Pseudotime informed GRN inference

Single-cell sequencing RNASeq estimates the abundance of RNA in individual cells. Frameworks such as Scanpy and Seurat have been used with great success to analyse the differences between cell types and differences between health and disease. Gene regulatory networks are graphical representations of the relationships between transcription factors, proteins that bind the DNA and induce or repress transcription, and other genes. These networks are inferred using tools like SCENIC+ or GRNBoost2 (arboreto) or other methods. As we can capture the RNA of single cells, we can now better understand developmental processes on the transcriptional level. During development and cell differentiation the transcriptional pattern in the cells changes, leading to their unique cell type identity. Using diverse tools we can compute so called trajectories, where the cells are ordered by their inferred pseudotime. Usually, GRN inference is performed on a single data set with one timepoint. This approach has a weakness, as the transcriptional activation of genes is delayed compared to the activity of the transcription factors. This project aims to overcome this weakness by first ordering the cells on a pseudotime trajectory before then predicting the gene expression of the cells in cells at timepoint t1 based on the transcriptional activity of cells at timepoint t0.

The project will proceed in the following steps.

- Recommended reading: Molecular Biology of the cell
- Familiarisation with single cell analysis, specifically pseudotime analysis.
- Standard trajectory inference on a cell lineage differentiation dataset.
- Standard GRN inference (cell type based, or whole dataset)
- Novel data analysis: GRN analysis of pseudotime aggregated cells.
- Comparison of the standard GRNs with the new GRNs using statistical measures (predictive error, …) and biological interpretation tools (GSEA, biological interpretation)

Requirements:

- Python
- Familiarity with using UNIX environments and usage of the command line.
- Interest in biological questions and willingness to independently read the relevant chapters in an undergrad molecular biology textbook.

May be extended to thesis project, if approach seems promising