



# Pseudo-temporal ordering and RNA velocity

Zhisong He, PhD

Senior researcher and Lecturer

Treutlein lab, D-BSSE, ETH Zurich

Basel, Switzerland

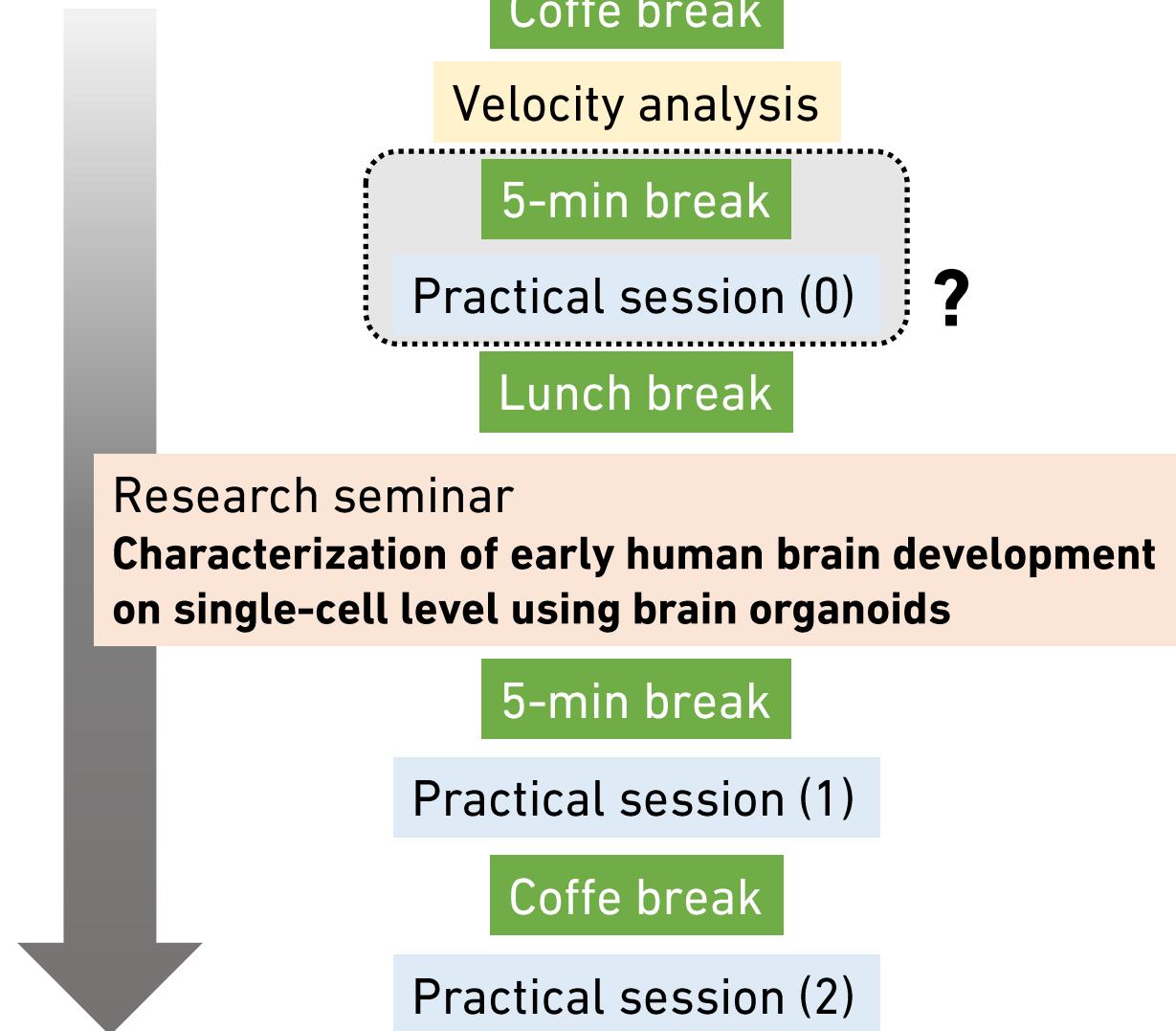
11.07.2024



Workshop  
**The  
Hitchhiker's  
Guide  
to scRNA-seq**

The Hitchhiker's Guide to scRNA-seq					
Workshop program					
	Monday	Tuesday	Wednesday	Thursday	
9h00	Welcome	Good morning	Good morning	Good morning	Good morning
9h30	Get to know each other	Integration António Sousa	Adaptive immune receptor Lisa Dratva	Velocity Zhisheng He	Spatial Transcriptomics Anna Maguza
10h40	Coffee break	Coffee break	Coffee break	Coffee break	Coffee break
11h00	Fundamentals of scRNA-seq Tomás Gomes	Integration António Sousa	Adaptive immune receptor Lisa Dratva	Velocity Zhisheng He	Spatial Transcriptomics Anna Maguza
12h30	Lunch break	Lunch break	Lunch break	Lunch break	Lunch break
13h30	Fundamentals of scRNA-seq Tomás Gomes	Seminar Yang Li HZI Braunschweig	Seminar Nuno Morais IMM	Seminar Zhisheng He ETH Zurich	Spatial Transcriptomics Anna Maguza
14h00	Integration António Sousa	Adaptive immune receptor Lisa Dratva	Velocity Zhisheng He		
15h30	Coffee break	Coffee break	Coffee break	Coffee break	Coffee break
16h30	Fundamentals of scRNA-seq Tomás Gomes	Integration António Sousa	Adaptive immune receptor Lisa Dratva	Velocity Zhisheng He	Closing session
17h00					
19h00	Art Gallery	Dinner with speakers	Dinner with speakers		
20h00	All group dinner				
21h00					
More info: <a href="mailto:scworkshop2024@medicina.ulisboa.pt">scworkshop2024@medicina.ulisboa.pt</a>					

## Pseudotime, trajectory analysis and fate probability



## Interior Design Style



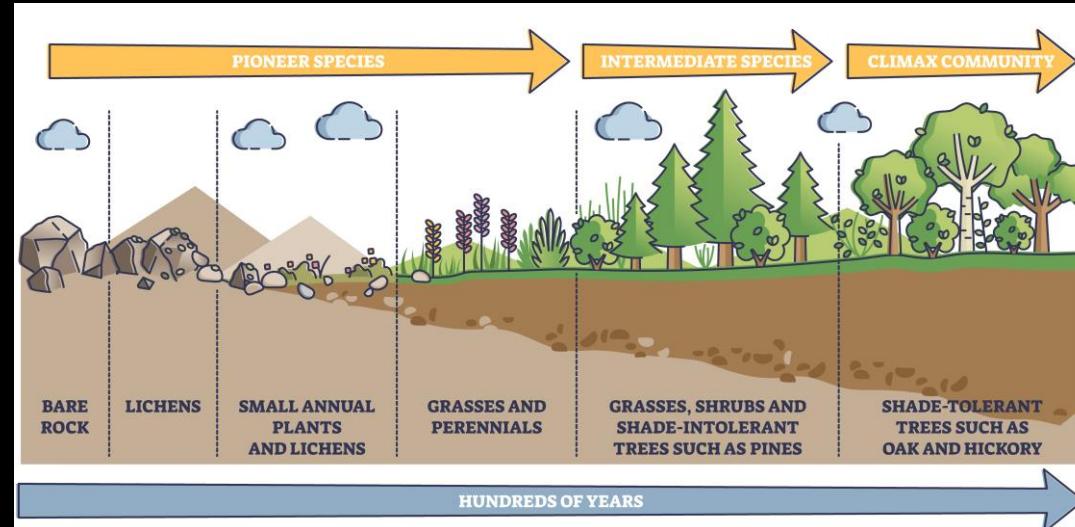
## Urban Expansion



**"The Only Constant in Life Is Change"**

- Heraclitus

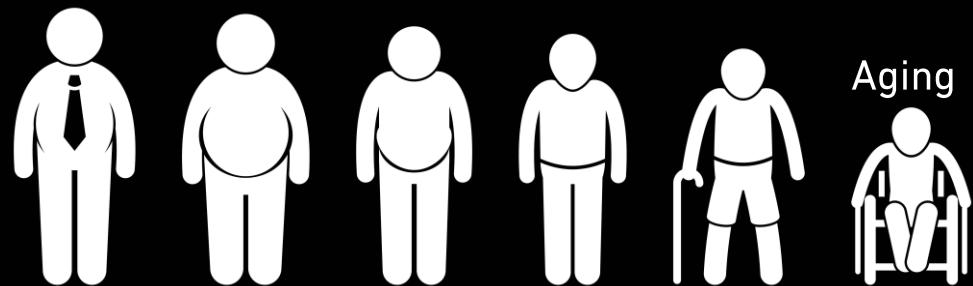
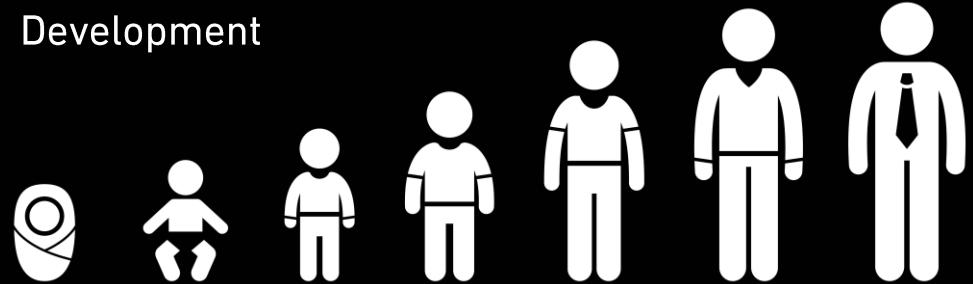
## Ecological Primary Succession



## Insect Metamorphosis



Development

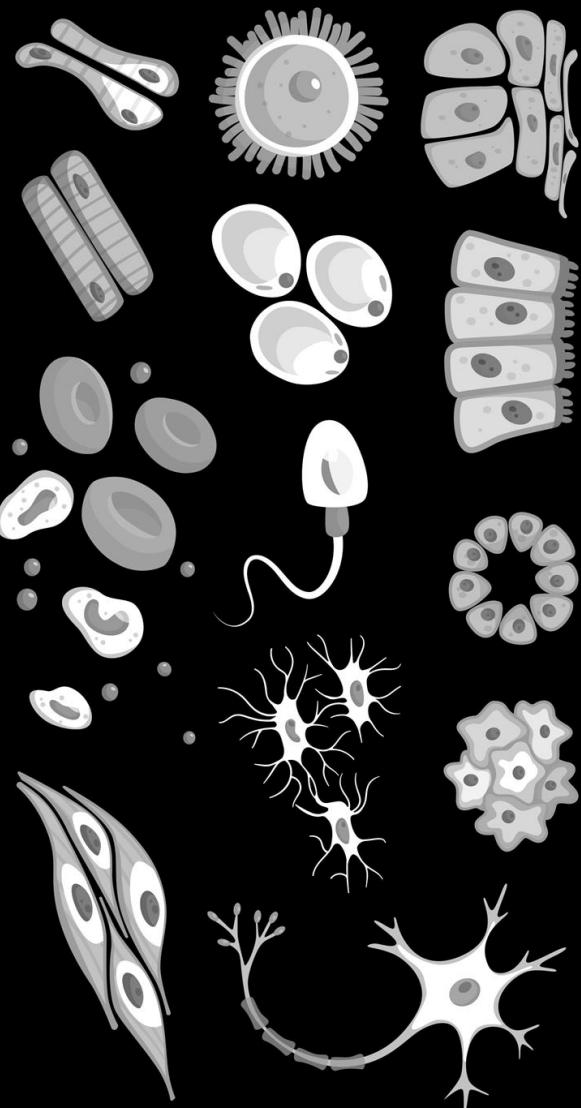


Aging

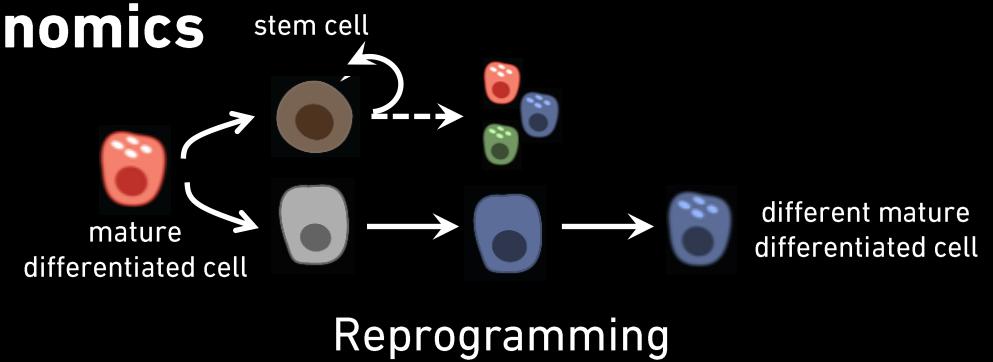
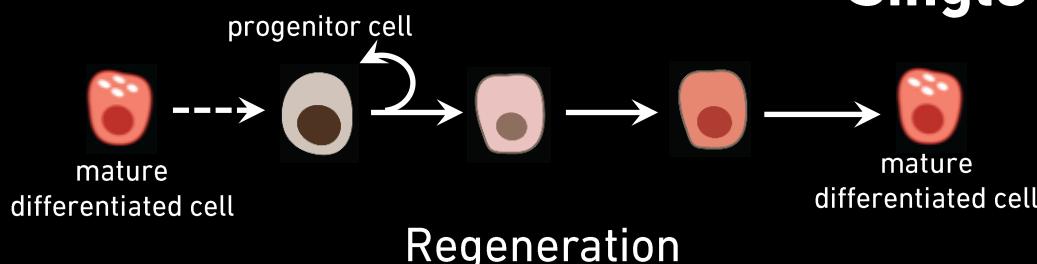
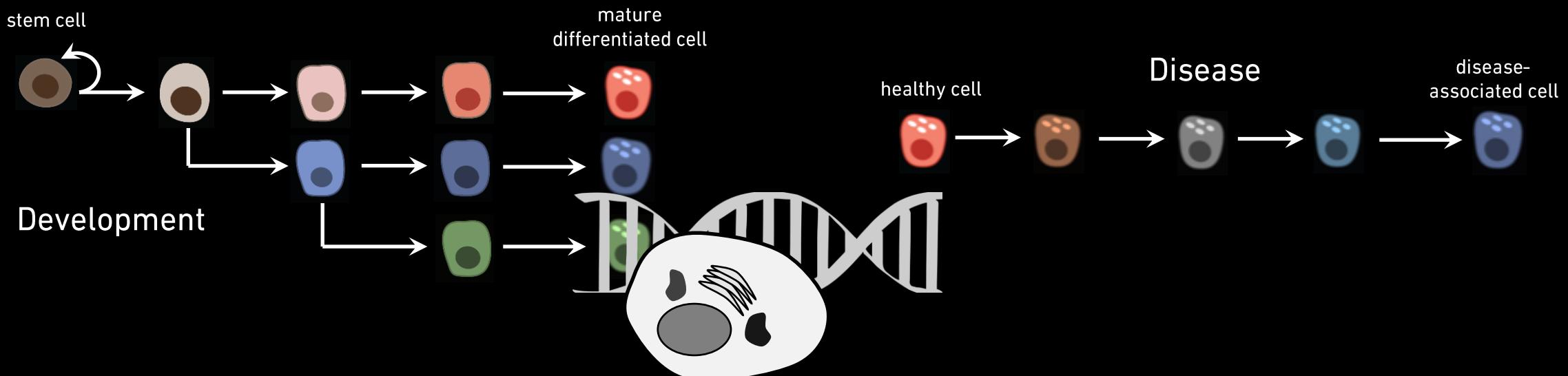
Diseases



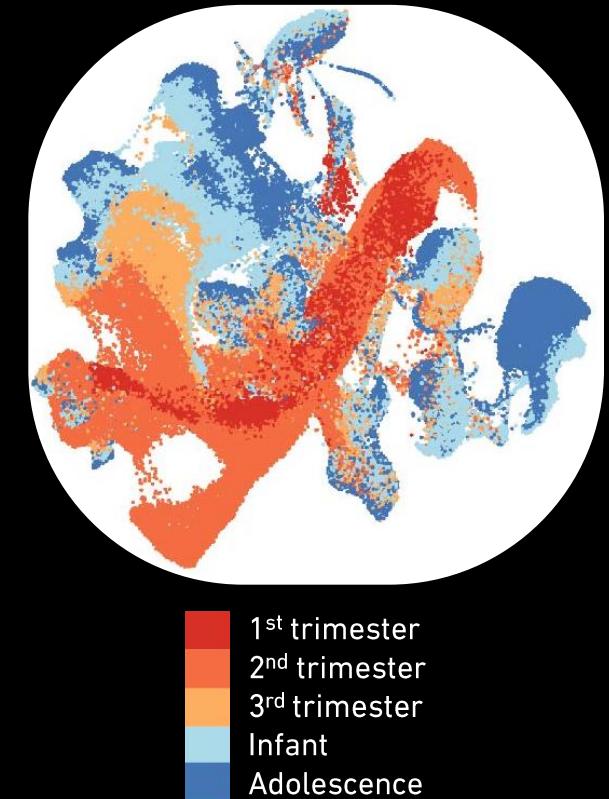
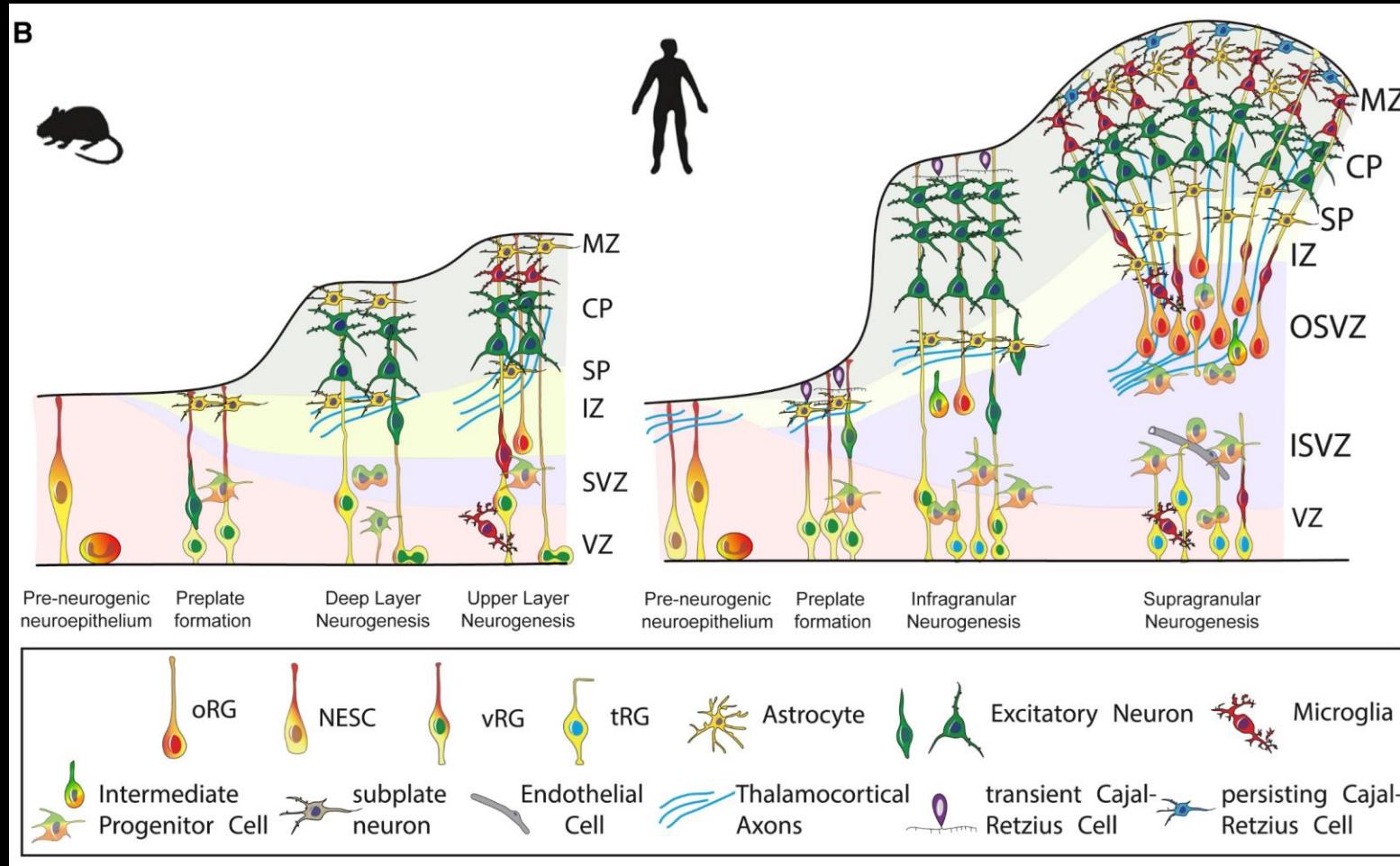
Changes in cells



# Cell state transitions



# Cell state transitions: Neurodevelopment



Re-processed data from Wang, et al. (2024)  
Molecular and cellular dynamics of the developing  
human neocortex at single-cell resolution. bioRxiv.  
<https://doi.org/10.1101/2024.01.16.575956>.

# scRNA-seq analysis to study cell state transitions



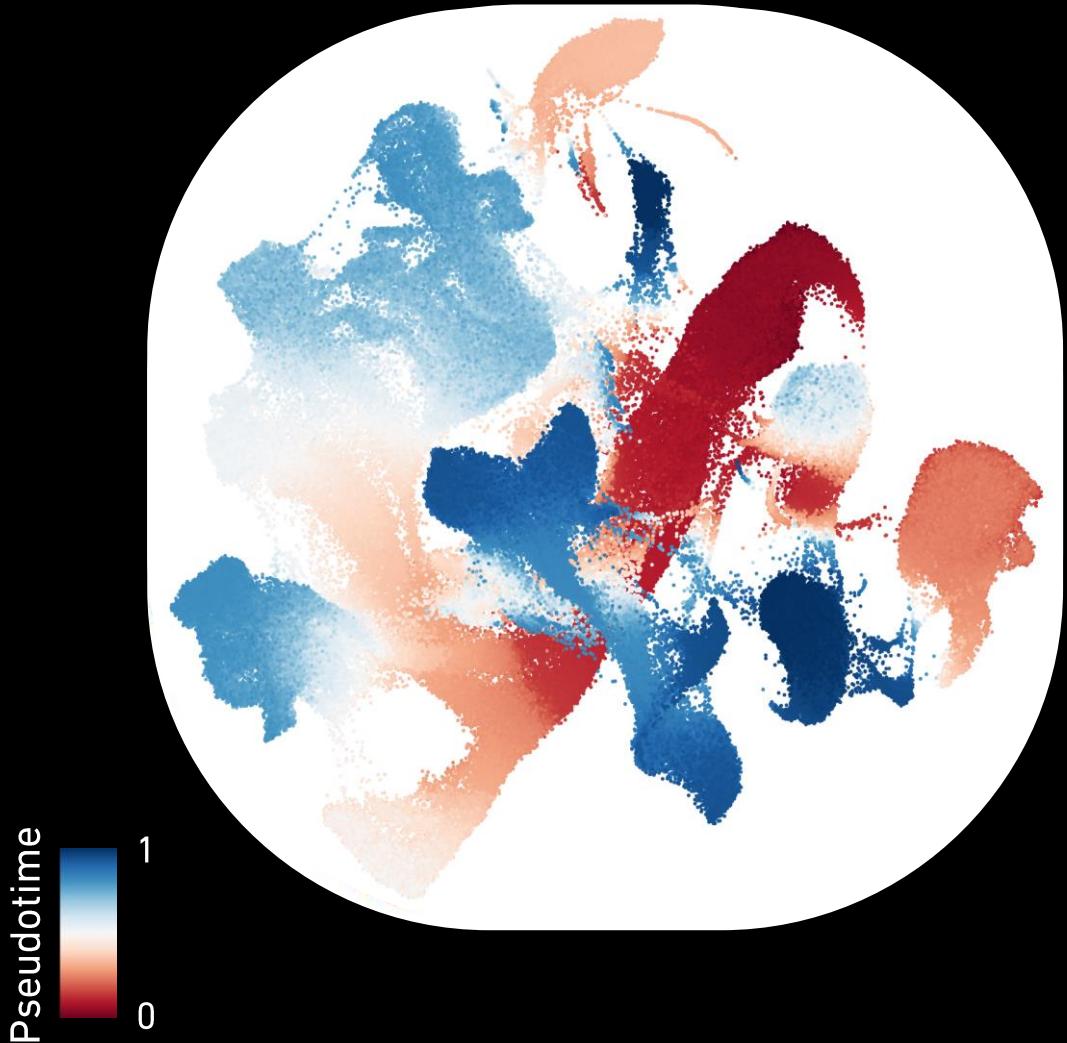
1<sup>st</sup> trimester  
2<sup>nd</sup> trimester  
3<sup>rd</sup> trimester  
Infant  
Adolescence

## Trajectory identification and characterization

Trajectories represent paths of cell state transitions

- Look for trajectory backbone
- Look for branching points
- Assign cells to the trajectory

# scRNA-seq analysis to study cell state transitions



## Trajectory identification and characterization

Trajectories represent paths of cell state transitions

- Look for trajectory backbone
- Look for branching points
- Assign cells to the trajectory

## Pseudotime / pseudo-temporal cell ordering

Pseudotime quantifies progression of cell state transitions

# scRNA-seq analysis to study cell state transitions



## Trajectory identification and characterization

Trajectories represent paths of cell state transitions

- Look for trajectory backbone
- Look for branching points
- Assign cells to the trajectory

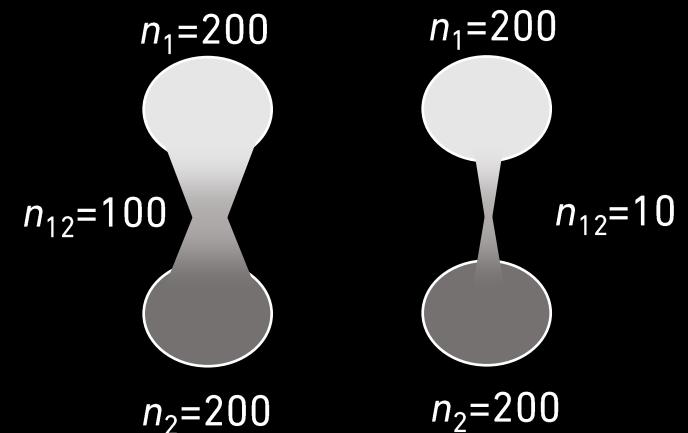
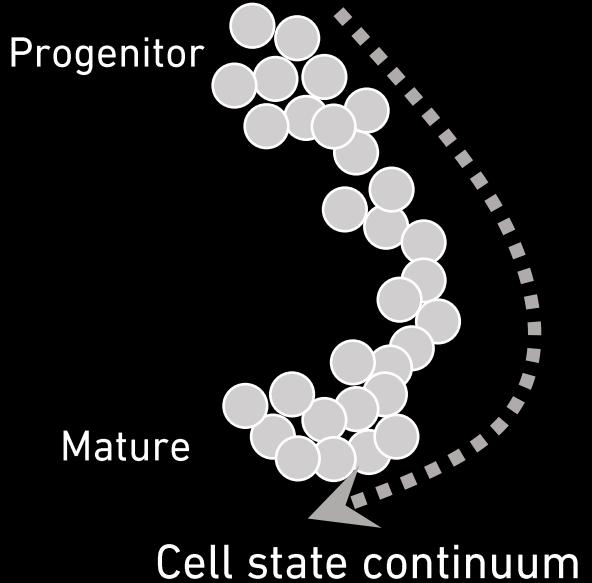
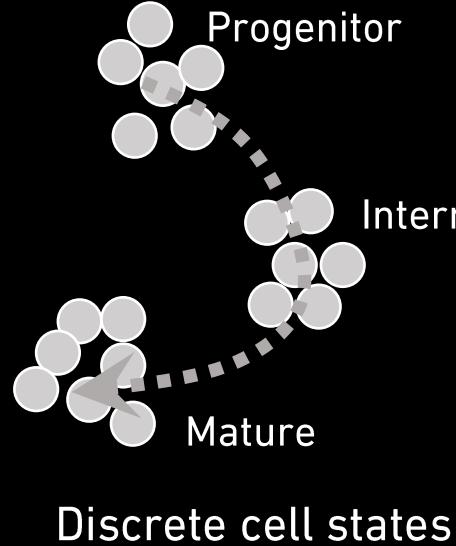
## Pseudotime / pseudo-temporal cell ordering

Pseudotime quantifies progression of cell state transitions

## Cell fate probability estimation

Estimate the fate commitment of cells

# Trajectory & Pseudotime: WHEN

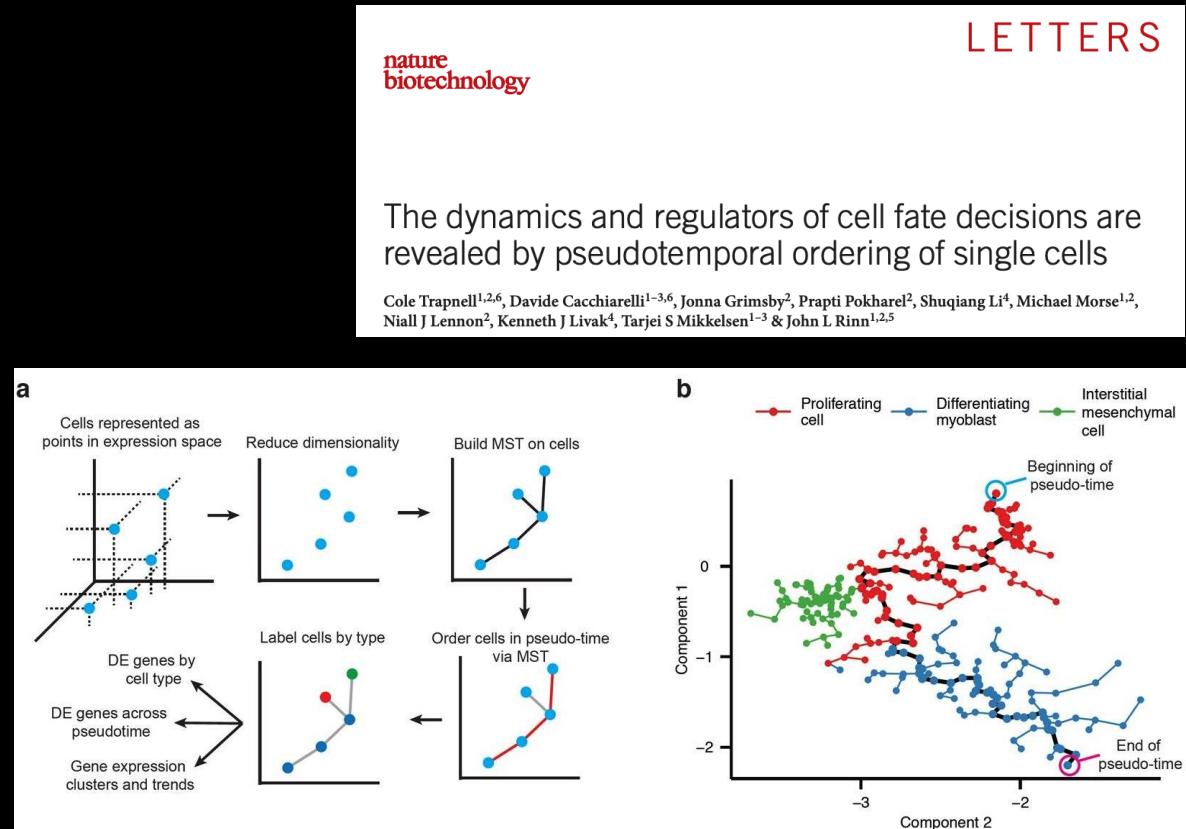
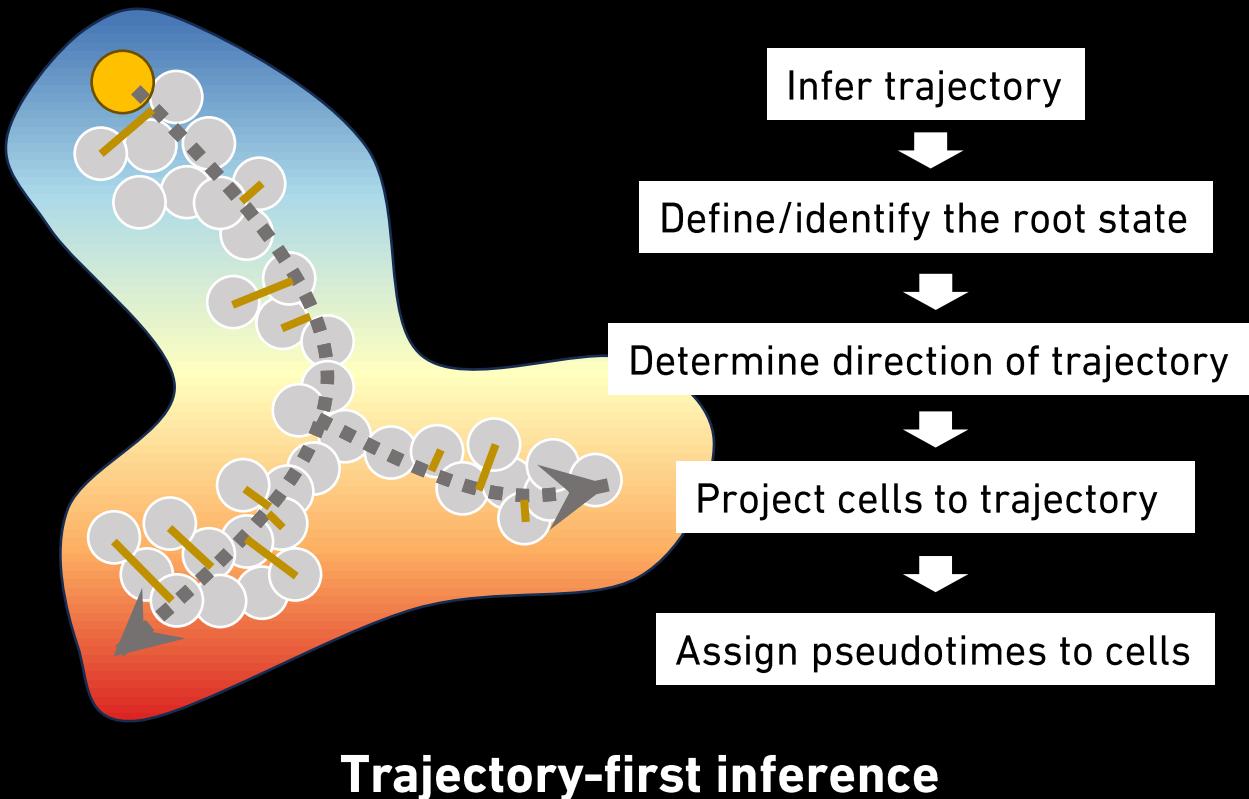


Artificial  
continuum by  
doublets?

- Doublet detection algorithms (e.g. *DoubletFinder*, *Scrublet*)
- TF markers for intermediates

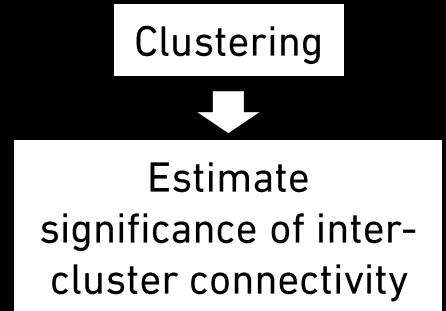
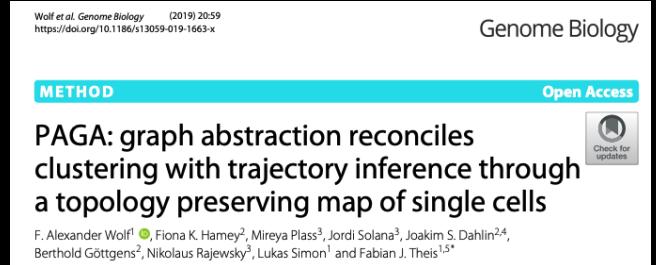
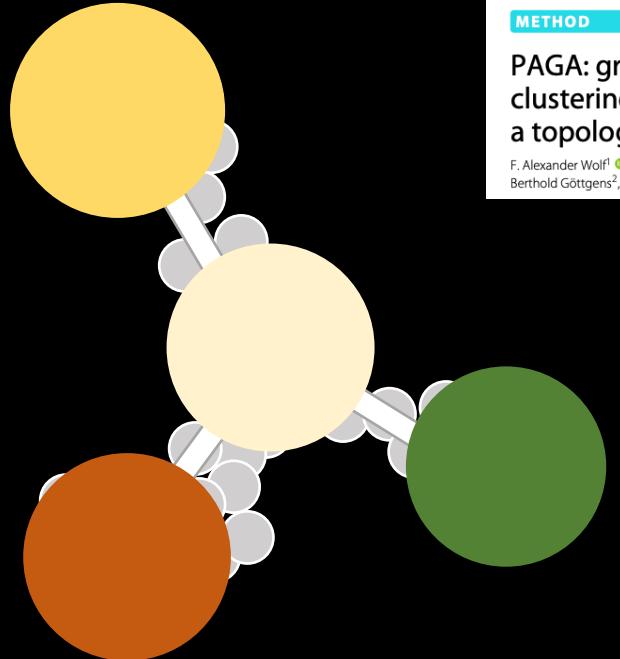
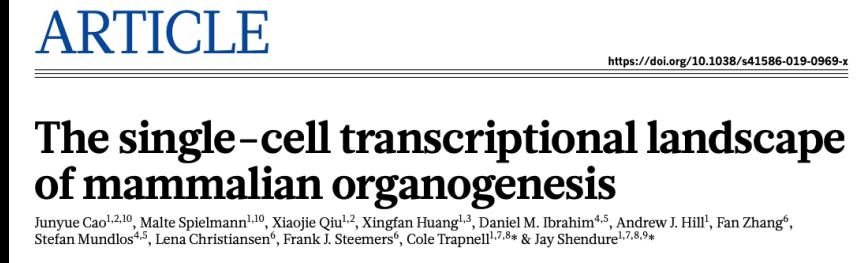
# Trajectory & Pseudotime: HOW

Monocle (2014)  
(*monocle* in R)



# Trajectory & Pseudotime: HOW (2)

Monocle 3 (2019)  
(*monocle3* in R, *py-monocle* in Python)



# Trajectory & Pseudotime: HOW (2)

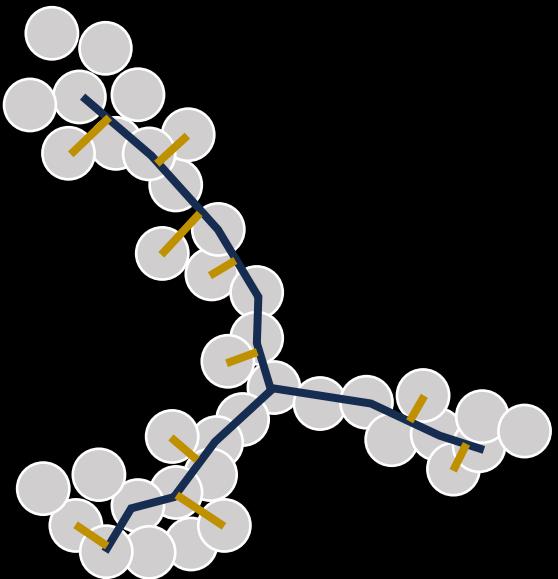
Monocle 3 (2019)  
(*monocle3* in R , *py-monocle* in Python)

ARTICLE

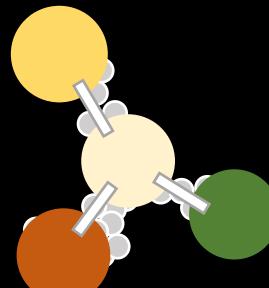
<https://doi.org/10.1038/s41586-019-0969-x>

## The single-cell transcriptional landscape of mammalian organogenesis

Junyue Cao<sup>1,2,10</sup>, Malte Spielmann<sup>1,10</sup>, Xiaojie Qiu<sup>1,2</sup>, Xingfan Huang<sup>1,3</sup>, Daniel M. Ibrahim<sup>3,5</sup>, Andrew J. Hill<sup>1</sup>, Fan Zhang<sup>6</sup>, Stefan Mundlos<sup>4,5</sup>, Lena Christiansen<sup>6</sup>, Frank J. Steemers<sup>6</sup>, Cole Trapnell<sup>1,7,8,\*</sup> & Jay Shendure<sup>1,7,8,9,\*</sup>



Connectivity constraint



## Principal graph

Maximize similarities between a cell and the average of its connected cells in the graph

## SimplePPT: A Simple Principal Tree Algorithm

Qi Mao\*

Le Yang†

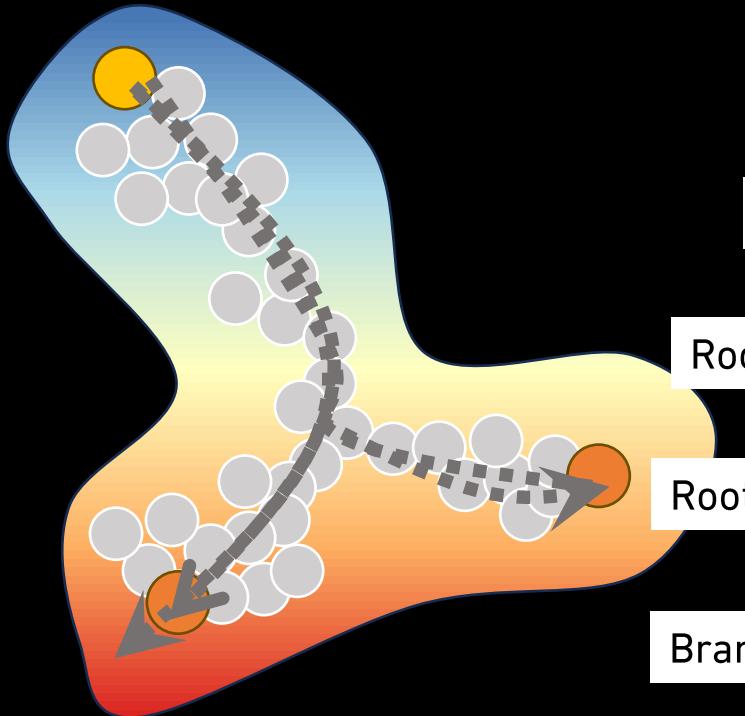
Li Wang‡

Steve Goodison§

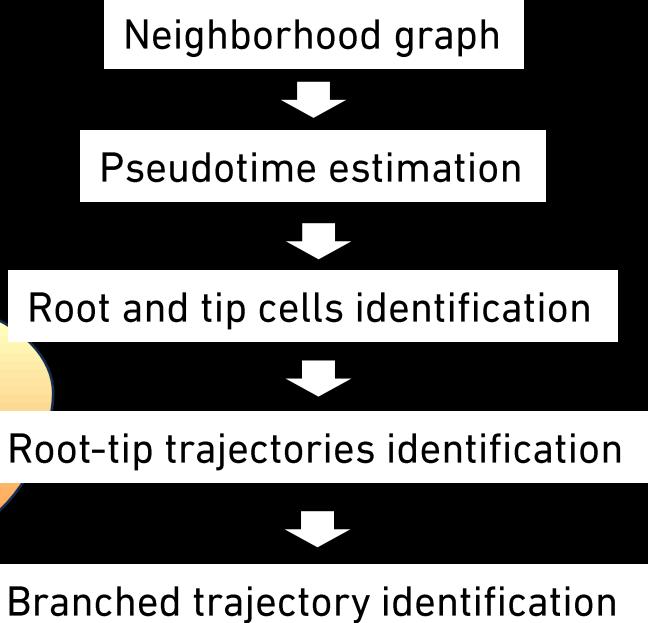
Yijun Sun¶

$$\min_{f_{\mathcal{G}} \in \mathcal{F}_{\mathcal{G}, \ell}} \frac{1}{N} \sum_{i=1}^N \min_{\mathbf{z} \in \mathcal{Z}} c(\mathbf{x}_i, f_{\mathcal{G}}(\mathbf{z})), \quad \mathcal{F}_{\mathcal{G}, \ell} = \{f_{\mathcal{G}} \in \mathcal{F}, \ell(\mathcal{G}) \leq \ell\},$$
$$\ell(\mathcal{G}) = \sum_{(V_i, V_j) \in \mathcal{E}} w_{i,j} \|f_{\mathcal{G}}(\mathbf{z}_i) - f_{\mathcal{G}}(\mathbf{z}_j)\|^2.$$

# Trajectory & Pseudotime: HOW (3)



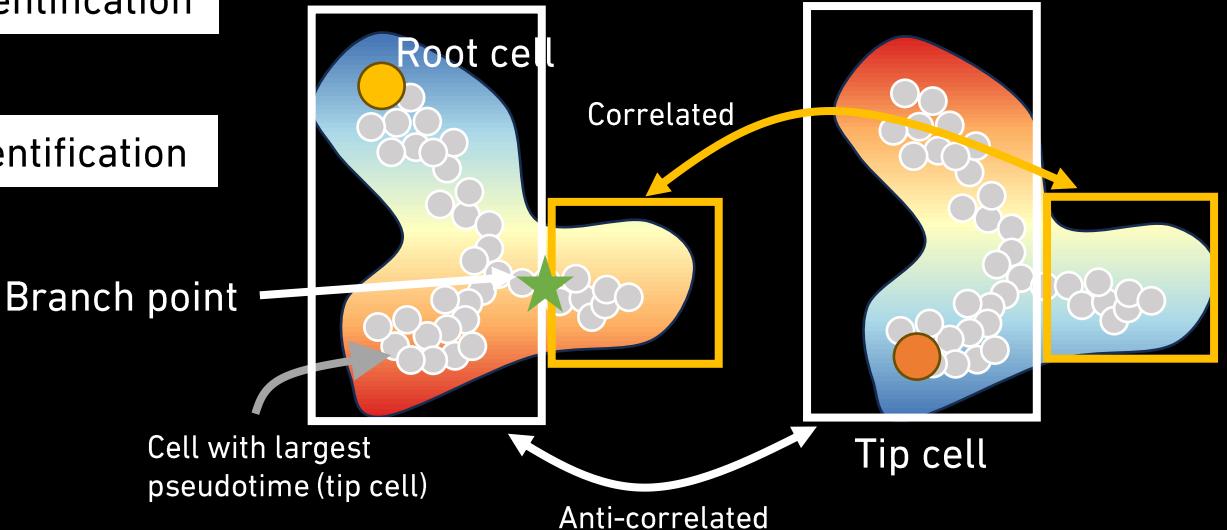
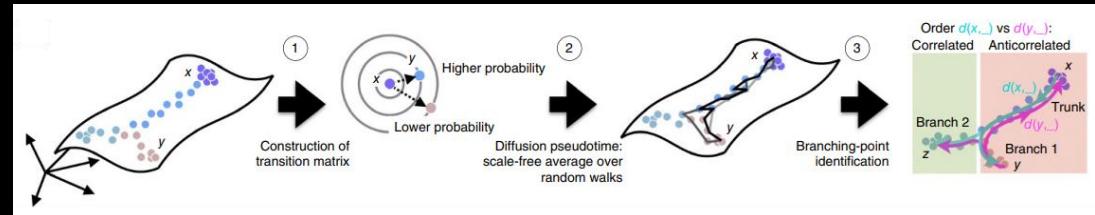
**Pseudotime-first inference**



Diffusion Pseudotime (2016)  
(*destiny* in R, *scanpy* in Python)

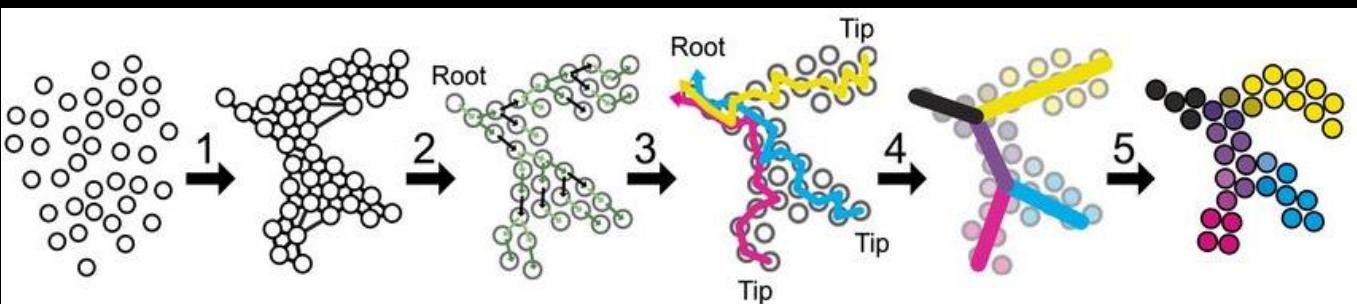
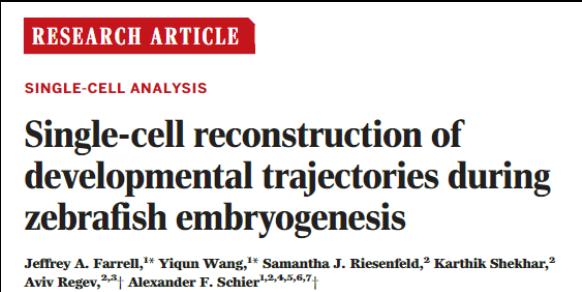
**Diffusion pseudotime robustly reconstructs lineage branching**

Laleh Haghverdi<sup>1,2</sup>, Maren Büttner<sup>1</sup>, F Alexander Wolf<sup>4</sup>,  
Florian Büttner<sup>1,3</sup> & Fabian J Theis<sup>1,2</sup>

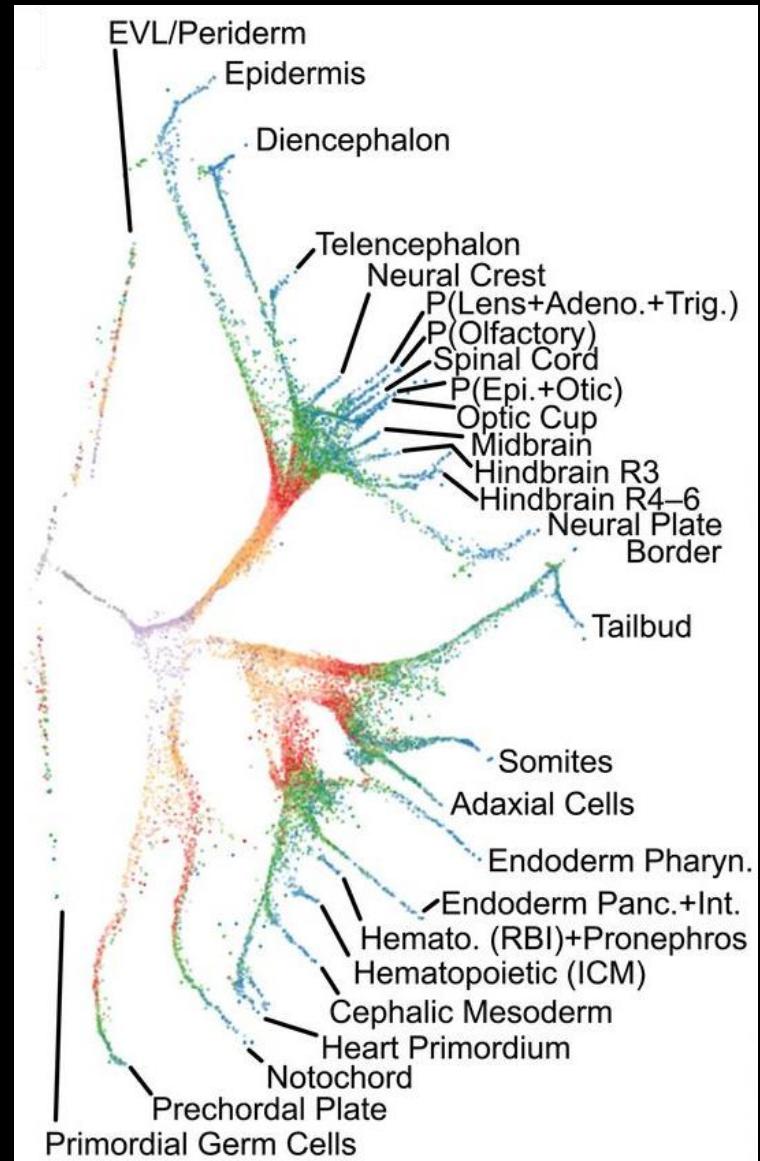


# Trajectory & Pseudotime: HOW (4)

URD (2018)  
(URD in R)



Cells	Transition matrix	Pseudotime calculation	Trajectory inference	Trajectory joining	Visualization
Diffusion map ( <i>destiny</i> )	Probabilistic breadth-first search (flood simulation)	Random walk simulation on biased transition matrix	Segment-and-join	Force-directed layout based on visitation frequency from tips	

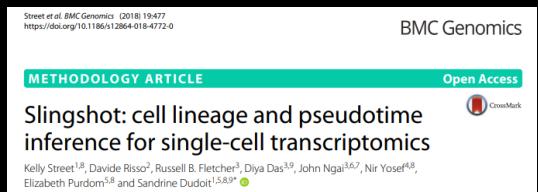


# Trajectory & Pseudotime: MORE

## Slingshot (2018)

(*slingshot* in R, *pyslingshot* in Python)

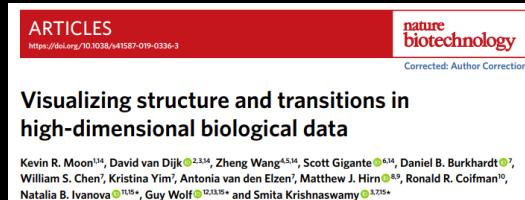
- Trajectory-first inference of trajectory and pseudotime



## PHATE (2019)

(*phate* in Python, *phateR* in R as a wrapper to the Python package)

- Diffusion-based pseudotime-first inference of pseudotime and trajectories
- Data visualization approach which can highlight cell state transitions



## Palantir (2019)

(*palantir* in Python)

- Pseudotime-first inference of pseudotime and fate probability



## moscot (2023)

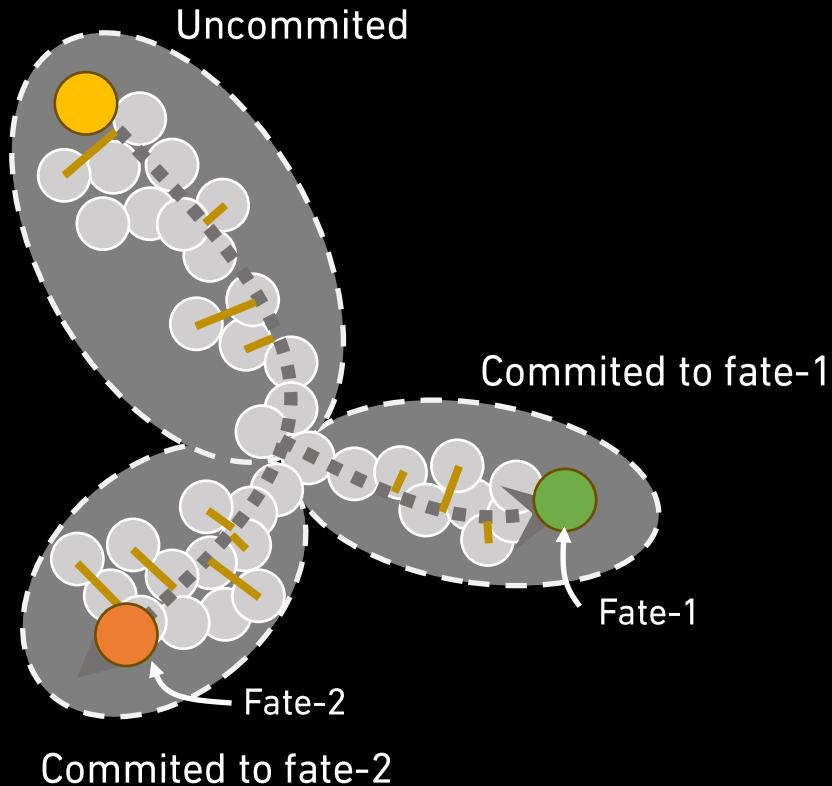
(*moscot* in Python)

- A general framework based on optimal transport to map cells across time and space
- Incorporate real time information for pseudotime estimation

## Mapping cells through time and space with moscot

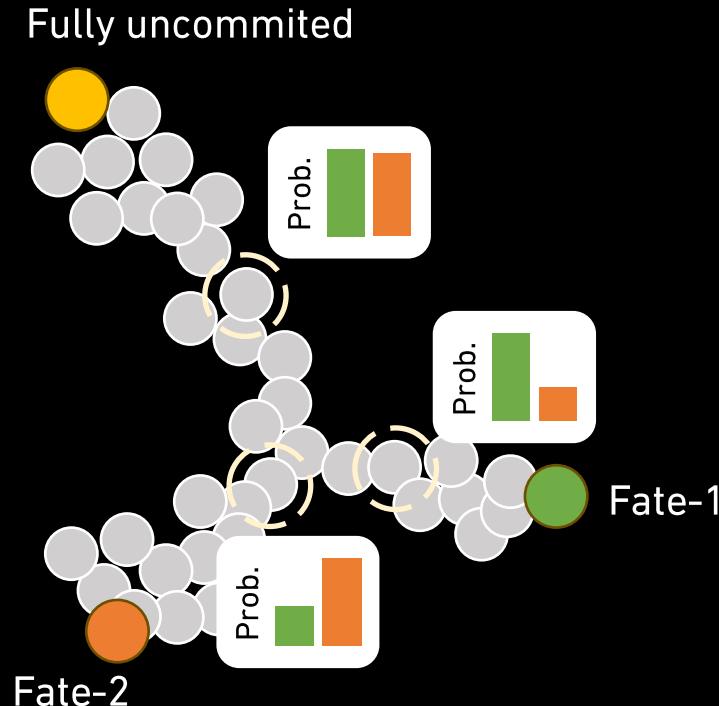
Dominik Klein\*, Giovanni Palla\*, Marius Lange\*, Michal Klein\*, Zoe Piran\*, Manuel Gander\*, Laetitia Meng-Papaxanthos\*, Michael Stern\*, Aimée Bastidas-Ponce\*, Marta Tarquis-Medina\*, Heiko Lickert\*, Mostafa Bakhti\*, Mor Nitzan\*, Marco Cuturi\*, Fabian J. Theis\*

# Trajectory vs. fate probability



## Trajectory analysis

Assume one decision per branch point  
Intuitive and simple



## Fate probability estimation

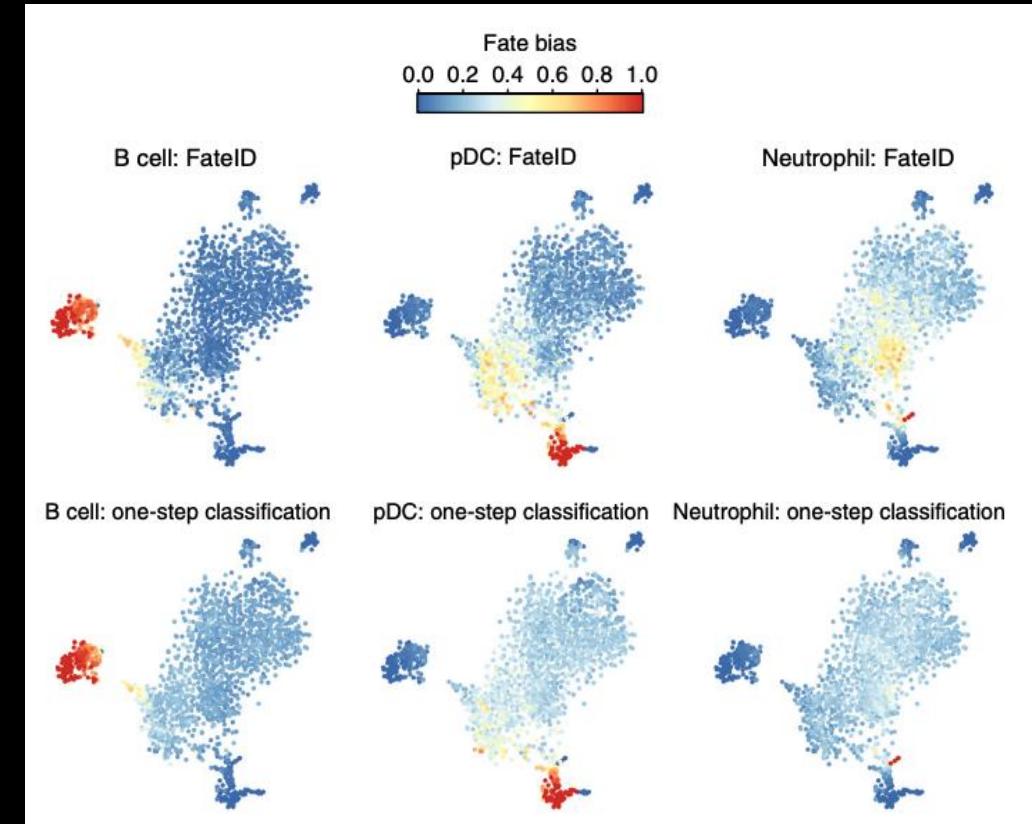
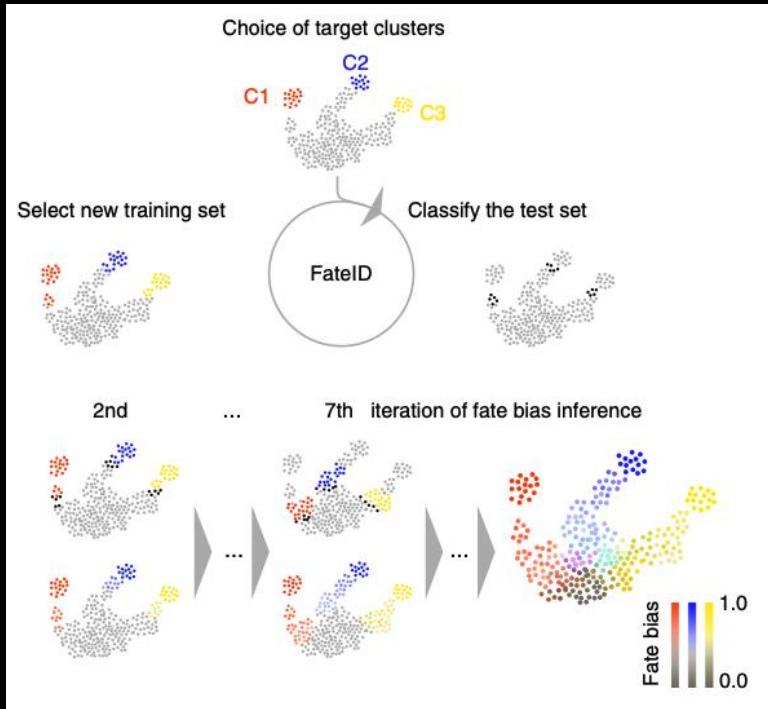
Assume continuous decision making  
Flexible and likely better reflect the real biology

# Fate probability estimation: HOW

FateID (2018)  
(*FateID* in R)

**FateID infers cell fate bias in multipotent progenitors from single-cell RNA-seq data**

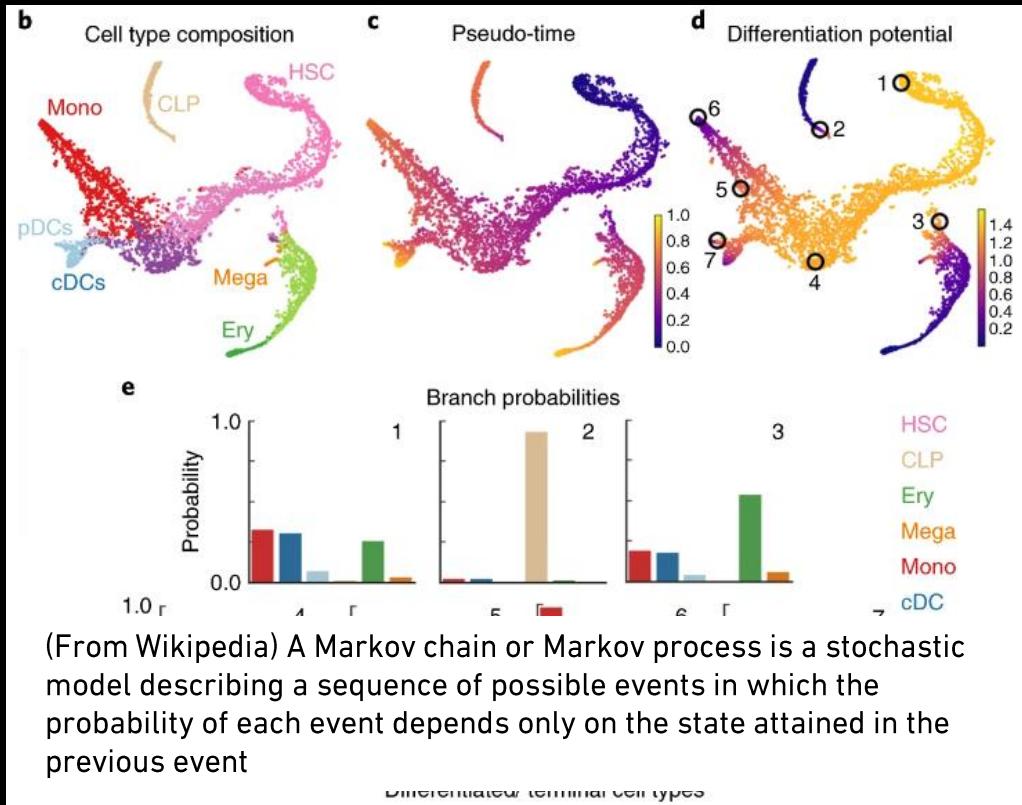
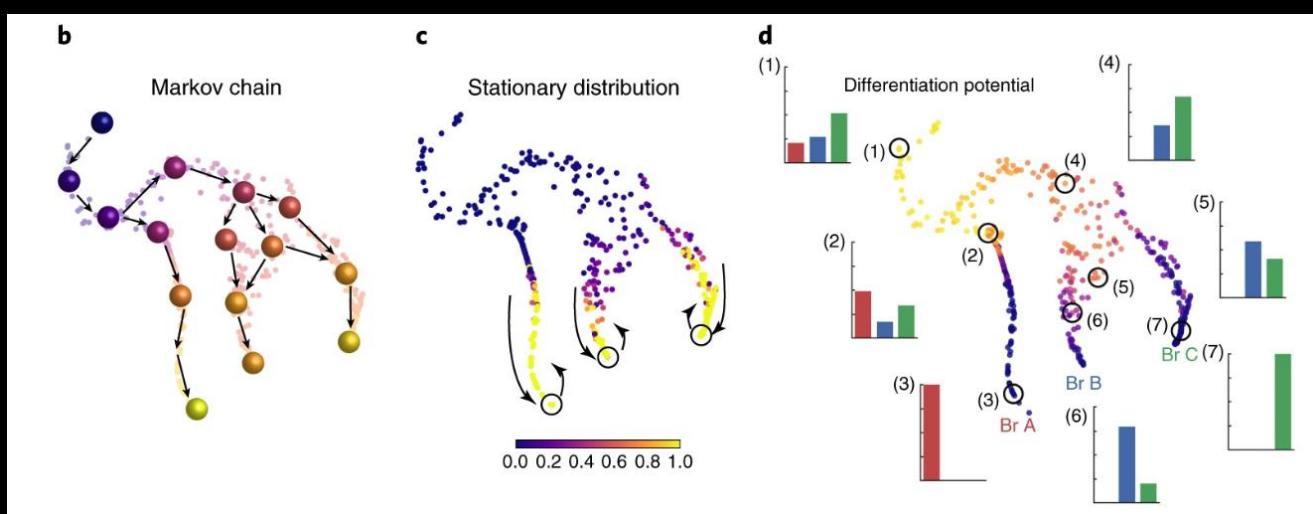
Josip S Herman<sup>1-3</sup>, Sagar<sup>1</sup> & Dominic Grün<sup>1</sup>



Pseudotime-free iterative cell type classification with random forest

# Fate probability estimation: HOW (2)

Palantir (2019)  
(*palantir* in Python)



Pseudotime  
(by shortest paths on diffusion neighborhood graph)

Pseudotime-biased transition matrix

Terminal states specification

Random walk through the absorbing Markov chain on "waypoint cells"

Propagate the information from "waypoint cells" to all cells

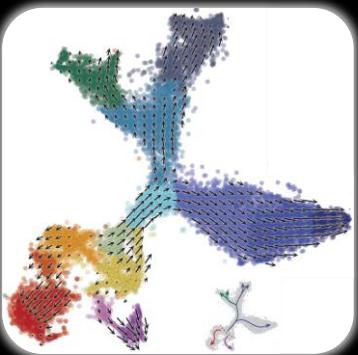
# Limitations of trajectory and pseudotime analysis

## scRNA-seq: a snapshot measure on every cell

- No real measurement on cell and cell state dynamics
  - Cell state dynamics are inferred based on transcriptomic similarity
- Domain-specific knowledge is needed to define root/initial (and terminal) states
  - Difficult to characterize unknown initial/terminal states

Can we recover cell dynamics from scRNA-seq data?

RNA velocity



# RNA velocity

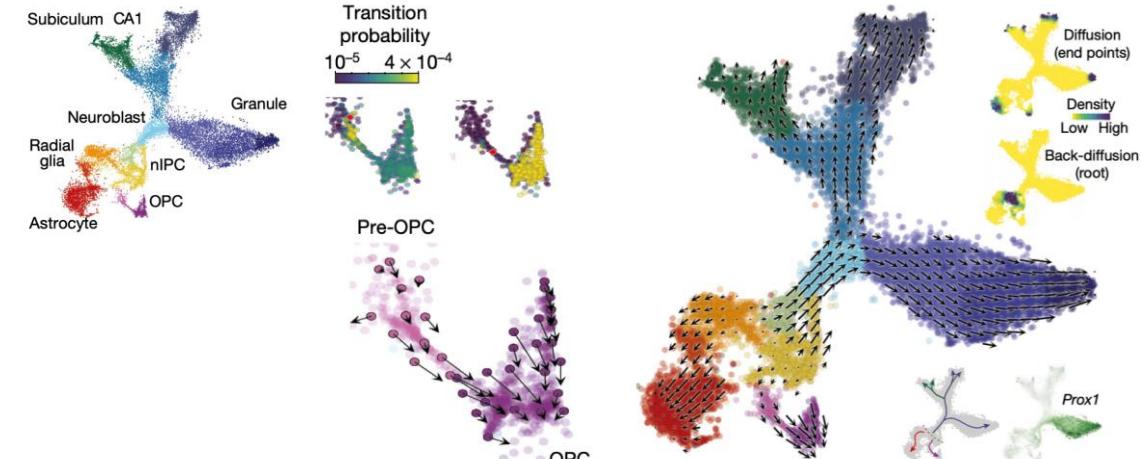
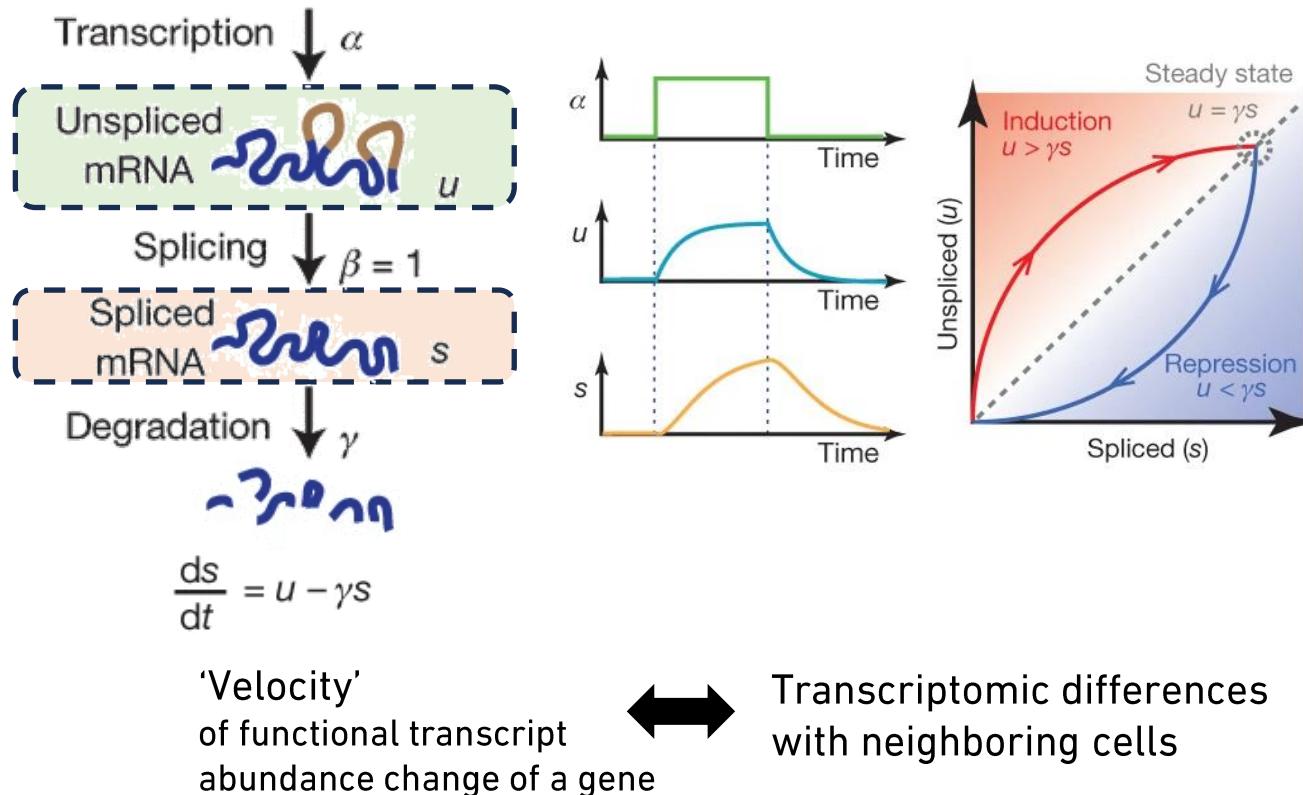
LETTER

La Manno et al. 2018

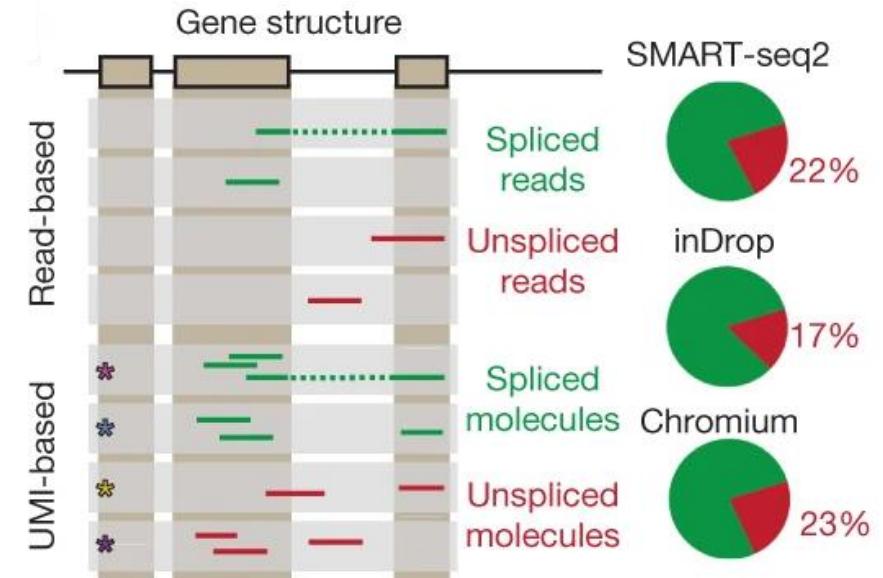
<https://doi.org/10.1038/s41586-018-0414-6>

## RNA velocity of single cells

Gioele La Manno<sup>1,2</sup>, Ruslan Soldatov<sup>3</sup>, Amit Zeisel<sup>1,2</sup>, Emelie Braun<sup>1,2</sup>, Hannah Hochgerner<sup>1,2</sup>, Viktor Petukhov<sup>3,4</sup>, Katja Lidschreiber<sup>5</sup>, Maria E. Kastriti<sup>6</sup>, Peter Lönnerberg<sup>1,2</sup>, Alessandro Furlan<sup>1</sup>, Jean Fan<sup>3</sup>, Lars E. Borm<sup>1,2</sup>, Zehua Liu<sup>3</sup>, David van Bruggen<sup>1</sup>, Jimin Guo<sup>3</sup>, Xiaoling He<sup>7</sup>, Roger Barker<sup>7</sup>, Erik Sundström<sup>8</sup>, Gonçalo Castelo-Branco<sup>1</sup>, Patrick Cramer<sup>5,9</sup>, Igor Adameyko<sup>6</sup>, Sten Linnarsson<sup>1,2\*</sup> & Peter V. Kharchenko<sup>3,10\*</sup>



(velocyto in R, velocyto.py in Python)



Oligo-T primer enrichment?

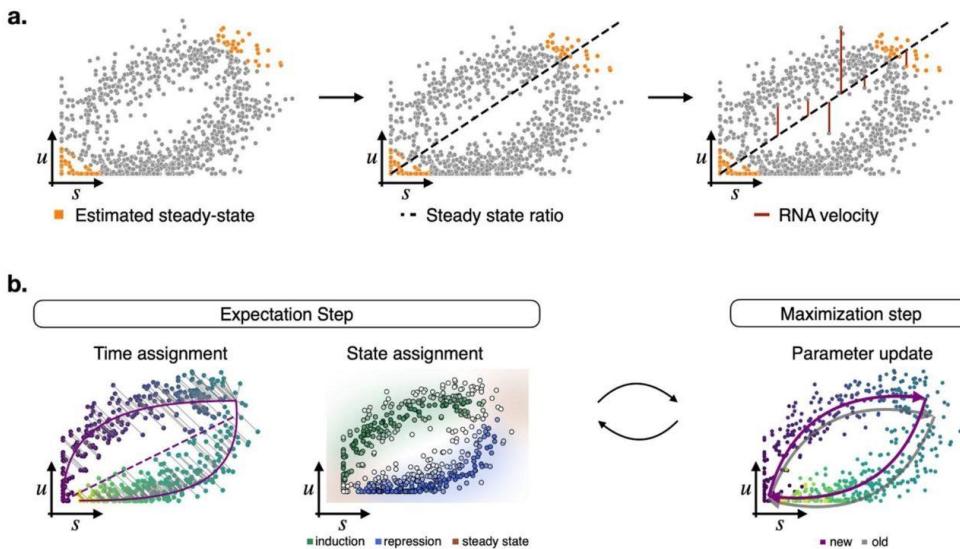
# RNA velocity (2)

## Limitations of *velocyto*

Assumptions must be made to infer the parameters

- Constant splicing rate for all genes
- Steady-state assumption for degradation rates
- Constant change rate of spliced molecule; or constant unspliced molecules

The R implementation is slow



Weiler et al. (2021) A guide to trajectory inference and RNA velocity. bioRxiv.  
<https://doi.org/10.1101/2021.12.22.473434>

Bergen et al. 2020  
(*scvelo* in Python)

## ARTICLES

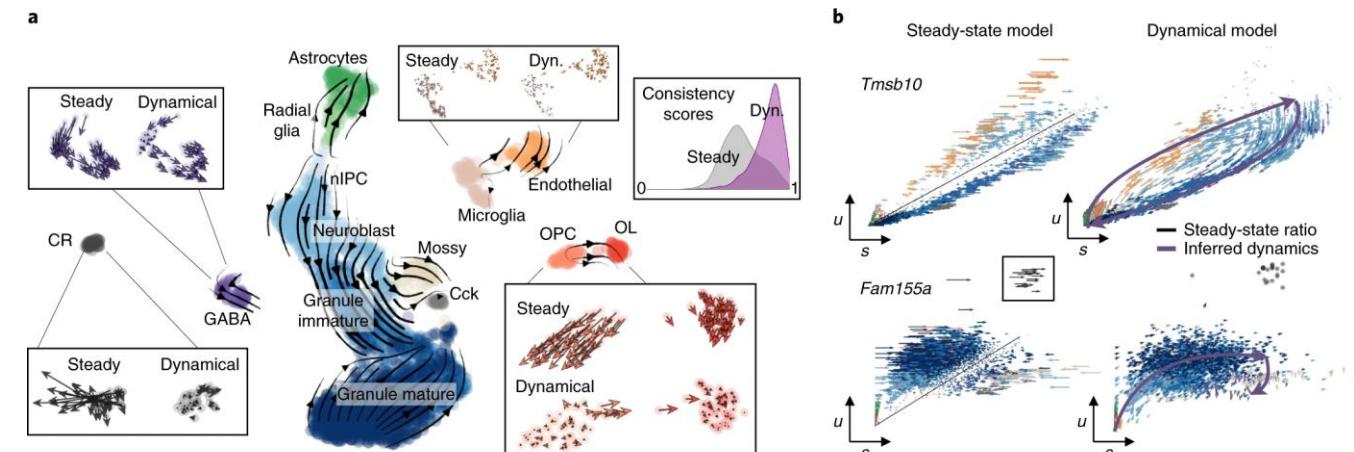
<https://doi.org/10.1038/s41587-020-0591-3>

nature  
biotechnology

Check for updates

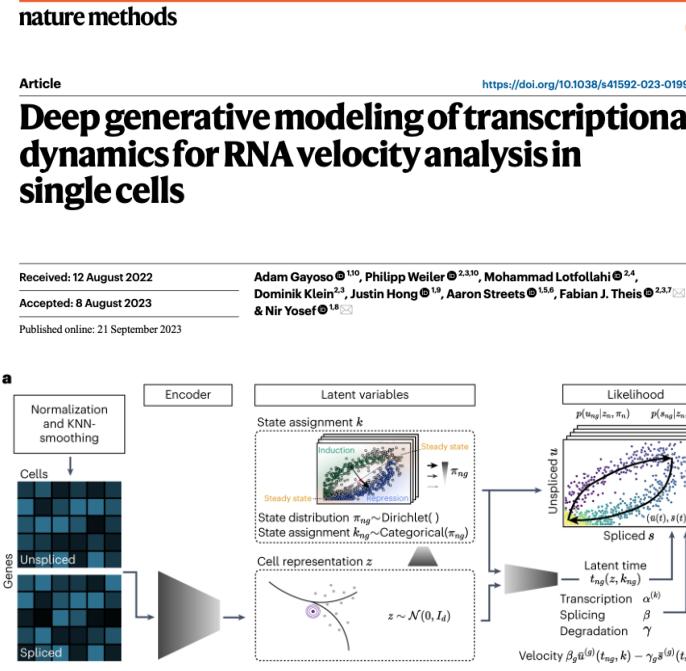
## Generalizing RNA velocity to transient cell states through dynamical modeling

Volker Bergen<sup>1,2</sup>, Marius Lange<sup>1,2</sup>, Stefan Peidli<sup>1</sup>, F. Alexander Wolf<sup>1</sup> and Fabian J. Theis<sup>1,2</sup>



# RNA velocity (3)

veloVI (2023)  
(*velovi* in Python)

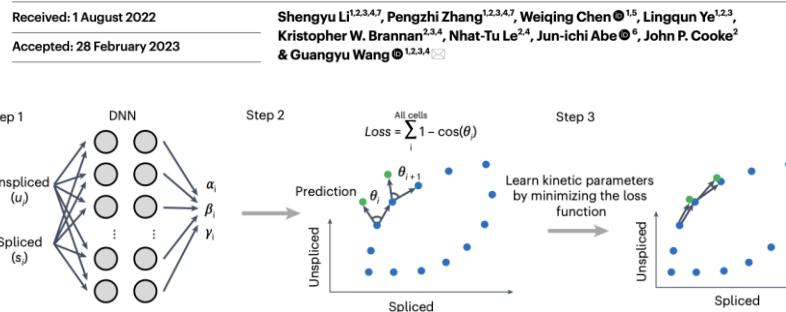


cellDancer (2023)  
(*celldancer* in Python)

**nature biotechnology** 

**Article** <https://doi.org/10.1038/s41587-023-01728-5>

**A relay velocity model infers cell-dependent RNA velocity**



DeepVelo (2024)  
(*deepvelo* in Python)

Cui et al. *Genome Biology* (2024) 25:27  
<https://doi.org/10.1186/s13059-023-03148-9>

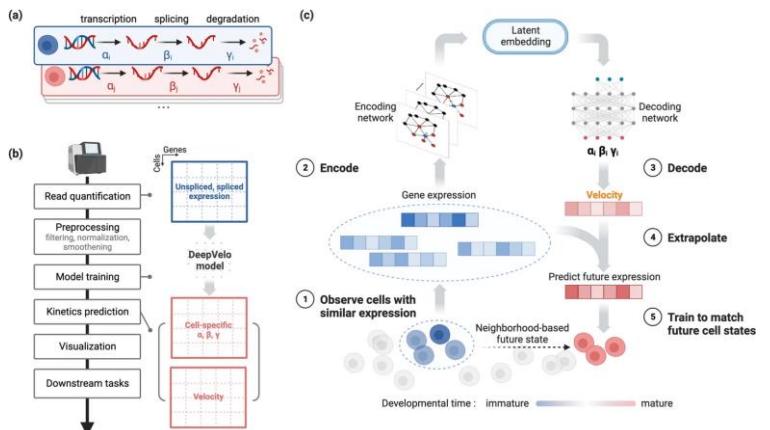
Genome Biology

## METHOD

Open Access

**DeepVelo: deep learning extends RNA velocity to multi-lineage systems with cell-specific kinetics**

Haotian Cui<sup>1,2,3†</sup>, Hassaan Maan<sup>1,3,4†</sup>, Maria C. Vladoiu<sup>5</sup>, Jiao Zhang<sup>6,7</sup>, Michael D. Taylor<sup>6,7,8,9</sup> and Bo Wang<sup>1,2,3,4,10</sup>

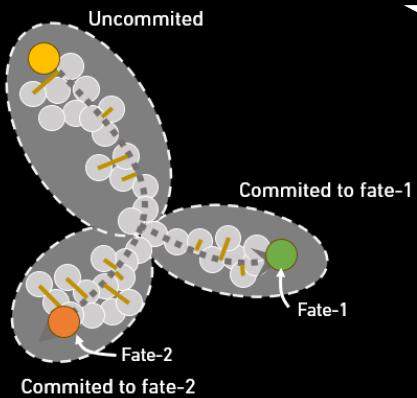
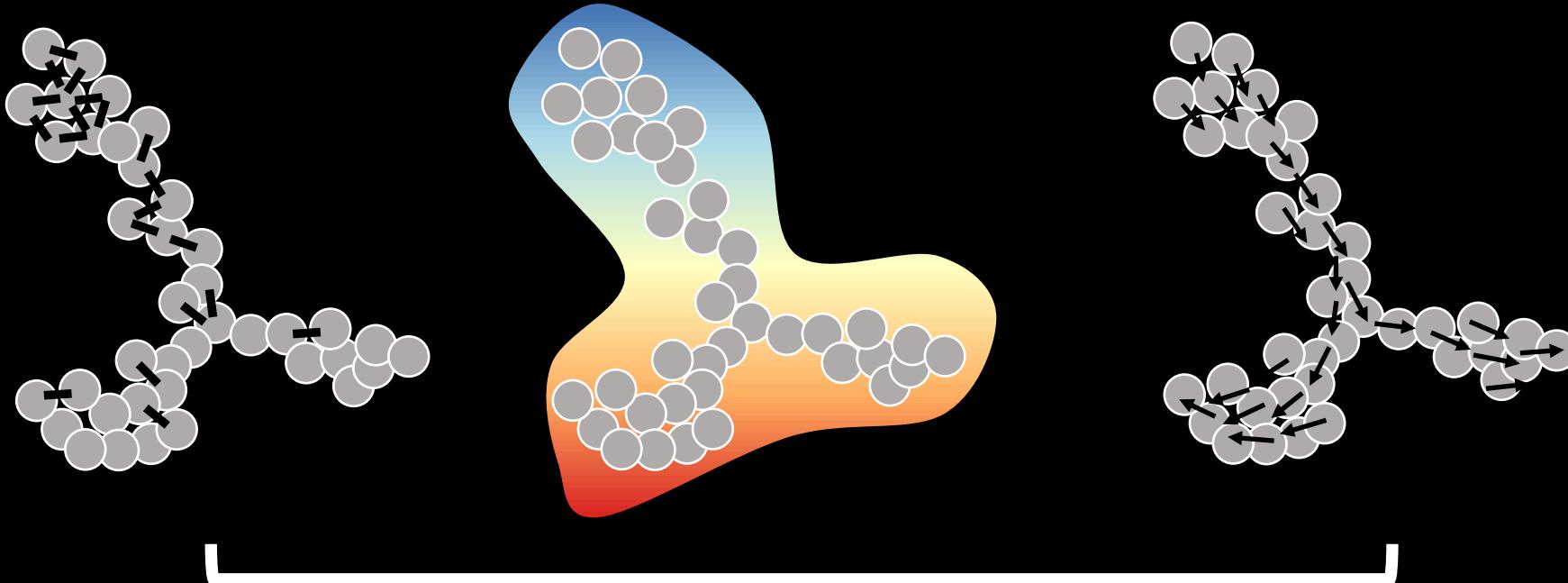


Use deep learning algorithms to infer gene-specific reaction rates for each cell

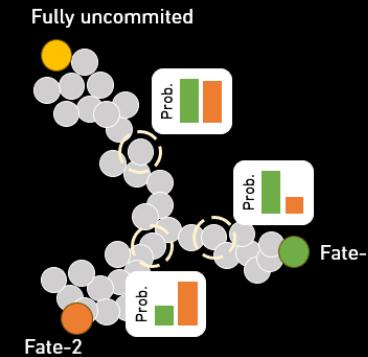
Transcriptomic  
similarities

Pseudotimes

RNA velocities



Trajectory analysis



Fate probability estimation

# Fate mapping combining different information

CellRank 2 (2024)  
(*cellrank* in Python)

nature methods

Article

<https://doi.org/10.1038/s41592-024-02303-9>

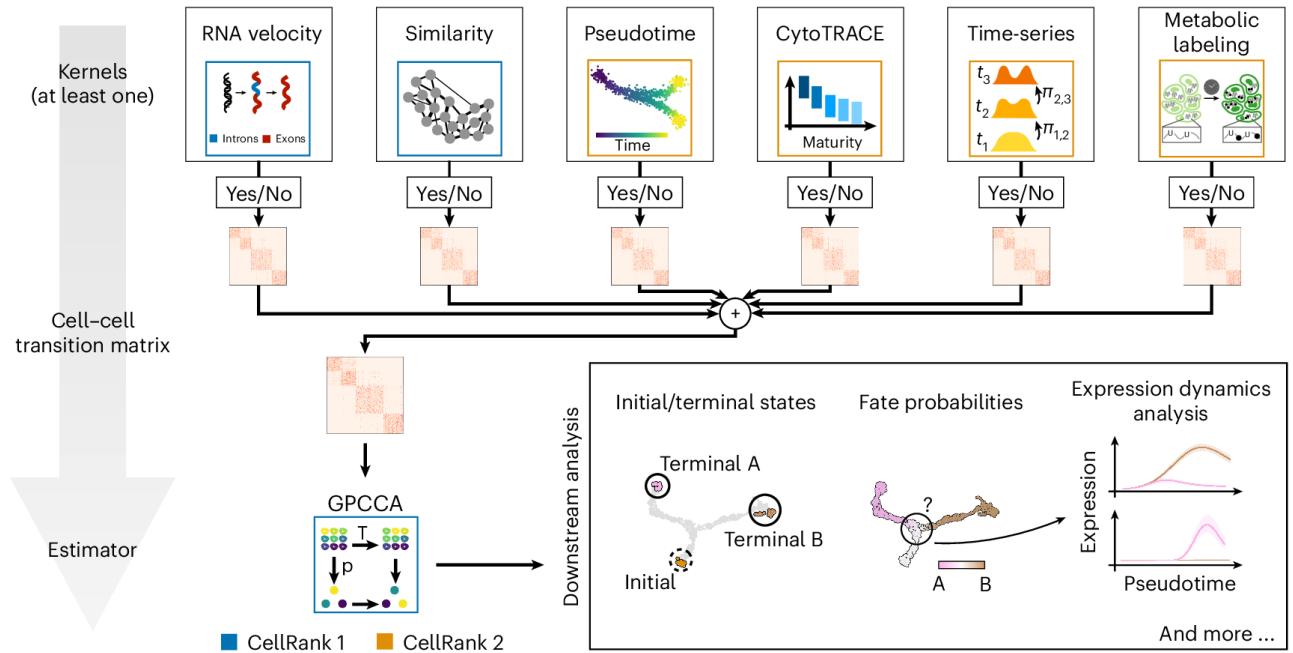
## CellRank 2: unified fate mapping in multiview single-cell data

Received: 18 July 2023

Philipp Weiler<sup>1,2,8</sup>, Marius Lange<sup>1,2,3,8</sup>, Michal Klein<sup>1,4</sup>, Dana Pe'er<sup>5,6</sup> & Fabian Theis<sup>1,2,7</sup>

Accepted: 9 May 2024

- Integrate different transition matrices based on different information
- Random walk on an absorbing Markov chain (very efficient implementation)



nature biotechnology

<https://doi.org/10.1038/s41587-023-01747-2>

## Multimodal spatiotemporal phenotyping of human retinal organoid development

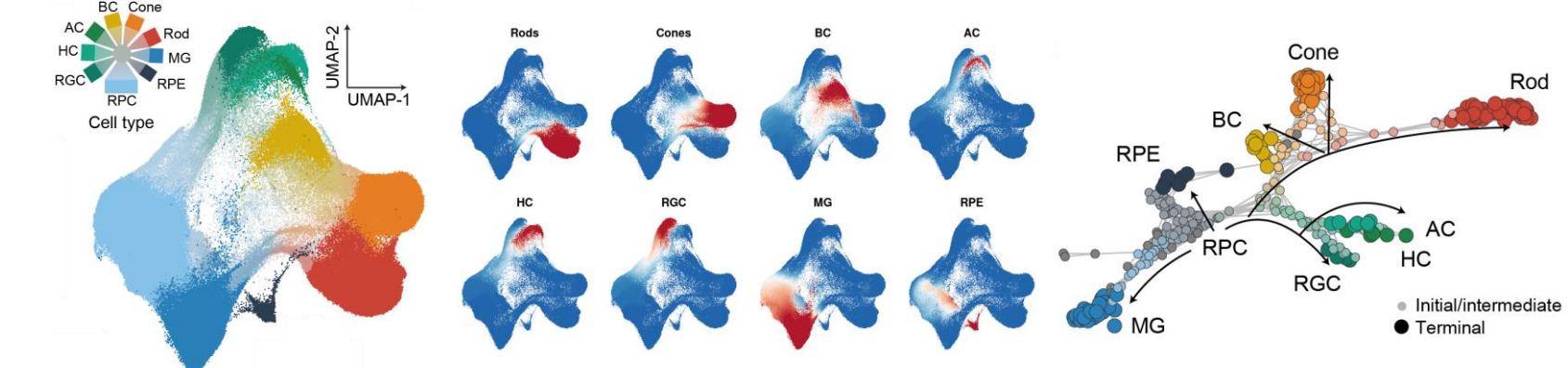
Received: 3 March 2022

Philipp Wahle<sup>1,2,3,6</sup>, Giovanna Brancati<sup>1,2,3,6</sup>, Christoph Harmel<sup>1,2,3,5,6</sup>, Zhisong He<sup>1,6</sup>, Gabriele Gut<sup>1</sup>, Jacobo Saravia del Castillo<sup>1,4</sup>, Aline Xavier da Silveira dos Santos<sup>1,2,3,5</sup>, Qianhui Yu<sup>2,3,5</sup>, Pascal Noser<sup>1</sup>, Jonas Simon Fleck<sup>1</sup>, Bruno Cjetz<sup>2,3,5</sup>, Dinko Pavlinić<sup>2,3</sup>, Simone Picelli<sup>1,2,3</sup>, Max Hess<sup>1,4</sup>, Gregor W. Schmidt<sup>1</sup>, Tom T. A. Lummen<sup>1</sup>, Yanyan Hou<sup>2,3</sup>, Patricia Gallikker<sup>2,3</sup>, David Goldblum<sup>2,3</sup>, Marton Balogh<sup>2,3</sup>, Cameron S. Cowan<sup>2</sup>, Hendrik P. N. Scholl<sup>2,3</sup>, Botond Roska<sup>2,3</sup>, Magdalena Renner<sup>2,3</sup>, Lucas Pelkmans<sup>1,4</sup>, Barbara Treutlein<sup>1,2,3</sup> & J. Gray Camp<sup>1,2,3,5</sup>

Accepted: 13 March 2023

Published online: 8 May 2023

Check for updates



# Last but not the least...

All analysis have hyperparameters, which can greatly affect the results

- Iterations and hyperparameter finetuning are sometimes necessary
- Be careful of the result robustness to the hyperparameters and input!

All analysis can output wrong and misleading results, even if you don't make any technical mistake

Limitations on the computational methods

Data quality

- Be cautious to any result you get from any analysis
- Don't easily make any conclusion based on a single analysis.  
Evidences from complementary analysis and/or more experiments are critical!

# Questions?