

Cold Spring Harbor Advanced Sequencing Technologies and Bioinformatics Analysis 2024

Full Course Days: Monday November 11 - Friday November 22

Application Deadline: August 15, 2024

Arrival: Sunday November 10th by 6pm EST

Departure: Saturday November 23th

Zoom link: <https://zoom.us/j/99539016173?pwd=UUgD5xhd4McRThX2zJe47sfcJbrMfl.1>

Meeting ID: 995 3901 6173

Passcode: 531181

Instructors: Felicia Gomez (fgomez@wustl.edu), Malachi Griffith (mgriffit@wustl.edu), Obi Griffith (obigriffith@wustl.edu), Elaine Mardis (elaine.mardis@nationwidechildrens.org), W. Richard McCombie (mccombie@cshl.edu), Chris Miller (c.a.miller@wustl.edu), Jonathan Preall (jpreall@cshl.edu)

TAs: Arpad Danos (adanos@wustl.edu), My Hoang (hmy@wustl.edu), Mariam Khanfar (k.mariam@wustl.edu), Jason Kunisaki (Jason.Kunisaki@hci.utah.edu), Claire Regan (cregan@cshl.edu), Kartik Singhal (k.singhal@wustl.edu), Daniel Puthawala (Daniel.Puthawala@nationwidechildrens.org), Jeremy Arbesfeld (Jeremy.Arbesfeld@nationwidechildrens.org), Corinne Strawser (Corinne.Haines@NationwideChildrens.org)

Guest lecturers (invited lectures will be live (in person or virtual), and recorded for later review):

Sequencing Technology Lectures 2024p

- Nov 11, 2:00pm: *Spatial Transcriptomics Alumni lecture* - Corinne Strawser (corinne.haines@nationwidechildrens.org)
- Nov 11, 3:00pm: *Ethics and Regulation of Genomic Research* - Jessica Mozersky (jmozersky@wustl.edu)
- Nov 13, 9:00am: *Genomics in agriculture* - Zach Lippman (lippman@cshl.edu)
- Nov 20, 2:00pm: *Rapid WGS pediatric diagnostics* - Bimal Chaudhari (bimal.chaudhari@nationwidechildrens.org)
- Nov 22, 9:00am: *What's new in the human (pan)genome?* - Adam Phillippy (adam.phillippy@nih.gov)

Analysis Lectures 2024

- Nov 13, 2:00pm: *Intro to Python* - Jason Williams (williams@cshl.edu)
- Nov 14, 9:30am: *Functional Genomics Alumni Lecture: Massively Parallel Assays* - Justin Kinney (jkinney@cshl.edu)
- Nov 14, 10:30am: *Intro R / additional lecture topic* - Katie Campbell (KatieCampbell@mednet.ucla.edu)
- Nov 20, 9:00am: *Variant Representation and Annotation* - Alex Wagner (alex.wagner@nationwidechildrens.org)
- Nov 21, 5:00pm: *Single cell ATAC-seq* - Yang Li (yeli@wustl.edu)

Relevant CSHL Staff:

David Stewart (stewart@cshl.edu) - Executive Director, Meetings & Courses Program.

Barbara Zane (zane@cshl.edu) - Scientific equipment/operations

Christopher Choute / Adam Crespo (choute@cshl.edu / crespo@cshl.edu) - IT support, computer sourcing

Ed Campodonico/Coleen Jackman (avstaff@cshl.edu / campodon@cshl.edu / jackman@cshl.edu) - AV needs

Olivia Mulligan (mulliga@cshl.edu) - Students/faculty affairs, housing, transportation, etc

Coleene Jackman (jackman@cshl.edu) - Housing, special housing requests

Melissa Davis/Ann May (medavis@cshl.edu / amay@cshl.edu) - Culinary Services

Wendy Crowley & team (crowley@cshl.edu) - Grants, budget etc

Teri Schindler (jurksch@cshl.edu) - Safety info/training

Connie Brukin (cbruakin@cshl.edu, constance.brukin@gmail.com) - Course Photo (Nov 11, 2:00pm; Nov 17, 10:30am Confirmed)

Mealtimes at Blackford Hall

Breakfast: 7:30am - 9:00am

Lunch: 11:30am - 1:30pm

Dinner: 5:30pm - 7:00pm

Prerequisites

<https://rnabio.org/module-00-setup/0000/05/01/Prerequisites/>

Students will be provided with this list in advance of the course.

Email list (instructors, TAs, guest lecturers and assistants):

fgomez@wustl.edu; mgriffit@wustl.edu; obigriffith@wustl.edu; elaine.mardis@nationwidechildrens.org;
mccombie@cshl.edu; c.a.miller@wustl.edu; jpreall@cshl.edu; adanos@wustl.edu; hmy@wustl.edu;
k.mariam@wustl.edu; Jason.Kunisaki@hci.utah.edu; cregan@cshl.edu; k.singhal@wustl.edu;
Daniel.Puthawala@nationwidechildrens.org; Jeremy.Arbesfeld@nationwidechildrens.org;
Corinne.Haines@NationwideChildrens.org; corinne.haines@nationwidechildrens.org; lippman@cshl.edu;
adam.phillippy@nih.gov; jmozersky@wustl.edu; bimal.chaudhari@nationwidechildrens.org; jkinney@cshl.edu;
alex.wagner@nationwidechildrens.org; yeli@wustl.edu; KatieCampbell@mednet.ucla.edu; williams@cshl.edu

Course Slack Channel: <http://seqtec-2024.slack.com>

Invite link (Note: Instructors/TAs should also ask to be added to the instructors channel):

https://join.slack.com/t/seqtec-2024/shared_invite/zt-2pyq71q9f-KK9spCK7ic77z5ilI9k0sA

Cloud Details:

Console login: <https://cshlworkshops.signin.aws.amazon.com/console>

Region: Virginia (us-east-1)

Instance type: m6a.xlarge (\$0.1728 per Hour)

Example instructor instance login: `ssh -i cshl_2024_instructor.pem ubuntu@[instructor-ip-address]`

Example student instance login: `ssh -i cshl_2024_student.pem ubuntu@[student-ip-address]`

Classroom computers

24 x 27" iMacs 16GB/3TB

1 x Printer

Software: IGV, Zoom, Slack, R/RStudio, Loupe Browser

CSHL Leading Strand link (registration and login required): TBA

Note: Video recordings will be available for 12 weeks after the course

Note: Each individual video will also be linked below

Online course materials (will be updated in real time as the course progresses):

[RNAbio.org](https://rnabio.org)

Lecture slides (will be updated in real time as the course progresses):

[Slides](#)

SCHEDULE:

10th November (Sunday) - Travel day (Elaine, Dick)

5:30pm - 7:00pm: **DINNER (Blackford Cafeteria)**

7:00pm: Meet and greet with instructors Dick McCombie and Elaine Mardis, in the Blackford Hall Bar for introductions and course overview

11th November (Monday) - Intro to genomics and technologies - lecture day (Elaine, Dick)

9:00am - 9:30am: Introduction to the course and curriculum. Review of activities, group assignments, schedule. Q&A.

9:30am - 10:45am: “**Overview of Next-generation Short Read Sequencing Technologies, Emerging platforms, and Experimental Design**” lecture (Elaine Mardis)

10:45am - 11:00am: **BREAK**

11:00am - 12:00pm: “**Long read sequencing technology and research applications**” lecture (Richard McCombie)

12:00pm - 1:00pm: **LUNCH**

1:00pm - 2:00pm: “**Single cell sequencing approaches**” (Jonathan Preall)

2:00pm - 2:15pm: **short break**

2:15pm - 3:15pm: “**Spatial Transcriptomics Alumni lecture**” (Corinne Strawser)

3:15pm - 4:15pm: “**Ethics and Regulation of Genomic Research**” (Jessica Mozersky-virtual lecture)

4:15pm - 4:30pm: **BREAK**

4:30pm - 6:00pm: Q&A, Time to work on your presentations

6:00pm - 7:00pm: **DINNER**

7:00pm - 9:00pm: 10 Student presentations part I (5-7 minute talk focusing upon your research and planned use of NGS)

9:00pm - 10:00pm: **SOCIAL HOUR (after student talks)**

12th November (Tuesday) - Intro to cloud, command line, file formats, and alignment concepts (Felicia)

9:00am - 9:15am: Outline of informatics portion of course/Intro to Instructors (Felicia Gomez)

9:15am - 10:30am: “**An overview of bioinformatics**” lecture (Felicia Gomez)

10:30am - 10:45am: **BREAK**

10:45am - 12:00pm: [Introduction to Unix/command line lab](#), using Mac terminal (Felicia Gomez)i

12:00pm - 1:00pm: **LUNCH**

1:00pm - 2:15pm: [Introduction to Cloud Computing lab](#) (My)

2:15pm - 3:30pm: [More command line lab using Mac terminal](#) (Jason)

3:15pm - 3:30pm: **BREAK**

3:30pm - 4:30pm: [File formats](#) (FASTQ, BAM, CRAM, etc) (Mariam Khanfar)

4:30pm - 5:30pm: [DNA Alignment concepts](#) (references, indexing, alignment) (Felicia Gomez)

5:30pm - 7:00pm: **DINNER**

7:00pm - 9:00pm: 10 Student presentations part II (5-7 minute talk focusing upon your research and planned use of NGS)

9:00pm - 10:00pm: **SOCIAL HOUR (after student talks)**

13th November (Wednesday) - Genome arithmetic, bedtools and python (Jason K, Jason W, Felicia, Obi)

9:00am-10:00am: “**Genomics in agriculture**” (Zach Lippman)

10:00am - 10:15am: **BREAK**

10:15am-11:00am: [Intro to genome coordinates, interval, concepts of genome arithmetic](#) (Mariam Khanfar)

11:00am-12:00pm: [Bedtools tutorial and puzzles](#) in groups (Jason K)

12:00pm - 1:00pm: **LUNCH**

1:00pm - 2:00pm: [Bedtools tutorial and puzzles](#) in groups cont'd (Jason K)

2:00pm - 3:30pm: Jason Williams - Python (confirmed date/time)

3:30pm - 3:45pm: **BREAK**

3:45pm - 5:00pm: Jason Williams - Python

5:00pm - 5:30pm: Prepare for social

5:30pm - 7:00pm: **Wine and cheese (Hershey)**

7:00pm - 9:00pm: **Pizza Dinner Buffet (Hershey)**

14th November (Thursday) - Introduction to R, IGV, and alignment lab (Katie, Felicia, Malachi, Obi)

9:30am - 10:30am: Functional Genomics Alumni Lecture: “**Massively Parallel Assays**” (Justin Kinney)

10:30am - 10:45am: **BREAK**

10:45am - 12:00pm: [Intro to R](#) (Katie Campbell)

12:00pm - 1:00pm: **LUNCH**

1:00pm - 2:45pm: Intro to R (Katie Campbell)

2:45pm - 3:00pm: **BREAK**

3:00pm - 4:30pm: [IGV lab](#) (Kartik Singhal)

4:30pm - 5:30pm: [DNA Alignment lab](#) (Mariam Khanfar)

5:30pm - 6:30pm: **DINNER**

6:30pm - 8:00pm: Team formation, mini-assignment - [DNA alignment and IGV](#) (My Hoang)

15th November (Friday) - Bulk RNA-seq day 1 (Malachi, Obi)

9:00am - 10:30am: “**Introduction to bulk RNA-sequencing**” lecture (Malachi Griffith)

10:30am - 10:45am: **BREAK**

10:45am - 12:00pm: [Environment](#) and [tool installation](#) lab (Obi Griffith)

12:00pm - 1:00pm: **LUNCH**

1:00pm - 2:15pm: [reference](#), [annotations](#), [indexing](#), [data](#) and [QC](#) lab (Obi Griffith)

2:15pm - 2:45pm: Alignment and visualization concepts lecture (Arpad Danos)

2:45pm - 3:00pm: **BREAK**

3:00pm - 4:00pm: [Alignment lab](#) (Arpad Danos)

4:00pm - 5:00pm: [Alignment visualization lab](#) (Malachi Griffith)

5:00pm - 6:30pm: [Post alignment QC lab](#) (Malachi Griffith)

6:30pm - 7:15pm: **COCKTAILS (Blackford Bar) - join the *Biological Data Sciences* meeting cocktails**

7:15pm - 8:00pm: **BANQUET DINNER (Blackford Hall) - join the *Biological Data Sciences* meeting banquet**

16th November (Saturday) - Bulk RNA-seq day 2 (Malachi, Obi)

9:00am - 10:00am: [Expression and Differential Expression analysis](#) lecture (Obi Griffith)

10:00am - 11:00am: [Expression analysis](#) lab (Obi Griffith)

11:00am - 11:15am: **BREAK**

11:15am - 11:45am: [Expression analysis lab](#) cont'd (Obi Griffith)

11:45am - 1:15pm: **LUNCH on the beach (weather permitting)**

1:15pm - 2:15pm: [Differential expression lab](#) (My Hoang)

2:15pm - 2:30pm: **BREAK**

2:30pm - 4:00pm: [Differential expression visualization and interpretation lab](#) (Malachi Griffith)

4:00pm - 5:30pm: [Differential expression pathway analysis lab](#) (Obi Griffith)

5:30pm - 6:30pm: **DINNER**

6:30pm - 8:00pm: [Team assignment - Alignment](#) (Arpad Danos)

17th November (Sunday) - bulk RNA-seq day 3 (Malachi, Obi)

9:00am - 10:30am: [Batch correction lab](#) (Malachi Griffith)

10:30am - 10:45am: **COURSE PHOTO PART II**

10:45am - 12:00pm: [Batch correction lab](#) cont'd (Malachi Griffith)

12:00pm - 1:00pm: **LUNCH**

1:00pm - 3:00pm: [Alignment free analysis lab](#) (Malachi Griffith)

3:00pm - 3:15pm: **BREAK**

3:15pm - 5:30pm: [Team assignment - Expression/DE](#) (Kartik Singhal / Mariam Khanfar)

5:30pm - 6:30pm: **DINNER**

EVENING OFF

18th November (Monday) - Epigenomics, long reads and data visualization (Chris, Malachi)

9:00am - 10:00am: "Introduction to epigenomics - ATAC, ChIPseq and Bisulfite approaches" lecture (Chris Miller)

10:00am - 10:15am: **BREAK**

10:15am - 11:15am: [Epigenomics lab - peak calling, etc](#) (Kartik)

11:15am - 12:00pm: Long read sequencing lecture (Chris)

12:00pm - 1:00pm: **LUNCH**

1:00pm - 2:00pm: [Long read sequencing \(RNAseq\) lab](#)

2:00pm - 2:15pm: **BREAK**

2:15pm - 3:15pm: [Data visualization with R](#) (Chris)

3:15pm - 4:00pm: Data Visualization concepts lecture (Chris)

4:00pm - 5:00pm: [Data Munging/Cleaning](#) (Chris)

5:00pm - 6:00pm: **DINNER**

6:00pm - 6:30pm: Bash tricks

6:30pm - 7:30pm: Dataviz practical exercises

19th November (Tuesday) - Variants and more python (Chris, Alex)

9:00am - 9:45am: Intro to variant calling (Chris)

9:45am - 10:30am: Germline variant calling lab

10:30am - 10:45pm: **BREAK**

10:45am - 12:00pm: Somatic variant calling lecture/lab

12:00pm - 1:00pm: **LUNCH**

1:00pm - 2:00pm Troubleshooting bioinformatics - things that can go wrong, will go wrong (Chris)

2:00pm - 2:30pm Copy Number and Structural variants lecture (Chris)

2:30pm - 2:45pm **BREAK**

2:45pm - 3:45pm [CNV/SV lab](#)

3:45pm - 5:00pm: **Group walk around CSHL campus (optional)**

5:00pm - 6:00pm: **DINNER**

6:00pm - 6:30pm Bioinformatics with Python (Alex)

6:30pm - 8:00pm Bioinformatics with Python lab

20th November (Wednesday) - Variant annotation and python visualization (Alex, Chris)

9:00am - 10:00am: "Variant Representation and Annotation" Lecture (Alex Wagner)

10:00am - 10:15am: **BREAK**

10:00am - 12:00pm: Variant Representation and Annotation lab (Alex Wagner)

12:00pm - 1:00pm: **LUNCH**

1:00pm - 1:30pm Variant Interpretation lecture (Alex Wagner)

1:30pm - 2:00pm Variant Interpretation lab (Alex)

2:00pm - 3:00pm: "Rapid WGS pediatric diagnostics" (Bimal Chaudhari)

3:00pm - 3:15pm: **BREAK**

3:15pm - 5:00pm: Python Visualization (seaborn, matplotlib) (Claire Regan)

5:00pm - 6:00pm: **FREE TIME TO RELAX AND PREPARE FOR BANQUET**

6:00pm - 6:45pm: **COURSE COCKTAILS (Blackford Bar)**

6:45pm - 8:00pm: **COURSE BANQUET DINNER (Blackford Hall)**

21st November (Thursday) - Single-cell RNAseq (Jonathan)

9:00am - 10:30am: "Introduction to scRNA analysis" lecture (Jonathan Preall)

10:30am - 10:45am: **BREAK**

10:45am - 12:00pm: scRNA - raw data format, data QC, Cell Ranger, and basic visualization in Loupe/VDJ browser lab (Jonathan Preall)

12:00pm - 1:00pm: **LUNCH.**

1:00pm - 3:00 pm: scRNA/ScanPy lab (Jonathan Preall)

3:00pm - 3:30pm: **BREAK**

3:30pm - 4:30pm: scRNA/ScanPy lab continued (Jonathan Preall)

5:00pm - 6:00pm: "Single cell ATAC-seq" lecture (Yang Li)

6:00pm - 7:30pm: **DINNER**

22nd November (Friday) - Single-cell RNAseq (Jonathan)

9:00am - 10:00am “What’s new in the human (pan)genome?” (Adam Phillippy - virtual)

10:00am-10:15am: **Meet at Grace at 10:00 for van pick-up**

10:15am - 11:30am: WOODBURY TOUR

11:30am - 1:00pm: **LUNCH**

1:00pm - 3:00pm: Single Cell lab: data integration methods (Jonathan Preall)

3:00pm - 3:15pm: **BREAK**

3:15pm - 4:30pm: Single Cell Exercise TBD (Jonathan Preall)

4:30pm - 6:00pm: Open Q&A, course evaluations

6:00pm - 7:00pm: **DINNER**

23rd November (Saturday) - Travel day (Jonathan, Dick)

Students depart - No activities scheduled

Course Instructors

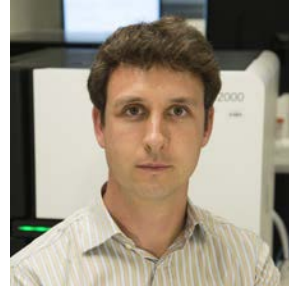
[Felicia Gomez](#), WASHU



[Malachi Griffith](#), WASHU



[Obi Griffith](#), WASHU



[Elaine Mardis](#), NCH



[W. Richard McCombie](#), CSHL



[Chris Miller](#), WASHU



[Jonathan Preall](#), CSHL

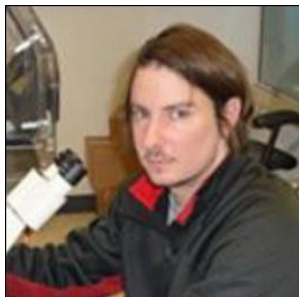


Teaching Assistants

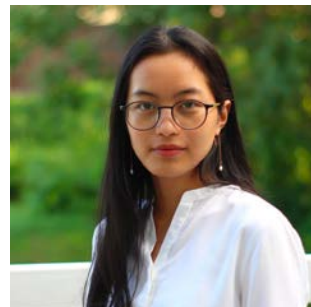
[Jeremy Arbesfeld](#), NCH



[Arpad Danos](#), WashU



[My Hoang](#), WashU



[Mariam Khanfar](#), WashU



[Jason Kunisaki](#), UU



[Daniel Puthawala](#), NCH



[Claire Regan](#), CSHL



[Kartik Singhal](#), WashU



Guest Lecturers

[Katie Campbell](#), UCLA



[Bimal Chaudhari](#), NCH



[Justin Kinney](#), CSHL



[Zach Lippman](#), CSHL



[Yang Li](#), WashU



[Jessica Mozersky](#), WashU



[Adam Phillippy](#), NHGRI



[Corinne Strawser](#), NCH



[Alex Wagner](#), NCH



[Jason Williams](#), CSHL

