Course Description:
This is a continuation of Bioinformatics I from last semester. In the previous course, you studied the basics of sequence comparisons, multiple sequence alignment, pattern recognition, and tree construction. In the current course, we will expand the application of these principles to real world genomic data. Specifically, we will start with the assembly of sequence data and proceed through identification and annotation of protein genes to analyzing mRNA and protein expression. Time permitting, we may also investigate current, specialized topics in bioinformatics and business. If there are topics of particular interest to you, please talk to me and we'll see if they can be incorporated into the syllabus.

The goal of this course is to further develop your skills in applied bioinformatics and to provide substantial hands on experience in sequence analysis real data.

Professor:
Dr. Stephen Aley is a molecular biologist specializing in infectious disease. He has been part of the genome project for *Giardia lamblia*, and is currently involved in the annotation of that genome as well as data mining and proteomics of that same organism. You can usually find him in one of B303 (office), B310 (lab) or with the DNA sequencers or LC/MSMS. He can be contacted by email (saley@utep.edu), WebCT, or phone (747-6997). His formal office hours are immediately following class (T 3:30 or R 4:30), but the best way to see him is to simply drop by his office or lab -- if there isn't time to talk right then, he'll schedule a mutually agreeable time. There is a rumor that he has a chocolate dispenser on his desk for visiting students. Don’t worry – he doesn’t normally talk about himself in the third person.

Teaching Assistant:
We are fortunate to have Nam Tonthat as a teaching assistant for this course. Nam will be assisting me with the computer laboratory exercises and will have grade many of the assignments. His office hours will be posted elsewhere.

Course resources and text:
The Required Text Book for this semester is the same text as for part one – Bioinformatics by David Mount. A programming guide on your programming language of choice will also be beneficial (although Google does know all...) – Beginning Perl for Bioinformatics, from first semester, will do fine. Readings from the text or from other sources will be assigned as needed.

The course will be coordinated through a WebCT course connection. If you are not familiar with WebCT, please see the instructor. WebCT will provide an online syllabus, course calendar, course bulletin board, and some supplemental web sites and notes for lectures. Grades will also be presented through WebCT. WebCT will also be used for pre and post class assessments.

Computational problems in this semester will emphasize command line interfaces (e.g., unix, linux, or even DOS) and customized programs or scripts for parsing output. My preference is Perl (recommended if you are looking for programming support from Dr. Aley!), but you may use any language with which you are comfortable. The programming components will be results oriented. Most of the software, including all major programs, will be available on STAR, the IBM S590 that serves as the University’s mainframe and is accessed over network connections by SSH. In addition, we will make use of Sun stations in Biostatistics and other unix boxes. That being said, you will find that that most bioinformaticians perform their day to day work on a personal laptop! If you have a laptop computer, I strongly recommend that you practice using it in and for the course, if only for access to mainframes. A few University laptop computers may be available as a loan for the semester. If you are interested, please see Dr. Aley.

Class components and Grades:
The primary goal of this course is to develop and practice approaches and skills that would apply to a full scale bioinformatics program. As such, the grading for the course will emphasize the practical application of acquired skills. In particular, at most 30% of course grade determined by traditional exams. Class
participation includes coming to class having read the assigned material, active participation in course
dialogue and participation in group and individual exercises. You should come to class with a good
understanding of the reading and have several questions ready to discuss. This activity will not be directly
graded, but will influence final grade assignment, UP OR DOWN.

Grade Point Distribution:

Computer Lab Exercises and Followups  30%.
Lab exercises are generally cookbook procedures designed to familiarize you with a particular analysis
or approach. They are expected to be completed during the laboratory portion of the course (Thursday,
1:30 to 4:20).

Most exercises will include a follow up exercise using real data, generally done outside of formal class
hours.

Software installation and manual development 10%.
These assignments are to mimic installing and maintaining public domain software for use by OTHER
research workers from diverse levels of bioinformatic experience. Each student will take responsibility
for the installation or upgrade of a major software program set on the STAR system. The final product
should be a functioning research package for use by the UTEP community, including scripts for
database maintenance (if appropriate), user manuals, training PowerPoint presentation, and tutorial
examples.

Informatics Research Project 30%
Each student will initiate and develop an informatics research project requiring topic research plus
application (or development) of bioinformatics tools, and interpretation and presentation of results.
Actual topic is flexible, but requires approval of professor. Deliverables will include a white paper,
midterm progress report, and a final full written report and oral presentation.

Individual Examination(s)  30%
Individual, written examinations will be used for assessment of understanding and application of basic
concepts format. Previous classes have opted for a single, comprehensive final examination, however a
mid-term plus a final examination, both comprehensive, could also be considered. If you have
preferences, please communicate them to the instructor! The Final Exam for this course is scheduled for
Thursday, 11 May 2006, 1- 4 pm

COURSE POLICIES POLICY ON MAKEUP EXAMINATIONS AND LATE WORK:
NO makeup exams or due date extensions will be given for reasons other than illness (doctor's note
required), absence with the instructor's prior approval, or travel on official University business
(documentation required BEFORE the absence). Makeup exams will be scheduled at instructors discretion.
The same policy will be followed for missed laboratory work.

WHOSE WORK IS IT:
For ALL ASSIGNMENTS and PROJECTS, you are to ASSUME that all work is to be done by the
INDIVIDUAL student whose name is on the cover page unless the WRITTEN instructions specifically state
otherwise. When teams are allowed and used, ALL PARTICIPATING TEAM MEMBERS must be listed
with their specific contributions. When outside work is incorporated (e.g., perl code from outside sources or
copied powerPoint slides) the source must be clearly credited. All written reports must clearly cite sources
in the format of a standard journal. Omissions will be reflected in the grading of that assignment.

POLICY ON ACADEMIC HONESTY:
Academic Dishonesty will not be tolerated. All university guidelines will be strictly followed. Please read
these guidelines carefully. If you have any questions regarding the university policy please contact the Dean
of Students.