RNA-seq miniCURE

January, 2023

Table of Contents

# About this Course

This miniCURE allows students to develop a project using RNA-seq datasets to introduce how model organisms, high-throughput sequencing, and the scientific process are used in both basic and clinical research.



Skills Level

*Genetics*  
**Novice**: Introduction to [central dogma of molecular biology](https://en.wikipedia.org/wiki/Central_dogma_of_molecular_biology)

*Programming skills*  
**Novice**: No programming experience needed

Learning Objectives

* Learn about model organisms
* Explore high-throughput sequencing datasets
* Practice the [scientific process](https://www.biointeractive.org/classroom-resources/how-science-works)

C-MOOR Collection

More coming soon!

# Onboarding

## 0.1 Join SciServer

*Providing access to big data resources to researchers worldwide*

### 0.1.1 Purpose

In this course we will use the online SciServer platform to do some data analysis for your research project. The purpose of this assignment is to register for a SciServer account, and then to inform the instructor of your username so that you can be added to the SciServer group for this course and access course materials.



### 0.1.2 Learning Objectives

1. Create an account on SciServer
2. Confirm your email address
3. Share your username with your instructor

### 0.1.3 Introduction

SciServer is an online platform for doing scientific data analysis. It is used by scientists studying astronomy, biology, oceanography, and more, and is free as long as you are using it for scientific research. Using SciServer means you do not need a fancy computer or need to install any special programs on your computer, you can just log in with your internet browser to start doing research. For this course, we have set up SciServer with customized collections of programs for RNA-seq analysis, as well as the data that we’ll be analyzing. Once you sign up for SciServer and are added to the group for this course, you will be able to access these tools and begin your data analysis journey!

### 0.1.4 Instructions

#### 0.1.4.1 Create an account on SciServer

This video ([video](https://link.c-moor.org/video-join-sciserver))([slides](https://docs.google.com/presentation/d/1kxbnBLoRsdPW4ZkjwNsAHS1XFPuJpQZ8I1aVqyZISW0)) shows you how to create a SciServer account. You can follow along with the video, or follow the steps below.

1. Open [sciserver.org](https://www.sciserver.org/) in a web browser
   1. It is a good idea to bookmark this page so that you can easily access it throughout the course.
2. Click “Login to SciServer”
3. Click “Create a new account”
4. Enter a User name, Email, etc. and click “Create account”
   1. Note that you cannot change your username

#### 0.1.4.2 Confirm your email address

1. **Important!**: Click the verification link in your email inbox.
   1. If you do not verify your account you will get locked out and will need to contact your instructor to unlock your account.
   2. If you do not see an email, try checking your spam.
2. After clicking the verification link, confirm that your username appears in the upper right hand corner.

##### 0.1.4.2.1 Resources

* [sciserver.org](https://www.sciserver.org/)
* [How to add a bookmark in Chrome](https://support.google.com/chrome/answer/188842)
* [SciServer help page](https://www.sciserver.org/support/how-to-use-sciserver/)

#### 0.1.4.3 Share your username with your instructor

1. Fill out [this form](https://docs.google.com/forms/d/e/1FAIpQLSdJva363KdHIVxI0jWBZzNdhz2M-u8Be3viKiy0Rboyzy4PPQ/viewform) with your SciServer username.

### 0.1.5 Footnotes

#### 0.1.5.1 Contributions and Affiliations

* Katherine Cox, Johns Hopkins University
* Frederick Tan, Carnegie Institution

Last Revised: January 25, 2022

## 0.2 First LearnR

*Interactive tutorials introducing various data science concepts*

### 0.2.1 Purpose

The purpose of this assignment is to (1) join the class SciServer group so you can access course materials, and (2) learn how to access the tutorials for this course on SciServer.



### 0.2.2 Learning Objectives

1. Accept invitation to join class SciServer Group
2. Start up a **C-MOOR LearnR** compute container
3. Complete your first LearnR tutorial
4. Delete your C-MOOR LearnR compute container

### 0.2.3 Introduction

Before beginning this assignment, you should have already created a SciServer account and submitted your SciServer username to your instructor. In this assignment you will learn how to set up a “compute container” on SciServer. Compute containers are how you use programs on SciServer. In this course you will use two compute containers: “C-MOOR LearnR” has tutorials that will teach you how to run data analyses, and “C-MOOR R-Studio” is where you will work on your own data analysis projects. This assignment shows you how to set up the C-MOOR LearnR compute container and start up your first tutorial.

### 0.2.4 Instructions

#### 0.2.4.1 Accept invitation to join class SciServer group

This video ([video](https://link.c-moor.org/video-join-sciserver-group))([slides](https://docs.google.com/presentation/d/1codot9UeUO7l0EDcEre7dJgyXurD_xyxpw6IJL_aEjM)) shows you how to join a SciServer group. You can follow along with the video, or follow the steps below.

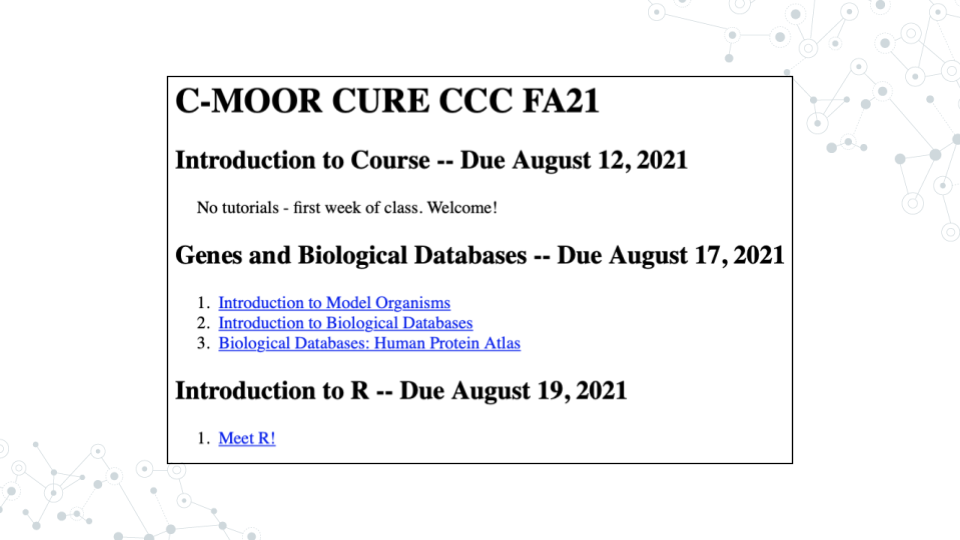
1. Open [sciserver.org](https://www.sciserver.org/) in a web browser and log in to your account.
2. Click “Groups”
3. On the left, you should see a list of all the groups you have joined or been invited to. Click on the name of the group for this course, then click “Accept invitation”.
   1. Your instructor must have your username to invite you to the group. If you do not see an invitation, contact your instructor with your SciServer username.
4. Confirm that you can access course data
   1. On the top menu bar, click “Files”
   2. On the left-hand menu, click “Data Volumes”
   3. Confirm that you see “C-MOOR-Data”
5. Confirm that you can access course computing resources
   1. Click “Home” in the top menu to return to the home page.
   2. Scroll down to the second set of boxes and click “Compute”
   3. Click “Create container”
   4. In the “Compute Image” drop-down menu, confirm that you can see “C-MOOR LearnR” and “C-MOOR R-Studio”
   5. Under “Data Volumes”, confirm that you can see “C-MOOR Data”
   6. You can close the Create Container dialog box (by clicking the “X” in the top right) once you’ve confirmed that you can see the C-MOOR content

##### 0.2.4.1.1 Resources

* [sciserver.org](https://www.sciserver.org/)
* [Get help with SciServer on the C-MOOR Discourse](https://help.c-moor.org/c/help/)

#### 0.2.4.2 Start up a “C-MOOR LearnR” compute container

This video ([video](https://link.c-moor.org/video-sciserver-create-learnr-container))([slides](https://docs.google.com/presentation/d/1Oaq8RzhaDANxkNh-tTKwme7e095pGgoiq5iZHbt7PLg)) shows you how to create and start up a C-MOOR LearnR compute container. You can follow along with the video, or follow the steps below.

1. Open [sciserver.org](https://www.sciserver.org/) in a web browser and log in to your account.
   1. If you are already logged in, click “Home” in the top menu to return to the home page.
2. Scroll down to the second set of boxes and click “Compute”
3. Click “Create container”
   1. Give your container a name. This can be anything you like, but it’s useful if it says something about the purpose of the container so you can tell your containers apart. You could name this container “Tutorials”, since you’ll be using it to access tutorials.
   2. In the “**Compute Image**” drop-down menu, select “**C-MOOR LearnR**”
   3. Under “**Data Volumes**”, check the box next to “**C-MOOR Data**”
   4. Click “Create”. This may take a moment.
4. You should now see a new entry in your list of containers
   1. “Created At” should be a few moments ago.
   2. “Name” should be the name you chose
   3. “Image” should be “C-MOOR LearnR”
5. Start your C-MOOR LearnR container by clicking on its name (whatever name you chose when you created it). This will open in a new tab.
   1. You should see a list of tutorials, organized by date. 
   2. If instead you see an error message, you most likely forgot to check the box next to “C-MOOR Data” when you created the container. 
   3. If you see something else, you may have picked the wrong “Compute Image” from the drop-down menu.

**If anything goes wrong, you can always delete your container by clicking the red “X” in the last column, and create a new container.**

##### 0.2.4.2.1 Resources

* [sciserver.org](https://www.sciserver.org/)
* [Get help with SciServer on the C-MOOR Discourse](https://help.c-moor.org/c/help/)

#### 0.2.4.3 Complete your first LearnR tutorial

1. If you’re not there already, go to the SciServer compute page and start up the C-MOOR LearnR container.
   1. Open [sciserver.org](https://www.sciserver.org/) in a web browser and log in to your account.
   2. If you are already logged in, click “Home” in the top menu to return to the home page.
   3. Scroll down to the second set of boxes and click “Compute”.
   4. Start your C-MOOR LearnR container by clicking on its name.
2. Click on “Biological Databases”. The tutorial will open in a new tab.
3. Complete the tutorial.

##### 0.2.4.3.1 Resources

* [sciserver.org](https://www.sciserver.org/)
* [Get help with SciServer on the C-MOOR Discourse](https://help.c-moor.org/c/help/)

#### 0.2.4.4 Delete your C-MOOR LearnR compute container

Compute containers are meant to be temporary, and you can only have 3 containers total on SciServer. So it’s generally a good idea to clean up after yourself and delete your containers when you’re done using them. Also, in this course, we will be updating the tutorials on the C-MOOR LearnR container, and **you will need to create a new container to get the latest updates.**

**Deleting your container will delete your progress in a tutorial**, so don’t delete the container until you have completed the tutorial and submitted any required items to your instructor. Later on in the course you will learn how to save things permanently on SciServer.

To delete a container:

1. If you’re not there already, go to the SciServer compute page.
   1. Open [sciserver.org](https://www.sciserver.org/) in a web browser and log in to your account.
   2. If you are already logged in, click “Home” in the top menu to return to the home page.
   3. Scroll down to the second set of boxes and click “Compute”.
   4. Start your C-MOOR LearnR container by clicking on its name.
2. Find the container you want to delete.
3. Click on the red “X” in the last column.

### 0.2.5 Footnotes

#### 0.2.5.1 Contributions and Affiliations

* Katherine Cox, Johns Hopkins University
* Frederick Tan, Carnegie Institution

Last Revised: May 13, 2021

# 1 Scientific Literature

# 2 Model Organisms and Databases

## 2.1 Pre-lab: Model Organisms



Image credit: [Max Westby](http://cubocube.com/dashboard.php?a=1179&b=1228&c=10). Some of the most important genetic model organisms in use today. Clockwise from top left: yeast, fruit fly, arabidopsis, mouse, roundworm, zebrafish. License: [CC ANS 2.5](https://creativecommons.org/licenses/by-nc-sa/2.5/)

### 2.1.1 Purpose

The two tutorials in this pre-lab will familiarize you with the concepts of model organisms, with an emphasis on Drosophila. In Lab 8, this information will help you look up information about Drosophila genes.

### 2.1.2 Learning Objectives

1. Explain the importance of model organisms and identify some of their desirable characteristics.
2. Describe the usefulness of Drosophila as a model system.

### 2.1.3 Introduction

Scientists frequently use a few specific organisms, called “model organisms” for their experiments. This tutorial will introduce you to a few of the most popular model organisms and will discuss why these organisms were chosen and what they are useful for.

The fruit fly (Drosophila melanogaster) is a popular model organism used to study a wide range of biological questions. The second tutorial will introduce you to some of the types of research being conducted with fruit flies, give a brief overview of Drosophila biology, and show you what it’s like to work with Drosophila in the lab.

### 2.1.4 Activity 1 - Model Organisms

Estimated time: 15 min

#### 2.1.4.1 Instructions

1. [Click here to open the Model Organisms Tutorial.](https://clovis.shinyapps.io/BIOL11A_Model_Organisms)
2. To move through the activities click “Continue” at the bottom of the screen. When you are done with a topic, click “Next Topic” to move on.
3. As you complete the tutorial, answer the questions below.

#### 2.1.4.2 Questions

| 1A. Explain what a “model organism” is and why they are useful. |
| --- |
|  |

| 1B. Define ortholog and explain how model organisms can be used to understand human genes. |
| --- |
|  |

| 1C. Name 4 commonly used model organisms |
| --- |
|  |

### 2.1.5 Activity 2 - Drosophila melanogaster

Estimated time: 15 min

#### 2.1.5.1 Instructions

1. [Click here to open the Drosophila melanogaster Tutorial.](https://clovis.shinyapps.io/BIOL11A_Drosophila/)
2. To move through the activities click “Continue” at the bottom of the screen. When you are done with a topic, click “Next Topic” to move on.
3. As you complete the tutorial, answer the questions below.

#### 2.1.5.2 Questions

| 2A. Provide 3 reasons why fruit flies are useful for scientific research. |
| --- |
|  |

| 2B. List 3 ways in which fruit flies are similar to humans. |
| --- |
|  |

| 2C. Compare and contrast the fruit fly genome to the human genome. |
| --- |
|  |

| 2D. Briefly describe the fruit fly life cycle. |
| --- |
|  |

### 2.1.6 Footnotes

#### 2.1.6.1 Resources

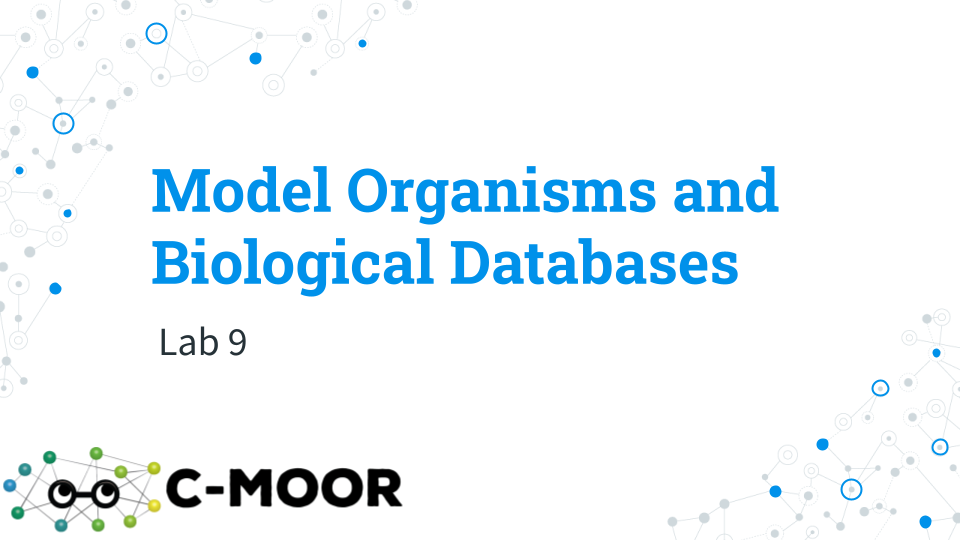
* [Google Doc](https://docs.google.com/document/d/1vFhm2XLMO9vjMDNT6CxZJ4VJuinw0BEo)

#### 2.1.6.2 Contributions and Affiliations

* Stephanie R. Coffman, Ph.D.

Last Revised: February 2022

## 2.2 Lab Slides



[[slides](https://docs.google.com/presentation/d/1kt0lW4D8AWqQm1j6FMo0rRDGM_G3zkLmPy3EZNDg3s0)]

## 2.3 Lab Activities: Biological Databases

### 2.3.1 Purpose

In this activity, students will learn to search the same online databases used by scientists to collect information about a set of genes and present them to your group. This will give you an opportunity to engage in inquiry-based learning and apply the concepts in molecular biology and genetics from this course.

### 2.3.2 Learning Objectives

1. Use online databases to look up information about a gene.

### 2.3.3 Activity 1 - Databases

Estimated time: 15 min

#### 2.3.3.1 Instructions

1. Follow [these directions](https://docs.google.com/document/d/1nQ-wd4hX0Xtd4mfloXcpbbtK_CqDzY874oqmBYql1M8) to launch your first LearnR tutorial: Biological Databases Tutorial.
2. To move through the activities click “Continue” at the bottom of the screen. When you are done with a topic, click “Next Topic” to move on.
3. As you complete the tutorial, fill in the table below. This will help you know which database to go back to later on.

#### 2.3.3.2 Questions

Table 1. Databases

| Database | Description |
| --- | --- |
| GenBank |  |
| OMIM |  |
| Human Protein Atlas |  |
| PDB |  |

### 2.3.4 Activity 2 - FlyBase

Estimated time: 15 min

#### 2.3.4.1 Instructions

1. Work through the FlyBase Tutorial in SciServer.
2. To move through the activities click “Continue” at the bottom of the screen. When you are done with a topic, click “Next Topic” to move on.

#### 2.3.4.2 Questions

| 2A. What is one question you have about using FlyBase? |
| --- |
|  |

| 2B. What is something that surprised you or that you found interesting about using FlyBase? |
| --- |
|  |

### 2.3.5 Activity 3 - Human Protein Atlas

Estimated time: 15 min

#### 2.3.5.1 Instructions

1. Work through the Human Protein Atlas in SciServer.
2. To move through the activities click “Continue” at the bottom of the screen. When you are done with a topic, click “Next Topic” to move on.

#### 2.3.5.2 Questions

| 3A. What is one question you have about using HPA? |
| --- |
|  |

| 3B. What is something that surprised you or that you found interesting about using HPA? |
| --- |
|  |

### 2.3.6 Activity 4 - Research a Gene!

Estimated time: 45 min

#### 2.3.6.1 Instructions

1. Before getting started on this activity, your instructor will assign your group a letter that corresponds to a group of 4 genes.

| Group Assigned Letter |
| --- |
|  |

1. [Look up your letter here](https://docs.google.com/spreadsheets/d/1GZtHz2GU3B4KMOd9yuKrcQLJJLn7qlnZgPN5c6SNqGg) and write the names of the four genes your group is assigned at the top of each column in the table below.
2. In your group, assign each student one of the four genes to research.

| Individual Assigned Gene |
| --- |
|  |

1. Use FlyBase to look up the information in Table 2 below.
2. Use HPA to look up the information in Table 3 below.

Table 2. FlyBase Information

| Category | Information |
| --- | --- |
| **General Information** |  |
| Full Gene Name |  |
| FlyBase ID |  |
| Sequence Location |  |
| **Function** |  |
| Biological Process |  |
| Cellular Component |  |
| **Expression Data** |  |
| Anatomical Expression |  |
| Developmental Stage |  |
| **Orthologs** |  |
| Orthologs in other species |  |
| Human Orthologs |  |

Table 3. Human Protein Atlas

| Category | Information |
| --- | --- |
| Function |  |
| Is the gene tissue specific? Which tissue? |  |
| Where is it localized in cells? |  |

### 2.3.7 Activity 5 - Present to your Group

Estimated time: 15 min

#### 2.3.7.1 Instructions

1. Present your gene to your group.
2. Take turns presenting your genes among your group and decide on one gene that you think is the most interesting.

### 2.3.8 Activity 6 - Class Presentation

Estimated time: 30 min

#### 2.3.8.1 Instructions

1. With your group, create a short presentation to present your chosen gene to the class.
2. Your presentation should have about four slides and be thorough:
   1. Slide 1: The GENE you picked to share with your group, your name and date
   2. Slide 2 - 4: Present the information you collected about the gene. For full credit, include relevant images/ diagrams on your slides.
3. One student in the group should post your slides on Canvas to the discussion board. Make sure you mention everyone in your group by name so they also get credit for the presentation.

### 2.3.9 Footnotes

#### 2.3.9.1 Resources

* [Google Doc](https://docs.google.com/document/d/1M7rtDzyGVUxO2GVBp09bTBu4fMDxocTs)

#### 2.3.9.2 Contributions and Affiliations

* Rosa Alcazar, Ph.D., Clovis Community College
* Katherine Cox, Ph.D., John Hopkins
* Stephanie R. Coffman, Ph.D., Clovis Community College

Last Revised: September 2021

# 3 RNA-seq Analysis

# 4 Differential Gene Expression

# 5 Project Work

# 6 Scientific Communication

# 7 Example miniCURE Projects

Read more about what these students did and how you can help

Zellweger Spectrum Disorder

[Drosophila Melanogaster a Good Model System of Zellweger Spectrum Disorder BIO11A SP2022](https://help.c-moor.org/t/326)



Look at This!

Explore other miniCURE and CURE projects in our [Look at This!](https://help.c-moor.org/c/look-at-this/8) category

# 8 C-MOOR Scholars

Meet the C-MOOR Scholars and learn how you can support them

* <https://www.cloviscollege.edu/alumni-and-community/c-moor/c-moor-scholars.html>

# 9 Online Community

Join the discussion at <https://help.c-moor.org>

# 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(s) | [FirstName LastName](link%20to%20personal%20website) |
| Lecturer(s) (include chapter name/link in parentheses if only for specific chapters) - make new line if more than one chapter involved | Delivered the course in some way - video or audio |
| Content Author(s) (include chapter name/link in parentheses if only for specific chapters) - make new line if more than one chapter involved | If any other authors besides lead instructor |
| Content Contributor(s) (include section name/link in parentheses) - make new line if more than one section involved | Wrote less than a chapter |
| Content Editor(s)/Reviewer(s) | Checked your content |
| Content Director(s) | Helped guide the content direction |
| Content Consultants (include chapter name/link in parentheses or word “General”) - make new line if more than one chapter involved | Gave high level advice on content |
| Acknowledgments | Gave small assistance to content but not to the level of consulting |
| **Production** |  |
| Content Publisher(s) | Helped with publishing platform |
| Content Publishing Reviewer(s) | Reviewed overall content and aesthetics on publishing platform |
| **Technical** |  |
| Course Publishing Engineer(s) | Helped with the code for the technical aspects related to the specific course generation |
| Template Publishing Engineers | [Candace Savonen](https://www.cansavvy.com/), [Carrie Wright](https://carriewright11.github.io/) |
| 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)) [Candace Savonen](https://www.cansavvy.com/), [John Muschelli](https://johnmuschelli.com/), [Carrie Wright](https://carriewright11.github.io/) |  |
| **Art and Design** |  |
| Illustrator(s) | Created graphics for the course |
| Figure Artist(s) | Created figures/plots for course |
| Videographer(s) | Filmed videos |
| Videography Editor(s) | Edited film |
| Audiographer(s) | Recorded audio |
| Audiography Editor(s) | Edited audio recordings |
| **Funding** |  |
| Funder(s) | Institution/individual who funded course including grant number |
| Funding Staff | Staff members who help with funding |

## ─ Session info ───────────────────────────────────────────────────────────────  
## setting value   
## version R version 4.0.2 (2020-06-22)  
## os Ubuntu 20.04.3 LTS   
## system x86\_64, linux-gnu   
## ui X11   
## language (EN)   
## collate en\_US.UTF-8   
## ctype en\_US.UTF-8   
## tz Etc/UTC   
## date 2023-01-20   
##   
## ─ Packages ───────────────────────────────────────────────────────────────────  
## package \* version date lib source   
## assertthat 0.2.1 2019-03-21 [1] RSPM (R 4.0.3)   
## bookdown 0.24 2022-02-15 [1] Github (rstudio/bookdown@88bc4ea)   
## callr 3.4.4 2020-09-07 [1] RSPM (R 4.0.2)   
## cli 2.0.2 2020-02-28 [1] RSPM (R 4.0.0)   
## crayon 1.3.4 2017-09-16 [1] RSPM (R 4.0.0)   
## desc 1.2.0 2018-05-01 [1] RSPM (R 4.0.3)   
## devtools 2.3.2 2020-09-18 [1] RSPM (R 4.0.3)   
## digest 0.6.25 2020-02-23 [1] RSPM (R 4.0.0)   
## ellipsis 0.3.1 2020-05-15 [1] RSPM (R 4.0.3)   
## evaluate 0.14 2019-05-28 [1] RSPM (R 4.0.3)   
## fansi 0.4.1 2020-01-08 [1] RSPM (R 4.0.0)   
## fs 1.5.0 2020-07-31 [1] RSPM (R 4.0.3)   
## glue 1.6.1 2022-01-22 [1] CRAN (R 4.0.2)   
## htmltools 0.5.0 2020-06-16 [1] RSPM (R 4.0.1)   
## knitr 1.33 2022-02-15 [1] Github (yihui/knitr@a1052d1)   
## lifecycle 1.0.0 2021-02-15 [1] CRAN (R 4.0.2)   
## magrittr 2.0.2 2022-01-26 [1] CRAN (R 4.0.2)   
## memoise 1.1.0 2017-04-21 [1] RSPM (R 4.0.0)   
## pkgbuild 1.1.0 2020-07-13 [1] RSPM (R 4.0.2)   
## pkgload 1.1.0 2020-05-29 [1] RSPM (R 4.0.3)   
## prettyunits 1.1.1 2020-01-24 [1] RSPM (R 4.0.3)   
## processx 3.4.4 2020-09-03 [1] RSPM (R 4.0.2)   
## ps 1.3.4 2020-08-11 [1] RSPM (R 4.0.2)   
## purrr 0.3.4 2020-04-17 [1] RSPM (R 4.0.3)   
## R6 2.4.1 2019-11-12 [1] RSPM (R 4.0.0)   
## remotes 2.2.0 2020-07-21 [1] RSPM (R 4.0.3)   
## rlang 0.4.10 2022-02-15 [1] Github (r-lib/rlang@f0c9be5)   
## rmarkdown 2.10 2022-02-15 [1] Github (rstudio/rmarkdown@02d3c25)  
## rprojroot 2.0.2 2020-11-15 [1] CRAN (R 4.0.2)   
## sessioninfo 1.1.1 2018-11-05 [1] RSPM (R 4.0.3)   
## stringi 1.5.3 2020-09-09 [1] RSPM (R 4.0.3)   
## stringr 1.4.0 2019-02-10 [1] RSPM (R 4.0.3)   
## testthat 3.0.1 2022-02-15 [1] Github (R-lib/testthat@e99155a)   
## usethis 2.1.5.9000 2022-02-15 [1] Github (r-lib/usethis@57b109a)   
## withr 2.3.0 2020-09-22 [1] RSPM (R 4.0.2)   
## xfun 0.26 2022-02-15 [1] Github (yihui/xfun@74c2a66)   
## yaml 2.2.1 2020-02-01 [1] RSPM (R 4.0.3)   
##   
## [1] /usr/local/lib/R/site-library  
## [2] /usr/local/lib/R/library