



Name: ..... Roll No.: .....

Branch: ..... Signature of Invigilator: .....

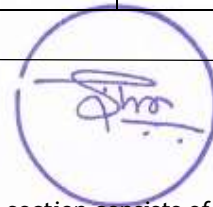
Semester: IVth

Date: 04/05/2022 (MORNING)

Subject with Code: BE205 BASICS OF BIOINFORMATICS

Marks Obtained	Section A (30)	Section B (20)	Total Marks (50)

INSTRUCTION TO CANDIDATE



1. The booklet (question paper cum answer sheet) consists of two sections. First section consists of MCQs of 30 marks. Candidates may mark the correct answer in the space provided / may also write answers in the answer sheet provided. The Second section of question paper consists of subjective questions of 20 marks. The candidates may write the answers for these questions in the answer sheets provided with the question booklet.
2. The booklet will be distributed to the candidates before 05 minutes of the examination. Candidates should write their roll no. in each page of the booklet.
3. Place the Student ID card, Registration Slip and No Dues Clearance (if applicable) on your desk. All the entries on the cover page must be filled at the specified space.
4. Carrying or using of mobile phone / any electronic gadgets (except regular scientific calculator)/chits are strictly prohibited inside the examination hall as it comes under the category of unfair means.
5. No candidate should be allowed to enter the examination hall later than 10 minutes after the commencement of examination. Candidates are not allowed to go out of the examination hall/room during the first 30 minutes and last 10 minutes of the examination.
6. Write on both side of the leaf and use pens with same ink.
7. The medium of examination is English. Answer book written in language other than English is liable to be rejected.
8. All attached sheets such as graph papers, drawing sheets etc. should be properly folded to the size of the answer book and tagged with the answer book by the candidate at least 05 minutes before the end of examination.
9. The door of examination hall will be closed 10 minutes before the end of examination. Do not leave the examination hall until the invigilators instruct you to do so.
10. Always maintain the highest level of integrity. Remember you are a BITian.
11. Candidates need to submit the question paper cum answer sheets before leaving the examination hall.

BIRLA INSTITUTE OF TECHNOLOGY, MESRA: RANCHI  
(END SEMESTER EXAMINATION)

**CLASS: BE**

**SEMESTER: IV**

**BRANCH: Open Elective**

**SESSION: 2021-22 (SP/22)**

**TIME: 2.00 Hours**

**FULL MARKS: 50**

**SUBJECT WITH CODE: BE205 Basics of Bioinformatics**

**INSTRUCTIONS:**

1. Questions number 1-22 are compulsory.
2. Answer any two questions from 29-31. If you answer all three, first two will be evaluated.
3. Pay attention to the marks of the questions.
4. The missing data, if any, may be assumed suitably.

**1** \_\_\_\_\_ is a tree in which branch lengths do not represent evolutionary time **1**

- A. Dendrogram
- B. Cladogram
- C. Phylogram
- D. None of the above

**2** The search for phylogenetic tree can be simplified by calculating \_\_\_\_\_ tree first **1**

**3** Calculate the Phylogenetic distances between the following sequences 1) ACGCGTTGG 2) ACACATTGA **1**

- A. 3
- B. 2
- C. 4
- D. 1

**4** Tips of the branches are present day species or sequences known as \_\_\_\_\_ **1**

**5** Which of the following is the character-based method? **1**

- A. UPGMA
- B. Maximum Likelihood
- C. Neighbor-Joining
- D. None of the above

**6** For unrooted trees, the number of unrooted tree topologies (NU) is \_\_\_\_\_ **2**

- A.  $NU = (2n-5)!/2n-3(n-5)!$
- B.  $NU = (2n-5)!/2n-3(n-3)!$
- C.  $NU = (2n-5)!/2-2(n-3)!$
- D.  $NU = (2n-5)!/2n(n-3)!$

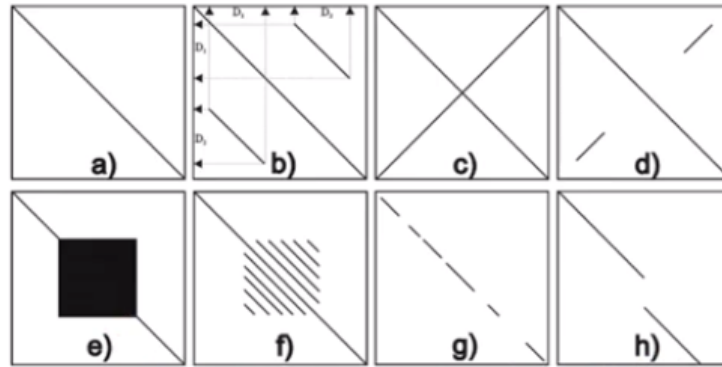
**7** The human genome contains approximately \_\_\_\_\_ **1**

- A. 6 billion bases

- B. 5 billion bases  
C. 3 billion bases  
D. 4 billion bases
- 8** Limitations with flat file format are **2**  
I. Only raw text; II. hidden instructions; III. Computers have to read all text to search for a specific entry  
A. I & II  
B. I & III  
C. II & III  
D. I, II & III
- 9** BankIt and Sequin are sequence submission tools in **1**  
A. DDBJ  
B. PDB  
C. GenBank  
D. EMBL
- 10** A gap opening penalty for any gap (g) and a gap extension penalty for each element in the gap (r) are most often used, to give a total gap score wx, according to the equation \_\_\_\_\_ **2**  
A.  $wx - rx = -g$   
B.  $wx = g - rx$   
C.  $wx = g + rx$   
D.  $wx + g + rx = 0$
- 11** Which of the following is untrue about Backbone Model Building Step? **1**  
A. Once optimal alignment is achieved, residues in the aligned regions of the target protein can assume a similar structure as the template proteins  
B. Coordinates of the corresponding residues of the template proteins can be simply copied onto the target protein  
C. If the two residues differ, everything other than the backbone atoms can be copied  
D. If the two aligned residues are identical, coordinates of the side chain atoms are copied along with the main chain atoms
- 12** In Smith- Waterman algorithm, in initialization step, the \_\_\_\_\_ row and \_\_\_\_\_ column are subject to gap penalty. **2**  
A. first, second  
B. second, first  
C. first, first  
D. second, second
- 13** In 3-D alignment cell, how many edges forms in each unit cube? **1**  
A. 3

- B. 5  
C. 7  
D. 9
- 14 How many total numbers of words (approximate) to search for 225 amino acids long sequence during BLAST search? **2**
- A. 11,000  
B. 11,250  
C. 12,000  
D. 12,500
- 15 Which of the following options contain only those amino acids that are likely to be present in the interior of a protein? **2**
- A. Arg, Val, Met  
B. Asp, Met, Phe  
C. Lys, Phe, Val  
D. Val, Met, Phe
- 16 Which of the following represents the two-dimensional structure of proteins? **1**
- A. Quaternary  
B. Tertiary  
C. Secondary  
D. Primary
- 17 Ramachandran plot can be used to predict which of the following structure? **1**
- A. Primary structure  
B. Secondary structure  
C. Tertiary structure  
D. Quaternary structure
- 18 Right-handed alpha-helix allowed region is present in which of the following quadrants of Ramachandran plot? **2**
- A. First quadrant  
B. Second quadrant  
C. Third quadrant  
D. Fourth quadrant
- 19 The angle about the  $C\alpha$  - N is denoted by **1**
- A.  $\omega$   
B.  $\Phi$   
C.  $\Psi$   
D.  $\pi$
- 20 The reemergence of “systems thinking” linked to the post-genomic era and the development of global molecular profiling methods collectively known as **1**

- A. Systemic property  
 B. Omics technologies  
 C. Molecular system  
 D. None of these
- 21 "Once an error, always an error" is the problem of \_\_\_\_\_ alignment method 1
- A. Dynamic Programming  
 B. Iterative  
 C. Progressive  
 D. Block-based
- 22 Arrange the steps of homology modeling in correct order 1) Template Selection 2) Model validation 3) Loop modeling 4) Backbone generation 2
- A. 1->2->3->4  
 B. 1->4->3->2  
 C. 1->4->2->3  
 D. 4->3->2->1
- 23 Explain the chou-fasman method for helix prediction. 2
- 24 what are the criteria to identify the informative sites in Maximum Parsimony Method? Explain with one example. 2
- 25 What is Homoplasy? Name the two statistical models used to correct Homoplasy. 2
- 26 Why protein sequences allow more sensitive alignment than DNA sequences? Explain with one example. 2
- 27 For the evolutionally closely related sequence, which of the following substitution matrix will you use and why? BLOSUM62, BLOSUM80, PAM40 and PAM250. 2
- 28 What is system biology? Discuss its importance in life science research. 2
- 29 Enlist all the steps of BLAST algorithms in sequential order and explain them in detail. Also explain the E-value and bit-score. 4
- 30 Generate the alignment for the following sequences using the Needleman-Wunsch algorithm. Use the BLOSUM-62 AA substitution matrix (attached with the question paper) for match and mismatch data and consider gap penalty -8. **Sequences: HEAGAWGHEE and PAWHEAE.** 4
- 31 Explain the following dot matrix plots with appropriate hypothetical sequences. 4



**Appendix: BLOSUM 62 AA Substitution Matrix**

C	9																			
S	-1	4																		
T	-1	1	5																	
P	-3	-1	-1	7																
A	0	1	0	-1	4															
G	-3	0	-2	-2	0	6														
N	-3	1	0	-2	-2	0	6													
D	-3	0	-1	-1	-2	-1	1	6												
E	-4	0	-1	-1	-1	-2	0	2	5											
Q	-3	0	-1	-1	-1	-2	0	0	2	5										
H	-3	-1	-2	-2	-2	-2	1	-1	0	0	8									
R	-3	-1	-1	-2	-1	-2	0	-2	0	1	0	5								
K	-3	0	-1	-1	-1	-2	0	-1	1	1	-1	2	5							
M	-1	-1	-1	-2	-1	-3	-2	-3	-2	0	-2	-1	-1	5						
I	-1	-2	-1	-3	-1	-4	-3	-3	-3	-3	-3	-3	1	4						
L	-1	-2	-1	-3	-1	-4	-3	-4	-3	-2	-3	-2	-2	2	2	4				
V	-1	-2	0	-2	0	-3	-3	-3	-2	-2	-3	3	2	1	3	1	4			
F	-2	-2	-2	-4	-2	-3	-3	-3	-3	-3	-1	-3	-3	0	0	0	-1	6		
Y	-2	-2	-2	-3	-2	-3	-2	-3	-2	-1	2	-2	-2	-1	-1	-1	-1	3	7	
W	-2	-3	-2	-4	-3	-2	-4	-4	-3	-2	-2	-3	-3	-1	-3	-2	-3	1	2	11
	C	S	T	P	A	G	N	D	E	Q	H	R	K	M	I	L	V	F	Y	W