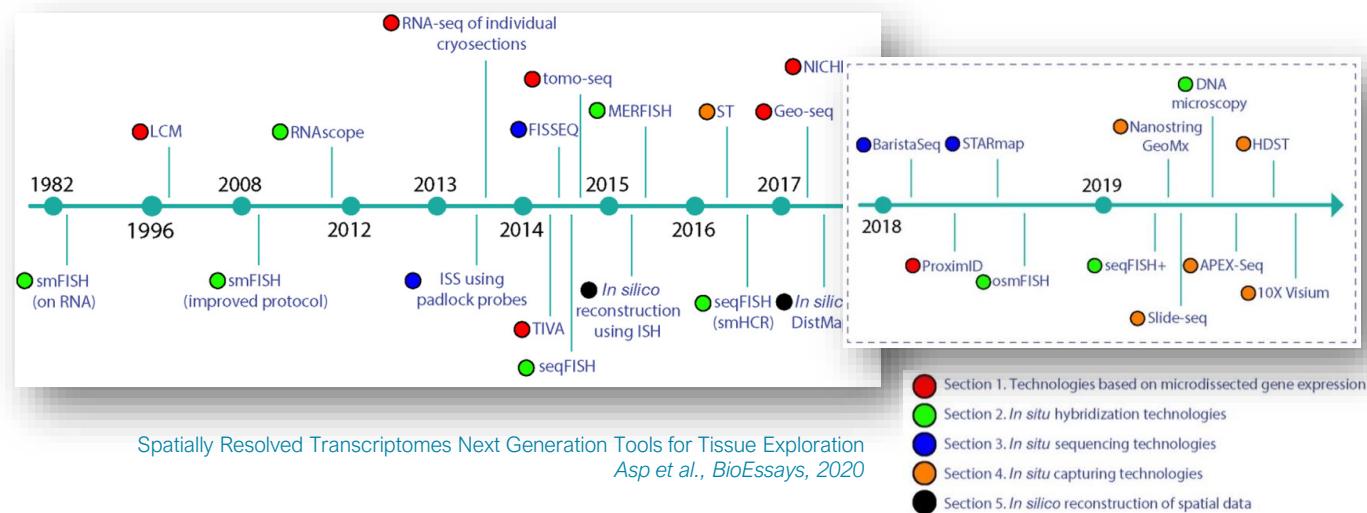


Spatial Transcriptomics approaches

Historical timeline

- Spatial transcriptomics aims at directly visualize gene expression in their original environment,
- It tackles 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,



Spatial isoform Transcriptomics (SiT)

Nucleic Acids Research, 2023

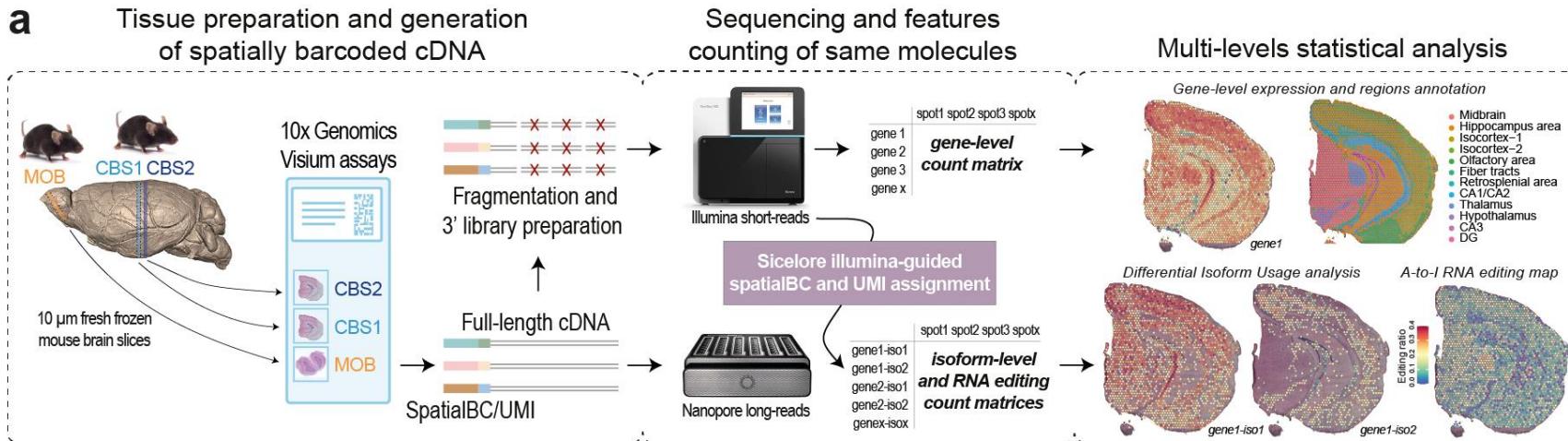
The spatial landscape of gene expression isoforms in tissue sections

Kevin Lebrigand, Joseph Bergenstråle, Kim Thrane, Annelie Mollbrink, Konstantinos Meletis, Pascal Barbuy , Rainer Waldmann, Joakim Lundeberg Author Notes

Nucleic Acids Research, Volume 51, Issue 8, 8 May 2023, Page e47, <https://doi.org/10.1093/nar/gkad169>

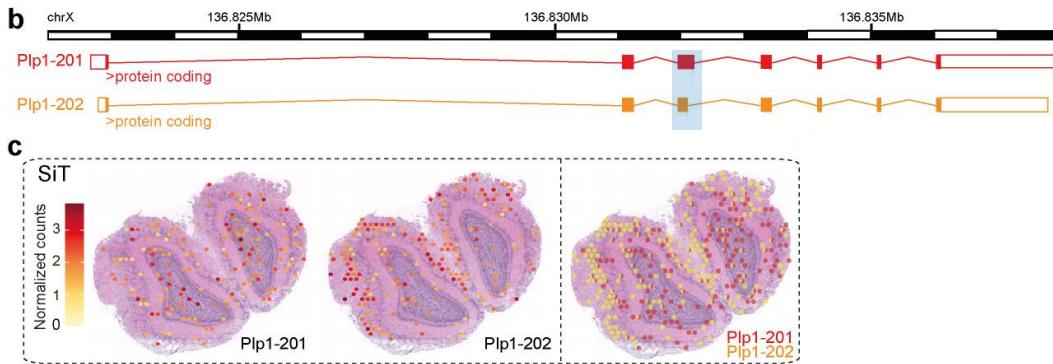
Published: 17 March 2023 Article history ▾

a



SiT reveals specific splicing pattern across MOB regions

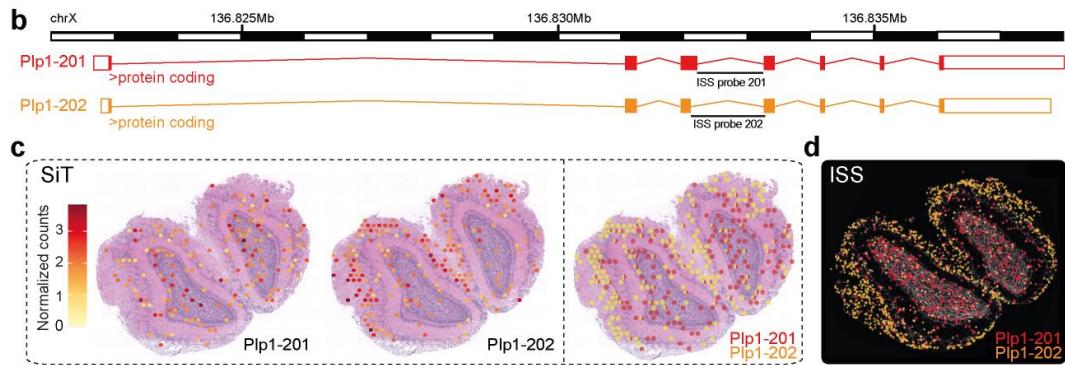
Plp1 Differential Transcript Usage (DTU)



Proteolipid Protein 1 (Plp1) is a gene involved in severe pathologies associated with CNS dysmyelination

SiT reveals specific splicing pattern across MOB regions

Plp1 Differential Transcript Usage (DTU)



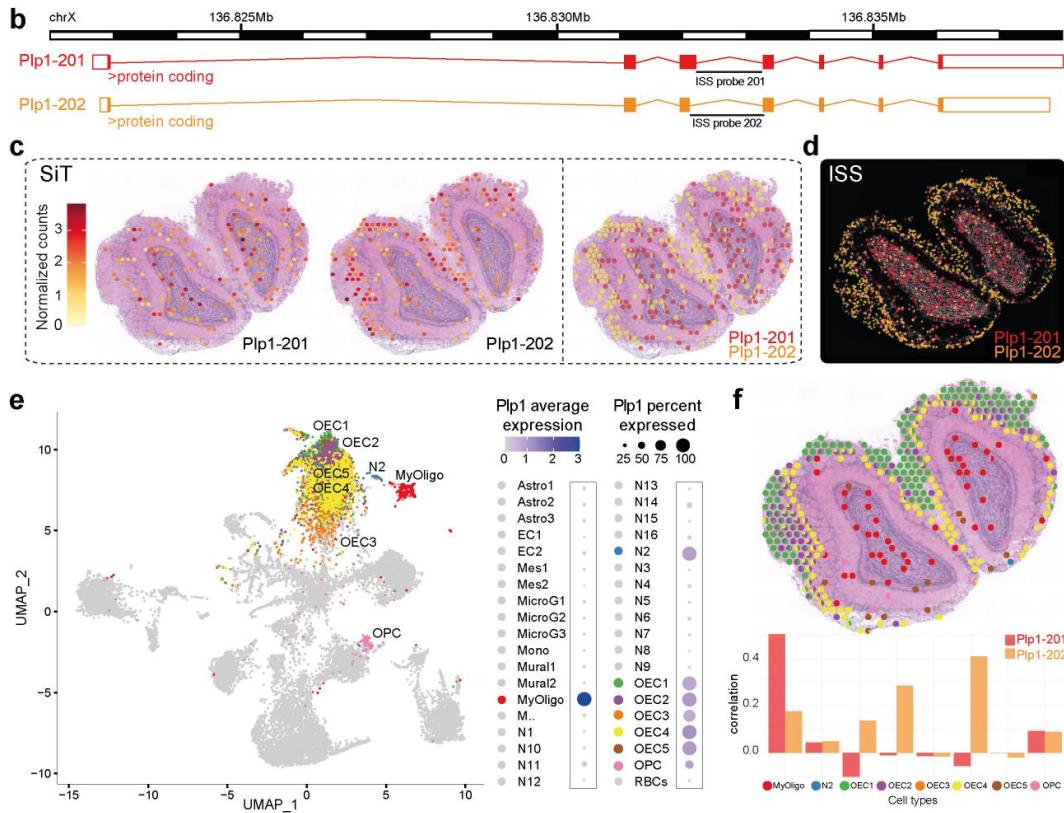
Proteolipid Protein 1 (Plp1) is a gene involved in severe pathologies associated with CNS dysmyelination



In Situ Sequencing Data

SiT reveals specific splicing pattern across MOB regions

Cell type deconvolution using single cell external dataset (Tepe et al., 2018)



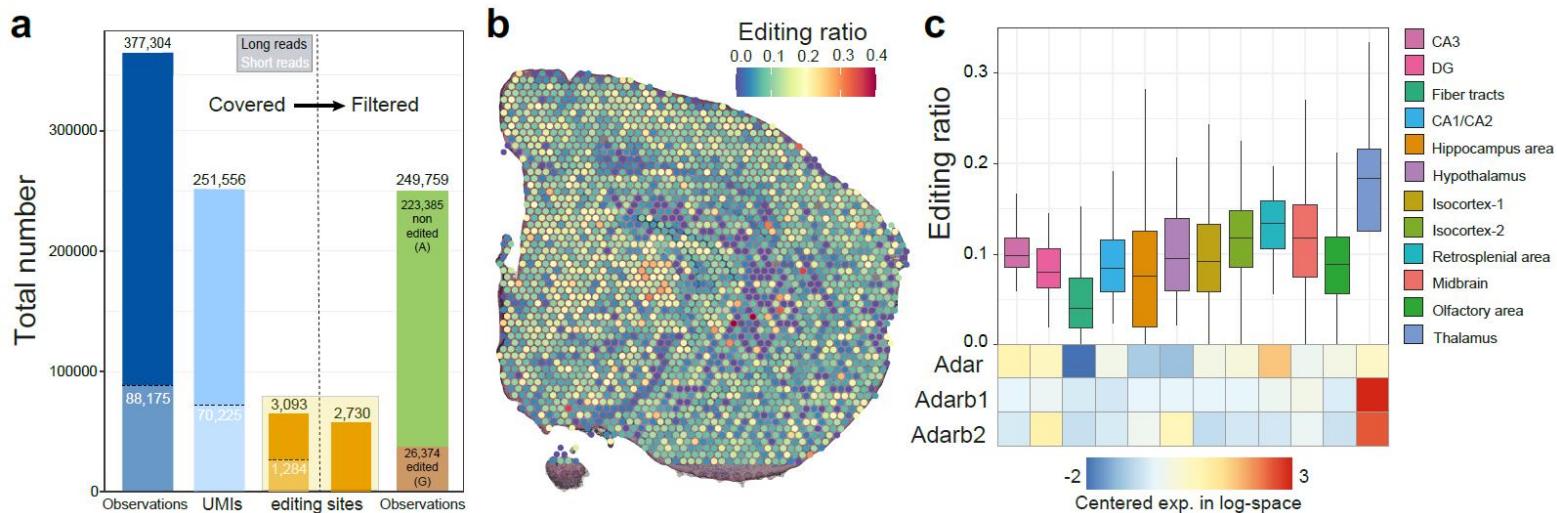
Proteolipid Protein 1 (Plp1) is a gene involved in severe pathologies associated with CNS dysmyelination

Spatial spot deconvolution of prominent *Plp1* expresser cell types. Correlation Deconvolution score / *Plp1* isoforms expression correlation shows that *Plp1* is predominantly expressed as *Plp1-202* by olfactory ensheathing cells (OEC) in the ONL and as *Plp1-201* isoform by myelinating-oligodendrocytes (MyOligo) in the GCL.

SiT reveals full-length sequence heterogeneity (CBS)

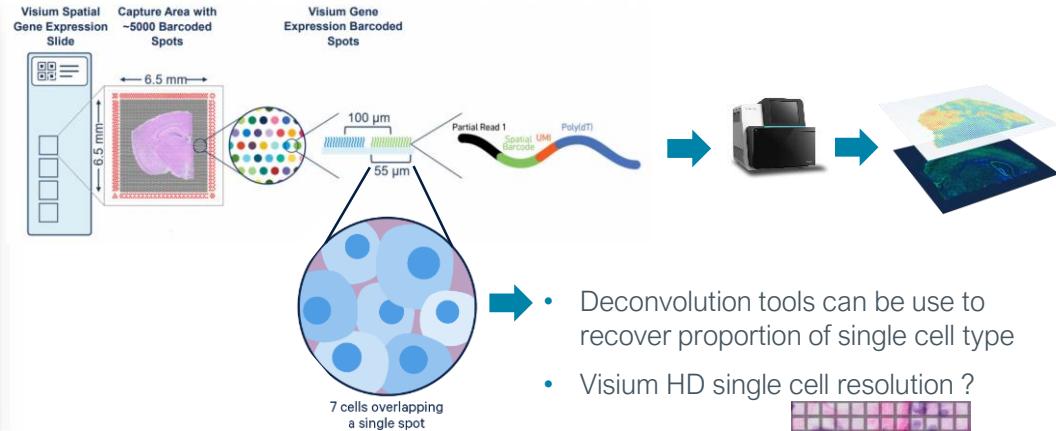
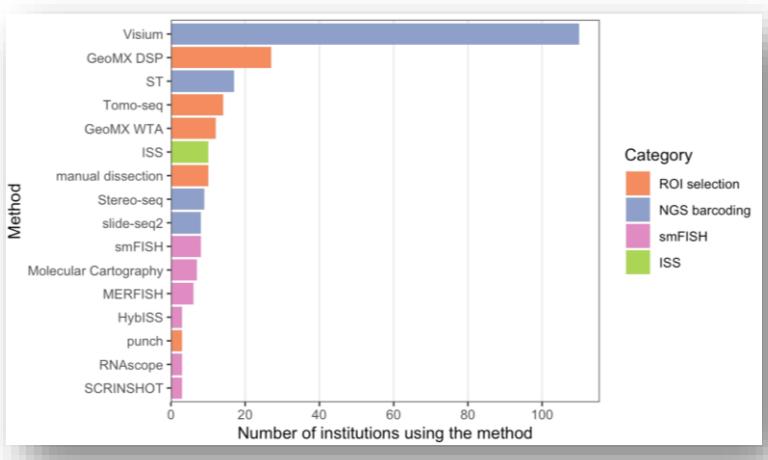
Global A-to-I RNA editing spatial map

- Exploration of 5,817 A-to-I RNA editing sites described in the literature (Ramaswami et al., 2013 (RADAR), Licht et al., 2019)
- Long read high confidence call thresholding, looking at agreement between long and short read base calls for 88,175 shared UMIs
 - number of reads per UMI ≥ 3
 - consensus Phred score QV ≥ 6

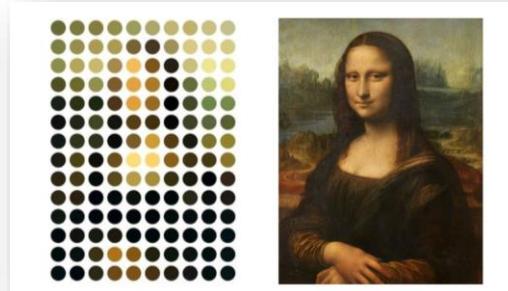
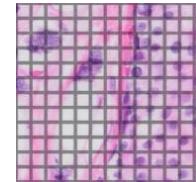


Spatial transcriptomics (2017-2022)

Visium is widely adopted by academics



- Deconvolution tools can be used to recover proportion of single cell type
- Visium HD single cell resolution ?

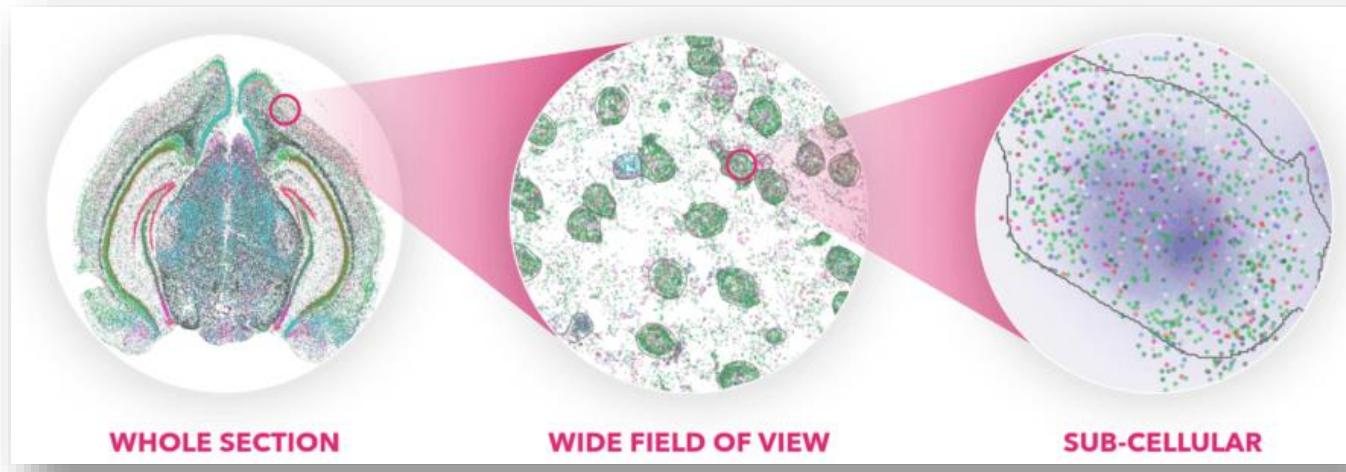


But is not the ideal readout for spatial biology
(Akoya credit rough caricature)

Spatial transcriptomics imaging (2023)

No more sequencing for direct single-cell resolution

- Lower gene targets (from whole transcriptome to 500-1,000 genes)
- Higher sensitivity (from ~6% to 30-80%)
- Same imaging area (40 to 236 mm²)
- Higher resolution (from 55 µm to subcellular)



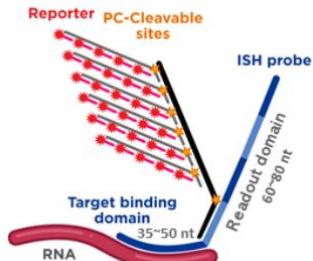
Spatial imaging technologies (2023)

No more sequencing for direct single-cell resolution



Nanostring CosMx
ISH-based

- Limited availability of system
- 960 targets
- Sensitivity : 30%
- Resolution: 200 nm
- Imaging area: 16 mm²



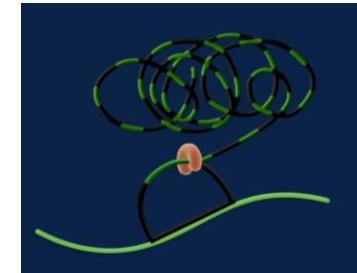
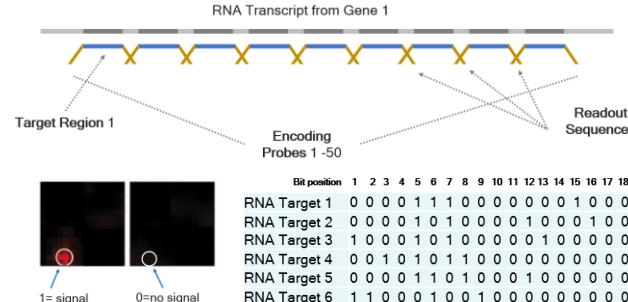
Vizgen Merscope
Merfish

- Available
- 500 targets
- Sensitivity : 30-80%
- Resolution: 100 nm
- Imaging area: 100 mm²



10xGenomics Xenium
Cartana ISS, padlock probes / RCA

- Available
- 400 targets
- Sensitivity : 30%
- Resolution: 200 nm
- Imaging area: 236 mm²



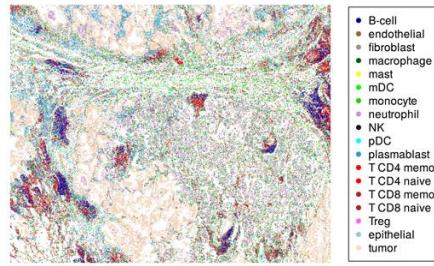
Spatial imaging technologies comparison

Compare available datasets

Vizgen Merscope

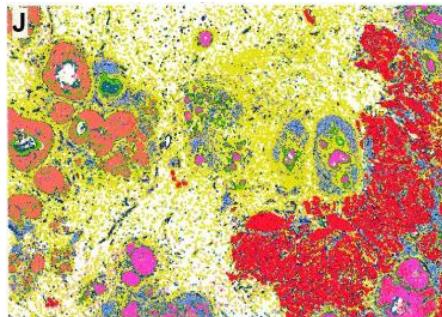
- Xiaowei Zhuang's lab merfish publications
 - Chen et al., Science (2015)
 - Moffitt et al., PNAS (2016), Science (2018)
 - Emanuel G et al., Nature Methods (2017)
 - Xia C. et al., PNAS (2019), Scientific Reports (2019)
 - Zhang M. et al., Nature (2021)
 - ...
- Internal data release program
 - Human Immuno-oncology (**breast**, colon, **lung**, liver, skin, prostate, uterine and ovarian) 500 genes, >4 billion transcripts, 9 million cells
 - Mouse Liver Map (347genes)
 - Mouse brain Receptor Map (483 genes)
- External labs publications
 - Dixon E. et al., Kidney Int. (2022): Kidney
 - Wang et al., Nat. Neuro. (2022): Mouse olfactory Glomerular map
 - Stogsdill et al., Nature (2022): Neocortex microglia
 - ...

Nanostring CosMx



- Release date: 11/2021
- FFPE Human NSCLC (Lung)
- 960 gene targets
- 8 sections for 800k cells
- Imaging area: 8 x 16 mm²
- 259,604,214 transcripts
- Mean transcripts/cell: 265

10xGenomics Xenium



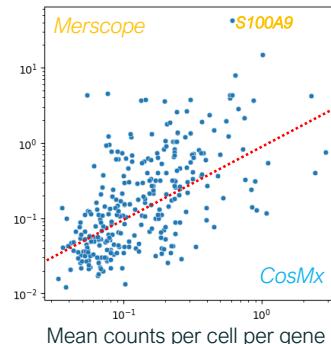
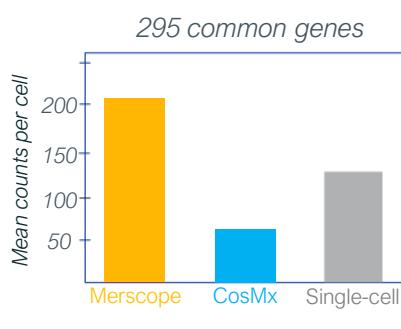
- Release date: 10/2022
- FFPE Human Breast cancer
- 313 gene targets
- 167,885 cells,
- 36,944,521 transcripts
- Imaging area: 40 mm²
- Mean transcripts/cell: 193

Spatial imaging technologies comparison

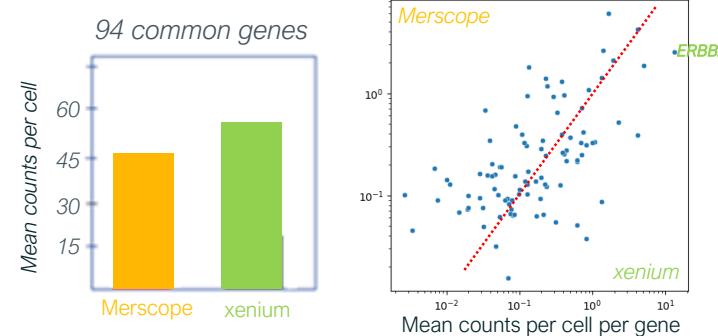
Compare available datasets: Lung and Breast cancer samples



FFPE Human Lung Cancer	Merscope	CosMx
Total cells	353 k (x4)	92 k
Detected transcripts	107 M (x4)	26 M
Gene targets	500	960 (x2)
Total RPKM	9,204	61,680 (x6)
Mean transcripts/cell	302	284



FFPE Human Breast Cancer	Merscope	Xenium
Total cells	713 k (x4)	168 k
Detected transcripts	353 M (x10)	32 M
Gene targets	500	313
Total RPKM	9,909	7,912
Mean transcripts/cell	495	193



<https://vizgen.com/wp-content/uploads/2022/12/Vizgen-Spatial-Genomics-Data-Quality-eBook-1.pdf>

MERSCOPE @ UCAGenomiX (Nice-Sophia-Antipolis)

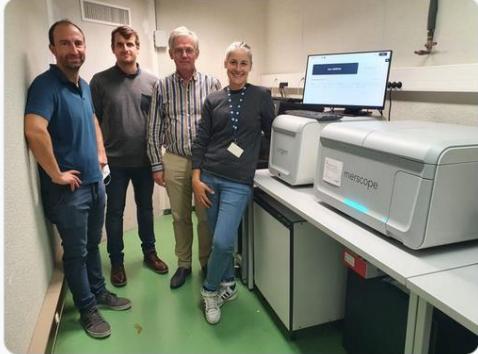
October 2022

↪ William Amoyal Retweeted
Pascal Barbuy @pbarbuy · Oct 12

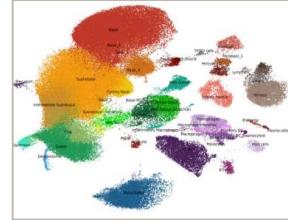
Happy to announce the installation of our first Merscope at @UCAGenomix.
Many thanks to @vizgen_inc people for amazing work and interactions.

Great spatial transcriptomics work to come

@fr_genomics @discovAIR_HCA @3IAcotedaduz @IPMC_sophia @CNRS
@Univ_CotedAzur @CanceropolePACA



- Human Lung Cell Atlas (CZI)  
Discovering the Cellular Landscape of the Airways and Lung Tissue



- 12 control / 2 IPF / 10 COPD patients
- 415,764 cells (117 samples)
- 48 cell types



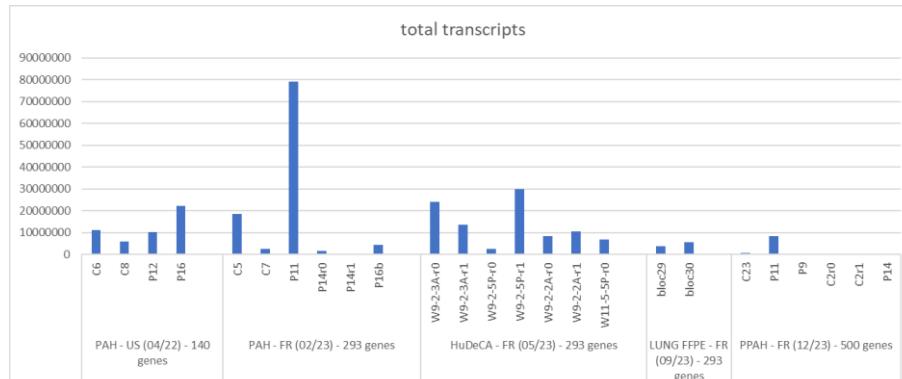
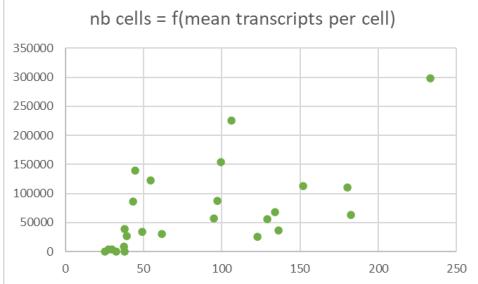
- HuDeCA - Paolo Giacobini, Inserm U837, Lille: Neuroendocrinology 
- Pulmonary Arterial Hypertension (Christophe Guignabert, Paris-Saclay)

MERSCOPE @ UCA GenomiX (Nice-Sophia-Antipolis)

Statistics

	PAH - US (04/22) - 140 genes				PAH - FR (02/23) - 293 genes					HuDeCA - FR (05/23) - 293 genes					LUNG FFPE - FR (09/23) - 293 genes			PPAH - FR (12/23) - 500 genes								
	C6	C8	P12	P16	C5	C7	P11	P14r0	P14r1	P16b	W9-2-3A-r0	W9-2-3A-r1	W9-2-5P-r0	W9-2-5P-r1	W9-2-2A-r0	W9-2-2A-r1	W11-5P-r0	bloc29	bloc30	C23	P11	P9	C2r0	C2r1	P14	
raw	total cells	121915	53054	213366	153498	277310	145253	343860	110709	26049	58206	132997	73755	58313	317814	65924	80178	146212	272885	244173	156556	228057	187809	20723	33392	141600
	total transcripts	11318606	6025439	10101918	22216691	18422575	2473759	79140205	1673178	478830	4285888	23985874	13623063	2497236	29878124	8329861	10667555	6858631	3852467	5542811	716078	8257288	449882	16534	17177	45247
	% in cells	76,9	84,4	72,4	79	86,2	84,9	88,3	83,3	83,6	75,073	84	84,6	83,6	82,5	87,1	85,6	84,2	62,7	84,6	85,6	86,3	87,2	74,2	69,3	68,7
	mean transcripts per cell	71,4	95,9	34,3	114,3	57,2	14,4	203,1	12,6	15,4	55,3	151,6	156,3	35,8	77,6	110,1	113,9	39,4	8,9	19,2	4	31,2	2,1	0,6	0,36	0,22
	median transcripts per cell	53	64	25	74	28	4	178	5	9	11	96	100	21	45	82	85	9	3	11	2	25	0	0	0	0
min>20	total cells	87902	36706	121897	113332	154142	33603	298993	26200	8205	25363	110857	62647	30618	225327	55501	67363	57066	38581	86636	4077	139703	3079	5	14	12
	mean transcripts per cell	97	136,1	54,4	152,2	99,6	49,3	233,2	39,4	37,3	122,9	180,2	182,6	61,6	106,3	129,2	133,9	95	38	43,4	27,6	44,6	30	32,6	38,1	25,2
	median transcripts per cell	79	111	44	120	78	40	201	33	32	85	124	123	46	72	99	103	67	32	36	25	37	26	36	37	25,5
	filtered cells	27,9	30,8	42,9	26,2	44,4	76,9	13,0	76,3	68,5	56,4	16,6	15,1	47,5	29,1	15,8	16,0	61,0	85,9	64,5	97,4	38,7	98,4	100,0	100,0	100,0
	filtered counts	1,6	1,5	6,8	1,3	2,8	17,6	0,1	21,7	19,9	2,4	0,7	0,7	8,1	2,4	1,1	1,1	4,9	25,0	16,7	71,7	10,7	67,1	74,2	66,9	68,2

	raw	filtered
slices	20	20
total cells	3,6M	1,8M
total transcripts	270M	220M
avg transcripts /cell	60	122



MERSCOPE @ UCAGenomiX (Nice-Sophia-Antipolis)

Data analysis

The image displays a collage of screenshots from several bioinformatics platforms:

- Seurat 5.0.1**: A user interface showing a heatmap of gene expression data and a sidebar with links to "GENERAL", "API", "GALLERY", and "Examples".
- Squidpy**: A storage format for spatial data, showing icons for tables, points, shapes, labels, and images, along with logos for OME and NGFF.
- SpatialData**: A Python library for spatially-aligned datasets, illustrating the process of reading data, applying transforms (translate, scale, rotate, chain), performing spatial queries, and aggregating observations.
- monkeybread**: A package for spatial data analysis, featuring a brain tissue image and a brief description of its capabilities.
- Giotto**: A spatial transcriptomic and proteomic analysis platform, showing a detailed description of its features and four representative figures (Figure 2A, 2B, 2C, 2D).
- cobioda / scisipy**: A GitHub repository page for the scisipy project, showing the repository's structure and a link to the public version.

<https://github.com/cobioda/scisipy/>

MERSCOPE @ UCAGenomiX (Nice-Sophia-Antipolis)

Cell label transfert

SCANVI cell typing

```
htap = sc.read_h5ad('htap_230227.h5ad')
htap.layers['counts'] = htap.raw.X.copy()
del htap.var['features']

adata.var.index = adata.var.index.str.upper() # Lower
htap.var.index = htap.var.index.str.upper()

# Select shared gene panel genes only
genes_Vizgen = adata.var.index
genes_10x = htap.var.index
genes_shared = genes_Vizgen.intersection(genes_10x) # List of shared genes

vizgen = adata[:, genes_Vizgen.isin(genes_shared)].copy()
htap = htab[:, genes_10x.isin(genes_shared)]
```

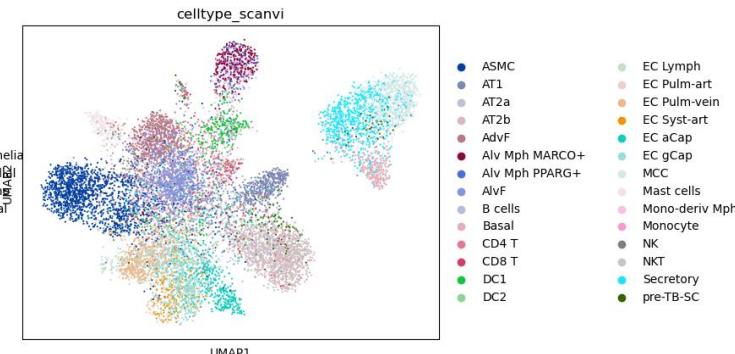
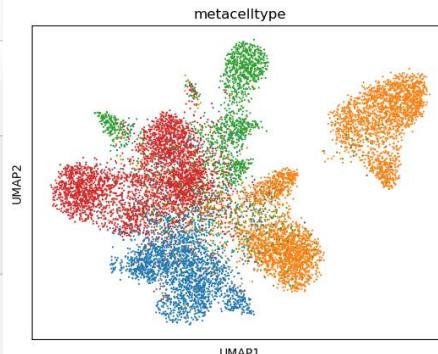
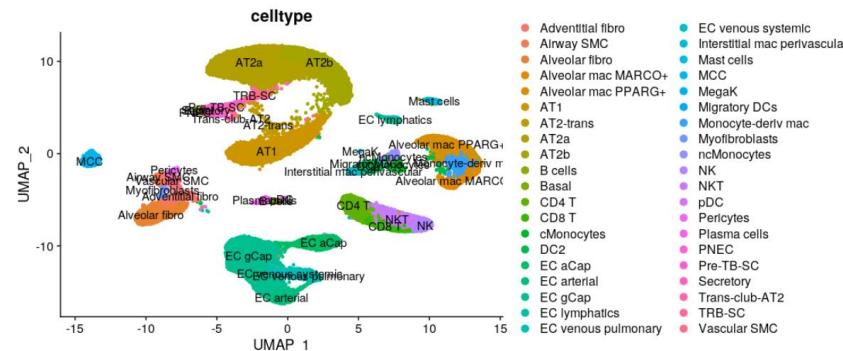
```
print('Data loaded and preprocessed.')
print(len(genes_shared), "common genes")

missed = list(set(genes_Vizgen) - set(genes_10x))
print("gene missed = ", missed)
```

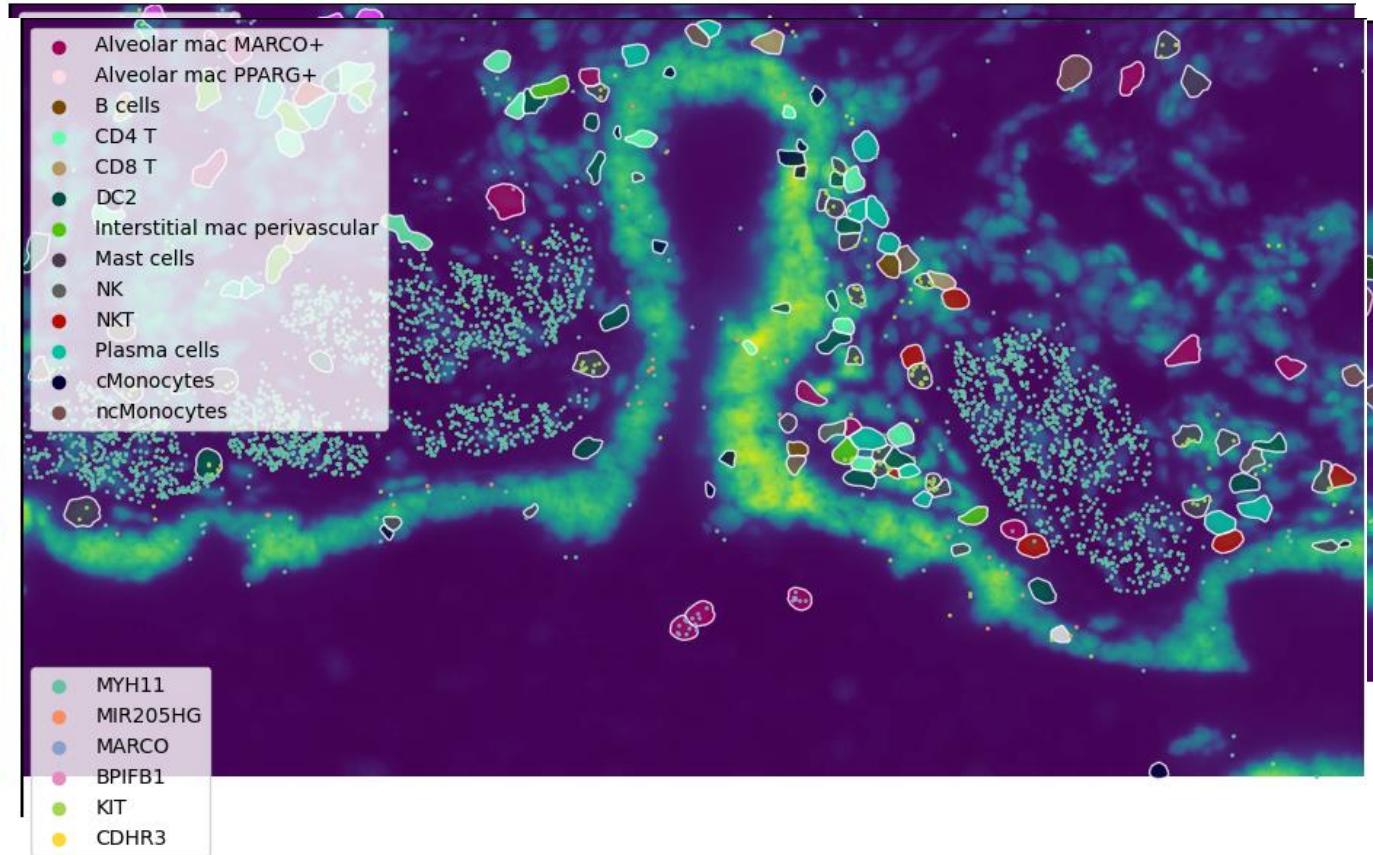
```
Data loaded and preprocessed.  
293 common genes  
gene missed = []
```

```
print("vizgen dims:",vizgen.shape)
print("htap dims : ",htap.shape)
sc.pp.filter_cells(htap, min_counts=10)
print("htap dims : ",htap.shape)
htap = htap[htap.obs.group == "PAH"]
print("htap dims : ",htap.shape)
```

```
vizgen dims: (12244, 293)  
htap dims : (77068, 293)  
htap dims : (76434, 293)  
htap dims : (34909, 293)
```



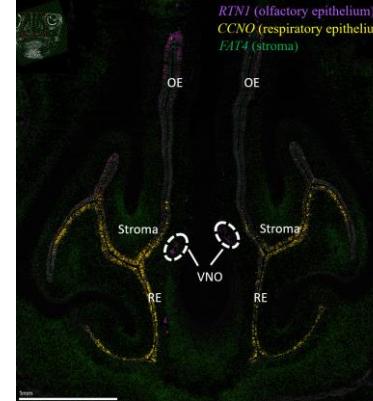
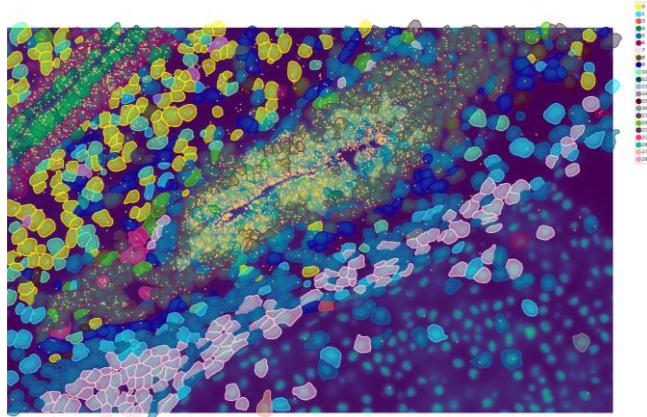
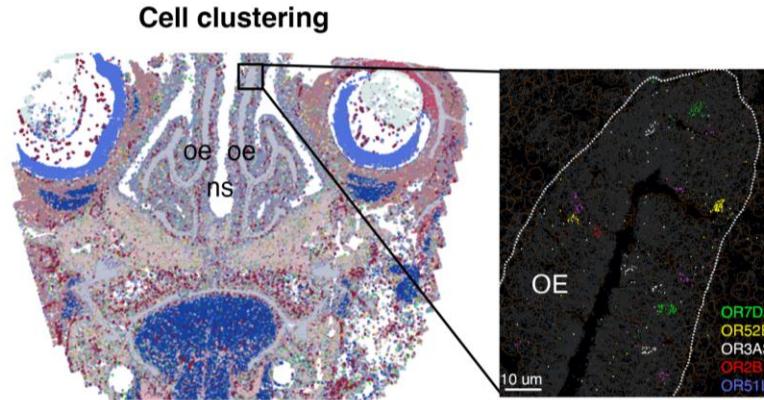
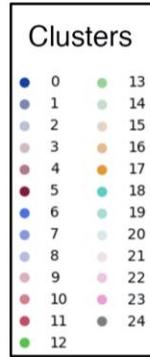
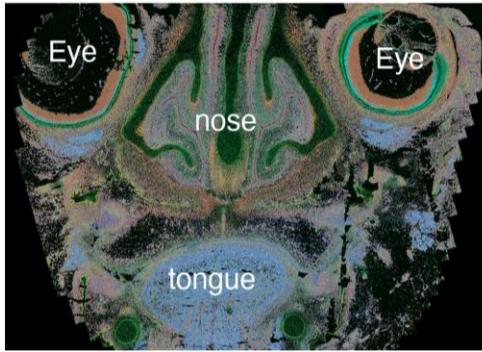
Human Cell Atlas Lung project



HuDeCa project

human fetal nose from 7 to 12 post-conceptional weeks (PCW) at single-cell resolution

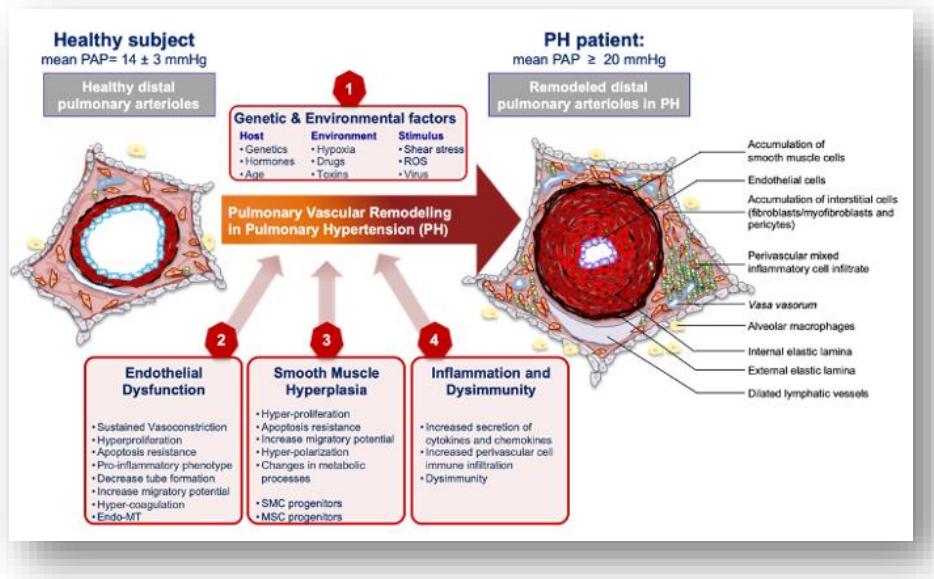
MERSCOPE-300 probes-gene panel



PAH : Pulmonary Arterial Hypertension

A rare vascular disorder

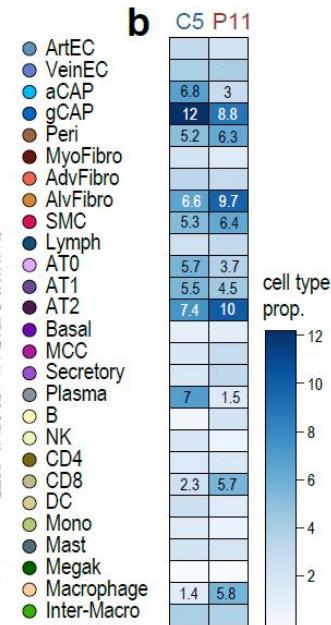
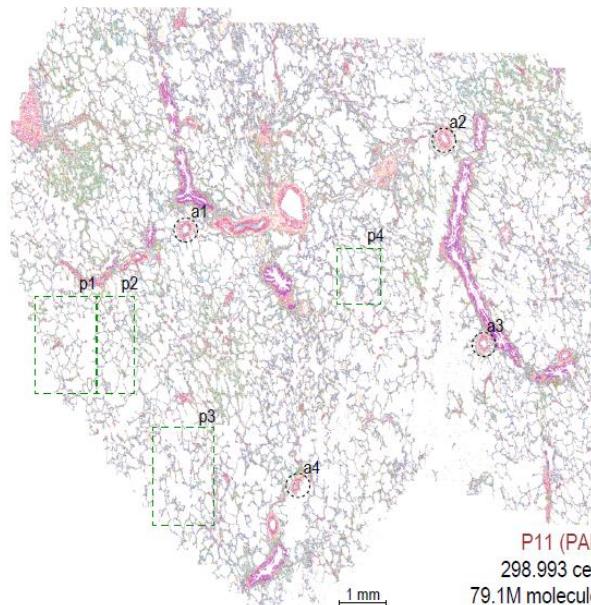
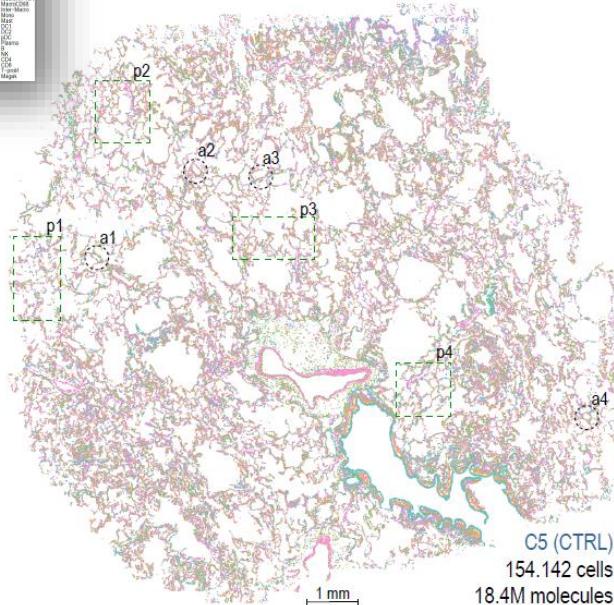
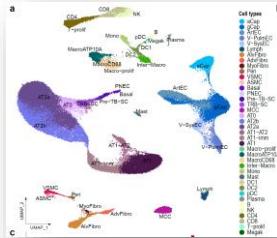
Characterized by the presence of occluded pulmonary arterioles resulting from the proliferation of pulmonary artery endothelial cells (PAECs), pulmonary artery smooth muscle cells (PASMCs) and fibroblasts, which leads to right heart hypertrophy and eventual cardiac failure



- Defined by a mean pulmonary arterial pressure >20 mmHg
- More frequent in women to men (2:1 to 4:1)
- Different origins:
 - IPAH (idiopathic or sporadic cases),
 - HPAH (heritable case family history) 6-10% monogenic autosomal-dominant - 14% ♂, 42% ♀
 - APAH (associated forms), anorexigens / liver / congenital heart / connective tissue disease

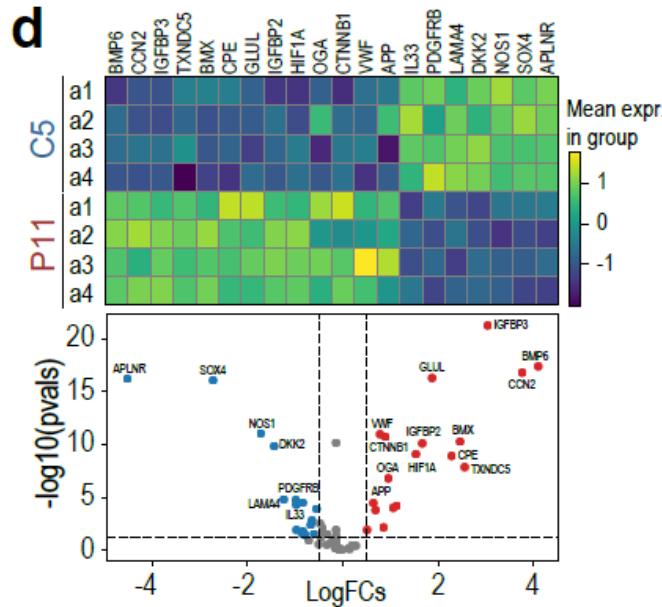
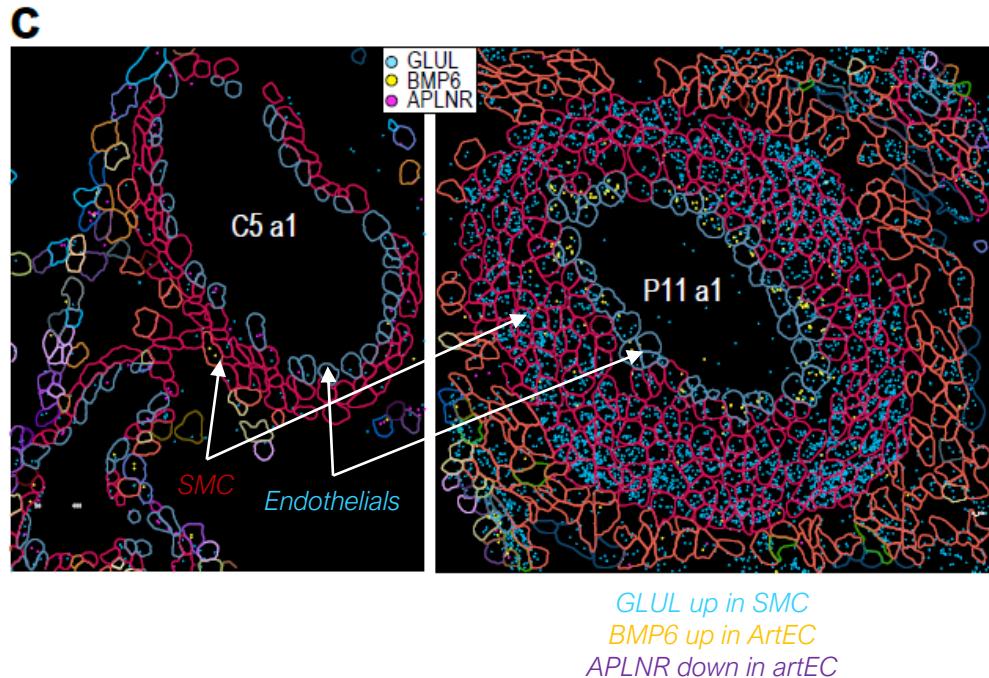
PAH : Pulmonary Arterial Hypertension

A rare vascular disorder



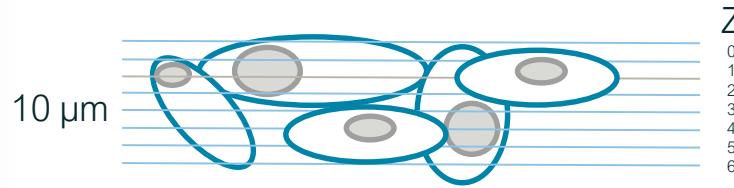
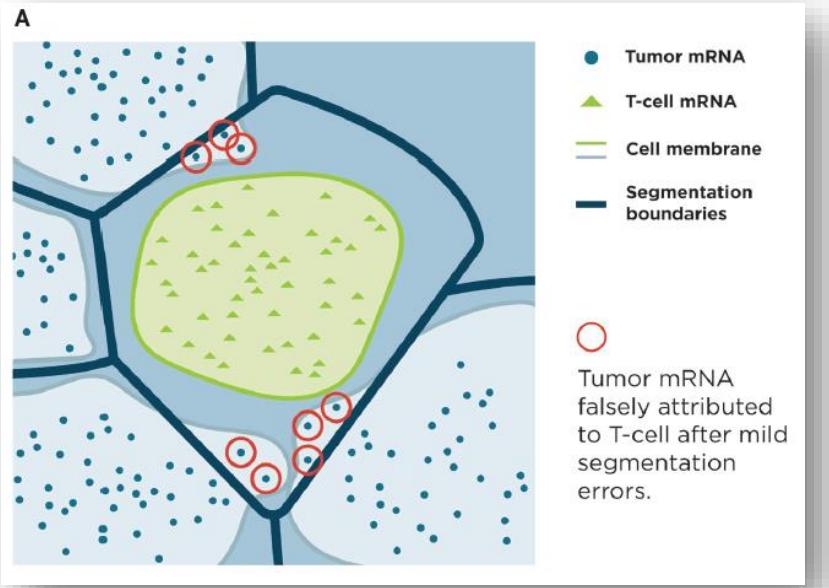
PAH : Pulmonary Arterial Hypertension

A rare vascular disorder

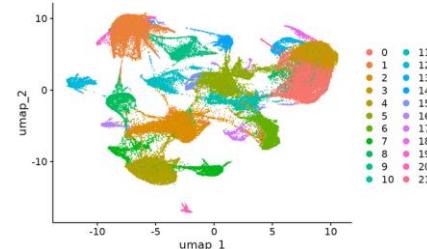


Cell segmentation is crucial

cell x gene matrix purity and 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 segmentation mask (nuclei of full cell)



Vizgen Postprocessing tool

Run on system output folder

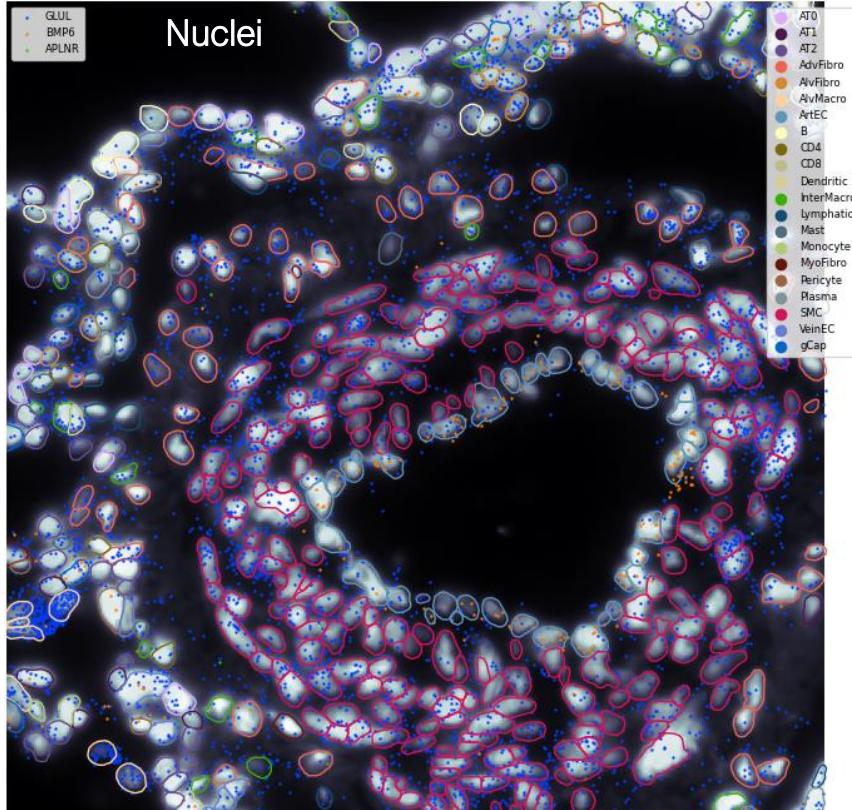
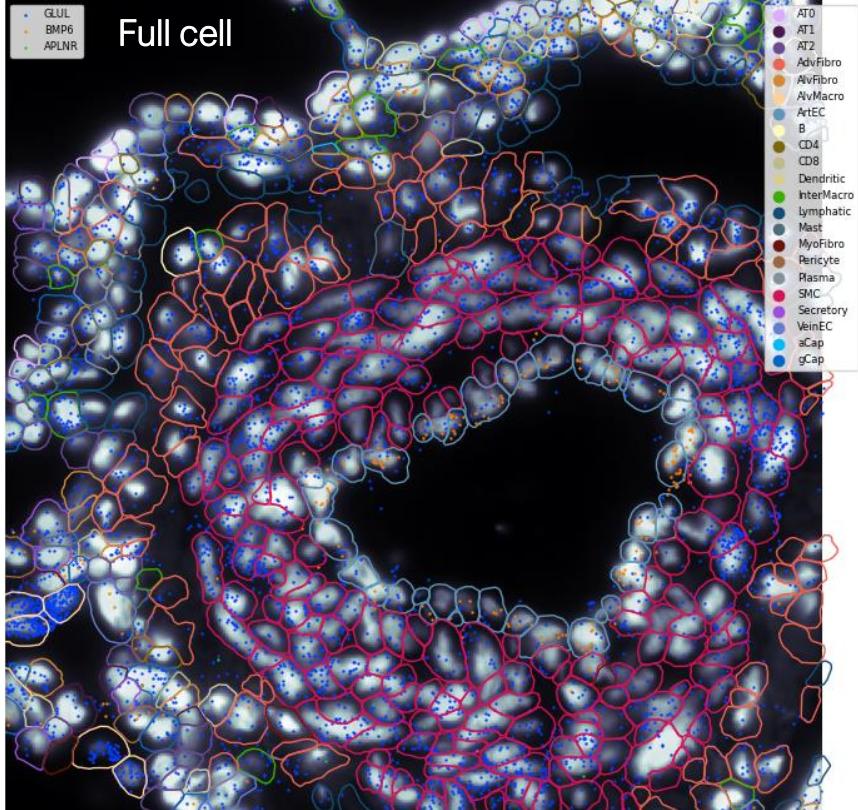
The Vizgen Post-processing Tool (VPT) enables users to reprocess and refine the single-cell results of MERSCOPE experiments. VPT is a command line tool that emphasizes scalable, reproducible analysis, and can be run on a workstation, a cluster, or be deployed in a cloud computing environment.

Features

- Perform cell segmentation
 - Reproduce standard Vizgen segmentation options
 - Perform reproducible custom segmentation
 - Import cell segmentation from other tools
 - Supports geojson and hdf5 formats
 - Regenerate single cell data with new segmentation
 - Cell by gene matrix
 - Cell spatial metadata
 - Image intensity in each cell
 - Update MERSCOPE Vizualizer file (vzg)
 - Image format conversion
 - Convert large tiff files to single or multi-channel Pyramidal OME-TIFF files
 - Nextflow compatible, example pipeline provided



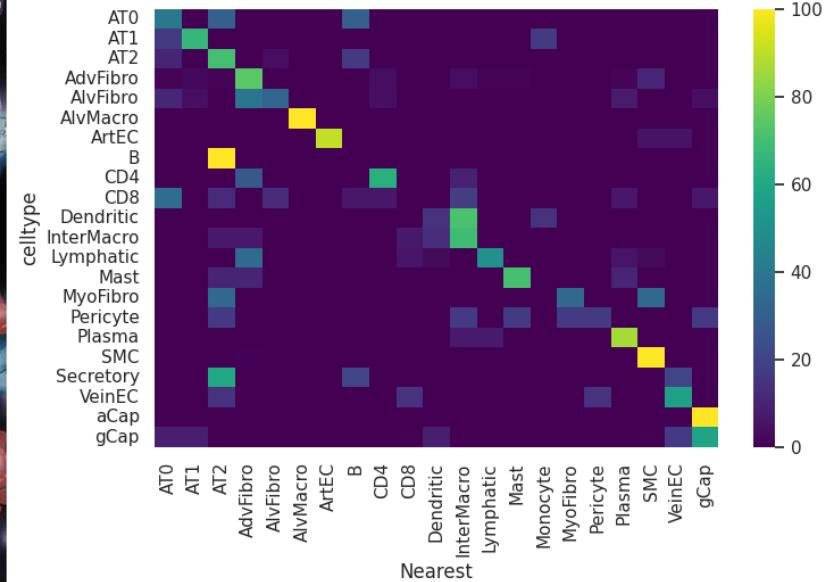
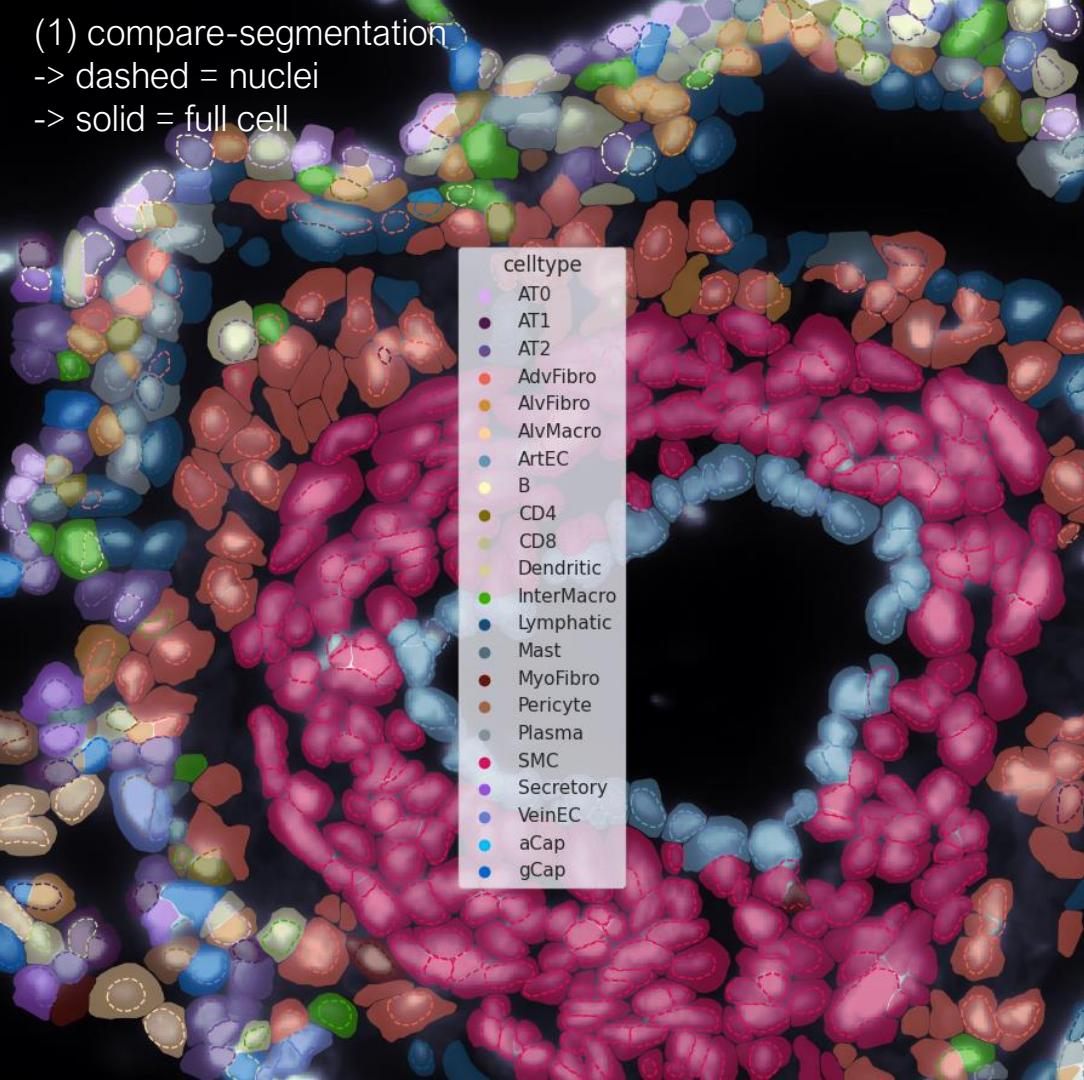
P11 cell versus nucleus segmentation



35% of detected transcript is lost but potentially a cleaner matrix for better biology

(1) compare-segmentation

- > dashed = nuclei
- > solid = full cell



cells = 572

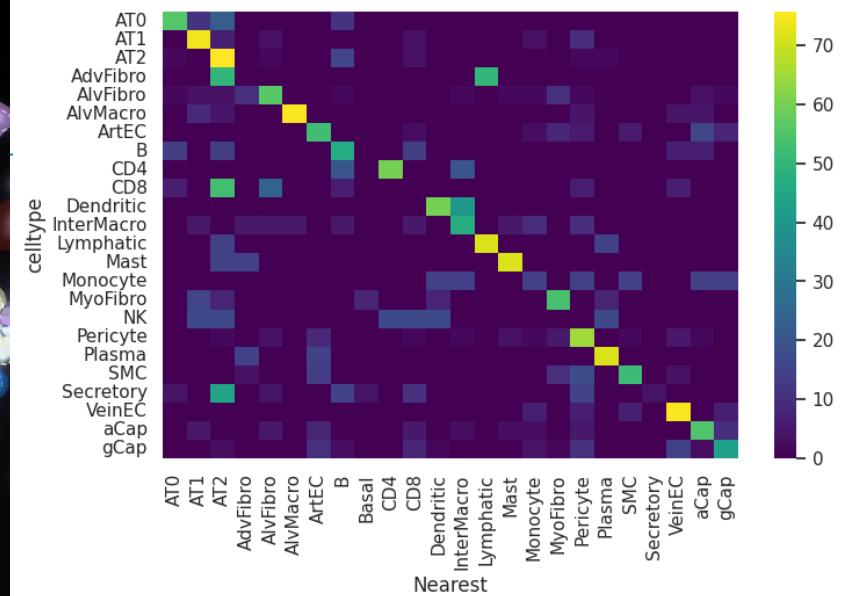
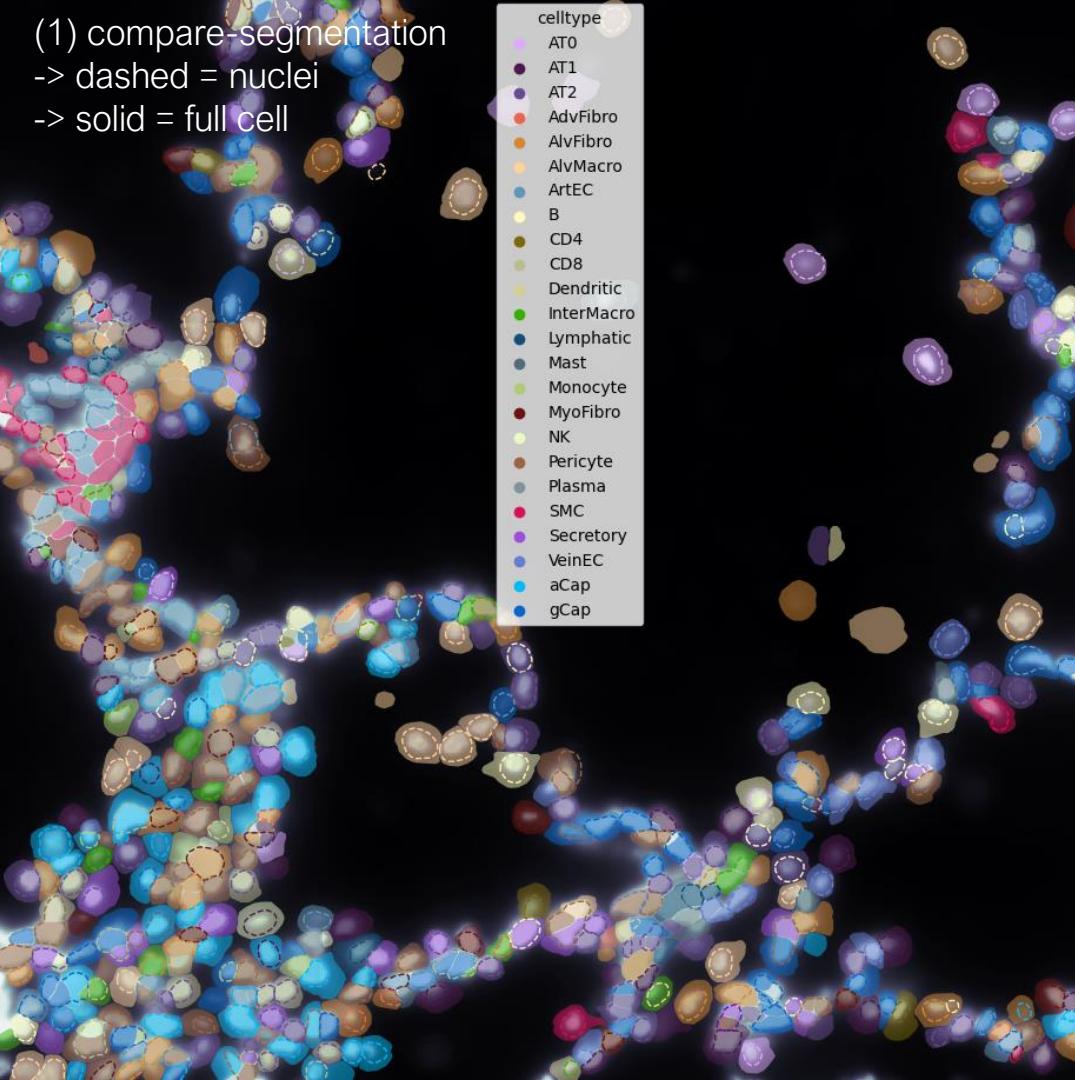
nuclei = 509

→ 428 label (84%)

(1) compare-segmentation

-> dashed = nuclei

-> solid = full cell



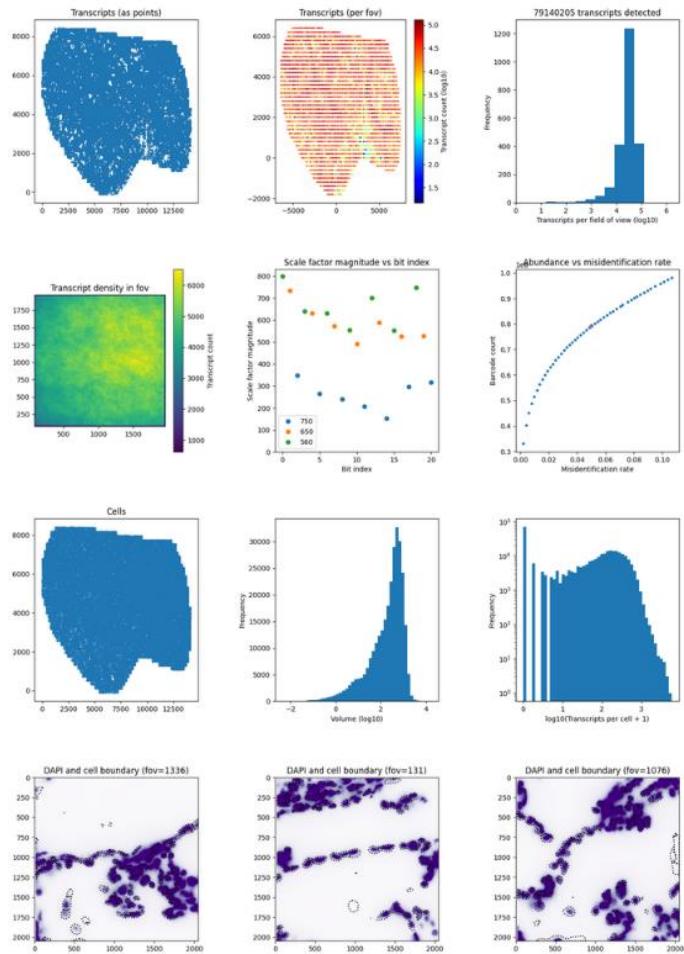
Cell : 586 cells

Nuclei : 452 cells

→ 316 same label (70%)

P11

202302201342_20230220-HTAP-P11_VMSC06001 (34969.7 transcripts per fov)



P11-2

202311241634_20232411-hts-wtr-l1330-TU-P11-2_VMSC06001 (5065.8 transcripts per fov)

