

Pseudotime / Trajectory Inference

Rhonda Bacher

University of Florida

ISMB 2019

Sunday, July 21, 2019

What is pseudotime?

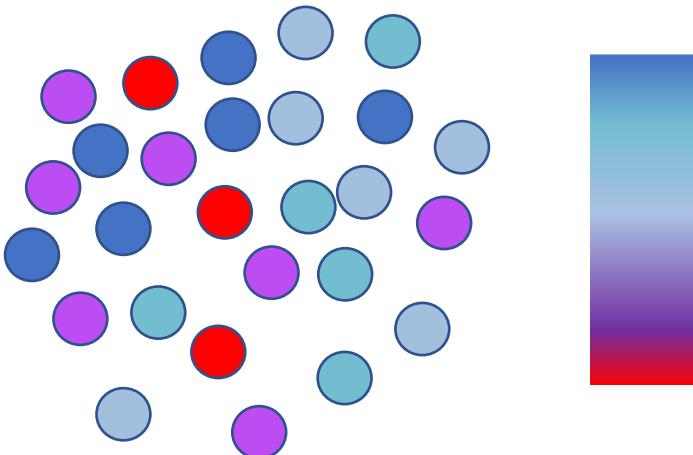
- More generally referred to as trajectory inference methods.
 - Over time, space, or any dynamic dimension.
- Cells can be computationally reordered based on similarities of their transcriptome profile.

What is pseudotime?

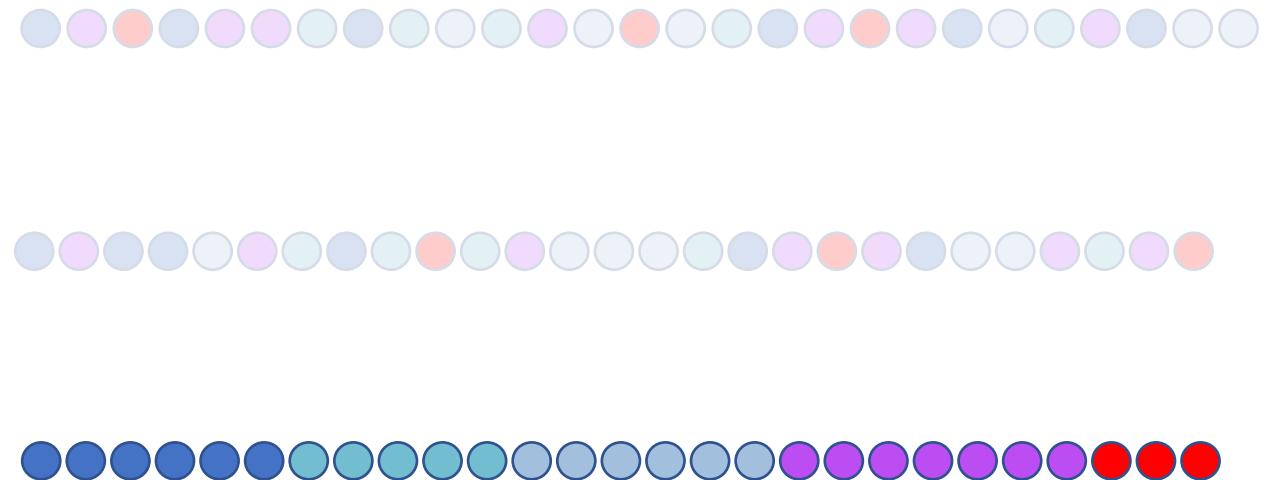
- For cells exposed to a signal (e.g. drug or differentiation factors) or during a dynamic process, cells will vary in their response rate.
- Snapshot scRNA-seq will reflect and capture those variations.
- Computationally reordering the cells based on the similarities/differences in gene expression, reconstructs the dynamics of the underlying process or transcriptional response.

Reconstructing cell order

Color represents the expression of Gene A



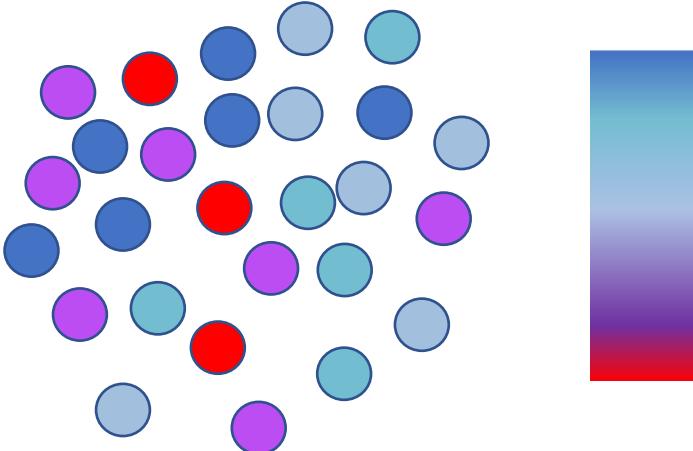
Reordering



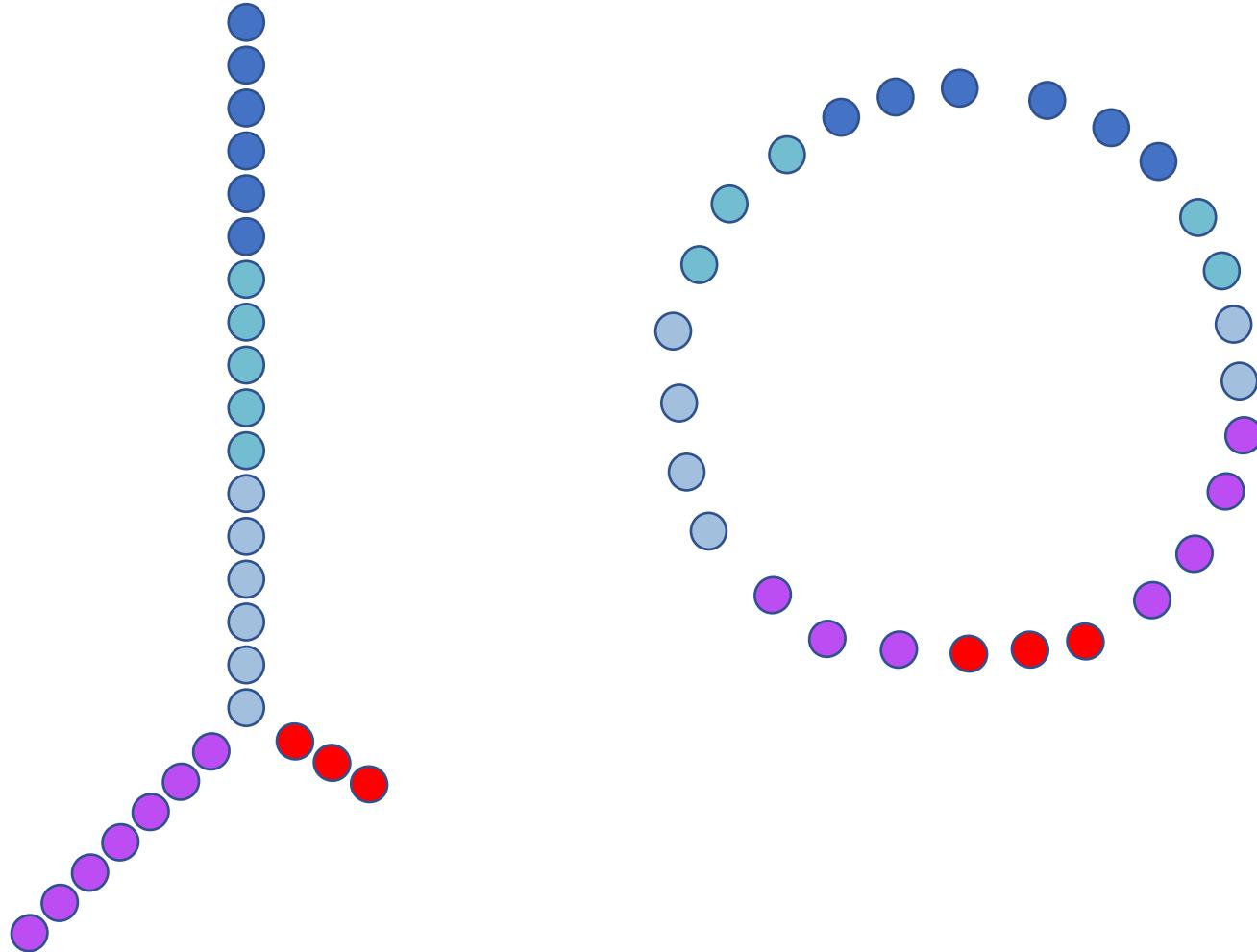
...but many dimensions with over 10,000 genes

Reconstructing cell order

Color represents the expression of Gene A



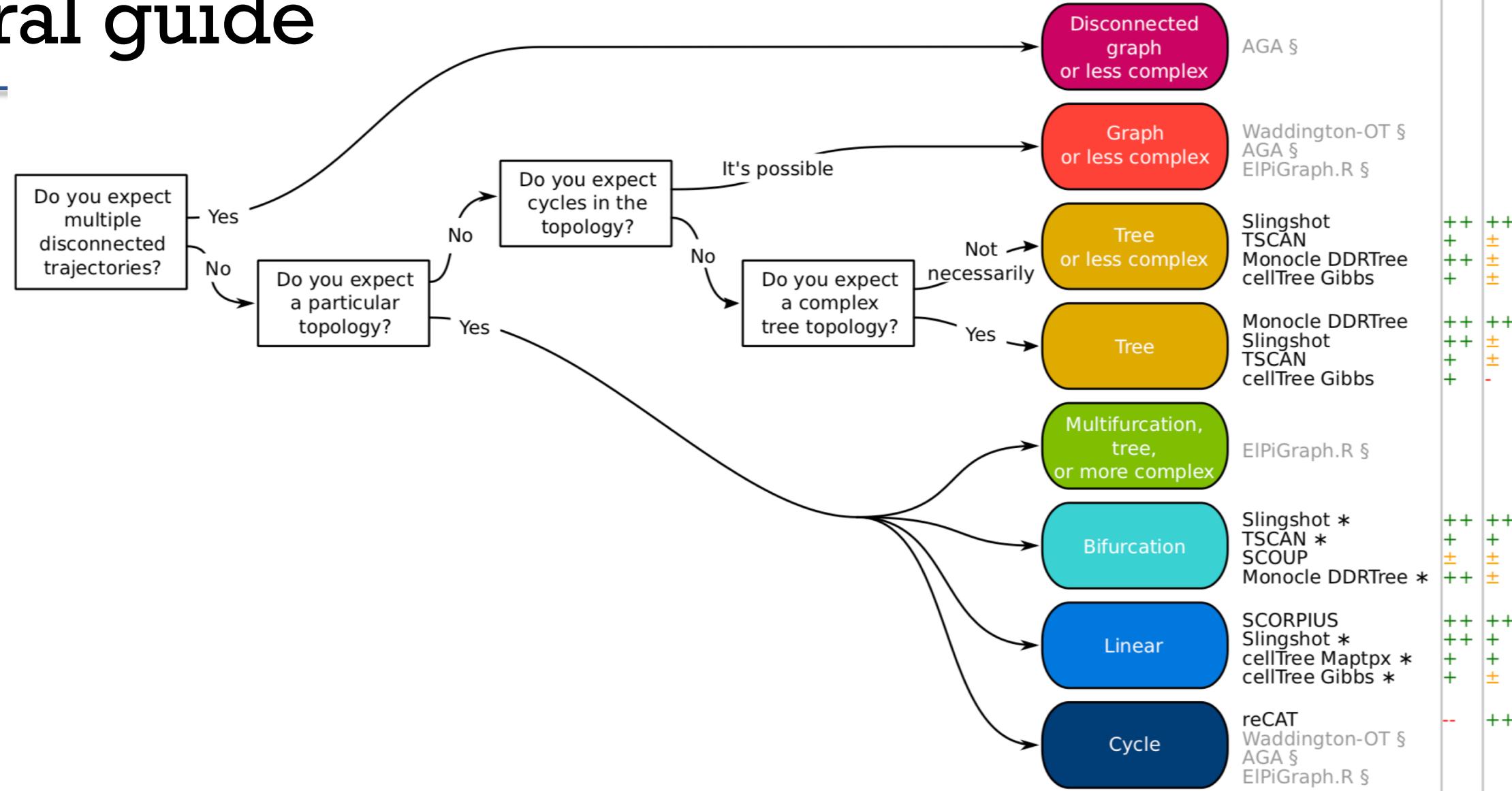
...and what if trajectory isn't linear?



Trajectory inference methods

- More than 50 methods for trajectory inference currently exist!
- <https://github.com/agitter/single-cell-pseudotime>
- Complete benchmarking and evaluation recently by Saelens et al., “A comparison of single-cell trajectory inference methods”. Nature Biotechnology. 2019.

General guide



* Method may return a different topology than requested
 § Not in the current version of the evaluation

Trajectory inference methods

Methods differ in their ability to infer specific topologies:

- Linear: [TSCAN](#)
- Tree: [Slingshot](#)
- Partitional Tree: [Monocle3](#), [PAGA](#)
- All methods implement some preprocessing to remove noisy/lowly expressed and/or lowly variable genes.

Trajectory inference methods

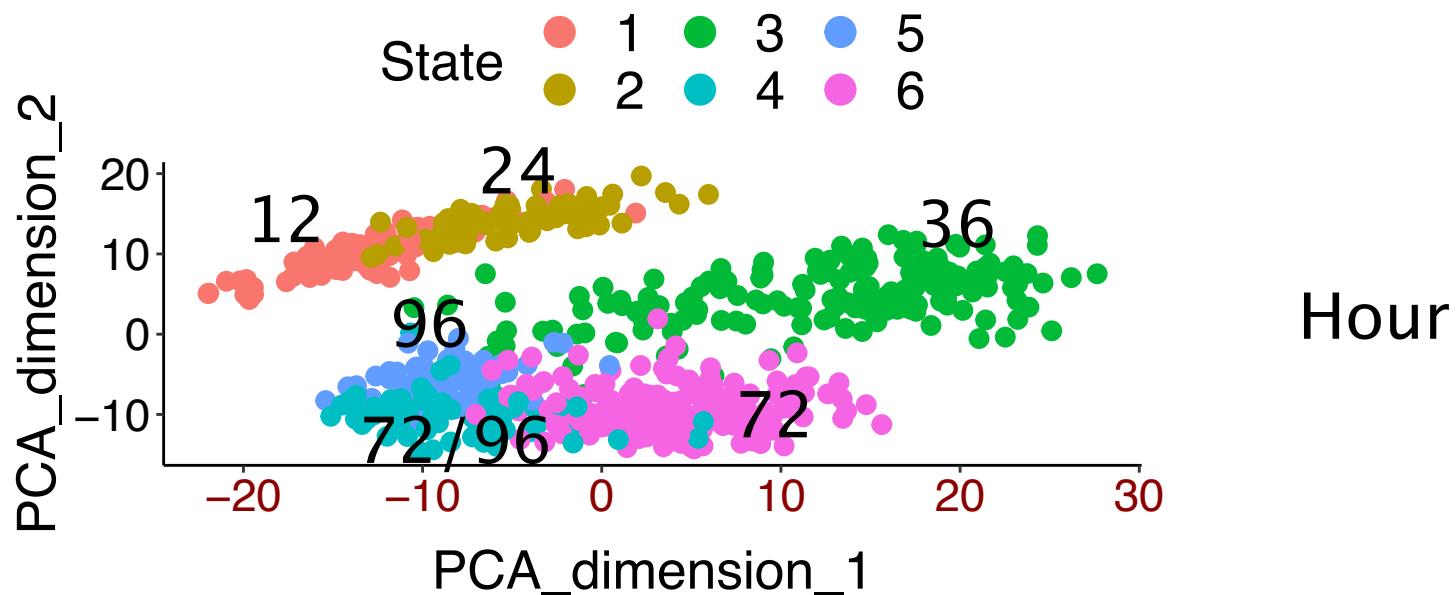
Additional differences between methods:

- Dimensionality reduction is done via PCA, TSNE, diffusion maps.
- Type of graph or tree is inferred on the cells or on clusters of cells and distance measure.
- Amount of supervision needed or able to accommodate.

Linear topology methods-TSCAN

TSCAN

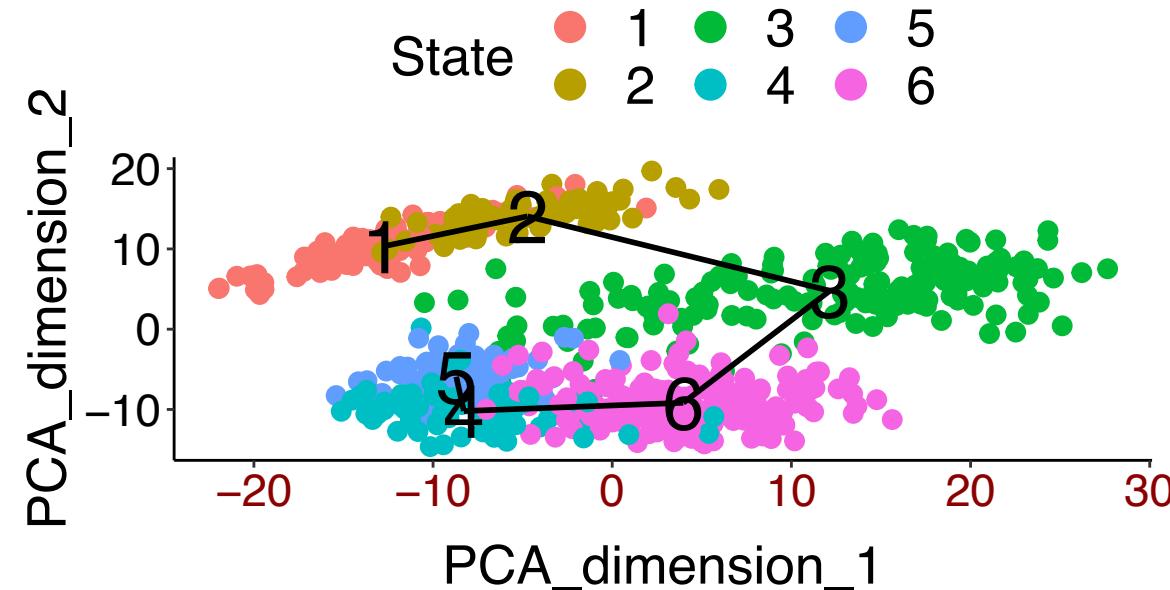
- Reduce dimensionality using PCA.
- Cluster the cells. TSCAN uses mclust, though other methods would be acceptable here.



	1	2	3	4	5	6
12	99	3	0	0	0	0
24	1	63	2	0	0	0
36	0	0	172	0	0	0
72	0	0	1	10	0	127
96	0	0	0	58	78	52

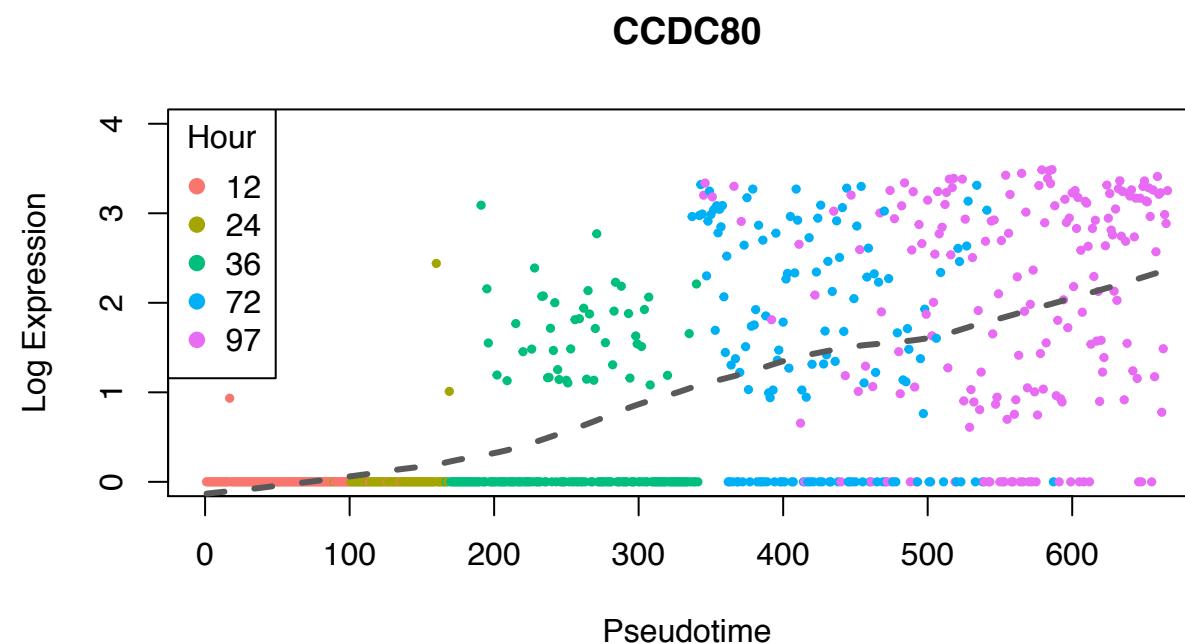
Linear topology methods-TSCAN

- TSCAN
 - Minimum spanning tree (MST) connects the cell cluster centers.



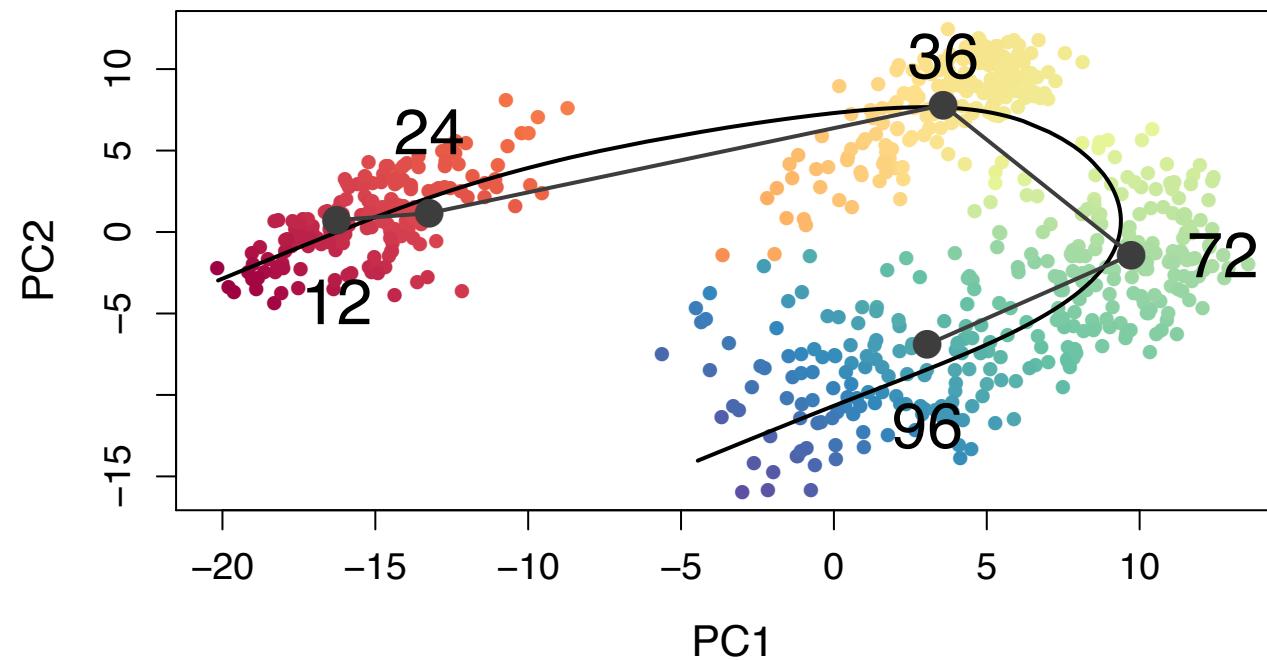
Linear topology methods-TSCAN

- TSCAN
 - Cells are projected onto the tree edge and to obtain pseudotime.

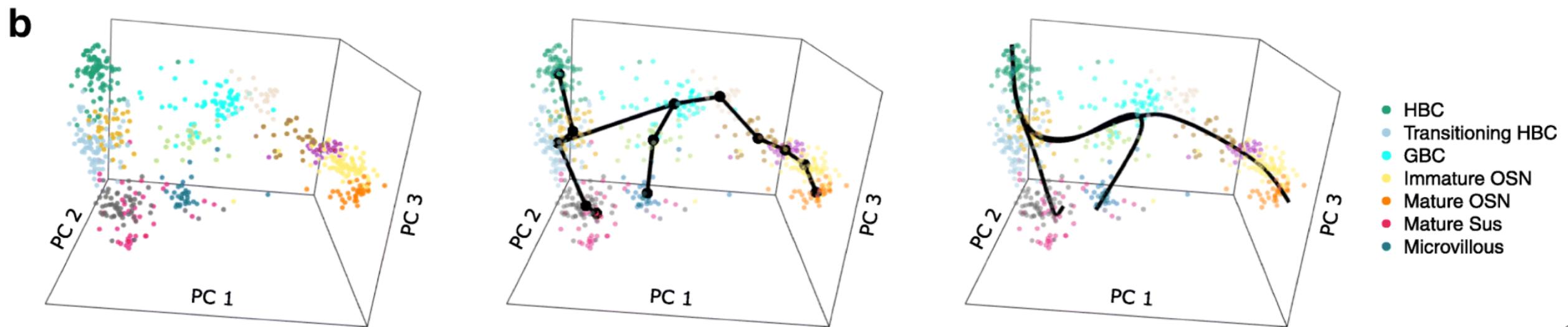
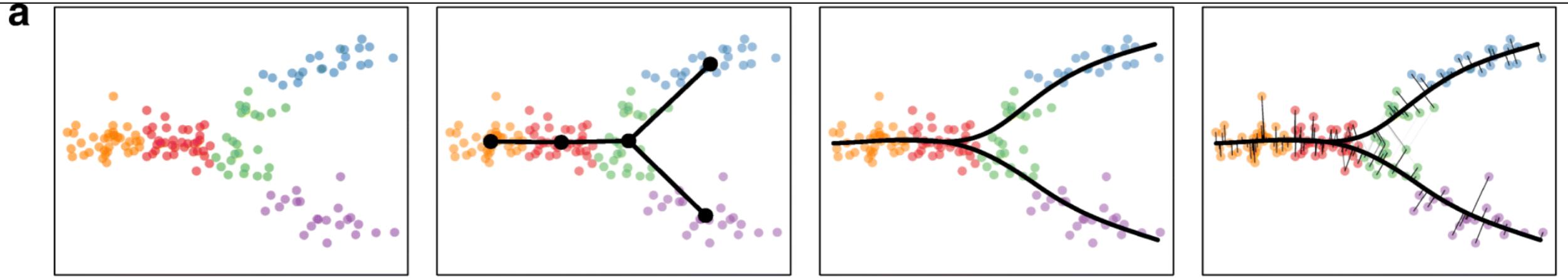


Tree topology methods - Slingshot

- Slingshot
 - Uses an MST like TSCAN with alternative distance metrics.
 - Flexible, allows incorporation of fixed starting and ending clusters.



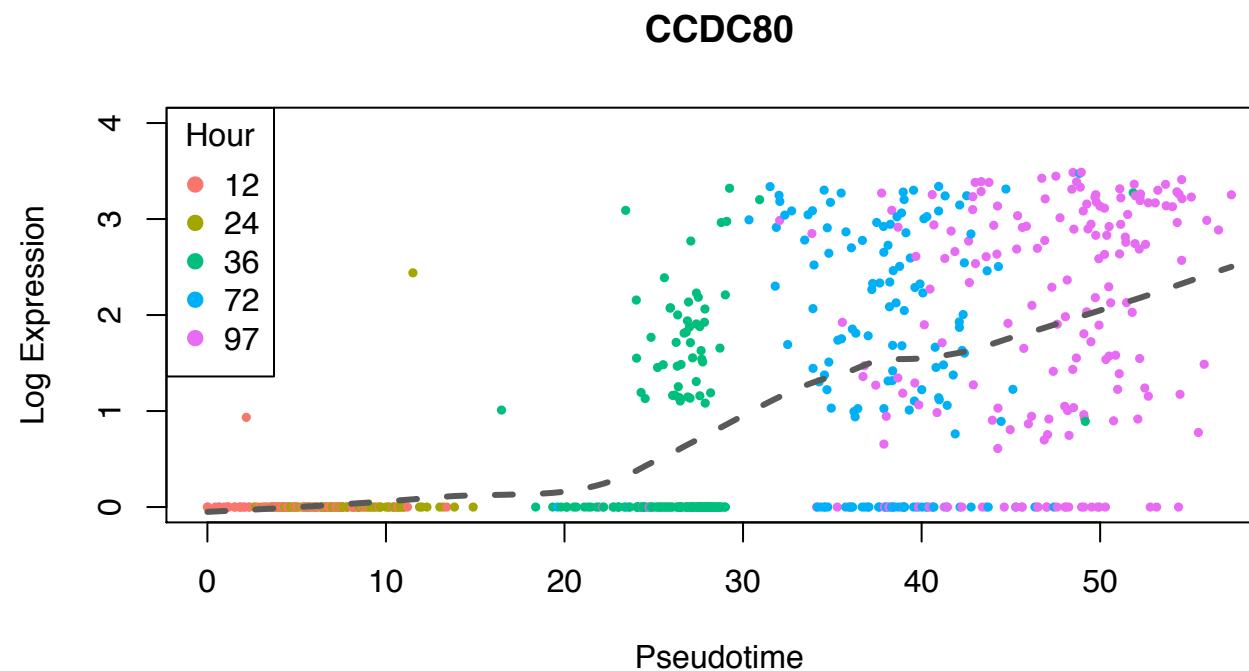
Tree topology methods - Slingshot



Tree topology methods - Slingshot

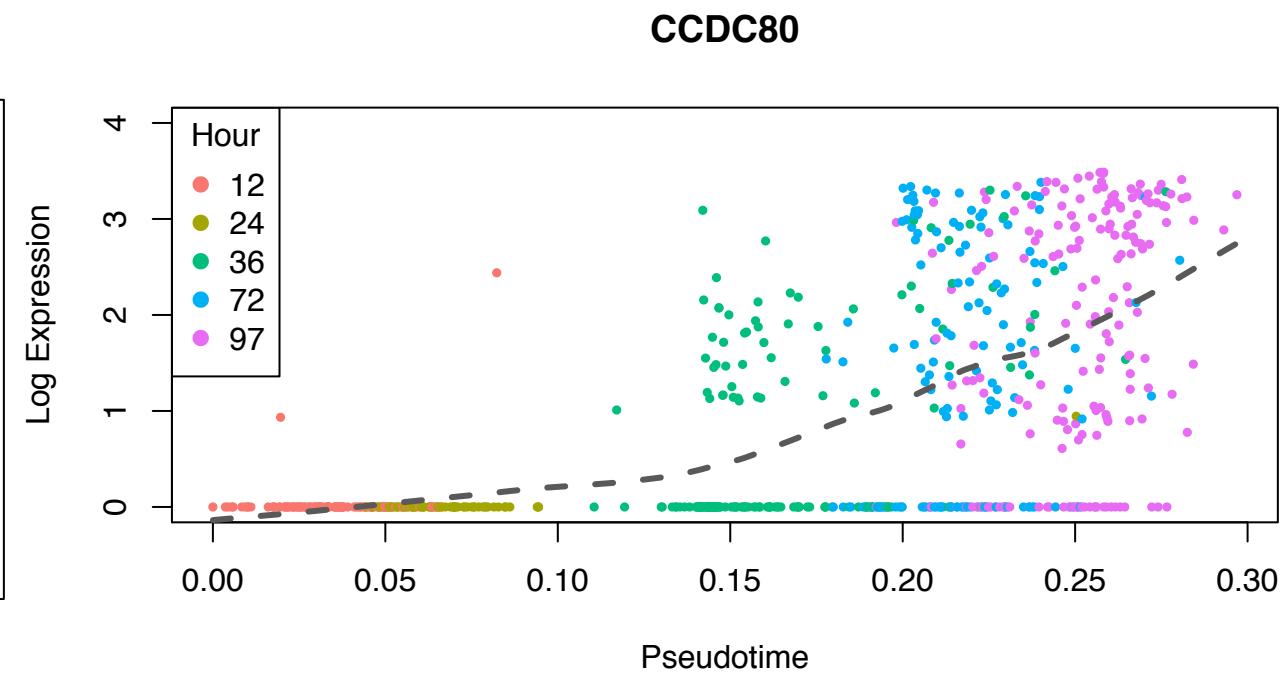
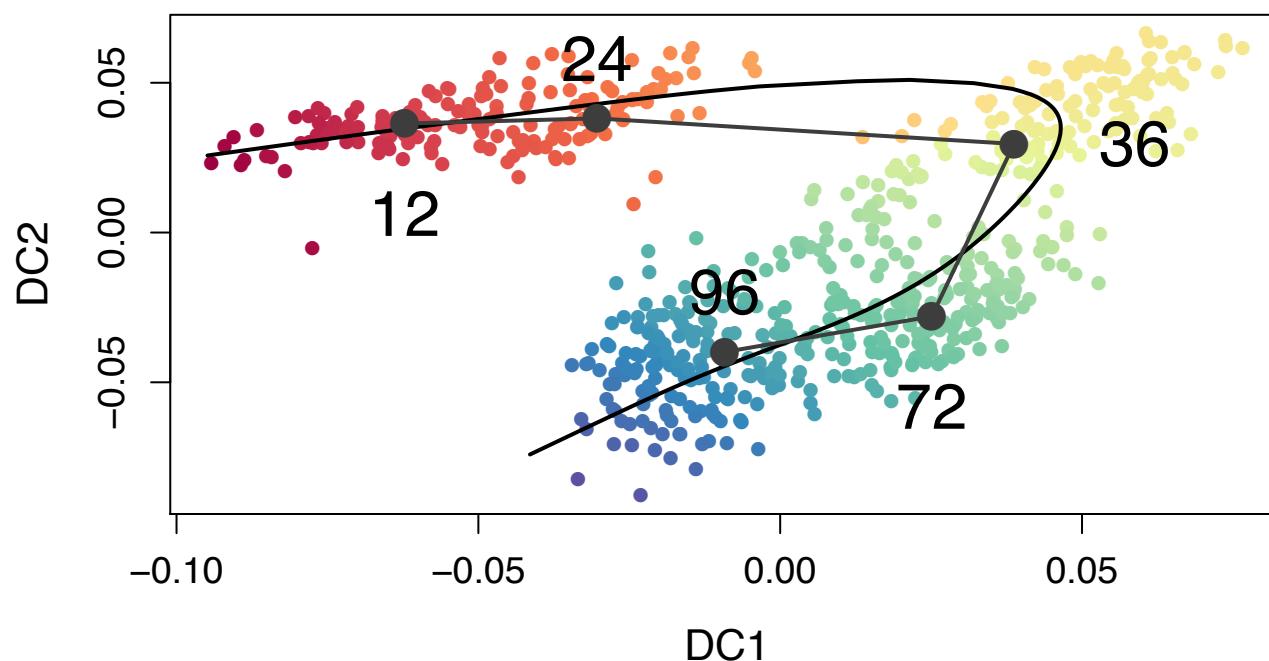
Slingshot

- Fits a principle curve through (each) lineage.
- Orthogonal projection and arc length distance to obtain pseudotime



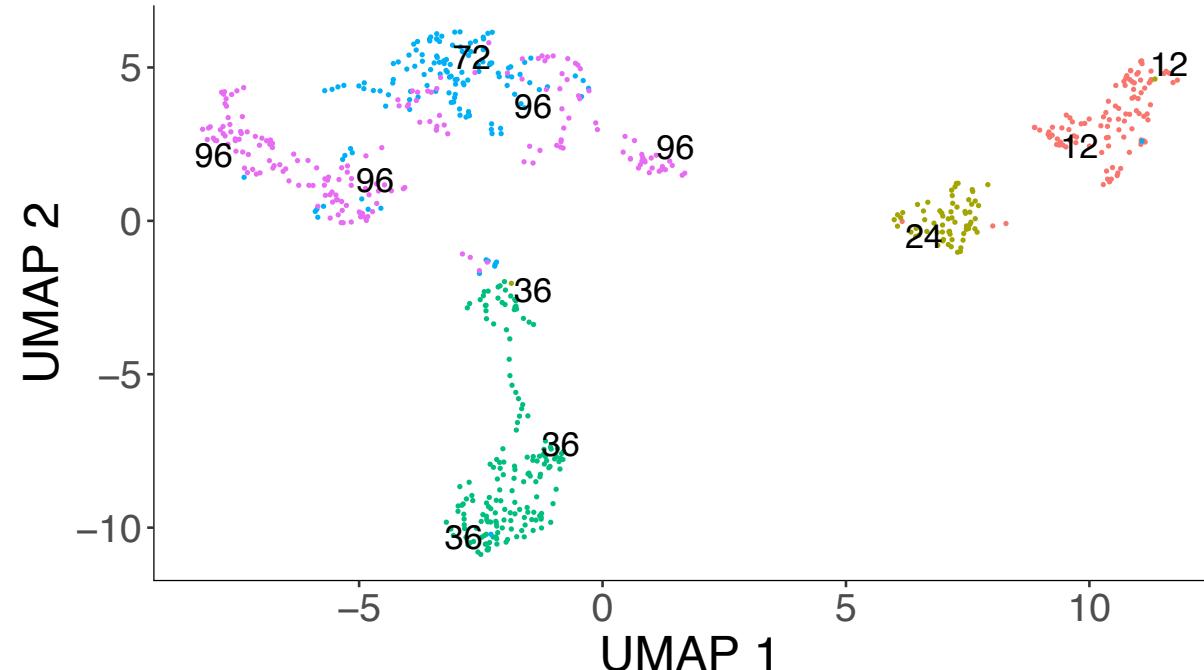
Diffusion Maps

- Diffusion Maps + Slingshot

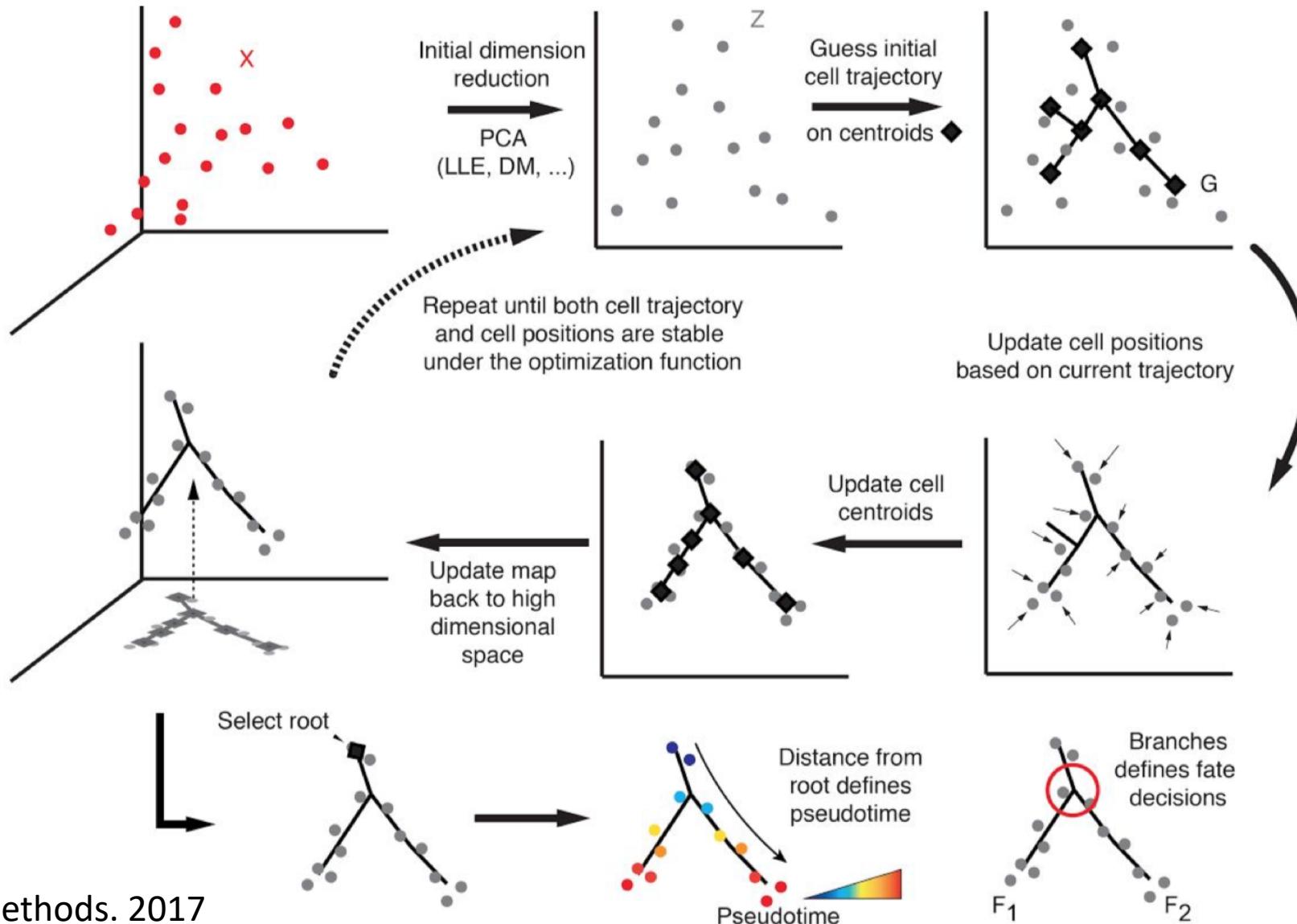


Tree topology methods – Monocle3

- Monocle3 (borrows clustering and graph learning ideas from PAGA)
 - Reduce dimensionality using UMAP.
 - Cluster the cells using Louvain clustering.

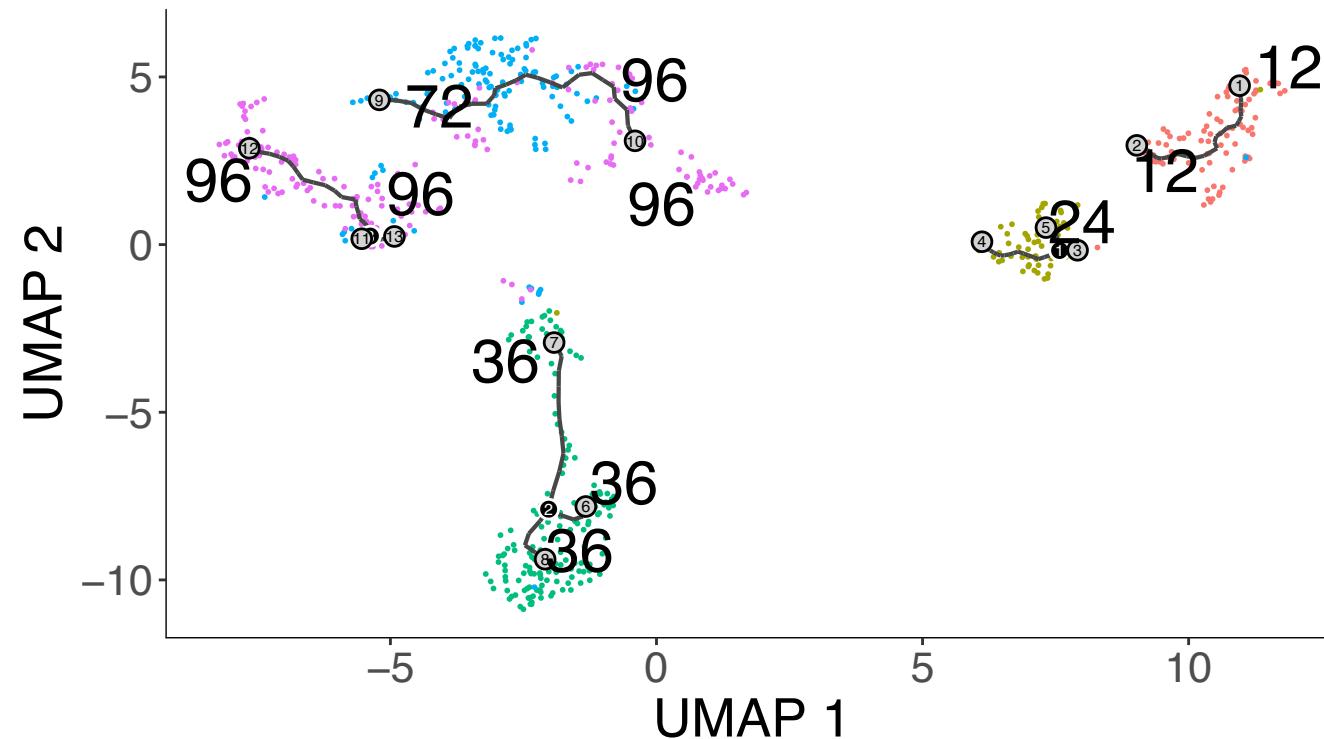


Tree topology methods – Monocle3



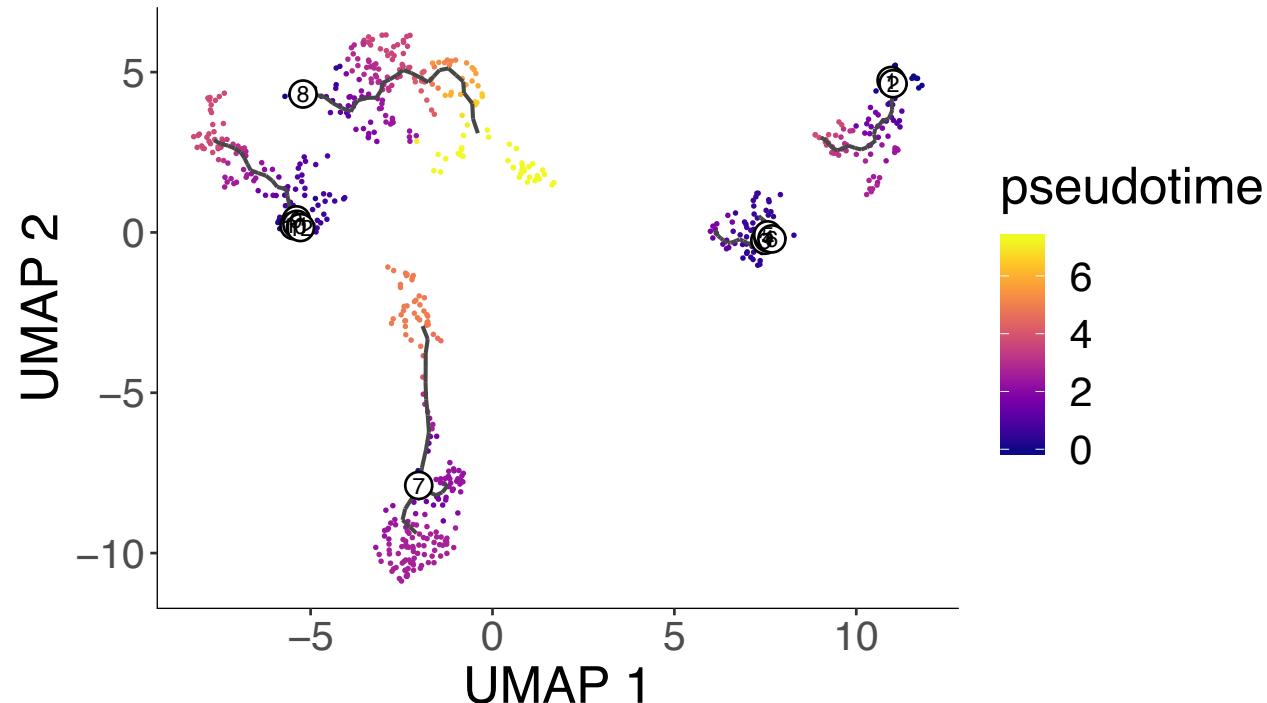
Tree topology methods – Monocle3

- Monocle3
 - Constructs trajectory using PAGA to partition and then principle graph embedding.

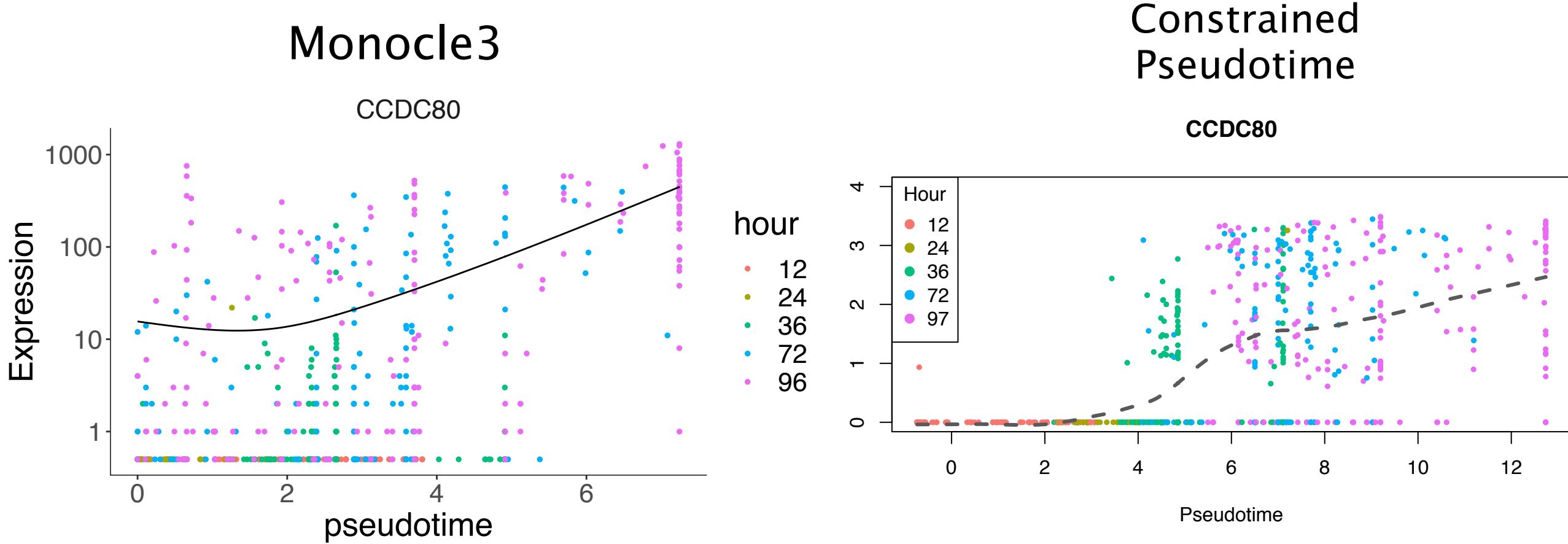


Tree topology methods – Monocle3

- Monocle3
 - Project each cell to the nearest location on tree and pseudotime is the distance from the root.

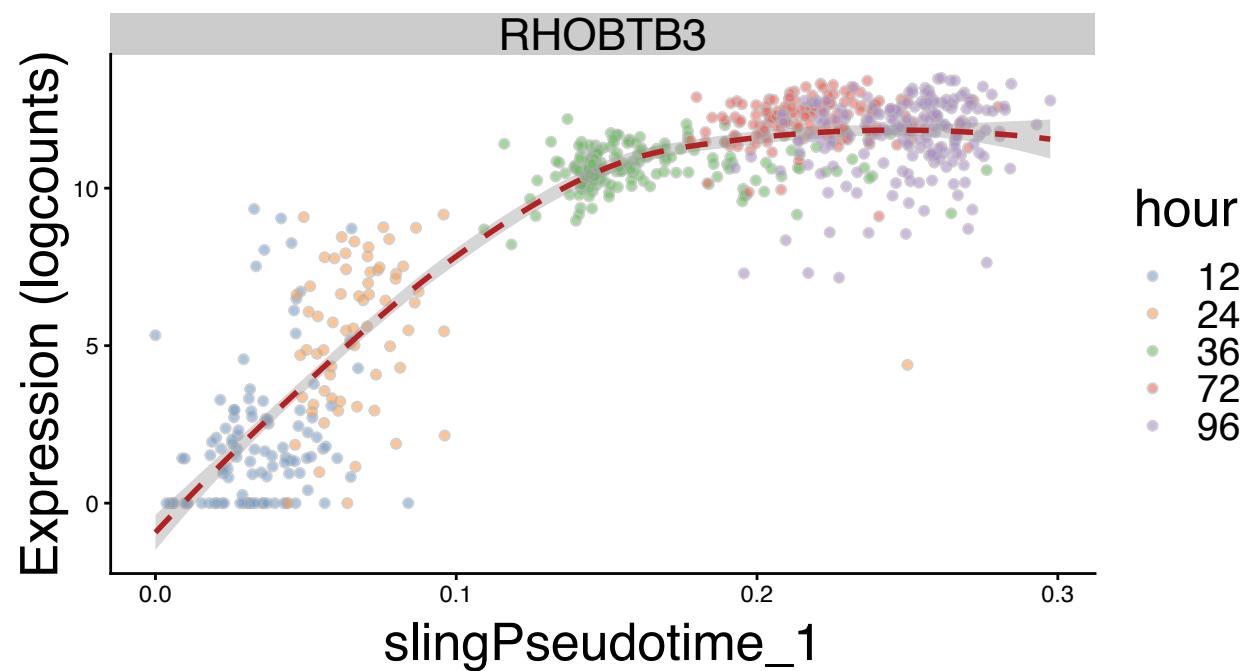


Tree topology methods – Monocle3



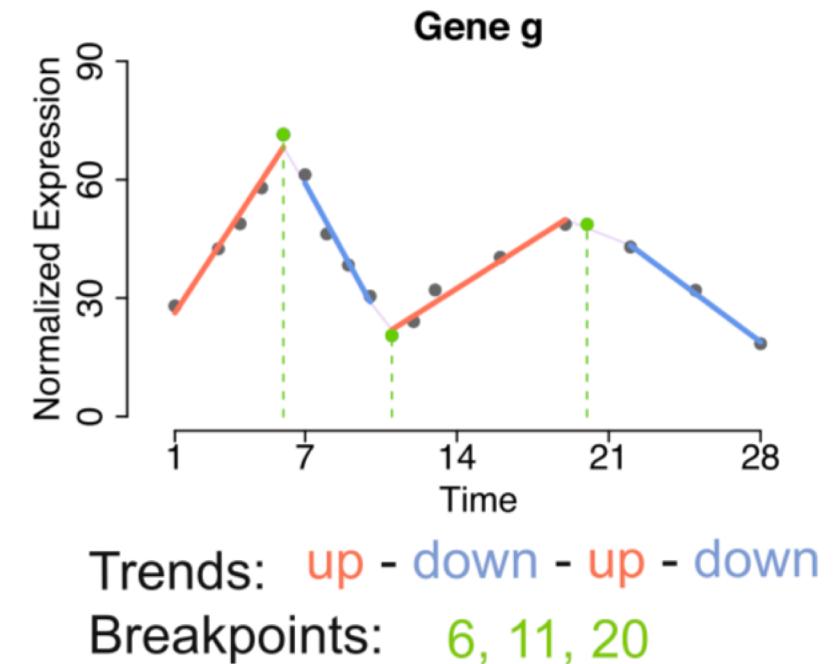
Dynamic/temporally expressed genes

- Identify genes that change as a function of pseudotime.
- Fit a Generalized Additive Model(GAM) with LOESS.



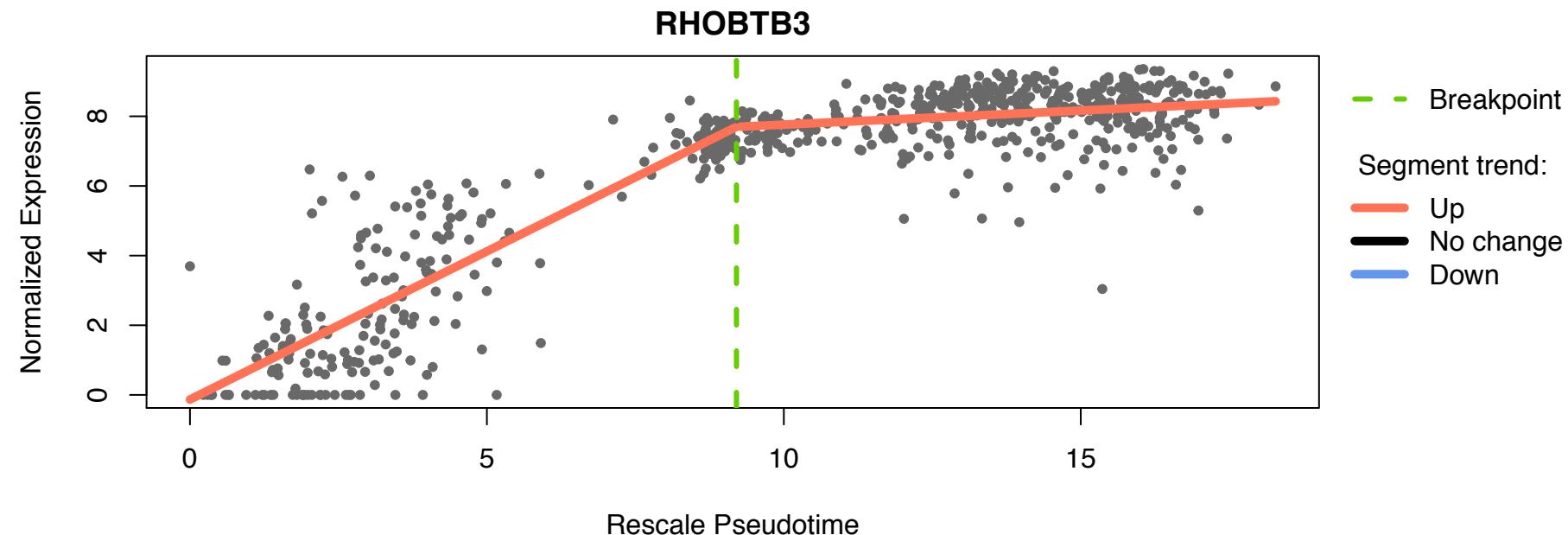
Dynamic/temporally expressed genes

- Trendy
 - Designed for general time-course experiments.
 - Fits a piece-wise (segmented) regression model for each gene, choosing the optimal number and location of break points.

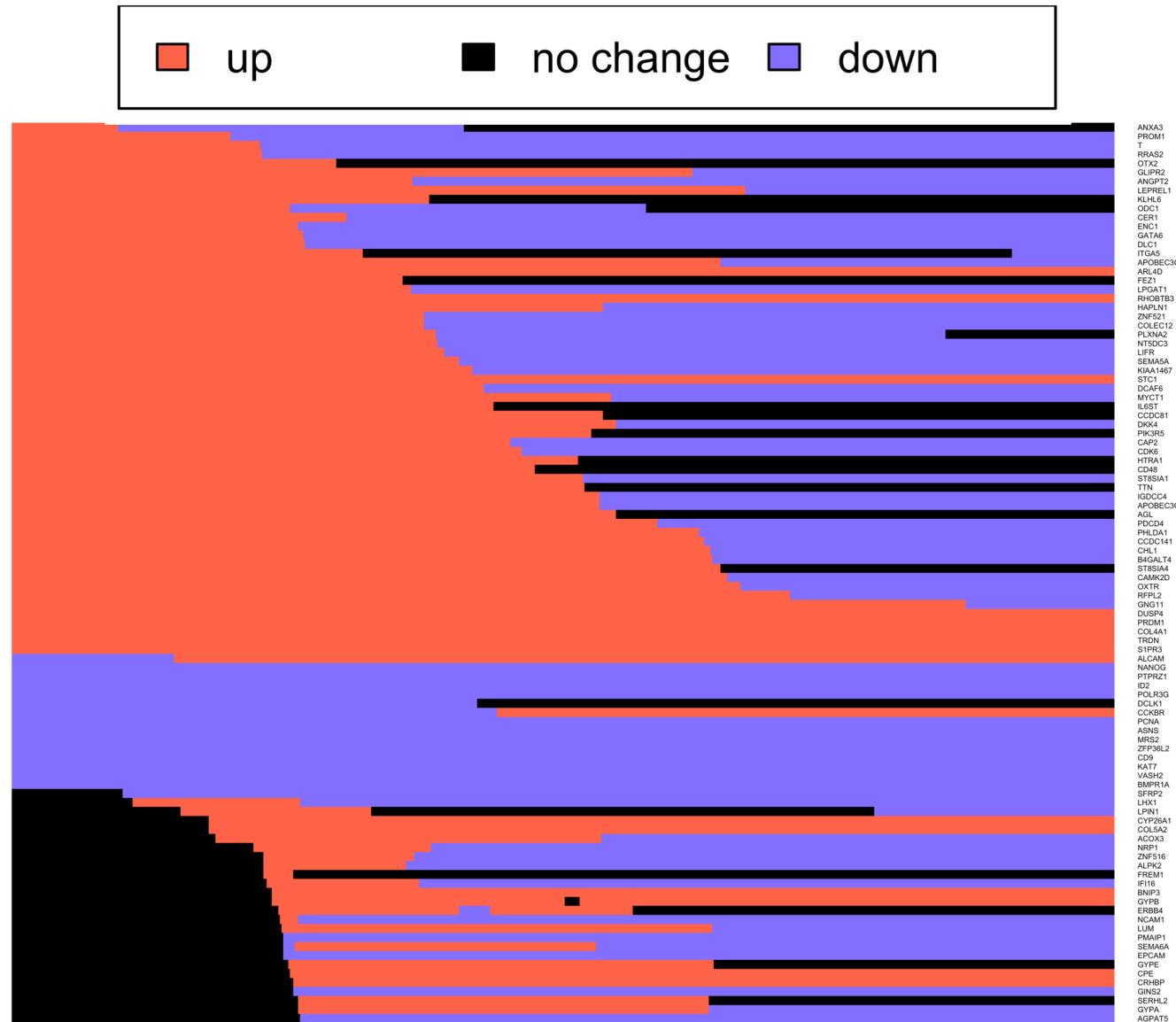


Dynamic/temporally expressed genes

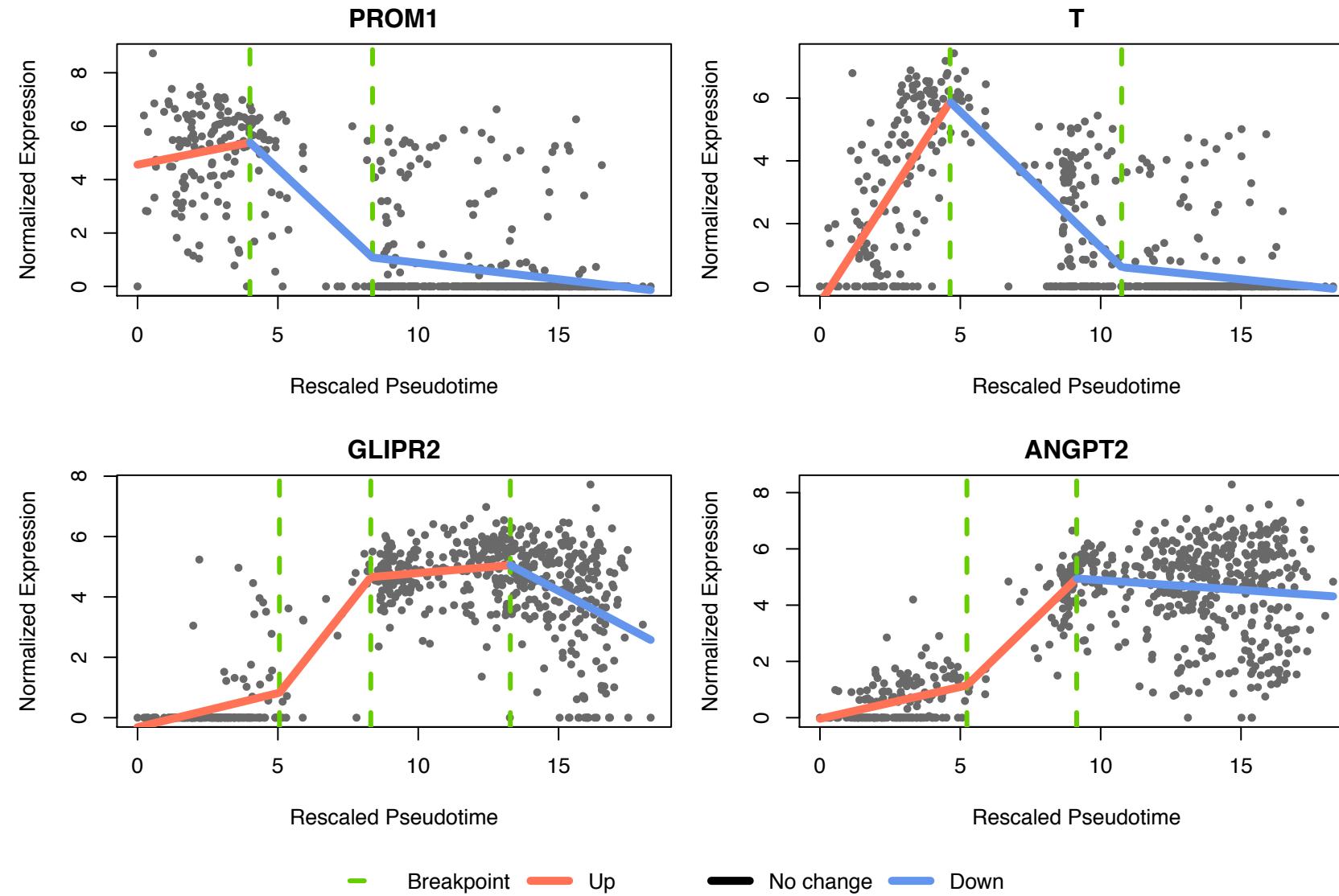
- Trendy
 - Summarize “time” and magnitude of gene expression changes.
 - Compare across genes.



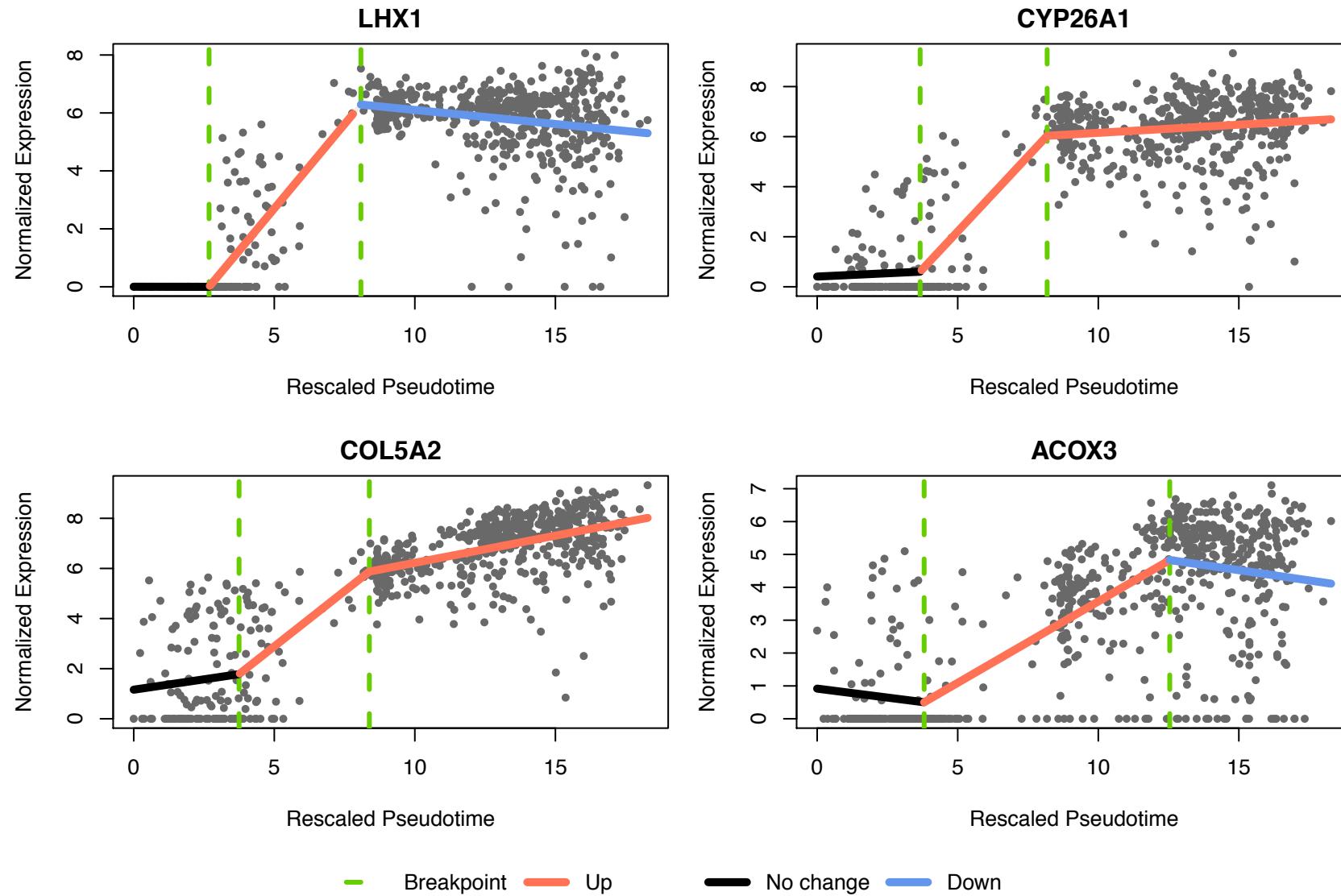
Dynamic/temporally expressed genes- Trendy



Dynamic/temporally expressed genes- Trendy



Dynamic/temporally expressed genes- Trendy



Dynamic/temporally expressed genes- Trendy

