

Trajectory inference

Mohammed Charrouf

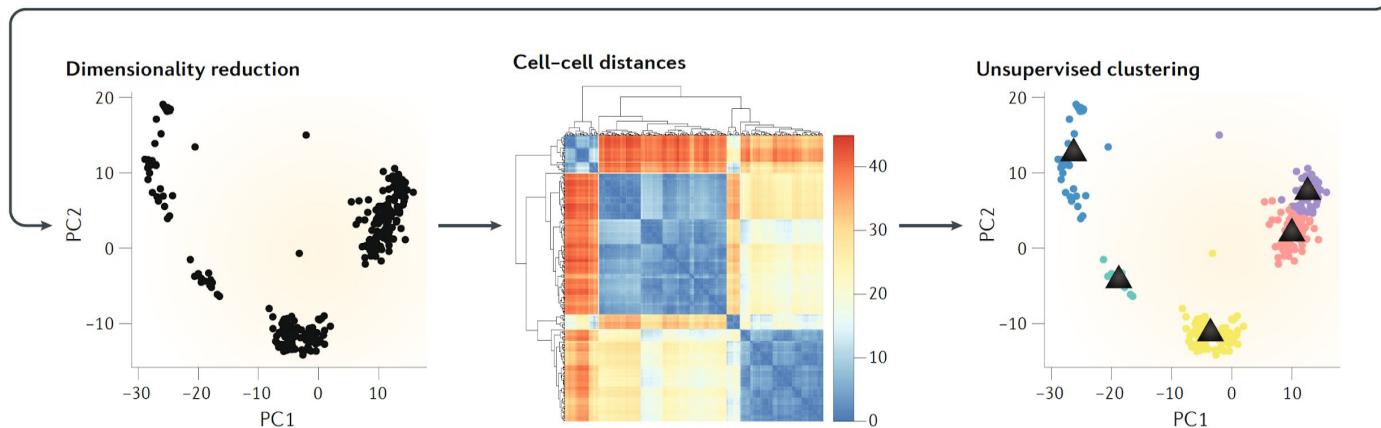
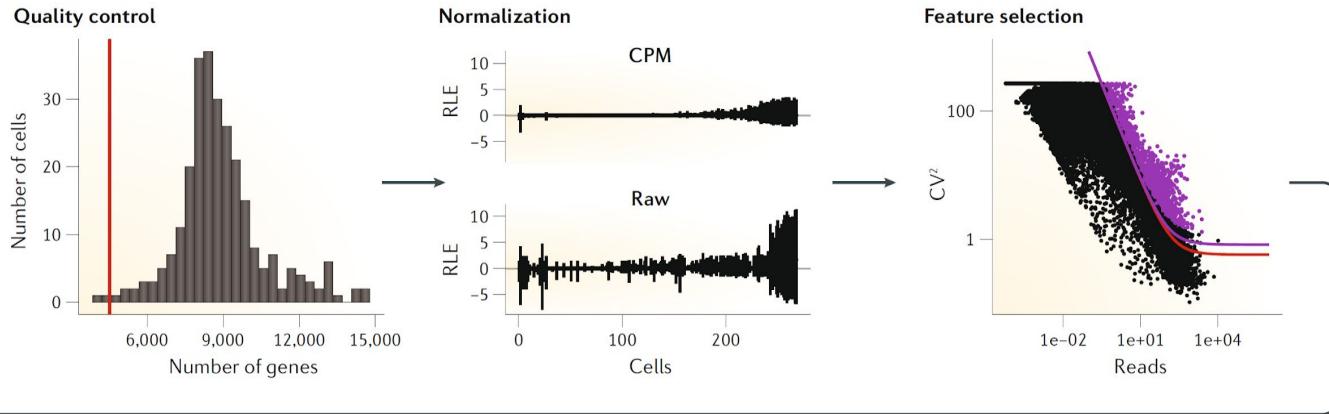
Leiden Computational Biology Center, LUMC
Delft Bioinformatics Lab, TU Delft



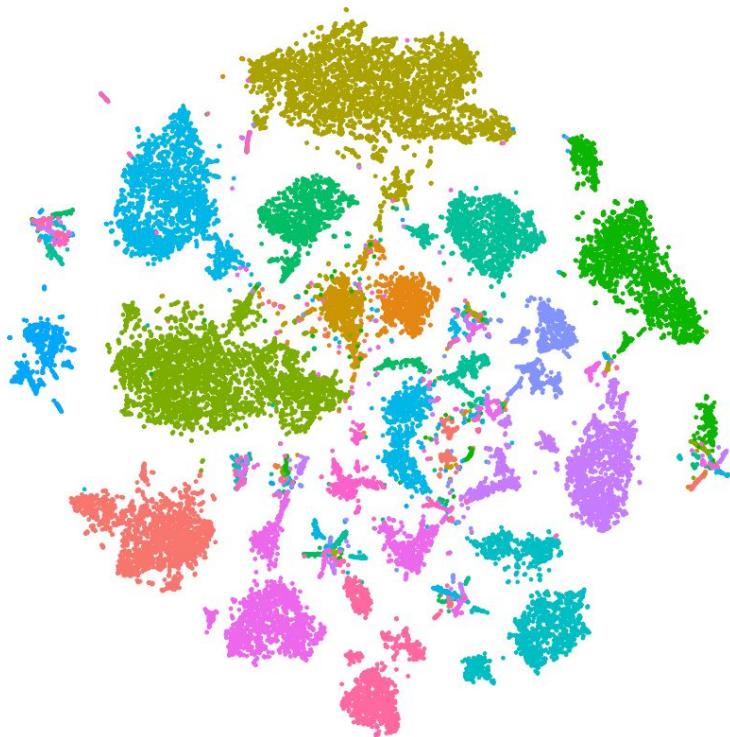
Leiden
Computational Biology Center



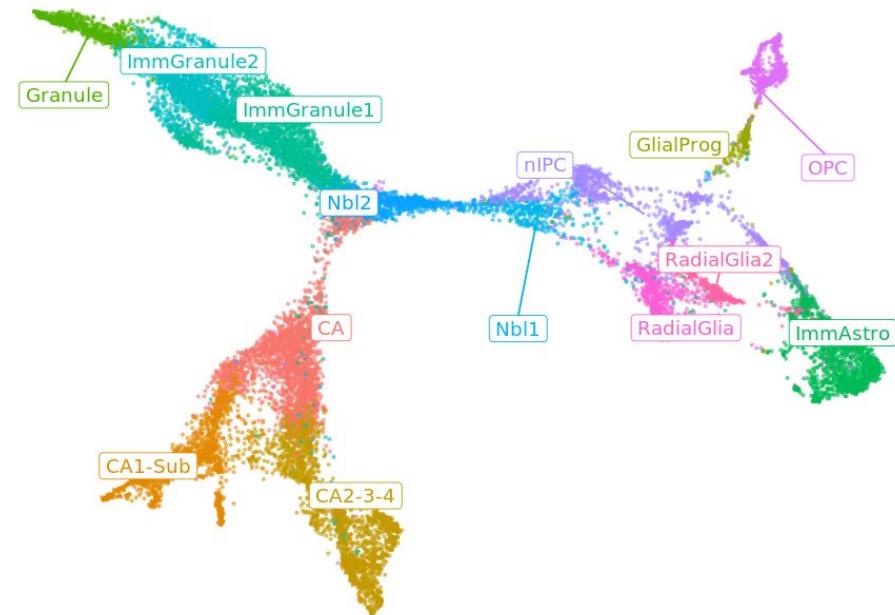
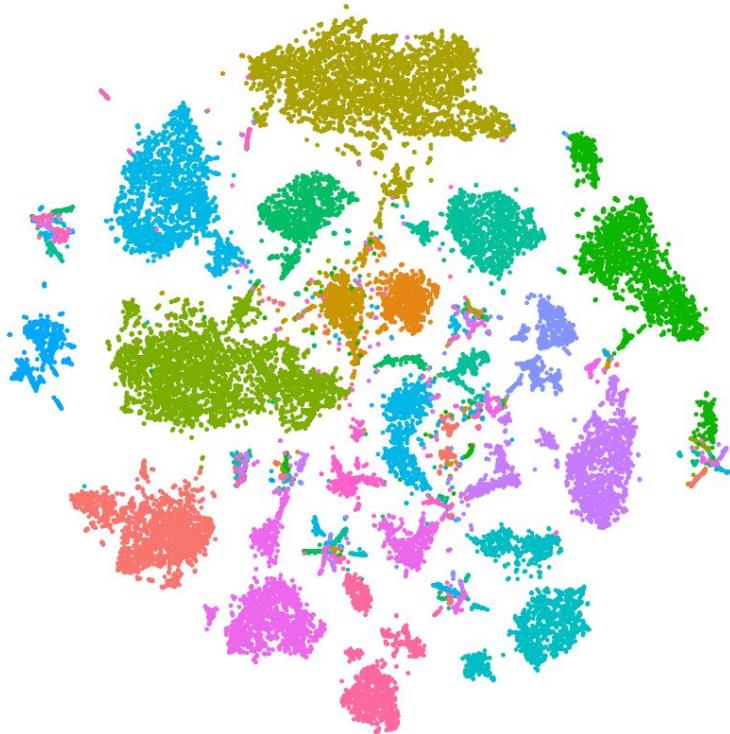
Single cell RNA-seq workflow



Clustering of differentiating cells

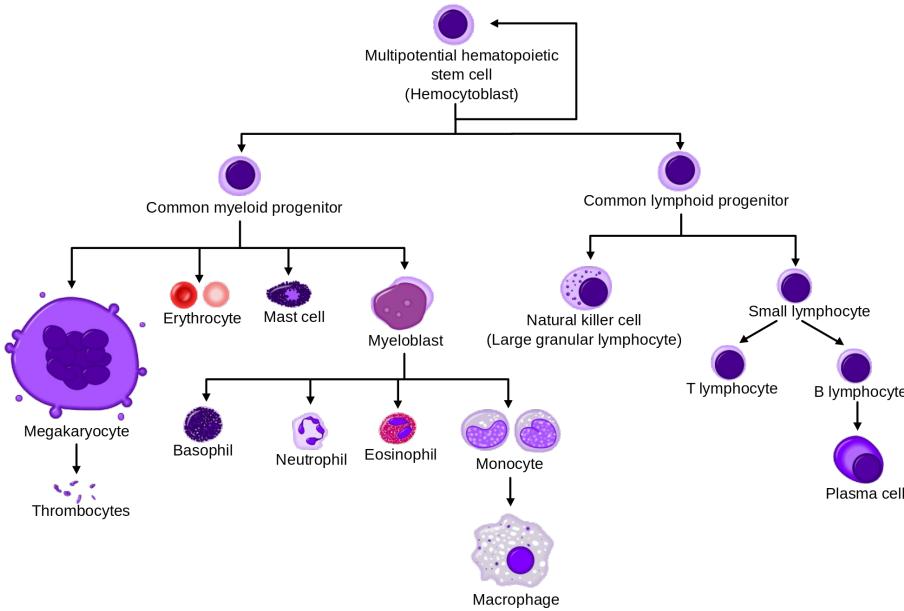


Clustering of differentiating cells



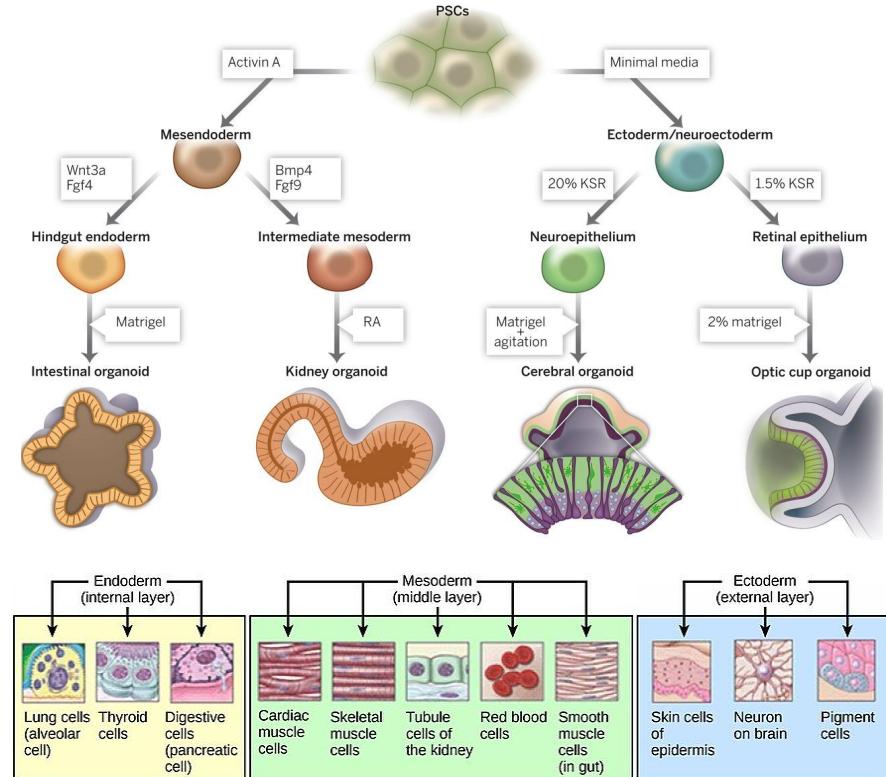
Differentiating cells in hematopoiesis

Hematopoiesis



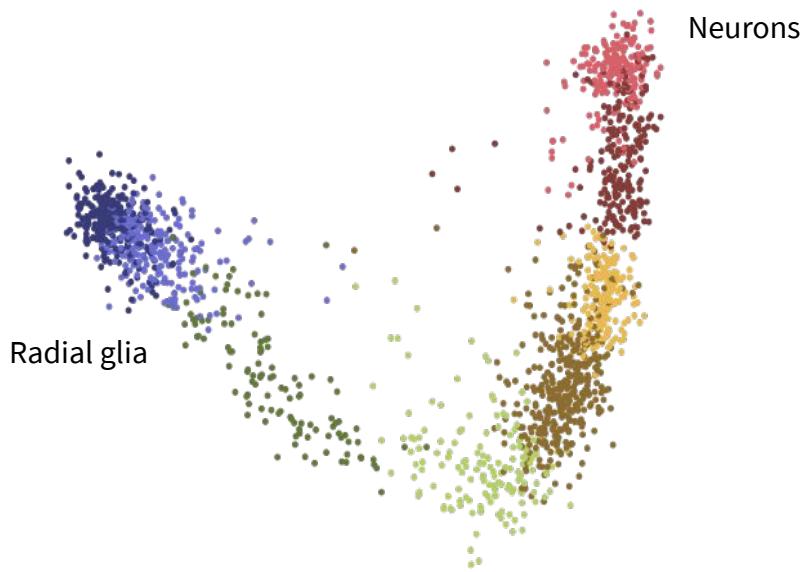
Lancaster et al. 2014

Organogenesis

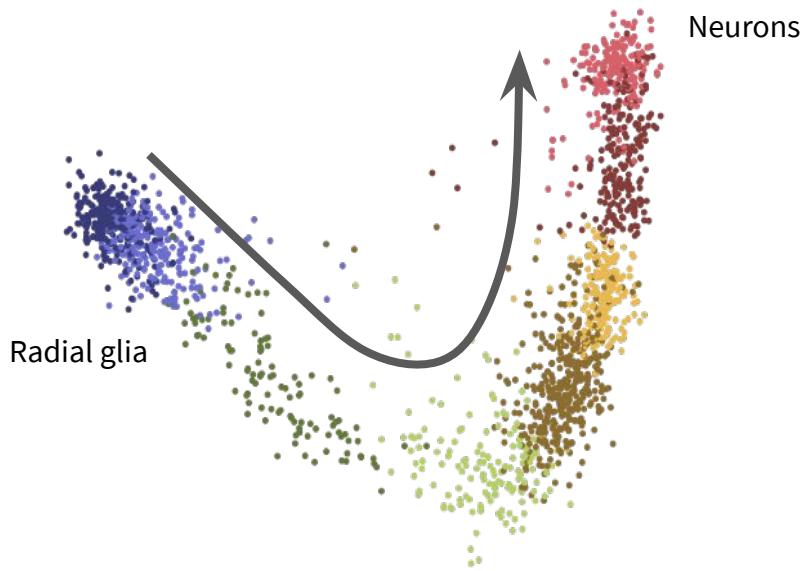


Trajectory inference

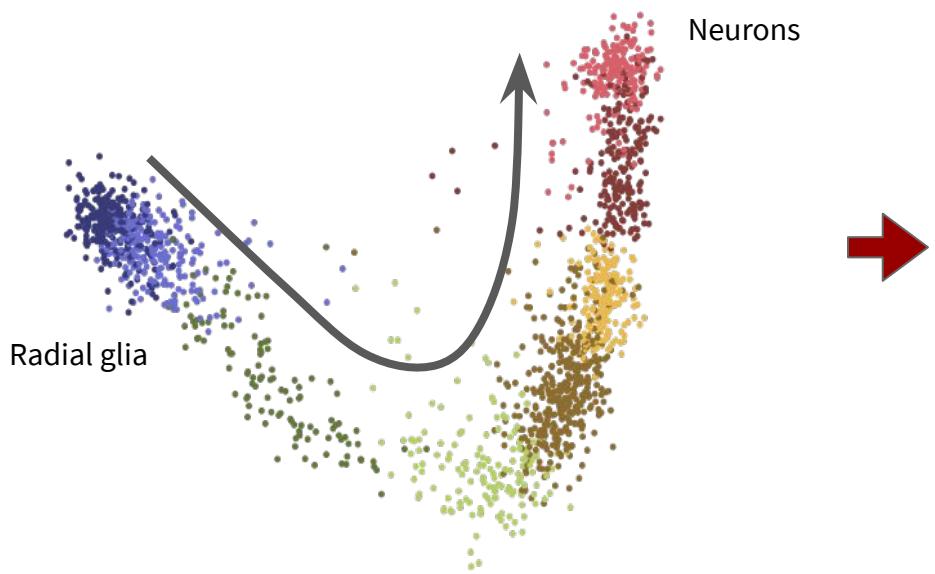
Trajectory inference / pseudotime inference



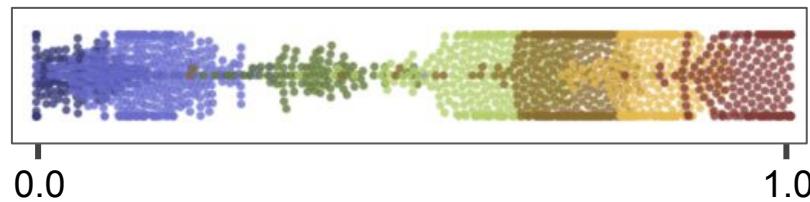
Trajectory inference / pseudotime inference



Trajectory inference / pseudotime inference

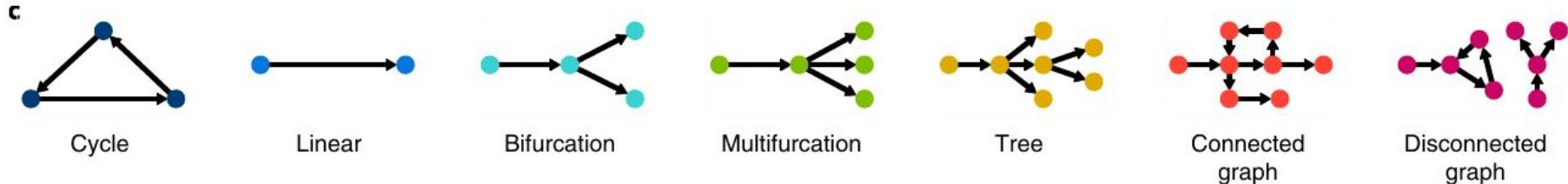


Pseudotime



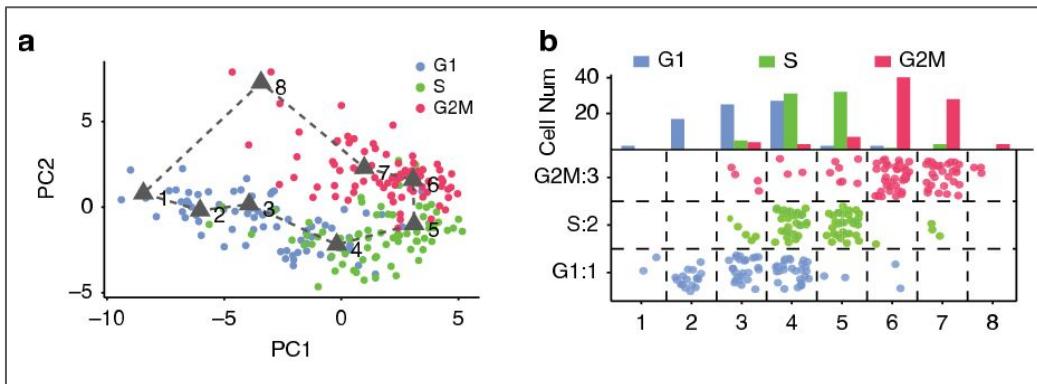
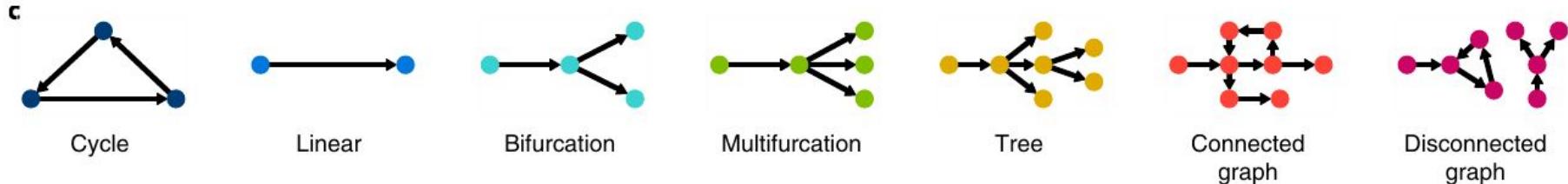
Trajectory structure

Saelens et al. (<https://doi.org/10.1038/s41587-019-0071-9>)



Trajectory structure

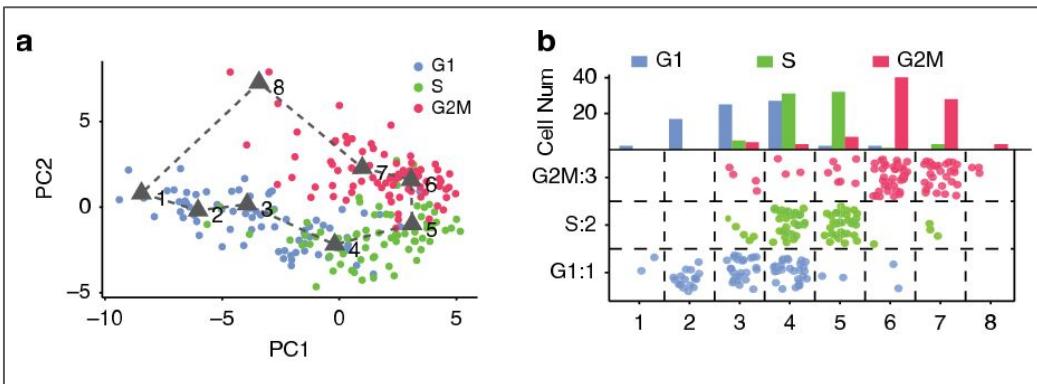
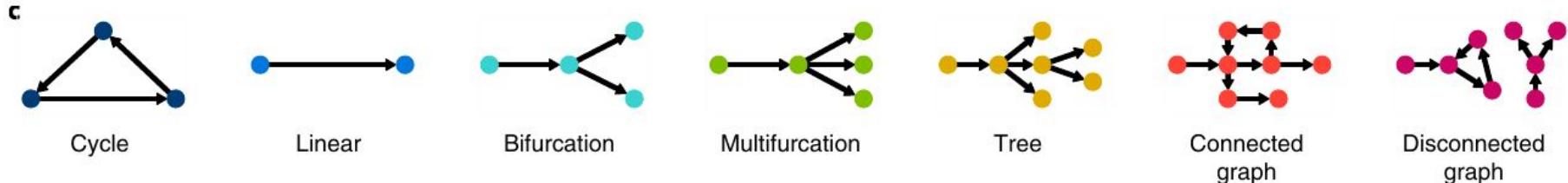
Saelens et al. (<https://doi.org/10.1038/s41587-019-0071-9>)



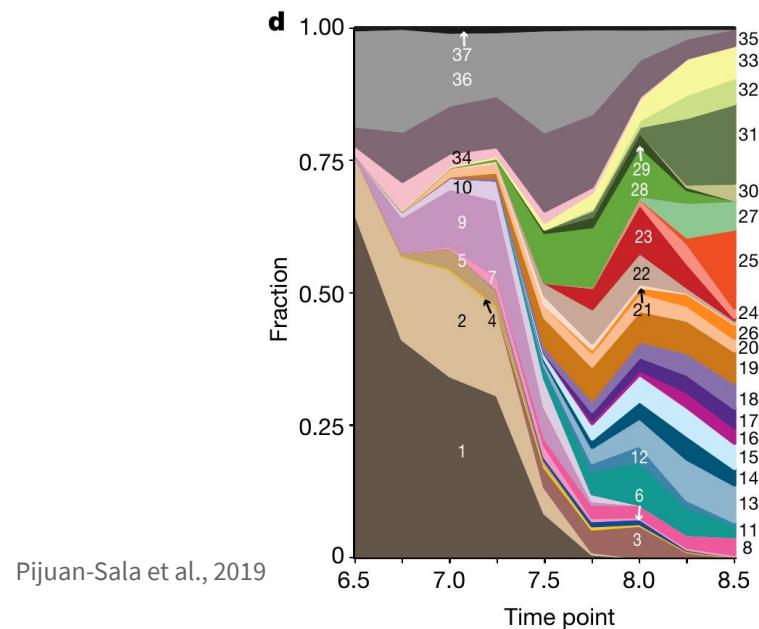
Liu et al. (<https://doi.org/10.1038/s41467-017-00039-z>)

Trajectory structure

Saelens et al. (<https://doi.org/10.1038/s41587-019-0071-9>)



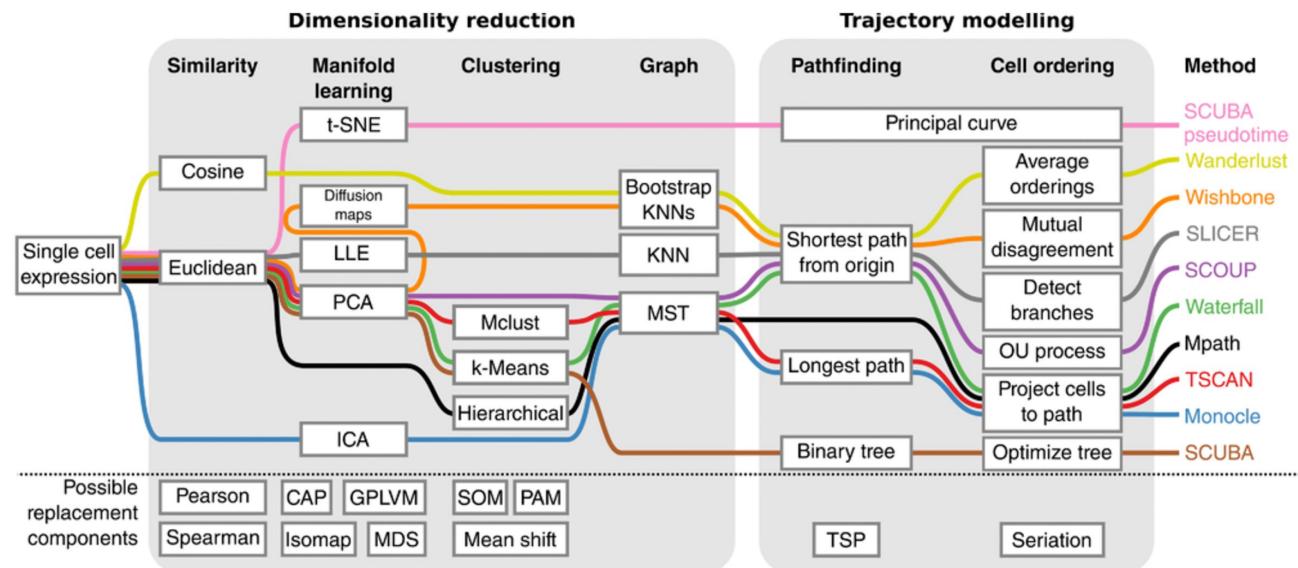
Liu et al. (<https://doi.org/10.1038/s41467-017-00039-z>)



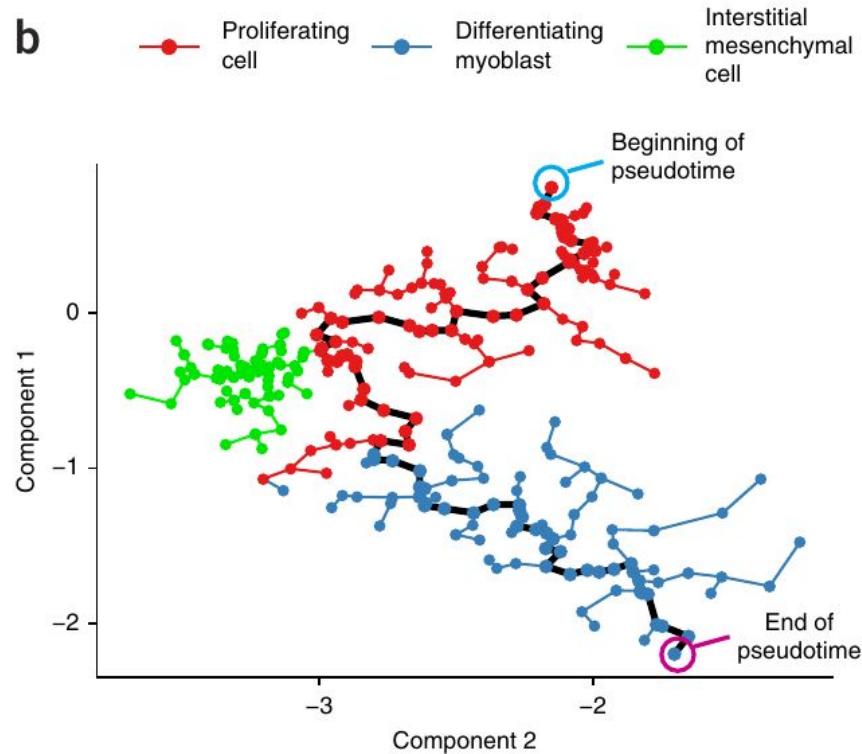
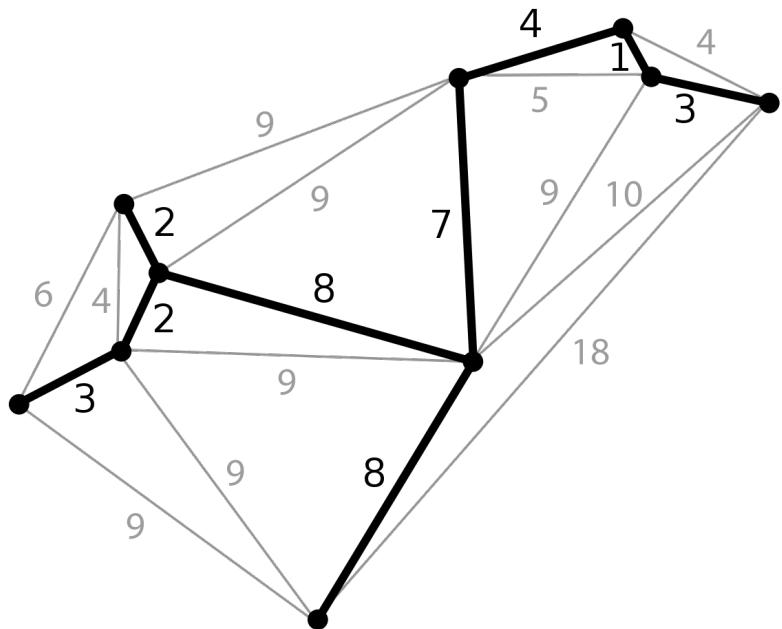
Methods

General trajectory inference pipeline

1. Dimensionality reduction
2. Trajectory fitting
3. Pseudotime assignment



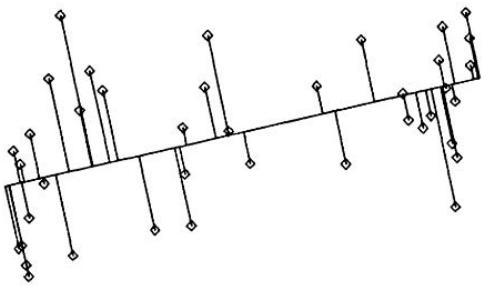
Minimum spanning trees (MST)



Trapnell et al., 2014

Principle curves

b Principal component



d Principal curve

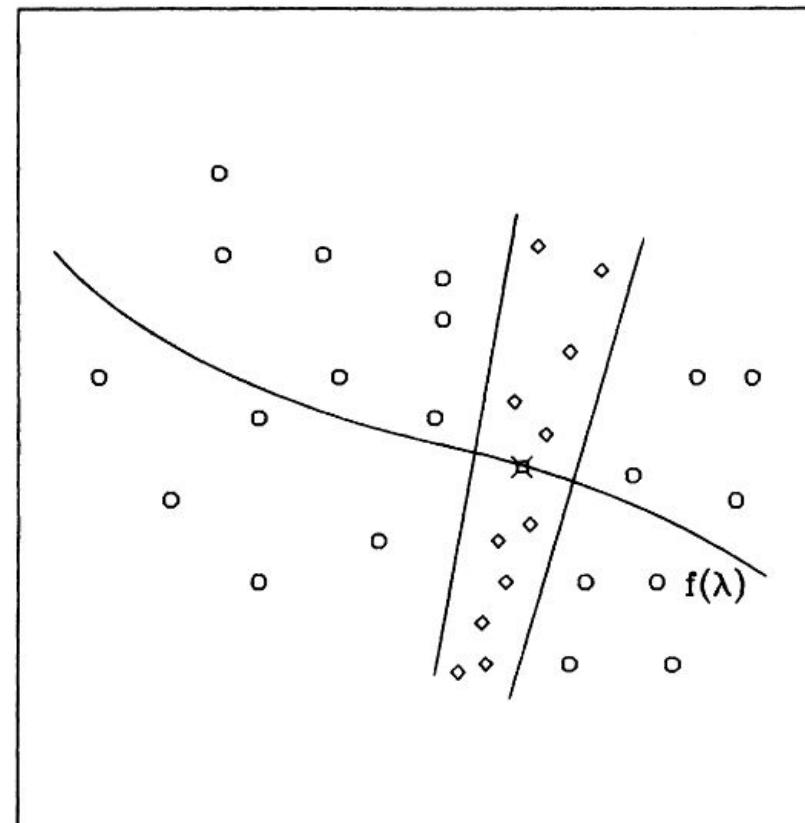
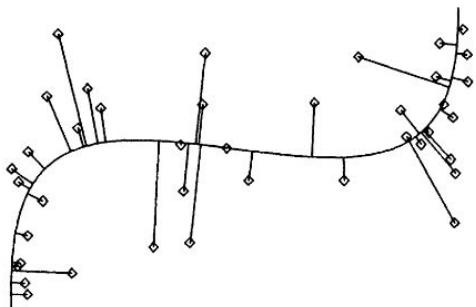


Figure 3. Each point on a principal curve is the average of the points that project there.

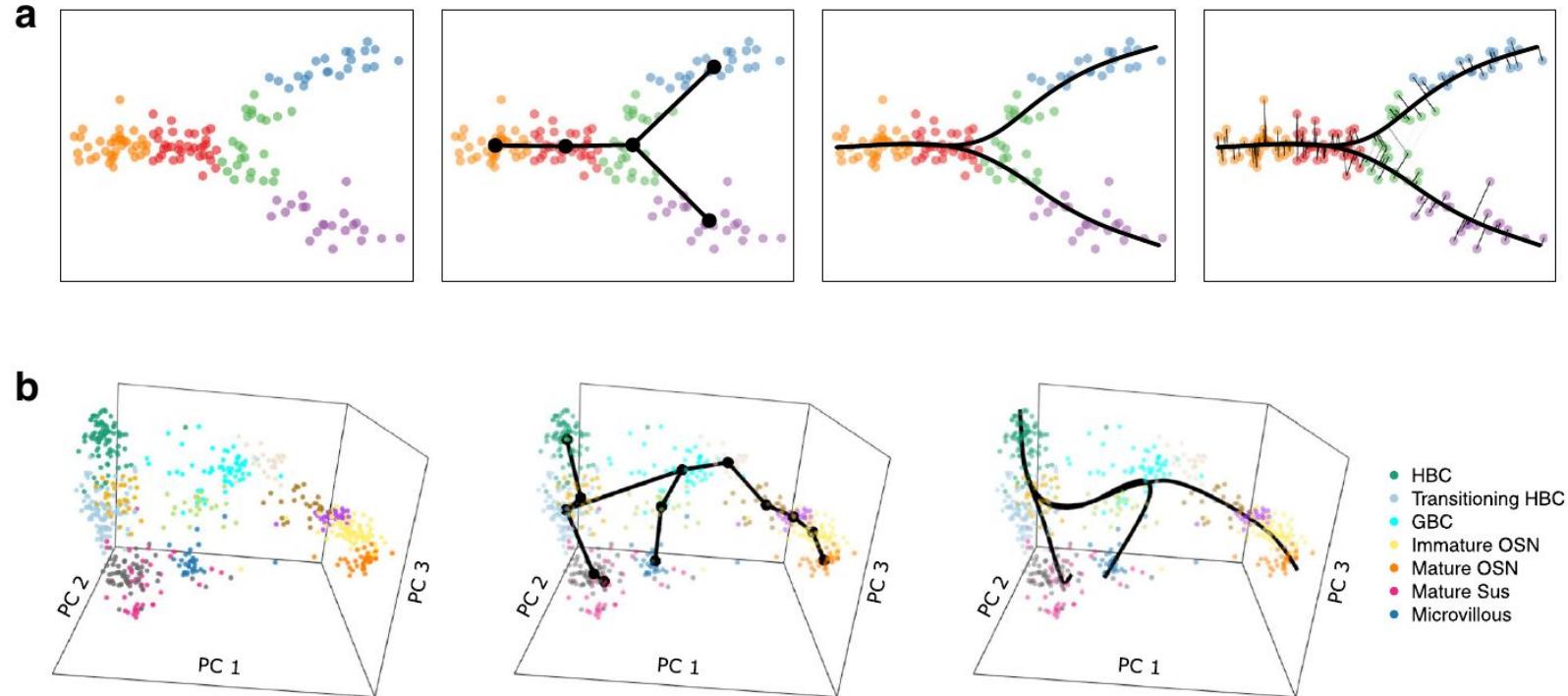
Trajectory inference methods

Three methods will be shown

- “General method”, exemplified with **Slingshot** Street et al., 2018 (doi: [10.1186/s12864-018-4772-0](https://doi.org/10.1186/s12864-018-4772-0))
- **Monocle** Cole et al., 2017 (doi: [10.1038/nmeth.4402](https://doi.org/10.1038/nmeth.4402))
- **Ouija** Campbell et al., 2018

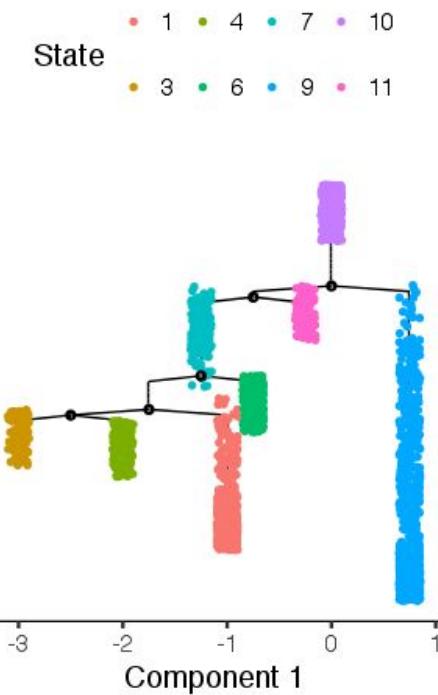
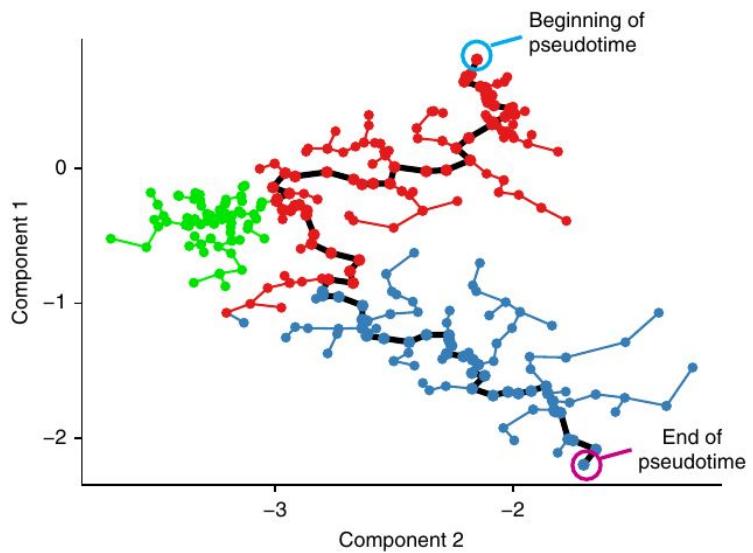
Slingshot

Fit multiple principle curves simultaneously, ensuring a shared trunk

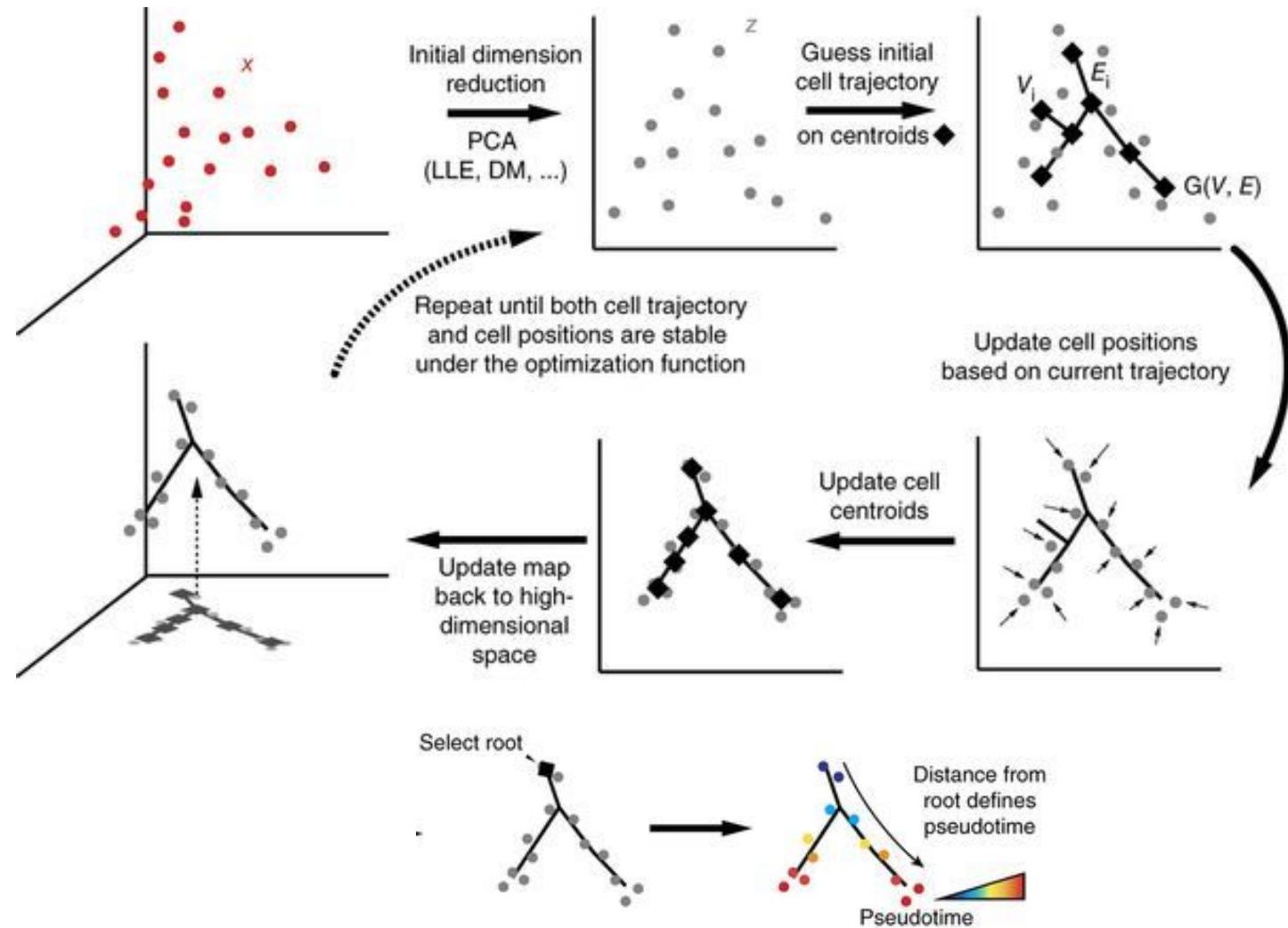


Monocle 2

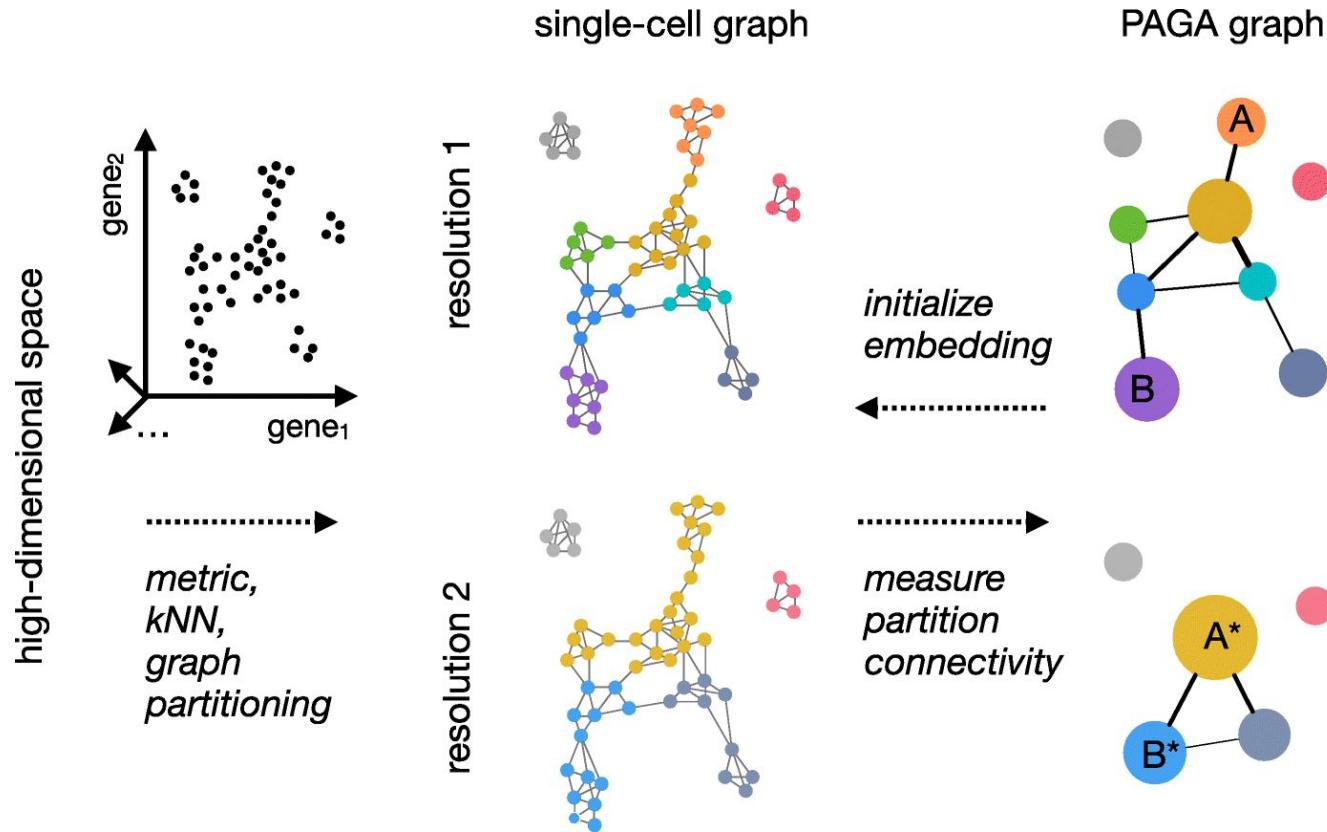
- Successor to Monocle 1
- End goal: Fit any arbitrary graph on the data



Monocle 2

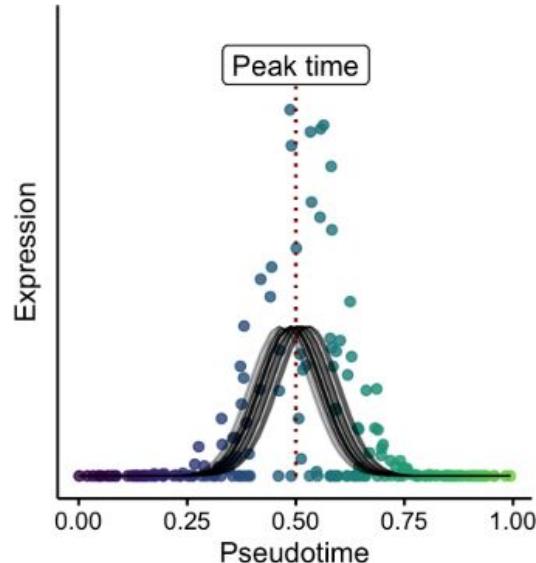
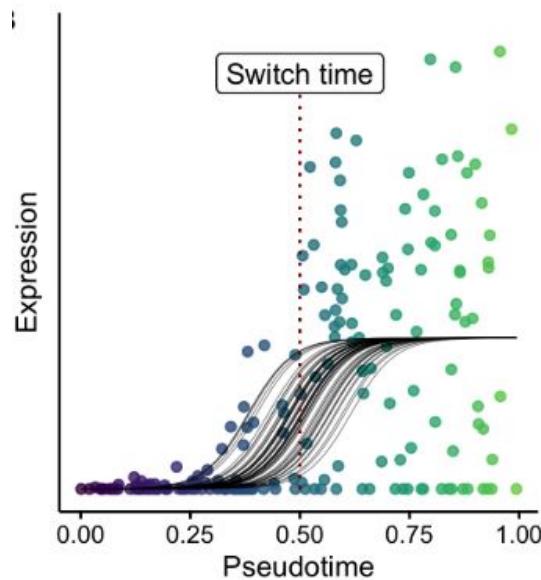


Monocle 3



Ouija

- Model a small set of marker genes instead of fitting trajectory on complete transcriptome
- Switch focus to interpretability



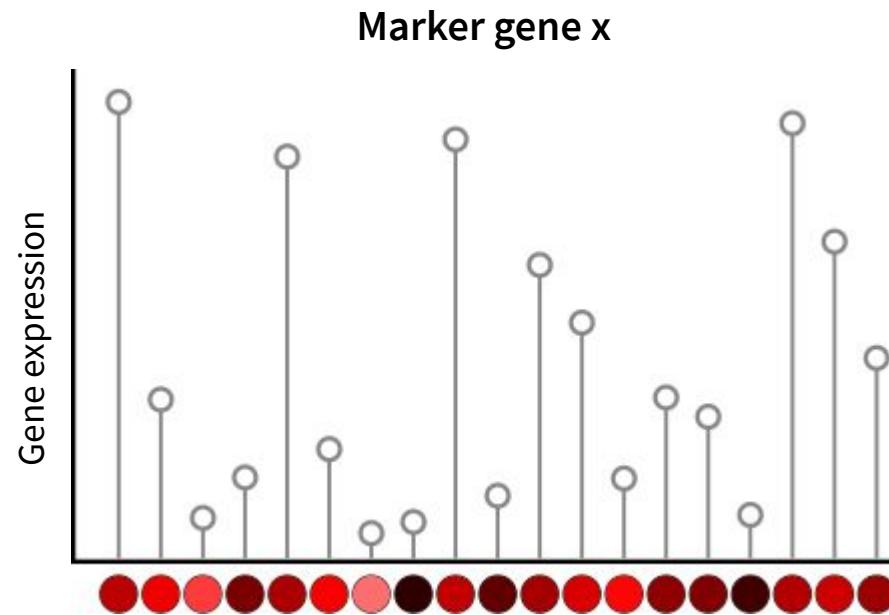
Ouija intuition

True ordering:



Random cell ordering:

Goodness-of-fit: **low**



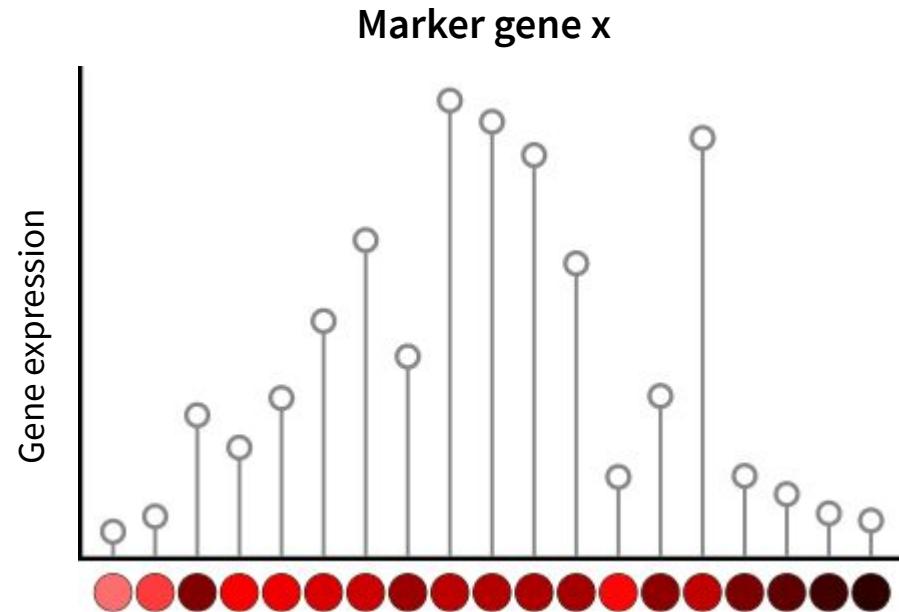
Ouija intuition

True ordering:



Optimize iteration: 100

Goodness-of-fit: mid



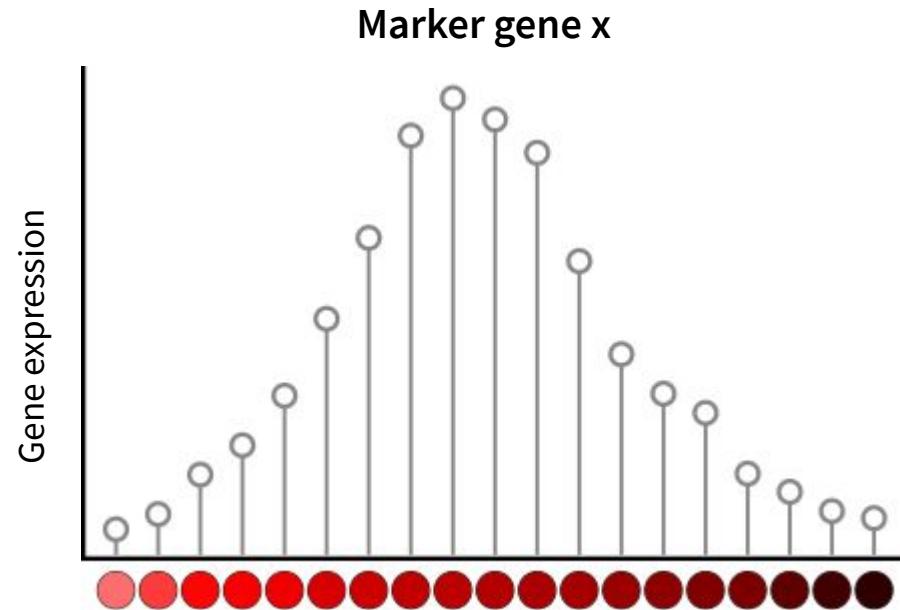
Ouija intuition

True ordering:

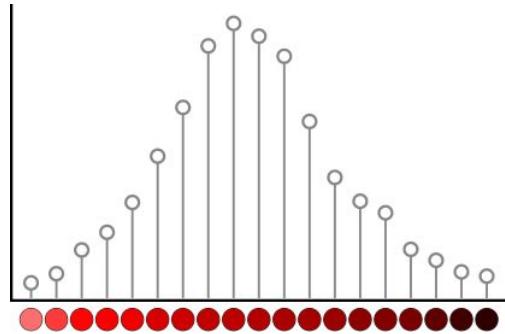


Optimize iteration: 500

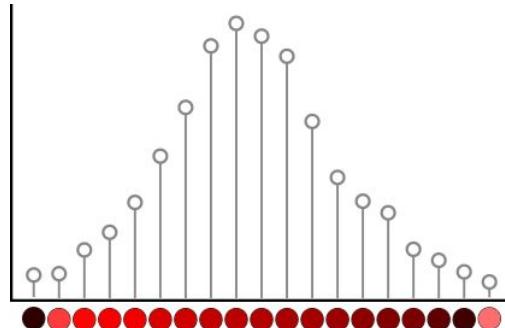
Goodness-of-fit: high



Marker gene x

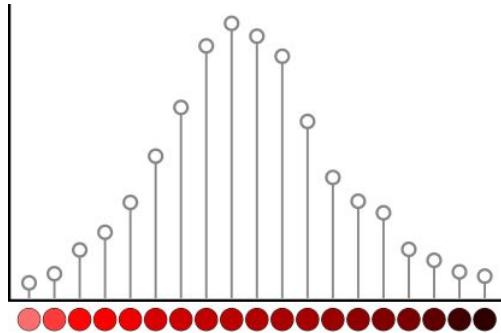


Goodness-of-fit: **high**



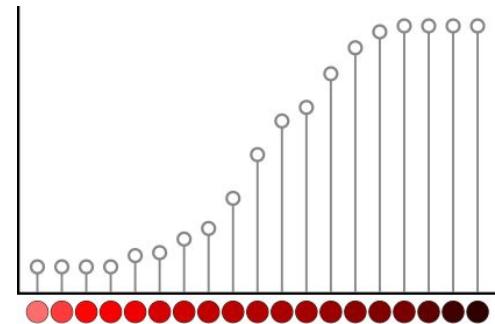
Goodness-of-fit: **high**

Marker gene x

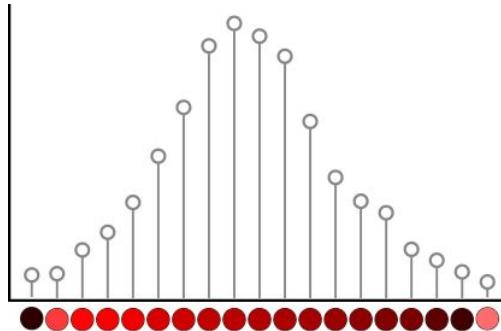


Goodness-of-fit: **high**

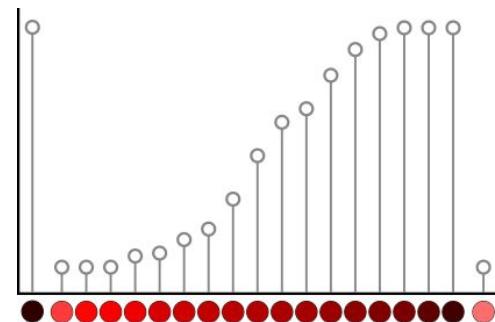
Marker gene y



Goodness-of-fit: **high**



Goodness-of-fit: **high**

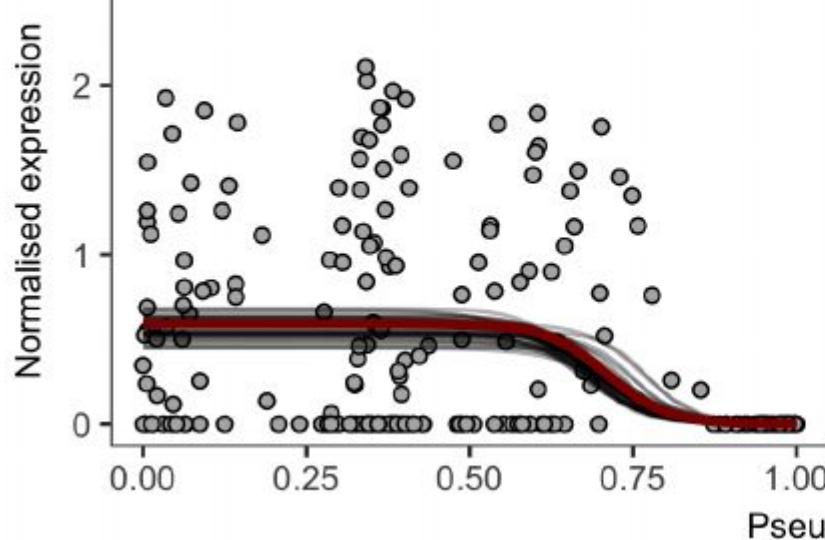


Goodness-of-fit: **mid**

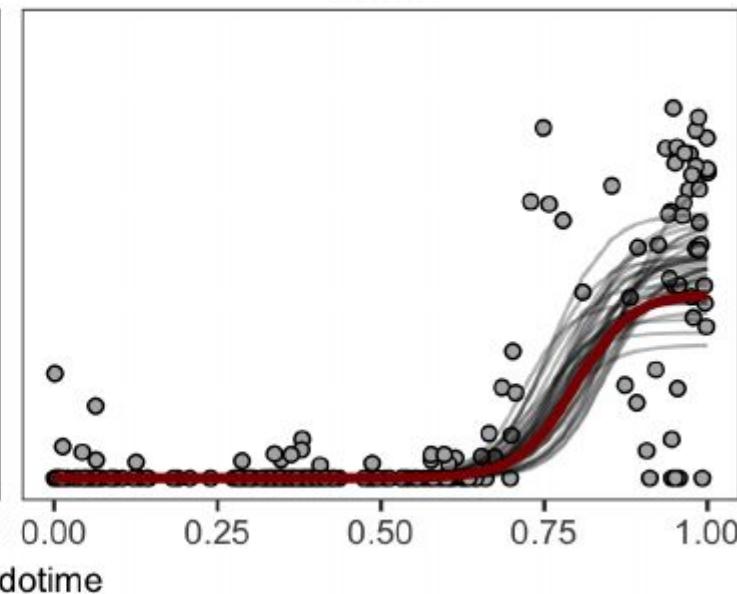
Ouija probabilistic modelling

C

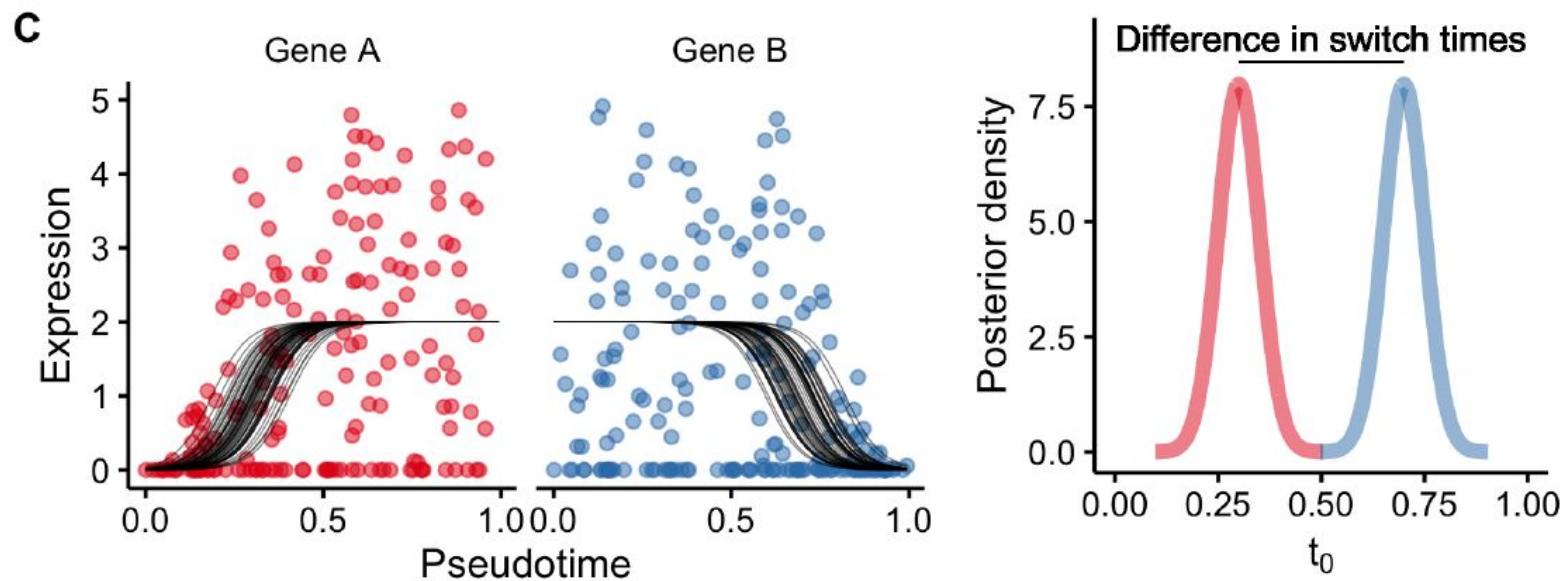
ID1



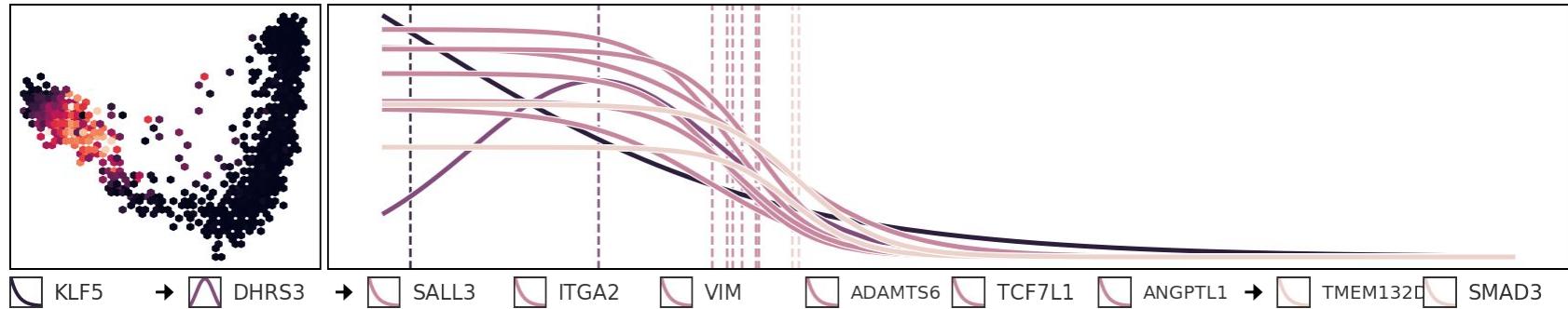
MYOG



Ouija probabilistic pseudotime assignment



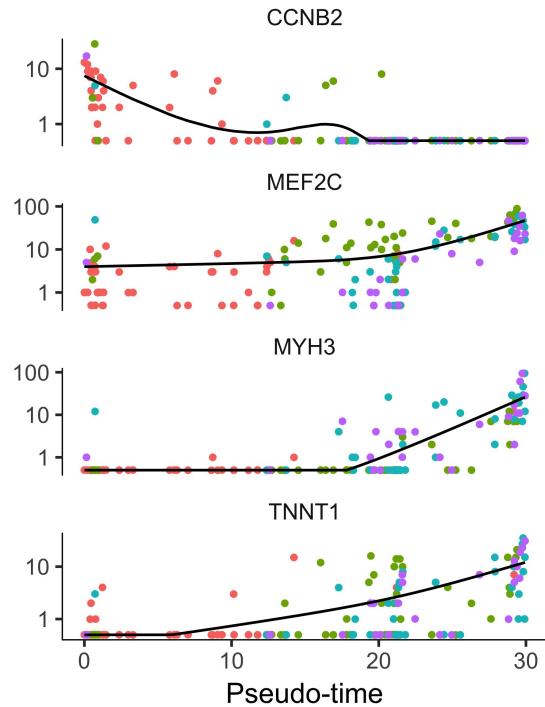
Gene ordering



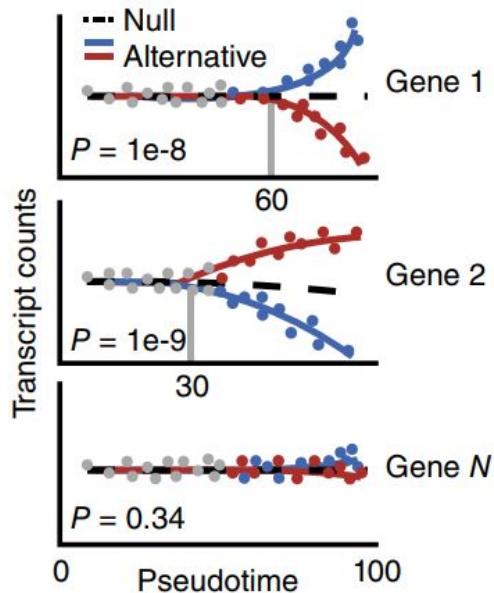
- KLF5 member of Kruppel-like family of transcription factors
 - Repressor of neurite growth, down-regulation linked to cell cycle arrest
- VIM, highly variable gene
 - Known marker of gliogenesis

Pseudotime analysis

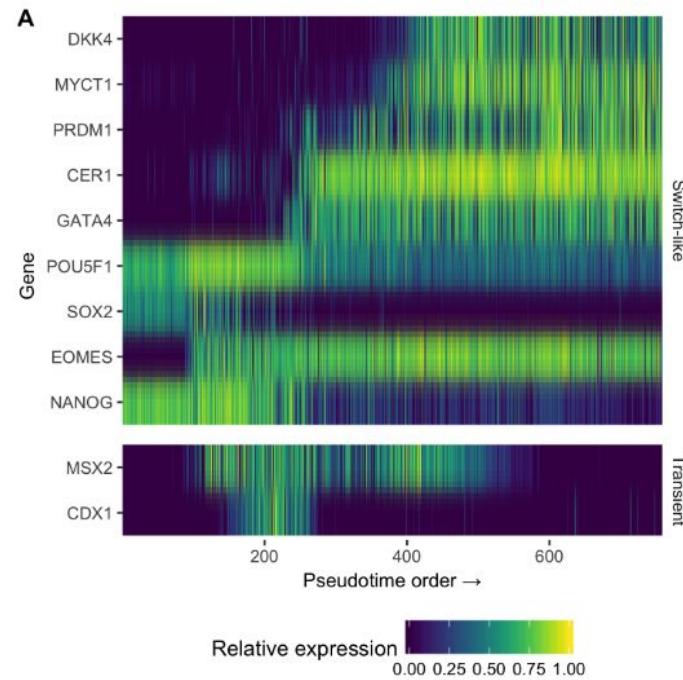
Pseudotime-gene relation



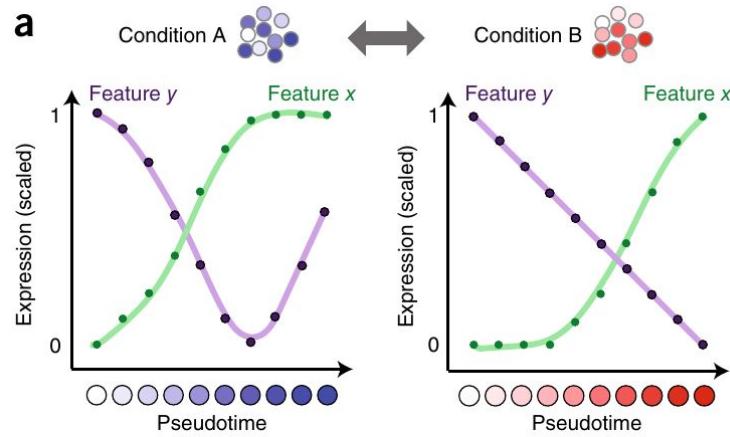
Detect branching genes



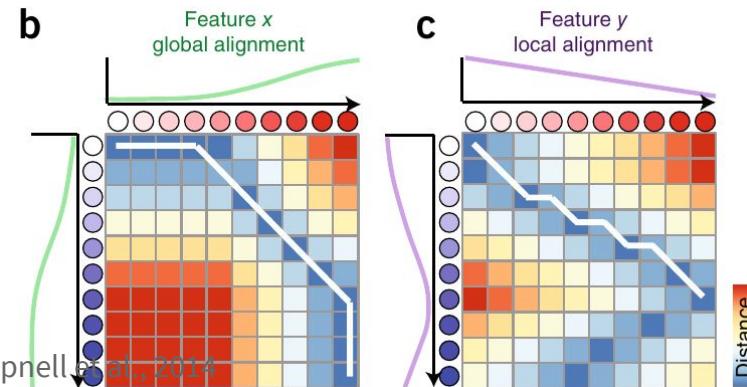
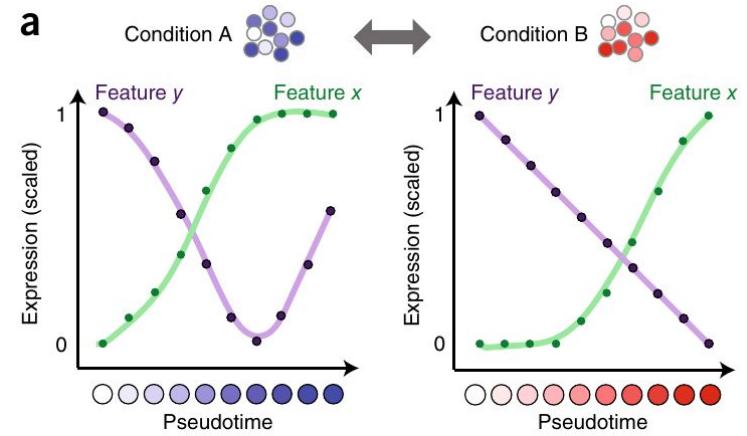
Interpretation of gene behaviour



Comparing trajectories

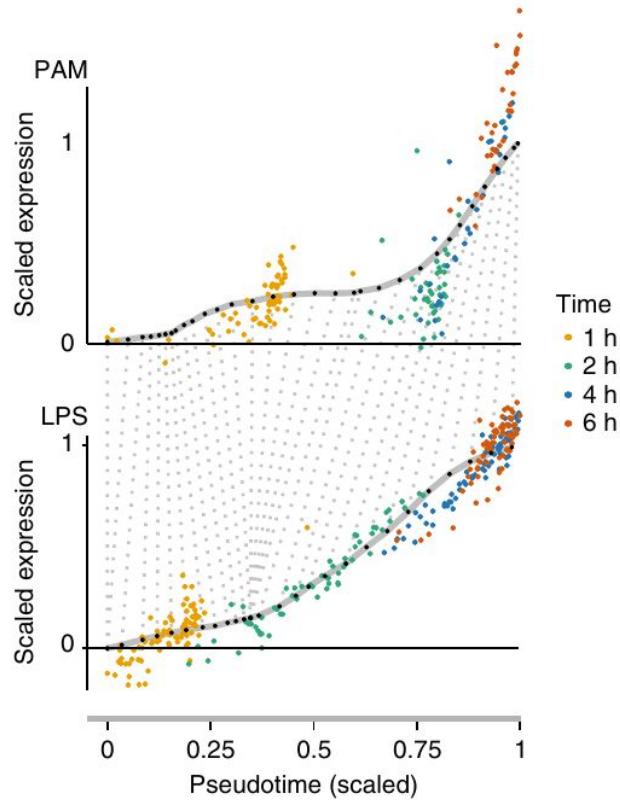
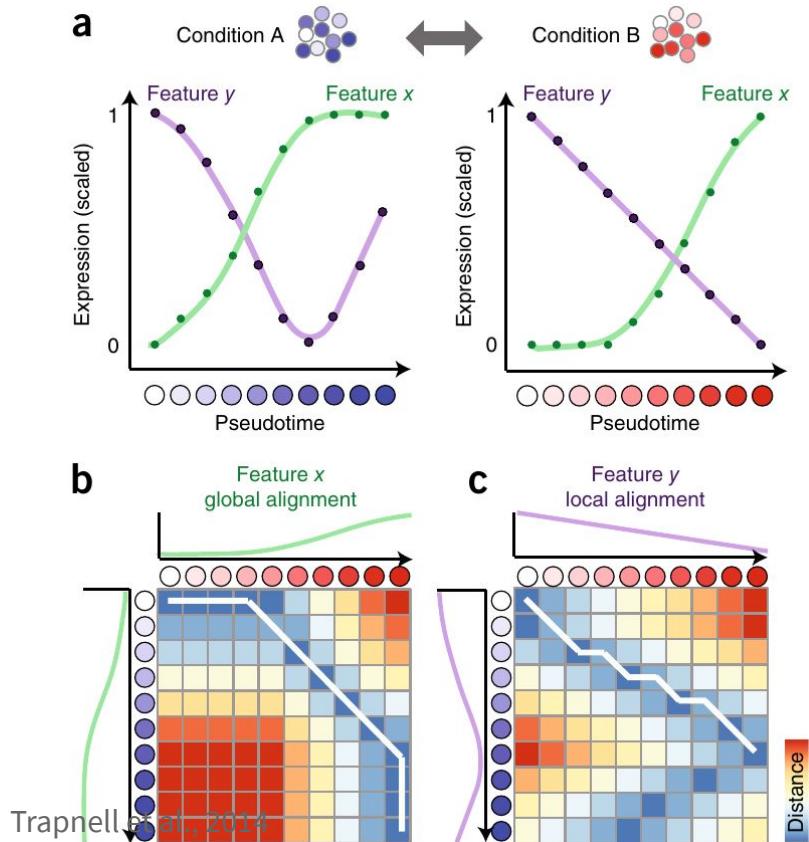


Comparing trajectories



“CellAlign”, Alpert et al., 2018

Comparing trajectories



“CellAlign”, Alpert et al., 2018

Discussion

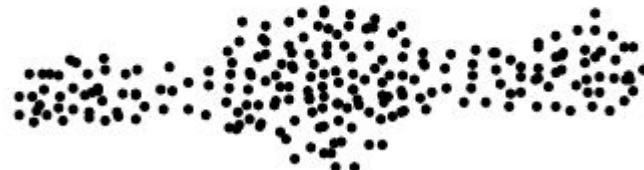
Is pseudotime inference a solved problem?

Fundamental limits on dynamic inference from single-cell snapshots

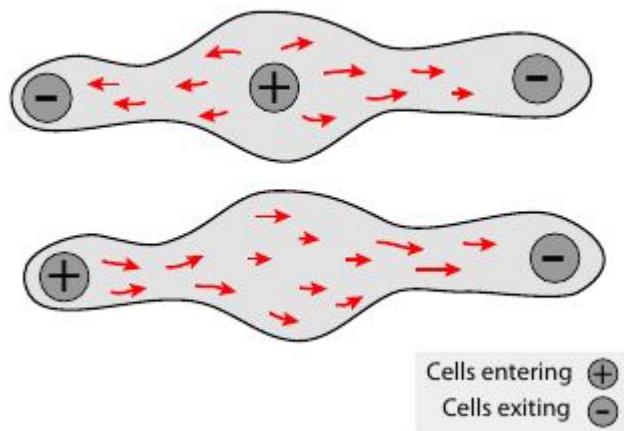
Caleb Weinreb^a, Samuel Wolock^a, Betsabeh K. Tusi^b, Merav Socolovsky^b, and Alon M. Klein^{a,1}

“The general challenge, even with perfect data, is that many regulatory mechanisms can generate the same dynamic process, and many dynamic processes can give rise to the same distribution.”

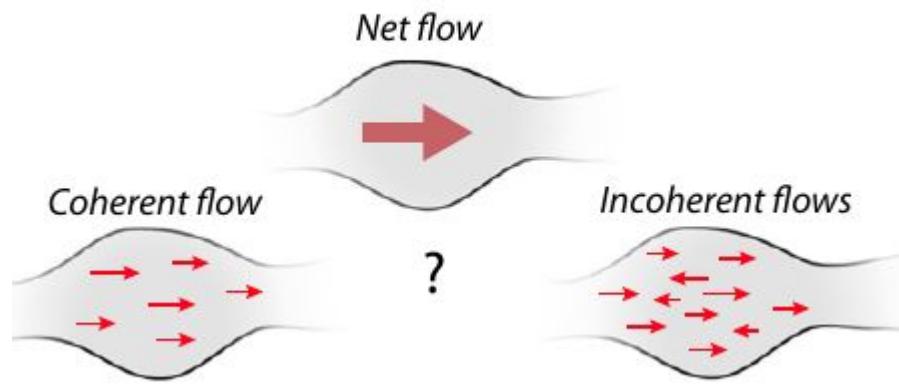
No unique solution



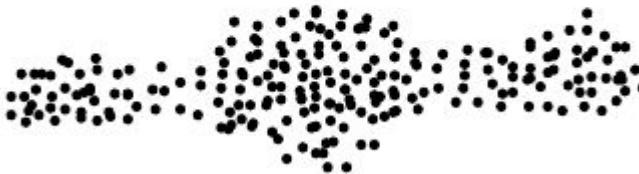
Entry and exit points direct the flow of cells



Net velocity may not equal actual velocity

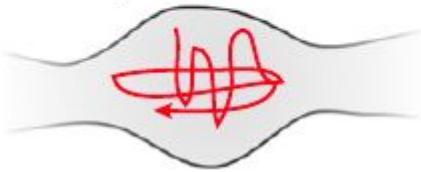


No unique solution

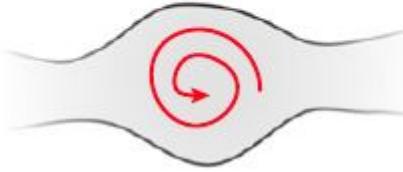


Rotations in state space do not alter cell density

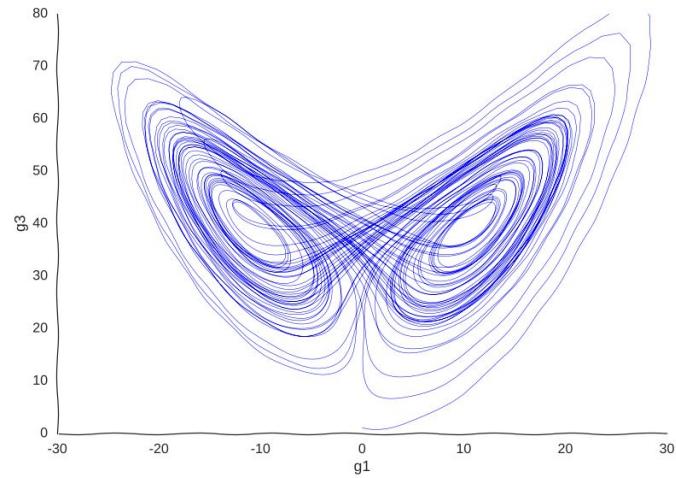
Simple fluctuations



Periodic oscillations

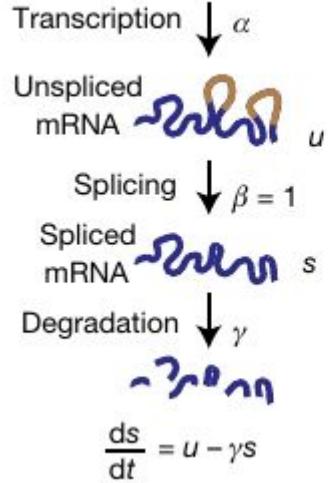


Two-state cyclical transition

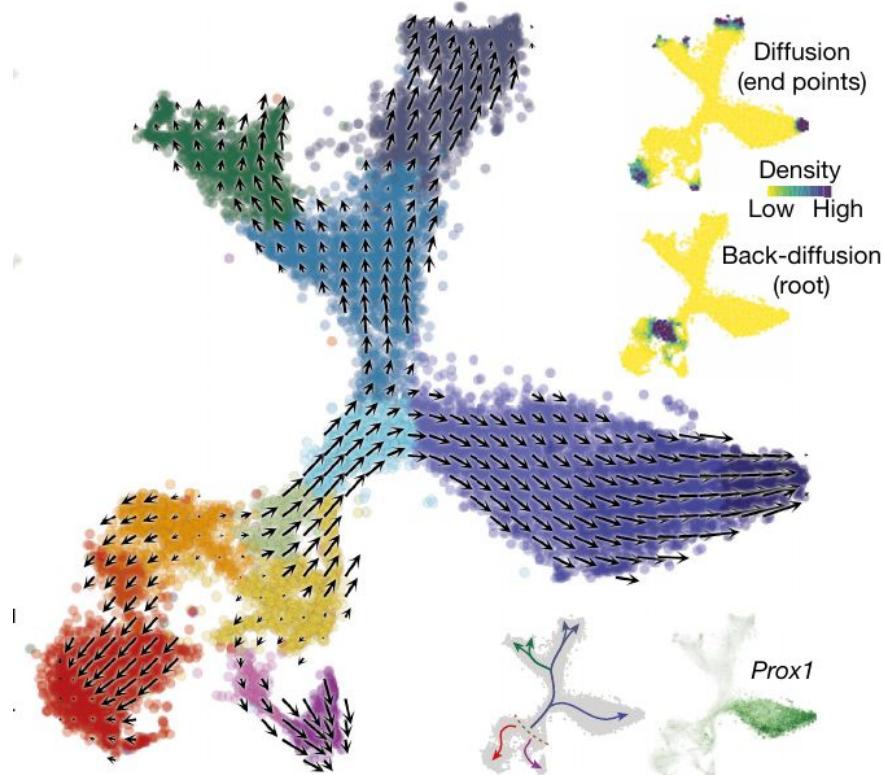
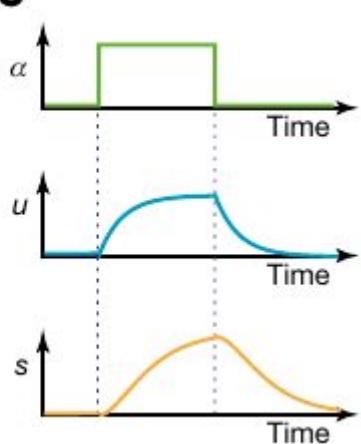


RNA velocity of single cells

b

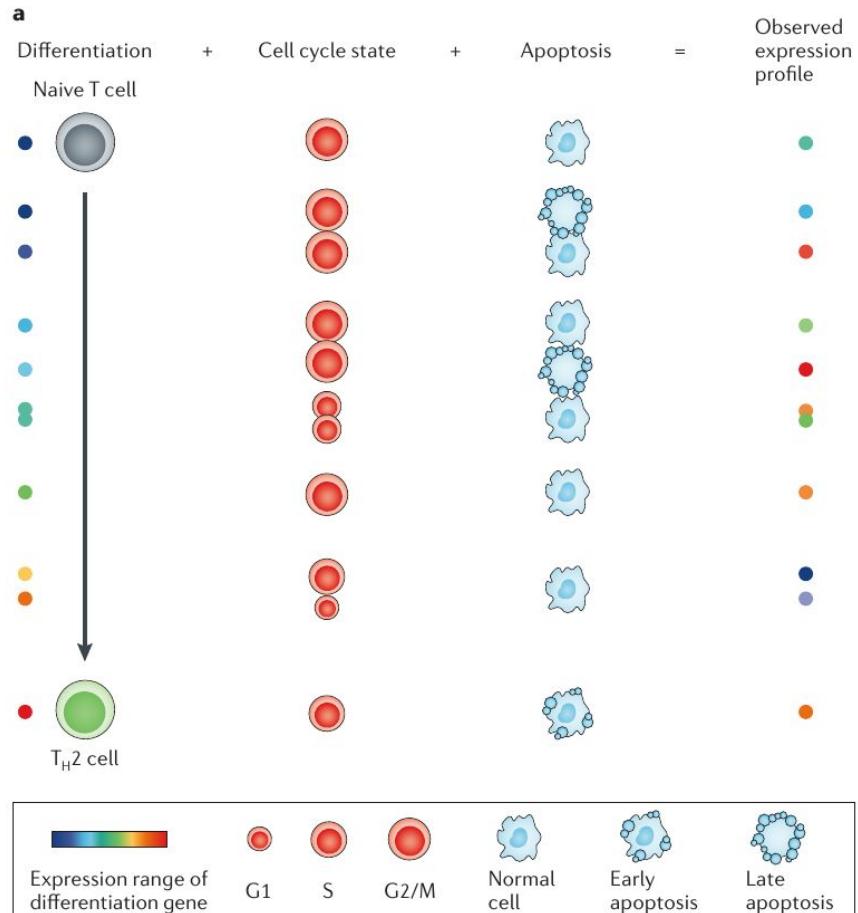


c



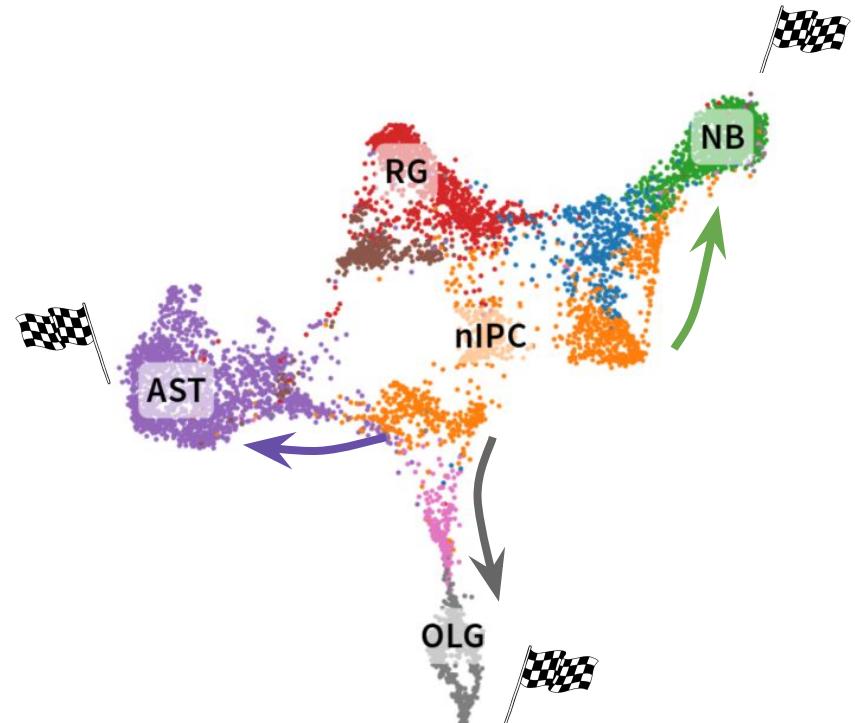
Confounding sources of variation

What biological processes can lead to the measured transcriptional profile?

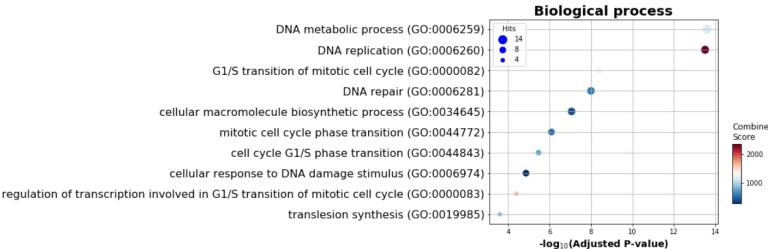
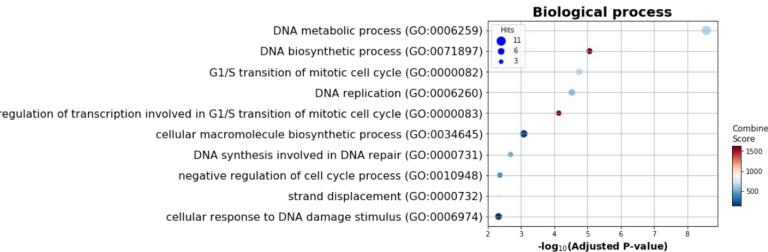
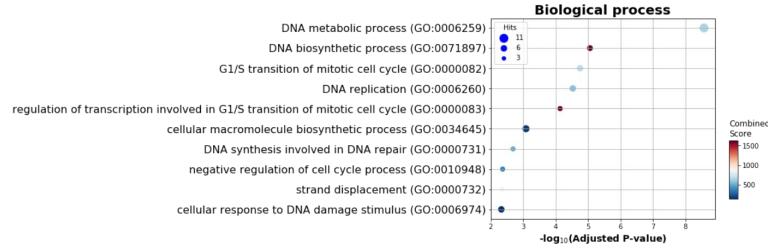
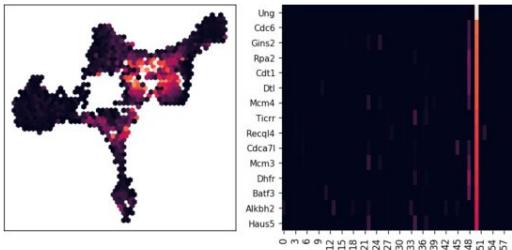
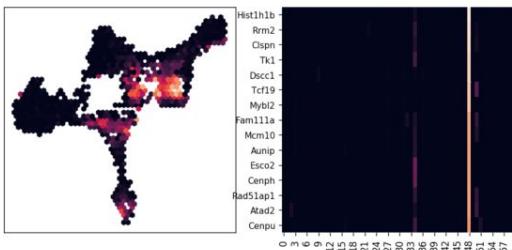
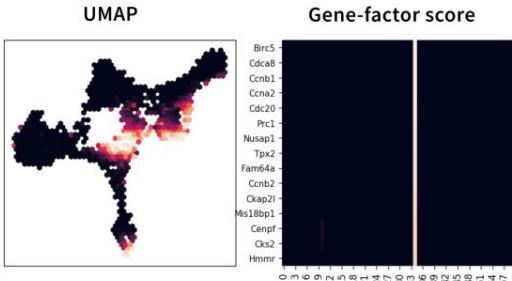


Developing mouse hippocampus dataset

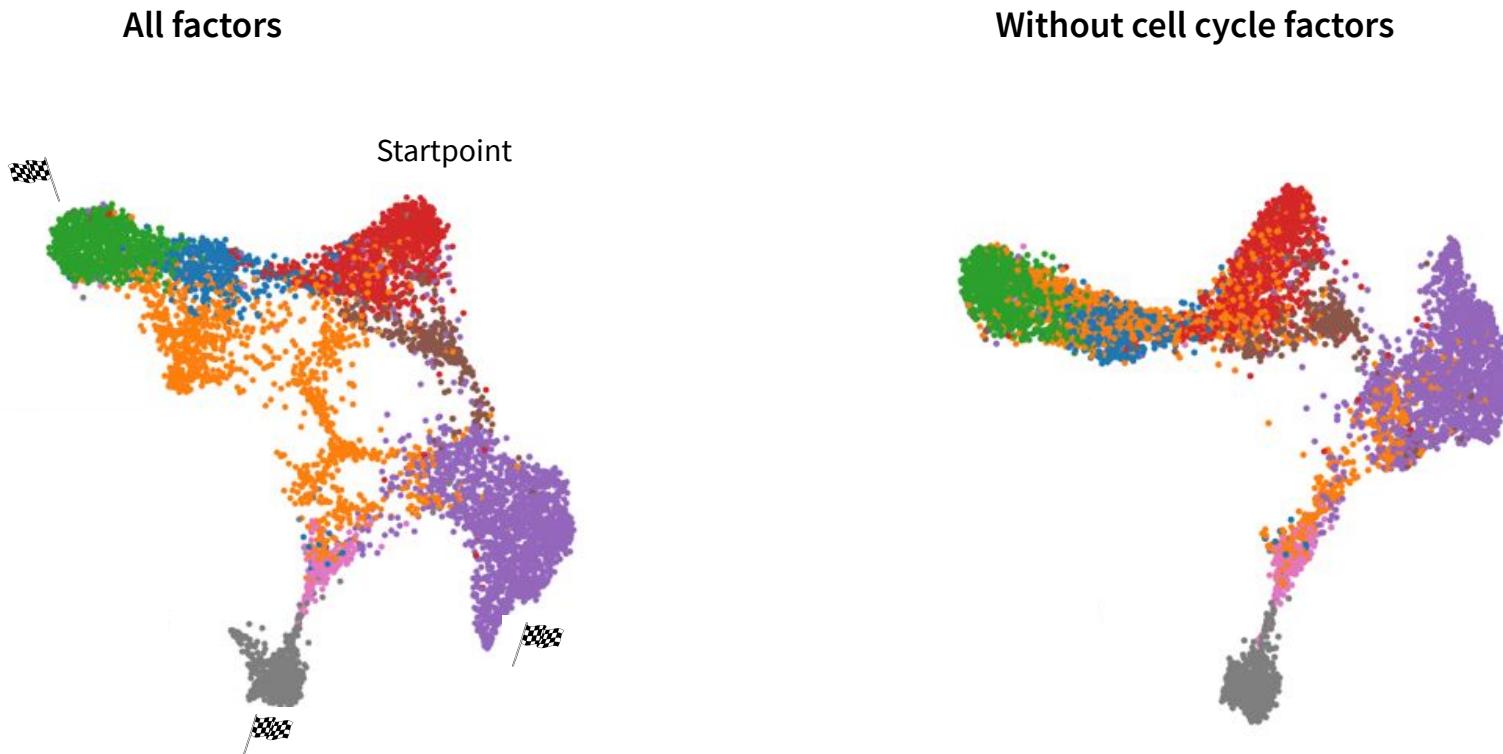
- Branching trajectory
- One common progenitor
 - Radial glia (RG)
- Intermediate cell state
 - Intermediate progenitor cells (nIPC)
- Multiple cell types
 - Neuroblasts (NB)
 - Oligodendrocytes (OLG)
 - Astrocytes (AST)



Cell cycle as a confounding process



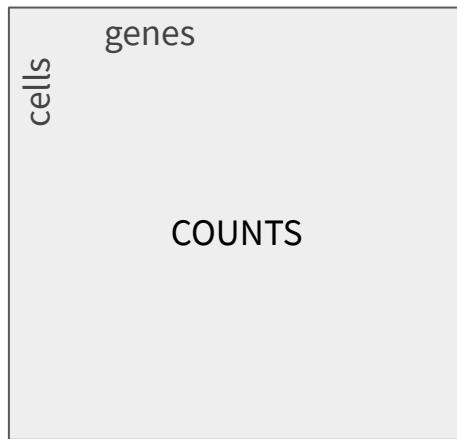
Cell cycle as a confounding process



Validation with RNA velocity



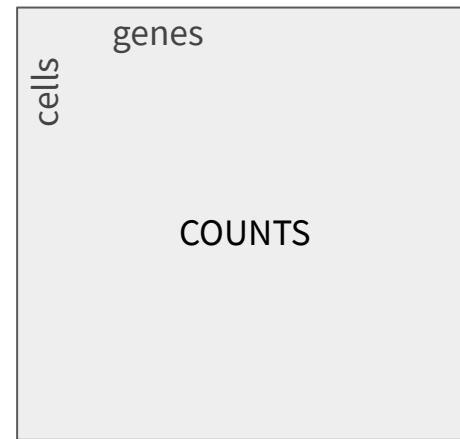
Timepoint 0



RNA velocity

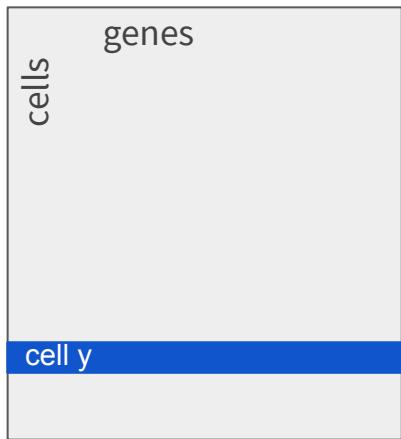


Timepoint 1

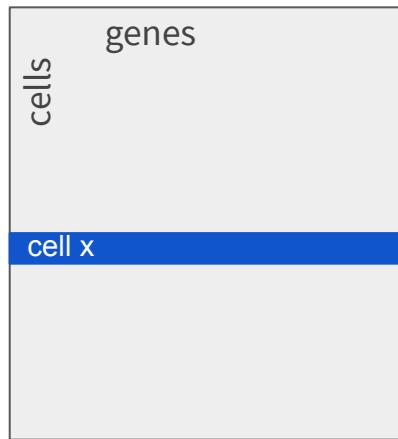


Transition probabilities

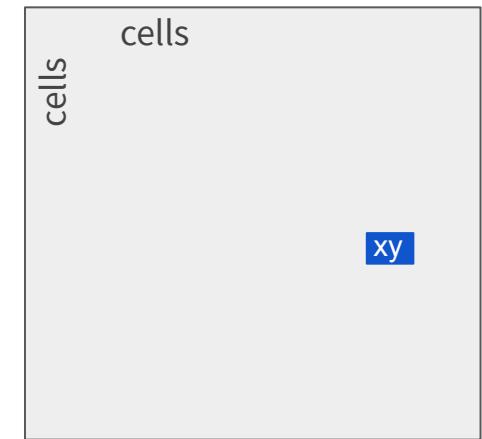
Counts t=0



Counts t=1



Transition probs.

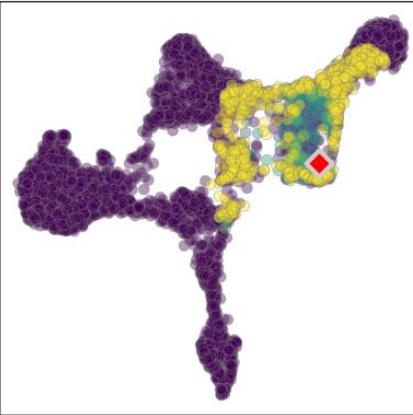


Precursor to

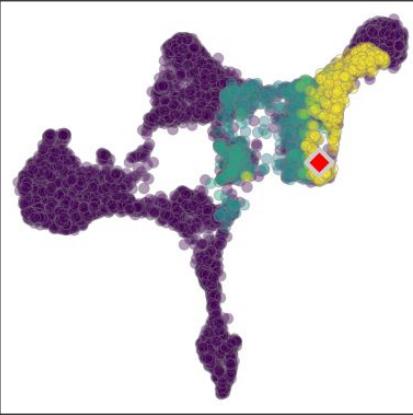
Cell cycle
factors

Neuroblasts

Included



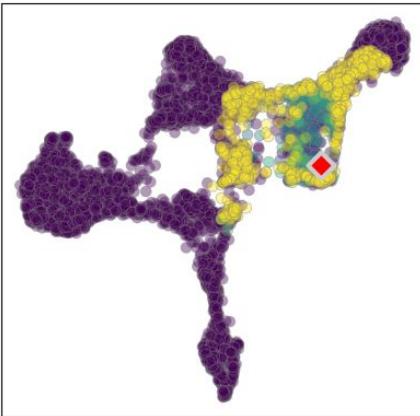
Excluded



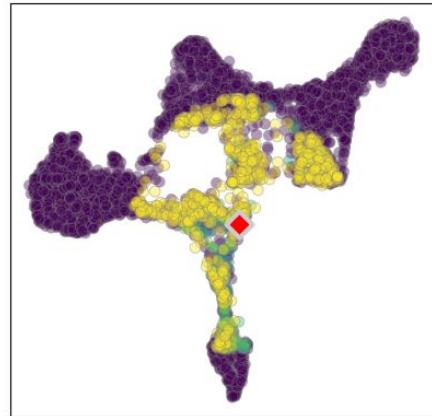
Precursor to

Cell cycle
factors

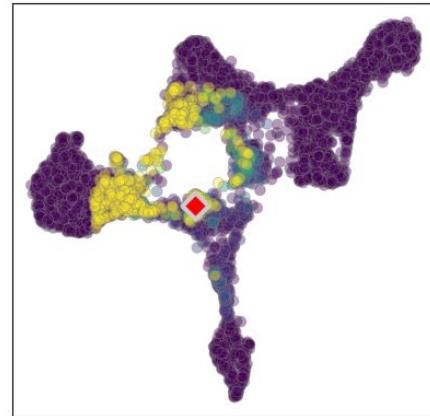
Neuroblasts



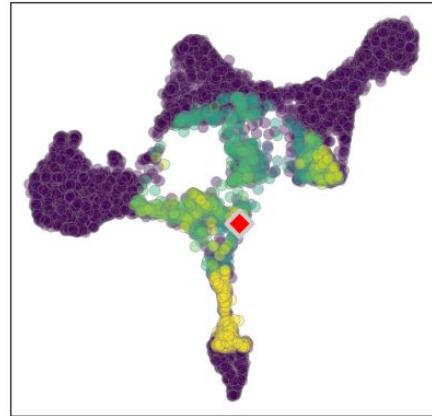
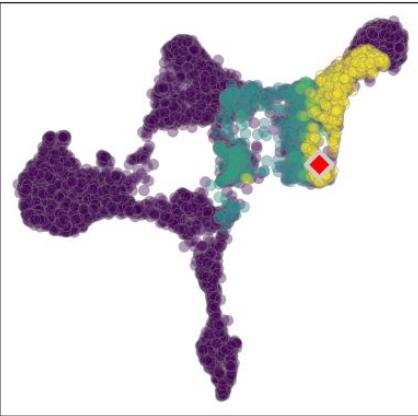
Oligodendrocytes



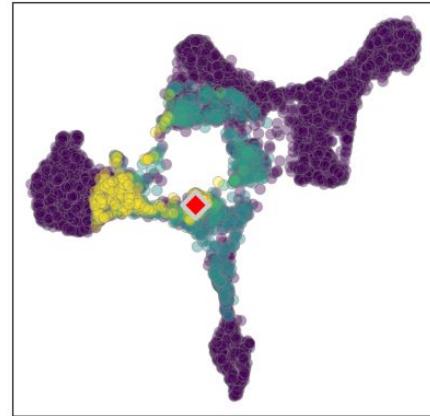
Astrocytes



Included



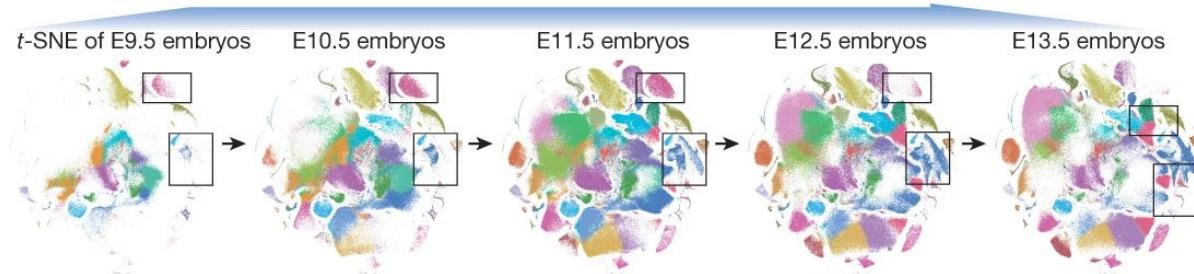
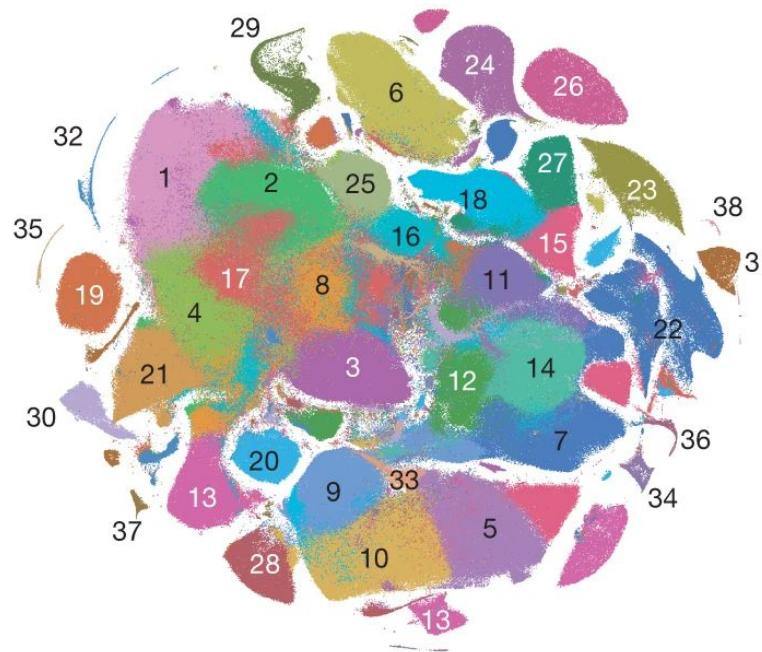
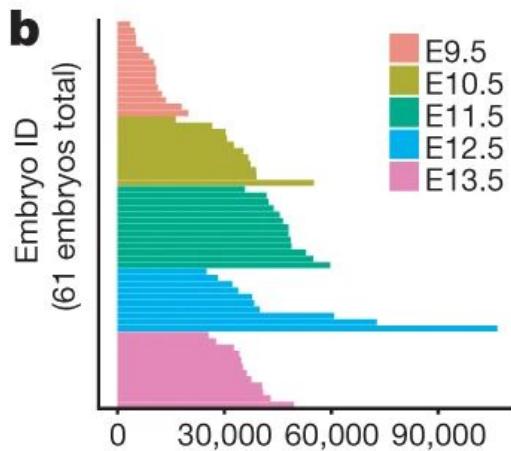
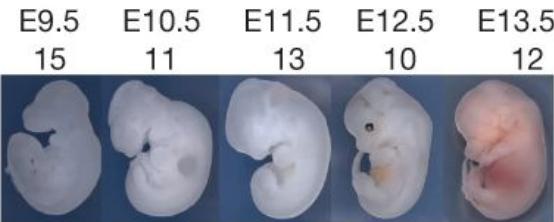
Excluded



Applications

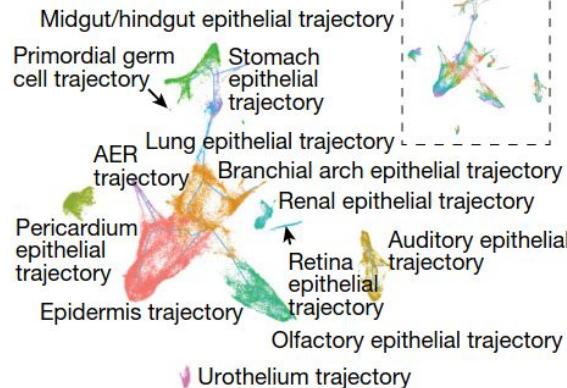
Mouse embryogenesis

Cao et al., 2019

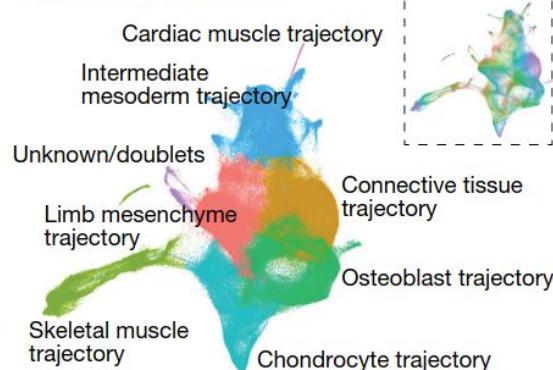


Mouse organogenesis

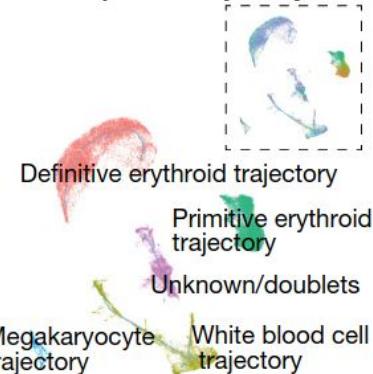
Epithelial trajectory



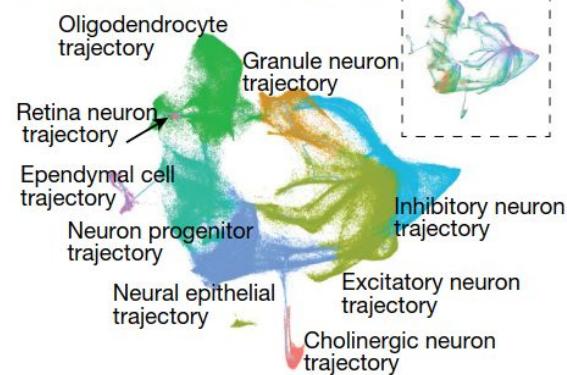
Mesenchymal trajectory



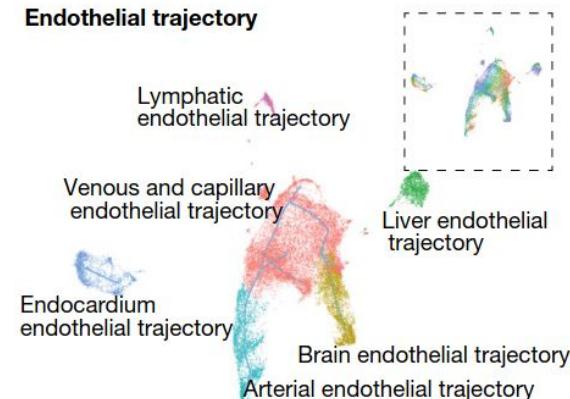
Haematopoiesis trajectory



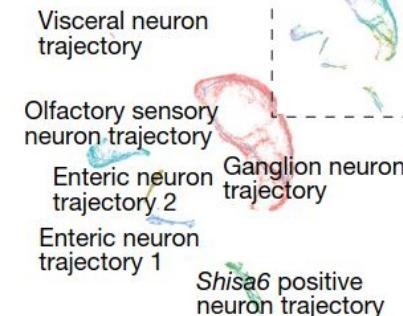
Neural tube/notochord trajectory



Endothelial trajectory



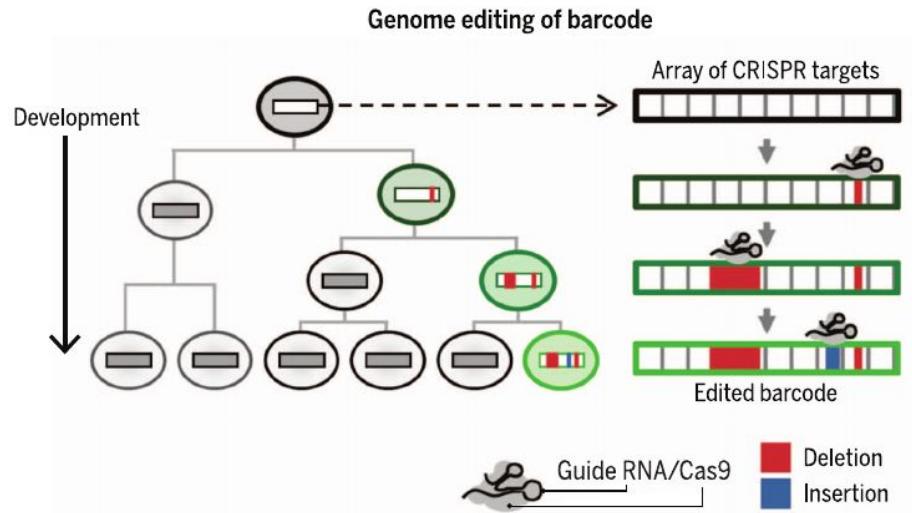
Neural crest (PNS neuron) trajectory 3



Lineage tracing

Pseudotime inference tracks cell differentiation “horizontally”

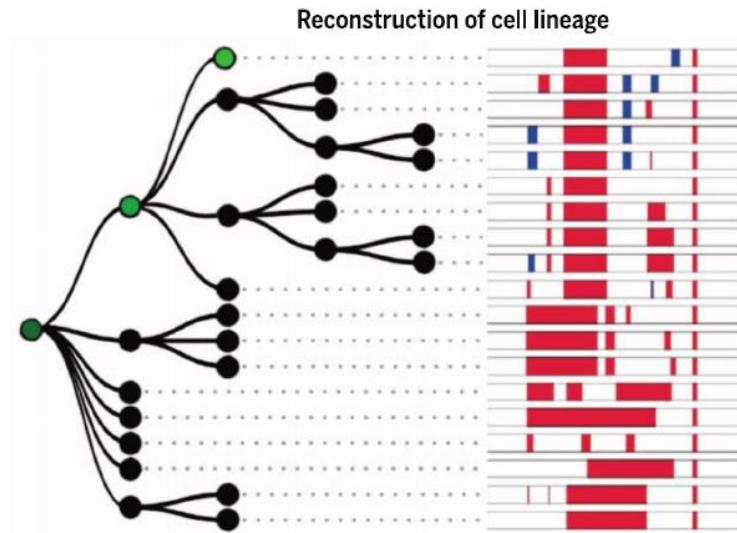
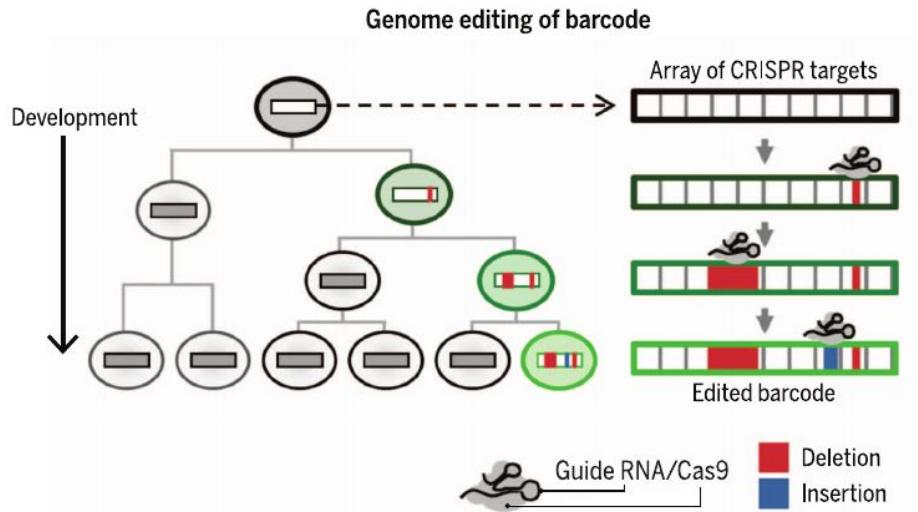
Lineage tracing tracks cell progeny:



Lineage tracing

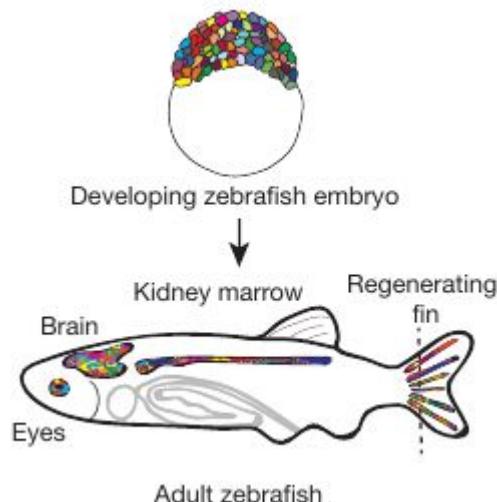
Pseudotime inference tracks cell differentiation “horizontally”

Lineage tracing tracks cell progeny:

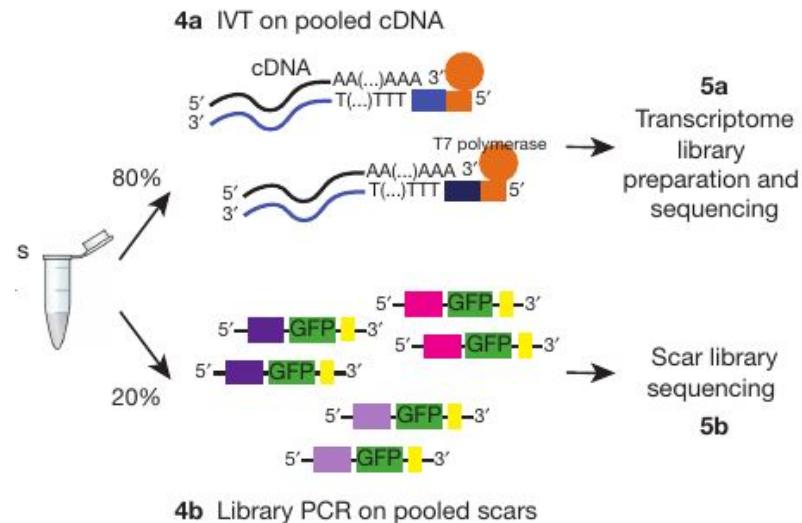


ScarTrace: Combining lineage tracing and scRNA-seq

Lineage tracing of zebrafish development. Span: 10 hpf

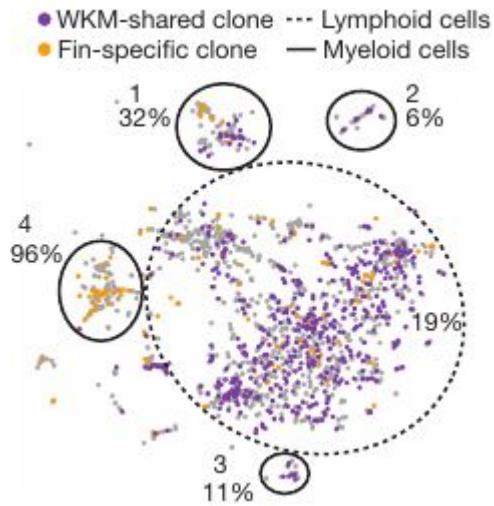


Cell material is split for scar library and transcriptome library preparation

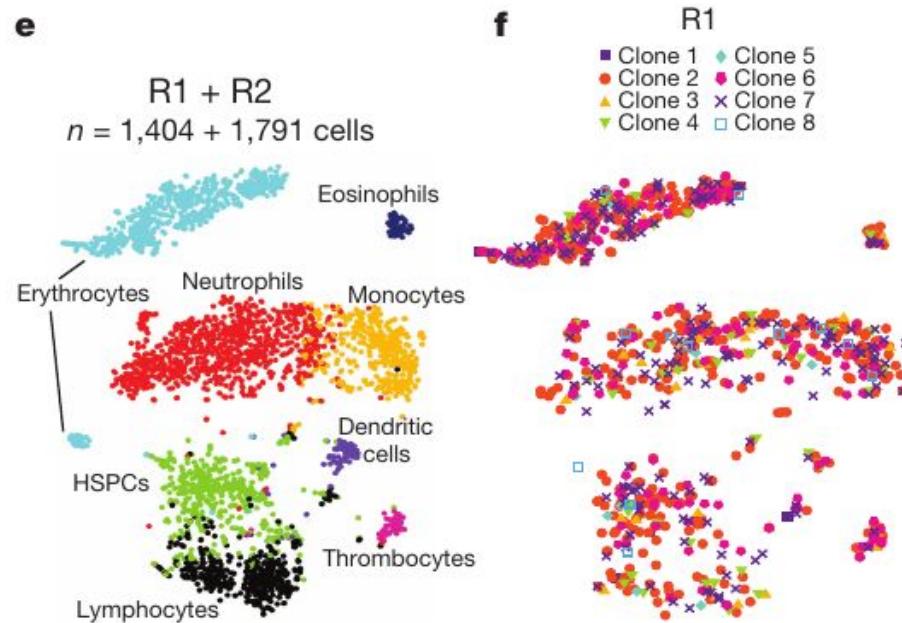


ScarTrace: Combining lineage tracing and scRNA-seq

Resident immune cells in fin



Clonal overlap in kidney marrow



Thank you!

- Folder: session-trajectories
- Use Monocle 2 and Destiny for pseudotime inference