Combining single-cell and Spatial Transcriptomics data: case study on human fetal heart

Stefania Giacomello

Why do we need spatial resolution?

Understanding how cell localization in the tissue influences gene expression

How adjacent regions in tissues interact at gene expression level

Cell fate decided by several morphogens whose gradients originate from different regions of the embryo

How do we achieve spatial resolution?

- Computational methods

- Spatial transcriptomics methods (ST, ISS, FISSEQ, imaging - smFISH)

Computational approaches

Seurat

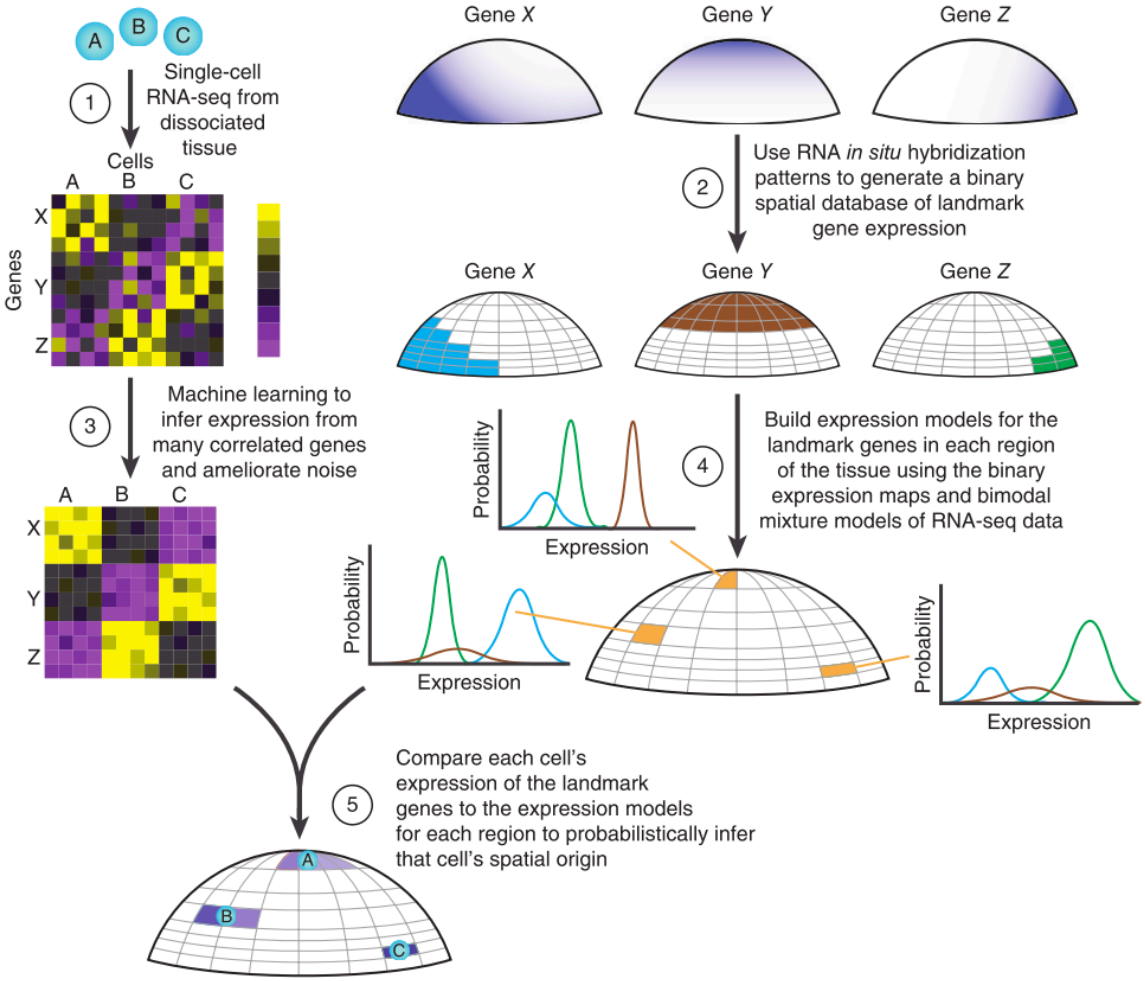
Spatial reconstruction of single-cell gene expression data

Rahul Satija[1,7,8], Jeffrey A Farrell[2,8], David Gennert[1], Alexander F Schier[1–5,9] & Aviv Regev[1,6,9]

NATURE BIOTECHNOLOGY VOLUME 33 NUMBER 5 MAY 2015

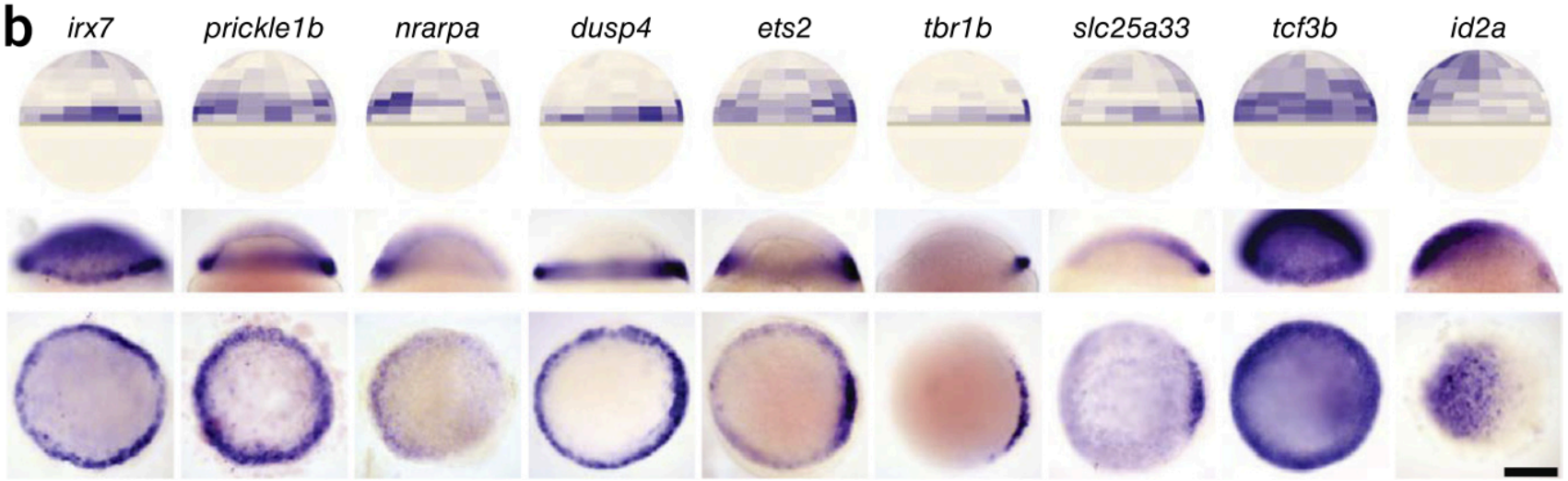
- Applied to zebrafish embryo
- Seurat combines cells' gene expression profiles (scRNA-seq) with a set of 'landmark' genes (*in situ* hybridization) to guide spatial assignment

Seurat



- 47 ISH genes
- 128 bins (64 L-R symmetry)
  ~40–120 cells per bin,
  from in situ expression domain
- 851 single cells

Seurat



- 47 ISH genes
- 128 bins (each ~40–120 cells), based on in situ expression domain → 64 bins due to left-right symmetry
- 851 single cells (no cells with less than 2000 genes)

Seurat – pros & cons

- Bins could be reduced to the single-cell level (each cell in each position has a distinct and reproducible gene expression identity and position)

- Seurat relies on the spatial segregation of gene expression patterns to construct a reference map → tissues such tumors (no guarantee of reproducible spatial patterning), or tissues where cells have highly similar expression patterns and are spatially scattered across a tissue (i.e. adult retina)?

DistMap

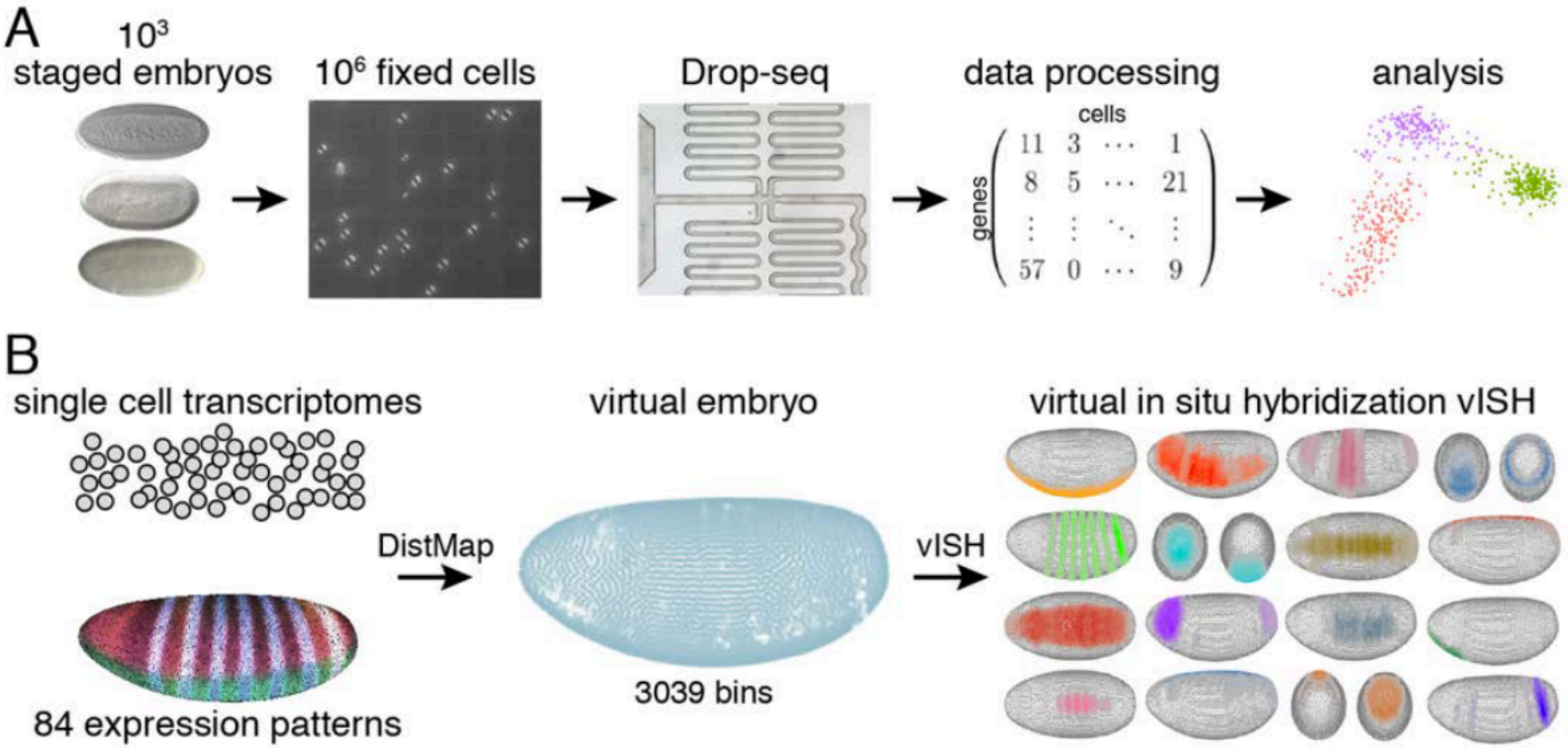
The *Drosophila* embryo at single-cell transcriptome resolution

Nikos Karaiskos[1,*], Philipp Wahle[2,*], Jonathan Alles[1], Anastasiya Boltengagen[1], Salah Ayoub[1], Claudia Kipar[2], Christine Kocks[1], Nikolaus Rajewsky[1,†], Robert P. Zinzen[2,†]

Science 31 Aug 2017:
eaan3235
DOI: 10.1126/science.aan3235

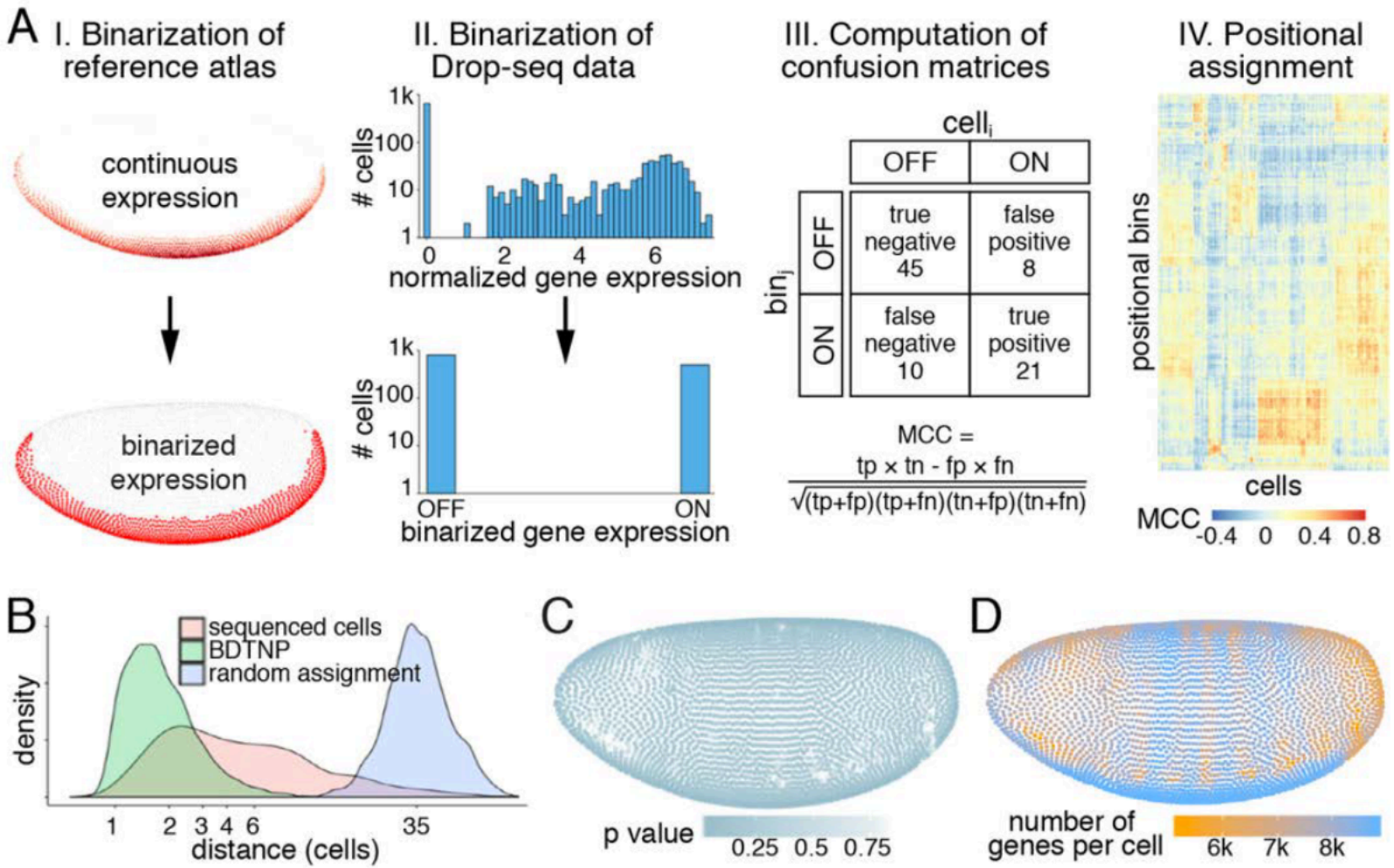
- Reconstruct the embryo and to predict spatial gene expression approaching single-cell resolution
- Seurat was not giving enough resolution→ obtained 87% of cells in the embryo are confidently resolved and depth (>8000 genes/cell)

DistMap



- in situ hybridization data for 84 genes, resulting in a quantitative high-resolution gene expression reference atlas with substantial combinatorial complexity

DistMap


A
I. Binarization of reference atlas
continuous expression
binarized expression
II. Binarization of Drop-seq data
cells
1k
100
10
1
0
2
4
6
normalized gene expression
OFF
ON
binarized gene expression
III. Computation of confusion matrices
cell_i
OFF
ON
bin_j
OFF
ON
true negative 45
false positive 8
false negative 10
true positive 21
MCC =
tp × tn - fp × fn
√(tp+fp)(tp+fn)(tn+fp)(tn+fn)
IV. Positional assignment
positional bins
cells
MCC
-0.4
0
0.4
0.8
B
sequenced cells
BDTNP
random assignment
density
1
2
3
4
6
35
distance (cells)
C
p value
0.25
0.5
0.75
D
number of genes per cell
6k
7k
8k

DistMap – pros & cons

- Bins are very small and the number of genes detected is high
- Spatial segregation of gene expression patterns to construct a reference map

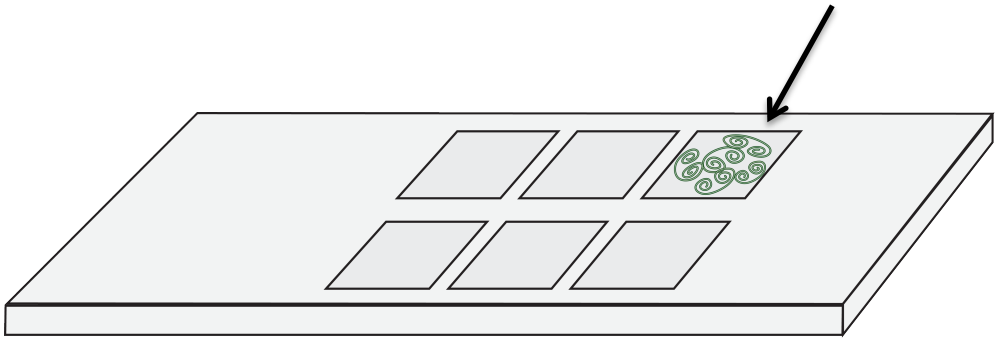
Wet lab approaches

Spatial Transcriptomics

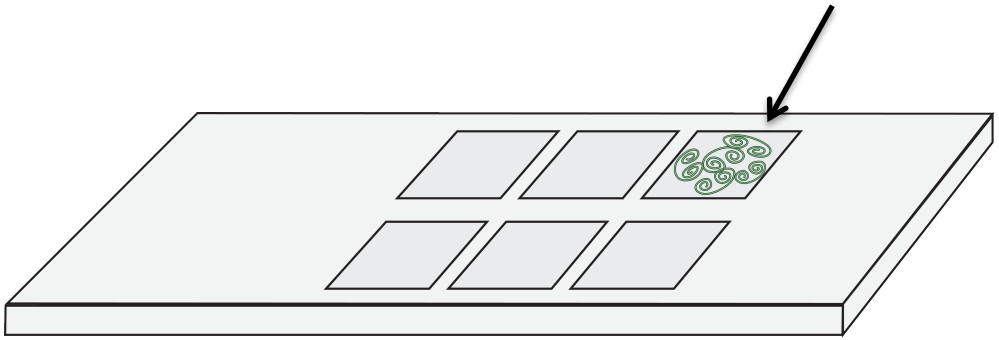
2D gene expression map of a tissue section

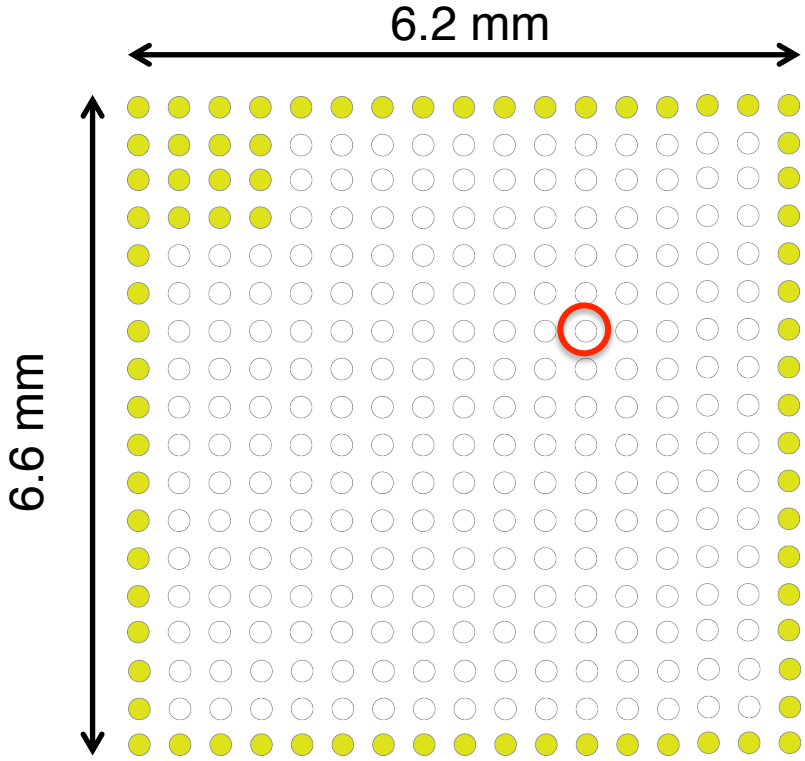
Study functional and developmental aspects

The concept



The concept



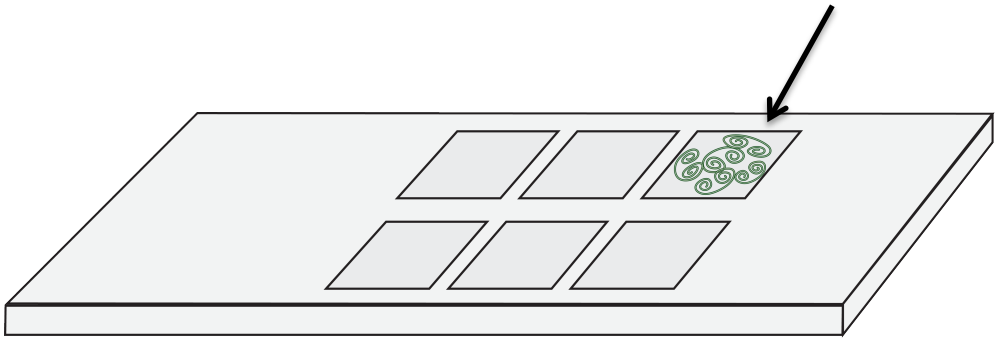


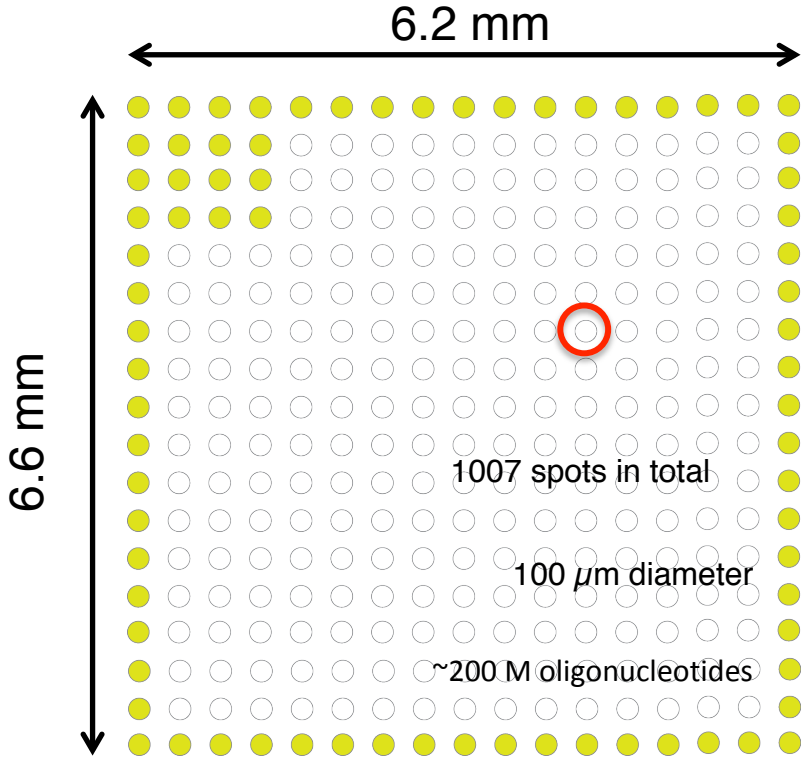
1007 spots in total

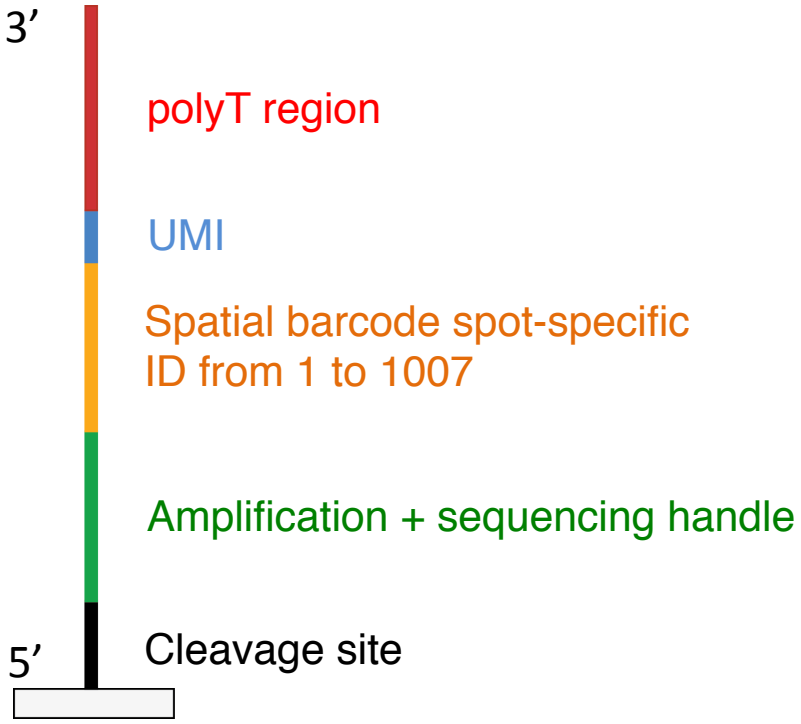
100 μm diameter

~200 M oligonucleotides

The concept

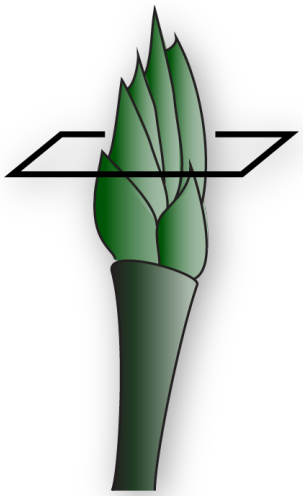


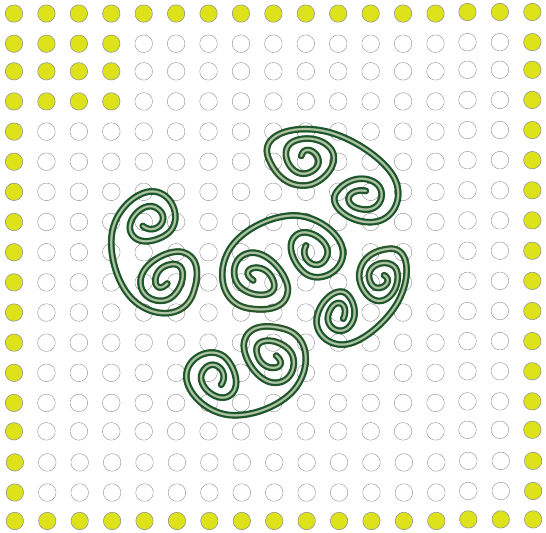

6.2 mm
6.6 mm
1007 spots in total
100 μm diameter
~200 M oligonucleotides


3’
polyT region
UMI
Spatial barcode spot-specific
ID from 1 to 1007
Amplification + sequencing handle
5’
Cleavage site

The method

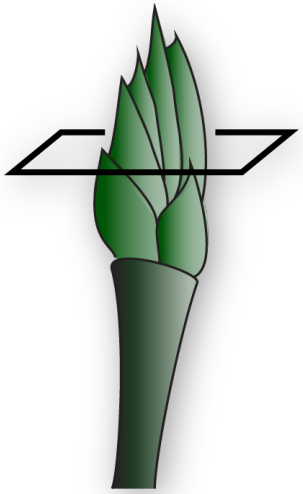
Cryosectioning



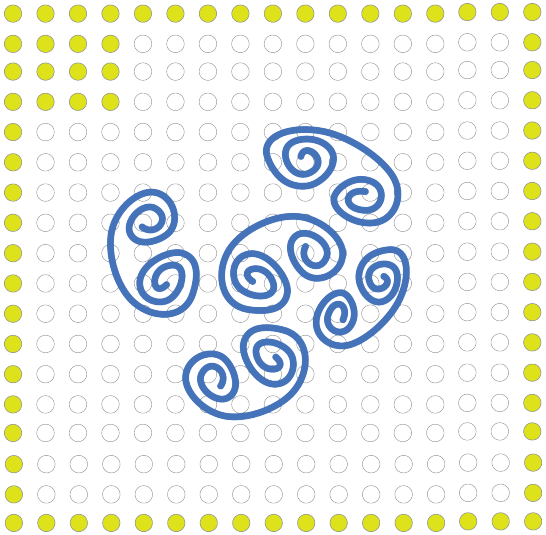


The method

Cryosectioning

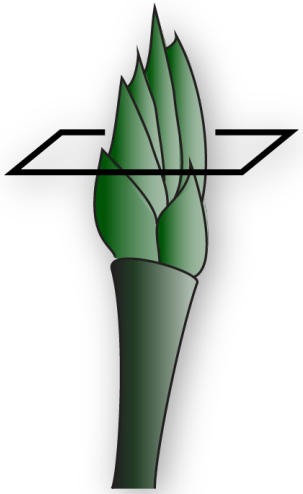


Staining

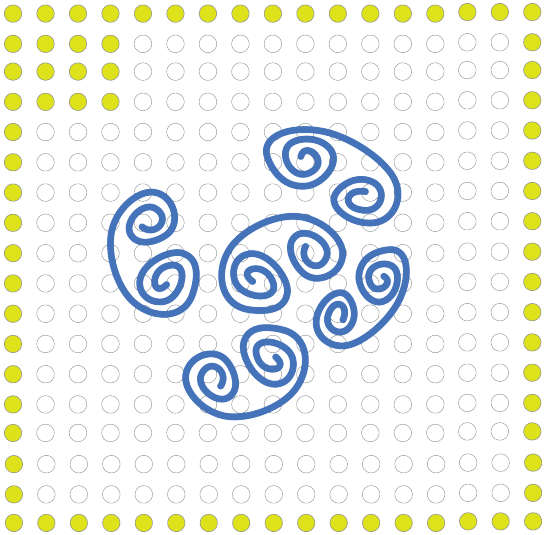


The method

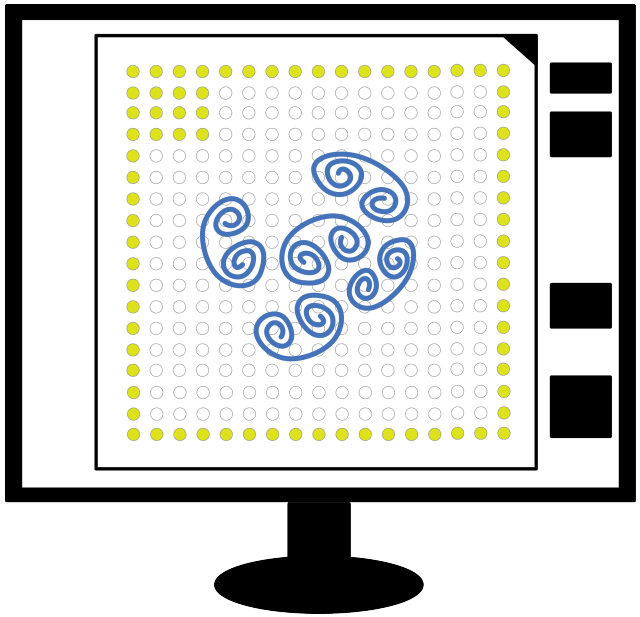
Cryosectioning



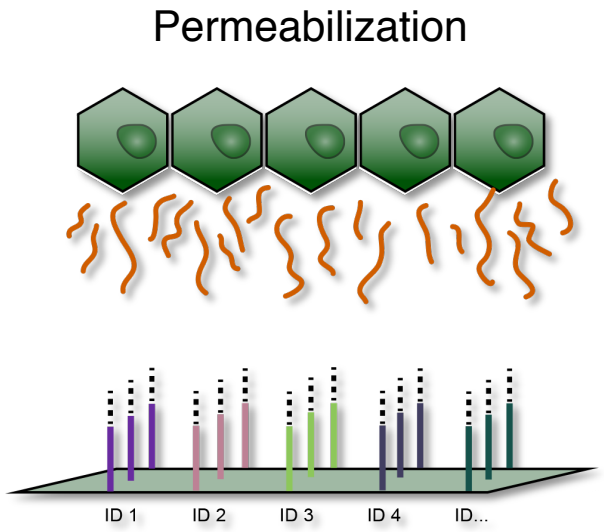
Staining



High resolution imaging

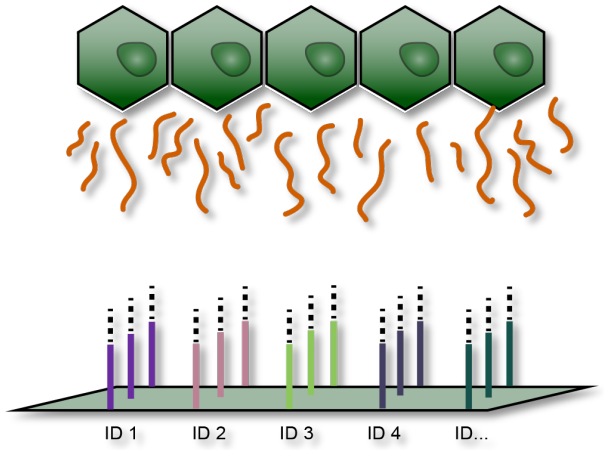


The method

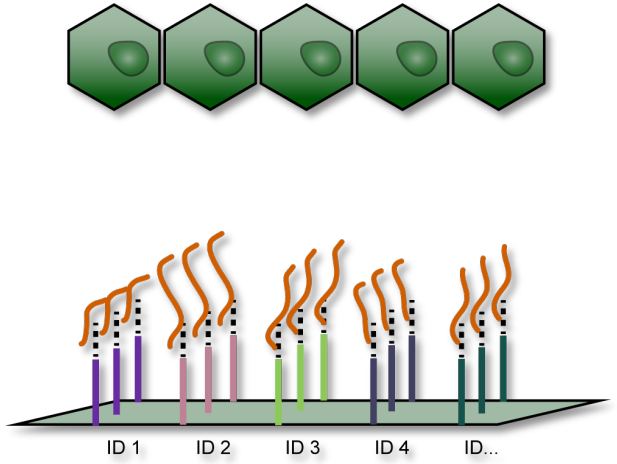


The method

Permeabilization

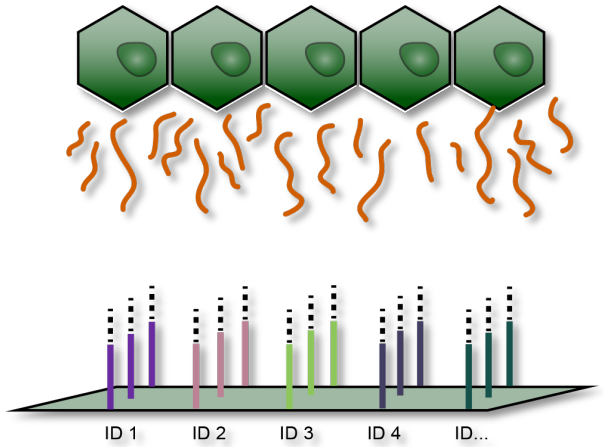


Poly-T capture of transcripts

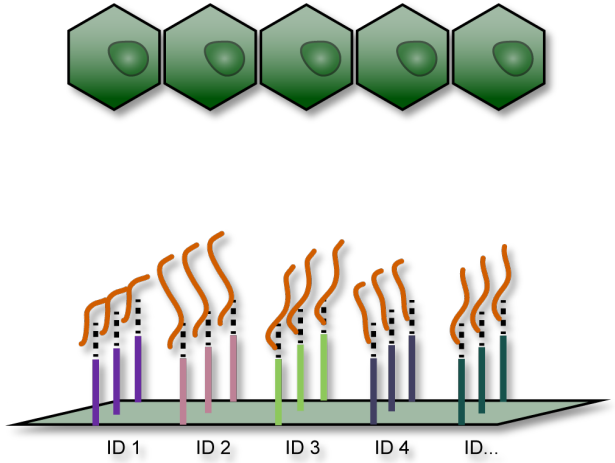


The method

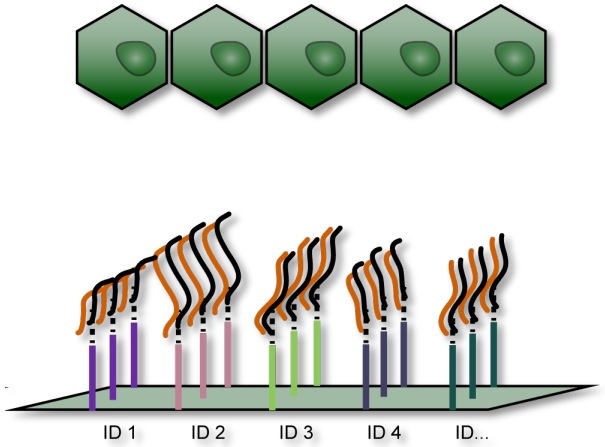
Permeabilization



Poly-T capture of transcripts

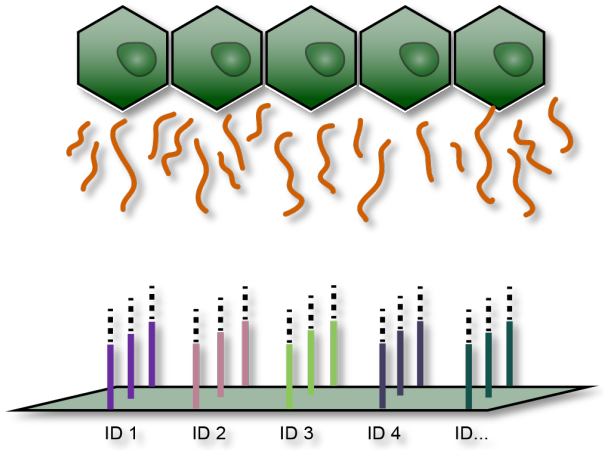


On surface cDNA synthesis

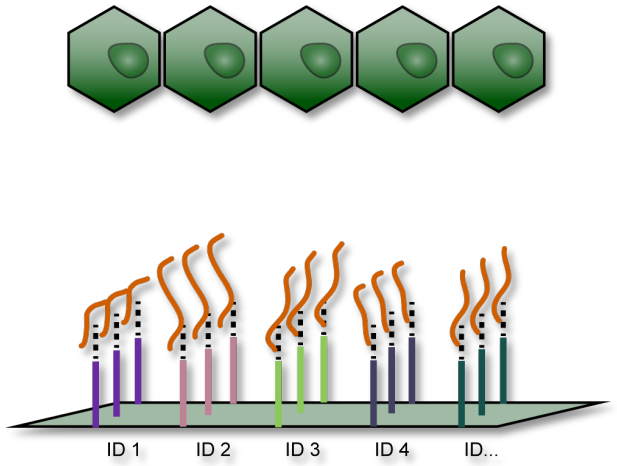


The method

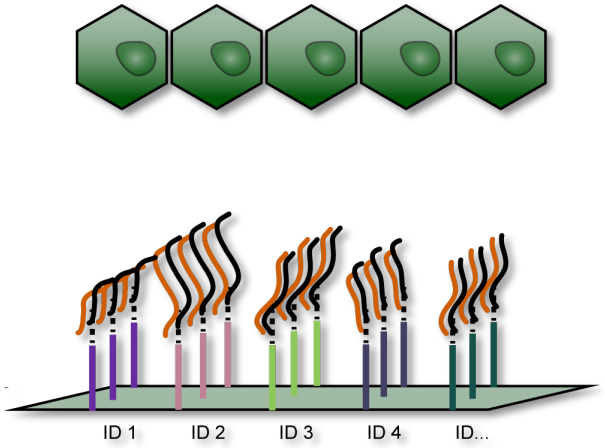
Permeabilization



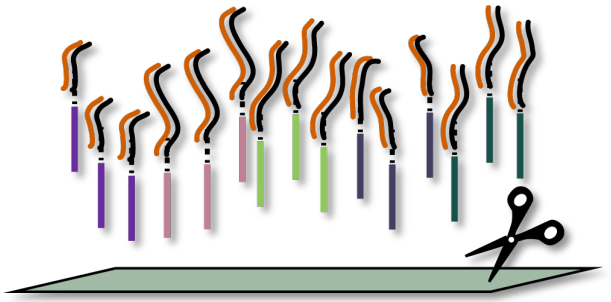
Poly-T capture of transcripts



On surface cDNA synthesis



Tissue removal and release

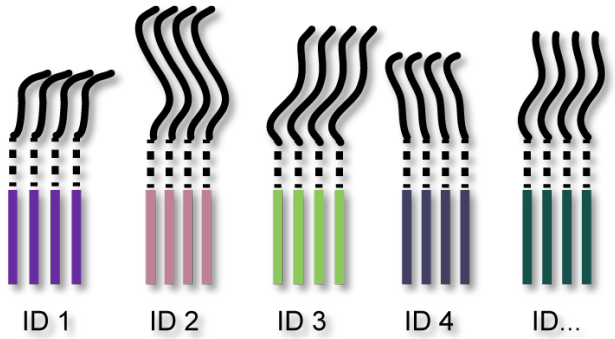


The method

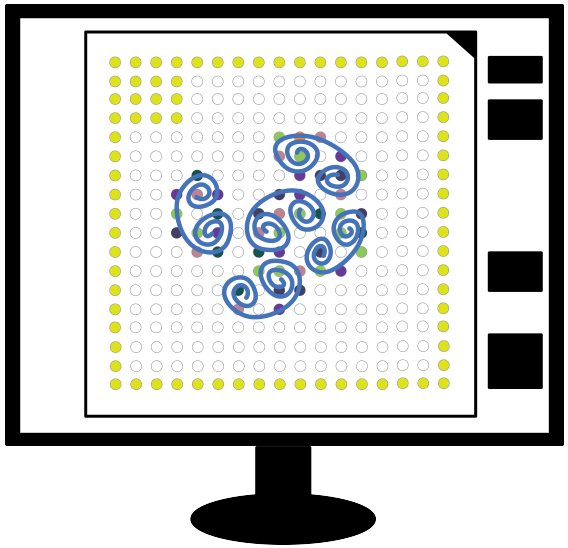
Illumina sequencing

GTACCTATTTAAGCGCGTATGCACCG
GCATGGCACGGCGCTCGCGTATGCAC
GTACCTATTTAAGCGCGTATGCACCG
TTAAGCGCGTATGCATTAGCCCACCG
GCCATATATATTCGCTATAATGCTGC
GCCACGGGCTACGATGCATTCGCTAT
GTACCTATTTAAGCGCGTATGCACCG
GCATGGCACGGCGCTCGCGTATGCAC
GTACCTATTTAAGCGCGTATGCACCG
TTAAGCGCGTATGCATTAGCCCACCG
GCCATATATATTCGCTATAATGCTGC
GCCACGGGCTACGATGCATTCGCTAT

Alignment and sorting of barcodes

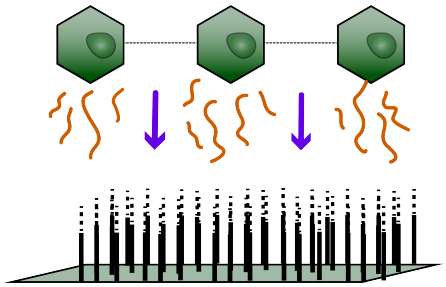


Alignment of image and barcoded transcripts



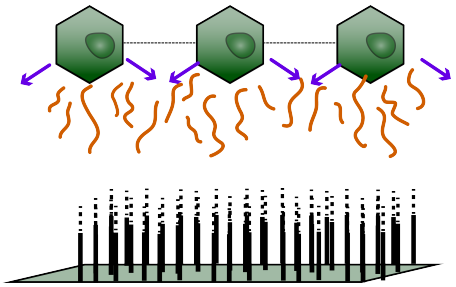
Proof of concept – later diffusion?

Vertical diffusion

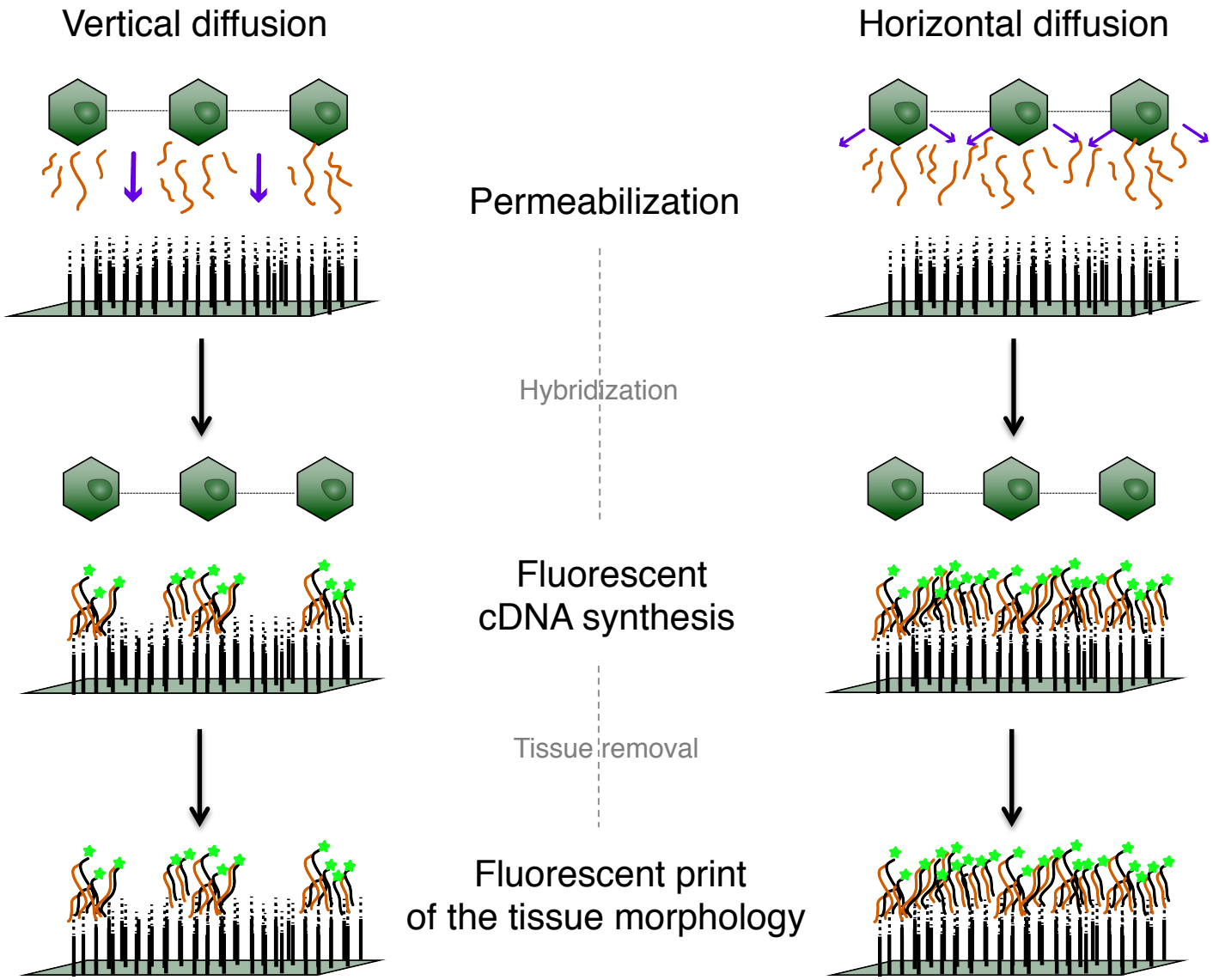


Permeabilization

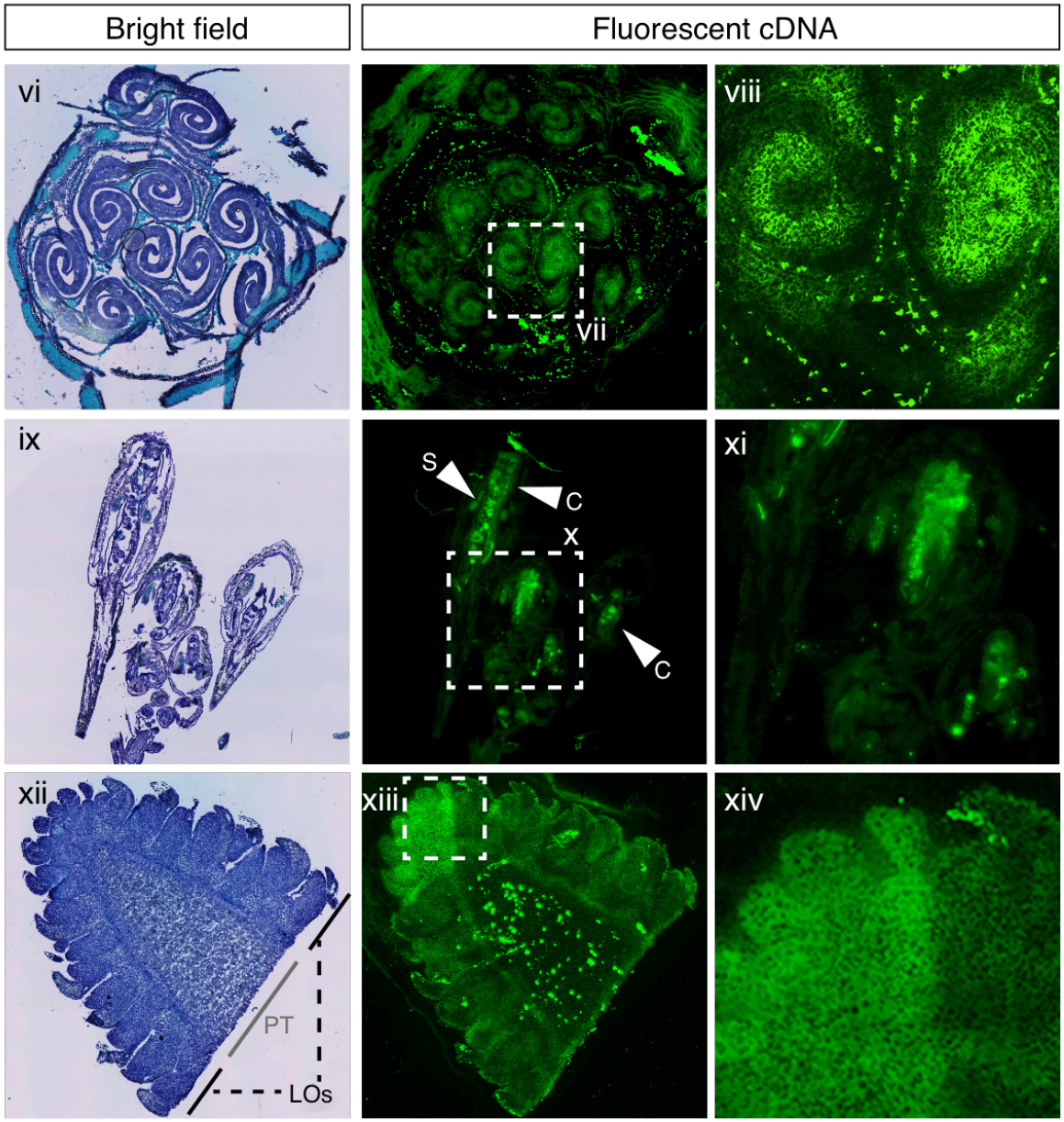
Horizontal diffusion



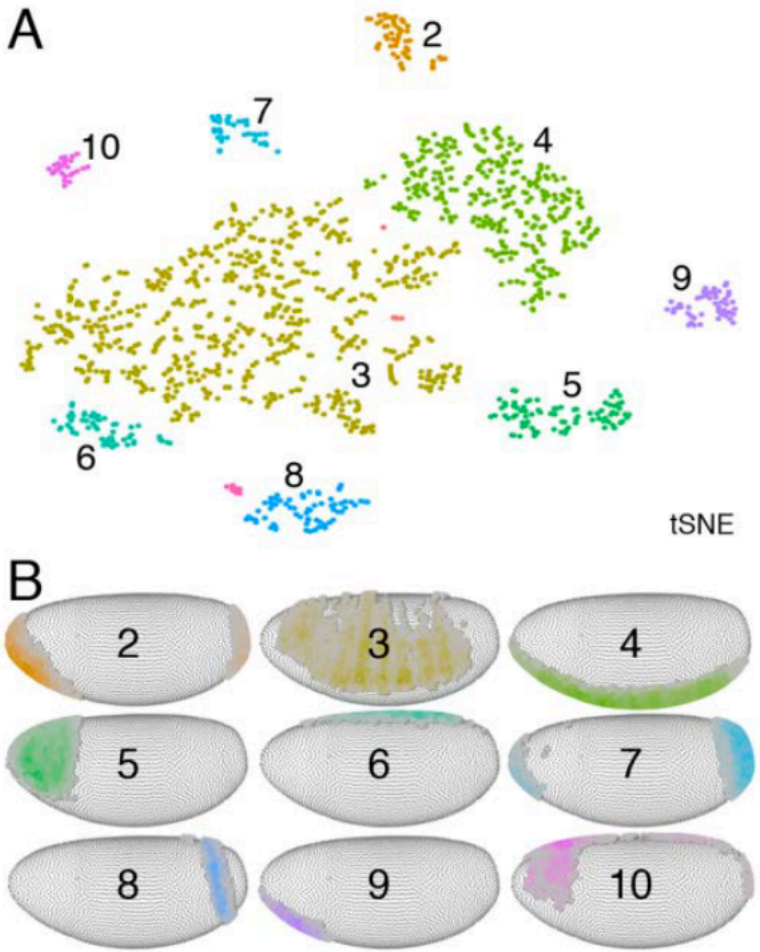
Proof of concept – later diffusion?



Proof of concept – no later diffusion

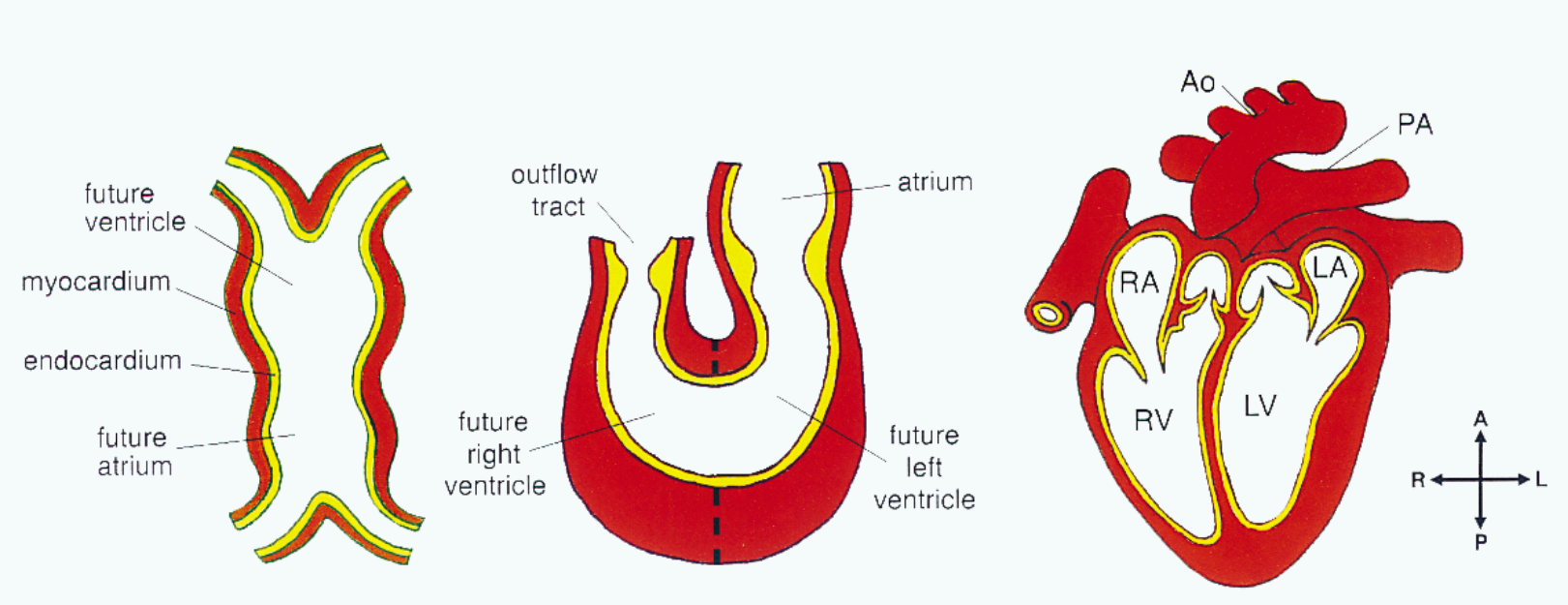


DistMap



Application on human fetal heart data

Main questions of the study



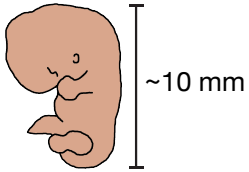
Cell, 1997

- Cardiomyocyte development
- Cardiac progenitor-/ stem cells
- Differentiation process

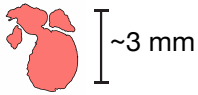
SciLifeLab

The approach

- single-cell RNA-seq (10X Chromium)
- Spatial Transcriptomics



Carnegie stage: 18
(44-48 days)

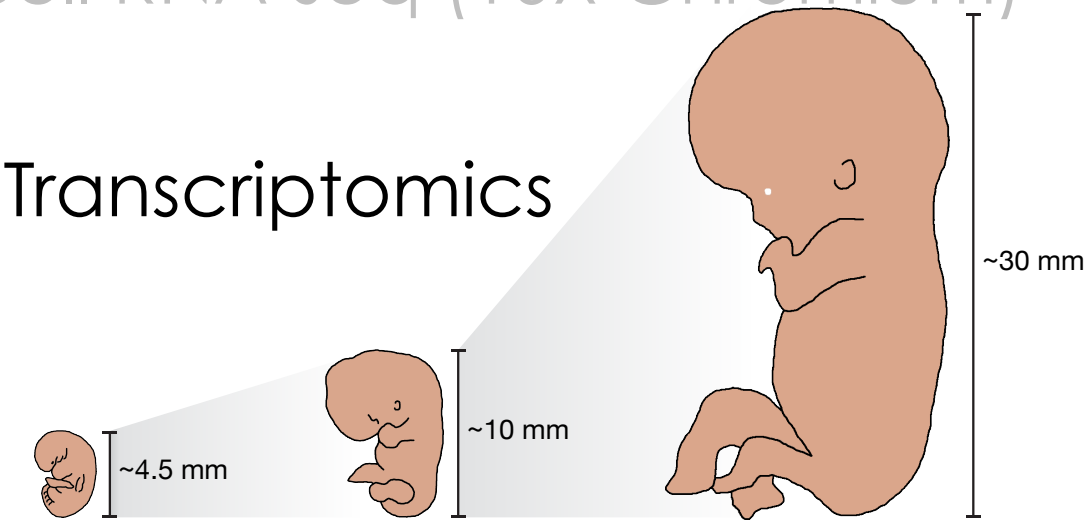


Clinical age: 6.5w
(~46 days)

SciLifeLab

The approach

- single-cell RNA-seq (10X Chromium)
- Spatial Transcriptomics



Carnegie stage: 13
(28-32 days)

Carnegie stage: 18
(44-48 days)

Carnegie stage: 23
(56-60 days)

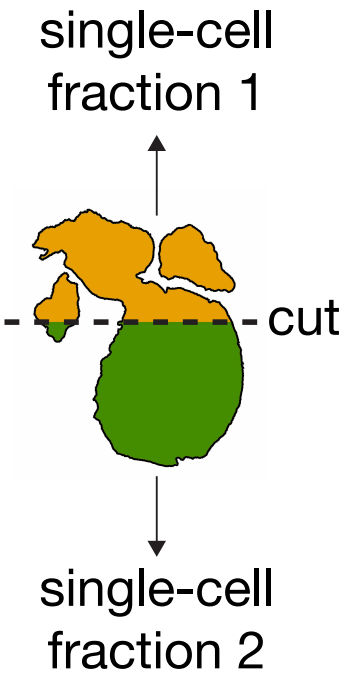
Clinical age: 4.5-5w
(~33 days)

Clinical age: 6.5w
(~46 days)

Clinical age: 9w
(~63 days)

SciLifeLab

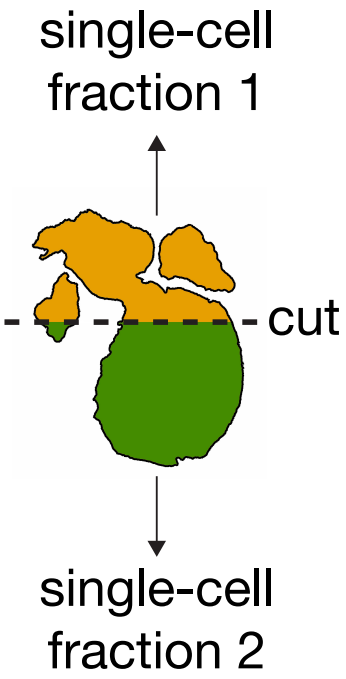
scRNA-seq dataset



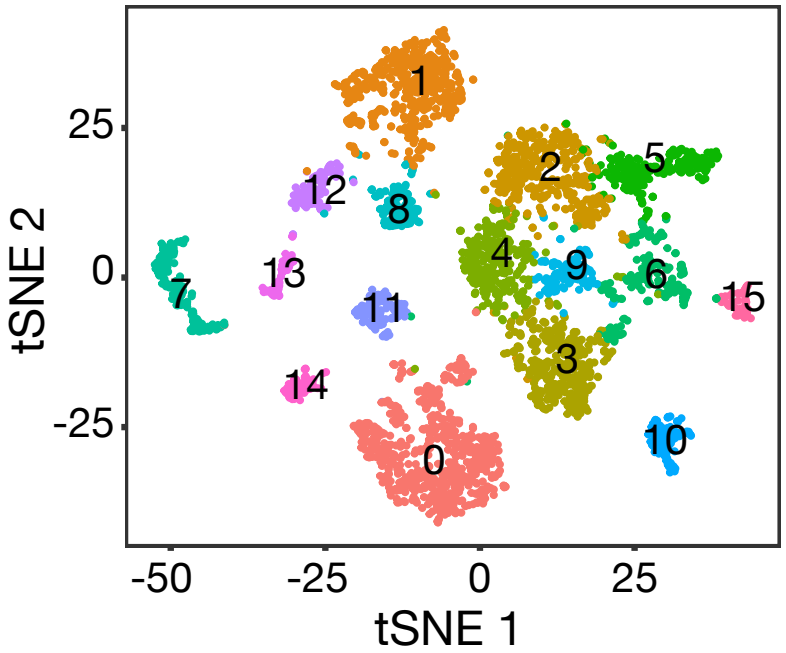
~4,000 cells
~3,000 genes per cell

SciLifeLab

scRNA-seq dataset



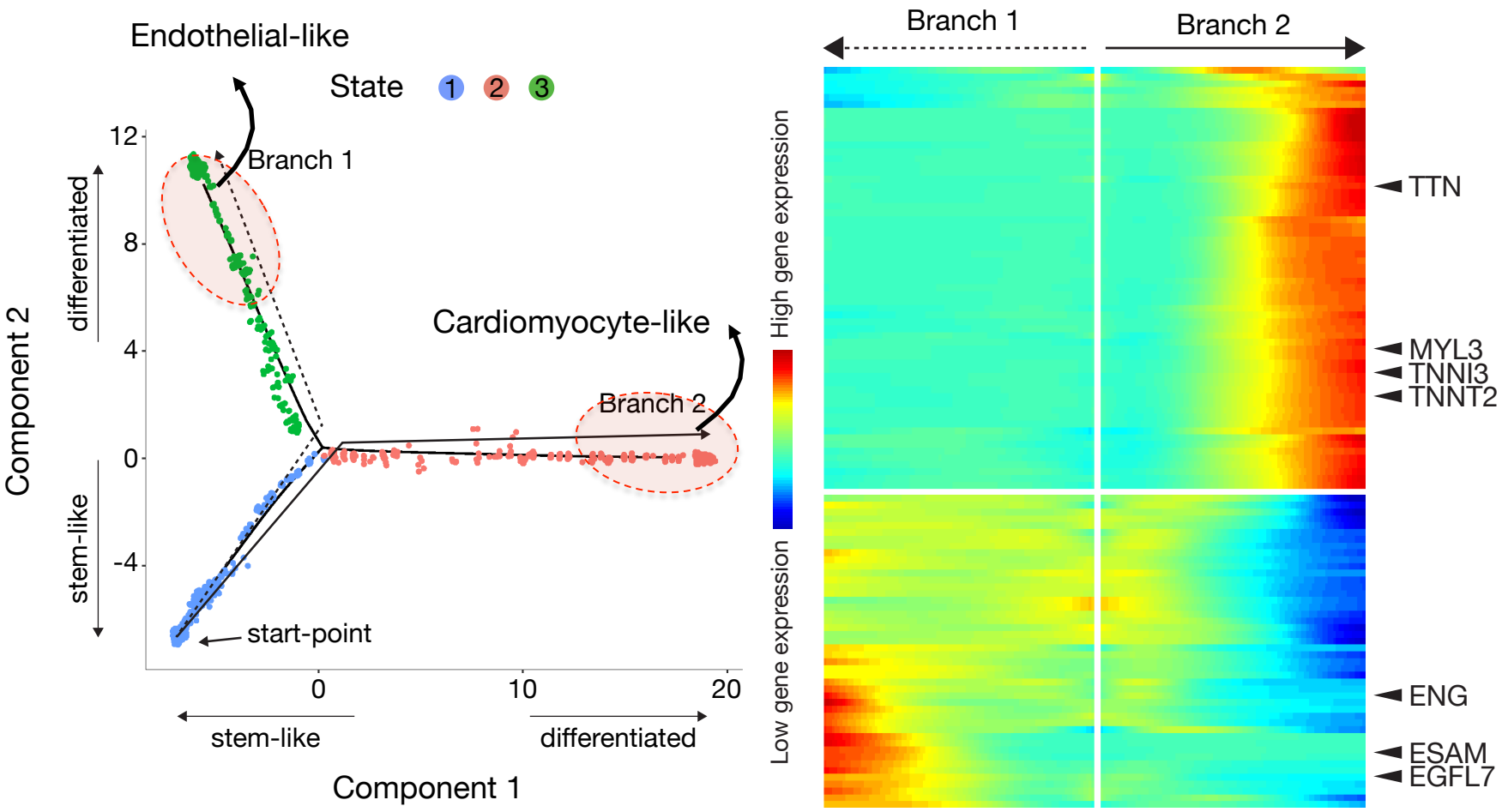
~4,000 cells
~3,000 genes per cell



0 Capillary endothelium
1 Ventricular cardiomyocytes
2 EPDCs
3 Fibroblast-like
4 Fibroblast-like
5 Smooth muscle cells /fibroblast-like
6 Fibroblast-like
7 Erythrocytes
8 Atrial cardiomyocytes
9 Fibroblast-like
10 EPDCs
11 Capillary endothelium /pericytes/adventitia
12 Cardiomyocytes
13 Erythrocytes
14 Immune cells
15 Cells related to cardiac neural crest

SciLifeLab

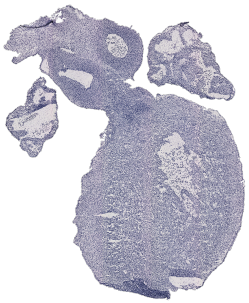
Fates of human fetal heart cells

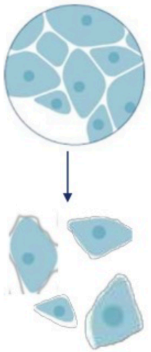


SciLifeLab

Spatial gene expression

Single cell sequencing
Embryonic heart 6.5-7w

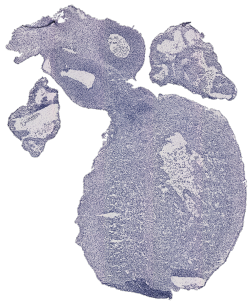


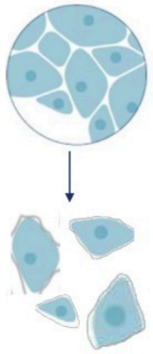


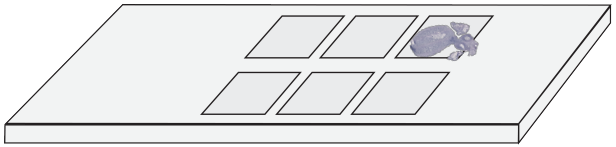
SciLifeLab

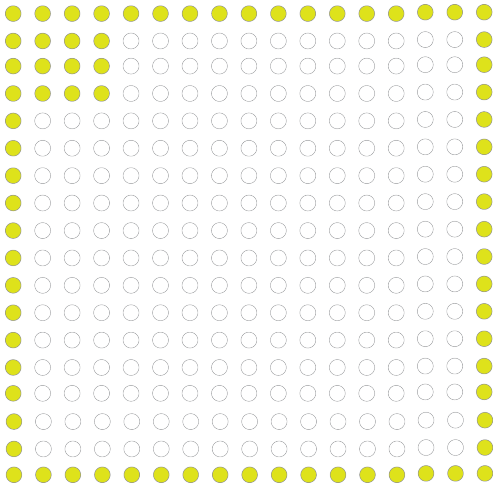
Spatial gene expression

Single cell sequencing
Embryonic heart 6.5-7w





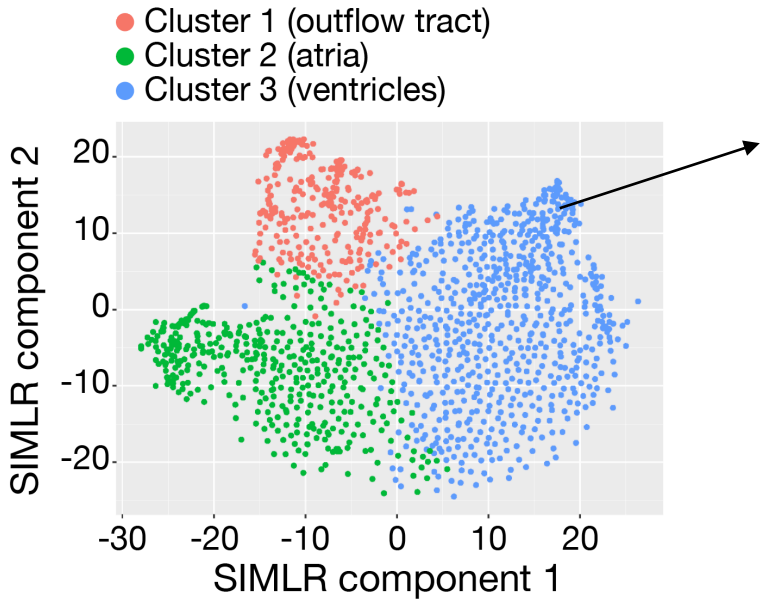




Spatial Transcriptomics
Embryonic heart 6.5w

SciLifeLab

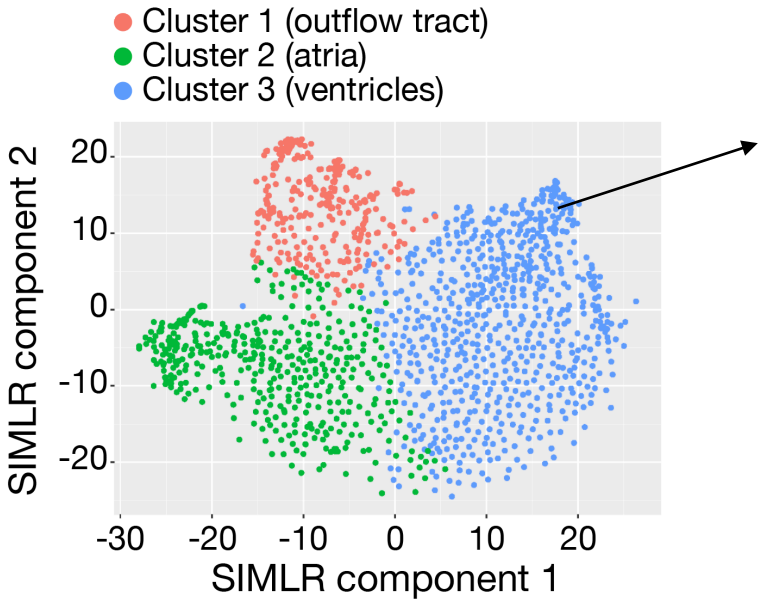
Spatial gene expression



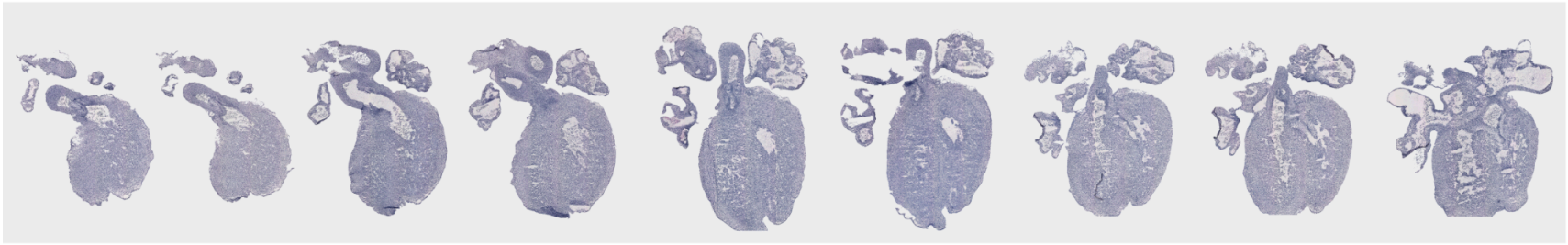
Each dot represents one "spot" i.e. microdissection

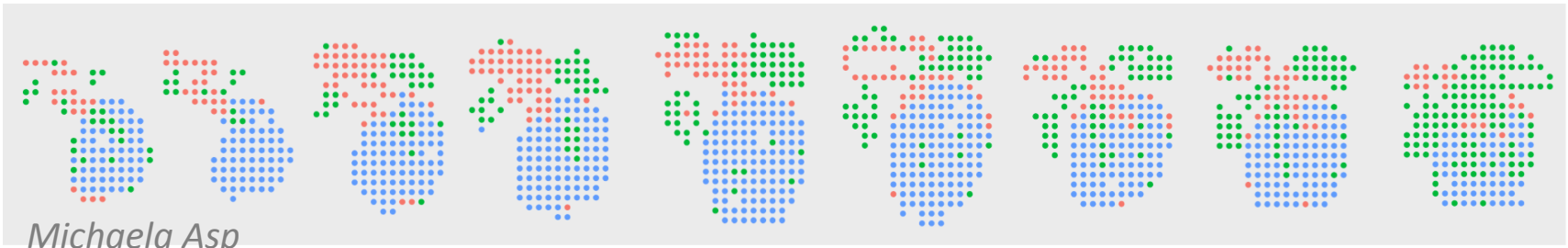
SciLifeLab

Spatial gene expression



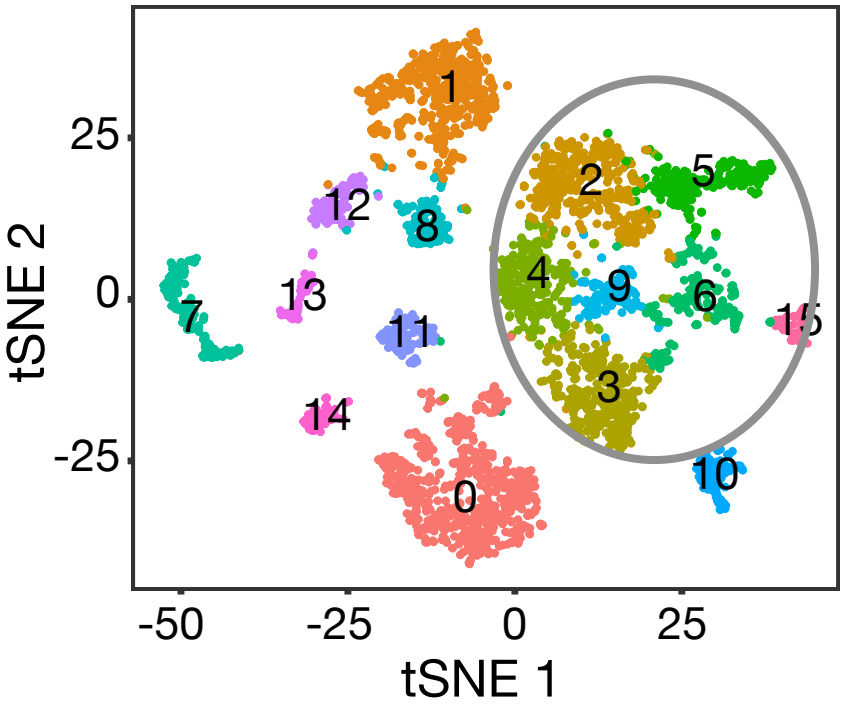
Each dot represents one "spot" i.e. microdissection

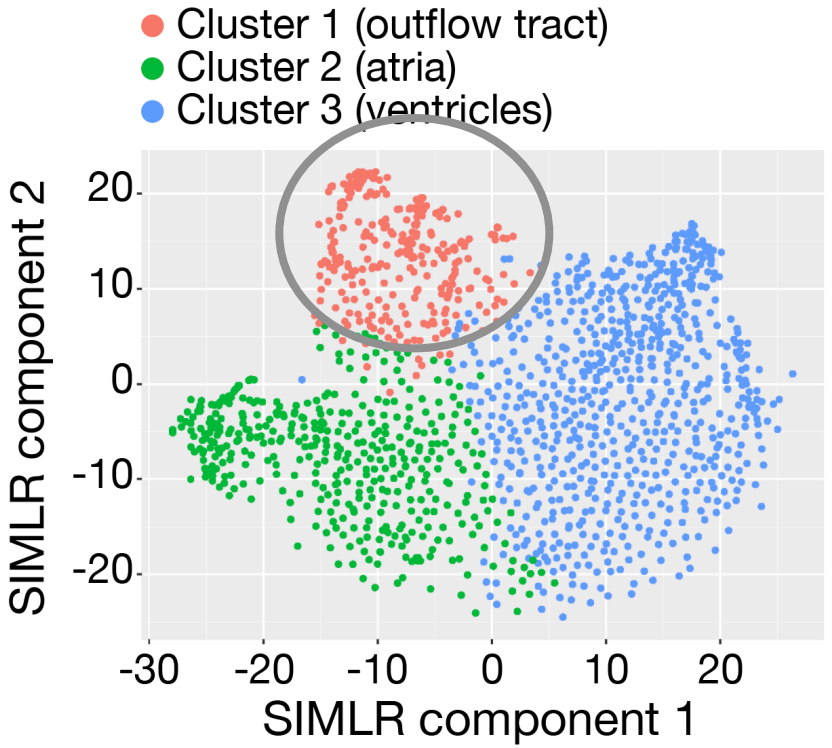




Michaela Asp

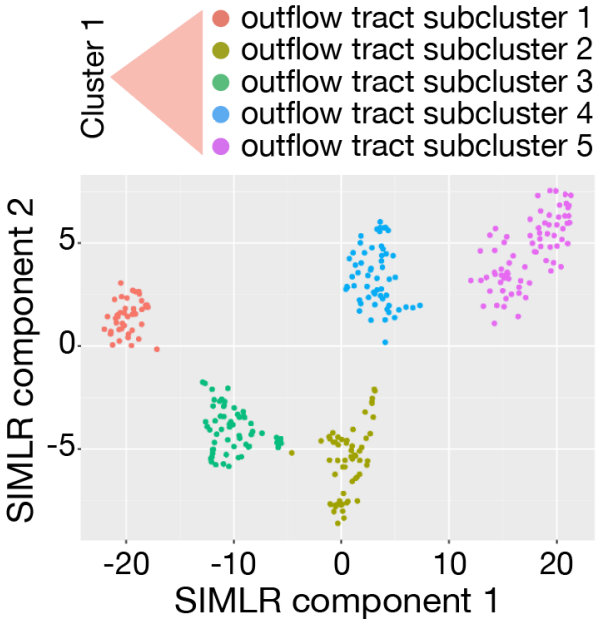
Subclustering of spatial transcriptomics data





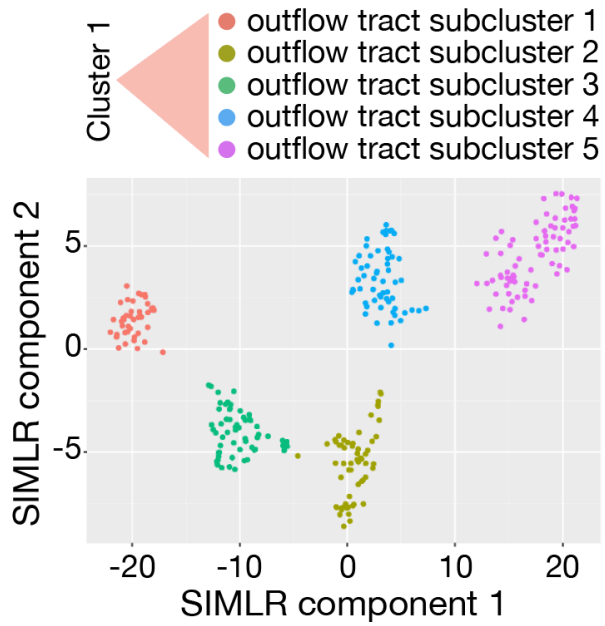
SciLifeLab

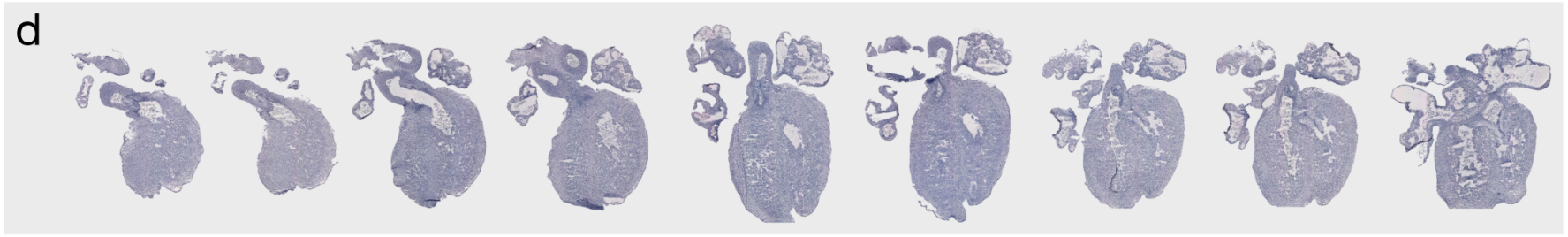
Spatial gene expression – subclustering of outflow tract

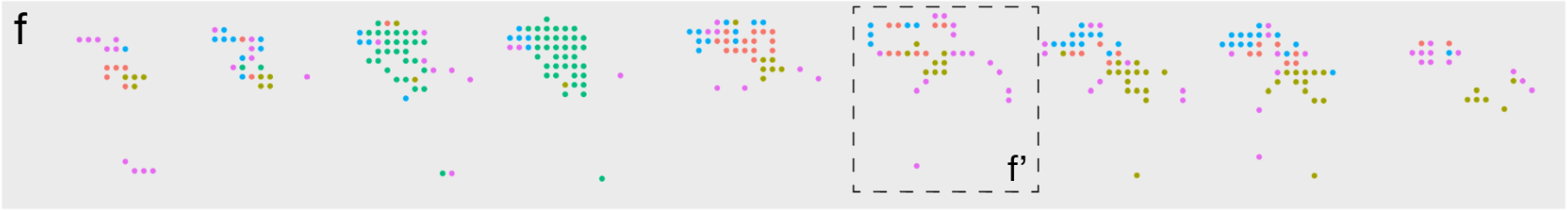


SciLifeLab

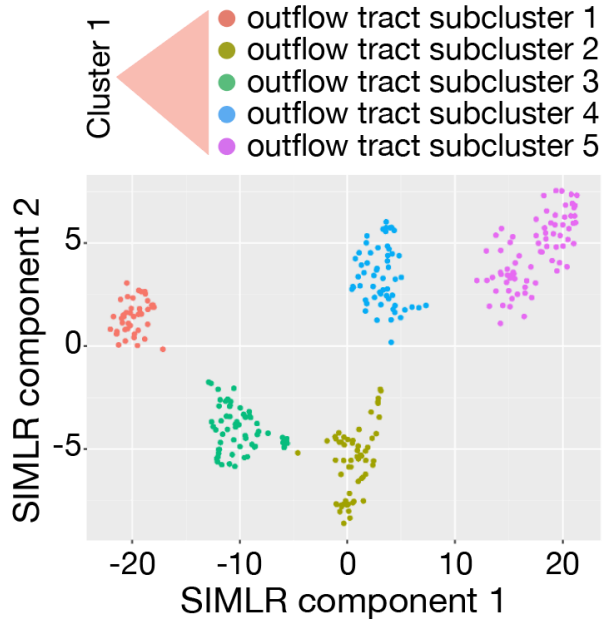
Spatial gene expression – subclustering of outflow tract

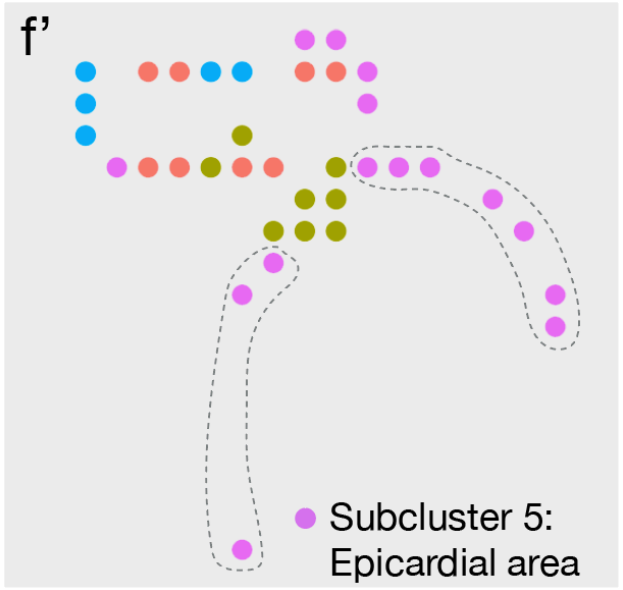


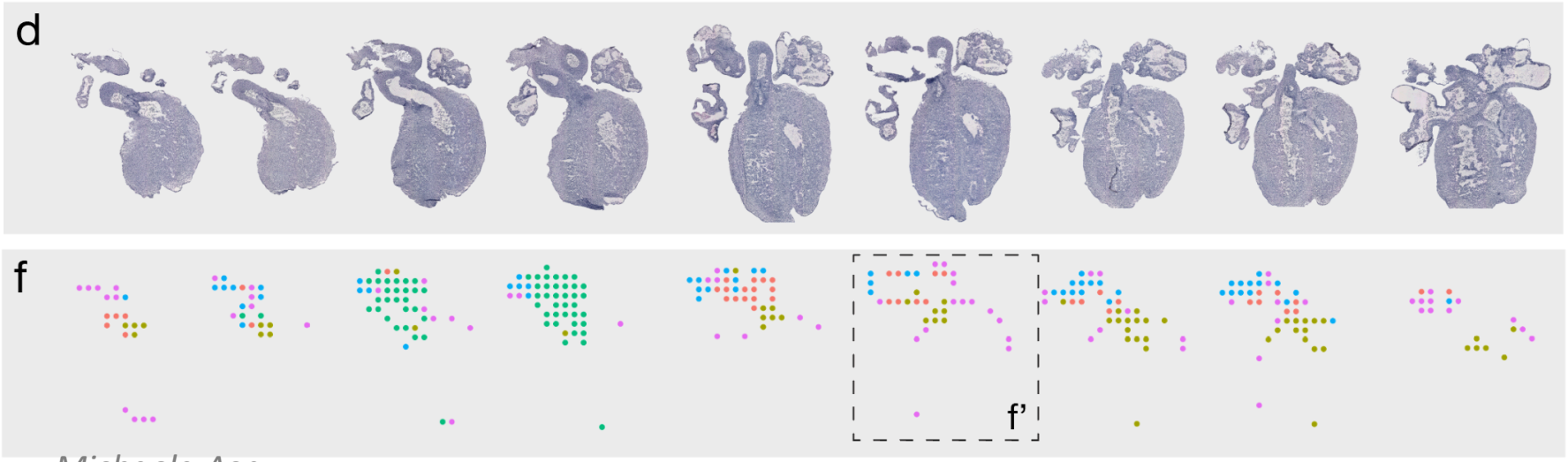




Spatial gene expression – subclustering of outflow tract

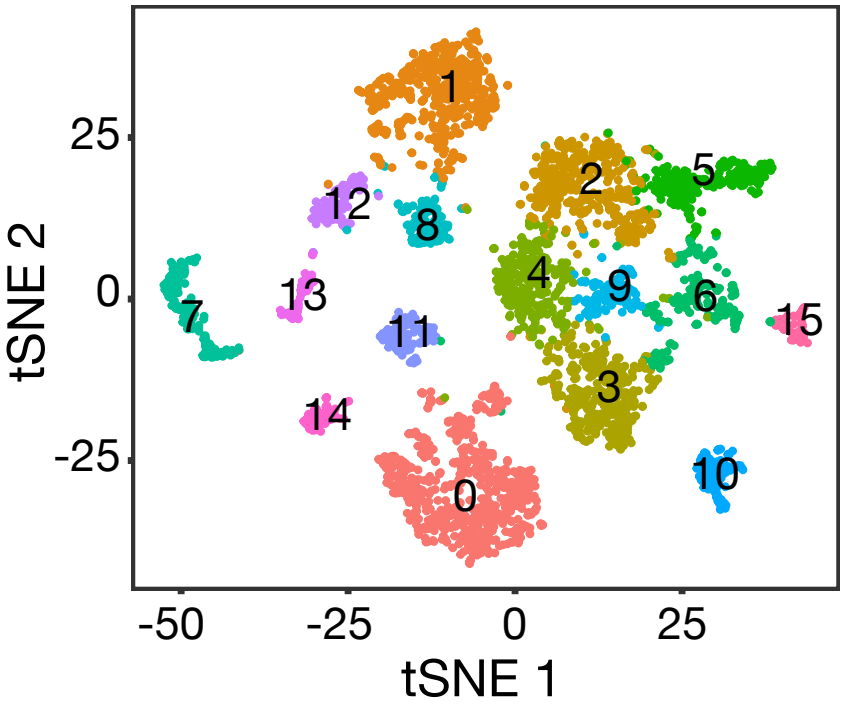




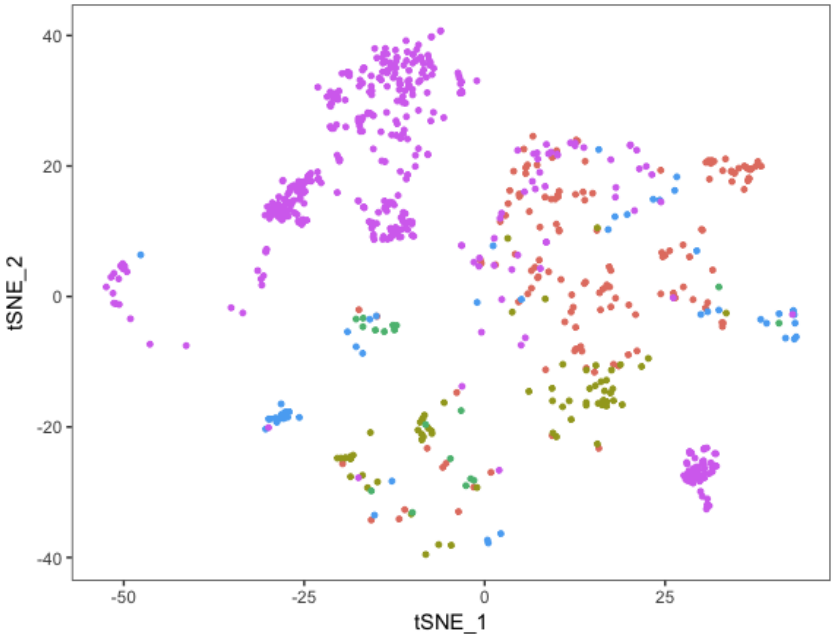


Michaela Asp

Mapping of single cells on spatial subclusters



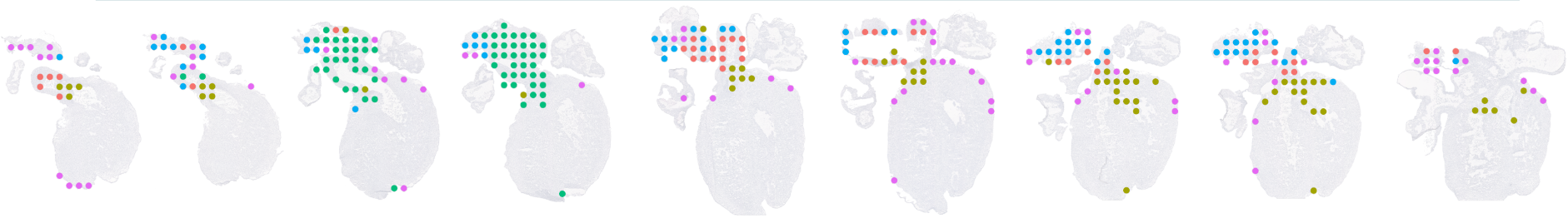
Mapping all single-cells to subclusters of the OFT



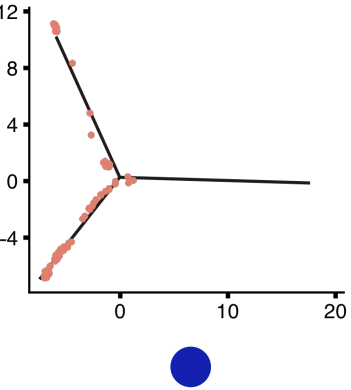
- Mapping uniquely to subcluster 1
- Mapping uniquely to subcluster 2
- Mapping uniquely to subcluster 3
- Mapping uniquely to subcluster 4
- Mapping uniquely to subcluster 5

SciLifeLab

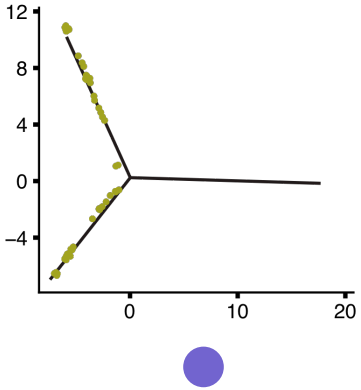
Spatial fate maps



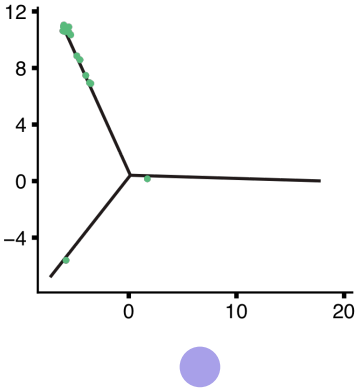
Cells uniquely mapping to subcluster 1



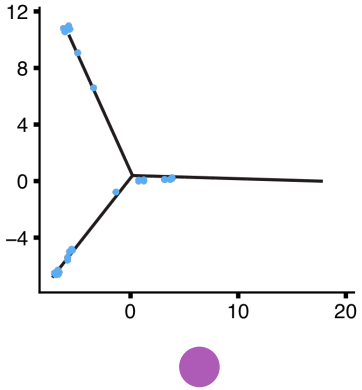
Cells uniquely mapping to subcluster 2



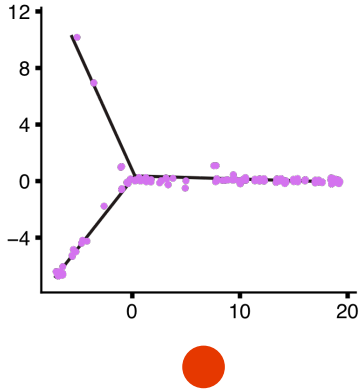
Cells uniquely mapping to subcluster 3



Cells uniquely mapping to subcluster 4

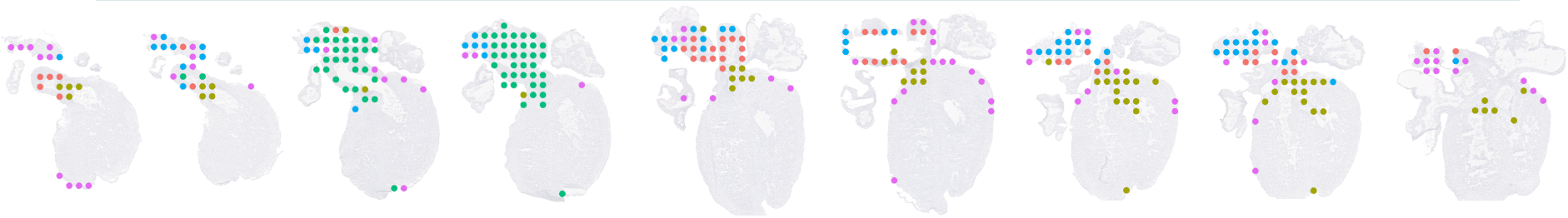


Cells uniquely mapping to subcluster 5

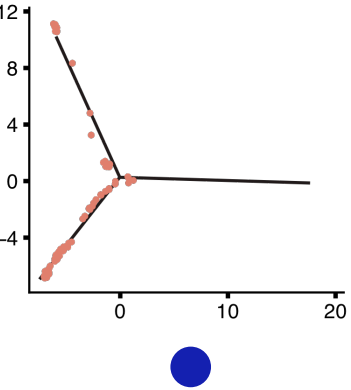


SciLifeLab

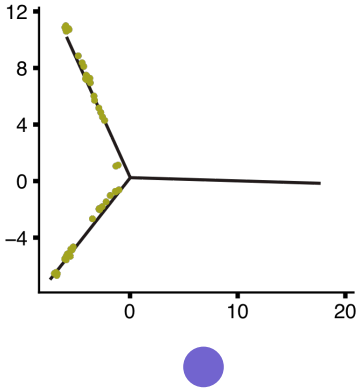
Spatial fate maps



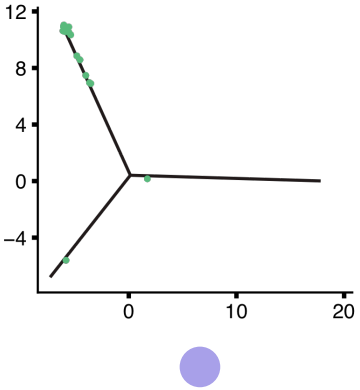
Cells uniquely mapping to subcluster 1



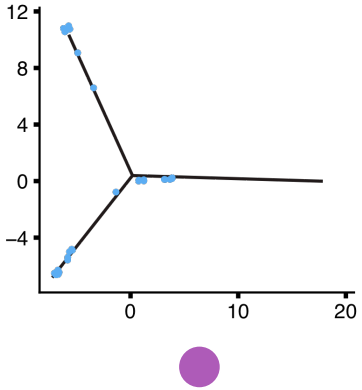
Cells uniquely mapping to subcluster 2



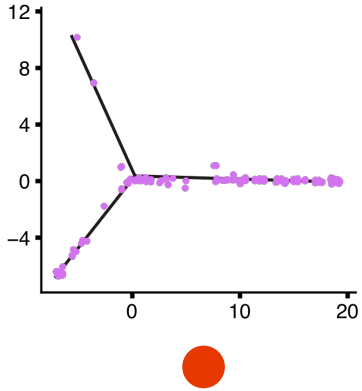
Cells uniquely mapping to subcluster 3

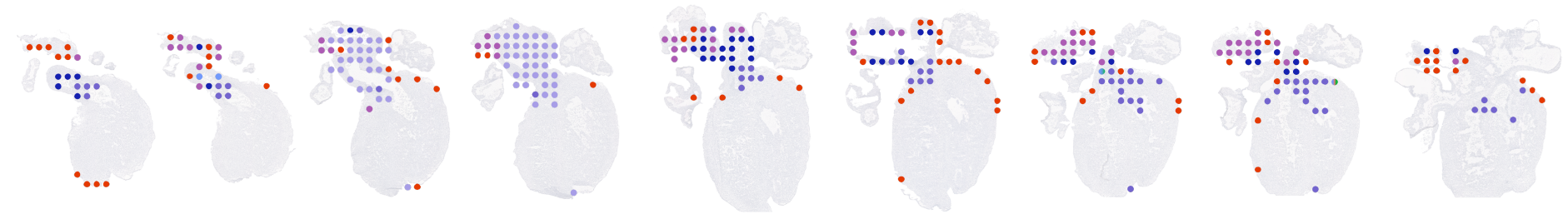


Cells uniquely mapping to subcluster 4



Cells uniquely mapping to subcluster 5





SciLifeLab

Conclusions

- Model organisms → computational methods can be useful
- Non model organisms → a spatial transcriptomics approach is more straightforward
- Overall, the best case scenario is a spatial transcriptomics approach with single-cell resolution

Acknowledgments

Michaela Asp

Joakim Lundeberg

SciLifeLab

Johan Reimegård

NBIS
NATIONAL BIOINFORMATICS INFRASTRUCTURE SWEDEN

Christer Sylvén

Eva Wärdell

Matthias Corbascio

KAROLINSKA INSTITUTET ANNO 1810
Karolinska Institutet