Syllabus for Introduction to Bioinformatics

BIOL 4395 Lecture: Monday and Wednesday: 4:30 pm - 5:50 pm
Liberal Arts Building – Room # 221

Course Description:
The objective of this course is to teach how computational techniques can help with solving biological problems. Students will learn to efficiently use multiple genomics and bioinformatics tools, that are freely available, for the analysis of DNA, RNA and protein sequences and structure. No programming skills are necessary for this course. This interdisciplinary course would be helpful for students in the department of Biology, Computer Science, Mathematics and Chemistry who aspire to go to either graduate school or medical school, or plan to work in the Bioinformatics industry that has experienced exponential growth within the last decade.

Course Goals:
With successful completion of this course, students will:

1. Gain knowledge about the current genome sequencing techniques.
2. Understand how bioinformatics is related to biology and medicine.
3. Gain knowledge about DNA, RNA and protein sequences and structures.
4. Learn to gather and use information UCSC and ENSEMBL genome browser.
5. Be able to perform pairwise and multiple sequence alignments.
6. Gain knowledge on searching multiple sequence databases.
7. Be able to perform phylogenetic analysis.
8. Be able to predict genes.
9. Be able to analyze RNA and protein structures.
10. Learn how to design and analyze primers.

Instructor:

Dr. Sourav Roy
BRB 2.172(0ffice)
sroy1@utep.edu
Office hours: By appointment

Dr. Elizabeth J. Walsh
BIOL 218 (0ffice)
ewalsh@utep.edu
Office hours: By appointment

Teaching Assistant
Abhijeet Patil
arpatil@miners.utep.edu

Course Resources:

The Required **Text Book** for this course is **Essentials of Bioinformatics by Jin Xiong.** The course will be coordinated through a Blackboard course connection (access via **my.utep.edu**). Blackboard will provide an online syllabus, course calendar, course bulletin board, and some supplemental web sites and notes for lectures. Grades will also be presented through Blackboard.

Grades

This course will employ a variety of teaching methods according to the type and content of material presented. The class is composed of quizzes, lab assignments, exams and a group project for this course. Applicable core concepts taught previously will be assessed throughout the length of this course. Note that the points for a given graded event (assignment, quizzes, exams, etc.) will be weighted and adjusted per the percentage allocations as noted in the table below.

Quizzes: Most lectures will have an open-book, open-notes quiz at the end of the lecture.

In-Class Exams: Two in-class exams (1 hr each) will be composed of a combination of multiple-choice, true or false, matching, and/or short answer questions. One lecture exam (Mid-term) will be taken during regularly scheduled class time. The final exam will be given during the time assigned by the Registrar's Office. Mid-term will have questions from materials covered in lecture and lab until the mid-term. Final exam would consist of questions from material covered after the mid-term.

Lab Assignments: Students will be given assignments for each lab section, they will use the bioinformatics tools (they will learn how to use these in the lectures) to complete the assignments. Their participation in these sections, as well as their demonstrated understanding of the readings' content and significance, will be assessed through their performance in the lab sections.

Assessments	Points	Percentage of Grade
Quizzes	50	20
Class Participation	50	5
Homework Assignments	50	25
Midterm Exam	25	25
Final Exam	25	25
Total	150	100

COURSE POLICIES

POLICY ON CLASS PARTICIPATION: You are expected to come to class prepared to answer questions about the assigned lab. Always bring a pen and paper. Pop quizzes may be given at any time during the lab period. The instructor will post grades electronically, but students are responsible for knowing their grades at all times.

POLICY ON CELL PHONES: Do NOT have them on or out in class….this includes texting! Cell phones can be confiscated for the class period if used in lab.

POLICY ON ALL OTHER ELECTRONIC DEVICES: You cannot surf the internet, watch movies, listen to

music, etc. in lab. You will be asked to leave if this happens.

POLICY ON 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].

POLICY ON MAKE-UP QUIZZES AND EXAMINATIONS: NO make-up quizzes or exams will be given for reasons other than illness (doctor's note required) or when a student is on official University business (documentation required). Make-ups must be scheduled within a week of when the quiz or test was given.

POLICY ON ACADEMIC INTEGRITY: Academic dishonesty is prohibited and is considered a violation of the UTEP Handbook of Operating Procedures. It includes, but is not limited to, cheating, plagiarism, and collusion. Cheating may involve copying from or providing information to another student, possessing unauthorized materials during a test, or falsifying research data on laboratory reports. Plagiarism occurs when someone intentionally or knowingly represents the words or ideas of another person's as ones' own. Collusion involves collaborating with another person to commit any academically dishonest act. Any act of academic dishonesty attempted by a UTEP student is unacceptable and will not be tolerated.

While you will be working in lab teams, the work you submit for assessment must be evaluated on its own merit. Therefore, team members' reports and work should reflect the individual's thoughts. Do NOT turn in 3 near-duplicate reports with different names or everyone involved will be sent to the Dean of Students for possible disciplinary action. Students may be suspended or expelled from UTEP for such actions. Yes, we have had to deal with this problem in the past and we are not lenient. You can calculate the consequences. All university guidelines will be strictly followed. Please read these guidelines carefully. The guidelines can be found on line at:
http://admin.utep.edu/Default.aspx?PageContentID=2084&tabid=30292

POLICY ON DISRUPTIVE BEHAVIOR: Any student who disrupts the class will be asked to leave and will be referred to the Dean of Students.

DISABILITY STATEMENT: If a student has or suspects he/she has a disability and needs an accommodation, he/she should contact the Center for Accommodation and Support Services (CASS) at 747-5148 or at cass@utep.edu or go to Room 106 Union East Building. The student is responsible for presenting to the instructor any CASS accommodation letters and instructions.

MILITARY STATEMENT: If you are a military student with the potential of being called into military service and/or training during the course of the semester you are encouraged to contact the instructor regarding these matters. You must let us know during the first week of the summer.

*** Disclaimer: This syllabus may be changed as necessary at the discretion of the instructor.

Course Outline

Week of Jan 21. Basics of genomics and bioinformatics:
Session 1
Central dogma (DNA –> Transcription –> RNA –> Translation -> Protein)
Describe gene and genome.
What is genomics?
What is Functional genomics and Comparative genomics?
Describe proteome.
What is proteomics?
Describe homologs, orthologs and paralogs
What is bioinformatics?

Week of Jan 27: Introduction to biological databases:
Session 2:
What is a database?
Types of databases
Biological databases
Pitfalls of biological databases
Information retrieval from biological databases
Session 3:
What are sequence databases?
Different types of sequence formats – FASTA, GENBANK, EMBL etc.
Major sequence databases
How to retrieve sequences from different databases
Sequence matrices
How to search for unknown gene using BLAST
How to search for unknown protein sequences using InterProScan

Week of Feb 3. Human genome and browsers:
Session 4
Genome sequencing methods
Sequencing of the Human genome
Genome browsers
Session 5
How to navigate and use the UCSC genome browser
ENSEMBL browser

Week of Feb 10. Pairwise Sequence alignment:
Session 6
What is sequence Alignment?
Global alignment and local alignment
What information can be gathered from sequence alignment
Methods of pairwise sequence alignment
Session 7
Scoring matrices and gap penalties in sequence alignment
Pairwise alignment with Blast2seq

Week of Feb 17. Multiple Sequence alignment:
Session 8
Methods of multiple sequence alignment
Structural or evolutionary alignment?
How to align multiple sequences

Session 9
Global multiple sequence alignment with CLUSTAL Omega, T-COFFEE
Local multiple sequence alignment

Week of Feb 24. Sequence polymorphism: Dr. Walsh
Session 10
Overview of molecular evolution
Origins of polymorphisms
Types of polymorphisms
Session 11
SNP discovery methods
Genotyping
The international haplotype map project

Week of March 2. Review and Midterm
Session 12
Review
Session 13
Midterm

Week of March 9. Phylogenetic analysis: Dr. Walsh
Session 15
Relation between phylogenetic analysis and multiple sequence alignment
Methods for phylogenetic prediction
Genome complexity and phylogenetic analysis
Methods of phylogenetic prediction
Session 16
Phylogenetic trees: reliability and uses
Forms of tree representation
Phylogenetic programs
Building phylogenetic trees – Phylogeny.fr

Week of March 16: Spring Break

Week of March 23. Gene finding/prediction:
Session 17
Introduction
Prokaryotic and Eukaryotic gene prediction
Open reading frame and its reliability
Gene prediction methods
Gene prediction Programs
Session 18
Exon, intron, UTRs, promoters
Gene prediction with GeneMark, GENSCAN

Week of March 30: Promoter and regulatory element prediction:
Session 19
In-class/Take-home assignment
Session 20
Promoter and regulatory elements in Prokaryotes
Promoter and regulatory elements in Eukaryotes
Prediction Algorithms

Week of April 6. RNA sequences: prediction and analysis of structures: Dr. Walsh
Session 21

Basics of RNA secondary structure
Features of secondary and tertiary structure
RNA folding: Sequence and base pairing patterns
Free energy determination
Levels of prediction (pseudoknot free; nested pseudoknots; pseudoknotted with nested pseudoknots)
Session 22
Methods for predicting RNA secondary structure
RNA secondary structure prediction and analysis with MFold and IPKnot

Week of April 13. Principles of protein structure:
Session 23
Description of protein structure
Amino Acids
Peptide Formation
Dihedral Angles
Structure and sequence similarity
Analyzing an active site
Session 24
Protein structural visualization
Molecular viewers: POLYVIEW-2D and Swiss PDB Viewer
Protein structure comparison
Protein structure classification
Prediction of protein secondary structure from amino acid sequences

Week of April 20. Primer designing:
Session 25
Primers and their uses
Basics of primer and oligo designing
Optimal size
Melting Temperature
GC content
Session 26
Self-complementarity
Hairpin Structures
Mis-priming
Primer designing using Primer3 and OligoAnalyzer 3.1

Week of April 27. Genomics and Proteomics:
Session 27
Challenges for genome analysis
Comparative genomics
Functional classification of genes Genomes
Functional genomics approaches
Session 28
Technology of protein expression analysis
Posttranslational modification
Protein–Protein interactions

Week of May 4. Review and Final Exams:
Session 29.
Importance of rigor and reproducibility in scientific research. Dr. Aguilera (guest lecture)
Session 30.
Review for final exam.

Final Exam: May 11 – 4:30 PM – 6:00 PM