

Bioinformatics

11:126:485/16:765:585

Fall 2023

Location:

Lecture: Foran Hall 124 (computer lab)

Lab: Foran Hall 124 (computer lab)

Journal club (grad section only): TBD

Instructors:

Alanna Cohen

Office: Foran Hall 339B

abc162@rutgers.edu

Anat Kreimer

Office: CABM 239

kreimer@cabm.rutgers.edu

Office hours: By appointment

Required text: No required text for lab or lecture

Suggested text: Applied Bioinformatics of Nucleic Acid Sequences by David A. Hendix

An Introduction to Applied Bioinformatics (IAB)

A Primer for computational Biology by Shawn T. O'Neil

(All of the above are best for those new to programming/coding and include additional help and exercises. They are free, open access resources)

Introduction:

Bioinformatics aims to build computational models of biological systems. This is accomplished primarily by the creating of algorithms, databases, and web applications to solve questions in biology.

Bioinformatics implements computational advances such as dynamic computing, artificial intelligence, and cloud computing. These methods are required to deal with the huge amounts of data being produced by high-throughput sequencing and experimental techniques.

We live at a time where sequencing costs are lower than ever. As a result, there are massive amounts of DNA, RNA, and protein sequences that we are not sure how to make sense of. Bioinformatics attempts to sift through the data, identifying patterns, and addressing biological questions and providing statistical significance.

What this course is:

This course is primarily for biologists, to introduce them to bioinformatics concepts and techniques. We will expand on previous courses that utilized web-based databases and tools and introduce command-

line programming. This will allow you to perform more complex and computationally expensive analyses with larger datasets. We will also explore how to write our own code to extract specific information or results. I strongly believe that there are good programs available, it is not usually necessary to write your own code from scratch. However, you must understand when and how to use certain tools and how to optimize parameters to best fit the needs of your project or biological question.

All the work will be done in a live development environment, using Amarel, the high-performance computing cluster used by many labs at Rutgers. Working in this environment requires a basic knowledge of how to navigate a command-line environment and some basic programming/scripting. These concepts should have been covered in prerequisite courses, but will be reviewed here as well. By the end of the course, students will be comfortable working with UNIX/LINUX systems, Python script development, and automating existing applications for large datasets.

Please sign up for an Amarel account prior to our first lab meeting. You may request access to the cluster at <https://oarc.rutgers.edu/amarel-cluster-access-request/>

Course objectives:

1. Introduce bioinformatic algorithms/concepts and their implementations
2. Introduce the basics of working in a Linux environment, SLURM environment job submissions for parallel computing, and python scripting
3. Teach how to approach a biological problem as one that can be answered/addressed with bioinformatics. Students will learn the skills necessary to select appropriate tools, install/use them, optimize them based on their particular data, and build pipelines from a combination of existing tools and *de novo* code
4. Prepare students for more advanced bioinformatic courses
5. Provide students with a bioinformatic skill set that includes knowledge of basic script development and the vocabulary to collaborate within a multi-disciplined research team

Tentative Course schedule (subject to change):

Date	Lecture	Lab
		Intro to linux/command line programming
5-Sep	Intro to Bioinformatics	7-Sep
12-Sep	Gene Finding	14-Sep
19-Sep	Pairwise sequence alignment, BLOSUM	21-Sep
26-Sep	BLAST, affine gap costs	28-Sep
3-Oct	MSAs and domain families	5-Oct
10-Oct	Sequence signatures and motifs	12-Oct
17-Oct	Midterm Exam	19-Oct
24-Oct	Phylogenetics I- Parsimony and NJ	26-Oct
31-Oct	Phylogenetics II- ML and Bayesian inference	2-Nov
7-Nov	Metagenomics **Dr. Ray Sullivan	9-Nov
14-Nov	Gene Expression analysis	16-Nov

21-Nov	Structural Bioinformatics Variation and molecular level natural	23-Nov	<i>Thanksgiving Break-no class</i>
28-Nov	selection	30-Nov	Genomes, annotations, Ka/Ks
5-Dec	Disease gene prioritization	7-Dec	Programming in R, GWAS
12-Dec	Precision medicine/review		<i>End of classes-no lab</i>
TBD	Final Exam/Final projects due		

Availability of lecture materials:

I will primarily use PowerPoint slides for lecture materials. Slides will be posted before class to the course Canvas site. Please remember that all lecture materials are my protected intellectual property and cannot be shared or posted without my permission.

Grading Policy:

Each category of coursework will be weighted as follows:

11:126:485 (Undergraduate)

Lab Homework/Quizzes	30%
Midterm	30%
Final	40%

16:765:585 (Graduate)

Lab Homework/Quizzes	25%
Midterm	25%
Final	35%
Journal Club	15%

Lecture:

Each week will consist of a lecture (Tuesdays) and laboratory (Thursdays) session. There is no required textbook for this course. However, you will be responsible for any information presented or discussed during lecture (regardless of if it is the slides). Slide presentations will be posted to Canvas prior to each lecture. Lectures will not be recorded.

I will not be taking attendance, and class participation is not included in your grade. But attendance is strongly encouraged as all test material will be taken directly from our lectures.

Lab:

Lab attendance is required. Missing a lab without a valid excuse will result in a zero grade for any quizzes or homework assignments associated with that laboratory session. Missing more than two labs without a valid excuse with proper documentation will result in an automatic failing grade for the course. Important lab information and coding quizzes will typically be given at the beginning of the lab session. Please be on time as lateness will not warrant a make-up or extra time given.

Homework/Quizzes

Homework assignments will be assigned during your lab period and due the following week, at the start of the next laboratory period, unless otherwise specified. Late submissions will not be accepted.

Assignments with coding components must run properly in a standard development environment. I will not debug your code, but partial credit will be given where appropriate. Lack of submission, empty submissions, or “fake” submissions will receive a score of 0%. Your code must be yours and properly commented.

Any quizzes will be given during the laboratory session, and may be written questions, coding assignments, or both. Quizzes are up to my discretion and may be announced, but do not have to be. The material may come from either lecture or lab, but will be appropriate to the current topic covered. There are no make-up quizzes or extra time given for late starts.

Midterm/Final

The midterm exam will consist of 1) a written portion taken during lecture time and 2) a take-home programming assignment. The midterm will be based on material covered in lecture AND lab. The take-home component will be assigned during lab and be due the following week

The final exam will consist of 1) a question-based portion taken during our assigned exam time during finals week that will include some combination of different tools and techniques learned throughout the semester and 2) a writing component that will include designing, running, and writing up a computational analysis of some biological data. Your final writing project should be modeled after primary literature articles including an introduction/background, results, materials and methods, and discussion. You will have at least three weeks to complete the take-home portion.

Journal Club (Graduate Component Only)

Graduate students will be required to attend journal club meetings. Each meeting, everyone is expected to have read the selected paper and participate in the discussion. Every student will be required to choose and present a bioinformatics paper of their choice. You will be expected to present a short presentation and lead the discussion surrounding the paper. Your presentation and weekly participation in journal club will account for 15% of your final grade.

Undergraduates are welcome and encouraged to attend journal club but will not receive extra credit for doing so.

Attendance

Attendance is not mandatory, and there will not be grades assigned to participation or attendance. However, what you get out of this class depends on what you put into it. The more you show up and ask questions, the more the entire class will benefit from a collaborative learning environment.

Lecture slides will be posted to Canvas by the start of each class. Lectures will not be recorded. Students are responsible for taking their own notes.

Absence Policy

Attendance in lecture not mandatory. However, in the event of illness or emergency, I will arrange to make up missed work/notes. Attendance in lab is required. If you expect to miss a lecture or lab, please let me know as soon as possible. In addition, you must contact your Dean of Students at deanofstudents@echo.rutgers.edu in order to verify your absence.

Academic Integrity

Academic dishonesty, in any form, will not be tolerated. This includes copying from another student or online resources. Students are allowed and encouraged to discuss homework problems. But should not copy any code directly. Use the “coffee break” rule. If you have a discussion about a problem, you should be able to walk away, have a cup of coffee, and then return and be able to come up with a solution from memory. The same applies to searching the web for solutions to problems. You can use online resources to learn from, but need to be able to write your own code from what you learn.

Working together is only allowed on homework assignments. There will be no group work accepted for midterm and final projects.

Exams are strictly closed book/notebook and students are not allowed to use outside sources or help on exams. Students need to comply with the Rutgers Policy on academic integrity found at <http://academicintegrity.rutgers.edu/academic-integrity-policy>

The principles of academic integrity require that a student:

- properly acknowledge and cite all use of the ideas, results, or words of others
- properly acknowledge all contributors to a given piece of work
- make sure that all work submitted as their own is produced without the aid of impermissible materials or impermissible collaboration.
- obtain all data or results by ethical means and report them accurately without suppressing any results inconsistent with their interpretation or conclusions.
- treat all other students in an ethical manner, respecting their integrity and right to pursue their educational goals without interference. This requires that a student neither facilitate academic dishonesty by others nor obstruct their academic progress.

Please note that all course material (power point slides, exams, etc.) are my intellectual property and cannot be shared to outside parties or sites.

Student Wellness Services:

Access helpful mental health information and resources for yourself or a friend in a mental health crisis on your smartphone or tablet and easily contact CAPS or RUPD.

Counseling, ADAP & Psychiatric Services (CAPS)

(848) 932-7884

17 Senior Street, New Brunswick, NJ 08901

www.rhscaps.rutgers.edu/

CAPS is a University mental health support service that includes counseling, alcohol and other drug assistance, and psychiatric services staffed by a team of professional within Rutgers Health services to support students' efforts to succeed at Rutgers University. CAPS offers a variety of services that include: individual therapy, group therapy and workshops, crisis intervention, referral to specialists in the community and consultation and collaboration with campus partners.

Violence Prevention & Victim Assistance (VPVA)

(848) 932-1181

3 Bartlett Street, New Brunswick, NJ 08901

www.vpva.rutgers.edu

The Office for Violence Prevention and Victim Assistance provides confidential crisis intervention, counseling and advocacy for victims of sexual and relationship violence and stalking to students, staff and faculty. To reach staff during office hours when the university is open or to reach an advocate after hours, call 848-932-1181.

Disability Services

(848) 445-6800

Lucy Stone Hall, Suite A145, Livingston Campus, 54 Joyce Kilmer Avenue, Piscataway, NJ 08854

<https://ods.rutgers.edu/>

Rutgers University welcomes students with disabilities into all of the University's educational programs. In order to receive consideration for reasonable accommodations, a student with a disability must contact the appropriate disability services office at the campus where you are officially enrolled, participate in an intake interview, and provide documentation. If the documentation supports your request for reasonable accommodations, your campus's disability services office will provide you with a Letter of Accommodations. Please share this letter with your instructors and discuss the accommodations with them as early in your courses as possible. To begin this process, please complete the Registration form on the ODS web site at <https://ods.rutgers.edu/students/registration-form> and visit <https://ods.rutgers.edu/students/documentation-guidelines> for more information

Scarlet Listeners

(732) 247-5555

<http://www.scarletlisteners.com>

Free and confidential peer counseling and referral hotline, providing a comforting and supportive safe space.