Uva Wellassa University, Sri Lanka
First Semester Examination – March 2011
CST 411-3 Bioinformatics

Time: **Three (03)** hours | Total **05** Questions | Answer **all** questions

---

Question 1. **(20 marks)**

(a) **"In DNA replication, the leading strand is synthesized continuously while the lagging strand is synthesized discontinuously"** Explain the above statement.

(b) Explain the **Central Dogma** of molecular biology.

(c) Mark the six possible reading frames of the following nucleotide sequence. Mark the 5' end and 3' end of your sequences.

5'- ATGCTAGCTGTAGCTGCATGA- 3'

(d) **Cloning** and **Polymerase Chain Reaction (PCR)** are two different techniques used in DNA amplification.

- I. Explain major steps of amplifying DNA by cloning.
- II. What are the key parts of a vector?
- III. Explain major steps of amplifying DNA by PCR.
- IV. What are the key ingredients of PCR?

(e) The gene for hemoglobin, one allele codes for normal hemoglobin, while another allele codes for altered hemoglobin, called sickle cell hemoglobin. When a person is homozygous for this sickle cell allele, this causes a serious disease called sickle cell anemia. Suppose that a person who is heterozygous for the sickle-cell allele marries a person who is also heterozygous for this allele.

- I. Draw a Punnett Square to show the expected genetic makeup of their children.
- II. What fraction of their children will suffer from sickle cell anemia?
- III. What fraction of their children will inherit the sickle cell allele, but will not suffer from sickle cell anemia?

23

Question 2. **(20 marks)**

(a) Briefly explain the biological motivations behind "**Sequence Alignment**".

(b) Using the **Smith and Waterman** dynamic programming method, construct the partial alignment score table for the following two sequences, using following scoring parameters: match score = +1, mismatch score = -1, gap penalty = -2.

MATCHES
THATCHER

Write down the local alignment of these two sequences.

(c) Using the **Needle and Wunsch** dynamic programming method, construct the partial alignment score table for the following two sequences, using following scoring parameters: match score = +1, mismatch score = -1, gap penalty = -2.

ACAGTAG
ACTCG

Write down the global alignment of these two sequences.

(d) "**Heuristic methods are much faster than the dynamic methods**". What are the pros and cons of heuristic methods and dynamic methods when searching similar sequences in bioinformatics databases?

(e) Briefly explain the use of **Scoring Matrices**. Name two Scoring Matrices use in Bioinformatics.

Question 3. **(20 marks)**

(a) What does it mean by a "**phylogenetic tree**"?

(b) Name two types of phylogenetic trees. What kind of information that can be extracted from each type?

(c) Briefly explain the parsimony based approach of developing phylogenetic trees.

(d) Construct a phylogenetic tree to explain the genetic relationship among taxa A, B, C, D, E and F, using ***UPGMA*** method.

	A	B	C	D	E	F
A	0	3	8	7	10	9
B		0	9	8	11	10
C			0	9	12	11
D				0	9	8
E					0	5
F						0

(e) Compute the quality of the above phylogenetic tree.

Question 4. **(20 marks)**

(a) What is the meaning of the term "**mutation**" in genome study?

(b) Explain how mutations cause genetic diseases/disorders.

(c) Name three databases/tools available at the **National Center for Biotechnology Information (NCBI)** website for analyzing genome data and disseminating biomedical information.

(d) Select two databases/tools you mentioned in part (c) and explain,

What type of information you can retrieve from those databases.
Or
What type of data you can analyze using those tools.

(e) Briefly explain the importance of the following primer design programs.
  I. **Oligo**
  II. **Primer3**

Question 5. **(20 marks)**

Write short notes about the following topics. Use diagrams where appropriate.

(a) Genetic Engineering Technology
(b) DNA Chip Technology
(c) Cell Culture Technology
(d) Medical Biotechnology
(e) Environmental Biotechnology

23
II