Gene regulatory network inference using nf-core and nextflow for community driven open-source analysis pipelines
Supervisor: Nicolai Meyerhöfer

There are many tools developed for gene regulatory network (GRN) inference, a central task in the analysis of single-cell RNA sequencing (scRNA-seq) data. However, almost all of these live as stand-alone python or R packages. This has multiple drawbacks. First, this makes it difficult for researchers to know which tool is best to use. Second, stand-alone python and R packages always demand extra effort to install and set-up because a lot of packages are poorly documented and not maintained. Lastly, stand-alone python and R packages are often not built for large-scale analysis of scRNA-seq data, which becomes more and more prominent with the increasing amount of data that is available.
There has been a recent push in the nf-core community to create open source pipelines for the scalable analysis of scRNA-seq data using nextflow in order to solve these problems.

This project has the goal of getting to know nextflow, a widely used programming language to create reproducible and scalable scientific workflows, and nf-core, a community driven effort to create open-source analysis pipelines using nextflow. For this, the aim of the project is to integrate new GRN inference tools into the already existing state-of-the-art nf-core pipeline *scdownstream*. The project is structured as follows:
- Learning about GRN inference
- Learning about nextflow and nf-core
- Learning about the nf-core pipeline *scdownstream*
- Choice of GRN inference tools to integrate
- Integrating the GRN inference tools into the *scdownstream* development branch
- Completing all necessary steps for publishing the GRN inference tools to the public *scdownstream branch* (these include tasks such as writing documentation, writing tests and getting in touch with the maintainers of the *scdownstream* pipeline)

Extension to a Master's thesis project possible.