

MBC

Medical Bioinformatics Centre

Turku Bioscience

University of Turku and Åbo Akademi University

FI-20520 Turku, Finland

Trajectory inference with Totem

Sini Junttila, PhD



**UNIVERSITY
OF TURKU**



Åbo Akademi

BF
Biocenter Finland



InFLAMES
Solution is in Immunity



**TURKU
BIOSCIENCE**

Trajectory inference analysis

- Biological processes can be seen as a continuum of dynamic changes within or between cellular states
- These processes can be characterised from scRNA-seq data by identifying a trajectory
- Pseudotime is the positioning of cells along the trajectory that quantifies the relative progression or differentiation of the biological process
- Several TI tools exist
 - Slingshot, TinGa, Monocle, PAGA...

Totem

- Designed to facilitate the inference of tree-shaped trajectories from single-cell data
- Generates a large number of clustering results with a k-medoids algorithm (CLARA) and constructs an MST for each clustering
- Estimates cell connectivity from the MSTs, which measures how many connections the cells have on average to other clusters in the MSTs
- From the ranked clusterings the user can select one or several clusterings and easily visualize the resulting MSTs
- This approach is useful because finding a good clustering automatically that will also generate a sensible MST is often difficult, requiring the user to try different tools and parameters
- Available in Github: <https://github.com/elolab/Totem>
- HTML vignette: <https://htmlpreview.github.io/?https://github.com/elolab/Totem-benchmarking/blob/main/Totem.html>

Trajectory exercise

- Demonstrate the use of Totem
- Script: `trajectory_inference_with_totem.R`
 - Figures are saved to a folder called “Totem_figures”
- Data available in Github: https://github.com/elolab/Olissipo_hands-on_session2
- 203 mouse iNKT cells, identity of the cells is known
- Engel et al (2016): <https://www.nature.com/articles/ni.3437>

[Published: 18 April 2016](#)

Innate-like functions of natural killer T cell subsets result from highly divergent gene programs

[Isaac Engel](#), [Grégory Seumois](#), [Lukas Chavez](#), [Daniela Samaniego-Castruita](#), [Brandie White](#), [Ashu Chawla](#), [Dennis Mock](#), [Pandurangan Vijayanand](#) ✉ & [Mitchell Kronenberg](#) ✉

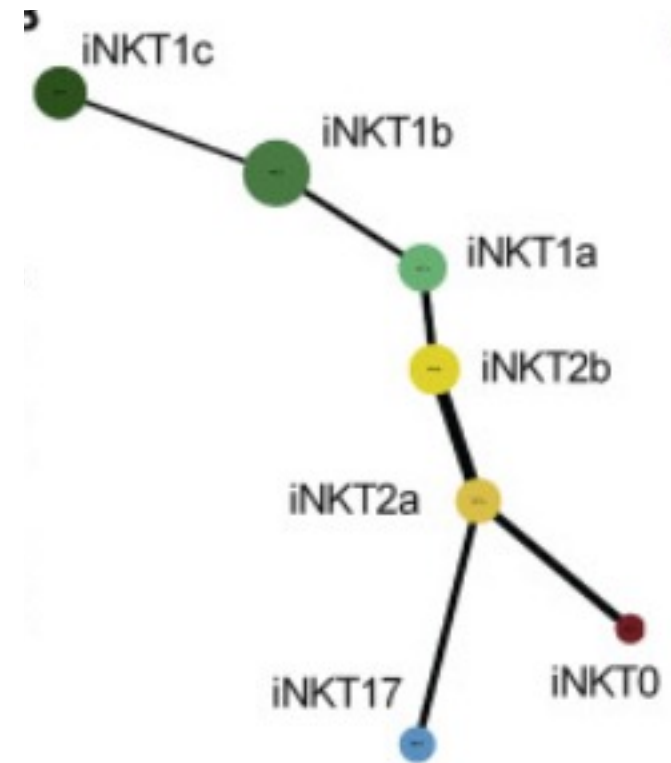
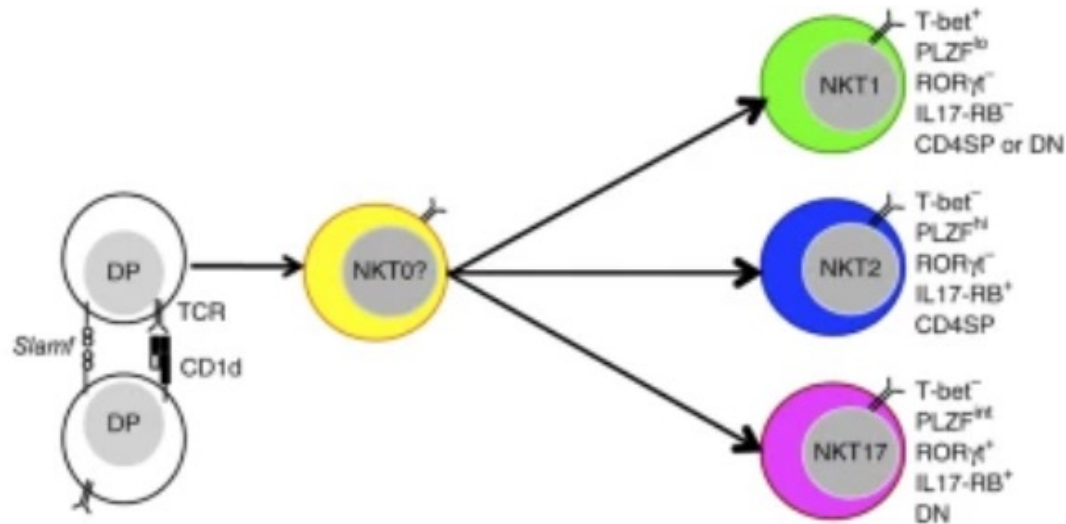
[Nature Immunology](#) **17**, 728–739 (2016) | [Cite this article](#)

9376 Accesses | **181** Citations | **14** Altmetric | [Metrics](#)

Introduction to iNKT cells

- Invariant natural killer T (iNKT) cells are considered to be an innate-like T lymphocyte population that can initiate or inhibit immune responses
- In the thymus, iNKT cells differentiate into three effector cell subsets, iNKT1, iNKT2, and iNKT17

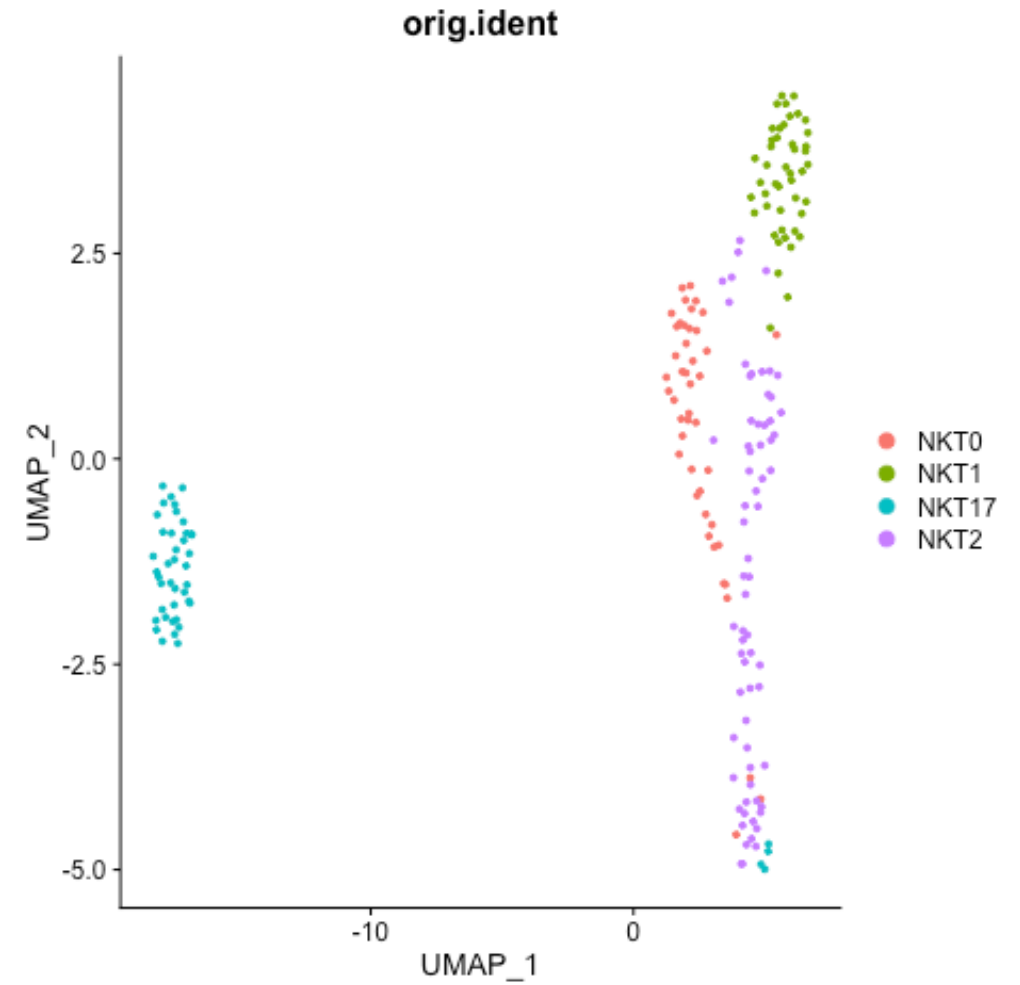
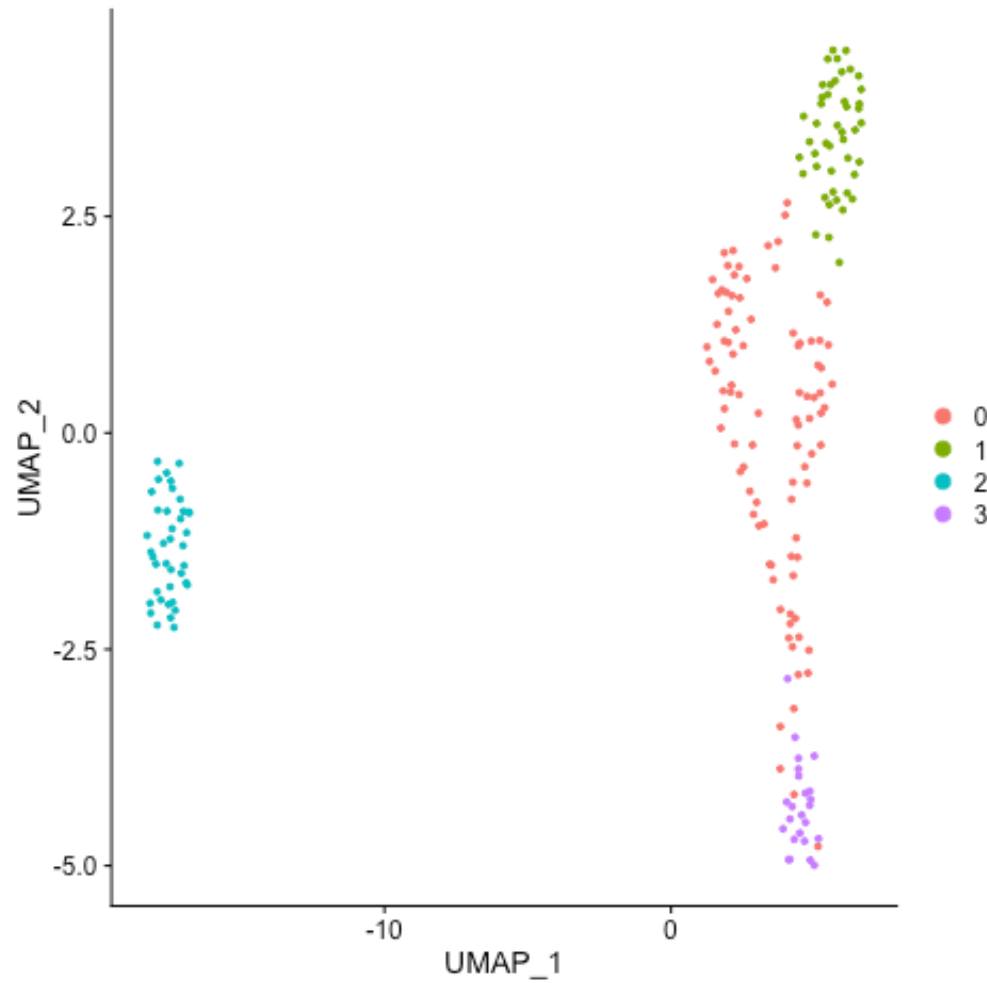
a



Reading in data and preprocessing

- Input data is raw counts
- Preprocessing done with Seurat
 - Totem requires normalised data
- If true cell identities were not known, annotation would be performed
- Run code until line 27
- How does the clustering look like?

Clustering with Seurat

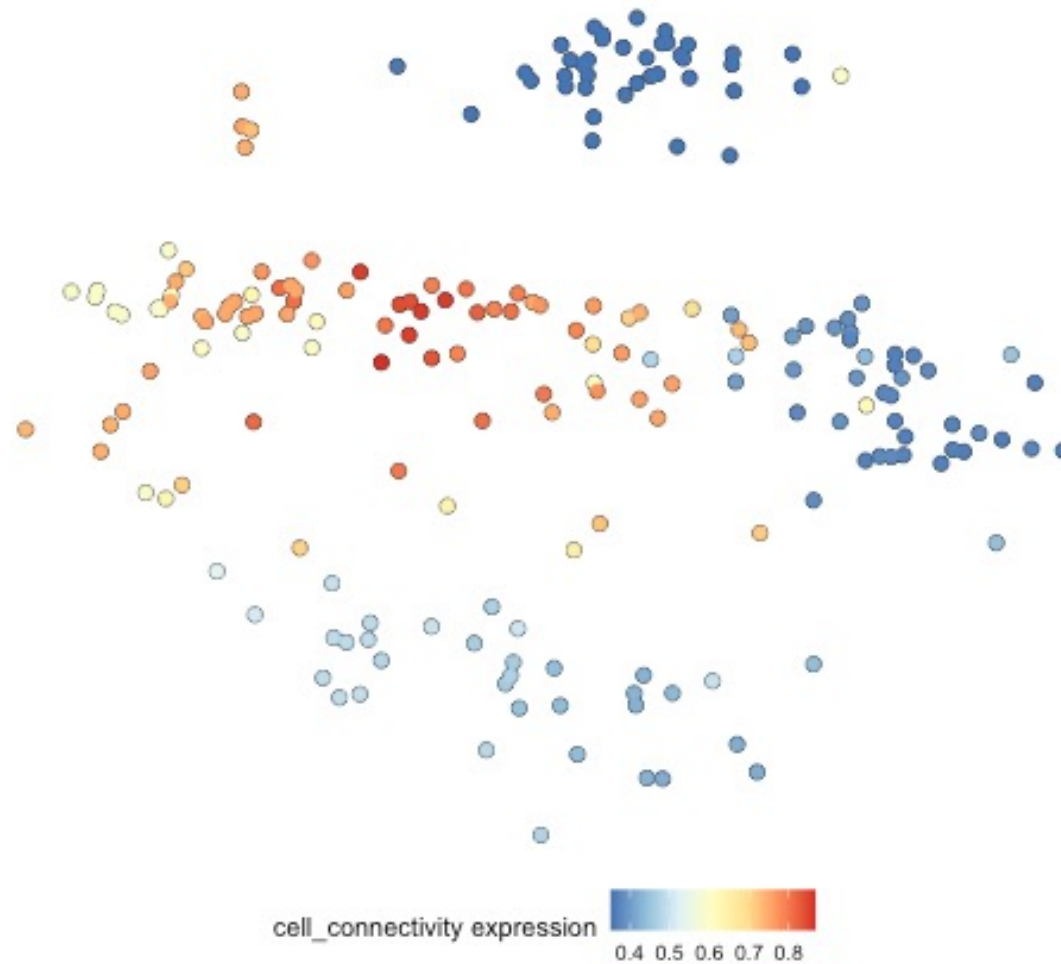


Totem – Dimensionality reduction and cell connectivity

- Required input is raw and normalised counts
- Default method for DR is LMDS
 - Other options are PCA, MDS, UMAP and t-SNE
- For large datasets, feature selection is highly recommended
- Selected features can be specified using the `dim.red.features` parameter in the `RunDimRed()` function

- Run code until line 48
- How does the cell connectivity look like?

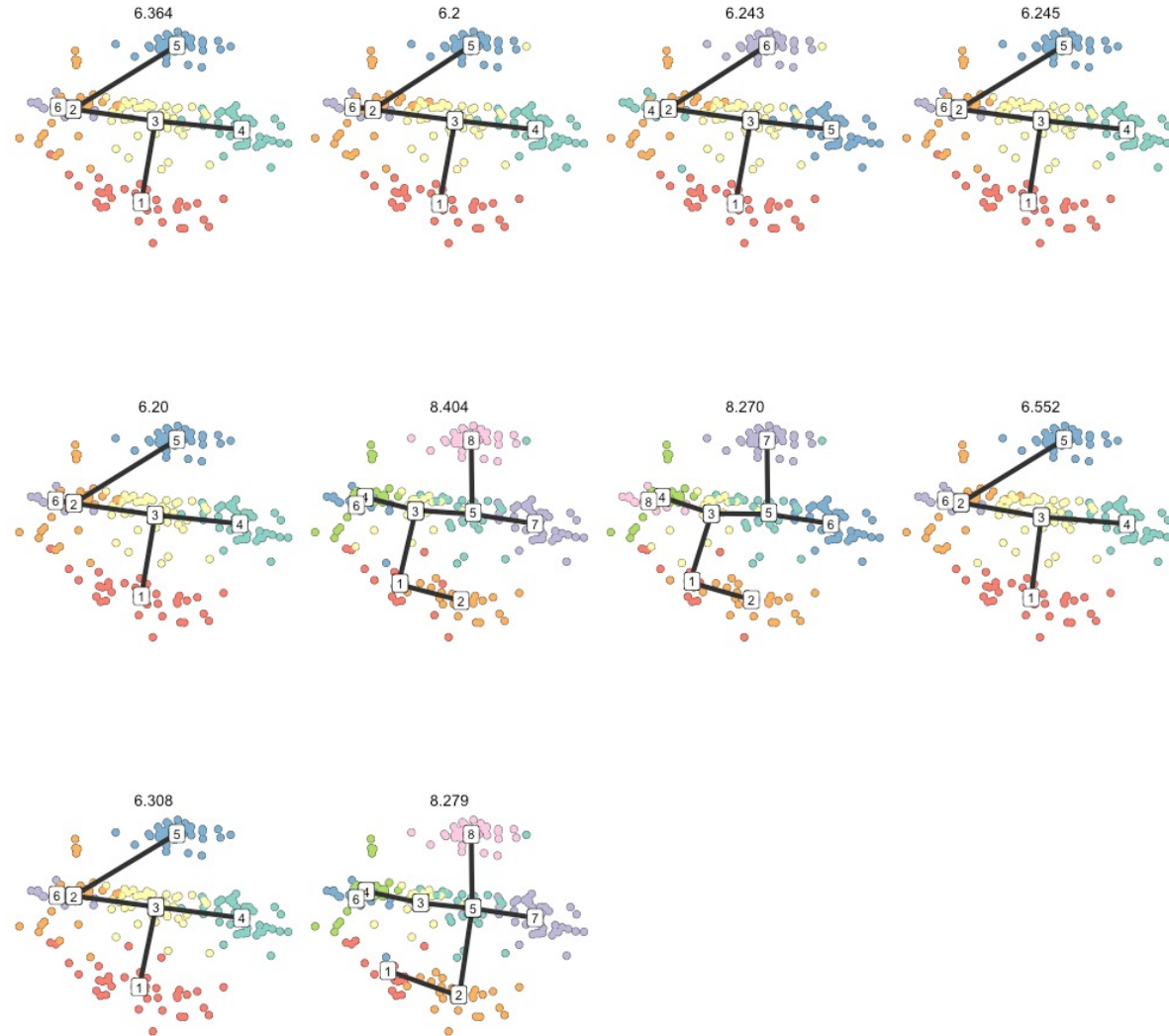
Dimensionality reduction and cell connectivity



Clustering selection and visualisation

- User can select how many clusterings to visualise and the method for selecting them
 - Five selection methods in Totem
- Run code until line 56
- Are there many differences between the clusterings?

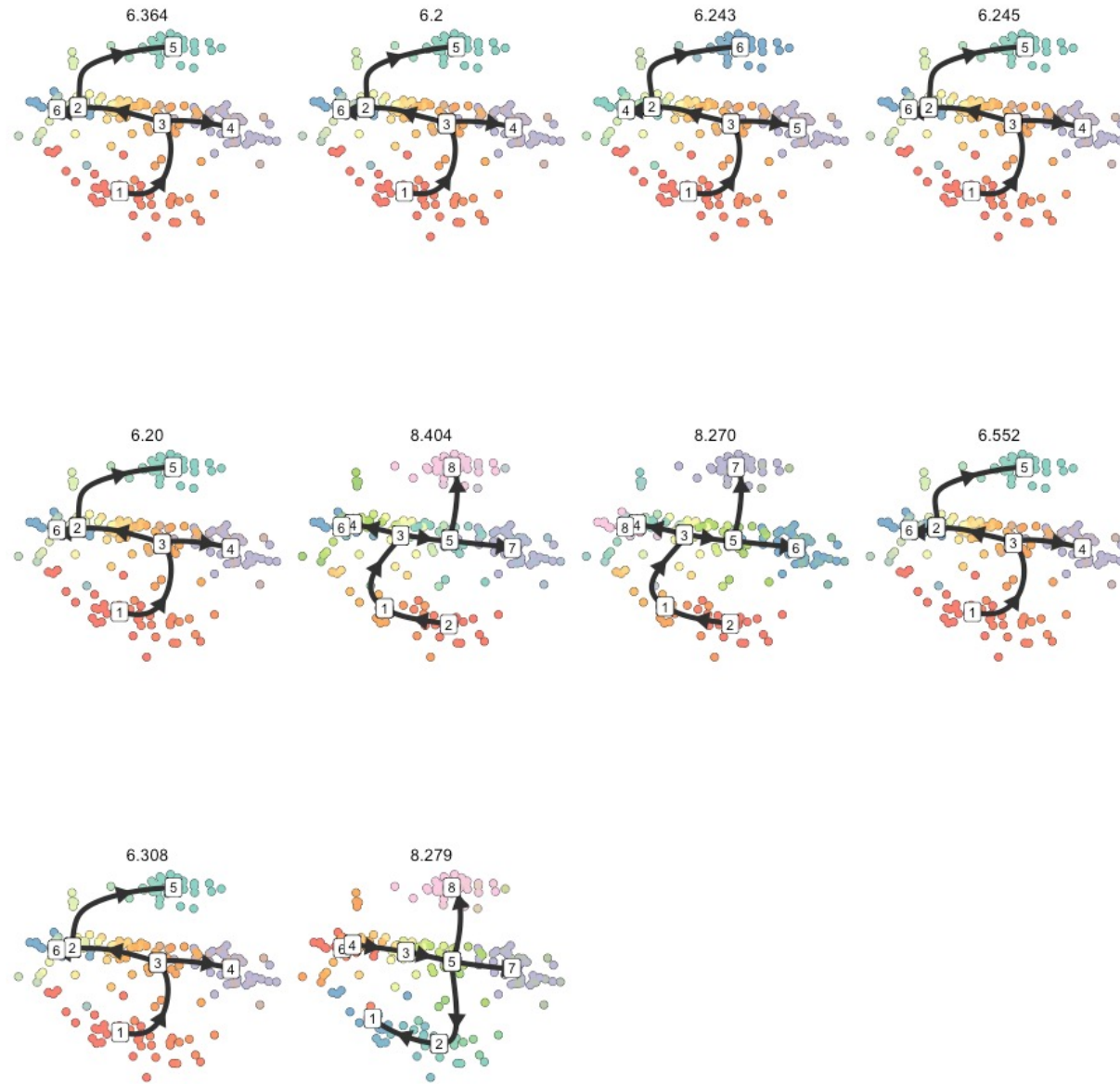
Clustering selection and visualisation



Smoothing

- Smoothing generates the pseudotime and a directed trajectory for the selected clusterings
- Run code until line 64
- How do the directions of the clusterings vary?

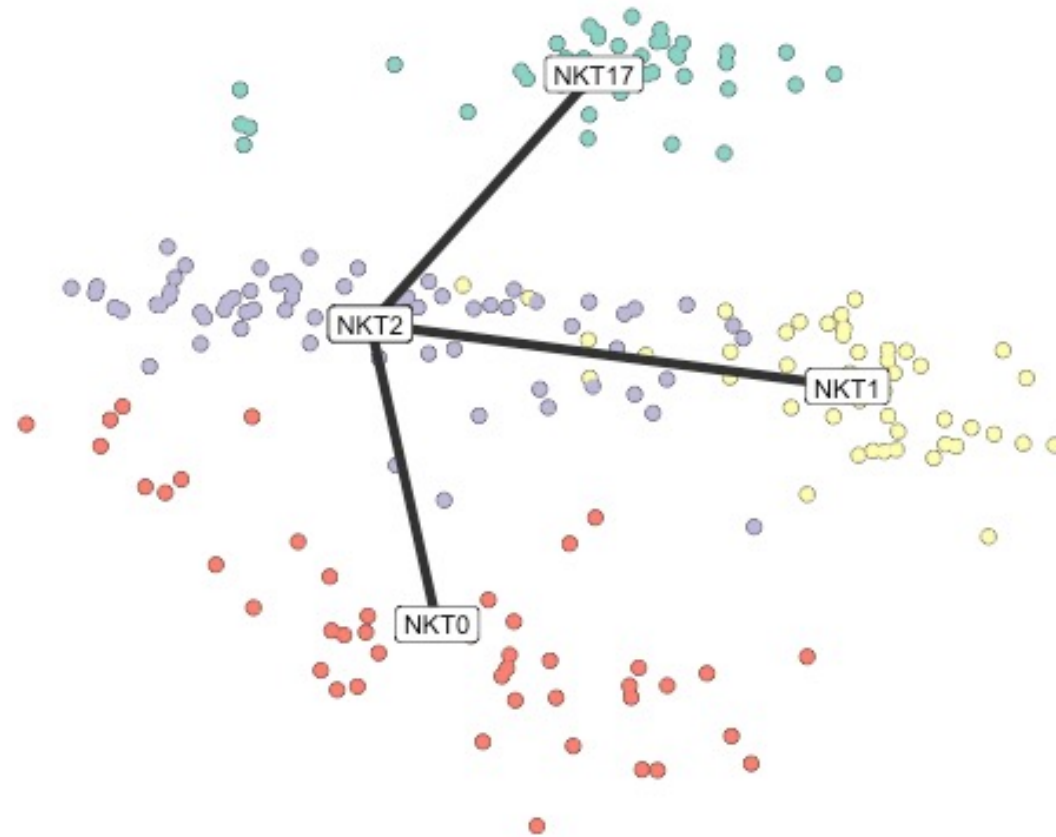
Smoothing



Visualising the cell identities

- Visualising the known cell identities will help in choosing the trajectory
- Run code until line 71
- Based on the cell identities and cluster directions, which trajectory will you choose?

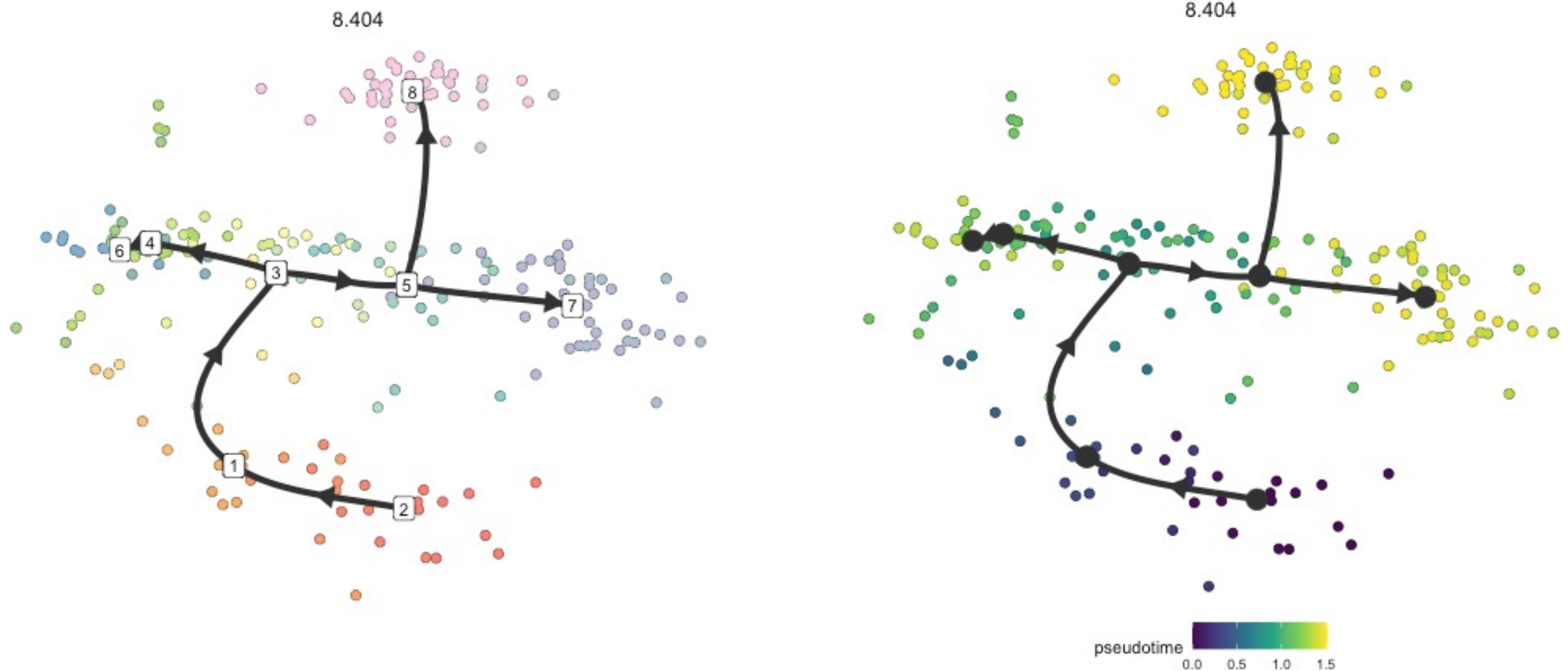
Visualising the cell identities



Visualising the selected trajectory

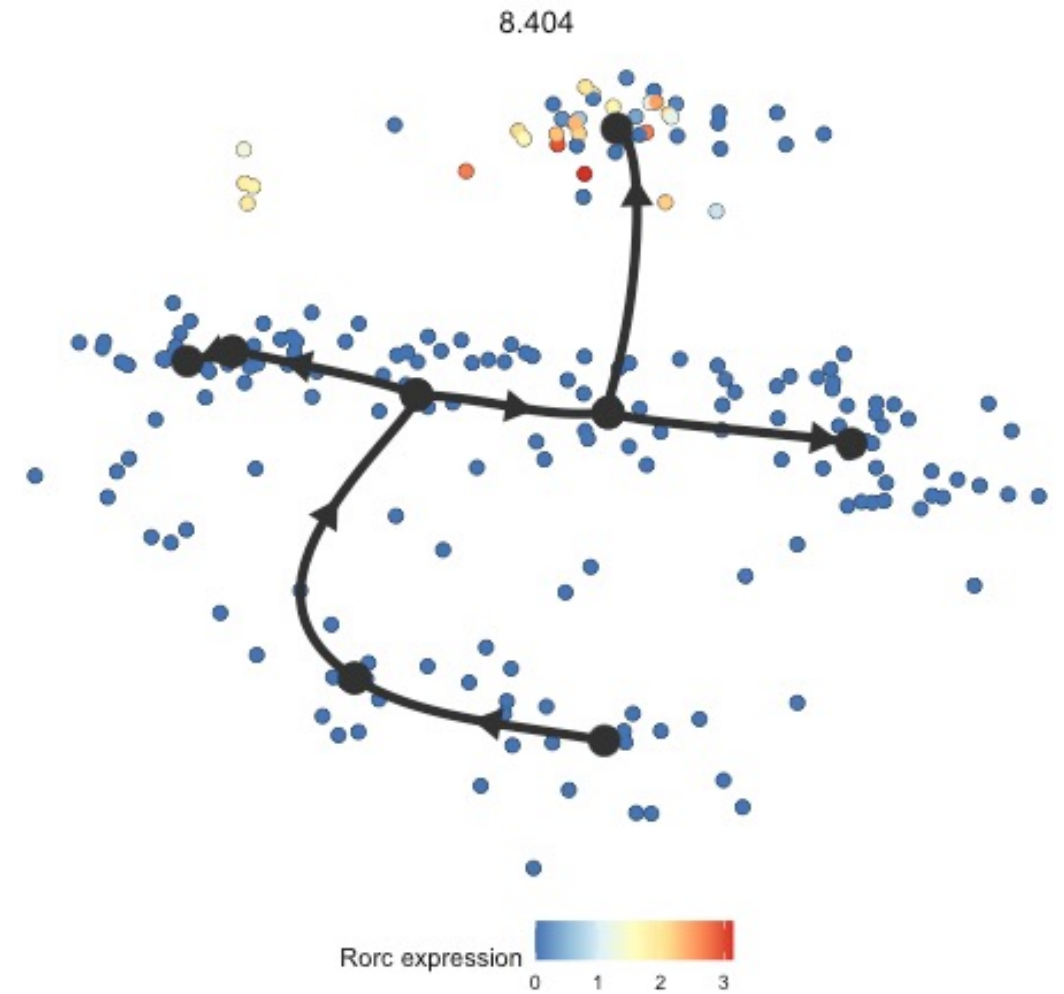
- Pseudotime can also be visualised for the trajectory
- The root cluster can be changed, if needed
- Run code until line 85
- Select one trajectory to use
- What do you think, is changing the root cluster necessary?

Visualising the selected trajectory



Visualising gene expression

- Run code until line 101
- Visualise the expression of Rorc gene
- Is Rorc expression localised to iNKT17 cells?



- Based on Totem trajectory inference, which hypothesis is more accurate?

a

