

Practical Assignment BSE322

Tanishq Gupta, 190894

Sequence

MESIIKLLKFCGFWGRPYQKFSWQPLCHIAVLVCLLAPGVIFIVRNSSNFASAIASAAIESMGFINTIL
LGTTMLYHRSALENAYGDIRIALRIGKSSSIGDVQRNIEFLEKSTNFLFKGYTVFQSVVGTGYALTIPSL
TVVYYVQTGQWPPLHGFIEADFFVDFDTTNVWLWVLVIAVGMFAMLCCLISVLVIVSSFNWSFLHYIIGLF
KLVHIRISRLNAFANPQSRQMELIEIVKLQELVYRCARTAEDTLNLFLLTQFGTCVVAICLTMMTLTLAS
NDQDLLIKMILMLAYILFNIFVYSMLGEELIATSTSLAEAAAGTQWYEWISIPEQRNILFIVRRSQKTAAL
TTGKFFAVNRSTFAATLQAAYSNFTVLRQMVHSH

Question 1:

My sequence represents full length of *odorant receptor 47* protein.
I used **Blast** tool and **NR** database to get the following results.

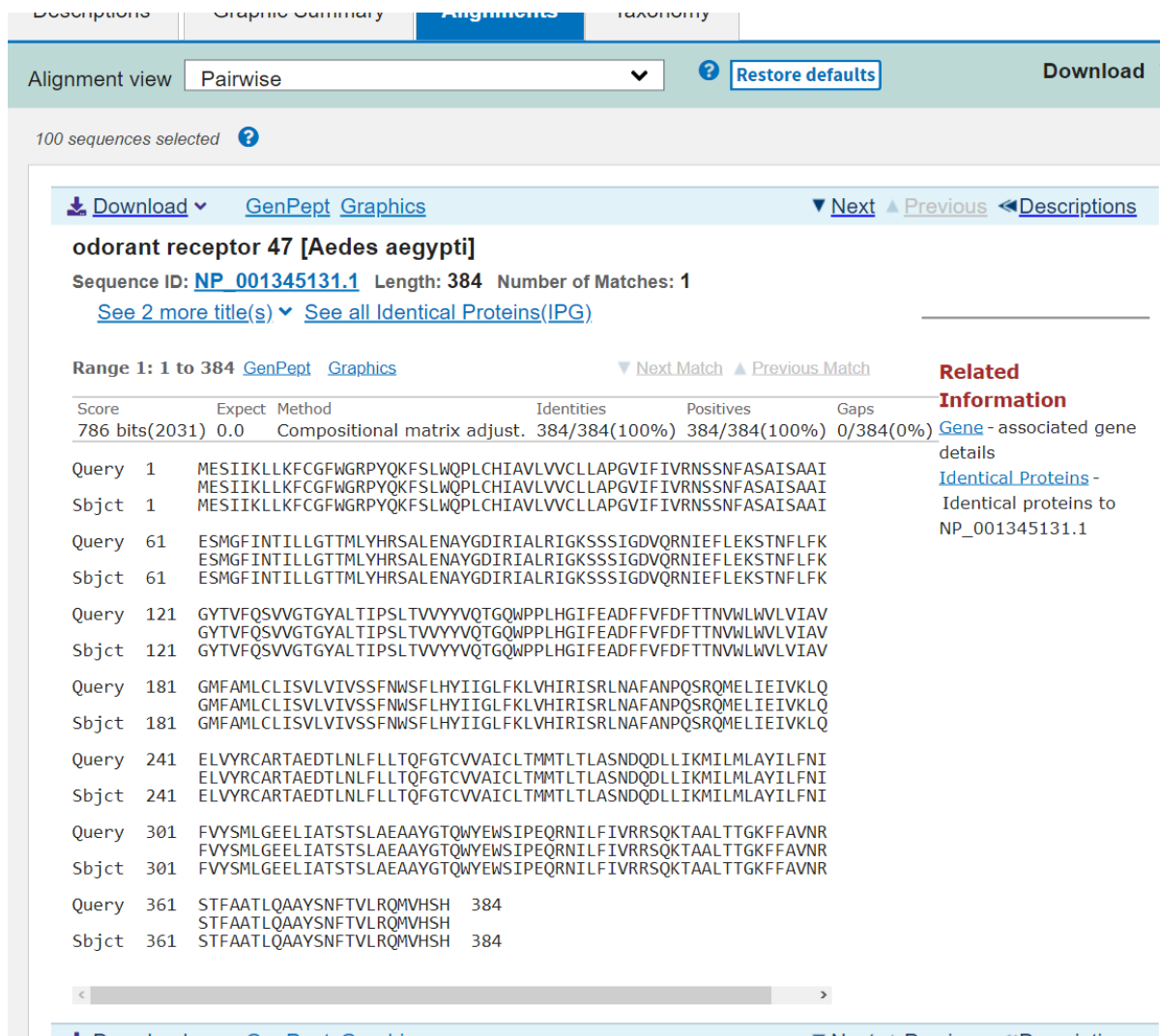


Figure 1: Alignment

Question 2:

NCBI Reference Sequence: **NP_001345131.1**

UniPort accession number: **J9HTA3**

It is found in *Aedes aegypti*, commonly known as the Yellow Fever mosquito.

Question 3:

As you can see from the image, I was unable to find anything relevant to the protein I searched. I think the reason could be that as my Protein is not found in HUMAN and the OMIM base.....

The screenshot shows the OMIM (Online Mendelian Inheritance in Man) search interface. The search term 'odorant receptor 47' is entered in the search bar. The results page shows 7,649 entries. The first 10 results are listed, each with a gene symbol, name, and genomic location. The results are as follows:

- 1: * 164342. **OLFACTORY RECEPTOR, FAMILY 1, SUBFAMILY D, MEMBER 2; OR1D2**
Cytogenetic location: 17p13.3, Genomic coordinates (GRCh38): 17:3,088,483-3,104,421
Matching terms: odorant, receptor
▶ Links
- 2: * 611538. **OLFACTORY RECEPTOR, FAMILY 7, SUBFAMILY D, MEMBER 4; OR7D4**
Cytogenetic location: 19p13.2, Genomic coordinates (GRCh38): 19:9,210,275-9,219,588
Matching terms: odorant, receptor
▶ Links
- 3: * 164320. **ODORANT-BINDING PROTEIN 2A; OBP2A**
Cytogenetic location: 9q34.3, Genomic coordinates (GRCh38): 9:135,544,811-135,549,968
Matching terms: odorant, receptor
▶ Links
- 4: * 602702. **PERILIPIN 3; PLIN3**
Cytogenetic location: 19p13.3, Genomic coordinates (GRCh38): 19:4,838,340-4,867,666
Matching terms: 47, receptor
▶ Links
- 5: * 608950. **MAGI2 INTRONIC TRANSCRIPT; MAGI2IT**
Cytogenetic location: 7q21, Genomic coordinates (GRCh38): 7:77,900,000-98,400,000
Matching terms: 47, receptor
▶ Links
- 6: * 615428. **DEAD-BOX HELICASE 47; DDX47**
Cytogenetic location: 12p13.1, Genomic coordinates (GRCh38): 12:12,813,345-12,829,980
Matching terms: 47, receptor
▶ Links
- 7: # 617166. **DEVELOPMENTAL AND EPILEPTIC ENCEPHALOPATHY 47; DEE47**
Cytogenetic location: 3q28-q29
Matching terms: 47
▶ Phenotype-Genotype Relationships ▶ Phenotypic Series ▶ ICD+ ▶ Links
- 8: * 600610. **GA-BINDING PROTEIN TRANSCRIPTION FACTOR, BETA SUBUNIT; GABPB**
GA-BINDING PROTEIN TRANSCRIPTION FACTOR, BETA SUBUNIT 1, INCLUDED; GABPB1, INCLUDED
Cytogenetic location: 15q21.2, Genomic coordinates (GRCh38): 15:50,275,388-50,359,305
Matching terms: 47
▶ Links
- 9: * 608803. **GAP JUNCTION PROTEIN, GAMMA-2; GJC2**
Cytogenetic location: 1q42.13, Genomic coordinates (GRCh38): 1:228,149,929-228,159,825
Matching terms: 47
▶ Gene-Phenotype Relationships ▶ Links
- 10: * 609335. **ODORANT RESPONSE ABNORMAL 4, C. ELEGANS, HOMOLOG OF; ODR4**
Cytogenetic location: 1q31.1, Genomic coordinates (GRCh38): 1:186,375,761-186,421,374
Matching terms: odorant
▶ Links

Navigation links: « First | < Previous | Next > | Last »

NOTE: OMIM is intended for use primarily by physicians and other professionals concerned with genetic disorders, by genetics researchers, and by scientists in science and medicine. While the OMIM database is open to the public, users seeking information about a personal medical or genetic condition are advised to consult with a qualified physician for diagnosis and for answers to personal questions.

OMIM® and Online Mendelian Inheritance in Man® are registered trademarks of the Johns Hopkins University.

Figure 2: Results of OMIM

Question 4:

As we can see from the picture below:

Chromosome number : 1

Location : NC_035107.1(268708662..268720133)

GeneID : 23687499

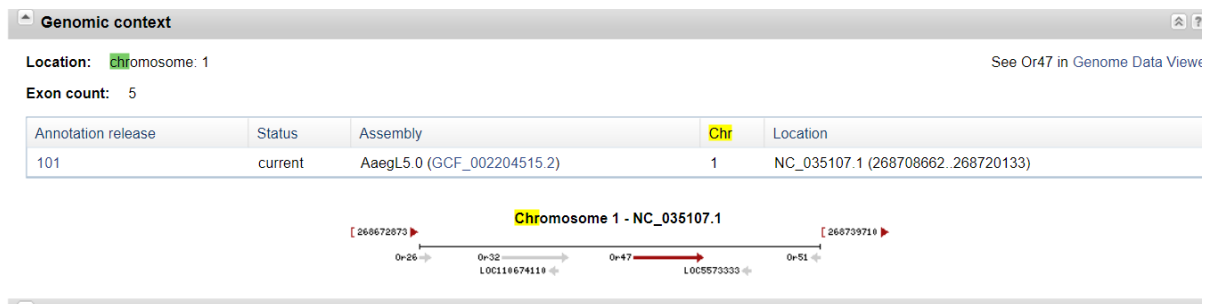


Figure 3: Genomic Context

Question 5:

Number of exons were 5.

Number of exons were 4.

I found this from NCBI database.

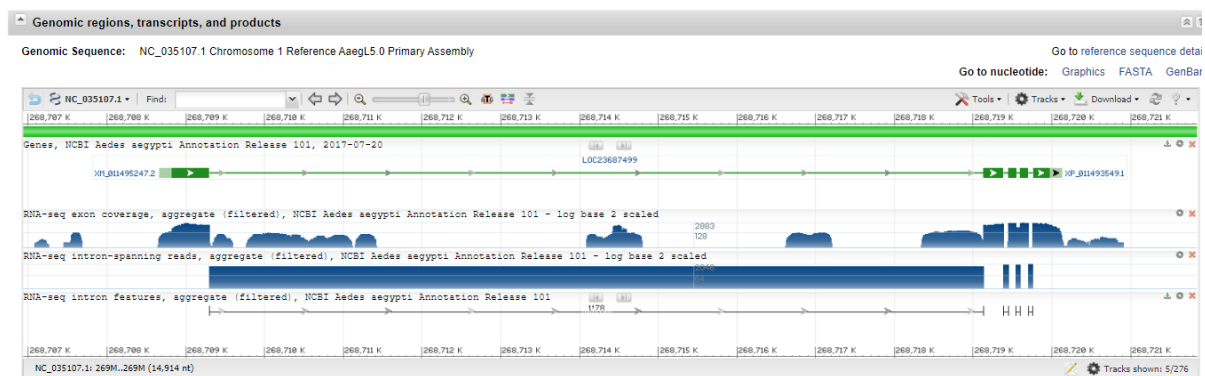


Figure 4: Pictorial view of the gene-structure

Question 6:

I tried to google and saw various results but I was unable to find my protein's structure. As you can see from the below image the structure is not present at Uniport as well.

Display [Help video](#) [BLAST](#) [Align](#) [Format](#) [Add to](#)

Entry **Protein** **Odorant receptor**
Gene **GPROR47**
Organism *Aedes aegypti* (Yellowfever mosquito)
Status [Unreviewed](#) - Annotation

None

☒ Function
☒ Names & Taxonomy
☒ Subcellular location
☐ Pathology & Biotech
☐ PTM / Processing
☐ Expression
☐ Interaction
☐ Structure

Functionⁱ

Caution

Lacks conserved residue(s) required for the

GO - Molecular functionⁱ

- odorant binding [Source: InterPro](#)
- olfactory receptor activity [Source: UniProt](#)

[Complete GO annotation on QuickGO ...](#)

GO - Biological processⁱ

- signal transduction [Source: UniProt](#)

Figure 5: Uniport Result

Question 7:

Query cover: 98%.

This means that 98% of our query sequence was aligned.

E-value: 0.0

Very good score, this shows that these sequence alignment is not random.

Percent Identity: 78%

This implies that 78% of the query sequence was identical to the matching sequence found.

[Download](#) [GenPept](#) [Graphics](#)

putative odorant receptor 92a [Aedes albopictus]
Sequence ID: [XP_029722553.1](#) Length: 384 Number of Matches: 1

Range 1: 1 to 384 [GenPept](#) [Graphics](#) [Next Match](#)

	Score	Expect	Method	Identities	Positives	Gaps
	635 bits (1639)	0.0	Compositional matrix adjust.	301/384 (78%)	345/384 (89%)	0/384 (0%)
Query	1	MESIIKLLKFCGFWGRPYOKFSLWOPLCHTAVLVCLLAPGVTFIVRNSSNFASASAAI				60
Sbjct	1	MESIIKLLKFCGFWGRPYOKFSLWOPLCHTAVLVCLLAPGVTFIVRNSSNFASASAAI				60
Query	61	ESMGFINILLGTTMLYHRSALENAYGDIRAL+KS++V+RNIFLEKSTNLFK				120
Sbjct	61	ESMGFINILLGTTMLYHRSALENAYGDIRAL+KS++V+RNIFLEKSTNLFK				120
Query	121	GYTVFQSVVGTGYALTIPSLTVVYVOTGOWPLHGFIEADFFVDFDTTMMWLVLIAV				180
Sbjct	121	GYTVFQSVVGTGYALTIPSLTVVYVOTGOWPLHGFIEADFFVDFDTTMMWLVLIAV				180
Query	181	GMFAMLCISLVIVSSFWNSFLHYIIGLFKLVHTRISRLNANFANPQSRQMLTEIVKLO				240
Sbjct	181	GMFAMLCISLVIVSSFWNSFLHYIIGLFKLVHTRISRLNANFANPQSRQMLTEIVKLO				240
Query	241	ELVYRCARTAEEDTLNLFLLTQGTGCVVACLTMWTLTASHNDQDLIKHMLAYILFNI				300
Sbjct	241	ELVYRCARTAEEDTLNLFLLTQGTGCVVACLTMWTLTASHNDQDLIKHMLAYILFNI				300
Query	301	FVYSMLGEELIATSTSLAAEAYGTQWYEWISPEQRNIFIVRSQKTAALTGKFFVNR				360
Sbjct	301	FVYSMLGEELIATSTSLAAEAYGTQWYEWISPEQRNIFIVRSQKTAALTGKFFVNR				360
Query	361	STFAATLQAAYSNFTVLRQWHS				384
Sbjct	361	STFAATLQAAYSNFTVLRQWHS				384

Figure 6: Hit 1

Query cover: 10%

This means that 10% of our query sequence was aligned.

E-value: 2.2

Very bad score, this shows that these sequence alignment can be random.

Percent Identity: 39%

This implies that 39% of the query sequence was identical to the matching sequence found.

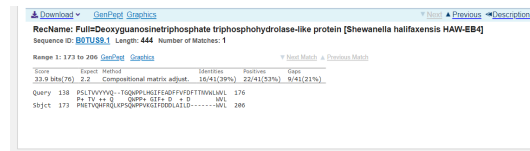


Figure 7: Hit 2

Question 8:

In Hit 1 we can see a diagonal with gaps in between the because sequences are quite similar.

In Hit 2 it is empty because the sequences are very different and we cannot see any 10 amino acid sequence matching.

Like we expected on reducing word size to 3 can see more points because probability of match's increase.

In hit 2 we see some dots as there are parts of sequences that match for at least 3 consecutive sites. However it is not much of use.

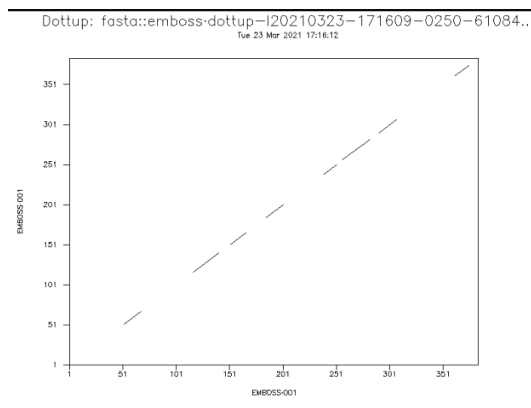


Figure 8: Hit 1—Win-
dow 10

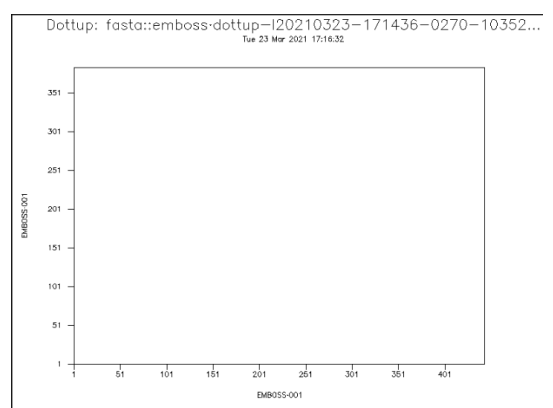


Figure 9: Hit 2—Win-
dow 10

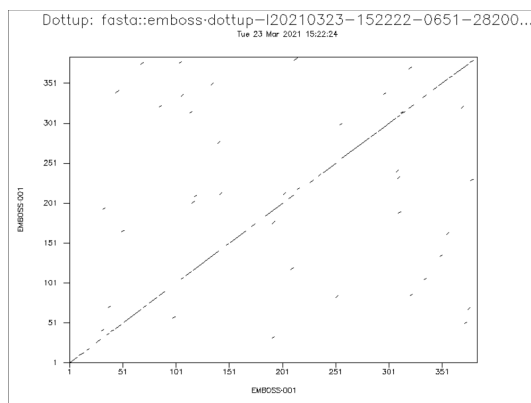


Figure 10: Hit 1—Win-
dow 3

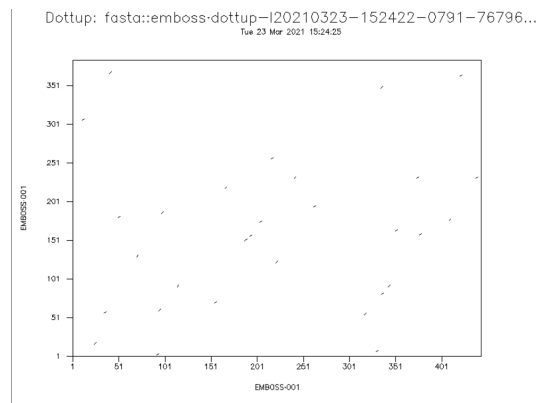


Figure 11: Hit 2—Win-
dow 3

Question 9:

Due to the fact that BLAST is based on a heuristic algorithm, the results received through BLAST, in terms of the hits found, may not be the best possible results, as it will not provide you with all the hits within the database. as it will not provide you with all the hits within the database.

It will not provide you with all the hits within the database however, on the other hand Smith-Waterman will catch them.

Also it provides better accuracy.

```
#####
# Program: water
# Runtime: Tue 23 Mar 2021 15:32:07
# Commandline: water
#
# -auto
# -stdout
# -sequence emboss_water-I20210323-153249-0090-94112711-plm.asequence
# -bsequence emboss_water-I20210323-153249-0090-94112711-plm.bsequence
# -datafile EBL0SUM62
# -gapopen 10.0
# -gapextend 0.5
# -aformat3 pair
# -sprotein1
# -sprotein2
#
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned sequences: 2
# 1: EMB0SS_001
# 2: EMB0SS_001
# Matrix: EBL0SUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 383
# Identity: 301/383 (78.6%)
# Similarity: 345/383 (90.1%)
# Gaps: 0/383 ( 0.0%)
# Score: 1592.0
#
#####

EMB0SS_001 1 MESIEXLKQCFGWSHPRQYKFLSLWQIQLZAVLVWICLAPGVFVIVNNS 50
EMB0SS_001 1 MESIEXLKQCFGWSHPRQYKFLSLWQIQLZAVLVWICLAPGVFVIVNNSF 50

EMB0SS_001 51 NFASIASAATESMGFINVTGLTMTLVHRSALENAYGDIALRLTGKSSS 100
EMB0SS_001 51 DFASIASAATESMGFINVTGLTMTLVHRSALENAYGDIALRFLAQKASAA 100

EMB0SS_001 101 IGVDRNIEELFEKSTNF.LFKGYTVFQSVVGTVGYALTZPSLTVVVVYGTQ 150
EMB0SS_001 101 NAEVLNRXKLFKESFTFLFKGYTVFQSVVGTVGYALTZPSLTVVHYTKTG 150

EMB0SS_001 151 WPLPHIGFEADFVFDFPTNNWLVLVIAVGFAMLLCSVLIVVSSFNW 200
EMB0SS_001 151 LPLPHIGFEADFVFDFPTTKFKWLVVIVIGSGFMLCLISVLIVVSSFNW 200

EMB0SS_001 201 SFLHYIIGLFLKLVHIRISRLNFAANFQSRQMLEIEIKVLQVLYRCARTA 250
EMB0SS_001 201 SLVHYIIGLFLKLVHIRISLSDLDTQDSQRKELNEVLQELVLYRCARTA 250

EMB0SS_001 251 EDTLNLFLLTFQGTCCAICLTMMTLT.LASNDDQLLXIMLMLAYILFNI 300
EMB0SS_001 251 ENALNLFLLTFQGTCCAICLTMMTLT.LASNDDQLLXIMLMLAYILFNI 300

EMB0SS_001 301 FVYSMLGQELISTSTSLAEAYGTQWYENIEPQRNLIIVRSKQTAAL 350
EMB0SS_001 301 FVYSMLGQELISTSTSVADAYGTQWYQNSLSEQRNLIIVRSKQMAAL 350

EMB0SS_001 351 TTGKFAANWRTFAATLQAAYSNFTLRQPHWS 383
EMB0SS_001 351 TTGKFFVNNRTFAATLQAAYSNFTLRQMLDS 383

#####
```

Figure 12: Hit 1


```

#####
# Program: water
# Runday: Tue 23 Mar 2021 15:40:45
# Commandline: water
# -auto
# -stdout
# -asequence emboss_water-120210323-154043-0418-95993958-p2m.asequence
# -bsequence emboss_water-120210323-154043-0418-95993958-p2m.bsequence
# -datafile EPM4250
# -gapopen 10.0
# -gapextend 0.5
# -eformat3 pair
# -sprotein1
# -sprotein2
# Align_format: pair
# Report_file: stdout
#####

#-----
#
# Aligned_sequences: 2
# 1: EMBOS_001
# 2: EMBOS_001
# Matrix: EPM4250
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 466
# Identity: 78/466 (16.7%)
# Similarity: 171/466 (36.7%)
# Gaps: 173/466 (37.1%)
# Score: 198.0
#
#-----
EMBOS_001 6 KLLKFCGW-----GRPYQKFSLWQPLCHTAVLVCLLAPGVIFIV 46
EMBOS_001 31 RILHSAAFRLQAKTQVLGVGNDF-YRTRLTH--SLEVSQIGTGIRAQL 77
EMBOS_001 47 RHSSNFASAIASAIESMGFTINTLLGTTHLVHRSALENAYGDRIALRIG 96
EMBOS_001 78 KQK--QPEFNPLNLSHSIESLCLANDI--GHPPFGHG-GEIALNYYMR 121
EMBOS_001 97 KSSSIGDVQRNIEFLEKSTNLFKGYTVF-----QSVVGT-GYALTIP 138
EMBOS_001 122 AHGGFEGNGQIFRLTR-----LEPYTEFYGMNLCRRTLLGLKYPAPYS 166
EMBOS_001 139 SL-----TVVYY--VQTGQMPPLHGIFEADFFVFDFTTNMLMVL--- 176
EMBOS_001 167 RLREQPHETVQHFRQLKPSQPPVKGIFFODDLAILD-----NMLEPL 209
EMBOS_001 177 -----VIAVG-----MFAML-CLT-----SV 191
EMBOS_001 210 SKADKQLFLSSYTVADGQHKRTRYKSLDLSIMELADDIAYAVHDEDAIV 259
EMBOS_001 192 LVIVSSFM-----SFLHYITGLF-----K 211
EMBOS_001 260 MGIVSEQWHNDVTQVLCNSEENLQIEFATHSLRLFSAKHQKDAIGT 309
EMBOS_001 212 LVH-----IRISRLNANPQSRQHELIEIVKLQELVYRCARTAEIDLNL 256
EMBOS_001 310 LVNGFVTATISINEVEGFDEPL-----LKYNAALEPAFDIAL 345
EMBOS_001 257 FLITQFGTCVAICLTMMTLTLASNDQDLXILMLAY----ILFNIEV 302
EMBOS_001 346 NVLKQF-----VFKYVIRKPE-----IQMLEYKGHQIWMELF- 377
EMBOS_001 303 YSMIGEELIATSTSLAAAYGTQYVENSIPQRNILFIVRRSQKTAALT 352
EMBOS_001 378 ----EAFISDPERLLPLNTQERN----IASEKQ----GENSHR-----VI 410
EMBOS_001 353 GKFFAVNRSTFAATLQ 368
EMBOS_001 411 ADYISGHTDEFAARLH 426

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

```

Figure 15: Hit 2

Question 11:

The difference in results can be very well predicted, as the gap penalty is reduced from 10 to 1 and they are made same as gap extension, the frequency of gaps are much more frequent and we can see a lot many extended gaps then previous alignment.

```
#####
# Program: water
# Rdate: Tue 23 Mar 2021 15:25:47
# Commandline: water
# -auto
# -stdout
# -asequence emboss_water-I20210323-152932-0977-31135918-plm.asequence
# -bsequence emboss_water-I20210323-152932-0977-31135918-plm.bsequence
# -datafile EBL0SUM62
# -gapopen 1.0
# -gapextend 1.0
# -aformat3 pair
# -sprotein1
# -sprotein2
# Align_format: pair
# Report_file: stdout
#####

#-----
#
# Aligned_sequences: 2
# 1: EMBOS_001
# 2: EMBOS_001
# Matrix: EBL0SUM62
# Gap_penalty: 1.0
# Extend_penalty: 1.0
#
# Length: 389
# Identity: 302/389 (77.6%)
# Similarity: 346/389 (88.9%)
# Gaps: 12/389 ( 3.1%)
# Score: 1598.0
#
#-----

EMBOS_001      1 MESIIKLLKCGFWGRPYQ--KFSLWQPLCHIAVLVCLLAPGVIFVRN 48
                ||| ||| ||| ||| ||| : ||| ||| ||| ||| ||| ||| |||
EMBOS_001      1 MESIIKLLQFCGFWSHPYQGHR--LWKPLCHIGVIMCLLGPVGFIVRN 48
                ||| ||| ||| ||| ||| : ||| ||| ||| ||| ||| ||| |||

EMBOS_001      49 SSNFASATSAATESMGFINILLGTTMLYHRSALENAYGDIRIALRIGKS 98
                |.: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
EMBOS_001      49 SDFASATSAATESMGFINIVLLGTTMLYHRSALSAAYGDIRFALQTAKS 98
                |.: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

EMBOS_001      99 SSI-GDVQRNIEFLEKSTNIFLFGYTVFQSVGTGYALTIPSLTVAYYVQ 147
                :.: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
EMBOS_001      99 AA-NAEVLRNIXFLEKSTFLFGYTVFQSVGTGYALTIPSLTLVHYTK 147
                :.: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

EMBOS_001      148 TGQWPLHGFIEADFFVDFDTTNWVLWLVIAVGMFAMLCILISVLIVSS 197
                |||.: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
EMBOS_001      148 TGQLPPLHGFIEADFFVDFDTTKFWLWVVIIVIGSFHMLCLISVLIVSS 197
                |||.: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

EMBOS_001      198 FNMSFL-HYIIGLFLVHIRISR-LNAFANPQSRQMELI-EIVKLQELVY 244
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
EMBOS_001      198 FNMS-LVHYLIGLFLKIVHSRLS-CLDDLTDQQRQKEL-NEIVLLQELVY 244
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

EMBOS_001      245 RCARTAEOTLNLFLLTQFGTCVVAICLTMMTLTLASNDQDLTKMILMLA 294
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
EMBOS_001      245 RCARTAEALNIFLLTQFGTCVVAICLTMMTLTLASNDRLLIKWLMLA 294
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

EMBOS_001      295 YILFNIFVYSMLGEELIATSTSLAEAYGTQWYEWSTPEQRNILFIVRRS 344
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
EMBOS_001      295 YILFNIFVYSMLGQELISTSTSVADAAYGTRWYDWSLSEQRNVLFVVSRS 344
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

EMBOS_001      345 QKTAALTGKFFAVNRSTFAATLQAAYSNTVLRQNVHS 383
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
EMBOS_001      345 QKMAALTGKFFVNRATFAATLQAAYSNTILRQMLDS 383
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

#-----
#-----
```

Figure 16: Hit 1

```
#####
# Program: water
# Rundate: Tue 23 Mar 2021 15:26:42
# Commandline: water
#
# -auto
# -stdout
# -asequence emboss_water-I20210323-152639-0507-98004192-p2n.asequence
# -bsequence emboss_water-I20210323-152639-0507-98004192-p2n.bsequence
# -datafile EBLOSUM62
# -gapopen 1.0
# -gapextend 1.0
# -format3 pair
# -sprotein1
# -sprotein2
# Align_format: pair
# Report_file: stdout
#####

#-----
#
# Aligned_sequences: 2
# 1: EMBOS_001
# 2: EMBOS_001
# Matrix: EBLOSUM62
# Gap_penalty: 1.0
# Extend_penalty: 1.0
#
# Length: 541
# Identity: 136/541 (25.1%)
# Similarity: 206/541 (38.1%)
# Gaps: 270/541 (49.9%)
# Score: 515.0
#
#-----

EMBOS_001      1 M-ESI-----I---KLLKFC-GFWG-R-PYQKF--S-LMQPLCH-IAV-- 32
      | :| | : | | : . . | ||| : : : | | .| .
EMBOS_001      1 MTQSIWIERLGEDK-LR--RN--DHRNPYQR-DRARI--L-HSAAFRR 40
      | . | | || : | . | :| |...:|.|. : |
EMBOS_001      33 L--V-VCLLAPGV-I--FI-VR---NS---SNFASATSAAT-----E-- 61
      | . | | || : | . | :| |...:|.|. : |
EMBOS_001      41 LQAKTV--L--GVGMIDFYRTLTHSLEVSQGTGIRAQKQKQPEFNP 86
      | | | | : | | . | :| |...:|.|. : |
EMBOS_001      62 ---SNGFINTI-L--LG-TTNLY-H--RSALENAYGD-IRIALRI-G-K 97
      | | | | : | | . | :| | | | :| | . | :
EMBOS_001      87 LLNISMSTIESLCLAHDIGHPP--FGHGGEIAL-N-Y--MWR-A-H-GGFE 127
      | | | | : | | . | :| | | | :| | . | :
EMBOS_001      98 SSSIGDQV--RNIEF--LEKSTNLFKGYTVF-QSVVG-TGY-ALTIP-S 139
      : : | | | . | | . | :| |...:|.|. :| | | |
EMBOS_001      128 -GN-G--QTFR-I-LTRLEPYTF-Y-GNMLCRRTLGLLKYPA---PYS 166
      | | | | : | | . | :| | | | :| | . | :
EMBOS_001      140 -L-----TVVVVV-Q--TQNPPLHGTFEADFVFDFTTMVL-WV-- 175
      | | | | : | | | | | | | | :| | . | | |
EMBOS_001      167 RLCREQPHETVQHF-RQLKPSQMPVKGIFD-D----DL-A-I-LDMVLE 207
      | | | | : | | | | | | | | :| | . | | |
EMBOS_001      176 -L-----V-I-A--V--GM-----F-AML-CLISV--LV--IVSSFMWS 201
      | : : : | | . : : | | | : | . | : : :
EMBOS_001      208 PLSKADKQLFLSSYTVADGQHKRTRYKS-LDC--SIMELADOT--AY--A 250
      | : : : | | . : : | | | : | . | : : :
EMBOS_001      202 FLH-Y---II-GLFKLV----HIR-ISR-LNAFANPQSRQHELIEIVK- 238
      :| . | :| | :| | | . : : | . | :| :|
EMBOS_001      251 -VHDEDAIVNG--IVSEQNMH-NDVTQVL--CN--S-E-D--E--W 283
      | | | | :| | . | :| | | . : : | . | :| :|
EMBOS_001      239 LQ-ELVYRCARTAE DT LNLFLLT----Q----FGTCVV-AICLTMMTLTL 278
      | | | | :| | . | :| | | . : : | . | :| :|
EMBOS_001      284 LQNE--F--A-TM--SLRLF--SAKHQKRDAGT-LVNGF-VT--AISI 320
      | | | | :| | . | :| | | . : : | . | :| :|
EMBOS_001      279 ASND-Q-D--LLIXMI-LML--AY-I-LFNI--FV--YSMLGE-ELIA 312
      | : : | | | | . . | :| | | :| | | : : : | |
EMBOS_001      321 --NIEVEGDEPLL-K-YIAALEPAFDIAL-NVLKQVFVKY-VIRKPE-I- 362
      | : : | | | | . . | :| | | :| | | : : : | |
EMBOS_001      313 TSTSLAEAAV-GTQ--W--YEMS-I--PEQRNITLF-I-VR-R---SQKTA 348
      . : | | | | . : : | | | | : : | | | :|
EMBOS_001      363 -Q-ML-E--YKGHQIVNELFE-AFISDPE-R-LL-PLNTQERKIASEK-- 401
      | : | | | | . : : | | | | : : | | | :|
EMBOS_001      349 ALTTGKFAVNRSTFAATLQAAV-S--N-FTVLRQMMHSH 384
      . : | | | | . : : | | | | : : | | | :|
EMBOS_001      402 ---QGE----N-S-H-RVI-ADYISGMTDEFAA-R-L-HQM 428
      | : | | | | . : : | | | | : : | | | :|

#-----
#-----
```

Figure 17: Hit 2