BI 694 Special Studies - Bioinformatics & Phylogenetics

Lecture Location: ML Classroom
Meeting time: Mo 1PM – 3PM & Thurs 10AM – 12PM
Office Hours: by appointment

Professor: Bastian Bentlage
Office/Lab: Marine Lab Rm 106

Overview
Bioinformatics and Phylogenetics is a graduate level course that will provide graduate students with hands-on experience in computational approaches to extracting information from molecular data and build their skillset to pursue their own research at UOG and their future careers.

Catalog course description

BI 694 SPECIAL STUDIES - BIOINFORMATICS & PHYLOGENETICS (4) FALL
This course provides an introduction to computational approaches to analyze molecular genetic data. The lecture will focus on the theoretical foundations of how genomes are sequenced and reconstructed, how genes and their functions are inferred, and how the evolution of organisms at the levels ranging from populations to species can be reconstructed using molecular data. Hands-on exercises will introduce students to introduce students to the usage of commonly-used bioinformatics tools. The UNIX/Linux operating systems will be used to automate repetitive tasks and scripts will be written in the computer programming language Python. The objective of this course is to gain the skills and conceptual understanding necessary to analyze molecular data with an emphasis on laying the foundation to facilitate extracting information from big datasets in the era of genomics.

Intended Course Learning Outcomes
In this course students will gain an understanding of how computers and software can be used to extract information from molecular genetic data. By the end of the course students will be able to interpret, judge and comprehend the results of modern genomics research and evolutionary biology. Students will demonstrate their understanding of the materials taught by applying them to analyses of realistic biological data. Students will create their own solutions to problems or modify existing approaches to data analysis.

The intended outcomes of this course will address the following Program Learning Outcomes set forth in the 2016/2017 Graduate Bulletin:

PROGRAM LEARNING OBJECTIVES
1) Demonstrate ability to analyze data and design experiments using standard statistical procedures
2) Demonstrate knowledge of basic cellular and molecular-level principles
3) Demonstrate knowledge of the latest advances in a variety of fields in biology
4) Demonstrate the ability to disseminate scientific concepts and research findings in a variety of formats

ASSESSMENT METHODS
Written exams: PLOs 2,3
Problem sets: PLOs 1,4
Hands-on in-class exercises: PLOs 1, 2, 4

KNOWLEDGE-BASED OUTCOMES
1) Students identify and apply appropriate strategies to extract biological information from genomic/molecular genetic datasets
2) Students understand the hierarchy of analytical approaches to investigate the structure, function, evolution of genes and genomes

SKILL-BASED OUTCOMES
1) Students gain experience with the practical aspects of handling and analyzing molecular genetic data using the Linux operating system, the shell, and the Python programming language
2) Students complete exercises pertaining to the major course themes
3) Students develop problem-solving strategies to solve take-home problem-sets

Course
Lectures and demonstrations of exercises in the Linux shell will provide the theoretical background and applied methods to introduce students to bioinformatic analyses. Textbooks and publications will serve as reading materials. In addition to learning the theoretical foundations of bioinformatic and phylogenetic analyses, students will follow demonstrations by the instructor, or work individually or in small groups on problem sets during class under guidance of the instructor. Students will analyze realistic datasets of sequence data as take-home problem-sets.

Attendance Policy:

I strive to create a course atmosphere that will be attractive enough for you wanting to attend every session. I do realize that life events may at times prevent you from attending class. Attendance is generally expected and absences should have an adequate excuse.

Texts and Required Reading:

Papers and book chapters will be provided but the following text book is recommended. The course will use the linux operating system (Debian or Ubuntu) and the Python programming language, but no prior knowledge of either is required.

Every student is expected to bring a laptop computer to class. A limited number of laptop computers may be provided but this requires prior consultation with the instructor.

The following book is recommended for the course

Haddock SHD & Dunn CW (2011) Practical computing for biologists. Sinauer, Sunderland MA

Prerequisites
BI315/BI315L General Genetics or instructor consent

Special needs
Students with special needs must make arrangements through the ADA office. The University makes every attempt to accommodate such requests (see below). Students who cannot meet the requirements of a particular field trip must discuss the problem with me several days in advance.

UOG Disabilities Policy
In accordance with the Americans with Disabilities Act (ADA) of 1990 and the Rehabilitation Act of 1973, the University of Guam does not discriminate against students and applicants on the basis of disability in the administration of its educational and other programs. The University offers reasonable accommodations for a student or applicant who is otherwise qualified, if the accommodation is reasonable, effective and will not alter a fundamental aspect of the University's program nor will otherwise impose an undue hardship on the University, and/or there are not equivalent alternatives. Students are expected to make timely requests for accommodation, using the procedure below*. If appropriate, the University may choose to consult with such individuals, at or outside the University, to provide expertise needed to evaluate the request for accommodation. Each student bears the responsibility for initiating and then documenting a disability-related request for accommodation in the manner requested in this Policy. Full text at: www.uog.edu/dynamicdata/EqualEmploymentOffice.aspx?siteid=1&p=66

UoG Tobacco-Free Policy
Pursuant to Board of Regents Resolution No. 13-24, the University of Guam (UOG) has a total ban on the sales, smoking and the distribution and use of tobacco and tobacco-based products on the UOG Campus, and properties.
The purpose of this policy is to protect the public health and welfare by prohibiting smoking and the use of tobacco products or simulated smoking devices, including but not limited to E-cigarettes, on the UOG campus and properties; to guarantee the right of nonsmokers to breathe smoke-free air, while recognizing that the need to breathe smoke-free air shall have priority over the desire to smoke; and to encourage a healthier, more productive living/learning environment for all members of our University community.

FERPA
UOG is bound by the policies of the Family Educational Rights & Privacy Act. This means your educational records and personal information are protected by law. Please consult the FERPA web site for more information:
http://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html

Evaluation Guidelines
Learning outcomes will be assessed in written examinations and during hands-on exercises and class discussions.

Attendance and participation (10%), problem sets (10 total) (40% plus 10% extra credit), midterm exam (20%), final exam (30%).

Final grades are assigned based on the percentage of total possible points received:
(A) 93-100%, (A-) 90-92%, (B+) 88-89%, (B) 83-87%, (B-) 80-82%, (C+) 78-79%, (C) 73-77%, (C-) 70-72%, (D+) 68-89%, (D) 63-67%, (D-) 60-62%, (F) <60%

Staffing and Resources
This course will be covered by Dr. Bastian Bentlage, a faculty member at the University of Guam Marine Laboratory. Students are expected to use their own personal laptop computers for hands-on exercises. A limited number of laptop computers may be provided by the instructor. All software used is open source and available free of charge. If necessary students will have to install a Linux/GNU operating system on their computers in a virtual machine. Guidance for this will be provided by the instructor, Dr. Bentlage. Computationally demanding problem sets will be analyzed on one of the servers housed in Rm 106 at the Marine Lab and administered by the course instructor. Students will get user accounts on these servers for the duration of the course and may access them remotely using secure shell through the campus network. This is already set up.

Course Schedule (DRAFT – subject to change as necessary)
Part I. Introduction (week 1)
1) Course outline and installation of Linux on laptops using virtual machines
Part II. Introduction to Linux and the shell (weeks 2-3)
1) Command-line operations and handling text files in the shell
2) Pattern searches using regular expressions and simple pipelines
3) basics of shell scripting
4) Programming concepts (flow control with loops and conditionals)
5) Genetic sequence, multiple sequence alignment, and phylogenetic tree file formats
6) Installing software (compiling from source and package managers)
Part III. Programming in Python (weeks 4-5)
1) Why use Python?
2) Reading and writing files
2) Variables, arrays, lists, and dictionaries (hash tables)
3) Mathematical and logical operators
4) Modules and libraries
5) Debugging scripts
MIDTERM EXAMINATION (week 6)
Part IV. Theory and practice of bioinformatic analyses (week 7-14)
1) Working on remote servers
2) Sequence similarity searches using BLAST and hidden markov models
3) Public sequence databases
4) Finding related sequences using clustering
5) multiple sequence alignments
6) inferring phylogenetic trees using sequence data
7) identifying genetic variation in populations using single nucleotide polymorphisms (SNPs)
8) Genome (and transcriptome) assembly
9) Relational databases
10) Version control and backing up your work
FINAL EXAM (week 16)