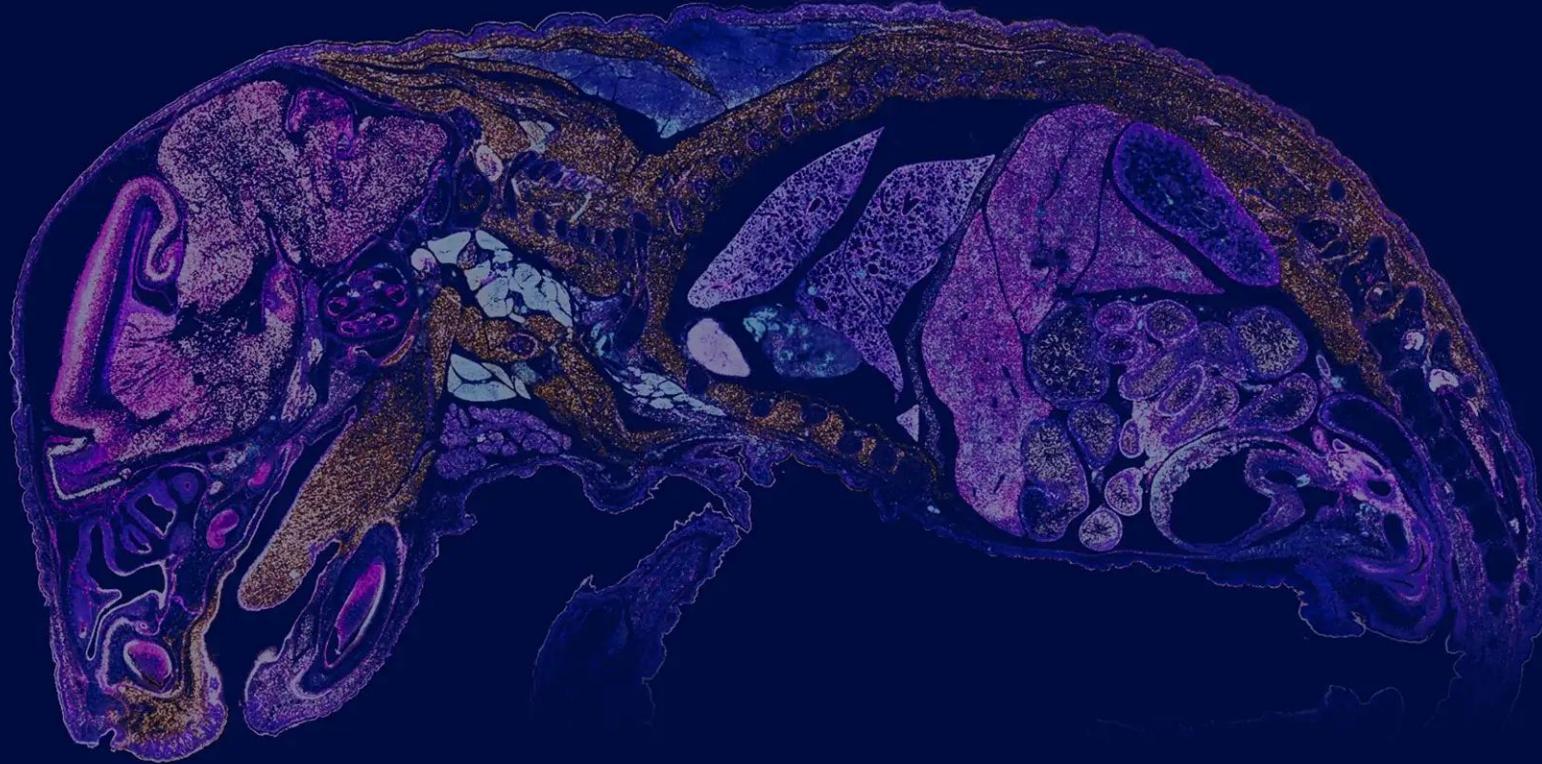
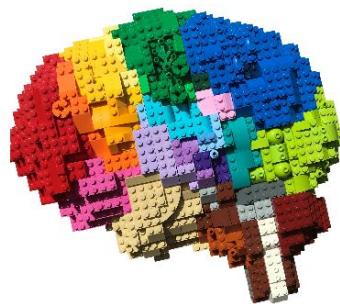


An introduction to spatial biology



What is spatial biology? And why we need it?

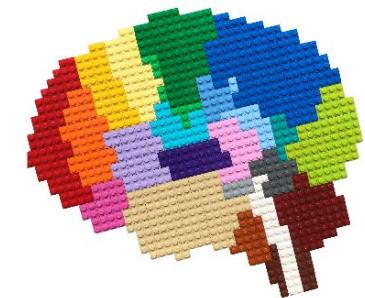
organ



Bulk RNA-seq

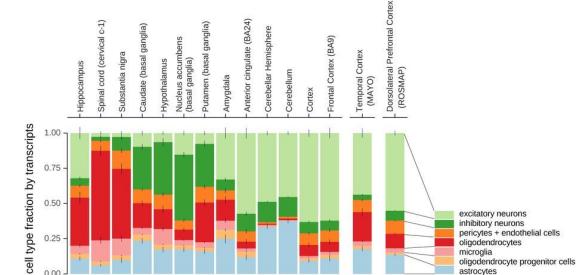


scRNA-seq



spatial RNA-seq

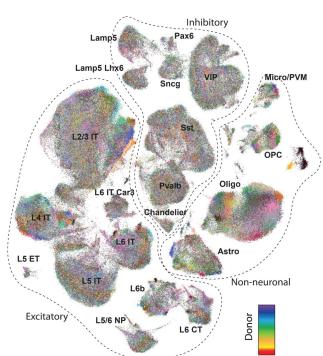
What is spatial biology? And why we need it?



[Park Y, 2021 bioRxiv](#)



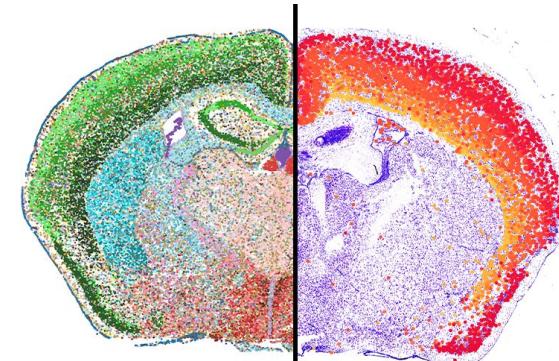
Bulk RNA-seq



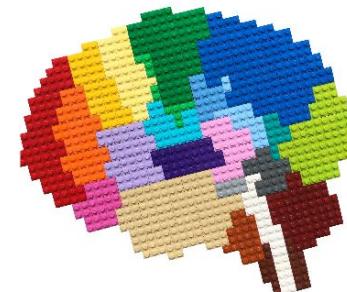
[Yao Z, 2023 Nature](#)



scRNA-seq

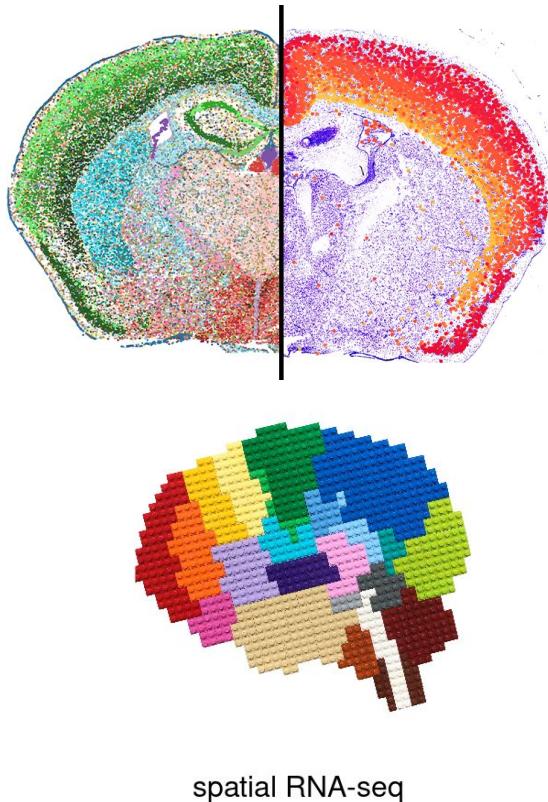


[Broad Institute](#)

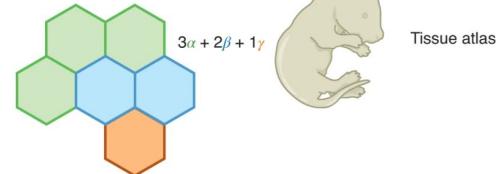


spatial RNA-seq

What is spatial biology? And why we need it?

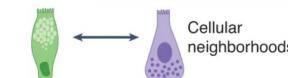
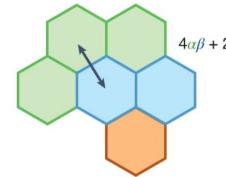


Cell composition



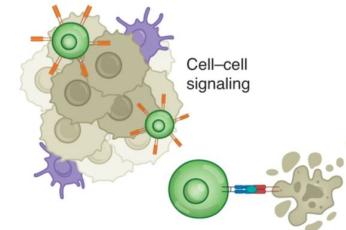
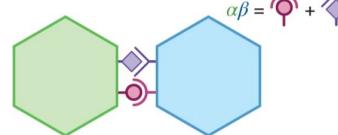
Tissue atlases

Cell-cell interactions



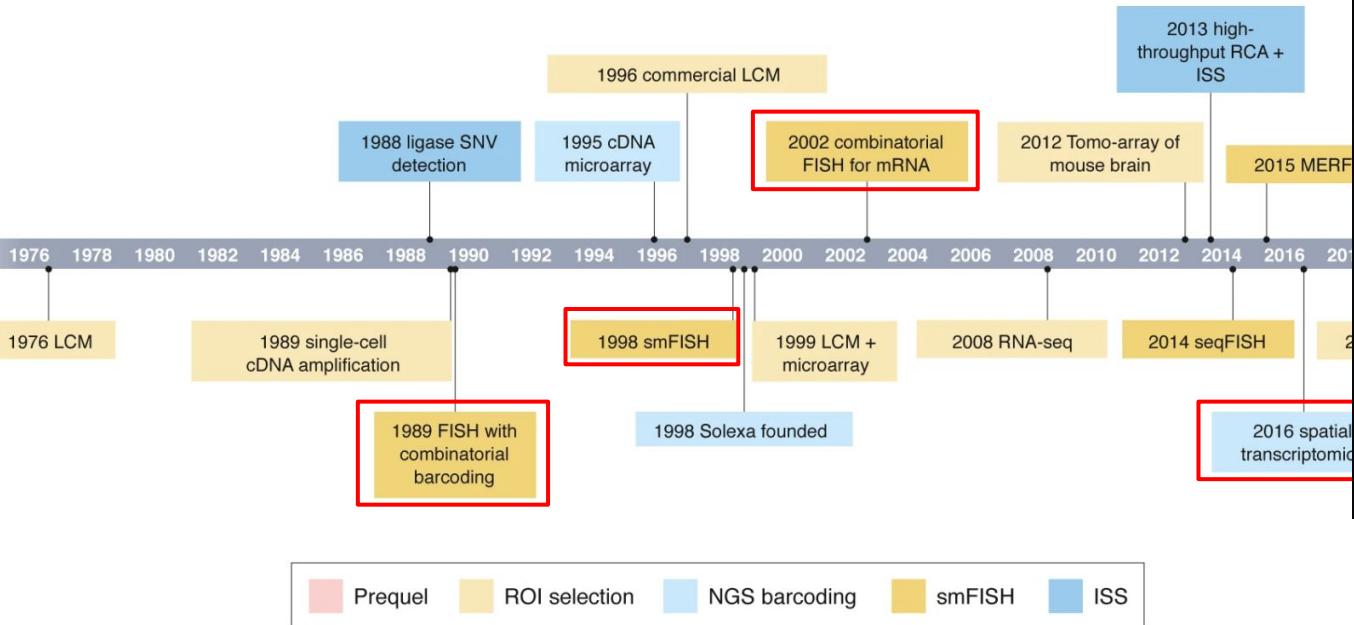
Cellular neighborhoods

Molecular interactions



Cell-cell signaling

Spatial biology: the beginnings



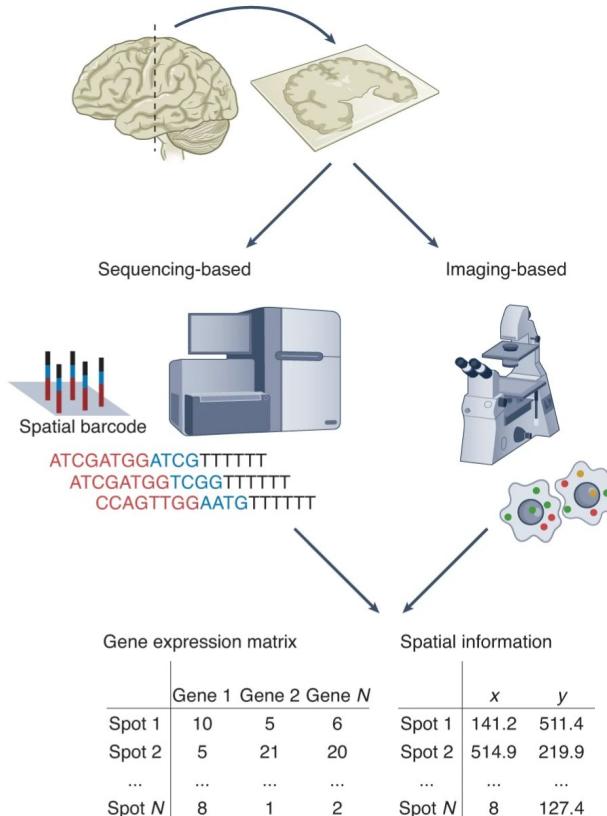
FISH - fluorescent in situ hybridization
smFISH - single-moleculeFISH
LCM - laser capture microdissection



1,789 publications on PubMed with 'spatial transcriptomics' in the abstract or title on 03.07.2024

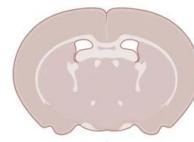
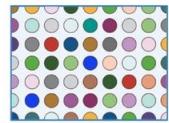
Source: [Moses L & Pachter L \(2020\)](#)

Spatial biology: current technologies



Sequencing-based technologies

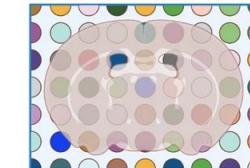
1. Array of spatially barcoded probes



2. Image barcode locations via ISS

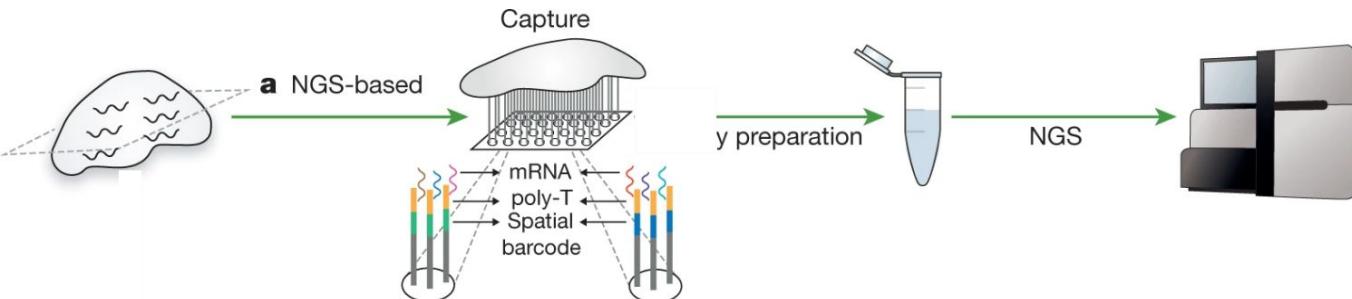
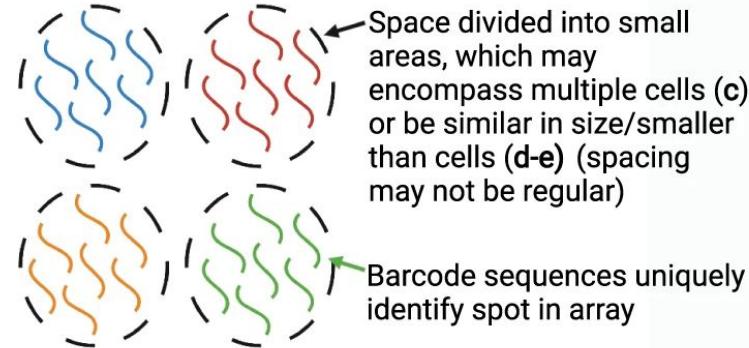


4. NGS of captured probes

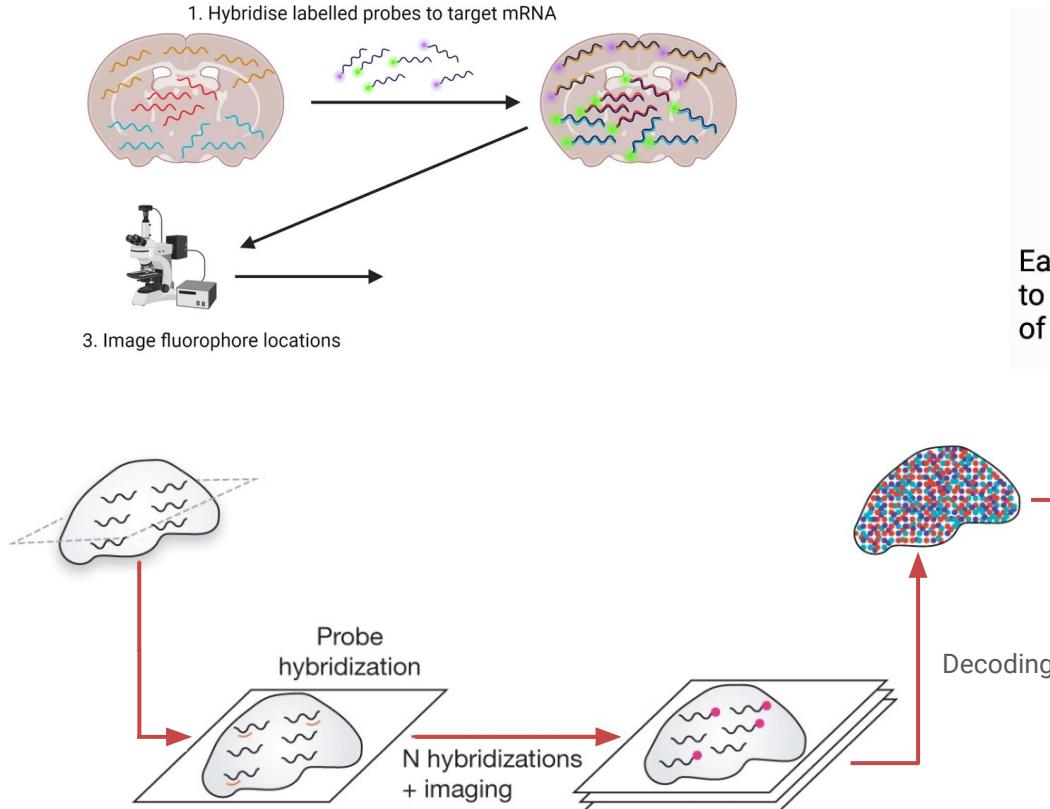


3. Overlay sample on array.
Ligate mRNA to probes.

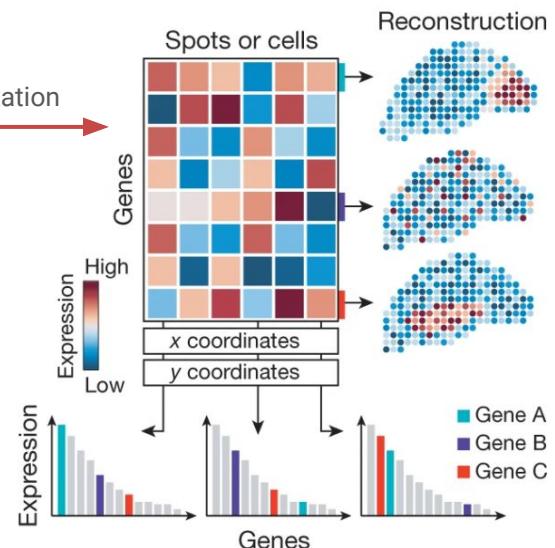
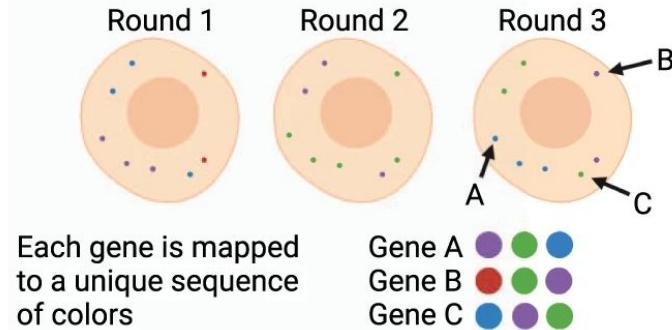
Spatial barcoding



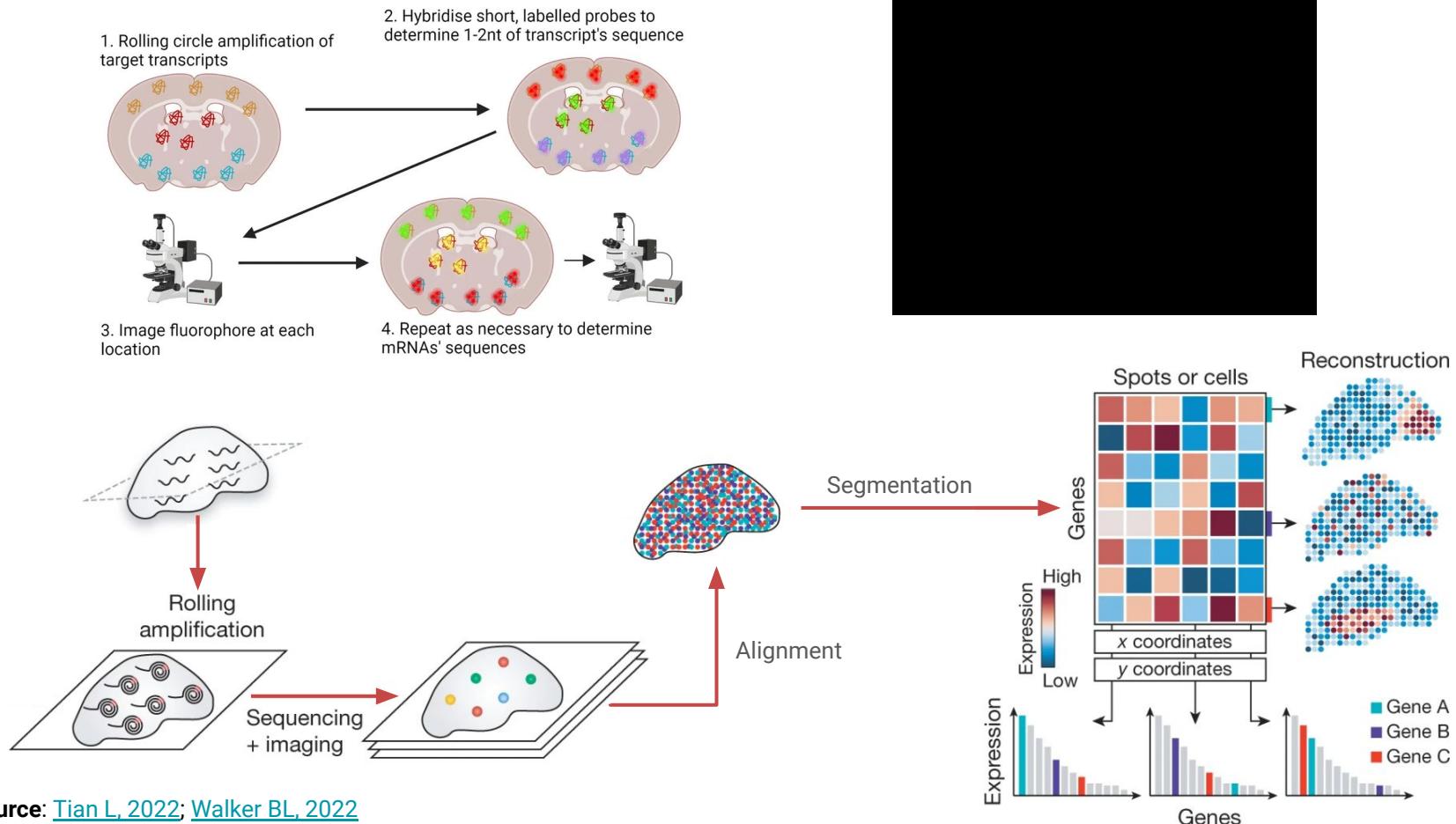
Imaging-based: *in-situ* hybridization



Sequential fluorescent imaging



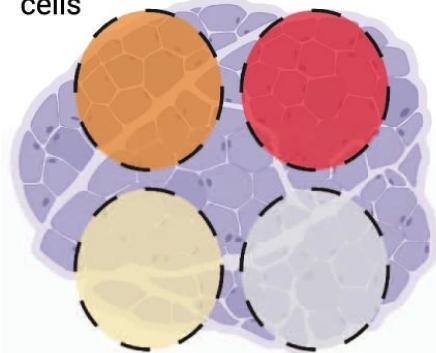
Imaging-based: *in-situ* sequencing



Spatial biology: resolution and methods

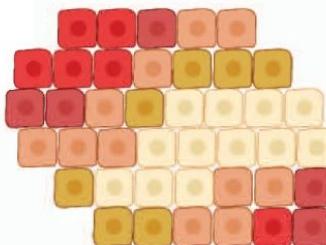
c Multi-cell resolution

Each spot collects expression among several cells



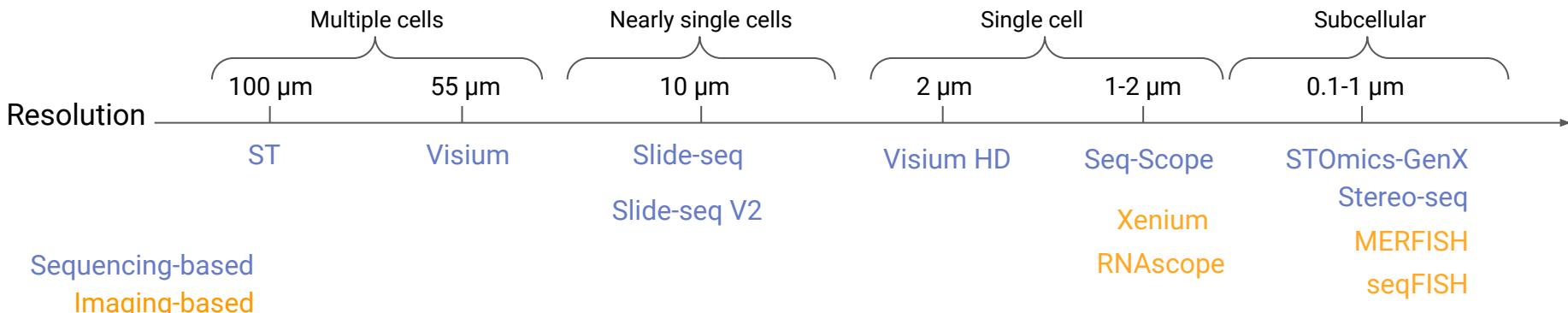
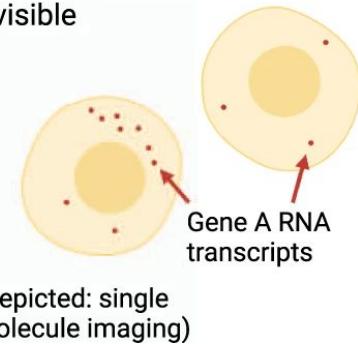
d Single-cell resolution

Data reveals expression patterns of individual cells



e Sub-cellular resolution

In addition to variation between cells, variation in RNA location within cells is visible



Trade-off between resolution and throughput

Source: [Walker BL, 2022](#); [Cheng M. 2023](#)

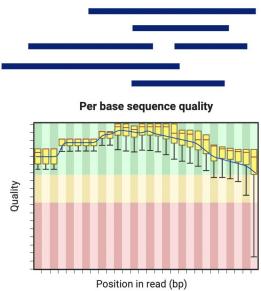
Technologies: outlook

- **Sequencing-based methods** (10x Visium, Slide-seq) capture RNA from tissue sections on a spatially barcoded array, followed by sequencing.
 - ✓ Enables whole-transcriptome analysis.
 - ✗ Limited spatial resolution (often capturing multiple cells per spot).
- **In situ hybridization methods** (smFISH, MERFISH, seqFISH) use fluorescently labeled probes to directly detect RNA molecules within intact cells or tissues.
 - ✓ Single-cell or subcellular resolution.
 - ✗ Limited to pre-selected sets of genes, challenges with tissue autofluorescence and signal amplification.
- **In situ sequencing methods** (STARmap, FISSEQ) perform sequencing reactions directly within the tissue to read RNA sequences in their original location.
 - ✓ Combines high spatial resolution with unbiased gene detection
 - ✗ Lower throughput, very complex experimental procedures, computational challenges in image analysis and base calling.

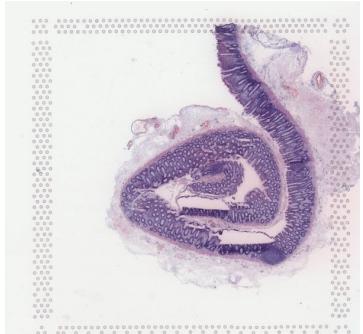
Any questions?

Spatial biology: first data

Sequencing-based methods



Raw sequencing reads

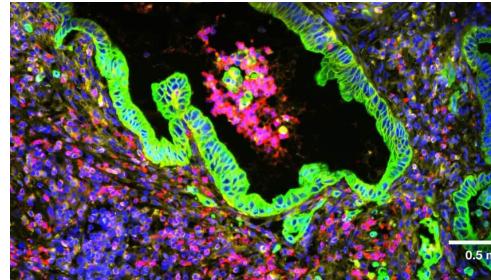


High-Resolution Tissue Image

1 ACGCCTGACACGGCT-1,0,0,0,747,1024
2 TACCGATCCAACACTT-1,0,1,1,819,1066
3 ATTAAGCGGAGCAGC-1,0,0,2,746,1108
4 GATAAAGGGACGATTAG-1,0,1,3,819,1150
5 GTGAAATACCAATA-1,0,0,4,746,1192
6 TGTGGCTGGCGGAAG-1,0,1,5,819,1234
7 GCATTCCTCTCTTATA-1,0,0,6,746,1275
8 GCGAGGGACTGCTAGA-1,0,1,7,819,1317
9 TGTTACCCGCACAGCC-1,0,0,8,746,1359
10 GCGGGTTAACATCGTA-1,0,1,9,818,1401
11 TGCCCTGCCCTTACGG-1,0,0,10,745,1443
12 GAGACTTCCAAGAA-1,0,1,11,818,1485
13 CCAGTGAGCTCTTGT-1,0,0,12,745,1526
14 ATACCTTGCTCAAAAT-1,0,1,13,818,1568
15 GGGTTCCGGCTTCCA-1,0,0,14,745,1610
16 TAACCGTCAGTTCAT-1,0,1,15,817,1652
17 AAACAACGAATAGTTC-1,0,0,16,745,1694

Spatial barcodes

Imaging-based methods



1 "cell_id","vertex_x","vertex_y"
2 "aaacagj-1",824.92505,807.075
3 "aaacagj-1",821.95,808.5625
4 "aaacagj-1",819.61255,810.2625
5 "aaacagj-1",816.0,811.9625
6 "aaacagj-1",816.6375,817.275
7 "aaacagj-1",821.3125,817.7
8 "aaacagj-1",823.86255,818.55005
9 "aaacagj-1",824.92505,818.55005
10 "aaacagj-1",827.05005,815.36255
11 "aaacagj-1",828.11255,812.8125
12 "aaacagj-1",827.6875,810.9
13 "aaacagj-1",825.35004,807.075
14 "aaacagj-1",824.02505,807.075

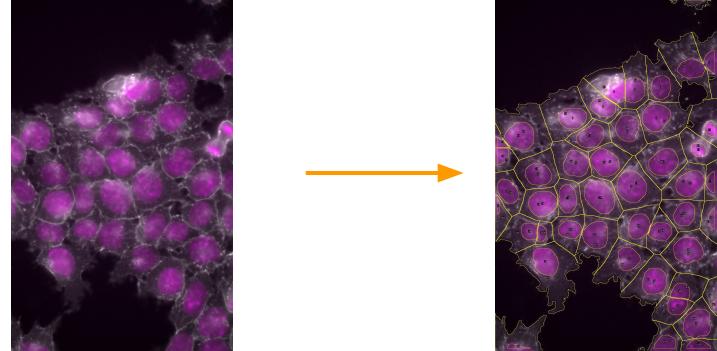
Multi-Channel Fluorescence Images Spot/Pixel Coordinates

"cell_id","x_centroid","y_centroid","transcript_counts","control_probe_counts","control_codeword_counts",
"aaacagj-1",822.3839721679688,813.380859375,24,0,0,0,0,24,96,7246910110116,15.353125557303429
"aaaaifa-1",848.04150390625,834.03271484375,31,0,0,0,0,31,108.69109769538045,20.410625740885735
"aaaajlbn-1",848.1946411132812,188.68264770507812,141,0,0,0,0,141,483.7137655833626,33.4156261295452
"aaaaapaa-1",820.1515502929688,421.4435729980469,96,0,0,0,96,243.7534463480115,56.3098457939928
"aaaakaii-1",853.015625,385.4767761230469,137,0,0,0,0,137,373.7131385654211,30.932032372802496
"aaaakcmb-1",824.2744750976562,854.9688720703125,40,0,0,0,0,40,276.4917287863791,41.99531402438879
"aaaaoijh-1",832.1690063476562,817.1231079101562,15,0,0,0,0,15,63.8960960936574,13.05015673070714
"aabahoh-1",826.76782265625,788.3306884765625,68,0,0,0,0,68,189.8820381425321,28.945157300680876
"aabainpd-1",829.824951171875,471.4559020996094,155,0,0,0,0,155,460.3679854609072,9.16671908274293
"aababmmia-1",821.2094116210938,825.931884765625,23,0,0,0,0,23,147.97703662142158,35.22187627851963
"aaaccejp-1",845.7847290039062,805.4239501953125,12,0,0,0,0,12,41.45189452694715,5.644531454890966
"aaaccnpn-1" 827.3102116007188 820.8207827118138 12 0 0 0 0 12 20 84828223246213 8 444210465616886

Fluorescence Intensity Data

Cell segmentation

- Process of identifying individual cells within tissue images.
- Essential for imaging-based methods.



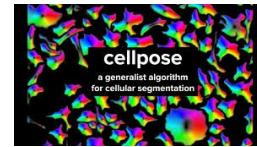
ImageJ
Image Processing and Analysis in Java

- Requires manual input and expertise for complex segmentations
- Good for general image analysis and simple segmentation tasks



ilastik

- Interactive machine learning-based segmentation
- User trains the model by annotating a few examples
- Good for complex tissues with varying cell types



cellpose
a generalist algorithm
for cellular segmentation



squidpy

- Deep learning-based tool for cell segmentation
- Generalist algorithm that works well on various cell types without retraining
- Can handle both 2D and 3D images

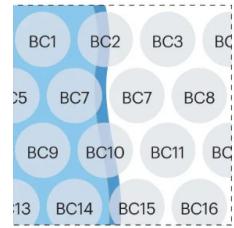


squidpy

- Specialized deep learning model for nuclei detection and segmentation
- Particularly effective for densely packed nuclei

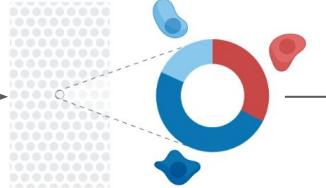
Spatial biology: the tools

Sequencing-based



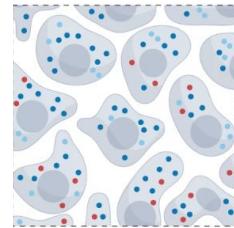
m_{raw} genes	→	x	y
0	5	2	...
10	0	0	...
15	0	0	...
...

Deconvolution



m_{raw} genes	→	x	y
15	2	...	-12
5	10	...	3
3	1	...	-13
...

Imaging-based

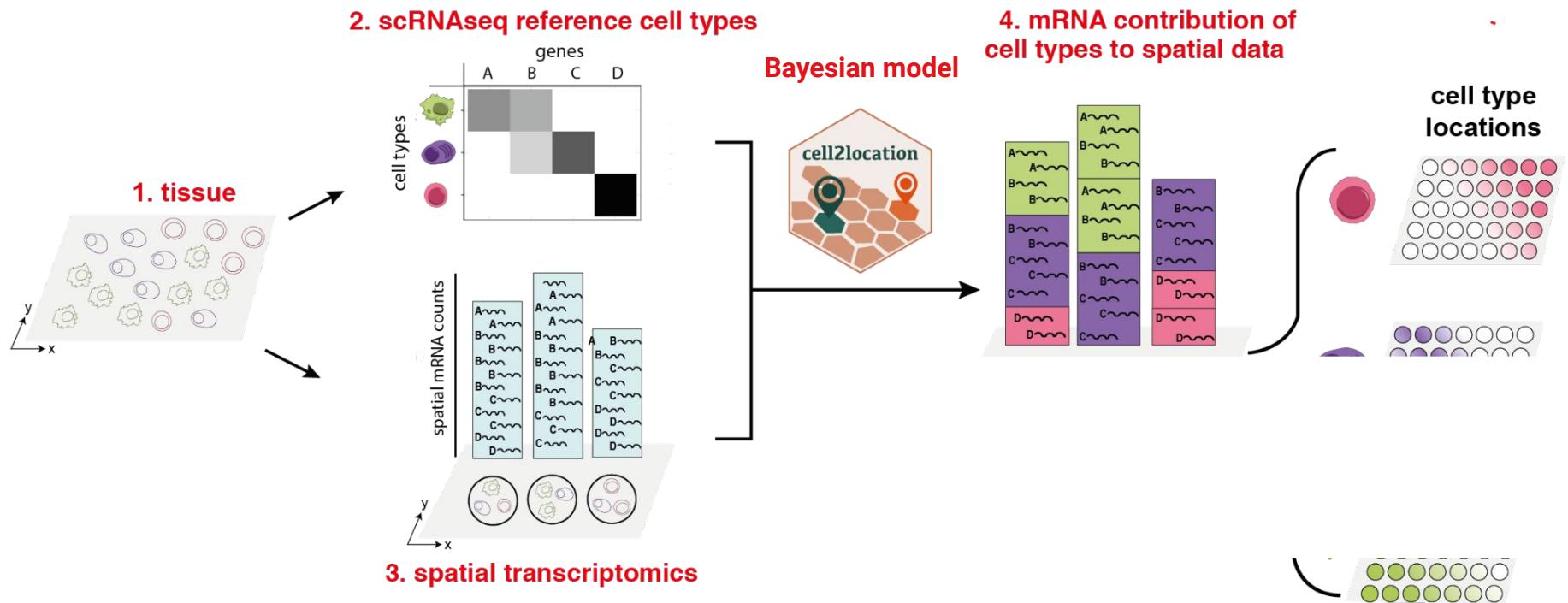


n_{raw} cells	→	x	y
15	2	...	-12
5	10	...	3
3	1	...	-13
...

Processing analogous to single-cell:

- QC (counts, genes, % of mitochondrial, etc)
- Normalization & log-transformation
- Dimensionality reduction (PCA, UMAP, tSNE)

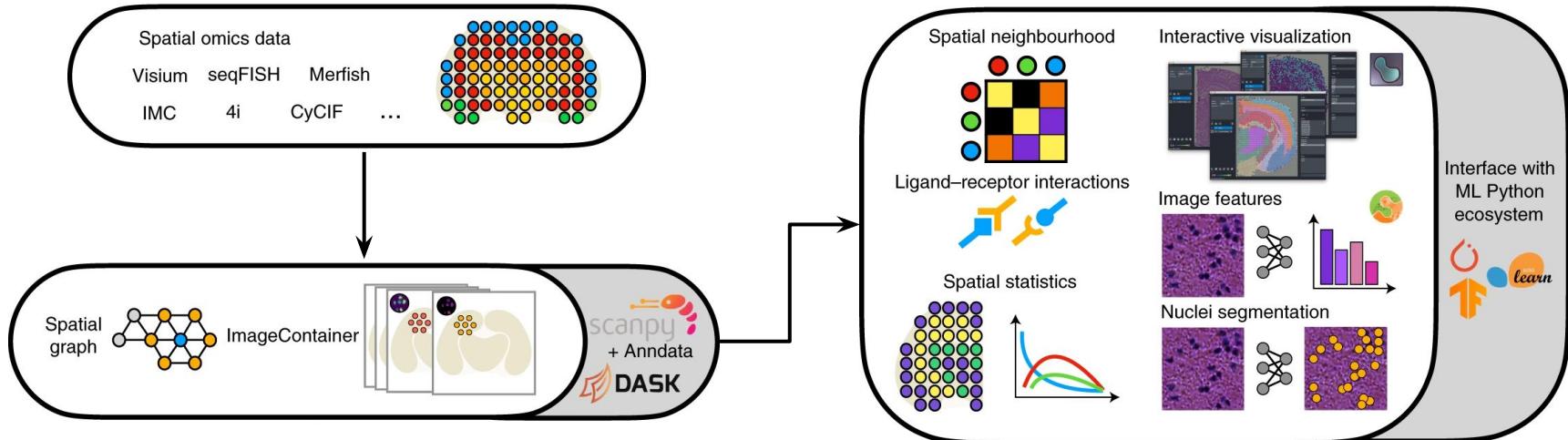
Spatial biology: deconvolution



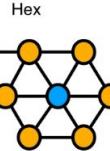
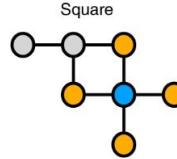
Spatial biology: R or Python



Spatial biology: squidpy



Grid graphs

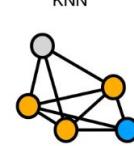
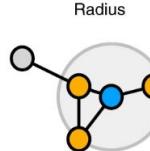


- ✓ Simple, uniform connectivity
- ✗ May not capture natural tissue structure

How to choose:

- Technology considerations (Visium - grid type and size are predetermined)
- Biological scale (size of cells or structures)
- Tissue density
- Local vs global tissue patterns
- Computational capabilities

Generic graphs



- ✓ Biologically intuitive (based on physical distance)
- ✗ Can lead to isolated points in sparse areas

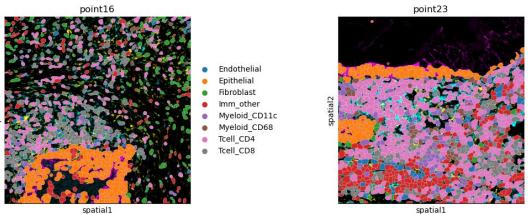
- ✓ Adapts well to varying density
- ✗ May create long-range connections in sparse areas

- ✓ Ensures each point has connections
- ✗ Can create unnatural long-range connections in variable density tissues

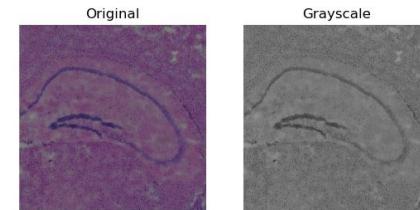
squidpy ImageContainer

- **ImageContainer** data structure designed to handle and process spatial image data efficiently.
- Integrates image to AnnData object.

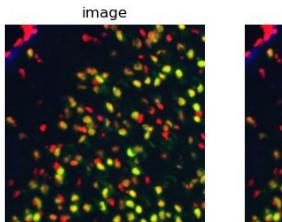
Show layers



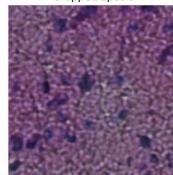
Convert to grayscale



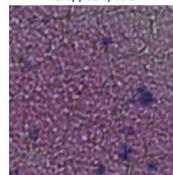
Crop images



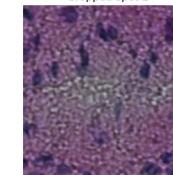
Cropped spot 0



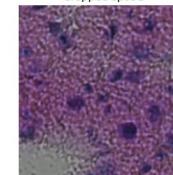
Cropped spot 1



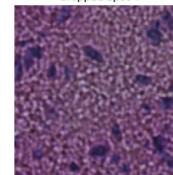
Cropped spot 2



Cropped spot 3



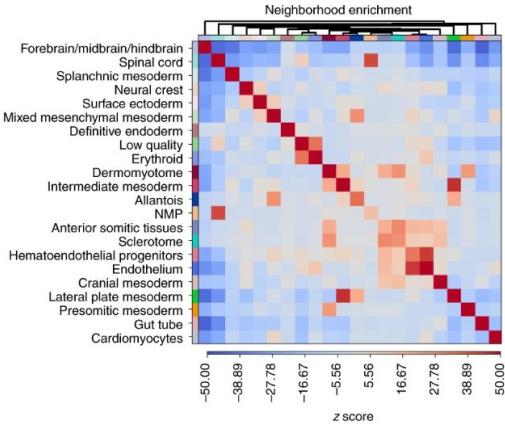
Cropped spot 4



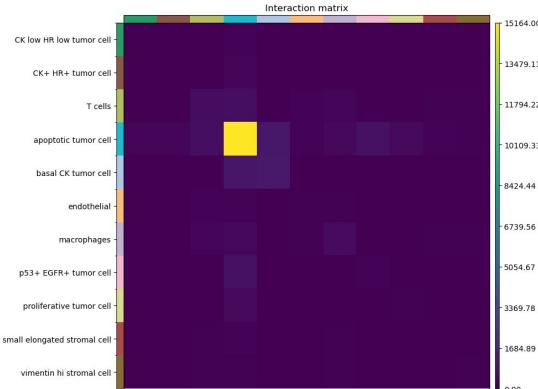
Generate cropped images from spots

Spatial biology: squidpy

Neighborhood enrichment analysis

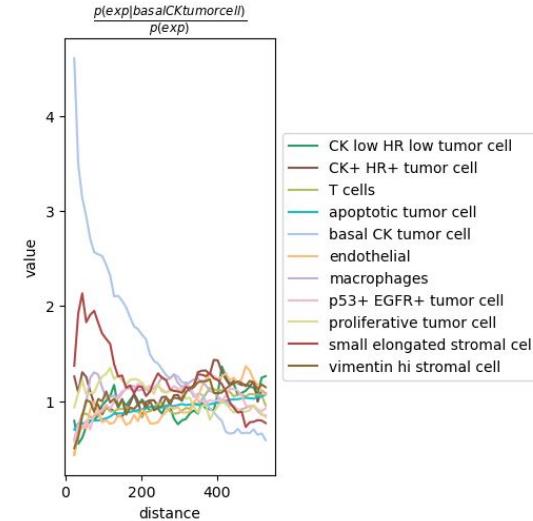


Interaction matrix



Based on proximity on the connectivity graph of cell clusters

Co-occurrence probability



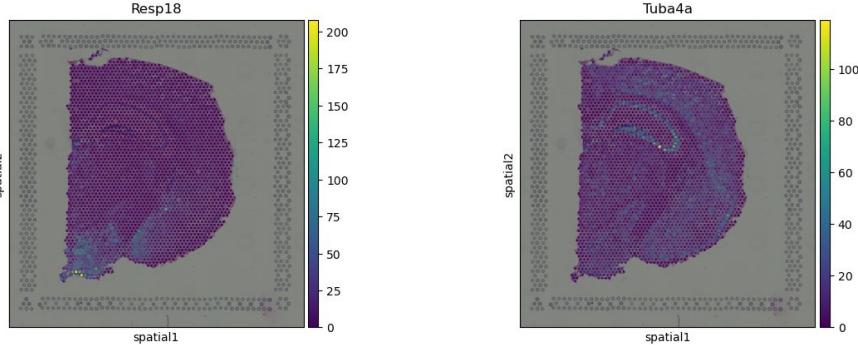
Number of edges that nodes belonging to a given annotation shares with the other annotations

$$\frac{p(exp|cond)}{p(exp)}$$

Likelihood of finding one cell type in the neighborhood of another, compared to what you'd expect by chance.

Spatial biology: squidpy

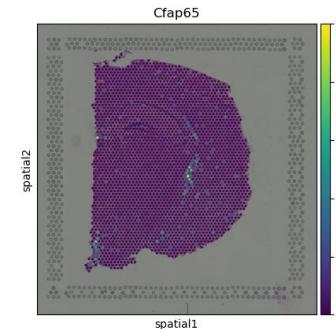
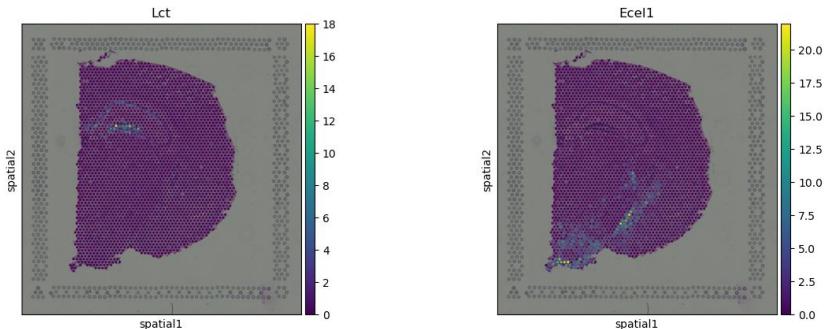
Moran's I score



- Spatial auto-correlation statistics.
- Evaluates whether genes shows a pattern that is clustered, dispersed or random in the tissue are under consideration.

Whether similar expression values are clustered together in space.

Sepal score

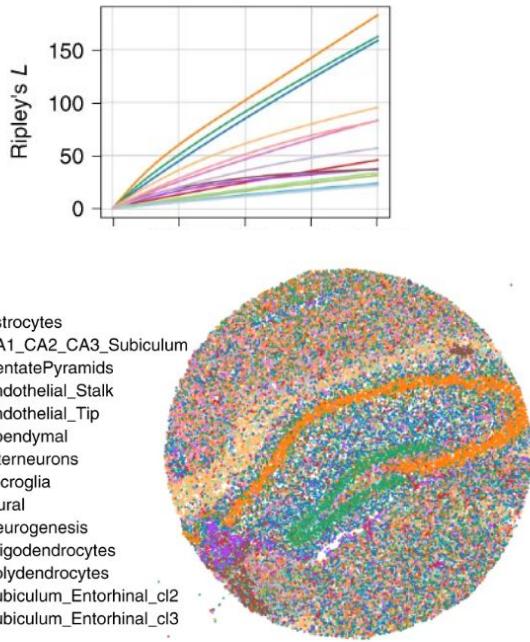


- Identifies spatially variable genes identification.
- Accepts only grid graphs.
- It is useful to filter out genes that are expressed in very few observations and might be wrongly identified as being spatially variable.

How much a gene's expression varies across space.

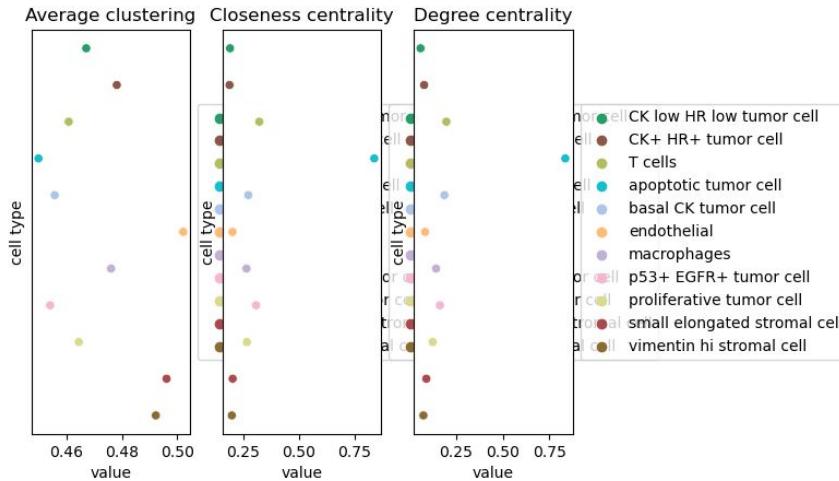
Spatial biology: squidpy

Ripley statistics



Determines whether points have a random, dispersed or clustered distribution pattern at certain scale.

Centrality scores



Centrality measures how important or central a node (cells or spatial locations) is in a network.

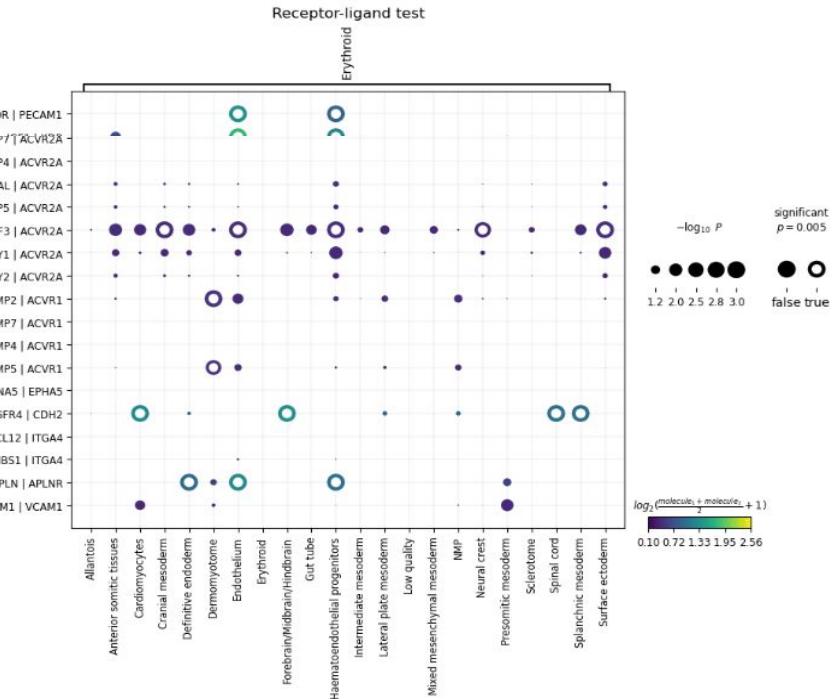
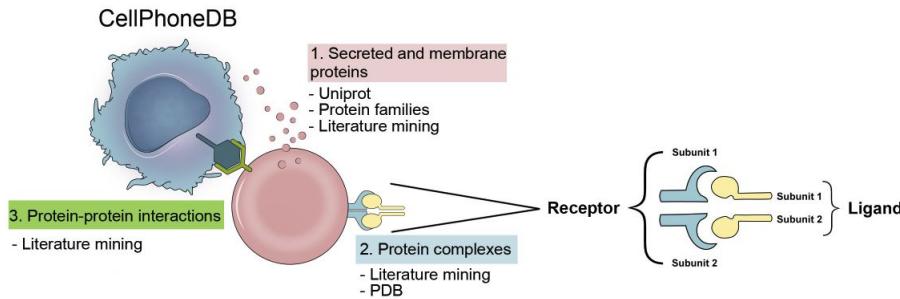
Closeness Centrality - how close a node is to all other nodes in the network → identifies potential signaling hubs.

Degree Centrality - measures the number of direct connections a node has → finds local communication center.

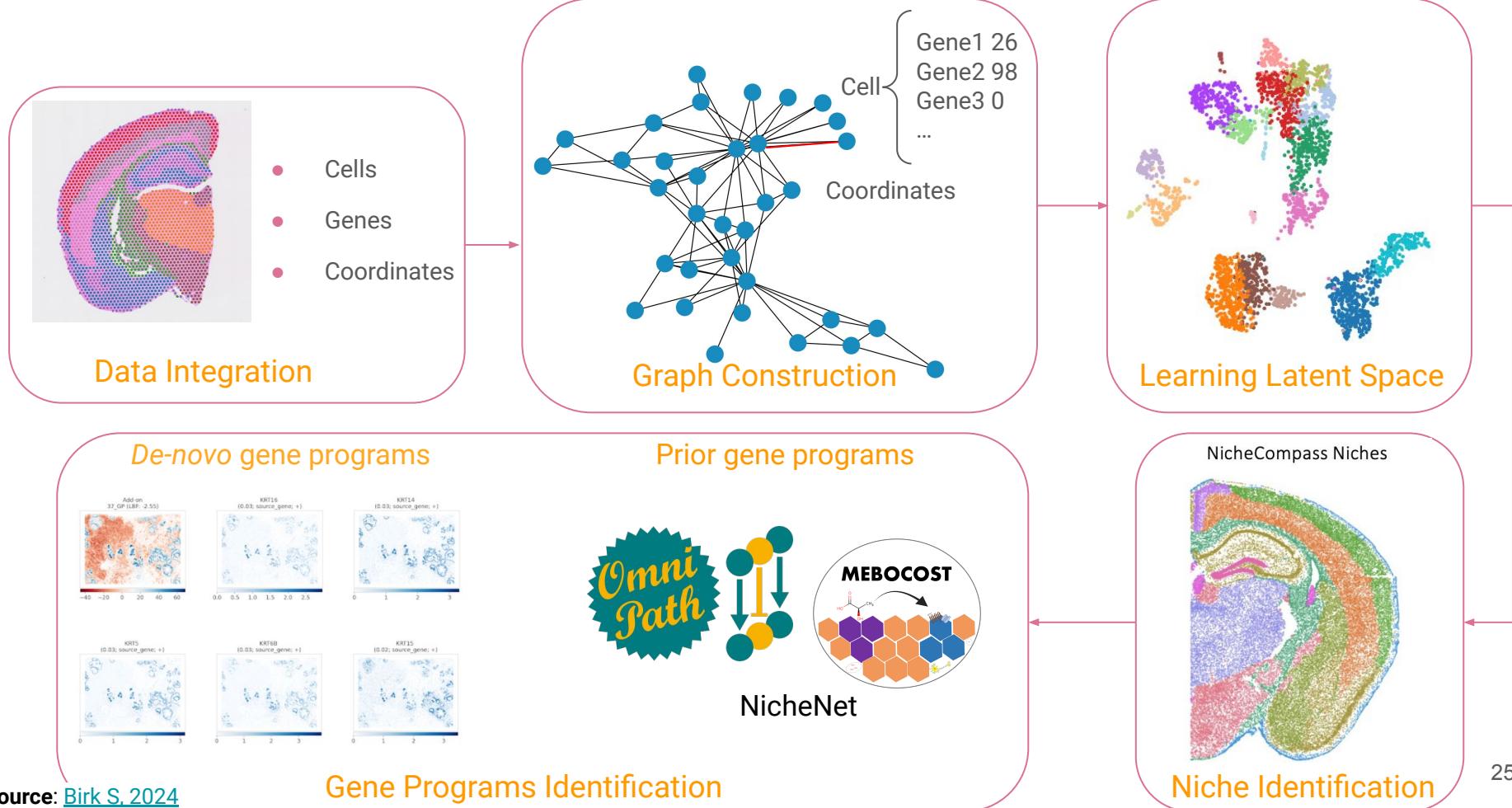
Clustering Coefficient - measures how much nodes tend to cluster together → identifies potential functional modules or niches in the tissue.

Spatial biology: squidpy

Ligand-receptor interactions



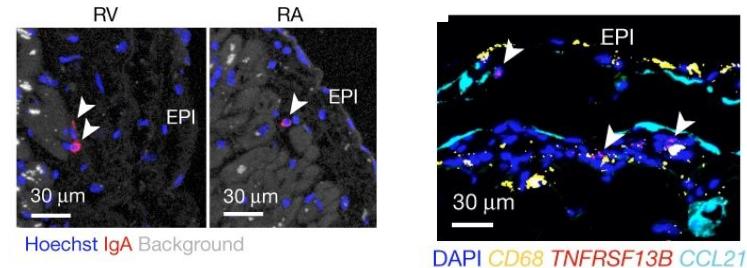
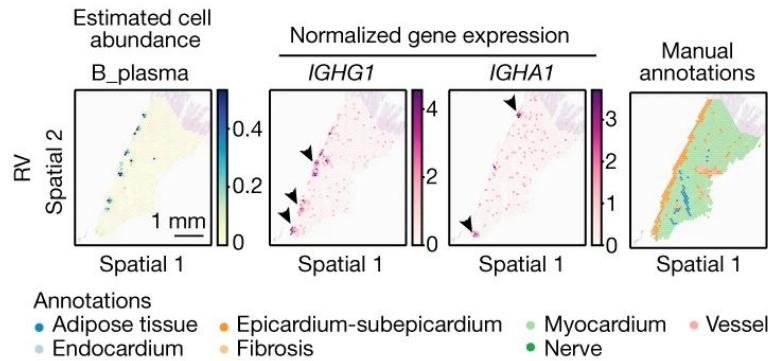
Advanced analysis: NicheCompass



Data analysis: outlook

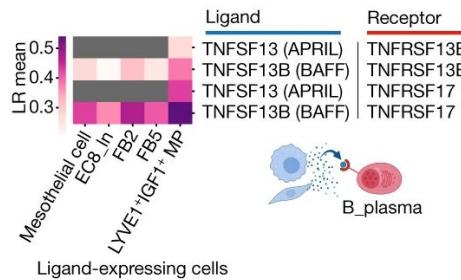
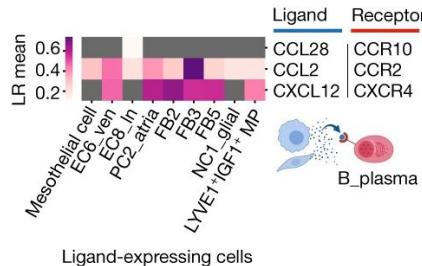
- **Cell segmentation** is the process of identifying individual cells in tissue images. It's essential for imaging-based spatial transcriptomics methods to achieve single-cell resolution.
- **Deconvolution** is a computational method used to infer cell-type specific gene expression from mixed signals. It's crucial for sequencing-based spatial transcriptomics methods where spots often contain multiple cells.
- **Squidpy** is a Python-based tool for analyzing spatial molecular data, integrating with existing single-cell analysis frameworks like Scanpy.
- **Spatial graph** in Squidpy represents the spatial relationships between spots or cells in the tissue. It's a fundamental structure for many spatial analyses.
- Spatial graph in Squidpy enables various spatial analyses including neighborhood enrichment, spatial autocorrelation (Moran's I), clustering, and co-occurrence patterns.
- Squidpy provides tools for integrating and analyzing image data alongside gene expression data.
- **NicheCompass** is an advanced computational method for identifying and characterizing cellular niches in spatial transcriptomics data. It integrates spatial information, gene expression, and prior knowledge to define functional microenvironments within tissues.

Applications of spatial biology: Cardio-Immune Niches



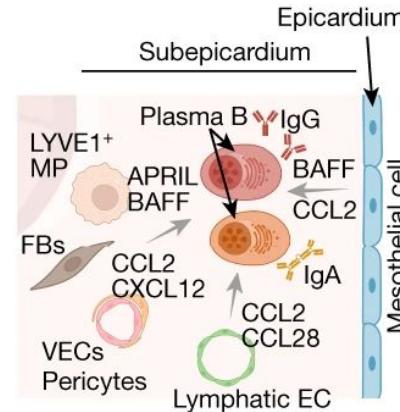
smFISH confirmed the findings.

Different mechanisms regulate the fate of IgG+ and IgA+ plasma B cells in epicardium.



Chemokine signalling recruits B cells in the niche.

TNF signalling for niche homeostasis.



Source: Kanemaru K, 2023

Thank you!

Any questions?



Anna Maguza



anna.maguza@uni-wuerzburg.de



anna-maguza-462968172/



AnnaMaguza/lisbon_spatial_workshop_2024