# Task1

J01415.2 - accession number

# Task2

D38113.1 - accession number

# Task3

Изображение выглядит как текст, снимок экрана, Прямоугольник, астрономия

Автоматически созданное описание

Изображение выглядит как текст, диаграмма, черно-белый, документ

Автоматически созданное описание

# Task4

Na wykresie jest widoczna prosta linia, co oznacza że pozycja w jednej sekwencji ma identyczną pozycje w drugiej sekwencji. Może to pokazywać na wspólne pochodzenie lub ewolucję.

# Task5

1. Brak informacji. Mamy tylko wykres podobieństwa pomiędzy dwoma sekwencjami, bez znaczenia.
2. Nie uwzględnia sekwencji niekodujących. Prowadzi analizę tylko między sekwencjami kodującymi (genami).
3. Zależność od podanych parametrów. Wybór odpowiednich parametrów, może mieć wpływ na wynik.

# Task6

* “treshold” - wartość progowa, która jest stosowana do analizy prawdopodobieństwa pomiędzy dwoma podanymi sekwencjami.
* “Window size over which to test threshold” - rozmiar okna w którym testujemy wartość "treshold". Określa jaka duża część jest uwzględniana w pojedynczym teście progowym. Czym wyżej, tym latwiej do analizy wykres.

# Task7

Marker: rs333

Oznaczenie: Odporność na HIV

Gen receptora chemokin CCR5 odgrywa ważną rolę w wielu procesach związanych z odpornością. Delta 32 rs333, oznaczająca delecję CCR5-delta32 32 nukleotydów z genu, jest prawdopodobnie najbardziej znanym allelem CCR5. 23andMe testuje to pod swoim prywatnym identyfikatorem/nazwą, I3003626. Osoby posiadające jedną kopię allelu delta 32 są w pewnym stopniu odporne na zakażenie wirusem HIV, który powoduje AIDS, a osoby z 2 kopiami (homozygoty delta 32, ~ 1% rasy kaukaskiej) są prawie całkowicie odporne na zakażenie wirusem HIV. [PMID 8898752] Allel delta 32 mógł zostać wyselekcjonowany w populacjach europejskich, ponieważ nadaje odporność na dżumę (czarną śmierć) lub ospę.

# Task8

>3 dna:chromosome chromosome:GRCh38:3:46373053:46373887:1 GTCATCCTCATCCTGATAAACTGCAAAAGGCTGAAGAGCATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTTCCTTCTTACTGTCCCCTTCTGGGCTCACTATGCTGCCGCCCAGTGGGACTTTGGAAATACAATGTGTCAACTCTTGACAGGGCTCTATTTTATAGGCTTCTTCTCTGGAATCTTCTTCATCATCCTCCTGACAATCGATAGGTACCTGGCTGTCGTCCATGCTGTGTTTGCTTTAAAAGCCAGGACGGTCACCTTTGGGGTGGTGACAAGTGTGATCACTTGGGTGGTGGCTGTGTTTGCGTCTCTCCCAGGAATCATCTTTACCAGATCTCAAAAAGAAGGTCTTCATTACACCTGCAGCTCTCATTTTCCATACAGTCAGTATCAATTCTGGAAGAATTTCCAGACATTAAAGATAGTCATCTTGGGGCTGGTCCTGCCGCTGCTTGTCATGGTCATCTGCTACTCGGGAATCCTAAAAACTCTGCTTCGGTGTCGAAATGAGAAGAAGAGGCACAGGGCTGTGAGGCTTATCTTCACCATCATGATTGTTTATTTTCTCTTCTGGGCTCCCTACAACATTGTCCTTCTCCTGAACACCTTCCAGGAATTCTTTGGCCTGAATAATTGCAGTAGCTCTAACAGGTTGGACCAAGCTATGCAGGTGACAGAGACTCTTGGGATGACGCACTGCTGCATCAACCCCATCATCTATGCCTTTGTCGGGGAGAAGTTCAGAAACTACCTCTTAGTCTTCTTCCAAAAGCACATTGCCAAACGCTTCTGCAAATGCTGTTCTATTTTCCAGC

# Task9

Chlorocebus pygerythrus CCR5 gene, complete cds: KU382465.1 - accession number

# Task10

>KU382465.1 Chlorocebus pygerythrus CCR5 gene, complete cds ATGGATTATCAAGTGTCAAGTCCAACCTATGACATCAATTATTATACATCGGAGCCCTGCCAAAAAATCAACGTGAAGCAAATTGCAGCCCGCCTCCTGCCTCCGCTCTACTCACTGGTGTTCATCTTTGGTTTTGTGGGCAACATACTGGTCGTCCTCATCCTGATAAACTGCAAAAGGCTGAAAAGCATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGCTTTTCCTTCTTACTGTCCCCTTCTGGGCTCACTATGCTGCTGCCCAGTGGGACTTTGGAAATACAATGTGTCAACTCTTGACAGGGCTCTATTTTATAGGCTTCTTCTCTGGAATCTTCTTCATCATCCTCCTGACAATCGATAGGTACCTGGCTATCGTCCATGCTGTGTTTGCTTTAAAAGCCAGGACAGTCACCTTTGGGGTGGTGACAAGTGTGATCACTTGGGTGGTGGCTGTGTTTGCCTCTCTCCCAAGAATCATCTTTACCAGATCTCAGAGAGAAGGTCTTCATTACACCTGCAGCTCTCATTTTCCATACAGTCAGTATCAATTCTGGAAGAATTTCCAGACATTAAAGATAGTCATCTTGGGGCTGGTCCTGCCGCTGCTTGTCATGGTCATCTGCTACTCGGGAATCCTGAAAACTCTGCTTCGGTGTCGAAACGAGAAGAAGAGGCACAGGGCTGTGAGGCTCATCTTCACCATCATGATTGTTTATTTTCTCTTCTGGGCTCCCTACAACATTGTCCTTCTCCTGAACACCTTCCAGGAATTCTTTGGCCTGAATAATTGCAGTAGCTCTAACAGGTTGGACCAAGCCATGCAGGTGACAGAGACTCTTGGGATGACACACTGCTGCATCAACCCCATCATCTATGCCTTCGTCGGGGAGAAGTTCAGAAACTACCTCTTAGTCTTCTTCCAAAAGCACATTGCCAAACGCTTCTGCAAATGCTGTTCCATTTTCCAGCAAGAGGCTCCCGAGCGAGCAAGTTCAGTTTACACCCGATCCACTGGGGAGCAGGAAACATCTGTGGGCTTGTGA

# Task11

########################################

# Program: needle

# Rundate: Thu 25 May 2023 22:21:05

# Commandline: needle

# -auto

# -stdout

# -asequence emboss\_needle-I20230525-222103-0642-58210278-p1m.asequence

# -bsequence emboss\_needle-I20230525-222103-0642-58210278-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 GGAGCCCTGCCAAAAAATCAACGTGAAGCAAATTGCAGCCCGCCTCCTGC 100

EMBOSS\_001 1 -------------------------------------------------- 0

EMBOSS\_001 101 CTCCGCTCTACTCACTGGTGTTCATCTTTGGTTTTGTGGGCAACATACTG 150

EMBOSS\_001 1 GTCATCCTCATCCTGATAAACTGCAAAAGGCTGAAGAGCATGACTGACAT 50

|||.|||||||||||||||||||||||||||||||.||||||||||||||

EMBOSS\_001 151 GTCGTCCTCATCCTGATAAACTGCAAAAGGCTGAAAAGCATGACTGACAT 200

EMBOSS\_001 51 CTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTTCCTTCTTACTGTCC 100

|||||||||||||||||||||||||||||||.||||||||||||||||||

EMBOSS\_001 201 CTACCTGCTCAACCTGGCCATCTCTGACCTGCTTTTCCTTCTTACTGTCC 250

EMBOSS\_001 101 CCTTCTGGGCTCACTATGCTGCCGCCCAGTGGGACTTTGGAAATACAATG 150

||||||||||||||||||||||.|||||||||||||||||||||||||||

EMBOSS\_001 251 CCTTCTGGGCTCACTATGCTGCTGCCCAGTGGGACTTTGGAAATACAATG 300

EMBOSS\_001 151 TGTCAACTCTTGACAGGGCTCTATTTTATAGGCTTCTTCTCTGGAATCTT 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 ATCACTTGGGTGGTGGCTGTGTTTGCGTCTCTCCCAGGAATCATCTTTAC 350

||||||||||||||||||||||||||.|||||||||.|||||||||||||

EMBOSS\_001 451 ATCACTTGGGTGGTGGCTGTGTTTGCCTCTCTCCCAAGAATCATCTTTAC 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 CCTAAAAACTCTGCTTCGGTGTCGAAATGAGAAGAAGAGGCACAGGGCTG 550

|||.|||||||||||||||||||||||.||||||||||||||||||||||

EMBOSS\_001 651 CCTGAAAACTCTGCTTCGGTGTCGAAACGAGAAGAAGAGGCACAGGGCTG 700

EMBOSS\_001 551 TGAGGCTTATCTTCACCATCATGATTGTTTATTTTCTCTTCTGGGCTCCC 600

|||||||.||||||||||||||||||||||||||||||||||||||||||

EMBOSS\_001 701 TGAGGCTCATCTTCACCATCATGATTGTTTATTTTCTCTTCTGGGCTCCC 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 GGGGAGAAGTTCAGAAACTACCTCTTAGTCTTCTTCCAAAAGCACATTGC 800

||||||||||||||||||||||||||||||||||||||||||||||||||

EMBOSS\_001 901 GGGGAGAAGTTCAGAAACTACCTCTTAGTCTTCTTCCAAAAGCACATTGC 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

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

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

# Task12

######################################## # Program: needle # Rundate: Sun 21 May 2023 19:32:36 # Commandline: needle # -auto # -stdout # -asequence emboss\_needle-I20230521-194446-0624-22337215-p2m.asequence # -bsequence emboss\_needle-I20230521-194446-0624-22337215-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: 3 # 2: KU382465.1 # Matrix: EBLOSUM62 # Gap\_penalty: 10.0 # Extend\_penalty: 0.5 # # Length: 835 # Identity: 818/835 (98.0%) # Similarity: 818/835 (98.0%) # Gaps: 0/835 ( 0.0%) # Score: 4916.0 # # #======================================= 3 1 GTCATCCTCATCCTGATAAACTGCAAAAGGCTGAAGAGCATGACTGACAT 50 |||.|||||||||||||||||||||||||||||||.|||||||||||||| KU382465.1 1 GTCGTCCTCATCCTGATAAACTGCAAAAGGCTGAAAAGCATGACTGACAT 50 3 51 CTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTTCCTTCTTACTGTCC 100 |||||||||||||||||||||||||||||||.|||||||||||||||||| KU382465.1 51 CTACCTGCTCAACCTGGCCATCTCTGACCTGCTTTTCCTTCTTACTGTCC 100 3 101 CCTTCTGGGCTCACTATGCTGCCGCCCAGTGGGACTTTGGAAATACAATG 150 ||||||||||||||||||||||.||||||||||||||||||||||||||| KU382465.1 101 CCTTCTGGGCTCACTATGCTGCTGCCCAGTGGGACTTTGGAAATACAATG 150 3 151 TGTCAACTCTTGACAGGGCTCTATTTTATAGGCTTCTTCTCTGGAATCTT 200 |||||||||||||||||||||||||||||||||||||||||||||||||| KU382465.1 151 TGTCAACTCTTGACAGGGCTCTATTTTATAGGCTTCTTCTCTGGAATCTT 200 3 201 CTTCATCATCCTCCTGACAATCGATAGGTACCTGGCTGTCGTCCATGCTG 250 |||||||||||||||||||||||||||||||||||||.|||||||||||| KU382465.1 201 CTTCATCATCCTCCTGACAATCGATAGGTACCTGGCTATCGTCCATGCTG 250 3 251 TGTTTGCTTTAAAAGCCAGGACGGTCACCTTTGGGGTGGTGACAAGTGTG 300 ||||||||||||||||||||||.||||||||||||||||||||||||||| KU382465.1 251 TGTTTGCTTTAAAAGCCAGGACAGTCACCTTTGGGGTGGTGACAAGTGTG 300 3 301 ATCACTTGGGTGGTGGCTGTGTTTGCGTCTCTCCCAGGAATCATCTTTAC 350 ||||||||||||||||||||||||||.|||||||||.||||||||||||| KU382465.1 301 ATCACTTGGGTGGTGGCTGTGTTTGCCTCTCTCCCAAGAATCATCTTTAC 350 3 351 CAGATCTCAAAAAGAAGGTCTTCATTACACCTGCAGCTCTCATTTTCCAT 400 |||||||||.|.|||||||||||||||||||||||||||||||||||||| KU382465.1 351 CAGATCTCAGAGAGAAGGTCTTCATTACACCTGCAGCTCTCATTTTCCAT 400 3 401 ACAGTCAGTATCAATTCTGGAAGAATTTCCAGACATTAAAGATAGTCATC 450 |||||||||||||||||||||||||||||||||||||||||||||||||| KU382465.1 401 ACAGTCAGTATCAATTCTGGAAGAATTTCCAGACATTAAAGATAGTCATC 450 3 451 TTGGGGCTGGTCCTGCCGCTGCTTGTCATGGTCATCTGCTACTCGGGAAT 500 |||||||||||||||||||||||||||||||||||||||||||||||||| KU382465.1 451 TTGGGGCTGGTCCTGCCGCTGCTTGTCATGGTCATCTGCTACTCGGGAAT 500 3 501 CCTAAAAACTCTGCTTCGGTGTCGAAATGAGAAGAAGAGGCACAGGGCTG 550 |||.|||||||||||||||||||||||.|||||||||||||||||||||| KU382465.1 501 CCTGAAAACTCTGCTTCGGTGTCGAAACGAGAAGAAGAGGCACAGGGCTG 550 3 551 TGAGGCTTATCTTCACCATCATGATTGTTTATTTTCTCTTCTGGGCTCCC 600 |||||||.|||||||||||||||||||||||||||||||||||||||||| KU382465.1 551 TGAGGCTCATCTTCACCATCATGATTGTTTATTTTCTCTTCTGGGCTCCC 600 3 601 TACAACATTGTCCTTCTCCTGAACACCTTCCAGGAATTCTTTGGCCTGAA 650 |||||||||||||||||||||||||||||||||||||||||||||||||| KU382465.1 601 TACAACATTGTCCTTCTCCTGAACACCTTCCAGGAATTCTTTGGCCTGAA 650 3 651 TAATTGCAGTAGCTCTAACAGGTTGGACCAAGCTATGCAGGTGACAGAGA 700 |||||||||||||||||||||||||||||||||.|||||||||||||||| KU382465.1 651 TAATTGCAGTAGCTCTAACAGGTTGGACCAAGCCATGCAGGTGACAGAGA 700 3 701 CTCTTGGGATGACGCACTGCTGCATCAACCCCATCATCTATGCCTTTGTC 750 |||||||||||||.||||||||||||||||||||||||||||||||.||| KU382465.1 701 CTCTTGGGATGACACACTGCTGCATCAACCCCATCATCTATGCCTTCGTC 750 3 751 GGGGAGAAGTTCAGAAACTACCTCTTAGTCTTCTTCCAAAAGCACATTGC 800 |||||||||||||||||||||||||||||||||||||||||||||||||| KU382465.1 751 GGGGAGAAGTTCAGAAACTACCTCTTAGTCTTCTTCCAAAAGCACATTGC 800 3 801 CAAACGCTTCTGCAAATGCTGTTCTATTTTCCAGC 835 ||||||||||||||||||||||||.|||||||||| KU382465.1 801 CAAACGCTTCTGCAAATGCTGTTCCATTTTCCAGC 835 #--------------------------------------- #---------------------------------------

# Task13

W drugim przypadku procent prawdopodobieństawa jest wyższy, ponieważ usunęliśmy fragmenty z sekwencji które nie miali odpowiednika. Oznacza to że teraz sekwencje są bardzo podobne i mają mniej różnic w porównaniu do pierwszej analizie.

# Task14

Match: CAGCA

|||||

CAGCA

Mismatch: TAGC

. . . .

ACAT

Gap: TGGAGAATA

- - - - - - - - -

# Task15

Duże litery "N", "M" i inne, jest oznaczeniami symboli specjanie używanych w kontekście sekwencji DNA lun RNA. "N": oznacza niezidentyfikowany nukleotyd. Używamy w sytuacji, gdy w danej pozycji sekwencji nie jest znana konkretna zasada azotowa. "M": oznacza że w danej pozycji sekwencji może występować adenina (A) lub cytozyna (C).

# Task16

Lysinibacillus sphaericus, Lysinibacillus fusiformis: gatunek bakterii.

# Task17

Lysinibacillus sphaericus zostało odkryte po raz pierwszy w 1965 r., a dalsze badania wykazały, że głównym celem tej bakterii są komary. Lysinibacillus fusiformis został początkowo wyizolowany z powierzchni beta vulgaris przez niemieckiego biologa dr O. Gottheila w 1901 roku.

# Task18

Ribosomal RNA gene

# Task19

Z próbki gleby (BG2-10)

# Task20

TGC-AGTCGAGCGAACAGAGAAGGAGCTTGCTCCTTCGACGTTAGCGGCGGACGGGTGAGTAACACGTG GGCAACCTACCTTATAGTTTGGGATAACTCCGGGAAACCGGGGCTAATACC

# Task21

Jeśli w miejscu kropki znajduję się nukleotyd, oznacza to że w danym miejscu występuje nukleotyd tymczasowy. "." - niedopasowania (mismatch) "-" - przerwy (garp)

# Task22

Изображение выглядит как текст, снимок экрана, Шрифт, число

Автоматически созданное описание

Изображение выглядит как текст, снимок экрана, число, линия

Автоматически созданное описание

Изображение выглядит как текст, снимок экрана, Шрифт

Автоматически созданное описание

# Task23

W analizowanym gatunku nukleotydy, które rozróżniają gatunek badany od pozostałych występuje nukleotyd T, podczas gdy w pozostałych C.

# Task24

G