Syllabus of BIOINF 528 (2017 Fall, Bioinformatics)

Course Name:

Advanced Applications of Bioinformatics

Course Description:

This course introduces fundamental concepts and methods for bioinformatics and the advanced applications. Topics covered include bioinformatics database, sequence and structure alignments, protein folding and protein structure prediction, protein-protein interaction, Monte Carlo simulation, and structural determination of macromolecules (X-ray crystallography and NMR). Emphasis is on the understanding of the concepts taught and their practical utilization, with the objective of helping student use cutting-edge bioinformatics tools/methods to solve problems in their own research.

Instructor:

Yang Zhang, Email: zhng@umich.edu, Phone: 734-647-1549

Invited lecturers:

Philip Andrews, Tomek Cierpicki, John Tesmer

Schedule and location:

9:00 am - 12:00 noon, Friday; #2036 Palmer Commons

Homework, lab & Exams:

There will be homework assignments, including code writing and literature reading. Homework and lab work constitutes 60% of your grade; the final exam constitutes another 40%.

Student evaluation and grades:

A = 90-100% B = 80-89 C = 70-79 D = 60-69 F = below 60

Textbook:

No textbook is required for this course. Assigned materials will be either handed out to the class or posted on the blackboard website.

Table of content

3.1. Neighbor-joining method and phylogenetic tree
3.2. How to construct multiple sequence alignment?
3.2.1. ClustalW
3.2.2. PSI-BLAST
3.2.3. Hidden Markov Models
3.2.3.1. Viterbi algorithm
3.2.3.2.HMM based multiple-sequence alignment
3.2.3.2.1. HMMER
3.2.3.2.2. SAM
3.3. Sequence profile & profile based alignment
3.3.1. Profile-to-sequence alignment
3.3.2. Profile-to-profile alignment

4. Protein structure alignments
4.1. What is sequence-dependent structure superposition?
4.1.1. RMSD
4.1.2. TM-score
4.2. What is sequence-independent structure alignment?
4.3. Different structure alignment algorithms
4.3.1. DALI
4.3.2. CE
4.3.3. TM-align
4.4. How to define the fold of proteins?
4.5. Number of protein folds in PDB

PART II: PROTEIN FOLDING AND PROTEIN-PROTEIN INTERACTIONS

5. Protein secondary structure predictions
5.1. What is protein secondary structure?
5.2. Hydrogen bond
5.3. How to define a secondary structure element?
5.4. Methods for predicting secondary structure
5.4.1. Chou and Fasman method
5.4.2. PHD
5.4.3. PSIPRED
5.4.4. PSSpred

6. Monte Carlo Simulation
6.1. Introduction: why Monte Carlo simulation?
6.2. Monte Carlo Sampling of Probabilities
6.2.1. Random number generator
6.2.1.1.How to test a random number generator?
6.2.2. Sampling of rectangular distributions
6.2.3. Sampling of probability distribution
6.2.3.1.Reverse transform method
6.2.3.2.Rejection sampling method
6.3. Boltzmann distribution
6.4. Metropolis method

6.5. Advanced Metropolis methods
    6.5.1. Replica exchange simulation
    6.5.2. Simulated annealing

7. Protein tertiary structure modeling

7.1. Basic concepts
7.2. Ab initio modeling
    7.2.1. Anfinsen thermodynamic hypothesis
    7.2.2. Molecular dynamics simulation
        7.2.2.1.CHARMM
        7.2.2.2.AMBER
    7.2.3. Knowledge-based free modeling
        7.2.3.1.Bowie-Eisenberg approach
        7.2.3.2.ROSETTA
        7.2.3.3.QUARK
        7.2.3.4.Why is beta-protein so difficult to fold?
7.3. Comparative modeling (homology modeling)
    7.3.1. Principle of homology modeling
    7.3.2. Modeller
7.4. Threading and fold-recognition
    7.4.1. What is threading?
    7.4.2. Threading programs
        7.4.2.1.Bowie-Luthy-Eisenberg
        7.4.2.2.HHpred
        7.4.2.3.MUSTER
    7.4.3. Meta-server threading
        7.4.3.1.3D-jury
        7.4.3.2.LOMETS
7.5. Combined modeling approaches
    7.5.1. TASSER/I-TASSER
7.6. CASP: A blind test on protein structure predictions

8. Principle of X-ray Crystallography & Molecular Replacement

8.1. What is X-ray Crystallography
8.2. Why can a wave be represented by $\exp(i\alpha)$?
8.3. How to calculate scattering on two electrons?
8.4. What is Laue condition?
8.5. What is Bragg's law?
8.6. How to calculate electron density of crystal?
8.7. What is Patterson function?
8.8. How to calculate electron density of crystal?
8.9. What is the idea of Molecular Replacement?
8.10. How to judge quality of MR?
8.11. What are often-used software for MR?

9. Introduction to nuclear magnetic resonance (NMR)

9.1. Basic magnetic property of nuclei
    9.1.1. Magnetic moment
    9.1.2. Nuclei in external magnetic field

9.1.3. Nuclear shielding of magnetic field
9.2. Chemical shift
9.3. NMR spectrum
9.3.1. Correlation spectroscopy (COSY)
9.3.2. Heteronuclear single-quantum correlation spectroscopy (HSQC)
9.3.3. Nuclear Overhauser effect spectroscopy (NOESY)
9.4. From NOE to 3D structure model

10. Protein function and structure-based function annotation

10.1. Gene ontology
10.2. Enzyme classification
10.3. Ligand-protein interaction
10.4. Structure-based function annotation
10.4.1. FindSite
10.4.2. COFACTOR
10.4.3. COACH

11. Modeling of protein-protein interactions

11.1. Experimental identification of protein-protein interactions
11.1.1. Yeast two-hybrid assay
11.1.2. High-throughput mass spectrometry
11.1.3. Interaction networks and system biology
11.2. Protein-protein quaternary structure modeling
11.2.1. Basic concepts
11.2.1.1. Degrees of freedom
11.2.1.2. Presentation of protein conformations
11.2.1.3. Hydrophobicity factor
11.2.1.4. Shape complementary
11.2.1.5. Docking Scoring function
11.2.2. Protein-protein docking algorithms
11.2.2.1. Fast Fourier Transformation (FFT)
11.2.2.2. Semi-flexible docking: Side-chain refinement
11.2.2.3. Clustering and refinement
11.2.3. Homology modeling
11.2.4. Multiple-chain threading
11.2.4.1. COTH
11.2.4.2. SPRING
11.2.5. Monte Carlo docking simulation
11.2.6. CAPRI: A blind test on protein-protein docking

PART III: Selected Topics (by invited speakers)

Prof. Philip Andrews:
Introduction to mass spectrometry and proteomics

Prof. Tomek Cierpicki:
NMR spectroscopy for protein structure determination

Prof. Daniel Southworth:
Cryo-electron microscopy for protein structure determination

Prof. Matthew Young:

X-ray crystallography for protein structure determination