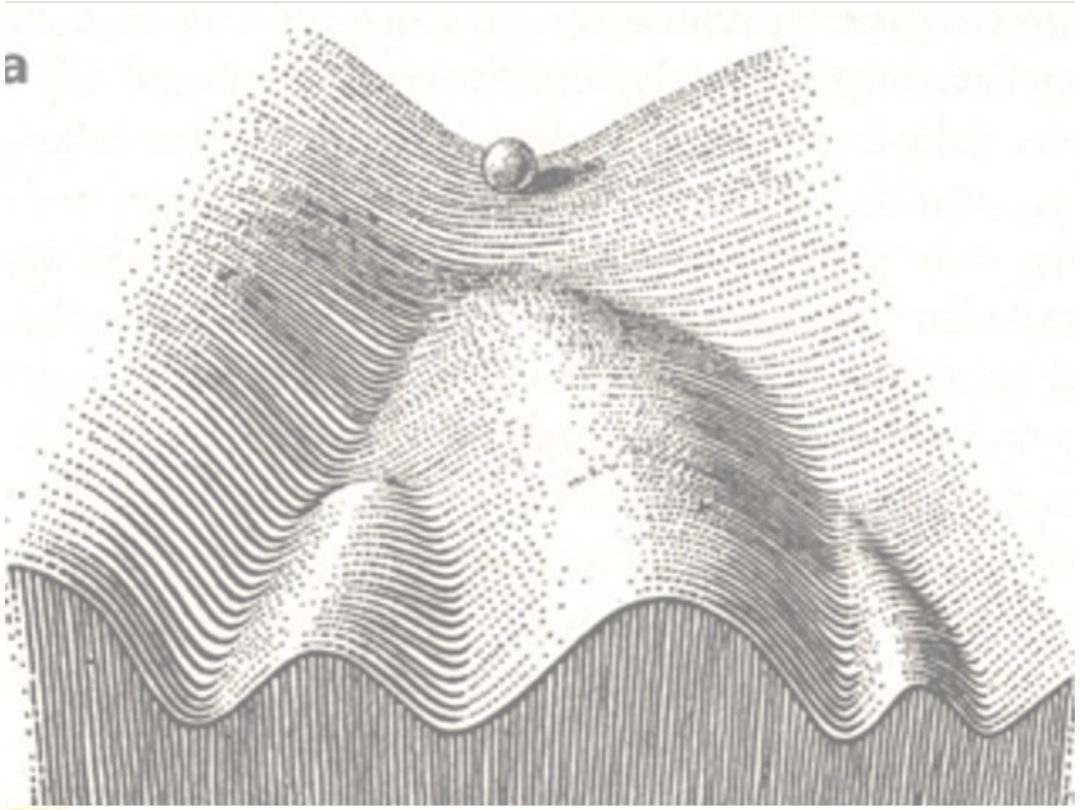
Generalizing RNA velocity to transient cell states through dynamical modeling

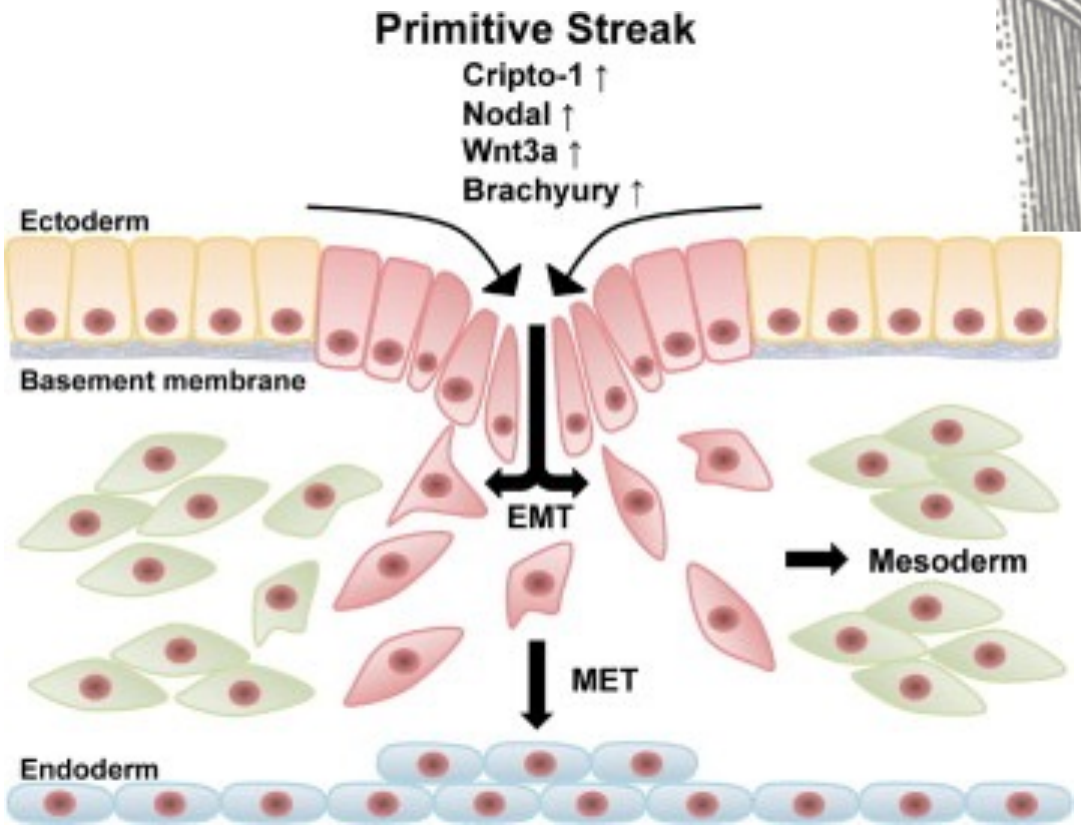
Bergen, V., Lange, M., Peidli, S. et al.

Presenter: Yiyu Pang

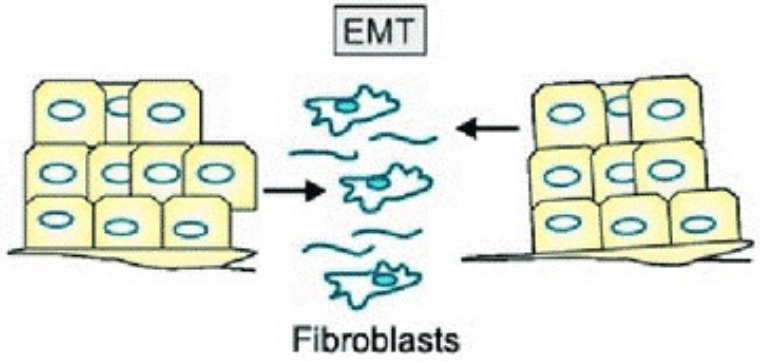
Cell fate decision



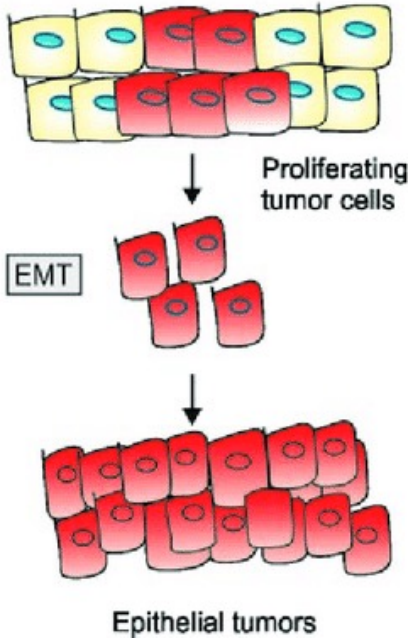
Embryonic development



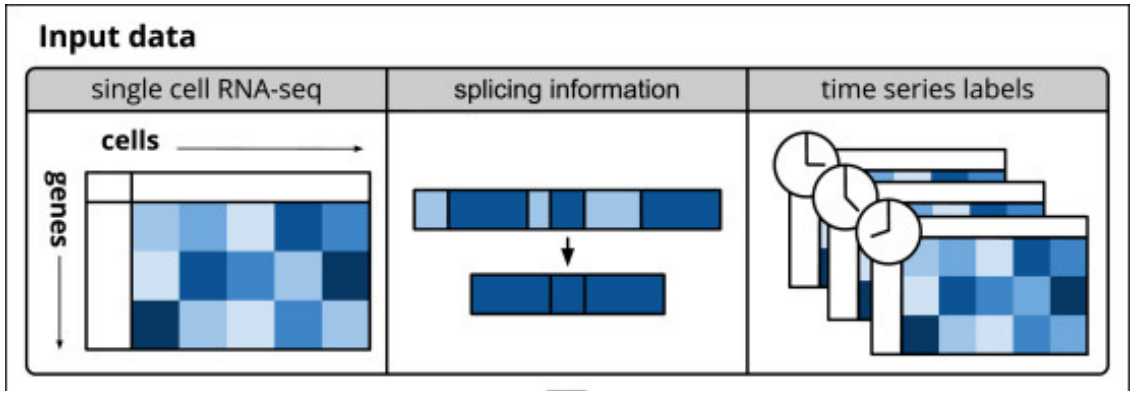
Tissue regeneration



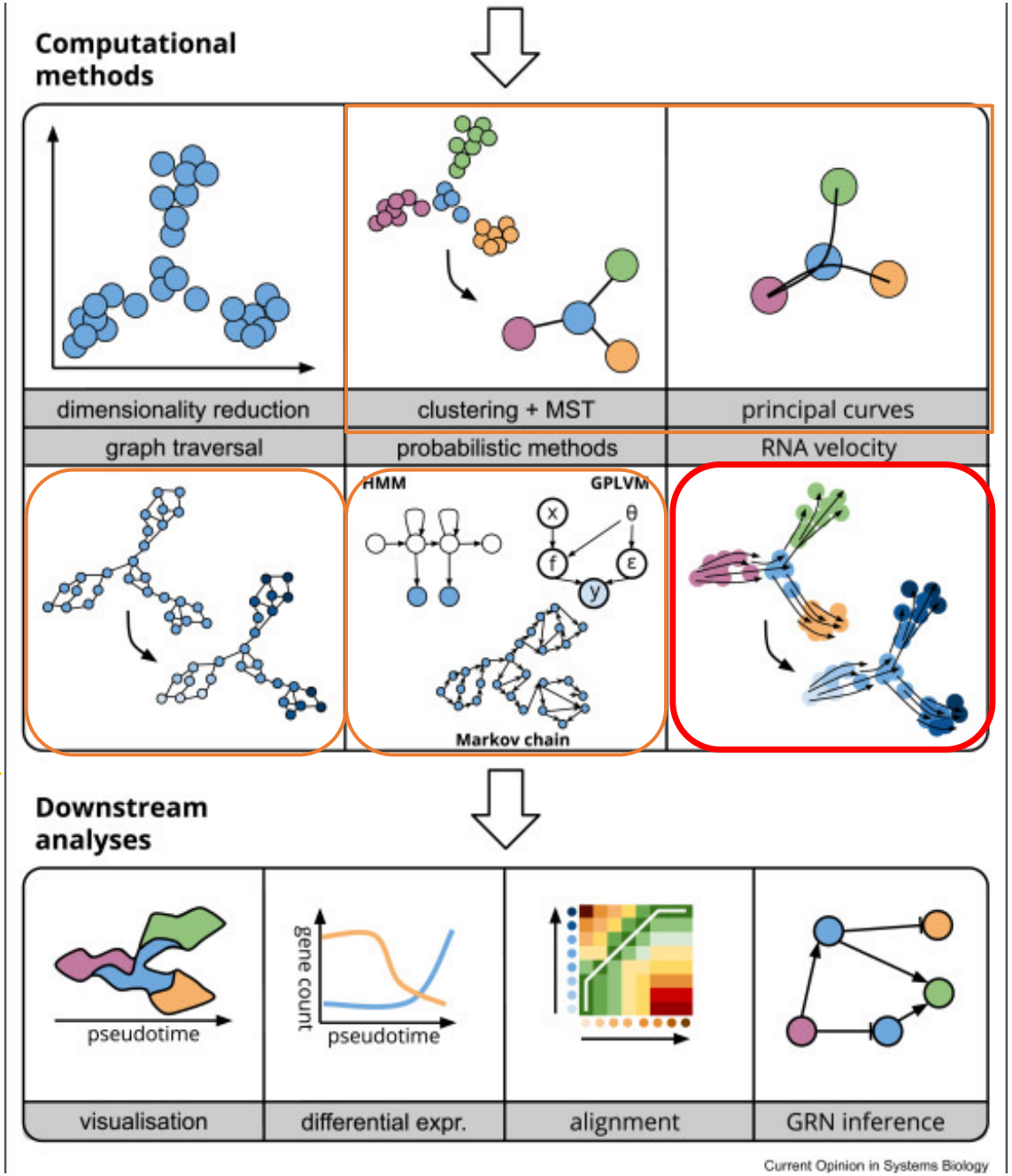
Cancer metastasiz



Cell trajectory inference



- MST: minimum weight spanning tree
- Principal curves: smooth one-dimensional curves
  - pass through the middle of a *p*-dimensional data set
  - providing a nonlinear summary of the data
- Graph traversal: the process of visiting (checking and/or updating) each vertex in a graph
  - Depth-first search
  - Breadth-first search
- Graph decomposition
- RNA velocity: an indicator of the future state of the cell



Deconinck, Louise, et al. "Recent advances in trajectory inference from single-cell omics data." Current Opinion in Systems Biology (2021)

Monocle

NATURE BIOTECHNOLOGY
Published: 23 March 2014

The dynamics and regulators of cell fate deci… revealed by pseudotemporal ordering of sing…

Cole Trapnell, Davide Cacchiarelli, Jonna Grimsby, Prapti Pokharel, Shuqiang Li, Michael… Lennon, Kenneth J Livak, Tarjei S Mikkelsen & John L Rinn

Nature Biotechnology **32**, 381–386 (2014) | Cite this article

65k Acc…

Genome Biology
Method | Open Access | Published: 19 March 2019

PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells

…a K. Hamey, Mireya Plass, Jordi Solana, Joakim S. Dahlin, Berthold Göttgens, …as Simon & Fabian J. Theis

…rticle number: 59 (2019) | Cite this article

…Citations | **136** Altmetric | Metrics

NATURE METHODS
Published: 29 August 2016

Diffusion pseudotime robustly reconstructs lineage branching

Laleh Haghverdi, Maren Büttner, F Alexander Wolf, Florian Buettner & Fabian J Theis

nature
Letter | Published: 08 August 2018

RNA velocity of single cells

Gioele La Manno, Ruslan Soldatov, Amit Zeisel, Emelie Braun, Hannah Lidschreiber, Maria E. Kastriti, Peter Lönnerberg, Alessandro Furlan, Jea… David van Bruggen, Jimin Guo, Xiaoling He, Roger Barker, Erik Sundstr… Cramer, Igor Adameyko, Sten Linnarsson & Peter V. Kharchenko

Nature **560**, 494–498 (2018) | Cite this article

137k Accesses | **625** Citations | **664** Altmetric | Metrics

BMC Genomics
Methodology article | Open Access | Published: 19 June 2018

Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics

Kelly Street, Davide Risso, Russell B. Fletcher, Diya Das, John Ngai, Nir Yosef, Elizabeth Purdom & Sandrine Dudoit

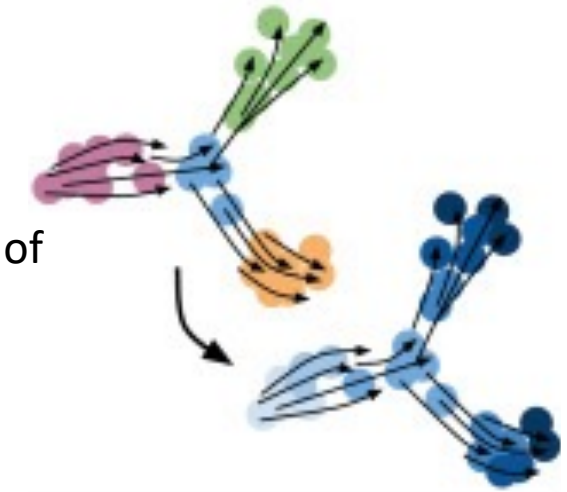
BMC Genomics **19**, Article number: 477 (2018) | Cite this article

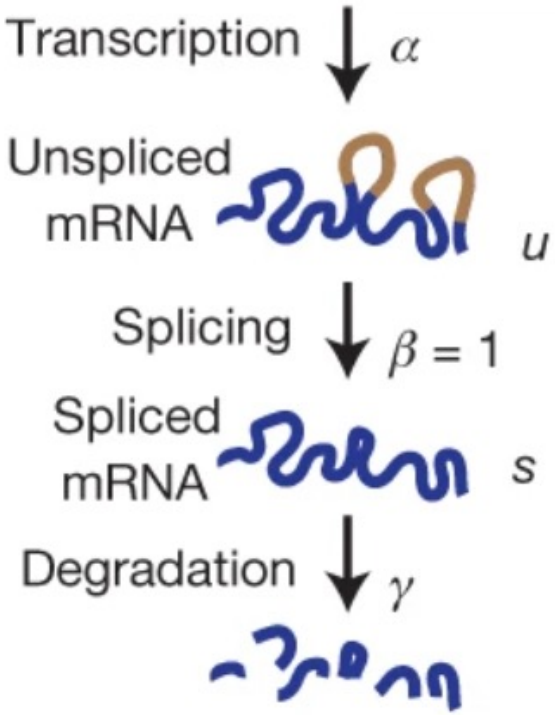
27k Accesses | **296** Citations | **20** Altmetric | Metrics

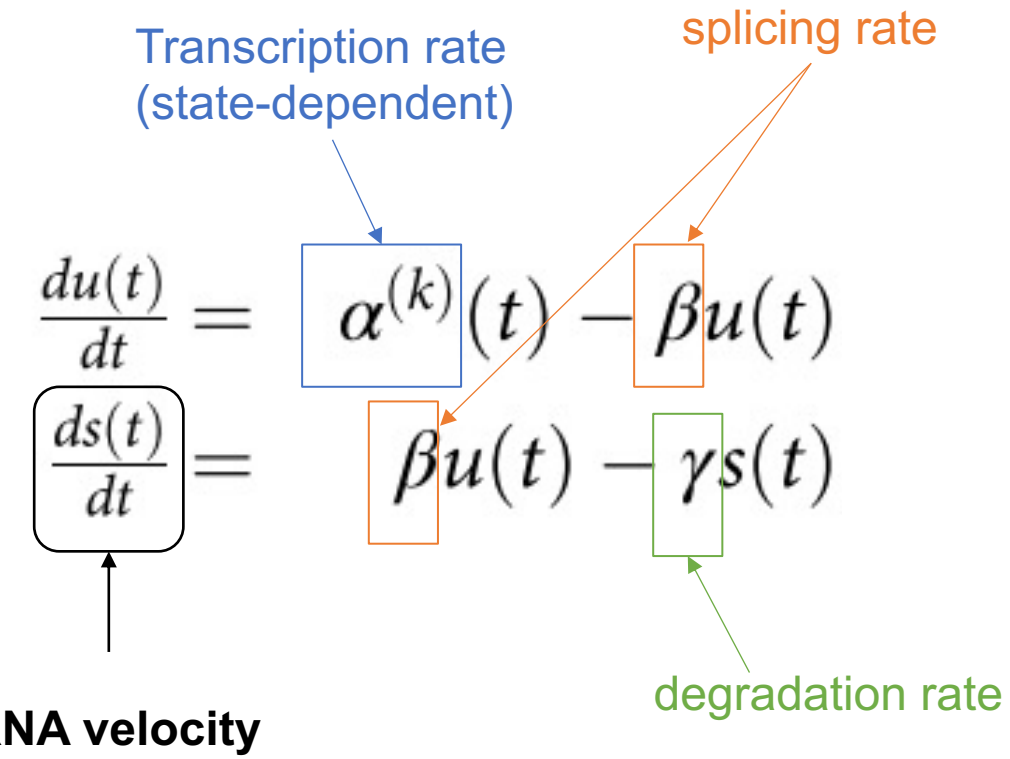
RNA velocity

An indicator for

- Future state of the abundance of mature mRNA
- Future state of the cell







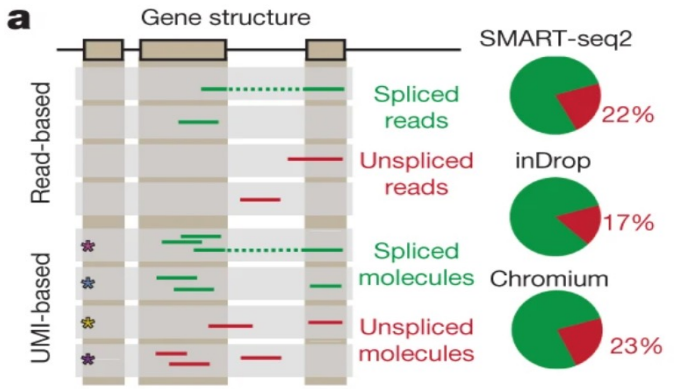
- u – unspliced /intronic/pre-mRNA abundance
- s – spliced/exonic/mature mRNA abundance

La Manno, G., Soldatov, R., Zeisel, A. *et al.* RNA velocity of single cells. *Nature* **560,** 494–498 (2018). https://doi.org/10.1038/s41586-018-0414-6

How to estimate RNA velocity?

- Estimate RNA reads
- Examine all data
  - Induction and repression
- Find steady state
- Estimate RNA velocity

Estimate unspliced and spliced reads



UMI : unique molecular identifier

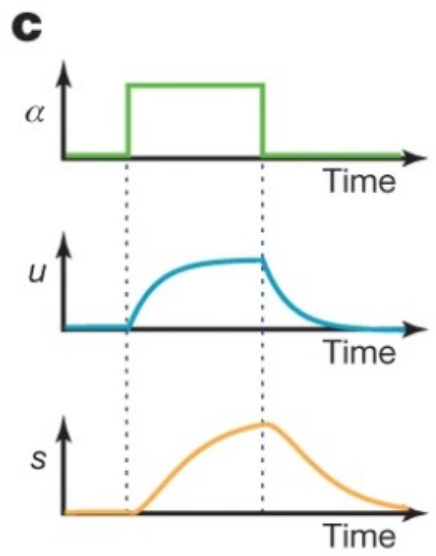
- single-cell sequencing protocol:
  - SMART-seq2
  - inDrop
  - Chromium

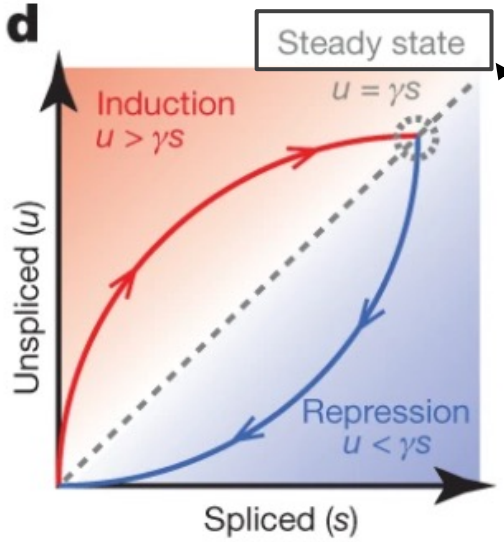
Transcription rate (state-dependent)

β assumed to be 1

$$\frac{du(t)}{dt} = \alpha^{(k)}(t) - \beta u(t)$$

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

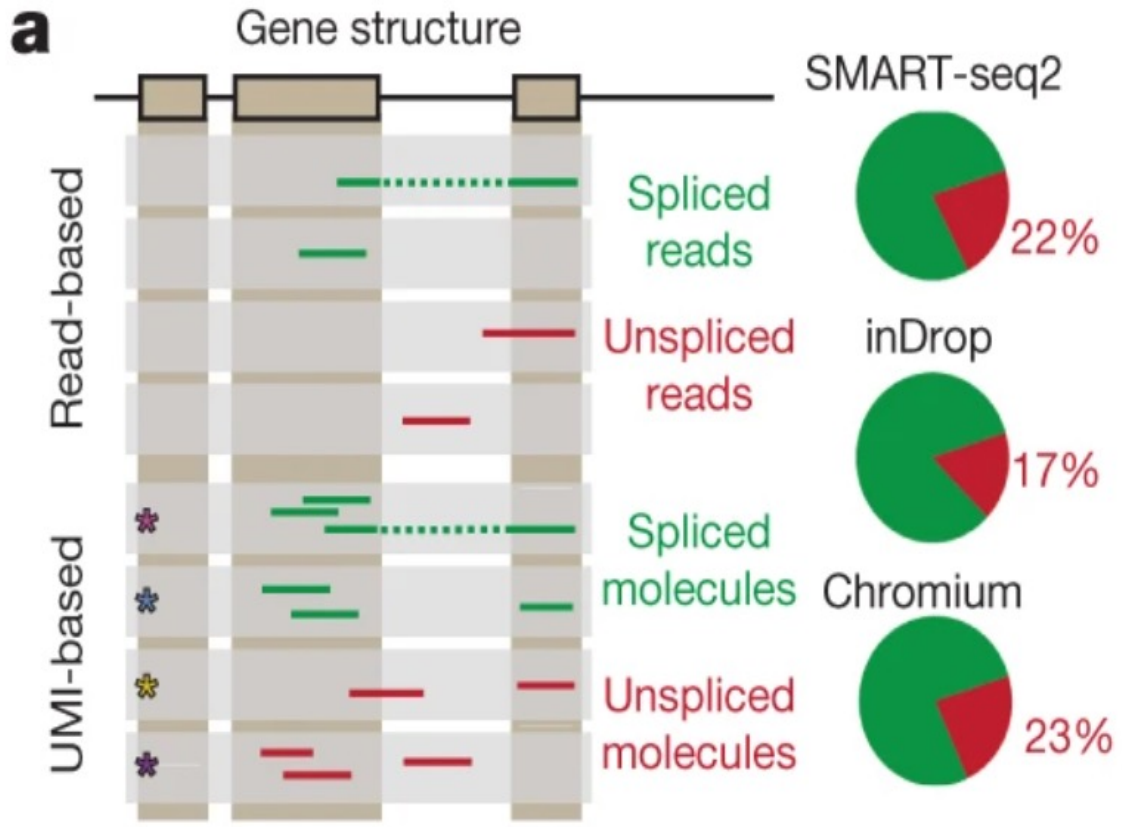




$$\frac{du(t)}{dt} = 0, \quad \frac{ds(t)}{dt} = 0$$

La Manno, G., Soldatov, R., Zeisel, A. *et al.* RNA velocity of single cells. *Nature* **560,** 494–498 (2018). https://doi.org/10.1038/s41586-018-0414-6

Estimate unspliced and spliced reads



UMI : unique molecular identifier

- single-cell sequencing protocol:
  - SMART-seq2
  - inDrop
  - Chromium

Estimate RNA velocity

- Governed by two equations

Transcription rate (state-dependent)

β assumed to be 1

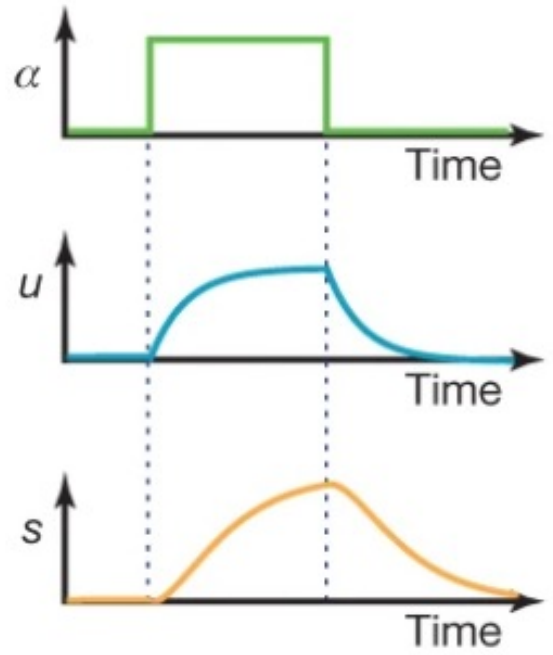
$$\frac{du(t)}{dt} = \alpha^{(k)}(t) - \beta u(t)$$

RNA velocity

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

degradation rate
Needs to be estimated

- Solution to equations
- Time course variation of u and s

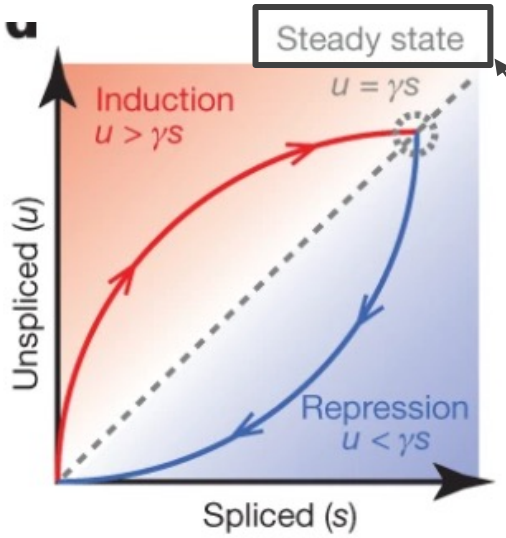


- Experimental data:
  - (u, s) from many measurements shown in one plot
- Data fall into two regimes:

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

  - Up-regulation/induction: ds/dt > 0
  - Down-regulation/repression: <0

- **Estimate degradation rate**
  - **Using data near steady state**



At steady state

$$\frac{du(t)}{dt} \quad \frac{ds(t)}{dt} \quad \} = 0$$

Dashed Line:

When no net change in **s**

Data near dashed line:

At steady state

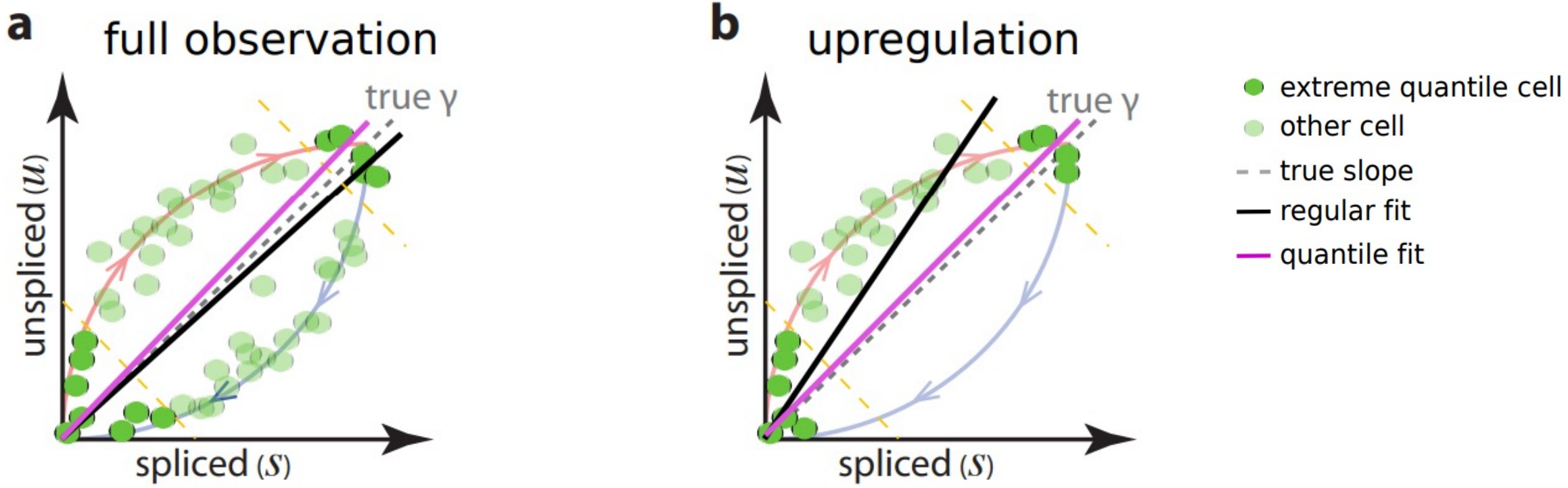
Estimate γ using extreme quantile fit

$$\boldsymbol{v} = \frac{ds}{dt} = u(t) - \gamma s(t)$$

Why extreme quantile fit will work?

Assume only partial gene regulation is observed

- Only has upregulation observation



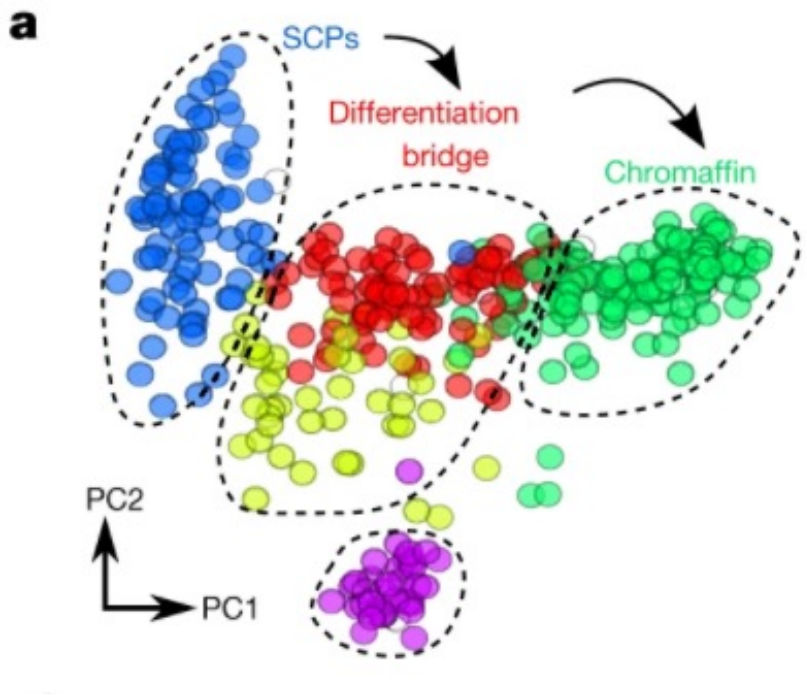
Extreme quantile fit:

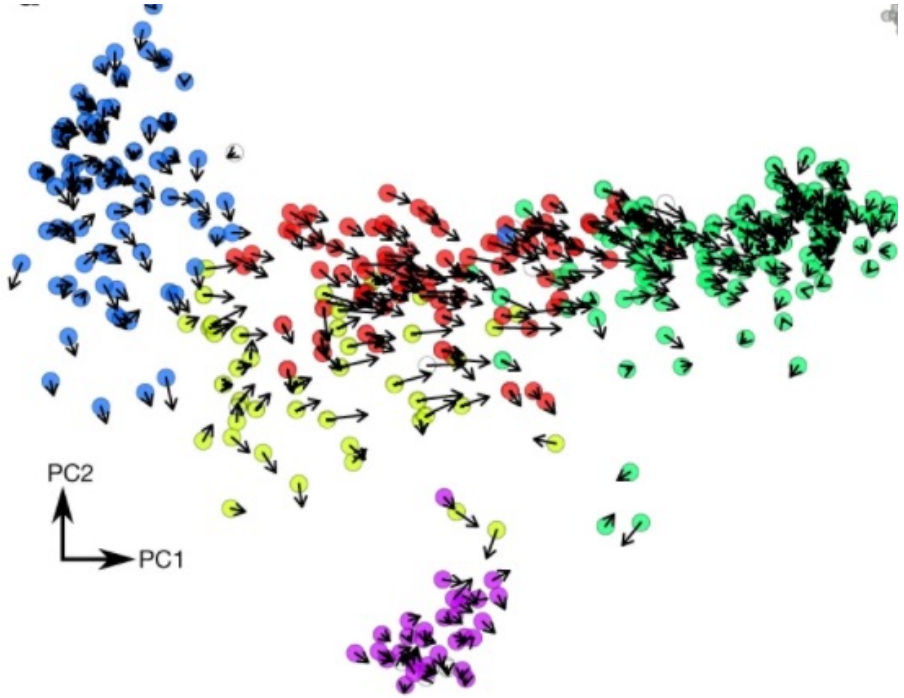
- including only cells near the origin and cells near the upper-right corner of the phase portrait
- Least square fit

La Manno, G., Soldatov, R., Zeisel, A. *et al.* RNA velocity of single cells. *Nature* **560,** 494–498 (2018). https://doi.org/10.1038/s41586-018-0414-6

Results: Mouse chromaffin cell differentiation

- SCPs: Schwann cell precursors
  - peripheral glial stem cells
- Chromaffin cells
  - neuroendocrine cells found in adrenal medulla
  - constitute the main hormonal component of the autonomic nervous system
  - the principal source for release of catecholamines, including adrenaline, in the systemic circulation
- sympathoblasts





Velocity arrow

- direction

  expression difference best correlated with the estimated velocity vector
- Length

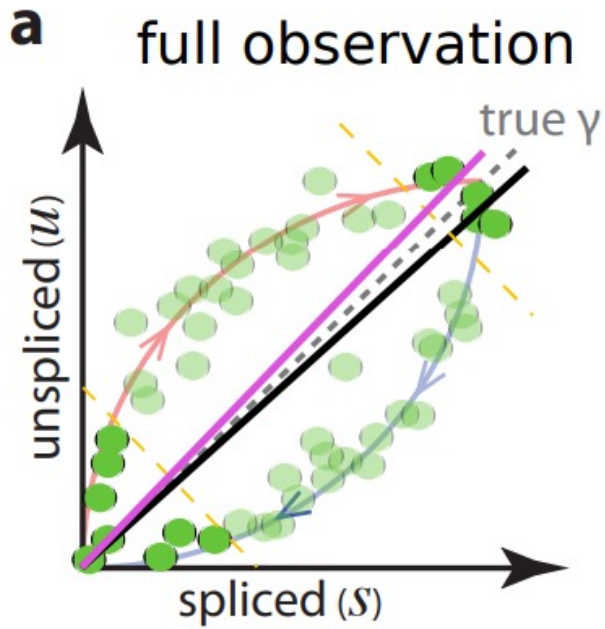
  Longer for higher transition probability to other cells

La Manno, G., Soldatov, R., Zeisel, A. *et al.* RNA velocity of single cells. *Nature* **560,** 494–498 (2018). https://doi.org/10.1038/s41586-018-0414-6

Limitation of the method

- A common splicing rate is assumed across different genes
- Assume the full splicing dynamics are measured
  - Transcriptional induction, repression and steady-state mRNA levels
  - Not realistic

$$\frac{du(t)}{dt} = \alpha^{(k)}(t) - \beta u(t)$$
$$\frac{ds(t)}{dt} = \beta u(t) - \gamma s(t)$$



NATURE BIOTECHNOLOGY

Article | Published: 03 August 2020

Generalizing RNA velocity to transient cell states through dynamical modeling

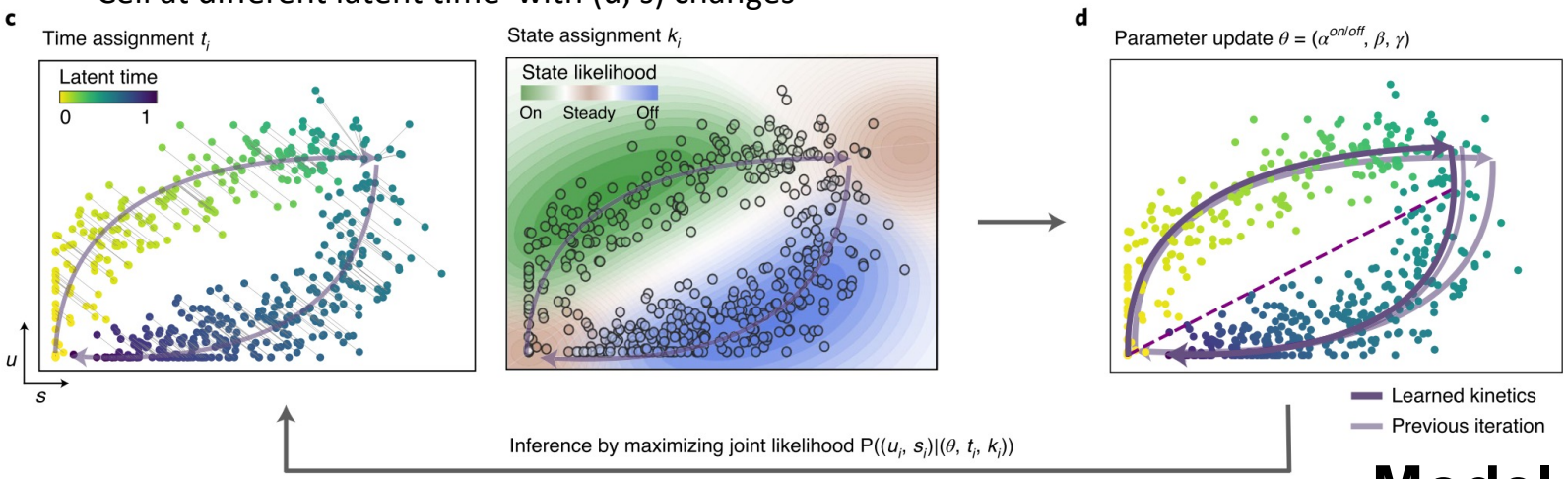
Volker Bergen, Marius Lange, Stefan Peidli, F. Alexander Wolf ✉ & Fabian J. Theis ✉

Nature Biotechnology **38**, 1408–1414 (2020) | Cite this article

Generalized RNA velocity

Latent time t_i

- Induction and repression process: a continuous time process
- the latent time:from the starting point right before the induction
- Cell at different latent time with (u, s) changes



β not 1 anymore

RNA velocity

$$\frac{du(t)}{dt} = \alpha^{(k)}(t) - \beta u(t)$$

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

Model parameters: $\theta = (\alpha^{(k)}, \beta, \gamma)$

Different regimes:

- 1: up-regulation
- 0: down-regulation
- SS_1 : active steady state
- SS_0 : inactive steady state

$$k_i \in \{1, 0, ss_1, ss_0\}$$

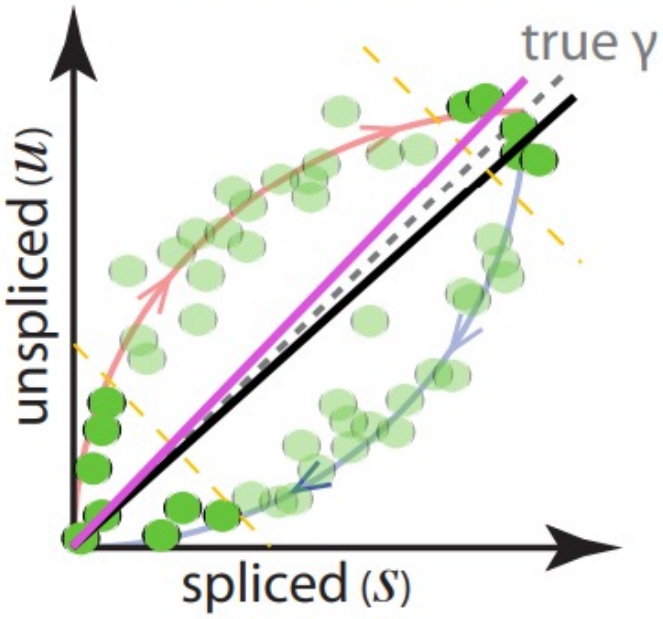
Model of continuous time process to estimate

- Model parameters: α, β, γ
- Latent time for each cell

Results: best fit summarized as gray curves.

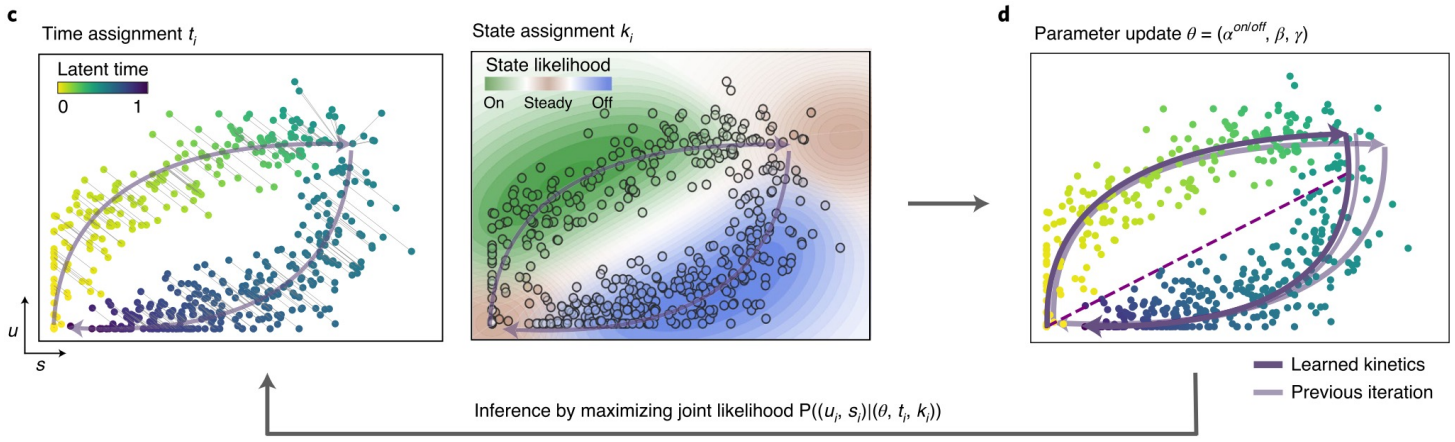
Bergen V, Lange M, Peidli S, Wolf FA, Theis FJ. Generalizing RNA velocity to transient cell states through dynamical modeling. Nat Biotechnol. 2020 Dec;38(12):1408-1414. doi: 10.1038/s41587-020-0591-3. Epub 2020 Aug 3. PMID: 32747759.

Steady-state model



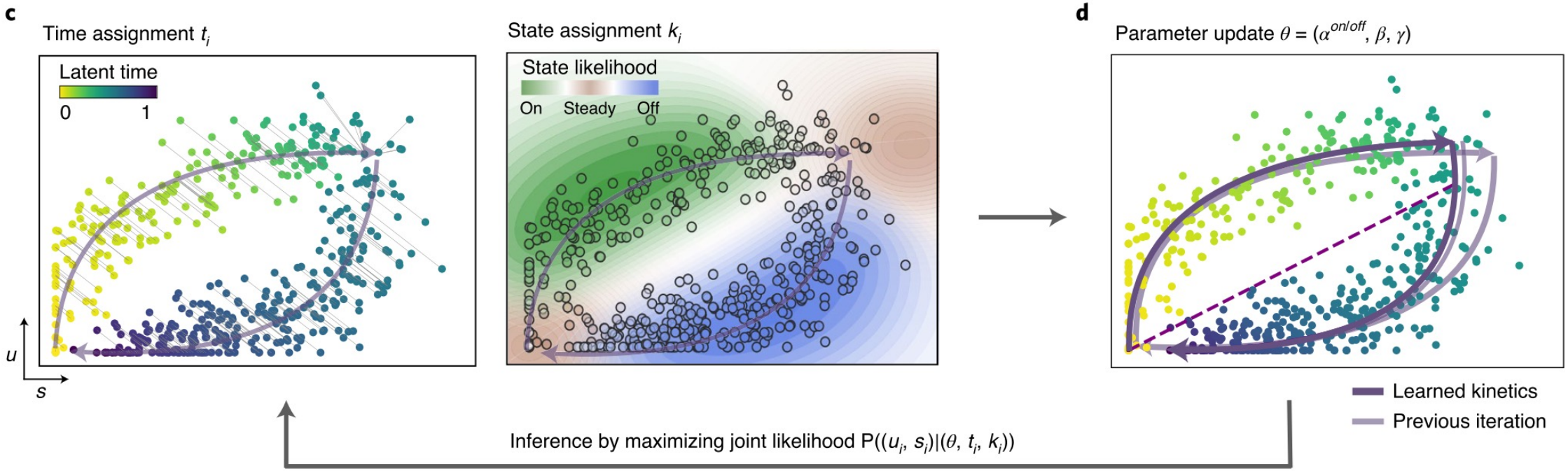
Paper 1

Dynamical model



Paper 2

Generalized RNA velocity



find a phase trajectory specified by $\hat{x}(t)$ that best describes the observations.

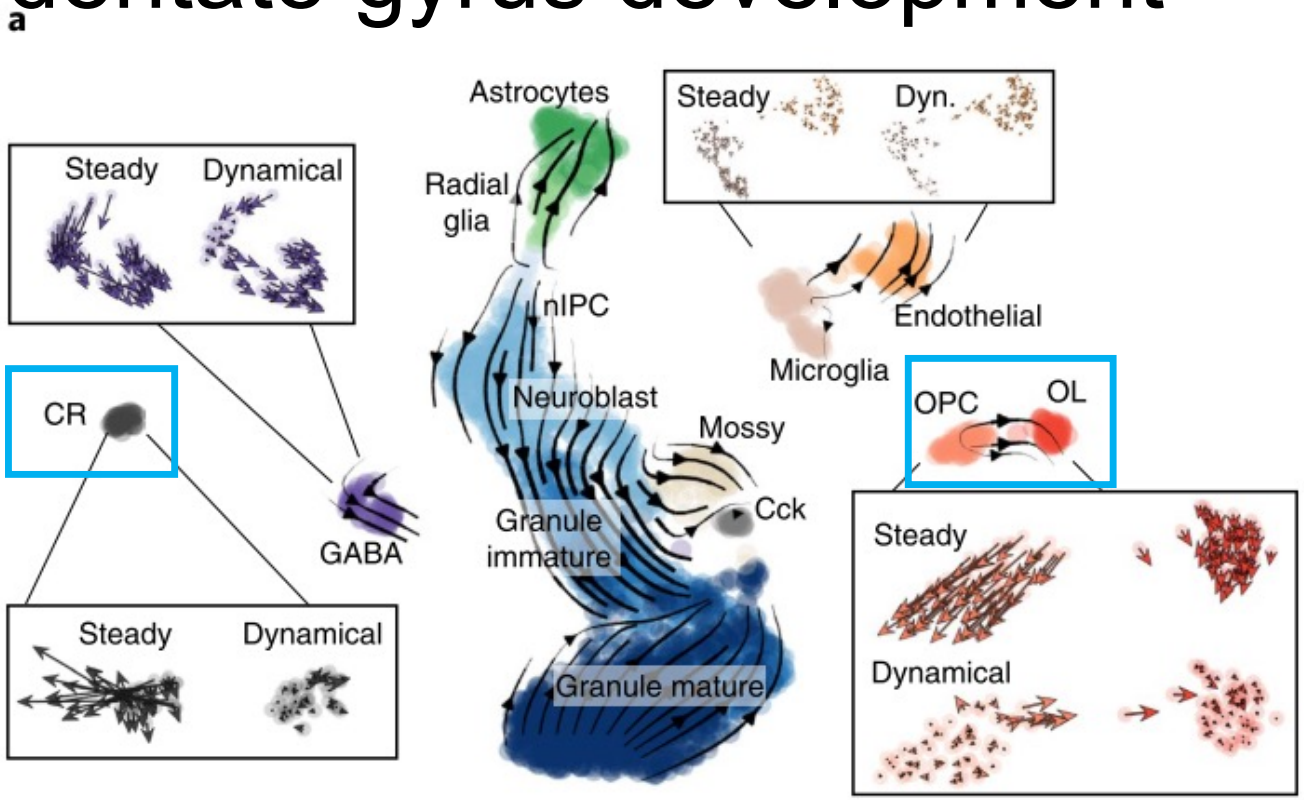
model estimate be $\hat{x}(t) = (\hat{u}(t), \hat{s}(t))$

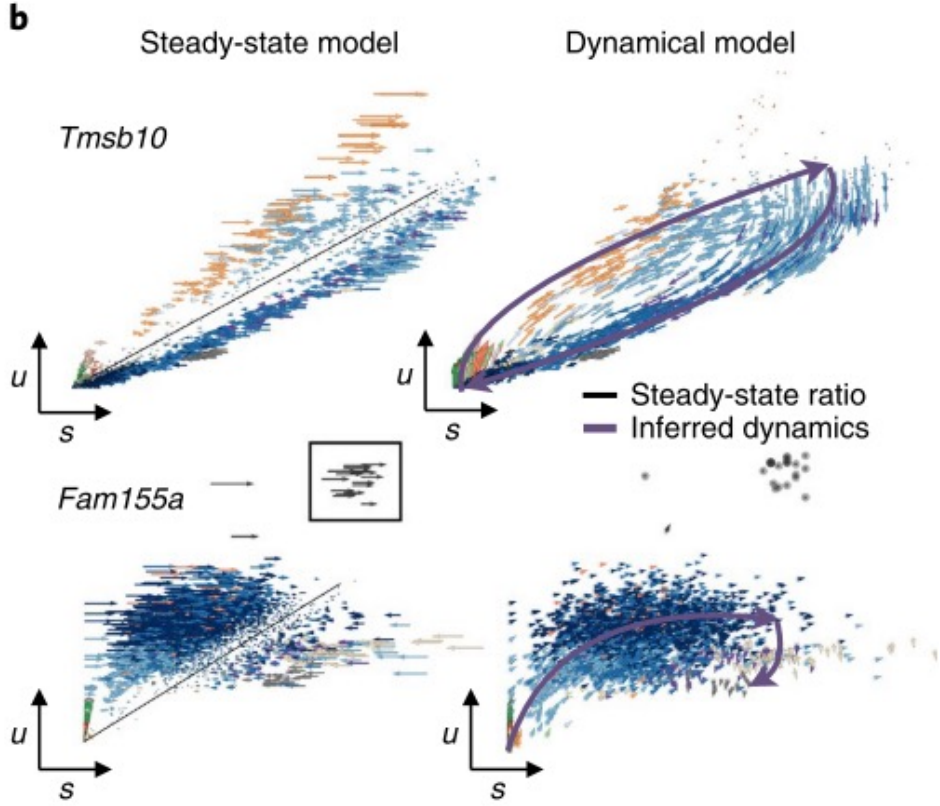
observations u_i^{obs} and s_i^{obs}

The likelihood for a particular gene $\mathcal{L}(\theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2n} \sum_i^n \frac{\left\| x_i^{obs} - x_{t_i}(\theta) \right\|^2}{\sigma^2} \right).$

Bergen V, Lange M, Peidli S, Wolf FA, Theis FJ. Generalizing RNA velocity to transient cell states through dynamical modeling. Nat Biotechnol. 2020 Dec;38(12):1408-1414. doi: 10.1038/s41587-020-0591-3. Epub 2020 Aug 3. PMID: 32747759.

Results: Resolving the heterogeneous population kinetics in dentate gyrus development





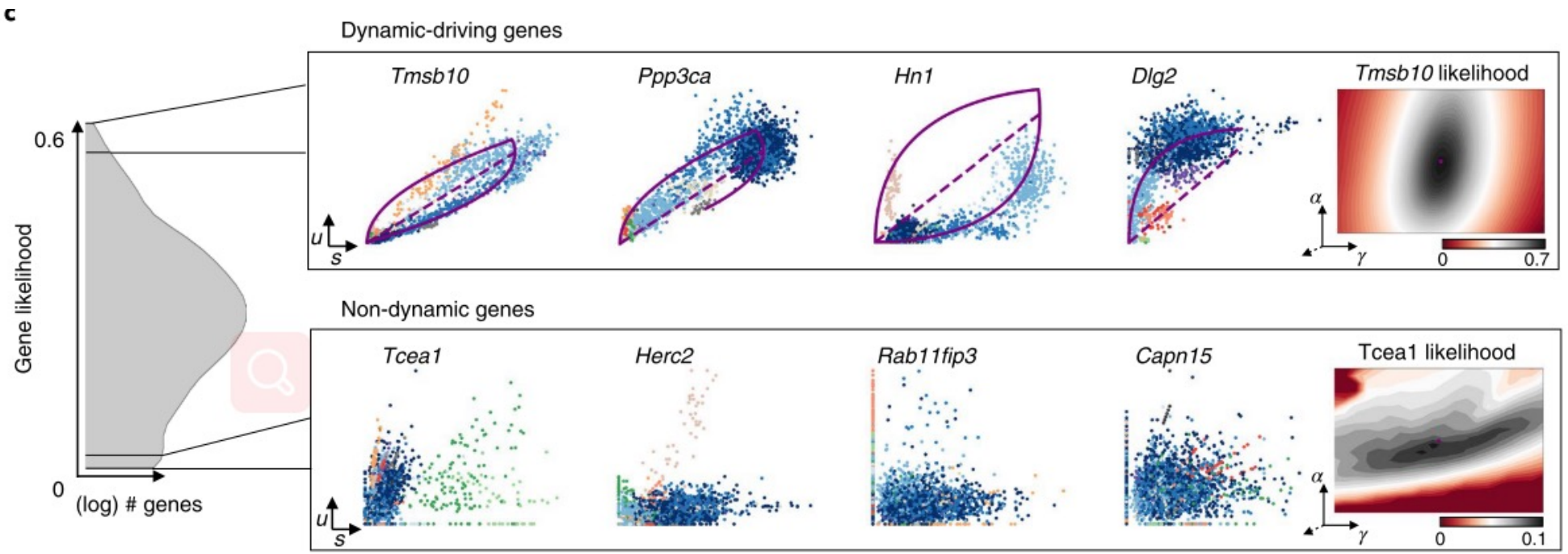
- Only **dynamical model** correctly identifies
  - the OPCs differentiating into OLs
  - CR cells as terminal.

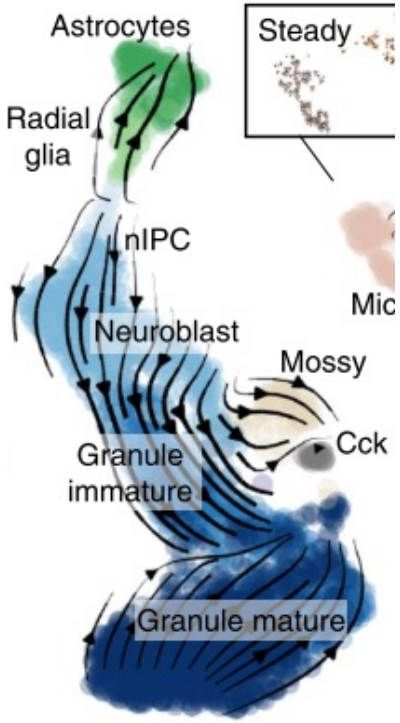
OPCs : oligodendrocyte precursor cells
OLs: myelinating oligodendrocytes
CRs: Cajal–Retzius cells

- Velocity of CR cells
  - Steady-state model
    - erroneously high velocities to CR cells
    - Fam155a phase portrait
  - while expression patterns show no evidence for any further maturation within the CR population.

Bergen V et al 2020 Nat Biotechnol.

Results: Identify putative driver genes

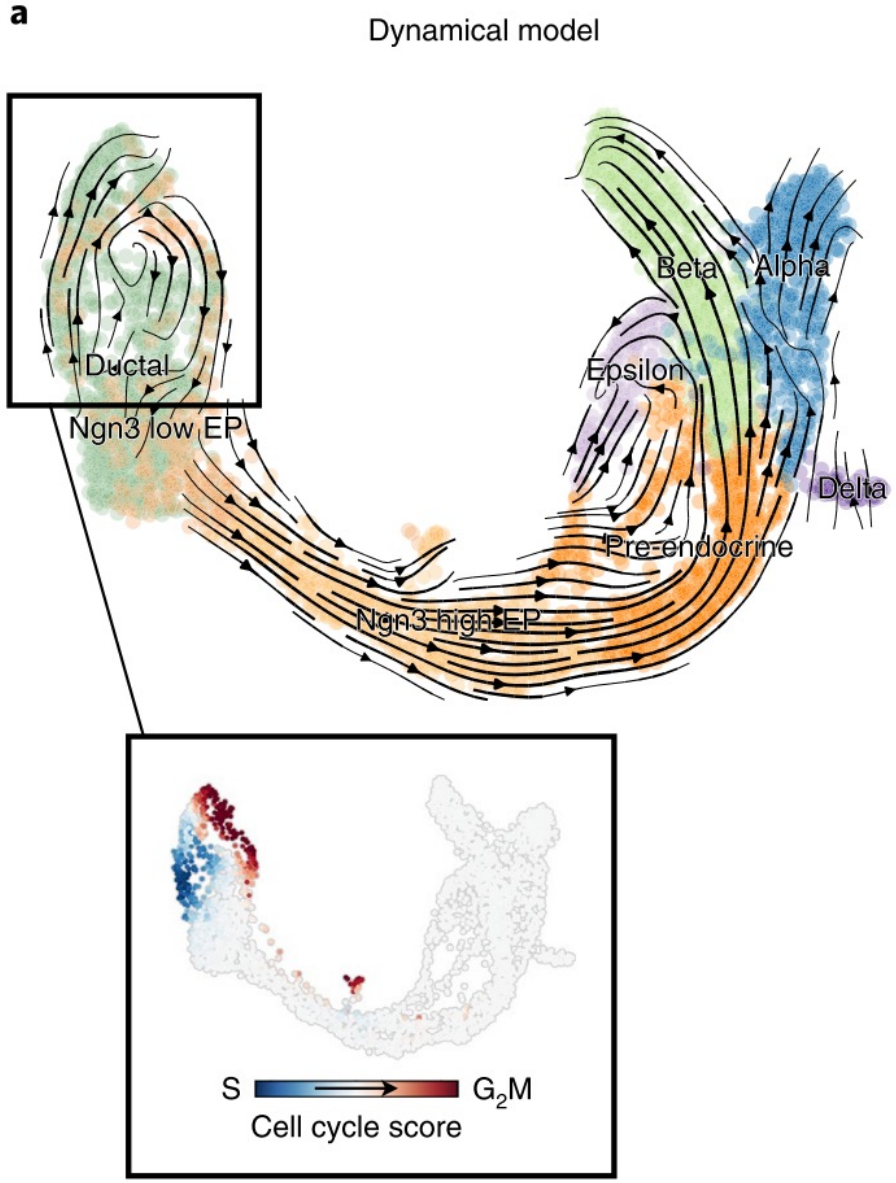


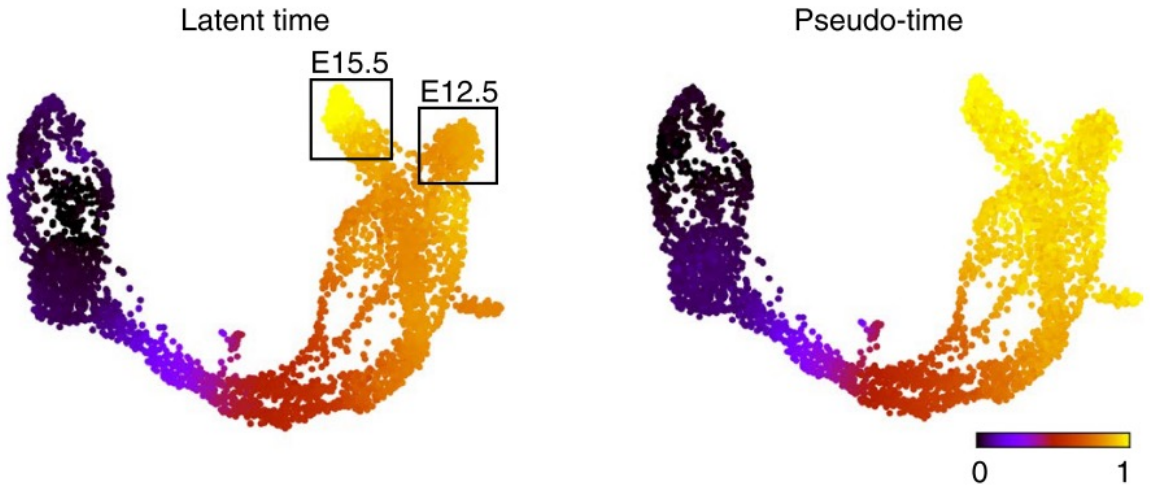


a likelihood for each gene

- explaining how well a cell is described by the learned spliced/unspliced phase trajectory.
- Aggregating over cells to obtain an overall gene likelihood

Results: Delineating cycling progenitors, commitment and fate transitions in endocrinogenesis

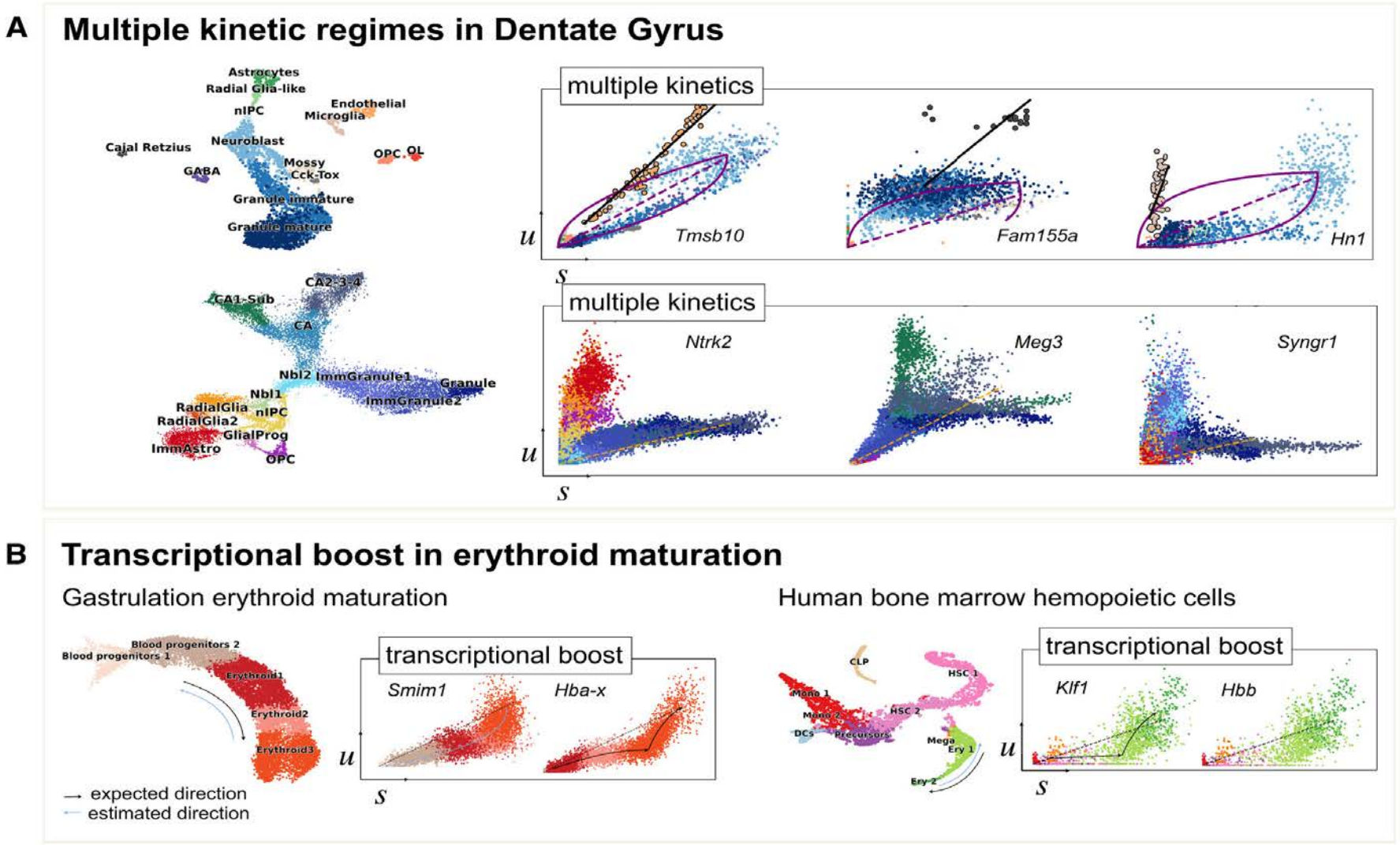




- Caputure cell cycle
- Predict real time better
  - α cells are produced earlier (before E12.5) then β cells(E12.5 – E15.5)

Bergen V et al 2020 Nat Biotechnol.

Potential pitfalls using current method



Multiple kinetics

- different genes have different kinetics

Transcription boost

- for one gene it can have severa different kinetics regime

Bergen, Volker, et al. "RNA velocity—current challenges and future perspectives." *Molecular Systems Biology* 17.8 (2021): e10282.

Conclusion

- Cell-specific RNA velocity estimates provide a natural basis for quantitative modelling of cell fates.
- Identification of dynamics driving genes
- Capturing cell cycles, complex branchings and heterogenous subpopulation kinetics
- Inference of shared latent time which enables relating lineages

Limitation and outlook

- Assumes constant reaction rates; full-length protocols allow accounting for:
  - Time- and state-dependent rates;
  - Alternative splicing
- Incorporate time and spatial information.

Q&A

Thank you!