BST282, Introduction to Bioinformatics and Computational Genomics
Spring 2023 – Tue & Thu 11:30-1 Kresge 502

Instructor Information
Faculty
Dr Martin Hemberg, Evergrande Center for Immunologic Diseases, Brigham and Women's Hospital and Harvard Medical School
Hale Building for Transformative Medicine 9016Z, mhemberg@bwh.harvard.edu
Office Hours: 4-5 PM Wednesday on zoom, https://broadinstitute.zoom.us/j/83379057723?pwd=ZDRtNjVWZ29LMHBJZTQ1M0NNblFLZz09. In person option available through prior appointment only.

Dr Luca Pinello, Molecular Pathology, Massachusetts General Hospital/Harvard Medical School
149 13th Street, 02129 Charlestown, 6.014, lpinello@mgh.harvard.edu
Office Hours: 4-5 PM Wednesday on zoom, https://broadinstitute.zoom.us/j/83379057723?pwd=ZDRtNjVWZ29LMHBJZTQ1M0NNblFLZz09. In person option available through prior appointment only.

Teaching Assistants
Yujie Guo, Research Assistant at Dana Farber Cancer Institute
Center for Life Science Boston 11007, yujie_guo@hms.harvard.edu
Office hours: TBD (will update)

Marie Zhang, Graduate Researcher at Harvard Medical School
Department of Biomedical Informatics, marie_zhang@hms.harvard.edu
Office hours: TBD (will update)

Credits
5 HSPH credits

Course Description
The course will cover basic technology platforms, data analysis problems and algorithms in computational biology. Topics include sequence alignment, high-throughput experiments for bulk and single-cell gene expression and chromatin accessibility, transcription factor binding and epigenetic profiling, motif finding, genome-wide association studies, spatial transcriptomics, multi-omics, CRISPR design, quantification and screens. Computational algorithms covered include hidden Markov models, dimensionality reduction, clustering and classification methods.

Pre-Requisites
- Molecular biology

- Computer science
- Statistics

Learning Objectives

Upon successful completion of this course, you should be able to:

- Be familiar with the most commonly used technologies for high-throughput profiling of the genome and transcriptome.
- Understand the computational and statistical challenges related to high-throughput sequencing data.
- Use existing computational tools for quality control and analysis of high-throughput sequencing data.
- Understand the algorithms underlying some of the most common computational tools.

Course Readings: There is no required textbook for this course. Required articles and additional reference material will be provided via the course website.

Course Structure

The course consists of three core components: lectures, labs, and homework problems/essays. Lectures are held in person on Tuesdays and Thursdays. Students are expected to attend lectures in person. Lectures will be recorded and available upon request. Lab times will be decided during the first week to ensure a time that works for as many students as possible, and they will be led by one of the Teaching Assistants. All labs are computer-based. Some of the guest lectures will be given over Zoom.

The course consists of four modules and at the end of each module students are required to submit an essay (see below). During the final lecture of each module a group of students will be tasked with presenting each of the three papers and leading a discussion.

Canvas Course Website: The Canvas site is an important learning tool for this course where students will access required articles, submit course assignments, and share other resources with the class. Course announcements will be posted on the site and students will be required to check the course website on a weekly basis.

Technical Information: You will get access to the Cannon cluster which will provide all hardware and software resources required for the labs and homework assignments.

Grading, Progress and Assessment

Each homework assignment will consist of problems to be solved by hand or via web applications, published algorithms, and coding (in Python or R). We will teach some programming in R and Python in the labs, but it is strongly recommended for students to have some programming background before taking this course. Students are welcome to discuss course material with each other and with the TAs, but the submitted work must be your own. Use of AI-based text generators is not allowed (e.g. ChatGPT) and will be detected.

As a general rule, students are expected to attend lectures in person, although we recognize that there may be circumstances in which remote participation would be admissible. Please notify one of the instructors of an absence before the class. Engagement and participation in discussions during class may be reflected in the final grade.

The final grade for this course will be based on:

- Six homework problem sets, each one can be worth up to ten points (60%)
- Four essays and associated presentations, each one can be worth up to nine points (36%)
- Four course evaluations, each one is worth one point. (4%)
- Up to ten bonus points can be awarded to those who create a new homework problem set.

Homework problem sets (10 points each)
Homework problem sets consist of discrete questions. They will typically require you to run existing software or write some code of your own. Answers should be submitted electronically, ideally as a markdown file or scientific notebook. All HW should be submitted through the Harvard Canvas web interface. A total of 10 late days is granted, which you can use towards any HW, although any single HW cannot be late for more than 3 days. After the 10 days are used up, 10% will be deducted from each additional late day on a HW, respectively, unless prior approval is obtained from one of the course instructors. Homework is due at midnight on specific Sundays at the end of the week.

Essays (9 points each)
At the end of each module you will be given a selection of three papers and a set of accompanying questions. Your task is to write a one-page essay about the paper where you discuss the three questions. The essay is due at midnight on the Sunday prior to the last lecture of each module. The same pool of late days that is granted for the HW sets can be used for the essays. Groups of 3-4 students will be giving a 10-15 minute presentation of each paper during the module review lecture. The presentation and the ensuing discussion will be counted for up to 3 points for this component of the grade.

Course evaluations (1 point each)
Questionnaires will be handed out at the end of each module. Successful and timely completion will reward you one point.

Bonus: create a new problem set (10 points)
Students can earn up to ten additional points by creating a new problem set based on the current course content that could be used for next year's course. Make sure to provide the full dataset, explain sufficient background, list the questions, and include the detailed solutions (code, answers) clearly. This problem set needs to be handed in prior to the last lecture, but it is **not** mandatory. You are allowed to

work in small groups for this assignment and submit jointly, but the number of points awarded will be divided by the number of members in the group.

Harvard Chan Policies and Expectations

Inclusivity Statement

Diversity and inclusiveness are fundamental to public health education and practice. Students are encouraged to have an open mind and respect differences of all kinds. We share responsibility with you for creating a learning climate that is hospitable to all perspectives and cultures; please contact us if you have any concerns or suggestions.

Bias Related Incident Reporting

The Harvard Chan School believes all members of our community should be able to study and work in an environment where they feel safe and respected. As a mechanism to promote an inclusive community, we have created an anonymous bias-related incident reporting system. If you have experienced bias, please submit a report here so that the administration can track and address concerns as they arise and to better support members of the Harvard Chan community.

Title IX
The following policy applies to all Harvard University students, faculty, staff, appointees, or third parties: Harvard University Sexual and Gender-Based Harassment Policy. Procedures For Complaints Against a Faculty Member Procedures For Complaints Against Non-Faculty Academic Appointees

Academic Integrity
Each student in this course is expected to abide by Harvard University and the Harvard T.H. Chan School of Public Health School's standards of Academic Integrity. All work submitted to meet course requirements is expected to be a student's own work. In the preparation of work submitted to meet course requirements, students should always take great care to distinguish their own ideas and knowledge from information derived from sources.

Collaboration in the completion of assignments is allowed. Students are expected to each write their own essays and code, but discussion and group work is allowed.

Should academic misconduct occur, the student(s) may be subject to disciplinary action as outlined in the Student Handbook. See the Student Handbook for additional policies related to academic integrity and disciplinary actions.

Accommodations for Students with Disabilities
Harvard University provides academic accommodations to students with disabilities. Any requests for academic accommodation should ideally be made before the first week of the semester, except for unusual circumstances, so arrangements can be made. Students must register with the Local Disability Coordinator in the Office for Student Affairs to verify their eligibility for appropriate accommodations. Contact Colleen Cronin ccronin@hsph.harvard.edu in all cases, including temporary disabilities.

Religious Holidays, Absence Due to
According to Chapter 151c, Section 2B, of the General Laws of Massachusetts, any student in an educational or vocational training institution, other than a religious or denominational training institution, who is unable, because of his or her religious beliefs, to attend classes or to participate in any examination, study, or work requirement on a particular day shall be excused from any such examination or requirement which he or she may have missed because of such absence on any particular day, provided that such makeup examination or work shall not create an unreasonable burden upon the School. See the student handbook for more information.

Course Evaluations
Constructive feedback from students is a valuable resource for improving teaching. The feedback should be specific, focused, and respectful. It should also address aspects of the course and teaching that are positive as well as those which need improvement.

Completion of the evaluation is a requirement for each course. Your grade will not be available until you submit the evaluation. In addition, registration for future terms will be blocked until you have completed evaluations for courses in prior terms.

Course Schedule & Assessment of Student Learning

Sample A: Snapshot of a COURSE SCHEDULE

Week of: 4/17-5/8				
Session topics	**Objectives**		**Readings**	**Assignments**
[illegible]	[illegible]			**Homework:** Homework problem set 6 due Sun 4/30 Essay 4 due Sun 5/7

Please note, session topics and activities may be subject to change during the course