

Human coronavirus 229E, complete genome

NCBI Reference Sequence: NC_002645.1

[FASTA Graphics](#)

[Go to:](#)

LOCUS NC_002645 27317 bp ss-RNA linear VRL 17-AUG-2021

DEFINITION Human coronavirus 229E, complete genome.

ACCESSION NC_002645

VERSION NC_002645.1

DBLINK BioProject: [PRJNA485481](#)

KEYWORDS RefSeq.

SOURCE Human coronavirus 229E

ORGANISM [Human coronavirus 229E](#)

Viruses; Riboviria; Orthornavirae; Pisuviricota; Pisoniviricetes; Nidovirales; Cornidovirineae; Coronaviridae; Orthocoronavirinae; Alphacoronavirus; Duvinacovirus.

REFERENCE 1 (bases 1 to 27317)

AUTHORS Ziebuhr,J., Thiel,V. and Gorbalenya,A.E.

TITLE The autocatalytic release of a putative RNA virus transcription factor from its polyprotein precursor involves two paralogous papain-like proteases that cleave the same peptide bond

JOURNAL J. Biol. Chem. 276 (35), 33220-33232 (2001)

PUBMED [11431476](#)

REFERENCE 2 (bases 1 to 27317)

AUTHORS Thiel,V., Herold,J., Schelle,B. and Siddell,S.G.

TITLE Infectious RNA transcribed in vitro from a cDNA copy of the human coronavirus genome cloned in vaccinia virus

JOURNAL J. Gen. Virol. 82 (Pt 6), 1273-1281 (2001)

PUBMED [11369870](#)

REFERENCE 3 (bases 1 to 27317)

AUTHORS Thiel,V., Herold,J. and Siddell,S.G.

TITLE Direct Submission

JOURNAL Submitted (11-SEP-2000) Virology, University of Wuerzburg, Versbacherstr.7, Wuerzburg 97078, Germany

REFERENCE 4 (bases 1 to 27317)

CONSRM NCBI Genome Project

TITLE Direct Submission

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

COMMENT REVIEWED [REFSEQ](#): This record has been curated by NCBI staff. The reference sequence was derived from [AF304460](#).

The mature peptides were added to the annotation at the NCBI.

COMPLETENESS: full length.

FEATURES Location/Qualifiers

source 1..27317

```
/organism="Human coronavirus 229E"
/mol_type="genomic RNA"
/db_xref="taxon:11137"
/clone="inf-1"

5'UTR      1..292
gene       293..20568
          /locus_tag="HCoV229Egp1"
          /db_xref="GeneID:918764"

CDS        join(293..12520,12520..20568)
          /locus_tag="HCoV229Egp1"
          /note="translated by -1 ribosomal frameshift"
          /codon_start=1
          /product="replicase polyprotein 1ab"
          /protein_id="NP_073549.1"
          /db_xref="GeneID:918764"
          /translation="MACNRVTLAVASDSEISANGCSTIAQAVRRYSEAASNGFRACRF
VSLDLQDCIVGIADDTYVMGLHGNQTLFCNIMKFSDRPFMLHGWLVSFNSNYLLEEFD
VVFGRGGGNVTYTDQYLCGADGKPVMSIDLWQFVDHFGENEEIIINGHTYVCAWLTK
RKPLDYKRQNNLAIEEIEYVHGDALHTLRNGSVLEMAKEVKTSSKVVLSDALDKLYKV
FGSPVMTNGSNILEAFTKPVFISALVQCTCGTKSWSVGDWTGFKSSCCNVISNKL CVV
PGNVKPGDAVITTTQAGAGIKYFCGMLKFVANIEGVSVWRVIALQSVDCFVASSTFV
EEEHVNRMDTFCFNVRNSVTDECR LAMLGAEMTSNVRQVASGVIDISTGWFDVYDDI
FAESKPWFVRKAEDIFGPCWSALASALKQLKVTTGELVRFVKSICNSAVAVVGGTIQI
LASVPEKFLNAFDVFVTAIQTVFDCAVETCTIAGKAFDKVFDYVLLDNALVKLVTTKL
KGVREGLNKVKYATVVVGSTEEVKSSRVERSTAVLTIANNYSKLFDEGYTVVIGDVA
YFVSDGYFRLMASPNSVLTTAVYKPLFAFNVMGMTRPEKFPTTVTCENLES AVL FVN
DKITEFQLDYSIDVIDNEIIVKPNISLCVPLYVRDYVDKWD DFCRQYSNESWFEDDYR
AFISVLDITDAAVKAAESKAFVDTI VPPCPSILKVIDGKKIWN GVIKVNVSVRDWLKS
LKLNLTTQQGLLGTCAKRFRWL GILLEAYNAFLDTVVSTVKIGGLTFKTYAFDKPYIV
IRDIVCKVENKTEAEWIELFPHNDRIKSFSTFESAYMPIADPTHFDIEEVELLDAEFV
EPGCGGILAVIDEHV FYKKDGVYYP SNGTNILPVAFTKAAGGKVSFSDDVEVKDIEPV
YRVKLCFEFEDEKLVDVCEKAIGKKIKHEGDWDSFCKTIQSALSVVSCYVNLPTYIY
DEEGGNDLSLPVMISEWPLSVQQAQQEATLPDIAEDVVDQVEEVNSIFDIETVDVKHD
VSPFEMPFEELNGLKILKQLDNNCWVNSVMLQIQLTGILDGDYAMQFFKMGRVAKMIE
RCYTAEQCIRGAMGDVGLCMYRL LKDLHTGFMVMDYKCSCTSGRLEESGAVLFCTPTK
KAFPYGTCLNCNAPRMCTIRQLQGTII FVQQKPEPVNPVSFVVKPVCSSIFRGAVSCG
HYQTNIYSQNL CVDGFGVNKIQPWTNDALNTICIKDADYNAKVEISVTPIKNTVDTTP
KEEFVVKEKLNAFLVHDNVAFYQGDVDTVVNGVDFDFIVNAANENLAHGGLAKALDV
YTKGKLQRLSKEHIGLAGKVKGVTGVMVECDSLRIFNVVGRPKGKHERDLLIKAYNTI
NNEQGTPLTPILSCGIFGIKLETSLEVLLDVCNTKEVKVFVYTDTEVCKVKDFVSGLV
NVQKVEQPKIEPKPVSIVKVPKPYRVDGKFSYFTEDLLCVADDKPIVLF TDSMLTLD
DRGLALDNALSGVLSAAIKDCVDINKAIPSGNLIKFDIGSVVVMCVVPSEKDKHLDN
NVQRCTRKLNRMLCDIVCTIPADYILPLVLSLTCNVSFV GELKAAEAKVITIKVTED
GVNVHDVTVTTDKSFEQQVGVIADKDKDLSGAVPSDLNTSELLTKAIDVDWVEFYGFK
```

DAVTFATVDHSAFAYESAVVNGIRVLKTSDNNCWVNAVCIALQYSKPHFISQGLDAAW
NKFVLGDVEIFVAFVYYVARLMKGDKGDAEDTLTKLSKYLANEAQVQLEHYSSCVECD
AKFKNSVASINSAIVCASVKRDGVQVGYCVHGIKYYSRVSVRGRAIIVSVEQLEPCA
QSRLLSGVAYTAFSGPVDKGHYTVYDTAKKSMYDGDREVKHDLSSLVTSVVMVGGYV
APVNTVKPKPVINQLDEKAQKFFDFGDFLIHNFVIFFTWLLSMFTLCKTAVTTGDVKI
MAKAPQRTGVVLKRSCLKYNLKASAAVLKSKWLLAKFTKLLLLIYTLYSVVLCCVRFG
PFNFCSETVNGYAKSNFVKDDYCDGSLGCKMCLFGYQELSQFSHLDVVKHITDPLFS
NMQPFIVMVLLLLIFGDNYLRCLLYFVAQMISTVGVLGYKETNWFLHFIPFDVICDE
LLVTVIVIKVISFVRHVLFGCENPDCIACSKSARLKRFPVNTIVNGVQRSFYVNANGG
SKFCKKHRFFCVDCCSYGYGSTFITPEVSRELGNITKTNVQPTGPAYVMIDKVEFENG
FYRLYSCETFWRYNFDITESKYSKEVFKNCNVLDDFIVFNNNGTNTVQKNASVYFS
QLLCRPIKLVDSSELLSTLSVDFNGVLHKAYIDVLRNSFGKDLNANMSLAECKRALGLS
ISDHEFTSAISNAHRCDVLLSDLSFNNFVSSYAKPEEKLSAYDLACCMRAGAKVVNAN
VLTKDQTFIVWHAKDFNSLSAEGRKYIVKTSKAKGLTFLLTINENQAVTQIPATSIVA
KQGAGDAGHSLTWLWLLCGLVCLIQFYLCFFMPYFMYDIVSSFEGYDFKYIENGQLKN
FEAPLKCVRNVFENFEDWHYAKFGFTPLNKQSCPIVGVSEIVNTVAGIPSNVYLVGK
TLIFTLQAAFGNAGVCYDIFGVTTPEKCIFTSACTRLEGLGNNVYCYNTALMEGSLP
YSSIQANAYYKYDNGNFIKLPEVIAQGFGRFTVRTIATKYCRVGECVESNAGVCFGFD
KWFVNDGRVANGYVCGTGLWNLVFNILSMFSSSFVAAMSGQILLNCALGAFAIFCCF
LVTKFRRMFGDLSVGVTVVAVLLNNVSYIVTQNLVTMIAYAILYFFATRSLRYAWI
WCAAYLIAYISFAPWWLCAWYFLAMLTGLLPSLLKLKVTNLFEGDKFVGTFESAAAG
TFVIDMRSYEKLANSISPEKLKSYAASYNRYKYYSGNANEADYRCACYAYLAKAMLDF
SRDHNDILYTPPTVSYGSTLQAGLRKMAQPSGFVEKCVVRVCYGNVTNLGLWLGDIVY
CPRHVIASNTTSAIDYDHEYSIMRLHNFSIISGTAFLGVVGATMHGVTLKI KVSQTNM
HTPRHSFRTLKSGEGFNILACYDGCAQGVFGVNMRTNWTIRGSFINGACGSPGYNLKN
GEVEFVYMHQIELGSGSHVGSSFDGVMYGGFEDQPNLQVESANQMLTVNVVAFLYAAI
LNGCTWWLKGEKLFVEHYNEWAQANGFTAMNGEDAFSILAAKTGVCVERLLHAIQVLN
NGFGGKQILGYSSLNDEFSINEVVKQMGFVNLQSGKTTSMFKSISLFAGFVVMFWAEL
FVYTTTIWVNPGLTPFMILLVALSLCLTFVVKHKVFLQVFLLPSSIIVAAIQNCAWD
YHVTKVLAEKFDYNVSVQMIDIQGFVNIFICLFVALLHTWRFAKERCTHWCTYLFSLI
AVLYTALYSYDYVSLLVMLLCAISNEWYIGAIIFRICRFGVAFLPVEYVSYFDGVKTV
LLFYMLLGFBVSCMYGGLLYWINRFCKCTLGVDYFCVSPAEFKYMVANGNLNPNPFD
LFLSFKLMGIGGPRTIKVSTVQSKLTDLKCTNVVLMGILSNMNIASNSKEWAYCVMH
NKINLCDDPETAQELLLALLAFLSKHSDFGGLDLVDSYFENDSILQSVASSFVGMP
FVAYETARQEYENAVANGSSPQIIKQLKKAMNVAKAEFDRESSVQKKINRMAEQAAAA
MYKEARAVNRKSKVVSAMHSLFLGMLRRLDMSSVDTIILNMARNGVVPLSVIPATSAAR
LVVVVPDHDSFVKMMVDGFVHYAGVVWTLQEVKDNDGKNVHLKDVTKENQEILVWPLI
LTCERVVKLQNEIMPGKMKVKATKGEKGGITSEGNALYNNEGGRAFMYAYVTTKPG
MKYVKWEHDSGVVTVELEPPCRFVIDTPTGPQIKYLYFVKNLNNLRRGAVLGYIGATV
RLQAGKQTEFVSNSHLLTHCSFAVDPAAYLDAVKQGAQPVGNCVKMLTNGSGSGQAI
TCTIDSNTTQDTYGGASVCIYCAHVAHPTMDGFCQYKKGWVQVPIGTNDPIRFCLN
TVCKVCGCWLNHGCTCDRTAIQSFDNSYLNVRGSSAARLEPCNGTDIDYCVRAFDVY
NKDASFIGKNLKSNCVRFKNVDKDDAFYIVKRCIKSVMDEQSMYNLLKGCNAVAKHD
FFTWHEGRTIYGNVSRQDLTKYTMDLFCALRNFDEKDCVFEKEILVLTGCCSTDYFE

MKNWFDPIENEDIHRVYAALGKVVANAMLKCVAFCEMVLKGVVGVLTLDNQDLNGNF
YDFGDFVLCPPGMGIPYCTSYYSYMPVVMGMTNCLASECFMKSDIFGQDFKTFDLLKY
DFTEHKEVLFNKYFKYWQDYHPDCVDCHDEMCIHCSNFNTLFATTIPNTAFGPLCR
KVFDIGVPVVATAGYHFKQLGLVWNKDVNTHSTRLTITELLQFVTDPTLIVASSPALV
DKRTVCFVSVAALSTGLTSQTVKPGHFNKEFYDFLRSQGFFDEGSELTLKHFFFTQKGD
AAIKDFDYRYNRPTMLDIGQARVAYQVAARYFDCYEGGCITSREVVTNLNKSAGWP
LNKFGKAGLYYESISYEEQDAIFSLTKRNILPTMTQNLNLKYAISGKERARTVGGVSLL
ATMTTRQFHQKCLKSIVATRNATVVIGTTKFYGGWDNMLKNLMADVDDPKLMGWDPK
CDRAMPSMIRMLSAMILGSKHVTCTASDKFYRLSNELAQVLTENVYSNGGFYFKPGG
TTSGDATTAYANSVFNIFQAVSSNINCVLSVNSSNCNNFENVKKLQRQLYDNCYRNSNV
DESFVDDFYGYLQKHFSMMILSDDSVVCYNKTYAGLGYIADISAFKATLYYQNGVFMS
TAKCWTEEDLSIGPHEFCSQHTMQIVDENGKYYLPYPDPSRIISAGVFVDDITKTDAV
ILLERYVSLAIDAYPLSKHPKPEYRKVFYALLDWVKHLNKTLEGVLESFSVTLLDEH
ESKFWDESFYASMYEKSTVLQAAGLCVVCQSQTVLRCGDCLRRPMLCTKAYDHFVGT
DHKFILAITPYVCNTSGCNVNDVTKLYLGGLNYCVDHKPHLSFPLCSAGNVFGLYKS
SALGSMDIDVFNKLSTSDWSDIRDYKLANDAKESLRLFAAETVKAKEESVKSSYAYAT
LKEIVGPKELLLLWESGKAKPPLNRNSVFTCFQITKDSKFQVGEFVFEKVDYGSDTV
YKSTATTKLVPGMLFILTSHNVAPLRAPTMANQEKYSTIYKLHPSFNVS DAYANLVPY
YQLIGKQRITTIQGPFGSGKSHCSIGIGVYYPGARIVFTACSHAADVSLCAKAVTAYS
VDKCTRIIPARARVECYSGFKPNNSAQYVFSTVNALPEVNADIVVDEVSMCTNYDL
SVINQRISYKHIVYVGDPQQLPAPRVLISKGVMEPIDYNVVTQRMCAIGPDVFLHKCY
RCPAEIVNTVSELVYENKFVPVKEASKQCFKIFERGSVQVDNGSSINRRQLDVVKRFI
HKNSTWSKAVFISPYNSQNYVAARLLGLQTQTVDSAQGSEYDYVIFAQTSDTAHACNA
NRFNVAITRAKKGIFCIMS DRTLFDALKFFEITMTDLQSESSCGLFKDCARNPIDLPP
SHATTYLSLSDRFKTSGDLAVQIGNNNVCTYEHVISYMGFRFDVSMPGSHSLFCTRDF
AMRHVRGWLGM DVEGAHVTGDNVGTNVPLQVGF SNGVDFVAQPEGCVLTNTGSVVKPV
RARAPPGEQFTHIVPLLRKGQPWSVLRKRIVQMIADFLAGSSDVLVFLWAGGLELTT
MRYFVKIGAVKHCQCGTVATCYNVSNDYCCFKHALGCDYVNPYVIDIQQWGYVGS
STNHHAICNVHRNEHVASGDAIMTRCLAVYDCFVKNVDWSITYPMIANENAINKGGRT
VQSHIMRAAIKLYNPKAIHDIGNPKGIRCAVTDKWKYCYDKNPINSNVKLTLEYDYMTH
GQMDGLCLFWNCNVDMPYEF SIVCRFDTRTRSTLNLEGVNGGSLYVNNHAFHTPAYDK
RAMAKLKPAPFFYYDDGSCEVVDQVNYVPLRATNCITKCNIGGAVCSKHANLYRAYV
ESYNIFTQAGFNIWVPTTFDCYNLWQTFTEVNLQGLENI AFNVVNKGSFVGADGELPV
AISGDKVFVRDGN TDNLVFNKTS LPTNIAFELFAKRKVGLTPPLSILKNLGVVATYK
FVLWDYEAERPLTSFTKSVCGYTDFAEDVCTCYDNSIQGSYERFTLSTNAVLF SATAV
KTGGKSLPAIKLNF GMLNGNAIATVKSEDGNIKNINWFVYVRKD GKPVDHYDGFYTQG
RNLQDFLPRSTMEEDFLNMDIGVFIQKYGLEDFNFEHVYVYGDVSKTTLGGLHLLISQV
RLSKMGILKAE EFVAASDITLKCCTV TYLNDPSSKTVCTYMDLLDDFVSVLKSLDLT
VVSKEVHEV IIDNKPWRWMLWCKDNAVATFY PQLSAEWKCGYSMPGIYKTQRMCLEPC
NLYNYGAGLKLPSGIMFNVVKYTQLCQYFNSTTLCVPHNM RVLHLGAGSDYGVAPGTA
VLKRWLPHDAIVVDNDVDYVSDADFSVTGDCATVYLEDKFDLLISDMYDGRTKAIDG
ENVSKEGFFTYINGFICEKLAIGGSIAIKVTEYSWNKKLYELVQRFSFWTMFCTSVNT
SSSEAFVVGINYLGDFAQGPFIDGNI IHANYVFWRNSTVMSLSYNSVLDLSKFNCKHK
ATVVVQLKDS DINEMVLSLVRSGKLLVRGNGKCLSF SNHLVSTK"

mat_peptide 293..625
 /locus_tag="HCoV229Egp1"
 /product="leader protein p9"
 /experiment="experimental evidence, no additional details recorded"
 /note="PL1-PRO cleavage product"
 /protein_id="[NP_835344.1](#)"

mat_peptide 626..2983
 /locus_tag="HCoV229Egp1"
 /product="p87"
 /note="counterpart of MHV p65"
 /protein_id="[NP_835357.1](#)"

mat_peptide 2984..7744
 /locus_tag="HCoV229Egp1"
 /product="coronavirus p195/p210 protein (nsp1)"
 /experiment="experimental evidence, no additional details recorded"
 /note="contains papain-like proteinase 1 domain (PL1-PRO), papain-like proteinase 2 domain (PL2-PRO), predicted phosphoesterase (similar to the Appr-1'-p processing enzyme), and putative metal- and membrane binding Y domain"
 /protein_id="[NP_835345.1](#)"

mat_peptide 7745..9187
 /locus_tag="HCoV229Egp1"
 /product="p44"
 /protein_id="[NP_835358.1](#)"

mat_peptide 9188..10093
 /locus_tag="HCoV229Egp1"
 /product="putative coronavirus nsp2 (3CL-PRO)"
 /protein_id="[NP_835346.1](#)"

mat_peptide 10094..10930
 /locus_tag="HCoV229Egp1"
 /product="putative coronavirus nsp3 (HD2)"
 /note="hydrophobic domain"
 /protein_id="[NP_835347.1](#)"

mat_peptide 10931..11179
 /locus_tag="HCoV229Egp1"
 /product="putative coronavirus nsp4"
 /protein_id="[NP_835348.1](#)"

mat_peptide 11180..11764
 /locus_tag="HCoV229Egp1"
 /product="putative coronavirus nsp5"
 /protein_id="[NP_835349.1](#)"

```

mat_peptide    11765..12091
                /locus_tag="HCoV229Egp1"
                /product="putative coronavirus nsp6"
                /protein_id="NP_835350.1"
mat_peptide    12092..12496
                /locus_tag="HCoV229Egp1"
                /product="formerly known as growth-factor-like protein"
                /note="putative coronavirus nsp7"
                /protein_id="NP_835351.1"
mat_peptide    join(12497..12520,12520..15276)
                /locus_tag="HCoV229Egp1"
                /product="RNA-dependent RNA polymerase"
                /note="putative coronavirus nsp9 (RdRp)"
                /protein_id="NP_835352.1"
mat_peptide    15277..17067
                /locus_tag="HCoV229Egp1"
                /product="putative coronavirus nsp10 (MB, NTPase/HEL)"
                /note="metal-binding domain, NTPase/helicase domain"
                /protein_id="NP_835353.1"
mat_peptide    17068..18621
                /locus_tag="HCoV229Egp1"
                /product="putative coronavirus nsp11"
                /protein_id="NP_835354.1"
mat_peptide    18622..19665
                /locus_tag="HCoV229Egp1"
                /product="putative coronavirus nsp12"
                /protein_id="NP_835355.1"
mat_peptide    19666..20565
                /locus_tag="HCoV229Egp1"
                /product="putative coronavirus nsp13"
                /protein_id="NP_835356.1"
CDS            293..12550
                /locus_tag="HCoV229Egp1"
                /codon_start=1
                /product="replicase polyprotein 1a"
                /protein_id="NP_073550.1"
                /db_xref="GeneID:918764"
                /translation="MACNRVTLAVASDSEISANGCSTIAQAVRRYSEASNGFRACRF
VSLDLQDCIVGIADDTYVMGLHGNQTLFCNIMKFSDRPFMLHGWLVSNSNYLLEEFD
VVFGRGGGNVTYTDQYLCGADGKPMSEDLWQFVDHFGENEIIINGHTYVCAWLTK
RKPLDYKRQNNLAIEEIEYVHGDALHTLRNGSVLEMAKEVKTSSKVVLSDALDKLYKV
FGSPVMTNGSNILEAFTKPVFISALVQCTCGTKSWSVGDWTGFKSSCCNVISNKLCSV
PGNVKPGDAVITTQQAGAGIKYFCGMLTKFVANIEGVSVWRVIALQSVDCFVASSTFV
EEEHVNRMDTFCFNVRNSVTDECRLAMLGAEMTSNVRQVQASGVIDISTGWFDVYDDI

```

FAESKPWFVRKAEDIFGPCWSALASALKQLKVTTGELVRFVKSICNSAVAVVGGTIQI
LASVPEKFLNAFDVFVTAIQTVFDCAVETCTIAGKAFDKVFDYVLLDNALVKLVTTKL
KGVREGLNKVKYATVVVGSTEEVKSSRVERSTAVLTIANNYSKLFDEGYTVVIGDVA
YFVSDGYFRLMASPNSVLTAVYKPLFAFNVNMGTREKFPTTVTCENLESAVLFVN
DKITEFQLDYSIDVIDNEIIVKPNISLCVPLYVRDYVDKWDDFCRQYSNESWFEDDYR
AFISVLDITDAAVKAAESKAFVDTIVPPCPSILKVIDGGKIWNGVIKNVNSVRDWLKS
LKLNLTTQQLLGTCAKRFRWLGILLEAYNAFLDTVVSTVKIGGLTFKTYAFDKPYIV
IRDIVCKVENKTEAEWIELFPHNDRIKSFSTFESAYMPIADPTHFDIEEVELLDAEFV
EPGCGGILAVIDEHVFYKKDGVYYPNGTNIIPVAFTKAAGGKVSFSDDDEVKIDIEPV
YRVKLCFEFEDEKLVDVCEKAIGKKIKHEGDWDSFCKTIQSALSVSVCYNLPTYIY
DEEGGNDLSLPVMISEWPLSVQQAQQEATLPDIAEDVVDQVEEVNSIFDIETVDVKHD
VSPFEMPFEELNGLKILKQLDNNCWVNSVMLQIQLTGILDGDYAMQFFKMGRVAKMIE
RCYTAEQCIRGAMGDVGLCMYRLLKDLHTGFMVMDYKCSCTSGRLEESGAVLFCTPTK
KAFPYGTCLNCNAPRMCTIRQLQGTIIIFVQQKPEPVNPVSFVVKPVCSSIFRGAVSCG
HYQTNIYSQNLCDVGFGVNKIQPWTNDALNTICIKDADYNAKVEISVTPIKNTVDTTP
KEEFVVKEKLNAFLVHDNVAFYQGDVDTVNVGVDFDFIVNAAENLAHGGGLAKALDV
YTKGKLQRLSKEHIGLAGKVKGVTGVMVECDLRIFNVVGPRKGKHERDLLIKAYNTI
NNEQGTPLTPILSCGIFGIKLETSLEVLLDVCNTKEVKVFVYTDTEVCKVKDFVSGLV
NVQKVEQPKIEPKPVSIVKVAPKPYRVDGKFSYFTEDLLCVADDKPIVLFTDSMLTLD
DRGLALDNALSGVLSAAIKDCVDINKAIPSGNLIKFDIGSVVVMCVVPSEKDKHLDN
NVQRCTRKLNRMLCDIVCTIPADYILPLVLSSLTNCVSVFVGLKAAEAKVITIKVTE
GVNVHDTVTTDKSFEQQVGVIADKDKDLGAVPSDLNTSELLTKAIDVDWVEFYGFK
DAVTFATVDHSAFAYESAVVNGIRVLKTSNNCWVNAVCIALQYSKPHFISQGLDAAW
NKFVLGDVEIFVAFVYYVARLMKGDKGDAEDTLTKLSKYLANEAQVQLEHYSSCVECD
AKFKNSVASINSAIVCASVKRDGVQVGVCVHGKIYYSRVSVRGRAIIVSVEQLEPCA
QSRLLSGVAYTAFSGPVDKGHYTVYDTAKKSMYDGRFVKHDLSSLSVTSVVMVGGYV
APVNTVKPKPVINQLDEKAQKFFDFGDFLIHNFVIFFTWLLSMFTLCKTAVTTGDVKI
MAKAPQRTGVVLKRSCLKYNLKASAAVLKSKWLLAKFTKLLLLLIYTLYSVLLCVRFG
PFNFCSETVNGYAKSNFVKDDYCDGSLGCKMCLFGYQELSQFSHLDVVWKHITDPLFS
NMQPFIVMVLLLIIFGDNYLRCLLYFVAQMISTVGVLFGYKETNWFLHFIPFDVICDE
LLVTVIVIKVISFVRHVLFGCENPDCIACSKSARLKRFPVNTIVNGVQRSFYVNANGG
SKFCKKHRFFCVDSDSYGYGSTFITPEVSRELGNITKTNVQPTGPAYVMIDKVEFENG
FYRLYSCETFWRYNFDITESKYSCKEVFKNCNVLDDFIVFNNGTNTVQVKNASVYFS
QLLCRPIKLVDSSELLSTLSVDFNGVLHKAYIDVLRNSFGKDLNANMSLAECKRALGLS
ISDHEFTSAISNAHRCDVLLSDLSFNNFVSSYAKPEEKLSAYDLACCMRAGAKVVNAN
VLTKDQTPIVWHAKDFNSLSAEGRKYIVKTSKAKGLTFLLTINENQAVTQIPATSIVA
KQGAGDAGHSLTWLWLLCGLVCLIQFYLCFFMPYFMYDIVSSFEGYDFKYIENGQLKN
FEAPLKCVRNVFENFEDWHYAKFGFTPLNKQSCPIVGVSEIVNTVAGIPSNVYLVGK
TLIFTLQAAFGNAGVCYDIFGVTTPEKCIFTSACTRLEGLGGNNVYCYNTALMEGSLP
YSSIQANAYYKYDNGNFIKLPEVIAQGGFRTVRTIATKYCRVGECVESNAGVCFGFD
KWFVNDGRVANGYVCGTGLWNLVFNILSMFSSSFVAAMSGQILLNCALGAFAIFCCF
LVTKFRRMFGDLSVGVCTVVAVLLNNVSYIVTQNLVTMIAYAILYFFATRSLRYAWI
WCAAYLIAYISFAPWWLCAWYFLAMLTGLLPSLLKLVSTNLFEGDKFVGTFESAAAG
TFVIDMRSYEKLANSISPEKLKSYAASYNRYKYYSGNANEADYRCACYAYLAKAMLDF

SRDHNDILYTPPTVSYGSTLQAGLRKMAQPSGFVEKCVVRVCYGNTVLNGLWLGDIVY
CPRHVIASNTTSAIDYDHEYSIMRLHNFSIISGTAFLGVVGATMHGVTLKIKVSQTNM
HPRHSFRTLKSGEGFNILACYDGCAQGVFGVNMRTNWTIRGSFINGACGSPGYNLKN
GEVEFVYMHQIELGSGSHVGSSFDGVMYGGFEDQPNLQVESANQMLTVNVVAFLYAAI
LNGCTWWLKGEKLFVEHYNEWAQANGFTAMNGEDAFSILAAKTGVCVERLLHAIQVLN
NGFGGKQILGYSSLNDEFSINEVVKQMFVNLQSGKTTSMFKSISLFAGFFVMFWAEL
FVYTTTIWVNPGLTPFMILLVALSLCLTFVVKHKVLFQLQVFLLPSSIIVAAIQNCWD
YHVTKVLAEKFDYNVSVMQMDIQGFVNIFICLFVALLHTWRFAKERCTHWCTYLFSLI
AVLYTALYSYDYVSLVMLLCAISNEWYIGAIIFRICRFGVAFLPVEYVSFYDGVKTV
LLFYMLLGFVSCMYGGLLYWINRFCKCTLGVYDFCVSPAEFKYMVANGLNAPNGPFDA
LFLSFKLMGIGGPRTIKVSTVQSKLTDLKCTNVVLMGILSNMNIASNSKEWAYCVMH
NKINLCDDPETAQELLLALLAFFLSKHSDFGLGDLVDSYFENDSILQSVASSFVGMP
FVAYETARQEYENAVANGSSPQIIKQLKKAMNVAKEFDRESSVQKKINRMAEQAAAA
MYKEARAVNRKSKVVSAMHSLFLGMLRRLDMSSVDTIILNMARNGVVPLSVIPATSAAR
LVVVVPDHDSFVKMMVDGFVHYAGVVWTLQEVKDNDGKNVHLKDVTENQEILVWPLI
LTCERVVKLQNNIMPCKMKVKATKGEKGGITSEGNALYNNEGGRAFMYAYVTTKPG
MKYVKWEHDSGVVTVLEPPCRFVIDTPTGPQIKYLYFVKNLNNLRGAVLGYIGATV
RLQAGKQTEFVSNSHLLTHCSFAVDPAAYLDVAVKQGAQKPGVNCVKMLTNGSGSGQAI
TCTIDSNTTQDTYGGASVCIYCAHVAHPTMDGFCQYKKGWVQVPIGTNDPIRFCLN
TVCKVCGCWLNHGCTCDRTAIQSFDNSYLNESGALVPLD"

gene

20570..24091
/gene="S"
/locus_tag="HCoV229Egp2"
/db_xref="GeneID:918758"

CDS

20570..24091
/gene="S"
/locus_tag="HCoV229Egp2"
/note="structural protein"
/codon_start=1
/product="surface glycoprotein"
/protein_id="NP_073551.1"
/db_xref="GeneID:918758"
/translation="MFVLLVAYALLHIAGCQTTNGLNTSYSVCNGCVGYSENVFAVES
GGYIPSDFAFNNWFLLTNTSSVVDGVVRSFQPLLLNCLWSVSGLRFTTGFFVYFNGTGR
GDCKGFSSDVLSDVIRYNLFEEENLRRTILFKTSYGVVVFYCTNNTLVSGDAHIFPG
TVLGNFYCFVNTTIGNETTSFVVGALPKTVREFVISRTGHFYINGRYFTLGNVEAVN
FNVTTAETTDFTVALASYADVLVNVSQTSIANIIYCNSVINRLRCDQLSFDVPDGFY
STSPIQSVELPVSIIVSLPVYHKHTFIVLYVDFKPKQSGGGKCFNCYPAGVNITLANFNE
TKGPLCVDTSHTTKYVAVYANVGRWSASINTGNCPFSGKVN NFVKFGSVCFSCLKDI
PGGCAMPIVANWAYSKEYTIGSLYVSWSDGDGITGVPQPVEGVSSFMNVTLDKCTKYN
IYDVSGVGIVRVSNDTFLNGITYTSTSGNLLGFKDVTGTIYSITPCNPDPQLVYVQQ
AVVGAMLSENFYSYGFNSNVVELPKFFYASNGTYNCTDAVLTYSSFGVCADGSIIAVQP
RNVSYDSVSAIVTANLSIPSNWTTSVQVEYLQITSTPIVVDCTYVCNGNVRCVELLK
QYTSACKTIEDALRNSARLESADVSEMLTFDKKAFTLANVSSFGDYNLSSVIPSLPTS

```

GSRVAGRSAIEDILFSKLVTSGLGTVDADYKKCTKGLSIADLACAQYYNGIMVLPGVA
DAERMAMYTGSLIGGIALGGLTSAVSIPFSLAIQARLNYVALQTDVLQENQKILAASF
NKAMTNIVDAFTGVNDAITQTSQALQTVATALNKIQDVVNQQGNSLNHLTSQLRQNFQ
AISSSIQAIYDRLDTIQADQQVDRDLITGRLAALNVFVSHLTLTKEYTEVRASRQLAQQKV
NECVKSQSKRYGFCGNGTHIFSIVNAAPEGLVFLHTVLLPTQYKDVEAWSGLCVDGTN
GYVLRQPNLALYKEGNYRITSRIMFEPRIPTMADFVQIENCNVTFVNISRSELQTIV
PEYIDVNKTLQELSYKLPNYTVPDLVVEQYNQTTILNLTSEISTLENKSAELNYTVQKL
QTLIDNINSTLVDLKWLNRVETIYKWPWWVWLCISVVLIFVVSMLLLCCCSTGCCGFF
SCFASSIRGCCESTKLPYYDVEKIHQ"

gene      24091..24492
          /gene="4a"
          /locus_tag="HCoV229Egp3"
          /db_xref="GeneID:918759"

CDS       24091..24492
          /gene="4a"
          /locus_tag="HCoV229Egp3"
          /codon_start=1
          /product="4a protein"
          /protein_id="NP_073552.1"
          /db_xref="GeneID:918759"
          /translation="MALGLFTLQLVSAVNQSLNNAKVSQVSRQVIQDVKGDTVTFNL
LAYTLMSLFVVYFALFKARSHRGRAALIVFKILILFVYVPLLYWSQAYIYATLIAVIL
LGRFFHTAWHCWLYKTWDFIVFNVTTLCYAR"

gene      24482..24748
          /gene="4b"
          /locus_tag="HCoV229Egp4"
          /db_xref="GeneID:918760"

CDS       24482..24748
          /gene="4b"
          /locus_tag="HCoV229Egp4"
          /codon_start=1
          /product="4b protein"
          /protein_id="NP_073553.1"
          /db_xref="GeneID:918760"
          /translation="MQGKCWFLENKALKPFVCFYGGDQFLYIGDRIVSYFSTNDLYVA
LRGRIDKDLSLSRKVELYNGECVYLFCEHPAVGIVNTDFKLEIH"

gene      24750..24983
          /gene="E"
          /locus_tag="HCoV229Egp5"
          /db_xref="GeneID:918761"

CDS       24750..24983
          /gene="E"
          /locus_tag="HCoV229Egp5"
          /note="structural protein; E protein"

```

	/codon_start=1 /product="envelope protein" /protein_id="NP_073554.1" /db_xref="GeneID:918761" /translation="MFLKLVDHALVNVLLWCVVLLVILLVCITIIKLIKLCFTCHMFCNRTVYGPIKNVYHIYQSYMHIDPFPKRVIDF"
gene	24995..25672 /gene="M" /locus_tag="HCoV229Egp6" /db_xref="GeneID:918762"
CDS	24995..25672 /gene="M" /locus_tag="HCoV229Egp6" /note="structural protein" /codon_start=1 /product="membrane protein" /protein_id="NP_073555.1" /db_xref="GeneID:918762" /translation="MSNDNCTGDIVTHLKNWNFGWNVILTIFIVILQFGHYKYSRLFYGLKMLVLWLLWPLVLALSIFDTWANWDSNWAFAVSFFMAVSTLVMWVMYFANSFRLFRRTTFWAWNPEVNAITVTTVLGQTYYPQPIQQAPTGITVTLISGVLYVDGHRLASGVQVHNLP EYMTVAVPSTTIIYSRVGRSVNSQNSTGWVFYVRVKHGDFS AVSSPMSNMTE NERLLHFF"
gene	25686..26855 /gene="N" /locus_tag="HCoV229Egp7" /db_xref="GeneID:918763"
CDS	25686..26855 /gene="N" /locus_tag="HCoV229Egp7" /note="structural protein" /codon_start=1 /product="nucleocapsid protein" /protein_id="NP_073556.1" /db_xref="GeneID:918763" /translation="MATVKWADASEPQRGRQGRIPYSLYSPLLVDSEQPWKVIPRNLV PINKKDKNKLIGYWNVQKRFRTRKGRVDLSPKLHFYYLGTGPHKDAKFRERVEGVWVAVDGAKTEPTGYGVRKRNSEPEIPHFNQKL PNGVTVVEEPDSRAPSRSQSRSGR GESKPQSRNPSSDRNHNSQDDIMKAVAAALKSLGFDKPQEKDKKSAKTGTPKPSRNQS PASSQTS AKSLARSQSSE TKEQKHMQPRWKRQPNDDVTSNVTQCFGPRDL DHNFGS AGVVANGVKAKGY PQFAELVPSTAAMLFDSHIVSKESGNTVVLTFTRTVTPKDHPLH GKFL EELNAFTREMQQHP LLNPSALEFNPSQTS PATAEPVRDEVSIETDIIDEVN"
3'UTR	26856..27317
ORIGIN	

1 acttaagtac cttatctatc tacagataga aaagttgctt tttagacttt gtgtctactt
61 ttctcaacta aacgaaatth ttgctatggc cggcatctth gatgctggag tctagtgtga
121 attgaaatth catttggtt gcaacagtht ggaagcaagt gctgtgtgtc ctagtctaag
181 ggtttctgtg tccgtcacga gattccattc tacaaacgcc ttactcgagg ttccgtctcg
241 tgtttgtgtg gaagcaaagt tctgtcttht tggaaaccag taactgttcc taatggcctg
301 caaccgtgtg acacttgccg tagcaagtga ttctgaaatt tctgcaaatt gctgttctac
361 tattgcgcaa gccgtccgcc gttatagcga ggccgctagc aatggthttt gggcatgccg
421 atttgtthca ttagattthc aggattgcat cgttggcatt gcagacgata catatgttat
481 gggctctgcat ggcaatcaga cgttgttht caacataatt aaatthtctg accgtcctth
541 tatgcttcat ggggtggttg tthtttccaa ttcaaattac cthttggagg aatthgatgt
601 tgtcttcggt aagagaggtg gtggaattgt gacatacact gaccagtatc tctgtggcgc
661 cgatggcaaa cctgttatga gtgaagattt atggcagtht gttgaccatt tccgtgagaa
721 cgaagaaatt atcatcaatt gtcatactta cgttctgtgt tggcttacta agcgtaaagg
781 cttagattac aaacgtcaga acaacctgac cattgaagag attgaattat tgcattgtga
841 tgctthtcat acactacgca atggttctgt tcttgaaatt gctaaggagg tgaagacatc
901 tagtaaatgt gtgttaagcg atgctcttga caaacctthc aaagtcttht gttctctgt
961 tatgacaaat ggttccaaca tcttagaggc cthtactaaa cctgtgttht ttagtgatt
1021 agttcaattgt acttgtggtt ccaagtcttg gtctgttggt gattggaccg gthttaaatc
1081 ctcttgtthc aacgtgatca gtaataaact gtgtgttggt cccggaattt taaacctgg
1141 tgatgctgtg attaccactc agcaagctgg tgctgggtatt aagtatthtt gtggcatgac
1201 tcttaagtht gttgcaaata ttgaaggtgt ctctgttht agagtgttg ctcttcagag
1261 tgtggattgc thtgttgctt cttccactth tgtagaagag gaacattgta atagaatgga
1321 tacattctgc ttcaattgtc gcaatagtgt tactgatgag tgcgtctgg ccatgttggg
1381 tgctgaaatt actagtaatt tcagaagaca agttgcttca ggtgtcatag acattagtac
1441 cggttggtth gatgtthatt atgacatctt tgctgaaagg aaacctggt thgttcgcaa
1501 ggctgaagac atthttggcc cttgttggtc cgctcttgct tctgacttht aacaacttht
1561 agtccactaca ggtgaacttg tgagatttht taagtctatt tgcaattcag ctgttgctgt
1621 cgtgggtggt actatacaaa thtctgctag tgtgcctgag aagtthttga atgcgtthga
1681 cgtgtthtgc acagctattc aaactgtctt tgactgtgct gttgaaactt gtactattgc
1741 cggtaaagca thtgacaagg thtttgacta tgtthtgcct gataatgcgc thgtaaaact
1801 tgtcaccaca aagcttaagg gtgttcgtga acgtggcctt aataaagtta agtatgcaac
1861 agttgttggt ggttccactg aagaagtht atcttcacgt gttgaacgta gcaactgtgt
1921 acttacaatc gccacaatt atccaaact thttgatgaa gggataactg thgtaatgg
1981 cgatgtggcg tactthttht gtgacggcta cttccgtctt atggccagtc caaatagtgt
2041 gttgactact gcagtctata aaccattgtt tgctthttht gtgaattgta tgggtactag
2101 acctgaaaaa thtccaacca ctgtgacttg tgaaaattth agtctgtctg thttgttht
2161 taatgacaaa attactgaat tccaattgga ttacagtatt gatgtcattg ataataaat
2221 aattgtcaaa cctaatatca gcctatgtgt tccacttht gtgagagact atgttgacaa
2281 atgggatgat thttgcagac aatatagtta cgagtcttg thtgaggatg attacagggc
2341 thttatcagt gthttggaca tcaactgatc tgctgtgaaa gctgcagagt ctaaagctth
2401 cgttgatact attgttccac cttgccatc thttthgaaa gttatagatg gaggcaaat
2461 atggaatggt gttattaaaa atgttaactc tgttagagac tggcttaagt cthtgaaagt
2521 aaatctcaca caacagggtt tgcttggaac atgtgcaaag cgtthttht gttggcttg
2581 cattthtcta gaggcctata atgcgtthtt agacactgtg gthtctactg thaaaattgg

2641 tggcttgacc tttaaaacat atgcttttga taaaccttac attgtgatac gtgatatcgt
2701 gtgtaagggt gaaaataaaa cagaagcaga atggattgag ctttttccac ataatgacag
2761 gattaagtct tttagtactt tcgagagtgc ttacatgcc aatgcagacc ctacacattt
2821 tgacattgaa gaagttgaac ttttagatgc agagtttgta gaaccaggct gtggtggtat
2881 tttggcagta atagatgagc acgtctttta taagaaggat ggtggttatt atccatcaaa
2941 tggactaac attctacctg ttgcatttac aaaagccgct ggtggtaaaag tttcattttc
3001 tgatgacgtt gaagtaaaag acattgaacc tgtttacaga gtcaagcttt gctttgagtt
3061 tgaagatgaa aaacttgtag atgtttgtga aaaggcaatt ggcaagaaaa ttaaacadga
3121 aggtgactgg gatagctttt gtaagactat tcaatcagca ctttctgttg tttcttgcta
3181 tgtaaactta cctacttatt acatttatga tgaagaaggc ggtaatgact tgagtttgcc
3241 cgttatgatt tctgaatggc ctctttctgt tcaacaagct caacaagaag ctactttacc
3301 tgatattgct gaggatgttg ttgaccaagt tgaagaagtc aatagcattt ttgacattga
3361 gacagtggat gttaaacadg atgtgagtc ttttgaaatg ccatttgaa agttaaatgg
3421 tttaaagata ctcaaacaat tggataacaa ctgctgggtt aactcagtta tgttacaaat
3481 acaattaact ggtatacttg atgggtgacta tgctatgcag ttttttaaaa tgggccgagt
3541 tgccaagatg attgaacgct gctacactgc tgagcaatgt atacgtgggtg ctatgggtga
3601 tgttggtttg tgtatgtata gactgcttaa agacttacac actgggtttta tgggttatga
3661 ttataaatgt agttgtacca gtggtaggct tgaagaatcg ggagctgttt tgtttgtac
3721 gcccactaag aaggcgtttc cttatggtag ttgtctaaat tgtaacgcac ctgcgatgtg
3781 tacaattagg cagttacaag gtaccataat atttgtgcaa caaaaaccag aacctgttaa
3841 tcctgtttct tttgttggtt aaaccagtctg ctcatcaatt tttcgtgggtg ctgtgtcttg
3901 tggtcattac cagactaaca tctattcaca aaatttgtgt gtggatgggt ttggtgttaa
3961 caagattcag ccctggacaa atgatgcact taatactatt tgtattaagg atgcagatta
4021 taatgcaaaa gttgaaatat ctgttacacc aattaaaaat acagttgata caacacctaa
4081 ggaagaattt gttgttaaag agaagttgaa cgccttcctc gttcatgaca atgtagcttt
4141 ctaccaaggt gatgttgata ctgttggtta tgggtgttgac tttgacttta ttgtaaagtc
4201 tgctaattgag aaccttgctc atgggtggagg acttgccaaa gcttttagatg tgtacactaa
4261 aggtaaactt caacgtttat ctaaagaaca cattggatta gcgggtaaaag taaaagttgg
4321 tacaggagtt atgggtgagt gtgatagcct tagaattttt aatgttggtg gtccacgcaa
4381 gggtaaacat gaacgtgatt tactcataaa agcttacaac actattaata atgaacaagg
4441 cacaccttta acaccaattt tgagctgttg tatttttgggt atcaaactcg aaacttcatt
4501 agaagttttg cttgatgttt gtaatacaaa agaagttaa gtttttgttt atacagacac
4561 agaggtttgt aaggtttaag attttgtgtc tggtttagtg aatgttcaaa aagttgagca
4621 acctaaaata gaacccaaac cagtgtccgt aattaaagtt gcacccaagc cttacagggt
4681 agatggtaaa tttagttact ttacagaaga cttgttgtgt gtcgctgatg acaaacccat
4741 tgttttgttt actgactcta tgcttacttt ggatgaccgt ggtttagctc tagacaatgc
4801 acttagtggt gtgcttagtg ctgctattaa ggattgtgtt gacataaata aagctatacc
4861 ttctggtaat cttattaagt ttgatatagg ttctgttgtt gtctacatgt gtgttggtcc
4921 atccgaaaag gacaaacatt tagataataa tgttcaacga tgcacacgta agttgaatag
4981 acttatgtgt gatatagttt gtactatacc agctgactac atcttgccat tgggtgtgtc
5041 tagtttgact tgtaatgttt cttttgttag tgaacttaaa gctgctgaag ctaaagttat
5101 aactataaag gtgacagagg atgggtgttaa tgttcatgat gtgaccgtga caacagacaa
5161 gtcatttgaa caacaagttg gtgttattgc tgataaggac aaagatcttt ctggtgcagt
5221 accaagtgat cttaacacat ctgaattgct tactaaagca atagatgttg attgggtcga

5281 attttatggc tttaaagatg ctgttacttt tgcaacagtt gatcatagtg cttttgccta
5341 tgaaagtgcgt gttgttaatg gtattagagt gttaaaaact agtgataata attggttgggt
5401 gaatgctggt tgtattgcac tacagtattc gaaaccccat tttatttcac aaggctcttga
5461 tgctgcgtgg aataaatttg ttttaggcga tgttgaaatt tttgttgcat ttgtttacta
5521 tgttgcaaga ctaatgaaag gtgacaaggg tgatgctgaa gacactttga ctaagttgtc
5581 taagtatctt gctaataag ctcaagttca attagaacat tatagttctt gtgttgaatg
5641 tgatgctaaa tttaaaaact ctgttgcatc tatcaattct gctatagttt gtgctagtgt
5701 caaacgtgat ggtgtgcaag ttggttattg tgtccatggt attaagtact attcacgtgt
5761 tagaagtgtt agaggtagag ctattatagt cagtgtcgaa cagcttgaac cgtgtgctca
5821 gtctagactt ttgagtgggt ttgcttatac tgctttttct ggacctgttg acaaagggtca
5881 ttatactggt tatgatactg caaagaaatc aatgtatgat ggtgatcgtt ttgttaaaca
5941 tgatctttct ctgctgtctg tcacatcagt tgttatgggt ggtggttatg ttgcacctgt
6001 taatacagtg aaacctaaac cagtcattaa tcaacttgat gaaaaggcac agaagttctt
6061 tgattttgggt gattttttga ttcataattt tgttattttt ttcacatggt tattgagtat
6121 gtttactttg tgtaaaactg cagtaactac aggtgatgtt aaaataatgg ccaaagcacc
6181 acaaaggacg ggtgttgttt taaaacgtag tcttaaatat aacttaaaag cgtcagcagc
6241 tgttcttaaa tctaagtgggt ggctgcttgc taagtttacg aaactactgt tactcatata
6301 tacattgtac tcagtagttt tgctttgtgt acgttttgga ccgtttaatt tttgtagtga
6361 gactgttaat ggttatgcta agtcaaactt tgtcaaggat gattactgtg atggttcatt
6421 gggctgcaag atgtgtcttt ttggttacca agagttaagt caatttagcc atttggatgt
6481 tgtgtggaag catataacag accctttgtt tagtaatatg caacctttca ttgtcatggt
6541 tttgctgctt atatttgggt acaattattt gagatgcttc ttgctgtatt ttgttgctca
6601 gatgataagc acagttgggt tttttctagg ttacaaggaa acaaattggt tcttgcaactt
6661 tattccattt gatgttattt gtgatgaact gcttgtcact gttattgtta ttaaggttat
6721 ttcttttgtc agacatgtgc tttttggttg tgaaaacca gattgtattg cgtgttctaa
6781 gagtgttaga cttaagagat tccctgttaa cacaattgtc aatggtgtgc aacgttcatt
6841 ttatgttaat gcaaatgggt gtagtaagtt ttgtaagaaa catagatttt tctgtgttga
6901 ttgtgactct tatggttatg gcagcacgtt tataacaccc gaagtttcta gagaacttgg
6961 taacattacc aaaacaaatg tgcaaccaac agggccggcc tatgtcatga ttgacaaagt
7021 ggagtttgaa aatggttttt acagattgta ttctgtgtaa acattttggc gttacaactt
7081 tgatataact gaaagcaagt attcttgcaa agaggttttt aaaaattgta atgttttggg
7141 tgatttcacg gtgtttaaca ataatgggac caatgtaacg cagggttaaaa atgctagtgt
7201 ttacttttca cagttgttgt gtaggcccat taaattagtt gacagtgaac ttttgtccac
7261 tttgtcagtt gattttaatg gtgtcttaca caaggcatac attgatgtac tacgtaatag
7321 ctttggtaaa gatcttaatg ctaatatgtc tttagccgag tgcaagagag ctttaggcct
7381 gtctattagt gatcatgaat ttactagtgc tatttctaata gcacatcgtt gtgacgtgtt
7441 gttatctgat ttgtcattta acaactttgt cagttcgtat gctaaacctg agggaaaatt
7501 atcagcttat gacttggcgt gttgtatgcg tgcaggtgct aaggttggtta atgccaatgt
7561 tctgacaaag gaccaaactc ctattgtttg gcatgcaaag gattttaaca gtctttctgc
7621 tgaaggctcg aagtatattg taaaaactag caaagctaag ggtttgactt tcttgttgac
7681 aattaatgaa aaccaagctg tcacgcaaat acctgcaact agcattgttg ctaagcaagg
7741 tgctgggtgat gctggccatt cattaacatg gctgtggcta ctgtgtggtc ttgtgtgttt
7801 gattcaattc tacttgtgct ttttcatgcc ctattttatg tacgatatcg tgagtagttt
7861 tgagggttat gattttaagt atatagaaaa tggtcagttg aagaattttg aagcgccact

7921 taaatgcgtc agaaacgttt ttgaaaactt tgaggactgg cattatgcta agtttggtt
7981 cacaccttta aacaagcaaa gctgtcctat tgtagttgga gtttctgaaa ttgttaatac
8041 tgtcgctggc attccatcta atgtgtatct tgttggtaaa actttaattt ttacactaca
8101 agctgctttt ggtaatgctg gtgtttgtta tgacattttt ggagtcacaa cacctgaaaa
8161 gtgcattttt acttctgctt gtactagatt agaaggtttg ggtggtaaca atgtttattg
8221 ttataacaca gcgcttatgg aaggttcttt gccttacagt tcaatacaag ctaatgcata
8281 ttataaatat gacaatggca attttattaa gttgccagaa gttattgcac aaggcttttg
8341 ttttagaaca gtgcgtacta ttgccaccaa atactgccgc gtaggtgaat gtgttgaatc
8401 caatgcaggt gtgtgttttg gctttgacaa gtggtttggt aacgatggac gtgttgccaa
8461 tggttacgtt tgtggtactg gtttgtggaa ccttgatttt aacatacttt ccatgttttc
8521 atcttcattc tctgttgctg caatgtcagg tcaaatttta ctttaattgtg cattaggtgc
8581 ttttgctatt ttttgttgtt ttcttgtagc aaagtttaga cgcattgttg gtgaccttc
8641 tgtaggtgtt tgcactgttg ttgtggctgt tttgcttaac aatgtctctt acattgtaac
8701 tcagaattta gtaacaatga ttgcttatgc catattgtat ttctttgcta ctagaagctt
8761 acgctatgca tggatttgggt gtgctgcata tttaattgcg tatatttctt ttgctccatg
8821 gtggttgtgt gcttggtact ttcttgctat gttgacaggt ttgttaccta gtttgctgaa
8881 gcttaaagtt tcgacaaatc ttttcgaagg tgacaaattt gtaggtacat ttgaaagtgc
8941 tgctgcagga acatttgtca ttgacatgcg ttcttatgag aaacttgcta atagcatctc
9001 tccagaaaag ttgaaaagtt atgctgctag ctataataga tataagtact atagtggtaa
9061 tgcaaatgaa gctgattacc gttgcgcttg ttatgcctat ttagcaaaag caatgttgga
9121 cttttcgcgt gatcataatg acatcttgta cacacctccg actgtcagtt atggttctac
9181 attacaggct ggtttgcgca aaatggcaca accatctggc tttgtggaga aatgtgttgt
9241 ccgtgtctgc tatggaaaca ctgtgttgaa tgggttgttg cttggtgata ttgtttattg
9301 cccacgtcat gttatcgcat ctaacacaac ttctgctata gattatgatc acgaatatag
9361 tattatgcgg ttgcataatt tttctataat atctggtaca gcatttcttg gtgtttagg
9421 tgctactatg catggagtaa ctcttaaaat taaggtttca cagactaaca tgcacacacc
9481 tagacattct tttagaacac taaaatctgg tgaaggtttt aacatcttag catgctatga
9541 tggttgtgct caagggtttt ttggtgtgaa catgagaact aattggacta tccgtggttc
9601 atttattaat ggtgcgtgtg gttcccctgg ctacaatctt aaaaatggcg aggtggaatt
9661 tgtttatatg catcaaattg aactcggaag tggtagccat gtaggttcta gctttgatgg
9721 tgttatgtat ggtggttttg aagaccaacc taatcttcaa gttgaatctg caaaccagat
9781 gttaacagtt aatgtggttg catttcttta tgctgctata ttgaatggtt gcacatggtg
9841 gcttaaaggt gaaaaattgt ttgtggagca ttataatgag tgggcacagg ctaatggttt
9901 cacagctatg aatggtgaag acgctttttc cattcttgct gctaaaactg gtgtctgtgt
9961 ggaaagatta cttcatgcta ttcaagtttt gaataatggc tttggtggta aacaaatttt
10021 gggttattct agtctcaatg atgagttcag tattaatgaa gttgtcaaac aaatgtttgg
10081 tgtaaacctg caaagtggta aaaccactag tatgtttaaa tccataagct tatttgctgg
10141 cttctttgtc atgttctggg ctgaattatt tgtttatacc accactatth gggtaaccc
10201 tggttttctt actccgttta tgattttgct tgttgctttg tcaactctgtc ttacatttgt
10261 tgtaaacat aaggttttgt ttttgcaagt gtttttggtt ccttcaatta ttgtggctgc
10321 tattcaaaac tgtgcttggg actaccatgt taaaaaggtg ttggcagaga agtttgatta
10381 taatgtttct gttatgcaaa tggacatcca gggttttgtt aacattttta tttgtctttt
10441 tgttgcactg ttgcatactt ggcgctttgc taaagagcgt tgtacacatt ggtgcactta
10501 tttgttctca ctcatgtctg ttttatacac tgcattgtat agttatgact acgttagttt

10561 gctggttatg ctactttgtg caatttctaa tgaatggtat attggtgcta ttatTTTTtag
10621 aatttgtcgt tttggtgttg catttttacc agtggaatac gtgtcttact ttgatgggtg
10681 taaaactgtg ctgttggtttt acatgttggtt aggctttgtt agctgtatgt actatgggtt
10741 gttgtactgg attaacaggt tctgtaagtg cacattaggt gtttatgatt tctgtgttag
10801 tccagccgaa ttttaagtata tgggttgctaa tggtttgaat gcaccaaagt gcccttttga
10861 tgcgctcttt ctgtctttta aactaatggg tattggcggg cctagaacca ttaaagtttc
10921 tactgtacag tctaaattga ctgatcttaa gtgcacaaac gtcgttctaa tgggcatttt
10981 gtctaactg aacatagctt ctaattcaaa ggagtgggca tattgtgttg aaatgcacaa
11041 taaaataaac ttgtgtgacg accctgaaac tgctcaagag ttattgctgg cgttggtggc
11101 ctttttcttg tctaagcata gtgattttgg tcttggtgat cttgtcgatt cttatTTTTga
11161 gaacgactcc attttgcaaa gtgttgcatc ttcttttgtt ggtatgccat cttttgttgc
11221 atatgaaaca gcaagacaag agtatgaaaa tgctgttgca aatgggtcct caccacaaat
11281 aatcaaacia ttgaagaagg ctatgaatgt tgcaaaagct gagtttgaca gggaatcatc
11341 tgttcaaaag aaaattaaca gaatggctga acaagctgct gcagctatgt acaagaagc
11401 acgtgctggt aatagaaaat caaaagttgt tagtgccatg catagtttac tctttggcat
11461 gctccgacgt ttggacatgt ctagtgttga cactatcctt aatatggcac gtaatgggtg
11521 tgtccctctt tccgttatcc ctgctacttc tgcagccagg ctgctcgctg tagtaccaga
11581 tcatgattca tttgtgaaaa tgatggtaga tggttttgtg cactacgctg gtgttgtttg
11641 gacattacag gaagttaagg ataatgatgg taagaatgtg catcttaaag atgttacaaa
11701 ggaaaaccag gaaataactt tttggcctct gatTTtgact tgtgaacgtg tcgttaaatt
11761 gcagaacaat gaaataatgc cgggcaagat gaagggtcaag gccaccaaag gtgaagggtga
11821 tggaggcatt actagtgaag gtaatgctct atacaacaat gaagggtggac gtgcattcat
11881 gtatgcatat gtgactacga agcctggcat gaagtatgtt aaatgggaac atgactctgg
11941 tgtggttaca gttgaattgg aaccaccttg cagatttgtt atagacacac ctactggacc
12001 ccaaattaag tatctttatt ttgttaagaa tcttaacaat ttaaggagag gtgctgtttt
12061 gggttacatt ggtgccactg tgagattgca agctggcaaa cagactgagt ttgtttcaaa
12121 ctccatttta ttaacacatt gttcttttgc tgttgacca gctgcagcct atcttgatgc
12181 tgttaaaaaa ggcgcaaac ctgttggtgaa ttgtgtaaag atgttgacta atggttctgg
12241 tagcggctag gctattactt gtaccattga ttccaacact acgcaggaca catatgggtg
12301 cgcgtctgtt tgtattttatt gcagagcaca tgttgacat ccaaccatgg acggtttttg
12361 tcagtacaaa ggcaagtggg tacaagtgcc tataggtaca aatgacccta taagattttg
12421 tcttgaaaat actgtttgta aagtttgtgg ttgttggtt aatcatggct gtacatgtga
12481 ccggaactgt atccaaagt ttgataacag ttattttaaag gagtccgggg ctctagtgcc
12541 gctcgactag agccctgtaa tggtagacag atagattact gtgtccgtgc atttgacgtt
12601 tacaataaag atgcgtcttt tatcggaaaa aatctgaagt ccaattgtgt gcgcttcaag
12661 aatgtagata aggatgacgc gttctatatt gttaaactgt gcattaagtc agttatggac
12721 cagcagcagt ccatgtataa ctactttaa ggctgtaatg ctgttgctaa gcatgatttc
12781 tttacttggc atgagggcag aaccatttat ggtaatgtta gtagacagga tcttactaaa
12841 tacaccatga tggatttgtg cttcgctctg cgtaactttg atgaaaaaga ctgtgaagtt
12901 ttttaaggaga tattggttct tactgggtgt tgtagtactg attactttga aatgaagaat
12961 tggtttgacc ccatagaaaa tgaggacata caccgtgtgt atgctgcttt aggtgaagga
13021 gttgcaaatg caatgcttaa gtgtgttgct ttttgcgacg aaatgggtgct caaaggagtt
13081 gttggtgttt tgacctaga caaccaagat cttaatggga atttctatga cttcgggtgac
13141 tttgtattgt gtcctcttgg aatgggaata ccctactgca cgtcatacta ttcttatatg

13201 atgcctgtta tgggtatgac taattgttta gctagtgagt gctttatgaa aagtgcacatc
13261 tttggtcaag acttcaaaac ttttgatttg ttgaaatatg atttcacaga acataaggag
13321 gttttgttta acaagtactt taagtattgg ggacaggatt atcatcctga ttgtgttgat
13381 tgccatgacg agatgtgtat tttgcattgt tcaaatttta acacactctt cgcaaccaca
13441 attccaaaca cggcttttgg acctctatgc agaaaagtgt ttattgatgg tgtacccgta
13501 gttgctactg ctggttacca ctttaaacaa ttaggacttg tgtggaacaa agatgttaac
13561 actcattcta ccagacttac tattactgaa ctcttacagt ttgtgacaga tccaacgctt
13621 atagtgcgt catgcctgc cttggtggat aaacgcactg tttgtttttc tgtcgctgct
13681 ttgagtacag gattaacatc ccaaacagta aaacctggcc attttaataa ggagttttat
13741 gacttccttac gttctcaggg gtttttcgat gaggggttcag aattaacatt gaagcatttc
13801 ttttttacac aaaaggtga tgctgcaatt aaagattttg attattatcg ttacaacaga
13861 cctactatgc tggatattgg acaagctcgc gtagcatatc aagtggcagc tcgctatttt
13921 gactgttacg aggggtggctg tattacatct agagagggtt ttgttacaaa ccttaataaa
13981 agcgctgggt ggccccctaa taagtttggg aaagctgggt tatattatga gtctattagt
14041 tatgaggaac aagatgctat tttttcatta acaaagcgta atattctccc tactatgact
14101 cagttaaadc ttaaatacgc catatctggt aaggaacgcg cacgtacagt ggggtggcgtc
14161 tctttattag ctactatgac tacaagacag tttcatcaga aatgtctgaa atccatagta
14221 gctaccagaa atgccaccgt tgttatcggc actaccaagt tttatggcgg gtgggataat
14281 atgttaaaga acctgatggc cgatgttgat gatcctaaat tgatgggatg ggactatcct
14341 aagtgtgata gagctatgcc ctcaatgatt cgtatgttgt cggctatgat cttaggttct
14401 aagcatgtca catgttgtac ggctagtgat aaattttata gacttagtaa tgagcttgct
14461 caagttttga ccgaggttgt ttattcaaat ggtgggtttt attttaaacc tgggtggtaca
14521 acttctggtg atgcaactac agcctacgcc aattctgtct ttaatatatt tcaggctgta
14581 agttctaaca ttaattgcgt tttgagcgtt aactcgtcaa attgcaataa ttttaatggt
14641 aagaagttac agagacaact ttatgataat tgctatagaa atagtaatgt tgatgaatct
14701 tttgtggatg acttttatgg ttatttgcaa aagcattttt ctatgatgat tctttctgat
14761 gatagtgttg tgtgctataa taaaacttat gctggacttg gttacattgc tgatattagt
14821 gcttttaag ccactttgta ttatcagaat ggtgtgttta tgagtacagc taagtgttg
14881 actgaggaag atctttctat aggacctcat gaattttgct cacagcacac tatgcagatt
14941 gtagatgaaa atggttaagta ttatctacca tatccagatc ctagccgtat tatttctgct
15001 ggtgtttttg tggatgacat cactaagact gatgctgtca ttcttttggg acgctatggt
15061 tctctggcta tagatgccta ccattgtct aagcatccta aacctgagta caggaagggtg
15121 ttttacgcat tgttagactg ggtcaaacat ctcaacaaga ctcttaacga aggtgttttg
15181 gagtcttttt ctgttacct tttagatgaa catgagtcta agttttggga tgaaagcttt
15241 tatgctagta tgtatgagaa gtctacagta ttacaagctg ctggtctttg tgtagtatgt
15301 ggttctcaaa cagttctaag atgcggtgat tgtttacgca gaccgatgtt gtgcactaag
15361 tgcgcctatg atcatgtgtt tggcactgat cataagttca ttttagctat tacaccatat
15421 gtgtgtaaca catctggctg caatgtaaat gacgttaca aactgtatct tggaggtttg
15481 aattattact gtgtagacca caaaccacat ctttcattcc cactgtgttc agctggtaat
15541 gtctttgggt tgtacaaaag ttctgctttg ggttccatgg acattgatgt ctttaacaaa
15601 ctttctacct ctgattgggtc tgacattcgc gactacaagc ttgctaataa tgcaaaagag
15661 tcaactaagg tgtttgcagc tgaaacgggtc aaggctaaag aggaaagtgt taagtcatca
15721 tacgcttatg ctaccctaaa ggagattgta ggtcctaagg aacttttgct cttatgggaa
15781 agtggaaaag ccaaaccacc gttaaaccgt aattctgttt ttacatgctt ccaaattaca

15841 aaagactcca agtttcaagt tggtagagttt gtgtttgaga aagtagatta cggttctgat
15901 acggttactt acaaatccac tgctactact aagttagtagt caggtagtgtt gtttattttg
15961 acttctcata atgttgctcc acttagagcg ccaacaatgg caaaccagga gaaatattct
16021 accatttaca agttgcaccc atcatttaat gttagttagt cttatgcaaa tcttgtagct
16081 tattaccaac ttattggcaa acagcgtata accacaatac agggctctcc tggtagtgga
16141 aaatcgcatg gttctatttg tattgggtgtg tattaccctg gagcgaggat cgtgttcacc
16201 gcttggttctc acgctgctgt tgattcgctc tgtgcaaaag ctgtcacagc ctatagtgtt
16261 gataagtgtg cacgtattat tcctgcacgt gccagagttg agtggtatag tggttttaaa
16321 cctaacaata atagtgcaca atacgtgttt agtactgtta atgcgttacc tgaagttaat
16381 gcagacattg ttgtcgtgga tgagggtgtct atgtgcacta actatgactt gtctgtgatt
16441 aaccagcgta tatcatataa acacattgta tatgttggtg atcctcaaca gcttcagct
16501 cctagagtgc ttatctctaa aggtgttatg gaaccaattg actataatgt tgtgacacaa
16561 cgtatgtgtg ctataggacc cgatgtcttt ttacacaagt gttacagatg tcctgctgaa
16621 atagttaaca ctgtttcaga gcttggttat gaaaacaagt ttgtacctgt caaagaagct
16681 agtaagcagt gcttcaaaat ctttgaacgc ggtagtggtc aggtagacaa tggctccagt
16741 ataaataggc gtcaacttga tgttggttaag cgatttatac ataaaaactc cacatggagc
16801 aaggctgtgt ttatctcacc ttacaatagt caaaattatg tagctgccag gcttttaggc
16861 ttacaaaactc agacagtgga ttctgctcaa ggtagtgaat atgactatgt tatattcgca
16921 cagacatcag atactgctca tgccgtgaat gccaatcgtt ttaacgttgc cactactaga
16981 gcaaagaaag gtattttctg tattatgtct gacagaactt tgtttgatgc acttaagttc
17041 tttgaaatca ctatgacaga ttacagctct gaaagtagtt gtggtttgtt taaggattgt
17101 gcacgtaacc ctattgattt accaccaagt catgccacta cttatttgtc attgtctgat
17161 agattttaaga ctagtggtga cttggctgtt caaataggta acaacaatgt ttgtacctat
17221 gaacatgtga tttcatatat gggtttcagg tttgatgtta gcatgcctgg tagtcatagt
17281 ttgttctgta ctagagactt tgccatgcgt catgtcagag gttgggttagg aatggatgtg
17341 gaagggtcac atgtcacagg tgacaatgtt ggcactaatg tacctctaca agttggtttt
17401 tccaatgggtg ttgattttgt agctcaacct gaagggtgtg ttctaacaaa cactggcagt
17461 gttgtaaaac ctgttcgtgc tcgtgcacca cctggagaac aattcactca cattgtacct
17521 ctgttacgca agggacaacc ttggagtgtg ttgagaaaac gtattgttca aatgatagca
17581 gattttcttg ctggctcatc tgatgtactg gtgtttgtac tttgggctgg cggtttagag
17641 ttgaccacta tgcgttattt tgttaagatt ggagctgtta aacattgcc aatgtgtact
17701 gttgcaacat gctacaattc tgttagtaat gactattgtt gctttaaaca tgcattgggc
17761 tgtgactatg tttataatcc atatgtcata gatattcaac aatgggggtta tgttggttca
17821 ctctccacta atcaccatgc aatttgtaat gttcatagaa atgagcatgt tgcttctggt
17881 gatgctatta tgactagatg tttggctgtg tatgactgct ttgttaagaa tgtggattgg
17941 tcaattacct accctatgat agctaataa aatgccataa acaagggcgg tcgcactgtg
18001 cagagtcata ttatgcgtgc tgctattaaa ttgtacaacc ctaaagcaat ccatgacatt
18061 ggtaatccta aggggtattcg ttgtgctgta actgatgcca agtggattgt ttatgacaag
18121 aaccttatta attctaattg gaaaacattg gagtatgatt acatgacaca tggccaaatg
18181 gatggcttgt gtttgttttg gaattgtaat gtggatatgt acctgaatt ctcaattgtt
18241 tgcaggtttg acacacgtac acgatctaca ttgaaccttg aagggtgtaa tgggtgggtca
18301 ttgtatgtca ataactatgc atttcacact cctgcttatg ataaacgtgc tatggctaaa
18361 ttgaaaccag caccgttttt ctactatgac gacggttcat gtgaggttgt tcacgatcaa
18421 gttaactatg ttcctttgag agccactaat tgcattacca agtgtaatat tgggtgggtgct

18481 gtatgttcta agcacgctaa tctctataga gcatatgttg agtcatataa cttttttact
18541 caagctgggtt ttaatatattg gggttcctacc acgtttgatt gttataattt gtggcagaca
18601 ttcacagagg tcaatttaca aggttttagag aacattgcctt ttaacgttgt taataaaggt
18661 tcatttggtg gtgctgatgg tgaattacca gtagccatta gtggtgataa agtgttcgta
18721 cgtgatggta acactgataa tttagtcttt gttaacaaaa catcactgcc tacaacata
18781 gcatttgaac tttttgctaa gaggaagggt gggttaacac cacctctcag tattctcaaa
18841 aaccttggtg ttgtcgccac atataagttt gtcttgtggg attatgaagc tgagcgtccc
18901 ttgacaagct ttactaagtc tgtttgtggt tatacagact ttgcagagga tgtttgtact
18961 tgttacgata atagtataca aggttcatac gaacgtttta ctctgtcaac taatgctgtg
19021 ttattctctg ctactgctgt gaaaacaggt ggtaagagtt tgccggctat taaattgaat
19081 tttggaatgc ttaatggtaa tgcaattgct actgtcaaat cagaagatgg taacataaaa
19141 aatattaact gggttggtta cgtacgcaaa gatggcaaac ctgttgatca ttatgatggt
19201 ttttataccc aaggctgtaa ttacaagac tttttgcctc gcagcacaat ggaagaagac
19261 tttttgaaca tggatatagg cgtgtttatt caaaagtatg gtctagagga tttcaacttc
19321 gagcacgttg tgtatggtga tgtttcaaaa actactctag gcgggttaca cttgttgatt
19381 tcacaagtac gtctgagtaa aatgggcac ttaaaggcag aggagtttgt ggcagcatct
19441 gacataacac tcaaatgttg tactgtgact tatcttaatg atcctagttc taagactggt
19501 tgtacttaca tggatttggt gttggatgat tttgtttctg tattgaagtc tttggatttg
19561 actgttgat ccaagggttc tgagggtcata attgacaaca aacctggag atggatgcta
19621 tgggtgtaaag ataatgccgt tgctacattc tctctcagt tgcagagtgc agaattgaaa
19681 tgccgggtatt ctatgcctgg tatttataag acacaacgta tgtgcttaga accatgtaat
19741 ttgtataatt atggtgcagg tttgaagttg cccagtggca ttatgttcaa tgttgttaaa
19801 tacactcaat tgtgtcaata ttttaacagt accacgttat gtgttcctca taatatgaga
19861 gtgttacact tgggtgctgg ctctgattat ggtgttgcac caggaactgc tgttcttaaa
19921 aggtgggttc cgcacgacgc aattgttggt gacaacgatg ttgttgacta tgtgagtgc
19981 gctgatttta gtgttactgg tgattgtgca accgtttatt tggaagacaa gtttgacttg
20041 ttaatctctg atatgtacga tggtaggaca aaggcaattg atggtgaaaa tgtttcgaaa
20101 gaaggatttt tcacttacat caatggtttc atttgtgaaa aacttgccat cggagggttcg
20161 attgctatta aagtaacaga gtatagctgg aataagaaat tgtatgaact tgtacaaaga
20221 ttttcttttt ggactatggt ttgcacttct gttaatacgt catcatcaga agcctttggt
20281 gtcggaatta actatcttgg tgatttcgca caaggacctt ttatagatgg taacataata
20341 cgcgcaaatt atgtattttg gcgtaactcc actgttatga gtttgccta caactctggt
20401 ttagacctga gtaaatttaa ttgcaaacac aaagcgactg ttgttggtgca attaaaggat
20461 agtgatatta atgaaatggt gcttagtctt gttaggagtg gtaagttgct tgtaaggggt
20521 aatggcaagt gtttgagttt tagtaatcat ttagtctcaa ctaaataaaa tgtttgtttt
20581 gcttggtgca tatgccttgt tgcataattgc tggttgtcaa actacaaatg ggctgaacac
20641 tagttactct gtttgcaacg gctgtgttgg ttattcagaa aatgtatttg ctgttgagag
20701 tgggtggttat ataccctccg actttgcatt caataattgg ttccttctaa ctaatacctc
20761 atctgttgta gatggtgttg tgaggagttt tcagcctttg ttgcttaatt gcttatggtc
20821 tgtttctggc ttgcgggtta ctactggttt tgtctatttt aatgggtactg ggagagggtga
20881 ttgtaaaggt ttttcctcag atgttttgtc tgatgtcata cgttacaacc tcaattttga
20941 agaaaacctt agacgtggaa ccattttggt taaaacatct tatggtgttg ttgtgtttta
21001 ttgtaccaac aacacttttag tttcagggtga tgctcacata ccatttggtg cagttttggg
21061 caatttttat tgctttgtaa atactactat tggcaatgaa actacgtctg cttttgtggg

21121 tgcactacct aagacagttc gtgagtttgt tatttcacgc acaggacatt tttatatataa
21181 tggctatcgc tatttcactt taggtaatgt agaagccgtt aatttcaatg tccactactgc
21241 agaaaccact gattttttgta ctgttgcggt agcttcttat gctgacgttt tgggttaatgt
21301 gtcacaaacc tctattgcta atataattta ttgcaactct gttattaaca gactgagatg
21361 tgaccagttg tcctttgatg taccagatgg tttttattct acaagcccta ttcaatccgt
21421 tgagctacct gtgtctattg tgcgctacc tgtttatcat aaacatacgt ttatttgttt
21481 gtacgttgac ttcaaaccct agagtggcgg tggcaagtgc tttaactgtt atcctgctgg
21541 tgttaatat acactggcca attttaatga aactaaaggg cctttgtgtg ttgacacatc
21601 acacttcaact accaaatacg ttgctgttta tgccaatgtt ggtaggtgga gtgctagtat
21661 taacacggga aattgccctt tttcttttgg caaagttaat aactttgtta aatttggcag
21721 tgtatgtttt tcgctaaagg atatacccg tggttgcgca atgcctatag tggctaattg
21781 ggcttatagt aagtactata ctataggctc attgtatgtt tcttgagtg atggtgatgg
21841 aattactggc gtcccacaac ctgttgaggg tgttagttcc tttatgaatg ttacattgga
21901 caaatgtact aaatataata tttatgatgt atctgggtgtg ggtgttattc gcgttagcaa
21961 tgacaccttt cttaatggaa ttacgtacac atcaacttca ggtaaccttc tgggttttaa
22021 agatgttact aagggcacca tctactctat cactccttgt aaccaccag atcagcttgt
22081 tgtttatcag caagctgttg ttggtgctat gttgtctgaa aattttacta gttacggctt
22141 ttctaattgt gtgaactgc cgaaattttt ctatgcgtcc aatggcactt ataattgcac
22201 agacgtgtt ttaacttatt ctagttttgg cgtttgtgca gatggttcta taattgctgt
22261 tcaaccacgt aatgtttcat atgatagtgt ttcagctatc gtcacagcta atttgtctat
22321 accttccaat tggaccactt cggtcaggt tgagtattta caaattacaa gtacacctat
22381 cgtagttgat tgctccactt atgtttgcaa tggtaatgtg cgctgtgttg aattgcttaa
22441 gcagtatact tctgcttgta aaactattga agacgcctta agaaatagcg ccaggctgga
22501 gtctgcagat gttagtgaga tgctcacttt tgacaagaaa gcgtttacac ttgctaattg
22561 tagtagtttt ggtgactaca accttagcag cgtcatacct agcttgccca caagtggtag
22621 tagagtggct ggtcgcagt ccatagaaga catacttttt agcaaacttg ttacttctgg
22681 acttggcact gtggacgcag actacaaaaa gtgcaactaag ggtctttcca ttgctgactt
22741 ggcttgtgct caatattata atggcattat ggttttgcct ggcgtcgtg atgctgaacg
22801 aatggccatg tatacaggtt ctttaatttg tggaaattgct ttaggaggtc taacatcagc
22861 cgtttcaata ccattttcat tagcaattca ggcacgttta aattatgttg cattgcagac
22921 tgatgtttta caagaaaatc agaaaattct tgctgcatct tttaacaaag caatgaccaa
22981 catagtagat gcctttactg gtgttaatga tgctattaca caaacttcac aagccctaca
23041 aacagttgct actgcactta acaagatcca ggatgttggt aatcaacaag gcaactcatt
23101 gaaccattta acttctcagt tgaggcagaa ttttcaagct atctctagct ctattcaggc
23161 tatctatgac agacttgaca ctattcaggc tgatcaacaa gtagataggc tgattactgg
23221 tagattggct gctttgaatg tattcgtttc tcatacattg actaagtaca ctgaagtctg
23281 tgcttccaga cagcttgcac aacaaaaagt gaatgagtgt gtcaaatccc agtctaagcg
23341 ttatggcttc tgtggaaatg gcactcacat tttctcaatt gttaatgctg ctccgtaggg
23401 gcttgttttt ctccacactg tcttgttgcc gacacaatat aaggatgttg aagcgtggtc
23461 tgggttgtgc gttgatggta caaacggta tgtgttgca caacctaatc ttgctcttta
23521 caaagaaggc aattattata gaatcacatc tcgcataatg tttgaaccac gtattcctac
23581 catggcagat tttgttcaaa ttgaaaattg caatgtcaca tttgttaaca tttctcgctc
23641 tgagttgcaa accattgtgc cagagtatat tgatgttaat aagacgctgc aagaattaag
23701 ttacaaattg ccaaattaca ctgttccaga cctagttgtc gaacagtaca accagactat

23761 tttgaatttg accagtgaat ttagcaccct tgaaaataaa tctgcgagc ttaattacac
23821 tgttcaaaaa ttgcaaacct tgattgacaa cataaatagc acattagtcg acttaaagtg
23881 gctcaaccgg gttgagactt acatcaagtg gccgtgggtg gtgtgggtgt gcatttcagt
23941 cgtgctcatc tttgtggtga gtatgttgc tttatgttgt tggttctactg gttgctgtgg
24001 cttctttagt tgttttgcat cttctattag aggttgttgt gaatcaacta aacttcctta
24061 ttacgacgtt gaaaagatcc acatacagta atggctctag gtttggtcac attgcaactt
24121 gtgtctgctg ttaatcaatc gcttagcaat gcgaaagtta gtgtgaagt ttcacgacag
24181 gttatccaag acgtgaaaga tggcactggt accttcaact tgctagcgta tacactaatg
24241 agcctctttg ttgtgtattt tgctttattt aaagcaagat cacaccgtgg cagagctgct
24301 cttatagtgt ttaaaattct aatccttttc gtttatgtgc cattgtgtga ttggtctcaa
24361 gcatatattt acgcaacttt gattgtgtga attttgcttg gaagattttt ccatacagct
24421 tggcactgct ggctctacaa gacatgggat ttcattgtct tcaatgtaac cacactttgc
24481 tatgcaaggt aagtgttggg ttcttgaaaa taaggctctg aaaccattcg tttgttttta
24541 cggaggggat caattccttt acataggcga cagaattggt tcttatttct caactaacga
24601 cttgtacgtt gctcttagag gacgtattga taaagacctc agcctttcta gaaaggttga
24661 gttatataac ggtgaatgtg tatacttggt ttgtgaacac ccagctgttg gaatagtcaa
24721 cacagatttc aaattagaaa tccactaaga tggttcctta gctagtggat gatcatgctt
24781 tgggtgttaa tgtactactc tgggtgtgtg tgcttatagt gataactacta gtgtgtatta
24841 caataattaa actaattaag ctttgtttca cttgccatat gttttgtaat agaacagttt
24901 atggcccat taaaaatgtg taccacattt accaatcata tatgcacata gaccctttcc
24961 ctaaacgagt tattgatttc taaactaaac gacaatgtca aatgacaatt gtacgggtga
25021 cattgtcacc catttgaaga attggaattt tgggttgaat gttattctaa ccatattcat
25081 tgttattctt cagtttgac actataaata ctccagattg ttttatggtt tgaagatgct
25141 tgtactgtgg cttctttggc cactcgctact tgctttgtca atctttgaca cctgggctaa
25201 ttgggattct aattgggcct ttgttgcat tagctttttt atggccgtat caacactcgt
25261 tatgtgggtg atgtacttcg caaacagttt cagacttttc cgacgtgctc gaactttttg
25321 ggcattggaat cctgagggtta atgcaatcac tgtcacaacc gtgttgggac agacatacta
25381 tcaaccatt caacaagctc caacaggcat tactgtgacc ttgctgagcg gcgtgcttta
25441 cggtgacgga catagattgg cttcaggtgt tcaggttcat aacctacctg aatacatgac
25501 agttgccgtg ccgagcacta ctataattta tagtagagtc ggaaggccg taaattcaca
25561 aaatagcaca ggctgggttt tctacgtacg agtaaaacac ggtgattttt ctgcagtga
25621 ctctcccatg agcaacatga cagaaaacga aagattgctt cattttttct aaactgaacg
25681 aaaagatggc tacagtcaaa tgggctgatg catctgaacc acaacgtggt cgctcagggt
25741 gaatacctta ttctctttat agccctttgc ttgttgatag tgaacaacct tgggaaggtga
25801 tacctcgtaa tttggtacct atcaacaaga aagacaaaaa taagcttata ggctattgga
25861 atgttcaaaa acgtttcaga actagaaagg gcaaacgggt ggatttgtca cccaagctgc
25921 atttttatta tcttggcaca ggaccccata aagatgcaaa atttagagag cgtgttgaag
25981 gtgtcgtctg ggttgctgtt gatggtgcta aaactgaacc tacaggttac ggtgttaggc
26041 gcaagaattc agaaccagag ataccacact tcaatcaaaa gctcccaaat ggtgttactg
26101 ttgttgaaga acctgactcc cgtgctcctt ccgggtctca gtogaggctg cagagtcgag
26161 gtcgtggtga atccaaacct caatctcgga atccttcaag tgacagaaac cataacagtc
26221 aggatgacat catgaaggca gttgctgcgg ctcttaaatac tttagggttt gacaagcctc
26281 aggaaaaaga taaaaagtca gcgaaaacgg gtactcctaa gccttctcgt aatcagagtc
26341 ctgcttcttc tcaaaattct gccaaagatc ttgctcgttc tcagagttct gaaacaaaag

26401 aacaaaagca tgaaatgcaa aagccacggt ggaaaagaca gcctaatagat gatgtgacat
26461 ctaatgtcac acaatgtttt ggccccagag accttgacca caactttgga agtgcagggtg
26521 ttgtggccaa tgggtgttaaa gctaaaggct atccacaatt tgctgagcct gtgccgtcaa
26581 cagctgctat gctgtttgat agtcacattg tttccaaaga gtcaggcaac actgtggtct
26641 tgactttcac tactagagtg actgtgccca aagaccatcc acacttgggt aagtttcttg
26701 aggagttaaa tgcattcact agagaaatgc aacaacatcc tcttcttaac cctagtgcac
26761 tagaattcaa cccatctcaa acttcacctg caactgctga accagtgcgt gatgaagttt
26821 ctattgaaac tgacataatt gatgaagtaa actaaacatg ccaactgtgtt gtttgaaatt
26881 caggcttttag ttggaatttt gcttttgttc tttcttttat tatctttctt ttgcctgttt
26941 ttagagagat ttggcgcctt ggtgccgtag atgaatacat tgcttttctc tgatctatgt
27001 atgatgttac gatcagagct gcttttaatt aacatgatcc cttgctttgg cttgacaagg
27061 atctagtctt atacacaatg gtaagccagt ggtagtaaag gtataagaaa tttgctacta
27121 tgttactgaa cctaggtgaa cgctagtata actcattaca aatgtgctgg agtaatcaaa
27181 gatcgattg acgagccaac aatggaagag ccagtcattt gtcttgagac ctatctagtt
27241 agtaactgct aatggaacgg ttctgatatg gatacacaaa aaaaaaaaaa aaaaaaaaaa
27301 aaaaaaaaaa aaaaaaa

//