Making the most of your data: Building a single-cell RNA-seq pipeline

Malte Lücken

Single-cell RNA-seq has enabled gene expression to be studied at an unprecedented resolution. The promise of this technology is attracting a growing user base for single-cell analysis methods. As more analysis tools are becoming available, it is becoming increasingly difficult to navigate this landscape and produce an up-to-date workflow to analyse one's data. In this talk, I introduce the steps of a typical single-cell RNA-seq analysis, including pre-processing (quality control, normalization, data correction, feature selection, and dimensionality reduction) and cell- and gene-level downstream analysis. We will explore some tool choices for these steps and elaborate how tool choice can affect the biological interpretation of transcriptomic data. Finally, we will go over the current best-practices for single-cell RNA-seq analysis based on independent comparison studies that we formulated in our recent molecular systems biology paper, and introduce our best-practices analysis pipeline that is available at `https://www.github.com/theislab/single-cell-tutorial`. This talk is intended to serve as a workflow tutorial for new entrants into the field, and help established users update their analysis pipelines.