

Wstęp do bioinformatyki  
Laboratorium 2  
Dopasowanie globalne par sekwencji  
Magdalena Trędak  
236712

1. Schemat blokowy algorytmu dopasowania lokalnego

Schematy blokowe algorytmów tworzenia macierzy punktowej oraz optymalnej ścieżki dopasowania ze względu na duży rozmiar i umożliwienia poprawy ich czytelności zamieszczono w repozytorium jako pliki graficzne o nazwach: SchematBlokowyGenerowaniaMacierzyPunktów.jpg, SchematBlokowyGenerowaniaŚcieżkiDopasowania.jpg.

Do wygenerowania schematów użyto programu online znajdującego się na stronie:

<http://www.algorytm.org/narzedzia/edytor-schematow-blokowych.html> (data dostępu 24.04.19)

2. Analiza złożoności obliczeniowej czasowej i pamięciowej

Analiza złożoności obliczeniowej czasowej i pamięciowej

- Oszacowanie złożoności czasowej dla poszczególnych funkcji f

scoringMatrix =  $m \cdot n$   
tracBackMatrix =  $m \cdot n$   
createInfo =  $m \cdot n + m$   
isSequenceCorrect =  $m + n$

$$\Sigma f = 3 \cdot m \cdot n + 2 \cdot m + n \leq 3 \cdot m \cdot n$$

$O(mn)$  – złożoność czasowa co najwyżej rzędu  $mn$

- Oszacowanie złożoności pamięciowej dla poszczególnych funkcji f

scoringMatrix =  $2 \cdot m \cdot n + m + n$   
tracBackMatrix =  $4 + (m+n) \cdot 6$   
createInfo =  $m \cdot n + m + n + m + n + m \cdot n + 5 \cdot (m+n) + m$   
makeFastaContent =  $5 \cdot (m+n)$   
readFasta =  $m + n$

$$\Sigma f = 19 \cdot m \cdot n + 15 \cdot (m+n) + m + 4 \leq 6 \cdot m \cdot n$$

$O(mn)$  – złożoność pamięciowa co najwyżej rzędu  $mn$

3. Porównanie przykładowych par sekwencji mitochondrialnego cytochromu b dla parametrów :  
match = -1, mismatch = 1, gap = 2, mode = distance

a) niepowiązanych ewolucyjnie

AJ009879.1 - Nubian ibex

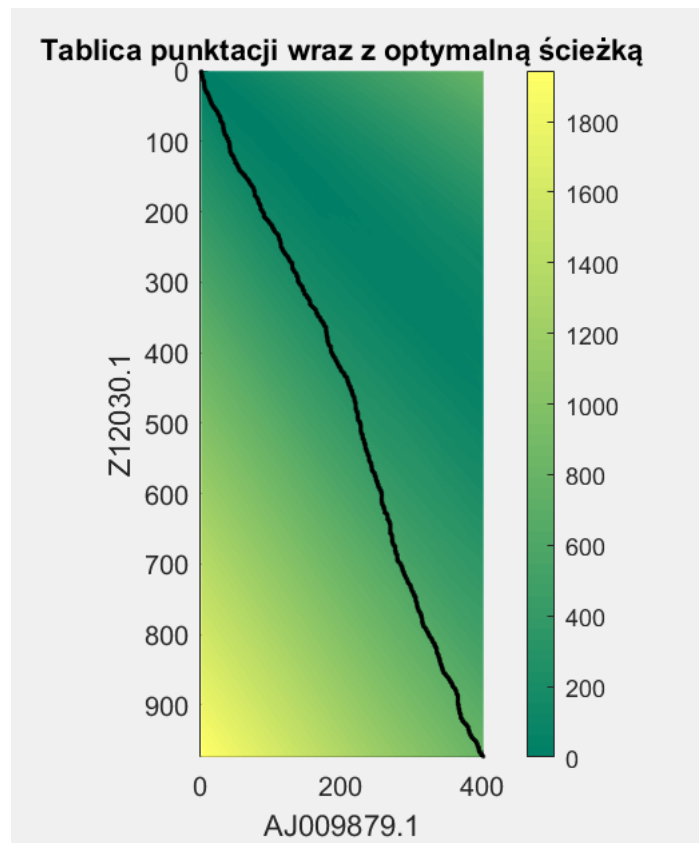
Z12030.1 - Tropheus annectens

```
Command Window
>> test1
#1: AJ009879.1
#2: Z12030.1
Mode: distance
#Match: -1
#Mismatch: 1
#Gap: 2
#Length: 972
#Score: 798
#Gaps: 572/972 (58.8477 %)
#Identity: 373/972 (38.3745 %)
>seq1 CAGATCCTAACAGGCCTATTCCTGGCAATACACTATACATCTGACACAACAACAGCATTC
>seq2 ---A---T---GG-----C---C-A-AC-C--T-C--C-G--A-AA-AAC-CCA--C
          *   *   **           *   * * * * *   * *   * *   * * * * *   * *   *
          *   *   **           *   * * * * *   * *   * *   * * * * *   * *   *

>seq1 TCCTCTGTAACTCACATTTGTCGAGATGTAAATTATGGCTGAATTATCCGATATATACAC
>seq2 -CC-C-----CTC-C-----T----A---AAA--AT---TG-----C-----A-A-AC
          ** *           *** *           *   *   ***   **   **           *           * * * *

>seq1 GCAAACGGAGCATCAATATTCTTTATCTGCCTATTCATACATGTCGGACGAGGCCTATAT
>seq2 G---AC---GCA-C--TA--GTTGATCT-CCCA-GC-T-CCT-TC--A--A-----ACA-
          *   **   *** *   **   ** * * * * * * *   * * * * *   *   *           * *
```

fx



b) powiązanych ewolucyjnie

X75584.1 - humpback whale

X75583.1 - pygmy Bryde's whale

# Command Window

```
>> Tester
#1: X75583.1
#2: X75584.1
Mode: distance
#Match: -1
#Mismatch: 1
#Gap: 2
#Length: 1140
#Score: -966
#Gaps: 0/1140 (0 %)
#Identity: 1053/1140 (92.3684 %)
>seq1 ATGACCAACATCCGAAAAACACACCCACTAATAAAGATTGTCAACGATGCATTTCGTTGAT
>seq2 ATGACCAACATCCGAAAAACACACCCACTAATAAAAATTATCAACGACACATTTCATTGAT
*****

>seq1 CTCCCCACCCCATCAAATATCTCCTCATGATGAAATTTTCGGCTCCCTACTCGGCCTCTGC
>seq2 CTACCCACCCCATCAAATATCTCCTCATGATGAAACTTTCGGTTCCTACTCGGCCTTTGC
** *****

>seq1 TTAATTACACAAATCCTAACAGGCCTATTCTAGCAATACACTACACACCAGACACAACA
>seq2 TTAATTATACAAATCCTAACAGGCCTATTCTAGCAATACACTACACACCAGACACAACA
*****
```

fx

