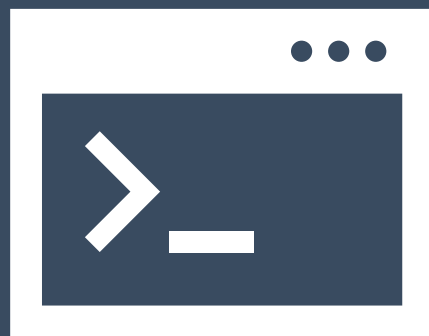


# Introduction to Single-cell RNA-seq Analysis

<https://tinyurl.com/hbc-scrnaseq-oct2024>



Harvard Chan Bioinformatics Core



# Introductions!





Shannan Ho Sui  
*Director*



Meeta Mistry  
*Associate Director*



Lorena Pantano  
*Director of Bioinformatics  
Platform*



John Quackenbush  
*Faculty Advisor*



Open Bhattarai



Heather Wick



Will Gammerdinger



Noor Sohail



Elizabeth  
Partan



Alex Bartlett



Emma Berdan



James Billingsley



Zhu Zhuo



Maria Simoneau



Shannan Ho Sui  
*Director*



Meeta Mistry  
*Associate Director*



Lorena Pantano  
*Director of Bioinformatics  
Platform*



John Quackenbush  
*Faculty Advisor*



Open Bhattarai



Heather Wick



Will Gammerdinger



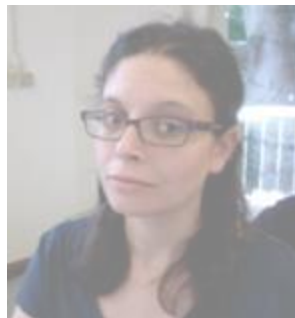
Noor Sohail



Elizabeth  
Partan



Alex Bartlett



Emma Berdan



James Billingsley



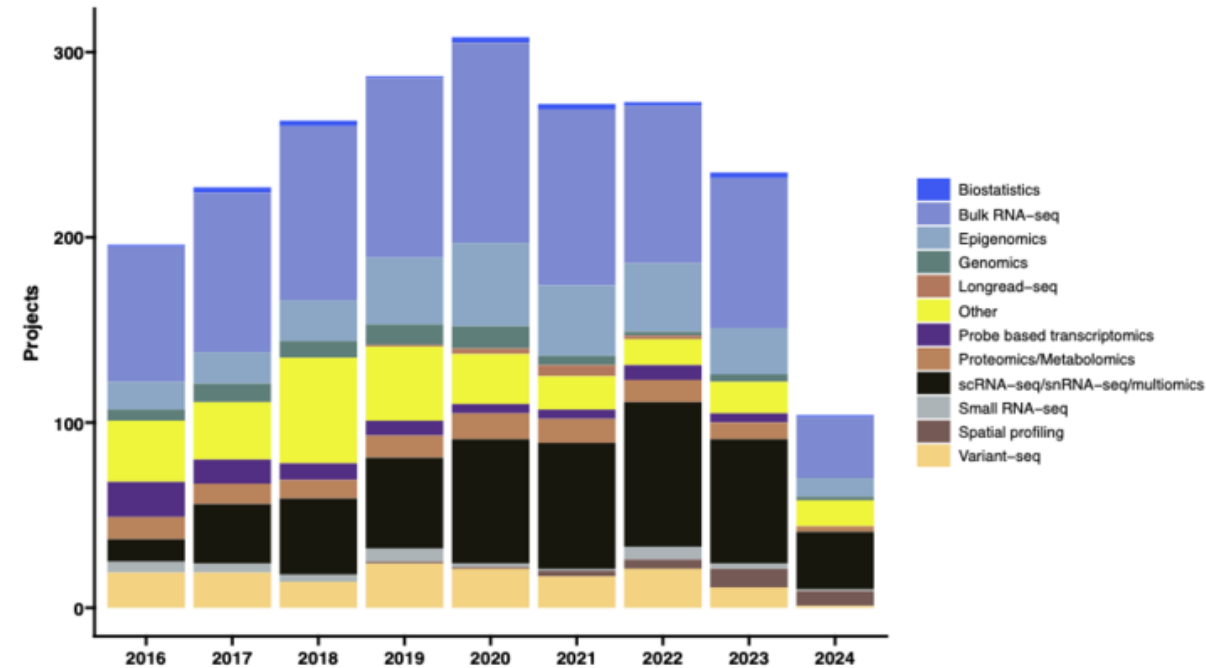
Zhu Zhuo



Maria Simoneau

# Consulting

- ❖ Transcriptomics: Bulk, single cell, small RNA
- ❖ Epigenomics: ChIP-seq, CUT&RUN, ATAC-seq, DNA methylation
- ❖ Variant discovery: WGS, resequencing, exome-seq and CNV
- ❖ Multiomics integration
- ❖ Spatial biology
- ❖ Experimental design and grant support



# Consulting

- ❖ Transcriptomics: Bulk, single cell, small RNA
- ❖ Epigenomics: ChIP-seq, CUT&RUN, ATAC-seq, DNA methylation
- ❖ Variant discovery: WGS, resequencing, exome-seq and CNV
- ❖ Multiomics integration
- ❖ Spatial biology
- ❖ Experimental design and grant support



**HARVARD**  
**T.H. CHAN**  
SCHOOL OF PUBLIC HEALTH

NIEHS

---



THE HARVARD CLINICAL  
AND TRANSLATIONAL  
SCIENCE CENTER

---



# Training

- ❖ Hands-on workshops design to reflect best practices, reproducibility and an emphasis on experimental design
  - ❖ Basic Data Skills
    - ❖ Shell
    - ❖ R
  - ❖ Advanced Topics: Analysis of high-throughput sequencing data
    - ❖ Chromatin Biology
    - ❖ Bulk RNA-seq
    - ❖ Differential Gene Expression
    - ❖ scRNA-seq
    - ❖ Variant Calling
  - ❖ Current Topics in Bioinformatics

# Training

- ❖ Hands-on workshops design to reflect best practices, reproducibility and an emphasis on experimental design
  - ❖ Basic Data Skills
    - ❖ Shell
    - ❖ R
  - ❖ Advanced Topics: Analysis of high-throughput sequencing data
    - ❖ Chromatin Biology
    - ❖ Bulk RNA-seq
    - ❖ Differential Gene Expression
    - ❖ scRNA-seq
    - ❖ Variant Calling
  - ❖ Current Topics in Bioinformatics

<https://bioinformatics.sph.harvard.edu/training>



**HARVARD**  
**T.H. CHAN**  
SCHOOL OF PUBLIC HEALTH

**DF/HCC**  
DANA-FARBER / HARVARD CANCER CENTER

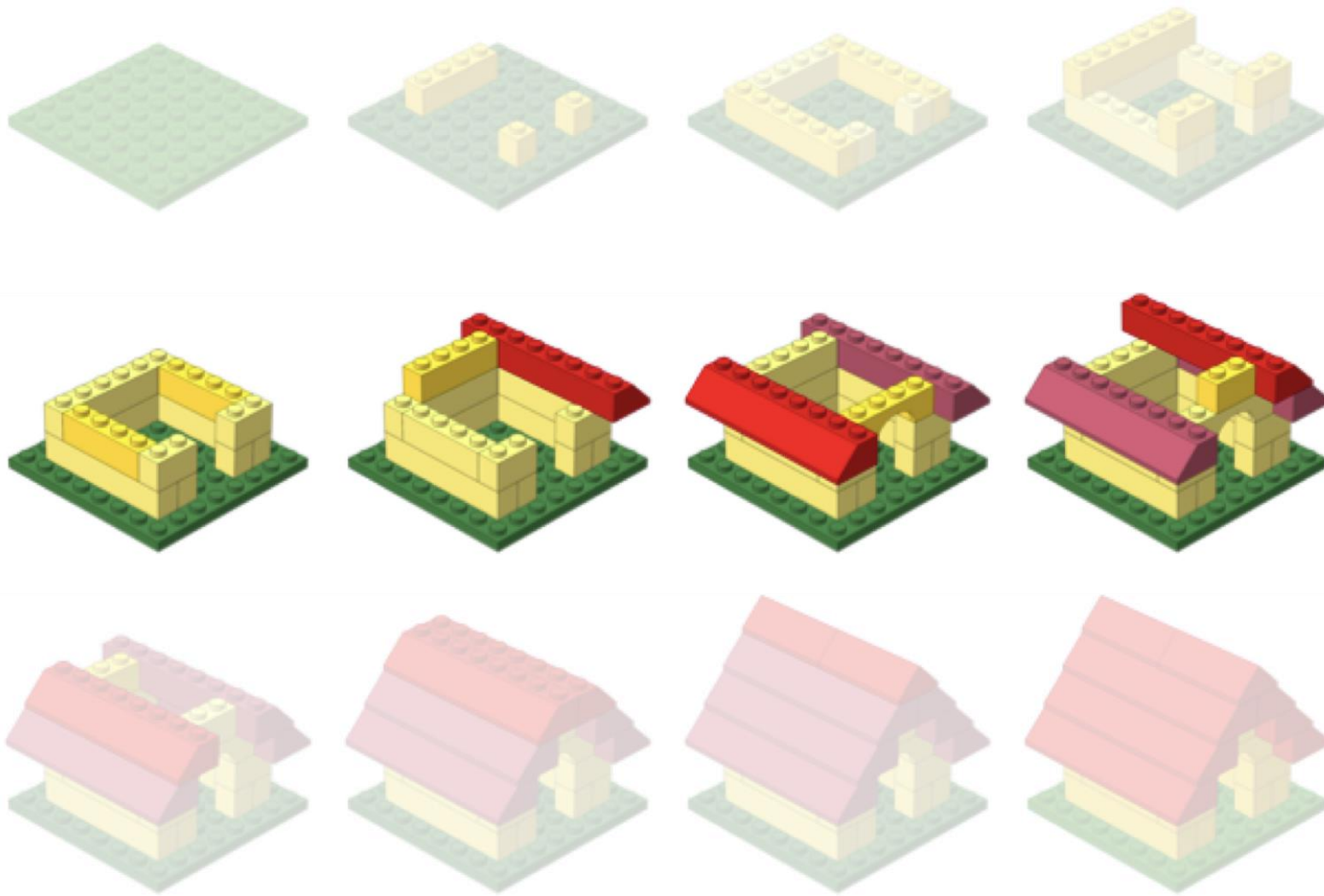


THE HARVARD CLINICAL  
AND TRANSLATIONAL  
SCIENCE CENTER



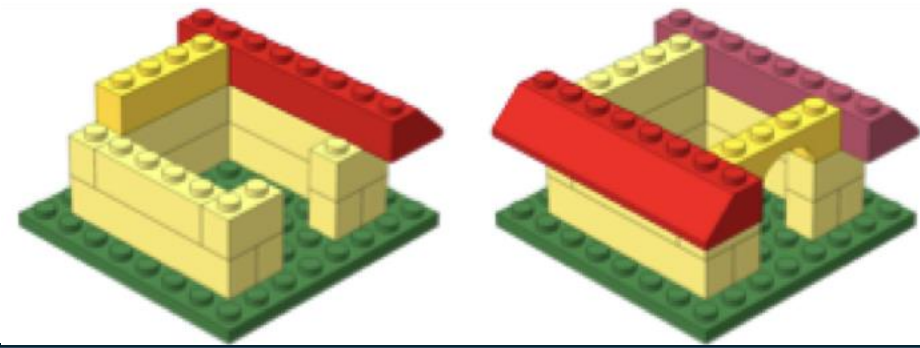


# **Workshop scope**



Bioinformatic Data Analysis

# Workshop Scope



- ❖ Describe best practices for designing a single-cell RNA-seq experiment
- ❖ Describe steps in a single-cell RNA-seq analysis workflow
- ❖ Use Seurat and associated tools to perform analysis of single-cell expression data, including data filtering, QC, integration, clustering, and marker identification
- ❖ Understand practical considerations for performing scRNA-seq, rather than in-depth exploration of algorithm theory

# Logistics



# Course schedule

## Day 1

Time	Topic	Instructor
09:30 - 09:45	Workshop introduction	Meeta
09:45 - 10:35	Introduction to Single Cell RNA-sequencing: a practical guide	Dr. Arpita Kulkarni
10:35 - 10:40	Break	
10:40 - 11:00	scRNA-seq pre-reading discussion	All
11:00 - 11:45	Quality control set-up	Noor
11:45 - 12:00	Overview of self-learning materials and homework submission	Meeta

<https://tinyurl.com/hbc-scrnaseq-oct2024>

# Course materials

- ❖ We continuously update our materials to reflect changes in the field/software

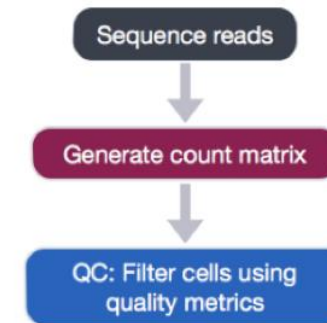


Approximate time: 90 minutes

## Learning Objectives:

- Understand how to bring in data from single-cell RNA-seq experiments
- Construct QC metrics and associated plots to visually explore the quality of the data
- Evaluate the QC metrics and set filters to remove low quality cells

## Single-cell RNA-seq: Quality control



# Single Screen & 3 Windows

The image displays a single screen with three overlapping windows:

- Zoom Meeting:** Shows a video call with three participants: Mary Piper, Troubleshooter (Radhika), and Jihe Liu. The interface includes a 'Participants (3)' list, a 'View Options' dropdown, and a 'You are viewing Jihe Liu's screen' notification.
- RStudio IDE:** Shows the R environment with the following code in the editor:

```
1 # Assignment operator
2 x <- 3
3
4 # Functions
5 getwd()
6
7 sqrt(81)
8
9 round(3.14159)
10 ?round
11
```

The console shows the execution of these commands:

```
> x <- 3
> # Functions
> getwd()
[1] "/Users/mariyaper/Desktop/R-testing"
> sqrt(81)
[1] 9
> round(3.14159)
[1] 3
> ?round
>
```

The Environment pane shows the values of variables:

```
Values
number 15
x       3
y      10
```
- Web Browser:** Shows a page with a blue background and white text, likely related to GitHub. The address bar shows the URL <https://hspH-bioc...>.

# Single Screen & 3 Windows

**Zoom**

*Our Recommendation*

```
1 # Assignment operator
2 x <- 3
3
4 # Functions
5 getwd()
6
7 sqrt(81)
8
9 round(3.14159)
10 ?round
11
```

```
> x <- 3
> # Functions
> getwd()
[1] "/Users/mariyaper/Desktop/R-testing"
> sqrt(81)
[1] 9
> round(3.14159)
[1] 3
> ?round
>
```

**Rounding of Numbers**

**Description**

ceiling takes a single numeric argument x and returns a numeric vector containing the smallest integers not less than the corresponding elements of x.

floor takes a single numeric argument x and returns a numeric vector containing the largest integers not greater than the corresponding elements of x.

trunc takes a single numeric argument x and returns a numeric vector containing the integers formed by truncating the values in x toward 0.

round rounds the values in its first argument to the specified number of decimal places (default 0). See 'Details' about "round to even" when rounding off a 5.

signif rounds the values in its first argument to the specified number of significant digits.

**Usage**

```
ceiling(x)
floor(x)
trunc(x, ...)
```



# Single Screen & 3 Windows

The screenshot displays a Zoom meeting interface. At the top, there are three video thumbnails for participants: Mary Piper, Troubleshooter (Radhika), and Jihe Liu. Below the thumbnails is a toolbar with icons for muting, video, and sharing. The main content area shows a shared RStudio window. The RStudio window has a dark theme and is divided into several panes. The top-left pane shows R code being edited. The bottom-left pane is the console, showing the execution of R code and its output. The right-hand pane is a 'Web Browser' window displaying the help page for the 'round' function in R. A green box highlights the 'Web Browser' window and the console output. The console output shows the following R code and its output:

```
1 # Assignment operator
2 x <- 3
3
4 # Functions
5 getwd()
6
7 sqrt(81)
8
9 round(3.14159)
10 ?round
11
```

The console output shows the following R output:

```
> x <- 3
> # Functions
> getwd()
[1] "/Users/mariyaper/Desktop/R-testing"
> sqrt(81)
[1] 9
> round(3.14159)
[1] 3
> ?round
```

The 'Web Browser' window shows the help page for the 'round' function, which includes the following text:

round(x, digits = 0, signif = 5)

round(x, digits = 0) rounds the values in its first argument to the specified number of decimal places (default 0). See 'Details' about "round to even" when rounding off a 5.

signif rounds the values in its first argument to the specified number of significant digits.

Usage

```
ceiling(x)
floor(x)
trunc(x, ...)
```

Our  
Recommendation

# Single Screen & 3 Windows

The image illustrates a single-screen setup for a Zoom meeting. The Zoom interface at the top shows three participants: Mary Piper (Co-host, me), Jihe Liu (Host), and Troubleshooter (Radhika) (Co-host). Below the Zoom window, the RStudio IDE is open, displaying R code in the editor and the console output. The code includes comments for assignment operators, functions, and rounding functions. The console output shows the results of the code execution. To the right, a browser window displays the R Documentation page for 'Rounding of Numbers', providing a description and usage for the ceiling, floor, trunc, and round functions.

```
483  
484  
485 getwd()  
486  
487 # square root function  
488 sqrt(81)  
489  
490 # round function  
491 round(3.14159)  
492 ?round  
493  
494  
495
```

```
1 # Assignment operator  
2 x <- 3  
3  
4 # Functions  
5 getwd()  
6  
7 sqrt(81)  
8  
9 round(3.14159)  
10 ?round  
11
```

```
> x <- 3  
> # Functions  
> getwd()  
[1] "/Users/mariyaper/Desktop/R-testing"  
> sqrt(81)  
[1] 9  
> round(3.14159)  
[1] 3  
> ?round  
>
```

**Our Recommendation**

**R Studio**

# Single Screen & 3 Windows

**Zoom**

**Web Browser**

**R Studio**

```
1 # Assignment operator
2 x <- 3
3
4 # Functions
5 getwd()
6
7 sqrt(81)
8
9 round(3.14159)
10 ?round
11
```

```
> x <- 3
> # Functions
> getwd()
[1] "/Users/mariyaper/Desktop/R-testing"
> sqrt(81)
[1] 9
> round(3.14159)
[1] 3
> ?round
>
```

Participants (3)

- Mary Piper (Co-host, me)
- Jihe Liu (Host)
- Troubleshooter (Radhika) (Co-host)

Values

number	15
x	5
y	10

Environment History Connections

Global Environment

Values

x	3
---	---

Files Plots Packages Help Viewer

R: Rounding of Numbers

Round (base)

Rounding of Numbers

Description

integers not less than the corresponding elements of x.

floor: takes a single numeric argument x and returns a numeric vector containing the largest integers not greater than the corresponding elements of x.

trunc: takes a single numeric argument x and returns a numeric vector containing the integers formed by truncating the values in x toward 0.

round: rounds the values in its first argument to the specified number of decimal places (default 0). See 'Details' about "round to even" when rounding off a 5.

signif: rounds the values in its first argument to the specified number of significant digits.

Usage

```
ceiling(x)
floor(x)
trunc(x, ...)
```

*Our  
Recommendation*

**R Studio**

# Course participation

- ❖ Mandatory review of self-learning lessons and assignments
- ❖ Attendance required for all classes
- ❖ Your questions and active participation drive learning
- ❖ **We look forward to all of your questions!**



# Course participation

- ❖ At-home lessons and exercises after each session
- ❖ Cover material not previously discussed
- ❖ Provides us feedback to help pace the course appropriately
- ❖ 3-5 hours to complete
- ❖ Homework load is heavier in the beginning of this workshop series and tapers off

# Using AI for Assignments

## ❖ Do

- ❖ Try to resolve error messages with it
- ❖ Test code written by AI on a dataset where you have expected results
- ❖ Take the time to review the generated code line-by-line

## ❖ Don't

- ❖ Implement it in replacement to learning
- ❖ Write code that you don't understand
- ❖ Assume the output from an AI process is correct

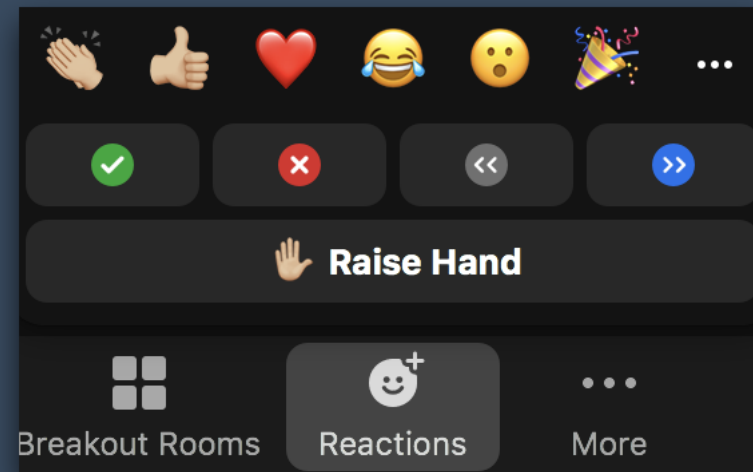
# Odds & Ends

❖ Quit/minimize all applications that are not required for class

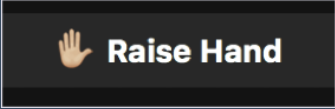
❖ Are you all set?

❖  = "agree", "I'm all set"

❖  = "disagree", "I need help"

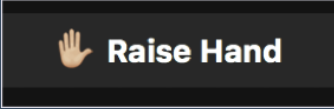


# Odds & Ends

- ❖ Questions for the presenter?
  - ❖ Post the question in the Chat window OR
  - ❖  when the presenter asks for questions
  - ❖ Let the Moderator know



# Odds & Ends

- ❖ Questions for the presenter?
  - ❖ Post the question in the Chat window OR
  - ❖  when the presenter asks for questions
  - ❖ Let the Moderator know
- ❖ Technical difficulties with software?
  - ❖ Start a private chat with the Troubleshooter with a description of the problem

# Contact Us

- ❖ *HBC training team:* [hbctraining@hsph.harvard.edu](mailto:hbctraining@hsph.harvard.edu)
- ❖ *HBC consulting:* [bioinformatics@hsph.harvard.edu](mailto:bioinformatics@hsph.harvard.edu)