



Date:

Registration number:

ST. JOSEPH'S COLLEGE (AUTONOMOUS), BANGALORE – 27  
M.SC(BIG DATA ANALYTICS) – III SEMESTER  
SEMESTER EXAMINATION – OCTOBER 2021  
(Examination conducted in January-March 2022)  
**BDADE 3621 - BIOINFORMATICS**

**TIME: 2.5 HOURS**

**MAX MARKS: 70**

**This Paper contain FOUR printed pages and THREE parts**

**PART A**

**Answer ALL questions**

**20 X 1 = 20**

1. \_\_\_\_\_ is a primary sequence database and accepts nucleotide sequences from authors.  
A) RefSeq  
B) GenBank  
C) SwissProt  
D) UniProt
2. What do you mean by 'Query coverage' in BLAST?  
A) Percentage of query covered in the subject  
B) Percentage of query covered in the database  
C) Percentage of sequences in the database  
D) Percentage of bases in the database
3. \_\_\_\_\_ is present only in eukaryotes.  
A) Cell membrane  
B) Ribosome  
C) Cell wall  
D) Endoplasmic reticulum
4. For which of the following functions, Bowtie tool can be used?  
A) Aligning reads to reference  
B) Aligning reads with each other  
C) Searching for similar sequences  
D) Searching for dissimilar sequences
5. \_\_\_\_\_ will help you in matching a nucleotide query sequence with a protein subject sequence.  
A) BLASTN  
B) BLASTP  
C) Both A and B  
D) BLASTX

6. Multiple sequence alignment helps in
  - A) Finding closely related species
  - B) Finding unrelated species
  - C) Both A and B
  - D) Finding new species
  
7. Which among the following is the first sequencing method to be invented?
  - A) Maxam and Gilbert method
  - B) Watson's method
  - C) Sanger's method
  - D) Craig venter's method
  
8. What would be your next step once you get the human DNA sequences from the sequencing machine?
  - A) Find exons in the sequences
  - B) Align them to the reference genome
  - C) Align them with each other
  - D) Find for the known DNA motifs
  
9. What is the full form of PDB?
  - A) Protein Data Bank
  - B) Protein Data Browser
  - C) Primary Data Bank
  - D) Primary Data Base
  
10. NM\_12345 identifier belongs to which database?
  - A) RefSeq
  - B) GenBank
  - C) SwissProt
  - D) UniProt
  
11. Emulsion-PCR amplification method is used in \_\_\_\_\_
  - A) Illumina sequencers
  - B) Roche 454 sequencers
  - C) Pacific Biosciences sequencers
  - D) All the above
  
12. Which of the following is also known as Third generation sequencers?
  - A) Roche 454
  - B) Illumina Genome Analyzer
  - C) Pacific Biosciences sequencers
  - D) Illumina HiSeq sequencers
  
13. \_\_\_\_\_ is NOT a protein structure prediction method.
  - A) Ab initio
  - B) In silico
  - C) Threading

- D) Homology modeling
14. Which of these is a multiple sequence alignment tool?  
A) ClustalW  
B) BLAST  
C) GeneCards  
D) PDB
15. Which E-value indicates the most significant match in BLAST?  
A)  $1.31E-5$   
B)  $3.31E-7$   
C) 0.01  
D)  $3.31E-8$
16. The term 'read' in an NGS experiment indicates  
A) A raw sequence that comes from a sequencing machine  
B) Complete DNA sequence  
C) Method of reading the bases in the DNA  
D) Mapped reference sequences
17. Bridge amplification method is used in in which sequencer?  
A) Illumina sequencers  
B) Roche 454 sequencers  
C) Pacific Biosciences sequencers  
D) All the above
18. Sanger sequencing is based on  
A) Chain termination  
B) Di-deoxy nucleotides  
C) Pyrosequencing  
D) All the above
19. GeneCards database provides only human related gene information  
A) TRUE  
B) FALSE
20. Genomics means \_\_\_\_\_  
A) Studying all genes and its contents  
B) Studying all transcripts  
C) Studying all mRNAs  
D) Studying all proteins

**PART B**

**Answer ANY SIX questions**

**6 X 5 = 30**

21. Write a detailed note on the usage of any 5 bioinformatics databases/tools in biological research.
22. Write a note on central dogma? Discuss how it is used by the cells and propagated?
23. Discuss the different applications of next generation sequencing methods in healthcare and research.
24. Define genomics, transcriptomics, and proteomics along with their applications. Write a note on various types of high-throughput data that are generated in these fields.
25. What are the major differences between a bacterial and human cell? Write the major functions of at least 3 cell organelles that are present in human.
26. Discuss three different protein structure prediction methods?
27. Write a note on the differences between RefSeq and GenBank databases.
28. Write a note on the various applications of multiple sequence alignment? List various features of ClustalOmega tool.

### **PART C**

**Answer ANY TWO questions**

**2 X 10 = 20**

29. Discuss in detail the different steps of homology modeling.
30. What is reference alignment? Discuss the methodology used by Illumina and Roche sequencers.
31. What is the importance of sequence analysis? Discuss various features of BLAST tool and important parameters for choosing the best alignment.