

CSC 857 Bioinformatics Computing, Fall 2013

HOME ASSIGNMENT 2 (Total 10 Points)

Assignment Date: Sept 10, Due Date: Oct 8, In Class

Question 1 (1 Point)

How frequently would you expect to find the sequence of nucleotides provided below in a DNA molecule simply as a result of random chance? Assume that each of the four nucleotides occurs with the same frequency.

5'-GGATCGTAGCCTA-3'

Question 2 (2.5 Points)

Determine the alignment score for the following sequence alignments:

a. Global alignment: match score = +1, mismatch score = 0, gap penalty = -1

T	G	T	A	C	G	G	C	T	A	T	A
T	C	--	--	C	G	C	C	T	--	T	A

b. Global alignment: match score = +1, mismatch score = 0, gap penalty = -1

--	--	T	C	T	G	T	A	C	G	C	G	A	T	C	A	T	G	T
T	A	G	C	--	G	T	C	C	G	A	T	A	T	--	A	--	--	--

c. Global alignment: match score = +1, mismatch score = -1, gap origination penalty = -2, gap length penalty = -1

A	G	A	T	A	G	A	A	A	C	T	G	A	T	A	T	A	T	A
A	G	A	--	A	--	A	--	A	C	A	G	A	G	--	T	--	--	--

d. Global alignment: match score = +1, mismatch score = -1, gap origination penalty = -2, gap length penalty = -1

A	G	A	T	A	G	A	A	A	C	T	G	A	T	A	T	A	T	A
A	G	--	--	--	A	A	A	A	C	A	G	A	G	T	--	--	--	--

e. Semiglobal alignment: match score = +1, mismatch score = -1, gap origination penalty = -2, gap length penalty = -1

A	G	A	T	A	G	A	A	A	C	T	G	A	T	A	T	A	T	A
A	G	--	--	--	A	A	A	A	C	A	G	A	G	T	--	--	--	--

Question 3 (2 Points)

Using the Needleman and Wunsch dynamic programming method, construct the alignment score table for the following two sequences, using the following scoring parameters. Mark all arrows clearly:

match score = +3, mismatch score = -1, gap penalty = -2

5'-ACAGTCGAACG-3'

5'-ACCGTCCG-3'

Question 4 (2 Points)

Using the same scoring parameters as in Question 3, use the modified Needleman and Wunsch method to compute the optimal semi-global alignment for the two sequences in Question 3. Mark all arrows clearly

Question 5 (2.5 Point)

Using the Smith-Waterman method, construct the partial alignment scoring table for a local alignment of the following two sequences using the parameters below:

match score = +3, mismatch score = -1, gap penalty = -2

5-'ACGTATCGCGTATA-3'

5-'GATGCTCTCGGAAA-3'

Mark all arrows clearly in the table.