

**BIOL/BINF 5352: Introduction to Bioinformatics II**  
**Spring 2020**  
**Monday, 3:30 pm to 5:30 pm (BE300)**  
**Wednesday, 3:00 pm to 6:00 pm (BE300, 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.

**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.**

**Professors:**

**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.

**Dr. Elizabeth Walsh** is the Director of the Ecology and Evolutionary Biology program. Her work focuses on ecology and evolutionary biology of freshwater invertebrates, especially phylogenetics and cryptic speciation. She has been teaching Bioinformatics 1, and occasionally Bioinformatics 2, since the inception of the program.

**Dr. Sourav Roy** is a computational biologist with a research background in genetics, genomics, bioinformatics, and molecular biology. His research has involved the utilization of diverse in-silico, in-vitro, and in-vivo methodologies to understand the regulation of gene expression. He joined UTEP from St. Mary's University, where he served as an assistant

professor and bioinformatics specialist at the School of Science, Engineering, and Technology, and played a key role in the development of the new Bachelor of Science in Bioinformatics program. Dr. Roy received his doctoral degree in genetics, genomics, and bioinformatics from the University of California, Riverside (UCR), in 2011. He joined the Department of Entomology at UCR after his Ph.D. and worked as a post-doctoral scholar and as an assistant professional researcher for seven years. His current research projects include regulatory genomic studies of vector mosquitoes; evolutionary and comparative genomic studies of molecular pathways in insect vectors; the role of miRNAs in various organisms; and investigation of oxidative stress induced cellular survival pathways in colorectal cancer health disparities – all of which involve analysis and validation of high throughput and next-generation sequencing data using bioinformatics tools and molecular techniques. Dr. Roy continues to collaborate with researchers at UCR and Loma Linda University School of Medicine on projects related to vector biology, virology, prostate cancer, diabetes, and thyroid cancer. He has coauthored more than twenty-five research and review articles in journals that include Nature, PNAS, Annual Review of Entomology, PLoS Pathogens, PLoS Genetics, BMC Genomics, Insect Biochemistry and Molecular Biology, etc. His research contributions have resulted in two RO1 grants from NIH, an ESA STEP Travel Award from the 25th International Congress of Entomology (ICE), a Fellowship Award from the Physiology, Biochemistry and Toxicology section of the ESA. The Roy lab currently has openings for graduate and undergraduate students and a post-doctoral researcher.

#### **Teaching Assistants:**

The Teaching Assistant for this semester is Abhijeet Patil, a Ph.D. student in Computational Science. They will be responsible for the computer laboratories and homework assignments. Their office hours will be posted separately from this syllabus in BE 302. They can also be reached by email at [arpatil@miners.utep.edu](mailto:arpatil@miners.utep.edu).

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).

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

Monday is nominally the Lecture portion of the class, running from 3:30 pm to 5:30 pm. Wednesday is usually a laboratory meeting, running from 3 pm to 6 pm. For administrative reasons, these classes are listed as held in Classroom Building C201 and C302, however they will normally be held in BE 300 or in specialized laboratories for the various “wet” labs on the schedule. The lecture/laboratory designations of a class, however, can vary depending on the schedules of the professors and guest lecturer/laboratory director, and in some circumstances we may trade class and laboratory times. 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, we 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. In our experience, when using software such as python, R, or mysql, we have found the most useful reference to be Google...

#### **You've mentioned “Blackboard course site” three times...**

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 and notes for lectures. 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 the instructor or TA earlier rather than later!!! If things get complicated, we 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 publicly 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. Dr. Aley is sufficiently old fashioned that he still insists on using perl, but don't panic. All other instructors and TAs use Python and R. 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 your personal computer 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 will include Class Presentations, Homework, a stand-alone White Paper, Review-Exam, Final-Exam, and Laboratory Assignments.

### **Grade Point Distribution:**

#### **Class Participation 10%**

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. This is particularly true when primary literature is assigned. Grading for this component is subjective and will be recorded in BlackBoard in a timely fashion.

#### **Homework Assignments 20%**

Homework assignments are weekly writing, research, or programming assignments to be done outside 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 Feb 6, a draft version of the white paper, due February 15. The final draft of the document is due February 24.

### **First Individual (Review) 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 Biological and informatic concepts from your earlier studies. The exam will be tightly timed on BlackBoard, the week of Feb 6, with the exact time to be determined by the instructor in consultation with class members.

### **Final Examination 25%**

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 11, 1 – 4:30 pm).

### **Computer and Wet Lab Exercises and Follow-up 30%.**

The Wednesday afternoon class will usually be the time for laboratory exercises. These exercises will be under the guidance of the laboratory teaching assistant, Mr. Patil, and the laboratories used will vary as needed. Computer laboratory exercises will be held in <BE 300>. Because of the computational time needed for many of the computer exercises, it is expected that you may need to complete the computer laboratories and write-ups outside of the scheduled lab hours. “Wet” laboratory locations will be scheduled as needed.

---

## **COURSE POLICIES:**

### **Class participation:**

You are expected to come to class prepared to answer questions about the assigned readings or other materials.

### **E-mail and Blackboard:**

You are required to provide the teaching team with a UTEP e-mail address and check your UTEP e-mail and Blackboard daily. We will use your UTEP e-mail to contact you and you will use your UTEP e-mail to contact us as well. *Do not use Blackboard for e-mail.*

### **Dropping the course:**

Students may drop the class and receive a W (withdrawal) on their transcript prior to April 3rd, 2020. You must consult the Instructor prior to dropping. Due to the University’s six-drop rule, dropping the course may not be in your best interest. After April 3rd, a drop will result in an F on your transcript. Receiving either a W or an F in any course may prevent you from meeting the satisfactory Academic Progress requirements necessary to receive financial aid.

### **Student Conduct:**

Each student is responsible for notice of and compliance with the provisions of the Regents’ [Rules and Regulations](http://www.utsystem.edu/bor/rules/homepage.htm), available at <http://www.utsystem.edu/bor/rules/homepage.htm>. All students are expected to behave as courteous, responsible adults. We will have frequent discussions and students are expected to tolerate and respect the opinions of others.

### **Cellular and electronic devices:**

Cell phones and other electronic and recording devices must be turned off during class time to minimize classroom disruptions and protect the integrity of test-taking situations. This means you cannot make calls, send text messages, or use social media during class. You may use your laptop or tablet to take notes in class, but this privilege will be revoked if the devices are used inappropriately.

Students who fail to follow this rule may incur disciplinary action up to and including dismissal from the class and upon repeated offenses, the course.

### **Final Exam:**

[From the online Schedule of Classes] “Exemption from final examination may not be given. Final examinations are scheduled to be two hours, forty-five minutes in length and take place during

the final examination period. It is the policy of the university not to administer a second final examination in the course. It is also university policy that students shall not have more than two final examinations in a single day. In the unlikely event that the examination schedule results in a student having three final examinations on a single day, the faculty member upon the request of the student shall reschedule the second of that student's three examinations."

**Academic Integrity:**

All graded assignments must be entirely the work of the *individual* student. "*Plagiarism*" means the appropriation, buying, receiving as a gift, or obtaining by any means another's work and the unacknowledged submission or incorporation of it in one's own academic work offered for credit, or using work in a paper or assignment for which the student had received credit in another course without direct permission of all involved instructors. (from the Regents' [Rules and Regulations](#)) Plagiarism is a serious violation of university policy and will not be tolerated. All cases of suspected plagiarism will be reported to the Dean of Students for further review.

**Disability accommodations:**

If you have or suspect you might have a disability and need an accommodation you should contact the Center for Accommodations and Support Services (CASS) at 747-5148 or at [cass@utep.edu](mailto:cass@utep.edu) or go to Room 106 Union East Building. Students are responsible for providing any CASS accommodation letters and instructions.

**Campus Carry:**

Persons holding a Concealed Handgun License can lawfully carry their handgun into a UTEP classroom as long as the gun remains concealed. Open carry remains prohibited on campus. In other words, none of us should see (or be able to tell that there is) a gun at UTEP. Call the University Police at 747-5611 or dial 911 if you see any individual on campus with a handgun or other type of weapon. For more information on campus carry, see [<http://sa.utep.edu/campuscarry/>]; for more information on overall campus safety, see [<http://admin.utep.edu/emergency>].

**Military Statement:**

If you are a military student with the potential of being called to military service and /or training during the course of the semester, you are encouraged to contact us no later than January 27.

Date	Day	Lect/Lab	Topic	Instructor
1/22/20	W	Lect 01	Introductions, Syllabus, Biochemistry Review	SBA
1/27/20	M	Lect 02	Sequencing, Assembly, & Finishing	SBA
1/29/20	W	Lab 02	Wet Lab: DNA Sequencing	EW
2/3/20	M	Lect 03	Environmental Sequencing	EW
2/5/20	W	Lab 03	Wet Lab: Next Gen Sequencing	EW
2/10/20	M	Lect 04	Gene Calling: de novo & empirical	SBA
2/12/20	W	Lab 04	Calling Genes ab initio	TA-AP
2/17/20	M	Lect 05	Genome Comparison I	EW
2/19/20	W	Lab 05	MUMMER, TaxPlot, exon comparisons	TA-AP
2/24/20	M	Lect 06	Annotation I	SR
2/26/20	W	Lab 06	Empirical calling of genes & Annotation I	TA-SR
3/2/20	M	Lect 07	Annotation II	SBA
3/4/20	W	Lab 07	Annotation II	TA-SBA
3/9/20	M	Lect 08	SNPs	SR
3/11/20	W	Lab 08	SNP comparison	TA-SR
3/16/20	M	Springbrk	no class	
3/18/20	W	Springbrk	no class	
3/23/20	M	Lect 09	RNA Hybridization and MicroArray I	SBA
3/25/20	W	Lab 09	Microarray Data download and Clustering (in R)	TA-AP
3/30/20	M	Lect 10	Analysis: MicroArray II	SR
4/1/20	W	Lab 10	Microarray, statistical comparisons (in R)	TA-SR
4/6/20	M	Lect 11	Regulatory genomics and motif discovery	SR
4/8/20	W	Lab 11	Motif discovery Lab	TA-SR
4/13/20	M	Lect 12	Epigenomics	SBA
4/15/20	W	Lab 12	Epigenomics Lab	TA-AP
4/20/20	M	Lect 13	Proteomics - MS/MS	Guest
4/22/20	W	Lab 13	Wet Lab: MS/MS laboratory	Guest
4/27/20	M	Lect 14	Protein folding and ligand binding	Guest
4/29/20	W	Lab 14	Protein folding lab	Guest
5/4/20	M	Lect 15	RNA folding, Protein folding	EW
5/6/20	W	Lab 15	RNA folding, non-coding gene detection, prokary vs eukary	TA-EW
5/11/20	M	Final Exam	1 pm, BE 300	TA-AP