UNIVERSITY
OF SKÖVDE

School of Bioscience

WRITTEN EXAMINATION

Course: Bioinformatics – Concepts and Methods

Examination: Module 5

Course code: BI760A

Credits for written examination: 1.5

Date: 26 March 2025

Examination time: 2 hours

Examination responsible: Zelmina Lubovac

Teachers concerned: Zelmina Lubovac

Aid at the exam/appendices: None

Other

Instructions

- ☐ Take a new sheet of paper for each teacher.
- ☒ Take a new sheet of paper when starting a new question.
- ☒ Write only on one side of the paper.
- ☒ Write your name and personal ID No. on all pages you hand in.
- ☒ Use page numbering.
- ☒ Don´t use a red pen.
- ☒ Mark answered questions with a cross on the cover sheet.

Grade points: 0-15 = F; 16-18 = E; 19-21 = D; 22-24 = C; 25-27 = B; 28-30 = A

Examination results should be made public within 18 working days

Good luck!

Total number of pages: 6

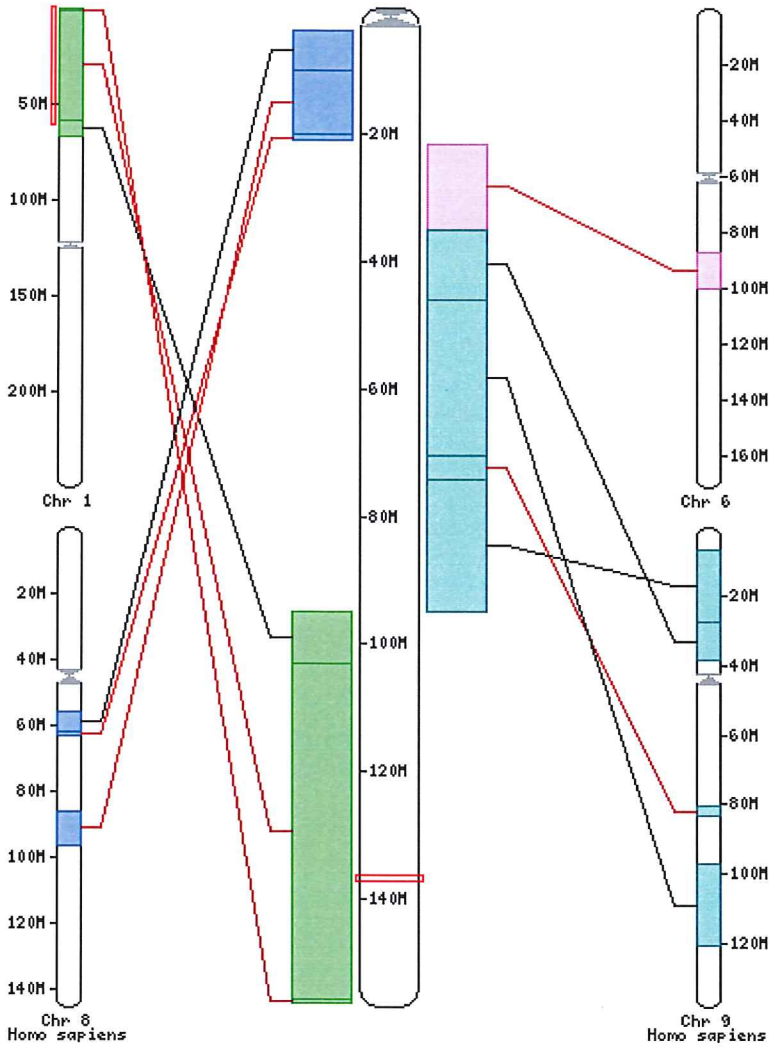
UNIVERSITY
OF SKÖVDE

Question 1 (6p)

The figure below shows a synteny map from Ensembl. The synteny map was generated by finding the gene entry for the mouse EPHA8 gene, which is located on chromosome 4, and then creating a synteny map for the region where EPHA8 is located. The comparison species is human, *Homo Sapiens*.

1a) Define the term "synteny" and explain its importance in comparative genomics. What can we learn from studying synteny between different genomes? **(3p)**

1b) Analyze the synteny map between mouse chromosome 4 and human. Your analysis should identify which chromosomes in the human genome show synteny with the different regions of mouse chromosome 4. Additionally, mention the probable location of the human ortholog of the mouse EPHA8 gene, based on the synteny map provided. **(3p)**



UNIVERSITY
OF SKÖVDE

Question 2 (6p)

For each of the following claims about Ensembl, state if the claim is true or false. You do not need to give any motivations in your answer, just writing (for each claim) "true" or "false" is sufficient.

2a) Ensembl only hosts vertebrate genomes. **(1p)**

2b) The Ensembl Variant Effect Predictor (VEP) cannot be used to analyze variants in non-human species. **(1p)**

2c) Information regarding single-nucleotide polymorphisms (SNPs) is available through Ensembl. **(1p)**

2d) BioMart is a tool within Ensembl that allows users to perform custom queries and export data. **(1p)**

2e) The only way to access the Ensembl database is through the Ensembl genome browser interface. **(1p)**

2f) Ensembl provides access to multiple types of entries, such as gene, transcript, and protein entries. **(1p)**

UNIVERSITY
OF SKÖVDE

Question 3 (6p)

ORF Finder is a tool used to identify potential Open Reading Frames (ORFs) within a nucleotide sequence.

For each of the following claims related to gene prediction, state if the claim is true or false.

3a) During ORF Finder-based ORF detection, any intermediate start codons are disregarded when searching for the largest possible open reading frames. **(1p)**

3b) tRNA molecules exhibit highly flexible secondary structures, with no fixed rules for stem and loop dimensions. **(1p)**

3c) Ab initio gene prediction relies only on the genome sequence without any additional information. **(1p)**

Following question requires a well-developed answer.

3d) Describe how you could confirm an ORF prediction made by ORF Finder using additional tools. Your answer should include examples of two approaches or tools that can be used to strengthen the prediction from ORF Finder. The accuracy and level of detail in your explanation will influence the points awarded. **(3p)**

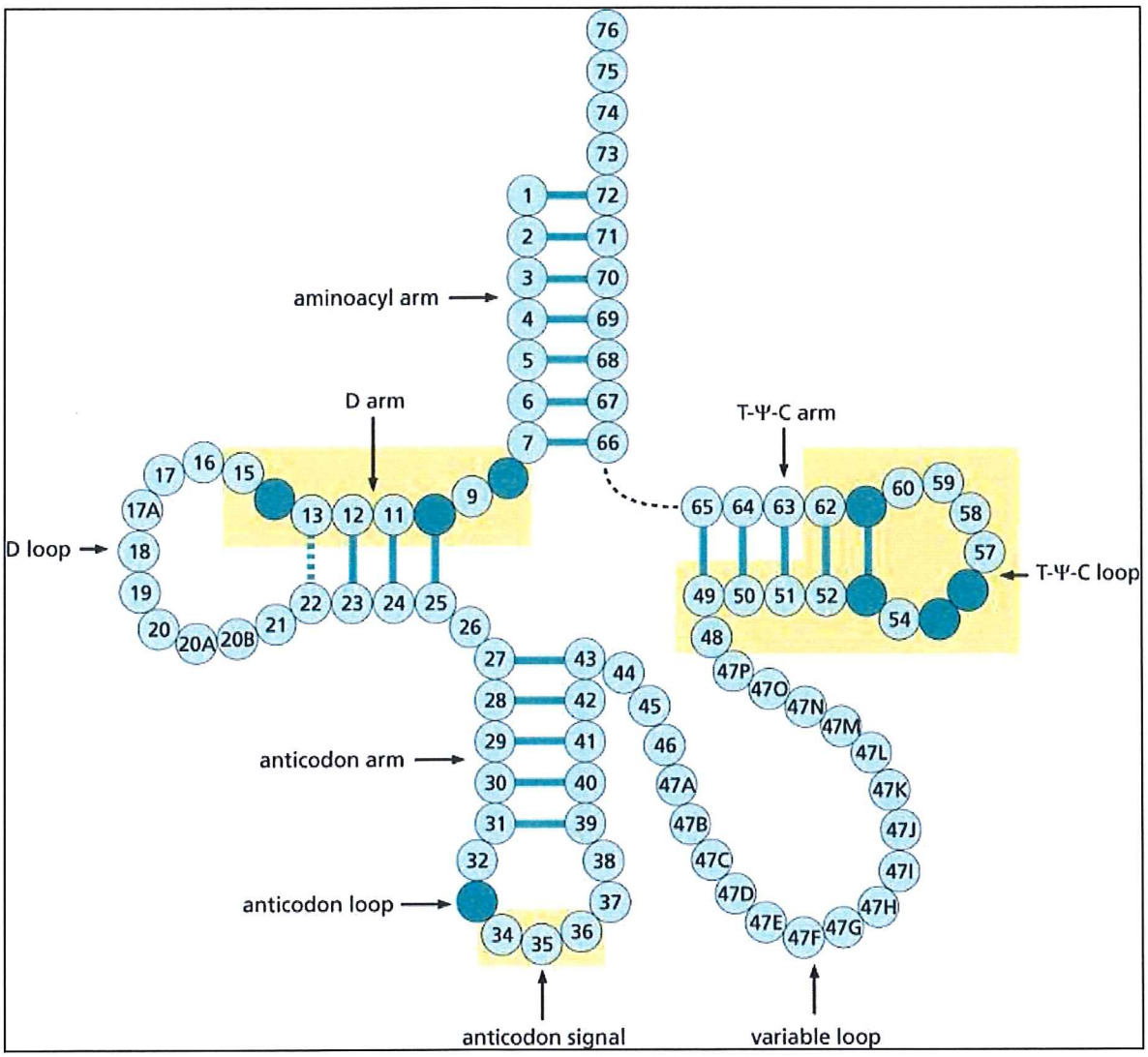
UNIVERSITY
OF SKÖVDE

Question 4 (6p)

The tRNAscan algorithm predicts tRNA genes using a rule-based approach.

4a) Describe **three** rules that tRNAscan applies to determine whether the sequence is a tRNA gene or not. Refer to the figure. (**3p**)

4b) In prokaryotes, gene prediction is typically more straightforward than in eukaryotes. Identify at least **three** major differences in genome organization or structure that account for this reduced complexity. (**3p**)



UNIVERSITY
OF SKÖVDE

Question 5 (6p)

5a) Describe what the figure shows and how it represents transitions between nucleotides. Your answer should include the description of key components of the model, the role of arrows and the concept of transition probabilities. (**3p**)

5b) Explain at least one application of such a model in computational genomics, and give an example of the tool/software that utilizes it. **(3p)**