Syllabus—Spring 2024
BIOL 4310—Advanced Bioinformatics
Utah Tech University

Instructor
Dr. Randy Klabacka
E-mail: randy.klabacka@utahtech.edu
Office: SET 513
Office hours: Tue 3:00-4:30
Wed 12:30-1:20* ; 3:00-4:30
Fri 1:00-2:00
**First Wed office hours in SET 403*

Course Teaching Assistant
Name: Seun Onileowo
E-mail: seun.abimbola.onileowo@utahtech.edu

Course Details

Credits: 3
Time: Tue/Thu 1:30 – 2:45 pm
Place: SET 501
Prerequisites:
BIOL 3030 (Introduction to Genetics)
BIOL 3300 (Introduction to Bioinformatics)
Course materials:
Laptop with access to Unix-based operating system

Course description

Examination, development, and application of scripting and genomic tools to perform genomic analyses for functional and evolutionary genomics.

Course Learning Outcomes

The Advanced Bioinformatics course at Utah Tech University has developed a set of program learning outcomes (CLOs). At the successful conclusion of this course, students will be able to:

- CLO 1: Describe a data analysis process.
- CLO 2: Create computer programs that facilitate biological data analysis.
- CLO 3: Interpret the analysis results biologically and explain the implications of them.

Program Learning Outcomes

The Biological Sciences Department at Utah Tech University has developed a set of program learning outcomes (PLOs). The key DBS student learning outcomes that we will address in BIOL 3030 are:

- PLO 1b: Students will describe and explain the centrality of genetic systems' governance of life's actions from the cellular to the phyletic.
- PLO 1d: Students will describe and explain the dynamic interaction of living systems with each other and their environments.
- PLO 1e: Students will describe and explain the transforming role of evolution in changing life forms and how evolution explains both the unity and diversity of life.
- PLO 2: Students will employ scientific methods to acquire, analyze and apply knowledge of biological phenomena.
- PLO 4: Reading Comprehension: Students will analyze and critique scientific literature: identifying hypotheses, critiquing methods, interpreting data and results, and articulating the context of discussions. predications.

Course Assignments and Grading

Course points

Item	Points
Coding Exercises	120
Code Review Paper	60
Exam	40
Class Project	80
Total:	300

Letter grade distribution

Percentage	Grade
>= 90.0	A
80.0 – 90.0	B
70.0 – 80.0	C
60.0 – 70.0	D
<60.0	F

Coding Exercises

Throughout the semester, you will complete assigned coding exercises to hone the skills you are learning in class. Each exercise is worth 10 points. Each exercise will be graded at the instructors discretion depending on the tasks required (some exercises only require completion, others require correct formatting and structure).

Exercise	Topic	Due date
Exercise 0	Pre test	Jan 11
Exercise 1	Mini Pipeline	Jan 24
Exercise 2	SARS-CoV2 Pipeline	Feb 2
Exercise 3	Cleaning and Mapping data	Feb 23
Exercise 4	Differential gene expression	Mar 2
Exercise 5	Calling variants	Mar 9
Exercise 6	Filtering SNPs	Mar 23
Exercise 7	Variant analysis	Apr 6
Exercise 8	GWAS	Apr 13
Exercise 9	Functional Genomics	Apr 20

Code Review

As a scientist, you will be asked to peer-review research manuscripts and associated code to assess the quality and reproducibility of the experiments to address the research question and the accuracy of their interpretations of the results. To practice these critical thinking skills and to assess your ability to apply the class concepts, you will choose a manuscript from the primary literature that utilized a significant scripting component to simulate, process, or analyze their data.

By January 31, you must have your chosen manuscript approved by the instructors. Failing to submit your selected manuscript will result in a deduction of 10 points.

You will write a two-page review of the article, focused on the associated code and your ability to reproduce the results presented in the published manuscript (papers significantly longer or shorter than two pages may lose points based on the instructor's discretion). Fully reproducing the results may not be feasible, but you should still discuss whether the authors provided what is necessary to do so (and if not, what was missing?). We will have review check-points where you will present the progress of your review with your peers. Each check-point will be announced one week before it occurs. The final review is **due on March 30**.

Code Review Rubric

Item	Points
Summary of manuscript objectives	20
Language(s) used	2
Number of scripts	2
Tools/packages used and why	6
Purpose of each script	10
Examination of code	30
Documentation critique	7
Code commenting documenting	7
Reproducibility critique	7
References to specific lines of code	10
Conclusion	5
Summary of critiques	5
Total:	60

You may need to look up best practices for other code if scripts aren't in python

Class Project

You will develop a script, or set of scripts, to simulate, process, and/or analyze biological data. Team members will work collaboratively via version-control software (Git) and host your code in a repository on GitHub. At the end of the semester, each team will give a short presentation in the form of a workshop/tutorial. The goal is for the team to guide the rest of the class through an example of how to use their script(s). The tutorial should have enough documentation to be self-contained; i.e., other students should be able to complete the tutorial without the presentation.

If possible, we want this project to be applicable to your research. To enable this, we will be flexible about how students assemble into teams. Teams can consist of 1–6 students.

By February 7, each team must submit a short project proposal to be approved by the instructor. The project can change after the proposal is approved, but a change MUST be approved by the instructor. The proposal will be submitted as documentation within a GitHub page, with all team members included as collaborators. Failure to submit a proposal by February 7 will result in a 5 point deduction from the final grade.

A link to the code and documentation for the project is **due by April 24**. Presentations will be in random order in class on **April 25 and April 27**.

A student's grade will be determined based on the quality of the project, the ability of others to replicate their results, and their individual contributions to the project (assessed by Git/GitHub's tracking of every line of code written or modified by each contributor). More specifically, the break down of your grade for the project will be as follows:

Class Project Rubric

Item	**Points**
Documentation	35
Summary of project objectives	10
Language(s) used	3
Number of scripts	2
External tools/packages used and why	10
Purpose of each script	10
Code	35
Following best practices (including commenting, documenting, modularity)	20
Reproducibility	15
Presentation	30
Documentation walk-through	10
Example run (using abbreviated data if needed)	7
Answering questions	5
Clarity and smoothness of presentation (no bugs)	5
Staying within time limit	3
Total:	100

You may need to look up best practices for other code if scripts aren't in python

Exam

There will be one exam in the course, which will be taken at the end of the semester. Students will be provided a study guide to prepare them for the exam.

Schedule

Week 1 (Jan 8 - 14): Review of basic bioinformatics		
Tuesday	In-class topic:	Syllabus, how to succeed in this class, and assessment
Thursday	In-class topic:	Discussing pre-test items
	Due before class:	Exercise 0
Week 2 (Jan 15 – 21): Introduction to scripting		
Tuesday	In-class topic:	Basics of Bash scripting
Thursday	In-class topic:	Basics of Python scripting
Week 3 (Jan 22 – 28): Overview of bioinformatics pipelines		
Tuesday	In-class topic:	Introduction to pipelines
Thursday	In-class topic:	Pipelines cont.
	Due before class:	Exercise 1
Week 4 (Jan 29 – Feb 4): Experimental design for genomics		
Tuesday	In-class topic:	What questions can be answered using genomics?
	Due before class:	Select manuscript for code review
Thursday	In-class topic:	Review of sequencing approaches
	Due before class:	Exercise 2
Week 5 (Feb 5 – 11): High-performance computing		
Tuesday	In-class topic:	Introduction to using a remote command line
	Due before class:	Select class project
Thursday	In-class topic:	Introduction to batch submission
Week 6 (Feb 12 – 18): Cleaning and mapping		
Tuesday	In-class topic:	Cleaning raw data for short-read sequencing
Thursday	In-class topic:	Cleaning raw data for long-read sequencing
Week 7 (Feb 19 – 25): Gene expression		
Tuesday	In-class topic:	Mapping reads for short-read sequencing
Thursday	In-class topic:	Mapping reads for long-read sequencing
	Due before class:	Exercise 3
Week 8 (Feb 26 – Mar 3): Variant calling and filtering		
Tuesday	In-class topic:	Differential gene expression pt. 1 (StringTie)
Thursday	In-class topic:	Differential gene expression pt. 2 (DeSeq2, EdgeR)
	Due before class:	Exercise 3
Week 9 (Mar 4 – 10): Variant analysis pt 1		
Tuesday	In-class topic:	Discovering variation (GATK, BCF tools, samtools)
Thursday	In-class topic:	Variant filtration (BCF tools, VCF tools, BED tools)
	Due before class:	Exercise 4
Week 10 (Mar 11 – 17): NO CLASS	Spring Break	
Week 11 (Mar 20 - 24): Variant analysis pt 2		
Tuesday	In-class topic:	Examining allele frequencies (VCF tools)
Thursday	In-class topic:	Population genetics (genomics general package, VCF tools)
	Due before class:	Exercise 5
Week 12 (Mar 25 – 31): GWAS		
Tuesday	In-class topic:	Introduction to genome-wide association studies (PLINK)
Thursday	In-class topic:	GWAS assumptions and caveats
	Due before class:	Code review
Week 13 (Apr 1 – 7): Functional genomics pt 1		
Tuesday	In-class topic:	Functional variant annotation (SIFT, Polyphen-2)
Thursday	In-class topic:	Locus reconstruction with allelic variation (GATK, BED tools)
	Due before class:	Exercise 7
Week 14 (Apr 8 – 14): Functional genomics pt 2		
Tuesday	In-class topic:	Peptide translation (Biopython)
Thursday	In-class topic:	Phylogenomics
	Due before class:	Exercise 8
Week 15 (Apr 15 – 21): Pipeline efficiency tools		
Tuesday	In-class topic:	Pipeline best practices (Snakemake)
Thursday	In-class topic:	Delayed progression
	Due before class:	Exercise 9
Week 16 (Apr 22 – 28): Class project presentations		
Tuesday	In-class topic:	Presentations
Thursday	In-class topic:	Presentations
Week 17 (Apr 29 - May 3): Exam		
Tuesday April 30 at 1:30 pm	**Final Exam**	SET 501

Course Policies

Class Inclusiveness

It is our intent that students from all backgrounds and perspectives be well-served by this course. We are committed to creating an inclusive space that fosters diversity along its many axes: ethnicity, race, sex, gender, disability, age, socioeconomic status, nationality, and culture. As your instructor and as a student in this class, it is our shared responsibility to develop and maintain a positive learning environment for everyone. Any type of discrimination or aggression toward your peers or instructor will not be tolerated. If you experience any form of discrimination in this course, please report your concern at https://utahtech.edu/report-a-concern/.

Late Assignment Policy

It is very important that students submit work on time, or they will find it very difficult to catch up. All work in the course (e.g., assignments, discussions, exams, quizzes, etc.) will be due by the date noted on Canvas for each assignment. Students should reach out to their instructor immediately to discuss any concerns.

Makeup Policy

Students who miss the normal exams will need to contact the instructor and turn in the valid excuse within 48 hours from the time that the exams were given. The makeup exam schedule is determined by the instructor and will need to be done within ONE week (5 work days) from the time that the exams were given. The format, questions and difficulty-level of make-up exams are not guaranteed to be same as the normal exam, which are at the discretion of the instructors. The student and the instructor will agree upon a date and time for the makeup exam. Valid excuses include: 1). illness documented by a physician. 2) evidence of personal or family emergency. 3) official university excuses (e.g., university functions).

Instructor Communication Policy

It is your responsibility to read course announcements sent by your instructor (over Canvas and through email). Your Utah Tech University email address is the university-approved form of communication between instructors and students. Please use email rather than the Canvas messenger feature, as email makes message threads easier to follow. Students should give the instructor 48 hours to get back to them on any communication, and one week for grading turnaround time on major assignments.

Academic Integrity

Academic dishonesty is an offense that will be reported to the UTU Academic Integrity Committee and handled by the instructor following University Policy 555. Acts of academic dishonesety include group work on quizzes/exams, sharing notes, communicating quiz/exam materials, and plagiarizing materials (including using machine learning natural language processing such as ChatGPT). Please refer to the following document for further information regarding academic honesty: UTU Academic Integrity Committee

Accessibility

Students who need accommodations are asked to electronically submit their approved accommodations to the UTU Disability Resource Center. If you need accommodations but have not established them, make an appointment with Disability Resource Center, North Plaza Building (Beside the Testing Center), 435-652-7516.

Generative AI

Generative AI models, such as language learning models, are recognized as valuable tools for enhancing the learning process in this course. These models are approved for guiding your learning experience, aiding in creating practice problems, and facilitating solution exploration. However, it's important to emphasize that while generative AI can support your understanding, it must not be employed to directly answer questions or complete assignments. Use of generative AI to develop practice problems and work through their solutions is encouraged as a study method. Misuse of generative AI will be considered an act of academic dishonesty in this course. If you have any uncertainties or inquiries regarding the appropriate utilization of generative AI, please reach out to the instructor. Open discussions about responsible AI implementation are welcomed, ensuring that we collectively uphold the educational value of these technologies. (This section was written with the help of chatGPT).

Important University Dates

See the UTU Academic Calendar

- Jan 8: Date classwork begins
- Jan 8: Tuition and fees due
- Jan 12: Last day to add without instructor signature
- Jan 15: Dr. Martin Luther King, Jr. Day (no classes)
- Jan 18: Drop/Audit fee begins
- Jan 20: End of 100% Refund Period
- Jan 29: Pell Grant Census
- Jan 29: Last day to drop
- Jan 29: Last day for refund

- Feb 5: Last day to add/audit
- Feb 19: President's Day Holiday (no classes)
- Mar 1: Midterm grades posted
- Mar 5: Last day to withdraw from class
- Mar 11-15: Spring Break (no classes)
- Apr 25: Date classwork ends
- Apr 26: Reading Day
- Apr 29-30 & May 1-2: Final Exams
- May 6: Final grades posted

Useful Links

Disability Resource Center
IT Help Desk
Library
Testing Center
Tutoring Center
Writing Center

Title IX Statement

Utah Tech University affirms its commitment to the promotion of fairness and equity in all aspects of the educational institution. Harassment and discrimination – including sex/gender discrimination, gender identity, gender expression, sexual harassment, sexual misconduct, gender-based violence, dating violence, domestic violence, stalking, pregnancy or parental , family or marital status and or retaliation –not only disrupts our commitment to maintaining an environment in which every member of the University community is treated with respect and dignity, but may also violate University policy and federal, state, and/or local law.

Should you or someone you know experience behavior that is coercive, discriminatory, harassing, and or sexually violent in nature, or if you or someone you know has questions about their rights and options regarding such behavior, you are encouraged to contact:

– Hazel Sainsbury, Dir. Of Equity Compliance, Title IX Coordinator: 435-652-7747 (ext. 7747) hazel.sainsbury@utahtech.edu ; titleix@utahtech.edu

Incidents may also be reported directly to law enforcement, either separately or in conjunction with any report made to the University's Title IX Coordinator, and the University will aid in making contact if requested.

– Utah Tech University Police: 435-275-4300 or by calling 9-1-1.

Maintaining a safe and inclusive University community is a shared responsibility. For more information on how Title IX protections can benefit you and help us keep a productive campus environment, visit titleix.utahtech.edu to learn more.

Utah Tech Email Disclaimer

You are required to frequently check your Utah Tech email account as important class and university information will be sent to this account, including bills, financial aid/scholarship notices, notices of canceled classes, reminders of important dates and deadlines, course information, and other information critical to your success at UT. To access your Utah Tech email ac-

count, visit mail.utahtech.edu. Your email account username is Digital-ID@utahtech.edu (e.g. D12345678@utahtech.edu). If you don't know or have forgotten your Digital-ID or password, please visit changepassword.utahtech.edu.

References and Changes to Syllabus

The instructor will reference the syllabus when students have questions on course objectives, policies, and grading. The instructor can make changes to the syllabus during the course of the semester, but these changes will always be reasonable and announced to the class.