22CGP086

Fundamentals of Biotechnology and Genetic Engineering

Semester 1 2022/23 In-Person Exam paper

This examination is to take place in-person at a central University venue under exam conditions. The standard length of time for this paper is **2 hours**.

You will not be able to leave the exam hall for the first 30 or final 15 minutes of your exam. Your invigilator will collect your exam paper when you have finished.

Help during the exam

Invigilators are not able to answer queries about the content of your exam paper. Instead, please make a note of your query in your answer script to be considered during the marking process.

If you feel unwell, please raise your hand so that an invigilator can assist you.

You may use a calculator for this exam. It must comply with the University's Calculator Policy for In-Person exams, in particular that it must not be able to transmit or receive information (e.g. mobile devices and smart watches are **not** allowed).

Answer **THREE** questions in total.

Each question carries 25 marks.

Candidates should show full working for calculations and derivations.

1. (a) Compare and contrast cellular catabolic and anabolic pathways. Give two examples of catabolic and anabolic pathways. [5 marks]

   (b) Explain why microorganisms that cannot use glucose as a growth substrate still require glucose. How do these organisms biosynthesise glucose if they are using C2 (acetate) or C5 (xylose) compounds as their growth substrates? [5 marks]

   (c) Evaluate the working principle of NAD^+/NADH-redox couple with the help of a schematic diagram. [6 marks]

   (d) If three bacteria, A, B, and C, were competing for glucose to use as the electron donor, and A was capable only of aerobic respiration, B anaerobic respiration with nitrate (NO_3^-) as the electron acceptor, and C anaerobic respiration with fumarate as the electron acceptor, which bacterium/ bacteria do you think would be most competitive in an oxic and anoxic environment? Provide adequate explanations to justify your answer. [9 marks]

   Use the following information if necessary:

   Standard free energy change of a redox reaction can be estimated using the equation:

$$\Delta G_0' = -nF\Delta E_0'$$

   where the symbols have their usual meanings and F = 96.5 $kJ.V^{-1}.mol^{-1}$.

Continued/…

Q1 Continued/…

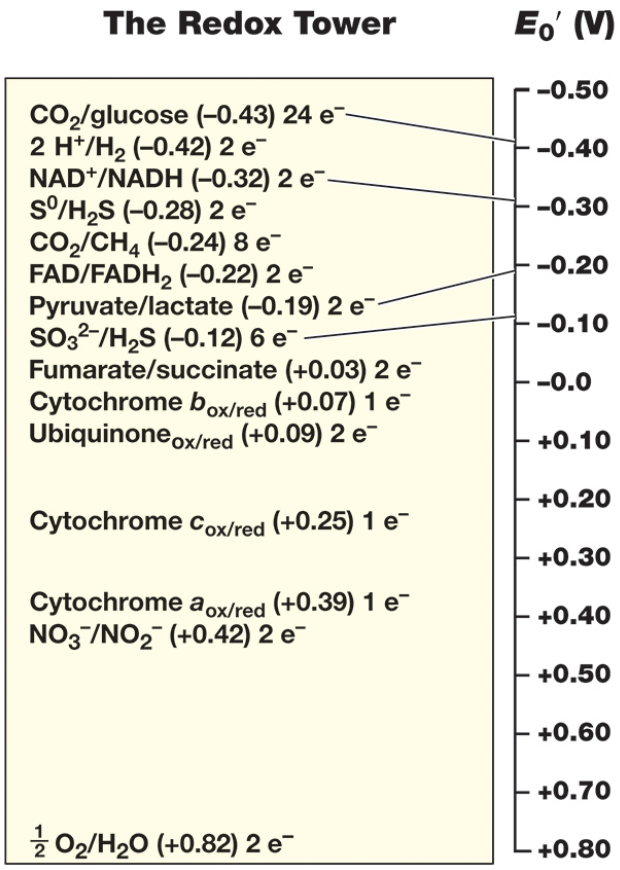


Figure Q1(d). Standard reduction potentials of some biologically important redox couples.

2. (a) Name the key molecules essential for Sanger sequencing. Provide different steps of the Sanger sequencing method with a schematic diagram. [7 marks]

(b) Define an open reading frame (ORF) and its typical structure. Explain the steps involved in the computational identification of possible ORFs from raw genome sequencing data. [7 marks]

(c) How do 20,000 protein-coding genes in the human genome help make roughly 200,000 to 1 million proteins by human cells? Analyse the two main differentiating factors contributing to the existence of 0.01% difference between the genomes of individuals of all nationalities. [6 marks]

(d) As a biotechnology researcher, you are tasked to identify and annotate the genes coding for noncoding RNAs (ncRNAs) in a newly sequenced genome. Identify and analyse the most abundant ncRNAs found in a typical genome. Briefly explain the ‘omics’ technology that would be required to identify and annotate ncRNAs in the sequenced genome. [5 marks]

3. The incidence of Type 1 diabetes is increasing globally, but the reasons are incompletely understood. The condition can lead to serious long-term complications and can have significant quality of life implications. In the UK, which has one of the highest rates in the world, the incidence of Type 1 diabetes is rising by 5% annually in children under five. In Type 1 diabetes, the immune system attacks and destroys the β cells in the pancreas that make the hormone insulin. Insulin is normally produced in response to an increase in blood glucose levels, but patients with Type 1 diabetes do not produce enough insulin and must replace it with insulin injections, typically more than 65,000 times within their lifetime. Genetic engineering techniques have advanced the production of artificial insulin, where previously insulin was extracted from animals such as pigs.

   (a) Compare and contrast why genetically engineered artificial insulin production is preferable to animal insulin. [4 marks]

   You are tasked with making artificial insulin from mRNA containing the insulin gene that was extracted from healthy β cells in the pancreas. You have been provided with a sample of extracted mRNA.

   (b) Design your process to genetically engineer artificial insulin:

      (i) Briefly explain your process stages. [5 marks]

      (ii) Justify your selections for amplification of the gene. [7 marks]

      (iii) Justify your selections for the type of vector delivery of the gene to the host cell and the type of host cell for production. [3 marks]

      (iv) Where are the critical points of analysis within the design of your process to genetically engineer artificial insulin? Briefly justify your answer. [6 marks]

4. (a) What are the contents of primary biological databases? Using three primary databases as examples, explain if these databases are sufficient to annotate a newly sequenced genome. [5 marks]

(b) Define the unique identifier that is required to retrieve information about a gene or protein from any biological databases. Categorise the following identifiers into ‘Gene’, ‘mRNA’, ‘RNA’, and ‘Protein’ molecules: [6 marks]

NC_000011, XP_015315799, NR_110467, XM_015460313, AC_000182, NP_000509, NM_203528, XR_002004952

(c) Illustrate the mechanisms behind the creation of orthologous and paralogous genes. Evaluate ‘orthologs’, ‘paralogs’, ‘in-paralogs’, and ‘out-paralogs’ in Figure Q4(c). [6 marks]

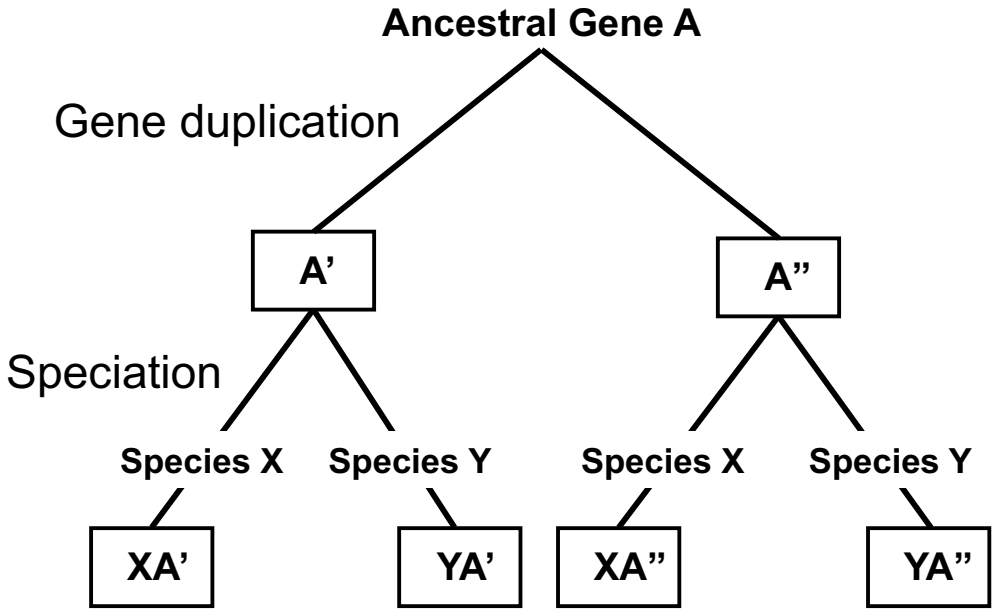
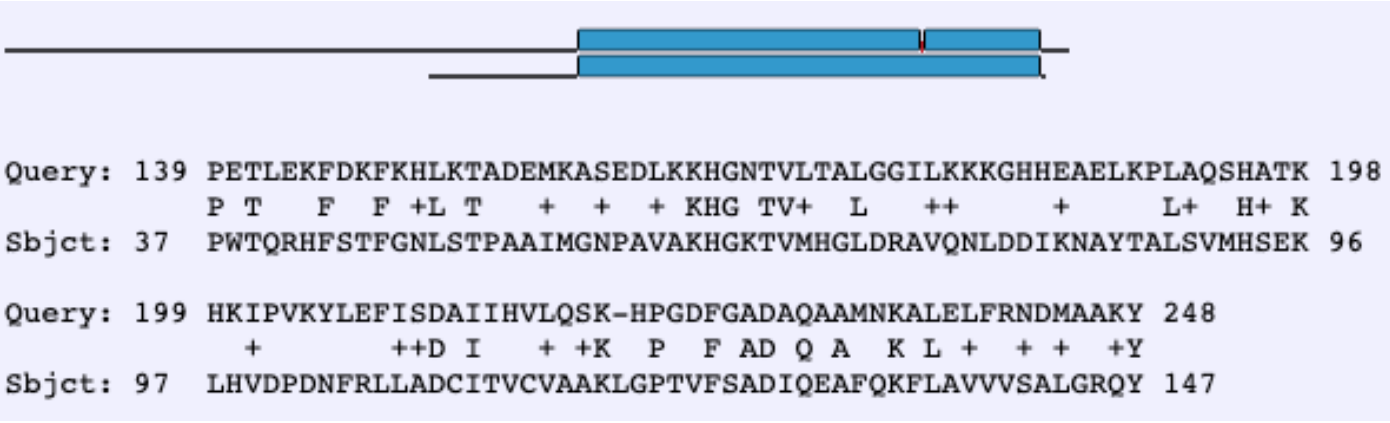


Figure Q4(c): Gene tree for an arbitrary gene A.

(d) Figure Q4(d) shows the pairwise local alignment of two protein sequences of 255 and 148 amino acids long. Analyse the similarities between these sequences. How would you determine whether the two sequences are homologous or not? Justify your answers. [8 marks]

Continued/…

Q4 Continued/…



```
Query: 139 PETLEKFDKFKHLKTADEMKASEDLKKHGNTVLTALGGILKKKGHHEAELKPLAQSHATK 198
           P T   F  F +L T   +  +  + KHG TV+  L   ++     +     L+  H+ K
Sbjct: 37  PWTQRHFSTFGNLSTPAAIMGNPAVAKHGKTVMHGLDRAVQNLDDIKNAYTALSVMHSEK 96

Query: 199 HKIPVKYLEFISDAIIHVLQSK-HPGDFGADAQAAMNKALELFRNDMAAKY 248
             +       ++D I   + +K  P  F AD Q A  K L +  + +  +Y
Sbjct: 97  LHVDPDNFRLLADCITVCVAAKLGPTVFSADIQEAFQKFLAVVVSALGRQY 147
```

Figure Q4(d): Pairwise alignment of two protein sequences.

Use the following information if required:

$$I(\%) = \frac{L_i}{L_a} \times 100\%$$

$$S(\%) = \frac{L_i + L_s}{L_a} \times 100\%$$

Where the symbols have their usual meanings.

END OF PAPER

A Islam, E Ratcliffe