



Second Semester Examination  
2023/2024 Academic Session

July/August 2024

**BOE202 – Introduction to Bioinformatic  
(*Pengenalan kepada Bioinformatik*)**

Duration : 2 hours  
(*Masa : 2 jam*)

---

Please ensure that this examination paper consists of FOUR (4) pages of printed material before you begin the examination.

*[Sila pastikan bahawa kertas peperiksaan ini mengandungi EMPAT (4) muka surat yang bercetak sebelum anda memulakan peperiksaan ini.]*

**Instructions:** Answer **FOUR (4)** out of **FIVE (5)** questions, in English or Bahasa Malaysia. Each question carries 25 marks.

**Arahan:** Jawab **EMPAT (4)** daripada **LIMA (5)** soalan, menggunakan Bahasa Inggeris atau Bahasa Malaysia. Setiap soalan bernilai 25 markah .]

In the event of any discrepancies, the English version shall be used.

*[Sekiranya terdapat sebarang percanggahan pada soalan peperiksaan, versi Bahasa Inggeris hendaklah digunapakai.]*

...2/-

- 2 -

1. [a] Explain the role of biological databases in supporting bioinformatics research, including the use of computational tools, sequence alignments, structural models, and phylogenetic information.  
*[Terangkan peranan pangkalan data biologi dalam menyokong penyelidikan bioinformatik, termasuk penggunaan aplikasi komputer, penjajaran jujukan, model struktur, dan maklumat filogenetik]*
- (10 marks / 10 markah)
- [b] Identify the steps involved in determining new insect species discovered during scientific expeditions at Endau-Rompin National Park, based on data retrieval from available online biological databases.  
*[Kenal pasti langkah-langkah yang terlibat dalam menentukan spesies serangga baharu yang ditemui semasa ekspedisi saintifik di Taman Negara Endau-Rompin, berdasarkan pengambilan data daripada pangkalan data biologi dalam talian yang tersedia.]*
- (15 marks / 15 markah)
2. [a] Explain the significance of data retrieval systems in molecular biology research and the use of biological databases to search for DNA and protein sequences. Provide examples of large databases for DNA sequences and discuss the benefits of database searching for proteins in molecular biology research.  
*[Terangkan kepentingan sistem perolehan data dalam penyelidikan biologi molekul dan penggunaan pangkalan data biologi untuk mencari jujukan DNA dan protein. Sediakan contoh pangkalan data besar untuk jujukan DNA dan bincangkan faedah pencarian pangkalan data untuk protein dalam penyelidikan biologi molekul.]*
- (10 marks / 10 markah)
- [b] Identify the importance of phylogenetics in biology and provide suitable examples.  
*[Kenal pasti kepentingan filogenetik dalam biologi dan berikan contoh yang sesuai .]*
- (15 marks / 15 markah)

...3/-

- 3 -

3. [a] Unweighted Pair Group Method with Arithmetic Mean (UPGMA) and Neighbour Joining (NJ) are two methods for constructing phylogenetic trees. Discuss their differences in terms of algorithmic approach, assumptions, computational complexity, and suitability for different types of data.

*[Kaedah kelompok berpasangan tak berpemberat dengan min aritmetik (UPGMA) dan penyambungan bersebelahan (NJ) adalah dua kaedah untuk membina pokok filogenetik. Bincangkan perbezaannya dari segi pendekatan algoritma, andaian, kekompleksan perkomputeran, dan kesesuaian untuk jenis data yang berbeza.]*

(10 marks / 10 markah)

[b]

Sequence 1 :	G A A T T C A G T T A
[Jujukan 1]	
Sequence 2 :	G G A T C G A
[Jujukan 2]	

Figure 1  
[Rajah 1]

By using the sequences shown in Figure 1, draw the sequence alignments based on the following scenarios;

*[Dengan menggunakan jujukan yang ditunjukkan dalam Rajah 1, lukiskan penjajaran jujukan berdasarkan senario berikut;]*

- [i] 2 matches, 4 mismatches and 0 gap  
*[2 padanan, 4 ketidakpadanan dan 0 jurang]*
- [ii] 4 matches, 3 mismatches and 0 gap  
*[4 padanan, 3 ketidakpadanan dan 0 jurang]*
- [iii] 5 matches, 1 mismatch and 2 gaps  
*[5 padanan, 1 ketidakpadanan dan 2 jurang]*

(15 marks / 15 markah)

...4/-

- 4 -

4. [a] Explain the prediction of genes through the ab initio method including the principles behind the approach, the challenges associated with gene prediction, and the strategies used to improve prediction accuracy.  
*[Terangkan ramalan gen melalui kaedah ab initio termasuk prinsip di sebalik pendekatan, cabaran yang berkaitan dengan ramalan gen, dan strategi yang digunakan untuk meningkatkan ketepatan ramalan.]*

(10 marks / 10 markah)

- [b] Write the **FIVE (5)** reasons that emphasize the importance of sequence alignment applications. Provide the example that demonstrates its importance.  
*[Tulis **LIMA (5)** sebab yang menekankan kepentingan aplikasi penjajaran jujukan. Berikan contoh yang menunjukkan kepentingannya.]*

(15 marks / 15 markah)

5. [a] Discuss the factors that limit the application of molecular clock models.  
*[Bincangkan faktor yang mengehadkan penggunaan model jam molekul.]*

(10 marks / 10 markah)

- [b] With the aid of a diagram, sketch and demonstrate the steps of protein modelling.  
*[Dengan bantuan gambar rajah, lakar dan tunjukkan langkah-langkah pemodelan protein.]*

(15 marks / 15 markah)

- oooOooo -