

Single-cell RNA-seq Analysis in Python

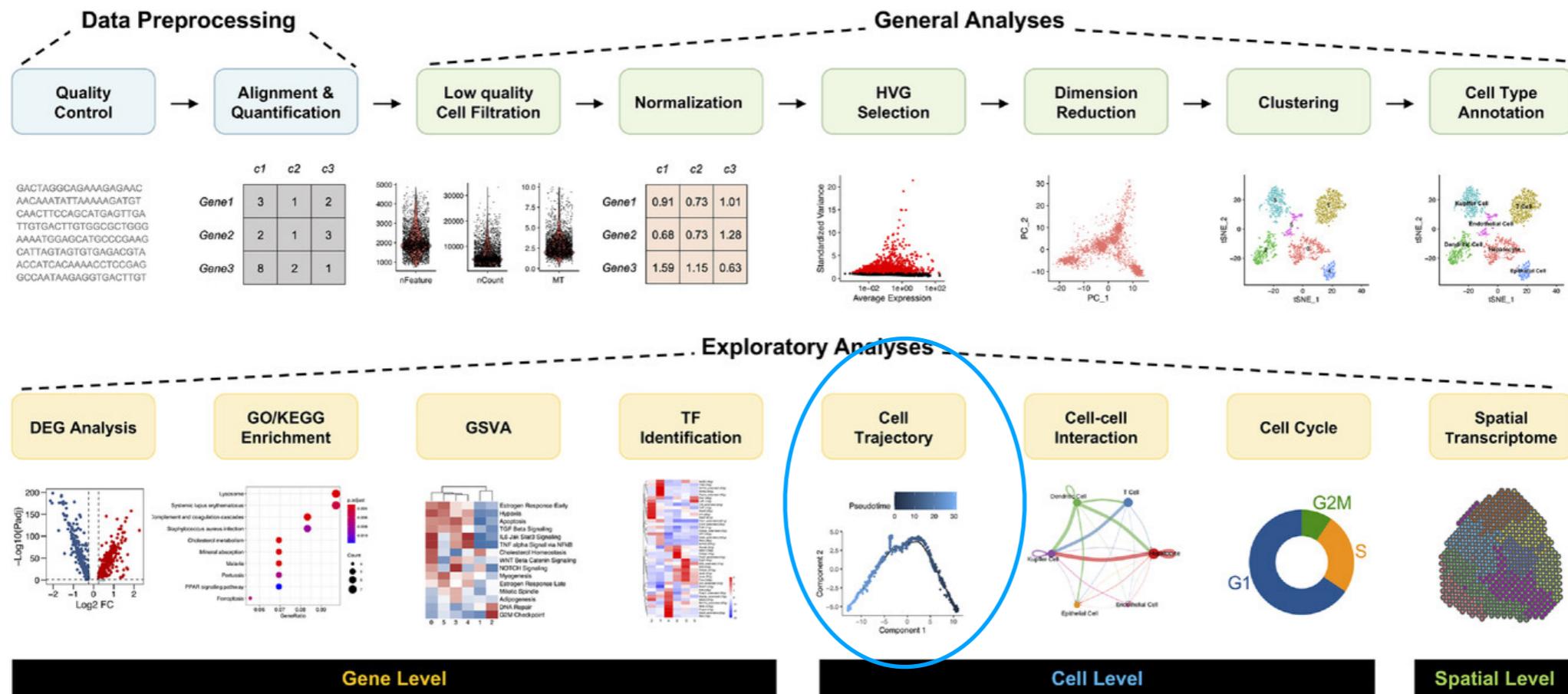
June 2024

Overview of Day #4

Cellular Dynamics and Differentiation

- Pseudotime analysis
- RNA velocity

Roadmap



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).

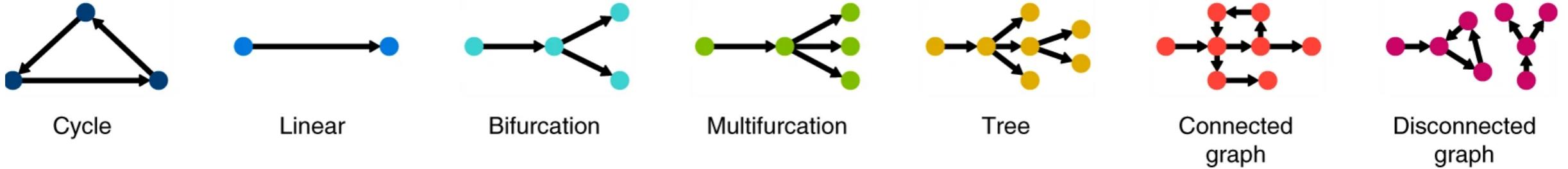
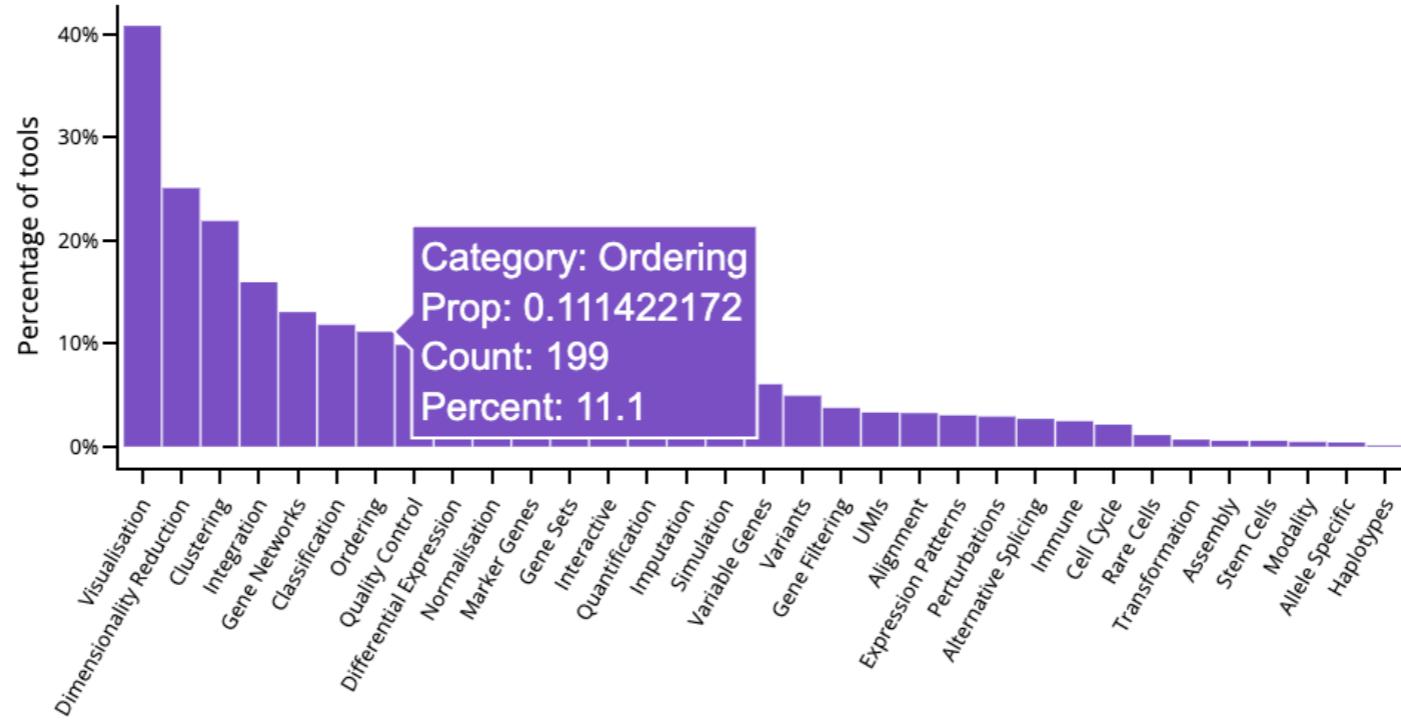
Should you run trajectory inference?

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

Trajectory Inference

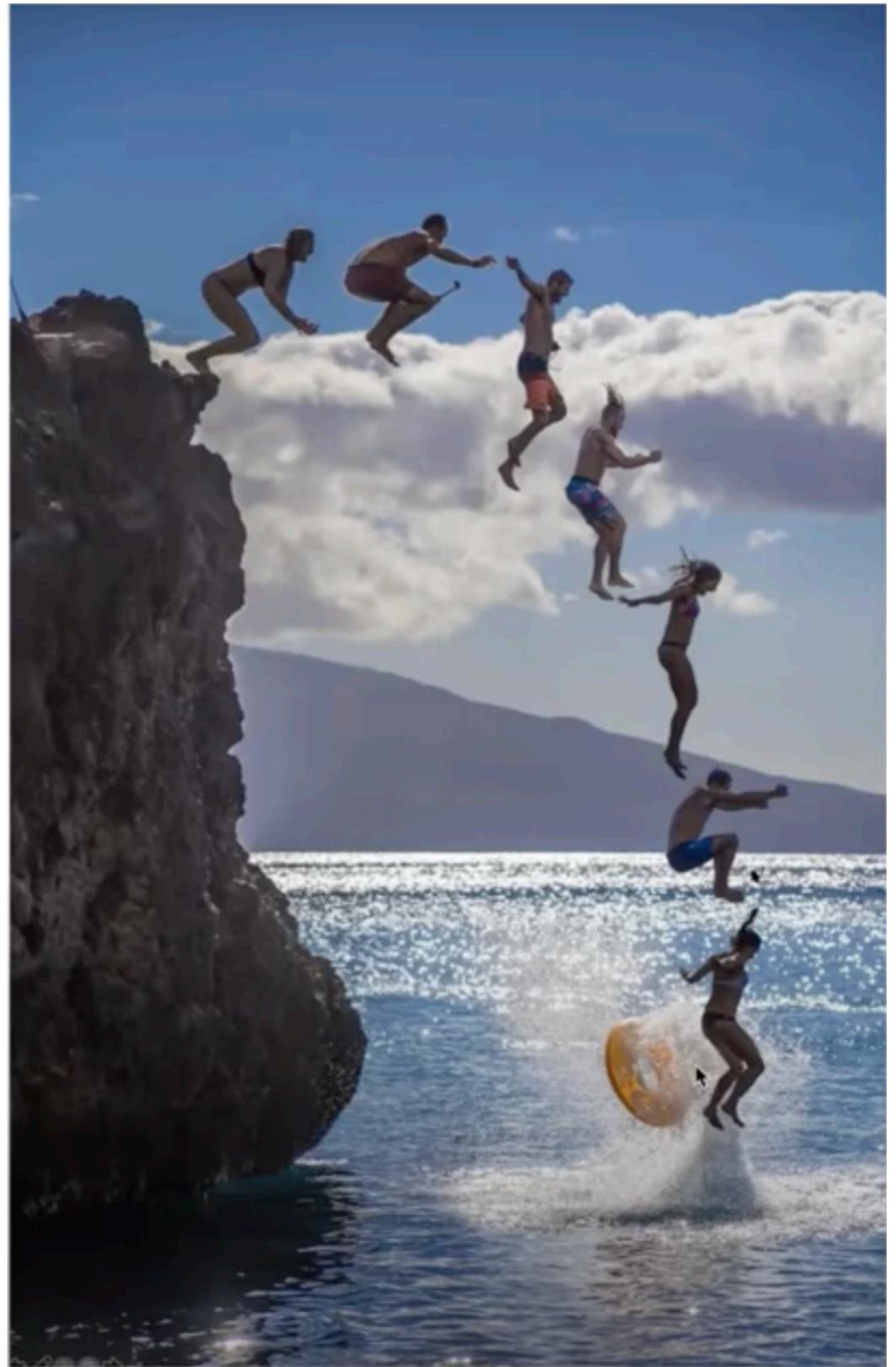


Trajectory Inference

- **The graph approach:** defines graph based on clusters and order the clusters (e.g. by estimating the connections between them in PAGA)
- **Manifold-learning-based approach:** uses principal curves or graphs to estimate the underlying trajectories (e.g. Slingshot)
- **Probabilistic approach:** assigns transition probabilities to ordered cell-cell pairs showing how likely it is that the reference cell is the ancestor of the other cell (e.g. DPT and Palantir)

RNA Velocity

ScRNA-seq provides only static snapshots of cellular states at the moment of the measurement.

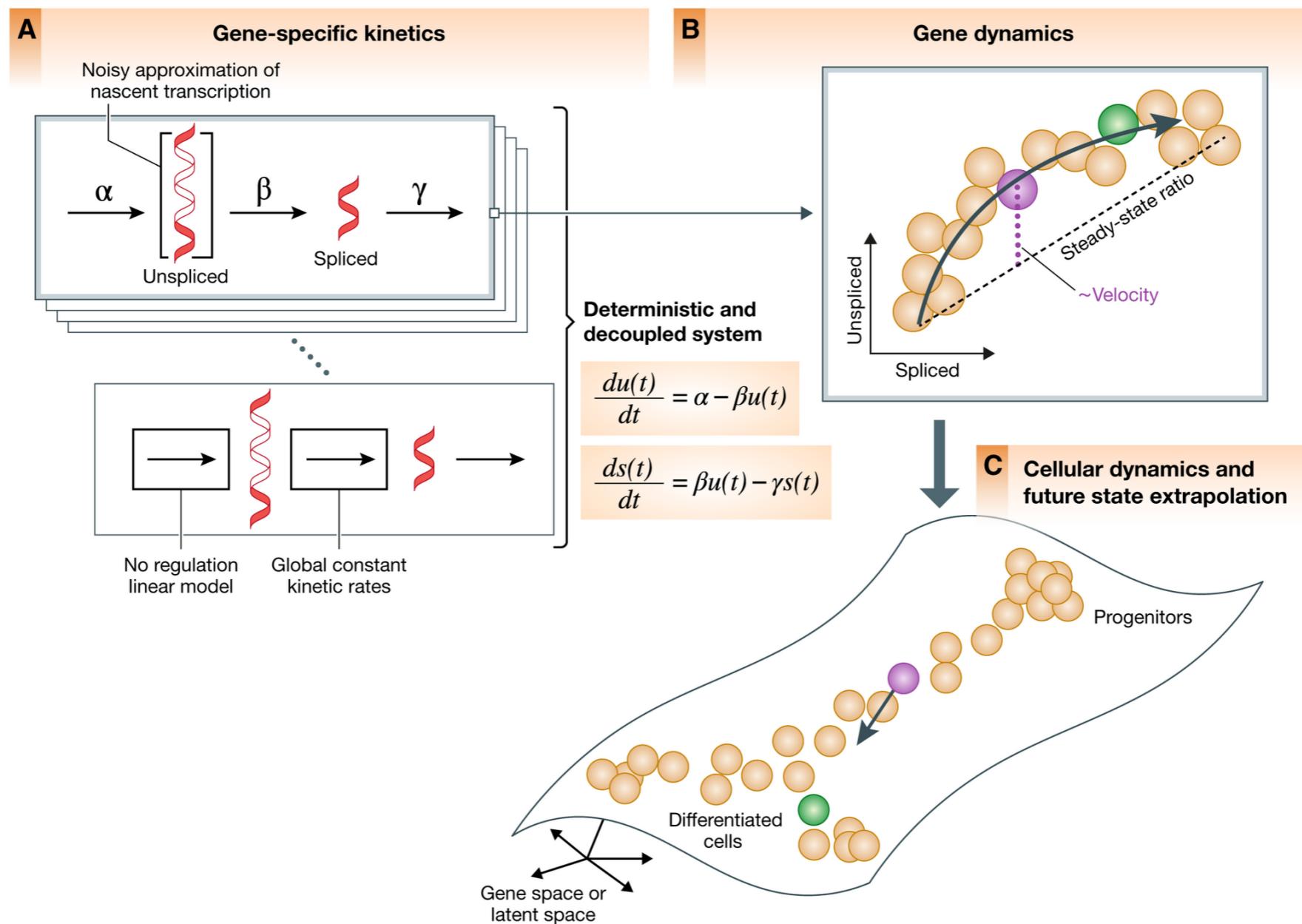


RNA Velocity

ScRNA-seq provides only static snapshots of cellular states at the moment of the measurement.



RNA Velocity



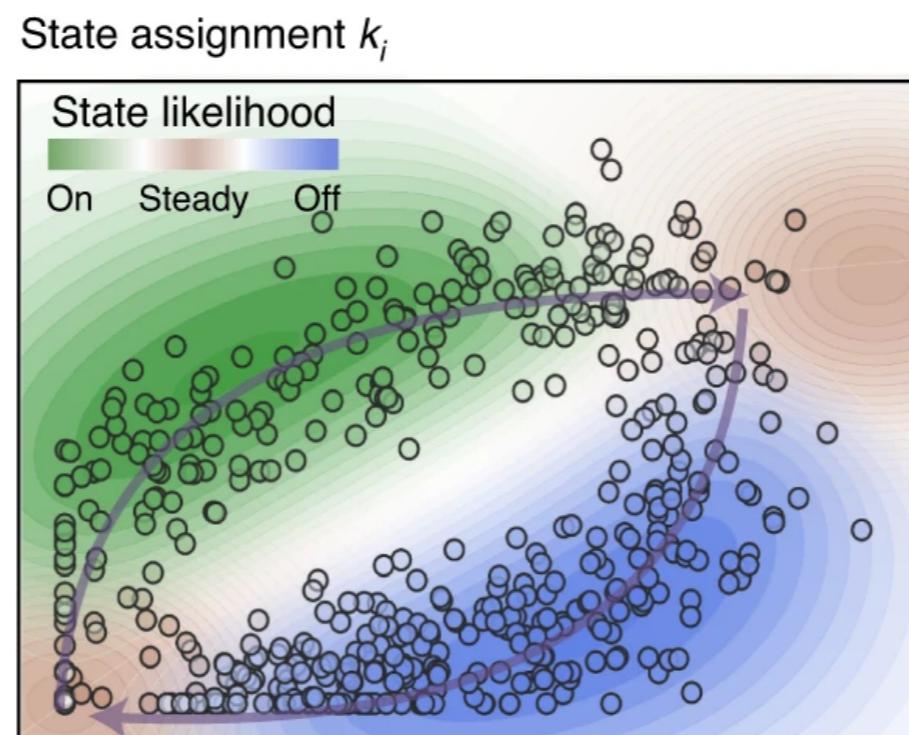
RNA Velocity

Steady-state model

- kinetics reached their equilibrium
- rates are constant
- there is a single, common splicing rate across all genes.

EM model

- all datapoints are used to infer the full set of parameters
- expectation-maximization (EM) framework to estimate parameters



RNA Velocity

- RNA velocity is a high-dimensional vector that predicts the future state of individual cells on a timescale of hours
- aids the analysis of developmental lineages and cellular dynamics
- Calculate the relative abundance of nascent (unspliced) and mature (spliced) mRNA to estimate the rates of gene splicing and degradation
- During a dynamic process:
 - increase in the transcription rate => rapid increase in unspliced mRNA => increase in spliced mRNA until a new steady state is reached.
 - A drop in the rate of transcription => drop in unspliced mRNA => reduction in spliced mRNA
- During induction of gene expression => unspliced mRNAs are present in excess
- During repression => unspliced mRNAs are present in lower amounts
- The balance of unspliced and spliced mRNA abundance is, therefore, an indication of the future state of mature mRNA abundance, and thus the future state of the cell

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/>)

Set up the environment

Create a new conda environment

In terminal (for Mac users) or in Anaconda Prompt miniconda3 (for Windows users):

```
conda create -n oldsc python=3.10
```

```
conda activate oldsc
```

```
pip install -r requirements02.txt
```

When you want to deactivate an activate conda use the following command

```
conda deactivate
```