QUESTION 1 :
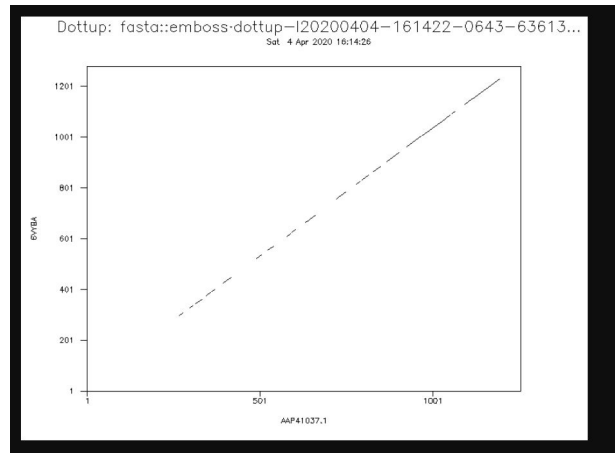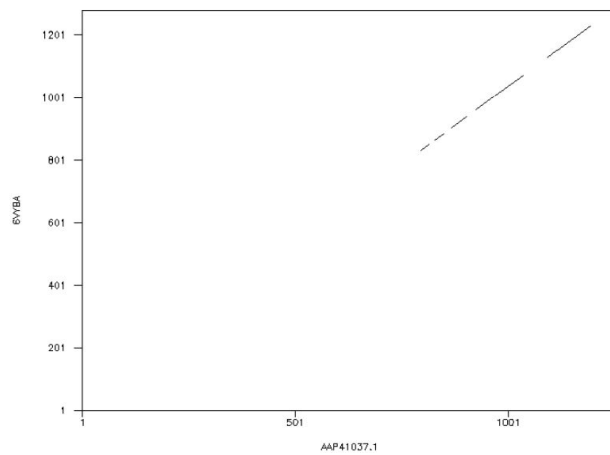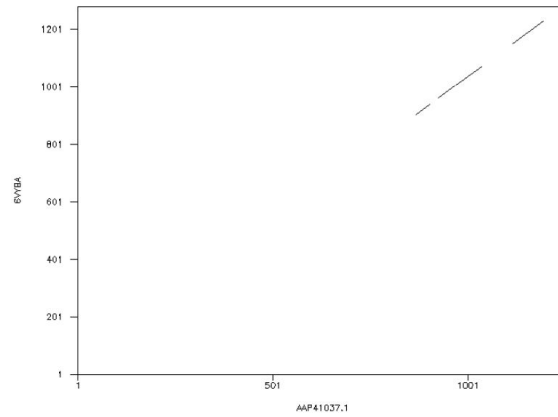
1. SARS COV 2 and SARS COV 1 : protein based Dottup graph : word size 10

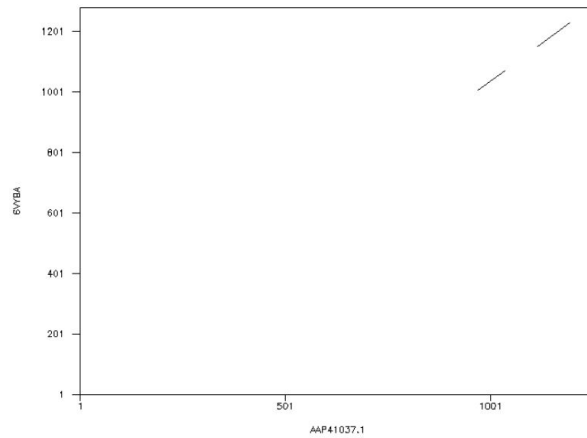2. SARS COV 2 and SARS COV 1 : protein based Dottup graph : word size 20

3. SARS COV 2 and SARS COV 1 : protein based Dottup graph : word size 30

4. SARS COV 2 and SARS COV 1 : protein based Dottup graph : word size 50

Now we shall plot the dottup graphs of SARS COV2 and MERS

1. SARS COV 2 and MERS : protein based Dottup graph : word size 50

The dottup plots show that SARS COV is more similar than MERS to SARS COV2

Now we shall plot Dottup graphs for SARS COV and SARS COV2 DNA based

1. SARS COV 2 and SARS COV : DNA based Dottup graph : word size 20
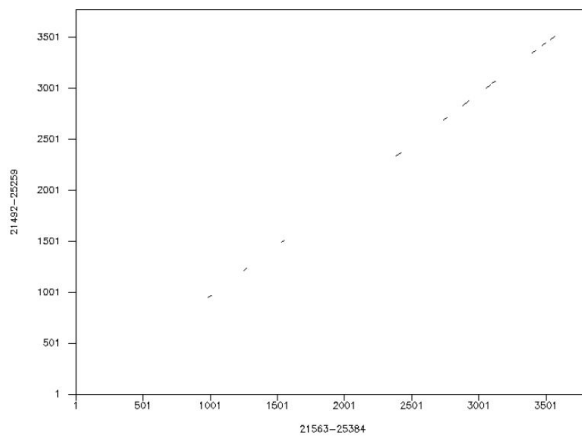
2. SARS COV 2 and SARS COV : DNA based Dottup graph : word size 30

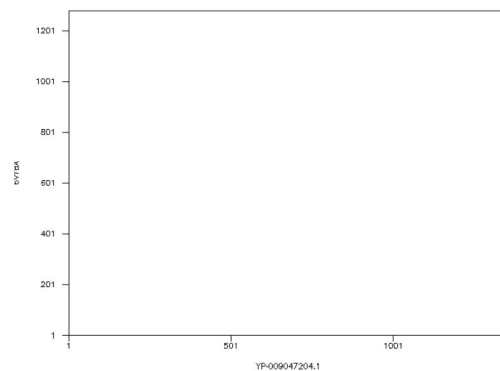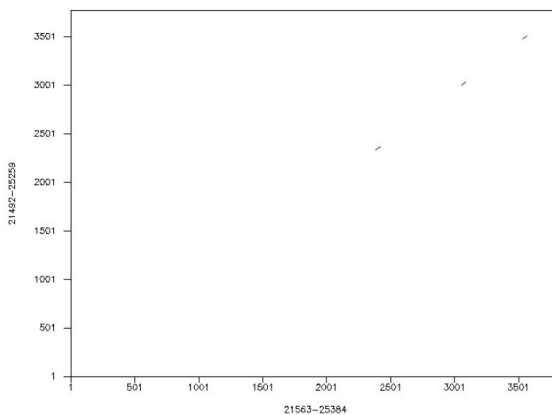Now we plot the dottup graph of SARS COV 2 and MERS COV : DNA based

1. SARS COV 2 and MERS COV : DNA based Dottup graph : word size 30
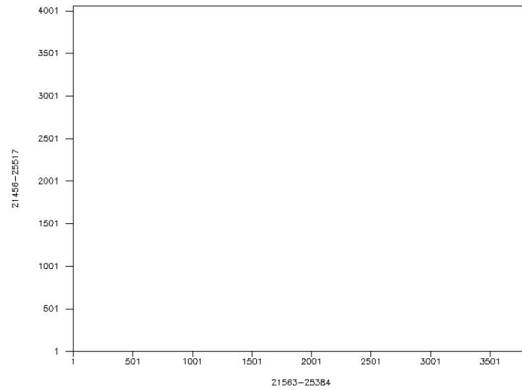
From these set of graphs we see that SARS COV 2 is more similar to SARS COV than MERS

QUESTION 2:

- **Following is the detail given for the SARS COV2 and SARS COV pairwise alignment at protein level**

```
########################################
# Program: needle
# Rundate: Sat  4 Apr 2020 17:47:22
# Commandline: needle
#    -auto
#    -stdout
#    -asequence emboss_needle-I20200404-174720-0808-87400211-p2m.aupfile
#    -bsequence emboss_needle-I20200404-174720-0808-87400211-p2m.bupfile
#    -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: AAP41037.1
# 2: 6VYBA
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 1308
# Identity:    918/1308 (70.2%)
# Similarity: 1057/1308 (80.8%)
# Gaps:        80/1308 ( 6.1%)
# Score: 4817.5
#
#
#=======================================

AAP41037.1      1 ------------------MFIFLLFLTLTSGSDLDRCTTFDDVQAPNYTQ     32
                        :.:.|:......|:.....||...: .|.||
6VYBA           1 MGILPSPGMPALLSLVSLLSVLLMGCVAETGTQCVNLTTRTQL-PPAYT-
48

AAP41037.1     33 HTSSMRGVYYPDEIFRSDTLYLTQDLFLPFYSNVTGFHTI-----NHT--     75
                    .|..||||||::|||..|:.|||||||:||||.||.|    |.|
6VYBA          49 -NSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKR
97

AAP41037.1     76 FGNPVIPFKDGIYFAATEKSNVVRGWVFGSTMNNKSQSVIIINNSTNVVI
125
                    |.|||:||.||:|||:|||||::|||:||:|:::|:||::|||:|||||
6VYBA          98 FDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVI
147

AAP41037.1    126 RACNFELCDNPFFAV----SKPMGTQTHTMIFDNAFNCTFEYISDAFSLD
171
                    :.|.|:.|::||..|    :.....::...::.:|.|||||:|...|.:|
6VYBA         148
KVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMD    197

AAP41037.1    172 VSEKSGNFKHLREFVFKNKDGFLYVYKGYQPIDVVRDLPSGFNTLKPIFK
221
                    :...|.||||:|||||||.||:...|..:|..||::||||.||:.|:|:..
```

```
6VYBA          198 LEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVD
247

AAP41037.1     222 LPLGINITNFRAIL----TAFSPAQDI--WGTSAAAYFVGYLKPTTFMLK   265
                   ||:|||||.|:.:|    :..:|....  |...||||:|||:|.||:||
6VYBA          248 LPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLK
297

AAP41037.1     266 YDENGTITDAVDCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPSGDVVR
315
                   |:|||||||||:..||:|.|||::|||.::||||||||||.|:...||
6VYBA          298 YNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVR
347

AAP41037.1     316 FPNITNLCPFGEVFNATKFPSVYAWERKKISNCVADYSVLYNSTFFSTFK
365
                   ||||||||||||||||:|.|||||.||:|||||||||||..|||||
6VYBA          348 FPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFK
397

AAP41037.1     366
CYGVSATKLNDLCFSNVYADSFVVKGDDVRQIAPGQTGVIADYNYKLPDD   415
                   |||||.|||||||:|||||||::||.||||||||||.||||||||||
6VYBA          398 CYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDD
447

AAP41037.1     416
FMGCVLAWNTRNIDATSTGNYNYKYRYLRHGKLRPFERDISNVPFSPDGK   465
                   |.|||:|||:.|:|:...|||||.||..|...|:|||||||..:.....
6VYBA          448 FTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGST
497

AAP41037.1     466 PCT-PPALNCYWPLNDYGFYTTTGIGYQPYRVVVLSFELLNAPATVCGPK
514
                   ||. ....|||:||..|||..|.|:||||||||||||||||:|||||||||
6VYBA          498 PCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPK
547

AAP41037.1     515
LSTDLIKNQCVNFNFNGLTGTGVLTPSSKRFQPFQQFGRDVSDFTDSVRD   564
                   .||:|:||:|||||||||||||||.|.|.:.||||||||::.|.||:|||
6VYBA          548 KSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRD
597
```

```
AAP41037.1      565 PKTSEILDISPCAFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQ
614
                    |:|.|||||:||:|||||||||||.|:||||||||||||:|..|||||
6VYBA           598 PQTLEILDITPCSFGGVSVITPGTNTSNEVAVLYQDVNCTEVPVAIHADQ
647


AAP41037.1      615 LTPAWRIYSTGNNVFQTQAGCLIGAEHVDTSYECDIPIGAGICASYHT--
662
                    |||.||:||||:|||||:|||||||||:.||||||||||||||||||.|
6VYBA           648 LTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQT
697


AAP41037.1      663 --VSLLRSTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNFSISITTEVMP
710
                    .|...|.:.:||:||||||::|:|||||:||||||:||:|||::|
6VYBA           698 NSPSGAGSVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILP
747


AAP41037.1      711
VSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQDRNT    760
                    |||.||||||.||||||||||:|||||||||||||||||:|||.|||:||
6VYBA           748 VSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNT
797


AAP41037.1      761 REVFAQVKQMYKTPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLFNKVTL
810
                    :|||||||:||||.:|.|.||||||||||||||.||:||||||||||||||
6VYBA           798 QEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTL
847


AAP41037.1      811 ADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMIAAYTAALV
860
                    |||||:||||:|||||.|||||||||||||||||||||||:|||.||:||:
6VYBA           848 ADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALL
897


AAP41037.1      861
SGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFN    910
                    :||.|:|:||||||||||||||||||||||||||||||||||.||||||
6VYBA           898 AGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFN
947
```

```
AAP41037.1      911 KAISQIQESLTTTSTALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLN
960
                    .||.:||:||::|::||||||||||||||||||||||||||||||
6VYBA           948 SAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLN
997

AAP41037.1      961 DILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATK
1010
                    |||||||··|||||||||||||||||||||||||||||||||||||
6VYBA           998 DILSRLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATK
1047

AAP41037.1     1011
MSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTYVPSQERNFTTA    1060
                    ||||||||||||||||||||||||:|||||||||||||:||:|||||
6VYBA          1048
MSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTA    1097

AAP41037.1     1061 PAICHEGKAYFPREGVFVFNGTSWFITQRNFFSPQIITTDNTFVSGNCDV
1110
                    |||||:|||:|||||||||·|||·||·||:||||:·|||||||||||||||
6VYBA          1098 PAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDV
1147

AAP41037.1     1111 VIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVN
1160
                    ||||:||||||||||||||||||||||||||||||||||||||||||||
6VYBA          1148 VIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVN
1197

AAP41037.1     1161 IQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLIAIV
1210
                    |||||||||||||||||||||||||||||||||||······|·|
6VYBA          1198 IQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKGSGRENLYF--------   1239

AAP41037.1     1211 MVTILLCCMTSCCSCLKGACSCGSCCKFDED-------DSEPVLKGVKLH
1253
                              :|....|...:...|     |.|.||....|.
6VYBA          1240 ----------------QGGGGSGYIPEAPRDGQAYVRKDGEWVLLSTFLG   1273

AAP41037.1     1254 YT------   1255
                         :.
6VYBA          1274 HHHHHHHH   1281
```

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

- ● **Following is the detail given for the SARS COV2 and SARS COV pairwise alignment at DNA level:**


```
########################################
# Program: needle
# Rundate: Sat  4 Apr 2020 17:49:03
# Commandline: needle
#    -auto
#    -stdout
#    -asequence emboss_needle-I20200404-174900-0067-26080698-p2m.aupfile
#    -bsequence emboss_needle-I20200404-174900-0067-26080698-p2m.bupfile
#    -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: 21563-25384
# 2: 21492-25259
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 3859
# Identity:    2833/3859 (73.4%)
```

```
# Similarity:  2833/3859 (73.4%)
# Gaps:         128/3859 ( 3.3%)
# Score: 15035.5
#
#
#=====================================

21563-25384     1 ATGTTTGTTTTTCTTGT--TTTATTGC---CACTAGT-------CTCTAG     38
                  |||||| ·|||||||·| |||·||·|  |||||||      ··||·|
21492-25259     1 ATGTTT-ATTTTCTTATTATTTCTTACTCTCACTAGTGGTAGTGACCTTG
49

21563-25384    39 TCAGTGTGTTAATCTTACAA-------CCAGAACTCAATTACCCCCTGCA
81
                   ·|·| |||·····|||···|     |·||··|··||||||·|·|··||
21492-25259    50
ACCG-GTGCACCACTTTTGATGATGTTCAAGCTCCTAATTACACTCAACA     98

21563-25384    82
TACACTAATTCTTTCACACGTGGTGTTTATTACCCTGACAAAGTTTTCAG     131
                   |||    |||·|··|···|·||·|||||·||·|||||··||·||||·||
21492-25259    99 TAC-----TTCATCTATGAGGGGGGTTTACTATCCTGATGAAATTTTTAG
143

21563-25384   132 ATC---CTCAGTTTTACATTCAACTCAGGACTTGTTCTTACCTTTCTTTT
178
                   |||  |·|··||| |||·|||||||||·||·||··|·||·||·|·|
21492-25259   144 ATCAGACACTCTTT---ATTTAACTCAGGATTTATTTCTTCCATTTTATT
190

21563-25384   179
CCAATGTTACTTGGTTCCATGCTATACATGTCTCTGGGACCAATGGTACT     228
                   |·||||||||··||||·|||·||||··||         ||   |||
21492-25259   191 CTAATGTTACAGGGTTTCATACTATTAAT-----------CA----TAC-     224

21563-25384   229
AAGAGGTTTGATAACCCTGTCCTACCATTTAATGATGGTGTTTATTTTGC     278
                   |||||··|||||||·||||·|||·|||||·|||||·||||||||||
21492-25259   225 -----GTTTGGCAACCCTGTCATACCTTTTAAGGATGGTATTTATTTTGC
269

21563-25384   279
TTCCACTGAGAAGTCTAACATAATAAGAGGCTGGATTTTTGGTACTACTT     328
```

```
              |·||||·|||||·||·||··|··|··|·||·|||·|||||||||·||||··
21492-25259    270
TGCCACAGAGAAATCAAATGTTGTCCGTGGTTGGGTTTTTGGTTCTACCA   319


21563-25384    329
TAGATTCGAAGACCCAGTCCCTACTTATTGTTAATAACGCTACTAATGTT   378
              |··|····|||·|·|||||··|·|·|||||·||||·||··||||||||||||
21492-25259    320
TGAACAACAAGTCACAGTCGGTGATTATTATTAACAATTCTACTAATGTT   369


21563-25384    379
GTTATTAAAGTCTGTGAATTTCAATTTTGTAATGATCCATTTTTGGGTGT   428
              |||||···||··|||·|·|||·||||·|||·|··|·||·||·||·|·|||
21492-25259    370
GTTATACGAGCATGTAACTTTGAATTGTGTGACAACCCTTTCTTTGCTGT   419


21563-25384    429
TTATTACCACAAAAACAACAAAAGTTGGATGGAAAGTGAGTTCAGAGTTT   478
              ||·|·|·|||        |||·|·|·||·|·|·|···|···|
21492-25259    420 TTCTAAACCCA--------------TGGGTACACAGACACATACTATGAT   455


21563-25384    479 ATTC--TAGTGCGAATAATTGCACTTTTGAATATGTCTCTCA-GCCTTTT  525
              ||||  ||·|||···|||||||||||·||·||··|·|||·| ||||||||
21492-25259    456
ATTCGATAATGCATTTAATTGCACTTTCGAGTACATATCTGATGCCTTTT   505


21563-25384    526
CTTATGGACCTTGAAGGAAAACAGGGTAATTTCAAAAATCTTAGGGAATT   575
              |···| ||··||··||··||·|||····|||||||·|||·|·|··|·||·||
21492-25259    506
CGCTT-GATGTTTCAGAAAAGTCAGGTAATTTTAAACACTTACGAGAGTT   554


21563-25384    576 TGTGTTTAAGAATATTGATGGTTATTTTAAAATATATTCTAAGCACAC-G  624
              ||||||||||·||||··|||||·|·|·|·|··|·||  |····||···|  ·
21492-25259    555
TGTGTTTAAAAATAAAGATGGGTTTCTCTATGTTTA-TAAGGGCTATCAA   603


21563-25384    625
CCTATTAATTTAGTGCGTGATCTCCCTCAGGGTTTTTCGGCTTTAGAACC   674
              |||||··||·||||·|||||||·|||···||||||····||||··||||
```

```
21492-25259    604
CCTATAGATGTAGTTCGTGATCTACCTTCTGGTTTTAACACTTTGAAACC    653


21563-25384    675
ATTGGTAGATTTGCCAATAGGTATTAACATCACTAGGTTTCAAACTTTAC    724
              ··|··|··|·|||||··|·||||||||||·||·|··|||··|··|·|
21492-25259    654
TATTTTTAAGTTGCCTCTTGGTATTAACATTACAAATTTTAGAGCCATTC    703


21563-25384    725 TTGCTTTACATAGAAGTTATTTGACTCCTGGTGATTCTTCTTCAGG---T
771
                  ||·|        ||···|||·|·||||        ||··||·  |
21492-25259    704 TTAC----------AGCCTTTT--CACCTG---------CTCAAGACATT    732


21563-25384    772
TGGACAGCTGGT--GCTGCAGCTTATTATGTGGGTTATCTTCAACCTAGG    819
              |||  ·||··||  |||||||||·||||·|||·||·|·|||·|··|·||·|··
21492-25259    733
TGG--GGCACGTCAGCTGCAGCCTATTTTGTTGGCTATTTAAAGCCAACT    780


21563-25384    820
ACTTTTCTATTAAAATATAATGAAAATGGAACCATTACAGATGCTGTAGA    869
              ||·|||·|··|·|·||·|||·||||||||||·||·||·||||||||||·||
21492-25259    781
ACATTTATGCTCAAGTATGATGAAAATGGTACAATCACAGATGCTGTTGA    830


21563-25384    870
CTGTGCACTTGACCCTCTCTCAGAAACAAAGTGTACGTTGAAATCCTTCA    919
              ·|||·|·|···|·||·||··|·|||···||·||··|·|·||···|||··
21492-25259    831
TTGTTCTCAAAATCCACTTGCTGAACTCAAATGCTCTGTTAAGAGCTTTG    880


21563-25384    920
CTGTAGAAAAAGGAATCTATCAAACTTCTAACTTTAGAGTCCAACCAACA    969
              ···|·||·||||||||·||·||·||·||||||·||·||·||····||·||
21492-25259    881
AGATTGACAAAGGAATTTACCAGACCTCTAATTTCAGGGTTGTTCCCTCA    930


21563-25384    970
GAATCTATTGTTAGATTTCCTAATATTACAAACTTGTGCCCTTTTGGTGA    1019
              |·|··|·||||·||||·||||||||||||||||||||·|||||||·||
21492-25259    931
GGAGATGTTGTGAGATTCCCTAATATTACAAACTTGTGTCCTTTTGGAGA    980
```

```
21563-25384    1020
AGTTTTTAACGCCACCAGATTTGCATCTGTTTATGCTTGGAACAGGAAGA  1069
           ·||||||||·||·||·|·|||··|·|||||·|||||·|||·|·||·||·|
21492-25259     981
GGTTTTTAATGCTACTAAATTCCCTTCTGTCTATGCATGGGAGAGAAAAA  1030


21563-25384    1070
GAATCAGCAACTGTGTTGCTGATTATTCTGTCCTATATAATTCCGCATCA  1119
           ·|||····||·|||||||||||·|||||·||·||·||·||··|||··
21492-25259    1031
AAATTTCTAATTGTGTTGCTGATTACTCTGTGCTCTACAACTCAACATTT  1080


21563-25384    1120
TTTTCCACTTTTAAGTGTTATGGAGTGTCTCCTACTAAATTAAATGATCT  1169
           ||||||·||·|||||||·|||||·||·|||·|·|||||·||·||||||||
21492-25259    1081
TTTTCAACCTTTAAGTGCTATGGCGTTTCTGCCACTAAGTTGAATGATCT  1130


21563-25384    1170
CTGCTTTACTAATGTCTATGCAGATTCATTTGTAATTAGAGGTGATGAAG  1219
           ·|||||··|·|||||||||||||||·||||||·|·|··||·|||||·|
21492-25259    1131
TTGCTTCTCCAATGTCTATGCAGATTCTTTTGTAGTCAAGGGAGATGATG  1180


21563-25384    1220
TCAGACAAATCGCTCCAGGGCAAACTGGAAAGATTGCTGATTATAATTAT  1269
           |·|||||||·||·|||||·|||||||····||||||||||||||||
21492-25259    1181
TAAGACAAATAGCGCCAGGACAAACTGGTGTTATTGCTGATTATAATTAT  1230


21563-25384    1270
AAATTACCAGATGATTTTACAGGCTGCGTTATAGCTTGGAATTCTAACAA  1319
           |||||·|||||||||||·|··||·||·||··|·|||||||·|||··||
21492-25259    1231
AAATTGCCAGATGATTTCATGGGTTGTGTCCTTGCTTGGAATACTAGGAA  1280


21563-25384    1320
TCTTGATTCTAAGGTTGGTGGTAATTATAATTACCTGTATAGATTGTTTA  1369
           ··|||||·|||·······|||||||||||||····||||·|···|||
21492-25259    1281
CATTGATGCTACTTCAACTGGTAATTATAATTATAAATATAGGTATCTTA  1330
```

```
21563-25384    1370
GG-AAGTCTAATCTCAAACCTTTTGAGAGAGATATTTCAACTGAAATCTA  1418
            |· |·|·| ||·||·|··||·||||||||||·||·||·|·||··· ||·
21492-25259    1331
GACATGGC-AAGCTTAGGCCCTTTGAGAGAGACATATCTAATGTGC-CTT  1378


21563-25384    1419
TCAGGCC-GGTAGCACACCTTGTAATGGTGTTGAAGGTTTTAATTGTTAC  1467
            ||···|| |·|·|||·||||||·|·······||·  |·||||||||||·
21492-25259    1379
TCTCCCCTGATGGCAAACCTTGCACCCCACCTGC---TCTTAATTGTTAT  1425


21563-25384    1468
TTTCCTTTACAATCATATGGTTTCCAACCCACTAATGGTGTTGGTTACCA  1517
            |··||·|||·|····|||||||··|··|||||||·|||··||||·|||||
21492-25259    1426
TGGCCATTAAATGATTATGGTTTTTACACCACTACTGGCATTGGCTACCA  1475


21563-25384    1518
ACCATACAGAGTAGTAGTACTTTCTTTTGAACTTCTACATGCACCAGCAA  1567
            |||·||||||||·|||||||||||||||||||·||·||||||||·||·|
21492-25259    1476
ACCTTACAGAGTTGTAGTACTTTCTTTTGAACTTTTAAATGCACCGGCCA  1525


21563-25384    1568
CTGTTTGTGGACCTAAAAAGTCTACTAATTTGGTTAAAAACAAATGTGTC  1617
            |·||||||||||·|||···||·|||·|·|·||||·|||·|·|||||||
21492-25259    1526
CGGTTTGTGGACCAAAATTATCCACTGACCTTATTAAGAACCAGTGTGTC  1575


21563-25384    1618
AATTTCAACTTCAATGGTTTAACAGGCACAGGTGTTCTTACTGAGTCTAA  1667
            |||||·||·||·|||||··|·||·||·||·|||||·|·|||···|||··
21492-25259    1576
AATTTTAATTTTAATGGACTCACTGGTACTGGTGTGTTAACTCCTTCTTC  1625


21563-25384    1668
CAAAAAGTTTCTGCCTTTCCAACAATTTGGCAGAGACATTGCTGACACTA  1717
            ·||·|··||||··||·||·|||||||||||·|·||··||·||||····|
21492-25259    1626
AAAGAGATTTCAACCATTTCAACAATTTGGCCGTGATGTTTCTGATTTCA  1675
```

```
21563-25384    1718
CTGATGCTGTCCGTGATCCACAGACACTTGAGATTCTTGACATTACACCA  1767
           |||||·|·||·||·|||||··|·|||··|||·||··|·||||||·||||·
21492-25259    1676
CTGATTCCGTTCGAGATCCTAAAACATCTGAAATATTAGACATTTCACCT  1725


21563-25384    1768
TGTTCTTTTGGTGGTGTCAGTGTTATAACACCAGGAACAAATACTTCTAA  1817
           ||··|||||||·|||||·|||||·||·|||||·|||||||||·||||···
21492-25259    1726
TGCGCTTTTGGGGGTGTAAGTGTAATTACACCTGGAACAAATGCTTCATC  1775


21563-25384    1818
CCAGGTTGCTGTTCTTTATCAGGATGTTAACTGCACAGAAGTCCCTGTTG  1867
           ··|·||||||||||·|||||·|||||||||||||·||·||··||···|
21492-25259    1776
TGAAGTTGCTGTTCTATATCAAGATGTTAACTGCACTGATGTTTCTACAG  1825


21563-25384    1868
CTATTCATGCAGATCAACTTACTCCTACTTGGCGTGTTTATTCTACAGGT  1917
           |·|||||||||||||||·||·||··|||||||··|·||||||||·||·
21492-25259    1826
CAATTCATGCAGATCAACTCACACCAGCTTGGCGCATATATTCTACTGGA  1875


21563-25384    1918
TCTAATGTTTTTCAAACACGTGCAGGCTGTTTAATAGGGGCTGAACATGT  1967
           ···|||||·||·||·||·|··||||||||·|·|||||·|||||·|||||
21492-25259    1876
AACAATGTATTCCAGACTCAAGCAGGCTGTCTTATAGGAGCTGAGCATGT  1925


21563-25384    1968
CAACAACTCATATGAGTGTGACATACCCATTGGTGCAGGTATATGCGCTA  2017
           |·|||··||·|||||||·|||||·||·|||||·||·||·||·||·|||||
21492-25259    1926
CGACACTTCTTATGAGTGCGACATTCCTATTGGAGCTGGCATTTGTGCTA  1975


21563-25384    2018
GTTATCAGACTCAGACTAATTCTCCTCGGCGGGCACGTAGTGTAGCTAGT  2067
           |||||·||·|  |||··|··||·|           ·|||||||  ·||||·
21492-25259    1976 GTTACCATA--CAGTTTCTTTAT----------TACGTAGT---ACTAGC  2010


21563-25384    2068
C---AATCCATCATTGCCTACACTATGTCACTTGGTGCAGAAAATTCAGT  2114
```

```
                 |   ||||·||··|·||·||·|||||||··|·|||||·||·|·||||·|
21492-25259    2011
CAAAAATCTATTGTGGCTTATACTATGTCTTTAGGTGCTGATAGTTCAAT   2060


21563-25384    2115
TGCTTACTCTAATAACTCTATTGCCATACCCACAAATTTTACTATTAGTG   2164
               ||||||||||||||·|·|||||·|||||·||·||·|||·|·|||||··
21492-25259    2061
TGCTTACTCTAATAACACCATTGCTATACCTACTAACTTTTCAATTAGCA   2110


21563-25384    2165
TTACCACAGAAATTCTACCAGTGTCTATGACCAAGACATCAGTAGATTGT   2214
               ||||·||||||·|··|·||·||·|||||·|·||·||·||·|||||||||
21492-25259    2111
TTACTACAGAAGTAATGCCTGTTTCTATGGCTAAAACCTCCGTAGATTGT   2160


21563-25384    2215
ACAATGTACATTTGTGGTGATTCAACTGAATGCAGCAATCTTTTGTTGCA   2264
               |··|||||||·||·||·|||||·|||||||····|||·|··|·|·|·||
21492-25259    2161
AATATGTACATCTGCGGAGATTCTACTGAATGTGCTAATTTGCTTCTCCA   2210


21563-25384    2265
ATATGGCAGTTTTTGTACACAATTAAACCGTGCTTTAACTGGAATAGCTG   2314
               |||||||·||·|||||·|||||·||||·|||||··|··|·||·||·||||
21492-25259    2211
ATATGGTAGCTTTTGCACACAACTAAATCGTGCACTCTCAGGTATTGCTG   2260


21563-25384    2315
TTGAACAAGACAAAAACACCCAAGAAGTTTTTGCACAAGTCAAACAAATT   2364
               ·||||||·||····|||||·|··|||||·||·||·|||||||||||||·
21492-25259    2261
CTGAACAGGATCGCAACACACGTGAAGTGTTCGCTCAAGTCAAACAAATG   2310


21563-25384    2365
TACAAAACACCACCAATTAAAGATTTTGGTGGTTTTAATTTTTCACAAAT   2414
               |||||||·|||·|··|·|||·|||||||||||||||||||||||||||
21492-25259    2311
TACAAAACCCCAACTTTGAAATATTTTGGTGGTTTTAATTTTTCACAAAT   2360


21563-25384    2415
ATTACCAGATCCATCAAAACCAAGCAAGAGGTCATTTATTGAAGATCTAC   2464
               ||||||·||·||···|||·||||··||||||||·|||||||·||··|·|
```

21492-25259   2361
ATTACCTGACCCTCTAAAGCCAACTAAGAGGTCTTTTATTGAGGACTTGC 2410

21563-25384   2465
TTTTCAACAAAGTGACACTTGCAGATGCTGGCTTCATCAAACAATATGGT 2514
          |·||·||·||·|||||||·||·|||||||||||||·||·||||||||·
21492-25259   2411
TCTTTAATAAGGTGACACTCGCTGATGCTGGCTTCATGAAGCAATATGGC 2460

21563-25384   2515
GATTGCCTTGGTGATATTGCTGCTAGAGACCTCATTTGTGCACAAAGTT 2564
          ||·|||||·|||||||||··|||||||||·|||||||||·||·|||||
21492-25259   2461
GAATGCCTAGGTGATATTAATGCTAGAGATCTCATTTGTGCGCAGAAGTT 2510

21563-25384   2565
TAACGGCCTTACTGTTTTGCCACCTTTGCTCACAGATGAAATGATTGCTC 2614
          ·||·||·|||||·||·|||||||||·|||||||·|||||·|||||||||·
21492-25259   2511
CAATGGACTTACAGTGTTGCCACCTCTGCTCACTGATGATATGATTGCTG 2560

21563-25384   2615
AATACACTTCTGCACT-GTTAGCGGGTACAATCACTTCTGGTTGGACCTT 2663
          ··|||||||·||||·||·|||||·|||||···||||·||||·|||||·||
21492-25259   2561
CCTACACTGCTGCTCTAGTTAGT-GGTACTGCCACTGCTGGATGGACATT 2609

21563-25384   2664
TGGTGCAGGTGCTGCATTACAAATACCATTTGCTATGCAAATGGCTTATA 2713
          |||||||·||·|||||··|·|·||||||||·||||||||||||||||·||||
21492-25259   2610
TGGTGCTGGCGCTGCTCTTCAAATACCTTTTGCTATGCAAATGGCATATA 2659

21563-25384   2714
GGTTTAATGGTATTGGAGTTACACAGAATGTTCTCTATGAGAACCAAAAA 2763
          |||||·|||||·|||||||||||||·||·||·|||||||||||||||||||||
21492-25259   2660
GGTTCAATGGCATTGGAGTTACCCAAAATGTTCTCTATGAGAACCAAAAA 2709

21563-25384   2764
TTGATTGCCAACCAATTTAATAGTGCTATTGGCAAAATTCAAGACTCACT 2813
          ···||·|||||||||||||·|··||·|||·|··|||||||||·|||||

21492-25259   2710
CAAATCGCCAACCAATTTAACAAGGCGATTAGTCAAATTCAAGAATCACT   2759

21563-25384   2814
TTCTTCCACAGCAAGTGCACTTGGAAAACTTCAAGATGTGGTCAACCAAA   2863
          |·|··|·|||·|||·||||·|·||·||·||·|||||·||·||·|||||·|
21492-25259   2760
TACAACAACATCAACTGCATTGGGCAAGCTGCAAGACGTTGTTAACCAGA   2809

21563-25384   2864
ATGCACAAGCTTTAAACACGCTTGTTAAACAACTTAGCTCCAATTTTGGT   2913
          |||||·|||||·|||||||·||||||||||||||||||·|||||||||
21492-25259   2810
ATGCTCAAGCATTAAACACACTTGTTAAACAACTTAGCTCTAATTTTGGT   2859

21563-25384   2914
GCAATTTCAAGTGTTTTAAATGATATCCTTTCACGTCTTGACAAAGTTGA   2963
          |||||||||||||··|||||||||||||·||·|||||·|||||·||
21492-25259   2860
GCAATTTCAAGTGTGCTAAATGATATCCTTTCGCGACTTGATAAAGTCGA   2909

21563-25384   2964
GGCTGAAGTGCAAATTGATAGGTTGATCACAGGCAGACTTCAAAGTTTGC   3013
          |||·||·||·|||||||·|||||·||·||||||||||||||··|·|
21492-25259   2910
GGCGGAGGTACAAATTGACAGGTTAATTACAGGCAGACTTCAAAGCCTTC   2959

21563-25384   3014
AGACATATGTGACTCAACAATTAATTAGAGCTGCAGAAATCAGAGCTTCT   3063
          |·||·|||||·||·||||||·||||·||·|||||·|||||||·||||||
21492-25259   2960
AAACCTATGTAACACAACAACTAATCAGGGCTGCTGAAATCAGGGCTTCT   3009

21563-25384   3064
GCTAATCTTGCTGCTACTAAAATGTCAGAGTGTGTACTTGGACAATCAAA   3113
          ||||||||||||||||||||||||||·|||||||·|||||||||||||
21492-25259   3010
GCTAATCTTGCTGCTACTAAAATGTCTGAGTGTGTTCTTGGACAATCAAA   3059

21563-25384   3114
AAGAGTTGATTTTTGTGGAAAGGGCTATCATCTTATGTCCTTCCCTCAGT   3163
          |||||||||·|||||||||||||||||·||·|||||||||||||·||··

21492-25259   3060
AAGAGTTGACTTTTGTGGAAAGGGCTACCACCTTATGTCCTTCCCACAAG 3109

21563-25384   3164
CAGCACCTCATGGTGTAGTCTTCTTGCATGTGACTTATGTCCCTGCACAA 3213
         |||||·||·|||||||||·||||||·|·|||||·||·|||||·||·|·|·||·
21492-25259   3110
CAGCCCCGCATGGTGTTGTCTTCCTACATGTCACGTATGTGCCATCCCAG 3159

21563-25384   3214
GAAAAGAACTTCACAACTGCTCCTGCCATTTGTCATGATGGAAAAGCACA 3263
         ||·|·||||||||·||·||·||·||·|||||||||||·||·||||||·|
21492-25259   3160
GAGAGGAACTTCACCACAGCGCCAGCAATTTGTCATGAAGGCAAAGCATA 3209

21563-25384   3264
CTTTCCTCGTGAAGGTGTCTTTGTTTCAAATGGCACACACTGGTTTGTAA 3313
         |||·||||||||||||||·|||||·|··||||||||····||||||·|·|
21492-25259   3210
CTTCCCTCGTGAAGGTGTTTTTGTGTTTAATGGCACTTCTTGGTTTATTA 3259

21563-25384   3314
CACAAAGGAATTTTTATGAACCACAAATCATTACTACAGACAACACATTT 3363
         |||||·|||||·||·|·|···|||||||·||||||||||||||·|||||
21492-25259   3260
CACAGAGGAACTTCTTTTCTCCACAAATAATTACTACAGACAATACATTT 3309

21563-25384   3364
GTGTCTGGTAACTGTGATGTTGTAATAGGAATTGTCAACAACACAGTTTA 3413
         ||·||·||·||·||||||||·||·||·||·||··|·||||||||||||||
21492-25259   3310
GTCTCAGGAAATTGTGATGTCGTTATTGGCATCATTAACAACACAGTTTA 3359

21563-25384   3414
TGATCCTTTGCAACCTGAATTAGACTCATTCAAGGAGGAGTTAGATAAAT 3463
         ||||||||·|||||||||··|·|||||||||||·||·|||·|·||·||·|
21492-25259   3360
TGATCCTCTGCAACCTGAGCTTGACTCATTCAAAGAAGAGCTGGACAAGT 3409

21563-25384   3464
ATTTTAAGAATCATACATCACCAGATGTTGATTTAGGTGACATCTCTGGC 3513
         |·||·||·||·||||||||||||||||||||·|·||·|||||·||·|||

21492-25259   3410
ACTTCAAAAATCATACATCACCAGATGTTGATCTTGGCGACATTTCAGGC   3459

21563-25384   3514
ATTAATGCTTCAGTTGTAAACATTCAAAAAGAAATTGACCGCCTCAATGA   3563
           |||||·|||||·||·||·||||||||||||||||||||||||
21492-25259   3460
ATTAACGCTTCTGTCGTCAACATTCAAAAAGAAATTGACCGCCTCAATGA   3509

21563-25384   3564
GGTTGCCAAGAATTTAAATGAATCTCTCATCGATCTCCAAGAACTTGGAA   3613
           |||·||·||·||||||||||||·||||·||·||·||||||·|·||||
21492-25259   3510
GGTCGCTAAAAATTTAAATGAATCACTCATTGACCTTCAAGAATTGGGAA   3559

21563-25384   3614
AGTATGAGCAGTATATAAAATGGCCATGGTACATTTGGCTAGGTTTTATA   3663
           |·|||||||·|||||·||||||||·|||||··|||||||·||·||·||·
21492-25259   3560
AATATGAGCAATATATTAAATGGCCTTGGTATGTTTGGCTCGGCTTCATT   3609

21563-25384   3664
GCTGGCTTGATTGCCATAGTAATGGTGACAATTATGCTTTGCTGTATGAC   3713
           |||||··|·||||||||·||·||||·||||··|||||||·||·|||||
21492-25259   3610
GCTGGACTAATTGCCATCGTCATGGTTACAATCTTGCTTTGTTGCATGAC   3659

21563-25384   3714
CAGTTGCTGTAGTTGTCTCAAGGGCTGTTGTTCTTGTGGATCCTGCTGCA   3763
           ·|||||·||·|||||·||||||||····||·|||||||·||·|||||||
21492-25259   3660
TAGTTGTTGCAGTTGCCTCAAGGGTGCATGCTCTTGTGGTTCTTGCTGCA   3709

21563-25384   3764
AATTTGATGAAGACGACTCTGAGCCAGTGCTCAAAGGAGTCAAATTACAT   3813
           |·|||||||·||·|||||||||||||·|||||·||·||||||||||||
21492-25259   3710
AGTTTGATGAGGATGACTCTGAGCCAGTTCTCAAGGGTGTCAAATTACAT   3759

21563-25384   3814 TACACATAA   3822
                |||||||||
21492-25259   3760 TACACATAA   3768

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