BINF 2111/2111L, INTRODUCTION TO BIOINFORMATICS COMPUTING
SYLLABUS

A. COURSE DESCRIPTION
This course introduces students to the UNIX environment, bash shell scripting, and python scripting. Fundamentals of each scripting language are introduced in the context of real-world bioinformatics data analysis problems.

B. PRE- OR CO-REQUISITES
Pre- or corequisite: BINF 1101/1101L.

C. OBJECTIVES OF THE COURSE
Having successfully completed this course, the student will be able to:
• Understand the UNIX command line environment.
• Use built-in UNIX commands to manipulate files and data.
• Use bash shell scripts to drive pipelines of bioinformatics programs.
• Use python scripts to read, manipulate and write bioinformatics data files.

D. INSTRUCTIONAL METHOD
The course is presented in a workshop format which will include the following elements as appropriate: presentation of factual material in a lecture format, interactive demonstrations of methods to be applied in assignments, and hands on coding exercises with the opportunity to ask questions and collaborate with other students.

E. MEANS OF STUDENT EVALUATION
Students will primarily be evaluated on their ability to successfully write assigned bash and python scripts. Understanding of bioinformatics workflows and programming core concepts presented in class will be assessed through short answer quizzes. Assignments and quizzes will be given weekly throughout the semester. The final examination will consist of both short answer questions and a coding exercise.

Grading: 25% quizzes and exams; 50% labs and assignments; 25% participation.

F. SPECIFY POLICIES THAT APPLY TO THIS COURSE:
The following policies apply to students in BINF 2111/2111L:

1. UNIVERSITY INTEGRITY
All students are required to read and abide by the Code of Student Academic Integrity. Violations of the Code of Student Academic Integrity, including plagiarism, will result in disciplinary action as provided in the Code. Definitions and examples of plagiarism are set forth in the Code. The Code is available from the Dean of Students Office or online at: http://www.legal.uncc.edu/policies/ps-105.html. A set of links to various resources on plagiarism and how to avoid it is available at the UNCC Library website: http://library.uncc.edu/display/?dept=instruction&format=open&page=920.
2. **Attendance**
   Attendance at lecture and lab sessions is mandatory, although exceptions can be made for documented reasons such as illness or family emergency. More than one absence will immediately impact the participation portion of the grade.

3. **Grading Policy**
   Grades will be assigned on the following scale:
   - A = 90-100%
   - B = 80-90%
   - C = 65-80%
   - U = 0-65%

4. **Additional Policies.**
   The use of cell phones, beepers, or other communication devices is disruptive, and is therefore prohibited during class. Except in emergencies, those using such devices must leave the classroom for the remainder of the class period. Students are permitted to use computers during class for note-taking and other class-related work only.

G. **Textbook**

   Textbooks:
   * Practical Computing for Biologists, Haddock and Dunn.

H. **Topical Outline of Course Content**

   - Week 1: Introduction to UNIX and biology review
   - Week 2: Introduction to genome assembly and command line software
   - Week 3: Bash shell scripting basics and the FASTX/SPAdes/QUAST pipeline
   - Week 4: Bash scripting and UNIX file manipulation
   - Week 5: Bash variables and control structures
   - Week 6: Introduction to variant calling and personal genomics
   - Week 7: Bash scripting and the bowtie2/samtools/picard/vcftools pipeline
   - Week 8: Python variables and syntax
   - Week 9: Python conditionals and comparators
   - Week 10: Python loops, lists, and basic file methods
   - Week 11: Python functions and dictionaries
   - Week 12: Python modules and libraries – introduction to BioPython
   - Week 13: Parsing and merging files
   - Week 14+: Manipulating personal genomics data with Python