

TASK 1

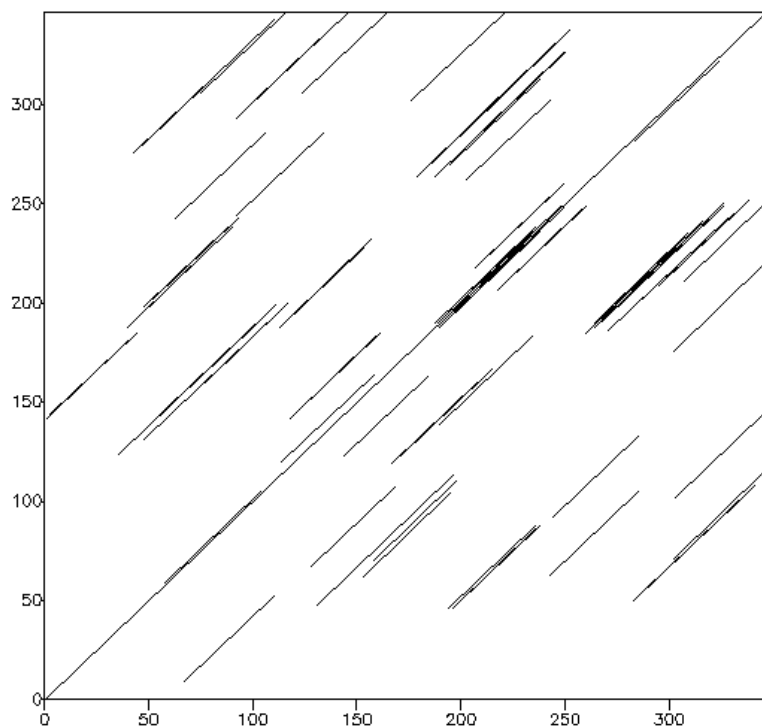
NC_012920.1

TASK 2

NC_001643

TASK 3

Dotmatcher: raw::/var/lib/emboss-explorer/output/023117/...
(windowsize = 40, threshold = 15.00 22/05/23)



TASK 4

Przez wykres przeciąga się jedna ukośna linia od 0,0 do 400,400. Im większa wartość "Window size over which to test threshold" tym mniej szczegółów/kropek jest widocznych, podobnie z "threshold".

TASK 5

TASK 6

Jako, że wartości zanikają i jest to porównanie dwóch sekwencji, opisałbym threshold (pol. próg), jako próg podobieństwa, przy którym jest "stawiana" kropka.

"Window size over which to test threshold" spróbowałbym określić jako *przybliżenie* brane pod uwagę podczas rysowania wykresu.

TASK 7

rs333

TASK 8

```
>3 dna:chromosome chromosome:GRCh38:3:46373053:46373887:1
GTCATCCTCATCCTGATAAACTGCAAAAGGCTGAAGAGCATGACTGACATCTACCTGCTC
AACCTGGCCATCTCTGACCTGTTTTTCCTTCTTACTGTCCCCTTCTGGGCTCACTATGCT
GCCGCCAGTGGGACTTTGGAAATACAATGTGTCAACTCTTGACAGGGCTCTATTTTATA
GGCTTCTTCTCTGGAATCTTCTTCATCATCCTCCTGACAATCGATAGGTACCTGGCTGTC
GTCCATGCTGTGTTTGCTTTAAAGCCAGGACGGTCACCTTTGGGGTGGTGACAAGTGTG
ATCACTTGGGTGGTGGCTGTGTTTGCCTCTCTCCAGGAATCATCTTTACCAGATCTCAA
AAAGAAGGTCTTCATTACACCTGCAGCTCTCATTTTTCCATACAGTCAGTATCAATTCTGG
AAGAATTTCCAGACATTAAAGATAGTCATCTTGGGGCTGGTCTGCCGCTGCTTGTCATG
GTCATCTGCTACTCGGGAATCTTAAAACTCTGCTTCGGTGTGCAAAATGAGAAGAAGAGG
CACAGGGCTGTGAGGCTTATCTTCACCATCATGATTGTTTATTTTCTTCTGCGCTCCC
TACAACATTGTCCTTCTCCTGAACACCTTCCAGGAATCTTTGGCCTGAATAATTGCAGT
AGCTCTAACAGGTTGGACCAAGCTATGCAGGTGACAGAGACTCTTGGGATGACGCACTGC
TGCATCAACCCCATCATCTATGCCTTTGTGCGGGAGAAGTTCAGAACTACCTCTTAGTC
TTCTTCCAAAAGCACATTGCCAAACGCTTCTGCAAATGCTGTTCTATTTTCCAGC
```

TASK 9

KU382465.1

TASK 10

```
>KU382465.1 Chlorocebus pygerythrus CCR5 gene, complete cds
ATGGATTATCAAGTGTCAAGTCCAACCTATGACATCAATTATTATACATCGGAGCCCTGCCAAAAATCA
ACGTGAAGCAAATTGCAGCCCGCTCCTGCCTCCGCTCTACTACTGGTGTTCATCTTTGGTTTTGTGGG
CAACATACTGGTCGTCCTCATCCTGATAAACTGCAAAAGGCTGAAAAGCATGACTGACATCTACCTGCTC
AACCTGGCCATCTCTGACCTGTTTTCTTCTTACTGTCCCCTTCTGGGCTCACTATGCTGCTGCCAGT
GGGACTTTGGAAATACAATGTGTCAACTCTTGACAGGGCTCTATTTTATAGGCTTCTTCTCTGGAATCTT
CTTCATCATCCTCCTGACAATCGATAGGTACCTGGCTATCGTCCATGCTGTGTTTGCTTTAAAGCCAGG
ACAGTCACCTTTGGGGTGGTGACAAGTGTGATCACTTGGGTGGTGGCTGTGTTTGCCTCTCTCCCAAGAA
TCATCTTTACCAGATCTCAGAGAGAAGGTCTTCATTACACCTGCAGCTCTCATTTTCCATACAGTCAGTA
TCAATTCTGGAAGAATTTCCAGACATTAAAGATAGTCATCTTGGGGCTGGTCTGCCGCTGCTTGTCATG
GTCATCTGCTACTCGGGAATCTGAAAACTCTGCTTCGGTGTGCAAAACGAGAAGAAGAGGCACAGGGCTG
TGAGGCTCATCTTCACCATCATGATTGTTTATTTTCTTCTTCTGGGCTCCCTACAACATTGTCCTTCTCCT
GAACACCTTCCAGGAATCTTTGGCCTGAATAATTGCAGTAGCTCTAACAGGTTGGACCAAGCCATGCAG
GTGACAGAGACTCTTGGGATGACACACTGCTGCATCAACCCCATCATCTATGCCTTCGTCGGGGAGAAGT
TCAGAACTACCTCTTAGTCTTCTTCCAAAAGCACATTGCCAAACGCTTCTGCAAATGCTGTTCCATTTT
CCAGCAAGAGGCTCCCGAGCGAGCAAGTTCAGTTTACACCCGATCCACTGGGGAGCAGGAAACATCTGTG
GGCTTGTGA
```

TASK 11

```
#####
# Program: needle
# Rundate: Sun 21 May 2023 23:58:23
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20230521-235805-0033-24972353-p1m.asequence
```

```

# -bsequence emboss_needle-I20230521-235805-0033-24972353-p1m.bsequence
# -datafile EDNAFULL
# -gapopen 10.0
# -gapextend 0.5
# -endopen 10.0
# -endextend 0.5
# -aformat3 pair
# -snucleotide1
# -snucleotide2
# Align_format: pair
# Report_file: stdout
#####

#=====
#
# Aligned_sequences: 2
# 1: EMBOSS_001
# 2: EMBOSS_001
# Matrix: EDNAFULL
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 1059
# Identity:      818/1059 (77.2%)
# Similarity:    818/1059 (77.2%)
# Gaps:          224/1059 (21.2%)
# Score: 4022.0
#
#
#=====

EMBOSS_001      1  ----- 0
EMBOSS_001      1  ATGGATTATCAAGTGTCAAGTCCAACCTATGACATCAATTATTATACATC 50
EMBOSS_001      1  ----- 0
EMBOSS_001     51  GGAGCCCTGCCAAAAATCAACGTGAAGCAAATTGCAGCCCGCCTCCTGC 100
EMBOSS_001      1  ----- 0
EMBOSS_001    101  CTCCGCTCTACTCACTGGTGTTCATCTTTGGTTTTGTGGGCAACATACTG 150
EMBOSS_001      1  GTCATCCTCATCCTGATAAACTGCAAAAGGCTGAAGAGCATGACTGACAT 50
      |||.|||||||||||||||||||||||||||.|||||||||||
EMBOSS_001    151  GTCGTCCTCATCCTGATAAACTGCAAAAGGCTGAAAAGCATGACTGACAT 200
EMBOSS_001     51  CTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTCTTCTTACTGTCC 100
      |||||||||||||||||||||||||.|||||||||||
EMBOSS_001    201  CTACCTGCTCAACCTGGCCATCTCTGACCTGCTTTCTTCTTACTGTCC 250
EMBOSS_001    101  CCTTCTGGGCTCACTATGCTGCCGCCAGTGGGACTTTGGAAATACAATG 150
      |||||||||||||||||.|||||||||||||
EMBOSS_001    251  CCTTCTGGGCTCACTATGCTGCTGCCAGTGGGACTTTGGAAATACAATG 300
EMBOSS_001    151  TGTCAACTCTTGACAGGGCTCTATTTATAGGCTTCTTCTCTGGAATCTT 200

```

EMBOSS_001	301	 TGTCAACTCTTGACAGGGCTCTATTTTATAGGCTTCTTCTCTGGAATCTT	350
EMBOSS_001	201	CTTCATCATCCTCCTGACAATCGATAGGTACCTGGCTGTCGTCCATGCTG	250
EMBOSS_001	351	 CTTCATCATCCTCCTGACAATCGATAGGTACCTGGCTATCGTCCATGCTG	400
EMBOSS_001	251	TGTTTGCTTTAAAAGCCAGGACGGTCACCTTTGGGGTGGTGACAAGTGTG	300
EMBOSS_001	401	 TGTTTGCTTTAAAAGCCAGGACAGTCACCTTTGGGGTGGTGACAAGTGTG	450
EMBOSS_001	301	ATCACTTGGGTGGTGGCTGTGTTTGCCTCTCTCCAGGAATCATCTTTAC	350
EMBOSS_001	451	 ATCACTTGGGTGGTGGCTGTGTTTGCCTCTCTCCAAGAATCATCTTTAC	500
EMBOSS_001	351	CAGATCTCAAAAAGAAGGTCTTCATTACACCTGCAGCTCTCATTTTCCAT	400
EMBOSS_001	501	 CAGATCTCAGAGAGAAGGTCTTCATTACACCTGCAGCTCTCATTTTCCAT	550
EMBOSS_001	401	ACAGTCAGTATCAATTCTGGAAGAATTTCCAGACATTAAAGATAGTCATC	450
EMBOSS_001	551	 ACAGTCAGTATCAATTCTGGAAGAATTTCCAGACATTAAAGATAGTCATC	600
EMBOSS_001	451	TTGGGGCTGGTCCTGCCGCTGCTTGTCATGGTCATCTGCTACTCGGGAAT	500
EMBOSS_001	601	 TTGGGGCTGGTCCTGCCGCTGCTTGTCATGGTCATCTGCTACTCGGGAAT	650
EMBOSS_001	501	CCTAAAACTCTGCTTCGGTGTGCGAAATGAGAAGAAGAGGCACAGGGCTG	550
EMBOSS_001	651	. CCTGAAAACTCTGCTTCGGTGTGCGAAACGAGAAGAAGAGGCACAGGGCTG	700
EMBOSS_001	551	TGAGGCTTATCTTCACCATCATGATTGTTTATTTTCTTCTCTGGGCTCCC	600
EMBOSS_001	701	. TGAGGCTCATCTTCACCATCATGATTGTTTATTTTCTTCTCTGGGCTCCC	750
EMBOSS_001	601	TACAACATTGTCCTTCTCCTGAACACCTTCCAGGAATTCTTTGGCCTGAA	650
EMBOSS_001	751	 TACAACATTGTCCTTCTCCTGAACACCTTCCAGGAATTCTTTGGCCTGAA	800
EMBOSS_001	651	TAATTGCAGTAGCTCTAACAGGTTGGACCAAGCTATGCAGGTGACAGAGA	700
EMBOSS_001	801	 TAATTGCAGTAGCTCTAACAGGTTGGACCAAGCCATGCAGGTGACAGAGA	850
EMBOSS_001	701	CTCTTGGGATGACGCACTGCTGCATCAACCCCATCATCTATGCCTTTGTC	750
EMBOSS_001	851	 CTCTTGGGATGACACACTGCTGCATCAACCCCATCATCTATGCCTTCGTC	900
EMBOSS_001	751	GGGGAGAAGTTCAGAACTACCTCTTAGTCTTCTTCCAAAAGCACATTGC	800
EMBOSS_001	901	 GGGGAGAAGTTCAGAACTACCTCTTAGTCTTCTTCCAAAAGCACATTGC	950
EMBOSS_001	801	CAAACGCTTCTGCAAATGCTGTTCTATTTTCCAGC-----	835
EMBOSS_001	951	 CAAACGCTTCTGCAAATGCTGTTCCATTTTCCAGCAAGAGGCTCCCGAGC	1000
EMBOSS_001	836	-----	835

```

EMBOSS_001      1001 GAGCAAGTTCAGTTTACACCCGATCCACTGGGGAGCAGGAAACATCTGTG      1050

EMBOSS_001      836 -----      835

EMBOSS_001      1051 GGCTTGTGA      1059

#-----
#-----

```

TASK 12

```

#####
# Program: needle
# Rundate: Mon 22 May 2023 00:05:18
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20230522-000515-0142-1745463-p2m.asequence
#   -bsequence emboss_needle-I20230522-000515-0142-1745463-p2m.bsequence
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

#=====
#
# Aligned_sequences: 2
# 1: EMOSS_001
# 2: EMOSS_001
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 1059
# Identity:      845/1059 (79.8%)
# Similarity:    845/1059 (79.8%)
# Gaps:          214/1059 (20.2%)
# Score: 5097.0
#
#
#=====

EMBOSS_001      1 -----      0

EMBOSS_001      1 ATGGATTATCAAGTGTCAAGTCCAACCTATGACATCAATTATTATACATC      50

EMBOSS_001      1 -----      0

```

EMBOSS_001	51	GGAGCCCTGCCAAAAAATCAACGTGAAGCAAATTGCAGCCCGCCTCCTGC	100
EMBOSS_001	1	-----CAACATACTG 	10
EMBOSS_001	101	CTCCGCTCTACTCACTGGTGTTCATCTTTGGTTTTGTGGGCAACATACTG	150
EMBOSS_001	11	GTCGTCCTCATCCTGATAAACTGCAAAAGGCTGAAAAGCATGACTGACAT 	60
EMBOSS_001	151	GTCGTCCTCATCCTGATAAACTGCAAAAGGCTGAAAAGCATGACTGACAT	200
EMBOSS_001	61	CTACCTGCTCAACCTGGCCATCTCTGACCTGCTTTTCCTTCTTACTGTCC 	110
EMBOSS_001	201	CTACCTGCTCAACCTGGCCATCTCTGACCTGCTTTTCCTTCTTACTGTCC	250
EMBOSS_001	111	CCTTCTGGGCTCACTATGCTGCTGCCAGTGGGACTTTGGAAATACAATG 	160
EMBOSS_001	251	CCTTCTGGGCTCACTATGCTGCTGCCAGTGGGACTTTGGAAATACAATG	300
EMBOSS_001	161	TGTCAACTCTTGACAGGGCTCTATTTTATAGGCTTCTTCTCTGGAATCTT 	210
EMBOSS_001	301	TGTCAACTCTTGACAGGGCTCTATTTTATAGGCTTCTTCTCTGGAATCTT	350
EMBOSS_001	211	CTTCATCATCCTCCTGACAATCGATAGGTACCTGGCTATCGTCCATGCTG 	260
EMBOSS_001	351	CTTCATCATCCTCCTGACAATCGATAGGTACCTGGCTATCGTCCATGCTG	400
EMBOSS_001	261	TGTTTGCTTTAAAAGCCAGGACAGTCACCTTTGGGGTGGTGACAAGTGTG 	310
EMBOSS_001	401	TGTTTGCTTTAAAAGCCAGGACAGTCACCTTTGGGGTGGTGACAAGTGTG	450
EMBOSS_001	311	ATCACTTGGGTGGTGGCTGTGTTTGCCTCTCTCCAAGAATCATCTTTAC 	360
EMBOSS_001	451	ATCACTTGGGTGGTGGCTGTGTTTGCCTCTCTCCAAGAATCATCTTTAC	500
EMBOSS_001	361	CAGATCTCAGAGAGAAGGTCTTCATTACACCTGCAGCTCTCATTTTCCAT 	410
EMBOSS_001	501	CAGATCTCAGAGAGAAGGTCTTCATTACACCTGCAGCTCTCATTTTCCAT	550
EMBOSS_001	411	ACAGTCAGTATCAATTCTGGAAGAATTTCCAGACATTAAGATAGTCATC 	460
EMBOSS_001	551	ACAGTCAGTATCAATTCTGGAAGAATTTCCAGACATTAAGATAGTCATC	600
EMBOSS_001	461	TTGGGGCTGGTCCTGCCGCTGCTTGTCATGGTCATCTGCTACTCGGGAAT 	510
EMBOSS_001	601	TTGGGGCTGGTCCTGCCGCTGCTTGTCATGGTCATCTGCTACTCGGGAAT	650
EMBOSS_001	511	CCTGAAAACCTGCTTCGGTGTGCGAACGAGAAGAAGAGGCACAGGGCTG 	560
EMBOSS_001	651	CCTGAAAACCTGCTTCGGTGTGCGAACGAGAAGAAGAGGCACAGGGCTG	700
EMBOSS_001	561	TGAGGCTCATCTTCACCATCATGATTGTTTATTTTCTTCTCTGGGCTCCC 	610
EMBOSS_001	701	TGAGGCTCATCTTCACCATCATGATTGTTTATTTTCTTCTCTGGGCTCCC	750
EMBOSS_001	611	TACAACATTGTCCTTCTCCTGAACACCTTCCAGGAATTCTTTGGCCTGAA	660

EMBOSS_001	751	TACAACTTCTCTCTGAACACCTTCCAGGAATTCTTTGGCCTGAA	800
EMBOSS_001	661	TAATTGCAGTAGCTCTAACAGTTGGACCAAGCCATGCAGGTGACAGAGA	710
EMBOSS_001	801	TAATTGCAGTAGCTCTAACAGTTGGACCAAGCCATGCAGGTGACAGAGA	850
EMBOSS_001	711	CTCTTGGGATGACACACTGCTGCATCAACCCCATCATCTATGCCTTCGTC	760
EMBOSS_001	851	CTCTTGGGATGACACACTGCTGCATCAACCCCATCATCTATGCCTTCGTC	900
EMBOSS_001	761	GGGGAGAAGTTCAGAACTACCTCTTAGTCTTCTTCCAAAAGCACATTGC	810
EMBOSS_001	901	GGGGAGAAGTTCAGAACTACCTCTTAGTCTTCTTCCAAAAGCACATTGC	950
EMBOSS_001	811	CAAACGCTTCTGCAAATGCTGTTCCATTTTCCAGC-----	845
EMBOSS_001	951	CAAACGCTTCTGCAAATGCTGTTCCATTTTCCAGCAAGAGGCTCCCGAGC	1000
EMBOSS_001	846	-----	845
EMBOSS_001	1001	GAGCAAGTTCAGTTTACACCCGATCCACTGGGGAGCAGGAAACATCTGTG	1050
EMBOSS_001	846	-----	845
EMBOSS_001	1051	GGCTTGTGA	1059
#-----			
#-----			

TASK 13

Podobieństwo jest wyższe, ponieważ usunęliśmy fragmenty, które odpowiadały za różnice w obydwu sekwencjach.

Świadczy o tym, że bez przycinania wynik porównywania dwóch różnych ciągów, jest niższy.

TASK 14

Wybieram dwie sekwencje DNA:

- Sekwencja 1: ATTAGGGGCGCCTGACAGTAAACGTTGGATAAGACTCCCAATTAGCCGAG
- Sekwencja 2: GAGTACGTAATATAATAGCGTTTAGCTATTCGGGTGCGGTCTTCAAGGAC

Wynik:

TASK 20

```
TGC-
AGTCGAGCGAACAGAGAAGGAGCTTGCTCCTTCGACGTTAGCGGCGGACGGGTGAGTAACACGTGGGCAACCTACCTTATA
GTTTGGGATAACTCCGGGAAACCGGGGCTAATACC
```

TASK 21

"." - dowolny nukleotyd, tożsamość/identity, zgodność z innymi

"-" - brak nukleotydu

Jeśli w miejscu kropki znajduje się nukleotyd, oznacza to, że jest to jednoliterowy skrót niezgodność pomiędzy badanymi sekwencjami.

TASK 22

Query: None Query ID: 1c1|Query_106413 Length: 119

Query range 1: 1 to 120

Query	1	TGC-
AGTCGAGCGAACAGAGAAGGAGCTTGCTCCTTCGACGTTAGCGGCGGACGGGTGAGTAACACGTGGGCAACCTACCTTATA		
GTTTGGGATAACTCCGGGAAACCGGGGCTAATACC	119	
NR_112628.1	29	
...	A
.....		148
NR_042072.1	39	
...	A
.....		158
NR_112569.1	29	
...	A
.....		148
NR_134074.1	21	
...	-T.....C
.....		139
NR_114207.1	29	
...	AT.....
.....		148
NR_118146.1	19	
...	-T.....C
.....		137
NR_114920.1	29	
...	AC
.....		148

TASK 23

Descriptions

Graphic Summary

Alignments

Taxonomy

Alignment view

Flat query-anchored with dots for identities

Line length: 120

Restore defaults

Download

Download

NextPreviousFirst Range

Query range 1: 1 to 120

Query	1	TGC-AGTCGAGCGAACAGAGAAGGAGCTTGCTCTTCGACGTTAGCGGGGACGGGTGAGTAACACGTGGGCAACCTACCTTATAGTTGGGATAACTCCGGGAAACCGGGGCTAATACC	1
NR_112628.1	29	...A.....	1
NR_042072.1	39	...A.....	1
NR_112569.1	29	...A.....	1
NR_134074.1	21T.....	1
NR_114207.1	29	...A.....	1
NR_118146.1	19T.....	1
NR_114920.1	29	...A.....	1

Są to:

- A
- A
- A
- T, C
- A, T
- T, C
- A, C

TASK 24

Jest to adenina (A). Wszystkie sekwencje się nim różnią od mojej, a w dwóch go brakuje (-).