

# BIOSC 1540 - Computational Biology

## Quiz 02

Feb 18, 2025

20 points

Please read the following instructions carefully before beginning your assessment.

- **Time limit:** You have 15 minutes to complete and turn in this assessment.
- **Closed note:** You may not use any notes or additional resources during this assessment.
- **No digital devices:** The use of digital devices, including calculators, is not allowed.

I agree to follow the above instructions. I affirm that all work on this assessment will be my own and that I will not give or receive any unauthorized assistance. To have your assessment graded, you must write your name, sign, and provide your student ID below.

---

Name

---

Signature

---

Student ID

### Problem 1

A researcher notices that a homology-based gene prediction pipeline misses many short open reading frames (ORFs) in bacterial genomes. What is the most likely reason for these false negatives?  
(2 points)

- Ⓐ Homology-based methods favor longer genes with strong sequence conservation.
- Ⓑ Short ORFs do not contain codon usage bias, making them undetectable.
- Ⓒ The genome assembly process removes all short genes.
- Ⓓ Short ORFs are not functionally relevant in bacteria.

### Problem 2

Why do standard Hidden Markov Models struggle with predicting genes in eukaryotic genomes?  
(2 points)

- Ⓐ They assume nucleotide transitions are independent.
- Ⓑ They rely on a constant, sliding  $k$ -mer windows.
- Ⓒ They cannot detect regulatory elements.
- Ⓓ They are exponentially costly for long sequences.

### Problem 3

The process of gene annotation includes two major components: \_\_\_\_\_ annotation, which identifies the locations of genes, and \_\_\_\_\_ annotation, which assigns gene ontology terms to the predicted genes.  
(3 points)

### Problem 4

In an optimal traceback path of a global alignment, which of the following conditions must hold?  
(2 points)

- Ⓐ The traceback path must pass through every cell in the matrix.
- Ⓑ The path must move in a perfectly diagonal direction unless gaps exist.
- Ⓒ The path must terminate in the top-left corner of the matrix.
- Ⓓ The path will be unique for any two sequences.

### Problem 5

Under which condition will the Smith-Waterman and Needleman-Wunsch algorithms produce the same final alignment for two sequences and is the least strict?

(1 point)

- ☐ (A) The two sequences are the same length and require no gaps to align perfectly.
- ☐ (B) The scoring matrix is symmetric and contains only positive values.
- ☐ (C) The sequences share at least one identical stretch of characters.
- ☐ (D) The best local alignment spans the entire length of both sequences.

### Problem 6

What color was Alex's shirt? \_\_\_\_\_

(1 bonus point)

### Problem 7

Given this scoring matrix from the \_\_\_\_\_ algorithm, show the traceback and the resulting alignment.

(2 points)

|   |    | A  | A  | T  | C  | G   |
|---|----|----|----|----|----|-----|
|   | 0  | -2 | -4 | -6 | -8 | -10 |
| A | -2 | 1  | -1 | -3 | -5 | -7  |
| A | -4 | -1 | 2  | 0  | -2 | -4  |
| C | -6 | -3 | 0  | 1  | 1  | -1  |
| G | -8 | -5 | -2 | -1 | 0  | 2   |

**Alignment:**

(2 points)

### Problem 8

Using the the scoring scheme of match = 1, mismatch = -1, and gap = -2, what is the value of the empty cell below?

(2 points)

|   | C | T |
|---|---|---|
| C | 0 | 1 |
| A | 2 |   |

### Problem 9

What is a key advantage of RNA-seq over microarrays?

(3 points)

- ☐ (A) RNA-seq can detect novel transcripts and isoforms.
- ☐ (B) RNA-seq is cheaper than microarrays.
- ☐ (C) Microarrays provide absolute quantification of gene expression.
- ☐ (D) RNA-seq cannot be used to analyze non-coding RNA.

### Problem 10

You are analyzing transcriptomic data from a bacterial population exposed to an antibiotic. A gene encoding a known antibiotic resistance protein shows a 100-fold increase in transcript abundance after antibiotic treatment. However, genome sequencing reveals that no new mutations have occurred in the gene's regulatory region. What is the most plausible explanation?

(1 point)

- ☐ (A) The sequencing depth in the genome study was insufficient to detect mutations.
- ☐ (B) The gene is under transcriptional regulation by an antibiotic-responsive activator.
- ☐ (C) The gene was horizontally transferred from another bacterial species.
- ☐ (D) The antibiotic directly binds to the RNA polymerase, enhancing transcription.

**Doodle area or word search:** Find as many words as you can hidden in the word "CONVERSATION" (e.g., "TONE" and "VISION").