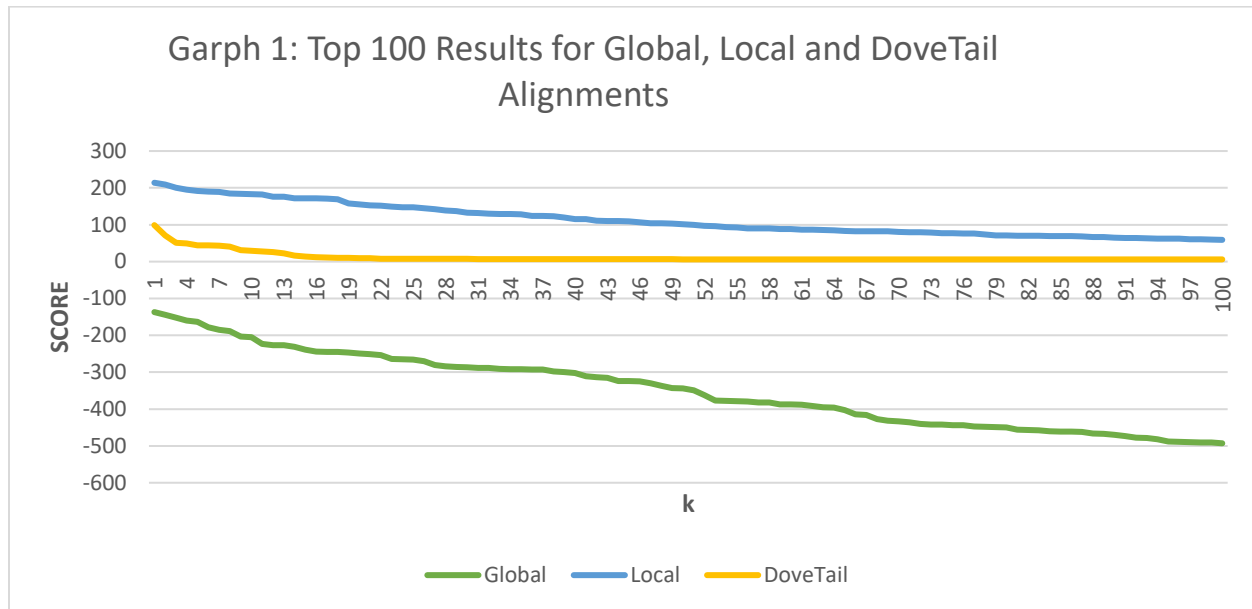


Name: Poornima Kumar

UFID: 5684-4925

Bioinformatics Assignment 1 Report

Graph 1:

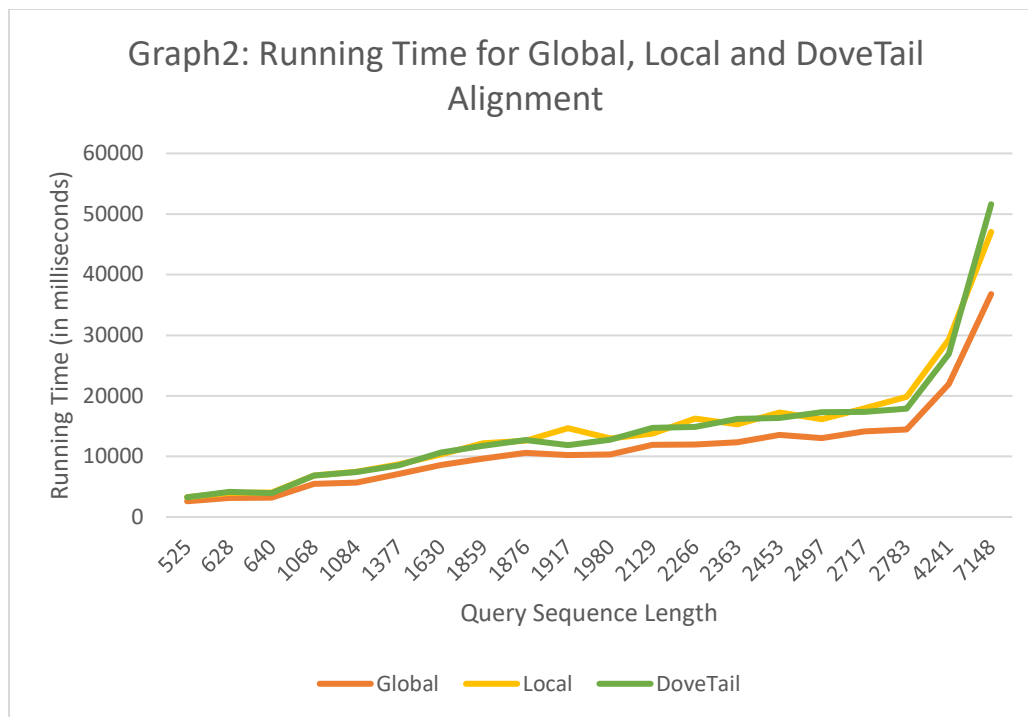


Score of the top 100 results for each alignment strategy.

Observations:

1. Global Alignment tries to align the entire sequence. When the query string is compared with the database string, there are many unmatched values. This results in more negative values in the output as the gap penalties are added with every unmatched pair.
2. Local Alignment tries to align subsequences. Unlike global, we compare the values in the dynamic programming matrix with 0. As a result, there can never be a negative score. The maximum value in the score matrix is taken as the highest score and is considered as a part of the local alignment. Hence, the substring will contain the maximum score.
3. Dovetail (or End space free) Alignment aligns opposite ends of the two sequences. The suffix of one string aligns with the prefix of the other or vice versa. For this approach, the maximum score is taken from the last column or the last row and then backtracking to the score 0 is done. The scores are between the global and the local alignment. Taking the maximum score from the last column/row ensures that the suffix of one string is considered and backtracking to 0 ensures the prefix of the other string is taken.

Graph2:



Alignment of each query in the query file to all the sequences in the sample database file.

Observations:

1. The general trend observed here is that with the increase in the length of the query string, the running time of all the three algorithms increases. This is because the database sequence has to be compared to large query sequences.
2. Global Alignment: This algorithm does not need to keep a track of the maximum score found in the dynamic programming table. This eliminates the comparison time taken to find maximum in the other algorithms. However, the algorithm takes considerable amount of time as it has to backtrack the entire matrix.
3. Local Alignment: This algorithm is comparatively slower than the global alignment algorithm. This is because it has to constantly maintain the maximum score found at any given point of time. The algorithm returns when it encounters the first zero while retracing the path.
4. Dovetail Alignment: This algorithm keeps track of the maximum element like local alignment. However, the overhead of the maximum is reduced as it searches for the maximum value only in the last row and the last column. Also, like global alignment, every entry in the matrix is compared to three values and the backtracking is done till the start of either of the string. Hence, the time for dovetail alignment is between global and local alignment.