R8365	Sub. Code
	502201

M.Sc. DEGREE EXAMINATION, APRIL – 2023

Second Semester

Bioinformatics

PHYLOGENY AND PHYLOGENOMICS

(CBCS – 2022 onwards)

Time : 3 Hours Maximum : 75 Marks

Part A $(10 \times 1 = 10)$

Answer **all** questions.

1. The most commonly used tool for phylogenetic analysis involves the sequencing of

   (a) Mitochondrial DNA.

   (b) Ribosomal RNA.

   (c) Nuclear DNA.

   (d) Mitochondrial RNA.

2. Who proposed the concept of the molecular clock?

   (a) Emile Zuckerkandl and Linus Pauling.

   (b) Susumo Ohno.

   (c) Motoo Kimura.

   (d) Charles Darwin and Alfred Russel Wallace.

3. Alignment algorithms, both global and local, are fundamentally similar and only differ in the optimization strategy used in aligning similar residues.

   (a) True (b) False

4. Which of the following is untrue regarding BLAST and FASTA?

(a) FASTA is faster than BLAST.

(b) FASTA is the most accurate.

(c) BLAST has limited choices of databases.

(d) FASTA is more sensitive to DNA-DNA comparisons.

5. Which of the following does not describe the global alignment algorithm?

(a) It attempts to align every residue in every sequence.

(b) It is most useful when the aligning sequences are similar and roughly the same size.

(c) It is useful when the aligning sequences are dissimilar.

(d) It can use the Needleman-Wunsch algorithm.

6. _____ is a short conserved sequence pattern associated with distinct functions of a protein or DNA.

(a) Profile (b) Blocks

(c) Domain (d) Motif

7. In regular expression, which of the following pair of pattern is wrongly matched with its significance?

(a) [ ] – Or (b) { } – Not

(c) ( )– Repeats (d) Z – Any

8. What is the typical unit that we use when describing estimates of evolutionary rate?

(a) Mutations/site/year

(b) Substitutions/site

(c) Substitutions/site/year

(d) Mutations/site

9. Which of the following is not a character-based phylogenetic tree construction method?

(a) Weighted Parsimony

(b) Maximum Likelihood

(c) Maximum Parsimony

(d) Minimum evolution

10. Which of these is the gene prediction algorithm?

(a) UPGMA

(b) Hidden Markov Model

(c) Maximum Parsimony

(d) None

Part B $(5 \times 5 = 25)$

Answer **all** questions, choosing either (a) or (b).

11. (a) Explain the concept of the Molecular clock and neutral evolution.

Or

(b) Describe convergent evolution with a suitable example.

12. (a) Write about the significance of sequence alignment and list some sequence alignment tools.

Or

(b) Compare BLOSUM and PAM matrices.

13. (a) Write a short note on regular expressions.

Or

(b) How would you implement the artificial neural network in gene discovery?

14. (a) What is a phylogenetic tree, and how would you construct it?

Or

(b) Differentiate distance-based and character-based methods of phylogenetic tree construction.

15. (a) List any five software available for phylogenetic analysis and mention its application.

Or

(b) Explain the steps in construction of phylogenetic trees .

Part C $(5 \times 8 = 40)$

Answer any **five** questions.

16. Explain the concept of speciation and its types with suitable examples.
17. Write about amino acid scoring matrices in detail with a suitable example.
18. Demonstrate Needleman-Wunch and Smith-Waterman algorithms with examples.
19. Discuss various methods of multiple sequence alignment.
20. Define patterns, motifs and blocks. Explain different sequence pattern representations.
21. Discuss the Hidden Markov model and its application in detail.
22. Explain any one distance-based method of phylogenetic tree construction with an example.
23. Write a detailed note on any two software available for phylogenetic analysis. Mention its algorithm, result interpretation along with its significance in sequence analysis.

R8366	Sub. Code
	502202

M.Sc. DEGREE EXAMINATION, APRIL – 2023

Second Semester

Bioinformatics

MOLECULAR MODELLING AND DRUG DESIGN

(CBCS – 2022 onwards)

Time : 3 Hours Maximum : 75 Marks

Part A $(10 \times 1 = 10)$

Answer **all** the questions.

1. What is meant by a lead compound in drug discovery?

   (a) A drug containing the element lead.

   (b) A leading drug in a particular area of medicine.

   (c) A compound that acts as the starting point for drug design and development.

   (d) A drug which is normally the first to be prescribed for a particular ailment.

2. Natural products are often used as lead compounds in the design and synthesis of novel drugs. Which of the following general characteristics of a natural product is most likely to be a disadvantage in synthesizing analogues?

   (a) Novelty of structure.

   (b) Complexity of structure.

   (c) Level of activity.

   (d) Availability.

3. Which of the following operations or calculations would generally be carried out using quantum mechanics?

(a) Energy minimization.

(b) Identifying stable conformations.

(c) Partial atomic charges.

(d) Energy calculations for specific conformations.

4. Which of the following statements regarding molecular mechanics is untrue?

(a) It treats atoms as spheres.

(b) It treats bonds as rigid features.

(c) It is a molecular modelling computational method.

(d) It uses equations that obey the laws of classical physics.

5. Beta-sheets allowed region is present in which of the following quadrants of the Ramachandran plot?

(a) First quadrant (b) Second quadrant

(c) Third quadrant (d) Fourth quadrant

6. The Chou-Fasman algorithm determines the propensity or intrinsic tendency of each residue to be in the helix, strand, and beta-turn conformation using observed frequencies found in protein crystal structures.

(a) True (b) False

7. One of the following is not used in QSAR:

(a) Molecular connectivity index

(b) Molecular similarity index

(c) Topological polar surface area

(d) Partition coefficient

8. The first set of molecules under investigation can be filtered out by

(a) Docking

(b) In vivo biological activity evaluation

(c) Lipinski's rule

(d) Toxicity evaluation

9. What does this mean if a drug has a 'small therapeutic index'?

(a) A bigger dose is required to get the drug affect.

(b) The drug is not effective.

(c) There is little difference between the dose for the right affect and the dose to be toxic.

(d) There is a big different between the dose for the right affect and the dose to be toxic.

10. Which of the following routes of administration do not bypass the first pass metabolism?

(a) Intravenous (b) Oral

(c) Sublingual (d) Transdermal

Part B (5 × 5 = 25)

Answer **all** questions, choosing either (a) or (b).

11. (a) How computational methods can aid in the lead optimization of a molecule? Explain.

Or

(b) Discuss different approaches for target identification and validation.

12. (a) Discuss about the features of Potential Energy Surface and the condition for its maximum and minimum points.

Or

(b) Elaborate on the types of bonded interactions revealed by a molecule.

13. (a) Elaborate on tools available for structure validation.

Or

(b) How will you predict secondary structure of a protein? Explain.

14. (a) How will you build a QSAR model? Explain.

Or

(b) Describe the basic concepts and the steps involved in de novo drug design.

15. (a) Write a note on Lipinski's Rule of five. How important is this rule in current drug discovery scenario?

Or

(b) Write a note on drug-receptor interactions.

Part C

$(5 \times 8 = 40)$

Answer any **five** questions.

16. What are the different stages through which a new molecule passes in drug development process? Explain with flow chart.

17. Write a detailed note on force field parameters.

18. Define energy minimization. Discuss any one method of energy minimization and its significance.

19. Explain the methods available for predicting tertiary structure of protein.

20. You are provided with a virtual molecular library of 1 lakh compounds. How will you find a lead molecule from the library? Explain in detail with flow charts.

21. Explain how docking is useful for computer-aided drug design. Discuss various docking algorithms in detail.

22. Discuss in detail the pharmacokinetics and pharmacodynamics of drug action.

23. (a) Explain the concept of hard and soft drugs.

(b) Elaborate on ADMET properties of drugs.

———————

R8367	Sub. Code
	502203

M.Sc. DEGREE EXAMINATION, APRIL – 2023

Second Semester

Bioinformatics

COMPUTATIONAL BIOLOGY

(CBCS – 2022 onwards)

Time : 3 Hours Maximum : 75 Marks

Part A (10 × 1= 10)

Answer **all** the questions.

1. What types of data are needed for approval of a biosimilar?
   (a) Comprehensive analytical studies, which form the foundation to demonstrate biosimilarity to an FDA-approved reference product. This includes data on the structure, purity, and bioactivity of the molecule.
   (b) Comparative animal studies to assess toxicity of the biosimilar product as compared to the reference product, if needed.
   (c) Comparative clinical study (or studies), if needed, to demonstrate safety, purity, and potency in one or more appropriate conditions of use for which licensure is sought for the biosimilar product. This often includes a comparison of pharmacokinetics (PK), pharmacodynamics (PD) if applicable; and may include comparison of clinical effect, safety, and immunogenicity.
   (d) Clinical trials to establish efficacy and safety in each indication for which licensure is sought.
   (e) a, b, and c.

2. Which statement is false about biosimilar and interchangeable biosimilar products?

(a) Approved biosimilar products can be used in patients who are treatment-experienced (i.e.. already used the reference product or biosimilar product) and treatment-naïve (i.e., never used the reference product or biosimilar product).

(b) Biosimilar products are expected to have the same rate of immunogenic response as the reference product. If a patient doesn't respond to the reference product or develops anti-drug antibodies, the same can be expected with the biosimilar.

(c) FDA may require additional studies (switching studies) to show that a biosimilar is interchangeable with its reference product.

(d) An interchangeable biosimilar product may be substituted without intervention of the healthcare provider subject to state laws.

(e) Interchangeable biosimilar products are safer and more effective than a regular biosimilar product.

3. Local alignments are more used when _____

(a) There are totally similar and equal length sequences

(b) Dissimilar sequences are suspected to contain regions of similarity

(c) Similar sequence motif with larger sequence context

(d) Partially similar, different length and conserved region containing sequences

4. Which of the following does not describe k-tuple methods?

   (a) k-tuple methods are best known for their implementation in the database search tools FASTA and the BLAST family.

   (b) They are also known as words methods

   (c) They are basically heuristic methods to find local alignment

   (d) They are useful in small scale databases

5. The more conserved amino acids in similar proteins from different species are ones that play an essential role in structure and function and the less conserved are in sites that can vary without having a significant effect on function.

   (a) True (b) False

6. Which of the following is not a description of dynamic programming algorithm?

   (a) A method of sequence alignment

   (b) A method that can take gaps into account

   (c) A method that requires a manageable number of comparisons

   (d) This method doesn't provide an optimal alignment

7. For palindromic sequences, what is the structure of the dot plot?

   (a) 2 intersecting diagonal lines at the midpoint

   (b) One diagonal

   (c) Two parallel diagonals

   (d) No diagonal

8. Which of the following is not a distance-based phylogenetic tree construction method?

(a) Neighbour-joining method.

(b) UPGMA.

(c) Maximum Parsimony.

(d) Minimum evolution.

9. What is meant by synthetic biology?

(a) The synthesis of naturally occurring compounds.

(b) The genetic modification of microbial cells such that they produce compounds that they would not normally produce.

(c) The study of how naturally occurring compounds are biosynthesized in cells

(d) The design and synthesis of enzymes capable of carrying out a specific reaction in organic synthesis

10. To solve Schrödinger equation we need potential and

(a) physical requirements of system

(b) boundary condition

(c) none of these

(d) both (a) and (b)

Part B $(5 \times 5 = 25)$

Answer **all** questions, choosing either (a) or (b).

11. (a) Define biosimilars and explain its molecular complexity.

Or

(b) Discuss Critical manufacturing parameters of biosimilars and Challenges.

12. (a) Write a note on nature and scope of computational biology.

Or

(b) Write a note on application of substitution matrices.

13. (a) Explain dynamic programming method with an example.

Or

(b) Write a note on dot plot and how will interpret it.

14. (a) Write a note on tools available for multiple sequence alignment and mention the application of MSA,

Or

(b) Define restriction map and write a short note on Cassettle transformation of restriction map.

15. (a) Describe Huckel theory with an example.

Or

(b) What is synthetic biology? List out the ethical issues of synthetic biology.

Part C $(5 \times 8 = 40)$

Answer any **five** questions.

16. Explain the development process of biosimilars.

17. Describe global and local alignment with an example and Compare them.

18. Explain any one of the substitution matrices with an example.

19. Elaborate on hidden markov model.

20. Discuss the algorithms used in multiple sequence alignment.

21. Elaborate on restriction maps.

22. Define Molecular Orbitals and explain Molecular Orbital theory.

23. Elaborate on computational synthetic biology and its application.

———————

R8368	Sub. Code
	502204

M.Sc. DEGREE EXAMINATION, APRIL – 2023

Second Semester

Bioinformatics

PROGRAMMING IN SCRIPTING LANGUAGES (PYTHON, PERL & R)

(CBCS – 2022 onwards)

Time : 3 Hours Maximum : 75 Marks

Part A $(10 \times 1 = 10)$

Answer **all** questions.

1. The creator of Perl programming language is ————.

   (a) James Gosling (b) Brendan Eich

   (c) Larry Wall (d) BjarneStroustrup

2. What is the syntax to create multiline comments in Perl?

   (a) = start
= end (b) # begin
cut

   (c) = begin
= cut (d) None of these

3. Which type of Programming does Python support?

   (a) object-oriented programming

   (b) structured programming

   (c) functional programming

   (d) all of the mentioned

4. Which of the following functions is a built-in function in python?

(a) factorial ( ) (b) print ( )

(c) seed ( ) (d) sqrt ( )

5. Which module in Python supports regular expressions?

(a) Re (b) Regex

(c) Pyregex (d) none of the mentioned

6. Which of the following creates a pattern object?

(a) re.create(str) (b) re.regex (str)

(c) re.compile (str) (d) re.assemble (str)

7. Which of the following creates a tuple?

(a) tuple1 = (5) * 2

(b) tuple1 = ("a","b")

(c) tuple1[2] = ("a","b")

(d) None of the above

8. To read two characters from a file object in file, we use _____

(a) infile.read (2) (b) infile.read ( )

(c) infile.readline ( ) (d) infile.readlines ( )

9. R functionality 1S divided into a number of _____

(a) Packages (b) Functions

(c) Domains (d) Classes

10. Which of the following is used to plot multiple histograms?

(a) multi.plot ( ) (b) multi.hist

(c) xyplot.multi ( ) (d) poly ( )

Part B (5 × 5 = 25)

Answer **all** questions, choosing either (a) or (b).

11. (a) Compare auto-increment and auto-decrement in Perl.

Or

(b) Describe the operators available in Perl.

12. (a) Discuss about variables of different types in Python.

Or

(b) How do you define function in Python? Explain.

13. (a) Compare call by value and call by reference.

Or

(b) Differentiate Local and Global special variables.

14. (a) Compare Positive and Negative indexing in tuple.

Or

(b) What are the methods that are used in python tuple?

15. (a) What are the clustering techniques available in data science? Explain.

Or

(b) Describe the procedure to read and write a data in R.

Part C (5 × 8 = 40)

Answer any **five** questions.

16. Illustrate about Control Structures in Perl with example.

17. Write a detailed note on Conditional Statements in Python.

18. Explain the procedure to create list as Array using python.

19. Discuss about special variables and its types in detail.

20. Write the procedure to use split and join functions with example.

21. Explain how to create dictionary in python with example.

22. Describe the methods used in python list with example.

23. Explain how to manipulate data in R.

———————