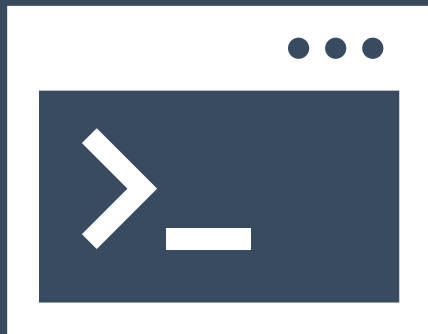


# Introduction to R

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



Harvard Chan Bioinformatics Core



# Introductions!





Shannan Ho Sui  
*Director*



Meeta Mistry  
*Associate Director*



Lorena Pantano  
*Director of Bioinformatics  
Platform*



John Quackenbush  
*Faculty Advisor*



Upen Bhattarai



Heather Wick



Will Gammerdinger



Noor Sohail



Alex Bartlett



Elizabeth  
Partan



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*



Upen Bhattarai



Heather Wick



Will Gammerdinger



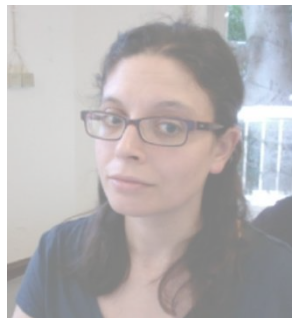
Noor Sohail



Alex Bartlett



Elizabeth  
Partan



Emma Berdan



James Billingsley



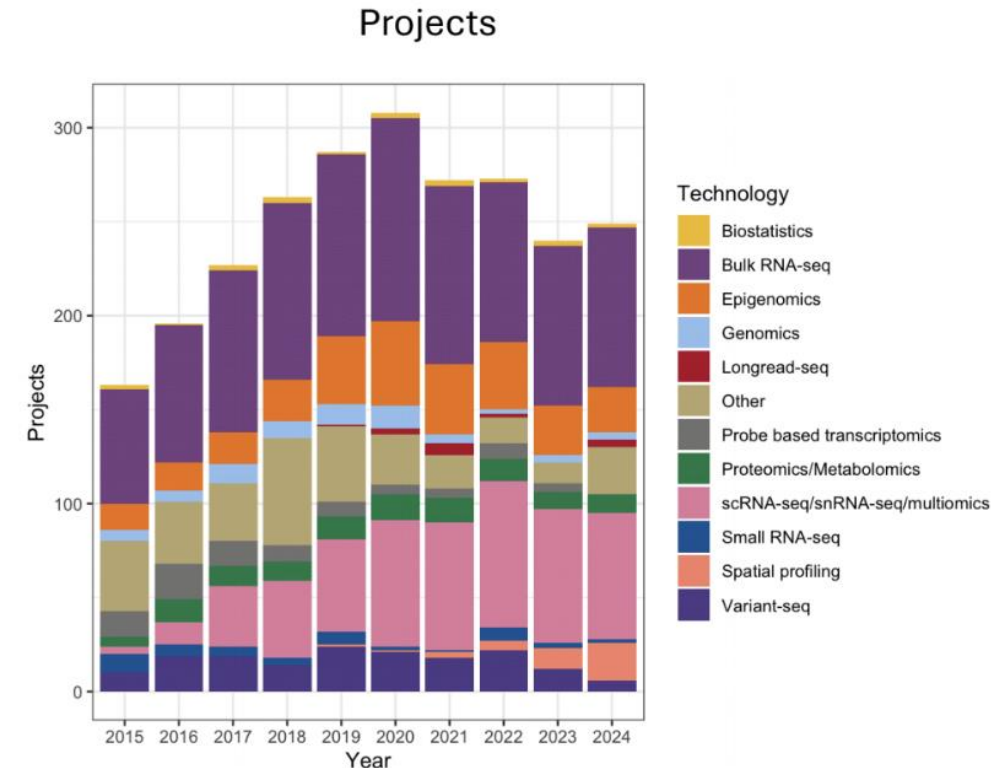
Zhu Zhuo



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

---



**HARVARD**  
**CATALYST**

THE HARVARD CLINICAL  
AND TRANSLATIONAL  
SCIENCE CENTER

---



**HARVARD**  
MEDICAL SCHOOL

# 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>

# 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



**HARVARD**  
**CATALYST**

THE HARVARD CLINICAL  
AND TRANSLATIONAL  
SCIENCE CENTER

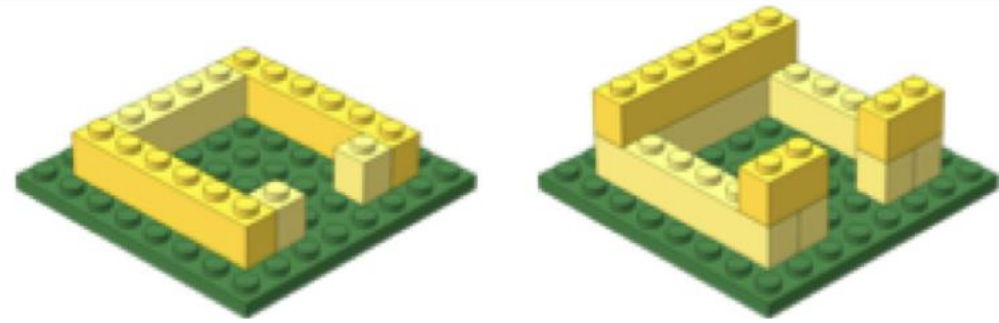


**HARVARD**  
MEDICAL SCHOOL

# **Workshop scope**



# 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



[CRAN](#)  
[Mirrors](#)  
[What's new?](#)  
[Task Views](#)  
[Search](#)

[About R](#)  
[R Homepage](#)  
[The R Journal](#)

[A3](#)  
[abbyyR](#)  
[abc](#)  
[ABCanalysis](#)  
[abc.data](#)  
[abcdeFBA](#)  
[ABCOptim](#)  
[ABCp2](#)  
[abcrf](#)

**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](#)

Accurate, Adaptable, and Accessible Error Metrics for Predictive Models  
Access to Abbyy Optical Character Recognition (OCR) API  
Tools for Approximate Bayesian Computation (ABC)  
Computed ABC Analysis  
Data Only: Tools for Approximate Bayesian Computation (ABC)  
ABCDE\_FBA: A-Biologist-Can-Do-Everything of Flux Balance Analysis with this package  
Implementation of Artificial Bee Colony (ABC) Optimization  
Approximate Bayesian Computational Model for Estimating P2  
Approximate Bayesian Computation via Random Forests

- ❖ Comprehensive R Archive Network
- ❖ 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

# 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 schedule

Introduction to R

Schedule

HBC

GitHub

Contact us

</> Code

## Workshop Schedule

### Day 1

Time	Topic	Instructor
10:00 - 10:30	<a href="#">Workshop Introduction</a>	Heather
10:30 - 11:45	<a href="#">Introduction to R and RStudio</a>	Will
11:45 - 12:00	Overview of self-learning materials and homework submission	Will

### Before the next class:

I. Please **study the contents** and **work through all the code** within the following lessons:

- [R Syntax and Data Structure](#)

# Course materials

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

[Introduction to R](#) [Schedule](#) [HBC](#) [GitHub](#) [Contact us](#)

Day 1:  
[Introduction to R and RStudio](#)

Day 1 Self-learning: >

Day 2 >

Day 2 Self-learning >

Day 3 >

Day 3 Self-learning >

Day 4 >

## Introduction to R and RStudio

AUTHOR  
Mary Piper, Meeta Mistry

PUBLISHED  
September 8, 2017

Approximate time: 45 minutes

### Learning Objectives

- Describe what R and RStudio are.
- Interact with R using RStudio.
- Familiarize various components of RStudio.
- Employ variables in R.

### What is R?

The common misconception is that R is a programming language but in fact it is much more than that. Think of R as an environment for statistical computing and graphics, which brings together a number of features to provide powerful functionality.

</> Code

On this page

[Learning Objectives](#)

[What is R?](#)

[Why use R?](#)

[What is RStudio?](#)

[Creating a new project directory in RStudio](#)

[RStudio Interface](#)

[Organizing your working directory & setting up](#)

[Interacting with R](#)

[The R syntax](#)

[Assignment operator](#)

[Variables](#)

[Interacting with data in R](#)

[Best practices](#)

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

# Single Screen & 3 Windows

The screenshot displays a web browser window with a video conference and a web application. The video conference shows three participants: Mary Piper, Troubleshooter (Radhika), and Jihe Liu. The web application is titled "Introduction To R / Intro-to-R" and shows a code editor with R code, a console with output, and a file explorer. The code includes a simple addition and a multi-line comment. The console shows the output of the code. The file explorer shows a project structure with folders for .Rhistory, data, project.Rproj, scripts, results, and figures.

**Participants (3)**

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

**On this page**

- [Code](#)
- [Learning Objectives](#)

**Introduction To R / Intro-to-R**

File Edit Code View Plots Session Build Debug Profile Tools Help

Intro-to-R.R

```
1 3 + 5
2
3 # Intro to R Lesson
4 # June 3rd, 2025
5
6 # Interacting with R
7
8 ## I am adding 3 and 5. R is fun!
9
10 # Top Level
```

Console

```
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> 3 + 5
[1] 8
> # Intro to R Lesson
> # June 3rd, 2025
> > # Interacting with R
> > ## I am adding 3 and 5. R is fun!
> 3 + 5
[1] 8
>
```

Files

- .Rhistory
- data
- project.Rproj
- scripts
- results
- figures

*Our  
Recommendation*

# Single Screen & 3 Windows

**Zoom**

**Our Recommendation**

**RStudio**

```
1 3 + 5
2
3 # Intro to R Lesson
4 # June 3rd, 2025
5
6 # Interacting with R
7
8 ## I am adding 3 and 5. R is fun!
9
10 3 + 5
11 [1] 8
12 >
```

Console

```
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> 3 + 5
[1] 8
> # Intro to R Lesson
> # June 3rd, 2025
> > # Interacting with R
> > ## I am adding 3 and 5. R is fun!
> 3 + 5
[1] 8
>
```

**Web Browser**

HBC GitHub Contact us

On this page  
Learning Objectives

Code

Introduction To R / Intro-to-R

Environment

Name	Size	Modified
..		
.Rhistory	0 B	May 28, 2025, 4:43 PM
data		
project.Rproj	205 B	May 30, 2025, 2:51 PM
scripts		
results		
figures		

# Single Screen & 3 Windows

You are viewing Jihe Liu's screen

Participants (3)

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

HBC GitHub Contact us

## Web Browser

Code Learning Objectives

Introduction To R / Intro-to-R

File Edit Code View Plots Session Build Debug Profile Tools Help

```
1 3 + 5
2
3 # Intro to R Lesson
4 # June 3rd, 2025
5
6 # Interacting with R
7
8 ## I am adding 3 and 5. R is fun!
9
10 3 + 5
11 [1] 8
12 >
```

Console

```
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> 3 + 5
[1] 8
> # Intro to R Lesson
> # June 3rd, 2025
> # Interacting with R
> ## I am adding 3 and 5. R is fun!
> 3 + 5
[1] 8
>
```

Environment

R - Global Environment

Files Plots Packages Help Viewer Presentation

- .Rhistory 0 B May 28, 2025, 4:43 PM
- data 205 B May 30, 2025, 2:51 PM
- project.Rproj
- scripts
- results
- figures

Our Recommendation

# Single Screen & 3 Windows

The image illustrates a remote collaboration setup. On the left, a video conference window shows three participants: Mary Piper, Troubleshooter (Radhika), and Jihe Liu. The main window displays the Posit Cloud interface, which is divided into three panes. The top pane shows the R script editor with the following code:

```
1 3 + 5
2
3 # Intro to R Lesson
4 # June 3rd, 2025
5
6 # Interacting with R
7
8 ## I am adding 3 and 5. R is fun!
```

The bottom-left pane shows the console output:

```
> 3 + 5
[1] 8
> # Intro to R Lesson
> # June 3rd, 2025
> # Interacting with R
> ## I am adding 3 and 5. R is fun!
```

The bottom-right pane shows the file explorer with the following files:

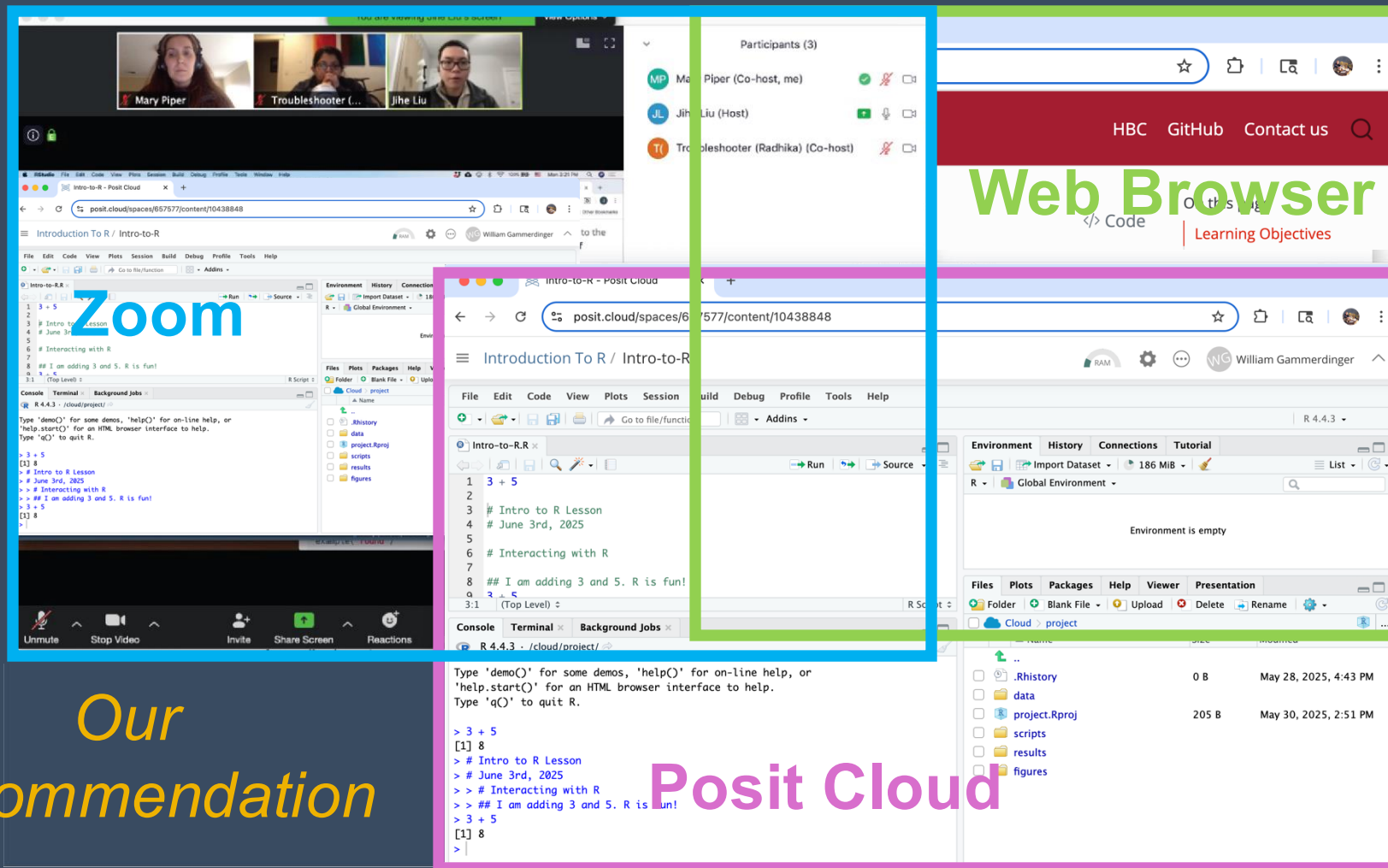
- .Rhistory (0 B, May 28, 2025, 4:43 PM)
- data
- project.Rproj (205 B, May 30, 2025, 2:51 PM)
- scripts
- results
- figures

The Posit Cloud interface also includes a top navigation bar with links to HBC, GitHub, and Contact us, and a sidebar with options for Code and Learning Objectives.

*Our Recommendation*

**Posit Cloud**

# Single Screen & 3 Windows



# 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

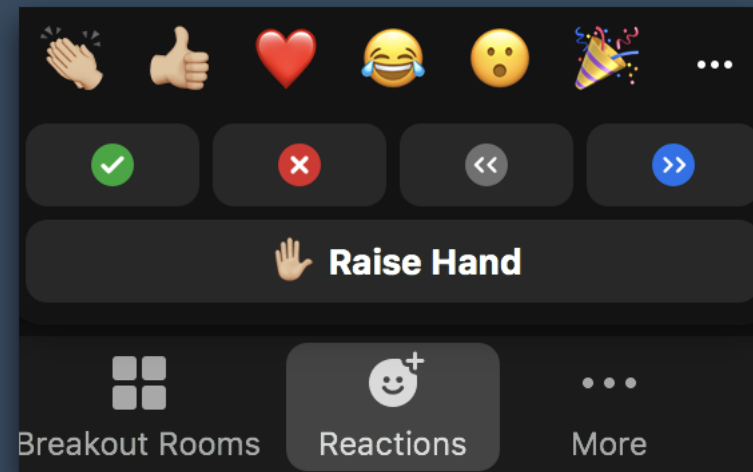
# Odds & Ends

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

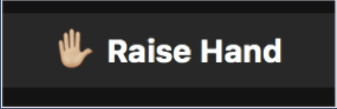
❖ Are you all set?

❖  = "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 Moderator 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)