

COSI-178A Computational Biology

Instructor: Pengyu Hong. Email: hongpeng@brandeis.edu

Description: Information and computing technologies are becoming indispensable to modern biological research due to the significant advances of high-throughput experimental technologies in recent years. This course presents an overview of the systematic development and application of computing systems and computational algorithms/techniques to the analysis of biological data (such as, sequences, gene expression, protein expression, biological networks, and so on). It is intended for advanced undergraduate students and graduate students. Hand-on training will be provided.

* This course will not teach programming. Some assignments (include the term project) require programming. Students can choose any of the following programming languages: C++, Java, Matlab, Python, or R.

Learning Objectives: At the end of the course, students should be able to describe the key molecular concepts that are relevant to bioinformatics analyses, locate online resources/databases and either manually or write programs to retrieve data from them for bioinformatics analyses, describe fundamental algorithms in sequence analysis and conduct biological sequence analyses (e.g., pair-wise sequence alignment, multiple sequence alignment, fast sequence search, interpret sequence analysis results, etc.) by using a variety of tools (e.g. BLAST, BWA, Clustal Omega, MUSCLE, COBALT, etc.), describe the key concepts and steps in molecular evolution/phylogeny and conduct phylogenetic analysis using online tools, describe the basic principles of genome-wide gene expression analyses (DNA microarray and RNA sequencing), conduct gene expression profile analyses (clustering, classification, differential expression analysis), perform functional analysis on gene lists using online tools (e.g., DAVID), describe the basic principles of whole genome shotgun sequencing, and describe the basic principles of mass spectrometry analysis of proteins/glycans (if time allows).

Course Plan:

- Biological Databases
- Sequence Analysis
 - Genome Sequencing
 - Pair-wise sequence alignment
 - Fast sequence searches
 - Multiple sequence alignment
 - Phylogeny
- Transcription Regulation and Gene Expression Profile Analysis
 - Transcription regulatory motif finding
 - Gene Ontology
 - Gene Expression Profiling Technologies
 - Gene expression profile analysis
 - Normalization and gene expression index
 - Differentially expressed gene detection and hypothesis testing
 - Clustering
 - Classification

- Dimension reduction techniques
- Whole genome sequencing technology (introduction)
- Proteomics/Glycomics (Mass spectrometry data analysis, if time allows)

Success in this 4-credit hour course is based on the expectation that students will spend a minimum of 9 hours of study time per week in preparation for class (readings, papers, discussion sections, preparation for exams, etc.).

Evaluation: 6-8 Homework Assignments (70%). A term project (30%)

Disabilities: Brandeis seeks to welcome and include all students. If you are a student who needs accommodations as outlined in an accommodations letter, please talk with me and present your letter of accommodation as soon as you can. I want to support you. In order to provide test accommodations, I need the letter more than 48 hours in advance. I want to provide your accommodations, but cannot do so retroactively. If you have questions about documenting a disability or requesting accommodations, please contact Student Accessibility Support (SAS) at 781.736.3470 or access@brandeis.edu.

Academic Integrity: You are expected to be honest in all of your academic work. Please consult [Brandeis University Rights and Responsibilities](#) for all policies and procedures related to academic integrity. Students may be required to submit work to [TurnItIn.com](#) software to verify originality. Allegations of alleged academic dishonesty will be forwarded to the director of academic integrity. Sanctions for academic dishonesty can include failing grades and/or suspension from the university. [Citation and research assistance](#) can be found on the [university library website](#).

Privacy: This class requires the use of tools that may disclose your coursework and identity to parties outside the class. To protect your privacy you may choose to use a pseudonym/alias rather than your name in submitting such work. You must share the pseudonym/alias with me and any teaching assistants as needed. Alternatively, with prior consultation, you may submit such work directly to me.

Communications: All communications (such as, course announcements, forum discussions, changes of syllabus, snow-day arrangements, and so on) will be on LATTE.