Statistics for Genomics: scRNAseq

May 08, 2025

Table of Contents

# About this Book

This book is intended to introduce concepts in single cell RNA-seq analysis (scRNA-seq).

## Skills Level

*Genetics*

*Programming skills*

## GDSCN Collection

This exercise is part of a collection of teaching resources developed through the *Genomic Data Science Community Network* (GDSCN). GDSCN works towards a vision where researchers, educators, and students from diverse backgrounds are able to fully participate in genomic data science research. Learn more about GDSCN by visiting <https://www.gdscn.org/home> or reading the [article in Genome Research](https://doi.org/10.1101/gr.276496.121).

Please check out our full collection of AnVIL and related resources: <https://hutchdatascience.org/AnVIL_Collection/>

# 1 Student Guide

## 1.1 Activity One

Many of the AnVIL books that are designed for teaching a class or workshop contain information aimed at the instructor.

You might want to create a “student guide” that contains a different subset of Rmd files from your book, or renders to a different output format (e.g. word document). You can specify the output and Rmd files that will be used for the student guide using the \_output.yml and \_bookdown.yml files in the student-guide directory.

By default, repositories based on AnVIL\_Template create a docx file containing this chapter as an example. Once you have adjusted student-guide/bookdown.yml to include the chapters you want in your student guide, you can link to it by replacing {repo-name} in the following URL:

https://github.com/jhudsl/{repo-name}/raw/main/student-guide/docs/Student\_Guide.docx

When someone clicks this link, they will download the docx file. For example, here is the link for the student guide from AnVIL\_Template:

<https://github.com/jhudsl/AnVIL_Template/raw/main/student-guide/docs/Student_Guide.docx>

## 1.2 Activity Two

Steps of the guide could go here.

# About the Authors

These credits are based on our [course contributors table guidelines](https://github.com/jhudsl/OTTR_Template/wiki/How-to-give-credits).

| Credits | Names |
| --- | --- |
| **Pedagogy** |  |
| Lead Content Instructor | Elizabeth Humphries |
| Content Idea Contributor | [Ava Hoffman](https://www.avahoffman.com/), [Frederick Tan](https://carnegiescience.edu/frederick-tan) |
| Content Directors | Kasper Hansen, [Jeff Leek](https://jtleek.com/), [Frederick Tan](https://carnegiescience.edu/frederick-tan) |
| **Production** |  |
| Content Publisher | [Ira Gooding](https://publichealth.jhu.edu/faculty/4130/ira-gooding) |
| **Technical** |  |
| Template Publishing Engineers | [Candace Savonen](https://www.cansavvy.com/), [Carrie Wright](https://carriewright11.github.io/), [Ava Hoffman](https://www.avahoffman.com/) |
| Publishing Maintenance Engineer | [Candace Savonen](https://www.cansavvy.com/) |
| Technical Publishing Stylists | [Carrie Wright](https://carriewright11.github.io/), [Candace Savonen](https://www.cansavvy.com/) |
| Package Developers ([ottrpal](https://github.com/jhudsl/ottrpal)) | [John Muschelli](https://johnmuschelli.com/), [Candace Savonen](https://www.cansavvy.com/), [Carrie Wright](https://carriewright11.github.io/) |
| **Funding** |  |
| Funder | [National Human Genome Research Institute (NHGRI)](https://www.genome.gov/) |
| Funding Staff | Fallon Bachman, Jennifer Vessio, Emily Voeglein |

## ─ Session info ───────────────────────────────────────────────────────────────  
## setting value  
## version R version 4.1.3 (2022-03-10)  
## os Ubuntu 20.04.5 LTS  
## system x86\_64, linux-gnu  
## ui X11  
## language (EN)  
## collate en\_US.UTF-8  
## ctype en\_US.UTF-8  
## tz Etc/UTC  
## date 2025-05-08  
## pandoc 2.17.1.1 @ /usr/local/bin/ (via rmarkdown)  
##   
## ─ Packages ───────────────────────────────────────────────────────────────────  
## package \* version date (UTC) lib source  
## bookdown 0.24 2023-03-22 [1] Github (rstudio/bookdown@88bc4ea)  
## cachem 1.0.6 2021-08-19 [1] RSPM (R 4.1.0)  
## callr 3.7.0 2021-04-20 [1] RSPM (R 4.1.0)  
## cli 3.2.0 2022-02-14 [1] RSPM (R 4.1.0)  
## crayon 1.5.1 2022-03-26 [1] RSPM (R 4.1.0)  
## desc 1.4.1 2022-03-06 [1] RSPM (R 4.1.0)  
## devtools 2.4.3 2021-11-30 [1] RSPM (R 4.1.0)  
## digest 0.6.29 2021-12-01 [1] RSPM (R 4.1.0)  
## ellipsis 0.3.2 2021-04-29 [1] RSPM (R 4.1.0)  
## evaluate 0.15 2022-02-18 [1] RSPM (R 4.1.0)  
## fastmap 1.1.0 2021-01-25 [1] RSPM (R 4.1.0)  
## fs 1.5.2 2021-12-08 [1] RSPM (R 4.1.0)  
## glue 1.6.2 2022-02-24 [1] RSPM (R 4.1.0)  
## htmltools 0.5.2 2021-08-25 [1] RSPM (R 4.1.0)  
## knitr 1.33 2023-03-22 [1] Github (yihui/knitr@a1052d1)  
## lifecycle 1.0.1 2021-09-24 [1] RSPM (R 4.1.0)  
## magrittr 2.0.3 2022-03-30 [1] RSPM (R 4.1.0)  
## memoise 2.0.1 2021-11-26 [1] RSPM (R 4.1.0)  
## pkgbuild 1.3.1 2021-12-20 [1] RSPM (R 4.1.0)  
## pkgload 1.2.4 2021-11-30 [1] RSPM (R 4.1.0)  
## prettyunits 1.1.1 2020-01-24 [1] RSPM (R 4.1.0)  
## processx 3.5.3 2022-03-25 [1] RSPM (R 4.1.0)  
## ps 1.6.0 2021-02-28 [1] RSPM (R 4.1.0)  
## purrr 0.3.4 2020-04-17 [1] RSPM (R 4.1.0)  
## R6 2.5.1 2021-08-19 [1] RSPM (R 4.1.0)  
## remotes 2.4.2 2021-11-30 [1] RSPM (R 4.1.0)  
## rlang 1.0.2 2023-03-22 [1] Github (r-lib/rlang@34b04a8)  
## rmarkdown 2.10 2023-03-22 [1] Github (rstudio/rmarkdown@02d3c25)  
## rprojroot 2.0.3 2022-04-02 [1] CRAN (R 4.1.3)  
## sessioninfo 1.2.2 2021-12-06 [1] RSPM (R 4.1.0)  
## stringi 1.7.6 2021-11-29 [1] RSPM (R 4.1.0)  
## stringr 1.4.0 2019-02-10 [1] RSPM (R 4.1.0)  
## testthat 3.0.1 2023-03-22 [1] Github (R-lib/testthat@e99155a)  
## usethis 2.1.6.9000 2023-03-22 [1] Github (r-lib/usethis@b4b8623)  
## withr 2.5.0 2022-03-03 [1] RSPM (R 4.1.0)  
## xfun 0.26 2023-03-22 [1] Github (yihui/xfun@74c2a66)  
## yaml 2.3.5 2022-02-21 [1] RSPM (R 4.1.0)  
##   
## [1] /usr/local/lib/R/site-library  
## [2] /usr/local/lib/R/library  
##   
## ──────────────────────────────────────────────────────────────────────────────

# 2 References

Tasic B et al. Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nat Neurosci. 2016 Feb;19(2):335-46. doi: 10.1038/nn.4216. Epub 2016 Jan 4. PMID: 26727548; PMCID: PMC4985242.

Van der Maaten L, Hinton G. Visualizing data using t-SNE. J. Mach. Learn. Res. 2008 Nov;9:2579-2605.

Zeisel A et al. Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science. 2015 Mar 6;347(6226):1138-42. doi: 10.1126/science.aaa1934. Epub 2015 Feb 19. PMID: 25700174.

Zeng H et al. Large-scale cellular-resolution gene profiling in human neocortex reveals species-specific molecular signatures. Cell. 2012 Apr 13;149(2):483-96. doi: 10.1016/j.cell.2012.02.052. PMID: 22500809; PMCID: PMC3328777.