UCLouvain

lbrmc2201 2023	Bioinformatics : DNA and protein sequence analysis

4.00 credits	30.0 h + 15.0 h	Q1

Teacher(s)	Ghislain Michel ;
Language :	English > French-friendly
Place of the course	Louvain-la-Neuve
Prerequisites	Introductory courses in biochemistry and molecular biology
Main themes	Bioinformatics refers to a set of concepts and tools that are required for the analysis of biological data and the interpretation of the results. This introductory course focuses on molecular biology databases (DNA and protein sequences), the algorithmic bases of the sequence analysis programs and on alignment score statistics. The course identifies the many pitfalls of interpreting data by giving a critical appraisal of the softwares used for sequence analysis.
Learning outcomes	**At the end of this learning unit, the student is able to :** 1 a. <u>Contribution de l'activité au référentiel AA (AA du programme)</u> Cohérence des AA cours en regard de ceux du programme 1.1, 1.2, 1.3 3.1, 3.2, 3.4, 3.5, 3.6 b. <u>Formulation spécifique pour cette activité des AA du programme (**maximum 10**)</u> At the end of this course, students will be able to perform a comprehensive and exhaustive sequence analysis, using appropriate computational programs tools and internet resources. This ability requires: - The understanding of the algorithmic bases of the computational programs - The description of the various molecular databases with emphasis on the positive and negative aspects of data structure and search tools - The discussion of the prediction results and eventually the proposition of a more appropriate analysis method -A strategy for protein function forecasting
Evaluation methods	Criteria including (i) the understanding of the algorithmic bases of the sequence analysis programs, (ii) the use of the most appropriate program and database, and (iii) the explanation of the statistical bases of prediction scores will be assessed via a written examination in an open-book format (70% of the final mark). Practical training is assessed during the quadrimester, via problems to solve individually or by a group of 2-3 students (30% of the final mark). Students can choose English or French for the written exam and the resolution of exercises
Teaching methods	The theoretical part consists of *ex cathedra* speeches in a classroom (30h). The training sessions (15h) consist of a set of problems to be resolved individually or by a group of 2 students, using free sequence analysis programs.
Content	1. Introduction : Overview of bioinformatics concepts 2. Sequence and 3-D structure databases, protein motif and family databases 3. Sequence comparison : dot plot, global and local alignment based on a dynamic programming method and score matrices 4. Database searching for similar sequences (matching word-based method), score statistics 5. Multiple sequence alignment, motif discovery,(patterns, profiles, Hidden Markov models) 6. Phylogenetic inference using phenetic and cladistic methods 7. High-throughput analysis of gene expression: RNA-seq
Inline resources	Moodle
Bibliography	Des copies papier des diaporamas et le manuel pour les exercices sont disponibles sur Moodle. Le cours ne fait appel à aucun support particulier qui serait payant et jugé obligatoire. Les ouvrages Bioinformatics de Mount (CSHL press) et Bioinformatics and functional genomics de Pevsner sont conseillés pour un apprentissage plus approfondi
Other infos	The lectures are given in English.

Faculty or entity in charge	AGRO

Programmes containing this learning unit (UE)				
Program title	Acronym	Credits	Prerequisite	Learning outcomes
Master [120] in Biochemistry and Molecular and Cell Biology	BBMC2M	4		🔍
Master [120] in Statistics: Biostatistics	BSTA2M	4		🔍
Master [60] in Biology	BIOL2M1	4		🔍
Master [120] in Chemistry and Bioindustries	BIRC2M	4		🔍