Faculty of Biological, Agronomic and Environmental Engineering

| BRMC2201 | Bioinformatics : DNA and protein sequences |

[30h+7.5h exercises] 3 credits

This course is taught in the 1st semester

Teacher(s): Michel Ghislain
Language: French
Level: Second cycle

Aims

This course in bioinformatics is for students who want to understand the methods used in sequence and genome analysis. The first objective is to help students appreciate the underlying algorithms used and assumptions made, as well as limitations of the predictions obtained. The second objective is to provide students with theoretical and practical knowledge of software commonly used in molecular biology. Thus, the students will be prepared to work in a research laboratory and to apply their knowledge of computational molecular biology to biological problems in academia or industry

Main themes

The theoretical themes are 1) the definition of the commonly used terms in computer analysis of biological information; 2) the collection and storage of information in a sequence database and retrieval by a database query program, via the World Wide Web; 3) simple types of sequence analysis, e.g., searching for restriction sites, translating sequences and compositional analysis; 4) the methods of sequence alignment used to compare two or more sequences in order to identify conserved regions, or motives, and to determine the phylogenetic relationships; 5) the methods used to predict protein's hydropathy and structure. Data interpretation and their statistical significance are examined carefully. The choice of the best method is discussed depending on the type of the results foreseen; experimental approaches that are suitable to test the predictions are discussed as well. Solved problems provide students with a practical knowledge of the bioinformatics tools used in molecular biology.

Content and teaching methods

The course begins with an overview of molecular biology databases and how to use them on the World Wide Web. This introduction is followed by a description of the tools for sequence comparison (dot matrix plot) and alignment (dynamic programming and hashing algorithms). The next chapter provides detailed descriptions of methods for homology searching of DNA and protein sequence databases using the BLAST program. The basic algorithm is compared to another common program, FASTA. Multiple sequence alignment allows the identification of conserved motifs or domains. The method of progressive alignment is illustrated with the CLUSTAL program, and compared to other methods known as profiles or hidden Markov models. One chapter is devoted to the description of the methods used to estimate evolutionary relationships between multiple sequences. Tree constructions based on parcimony, distance and maximal likelihood are explained. Finally, the last chapter provides methods of predictions of protein's hydropathy, tertiary structure and subcellular localization. The course includes many practical tutoring sessions. Students are evaluated for their ability to apply their knowledge to a real problem in molecular biology.

Other information (prerequisite, evaluation (assessment methods), course materials recommended readings, ...)

Precursory courses Elements of biochemistry and molecular biology
Evaluation Students' ability to apply their knowledge to a real problem in molecular biology
Support Powerpoint files and notes from the teacher
Teaching team The course includes many practical classes

Other credits in programs

BIR23/2C	Troisième année du programme conduisant au grade de bio-ingénieur : Chimie et bio-industries (Ingénierie biomoléculaire et cellulaire)	(3 credits)	Mandatory
SC3DA/B	Diplôme d'études approfondies en sciences (Biologie)	(3 credits)	