

Biol 47800/59500 Homework 3

Work individually on this assignment. Feel free to write a program to perform the alignments if you like (it took me about 1.5 hr to do so).

ACGCTAGCTG

TCCCCTGATG

1. Using the sequences above, show the global dynamic programming matrix (scores and pointers) and alignment using an identity scoring system

match = +1

mismatch = -0.5

gap open = -1 *nb*, length independent penalty

gap extend = 0

Hint: you can avoid fractions by multiplying by 10.

2. Show the dynamic programming score and traceback matrix for the local alignment of sequence 1, "ATG", and sequence 2, "AATCGTG", using an identity scoring system. This can be a single matrix with numbers for the scores and arrows for the traceback, as shown in class. In order to avoid the decimals, you may want to multiply all scores and penalties by 10.

match=+1

mismatch=-0.8

gap open=-1

gap extend = 0).

- a) Give the positions of all the maximum score(s) (e.g., position 1,1 would be the first letter in sequence 1 aligned with the first letter of sequence 2).
- b) For each maximum, show all the local alignments that can be traced using the traceback strategy discussed in class.