

## 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: <a href="#">122928</a> "                                                                                                                                                                                                                                                                                                                                                                   |
| <a href="#">5'UTR</a> | 1..4                                                                                                                                                                                                                                                                                                                                                                                                        |
| <a href="#">gene</a>  | 5..5374                                                                                                                                                                                                                                                                                                                                                                                                     |
|                       | /gene="ORF1"                                                                                                                                                                                                                                                                                                                                                                                                |
|                       | /locus_tag="NoVGI_gp1"                                                                                                                                                                                                                                                                                                                                                                                      |
|                       | /old_locus_tag="NVgp1"                                                                                                                                                                                                                                                                                                                                                                                      |
|                       | /db_xref="GeneID: <a href="#">1491970</a> "                                                                                                                                                                                                                                                                                                                                                                 |
| <a href="#">CDS</a>   | 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="MMMASKDVVPTAASSENANNSSIKSRLLARLKSGGATSPPN  
IKITNQDMALGLIGQVPAPKATSVDPKQQRDRPPRTVAEVQQNLRWTERPQDQNVKT  
WDELDHTTKQQILDEHAWEFDAGGLGPSTLPTSHERYTHENDEGHQVKWSAREGVDLG  
ISGLTTVSGPEWNMCPLPPVDQRSTTPATEPTIGDMIEFYEGHIYHYAIYIGQGKT  
VHSPQAAFSITRITIQPISAWVRVCYVPQPKQRLTYDQLKELENWPYAAVTNNCFE  
FCCQVMCLEDTWLQRKLISSGRFYHPTQDWSRDTPEFQQDSKLEMVRDAVLAAINGLV  
SRPFKDLLGKLKPLNVNLNLSNCDWTFMGVVEMVVLLELFGIFWNPPDVSNFIASLL  
PDFHLQGPEDLARDLVPIVLGGIGLAIGFTRDKVSKMMKNAVDGLRAATQLGQYGLEI  
FSLKKYFFGGDQTEKTLKDIESAVIDMEVLSSTSVTQLVRDKQSARAYMAILDNEEE  
KARKLSVRNADPHVVSSTNALISRISMARAAAKQAEMTSRMRPVVIMCGPPGIGK  
TKAAEHLAKRLANEIRPGGKVLVPREAVDHWGYPHGEVMLWDDYGMTKIQEDCNKL  
QAIADSAPLTLNCDRIENKGMQFVSDAIVITTNAPGPAPVDFVNLGPVCRRVDFLVYC  
TAPEVEHTRKVS PGDTTALKDCFKPDFSHLKMELAPQGGFDNQGNTPFKGVMKPTTI  
NRLLIQAVALTMERQDEFQLQGPTYDFDTRVAAFTRMARANGLGLISMASLGKKLRS  
VTTIEGLKNALSGYKISKCSIQWQSRVYIIESDGASVQIKEDKQALTPLOQTINTASL  
AITRLKAARAVAYASCFQSAITTIILQMAGSALVINRAVKRMFGTRTAAMALEGPGKEH  
NCRVHKAKEAGKGPIGHDDMVERFGLCETEEEESEDQIQMVPSDAVPEGKNKGKTKKG  
RGRKNYNNAFSRRGLSDEEYEEYKKIREEKNGNYSIQEYLEDQRQYEEELADEVQAGGD  
GGIGETEMEIRHRVIFYKSKSKKHQQEQRRQLGLVTGSDIRKRPIDWTTPKNEWADDD  
REVVDYNEKINFEAPPTLWSRVTKFGSGWGFVWSPTVFITTHVVP TGVEFFGEPLSS  
IAIHQAGEFTQFRFSKMRPDLTGMVLEEGCPEGTVCSVLIKRDSGELLPLAVRMGAI  
ASMRIQGRVLVHGQSGMLLTGANAKGMDLGTIPGDCGAPYVHKRGNDDWVCGVHAAATK  
SGNTVVCAVQAGEGETALEGGDKGHYAGHEIVRYGSGPALSTKTKFWRSSPEPLPPGV  
YEPAYLGKGDPRVQNGPSLQQVLRDQLKPFADPRGRMPEPGLLEAAVETVTSMLEQTM  
DTPSPWSYADACQSLDKTTSSGYPHHKRKNDDWNGTTFVGELGEQAAHANMNYENAKH  
MKPIYTAALKDELVKPEKIYQKVKRLLWGADLGTVVRAARAFGPFCD AIKSHVIKLP  
IKVGMNTIEDGPLIYAEHAKYKNHFDADYTAWDSTQNRQIMTESFSIMSRLTASPELA  
EVVAQDLLAPSEMDVG DYVIRVKEGLPSGFPCTSQVNSINHWIITLCALSEATGLSPD  
VVQSMSYFSFYGDDEIVSTDIDFDPARLTQILKEYGLKPTRPDKTEGPIQVRKNVDGL  
VFLRRTISRDAAGFQGRLDRASIERQIFWTRGPNHSDPSETLVPHTQRKIQLISLLGE  
ASLHGEKFYRKISSKVIHEIKTGGLEMYVPGWQAMFRWRFHDLGLWTGDRDLLPEFV  
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"
/protein_id="NP_786945.1"
mat_peptide    1199..2287
/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

```

|      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|      | <div> <div>/gene="ORF1"</div> <div>/locus_tag="NoVGI_gp1"</div> <div>/old_locus_tag="NVgp1"</div> <div>/product="RdRp"</div> <div>/note="the FCV polymerase (Pol) can exist in the form of a stable Pro-Pol precursor; RNA-dependent RNA polymerase"</div> <div>/protein_id="NP_786950.1"</div> </div>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| gene | 5358..6950                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|      | <div> <div>/gene="ORF2"</div> <div>/locus_tag="NoVGI_gp2"</div> <div>/old_locus_tag="NVgp2"</div> <div>/db_xref="GeneID:1491972"</div> </div>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| CDS  | 5358..6950                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|      | <div> <div>/gene="ORF2"</div> <div>/locus_tag="NoVGI_gp2"</div> <div>/old_locus_tag="NVgp2"</div> <div>/note="ORF2; 58kd capsid protein; major capsid protein"</div> <div>/codon_start=1</div> <div>/product="VP1"</div> <div>/protein_id="NP_056821.2"</div> <div>/db_xref="GeneID:1491972"</div> <div>/translation="MMASKDATSSVDGASGAGQLVPEVNASDPLAMDPVAGSSTAVATAGQVNPIDPWIINNWFVQAPQGEFTISPNNTPGDVLFDSLGLPHLNPFLHLHSQMYNGWVGNMRVRIMLAGNAFTAGKIIIVSCIPPFGFSGHNLITIAQATLFPHVIADVRLDPIEVPLEDVRNVLFHNNDRNQQTMLRVCMLYTPLRTGGGTGDSFVVAGRVMTCPSPDFNFLFLVPPTVEQKTRPFTLPNLPLSSLSNSRAPLPISSMGISPDNVQSVQFQNGRCTLDGRLVGTTPVSLSHVAKIRGTSNGTVINLTLDGTFFHPFEGPAPIGFDPDLGGCDWHINMTQFGHSSQTQYDVDTPDTFVPHLGSIQANGIGSGNYVGVLSWISPPSHPSGSQVDLWKIPNYGSSITEATHLAPSVYPPGFGEVLVFFMSKMPGPGAYNLPCLLPQEYISHLASEQAPTVGEAALLHYVDPDTGRNLGEFKAYPDGFLTCVPNGASSGPQQLPINGVFVFSWVS RFYQLKPVGTASSARGRLGLRR"</div> </div> |
| gene | 6950..7588                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|      | <div> <div>/gene="ORF3"</div> <div>/locus_tag="NoVGI_gp3"</div> <div>/old_locus_tag="NVgp3"</div> <div>/db_xref="GeneID:1491971"</div> </div>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| CDS  | 6950..7588                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|      | <div> <div>/gene="ORF3"</div> <div>/locus_tag="NoVGI_gp3"</div> <div>/old_locus_tag="NVgp3"</div> <div>/note="ORF3; minor structural protein; small basic protein"</div> <div>/codon_start=1</div> <div>/product="VP2"</div> </div>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |

/protein\_id="[NP\\_056822.1](#)"  
/db\_xref="GeneID:[1491971](#)"  
/translation="MAQAIIGAIAASTAGSALGAGIQVGGEAALQSORYQQNLQLQEN  
SFKHREMIGYQVEASNQLLAKNLATRYSLRAGGLTSADAARSVAGAPVTRIVDWNQ  
VRVSAPESSATTLRSGGFMSVPIPFASKQKQVQSSGISNPNYSPSSISRRTTSWVESQN  
SSRFGNLSPIYHAEALNTVWLTPPGSTASSTLSSVPRGYFNTDRLPLFANNRR"  
[3'UTR](#) 7589..7654

ORIGIN

```
1  gtgaatgatg atggcgtcaa 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 acggactggt gccgaagttc aacaaaatth gcgttggact gagagaccac aagaccagaa
301 tgttaagacg tgggatgagc ttgaccacac aacaaaacaa cagatacttg atgaacacgc
361 tgagtggtht gatgccgtg gcttaggtcc aagtacacta cccactagtc atgaacggtg
421 cacacatgag aatgatgaag gccaccaggt aaagtggctg gctagggag gtgtagacct
481 tggcatatcc gggctcacga cgggtgtctg gcctgagtgg aatatgtgcc cgctaccacc
541 agttgaccaa aggagcacga cacctgcaac tgagcccaca attggtgaca tgatcgaatt
601 ctatgaaggc cacatctatc attatgctat atacataggt caaggcaaga cgggtgggtg
661 acactcccct caagcagcct tctcaataac gaggatcacc atacagccca tatcagcttg
721 gtggcgagtc tgttatgtcc cacaaccaa acagaggctc acatacgacc aactcaaaga
781 attagaaaat gaacctggc cgtatgccg agtcacgaac aactgcttcg aattttgttg
841 ccaggtcatg tgcttggaag atacttggtt gcaaaggaag ctcatctcct ctggccggtt
901 ttaccaccgc acccaagatt ggtcccgaga cactccagaa ttccaacaag acagcaagtt
961 agagatggtt agggatgcag tgctagccgc tataaatggg ttggtgtcgc ggccatttaa
1021 agatcttctg ggtaagctca aacccttgaa cgtgcttaac ttactttcaa actgtgattg
1081 gacgttcag ggggtcgtgg agatgggtgt cctcctttta gaactcttg gaactcttg
1141 gaaccacct gatgtttcca actttatagc ttcactcctg ccagatttcc atctacagg
1201 ccccgaggac cttgccagg atctcgtgcc aatagtattg ggggggatcg gcttagccat
1261 aggattcacc agagacaagg taagtaagat gatgaagaat gctgttgatg gacttcgtgc
1321 ggcaaccag ctcggtcaat atggcctaga aatatttca ttactaaaga agtacttctt
1381 cgggtggtgat caaacagaga aaaccctaaa agatattgag tcagcagtta tagatatgga
1441 agtactatca tctacatcag tgactcagct cgtgagggac aaacagtctg caggggctta
1501 tatggccatc ttagataatg aagaagaaaa ggcaaggaaa ttatctgtca ggaatgccga
1561 cccacacgta gtatctcta ccaatgctct catatcccgc atctcaatgg ctagggtgc
1621 attggccaag gctcaagctg aaatgaccag caggatgcgt cctgtggtca ttatgatgtg
1681 tgggccccct ggtataggta aaaccaaggc agcagaacat ctggctaaac gcctagccaa
1741 tgagatacgg cctggtggtg aggttgggtc ggtccacagg gaggcagtgg atcattggga
1801 tggatatcac ggagaggaag tgatgctgtg ggacgactat ggaatgacaa agatacagga
1861 agactgtaat aaactgcaag ccatagccga ctcagcccc ctaacactca attgtgaccg
1921 aatagaaaac aagggaatgc aatttgtgtc tgatgctata gtcacacca ccaatgctcc
1981 tggcccagcc ccagtggact ttgtcaacct cgggcctgtt tgccgaaggg tggacttctc
2041 tgtgtattgc acggcacctg aagttgaaca cagaggaaa gtcagtcctg gggacacaac
2101 tgcactgaaa gactgcttca agcccattt ctcacatcta aaaatggagt tggctcccca
```

2161 agggggccttt gataaccaag ggaatacccc gtttggttaag ggtgtgatga agcccaccac  
2221 cataaacagc ctgttaatcc aggctgtagc cttgacgatg gagagacagc atgagttcca  
2281 actccagggg cctacgtatg actttgatac tgacagagta gctgcgttca cgaggatggc  
2341 ccgagccaac gggttgggtc tcatatccat ggcctcccta ggcaaaaagc tacgcagtgt  
2401 caccactatt gaaggattaa agaatgctct atcaggctat aaaatatcaa aatgcagtat  
2461 acaatggcag tcaaggggtg acattataga atcagatggg gccagtgtac aaatcaaaga  
2521 agacaagcaa gctttgacct ctctgcagca gacaattaac acggcctcac ttgccatcac  
2581 tcgactcaaa gcagctaggg ctgtggcata cgcttcatgt ttccagtccg ccataactac  
2641 catactacaa atggcgggat ctgcgctcgt tattaatcga gcggccaagc gtatgtttgg  
2701 taccctgaca gcagccatgg cattagaagg acctgggaaa gaacataatt gcagggcca  
2761 taaggctaag gaagctggaa agggggccat aggtcatgat gacatggtag aaaggtttgg  
2821 cctatgtgaa actgaagagg aggagagtga ggaccaaatt caaatggtac caagtgtatg  
2881 cgtcccagaa ggaaagaaca aaggcaagac caaaaaggga cgtgggtcgca aaaataacta  
2941 taatgcattc tctcgccgtg gtctgagtga tgaagaatat gaagagtaca aaaagatcag  
3001 agaagaaaag aatggcaatt atagtataca agaatacttg gaggaccgcc aacgatatga  
3061 ggaagaatta gcagaggtac aggcaggtgg tgatgggtggc ataggagaaa ctgaaatgga  
3121 aatccgtcac agggctcttct ataaatccaa gagtaagaaa caccaacaag agcaacggcg  
3181 acaacttggg ctagtgtactg gatcagacat cagaaaacgt aagcccattg actggacccc  
3241 gccaaagaat gaatgggcag atgatgacag agaggtggat tataatgaaa agatcaattt  
3301 tgaagctccc ccgacactat ggagccgagt cacaaagttt ggatcaggat ggggcttttg  
3361 ggtcagcccg acagtgttca tcacaaccac acatgtagtg ccaactggtg tgaaagaatt  
3421 ctttggtgag cccctatcta gtatagcaat ccaccaagca ggtgagttca cacaattcag  
3481 gttctcaaag aaaatgcgcc ctgacttgac aggtatggct cttgaagaag gttgccctga  
3541 agggacagtc tgctcagtcc taattaaacg ggattcgggt gaactacttc cgctagccgt  
3601 ccgtatgggg gctattgcct ccatgaggat acaggggtcgg cttgtccatg gccaatcagg  
3661 gatgttactg acagggggcca atgcaaaggg gatggatctt ggcactatac caggagactg  
3721 cggggcacca tacgtccaca agcgcgggaa tgactgggtt gtgtgtggag tccacgtgc  
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 acaacaggtg ctacgtgacc aactgaaacc ctttgcgac cccgcggcc gcattgctga  
4081 gcctggccta ctggaggctg cggttgagac tgtaacatcc atgttagaac agacaatgga  
4141 taccccaagc ccgtggtctt acgctgatgc ctgccaatct cttgacaaaa ctactagttc  
4201 ggggtaccct caccataaaa ggaagaatga tgattggaat ggcaccacct tcgttgaga  
4261 gctcgggtgag caagctgcac acgccaacaa tatgtatgag aatgctaaac atatgaaacc  
4321 catttacact gcagccttaa aagatgaact agtcaagcca gaaaagattt atcaaaaagt  
4381 caagaagcgt ctactatggg gcgccgatct cggaacagtg gtcagggccg cccgggcttt  
4441 tggccattt tgtgacgcta taaaatcaca tgatcatcaa ttgccaataa aagttggcat  
4501 gaacacaata gaagatggcc cctcatcta tgctgagcat gctaaatata agaattcttt  
4561 tgatgcagat tatacagcat gggactcaac acaaaataga caaattatga cagaatcctt  
4621 ctccattatg tcgcgcctta cggcctcacc agaattggcc gaggttgtgg cccaagattt  
4681 gctagcacca tctgagatgg atgtaggatga ttatgtcatc agggtaaaag aggggctgcc  
4741 atctggattc ccatgtactt ccaggtgaa cagcataaat cactggataa ttactctctg

4801 tgcactgtct gaggccactg gtttatcacc tgatgtggtg caatccatgt catatttctc  
4861 attttatggt gatgatgaga ttgtgtcaac tgacatagat tttgaccag cccgcctcac  
4921 tcaaatcttc aaggaatatg gcctcaaacc aacaaggcct gacaaaacag aaggaccaat  
4981 acaagtgag aaaaatgtgg atggactggt cttcttgagg cgcaccattt cccgtgatgc  
5041 ggcagggttc caaggcaggt tagatagggc ttcgattgaa cgccaaatct tctggaccgc  
5101 cgggcccatt cattcagatc catcagagac tctagtgcc acaactcaaa gaaaaataca  
5161 gttgatttca cttctagggg aagcttcact ccatgggtgag aaattttaca gaaagatttc  
5221 cagcaaggtc atacatgaaa tcaagactgg tggattggaa atgtatgtcc caggatggca  
5281 ggccatgttc cgctggatgc gcttccatga cctcggattg tggacaggag atcgcgatct  
5341 tctgcccga ttcgtaaatg atgatggcgt ctaaggacgc tacatcaagc gtggatggcg  
5401 ctagtggcgc tggtcagttg gtaccggagg ttaatgcttc tgaccctctt gcaatggatc  
5461 ctgtagcagg ttcttcgaca gcagtcgcga ctgctggaca agttaatcct attgatccct  
5521 ggataattaa taattttgtg caagccccc aaggtgaatt tactatttcc ccaataata  
5581 cccccggtga tgttttgttt gatttgagtt tgggtcccca tottaatcct ttcttgctcc  
5641 atctatcaca aatgtataat ggttgggttg gtaacatgag agtcaggatt atgctagctg  
5701 gtaatgcctt tactgcgggg aagataatag tttcctgcat accccctggt tttggttcac  
5761 ataatcttac tatagcacia gcaactctct ttccacatgt gattgctgat gttaggactc  
5821 tagaccccat tgaggtgcct ttggaagatg ttaggaatgt tctctttcat aataatgata  
5881 gaaatcaaca aaccatgcgc cttgtgtgca tgctgtacac cccctccgc actggtggtg  
5941 gtactggtga ttcttttgta gttgcagggc gagttatgac ttgccccagt cctgatttta  
6001 atttcttgtt tttagtccct cctacggtgg agcagaaaac caggcccttc acaactccaa  
6061 atctgccatt gagttctctg tctaactcac gtgcccctct cccaatcagt agtatgggca  
6121 tttcccaga caatgtccag agtgtgcagt tccaaaatgg tcggtgtact ctggatggcc  
6181 gcctggttgg caccacccca gtttcattgt cacatgttgc caagataaga gggacctcca  
6241 atggcactgt aatcaacctt actgaattgg atggcacacc ctttcaccct tttgagggcc  
6301 ctgcccccat tgggtttcca gacctcgggt gttgtgattg gcataatcaat atgacacagt  
6361 ttggccattc tagccagacc cagtatgatg tagacaccac cctgacact tttgtcccc  
6421 atcttggttc aattcaggca aatggcattg gcagtggtaa ttatgttggt gttcttagct  
6481 ggatttcccc cccatcacac ccgtctggct cccaagtga cctttggaag atccccatt  
6541 atgggtcaag tattacggag gcaacacatc tagcccttc tgtatacccc cctggtttcg  
6601 gagaggtatt ggtctttttc atgtcaaaaa tgccagggtc tgggtgcttat aatttgccct  
6661 gtctattacc acaagagtac atttcacatc ttgctagtga acaagcccct actgtagggtg  
6721 aggtgtccct gctccactat gttgaccctg ataccggtcg gaatcttggg gaattcaaa  
6781 cataccctga tgggtttcct acttgtgtcc ccaatggggc tagctcgggt ccacaacagc  
6841 tgccgatcaa tggggtcttt gtctttgttt catgggtgtc cagattttat caattaaagc  
6901 ctgtgggaac tgccagctcg gcaagaggtg ggcttggtct gcgccgataa tggcccaagc  
6961 cataattggt gcaattgctg cttccacagc aggtagtgtc ctgggagcgg gcatacaggt  
7021 tgggtggcga gcggccctcc aaagccaaag gtatcaacaa aatttgcaac tgcaagaaaa  
7081 ttcttttaaa catgacaggg aaatgattgg gtatcagggt gaagcttcaa atcaattatt  
7141 ggctaaaaat ttggcaacta gatattcact cctccgtgct gggggtttga ccagtgtga  
7201 tgcagcaaga tctgtggcag gagctccagt caccgcatt gtagattgga atggcgtgag  
7261 agtgtctgct cccgagtcct ctgctaccac attgagatcc ggtggcttca tgtcagttcc  
7321 cataccattt gcctctaagc aaaaacaggt tcaatcatct ggtattagta atccaaatta  
7381 ttccccttca tccatttctc gaaccactag ttgggtcgag tcacaaaact catcgagatt

7441 tggaaatctt tctccatacc acgcggaggc tctcaataca gtgtggttga ctccaccg  
7501 ttcaacagcc tcttctacac tgtcttctgt gccacgtggt tatttcaata cagacagg  
7561 gccattattc gcaaataata ggcgatgatg ttgtaatatg aaatgtgggc atcatatt  
7621 ttttaattagg ttttaattagg ttttaattga tggt