

Methods for temporal analysis of single-cell data snapshots

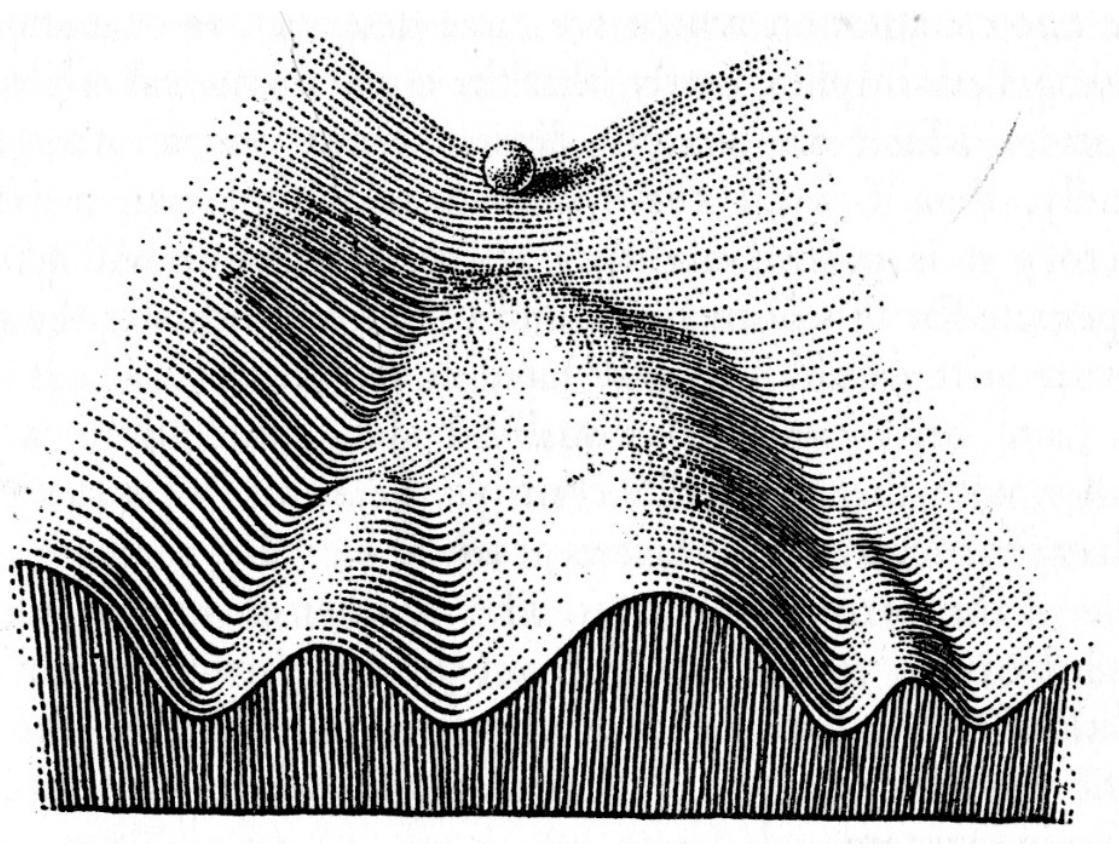
Alex Lederer
alex.lederer@epfl.ch

What is a cell state transition?

... the process by which cells change in the multi-dimensional feature space over time

What is a cell state transition?

... the process by which cells change in the multi-dimensional feature space over time



“Part of an epigenetic landscape”

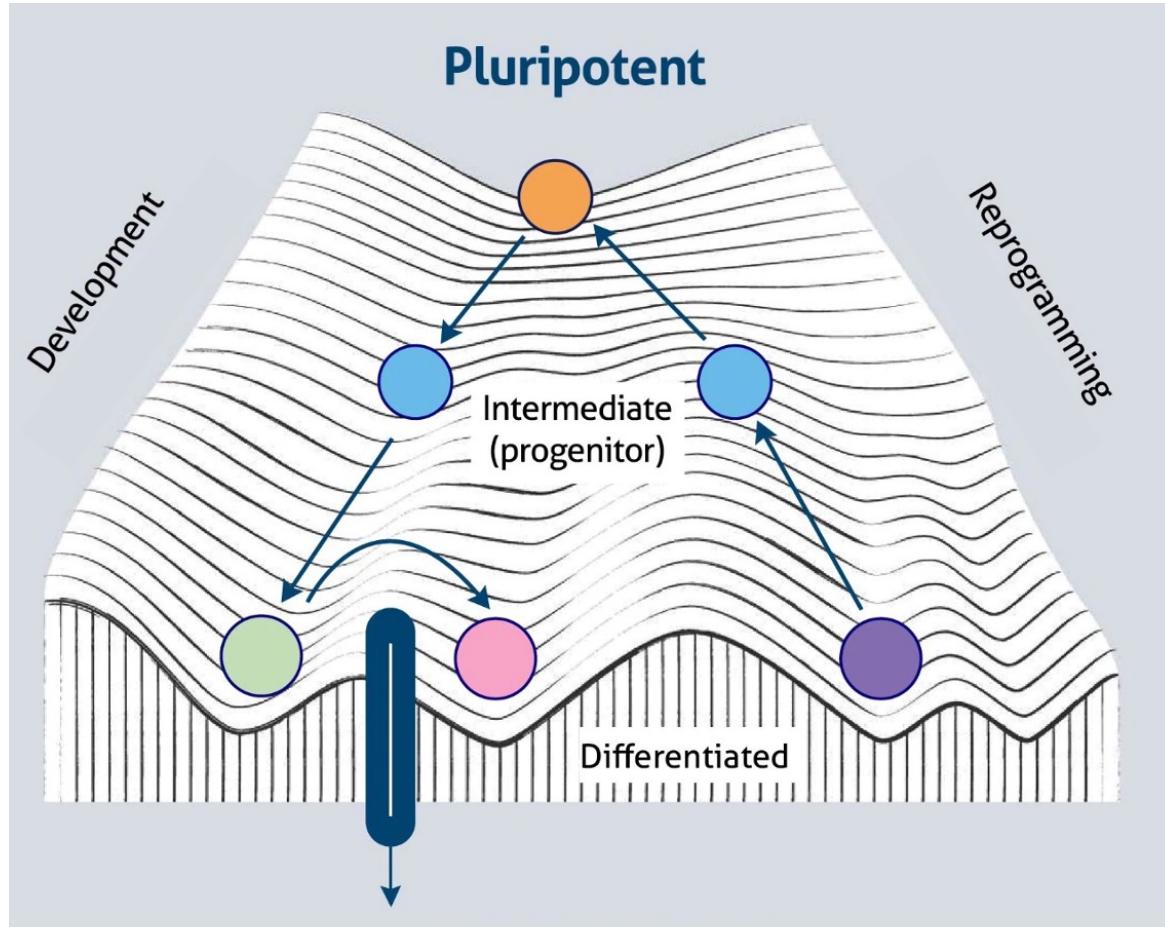
“A system containing many components can be represented by a point in multidimensional space. [...]”

In the study of development, we are interested not only in the final state to which the system arrives, but also the course by which it gets there [...]”

Conrad H. Waddington
The strategy of genes (1952)

What is a cell state transition?

... the process by which cells change in the multi-dimensional feature space over time



"A system containing many components can be represented by a point in multidimensional space. [...]

In the study of development, we are interested not only in the final state to which the system arrives, but also the course by which it gets there [...]

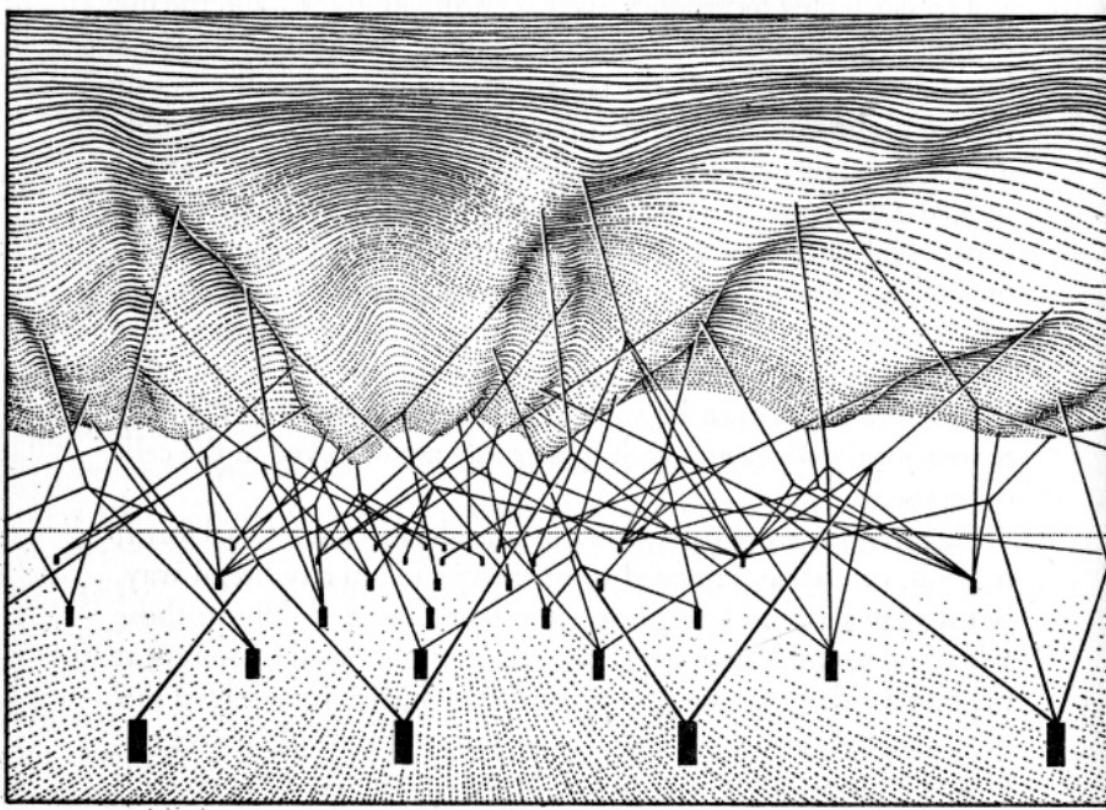
Conrad H. Waddington
The strategy of genes (1952)

What defines the features of the multi-dimensional space?

... usually, the genes!

What defines the features of the multi-dimensional space?

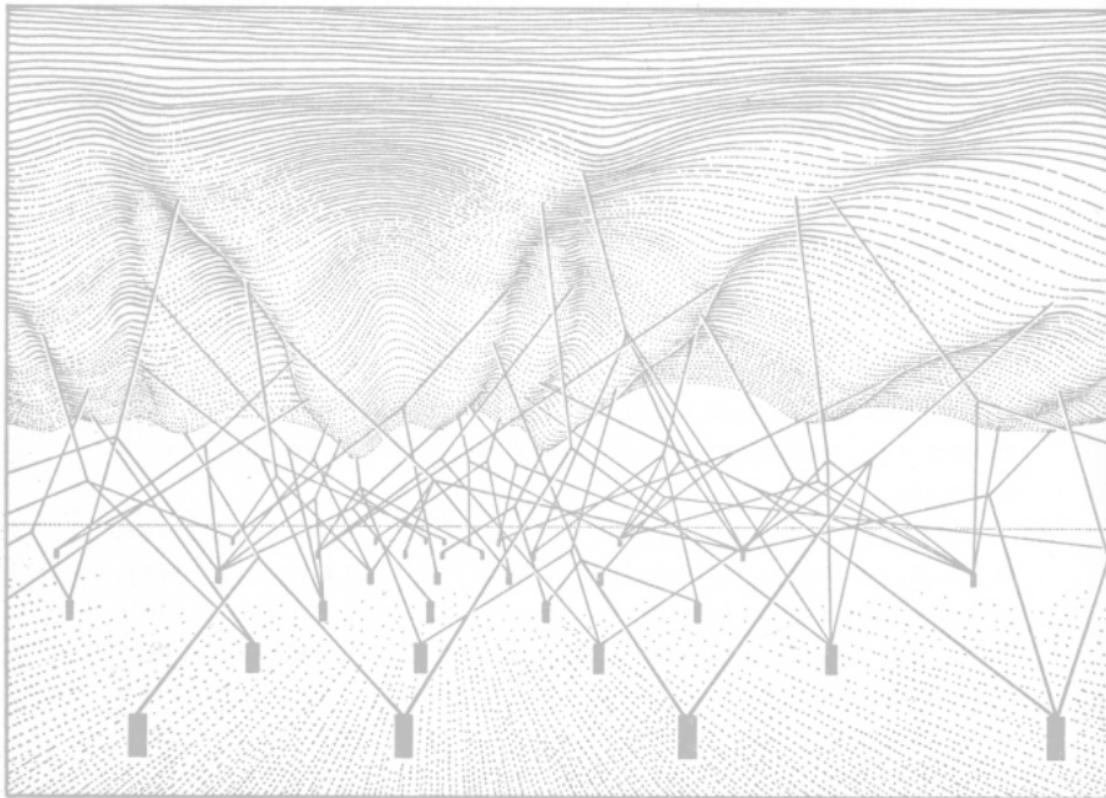
... usually, the genes!



“The complex system of interactions
underlying the epigenetic landscape”

**Changes in gene abundances
define the trajectories taken by
cells during cell state transitions**

What defines the features of the multi-dimensional space?



“The complex system of interactions underlying the epigenetic landscape”

We need technologies to **quantify** the abundance of molecular features

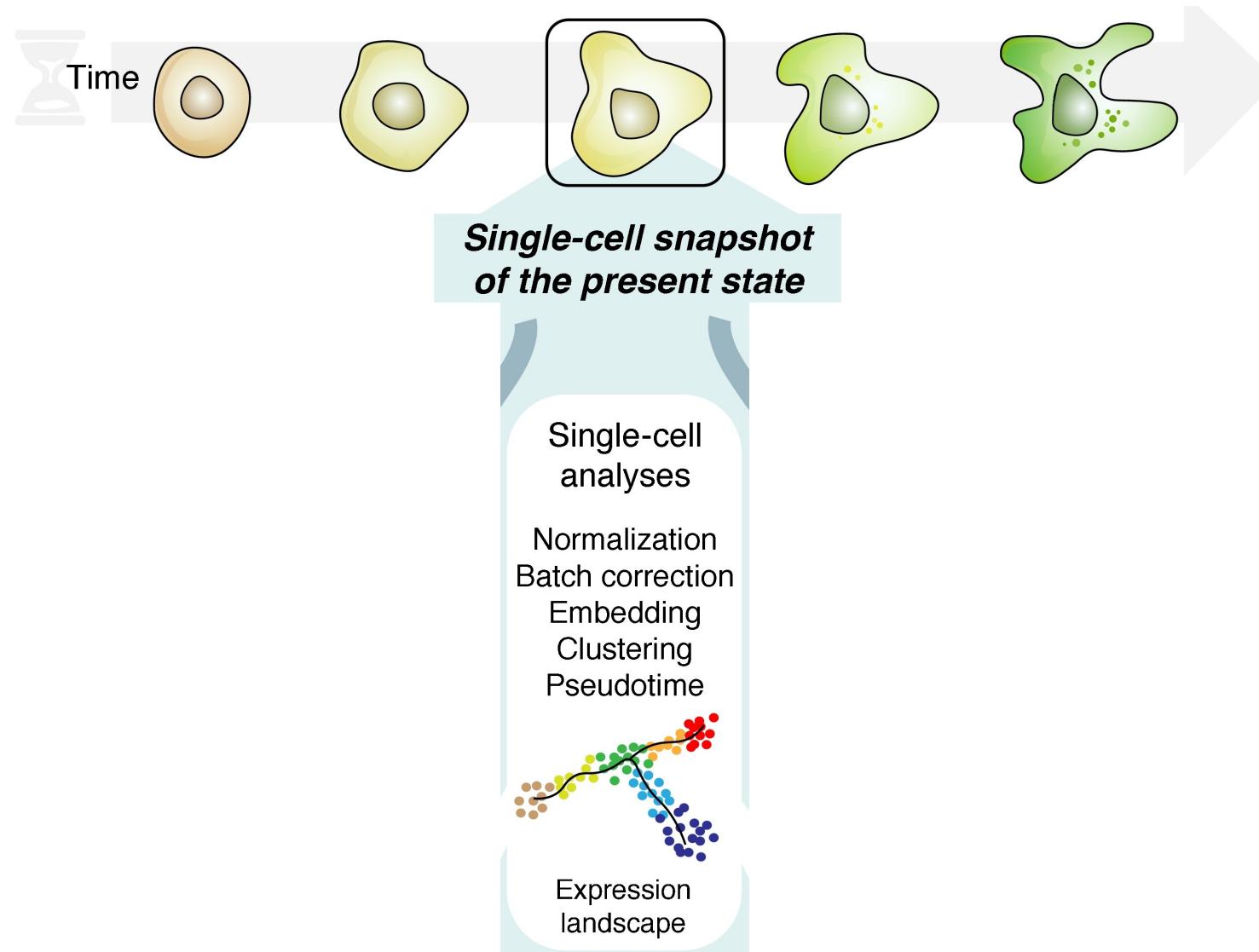
Single cell RNA sequencing!

10X
GENOMICS®

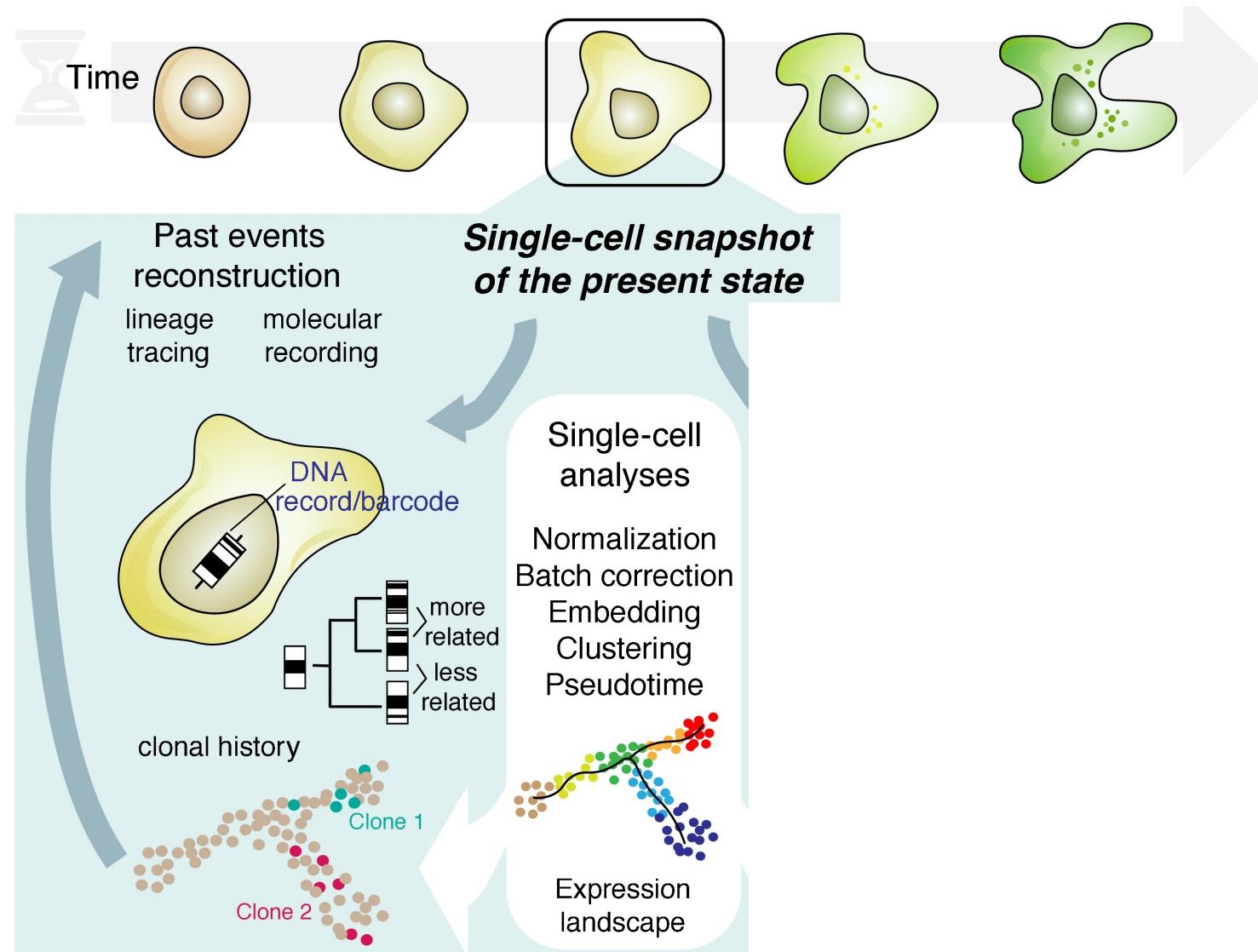
Single-cell RNA sequencing is a **destructive** technology

A cell can be profiled only one time,
providing a static **snapshot**

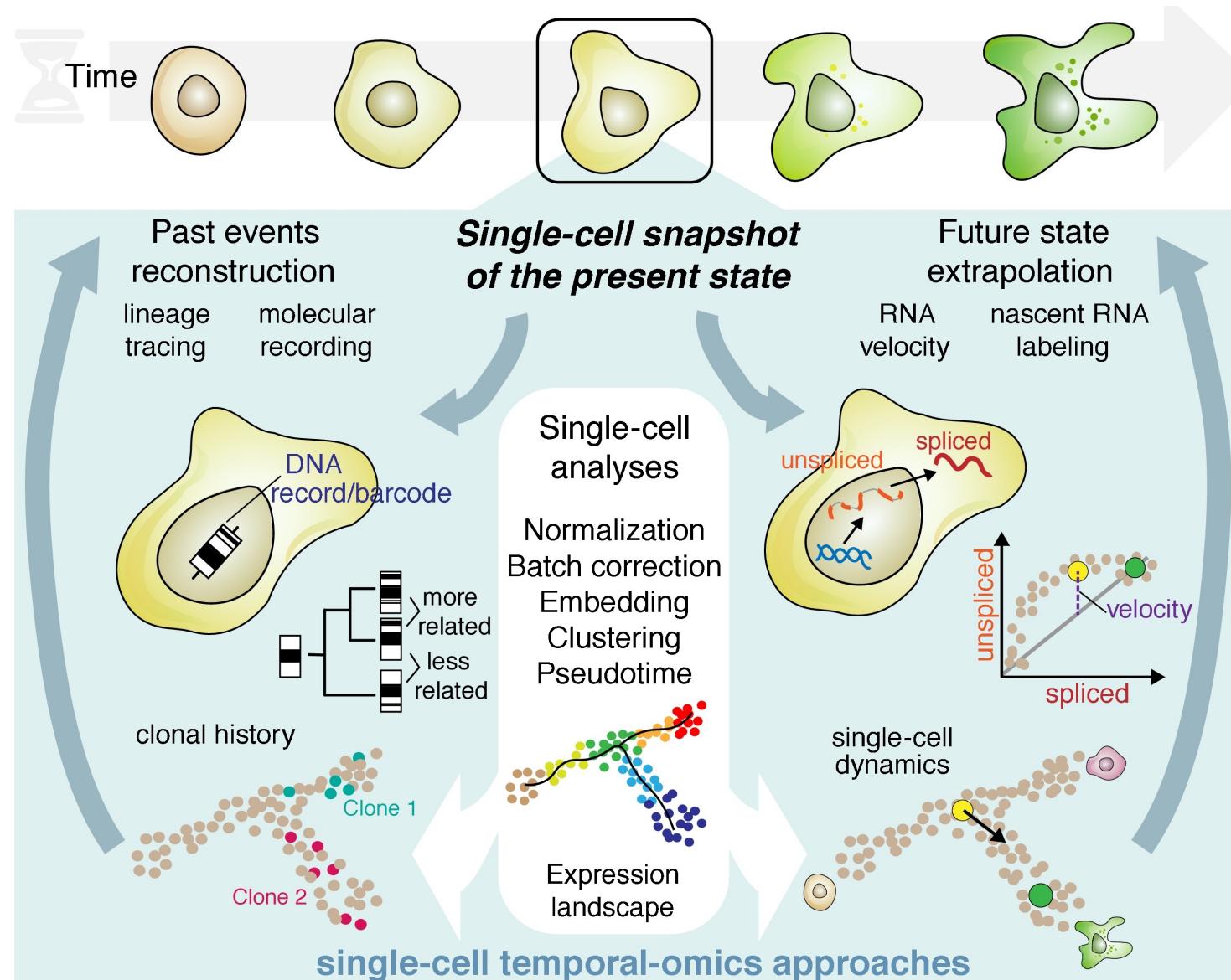
Emergence of single-cell temporal-omics approaches



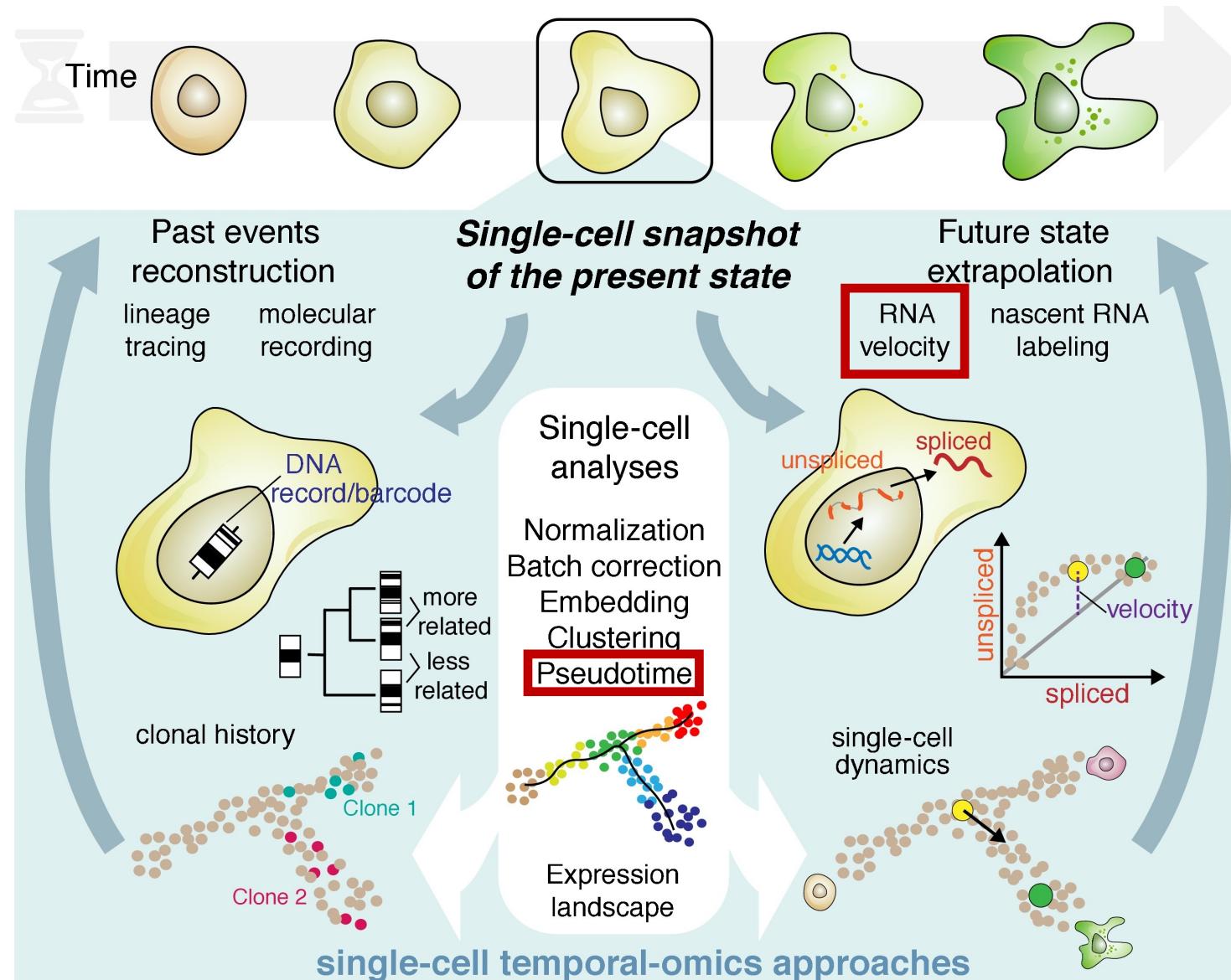
Emergence of single-cell temporal-omics approaches



Emergence of single-cell temporal-omics approaches



Emergence of single-cell temporal-omics approaches



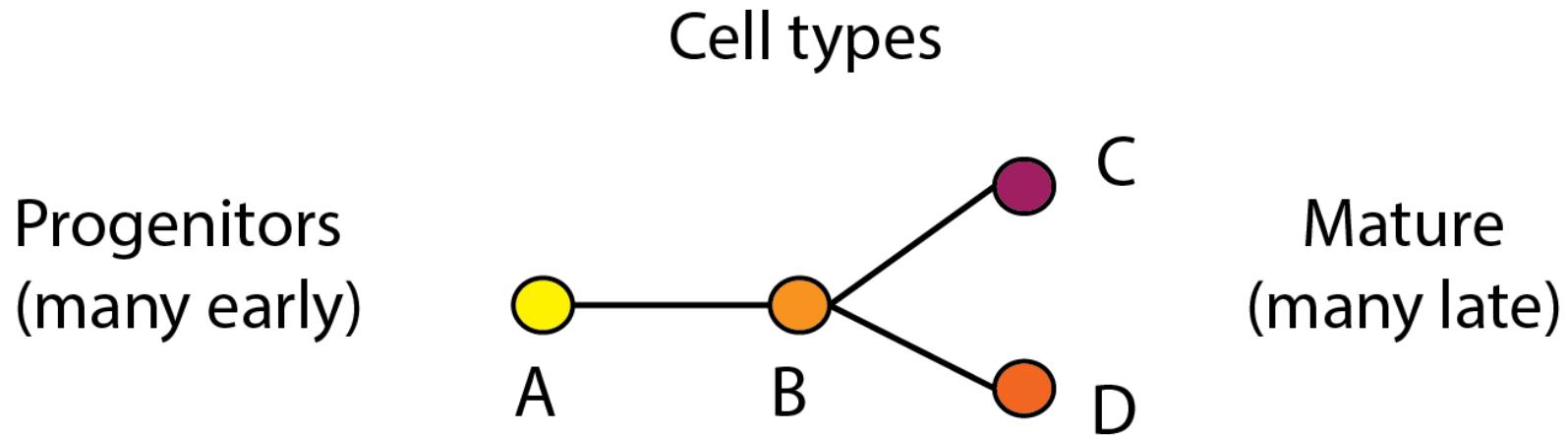
Question

Pseudotime trajectory inference

Pseudotime trajectory inference

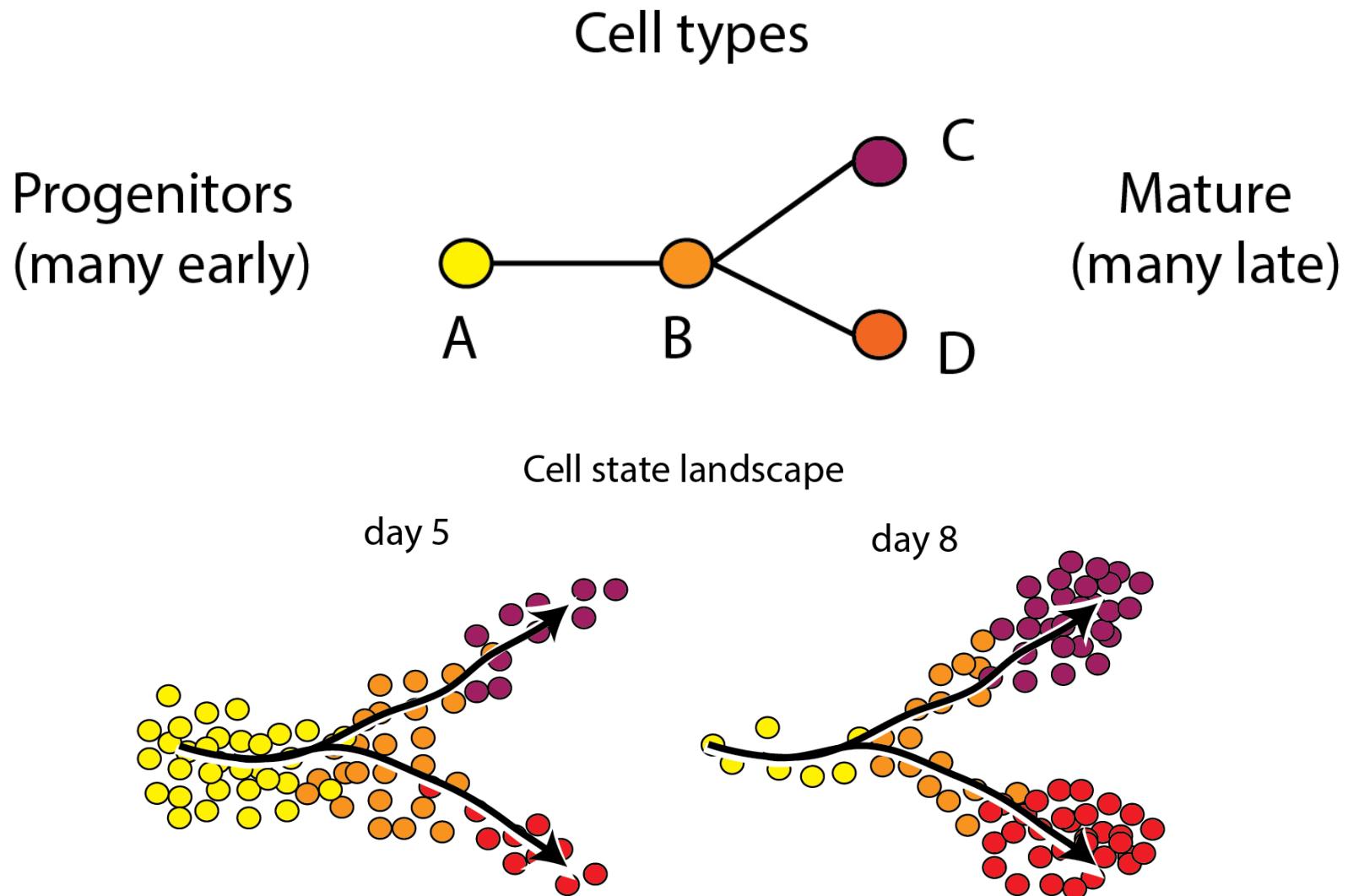
- Differences in gene expression between cells might be attributed to dynamic processes:
 - Cell cycle
 - Development or differentiation
 - Response to a stimuli (environmental change, drug treatment)
- Trajectory inference orders a population of individual cells along a path or lineage
- Cells on the learned path can be assigned a “pseudotime”, which is a measure of their amount progress along the path.
- Can be a good starting point for further analysis:
 - Determine gene expression programs driving changes in conditions that drive cells towards more or less differentiated states or phenotypes (i.e., number of cells in the beginning vs end of the pseudotime axis).

Pseudotime trajectory inference

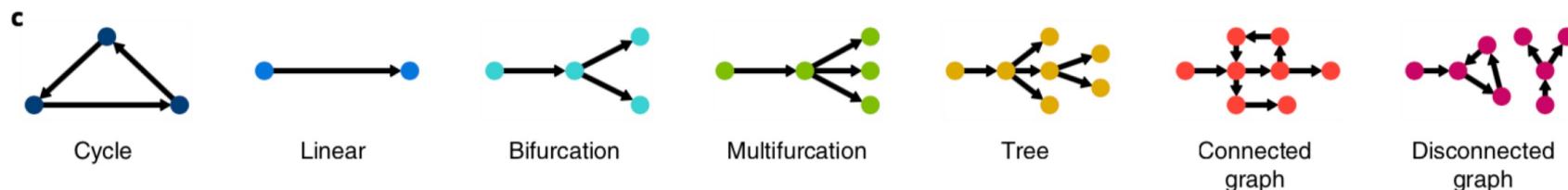
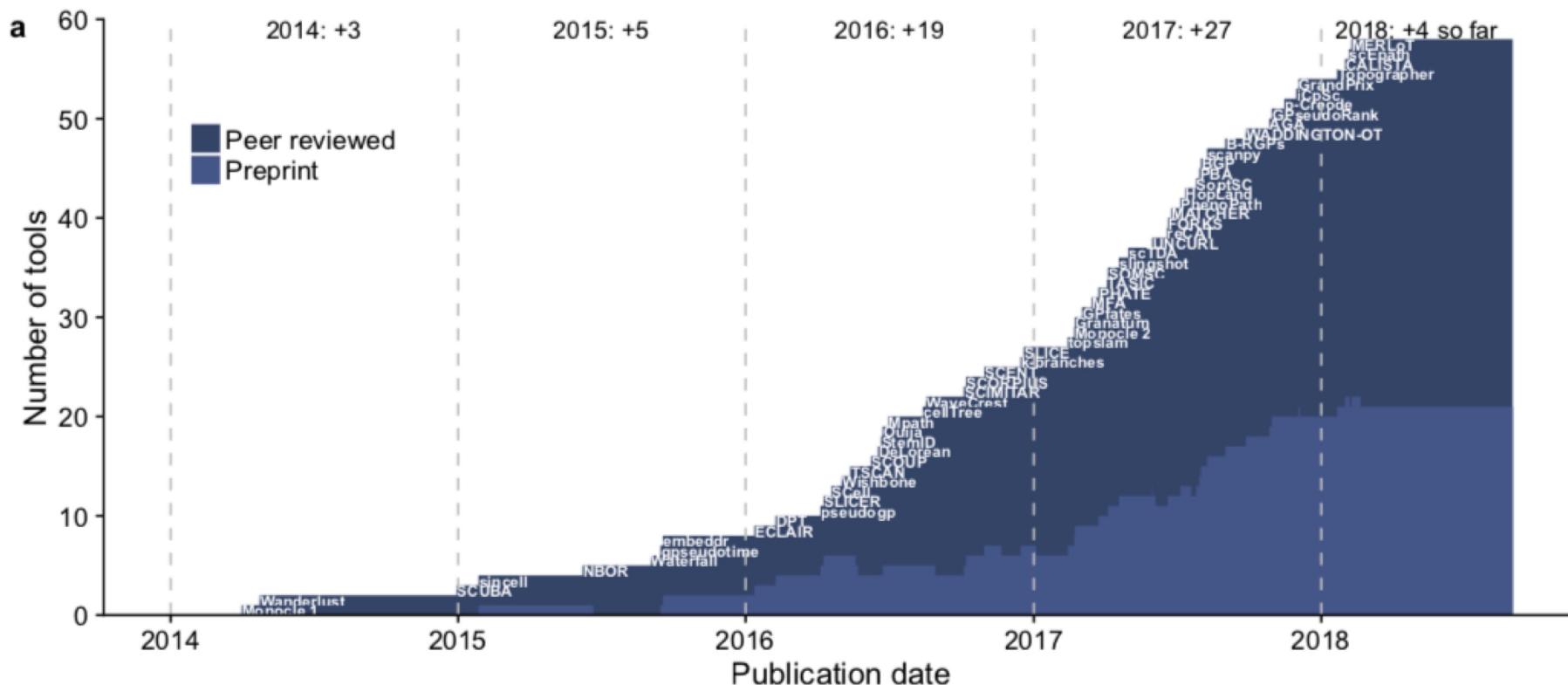


let's say this process takes 7 days to unfold...

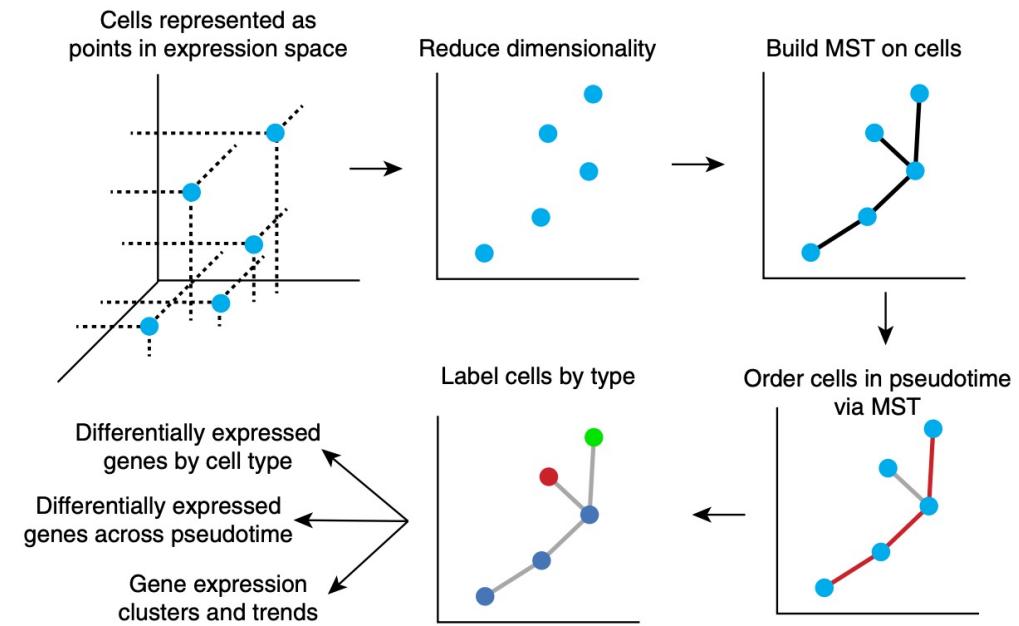
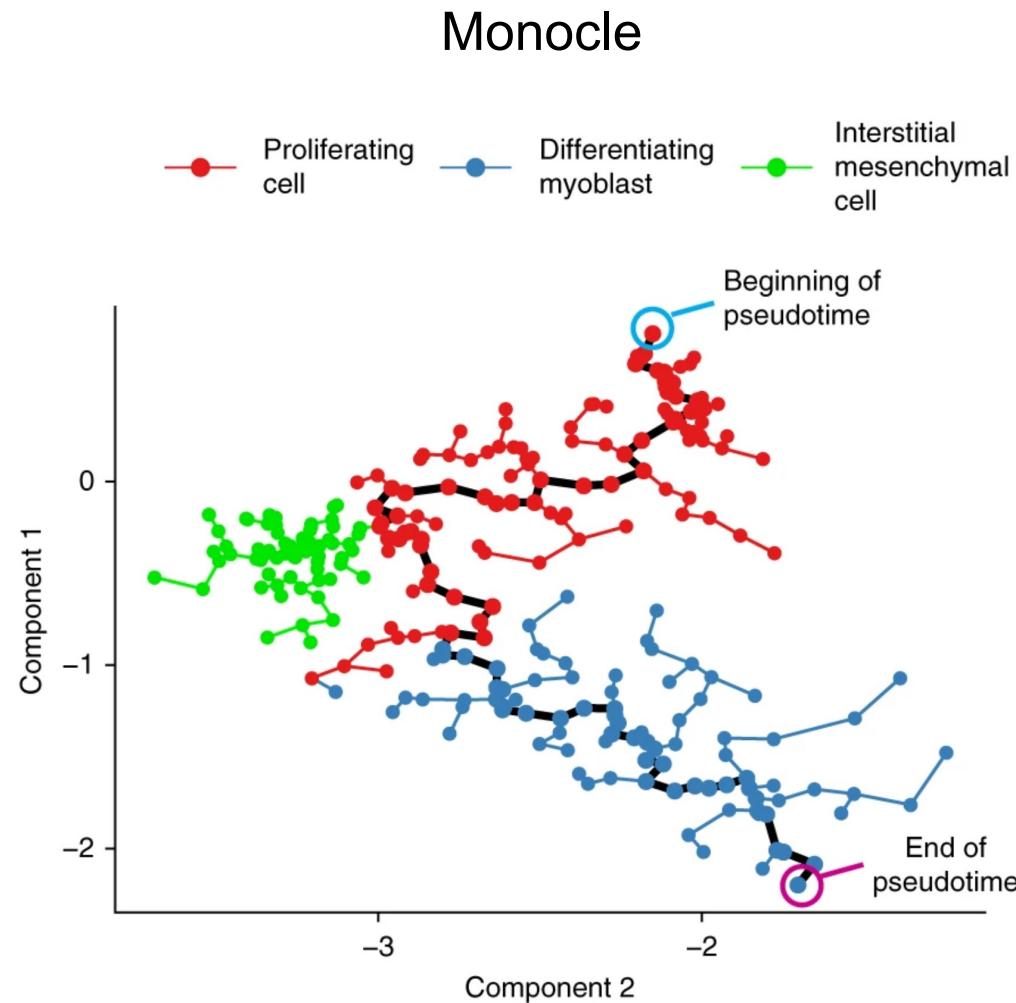
Pseudotime trajectory inference



There are many trajectory inference methods to choose from!

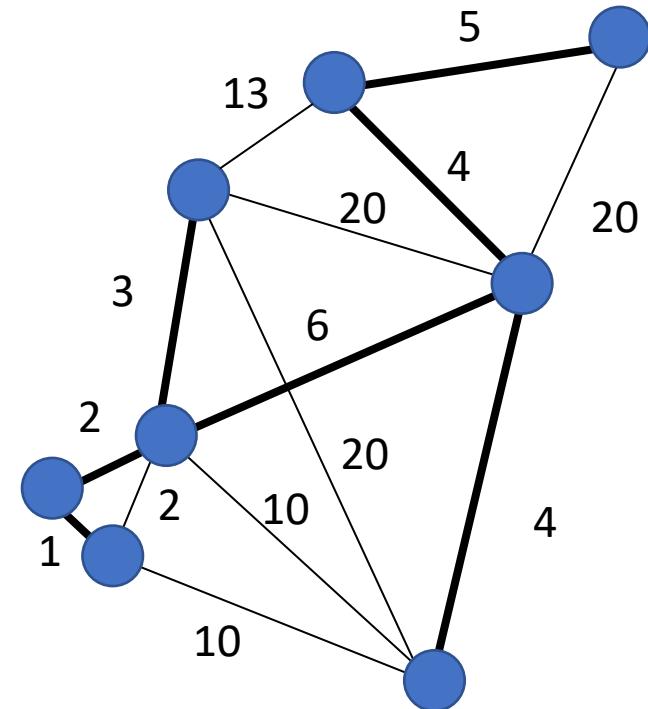


The first pseudotime algorithm: Monocle



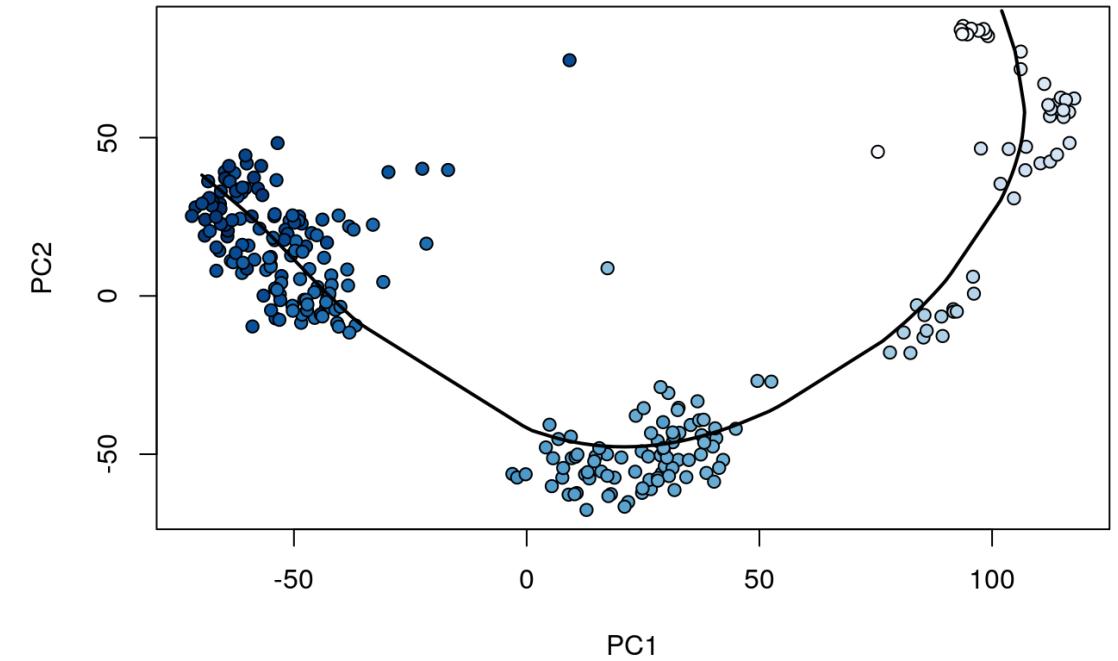
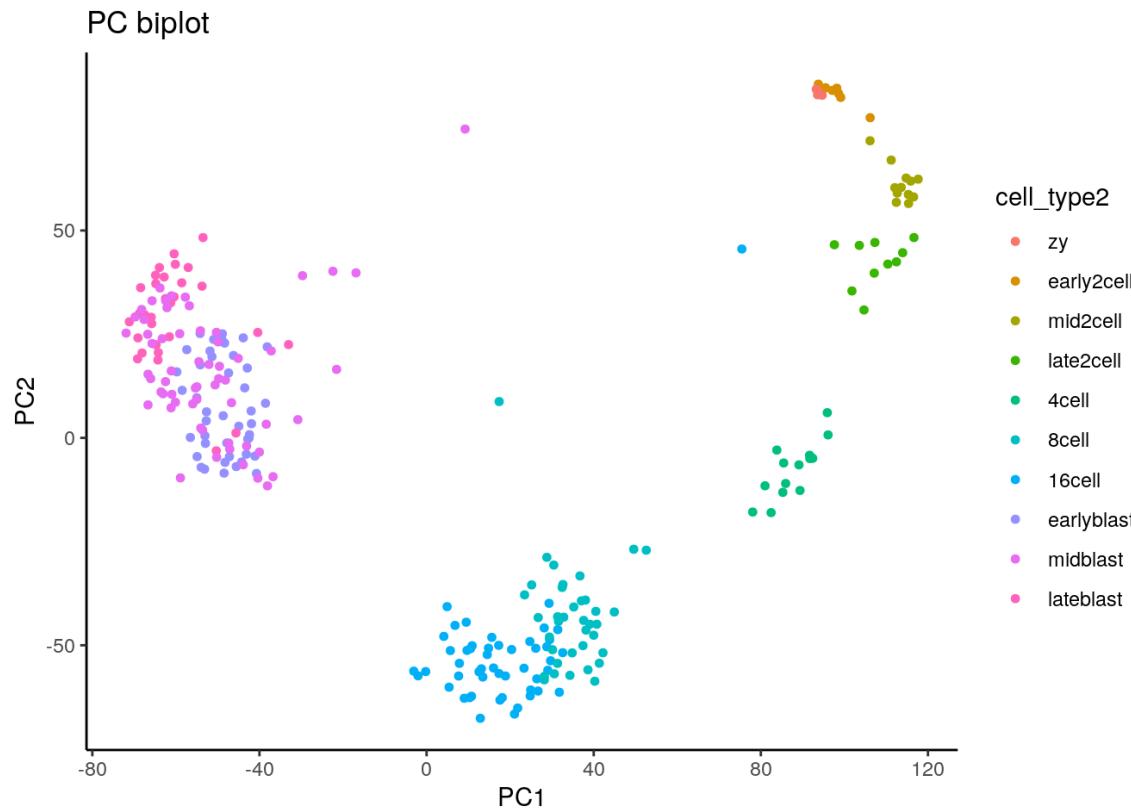
Minimum spanning tree

- Sum of all distances in the tree (graph) among single cells is at its minimum
- Having more intermediate cells improves the definition of the tree
- The weights are usually a distance in the dimensionality reduction space (PCA, UMAP)
- MST has no cycles, cell cycles will not work in here



Optional exercise on pseudotime trajectory inference

- Tutorial for estimating a pseudotime trajectory using Monocle3 and slingshot



https://sib-swiss.github.io/single-cell-training/day3/day3-3_trajectory_analysis.html

Dr. Rachel Marcone

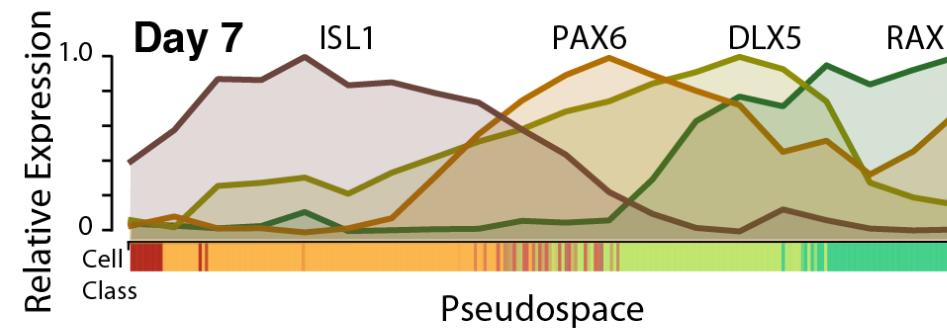
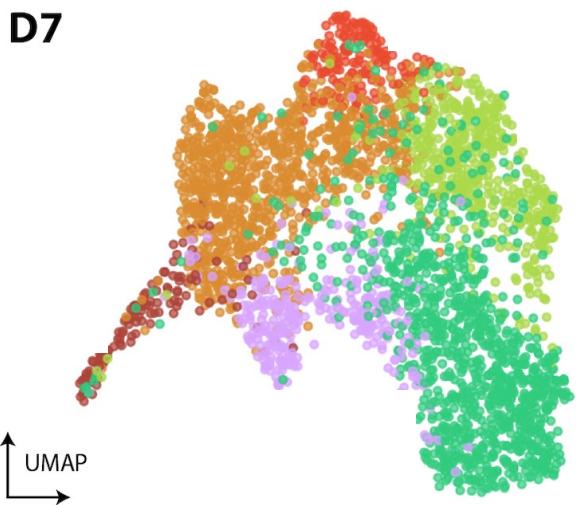
Should you run trajectory inference?

Questions to ask:

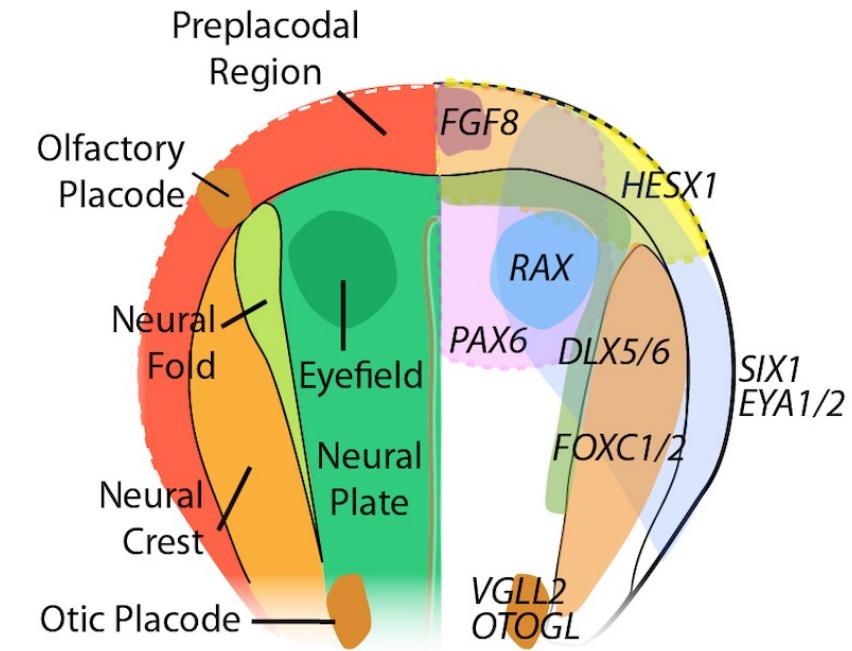
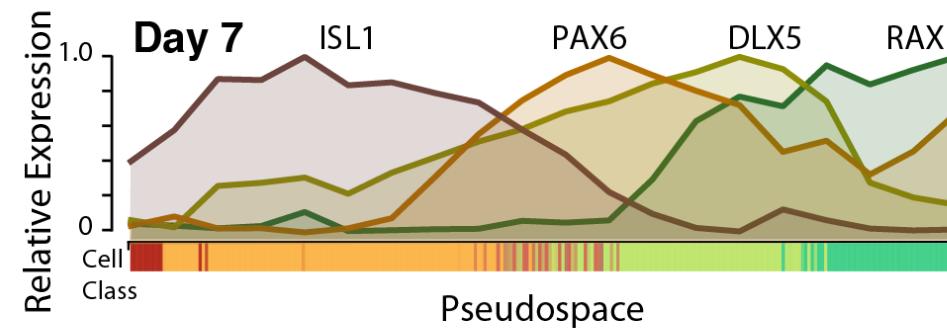
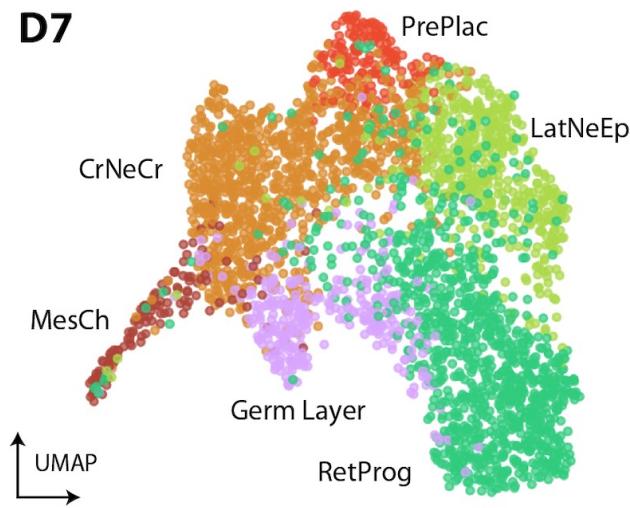
- Are you sure that you expect a trajectory?
- Do you have intermediate states?
- Do you think you have branching in your trajectory?
- Do you have a time scale on your cells?
- Do you know your start or end state?

Be aware, any dataset can be forced into a trajectory without any biological meaning!

An example where pseudotime can be misinterpreted



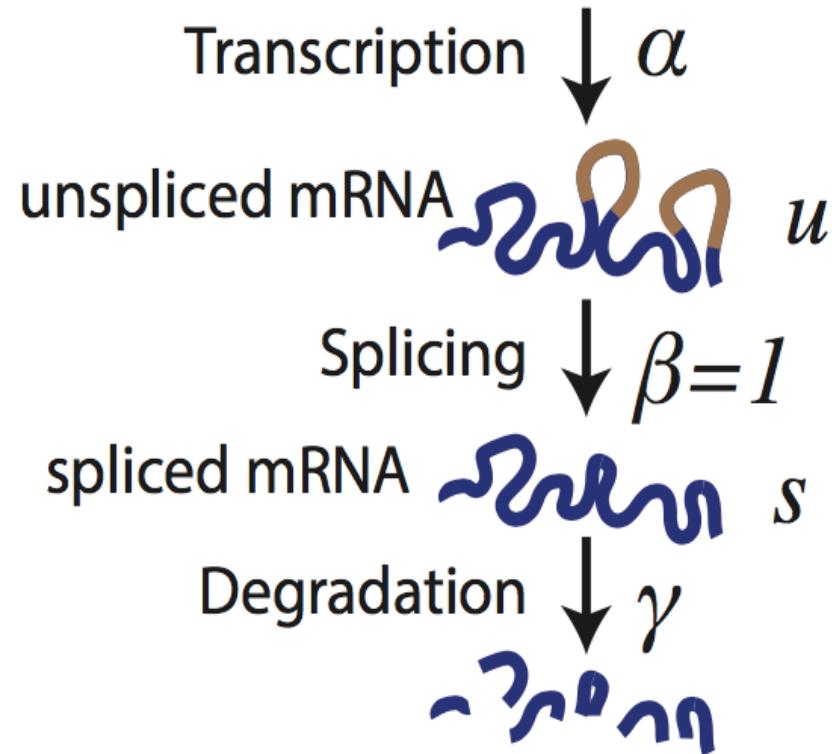
An example where pseudotime can be misinterpreted



- Cell types are similar to rostral embryonic tissues
- Spatial patterning (rather than a temporal axis of variation)

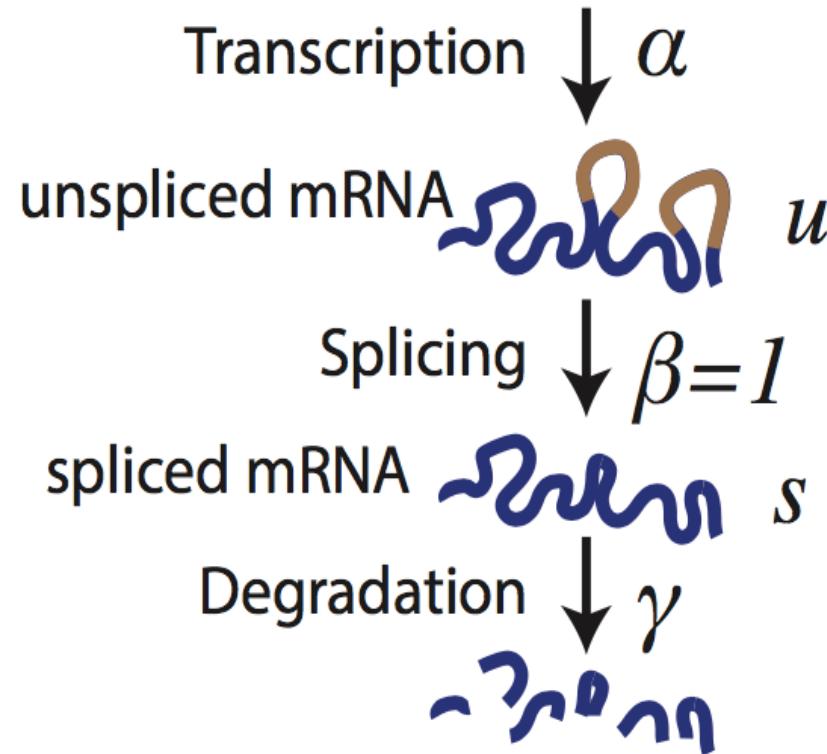
RNA velocity

RNA velocity can be estimated from RNA metabolism



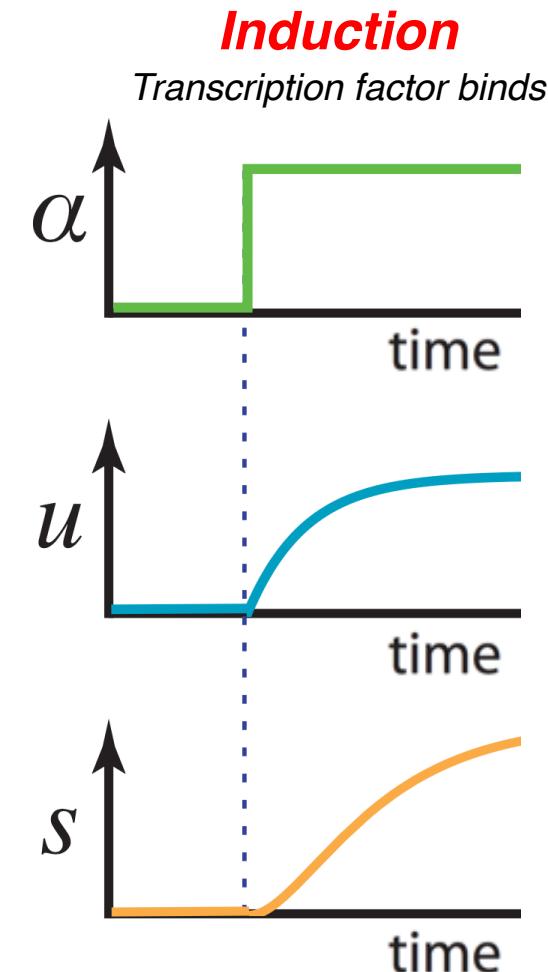
$$\frac{du}{dt} = \alpha - \beta u \quad \frac{ds}{dt} = \beta u - \gamma s$$

RNA velocity can be estimated from RNA metabolism

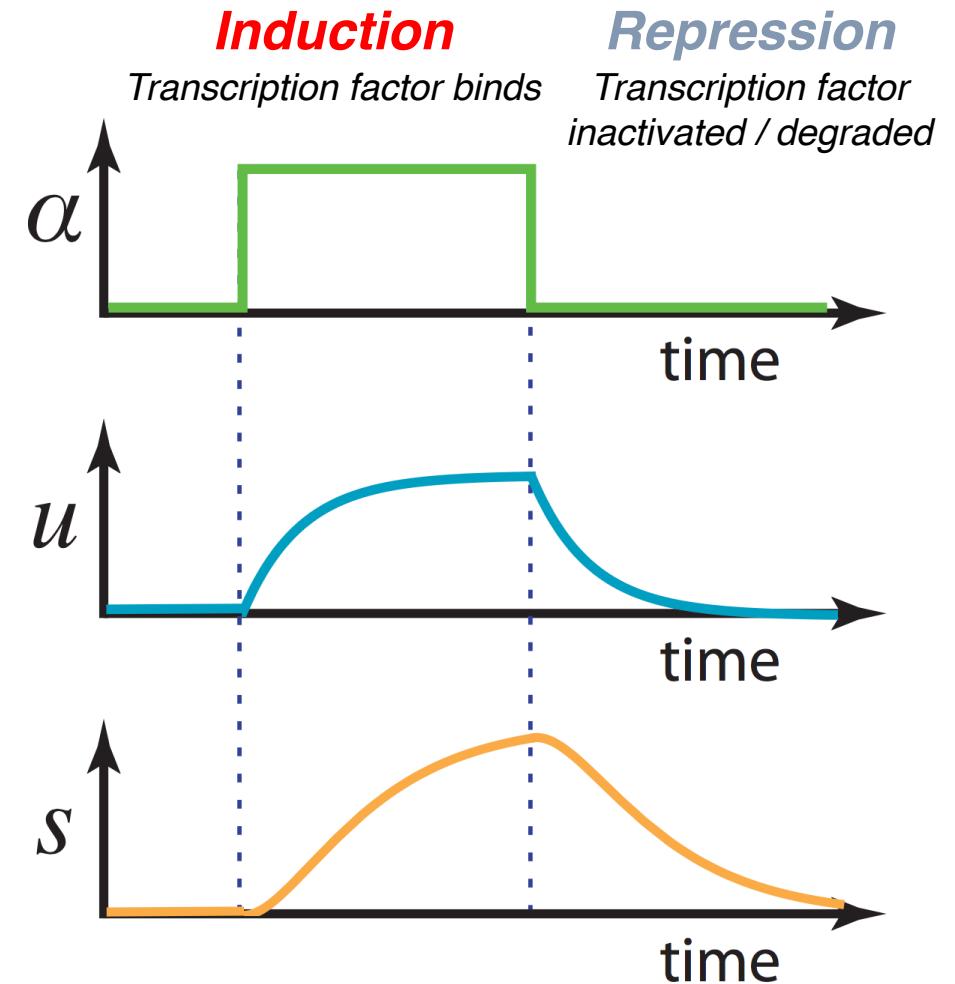
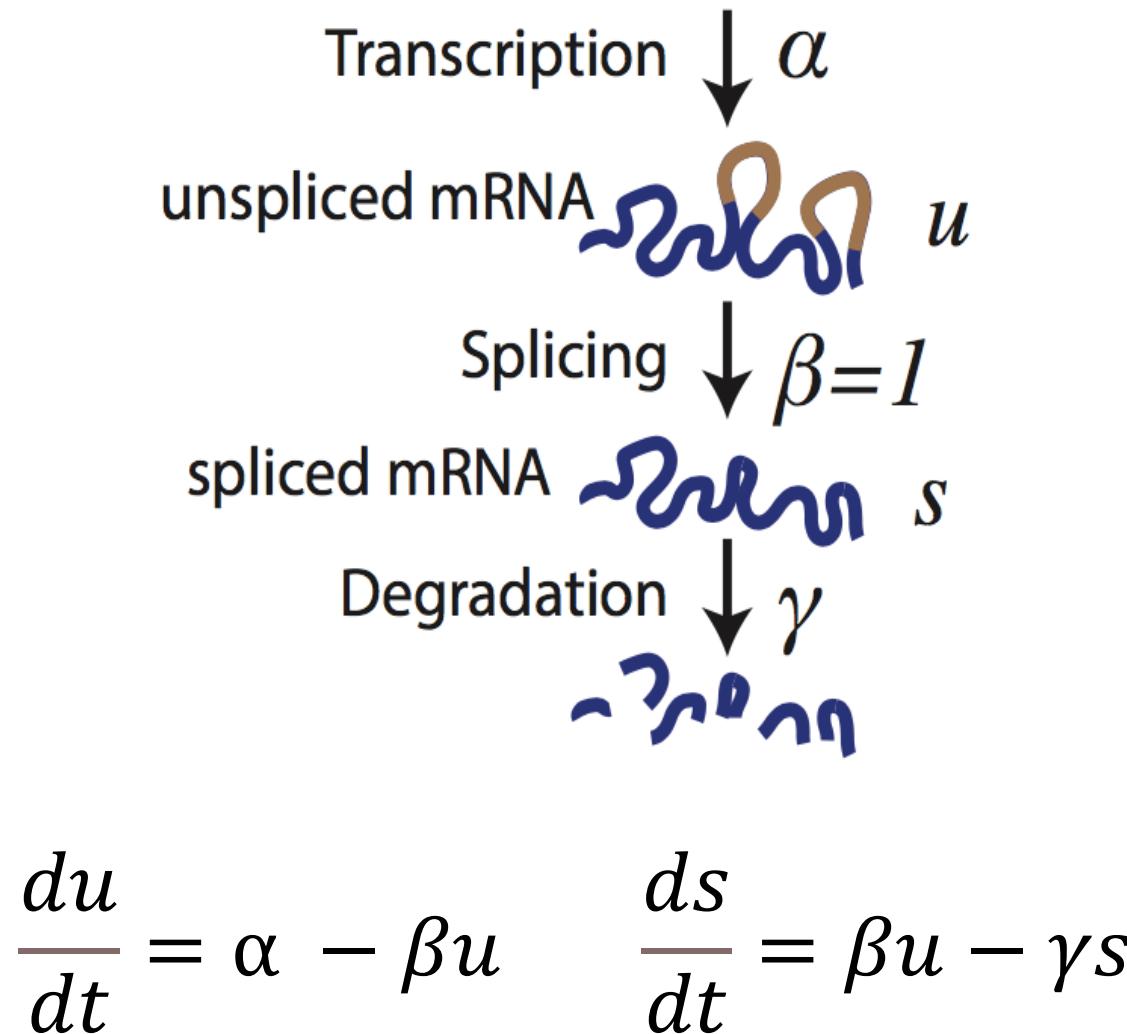


$$\frac{du}{dt} = \alpha - \beta u$$

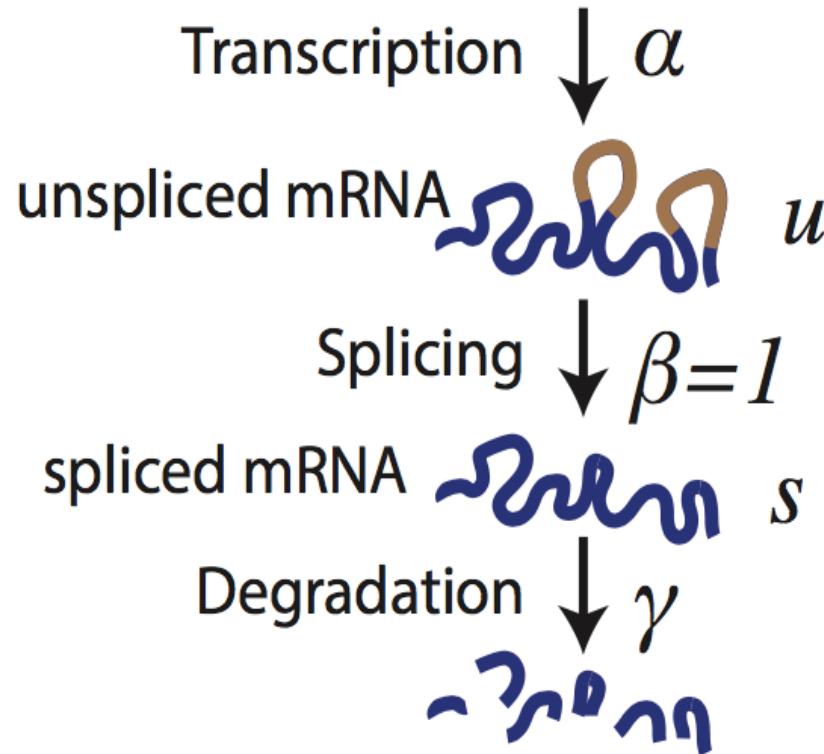
$$\frac{ds}{dt} = \beta u - \gamma s$$



RNA velocity can be estimated from RNA metabolism

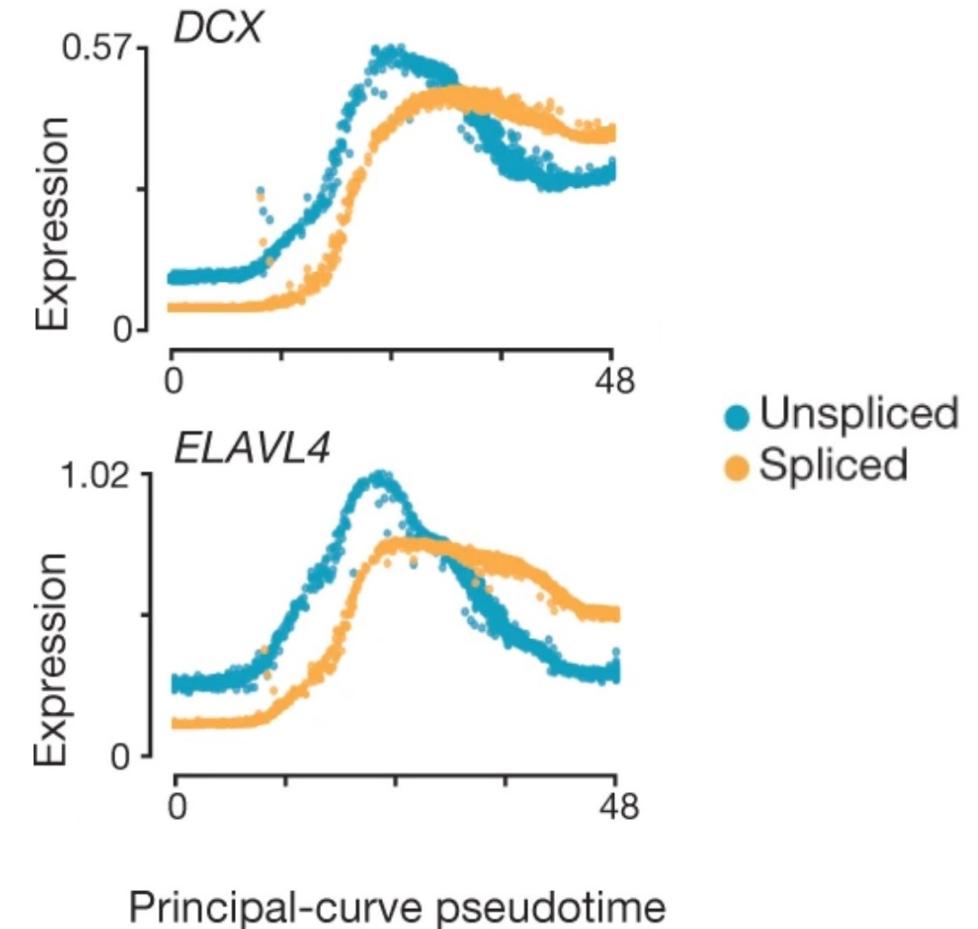


RNA velocity can be estimated from RNA metabolism

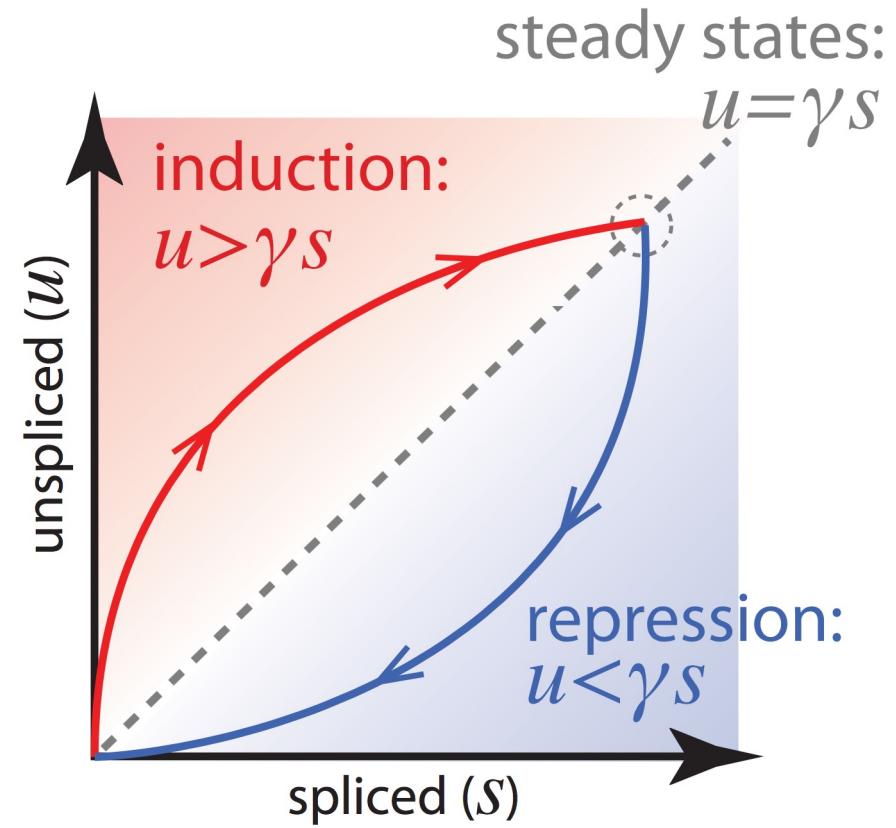
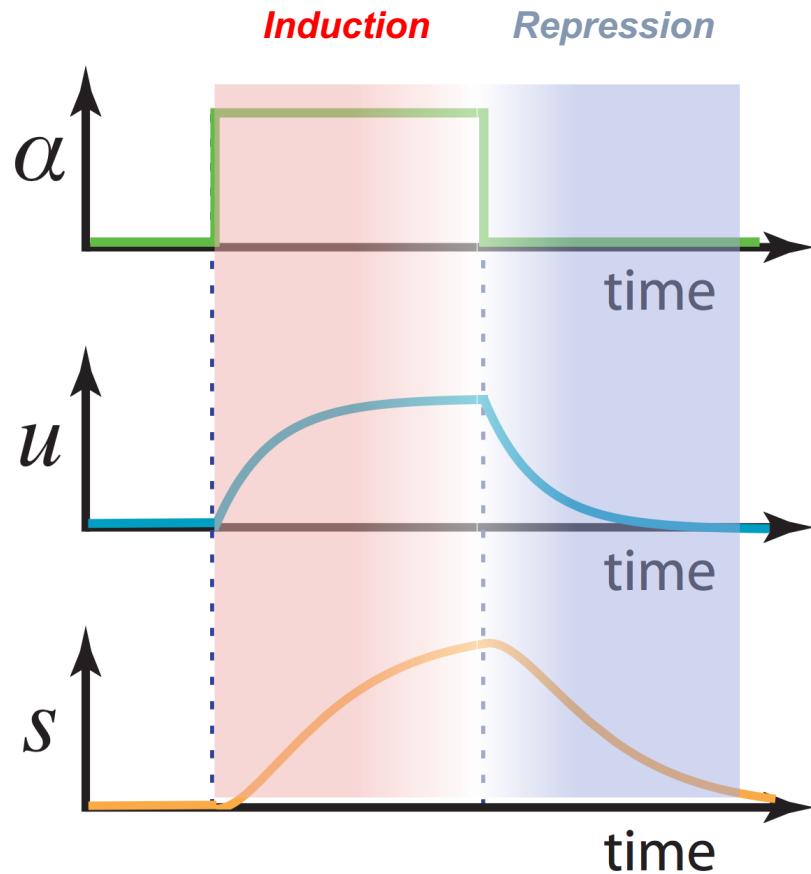


$$\frac{du}{dt} = \alpha - \beta u$$

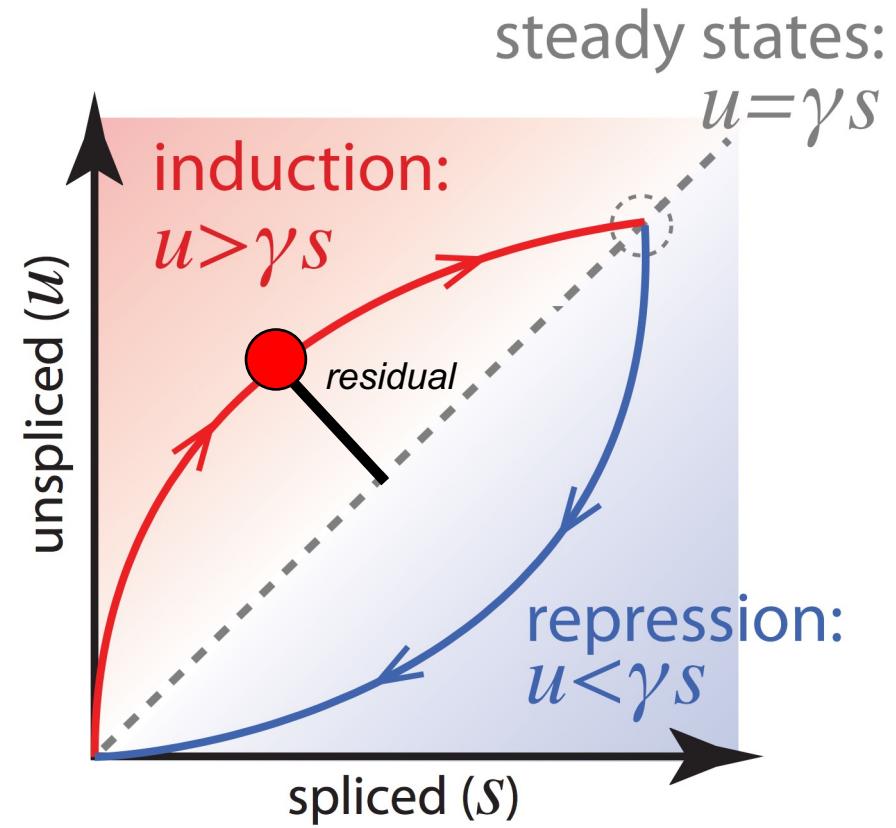
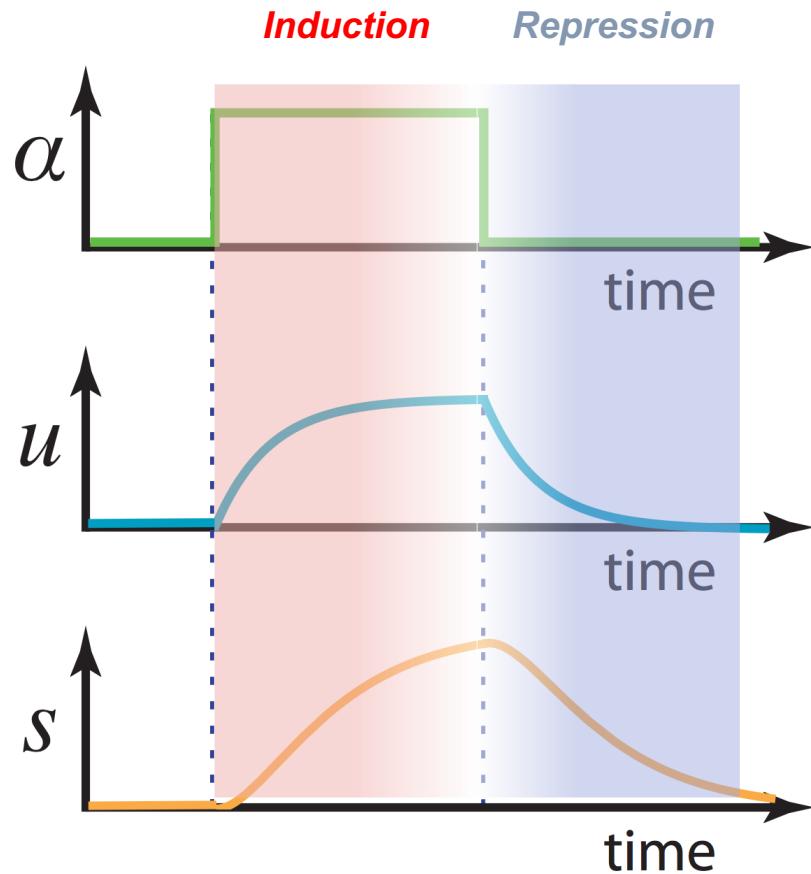
$$\frac{ds}{dt} = \beta u - \gamma s$$



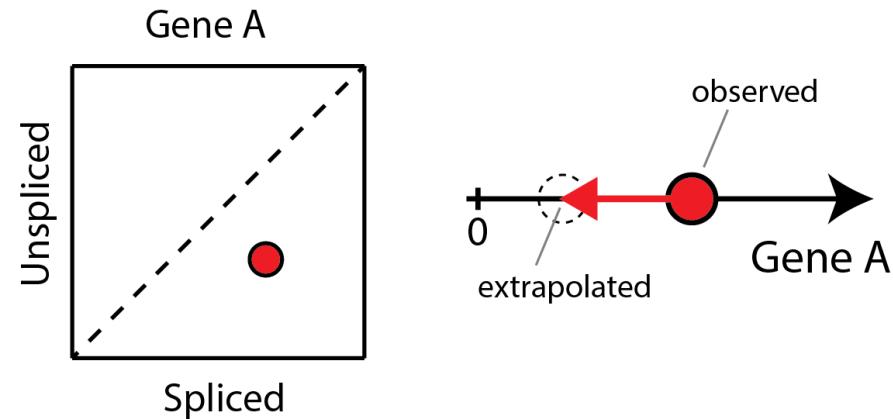
Phase portraits show changes to relative unsPLICED and sPLICED RNA abundances



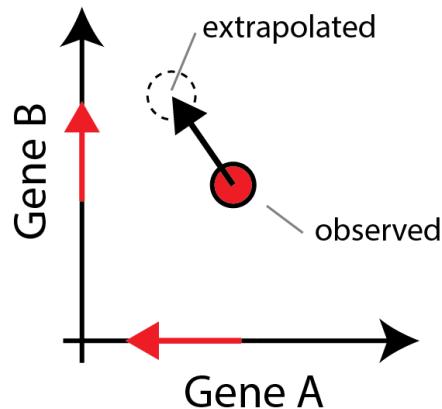
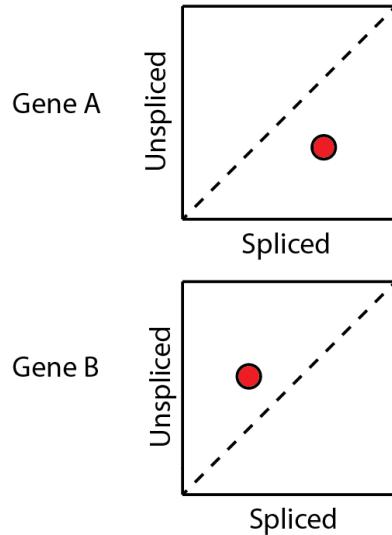
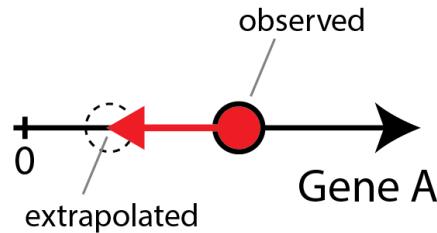
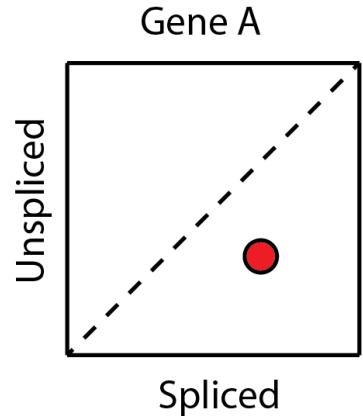
Phase portraits show changes to relative unsPLICED and sPLICED RNA abundances



RNA velocity is a gene-specific estimate

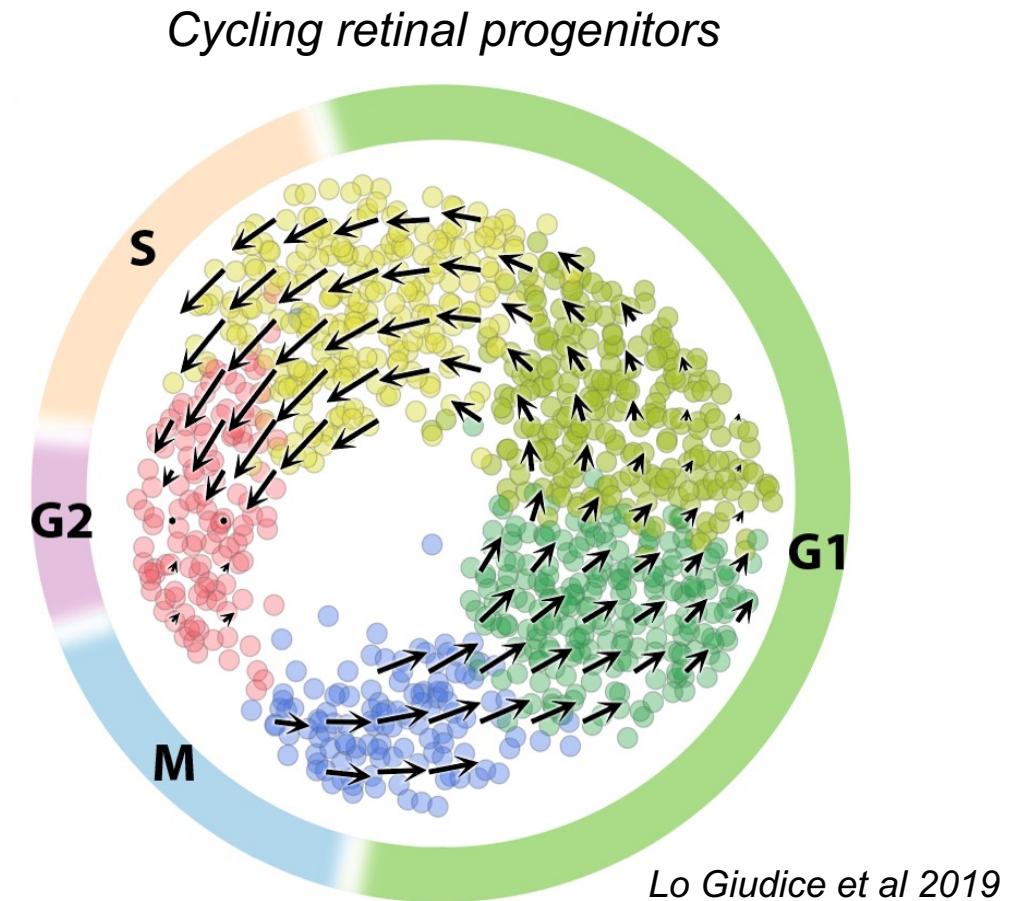
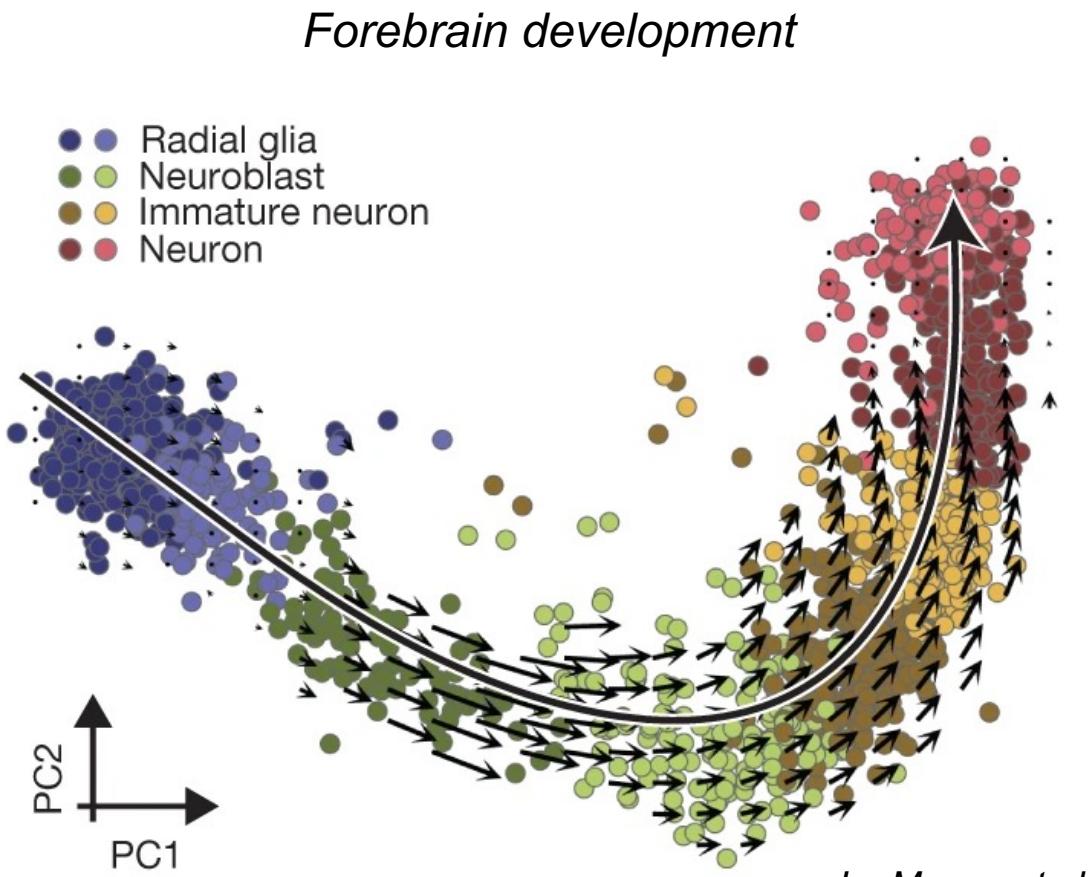


RNA velocity is a gene-specific estimate



Problem: each gene is on a different time scale!

Examples of RNA velocity applied to trajectories

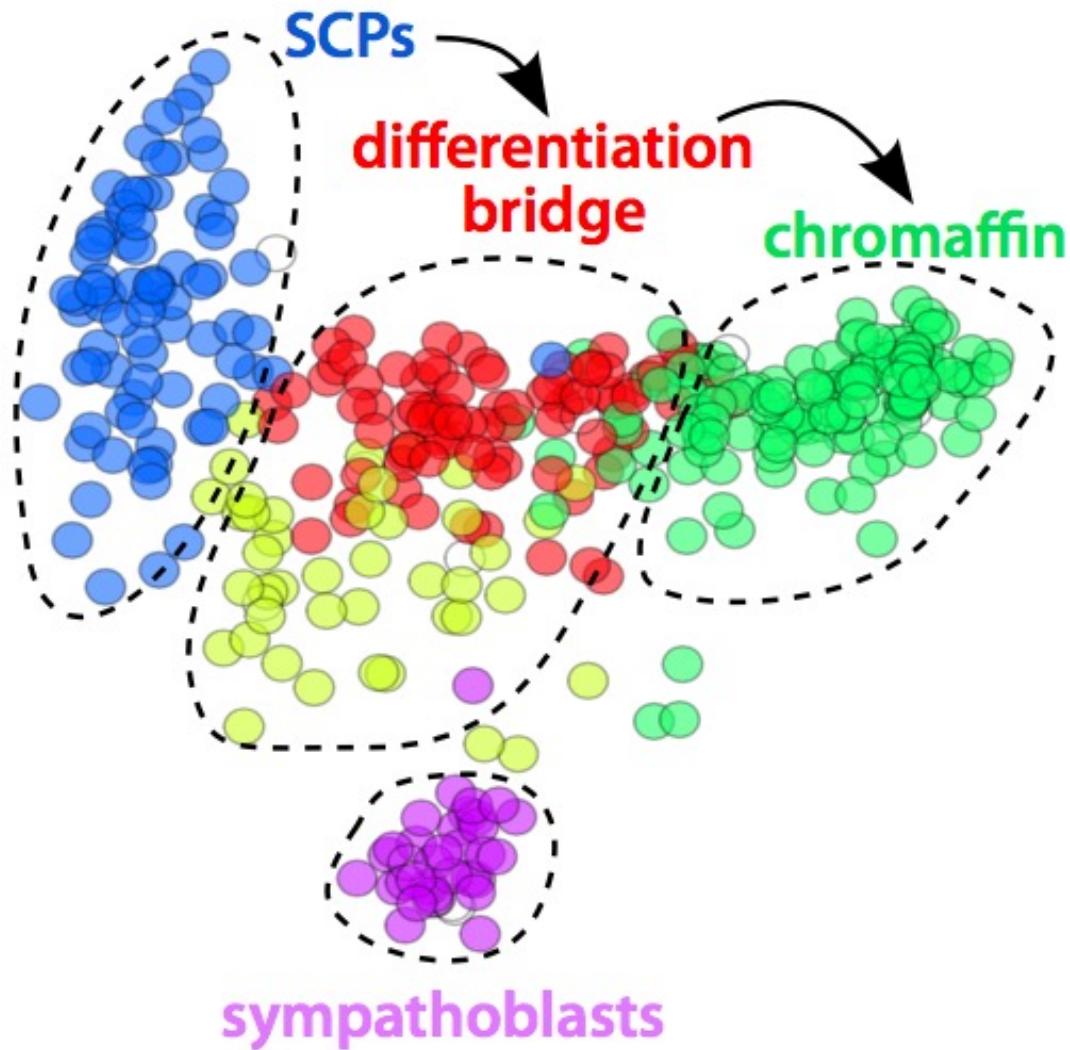


When should I use RNA velocity?

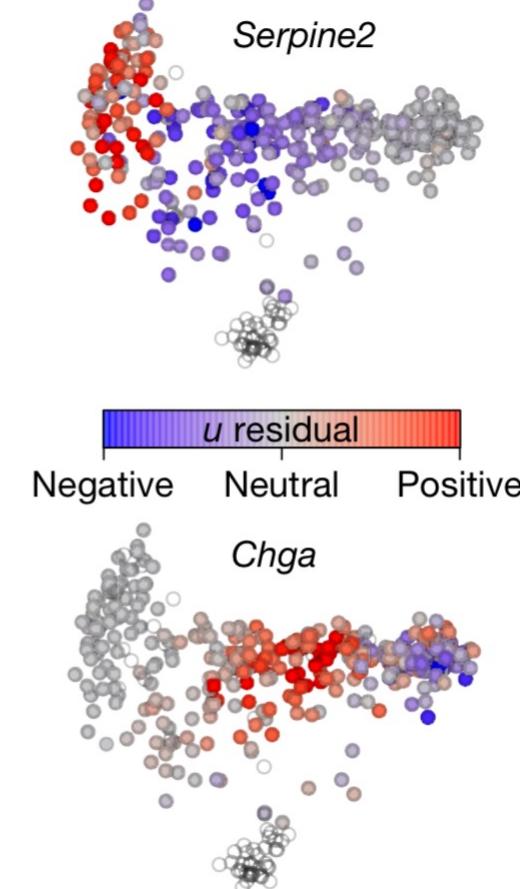
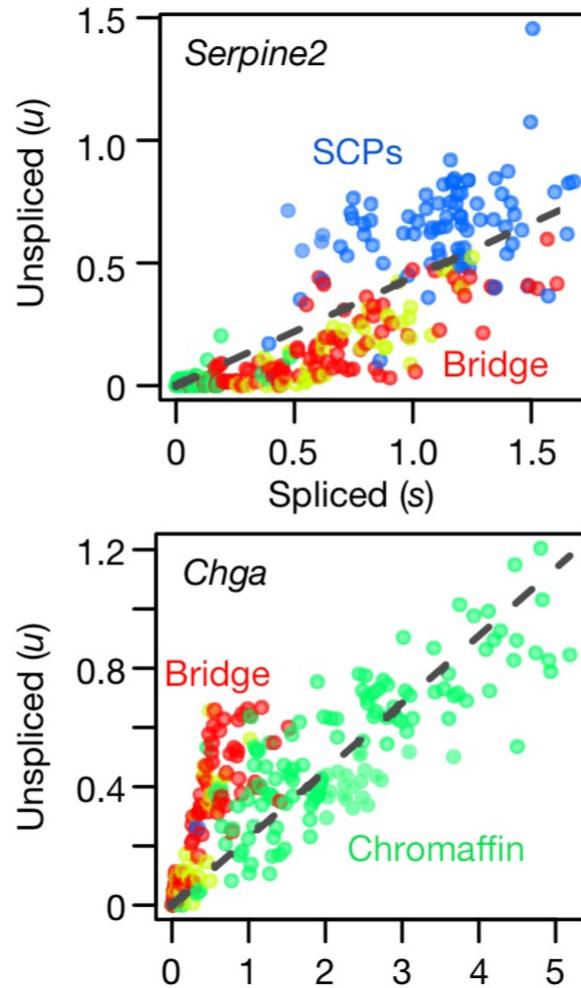
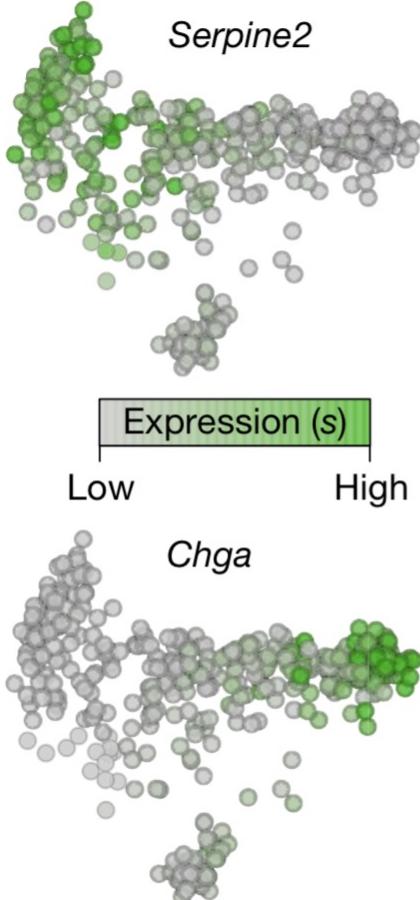
- I expect there to be a trajectory in my single-cell dataset, but:
 - I do not know the **direction** along which cells should change
 - I expect the **trajectories to change** for a similar population of cells but in different dataset conditions (in response to gene perturbations, environmental stimuli, etc)

Note: the changes between cell states should unfold on a relatively short time scale.

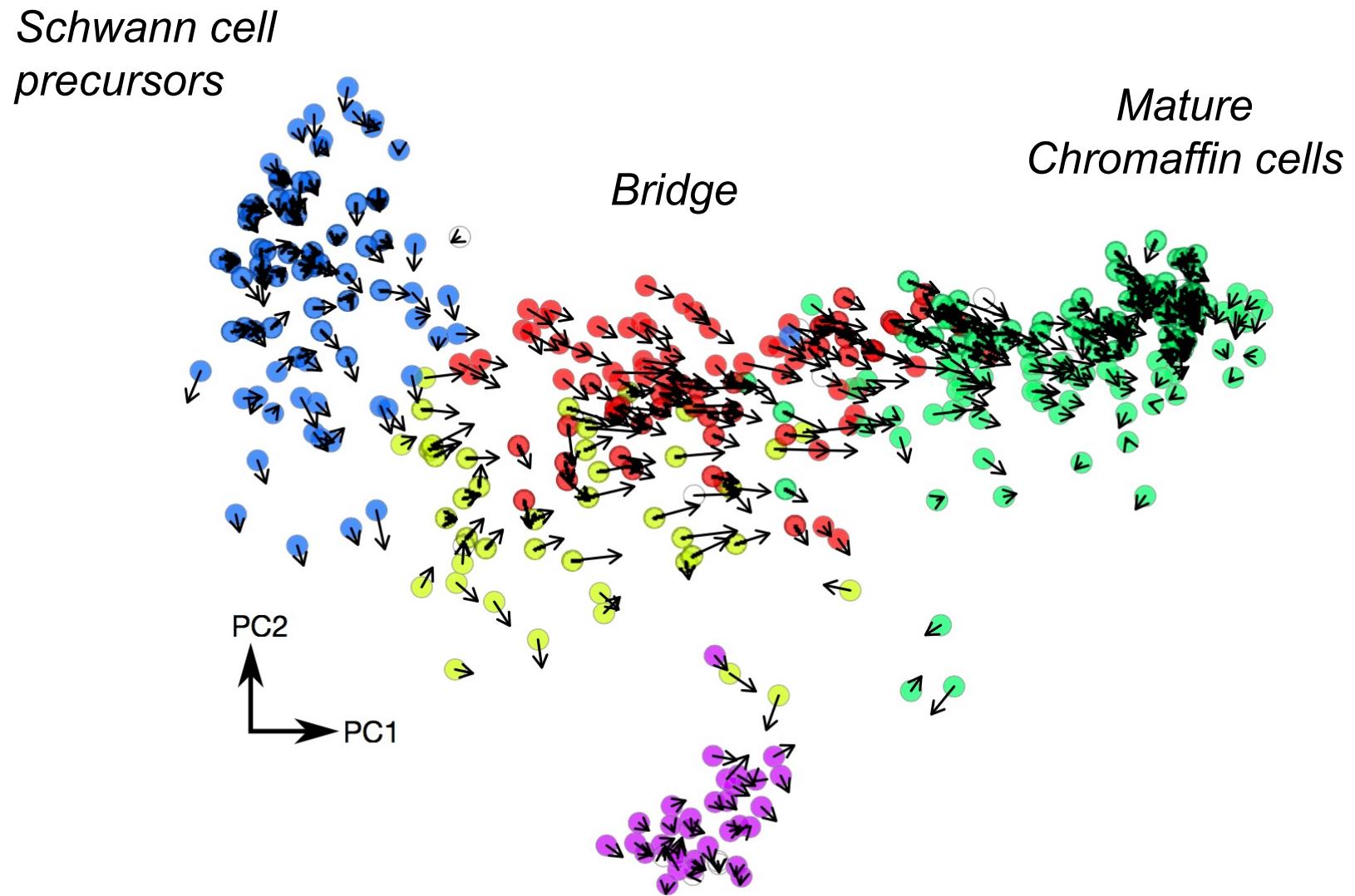
RNA velocity to describe a simple differentiation process



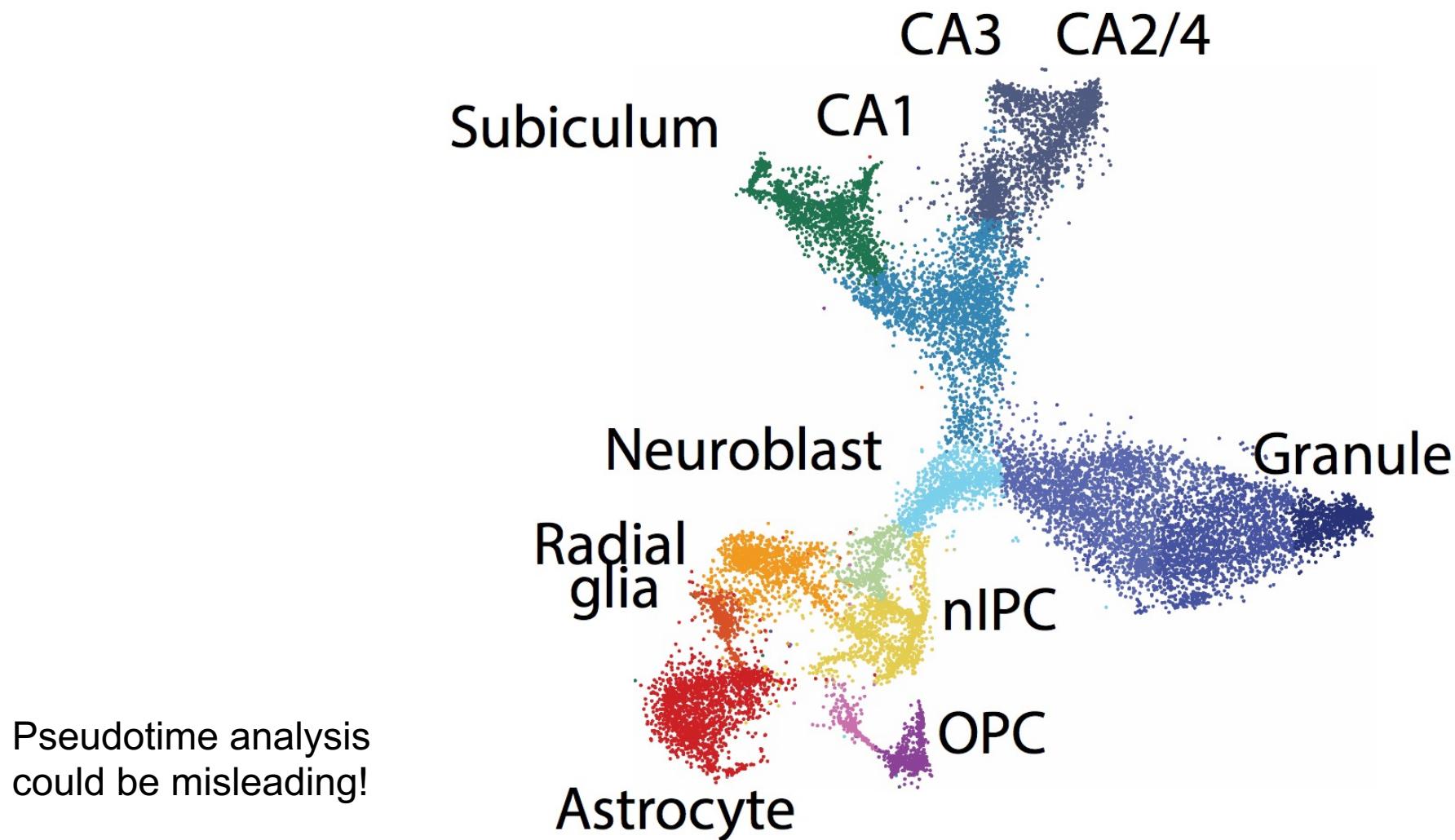
RNA velocity to describe a simple differentiation process



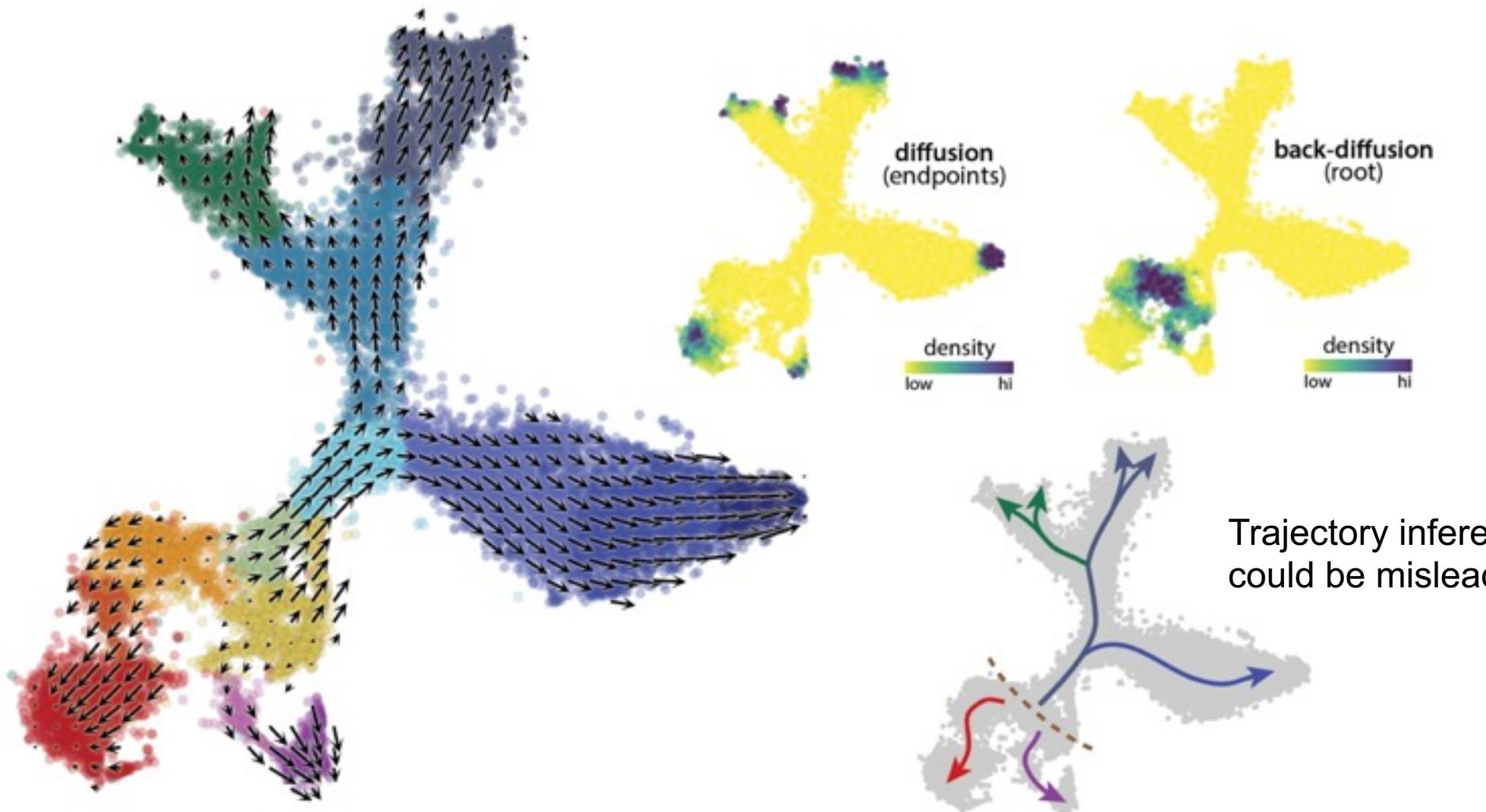
RNA velocity to describe a simple differentiation process



A more complex application of RNA velocity

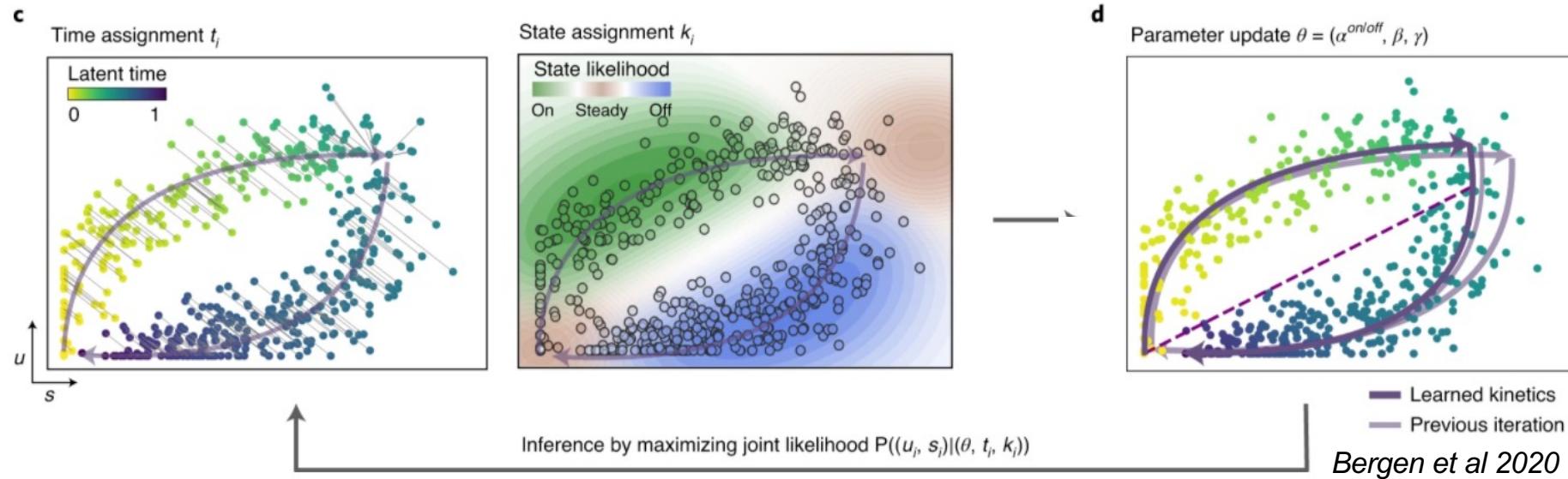


A more complex application of RNA velocity



Trajectory inference alone
could be misleading!

scvelo and other recent extensions to RNA velocity



Chromatin Potential Identified by Shared Single-Cell Profiling of RNA and Chromatin

Protein velocity and acceleration from single-cell multiomics experiments

Representation learning of RNA velocity reveals robust cell transitions

UniTVelo: temporally unified RNA velocity reinforces single-cell trajectory inference

A relay velocity model infers cell-dependent RNA velocity

Multi-omic single-cell velocity models epigenome-transcriptome interactions and improves cell fate prediction

DeepVelo: Single-cell transcriptomic deep velocity field learning with neural ordinary differential equations

Should you perform RNA velocity analysis on your data?

RNA velocity unraveled:

<https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1010492>



Lior Pachter
@lpachter

...

If you work w/ single-cell RNA-seq & are performing RNA velocity analyses, you might find this @GorinGennady et al. preprint w/ Meichen Fang & Tara Chari of interest. It's a deep dive into the method, and navigation of the 67 pages may be aided w/ this  1/



biorxiv.org

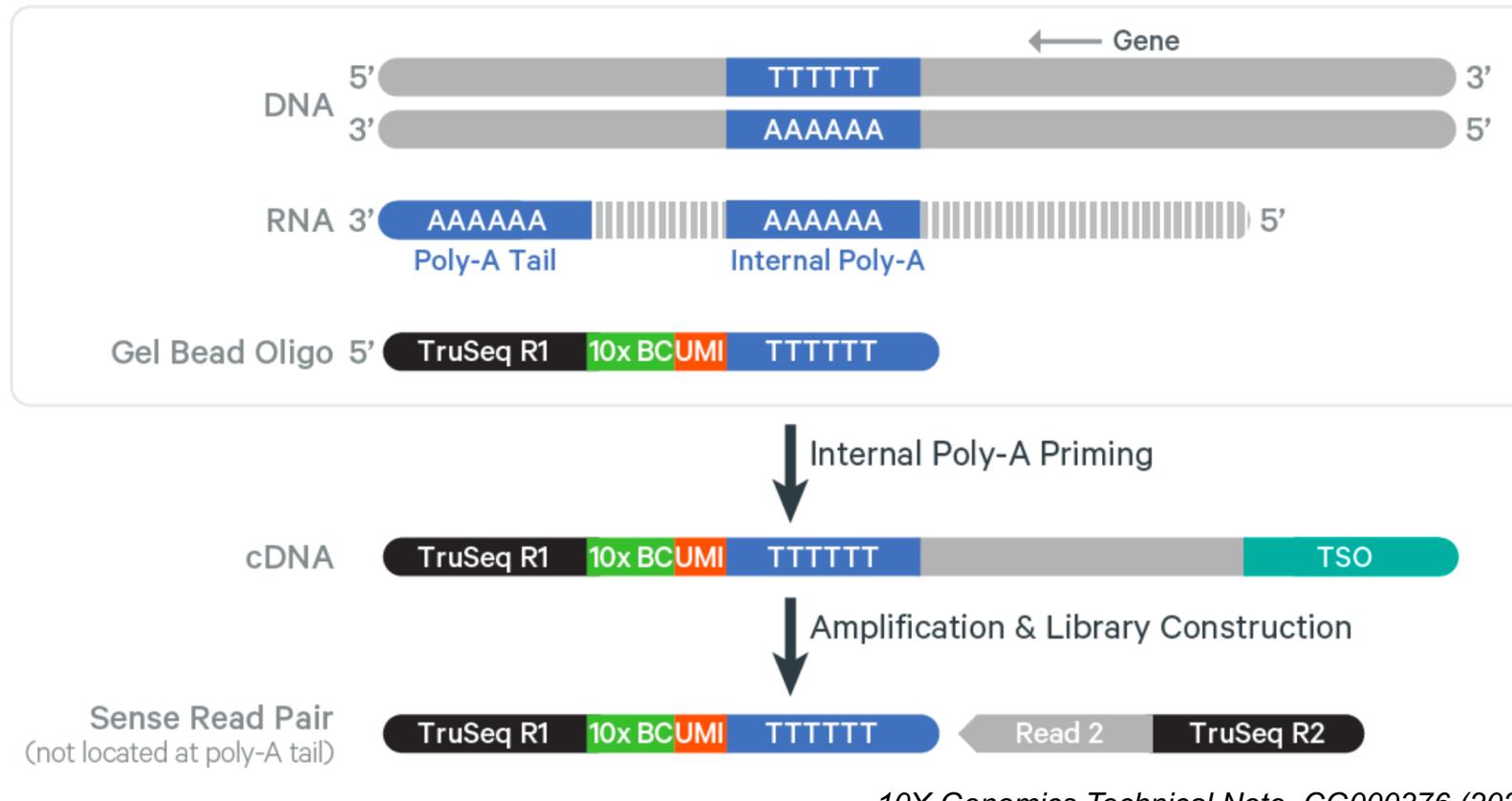
RNA velocity unraveled

We perform a thorough analysis of RNA velocity methods, with a view towards understanding the suitability of the ...

Should you perform RNA velocity analysis on your data?

- Do I expect a temporal axis of variation in your data?
- Do I expect the time scale of temporal change to unfold in hours/days or weeks?
- Which single-cell technology do you use (single cell vs single nuclei) and how much detection of intronic reads do you expect?

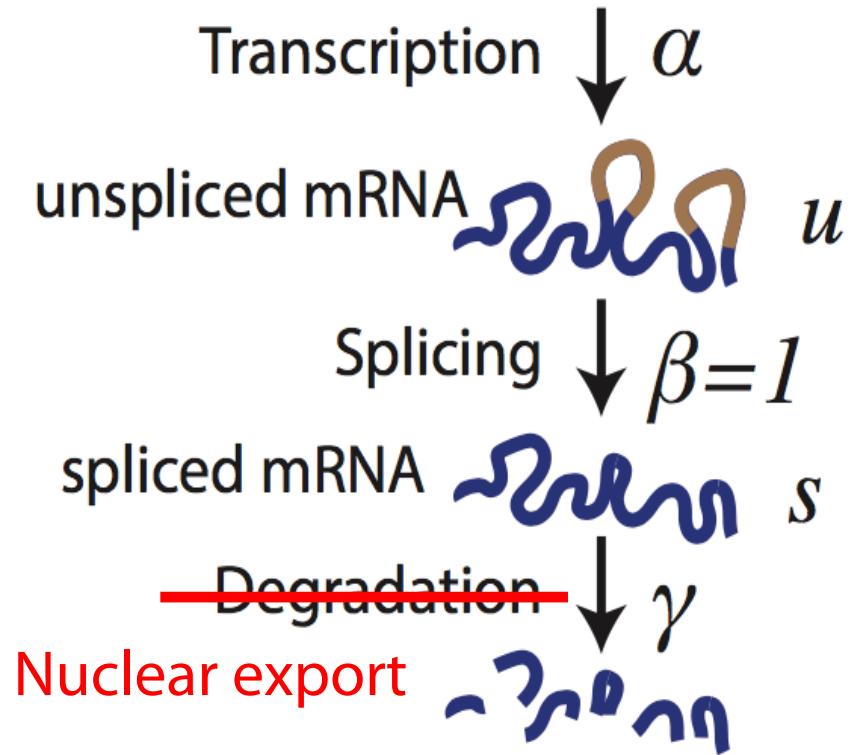
Internal priming and measurement of additional intronic regions



- The poly(dT) primer can prime to an internal poly-A stretch
- In humans, there are 21x more poly-A stretches found in introns than exons

Conclusion: there may be intron detection biases between different genes in the data!

The RNA life cycle captured by single nucleus RNA-seq is different



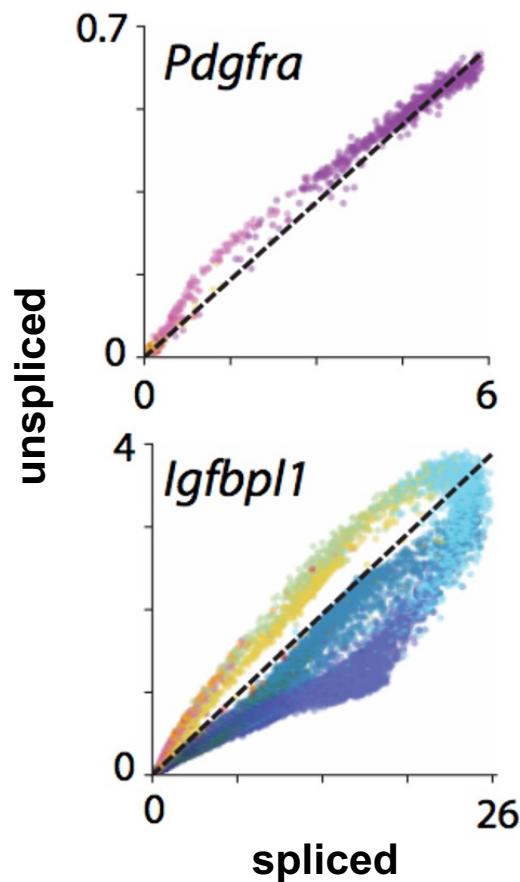
In single-nuclei data, it is not possible to measure the degradation rate for a gene.

The nuclear export rate may be instead described by changes to spliced abundance.

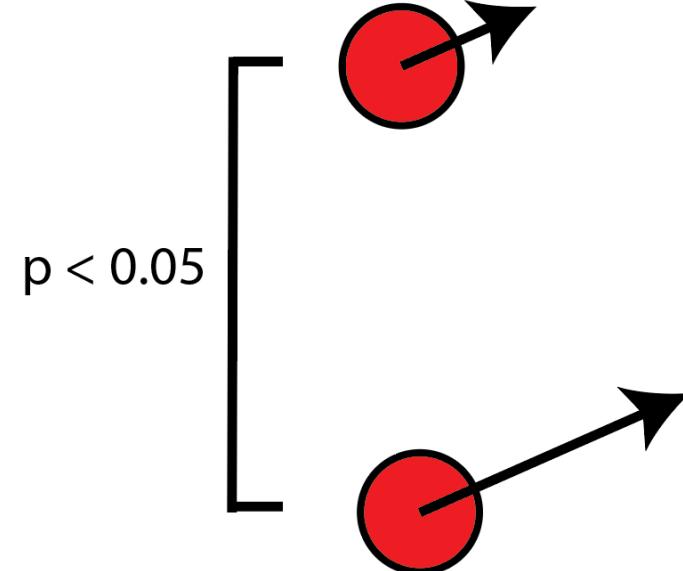
Nuclear export occurs much more quickly than the degradation – be careful!

Ongoing limitations of RNA velocity

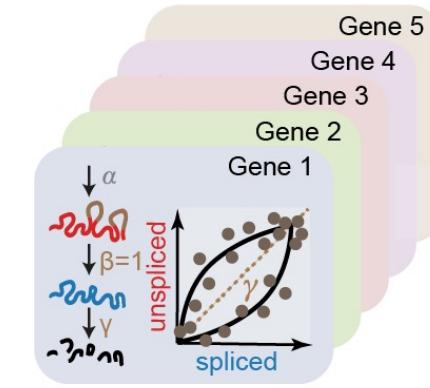
Necessity to run
smoothing methods



Inference on velocity
(i.e. differential test)
is currently not possible

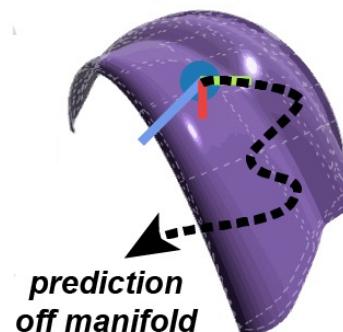


Gene-wise velocity estimation
yields incorrectly scaled components



gene-wise kinetics

Incorrectly-scaled
velocity components



*prediction
off manifold*

Two steps for RNA velocity analysis

Step 1. Intron/exon counting

- Command line tool that takes the output bam files from CellRanger and generates separate count matrices for spliced and unspliced mRNA
- Original tool (*velocyto*): <https://velocyto.org/velocyto.py/tutorial/cli.html>
- Also possible with STARsolo and alevin-fry methods for read alignment

Step 2. Velocity estimation and visualization

Most widely-used tools are:

- *velocyto* (<https://velocyto.org/velocyto.py/tutorial/analysis.html>)
- *scvelo* (<https://scvelo.readthedocs.io/>)

Most RNA velocity tools are implemented and better-supported in Python.

- The original “*velocyto*” implementation of RNA velocity analysis in R:
<https://github.com/velocyto-team/velocyto.R>
- Use the reticulate package to run python code in R Studio environment. A detailed tutorial for doing this for *scvelo* package is available here:
<https://jef.works/blog/2020/08/25/using-scvelo-in-R-using-reticulate/>

Question