**MIP 280A4: Computational Microbiology**

**Pairwise Sequence Alignments, Day #2, In-class exercise questions**

**N-W exercise**

1. Use the Needleman-Wunsch algorithm and the provided worksheet to determine the highest scoring global alignment of the two sequences listed below given the scoring system listed below. (5 pts).

Seq 1: ACTG

Seq 2: ACG

**Scoring system:**

|  |  |
| --- | --- |
| Match | +1 |
| Mismatch | -1 |
| Gap | -1 |

Highest scoring alignment:

Score of highest scoring alignment:

Please upload your worksheet (matrix) also.

1. On a scale of 1-10, how much would you enjoy filling out a N-W matrix for 2 sequences of length 100 bases? (1 = no enjoyment; 10 = maximum enjoyment) (1 pt)

**Protein scoring matrices**

1. How could you use the fact that different amino acids have different chemical properties in a protein alignment scoring system? (1 pt)

**BLOSUM exercise**

1. Use the BLOSUM62 matrix to score the following gapless alignment, which is a partial alignment of the West Nile virus and Zika virus envelope protein sequences (2 pts)

C R V K M E K L Q L K

| | | | | | | |

C R L K M D K L R L K

1. Which match (or mismatch) in the above alignment contributed most positively to the alignment’s score? What is the biological basis for this? (2 pts)

**Protein sequence alignment exercise**



Electron micrograph of *H. pylori* possessing multiple flagella, image: Yutaka Tsutsumi

1. Download into Geneious the flagellin protein sequences for *E. coli* (NP\_310689) and for *Shigella dysenteriae* (WP\_119176587). Perform a global alignment (Needleman-Wunsch) using default settings.
   1. How long is the alignment? (1 pt)
   2. What scoring matrix did this alignment use? (1 pt)
   3. What is the pairwise percent identity of this alignment? (1 pt)
   4. What is the pairwise similarity of this alignment? (1 pt)
   5. Is identity even across this alignment? Where is identity higher or lower? (1 pt)
2. Perform a local alignment (Smith-Waterman) of these two sequences using default settings in Geneious.
   1. How long is the alignment? (1 pt)
   2. What scoring matrix did this alignment use? (1 pt)
   3. What is the pairwise percent identity of this alignment? (1 pt)
   4. What is the pairwise similarity of this alignment? (1 pt)