



Sergio Sarnataro, PhD

Bioinformatics Awareness Days
July 13th, 2022

***Spatial Transcriptomics:
Technology and data analysis***

Overview

- Why Spatial Transcriptomics?
- Different kind of techniques
- Historical development of Spatial Transcriptomics technologies
- Focus on Visium and MERFISH
- Computational workflow and examples

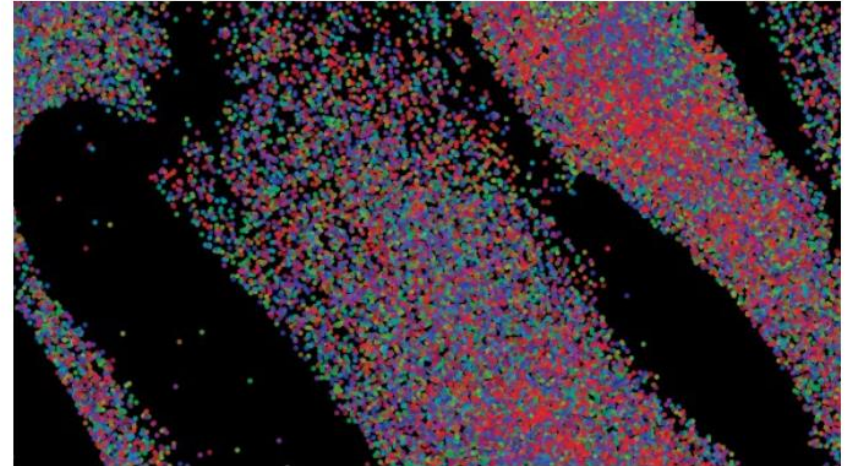
Method of the year 2020 according to Nature

nature > nature methods > focus

Focus | 06 January 2021

Method of the Year 2020: spatially resolved transcriptomics

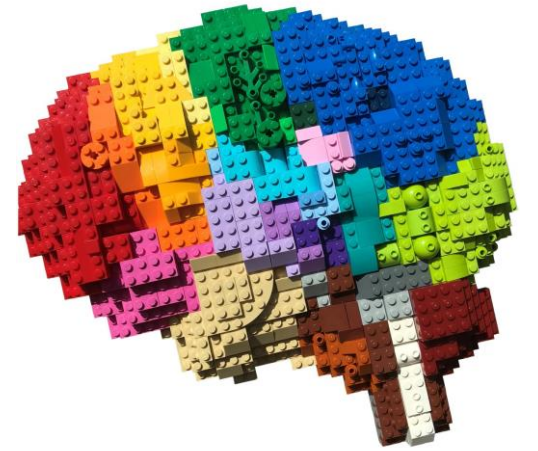
Spatially resolved transcriptomics is our Method of the Year 2020, for its ability to provide valuable insights into the biology of cells and tissues while retaining information about spatial context.



”Understanding the organization of cells and tissues and how this organization influences function is a fundamental pursuit in life sciences research”

Spatial Transcriptomics vs other techniques

Original organ

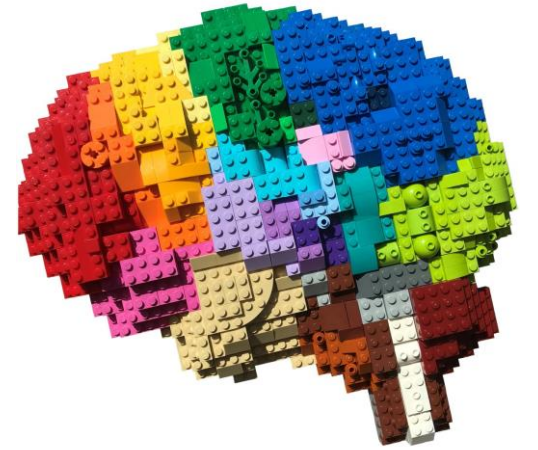


Spatial Transcriptomics vs other techniques

Bulk RNA-Seq



Original organ



Spatial Transcriptomics vs other techniques

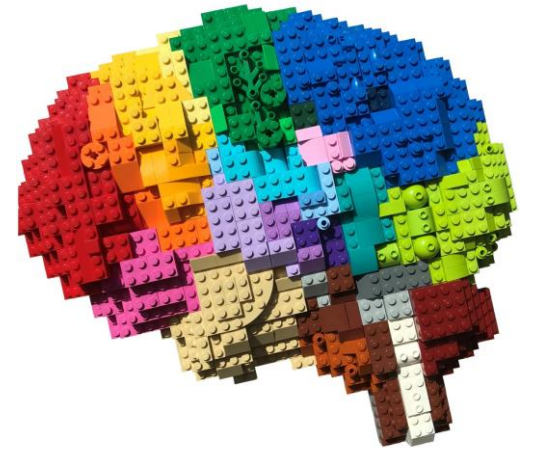
Bulk RNA-Seq



Single-cell RNA-Seq



Original organ



Spatial Transcriptomics vs other techniques

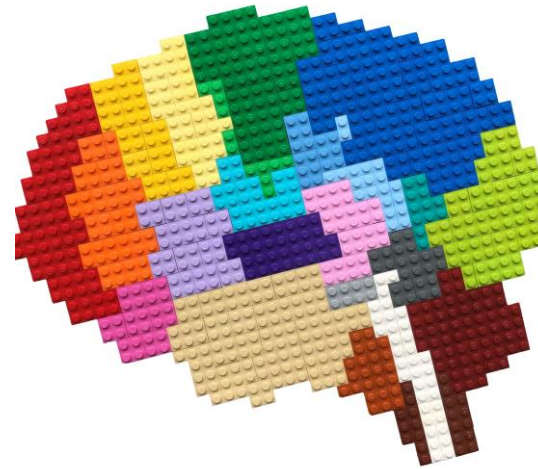
Bulk RNA-Seq



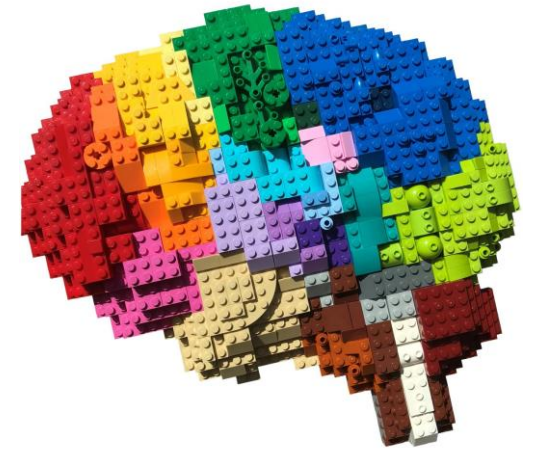
Single-cell RNA-Seq



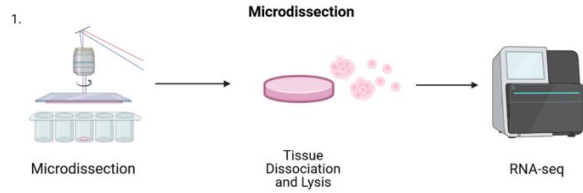
Spatial Transcriptomics



Original organ

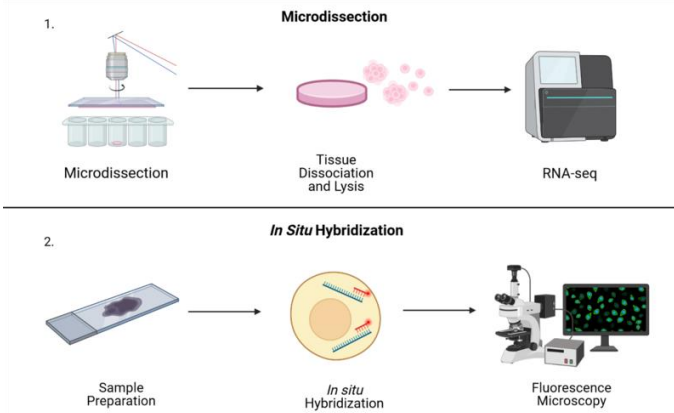


Different types of Spatial Transcriptomics



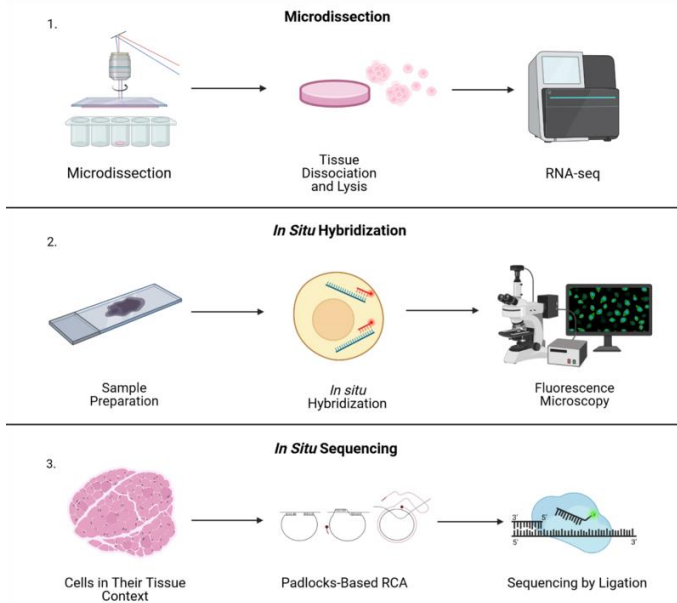
- Tissues are dissociated and the different parts are then sequenced

Different types of Spatial Transcriptomics



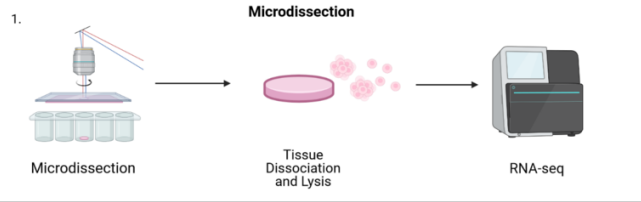
- Tissues are dissociated and the different parts are then sequenced
- Use of fluorescence microscopy to tag different transcripts inside cells. Example: smFISH and MERFISH

Different types of Spatial Transcriptomics

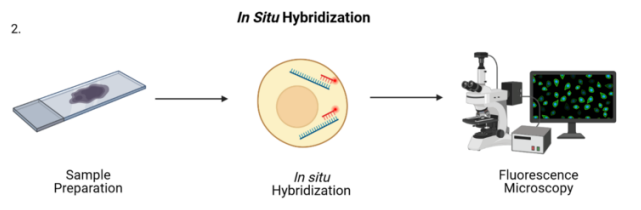


- Tissues are dissociated and the different parts are then sequenced
- Use of fluorescence microscopy to tag different transcripts inside cells. Example: smFISH and MERFISH
- ISS methods yield spatial transcriptome information by sequencing, typically by ligation, gene barcodes, or short fragments of cDNAs in situ.

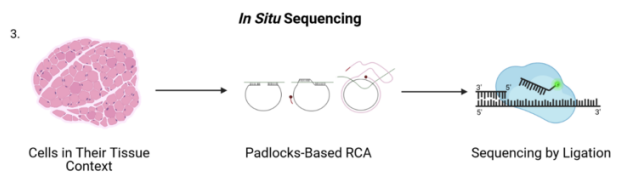
Different types of Spatial Transcriptomics



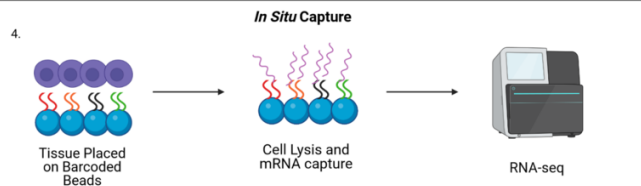
- Tissues are dissociated and the different parts are then sequenced



- Use of fluorescence microscopy to tag different transcripts inside cells. Example: smFISH and MERFISH

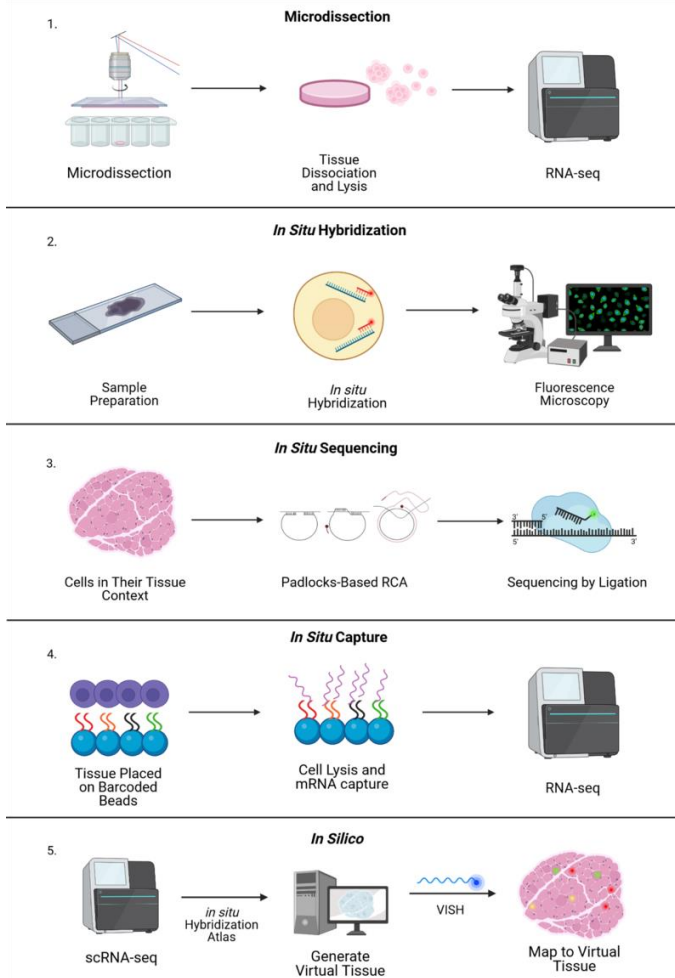


- ISS methods yield spatial transcriptome information by sequencing, typically by ligation, gene barcodes, or short fragments of cDNAs in situ.



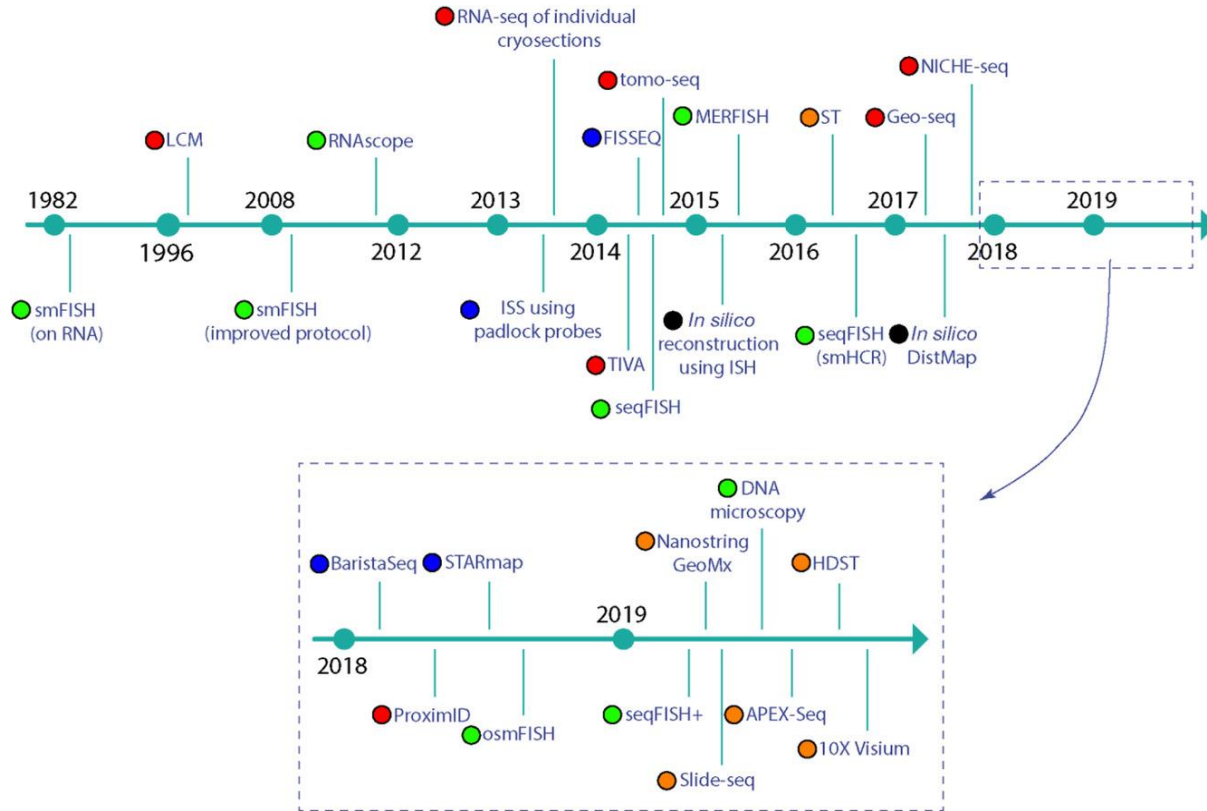
- Spatial locations of transcripts is preserved by capturing them from tissue sections on in situ array. Example: 10x Visium

Different types of Spatial Transcriptomics



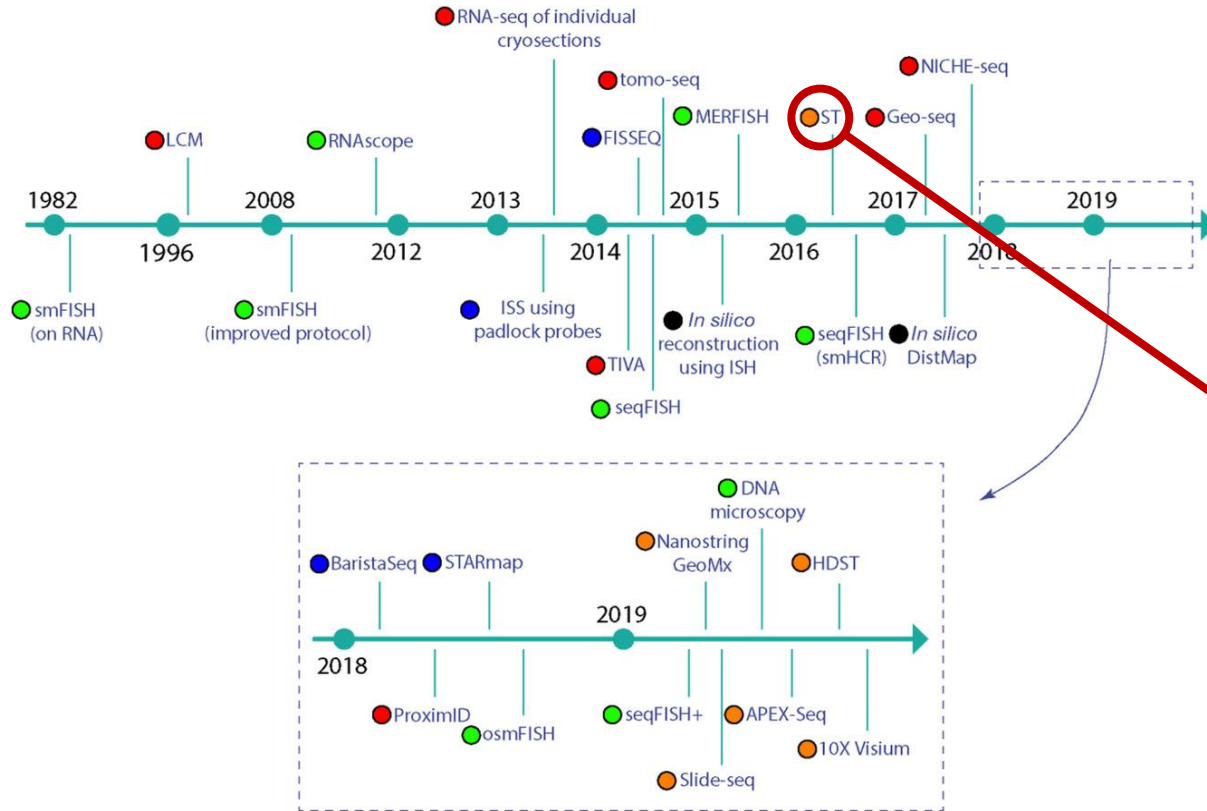
- Tissues are dissociated and the different parts are then sequenced
- Use of fluorescence microscopy to tag different transcripts inside cells. Example: smFISH and MERFISH
- ISS methods yield spatial transcriptome information by sequencing, typically by ligation, gene barcodes, or short fragments of cDNAs in situ.
- Spatial locations of transcripts is preserved by capturing them from tissue sections on in situ array. Example: 10x Visium
- Computational spatial reconstruction of cells' locations according to their expression profiles

Historical overview



- Section 1. Technologies based on microdissected gene expression
- Section 2. *In situ* hybridization technologies
- Section 3. *In situ* sequencing technologies
- Section 4. *In situ* capturing technologies
- Section 5. *In silico* reconstruction of spatial data

Historical overview



- Section 1. Technologies based on microdissected gene expression
- Section 2. *In situ* hybridization technologies
- Section 3. *In situ* sequencing technologies
- Section 4. *In situ* capturing technologies
- Section 5. *In silico* reconstruction of spatial data

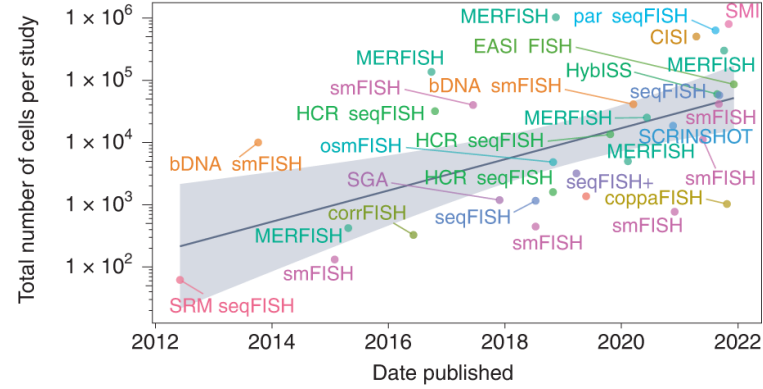
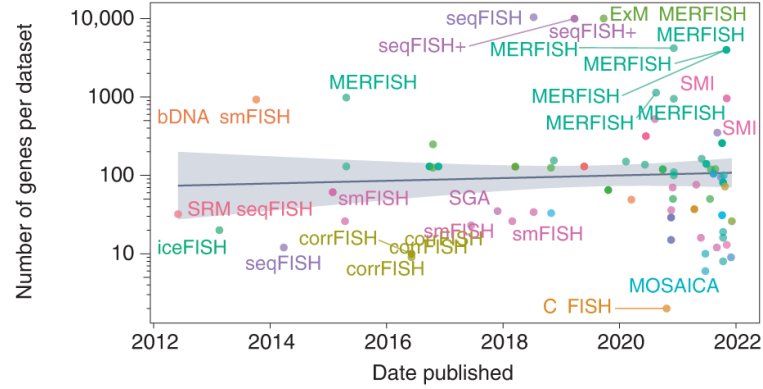
TRANSCRIPTION

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

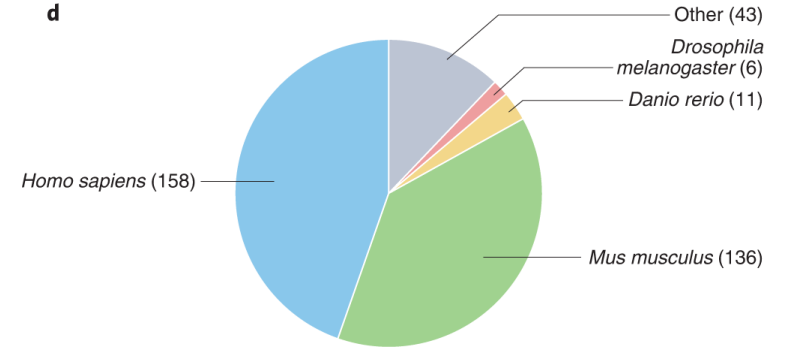
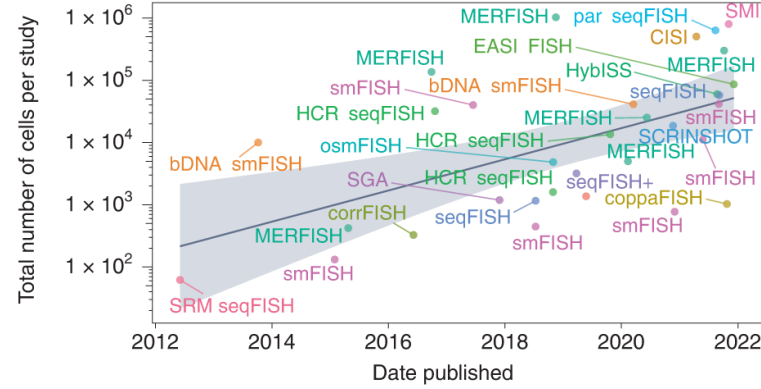
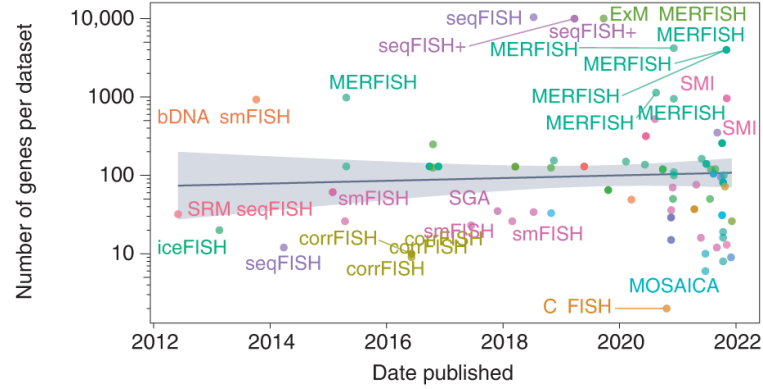
Patrik L. Ståhl,^{1,2*} Fredrik Salmén,^{2*} Sanja Vickovic,^{2†} Anna Lundmark,^{2,3†} José Fernández Navarro,^{1,2} Jens Magnusson,¹ Stefania Giacomello,² Michaela Asp,² Jakub O. Westholm,⁴ Mikael Huss,⁴ Annelie Mollbrink,² Sten Linnarsson,⁵ Simone Codeluppi,^{5,6} Åke Borg,⁷ Fredrik Pontén,⁸ Paul Igor Costea,² Pelin Sahlén,² Jan Mulder,⁹ Olaf Bergmann,¹ Joakim Lundeberg,^{2†} Jonas Frisén¹

Analysis of the pattern of proteins or messenger RNAs (mRNAs) in histological tissue sections is a cornerstone in biomedical research and diagnostics. This typically involves the visualization of a few proteins or expressed genes at a time. We have devised a strategy, which we call “spatial transcriptomics,” that allows visualization and quantitative analysis of the transcriptome with spatial resolution in individual tissue sections. By positioning histological sections on arrayed reverse transcription primers with unique positional barcodes, we demonstrate high-quality RNA-sequencing data with maintained two-dimensional positional information from the mouse brain and human breast cancer. Spatial transcriptomics provides quantitative gene expression data and visualization of the distribution of mRNAs within tissue sections and enables novel types of bioinformatics analyses, valuable in research and diagnostics.

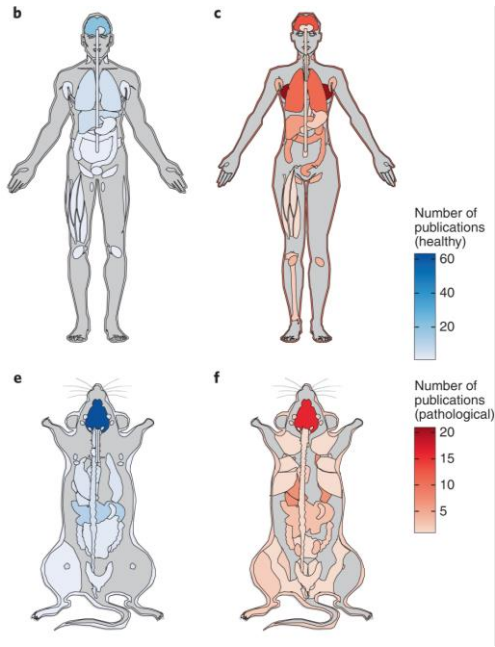
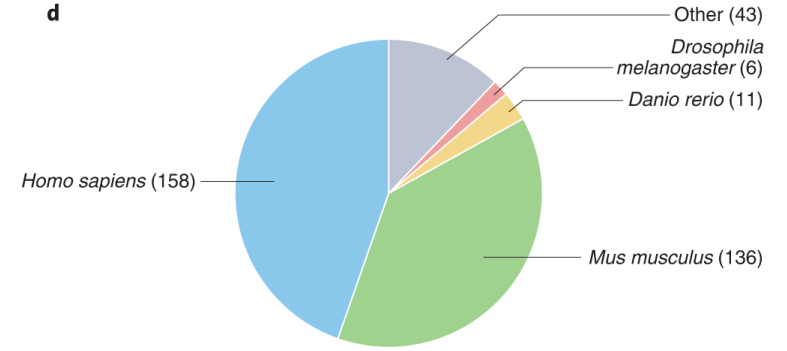
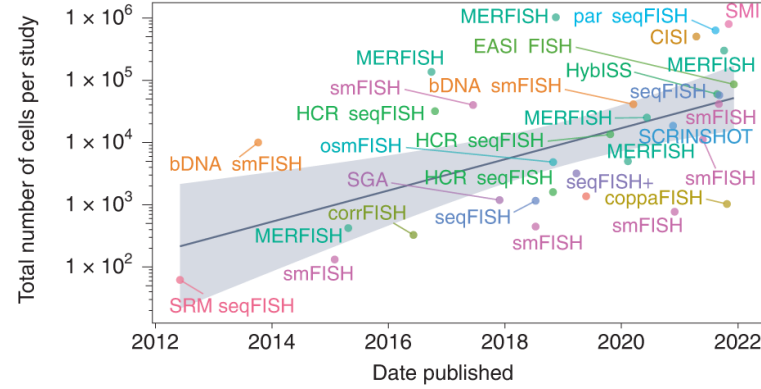
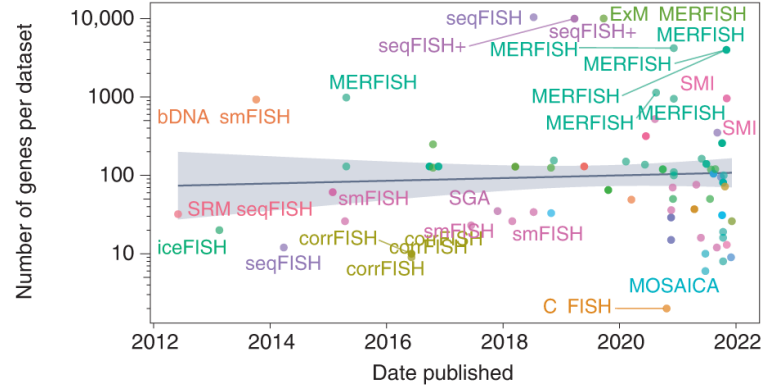
Historical development and state of the art



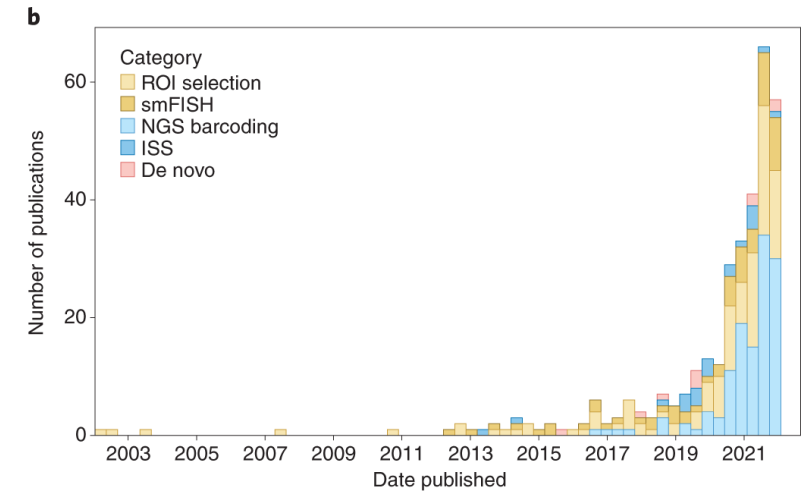
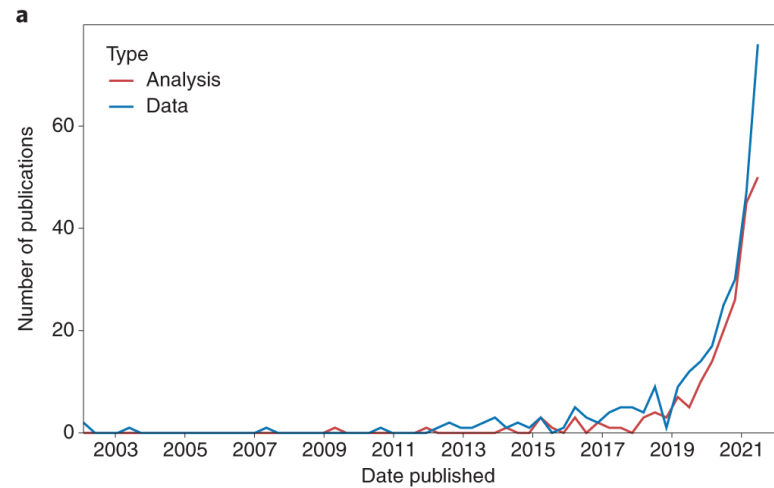
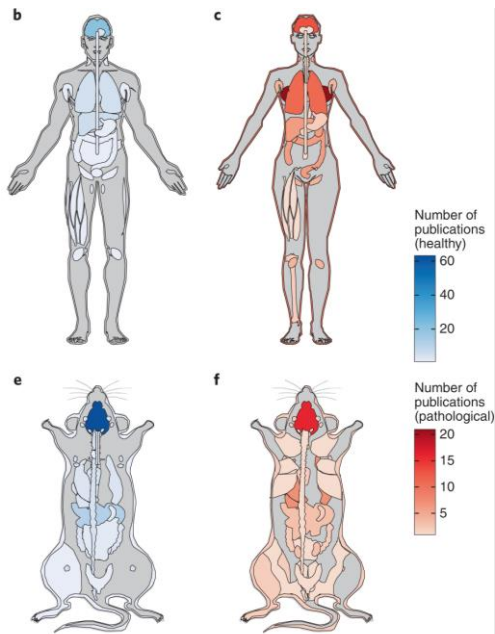
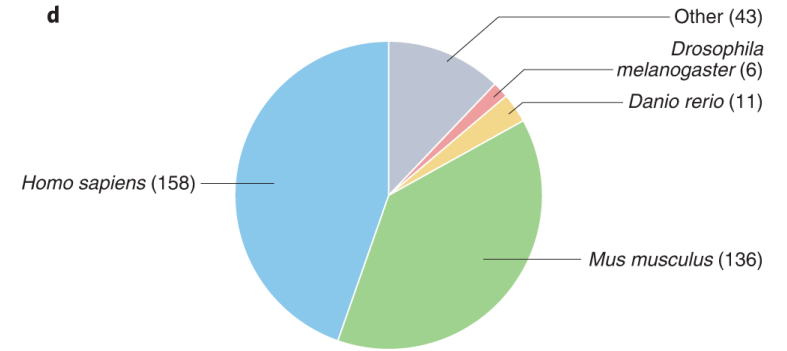
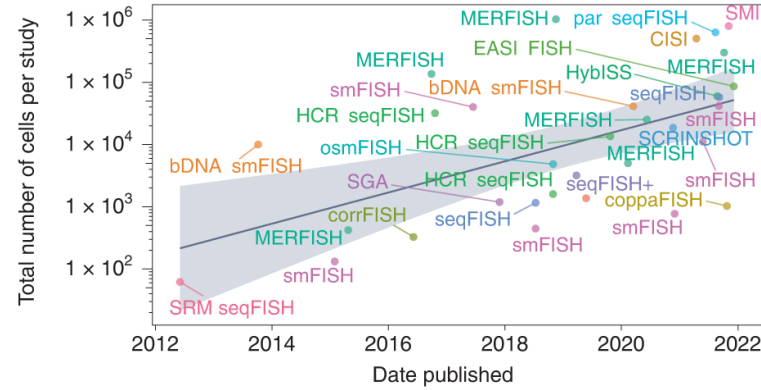
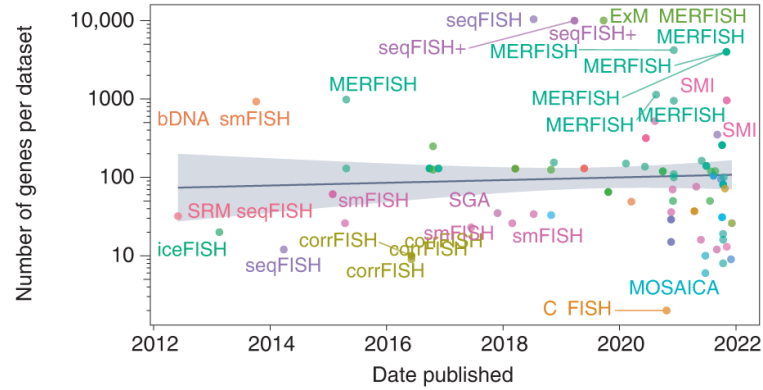
Historical development and state of the art



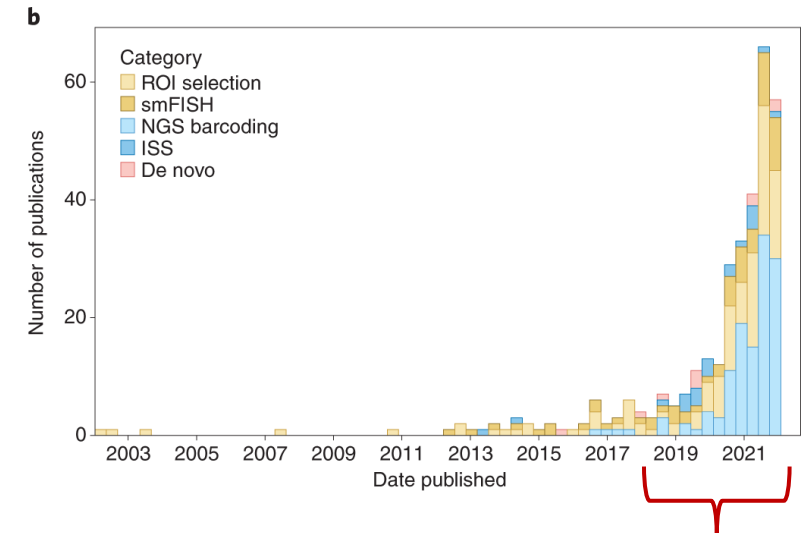
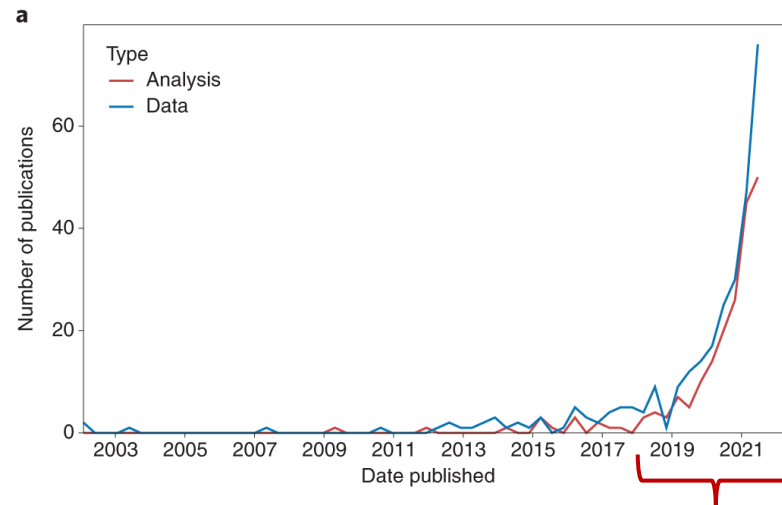
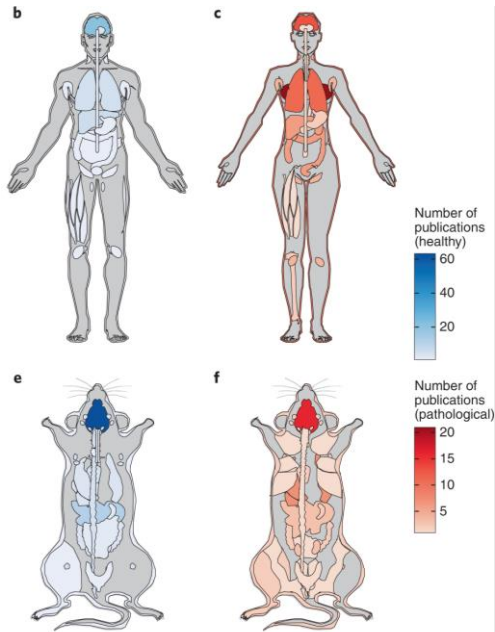
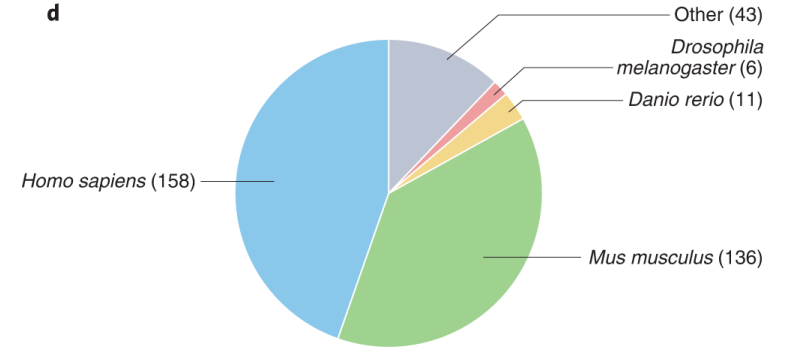
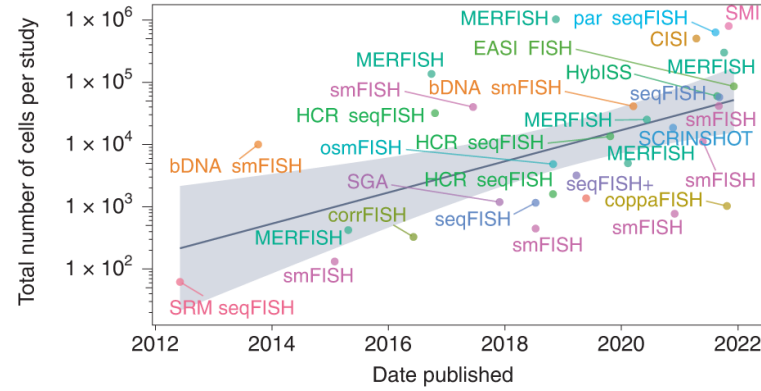
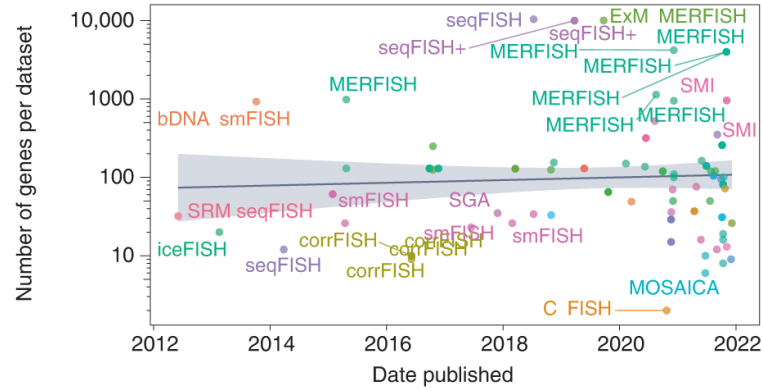
Historical development and state of the art



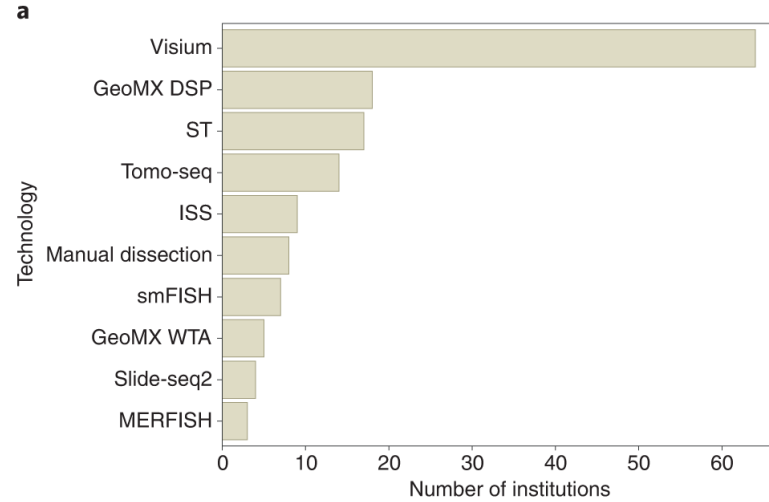
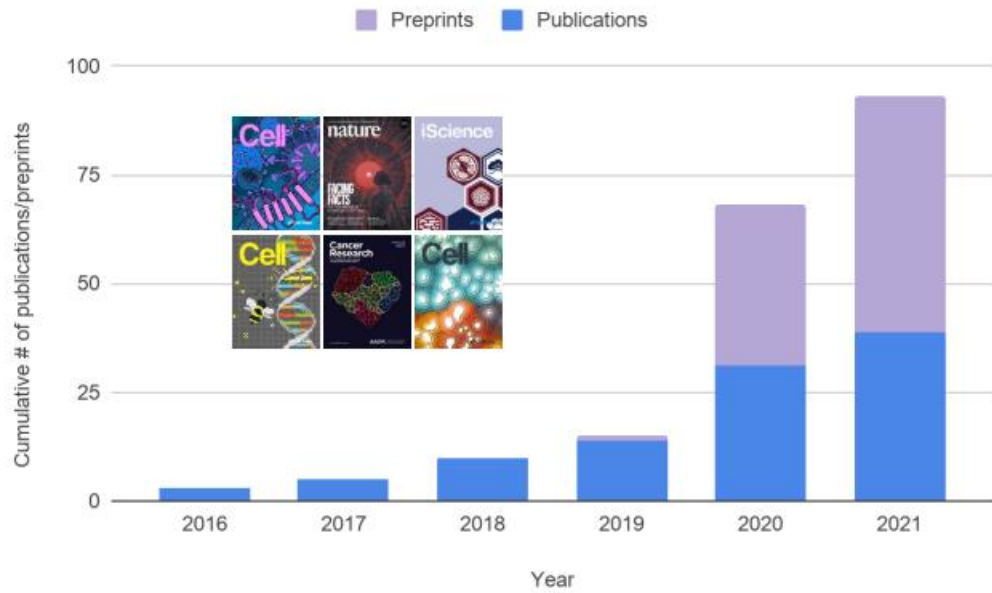
Historical development and state of the art



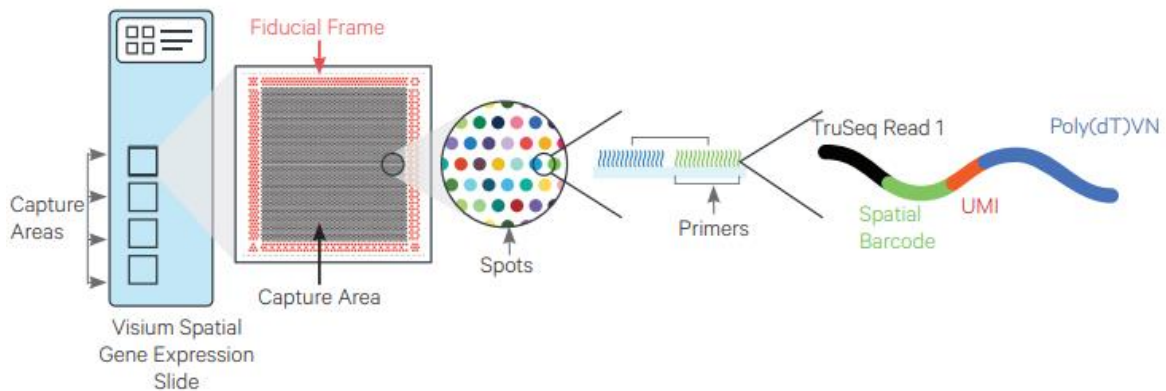
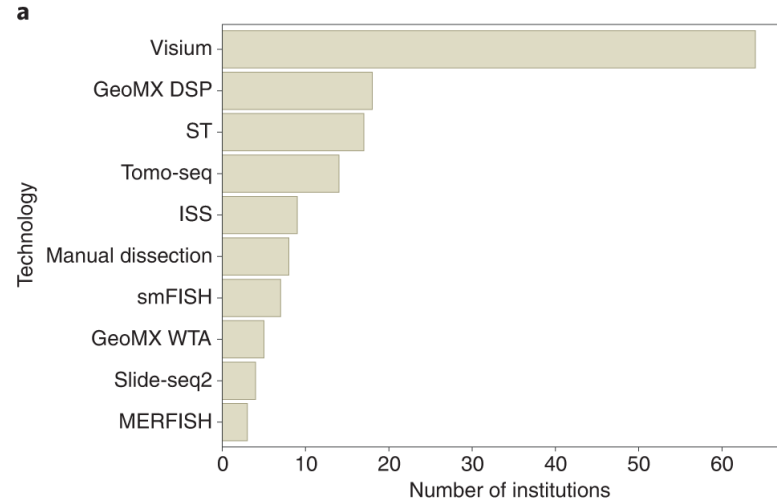
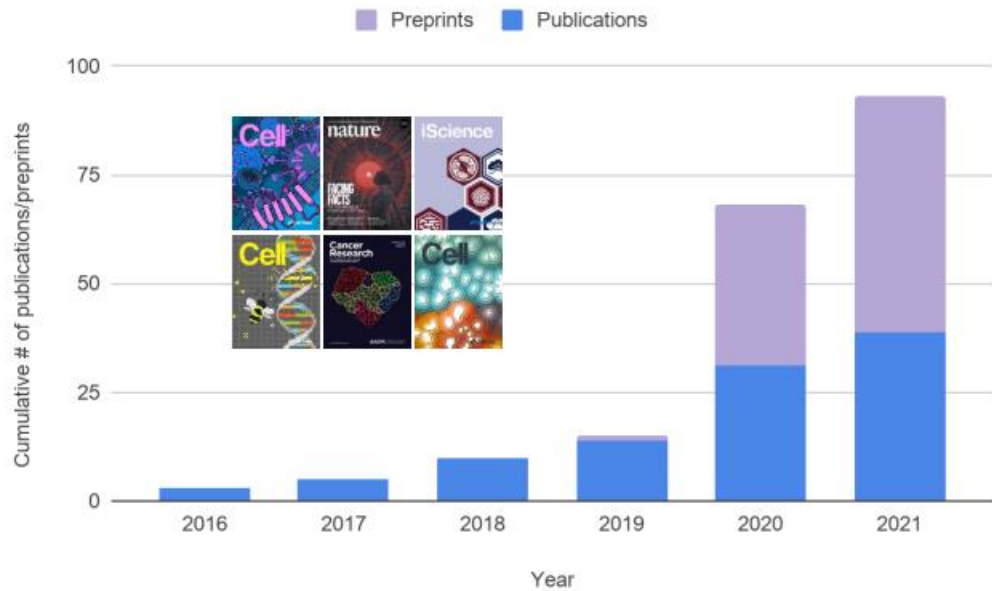
Historical development and state of the art



Focus on 10x Visium



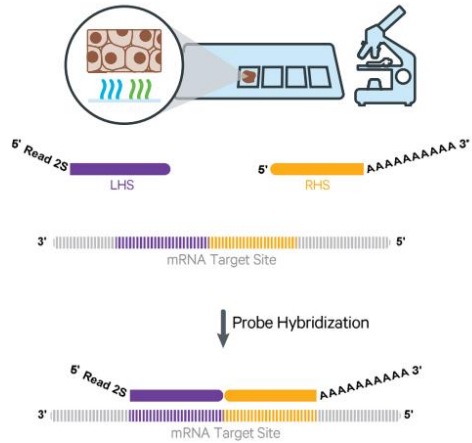
Focus on 10x Visium



- 5000 spots per area
- Order of millions of oligos per spot
- 18,000 unique genes detectable

10x Visium Technical overview

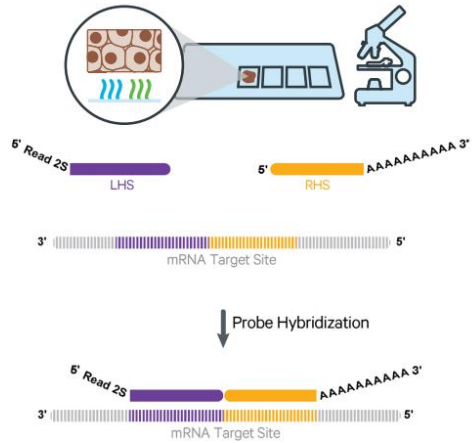
1. Probe hybridization



The human or mouse whole transcriptome probe panel, consisting of a pair of specific probes for each targeted gene, is added to the deparaffinized, stained, and decrosslinked tissues. Together, probe pairs hybridize to their complementary target RNA.

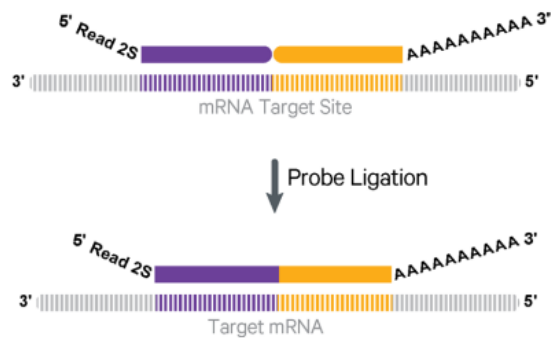
10x Visium Technical overview

1. Probe hybridization



The human or mouse whole transcriptome probe panel, consisting of a pair of specific probes for each targeted gene, is added to the deparaffinized, stained, and decrosslinked tissues. Together, probe pairs hybridize to their complementary target RNA.

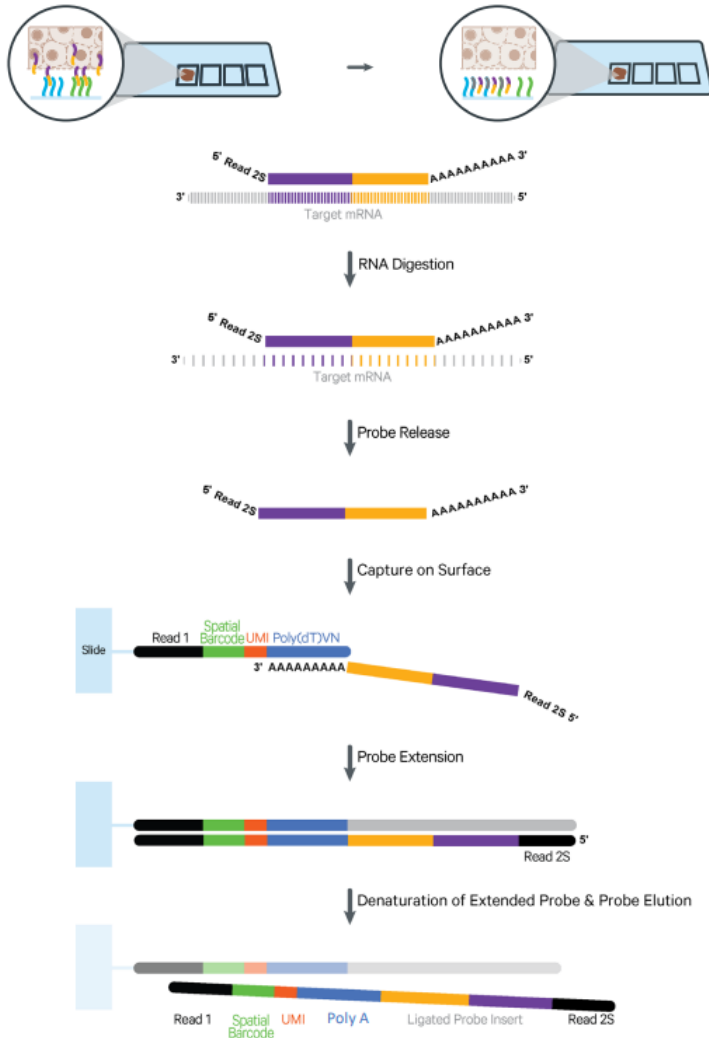
2. Probe ligation



After hybridization, a ligase is added to seal the junction between the probe pairs that have hybridized to RNA, forming a ligation product.

10x Visium Technical overview

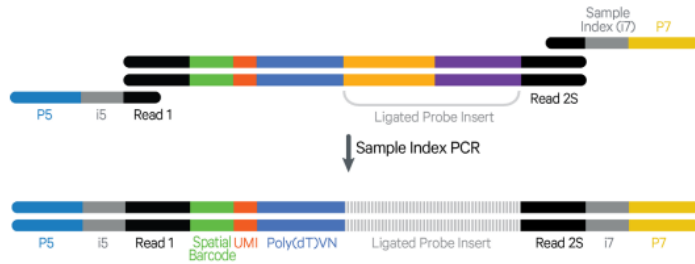
3. Probe release & extension



The single stranded ligation products are released from the tissue upon RNase treatment and permeabilization, and then captured on the Visium slides. Once ligation products are captured, probes are extended by the addition of UMI, Spatial Barcode and partial Read 1. This generates spatially barcoded, ligated probe products, which can then be carried forward for library preparation.

10x Visium Technical overview

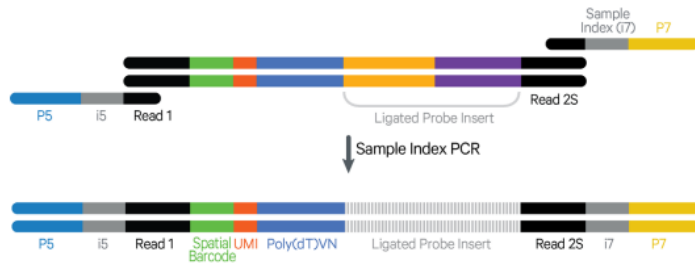
4. Library construction



The spatially barcoded, ligated probe products are released from the slide and harvested for qPCR to determine Sample Index PCR cycle number. The products then undergo indexing via Sample Index PCR. This, in turn, generates final library molecules that are cleaned up by SPRIselect, assessed on a bioanalyzer or a similar instrument, quantified, and then sequenced

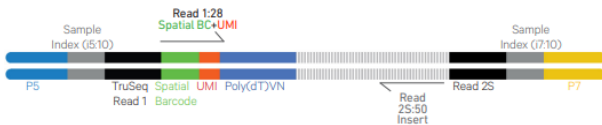
10x Visium Technical overview

4. Library construction



The spatially barcoded, ligated probe products are released from the slide and harvested for qPCR to determine Sample Index PCR cycle number. The products then undergo indexing via Sample Index PCR. This, in turn, generates final library molecules that are cleaned up by SPRIselect, assessed on a bioanalyzer or a similar instrument, quantified, and then sequenced

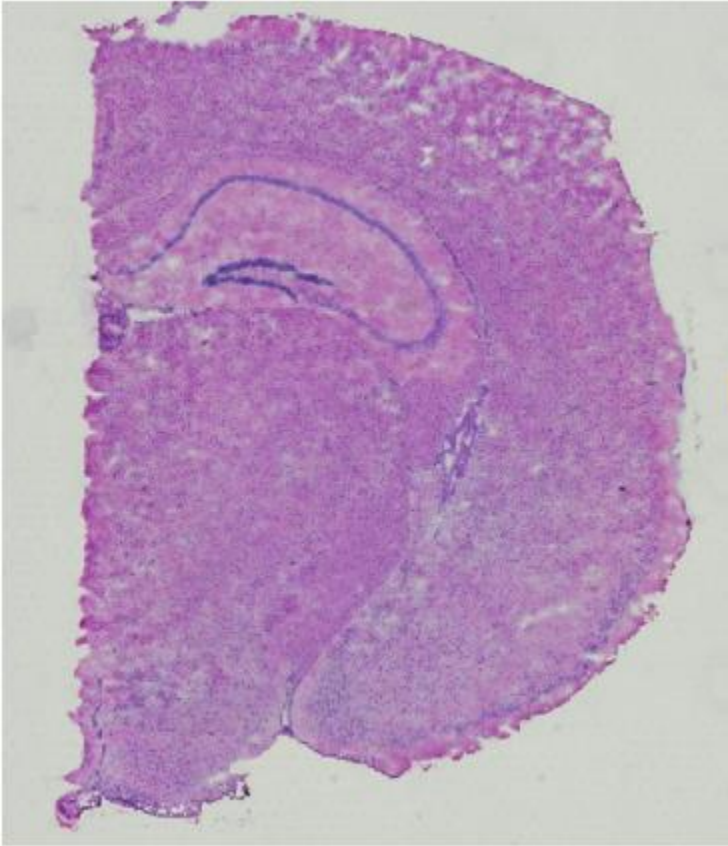
5. Sequencing



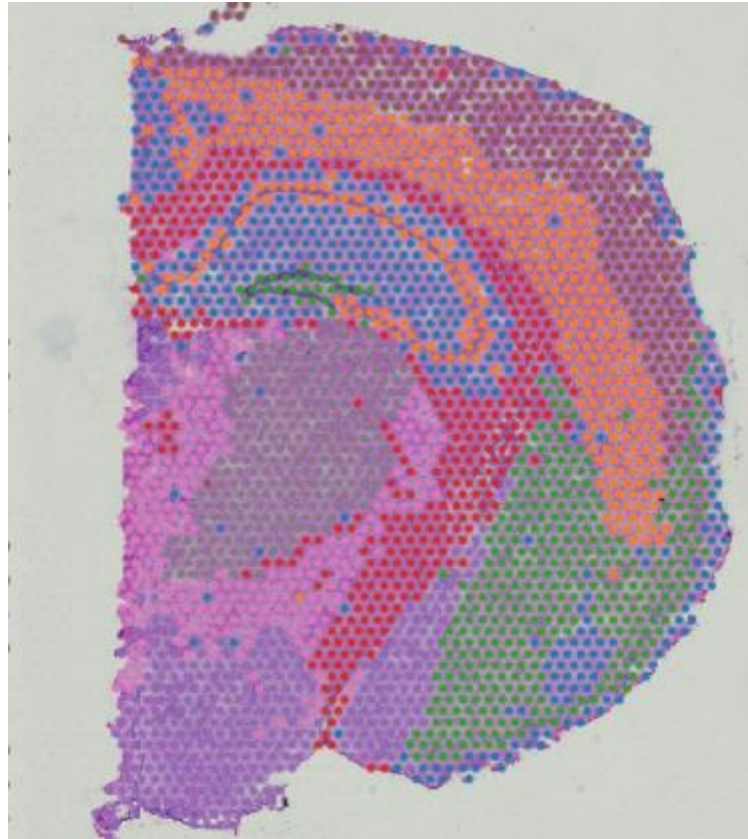
A Visium Spatial Gene Expression – FFPE library comprises standard Illumina paired-end constructs which begin and end with P5 and P7. The 16 bp Spatial Barcode and 12 bp UMI are encoded in Read 1, while Small RNA Read 2 (Read 2S) is used to sequence the ligated probe insert. Illumina sequencer compatibility, sample indices, library loading and pooling for sequencing are summarized in step 5.

Information contained in 10x Visium data

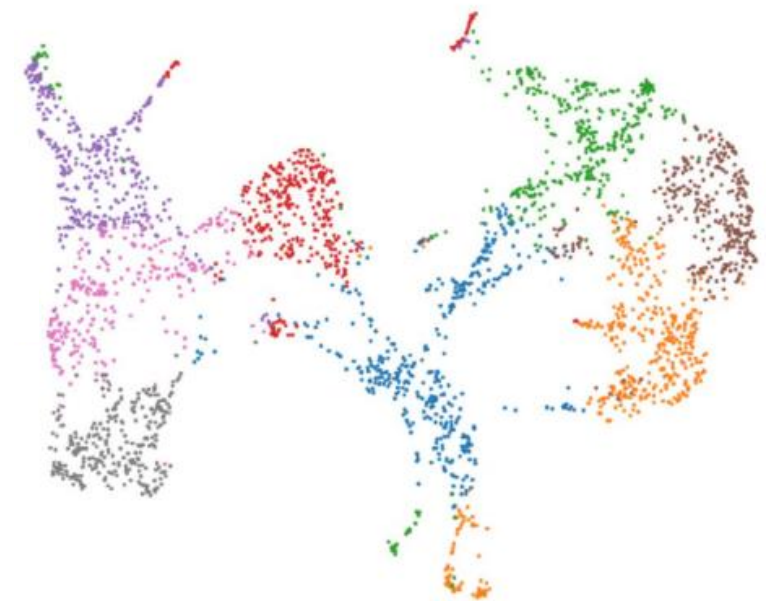
Histology image



Histology + gene expression



Gene expression



10x Visium data preprocessing

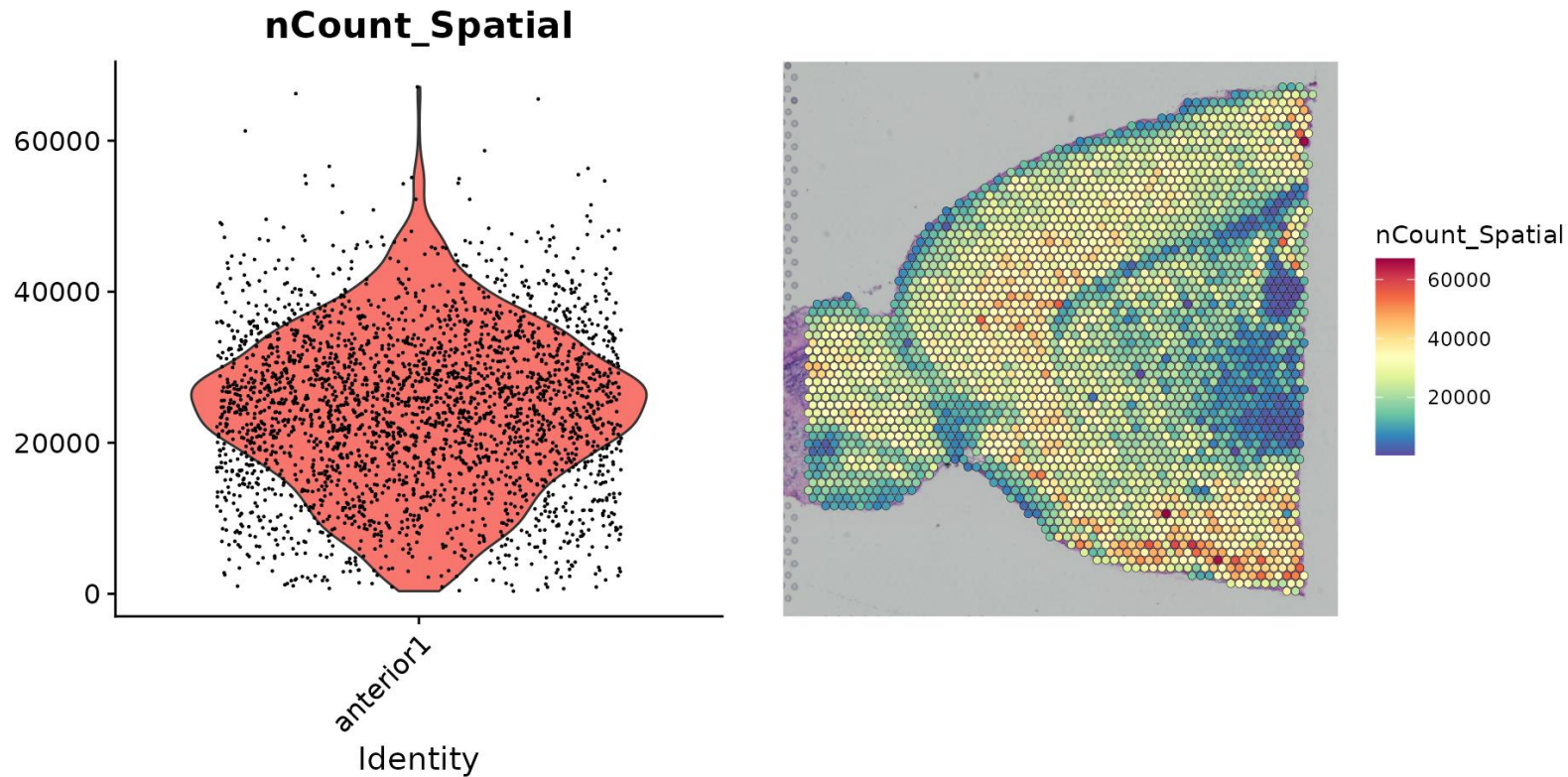
Alignment to a reference genome. Usually, it is done by using spaceranger tool (of note: a very similar tool, cellranger, has been developed for 10x single-cell)

Alignment script example

```
#SBATCH --job-name=cca_13_SP_alignment # Job name
#SBATCH --ntasks=1
#SBATCH --mem=80gb # Job memory request
#SBATCH --cpus-per-task 10
#SBATCH --output=cca_13_SP.log # Standard output and error log

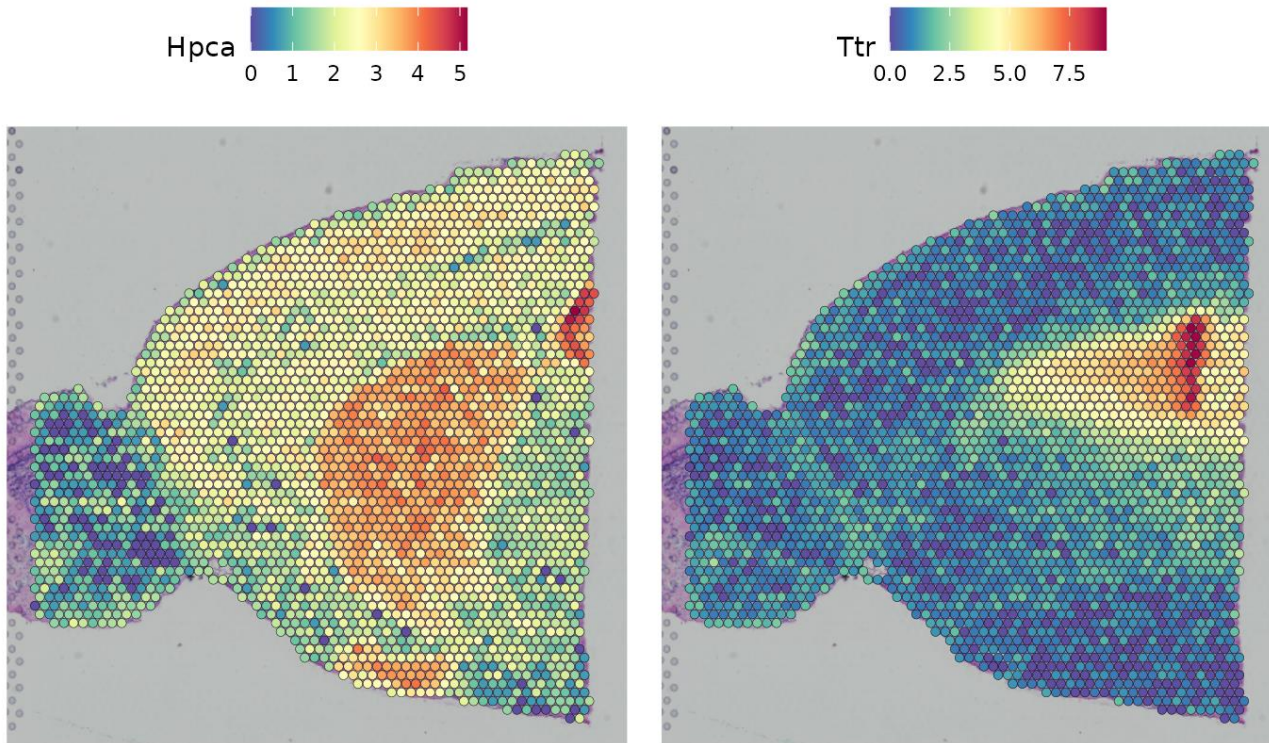
/opt/spaceranger-1.3.1/spaceranger count --id=cca_13 \
--transcriptome=/home/tigem/s.sarnataro/spaceranger_reference/refdata-gex-GRCh38-2020-A \
--probe-set=/home/tigem/s.sarnataro/spaceranger_reference/Visium_Human_Transcriptome_Probe_Set_v1.0_GRCh38-2020-A.csv \
--fastqs=/home/tigem/s.sarnataro/gbioinfo/Carotenuto/spatial_transcriptomics_2022 \
--sample=cca_13_SP \
--image=/home/tigem/s.sarnataro/gbioinfo/Carotenuto/spatial_transcriptomics_2022/CCA_SPATIAL_IMAGE/CCA_SPATIAL_IMAGE/A2_anna_image_2.tif \
--area=B1 \
--localcores=8 \
--localmem=64 \
--slide=V11A20-288
```

10x Visium data analysis with Seurat



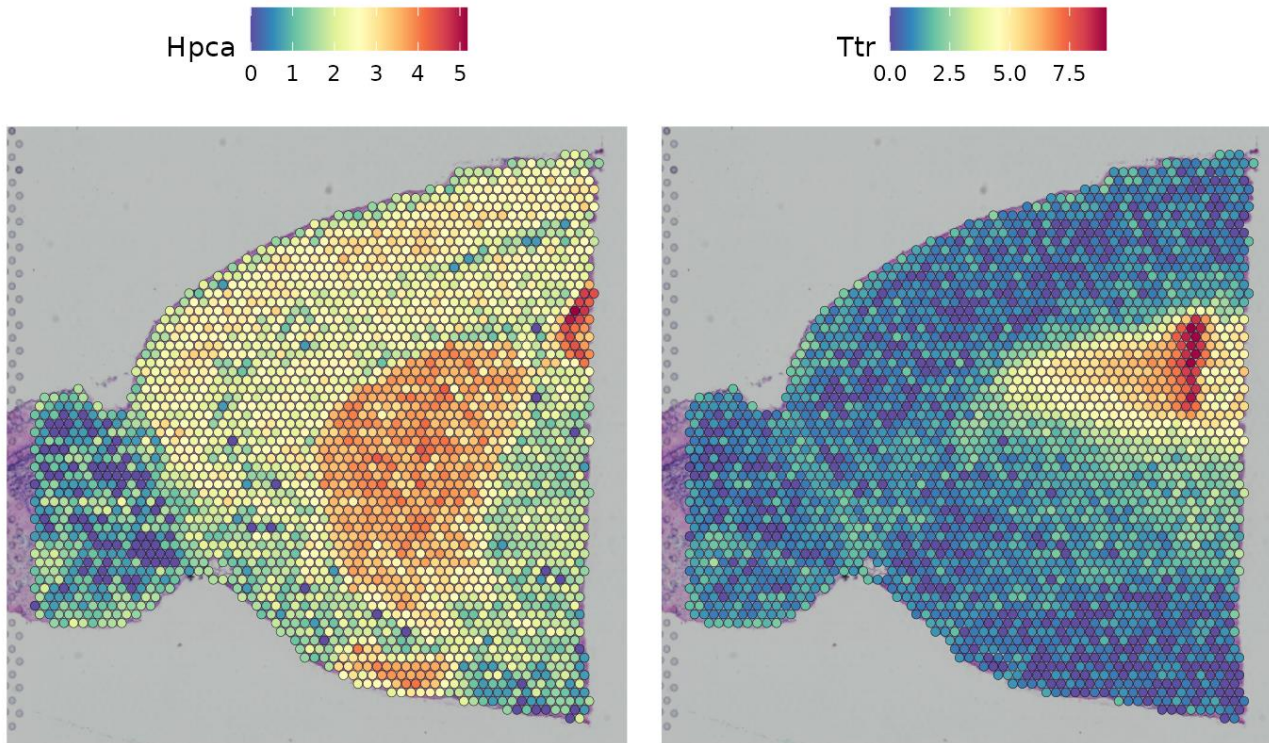
```
brain <- SCTransform(brain, assay = "Spatial", verbose = FALSE)
```


Gene expression visualization

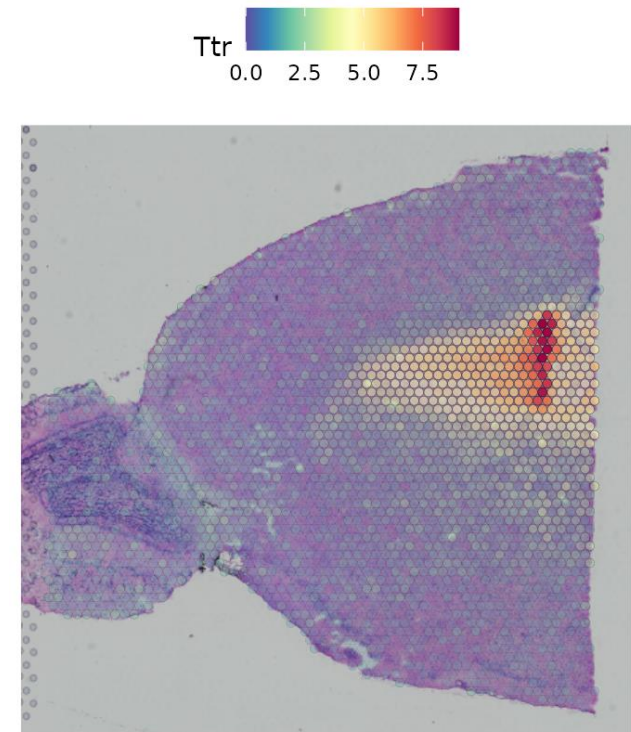


```
SpatialFeaturePlot(brain, features = c("Hpca", "Ttr"))
```

Gene expression visualization



```
SpatialFeaturePlot(brain, features = c("Hpca", "Ttr"))
```



```
SpatialFeaturePlot(brain, features = "Ttr", alpha = c(0.1, 1))
```

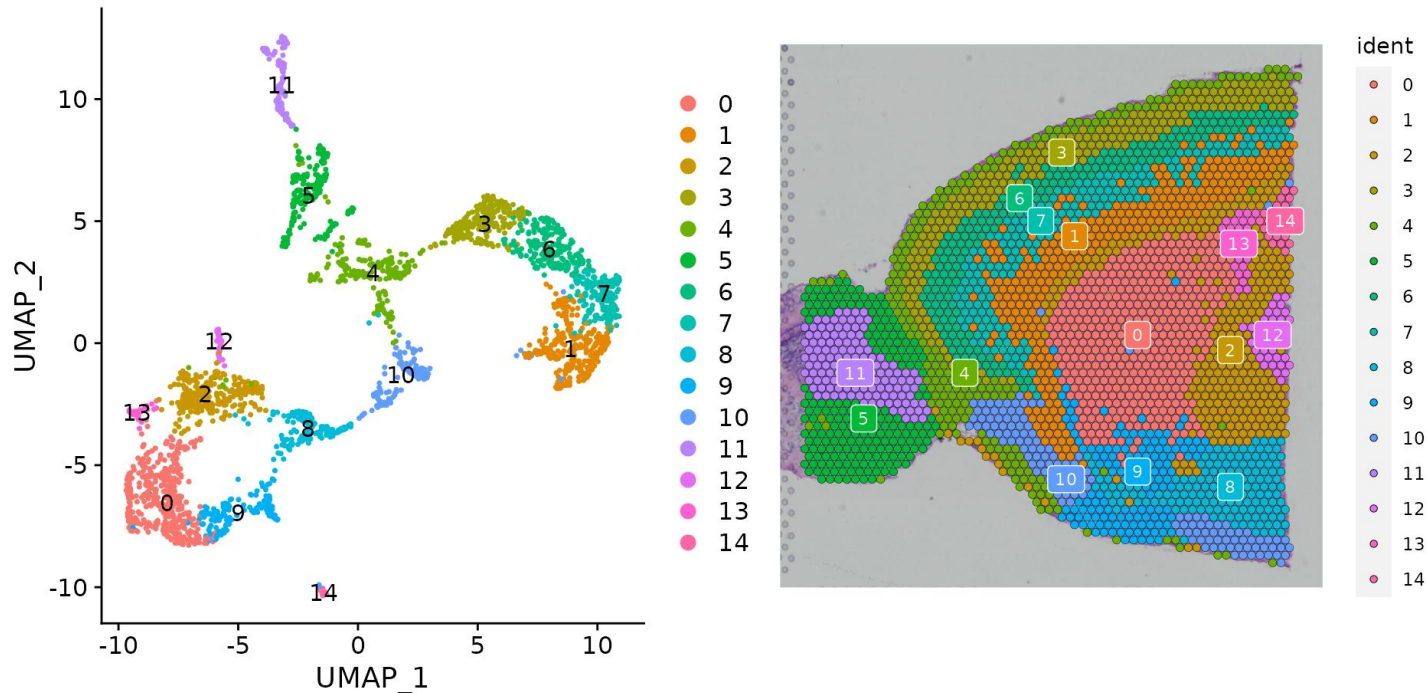
Dimensionality reduction, clustering and visualization

Dimensionality reduction and clustering

```
brain <- RunPCA(brain, assay = "SCT", verbose = FALSE)
brain <- FindNeighbors(brain, reduction = "pca", dims = 1:30)
brain <- FindClusters(brain, verbose = FALSE)
brain <- RunUMAP(brain, reduction = "pca", dims = 1:30)
```


Dimensionality reduction, clustering and visualization

Visualization

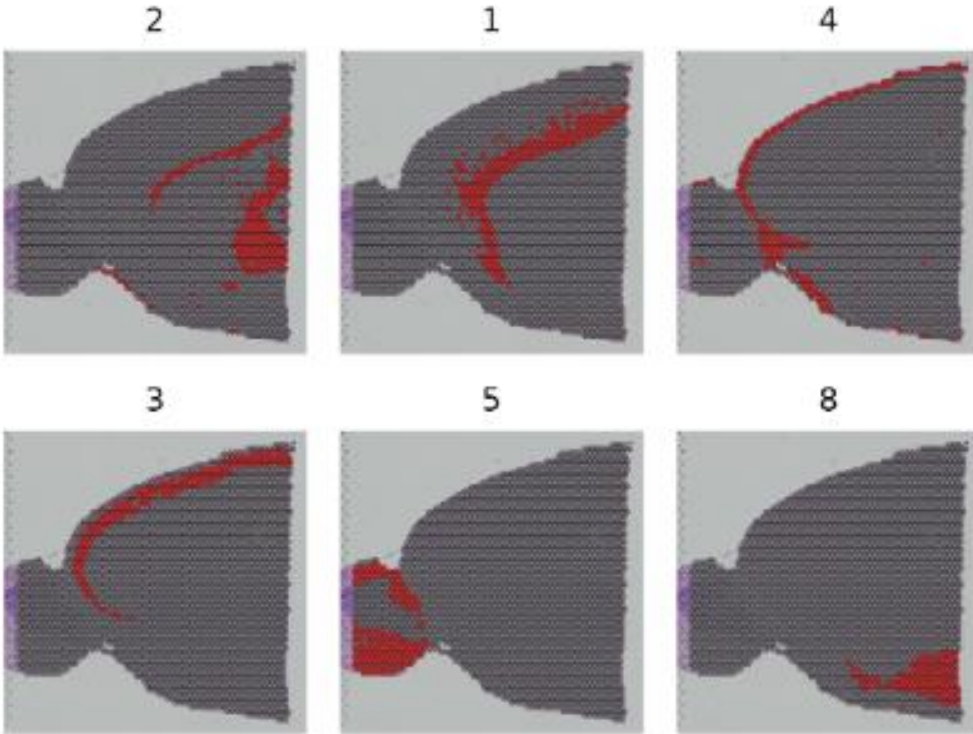


Dimensionality reduction and clustering

```
brain <- RunPCA(brain, assay = "SCT", verbose = FALSE)
brain <- FindNeighbors(brain, reduction = "pca", dims = 1:30)
brain <- FindClusters(brain, verbose = FALSE)
brain <- RunUMAP(brain, reduction = "pca", dims = 1:30)
```

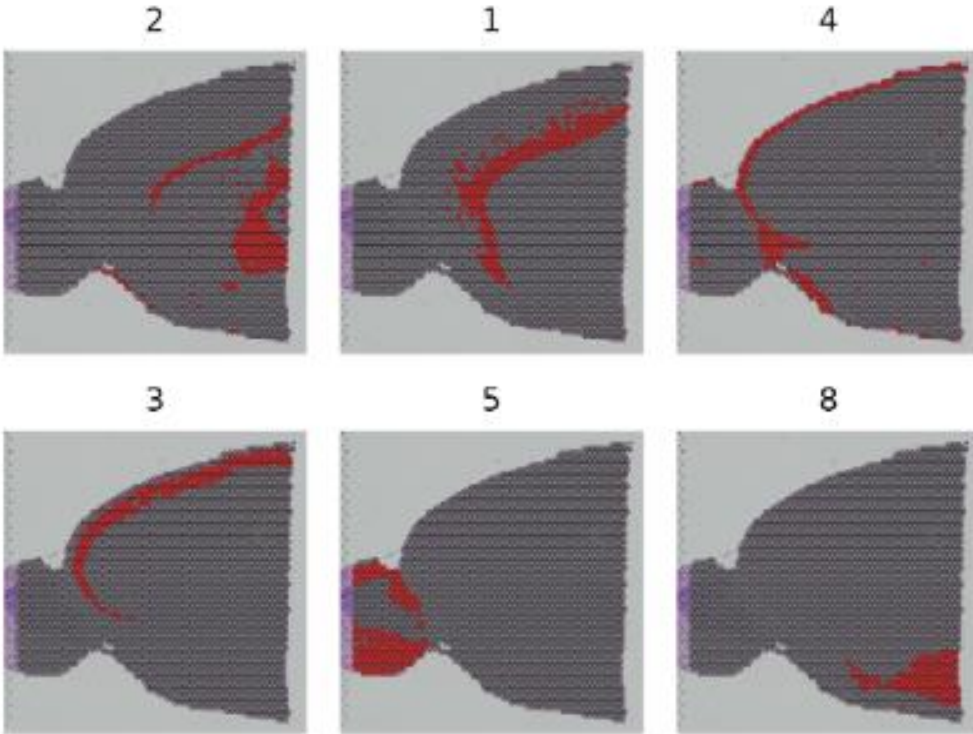
```
p1 <- DimPlot(brain, reduction = "umap", label = TRUE)
p2 <- SpatialDimPlot(brain, label = TRUE, label.size = 3)
p1 + p2
```

Single clusters visualization and DGE

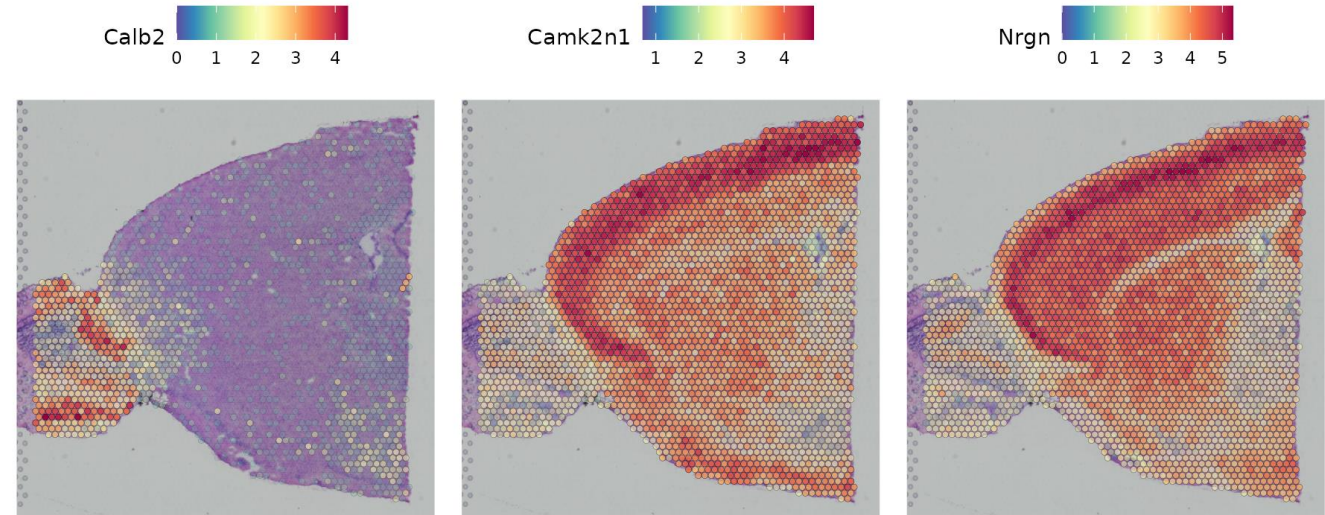


```
SpatialDimPlot(brain, cells.highlight = CellsByIdentities(object = brain, idents = c(2, 1, 4, 3, 5, 8)), facet.highlight = TRUE, ncol = 3)
```

Single clusters visualization and DGE



Identification of Spatially Variable Features



```
de_markers <- FindMarkers(brain, ident.1 = 5, ident.2 = 6)
SpatialFeaturePlot(object = brain, features = rownames(de_markers)[1:3], alpha = c(0.1, 1), ncol = 3)
```

```
SpatialDimPlot(brain, cells.highlight = CellsByIdentities(object = brain, ids = c(2, 1, 4, 3,
5, 8)), facet.highlight = TRUE, ncol = 3)
```


Typical orders of magnitude

2,264

Number of Spots Under Tissue

28,826

Mean Reads per Spot

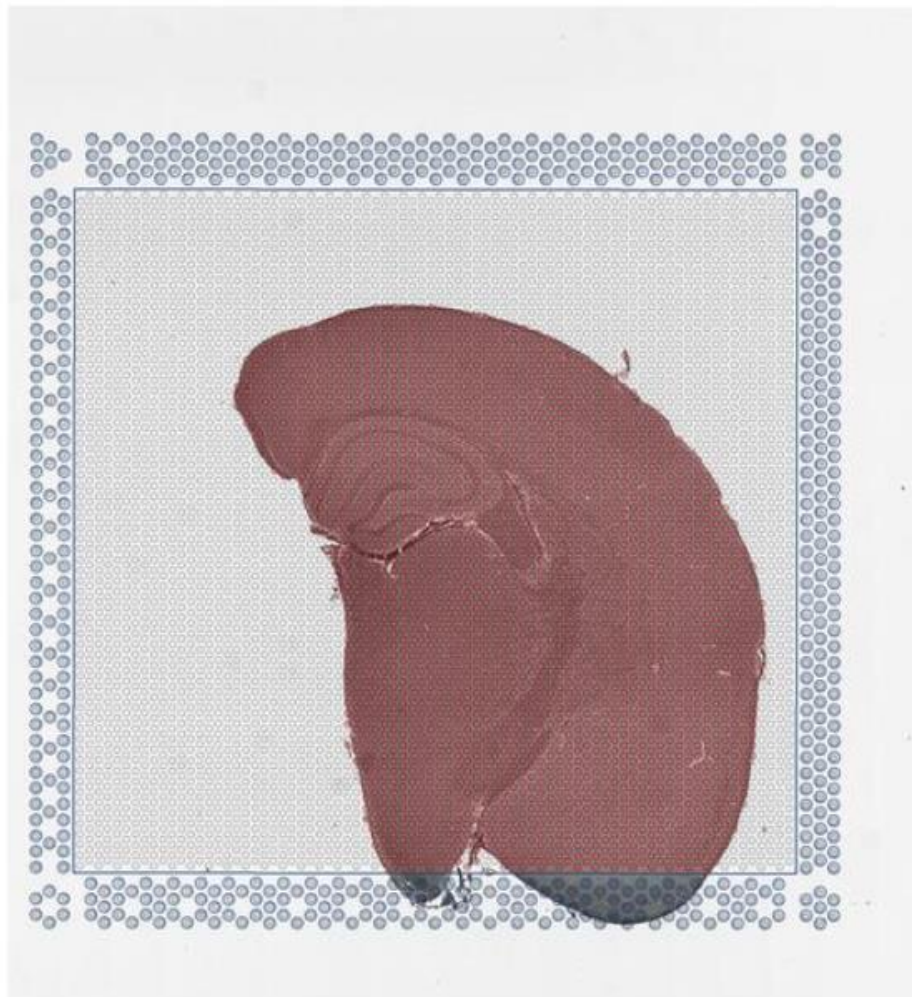
6,160

Median Genes per Spot

Sequencing ?

Number of Reads	65,261,954
Valid Barcodes	98.7%
Valid UMIs	100.0%
Sequencing Saturation	21.5%
Q30 Bases in Barcode	96.9%
Q30 Bases in Probe Read	96.3%
Q30 Bases in UMI	96.9%

Spots ?



Focus on MERFISH

Based on **single molecule FISH** (smFISH)

Binds specifically to target
transcript



Focus on MERFISH

Based on **single molecule FISH** (smFISH)

Binds specifically to target transcript



Directly count RNA targets with high-detection efficiency. But...

Focus on MERFISH

Based on **single molecule FISH** (smFISH)

Binds specifically to target transcript

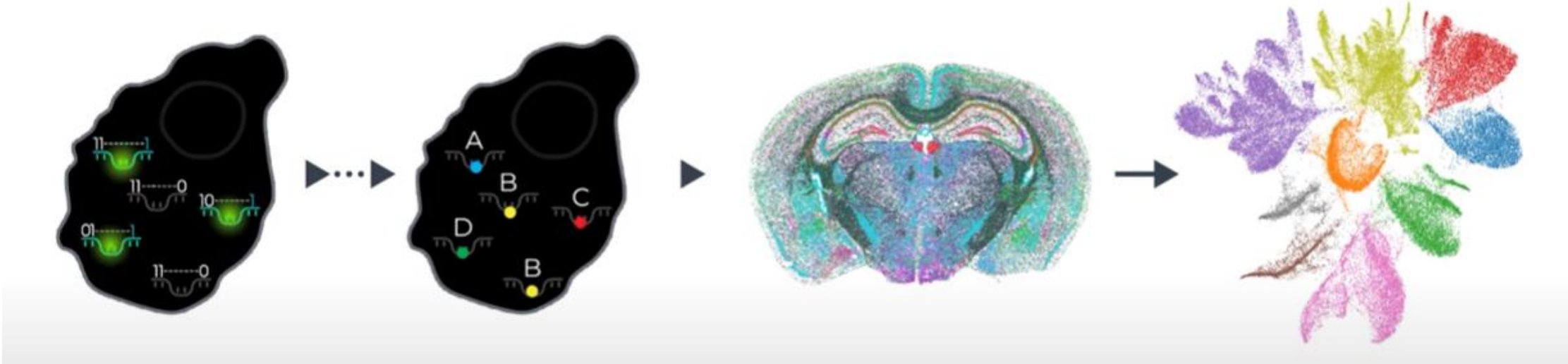


Directly count RNA targets with high-detection efficiency. But...

...smFISH is limited by multiplexing capability. You can measure few genes at a time.

Focus on MERFISH

Multiplexed error-robust fluorescence in situ hybridization



Based on **single molecule FISH** (smFISH)

Binds specifically to target transcript

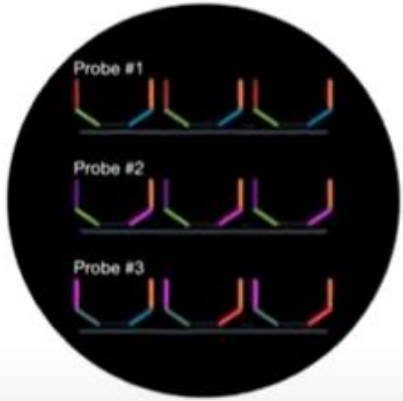


Directly count RNA targets with high-detection efficiency. But...

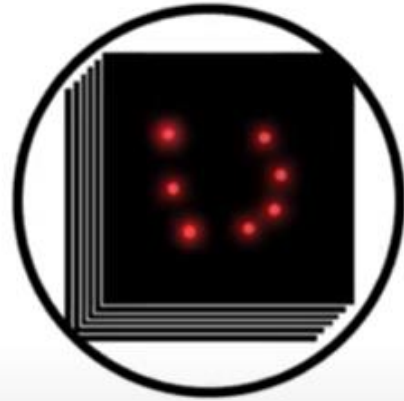
...smFISH is limited by multiplexing capability. You can measure few genes at a time.

MERFISH: how it works

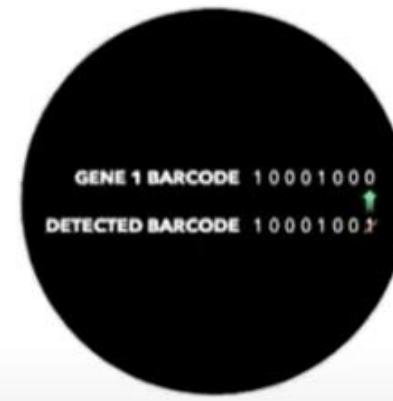
Combinatorial labeling



Sequential Imaging

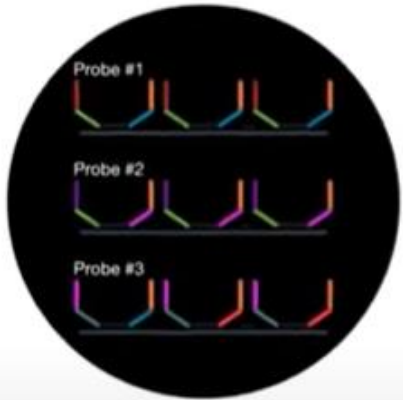


Error Robust Barcoding

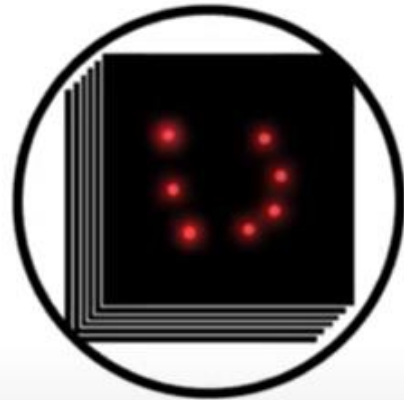


MERFISH: how it works

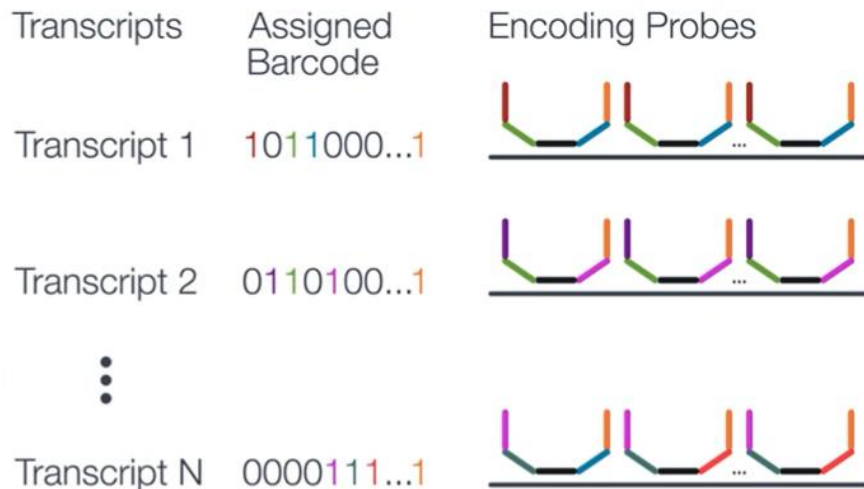
Combinatorial labeling



Sequential Imaging



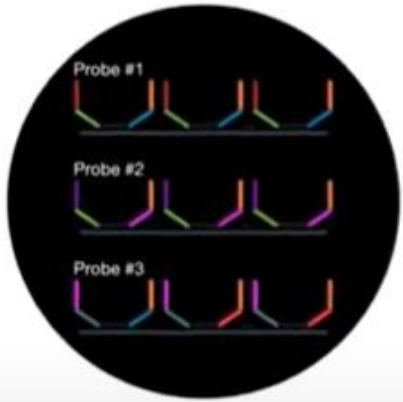
Error Robust Barcoding



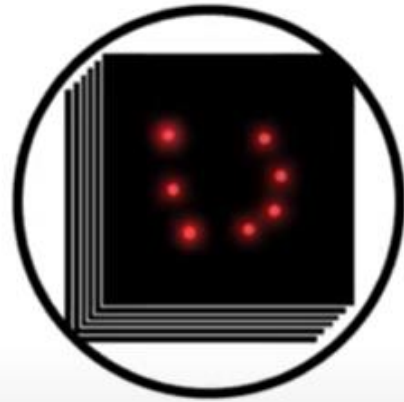
Sequential rounds of imaging are used to readout the barcode, that is a sequence of zeros and ones.

MERFISH: how it works

Combinatorial labeling



Sequential Imaging



Error Robust Barcoding



Transcripts	Assigned Barcode	Encoding Probes
Transcript 1	1011000...1	
Transcript 2	0110100...1	
⋮		
Transcript N	0000111...1	

Sequential rounds of imaging are used to readout the barcode, that is a sequence of zeros and ones.

Watch the video:
https://www.youtube.com/watch?v=00QekKSscjA&t=16s&ab_channel=Vizgen