ADVANCED BIOCOMPUTING: BIOL SPECIAL TOPICS 7970-004 (CRN: 17894)
FALL 2011

INSTRUCTOR: Dr. Scott R. Santos, Department of Biological Sciences
(santos@auburn.edu)

DATES: September 12-23

LECTURES: Lectures every day 9-12.30 at conference room D (first week) and A (second week, Mon-Wed) at the VilVite Center, and at N-terminalen conference room, 5th floor, Bioblokken (Department of molecular biology) Thu-Fri the second week.

REQUIRED TEXTS:
Barrett, D.J. - "Linux Pocket Guide" by O'Reilly Book (Amazon for $7.99; Book-A-Million for $9.95)

Haddock, S.H.D. and Dunn, C.W. - "Practical Computing for Biologists" (Paperback) ($60.00 from Amazon)

ALTERNATIVE TEXT:
Wünschiers, R. - "Computational Biology: Unix/Linux, Data Processing and Programming" (Paperback) ($32.00 from Amazon)

STUDENTS ARE REQUIRED TO HAVE THEIR OWN WINDOWS OR APPLE LAPTOP FOR USE IN CLASS. ALL FILES ON THE LAPTOP SHOULD BE BACKED UP PRIOR TO THE FIRST DAY OF CLASS AND THERE SHOULD BE (AT LEAST) 20GB OF FREE HARD DRIVE SPACE ON THE MACHINE.

PREREQUISITES: Background in molecular genetics and statistics is strongly required. It is expected that everyone has already taken some type of statistics course as well as courses in molecular biology, etc. Lectures will assume that everyone is comfortable with these subjects. If you are not, I would suggest dropping the course and taking it at a later date.

COURSE DESCRIPTION: Over the last decade, biological data sets have been rapidly growing in size and complexity. This course focuses on how to use computers to streamline the analysis of biological data, with an emphasis on "working smart" rather then hard. Developing a solid background in how computers can facilitate biological research will not only help with your thesis projects, but will also make you more "marketable" for post-docs and faculty positions in the future.

ADDITIONAL STUDY AIDS FOR THIS COURSE:
Additional reading materials, in the form of PDFs and HTML documents, will be used to supplement readings from the above books. These will be distributed either via email attachments or posted links from a computer server. These materials will be made available on the Friday following the class that they are assigned in.

IMPORTANT INFORMATION OF SPECIAL NOTE:
• **Read the assigned materials prior to coming to lecture.** The lectures are meant to clarify and discuss concepts, not to serve as your first exposure to them. Each lecture is presented with the assumption that you have read the material and are at least vaguely familiar with it.

Since the lectures will tend to build upon the previously covered materials, it is highly recommended that

you review your notes before the next lecture (if possible). ATTENDANCE IS MANDATORY.
GENERAL POLICY and PROCEDURES: You should retain this schedule of lecture topics and relevant instructions for reference. You are responsible for learning the material that will be covered, for preparing for lectures by reading assignments beforehand, and for being present at all lectures without further notice or additional reminders.

Special Request: Cell phones and pagers should be turned off for the duration of the lecture. Students will be asked to leave the classroom for the remainder of the lecture in the event one of these devices is activated during the lecture.

GRADING: There will be no grading of the course. Credits will be awarded on the basis of attendance at all course activities.

BIOL 7970-004 Advanced Biocomputing Fall 2010

LECTURE SCHEDULE BY SUBJECT MATTER (TENTATIVE AND SUBJECT TO CHANGE)

Installation of Linux (PCs) and Developer Tools (Apple)

The "shell" and basic UNIX commandline

The structure of the UNIX file system

Simple programs and how to automate commands via scripts (shell, PERL, etc)

Compiling programs from scratch (using EMBOSS as an example)

Analysis of molecular data

Statistics with R

Graphics with R

Text manipulation using grep, sed and awk

Geographic mapping using the Generic Mapping Tools (GMT)

Image formats: what's the difference between them and manipulation using Imagemagick

Creation of graphs and figures for publications using ploticus, GIMP and Inkscape

Web servers and databases