

SYLLABUS

Name: Principles of bioinformatics (25-BI-S2-W-PB-AN)

Name in Polish: Podstawy bioinformatyki

Name in English: Principles of bioinformatics

Information on course:

Course offered by department: Faculty of Biological Sciences

Course for department: Faculty of Biological Sciences

Default type of course examination report:

Grading

Language:

English

Short description:

Prerequisites regarding knowledge, skills, and social competences for the course/module:

Basic knowledge in genetics and computer sciences

Student's own work:

- reading scientific publications: 15h
- preparing homework assignments: 15h
- preparing for tests: 15h

Description:

Educational aims:

Transfer of knowledge in the field of theoretical and practical foundations of various bioinformatics research at all levels of biological information organization, including recognition of coding sequences and other nucleotide sequence analyses, protein sequence analyses, sequence alignment, searching for similar sequences in databases, phylogenetic analyses; basic databases related to bioinformatics and genomics.

Course content:

- Lectures:
- subject and levels of bioinformatics analysis. Basic terms related to bioinformatics.
- types of databases. Basic databases of nucleotide and amino acid sequences: GenBank, EMBL, UniProt. NCBI database. Errors in databases.
- computer identification of protein coding sequences: search for Open Reading Frames, search for genes in Prokaryota and Eukaryota, methods of gene recognition based on composition, signals, similarity to other sequences. Integrated gene search methods (discriminant analysis, dynamic programming, neural networks, Markov models). Problems and pitfalls in the search for genes. Examples of gene search programs. Evaluation of gene recognition algorithms. Computer analysis of RNA sequences, prediction of secondary structure.
- pairwise alignment: definition, application, types of methods (dot matrix, dynamic programming - local and global alignment), scoring systems (PAM and BLOSUM amino acid matrices, gap penalties), evaluation of alignment significance, examples of programs.
- multiple sequence alignment: definition, application, types of algorithms and programs: progressive global alignment.
- searching for similar sequences in databases: basic terminology and goals, search methods, types of programs and algorithms (FASTA, BLAST), criteria and significance of sequence similarity, types of programs from the BLAST family and their options, e.g. MegaBLAST, PSI-BLAST, PHI-BLAST.
- computer analysis of protein sequences: translation of nucleotide sequences into amino acids, analysis of basic physicochemical properties of proteins, search for transmembrane regions, post-translational modifications of proteins, determination of subcellular localization of proteins, search for motifs and domains, determination of protein secondary structure. Spatial structure databases (PDB), structural comparison, methods for predicting spatial structures, structural classification of proteins, programs for viewing spatial structures.
- molecular phylogenetics: molecular and classical phylogenetics, a brief history, the molecular basis of evolution, types of mutations, stages of phylogenetic analysis (gathering of homologous sequences, alignment of sequences, determination of the substitution model, construction of a phylogenetic tree, analysis and evaluation of the phylogenetic tree), types of homologous sequences (orthologs, paralogs, xenologs), types of substitution models, variability of substitution rates within sequences and between different sequences, molecular clock, types of trees, methods of creating phylogenetic trees, bootstrap method.
- Classes:
- searching for information contained in literature databases as well as databases of nucleotide and amino acid sequences.
- basic bioinformatic analyses of nucleotide sequences: composition analysis, codon usage determination. Search for protein coding sequences in prokaryotic and eukaryotic genomes. Computer analysis of RNA sequences.
- computer analysis of protein sequences: translation of nucleotide sequences into amino acids, analysis of basic physicochemical properties of proteins, computer digestion of proteins, search for transmembrane regions, post-translational modifications of proteins, determination of subcellular localization of proteins, search for motifs and domains in proteins, determination of secondary and tertiary structure of proteins. PDB database.
- searching for similar sequences in databases using the BLAST algorithm.
- sequence pair alignment: matrix of points, global and local alignment. Aligning of multiple sequences. Editing of an alignment.
- gathering a set of sequences to create phylogenetic trees. Determining the substitution model. Constructing phylogenetic trees. Estimating the significance of tree branches.

Bibliography:

Mandatory and recommended reading list:

- A.D. Baxevanis, B.F.F. Ouellette (red.), Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins. Wiley-Liss, Inc. 2005
J.-M. Claverie, C. Notredame, Bioinformatics For Dummies. Wiley Publishing, Inc. 2006.
P.G. Higgins, T.K. Attwood, Bioinformatics and molecular evolution, Wiley Online Books, 2004
B. Hall, Phylogenetic Trees Made Easy: A How-to Manual. Oxford University Press, 2017
X. Jin, Essential Bioinformatics, Cambridge University Press, 2006

Learning outcomes:

Intended learning outcomes

Student:

- in terms of knowledge:

<p>K_W02 knows and understands what objects and levels of organization of biological information are the subject of bioinformatics analyses</p> <p>K_W05 knows and understands the basic and/or specialized methods and algorithms used in the bioinformatics of nucleic acids and proteins</p> <p>K_W08 knows and understands databases and computer programs used in bioinformatics analyses</p> <p>- in terms of skills: K_U05 is able to use appropriate programs and databases and set parameters in bioinformatics analyses</p> <p>K_U07 is able to interpret the results of bioinformatics analyzes from the point of view of the analyzed objects (sequences)</p> <p>- in terms of social competences: K_K07 is ready to cope with increasingly complex bioinformatics analyzes and combine partial results into a single whole</p>
<p>Assessment methods and assessment criteria:</p> <p>Assessment methods for the intended learning outcomes: Lectures: - written test - minimum 50% of points; Classes: - execution of assessments, partial tests - minimum 50% of points.</p> <p>Credit requirements for individual components of the course/module: Lectures: - written test - minimum 50% of points to pass; Classes: - attendance control, assessments, partial tests - minimum 50% of points to pass.</p>

Course credits in various terms:

<without a specific program>			
Type of credits	Number	First term	Last term
European Credit Transfer System (ECTS)	3	2023/24-Z	