Syllabus of BIOINF 528 (2017 Fall, Bioinformatics)

Course Name:

Structural Bioinformatics

Course Description:

This course introduces fundamental concepts and methods for structural bioinformatics and the advanced applications. Topics covered include sequence, structure and function databases of DNA and protein molecules, advanced sequence and structure alignment methods, methods of protein folding and protein structure prediction (homologous modeling, threading and *ab initio* folding), basics of molecular dynamics and Monte Carlo simulation, principle and application of machine learning, and techniques of protein structure determination (X-ray crystallography, NMR and cryo-EM). Emphasis is on the understanding of the concepts taught and the practical utilization, with the objective to help students to use the cutting-edge bioinformatics tools/methods to solve problems in their own research. For this term, four world's top experts are invited to give lectures on mass spectrometry and proteomics (Prof. Philip Andrews), NMR spectroscopy (Prof. Tomek Cierpicki), Cryo-electron microscopy (Prof. Melanie Ohi), and X-ray crystallography (Prof. Mark Saper).

Instructor:

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

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

2.4. Heuristic methods
    2.4.1. FASTA
    2.4.2. BLAST
2.5. Statistics of sequence alignment score
    2.5.1. E-Value
    2.5.2. P-Value

3. Phylogenic tree & multiple sequence alignments

3.1. Neighbor-joining method and phylogenetic tree
3.2. How to construct multiple sequence alignments?
    3.2.1. ClustalW
    3.2.2. PSI-BLAST
        3.2.2.1. PSI-Blast pipeline
        3.2.2.2. Profile pseudocount
        3.2.2.3. PSSM-position specific scoring matrix
        3.2.2.4. Installing and running PSI-Blast programs
        3.2.2.5. Interpret PSI-Blast out
    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. Creating HMM by iteration
            3.2.3.2.2. HMMER
            3.2.3.2.3. SAM
3.3. Sequence profile & profile based alignments
    3.3.1. What is sequence profile?
    3.3.2. Henikoff weighting scheme
    3.3.3. Profile-to-sequence alignment
    3.3.4. Profile-to-profile alignment

4. Protein structure alignements

4.1. Structure superposition versus structural alignment
4.2. Structure superposition methods
    4.2.1. RMSD
    4.2.2. TM-score
4.3. Structure alignment methods
    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 the PDB

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. Basics of machine learning and neural network methods
5.5. Methods for predicting secondary structure
    5.5.1. Chou and Fasman method

5.5.2. PHD
5.5.3. PSIPRED
5.5.4. PSSpred

6. Introduction to 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 folding and protein 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. PSI-BLAST
7.3.3. 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
7.5.1.1. Force field design
7.5.1.2. Search engine: replica-exchange Monte Carlo simultion

7.5.1.3. Major issues and recent development
7.6. CASP: A blind test on protein structure predictions

8. Protein function and structure-based function annotation
8.1. Gene ontology
8.2. Enzyme classification
8.3. Ligand-protein interaction
8.4. Structure-based function prediction
8.4.1. Concavity
8.4.2. FindSite
8.4.3. COFACTOR
8.4.4. COACH

9. Principle of X-ray Crystallography & Molecular Replacement
9.1. What is X-ray Crystallography
9.2. Why can a wave be represented by exp(iα) ?
9.3. How to calculate scattering on two electrons?
9.4. What is Laue condition?
9.5. What is Bragg's law?
9.6. How to calculate electron density of crystal?
9.7. What is Patterson function?
9.8. How to calculate electron density of crystal?
9.9. What is the idea of Molecular Replacement?
9.10. How to judge quality of MR?
9.11. What are often-used software for MR?

10. Introduction to nuclear magnetic resonance (NMR)
10.1. Basic magnetic property of nuclei
10.1.1. Magnetic moment
10.1.2. Nuclei in external magnetic field
10.1.3. Nuclear shielding of magnetic field
10.2. Chemical shift
10.3. NMR spectrum
10.3.1. Correlation spectroscopy (COSY)
10.3.2. Heteronuclear single-quantum correlation spectroscopy (HSQC)
10.3.3. Nuclear Overhauser effect spectroscopy (NOESY)
10.4. From NOE to 3D structure model

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. Melanie Ohi:
Cryo-electron microscopy for protein structure determination

Prof. Mark Saper:
X-ray crystallography for protein structure determination