JOHNS HOPKINS
WHITING SCHOOL
of ENGINEERING

Computer Science EN.601.449/EN.601.649
Computational Genomics: Applied Comparative Genomics
Fall, 2025 (3 credits, Engineering/FA2)

Instructor
Professor Michael Schatz, mschatz@cs.jhu.edu. http://schatz-lab.org
Office hours: TBD; and by appointment
Office hours location: Malone 323

Teaching Assistant
Mahler Revsine, mrevsin1@jh.edu
Office hours: TBD
Office hours location: Malone Hall

Meetings
Monday, Wednesday, 3:00pm – 4:15pm, Hodson 316
Recorded lectures can be accessed on the course Canvas site

Textbook
No textbooks are required

Online Resources
The following online resources are essential:

- The course web site contains a schedule of topics, class notes, and assignment details: https://github.com/schatzlab/appliedgenomics2025
- The course Piazza site at https://piazza.com/class/meogfdbmu7x7hf
- Piazza will serve as our discussion site for the course. Please use Piazza to ask questions of the instructor, TA and fellow students.
- The graded materials will be collected through Gradescope: https://www.gradescope.com/courses/1097756. Entry Code: GVXGV2
-

Course Information

- The goal of this course is to study the leading computational and quantitative approaches for comparing and analyzing genomes starting from raw sequencing data. The course will focus on human genomics and human medical applications, but the techniques will be broadly applicable across the tree of life. The topics will include genome assembly & comparative genomics, variant identification & analysis, gene expression & regulation, personal genome analysis, and cancer

genomics. The grading will be based on assignments, a midterm exam, class presentations, and a significant class project.

- **Prerequisites:** knowledge of the Unix operating system and programming expertise in a language such as R or Python. [Applications]

Course Goals

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

1. Understand the theoretical foundations for several of the most important genomic analysis tools
2. Have hands-on experience running several of the most important genomic tools
3. Perform novel research and analysis in computational biology

This course will address the following CSAB ABET Criterion 3 Student Outcomes

Graduates of the program will have an ability to:

SO1. Analyze a complex computing problem and apply principles of computing and other relevant disciplines to identify solutions.

SO2. Design, implement, and evaluate a computing-based solution to meet a given set of computing requirements in the context of the program's discipline.

SO3. Communicate effectively in a variety of professional contexts.

SO4. Recognize professional responsibilities and make informed judgments in computing practice based on legal and ethical principles.

SO6. Apply computer science theory and software development fundamentals to produce computing-based solutions.

Course Topics

We will study the leading computational and quantitative approaches for comparing and analyzing genomes starting from raw sequencing data. The course will focus on human genomics and human medical applications, but the techniques will be broadly applicable across the tree of life. The topics will include genome assembly & comparative genomics, variant identification & analysis, gene expression & regulation, personal genome analysis, and cancer genomics. Please see the main course website for a more detailed schedule, which will be updated as the semester progresses

Course Expectations & Grading

Course grades will be based on assignments (typically running and analyzing existing tools and/or analyzing a dataset), an examination, and a class project, according to the proportions below. Each homework assignment will be assigned a point value; the overall homework assignment grade will be computed as your total points earned divided by the total achieved in the class.

- 25% - Assignments (5% for each of 5 assignments)

- 40% - Exams (2 in-class exams; See class schedule for dates)
- 35% - Class Project
    - 2% - Project proposal
    - 3% - Preliminary report
    - 10% - In class presentation
    - 10% - Final written report
    - 10% - Oral Defense of written report

All grades will be distributed via email and recorded on GradeScope. Please keep your own record of your grades so that you will know your standing in the course. Letter grades for the course will be assigned on a standard scale, subject to the instructor's evaluation of your overall class performance. Students are allowed a total of 96 hours to extend the deadline for assignments, but not the class project, without any penalty. No further extensions will be allowed without a doctor's note or a note from the university.

Assignment Logistics. The assignments and projects in this course will require you to execute command line programs, write code in the language of your choice, or carry out a calculation. You must write all code independently unless the assignment specifically states that you can work in groups. Assignments will be submitted via Gradescope.

Attendance. All students are expected to attend all meetings of this course, and actively participate in all course meetings. If you miss a class meeting for any reason, you are responsible for material presented, and it is your responsibility to obtain any missed handouts or other materials. If you will be missing more than 1 class, please contact the instructor to discuss how to best review the missed materials.

Key Dates
See the schedule on the class website

Assignments & Readings
See the schedule on the class website

Ethics
The strength of the university depends on academic and personal integrity. In this course, you must be honest and truthful, abiding by the *Computer Science Academic Integrity Policy*:

> Cheating is wrong. Cheating hurts our community by undermining academic integrity, creating mistrust, and fostering unfair competition. The university will punish cheaters with failure on an assignment, failure in a course, permanent transcript notation, suspension, and/or expulsion. Offenses may be reported to medical, law or other professional or graduate schools when a cheater applies.
>
> Violations can include cheating on exams, plagiarism, reuse of assignments without permission, improper use of the Internet and electronic devices, unauthorized collaboration, alteration of graded assignments, forgery and falsification, lying, facilitating academic dishonesty, and unfair competition. Ignorance of these rules is not an excuse.

Academic honesty is required in all work you submit to be graded. Except where the instructor specifies group work, you must solve all homework and programming assignments without the help of others. For example, you must not look at anyone else's solutions (including program code) to your homework problems. However, you may discuss assignment specifications (not solutions) with others to be sure you understand what is required by the assignment.

If your instructor permits using fragments of source code from outside sources, such as your textbook or on-line resources, you must properly cite the source. Not citing it constitutes plagiarism. Similarly, your group projects must list everyone who participated.

Falsifying program output or results is prohibited.

Your instructor is free to override parts of this policy for particular assignments. To protect yourself: (1) Ask the instructor if you are not sure what is permissible. (2) Seek help from the instructor, TA or CAs, as you are always encouraged to do, rather than from other students. (3) Cite any questionable sources of help you may have received.

On every exam, you will sign the following pledge: "I agree to complete this exam without unauthorized assistance from any person, materials or device. [Signed and dated]". Your course instructors will let you know where to find copies of old exams, if they are available.

In addition, the specific ethics guidelines for this course are:

1. *AI systems may be used for the homeworks and for the project, although you must indicate when you used AI*
2. *For exam days, a single sheet of paper (8.5"x11") may be used for notes (either hand written or typed). No other written or electronic resources may be used during the examinations.*

Report any violations you witness to the instructor.

You can find more information about university misconduct policies on the web at these sites:

- For undergraduates: https://studentaffairs.jhu.edu/policies-guidelines/undergrad-ethics/
- For graduate students: http://e-catalog.jhu.edu/grad-students/graduate-specific-policies/