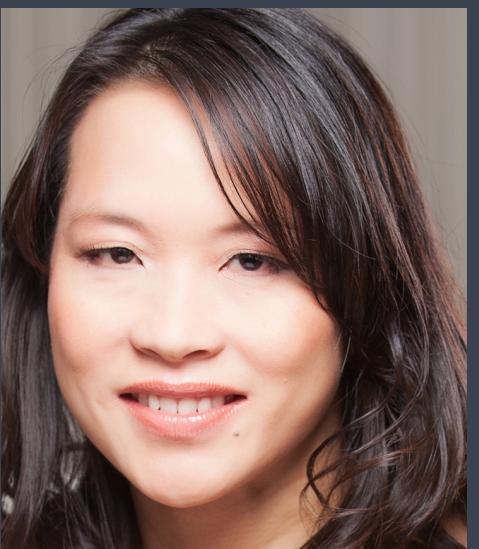


# NGS data analysis course: Winter 2016

Harvard Chan Bioinformatics Core (HBC)

<intro to core>

Harvard Chan Bioinformatics Core



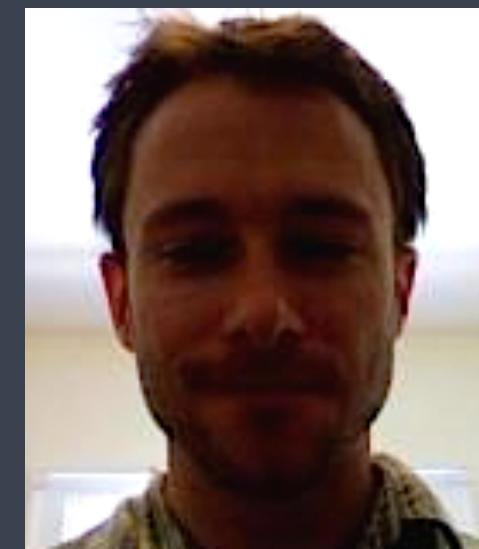
Shannan Ho Sui



John Hutchinson



Brad Chapman



Rory Kirchner



Meeta Mistry



Radhika Khetani



Mary Piper



Lorena Pantano



Oliver Hofmann



Peter Kraft

**Harvard Chan Bioinformatics Core**

# Services offered by HBC

## Consulting:

RNA-seq, small RNA-seq and ChIP-seq

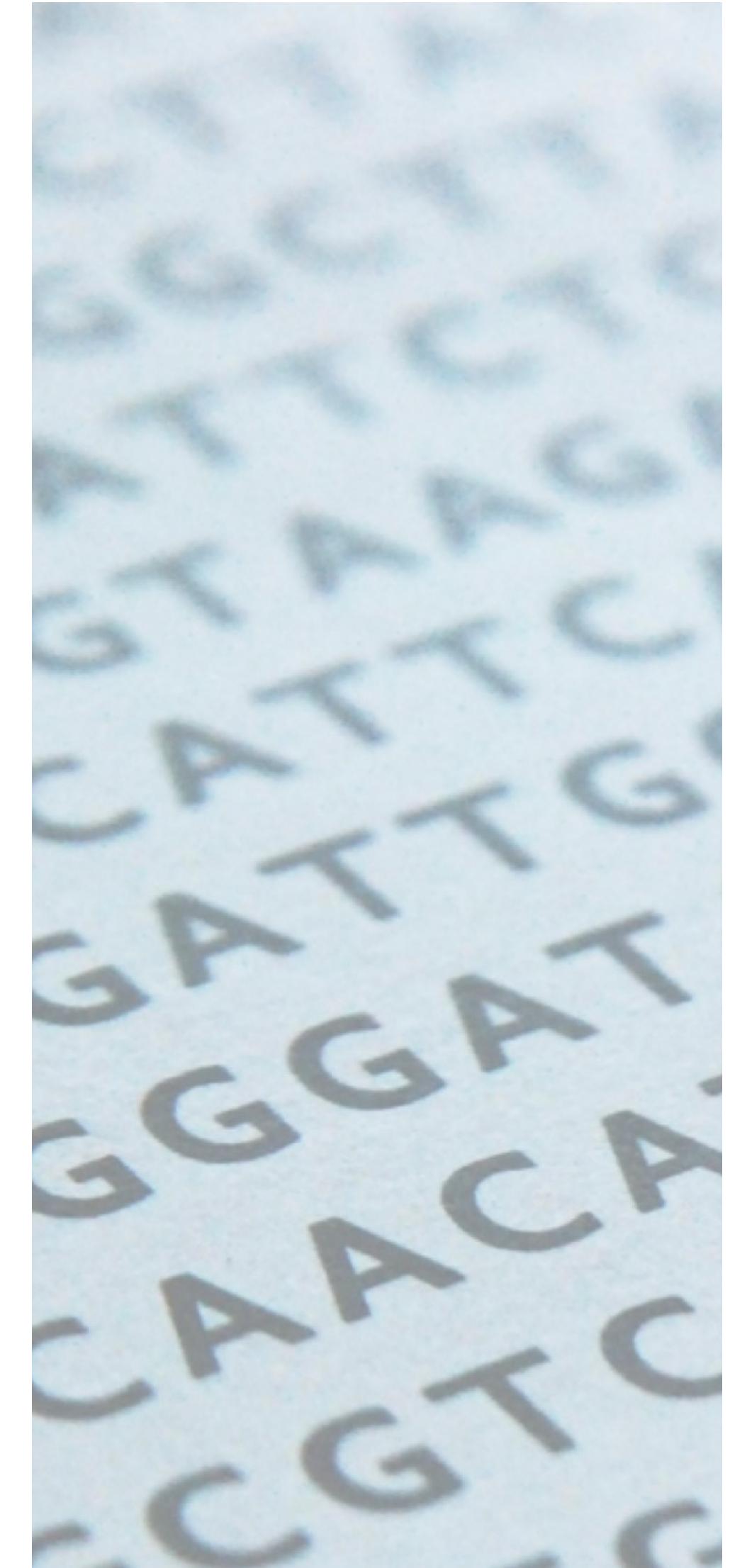
Genome-wide methylation

WGS, resequencing, exome-seq and structural variation

Gene expression arrays (microarrays)

Functional enrichment

Grant support



# Services offered by HBC

## Consulting:

RNA-seq, small RNA-seq and ChIP-seq

Genome-wide methylation

WGS, resequencing, exome-seq and structural variation

Gene expression arrays (microarrays)

Functional enrichment

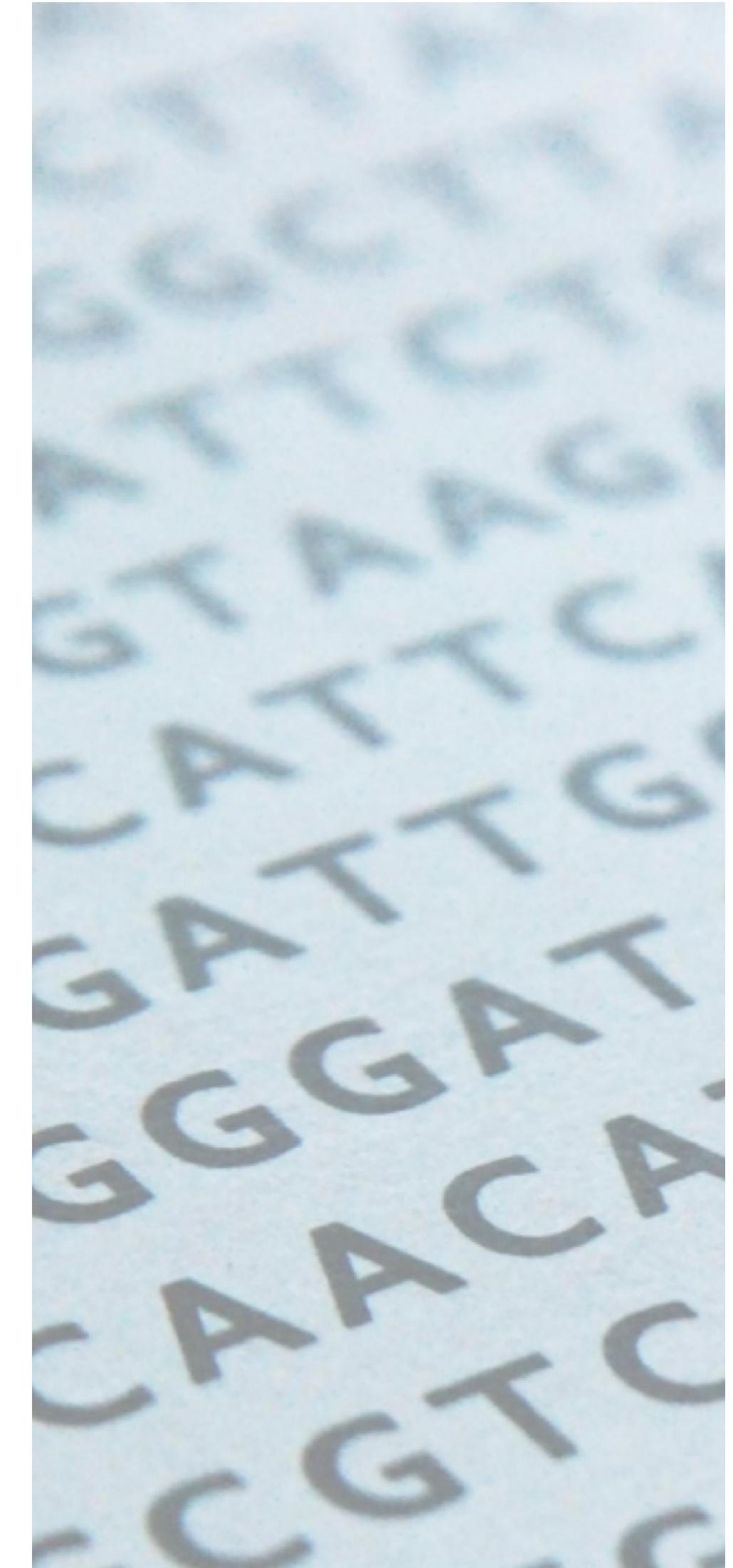
Grant support

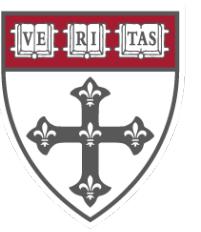
## NGS-focused bioinformatics training:

Galaxy-based NGS analysis, Introductory and intermediate R,

Introductory Python, Introduction to Unix and HPC,

In-depth courses, and other.



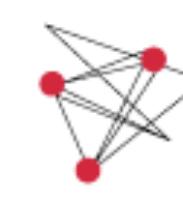


**HARVARD**  
**T.H. CHAN**  
SCHOOL OF PUBLIC HEALTH

NIEHS / CFAR  
Bioinformatics  
Core

**HSCI**  
HARVARD STEM CELL  
INSTITUTE

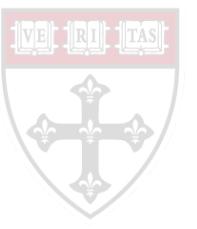
Center for  
Stem Cell  
Bioinformatics

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

Harvard  
Catalyst  
Bioinformatics  
Consulting

 **HARVARD**  
MEDICAL SCHOOL

HMS  
Tools &  
Technology  
  
Harvard  
NeuroDiscovery  
Center



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

NIEHS  
Bioinformatics  
Core

**HSCI**  
HARVARD STEM CELL  
INSTITUTE

Center for  
Stem Cell  
Bioinformatics

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

Harvard  
Catalyst  
Bioinformatics  
Consulting

 **HARVARD**  
MEDICAL SCHOOL

HMS  
Tools &  
Technology  
  
Harvard  
NeuroDiscovery  
Center

# Want more information?

Consult email: [bioinformatics@hsph.harvard.edu](mailto:bioinformatics@hsph.harvard.edu)

Training email: [hbctraining@hsph.harvard.edu](mailto:hbctraining@hsph.harvard.edu)

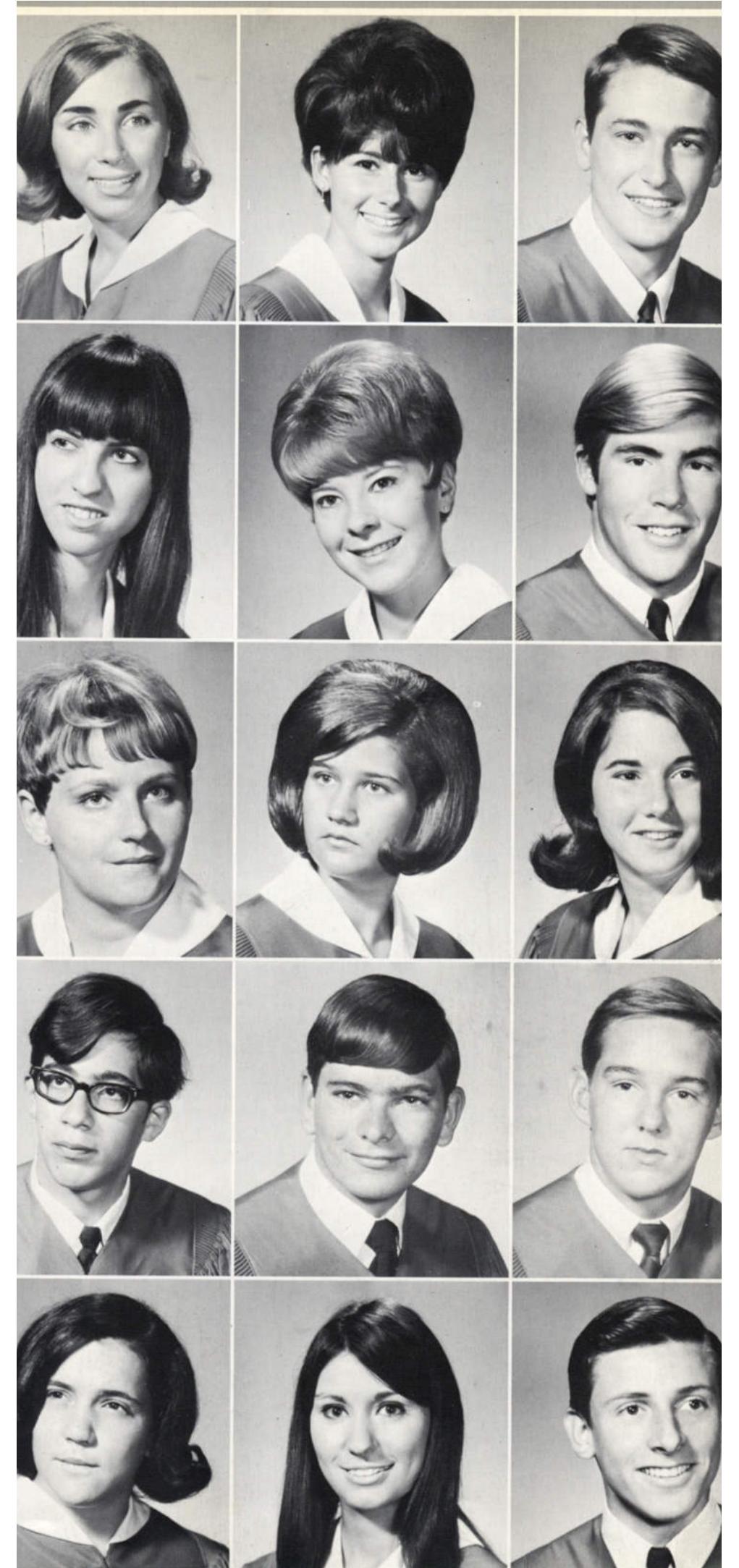
Website: <http://bioinformatics.sph.harvard.edu>

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



</intro to core>

# Course: Introductions!





Meeta Mistry



Radhika Khetani



Mary Piper

## Harvard Chan Bioinformatics Core's teaching team

Boswell, Sarah

Chatterjee, Nirmalya

Chopra, Sameer

de Esch, Celine

Doupe, David

Ettou, Sandrine

Hu, Dan

Kathrein, Katie

Lobbardi, Riadh

Malleshaiah, Mohan

Paschini, Margherita

Renthal, William

Rood, Benjamin

Rupaimoole, Rajesha

Scheffer, Deborah

Shah, Manasvi

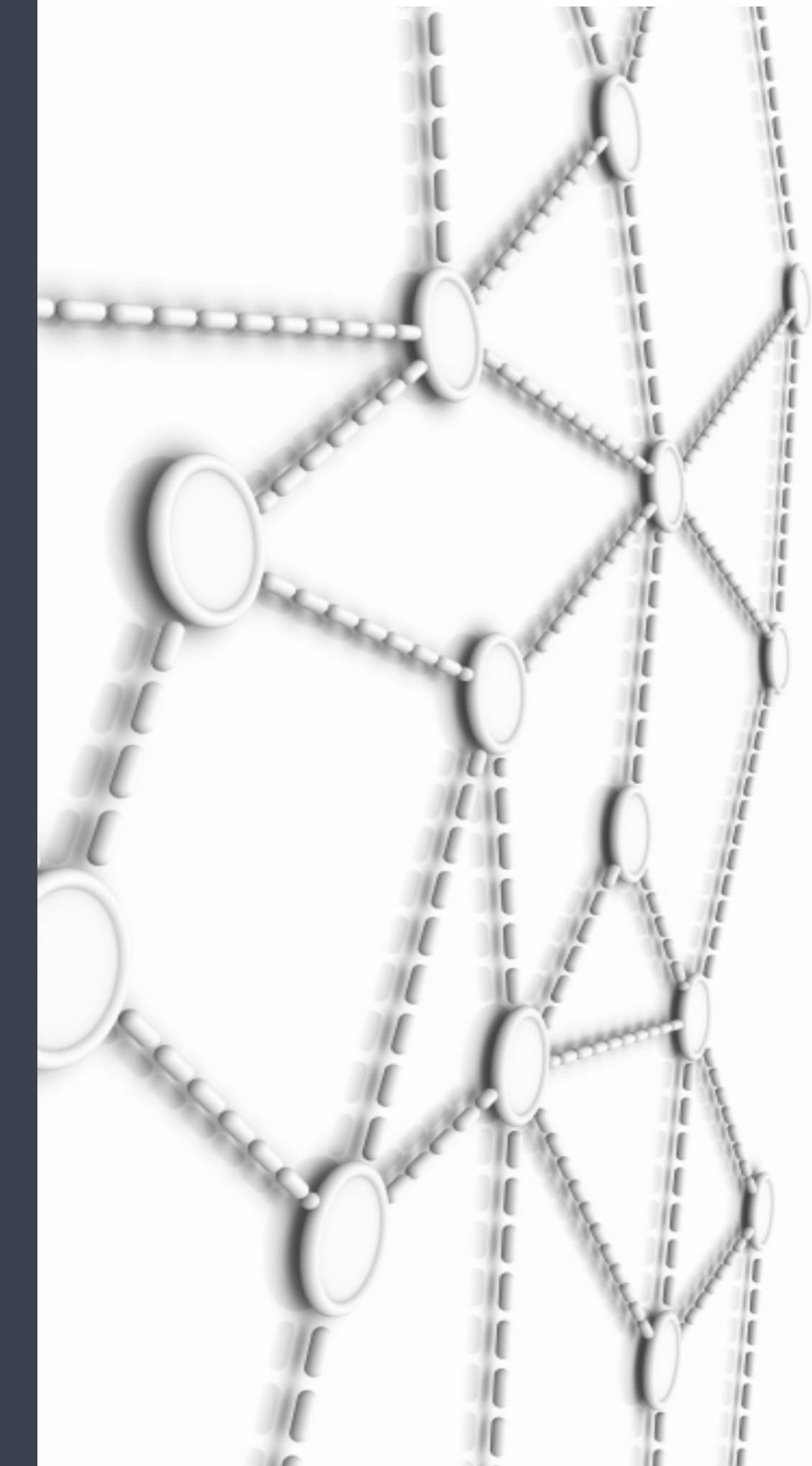
Tan, Catherine

Zaborowski, Mikolaj

Zerbato, Madeleine

Zuccaro, Emanuela

# Course: Learning Objectives



- ✓ Comprehend the nature of Next-Generation Sequencing (NGS) data
  - ◆ Multiple technologies
  - ◆ Caveats
  - ◆ Options and strategies

Learning objectives

- ✓ Comprehend the nature of Next-Generation Sequencing (NGS) data
  - ◆ Multiple technologies
  - ◆ Caveats
  - ◆ Options and strategies
- ✓ Understand how tools and workflows for NGS-based analysis work

Learning objectives

- ✓ Comprehend the nature of Next-Generation Sequencing (NGS) data
  - ◆ Multiple technologies
  - ◆ Caveats
  - ◆ Options and strategies
- ✓ Understand how tools and workflows for NGS-based analysis work
- ✓ Utilize these tools and workflows
  - ◆ Big data = Big Computational Requirements; what does this really entail?
  - ◆ UNIX command-line interface
  - ◆ R

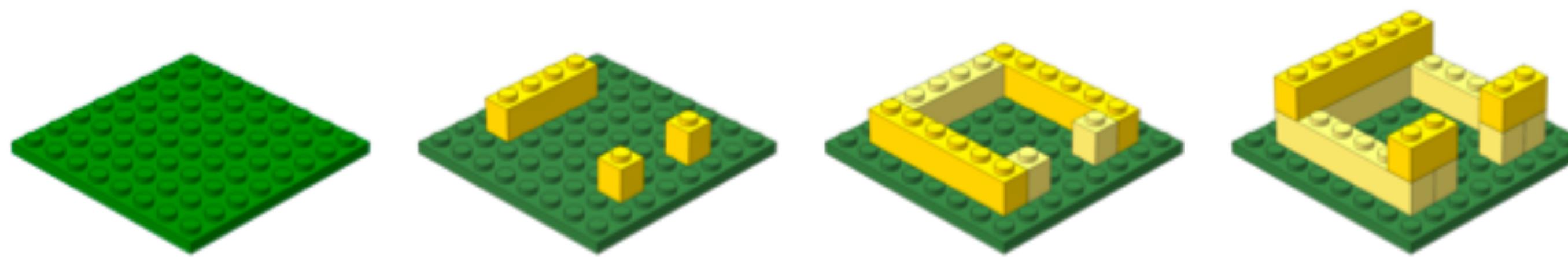
Learning objectives

- ✓ Implement best practices
  - ◆ Experimental design
  - ◆ Quality control and Assessment
  - ◆ Reproducibility

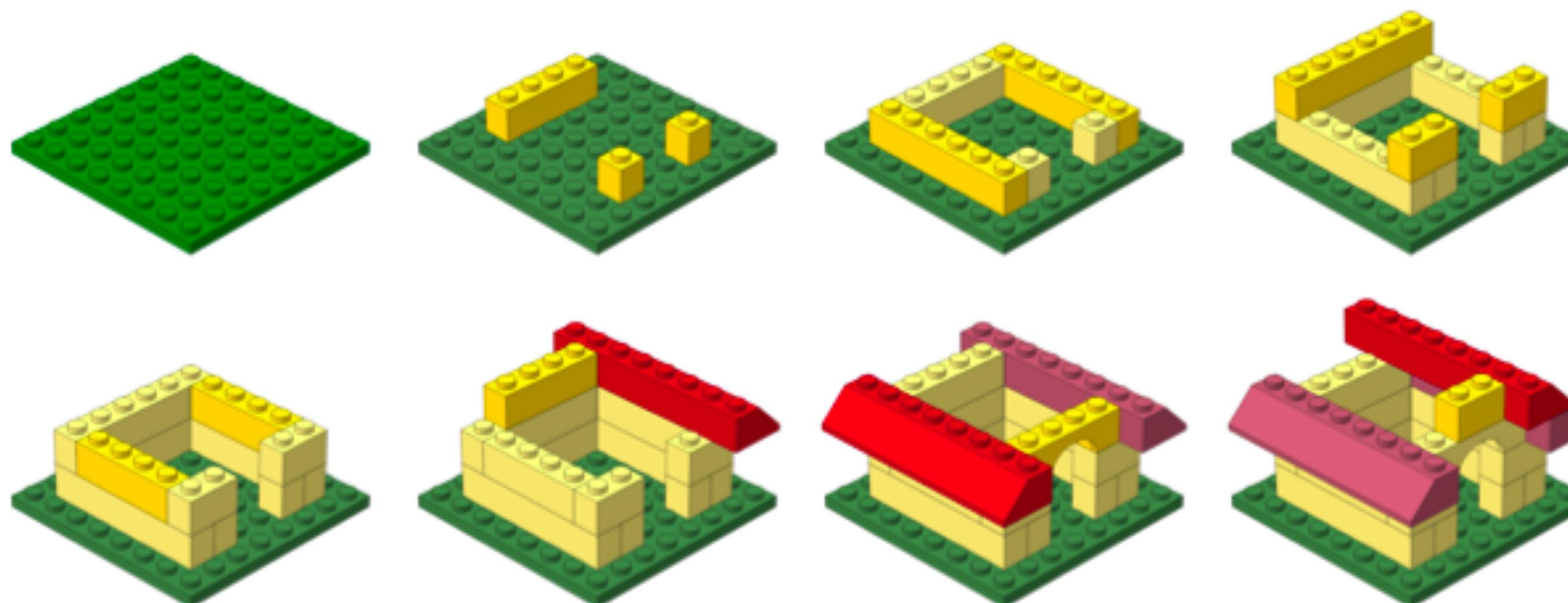
Learning objectives

- ✓ Implement best practices
  - ◆ Experimental design
  - ◆ Quality control and Assessment
  - ◆ Reproducibility
- ✓ Become a resource for your group

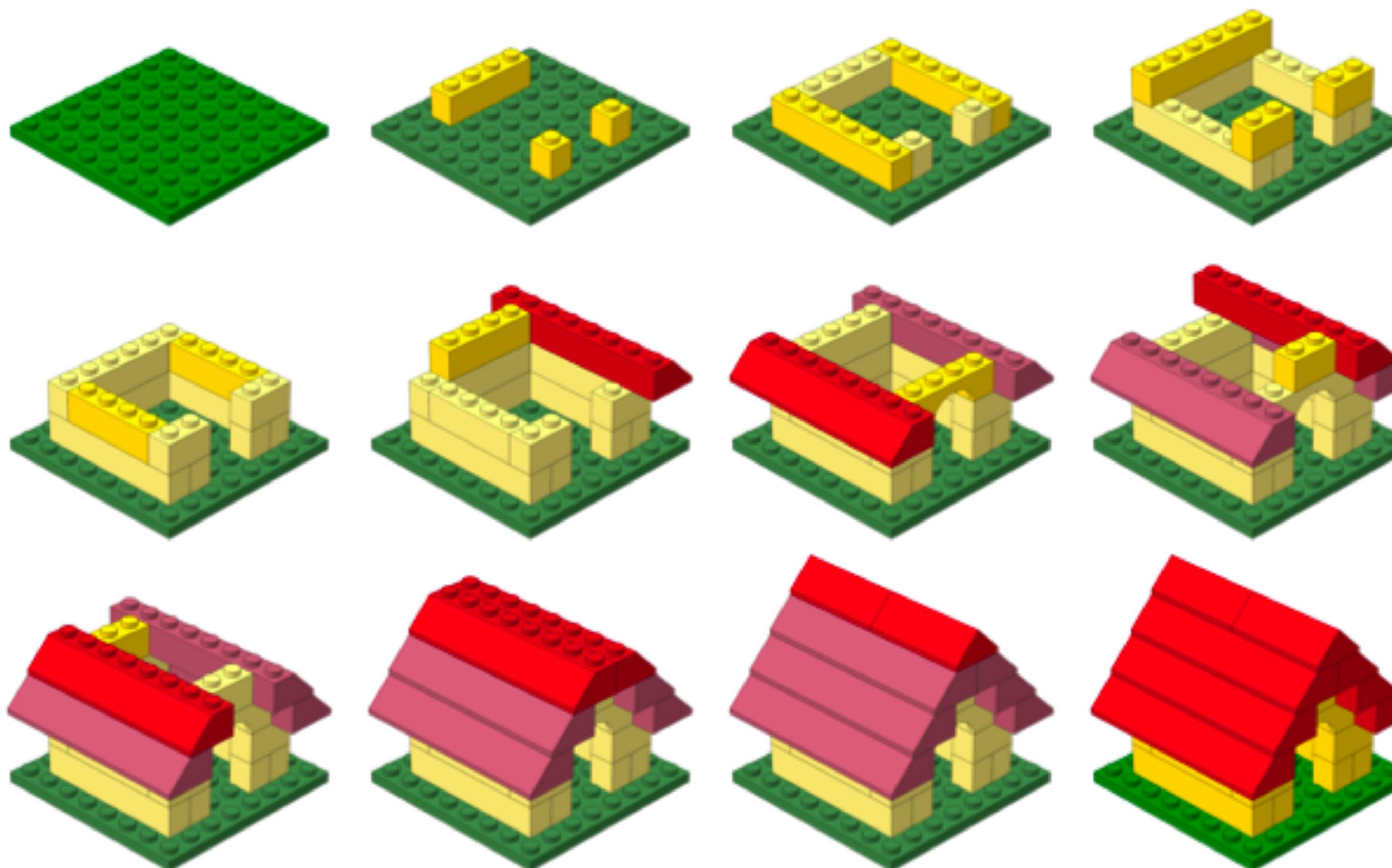
Learning objectives



# Building blocks



# Building blocks



# Building blocks

# Course: Logistics etc.



<b>Dates</b>	<b>Days</b>	<b>Session</b>	<b>Countway Room (5th floor)</b>	<b>Time</b>	<b>Wiki</b>
Feb 8th and 9th	Monday+Tuesday	Session I	Minot room / Ware room	9am - 5pm	<a href="http://tinyurl.com/hbc-session1">http://tinyurl.com/hbc-session1</a>
TBD	Thursday/Friday	Office hours	TBD	2pm - 4pm	
Feb 16th and 17th	Tuesday+Wednesday	Session II	Ballard room	9am - 5pm	<a href="http://tinyurl.com/hbc-session2">http://tinyurl.com/hbc-session2</a>
TBD	Thursday/Friday	Office hours	TBD	2pm - 4pm	
Feb 22nd and 23rd	Monday+Tuesday	Session III	Ware room / Minot room	9am - 5pm	<a href="http://tinyurl.com/hbc-session3">http://tinyurl.com/hbc-session3</a>
TBD	Thursday/Friday	Office hours	TBD	2pm - 4pm	
Feb 29th and March 1st	Monday+Tuesday	Session IV	Minot room	9am - 5pm	<a href="http://tinyurl.com/hbc-session4">http://tinyurl.com/hbc-session4</a>
TBD	Thursday/Friday	Office hours	TBD	2pm - 4pm	
March 7th and 8th	Monday+Tuesday	Session V	Ware room	9am - 5pm	<a href="http://tinyurl.com/hbc-session5">http://tinyurl.com/hbc-session5</a>

Optional: 2-day session on Variant calling in March or April

Dates	Days	Session	Countway Room (5th floor)	Time	Wiki
Feb 8th and 9th	Monday+Tuesday	Session I	Minot room / Ware room	9am - 5pm	<a href="http://tinyurl.com/hbc-session1">http://tinyurl.com/hbc-session1</a>
TBD	Thursday/Friday	Office hours	TBD	2pm - 4pm	
Feb 22nd and 23rd	Monday+Tuesday	Session III	Ware room / Minot room	9am - 5pm	<a href="http://tinyurl.com/hbc-session2">http://tinyurl.com/hbc-session2</a>
Feb 22nd and 23rd	Monday+Tuesday	Session III	Ware room / Minot room	9am - 5pm	<a href="http://tinyurl.com/hbc-session3">http://tinyurl.com/hbc-session3</a>
TBD	Thursday/Friday	Office hours	TBD	2pm - 4pm	
Feb 29th and March 1st	Monday+Tuesday	Session IV	Minot room	9am - 5pm	<a href="http://tinyurl.com/hbc-session4">http://tinyurl.com/hbc-session4</a>
TBD	Thursday/Friday	Office hours	TBD	2pm - 4pm	
March 7th and 8th	Monday+Tuesday	Session V	Ware room	9am - 5pm	<a href="http://tinyurl.com/hbc-session5">http://tinyurl.com/hbc-session5</a>
Optional: 2-day session on Variant calling in March or April					

<http://tinyurl.com/hbc-course-main>

- Log in with either your HUID, HarvardKey or eCommons credentials
- To see all accessible pages use the collapsible menu on the left side

Course wiki page

<http://tinyurl.com/hbc-course-main>

Harvard Wiki Spaces Browse Create

NGS Data Analysis Course, Spring 2016

Bioinformatics Training at Harvard Chan Bioinformatics Core / Bioinformatics Training at HBC

Created by Mary Piper, last modified on Jan 13, 2016

**Welcome to the NGS Data Analysis Course, Spring 2016!**

We are excited to have you join us for the NGS Data Analysis Course! The course starts **Monday, February 8th at 9:00am in the Countway Library, Minot Room (506)**. We will be using this course wiki to provide you with current announcements, the up-to-date course schedule, the course materials, and a discussion space.

**General Course Information**

The course is comprised of five mandatory hands-on sessions and one optional session, with each session running for 2 consecutive days. **Attendance is required for all mandatory sessions**, and all sessions will be held from **9am to 5pm on the 5th floor of the Countway Library** (see map). Please see the course agenda or session pages to find the specific meeting rooms for each day of the course. While we will provide snacks and coffee/tea during breaks, we will not be providing lunch, so please plan accordingly.

Homework assignments for practicing key concepts will be assigned at the end of each session and will be due at the beginning of the following session (homework assignments will take no more than roughly 4 hours to complete). **Office hours will be held on Thursdays and/or Fridays in Room 438B in Building 2, Harvard Chan SPH** to help with homework questions or general questions.

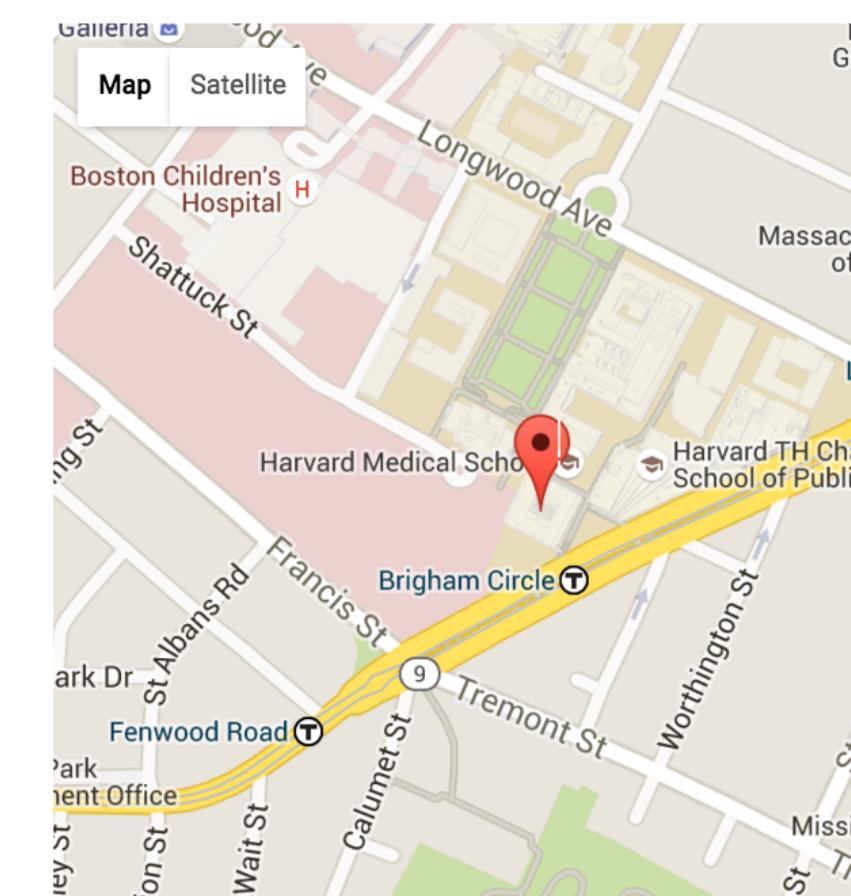
**Course Wiki Navigation**

The course wiki page navigation is easiest using the hierarchical tree on the upper left-hand side of all course pages. This welcome page is the parent page for all other course content. You can view the direct child pages by clicking on the ">" next to the page title in the hierarchical tree. The child pages with ">" icons next to the titles also have child pages, which you can view by clicking on the ">". Listed below are the descriptions of the main course pages:

- **Announcements:** critical announcements related to our course
- **Course Agenda:** course schedule, daily timelines
- **Course Preparation:** describes all installations necessary prior to start of course
- **Homework:** provides homework assignments and a space for each participant to upload homework assignments for each course session

**Course Venue**

**Countway Library of Medicine**  
10 Shattuck St., Boston, MA 02115



<http://tinyurl.com/hbc-course-main>

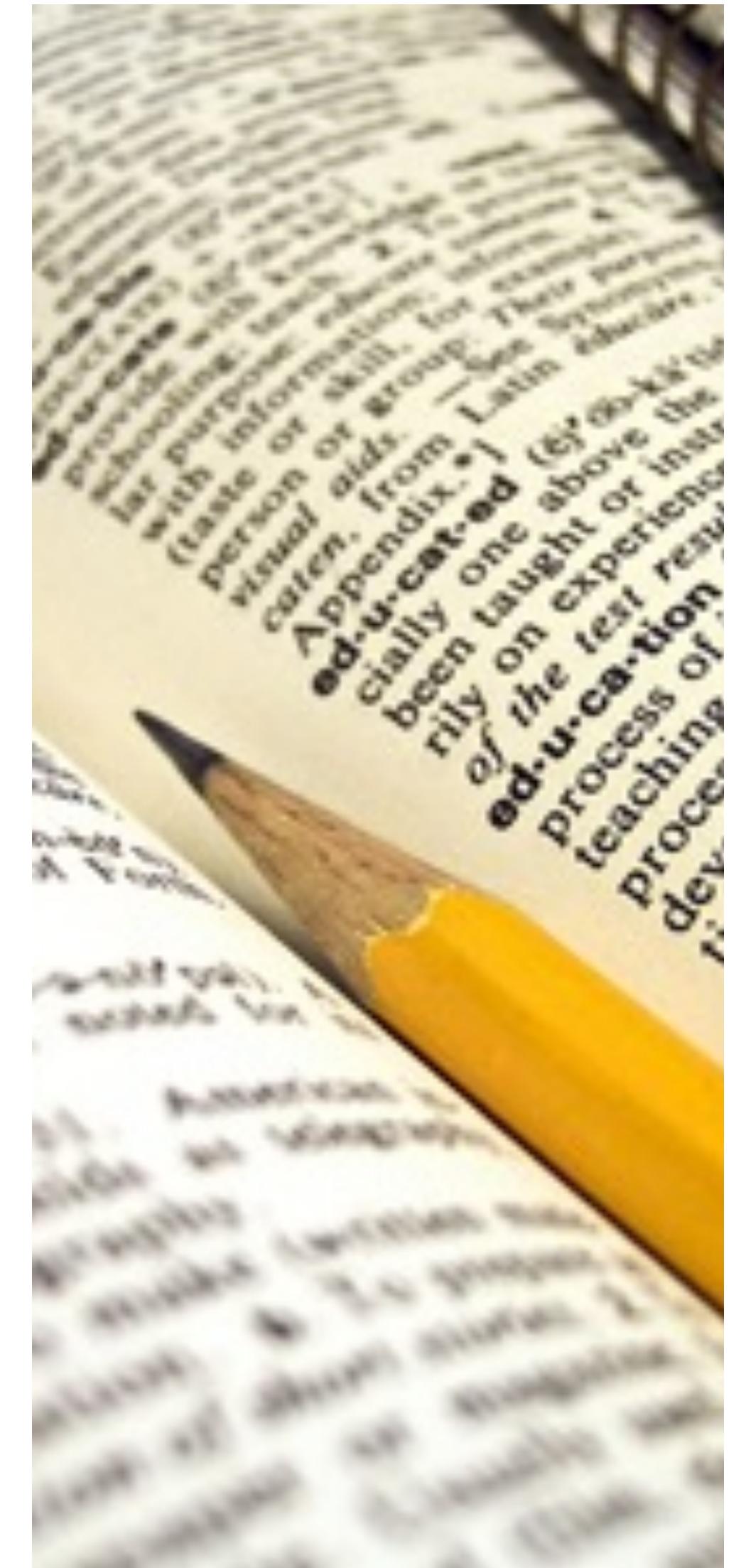
- ▼ **NGS Data Analysis Course, Spring 2016**
  - Announcements (2016)
  - Course Agenda (2016)
  - Course Preparation (2016)
  - ▼ Homework (2016)
    - Homework Submission (2016)
  - ▼ Session I: Introduction to Unix / Orchestra and NGS Data Analysis (2016)
    - Session I Homework (2016)

Course wiki page

# Weekly homework

- ⌄ [Homework \(2016\)](#)
  - [Homework Submission \(2016\)](#)
  - [Session I: Homework Assignment, 2016](#)

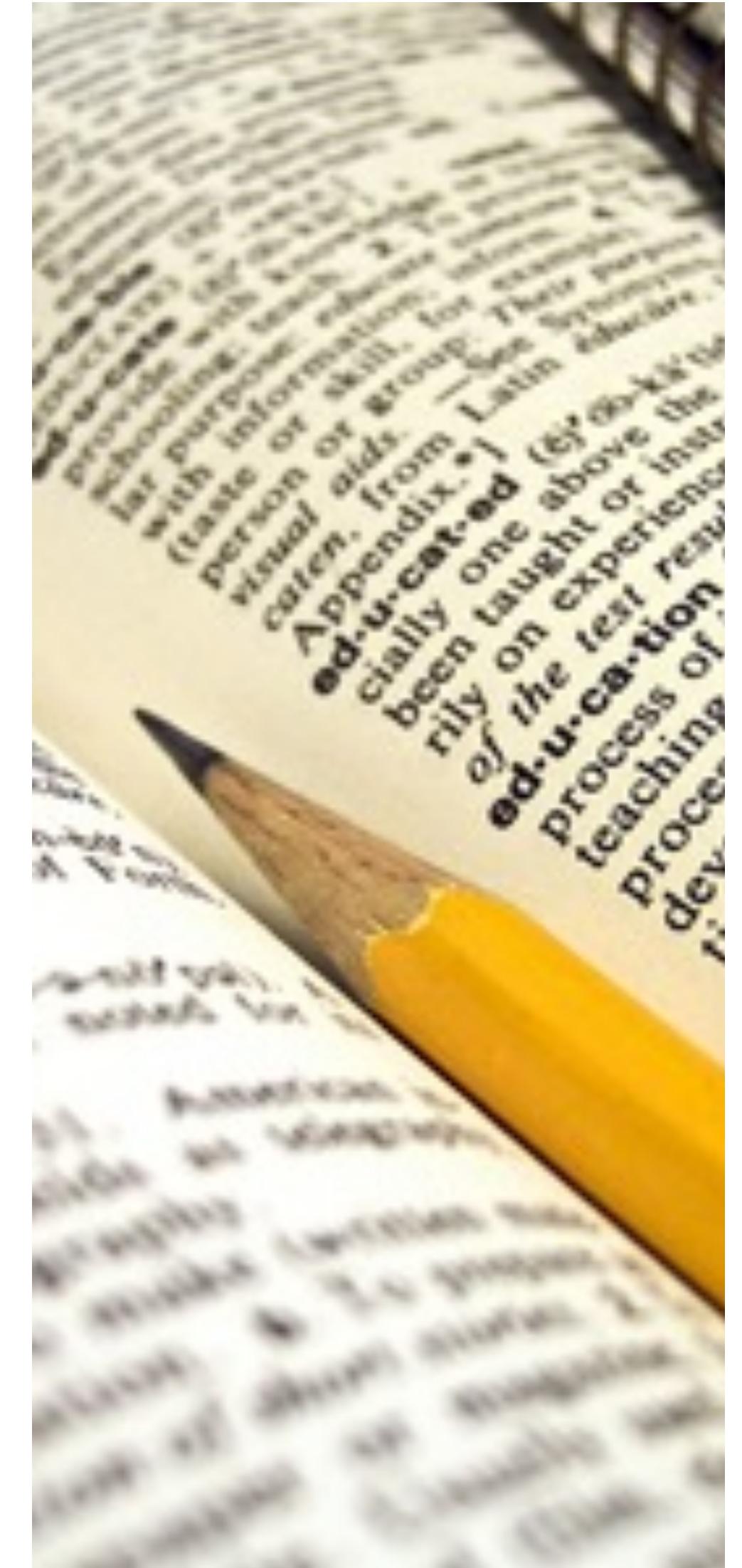
✓ Mandatory



# Weekly homework

- ⌄ [Homework \(2016\)](#)
  - [Homework Submission \(2016\)](#)
  - [Session I: Homework Assignment, 2016](#)

- ✓ Mandatory
- ✓ Help available during office hours\*\*, between sessions



# Weekly homework

- ⌄ [Homework \(2016\)](#)
  - [Homework Submission \(2016\)](#)
  - [Session I: Homework Assignment, 2016](#)

- ✓ Mandatory
- ✓ Help available during office hours\*\*, between sessions
- ✓ Comment and discuss on the homework page



# Weekly homework

- ⌄ [Homework \(2016\)](#)
  - [Homework Submission \(2016\)](#)
  - [Session I: Homework Assignment, 2016](#)

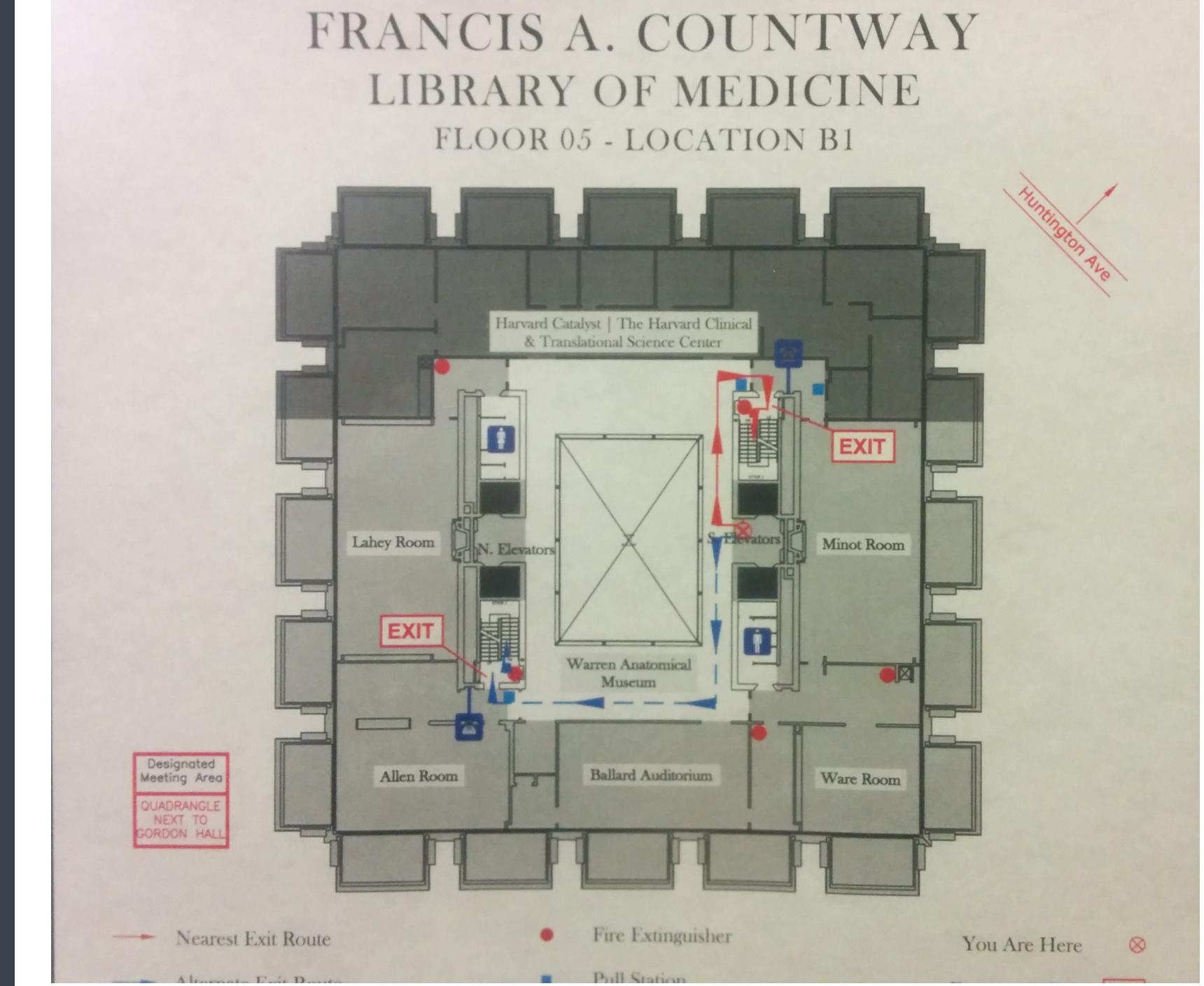
- ✓ Mandatory
- ✓ Help available during office hours\*\*, between sessions
- ✓ Comment and discuss on the homework page

*\*\* poll to schedule office hours*



# Countway Library

- ✓ Minot, Ware and Ballard rooms
- ✓ Bathrooms
- ✓ Water fountain



# Odds and Ends

- ✓ Name tags
- ✓ Post its
- ✓ Lunch locations



*These materials have been developed by members of the teaching team at the [Harvard Chan Bioinformatics Core \(HBC\)](#). These are open access materials distributed under the terms of the [Creative Commons Attribution license \(CC BY 4.0\)](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.*

