



SCHOOL OF
**COMPUTING &
DATA SCIENCE**
The University of Hong Kong

Molecular Biomedical Imaging in the Era of Spatial Omics

Zhenqin Wu

School of Computing and Data Science
The University of Hong Kong

2025.11.28, Croucher Advanced Study Institute (ASI)

Spatial Omics



Bulk



Single cell

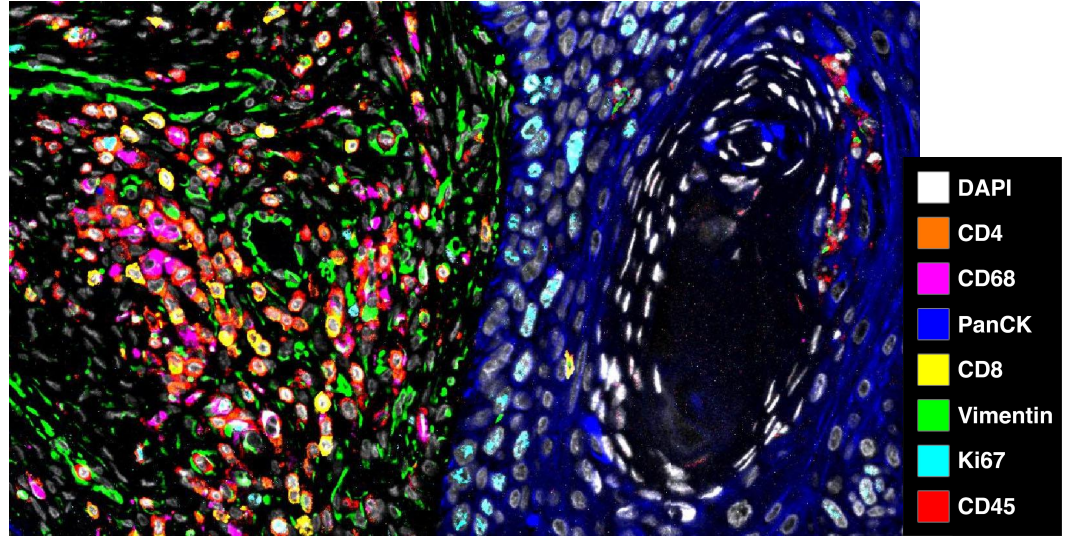


Spatial omics

Opportunities and Challenges

High-resolution, multi-dimensional in situ characterization of molecules

- Spatial context of:
 - Molecules in cell
 - Cells in tissue
- Cell-cell interaction
- Spatial signatures as disease biomarkers

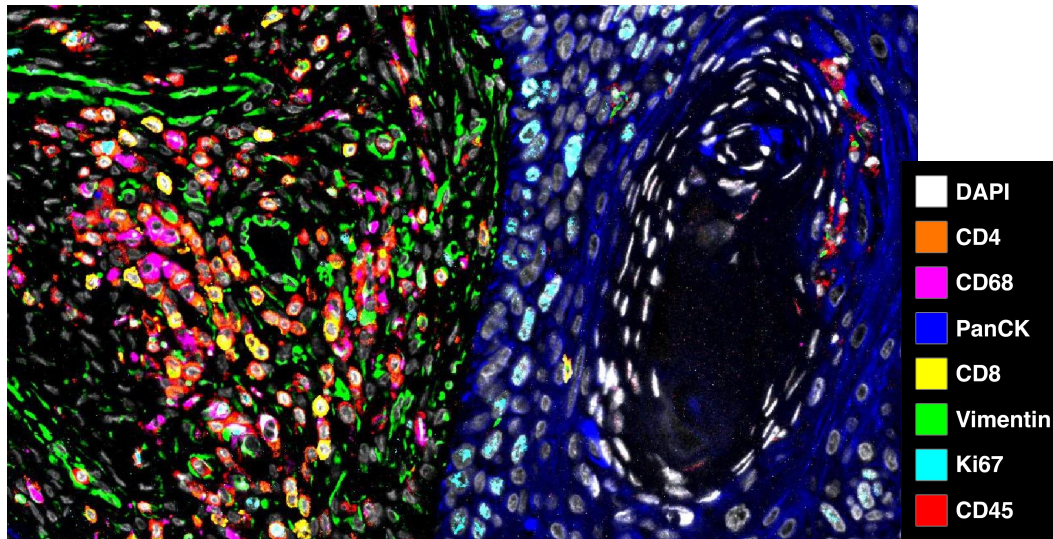


CODEX image of a head-and-neck cancer sample

Opportunities and Challenges

High-resolution, multi-dimensional in situ characterization of molecules

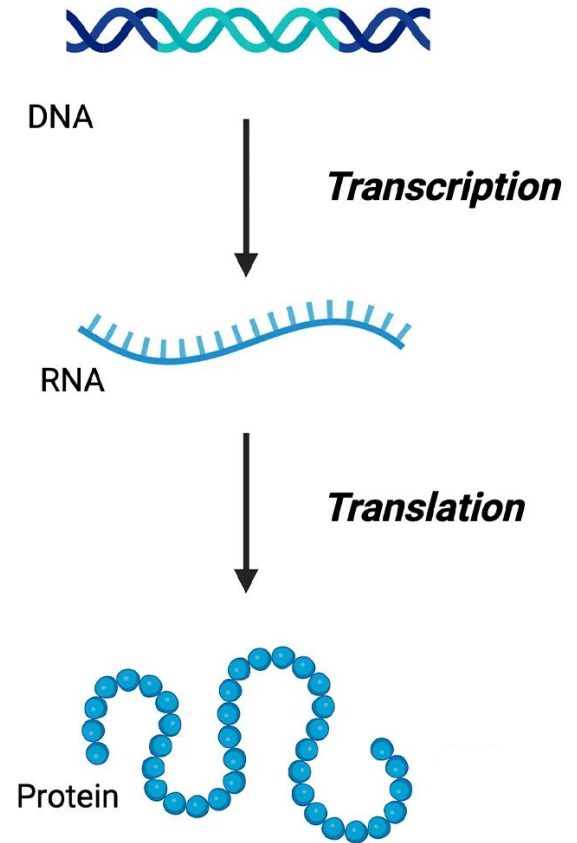
- Data volume
- Convoluted data analysis pipeline
- Complex spatial relations between cells
- Difficulties in data integration
- Spatially aware analytical tools



CODEX image of a head-and-neck cancer sample

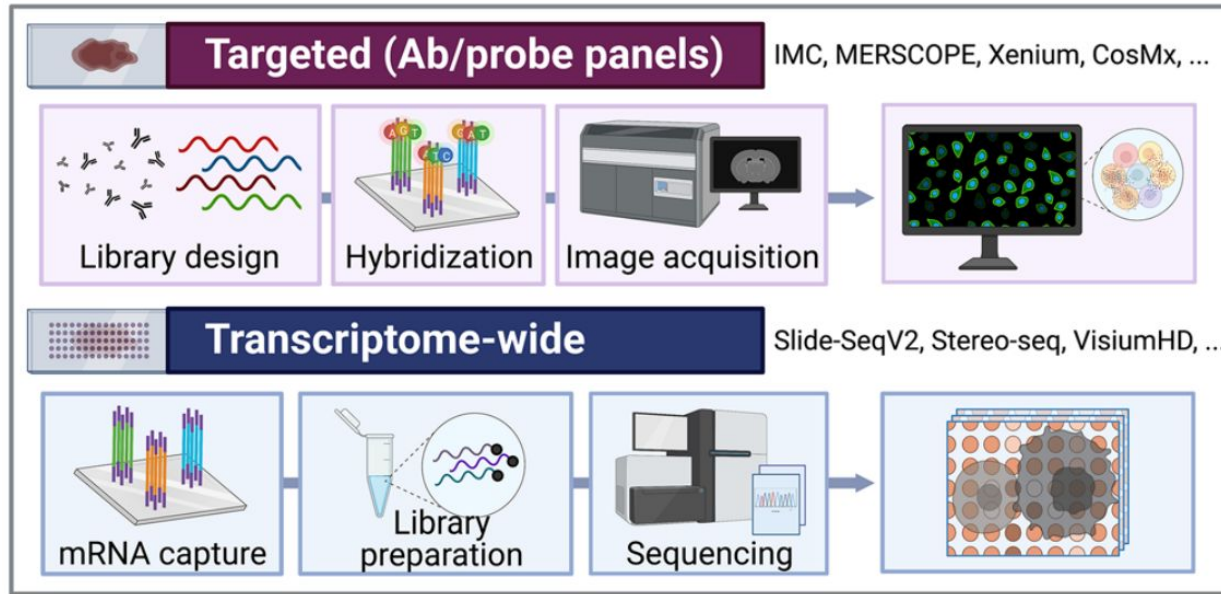
Data modalities

- Transcriptomics
 - Visium, Visium HD, Xenium (10x)
 - MERSCOPE (Vizgen)
 - Stereo-seq (BGI)
 - GeoMx, CoxMx (NanoString/Bruker)
 - Slide-seq, Slide-tag, HDST, Spatial-ATAC-seq, etc.
- Proteomics
 - PhenoCycler/CODEX (Akoya)
 - Hyperion/IMC (Standard Biotools)
 - MIBIscope (Ionpath)

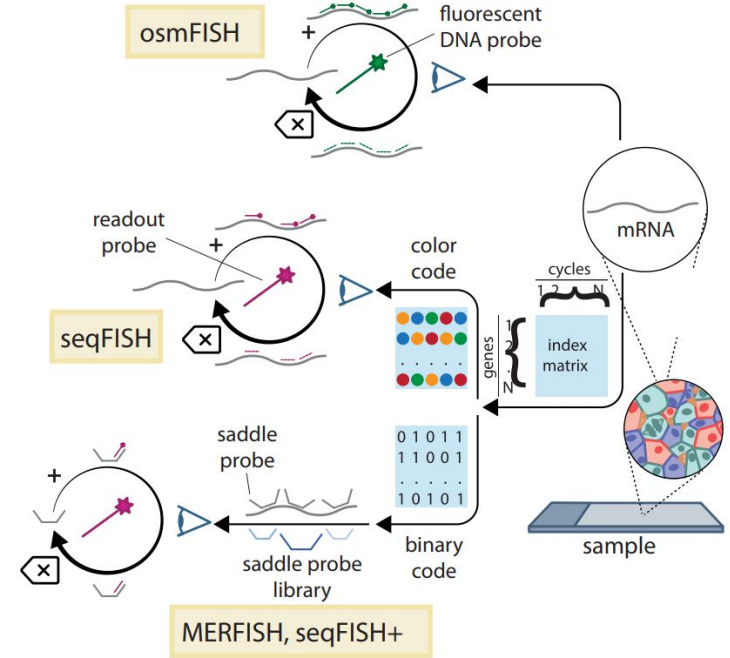
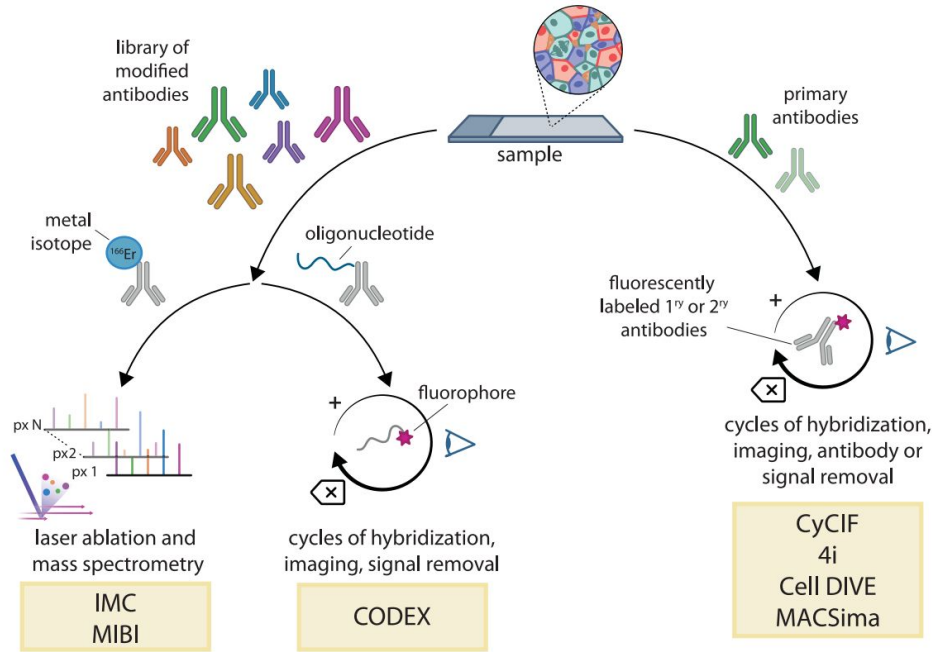


Data modalities

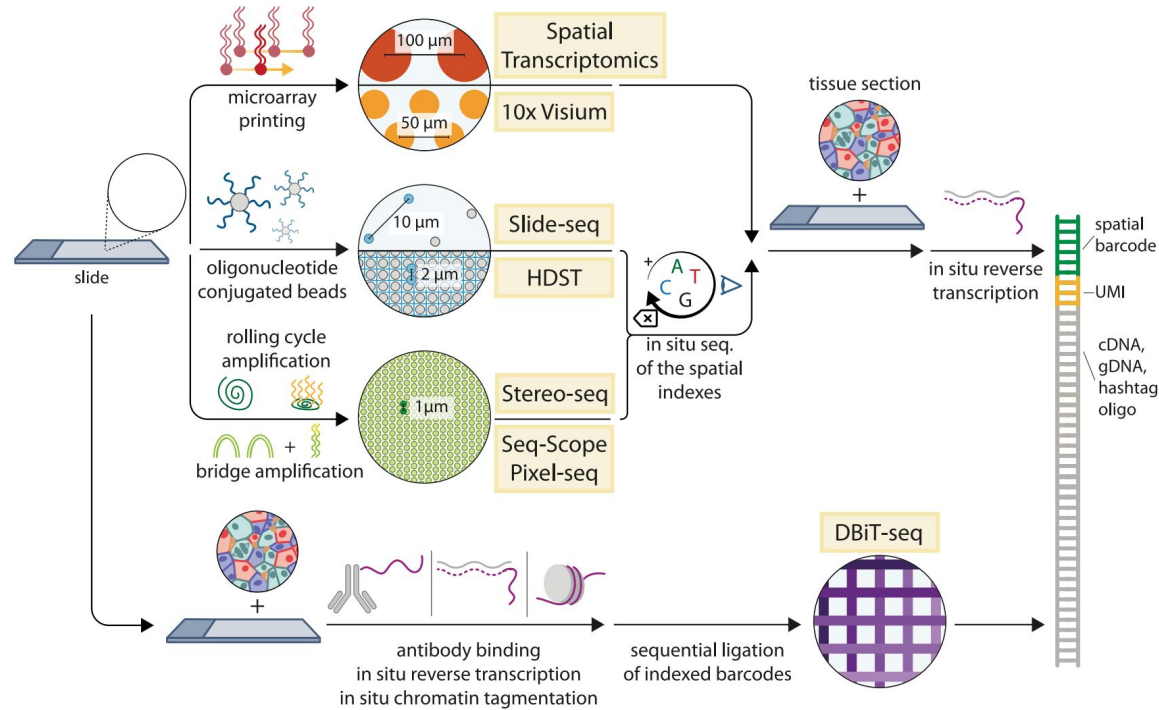
- Targeted versus untargeted (transcriptome-wide)

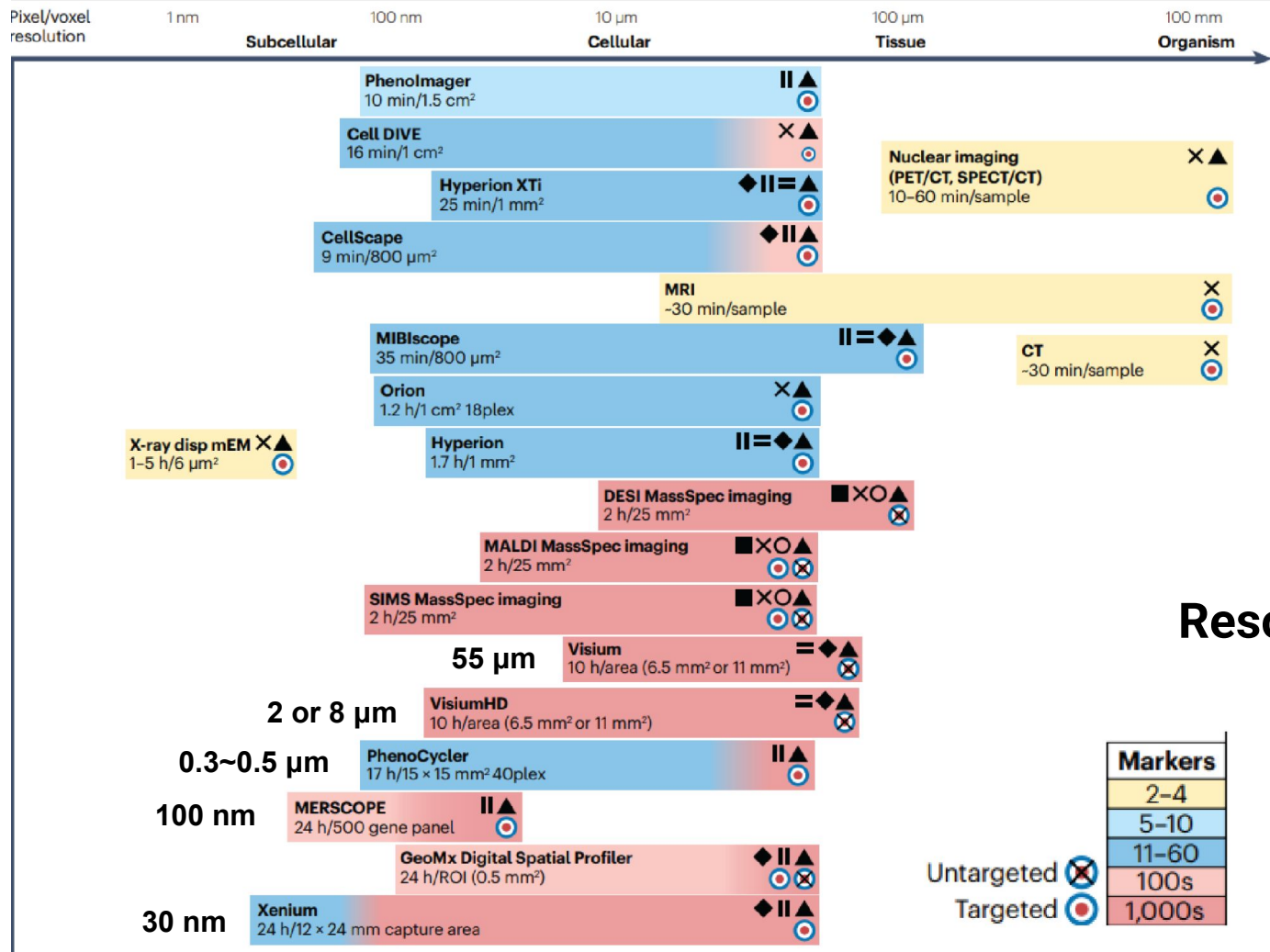


Imaging-based (targeted) acquisition



Sequencing-based (full-transcriptome) acquisition



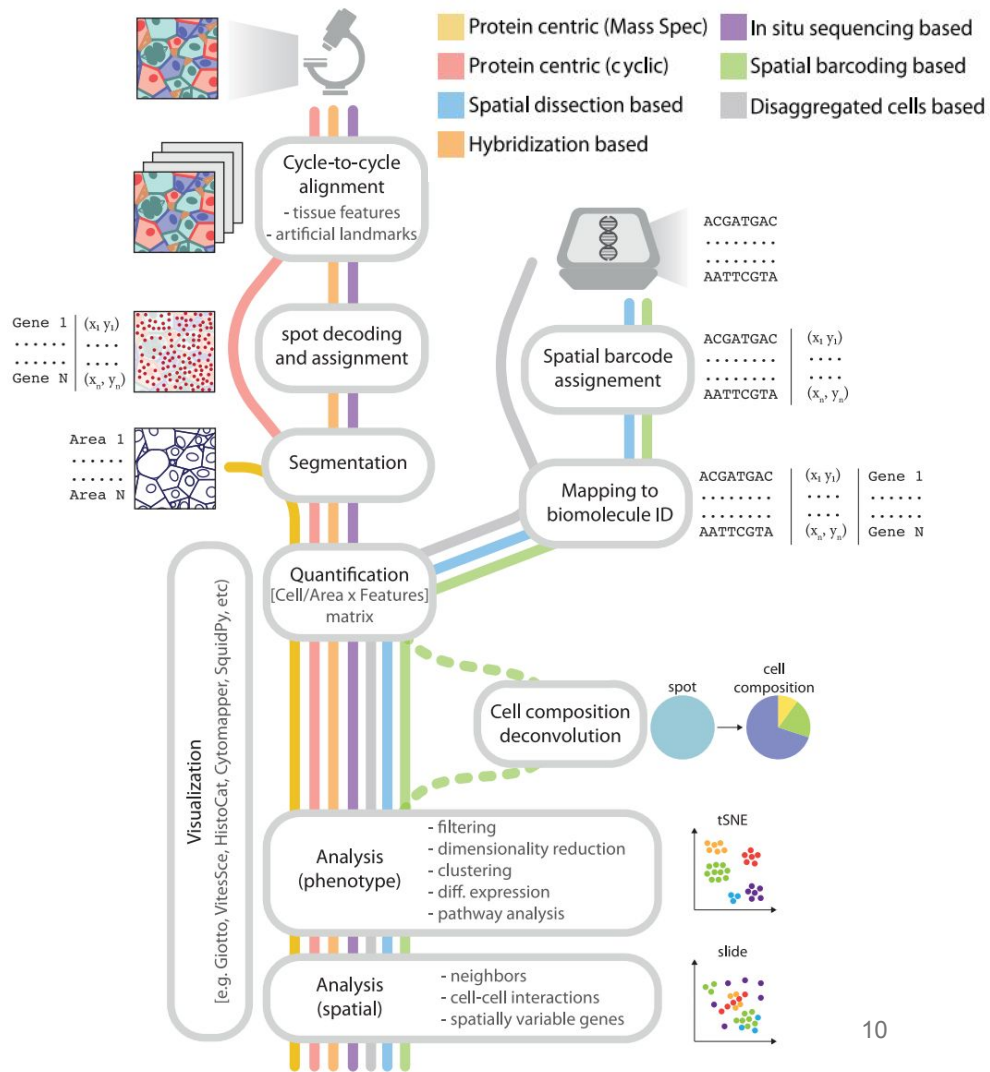


Resolution and Plex

Carstens, Julianne L., et al. "Spatial multiplexing and omics." *Nature Reviews Methods Primers* 4.1 (2024): 54.

Data analysis

- **Goals:** identify spatial entities of interest from the data

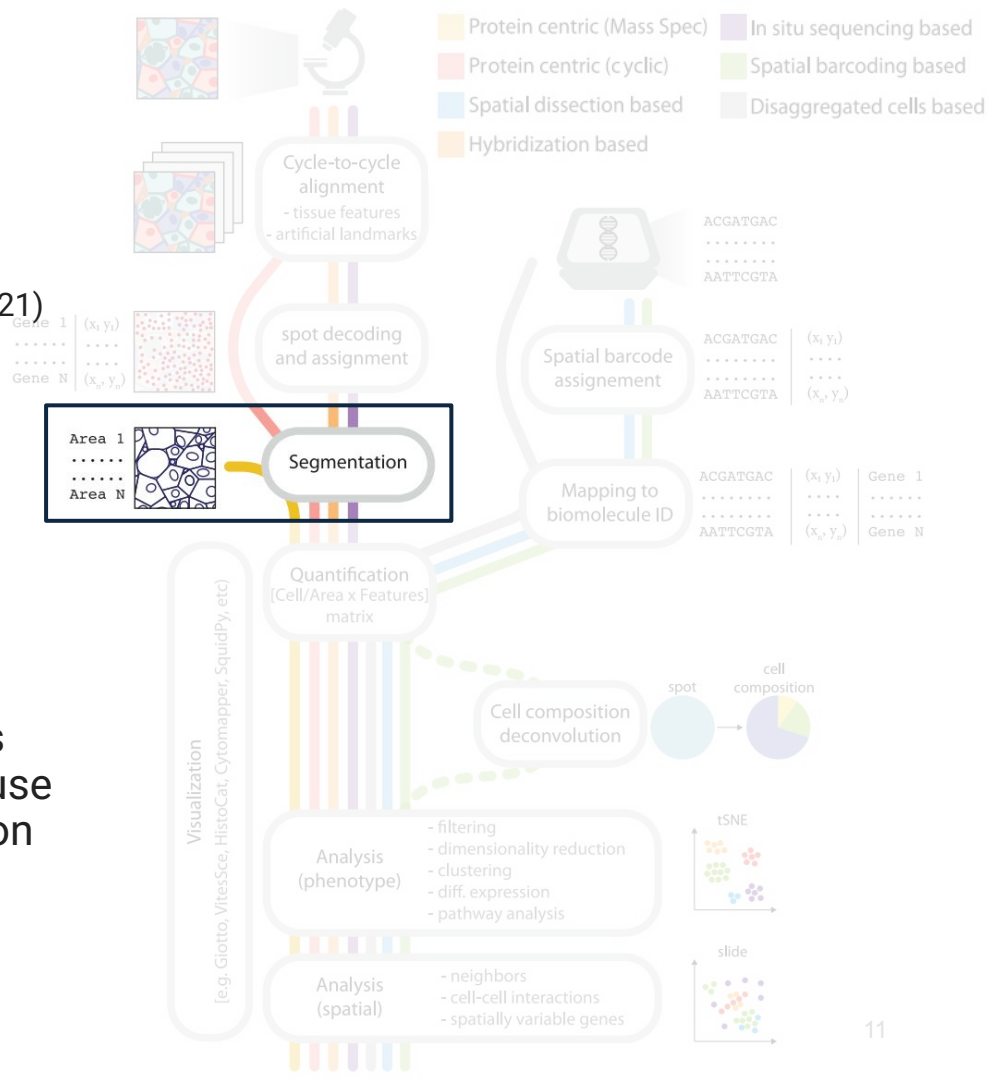


Segmentation

- DeepCell / Mesmer (Greenwald et al., 2021)
- StarDist (Schmidt et al., 2018)
- Cellpose (Pachitariu & Stringer, 2024)
- Xenium's cell segmentation

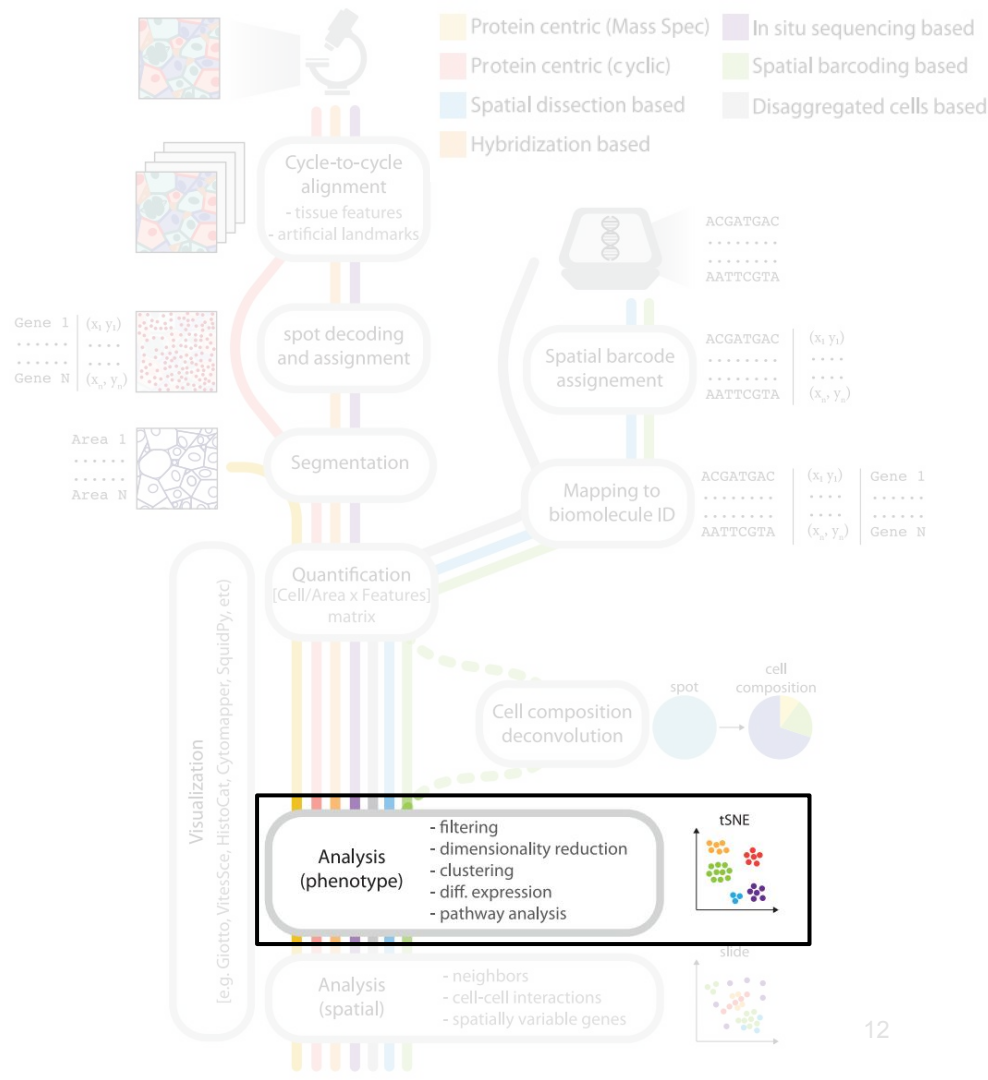
etc.

Accurate whole-cell segmentation needs membrane markers; many applications use watershed / dilated nucleus segmentation instead => **spill-over** issues

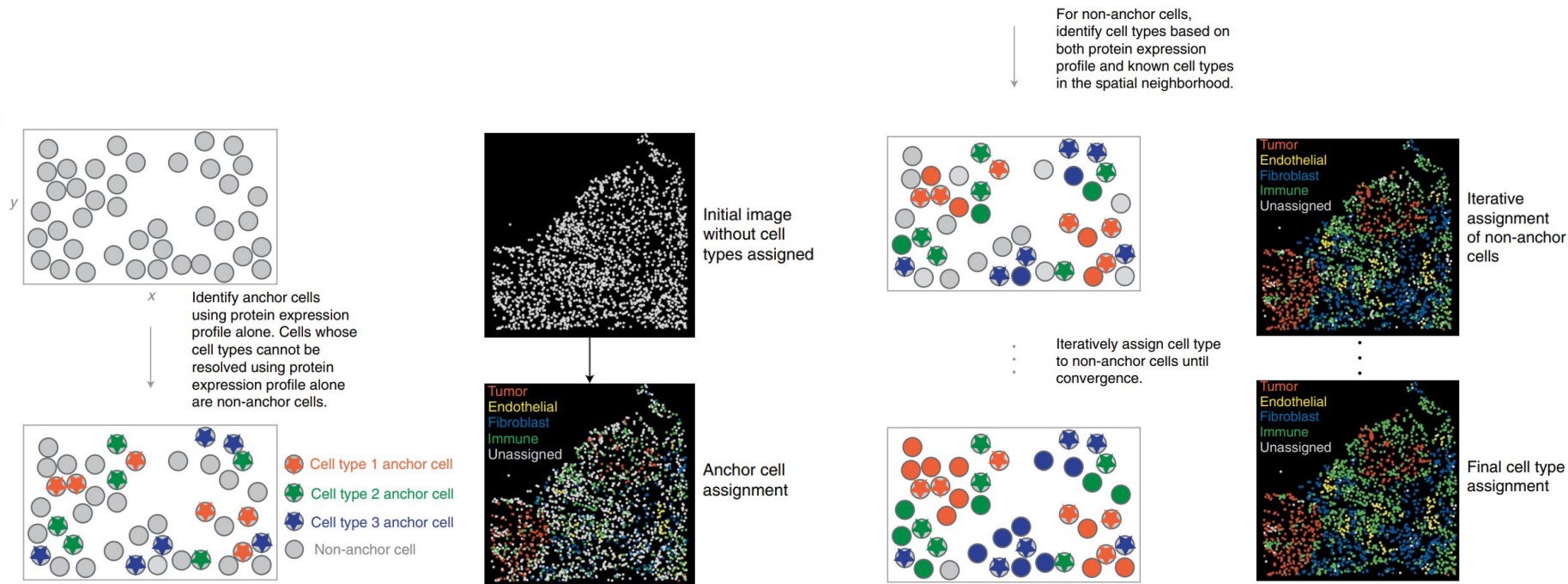


Cell Type Annotation

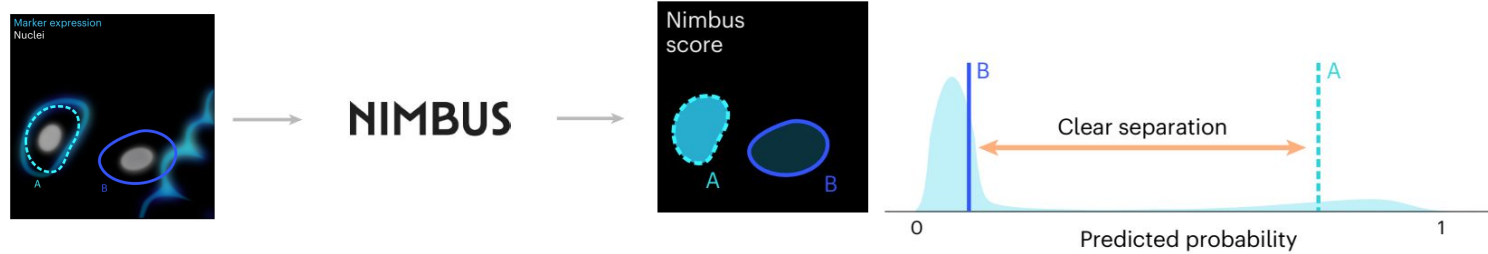
- Gating on lineage markers
- Unsupervised clustering via Leiden (Traag et al., 2019) or Louvain (Blondel et al., 2008) community detection
- Reference mapping via machine learning tools: CellTypist (Conde et al., 2022), CellHint (Xu et al., 2023)
- Direction prediction via machine learning: MAPS (Shaban et al. 2024)
- Ask GPT (Hou and Ji, 2024)



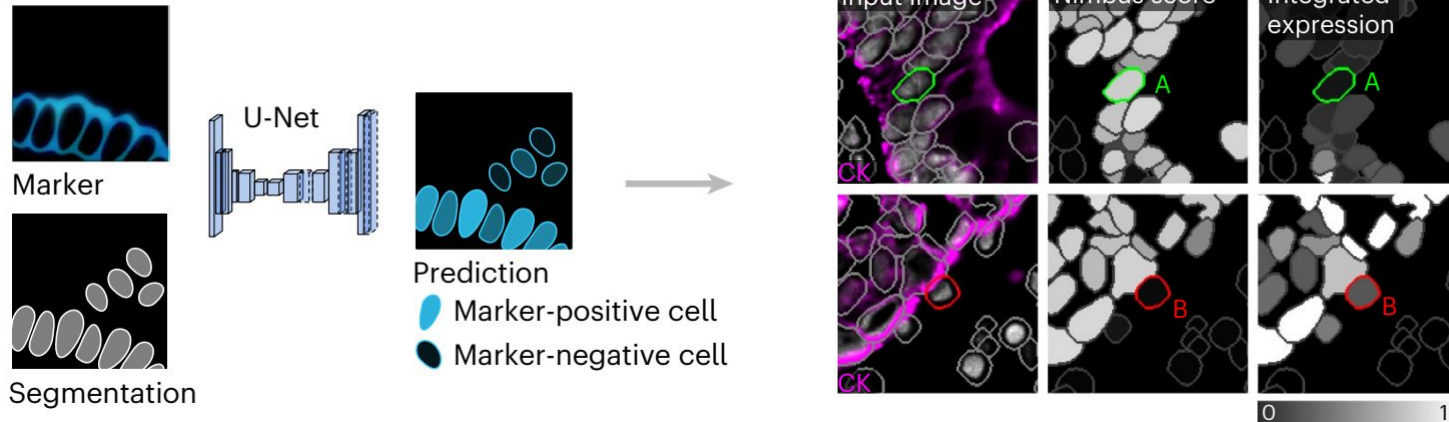
CELESTA: Iterative refinement of cell type annotations based on protein expression and spatial information



NIMBUS: classification of cellular expression in multiplexed imaging



(ii) Train Nimbus model



Neighborhood Annotation

Based on cell types

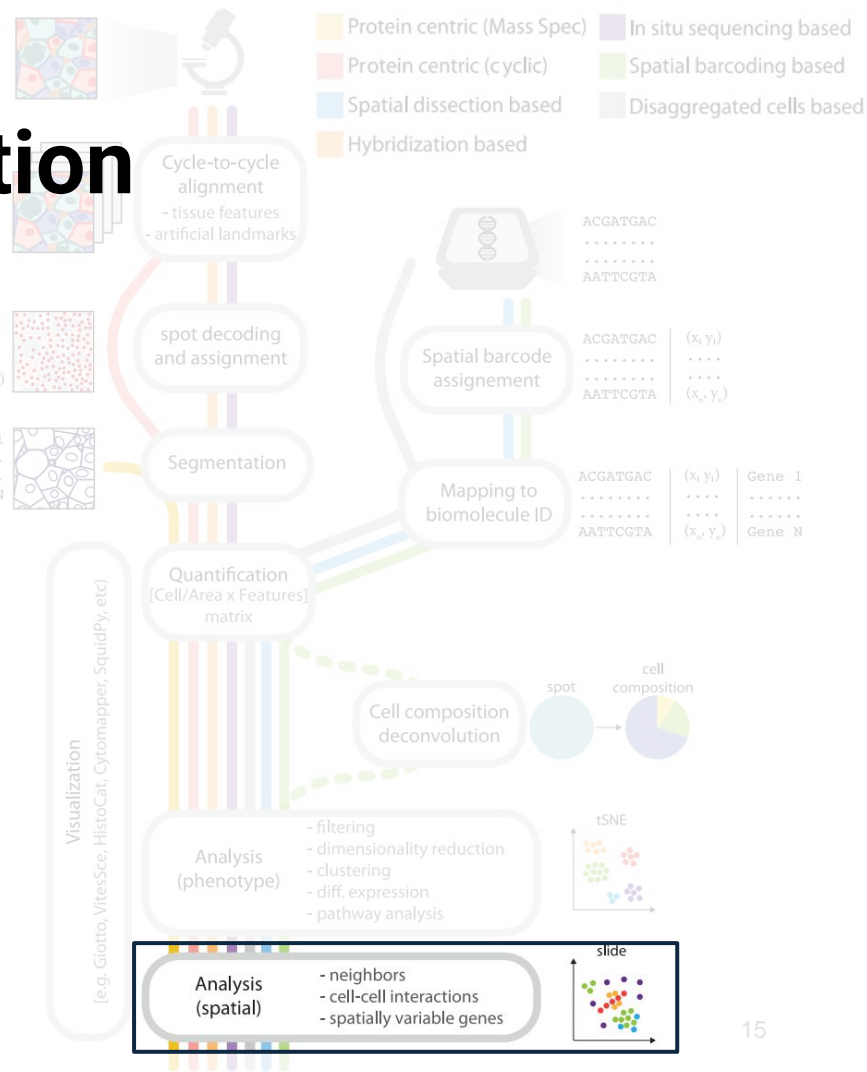
- Cellular Neighborhood (Schürch et al., 2020)
- CytoCommunity (Hu et al., 2023)
- scNiches (Qian et al., 2025)

Based on biomarker expression: single-cell, smoothed, covariance

- UTAG (Kim et al., 2022)
- SpiceMix (Chidester et al., 2023)
- CellCharter (Varrone et al., 2024)
- SCGP (Wu et al., 2024)
- stDyer (Xu et al., 2025)

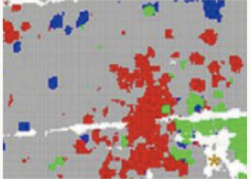
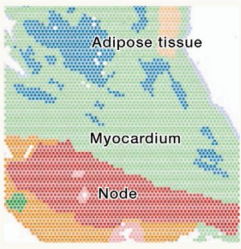
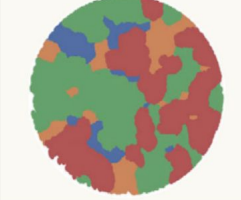
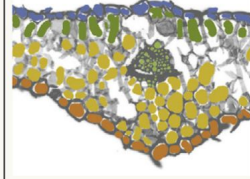
With addition of morphology:

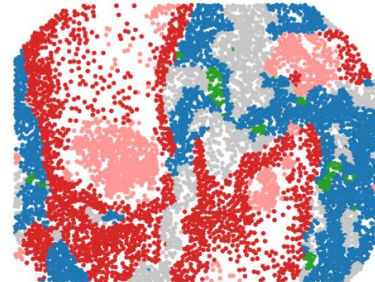
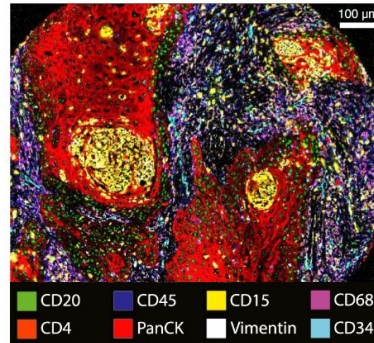
- SpaGCN (Hu et al., 2021)
- stLearn (Pham et al., 2023)
- GIST (Ge et al., 2025)



Neighborhood Annotation

- To elucidate the structure and function of the organs
- To understand the heterogeneity of tumor microenvironment and tumor-immune interplay

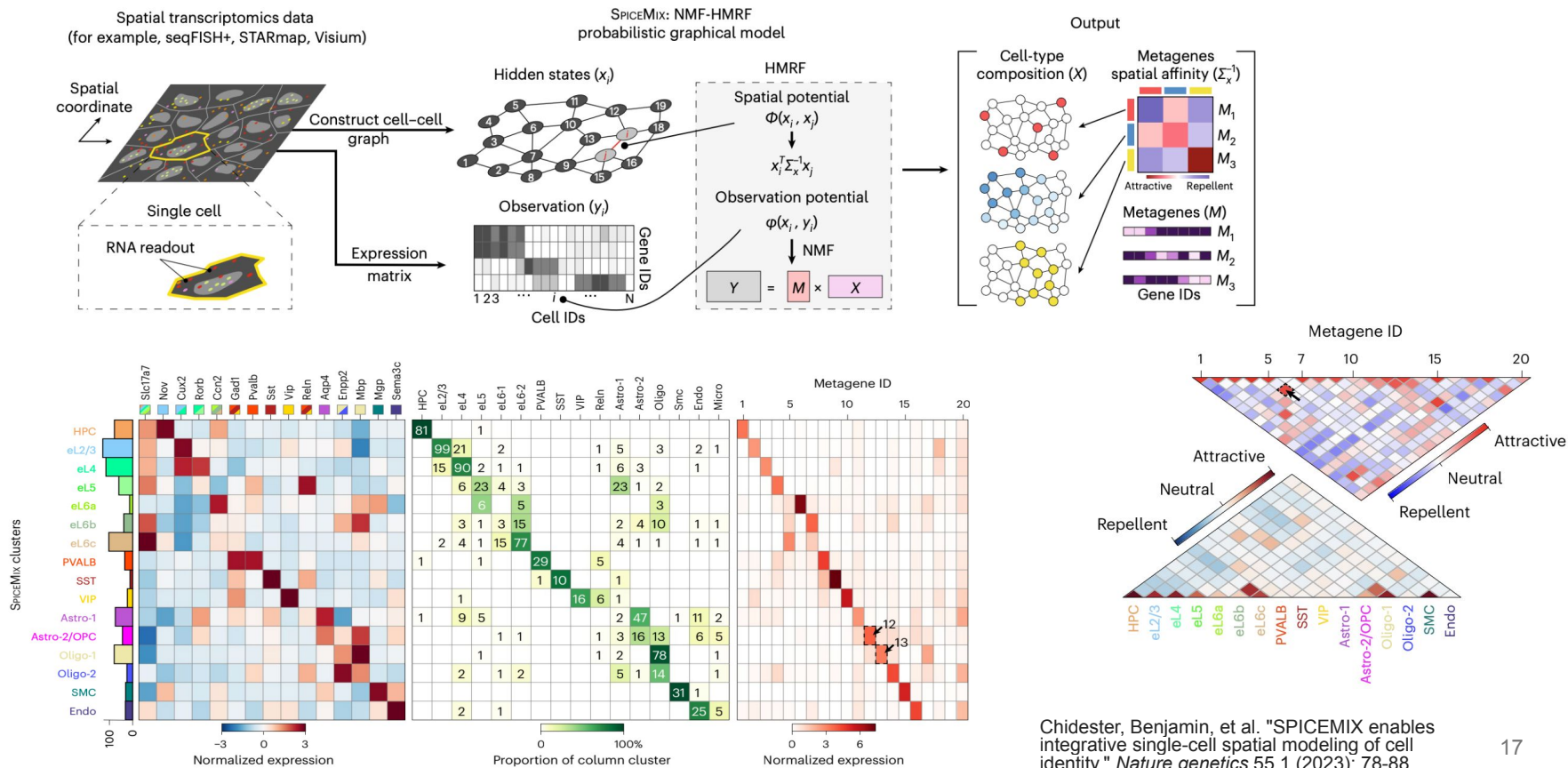
	Liver seq-Scope	Heart Visium	Testis Slide-seq	Leaves Stereo-seq
Spatial organization	 ● MΦ-Kupffer ● HSC-A ● MΦ-inflamed	 Adipose tissue Myocardium Node	 Stages ● I-III ● IV-VI ● VII-VIII ● IX-XII	 ● Upper epidermal wall ● Lower epidermal wall
Function	Inflammation and injury	Impulse conduction	Spermatogenesis	Protection and transportation



Tissue Structure Partitions

- Immune cell-enriched stroma
- Stroma / Fibrosis
- Vessels
- Tumor Mass
- Granulocyte/Tumor

SpiceMix: joint modeling of cell identities, gene modules (metagenes), and spatial clusters



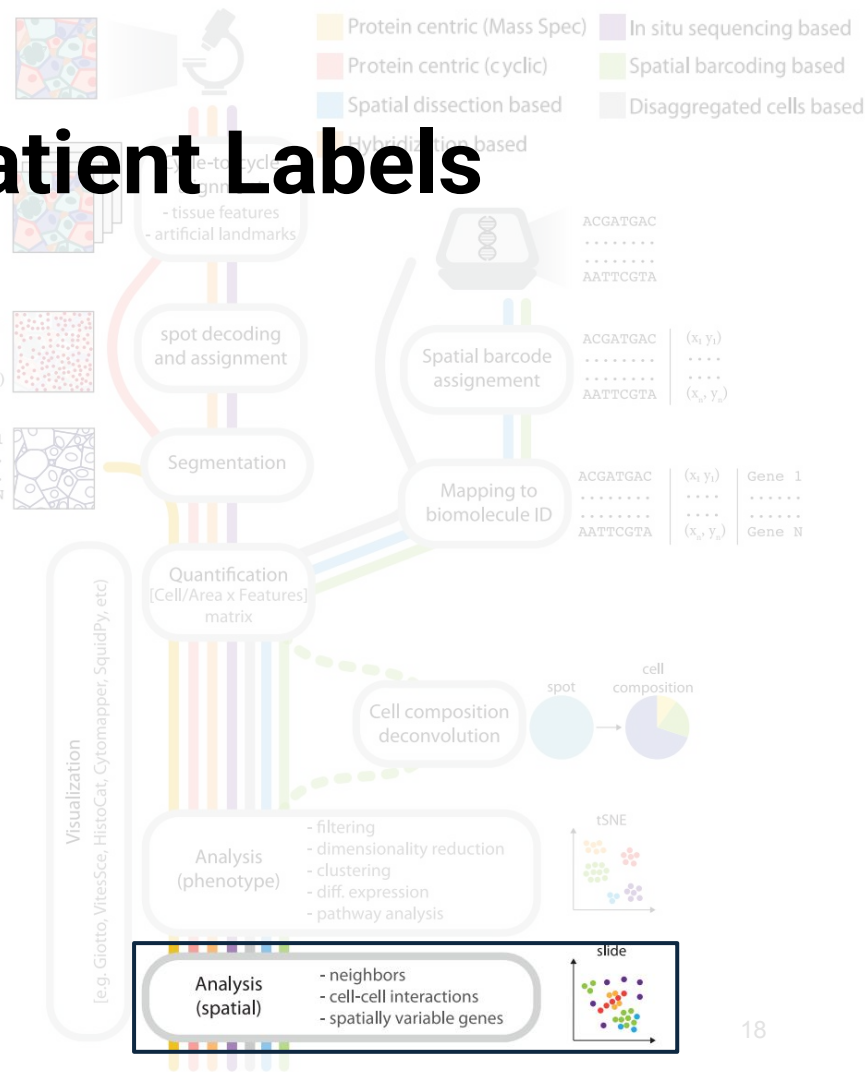
Neighborhood Annotation



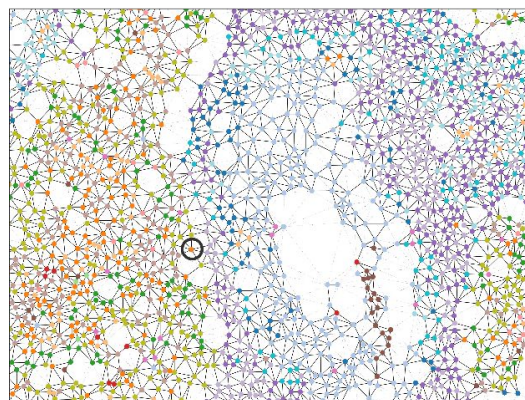
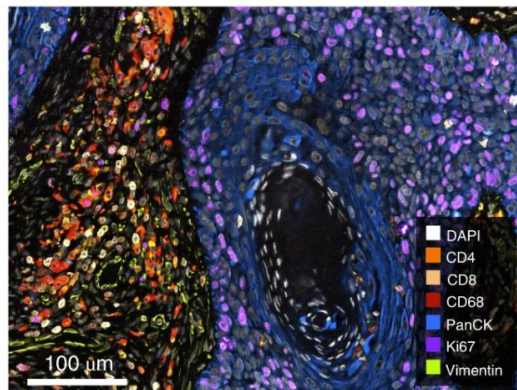
Patient Labels

Which neighborhood is relevant to a specific patient class? (tumor grades, prognosis, responders to therapy, etc.)

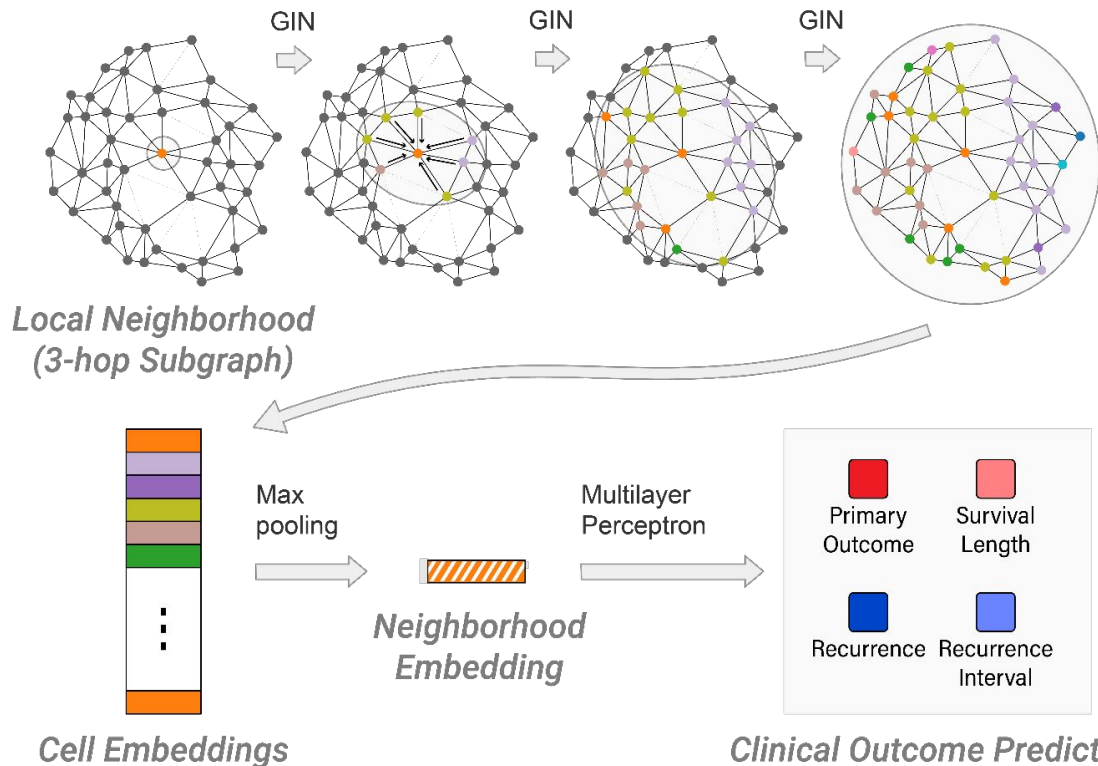
- Counts & proportions of cell types and/or different neighborhood classes
- Supervised modeling followed by feature interpretations
 - Cell graph (Wang et al., 2022)
 - SPACE-GM (Wu et al., 2022)
 - SORBET (Shimonov et al., 2024)
 - GNN (Ali et al., 2025)



SPACE-GM: Supervised Learning for Clinical Outcome Prediction



Spatial Graph of Cells

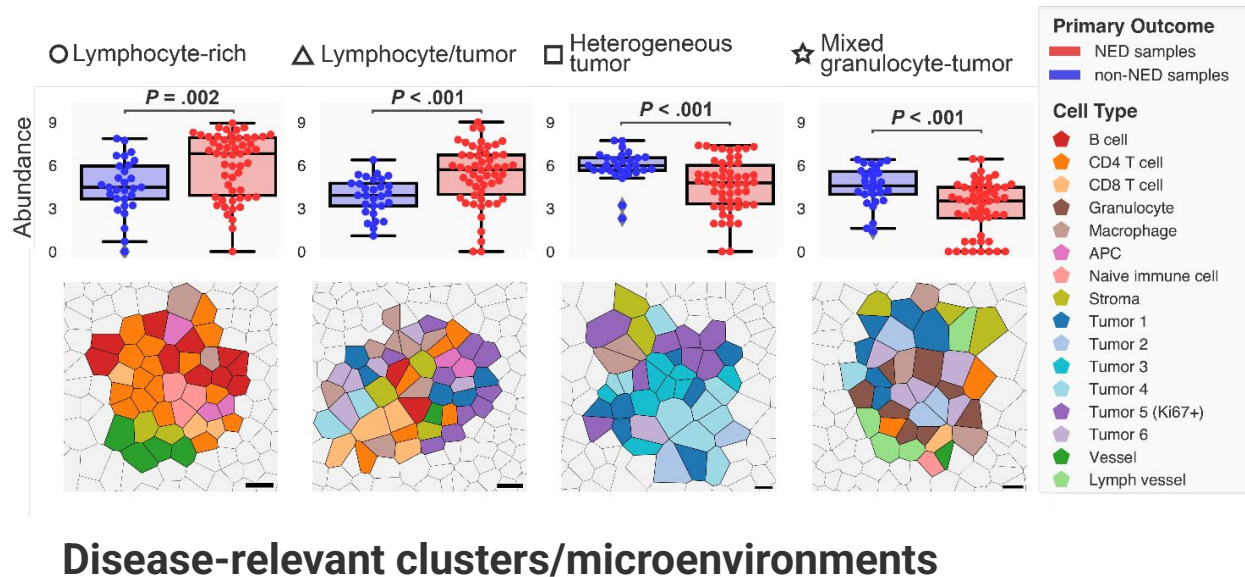
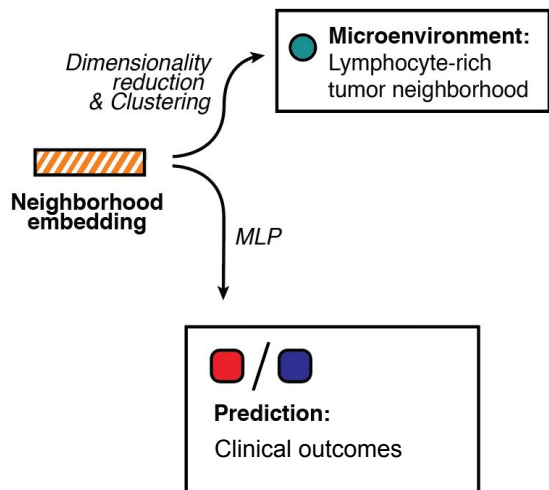


SPACE-GM: Supervised Learning for Clinical Outcome Prediction

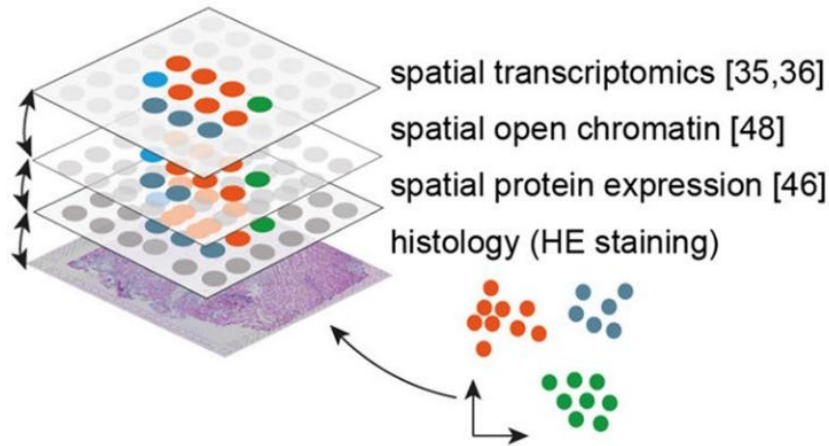
Neighborhoods

⇒ Embeddings

⇒ Clusters



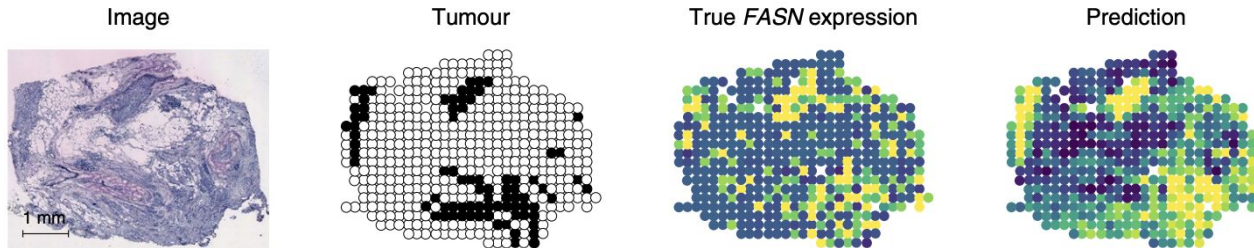
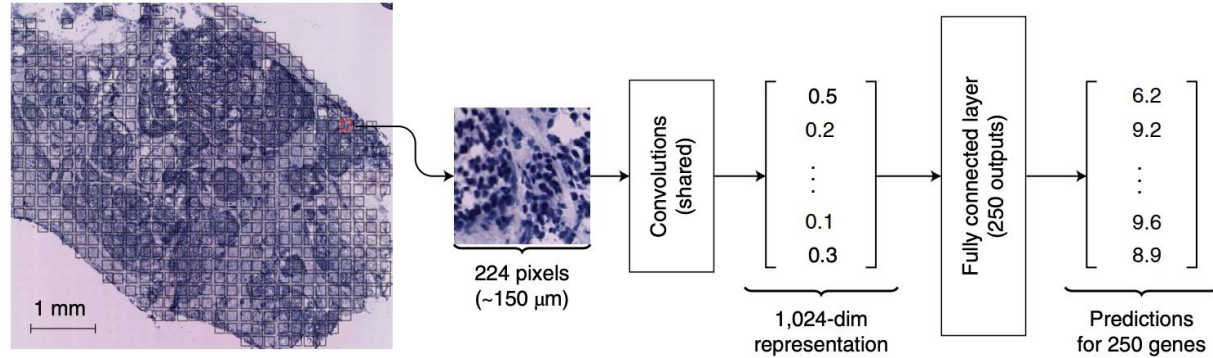
Spatial multi-omics



Key challenges:

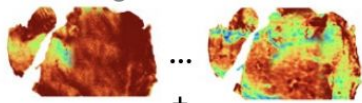
- Modality translation: H&E to spatial omics, H&E to cell type, etc.
- Data integration
- Joint analysis to reveal multi-omics insights

Predict gene expression with histopathology images



iScale: Scaling up spatial transcriptomics with H&E

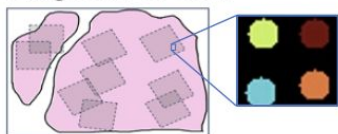
H&E image features



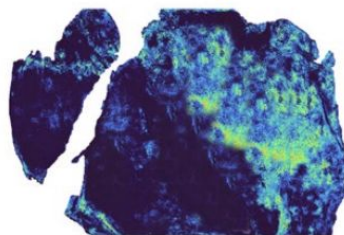
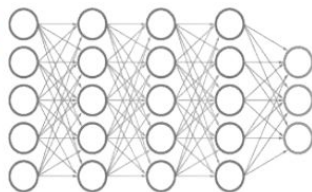
...

+

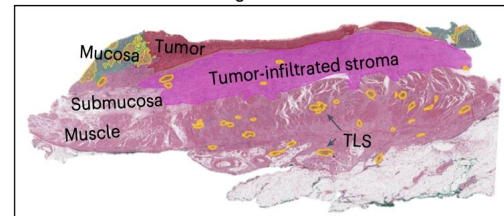
Integrated ST data



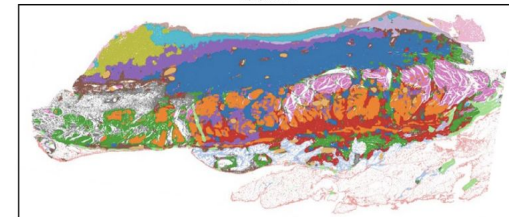
Super-resolution gene expression
across entire large tissue



Pathologist annotation

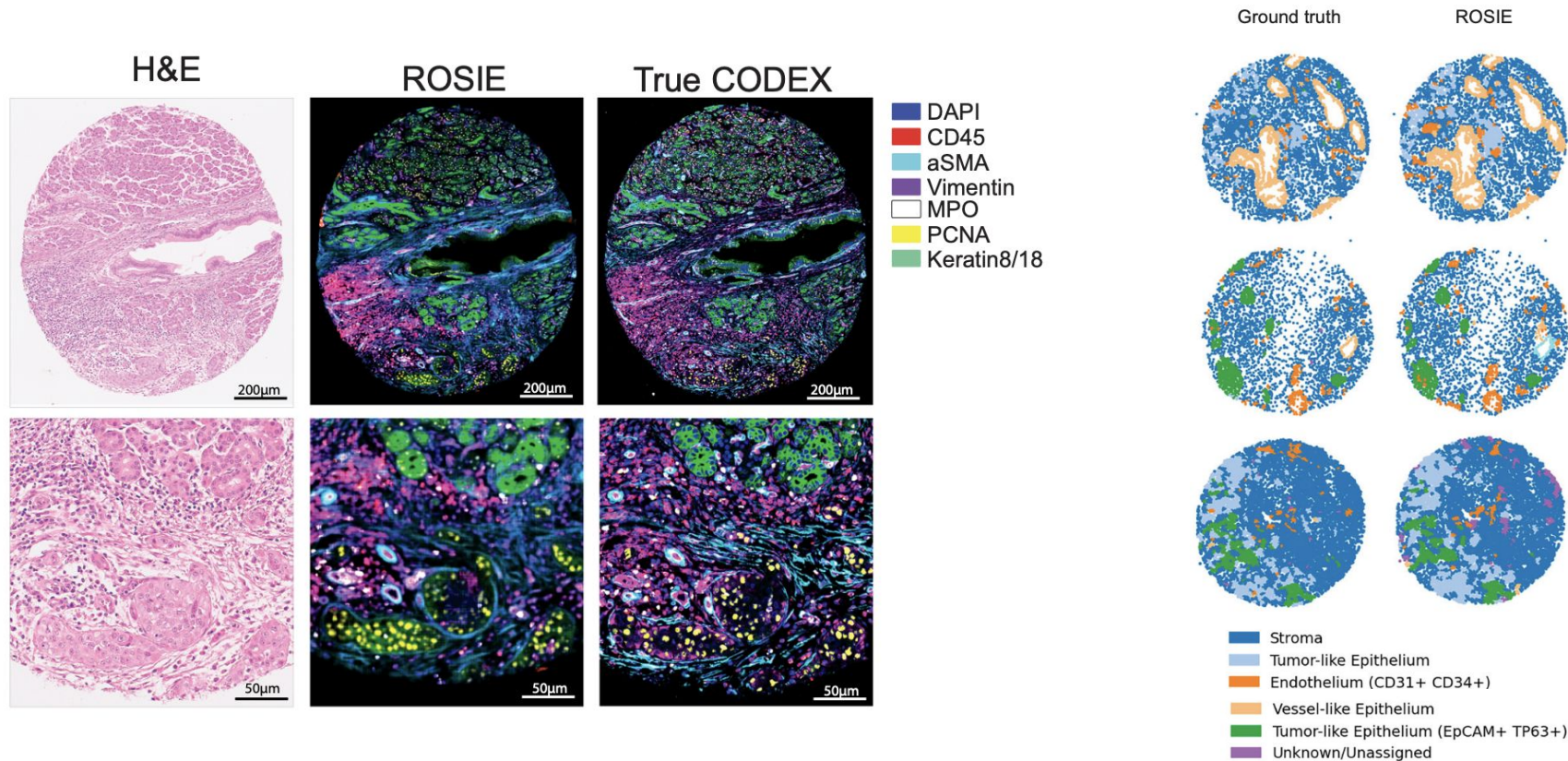


iSCALE

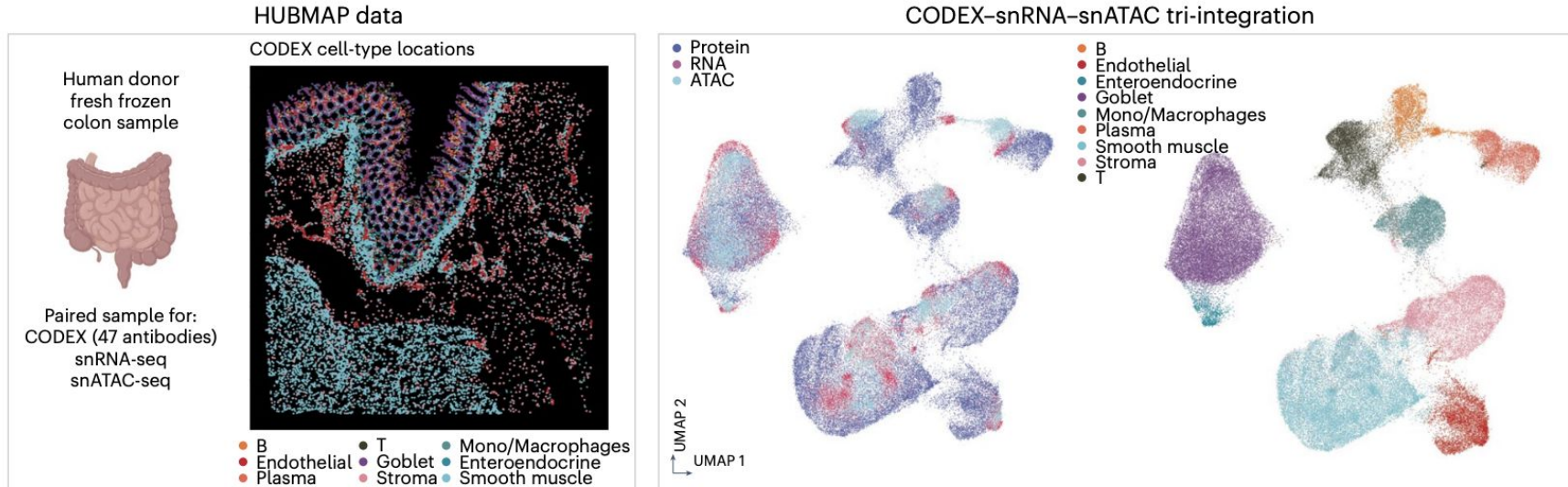


	MYH11	TFF2	ACTA2	CAPN8
Ground truth				
iSCALE-Seq	$r = 0.5037$ 	$r = 0.3680$ 	$r = 0.2703$ 	$r = 0.3868$

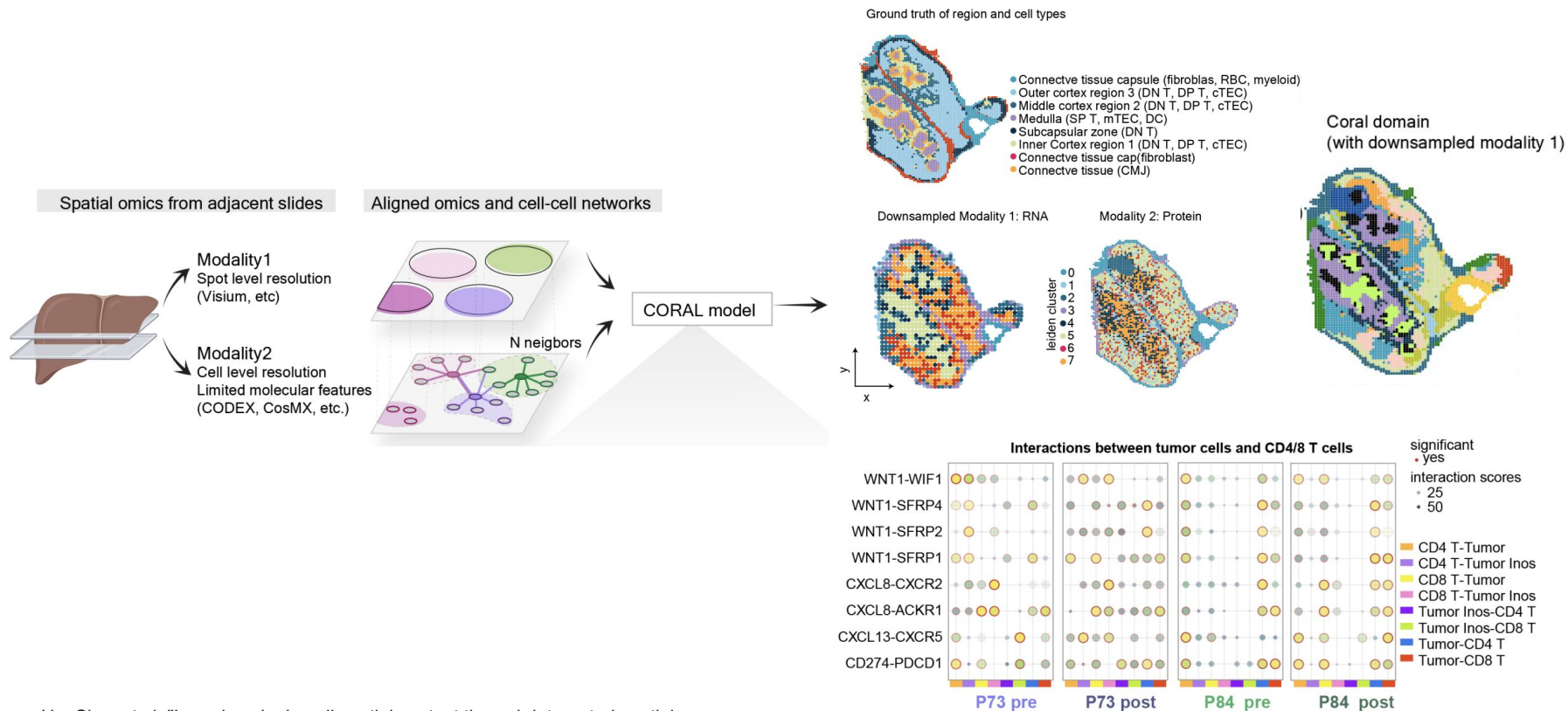
ROSIE: in silico multiplexed immunofluorescence staining



MaxFuse: cross-modal integration of spatial and single-cell omics data

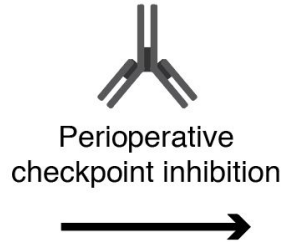


CORAL: integrating spatial modalities with varying resolutions



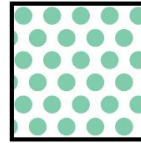
Biological Discoveries enabled by Spatial Multi-omics

HCC Patients
(N=13, 6 responders, 7 nonresponders)

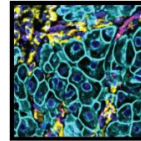


**Pretreatment biopsy +
Posttreatment resection**

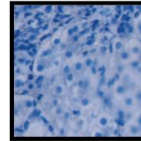
Multimodal Profiling



Spatial
transcriptomics
(10X Visium)

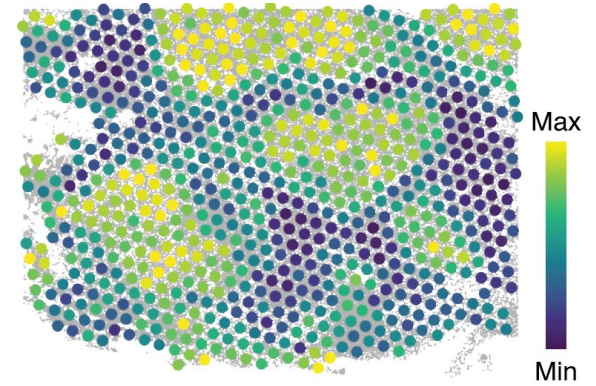


Spatial
proteomics
(Phenocycler)

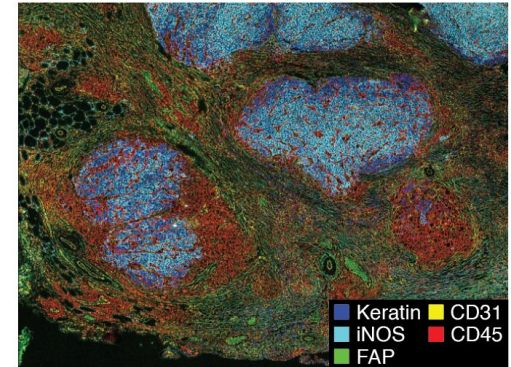


Histopathology
staining

ST (Visium)

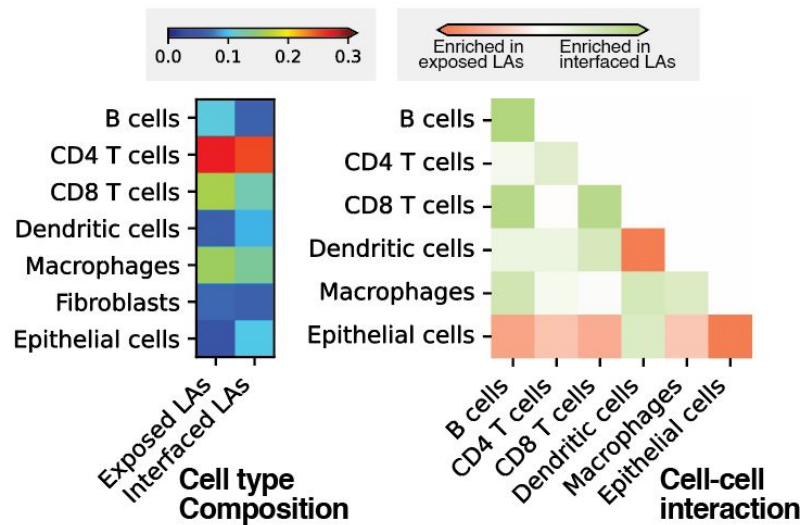


MIF (CODEX)

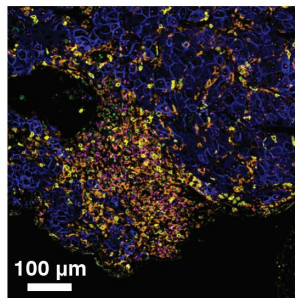


Biological Discoveries enabled by Spatial Multi-omics

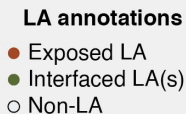
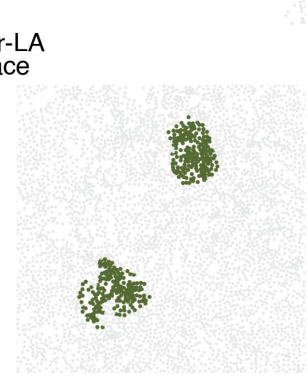
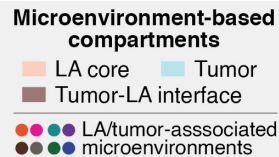
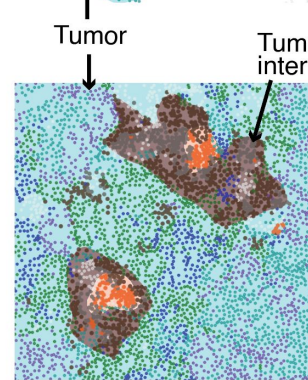
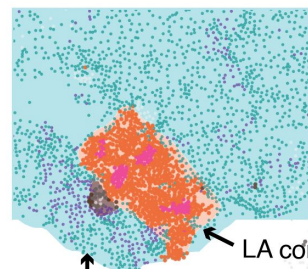
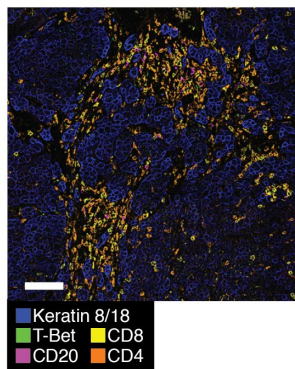
- Discovery of **Exposed LA** (responders) and **Interfaced LA** (non-responders)



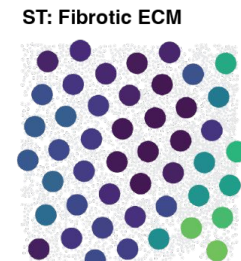
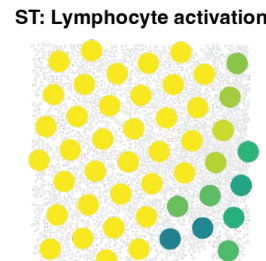
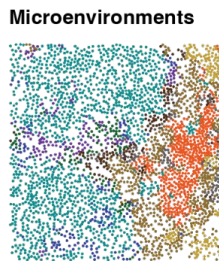
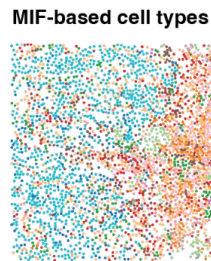
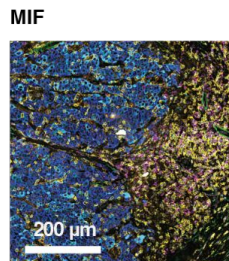
Exposed LA



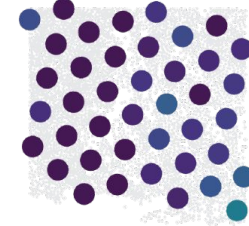
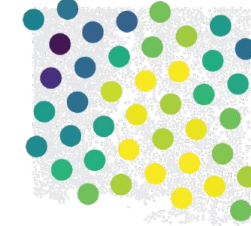
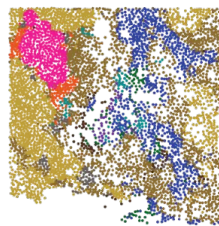
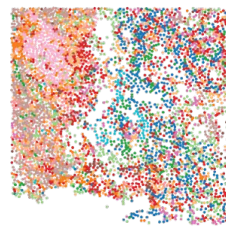
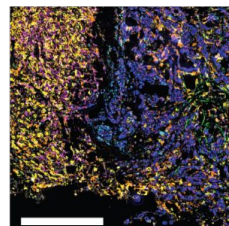
Interfaced LA



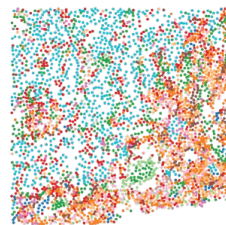
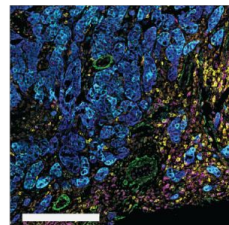
Exposed LA example



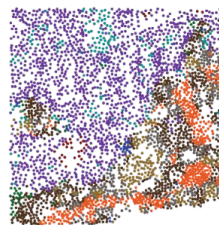
Exposed LA example



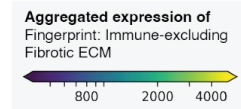
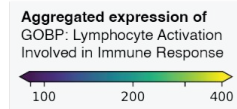
Interfaced LA example



- Epithelial cells (INOS-)
- Epithelial cells (INOS+)
- Endothelial cells
- Fibroblasts
- CD4 T cells
- CD8 T cells
- Macrophages 1
- Macrophages 2
- Neutrophils
- B cells
- Dendritic cells
- Plasma cells
- Unknown/Uncharacterized



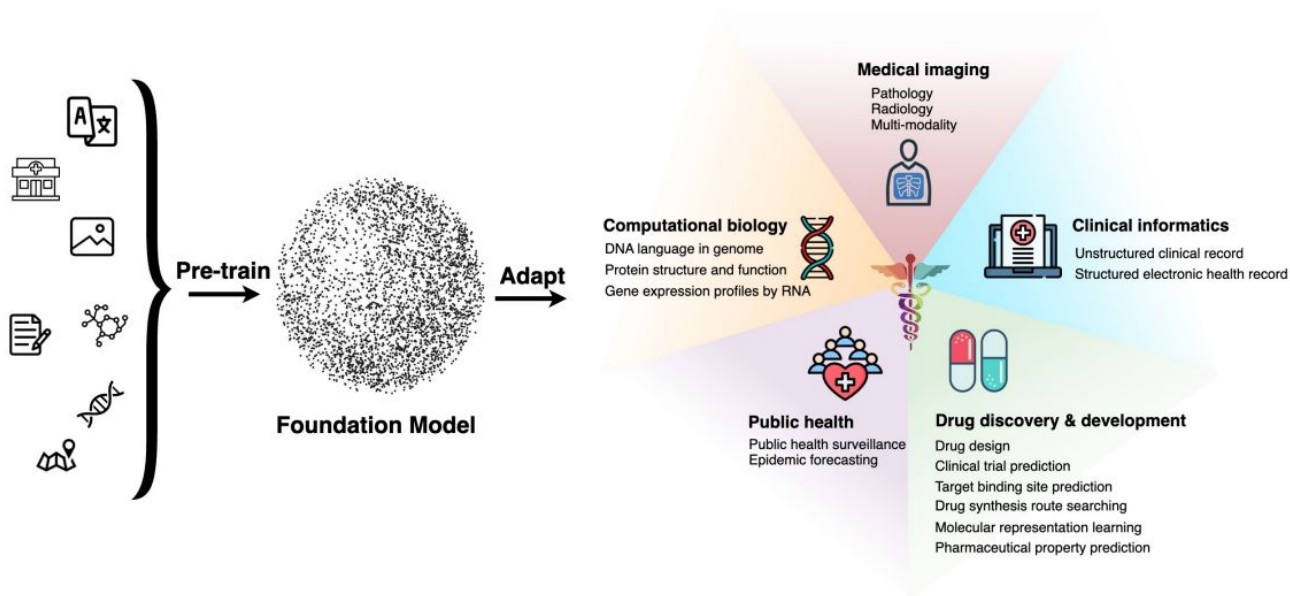
- CD4 T-rich LA
- B-rich LA
- Tumor interface at LA rim
- Fibroblast interface at LA rim
- Tumor INOS-
- Endothelial cell-rich tumor
- Adjacent liver INOS+
- Lymphocyte-rich tumor
- Myeloid cell-rich tumor
- T/Macrophage-rich stroma
- IFN-γ activated stroma



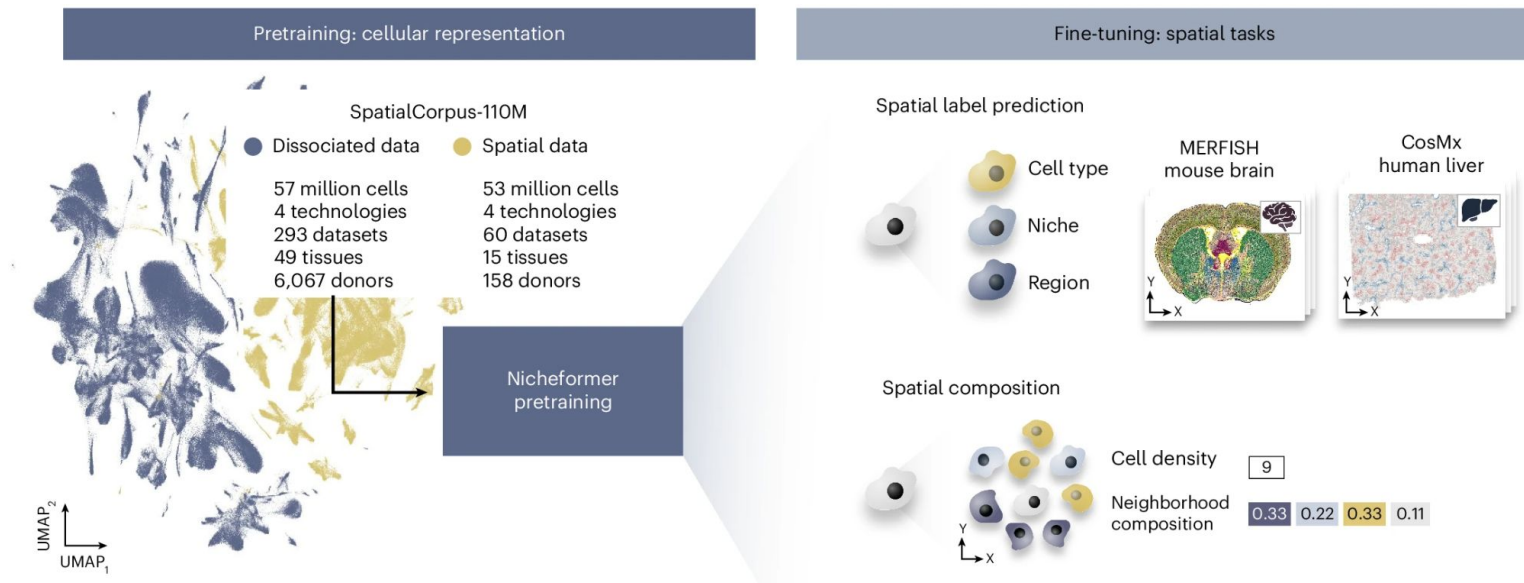
(4th column, ST: Lymphocyte activation): T Cell Activation Involved In Immune Response (GO:0002286)

(5th column, ST: Fibrotic ECM): COL1A1, COL3A1, COL4A1, COL6A1, DCN, ELN, FBLN5, LUM, SPARC, THBS1, etc.

Foundation models for spatial omics



Nicheformer: cell-based foundation model for single-cell and spatial omics

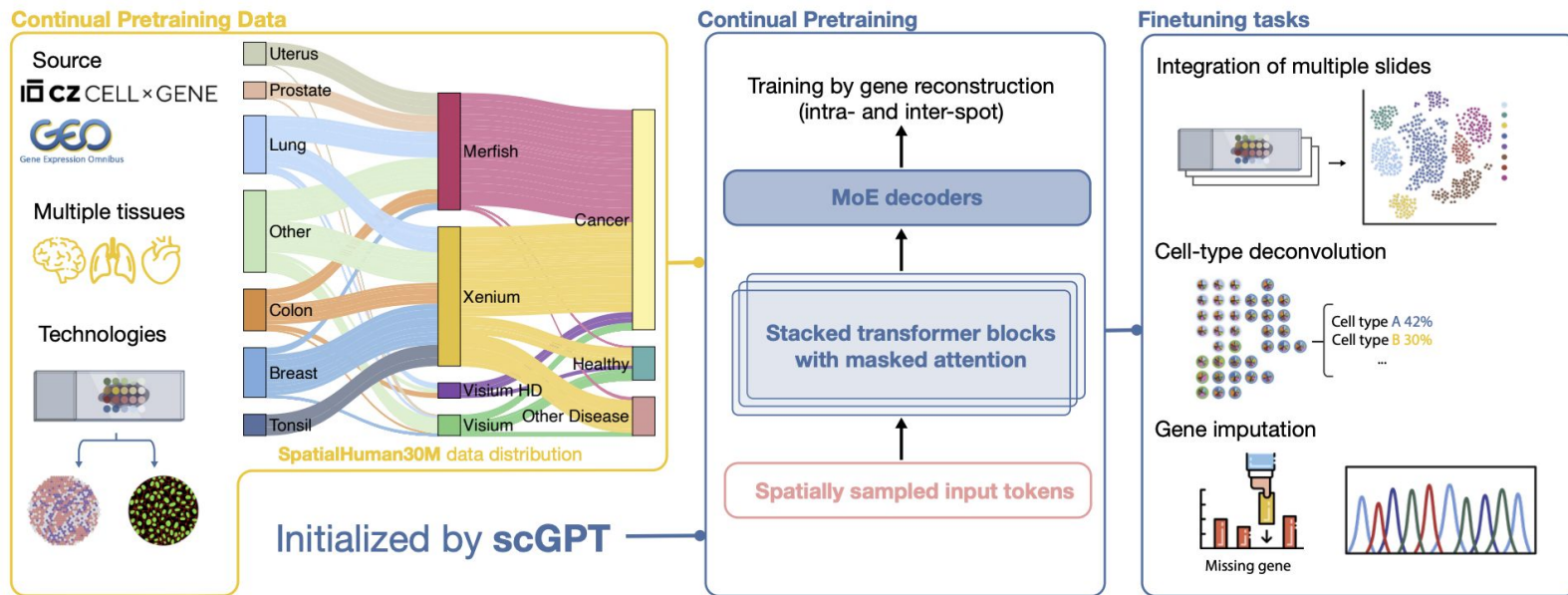


Cells are represented by rank-ordered genes, pre-trained via masked gene reconstruction.

Downstream tasks:

- Cell type / composition / niche label predictions

scGPT-spatial: Continual pretraining of a single-cell foundation model

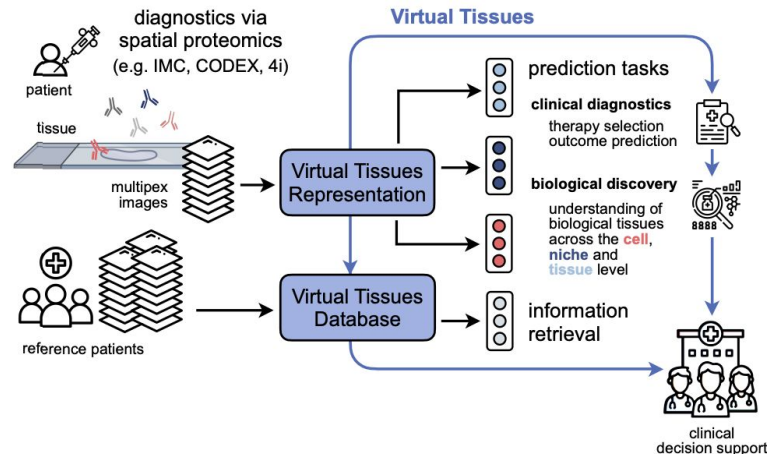
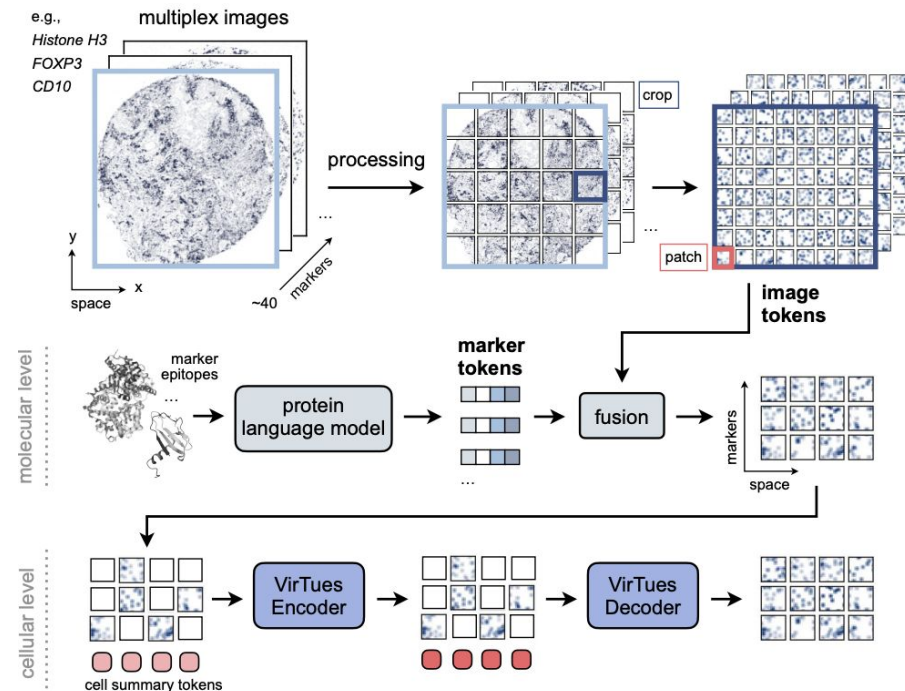


Cells are represented by gene tokens + binned expression, pre-trained via reconstructing self-expression & neighbor-expression

Downstream tasks:

- Missing gene imputation
- Cell type / composition / niche label predictions

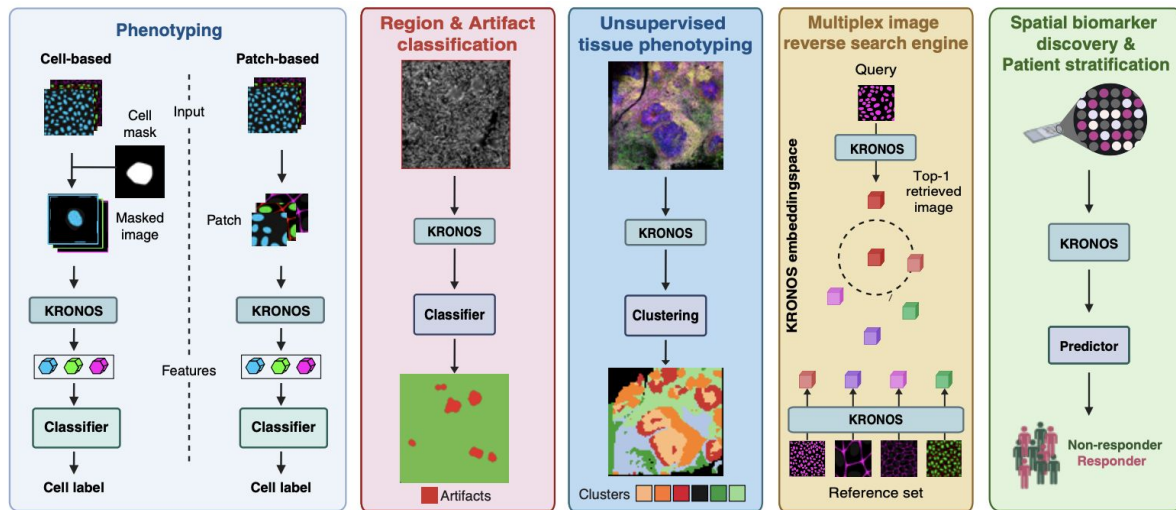
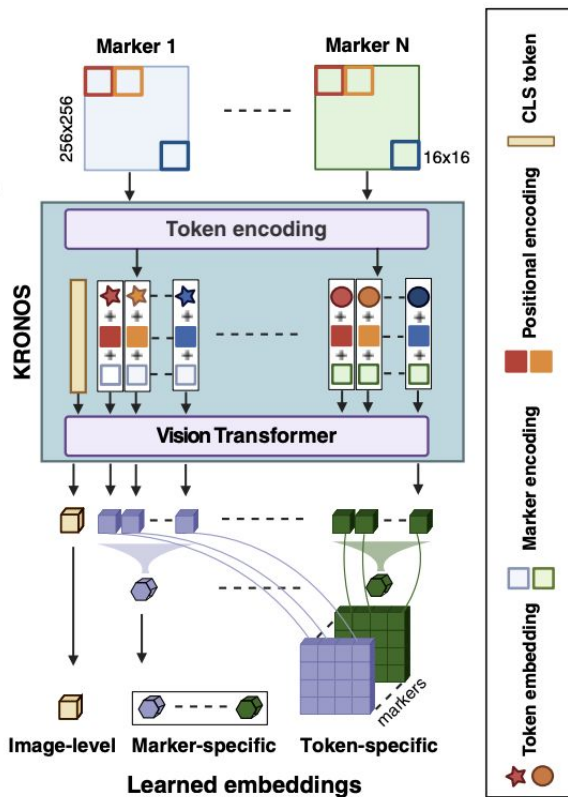
VirTues: Masked autoencoder for MIF



Downstream tasks:

- Cell type / composition / niche label predictions
- Clinical diagnostics
- Zero-shot retrieval of patients

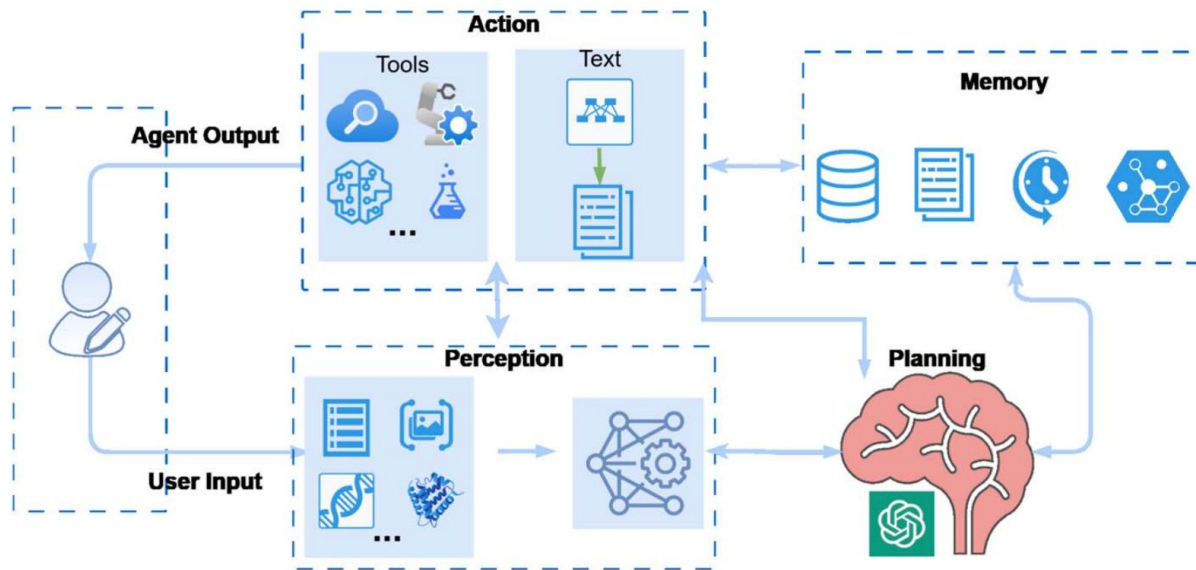
KRONOS: Foundation model (DINOv2) for spatial proteomics



Downstream tasks:

- Cell type / niche label predictions, unsupervised spatial clustering
- Quality control
- Clinical diagnostics / prognostics
- Zero-shot retrieval of cells / patch images

LLM-powered data analytics



Cell typing with Large Language Models



Identify cell types of human prostate cells using the following markers. Identify one cell type for each row. Only provide the cell type name.

CR2, CD24, FAS, CXCR3, CD1c
 KLK3, KRT8, KLK2, MSMB, ACP, KLK1, KLK4
 MMRN1, FLT4, RELN, CCL21, PROX1, LYVE1
 TPSAB1, FCER1A, TPSB2, KIT, CD69, HDC
 ACTA2, MYO1B, ACTA2, ANPEP, DES, MCAM, PDGFRB, CSPG4



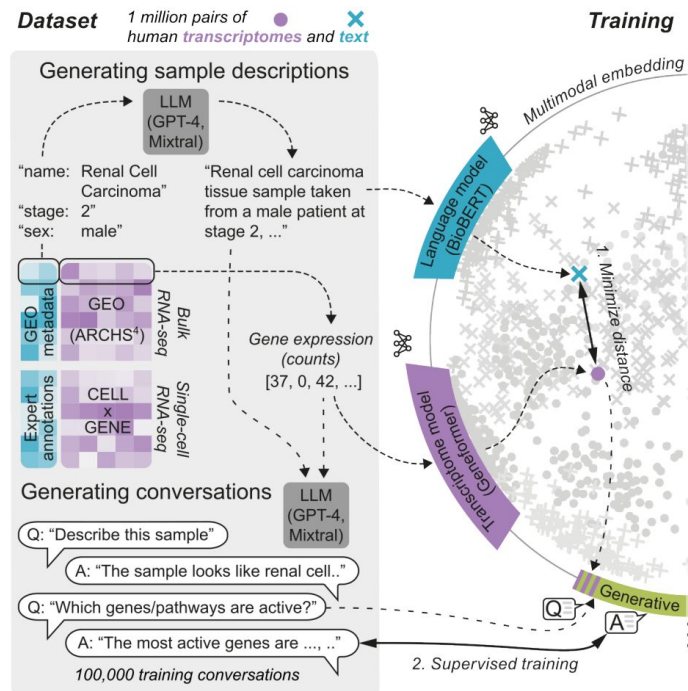
- Dendritic cells
- Luminal epithelial cells
- Lymphatic endothelial cells
- Mast cells
- Pericytes

Cell sentences

Sentence 1: " $g_4 \cdot g_6 \dots g_3$,"
 Sentence 2: " $g_2 \cdot g_{15} \dots g_{27}$,"
 ⋮
 Sentence N: " $g_{13} \cdot g_1 \dots g_3$,"
 Example cell sentence: " $RPL1 \cdot RPL2 \dots RPLX_i$,"

Textual annotation

"CD4 T-cell $\cdot g_4 \cdot g_6 \dots g_3$,"
 "T-cell in melanoma $\cdot g_2 \cdot g_{15} \dots g_{27}$,"
 "Parkinson's disease $\cdot g_{13} \cdot g_1 \dots g_3$,"
 "Liver tissue $\cdot RPL1 \cdot RPL2 \dots RPLX_i$,"



Hou, Wenpin, and Zhicheng Ji. "Assessing GPT-4 for cell type annotation in single-cell RNA-seq analysis." *Nature methods* 21.8 (2024): 1462-1465.

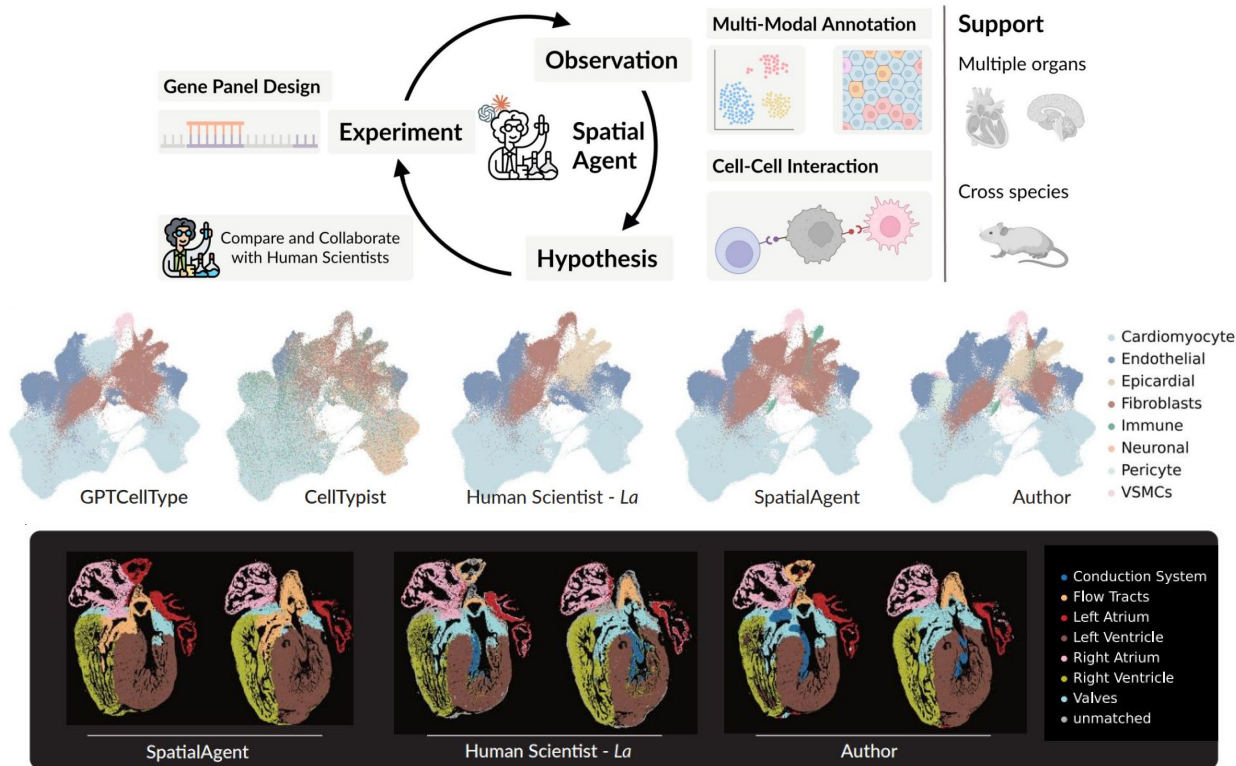
Levine, Daniel, et al. "Cell2Sentence: teaching large language models the language of biology." *BioRxiv* (2024): 2023-09.

Schaefer, Moritz, et al. "Multimodal learning of transcriptomes and text enables interactive single-cell RNA-seq data exploration with natural-language chats." *bioRxiv* (2024): 2024-10.

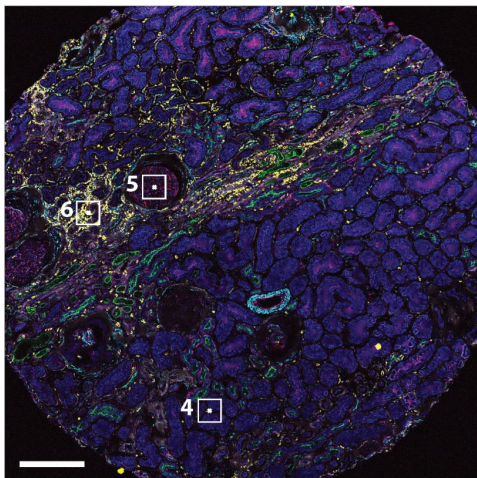
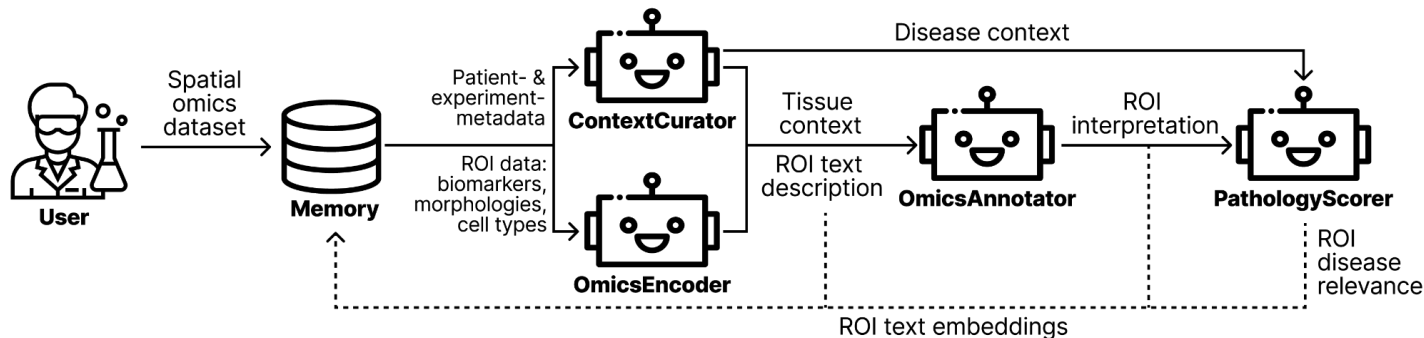
SpatialAgent: autonomous AI agent for spatial biology

Agents enabled automated:

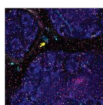
- Biomarker panel design
- Cell type / niche annotation
- Proposing new hypotheses for cell-cell interactions



OmicsNavigator: in-depth annotation and exploration of spatial omics data



4.



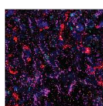
****Proximal Tubules**:**

The dominant cell type, "Tubular cells (CD183+/CD227-)" (87% composition), strongly supports localization to the proximal tubules. ... [More]

Diabetic Kidney Disease (DKD2B), Relevance score: ****40****

Tubular cells dominate the ROI, comprising 87% of the observed cell types. While tubular atrophy and interstitial fibrosis are hallmark features of DKD2B, there is **no direct evidence of these changes in this ROI** based on the observed biomarkers, spatial metrics, or morphology. ... [More]

5.



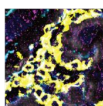
****Glomeruli**:**

The high expression of ****CD196**** and ****CD31****, coupled with the dominance of ****Endothelial cells (CD31+/CD196+)**** (49% of total composition), strongly suggests this region is associated with glomeruli. ... [More]

Diabetic Kidney Disease (DKD2B), Relevance score: 60

****Glomerular Changes**:** The overexpression of ****CD31**** and ****CD196****, alongside densely packed and uniformly oriented endothelial cells, strongly supports the presence of glomerular capillaries. **Moderate expression of **Collagen IV** aligns with glomerular basement membrane thickening, a feature of DKD.** ... [More]

6.



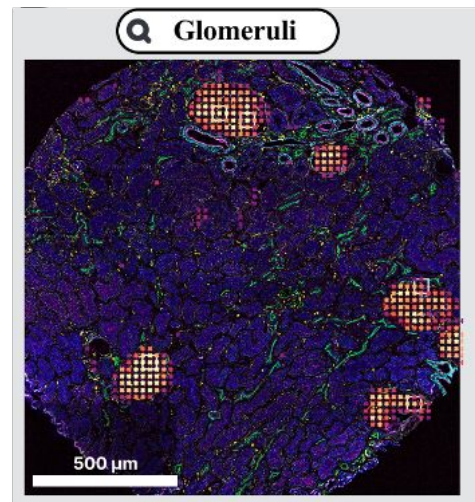
****Inflammatory Domains**:**

Supporting Evidence:

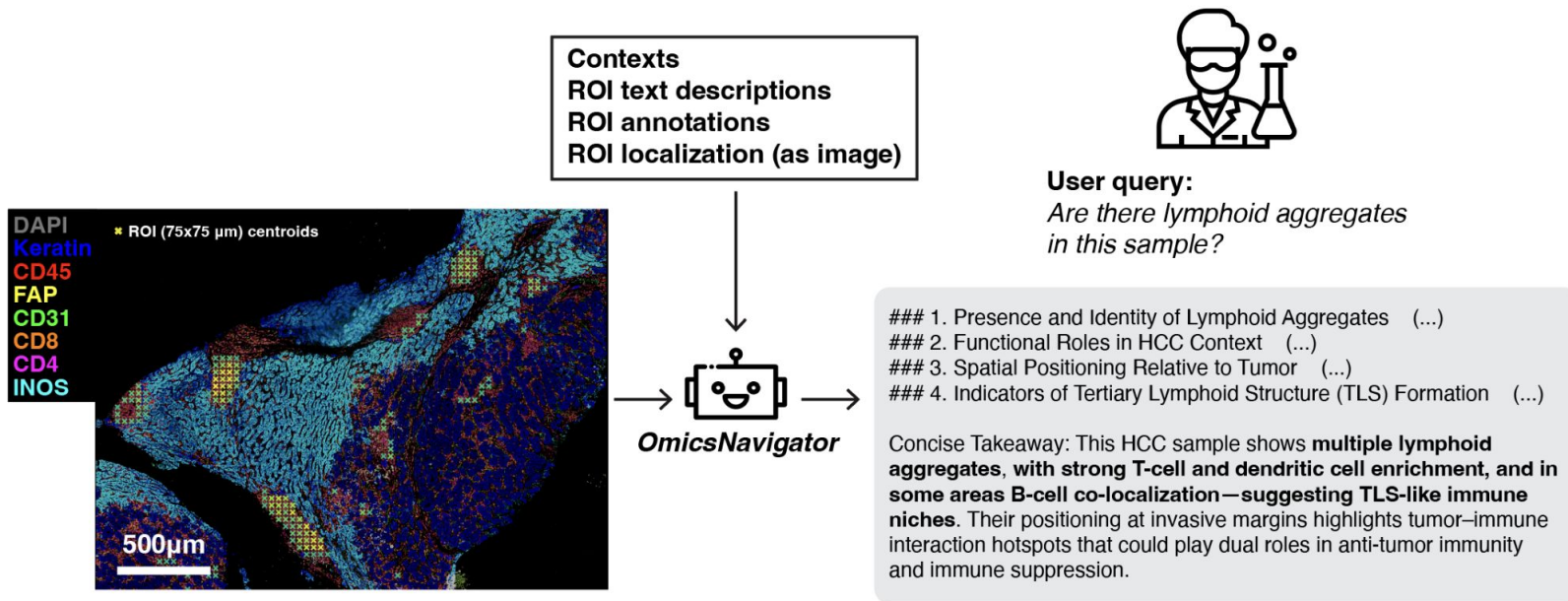
- High expression of immune-related biomarkers: ****CD45**** (z-score = 4.9, percentile = 100%) and ****CD68**** (z-score = 1.7, percentile = 93%), which are markers for immune cells such as hematopoietic cells and macrophages. ... [More]

Diabetic Kidney Disease (DKD2B), Relevance score: 85

****Inflammatory Cell Infiltration**:** The ROI demonstrates significant immune cell infiltration, with ****CD45+ immune cells**** comprising 77% of the population and elevated expression of ****CD68****. **This correlates strongly with the tubulointerstitial changes described in DKD, particularly the presence of monocytes and macrophages contributing to inflammation.** ... [More]



OmicsNavigator: in-depth annotation and exploration of spatial omics data



Looking ahead

- Multi-modal integrative analysis: beyond cell / niche annotations
- 3D extension via deep learning, 3D analytics of spatial biology
- Longitudinal modeling of spatial biology
- In-depth use of foundation models for biological discovery
- Automated insight discovery via agentic systems