

Bioinformatics exam items 2017

(Numbering is based on the order of lectures.)

1: 2-7

- 2.a) Review the main bioinformatics *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 the *dot-plot* figure is made? What informations we can 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 different *nucleotide substitution* types do exist? Why do we apply nucleotide substitution models? Review the *distance* based phylogenetic method.
- 5.b) Review the *Maximum Parsimony* phylogenetic method.
- 5.c) What different phylogenetic tree shapes and topologies you know? How to calculate a *consensus tree*? Review the *bootstrap* analysis used in phylogenetic researches.
6. Review the Maximum Likelihood and Bayesian inference phylogenetic methods.

2: 8-12

- 8.a) Review the *Next Generation Sequencing* methods. What does genome *assembly* and *annotation* mean? Review the analysis of *Next Generation Sequencing* data.
- 8.b) Review the *microarray* technique and the analysis of expression array data.
- 8.c) Review the RNA-Seq technique the analysis of RNA-Seq expression data.
- 9.a) Review the history and the basics of graph theory, the type of graphs, and network science.
- 9.b) Review a variety of network theoretical metrics (degree, degree distribution, shortest path, diameter), and classify the networks by degree of distribution.
- 10.a) Review molecular networks (nodes, interactions, origins of relationships, opportunities to collect).
- 10.b) Introduce the methods of detection of protein-protein interactions.
- 11.a) Review a probabilistic and context-specific networks and differences between them and types of nodes.

11.b) Review the robustness of scale-free networks.

12.a) How can the structure of proteins be determined? Where and in what form can you access the already defined structures?

12.b) How can secondary structures be assigned to proteins? What does the structure fit and cover? How can structural domains be identified?