SAHLGRENSKA ACADEMY
CORE FACILITIES

UNIVERSITY OF GOTHENBURG

Benchmarking of 5 tools designed for pseudotime analysis of scRNA-seq data

Vanja Börjesson

Bioinformatics Core Facility, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

Introduction

Cells differentiate and specialize to achieve different functions within the organism. Studies on how cells change over time, also called pseudotime analysis, provide researchers with better insight into the cell process and the cause for different phenotypes. Single cell RNA sequencing (scRNA-seq) methods and analysis pipelines have gradually been improved and several bioinformatic tools for pseudotime analysis exist. However, the performance of these tools vary and it can be difficult to decide which tool to use.

We have benchmarked five tools for pseudotime analysis with respect to memory and time efficiency, the ability to control filtering, normalization, the look of the output figures and if the tool is easy to install and use.

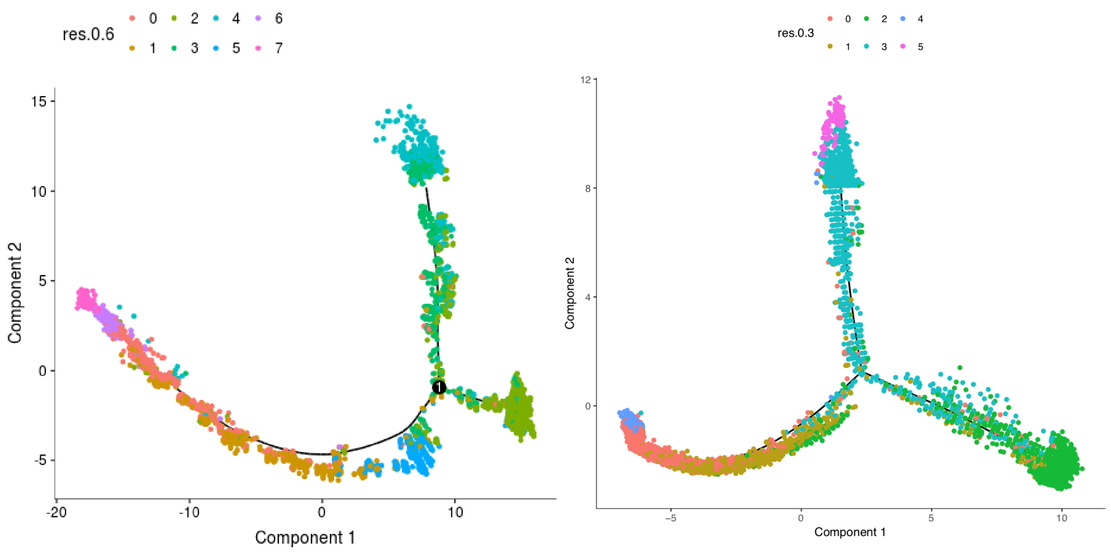
The method

For this benchmarking we used an inhouse drop-seq scRNA-seq dataset[1]. The five tools benchmarked were; 1) Destiny[2], 2) Monocle2[3], 3) Monocle3, 4) Slicer[4] and 5) Scanpy[5].

Mapping and counting was performed using CellRanger, filtering and normalizing using Seurat and clustering was done using tSNE. Then we imported the counts to run these five tools with default settings to predict the pseudotime.

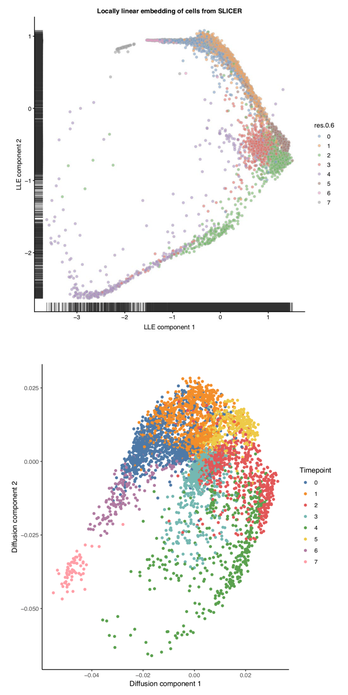
Monocle2 and Monocle3

Monocle2 and the newest version beta 3 give the user many alternatives for pre processing the data, but also to visualize the output.



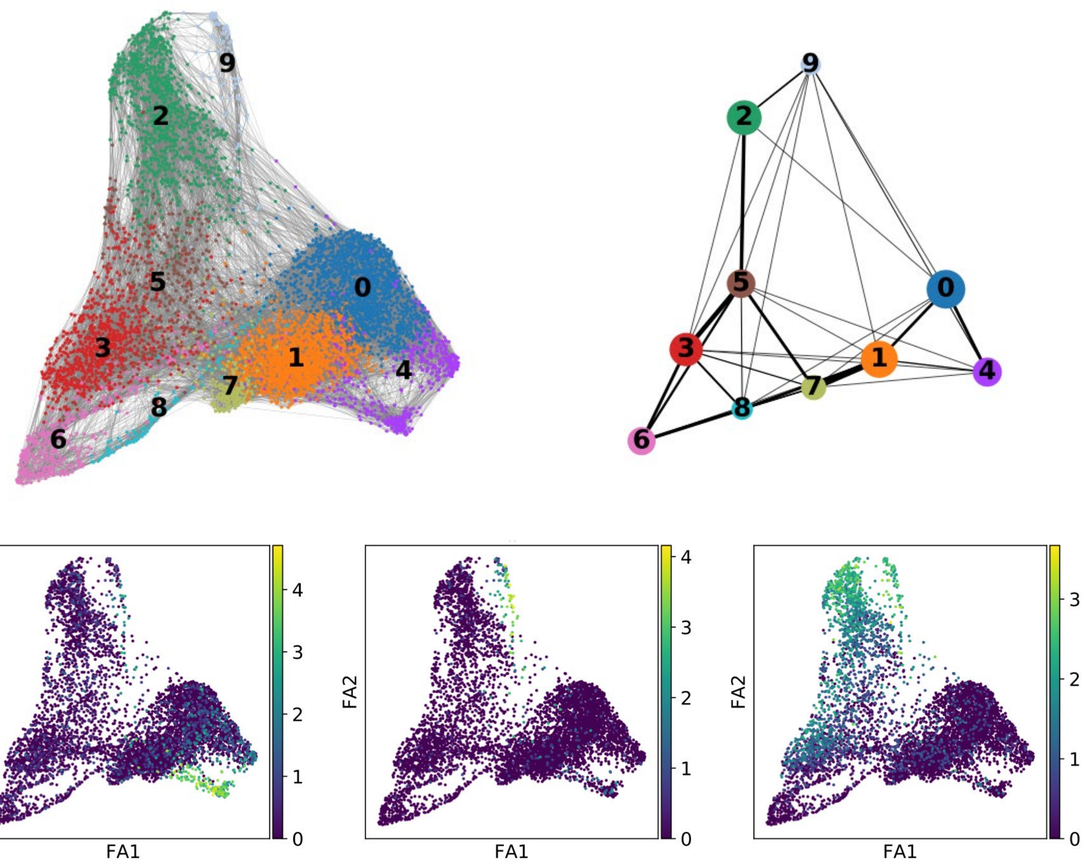
Slicer and Destiny

Slicer and Destiny were easily installed and executed in R. Slicer was the tool taking longest time to run. Running on a high performance computer using 4 cores, the analysis was completed after 24 hours. The other tools used in this comparison took about 5 to 30 minutes to run. Neither Slicer nor Destiny provided functions to calculate differentially expressed genes, functions to plot genes or groups of interest.



Scanpy

A time efficient tool that provides all functions needed for pre-processing data and visualizing the results. Written in python.



Results

All five tools predict the same trajectory for our dataset. All tools were easy to install and use. Slicer was by far the most time consuming tool. Monocle and Scanpy implement different algorithms for trajectory development, filtering and normalizing the data and have many options for visualizing differentially expressed genes.

This project is financially supported by the Swedish Foundation for Strategic Research

Contact: vanja.borjesson@gu.se

References

1. Malin Johansson, Department of Medical Biochemistry and Cell biology, University of Gothenburg, Sweden
2. Philipp, A., et al. ”*destiny*: diffusion maps for large-scale single-cell data in R”. Bioinformatics (2016)
3. Cole Trapnell and Davide Cacchiarelli et al. ”The dynamics and regulators of cell fate decisions are revealed by pseudo-temporal ordering of single cells. Nature Biotechnology (2014)
4. Welch, J. D., et al. ”SLICER: inferring branched, nonlinear cellular trajectories from single cell RNA-seq data.” Genome Biology (2016)
5. Wolf, F. and Angerer, P. ”SCANPY: Large-scale single–cell gene expression data analysis”. Genome Biology (2018)