TO KNOW. TO ACT.

Spatially resolved sub-cellular transcriptomics (Stereo-seq) and its applications

Vikrant Kumar, PhD

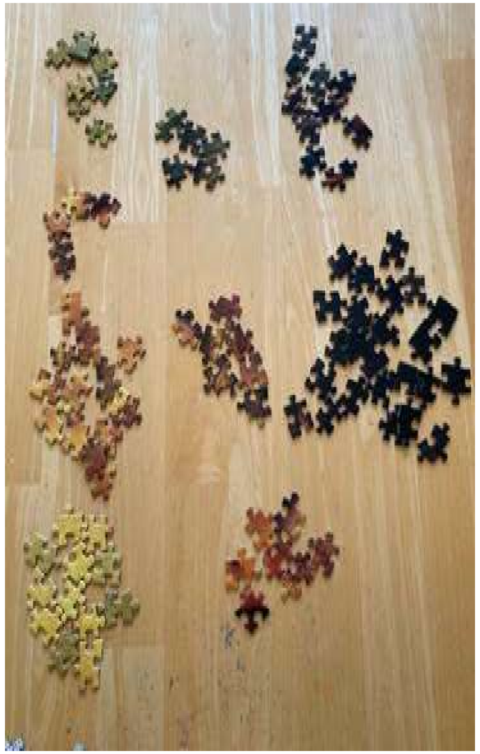
Outline

- **01** **Stereo-seq Intro & Workflow**
- 02 Stereo-seq Advantages
- 03 Stereo-seq Application

Genomics research is shifting from bulk to single cell to spatial

Bulk Transcriptomics

Single Cell Transcriptomics

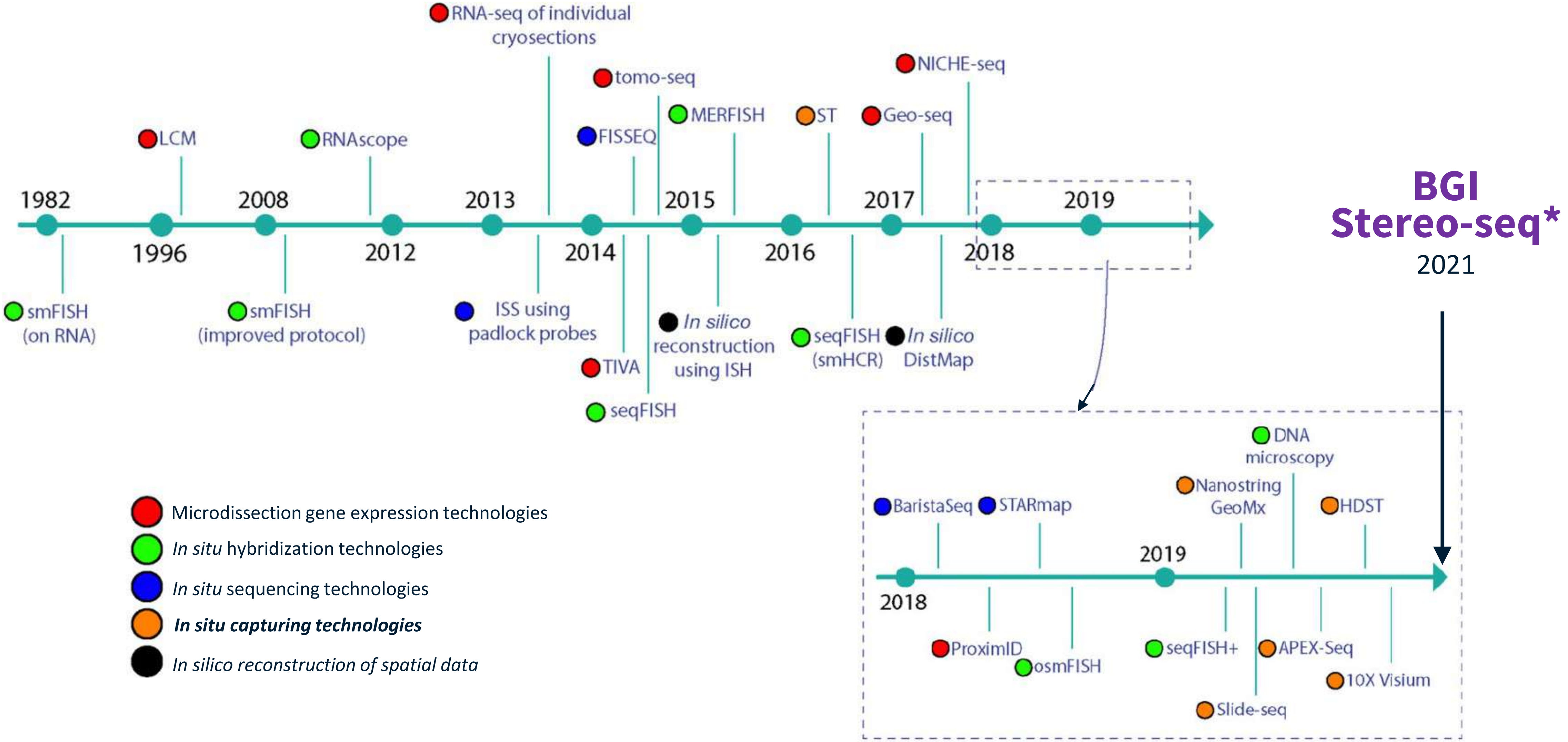


Low Resolution Spatial Transcriptomics

High Resolution Spatial Transcriptomics

mirxes
TO KNOW. TO ACT.

Development of Spatial Technologies



1982
🟢 smFISH (on RNA)

1996
🔴 LCM

2008
🟢 smFISH (improved protocol)

2012
🟢 RNAscope

2013
🔴 RNA-seq of individual cryosections
🔵 ISS using padlock probes

2014
🔵 FISSEQ
🔴 tomo-seq
🔴 TIVA
🟢 seqFISH

2015
🟢 MERFISH
⚫ *In silico* reconstruction using ISH

2016
🟠 ST
🟢 seqFISH (smHCR)

2017
🔴 Geo-seq
🔴 NICHE-seq
⚫ *In silico* DistMap

2018
🔵 BaristaSeq
🔵 STARmap
🔴 ProximID
🟢 osmFISH

2019
🟠 Nanostring GeoMx
🟢 DNA microscopy
🟠 HDST
🟢 seqFISH+
🟠 Slide-seq
🟠 APEX-Seq
🟠 10X Visium

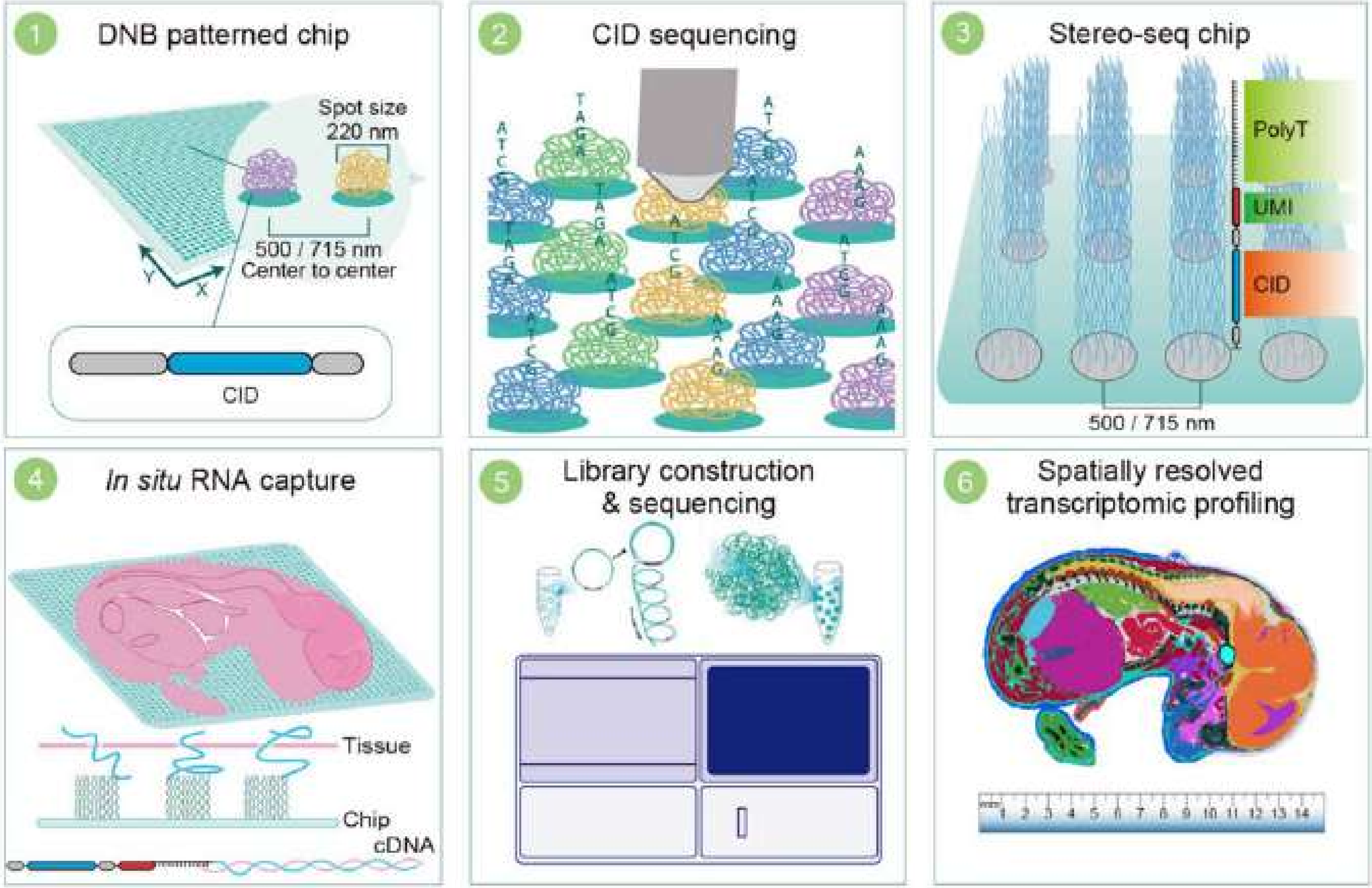
BGI Stereo-seq*
2021

🔴 Microdissection gene expression technologies
🟢 *In situ* hybridization technologies
🔵 *In situ* sequencing technologies
🟠 ***In situ capturing technologies***
⚫ *In silico reconstruction of spatial data*

***SpaTial Enhanced REsolution Omics-sequencing**

Asp et al., 2020 (BioEssays)

Stereo-seq: SpaTial Enhanced REsolution Omics-Sequencing



STOmics Fluorescent Set-F1 workflow

TOTAL TIME: ~4 HRS

EXPERIMENTAL PREPARATION	CHIP PREPARATION & MOUNTING	TISSUE FIXATION	TISSUE PERMEABILIZATION TESTING
30 min	15 min	30 min	*variable

REVERSE TRANSCRIPTION	TISSUE REMOVAL	IMAGING & PERMEABILIZATION TIME DETERMINATION
1 hr	1 hr	*variable

STOmics Gene Expression Set-S1 workflow

TOTAL TIME: ~ 1.5 DAYS

DAY 1

EXPERIMENTAL PREPARATION	TISSUE SECTIONING & MOUNTING	TISSUE FIXATION	TISSUE STAINING & IMAGING
30 min	15 min	30 min	20 min

TISSUE PERMEABILIZATION	REVERSE TRANSCRIPTION	TISSUE REMOVAL	cDNA RELEASE & COLLECTION ⛔
*variable	3 hr	10 min	3 hr–18 hr

DAY 2

cDNA PURIFICATION & AMPLIFICATION ⛔	LIBRARY PREPARATION
140 min	90 min

Proceed to Sequencing (DNBSEQ-G400RS/ T7RS)

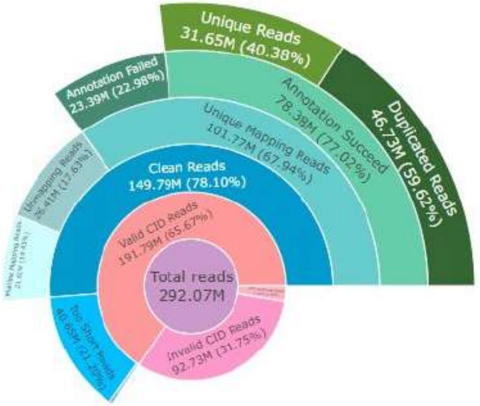
Data Analysis

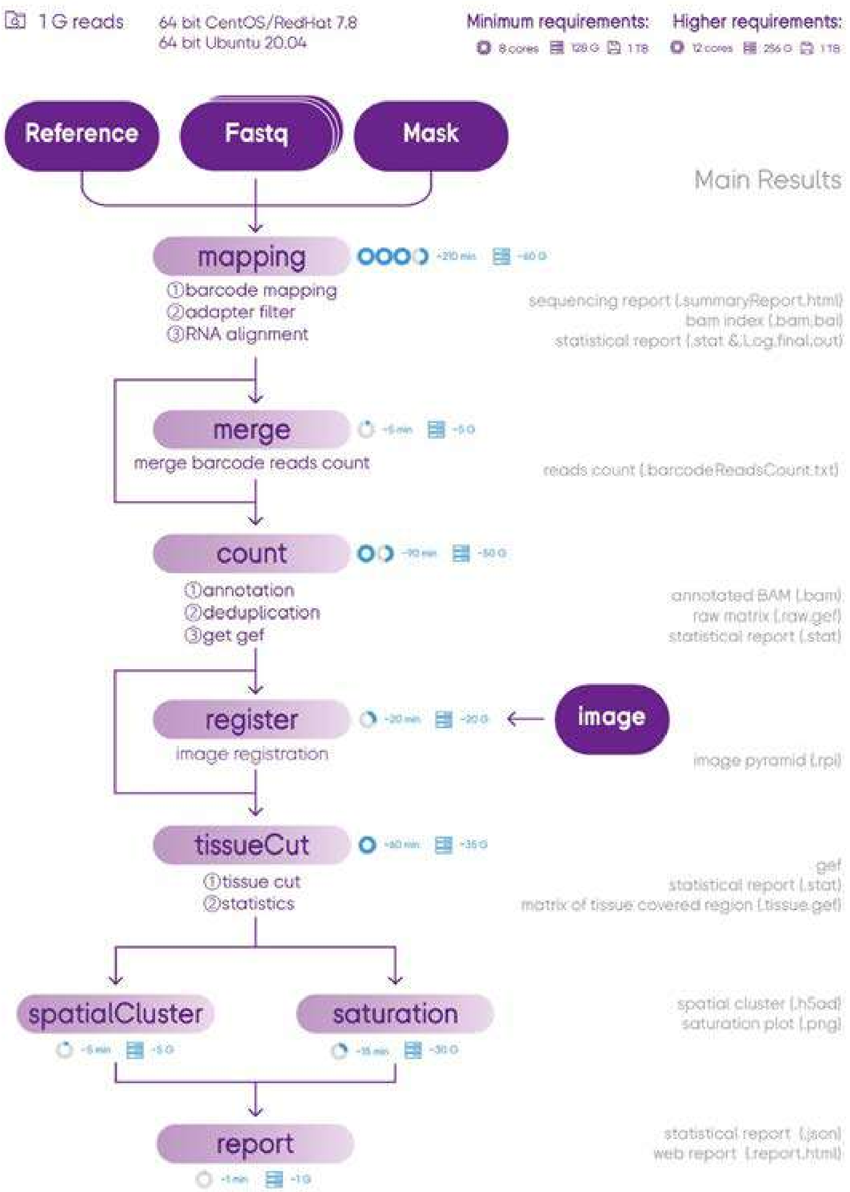
STOmics Analysis Workflow (SAW)

Four Inputs:

1. **Fastq** : sequencing data
2. **Mask** : CID information
3. **Reference** : Organism genome
4. **Image** : nuclear stained image

Final Output: GEF / GEM



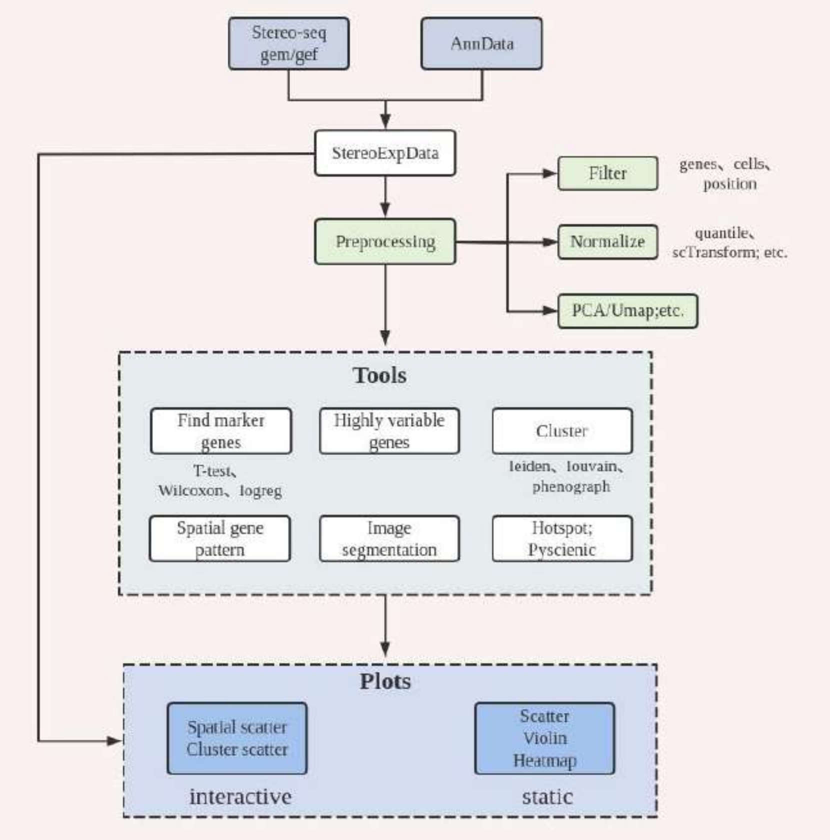


Spatial Transcriptomics Analysis in Python (Stereopy)

In-house visualization pipelines

STOmics Analysis Platform (SAP)

Cloud-based visualization website


Stereo-seq gem/gef

AnnData

StereoExpData

Preprocessing

Filter

genes、cells、position

Normalize

quantile、scTransform; etc.

PCA/Umap;etc.

Tools

Find marker genes

T-test、Wilcoxon、logreg

Highly variable genes

Cluster

leiden、louvain、phenograph

Spatial gene pattern

Image segmentation

Hotspot; Pyscienic

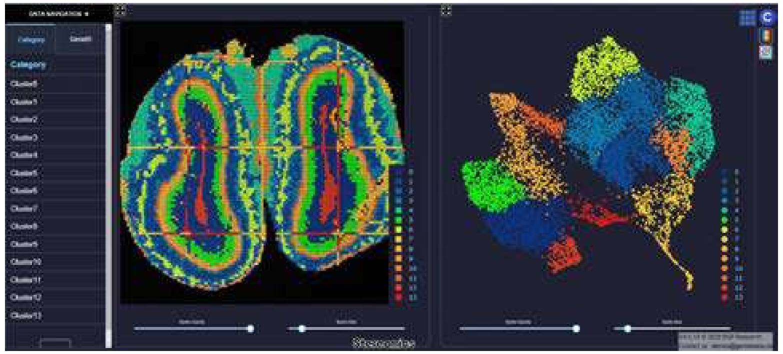
Plots

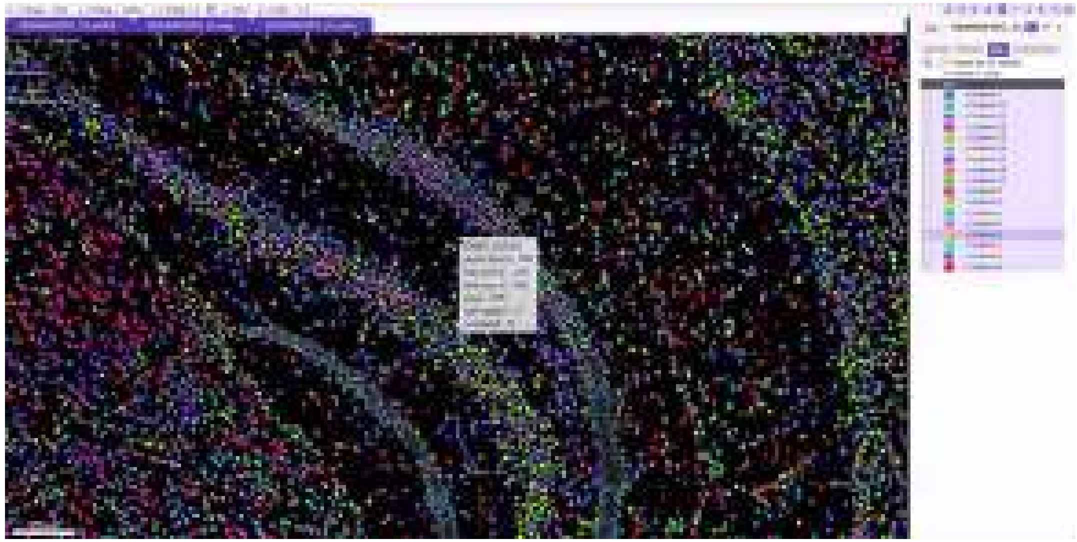
Spatial scatter Cluster scatter

interactive

Scatter Violin Heatmap

static





Outline

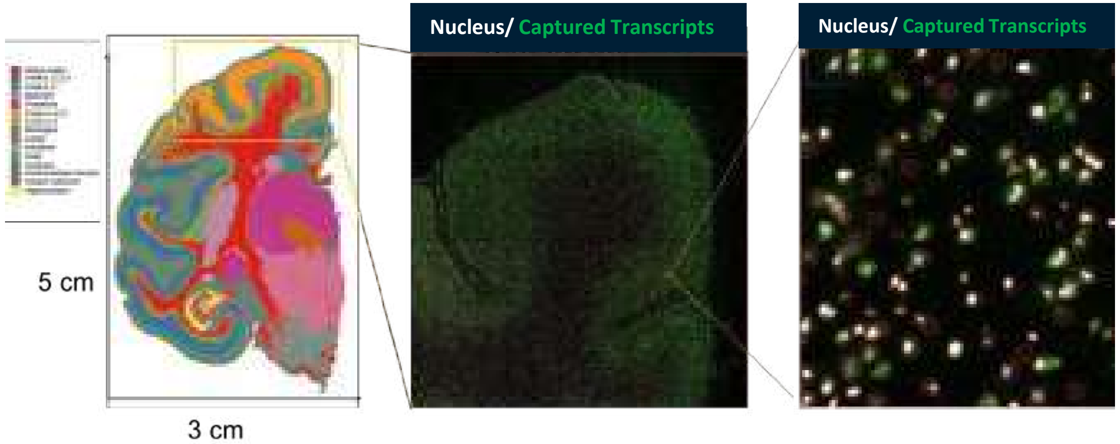
01 Stereo-seq Intro & Workflow

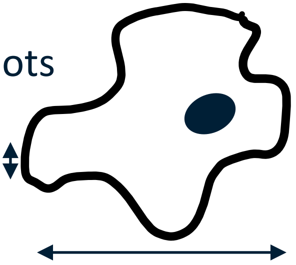
02 Stereo-seq Advantages

03 Stereo-seq Application

Nanoscale Resolution

- ✓ **Subcellular Resolution**
- ✓ **Single cell segmentation**

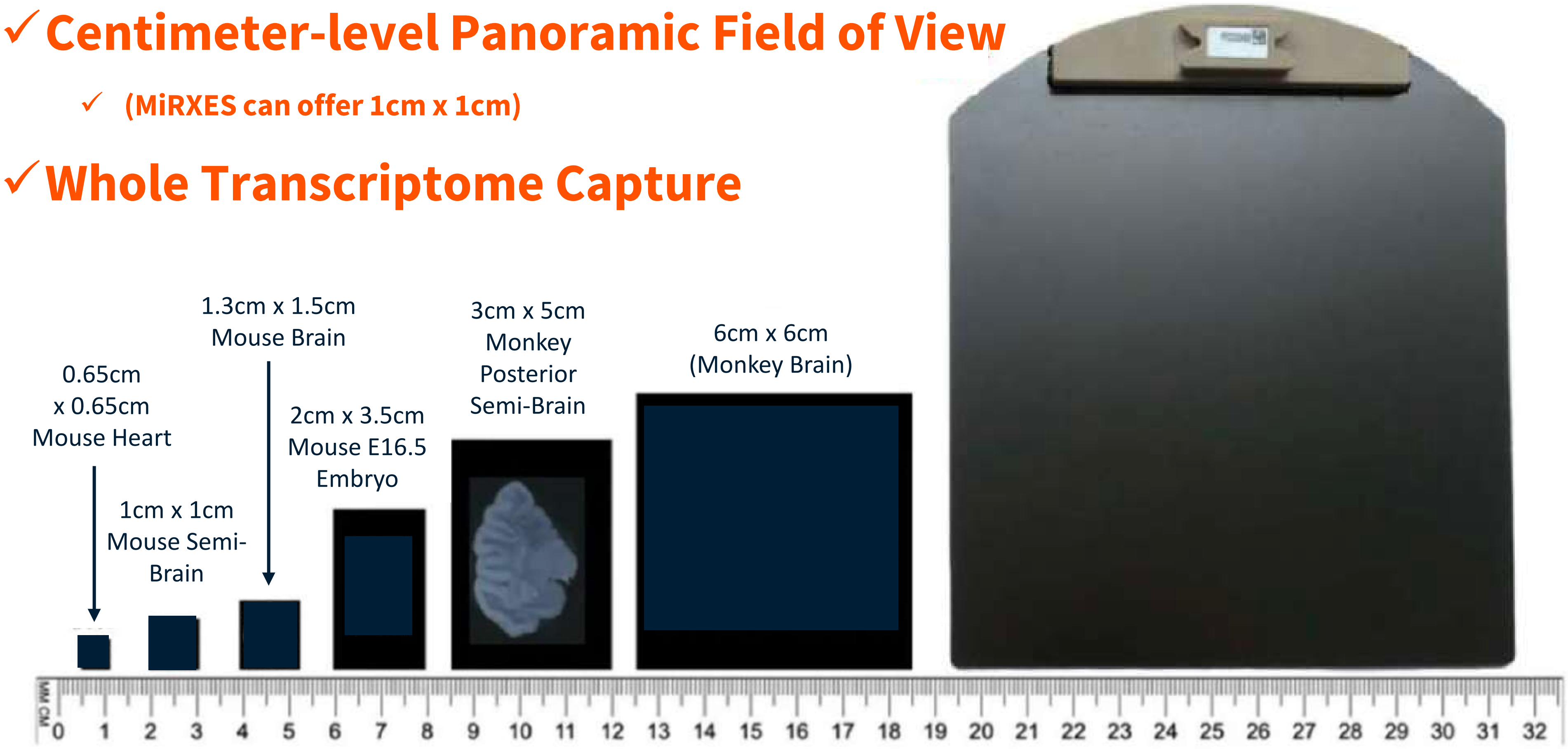




Chen et al., 2022 (Cell)

Panoramic Field of View and Full Capture

- ✓ **Centimeter-level Panoramic Field of View**
  - ✓ **(MiRXES can offer 1cm x 1cm)**
- ✓ **Whole Transcriptome Capture**



0.65cm
x 0.65cm
Mouse Heart

1cm x 1cm
Mouse Semi-Brain

1.3cm x 1.5cm
Mouse Brain

2cm x 3.5cm
Mouse E16.5
Embryo

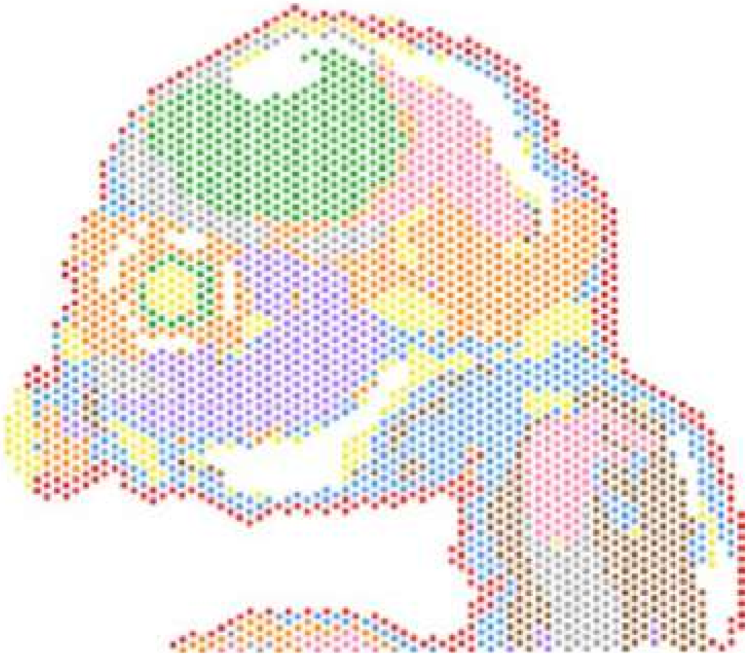
3cm x 5cm
Monkey
Posterior
Semi-Brain

6cm x 6cm
(Monkey Brain)

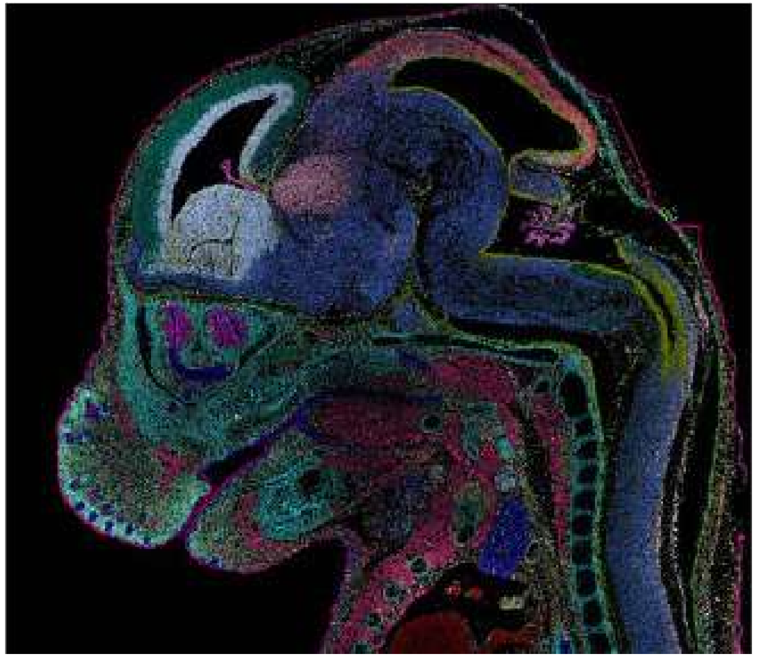
13.2cm x 13.2cm
(Human Brain)

Stereo-seq sample chip sizes

Comparison of V to StereoSeq



V



StereoSeq

Outline

01 Stereo-seq Intro & Workflow

02 Stereo-seq Advantages

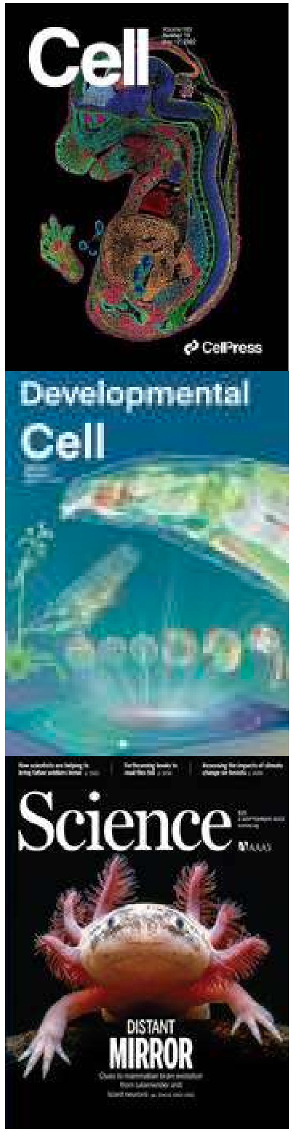
03 Stereo-seq Application

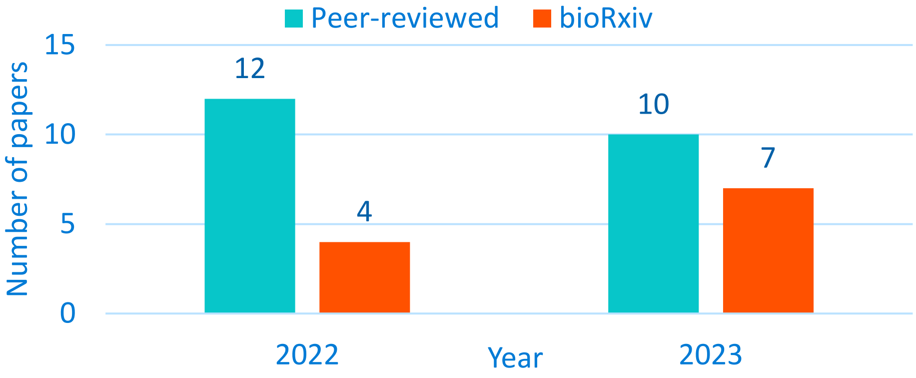
Stereo-seq publications

Published in highly impact factor journals:

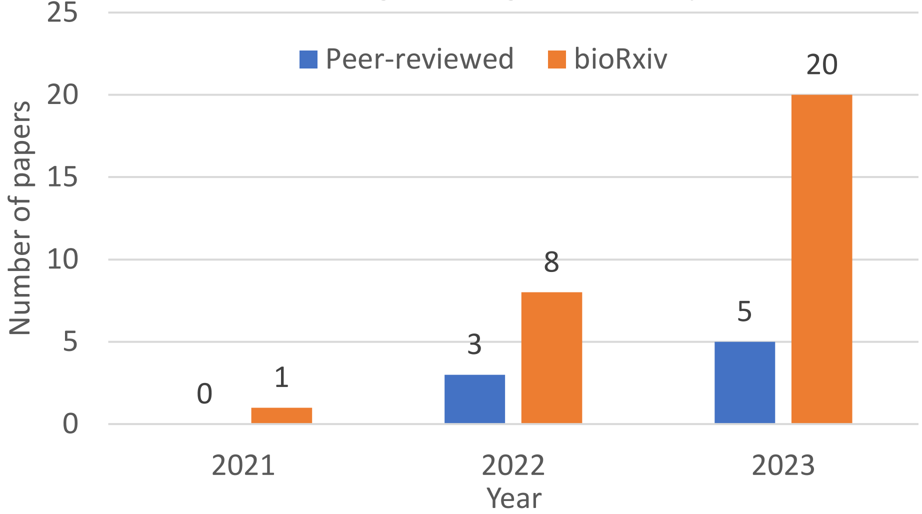
Science, Nature/ Nature Communications/ Cell Research/npj Precision Oncology,

Cell/ Developmental Cell/ Cell Reports Medicine, Plant Biotechnology Journal, Advance Science





Number of published papers on computational methods using existing Stereo-seq dataset



The single-cell stereo-seq reveals region-specific cell subtypes and transcriptome profiling in *Arabidopsis* leaves

High-resolution 3D spatiotemporal transcriptomic maps of developing *Drosophila* embryos and larvae

Spatiotemporal mapping of gene expression landscapes and developmental trajectories during zebrafish embryogenesis

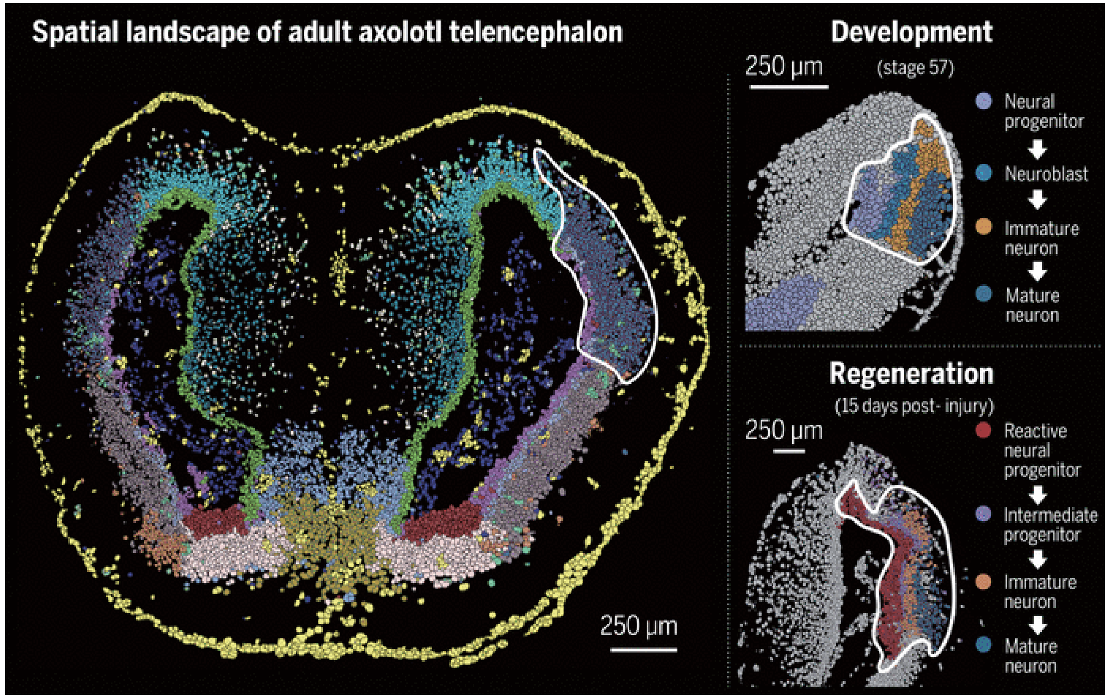
Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays

A cellular hierarchy in melanoma uncouples growth and metastasis

Single-Nucleus RNA Sequencing and Spatial Transcriptomics Reveal the Immunological Microenvironment of Cervical Squamous Cell Carcinoma

Spatially resolved gene regulatory and disease-related vulnerability map of the adult Macaque cortex

Spatial resolution of axolotl brain



Wei et al., Science (2022) Volume: 377, Issue: 6610

- ➢ Axolotl brain injury model to capture the cell lineage transition during regeneration and to compare with developmental process
- ➢ **High-definition** and **large-field** Stereo-seq technology to generate spatial transcriptomic data at single-cell resolution from sections that cover both hemispheres of the axolotl telencephalon

Neuroscience | News | Translational Medicine | Regenerative Medicine/Tissue Engineering | Spatial Biology | Traumatic Brain Injury

Axolotl Spatial Questions? Stereo-Seq Reveals Brain Regeneration Insights on Salamanders

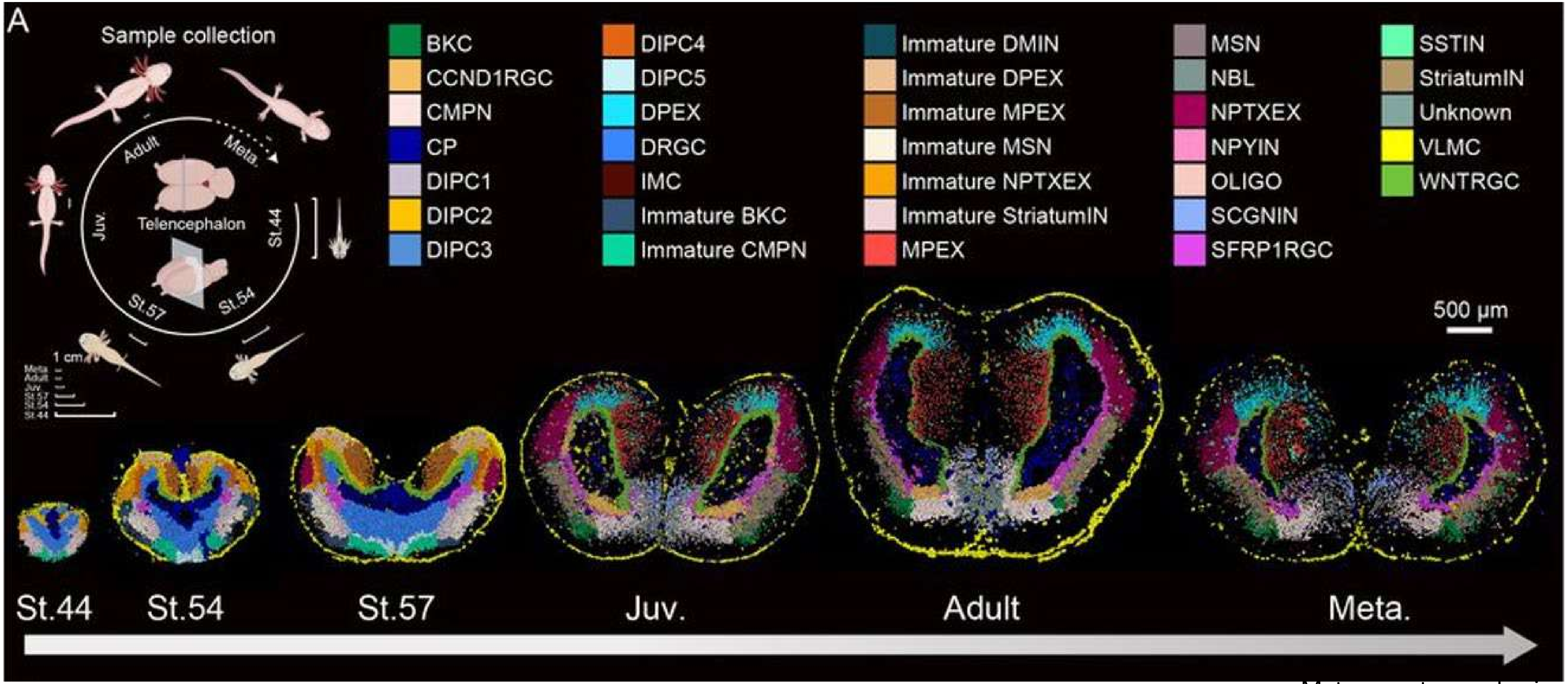
September 7, 2022

SCIENCE
2 Sep 2022
Vol 377, Issue 6610

https://pubmed.ncbi.nlm.nih.gov/36048929/

Wei X, et. al. **Science**. 2022 Sep 2;377(6610)

Development related radial gilial cells (DRGCs) identified- the subpopulation was present in dominance throughout developmental stages but decreased in number and disappeared after juvenile stage.



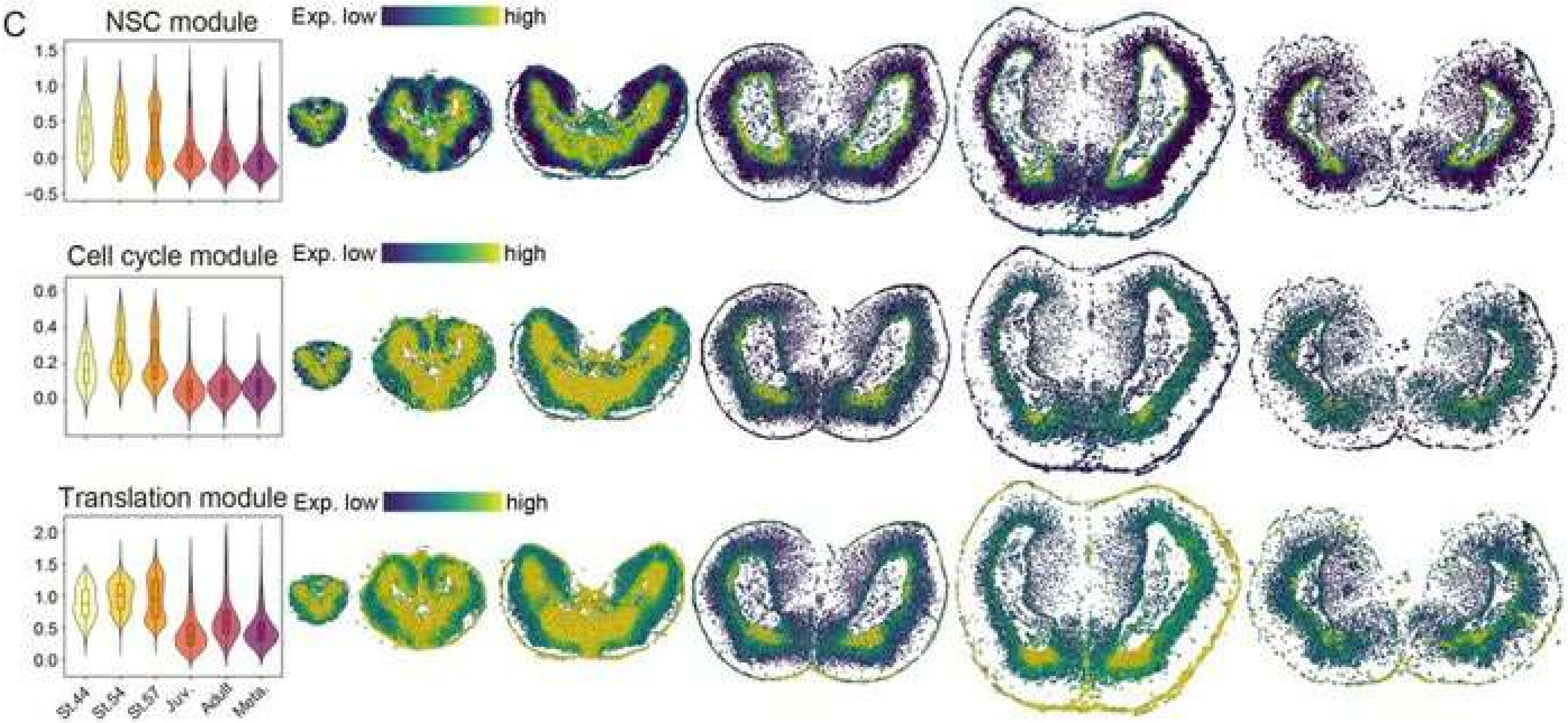
St.= stage

Meta. = metamorphosis

Wei X, et. al. **Science**. 2022 Sep 2;377(6610)

Stemness and proliferation activity of RGCs

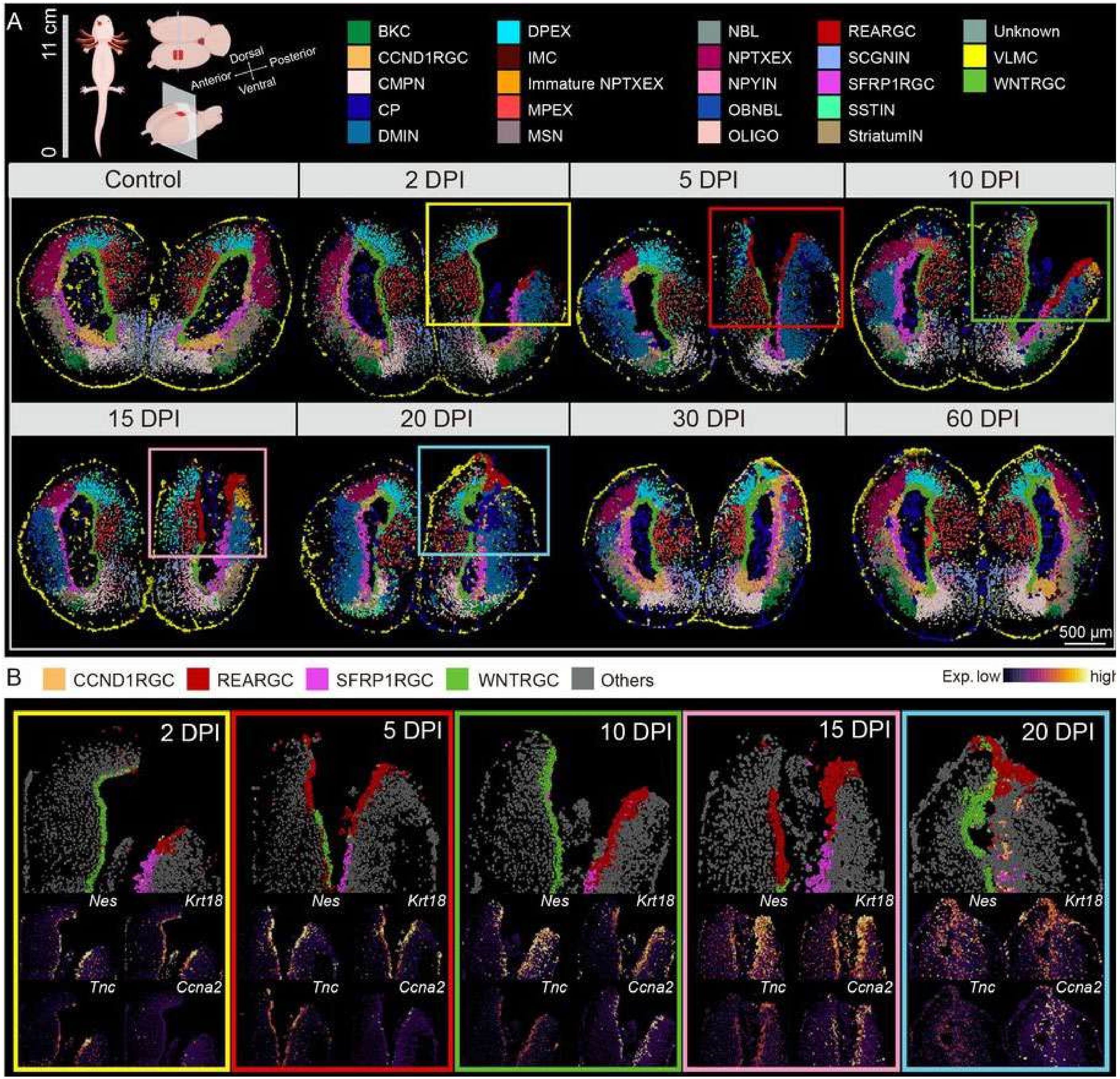
- Stereo-seq data clearly shows that cells positive for neural stemness, cell cycle and translation module significantly dropped in number and became restricted to the ventricular zone (VZ) region.



Wei et al., 2022
(Science)

Identification of novel RGC type (REARGC) after axolotl injury

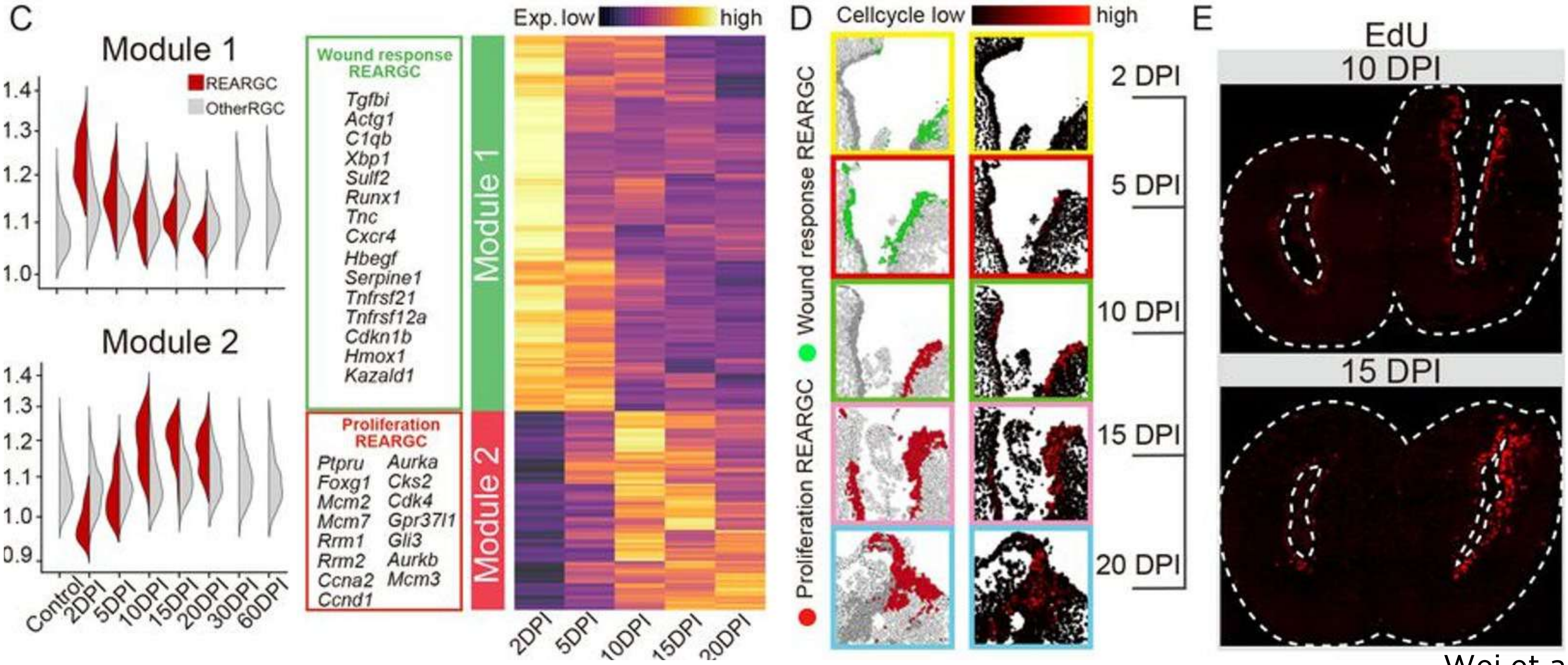
- ➢ A novel type of RGC (named REARGC for reactive RGC) enriched at the edge of region but not in intact right telencephalon sections was identified.
- ➢ This injury-specific RGCs showed high expression level of genes e.g., Nes, Krt18, Tnc and Ccna2, indicating stem cell state and high proliferation activity.
- ➢ REARGCs were the dominant cell type in wound area from 2-20 DPI.



Wei et al., 2022 (Science)

REARGCs has dual roles of inflammatory response and cell propagation for tissue regeneration

- Early and transient expression of wound response related genes (Module1) was observed at 2 DPI at wound edge
- Second wave showed continuous induction of proliferation induced genes (Module 2) in REARGCs from 10 to 20 DPI when REARGCs expanded in number.

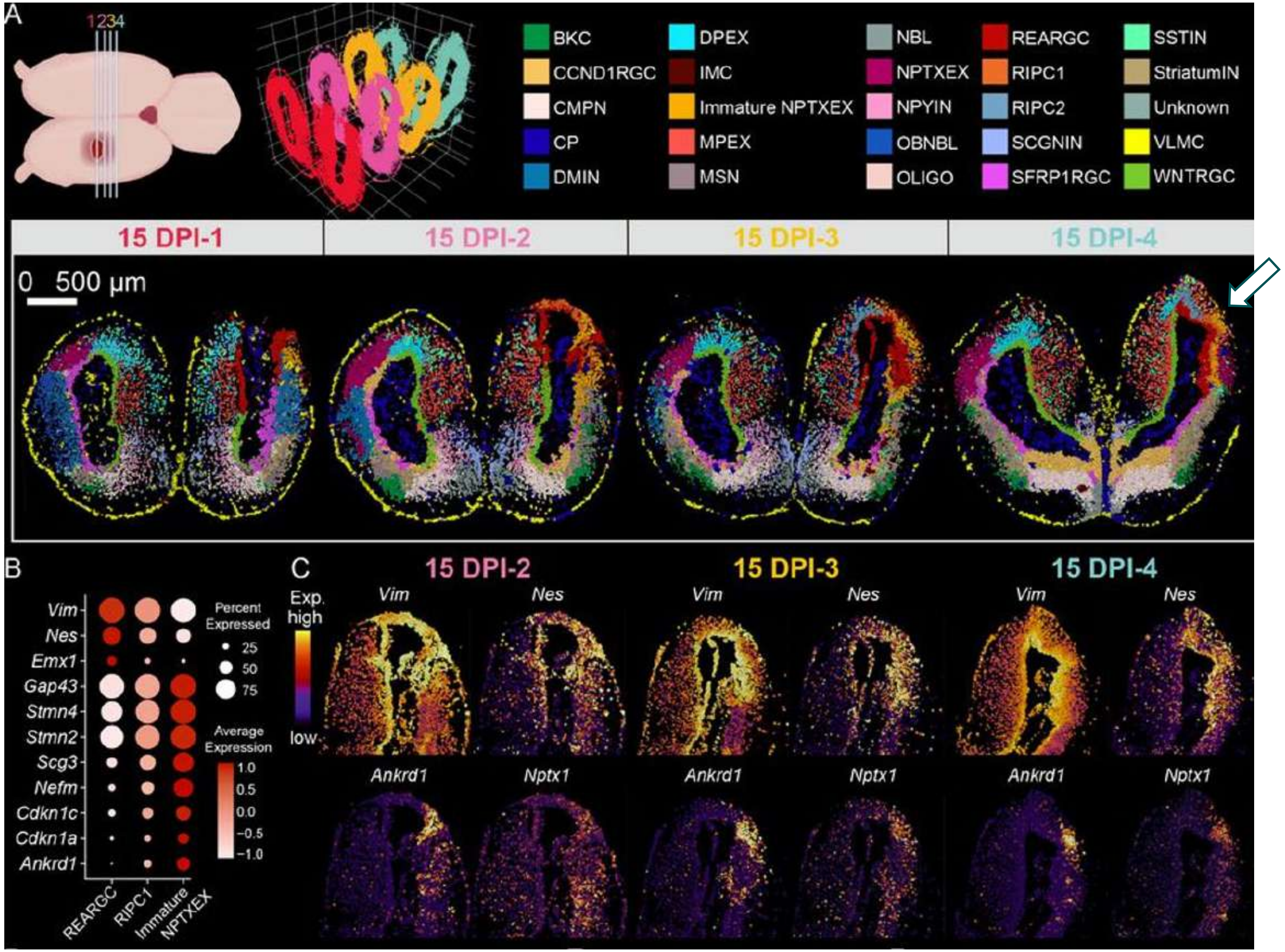


Proliferation of REARGCs confirmed by EdU staining from injury site along VZ region

Wei et al., 2022 (Science)

Regeneration intermediate progenitor cells 1 (RIPC1) identified as intermediate state between REARGCs and immature NPTXEX

- Early and transient expression of wound response related genes (Module1) was observed at 2 DPI at wound edge
- Second wave showed continuous induction of proliferation induced genes (Module 2) in REARGCs from 10 to 20 DPI when REARGCs and immature NPTXEXs (Nptx+ lateral pallium excitatory neuron) expanded in number.

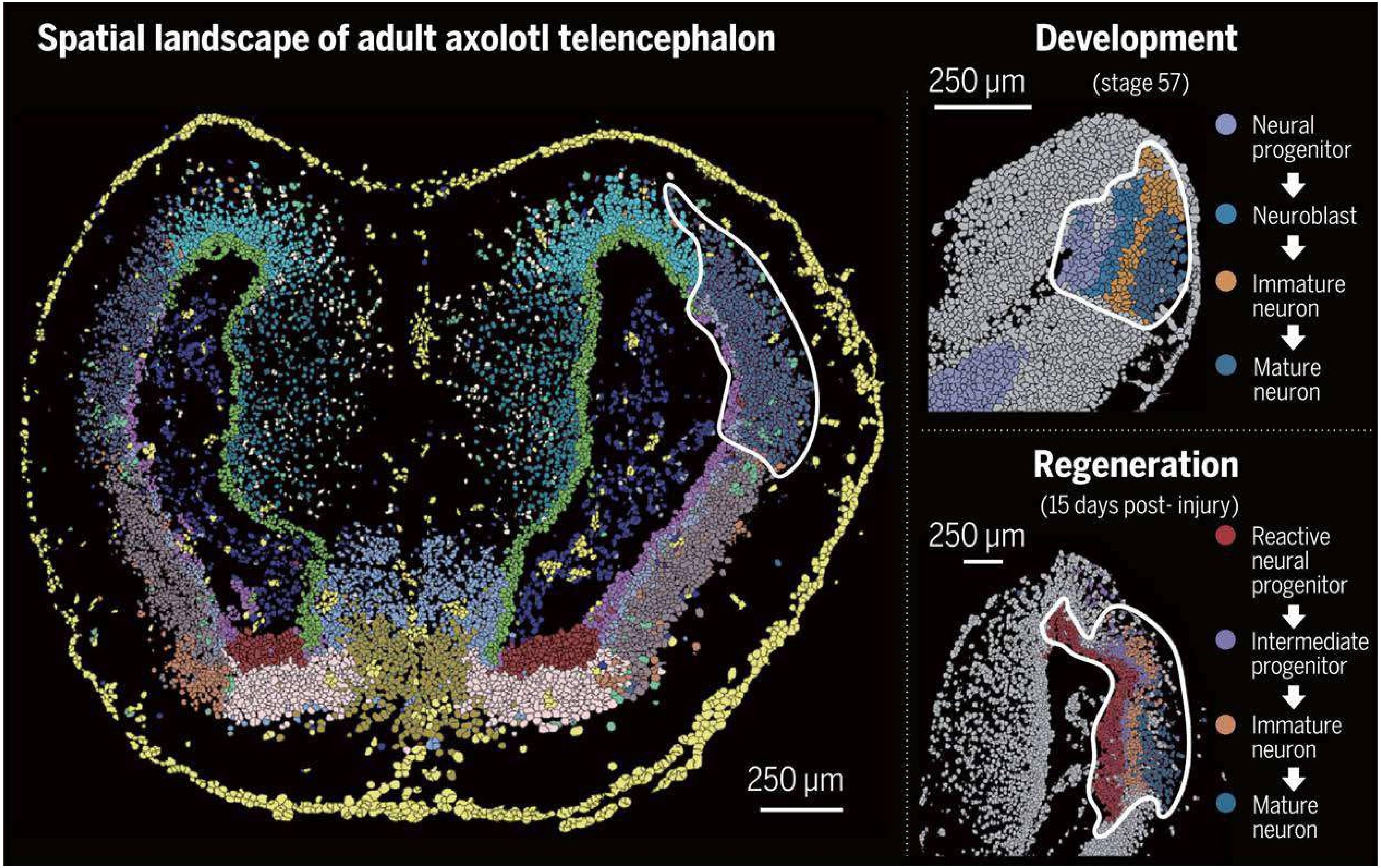


Four adjacent cell layers of REARGCs, RIPC1s, immature NPTXEXs, and mature NPTXEXs

RIPC1 expressed both REARGC and NPTXEX markers

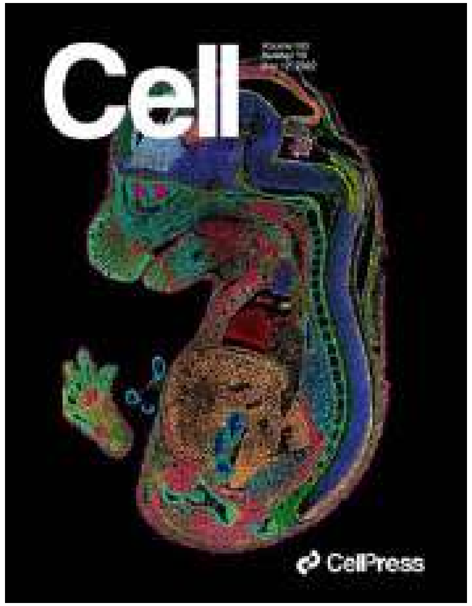
Wei et al., 2022 (Science)

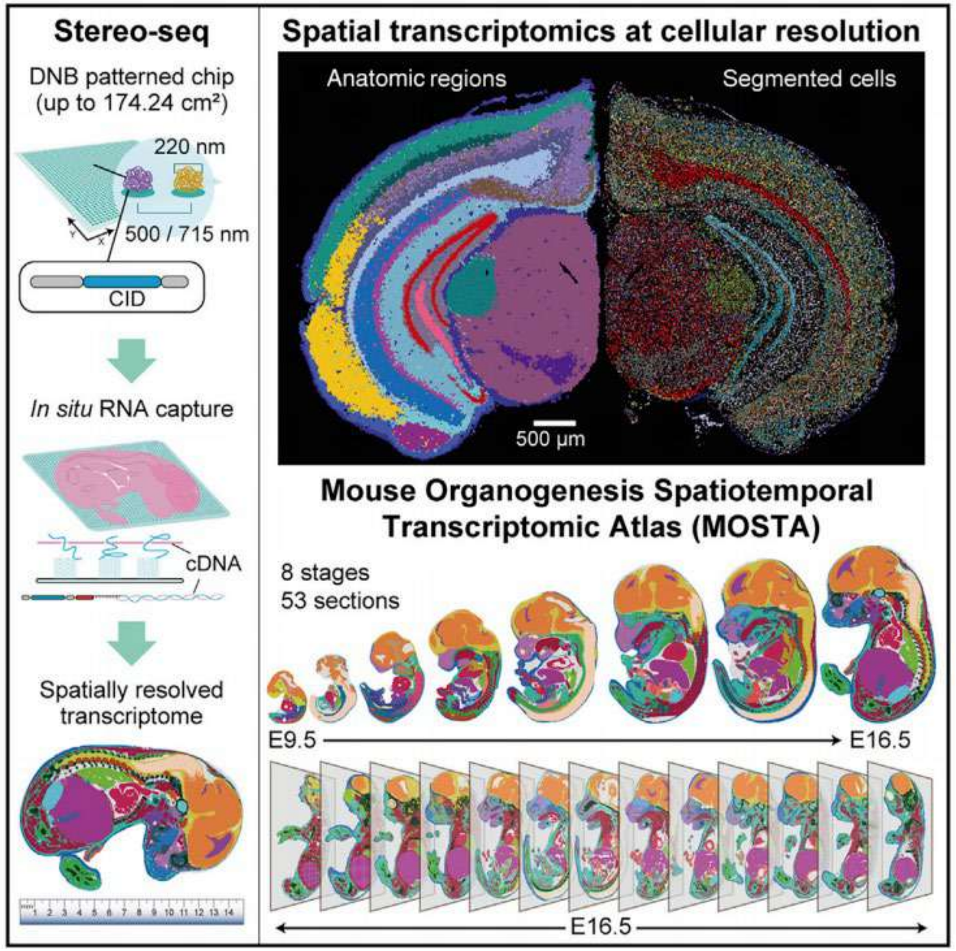
Axolotl development vs regeneration



Wei X, et. al.
Science. 2022

On the cover: Stereo-seq dissects cell-type composition in an E16.5 mouse embryo sagittal section

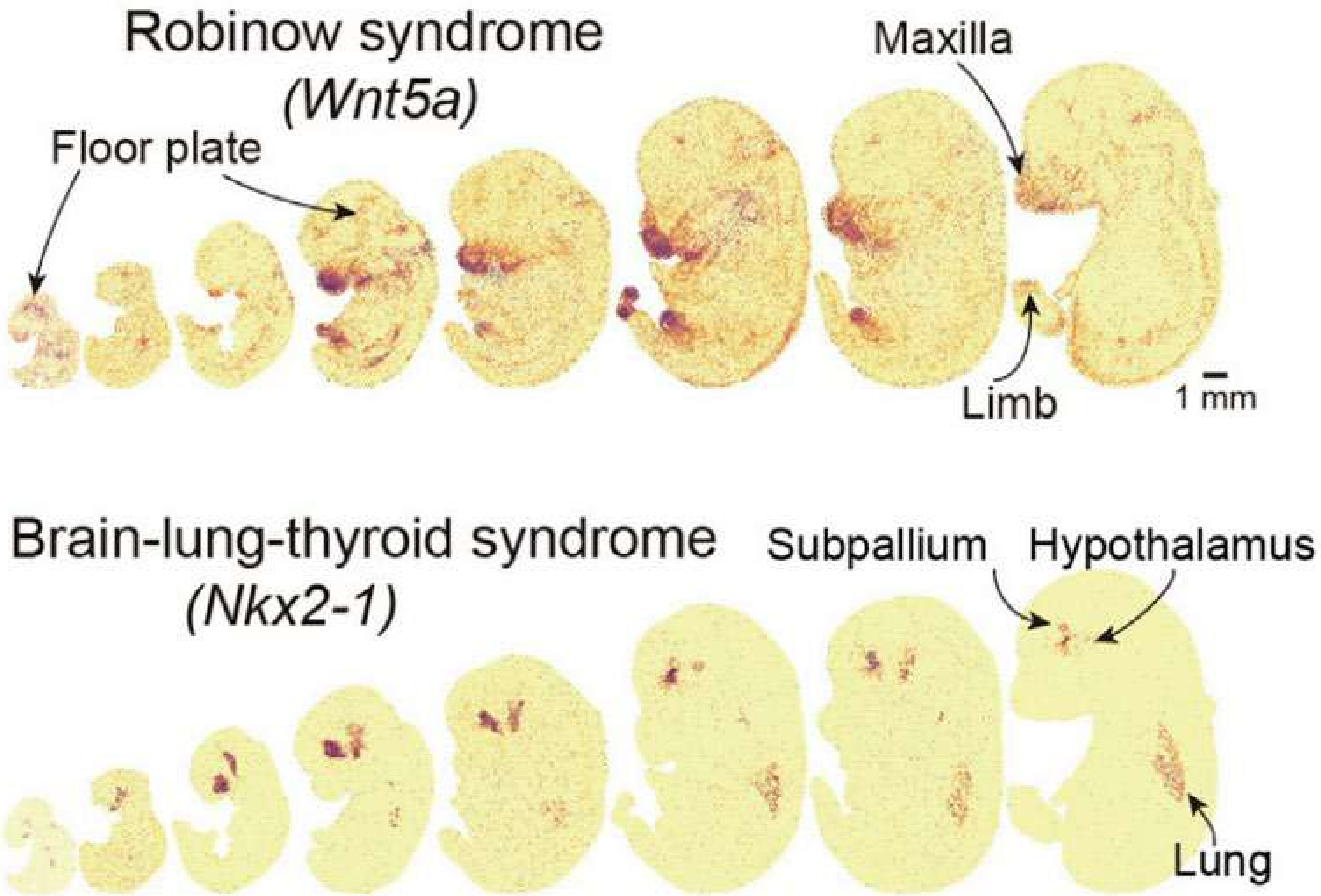




Chen et al., 2022 (Cell)

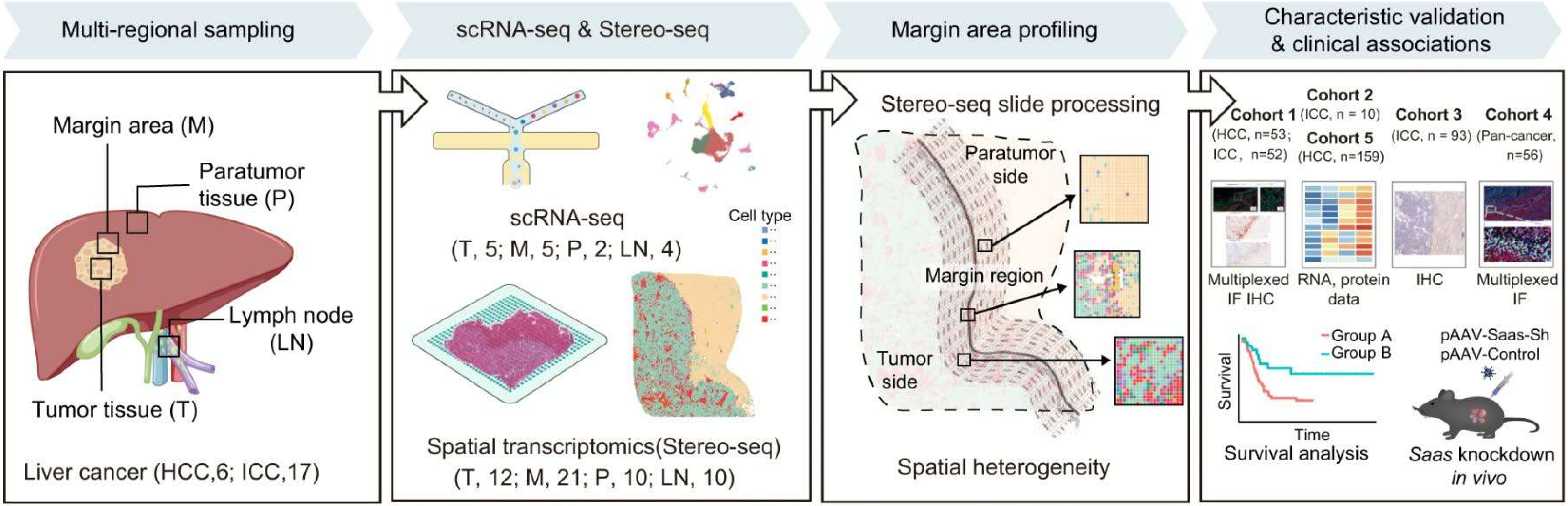
- **Spatial** and **Temporal** transcriptomic dynamics during mouse organogenesis at **cellular resolution**
- Cell-type heterogeneity: identification of previously undefined cell subtypes or cell states
- Decipher developmental disease susceptibility

Mouse Embryo Development



Chen et al., 2022 (Cell)

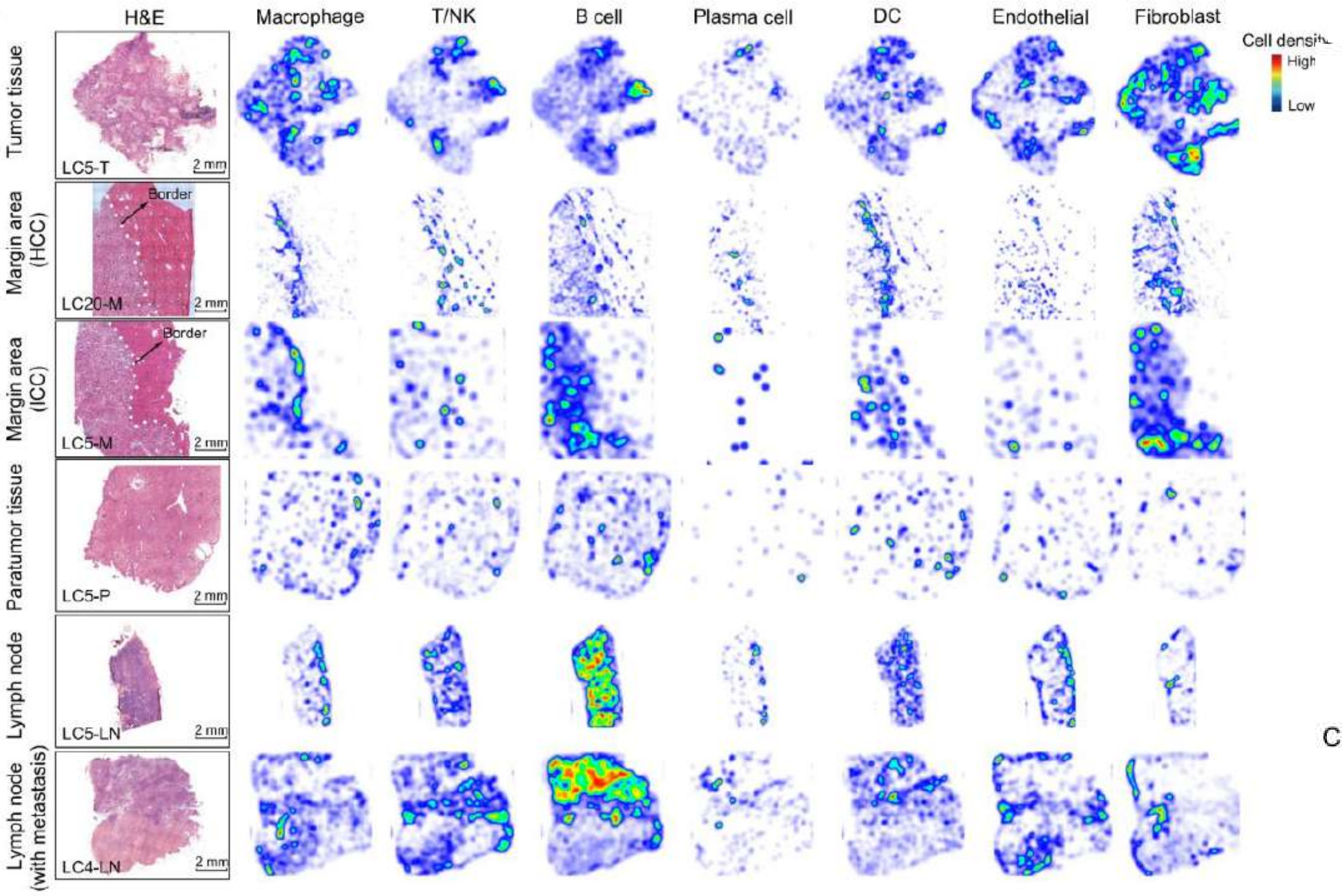
Spatially-resolved transcriptomic profiles in multiple regional sites in human liver cancer

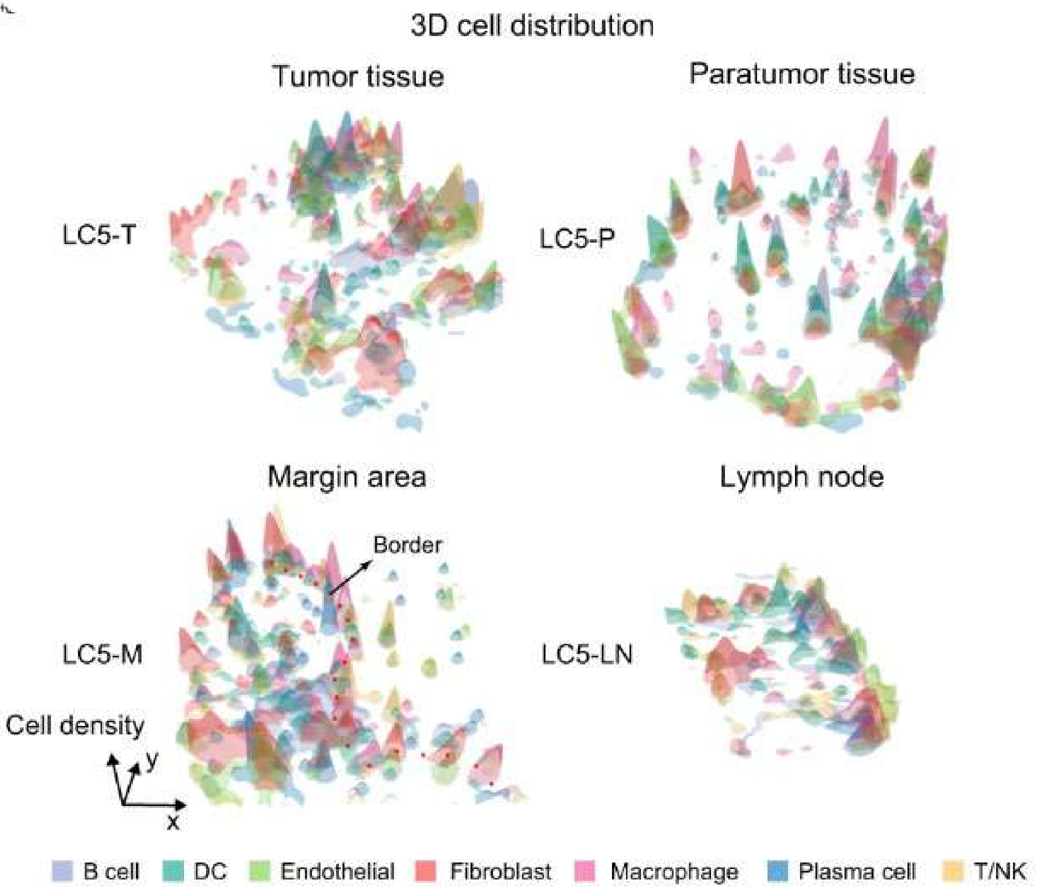


Wu, L., *Cell Res* **33**, 585–603 (2023)

Spatial heterogeneities of Cell Types in Liver Cancer Tissues

- Spatially, immune cells such as macrophages were observed to accumulate primarily around the tumor border in margin areas, indicating a distinct immune microenvironment around the border

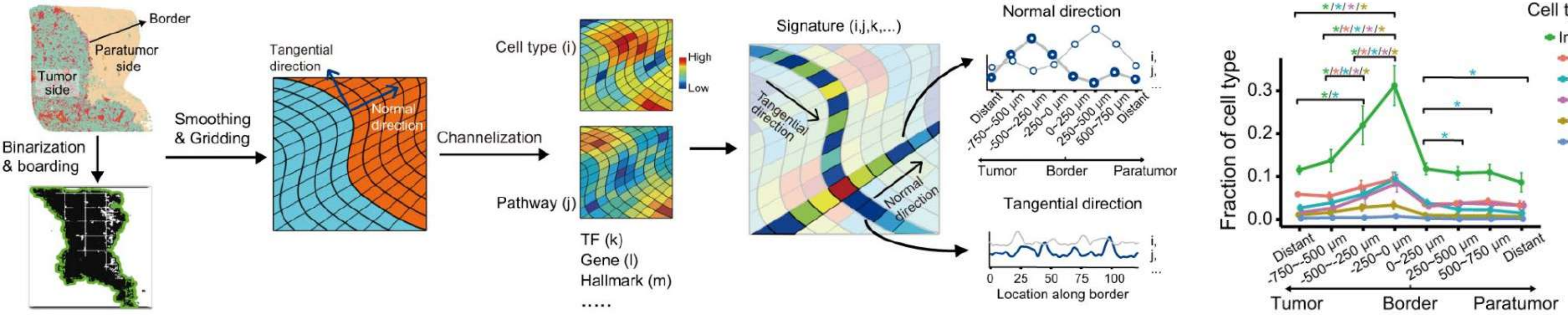




Wu et al., 2023 (Cell Research)

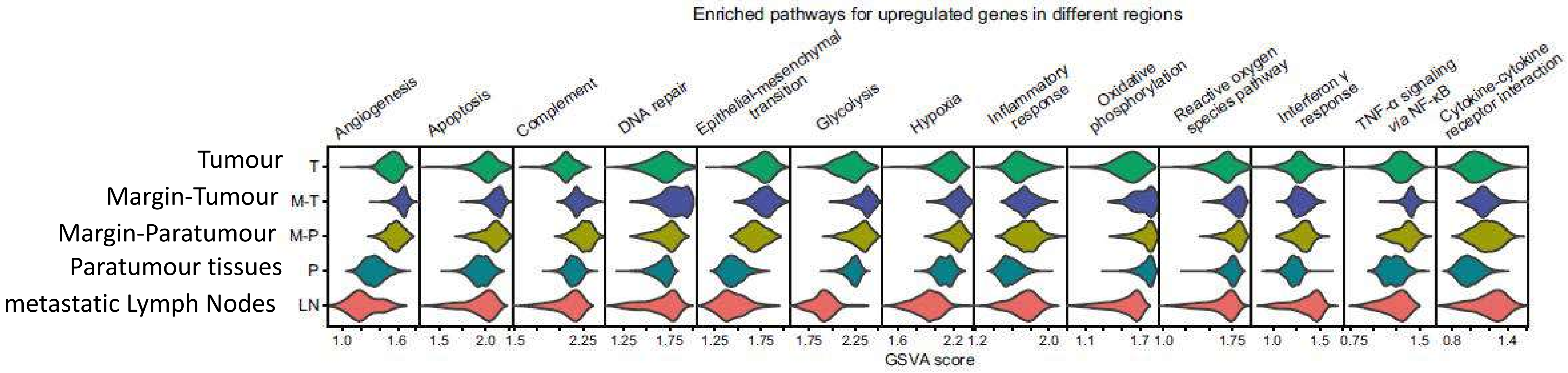
Tumour Border Segmentation Strategy Based on Stereo-seq

- Based on Stereo-seq, the margin areas along the border was segmented into several layers (layer representing a 250 µm-wide zone from the border).
- Each layer was divided into 100 equal parts along the normal direction of the border
- Revealed significant enrichment of immune cells and fibroblasts on the tumor side



Wu et al., 2023 (Cell Research)

Differentially Enriched Pathways in different parts of Liver Cancer Tissue



- ➢ Pathways related to epithelial-mesenchymal transition (EMT), the TMEs (e.g., hypoxia, angiogenesis, TGF-β signaling), inflammation and apoptosis, glycolysis and oxidative phosphorylation, and immune response were highly enriched in the tumor-adjacent part of margin areas compared to tumor tissues.
- ➢ Implies tumor margin areas might be a complex region characterized by a hypoxic microenvironment, robust inflammatory responses, and high immune escape.

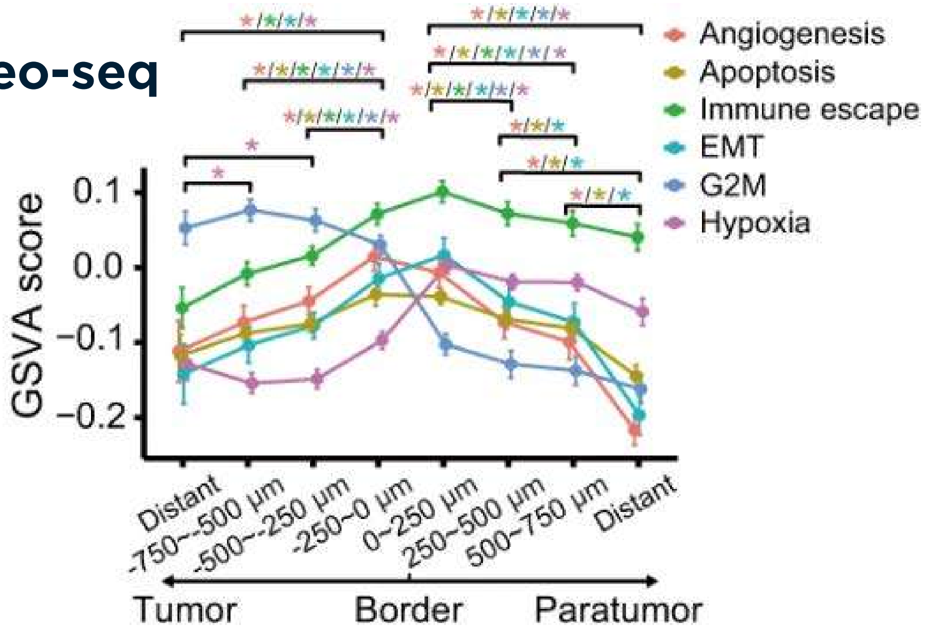
Wu et al., 2023 (Cell Research)

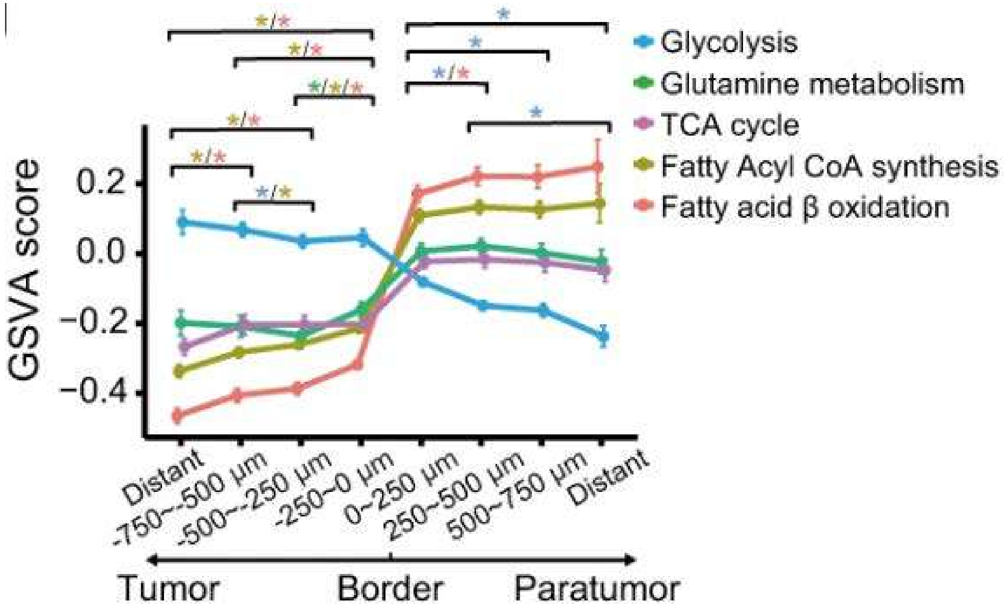
Invasive zone in human liver cancer identified by Stereo-seq

- Tumour cells in the first layer from the border (-250~0µm) exhibited the following compared to other layers:
    - ↑ activation of hypoxic response pathways, angiogenesis, and EMT signatures
    - ↑ apoptosis
    - ↓ proliferation capacity
    - ↑ pathways related to fatty acid metabolism, including fatty acyl CoA synthesis and fatty acid β-oxidation

  Suggests that such tumor cells close to the border could actively initiate metabolic reprogramming, important for tumour invasion

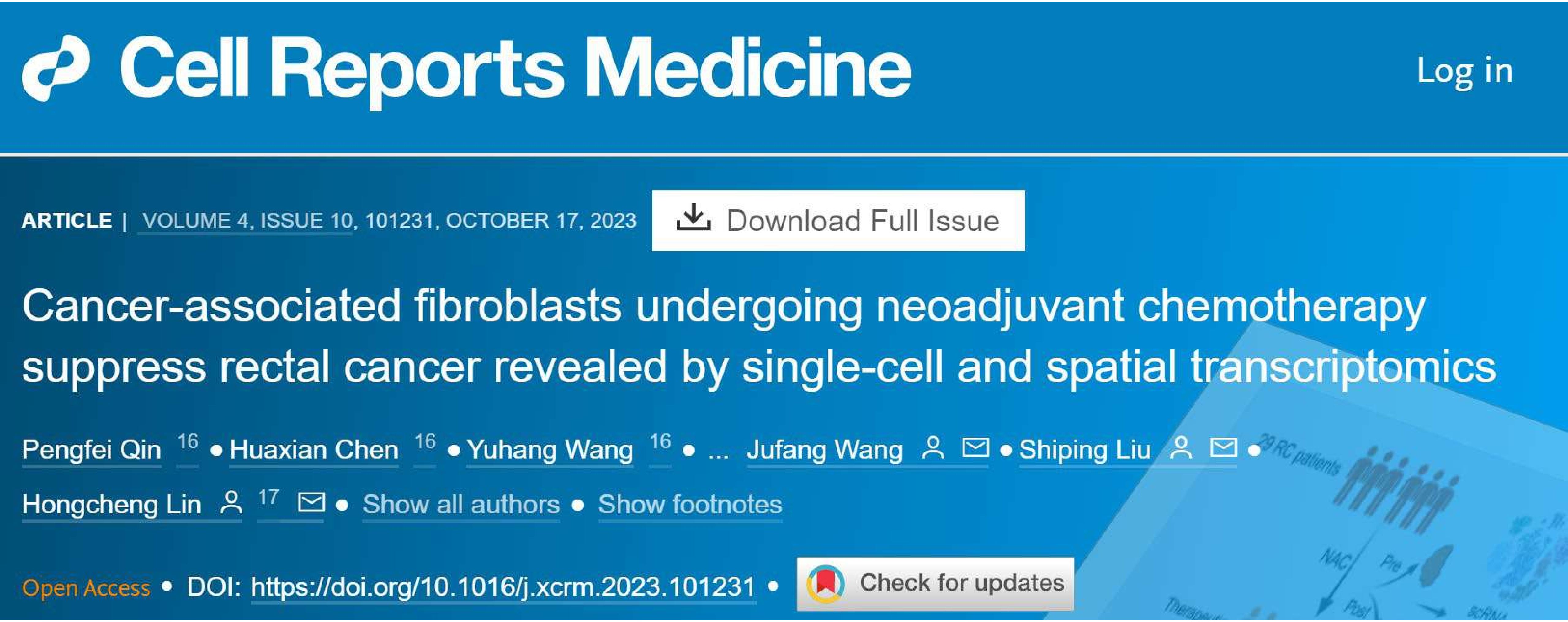
- Cells in the paratumour layer closest to the border:
    - Show enrichment in apoptosis, angiogenesis, proliferation capacity, hypoxia, and glycolysis were significantly enriched





Wu et al., 2023 (Cell Research)

Stereo-seq (SpaTial Enhanced REsolution Omics-sequencing)



ARTICLE | VOLUME 4, ISSUE 10, 101231, OCTOBER 17, 2023

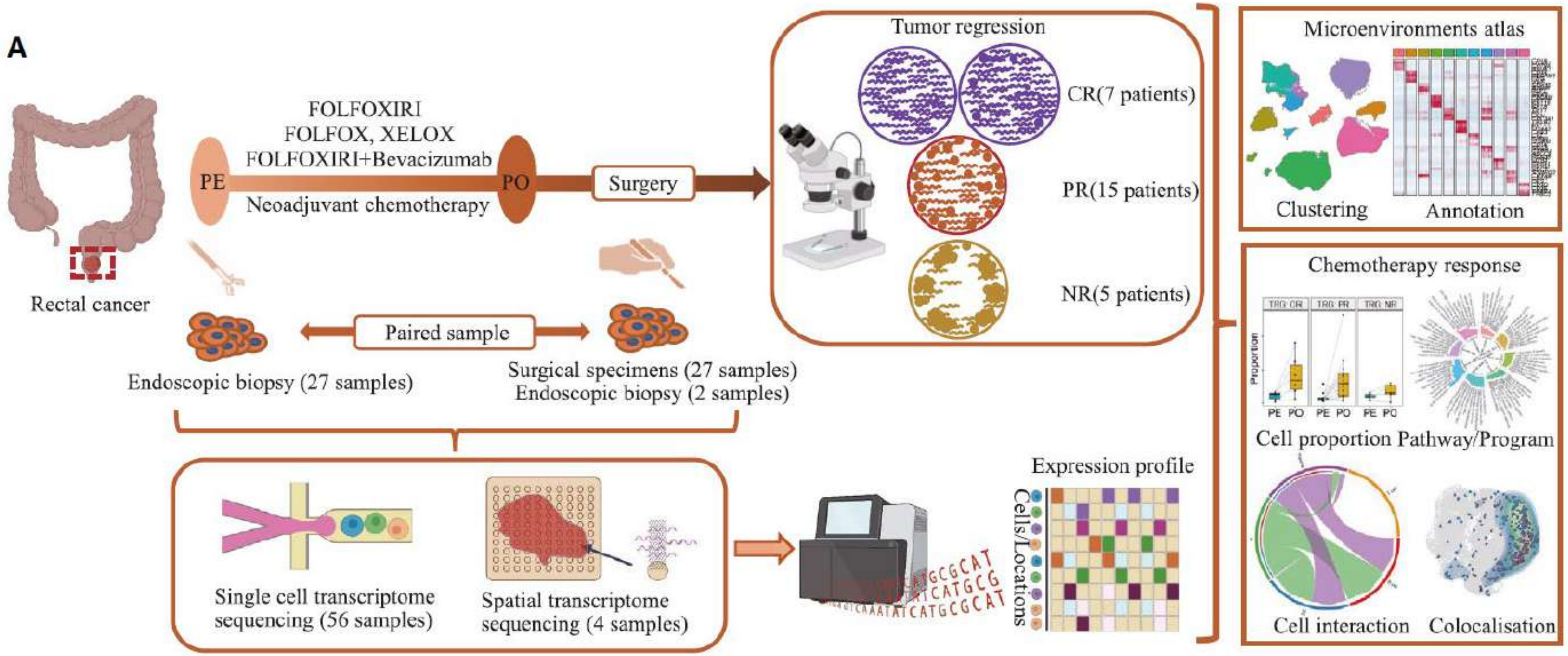
Cancer-associated fibroblasts undergoing neoadjuvant chemotherapy suppress rectal cancer revealed by single-cell and spatial transcriptomics

Pengfei Qin [16] • Huaxian Chen [16] • Yuhang Wang [16] • ... Jufang Wang • Shiping Liu • Hongcheng Lin [17] • Show all authors • Show footnotes

Open Access • DOI: https://doi.org/10.1016/j.xcrm.2023.101231 •

Qin et al., 2023 (Cell Reports Medicine)

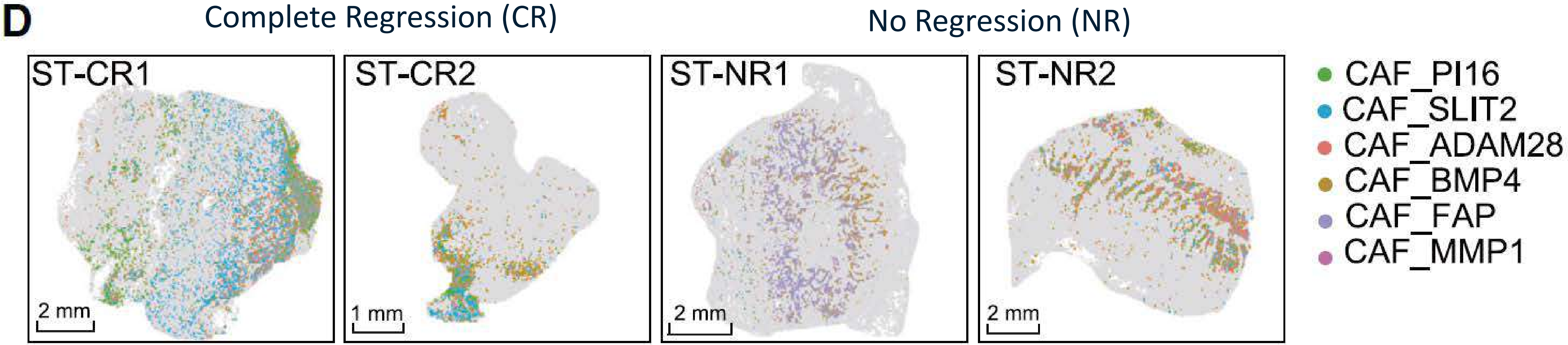
Cancer-associated fibroblast (CAF) undergoing neoadjuvant chemotherapy suppress rectal cancer



- ➢Clear viewing field and robust confirmation of cell composition, proximity, interaction and signaling pathways.
- ➢Confirmed cell-cell interactions by correlation analysis (scRNA-seq) and spatial co-localization (Stereo-seq).
- ➢Revealed CAF subset complexity and roles of tumour promoters and immune activators.

Qin et al., 2023 (Cell Reports Medicine)

Cell distribution of CAF subsets in Complete vs No Regression

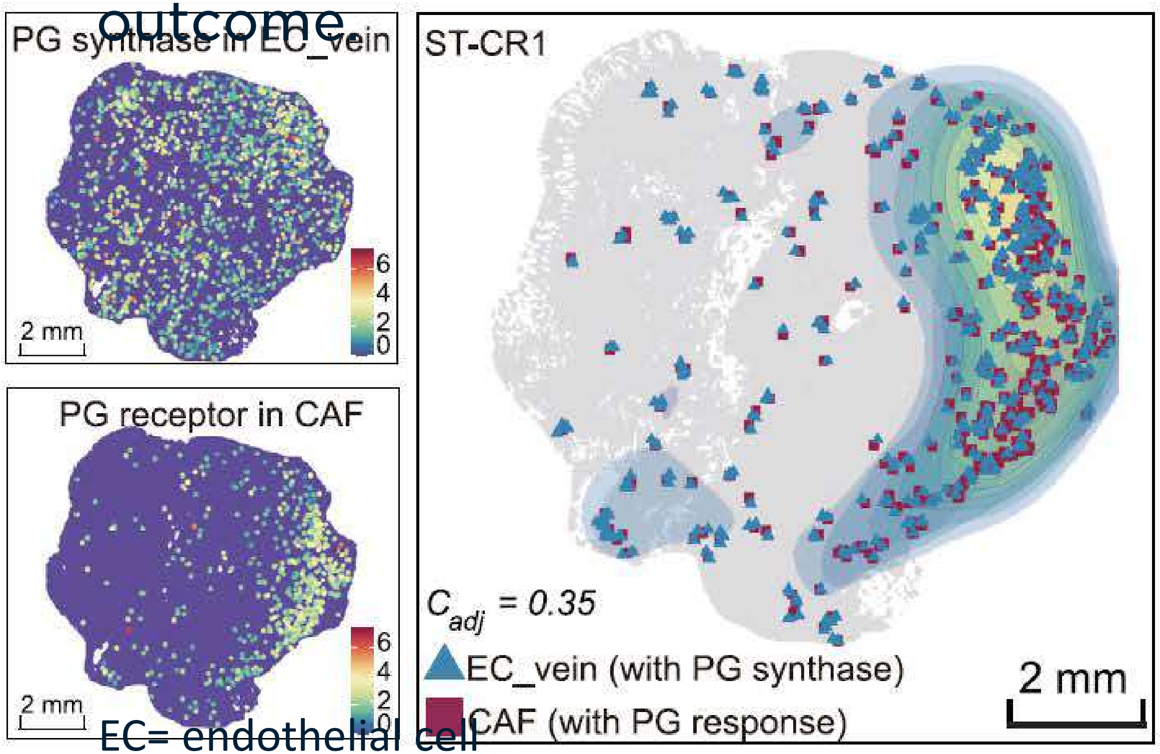


- CAF_PI16, CAF_SLIT2, and CAF_ADAM28 dominated the TME space of CR samples, whereas CAF_BMP4, CAF_FAP, and CAF_MMP1 were the major components in the TME space of NR samples
- Cell distribution of CAF subsets (CR vs NR) in Stereo-seq data is consistent with the abundance observed in the scRNA-seq data.

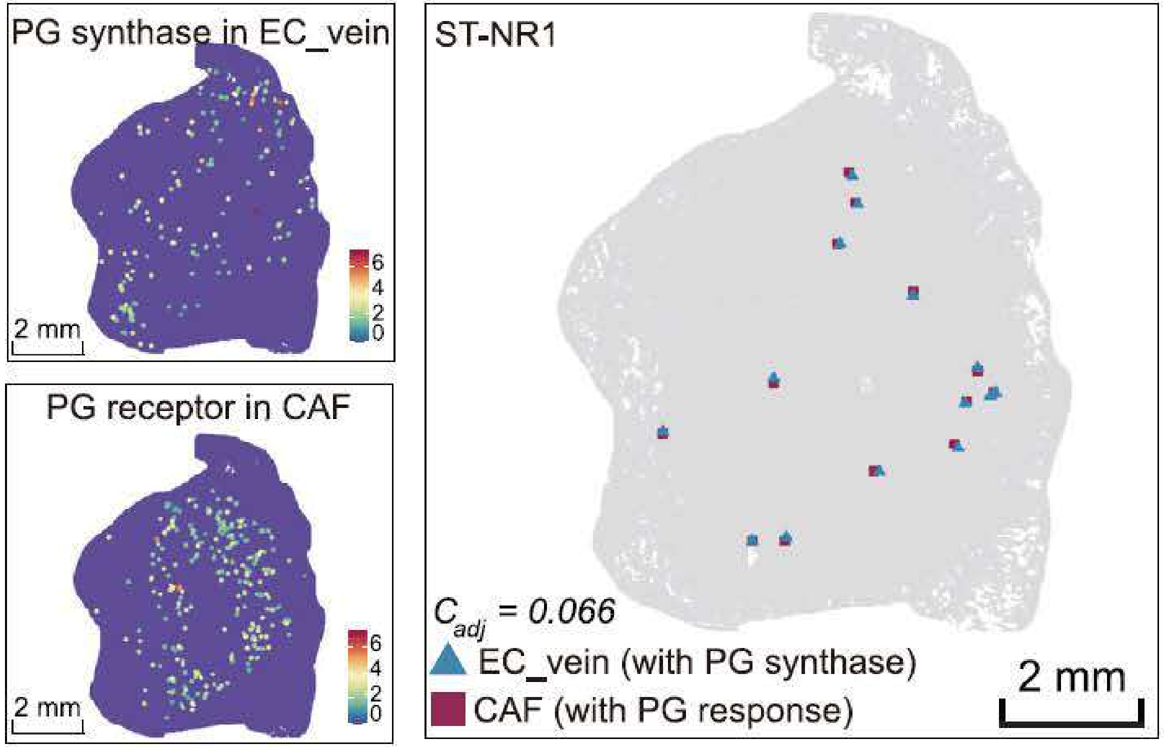
Qin et al., 2023 (Cell Reports Medicine)

Spatial co-localization of Prostagladin (PG) synthase and receptors observed in CR but not in NR

- ➢ ***Background:*** *Based on GSEA data, the most upregulated pathway in pCAF (positive response CAF) was the prostaglandin (PG) response, suggesting that PG is involved in reshaping CAF populations*
- ➢ **Significance:** PG interactions between ECs and CAFs affect rectal cancer outcome.



EC= endothelial cell



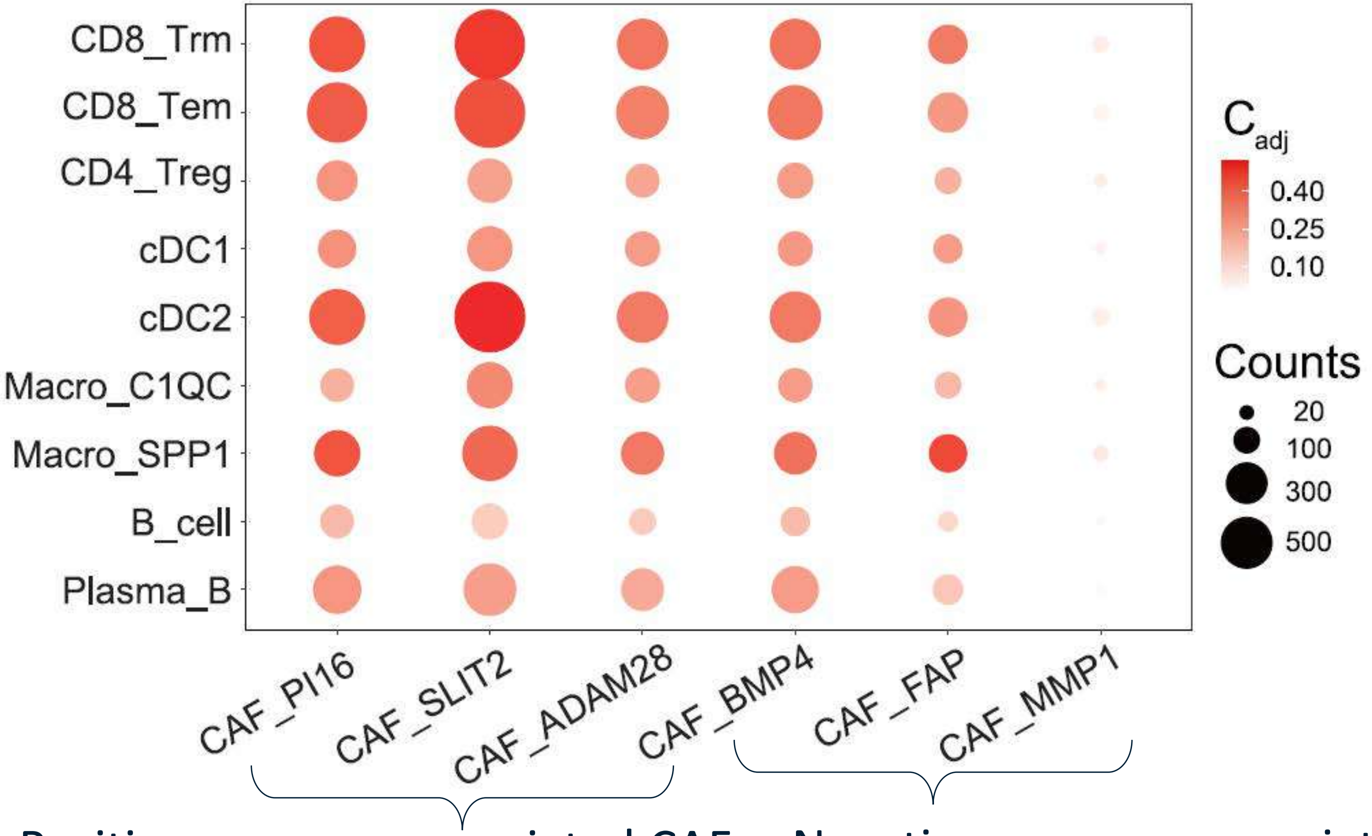
Cadj - intensity of colocalization between two cell sets

Qin et al., 2023 (Cell Reports Medicine)

mirxes
TO KNOW. TO ACT.

Spatially diverse colocalization of CAF subsets with immune cells

- CAF_PI16 strongly enriched CD8+Trm,CD8+Tem, and cDC2s in CR samples, with higher levels than negative response associated CAF (nCAF) subsets.



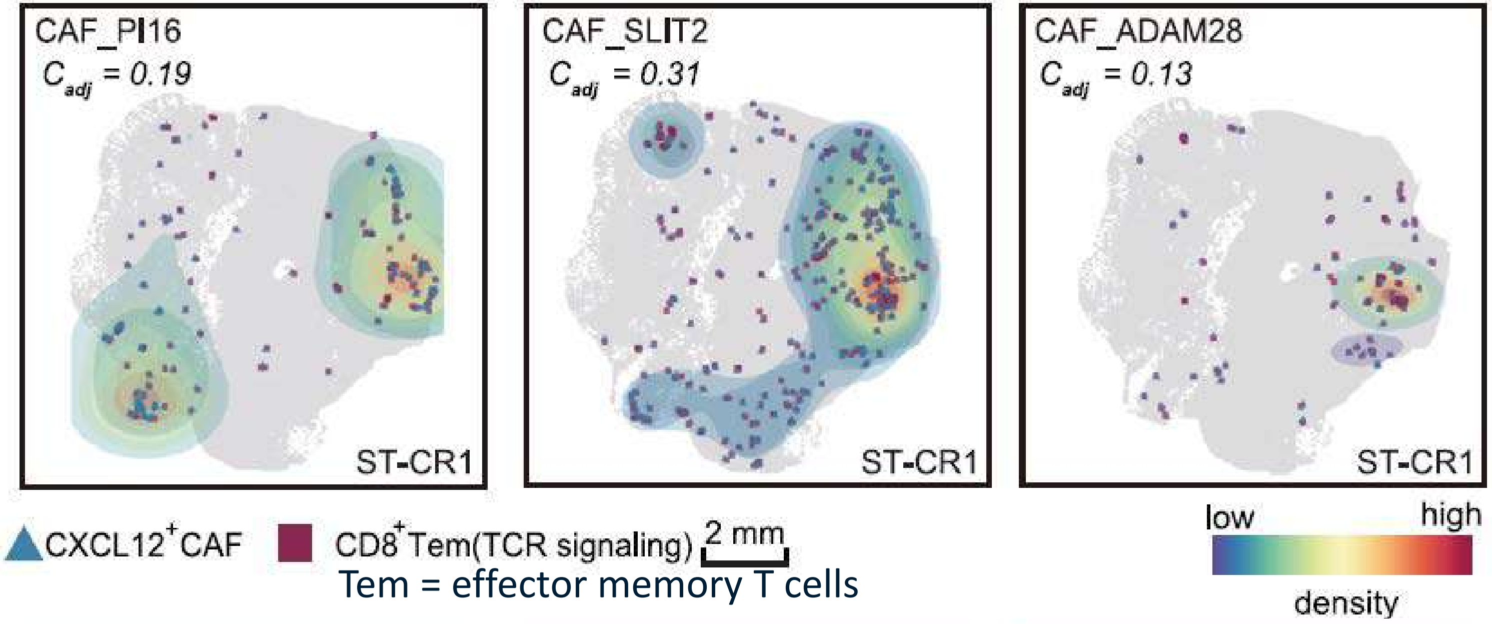
Positive response associated CAFs Negative response associated CAFs

"Counts" - number of immune cells surrounding CAFs in less than a 3-bin range.

Cadj - intensity of colocalization between two cell sets

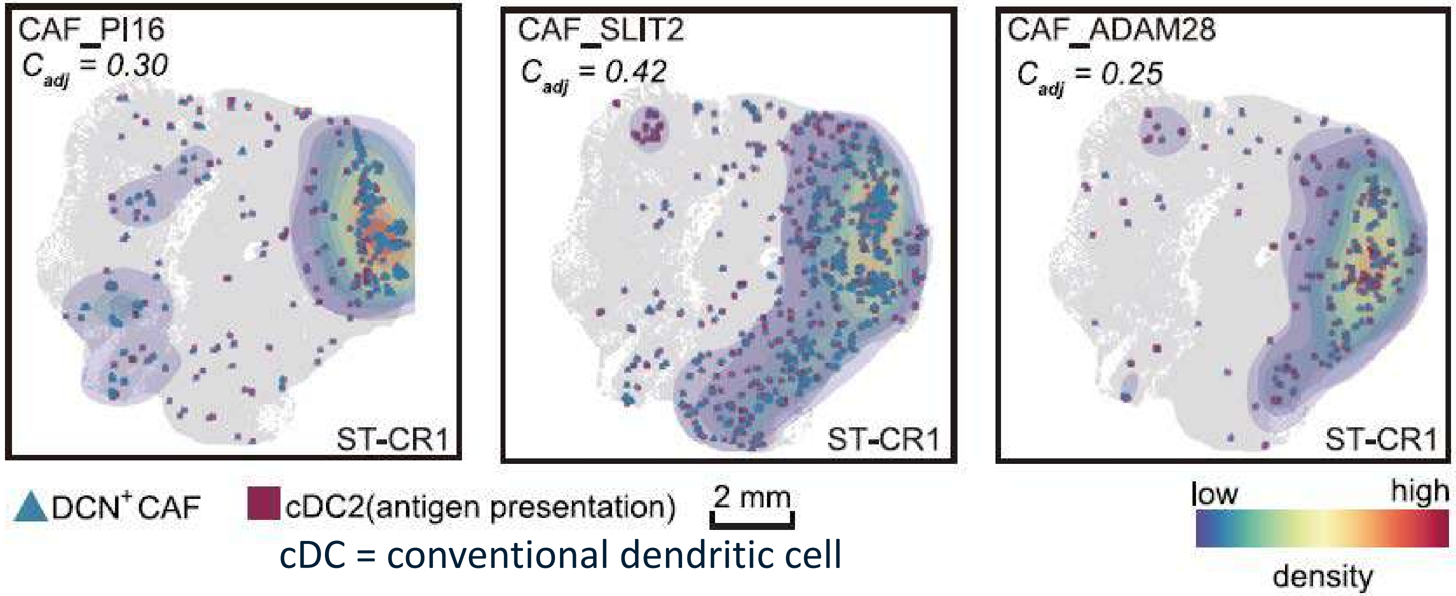
Qin et al., 2023 (Cell Reports Medicine)

Immunoregulation of positive response associated CAFs in T cells



Tem = effector memory T cells

- ➢ Spatial colocalization of CAF subsets expressing CXCL12 and CD8+Tem with active TCR pathway in the ST-CR1 sample.

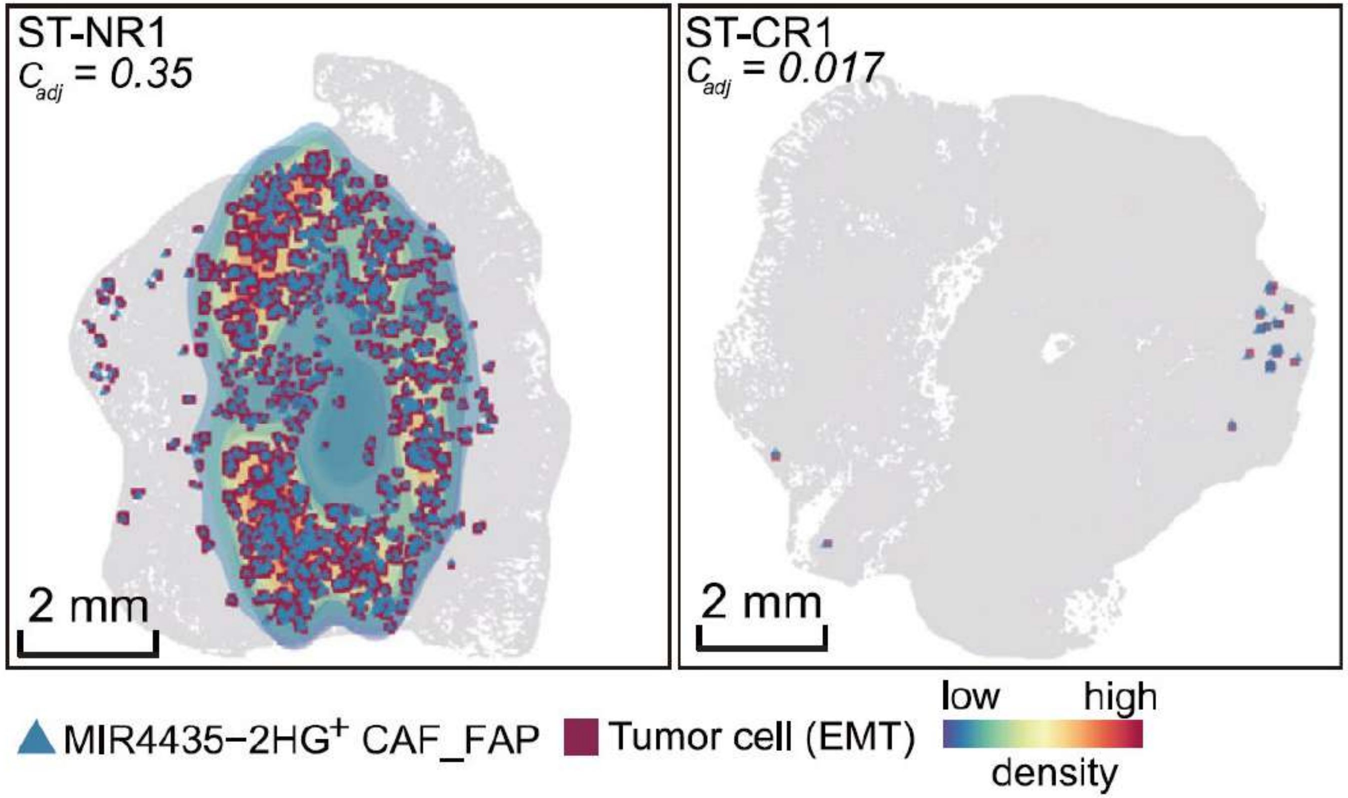


cDC = conventional dendritic cell

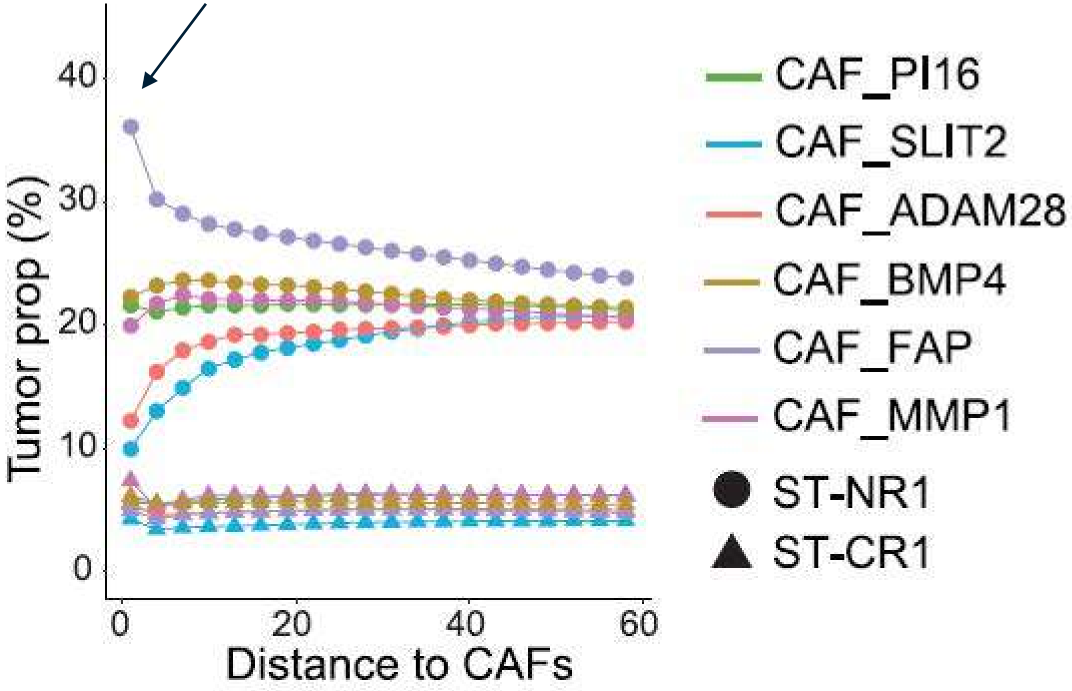
- ➢ Spatial colocalization of CAF subsets expressing DCN and cDC2s with active antigen presentation pathway in the ST-CR1 sample.

Qin et al., 2023 (Cell Reports Medicine)

Interaction between CAF_FAPs and malignant cells in NR samples but not CR samples

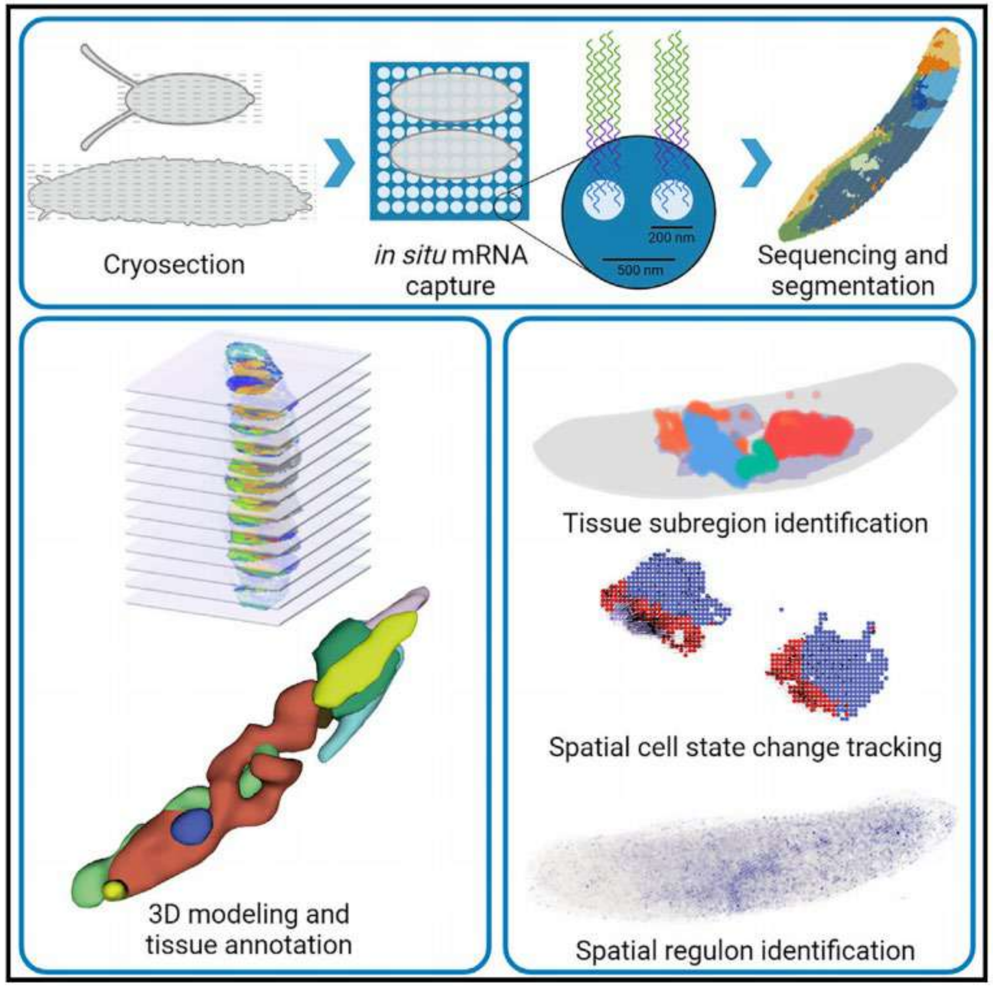


Highlight interaction between CAF_FAPs and malignant cells



Qin et al., 2023 (Cell Reports Medicine)

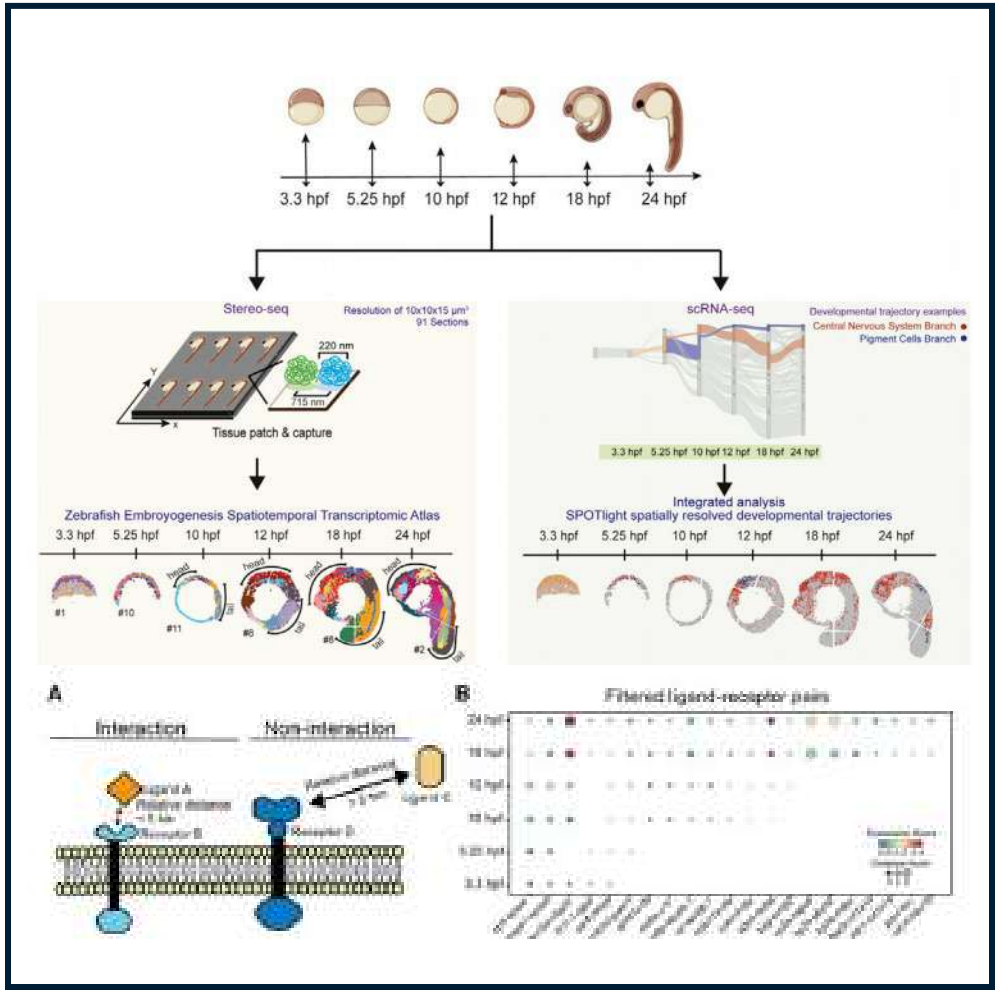
Drosophila Embryo & Larvae Development



Wang et al., 2022 (Developmental Cell)

3D reconstruction of spatial transcriptomes of *Drosophila* late-stage embryos & larvae

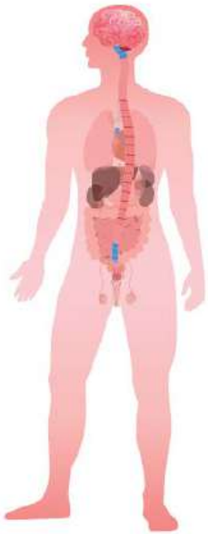
Zebrafish Embryogenesis



Liu et al., 2022 (Developmental Cell)

Identification of potential interactions (signalling pathways) based on spatial distribution of **ligand-receptor pairs**

List of Species & Tissues



- Brain
- Thyroid cancer
- Nasopharyngeal cancer
- Hypophysis
- Lymph node
- Parietal lobe
- Lung cancer
- Hepatic cyst
- Hepatic hemangioma
- Fatty liver
- Cervical cancer
- Colorectal cancer
- Lymph node cancer
- Ovarian cancer
- Brain cancer

- Ovary
- Vagal Complex
- Prefrontal lobe
- Embryonic heart
- Kidney
- Liver
- Whole brain
- Pancreas
- Heart
- Thymus
- Colon
- Cecum
- Testis
- Jejunum
- Foot Pad

Arabidopsis

Soybean

Drosophila

Zebra fish

Salamander

Macaque

mirxes

TO KNOW. TO ACT.

Thank you!

email: craigmorley@mirxes.com
email: michaelfanelli@mirxes.com

CID: Coordinate Identity Barcode

Captures the XY coordinates of the chip, Each chip has a unique set of barcodes.

DNB: DNA Nanoball

Capture spot on stereo-seq chip

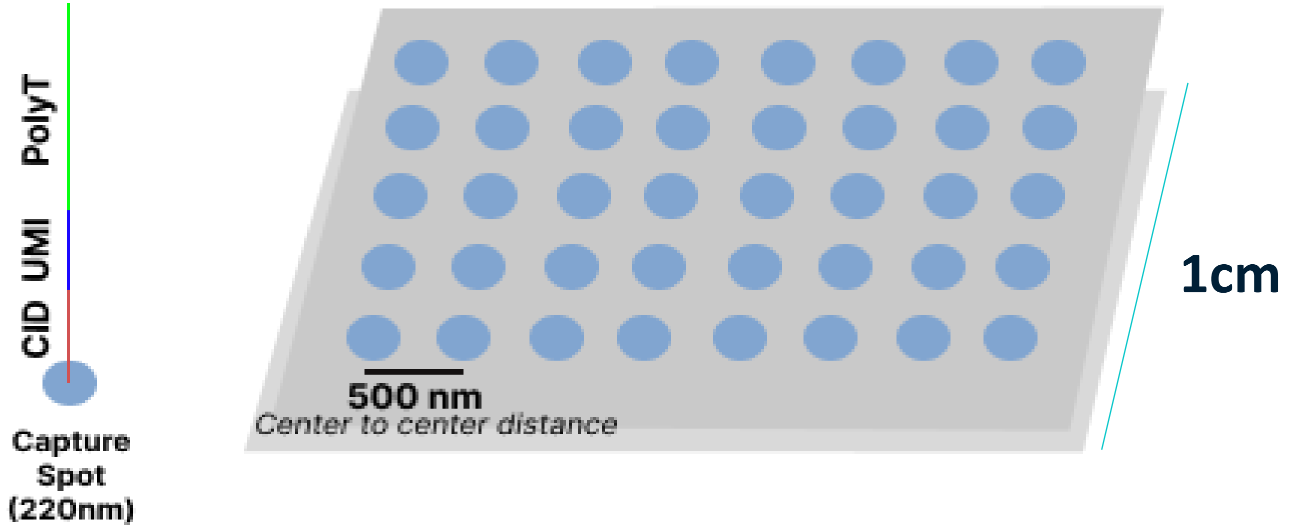
Mask File

Vendor generated file that contains CID Barcode information. Identified by unique Chip ID.

MID: Molecular Identity Barcode

Short barcode sequences used for uniquely tagging each molecule. UMI and MID terms are inter-changeably used.

STEREO-SEQ CHIP



FASTQ Files

Q30 Scores

Q30%: Percentage of bases with phred quality score greater than 30, indicating 1 in 1000 chance of base calling error.

- CID (Coordinate Identity) Q30
- UMI (Unique Molecular Identity) Q30
- Seq Q30

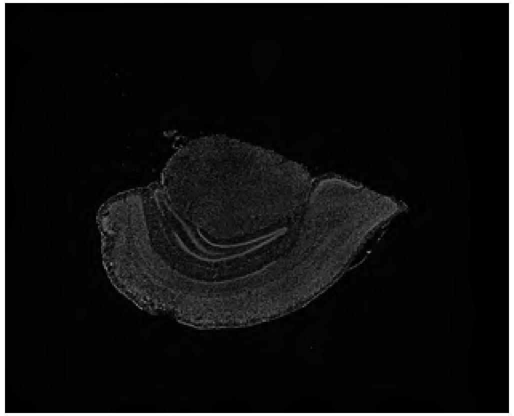
A preliminary QC is performed on the fastq file to determine the quality of the data.

Q30 cutoff is 85%. The scores are consistently achieved above the cutoff, for almost all cases.

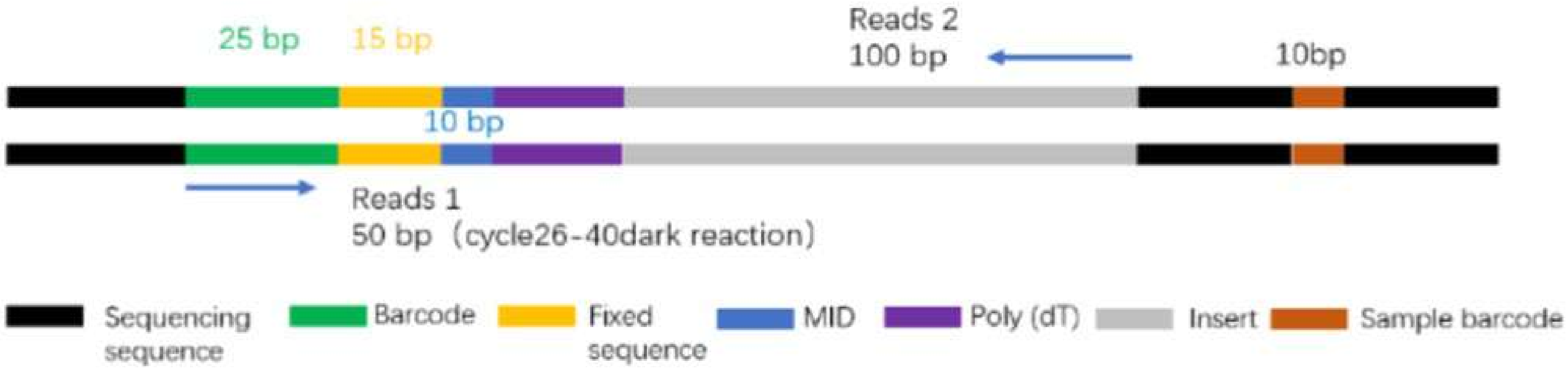
ImageQC software is used for tissue image QC.

It provides the score, indicating the ability of the software to detect the track lines captured in the image. The detected track lines enables the superimposition of tissue image on the count matrix.

Tissue Image



QC Score	Pass/Fail
Image Track line QC Score	Required >= 0.5



FASTQ FILES

Files containing paired end sequencing data.

Read1 – CID+ UMI
Read2 – PolyA transcript

MASK FILES

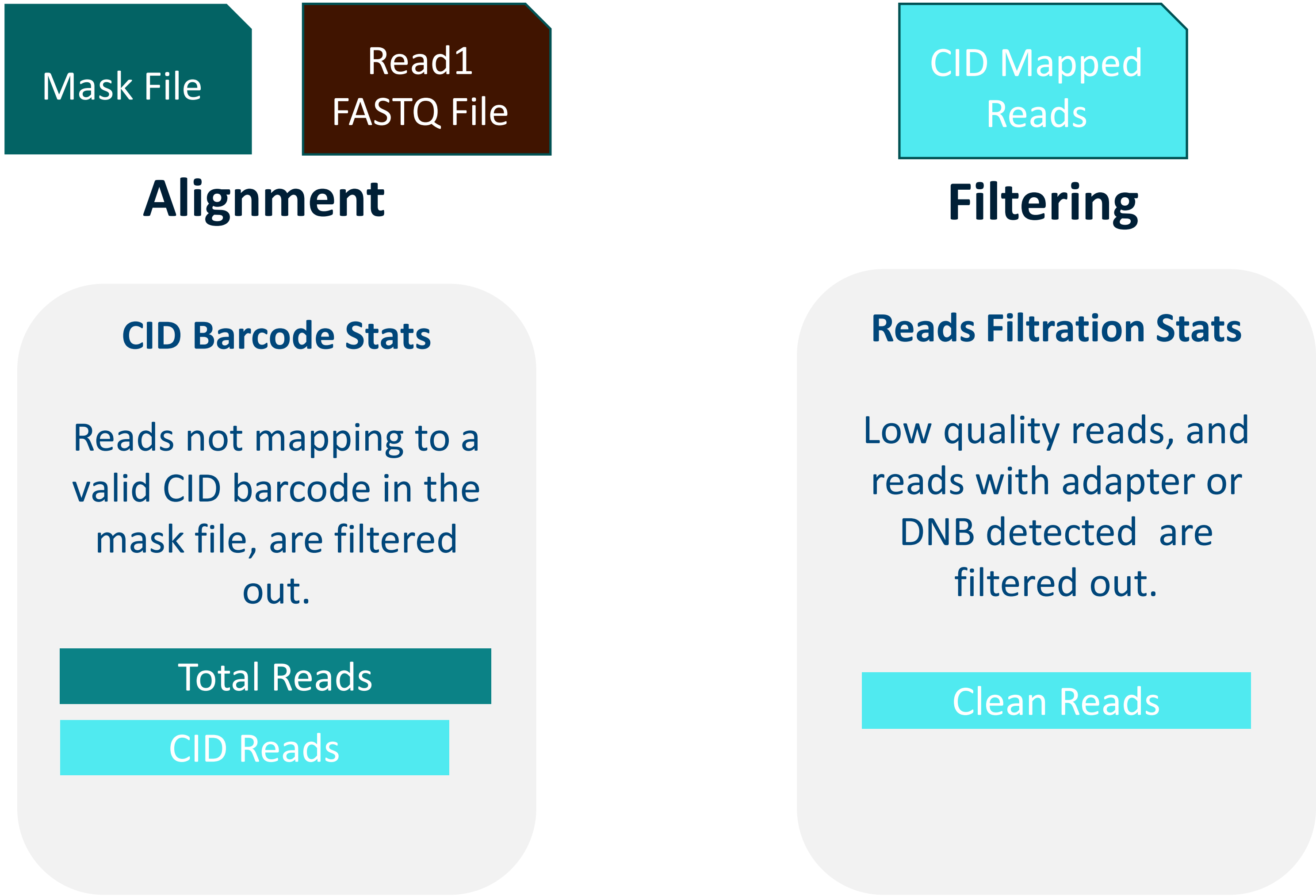
Vendor generated file identified by unique chip ID number.

Mask file contains spatial coordinates information

STAINED TISSUE IMAGE

Stained tissue image in .tif format.

It is used for obtaining tissueCut gene count matrix.



Mask File

Read1 FASTQ File

Alignment

CID Barcode Stats

Reads not mapping to a valid CID barcode in the mask file, are filtered out.

Total Reads

CID Reads

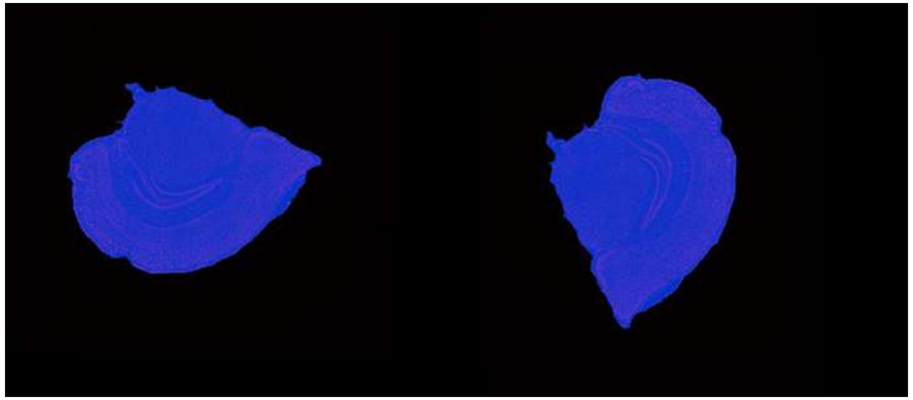
CID Mapped Reads

Filtering

Reads Filtration Stats

Low quality reads, and reads with adapter or DNB detected are filtered out.

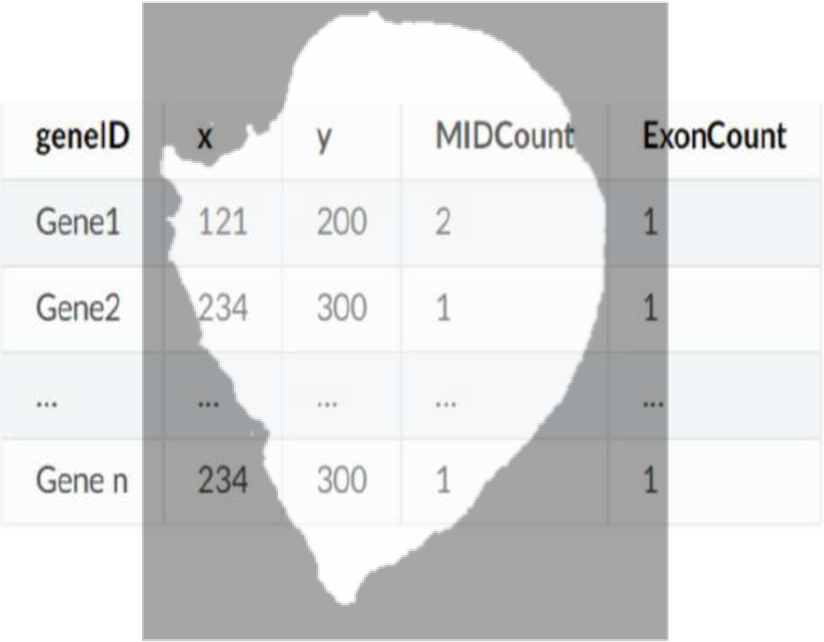
Clean Reads



Alignment of Image & the Gene Count Matrix

Microscopic stained tissue image is aligned to the Gene Expression Matrix (GEM), with the help of the track lines detected.

The image is rotated, flipped and/or scaled to achieve the alignment between the coordinates in the GEM file and the image track lines.



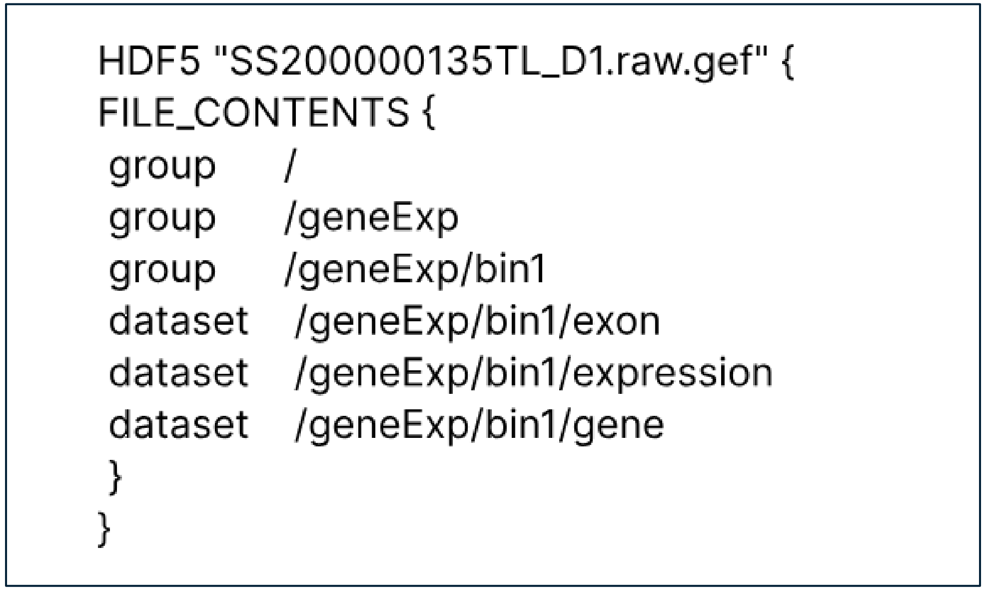
geneID	x	y	MIDCount	ExonCount
Gene1	121	200	2	1
Gene2	234	300	1	1
...	...	...	...	...
Gene n	234	300	1	1

Overlay Image the gene count matrix to extract expression from tissue covered area

GeneID	X	Y	MIDCount	ExonCount
Gene1	1000	1200	2	0
Gene1	1200	1200	3	1
Gene2	1300	1350	1	1

GEM Format – Gene Expression Matrix Format

It is a tab separated format, describing the gene count information at each capture spot.

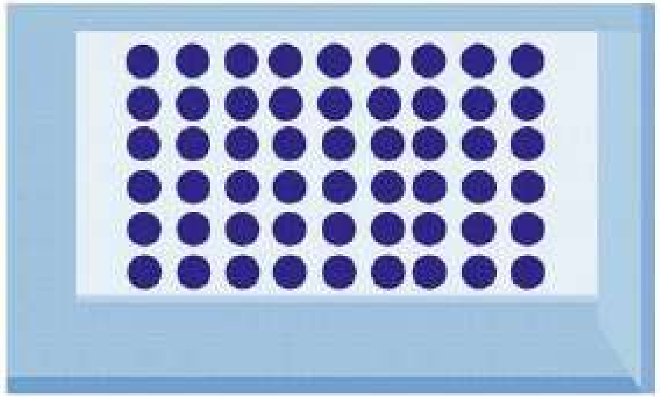


```
HDF5 "SS200000135TL_D1.raw.gef" {
FILE_CONTENTS {
 group      /
 group      /geneExp
 group      /geneExp/bin1
 dataset    /geneExp/bin1/exon
 dataset    /geneExp/bin1/expression
 dataset    /geneExp/bin1/gene
 }
}
```

GEF Format – Gene Expression File Format

It is a hierarchical format, describing the same information as GEM format. This format is widely used for downstream analysis, as it requires shorter time to read and parse the file contents.

SQUARE BINNING



Representing Bin3 example, demarcating 9 spots as one cell.

Aggregating the expression of sub-cellular capture spots, in a square pattern, mimicking the average cell size. For e.g. – Bin20 corresponds to approximately 10 microns size.

CELL SEGMENTATION

Representing cell boundaries determined by the cell segmentation algorithm

Defining single cell boundaries, using stained tissue image and gene expression data, using **Stereocell**.

ImageJ software can be used for Image visualization.

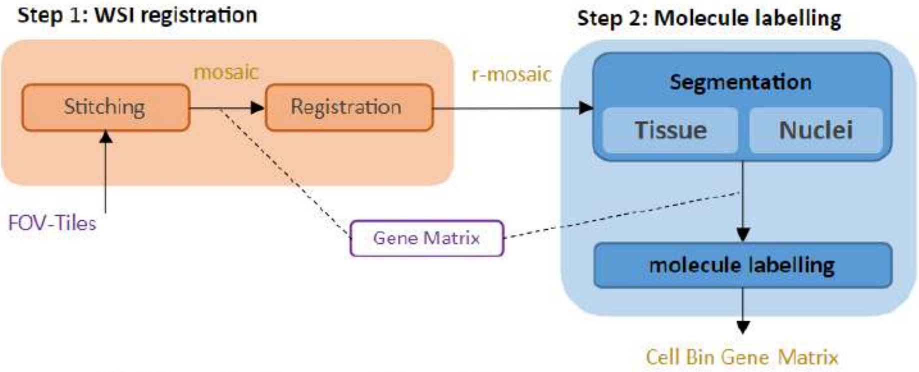


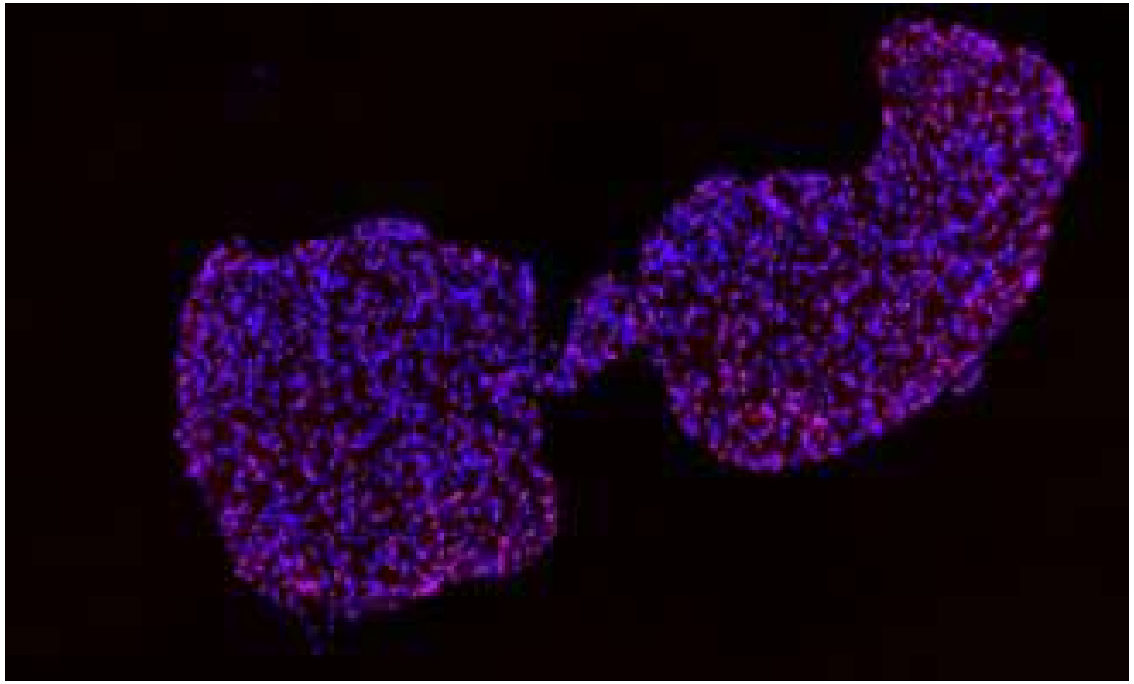
Fig Schematic diagram of StereoCell

The algorithm helps generate cellular mask using the gene expression map and the nuclei-stained image.

Other Tools

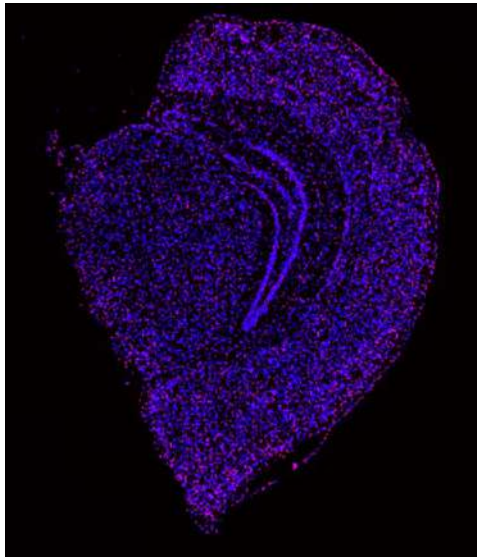
Cellpose

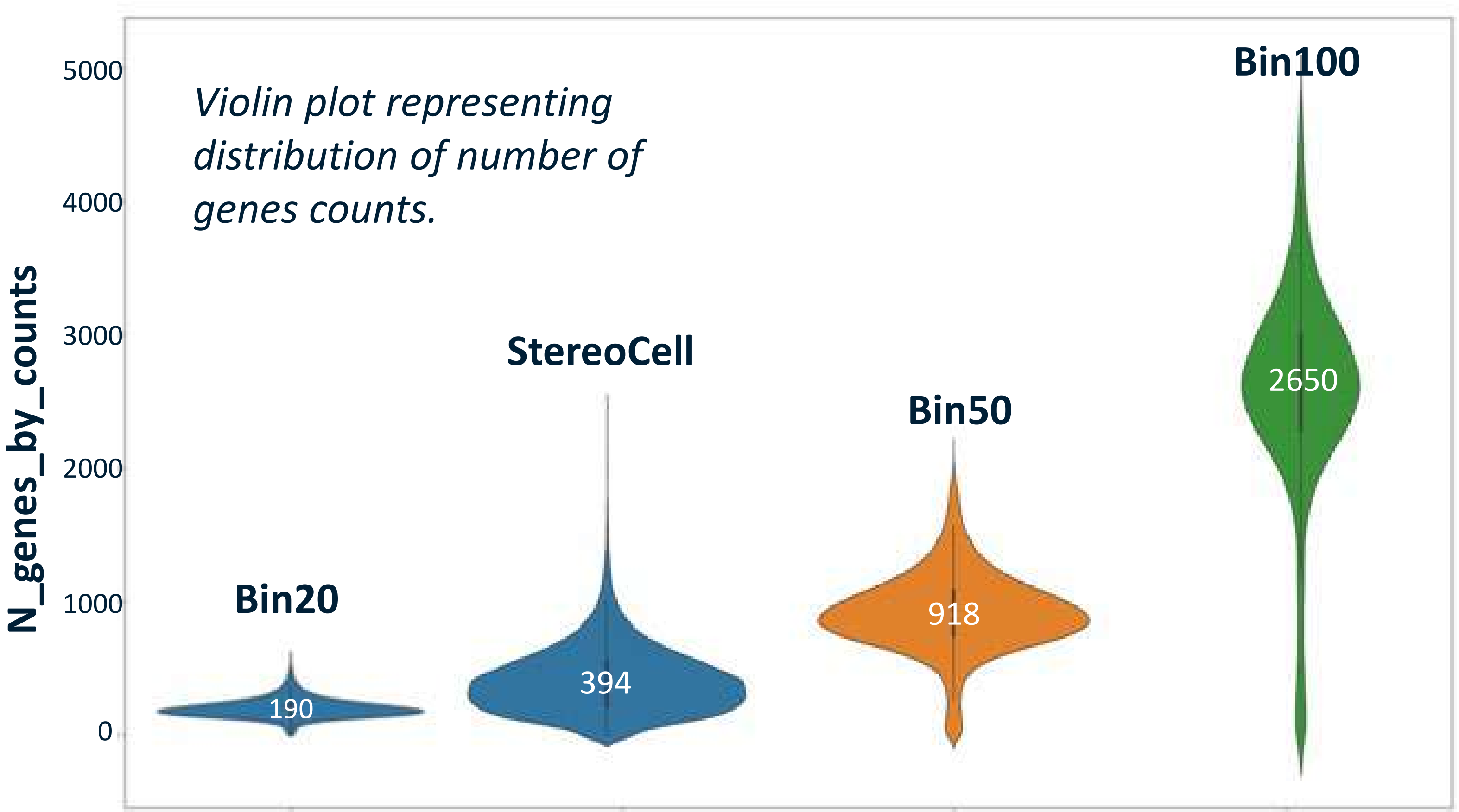
Spateo



Overlap map of nuclei-stained image and cell mask generated by the cell segmentation algorithm

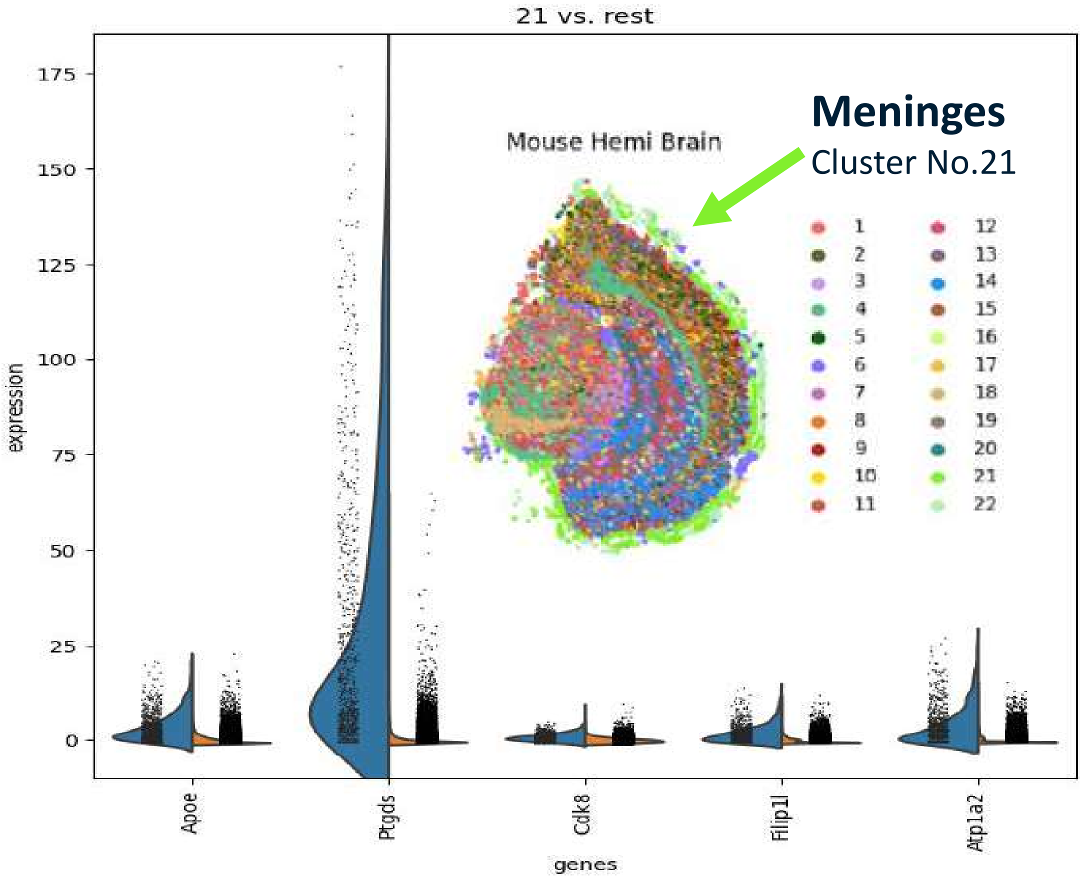
The map is an indicator that all the nuclei stains are accurately labelled by the cell mask





Violin plot representing distribution of number of genes counts.

	Mean Gene Count	Mean MID Count	Number of Cells
Bin20	199.1	386.8	120,210
Bin50	933	2377.5	19,557
Bin100	2594	9249.6	5,027
Stereo Cell	425.3	943.5	31,345

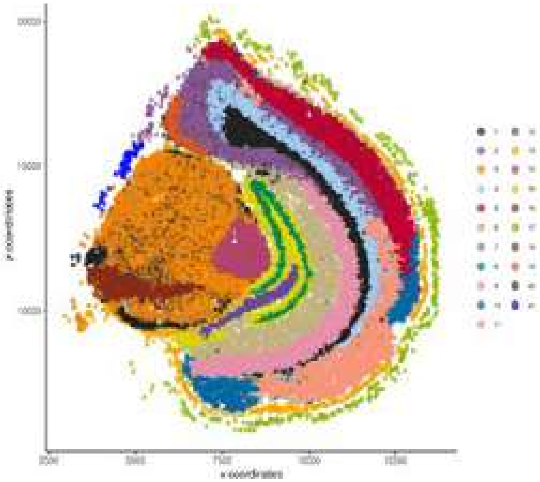


Meninges
Cluster No.21

Spatially aware clustering, and identification of highly variable gene in cluster no. 21, confirms the identification of Meninges.

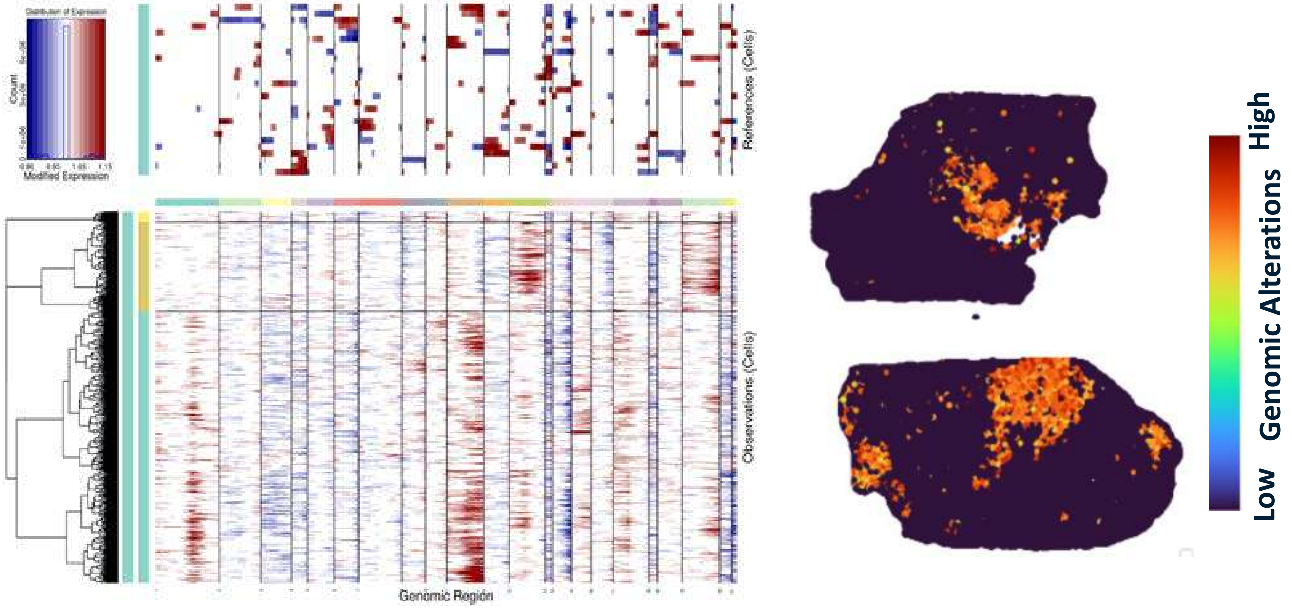
Tools used: DeepST & Scanpy

Other Tools



Banksy

Ref - Chen A et al, Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. Cell. 2022 May 12;185(10):1777-1792.e21. doi: 10.1016/j.cell.2022.04.003. Epub 2022 May 4. PMID: 35512705.

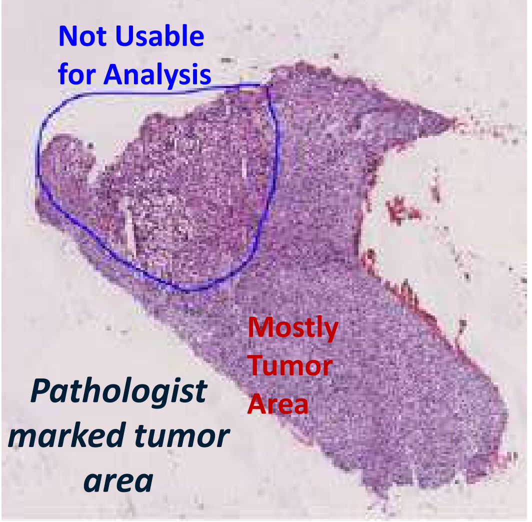


🟩 **Reference cells**

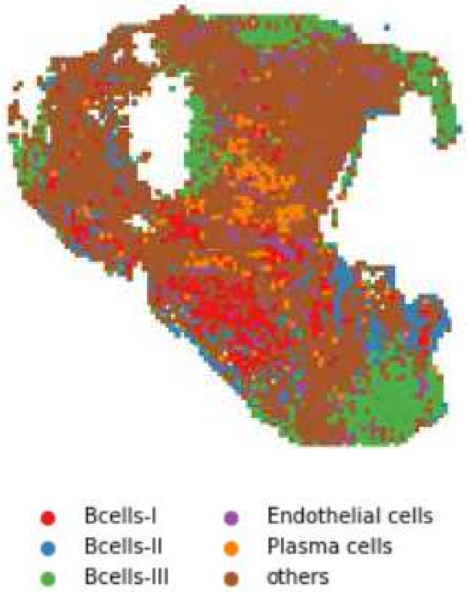
🟩 **Tumor -I** 🟨 **Tumor-II** 🟨 **Negative control**

Identification of Copy Number Variations in cells across a tissue & mapping the cells back to their original location

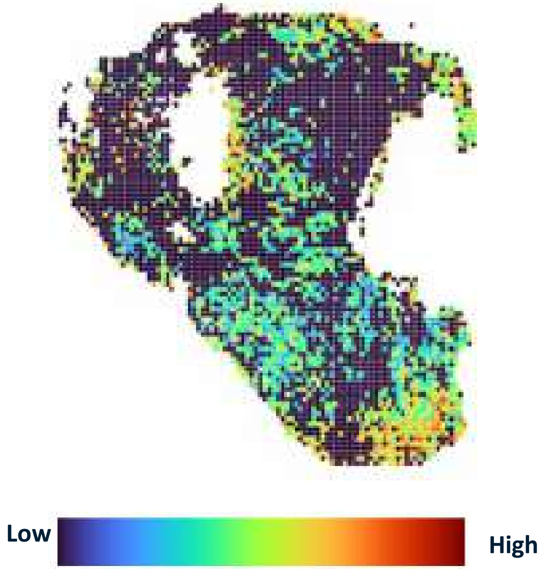
Corroborating CNV results with Pathologists annotated HnE tissue images



Annotation of cell types of interest



Spatial InferCNV Analysis



Erickson, A., He, M., Berglund, E. et al. Spatially resolved clonal copy number alterations in benign and malignant tissue. Nature 608, 360–367 (2022). https://doi.org/10.1038/s41586-022-05023-2

A GUI based tool for preliminary visualization of secondary files

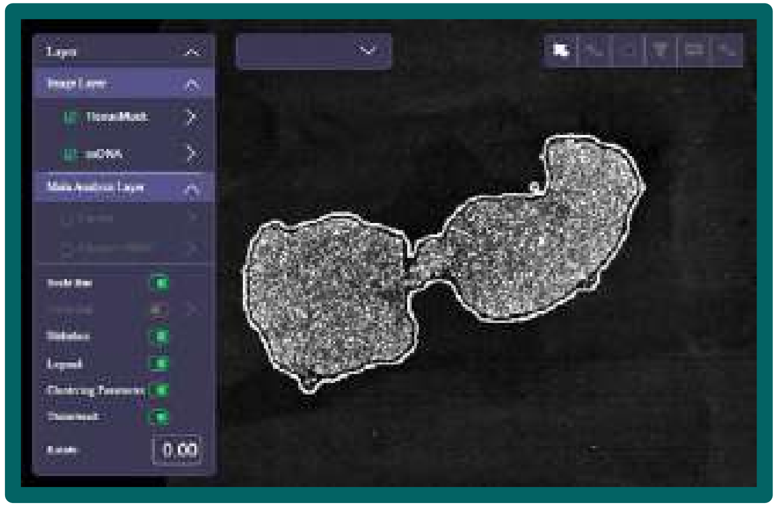
Download the tool to visualize

- Gene Count Matrix (GEF Files only)
    - Spatial scatter plots at different bin sizes
    - Expression of genes of interest
- Image RPI files

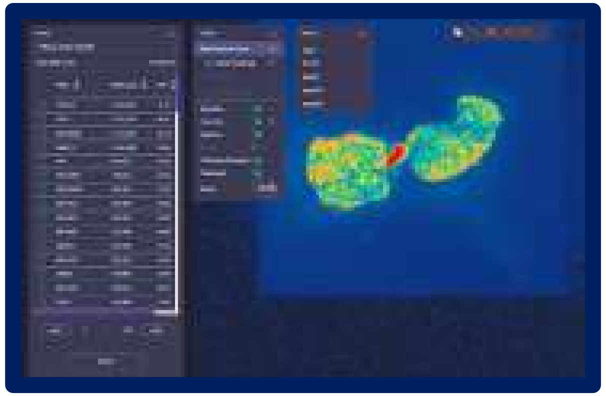
This file is generated during the image registration step, in secondary analysis workflow

    - Image layers like ssdna stain, tissue mask, other stained image layers (if available)
- H5ad combined cellbin GEF
    - Visualize clustering results

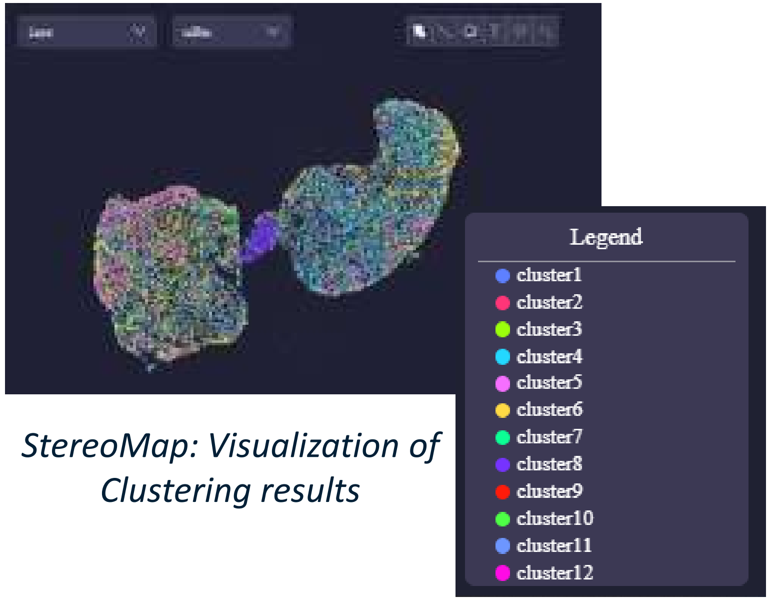
Kindly note – StereoMap can not perform downstream analysis, however, can only be used for visualization



StereoMap: Overlay of tissue mask with ssdna image

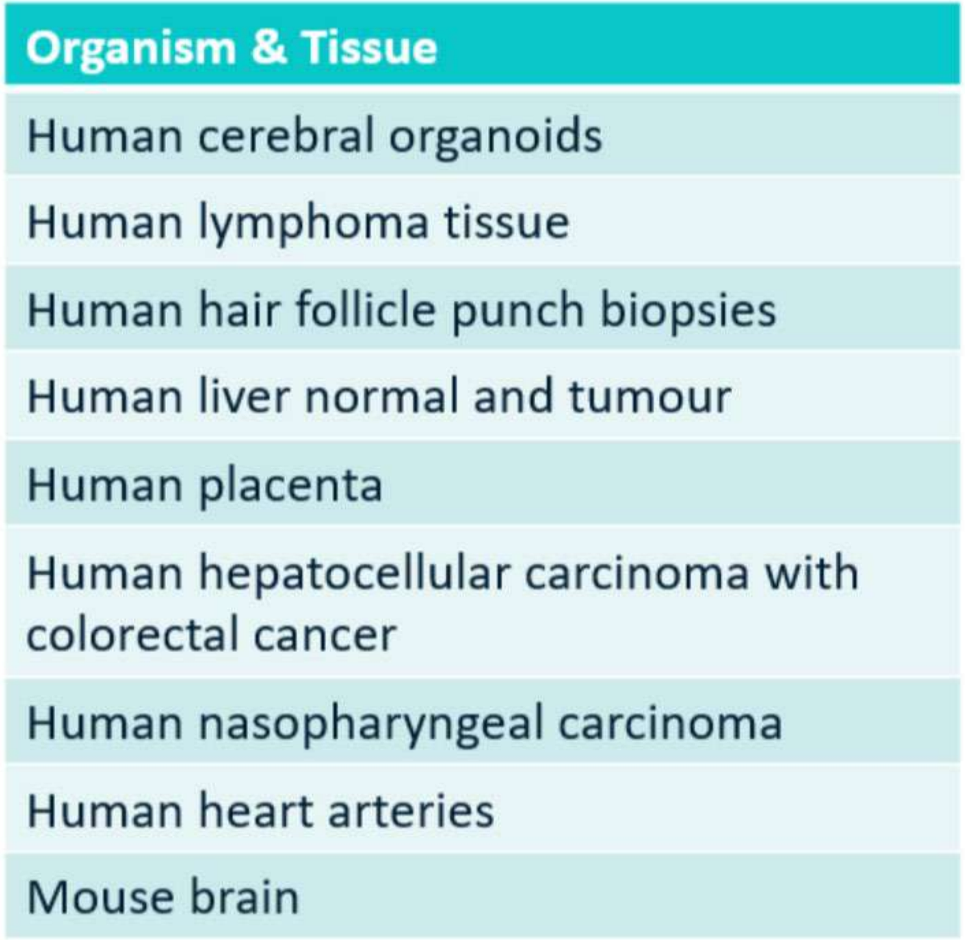


StereoMap: Visualization of GEF file



StereoMap: Visualization of Clustering results

100% Stereo-seq success rate



Organism & Tissue
Human cerebral organoids
Human lymphoma tissue
Human hair follicle punch biopsies
Human liver normal and tumour
Human placenta
Human hepatocellular carcinoma with colorectal cancer
Human nasopharyngeal carcinoma
Human heart arteries
Mouse brain