



**M.TECH. DEGREE EXAMINATIONS JUNE 2015**

(Regulation 2014)

Second Semester

**BIOTECHNOLOGY**

P14BTT203: Computational Biology

**Time: Three Hours**

**Maximum Marks: 100**

**Answer all the Questions:-**

**PART A (10 x 1 = 10 Marks)**

1. **Assertion (A):** Under some circumstances, the similarity (%) value expressed by any tool is [K<sub>3</sub>] greater than the identical (%) value  
**Reason (R):** This is due to the summation (%) of the identical characters and acceptable mutatable characters.
- a) Both (A) and (R) are true, but (R) is the not the correct explanation of (A)      b) (A) is true and (R) is false  
c) Both (A) and (R) are false      d) Both (A) and (R) are true, and (R) is the correct explanation of (A)
2. Identify the correct combination of the matrices adopted for finding the sequence dissimilarity. [K<sub>3</sub>]  
a) BLOSUM 30 AND PAM 450      b) PAM 1 AND BLOSUM 62  
b) PAM 20 AND BLOSUM 80      d) BLOSUM 90 AND PAM 250
3. Which program use biological recombination procedure for multiple sequence alignment? [K<sub>2</sub>]  
a) CLUSTAL OMEGA      b) SOPMA  
c) SAGA      d) PHYLIP
4. Match the following and pick the right choice: [K<sub>4</sub>]

List I	List II
A. Strong sequence similarity	i. Boot strapping method
B. Poor sequence similarity	ii. NJ method
C. Moderate sequence similarity	iii. Maximum parsimony method
D. Quality check of predicted trees	iv. Maximum likelihood method

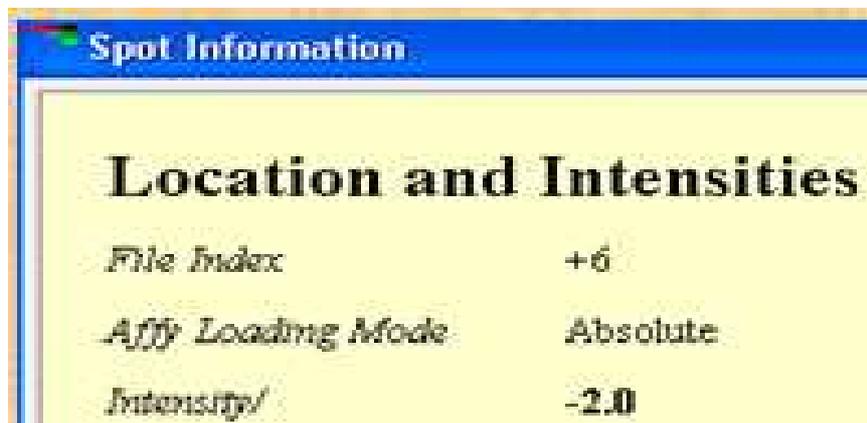
	A	B	C	D
a)	iv	iii	ii	i
b)	i	iv	iii	ii
c)	iii	iv	ii	i
d)	iii	iv	i	ii

5. **Assertion (A):** Perl is a portable scripting language [K<sub>2</sub>]

**Reason (R):** It can't be run in a Windows platform

- a) Both (A) and (R) are true, and (R) is the correct explanation of (A)      b) (A) is false, but (R) is true
- c) (A) is true and (R) is false      d) Both (A) and (R) are true, but (R) is not the correct explanation of (A)

6. Observe the output of a microarray analysis performed by MeV tool and select the correct choice: [K<sub>4</sub>]



- a) Genes are highly expressed      b) Genes are near moderately expressed
- c) Genes are moderately expressed      d) Genes are under expressed

7. Which of the following tool(s) is/ are used to predict the gene sequence? [K<sub>1</sub>]

1. HMM;      2. TMPred;      3. Geno3D;      4. GOR

- a) 1 only      b) 4 only
- c) 3 & 4      d) 1 & 3

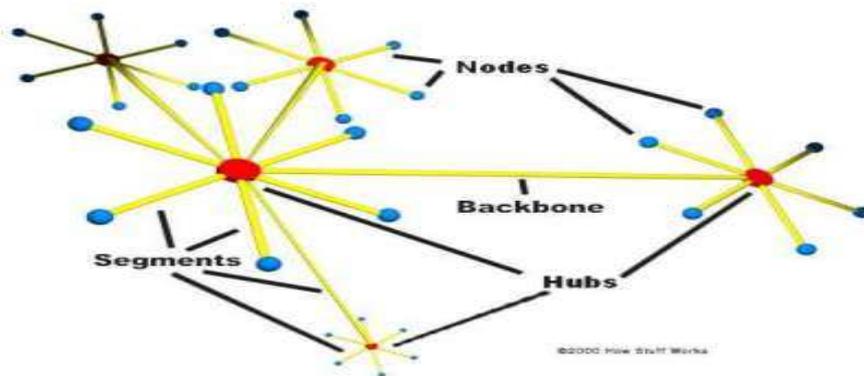
8. **Assertion (A):** In Chou Fasman method, amino acid propensity values are used to find the nucleation site [K<sub>4</sub>]

**Reason (R):** Propensity values are formulated based upon the relative molecular weight ( $M_r$ )

- a) Both (A) and (R) are true, and (R) is the correct explanation of (A)      b) (A) is true and (R) is false
- c) (A) is false, but (R) is true      d) Both (A) and (R) are true, but (R) is not the correct explanation of (A)

[K<sub>1</sub>]

9. Identify the below depicted network topology:



- a) Star
  - b) Star ring
  - c) Star bus
  - d) Star tree
10. Select the correct flow of the message in the layers of an OSI model [K<sub>3</sub>]
- 1. Session; 2. Application; 3. Physical; 4. Network; 5. Transport
  - a) 2 → 3 → 5 → 1 → 4
  - b) 2 → 1 → 5 → 4 → 3
  - c) 3 → 1 → 5 → 4 → 2
  - d) 2 → 1 → 4 → 3 → 5

**PART B (10 x 2 = 20 Marks)**

- 11. List out few basic UNIX commands. [K<sub>2</sub>]
- 12. How microarray analysis aid in the interpretation of gene expression? [K<sub>3</sub>]
- 13. Define the ftp with an example. [K<sub>2</sub>]
- 14. Give few applications of bioinformatics. [K<sub>1</sub>]
- 15. How many 3-tuples are possible for the given oligonucleotide sequence? [K<sub>3</sub>]

5'ATGCTCTTACT3'

- 16. Calculate the identity (%) and similarity (%) of the following string alignment [K<sub>4</sub>]

<b>A</b>	<b>S</b>	<b>-</b>	<b>Y</b>	<b>S</b>	<b>E</b>	<b>R</b>
			:	.		:
<b>-</b>	<b>S</b>	<b>R</b>	<b>F</b>	<b>W</b>	<b>E</b>	<b>K</b>

- 17. Compare and contrast Kimura with Jukes – Cantor model. [K<sub>3</sub>]
- 18. Spell out few iterative msa tools. [K<sub>1</sub>]
- 19. Write few applications of artificial neural nets (ANN) in computational biology. [K<sub>2</sub>]
- 20. Define homology modeling. [K<sub>2</sub>]

**PART C (10 x 5 = 50 Marks)**

- 21. Write the explanation of the following Perl operators' adopted for regular expressions of DNA and protein strings: [K<sub>4</sub>]

- a) /G....A/
- b) /GC{6,15}A/
- c) /[GCG]/
- d) /[^TA]/
- e) /A.{10}T/

22. Give the applications of systems biology. [K<sub>2</sub>]
23. Compare the advantages and limitations of various network topologies. [K<sub>3</sub>]
24. Outline about the physical media of network based communication. [K<sub>2</sub>]
25. Classify biological databases with different perspectives. [K<sub>3</sub>]
26. Discuss a suitable heuristics based database search algorithm of pairwise string alignment. [K<sub>3</sub>]
27. Write a short note 7<sup>th</sup> order HMM. [K<sub>3</sub>]
28. Summarize the algorithm proposed by Chou and Fasman for the prediction of protein secondary structures. [K<sub>2</sub>]
29. Write the algorithm of msa performed by CLUSTAL W. [K<sub>3</sub>]
30. What is SAGA? Give its algorithm. [K<sub>3</sub>]

**PART D (2 x 10 = 20 Marks)**

31. Assume that you have isolated, purified and sequenced a heptapeptide that possesses a significant anti-cancer activity from *Rhynchosia volubilis* (Black soybean, கருப்பு சோயா). According to the modified dynamic program of global alignment (EMBOSS: NEEDLE), you got the result furnished below which you felt unsatisfactory. [K<sub>5</sub>]

```
#-----
#
# Aligned_sequences: 2
# 1: EMBOSS_001
# 2: EMBOSS_001
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 7
# Identity:      2/7 (28.6%)
# Similarity:   2/7 (28.6%)
# Gaps:         3/7 (42.9%)
# Score: 8.0
..
```

So, set up your own base conditions for match, mismatch and gap, and perform the global alignment proposed by Needleman and Wunsch original algorithm for the below mentioned sequences. Compare and interpret the results of the modified program with the original program.

Query sequence = **GYTRRSE** and Db sequence = **GRTS**

32. Select six different pentanucleotide sequences of your own and construct a phylogenetic tree based upon the algorithm proposed as per UPGMA method [K<sub>6</sub>]

\*\*\*\*\*