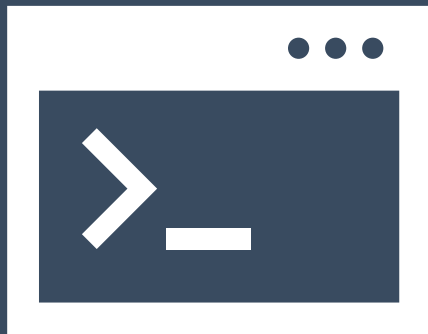


Introduction to Peak Analysis

<https://tinyurl.com/hbc-peak-analysis>



Harvard Chan Bioinformatics Core



Introductions!





Shannan Ho Sui
Director



Lorena Pantano
*Director of Bioinformatics
Platform*



John Quackenbush
Faculty Advisor



Upen Bhattarai



Will Gammerdinger



Noor Sohail



Alex Bartlett



Elizabeth
Partan



Emma Berdan



Zhu Zhuo



James Billingsley



Shannan Ho Sui
Director



Lorena Pantano
*Director of Bioinformatics
Platform*



John Quackenbush
Faculty Advisor



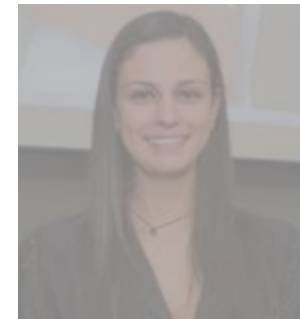
Upen Bhattarai



Will Gammerdinger



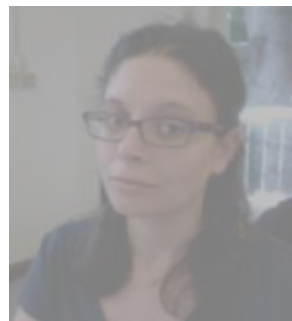
Noor Sohail



Alex Bartlett



Elizabeth
Partan



Emma Berdan



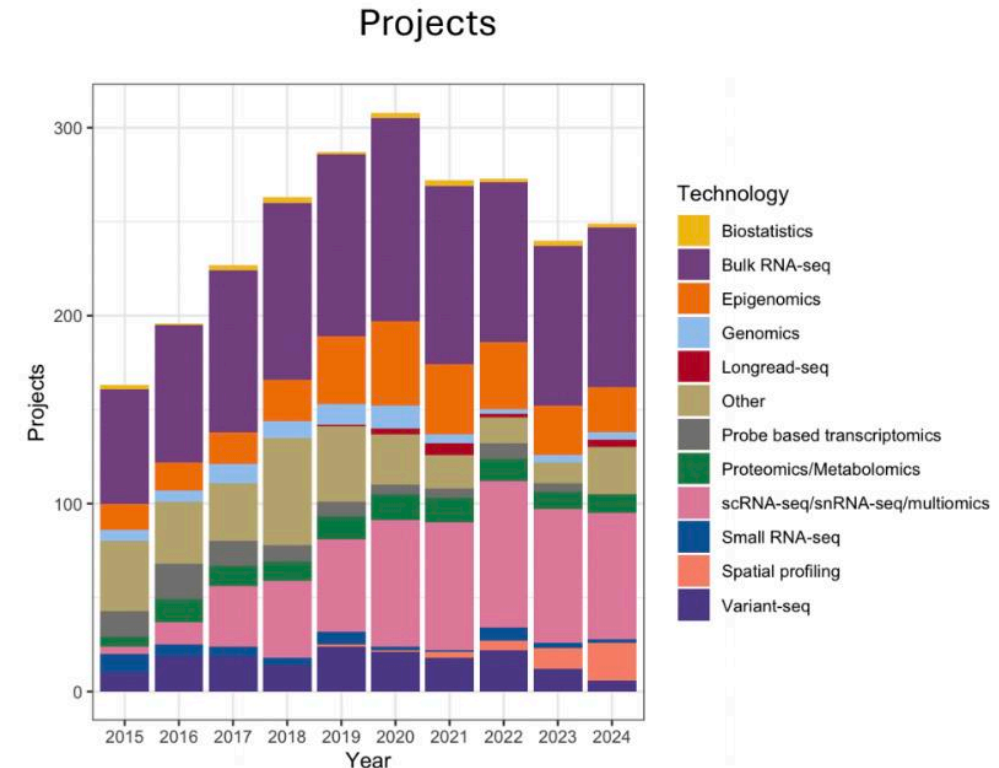
Zhu Zhuo



James Billingsley

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



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 & Peak Analysis

- ❖ Bulk RNA-seq

- ❖ Differential Gene Expression (Bulk & scRNA)

- ❖ 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



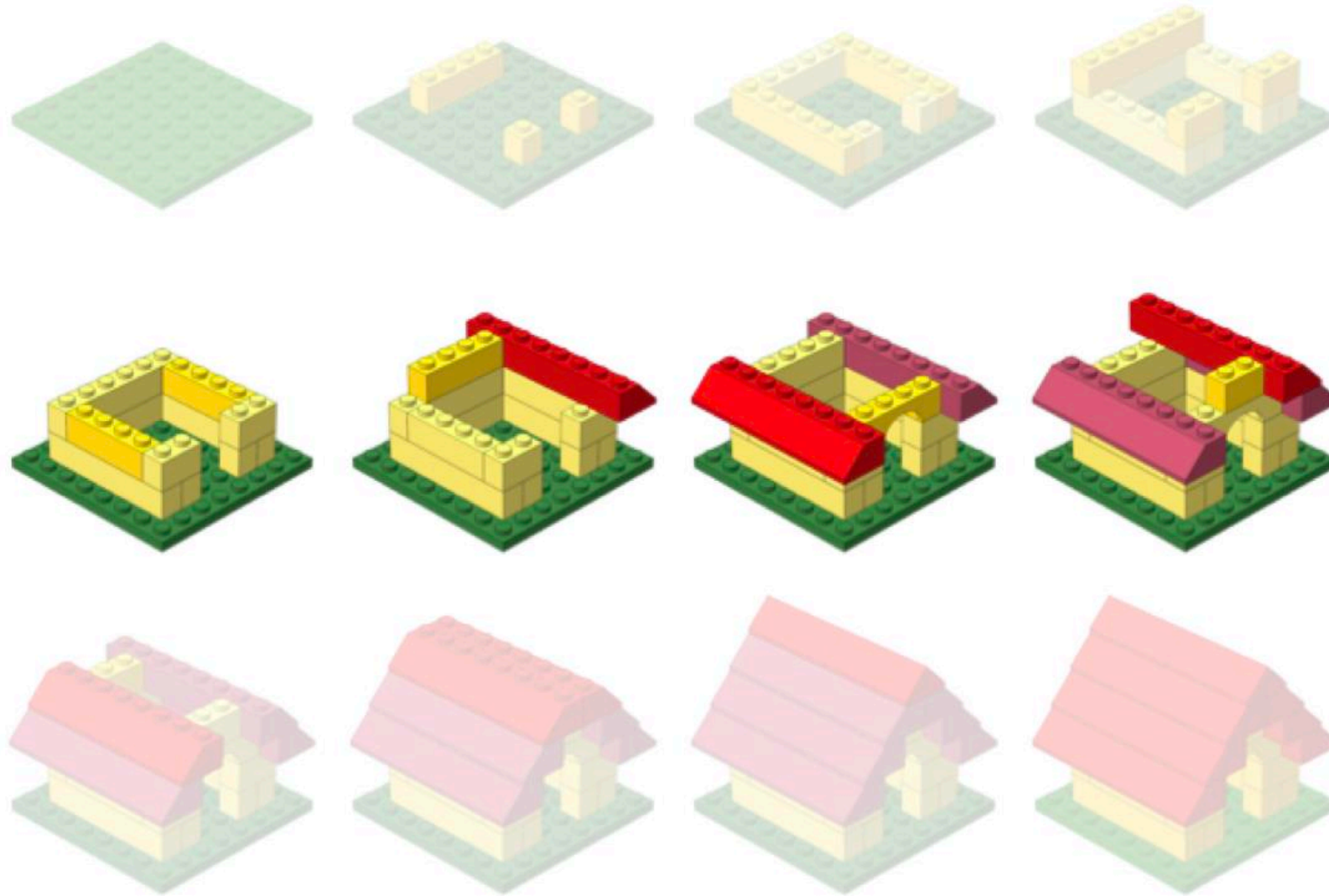
HARVARD
CATALYST

THE HARVARD CLINICAL
AND TRANSLATIONAL
SCIENCE CENTER



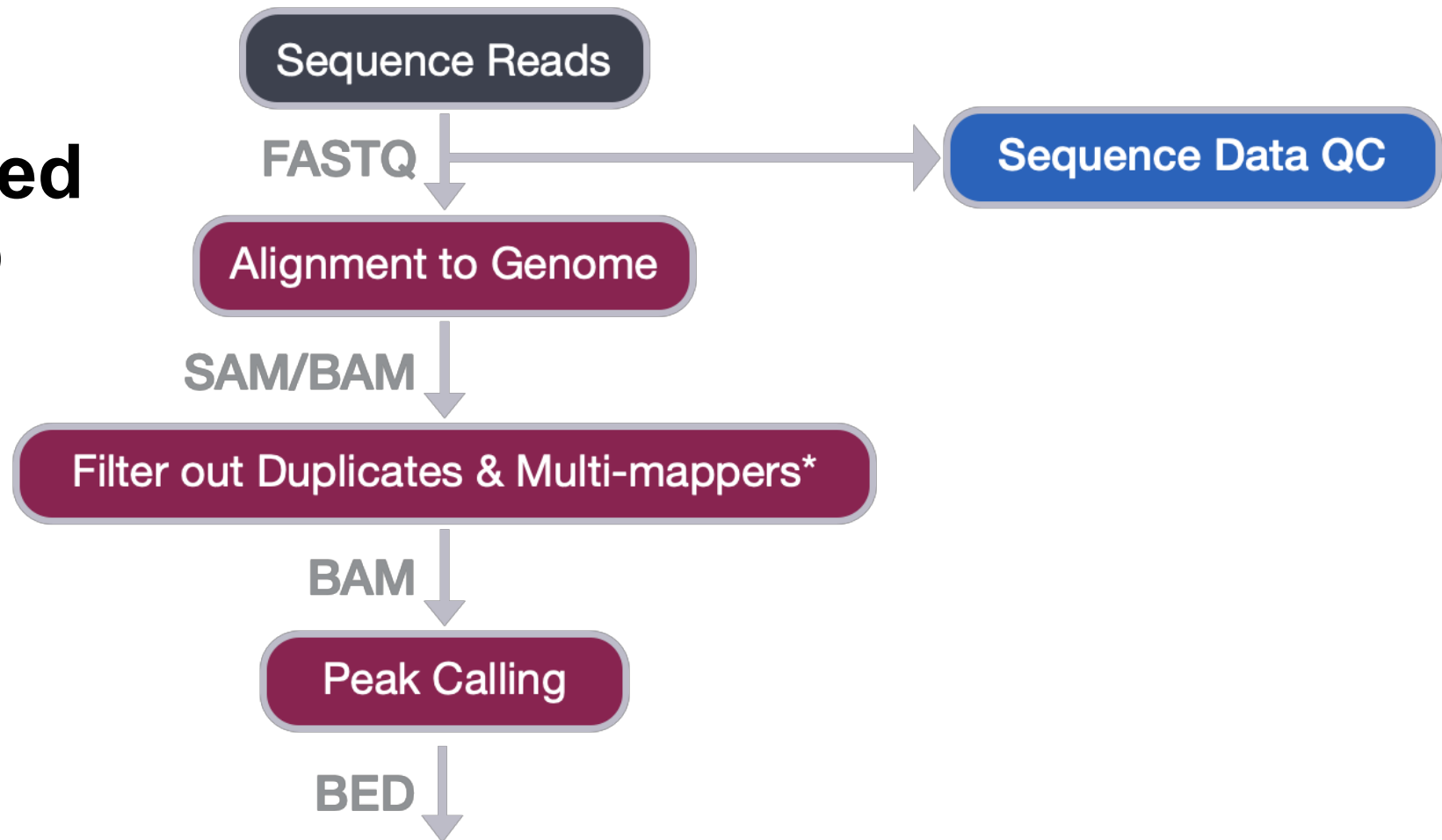
HARVARD
MEDICAL SCHOOL

Workshop scope

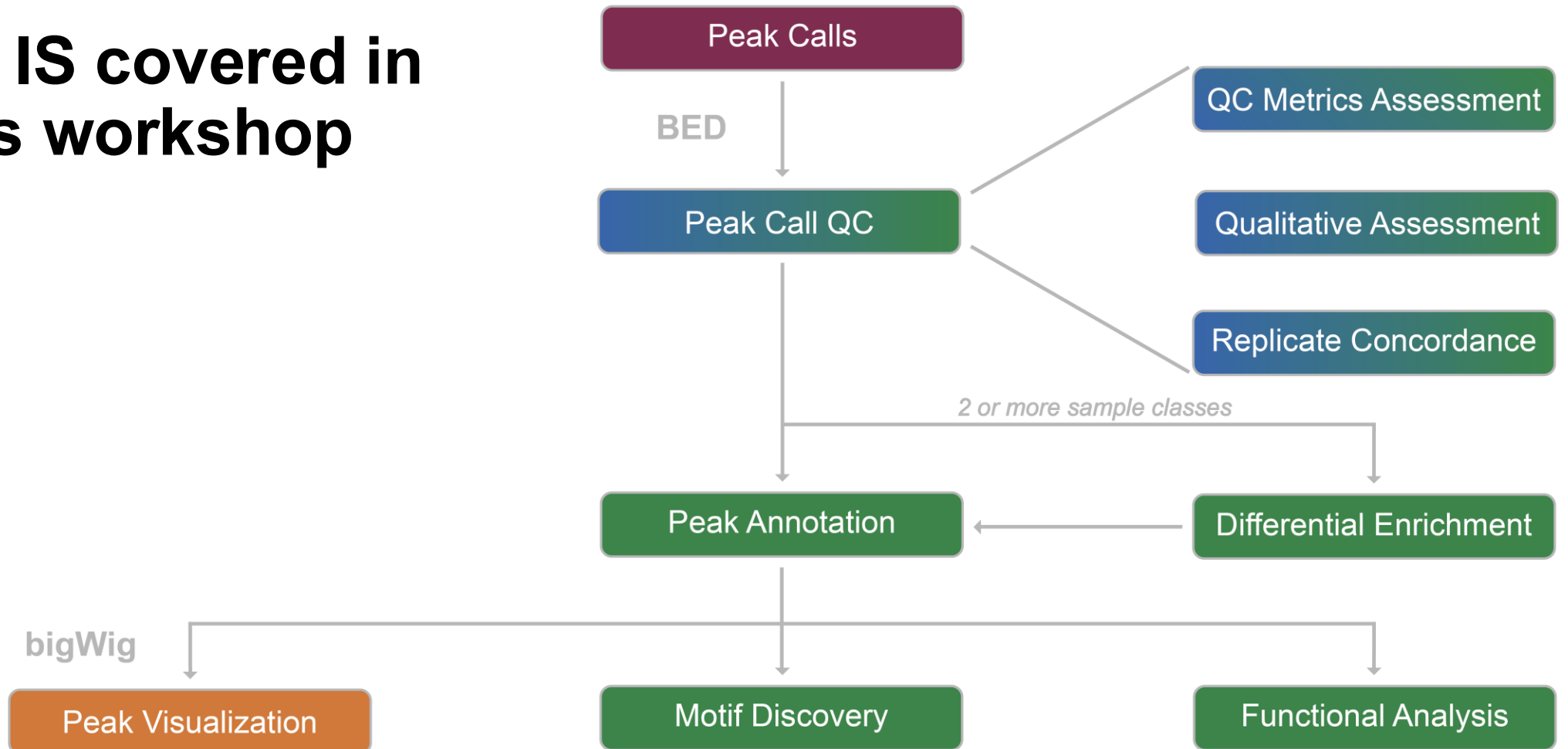


Bioinformatic Data Analysis

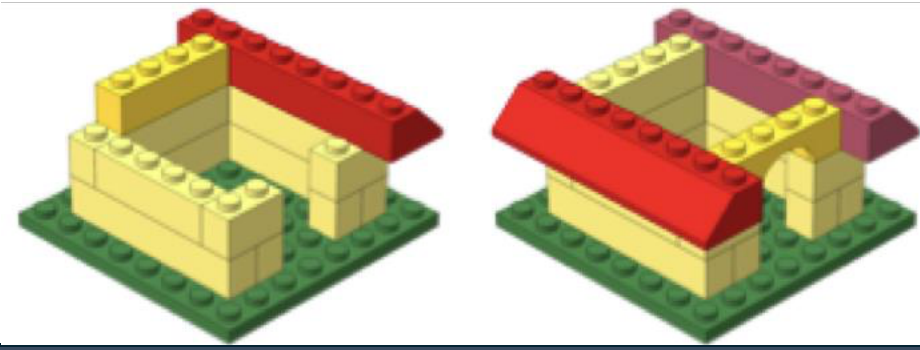
What is NOT covered in this workshop



What IS covered in this workshop



Workshop Scope



- ❖ Describe peak data and different file formats generated from peak calling algorithms
- ❖ Assess various metrics used to assess the quality of peak calls
- ❖ Compare peak calls across samples within a dataset
- ❖ Create visualizations to evaluate peak annotations
- ❖ Evaluate differentially enriched regions between two sample groups

Logistics



Course schedule

Day 1

Time	Topic	Instructor
09:30 - 09:45	Workshop Introduction	Will
09:45 - 10:15	Pre-reading discussion	Upen
10:15 - 11:00	Understanding peaks and peak file formats	Will
11:00- 11:05	Break	
11:05 - 12:00	Assessing peak quality metrics	Upen

Before the next class:

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

1. [Assessing sample similarity and identifying potential outliers](#)

[Click here for a preview of this lesson](#)

2. [Concordance across replicates using peak overlaps](#)

[Click here for a preview of this lesson](#)

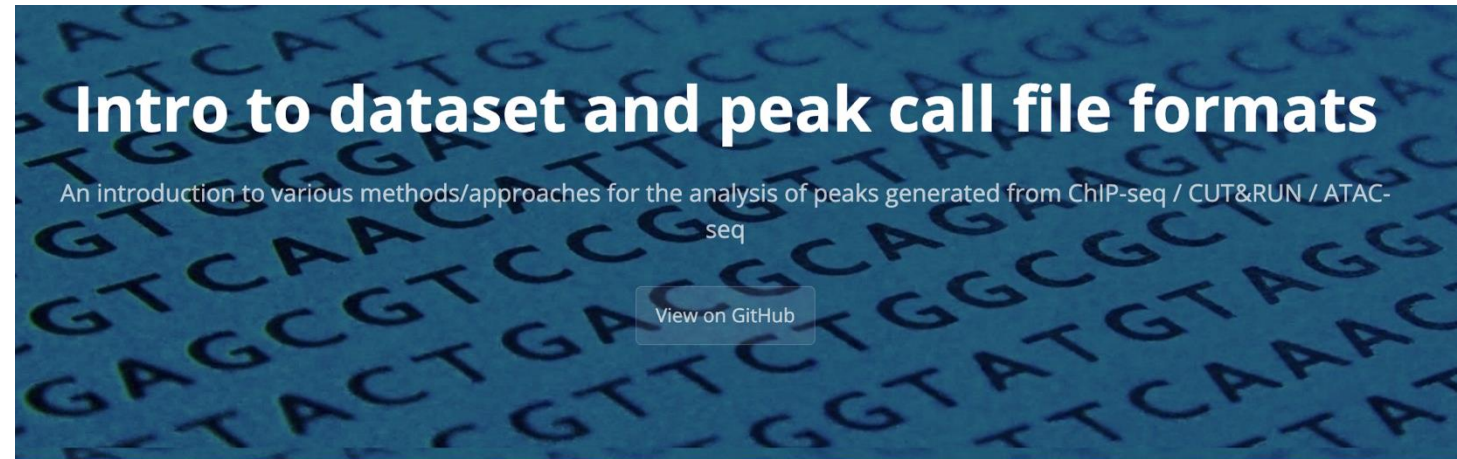
3. **Complete the exercises:**

- Each lesson above contains exercises; please go through each of them.
- Copy over your solutions into the [Google Forms](#) the **day before the next class**.

<https://tinyurl.com/hbc-peak-analysis>

Course materials

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



Contributors: Heather Wick, Upendra Bhattarai, Meeta Mistry, Will Gammerdinger

Approximate time: 40 minutes

Learning Objectives

In this lesson, we will:

- Explain the dataset and the biological context
- Define peaks as genomic coordinate data
- Describe file formats for peak data

Introduction to the dataset

For this workshop we will be working with ChIP-seq data from a publication in Neuron by *Baizabal et al., 2018* ¹.

<https://tinyurl.com/hbc-peak-analysis>

Single Screen & 3 Windows

The image shows a web browser window displaying a video conference and a software interface. The video conference is titled "You are viewing Jihe Liu's screen" and shows three participants: Mary Piper, Troubleshooter (Radhika), and Jihe Liu. The software interface is the RStudio IDE, showing a script editor with R code, a console with output, and a file explorer. The browser address bar shows "posit.cloud/spaces/657577/content/10438848".

Participants (3)

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

Introduction To R / Intro-to-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 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
>
```

Files

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

*Our
Recommendation*

Single Screen & 3 Windows

Zoom

Participants (3)

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

Introduction To R / Intro-to-R

File Edit Code View Plots Session Build Debug Profile Tools Help

Environment History Connections

Files Plots Packages Help View

Console

```
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 is empty

Files Plots Packages Help Viewer Presentation

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		

Our Recommendation

```
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
>
```

Single Screen & 3 Windows

Web Browser

HBC GitHub Contact us

Code Learning Objectives

Our Recommendation

```
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 >
13 > # Intro to R Lesson
14 > # June 3rd, 2025
15 > # Interacting with R
16 > ## I am adding 3 and 5. R is fun!
17 > 3 + 5
18 [1] 8
19 >
```

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

The image illustrates a recommendation for using Posit Cloud. It is divided into two main sections: 'Our Recommendation' and 'Posit Cloud'.

Our Recommendation

This section features a video conference interface. At the top, it says "You are viewing Jihe Liu's screen". Below this, there are three video feeds of participants: Mary Piper, Troubleshooter (Radhika), and Jihe Liu. The interface also includes a "Participants (3)" list on the right, showing Mary Piper (Co-host, me), Jihe Liu (Host), and Troubleshooter (Radhika) (Co-host). At the bottom, there are controls for "Unmute", "Stop Video", "Invite", "Share Screen", and "Reactions".

Posit Cloud

This section shows a screenshot of the Posit Cloud workspace. The browser address bar displays the URL: `posit.cloud/spaces/657577/content/10438848`. The workspace title is "Introduction To R / Intro-to-R". The interface includes a menu bar with options like File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. The main area shows an 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 console output shows the results of the R script execution:

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

The file explorer on the right shows the project structure:

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

The environment panel shows "Environment is empty". The bottom status bar indicates the R version is 4.4.3 and the current directory is `/cloud/project/`.

Single Screen & 3 Windows

Zoom

Web Browser

Posit Cloud

Our Recommendation

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
>
```

File 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		

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

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

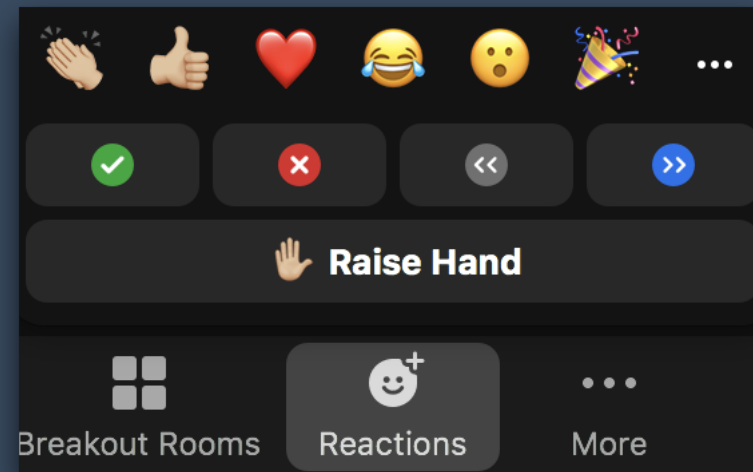
Odds & Ends

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

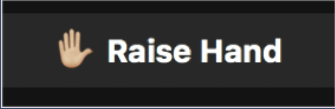
❖ Are you all set?

❖  = "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
- ❖ *HBC consulting:* bioinformatics@hsph.harvard.edu