

TRƯỜNG ĐẠI HỌC VĂN LANG
GHI TÊN ĐƠN VỊ CHỦ QUẢN MÔN HỌC

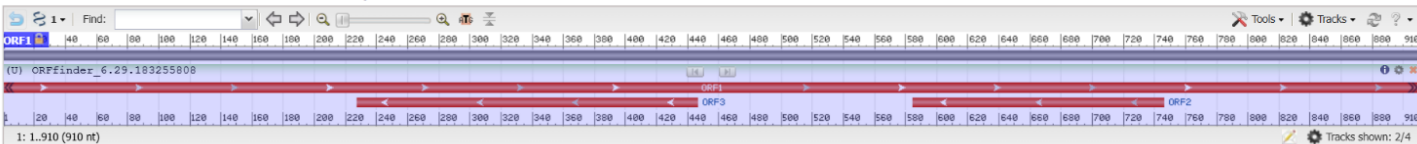
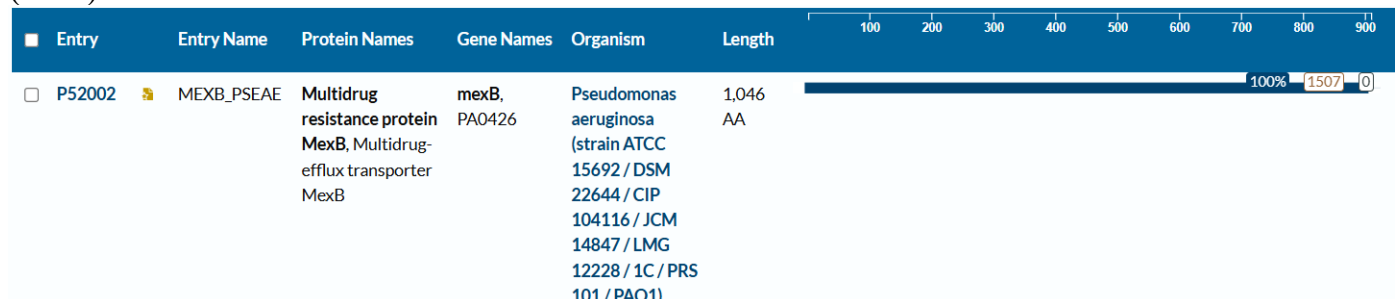
ĐỀ THI, ĐÁP ÁN/RUBRIC VÀ THANG ĐIỂM
THI KẾT THÚC HỌC PHẦN
Học kỳ 3, năm học 2024-2025

I. Thông tin chung

Tên học phần:	Tin sinh học I		
Mã học phần:	71GENE40112	Số tín chỉ:	2
Mã nhóm lớp học phần:	243_71GENE40112_01		
Hình thức thi: Tự luận	Thời gian làm bài:	95	phút
☐ Đề thi có sử dụng phần mềm riêng	GV ghi cụ thể tên phần mềm:		
<i>Thí sinh được tham khảo tài liệu:</i>	☒ Có	☐ Không	

ĐÁP ÁP VÀ THANG ĐIỂM

Phần câu hỏi	Nội dung đáp án	Thang điểm	Ghi chú
Tự luận			
Câu 1	Một nhóm nghiên cứu phát hiện một đoạn DNA dài 1.5 kb trong bộ gen của một chủng <i>Pseudomonas aeruginosa</i> kháng kháng sinh. Họ nghi ngờ đây là một gen mã hóa protein vận chuyển có thể liên quan đến kháng thuốc. ATGTCTGAAGTTTTTCATTGATAGGCCCATTTTTGCGTGGGTGATCGCCTTGGTGATCATGCTCGCGGGCG GCCTGTCTGATCCTCAGTCTGCCGGTCAACAGTACCCGGCCATCGCCCCGCCGGCCATCGCCGTGCAGGT GAGTACCCGGGCGCCTCGGCCGAGACGGTGCAGGACACCGTGGTCCAGGTGATCGAGCAGCAGATGAAC GGGATCGACAATCTGCGCTACATCTCCTCGGAGAGTAACTCCGACGGCAGCATGACCATCACCGTGACCT TCGAACAGGGCACCGACCCCGACATCGCCCAGGTCCAGGTGCAGAACAAGCTGCAACTGGCCACCCCGCT ACTGCCGCGAGGAAGTGCAGCGCCAGGGGATCCGGGTGACCAAGGCGGTGAAGAACTTCCTCATGGTGGTC GGTGTGGTTTTCCACCGACGGCAGCATGACCAAGGAAGACCTGTCGAACCTACATCGTTTTCCAACATCCAGG ACCCGCTCTCGCGGACCAAGGGCGTCGGTGACTTCCAGGTGTTTCGGCTCGCAGTACTCGATGCGCATCTG GCTCGACCCGGCCAAGCTGAACAGCTACCAGCTGACCCCGGCGACGTGAGCAGCGCGATCCAGGCGCAG AACGTGCAGATTTTCCTCCGGCCAGCTCGGCGGCTTGCCGGCGGTCAAGGGCCAGCAGCTCAACGCCACCA TCATCGGCAAGACCCGCTGCGAGCCGCGGAGCAATTGAGAACATCCTGCTCAAGGTCAATCCCGACGG TTCCCAGGTGCGCCTGAAGGACGTGCGCGATGTAGGCCTGGGCGGCCAGGACTACAGCATCAACGCGCAG TTCAACGGCAGCCCGGCGTCCGGTATCGCGATCAAGCTGGCCACCGGCGCCAACGCGCTGGATACCGCCA Hãy trình bày kết quả xác định khung đọc mở (ORF), dự đoán chức năng, và xác định miền chức năng protein mã hóa từ đoạn DNA trên bằng cách sử dụng ORF Finder và ExPASy/BLASTP.	2.0	

Nội dung a.	Phát hiện ORF: Xác định khung đọc mở bằng orffinder Open Reading Frame Viewer Sequence ORFs found: 3 Genetic code: 1 Start codon: 'ATG' only  - Tổng số ORF được tìm thấy: 3 khung đọc mở (ORF1, ORF2, ORF3). (0.25) - ORF1 nằm trên mạch xuôi (+), dài nhất, có chiều dài đáng kể (~900 nt, tương đương 302 amino acid), nên là ứng viên chính để phân tích chức năng. (0.25)	0.5																									
Nội dung b.	Sử dụng công cụ SmartBlast vào UniProtKB/Swiss-Prot (swissprot) - Protein mã hóa bởi ORF1 được đưa vào BLASTP để so sánh với cơ sở dữ liệu Swiss-Prot. (0.25) <table><thead><tr><th>Tên protein khớp</th><th>Loài</th><th>% Identity</th><th>Query cover</th><th>E-value</th><th>Accession</th></tr></thead><tbody><tr><td>MexB</td><td><i>Pseudomonas</i></td><td>100%</td><td>100%</td><td>0.0</td><td>WP_003107312.1</td></tr><tr><td>AcrB</td><td><i>E. coli</i></td><td>68.65%</td><td>100%</td><td>3e-146</td><td>WP_001132452.1</td></tr><tr><td>MtrD</td><td><i>Neisseria meningitidis</i></td><td>50.17%</td><td>100%</td><td>9e-96</td><td>WP_082308689.1</td></tr></tbody></table> - Protein mã hóa bởi ORF1 được đưa vào BLASTP để so sánh với cơ sở dữ liệu ExpaSy. (0.25) 	Tên protein khớp	Loài	% Identity	Query cover	E-value	Accession	MexB	<i>Pseudomonas</i>	100%	100%	0.0	WP_003107312.1	AcrB	<i>E. coli</i>	68.65%	100%	3e-146	WP_001132452.1	MtrD	<i>Neisseria meningitidis</i>	50.17%	100%	9e-96	WP_082308689.1	0.5	
Tên protein khớp	Loài	% Identity	Query cover	E-value	Accession																						
MexB	<i>Pseudomonas</i>	100%	100%	0.0	WP_003107312.1																						
AcrB	<i>E. coli</i>	68.65%	100%	3e-146	WP_001132452.1																						
MtrD	<i>Neisseria meningitidis</i>	50.17%	100%	9e-96	WP_082308689.1																						

Nội dung c.	Query: unnamed protein product, partial Query length: 303 aa multidrug efflux RND transporter permease subunit MexD multidrug efflux RND transporter permease subunit MtrD efflux RND transporter permease subunit efflux RND transporter permease AcrB Your query: unnamed protein product, partial multidrug efflux RND transporter permease subunit MexB About the database See full multiple alignment Legend - Trình tự có 100% giống với MexB, một protein vận chuyển thuộc họ RND efflux pump, giúp vi khuẩn tổng xuất kháng sinh ra ngoài tế bào. (0.25) - Đây là bằng chứng rõ ràng cho thấy đoạn DNA đã cho mã hóa gen <i>mexB</i>, đóng vai trò trong cơ chế kháng kháng sinh đa dạng ở <i>Pseudomonas aeruginosa</i>. (0.25) - Các kết quả tương đồng với AcrB, MtrD từ các loài khác cũng củng cố thêm rằng đây là thành viên của họ multidrug transporter RND-type. (0.25)	0.75																															
Nội dung d.	- Tổng hợp và kết luận logic, rõ ràng	0.25																															
Câu 2	Trình tự mã hóa có mã accession NCBI là BC000292 -Trình tự này có nguồn gốc từ loài sinh vật nào? -Cho biết các vùng chức năng của protein này là gì?	3.0																															
Nội dung a.	Trình tự XP_047299549 thuộc về loài <i>Homo sapiens</i>.	0.5																															
Nội dung b.	Các vùng chức năng (domain) của protein (0.25) Graphical summary ☐ Zoom to residue level show extra options Query seq. putative methyl-histone binding site PWWP sequence motif Zn binding site peptide binding site Specific hits PHA03169 PWWP_DNMT3A ADDZ_Dnmt3a FYVE_like_SF superfamily Dcm Superfamilies PWWP superfamily Search for similar domain architectures Refine search List of domain hits <table><tr><th>#</th><th>Name</th><th>Accession</th><th>Description</th><th>Interval</th><th>E-value</th></tr><tr><td>[H]</td><td>PWWP_DNMT3A</td><td>cd20154</td><td>PWWP domain found in DNA (cytosine-5)-methyltransferase 3A (DNMT3A), DNMT3A, also called DNA ...</td><td>285-418</td><td>2.77e-102</td></tr><tr><td>[H]</td><td>ADDZ_Dnmt3a</td><td>cd11729</td><td>ADDZ domain found in DNA (cytosine-5) methyltransferases (C5-MTases) 3a (Dnmt3a), Dnmt3a is a ...</td><td>487-614</td><td>5.06e-89</td></tr><tr><td>[H]</td><td>Dcm super family</td><td>cd43082</td><td>DNA-cytosine methylase [Replication, recombination and repair]</td><td>631-794</td><td>7.22e-16</td></tr><tr><td>[H]</td><td>PHA03169 super family</td><td>cd27451</td><td>hypothetical protein, Provisional</td><td>15-188</td><td>4.38e-03</td></tr></table> Domain 1: PWWP domain Tên domain: PWWP domain. (0.25) Vị trí xấp xỉ: amino acid 67–153. (0.25)	#	Name	Accession	Description	Interval	E-value	[H]	PWWP_DNMT3A	cd20154	PWWP domain found in DNA (cytosine-5)-methyltransferase 3A (DNMT3A), DNMT3A, also called DNA ...	285-418	2.77e-102	[H]	ADDZ_Dnmt3a	cd11729	ADDZ domain found in DNA (cytosine-5) methyltransferases (C5-MTases) 3a (Dnmt3a), Dnmt3a is a ...	487-614	5.06e-89	[H]	Dcm super family	cd43082	DNA-cytosine methylase [Replication, recombination and repair]	631-794	7.22e-16	[H]	PHA03169 super family	cd27451	hypothetical protein, Provisional	15-188	4.38e-03	2.5	
#	Name	Accession	Description	Interval	E-value																												
[H]	PWWP_DNMT3A	cd20154	PWWP domain found in DNA (cytosine-5)-methyltransferase 3A (DNMT3A), DNMT3A, also called DNA ...	285-418	2.77e-102																												
[H]	ADDZ_Dnmt3a	cd11729	ADDZ domain found in DNA (cytosine-5) methyltransferases (C5-MTases) 3a (Dnmt3a), Dnmt3a is a ...	487-614	5.06e-89																												
[H]	Dcm super family	cd43082	DNA-cytosine methylase [Replication, recombination and repair]	631-794	7.22e-16																												
[H]	PHA03169 super family	cd27451	hypothetical protein, Provisional	15-188	4.38e-03																												

	Chức năng: Gắn vào DNA hoặc histone đã methyl hóa → hướng enzym đến vị trí cần methyl hóa. (0.25) Domain: ADDz_Dnmt3a domain Tên domain: ADD domain (ATRX-DNMT3-DNMT3L). (0.25) Vị trí xấp xỉ: amino acid 264–391. (0.25) Chức năng: Điều hòa hoạt tính enzyme và tương tác với histone. (0.25) Domain 3: Cyt_C5_DNA_methylase domain Tên domain: DNA methyltransferase catalytic domain (Cyt_C5_DNA_methylase). (0.25) Vị trí xấp xỉ: amino acid 409–572. (0.25) Chức năng: Xúc tác quá trình methyl hóa cytosine tại vị trí C5 → thay đổi biểu hiện gen (epigenetic regulation). (0.25)		
Câu 3	Trong một nghiên cứu về thích nghi với độ cao, bạn phát hiện biến thể SNP rs11549465 trong gen HIF1A có tần số cao ở người Tây Tạng. Sử dụng UCSC Genome Browser, hãy mô tả cách xác định vị trí của SNP này và đánh giá xem nó có ảnh hưởng đến vùng mã hóa hay không. Kết luận ý nghĩa của SNP trong thích nghi sinh học.	2.0	
Nội dung a.	Cách xác định vị trí của SNP trên UCSC Genome Browser Truy cập UCSC Genome Browser và chọn bộ gen người (Human, ví dụ: GRCh38/hg38). (0.25) Nhập mã biến thể rs11549465 vào ô tìm kiếm (search box) và nhấn “go”. (0.25) Trình duyệt chuyển đến vị trí SNP trên nhiễm sắc thể: chr14:62,178,890 (GRCh38/hg38). (0.25) Quan sát vùng hiển thị, đặc biệt là track Genes and Gene Predictions, dbSNP, và RefSeq Genes để xác định vị trí SNP trong gen. (0.25)	1.0	
Nội dung b	Đánh giá SNP có nằm trong vùng mã hóa hay không - SNP rs11549465 nằm trong gen HIF1A (Hypoxia-Inducible Factor 1-Alpha). (0.25) - Theo RefSeq, SNP nằm trong exon 12 của HIF1A → là vùng mã hóa. (0.25) - Đây là SNP không đồng nghĩa (missense): gây thay đổi amino acid từ Proline → Serine tại vị trí 582 (P582S mutation). (0.25)	1.0	

	- Thay đổi này có thể ảnh hưởng đến ổ điều hòa oxy hóa và phân hủy protein HIF1A → ảnh hưởng đến ổn định protein dưới điều kiện thiếu oxy. (0.25)		
Câu 4	Loài <i>Helicobacter</i> là vi khuẩn có khả năng xâm nhiễm và gây bệnh ở nhiều loài vật chủ khác nhau như <i>Helicobacter pylori</i> gây bệnh ở người. Hãy vẽ cây phát sinh chủng loài bằng trình tự gen 16S <i>rRNA</i> của các loài sau: <i>Helicobacter cinaedi</i> (NR_025941), <i>Helicobacter pylori</i> (U08906), <i>Helicobacter acinonychis</i> (NR_025940), <i>Helicobacter felis</i> (M37643), và <i>Campylobacter jejuni</i> (NR_118520). - Thực hiện sắp xếp cột các trình tự này và - Biện luận mối quan hệ di truyền giữa các loài vi khuẩn <i>Helicobacter</i> - Tính khoảng cách di truyền giữa các loài này.	3.0	
Nội dung a	Sắp dòng cột trình tự 16S <i>rRNA</i> Tải các trình tự từ NCBI theo mã accession đã cho dưới định dạng FASTA. (0.25) Sử dụng công cụ căn chỉnh (multiple sequence alignment) như Clustal Omega, MAFFT hoặc MEGA. (0.25) Kết quả hiển thị các vùng giống và khác nhau trong 16S <i>rRNA</i> → tạo cơ sở phân tích mối quan hệ tiến hóa. (0.25) Kiểm tra xem có sai khác lớn tại vùng V3–V4 hay vùng biến thiên (variable region) trong 16S không. (0.25)	1.0	
Nội dung b	Vẽ cây phát sinh chủng loài (phylogenetic tree) Sử dụng dữ liệu căn chỉnh để tạo cây bằng phần mềm như MEGA, PhyML hoặc iTOL. (0.25) Chọn phương pháp xây dựng cây: Neighbor-Joining (NJ) hoặc Maximum Likelihood (ML). (0.25) Cây phát sinh cho thấy: (0.25) • <i>H. pylori</i> và <i>H. acinonychis</i> nằm gần nhau nhất. • <i>Campylobacter jejuni</i> nằm ngoài (outgroup). Đảm bảo cây có nhánh bootstrap (>70%) để đảm bảo độ tin cậy phân nhóm. (0.25)	1.0	
Nội dung c	Biện luận mối quan hệ di truyền <i>H. pylori</i> gần với <i>H. acinonychis</i> → chứng tỏ mối quan hệ tiến hóa gần, có thể do cùng nguồn gốc tổ tiên. (0.25)	0.5	

	<i>H. felis</i> và <i>H. cinaedi</i> xa hơn → cho thấy sự khác biệt tiến hóa do thích nghi ở các vật chủ khác nhau (mèo, người, linh trưởng). (0.25)		
Nội dung d	Tính khoảng cách di truyền Khoảng cách di truyền có thể tính bằng mô hình Kimura 2-parameter hoặc Jukes-Cantor trong phần mềm MEGA. (0.25) 11. Kết quả cho thấy: (0.25) - Khoảng cách <i>H. pylori</i> – <i>H. acinonychis</i> ≈ 0.005 - Khoảng cách với <i>H. felis</i> ≈ 0.02 <i>Campylobacter jejuni</i> có khoảng cách lớn nhất (ngoài nhóm): >0.1	0.5	
	Điểm tổng	10.0	