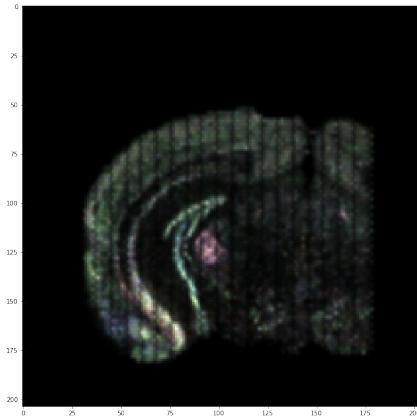
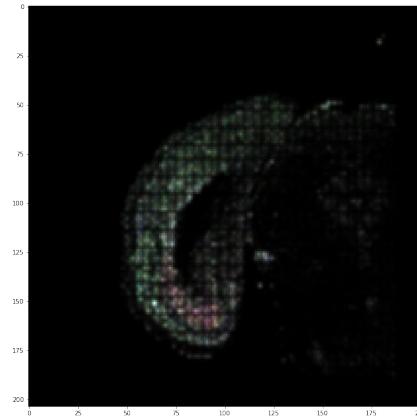
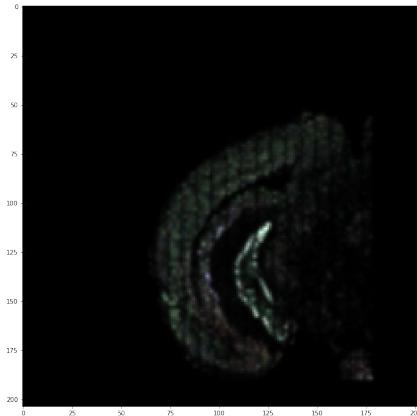
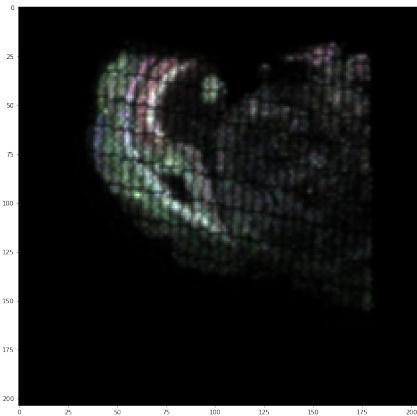


Spatially map single cell RNA-seq to spatial coordinates

without prior knowledge of cell locations for scRNASeq (BAR-seq2)



Motivation

Growing better organoids to form organ tissues for organ repair and replacement

Mapping single cells to spatial structures improves interpretation of spatial developmental biology

For example, key kidney paper on SIX2 gene in progenitor kidneys involved spatial knowledge:

Six2 activity is required for maintaining the mesenchymal progenitor population in an undifferentiated state by opposing the inductive signals emanating from the ureteric bud.

Six2-expressing cells give rise to all cell types of the main body of the nephron during all stages of nephrogenesis.



[EMBO J.](#) 2006 Nov 1; 25(21): 5214–5228.

Published online 2006 Oct 12. doi: [10.1038/sj.emboj.7601381](https://doi.org/10.1038/sj.emboj.7601381)

Six2 is required for suppression of nephrogenesis and progenitor renewal in the developing kidney

[Michelle Self](#),¹ [Oleg V Lagutin](#),¹ [Beth Bowling](#),¹ [Jaime Hendrix](#),¹ [Yi Cai](#),² [Gregory R Dressler](#),² and [Guillermo Oliver](#),^{1,a}

[Author information](#) [Article notes](#) [Copyright and license information](#) [Disclaimer](#)

PMCID: PMC1630416

PMID: [17036046](#)

Cell Stem Cell

Volume 3, Issue 2, 7 August 2008, Pages 169-181



Article

Six2 Defines and Regulates a Multipotent Self-Renewing Nephron Progenitor Population throughout Mammalian Kidney Development

Akiyo Kobayashi¹, M. Todd Valerius¹, Joshua W. Mugford¹, Thomas J. Carroll^{1,3}, Michelle Self², Guillermo Oliver², Andrew P. McMahon^{1,2,3}

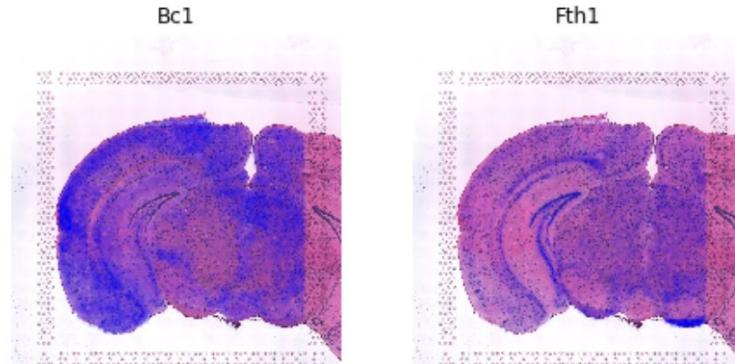
Datasets

Spatial Transcriptomic data from the Allen Brain Institute is used from:

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

This is purposed for the **BAR-seq2** data of the **Competition from the Allen Brain Institute**:

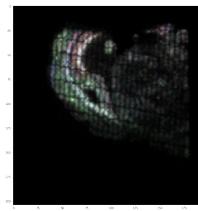
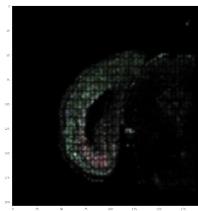
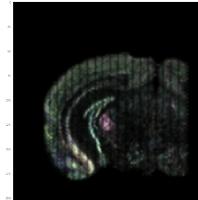
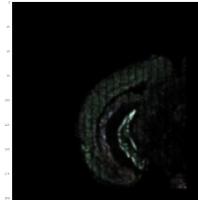
https://labshare.cshl.edu/shares/gillislab/resource/CellTypeMapping_2022/



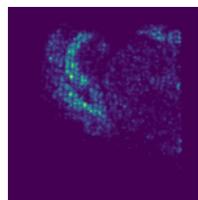
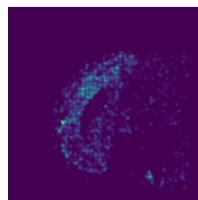
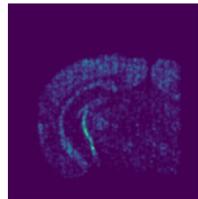
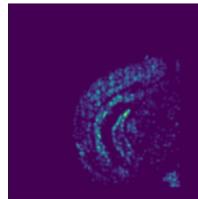
<u>Name</u>	<u>Last modified</u>	<u>Size</u>	<u>Description</u>
Parent Directory		-	
barseq2.hdf5	2022-04-16 11:27	29M	
barseq2.rds	2022-04-16 11:27	27M	
readme.txt	2022-04-16 11:27	1.2K	

Training set

Predicted image tissue histology pixels matched to observed image tissue histology pixels

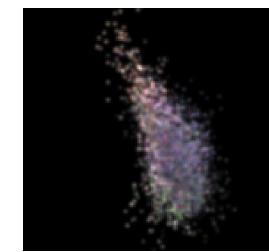
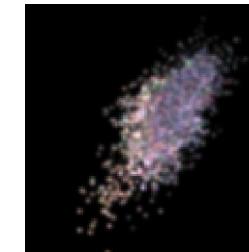
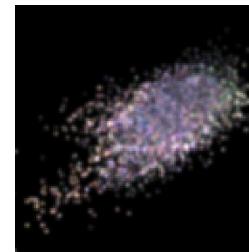


Predicted gene expression of SLC24A2 on spatial histology matched to observed spatial gene expression

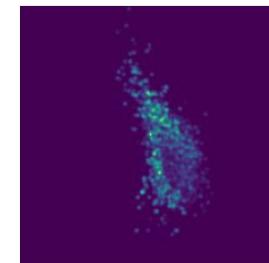
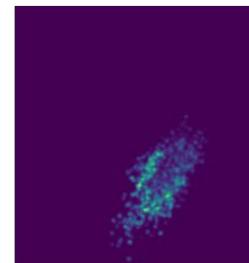
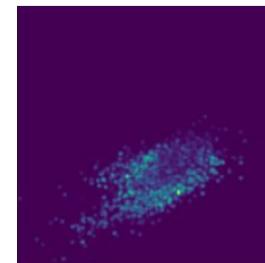


Test set

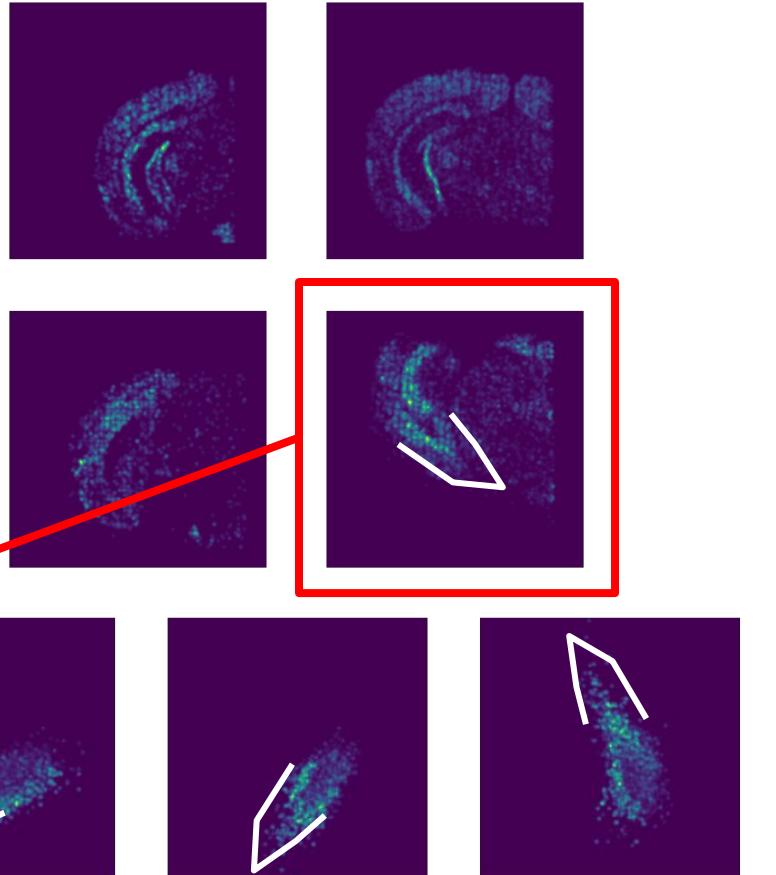
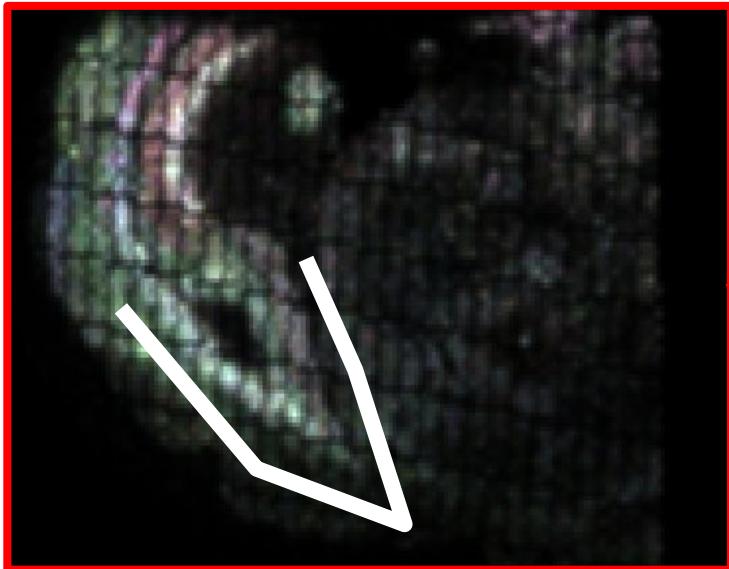
Predicted image tissue histology based on scRNAseq + model

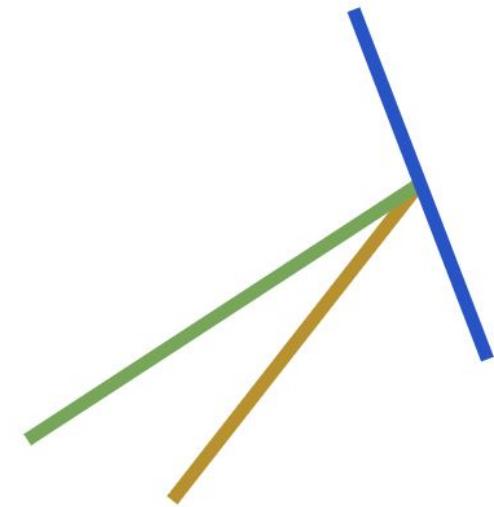
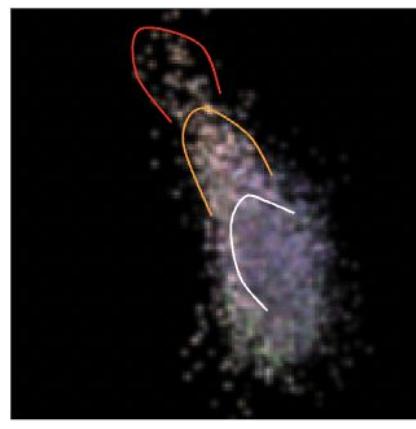
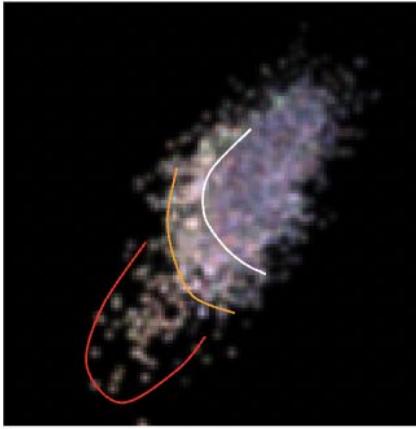
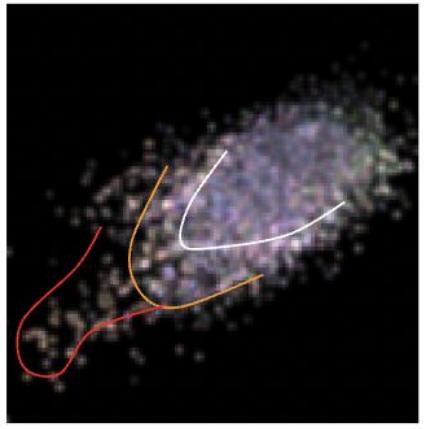
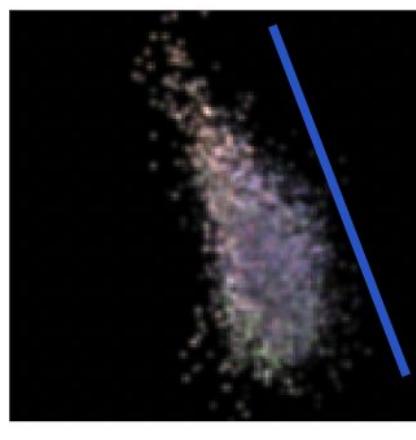
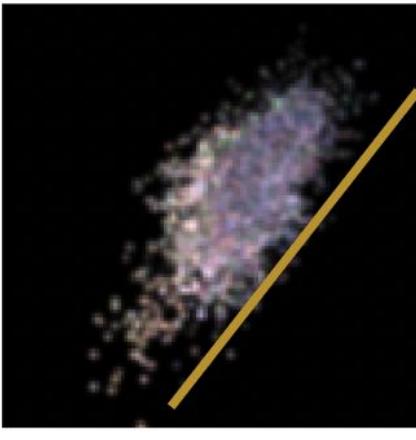
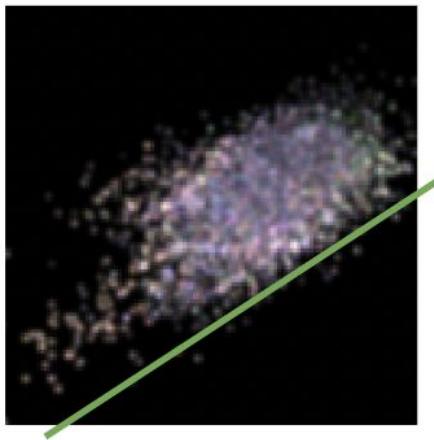


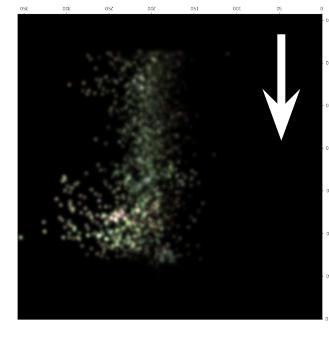
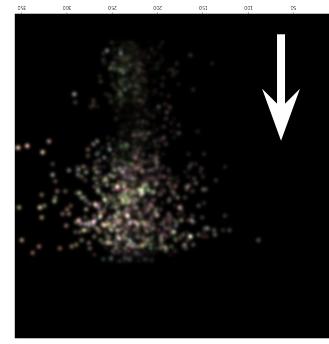
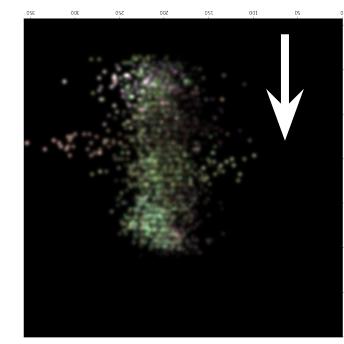
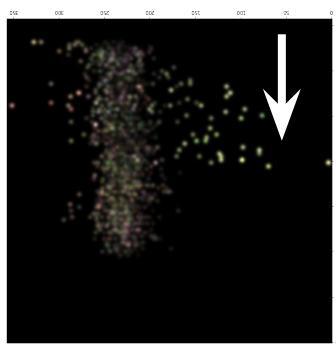
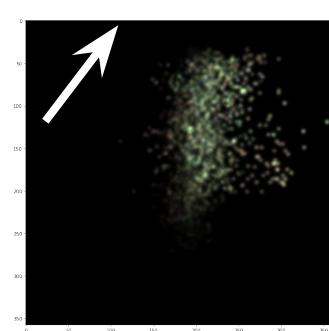
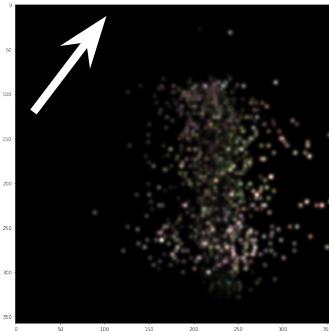
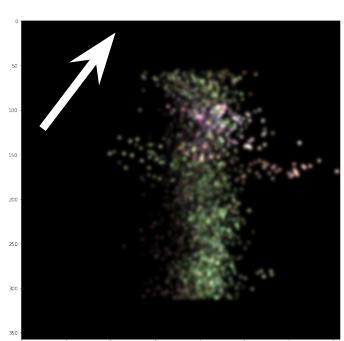
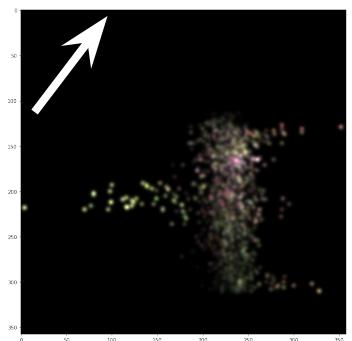
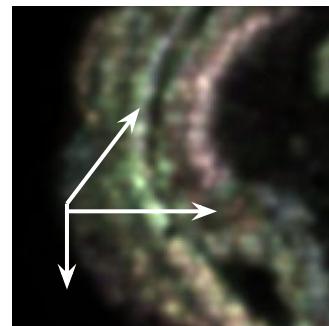
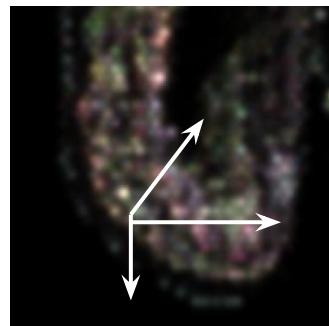
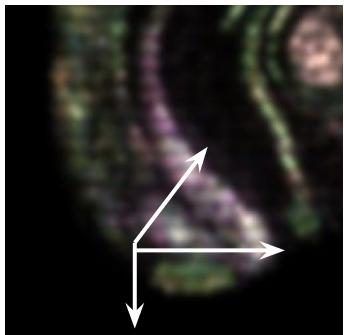
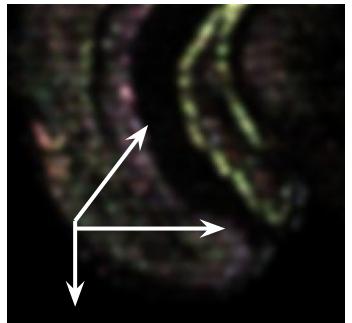
Predicted gene expression for SLC24A2 based on scRNAseq + model



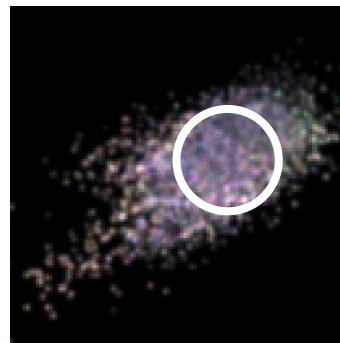
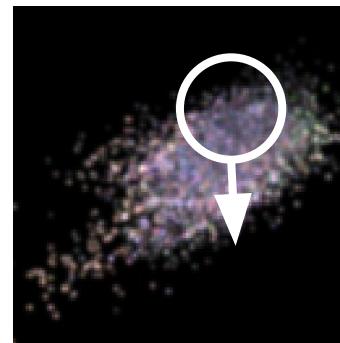
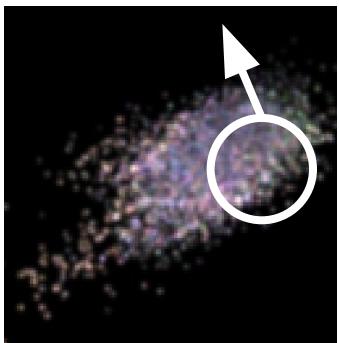
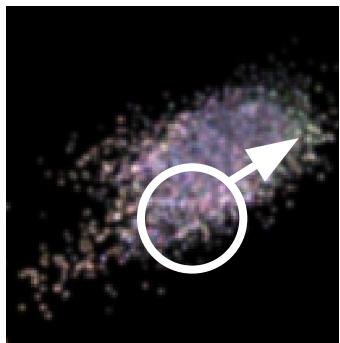
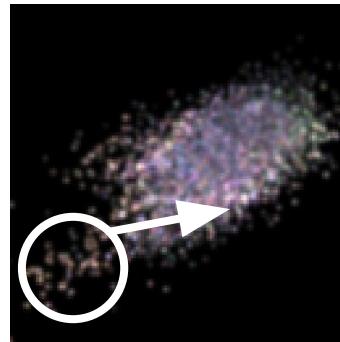
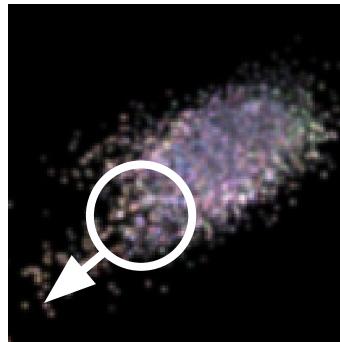
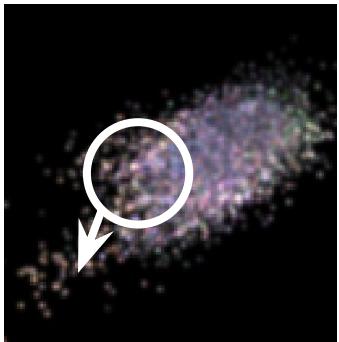
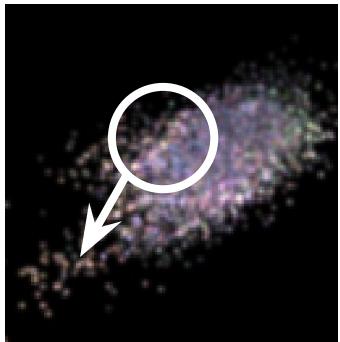
The location seems to visibly cover the **hippocampal / entorhinal cortex** region with relevant genes



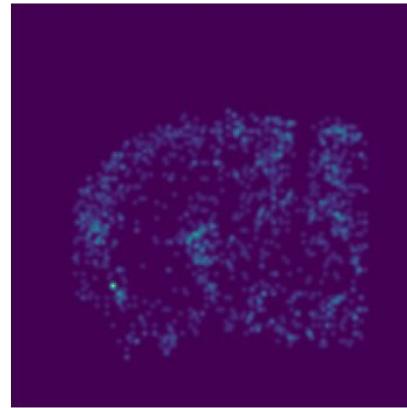
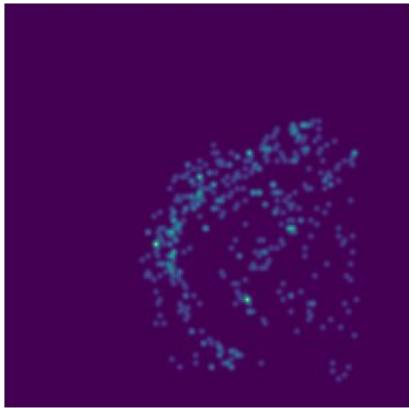




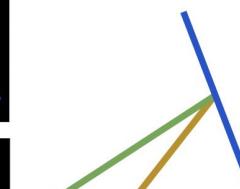
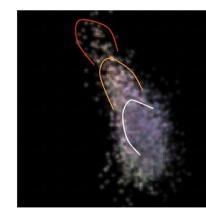
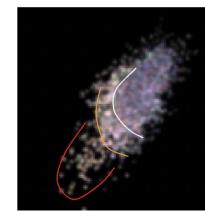
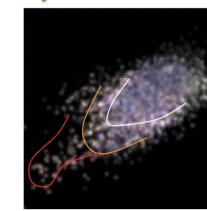
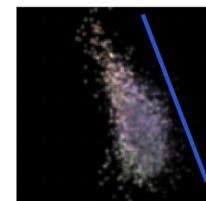
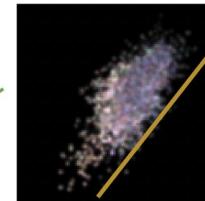
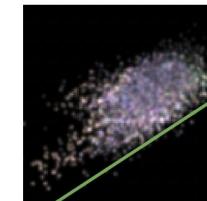
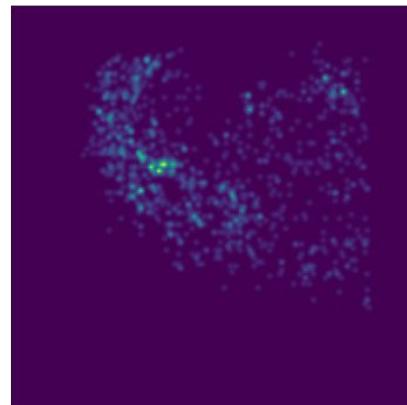
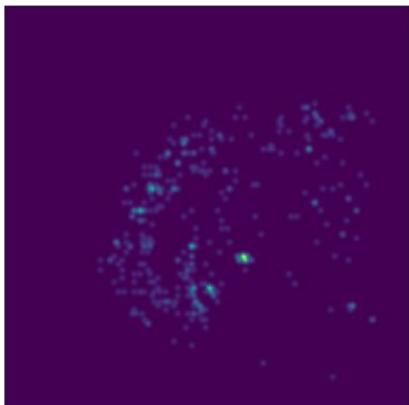
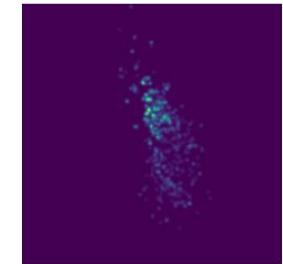
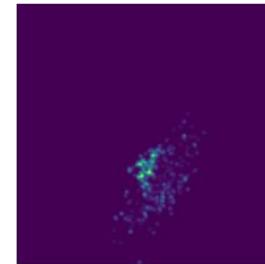
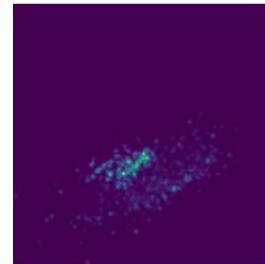
Next set of slides show the trajectory of gene expression based on location for different genes



Real GEX Olfm3

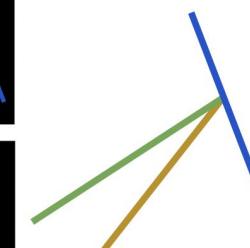
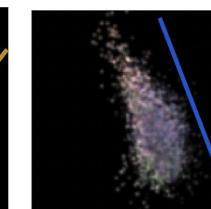
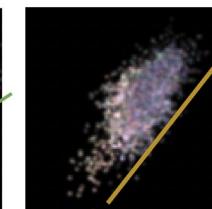
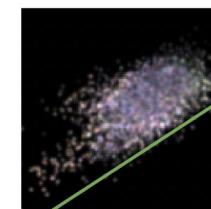
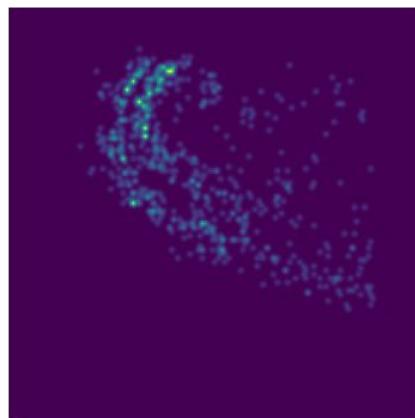
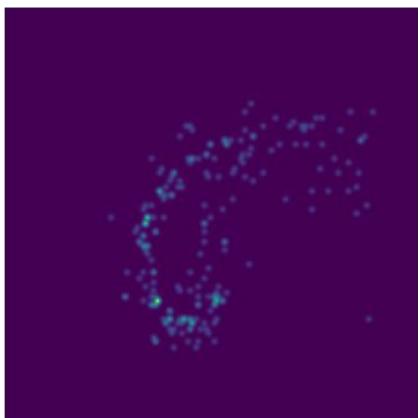
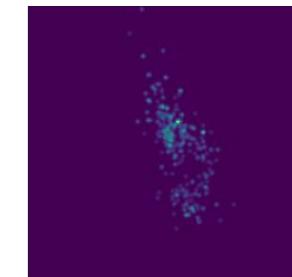
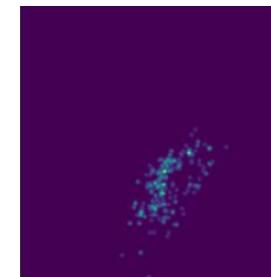
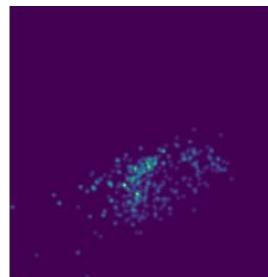
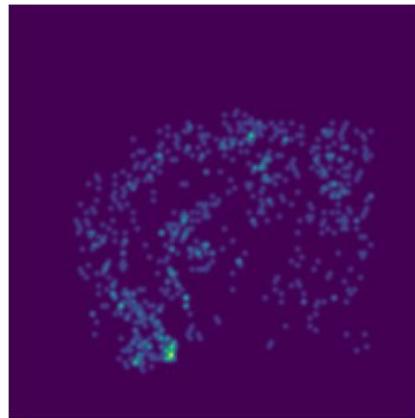
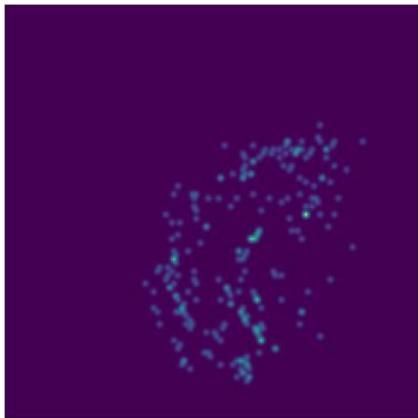


In-Silico GEX - X-Y-Z-axis Olfm3

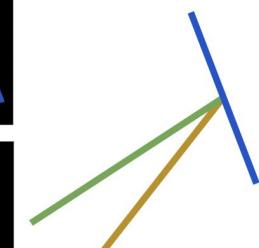
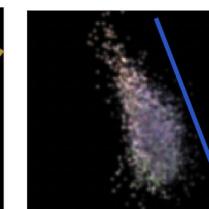
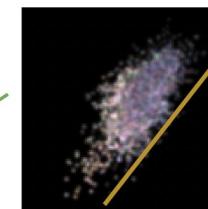
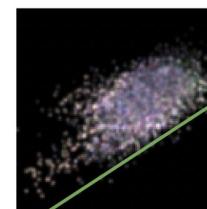
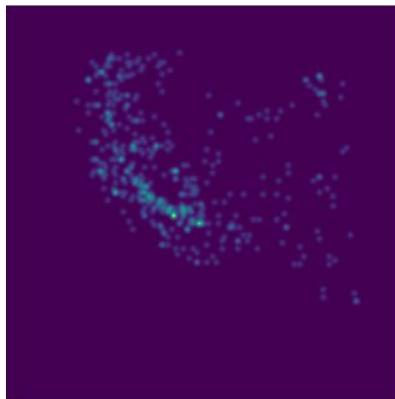
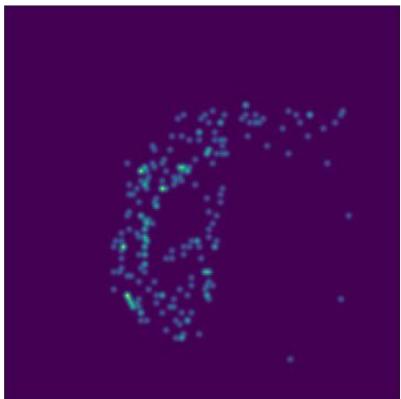
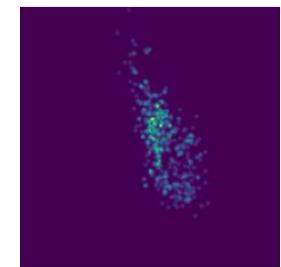
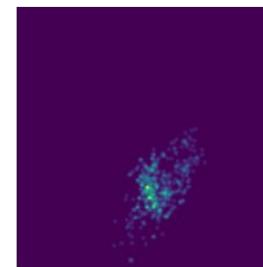
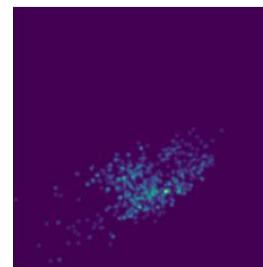
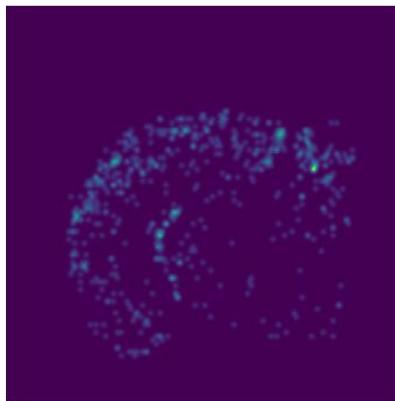
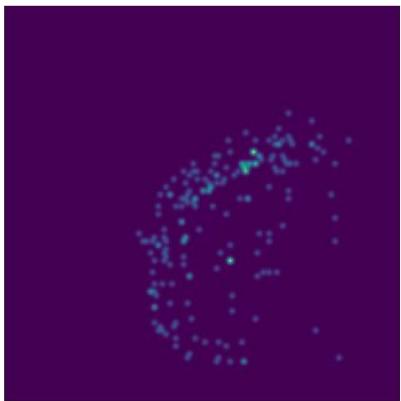


Real GEX Galnt14

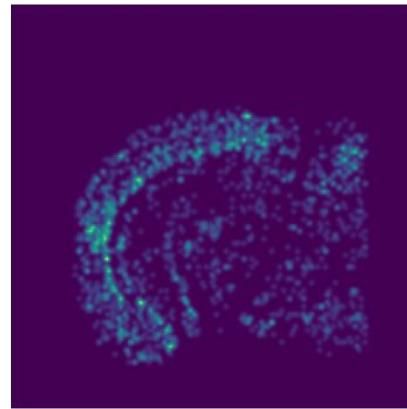
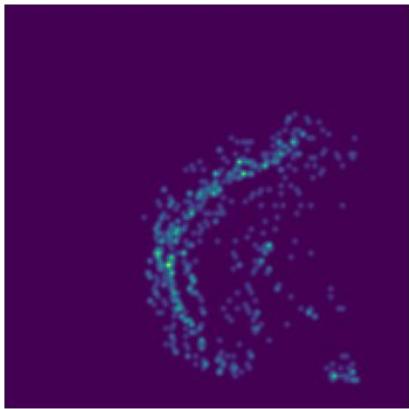
In-Silico GEX - X-Y-Z-axis Galnt14



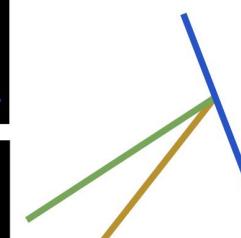
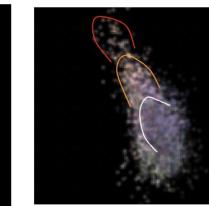
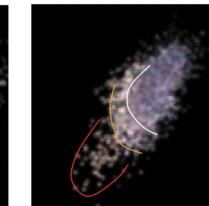
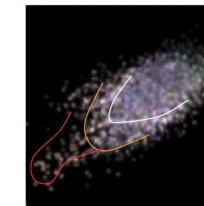
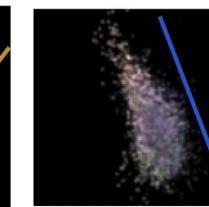
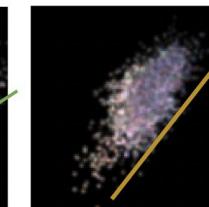
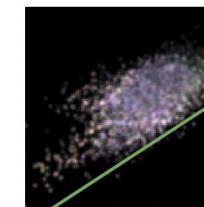
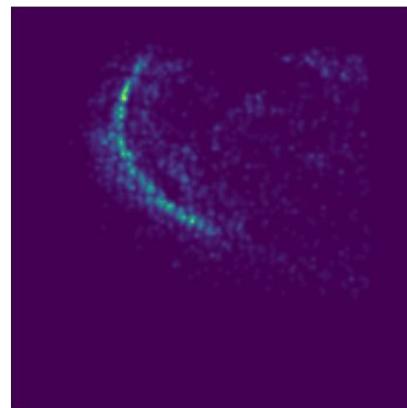
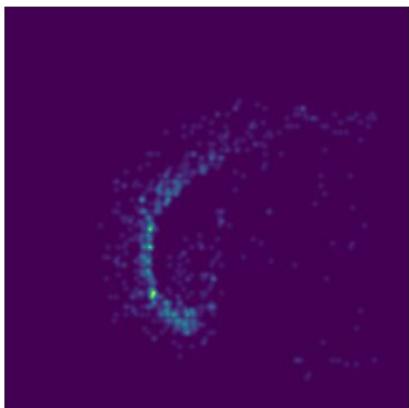
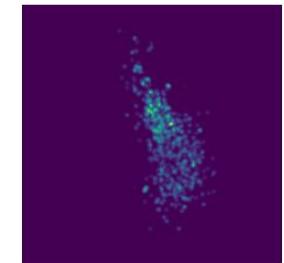
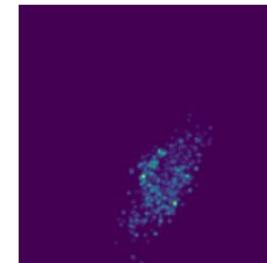
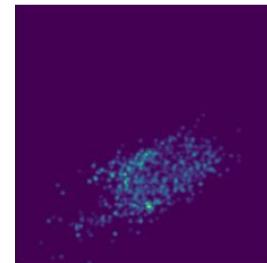
Real GEX Trhde



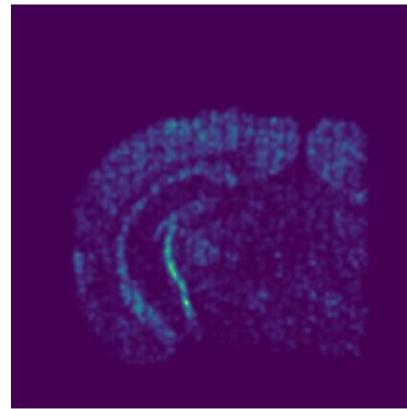
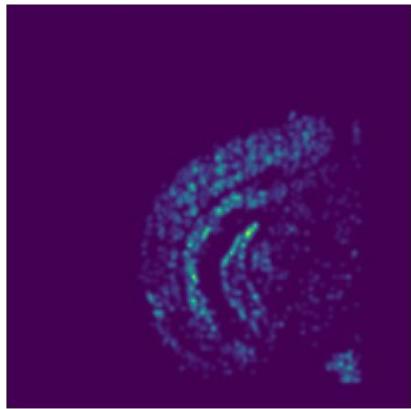
Real GEX lgsf21



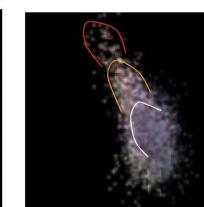
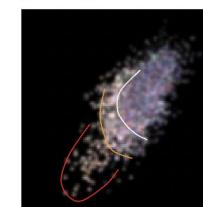
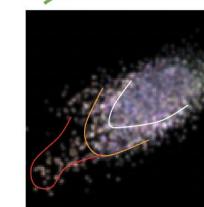
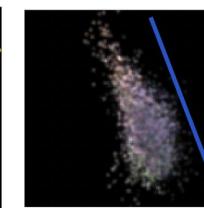
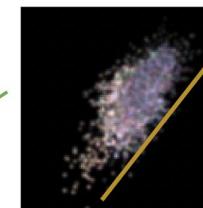
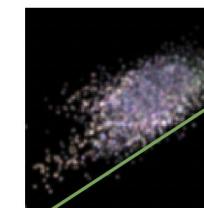
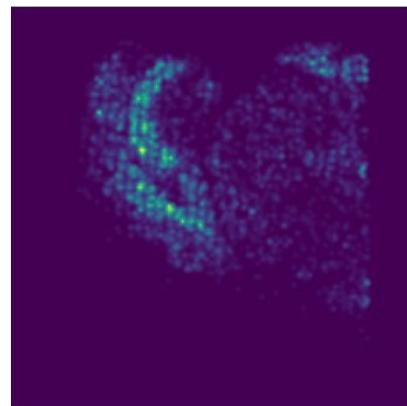
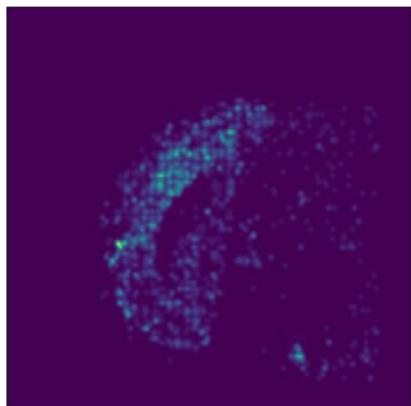
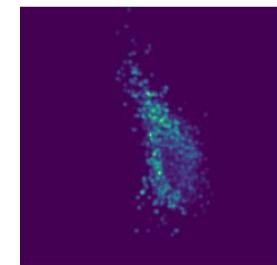
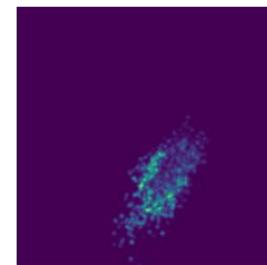
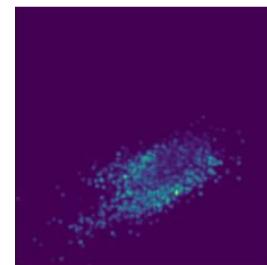
In-Silico GEX - X-Y-Z-axis lgsf21



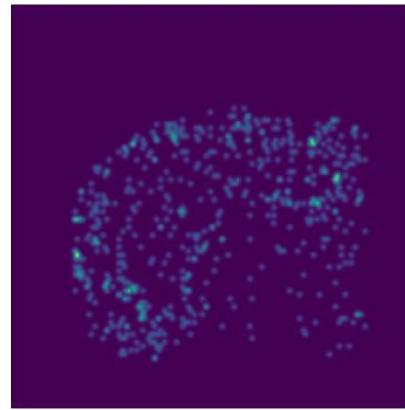
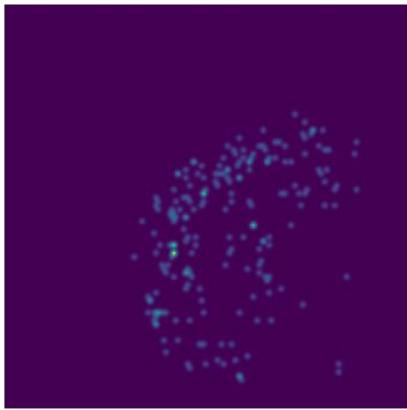
Real GEX Slc24a2



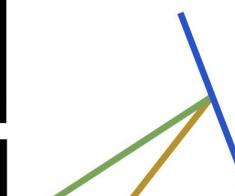
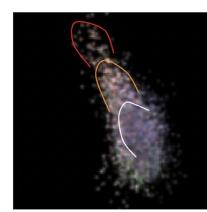
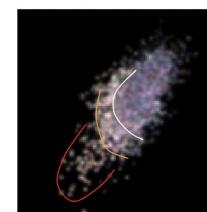
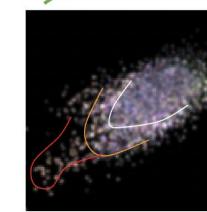
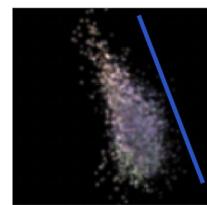
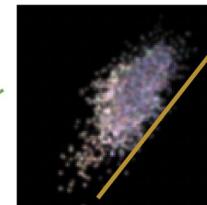
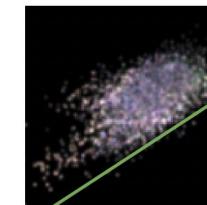
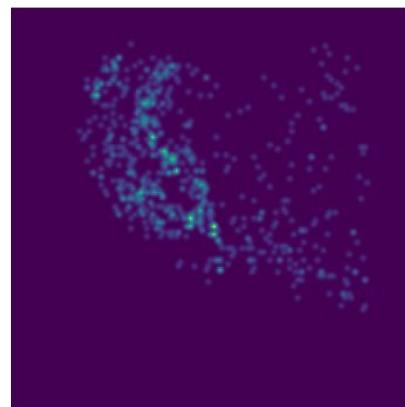
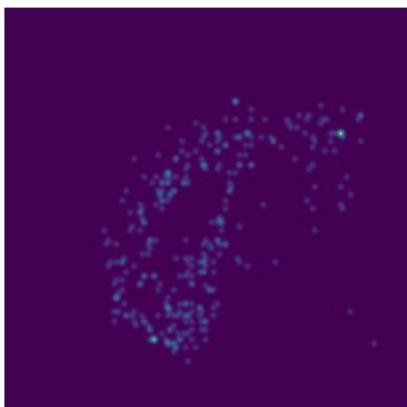
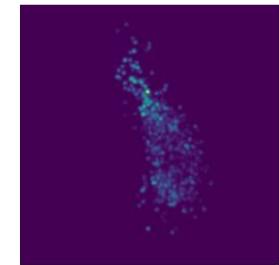
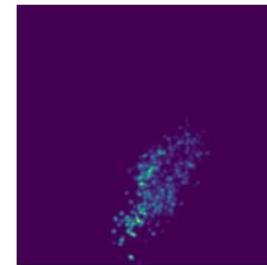
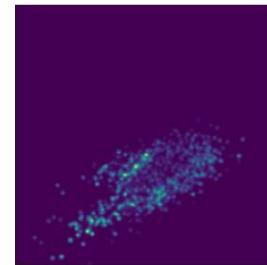
In-Silico GEX - X-Y-Z-axis Slc24a2



Real GEX Tenm3

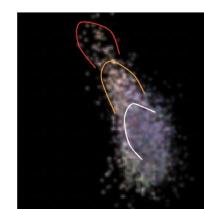
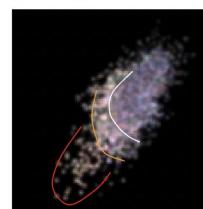
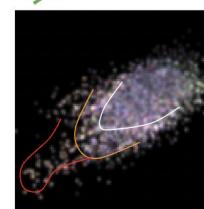
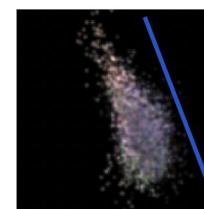
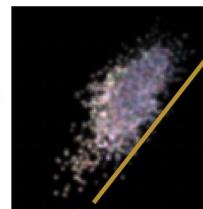
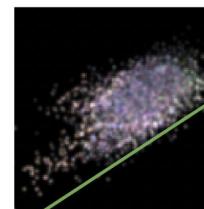
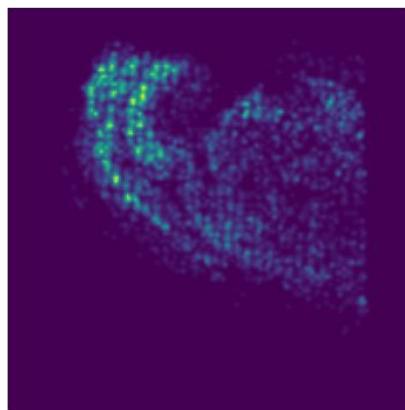
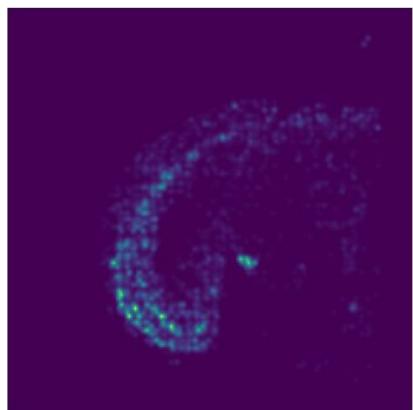
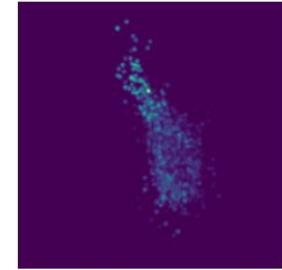
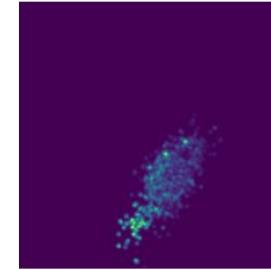
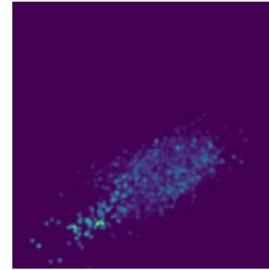
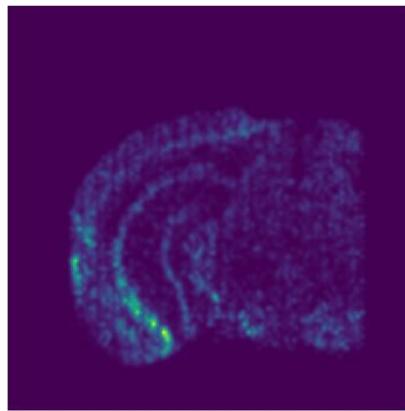
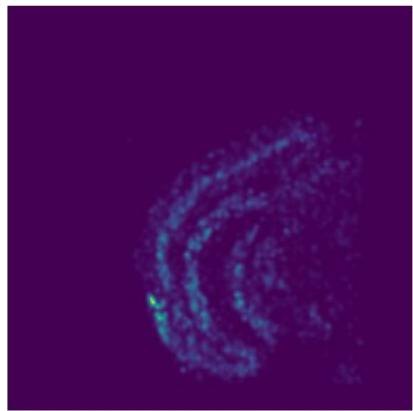


In-Silico GEX - X-Y-Z-axis Tenm3

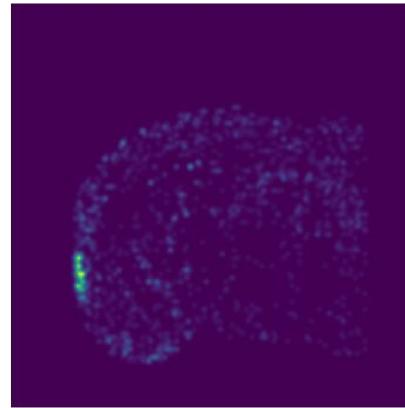
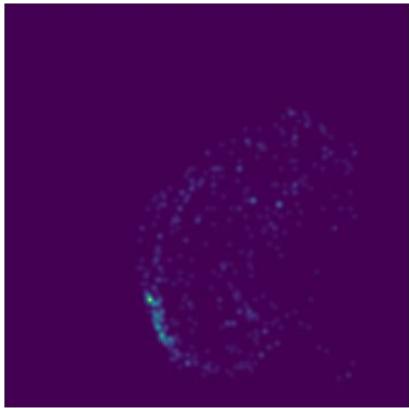


Real GEX Rab3c

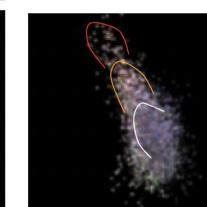
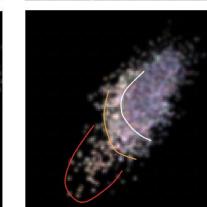
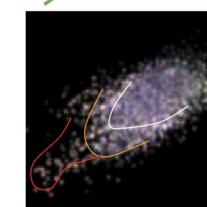
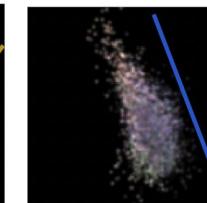
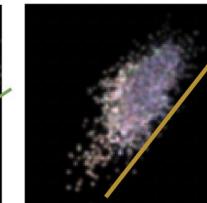
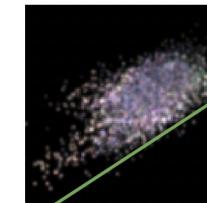
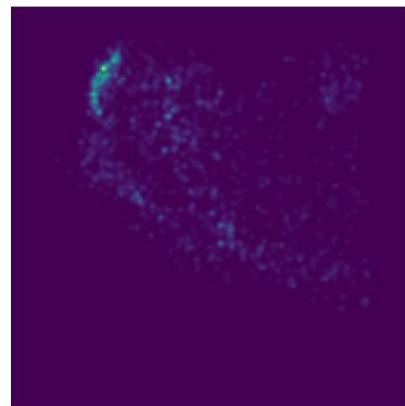
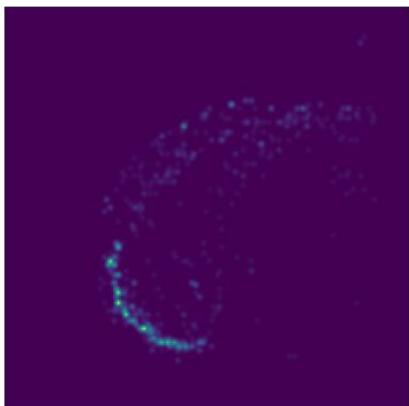
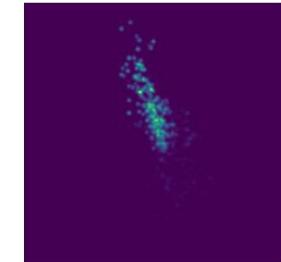
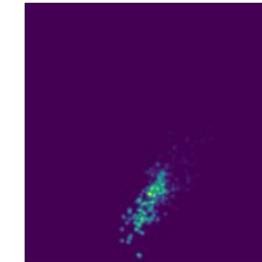
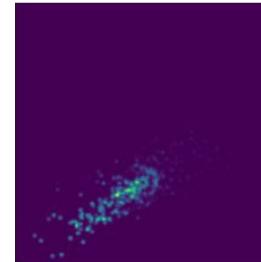
In-Silico GEX - X-Y-Z-axis Rab3c



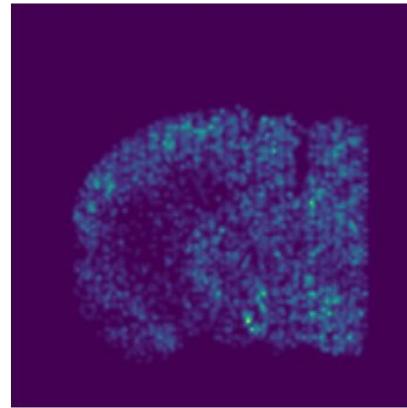
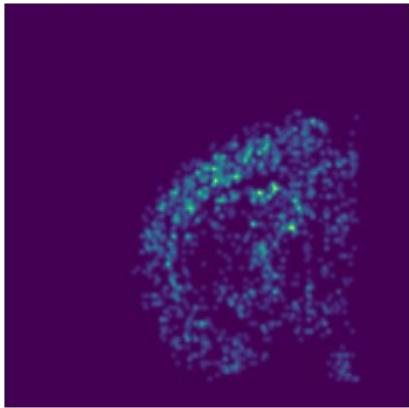
Real GEX Reln



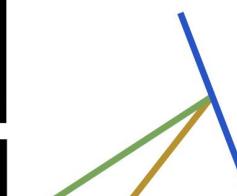
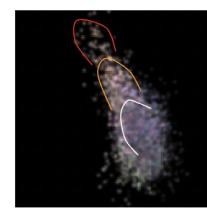
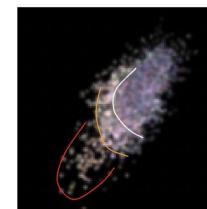
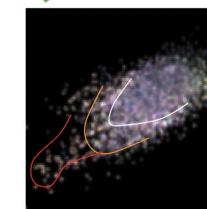
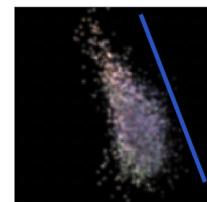
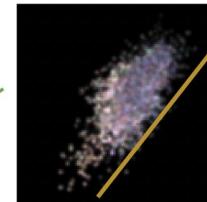
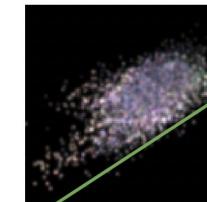
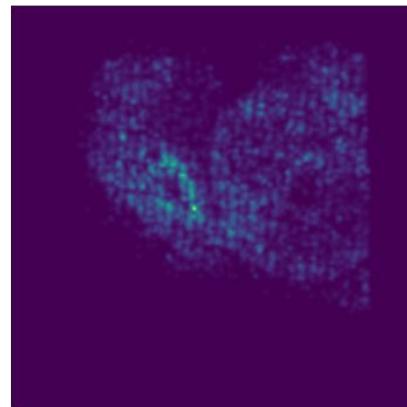
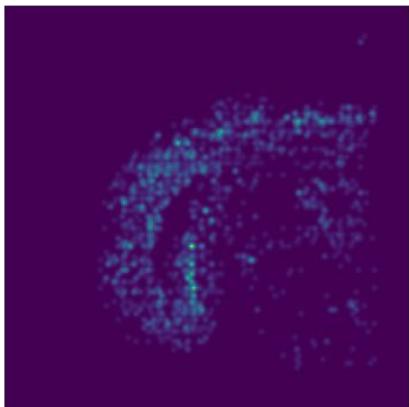
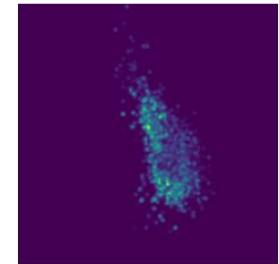
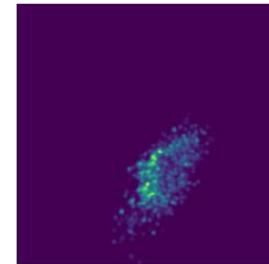
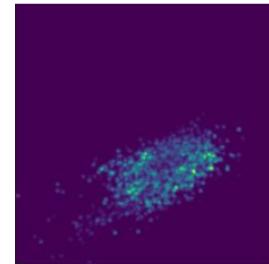
In-Silico GEX - X-Y-Z-axis Reln



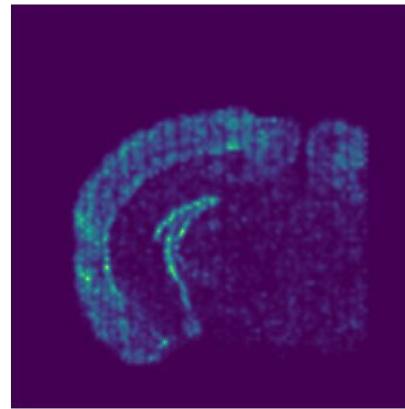
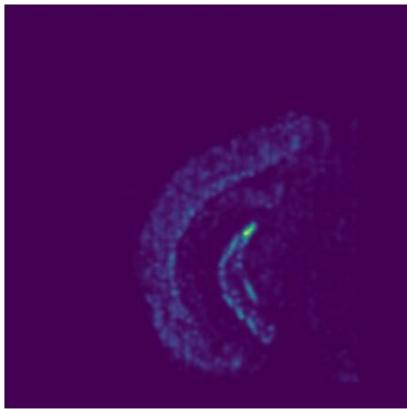
Real GEX Rcan2



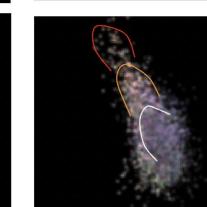
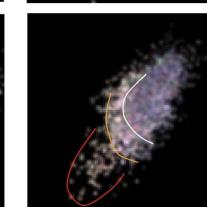
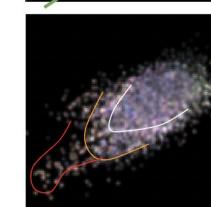
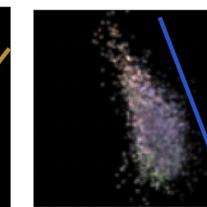
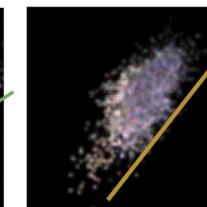
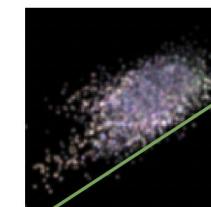
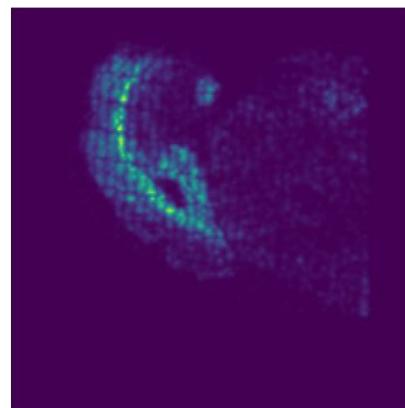
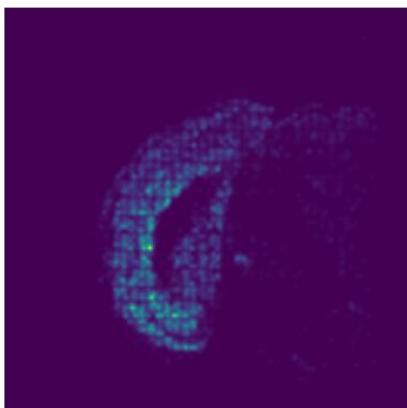
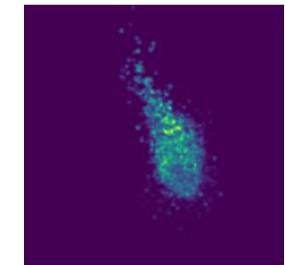
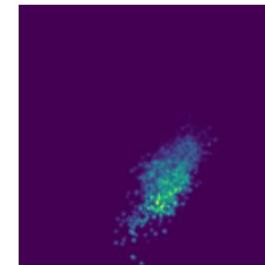
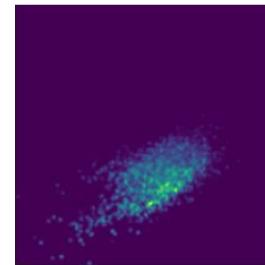
In-Silico GEX - X-Y-Z-axis Rcan2



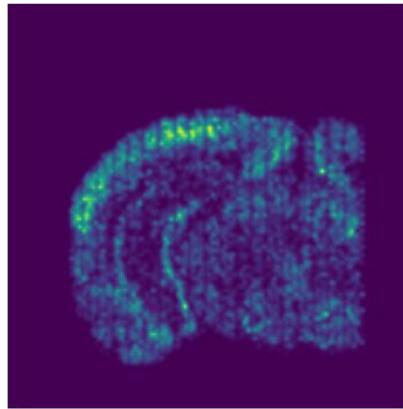
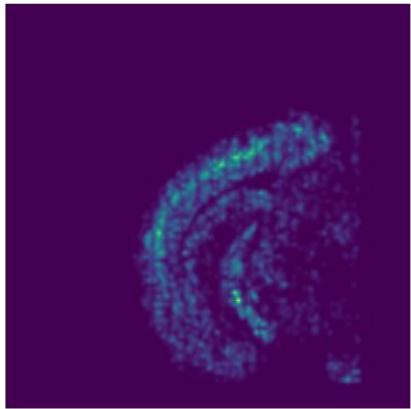
Real GEX Ncald



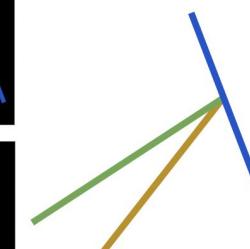
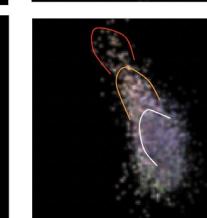
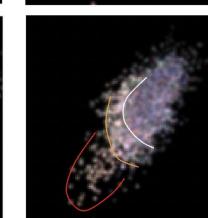
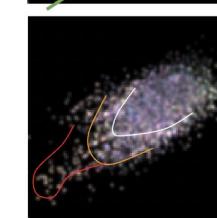
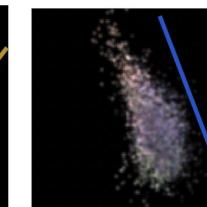
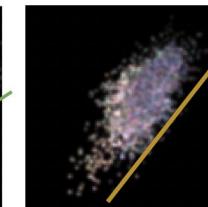
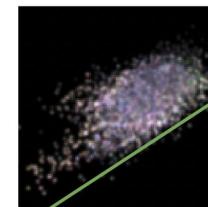
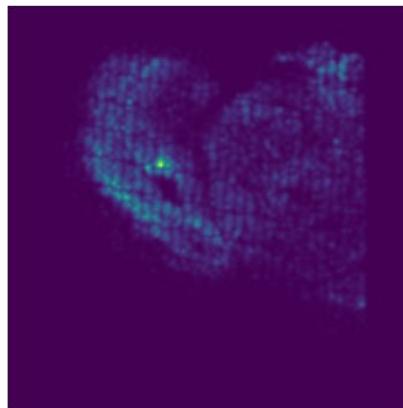
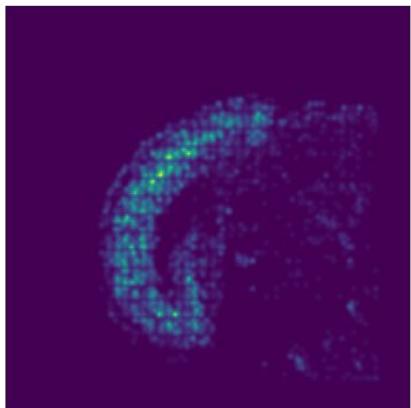
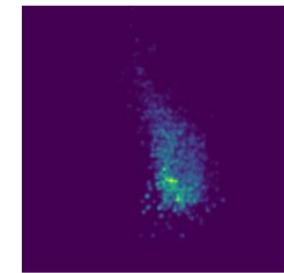
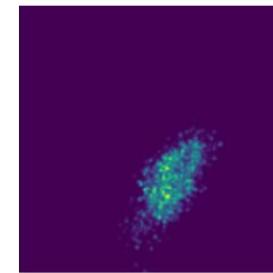
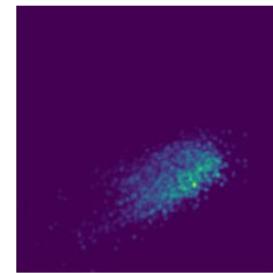
In-Silico GEX - X-Y-Z-axis Ncald



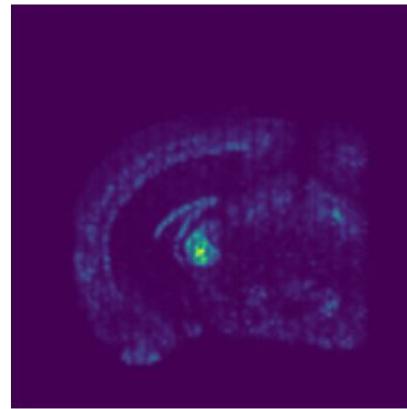
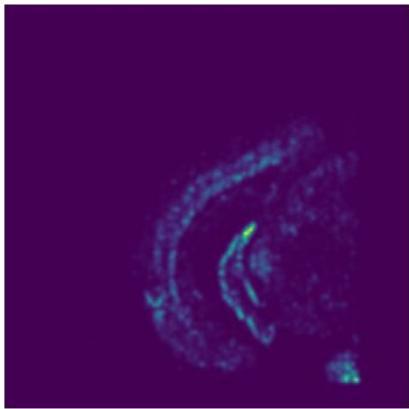
Real GEX Nrsn1



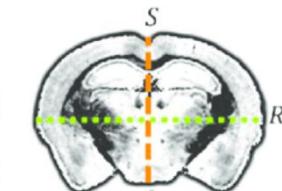
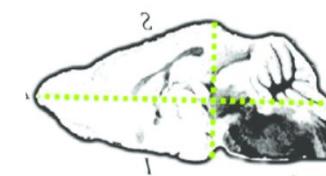
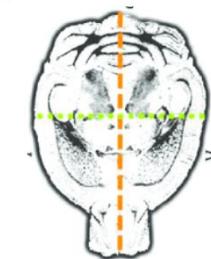
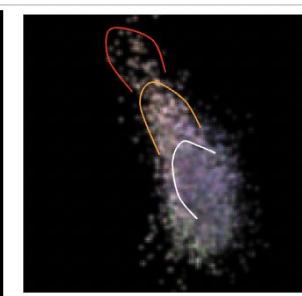
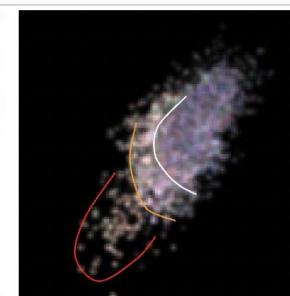
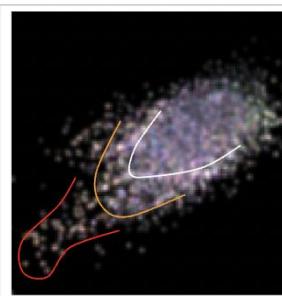
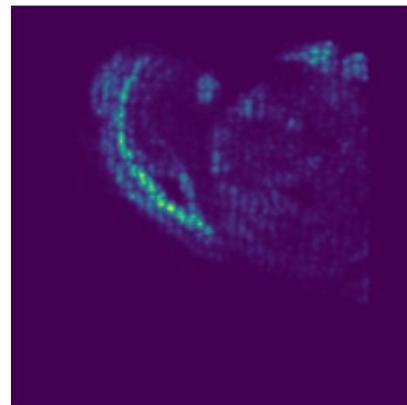
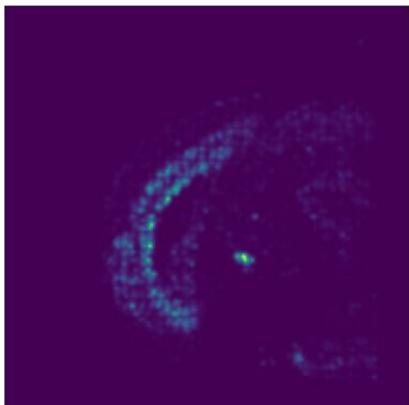
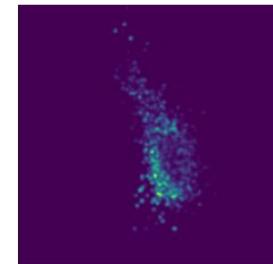
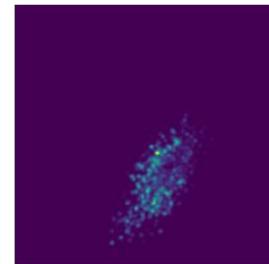
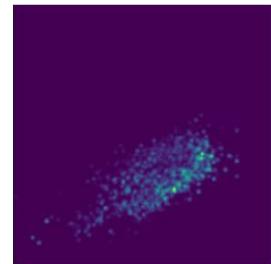
In-Silico GEX - X-Y-Z-axis Nrsn1



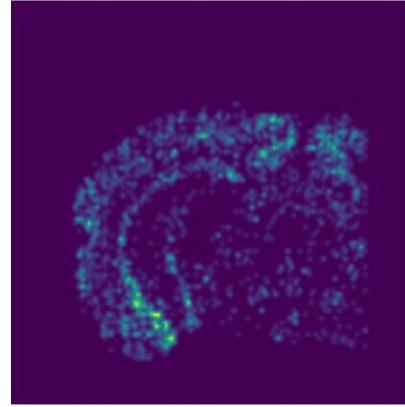
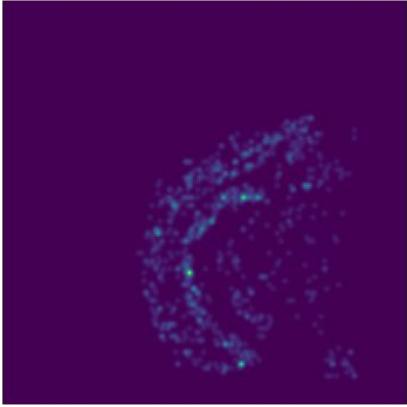
Real GEX Pcp4



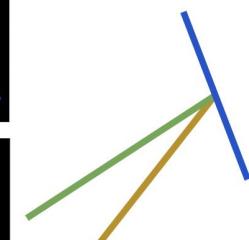
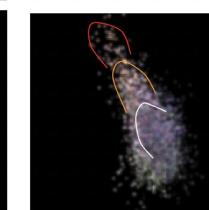
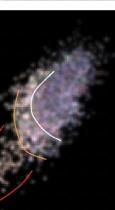
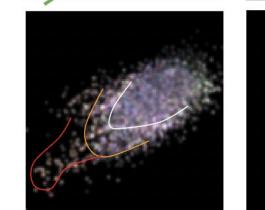
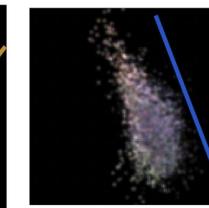
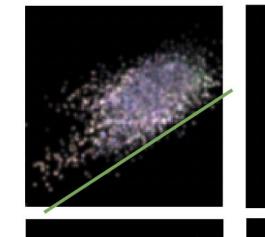
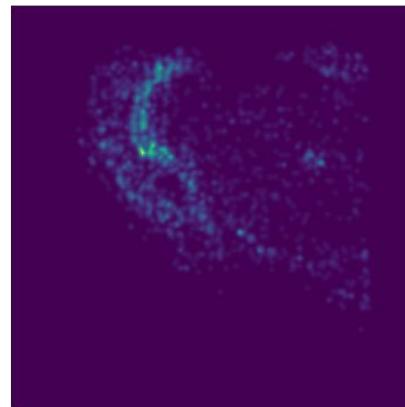
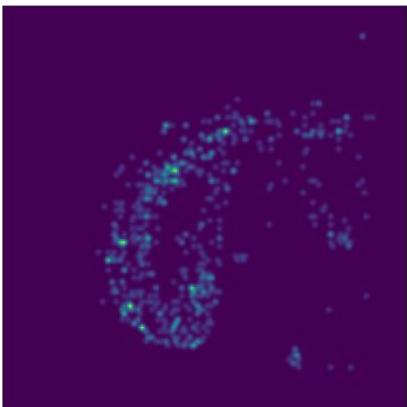
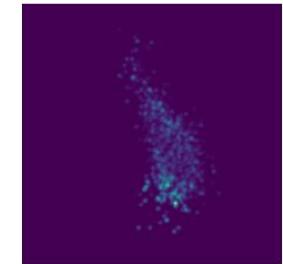
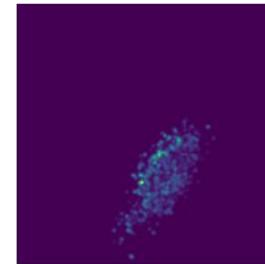
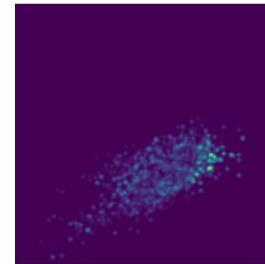
In-Silico GEX - X-Y-Z-axis Pcp4



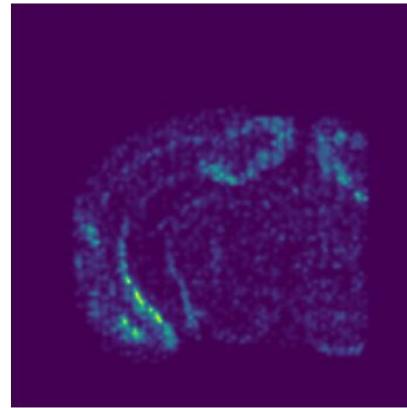
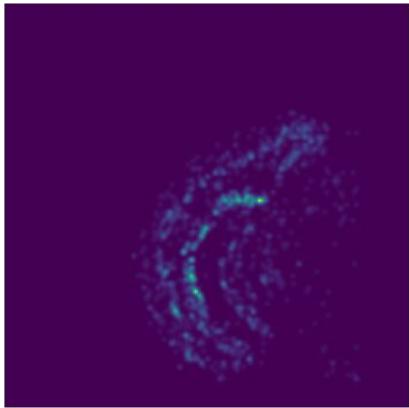
Real GEX Hcn1



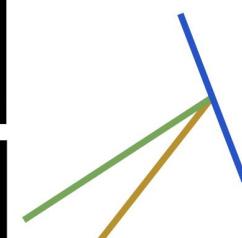
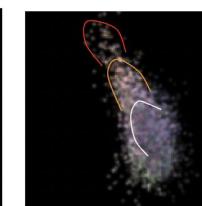
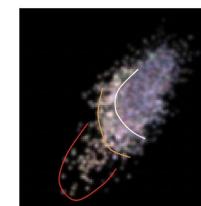
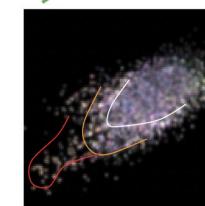
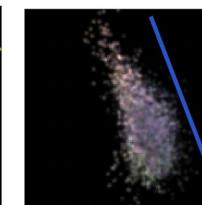
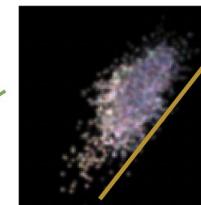
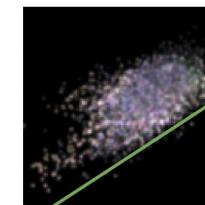
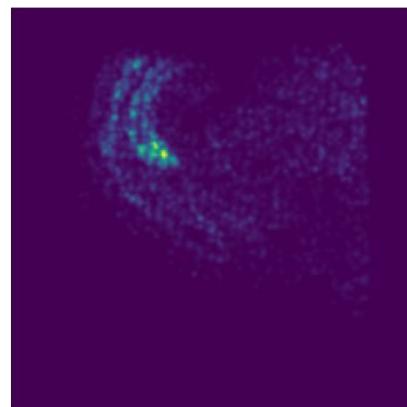
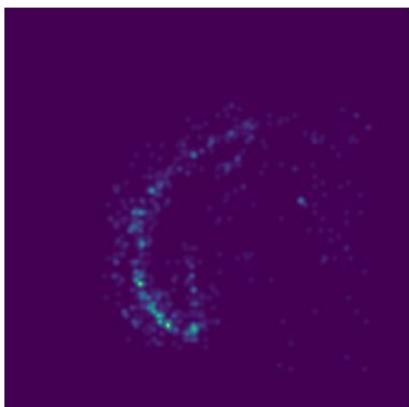
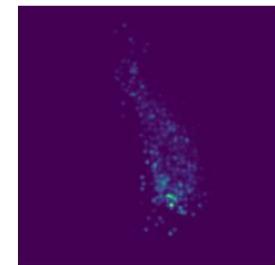
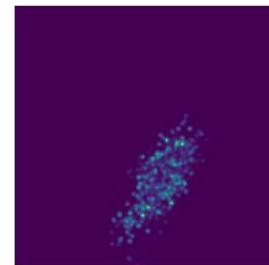
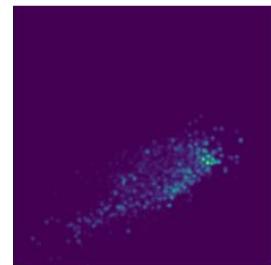
In-Silico GEX - X-Y-Z-axis Hcn1



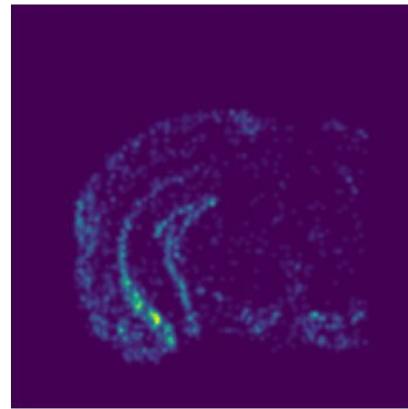
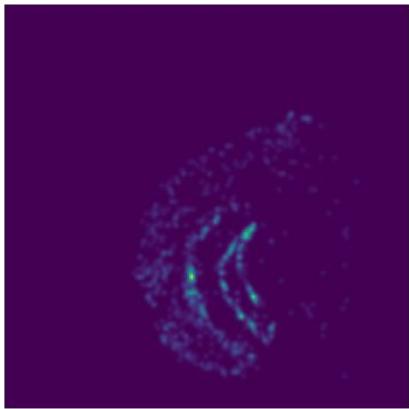
Real GEX Etv1



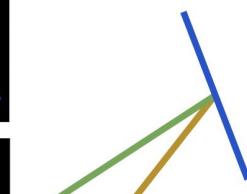
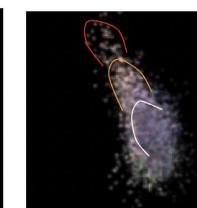
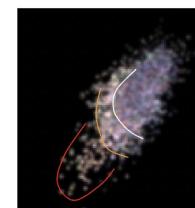
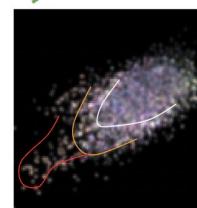
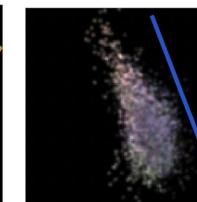
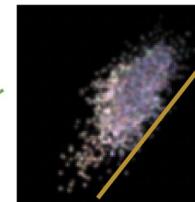
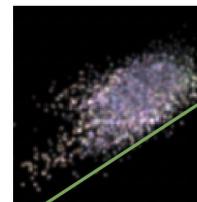
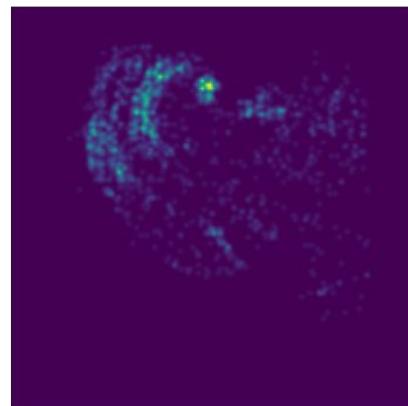
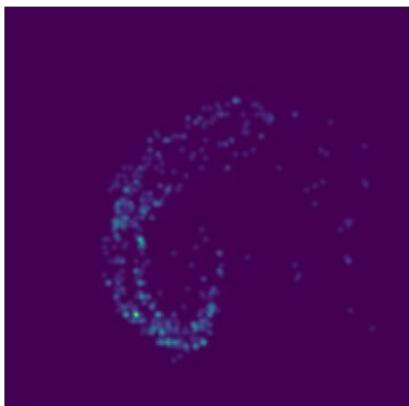
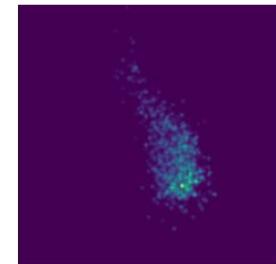
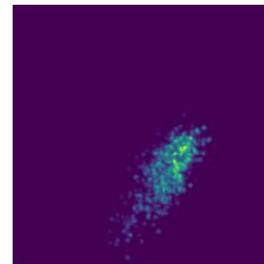
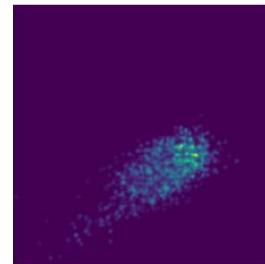
In-Silico GEX - X-Y-Z-axis Etv1



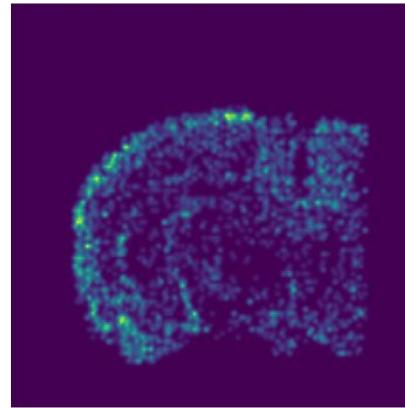
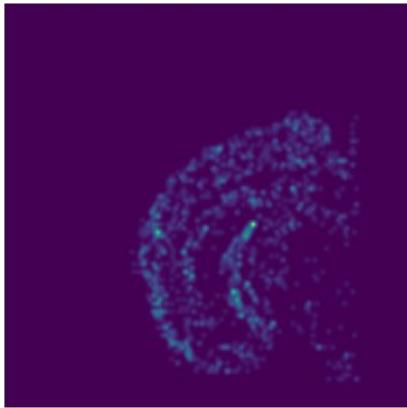
Real GEX Syt17



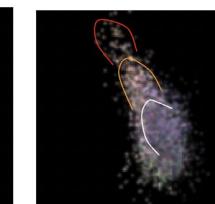
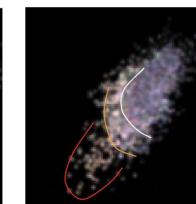
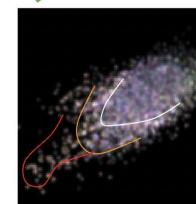
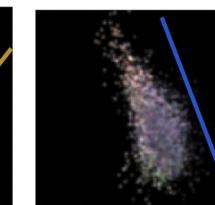
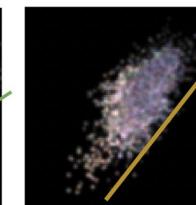
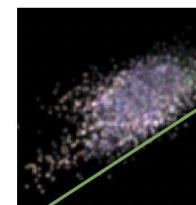
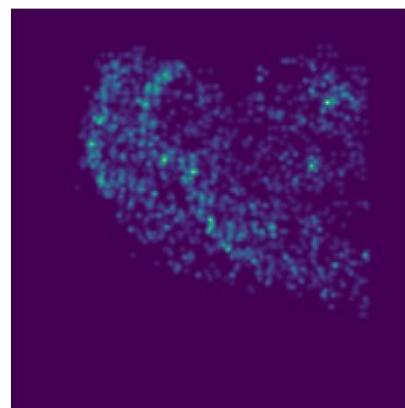
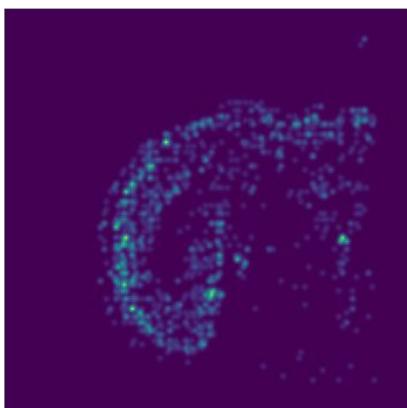
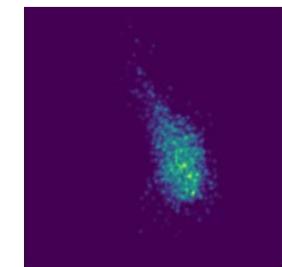
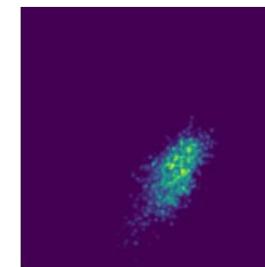
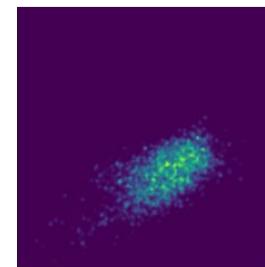
In-Silico GEX - X-Y-Z-axis Syt17



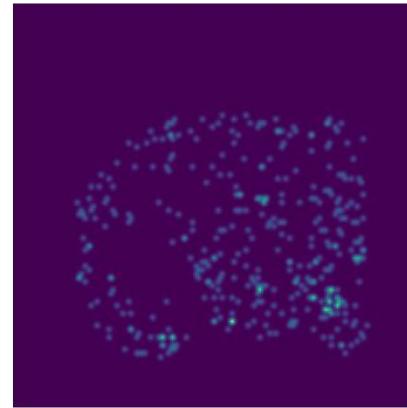
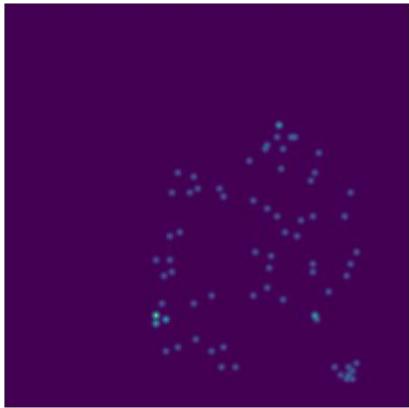
Real GEX Pam



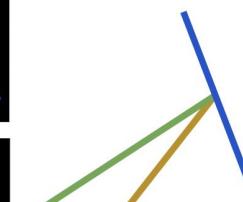
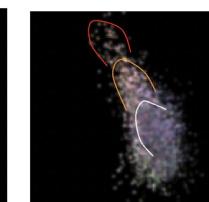
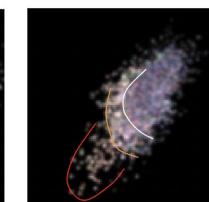
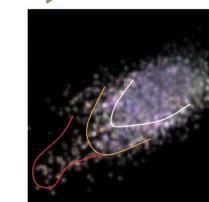
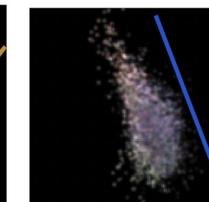
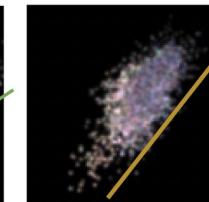
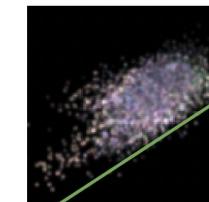
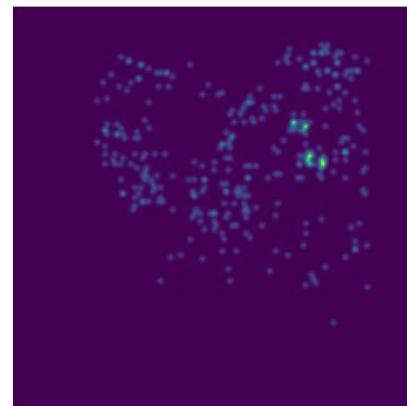
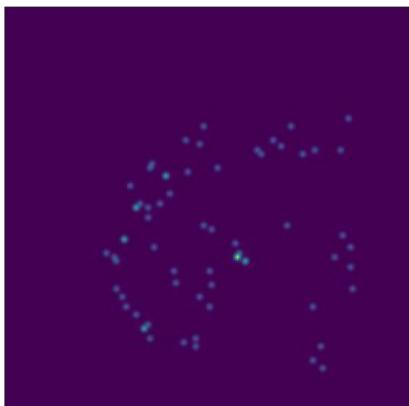
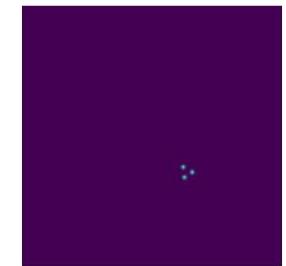
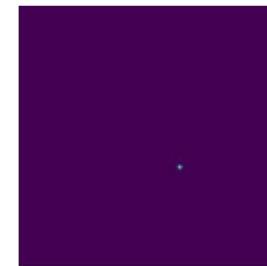
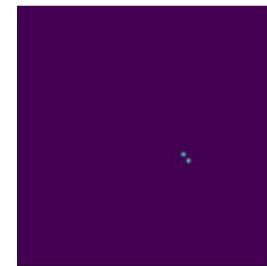
In-Silico GEX - X-Y-Z-axis Pam



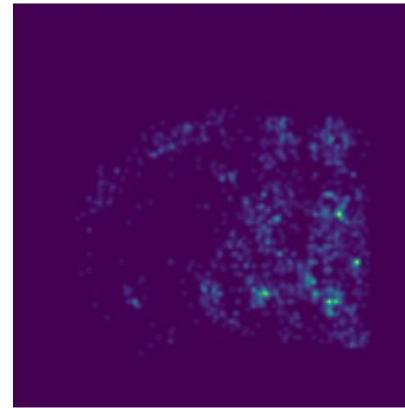
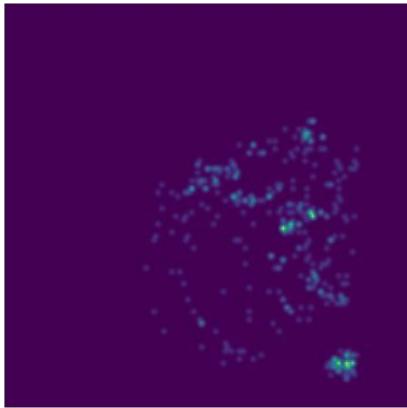
Real GEX Nrg1



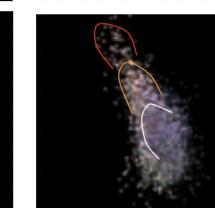
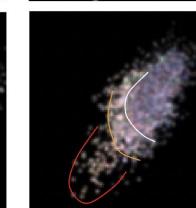
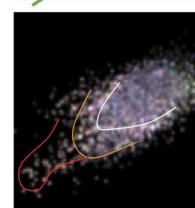
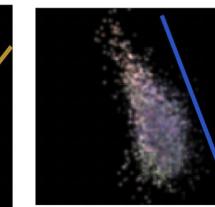
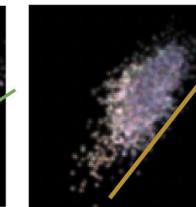
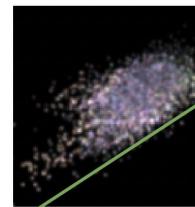
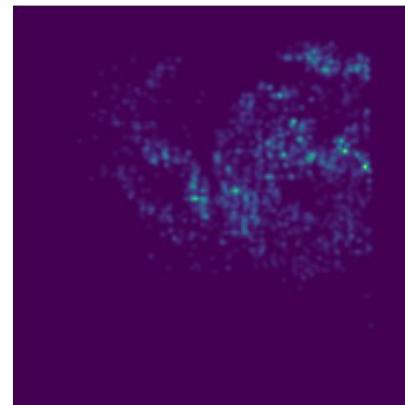
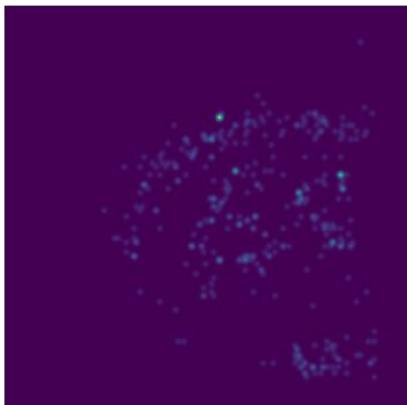
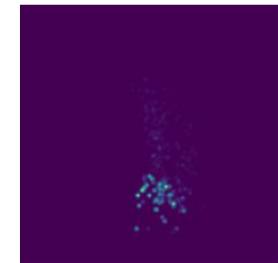
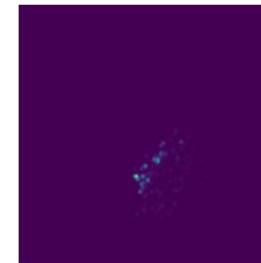
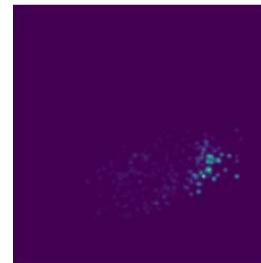
In-Silico GEX - X-Y-Z-axis Nrg1



Real GEX Syt2



In-Silico GEX - X-Y-Z-axis Syt2



Model primer

arg min (

{ {

A: How accurate are cell distributions compared to training?

A +

B: Whether the parameters for the new cell locations are actual signals?

B +

C: Whether gene to pixel transformations are learned correctly?

D +

D: Whether pixel image prediction is aligned with reality on a local predicted imaging spot level?

E +

F: Whether pixel image prediction is aligned with reality on a global predicted imaging tissue level?

}² }^{1/2}

F: Whether the movement of cells around fits well with the idea of Waddington's marble analogy?

)

Actual model layers ...

```
# beta is to learn gene expression  
# spatial signals from scRNAseq only  
self.summary_beta = nn.Linear(10,window_size*2*window_size*2*106,bias=False)  
  
# gex and pix are to learn transformation parameters that  
# assign based on image similarity to real histology  
self.summary_gex = nn.Linear(len(genes),100,bias=False)  
self.summary_pix = nn.Linear(32*32*3,100,bias=False)
```

... and that's it!

A fully interpretable linear neural network

- beta : maps KNN gene convolutions to **predict coordinates**
- gex : maps gene to **shared space** with images
- pix : maps images to **shared space** with genes

How are coordinates predicted?

1

Cells are randomly distributed onto a grid (e.g. $1000 \times 1000 \times G$) with G gene channels and a spatial kernel applied to diffuse genes

2

Convolutions around each cell are taken, giving a small window of about $5 \times 5 \times G$, where G are the gene channels

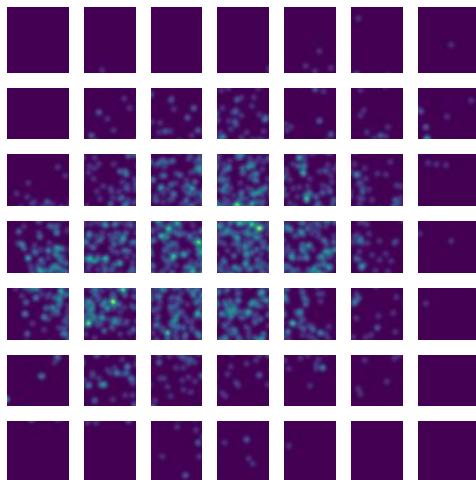
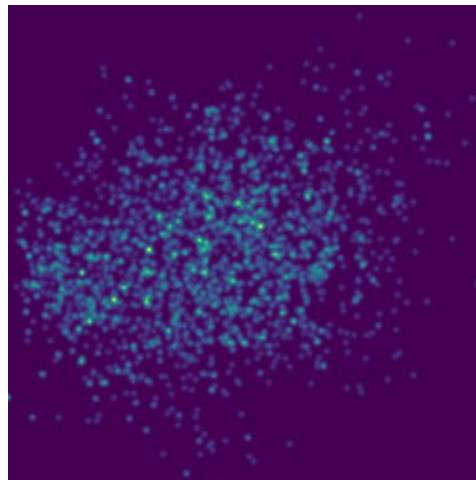
3

Convolutions of cells are put through a Neural Network to predict where the cell is next likely to occur

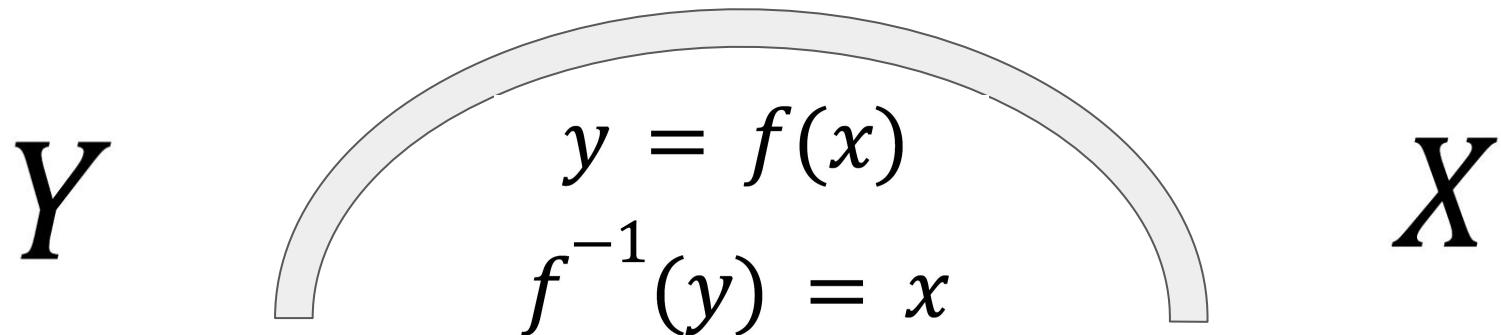
4

This is a repeated process to continuously update a walking process of the cell to the next likely location based on gene expression around the grid ...

... like Waddington's marble analogy!



What is a shared space?

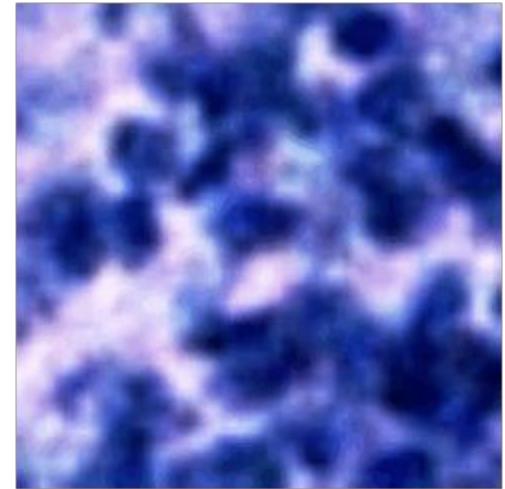
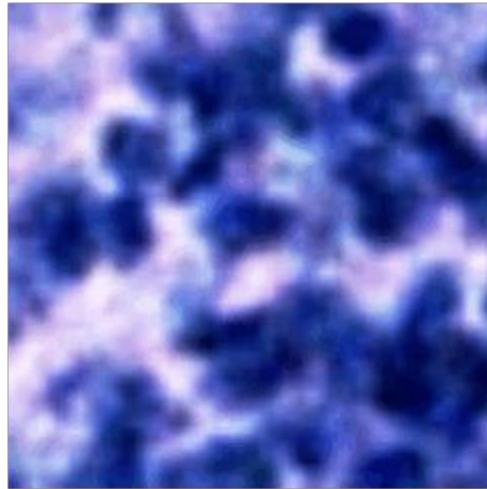
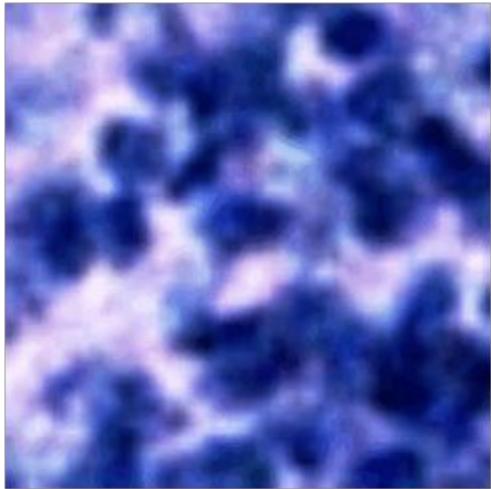


Spatial Histology to Spatial Transcripts

Methylation to Chromatin Accessibility to Chromatin Conformation to Chromatin Profiles

Genomics to Epigenomics to Transcriptomics to Proteomics

What can a shared space do?





https://github.com/AskExplain/scRNA_maps/

Acknowledgements for inspiration

Fei Yang

University of Queensland

Dr Alan Huang

University of Queensland

Professor Geoffrey McLachlan

University of Queensland

Professor Cinzia Viroli

University of Bologna

Dr Quan Nguyen

University of Queensland

Dr Jessica Mar

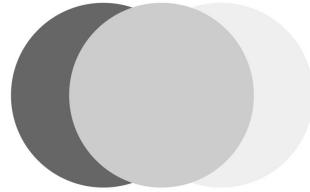
University of Queensland

Dr Ruth Pidsley

Garvan Institute of Medical Research

Our vision is to create great collaborations from science to industry

Through an interest and passion for science to progress, our commitment is to
work on, communicate around, and support collaborations



AskExplain
david.b@askexplain.com