Computational Biology

Jiankui He

Office: room 336, The 1st faculty research building

Phone: 88015405; email: hejk@sustc.edu.cn

Semester: Fall 2016

Course description

Computational Biology is a practical, hands-on approach to the field of computational molecular biology. The course is recommended for both molecular biologists and computer scientists desiring to understand the major issues concerning analysis of genomes, sequences and structures. Various existing methods will be critically described and the strengths and limitations of each will be discussed. There will be practical assignments utilizing the tools described. All homework and coursework will be submitted electronically. Prerequisites include an introductory molecular biology course or permission of the instructor.

Course Requirements

There will be several homework assignments utilizing the tools described in the lectures. All homework and final projects will be submitted in electronic form, as e-mail or e-mail attachments to hejk@sustc.edu.cn. A final paper will be required for the course that critically and constructively analyzes any area of computational molecular biology, bioinformatics or genomics. The final project may also present a novel application of existing tools or the development of some new or improved method.

Course grading

|  |  |
| --- | --- |
|  | Percentage |
| Class performance | 10 |
| Homework | 40 |
| Final project | 50 |

Class Performance will be evaluated by: (1) attending the class and (2) asking questions. If you attend almost all classes, and ask at least 2 questions per class, you will get full credit.

Homework: all homework should be submitted before 11pm, Thursday (the day before the next class). Late submission will get a 30% penalty. No submission will be accepted after the beginning of next class. If you have a valid reason for being late (illness, conference, death in family etc.) and let us know ahead of time we can arrange a later due date for you. Please submit your homework in any of the following formats: Word, text, PDF, or HTML document. Please put all copied material in quotes with a full URL or bibliographic reference to the source. Failure to quote and reference copied material is considered. If you do not get 100% correct on an assignment you will be given a description of what you did wrong and you will then have one week to submit a revised homework for full credit. The final paper may only be submitted once.

Lecture Schedule

|  |  |
| --- | --- |
| Lecture | Topic |
| 1 | Introduction and literature search |
| 2 | Human Genome |
| 3 | Genome and sequence database |
| 4 | Protein sequence and motif databas |
| 5 | Unix |
| 6 | Perl |
| 7 | Sequence alignment |
| 8 | Rapid similarity search |
| 9 | Multiple alignment |
| 10 | Molecule phylogenetics |
| 11 | Clustering |
| 12 | Population evolution |
| 13 | Protein structure and folding |