

Introduction to Single-cell RNA-seq analysis

Harvard Chan Bioinformatics Core



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



Shannan Ho Sui
Director



Victor Barrera



Amélie Julé



Zhu Zhuo



James Billingsley



Radhika Khetani
Director of Education



Meeta Mistry



Jihe Liu



Will Gammerdinger



Emma Berdan



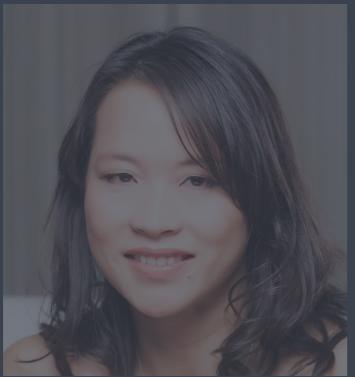
Sergey Naumenko



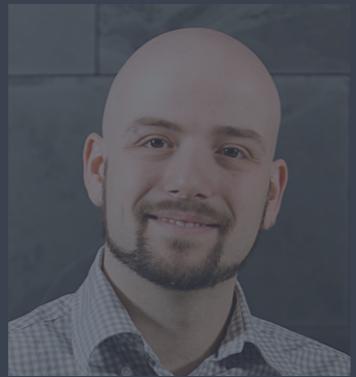
Maria Simoneau



Peter Kraft
Faculty Advisor



Shannan Ho Sui
Director



Victor Barrera



Amélie Julé



Zhu Zhuo



James Billingsley



Radhika Khetani
Director of Education



Meeta Mistry



Jihe Liu



Will Gammerdinger



Emma Berdan



Sergey Naumenko



Maria Simoneau



Peter Kraft
Faculty Advisor

Consulting

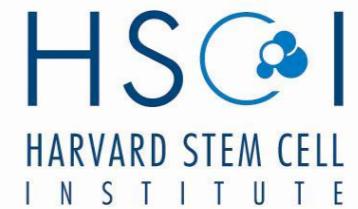
- Experimental design help
- RNA-seq analysis: bulk, single cell, small RNA
- ChIP-seq and ATAC-seq analysis
- Genome-wide methylation
- WGS, resequencing, exome-seq and CNV studies
- QC & analysis of gene expression arrays
- Functional enrichment analysis
- Grant support

<http://bioinformatics.sph.harvard.edu/>



HARVARD
T.H. CHAN
SCHOOL OF PUBLIC HEALTH

NIEHS



Training

A key component of the HBC's mission is its training initiative. Our dedicated training team holds workshop to help researchers at Harvard better understand analytical methods for NGS data.

HBC's training team is made up of four PhD-level scientists who devote substantial time to material development, training and community building/outreach. All members of the training team also participate in consultations on research projects to ensure they remain up-to-date on current best practices in NGS analysis.

Our hands-on workshops focus on **basic data skills** and **analysis of high-throughput sequencing data**, with an emphasis on **experimental design**, current **best practices** and **reproducibility**. Our workshops are designed for **wet-lab biologists** aiming to independently design sequencing-based experiments and analysing the resulting data.

We offer three types of workshops:

1. Short, 3-hour monthly workshops (*Current topics in bioinformatics*)
2. Basic Data Skills**
3. Advanced Topics: Analysis of high-throughput sequencing (NGS) data**

***The basic data skills workshops serve as the foundation for the advanced workshops.*

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

<https://hbctraining.github.io/main/>

Training

A key component of the HBC's mission is to train researchers at Harvard and beyond.

HBC's training team is made up of scientists, educators, and community leaders who work on research projects to ensure our training is relevant.

Our hands-on workshops focus on **bioinformatics**, with an emphasis on **experimental design** and **data analysis**. We also offer **wet-lab biologists** and **computational biologists** training in **NGS data analysis**.

We offer three types of workshops:

1. Short, 3-hour monthly workshops
2. Basic Data Skills**
3. Advanced Topics: Analysis of high-throughput sequencing data

***The basic data skills workshops are designed for the non-computer scientist.*



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

DF/HCC
DANA-FARBER / HARVARD CANCER CENTER



 **HARVARD CATALYST**
THE HARVARD CLINICAL
AND TRANSLATIONAL
SCIENCE CENTER



**HARVARD
MEDICAL SCHOOL**

Our dedicated training team holds workshops to help researchers learn how to analyze **bioinformatics** or **NGS** data.

We also devote substantial time to material development, including **workshop materials**. Our training team also participate in consultations on **bioinformatics** and **NGS analysis**.

Workshops focus on **bioinformatics**, with an emphasis on **experimental design** and **data analysis**. Our workshops are designed for **wet-lab biologists** and **computational biologists** to learn about **high-throughput sequencing data**, with an emphasis on **accuracy**, **precision**, and **reproducibility**. Our workshops are designed for **bioinformatics** and **NGS** analysis.

bioinformatics)

(**bioinformatics** and **NGS**) data**

and **bioinformatics** for the advanced workshops.

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

<https://hbctraining.github.io/main/>

Training

A key component of the HBC's mission is to train researchers at Harvard and beyond.

HBC's training team is made up of scientists, educators, and community leaders who work on research projects to ensure our training is relevant.

Our hands-on workshops focus on **bioinformatics**, with an emphasis on **experimental design** and **data analysis**. We also offer **wet-lab biologists** and **clinicians** training in how to analyze their data.

We offer three types of workshops:

1. Short, 3-hour monthly workshops
2. Basic Data Skills**
3. Advanced Topics: Analysis of high-throughput sequencing data

***The basic data skills workshops are designed for the non-bioinformatician.*



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

DF/HCC
DANA-FARBER / HARVARD CANCER CENTER



 **HARVARD CATALYST**
THE HARVARD CLINICAL
AND TRANSLATIONAL
SCIENCE CENTER



**HARVARD
MEDICAL SCHOOL**

Our dedicated training team holds workshops to help researchers learn how to analyze their data, including NGS data.

We also devote substantial time to material development, including **workshop materials**. Our training team also participate in consultations on **best practices in NGS analysis**.

Analysis of high-throughput sequencing data, with an emphasis on **experimental design**, **bioinformatics**, and **reproducibility**. Our workshops are designed for **wet-lab biologists** and **clinicians** performing sequencing-based experiments and analysing the resulting data.

bioinformatics)

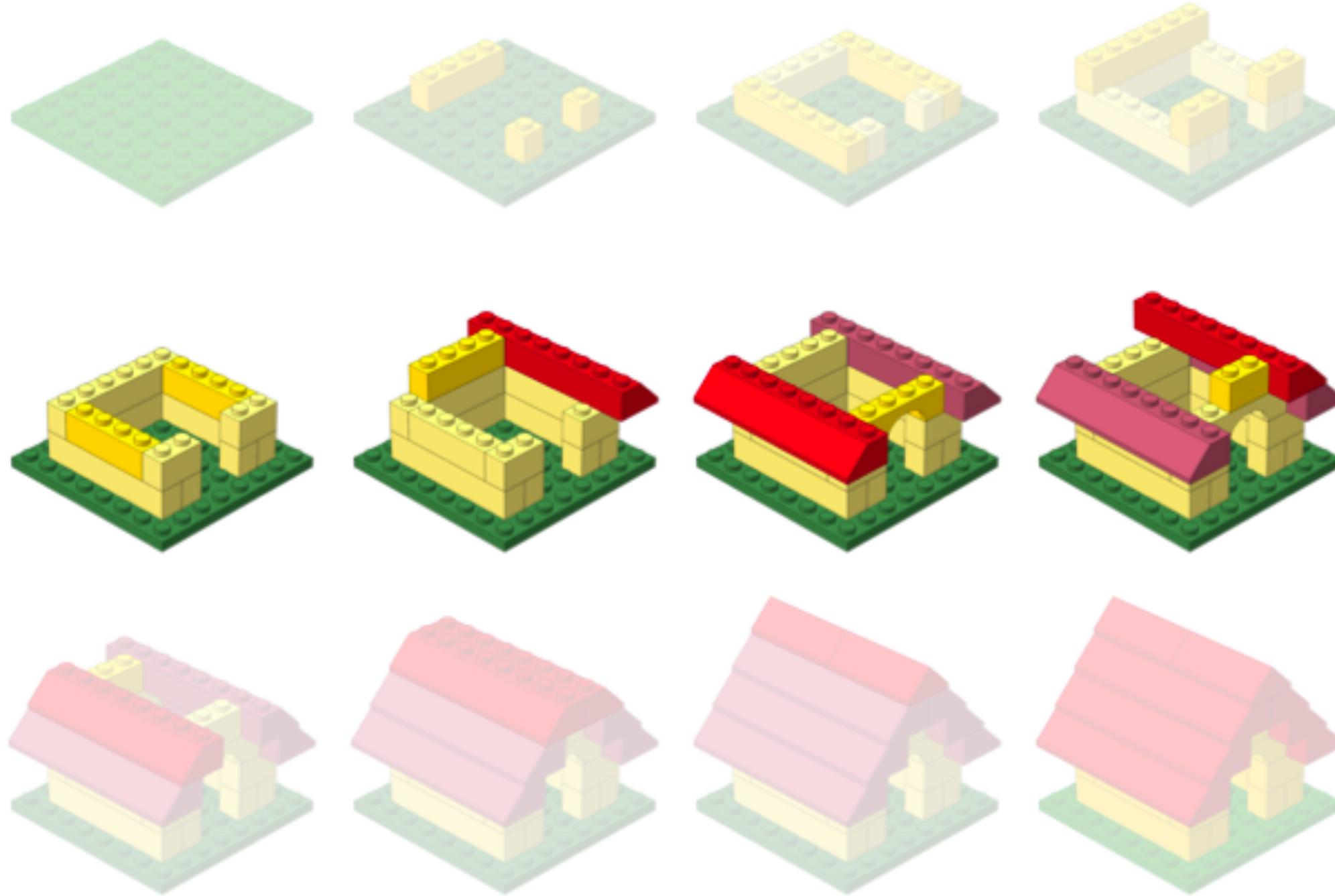
NGS) data**

or the advanced workshops.

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

<https://hbctraining.github.io/main/>

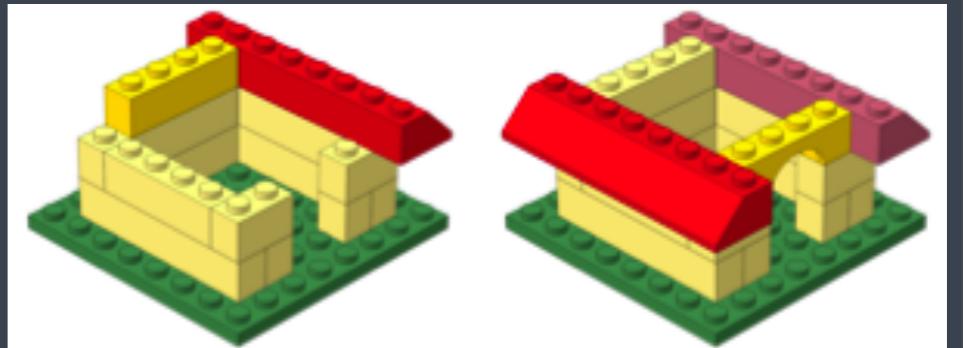
Workshop scope



<http://anoved.net/tag/lego/page/3/>

Bioinformatics data analysis

Learning Objectives



- ✓ 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 webpage

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

Course materials online

Introduction to Single-cell RNA-seq

[View on GitHub](#)

Approximate time: 90 minutes

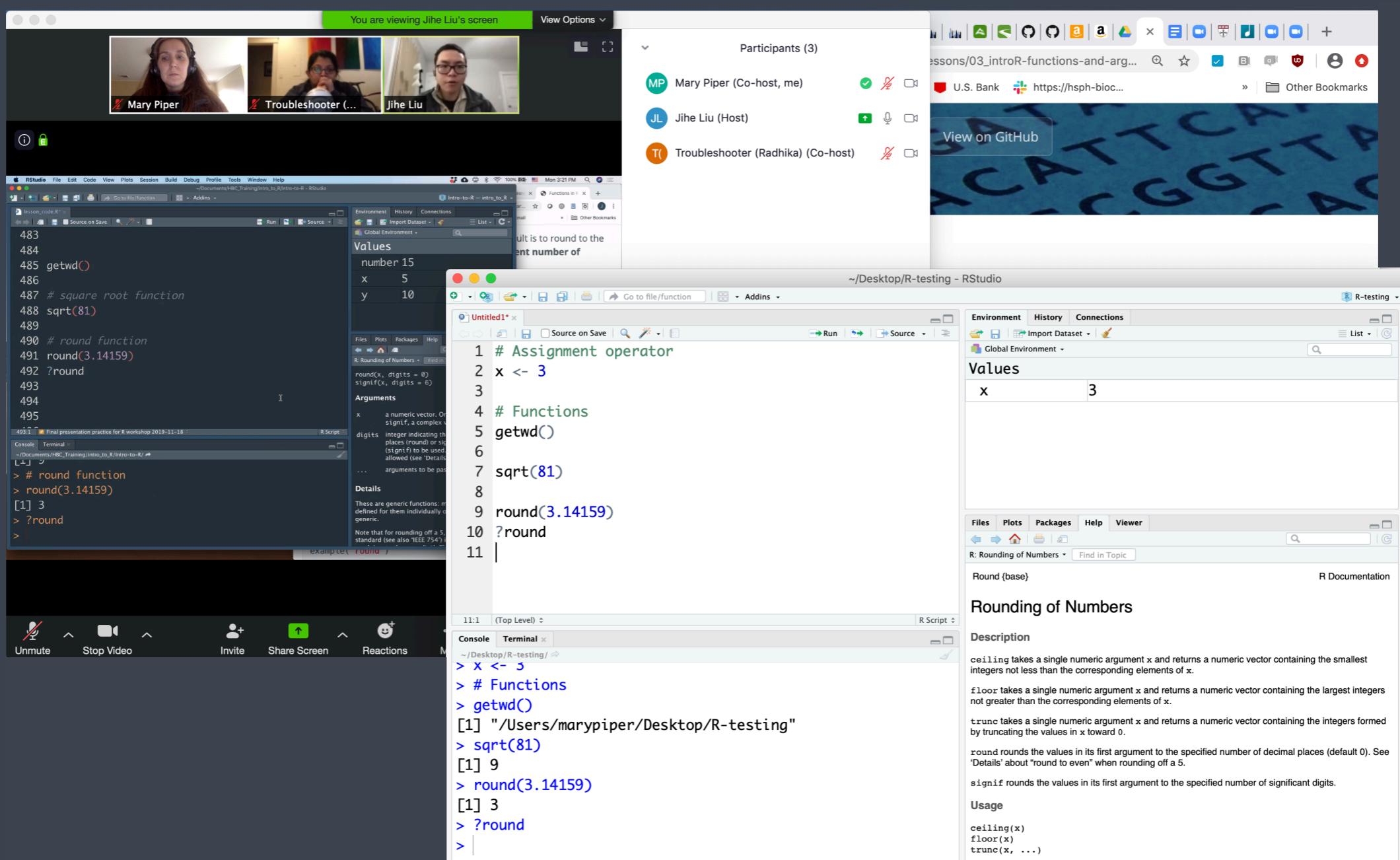
Learning Objectives:

- 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

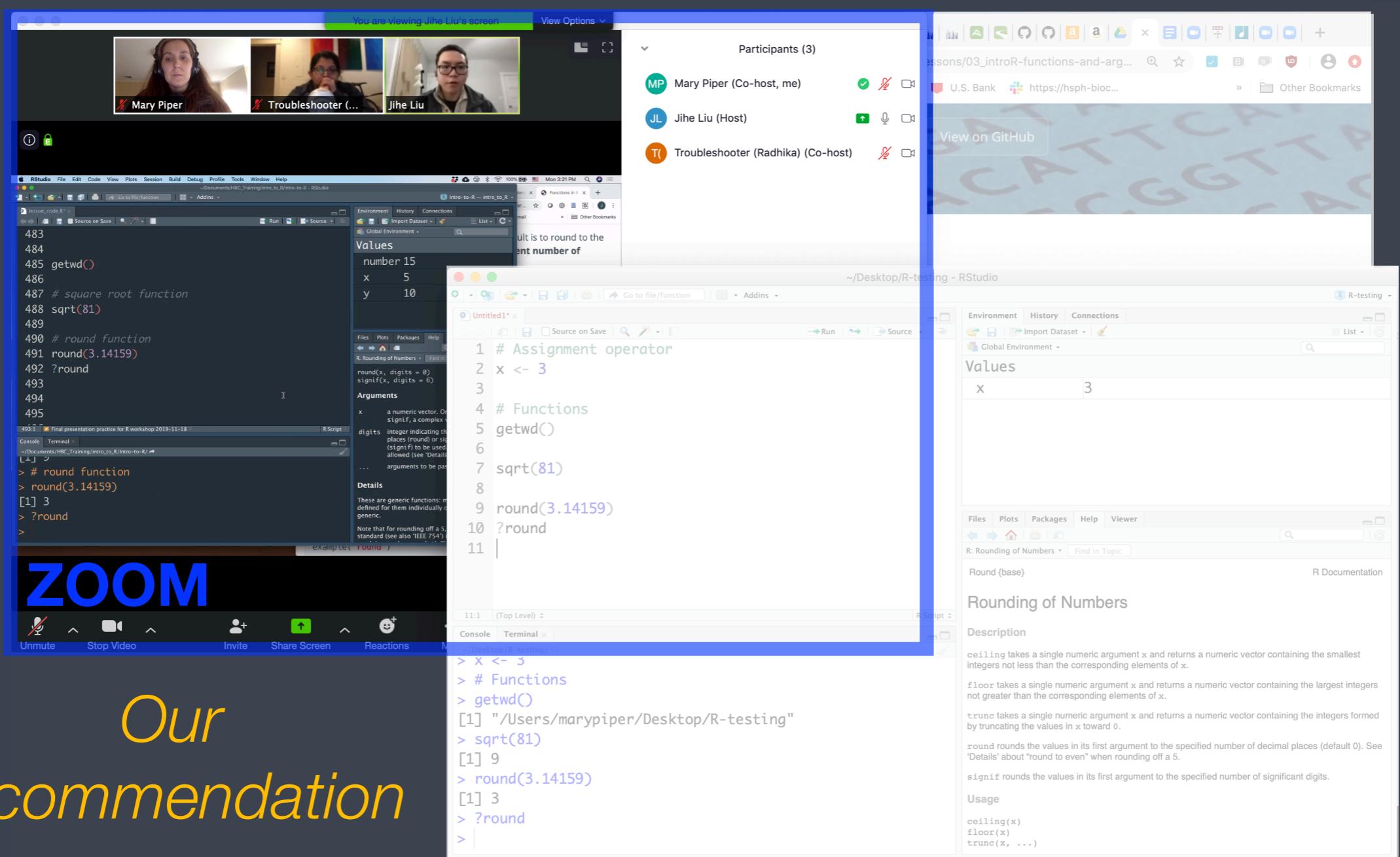
Single-cell RNA-seq: Quality control

```
graph TD; A[Sequence reads] --> B[Generate count matrix]; B --> C["QC: Filter cells using quality metrics"]
```

Single screen & 3 windows?

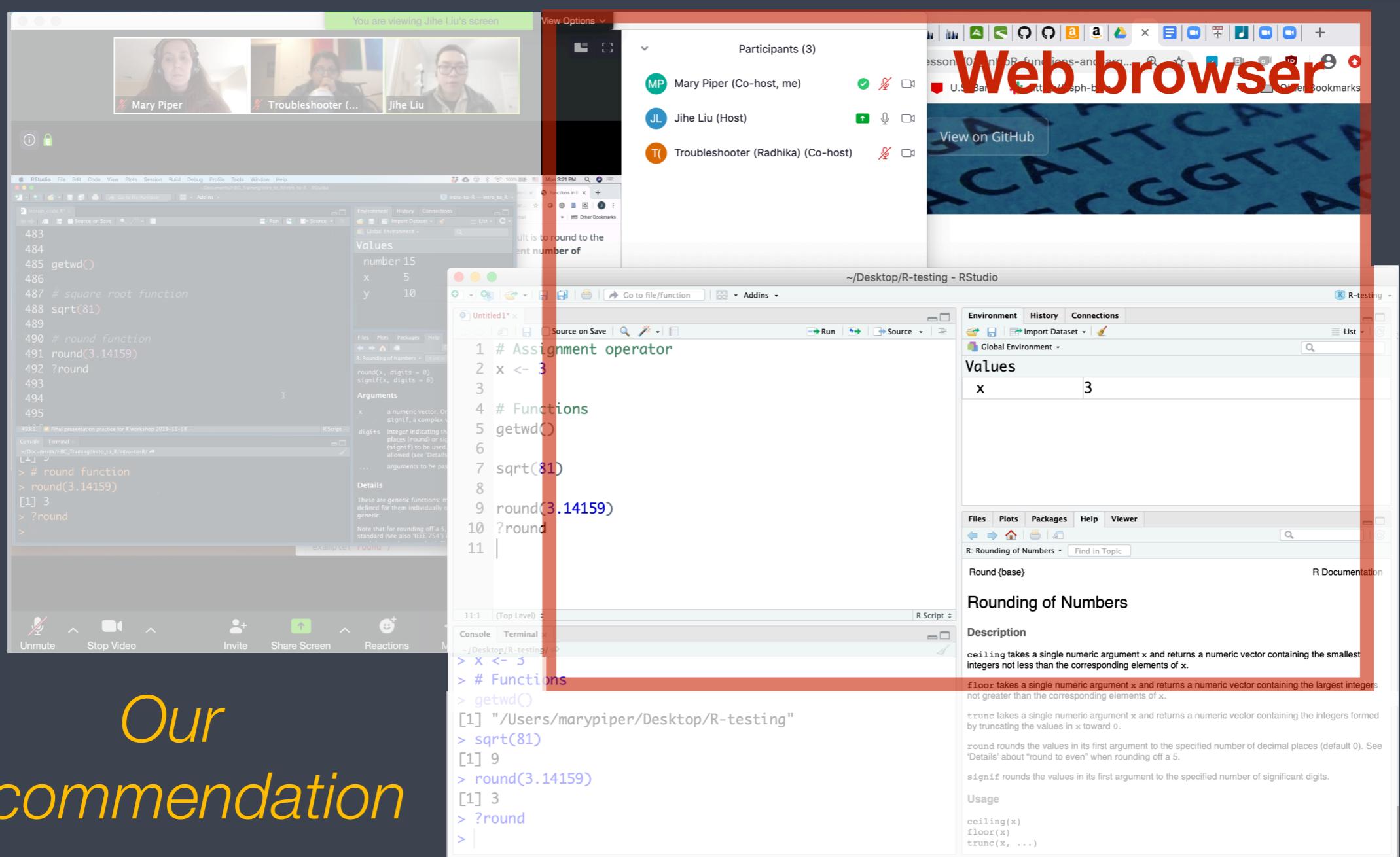


Single screen & 3 windows?



*Our
recommendation*

Single screen & 3 windows?



Single screen & 3 windows?

The screenshot shows a video conference interface with three windows:

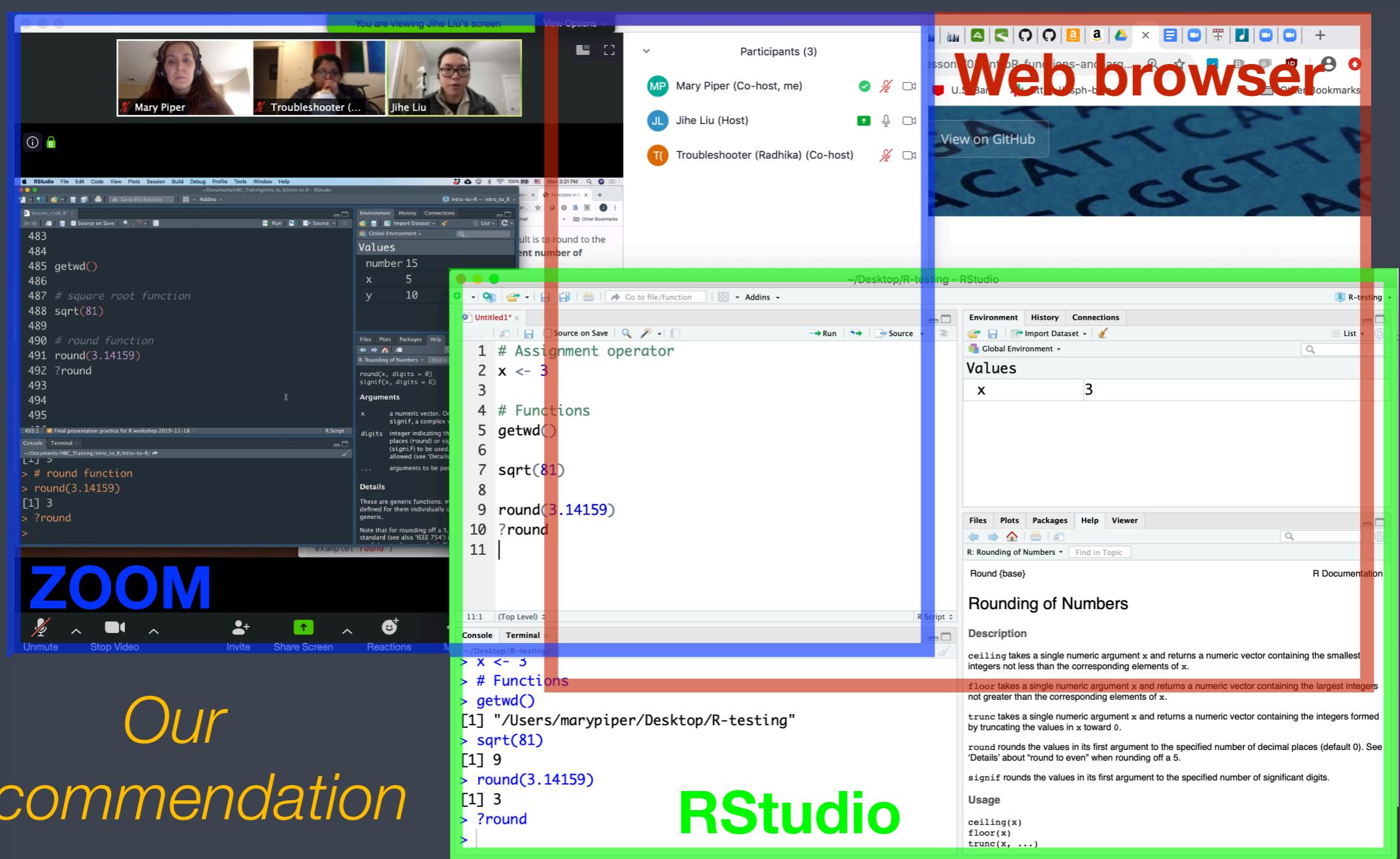
- Top Left Window:** A video feed showing three participants: Mary Piper, Troubleshooter (Radhika), and Jihe Liu.
- Middle Left Window:** An RStudio session titled "intro_to_R -- intro_to_R.R" containing the following R code:

```
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
```
- Bottom Left Window:** An RStudio session titled "Untitled1*" containing the following R code:

```
1 # Assignment operator  
2 x <- 3  
3  
4 # Functions  
5 getwd()  
6  
7 sqrt(81)  
8  
9 round(3.14159)  
10 ?round  
11
```
- Top Right Window:** A web browser window showing a presentation slide with the title "R: Rounding of Numbers". The slide includes sections for "Description", "Details", and "Usage". It also contains a screenshot of the RStudio session from the bottom left window.
- Bottom Right Window:** A large green box highlighting the RStudio session from the bottom left window, specifically the "Assignment operator" section.

Our recommendation: RStudio

Single screen & 3 windows?



*Our
recommendation*

Course participation

- ▶ Please keep your videos on, we would love to see your faces!
- ▶ 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!

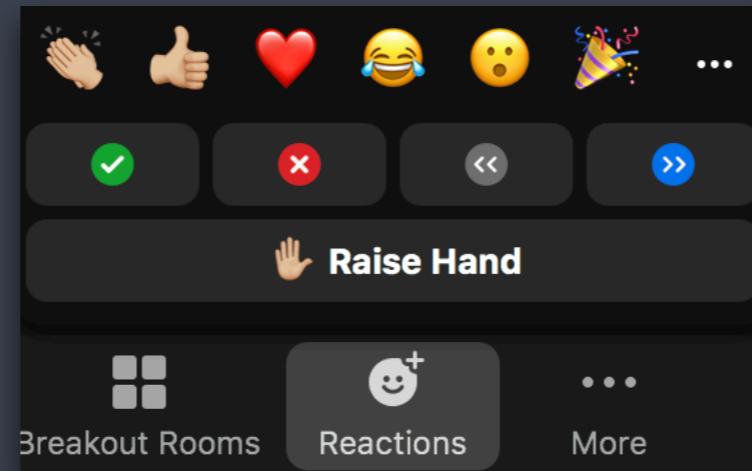


Odds and Ends

- ❖ Quit/minimize all applications that are not required for class
- ❖ Captioning is available upon request

Odds and Ends (1/2)

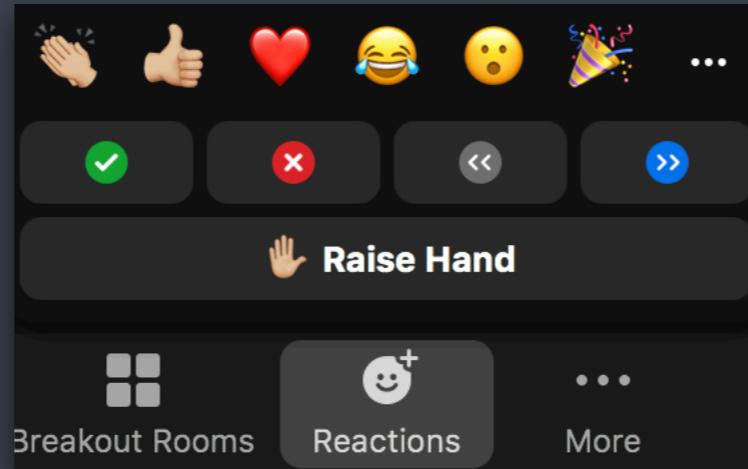
- ❖ Quit/minimize all applications that are not required for class
- ❖ Captioning is available upon request
- ❖ Are you all set?
 - ▶  = "agree", "I'm all set" (equivalent to a **green post-it**)
 - ▶  = "disagree", "I need help" (equivalent to a **red post-it**)



Odds and Ends (2/2)

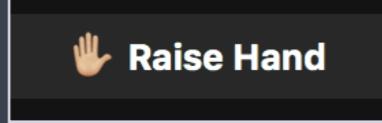
- ❖ Questions for the presenter?

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



Odds and Ends (2/2)

- ❖ 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 RStudio or R?

- Start a private chat with the Troubleshooter with a description of the problem.

Contact us!

Training team  : hbctraining@hsph.harvard.edu

Consulting  : bioinformatics@hsph.harvard.edu

 [@bioinfocore](https://twitter.com/bioinfocore)