TRƯỜNG ĐẠI HỌC VĂN LANG
Khoa Công nghệ ứng dụng

ĐỀ THI, ĐÁP ÁN/RUBRIC VÀ THANG ĐIỂM
THI KẾT THÚC HỌC PHẦN
Học kỳ 3, năm học 2023-2024

I. Thông tin chung

Tên học phần:	Tin sinh học I		
Mã học phần:	71GENE40112	Số tín chỉ:	2
Mã nhóm lớp học phần:	233_71GENE40112_01		
Hình thức thi: **Tự luận**	Thời gian làm bài:	**90**	phút
Thí sinh được tham khảo tài liệu:	☒ Có	☐ Không	

II. Các yêu cầu của đề thi nhằm đáp ứng CLO

(Phần này phải phối hợp với thông tin từ đề cương chi tiết của học phần)

- Ngành CNSH Y Dược

Ký hiệu CLO	Nội dung CLO	Hình thức đánh giá	Trọng số CLO trong thành phần đánh giá (%)	Câu hỏi thi số	Điểm số tối đa	Lấy dữ liệu đo lường mức đạt PLO/PI
(1)	(2)	(3)	(4)	(5)	(6)	(7)
CLO1	Áp dụng kiến thức cơ sở khoa học của hóa sinh, vi sinh, sinh học phân tử, miễn dịch, công nghệ gene…để giải quyết các vấn đề trong chẩn đoán, xét nghiệm và liệu pháp điều trị.	Tự luận	20%	1	2.5	2.1
CLO2	Sử dụng hiệu quả các phần mềm chuyên dụng để giải quyết các vấn đề chuyên môn.	Tự luận	20%	2	2.5	4.2
CLO3	Sử dụng thành thạo các phần mềm phân tích dữ liệu trong việc phân tích và đánh giá kết quả	Tự luận	20%	3	2.5	5.2

	trong lĩnh vực CNSH YD.					
CLO4	Thực hiện thành thạo công nghệ giải trình tự DNA trong việc giám định gene, chẩn đoán đồng hành.	Tự luận	20%	3	3	5.3
CLO5	Tuân thủ các chuẩn mực đạo đức nghề nghiệp để gìn giữ sự trong sáng, thanh cao của nghề nghiệp, sự nghiệp.	Tự luận	20%	1-3	3	7.2

- Ngành CNSH

Ký hiệu CLO	**Nội dung CLO**	**Hình thức đánh giá**	**Trọng số CLO trong thành phần đánh giá (%)**	**Câu hỏi thi số**	**Điểm số tối đa**	**Lấy dữ liệu đo lường mức đạt PLO/PI**
(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
CLO1	Tổng quan tài liệu liên quan đến các vấn đề nghiên cứu và phát triển các sản phẩm Công nghệ Sinh học.	Tự luận	20%	1	2.5	3.1
CLO2	Triển khai các nội dung nghiên cứu dựa trên các phương pháp thực nghiệm đã xác định	Tự luận	20%	2	2.5	3.2
CLO3	Sử dụng hiệu quả các phần mềm tin sinh học, phân tích thống kê để giải quyết các vấn đề chuyên môn	Tự luận	20%	3	3	4.2
CLO4	Vận dụng hiệu quả kỹ năng tư duy (tư duy: sáng tạo, phản biện, logic, tích cực,…, tùy từng ngành thiết kế) nhằm giải quyết một cách tối ưu các vấn đề trong công việc	Tự luận	20%	4	3	4.6

CLO5	Thể hiện khả năng thích nghi với môi trường học tập và làm việc thay đổi.	Tự luận	20%	1-4	2.5	7.1

III. Nội dung câu hỏi thi

Yêu cầu: Sinh viên làm bài thi tại phòng máy tính, được sử dụng mạng Internet để thực hiện bài làm.

- Upload file bài làm (word, excel, pdf…);
- Upload hình ảnh bài làm (chỉ những trường hợp vẽ biểu đồ, công thức tính toán đặc biệt).

Câu hỏi 1: (3 điểm) Cho trình tự Bacteriophage phi-X174 có accession number là Z35638.1
- Hãy dự đoán có bao nhiêu đoạn ORF (open reading frame) và CDS có kích thước tối thiểu 300 nucleotide trên phage này?
- Tìm vùng mã hóa gen kháng kháng sinh beta-lactam trên phage này?

(Sử dụng các công cụ OFR Finder, BLAST, CD search)

Câu hỏi 2: (3 điểm) Xây dựng một cây phả hệ (rooted tree) từ trình tự protein Cytochrome c từ các loài eukaryote khác nhau. Các trình tự DNA được cung cấp bằng đường link phía dưới
https://www.dropbox.com/s/fw3dq0wsjh78dr9/cytochrome%20c.txt?dl=0
- Thực hiện sắp dóng cột các trình tự này.
- Hãy upload hình cây phả hệ vào khung trả lời của hệ thống thi.
- Cho biết loài nào có mối quan hệ gần nhất với người? Giải thích tại sao?

(Gợi ý: MAFFT, MEGA11 sử dụng phương pháp Neighbor joining, chỉ số bootstrap là 100)

Câu 3 (2 điểm): Cho một đoạn trình tự sau đây

TGCAGTCGATCATCAGCATACCTAGGTTTCGTCCGGGTGTGACCGAAAGGTAAG
ATGGAGAGCCTTGTTCTTGGTGTCAACGAGAAAACACACGTCCAACTCAGTTTG
CCTGTCCTTCAGGTTAGAGACGTGCTAGTGCGTGGCTTCGGGGACTCTGTGGAA
GAGGCCCTATCGGAGGCACGTGAACACCTCAAAAATGGCACTTGTGGTCTAGTA
GAGCTGGAAAAAGGCGTACTGCCCCAGCTTGAACAGCCCTATGTGTTCATTAAA
CGTTCTGATGCCTTAAGCACCAATCACGGCCACAAGGTCGTTGAGCTGGTTGCA
GAAATGGACGGCATTCAGTACGGTCGTAGCGGTATAACACTGGGAGTACTCGT
GCCACATGTGGGCGAAACCCCAATTGCATACCGCAATGTTCTTCTTCGTAAGAA
CGGTAATAAGGGAGCCGGTGGTCATAGCTATGGCATCGATCTAAAGTCTTATGA
CTTAG

- Đây là vật liệu di truyền từ vi khuẩn hay virus
- Đoạn trình tự sau đây có mã hóa protein hay không?
- Chức năng sinh học của trình tự này là gì?

Câu hỏi 4: (2 điểm) Download một file fasta chứa toàn bộ gen mã hóa protein của vi khuẩn E. coli từ đường link sau về máy tính:

1/
https://ftp.ncbi.nlm.nih.gov/genomes/refseq/bacteria/Escherichia_coli/reference/GCF_000005845.2_ASM584v2/GCF_000005845.2_ASM584v2_cds_from_genomic.fna.gz

2/ gunzip GCF_000005845.2_ASM584v2_cds_from_genomic.fna.gz

- Hãy cho biết vi khuẩn này có bao nhiêu gen mã hóa protein
- Hãy cho biết có bao nhiêu gen mã hóa cho protein “DNA gyrase subunit”.

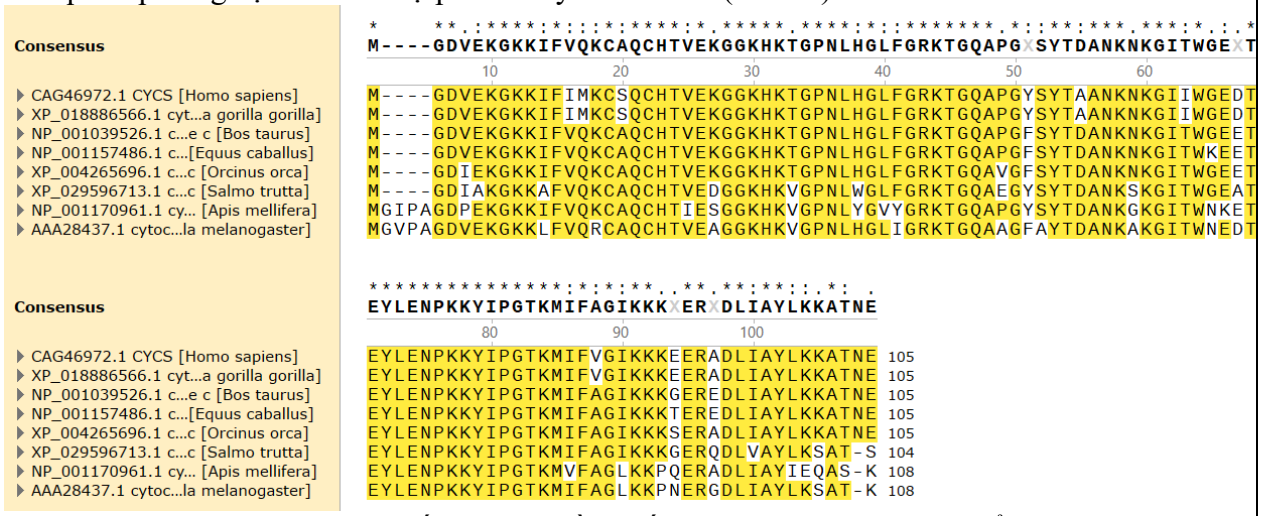
ĐÁP ÁP VÀ THANG ĐIỂM

<table>
<tr><th>Phần câu hỏi</th><th>Nội dung đáp án</th><th>Thang điểm</th><th>Ghi chú</th></tr>
<tr><td colspan="2">Tự luận</td><td></td><td></td></tr>
<tr><td>Câu 1: (3 điểm)</td><td>Cho trình tự Bacteriophage phi-X174 có accession number là Z35638.1
- Hãy dự đoán có bao nhiêu đoạn ORF (open reading frame) và CDS có kích thước tối thiểu 300 nucleotide trên phage này?
- Tìm kiếm chức năng gen của các ORF của phage này?
- Tìm vùng mã hóa gen kháng kháng sinh beta-lactam trên phage này?
- Tìm vùng promotor của phage
Đáp án:
- Có 2 đoạn ORF và CDS có kích thước tối thiểu 300 nucleotide trên phage này (1.5 điểm)
<table><tr><th>Label</th><th>Strand</th><th>Frame</th><th>Start</th><th>Stop</th><th>Length (nt | aa)</th></tr><tr><td>ORF1</td><td>+</td><td>3</td><td>201</td><td>1061</td><td>861 | 286</td></tr><tr><td>ORF2</td><td>-</td><td>3</td><td>2979</td><td>2674</td><td>306 | 101</td></tr></table>
- Đoạn ORF1 mã hóa cho protein TEM family class A beta-lactamase chịu trách nhiệm cho sự đề kháng kháng sinh beta-lactam. (1.5 điểm)</td><td>4 điểm</td><td></td></tr>
<tr><td>Câu 2: (3 điểm)</td><td>Xây dựng một cây phả hệ (rooted tree) từ trình tự protein Cytochrome c từ các loài eukaryote khác nhau. Các trình tự DNA được cung cấp bằng đường link phía dưới
https://www.dropbox.com/s/fw3dq0wsjh78dr9/cytochrome%20c.txt?dl=0
- Thực hiện sắp dóng cột các trình tự này.
- Upload hình cây phả hệ vào khung trả lời của hệ thống thi.
- Cho biết loài nào có mối quan hệ gần nhất với người? Giải thích tại sao?</td><td>3 điểm</td><td></td></tr>
</table>

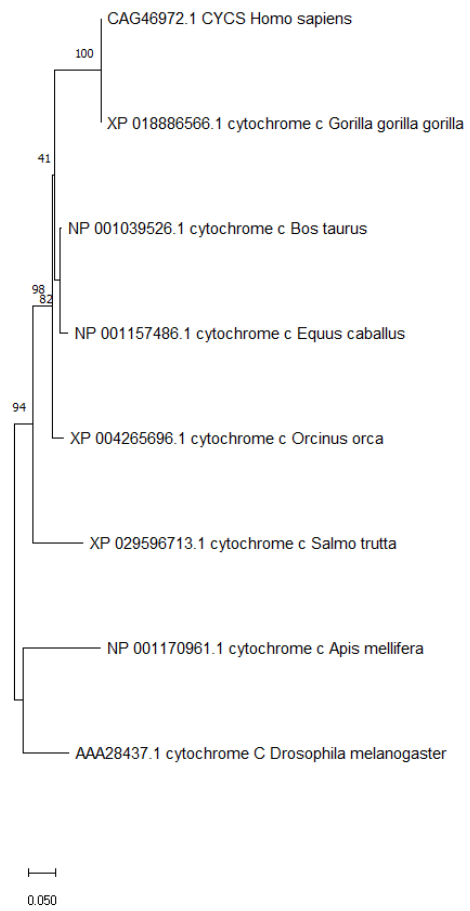
(Gợi ý: sử dụng phương pháp Neighbor joining, chỉ số bootstrap là 100)

Đáp án:

- Kết quả sắp dóng cột các trình tự protein Cytochrome c (1 điểm)



- Loài Gorilla gorilla gorrila có mối quan hệ gần nhất với Homo sapiens. (1 điểm)
- Càng nhiều acid amin mà một sinh vật có điểm chung, cả về loại và trật tự, cho thấy mối quan hệ càng chặt chẽ. (1 điểm)



Câu 3: (2 điểm)	Cho một đoạn trình tự sau đây TGCAGTCGATCATCAGCATACCTAGGTTTCGTCCGGGTGTGACCGAAAGGTAAGAT GGAGAGCCTTGTTCTTGGTGTCAACGAGAAAACACACGTCCAACTCAGTTTGCCTG TCCTTCAGGTTAGAGACGTGCTAGTGCGTGGCTTCGGGGACTCTGTGGAAGAGGCC CTATCGGAGGCACGTGAACACCTCAAAAATGGCACTTGTGGTCTAGTAGAGCTGGA AAAAGGCGTACTGCCCCAGCTTGAACAGCCCTATGTGTTCATTAAACGTTCTGATG	**3 điểm**	

CCTTAAGCACCAATCACGGCCACAAGGTCGTTGAGCTGGTTGCAGAAATGGACGG
CATTCAGTACGGTCGTAGCGGTATAACACTGGGAGTACTCGTGCCACATGTGGGCG
AAACCCCAATTGCATACCGCAATGTTCTTCTTCGTAAGAACGGTAATAAGGGAGCC
GGTGGTCATAGCTATGGCATCGATCTAAAGTCTTATGACTTAG

- Đây là vật liệu di truyền từ vi khuẩn hay virus (0.5 điểm)
- Đoạn trình tự sau đây có mã hóa protein hay không? (0.5 điểm)
- Chức năng sinh học của trình tự này là gì? (1 điểm)

Đáp án:

- Kết quả BLAST cho thấy trình tự trên là của virus SARS-Cov2 (0.5 điểm)

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Mutant SARS coronavirus Urbani clone SARS-Urbani-MA_SHC014-spike, complete genome	SARS coronavir…	905	905	100%	0.0	100.00%	29730	MT308984.1
SARS coronavirus Urbani isolate icSARS-C7-MA, complete genome	SARS coronavir…	905	905	100%	0.0	100.00%	29874	MK062184.1
SARS coronavirus Urbani isolate icSARS-C7, complete genome	SARS coronavir…	905	905	100%	0.0	100.00%	29874	MK062183.1
SARS coronavirus Urbani isolate icSARS-C3-MA, complete genome	SARS coronavir…	905	905	100%	0.0	100.00%	29727	MK062182.1
SARS coronavirus Urbani isolate icSARS-C3, complete genome	SARS coronavir…	905	905	100%	0.0	100.00%	29727	MK062181.1

- Kết quả đối chiếu trên GENBANK cho thấy đây là đoạn trình tự mã hóa cho protein có tên là ORF1ab (1 điểm)

	FEATURES Location/Qualifiers source 1..490 /organism="SARS coronavirus Urbani" /mol_type="genomic RNA" /host="mouse" /db_xref="taxon:228330" /clone="SARS-Urbani-MA_SHC014-spike" /note="mouse-adapted; spike protein derived from bat coronavirus SHC014" 5'UTR <1..54 CDS 55..>490 /ribosomal_slippage /codon_start=1 /product="ORF1ab" /protein_id="QJE50587.1" - Chức năng của ORF1ab là protein phi cấu trúc, và có vai trò trong phiên mã virus, nhân bản, ức chế hệ miễn dịch của tế bào chủ (0.5 điểm)		
Câu 4: (2 điểm)	Download một file fasta chứa toàn bộ gen mã hóa protein của vi khuẩn *E. coli* từ đường link sau về máy tính: 1/ https://ftp.ncbi.nlm.nih.gov/genomes/refseq/bacteria/Escherichia_coli/reference/GCF_000005845.2_ASM584v2/GCF_000005845.2_ASM584v2_cds_from_genomic.fna.gz 2/ gunzip GCF_000005845.2_ASM584v2_cds_from_genomic.fna.gz - Hãy cho biết vi khuẩn này có bao nhiêu gen mã hóa protein - Hãy cho biết có bao nhiêu gen mã hóa cho protein "DNA gyrase subunit". Đáp án: a) grep -c "protein" GCF_000005845.2_ASM584v2_cds_from_genomic.fna Có 4315 gen mã hóa protein từ vi khuẩn E. coli b) grep -c "DNA gyrase subunit" GCF_000005845.2_ASM584v2_cds_from_genomic.fna Có 2 gen mã hóa cho DNA gyrase subunit A và B	**2 điểm**	
	Điểm tổng	**10.0**	

TP. Hồ Chí Minh, ngày ..... tháng ..... năm 2024

Người duyệt đề	**Giảng viên ra đề**
TS. Vũ Thị Quyền	**TS. Bùi Hoàng Phúc**