Strathmore
UNIVERSITY

STRATHMORE INSTITUTE OF MATHEMATICAL SCIENCES

MASTER OF SCIENCE IN BIOMATHEMATICS

END OF SEMESTER EXAMINATIONS

BMA 8204: INTRODUCTION TO COMPUTATIONAL STRUCTURAL BIOLOGY

DATE: 6th DECEMBER 2023 **TIME: 3 HOURS**

INSTRUCTIONS: Answer Question ONE and ANY other TWO questions.

QUESTION ONE (30 MARKS)

(a) Define the following terms.

(i) Codon **[1 mark]**

(ii) Open reading frame **[1 mark]**

(iii) Nucleotide **[1 mark]**

(b) Show that for a simple brute force algorithm, that the number of possible alignments of two sequences of lengths n and m is approximated by $\frac{2^{2n}}{\sqrt{\pi n}}$. **[3 marks]**

(c) Consider the following sequences

AGGCTATCACCTGACCTCCAGGCCGATGCCC

TAGCTATCACGACCGCGGTCGATTTGCCGAC

Determine

(i) Sequence identity. **[3 marks]**

(ii) Sequence similarity. **[3 marks]**

(d) Distinguish the following terms.

(i) Paralogous sequence and orthologous sequence. **[2 marks]**

(ii) global alignment and multiple alignment. **[2 marks]**

(e) Outline how random walk concept is used in cellular level simulation. **[3 marks]**

(f) Outline how to calculate the log odds scoring of an alignment. **[3 marks]**

(g) Briefly describe some of the limitations of gene predicting tools. **[2 marks]**

(h) Compare accepted point mutation model (PAM) and block substitution matrices (BLOSUM) used in sequence alignment.

[3 marks]

(i) Describe the Hidden Markov chain models. **[3 marks]**

QUESTION TWO (15 MARKS)

(a) Determine the optimal global alignment of the two sequences given by

CTCGCAGC

CATTTCAC

If the scoring is -1 for mismatch, +1 for a match and -2 for a gap. Show the scoring matrix.

[7 marks]

(b) Define the following terms and give an example for each

(i) Sub-sequence. **[2 marks]**

(ii) Substring. **[2 marks]**

(c) Outline the FASTA procedure. **[4 marks]**

QUESTION THREE (15 MARKS)

(a) Outline the equations of motion of molecular dynamics and outline the limitations of molecular dynamics simulations.

[6 marks]

(b) Describe the process of protein synthesis

[5 marks]

(c) Describe any three next generation sequencing technologies. **[4 marks]**

QUESTION FOUR (15 MARKS)

(a) Describe the steps of accepted point mutation model development. **[5 marks]**

(b) Outline the Needleman-Wunsch algorithm for sequence alignment. **[5 marks]**

(c) Define the term memoization. **[1 marks]**

(d) Describe how memoization has been achieved in this code below. **[4 marks]**

```
def fib(n):
    fib_table = [1, 1] + (n-1)*[0]
    def get(i):
        if fib_table[i] != 0:
            return fib_table[i]
        else:
            fib_table[i] = get(i-1) + get(i-2)
        return fib_table[i]
        if n == 0 or n == 1:
            return 1
        else:
            return get(n-1) + get(n-2)
```