DE with SLICER

Group Members

- Justin Landis
- Victor Adediwura

Abstract

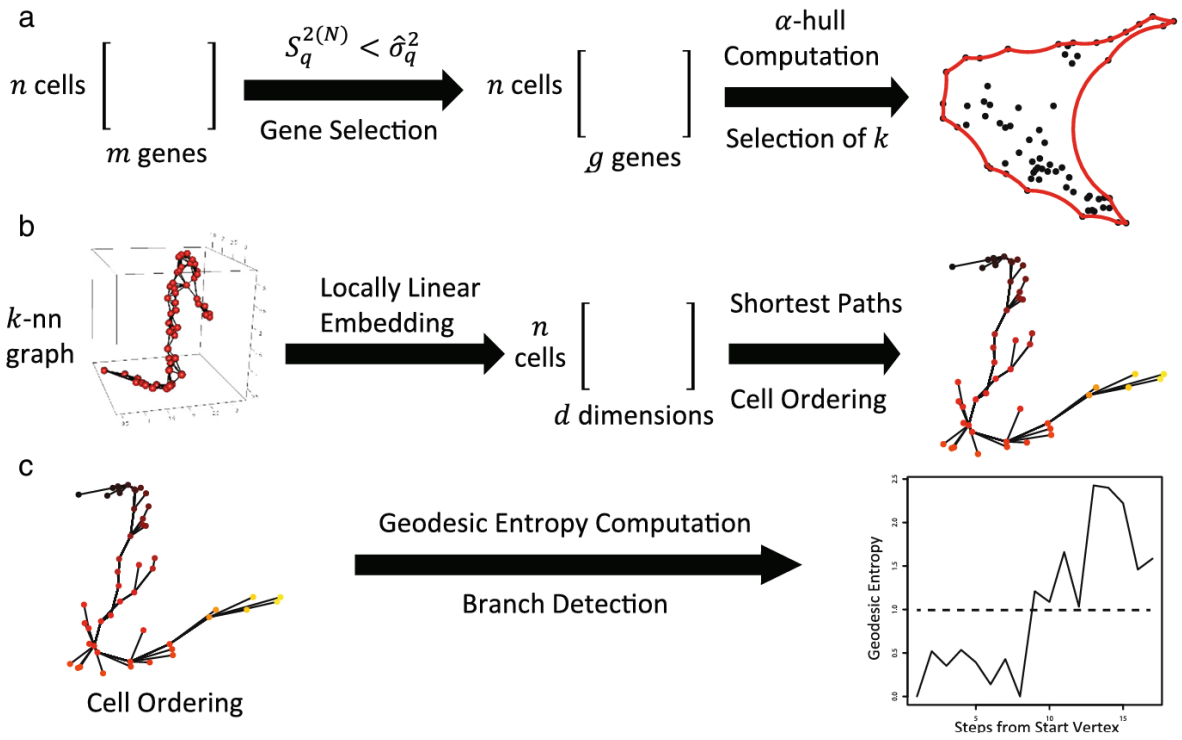
SLICER is a method to select features (genes) to build a trajectory of cells. In Single Cell RNAseq (scRNAseq), this method may be helpful in the context of cell differentiation analyses. The goal of our project is to investigate Differential Expression (DE) approaches to the features selected by SLICER.

Formal Statement of the Problem

While SLICER automatically selects genes that are important for defining a trajectory among the data, it does not associate which features are most important to defined cell types.

Related Work

- Directly related works are `SLICER` (Welch, Hartemink, and Prins 2016)
- `DESeq2` (Love, Huber, and Anders 2014)



Contributions

The overall goal of our work is to identify if performing DE based on branches assignments by SLICER will lead to biologically significant results.

Datasets

We will be using data sets from Single-cell datasets for temporal gene expression integration, specifically utilizing a few Hematopoiesis differentiation dataset (as there are 2).

- (Nestorowa et al. 2016)
- (Olsson et al. 2016); (Olsson et al. 2019)

Intended Experiments

As a validation, we will perform Differential Expression analysis on input data prior to SLICER and then compare the results of the same DE pipeline, except only on features selected by SLICER. In practicality, this requires a dataset with some experimental labels which will be used for differential expression comparisons.

Intended Experiments Continued

Validation Results:

- DE on ALL Features, comparing against original experimental labels

Experiments:

- DE on SLICER Features, comparing against original experimental labels
- DE on SLICER Features, comparing against SLICER branch labels
- DE on ALL Features, comparing against SLICER branch labels

In experiments where SLICER Features are used, we intend to do a set comparison between DE genes in validation versus the experiment.

In experiments where SLICER branch labels are used, we NMI to assess if branches correspond to experimental labels (they may not!)

Expected Challenges

Immediate challenges will be the disparity between softwares. Data is stored in a `h5ad` format that can be read into memory via `scanpy.read_h5ad(...)`. However `SLICER` is implemented in R and will need to be locally installed as it was removed from CRAN in `2022`.

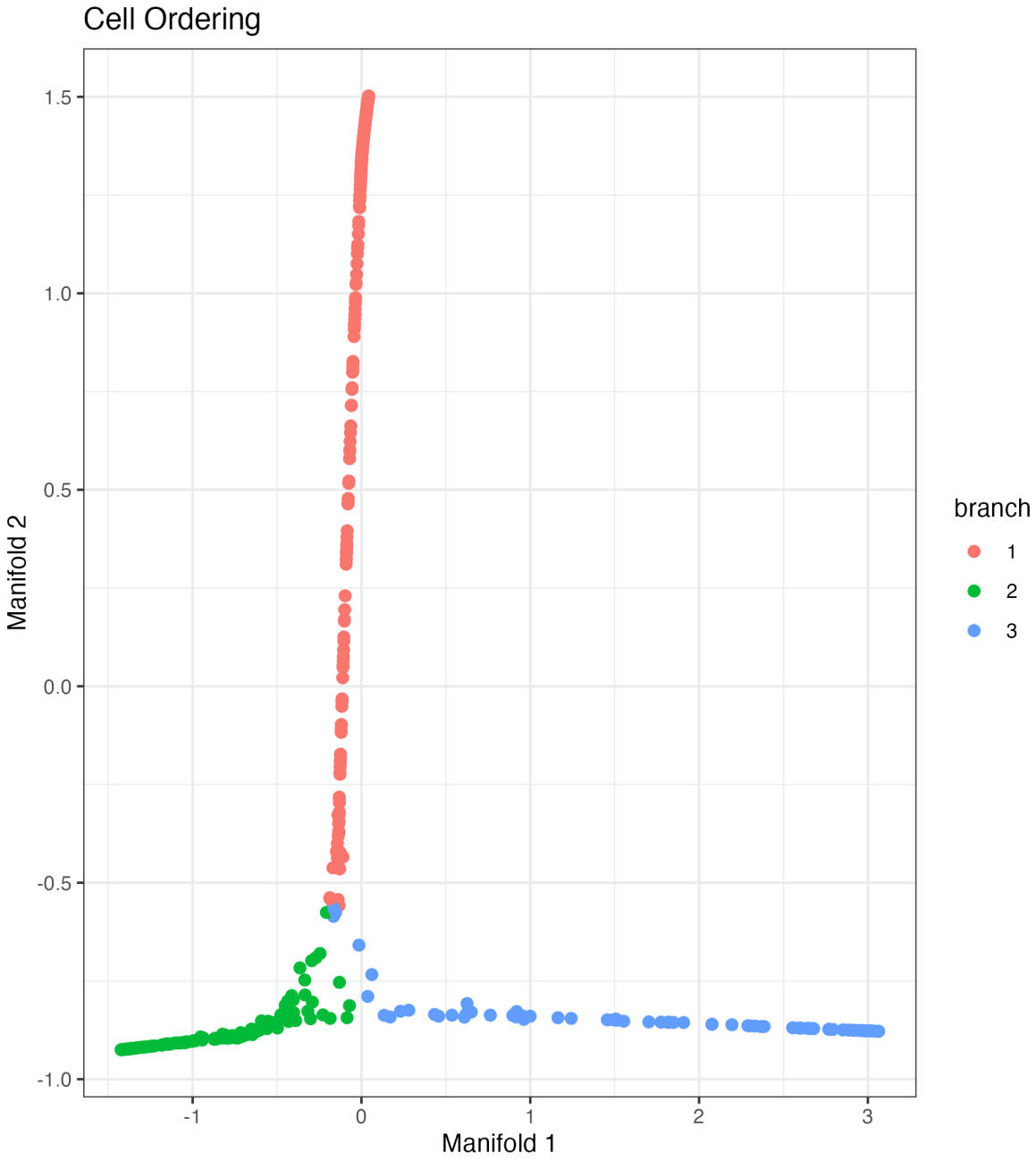
Furthermore, branch assignments seem to be based on the Dimensionality Reduction of `LLE`, but the actual trajectories through the graph may bounce between branch assignments (see Preliminary Results). Additionally, there is no guarantee to the size of branch assignments given by SLICER. Assuming one branch is sufficiently small, this may lead to under powered DE results. DE results may not be comparable to SLICER results due to the nature of SLICER feature selection (selecting genes with low neighborhood variance versus global variance).

Implementation

Since `SLICER` is implemented in R, we will be implementing our DE in R as well. We will ideally provide an R function that takes a cell by gene matrix and returns SLICER results along with supplemental data relating to DE analysis. Our code will be posted on GitHub.

Preliminary Results

Our preliminary results at the moment just involve running `SLICER`'s workflow on their own toy dataset. Please enjoy the following gif revealing cells along `SLICER`'s defined trajectory.



References

Love, Michael I., Wolfgang Huber, and Simon Anders. 2014. "Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2" 15: 550. https://doi.org/10.1186/s13059-014-0550-8.

Nestorowa, Sonia, Fiona K. Hamey, Blanca Pijuan Sala, Evangelia Diamanti, Mairi Shepherd, Elisa Laurenti, Nicola K. Wilson, David G. Kent, and Berthold Göttgens. 2016. "A Single-Cell Resolution Map of Mouse Hematopoietic Stem and Progenitor Cell Differentiation." *Blood* 128 (8): e20–31. https://doi.org/10.1182/blood-2016-05-716480.

Olsson, Andre, Meenakshi Venkatasubramanian, Viren K. Chaudhri, Bruce J. Aronow, Nathan Salomonis, Harinder Singh, and H. Leighton Grimes. 2016. "Single-Cell Analysis of Mixed-Lineage States Leading to a Binary Cell Fate Choice." *Nature* 537 (7622): 698–702. https://doi.org/10.1038/nature19348.

Olsson, Andre, Meenakshi Venkatasubramanian, Virendra K. Chaudhri, Bruce J. Aronow, Nathan Salomonis, Harinder Singh, and H. Leighton Grimes. 2019. "Author Correction: Single-Cell Analysis of Mixed-Lineage States Leading to a Binary Cell Fate Choice." *Nature* 569 (7755): E3–3. https://doi.org/10.1038/s41586-019-1107-5.

Welch, Joshua D., Alexander J. Hartemink, and Jan F. Prins. 2016. "SLICER: Inferring Branched, Nonlinear Cellular Trajectories from Single Cell RNA-Seq Data." *Genome Biology* 17 (1). https://doi.org/10.1186/s13059-016-0975-3.