

Pairwise Sequence Alignment Assignment

Task 1: Choose a nucleotide or protein sequence from GenBank.

Chosen sequence – Human Insulin (Accession Number- AJ009655.1)

Advanced

FASTA ▾

Homo sapiens ins gene, partial

GenBank: AJ009655.1

[GenBank](#) [Graphics](#)

>AJ009655.1 Homo sapiens ins gene, partial
AGCAGGCTGTTCACAGGGCTTTGCGTCAGGTGGGCTCAGGGTTCCAGGGTGGCTGGACCCAGGCCCC
AGCTGTGCAGCAGGGAGGACGTGGCTGGGCTCGTAAGCATGTGGGGTGAGCCAGGGGCCCAAGGCA
GGGCACATGGCCTTACGCTGCTCAGCCCTGCCTGTCTCCAGATCACTGCTCTTGCCTAGGCCCTG
TGGATGGCTCTCTGCCCCTGCTGGCGCTGCTGGCCCTTGGGGACCTGACCCAGCCGAGCCTTTGTGA
ACCAACACCTGTGGGCTCACACCTGGTGAAGCTCTTACCTAGTGTGGGGGAACGAGGCTTCTTCTA
CACACCAGACCCCGGGAGGCAGAGGACCTGCAGGGTGAGCCAGCCGCCATTGCTGCCCTGGCCG
CCCCAGCCACCCCTGCTCTGGCGCTCCACCCAGCATGGGCAGAGGGGCGAGGAGCTGCCACCCA
GCAGGGGTCAGGTGCACCTTTTAAAAAAGTTCTTGGTCAGCTCTTAAAGTGACAGCTCCCTG
TGGCCAGTCAGAATCTCAGCCTGAGGACGGTTGGCTTCGGCAGCCCGAGATACATCAGAGGGTGGG
CAGCTCTCCCTCACTGCCCCCAAAATAAGTCCCGCAGCCATTCTCCACCTCATTTGATGAC
CGCAGATTCAAGTGTTTTGTAAATAAGTCTGGGTGACCTGGGGTACAGGGTGCCCGACGCTGCTG
CCTCTGGGGAACACCCATCAGCCCCGAGGAGGGCTGGCTGCTGCTGAGTGGGCGAGCCCTGT
CGCCAGCCTCACGGCAGCTCCATAGTCAGGAGATGGGGAAGATGCTGGGGACAGGCCCTGGGGAGAAGTA
CTGGGATCAGCTGTTCAAGCTCCCACTGTGACGTGCCCGGGGCGGGGAAGGAGGTGGGACATGTGGG
CGTTGGGCTGAGTCCACACCCAGTGGGTGACCTCCCTCTAACCTGGGTCAGCCCGGCTGGAG
ATGGGTGGGAGTGCAGCTAGGGCTGGCGGCGAGCGGGCACTGTGTCTCCCTGACTGTCTCTCTGTG
TCCCTGCTCTCGCGCTGTTCGGAACTGCTGTGCGCGGACGTCCTGGCAGTGGGCGAGGTGGAGCT
GGGCGGGGCTGTGTCAGGAGCTGACGCTTGGCCCTGGAGGGTCCCTGCAGAGCGTGCGATT
GTGGAACAATGCTGTACAGCATCTGCTCCCTTACCAGCTGGAGAACTACTGCACTAGACGAGCCTG
CAGGAGCCCCACACCCCGCGCTCTGACCCGAGAGAGATGGAATAAGCCCTTGAACCAAC

Task 2: Perform BLASTN (for nucleotides) or BLASTP (for proteins) against the NCBI database

Standard nucleotide blast (blastn) of Human Insulin was performed

Standard Nucleotide BLAST

blastn blastp blastx tblastn tblastx

BLASTN programs search nucleotide databases using a nucleotide query. more...

Reset page Bookmark

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) Clear

Query subrange

From To

Or, upload file

Choose File No file chosen

Job Title

AJ009655.1 Homo sapiens ins gene, partial

Enter a descriptive title for your BLAST search

Align two or more sequences

Choose Search Set

Database

Standard databases (nr etc.) rRNA/ITS databases Genomic + transcript databases Betacoronavirus Experimental databases

Core nucleotide database (core nt)

Organism

Optional

Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown

Exclude

Optional

Models (KMP) Uncultured/environmental sample sequences

Limit to

Optional

Sequences from type material

Entrez Query

Optional

Enter an Entrez query to limit search

Program Selection

Optimize for

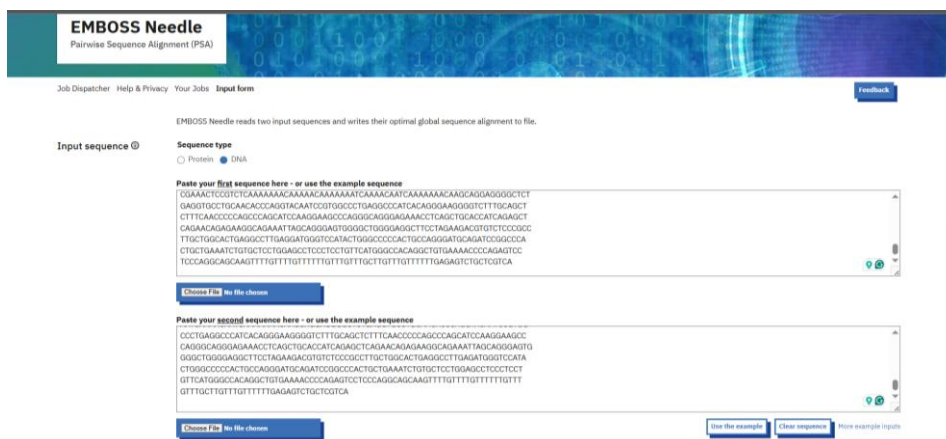
Highly similar sequences (megablast) More dissimilar sequences (discontinuous megablast) Somewhat similar sequences (blastn) Choose a BLAST algorithm

BLAST

Search database core_nt using Megablast (Optimize for highly similar sequences)

Show results in a new window

Feedback



EMBOSS Water
Pairwise Sequence Alignment (PSA)

Job Dispatcher Help & Privacy Your Jobs **Input form** [feedback](#)

Welcome to the Job Dispatcher website! If you need assistance or have feedback, please [contact us](#).

EMBOSS Water uses the Smith-Waterman algorithm (modified for speed enhancements) to calculate the local alignment of two sequences.

Input sequence ☐ Protein ☒ DNA

Paste your **first** sequence here - or use the example sequence

```

TCCAGGCTTCACGACGCTCCATAGTCAGGAGATGGGAGAGATGCTGGGACAGGCTCTGGGAGAGATG
CTGGGATCACCTGTTTCAGGCTCCCACTGTGACGCTGCCCCGGGGCGGGGAGAGGTTGGGACATGTGGG
CGTTGGGCTGTGTAGGTCCACGACGATGTGTGTGACCTCTCTTAACTGTGGTCAAGCGCGCTGGAG
ATGGGTGGGAGTGGACCTAGGGCTGGGGGGCAGCGGGGACTGTGTCTCCCTGACTGTGTCTCTCTGTG
TCCCTCTGCTCCCGCGCTGTTCCGGAACTGCTCTGCGCGGACGCTCTGGGAGTGGGGCAGGTGGAGCT
GGGCGGGGGCGCTGTGTCAGGACGCTGACGCTTGGCCCTGGAGGGTCCCTGCAGAGCGTGGGATT
GTGGACATGCTGTACCACTCTGCTCCCTCTACCACTGGAGAACTGCACTAGACGCGAGCTG
CAGGCGCCCGACACCGCGCGCTCTGACCGAGAGAGATGAATAAGCCCTTGAACAGC

```

[Choose File](#) [No file chosen](#)

Paste your **second** sequence here - or use the example sequence

```

CAGGCGAGGAGAACTCAGCTGCACCATCAGGCTCAGACAGAGAGAGGAGAAATTAGCAGGGAGTG
GGGCTGGGGAGGCTTCTTAGAAGACGTGTCTCCCGCTTGTGGCACTGAGGCTTGAGATGGGTCCATA
CTGGGCGCCGACGTCCAGGATCAGATCGGCGCACTGTGAATCTGTGCTCTGGAGCTGCTCTCT
GTTCATGGGCGACAGCTGTGAACCCGAGAGTCTCCAGGAGCAAGTTTGTGTTTGTGTTT
GTTTGTCTGTTTGTGTTTGTGAGAGTCTGCTGTCA

```

```

#####
#
# Aligned_sequences: 2
# 1: AJ009655.1
# 2: L15440.1
# Matrix: EDNAFULL
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 1393
# Identity:   1392/1393 (99.9%)
# Similarity: 1392/1393 (99.9%)
# Gaps:      0/1393 ( 0.0%)
# Score: 6956.0
#
#
#####

```

Local Alignment

```

#####
#
# Aligned_sequences: 2
# 1: AJ009655.1
# 2: L15440.1
# Matrix: EDNAFULL
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 12565
# Identity:   1392/12565 (11.1%)
# Similarity: 1392/12565 (11.1%)
# Gaps:      11172/12565 (88.9%)
# Score: 6956.0
#
#
#####

```

Global Alignment

Khushi Kiran PG Diploma Bioinformatics Class (7 -8pm batch)

| | | | |
|------------|------|--|------|
| L15440.1 | 4351 | .. | 4400 |
| AJ009655.1 | 118 | GTGAGCCCAGGGGCCCAAGGCAGGGCACCTGGCCTTCAGCCTGCCTCAG | 167 |
| L15440.1 | 4401 | | 4450 |
| AJ009655.1 | 168 | CCCTGCCTGTCTCCAGATCACTGTCTTCTGCCATGGCCCTGTGGATGC | 217 |
| L15440.1 | 4451 | | 4500 |
| AJ009655.1 | 218 | GCCTCCTGCCCTGCTGGCGCTGCTGGCCCTCTGGGGACCTGACCCAGCC | 267 |
| L15440.1 | 4501 | | 4550 |
| AJ009655.1 | 268 | GCAGCCTTTGTGAACCAACACCTGTGCGGCTCACACCTGGTGGAAGCTCT | 317 |
| L15440.1 | 4551 | | 4600 |
| AJ009655.1 | 318 | CTACCTAGTGTGCGGGGAACGAGGCTTCTTCTACACACCAAGACCCGCC | 367 |
| L15440.1 | 4601 | | 4650 |
| AJ009655.1 | 368 | GGGAGGCAGAGGACCTGCAGGGTGAGCCAACCGCCATTGCTGCCCTGG | 417 |
| L15440.1 | 4651 | | 4700 |
| AJ009655.1 | 418 | CCGCCCCAGCCACCCCTGCTCCTGGCGCTCCACCCAGCATGGGCAGA | 467 |
| L15440.1 | 4701 | | 4750 |
| AJ009655.1 | 468 | AGGGGGCAGGAGGCTGCCACCCAGCAGGGGGTCAGGTGCACTTTTTTAAA | 517 |
| L15440.1 | 4751 | | 4800 |
| AJ009655.1 | 518 | AAGAAGTTCTCTTGGTCACGTCTAAAGTGACCAGCTCCCTGTGGCCCA | 567 |
| L15440.1 | 4801 | | 4850 |
| AJ009655.1 | 568 | GTGAGCTGTACGCTACGACCTCTTCCCTTCCGACGCGGACATAC | 617 |

emboss_water-I20250330-055215-0537-18184253-p1m - Notepad

File Edit Format View Help

```
#
# Length: 1393
# Identity: 1392/1393 (99.9%)
# Similarity: 1392/1393 (99.9%)
# Gaps: 0/1393 ( 0.0%)
# Score: 6956.0
#
#
#=====
```

| | | | |
|------------|------|---|------|
| AJ009655.1 | 1 | AGCAGGTCTGTTCCAAGGGCCTTTGCGTCAGGTGGGCTCAGGGTTCCAGG | 50 |
| L15440.1 | 4284 | | 4333 |
| AJ009655.1 | 51 | GTGGCTGGACCCAGGCCCCAGCTGTGCAGCAGGGAGGACGTGGCTGGGC | 100 |
| L15440.1 | 4334 | | 4383 |
| AJ009655.1 | 101 | TCGTGAAGCATGTGGGGGTGAGCCCAGGGGGCCCCAAGGCAGGGCACCTGG | 150 |
| L15440.1 | 4384 | | 4433 |
| AJ009655.1 | 151 | CCTTCAGCCTGCCTCAGCCCTGCCTGTCTCCAGATCACTGTCTTCTGC | 200 |
| L15440.1 | 4434 | | 4483 |
| AJ009655.1 | 201 | CATGGCCCTGTGGATGCGCCTCCTGCCCTGCTGGCGCTGCTGGCCCTCT | 250 |
| L15440.1 | 4484 | | 4533 |
| AJ009655.1 | 251 | GGGGACCTGACCCAGCCGCAGCCTTTGTGAACCAACACCTGTGCGGCTCA | 300 |
| L15440.1 | 4534 | | 4583 |
| AJ009655.1 | 301 | CACCTGGTGGAAGCTCTCTACCTAGTGTGCGGGGAACGAGGCTTCTTCTA | 350 |
| L15440.1 | 4584 | | 4633 |
| AJ009655.1 | 351 | CACACCCAAGACCCGCCGGGAGGCAGAGGACCTGCAGGGTGAGCCAACCG | 400 |
| L15440.1 | 4634 | | 4683 |

emboss_water-I20250330-060025-0873-21332805-p1m - Notepad

File Edit Format View Help

```
#=====
#
# Aligned_sequences: 2
# 1: AJ009655.1
# 2: AH002844.2
# Matrix: EDNAFULL
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 1393
# Identity: 1392/1393 (99.9%)
# Similarity: 1392/1393 (99.9%)
# Gaps: 0/1393 ( 0.0%)
# Score: 6956.0
#
#=====
```

| | | | |
|------------|------|--|------|
| AJ009655.1 | 1 | AGCAGGTCTGTTCCAAGGGCCTTTGCGTCAGGTGGGCTCAGGGTTCCAGG | 50 |
| | | | |
| AH002844.2 | 2223 | AGCAGGTCTGTTCCAAGGGCCTTTGCGTCAGGTGGGCTCAGGGTTCCAGG | 2272 |
| AJ009655.1 | 51 | GTGGCTGGACCCAGGCCCCAGCTGTGCAGCAGGGAGGACGTGGCTGGGC | 100 |
| | | | |
| AH002844.2 | 2273 | GTGGCTGGACCCAGGCCCCAGCTGTGCAGCAGGGAGGACGTGGCTGGGC | 2322 |
| AJ009655.1 | 101 | TCGTGAAGCATGTGGGGGTGAGCCCAGGGGCCCAAGGCAGGGCACCTGG | 150 |
| | | | |
| AH002844.2 | 2323 | TCGTGAAGCATGTGGGGGTGAGCCCAGGGGCCCAAGGCAGGGCACCTGG | 2372 |
| AJ009655.1 | 151 | CCTTCAGCCTGCCTCAGCCCTGCCTGTCTCCCAGATCACTGTCCTTCTGC | 200 |
| | | | |
| AH002844.2 | 2373 | CCTTCAGCCTGCCTCAGCCCTGCCTGTCTCCCAGATCACTGTCCTTCTGC | 2422 |
| AJ009655.1 | 201 | CATGGCCCTGTGGATGCGCCTCCTGCCCTGCTGGCGCTGCTGGCCCTCT | 250 |
| | | | |
| AH002844.2 | 2423 | CATGGCCCTGTGGATGCGCCTCCTGCCCTGCTGGCGCTGCTGGCCCTCT | 2472 |
| AJ009655.1 | 251 | GGGGACCTGACCCAGCCGCGAGCCTTTGTGAACCAACACCTGTGCGGGCTCA | 300 |
| | | | |
| AH002844.2 | 2473 | GGGGACCTGACCCAGCCGCGAGCCTTTGTGAACCAACACCTGTGCGGGCTCA | 2522 |

emboss_water-I20250330-060314-0001-26192335-p1m - Notepad

File Edit Format View Help

```
#
#
# Aligned_sequences: 2
# 1: AJ009655.1
# 2: NG_050578.1
# Matrix: EDNAFULL
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 1394
# Identity: 1389/1394 (99.6%)
# Similarity: 1389/1394 (99.6%)
# Gaps: 1/1394 ( 0.1%)
# Score: 6919.0
#
#=====
```

| | | | |
|-------------|------|--|------|
| AJ009655.1 | 1 | AGCAGGTCTGTTCCAAGGGCCTTTGCGTCAGGTGGGCTCAGGGTTCCAGG | 50 |
| | | | |
| NG_050578.1 | 5038 | AGCAGGTCTGTTCCAAGGGCCTTTGCGTCAGGTGGGCTCAGGATTCCAGG | 5087 |
| AJ009655.1 | 51 | GTGGCTGGACCCAGGCCCCAGCTGTGCAGCAGGGAGGACGTGGCTGGGC | 100 |
| | | | |
| NG_050578.1 | 5088 | GTGGCTGGACCCAGGCCCCAGCTGTGCAGCAGGGAGGACGTGGCTGGGC | 5137 |
| AJ009655.1 | 101 | TCGTGAAGCATGTGGGGGTGAGCCCAGGGGCCCAAGGCAGGGCACCTGG | 150 |
| | | | |
| NG_050578.1 | 5138 | TCGTGAAGCATGTGGGGGTGAGCCCAGGGGCCCAAGGCAGGGCACCTGG | 5187 |
| AJ009655.1 | 151 | CCTTCAGCCTGCCTCAGCCCTGCCTGTCTCCCAGATCACTGTCCTTCTGC | 200 |
| | | | |
| NG_050578.1 | 5188 | CCTTCAGCCTGCCTCAGCCCTGCCTGTCTCCCAGATCACTGTCCTTCTGC | 5237 |
| AJ009655.1 | 201 | CATGGCCCTGTGGATGCGCCTCCTGCCCTGCTGGCGCTGCTGGCCCTCT | 250 |
| | | | |
| NG_050578.1 | 5238 | CATGGCCCTGTGGATGCGCCTCCTGCCCTGCTGGCGCTGCTGGCCCTCT | 5287 |
| AJ009655.1 | 251 | GGGGACCTGACCCAGCCGCGAGCCTTTGTGAACCAACACCTGTGCGGGCTCA | 300 |
| | | | |
| NG_050578.1 | 5288 | GGGGACCTGACCCAGCCGCGAGCCTTTGTGAACCAACACCTGTGCGGGCTCA | 5337 |
| AJ009655.1 | 301 | CACCTGGTGAAGCTCTCTACCTAGTGTGCGGGGAACGAGGCTTCTTCTA | 350 |

```

emboss_water-I20250330-060731-0076-88674023-p1m - Notepad
File Edit Format View Help
#=====
#
# Aligned_sequences: 2
# 1: AJ009655.1
# 2: AC132217.15
# Matrix: EDNAFULL
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 1394
# Identity: 1389/1394 (99.6%)
# Similarity: 1389/1394 (99.6%)
# Gaps: 1/1394 ( 0.1%)
# Score: 6919.0
#
#
#=====
AJ009655.1      1  AGCAGGTCTGTTCCAAGGGCCTTTGCGTCAGGTGGGCTCAGGGTTCCAGG      50
      |||
AC132217.15    86456 AGCAGGTCTGTTCCAAGGGCCTTTGCGTCAGGTGGGCTCAGGATTCCAGG    86505

AJ009655.1      51  GTGGCTGGACCCAGGCCAGCTGTGCAGCAGGGAGGACGTGGCTGGGC      100
      |||
AC132217.15    86506 GTGGCTGGACCCAGGCCAGCTGTGCAGCAGGGAGGACGTGGCTGGGC      86555

AJ009655.1     101  TCGTGAAGCATGTGGGGGTGAGCCAGGGGCCCAAGGCAGGGCACCTGG      150
      |||
AC132217.15    86556 TCGTGAAGCATGTGGGGGTGAGCCAGGGGCCCAAGGCAGGGCACCTGG      86605

AJ009655.1     151  CCTTCAGCCTGCCTCAGCCCTGCCTGTCTCCAGATCACTGTCTTCTGC      200
      |||
AC132217.15    86606 CCTTCAGCCTGCCTCAGCCCTGCCTGTCTCCAGATCACTGTCTTCTGC      86655

AJ009655.1     201  CATGGCCCTGTGGATGCGCCTCCTGCCCTGCTGGCGCTGCTGGCCCTCT      250
      |||
AC132217.15    86656 CATGGCCCTGTGGATGCGCCTCCTGCCCTGCTGGCGCTGCTGGCCCTCT      86705

AJ009655.1     251  GGGGACCTGACCCAGCCGAGCCTTTGTGAACCAACACCTGTGCGGCTCA      300
      |||
AC132217.15    86706 GGGGACCTGACCCAGCCGAGCCTTTGTGAACCAACACCTGTGCGGCTCA      86755

```

| Query with Targets- | Global Alignment(Smith Waterman) | | | | Local Alignment (Needleman Wunsch) | | | |
|-----------------------------|----------------------------------|-----------|--------|-------|------------------------------------|-----------|--------|-------|
| | Gaps | Identity% | Length | Score | Gaps | Identity% | Length | Score |
| (V00565.1) | 72.1% | 27.9% | 4992 | 6956 | 0% | 99.9% | 1393 | 6956 |
| L15440.1 | 88.9% | 11.1% | 12565 | 6956 | 0% | 99.9% | 1393 | 6956 |
| AH002844.2 | 72.0% | 28.0% | 4969 | 6956 | 0% | 99.9% | 1393 | 6956 |
| NG_050578.1 | 96.4% | 3.6% | 39098 | 6919 | 0.1% | 99.6% | 1394 | 6919 |
| AC132217.15 | 99.2% | 0.8% | 170027 | 6919 | 0.1% | 99.6% | 1394 | 6919 |

Table 1: Comparison of Results obtained when query was aligned with first 5 hits of blast using global and local alignment

From the table 1, the following comparison can be made:-

| | Global Alignment(Needleman Wunsch) | Local Alignment (Smith - Waterman) |
|--------------------|---|---|
| Function | Aligns entire sequence | Covers only the sequence that matches |
| Gaps | More Gaps are added in Global Alignment | Fewer gaps are added |
| Length of sequence | Length of sequence in Global Alignment is greater | Length of sequence in Global Alignment is less |
| Use | To identify closely related species | To identify conserved regions, and functional domains |
| Identity | Has lower identity and similarity percentage | Has higher identity and similarity percentage |

Approach 2: All 5 sequences are added as query in EMBOSS Clustal Omega

clustalo-l20250329-173124-0455-68786477-p1m - Notepad

File Edit Format View Help

| | | |
|-------------|--|-------|
| V00565.1 | -----AGGTGGGCTCAGGGTTCCAGGGTGGCTGGACCCAGGGCCCAGCTCTGCAGCAG | 2305 |
| | *** * **** * * ** * | |
| AC132217.15 | GG-----GTGGCTGCTCTCCTGCACCCTCTGCAGACAGCTACCTCACCTCACGTCT--- | 31726 |
| NG_050578.1 | GGAGGACGTGGCTGGGCTCGTGAAGCATGTGGGGTGAGCCAGGGGCCCCAAGGCAGGG | 5180 |
| AH002844.2 | GGAGGACGTGGCTGGGCTCGTGAAGCATGTGGGGTGAGCCAGGGGCCCCAAGGCAGGG | 2365 |
| L15440.1 | GGAGGACGTGGCTGGGCTCGTGAAGCATGTGGGGTGAGCCAGGGGCCCCAAGGCAGGG | 4426 |
| AJ009655.1 | GGAGGACGTGGCTGGGCTCGTGAAGCATGTGGGGTGAGCCAGGGGCCCCAAGGCAGGG | 143 |
| V00565.1 | GGAGGACGTGGCTGGGCTCGTGAAGCATGTGGGGTGAGCCAGGGGCCCCAAGGCAGGG | 2365 |
| | ** ***** ** * * * * * * | |
| AC132217.15 | CCCCAAGCCCTGGTCACACACTTGGCCTGTCCACTTGAAGGCCCTGCGGGCTTCTGAAAA | 31786 |
| NG_050578.1 | CACCTGGCCTTCAGCCTGCCTCAGCCCTGCCTGTCTCCAGATCACTGTCTTCTGCCAT | 5240 |
| AH002844.2 | CACCTGGCCTTCAGCCTGCCTCAGCCCTGCCTGTCTCCAGATCACTGTCTTCTGCCAT | 2425 |
| L15440.1 | CACCTGGCCTTCAGCCTGCCTCAGCCCTGCCTGTCTCCAGATCACTGTCTTCTGCCAT | 4486 |
| AJ009655.1 | CACCTGGCCTTCAGCCTGCCTCAGCCCTGCCTGTCTCCAGATCACTGTCTTCTGCCAT | 203 |
| V00565.1 | CACCTGGCCTTCAGCCTGCCTCAGCCCTGCCTGTCTCCAGATCACTGTCTTCTGCCAT | 2425 |
| | * ** * * * * * * * * * * * | |
| AC132217.15 | CTCACTTGTCTGGGACTTGTCTCCCGTCCCCCTCACCTACCCACAGTGGCCGTCTCTGG | 31846 |
| NG_050578.1 | GGCCCTGTGGATGCGCCTCCTGCCCTGCTGGCGCTGCTGGCCCTCTGGGGACCTGACCC | 5300 |
| AH002844.2 | GGCCCTGTGGATGCGCCTCCTGCCCTGCTGGCGCTGCTGGCCCTCTGGGGACCTGACCC | 2485 |
| L15440.1 | GGCCCTGTGGATGCGCCTCCTGCCCTGCTGGCGCTGCTGGCCCTCTGGGGACCTGACCC | 4546 |
| AJ009655.1 | GGCCCTGTGGATGCGCCTCCTGCCCTGCTGGCGCTGCTGGCCCTCTGGGGACCTGACCC | 263 |
| V00565.1 | GGCCCTGTGGATGCGCCTCCTGCCCTGCTGGCGCTGCTGGCCCTCTGGGGACCTGACCC | 2485 |
| | * ** * * * * * * * * * * * | |
| AC132217.15 | TCCCCCTTCCCCATCCCCAAGCGCCCTGCCGTCCTCCCT-----CACCTACCCCCC | 31901 |
| NG_050578.1 | AGCCGACGCTTTGTGAACCAACACCTGTGCGGCTCACACCTGGTGAAGCTCTCTACCT | 5360 |
| AH002844.2 | AGCCGACGCTTTGTGAACCAACACCTGTGCGGCTCACACCTGGTGAAGCTCTCTACCT | 2545 |
| L15440.1 | AGCCGACGCTTTGTGAACCAACACCTGTGCGGCTCACACCTGGTGAAGCTCTCTACCT | 4606 |
| AJ009655.1 | AGCCGACGCTTTGTGAACCAACACCTGTGCGGCTCACACCTGGTGAAGCTCTCTACCT | 323 |
| V00565.1 | AGCCGACGCTTTGTGAACCAACACCTGTGCGGCTCACACCTGGTGAAGCTCTCTACCT | 2545 |
| | * ** * * * * * * * * * * * | |
| AC132217.15 | AGCACCCGTCCCAGTCCCCTTACCTACTCCCCAGCGGCTTCCGGGTCCCCTCTCACCT | 31961 |
| NG_050578.1 | AGTGTGCGGGGAACGAGGCTTCTTCTACACACCAAGACCCGCCGGGAGGCAGAGGACCT | 5420 |
| AH002844.2 | AGTGTGCGGGGAACGAGGCTTCTTCTACACACCAAGACCCGCCGGGAGGCAGAGGACCT | 2605 |
| L15440.1 | AGTGTGCGGGGAACGAGGCTTCTTCTACACACCAAGACCCGCCGGGAGGCAGAGGACCT | 4666 |
| AJ009655.1 | AGTGTGCGGGGAACGAGGCTTCTTCTACACACCAAGACCCGCCGGGAGGCAGAGGACCT | 383 |
| V00565.1 | AGTGTGCGGGGAACGAGGCTTCTTCTACACACCAAGACCCGCCGGGAGGCAGAGGACCT | 2605 |
| | ** ** * **** * * * * * * | |

The following identity matrix from EMBOSS Clustal Omega was extracted to identify percentage identity

| | 1 | 2 | 3 | 4 | 5 |
|----------------|--------|--------|--------|--------|--------|
| 1: NG_050578.1 | 100.00 | 88.66 | 99.28 | 99.71 | 98.53 |
| 2: AH002844.2 | 88.66 | 100.00 | 89.41 | 99.93 | 89.91 |
| 3: L15440.1 | 99.28 | 89.41 | 100.00 | 99.93 | 99.08 |
| 4: AJ009655.1 | 99.71 | 99.93 | 99.93 | 100.00 | 99.93 |
| 5: V00565.1 | 98.53 | 89.91 | 99.08 | 99.93 | 100.00 |

The highlighted part (yellow) is the alignment of query insulin with different targets using Clustal Omega

| Query - AJ009655.1 | Clustal Omega Percentage Identity of queries with AJ009655.1 | Percentage Identity observed in Smith Waterman algorithm |
|-----------------------|---|--|
| NG_050578.1 | 99.71 | 99.6% |
| AH002844.2 | 99.93 | 99.9% |
| L15440.1 | 99.93 | 99.9% |
| V00565.1 | 99.93 | 99.9% |

Table 3:Percentage Identity obtained from ClustalW and Smith Waterman Algorithm

Conclusion:

Hence the comparison showed that results obtained from both approaches are similar

Task 5: Discuss the significance of high vs. low E-values.

| Criteria | Low E value | High E values |
|--------------|---------------------------------|--|
| Significance | Alignment is homologous | Alignment not biologically significant |
| Seen in | Related species, conserved gene | Unrelated sequences |
| Meaning | Strong Match | Weak match |
| Example | 1e-10, value close to 0,0 | Above 1 |