1   Substitution Matrix

Given that, A ←→ G and T ←→ C) are less common than tranversions (A ←→ T , A ←→ C, G ←→ T , and G ←→ C)
The substitution matrix for above changes is
[[0 A G T C]
[A 1 -2 -1 -1]
[G -2 1 -1 -1]
[T -1 -1 1 -2]
[C -1 -1 -2 1]]

2   Global Alignment

The Needleman-Wunsch algorithm was used to create a 2D matrix D using the global alignment function. Sequence A, Sequence B, the substitution matrix, and the gap penalty
Example: Sequence A="GATA" Sequence="CTAC" with a match score of 1, mismatch score of -1, and gap score of -2.
The D matrix is generated as below:
[[0 -2 -4 -6 -8]
[-2 -1 -3 -5 -5]
[-4 -3 -2 -2 -4]
[-6 -5 -2 -3 -3]
[-8 -7 -4 -1 -3]]
As we traverse the input sequence "GATA" and "CTAC" we get the global alignments [("GATA-","C-TAC"),("GATA-","-CTAC")]

Example: Sequence A="ACTAG" Sequence B="AGCT" with a match score of 1, mismatch score of -1, and gap score of -2.
The D matrix is generated as below:
[[0 -2 -4 -6 -8]
[-2 1 -1 -3 -5]
[-4 -1 0 0 -2]
[-6 -3 -2 -1 1]
[-8 -5 -4 -3 -1]
[-10 -7 -4 -5 -3]]
we get the global alignments such as [("A-CTAG", "AGCT--")]

3   Local Alignment

The Smith-Waterman algorithm was used to create a 2D matrix D with the help of a function called local alignment, which takes variables. Sequence A, Sequence B, the substitution matrix, and the gap penalty

Example: Sequence A="ACTG" Sequence B="GTCA" with a match score

of 1, mismatch score of -1, and gap score of -2.
The D matrix is generated as below:
[0 0 0 0 0]
[0 0 0 0 1]
[0 0 0 1 0]
[0 0 1 0 0]
[0 1 0 0 0]
Traversing for the input sequence "ACTG" and "GTCA" we get the global alignments such as [("A", "A")]

Example: Sequence A="GTACT" Sequence B="CTAGA" with a match score of 3, mismatch score of -1, and gap score of -2.
The D matrix is generated as below:
[0 0 0 0 0 0]
[0 0 0 0 1 0]
[0 0 1 0 0 0]
[0 0 0 2 0 1]
[0 1 0 0 1 0]
[0 0 2 0 0 0]
As we traversing D,for the input sequence "GTACT" and "CTAGA" we get the global alignments such as [("TA", "TA")]

4   Custom Alignment

Taking first name and last name in lowercase and concatenating them, first-name="venkatankit" ,lastname="gundala"
when we concatenate string is "venkatankitgundala" ,
With the help of function named local alignment from above problem which takes variables,Sequence A,Sequence B,substitution matrix,gap penalty
match=2,semi-match=+1,mismatch=-1,gap=-2 Sequence A="venkatankitgundala"
Sequence B="hequickbrownfoxjumpsoverthelazydog"

The score matrix is included in file 1002069069_S.txt

The pretty print is included in 1002069069_D.txt

The local alignment for the above strings is [("ve-nkata", "verthela"), ("venkatan", "verthela"), ("ve-nkatankitg","verthelazydog"), ("tgundala", "verthela")]

5   Difficulty Adjustment

It took me around 20 hours to complete this assignment and I felt that the custom alignment part was a bit tough to complete