

Assignment – III

(Sequence Alignment)

Deadline: 8th April

1. You are given 2 nucleotide sequences:

GGCTGCAACTAGCTC
GGGTAAGCTTGC

and the transition-transversion scoring matrix (expressed in similarity):

	A	C	G	T
A	4	-1	1	-1
C	-1	4	-1	1
G	1	-1	4	-1
T	-1	1	-1	4

and gap penalty -3.

Carry out the global and local alignment (dynamic programming algorithm), and indicate the final similarity score and the best alignment.

2. Identify the dinucleotide CA repeat region and the score in the following sequence (using the scoring scheme (1, 0, -1) for match, mismatch, indel, respectively):

TGGCACACTCACACCACACAGACAGTTA

3. When would you encounter a situation for using DP for overlap regions? Give the boundary conditions and recursive relations in this case?
4. What is the advantage of using affine gap scores?
5. Give the time and space complexity of dynamic programming algorithm for sequence comparison. Under what conditions does time and space can become an issue?
6. How is PSI-BLAST different from blastp program?
7. Find self-complementary regions in the following RNA sequence using dotplots approach: AUGUGGCAUGCCAGG