

CS2220: Introduction to Computational Biology Course Briefing, 13/1/05

Limsoon Wong



Recommended “Pre-requisites”

- **CS1102: Data Structures and Algorithms**
- **LSM1102: Molecular Genetics**

Objectives

- **Develop flexible and logical problem solving skill**
- **Understand bioinformatics problems**
- **Appreciate techniques and approaches to bioinformatics**

To achieve the goals above, we expose the students to a series of case studies spanning gene feature recognition, gene expression and proteomic analysis, gene finding, sequence homology interpretation, phylogeny analysis, physical mapping, and genome sequencing

What to Expect

- **Time Table**
- **Course Syllabus**
- **Course Homepage**
- **Teaching Style**
- **Project, Assignments, Exams**
- **Readings**
- **Assessment**

Time Table

- **Lecture**
 - Friday 2:00pm – 4:00pm, LT33
- **Tutorial**
 - Thursday 10:00am – 11:00am, S16-432
- **Consultation**
 - Friday 4:00pm – 6:00pm
- **Office**
 - S16 Level 6 Room 5
- **Email**
 - wongls@comp.nus.edu.sg

Course Syllabus

- **Essence of Bioinformatics**
 - molecular biology
 - tools and instruments for molecular biology
 - themes and applications of bioinformatics
- **Essence of Knowledge Discovery**
 - classification performance measures
 - feature selection techniques
 - machine learning techniques
- **Gene Feature Recognition from Genomic DNA**
 - feature generation, selection, & integration
 - translation initiation site (TIS) recognition
 - Transcription start site (TSS) recognition
- **Gene Expression and Proteome Analysis**
 - Microarray and mass-spec basics
 - classification of gene expression profiles
 - classification of proteomic profiles
 - clustering of gene expression profiles
 - molecular network reconstruction
- **Essence of Seq Comparison**
 - Dynamic programming basics
 - Sequence comparison and alignment basics
 - Needleman-Wunsh global alignment algorithm
 - Smith-Waterman local alignment algorithm
- **Seq Homology Interpretation**
 - protein function prediction by sequence alignment
 - protein function prediction by phylogenetic profiling
 - active site and domain prediction
 - key mutation sites prediction
- **Gene Finding**
 - Overview of gene finding
 - GRAIL
 - Handling of frame shifts and in-dels
- **Phylogenetic Trees**
 - Phylogeny reconstruction method basics
 - origin of Polynesians & Europeans
 - Large-scale sequencing basics
- **Physical Mapping and Genome Sequencing**
 - Physical mapping basics
 - sequence assembly algorithm
 - shortest common superstring problem

Course Homepage

- **IVLE**
 - <http://ivle.nus.edu.sg/search/internet/search.asp?code=CS2220&title=&lec=wong>
- **Lecture Slides & etc**
 - <http://www.comp.nus.edu.sg/~wongls/courses/cs2220>

Teaching Style

- **Bioinformatics is a broad area**
- **Need to learn a lot of material by yourself**
 - Reading books
 - Reading papers
 - Practice on the web
- **Don't expect to be told everything**

Assignments, Project, & Exam

- **Assignments**
 - Probably 4 assignments
 - Some are programming assignments
- **Project**
 - Based on a case study in the class
 - 8-10 pages of report expected
- **Exam**
 - No mid-term exam ... I hope!
 - 1 final open-book exam

Be Honest

- **Exam**
 - Absence w/o good cause results in ZERO mark
 - Cheating results in ZERO mark
- **Discussion on assignments is allowed**
- **Blatant plagiarism is not allowed**
 - Offender gets ZERO mark for assignment or exam
 - Penalty applies to those who copied AND those who allowed their assignments to be copied

Background Readings

- Peter Clote and Rolf Backofen, *Computational Molecular Biology: An Introduction*, John Wiley, 2000
- Pierre Baldi and Soren Brunak, *Bioinformatics: the Machine Learning Approach*, MIT Press, 1998
- Pavel Pevner, *Computational Molecular Biology: An Algorithmic Approach*, MIT Press, 2000.
- Limsoon Wong, *The Practical Bioinformatician*, WSPC, 2004
- Dan Gusfield, *Algorithms on Strings, Trees, and Sequences - Computer Science and Computational Biology*, Cambridge University Press, 1997

Assessment

- **Continuous Assessment: 50%**
- **Final Exam: 50%**

Any questions?

I hope you will enjoy this class 😊