

Intro to Single-cell Genomics Handout

Research Informatics Solutions
Minnesota Supercomputing Institute
2022-04-28

General Resources

MSI Tutorial - Introduction to Linux:

<https://www.msi.umn.edu/tutorials/introduction-linux>

This tutorial addresses the basics of interacting with a Linux system via the command line. Single-cell analyses are computationally intense, and are best run on a managed HPC system (rather than a personal laptop). MSI servers are Linux-based, so understanding the basics of interacting with a Linux system is important.

MSI Tutorial - Job Submission and Scheduling:

<https://www.msi.umn.edu/tutorials/job-submission-and-scheduling-msi>

This tutorial will cover how to submit both interactive and non-interactive (batch) jobs to MSI's high-performance compute clusters. Single-cell genomics analyses generally require computing resources that exceed the capacity of personal workstations, and so should be run on a high-performance compute system.

UC Riverside Manual - Programming in R:

<http://manuals.bioinformatics.ucr.edu/home/programming-in-r>

This manual page gives working examples and explanations of R programming. Many single-cell genomics analysis tools require knowledge of R.

MSI Tutorial - Python for Scientific Computing:

<https://www.msi.umn.edu/tutorials/python-scientific-computing>

This tutorial addresses techniques in Python for high-performance computing, data visualization, and scientific analyses. Many single-cell genomics analysis tools require Python knowledge.

MSI Tutorial - Software Installation and Management at MSI:

<https://www.msi.umn.edu/tutorials/software-installation-and-management-msi> (Registration)

<https://pages.github.umn.edu/dunn0404/software-management-tutorial/> (Tutorial materials)

This tutorial addresses installation and management of custom software modules and Anaconda environments. Single-cell genomics analysis tools can have complicated and easily-broken networks of software dependencies, so knowing how to manage software environments is very helpful.

Experimental Design and QA/QC Resources

10X Genomics Single Cell ATAC Knowledgebase:

<https://kb.10xgenomics.com/hc/en-us/categories/360001072491>

10X Genomics Single Cell Gene Expression Knowledgebase:

<https://kb.10xgenomics.com/hc/en-us/categories/360000149952-Single-Cell-3-Gene-Expression>

10X Genomics Single Cell Multi-ome (Expression+ATAC) Knowledgebase:

<https://kb.10xgenomics.com/hc/en-us/categories/360004142131-Single-Cell-Multiome-ATAC-Gene-Expression>

Shallow sequencing run as library QC:

<https://kb.10xgenomics.com/hc/en-us/articles/360054613831-Can-I-perform-shallow-sequencing-to-assess-the-quality-of-Single-Cell-3-Gene-Expression-libraries->

This article from 10X Genomics discusses the accuracy of various QC metrics when performing shallow sequencing (e.g., MiSeq or iSeq) of a single-cell library.

SCOPIT:

https://alexdavisscs.shinyapps.io/scs_power_multinomial/

This Shiny app will help you estimate the number of cells required for a successful single-cell experiment, assuming a certain abundance of a rare cell type. It can also perform a retrospective power calculation for a dataset that has already been collected.

scPower:

<http://scpower.helmholtz-muenchen.de/>

This app will help design a multi-sample single-cell transcriptomics experiment. It has models for cell type detection and differential expression.

Single-cell Analysis Resources

Seurat - single cell RNA-seq analysis (vignette):

https://satijalab.org/seurat/articles/pbmc3k_tutorial.html

This tutorial is a step-by-step guide demonstrating clustering, marker gene identification, and cell type labeling of PBMCs with Seurat (R).

Seurat - single cell RNA-seq analysis with HTOs (vignette):

https://satijalab.org/seurat/articles/hashing_vignette.html

This tutorial is a step-by-step guide demonstrating how to handle hashed (multiplexed) cells with Seurat (R).

Clustree (vignette):

<https://cran.r-project.org/web/packages/clustree/vignettes/clustree.html>

This R vignette demonstrates how to use the Clustree package to assist in choosing a “resolution” value for single-cell clustering.

SingleR - Automated Cell Type Annotation Tool (vignette):

<https://bioconductor.org/packages/release/bioc/vignettes/SingleR/inst/doc/SingleR.html>

This R vignette demonstrates how to use SingleR to use a reference gene expression dataset to label cell types in a single-cell dataset.

cellDex Bioconductor package:

<https://bioconductor.org/packages/release/data/experiment/html/cellDex.html>

This package contains reference datasets for cell type labeling of single-cell transcriptomic data. The reference datasets are for both human and mouse and are derived from high-quality (sorted) bulk gene expression data.

Seurat - Integration (vignette):

https://satijalab.org/seurat/articles/integration_introduction.html

This tutorial demonstrates how to integrate multiple (non-hashed) single-cell datasets with Seurat.

MSI Video - Getting Started with Vision:

<https://www.youtube.com/watch?v=jWOebw4VF1Y>

This video shows a step-by-step guide to using Vision (<https://yoseflab.github.io/VISION/>) for interactively visualizing single-cell RNAseq data. This process assumes familiarity with R.

10X Cell Ranger ATAC:

<https://support.10xgenomics.com/single-cell-atac/software/pipelines/latest/what-is-cell-ranger-atac>

Seurat - Integration of RNA and ATAC (vignette):

https://satijalab.org/seurat/articles/atacseq_integration_vignette.html

10X Cell Ranger V(D)J:

<https://support.10xgenomics.com/single-cell-vdj/software/pipelines/latest/using/vdj>

Review of Trajectory Inference Tools (article):

<https://www.nature.com/articles/s41587-019-0071-9>

RNA velocity paper (article):

<https://www.nature.com/articles/s41586-018-0414-6>

RNA velocity:

<http://velocityto.org>

inferCNV:

<https://data.humancellatlas.org/analyze/methods/infer>

Catalog of scRNA-seq tools (website):

<https://www.scrna-tools.org>