

**BIOL/BINF 5352: Introduction to Bioinformatics II**  
**Spring 2019**  
**Monday, 3:30 pm to 5:30 pm (UGLC 338)**  
**Wednesday, 3:00 pm to 6:00 pm (UGLC 338 wet labs TBN)**

**Unofficial Course Description:**

This is a continuation of BINF 5351: Bioinformatics I, offered last semester. In the previous course, you studied the basics of sequence comparisons, multiple sequence alignment, pattern recognition, and tree construction on a limited number of identified sequences. In the current course, we will expand the application of these principles to real world, genomic scale data and apply them within the context of the cell and Cell Biology. Topics include genome assembly and finishing, identification and annotation of protein genes and metabolic pathways, determination of mRNA and protein expression, and advanced RNA analysis and folding. As time permits, we will also investigate current, specialized topics of bioinformatics in research and industry. If there are topics of particular interest to you, please talk to me and we'll see if they can be incorporated into the syllabus. Throughout the semester we will also be modeling the requirements for independent work in the bioinformatics field: reading and understanding current research presented in the primary literature; and the development, implementation, and presentation of an independent research project.

**Student Learning Objectives:**

**The Successful student will be able to:**

- 1). Read, understand, and apply bioinformatic data and tools gleaned from primary literature.**
- 2). Understand and apply bioinformatic sequence analysis in the context of the whole cell.**
- 3). Demonstrate practical experience in the implementation of specialized bioinformatics software for genome scale analysis and higher.**
- 4). Synthesize and cogently present on scientific topics and systems.**
- 5). Design, implement, and present an independent bioinformatic research project.**

**Lecture Professor:**

**Dr. Stephen Aley** is a molecular biologist and biochemist specializing in infectious disease. He has been part of the genome project for *Giardia lamblia*, and still consults on related projects in infectious disease and bioinformatics. Recently, his day-job changed (again), and he now serves as the Associate Vice President for Research. His primary residence is in the Office of Research and Sponsored Projects (Administration Building, Rm 209C). The good news is that the staff at the front desk in ORSP can usually tell you where he is and when he might be back in his office! He is most reliably contacted by email ([saley@utep.edu](mailto:saley@utep.edu)), however he has been known to respond to the course BlackBoard discussion group or even answer his direct phone line (747-8813). Given the small size of this class, he will not have fixed office hours, but will instead encourage students to schedule individual or group meetings as needed. There is a rumor that he has a new candy dispenser and has stocked it with chocolate instead of sugar free candies, but please don't tell his wife. Don't worry – he doesn't normally talk about himself in the third person.

**Course Structure:**

This semester's class will be organized more like a seminar or small group discussion. We will still meet each week as a class, but BEFORE EACH face to face meeting, you are required to read

and review the foundational content for that topic, provided on the Blackboard course site. During our meeting time, we can, as a group, explore the material in greater depth and discuss any questions raised by students or professor. In addition, we will be reviewing progress on computer and written assignments at this time. Monday is nominally the discussion portion of the class, running from 3:30 pm to 5:30 pm. Wednesday is scheduled a laboratory meeting, running from 3 pm to 6 pm, however, the scheduling of the discussion portion of the class will vary depending on the schedules of the professors and guest lecturer/laboratory director. The course schedule, on the Blackboard course site, contains the current schedule. As with all experiments, there will likely be small modifications to the structure as the semester progresses.

### **What happened to the Laboratory?**

The experiential computer projects are still a key part of the course. Most weeks will include a new or continuing computer assignment, but generally as a take home project rather than a fixed laboratory session, with progress to be reviewed in our regular meetings. Wet Lab projects (e.g., Mass Spectrometry, DNA sequencing, etc) will be individually schedule, however, they will be held during one of the regular course times, as scheduled.

### **Teaching Assistants:**

Because of the small size of this class, we will not have an assigned teaching assistant for this semester. For issues with the Bioinformatics computer systems and software, you may also seek help from Jon Mohl, systems analyst for the Bioinformatics program. His email is [jemohl@utep.edu](mailto:jemohl@utep.edu). All other questions should be directed to me.

### **If this is Monday, Where am I supposed to be!?**

Most classes are scheduled to be held in UGLC 338, however, we have the option of changing locations to meet specific needs (e.g., we could use BE 300 when computers are needed or other specialized laboratories for the various “wet” labs on the schedule. The table below gives the current schedule, but this will likely change!! Please check your email and the Blackboard course site regularly for updates!!

### **So, what book do I need to buy?**

At this level of course, books are primarily for reference use. The textbook from Bioinformatics part I, **Bioinformatics** by David Mount, will still be useful, so don't give it away. In addition, you will probably want access to an undergraduate level biochemistry or molecular biology text, but you may choose which text you use. In particular, I would suggest that you check with your classmates or a friendly professor to see if they have a book they could loan (or give) you for the semester. Finally, you will find use for a programming guide on your programming language of choice, and a basic guide on mysql. I can tell you that for me, using perl and mysql, I have found the most useful reference to be Google...

### **You've mentioned “Blackboard course site” four times, already...**

The course will be coordinated through a **Blackboard** course site, available through your logon to “my.utep.edu”. This site is the primary source of course readings and other handouts (including a copy of this syllabus), course calendar, and some supplemental web sites. There are also discussion boards for after class interaction as well as quizzes and other assessments. All course grades will also be presented through Blackboard. The site is reasonably self-explanatory,

however if questions arise, please talk to me earlier rather than later!!! If things get complicated, I recommend chamomile tea to soothe the nerves...

### **What Computer will I use?**

The short answer is essentially whichever one you want to! Almost all of the software we use in this portion of the course is open source and/or publically available and able to be compiled on most systems. The scale of the computational problems in this semester will be of a scope that is best handled by **command line interfaces** (e.g., unix, linux, or even DOS), customized programs, or scripts for parsing output. My preference (out of habit) is perl, but every year there are a few class members who try to convince me of the superiority of PYTHON. Some day they might succeed. Regardless, you may use the language of your choice. You will also want to be able to query items stored in a mySQL database – but this, too, is open source, and all components are also available on any major platform.

All computational labs may be done in the Bioinformatics computer lab, in Physical Science Annex, on the Linux and Unix platform machines there. In addition to times when you have direct access to these machines, they are also accessible via the Internet through SSH. That being said, you will find that that most bioinformaticians perform much of their day-to-day work on a personal laptop! If you have a laptop computer, I strongly recommend that you practice using it in and for the course. Even when your complete program may require a mainframe, you can often accomplish much of the setup and debugging – and analysis! -- on your personal machine.

### **Class Components and Grades:**

The primary goal of this course is to develop and practice approaches and skills that would apply to a full-scale bioinformatics project. As such, the grading for the course will emphasize the practical application of acquired skills. Specific grade components of the lecture will include Class Presentations, Homework, a stand-alone White Paper, review Exam, Final Exam; and major Class Project, including written and oral presentations.

### **Grade Point Distribution:**

#### **Class Participation 15%**

For each class, each student is expected to have prepared for the class by completing assigned reading or other work and to actively participate in the class discussion and exploration. Grading for this component is subjective and will be recorded in BlackBoard in a timely fashion.

#### **Homework Assignments 10%**

Homework assignments are weekly writing or other research assignments to be done out-side of class and turned in according to individual deadlines. In some cases, these assignments will overlap with the laboratory portion of the course.

#### **White Paper on Research Project 10%**

A “White Paper” is a one to two-page summary of a proposed Research Project (topic area to be assigned). Full formatting details are available on BlackBoard, under “Handouts.” Grading will be by the “White Paper Rubric” found on Blackboard. Assignments include White paper hypothesis, due (on or before the start of class) on Monday, Feb 5, and a draft version of the white paper, due February 14. The final draft of the document is due February 23.

#### **First Individual Exam 5%**

This Exam will review Biochemistry and Molecular Biology knowledge. There will be a review outline on BlackBoard, but this really should be only a review of concepts from your earlier studies. The exam will be tightly timed on BlackBoard, the week of Feb 5, with the exact time to be determined by the instructor in consultation with class members.

### **Informatics Research Project 20%**

Each student will develop, implement, and present orally and in writing, an informatics research project requiring topic research plus application (or development) of bioinformatics tools, and interpretation and presentation of results. Actual topic is flexible, but requires approval of professor. An oral presentation of the project and results will be presented in class in the last two weeks of the course (order determined later). The final written report is due the last day of class. Earlier submissions are encouraged!

### **Final Examination 20%**

The final examination will be used for assessment of understanding and application of basic concepts and applications. The actual format is flexible (start lobbying now!), but please reserve the final exam time scheduled by the registrar for this class (Monday, May 13, 1 – 4:30 pm).

### **Computer and Wet Lab Exercises 20%.**

Each week, there will be either a computer assignment or a wet lab assignment. Write-ups for weekly computer project and wet lab assignments should be submitted to the instructor through email, in pdf format. Due date/time will be 3 pm on the Monday following the assignment or lab.

### **COURSE POLICIES POLICY ON MAKEUP EXAMINATIONS AND LATE WORK:**

NO makeup exams or due date extensions will be given for reasons other than illness (doctor's note required) or excused absence (e.g., travel on official University business) with the instructor's **prior** approval, or travel on official University business. Makeup exams for excused absences will be scheduled at the instructor's discretion. The same policy will be followed for missed laboratory work.

### **WHOSE WORK IS IT:**

For ALL ASSIGNMENTS and PROJECTS, you are to ASSUME that all work is to be done by the INDIVIDUAL student whose name is on the cover page unless the WRITTEN instructions specifically state otherwise. When teams are allowed and used, ALL PARTICIPATING TEAM MEMBERS must be listed on the cover page of the assignment and list the specific contributions of each team member. When outside work is incorporated into weekly or project presentations (e.g., perl code from outside sources or copied illustrations on PowerPoint slides) the source must be clearly credited. All written reports must clearly cite sources in the format of a standard Scientific Journal (your choice!). To do otherwise is academically dishonest and will not be tolerated.

### **POLICY ON ACADEMIC HONESTY:**

Academic Dishonesty will not be tolerated. All university guidelines will be strictly followed. Please read the university guidelines carefully. If you have any questions regarding the university policy please contact the Dean of Students.

**DISABILITY STATEMENT:**

If a student has or suspects s/he has a disability and needs an accommodation, s/he should contact the Disabled Student Services Office (DSSO) at 747-5148 or at <dss@utep.edu> or go to Room 106 Union East Building. The student is responsible for presenting to the instructor any DSS accommodation letters and instructions.

The schedules below may be subject to change, depending on the needs of the course and requests of the course participants:

Date	Day	Lect	Topic
1/23/19	W	Lect 01	Introductions; Syllabus and major assignments; Data formatting and data mining. Personal Databases
1/28/19	M	Lect 02	Sequencing, Assembly, & Finishing
2/4/19	M	Lect 03	Gene Calling: de novo & empirical
2/13/19	W	Lect 04	Annotation I
2/18/19	M	Lect 05	Annotation II
2/25/19	M	Lect 06	Genome Comparison I
3/4/19	M	Lect 07	SNPs
3/11/19	M	Lect 08	RNA Hybridization and MicroArray I
3/25/19	M	Lect 09	Analysis: MicroArray II
4/8/19	M	Lect 11	Epigenomics
4/15/19	M	Lect 12	Proteomics - MS/MS
4/22/19	M	Lect 13	Protein folding and ligand binding
4/29/19	M	Lect 14	RNA Folding
5/6/19	M	Projects	Projects

Date	Day	Lab	Topic
4/1/19	M	Lab 10	Wet Lab: ABI sequencing
4/17/19	W	Lab 12	Wet Lab: MS/MS laboratory
4/24/19	W	Lab 13	Protein folding lab
5/1/19	W	Lab 14	Wet Lab: Next Gen Sequencing

Date	Day	Final Exam	Topic
5/13/19	M	Fin Exam	1 pm UGLC 348