

An Introduction to Spatial Data Analysis

Charting Life at Single-Cell Resolution: A Comprehensive Summer School in Spatial Omics
Hands-On Workshop

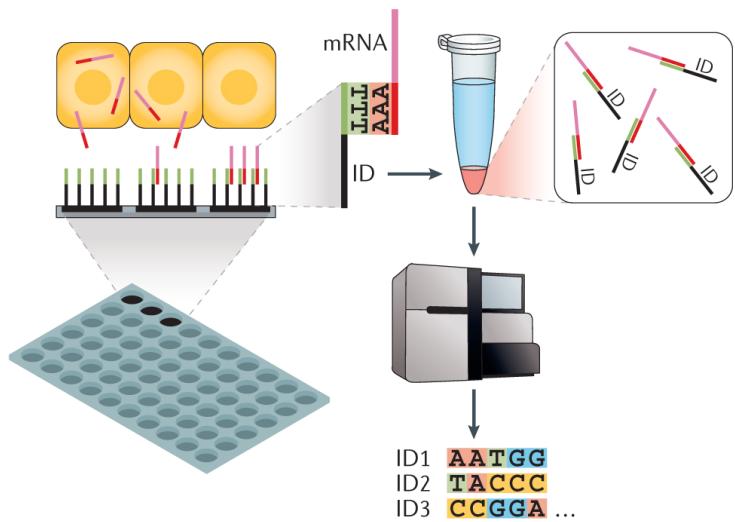
Mattia Toninelli

09/09/25



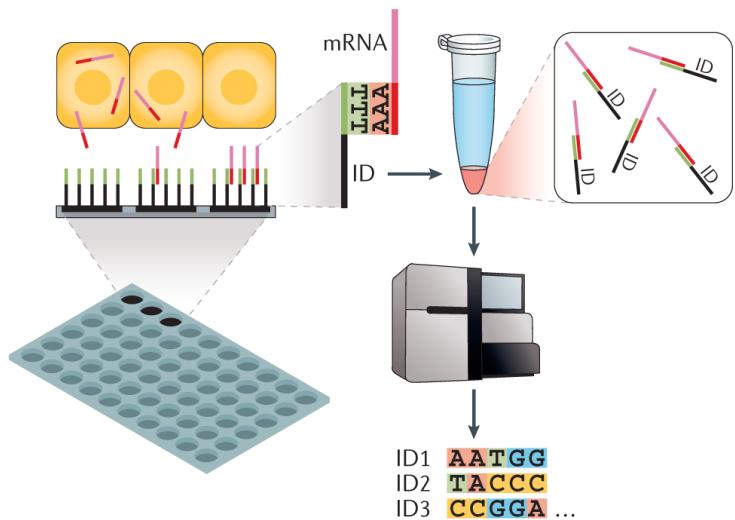
How to look at Spatial Profiling Technologies?

Sequencing

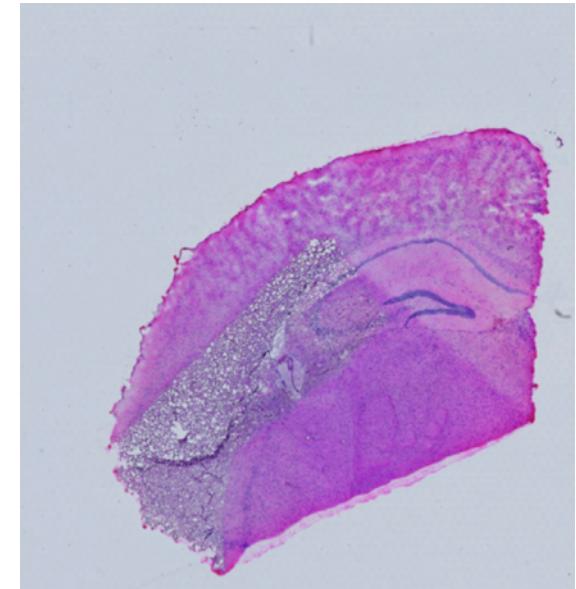


How to look at Spatial Profiling Technologies?

Sequencing

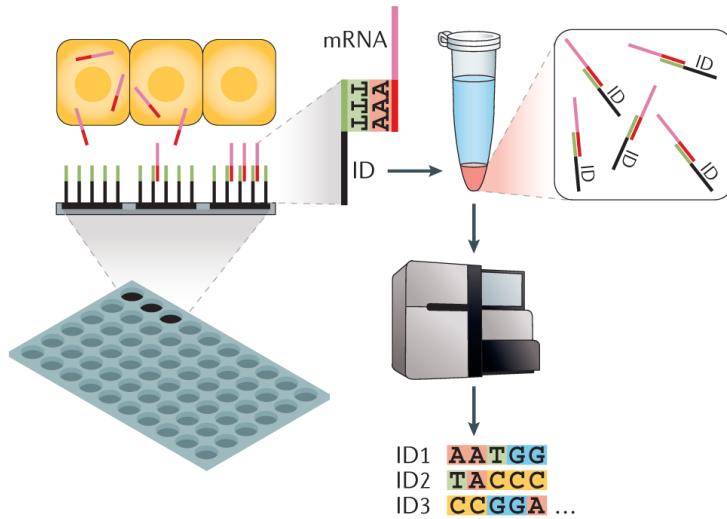


Classic Imaging/Pathology

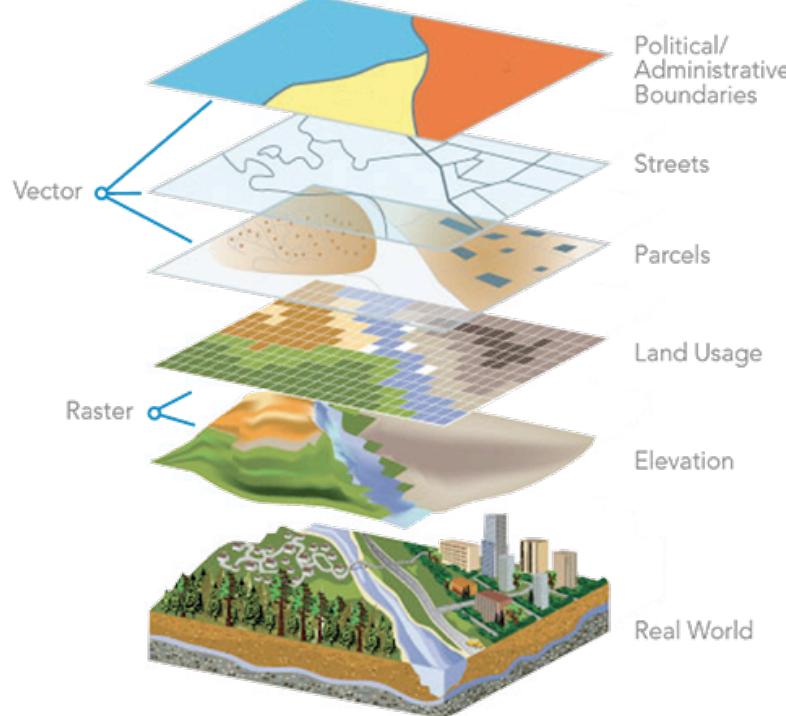


How to look at Spatial Profiling Technologies?

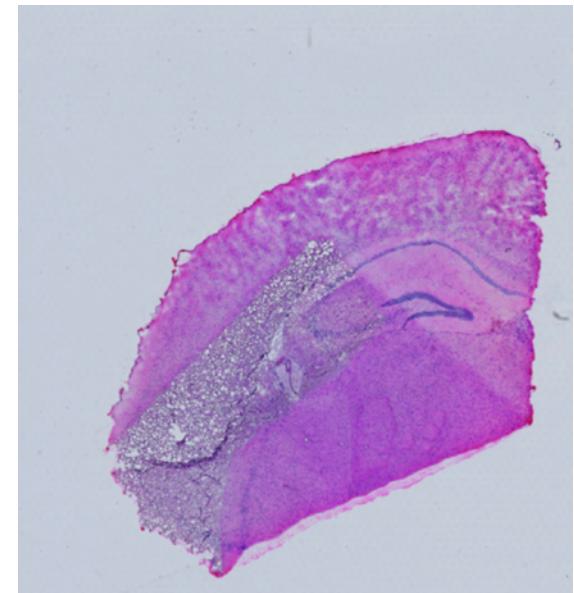
Sequencing



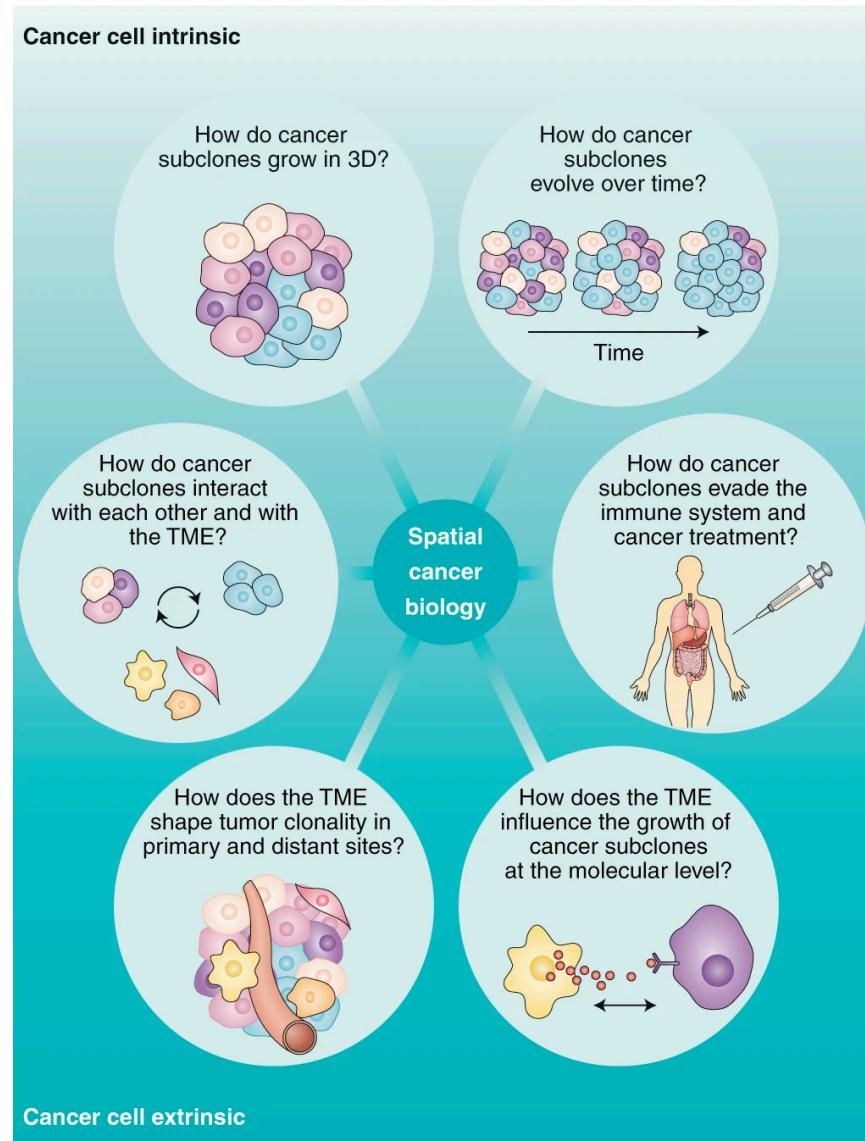
Spatial Data Science



Classic Imaging/Pathology



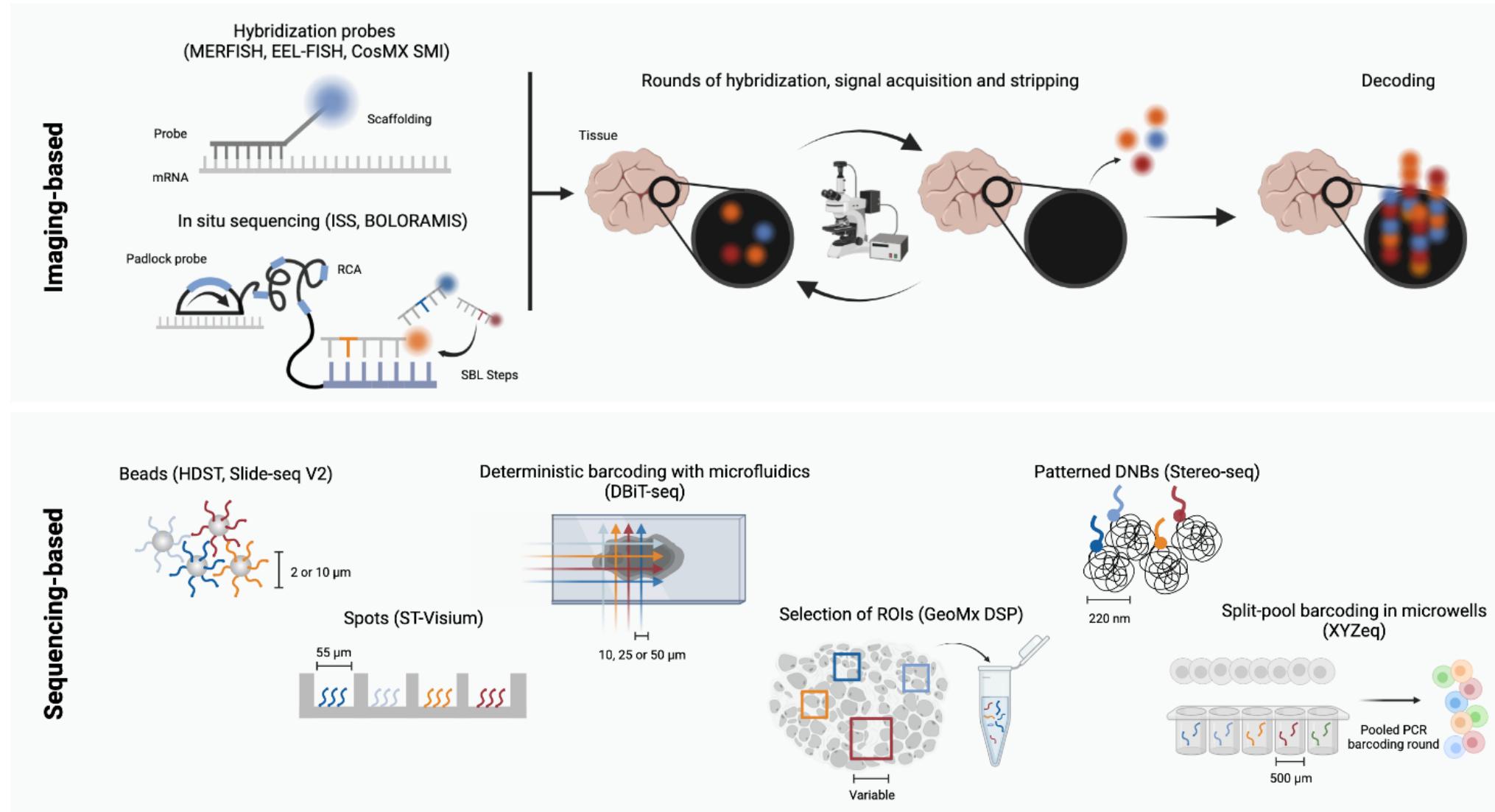
What sorts of questions can be answered?



Examples of biological phenomena occurring at **various spatial scales** come from cancer:

- Tumor cells interact with other microenvironment cells
- Importance of local/broad invasion foci
- Different microenvironmental stimuli (nutrients, hypoxia)

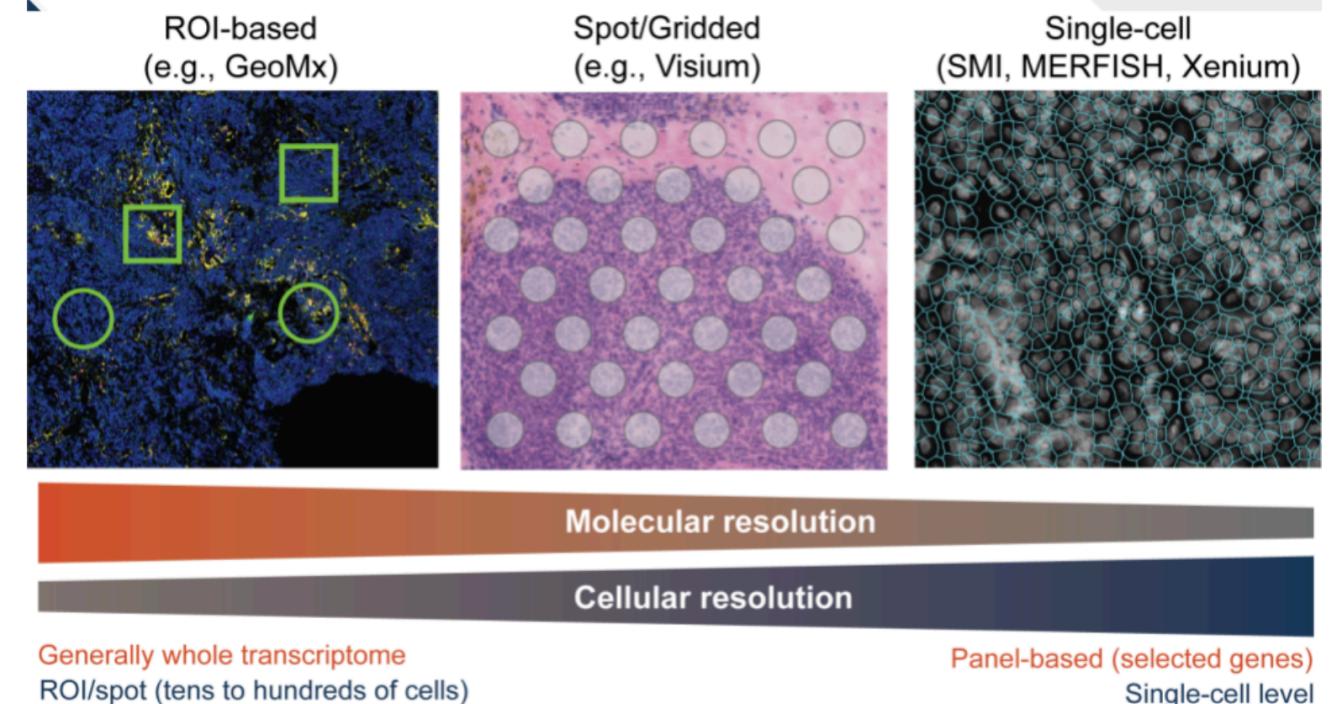
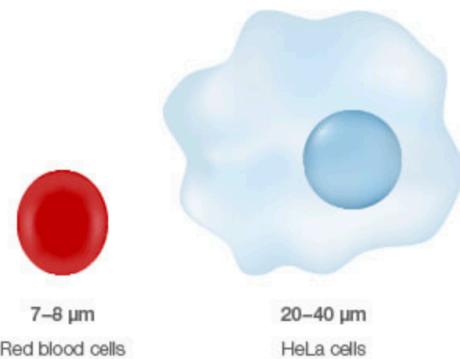
An overview of spatial profiling technologies



Trade-offs in spatial technologies

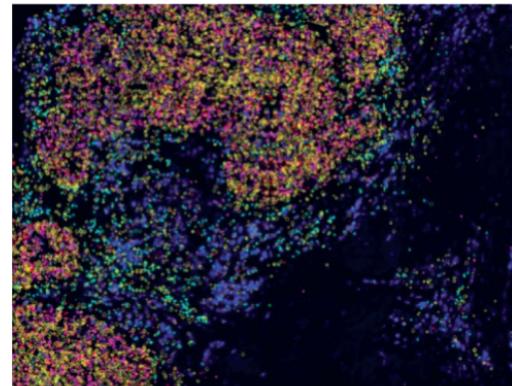
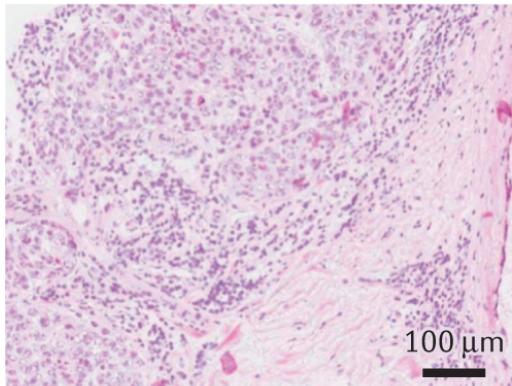
Trade-offs between:

- Transcriptome-wide profiling
- Single-cell (spatial) resolution
- Detection efficiency
- Tissue area
- Usability
- Data analysis



Trade-offs in spatial technologies

b Localized transcripts coloured by specific gene



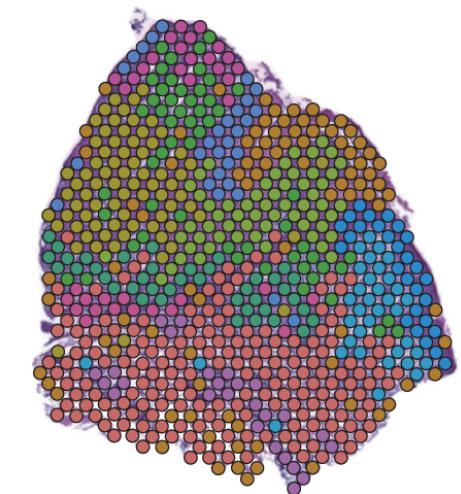
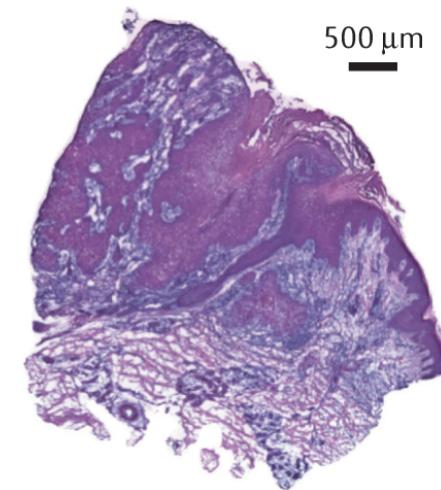
Strengths

- Single-cell resolution
- Greater depth (per transcript)
- Better suited to capture subtype change due to spatial influence

Drawbacks

- Biased (pre-selected gene targets required)
- Lower coverage
- Smaller field of view
- More read-out noise
- Requires more specialized equipment

b Capture spot transcript mixtures deconvolved by dominant cell type



Strengths

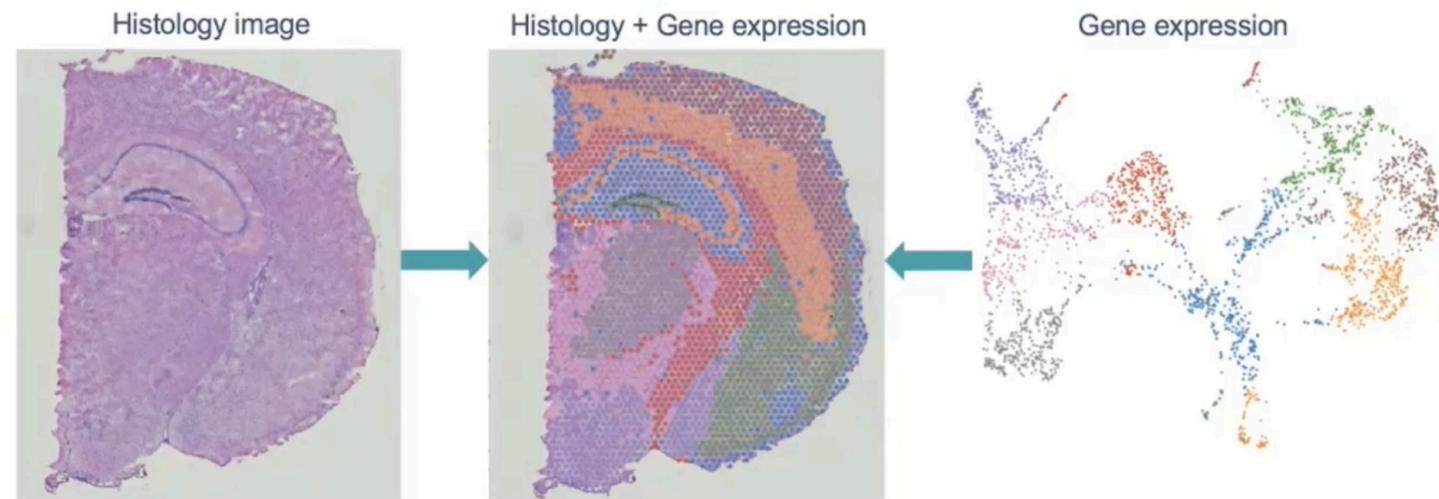
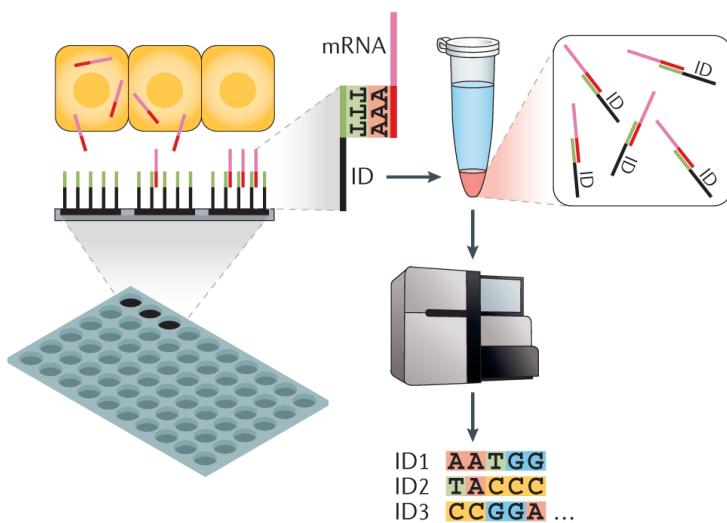
- Unbiased
- Greater coverage
- Greater field of view
- More accessible (typically sequenced using standard NGS machine)

Drawbacks

- Limited to capture spot resolution
- Lower depth (per transcript)

Next Generation Sequencing with spatial dimensions

Sequencing based methods



Preserve spatial location of transcripts by capturing the transcripts from tissue sections on *in situ arrays*

Spatial Gene Expression with Visium

Visualization and analysis of gene expression in tissue sections by spatial transcriptomics

PATRIK L. STÅHL, FREDRIK SALMÉN, SANJA VICKOVIC, ANNA LUNDMARK, JOSÉ FERNÁNDEZ NAVARRO, JENS MAGNUSSON, STEFANIA GIACOMELLO, MICHAELA ASP,

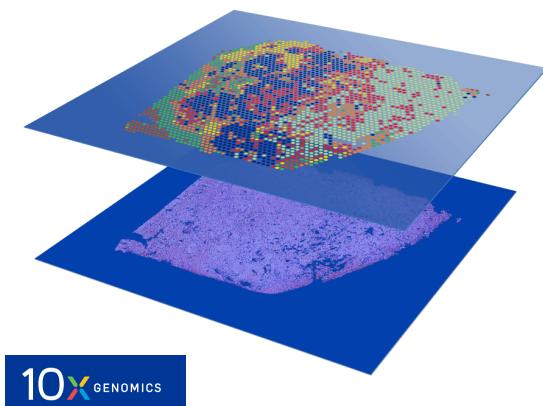
JAKUB O. WESTHOLM, [...] , AND JONAS FRISÉN

+11 authors

[Authors Info & Affiliations](#)

SCIENCE • 1 Jul 2016 • Vol 353, Issue 6294 • pp. 78-82 • DOI: 10.1126/science.aaf2403

Built on the original “Spatial Transcriptomics” concept
(Ståhl et al, 2016)



- Fresh frozen & FFPE
- Whole tissue section profiling
- Whole transcriptome
- Cellular resolution of 1-10 cell per spot

nature
protocols

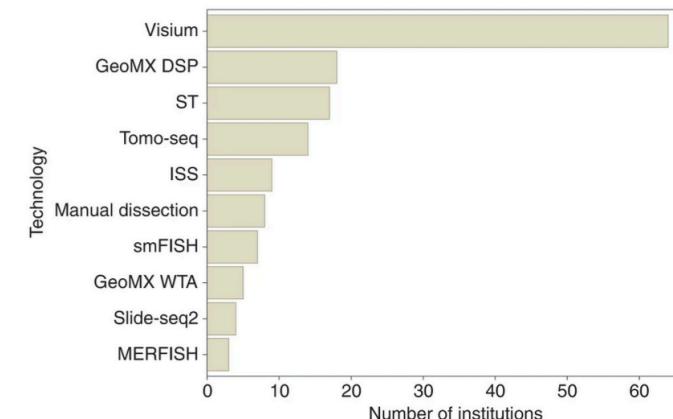
PROTOCOL

<https://doi.org/10.1038/s41596-018-0045-2>

Barcoded solid-phase RNA capture for Spatial Transcriptomics profiling in mammalian tissue sections

Fredrik Salmén^{1,2,5}, Patrik L. Ståhl^{1,5*}, Annelie Mollbrink¹, José Fernández Navarro¹, Sanja Vickovic^{1,3}, Jonas Frisén⁴, Joakim Lundeberg^{1*}

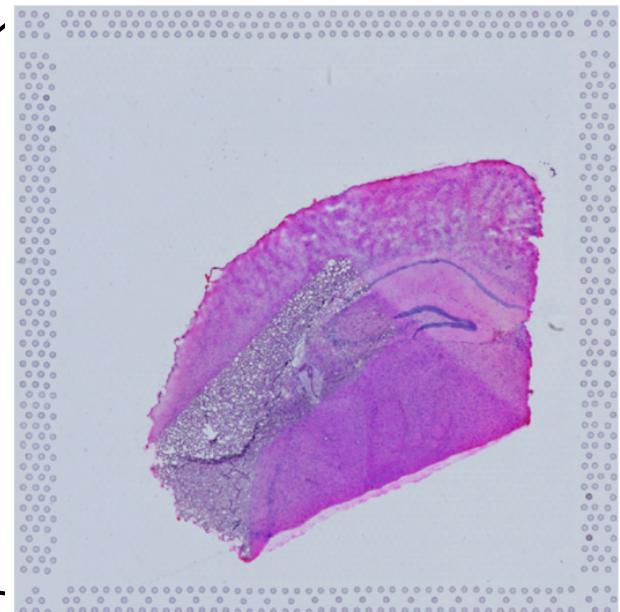
Spatial resolution of gene expression enables gene expression events to be pinpointed to a specific location in biological tissue. Spatially resolved gene expression in tissue sections is traditionally analyzed using immunohistochemistry (IHC) or *in situ* hybridization (ISH). These technologies are invaluable tools for pathologists and molecular biologists; however, their throughput is limited to the analysis of only a few genes at a time. Recent advances in RNA sequencing (RNA-seq) have made it possible to obtain unbiased high-throughput gene expression data in bulk. Spatial Transcriptomics combines the benefits of traditional spatially resolved technologies with the massive throughput of RNA-seq. Here, we present a protocol describing how to apply the Spatial Transcriptomics technology to mammalian tissue. This protocol combines histological staining and spatially resolved RNA-seq data from intact tissue sections. Once suitable tissue-specific conditions have been established, library construction and sequencing can be completed in ~5–6 d. Data processing takes a few hours, with the exact timing dependent on the sequencing depth. Our method requires no special instruments and can be performed in any laboratory with access to a cryostat, microscope and next-generation sequencing.



Analysing whole tissues with morphology

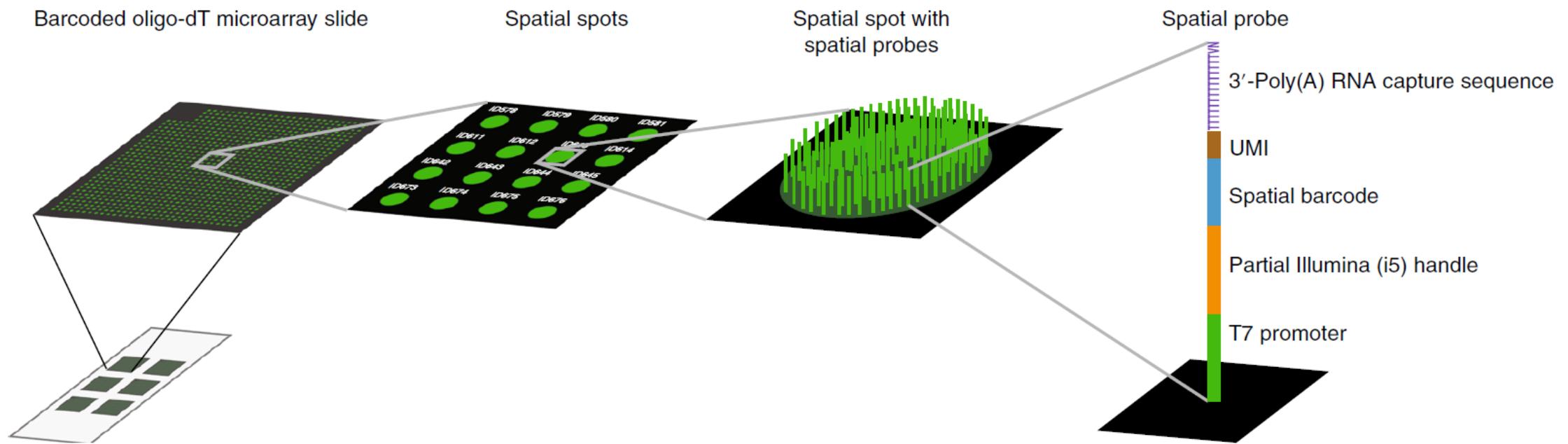
Visium Spatial Gene Expression incorporates unbiased **total mRNA analysis** for intact tissue sections **with morphological context**.

Visium "Slide"



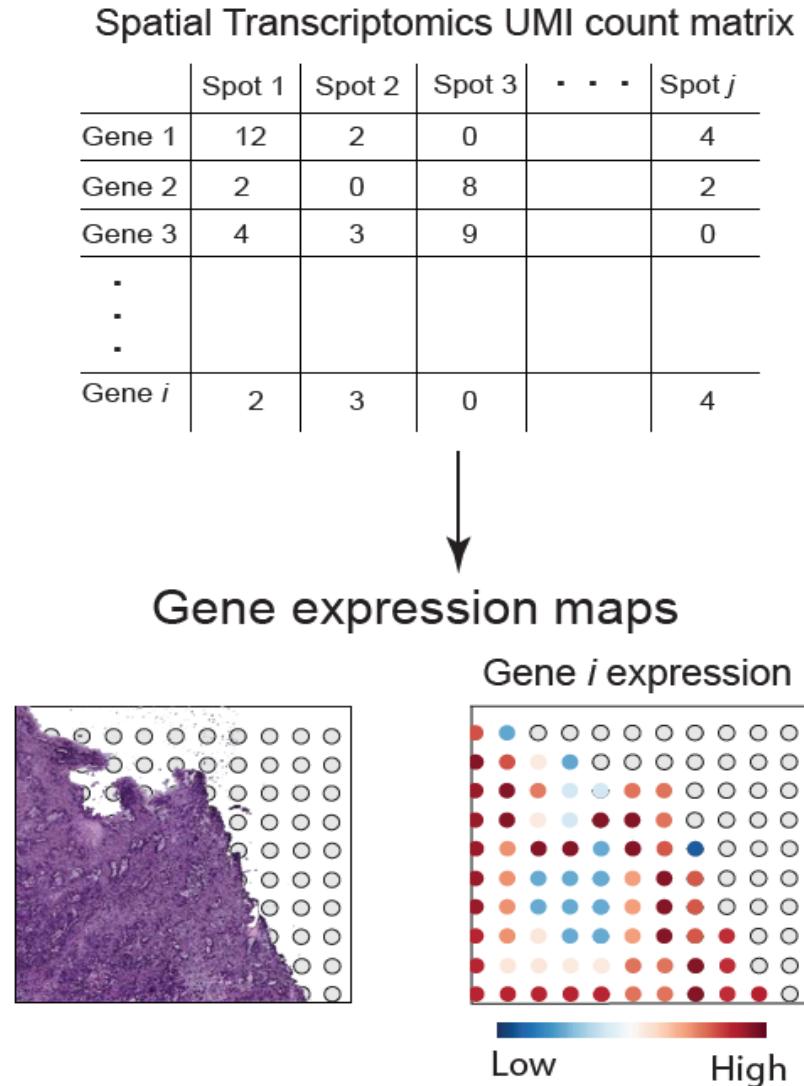
- Four or two Capture Areas of either 6.5 x 6.5 mm or 11 x 11 mm
- Each Capture Area is arrayed with ~5,000 or ~14,000 capture spots

Unbiased capture of polyA tails across whole tissues



Spatial information of spots is retained by each one's **unique barcode sequence**, this is analogous to what happens in the case of single cell RNA-sequencing

Mapping of transcripts over entire tissue sections



- Individual sequencing reads are assigned to their spot of origin
- This is then aligned with a high-resolution image of the tissue
- Visualize the expression of any mRNA and perform analysis with a new spatial dimension
- Spatial data is by definition “multi-modal”

The data flow from acquisition to downstream analyses

1

Transform raw sequencing data into
gene expression data aligned with
microscope images

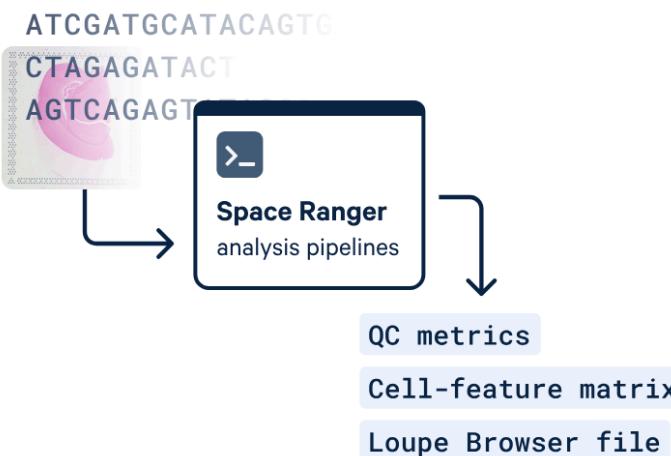
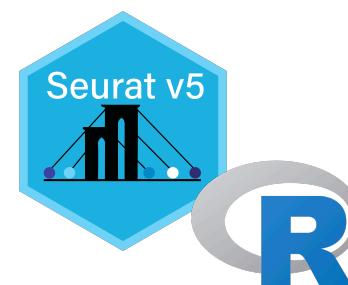
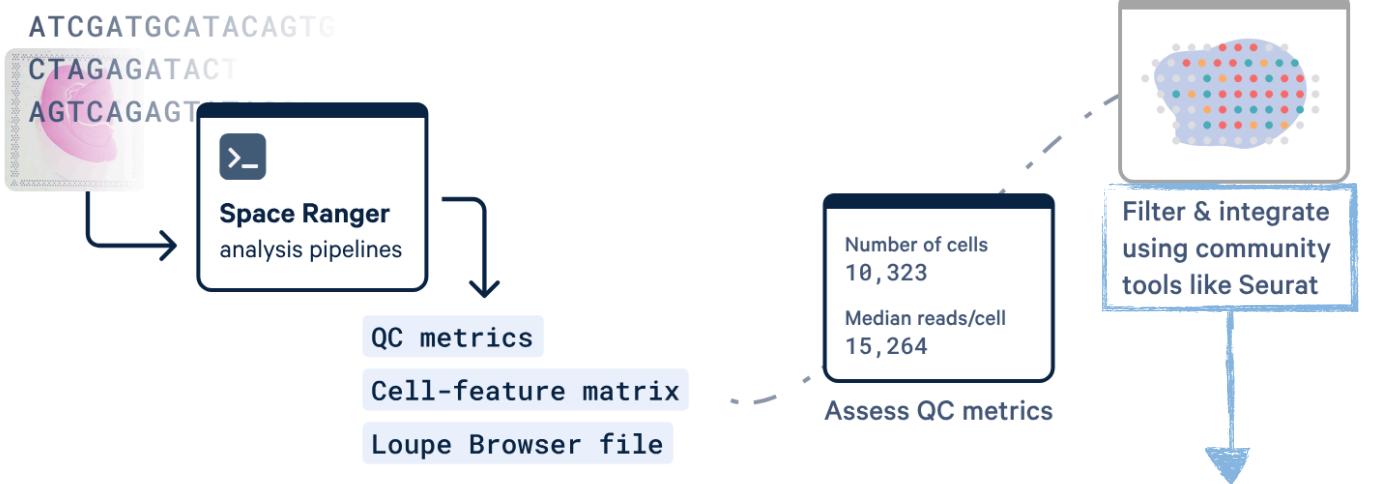


Image data and transcriptional data
are concomitantly analysed

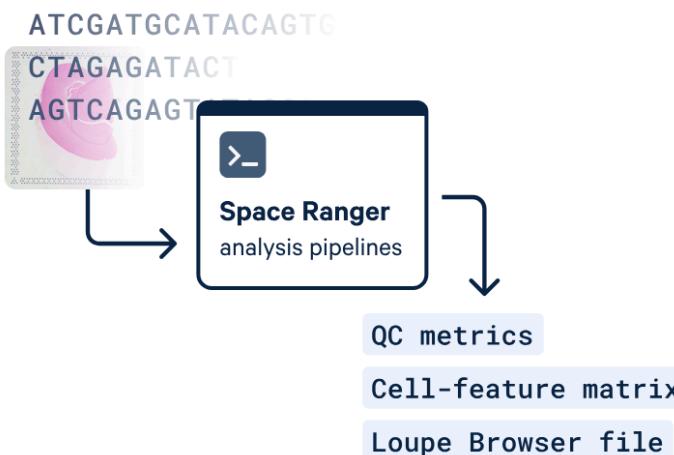
The data flow from acquisition to downstream analyses

- 1 Transform raw sequencing data into gene expression data aligned with microscope images
- 2 Process and refine your data

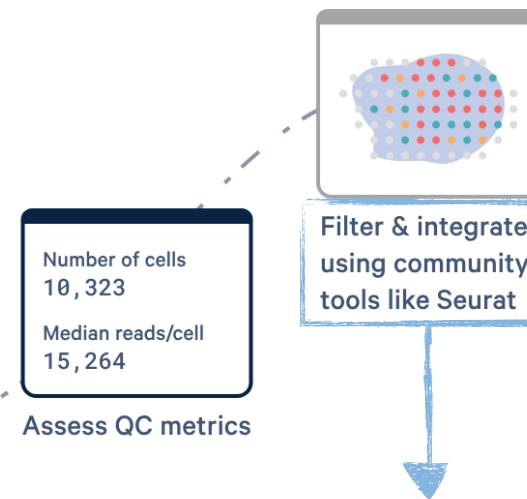


The data flow from acquisition to downstream analyses

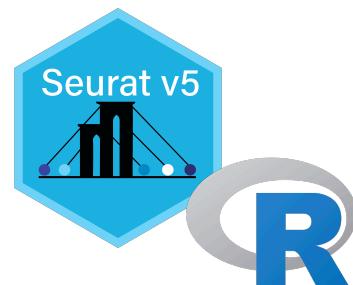
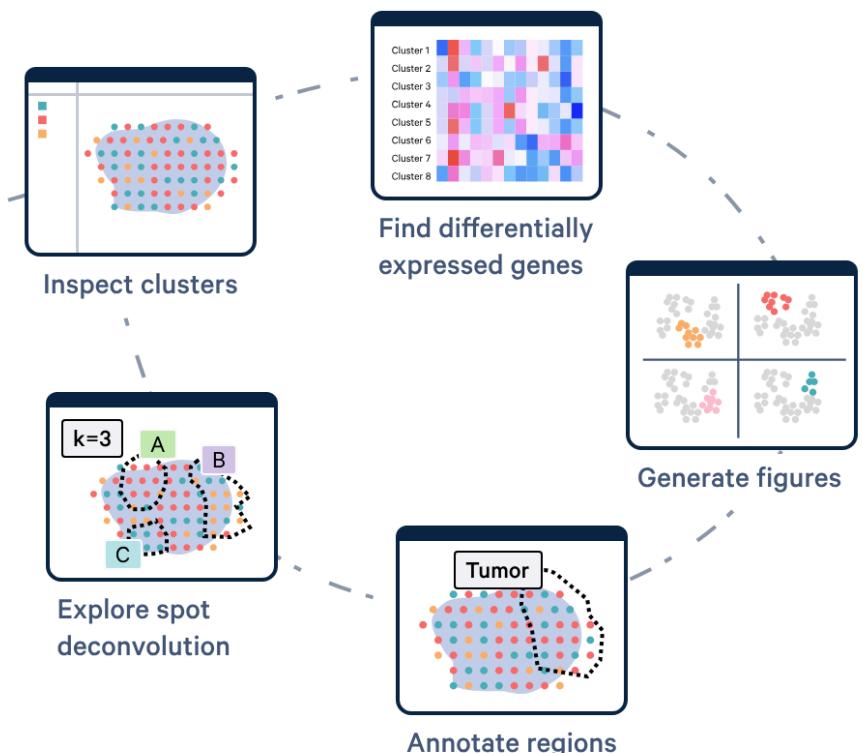
1 Transform raw sequencing data into gene expression data aligned with microscope images



2 Process and refine your data



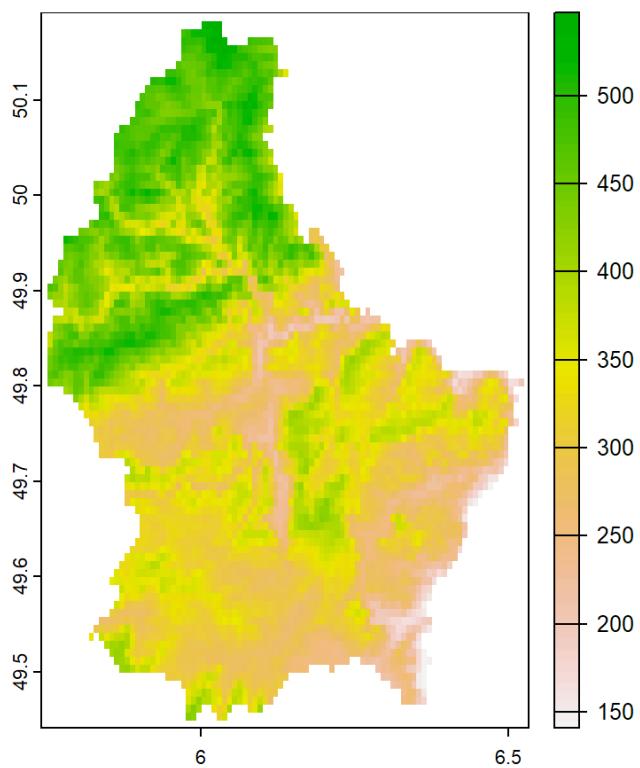
3 Visualize and explore your data



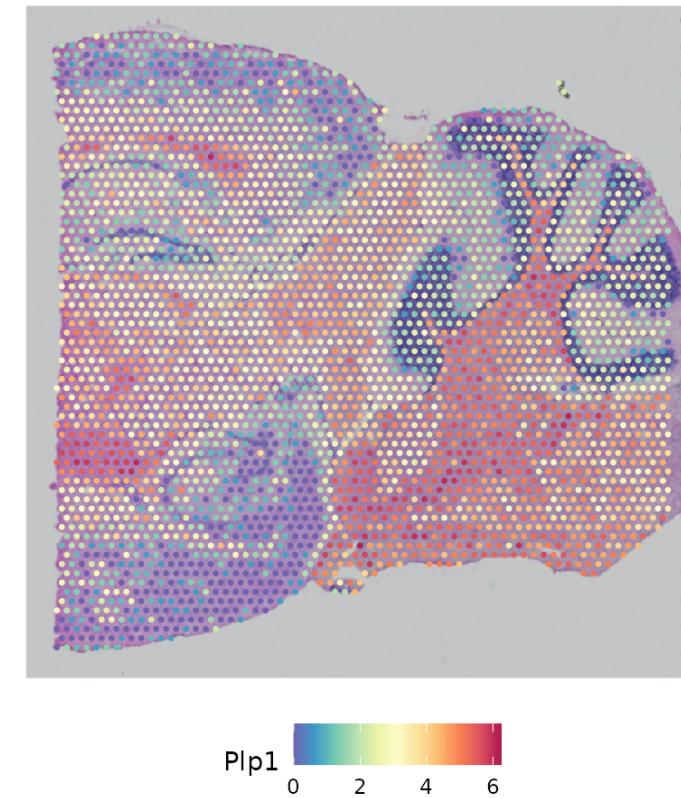
Our workshop's entry point!

Can we just talk about geographical data?

Elevation measure in Luxembourg



Pip1 expression levels in a mouse brain (10X Visium)



The only real difference between the two is a matter of scale, on the left we measure at the level of 100s of Kms whereas on the right we are talking about a few mm at most! At their core, the data structure is the same!



A new view of single-cell data

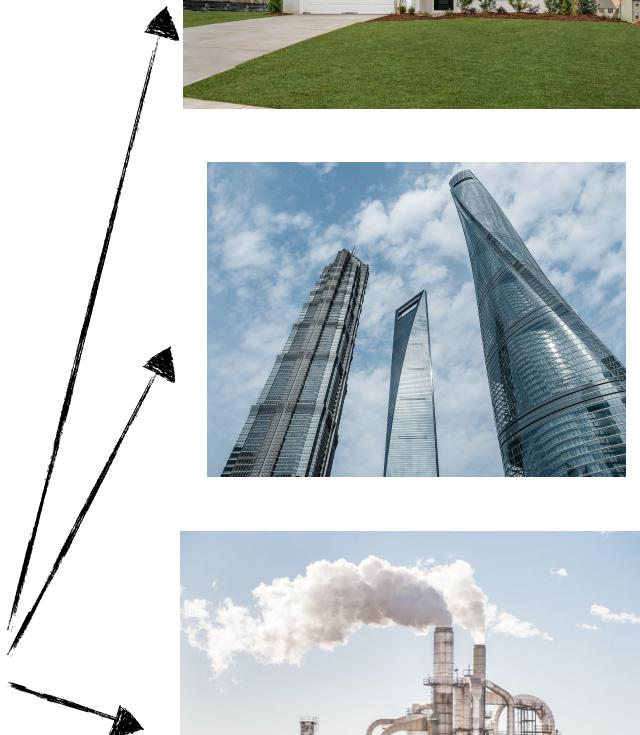


Bulk
RNA-sequencing



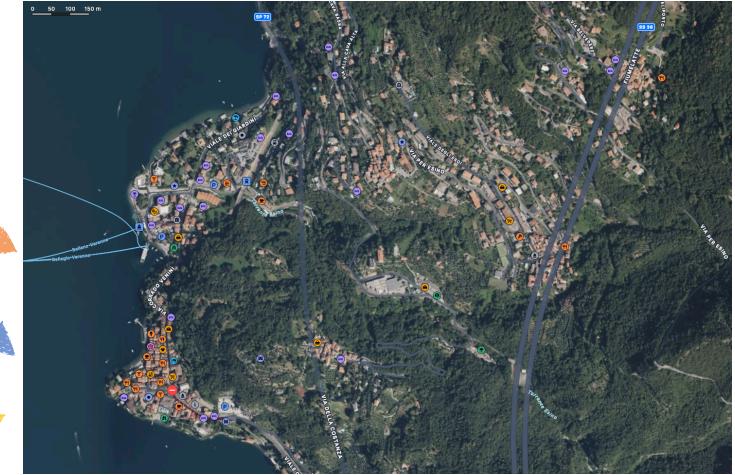
Single-cell
RNA-sequencing

A new view of single-cell data



Spatial Resolution
Functional Organization

A new view of single-cell data

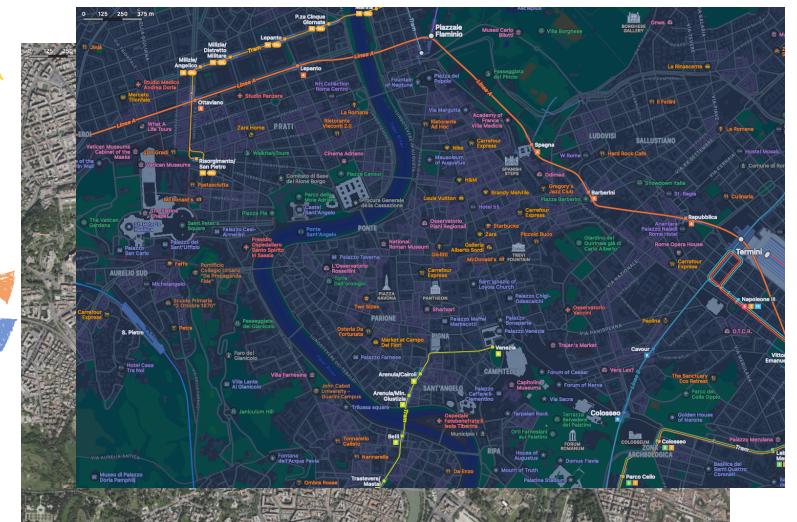
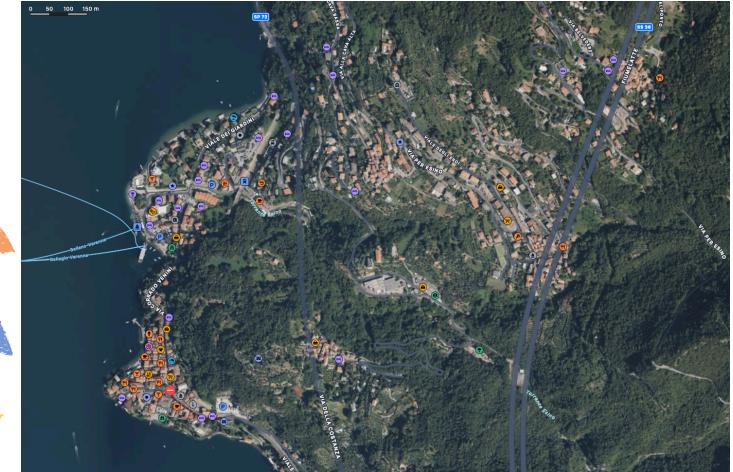


Broader functional units



A new view of single-cell data

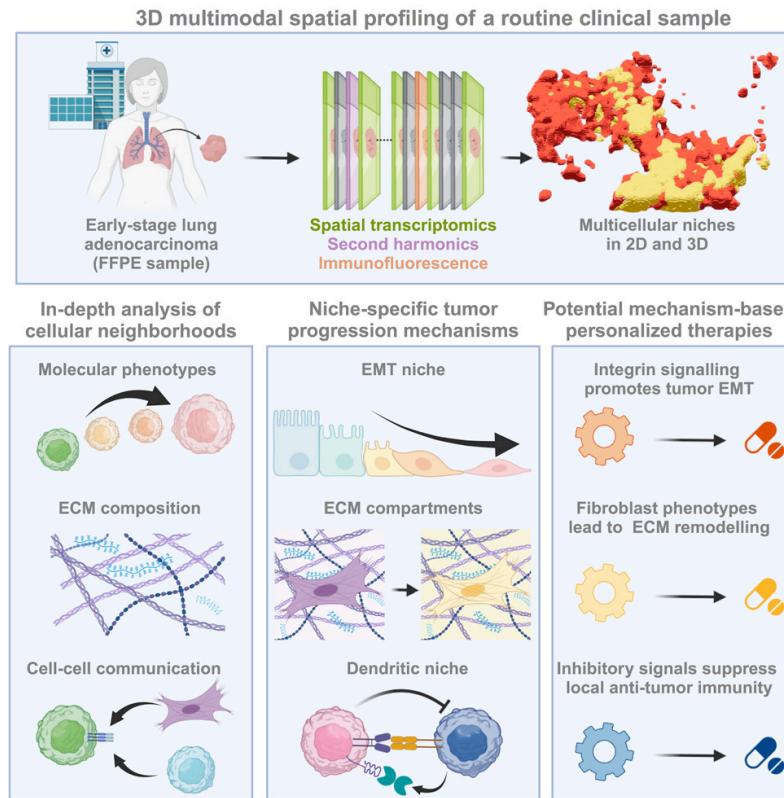
Layering information increases detail and insights



What will future spatial data look like?

Combining spatial transcriptomics and ECM imaging in 3D for mapping cellular interactions in the tumor microenvironment

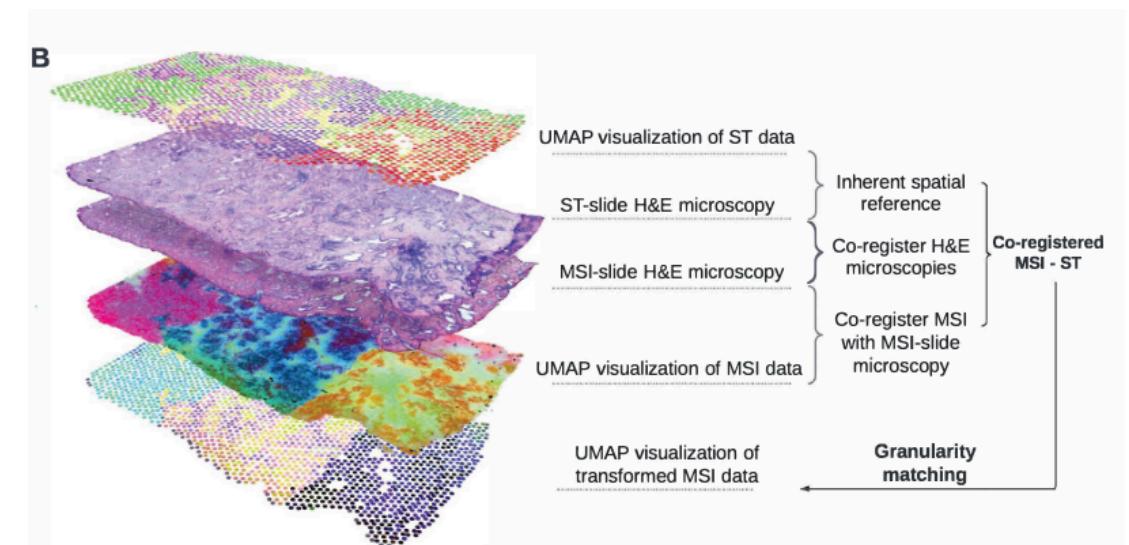
Tancredi Massimo Pentimalli^{1,2} · Simon Schallenberg³ · Daniel León-Periñán¹ · ... · Nikos Karaikos¹ · Frederick Klauschen^{2,7,8}
· Niklaus Rajewsky¹,^{2,14,15,16,17,19}  ... Show more



Integration of Multiple Spatial Omics Modalities Reveals Unique Insights into Molecular Heterogeneity of Prostate Cancer

Wanqui Zhang, Xander Spotbeek, Sebastiaan Vanuytven, Sam Kint, Tassiani Sarreto, Fabio Socciaelli, Katy Vandereyken, Jonas Dehairs, Jakub Idkowiak, David Wouters, Jose Ignacio Alvira Larizgoitia, Gabriele Partel, Alice Ly, Vincent de Laat, Maria José Q Mantas, Thomas Gevaert, Wout Devlieg, Chui Yan Mah, Lisa M Butler, Massimo Loda, Steven Joniau, Bart De Moor, Alejandro Sifrim, Shane R. Ellis, Thierry Voet, Marc Claesen, Nico Verbeeck, Johannes V. Swinnen

doi: <https://doi.org/10.1101/2023.08.28.555056>



Hands-On Overview: Spatial data analysis using Seurat

Data from High-Grade Serous Ovarian Carcinoma (HGSOC)

What data will you be analysing?

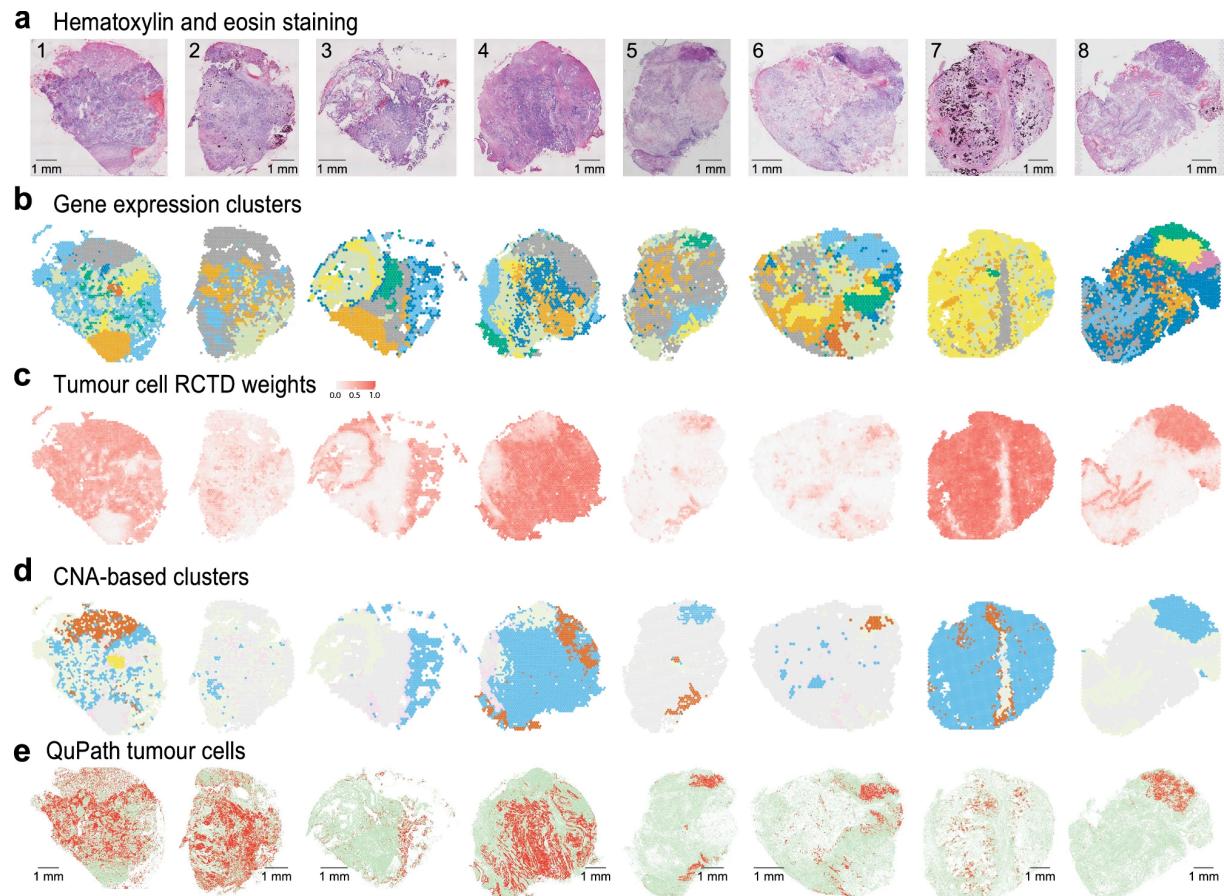
nature communications

Article | [Open access](#) | Published: 03 April 2024

Spatial transcriptomics reveals discrete tumour microenvironments and autocrine loops within ovarian cancer subclones

Elena Denisenko , Leanne de Kock, Adeline Tan, Aaron B. Beasley, Maria Beilin, Matthew E. Jones, Rui Hou, Dáithí Ó Muirí, Sanelia Bilic, G. Raj K. A. Mohan, Stuart Salfinger, Simon Fox, Khaing P. W. Hmon, Yen Yeow, Youngmi Kim, Rhea John, Tami S. Gilderman, Emily Killingbeck, Elin S. Gray, Paul A. Cohen , Yu Yu  & Alistair R. R. Forrest 

- HGSOC is genetically unstable and characterised by the presence of subclones with distinct genotypes
- Intratumoral heterogeneity is linked to recurrence, chemotherapy resistance, and poor prognosis
- Study with primary tumor samples ($n = 8$) profiled with Visium
- Reveal the relationship between tumour subclonal genotypes and infiltration patterns by non-malignant cell types
- Prediction of a link between subclonal genotype differences and differential infiltration patterns



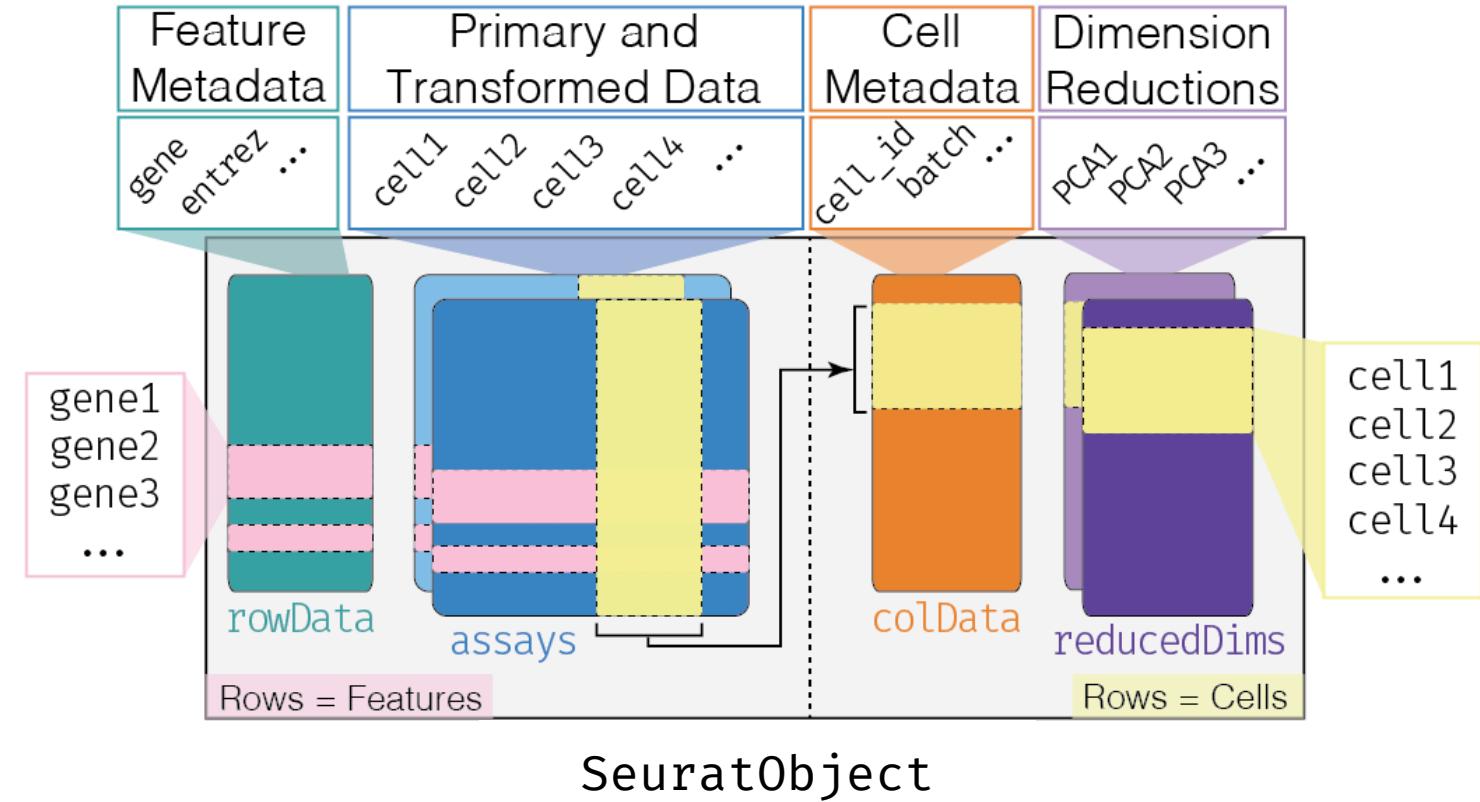
Overall Flow

- Downloading data (1 slide) and creating the ***Seurat*** object
- Quality metric inspection
- Normalization
- Filtering of spots
- Dimensional reduction and clustering
- Spot cell type enrichment with ***clustermole***
- Detecting spatially-variable features and gene modules with ***Giotto***
- Spatially-aware spot clustering
- Cell-cell communication analysis with ***CellChat***

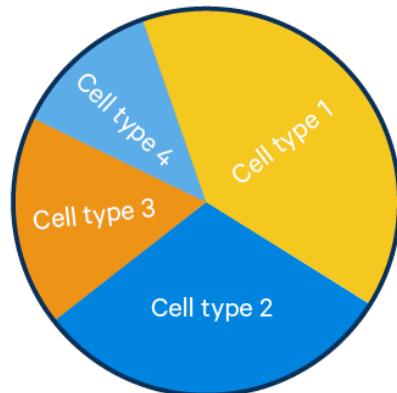
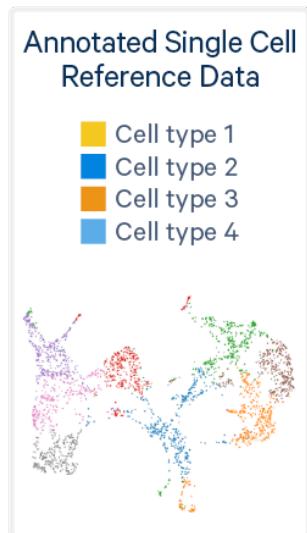
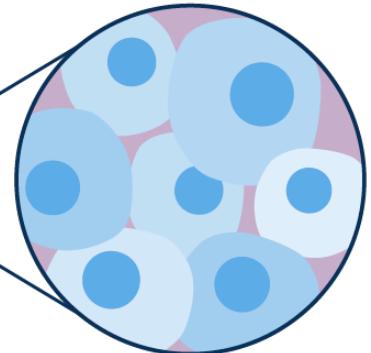
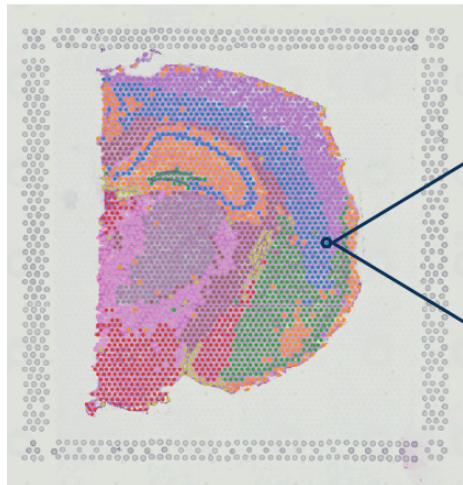
Visium Data in Seurat: Representation in an “object”



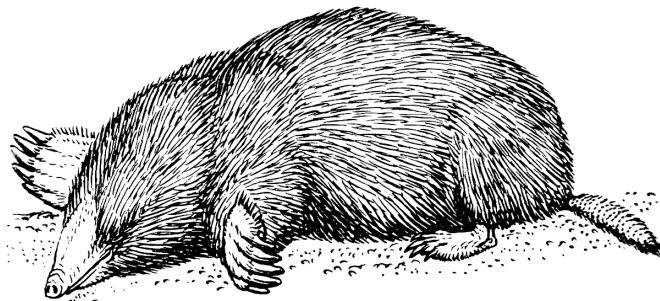
Spot-level expression data
+
Gene information
+
Cell additional information (metadata)
+
Cell locations in reduced dimensionality space (PCA/UMAP/Spatial 2D)



Celltype enrichment estimation in spots using *clustermole*



Clustermole R Package

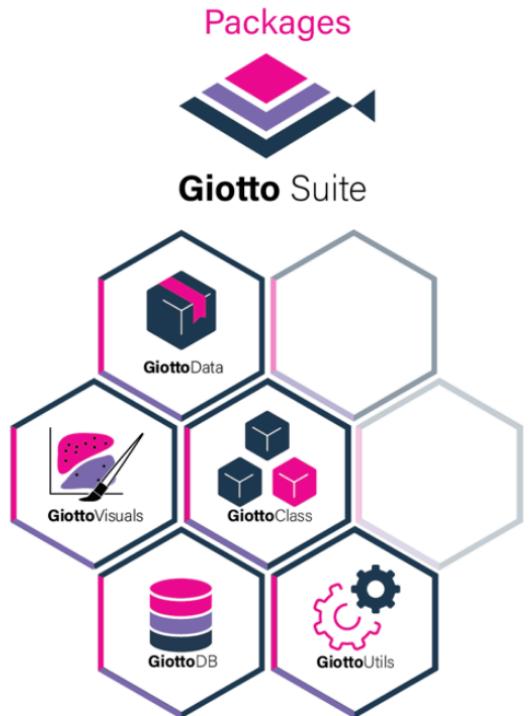


Dig up cell type markers easily!

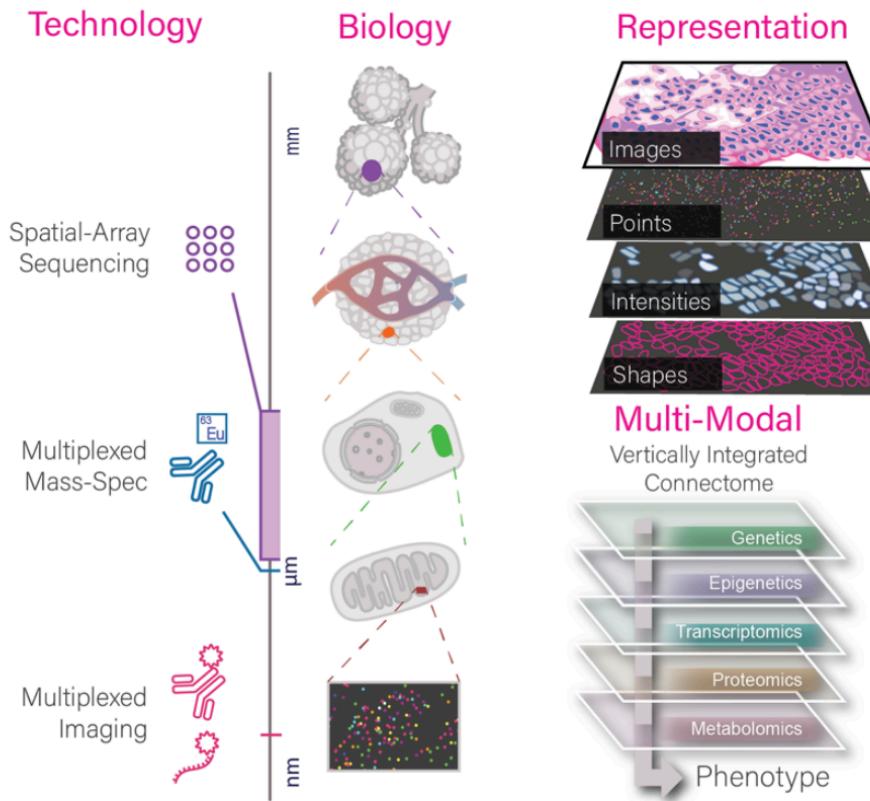
- Visium data is a mini-bulk mix of spots
- Diverse cell types can equally or unequally contribute to signals
- We can either “deconvolve” the spot signal into single cell type contributions...
- ...or check for the “enrichment” of gene signatures from cell types in each spot
- Former requires scRNA-seq, the latter is knowledge based but more limited

The *Giotto* suite for spatial data analysis

A

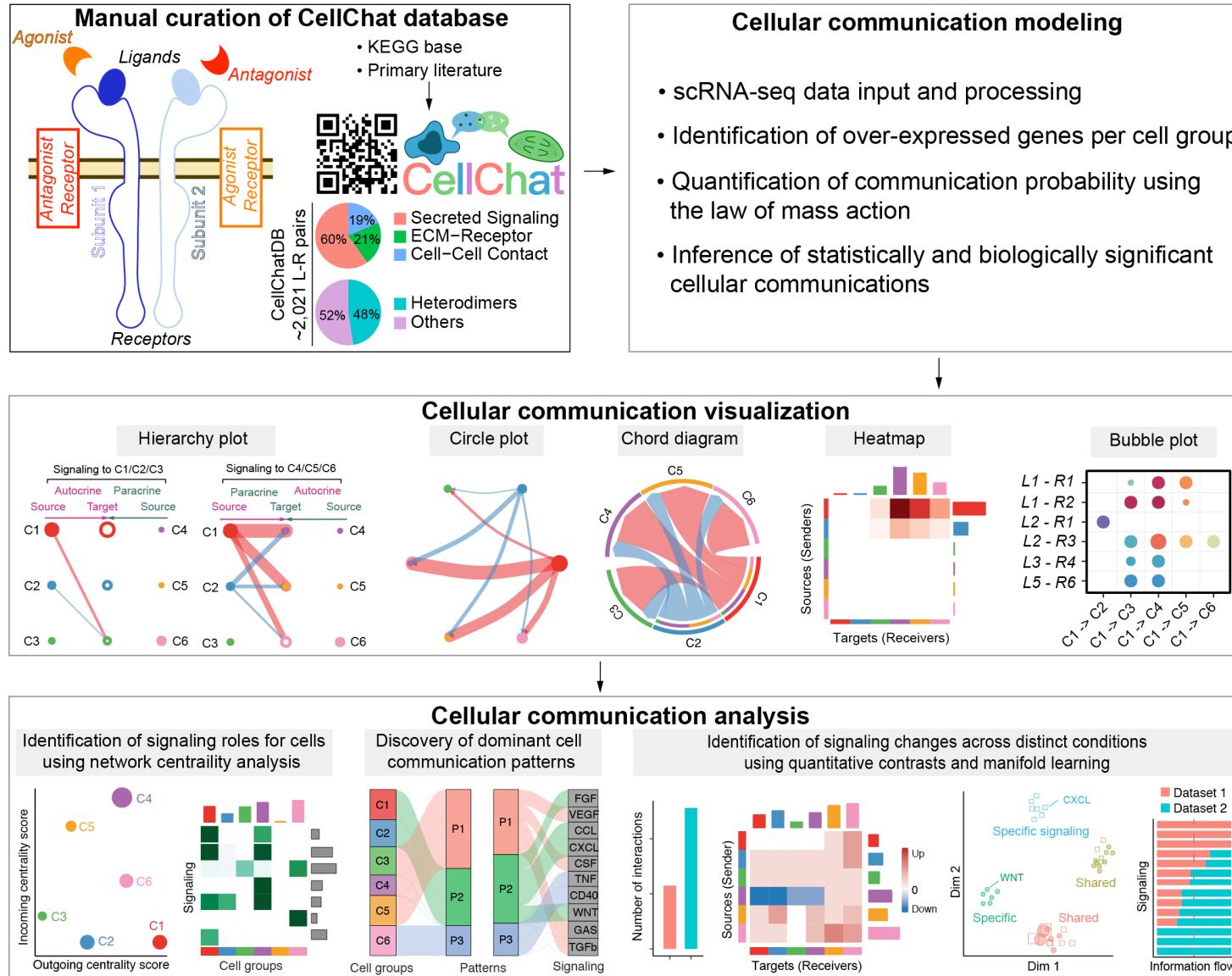


B



- Comprehensive toolkit to analyse spatial data
- Spatially-variable genes identification
- Spatial domains (niches) identification
- Image processing and registration
- Facilitates cross-modality integration

Ligand-receptor analysis with CellChat



- Spatially-weighted ligand-receptor (LR) activity
- Starts from a LR database with > 1,000 annotated interactions
- Scores interactions at the single-spot level and aggregates them across clusters
- All-in-one analysis and visualization package

It's up to you now!