatial omics data.

5. Discussion

We present a novel LLM-driven multi-agent system that autonomously ingests spatial omics data and derives expert-level biological insights without requiring domain-specific fine-tuning. OmicsNavigator effectively encodes arbitrary ROIs from spatial omics studies into text representations, enabling LLM-driven agents to reason over high-dimensional omics measurements. Evaluated across three spatial omics datasets of different kidney diseases, OmicsNavigator excels in tasks such as structural annotations and pathology assessments.

Beyond annotations, the language-native format of OmicsNavigator's outputs greatly expands its applicability in downstream tasks. We demonstrated applications including semantic search, identification of disease-relevant hotspots, and summarization of spatial patterns. More ver-
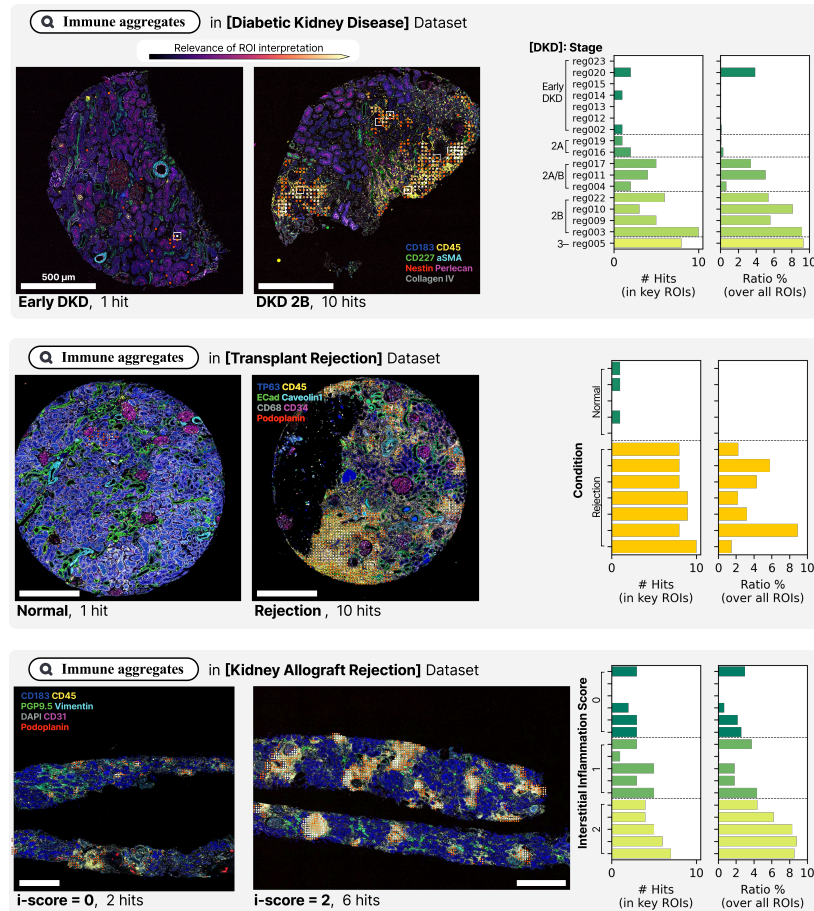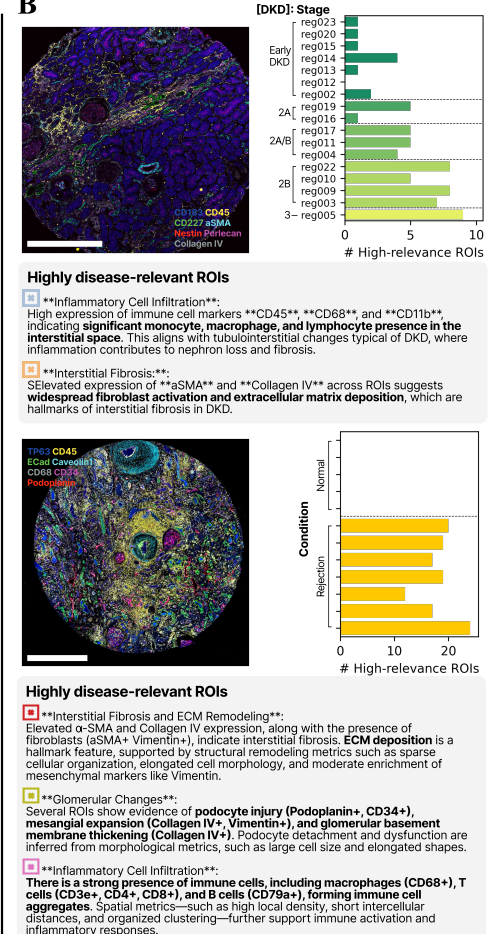
Figure 7. Exploration of disease-relevant spatial patterns with OmicsNavigator. **(A)** Targeted retrieval using semantic search. ROIs matching the keyword "Immune aggregates" were retrieved across three datasets. Retrieved ROIs exhibited high expression of immune-related biomarkers (e.g., CD45, CD68, CD3e), and the abundance of retrieved ROIs closely reflected region-level inflammation levels and disease stages (right panels). For each dataset, two contrasting regions are shown (left and middle panels). **(B)** Untargeted, data-driven exploration with OmicsNavigator. ROIs with high disease relevance scores—automatically inferred by OmicsNavigator—were retrieved and summarized to curate disease-associated patterns. Visualizations from two late-stage disease samples illustrate how retrieved ROIs capture diverse modes of disease-associated tissue alteration, with the number of high-relevance ROIs reflecting disease severity across datasets.

satile usages of OmicsNavigator's textual interpretations will be explored in subsequent works. Notably, OmicsNavigator's interpretability and modularity make it especially suitable for human-in-the-loop analysis, where OmicsNavigator, applied in conjunction with other LLM-driven tools, can serve as co-pilot for tasks such as guided tissue annotation, cross-condition comparative analysis, and hypothesis generation/refinement for disease-relevant studies. Additionally, OmicsNavigator could be deployed as a first-pass analysis tool for pathological assessment in high-throughput settings. It can autonomously identify and localize major pathological signatures to prioritize findings for further human evaluations.

Despite these advances and promises, several limitations remain. First, OmicsNavigator relies heavily on tissue-specific and disease-specific contexts, which are curated from the prior knowledge of LLMs and/or external reference databases. In highly specialized or emerging domains, the scarcity of existing knowledge may constrain its performance. Second, OmicsNavigator's outputs may exhibit uncertainty or mistakes. While error-tolerant mechanisms have been introduced in some downstream tasks in this work, the supervision of potentially erroneous outputs—especially to mitigate risks in clinical or research settings—remains a critical focus for future work. Last but not the least, the omics-to-text transformation in OmicsNaviga-

tor may lead to compression-related information loss, potentially omitting subtle but biologically important details. Future work could explore hybrid architectures that combine text-based reasoning with spatially-aware modeling or vision-language frameworks to better capture spatial features while maintaining interpretability.

Taken together, OmicsNavigator is an autonomous, scalable, interpretable, and versatile approach that derives biological insights from spatial omics data, with strong potential to advance both research and clinical applications of spatial omics studies.

Data and code availability

All codes and data used in this study are publicly available at https://github.com/yyli-leo/OmicsAnnotator.

References

[1] Benedikt Boecking, Naoto Usuyama, Shruthi Bannur, Daniel C Castro, Anton Schwaighofer, Stephanie Hyland, Maria Wetscherek, Tristan Naumann, Aditya Nori, Javier Alvarez-Valle, et al. Making the most of text semantics to improve biomedical vision–language processing. In *European conference on computer vision*, pages 1–21. Springer, 2022. 1

[2] Nathan A Bracey, Jonathan Maltzman, Adrienne Long, Renu Dhanasekeran, Vishnu Shankar, Azam Mohsin, Neeraja Kambham, Gerlinde Wernig, Andrew Gentles, Mark M Davis, et al. The immune microenvironment of transplant glomerulitis. *Kidney International Reports*, 2025. 4

[3] Chengkuan Chen, Ming Y Lu, Drew FK Williamson, Tiffany Y Chen, Andrew J Schaumberg, and Faisal Mahmood. Fast and scalable search of whole-slide images via self-supervised deep learning. *Nature Biomedical Engineering*, 6(12):1420–1434, 2022. 4

[4] Kok Hao Chen, Alistair N Boettiger, Jeffrey R Moffitt, Siyuan Wang, and Xiaowei Zhuang. Spatially resolved, highly multiplexed rna profiling in single cells. *Science*, 348 (6233):aaa6090, 2015. 1

[5] Yury Goltsev, Nikolay Samusik, Julia Kennedy-Darling, Salil Bhate, Matthew Hale, Gustavo Vazquez, Sarah Black, and Garry P Nolan. Deep profiling of mouse splenic architecture with codex multiplexed imaging. *Cell*, 174(4):968–981, 2018. 1, 4

[6] Felix J Hartmann, Dunja Mrdjen, Erin McCaffrey, David R Glass, Noah F Greenwald, Anusha Bharadwaj, Zumana Khair, Sanne GS Verberk, Alex Baranski, Reema Baskar, et al. Single-cell metabolic profiling of human cytotoxic t cells. *Nature biotechnology*, 39(2):186–197, 2021. 1

[7] Toshihito Hirai, Ayano Kondo, Tomokazu Shimizu, Hironori Fukuda, Daisuke Tokita, Toshio Takagi, Aaron T Mayer, and Hideki Ishida. Unveiling spatial immune cell profile in kidney allograft rejections using 36-plex immunofluorescence imaging. *Transplantation*, 108(12):2446–2457, 2024. 5

[8] Wenpin Hou and Zhicheng Ji. Assessing gpt-4 for cell type annotation in single-cell rna-seq analysis. *Nature methods*, 21(8):1462–1465, 2024. 2

[9] Zhi Huang, Federico Bianchi, Mert Yuksekgonul, Thomas J Montine, and James Zou. A visual–language foundation model for pathology image analysis using medical twitter. *Nature medicine*, 29(9):2307–2316, 2023. 1

[10] Kayla C Jackson and Lior Pachter. A standard for sharing spatial transcriptomics data. *Cell Genomics*, 3(8), 2023. 1

[11] Shivam Kalra, Hamid R Tizhoosh, Charles Choi, Sultaan Shah, Phedias Diamandis, Clinton JV Campbell, and Liron Pantanowitz. Yottixel–an image search engine for large archives of histopathology whole slide images. *Medical Image Analysis*, 65:101757, 2020. 4

[12] ELHAM Karimi, N Simo, N Milet, W TE, A ALSH, ND QU, L AIL, R ABS, A ALIND, ND MORRIS GOODMA, et al. Method of the year 2024: spatial proteomics. *Nat Methods*, 21:2195–2196, 2024. 1

[13] Sumeer Ahmad Khan, Xabier Martinez de Morentin, Vincenzo Lagani, Robert Lehmann, Abdel Rahman Alsabbagh, Mahmoud Zahran, Narsis A Kiani, David Gomez-Cabrero, and Jesper Tegnér. Spell: Spatial prompting with chain-of-thought for zero-shot learning in spatial transcriptomics. In *ICLR 2025 Workshop on Machine Learning for Genomics Explorations*, 2025. 2

[14] Wasif Khan, Seowung Leem, Kyle B See, Joshua K Wong, Shaoting Zhang, and Ruogu Fang. A comprehensive survey of foundation models in medicine. *IEEE Reviews in Biomedical Engineering*, 2025. 1

[15] Junbum Kim, Samir Rustam, Juan Miguel Mosquera, Scott H Randell, Renat Shaykhiev, André F Rendeiro, and Olivier Elemento. Unsupervised discovery of tissue architecture in multiplexed imaging. *Nature methods*, 19(12):1653–1661, 2022. 4, 6

[16] Ayano Kondo, Monee McGrady, Dhiraj Nallapothula, Hira Ali, Alexandro E Trevino, Amy Lam, Ryan Preska, H Blaize D'Angio, Zhenqin Wu, Lauren N Lopez, et al. Spatial proteomics of human diabetic kidney disease, from health to class iii. *Diabetologia*, 67(9):1962–1979, 2024. 4

[17] Ming Y Lu, Bowen Chen, Drew FK Williamson, Richard J Chen, Ivy Liang, Tong Ding, Guillaume Jaume, Igor Odintsov, Long Phi Le, Georg Gerber, et al. A visual-language foundation model for computational pathology. *Nature Medicine*, 30(3):863–874, 2024. 1

[18] Renqian Luo, Liai Sun, Yingce Xia, Tao Qin, Sheng Zhang, Hoifung Poon, and Tie-Yan Liu. Biogpt: generative pre-trained transformer for biomedical text generation and mining. *Briefings in bioinformatics*, 23(6):bbac409, 2022. 1

[19] Sergio Marco Salas, Louis B Kuemmerle, Christoffer Mattsson-Langseth, Sebastian Tismeyer, Christophe Avenel, Taobo Hu, Habib Rehman, Marco Grillo, Paulo Czarnewski, Saga Helgadottir, et al. Optimizing xenium in situ data utility by quality assessment and best-practice analysis workflows. *Nature Methods*, pages 1–11, 2025. 1

[20] Syed Asad Rizvi, Daniel Levine, Aakash Patel, Shiyang Zhang, Eric Wang, Sizhuang He, David Zhang, Cerise Tang, Zhuoyang Lyu, Rayyan Darji, et al. Scaling large language

models for next-generation single-cell analysis. *bioRxiv*, pages 2025–04, 2025. 2

[21] Anna C Schaar, Alejandro Tejada-Lapuerta, Giovanni Palla, Robert Gutgesell, Lennard Halle, Mariia Minaeva, Larsen Vornholz, Leander Dony, Francesca Drummer, Mojtaba Bahrami, et al. Nicheformer: a foundation model for single-cell and spatial omics. *bioRxiv*, pages 2024–04, 2024. 1

[22] Karan Singhal, Shekoofeh Azizi, Tao Tu, S Sara Mahdavi, Jason Wei, Hyung Won Chung, Nathan Scales, Ajay Tanwani, Heather Cole-Lewis, Stephen Pfohl, et al. Large language models encode clinical knowledge. *Nature*, 620 (7972):172–180, 2023. 1

[23] Karan Singhal, Shekoofeh Azizi, Tao Tu, S Sara Mahdavi, Jason Wei, Hyung Won Chung, Nathan Scales, Ajay Tanwani, Heather Cole-Lewis, Stephen Pfohl, et al. Large language models encode clinical knowledge. *Nature*, 620 (7972):172–180, 2023. 2

[24] Karan Singhal, Tao Tu, Juraj Gottweis, Rory Sayres, Ellery Wulczyn, Mohamed Amin, Le Hou, Kevin Clark, Stephen R Pfohl, Heather Cole-Lewis, et al. Toward expert-level medical question answering with large language models. *Nature Medicine*, pages 1–8, 2025. 1, 2

[25] Patrik L Ståhl, Fredrik Salmén, Sanja Vickovic, Anna Lundmark, José Fernández Navarro, Jens Magnusson, Stefania Giacomello, Michaela Asp, Jakub O Westholm, Mikael Huss, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*, 353 (6294):78–82, 2016. 1

[26] Mathias Uhlén, Linn Fagerberg, Björn M Hallström, Cecilia Lindskog, Per Oksvold, Adil Mardinoglu, Åsa Sivertsson, Caroline Kampf, Evelina Sjöstedt, Anna Asplund, et al. Tissue-based map of the human proteome. *Science*, 347 (6220):1260419, 2015. 3

[27] Marco Varrone, Daniele Tavernari, Albert Santamaria-Martínez, Logan A Walsh, and Giovanni Ciriello. Cellcharter reveals spatial cell niches associated with tissue remodeling and cell plasticity. *Nature genetics*, 56(1):74–84, 2024. 4, 6

[28] Britta Velten and Oliver Stegle. Principles and challenges of modeling temporal and spatial omics data. *Nature Methods*, 20(10):1462–1474, 2023. 1

[29] Chloe Wang, Haotian Cui, Andrew Zhang, Ronald Xie, Hani Goodarzi, and Bo Wang. scgpt-spatial: Continual pretraining of single-cell foundation model for spatial transcriptomics. *bioRxiv*, pages 2025–02, 2025. 1, 2

[30] Hanchen Wang, Yichun He, Paula P Coelho, Matthew Bucci, Abbas Nazir, Bob Chen, Linh Trinh, Serena Zhang, Kexin Huang, Vineethkrishna Chandrasekar, et al. Spatialagent: An autonomous ai agent for spatial biology. *bioRxiv*, pages 2025–04, 2025. 2

[31] Xuezhi Wang, Jason Wei, Dale Schuurmans, Quoc Le, Ed Chi, Sharan Narang, Aakanksha Chowdhery, and Denny Zhou. Self-consistency improves chain of thought reasoning in language models. *arXiv preprint arXiv:2203.11171*, 2022. 7

[32] Zifeng Wang, Zhenbang Wu, Dinesh Agarwal, and Jimeng Sun. Medclip: Contrastive learning from unpaired medical images and text. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing. Conference on Empirical Methods in Natural Language Processing*, page 3876, 2022. 1

[33] ER Watson, A Taherian Fard, and JC Mar. Computational methods for single-cell imaging and omics data integration. front. mol. biosci. 8, 768106, 2021. 1

[34] Ebony Rose Watson, Atefeh Taherian Fard, and Jessica Cara Mar. Computational methods for single-cell imaging and omics data integration. *Frontiers in molecular biosciences*, 8:768106, 2022. 1

[35] Yingcheng Wu, Yifei Cheng, Xiangdong Wang, Jia Fan, and Qiang Gao. Spatial omics: Navigating to the golden era of cancer research. *Clinical and Translational Medicine*, 12(1): e696, 2022. 1

[36] Zhenqin Wu, Ayano Kondo, Monee McGrady, Ethan AG Baker, Benjamin Chidester, Eric Wu, Maha K Rahim, Nathan A Bracey, Vivek Charu, Raymond J Cho, et al. Discovery and generalization of tissue structures from spatial omics data. *Cell Reports Methods*, 4(8), 2024. 4, 6

[37] A Xiaowei. Method of the year 2020: Spatially resolved transcriptomics. *Nat. Methods*, 18(1):10–1038, 2021. 1

[38] Fan Yang, Wenchuan Wang, Fang Wang, Yuan Fang, Duyu Tang, Junzhou Huang, Hui Lu, and Jianhua Yao. scbert as a large-scale pretrained deep language model for cell type annotation of single-cell rna-seq data. *Nature Machine Intelligence*, 4(10):852–866, 2022. 1

[39] Sheng Zhang, Yanbo Xu, Naoto Usuyama, Hanwen Xu, Jaspreet Bagga, Robert Tinn, Sam Preston, Rajesh Rao, Mu Wei, Naveen Valluri, et al. Biomedclip: a multimodal biomedical foundation model pretrained from fifteen million scientific image-text pairs. *arXiv preprint arXiv:2303.00915*, 2023. 1

[40] Weiruo Zhang, Irene Li, Nathan E Reticker-Flynn, Zinaida Good, Serena Chang, Nikolay Samusik, Saumyaa Saumyaa, Yuanyuan Li, Xin Zhou, Rachel Liang, et al. Identification of cell types in multiplexed in situ images by combining protein expression and spatial information using celesta. *Nature methods*, 19(6):759–769, 2022. 1