

# Single-Cell Environment and Proximal Trajectory Inference using Collaborative Reinforcement Learning (scEPTIC-RL)

Noah Cohen Kalafut<sup>1,2</sup> and Daifeng Wang<sup>1,2,3,\*</sup><sup>1</sup>Department of Computer Sciences, Wisconsin, US. <sup>2</sup>Waisman Center, University of Wisconsin-Madison, Wisconsin, US. <sup>3</sup>Department of Biostatistics and Medical Informatics, Wisconsin, US

## Abstract

Recent techniques enable functional characterization of single-cells, which allows the study of cellular and molecular mechanisms in complex biological processes including cell development. Many methods have been developed to utilize these single-cell datasets to reveal cell developmental trajectories such as dimensionality reduction and pseudotime. However, these methods generally produce static snapshots of the data, challenging a deeper understanding of the mechanistic dynamics underlying cell development.

To address this, we have developed scEPTIC-RL (single-cell Environment and Proximal Trajectory Inference using Collaborative Reinforcement Learning), a multi-agent reinforcement learning model to recapitulate the dynamic progression of cells during development. scEPTIC-RL takes single-cell data, either single or multimodality, and trains a collaborative reinforcement learning model that governs cell-cell dynamic interactions driving development. Particularly, it models single cells as individual agents which coordinate progression on a latent space through interacting with neighboring cells. The trained model can further prioritize cellular features and in-silico predict the dependencies of cell development from feature perturbations (e.g., gene knockout). We apply scEPTIC-RL to both simulation and real-word single-cell multiomics datasets including brain development and cancers, revealing potential novel mechanistic insights on gene expression and regulation in those complex developmental processes.



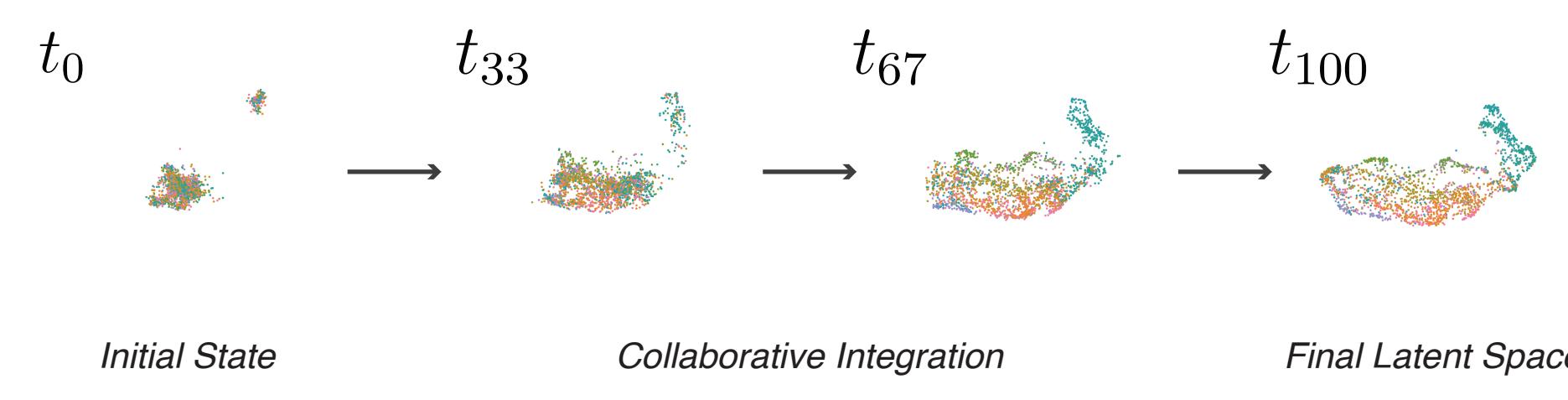
## Datasets

Citation	Dataset Description	Cells	Features
Liu et al. [1]	Simulation data consisting of 3 distinct cell types	300	RNA-seq 1,000, ATAC-seq 2,000
*Trevino et al. [2]	Developing human brain data over 21 post-conceptual weeks	8,981	34,104, 19,836
*Cao and Gao [3]	Single-nucleus profiling of the adult mouse cortex	9,190	28,930, 241,757

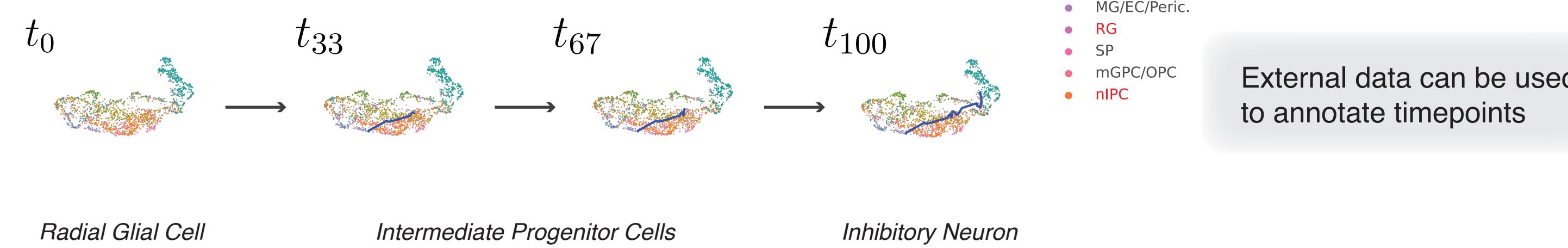
\*Features were preprocessed using 512 PCA features before input to scEPTIC-RL

## Recovery of Human Brain Developmental Trajectories

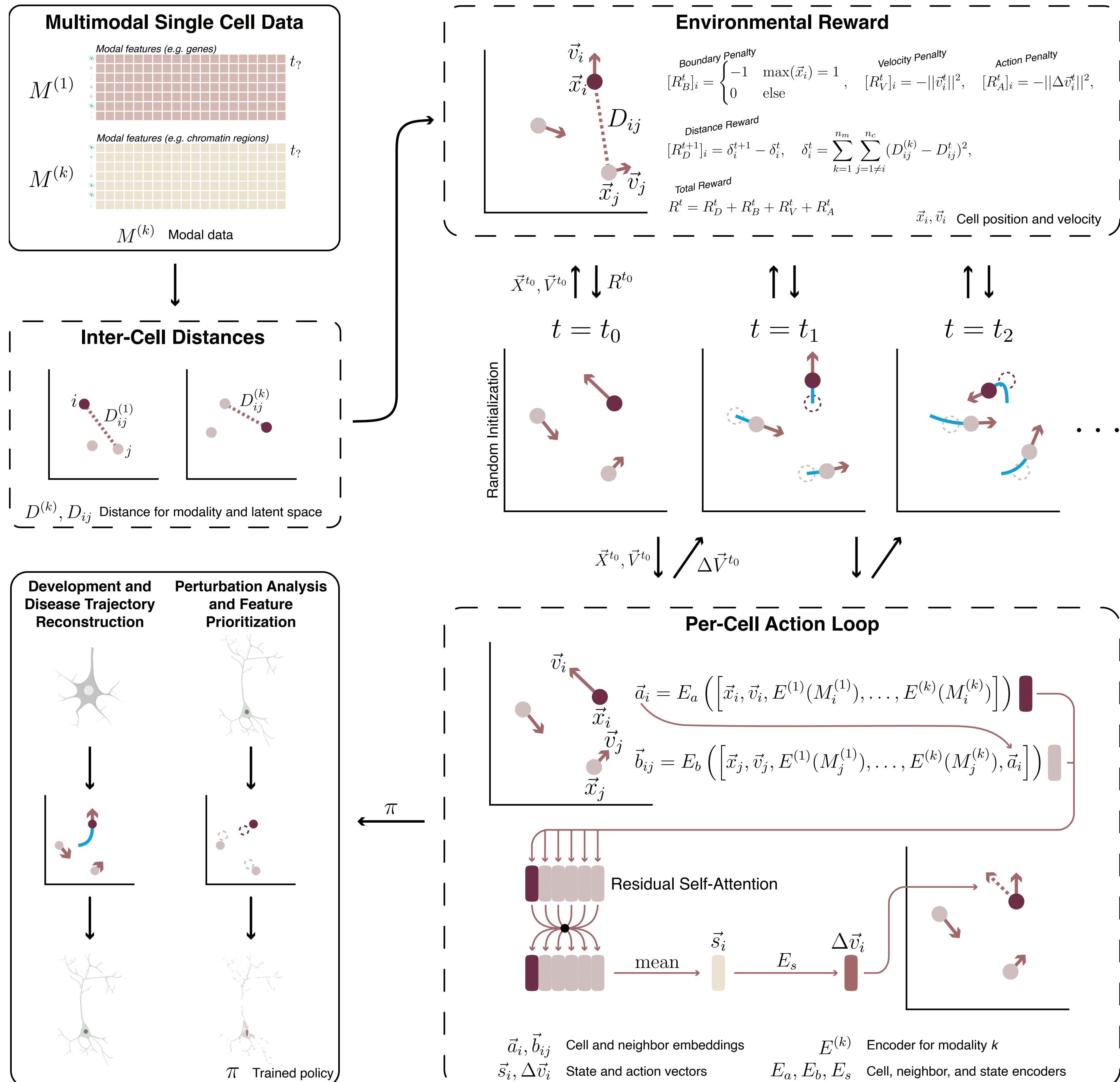
### Latent Space Formulation



### Developmental Trajectory Recovery



## scEPTIC-RL Recovers Developmental Trajectory of Single Cells



## Algorithm

### Algorithm 1 scEPTIC-RL

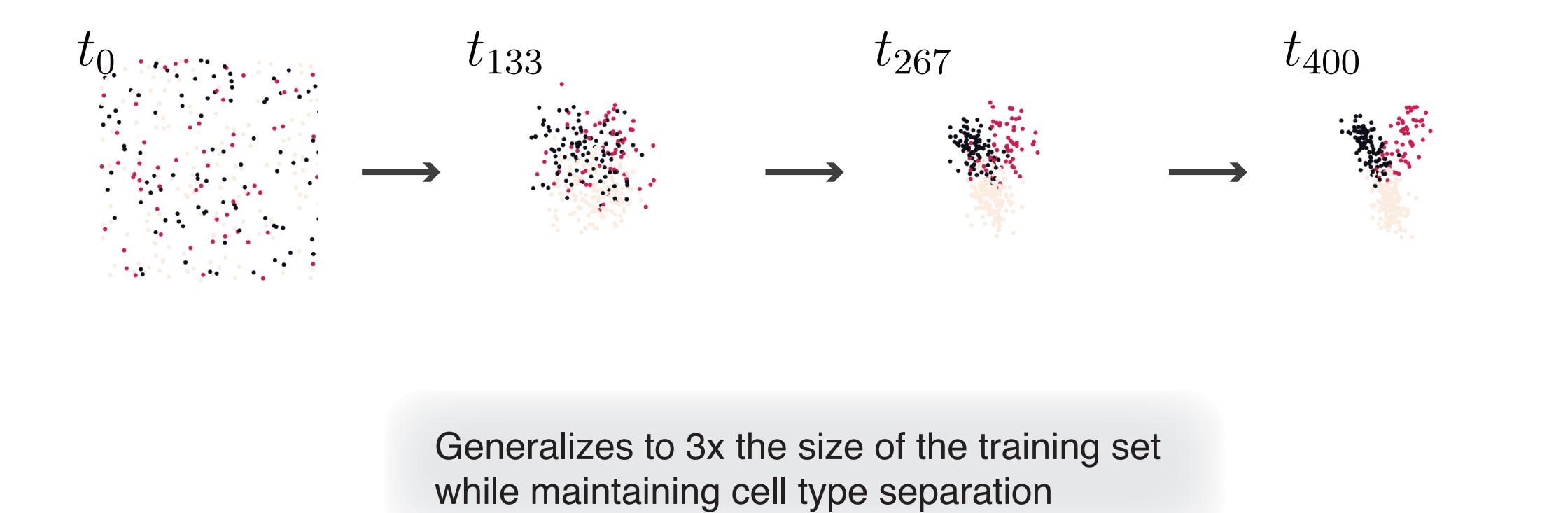
```

Input:  $M^{(1)}, M^{(2)}, \dots, M^{(n_m)}$ 
Parameters: #max_timesteps, #update_timesteps, #epochs,  $\gamma$ 
Output:  $\pi_\theta$ 
1: Randomly initialize environment  $e$  and model  $\pi_\theta$ 
2:  $t \leftarrow 1$ 
3: for #max_timesteps do
4:    $s^t \leftarrow e$ 's state
5:    $a^t \sim \pi_\theta(s^t)$ 
6:    $R^t, \text{terminal} \leftarrow$  perform actions  $a^t$  on environment  $e$ 
7:   Record  $(s^t, a^t, R^t)$  into memory
8:   if #update_timesteps | t then
9:     Compute episode rewards using  $\gamma$  decay
10:    for #epochs do
11:      Compute  $L_{\text{CLIP}}(\theta)$  using Proximal Policy Optimization
12:      Backpropagate on  $\pi_\theta$  from memory
13:    end for
14:   end if
15:   if terminal then
16:     Reset environment
17:   end if
18:    $t \leftarrow t + 1$ 
19: end for

```

$$L_{\text{CLIP}}(\theta) = \text{MIN} \left( \frac{\pi_\theta(a^t|s^t)}{\pi_{\theta_{\text{old}}}(a^t|s^t)} \hat{A}^t, \text{CLIP} \left( \frac{\pi_\theta(a^t|s^t)}{\pi_{\theta_{\text{old}}}(a^t|s^t)}, 1 - \epsilon, 1 + \epsilon \right) \hat{A}^t \right)$$

## Simulation



## Conclusions

- ★ Easily distributable and highly generalizable
- ★ Reliable trajectory reconstructions for cell development and disease progression
- ★ Future applications include extension to further datasets as well as imputation of spatial data.

## Acknowledgements and Contact

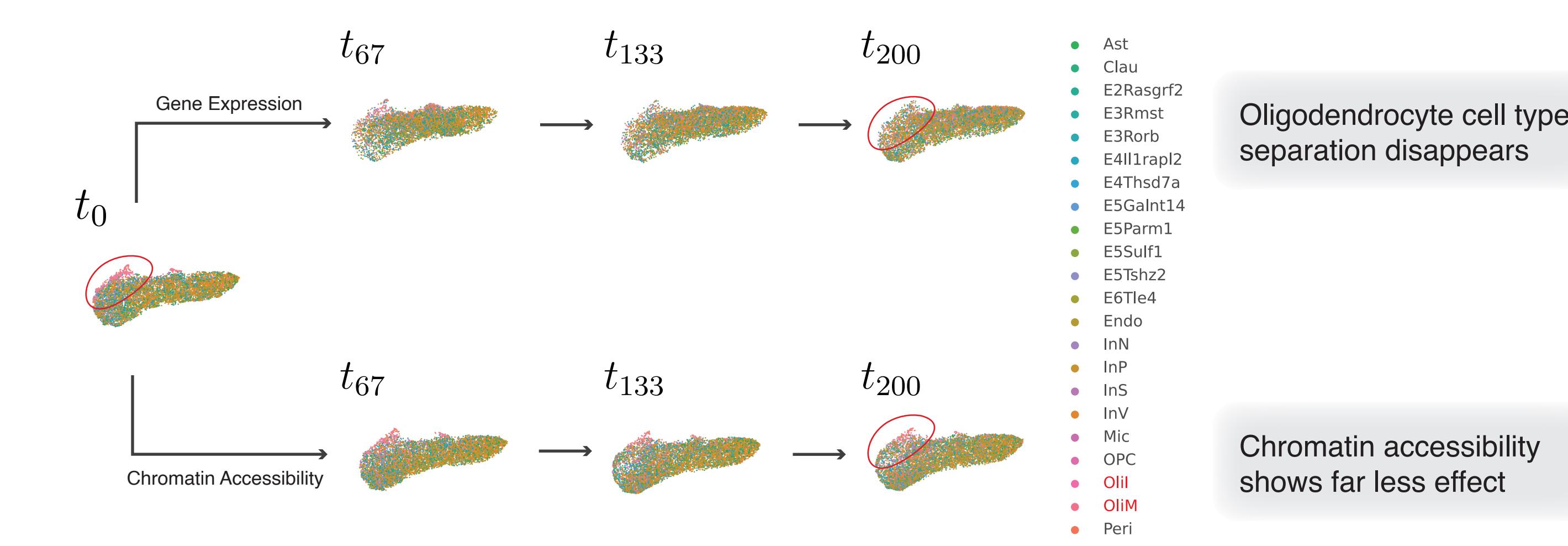
This work is supported by National Institutes of Health grants, RF1MH128695, R01AG067025, R21NS128761, R21NS127432, and National Science Foundation Career Award 2144475.



<https://daifengwanglab.org/>  
nkalafut@wisc.edu  
daifeng.wang@wisc.edu

## Feature Prioritization in the Adult Mouse Cortex

### RNA Perturbation



### ATAC Perturbation

A note says: 'Chromatin accessibility shows far less effect'

[1] Jie Liu, et al. "Jointly Embedding Multiple Single-Cell Omics Measurements". en: In: Algorithms Bioinform 143 (Sept. 2019).  
[2] Alejandro E. Trevino, et al. "Chromatin and gene-regulatory dynamics of the developing human cerebral cortex at single-cell resolution". In: Cell 184,19 (Sept. 2021), 505-515.e22. issn: 0092-8674. doi: 10.1016/j.cell.2021.07.039.  
[3] Zhi-Jie Cao and Ge Gao. "Multi-omics single-cell data integration and regulatory inference with graph-linked embedding". In: Nature Biotechnology 40,10 (Oct. 2022), pp. 1458-1466. issn: 1546-1696. doi: 10.1038/s41587-022-01284-4. url: https://doi.org/10.1038/s41587-022-01284-4.