ourdiu.com | Shared by Humaira Afrin Prity

**Quiz-2, CSE115-IBCC, Sum'26, Time: 45minutes, Marks: (8+7), ID:**

You are a comparative genomics researcher studying gene duplication events. You are comparing a functional metabolic gene in *E. coli* to a suspected non-functional pseudogene in a closely related symbiotic bacterium. Because pseudogenes accumulate mutations rapidly but rarely tolerate massive structural gaps without breaking completely, you set a strict penalty for gaps to focus purely on sequence drift.

| **Given the sequences:** | **Scoring scheme:** |
|---|---|
| S1 = CGCTAC | Match = +1 |
| S2 = CTGTAG | Mismatch = -1 |
| | Gap = -2 |

Which of these scenarios would you like to solve step-by-step using a dynamic programming matrix?

A clinical research team at a city hospital is monitoring a cohort of patients undergoing long-term antiviral therapy for Hepatitis B. To track the emergence of drug-resistance mutations in the viral reverse transcriptase domain, they sequence a highly conserved 6-base motif from six patient isolates (A–F).

| | | | | | | |
|---|---|---|---|---|---|---|
| A: | G | T | C | A | A | G |
| B: | G | T | G | A | A | T |
| C: | G | A | C | A | G | G |
| D: | C | A | C | A | G | G |
| E: | C | A | G | T | G | C |
| F: | G | A | C | G | T | C |

Now, Generate Distance matrix using P-Distance method. Then, apply UPGMA method to generate first updated cluster table.