

Schedule for Lectures and Practicals in Bioinformatics - 2017

1. What does bioinformatics mean?

12 Sep

Introduction to the course: When do we need computers for biological research? The meaning of bioinformatics. The history of bioinformatics. Bioinformatics applications. Trends in biological research. Bioinformatics on the web. Most used bioinformatic methods and softwares. Introduction to Linux.

Dávid Fazekas

Practical: Introduction of computer usage in biological research

The schedule and the requirements of the practicals. Introduction of the computer room. Introduction of Linux operating system: graphical interface, text editor, how to start a software or a command. Usage of simple Linux commands to solve bioinformatic problems.

Eszter Ari

2. Databases in molecular biology

19 Sep

The structure and content of databases. Relational databases. How to apply molecular biology databases in biological research? Nucleotide (EMBL, GenBank) and protein databases (PIR, UniProt, (SWISS-PROT, TrEMBL)). Entry formats of sequence databases. ID mapping. Other databases (Gene Ontology, Tree of life).

Dávid Fazekas

Practical: Application of molecular biology databases

Introduction to the usage of NCBI databases and the Uniprot protein database. How to search in databases? Searching in Uniprot database: Retrieve, ID mapping.

Eszter Ari

3. Sequence comparison and alignment

26 Sep

Clarification of conceptions: sequence similarity vs. homology. Pairwise sequence alignments, pairwise comparisons: dot-plots. Scoring system: substitution matrices: PAM, BLOSUM. Optimal global and local alignments, dynamic programming algorithm. Gap-penalties. Multiple sequence alignment methods: progressive alignment, ClustalW.

Eszter Ari

Practical: Basics of sequence analyses, sequence alignment

Manipulation of sequences, different sequence file formats. Introduction to the usage of the UGENE software. Using UGENE for database searching and downloading sequences. Analysis of sequences using UGENE: Dot-plots, pairwise and multiple alignments.

Eszter Ari

4. Sequence similarity searching

3 Oct

Heuristic algorithm for similarity searching: BLAST software. Statistical bases: evaluate the significance of hits. Filtering false positive hits: low complexity and repetitive regions. Different BLAST versions and other similarity searching softwares.

Dávid Fazekas

Practical: Searching sequences in databases based on their similarity

Similarity searches in online sequence databases. Using the online version of BLAST softwares on NCBI webpage. Evaluating the results. Downloading, aligning and editing the hit sequences.

Eszter Ari

5. Molecular phylogenetics I

10 Oct

What is phylogenetics? How can we apply it? Phylogenetic signal, phylogenetic trees. Prediction of phylogenetic changes and distances, nucleotide substitution models. Methods for phylogenetic

reconstruction. Distance based methods: neighbor-joining. Character based methods: maximum parsimony. Consensus trees. Statistical tests: bootstrap analysis.

Eszter Ari

Practical: Molecular phylogenetics in practice I

Introduction to the usage of PHYLIP software package. Calculating maximum parsimony tree using PHYLIP. Reconstructing distance based phylogenetic tree with UGENE. Tree visualization using the Figtree software.

Eszter Ari

6. Molecular phylogenetics II

17 Oct

Theoretical bases of the maximum likelihood method. Introduction to Bayesian statistics and Markov chain Monte Carlo. How to use these to reconstruct a phylogenetic tree? How to choose the best nucleotide substitution model?

Eszter Ari

Practical: Molecular phylogenetics in practice II

Choosing the nucleotide substitution model. Calculating phylogenetic trees with maximum likelihood and Bayes statistic methods using the UGENE software as interface. Downloading taxonomical trees from NCBI Taxonomy database. Comparing the calculated trees with the taxonomic tree.

Eszter Ari

7. Genomics and transcriptomics I

24 Oct

The history of sequencing. Genome sequencing methods. Analysis of high-throughput sequencing data with bioinformatic methods. Genome assembly, mapping reads to reference genome, annotation, SNP analysis.

Practical: Genome databases

Introduction to the Ensembl genome database. What is Biomart and how to use it?

Eszter Ari

fall break

30 Oct – 3 Nov

8. Genomics and transcriptomics II

7 Nov

Introduction to the microarray technique. Analysis and possible applications of microarray data. Introduction to the RNA-seq method and data analysis. *De novo* assembly, mapping reads to reference genome, differential expression analysis. Functional enrichment analysis.

Eszter Ari

Practical: Differential expression analysis of microarray and RNA-seq data

Introduction to the NCBI GEO database. Differential expression analysis. Functional enrichment analysis using the GeneOntology database.

Eszter Ari

9. Network and systemsbiology I + Mid term test I (Lecture: 1-6)

14 Nov

Introduction to network and systemsbiology. What kind of data can be used? Bioinformatic approaches in network and systemsbiology.

Dávid Fazekas

Practical: Cytoscape I

Introduction to the usage of Cytoscape network analyzer and visualization software.

Eszter Ari

10. Network and systemsbiology II

21 Nov

Introduction to the “world of networks”: topology, hierarchy and dynamics. Different types and sources of molecular networks. Creation of protein-protein interaction (PPI) networks, high-throughput methods, their advantages and disadvantages.

Dávid Fazekas

Practical: Cytoscape II

Creating a tissue specific network using the Cytoscape software.

Eszter Ari

11. Network and systemsbiology III

28 Nov

Molecular networks of the cell. Investigating network dynamics using bioinformatic methods.

Dávid Fazekas

Practical: Cytoscape III.

Simulation of a dynamic molecular network.

Eszter Ari

12. Structural bioinformatics

5 Dec

Structural identification methods. The usage and structure of the Protein Data Bank (PDB). The PDB format. Protein structures as models. Analysis of the quality of experimentally solved structures. Secondary structure elements, and domains in three dimensional structures. Basics of structural classifications, hierarchical databases. Comparison of structures, algorithms and their usage.

Zsuzsa Dosztányi

Practical: Bioinformatic analysis of protein structures

Familiarizing with the PDB database: finding structures, additional information, and downloading. Visualizing protein structures. Highlighting regions, side chains, secondary structure elements. Structural comparison.

Zsuzsa Dosztányi

13. Mid term test 2 (Lecture 7-12)

12 Dec

Practical: Consultation about the essay tasks

Eszter Ari, Dávid Fazekas

Evaluation

Evaluation of the theoretical part (lecture grade):

The average grade of the 2 mid term tests (written, 45 min.).

You have to take an oral exam during the examination period if one (or both) mid term tests resulted with mark 1 or you missed the mid term exam. Oral exam will be about that part in which you failed or missed. You also can improve your mark (what you got as an average) if you wish in an oral exam – but be aware that you can also decrease the mark.

x.5 averages will be rounded upwards.

Evaluation of the practical part (term grade):

Criterion: Being active on the practicals. You can miss maximum 3 practicals.

Write a project work essay in groups of 3 people (using 3 different proteins) at the end of the semester using the results and the knowledge you have learned during the practicals.