

Bioinformatics exam items 2019

(The numbering is based on the order of lectures.)

1: 2-6

- 2.a) Review the main DNA sequence databases.
- 2.b) Review the *Uniprot* database, describe its structure, the source and quality of its data.
- 3.a) What you can conclude about the similarity of two sequences? What are the aims of sequence alignment? What does global and local alignment mean? What does gap penalty mean?
- 3.b) How to construct a dot-plot figure? What information can we learn about the sequences using this technique? What are the amino acid substitution matrices and how these can be applied?
- 4.) Review the *BLAST* similarity searching method.
- 5.a) What are the basic features of protein structures? Describe the different levels of protein structure. What kind of forces stabilize proteins structures? What is the Ramachandram plot and what does it characterize?
- 5.b) What are the methods used for structure determination and how do they work? What is the database for protein structures?
- 6.a) What computational methods can you use to characterize your protein if it does not have a known structure? What are secondary structure prediction methods and how do they work?
- 6.b) What are the main steps of homology modelling? How can you evaluate the quality of the predictions?

2: 7-12

- 7.a) Review different nucleotide substitution types? Why do we apply nucleotide substitution models? Review the distance based phylogenetic methods.
- 7.b) Review the Maximum Parsimony phylogenetic method.
- 7.c) Review the different phylogenetic tree shapes and topologies? How would you calculate a consensus tree? Review the bootstrap analysis used in phylogenetic researches.
- 8.a) Review the *Next Generation Sequencing* methods. What does genome *assembly* and *annotation* mean? Review the analysis of *Next Generation Sequencing* data.
- 9.a) Review the *microarray* technique and the analysis of expression array data.
- 9.b) Review the RNA-Seq technique the analysis of RNA-Seq expression data.
- 10.a) Network and graph definitions: network properties

10.b) Network types, file formats and Cytoscape

11.a) Small and large scale study comparison (methods)

11.b) Network stability and data structures

12.a) Haplotype: definition, the relationship of haplotype with chromosome and gene. The haplotypes of ABO gene.

12.b) Haplotype reconstruction: algorithms, performance and impact.