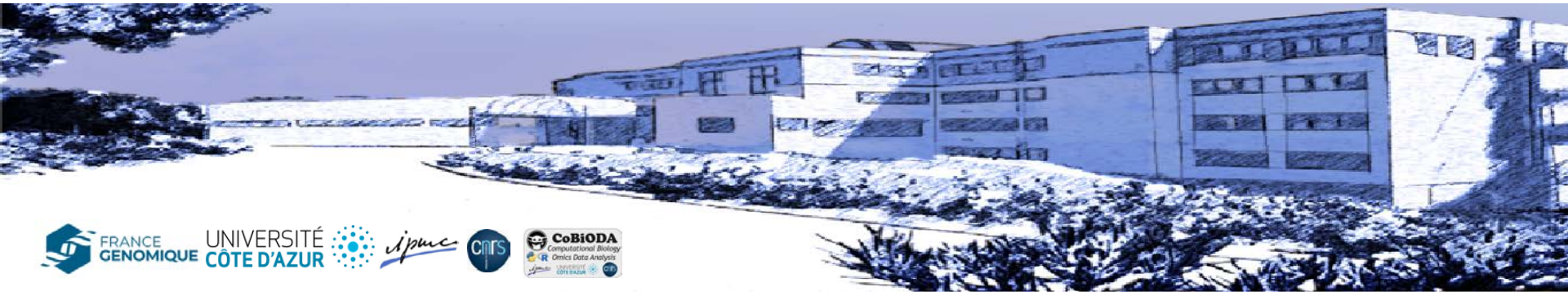
Imaging-based Spatial Transcriptomics

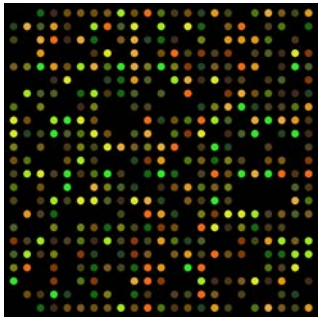
Biological Image Processing and Analysis (BioIAS)
Nice, June 14th. 2024

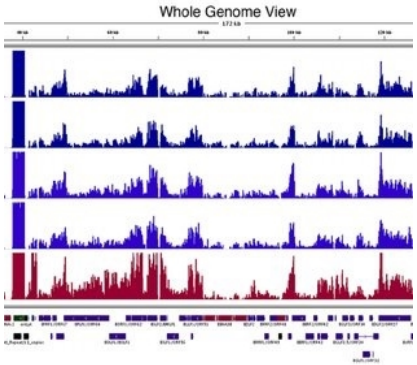
Kévin Lebrigand
Computational Biology and Omics Data Analysis
IPMC, CNRS, Côte d'Azur University, France
https://cobioda.github.io
lebrigand@ipmc.cnrs.fr
@kevinlebrigand

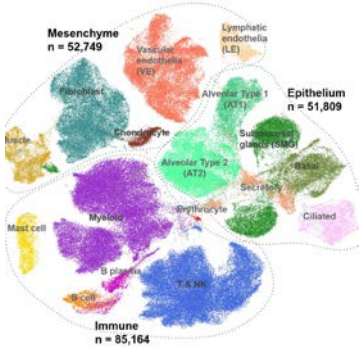


20 years of transcriptomics

Driven by microfluidics technological developments







Early 2000's: DNA microarray

- Large-scale transcriptome
- Oligonucleotide probe tilling
- Fluorochrome signal analysis
- Bulk resolution

Cost : 4k€
20 samples
25k genes
0,5M matrix

Late 2000's : RNA sequencing

- Whole transcriptome
- Next Generation Sequencing
- Full-transcript coverage
- Bulk resolution

Cost : 4k€
20 samples
50k genes
1M matrix

Mid 2010's: Single -cell

- Whole transcriptome
- Microfluidics + NGS
- 3p-end gene signal (UMI)
- Sensitivity (6%)
- Single-cell / state resolution

Cost : 4k€
5k cells
50k genes
250M matrix

2020's : Spatial

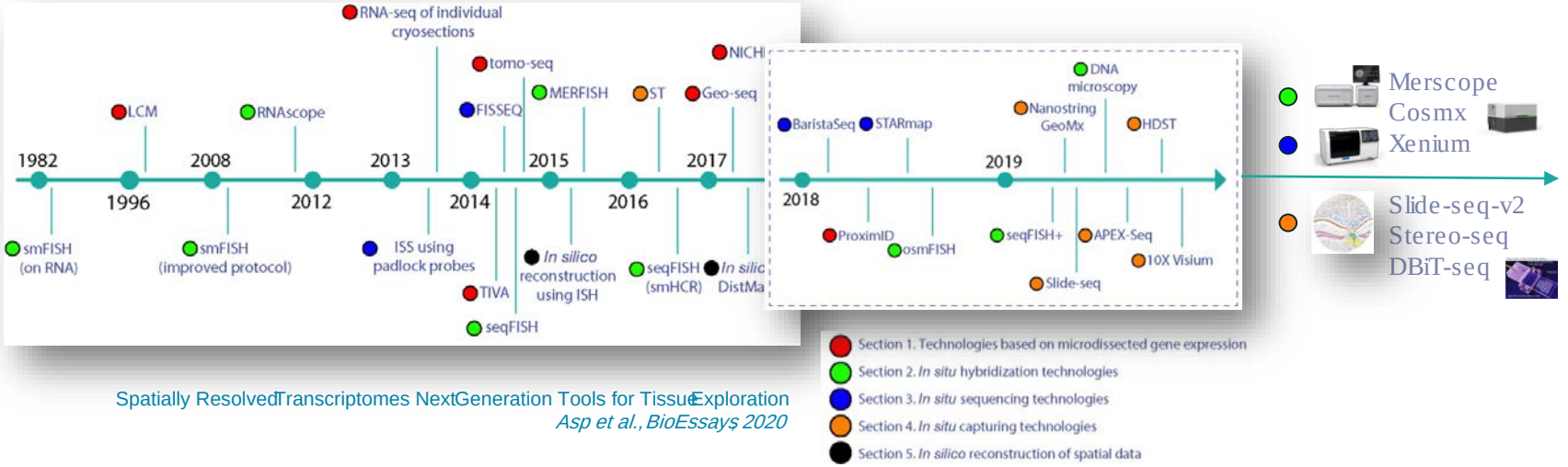
- 500-1000 gene targets
- Imaging analysis
- Multiplexing FiSH (single molecule)
- Sensitivity (30-80%)
- Sub-cellular resolution

Cost : 4k€
250k cells
1k genes
250M matrix
+ Spatial dimension

Spatial Transcriptomics approaches

Historical timeline

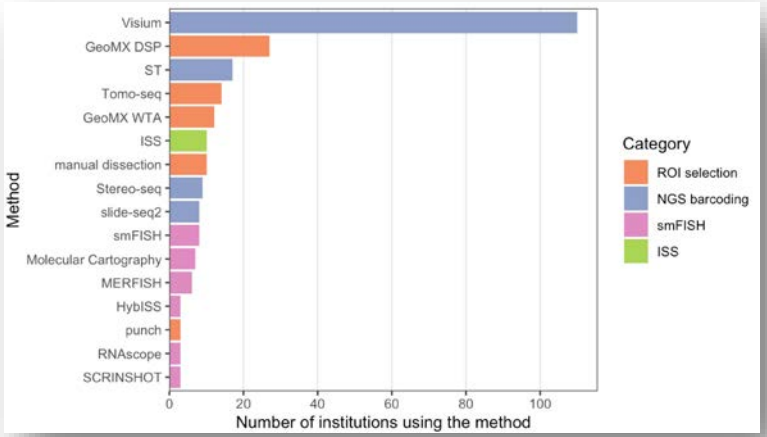
- Spatial transcriptomics aims at directly visualize gene expression in their original environment
- Tackle the main limitation of single cell experiment missing the spatial organization
- A lot of developments in the last years thanks to recent advances in different fields

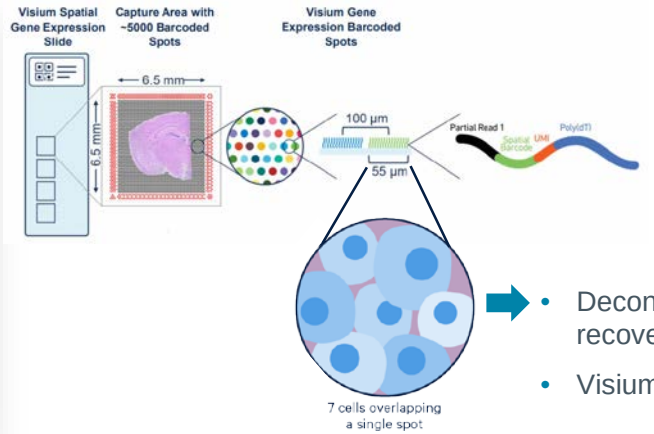


Spatially Resolved Transcriptomes Next Generation Tools for Tissue Exploration
Asp et al., BioEssays 2020

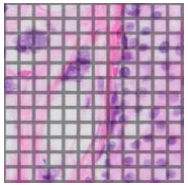
In-situ capture Spatial Transcriptomics (2017 -2022)

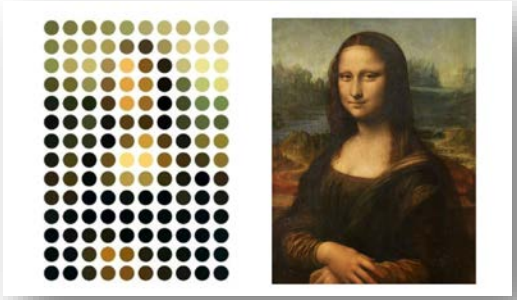
Visiumis widely adopted by academics





- Deconvolution tools can be use to recover proportion of single cell type
- Visium HD single cellresolution (2µm)





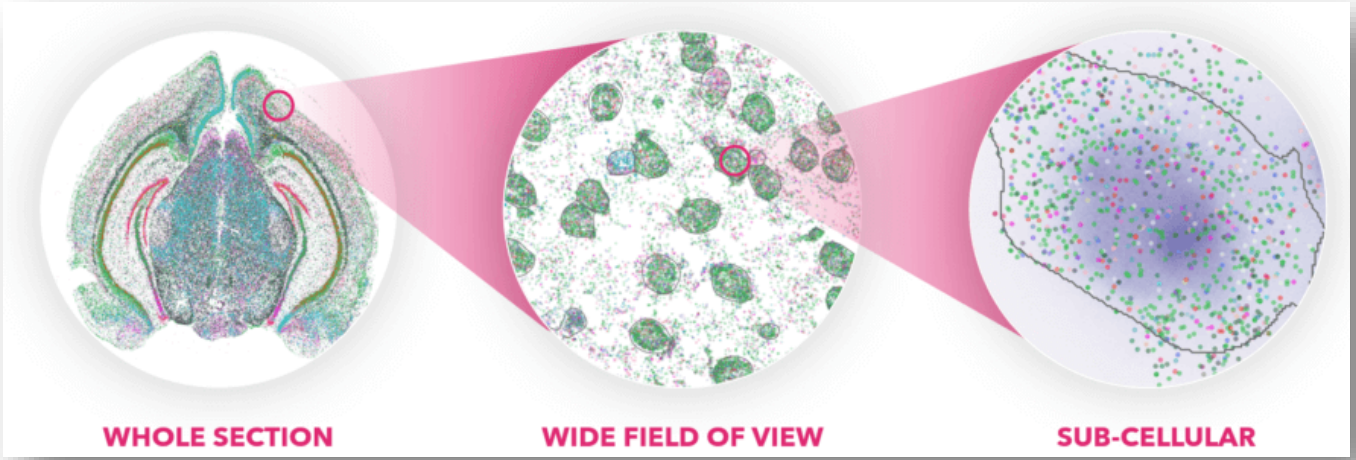
But is not the idealreadout for spatialbiology

(Akoya credit rough caricature)

Imaging -based Spatial Transcriptomics (2022)

No more sequencing for direct single-cell resolution

- Lower gene panel targets (from whole transcriptome to ~1,000 genes)
- Higher sensitivity (from ~6% to 30-80%)
- Larger imaging area (42 to 236 mm2)
- Higher resolution (from 55 µm to subcellular)

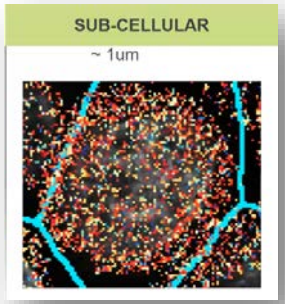


Imaging -based Spatial Transcriptomics (2022)

No more sequencing for direct single-cell resolution

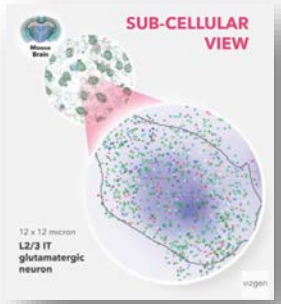
Nanostring CosMx

- 960 targets (panel 20k, AGBT24)
- Sensitivity: << 30-80% (+)
- Imaging area:16 mm2 (2 days)
- Resolution:200 nm



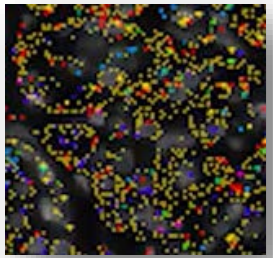
Vizgen Merscope

- 1.000 targets
- Sensitivity: 30-80% (+++)
- Imaging area: 100 mm2 (2 days)
- Resolution 100 nm



10xGenomics Xenium

- 400 - 6,000 targets
- Sensitivity : 5-30% (++)
- Imaging area:236 mm2 (4 days)
- Resolution 200 nm



Imaging -based Spatial Transcriptomics (2022)

No more sequencing for direct single-cell resolution

Nanostring CosMx

ISH-based

Vizgen Merscope

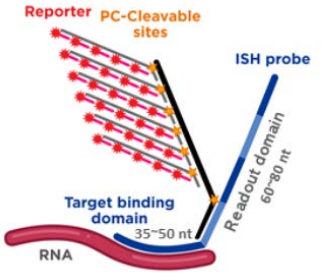
Multiplex Error-Robust FISH

Available (oct.2022)

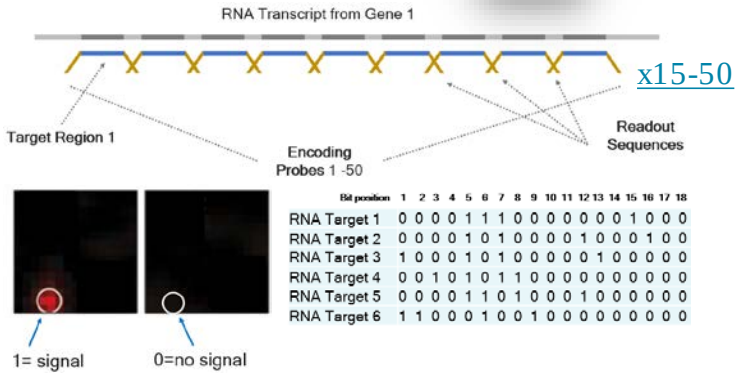
10xGenomics Xenium

Cartana ISS, padlock probes / RCA

Available (jan.2024)

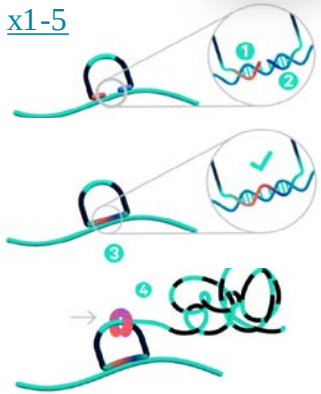


x4-8 / target gene



Cyclic *in situ* Hybridization Chemistries

x1-5



Imaging -based Spatial Transcriptomics platforms comparison

2 recent bioRxiv comparative studies

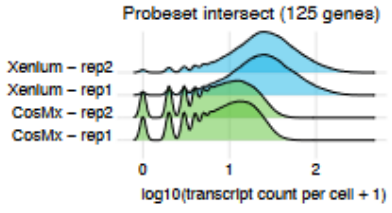
A Comparative Analysis of Imaging-Based Spatial Transcriptomics Platforms

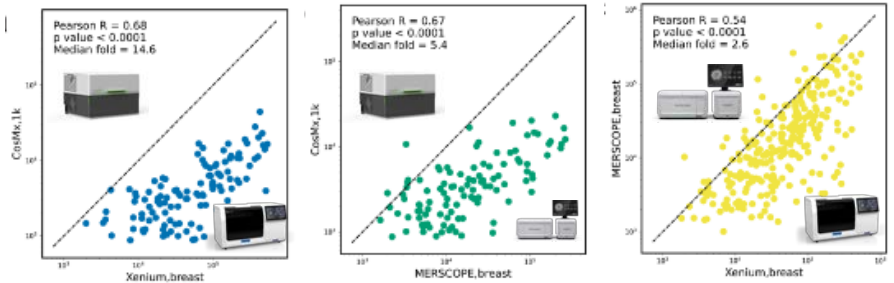
David P. Cook[1], Kirk B. Jensen[2,3,4], Kellie Wise[2,3], Michael J. Roach[2,3], Felipe Segato Dezem[6,7], Natalie K. Ryan[3,5], Michel Zamojski[9], Ioannis S. Vlachos[10,11,12], Simon R. V. Knott[13,14], Lisa M. Butler[3,5], Jeffrey L. Wrana[1,15], Nicholas E. Banovich[16], Jasmine T. Plummer[6,7,8*], Luciano G. Martelotto[2,3*]

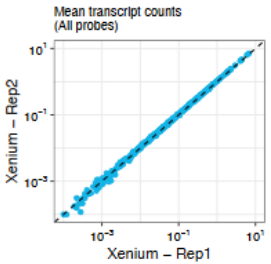
Systematic benchmarking of imaging spatial transcriptomics platforms in FFPE tissues

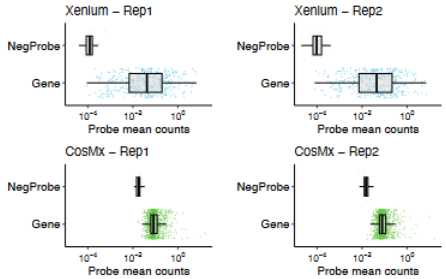
Huan Wang[1,*], Ruixu Huang[2,*], Jack Nelson[1,*], Ce Gao[3], Miles Tran[3], Anna Yeaton[4], Kristen Felt[5], Kathleen L. Pfaff[6], Teri Bowman[7], Scott J. Rodig[6,7], Kevin Wei[3,7], Brittany A. Goods[2,**], Samouil L. Farhi[1,**]

	Xenium Rep 1	Xenium Rep 2	CosMx Rep 1	CosMx Rep 2
Gene target #	377	377	1000	1000
Total cell count	99,852	102,508	96,139	98,767
Median gene count per cell	33	34	75	71
Median transcript count per cell	88	92	113	99
Median transcript count / gene target count	0.23	0.24	0.11	0.10
Median transcript count (intersecting targets only)	23	24	8	7









- CosMx is much less sensitive (high FPR)
- Merscope / Xenium for Fresh frozen slice
- Xenium optimal for FFPE slice

Gene targets panel design

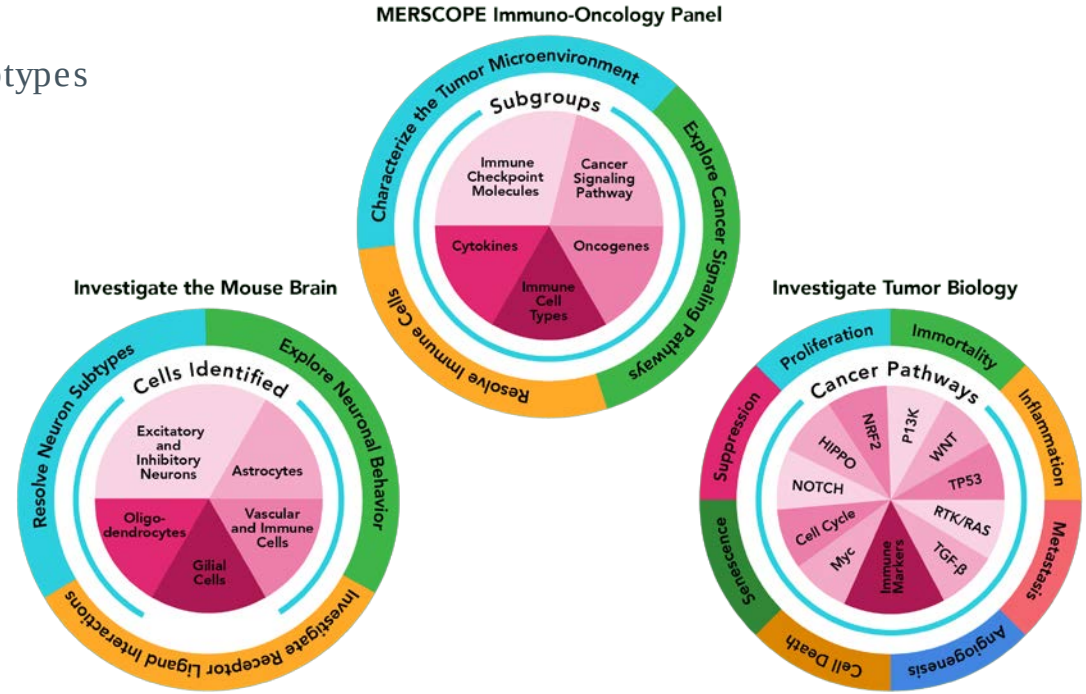
Depending on the biological question !

Depending of your specific scientific focus

- Identify all major cell types, resolve cell subtypes
- Explore functional information
- Investigate interactions between cell types
- Ligand-receptors analysis
- Explore canonical signaling pathways
- Profile immune checkpoint molecules
- …

Satisfy technological system limitations

- Number of targets available
- Range of gene targets expression
- Total gene targets expression
- Budget around 15 k€ for 10 reactions



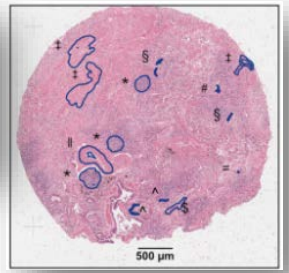
https://portal.vizgen.com/
https://cloud.10xgenomics.com/xenium-panel-designer

Experimental design

Take advantage of the large imaging area

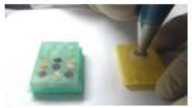
Image-based spatial transcriptomics identifies molecular niche dysregulation associated with distal lung remodeling in pulmonary fibrosis

Annika Vannan[1,a], Ruqian Lyu[2,3,a], Arianna L. Williams[1], Nicholas M. Negretti[4], Evan D. Mee[1], Joseph Hirsh[4], Samuel Hirsh[4], David S. Nichols[5], Carla L. Calvi[5], Chase J. Taylor[5], Vasiliy. V. Polosukhin[5], Ana PM Serezani[5], A. Scott McCall[5], Jason J. Gokey[5], Heejung Shim[3], Lorraine B. Ware[5,7], Matthew J. Bacchetta[8], Ciara M. Shaver[5], Timothy S. Blackwell[5,9,10], Rajat Walia[11], Jennifer MS Sucre[4,9], Jonathan A. Kropski[5,9,10,b], Davis J McCarthy[2,3,b], Nicholas E. Banovich[1,b,*]



Quick-Ray

Recipient Blocks



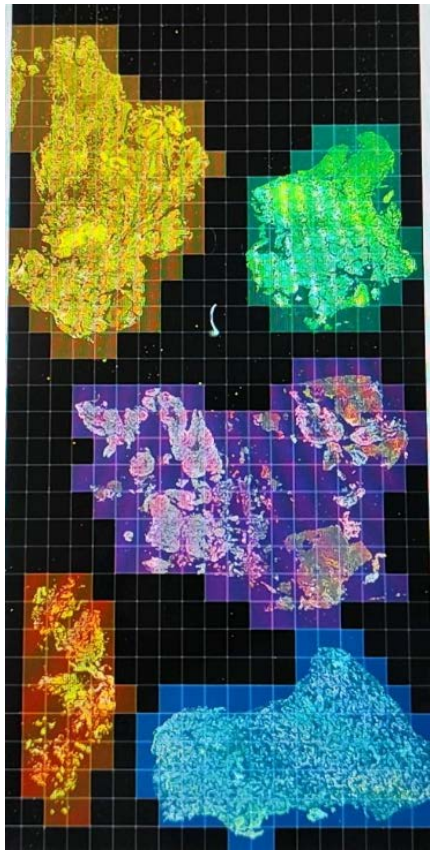
https://www.ihcworld.com/products/Quick-Ray-Mold.htm

Each slide cost around5 k€

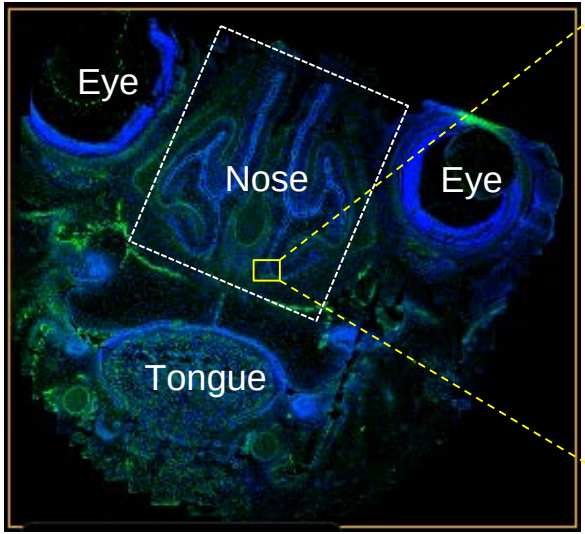
multiplexing to remove batch effect and increase replicates for robust statistical analysis

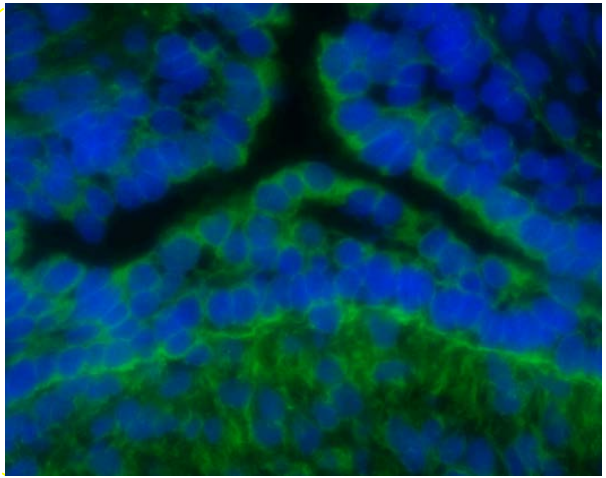
Data acquisition

DAPI and cell boundaries staining for cell segmentation



Human fetal head section (PCW9)



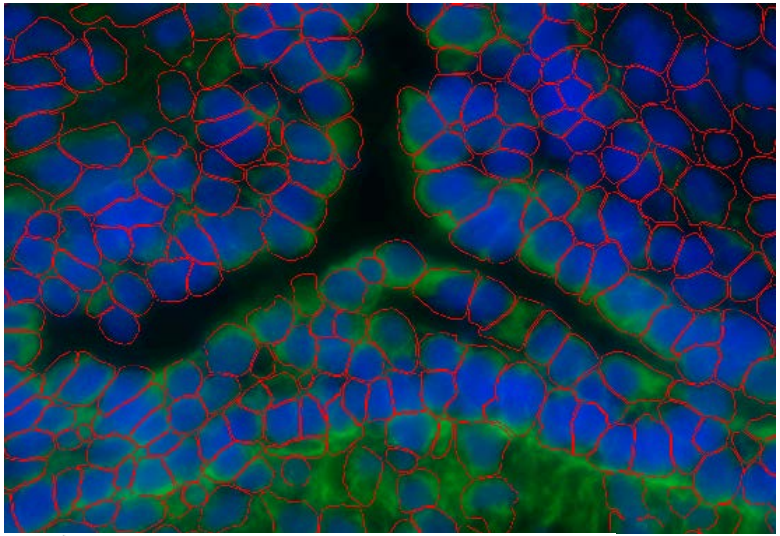


DAPI chanel

Cell boundaries chanel

Data acquisition

Cell segmentation



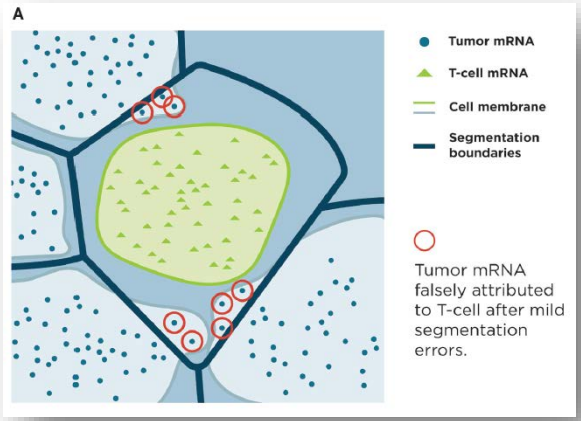
Article | Published: 14 December 2020

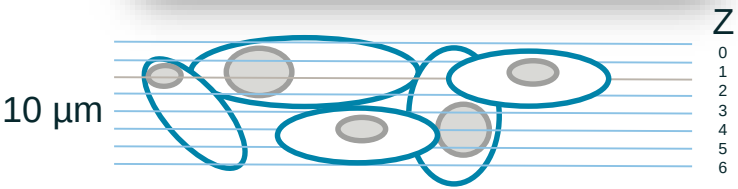
Cellpose: a generalist algorithm for cellular segmentation

Carsen Stringer, Tim Wang, Michalis Michaelos & Marius Pachitariu

Nature Methods **18**, 100–106 (2021) | Cite this article

Cell segmentation is crucial to ensure cell x gene matrix purity for good subsequent biology

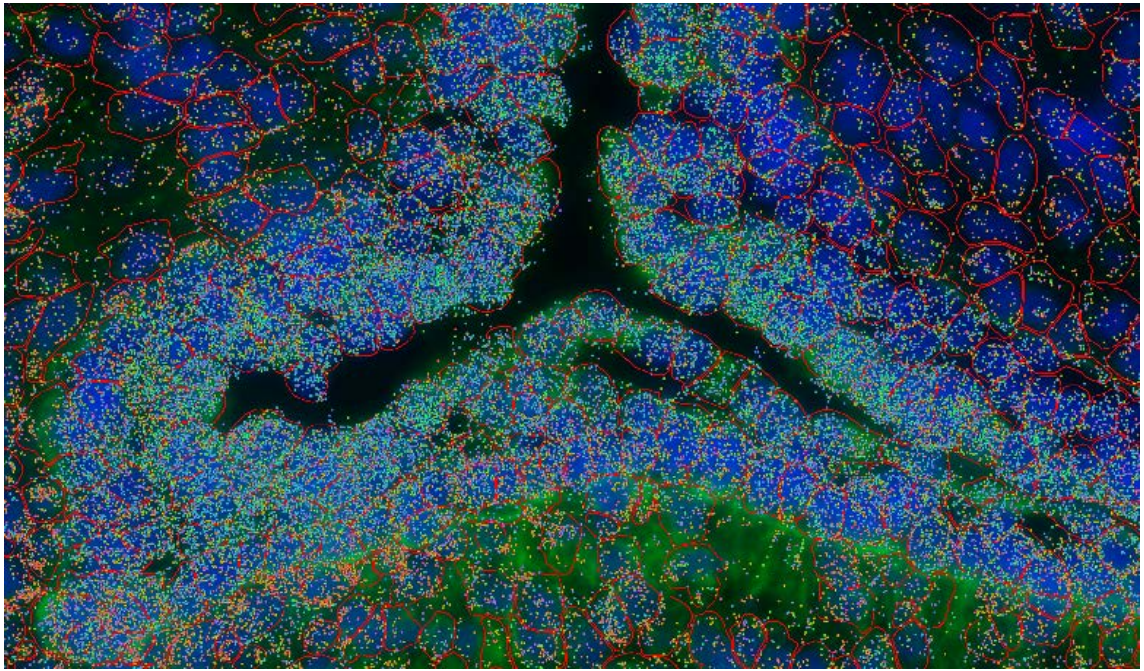




3D segmentation required, actually not used, 2D segmentation per Z then harmonizing and summing the detected transcripts for all Z into the harmonized segmentationmask (nuclei of full cell)

Raw data

Cell x genematrix

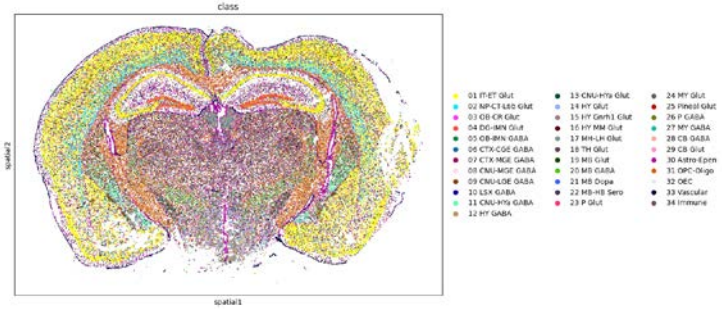


Gene-level matrix

100k's cells

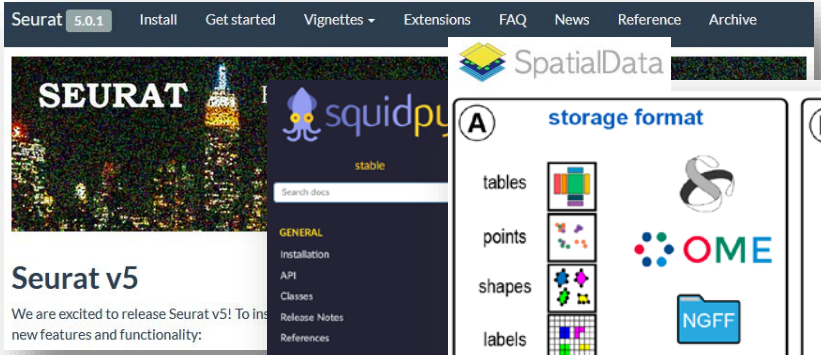
Clta	5	4	6	7	1	4	3	9	5	5	.	1	4	3	6	4	.	7	5	2	4
Myl6	5	2	5	1	2	4	13	2	5	2	4	4	1	4	8	4	2	3	3	1	1
Pkm	3	2	1	.	.	8	3	9	4	10	5	1	.	4	12	5	1	3	9	1	7
Tecr	3	1	5	2	4	6	5	9	6	3	2	2	1	4	7	1	1	1	3	2	8
Meis2	8	29	3	.	24	6	6	21	25	16	6	1	.	2	29	.	.	6	18	.	10
...																					

1.000 Genes



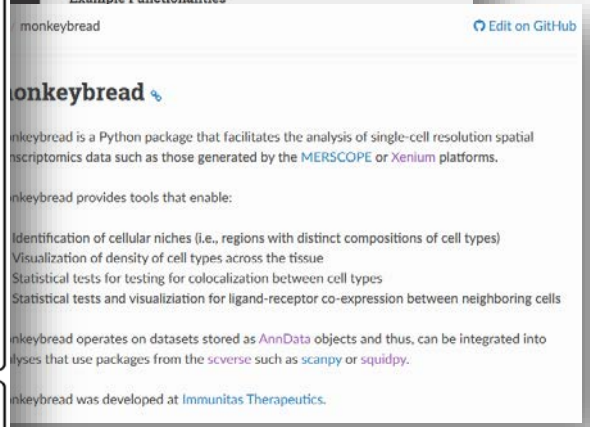
Statistical data analysis

Standardized workflows + packages development



Satija'slab, NYGC

Theis'slab, h



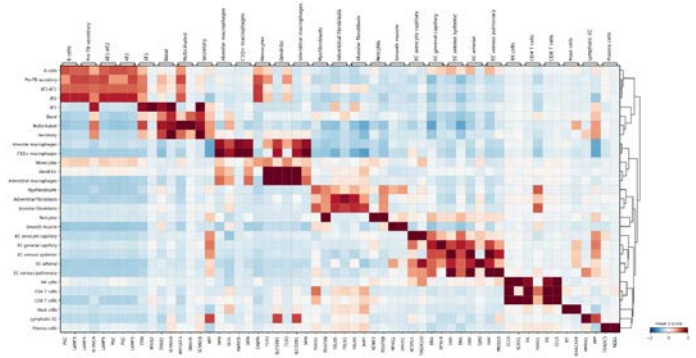
cs

Scverse ecosystem, Oliver Stegle & Fabian J. Theis

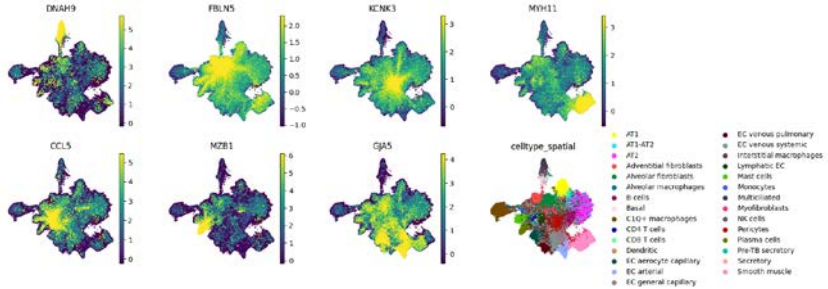
Single -cell data analysis

Scanpy and Squidpy toolkits

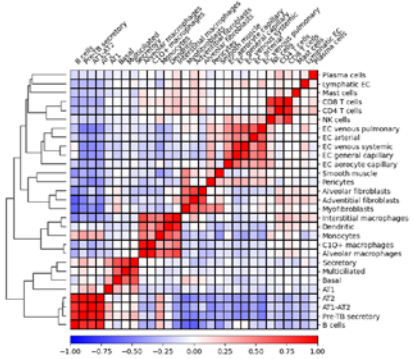
Gene marker detection, manual or automatic cell type identification



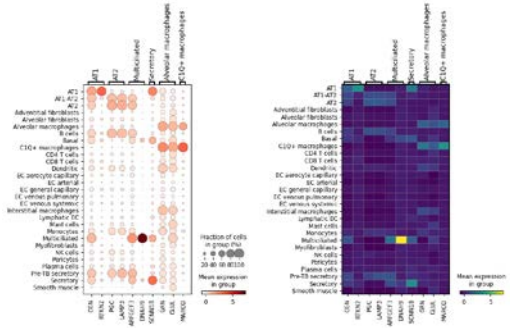
Batch effect correction, sample integration, cell type labeling transfer from single-cell references dataset



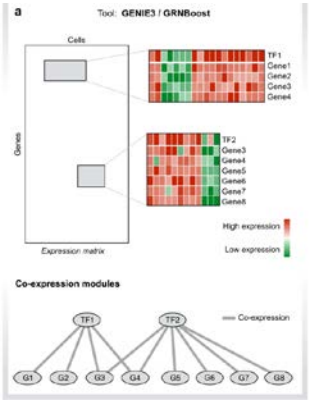
Cell type correlation



Differential expression analysis
Gene set functional enrichment

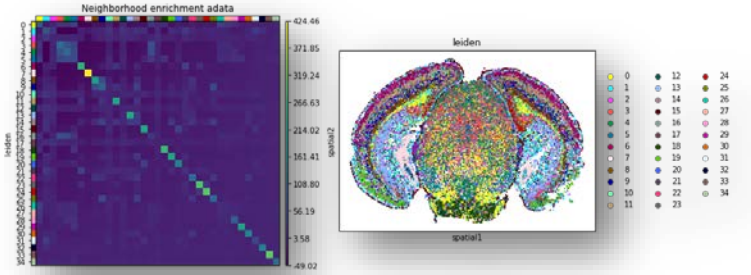


Transcription Regulatory Network



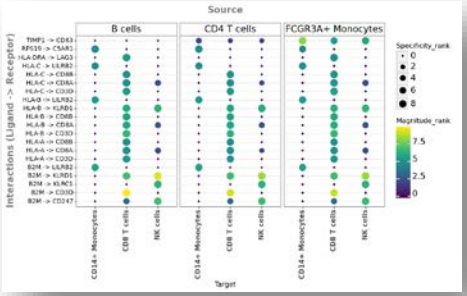
Single -cell data analysis including spatial resolution

New vast area for computational biologists (just like single-cell 5 years ago)



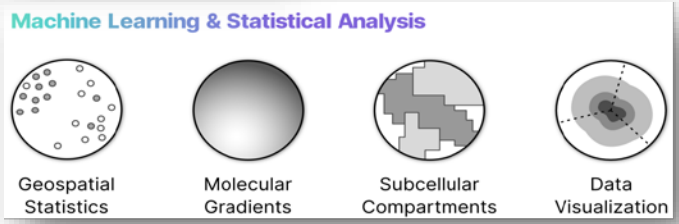
Neighbors enrichment analysis

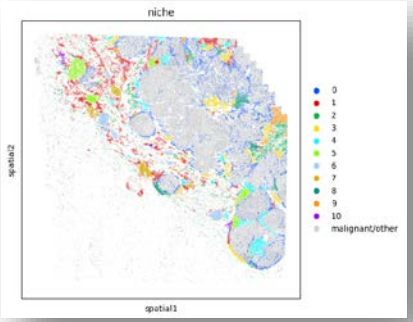
Test if cells belonging to 2 clusters are close to each other more often than expected (co-occurrence probability)



Cell-cell communication Ligand-Receptor analysis

- Need to be in gene panel or inferred
- CellPhoneDB [Efremova et al., 2020]
- Omnipath [Türei et al., 2016].



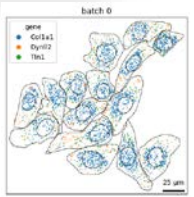


Cellular niches analysis

for each cell, we count the number of neighbors that are of each cell type thus forming a “neighborhood profile” vector of length C, where C is the number of cell types. We then cluster all neighborhood profiles and call each cluster a “niche”.

Sub-cellular exploration

Bento is a Python toolkit for performing subcellular analysis of spatial transcriptomics



Acknowledgments

Institut de Pharmacologie Moléculaire et Cellulaire

Pascal Barbry's Lab (IPMC, CNRS, France)

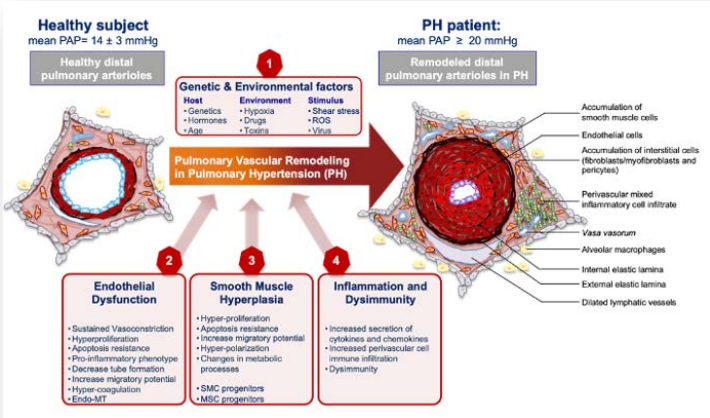
- Virginie Magnone
- Géradine Rios
- Marie Couralet
- **Valentine Freschi**
- **Marie-Jeanne Arguel**

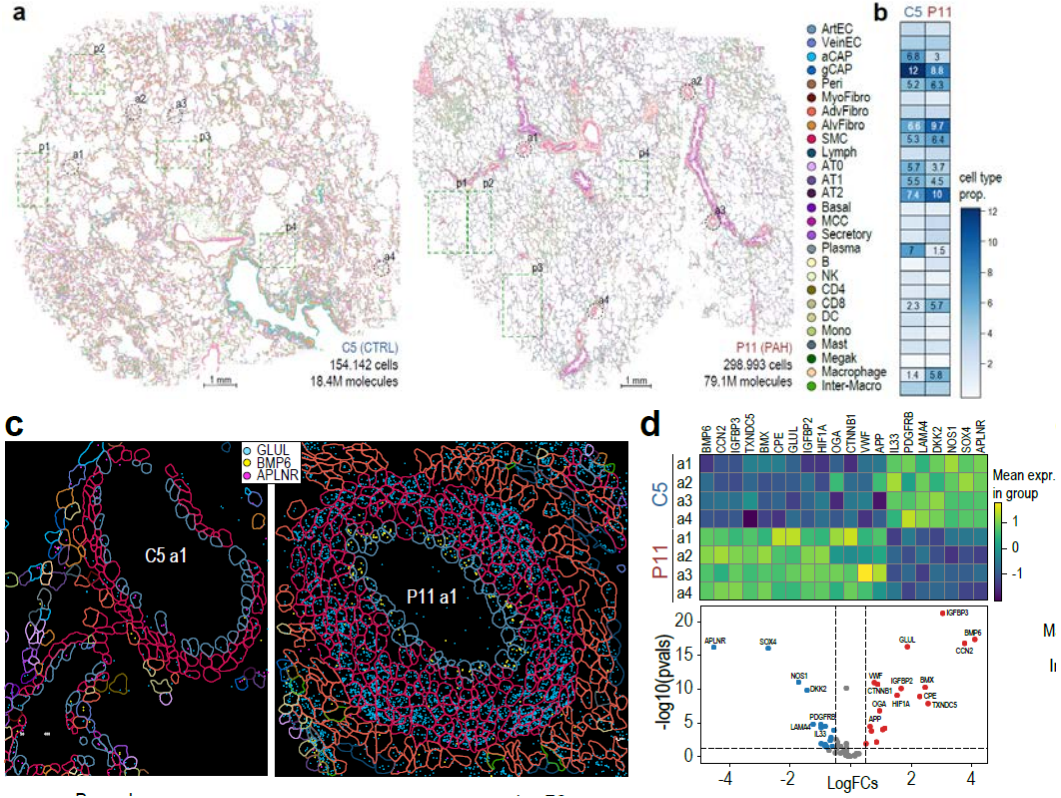
CoBiODA IPMC bioinformatics

- Kévin Lebrigand
- **Morgane Fierville**
- Marin Truchi
- Eamon McAndrew

PAH : Pulmonary Arterial Hypertension

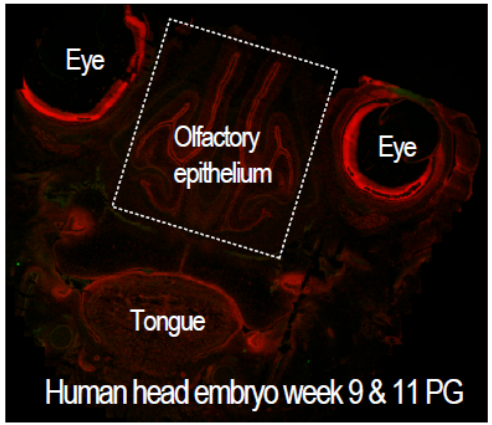
A rare vascular disorder

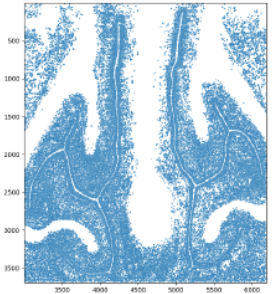




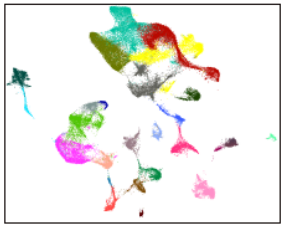
HuDeCa project

human fetal nosefrom 7 to 12 post-conceptionalweeks (PCW) at singlecell resolution

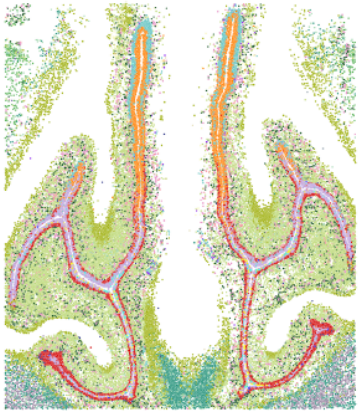




Extraction
Rotation
Segmentation

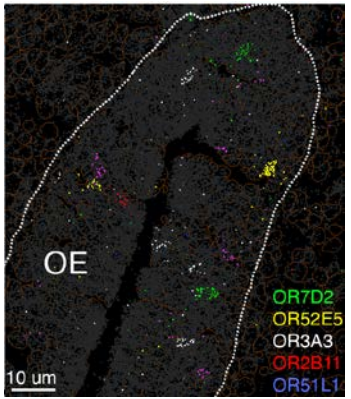


Automatic classification using snRNA-seq



Cell types
- Cartilage
- Stromal
- Lymphatic EC
- Vascular EC
- Pericytes
- Respiratory HBCs
- Olfactory HBCs
- Duct/MUC
- Multiciliated
- Deuterosomal
- Sustentaculars
- GBCs
- Early OSNs
- Excitatory neurons
- Inhibitory neurons
- GnRH neurons

Spatial data exploration



MERSCOPE raw data

Standard Merscope output files

```
000-giacobini/000-data/202304281610_20230428-HuDeCa-Giacobini-PGW9-2-3A_VMSC06001/region_0:
total 13G
-rw-r--r-- 1 lebrigand solid  11G  3 mai    2023 202304281610_20230428-HuDeCa-Giacobini-PGW9-2-3A_VMSC06001_region_0.vzg
-rw-r--r-- 1 lebrigand solid 447M  3 mai    2023 cell_boundaries.parquet
-rw-r--r-- 1 lebrigand solid  83M  3 mai    2023 cell_by_gene.csv
-rw-r--r-- 1 lebrigand solid  46M  3 mai    2023 cell_metadata.csv
-rw-r--r-- 1 lebrigand solid 2,1G  3 mai    2023 detected_transcripts.csv
drwxr-xr-x 2 lebrigand solid 4,0K  3 mai    2023 images
-rw-r--r-- 1 lebrigand solid 855K  3 mai    2023 summary.png
```

```
8,9K  3 mai    2023 manifest.json
 227  3 mai    2023 micron_to_mosaic_pixel_transform.csv
6,6G  3 mai    2023 mosaic_Cellbound2_z2.tif
6,6G  3 mai    2023 mosaic_DAPI_z2.tif
```

0,5-2 Tb

```
1  sdata.shapes['P11_region_0_polygons']
✓ 0.0s
```

	geometry
3613852200009100002	POLYGON ((4702.379 302.065, 4704.864 304.790, ...
3613852200009100005	POLYGON ((4709.067 304.517, 4709.887 305.673, ...
3613852200009100006	POLYGON ((4767.177 301.994, 4767.475 303.202, ...
3613852200009100007	POLYGON ((4723.735 306.300, 4723.939 307.090, ...
3613852200009100008	POLYGON ((4764.231 307.210, 4764.805 312.376, ...

```
1  sdata.table.layers['counts']
✓ 0.0s
array([[0, 1, 1, ..., 1, 0, 0],
       [0, 1, 0, ..., 2, 0, 0],
       [0, 3, 4, ..., 0, 0, 0],
       ...,
       [0, 0, 0, ..., 0, 0, 0],
       [0, 0, 0, ..., 0, 0, 0],
       [0, 2, 6, ..., 4, 0, 0]])
```

```
1  sdata.table.obs
✓ 0.1s
```

	fov	volume	center_x	center_y	min_x	min_y	max_x	max_y	anisotropy
3613852200009100002	NaN	841.761444	4707.689677	298.299088	4702.379316	292.031852	4713.603506	305.207677	1.990932
3613852200009100005	NaN	762.740701	4713.532399	299.961526	4709.066587	293.764119	4718.738288	306.280330	1.812676
3613852200009100006	NaN	1243.613805	4774.139602	300.963756	4767.176971	293.701627	4780.340526	307.566412	1.069031
3613852200009100007	NaN	391.378917	4727.286292	305.141647	4723.734863	301.881750	4730.984004	308.291885	1.194130
3613852200009100008	NaN	680.804739	4768.106763	308.764922	4764.065981	303.609728	4772.968368	313.933532	1.204513

```
1  sdata.points['P11_region_0_transcripts'].compute()
✓ 15.7s
```

	x	y	gene	Unnamed: 0	global_z	transcript_id	fov	barcode_id	cell_id
0	76.288345	6615.9116	CFTR	138	0.0	ENST00000003084	0	0	-1
1	-0.706320	6656.7720	CFTR	480	0.0	ENST00000003084	0	0	-1
2	56.071490	6741.6104	CFTR	1284	0.0	ENST00000003084	0	0	3613852200420100110
3	-6.401468	6764.1587	CFTR	1479	0.0	ENST00000003084	0	0	-1
4	12.010611	6604.1650	CFTR	1808	1.0	ENST00000003084	0	0	3613852200390100641