

# Introduction to R

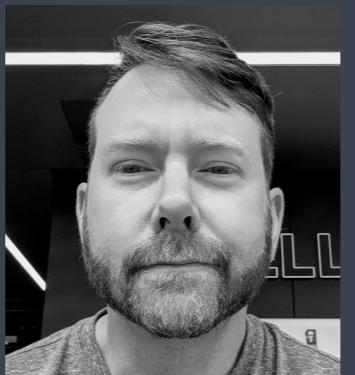
Harvard Chan Bioinformatics Core

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

Sponsored by DF/HCC



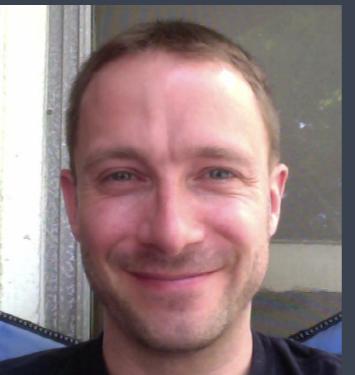
Shannan Ho Sui  
*Director*



John Hutchinson  
*Associate Director*



Victor Barrera



Rory Kirchner



Zhu Zhuo



Preetida Bhetariya



Meeta Mistry



Mary Piper



Jihe Liu



Radhika Khetani  
*Training Director*



Ilya Sytchev



James Billingsley



Sergey Naumenko



Joon Yoon

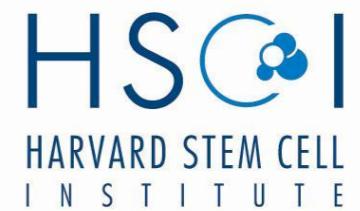


Peter Kraft  
*Faculty Advisor*

# Consulting

- 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/>



# Training

We have divided our short workshops into 2 categories:

1. Basic Data Skills - No prior programming knowledge needed (no prerequisites)
2. Advanced Topics: Analysis of high-throughput sequencing (NGS) data - Certain “Basic” workshops required as prerequisites.

*Any participants wanting to take an advanced workshop will have to have taken the appropriate basic workshop(s) within the past 6 months.*

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

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



# Training

We have divided our short workshops into 2 categories:

1. Basic Data Skills - No prior programming knowledge needed (no prerequisites)
2. Advanced Topics: Analysis of high-throughput sequencing (NGS) data - Certain “Basic” workshops required as prerequisites.

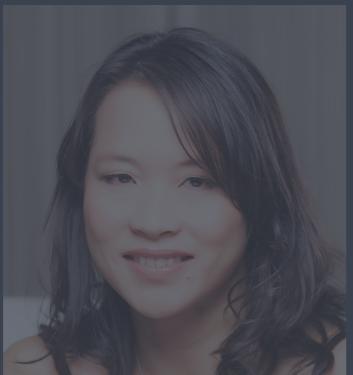
*Any participants wanting to take an advanced workshop will have to have taken the appropriate basic workshop(s) within the past 6 months.*

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

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



# Introductions!



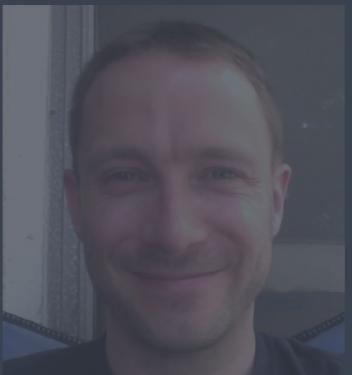
Shannan Ho Sui  
*Director*



John Hutchinson  
*Associate Director*



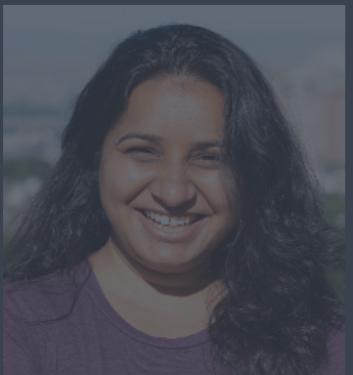
Victor Barrera



Rory Kirchner



Zhu Zhuo



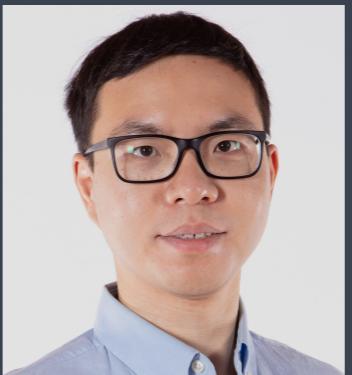
Preetida Bhetariya



Meeta Mistry



Mary Piper



Jihe Liu



Radhika Khetani  
*Training Director*



Ilya Sytchev



James Billingsley



Sergey Naumenko

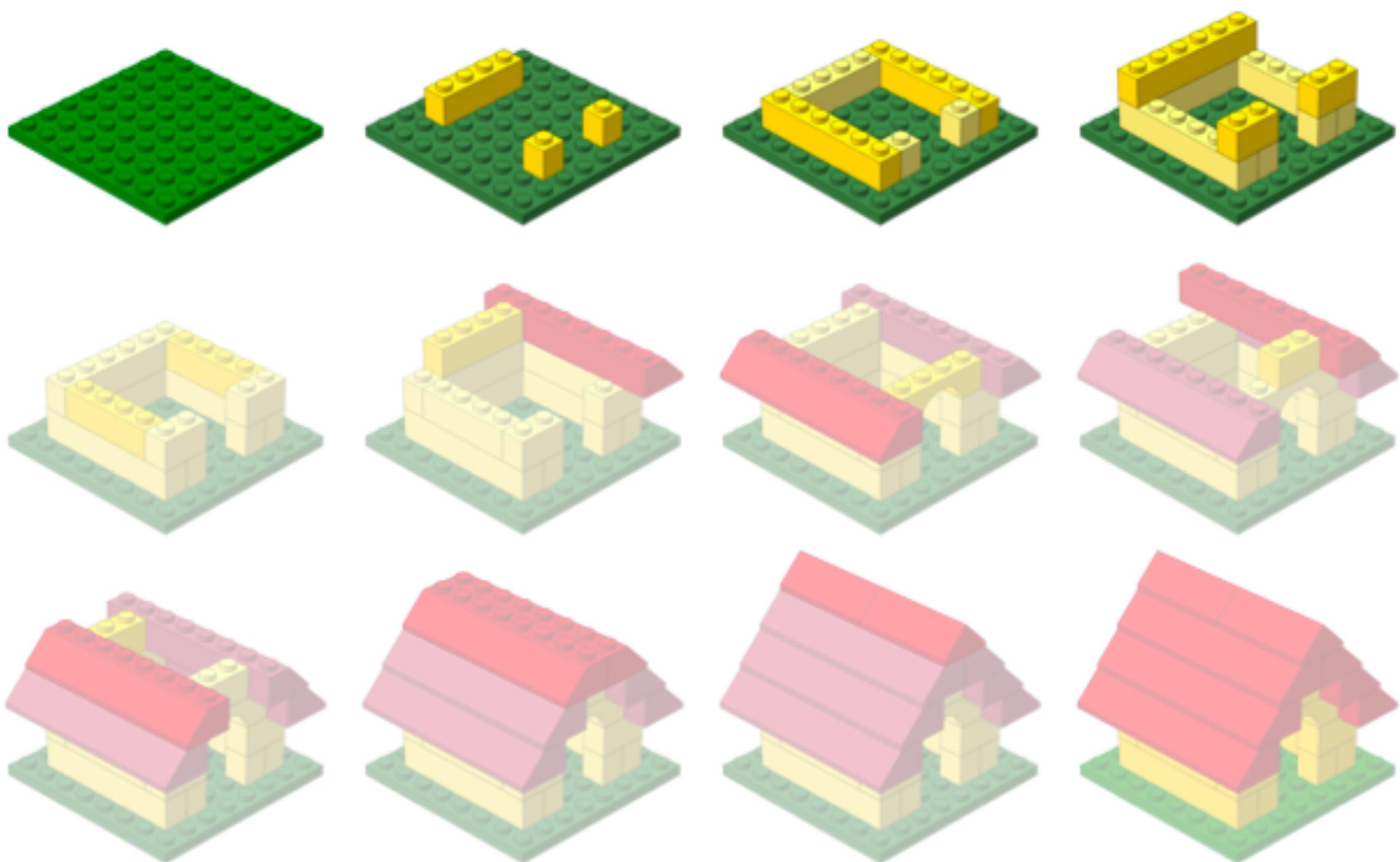


Joon Yoon



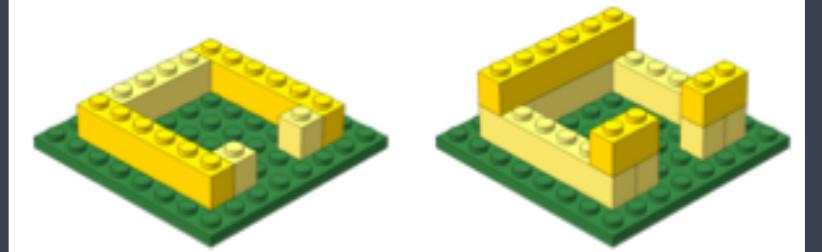
Peter Kraft  
*Faculty Advisor*

# Workshop Scope...



# Learning R

# Workshop Scope



- ✓ Comfortably use RStudio (a graphical interface for R)
- ✓ Fluently interact with R using RStudio
- ✓ Become familiar with R syntax
- ✓ Understand data structures in R
- ✓ Inspect and manipulate data structures
- ✓ Install packages and use functions in R

# CRAN

# (Comprehensive R Archive Network)



**Available CRAN Packages By Name**

[A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [J](#) [K](#) [L](#) [M](#) [N](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [T](#) [U](#) [V](#) [W](#) [X](#) [Y](#) [Z](#)

<a href="#">A3</a>	Accurate, Adaptable, and Accessible Error Metrics for Predictive Models
<a href="#">abbyyR</a>	Access to Abbyy Optical Character Recognition (OCR) API
<a href="#">abc</a>	Tools for Approximate Bayesian Computation (ABC)
<a href="#">ABCanalysis</a>	Computed ABC Analysis
<a href="#">abc.data</a>	Data Only: Tools for Approximate Bayesian Computation (ABC)
<a href="#">abcdeFBA</a>	ABCDE_FBA: A-Biologist-Can-Do-Everything of Flux Balance Analysis with this package
<a href="#">ABCOptim</a>	Implementation of Artificial Bee Colony (ABC) Optimization
<a href="#">ABCp2</a>	Approximate Bayesian Computational Model for Estimating P2
<a href="#">abcrf</a>	Approximate Bayesian Computation via Random Forests

*CRAN  
Mirrors  
What's new?  
Task Views  
Search  
  
About R  
R Homepage  
The R Journal*

- The main repository for R packages
- Easy to install

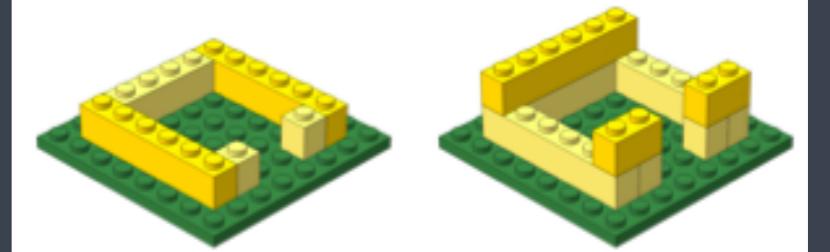
<https://cran.r-project.org/>



- An alternative package repository; “..provides tools for the analysis and comprehension of *high-throughput genomic data*.”
- Includes (but is not limited to) tools for:
  - ♦ performing statistical analysis
  - ♦ accessing public datasets
- Open source and open development
- Free

[www.bioconductor.org](http://www.bioconductor.org)

# Workshop Scope



- Comfortably use RStudio (a graphical interface for R)
  - Fluently interact with R using RStudio
  - Become familiar with R syntax
  - Understand data structures in R
  - Inspect and manipulate data structures
  - Install packages and use functions in R
- ✓ Visualize data using *ggplot2*
- ✓ Utilize pipes, tibbles and functions from the Tidyverse package suite

# Logistics

# Course webpage

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

# Course schedule online

## Workshop Schedule

### Day 1

Time	Topic	Instructor
10:00 - 10:30	Workshop Introduction	Radhika
10:30 - 11:15	Introduction to R and RStudio	Radhika
11:15 - 11:30	Overview of self-learning materials and homework submission	Mary

### Self-Learning Part 1

1. [R Syntax and Data Structure](#)
2. [Functions and Arguments](#)

### Assignment #1

- All exercises from above lessons have been put together in [html format](#) (web access) and [R script format](#) (download for local access).
- Add your solutions to the exercises in the downloaded [.R](#) file and **upload the saved file** to [Dropbox](#) on **Wednesday before the next class**.
- [Email us](#) about questions regarding the homework that you need answered before the next class.
- Post questions that you would like to have reviewed in class [here](#).

# Course materials online



## Introduction to R

[View on GitHub](#)

Approximate time: 70 min

### Learning Objectives

- Employ variables in R.
- Describe the various data types used in R.
- Construct data structures to store data.

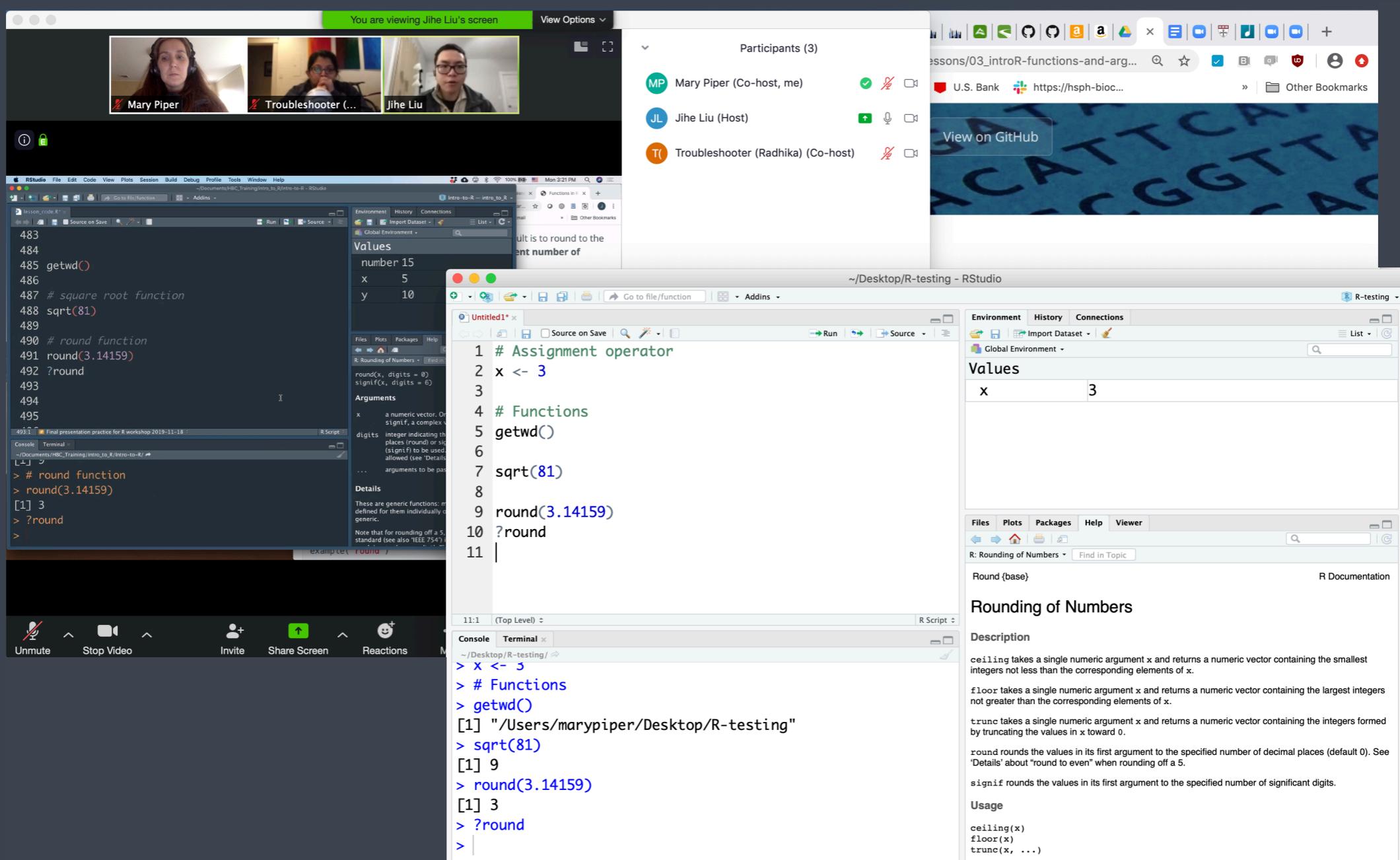
### The R syntax

Now that we know how to talk with R via the script editor or the console, we want to use R for something more than adding numbers. To do this, we need to know more about the R syntax.

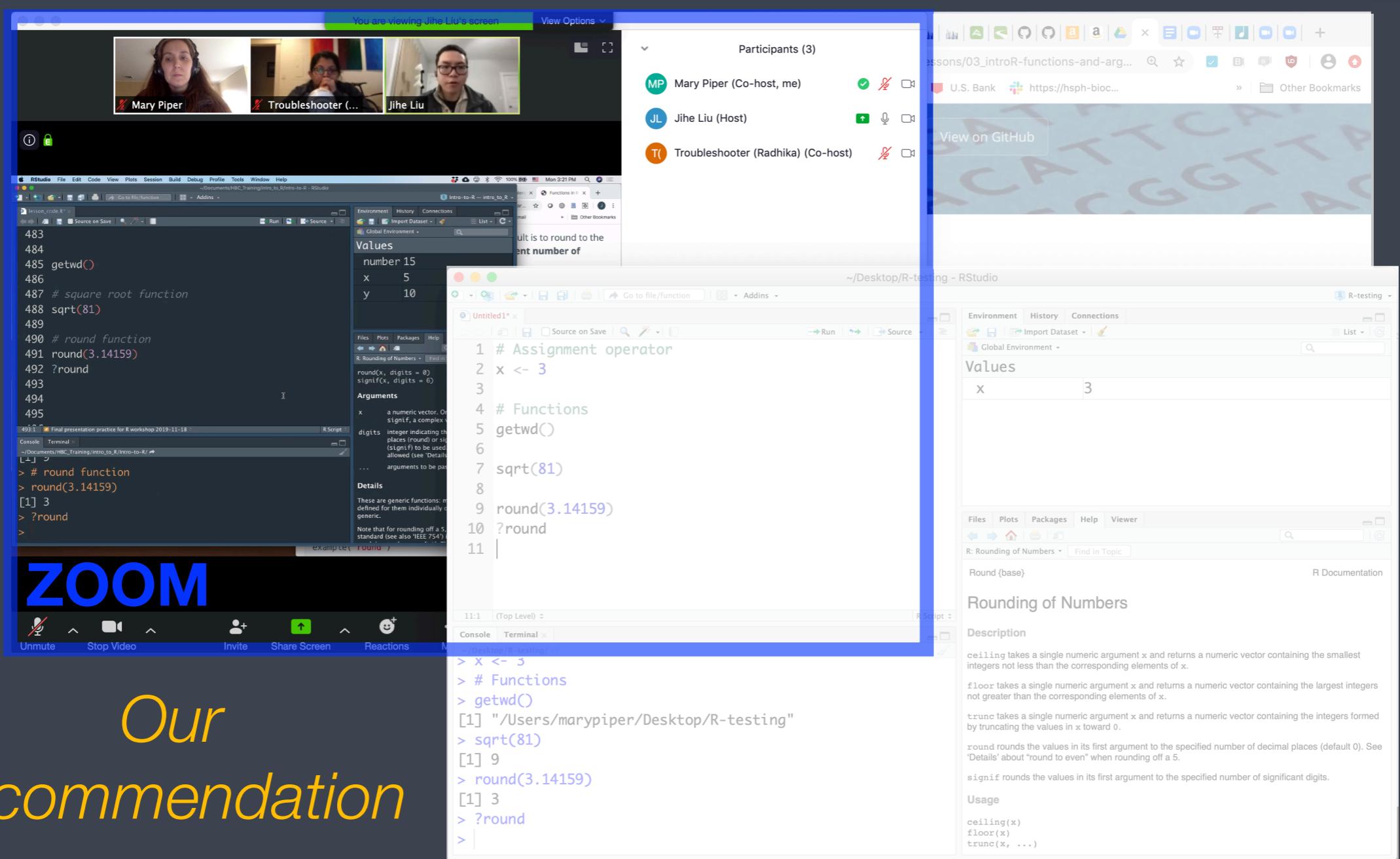
Below is an example script highlighting the many different “parts of speech” for R (syntax):

- the **comments** `#` and how they are used to document function and its content
- **variables and functions**
- the **assignment operator** `<-`

# Single screen & 3 windows?

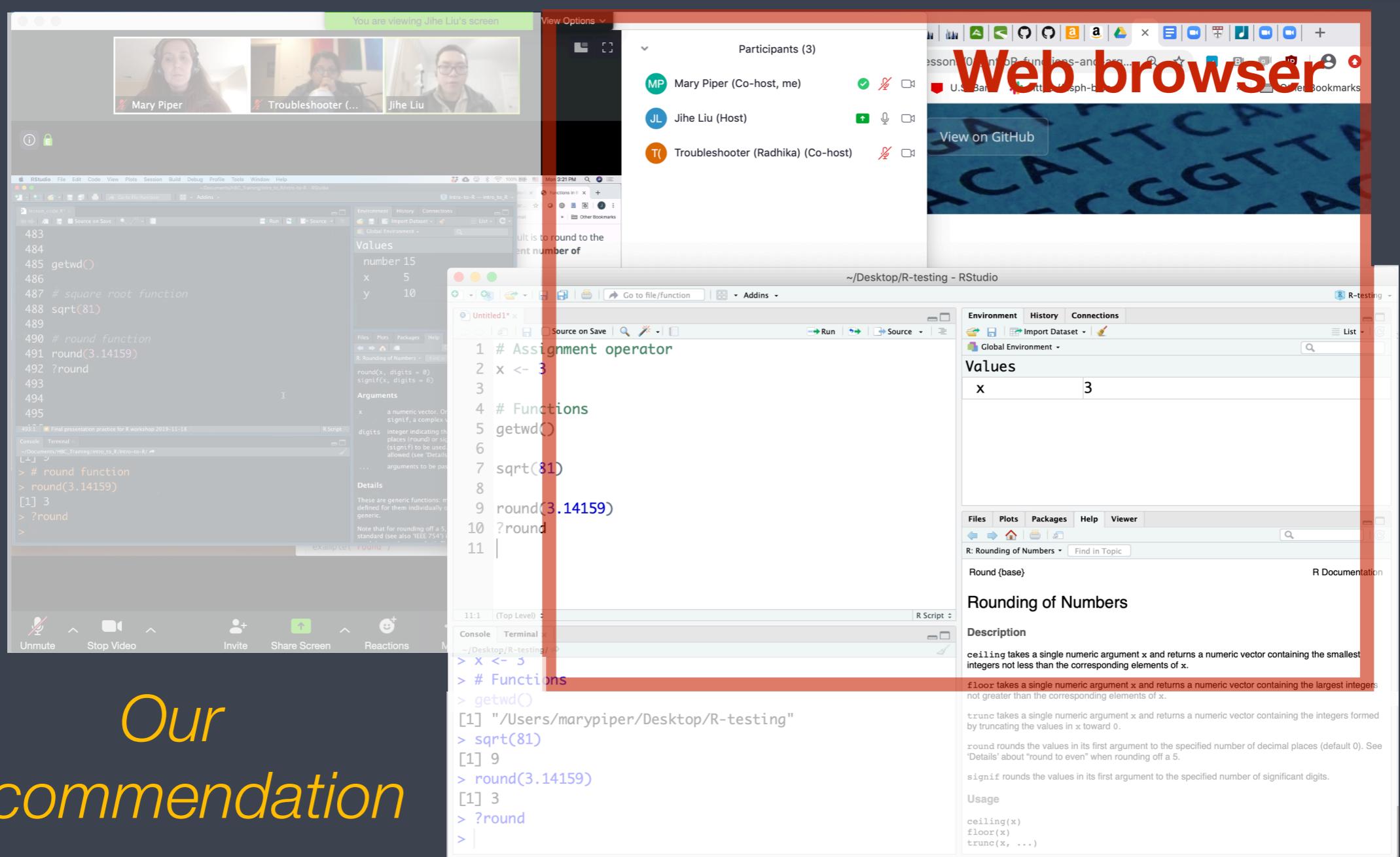


# Single screen & 3 windows?



*Our  
recommendation*

# Single screen & 3 windows?



# Single screen & 3 windows?

The image 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". It contains the following R code:

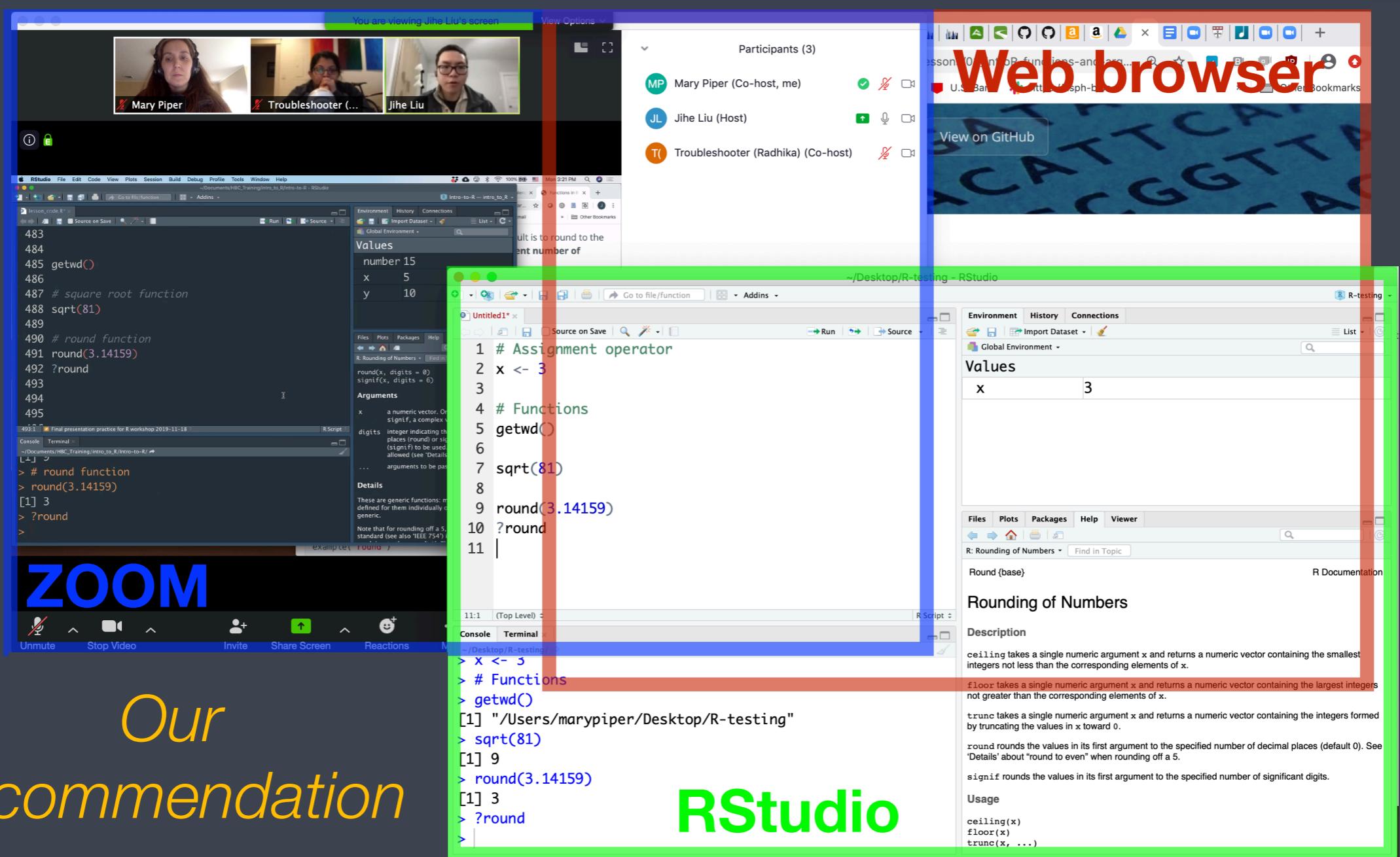
```
483  
484  
485 getwd()  
486  
487 # square root function  
488 sqrt(81)  
489  
490 # round function  
491 round(3.14159)  
492 ?round  
493  
494  
495
```
- Bottom Left Window:** An RStudio session titled "Untitled1\*". It contains the following R code:

```
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 GitHub page for "introR-functions-and-args".
- Middle Right Window:** An RStudio session titled "~/Desktop/R-testing - RStudio". It contains the following R code:

```
11:1 (Top Level) >  
Console Terminal x  
~/Desktop/R-testing/ >  
> x <- 3  
> # Functions  
> getwd()  
[1] "/Users/marypiper/Desktop/R-testing"  
> sqrt(81)  
[1] 9  
> round(3.14159)  
[1] 3  
> ?round  
>
```
- Bottom Right Window:** A detailed view of the "round" function documentation from the R documentation website.

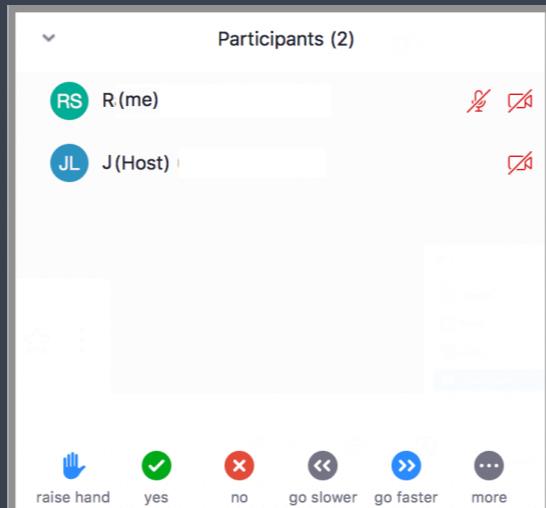
**Our recommendation:** RStudio

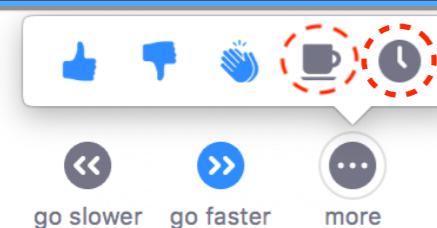
# Single screen & 3 windows?



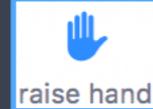
# Odds and Ends (1/2)

- ❖ Quit/minimize all applications that are not required for class
- ❖ Click on “Participants” to open that panel in Zoom



- ▶  = "agree", "I'm all set" (equivalent to a **green post-it**)
- ▶  = "disagree", "I need help" (equivalent to a **red post-it**)
- ▶  If you are away from the computer use the coffee cup or clock icon

# Odds and Ends (2/2)

- ❖ Questions for the presenter?
  - Post the question in the Chat window OR
  - Raise your hand  when the presenter asks for questions
- ❖ Technical difficulties with R or RStudio?
  - 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)

Twitter

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