ANSC 542 CPSC 569
APPLIED BIOINFORMATICS

Semester: SUMMER I 2012
Credit: 4 hours

Course Information

Instructor (first section): Gustavo Caetano-Anollés, Ph. D.
May 14 to May 18, 2012
Classroom: 607 IGB
Office: Room 332, National Soybean Research Center
Email: gca@illinois.edu
Phone: (217) 333-8172
Office hours: T Th 11:45-12:15 AM
Office hours: Before lectures

Instructor (second section): Sandra Rodriguez Zas, Ph. D.
May 21 to May 25, 2012
Classroom: 23 LIAC (ACES Library Computer Lab, basement)
Email: rodrgzzs@illinois.edu
Office: Room 30, Animal Sciences Laboratory
Phone: (217) 333-8810
Office hours: After lectures or by appointment

Instructor (third section): Maria Villamil, Ph. D.
May 29 to June 1, 2012
Classroom: 607 IGB
Email: villamil@illinois.edu
Office N323, Turner Hall
Phone: (217) 333-4690
Office hours: Before lectures

Classroom lecture-laboratory:
M, T, W, TH, F 10:30 AM - 12:20 AM.
Classrooms:
Computer lab, 607 Institute for Genomic Biology (IGB) May 14 to May 18, May 29 to June 1, 2012
Computer lab, 23 LIAC (ACES Library Computer Lab, basement) May 21 to May 25, 2012

Course name: ANSC 542 / CPSC 569 / IB 506 Applied Bioinformatics
Course web address: https://compass2g.illinois.edu/
Compass web page login and password: NetID and associated password, respectively.
Compass support: http://www.cites.illinois.edu/illinoiscompass/studentresources.html

Prerequisites: Graduate or undergraduate status and basic knowledge on molecular biology.

Reading requirements: Class notes corresponding to each class, before class.

Objectives: Genomic and proteomic projects generate large amounts of complex data that challenge the effective storage, analysis and interpretation of biological information. Applied bioinformatics integrates concepts in biology, statistics, computer and information sciences to effectively use the bioinformatic resources available. The course combines lectures and hands-on experiences that facilitate the students' understanding of concepts and a wide range of commonly used databases and tools. The students will be introduced to a variety of bioinformatic databases and tools and will apply the concepts to individual projects related to their own research or interest.

Applied Bioinformatics includes hands on exercises in databases, sequence alignment, phylogenetic analysis, statistical analysis using R, biomolecular folding, and the use of web resources in bioinformatics. Classes include lectures and demonstrations, and students will have the opportunity to use computer resources in every class.

Outline of topics

Section 1 (Instructor: Dr. Caetano-Anollés)

May 14 - May 18, 2012

1. Introduction: Patterns, processes, and science in bioinformatics. Genes in genomes and populations. Trees and genomic archaeology
2. Patterns in molecular evolution: Inferring molecular phylogeny. Phylogeny and Bayesian inference. Processes in molecular evolution: Evolutionary models
3. Comparative genomics and phylogenomics
4. Cladistics and molecular evolution. Experimental focus: Phylogenetic reconstruction with PAUP* *Assignment 1*
5. Molecules and structure: Patterns and form in RNA. RNA folding, evolutionary processes, and the history of our natural world. Experimental focus: RNA folding.
6. Molecules and structure: Patterns and form in protein structure. Structural diversity in proteins. Protein domains and evolution. *Assignment 2*

Section 2 (Instructor: Dr. Rodriguez Zas)

May 21 - May 25, 2012

1. Introduction to bioinformatics and Biology Workbench *Assignment selection of gene class*
2. Global and local sequence alignments and PAM and BLOSUM scoring schemes
3. BLAST
4. Statistics of alignments and multiple sequence alignment *Assignment 3*
5. Patterns and profiles. PSIBLAST, detection of restriction sites and primer design
6. Hidden Markov models (HMM) applications. *Assignment 4*

Section 3 (Instructor: Dr. Villamil)

May 29 – June 1, 2012

1. Introduction to R
7. Applications of R to statistical analysis
8. Applications of R to statistical analysis
9. Applications of R to bioinformatics. *Assignment 5*

Attendance of lectures and laboratories

Class materials will be complemented with additional information provided during the class and laboratory hours by the instructors. Attendance to lectures is obligatory. Students that miss a class are expected to read the corresponding class materials and obtain the additional information provided in class from a student that attended the class.

Grading

Grading will be based on 5 assignments. Each assignment has a value of 200 points and all assignments total 1000 points. Attendance of lectures is required and demonstration of work on laboratory activities is a plus. A total points-to-grade scale will be based on groups in the total final cumulative score of all the assignments.

Deadlines: Assignments will be posted Wednesday May 16, Friday May 18, Wednesday May 23, Friday May 25, and Friday June 1, 2012. Assignments will be due May 18, May 21, May 25, May 29 and June 6, 2012 at 10:30 AM sharp. An additional deadline is May 23, by when all students must have submitted the name of the selected class of gene or protein that will be used in the assignments due May 25 and May 29. Gene class names must be submitted through the course Compass2g website.

Assignment	Posted	Deadline	Points	Instructor
1	May 16	May 18	200	Caetano-Anollés
2	May 18	May 21	200	Caetano-Anollés
Gene class	May 21	May 25	0	Rodriguez Zas
3	May 23	May 25	200	Rodriguez Zas
4	May 25	May 29	200	Rodriguez Zas
5	June 1	June 6	200	Villamil

Assignment submission: All assignments must be electronically submitted using the

course *Compass2g* webpage by the respective deadlines. Electronic mail attachments or printed assignment submissions will not be accepted. Only one file will be accepted per homework assignment and student. The assignment file must be named with the *netid* (or last name) of the student followed by the homework assignment number (e.g. rodrgzzs_hwk3, gca_hwk1, or villamil_hwk5). Only one assignment submitted no later than 3 days (72 hours) after the deadline will be accepted. This exceptional late submission must be previously arranged with the instructor responsible for the particular assignment.

IMPORTANT: Students are responsible of ensuring that their work is correctly and successfully submitted electronically and should notify the instructor of any problems with the internet connection or website at least 10 minutes before the assignment deadline. Students are encouraged to submit their assignments at least 20 minutes before the deadline.

Basic outline of assignments and final project: The assignments consist of one or two components. One component includes general questions and exercises based on the materials covered in the lectures, labs and class notes. The second component includes the application of the materials covered in the lectures, labs and class notes to real life scenarios.

Homework handouts that describe the assignment and steps required for its completion and will be available for download from the Assignment link in the course *Compass2g* webpage. Assignments will encourage problem solving and will include a report that must be legibly written and should express the findings and opinion of each student.

Academic Integrity

The Code on Campus Affairs and Handbook of Policies and Regulations Applying to All Students (available at http://www.uiuc.edu/admin_manual/code/) gives complete details on the students' rights and responsibilities. Students are responsible for knowing and abiding by these rules.

Policies on computer resources and copyrights

All students must adhere to the rules and policies indicated by the software, websites and computer laboratories used for course related purposes. The policy on course notes and related printed and internet materials (e.g. published articles, website information) copyrights follows The General Rules Concerning University Organization and Procedure (University of Illinois Board of Trustees, 1998) and can be found at http://www.vpaa.uillinois.edu/policies/ and any other rule mentioned in the materials.

Recommended reading

Higgs PG, Attwood TK (2005) Bioinformatics and molecular evolution. Blackwell Publishing.

Suggested readings

Caetano-Anollés G (2010) Evolutionary genomics and systems biology. Wiley

Gibas C, Jambeck P (2001) Developing bioinformatics computer skills. O'Reilly.

Li W-H (1997) Molecular evolution. Sinauer Associates.

Waterman MS (1996) Introduction to computational biology: Maps, sequences and genomes. Chapman & Hall.

Bishop MJ (1999) Genetics Databases. Academic Press.

Setubal J, Meidanis J (1997) Introduction to computational molecular biology. PWS Pub.

Baxevanis AD, Ouellette BFF (1998) Bioinformatics. A practical guide to the analysis of genes and proteins. John Wiley & Sons.

Weir BS (1996) Genetic data analysis II: methods for discrete population genetic data. Sinauer.

Page RDM, Holmes EC (1998) Molecular evolution:a phylogenetic approach. Blackwell Sci.

Baldi P, Brunak S, Brunak S (2001) Bioinformatics. MIT Press.

Campbell AM, Heyer LJ (2002) Discovering genomics, proteomics, and bioinformatics. Pearson Educ.

Ewens WJ, Grant GR and Grant G (2001) Statistical methods in bioinformatics: an introduction. Springer-Verlag.

Mount DW, Mount D (2002) Bioinformatics: sequence and genome analysis. Cold Sping Harbor.

Krane DE, Raymer ML (2002) Fundamental concepts of bioinformatics. Pearson Educ.

Claverie J, Notredame C (2003) Bioinformatics for Dummies. John Wiley & Sons.