NORTHERN ARIZONA UNIVERSITY
College of Engineering, Forestry & Natural Sciences
Department of Electrical Engineering & Computer Science
Department of Biological Sciences

Syllabus for Bio599
Computational Biology

General Information

- Fall 2012
- Tu/Th 8:00-9:15 (3 credit hours) in Engineering (building 69), Room 224
- Professor: Greg Caporaso (gregcaporaso@gmail.com)
  - Office location: Engineering (building 69) room 265
  - Office hours: Tu/Th 9:15am-10:15am.
- TA: Andrew Krohn (alk224@nau.edu)
  - Office location: SLF 307
  - Office hours: M/W 11am-12:00pm
- Course website: http://bit.ly/bio599-web
- Course schedule: http://bit.ly/bio599-sch [The schedule is dynamic - check it regularly!]

Course Prerequisites

Instructor consent.
This course assumes that you have some knowledge of basic molecular biology, as well as basic techniques in molecular biology (things like PCR, DNA sequencing, gel electrophoresis). You should also have a basic familiarity with using bioinformatics tools such as BLAST, muscle, FastTree, uclust, and QIIME.

Course Description

This course will focus on many of the tools that are necessary to initiate, implement, document, and complete bioinformatics projects. The course will be project-oriented, with small in-class projects, and larger take-home assignments. These will involve interacting with biological databases, running bioinformatics analyses, interpreting results, and presenting those results both in writing and through oral presentations; developing scripts in python; and running parallel jobs on the Amazon Web Services cloud. We will cover computing topics such as interacting with command line tools, developing and using regular expressions, interacting with remote Linux systems (via the Amazon cloud), and python programming. We will also explore specific topics in bioinformatics, including high-throughput microbial ecology and metagenomics, genome assembly, and topics of interest chosen by the class.

Student Learning Expectations/Outcomes for this Course

After taking this course, students should be comfortable with developing, documenting and testing basic python software and working on remote Linux systems in a cloud-computing environment. Students will have experience with software tools that are used in real-world bioinformatics, and will understand how to initiate, perform, carry out and present a bioinformatics experiment.

Course Structure/Approach

This course will be focused on lectures, projects, student presentations of those projects, paper discussions and short quizzes. Regular in-class time will be scheduled for hands-on work, such as interacting with the IPython interpreter to accomplish computational tasks, identifying the source of erroneous sequence reads from "next-generation" DNA sequencers, and reconstructing phylogeny from sequence data. Homework assignments and reading should be completed by the due date provided on the course schedule.

There will be regular quizzes throughout the semester. These will be announced one-day prior to the quiz in order to ensure that you are keeping up on the material, but also give you a small amount of preparation time. These will cover material that was recently presented, and should not be difficult if you are attending class and doing the readings. Quizzing on the homework and reading assignments due the same day is fair game.

This course covers a lot of diverse material, and office hours are your chance to directly interact with the instructor and TA to ensure that you understand the concepts being covered in class. Coming to office hours is also a great way to boost your class participation score.

Course projects will be focused on using the PyCogent and QIIME bioinformatics software packages. These are widely used open-source bioinformatics packages. Contributions that you make to these software packages, in the form of new documentation or new features, can be applied as extra credit in the course. If you're interested in making a contribution, talk with me ahead of time and we will discuss whether it is a contribution worthy of extra credit, and if so how much it will be worth. Additionally, submitting bug reports or contributing to the QIIME forum (in the form of correctly answering questions) will count toward increasing your participation grade. If you submit bug reports or respond on the forum, forward that information to me by email.

Required Materials
Practical Computing for Biologists by Haddock and Dunn (ISBN: 978-0-87893-391-4)

A computer: You will need access to a computer for this course. If you have a laptop, get in the habit of bringing it to class. Mac computers are ideal for Bioinformatics. If you're in the market for a new laptop, a Mac *will* be the most cost-effective option as you progress in your career as a biologist.

Recommended References
A list of references is available on my teaching website here: http://caporaso.us/teaching.

Course Schedule
http://bit.ly/bio599-sch
A tentative course schedule is provided at the link above. The schedule will change over the course of the semester as we identify certain areas that we want to spend additional time on. For that reason the outline is provided in a dynamic format. The spreadsheet at the above link will always be the definitive source for the outline, reading assignments, and homework assignments. **Check this link regularly!**

Assessment of Student Learning Outcomes
You will be assessed primarily thorough graded material (see *Grading System*). A small portion of your assessment will be based on participation. This allows me to determine if you understand the material well enough to ask questions and engage in discussion in class and in office hours.

I will strive to return all graded materials within seven days, and hopefully sooner. Under some circumstances I may need a little more time. Check with me at any time during the semester on your current participation score.

Grading System

- 10% Participation
- 20% Short quizzes (around 20-30 minutes), you will have one-day warning and lowest quiz score will be dropped.
- 20% Presentations
- 50% Graded homework assignments and projects

Grading Scale

A: 90-100
B: 80-90
C: 70-80
D: 60-70
F: 0-59

Course Policy

- If you have a legitimate reason why you cannot make one of the quiz times or why you won't be able to complete an assignment by the due date, talk with me early on and something can be arranged. Students who do not show up for quizzes without making prior arrangements will get a zero.
- Students are not graded on attendance, but there will be **no make-ups of quizzes** so it is highly recommended that you don't miss classes when a quiz is scheduled. If you have to miss a lecture and there is a quiz that day, it will count as your dropped quiz score.
- Plagiarism and cheating will not be tolerated. Any students found guilty of either will receive a failing grade in the class. We will discuss what counts as plagiarism of software on the first day of the course.
- No computers, cell phones, headphones, books or papers may be used during quizzes. If a student is observed looking at or touching any of these items during a quiz it will be considered cheating – this includes checking to see who just texted you! To be safe, I recommend turning off cell phones during quizzes.
- Computers may be used for course-related work during the class such note taking or working along with programming examples that are being done on the overhead. You may not however use computers for non-course-related work such as checking e-mail or Facebook. If I notice you doing this I will document it and count it against your participation grade in the course.
- On the first week of class you'll sign up for the course e-mail list. I expect you to check your e-mail at least once per day.

University Policies

The Safe Environment, Students with Disabilities, Institutional Review Board, Academic Integrity, Academic Contact Hour, Classroom Management and Professional Ethics and Code of Conduct policies are available at http://www4.nau.edu/avpaa/policy1.html. Students are responsible for reviewing and understanding these policies.