Department of Biology
BIOL 435/535: Bioinformatics
Spring 2021; MSEC 195; https://zoom.us/j/99063029445?pwd=S0NKbzJFb0haVWIvRkUzSmFML1RkUT09
MWF: 9:50am – 10:40am

Instructor: Dr. Joel Sharbrough
Office: 313 Jones Annex
Phone: (575) 835-5347
Email: Joel.Sharbrough@nmt.edu

Office Hours: Wednesday: 11 am – 12 pm; https://zoom.us/j/95601690680?pwd=V0ZnMEIUdUwwSTdnV1NJb3ZBZlhtUT09
Thursday: 9 am – 11 am; https://zoom.us/j/96441419593?pwd=UGo5R3ZZL2hkZDJES3hydVdQam5yZz09

Course Description:

Bioinformatics represents one of the fastest growing fields in the life sciences and dealing with the massive data sets that are now commonplace requires comfort and confidence in computational techniques couched in biological understanding. This course will take a Biology-centered approach to the discipline, in which students learn the *why* along with the *how*. To accomplish this, the course will begin with low-throughput, high-quality methods and gradually progress to high-throughput techniques that are necessary for dealing with Big Data, which will be implemented on the newly installed Discovery High Performance Computing cluster housed at NMSU. The course is designed for upper-division undergraduate (BIOL 435) and graduate students (BIOL 535/BIOT 535) of all skill levels. Students will complete homeworks designed to ensure basic competence and familiarity with essential bioinformatic functions (e.g., sequence alignment, quality control, read mapping, variant calling) and to introduce exciting new methodologies. Students will also write two papers during the semester, and BIOL 535 enrollees will participate in a class project and present their portion of the project to the class. If the results of the class project are sufficiently novel, the results will be written up and submitted for publication.

Mode of Instruction:

The course will be delivered both in person and via zoom. In-person attendance is limited to the room capacity defined in accordance with COVID-19 safety protocols. All lectures and associated slides will be recorded and posted on Canvas.

Pre-requisites:

BIOL 311; Genetics

Place in Curriculum:

Bioinformatics is intended for upper-division undergraduate students in the Biology or Biomedical Sciences programs and graduate students from the Biology or Biotechnology programs. Students from other programs should speak with the instructor to ensure they have the necessary foundation in Genetics.

Learning Objectives:

By the end of this course, students should be able to:

- Understand the distinction between different types of homology (i.e., orthology and paralogy) and assess whether nucleotide and amino acid sequences are homologous
- Perform and evaluate the quality of multiple sequence alignments from nucleotide and amino acid sequence data
- Manage, manipulate, and assess quality of large datasets, especially including next-generation-sequencing datasets
- Understand the problems and pitfalls associated with variant calling, as well as how to interpret genotype likelihoods
- Have the confidence to learn new and useful techniques that they may encounter in their bioinformatic exploits

Program learning outcomes:

https://nmt.edu/academics/biology/undergrad.php

Grading:

Item	BIOL 435	BIOL 535
Participation	20%	20%
Homeworks (16)	60%	50%
Scientific Writing Assignments* (2)	20%	20%
Project Presentation**	-	10%

*** – BIOL 435 enrollees will write a description of a proposed method to solve a biological question for the second writing assignment. BIOL 535 enrollees will complete a project report for their section of the class project.**

**** – Only BIOL 535 enrollees will be required to give a 7-10 minute Project Presentation, which will focus on the results of their portion of the class project.**

Participation:

Participation includes attending classes, asking questions, contributing to group/class discussions, reading assigned papers, and providing answers to individual reading prompts.

Homeworks:

Homeworks will be due every Sunday by 11:59pm and submitted through Canvas. New assignments will be made available one (1) week in advance and will be announced to the class via Canvas and email. Difficulties interfacing with Canvas should be brought up with the instructor so that appropriate accommodations can be made.

Scientific Writing Assignments:

There will be two scientific writing assignments for both BIOL 435 and BIOL 535. The first assignment will be shared by both courses and will consist of a 4-5 page paper describing a popular bioinformatic application, the tools that are commonly employed to accomplish that application, and the scientific work that has relied upon them. The first writing assignment is due **at 11:59pm on SUNDAY, MARCH 14th, 2021** and will be submitted through Canvas.

The second writing assignment is due at **5 PM on Friday May 7th, 2021** and will be different across the two sections. For **BIOL 435**, students will write a 2-3 page methodological proposal for answering one of the biological questions provided by the instructor. The proposal should include a Background/Justification section, a Specific Aims section, a Research Plan section, and a Potential Pitfalls and Alternative Strategies section. Figures are not mandatory, but are encouraged. Students wishing to design a bioinformatic approach to a biological question of their

own design must consult with the instructor by **5pm** on **Friday, April 16th, 2021.** An outline and/or first draft must be submitted via Canvas by **5pm on Friday April 23rd, 2021.**

Students enrolled in **BIOL 535** will complete a 2-3 page write up of their portion of the class project (see below). The write-up should include an Introduction, a Methods section, a Results section, and a Discussion/Conclusions section. Additionally, students must provide at least one Figure and Figure Legend displaying the results of their analyses. All methods and data files must be deposited on Canvas by **5pm on Friday April 23rd, 2021** to allow for verification by fellow classmates.

Project Presentation:

Students enrolled in **BIOL 535** will also present their portion of the class project in the final week of classes (**April 26-30th, 2021**). Presentations will be 7-10 minutes long as will follow the same general format as the second writing assignment.

Class project:

Students will together choose one of three class projects presented by the instructor. Specific roles in the project will be chosen by volunteer, and/or random assignment. Students will need to meet with the instructor during Office Hours to discuss the specific steps involved in their assigned project **no later than April 1st, 2021**. Because of the nature of Bioinformatic pipelines, some students will need to complete their portion of the project earlier in the semester than others, so students with prior experience in Bioinformatics/coding will be assigned to earlier portions of the project. If deadlines are not met, students will need to work with instructor to finish necessary steps so that other students can begin downstream aspects of the project.

Schedule:

Week	Day	Topic	Reading	Homework	Notes
1	W – 1/20	Homology	-	BLAST	
	F – 1/22	BLAST	Altschul et al., 1990		
2	M – 1/25	Pairwise alignment	Needleman & Wunsch 1970	Alignment	
	W – 1/27	Multiple Sequence Alignment	Smith & Waterman 1981		
	F – 1/29	Bad Alignments	Venkat et al., 2018		
3	M – 2/1	Gene Architecture	-	Gene Models	
	W – 2/3	Finding Genes/ Primer Design	ENCODE 2011		
	F – 2/5	Transposable Elements	Doolittle 2013		
4	M – 2/8	Mutations & Substitutions	-	SNPs	
	W – 2/10	SNPs	Amberger et al., 2015		
	F – 2/12	GWAS & Racism in Genetics	Lee et al., 2018		
5	M – 2/15	Phylogenetics	Baum et al., 2005	Tree Thinking	
	W – 2/17	Maximum Likelihood vs. Bayes	Felsenstein 1981		
	F – 2/19	No Class, Academic Holiday			
6	M – 2/22	Molecular Evolution	Yang 2007	Phylogenetics/ Population Genetics	
	W – 2/24	Population Genetics I	Tajima 1989		
	F – 2/26	Population Genetics II	Charlesworth 2009		
7	M – 3/1	Intro to Unix	-	Discovery On-Boarding	
	W – 3/3	Parallelized Computing	-		
	F – 3/5	Data Management	Bowman & Woo 2018		
8	M – 3/8	NCBI and Other Data Repositories/ Project I	-	NCBI SRA/ Rclone	**** First writing assignment due 3/14****
	W – 3/10	GitHub/ Project II	-		
	F – 3/12	Rclone/ Project III	-		
9	M – 3/15	Next-Generation Sequencing Data	-	Burrows-Wheeler Transform/ Read Mapping	
	W – 3/17	NGS Trimming, Filtering, and Quality Control	Burrows & Wheeler 1994		
	F – 3/19	Short-Read Alignment I	-		
10	M – 3/22	Short-Read Alignment II	-	Variant Calling	
	W – 3/24	Principles of Variant Calling	McKenna et al., 2010		
	F – 3/26	Visualizing Variants with IGV	-		
11	M – 3/29	Genome and Transcriptome Assembly	-	Transcriptome Assembly	**BIOL 535 students need to have met with instructor by 4/1 to discuss class project role**
	W – 3/31	RNASeq	Haas et al., 2013		
	F – 4/2	No Class, Academic Holiday			
12	M – 4/5	Quantifying Gene Expression	-	Gene Expression	
	W – 4/7	Splice-Aware Mapping	Kim et al., 2019		
	F – 4/9	Metagenomics	-		

13	M – 4/12	Introduction to PacBio and Nanopore Sequencing	-	Long Reads/ Structural Variants	
	W – 4/14	Structural Variant Detection	-		
	F – 4/16	Finding Genes: A Reappraisal	Sharon et al., 2013		
14	M – 4/19	Population Structure	-	GWAS	**BIOL 435 submit outline/first draft of 2nd writing assignment by 5 PM 4/23** **BIOL 535 submit methods/data files by 5 PM 4/23**
	W – 4/21	GWAS II – Doing It Right	Coop 2019		
	F – 4/23	Proteomics	-		
15	M – 4/26	HMMs/Machine Learning/AI	-	AI/Machine Learning	**BIOL 535 Presentations**
	W – 4/28	Class Project Presentations	-		
	F – 4/30	Class Project Presentations	-		
Finals Week	F – 5/7	**** Second Writing Assignment due 5:00 PM 5/7****			

Policies on attendance, late assignments, and missed in-class discussion:

This semester presents unique challenges as a result of COVID-19. As such absences, late assignments, and missed discussion will be handled on a case-by-case basis, but communication with the Instructor is required. Students are expected to do their best to attend classes, participate in group discussions, and turn in homeworks and writing assignments on time. For the most up-to-date information on COVID-19, visit the NMT COVID-19 page at https://www.nmt.edu/covid19/.

Academic Honesty:

New Mexico Tech's Academic Honesty Policy for undergraduate and graduate students is found in the student handbook, which can be found at: https://www.nmt.edu/studentlife/dos/NMT%20Student%20Handbook%202019-20.pdf. You are responsible for knowing, understanding, and following this policy.

Reasonable Accommodations:

New Mexico Tech is committed to protecting the rights of individuals with disabilities. Qualified individuals who require reasonable accommodations are invited to make their needs known to the Office for Disability Services (ODS) as soon as possible. To schedule an appointment, please call 835-6209, or email disability@nmt.edu.

Counseling Services:

New Mexico Tech offers individual and couples counseling, safety assessments, crisis intervention and consultations through The Counseling Center. These confidential services are provided free of charge by licensed professionals. For more information, please call 835-6619, email counseling@nmt.edu or complete an Intake Form on our website at https://www.nmt.edu/cds/. All services are provided via phone or Zoom during the Covid-19 pandemic.

Respect Statement:

New Mexico Tech supports freedom of expression within the parameters of a respectful learning environment. As stated in the New Mexico Tech Guide to Conduct and Citizenship: “New Mexico Tech’s primary purpose is education, which includes teaching, research, discussion, learning, and service. An atmosphere of free and open inquiry is essential to the pursuit of education. Tech seeks to protect academic freedom and build on individual responsibility to create and maintain an academic atmosphere that is a purposeful, just, open, disciplined, and caring community.”

COVID-19 Safety Issues for Face-to-Face Instruction:

Students must follow campus-wide safety protocols, including mandatory use of face coverings and maintaining a minimum of 6 ft social distance from other students and faculty. Students should not enter the classroom earlier than 10 minutes prior to start of class, and should exit the classroom within 10 minutes of the end of class. Students who fail to comply are subject to disciplinary procedures.

Title IX Reporting:

Sexual misconduct, sexual violence and other forms of sexual misconduct and gender-based discrimination are contrary to the University’s mission and core values, violate university policies, and may also violate state and federal law (Title IX). Faculty members are considered “Responsible Employees” and are required to report incidents of these prohibited behaviors. Any such reports should be directed to Tech’s Title IX Coordinator (Dr. Peter Phaiah, 20D Brown Hall, 575-835-5187, titleixcoordinator@nmt.edu). Please visit Tech’s Title IX Website (www.nmt.edu/titleix) for additional information and resources.