

lab6__Biology

Batuhan Seyhan S20343

TASK 1

AAB58943.1

TASK 2

NC_001643

```
for i in range(10):  
    if i & 2 == 0:  
        print("even")  
    else:  
        print("odd")
```

TASK 3

<https://ibb.co/746N6GC>

TASK 4

As far as i see from the generated graph it shows the similarity of our genes. So between human and chimpanze genes are very similar

TASK 5

Lack of information on the biological significance of the matches identified, requiring further analysis to interpret the results.

Increased computation time and resource requirements for comparing large sequences or multiple sequences.

TASK 6

In terms of the “threshold” and “window size over which to test threshold”, these parameters affect the sensitivity and specificity of the dotmatcher algorithm. The threshold sets the minimum score required for a match to be considered significant, while the window size determines the length of the sequence that is being compared at any given time. By adjusting these parameters, users can control the balance between false positives and false negatives in their sequence comparisons.

TASK 7

rs333

TASK 8

3 dna:chromosome chromosome:GRCh38:3:46373053:46373887:1 GTCATCCT-
CATCCTGATAAACTGCAAAAGGCTGAAGAGCATGACTGACATCTAC-
CTGCTC AACCTGGCCATCTCTGACCTGTTTTTCCTTCTTACTGTCC-
CCTTCTGGGCTCACTATGCT GCCGCCCAGTGGGACTTTGGAAATA-
CAATGTGTCAACTCTTGACAGGGCTCTATTTTATA GGCTTCTTCTCTG-
GAATCTTCTTCATCATCCTCCTGACAATCGATAGGTACCTGGCTGTC
GTCCATGCTGTGTTTGCTTTAAAAGCCAGGACGGTCACCTTTGGGGTG-
GTGACAAGTGTG ATCACTTGGGTGGTGGCTGTGTTTGCGTCTCTC-
CCAGGAATCATCTTTACCAGATCTCAA AAAGAAGGTCTTCATTACAC-
CTGCAGCTCTCATTTTCCATACAGTCAGTATCAATTCTGG AAGAATTTCCA-
GACATTAAAGATAGTCATCTTGGGGCTGGTCCTGCCGCTGCTTGTCATG
GTCATCTGCTACTCGGGAATCCTAAAAACTCTGCTTCGGTGTCGAAATGA-
GAAGAAGAGG CACAGGGCTGTGAGGCTTATCTTCACCATCATGATTGTT-
TATTTTCTCTTCTGGGCTCCC TACAACATTGTCCTTCTCCTGAACAC-
CTTCCAGGAATTCTTTGGCCTGAATAATTGCAGT AGCTCTAACAGGTTG-
GACCAAGCTATGCAGGTGACAGAGACTCTTGGGATGACGCACTGC TGCAT-
CAACCCCATCATCTATGCCTTTGTCTGGGGAGAAGTTCAGAACTACCTCT-
TAGTC TTCTTCCAAAAGCACATTGCCAAACGCTTCTGCAAATGCTGTTT-
TATTTTCCAGC

TASK 9

KU382465.1

TASK 10

KU382465.1 *Chlorocebus pygerythrus* CCR5 gene, complete cds ATGGATTAT-
CAAGTGTCAAGTCCAACCTATGACATCAATTATTATACATCGGAGCCCT-
GCCAAAAAATCAACGTGAAGCA AATTGCAGCCCGCCTCCTGCCTCCGCTC-
TACTCACTGGTGTTCATCTTTGGTTTTGTGGGCAACATACTGGTCGTCCTCA
TCCTGATAAACTGCAAAAGGCTGAAAAGCATGACTGACATCTACCTGCT-
CAACCTGGCCATCTCTGACCTGCTTTTCCTT CTTACTGTCCCCTTCTGGGGCT-
CACTATGCTGCTGCCAGTGGGACTTTGGAAATACAATGTGTCAACTCTTGACAGGGCT
CTATTTTATAGGCTTCTTCTCTGGAATCTTCTTCATCATCCTCCTGA-
CAATCGATAGGTACCTGGCTATCGTCCATGCTG TGTTTGCTTTAAAAGCCAGGACAGT-
CACCTTTGGGGTGGTGACAAGTGTGATCACTTGGGTGGTGGCTGTGTTTGCCTCT
CTCCCAAGAATCATCTTTACCAGATCTCAGAGAGAAGGTCTTCATTACAC-
CTGCAGCTCTCATTTTCCATACAGTCAGTA TCAATTCTGGAAGAATTTCCA-
GACATTAAAGATAGTCATCTTGGGGCTGGTCCTGCCGCTGCTTGTCATGGTCATCTGCT
ACTCGGGAATCCTGAAAACCTCTGCTTCGGTGTCGAAACGAGAAGAAGAG-
GCACAGGGCTGTGAGGCTCATCTTCACCATC ATGATTGTTTTATTTTCTCTTCTGGGCTC-
CCTACAACATTGTCCTTCTCCTGAACACCTTCCAGGAATTCTTTGGCCTGAA
TAATTGCAGTAGCTCTAACAGGTTGGACCAAGCCATGCAGGTGACA-
GAGACTCTTGGGATGACACACTGCTGCATCAACC CCATCATCTATGC-
CTTCGTCGGGGAGAAGTTCAGAACTACCTCTTAGTCTTCTTCCAAAAGCACATTGCCAAA
TGCAAATGCTGTTCCATTTTCCAGCAAGAGGCTCCCGAGCGAGCAAGTTCAGTT-
TACACCCGATCCACTGGGGAGCAGGA AACATCTGTGGGCTTGTGA

TASK 11

Program: needle

Rundate: Tue 9 May 2023 13:58:09

Commandline: needle

-auto

-stdout

-asequence emboss__needle-I20230509-135806-0946-40726824-
p2m.asequence

-bsequence emboss__needle-I20230509-135806-0946-40726824-
p2m.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 CCTTCTGGGCTCACTATGCTGCCGCCAGTGGGACTTTGGAAATACAATG 150 |||||||||||||||.||||||||| EMBOSS_001 251 CCTTCTGGGCTCACTATGCTGCTGCCCAGTGGGACTTTGGAAATACAATG 300

EMBOSS_001 151 TGTCAACTCTTGACAGGGCTCTATTTTATAGGCTTCTTCTCTGGAATCTT 200 |||||||||||||||||||.||||||||| EMBOSS_001 301 TGTCAACTCTTGACAGGGCTCTATTTTATAGGCTTCTTCTCTGGAATCTT 350

EMBOSS_001 201 CTTTCATCATCCTCCTGACAATCGATAGGTACCTGGCTGTCGTCCATGCTG 250 |||||||||||||||||||.||||||||| EMBOSS_001 351 CTTTCATCATCCTCCTGACAATCGATAGGTACCTGGCTATCGTCCATGCTG 400

EMBOSS_001 251 TGTTTGCTTTAAAAGCCAGGACGGTCACCTTTGGGGTGTTGACAAAGTGTG 300 |||||||||||||||.||||||||| EMBOSS_001 401 TGTTTGCTTTAAAAGCCAGGACAGTCACCTTTGGGGTGTTGACAAAGTGTG 450

EMBOSS_001 301 ATCACTTGGGTGGTGGCTGTGTTTGCGTCTCTCCCAGGAATCATCTTTAC 350 |||||||||||||||.|||||.||||||||| EMBOSS_001 451 ATCACTTGGGTGGTGGCTGTGTTTGCGTCTCTCCCAAGAATCATCTTTAC 500

EMBOSS_001 351 CAGATCTCAAAAAGAAGGTCTTCATTACACCTGCAGCTCTCATTTTCATCAT 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 CCTAAAACTCTGCTTCGGTGTCGAAATGAGAAGAAGAGGCACAGGGCTG
550 |||.|||||||||||||||||.||||||||||||||||| EMBOSS_001 651 CCTGAAAACCTCTGCTTCGGTGTC-
GAAACGAGAAGAAGAGGCACAGGGCTG 700

EMBOSS_001 551 TGAGGCTTATCTTCACCATCATGATTGTTTATTTTCTCTTCTGGGCTCCC
600 |||||.||||||||||||||||||||||||||||||||| EMBOSS_001 701 TGAGGCTCATCTTCACCATCAT-
GATTGTTTATTTTCTCTTCTGGGCTCCC 750

EMBOSS_001 601 TACAACATTGTCCTTCTCCTGAACACCTTCCAGGAATTCTTTGGC-
CTGAA 650 ||||||||||||||||||||||||||||||||||| EMBOSS_001 751 TACAACATTGTCCTTCTC-
CTGAACACCTTCCAGGAATTCTTTGGCCTGAA 800

EMBOSS_001 651 TAATTGCAGTAGCTCTAACAGGTTGGACCAAGCTATGCAGGTGACA-
GAGA 700 |||||||||||||||||||.||||||||||||| EMBOSS_001 801 TAATTGCAGTAGCTC-
TAACAGGTTGGACCAAGCCATGCAGGTGACAGAGA 850

EMBOSS_001 701 CTCTTGGGATGACGCACTGCTGCATCAACCCCATCATCTATGC-
CTTTGTC 750 |||||||||.|||||||||||||||||||||.||| EMBOSS_001 851 CTCTTGGGATGACA-
CACTGCTGCATCAACCCCATCATCTATGCCTTCGTC 900

EMBOSS_001 751 GGGGAGAAGTTCAGAAACTACCTCTTAGTCTTCTTCCAAAAGCA-
CATTGC 800 ||||||||||||||||||||||||||||||||||| EMBOSS_001 901 GGGGAGAAGTTCA-
GAAACTACCTCTTAGTCTTCTTCCAAAAGCACATTGC 950

EMBOSS_001 801 CAAACGCTTCTGCAAATGCTGTTCTATTTTCCAGC————— 835
|||||||||||||||||.|||||||
EMBOSS_001 951 CAAACGCTTCTGCAAATGCTGTTCCATTTTCCAGCAAGAGGCTCCC-
GAGC 1000

EMBOSS_001 836 ————— 835

EMBOSS_001 1001 GAGCAAGTTCAGTTTACACCCGATCCACTGGGGAGCAGGAAA-
CATCTGTG 1050

EMBOSS_001 836 ——— 835

EMBOSS_001 1051 GGCTTGTGA 1059

#—————#

TASK 12

Program: needle

Rundate: Tue 9 May 2023 14:10:07

Commandline: needle

-auto

-stdout

-asequence emboss__needle-I20230509-141004-0493-9055704-
p1m.asequence

-bsequence emboss__needle-I20230509-141004-0493-9055704-
p1m.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: EMBOSS_001 # 2: EMBOSS_001 # Matrix: EBLOSUM62 # Gap_penalty: 10.0 # Ex-
tend_penalty: 0.5 # # Length: 909 # Identity: 807/909 (88.8%) # Similarity: 807/909 (88.8%) #
Gaps: 86/909 ( 9.5%) # Score: 4841.0 # # #=====
EMBOSS_001 1 GTCATCCTCATCCTGATAAACTGCAAAAGGCTGAAGAGCATGACTGA-
CAT 50 |||||.....| EMBOSS_001 1 -----CTGATAAACTGCAAAAGGCTGAAAAGCAT
```

EMBOSS_001 51 CTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTTCCTTCTTACTGTCC 100 ||||| EMBOSS_001 39 CTACCTGCTCAACCTGGCCATCTCTGACCTGCTTTTCCTTCTTACTGTCC 88

EMBOSS_001 101 CCTTCTGGGCTCACTATGCTGCCGCCAGTGGGACTTTGGAAATACAATG 150 ||||| EMBOSS_001 89 CCTTCTGGGCTCACTATGCTGCTGCCCAGTGGGACTTTGGAAATACAATG 138

EMBOSS_001 151 TGTCAACTCTTGACAGGGCTCTATTTTATAGGCTTCTTCTCTGGAATCTT 200 ||||| EMBOSS_001 139 TGTCAACTCTTGACAGGGCTCTATTTTATAGGCTTCTTCTCTGGAATCTT 188

EMBOSS_001 201 CTTTCATCATCCTCCTGACAATCGATAGGTACCTGGCTGTCGTCCATGCTG 250 ||||| EMBOSS_001 189 CTTTCATCATCCTCCTGACAATCGATAGGTACCTGGCTATCGTCCATGCTG 238

EMBOSS_001 251 TGTTTGCTTTAAAAGCCAGGACGGTCACCTTTGGGGTGGTGACAAGTGTG 300 ||||| EMBOSS_001 239 TGTTTGCTTTAAAAGCCAGGACAGTCACCTTTGGGGTGGTGACAAGTGTG 288

EMBOSS_001 301 ATCACTTGGGTGGTGGCTGTGTTTGCGTCTCTCCCAGGAATCATCTTTAC 350 ||||| EMBOSS_001 289 ATCACTTGGGTGGTGGCTGTGTTTGCGTCTCTCCCAGGAATCATCTTTAC 338

EMBOSS_001 351 CAGATCTCAAAAAGAAGGTCTTCATTACACCTGCAGCTCTCATTTTCAT 400 ||||| EMBOSS_001 339 CAGATCTCAGAGAGAAGGTCTTCATTACACCTGCAGCTCTCATTTTCAT 388

EMBOSS_001 401 ACAGTCAGTATCAATTCTGGAAGAATTTCCAGACATTAAAGATAGTCATC 450 ||||| EMBOSS_001 389 ACAGTCAGTATCAATTCTGGAAGAATTTCCAGACATTAAAGATAGTCATC 438

EMBOSS_001 451 TTGGGGCTGGTCCTGCCGCTGCTTGTCATGGTCATCTGCTACTCGGGAAT 500 ||||| EMBOSS_001 439 TTGGGGCTGGTCCTGCCGCTGCTTGTCATGGTCATCTGCTACTCGGGAAT 488

EMBOSS_001 501 CCTAAAACTCTGCTTCGGTGTCGAAATGAGAAGAAGAGGCACAGGGCTG 550 ||||| EMBOSS_001 489 CCTGAAAACCTCTGCTTCGGTGTCGAAACGAGAAGAAGAGGCACAGGGCTG 538

EMBOSS_001 551 TGAGGCTTATCTTCACCATCATGATTGTTTATTTTCTCTTCTGGGCTCCC 600 ||||| EMBOSS_001 539 TGAGGCTCATCTTCACCATCATGATTGTTTATTTTCTCTTCTGGGCTCCC 588

EMBOSS_001 601 TACAACATTGTCCTTCTCCTGAACACCTTCCAGGAATTCTTTGGCCTGAA 650 ||||| EMBOSS_001 589 TACAACATTGTCCTTCTCCTGAACACCTTCCAGGAATTCTTTGGCCTGAA 638

EMBOSS_001 651 TAATTGCAGTAGCTCTAACAGGTTGGACCAAGCTATGCAGGTGACAGAGA 700 ||||| EMBOSS_001 639 TAATTGCAGTAGCTCTAACAGGTTGGACCAAGCCATGCAGGTGACAGAGA 688

Program: needle

Rundate: Tue 9 May 2023 14:20:09

Commandline: needle

-auto

-stdout

-asequence emboss__needle-I20230509-142006-0332-20427209-
p1m.asequence

-bsequence emboss__needle-I20230509-142006-0332-20427209-
p1m.bsequence

-datafile EBLOSUM62

-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: EMBOSS_001 # 2: EMBOSS_001 # Matrix: EBLOSUM62 # Gap_penalty: 10.0 # Ex-
tend_penalty: 0.5 # # Length: 55 # Identity: 30/55 (54.5%) # Similarity: 30/55 (54.5%) # Gaps:
10/55 (18.2%) # Score: 113.5 # # #=====
```

EMBOSS_001 1 ATC-GTTAGGTACCTAGGACGACTGTCAGTAC-GAGGCTTAAGT-
GGC 45 ..| |.|||| |||.||| |||...||. |||| |..||| |. EMBOSS_001 1 TACAGGCTAGG-ACTTTAGAC-
CTGATGGTGTGCTGAGG-TCCAGTACGAC 46

EMBOSS_001 46 TTGTT 50 ||.. EMBOSS_001 47 TTAC- 50

#-----#

TASK 15

Guanine or adenine or thymine or cytosine: N Adenine or cytosine: M Guanine or cytosine: S ##
TASK 16

Select seq OP986726.1 Lysinibacillus ## TASK 17 Isolated in 1970. Obtained from mosquito
larvae.

TASK 18

RNA

TASK 19

Bacteria ## TASK 20 CGAACAGAGAAGGAGCTTGCTCCTTCGACGTTAGCGGCG-
GACGGGTGAGTAACACGTGGGCAACCTACCTTATAGTTTGGGATAACTCCGGGAAACCGGGGGCTA
TASK 21 The symbols “.” (dot) and “-” (dash) are used to represent gaps or insertions
in a sequence alignment. A dot (.) indicates a position in the alignment where there is
no residue (nucleotide or amino acid) present. A dash (-) is also used to represent a gap
or insertion, but it typically has a different meaning. In the context of BLAST, a dash
usually represents an alignment position where there is an insertion in the query sequence
compared to the subject sequence. ## TASK 22 Query 24 CGAACAGAGAAGGAGCTTGC-TC-
CTTCGACGTTAGCGGCGGACGGGTGAGTAACACGTGGGCAACCTACCTTATAGTTTGGGATAACT
141 NR_042072.1 50-.-..... 167
NR_112628.1 40-.-..... 157 NR_112569.1
40-.-..... 157 NR_118146.1 29-
..-...T.....C..... 146 NR_042073.1 50-.-
...T.....C..... 167 NR_112627.1 40-.-...T.....
157 NR_175531.1 102 ..A.....T..T...-.-T.....C..... 209

TASK 23

C

TASK 24

A