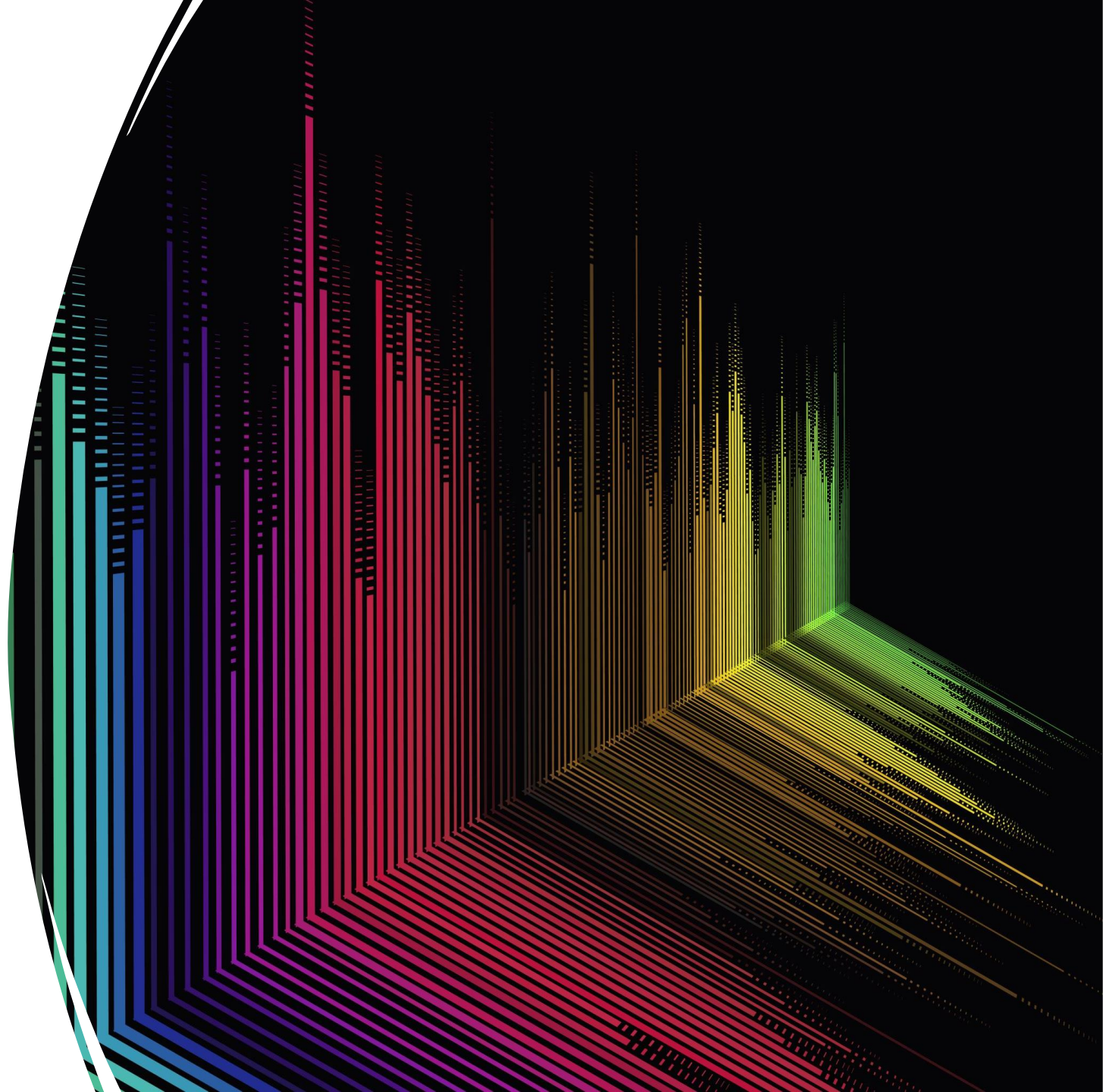


# SlideSeq Update

---

by Kian

4/22/24



# Outline



HPC3

Usage and memory allocation



Why Python?

Data formatting



Fully integrated e11 and e14

SCdreamer



ScanPy Pipeline for SlideSeq

Pipeline demo



Tangram Algorithm

All pucks are referenced mapped

# HPC3

## Job submissions

- sbatch commands

## Memory Allocation

- GPU use is necessary for certain algorithms



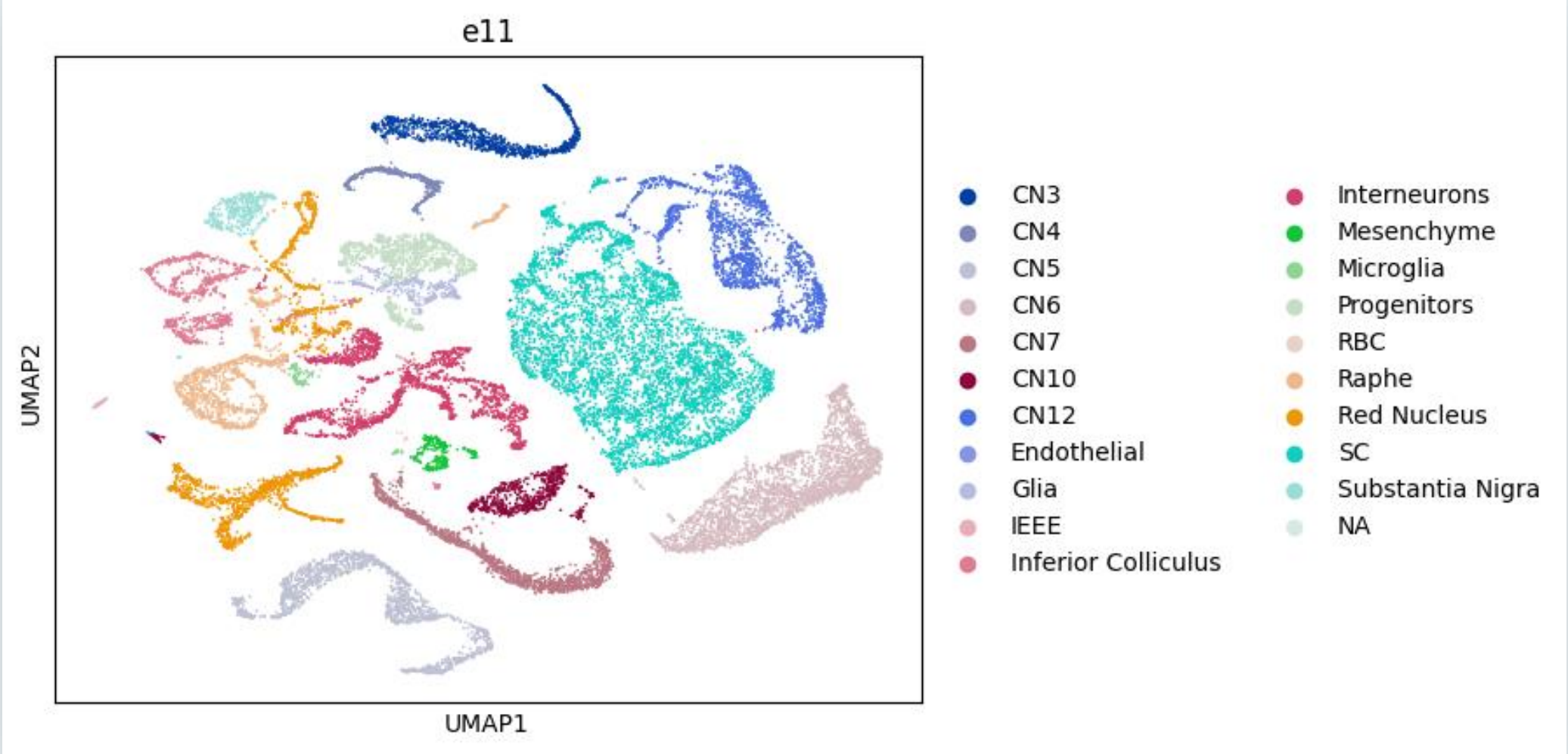
# Why Python?

R leaves data in  
.rds format

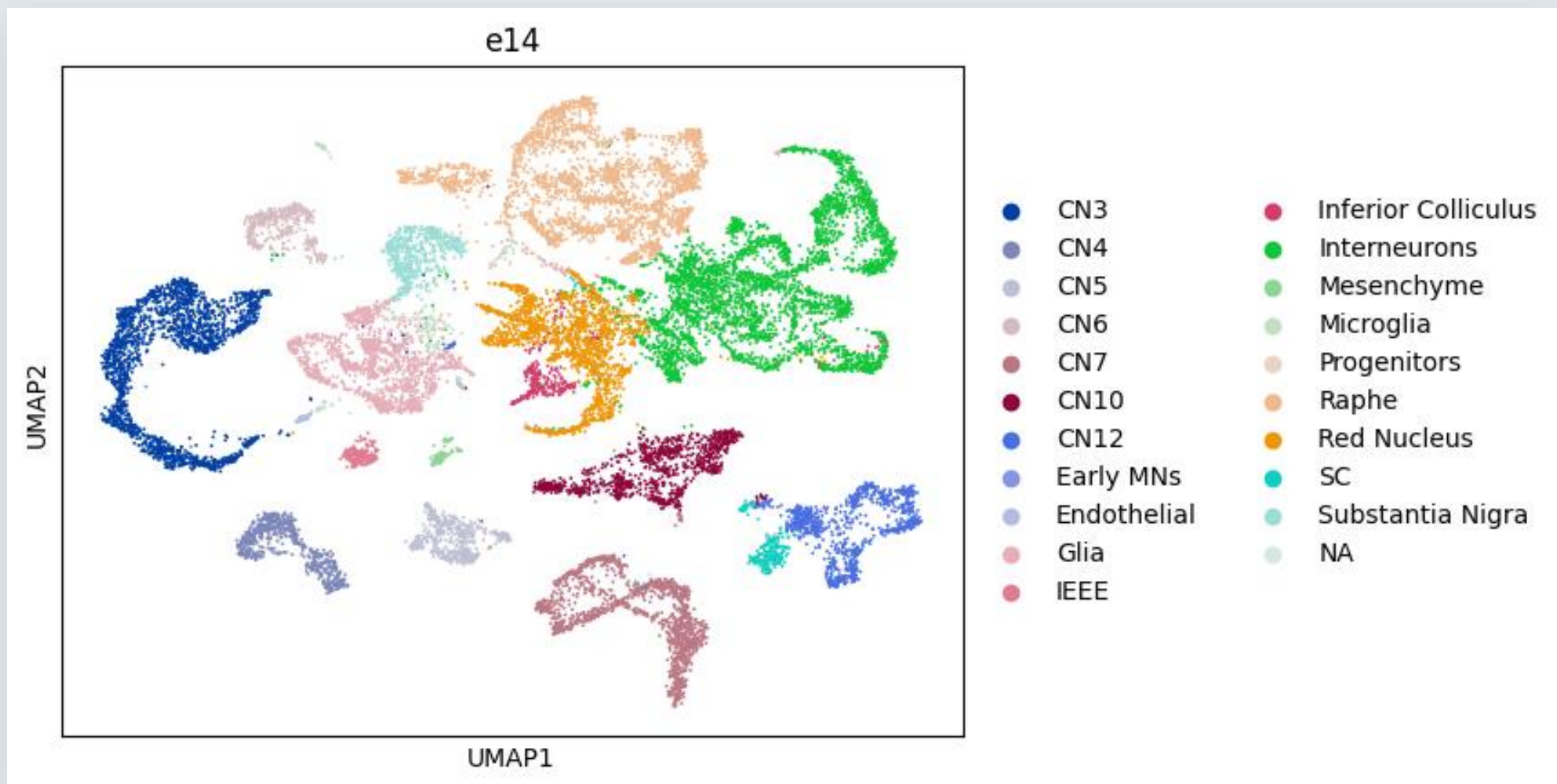
Python reference  
mapping  
algorithms seem  
to perform better

---

Fully integrated e11 and e14



## Fully integrated e11 and e14





# • ScanPy Pipeline for SlideSeq

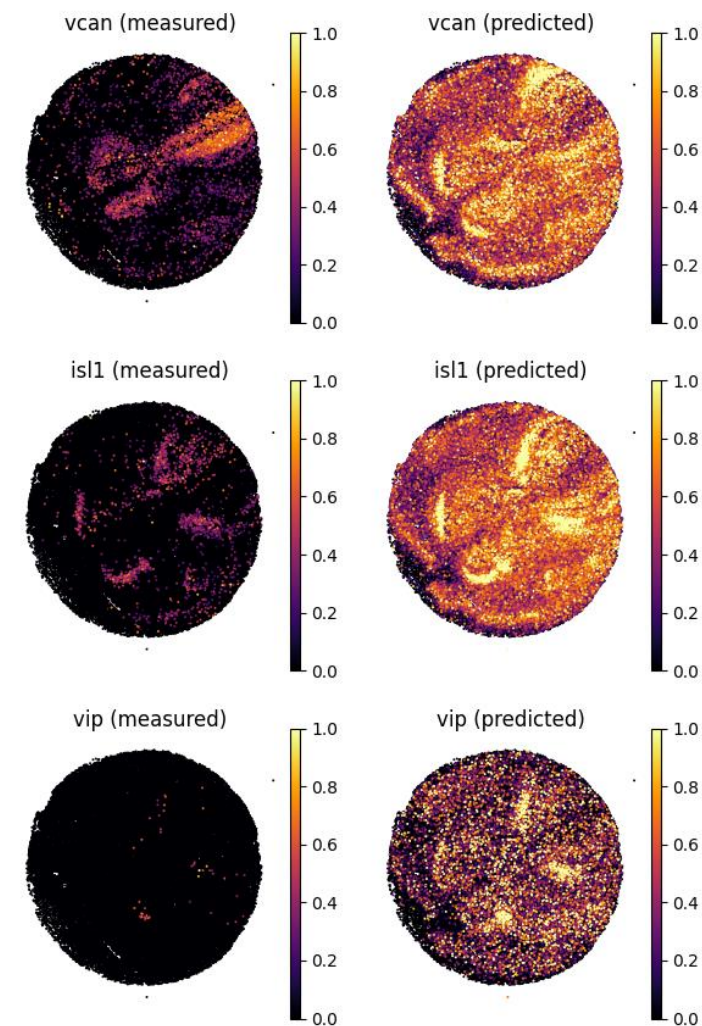
Quick run through code



# Tangram Results



Measured vs Predicted  
(Puck 03)



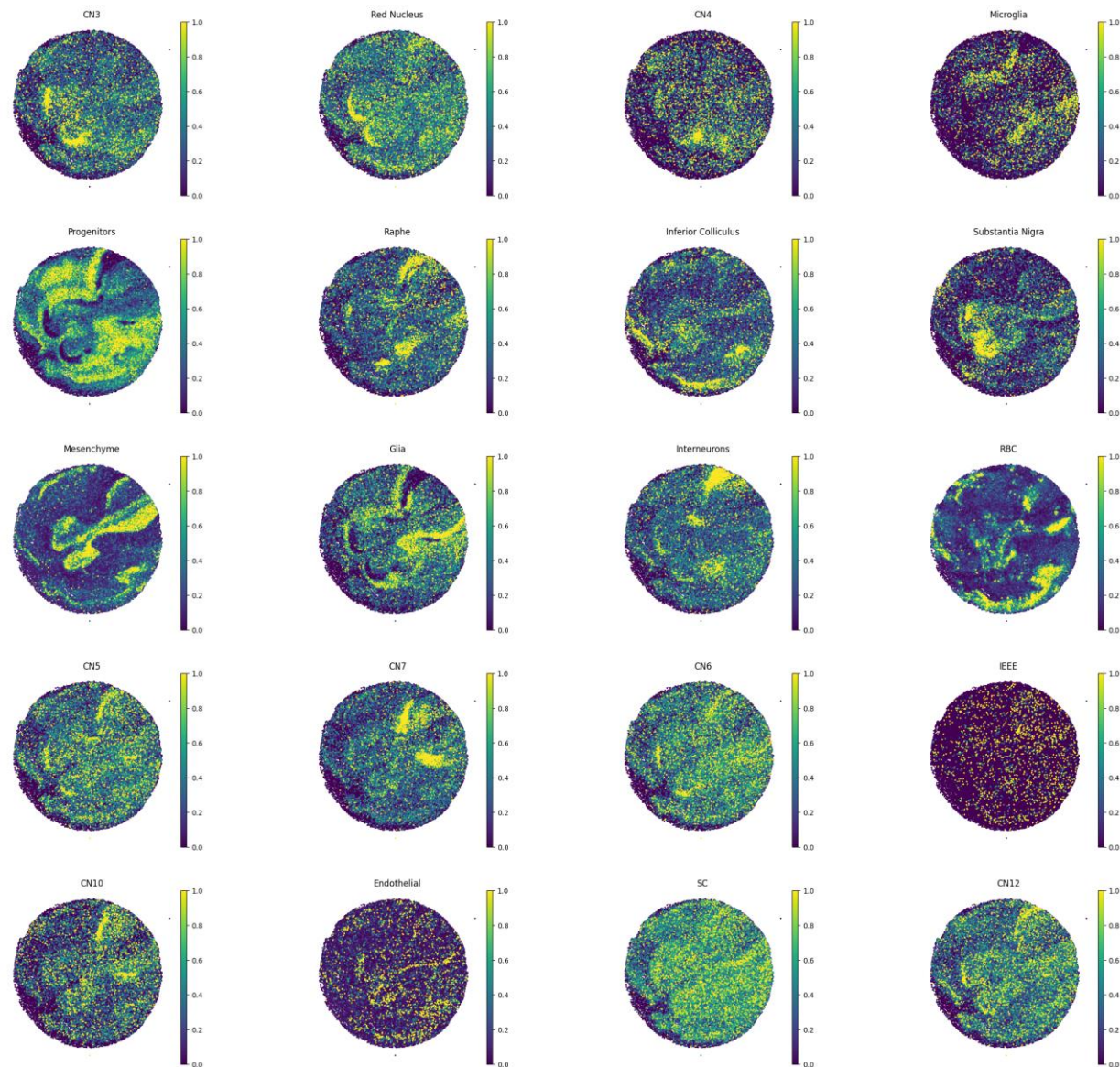
Tangram Algorithm scores (Puck 03)

	train_score	sparsity_sc	sparsity_sp	sparsity_diff
vcan	0.320837	0.532377	0.919604	0.387227
isl1	0.289753	0.382663	0.967144	0.584481
vip	0.771695	0.969076	0.998941	0.029864



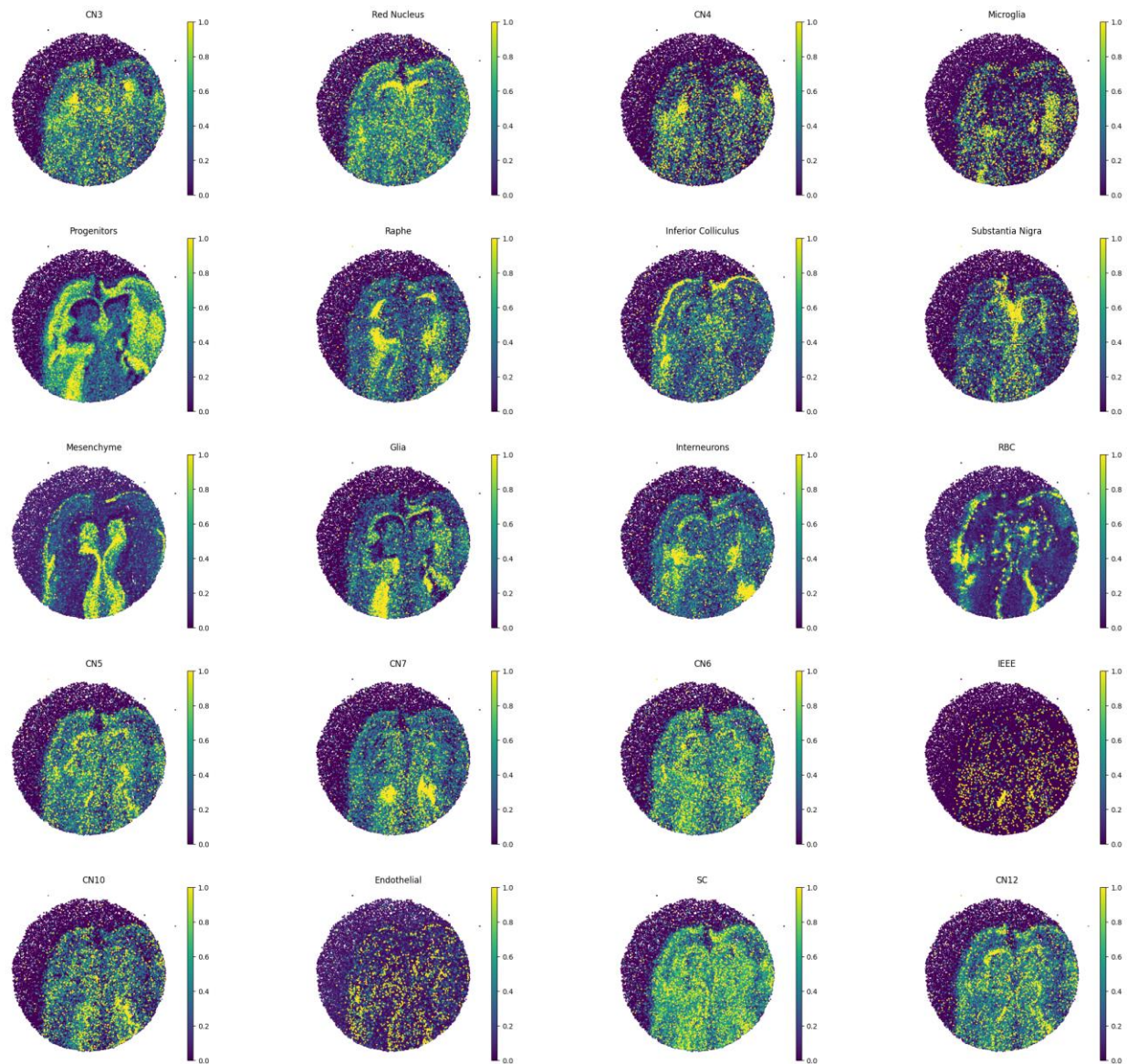
e11  
mapped to  
puck 03

---



e11  
mapped to  
puck 05

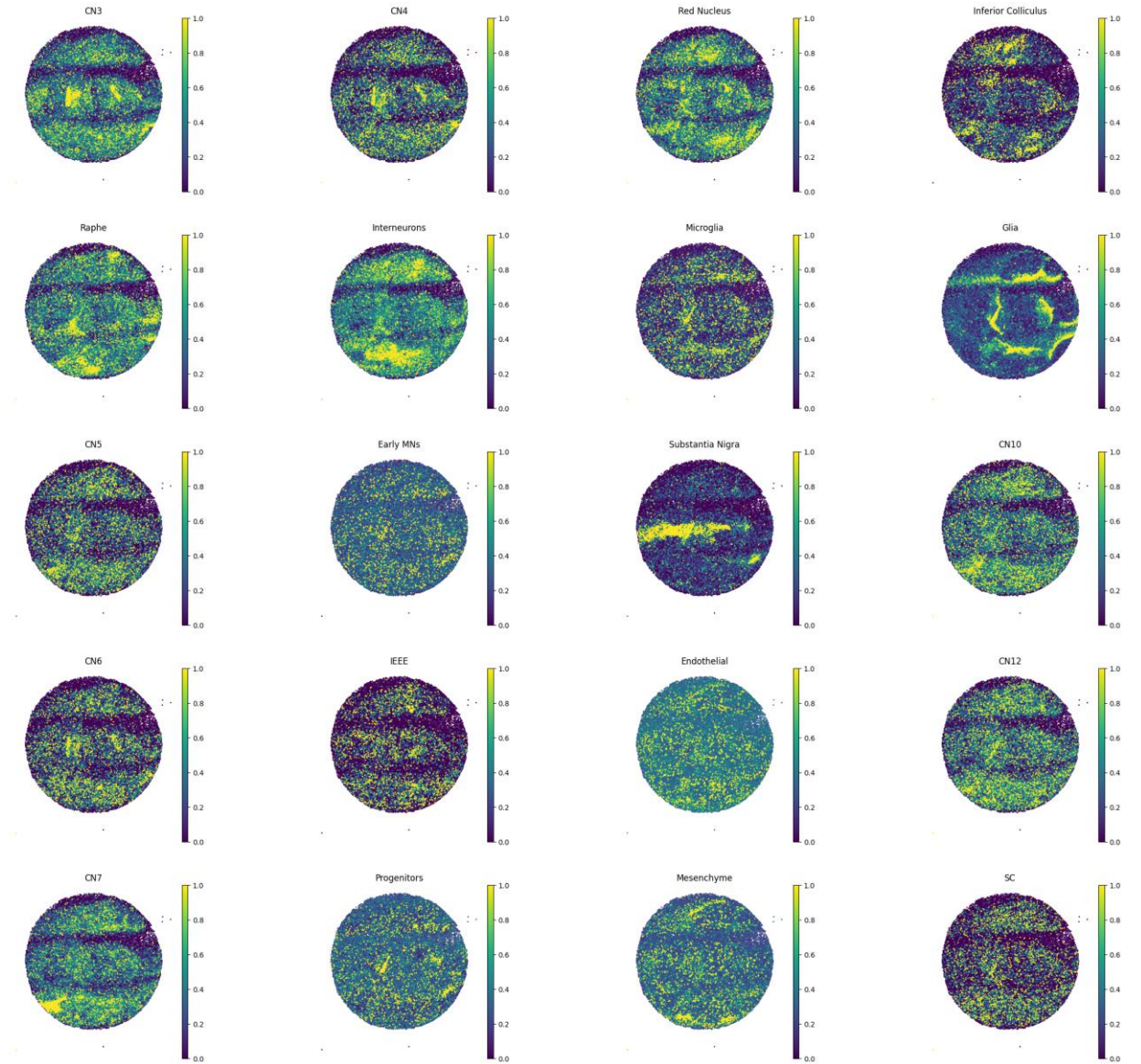
---





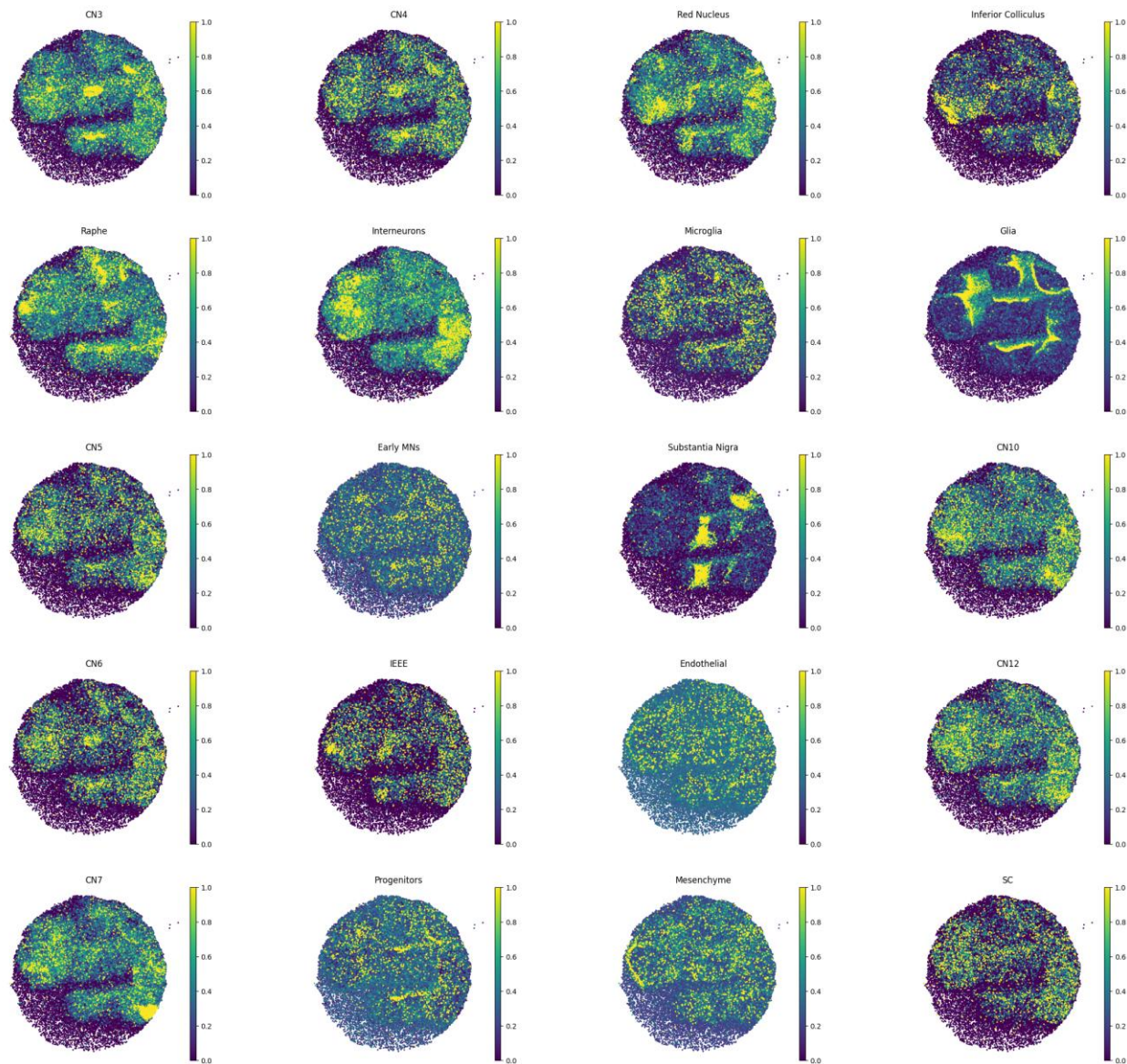
e14  
mapped to  
puck 06

---



e14  
mapped to  
puck 07

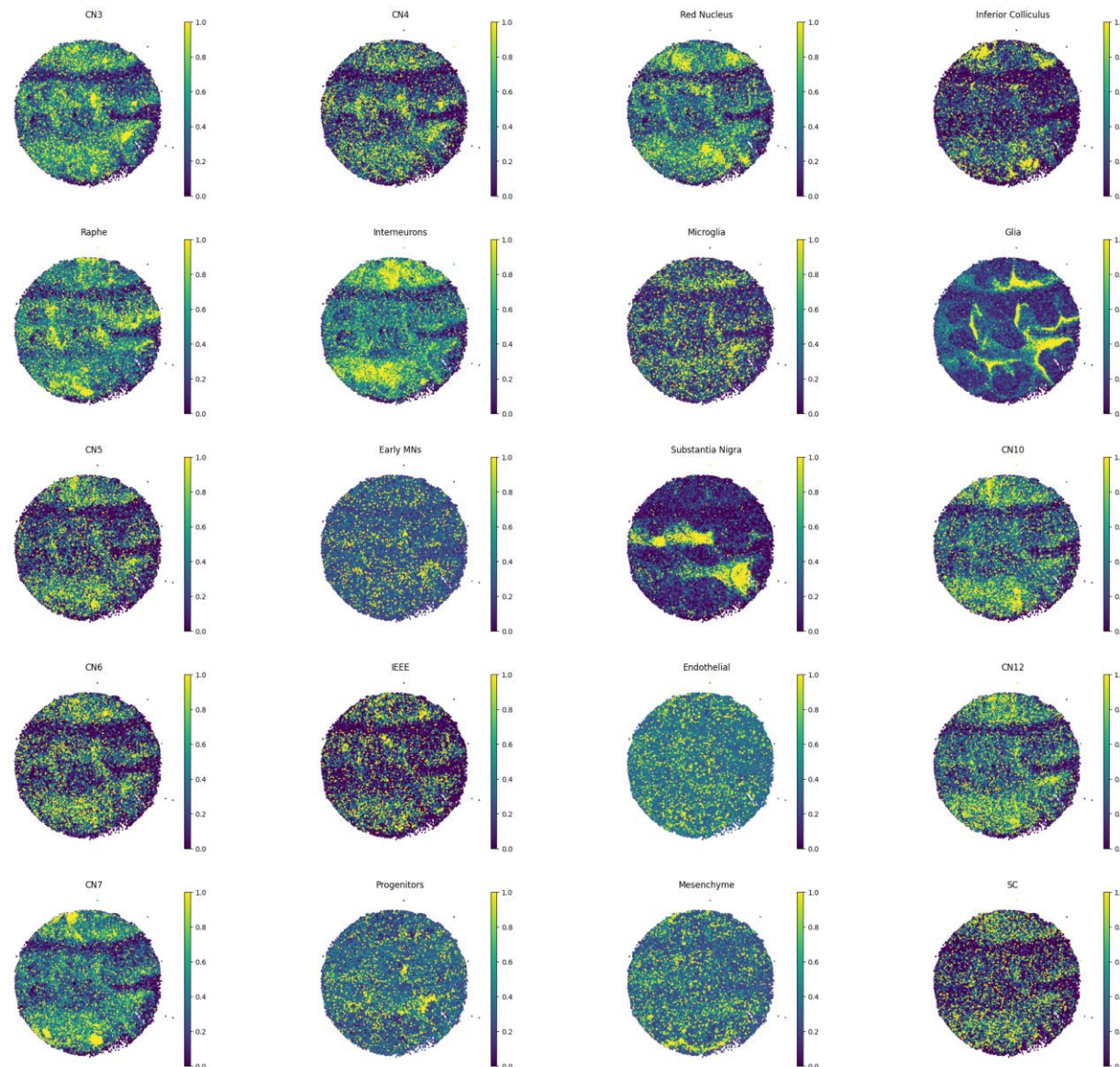
---





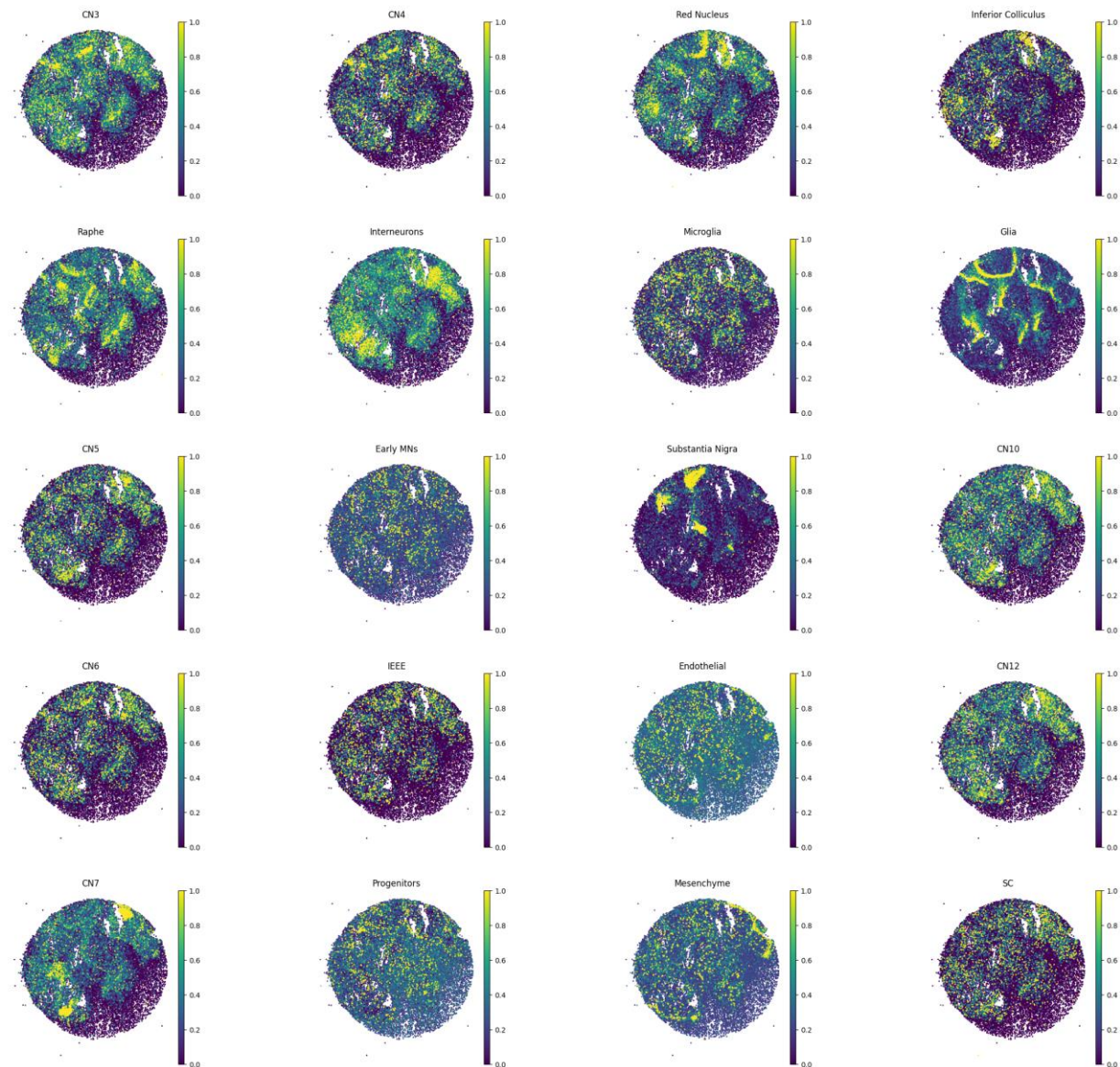
e14  
mapped to  
puck 08

---



e14  
mapped to  
puck 09

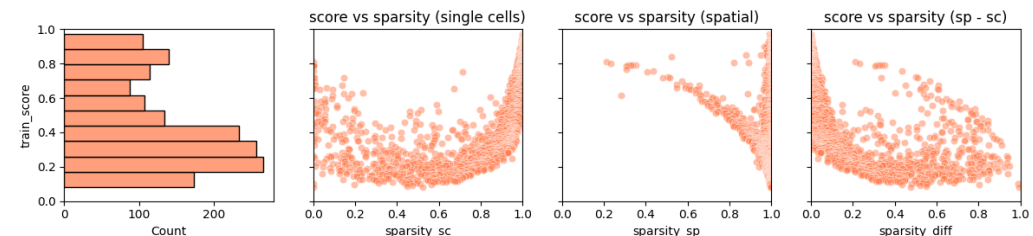
---



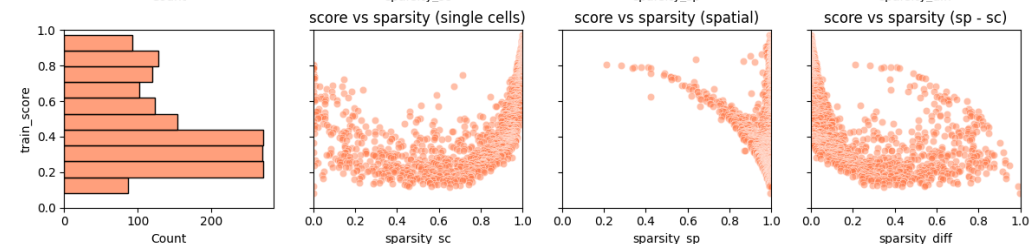


# Training scores

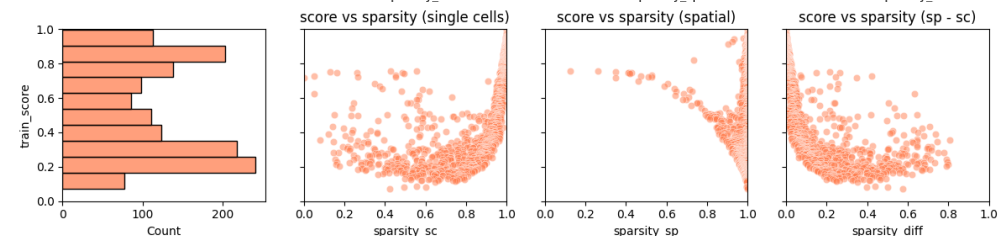
Puck 03



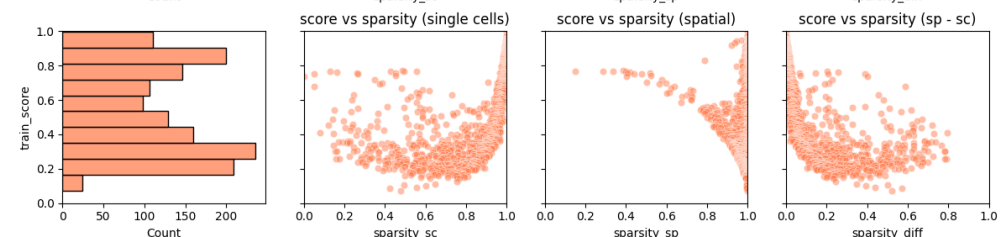
Puck 05



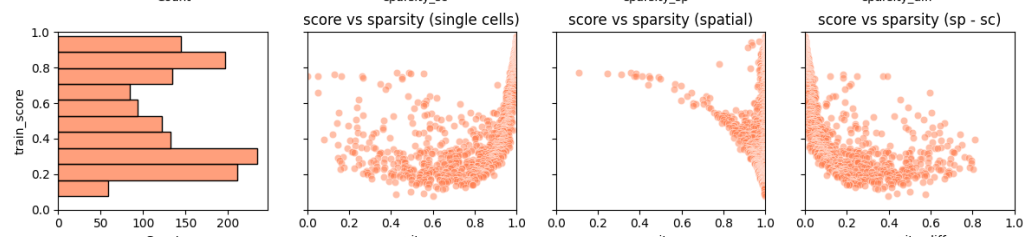
Puck 06



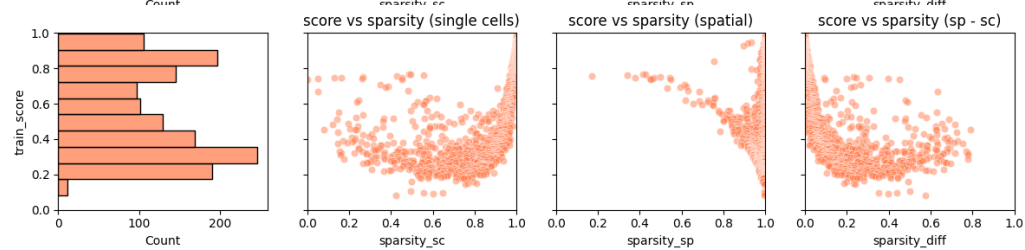
Puck 07



Puck 08



Puck 09



# Tangram issues and to-dos



## High memory usage

Need to run on GPU



## Fine tune parameters

Trained on whole transcriptome of the single cell data

Create list of top 200-600 most variable genes