

Virus Report

21-01-2015 14:50:48

Secondary structure results

Virus name:
Coca-Cola

Sequence:

```
TGGTATCAGAGCAAGGTTTTCTGTAAGTAATTTATGGCTTTCATGGGGTAAAACCCCTTAGA
TAGGAGCCTGAGGGCTCTGCTATGTTCTATTTTGAGAAAATAACTGTGTAAGTAGTAATGC
ATGATAAATCGAATAAGTTCTTTACGGATGGTCCTAAATTTATGAAATCATTCTTTTACTAT
CTGAGTGAGACCCTTGTGTTTCCAAGTTCTACTGAGAAATCATGAGTAAATGGACGTTGA
CACACTGACTTAAGGAAAAGAAAGATGTGCCTAGGATGAAGTCGAGCACTGTTCTAGGCT
GAAGGGTGAAAAGAAAAATGAGTTGAGTTGGGAGTAAGACTCTGAGAGGCTTGAGGCAA
AAGAACCGGAACCTTACCTAGCAAAGTAAGACCCAAGGGAGGGCAGTAGCTGGAGAACC
CCGGTGTGAGTTTCTTTACCTAGCAGAAGATATACTCTAGTATCCTCAAAACACCGTGAAG
TATTTACAACCCTTTATGAACAAGTATTTTCTGGAGTTAGATGTCAGCCACCAAAGATCAT
GATCTATCACAAACAAACTTGAGTCATACAGAGAAGAACTCAGATGTAACCTGAATGACTAA
TGATCAGGGACCTCTATTAAGAGCTTAGTAAGAAGTAAGTGCATAAACTACCTCTGGAAC
TGGTTATAATCTTGAAATGAGCATATGTATCCCAAAGGTTACTCCTATCTGTCCGTGACCA
TAAGAAACCTTCTTTAAAGAAAAACTCTGGATTCTACTGGGATCTTGAATTTGAAAAGTA
CAAAAACCTCTCATTCCAAATCTGTTTCAGATCTTTCTTATTTAGACTTAGCATCTGCTGATA
AGGTCTCTAATAAAGACCTTGCATTTAATTTACATATAAACACTTATCGTAGTGATCTTGGT
TTTAAAGTCGCAATTCACCTCTTTATCAAAAAATCGTGAGCTGCTAATACAGAACAGGAAGC
TTTTGGAAGAACAAAAGCAGCAGCTGTCTGAAATTAACAACCTATCTAAGGTTGTGCGTCT
CCAACGAGCTGATTTAAAGGAGACCTTGAGAAGACAGGACGTCCTTGCGAAGGAGCTTC
AGGCTCTTCGTAAAGACTATCTGGAAGGCGCCCTCTTAGCAAGGAGGACGTGGAAGAA
TTGGTTGTCCGCATCTCTGAACAACCAAAAATTTATTGAAAAAACAGACGGAAGCTCTTAC
TGAGGAGCTTACAAAGGAGGTTGAGGCACTGAAGCTAATCATCCATAACTTCGAACAGAA
GTTACTGGGATGAGCCTGGCCGGAAGCAAGGCCACAGCAATCTATCAAGAAGCCTTGCA
ATCTACTGCTGCTGGTTGGGAAGACACCGGCATCGGATTTACTGATAAAGGAGTTATCAG
TACCACTACCCTATCAAGGCAGCTAAATTCGCTGTTATTTCTGGTGATCAAAACAAAATCAG
CAGATCACCACCATCTCGGAAGAACTAACCCAACCTTCACAACCGGGTTAAAAACCTTGAA
GGAAGGACAGGGGTATCTGCAAGCCCCCTGTATAAATCGGAGATTGAATCCATCAACGC
CAAGCTCAAAAGTATTCAAGATATACAAGGGAGTCACCCTCCAAAAGAGACCCCCCTCCGG
AGTTATCAAAGTTTTTGAGGATCCCTACAGCATCCTCCGCAGACTATAAATGGCGTCAAGA
CCTAGAGTTTCTGGGAGCACTACCAGAACCATGATAGCAGAACCAGGGGTTCCCTGGT
GGATGATCAGATCCGGGAATACCGGAGTGCGGCCCGAGTCGCATACGAAGCACAAAGG
ATTGCCAGACGCACTGGCAATATTCTTGGAAGAATTGTTGGACGACAACCAAGGGAGCAT
ACCCTTGCGATGGTCGTGGACCCCAACAGTGAGCTGGAACGCTCACTGGCTCACCGGGC
ACGAACAATACCTGCTGAAGTCCTTTACATGACCCAGCGCGGTGAGCCTACAAATCGGGT
CTATCGGAACAGAACTGAAGAAAGGATGCTAGTGACTCATGGTCAGCAAGATAGGACTTT
CATCTTGCCAGAGTCATATGAAGAGCTCAGGGAGGCTGGATTTGAGTACATCCATCTAGG
GGTACTCCAGGTGCGCATCCAAATTATGCACCGAACATATGATGGTACAATGGCATTAGT
GGTCTTCCGTGACACCAGATGGACACAGGAGAACCACCAAGACAGATCCATTATAGCAAC
CATGGAGGCTGATCTCTCCAAGGGCATCAGCTAATTTACGTCATCCCAGATATAATGAT
GACTATTCGGGATTTCTACCAGCACATTGAGATAAGCATCCTTACCAAAGGATATGAAGG
CTTTCAAGGAGAAGCAAATCTCCTTATTACAAGAAGCTGCCGGTGACAGATTAAGCAACGT
TCCTAACGTTGGCTTCCAATACAACATCCAGAACGTCGTGGAATTCTTAAAATCTAGGGGA
GTAAAAGCCCTAAATGCGACGAAATTAAGCACCAGGAGGTTCCAAGGAGGAGAATGGAA
CATCCGACCATCAGAAGTGGTGGTCCCTATGCAGCCAACAACCTATGATAGTACGAGTTAA
TTATGACTCGTCACGAAGCATCAGATTCGGGGATTATGAAGCTAGCACATCATCATCAGC
CCCGAGATATGAGCAAGACGGTGATGAAGATGAAGCACTTGGTGATATCCATCAAGTAA
TATGATCACCATCATAGAAGATGATGCAGAAGATGACTACCCACGGTTATCAGCTTTAGA
GCGAATAATCGCTCCAGAAAGCATGGTGGGAGAGGAGGACACAATTGCTGAATTTTTAAG
CAATCTGTCTTAGATTCTCCACTGATGAGGAATTCTACGACGCCGACAACCTCATTATTC
GAAGAAGAAGAATATGATGGGGATTCCGAAGTTAGCACCCCAAGGAGCAAATACAATATA
TTTGCGCTAGAAGATGAGTACCCAAAACCTTCAACAGCTGGAAAGCCTGGTACTCTCCACA
ACAGAGTCTGCTATCAGTCGCTTTAGACCAGCAGACACAGATATGACTGGCGTAGGCCCT
GGCTACGCACCAGCAACTGGAACAGCTGGCTATACTGGAGCCAGTTCATCTGATTTCCCT
TACCCTAGAAGACCAAGAAAGTGGGACAACAACCTCGGAGTGGTTCAACCTACCCACCGC
TAACGCGAGGCAAGCATCAATATTCGTTATGCCTCAGGATTTTGACACAAAGGTCTTTGAA
```

AGATGGGAAAGTTCTGTTCTCCTTCACATGTCTGACAAGGTTTTTGGATGATCCACAAGACA
AGTTAACTTATGTGGAAAACCTACTAGGAGAATCAGAAAAGAAGATGTTTATCACTTGGAG
GATGATGTTACAGCTGAGTACGAAGAAATGAAGAACAACGCTCTCGGCTCAAATGGAAC
TCAAATATCCTGAACCAGATCAGGATGATATTCTTTCTGGAAAATCCCCAAGTTGGAACC
ACCAACACTCAAGACGCAGCATAACAAGACACTCAAACAAGTCTGCACAGAGATGTCT
GGACCTGCGATCTACAGGTACCTAAATGATTACTTCCATTTAGCGGCAAAATCTGGAAGA
GCGTGGGCATCTGATGAGCTGTCCAAGGAATTCTTTACAAAAGTCCAAGGGGATTAGG
GGACAGAGTTGAAAAGAAATTCAAAGAAAAGTACCCCAACAACACTATTGGAGTAGCCCC
CAGAATCACCTTCACAAGAAATTATATAAAGGAAATATGCCAAGAGGCTGTATTCCAAAGC
CAACTGAAAAGGCTAGATTTTTTGCAAAGGGACACCCGTCACGGCTTATATGGTAAAGAG
AAGGCATATGGGAGGAAATACGGAGTCAGGAAAAGCACTTCGTACAAAGGAAAGCCTCA
CAAGTCACATGTGAGGATAGATAAAAAGAAGCACCTCCTGATGAAACGAAAAGACTGCAA
GTGTTTTGCTTGCGGAGATATCGGTCACTTCGCATCAGAGTGTCCGAATCCCAAGAAGCT
GATGCACCGAGTTCAAATTCTACAATCCTTAGAGCTTGATGACGGAATCGACGTAATCTC
CGTAGGATTTGATGAGTCAGATGTCTCAGACATCTATTCGGTATCTGAAGGCGAGGACAG
CTACCAGTTCAATAATGAGGATTTTGACGTTATAGGACATGACGTGTTTCATGTTCAACCATT
GAAGAACAAGAAAGTGCCTGGTAGAAACCACGTCAGCCTGGAGAAGTGCAATGAAAGTT
ACTCCAGAAGAAAAGAATTGCTTACACACATGGAGCTTTGAGGAAAAGACAACAGACCAC
TGCAGAGCCTGCAAAAATCTGGCCCTACGAGGGAGCAGAGCTGACTGCACTCAGTGTA
GATAATCATTGCTCTTTATGCAAACCATACTATTTCCAGGATGGTTCACCTATCCCAGCA
CAAAGTTCAACCCCATCAGGATACTCCTACGATGATTGGATGGGATCAGCAAATAGGTGG
AAGGCACACTATGAGTTCTCTCAGGCAAGAAGGAAGAGCCTGAAGGCAGACCTTGAAAG
AGCGGAAGAAGAACTAAAATTTTATAAGCAAAAAGGAAAAGGAAAAGGCCAAATTAAGGA
TCAAATTCAGAAAGCAGTACAAGCAAAAGTGGATGATCTGGAAAAAGAAAAAGAGCTCAA
TAACATCCTAAGAATAGAGGCTGAGACAGAGCTAAAAGCATTAAAGGAAAGCTTCAAAGA
AAAGGAAGAAGCCCTGAAGGAAGAAATCACAGCTCTAGAAGAAGAAGTGAGGATACACA
AAGAGGAGGCTGAAGAAGTTCAGGAAGAAAATCAAAAAGTCAAAGAGAAGATCATAGCCT
TCGAAAAAGACGTAACACAAGGACCAGAAGAAGTGATCGAATTGGTCAATAACGTGGAGG
AGCACCTGGTACTAACAGGACAACAGAAGAACAATCTCCTCAACATCAAGATAACTCTGG
AAGTTAAGGAAAAAAGGATTACCATGAACGCGATACTAGACACTGGAGCCGCAATCTGTG
TCTGTGATGGGCAAATGGTAAACGAATATTTTCAAGGAGACCATCAATGATGAATGCGTTCAT
AAAGGGAGTTAATGGGATTACCAATGTCAAAGAAATCTTAGAGGAAGGAAAAATCTGGAT
TGGTAATCAATGGTTCAGAATCCCAAGGACATACATCATGCCCCAGCTATCAGAAGGCCT
TCATTTTCATCATCGGAATGAACCTTTATCAGGGCAATGGAAGGAGGGATTGCAATTGAGCA
AGGAACGGTAACCTTCTACAAAATGGTTACACAAGCACAGGCACCCCCCTATGGTACACGA
TATTTCTTACCTTGAAGAATTAGAATTAGAAGTGCCGATTTACTATGATATCTGTGCAACTA
ACCCCTCAGGAGGAGAAATCAATAGTGACCTAATATCCCCCTCGGAGATCCGGAAACTAA
AGGACTTAGGCTACATTGGGGAAGAGCCCTTGAAACATTGGGCTAAGAATCAAGTCAAGT
GCCGAATTGAAATCAAGAACCCTGACTTAATCATTGAAGATAGGCCCTTGAAAGCACGTCA
CCCCAGCAATGAAAGAATCCATGAAGAAACATGTGGACAAGCTGCTGGAAGTTAAGGTAA
TCAGGCCATCCACAAGCAAAACACCGGACGACTGCGATAATTGTTCAATCCGGTACAGAAA
TTGACCCCTCACTGGAAAAGAGAAAAGAGGGGAAAGAGAGGCTTGATTCAACTACAAAC
GCCTCAACGACAATACCGAAAAAGACCAGTATTCACCTACCTGGCATCAATACAATTATCAG
CAGGATCGGCAAGTCAAAAATTTACAGCAAATTTGACTTAAAATCCGGATTCCACCAGGTA
GCCATGGATCCAGAAAGCATCCCATGGACGGCCTTTTGGGCCATAGATGGACTTTATGAA
TGGCTAGTTATGCCATTTGGTCTGAAGAATGCGCCCGCTATATTTTCAGAGAAAGATGGAC
AACTGCTTCCGAGGAACGGAGGAATTCATAGCGGTATATATTGATGATATTTTAATATTCT
CTGACAATATCTCTGATCACAGGAAACATCTGTCAAATTCCTGGAGATCTGCAAGGCGA
ATGGGCTGGTATTAAGCCCAACAAAAATGAAGATAGGCGCAAAGGAAATTGATTTCTAG
GAGCAACTATTGGAAACTCCAAGATCAAGCTTCAACCTCATATAATCAAGAAGATCATCGA
GACAAAGGACGAGGAGCTAAAGGAAACAAAGGGGCTCAGAAAATGGTTGGGAGTCCTTA
ACTATGCACGGGCATACATTCCAAATTTAGGAAAAACATTAGGCCCGCTCTACTCCAAGA
CGTCAATTAATGGAGAGAAGAAGATGAACAGCCAAGATTGGAAGGTTGTTCAACTGATCA
AAAATCAGGTACAAAATTTACCTGACCTTGACATACCCCTGCGAGAGGCAACTATGGTCTT
AGAGACTGACGGGTGTATGGAAGGATGGGGAGGAGTATGCAAATGGAAGCTCCATCCCT
CTGACACAAGACTGGCAGAAAAGGTCTGTGCGTACGCAAGTGGAAGGTATCACCCCATC
AAGAGCACAATTGATGCAGAGGTACACGCAAGTATCAACAGCTTAGAAAAATTCAAATTT
ATTATCTGGATAAAAAGGAGTTGATCATAAGAACAGACAGCCAGGCTATTGTGGCTTTTTA
TAAGAAGCAAGCTGATCATAAGCCCTCAAGAACAAGGTGGCTTATGCTAATTGACTACATT
ACTGGGCTCGGAATCAACGTCAAATTCGAGCATATTGACGGGAAAGAAAATGTTCTAGCT
GATACTCTATCAAGGCTGGTCCAAGTGCTGATACCAAGGTTTCATCATCCAGCAGAAACC
CAGCTAGTCGAAGCCGTGATGGAAGTTATAAGCAATCCAAAGAAAGAAGCCTTGGACAAG
GTAAACCATTTTATCTTCTTAACCCAACAGTGGATTGCAGAACGCAAGGAGGAGCACACG
GTGAACACGCTACTCCAGTTGGAAGAACCACAGCTGCATTGTGGTTGTAGAAACTATGAA

15

Ending aminoacid position:
145

Domain description:
Protein of unknown function (DUF1319)

Identifier:
DUF1319

Protein 2

Number of domains:
0

Aminoacidsequence:
MSLAGSKATAIYQEALQSTAAGWEDTGIGFTDKGVISTTTLSRQLNSLLFLVIKQNQQITTISEE
LTQLHNRVKNLEGRTGVSASPLYKSEIESINAKLKSIQDIQGSHPKETPSGVIKVFEDPYSILR
RL

Starting nucleotide position:
1282

Ending nucleotide position:
1681

Protein 3

Number of domains:
4

Aminoacidsequence:
MASRPRVSGSTTRTMIAEPGVPLVDDQIREYRSAARVAYEAQRIARRTGNILGRIVGRQPREH
TLAMVVDPNSELESLAHRARTIPAEVLYMTQRGEPTNRVYRNRTTEERMLVTHGQQDRTFIL
PESYEELREAGFEYIHLGVLQVRIQIMHRTYDGTMALVVFRDTRWTQENHQDRSIIATMEADL
SQGHQLIYVIPDIMMTIRDFYQHIQISILTKGYEGFQGEANLLITRSCRCRLSNVPNVGFQYNIQ
NVVEFLKSRGVKALNATKLSTRRFQGGGEWNIRPSEVVVPMQPTTMIVRVNYDSSRSIRFGDY
EASTSSSAPRYEQDGEDEALGDIHQVNMITIIEDDAEDDYPRLSALERIIAPESMVGEEDTIAE
FLSNLSLDSSTDEEFYDADNSLFEEEEYDGDSEVSTPRSKYNIFALEDEYPKLQQLESVLSTT
ESAISRFRPADTDMTGVGPGYAPATGTAGYTGASSSDFPYPRRPRKWDNNSEWFNLPTANA
RQASIFVMPQDFDTKVFERWESSVLLHMSDKVFDDPQDKLTYVENLLGESEKKMFITWRMM
FTAEEEMKNNALGSNGTQNILNQIRMIFFLENPQVGTTNTQDAAYKTLKQLVCTEMSGPAIY
RYLNDYFHAAKSGRAWASDELSKEFFTKLPRGLGDRVEKKFKEKYPNNTIGVAPRITFTRNY
IKEICQEA VFQSQLKRLDFCKGTPVHGLYGKEKAYGRKYGVRKSTSYKGKPKHKS HVRIIDKKK
HLLMKRKDCKCFACGDIGHFASECPNPKLMHRVQILQSLELDDGIDVISVGFDESVDIYSV
SEGEDSYQFNNEFDVIGHDVFMFTIEEQRNCLVETTSAWRSAMKVTPPEEKNCLHTWSFEED
TTDHCRACKNLALRGSRADCTQCKIIICSLCKPYYFQDGSPIPAQSSTPSGYSYDDWMGSAN
RWKAHYEFSQARRKSLKADLERAEELKFYKQKEKEKAKLKDQIPEAVQAKLDDLEKEKELN
NILRIEAETELKALKESFKEKEEALKEEITALEEEVRIHKEEAEELQEENQKLKEKIIAFEKDVTQ
GPREEVIELVNNVEEHLVLTGQQKNNLLNIKITLEVKEKRITMNAILDGTGAICVCDGQMVNEYFR
RPSMMAFIKGVNGITNVKEILEEGKIWIGNQWFRIPRTYIMPQLSEGLHFIIGMNFIRAMEGGI
RIEQQGTVTFYKMTQAQAPPMVHDISYLEELELELPIYYDICATNPSGGEINSDLPSEIRKLKD
LGYIGEEPLKHWAKNQVKCRIEIKNPDLIEDRPLKHVTPAMKESMKKHVDKLELKVIRPSTSK
HRTTAIVQSGTEIDPLTGKEKRGKERLVFNKYKRLNDNTEKDQYSLPGINTIISRIGKSKIYSKFD
LKSGFHQVAMDPEIPWTAFAWIDGLYEWLVMPFGLKNAPAIQQRKMDNCFRGTEEFIAVYID
DILIFSDNISDHRKHLSKFLEICKANGLVLSPTKMKIGAKEIDFLGATIGNSKIKLQPHIIKKIIETKD
EELKETKGLRKWLGVNLNARAYIPNLGKTLGPLYSKTSINGEKKMNSQDWKV VQLIKNQVQNL
PDLIPPAEATMVLETDGCMEGWGGVCKWKLHPSDTRLAEKVCAYASGRYHPIKSTIDAEVH
AVINSLEKFKIYYLDKKELIIRTD SQAIVAFYKKQADHKPSRTRWLMLIDYITGLGINVKFEHIDGK

ENVLADTL SRLVQVLITKVHHPAETQLVEAVMEVISNPKEALDKVNHFIFLTQQWIAERKEEH
TVNTLLQLEEPQLHCGCRNYETGERRNAILLSHTSANPNRWFKCAENKCHIWIWKDILDQ
YAEDYATYTRIGLEALNLEDWFEEPEPDPPNPVDRQRIEDILLNVSND

Starting nucleotide position:
1681

Ending nucleotide position:
7378



Domain 1

Starting aminoacid position:
125

Ending aminoacid position:
246

Domain description:
Viral movement protein (MP)

Identifier:
MP

Domain 2

Starting aminoacid position:
763

Ending aminoacid position:
780

Domain description:
Zinc knuckle

Identifier:
zf-CCHC

Domain 3

Starting aminoacid position:
1090

Ending aminoacid position:
1201

Domain description:
Retroviral aspartyl protease

Identifier:
RVP

Domain 4

Starting aminoacid position:
1349

Ending aminoacid position:
1509

Domain description:

Reverse transcriptase (RNA-dependent DNA polymerase)

Identifier:

RVT_1