BIOS 10602: Multiscale Modeling of Biological Systems I

Instructor: Esmael J. Haddadian, Ph.D. haddadian@uchicago.edu
Office hours: Thursday and Friday 3:00 – 5:00 pm or by appointment, BSLC 222

Lab TA, Tue: Albert Nam albertnam@me.com
Lab TA, Wed: Christina Cheng chcheng@uchicago.edu

Class TA: Davina Moossazadeh davina@uchicago.edu
Office hours: Mon, Wed, 5:30-6:30 pm, Second floor of BSLC

Lectures: Tue, Th; 11:00 AM-12:20 PM; BSLC 008
Computer Labs: Tue, Wed 2:00-5:20 PM; BSLC 401

Goals: Modern biology generates massive amounts of data; handling and analyzes of this data requires mathematical and computational methods. The first course in the sequence is devoted to biological information and the models and computational techniques used to make sense of it. The course begins with an introduction to the chemical context of life, water and life, and the organization of life at the molecular level. Students will learn about the structure and function of large biological molecules (such as DNA, RNA and proteins), energy and matter transformations and evolution. Students learn about biological databases, algorithms for sequence alignment and phylogenetic tree building. Students will also be introduced to basics of high performance computation and its application to the field of bioinformatics. They will learn how to use our in-house supercomputer to process and analyze next generation gene sequencing data in order to identify disease-relevant variants. Students implement computational algorithms using R and Unix.
After taking this sequence a student will: a) understand biological concepts commonly encountered in popular and technical literature; b) have a background in mathematical and computational tools used in biological settings; c) be prepared to work in collaborative teams with biologists, either in basic research or in applications, for example in pharmaceutical, information retrieval, or medical device industries; d) learn how to write a scientific report.

Grading:
25% Computer lab assignments and homework (lab reports are due at midnight of the day of the lab, lab homework is due on the Sunday midnight after the lab)
25% Weekly in class quizzes on last week lectures, current week topic and assigned papers; quizzes are given on Th
15% Weekly homework assignments, class discussions and presentations
35% Final project

Final project:
Students will use the super computer Midway to process and analyze whole-exome next generation sequencing data. The exome is the part of the genome formed by exons, the sequences which when transcribed remain within the mature RNA after introns are removed by RNA splicing. Exome sequencing is a process by which the DNA from canonical coding regions (genes) is selectively captured and sequenced. Using state of the art tools, students will align and genotype samples of exomes from the 1000 genome project in order to identify disease-relevant variants. They will subsequently provide a report of their findings.

Textbooks:
1. **(Required)** Bioinformatics for Beginners: Genes, Genomes, Molecular Evolution, Databases and analytical Tools, Supratim Choudhuri. ISBN: 978-0-12-410471-6.

2. Campbell, Biology (11th edition), Pearson Benjamin Cummings Publishers. ISBN: 978-0-13-409341-3.

3. Bioinformatics and Functional Genomics, Jonathan Pevsner. Third edition. ISBN: 978-0-470-08585-1.

Lab Software:

R:	http://www.r-project.org
R-studio:	http://www.rstudio.com
Unix commands:	https://kb.iu.edu/d/afsk
	http://mally.stanford.edu/~sr/computing/basic-unix.html

Policy on late/missing work: Computer assignments must be completed during scheduled lab time. If you cannot come to your assigned lab please contact me to arrange to participate in another session in the same week, otherwise makeup labs are not allowed. Quizzes have to be completed in the allocated time during class. No late homework will be accepted, unless an arrangement is made with the instructor ahead of time. Retroactive arrangements may be made only in cases of genuine emergency. Students are responsible to check their grades on Canvas and report a missing grade no later than a week after the graded assignment is returned.

Attendance: Attendance is required, but 3 absences are absolutely free. Save those allowed absences for when you really need them! Beyond the 3 pre-excused absences, there will be a penalty of 1% of the total course grade for any extra absences. If you don't sign the attendance sheet, you won't be considered present.

Extra readings will be posted on chalk, accompanied by an announcement.

	Date	Topic & Reading	Lab
Week-1	25-Sept to 29- Sept	Introduction, Science of learning The chemical context of life, Water and life *Book: Chapters 2, 3, Campbell biology*	Introduction to R
Week-2	2-Oct to 6-Oct	Carbon and the molecular diversity of life The structure and function of large biological molecules *Book: Chapters 4 and 5, Campbell biology*	Introduction to graphing in R
Week-3	9-Oct to 13-Oct	The molecular base of inheritance *Book: Chapters 16, Campbell biology* Data and databases search *Book: Chapter 5, Bioinformatics for Beginners*	Retrieving NCBI-data using R and sequence analysis, part 1
Week-4	16-Oct to 20-Oct	Genomic technologies *Book: Chapter 3 and parts of Chapter7, Bioinformatics for Beginners* Visiting DNA Sequencing & Genotyping Core; **TBA** Address: KCBD 1230 H&I, 900 E 57th Street	Sequence analysis, part 2
Week-5	23-Oct to 27- Oct	Sequence alignment and similarity *Book: Bioinformatics for Beginners (chapter 6) and Bioinformatics and Functional Genomics (chapter 3)*	R-sequence alignment, part 1
Week-6	30-Oct to 3-Nov	Sequence alignment and similarity (continued) *Book: Bioinformatics for Beginners (chapter 6) and Bioinformatics and Functional Genomics (chapter 3,4)*	R-sequence alignment, part 2
Week-7	6-Nov to 10-Nov	Sequence alignment and similarity (continued) An introduction to super computing *Bioinformatics and Functional Genomics (chapter 3,4)* Visiting research computing center; TBA	Basics of Unix Commands
Week-8	13-Nov to 17-Nov	An introduction to evolution and phylogenetic analysis *Book: Bioinformatics for Beginners (chapter 9)* *Campbell (chapter 26)*	Building phylogeny trees in R
Week-9	20-Nov to 24-Nov	Energy and matter transformations An introduction to metabolism *Book: Chapters 8, Campbell biology*	Exome sequencing
Week-10	27-Nov to 1-Dec	Human Genetic Diseases *Book: Bioinformatics and Functional Genomics (chapter 20)*	No-LAB