



Exploring Spatial Transcriptomics

2020-05-28
Alma Andersson



<https://github.com/almaan>



SPATIAL
research
<https://www.spatialresearch.org>

A brief Introduction

- Alma Andersson
- From : Utterbäck, Sweden
 - Population : 69
- Now : Stockholm, Sweden
 - Population : 1,605,030
- MSc.Eng from KTH
- 2017-2018 : Delemotte Lab
 - Molecular Dynamics
 - Membrane proteins (Ion Channels)
- 2018-Current : Lundeberg Lab
 - Spatial Transcriptomics (ST)
 - Computational Method Development
- github : almaan
- Small disclaimer : first online teaching experience

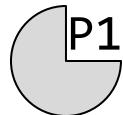


Agenda

- ~~Introduction~~
- Notation
- Background
- Data Orientation
- Break
- Questions
- Methods of Analysis
- Examples
- Information Regarding Exercises
- Questions

Notation

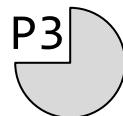
- Workshop consists of 3 parts
- Symbols below used to indicate when material is included in one of these



- Material in Part 1



- Material in Part 2



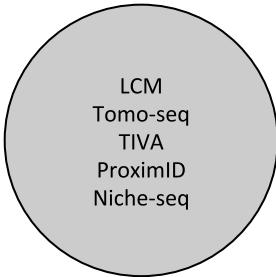
- Material in Part 3

|| ■■ Background ■■||

The spatial space

The spatial space

Microdissection

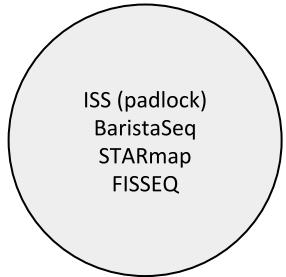


Isolate a region of interest, place isolate in separate well and sequence (either by bulk or single-cell methods).

“Brute Force” approach.

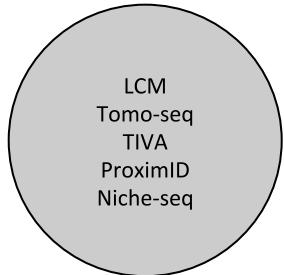
The spatial space

In situ sequencing



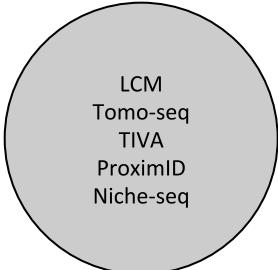
Sequence the transcripts in place. Provides sub-cellular resolution. Tend to rely on gene panels, “*a priori*” defined targets.

Microdissection

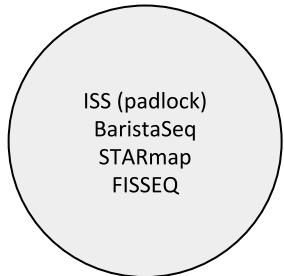


The spatial space

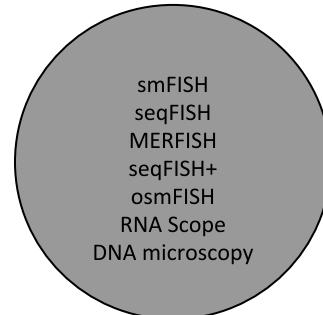
Microdissection



In situ sequencing



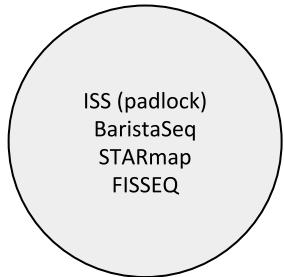
In situ hybridization



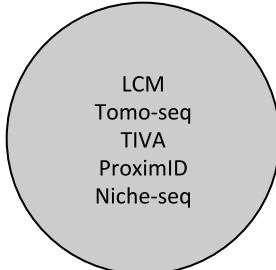
Labeled probes for specific targets, hybridize in place and visualize for spatial information.

The spatial space

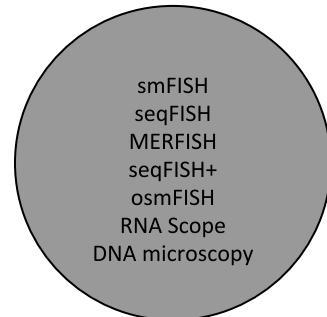
In situ sequencing



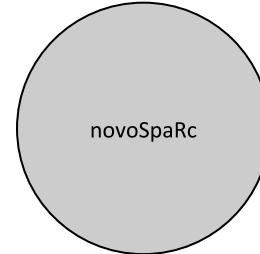
Microdissection



In situ hybridization

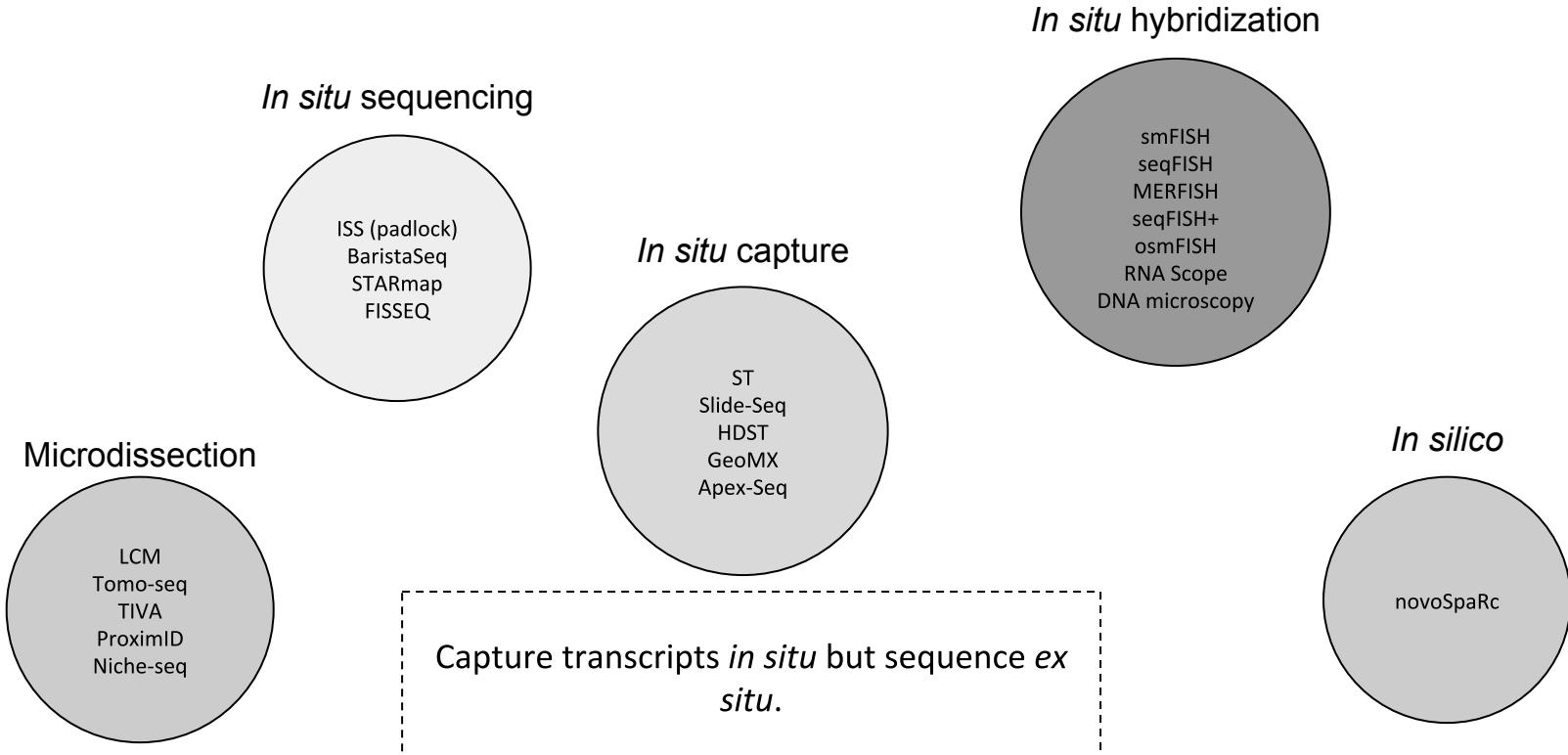


In silico



Infer and reconstruct spatial structure from
single cell data.

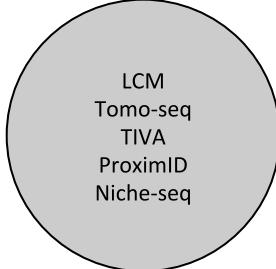
The spatial space



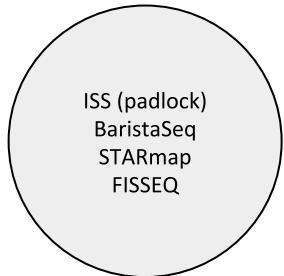
Credit to J.Bergenstråhle and M.Asp for categorization of techniques

The spatial space

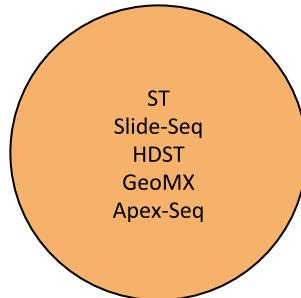
Microdissection



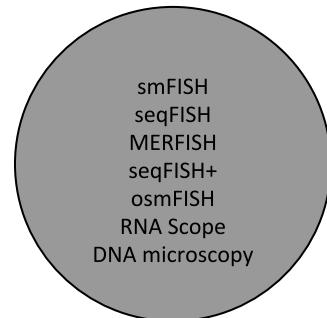
In situ sequencing



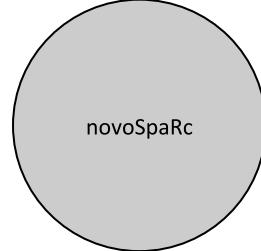
In situ capture



In situ hybridization



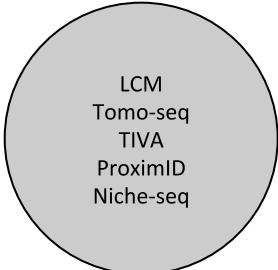
In silico



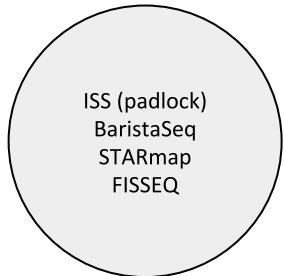
Credit to J.Bergenstråhle and M.Asp for categorization of techniques

The spatial space

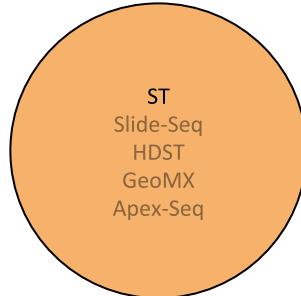
Microdissection



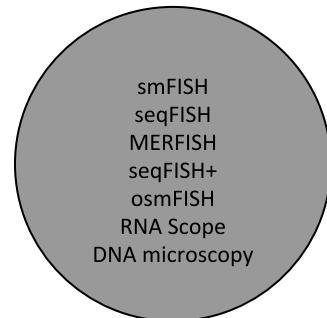
In situ sequencing



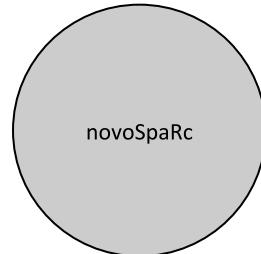
In situ capture



In situ hybridization



In silico



Credit to J.Bergenstråhle and M.Asp for categorization of techniques

Spatial Transcriptomics (ST)

TRANSCRIPTION

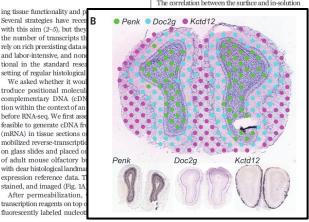
Visualization and analysis of gene expression in tissue sections by spatial transcriptomics

Patric L. Ståhl,^{1,2,*} Fredrik Salminen,^{3,7} Sanja Vicković,^{2,8} Anna Lundmark,^{2,9,2*} José Fernández-Nebro,^{1,3,10} Jens Magnusson,^{1,3,10} Stefania Giannella,^{1,3,10} Michaela Asplund,² Joakim Sandelin,^{1,3,10} Åke Bergström,^{1,3,10} Anders Sandberg,^{1,3,10} Kristoffer Wiktorin,^{1,3,10} Simon Cudrup,^{1,3,10} Åke Borg,² Fredrik Pontén,² Paul Iggo,¹⁰ Petter Sahlén,² Jan Muster,² Olaf Bergmann,¹ Joakim Landenberg,² Jonas Frisell,²

Analysis of the distribution of proteins or messenger RNAs (mRNAs) in tissue sections by spatial transcriptomics. We have developed a novel strategy, which we call "spatial transcriptomics," that allows visualization and quantitative analysis of the transcription with spatial resolution down to the level of individual cells. By using a combination of mRNA reverse sequencing primers with unique political barcodes, we demonstrate high-quality RNA sequencing data with maintained two-dimensional positional information from the mouse brain. This approach allows us to simultaneously measure quantitative gene expression and visualization of the distribution of mRNAs within tissue sections and enables novel types of bioinformatics analysis, valuable in research and diagnostics.

Tissue sequencing is typically studied by RNA sequencing (RNA-seq) (*J*). However, this approach has several disadvantages. First, it requires large amounts of averaged transcriptome and loss of spatial information. The positional context of gene expression is at best dependent on understanding

Department of Cell and Molecular Biology, Karolinska Institutet, Stockholm, Sweden; ²KTH Royal Institute of Technology, Stockholm, Sweden; ³Science for Life Laboratory, Department of Gene Technology, KTH Royal Institute of Technology, Stockholm, Sweden; ⁴Department of Oral Medicine, Division of Periodontology, Karolinska Institutet, Stockholm, Sweden; ⁵Department of Cell Biology, Karolinska Institutet, Stockholm, Sweden; ⁶Department of Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden; ⁷Department of Molecular Neuroscience, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden; ⁸Department of Physiology and Pharmacology, Karolinska Institutet, Stockholm, Sweden; ⁹Swedish Translational Neuroscience Center, Uppsala University, Uppsala, Sweden; ¹⁰Swedish Translational Neuroscience Center, Karolinska Institutet, Stockholm, Sweden. *These authors contributed equally to this work. ¹These authors contributed equally to this work. Corresponding author: jonasfrisell@kth.se



anchored cDNA (Fig. 1A, and fig. S1). The tissue was then enzymatically removed, which left cDNA coupled to the surface oligonucleotides on the slide (Fig. 1B). After fragmentation, mRNA was isolated corresponding to the tissue structure revealed by the general histology (Fig. 1B and C), and the cDNA was used for mRNA sequencing (Fig. 1C, Fig. S1, D to G). By comparing the hematoxylin-and-eosin and fluorescence signals, we could measure the average size of the tissue structures (the size of a cell to $1.7 \times 2 \mu\text{m}$ (mean \pm SD) (Fig. S1, E to H).

We have developed a strategy to measure mRNA in tissue sections with minimal diffusion and maintained positional representation using unique political barcodes (Fig. 1D) and barcodes (Fig. 2A), and we denoted this strategy "spatial transcriptomics." We deposited ~20 million mRNA molecules onto each slide (Fig. 2A) with a diameter of 100 μm and a center-to-center distance of 200 μm over an area of 0.15 cm^2 (Fig. S2). We generated sequencing libraries based on mRNA extracted from the tissue sections (Fig. 2A and B) (Fig. 2C). Comparison with data from RNA extracted and fragmented in solution revealed a high correlation between the two methods (Fig. 2D). The correlation was also found with the other (fig. S2C).

The correlation between the surface and tissue

is

high

and

stable

over

time

(fig. S2E)

Several strategies have recently been developed to increase the number of transcripts in the standard next-generation sequencing setting of regular histological samples. We believe that our strategy introduce positional molecular information to RNA sequencing data before RNAseq. We first use a combination of mRNA and cDNA (mRNA) in tissue sections to mobilized reverse transcriptase and polymerase chain reaction (PCR) to avoid loss of adult mouse olfactory bulb tissue sections. We have generated a comprehensive expression reference data. Tissue sections were fixed, permeabilized, and stained, and imaged (Fig. 1A to C). mRNA was then extracted, fragmented, and labeled with unique political barcodes (Fig. 1D to G). These political barcodes are complementary to the sequencing

Science; Ståhl et al. . 2016

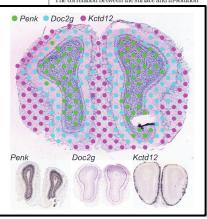
DOI: 10.1126/science.aaf2403

Spatial Transcriptomics (ST)

TRANSCRIPTION

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

Issue transcriptomes are typically studied by RNA-seq sequencing (RNA-seq) [7] of *in vitro* or *in vivo* samples to determine the presence and quantity of gene expression information. The potential context of gene expression is of key importance to understand its function. Several strategies have been developed to increase the number of transcriptomes analyzed by RNA-seq, including the use of RNA pools [8], the use of RNA barcodes [9], and the use of RNA index genes [10]. RNA index genes are used to standardize the expression of genes across different samples. RNA index genes are usually selected from the transcriptome of a reference cell type, such as fibroblasts, and are used to normalize the expression of other genes. RNA index genes are usually selected from the transcriptome of a reference cell type, such as fibroblasts, and are used to normalize the expression of other genes.



Late 2018

10x Genomics™ acquisition

Science; Ståhl et.al . 2016

DOI: 10.1126/science.aaf2403

SciLifeLab

Spatial Transcriptomics (ST)

Mid 2016

TRANSCRIPTION

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

Patric L. Ståhl,^{1,2,*} Fredrik Salminen,^{3,7} Sanja Vicković,⁷ Anna Lundmark,^{2,3} José Fernández Navarro,^{1,3} Jens Magnusson,¹ Stefania Giannella,⁷ Michaela Asplund,² Joakim Söderström,¹ Mats Jansson,¹ Åsa Berglund,¹ Annika Söderström,¹ Simonne Cudkovich,^{4,5} Åke Borg,² Fredrik Pontén,² Paul Igro Coates,² Petter Salminen,² Jan Musterer,² Olaf Bergmann,² Joakim Lundeberg,² Jonas Frisell,²

Analysis of the distribution of proteins or messenger RNAs (mRNAs) in tissue sections is a key technique in biological research and diagnostics. This study presents the distribution of few proteins or expressed genes at a time. We have developed a strategy, which we call "spatial transcriptomics," that allows visualization and quantitative analysis of the transcripts with unique molecular barcodes. By using a combination of standard reverse transcription and reverse transcription primers with unique political barcodes, we demonstrate high-quality RNA sequencing data with maintained two-dimensional positional information from the mouse brain. We also show that the method can be used to analyze gene expression in tissue sections and visualization of the distribution of mRNAs within tissue sections and enables novel types of bioinformatics analyses, valuable in research and diagnostics.

Tissue transcriptomics are typically studied by RNA sequencing (RNA-seq) (*i*) of homogenized tissue samples. This approach has averaged transcriptome and loss of spatial information. The positional context of gene expression is of key importance to understand-

ing tissue functionality and disease processes.

Several strategies have recently been developed to maintain the number of transcripts at the single-cell level and to extend the application to the standard research setting of regular histological tissue sections (*ii*). These approaches rely on rich preexisting molecular reference data and often focus on a specific cell type or tissue within the context of an entire transcriptome. We first analyzed the distribution of mRNA (mRNA) in tissue sections of mobilized reverse transcriptase (RT)-labelled adult mouse olfactory bulb (OB) tissue using a set of mRNA expression reference data. Tissue was fixed, paraffin-embedded, sectioned, and imaged (Fig. 1A). The mRNA was then converted to cDNA using RT and fluorescently labeled with

*These authors contributed equally to this work.
†These authors contributed equally to this work.
¹Corresponding author. E-mail: jonas.frisell@ki.se
²Department of Cell and Molecular Biology, Karolinska Institutet, Stockholm, Sweden; ³Department of Oral and Maxillofacial Surgery, Karolinska Institutet, Stockholm, Sweden; ⁴Department of Dentistry, Karolinska Institutet, Stockholm, Sweden; ⁵Department of Oral Medicine, Division of Periodontology, Karolinska Institutet, Stockholm, Sweden; ⁶Department of Radiology, Karolinska Institutet, Stockholm, Sweden; ⁷Department of Biostatistics, Karolinska Institutet, Stockholm, Sweden; ⁸Department of Physiology and Pharmacology, Karolinska Institutet, Stockholm, Sweden; ⁹Swedish Institute of Dermatology, Karolinska Institutet, Stockholm, Sweden; ¹⁰Department of Oncology and Pathology, Department of Clinical Genetics, Karolinska Institutet, Stockholm, Sweden; ¹¹Department of Immunobiology, Genetics and Pathology, Uppsala University, Uppsala, Sweden; ¹²Umeå Center for Science in Life, Umeå University, Umeå, Sweden

Science Publication
Ståhl et.al

DOI: 10.1126/science.aaf2403

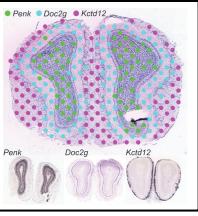
synthesized cDNA (Fig. 1A, and fig. S1). The tissue was then enzymatically removed, which left cDNA coupled to the service oligonucleotides on the slide (Fig. 1B). The cDNA was then fragmented and corresponding to the tissue structure revealed by the general histology (Fig. 1B and C), and the cDNA was sequenced (Fig. 1C, and Fig. 2A, B, C, D, E, F, G, H, I, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z, D to G). By comparing the hematoxylin-and-eosin and fluorescence signals, we could measure the average distance between the center of a cell to a cell to 1.7 ± 2 μm (mean ± SD) (Fig. S1, F to H).

We next used the same strategy to measure mRNA in tissue sections with minimal diffusion and maintained positional representation notwithstanding the use of standard histological procedures (Fig. 2A), and we denoted this strategy "spatial transcriptomics."

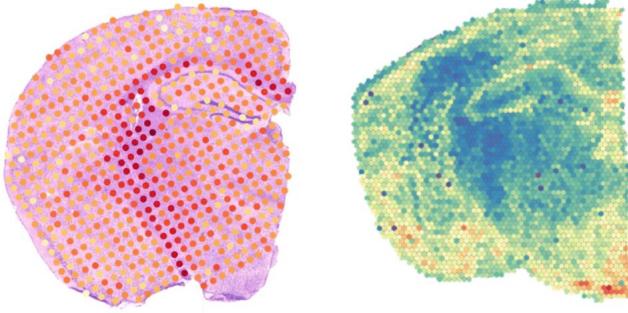
We deposited ~20 million

transcripts in the public repository (Fig. 2A)

(Fig. 2B) and found a correlation between the



Late 2018
10x Genomics™
acquisition



Launch of Visium Spatial Gene Expression Platform

Spatial Transcriptomics (ST) → Visium

Mid 2016

TRANSCRIPTION

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

Patrik L. Ståhl,^{1,2,*} Fredrik Salhün,^{2,3} Sanja Vicković,^{2,4} Anna Lundmark,^{2,5,6} José Fernández-Nebro,^{1,2} Jens Magnusson,¹ Stefania Giannouli,² Michaela Aspli,² Joakim Lundeberg,^{1,2} Mikael Malmström,¹ Åsa Berg,² Anders Sandelin,² Simon Cudkovich,^{2,7} Åke Borg,² Fredrik Pontén,² Paul Igivu Costa,² Petter Salhün,² Jan Muster,² Olaf Bergmann,² Joakim Lundeberg,² Jonas Frisell,²

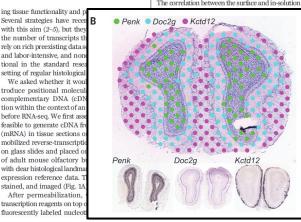
Analysis of the distribution of proteins and messenger RNAs (mRNAs) in *Neurogranin* (Ngn2) sections is a key step in the visualization and research of gene expression. This study provides the distribution of few proteins or expressed genes at a time. We have developed a strategy, which we call "spatial transcriptomics," that allows visualization and quantitative analysis of the transcripts with spatial resolution down to 10 µm. By using a combination of spatial transcriptomics and barcoded reverse transcription primers with unique political barcodes, we demonstrate high-quality RNA sequencing data with maintained two-dimensional positional information from the mouse brain. The method is general and can be applied to any tissue section. We also show that mRNA sequencing data with maintained two-dimensional positional information can be used to predict gene expression of brain tissue sections. This study provides a new way to visualize and visualization of the distribution of mRNAs within tissue sections and enables novel types of bioinformatics analyses, valuable in research and diagnostics.

Tissue transcriptomes are typically studied by RNA sequencing (RNA-seq) (*i*) of homogenized samples. This approach loses averaged transcriptome and loss of spatial information. The positional context of gene expression is of key importance to understand-

ing tissue functionality and disease processes.

Several strategies have recently been developed to maintain the spatial context of the number of transcripts of individual genes in tissue sections. These approaches are typically based on the use of barcoded reverse transcription primers and sequencing of cDNA in the standard sequencing setting of regular biological samples.

We have developed a strategy to introduce positional molecular information to the transcriptome within the context of barcoded RNA-seq. We first use a combination of total RNA (tRNA) and mobilized reverse transcriptase to generate cDNA in tissue sections. tRNA is used to reduce the risk of adult mouse olfactory bulb-derived RNA interference with the gene expression reference data. tRNA is stained, and imaged (Fig. 1A). The tRNA is then used as a template for reverse transcription reaction with top-sequencing barcodes fluorescently labeled on nucleotides.

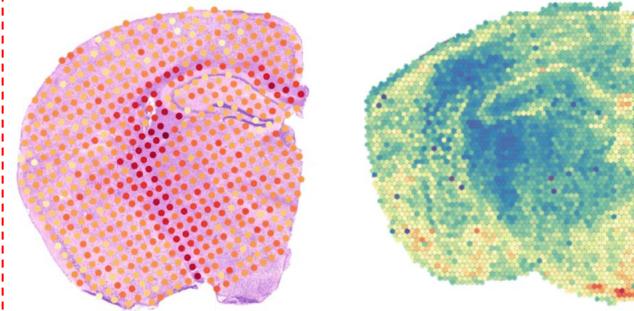


Science Publication
Ståhl et.al

DOI: 10.1126/science.aaf2403

Late 2018
10x Genomics™
acquisition

Late 2019



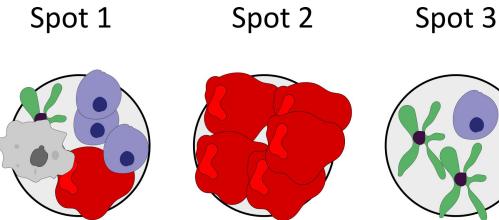
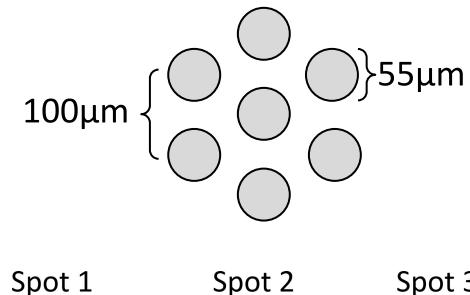
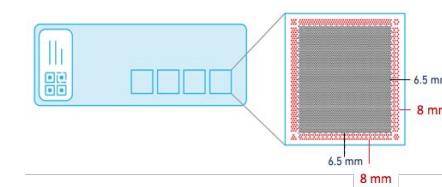
Launch of Visium Spatial Gene Expression Platform

SciLifeLab

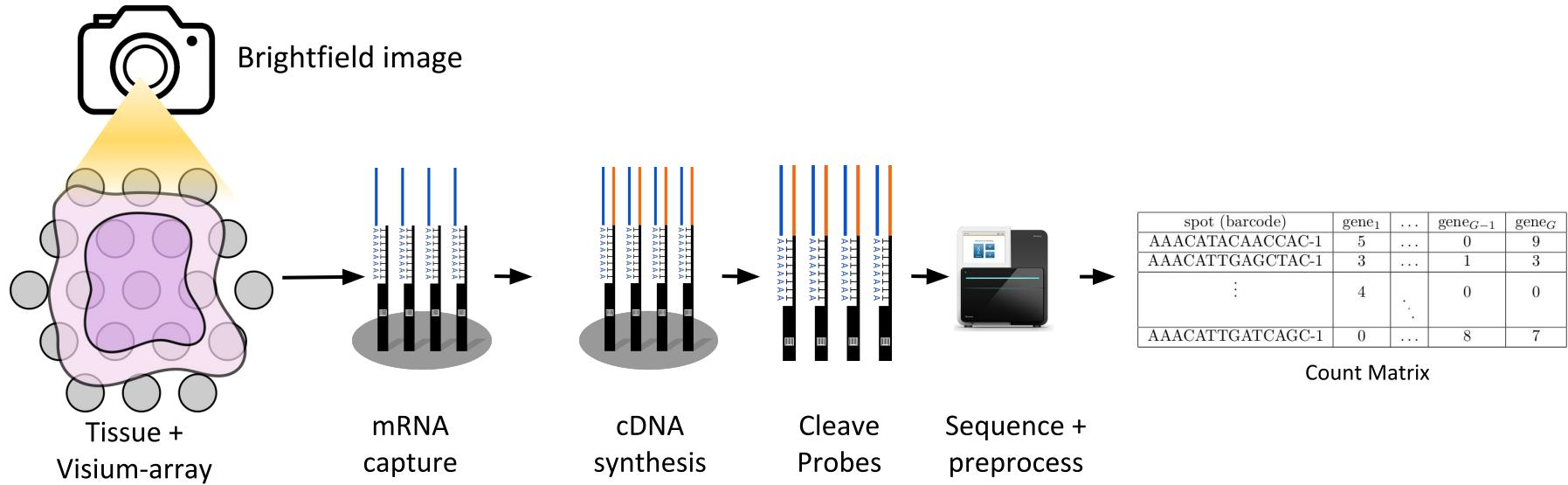


Visium Platform

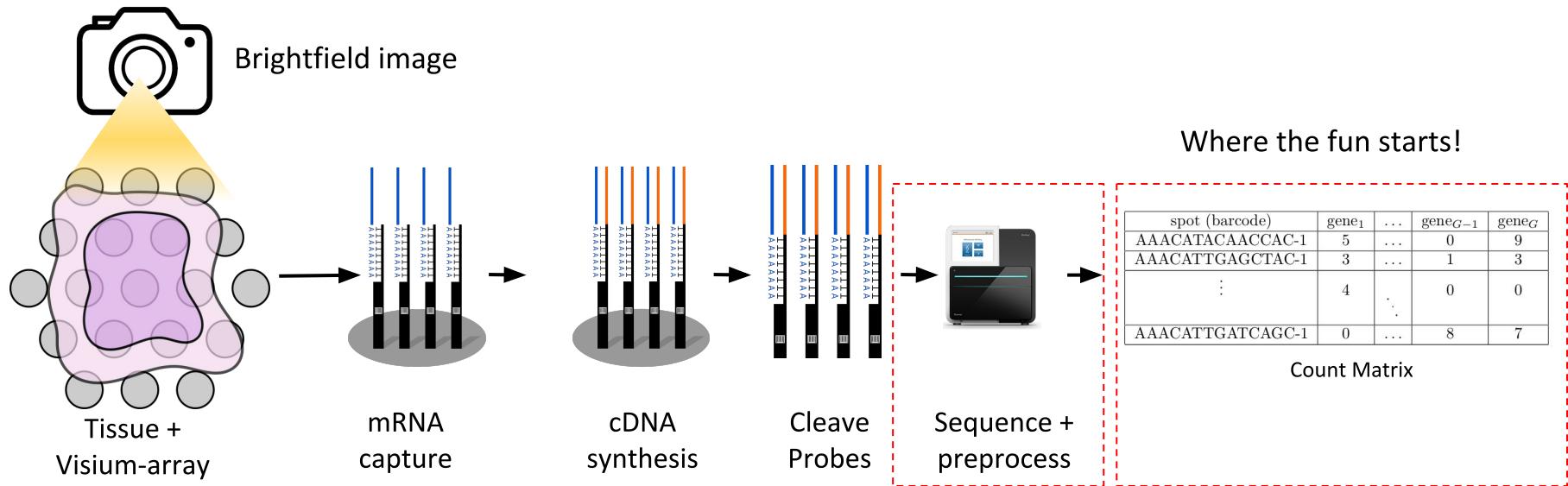
- Array based technique
- 6.5mm x 6.5mm area to put sample on
- 4992 spots arranged in regular grid
- Spot specs:
 - Spot diameter : 55 μm
 - Center to center distance : 100 μm
- Each spot has millions of capture probes
 - spatial barcode
 - polyT sequence
 - captures polyadenylated mRNA
 - Full transcriptome(-ish)
- ~ 1-10 cells contribute to each spot
 - **NOTE :** Not single cell resolution!



The experimental workflow (in a nutshell)



The experimental workflow (in a nutshell)





|| ■ ■ Data Orientation ■ ■ ||

After sequencing (brief)

- **spaceranger mkfastq** | BCL files to FASTQ
- **spaceranger count** | tissue detection/alignment, UMI counting



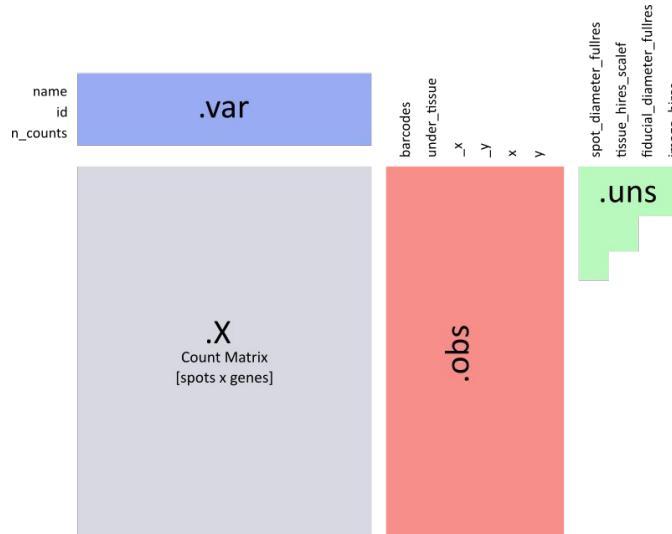
```
-bash-4.2$ tree -L 2
.
├── analysis
│   ├── clustering
│   ├── diffexp
│   ├── pca
│   ├── tsne
│   └── umap
└── cloupe.clope
    ├── filtered_feature_bc_matrix
    │   ├── barcodes.tsv.gz
    │   ├── features.tsv.gz
    │   └── matrix.mtx.gz
    ├── filtered_feature_bc_matrix.h5
    ├── metrics_summary.csv
    ├── molecule_info.h5
    ├── possorted_genome_bam.bam
    ├── possorted_genome_bam.bam.bai
    ├── raw_feature_bc_matrix
    │   ├── barcodes.tsv.gz
    │   ├── features.tsv.gz
    │   └── matrix.mtx.gz
    └── raw_feature_bc_matrix.h5
        ├── spatial
        │   ├── aligned_fiducials.jpg
        │   ├── detected_tissue_image.jpg
        │   ├── scaleFactors.json
        │   ├── tissue_hires_image.png
        │   ├── tissue_lowres_image.png
        │   └── tissue_positions_list.csv
        └── web_summary.html
```

The terminal output shows a directory tree structure generated by spaceranger count. The tree includes automatically generated analysis (clustering, diffexp, pca, tsne, umap), cloupe.clope (filtered and raw feature matrices, metrics, genome, and raw matrices), and spatial (aligned fiducials, detected tissue images, scale factors, resized images, and a positions list). A large bracket on the right side groups the cloupe.clope directory and its contents under the label "Automatically generated analysis". Another bracket groups the raw and spatial directories under the label "raw = all spots". Arrows point from specific files in the spatial directory to their descriptions: "For mapping between image coordinates" points to scaleFactors.json, "Resized images" points to tissue_hires_image.png and tissue_lowres_image.png, and "Maps barcode to coordinate" points to tissue_positions_list.csv.

Example of **spaceranger count** output

Processed data

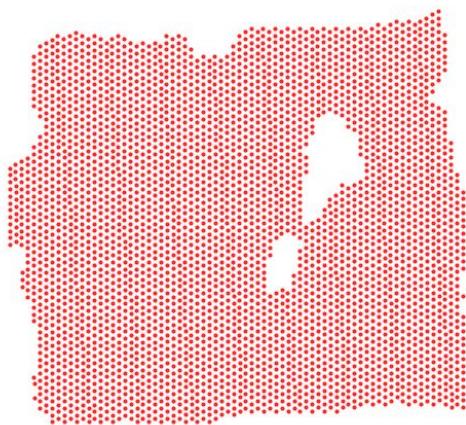
- Either use .mtx files or .h5 files to assemble a data object to work with
 - No standardized format
- Personal preference : convert to .h5ad file (will be using in exercises)
 - scanpy/anndata teams working on - soon to release - their own (similar) format



An initial assessment

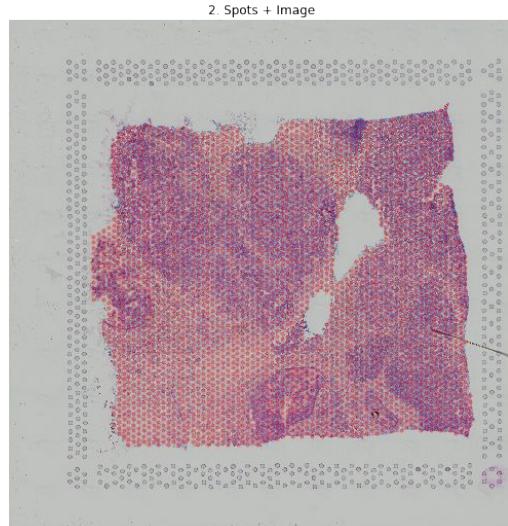
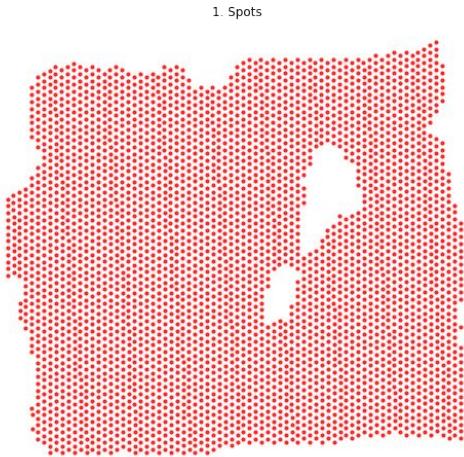
- Example with Human Breast cancer data
 - Public data : Available at 10x website

1. Spots



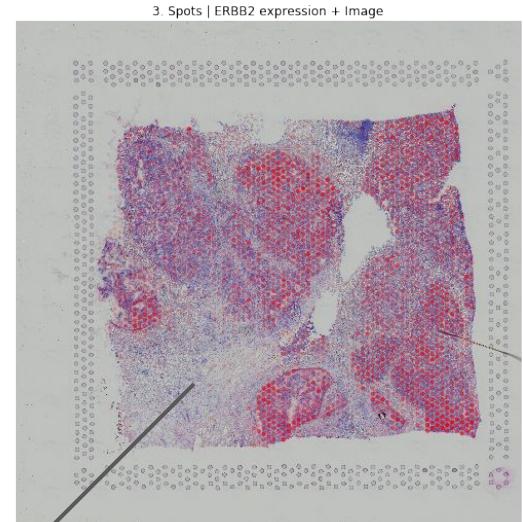
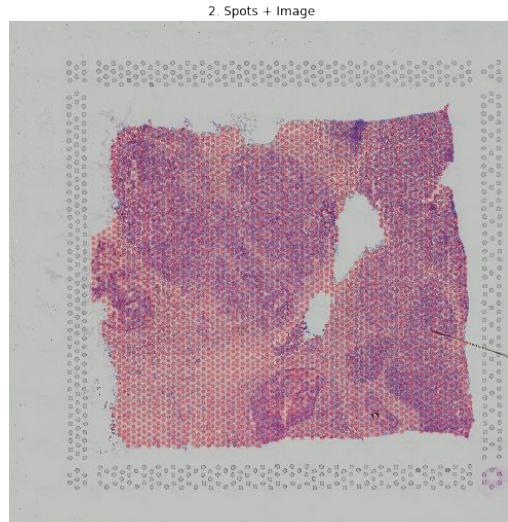
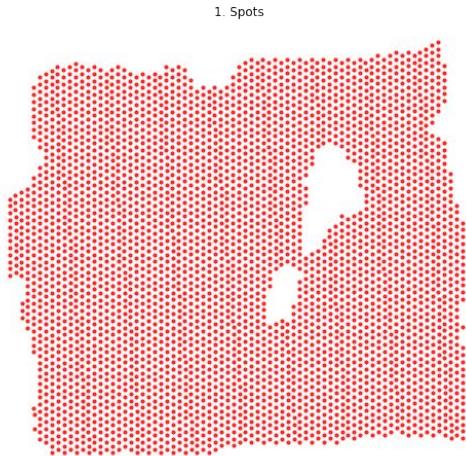
An initial assessment

- Example with Human Breast cancer data
 - Public data : Available at 10x website



An initial assessment

- Example with Human Breast cancer data
 - Public data : Available at 10x website



Facecolor intensity proportional
to gene expression value

■ ■ || Break

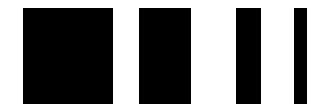
- 10 min Break
- During :
 - Read through questions (for me)
- After :
 - Discuss (some) questions when everyone is back
 - More on Data Analysis
 - Information about the exercises
 - One more question session

(()) C []

■ ■ | Question time

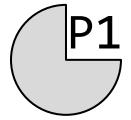


Data Analysis

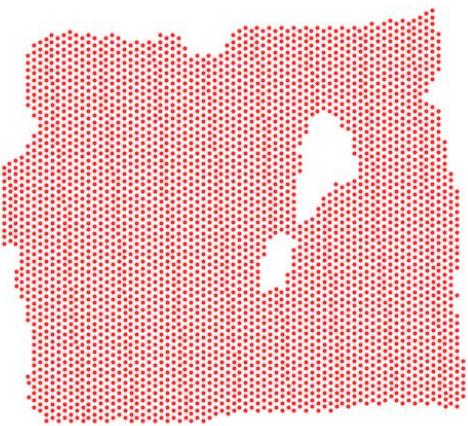




An initial assessment



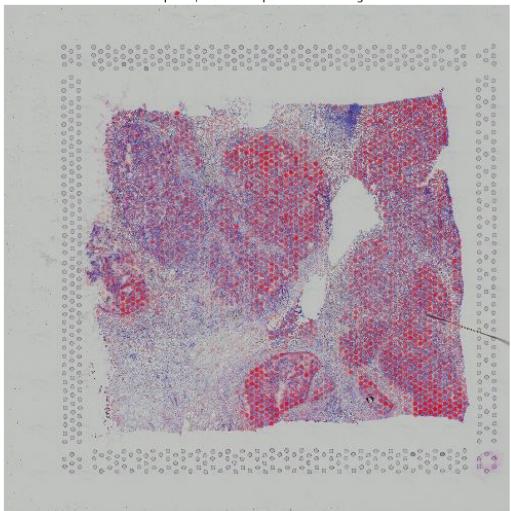
1. Spots



2. Spots + Image



3. Spots | ERBB2 expression + Image



■ ■ | Visualizing high dimensional spatial data

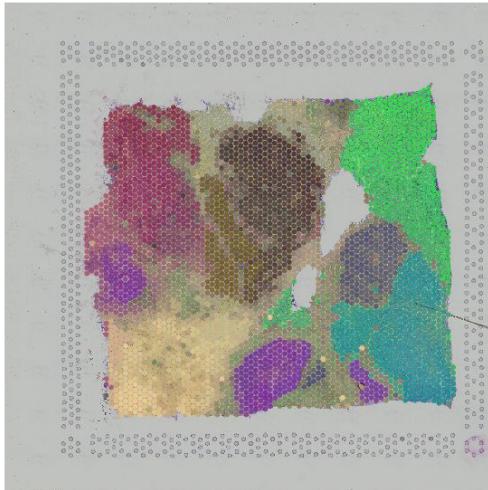
- We visualized one feature/gene (*ERBB2*)
- How do we handle multiple non-mutually exclusive features?

■ ■ | Visualizing high dimensional spatial data

- We visualized one feature/gene (*ERBB2*)
- How do we handle multiple non-mutually exclusive features?
- One idea :
 - Embed gene expression data in 3 dimensional space (e.g. using UMAP)
 - Do affine transformation to unit cube
 - Consider values as RGB values (or other colorspace) and color spots accordingly

■ ■ ■ Visualizing high dimensional spatial data

- We visualized one feature/gene (*ERBB2*)
- How do we handle multiple non-mutually exclusive features?
- One idea :
 - Embed gene expression data in 3 dimensional space (e.g. using UMAP)
 - Do affine transformation to unit cube
 - Consider values as RGB values (or other colorspace) and color spots accordingly



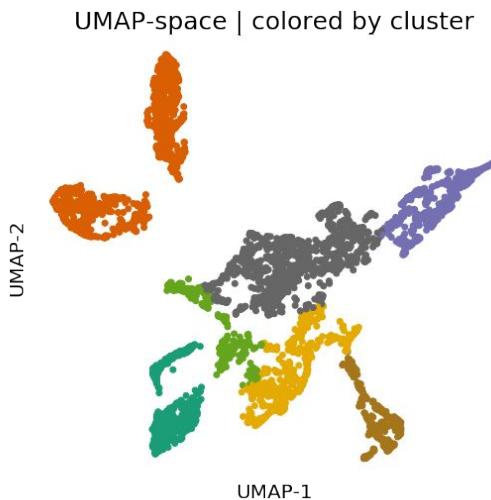
Regions with similar colors have similar gene expression.

Before the analysis

- No magic recipe to give
 - How to process your data is very much dependent on the samples and objective
 - Much can be learnt from analysis of single cell data
 - Will however give some general guidelines
- Parameters to consider for filtering :
 - Genes based on expression levels (total expression > thrs)
 - Genes based on spot presence (#spots gene is observed at > thrs)
 - Spots based on expression levels (total gene expression at spot > thrs)
 - Ribosomal and mitochondrial genes tend to exhibit spurious expression patterns. Exclusion of these is common.
- During normalization / batch correction :
 - Recommend to account for spot “library size”
 - Include slide/array as covariate (sometimes big variation has been observed)
 - Tools that have performed well : e.g., scTransform (Seurat) and Harmony

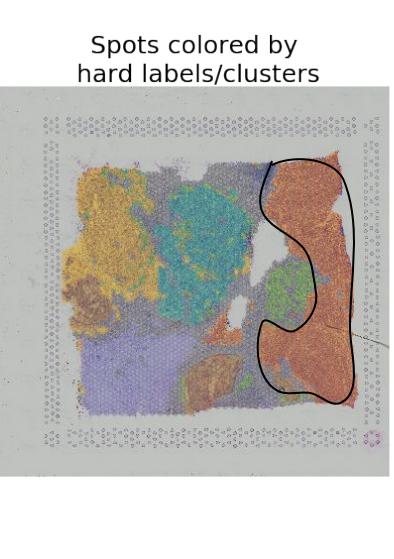
Basic Analysis : Expression based clustering

- Cluster the spots based on gene expression
 - Normalize → PCA → UMAP → GMM
- Backmap clusters onto tissue
- Use HE-image as reference
 - Sanity check - does it make sense?
 - Valuable resource!

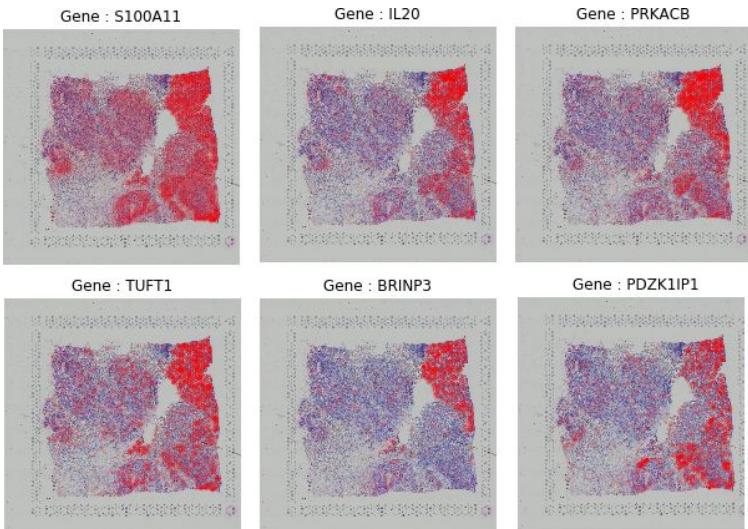




Basic Analysis : DE-analysis



Contrast:
Cluster 1 vs. All



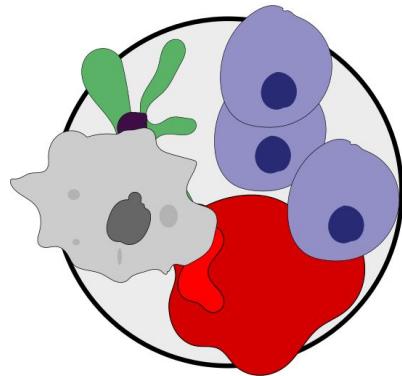


Expression based clustering is informative, BUT

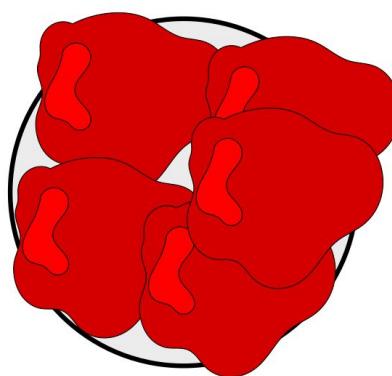
Remember :

each spot is a mixture of multiple cells, i.e., one spot may contain multiple cell types

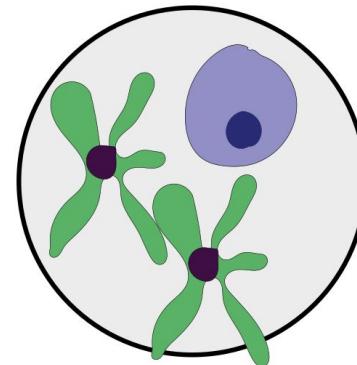
Spot 1



Spot 2

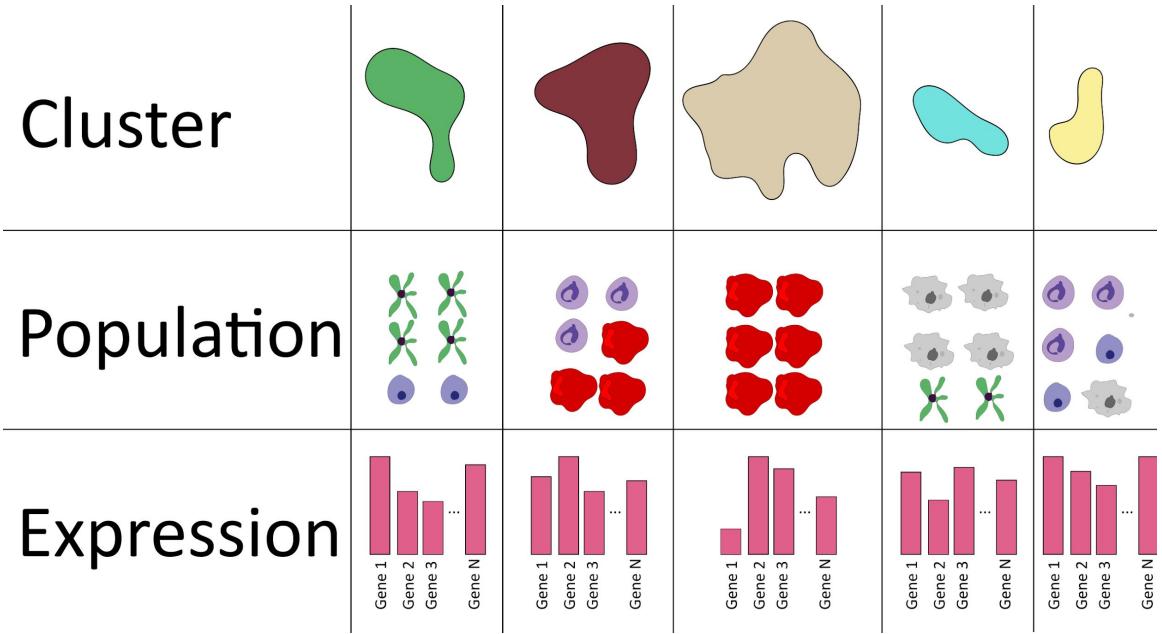
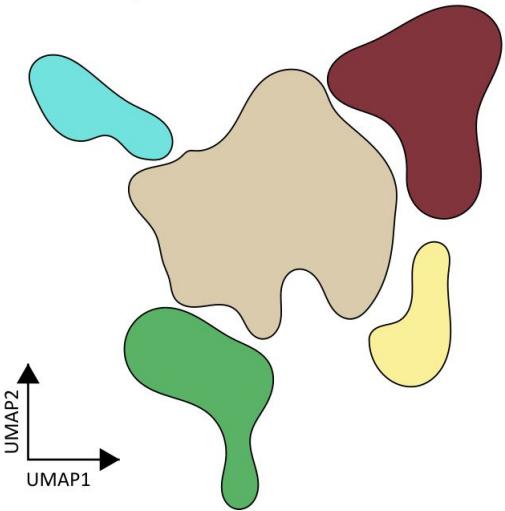


Spot 3



█████ Cluster ≠ cell type

Clustered Spatial Gene Expression Data



So where are my cell types?

- Marker genes?

So where are my cell types?

- Marker genes?
 - Easy and straightforward
 - Requires known marker genes
 - Overlapping among marker genes
 - How to interpret high expression values
 - Lowly expressed markers genes may not always be observed

So where are my cell types?

- Marker genes?
 - Easy and straightforward
 - Requires known marker genes
 - Overlapping among marker genes
 - How to interpret high expression values
 - Lowly expressed markers genes may not always be observed
- Alternative solution - Integrate single cell (SC) and spatial data!

So where are my cell types?

- Marker genes?
 - Easy and straightforward
 - Requires known marker genes
 - Overlapping among marker genes
 - How to interpret high expression values
 - Lowly expressed markers genes may not always be observed
- Alternative solution - Integrate single cell (SC) and spatial data!
 - Extract information of cell types from SC data and apply to spatial data
 - Multiple approaches have emerged, (e.g., MIA by Itai and Seurat's module)
 - **Big challenge : deconvolution required (on Visium data)**



Integration of Single Cell and Spatial Data

Single Cell Data
Clusters of cell types

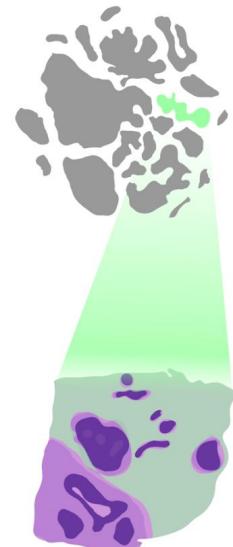
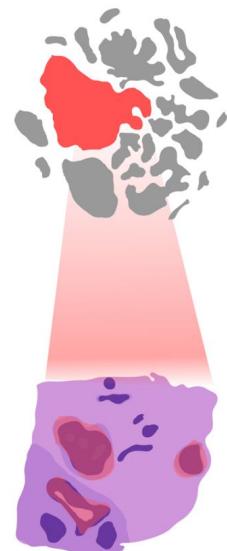


Integration

Spatial Data



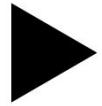
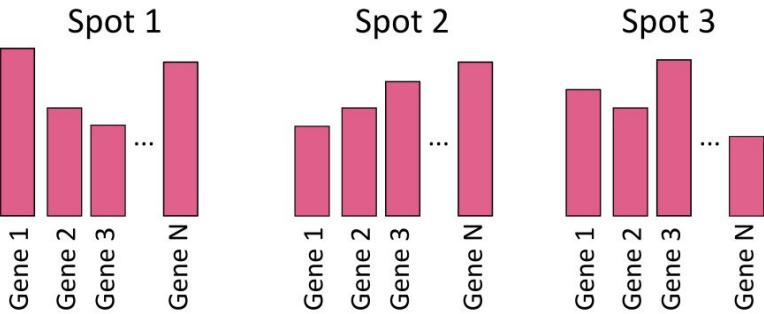
Spatially map each cell type



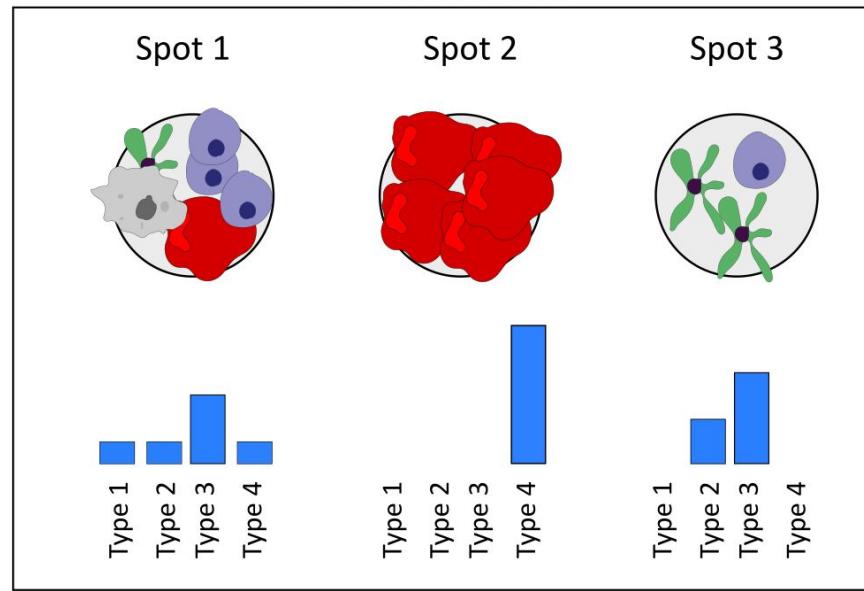


Our objective : deconvolve expression data

From this



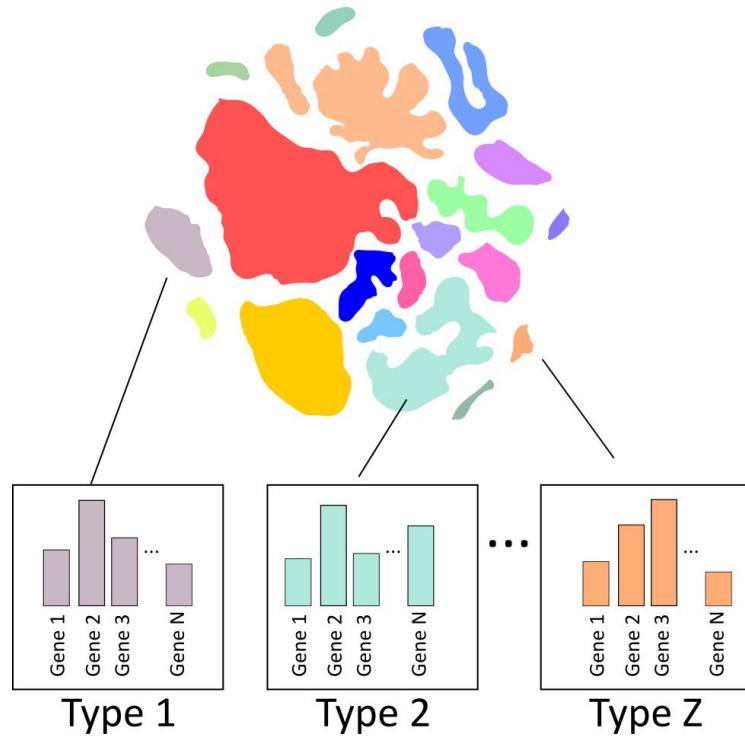
We want this





It's as easy as 1-2-3

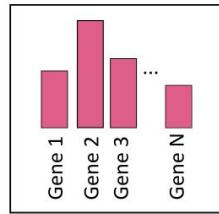
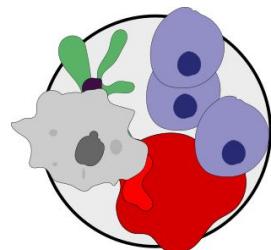
1. Infer cell type expression parameters from SC data





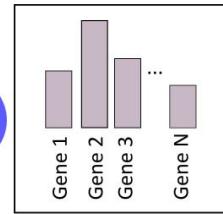
It's as easy as 1-2-3

2. Use inferred parameters to find optimal combination **combination** of cell types in spot



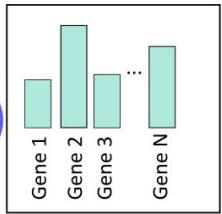
=

W_1



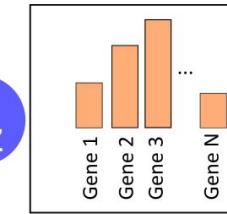
+

W_2



+

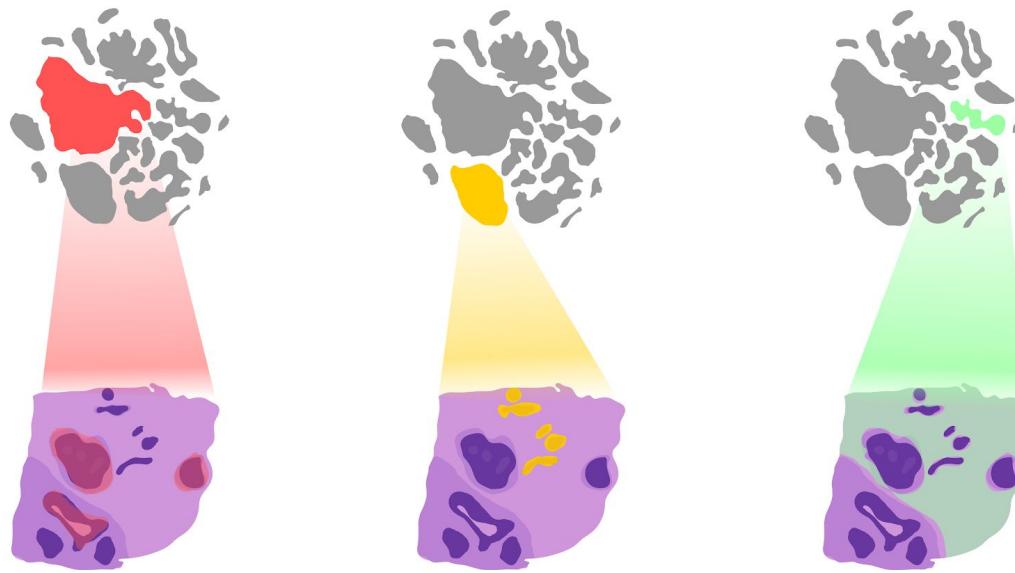
$... + W_z$





It's as easy as 1-2-3

3. Map cell type proportions back onto the tissue



Slightly more complex than adding bar graphs..

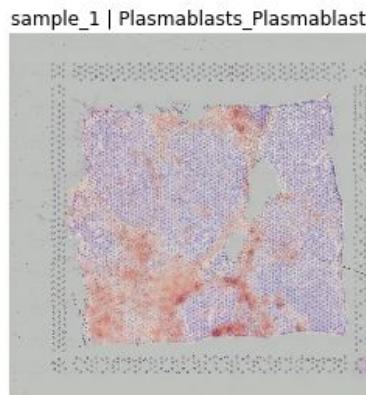
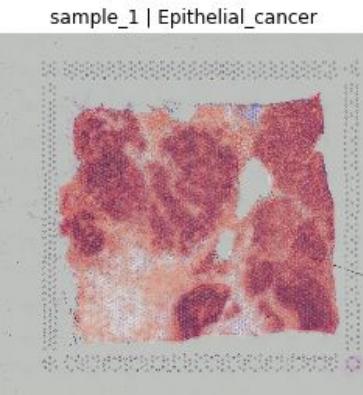
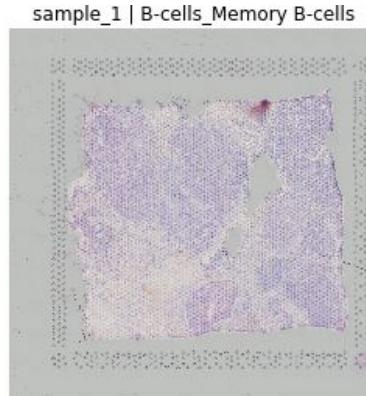
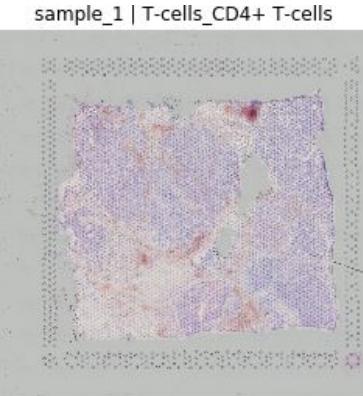
- Probabilistic Model
 - Negative Binomial distribution
 - Stochastic optimization (PyTorch)
- Output : [spot] x [cell_type] matrix
 - Elements are proportion of cell belonging to the given cell type at each spot
- Tool : *stereoscope*



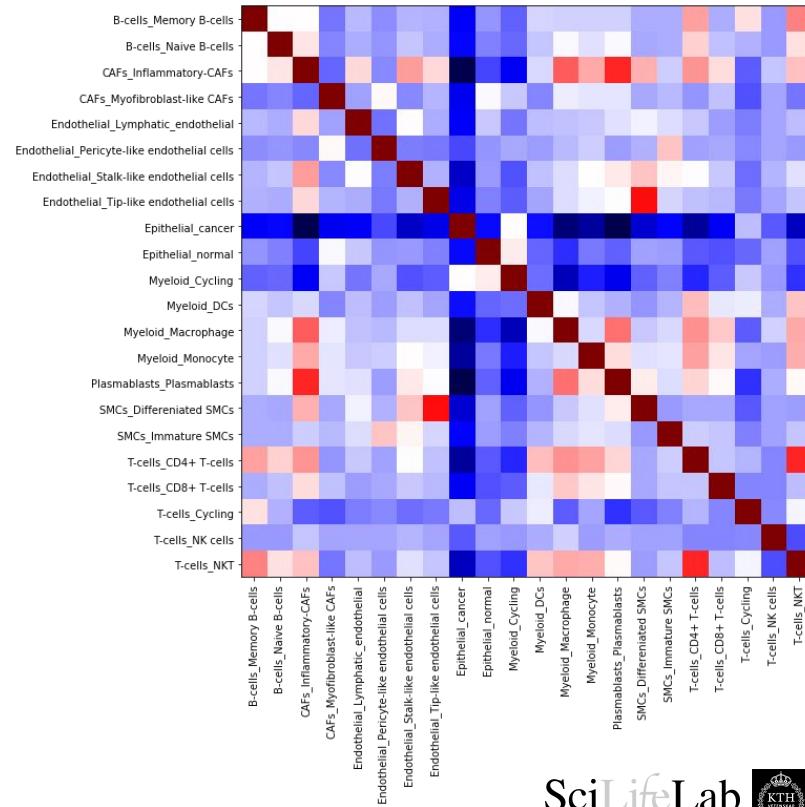
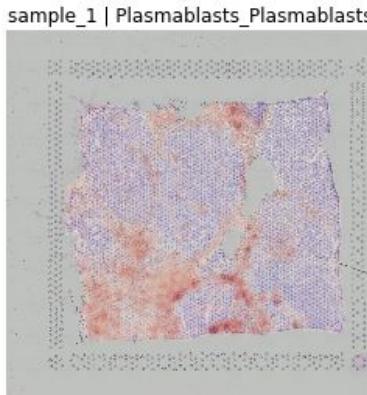
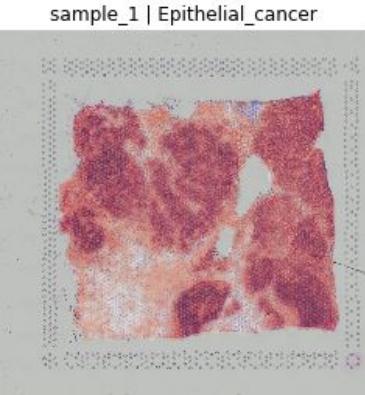
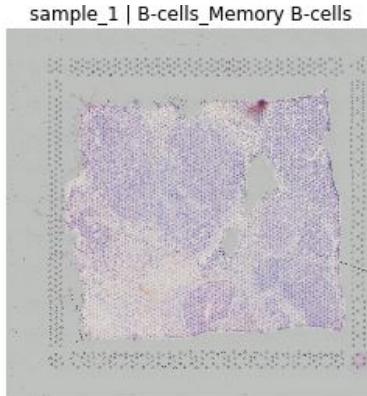
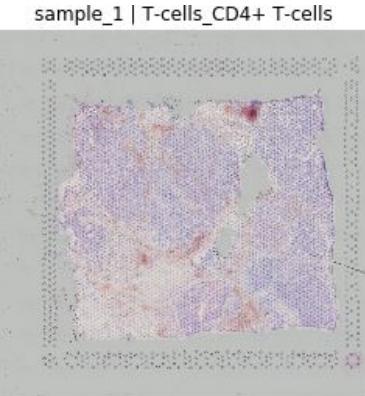
<https://github.com/almaan/stereoscope>

#shameless self-advertising

Applying it to our breast cancer data

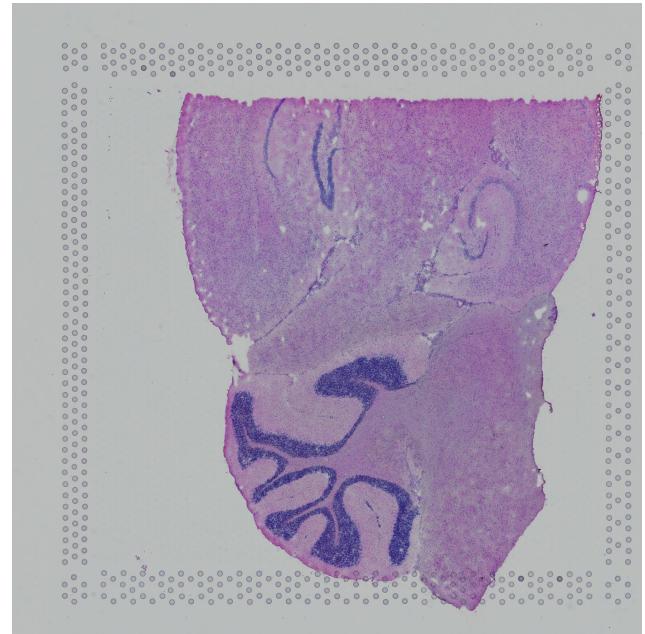


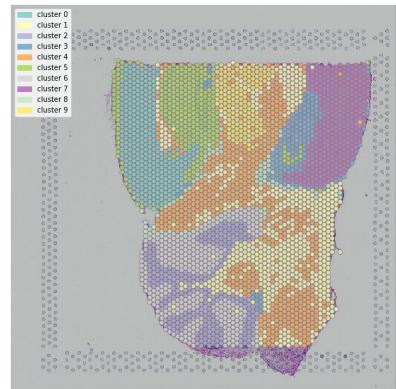
Applying it to our breast cancer data





More Spatially specific analysis







Exercise Session information

- Ambitions:
 - Getting familiar with spatial data
 - Exposure to different ways spatial data can be analyzed
 - Work with Visium, but tried to keep “general”
 - More Focus on concepts than maybe fine-tuning creating the perfect analysis
- Three Parts
 - Part 1 - “*Getting Comfy with Spatial Data*”
 - Loading, Inspecting and visualizing spatial data
 - General workflow
 - Part 2 - “



Thank you for the attention