

Norovirus GI, complete genome

NCBI Reference Sequence: NC_001959.2

[FASTA Graphics](#)

[Go to:](#)

LOCUS NC_001959 7654 bp ss-RNA linear VRL 13-AUG-2018

DEFINITION Norovirus GI, complete genome.

ACCESSION NC_001959

VERSION NC_001959.2

DBLINK BioProject: [PRJNA485481](#)

KEYWORDS RefSeq.

SOURCE Norovirus GI

ORGANISM [Norovirus GI](#)
Viruses; Riboviria; Orthornavirae; Pisuviricota; Pisoniviricetes;
Picornavirales; Caliciviridae; Norovirus.

REFERENCE 1 (bases 1 to 7654)

AUTHORS Fernandez-Vega,V., Sosnovtsev,S.V., Belliot,G., King,A.D.,
Mitra,T., Gorbalenya,A. and Green,K.Y.

TITLE Norwalk virus N-terminal nonstructural protein is associated with
disassembly of the Golgi complex in transfected cells

JOURNAL J. Virol. 78 (9), 4827-4837 (2004)

PUBMED [15078964](#)

REFERENCE 2 (bases 1 to 7654)

AUTHORS Liu,B.L., Viljoen,G.J., Clarke,I.N. and Lambden,P.R.

TITLE Identification of further proteolytic cleavage sites in the
Southampton calicivirus polyprotein by expression of the viral
protease in E. coli

JOURNAL J. Gen. Virol. 80 (Pt 2), 291-296 (1999)

PUBMED [10073687](#)

REFERENCE 3 (bases 1 to 7654)

AUTHORS Liu,B., Clarke,I.N. and Lambden,P.R.

TITLE Polyprotein processing in Southampton virus: identification of
3C-like protease cleavage sites by in vitro mutagenesis

JOURNAL J. Virol. 70 (4), 2605-2610 (1996)

PUBMED [8642693](#)

REFERENCE 4 (bases 1 to 12)

AUTHORS Hardy,M.E. and Estes,M.K.

TITLE Completion of the Norwalk virus genome sequence

JOURNAL Virus Genes 12 (3), 287-290 (1996)

PUBMED [8883366](#)

REFERENCE 5 (bases 13 to 7654)

AUTHORS Jiang,X., Wang,M., Wang,K. and Estes,M.K.

TITLE Sequence and genomic organization of Norwalk virus

JOURNAL Virology 195 (1), 51-61 (1993)

PUBMED [8391187](#)

REFERENCE 6 (bases 1 to 7654)

CONSRTM NCBI Genome Project

TITLE Direct Submission

JOURNAL Submitted (01-AUG-2000) National Center for Biotechnology Information, NIH, Bethesda, MD 20894, USA

REFERENCE 7 (bases 1 to 12)

AUTHORS Hardy,M.E.

TITLE Direct Submission

JOURNAL Submitted (07-NOV-1995) Michele E. Hardy, Molecular Virology, Baylor College of Medicine, Houston, TX 77006, USA

REFERENCE 8 (bases 13 to 7654)

AUTHORS Estes,M.K.

TITLE Direct Submission

JOURNAL Submitted (09-AUG-1993) M.K. Estes, Molecular Virology, Baylor College of Medicine, Houston, TX 77006, USA

COMMENT REVIEWED [REFSEQ](#): This record has been curated by NCBI staff. The reference sequence was derived from [M87661](#).
On Jun 1, 2006 this sequence version replaced [NC_001959.1](#).
COMPLETENESS: full length.

FEATURES Location/Qualifiers

source	1..7654
	/organism="Norovirus GI"
	/mol_type="genomic RNA"
	/db_xref="taxon: 122928 "
5'UTR	1..4
gene	5..5374
	/gene="ORF1"
	/locus_tag="NoVGI_gp1"
	/old_locus_tag="NVgp1"
	/db_xref="GeneID: 1491970 "
CDS	5..5374
	/gene="ORF1"
	/locus_tag="NoVGI_gp1"
	/old_locus_tag="NVgp1"
	/note="The mature peptides were added by the NCBI staff following publications Liu et al (1996, 1999), that (tentatively) determined processing map by site-directed mutagenesis or by direct sequencing of the cleavage products for closely related Southampton virus (a Norwalk-like virus); ORF1; sequence homologies to 2C helicase, 3C protease, and 3D RNA-dependent RNA polymerase of picornavirus"
	/codon_start=1
	/product="nonstructural polyprotein"

/protein_id="[NP_056820.1](#)"
/db_xref="GeneID:[1491970](#)"
/translation="MMASKDVVPTAASSENANNSSIKSRLLARLKSGGATSPPN
IKITNQDMALGLIGQVPAPKATSVDVPKQQRDRPPRTVAEVQQLRWTERPDQNVKT
WDELDHTTKQQILDEHAEWFDAGGLGPSTLPTSHERYTHENDEGHQVKWSAREGV
DLGISGLTTVSGPEWNMCPLPPVDQRSTTPATEPTIGDMIEFYEGHIYHYAIYIGQKTVG
VHSPQAAFSITRITIQPISAWWRVCYVPQPKQRLTYDQLKELENEPWPYAAVTNNCFE
FCCQVMCLEDTWLQRKLISGRFYHPTQDWSRDTPEFQQDSKLEMVRDAVLAAINGLV
SRPFKDLLGKLGKPLNVLNLLSNCDWTFMGVVEMVLLLELFGIFWNPPDVSNFIASLL
PDFHLQGPEDLARDLVPIVLGGIGLAIGFTRDKVSKMMKNAVDGLRAATQLGQYGLEI
FSLKKYFFGGDQTEKTLKDIESAVIDMEVLSSTSVTQLVRDKQSARAYMAILDNEEE
KARKLSVRNADPHVVSSTNALISRISMARAAALAKAQAEMTSRMRPVVIMCGPPGIGK
TKAAEHLAKRLANEIRPGGKVGLVPREAVDHWGYPHGEVMLWDDYGMTKIQEDCNKL
QAIADSAPLTLNCDRIENKGMQFVSDAIVITTNAPGAPVDFVNLGPVCRRVDFLVYC
TAPEVEHTRKVS PGDTTALKDCFKPDFSHLKMELAPQGGFDNQGNTPF GKGVMPKPTTI
NRLLIQAVALTMERQDEFQLQGPTYDFDTRVAAFTRMARANGLGLISMASLGKKLRS
VTTIEGLKNALSGYKISKCSIQWQSRVYIESDGASVQIKEDKQALTPLQQTINTASL
AITRLKAARAVAYASCFQSAITITILQMAGSALVINRAVKRMFGTRTAAMALEGPGKEH
NCRVHKAKEAGKGPIGHDDMVERFGLCETEEEESEDQIQMVPSDAVPEGKNKGKTKKG
RGRKNYNNAFSRRGLSDEEYEEYKKIREEKNGNYSIQEYLEDQRQYEEELADEVQAGGD
GGIGETEMEIRHRVIFYKSKSKKHQQEQRRQLGLVTGSDIRKRKPIDWTPPKNEWADDD
REVDYNEKINFEAPPTLWSRVTKFGSGWGFVVSPTVFITTHVVP TG VKEFFGEPLSS
IAIHQAGEFTQFRFSKMRPDLTGMVLEEGCPEGTVC SVLIKRDSGELLEPLAVRMGAI
ASMRIQGRLVHGQSGMLLTGANAKGMDLGTIPGDCGAPYVHKRGNDDWVCGVHAAATK
SGNTTVCAVQAGEGETALEGGDKGHYAGHEIVRYGSGPALSTKTKFWRSSPEPLPPGV
YEPAYLGKDP RVQNGPSLQQVLRDQLKPFADPRGRMPEPGLLEAAVETVTSMLQTM
DTPSPWSYADACQSLDKTTSSGYPHHKRKNDDWNGTTFVGELGEQAAHANMNYENAKH
MKPIYTAALKDELVKPEKIYQKVKRLLWGADLGTVVRAARAFGPFCDAIKSHVIKLP
IKVGMNTIEDGPLIYAEHAKYKNHFDADYTAWDSTQNRQIMTESFSIMSRLTASPELA
EVVAQDLLAPSEMDVGDYVIRVKEGLPSGFPCTSQVNSINHWIITLCAISEATGLSPD
VVQSMSYFSFYGDDEIVSTDIDFDPARLTQILKEYGLKPTRPKTEGPIQVRKNVDGL
VFLRRTISRDAAGFQGRLDRASIERQIFWTRGPNHSDPSETLVPHTQRKIQLISLLGE
ASLHGEKFYRKISSKVIHEIKTGGLMYVPGWQAMFRWRFHDLGLWTGDRDLLPEFV
NDDGV"

[mat_peptide](#)

5..1198
/gene="ORF1"
/locus_tag="NoVGI_gp1"
/old_locus_tag="NVgp1"
/product="p48"
/note="N-terminal leader p48; 2A2-2B-like membrane
protein; distant homolog of H-rev107-like parechovirus
2a2, a putative regulator of cell proliferation; the
c-terminal membrane-binding portion contributes to the
Golgi disassembly and, therefore, functionally similar to

the picornavirus 2B protein"

mat_peptide 1199..2287
 /protein_id="[NP_786945.1](#)"
 /gene="ORF1"
 /locus_tag="NoVGI_gp1"
 /old_locus_tag="NVgp1"
 /product="NTPase"
 /note="2C-L; probable ortholog of the 2C protein of picornaviruses; the calicivirus NTPase was found in membranous replication complexes"
 /protein_id="[NP_786946.1](#)"

mat_peptide 2288..2890
 /gene="ORF1"
 /locus_tag="NoVGI_gp1"
 /old_locus_tag="NVgp1"
 /product="p22"
 /note="3A-L; located in the polyprotein between NTPase and VPg; second most variable region of the calicivirus polyprotein; the FCV ortholog was detected in membranous replication complexes"
 /protein_id="[NP_786947.1](#)"

mat_peptide 2891..3304
 /gene="ORF1"
 /locus_tag="NoVGI_gp1"
 /old_locus_tag="NVgp1"
 /product="VPg"
 /note="For Southampton calicivirus, both N-terminal and C-terminal cleavage sites have been confirmed by direct sequencing. In caliciviruses, VPg can also exist in the form of precursor with the upstream protein which is thought to function as a membrane anchor"
 /protein_id="[NP_786948.1](#)"

mat_peptide 3305..3847
 /gene="ORF1"
 /locus_tag="NoVGI_gp1"
 /old_locus_tag="NVgp1"
 /product="Pro"
 /note="Chymotrypsin-like cysteine proteinase; For Southampton calicivirus, both N-terminal and C-terminal cleavage sites have been confirmed by direct sequencing. The FCV proteinase (Pro) can exist in the form of a stable Pro-Pol precursor"
 /protein_id="[NP_786949.1](#)"

mat_peptide 3848..5371

```

/ gene="ORF1"
/ locus_tag="NoVGI_gp1"
/ old_locus_tag="NVgp1"
/ product="RdRp"
/ note="the FCV polymerase (Pol) can exist in the form of a
stable Pro-Pol precursor; RNA-dependent RNA polymerase"
/ protein_id="NP\_786950.1"
5358..6950
/ gene="ORF2"
/ locus_tag="NoVGI_gp2"
/ old_locus_tag="NVgp2"
/ db_xref="GeneID:1491972"
5358..6950
/ gene="ORF2"
/ locus_tag="NoVGI_gp2"
/ old_locus_tag="NVgp2"
/ note="ORF2; 58kd capsid protein; major capsid protein"
/ codon_start=1
/ product="VP1"
/ protein_id="NP\_056821.2"
/ db_xref="GeneID:1491972"
/ translation="MMMASKDATSSVDGASGAGQLVPEVNASDPLAMPVAGSSSTAVA
TAGQVNPIDPWIINNFBVQAPQGEFTISPNTPGDVLFDLSLGPLNPFLLHLSQMYNG
WVGNMRVRIMLAGNAFTAGKIIVSCIPPGFGSHNLITIAQATLFPVHVIADVRTLDPIEV
PLEDVRNVLFHNDRNQQTMRLLVCMLYTPLRTGGGTGDSFVVAGRVMTCSPDFNFLF
LVPPTVEQKTRPFTLPNLPLSSLSNSRAPLPISSMGISPDNVQSVQFQNGRCTLDGRL
VGTTTPVSLSHVAKIRGTSNGTVINLTLDGTPFHPFEGPAPIGFPDLGGCDWHINMTQ
FGHSSQTQYDVTTPDTFVPHLGSIQANGIGSGNYVGVLSWISPPSHPSGSQVDLWKI
PNYGSSITEATHLAPSVYPPGFGEVLVFFMSKMPGPGAYNLPCLLPQEYISHLASEQA
PTVGEAALLHYVDPDTGRNLGEFKAYPDGFLTCVPNGASSGPQQLPINGVFVFSWVS
RFYQLKPVGTASSARGRLGLRR"
6950..7588
/ gene="ORF3"
/ locus_tag="NoVGI_gp3"
/ old_locus_tag="NVgp3"
/ db_xref="GeneID:1491971"
6950..7588
/ gene="ORF3"
/ locus_tag="NoVGI_gp3"
/ old_locus_tag="NVgp3"
/ note="ORF3; minor structural protein; small basic
protein"
/ codon_start=1
/ product="VP2"

```

/protein_id="[NP_056822.1](#)"
/db_xref="GeneID:[1491971](#)"
/translation="MAQAIIGAIAASTAGSALGAGIQVGGEAALQSQRYQQNLQLQEN
SFKHREMIGYQVEASNQLLAKNLATRYSLLRAGGLTSADAARSVAGAPVTRIVDWNG
VRVSAPESSATTLRSGGFMSVPIPFASKQKQVQSSGISNPNYSPSSISRRTTSWVESQN
SSRFGNLSPIYHAEALNTVWLTPPGSTASSTLSSVPRGYFNTDRLPLFANNRR"
[3'UTR](#) 7589..7654

ORIGIN

```
1  gtgaatgatg atggcgtaa aagacgtcgt tcctactgct gctagcagtg aaaatgctaa
61  caacaatagt agtattaagt ctcgtctatt ggcgagactc aagggttcag gtggggctac
121 gtccccaccc aactcgataa agataaccaa ccaagatatg gctctggggc tgattggaca
181 ggtcccagcg ccaaaggcca catccgtcga tgtccctaaa caacagaggg atagaccacc
241 acggactgtt gccgaagtgc aacaaaatth gcgttggact gagagaccac aagaccagaa
301 tgttaagacg tgggatgagc ttgaccacac aacaaaacaa cagatacttg atgaacacgc
361 tgagtgggtt gatgccgtg gcttaggtcc aagtacacta cccactagtc atgaacggta
421 cacacatgag aatgatgaag gccaccaggt aaagtggctg gctagggaag gtgtagacct
481 tggcatatcc gggctcacga cgggtgtctg gcctgagtgg aatatgtgcc cgctaccacc
541 agttgaccaa aggagcacga cacctgcaac tgagcccaca attggtgaca tgatcgaatt
601 ctatgaaggg cacatctatc attatgctat atacataggt caaggcaaga cgggtgggtg
661 acactcccct caagcagcct tctcaataac gaggatcacc atacagcca tatcagcttg
721 gtggcgagtc tgttatgtcc cacaacaaaa acagaggctc acatacgacc aactcaaaga
781 attagaaaat gaaccatggc cgtatgccgc agtcacgaac aactgcttcg aattttgttg
841 ccaggctatg tgcttggaag atacttggtt gcaaaggaag ctcatctcct ctggccggtt
901 ttaccacccg acccaagatt ggtcccgaga cactccagaa ttccaacaag acagcaagtt
961 agagatgggt agggatgcag tgctagccgc tataaatggg ttggtgtcgc ggccatttaa
1021 agatcttctg ggtaagctca aacccttgaa cgtgcttaac ttactttcaa actgtgattg
1081 gacgttcatt ggggtcgtgg agatgggtgt cctcctttta gaactctttg gaactctttg
1141 gaaccacact gatgtttcca actttatagc ttactcctg ccagatttcc atctacaggg
1201 ccccgaggac cttgccaggg atctcgtgcc aatagtattg ggggggatcg gcttagccat
1261 aggattcacc agagacaagg taagtaagat gatgaagaat gctgttgatg gacttcgtgc
1321 ggcaaccagc ctgggtcaat atggcctaga aatattctca ttactaaaga agtacttctt
1381 cgggtggtgat caaacagaga aaaccctaaa agatattgag tcagcagtta tagatatgga
1441 agtactatca tctacatcag tgactcagct cgtgagggac aaacagctcg cacgggctta
1501 tatggccatc ttagataatg aagaagaaaa ggcaaggaaa ttatctgtca ggaatgccga
1561 cccacacgta gtatcctcta ccaatgctct catatcccgc atctcaatgg ctagggctgc
1621 attggccaag gctcaagctg aaatgaccag caggatgcgt cctgtggtca ttatgatgtg
1681 tgggccccct ggtataggtg aaaccaaggc agcagaacat ctggctaaac gcctagccaa
1741 tgagatacgg cctggtggta aggttgggct ggtccacagc gaggcagtgg atcattggga
1801 tggatatcac ggagaggaag tgatgctgtg ggacgactat ggaatgacaa agatacagga
1861 agactgtaat aaactgcaag ccatagccga ctcagcccc ctaacactca attgtgacct
1921 aatagaaaac aagggaatgc aatttgtgtc tgatgctata gtcacacca ccaatgctcc
1981 tggcccagcc ccagtggact ttgtcaacct cgggcctgtt tgccgaaggg tggacttcct
2041 tgtgtattgc acggcacctg aagttgaaca cagaggaaa gtcagtctg gggacacaac
2101 tgcactgaaa gactgcttca agcccattt ctcacatcta aaaatggagt tggctcccca
```

2161 agggggccttt gataaccaag ggaatacccc gtttggttaag ggtgtgatga agcccaccac
2221 cataaacagg ctgttaatcc aggctgtagc cttgacgatg gagagacagg atgagttcca
2281 actccagggg cctacgtatg actttgatac tgacagagta gctgcgttca cgaggatggc
2341 ccgagccaac gggttgggtc tcatatccat ggccctcccta ggcaaaaagc tacgcagtgt
2401 caccactatt gaaggattaa agaatgctct atcaggctat aaaatatcaa aatgcagtat
2461 acaatggcag tcaagggtgt acattataga atcagatggt gccagtgtac aaatcaaaga
2521 agacaagcaa gctttgacct ctctgcagca gacaattaac acggcctcac ttgccatcac
2581 tcgactcaaa gcagctaggg ctgtggcata cgcttcatgt ttccagtccg ccataactac
2641 catactacaa atggcgggat ctgcgctcgt tattaatcga gcggtcaagc gtatgtttgg
2701 taccctgaca gcagccatgg cattagaagg acctgggaaa gaacataatt gcagggtcca
2761 taaggctaag gaagctggaa agggggccat aggtcatgat gacatggtag aaaggtttgg
2821 cctatgtgaa actgaagagg aggagagtga ggaccaaatt caaatggtac caagtgatgc
2881 cgtcccagaa ggaaagaaca aaggcaagac caaaaaggga cgtggtcgca aaaataacta
2941 taatgcattc tctcgccgtg gtctgagtga tgaagaatat gaagagtaca aaaagatcag
3001 agaagaaaag aatggcaatt atagtataca agaatacttg gaggaccgcc aacgatatga
3061 ggaagaatta gcagaggtac aggcaggtgg tgatggtggc ataggagaaa ctgaaatgga
3121 aatccgtcac aggggtcttct ataaatccaa gagtaagaaa caccaacaag agcaacggcg
3181 acaacttggt ctagtgactg gatcagacat cagaaaacgt aagcccattg actggacccc
3241 gccaaagaat gaatgggcag atgatgacag agaggtggat tataatgaaa agatcaattt
3301 tgaagctccc ccgacactat ggagccgagt cacaaagttt ggatcaggat ggggcttttg
3361 ggtcagcccc acagtgttca tcacaaccac acatgtagtg ccaactggtg tgaaagaatt
3421 ctttggtgag ccctatcta gtatagcaat ccaccaagca ggtgagttca cacaattcag
3481 gttctcaaag aaaatgcgcc ctgacttgac aggtatggtc cttgaagaag gttgccctga
3541 agggacagtc tgctcagtc taattaaacg ggattcgggt gaactacttc cgctagccgt
3601 ccgtatgggg gctattgcct ccatgaggat acagggtcgg cttgtccatg gccaatcagg
3661 gatgttactg acaggggcca atgcaaaggg gatggatctt ggcaactatac caggagactg
3721 cggggcacca tacgtccaca agcgcgggaa tgactgggtt gtgtgtggag tccacgctgc
3781 agccacaaag tcaggcaaca ccgtggtctg cgctgtacag gctggagagg gcgaaaccgc
3841 actagaaggt ggagacaagg ggcattatgc cggccacgag attgtgaggt atggaagtgg
3901 ccagcactg tcaactaaaa caaatctctg gaggtcctcc ccagaaccac tgccccccgg
3961 agtatatgag ccagcatacc tggggggcaa ggacccccgt gtacagaatg gcccatccct
4021 acaacaggt aactgaaacc ctttgcgac ccccgcgcc gcattgcctga
4081 gcctggccta ctggaggctg cggttgagac tgtaacatcc atgttagaac agacaatgga
4141 taccccaagc ccgtggtctt acgctgatgc ctgccaatct cttgacaaaa ctactagttc
4201 ggggtaccct caccataaaa ggaagaatga tgattggaat ggcaccacct tcgttgagga
4261 gctcgggtgag caagctgcac acgccaacaa tatgtatgag aatgctaaac atatgaaacc
4321 catttacact gcagccttaa aagatgaact agtcaagcca gaaaagattt atcaaaaagt
4381 caagaagcgt ctactatggg gcgccgatct cggaacagtg gtcagggccg cccgggcttt
4441 tggcccatth tgtgacgcta taaaatcaca tgtcatcaaa ttgccaataa aagttggcat
4501 gaacacaata gaagatggcc ccctcatcta tgctgagcat gctaaatata agaatacttt
4561 tgatgcagat tatacagcat gggactcaac acaaaataga caaattatga cagaatcctt
4621 ctccattatg tcgcgcctta cggcctcacc agaattggcc gaggttgtgg cccaagattt
4681 gctagcacca tctgagatgg atgtaggtga ttatgtcatc agggtaaaag aggggctgcc
4741 atctggattc ccatgtactt ccaggtgaa cagcataaat cactggataa ttactctctg

4801 tgcactgtct gaggcactg gtttatcacc tgatgtggtg caatccatgt catattttctc
4861 attttatggt gatgatgaga ttgtgtcaac tgacatagat tttgaccag ccgcctcac
4921 tcaaattctc aaggaatatg gcctcaaacc aacaaggcct gacaaaacag aaggaccaat
4981 acaagtgagg aaaaatgtgg atggactggt cttcttgcg cgaccattt cccgtgatgc
5041 ggcagggttc caaggcaggt tagatagggc ttcgattgaa cgccaaatct tctggaccgg
5101 cgggcccaat cattcagatc catcagagac tctagtgcc cactcaca gaaaaatata
5161 gttgatttca cttctagggg aagcttact ccatggtgag aaattttaca gaaagatttc
5221 cagcaaggtc atacatgaaa tcaagactgg tggattggaa atgtatgtcc caggatggca
5281 ggccatgttc cgctggatgc gcttccatga cctcggattg tggacaggag atcgcgatct
5341 tctgcccgaa ttcgtaaatg atgatggcgt ctaaggacgc tacatcaagc gtggatggcg
5401 ctagtggcgc tggtcagttg gtaccggag ttaatgcttc tgaccctctt gcaatggatc
5461 ctgtagcagg ttcttcgaca gcagtcgca ctgctggaca agttaatcct attgatccct
5521 ggataattaa taattttgtg caagcccccc aagggtgaatt tactatttcc ccaaataata
5581 ccccggtga tgttttgtt gatttgagtt tgggtcccca tcttaatcct ttcttgetcc
5641 atctatcaca aatgtataat ggttgggttg gtaacatgag agtcaggatt atgctagctg
5701 gtaatgcctt tactgcgggg aagataatag tttcctgcat acccctggt tttggttcac
5761 ataatcttac tatagcaca gcaactctct ttccacatgt gattgctgat gttaggactc
5821 tagaccccat tgaggcgcct ttggaagatg ttaggaatgt tctctttcat aataatgata
5881 gaaatcaaca aaccatgcgc cttgtgtgca tgctgtacac cccctccgc actggtggtg
5941 gtactggtga ttcttttga gttgcagggc gagttatgac ttgccccagt cctgatttta
6001 atttcttgtt tttagtccct cctacggtgg agcagaaaac caggcccttc acactccaa
6061 atctgccatt gagttctctg tctaactcac gtgcccctct cccaatcagt agtatgggca
6121 tttccccaga caatgtccag agtgtgcagt tccaaaatgg tcggtgtact ctggatggcc
6181 gcctggttgg caccaccca gtttcattgt cacatgttgc caagataaga gggacctcca
6241 atggcactgt aatcaacctt actgaattgg atggcacacc ctttcacctt tttgagggcc
6301 ctgcccccat tgggtttcca gacctcggtg gttgtgattg gcataatcaat atgacacagt
6361 ttggccattc tagccagacc cagtatgatg tagacaccac ccctgacact tttgtcccc
6421 atcttgggtc aattcaggca aatggcattg gcagtggtaa ttatgttgggt gttcttagct
6481 ggatttcccc cccatcacac ccgtctggct cccaagttga cctttggaag atccccaatt
6541 atgggtcaag tattacggag gcaacacatc tagccccttc tgtatacccc cctggtttcg
6601 gagaggatatt ggtctttttc atgtcaaaaa tgccagggtcc tgggtgcttat aatttgcctt
6661 gtctattacc acaagagtac atttcacatc ttgctagtga acaagcccct actgtagggtg
6721 aggtgcctt gctccactat gttgacctg ataccggtcg gaatcttggg gaattcaaag
6781 cataccctga tggtttctc acttgtgtcc ccaatggggc tagctcgggt ccacaacagc
6841 tgccgatcaa tggggtctt gtctttgtt catgggtgtc cagattttat caattaaagc
6901 ctgtgggaac tgccagctcg gcaagaggta ggcttggct gcgccgataa tggcccaagc
6961 cataattggt gcaattgctg cttccacagc aggtagtgtc ctgggagcgg gcatacagggt
7021 tgggtggcga gcggccctcc aaagccaaag gtatcaacaa aatttgcaac tgcaagaaaa
7081 ttcttttaaa catgacaggg aaatgattgg gtatcaggtt gaagcttcaa atcaattatt
7141 ggctaaaaat ttggcaacta gatattcact cctccgtgct gggggttga ccagtgtga
7201 tgcagcaaga tctgtggcag gagctccagt caccgcatt gtagattgga atggcgtgag
7261 agtgtctgct ccgagtcct ctgctaccac attgagatcc ggtggcttca tgtcagttcc
7321 cataccattt gcctctaagc aaaaacagggt tcaatcatct ggtattagta atccaaatta
7381 ttccccttca tccatttctc gaaccactag ttgggtcgag tcacaaaact catcgagatt

7441 tggaaatcctt tctccatacc acgcggaggc tctcaataca gtgtggttga ctccaccgg
7501 ttcaacagcc tcttctacac tgtcttctgt gccacgtggt tatttcaata cagacaggtt
7561 gccattattc gcaaataata ggcgatgatg ttgtaatatg aaatgtgggc atcatattca
7621 ttttaattagg ttttaattagg ttttaatttga tggt

//