

Zadanie 2

Wojciech Miśta, 236453

Algorytm znajduje się w osobnym pliku na repozytorium.

Analiza złożoności obliczeniowej:

scoreMatrix.m

```
function [scoredMatrix, indexMatrix] = scoreMatrix(scoredMatrix, match, mismatch, gap, m, n)
    indexMatrix = zeros(size(scoredMatrix));
    for k = 1:m%wiersze
        for p = 1:n%kolumny
            if (scoredMatrix(k+1,p+1) == 1)
                val1 = scoredMatrix(k,p) + match;
            else
                val1 = scoredMatrix(k,p) + mismatch;
            end
            valGap1 = scoredMatrix(k,p+1) + gap; %one up
            valGap2 = scoredMatrix(k+1,p) + gap; %one left
            [maxValue, index] = max([val1 valGap1 valGap2]);
            scoredMatrix(k+1,p+1) = maxValue;
            %if 1 - match/mismatch
            %if 2 - gap up
            %if 3 - gap left
            indexMatrix(k+1,p+1) = index;
        end
    end
end
```

m,n – długość sekwencji wejściowej

Złożoność obliczeniowa: $O(mn)$

Złożoność przestrzenna: $O(mn)$

createMatrixPath.m

```
function [matrixPath, length, matchCount, gapCount, seq1Array, seq2Array] = createMatrixPath
%this function creates the path and creates a sequence output

length = 0;
matchCount = 0;
gapCount = 0;
matrixPath = zeros(size(indexMatrix));
i = 1;

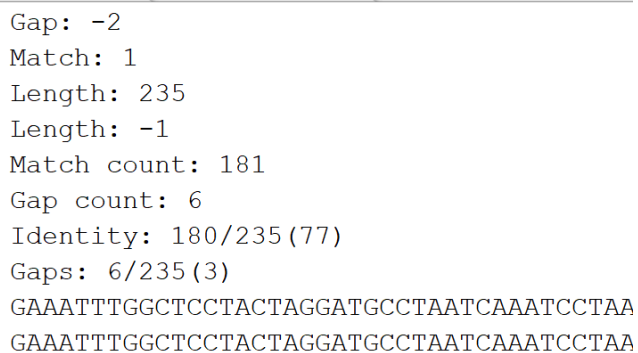
while (curColumn > 1) || (curRow > 1)
    matrixPath(curRow,curColumn) = 1;
    if(indexMatrix(curRow, curColumn) == 1)
        if(matrixCompared(curRow-1,curColumn-1) == 1) %bo nie wiem czy to match czy mism:
            matchCount = matchCount + 1;
            seq1Array(i)=seq1(curRow-1);
            seq2Array(i)=seq2(curColumn-1);
        end
        curColumn = curColumn - 1;
        curRow = curRow - 1;
    elseif(indexMatrix(curRow,curColumn) == 2)
        seq1Array(i)=seq1(curRow-1);
        seq2Array(i)='_';
        curRow = curRow - 1;
        gapCount = gapCount + 1;
    else
        seq1Array(i)='_';
        seq2Array(i)=seq2(curColumn-1);
        curColumn = curColumn - 1;
        gapCount = gapCount + 1;
    end
    length = length +1;
    i = i + 1;
end
end
```

p – rozmiar zmiennej 'matrixPath'

Złożoność czasowa: $O(p)$

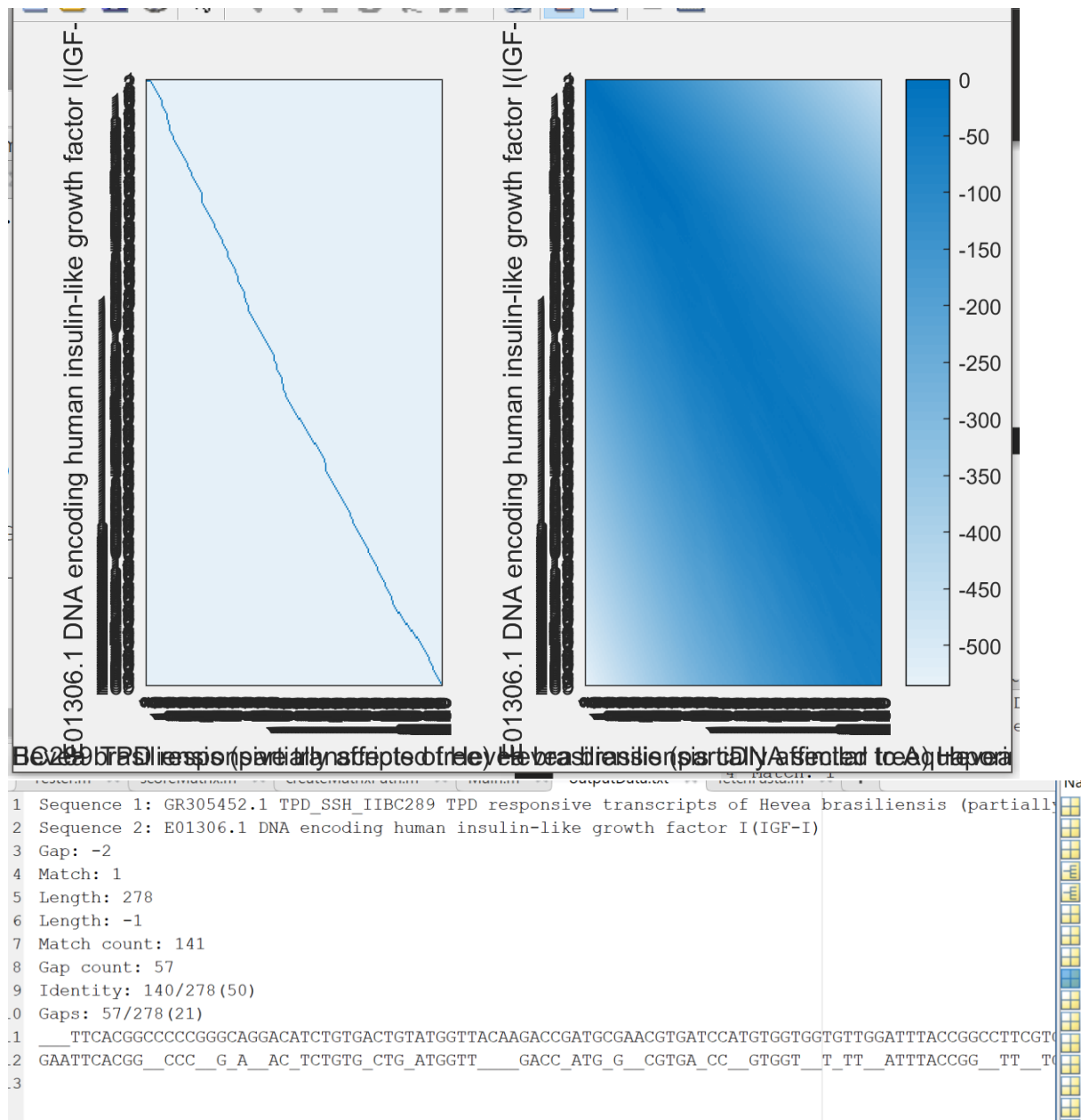
Złożoność przestrzenna: $O(p^2)$

Powiązanych:



Porównanie mamuta oraz nosorożca.

Niepowiązanych:



Wnioski:

- Zauważamy znaczą różnicę w drodze ścieżki przy porównaniu sekwencji powiązanych oraz niepowiązanych. Sekwencje powiązane tworzą znacznie prostszą ścieżkę, w której znajduje się znikoma ilość insercji bądź delecji (w przykładzie – 6, 'identity' równe 77%). Sekwencje niepowiązane mają znacznie niższą wartość 'identity' oraz znacznie zwiększona liczbę gap'ów. Na przykładzie, 57 gap'ów przy 'identity' równym 50%.