

# Biology 4425/5425 – Fall 2021

## Bioinformatics



---

### Instructor: Dr. Greg Krukonis

Email: [gkrukonis@angelo.edu](mailto:gkrukonis@angelo.edu)

Phone: (cell) 717-417-8717

Office: Cavness 003A

**Office Hours: By appointment, seven days a week, evenings included. All meetings will be via video.**

## Course Information

### Course Description

Lecture: Introduction to methods for acquiring, analyzing, and employing biological sequence information. Topics will include – Theory and process of PCR, mass spectroscopy, and DNA microarrays. Algorithms for searching and clustering sequences. Applications of bioinformatic data to questions such as the geographical movement of human pathogens, horizontal gene transfer in viruses, and changes in human gene expression in response to disease and treatment.

Lab: Students will access remote sequence databases (NCBI, EMBL-EBI) and analyze sequences with open-source bioinformatics software running natively, in a Linux virtual machine, and on remote servers. Analyses will include protein structure prediction, phylogenetics using molecular data, and genome annotation. This portion of the course offers you the opportunity to explore and apply concepts to answer research questions. Success in the laboratory involves teamwork in designing and conducting analyses and reporting your findings. In addition, you will conduct activities designed to develop and improve critical thinking and problem-solving skills related to the topics discussed in lectures. Students will annotate a novel viral genome and submit the completed annotation to NCBI.

### Prerequisite and Co-requisite Courses

Genetics

## Student Learning Outcomes

Upon completion of this course, students will understand/be able to:

- Understand the role of computation in hypothesis driven discovery processes within the life sciences.
- Explain the role of wet-bench techniques in verifying computational results in life science research.
- Compare and contrast computer-based research with wet-lab research.
- Read a scientific article and evaluate how bioinformatics methods were employed by the authors to explore a particular hypothesis.
- Given a scientific question, develop a hypothesis and propose computational approaches that could be used to explore the hypothesis.
- Evaluate the social, legal, and ethical implications of computational approaches to understanding biology.
- Define the term algorithm.
- Explain the difference between a heuristic (approximate) algorithm and an exact algorithm.
- What statistical concepts are important in bioinformatics.
- Perform elementary statistical analysis on an “omics” dataset (e.g. using Excel or Weka).
- Interpret the biological significance of an e-value.
- The types of bioinformatic data (e.g., nucleotide sequence, epigenomics) and how these data are stored and accessed.
- Describe how nucleotide sequence data are represented (FASTA, FASTQ, GenBank).
- Describe the nucleotide databases available at NCBI.
- Describe how protein sequence data are represented (e.g., FASTA, GenBank, etc.)
- Describe the different protein databases available at NCBI (sequence, structure, function).
- Describe how the NCBI nucleotide databases intersect with other nucleotide databases (EBI, DDBJ, UniProt, etc.).
- Search for a sequence record in a nucleotide database with a given accession number.
- Create a collection of nucleotide sequence records that meet a specified criterion (e.g., gene name or symbol).
- Determine the DNA methylation state of a particular region of a genome.

- Describe the types of metadata that accompany sequence data to make for useful biological interpretation (e.g. biological source, accession number, GeneID, journal articles, etc.).
- How bioinformatics tools can be employed to analyze genetic information.
- Calculate the alignment score between two DNA sequences using a provided scoring matrix.
- Perform a BLASTN search and interpret the results.
- Create and interpret a multiple sequence alignment (e.g., T-COFFEE, MUSCLE, etc.).
- For a genomic region of interest (e.g., the neighborhood of a particular gene), use a genome browser to view nearby genes, transcription factor binding regions, epigenetic information, etc.
- Compare and contrast data contained in different databases.
- Explain the BLASTP, BLASTX, tBLASTn, tBLASTx algorithms for protein sequence information
- How can bioinformatics tools be employed to examine protein structure and function?
- Query a dataset with a specific protein sequence to learn about potential functions (e.g. Pfam, CDD,
- SwissProt, UniProt, etc.).
- View and interpret the structure output from Protein Data Bank (e.g. Cn3D, Jmol, etc.).
- Propose potential functions for a give protein structure.
- Explain the outputs from protein-folding algorithms to predict structure from sequence.
- Create and interpret a multiple sequence alignment (e.g., T-COFFEE, MUSCLE, etc.).
- Describe the components of a phylogenetic tree (e.g., root, node, leaf).
- Explain the various types of phylogenetic trees (e.g., rooted, unrooted).
- Interpret a phylogenetic tree (e.g., which organism is most closely related to a given organism in the tree)
- Use bootstrapping to assess the quality of a phylogenetic tree.
- Create a phylogenetic tree for a set of related sequences (nucleotide or amino acid) (e.g., MEGA).
- Use a phylogenetic tree to address a research question.

## Course Delivery

This is a face-to-face course with online components that students are expected to access in Blackboard.<sup>1</sup> Changes in the severity of the coronavirus pandemic may

require modifications, for example a switch to a hybrid format should the University decide such changes are warranted.

## Required Texts and Materials

All readings will be from the primary literature

## Technology Requirements

Laboratory work will computer based. All software will be available in University teaching labs as well as being available, at no cost, for download to students personal machines.

## Communication

Faculty will respond to email and/or telephone messages within 24 hours during working hours Monday through Friday. Weekend messages may not be returned until Monday.

**Written communication via email:** All private communication will be done exclusively through your ASU email address. Check frequently for announcements and policy changes. In your emails to faculty, include the course name and section number in your subject line.

**Virtual communication:** Office hours and/or advising will be done with the assistance of the telephone, Collaborate, etc.

Office hours are whenever you need them to be. Days/Nights/Weekends all work. These are extraordinary times. Never hesitate to reach out to arrange a collaborate session.

## Grading

### Evaluation and Grades

Course grades will be determined as indicated in the table below.

Assessment	Percent of Total Grade	
Exam 1	100pts	10%
Exam 2	150pts	15%
Exam 3	150pts	15%
Short Format lecture/lab assignments	150pts	15%
Bacteriophage Annotation	300pts	30%
Oral presentations	150pts	15%
Total	1000pts	

## Grading System

Course grades will depend on completing course requirements and meeting the student learning outcomes.

This course uses the following grading scale:

A = 90.00-100 points

B = 80.00-89.99 points

C = 70.00-79.99 points

D = 60.00-69.99 points

F = 0-59.99 points (Grades are not rounded up)

## Assignment and Activity Descriptions

Posted on Blackboard.

## General Policies Related to This Course

All students are required to follow the policies and procedures presented in these documents:

- [Angelo State University Student Handbook](#)<sup>2</sup>
- [Angelo State University Catalog](#)<sup>3</sup>

## Academic Integrity

Students are expected to maintain complete honesty and integrity in all work. Any student found guilty of any form of dishonesty in academic work is subject of disciplinary action and possible expulsion from ASU.

The College of Science and Engineering adheres to the university's [Statement of Academic Integrity](#).<sup>4</sup>

## Accommodations for Students with Disabilities

ASU is committed to the principle that no qualified individual with a disability shall, on the basis of disability, be excluded from participation in or be denied the benefits of the services, programs or activities of the university, or be subjected to discrimination by the university, as provided by the Americans with Disabilities Act of 1990 (ADA), the Americans with Disabilities Act Amendments of 2008 (ADAAA) and subsequent legislation.

Student Disability Services is located in the Office of Student Affairs, and is the designated campus department charged with the responsibility of reviewing and authorizing requests for reasonable accommodations based on a disability. It is the

student's responsibility to initiate such a request by contacting an employee of the Office of Student Affairs, in the Houston Harte University Center, Room 112, or contacting the department via email at [ADA@angelo.edu](mailto:ADA@angelo.edu). For more information about the application process and requirements, visit the [Student Disability Services website](#).<sup>5</sup> The employee charged with the responsibility of reviewing and authorizing accommodation requests is:

Dr. Dallas Swafford  
Director of Student Disability Services  
Office of Student Affairs  
325-942-2047  
[dallas.swafford@angelo.edu](mailto:dallas.swafford@angelo.edu)  
Houston Harte University Center, Room 112

### **Incomplete Grade Policy**

It is policy that incomplete grades be reserved for student illness or personal misfortune. Please contact faculty if you have serious illness or a personal misfortune that would keep you from completing course work. Documentation may be required. See ASU Operating Policy 10.11 [Grading Procedures](#)<sup>6</sup> for more information.

### **Plagiarism**

Plagiarism is a serious topic covered in ASU's [Academic Integrity policy](#)<sup>7</sup> in the Student Handbook. Plagiarism is the action or practice of taking someone else's work, idea, etc., and passing it off as one's own. Plagiarism is literary theft.

In your discussions and/or your papers, it is unacceptable to copy word-for-word without quotation marks and the source of the quotation. It is expected that you will summarize or paraphrase ideas giving appropriate credit to the source both in the body of your paper and the reference list.

Papers are subject to be evaluated for originality. Resources to help you understand this policy better are available at the [ASU Writing Center](#).<sup>8</sup>

### **Student Absence for Observance of Religious Holy Days**

A student who intends to observe a religious holy day should make that intention known in writing to the instructor prior to the absence. See ASU Operating Policy 10.19 [Student Absence for Observance of Religious Holy Day](#)<sup>9</sup> for more information.

## **Title IX at Angelo State University**

Angelo State University is committed to providing and strengthening an educational, working, and living environment where students, faculty, staff, and visitors are free from sex discrimination of any kind. In accordance with Title VII, Title IX, the Violence Against Women Act (VAWA), the Campus Sexual Violence Elimination Act (SaVE), and other federal and state laws, the University prohibits discrimination based on sex, which includes pregnancy, and other types of Sexual Misconduct. Sexual Misconduct is a broad term encompassing all forms of gender-based harassment or discrimination and unwelcome behavior of a sexual nature. The term includes sexual harassment, nonconsensual sexual contact, nonconsensual sexual intercourse, sexual assault, sexual exploitation, stalking, public indecency, interpersonal violence (domestic violence or dating violence), sexual violence, and any other misconduct based on sex.

You are encouraged to report any incidents involving sexual misconduct to the Office of Title IX Compliance and the Director of Title IX Compliance/Title IX Coordinator, Michelle Miller, J.D. You may submit reports in the following manner:

Online: [Incident Reporting Form](#)<sup>10</sup>

Face to Face: Mayer Administration Building, Room 210

Phone: 325-942-2022

Email: [michelle.miller@angelo.edu](mailto:michelle.miller@angelo.edu)

Note, as a faculty member at Angelo State, I am a mandatory reporter and must report incidents involving sexual misconduct to the Title IX Coordinator. Should you wish to speak to someone in confidence about an issue, you may contact the University Counseling Center (325-942-2371), the 24-Hour Crisis Helpline (325-486-6345), or the University Health Clinic (325-942-2171).

For more information about resources related to sexual misconduct, Title IX, or Angelo State's policy please visit the [Title IX website](#).<sup>11</sup>

## **Information About COVID-19**

Please refer to ASU's [COVID-19 \(Coronavirus\) Updates](#)<sup>12</sup> web page for current information about campus guidelines and safety standards as they relate to the COVID-19 pandemic.

## **Modifications to the Syllabus**

This syllabus, including grade evaluation and course schedule, is subject to modification on potentially short notice based on developing circumstances.

## Course Schedule

### Week

- Aug 23 Introduction to bioinformatics. Jagger et al. An Overlapping Protein-Coding Region in Influenza A Virus Segment 3 Modulates the Host Response. No Lab
- Aug 30 Continue Jagger. Intro to Blast, DNA MASTER, Genome analysis, and annotation of bacteriophage
- Sept 6 Continue Jagger. Intro to MegaX, Phylogenetic trees, DNA sequencing techniques
- Sept 13 Continue Jagger. Next Gen sequencing techniques. Sequence alignment methodologies. HHPRED
- Sept 19 **Exam 1 WED Sept 22.** Martínez et al. Synonymous Virus Genome Recoding as a Tool to Impact Viral Fitness. Installation of Linux virtual machine, HHMI software package, Phamerator.
- Sept 26 Epigenetics and DNA methylation. Enard et al. Evidence that RNA Viruses Drove Adaptive Introgression between Neanderthals and Modern Humans. Protein structure prediction with PHYRE THMM
- Oct 4 ChIP-chip, and ChIP-seq. Origin and evolution of the SARS-CoV2. Identifying horizontal gene transfer with Dot plot analysis.
- Oct 11 Bioinformatics of vaccine design. Analysis of reassortment events in Influenza A. Building gene compliment-based trees with Splitstree.
- Oct 18 **Exam 2 Oct 20.** Spatial-omics and Cancer.
- Oct 25 Mass Spectroscopy theory and application. Oral presentations. Continue Bacteriophage Annotation.
- Nov 1 CRISPR-Cas bioinformatics. Oral presentations. Continue Bacteriophage Annotation.
- Nov 8 Horizontal Gene Transfer: From Evolutionary Flexibility to Disease Progression. Oral presentations. Continue Bacteriophage Annotation.
- Nov 15 **Exam 3 Nov 17.** Oral presentations. Continue Bacteriophage Annotation
- Nov 29 Final preparation and submission of Bacteriophage Genome to NCBI. Oral presentations

- 
- 1 <https://blackboard.angelo.edu/>
  - 2 <https://www.angelo.edu/current-students/student-handbook/>
  - 3 <https://www.angelo.edu/academics/catalog/>
  - 4 <https://www.angelo.edu/live/files/27603-student-handbook-2020-21#page=96>
  - 5 <https://www.angelo.edu/current-students/disability-services/>
  - 6 <https://www.angelo.edu/content/files/14197-op-1011-grading-procedures>
  - 7 <https://www.angelo.edu/live/files/27603-student-handbook-2020-21#page=96>
  - 8 [https://www.angelo.edu/current-students/writing-center/academic\\_honesty.php](https://www.angelo.edu/current-students/writing-center/academic_honesty.php)
  - 9 <https://www.angelo.edu/content/files/14206-op-1019-student-absence-for-observance->  
[of](#)
  - 10 <https://www.angelo.edu/incident-form>
  - 11 <https://www.angelo.edu/title-ix>
  - 12 <https://www.angelo.edu/covid-19/>