Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Systems Biology

# Recent advances in trajectory inference from single-cell omics data

Louise Deconinck[1,2], Robrecht Cannoodt[1,2,3], Wouter Saelens[4,5], Bart Deplancke[4,5] and Yvan Saeys[1,2]

**Abstract**

Trajectory inference methods have emerged as a novel class of single-cell bioinformatics tools to study cellular dynamics at unprecedented resolution. Initial development focused on adapting methods based on clustering or graph traversal, but recent advances extend the field in different directions. A first class of methods includes novel probabilistic methods that report uncertainties about their outputs, and new methods that consider complementary knowledge, such as unspliced mRNA, time point information, or other types of omics data to construct the trajectory. A second class of methods uses the obtained trajectories as a starting point for novel analyses, such as visualization approaches, new types of statistical analyses, and the possibility to render static analyses more dynamic, such as dynamic gene regulatory network inference.

**Addresses**

[1]Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Ghent, Belgium
[2]Data Mining and Modelling for Biomedicine, VIB Center for Inflammation Research, Ghent, Belgium
[3]Data Intuitive, Lebbeke, Belgium
[4]School of Life Sciences, École Polytechnique Fédérale de Lausanne, Lausanne, Vaud 1015, Switzerland
[5]Swiss Institute of Bioinformatics, Lausanne, Vaud, Switzerland

Corresponding author: Saeys, Yvan (yvan.saeys@ugent.be)

**Current Opinion in Systems Biology** 2021, **27:**100344

This review comes from a themed issue on **Theoretical approaches to analyze single-cell data (2021)**

Edited by **Olivier Gandrillon** and **Michael Stumpf**

For complete overview of the section, please refer the article collection - Theoretical approaches to analyze single-cell data (2021)

Available online 24 May 2021

**https://doi.org/10.1016/j.coisb.2021.05.005**

2452-3100/© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

**Keywords**

Single-cell omics, Trajectory inference, Cell developmental modeling.

## Introduction

Single-cell technologies have emerged as the next-generation microscopes, allowing to characterize tissues and organisms in ever greater detail. While tracking the transcriptomic profile of a single cell over time is challenging, studying large amounts of single cells sampled from a dynamic cellular process allows computationally reconstructing cell developmental processes. Trajectory inference (TI) methods have emerged as a novel subfield within computational biology to better study the underlying dynamics of a biological process of interest, such as cellular development, differentiation, or immune responses [1]. These methods aim to infer a graph-like structure underlying the dynamic process from which the cells are sampled. By mapping the cells to this inferred structure, their properties can be compared over pseudotime, an abstract unit of progress through the dynamic process [2]. TI thus allows studying how cells evolve from one cell state to another, and subsequently when and how cell fate decisions are made. While TI methods can be in principle applied to any dynamical process that is sampled, the bulk of the methods have been developed specifically for single-cell transcriptomics data. However, as this field is evolving very rapidly, so is the field of TI method development, adapting to a wide variety of novel technological possibilities.

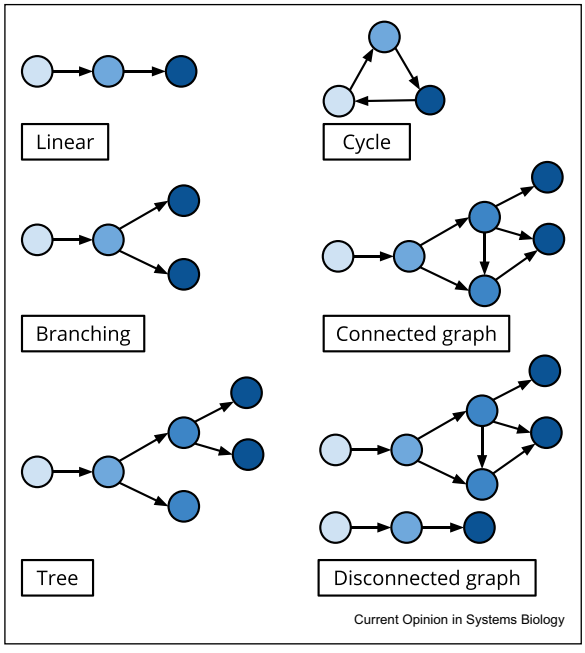
To study cell developmental dynamics using TI, several general as well as method-specific assumptions should be checked. General assumptions include that (1) the biological process of interest is dynamic, and the appropriate cells are sampled; (2) the biological data are sampled to sufficient depth, so that the entire developmental process, including very transient states is presented; and (3) the changes in gene expression are gradual during the developmental process. Further assumptions are specific to the methods used, such as confinement to particular trajectory types (e.g. linear or branching) or additional prior knowledge needed (e.g. a starting cell that is representative of the starting state of the dynamic process). In addition, many tools implicitly assume that all cells in the data set belong to the trajectory, calling for rigorous preprocessing to make sure noise or outlier cells that should not be used to construct the trajectory are removed. TI is generally performed after clustering and annotation of the data. An overview of general single-cell analyses, and where TI methods fit in, can be found in Luecken and Theis, who compiled a

comprehensive overview of the single-cell analysis pipeline [3].

## Trajectory topologies

As different TI methods make different assumptions about the data, a first choice to make is based on which biological process is to be expected: not all TI methods are designed to infer all kinds of biological processes. To categorize the TI methods with regard to this ability, we use the model proposed by Saelens et al. [4]. A trajectory is modeled by a graph-based network topology, consisting of connected milestones and cells placed on the connections between these milestones. A graphical overview of the different types can be found in Figure 1. The simplest processes can be represented using linear trajectories, for example, representing an ordering from immature cells, through intermediary stages, finally ending up in a terminal state. Methods specific to linear trajectories include the pioneering method Wanderlust [5], as well as more recent methods such as MATCHER [6] and SCORPIUS [7]. Similarly, cyclical trajectories, for example, modeling the cell cycle, can be inferred using tools such as ElPiGraph [8] and reCAT [9]. However, developmental processes are often not linear, and may include branching points, where cells make a commitment to follow one of a few possible routes. Slingshot [10] and Monocle [2] allow the modeling of these kinds of processes as bifurcating or tree-shaped topologies. Finally, methods such as PAGA [11], RaceID/StemID [12], and TinGa [13] are among the 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 [4]. To that end, methods that predetermine a certain topology (linear, bifurcating) can be used in combination with ones that do not. For more information about which particular TI method will suit a certain use case best we refer to Saelens et al. (Saelens et al., 2019), who give detailed information about both the accuracy, memory requirements, runtime, and prior information needed for a wide variety of methods and bundled these guidelines in an interactive tool (dynguidelines; URL: https://github.com/dynverse/dynguidelines; Figure 2).

**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.

## Computational approaches for trajectory inference

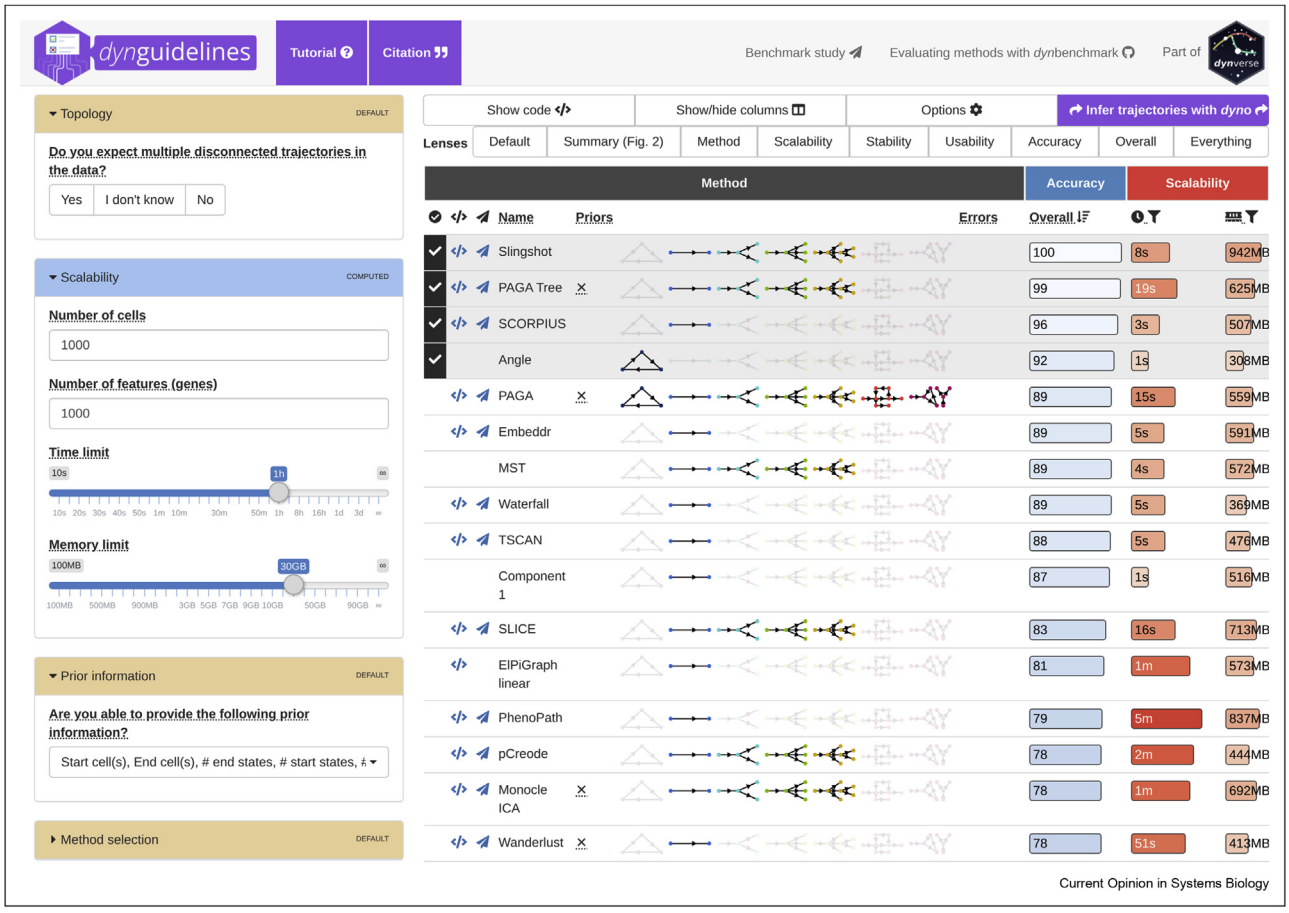
To approach the problem of trajectory inference, a wide variety of computational approaches is used. A unifying framework to characterize TI methods has been proposed by Cannoodt et al. [1]. In a first step, most methods apply some sort of dimensionality reduction method to reduce the high-dimensional gene space. Subsequently, several approaches can be considered to reconstruct the trajectory. These can be largely divided into clustering- and graph-based approaches. Clustering-based approaches first identify cell states (clusters) and subsequently connect these clusters into trajectory structures. Graph-based approaches construct a similarity graph, and subsequently apply graph-based methods to find connected components and organize them into trajectories. Many methods also combine elements from both approaches. An overview of these methods and subsequent analysis can be found in Figure 3.

### Current approaches

The most common types of TI methods start by performing a dimensionality reduction, a necessary step due to the inherent high dimensionality of the gene space. PCA is commonly used [14,11,15], but diffusion maps [16,17] and local linear embedding [18] are some of the other options. These methods try to represent the true dimensionality of the data, by attempting to preserve a combination of the local and global structure found in the data.

In this reduced space, the trajectory will be reconstructed. Clustering-based approaches first try to find stable cell states in the data and afterward connect these states to form a trajectory and map the cells to the formed graph structure. Many clustering approaches are used, such as soft K-means clustering [19], Louvain clustering [20], non-negative matrix factorization [18], or

**Figure 2**




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

hierarchical clustering [21]. In some cases, the method requires clusters as input and is agnostic to the actual clustering method used [22]. To connect the clusters to each other, usually a minimum weight spanning tree (MST) is constructed or edges are selected so that the most similar clusters are connected. MSTs are guaranteed to be unique if each edge in the graph has a distinct weight. Owing to the inherent heterogeneity of the cells in this kind of data, this is generally the case.

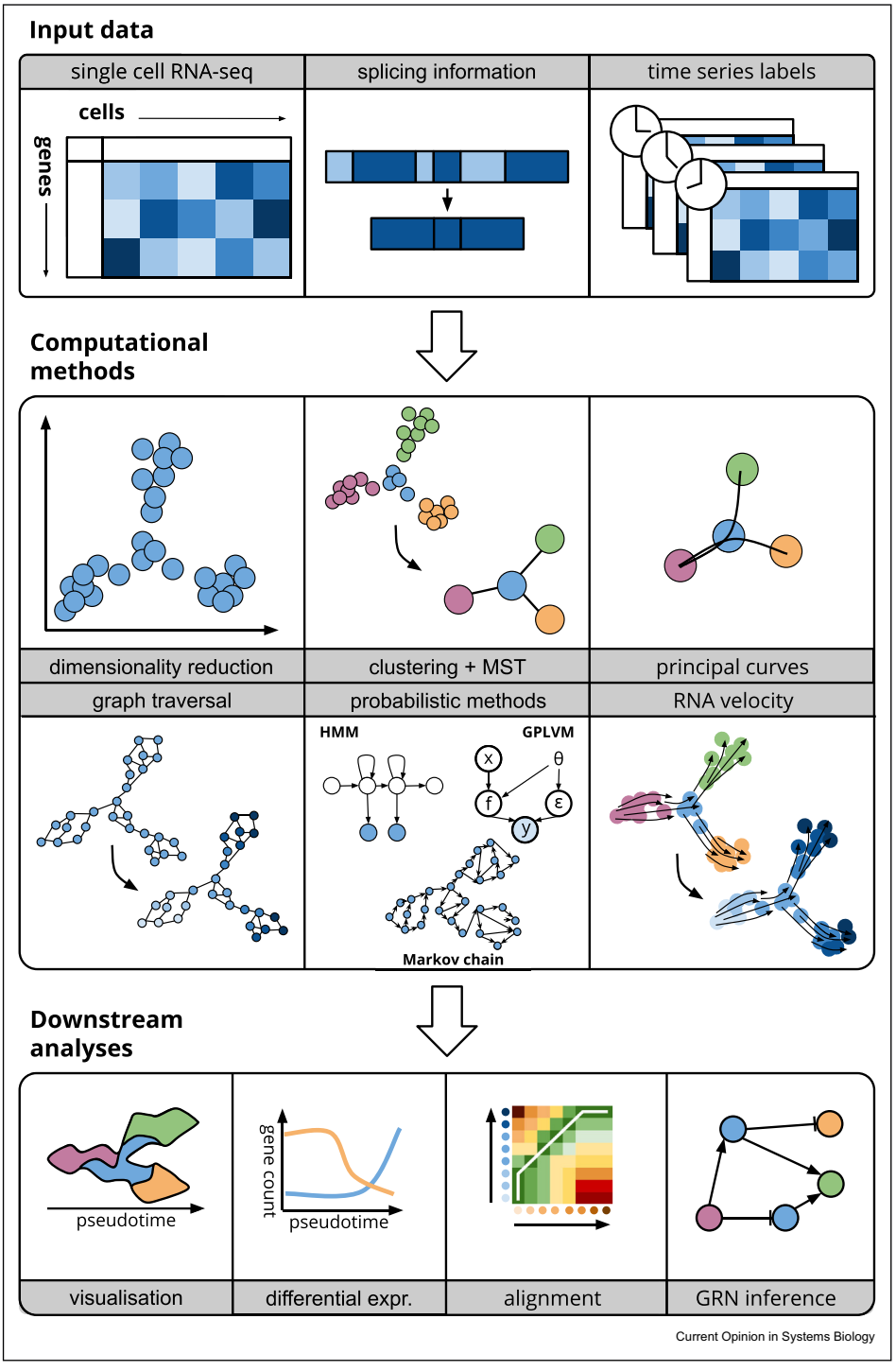
Graph-based approaches construct a graph representation of the cells and either use graph decomposition to reveal connected and disconnected components, or graph diffusion or traversal methods to construct the trajectory topology. Most methods start by building a k-nearest neighbors (kNN) graph using the Euclidean distance. PAGA [11] uses Louvain clustering to partition the graph and uses a statistical model to determine the strength between the different clusters to reveal connected and disconnected regions at the chosen resolution. DPT [23] calculates the probability of cells transitioning into each other using random walks from a user-provided root cell and uses the differences between these probabilities as pseudotime. Wanderlust [5] infers pseudotime for each cell based on the distance from a user-provided root cell.

Other methods use manifold learning-based approaches such as principal curves and graphs to infer the trajectory structure. In addition to the previous methods that use an MST to model the trajectory structure, SCORPIUS [7] and Slingshot [10] use principal curves to obtain smooth trajectories. They both still cluster and construct a trajectory by connecting the closest clusters or computing an MST as previous steps. Cells can then be conveniently projected onto the smoothed trajectory obtained by principal curves.

## Novel probabilistic approaches

Several methods that try to model the developmental process in a probabilistic way have been recently introduced. Instead of single point pseudotime estimates,

**Figure 3**




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.

they can also provide uncertainty about the pseudotime or uncertainty of a cell belonging to a certain branch of the trajectory. Some of these methods build further on certain elements used in early TI research, such as neighborhood graphs and diffusion components in addition to a probabilistic model.

Palantir [17] models the trajectory as a Markov process. First, a nearest neighbor graph is constructed using diffusion maps. Each cell's pseudotime is iteratively computed. First, the shortest path from the user-defined root cell is used. Then this value is refined using the distance from certain waypoints, sampled to encompass the whole differentiation trajectory. This neighborhood graph and the associated pseudotime values are then used to construct a Markov chain in which each node represents a cell. The pseudotime provides directionality between the edges connecting the nodes. This chain is used to calculate the transition probabilities of each cell to reach a neighboring cell in one step. After determining which states are terminal, a pseudotime value and a branch probability are assigned to each cell.

A different method, CSHMM [24], uses a continuous state Hidden Markov Model to infer a trajectory. This allows for an infinite number of states, allowing cells to be assigned to a fine-grained trajectory, based on both the inferred pseudotime and cell state. The method works well with the integration of time-series RNA-Seq data. GPseudoRank [25] uses Markov Chain Monte Carlo methods to sample from a posterior distribution of cell orderings. They do, at first, not infer a continuous pseudotime value for the cells, but they infer a discrete ordering of the cells. This avoids exploring pseudotime assignments that map to the same ordering. The continuous pseudotime values assigned to the cells are then derived using a deterministic transformation.

Ouija [26] uses a Bayesian latent variable model to learn pseudotime from a small set of known or suspected marker genes. It also explicitly models the expression peaks and the on/off switching of these genes during the developmental process. This allows determining easily which of these genes are involved in the regulation of certain parts of the trajectory.

Grandprix [27] uses the Gaussian Process Latent Variable Model (GPLVM) to estimate pseudotime values. The main contribution of the work is the implementation of a more efficient model, which allows the GPLVM to be used with larger data sets. As is often the case with probabilistic models, these are computationally intensive models to run, making them less scalable to large-scale datasets [4].

## Extensions of trajectory inference

### Integration of other data sources

To improve the biological relevance of the inferred trajectories, several methods use new sources of information instead of, or in addition to, the gene expression counts. By far the most popular source of information to add is RNA velocity [28]. RNA velocity methods estimate the future state of a single cell captured in a static snapshot by looking at the ratios between spliced mRNA, unspliced mRNA, and mRNA degradation. scVelo [29] and CellRank [30] use this extra velocity information to construct a directed k-nearest neighbor graph, as a starting step for the TI method. This has the advantage that root cell specification is not necessary, and adds directional information to the trajectory.

The dynamics of the cellular state is not only represented in the transcriptome but also in other modalities such as the epigenome and the proteome. Given that these modalities are often related, for example, with a time delay, it is of great interest to profile and infer trajectories from multiple modalities. MATCHER [6] is a TI method integrating both transcriptomic and epigenetic information, of different samples belonging to the same type, using manifold learning. Other methods that do not infer trajectories, but embed multimodal data in a common space are also of interest. UnionCom [31] constructs distance matrices of each modality and matches them before embedding in a common feature space. Seurat v4 [32] uses weighted nearest neighbor analysis and shows that the inclusion of protein information can improve trajectory characterization. TI methods specifically leveraging multimodal information can thus uncover a more accurate picture of the developmental process.

Some methods focus on combining data sets from time-series experiments. Multiple single-cell transcriptomics experiments are sequentially combined, so that a process can be studied over time. Tempora [33] constructs a network of cell clusters and uses ARACNE [34] to find vertices between clusters based on the pathway enrichment profiles of each cluster. To determine the directionality of these vertices, the time-series labels are used.

### Quality control and benchmarking

Owing to the scarcity of large benchmarking data sets during the advent of scRNA-seq technologies, early computational methods were developed in the context of a specific single-cell data set so that generalization beyond this data set was rarely assessed. The lack of consensus guidelines regarding how trajectory information should be stored and which metrics should be used to compare studies rendered the evaluation of TI methods particularly challenging.

A comprehensive benchmark of TI methods was performed in Ref. [4], using multiple metrics to not only assess the accuracy on multiple different simulated and real data sets but also the stability, scalability, and usability of each method. Several of the metrics proposed in this study could be used to quantitatively evaluate TI methods. Guidelines for selecting the most appropriate method for a particular data set were also provided. Novel single-cell simulation engines have been developed, allowing to generate ground truth models that can be used to benchmark TI methods [35–37].

## Downstream analysis of inferred trajectories

Inferring trajectories from single-cell omics data allows to extract dynamics, typically from a single snapshot of a biological system. However, the topological structure and its associated pseudotime offer an additional time dimension that allows a temporal mapping of every cell. This novel time component subsequently allows applying many proven techniques derived from time series analysis, as well as developing interesting novel approaches, as the resulting trajectory can be interpreted as a very granular time-course experiment.

### Visualization

Some methods provide a dedicated way to visualize the trajectory, but these are mostly very simple visualizations: the pseudotime values are added as a color scale onto a common visualization method such as t-SNE or UMAP, where sometimes nodes and edges are drawn as well to visualize the backbone of the trajectory. STREAM [38] implemented multiple visualizations for linear and branching trajectories that facilitate visualizing cell density, proportions of cell states, and expression of a single gene across the trajectory. dyno (dyno; URL: https://github.com/dynverse/dyno) presents an entire TI pipeline, including a variety of visualization options for any trajectory type.

### Trajectory differential expression

Inferring trajectories from single-cell data offers novel ways to extract developmental information, such as finding out which transcription factors drive and regulate the developmental process. It is possible to perform a differential expression analysis along these trajectories: some TI methods, such as Monocle, incorporate this in their workflows. Generic approaches have also been proposed that work with different TI methods in a common statistical framework: switchde [39] allows testing for differential expression along a single-cell trajectory, whereas tradeSeq [40] can test for differential expression within a particular branch of a trajectory, but also provides a framework to test differential expression of genes between different branches of a trajectory, which can be particularly useful to compare trajectories resulting from different conditions.

However, a concern with this kind of analysis is circularity, as the same data points and features are used to perform the TI and the differential expression analysis. The TI step enforces a certain optimized ordering upon the cells, potentially enhancing expression differences along trajectories, leading to artificially low p-values and an inflated number of false positives. This is an issue present in other steps of the single-cell analysis pipeline and has gained attention in the field of cluster-based differential expression [41]. One possible solution is to incorporate the uncertainty of the inferred trajectory within the differential expression to correct the p-values. Another possible solution is to use separate data sets for each, for example, by means of multimodal data such as CITESeq estimate the trajectory on gene expression data and test for differential protein expression.

### Trajectory alignment and comparison

It can be insightful to compare trajectories inferred from samples belonging to certain conditions, for example, healthy versus diseased patients. PhenoPath [42] uses a Bayesian statistical model capable of inferring different pseudo temporal trajectories for different patient subgroups. A different method to compare existing pseudo temporal trajectories of different patient groups, or even species, is Dynamic Time Warping [43,44]. This method aligns the trajectories to a similar timescale so that comparisons can be made easier. A dedicated method, cellAlign [45], was developed to compare linear non-branching single-cell trajectories.

### Dynamic gene regulatory network inference

The expression level of a single gene is determined by a complex regulatory network, detailing which gene and molecule interactions influence the expression level of that gene. Uncovering this regulatory network can be performed using gene regulatory network (GRN) inference methods [46]. Typically, GRN inference methods are applied on gene expression data to yield a static network. By combining GRN inference methods with the additional time dimension uncovered by TI methods dynamic gene regulatory networks can be reconstructed, where the transcription factor — target predictions are not static, but can evolve during the course of the trajectory [47–49]. However, pseudotime information might not work as well as true time-series data [50].

## A future outlook for trajectory inference methods

TI methods have emerged as a novel set of computational techniques to study cellular dynamical processes. By allowing modeling gradual transitions between cell states, these methods generalize the concept of clustering, resulting in a topological map of the dynamic process of interest, where cells are assigned to different

branches in this topology and mapped to their progress along this dynamic process. Although methods in the past mostly tried to accommodate increasing levels of complexity of the underlying topology, current techniques diversify the landscape of TI methods. Recently introduced methods focus more on providing uncertainty estimates about their predictions, and try to incorporate complementary information to construct the trajectory. This includes directional information, as, for example, obtained from RNA velocity estimates, prior information (e.g. known time point information) or multiple complementary data sources (e.g. single-cell multi-omics). As the field is moving toward measuring more complementary information at the single-cell level, a novel wave of multimodal TI methods can be anticipated. New benchmarking studies can be expected to help quantify the importance of these additional sources of information on the resulting trajectory. Similarly, the creation of an unsupervised metric measuring how well a trajectory fits the dataset used could serve as a quality control before performing downstream analyses based on these inferred trajectories. As different data sources might shed light on different dynamical aspects of a biological process, there is a high need for methods that are able to better extract multiple parallel aspects of a dynamic biological process (e.g. differentiation combined with cell cycle status and cell migration). Finally, it can be expected that many more downstream analysis methods will be developed, taking advantage of the time component returned by TI methods. The inferred trajectory can be regarded as a very granular time series, transforming many static gene expression-based tools into their dynamic counterpart, an example being dynamic gene regulatory network inference.

## Conflict of interest statement

Nothing declared.

## Acknowledgements

This project received funding from the Flemish Government (AI Research Program).

## References

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest
** of outstanding interest

1. Cannoodt R, Saelens W, Saeys Y: **Computational methods for trajectory inference from single-cell transcriptomics**. *Eur J Immunol* 2016, **46**:2496–2506.

2. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS, Rinn JL: **The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells**. *Nat Biotechnol* 2014, **32**: 381–386.

3. ** Luecken MD, Theis FJ: **Current best practices in single-cell RNA-seq analysis: a tutorial**. *Mol Syst Biol* 2019, **15**, e8746.
A comprehensive tutorial and an overview of the complete single-cell analysis process. It presents the current best practices in the field and how different tools are chained together to form the standard single-cell analysis pipeline.

4. ** Saelens W, Cannoodt R, Todorov H, Saeys Y: **A comparison of single-cell trajectory inference methods**. *Nat Biotechnol* 2019, **37**:547–554.
A comprehensive benchmark measures cell ordering, found topology, scalability and usability on 45 different trajectory inference tools using synthetic and real data. Provides a complete pipeline for the evaluation of trajectory inference methods and gives useful guidelines as to how to best use these them.

5. Bendall SC, Davis KL, Amir E-AD, Tadmor MD, Simonds EF, Chen TJ, Shenfeld DK, Nolan GP, Pe'er D: *Single-Cell trajectory detection uncovers progression and regulatory coordination in human B cell development*. 2014.

6. Welch JD, Hartemink AJ, Prins JF: **MATCHER: manifold alignment reveals correspondence between single cell transcriptome and epigenome dynamics**. *Genome Biol* 2017, **18**: 138.

7. R. Cannoodt, W. Saelens, D. Sichien, S. Tavernier, S. Janssens, M. Guilliams, B. Lambrecht, K. De Preter, Y. Saeys, SCORPIUS improves trajectory inference and identifies novel modules in dendritic cell development.

8. Albergante L, Mirkes E, Bac J, Chen H, Martin A, Faure L, Barillot E, Pinello L, Gorban A, Zinovyev A: *Robust and scalable learning of complex intrinsic dataset geometry via ElPiGraph*. 2020.

9. Ye Y, Gao L, Zhang S: **Circular trajectory reconstruction uncovers Cell-Cycle progression and regulatory dynamics from Single-Cell Hi-C maps**. *Adv Sci* 2019, **6**:1900986.

10. * Street K, Risso D, Fletcher RB, Das D, Ngai J, Yosef N, Purdom E, Dudoit S: **Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics**. *BMC Genom* 2018, **19**:477.
Slingshot is a widely used trajectory inference method capable of inferring multiple branching lineages.They developed simultaneous principal curves to better estimate the pseudotime values for each lineage.

11. * Wolf FA, Hamey FK, Plass M, Solana J, Dahlin JS, Göttgens B, Rajewsky N, Simon L, Theis FJ: **PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells**. *Genome Biol* 2019, **20**: 59.
PAGA is a widely used trajectory inference method allowing to analyse data at multiple resolutions and developed a statistical model measuring the connectivity of groups of cells, thereby denoising typical k-nearest neighbor graphs.

12. Grün D, Muraro MJ, Boisset J-C, Wiebrands K, Lyubimova A, Dharmadhikari G, van den Born M, van Es J, Jansen E, Clevers H, de Koning EJP, van Oudenaarden A: **De novo prediction of stem cell identity using Single-Cell transcriptome data**. *Cell Stem Cell* 2016, **19**:266–277.

13. Todorov H, Cannoodt R, Saelens W, Saeys Y: **TinGa: fast and flexible trajectory inference with growing neural gas**. *Bioinformatics* 2020, **36**(Supplement_1):i66–i74.

14. Klimovskaia A, Lopez-Paz D, Bottou L, Nickel M: **Poincaré maps for analyzing complex hierarchies in single-cell data**. *Nat Commun* 2020, **11**:2966.

15. Ko ME, Williams CM, Fread KI, Goggin SM, Rustagi RS, Fragiadakis GK, Nolan GP, Zunder ER: **FLOW-MAP: a graph-based, force-directed layout algorithm for trajectory mapping in single-cell time course datasets**. *Nat Protoc* 2020, **15**: 398–420.

16. Setty M, Tadmor MD, Reich-Zeliger S, Angel O, Salame TM, Kathail P, Choi K, Bendall S, Friedman N, Pe'er D: **Wishbone identifies bifurcating developmental trajectories from single-cell data**. *Nat Biotechnol* 2016, **34**:637–645.

17. Setty M, Kiseliovas V, Levine J, Gayoso A, Mazutis L, Pe'er D: **Characterization of cell fate probabilities in single-cell data with palantir**. *Nat Biotechnol* 2019, **37**:451–460.

18. Wang S, Karikomi M, MacLean AL, Nie Q: **Cell lineage and communication network inference via optimization for single-cell transcriptomics**. *Nucleic Acids Res* 2019, **47**:e66.

19. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, Trapnell C: **Reversed graph embedding resolves complex single-cell trajectories**. *Nat Methods* 2017, **14**:979–982.

20. Hie B, Cho H, Bryson B, Berger B: **Coexpression enables multi-study cellular trajectories of development and disease**. *bioRxiv* 2020, https://doi.org/10.1101/719088. preprint.

21. Chawla S, Samydurai S, Kong SL, Wu Z, Wang Z, Tam WL, Sengupta D, Kumar V: **UniPath: a uniform approach for pathway and gene-set based analysis of heterogeneity in single-cell epigenome and transcriptome profiles**. *Nucleic Acids Res* 2021, **49**:e13, https://doi.org/10.1093/nar/gkaa1138.

22. Wei J, Zhou T, Zhang X, Tian T: **SCOUT: a new algorithm for the inference of pseudo-time trajectory using single-cell data**. *Comput Biol Chem* 2019, **80**:111–120.

23. Haghverdi L, Büttner M, Wolf FA, Buettner F, Theis FJ: **Diffusion pseudotime robustly reconstructs lineage branching**. *Nat Methods* 2016, **13**:845–848.

24. Lin C, Bar-Joseph Z: **Continuous-state HMMs for modeling time-series single-cell RNA-Seq data**. *Bioinformatics* 2019, **35**:4707–4715.

25. Strauß ME, Reid JE, Wernisch L: **GPseudoRank: a permutation sampler for single cell orderings**. *Bioinformatics* 2019, **35**:611–618.

26. Campbell KR, Yau C: **A descriptive marker gene approach to single-cell pseudotime inference**. *Bioinformatics* 2019, **35**:28–35.

27. Ahmed S, Rattray M, Boukouvalas A: **GrandPrix: scaling up the bayesian GPLVM for single-cell data**. *Bioinformatics* 2019, **35**:47–54.

28. La Manno G, Soldatov R, Zeisel A, Braun E, Hochgerner H, Petukhov V, Lidschreiber K, Kastriti ME, Lönnerberg P, Furlan A, Fan J, Borm LE, Liu Z, van Bruggen D, Guo J, He X, Barker R, Sundström E, Castelo-Branco G, Cramer P, Adameyko I, Linnarsson S, Kharchenko PV: **RNA velocity of single cells**. *Nature* 2018, **560**:494–498.

29. * Bergen V, Lange M, Peidli S, Alexander Wolf F, Theis FJ: **Generalizing RNA velocity to transient cell states through dynamical modeling**. *Nat Biotechnol* 2020, **38**:1408–1414.
scVelo generalizes the RNA velocity concept to one that can be applied to developmental processes containing transient cell states.

30. F. Theis, M. Lange, V. Bergen, M. Klein, M. Setty, B. Reuter, M. Bakhti, H. Lickert, M. Ansari, J. Schniering, H. Schiller, D. Pe'er, CellRank for directed single-cell fate mapping.

31. Cao K, Bai X, Hong Y, Wan L: **Unsupervised topological alignment for single-cell multi-omics integration**. *Bioinformatics* 2020, **36**(Supplement_1):i48–i56.

32. ** Y. Hao, S. Hao, E. Andersen-Nissen, W. M. Mauck, S. Zheng, A. Butler, M. J. Lee, A. J. Wilk, C. Darby, M. Zagar, P. Hoffman, M. Stoeckius, E. Papalexi, E. P. Mimitou, J. Jain, A. Srivastava, T. Stuart, L. B. Fleming, B. Yeung, A. J. Rogers, J. M. McElrath, C. A. Blish, R. Gottardo, P. Smibert, R. Satija, Integrated analysis of multimodal single-cell data,Implements weighted nearest neighbor, allowing to define cellular states based on multi-modal information. They combine CITE-seq data with transcriptomics data to develop a multimodal atlas of human PBMC's.

33. Tran TN, Bader GD: **Tempora: cell trajectory inference using time-series single-cell RNA sequencing data**. *PLoS Comput Biol* 2020, **16**, e1008205.

34. Margolin AA, Nemenman I, Basso K, Wiggins C, Stolovitzky G, Favera RD, Califano A: *ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context*. 2006.

35. Zappia L, Phipson B, Oshlack A: **Splatter: simulation of single-cell RNA sequencing data**. *Genome Biol* 2017, **18**:174.

36. Papadopoulos N, Gonzalo PR, Söding J: **PROSSTT: probabilistic simulation of single-cell RNA-seq data for complex differentiation processes**. *Bioinformatics* 2019, **35**:3517–3519.

37. R. Cannoodt, W. Saelens, L. Deconinck, Y. Saeys, dyngen: a multi-modal simulator for spearheading new single-cell omics analyses.

38. * Chen H, Albergante L, Hsu JY, Lareau CA, Lo Bosco G, Guan J, Zhou S, Gorban AN, Bauer DE, Aryee MJ, Langenau DM, Zinovyev A, Buenrostro JD, Yuan G-C, Pinello L: **Single-cell trajectories reconstruction, exploration and mapping of omics data with STREAM**. *Nat Commun* 2019, **10**:1903.
This paper presents, besides a trajectory inference method, an interesting and insightful way to visualise trajectories and the information they contain.

39. Campbell KR, Yau C: **switchde: inference of switch-like differential expression along single-cell trajectories**. *Bioinformatics* 2017, **33**:1241–1242.

40. * Van den Berge K, Roux de Bézieux H, Street K, Saelens W, Cannoodt R, Saeys Y, Dudoit S, Clement L: **Trajectory-based differential expression analysis for single-cell sequencing data**. *Nat Commun* 2020, **11**:1201.
tradeSEQ expands differential expression analysis from a cluster-based technique to one that takes the continuous resolution of trajectory inference into account. It can both find differential expression within and between lineages.

41. Zhang JM, Kamath GM, Tse DN: **Valid post-clustering differential analysis for Single-Cell RNA-Seq**. *Cell Syst* 2019, **9**:383–392.e6.

42. Campbell KR, Yau C: **Uncovering pseudotemporal trajectories with covariates from single cell and bulk expression data**. *Nat Commun* 2018, **9**:2442.

43. * Cacchiarelli D, Qiu X, Srivatsan S, Manfredi A, Ziller M, Overbey E, Grimaldi A, Grimsby J, Pokharel P, Livak KJ, Li S, Meissner A, Mikkelsen TS, Rinn JL, Trapnell C: **Aligning Single-Cell developmental and reprogramming trajectories identifies molecular determinants of myogenic reprogramming outcome**. *Cell Syst* 2018, **7**:258–268. e3.
This method allows shifting the pseudotimes of samples to a common axis, enabling comparisons of expression dynamics along trajectories.

44. Kanton S, Boyle MJ, He Z, Santel M, Weigert A, Sanchís-Calleja F, Guijarro P, Sidow L, Fleck JS, Han D, Qian Z, Heide M, Huttner WB, Khaitovich P, Pääbo S, Treutlein B, Camp JG: **Organoid single-cell genomic atlas uncovers human-specific features of brain development**. *Nature* 2019, **574**:418–422.

45. Alpert A, Moore LS, Dubovik T, Shen-Orr SS: **Alignment of single-cell trajectories to compare cellular expression dynamics**. *Nat Methods* 2018, **15**:267–270.

46. Huynh-Thu VA, Sanguinetti G: **Gene regulatory network inference: an introductory survey**. *Methods Mol Biol* 2019, **1883**:1–23.

47. Ocone A, Haghverdi L, Mueller NS, Theis FJ: **Reconstructing gene regulatory dynamics from high-dimensional single-cell snapshot data**. *Bioinformatics* 2015, **31**. i89–96.

48. Matsumoto H, Kiryu H, Furusawa C, Ko MSH, Ko SBH, Gouda N, Hayashi T, Nikaido I: **SCODE: an efficient regulatory network inference algorithm from single-cell RNA-Seq during differentiation**. *Bioinformatics* 2017, **33**:2314–2321.

49. Matsumoto H, Kiryu H: **SCOUP: a probabilistic model based on the Ornstein-Uhlenbeck process to analyze single-cell expression data during differentiation**. *BMC Bioinf* 2016, **17**:232.

50. Qiu X, Rahimzamani A, Wang L, Ren B, Mao Q, Durham T, McFaline-Figueroa JL, Saunders L, Trapnell C, Kannan S: **Inferring causal gene regulatory networks from coupled Single-Cell expression dynamics using scribe**. *Cell Syst* 2020, **10**:265–274. e11.