

KATHMANDU UNIVERSITY  
End Semester Examination  
February, 2025

Marks Scored:

Level : B.Tech.

Year : III

Exam Roll No. :

Time: 30 mins.

Registration No.:

Course : BIOT 312

Semester : II

F. M. : 20

Date : 24 FEB 2025

SECTION "A"

23

[20 Q. × 1 = 20 marks]

**Choose and Mark [X] in the most appropriate option from each set of choices**

- Which of the following is **NOT** a component of functional analysis in bioinformatics?  
 Genome comparison                       Metabolic Pathway Modeling  
 Gene expression profiling                       Protein Interaction Prediction
- Addition of unlabelled nucleotides and enzyme to initiate solid phase bridge amplification occurs in  
 Sanger Sequencing                       Pyrosequencing  
 Illumina Genome Analyzer                       SOLiD sequencing
- Which of the database stores information in cataloging human genes and genetic diseases?  
 KEGG                       TrEMBL                       RefSeq                       OMIM
- The 3<sup>rd</sup> column and 4<sup>th</sup> column of the Browser Extensible Data and Gene Feature Format files hold the information on  
 chromosome name                       feature start coordinate  
 feature stop coordinate                       feature name
- A Line that runs parallel to the main continuous diagonal in a dotplot can be referred as:  
 Repeats                       Inverted repeats                       Palindrome                       Partial palindrome
- PAM 80 matrix was created using sequences that had \_\_\_\_\_ percentage sequence similarity:  
 80 %                       50%                       20 %                       100%
- During BLAST searches, if the word size is increased then,  
 sensitivity increases & speed decreases                       sensitivity decreases & speed increases  
 sensitivity increases & speed increases                       sensitivity decreases & speed decreases
- The number of possible rooted tree the can be created for 7 OTUs is  
 945                       10395                       135135                       2027025
- Sequences A and B differ by 30%. If 20% of changes are due to transitions and 10% due to transversions, the evolutionary distance according to Jukes-Cantor model is given as:  
 0.38                       0.40                       0.48                       0.30
- If there are two changes but only one difference is observed, then it is a type of  
 Coincidental substitution                       Single substitution  
 Parallel substitution                       Back substitution

11. A large set of pseudosamples all create the same tree topology, this is referred as  
 Accuracy       Repeatability       Precision       Variance
12. The term that specifies incorrectly predicted feature is called as  
 True Positive       False Positive       False Negative       True Negative
13. Which of the amino acids participates in Hydrogen bond or in enzyme active sites?  
 Glutamine       Glycine       Alanine       Proline
14. The major class of secondary structure found in Cytochrome C (2CCY) is  
 mainly  $\alpha$ -helix       mainly  $\beta$ -sheet  
 mixture of  $\alpha$ -helix &  $\beta$ -sheet        $\alpha$ -helix &  $\beta$ -sheet in separate domains
15. Which of the following is an example of tertiary secondary structure elements of RNA  
 bulge loop       kissing hairpins  
 multi-branch loop       hairpin loop
16. Which function in Biopython is used to read multiple sequences from a file?  
 SeqIO.read()       SeqIO.fetch()       SeqIO.parse()       SeqIO.extract()
17. Which function in Bio.SeqUtils is used to calculate the GC content of a sequence?  
 GCContent()       gc\_fraction()       gc\_ratio()       GC\_percentage()
18. Which of the following can lead to incorrect multiple sequence alignment for phylogeny tree construction?  
 Sequence are homologous       Incorporating gap penalties  
 Domains in same order       Proteins contain repeated domains
19. Find the correct order for the construction of position specific scoring matrix for:  
a. fit a sequence into the matrix calculate raw frequencies  
b. normalize raw frequencies  
c. calculate raw frequencies  
d. take log base 2  
 a, b, c, d       d, c, b, a       c, b, d, a       b, c, d, a
20. In a phylogenetic analysis, long branch attraction (LBA) can cause which of the following issues?  
 Closely related taxa appear more distantly related  
 Distantly related taxa are incorrectly grouped together  
 Increased support for the true evolutionary relationships  
 No effect on tree topology

KATHMANDU UNIVERSITY  
End Semester Examination  
February, 2025

Level : B.Tech.  
Year : III  
Time : 2 hrs. 30mins.

23  
~~24~~ FEB 2025

Course : BIOT 312  
Semester : II  
F. M. : 55

SECTION "B"

[3 Q. × 7 = 21 marks]

Attempt **ANY THREE** questions

1. Differentiate between Distance based and Character based Tree Building Method. Create a UPGMA tree and indicate branch lengths between OTUs and all internal nodes. [2+5]

	A	B	C	D	E	F
A						
B	17					
C	21	30				
D	31	34	28			
E	23	21	39	43		
F	25	28	15	35	18	

2. Explain how Sequencing by Oligonucleotide Ligation and Detection works. Explain how dibase decoding and the primer reset decodes the dibase signal. [4+3]
3. Explain the differences between local and global alignment procedure for sequence alignment. Create **TWO** possible global alignment for the below provided sequences; considering score for match = 2, mismatch = (-3) and gap penalty = (-1) [2+5]  
Seq #1: ATCGCGATTA  
Seq #2: ATGGCGCTTA
4. Define the SeqRecord object in Biopython. Discuss its primary attributes and explain how Bio.SeqIO utilizes SeqRecord objects for handling biological sequence files with an example. Write a Python program using Biopython to perform a pairwise sequence alignment between the following two DNA sequences and print the alignment score. [2+3+3]

SECTION "C"

[6 Q. × 4 = 24 marks]

Attempt **ANY SIX** questions

5. For the following sequence 5'-CGAATGCTCAGG-3'; explain the gel electrophoresis profile generated after Sanger Sequencing.
6. Calculate the log odds occurrence of CTGGCCTCT if the multiple sequence alignment is

C	T	G	-	-	-	C	T	C	T
A	C	G	G	-	-	A	T	A	T
A	C	C	G	-	C	A	T	A	T
A	G	C	-	-	-	T	T	C	T
C	A	G	G	-	C	C	T	C	T

**P.T.O.**

7. Construct a most parsimonious unrooted tree for four taxa using the provided alignment positions generated after a multiple sequence alignment.

Seq ID	Alignment Positions			
	131	132	133	134
Human	A	A	T	G
Chimp	C	G	T	C
Gorilla	G	A	A	C
Orangutan	G	G	A	C

8. Explain how operons are predicted in prokaryotes.
9. Create a PSSM matrix based on the position of the nucleotides in the provide sequence and evaluate the probability of sequence ACCGCGATAT fitting to that matrix.

Position	1	2	3	4	5	6	7	8	9	10
Seq. 1	A	C	C	G	C	G	A	T	A	T
Seq. 2	A	G	C	G	C	A	T	T	C	T
Seq. 3	C	A	G	G	C	A	C	T	C	T
Seq. 4	C	T	G	G	C	G	C	T	C	T
Seq. 5	A	C	G	G	C	G	A	T	A	T

10. Write a Python script using Biopython to fetch a nucleotide sequence from GenBank using Entrez.efetch(), parse it using SeqIO, and print the retrieved sequence record. Explain the individual steps used.
11. Explain Nussinov Algorithm. Derive the RNA secondary structure of AUCGGCAUUCAGU

SECTION "D"  
[5 Q. × 2 = 10 marks]

Answer *ANY FIVE* questions

12. Describe the various applications of Bioinformatics.
13. Differentiate between shotgun and hierarchical sequencing approaches.
14. State the significance of various matrices used for scoring DNA and protein alignments.
15. How does gene phylogeny and species phylogeny differ from each other?
16. How does gene prediction in prokaryotes work?
17. Explain Ramachandran plot.