Transcriptómica II 2025 single-cell RNA-seq

PRÁCTICO 4

Natalia Rego

nrego@fcien.edu.uy

bajar datos del DRIVE y arrancar script Colab

DRIVE

https://drive.google.com/drive/folders/1KMcmRoNGnpzze1k7AUEbg91xy_4tSck2?usp=sharing

DEJAR CORRIENDO INSTALACIÓN

Trajectory Analysis (TI)

- Instead of tracking a single cell over time, these methods study large amounts of single cells sampled from a dynamic cellular process, allowing their computational reconstruction
  - development
  - differentiation
  - immune responses
- These methods aim to infer a graph-like structure underlying the dynamic process from which the cells are sampled
- Then, by mapping the cells to this inferred structure, the cell properties can be compared over pseudotime, an abstract unit of progress through the dynamic process
- TI thus allows studying how cells evolve from one cell state to another, and subsequently when and how cell fate decisions are made

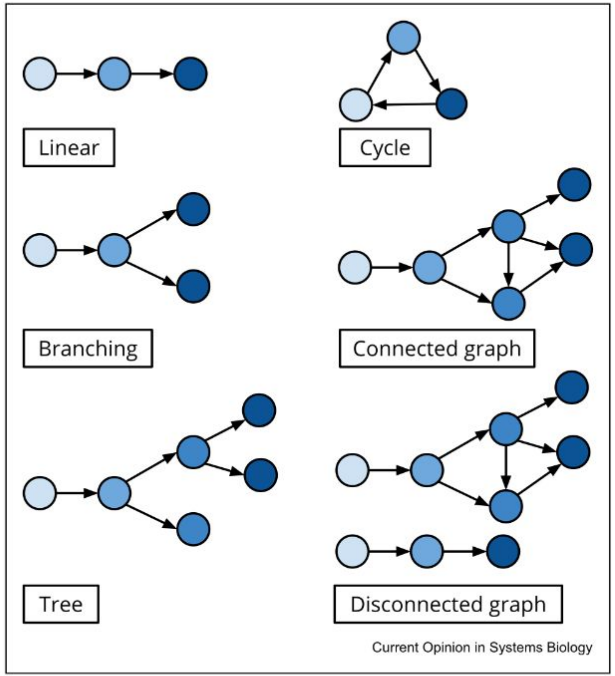
Assumptions for TI

- The biological process of interest is dynamic, and the appropriate cells are sampled
- The biological data are sampled to sufficient depth, so that the entire developmental process, including very transient states is presented
- The changes in gene expression are gradual during the developmental process.

Further assumptions are specific to the methods: e.g. trajectory type (linear or branching); additional prior knowledge needed (e.g. starting point cells)

TI is generally conducted after carefully analysis and annotation of the data: many tools assume that all cells in the dataset belong to the trajectory (and can't deal with noisy and/or outlier cells)

Figure 1

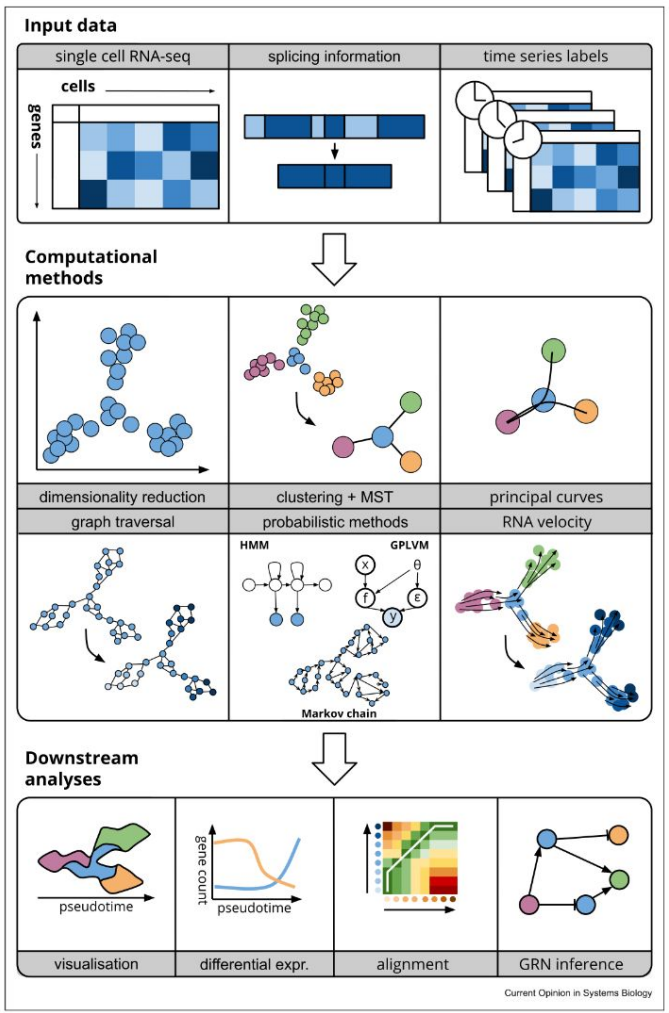


The most common trajectory topologies vary from simple linear and branching, to tree like and cyclical ones. Graph-like structures are a possibility as well. Not all TI methods can correctly infer all possible topologies.

Some methods able to extend the modeling of trajectory topologies toward general graphs and allow the inclusion of loops or even multiple separated trajectories.

As often the true underlying process is unknown, current best practices include comparing multiple methods to confirm the general topology structure.

To that end, methods that predetermine a certain topology (linear, bifurcating) can be used in combination with ones that do not.

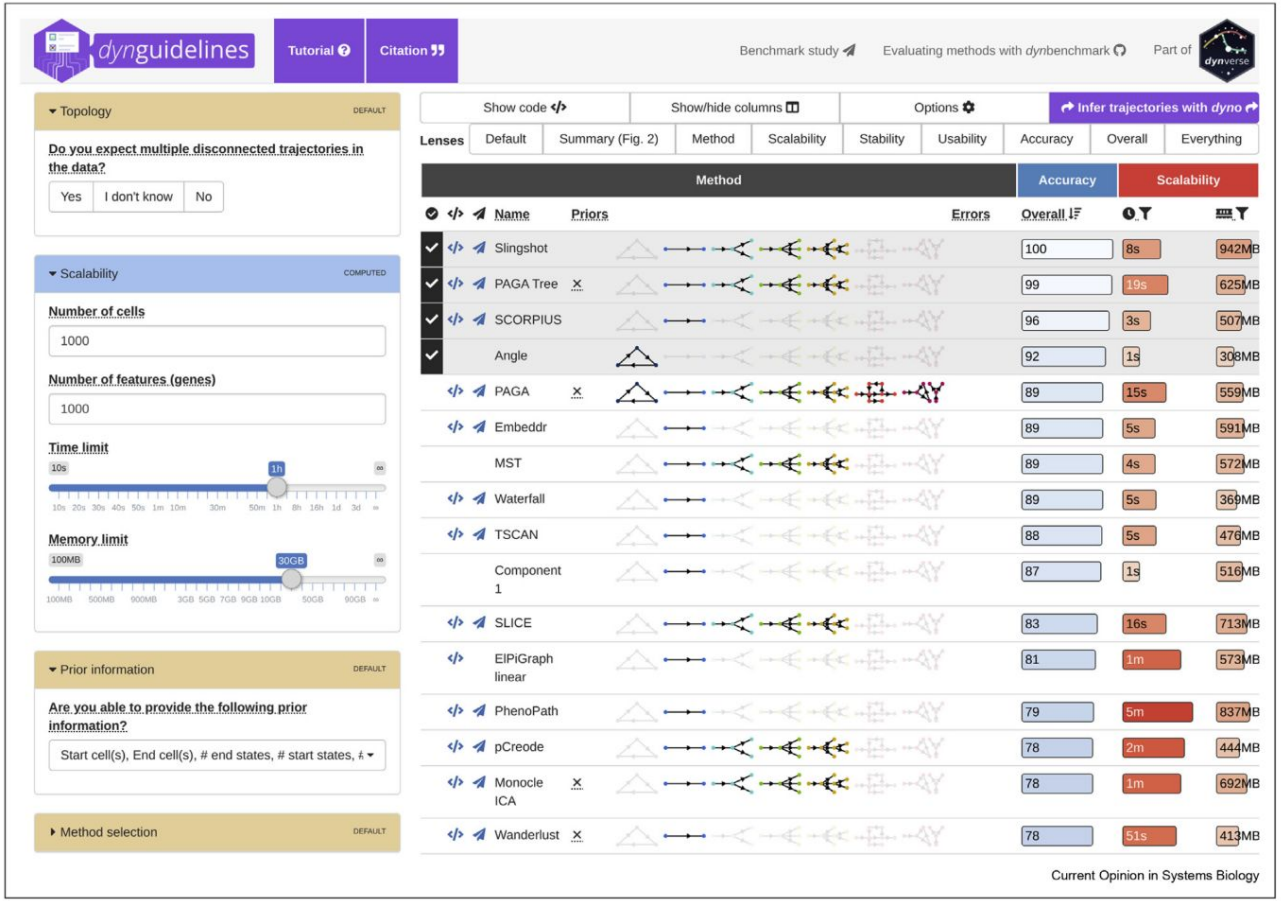


Input data can consist of a single-cell RNA-seq data set, splicing information or a time-series experiment comprising of multiple single-cell RNA-seq data sets.

The computational methods that actually perform the trajectory inference always include a dimensionality reduction. The following steps differ: clustering and MST construction, possibly in combination with principal curves is an option. Another approach is graph traversal on the individual cells. Probabilistic methods use a variety of techniques such as a Hidden Markov Model, a Markov chain, or a Gaussian Process Latent Variable Model. RNA velocity information can also be added to improve the TI.

Downstream analyses consist of visualization, trajectory-based differential expression, alignment between multiple data sets, or gene regulatory network inference.

Figure 2



An interactive tool that helps the user decide which TI method suits a particular use case best.

3 Teóricos, 4 Prácticos, 2 jornadas de seminarios, 3 hs c/u

TRANSCRIPTÓMICA II, SINGLE-CELL RNA-seq código B0058				
TEÓRICO 03/11 - 06/11/25	lunes	16:00 - 19:00	salón 201/203	
	miércoles	16:00 - 19:00	salón 209	
	jueves	16:00 - 19:00	salón 209	
PRÁCTICO 10/11 - 20/11	lunes	16:00 - 19:00	salón 107	
	miércoles	16:00 - 19:00	salón 109	
	jueves	16:00 - 19:00	salón 107	