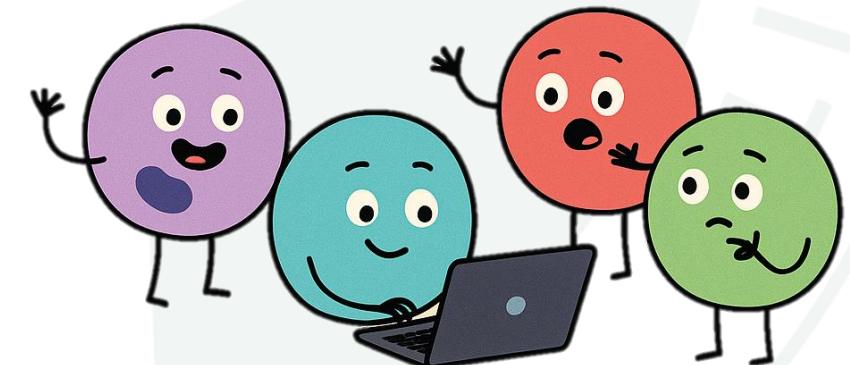


scTalk 3: How to Compute Trajectories and Pseudotime

Arash Bagherabadi
Feb18, 2026



dana_peer @dana_peer · Feb 6
This is the way to compute trajectories and pseudotime

Fabian Theis @fabian_theis · Feb 5
CellRank protocol now in @NatureProtocols! 🎉

@PhilippWeiler's detailed guide shows how to:
- Integrate velocity, pseudotime & time points
- Infer cell fate probabilities...

The figure displays two UMAP plots illustrating cell trajectory analysis. The left plot shows a complex, branching cluster of cells colored by cell type, with axes labeled UMAP₁ and UMAP₂. A legend identifies the following cell types: HSC (pink), MEP (light green), Ery (green), CLP (yellow), HMP (purple), Monocyte (red), Pre-DC (light blue), cDC (dark blue), and pDC (light blue). An arrow labeled "CellRank" points to the right plot. The right plot shows the same UMAP space, but the cells have been projected onto a simplified, more linear trajectory, representing the inferred pseudotime path. The final "Terminal states" are highlighted in red, green, yellow, and blue.

Single Cell 101 ?

BOX1

Common approaches for inferring cellular trajectories

Pseudotime

Pseudotime algorithms align single-cell data along a one-dimensional manifold to reconstruct dynamical processes from snapshot data, commonly relying on prior information to direct the inference. Immature cells score a low pseudotime that increases with maturity^{13,26,39,40,63,64}.

RNA velocity

RNA velocity relates unspliced and spliced mRNA through a dynamical model describing splicing dynamics; RNA velocity can be inferred from single-cell transcriptomics data because common experimental assays contain the relevant information for reconstructing unspliced and spliced counts. Various inference algorithms have been proposed, each posing method-specific assumptions^{30,35,36,46}.

OT

OT leverages time point information to match cells at one time point with their putative future state at a later time point in a probabilistic fashion. On a high level, OT transforms the original distribution into the distribution of progenitor states in an optimal manner^{15,24,60,65}.

Metabolic labeling

Metabolic labeling experiments label newly transcribed mRNA molecules with nucleotide analogs, thereby introducing temporally related quantities such as unspliced and spliced mRNA. Compared with classical sequencing data from distinct time points, metabolic labeling uses shorter time scales, allowing it to reveal biological mechanisms and distinguish regulatory effects^{16,31,51,66–69}.

BOX 1

Common approaches for inferring cellular trajectories

Pseudotime ↗ Monocle!, Slingshot!, PAGA!, Palantir!

Pseudotime algorithms align single-cell data along a one-dimensional manifold to reconstruct dynamical processes from snapshot data, commonly relying on prior information to direct the inference. Immature cells score a low pseudotime that increases with maturity^{13,36,39,40,63,64}.

RNA velocity 🚀 velocyto!, scVelo!

RNA velocity relates unspliced and spliced mRNA through a dynamical model describing splicing dynamics; RNA velocity can be inferred from single-cell transcriptomics data because common experimental assays contain the relevant information for reconstructing unspliced and spliced counts. Various inference algorithms have been proposed, each posing method-specific assumptions^{30,35,36,46}.

OT ↗ WOT! scOT!

OT leverages time point information to match cells at one time point with their putative future state at a later time point in a probabilistic fashion. On a high level, OT transforms the original distribution into the distribution of progenitor states in an optimal manner^{15,24,60,65}.

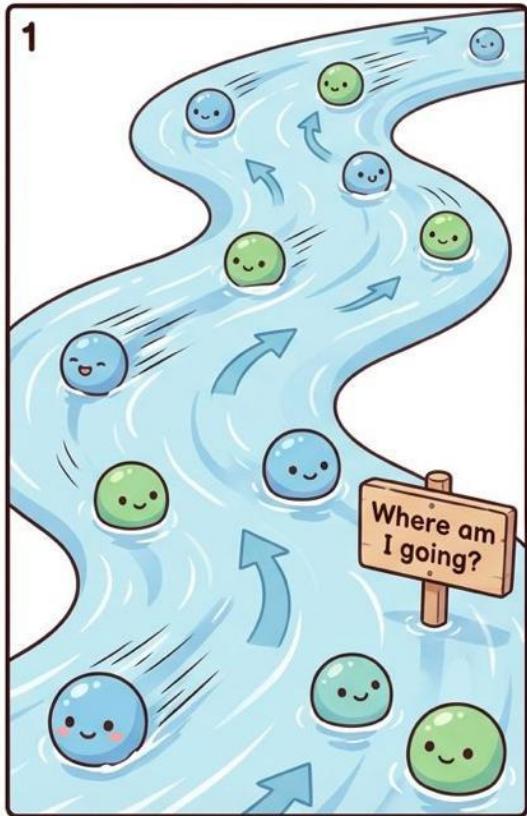
Metabolic labeling scSLAM-seq!, TimeLapse-seq!

Metabolic labeling experiments label newly transcribed mRNA molecules with nucleotide analogs, thereby introducing temporally related quantities such as unspliced and spliced mRNA. Compared with classical sequencing data from distinct time points, metabolic labeling uses shorter time scales, allowing it to reveal biological mechanisms and distinguish regulatory effects^{16,31,51,66–69}.

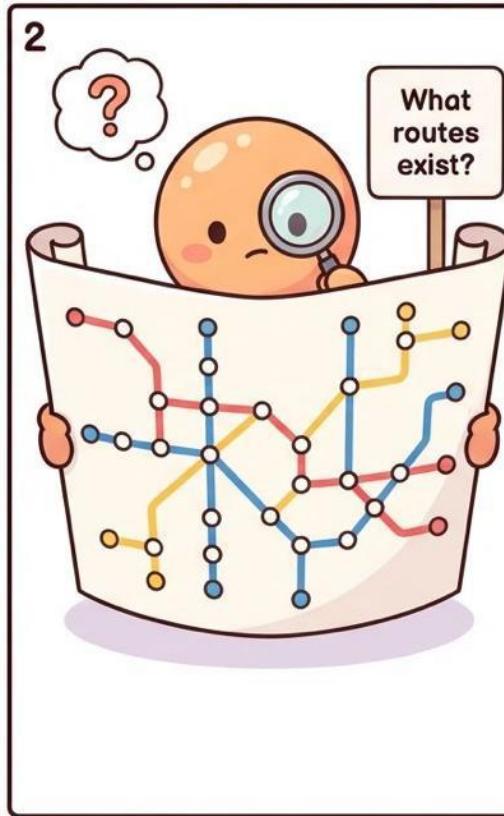
where cells are heading

+ how likely they are to end up there

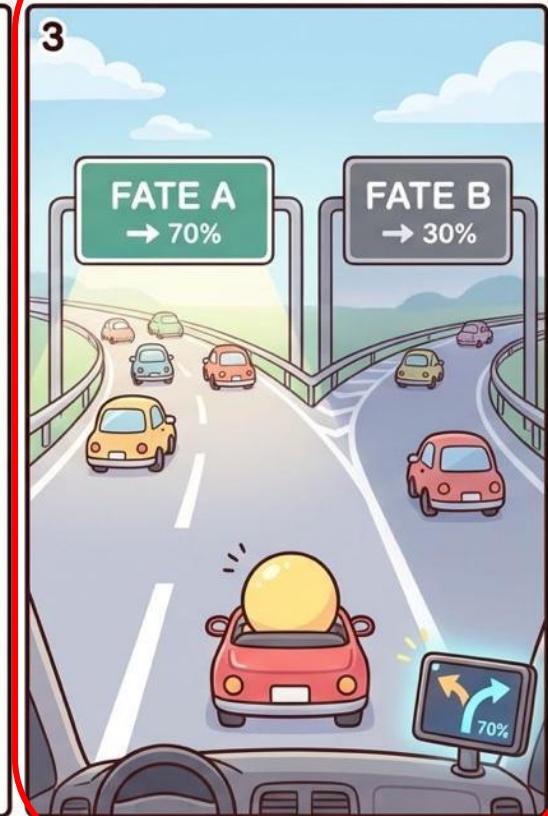
scVelo - The River



Monocle3 - The Map



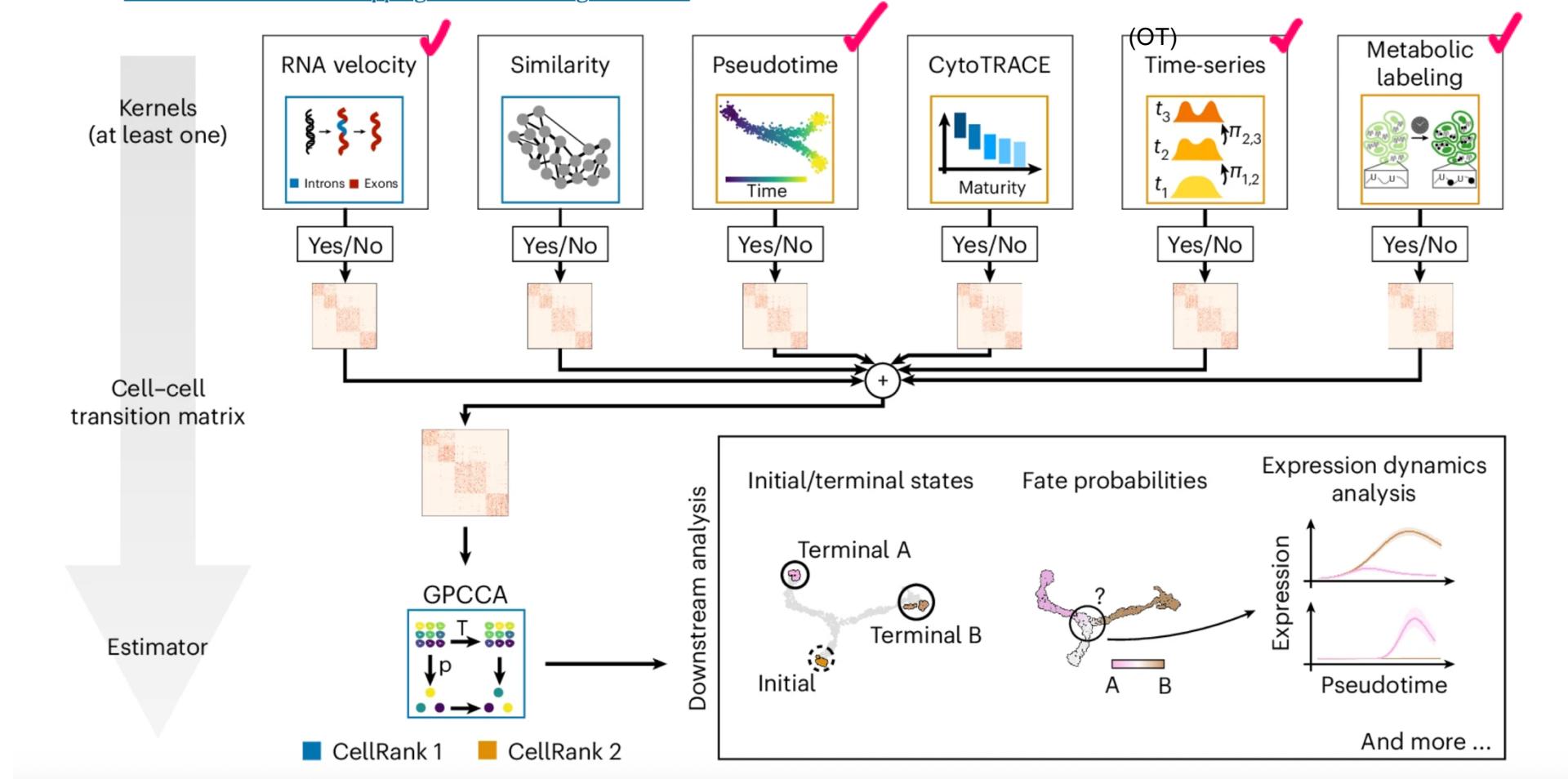
CellRank2 - GPS Navigation



Xinru Qiu, LinkedIn

Fig. 1: CellRank 2 provides a unified framework for studying single-cell fate decisions using Markov chains.

From: [CellRank 2: unified fate mapping in multiview single-cell data](#)



[Explore content](#) [About the journal](#) [Publish with us](#)

nature > nature methods > articles > article

Article [Open access](#) Published: 13 January 2022

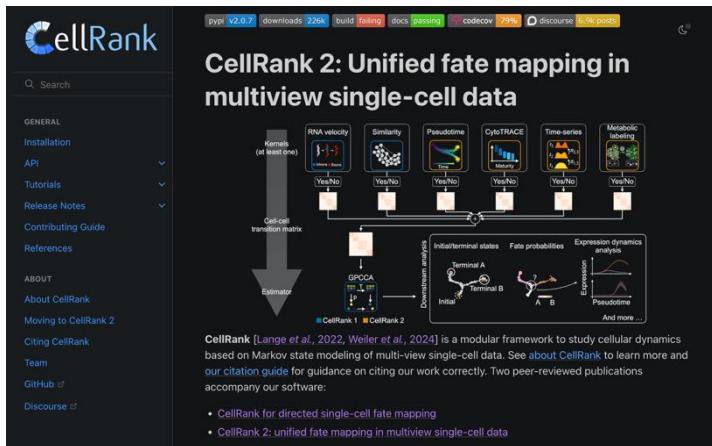
CellRank for directed single-cell fate mapping

Marius Lange, Volker Bergen, Michal Klein, Manu Setty, Bernhard Reuter, Mostafa Bakhti, Heiko Lickert
Meshal Ansari, Janine Schniering, Herbert B. Schiller, Dana Pe'er✉ & Fabian J. Theis✉

Nature Methods 19, 159–170 (2022) | Cite this article

155k Accesses | 686 Citations | 445 Altmetric | Metrics

<https://www.nature.com/articles/s41592-021-01346-6>



<https://cellrank.readthedocs.io/en/latest/>

1

[Explore content](#) ▾ [About the journal](#) ▾ [Publish with us](#) ▾

nature > nature methods > articles > article

Article | Open access | Published: 13 June 2024

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

Philipp Weiler, Marius Lange, Michal Klein, Dana Pe'er & Fabian Theis

Nature Methods 21: 1196–1205 (2024) | Cite this article

49k Accesses | **155** Citations | **81** Altmetric | Metrics

<https://www.nature.com/articles/s41592-024-02303-9>

3

nature protocols

[Explore content](#) | [About the journal](#) | [Publish with us](#)

For more information about the study, please contact Dr. Michael J. Hwang at (310) 794-3000 or via email at mhwang@ucla.edu.

CellRank: consistent and d

CellRank: consistent and data view agnostic fate mapping for single-cell genomics

Philip Weiler & Fabian J. Theis 

Nature Protocols (2026) | Cite this article

4837 Accesses | 64 Altmetric | Metrics

<https://www.nature.com/articles/s41596-025-01314-w>

Other Kernels Added

Protocol for Cell Rank2

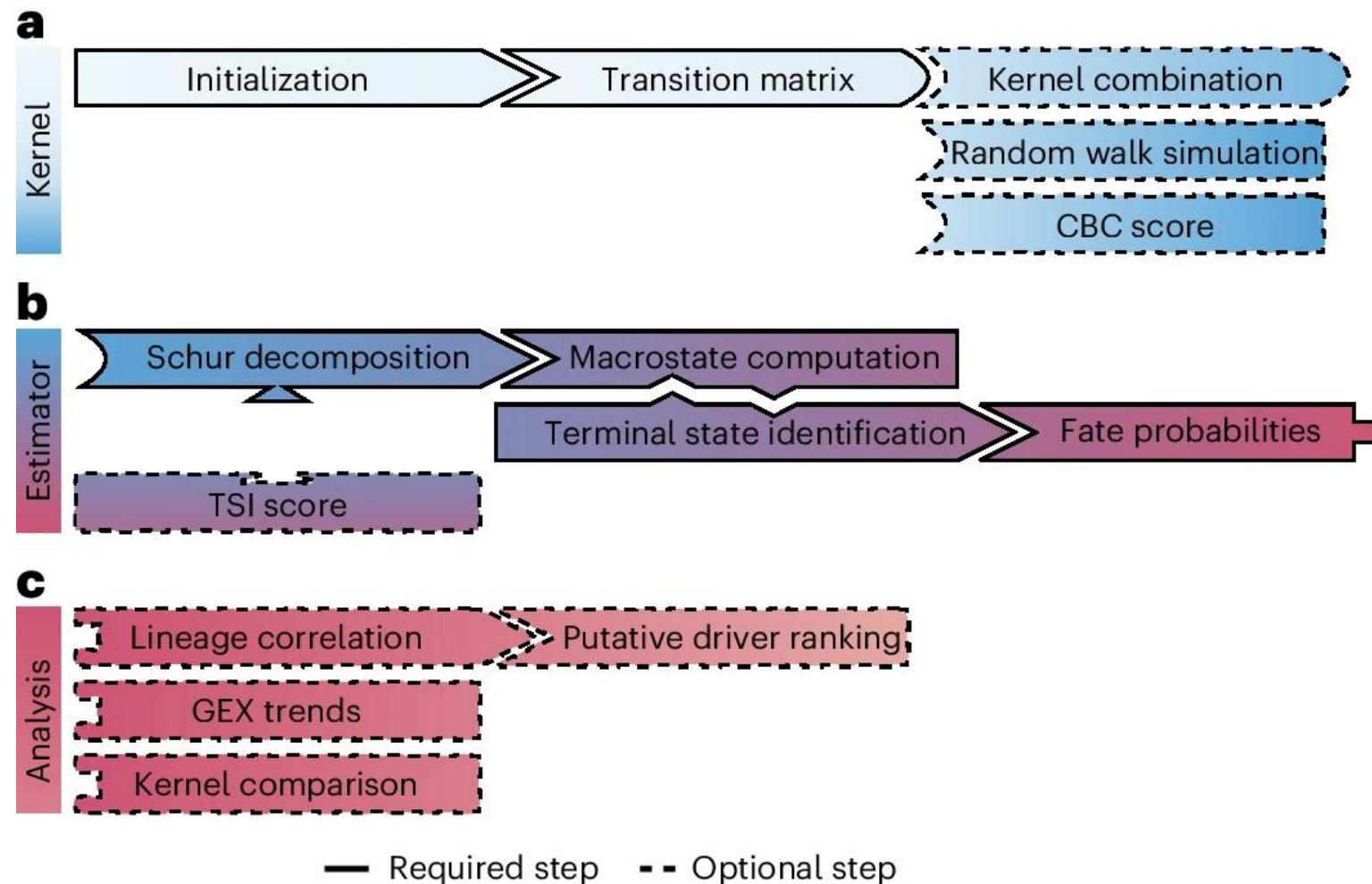
- Discussion-driven single-cell journal club
- Break down key/new papers
- Shared standards and critical thinking in analysis
- Open questions, interruptions, and debate encouraged



Please Sign-in +
Comments & Questions

Fig. 1: CellRank is a modular and scalable framework for cellular fate mapping.

| 10



BOX 2

Details on kernel inputs

ConnectivityKernel

The ConnectivityKernel requires a precomputed, symmetric cell-cell similarity graph. The kernel is independent of the computation of the similarity matrix and only requires it as a field in the obsp slot of the AnnData object.

PseudotimeKernel

Although each pseudotime method has its own assumptions, CellRank works with any pseudotime. The pseudotime must be saved as a column in the obs slot of the AnnData object and the column name specified when initializing the PseudotimeKernel. In addition, the PseudotimeKernel requires a precomputed, symmetric cell-cell similarity graph, saved in the obsp slot of the AnnData object.

CytoTRACEKernel

Inferring the CytoTRACE score as a stemness potential does not require any additional input beyond GEX counts. The kernel converts these cell scores into a cell-cell transition matrix, by biasing an undirected cell-cell similarity graph, saved in the obsp slot of the AnnData object, into the direction of decreased potential.

VelocityKernel

CellRank decouples the inference of cell-cell transition probabilities on the basis of cellular velocity estimates from the actual velocity inference. Estimating cell-cell transition probabilities requires both cell-specific GEX representation and velocity estimates, saved either in the layers or obsm slot of the AnnData object, and a precomputed, symmetric cell-cell similarity graph, saved in the obsp slot; the corresponding field names are passed to the VelocityKernel upon initialization.

RealTimeKernel

The RealTimeKernel connects cells across distinct experimental time points using OT. It reads the experimental time point of each cell from a corresponding column in the obs slot of the AnnData object; the corresponding column name is passed to the kernel when initializing it.

PrecomputedKernel

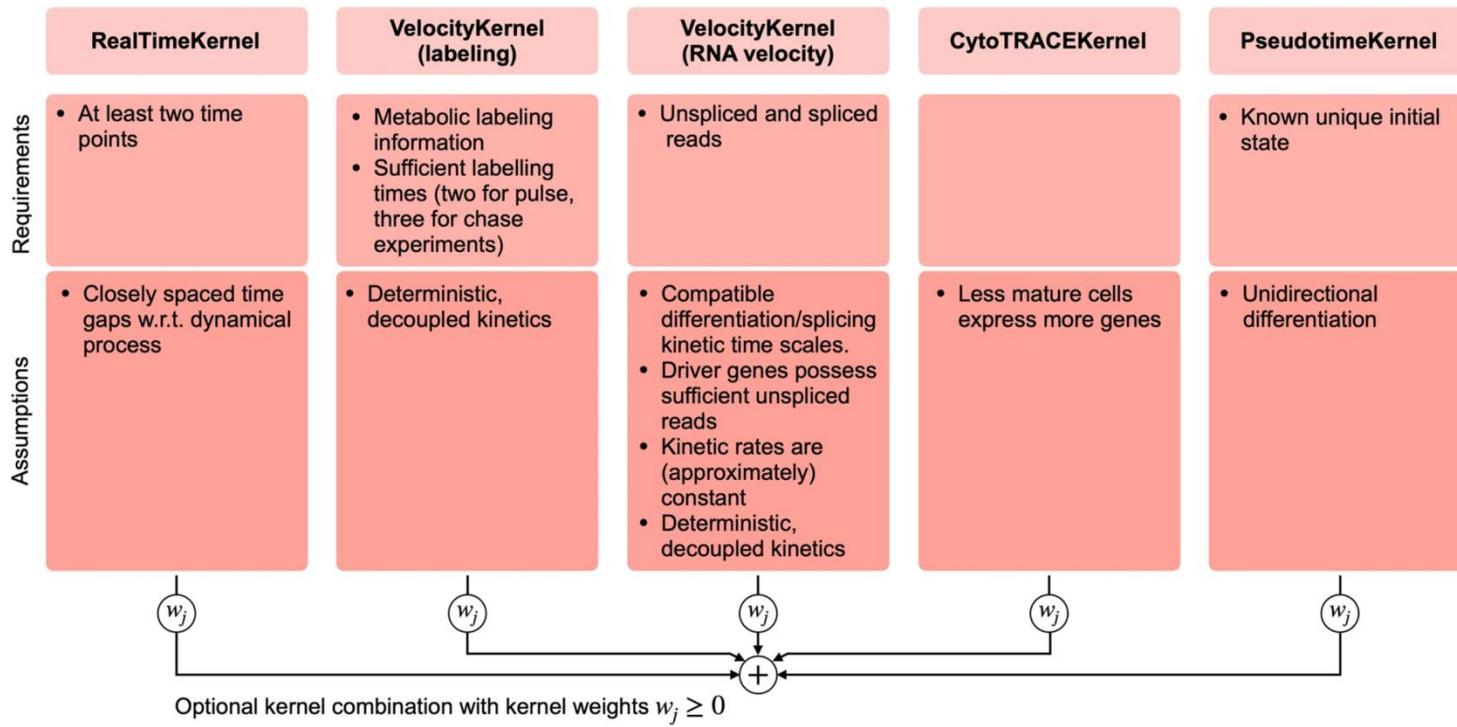
Users computing cell-cell transition matrices independent of CellRank can still rely on the CellRank framework for trajectory inference through the PrecomputedKernel. The kernel requires the transition matrix in the obsp slot of the AnnData object.

Decision Tree for Choosing Kernels

12

Extended Data Fig. 1: Guiding kernel choice in CellRank 2.

From: [CellRank 2: unified fate mapping in multiview single-cell data](#)



CellRank 2 implements various kernels suitable for different data modalities and experimental designs. The diagram may be used as a guide to identify the most suitable kernel. Note that assumptions change as methods evolve; for example, more recent inference schemes for RNA velocity account for non-constant kinetic rates⁹³.

CellRank Kernels have also been used in the Moscot paper

13

Mapping cells through time and space with moscot

Trajectory inference with different trajectory-inference methods

We used diffusion pseudo-time¹³⁴, scVelo³⁹, veloVI¹³⁵, MultiVelo¹³⁶, CytoTrace¹³⁷ and the ConnectivityKernel¹⁶ in CellRank to predict trajectories in the pancreatic endocrinogenesis dataset. As we were interested in the endocrine-cell trajectories, we filtered the dataset to endocrine cells and their progenitors. We then applied the GPCCA estimator in CellRank to each corresponding trajectory-inference kernel. To compute fate probabilities, we used compute_fate_probabilities and aggregate_fate_probabilities to plot the fate probabilities and aggregated cell-type to cell-type transition matrices. We used the plot_projection method to generate the stream embedding plots. We used all default arguments provided for the methods, and only increased max_epochs in VeloVI to 50. When using graphs, we used the WNN graph as described above. For building the RealTimeKernel with moscot, we set the weight of the ConnectivityKernel to 0.001 to strengthen the influence of moscot and weaken the influence of the ConnectivityKernel. We highlight that the transition probabilities computed with CellRank rely on a different procedure than the transition probabilities we computed with moscot.

<https://www.nature.com/articles/s41586-024-08453-2#Sec8>

Different Velocity Methods

14



New Results

Benchmarking RNA velocity methods across 17 independent studies

Ya Luo, Jun Ren, Qian Yang, Ying Zhou, Zhiyu You, Qiyuan Li

doi: <https://doi.org/10.1101/2025.08.02.668272>

This article is a preprint and has not been certified by peer review [what does this mean?].

HOME | SUBMIT | FAQ | BLOG | ALERTS / RSS |

Search

Follow this preprint ← Previous

Posted August 02, 2025.

Download PDF

Print/Save Options

Supplementary Material

<https://www.biorxiv.org/content/10.1101/2025.08.02.668272v1>

C

14 methods

Differential equation-based:

velocyto, scvelo-sto,
scvelo-dyn, Multivelox

Machine learning-based:

uniTvelo, Dynamo,
Pyro-velocity, cell2fate

Deep learning- based:

veloAE, veloVI,
veloVAE, latentVelo,
cellDancer, Deepvelo

17 scRNA-seq datasets



5 human datasets



12 mouse datasets

BOX3

Markov chain-related concepts and definitions

Markov chain

Markov chains describe dynamic processes by connecting a set of states probabilistically. In the context of CellRank, the states are the sequenced cells.

Transition probability

Transition probabilities quantify how likely one cell is the ancestor state of another.

Random walks

Starting from a given cell and moving to one of its putative future states on the basis of the inferred cell-cell transition probabilities defines a random walk. Random walks provide a simulation framework to assess the induced Markov chain.

Macrostate

States and the underlying dynamics described by a Markov chain may be noisy. Macrostates represent relevant biological states, that is, a denoised and coarse-grained version of the original Markov chain.

Initial state

The initial state of a system is the state where the underlying dynamical process starts.

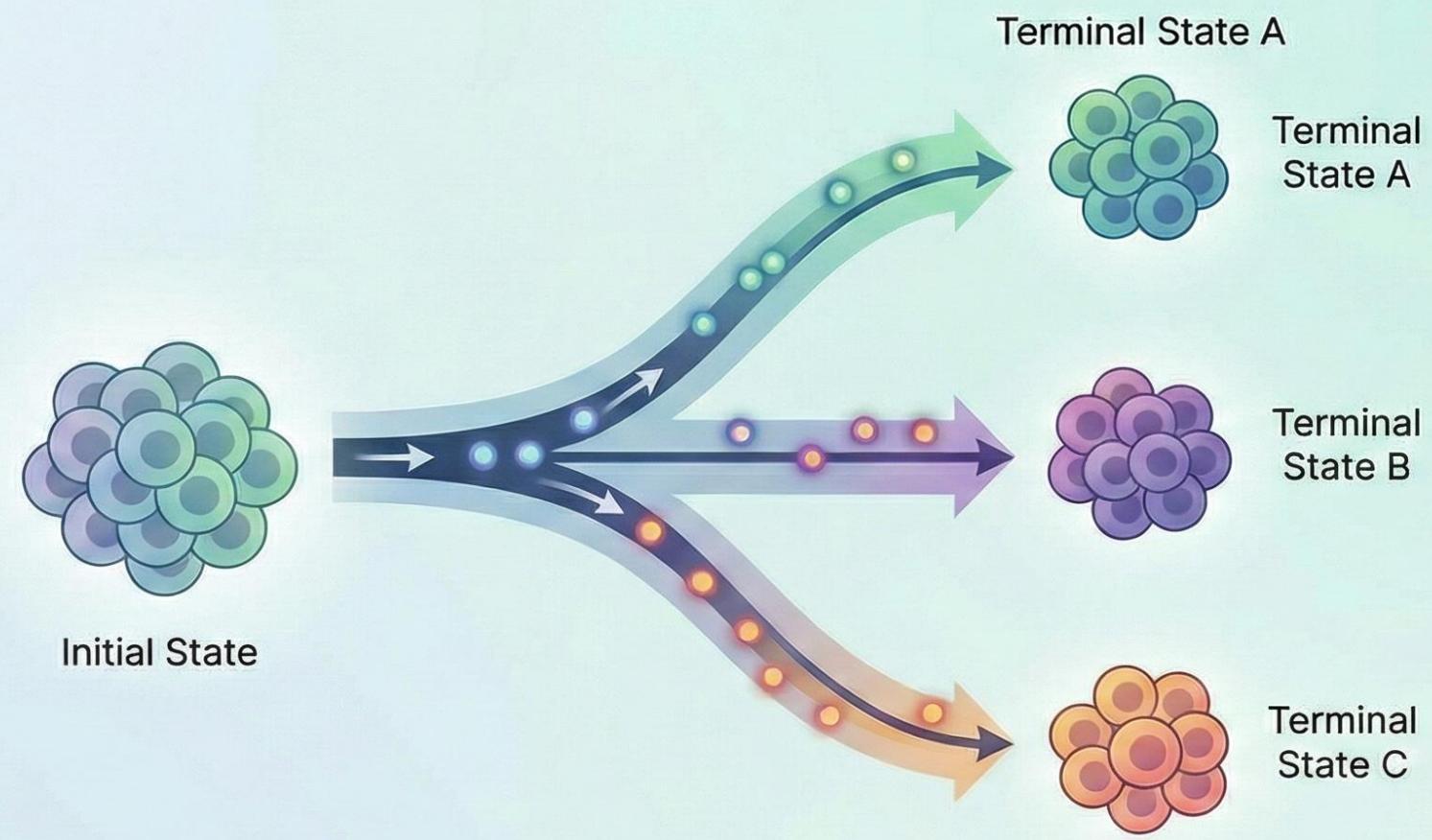
Terminal state

Terminal states are states present in the data for which no progenitor states exist. These states may be fully differentiated cell types or intermediate states for which no future states have been captured.

Fate probability

Fate probabilities describe the probability of a cell to differentiate into a terminal state.

- *CellRank models average cellular behavior, assuming incremental, memoryless change along the phenotypic manifold.*
- *CellRank collects the cell-cell transition probabilities in a cell-cell transition matrix that defines/induces a Markov chain and for each matrix an estimator analyzes the induced Markov chain to infer the initial and terminal states of a biological process, quantify the corresponding fate probabilities, describe their lineage formation and enable further downstream analyses.*
- *Generalized Perron cluster cluster analysis (GPCCA) estimator coarse grains the transition matrix to define macro states and transition probabilities between them, thereby defining terminal states and fate probabilities.*



Thank You!

$$\mathcal{P}(x_{t+1} | x_t)$$



Questions?



Supp Slides

Single Cell 101

The Body as a Country!

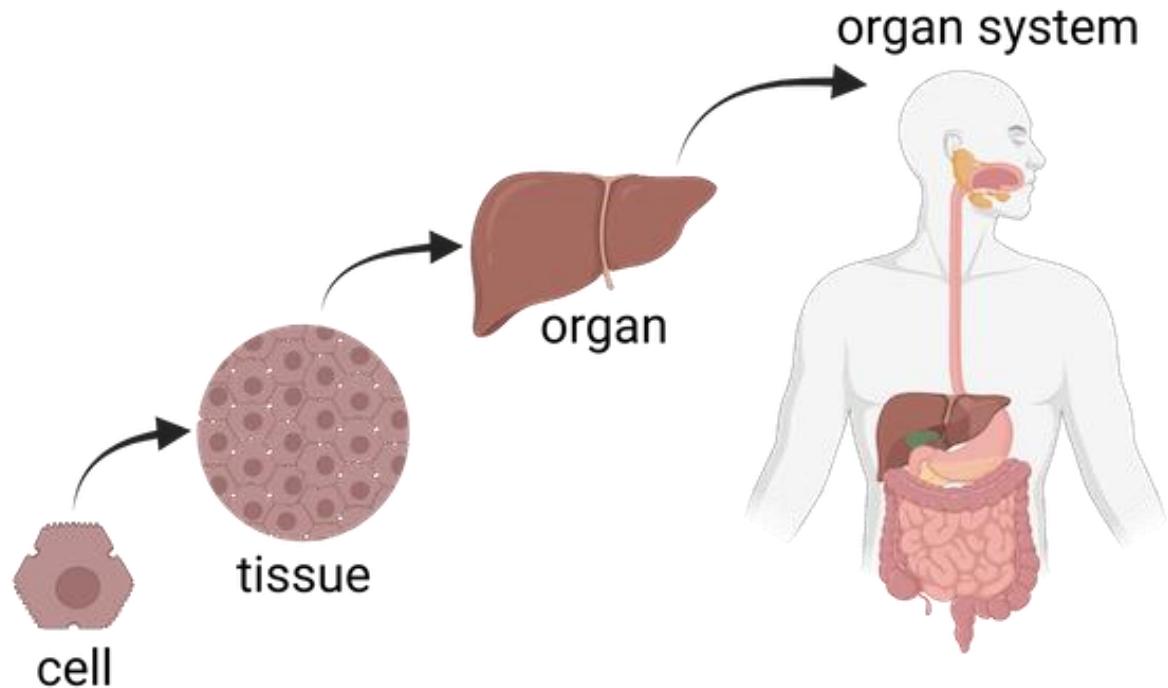
- **Body** = USA
- **Organs** = States with specialized roles (Brain, Heart, Lungs, etc.)
- **Tissues** = Cities (Cardiac muscle, Alveolar tissue, etc.)
- **Niche** = Neighborhoods
- **Cells** = Buildings / Organizations / Companies



Cells coordinate to keep each “city” functioning!

| 20

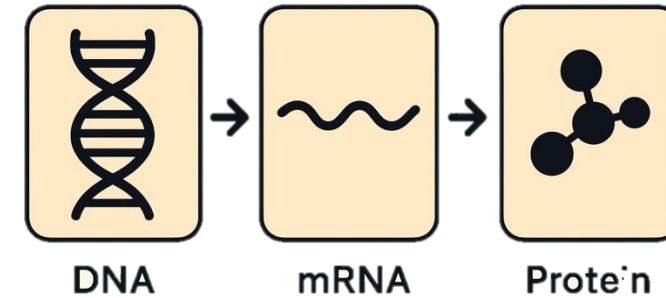
- A cell is the **basic membrane-bound unit that contains the fundamental molecules of life** and of which all living things are composed.
- A single cell is often a **complete organism in itself**.



<https://www.britannica.com/science/cell-biology>

<https://www.khanacademy.org/>

- **Every cell** = a self-run organization inside the city
- **DNA** = full archive of laws and manuals
- **Genes** = individual regulations
- **mRNA** = copied memos for active tasks
- **Proteins** = workers, machines, and tools
- **Signals** = inter-building communication



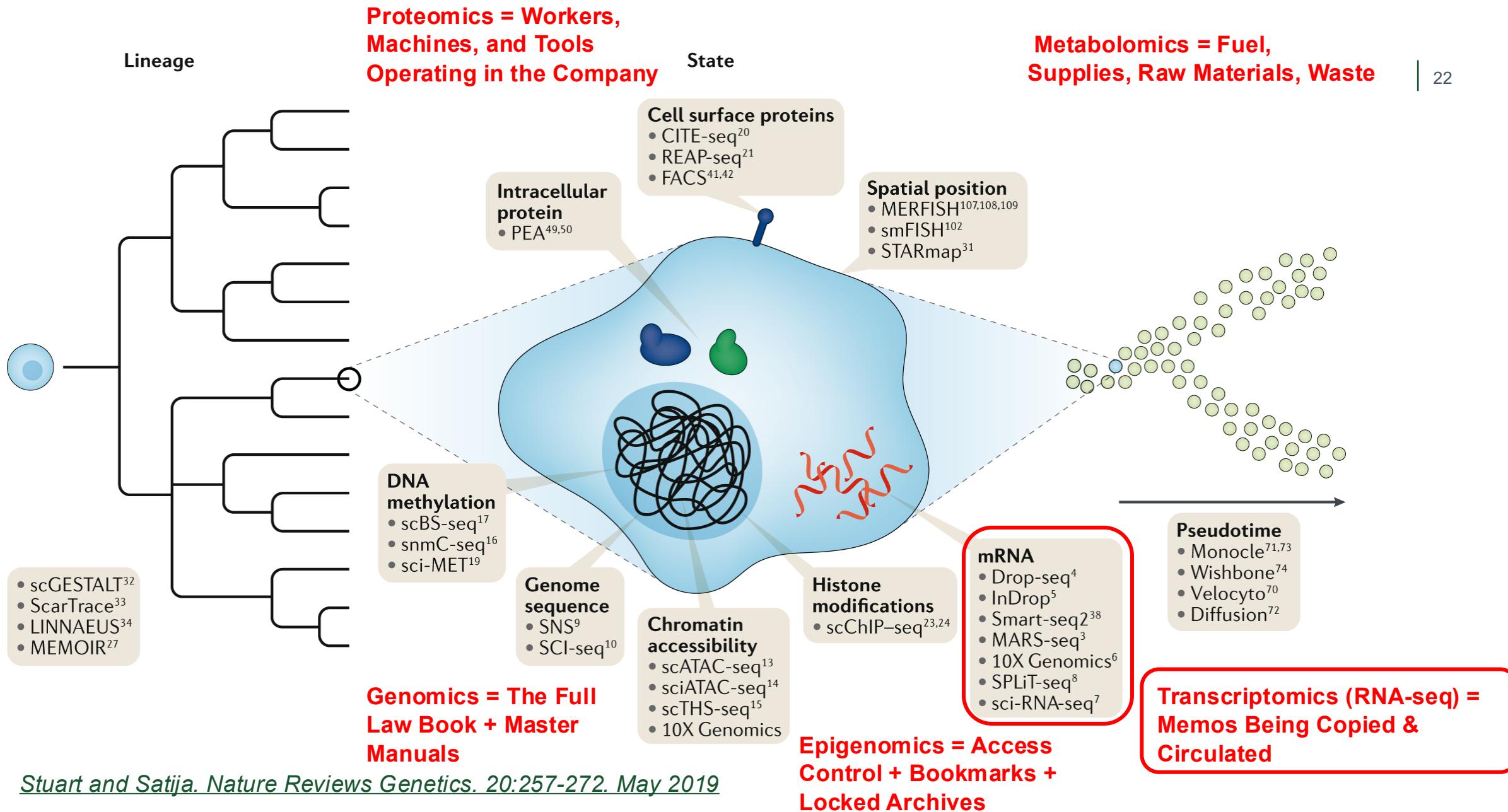
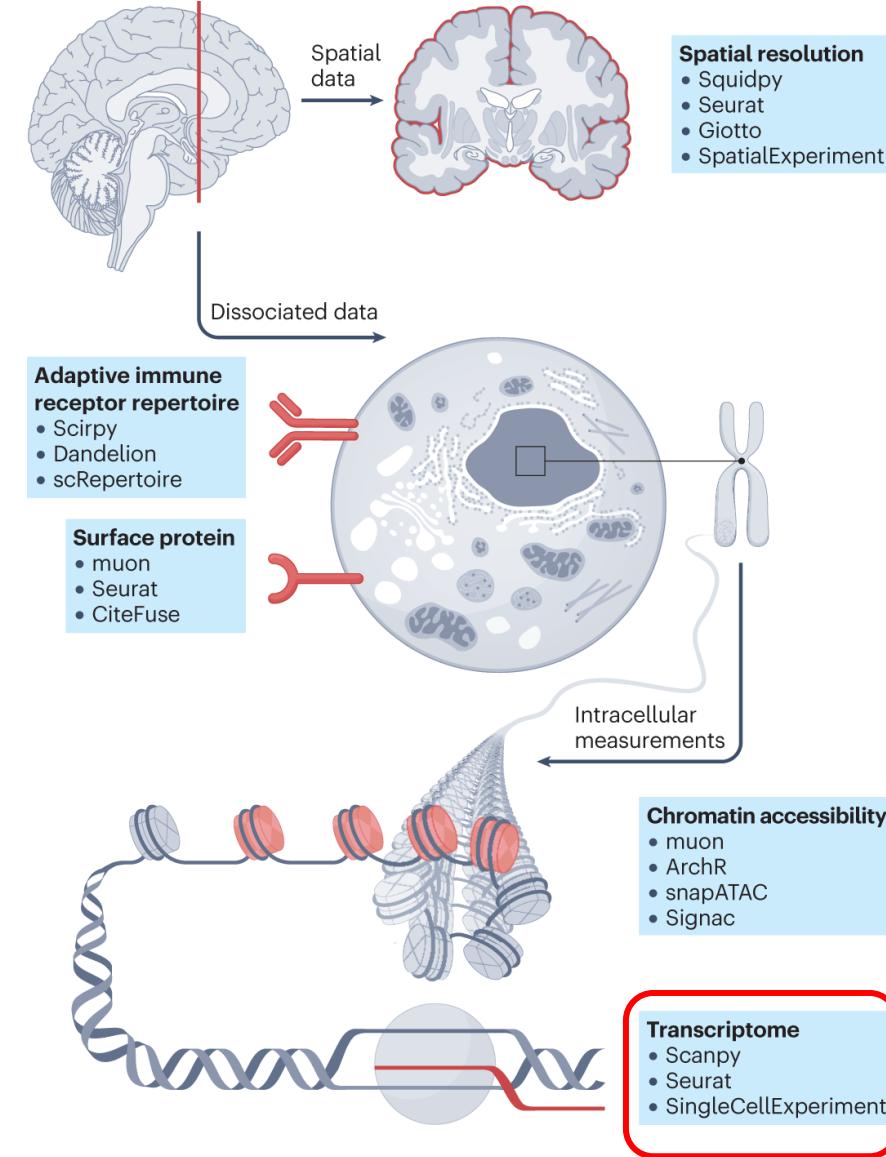


Fig1. Single-cell analysis across modalities



nf-core/scdownstream

24

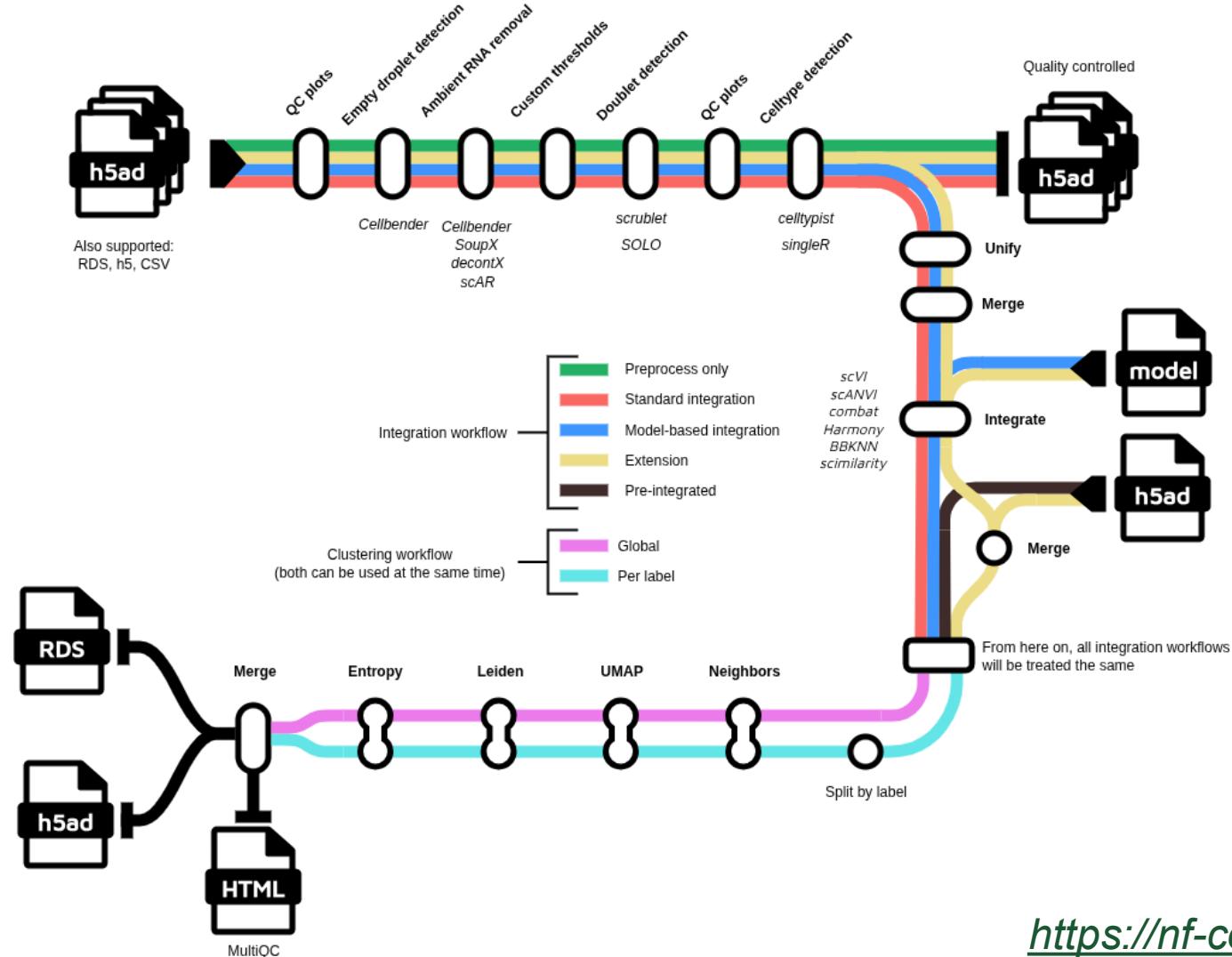


Fig2. Overview of unimodal analysis steps for scRNA-seq

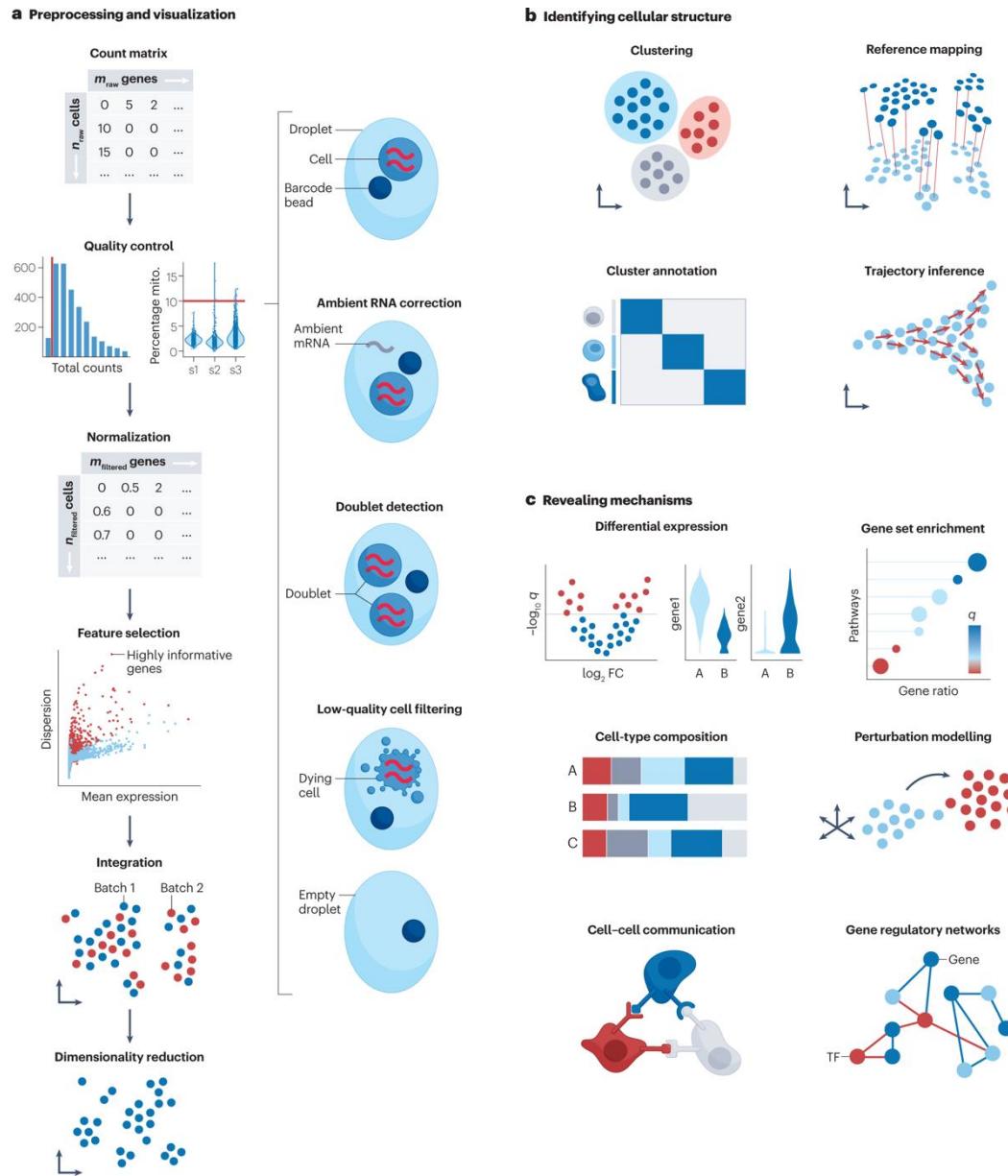


Fig. 6.1 low-quality cells, cell-free RNA and doublets (Book)

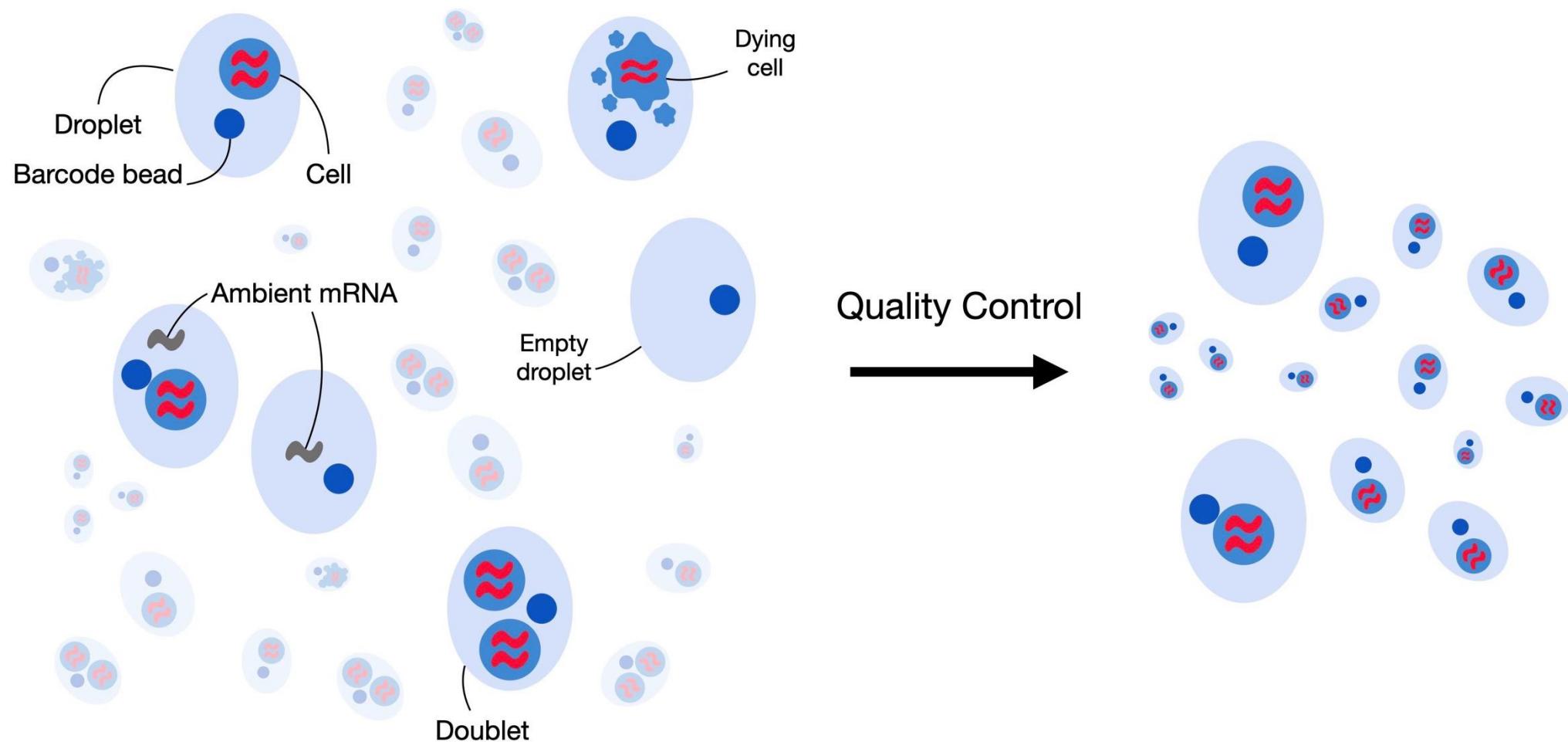
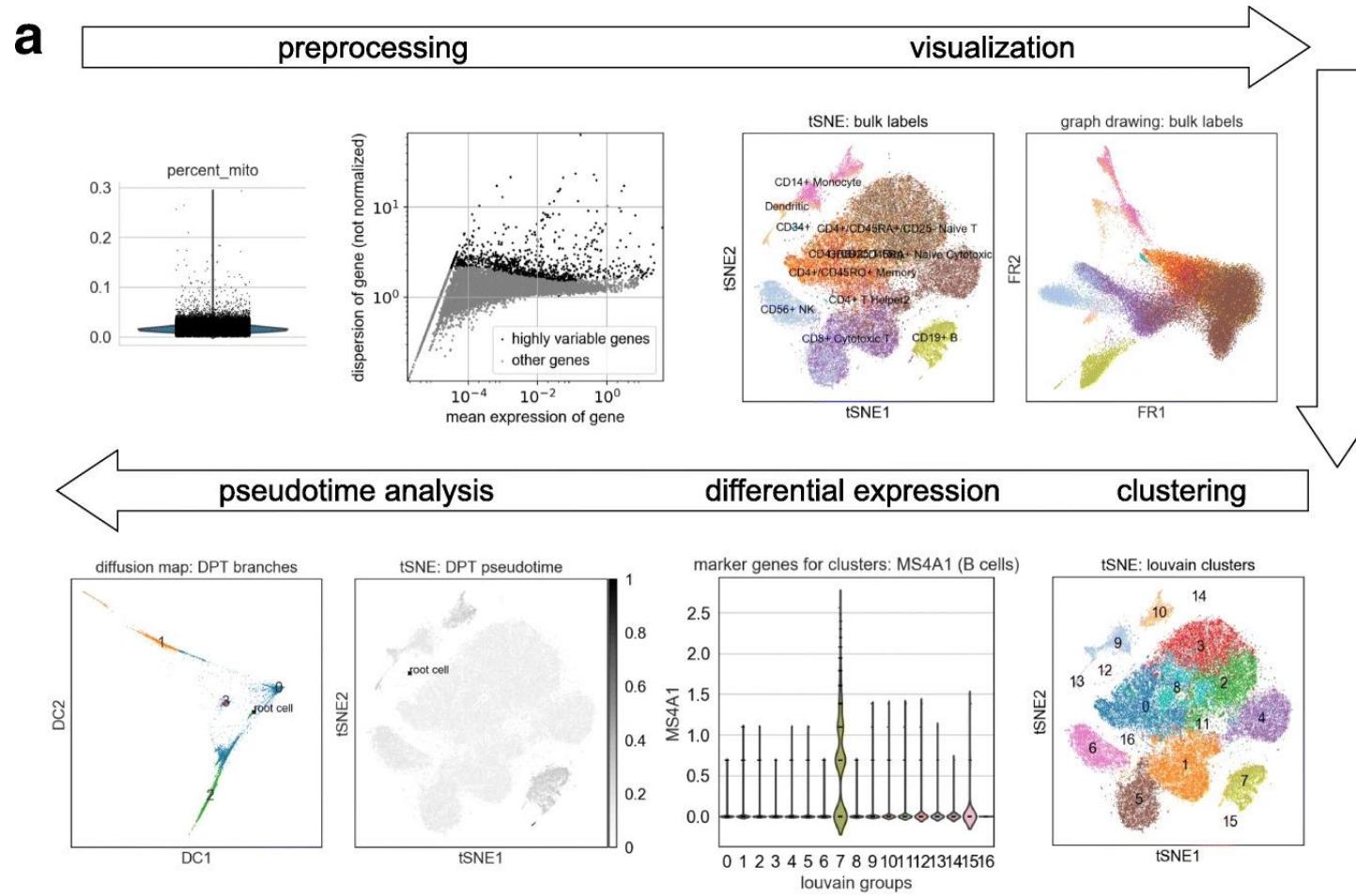
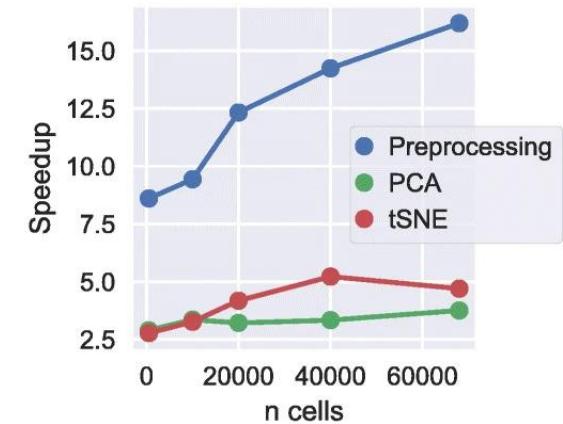


Fig. 4.2 scanpy overview (Book)



b Speedup: Scanpy vs. Cell Ranger R



c tSNE of clustered 1.3 million cells

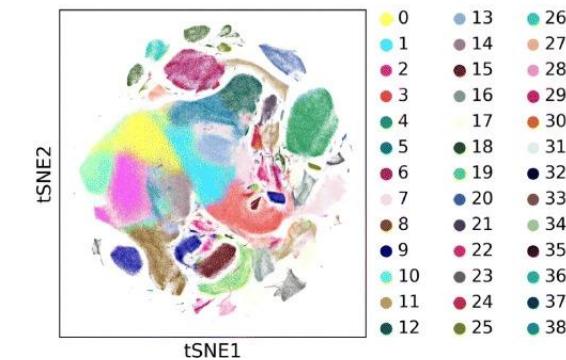
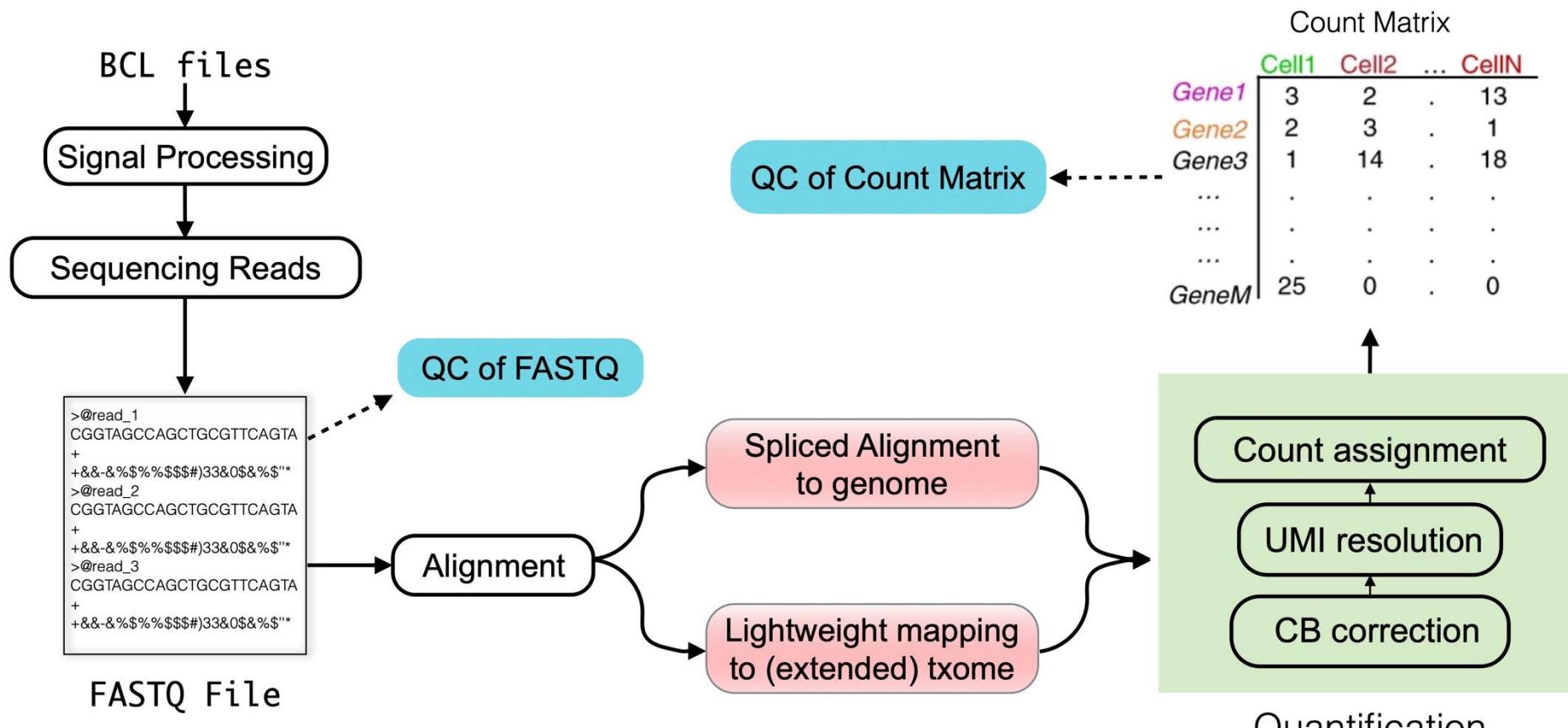


Fig. 3.1 (Book)



“txome” stands for transcriptome

Fig. 3: Overview of scATAC-seq analysis steps

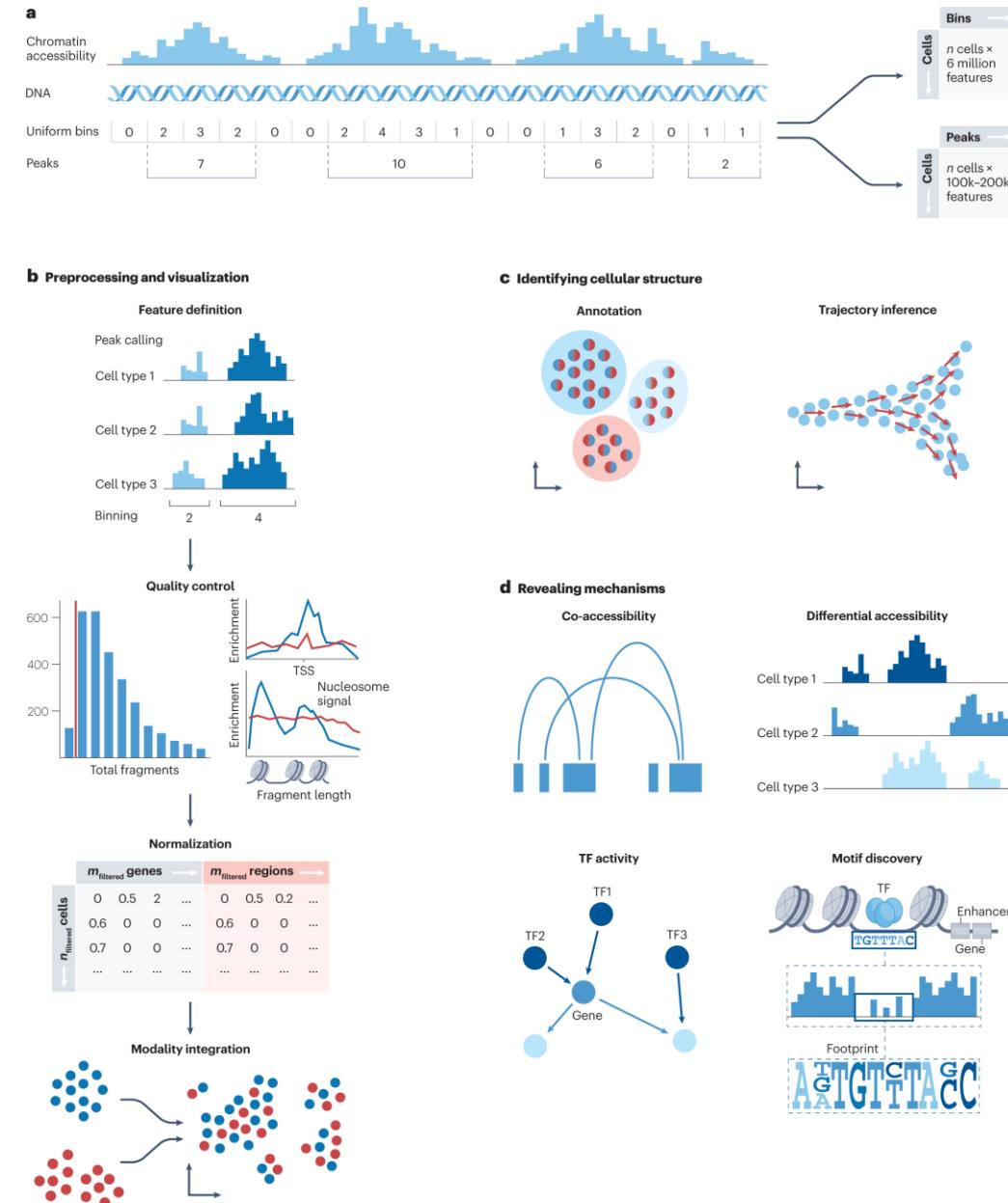


Fig. 4: Overview of CITE-seq data processing

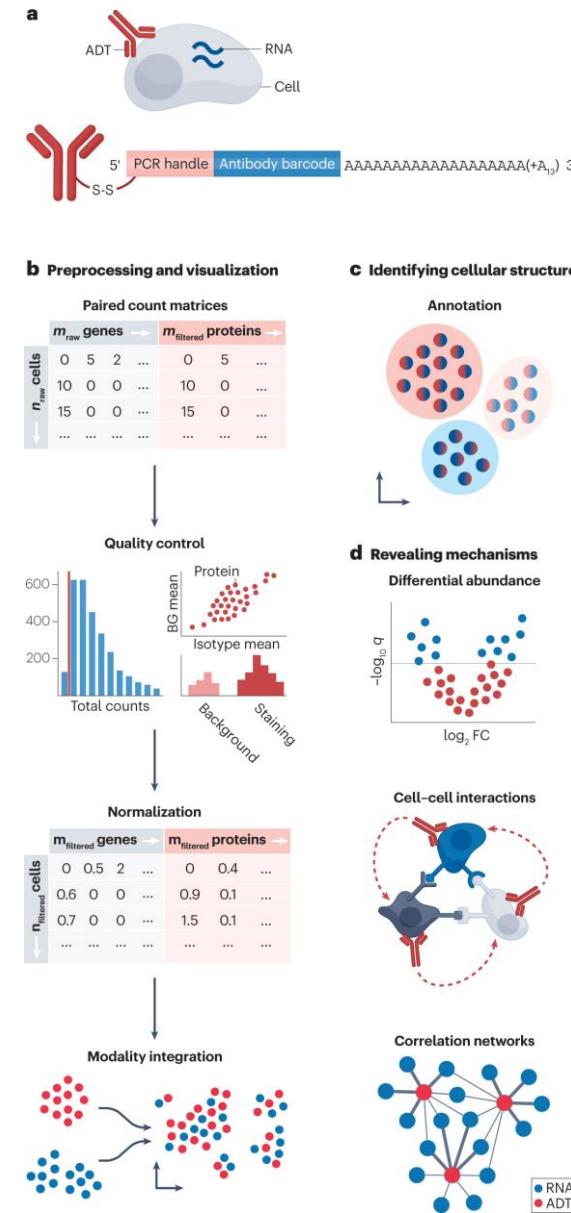


Fig. 6: Overview of spatial transcriptomics preprocessing and downstream analysis steps

Dissociated
vs. Spatial
Single-Cell

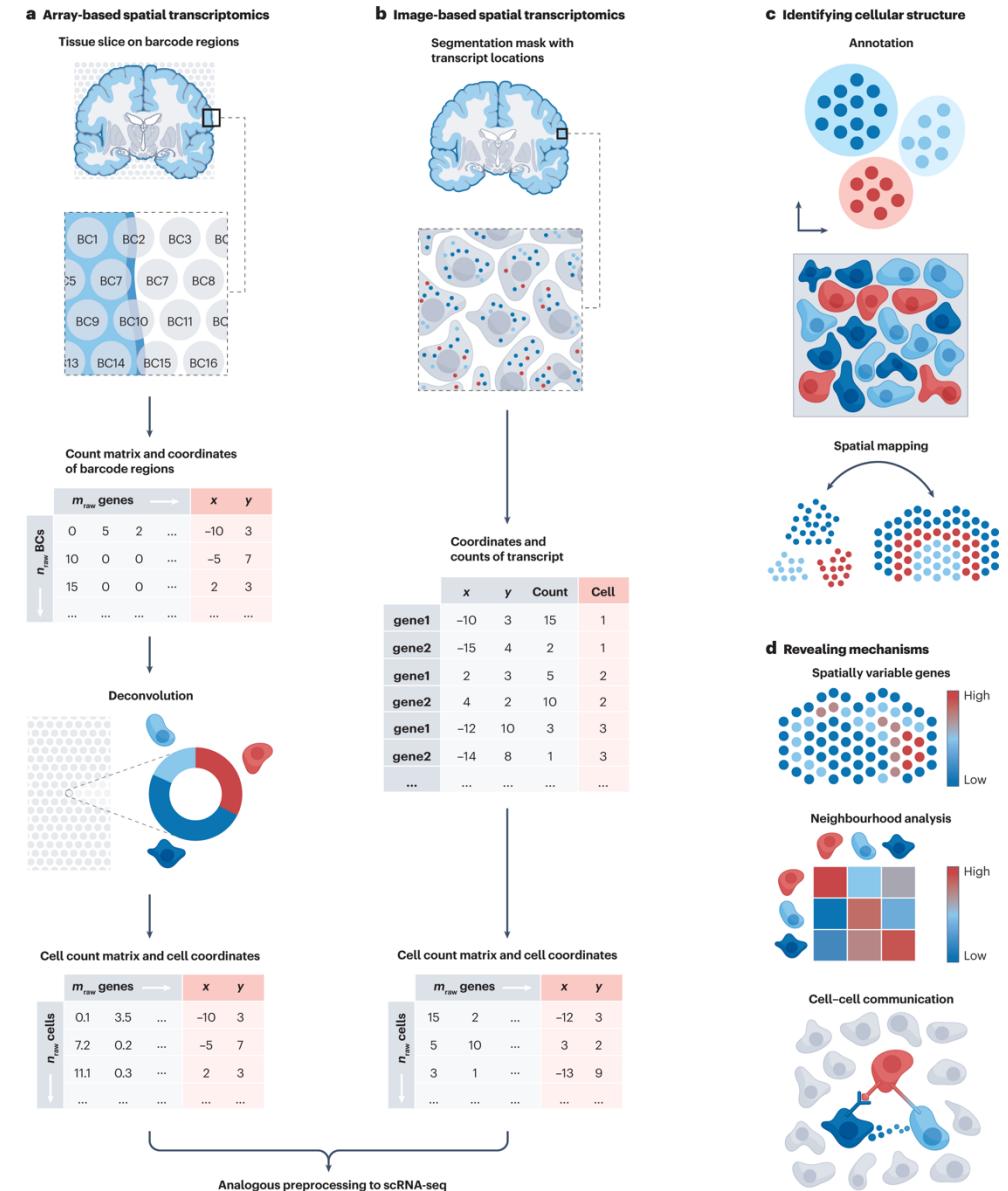
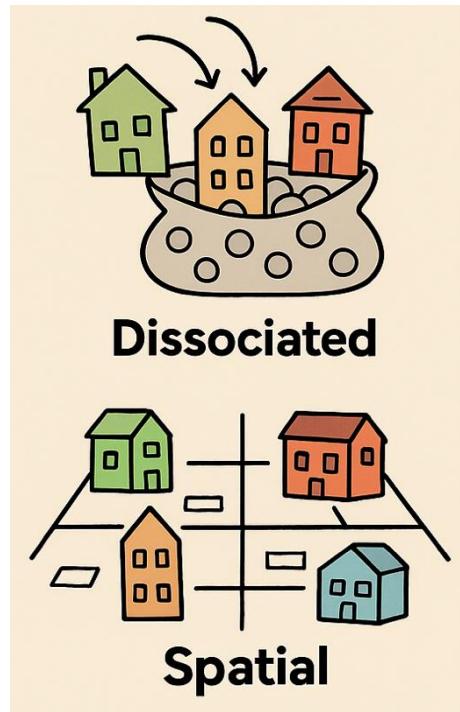
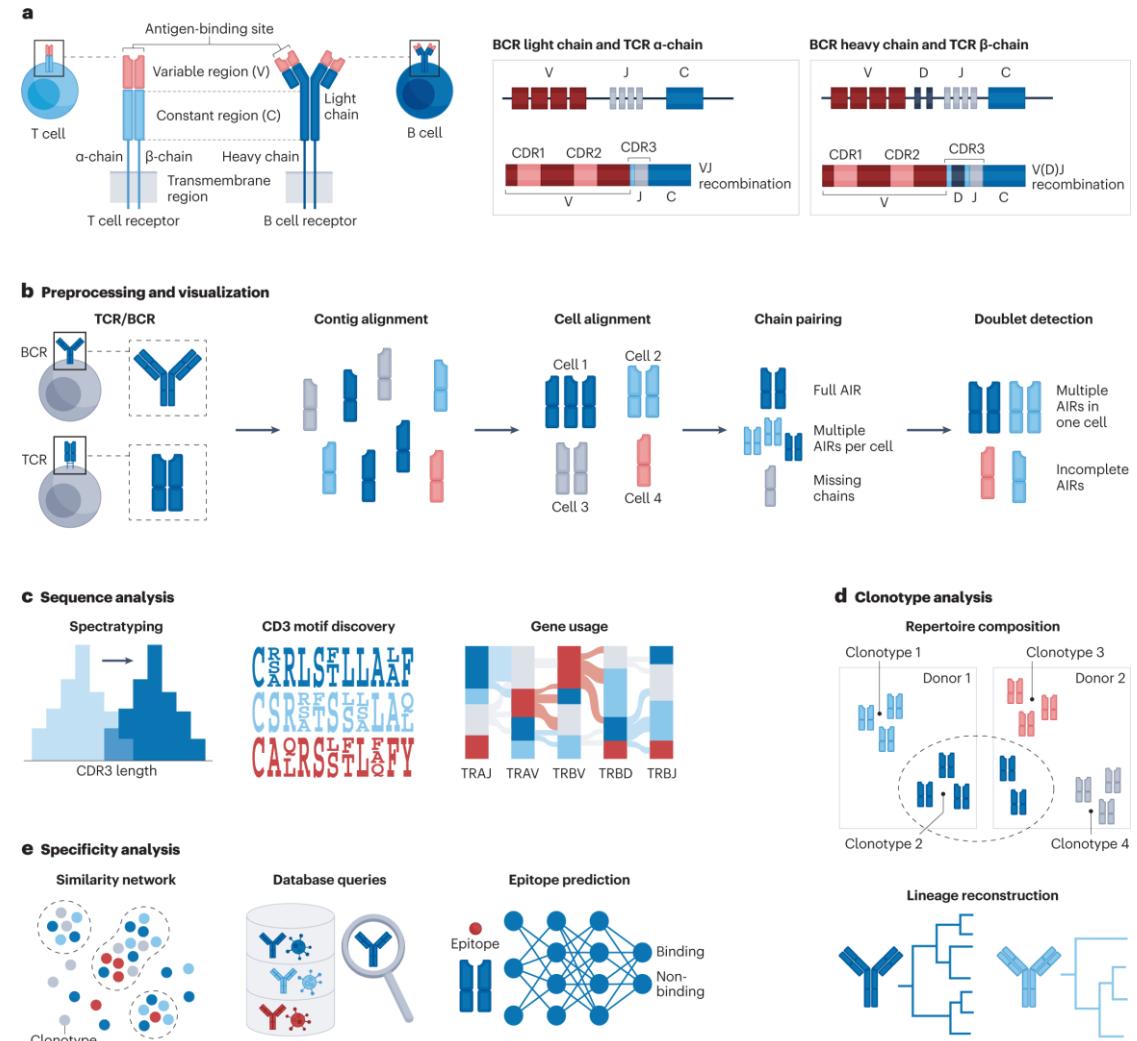


Fig. 5: Overview of the adaptive immune receptor analysis



scTalk1

nature reviews genetics

Explore content ▾ About the journal ▾ Publish with us ▾

[nature](#) > [nature reviews genetics](#) > [expert recommendation](#) > [article](#)

Expert Recommendation | Published: 31 March 2023

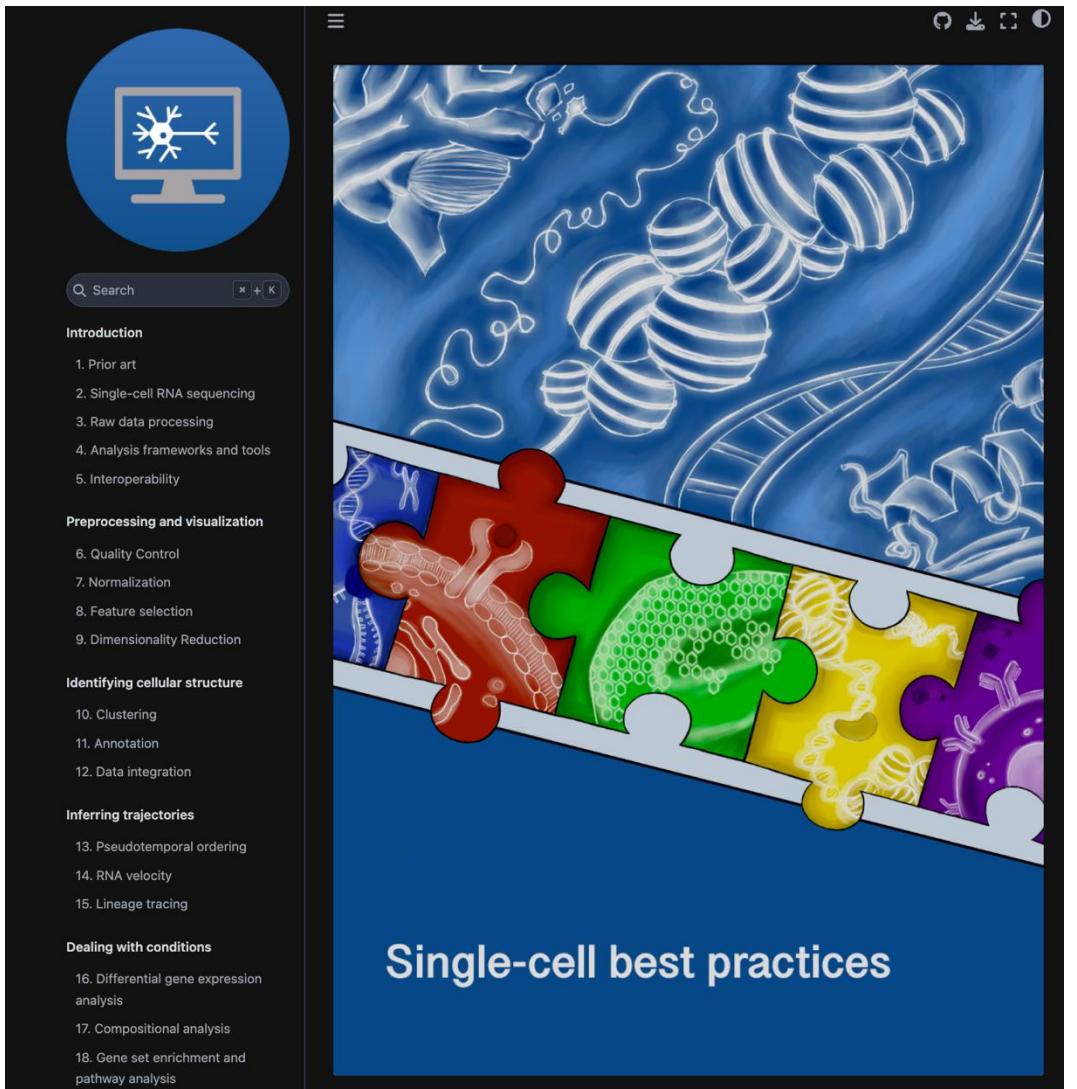
Best practices for single-cell analysis across modalities

Lukas Heumos, Anna C. Schaar, Christopher Lance, Anastasia Litinetskaya, Felix Drost, Luke Zappia, Malte D. Lücke, Daniel C. Strobl, Juan Henao, Fabiola Curion, Single-cell Best Practices Consortium, Herbert B. Schiller & Fabian J. Theis 

[Nature Reviews Genetics](#) 24, 550–572 (2023) | [Cite this article](#)

272k Accesses | 808 Citations | 358 Altmetric | [Metrics](#)

<https://www.nature.com/articles/s41576-023-00586-w>



nature biotechnology

Defining and benchmarking
open problems in single-cell
analysis.

[openproblems.bio](#)

Explore content ▾ About the journal ▾ Publish with us ▾

[nature](#) > [nature biotechnology](#) > [correspondence](#) > [article](#)

Correspondence Published: 01 July 2025

Defining and benchmarking open problems in single-cell analysis

[Malte D. Luecken](#), [Scott Gigante](#), [Daniel B. Burkhardt](#), [Robrecht Cannoodt](#), [Daniel C. Strobl](#), [Nikolay S. Markov](#), [Luke Zappia](#), [Giovanni Palla](#), [Wesley Lewis](#), [Daniel Dimitrov](#), [Michael E. Vinyard](#), [D. S. Magruder](#), [Michaela F. Mueller](#), [Alma Andersson](#), [Emma Dann](#), [Qian Qin](#), [Dominik J. Otto](#), [Michał Klein](#), [Olga Borisovna Botvinnik](#), [Louise Deconinck](#), [Kai Waldrant](#), [Sai Nirmayi Yasa](#), [Artur Szałata](#), [Andrew Benz](#), [Zhijian Li](#), [Open Problems Jamboree Members](#), [Jonathan M. Bloom](#), [Angela Oliveira Pisco](#), [Julio Saez-Rodriguez](#), [Drausin Wulsin](#), [Luca Pinello](#), [Yvan Saeys](#), [Fabian J. Theis](#)✉ & [Smita Krishnaswamy](#)✉

— Show fewer authors

Helmholtz Munich & Yale

[Nature Biotechnology](#) 43, 1035–1040 (2025) | [Cite this article](#)

11k Accesses | 13 Citations | 121 Altmetric | [Metrics](#)

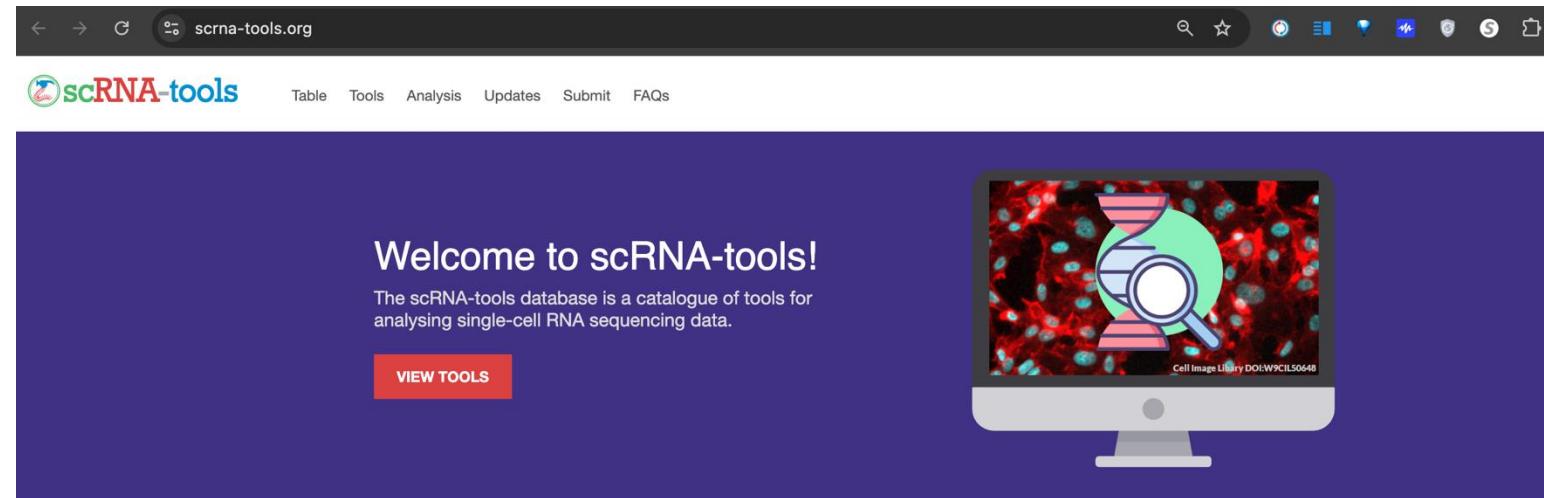
Why we need uniform, live, independent, and standardized benchmarks?

| 35

- Bespoke benchmarks set up by method developers to evaluate newly developed algorithms often include **datasets and metrics chosen to highlight the advantages of their tools**, which has been shown to lead to **less objective assessments^{6,7}**. Even if datasets and metrics are standardized, historical analysis shows that when benchmarks are implemented by the same groups introducing new methods, **the evaluations tend to inflate performance of the newest models via custom hyperparameter selection and data processing⁸**.
- ... efforts aim to systematically evaluate the current state of the art in a given area and may be less biased. However, their results are **static and inevitably age**.
- ... datasets and metrics typically have **less than 10% overlap** between benchmarks¹³.

Over 1800 Published Algorithms for single-cell Analysis

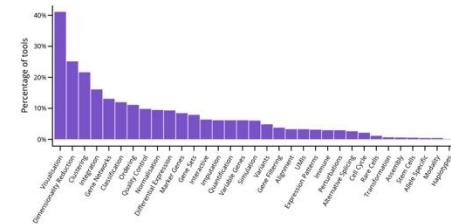
| 36



The scRNA-tools database records details of software tools designed for analysing scRNA-seq data. Each tool is categorised according to the analysis tasks it can be used for.



We currently track **1837** tools...



...in over **30** categories

<https://www.scRNA-tools.org>

“awesome-single-cell” GitHub collection as a common compilation of single-cell methods

37

<https://github.com/seandavi/awesome-single-cell>

<https://notebooklm.google.com/notebook/ab5c5e5c-6907-4883-8918-f4868f15ff88>

The screenshot shows the GitHub repository page for 'awesome-single-cell'. The repository is public, has 6 branches, and 6 tags. The README file is the active tab, displaying the title 'awesome-single-cell' and a brief description: 'List of software packages (and the people developing these methods) for single-cell data analysis, including RNA-seq, ATAC-seq, etc. Contributions welcome...'. Below the README are sections for Contributors (showing profile icons of several individuals) and Activity (showing metrics: 3.6k stars, 246 watching, 1.1k forks). The right sidebar contains an 'About' section with a community-curated list of software packages and data resources for single-cell analysis, including RNA-seq, ATAC-seq, etc., and a list of tags: python, bioinformatics, analysis, clustering, gene-expression, data-visualization, dimensionality-reduction, awesome-list, data-integration, atac-seq, single-cell, rna-seq-data, scRNA-seq-data, cell-cycle, cell-differentiation, gene-expression-profiles, analysis-pipeline, cell-populations, rna-seq-experiments, and cell-clusters.