Overview

- Using trajectory inference to study cellular dynamics
- Benchmarking trajectory inference methods
- Benchmarking single-cell omics tools in general
- dyno: inferring and interpreting single-cell trajectories
- dyno and anndata
- Customising dyno plots
Characterise dynamic processes by inferring trajectories

Single-cell omics

Genes

Cells

Trajectory inference

Milestones

Cellular states

Cells positioned along transitions between cell states

Differential Expression

Label

Comp 2

Comp 1

MEF

Induced

Myocyte

Neuron

Hmga2

Cdkn1c

Tnnc2

Syp
Trajectory inference in 2014

**Wanderlust**

- **A**: Early/Young vs. Late/Mature.
- **B**: Short Circuit vs. Graph Walk.
- **C**: Wanderlust Analysis.
  - n-dimensional plot.
  - Average each cell's position across all k INN graphs.
  - For every cell, plot distance relative to start and waypoints.
  - Repeat and refine until convergence.

**Monocle (v1)**

- **a**: Cells represented as points in expression space.
  - Reduce dimensionality.
  - Build MST on cells.
- **b**: Dendrogram of cells.
- **c**: Relative expression of genes.
- **d**: Violin plots of gene expression over pseudo-time.
- **e**: Heatmap and expression plots.
- **f**: Violin plots of gene expression over pseudo-time.

**References**

- *Bendall et al.* Cell 2014. doi: [10.1016/j.cell.2014.04.005](https://doi.org/10.1016/j.cell.2014.04.005).
- *Trapnell et al.* Nat Biotech 2014. doi: [10.1038/nbt.2859](https://doi.org/10.1038/nbt.2859).
TI methods are very diverse
2014-2020: >70 TI methods developed

However, only ±40% of authors benchmark their own tool
Overview

- Using trajectory inference to study cellular dynamics
- Benchmarking trajectory inference methods
- Benchmarking single-cell omics tools in general
- dyno: inferring and interpreting single-cell trajectories
- dyno and anndata
- Customising dyno plots
A large-scale benchmarking study of TI methods

110 real & 229 synthetic datasets + 45 trajectory inference methods + 4 metrics

Accuracy + Scalability + Stability + Usability

User guidelines + New possibilities for developers

Method wrappers
Guidelines app
Benchmarking pipeline

methods.dynverse.org
guidelines.dynverse.org
benchmark.dynverse.org

Saelens*, Cannoodt*, Todorov, Saeys.
Nat Biotechnol 2019. doi:10.1038/s41587-019-0071-9
Assessing the accuracy of a trajectory
## a) Method

### Graph methods

| Method          | Priors required | Wrapper type | Platform | Topology inference | Cycle | Linear | Bilirucation | Multifurcation | Connected | Disconnected | Overall | Accuracy | Scalability | Stability | Usability |
|-----------------|-----------------|--------------|----------|--------------------|-------|--------|-------------|----------------|-----------|-------------|---------|----------|-------------|-----------|-----------|
| PAGA            | ×               | Direct       | Python   | Free               |       |        |             |                |           |             |         |          |             |           |           |
| RaceID / StemID | Proj            | R            | Free     |                    |       |        |             |                |           |             |         |          |             |           |           |
| SLICER          | ×               | Cell         | R        | Free               |       |        |             |                |           |             |         |          |             |           |           |

### Tree methods

| Method          | Priors required | Wrapper type | Platform | Topology inference | Cycle | Linear | Bilirucation | Multifurcation | Connected | Disconnected | Overall | Accuracy | Scalability | Stability | Usability |
|-----------------|-----------------|--------------|----------|--------------------|-------|--------|-------------|----------------|-----------|-------------|---------|----------|-------------|-----------|-----------|
| Slingshot       | ×               | Direct       | R        | Free               |       |        |             |                |           |             |         |          |             |           |           |
| PAGA Tree       | ×               | Direct       | Python   | Free               |       |        |             |                |           |             |         |          |             |           |           |
| MST             | Proj            | R            | Free     |                    |       |        |             |                |           |             |         |          |             |           |           |
| pCreode         | Proj            | Python       | Free     |                    |       |        |             |                |           |             |         |          |             |           |           |
| SCUBA           | ×               | Cluster      | Python   | Free               |       |        |             |                |           |             |         |          |             |           |           |
| Monocle DDRTree | Cell            | R            | Free     |                    |       |        |             |                |           |             |         |          |             |           |           |
| Monocle ICA     | ×               | Cell         | R        | Param              |       |        |             |                |           |             |         |          |             |           |           |
| cellTree mappx  | Cell            | R            | Free     |                    |       |        |             |                |           |             |         |          |             |           |           |
| SLICE           | Direct          | R            | Free     |                    |       |        |             |                |           |             |         |          |             |           |           |
| cellTree VEM    | Cell            | R            | Free     |                    |       |        |             |                |           |             |         |          |             |           |           |
| EPiGraph        | Direct          | R            | Free     |                    |       |        |             |                |           |             |         |          |             |           |           |
| Sincell         | Cell            | R            | Free     |                    |       |        |             |                |           |             |         |          |             |           |           |
| URD             | ×               | Direct       | R        | Free               |       |        |             |                |           |             |         |          |             |           |           |
| CellTraits      | ×               | Cell         | R        | Free               |       |        |             |                |           |             |         |          |             |           |           |
| Mpath           | ×               | Cluster      | R        | Free               |       |        |             |                |           |             |         |          |             |           |           |
| CellRouter      | ×               | Cell         | R        | Free               |       |        |             |                |           |             |         |          |             |           |           |

### Multifurcation methods

| Method          | Priors required | Wrapper type | Platform | Topology inference | Cycle | Linear | Bilirucation | Multifurcation | Connected | Disconnected | Overall | Accuracy | Scalability | Stability | Usability |
|-----------------|-----------------|--------------|----------|--------------------|-------|--------|-------------|----------------|-----------|-------------|---------|----------|-------------|-----------|-----------|
| STEMNET         | ×               | Prob         | R        | Param              |       |        |             |                |           |             |         |          |             |           |           |
| FateID          | ×               | Prob         | R        | Param              |       |        |             |                |           |             |         |          |             |           |           |

## b) Summary

### Aggregated scores per experiment
| Method           | Accuracy | Scalability | Stability | Usability          |
|------------------|----------|-------------|-----------|--------------------|
|                  | Topology | Branch      | Cell positions | Per metric | Per dataset source | Per trajectory type | Predicted time | Similarity between runs | Quality of software and paper |
| PAGA             | Topology | Branch      | Cell positions | Gold | Silver | Dyogen | Dyogen | PROGSTT | Splatter | Cycle | Linear | Bifurcation | Convergence | Multifurcation | Tree | Acryc | Conic | 1m | 100 | 1k | 10k | 100k | 1m | 100k | 1k | 10k | 100k | Pred vs. real | Topology | Branch | Assignment | Per metric | Per dataset source | Per trajectory type |
| RaceID / StemID  |          |             |             |       |       |       |       |       |       |       |       |       |       |       |     |     |     | 1h | 7d | 1h | 1h | 1h | 4h | 1h | 7d | 2h | 1h | 31s | <1s | 0.77 |          |         |         |         |               |               |
| SLICER           |          |             |             |       |       |       |       |       |       |       |       |       |       |       |     |     |     | 1h | 7d | 1h | 1h | 1h | 4h | 1h | 7d | 2h | 1h | 31s | <1s | 0.99 |          |         |         |         |               |               |
|                  | Topology | Branch      | Cell positions |       |       |       |       |       |       |       |       |       |       |       |     |     |     | 1h | 7d | 1h | 1h | 1h | 4h | 1h | 7d | 2h | 1h | 31s | <1s |      |          |         |         |         |               |               |
| Slingshot        | Topology | Branch      | Cell positions |       |       |       |       |       |       |       |       |       |       |       |     |     |     | 1h | 7d | 1h | 1h | 1h | 4h | 1h | 7d | 2h | 1h | 31s | <1s |      |          |         |         |         |               |               |
| PAGA Tree        | Topology | Branch      | Cell positions |       |       |       |       |       |       |       |       |       |       |       |     |     |     | 1h | 7d | 1h | 1h | 1h | 4h | 1h | 7d | 2h | 1h | 31s | <1s |      |          |         |         |         |               |               |
| MST              | Topology | Branch      | Cell positions |       |       |       |       |       |       |       |       |       |       |       |     |     |     | 1h | 7d | 1h | 1h | 1h | 4h | 1h | 7d | 2h | 1h | 31s | <1s |      |          |         |         |         |               |               |
| pCreode          | Topology | Branch      | Cell positions |       |       |       |       |       |       |       |       |       |       |       |     |     |     | 1h | 7d | 1h | 1h | 1h | 4h | 1h | 7d | 2h | 1h | 31s | <1s |      |          |         |         |         |               |               |
| SCUBA            | Topology | Branch      | Cell positions |       |       |       |       |       |       |       |       |       |       |       |     |     |     | 1h | 7d | 1h | 1h | 1h | 4h | 1h | 7d | 2h | 1h | 31s | <1s |      |          |         |         |         |               |               |
| Monocle DDRTree  | Topology | Branch      | Cell positions |       |       |       |       |       |       |       |       |       |       |       |     |     |     | 1h | 7d | 1h | 1h | 1h | 4h | 1h | 7d | 2h | 1h | 31s | <1s |      |          |         |         |         |               |               |
| Monocle ICA      | Topology | Branch      | Cell positions |       |       |       |       |       |       |       |       |       |       |       |     |     |     | 1h | 7d | 1h | 1h | 1h | 4h | 1h | 7d | 2h | 1h | 31s | <1s |      |          |         |         |         |               |               |
| cellTree mappix  | Topology | Branch      | Cell positions |       |       |       |       |       |       |       |       |       |       |       |     |     |     | 1h | 7d | 1h | 1h | 1h | 4h | 1h | 7d | 2h | 1h | 31s | <1s |      |          |         |         |         |               |               |
| SLICE            | Topology | Branch      | Cell positions |       |       |       |       |       |       |       |       |       |       |       |     |     |     | 1h | 7d | 1h | 1h | 1h | 4h | 1h | 7d | 2h | 1h | 31s | <1s |      |          |         |         |         |               |               |
| cellTree VEM     | Topology | Branch      | Cell positions |       |       |       |       |       |       |       |       |       |       |       |     |     |     | 1h | 7d | 1h | 1h | 1h | 4h | 1h | 7d | 2h | 1h | 31s | <1s |      |          |         |         |         |               |               |
| EIPiGraph        | Topology | Branch      | Cell positions |       |       |       |       |       |       |       |       |       |       |       |     |     |     | 1h | 7d | 1h | 1h | 1h | 4h | 1h | 7d | 2h | 1h | 31s | <1s |      |          |         |         |         |               |               |
| Sincell          | Topology | Branch      | Cell positions |       |       |       |       |       |       |       |       |       |       |       |     |     |     | 1h | 7d | 1h | 1h | 1h | 4h | 1h | 7d | 2h | 1h | 31s | <1s |      |          |         |         |         |               |               |
| URD              | Topology | Branch      | Cell positions |       |       |       |       |       |       |       |       |       |       |       |     |     |     | 1h | 7d | 1h | 1h | 1h | 4h | 1h | 7d | 2h | 1h | 31s | <1s |      |          |         |         |         |               |               |
| CellTraals       | Topology | Branch      | Cell positions |       |       |       |       |       |       |       |       |       |       |       |     |     |     | 1h | 7d | 1h | 1h | 1h | 4h | 1h | 7d | 2h | 1h | 31s | <1s |      |          |         |         |         |               |               |
| Mpath            | Topology | Branch      | Cell positions |       |       |       |       |       |       |       |       |       |       |       |     |     |     | 1h | 7d | 1h | 1h | 1h | 4h | 1h | 7d | 2h | 1h | 31s | <1s |      |          |         |         |         |               |               |
| CellRouter       | Topology | Branch      | Cell positions |       |       |       |       |       |       |       |       |       |       |       |     |     |     | 1h | 7d | 1h | 1h | 1h | 4h | 1h | 7d | 2h | 1h | 31s | <1s |      |          |         |         |         |               |               |

**Multifurcation methods**

- Topology:
  - Gold: Gold Standard
  - Silver: Silver Standard
  - Dyogen: Dyogen Standard
  - PROGSTT: PROGSTT Standard

- Branch:
  - Linear: Linear Branch
  - Bifurcation: Bifurcation Branch

- Cell positions:
  - Connected: Connected Cell Positions
  - Disconnected: Disconnected Cell Positions

- Convergence:
  - Tree: Tree Convergence
  - Acryc: Acryc Convergence

- Predict vs. real:
  - Pred: Predicted Time
  - vs. real: Real Time

- Availability:
  - Code assurance: Code Assurance

- Quality of paper:
  - Quality: Paper Quality
Large variability in scores
Guidelines for trajectory inference

- **Disconnected**
  - Start cell(s)
  - Accuracy: +
  - Usability: ±
  - Estimated running time (100k = 100k) = 7m 55s 19s
  - Required priors

- **Graph**
  - Start cell(s)
  - Accuracy: +
  - Usability: ±
  - Estimated running time (100k = 100k) = 7m 55s 19s
  - Required priors

- **Tree**
  - Start cell(s)
  - Accuracy: ±
  - Usability: ±
  - Estimated running time (100k = 100k) = 7m 55s 19s
  - Required priors

- **Multifurcation**
  - Start cell(s)
  - Accuracy: ±
  - Usability: ±
  - Estimated running time (100k = 100k) = 7m 55s 19s
  - Required priors

- **Bifurcation**
  - Start cell(s)
  - Accuracy: ±
  - Usability: ±
  - Estimated running time (100k = 100k) = 7m 55s 19s
  - Required priors

- **Linear**
  - Start cell(s)
  - Accuracy: ±
  - Usability: ±
  - Estimated running time (100k = 100k) = 7m 55s 19s
  - Required priors

- **Cycle**
  - Start cell(s)
  - Accuracy: ±
  - Usability: ±
  - Estimated running time (100k = 100k) = 7m 55s 19s
  - Required priors

**Confirm expectations using a method with free topology**
**Confirm results using at least 2 methods**
**Check out the interactive guidelines at guidelines.dynverse.org**

- **Yes**
  - Do you expect multiple disconnected trajectories?
  - Yes / I don't know

- **No**
  - Do you expect cycles in the topology?
  - No / I don't know

- **Yes**
  - Do you expect a tree with two or more bifurcations?
  - Yes

- **Fixed topology**

- **No**
  - Do you expect a particular topology?
  - Yes

- **Free topology**

**End cell(s), Cell Clustering**
**Cell clustering, Start & end cells**
**End cell(s), Cell Clustering**

- **STEMNET**
  - Accuracy: +
  - Usability: ±
  - Estimated running time (100k = 100k) = 7m 55s 19s
  - Required priors

- **Slingshot**
  - Accuracy: +
  - Usability: ±
  - Estimated running time (100k = 100k) = 7m 55s 19s
  - Required priors

- **FateID**
  - Accuracy: +
  - Usability: ±
  - Estimated running time (100k = 100k) = 7m 55s 19s
  - Required priors

- **GrandPrix**
  - Accuracy: ±
  - Usability: ±
  - Estimated running time (100k = 100k) = 7m 55s 19s
  - Required priors

- **STEMNET**
  - Accuracy: ±
  - Usability: ±
  - Estimated running time (100k = 100k) = 7m 55s 19s
  - Required priors

- **SCORPIUS**
  - Accuracy: ±
  - Usability: ±
  - Estimated running time (100k = 100k) = 7m 55s 19s
  - Required priors

- **Embeddr**
  - Accuracy: ±
  - Usability: ±
  - Estimated running time (100k = 100k) = 7m 55s 19s
  - Required priors

- **TSCAN**
  - Accuracy: ±
  - Usability: ±
  - Estimated running time (100k = 100k) = 7m 55s 19s
  - Required priors

- **Slingshot**
  - Accuracy: ±
  - Usability: ±
  - Estimated running time (100k = 100k) = 7m 55s 19s
  - Required priors

- **Angle**
  - Accuracy: ±
  - Usability: ±
  - Estimated running time (100k = 100k) = 7m 55s 19s
  - Required priors
Want to learn about computational tools for single-cell omics data? Read benchmarking papers!

Clustering
- A systematic performance evaluation of clustering methods for single-cell RNA-seq data. doi.org/gfgp3k
- Comparison of clustering tools in R for medium-sized 10x Genomics single-cell RNA-sequencing data. doi.org/gd9grv

Normalization
- scRNA-seq mixology: towards better benchmarking of single cell RNA-seq protocols and analysis methods. doi.org/10.1101/433102

Differential expression
- Bias, robustness and scalability in single-cell differential expression analysis. doi.org/10.1038/nmeth.4612

Trajectory inference
- A comparison of single-cell trajectory inference methods. doi.org/10.1038/s41587-019-0071-9

Network inference
- Benchmarking algorithms for gene regulatory network inference from single-cell transcriptomic data. doi.org/10.1038/s41592-019-0690-6
Overview

- Using trajectory inference to study cellular dynamics
- Benchmarking trajectory inference methods
- **Benchmarking single-cell omics tools in general**
- dyno: inferring and interpreting single-cell trajectories
- dyno and anndata
- Customising dyno plots
dyngen: a multi-modal simulator for spearheading single-cell omics analyses

Output of dyngen

Applications of dyngen

G

Trajectory inference

H

Trajectory alignment

Cannoodt*, Saelens*, Deconinck, Saeys
bioRxiv 2020. doi:10.1101/2020.02.06.936971.
dyngen: a multi-modal simulator for spearheading single-cell omics analyses

Cannoodt*, Saelens*, Deconinck, Saeys
bioRxiv 2020. doi:10.1101/2020.02.06.936971.
Synthetic benchmarking with dyngen

**Trajectory alignment**

**RNA velocity**

**Case-wise GRN inference**

Cannoodt*, Saelens*, Deconinck, Saeys
bioRxiv 2020. doi:10.1101/2020.02.06.936971.
Overview

- Using trajectory inference to study cellular dynamics
- Benchmarking trajectory inference methods
- Benchmarking single-cell omics tools in general
- dyno: inferring and interpreting single-cell trajectories
- dyno and anndata
- Customising dyno plots
A common interface for 55 TI methods

Required input
- Raw counts
- Normalised expression
- Start cell(s)
- End cell(s)
- Grouping / clustering
- Dimensionality reduction

Optional prior input

Trajectory inference method

Required output
- Topology
- Cell ordering
- Feature importances

Optional output

- Python containers
dockerhub: python
- R containers
dockerhub: rocker/tidyverse
- R and Python containers
dockerhub: dataintuitive/randpy
dyno: Inferring and interpreting single-cell trajectories

1. Selecting the most optimal method(s)

 guidelines.dynverse.org

2. Inferring trajectories

- A common input and output interface for 55 methods

  | Input | Method | Output |
  |-------|--------|--------|
  | Raw counts | Topology | Cell positions |
  | Normalised | Start cell(s) | |
  | End cell(s) | | |
  | Cell groups | | |

- One line and run any method:

  ```python
  infer_trajectory(dataset, "paga")
  infer_trajectory(dataset, "slingshot")
  infer_trajectory(dataset, "my_favorite_ti_method")
  ```

- Easy to include a new method

  methods.dynverse.org
dyno: Inferring and interpreting single-cell trajectories

3. Annotating the trajectory

- Labelling milestones
- Rooting the trajectory
- Determining the directionality

4. Detecting differential expression

- Overall
- Branch
- Branch point

5. Visualizing the trajectory

- Visualize a model in multiple ways
- Compare models on a common embedding
Frequently asked questions

● How to run dyno with my 'dataset.h5ad' file?
→ add support for importing h5ad files into dyno
● How can I customise my plot?
→ dynplot2: provide ggplot2-like interface
● Can I use RNA velocity to determine the directionality?
→ add support for visualising RNA velocity & reorienting trajectory
● What about new TI method Monocle 3 / STREAM / Palantir?
→ add new and update existing TI methods
Overview

- Using trajectory inference to study cellular dynamics
- Benchmarking trajectory inference methods
- Benchmarking single-cell omics tools in general
- dyno: inferring and interpreting single-cell trajectories
- dyno and anndata
- Customising dyno plots
How to run dyno with my 'dataset.h5ad' file?
How to use h5ad files in R?

- **Directly interface with h5file**
  - `hdf5r::h5file(filename = "dataset.h5ad", mode = 'r')`
  - a lot of manual work (e.g. converting sparse and dense matrices)
  - modified h5 files sometimes cannot be read back into anndata
    (due to incompatibilities between hdf5 API in R and Python)
  - a lot of copy pasted code

- **Use reticulate to access anndata/scanpy**
  - `ad <- reticulate::import("anndata"); ad$read_h5ad("dataset.h5ad")`
  - still a lot of manual work (e.g. converting custom anndata classes)
  - a lot of copy pasted code

- **Convert to Seurat, or to SingleCellExperiment with zellkonverter**
  - best approach
  - if you want an h5ad as output, you need to convert h5ad → SCE/Seurat → h5ad
How to use h5ad files in R?

- Use anndata for R!
  - Uses reticulate to provide anndata interface but supporting common R constructs
  - On GitHub at rcannood/annndata
  - Also on CRAN (But wait until release > 0.7.5.1)
  - Experimental: working through porting the python unit tests to R to detect discrepancies
How to run dyno with my 'dataset.h5ad' file?

```r
library(dyno)
library(tidyverse)
set.seed(1)

ad <- anndata::read_h5ad("ginhoux.h5ad") # read h5ad
dataset <- dynio::from_h5ad(ad) # import to dyno

traj <- infer_trajectory(dataset, ti_slingshot(ndim = 2)) # run slingshot

plot_dimred(traj, label_milestones = TRUE) # plot dimred
plot_heatmap(traj, expression_source = dataset, features_oi = 100) # plot heatmap

dynio::to_h5ad(traj, ad) # add to ad obj
ad$write_h5ad("output.h5ad") # write to file
```
Overview

● Using trajectory inference to study cellular dynamics
● Benchmarking trajectory inference methods
● Benchmarking single-cell omics tools in general
● dyno: inferring and interpreting single-cell trajectories
● dyno and anndata
● Customising dyno plots
How can I customise my plot?

library(dynplot2)
# ... other imports, load dataset, perform trajectory inferen
dynplot_dimred(dataset, traj) +

geom_cell_point(aes(colour = select_feature_expression("Cd
scale_colour_distiller(palette = "RdBu") +
labs(colour = "Log Cd34") +

new_scale_colour() +
geom_trajectory_segments(aes(colour = edge_id), size = 3) 
scale_colour_brewer(palette = "Dark2") + 
labs(colour = "Edge")

- Familiar ggplot2 interface
- Feature parity with dynplot and dynplot2 is a lot of work
