

**CSC / BIO 310**  
**Bioinformatics**  
**Instructor: Dr. Laurie J. Heyer**  
**Assignment #7**  
**Due Thursday, Feb 7**

1. (a) Perform a local alignment **by hand** of the following two sequences:

ATGGCCTC

ACGGCTCA

using the same 1, -3, -4 scoring vector we have used in class. Show your work. You may check your answer by modifying the Excel implementation of global alignment to perform local alignment instead.

- (b) What **single change** could you make to the scoring vector (i.e. replace one of the three numbers with a new number) that would make the optimal local alignment significantly longer? Specify what this new local alignment would be, and what score it would have. (Note that there is more than one right answer to this question, and that you can answer it without redoing the entire alignment.)

2. Show that the product of two Jukes-Cantor matrices is again a Jukes-Cantor matrix. In other words, if  $M(\alpha_1)$  is the Jukes-Cantor matrix with parameter  $\alpha_1$  and  $M(\alpha_2)$  is the Jukes-Cantor matrix with parameter  $\alpha_2$ , show that  $M(\alpha_1) M(\alpha_2)$  has the form  $M(\alpha_3)$ , and give a formula for  $\alpha_3$  in terms of  $\alpha_1$  and  $\alpha_2$ .

3. The phylogenetic distances between pairs of sequences S1, S2, S3 and S4 are given in the table below. Use the UPGMA algorithm to construct a rooted tree showing the relationships among these sequences. Show the details of each step of the algorithm.

	S1	S2	S3	S4
S1		1.2	.9	1.7
S2			1.1	1.9
S3				1.6