NTNU – Trondheim
Norwegian University of
Science and Technology

Department of Biology

Examination paper for BI3019
Systems Biology: Resources, standards and tools

Academic contact during examination: Martin Kuiper

Phone: 91897322

Examination date: 25 May 2016

Examination time (from-to): 9:00 – 13:00

Permitted examination support material: Dictionary of English

Other information: Credits 7.5

Language: English

Number of pages: 3

Checked by:

_____

Date Signature

Exam Questions BI3019 – 25 May 2016 - Systems Biology

Your answers should typically be at least ½ page of written text.
You may skip one of the 13 questions (optional)

1. Systems biology is often based on some sort of model of a system. Describe two different types of such models (static and dynamic), how they can be constructed and how they can be used to come to new insights about experimental results and pathway behaviour. Explain which approach you would use if you are interested in studying a regulatory pathway related to the development of for instance cancer (multiple good answers are possible here, it is your argumentation that counts).

2. Describe what is meant with the name 'interactome'; describe how interactome information can be generated: mention experimental designs and approaches as well as ways to describe such information (e.g. in the form of tables, and what the role of ontologies and controlled vocabularies could be for this). Explain how this type of information can be integrated with other 'omics' data.

3. Describe the UniProt database, explain what its purpose is, how it is built, what different parts it has (e.g. with respect to its curation and annotation), and discuss some of the different types of information that it contains.

4. The Gene Ontology (GO) is important for the analysis of experimental results, for instance a set of regulated genes. Describe the following aspects of the GO: which specific ontology branches does it have, and what do they describe; explain the hierarchical structure of the ontology and how this can be used for 'inferencing'; and describe how you can used GO to shed light on your experimental results.

5. You have found 5 genes that seem to be highly regulated in your experiment. Explain at least 3 approaches using internet tools and resources to find more information about these genes, how they would be connected and how you could include other published experimental data to check your results.

6. You are interested in the human RB1 gene, describe the steps that you would take using the iHOP web tool to find out more about its function, and how you would use iHOP to assemble a small network of RB1 together with several other genes that it is related to, in a step-by-step approach.

7. Describe the MIAME checklist, explain what it was developed for, and how it paved the way for a large series of similar checklists. Discuss also the MIBBI initiative and how checklists in general are important for providing more value to data that is uploaded to public databases.

8. You want to download a large protein-protein interaction network from IntAct. Describe how you would do that, what selection criteria you could use, and how the NetworkAnalyzer tool in Cytoscape can be used to perform a graph-based analysis of this network. Explain some of the results that you would expect to see in an analysis using NetworkAnalyzer, focusing on network structure and node connectivity. You may use drawings in your explanation.

9. Describe the GeneMANIA resource, describe what types of information it contains, and how you may use it to get more insight into a particular experimental result, for instance using a hypothetical use case in which you have a gene set of 50 genes.

10. High throughput technologies are excellent for producing large amounts of data, but the data that they produce is often 'noisy', meaning that they not only concern biological information but also include random signals: noise. Describe this noise for a transcriptome experiment and a protein interaction experiment. Suggest refinements of analysis approaches that you could use to become more confident about the data. In your answer you may mention reliability, confidence, sensitivity and false positives/false negatives.

11. Compare Boolean modelling and Ordinary Differential Equation modelling with respect to various strengths and weaknesses: for instance the amount of data that you need, the type of data, the size of the models that you can use these for and provide arguments concerning which of these approaches you would chose if you would have a network map with 50 components (multiple good answers are possible here, it is your argumentation that counts).

12. You have performed a microarray experiment that resulted in a list of 100 significantly regulated genes. Provide a step-by-step description how you can find evidence for overrepresentation of specific gene classes in Cytoscape using BiNGO, and discuss the important statistical decisions to consider in this analysis.

13. Provide brief descriptions of the REACTOME knowledge base, the KEGG knowledge base and the ENSEMBL genome browser. You may consider for instance the type of knowledge in these resources, the functionality provided and the integration of the knowledge bases with other resources.