(University of Choice)

MASINDE MULIRO UNIVERSITY OF SCIENCE AND TECHNOLOGY (MMUST)

MAIN CAMPUS

UNIVERSITY EXAMINATIONS
2019/2020 ACADEMIC YEAR

THIRD YEAR SECOND SEMESTER EXAMINATIONS

FOR THE DIPLOMA
IN
MEDICAL BIOTECHNOLOGY

MAIN EXAM

COURSE CODE: BBD 323

COURSE TITLE: FUNDAMENTALS OF BIOINFORMATICS & BIOCOMPUTING

DATE: 7TH DECEMBER 2020   **TIME**: 8.00 -10.00AM

INSTRUCTIONS TO CANDIDATES

This paper is divided into three sections, **A B** and **C,** carrying respectively: Multiple Choice Questions (**MCQs**), short answer questions (**SAQs**) and long answer questions (**LAQs**).

 TIME: 2 Hours

MMUST observes ZERO tolerance to examination cheating

This Paper Consists of 5 Printed Pages. Please Turn Over.

▶

SECTION A: MULTIPLE CHOICE QUESTIONS (MCQs)

1. Which of the following is a nucleotide repository program
    a) SwissProt
    b) Uniprot-TrEMBL
    c) GenBank
    d) Entrez

2. Which of the following describes the main reason for doing a sequence alignment?
    a) Allow phylogenetic tree estimation.
    b) Illuminate conserved elements important to function
    c) Inform structural predictions
    d) To judge the similarities and differences between two or more sequences.

3. ClustalW is a suitable software for computational analysis of
    a) Pairwise sequence alignment
    b) Multiple sequence alignment
    c) Gene expression profiles
    d) Protein folding into tertiary structures

4. BLOSUM model is
    a) Database for expression profiles of genes
    b) Based on log-odds (lod) scores.
    c) Computational Software for protein analysis
    d) Bank for expression regulatory molecules

5. An algorithm is best defined as;
    a) Are a set or ordered steps, essentially a recipe, for solving a problem.
    b) Is a way of solving problems where you make the best decision at each step along the way to your solution.
    c) A short cut that makes a problem tractable and that will give us a good solution most of the time.
    d) A statistical score, based on the extreme value distribution, used to interpret the results of database searches.

6. What does OMIM represent in bioinformatics?
    a) Multiple sequence alignment tool
    b) Organization of mammalian genetic trees
    c) Database for Mendelian inheritance in man
    d) Objects for microorganism inheritance

7. HSP as applied in bioinformatics is a
    a) Primer design program
    b) Data bank for proteins and their structures
    c) A high scoring segment pair
    d) Phosphorylating dimer of biomolecules

8. Orthologous genes

a) Similar genes in the same genus that result from gene duplication.
b) Genes that have lost function over time in ancestry.
c) Similar genes in different genus that results from gene ancestral inheritance.
d) Acquired genes from other organisms that were not part of the genus ancestry.

9. Which of the following best describes a reference sequence?
a) Is a curated collection of DNA, RNA, and protein sequences built by the NCBI.
b) Set of unique genes that are conserved across the genus.
c) A set of non-essential genes in an organism
d) The genome that is found in all organisms

10. Which of the following is a BLAST program for a protein sequence in a protein database?
a) BLASTp
b) BLASTn
c) BLASTx
d) tBLASTn

11. Which of the following is NOT a characteristic of entries in a reference sequence
a) NCBI staff hand-curate entries using data sources including GenBank.
b) Records are revised by NCBI staff as additional information becomes available.
c) The genome that is found in all organisms is available
d) Only information from major model organisms are included in Reference sequence.

12. A bioinformatician would like to study expressed genes and their abundance. Where can he find this information
a) ENTREZ
b) UNI GENE
c) BLAST
d) Protein data bank

13. During computational translation the software will stop the peptide sequence translation when it encounters the following EXCEPT?
a) AUG
b) UGA
c) UAG
d) UAA

14. The following statement is true about Hidden Markov Model
a) Align two groups of sequences which share a common domain
b) It tends to be successful at detecting and aligning critical motifs and conserved core structure of protein families.
c) It is a post-transcription process
d) Subgroups with significant divergence may not be correctly aligned to the consensus structure

15. Which of the following is a BLAST program for a translated nucleotide sequence in protein database?

a) BLASTp
b) BLASTn
c) BLASTx
d) tBLASTn

16. Bioinformatics and biocomputation field involving the following EXCEPT?
    a) The discovery of therapeutic products
    b) Biotechnology and recombinant technology
    c) The study of molecular pathways and drug resistance
    d) Improving the genetic code

17. Which of the following is not TRUE about ClustalW?
    a) Involves construction of pairwise alignments.
    b) It was developed by Feng and Doolittle
    c) Initial step involves progressive alignment guided by the tree.
    d) 'W' stands for weighted

18. Which of the following is NOT an example of a nucleic acid models?
    a) Jukes-Cantor model (JC).
    b) Kimura two-parameter model (K2P).
    c) Hasegawa, Kishino and Yano model (HKY).
    d) PAM model

19. Which of the following is NOT a functional macromolecule in cellular processes?
    a) mRNA
    b) rRNA
    c) tRNA
    d) Ribosomal RNA

20. Which of the following is a BLAST program for a protein sequence in translated nucleotide database?
    a) tBLASTx
    b) BLASTn
    c) BLASTx
    d) tBLASTn

SECTION B: SHORT ANSWER QUESTIONS

1) Using specific examples, discuss the three main nucleotide repositories   (**7 Marks**)

2) Compare and contrast between the PAM and BLOSUM amino acid sequence alignment model.   (**6 Marks**)

3) Describe the following;   (**8 Marks**)
    a) Evolutionary models
    b) Amino Acid models

4) Briefly, write short notes on heuristics and dynamic programming sequence alignment (**10 Marks**)

5) Define the following terminologies;
     A Query sequence
     B Subject sequence
     C Raw score   (**9 Marks**)

SECTION C: LONG ANSWER QUESTIONS (40 Marks)

1) Discuss the major points to consider before submitting sequences to databases (**20 Marks**)

2) With the help of examples, discuss the genomic map elements   (**20 Marks**)

3) Compare and contrast the various common algorithms used in bioinformatics **(20 Marks).**