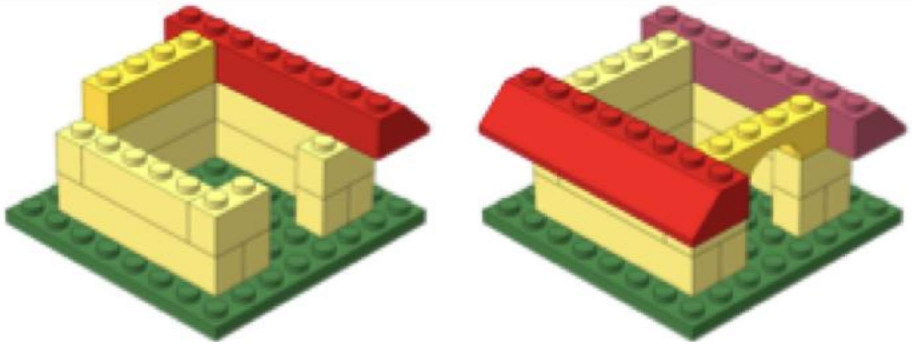
Pseudobulk and related approaches for scRNA-seq analysis

https://tinyurl.com/hbc-pseudobulk-quarto

Harvard Chan Bioinformatics Core

Workshop Scope

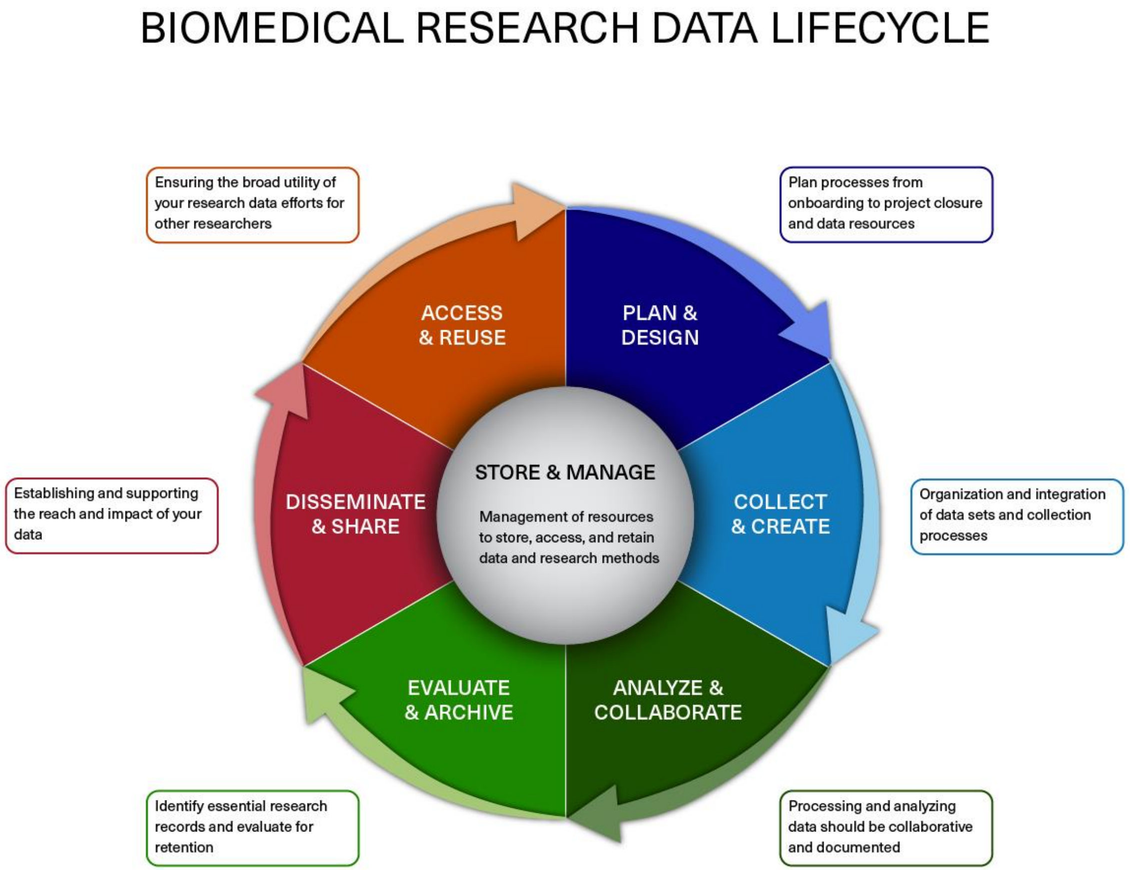


- Understanding considerations for when to use different DGE algorithms on scRNA-seq data
- Using FindMarkers to evaluate significantly DE genes
- Aggregating single cell expression data into a pseudobulk counts matrix to run a DESeq2 workflow
- Evaluating expression patterns of differentially expressed genes at the pseudobulk and single cell level
- Application of methods for evaluating differential proportions of cells between conditions

Exit survey

https://tinyurl.com/hbc-scRNAseq-DGE-exit-survey

Research Data Management (RDM)



BIOMEDICAL RESEARCH DATA LIFECYCLE

STORE & MANAGE

Management of resources to store, access, and retain data and research methods

PLAN & DESIGN

Plan processes from onboarding to project closure and data resources

COLLECT & CREATE

Organization and integration of data sets and collection processes

ANALYZE & COLLABORATE

Processing and analyzing data should be collaborative and documented

EVALUATE & ARCHIVE

Identify essential research records and evaluate for retention

DISSEMINATE & SHARE

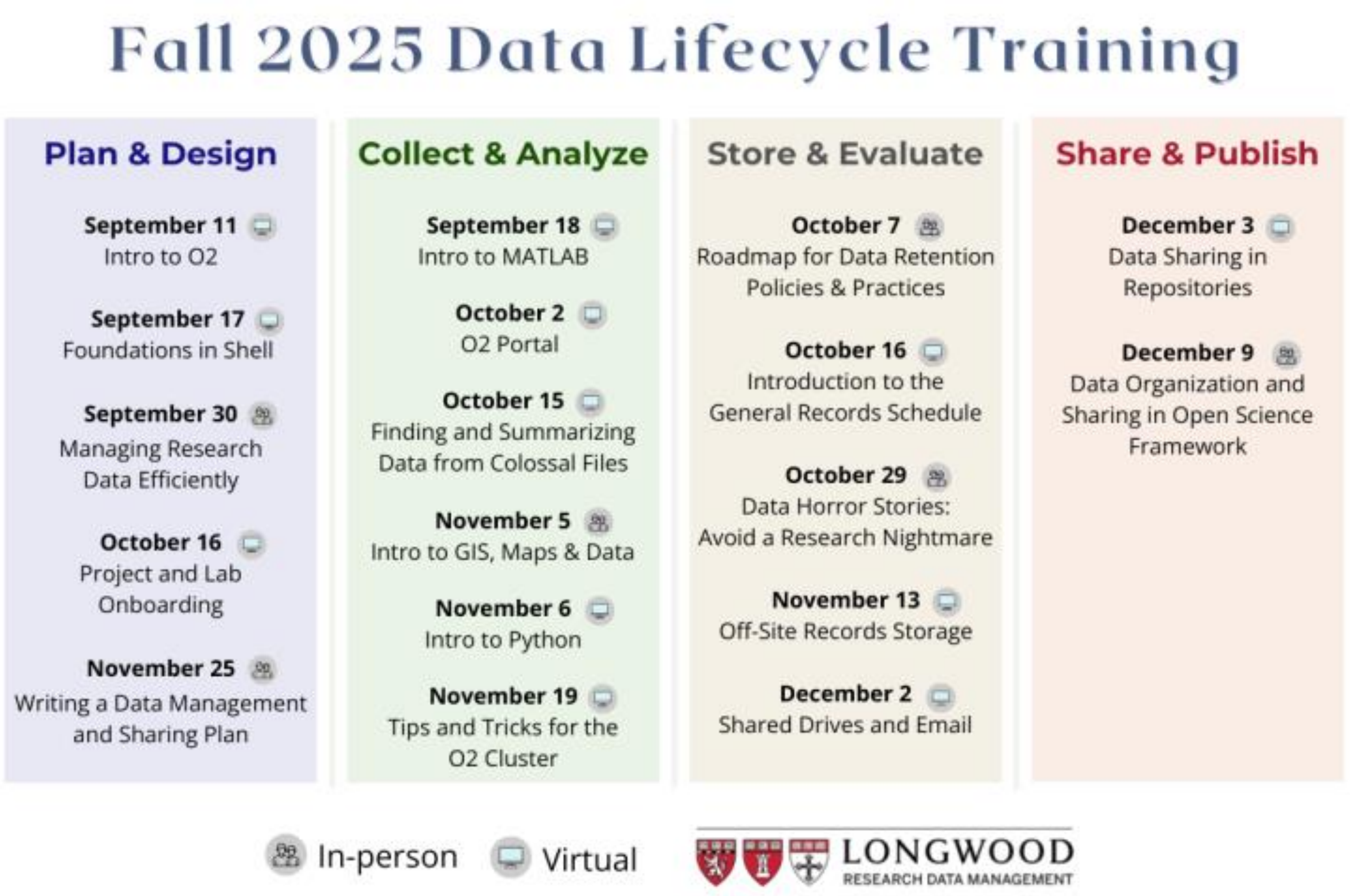
Establishing and supporting the reach and impact of your data

ACCESS & REUSE

Ensuring the broad utility of your research data efforts for other researchers

Better RDM practice benefits you

- **HMS Data Management LMA**
  - **Webpage:** https://datamanagement.hms.harvard.edu
  - **Sign up for quarterly email updates**
- **Harvard-wide Research data Management**
  - https://researchdatamanagement.harvard.edu/



Fall 2025 Data Lifecycle Training

Plan & Design

September 11 🖥
Intro to O2

September 17 🖥
Foundations in Shell

September 30 👥
Managing Research Data Efficiently

October 16 🖥
Project and Lab Onboarding

November 25 👥
Writing a Data Management and Sharing Plan

Collect & Analyze

September 18 🖥
Intro to MATLAB

October 2 🖥
O2 Portal

October 15 🖥
Finding and Summarizing Data from Colossal Files

November 5 👥
Intro to GIS, Maps & Data

November 6 🖥
Intro to Python

November 19 🖥
Tips and Tricks for the O2 Cluster

Store & Evaluate

October 7 👥
Roadmap for Data Retention Policies & Practices

October 16 🖥
Introduction to the General Records Schedule

October 29 👥
Data Horror Stories: Avoid a Research Nightmare

November 13 🖥
Off-Site Records Storage

December 2 🖥
Shared Drives and Email

Share & Publish

December 3 🖥
Data Sharing in Repositories

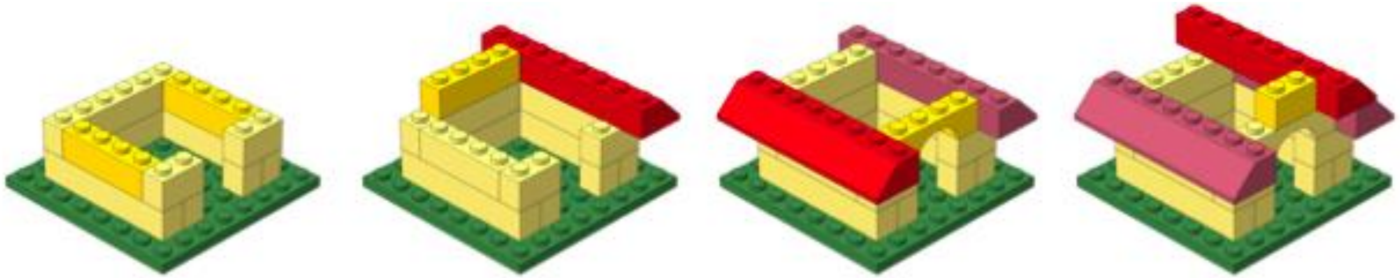
December 9 👥
Data Organization and Sharing in Open Science Framework

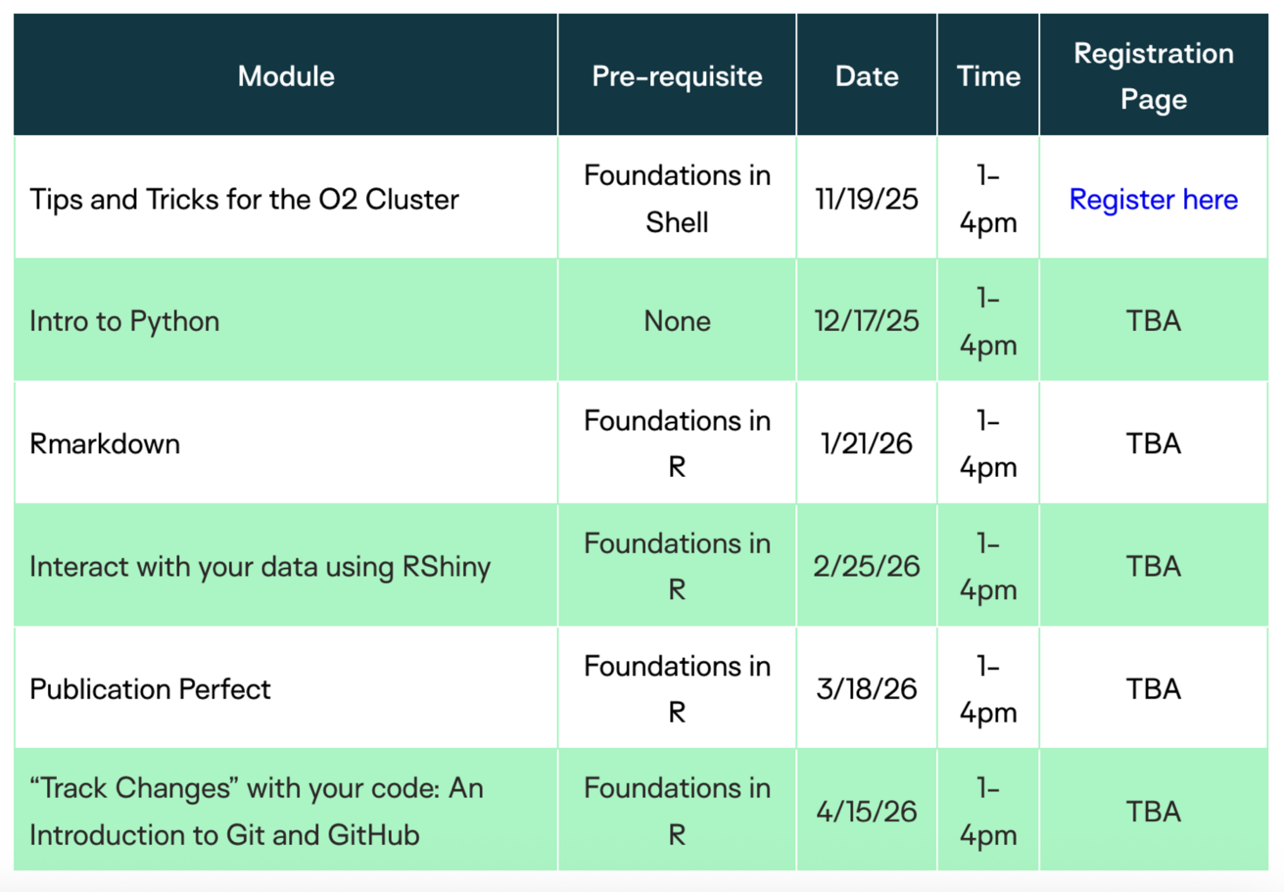
👥 In-person 🖥 Virtual

LONGWOOD
RESEARCH DATA MANAGEMENT

https://datamanagement.hms.harvard.edu/about/news-events/rdmwg-calendar

Keep building!

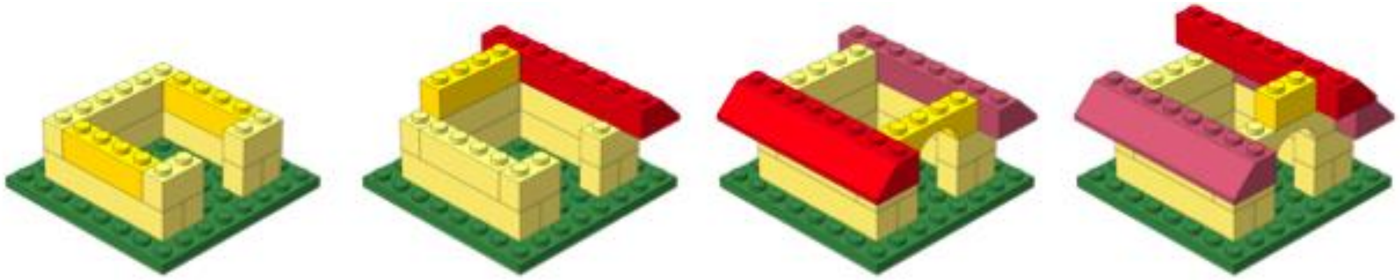




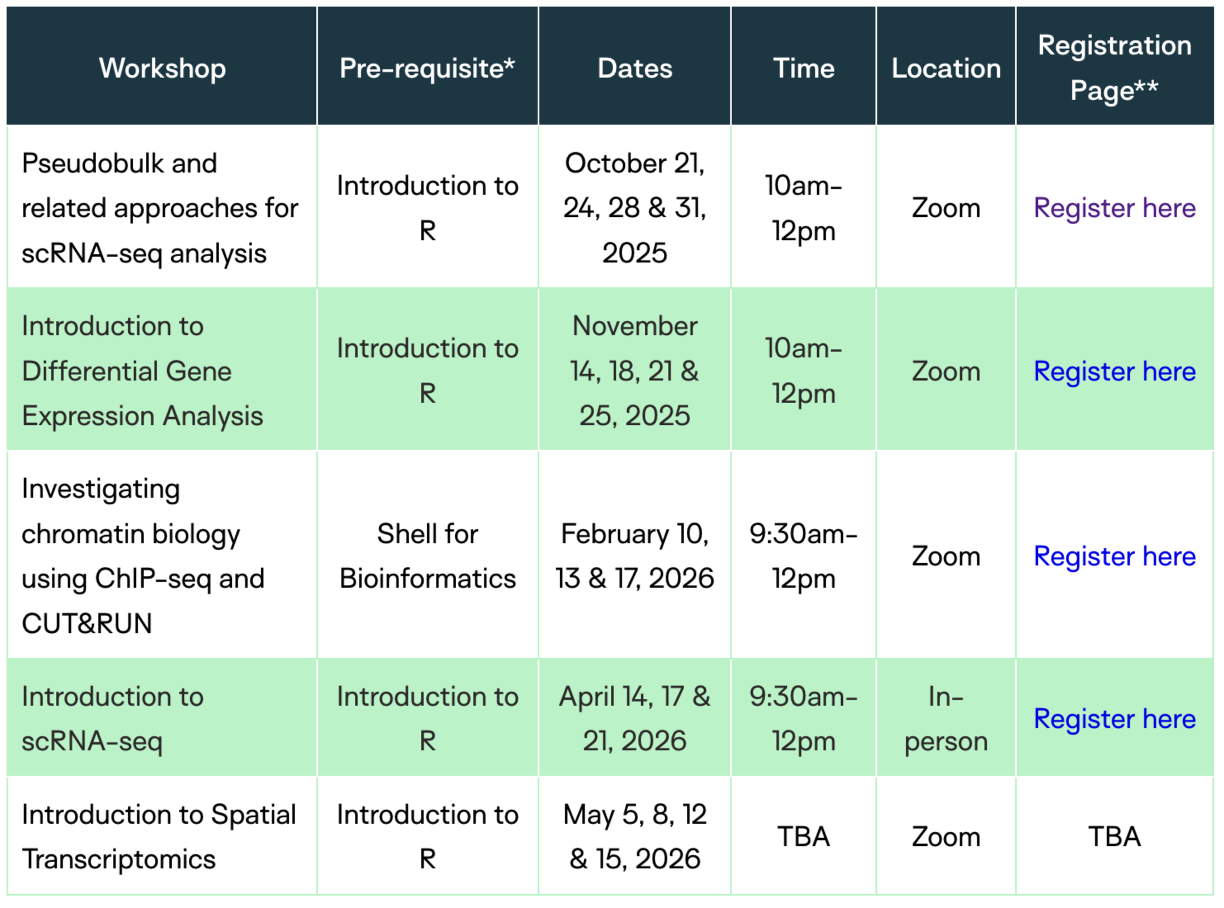
Module	Pre-requisite	Date	Time	Registration Page
Tips and Tricks for the O2 Cluster	Foundations in Shell	11/19/25	1-4pm	Register here
Intro to Python	None	12/17/25	1-4pm	TBA
Rmarkdown	Foundations in R	1/21/26	1-4pm	TBA
Interact with your data using RShiny	Foundations in R	2/25/26	1-4pm	TBA
Publication Perfect	Foundations in R	3/18/26	1-4pm	TBA
“Track Changes” with your code: An Introduction to Git and GitHub	Foundations in R	4/15/26	1-4pm	TBA

https://hsph.harvard.edu/research/bioinformatics/training/training-schedule/

Keep building!



Workshop	Dates	Time	Location	Registration Page*
Shell for Bioinformatics	January 23, 27 & 30, 2026	9:30am-12pm	In-person	Register here
Introduction to R	March 6, 10, 13 & 17, 2026	10am-12pm	In-person	Register here
Shell for Bioinformatics	June 2, 5 & 9, 2026	9:30am-12pm	Zoom	TBA



Workshop	Pre-requisite*	Dates	Time	Location	Registration Page**
Pseudobulk and related approaches for scRNA-seq analysis	Introduction to R	October 21, 24, 28 & 31, 2025	10am-12pm	Zoom	Register here
Introduction to Differential Gene Expression Analysis	Introduction to R	November 14, 18, 21 & 25, 2025	10am-12pm	Zoom	Register here
Investigating chromatin biology using ChIP-seq and CUT&RUN	Shell for Bioinformatics	February 10, 13 & 17, 2026	9:30am-12pm	Zoom	Register here
Introduction to scRNA-seq	Introduction to R	April 14, 17 & 21, 2026	9:30am-12pm	In-person	Register here
Introduction to Spatial Transcriptomics	Introduction to R	May 5, 8, 12 & 15, 2026	TBA	Zoom	TBA

https://hsph.harvard.edu/research/bioinformatics/training/training-schedule/

Talk to us early!

Involvement in study design to optimize experiments

More Information

- *HBC training materials: https://hbctraining.github.io/main*
- *HBC website: http://bioinformatics.sph.harvard.edu*

Contact Us

Sign up for our mailing list:

https://tinyurl.com/hbc-training-mailing-list

- *HBC training team:* hbctraining@hsph.harvard.edu
- *HBC consulting:* bioinformatics@hsph.harvard.edu