

Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome

NCBI Reference Sequence: NC_045512.2

Comparison to Bat-SARS-ZXC21 marked red. Region does not match Bat-SARS-ZC45.

Bat SARS-like coronavirus Isolate "bat-SL-CoVZXC21" Host "Rhinolophus pusillus"

Country "China" Collection date="Jul-2015" SOURCE: NIH GenBank: MG772934.1

HIV-1 analogs believed to be the binding sites of the AIDS gp120 antibodies obtained from the NIH AIDS Reagent program. <https://www.nature.com/articles/s41598-021-91746-7> and <https://doi.org/10.1101/2020.01.30.927871>

HIV-1 sequences in question: GTNGTKR, YYHKNNKS, GDSSSG, and QTNSPRRA.

aagaa - suspected RNA modification precursors
<https://doi.org/10.1016/j.cell.2020.04.011> and
<https://doi.org/10.1016/j.bbadis.2020.165878>

Spike protein coding region of the RNA is bolded: 21563 – 25384.

Furin cleavage site **cgg cgg which codes for a pair of Arginines is frequently found in genetic manipulation. See <https://pubmed.ncbi.nlm.nih.gov/32869021/>**

ctcctcggcg ggcacgtag Furin cleavage site found at 23601-23619. The reverse complement form of this sequence is ctacgtgccgcccgcgaggag which is found in 5 U.S. Patents issued to ModernaTx Inc: US9149506B2, US9216205B2, US9255129B2, US9301993B2, and US9587003B2. The first four were filed for patent on Dec 16, 2013 and the last was filed for patent on Feb 4, 2016. This Furin cleavage site is believed by many to be the region causing severe disease in humans.

SARS-CoV-2 Wuhan 20712-20771 compared to

SARS-CoV CDC Patent US7220852 20701-20760

aagaatgctattagaaaagtgtgaccttcaaaattatggtgatagtcaacattacctaa

aagaatgcttcttgaagaagtgtgaccttcagaattatggtgaaaatgctgttataccaaa

Region matching spike of original SARS-CoV spike: acaa aggaattt

[FASTA Graphics](#)

[Go to:](#)

LOCUS NC_045512 29903 bp ss-RNA linear VRL 18-JUL-2020
DEFINITION Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1,
complete genome.
ACCESSION NC_045512
VERSION NC_045512.2
DBLINK BioProject: [PRJNA485481](#)
KEYWORDS RefSeq.
SOURCE Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)
ORGANISM [Severe acute respiratory syndrome coronavirus 2](#)
Viruses; Riboviria; Orthornavirae; Pisuviricota; Pisoniviricetes;
Nidovirales; Coronidovirineae; Coronaviridae; Orthocoronavirinae;
Betacoronavirus; Sarbecovirus.
REFERENCE 1 (bases 1 to 29903)
AUTHORS Wu,F., Zhao,S., Yu,B., Chen,Y.M., Wang,W., Song,Z.G., Hu,Y.,
Tao,Z.W., Tian,J.H., Pei,Y.Y., Yuan,M.L., Zhang,Y.L., Dai,F.H.,
Liu,Y., Wang,Q.M., Zheng,J.J., Xu,L., Holmes,E.C. and Zhang,Y.Z.
TITLE A new coronavirus associated with human respiratory disease in
China
JOURNAL Nature 579 (7798), 265-269 (2020)
PUBMED [32015508](#)
REMARK Erratum:[Nature. 2020 Apr;580(7803):E7. PMID: 32296181]
REFERENCE 2 (bases 13476 to 13503)
AUTHORS Baranov,P.V., Henderson,C.M., Anderson,C.B., Gesteland,R.F.,
Atkins,J.F. and Howard,M.T.
TITLE Programmed ribosomal frameshifting in decoding the SARS-CoV
genome
JOURNAL Virology 332 (2), 498-510 (2005)
PUBMED [15680415](#)
REFERENCE 3 (bases 29728 to 29768)
AUTHORS Robertson,M.P., Igel,H., Baertsch,R., Haussler,D., Ares,M. Jr.
and
Scott,W.G.
TITLE The structure of a rigorously conserved RNA element within the
SARS
virus genome
JOURNAL PLoS Biol. 3 (1), e5 (2005)
PUBMED [15630477](#)
REFERENCE 4 (bases 29609 to 29657)
AUTHORS Williams,G.D., Chang,R.Y. and Brian,D.A.
TITLE A phylogenetically conserved hairpin-type 3' untranslated region
pseudoknot functions in coronavirus RNA replication
JOURNAL J. Virol. 73 (10), 8349-8355 (1999)
PUBMED [10482585](#)
REFERENCE 5 (bases 1 to 29903)
CONSRM NCBI Genome Project
TITLE Direct Submission
JOURNAL Submitted (17-JAN-2020) National Center for Biotechnology
Information, NIH, Bethesda, MD 20894, USA
REFERENCE 6 (bases 1 to 29903)
AUTHORS Wu,F., Zhao,S., Yu,B., Chen,Y.-M., Wang,W., Hu,Y., Song,Z.-G.,
Tao,Z.-W., Tian,J.-H., Pei,Y.-Y., Yuan,M.L., Zhang,Y.-L.,
Dai,F.-H., Liu,Y., Wang,Q.-M., Zheng,J.-J., Xu,L., Holmes,E.C.
and

Zhang, Y.-Z.
TITLE Direct Submission
JOURNAL Submitted (05-JAN-2020) Shanghai Public Health Clinical Center &
School of Public Health, Fudan University, Shanghai, China
COMMENT REVIEWED [REFSEQ](#): This record has been curated by NCBI staff. The
reference sequence is identical to [MN908947](#).
On Jan 17, 2020 this sequence version replaced [NC_045512.1](#).
Annotation was added using homology to SARSr-CoV NC_004718.3. ###
Formerly called 'Wuhan seafood market pneumonia virus.' If you

have

questions or suggestions, please email us at
info@ncbi.nlm.nih.gov
and include the accession number NC_045512.### Protein structures
can be found at
<https://www.ncbi.nlm.nih.gov/structure/?term=sars-cov-2.###> Find
all other Severe acute respiratory syndrome coronavirus 2
(SARS-CoV-2) sequences at
<https://www.ncbi.nlm.nih.gov/genbank/sars-cov-2-seqs/>

##Assembly-Data-START##
Assembly Method :: Megahit v. V1.1.3
Sequencing Technology :: Illumina
##Assembly-Data-END##
COMPLETENESS: full length.

FEATURES Location/Qualifiers
source 1..29903
/organism="Severe acute respiratory syndrome coronavirus
2"
/mol_type="genomic RNA"
/isolate="Wuhan-Hu-1"
/host="Homo sapiens"
/db_xref="taxon:[2697049](#)"
/country="China"
/collection_date="Dec-2019"
[5'UTR](#) 1..265
[gene](#) 266..21555
/gene="ORF1ab"
/locus_tag="GU280_gp01"
/db_xref="GeneID:[43740578](#)"
[CDS](#) join(266..13468,13468..21555)
/gene="ORF1ab"
/locus_tag="GU280_gp01"
/ribosomal_slippage
/note="pplab; translated by -1 ribosomal frameshift"
/codon_start=1
/product="ORF1ab polyprotein"
/protein_id="[YP_009724389.1](#)"
/db_xref="GeneID:[43740578](#)"

/translation="MESLVPGFNEKTHVQLSLPVLQVRDVLVRGFGDSVEEVLSEARQ
HLKDGTCLVEVEKGVLPQLEQPYVFIKRS DARTAPHGHVMVELVAELEGIQYGRSGE
TLGVLVPHVGEIPVAYRKVLLRKNNGKAGGHSYGADLKSFDLGDELGTDPYEDFQEN
WNTKHSSGVTRELMRELNGGAYTRYVDNNFCGPDGYPLECIKDLLARAGKASCTLSEQ

LD FIDTKRGVYCCREHEHEIAWYTERSEKSYELQTPFEIKLAKKFDTFN GECPNFVFP
LNSIIKTIQPRVEKKKLDGFMGRIRSVYPVSPNECNQMCLSTLMKCDHCGETSWQTG
DFVKATCEFCGTENLTKEGATTCGYLPQNAVVKIYCPACHNSEVGP EHS LAEYHNESG
LKTILRKGGRTIAFGGCVFSYVGCHNKAYWVPRASANIGCNHTGVV GEGSEGLNDNL
LEILQKEKVNINIVGDFKLN E E I A I I L A S F S A S T S A F V E T V K G L D Y K A F K Q I V E S C G N
FKVTKGKAKKGAWNIGE Q K S I L S P L Y A F A S E A A R V V R S I F S R T L E T A Q N S V R V L Q K A A
ITILDGISQYSLRLIDAMMFTSDLATNNLVVMAYITGGVVQLTSQWL TNIFGTVYEKL
KPVLDWLEEKFKEGVEFLRDGWEIVKFISTCACEIVGGQIVTCAKEIKESVQTFFKLV
NKFLALCADSIIIGGAKLKALNLGETFVTHSKGLYRKCVKSREETGLLMPLKAPKEII
FLEGETLPTEVLTEE VVLKTGDLQPLEQPTSEAVEAPLVGTPVCINGLM LLEIKDTEK
YCALAPNMMVTNNTFTLKG GAPT K V T F G D D T V I E V Q G Y K S V N I T F E L D E R I D K V L N E K
CSAYTVELGTEVNEFACVVADAVIKTLQPVSELLTPLGIDLDEWSMATYYLFD E S G E F
KLASHMYCSFYPPDEDEEEGDCEEEEFEPSTQYEYGTEDDYQ GKPLEFGATSAAALQPE
EEQEEDWLDDDSQQT V G Q Q D G S E D N Q T T T I Q T I V E V Q P Q L E M E L T P V V Q T I E V N S F S G
YLKLTDNVYIKNADIVEEAKVKVPTVVVNAANVYLKHGGGVAGALNKATNNAMQVESD
DYIATNGPLKVG G S C V L S G H N L A K H C L H V V G P N V N K G E D I Q L L K S A Y E N F N Q H E V L L A
PLLSAGIFGADPIHSLRVCVDTVR TNVYLA V F D K N L Y D K L V S S F L E M K S E K Q V E Q K I A
EIPKEEVKPFITESKPSVEQRKQDDKKIKACVEEVT T T T L E E T K F L T E N L L L Y I D I N G N
LHPDSATLVSDIDITFLKKDAPYIVGDVVQEGVLTAVVIPTKKAGGTTEMLAKALRKV
PTDNYITTYPGQGLNGYTV E E A K T V L K K C K S A F Y I L P S I I S N E K Q E I L G T V S W N L R E M
LAHAETRKLMPVCVETKAIVSTIQRKYKGIKI Q E G V V D Y G A R F Y F Y T S K T T V A S L I N
TLNDLNETLVTMPLGYVTHGLNLEEAARYMRS L K V P A T V S V S S P D A V T A Y N G Y L T S S S
KTPEEHFIETISLAGSYKDWSYSGQSTQLGIEFLKRGDKSVYYT SNPTTFHLDGEVIT
FDNLKTL L S L R E V R T I K V F T T V D N I N L H T Q V V D M S M T Y G Q Q F G P T Y L D G A D V T K I K P H
NSHEGKTFYVLPND D T L R V E A F E Y Y H T T D P S F L G R Y M S A L N H T K K W K Y P Q V N G L T S I K
WADNNCYLATALLTQQIELKFNPPALQDAYYRARAGEAANFCALILAYCNKTVGELG
DVRETMSYLFQHANLD SCKRVLNVVCKTCGQQQTTLKGVEAVMYMG T L S Y E Q F K K G V Q
IPCTCGKQATKYL V Q Q E S P F V M M S A P P A Q Y E L K H G T F T C A S E Y T G N Y Q C G H Y K H I T S K

ETLYCIDGALLTKSSEYKGPITDVFYKENSYTTTTIKPVTYKLDGVVCTEIDPKLDNYY
KKDNSYFTEQPIDLVPNQYPYNASFDNFKFVCDNIKFADDLNQLTGYKKPASRELKVT
FFPDLNGDVVAIDYKHYTPSFKKGAKLLHKPIVWHVNNATNKATYKPNWTCIRCLWST
KPVETSNSFDVLKSEDAQGMDNLACEDLKPVSEEVVENPTIQKDVLECNVKTTEVVGD
IILKPANNSLKITEEVGHTDLMAAYVDNSSLTIKKPNELSRVLGLKTLATHGLAAVNS
VPWDTIANYAKPFLNKVVSTTTNIVTRCLNRVCTNYMPYFFTLQLCTFTRSTNSRI
KASMPPTIAKNTVKSVMGKFCLEASFNYLKSPNFSKLINIIWFLLLSVCLGSLIYSTA
ALGVLMNSNLGMPYCTGYREGYLNSTNVTIATYCTGSI PCSVCLSGLDSDLTYP SLET
IQITISSFKWDLTAFGLVAEWFLAYILFTRFFYVLGLAAIMQLFFSYFAVHFISNSWL
MWLIINLVQMAPISAMVRMYIFFASFYVWKS YVHVVDGCNSSTCMMCYKRN RATRVE
CTTIVNGVRRSFYVYANGGKGFCKLHNWNCVNCDTFCAGSTFISDEVARDSLQFKRP
INPTDQSSYIVDSVTVKNGSIHLYFDKAGQKTYERHSLSHFVNLDNLRANNTKGS LPI
NVIVFDGKSKCEESSAKSASVYYSQLMCQPILLDDQALVSDVGDSA EVAVKMF DAYVN
TFSSTFNVPMEKLT LVATAEAE LAKNVSLDNVLS TFI SAARQGFVDS DVETKDVVEC
LKLSHQSDIEVTGDSCNNYMLTYNKVENMTPRDLGACIDCSARHINAQVAKSHNIALI
WNVKDFMSLSEQLRKQIRSAAKNNLPFKLTCATTRQVVNVVTTKIALKGGKIVNNWL
KQLIKVTLVFLFVA AIFYLITPVHVMSKHTDFSSEIIGYKAIDGGVTRDIAS TDT CFA
NKHADFDTWFSQRGGSY TNDKACPLIAAVITREVG FVVPGLPGTILRTTNGDFLHFLP
RVFSAVGNICYTPSKLIEYTD FATSACVLAAECTIFKDASGKVPYCYDTNVLEGSVA
YESLRPDTRYV LMDGSI IQFPNTYLEGSRVVTTFDSEYCRHGTCERSEAGVCVSTSG
RWVLNNDYYRSLPGVFCGVDAVNLLTNMFTPLIQPIGALDISASIVAGGIVAI VVTCL
AYYFMRFRRAFGEYSHVVA FNTLLFLMSFTVLCLTPVYSFLPGVYSVIYLYLTFYLTN
DVSFLAHIQWVMVFTPLVPFWIT IAYI ICISTKH FYWFFSNYLRKRVVFN GVSFSTFE
EAALCTFLLNKEMYLKLRSDVLLPLTQYNRYLALYNKYKYFSGAMD TTSYRE AACCHL
AKALNDFSNSGSDVLYQPQTSITSAVLQSGFRKMAFSPGKVEGCMVQVTCGTTTLNG
LWLDDVVYCPRHVICTSEDMLNP NYEDLLIRKSNHNFLVQAGNVQLRVIGHSMQNCVL
KLKVD TANPKTPKYKFVRIQPGQTF SVLACYNGSPSGVYQCAMPNFTIKGSFLNGSC
GSVGFNIDYDCVSFCYMHMELPTGVHAGTDLEGNFYGPFVDRQTAQAAGTDTTITVN

VLAWLYAAVINGDRWFLNRFTTTLNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGI
LDMCASLKELLQNGMNGRTILGSALLEDEFTPFDVVRQCSGVTFQSAVKRTIKGTHHW
LLLTILTSLLVLVQSTQWSLFFFLYENAFLEPFAMGIIAMSAFAMMFVKHKAFLCLFL
LPSLATVAYFNMVYMPASWVMRIMTWLDMVDTSLSGFKLKDCVMYASAVVLLILMTAR
TVYDDGARRVWTLMNVLTLVYKVYYGNALDQAI SMWALI I SVTSNYSGVVTTVMFLAR
GIVFMCVEYCPIFFITGNLQOCIMLVYCFGLGYFCTCYFGLFCLLNRYFRLTLGVYDYL
VSTQEFRYMNSQGLLPPKNSIDAFKLNKLLGVGGKPCIKVATVQSKMSDVKCTSVVL
LSVLQQLRVESSSKLWAQCVQLHNDILLAKDTTEAFEKMSVLLSVLLSMQGAVDINKL
CEEMLDNRATLQAIASEFSSLSYAAFATAQEAYEQAVANGDSEVVLLKCLKKSLNVAK
SEFDRDAAMQRKLEK MADQAMTQMYKQARSEDKRAKVT SAMQTMLFTMLRKLNDALN
NI INNARDGCVPLNI I PLTTAAKLMVVI PDYNTYKNTCDGTTFTYASALWEIQQVDA
DSKIVQLSEI SMDNSPNLAWPLIVTALRANSVAVKLQNNELSPVALRQMSCAAGTTQTA
CTDDNALAYYNTTKGGRFVLALLSDLQDLKWARFPKSDGTGTIYTELEPPCRFVTDTP
KGPVKYLYFIKGLNNLNRMVGLS LAATVRLQAGNATEVPANSTVLSFCAFAVDAAK
AYKDYLASGGQPI TNCVKMLCTHTGTGQAITVTPEANMDQESFGGASCCLYCRCHIDH
PNPKGFCDLKGYVQI PTTCANDPVGFTLKNTVCTVCGMWKGYGCSCDQLREPMLQSA
DAQSFLNRVCGVSAARLTPCGTGTSTDVVYRAFDIYNDKVAGFAKFLKTNCCRFQEKD
EDDNLIDSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHDFKFRIDGDMVPHISRQR
LTKYTMADLVYALRHFDEGNCDTLKEILVTYNCCDDYFNKKDWYDFVENPDILRVYA
NLGERVRQALLKTVQFC DAMRNAGIVGVLTLDNQLNGNWYDFGDFIQTTPGSGVPVV
DSYYSLLMPILTLTRALTAESHVDTDLTKPYIKWDLKDYDFTEERLKLFDYFKYWDQ
TYHPNCVNCLDDRCILHCANFNVLSTVFPPTSFGPLVRKIFVDGVPFVVSTGYHFRE
LGVVHNQDVNLHSSRSLFKELLVYAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQ
TVKPGNFNKDFYDFAVSKGFFKEGSSVELKHFFFAQDGNAAISDYDYRYNLPTMCDI
RQLLFVVEVVDKYFDCYDGGCINANQVIVNNLDKSAGFPFNKWKARLYYDSMSYEDQ
DALFAYTKRNVIP TITQMNLYAISAKNRARTVAGVSICSTMTNRQFHQKLLKSIAAT
RGATVVIGTSKFYGGWHNMLKTVYSDVENPHLMGWDYPKCDRAMPNMLRIMASLVLAR
KHTTCCSLSHRFYRLANECAQVLSEMVMCGGSLYVKPGGTSSGDATTAYANSVFNICQ

AVTANVNALLSTDGNKIADKYVRNLQHRLYECLYRNRDVDTDFVNEFYAYLRKHFSMM
ILSDDAVVCFNSTYASQGLVASIKNFKSVLYYQNNVFMSEAKCWTETDLTKGPHEFCS
QHTMLVKQGDDYVYLPYPDPSRILGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKH
PNQEYADVFLYLQYIRKLHDELTDGHMLDMYSVMLTNDNTSRYWEPEFYEAMYTPHTV
LQAVGACVLCNSQTSLRGACIRRPFLCCKCCYDHVISTSHKLVLSVNPYVCNAPGCD
VTDVTQLYLGGMSYYCKSHKPPISFPLCANGQVFGLYKNTCVGSDNVTDFNAIATCDW
TNAGDYILANTCTERLKLFAAETLKATEETFKLSYGIATVREVLSDRELHLSWEVGKP
RPPLNRNYVFTGYRVTKNSKVQIGEYTFEKGDYGDVAVYRGTTTTYKLVNGDYFVLTSH
TVMPLSAPTLVPOEHYVRITGLYPTLNI SDEFSSNVANYQKVGMQKYSTLQGPPGTGK
SHFAIGLALYYP SARIVYTACSHA AVDALCEKALKYLPIDKCSRIIPARARVECFDKF
KVNSTLEQYV FCTVNALPETTADIVVFDEISMATNYDLSVVNARLRAKHVYVIGDPAQ
LPAPRTLLTKGTLEPEYFNSVCRLMKTIGPDMFLGTCRRCPAEIVDTV SALVYDNK LK
AHKDKSAQC FKM FYKGVITHDVSSAINRPQIGVVREFLTRNPAWRKAVFISPYN SQNA
VASKILGLPTQTVDSSQ GSEYDYVIFTQT TETAHSCNVNRFNVAITRAKVGILCIMS D
RDLYDKLQFTSLEIPRRNVATLQAENVTGLFKDCSKVITGLHPTQAPTHLSVDTKFKT
EGLCVDIPGIPKDMTYRRLISMMGFKMNYQVNGYPNMFITREEAIRHVRAWIGFDVEG
CHATREAVGTNLPLQLGFSTGVNVLVAVPTGYVDTPNNTDFSRVSAKPPPGDQFKHLIP
LMYKGLPWNVVRIKIVQMLSDTLKNLSDRVVFVLWAHGFE L TSMKYFVKIGPERTCCL
CDRRATCFSTASDTYACWHHSIGFDYVYNPFMIDVQQWGFTGNLQSNHDLYCQVHGNA
HVASCDAIMTRCLAVHECFVKRVDWTIEYPIIGDELKINAACRKVQH MVVKAALLADK
FPVLHDIGNPKAIKCVPOADVEWKFYDAQPCSDKAYKIEELFY SYATHSDKFTDGVCL
FWNCNVDRYPANSIVCRFDTRVLSNLNLP GCDGGS LYVNKHAFTPAFDKSAFVNLKQ
LPPFFYSDSPCESHGKQVVSDIDYVPLKSATCITRCNLGGAVCRHHANEYRLYLDAYN
MMISAGFSLWVYKQFDTYNLWNTFTRLQ SLENVAFNVVNKGHFDGQQGEVPVSIINNT
VYTKVDGVDVELFENKTTLPVNVA FELWAKRNIKPVPEVKILNNLGVDIAANTVIWDY
KRDAPAHISTIGVCSMTDIAKKPTETICAPLTVFFDGRVDGQVDLFRNARNGVLITEG
SVKGLQPSVGPQASLNGVTLIGEAVKTQFNYYKKVDGVVQQLPETYFTQSRNLQEFK
PRSQMEIDFLELAMDEFIERYKLEGYAFEHIVYGDFSHSQLGGLHLLIGLAKRFKESP

FELEDVIPMDSTVKNYFITDAQTGSSKCVCSVIDLLLLDDFVEI IKSQDLSVVSQVVKV
TIDYTEISFMLWCKDGHVETFYPKLQSSQAWQPGVAMPNLYKMQRMLLEKCDLQNYGD
SATLPKGIMMNVAKYTQLCQYLNTLTLAVPYNMRVIHFGAGSDKGVAPGTAVLRQWLP
TGTLVDSLDNDFVSDADSTLIGDCATVHTANKWDLIISDMYDPKTKNVTKENDSKEG
FFTYICGFIQQKLALGGSVAIKITEHSWNADLYKLMGHFAWWTAFVTNVNASSSEAF
IGCNYLGKPREQIDGYVMHANYIFWRNTNPIQLSSYSLFDMSKFPLKLRGTAVMSLKE

GQINDMILSLLSKGRLIIRENNRVVIVSSDVLVNN"
mat peptide 266..805
/gene="ORF1ab"
/locus_tag="GU280_gp01"
/product="leader protein"
/note="nsp1; produced by both ppla and pplab"
/protein_id="[YP_009725297.1](#)"
mat peptide 806..2719
/gene="ORF1ab"
/locus_tag="GU280_gp01"
/product="nsp2"
/note="produced by both ppla and pplab"
/protein_id="[YP_009725298.1](#)"
mat peptide 2720..8554
/gene="ORF1ab"
/locus_tag="GU280_gp01"
/product="nsp3"
/note="former nsp1; conserved domains are: N-terminal
acidic (Ac), predicted phosphoesterase, papain-like
proteinase, Y-domain, transmembrane domain 1 (TM1),
adenosine diphosphate-ribose 1''-phosphatase (ADRP);
produced by both ppla and pplab"
/protein_id="[YP_009725299.1](#)"
mat peptide 8555..10054
/gene="ORF1ab"
/locus_tag="GU280_gp01"
/product="nsp4"
/note="nsp4B_TM; contains transmembrane domain 2 (TM2);
produced by both ppla and pplab"
/protein_id="[YP_009725300.1](#)"
mat peptide 10055..10972
/gene="ORF1ab"
/locus_tag="GU280_gp01"
/product="3C-like proteinase"
/note="nsp5A_3CLpro and nsp5B_3CLpro; main proteinase
(Mpro); mediates cleavages downstream of nsp4. 3D
structure of the SARSr-CoV homolog has been determined
(Yang et al., 2003); produced by both ppla and pplab"
/protein_id="[YP_009725301.1](#)"
mat peptide 10973..11842
/gene="ORF1ab"
/locus_tag="GU280_gp01"
/product="nsp6"
/note="nsp6_TM; putative transmembrane domain; produced

by

[mat peptide](#) both ppla and pplab"
 /protein_id="[YP_009725302.1](#)"
 11843..12091
 /gene="ORF1ab"
 /locus_tag="GU280_gp01"
 /product="nsp7"
 /note="produced by both ppla and pplab"
[mat peptide](#) /protein_id="[YP_009725303.1](#)"
 12092..12685
 /gene="ORF1ab"
 /locus_tag="GU280_gp01"
 /product="nsp8"
 /note="produced by both ppla and pplab"
[mat peptide](#) /protein_id="[YP_009725304.1](#)"
 12686..13024
 /gene="ORF1ab"
 /locus_tag="GU280_gp01"
 /product="nsp9"
 /note="ssRNA-binding protein; produced by both ppla and
 pplab"
[mat peptide](#) /protein_id="[YP_009725305.1](#)"
 13025..13441
 /gene="ORF1ab"
 /locus_tag="GU280_gp01"
 /product="nsp10"
 /note="nsp10_CysHis; formerly known as growth-factor-
 like
 protein (GFL); produced by both ppla and pplab"
[mat peptide](#) /protein_id="[YP_009725306.1](#)"
 join(13442..13468,13468..16236)
 /gene="ORF1ab"
 /locus_tag="GU280_gp01"
 /product="RNA-dependent RNA polymerase"
 /note="nsp12; NiRAN and RdRp; produced by pplab only"
[mat peptide](#) /protein_id="[YP_009725307.1](#)"
 16237..18039
 /gene="ORF1ab"
 /locus_tag="GU280_gp01"
 /product="helicase"
 /note="nsp13_ZBD, nsp13_TB, and nsp_HEL1core; zinc-
 binding
 domain (ZD), NTPase/helicase domain (HEL), RNA
 5'-triphosphatase; produced by pplab only"
[mat peptide](#) /protein_id="[YP_009725308.1](#)"
 18040..19620
 /gene="ORF1ab"
 /locus_tag="GU280_gp01"
 /product="3'-to-5' exonuclease"
 /note="nsp14A2_ExoN and nsp14B_NMT; produced by pplab
 only"
[mat peptide](#) /protein_id="[YP_009725309.1](#)"
 19621..20658
 /gene="ORF1ab"
 /locus_tag="GU280_gp01"
 /product="endoRNase"
 /note="nsp15-A1 and nsp15B-NendoU; produced by pplab
 only"

mat peptide /protein_id="[YP_009725310.1](#)"
20659..21552
/gene="ORF1ab"
/locus_tag="GU280_gp01"
/product="2'-O-ribose methyltransferase"
/note="nsp16_OMT; 2'-o-MT; produced by pplab only"
/protein_id="[YP_009725311.1](#)"
CDS 266..13483
/gene="ORF1ab"
/locus_tag="GU280_gp01"
/note="ppla"
/codon_start=1
/product="ORF1a polyprotein"
/protein_id="[YP_009725295.1](#)"
/db_xref="GeneID:[43740578](#)"

/translation="MESLVPGFNEKTHVQLSLPVLQVRDVLVRGFGDSVEEVLSEARQ
HLKDGTCGLVEVEKGVLPQLEQPYVFIKRS DARTAPHGHVMVELVAELEGIQYGRSGE
TLGVLVPHVGEIPVAYRKVLLRKNNGKAGGHSYGADLKSFDLGDELGTDYPYEDFQEN
WNTKHSSGVTRELMRELNGGAYTRYVDNNFCGPDGYPLECIKDLLARAGKASCTLSEQ
LDFIDTKRGVYCCREHEHEIAWYTERSEKSYELQTPFEIKLAKKFDTFNGECPNFVFP
LNSIIKTIQPRVEKKKLDGFMGRIRSVYPVSPNECNQMCLSTLMKCDHCGETSWQTG
DFVKATCEFCGTENLTKEGATTCGYLPQNAVVKIYCPACHNSEVGPEHSLAEYHNESG
LKTILRKGGRTIAFGGCVFSYVGCHNK CAYWVPRASANIGCNHTGVVGESEGLNDNL
LEILQKEKVNINIVGDFKLNEEIAIILASFSASTSAFVETVKGLDYKAFKQIVESCGN
FKVTKGKAKKGAWNIGE QKSILSPLYAFASEAARVVRSIFSRTLETAQNSVRVLQKAA
ITILDGISQYSLRLIDAMMFTSDLATNNLVVMAYITGGVVQLTSQWLTNIFGTVYEKL
KPVLDWLEEKFKEGVEFLRDGWEIVKFISTCACEIVGGQIVTCAKEIKESVQTFKLV
NKFLALCADSIIIGGAKLKALNLGETFVTHSKGLYRKC VKSREETGLLMPLKAPKEII
FLEGETLPTEVLTEEVVLKTGDLQPLEQPTSEAVEAPLVGTPVCINGLMLEIKDTEK
YCALAPNMMVTNNTFTLKGGA PTKVTFGDDTVIEVQGYKSVNITFELDERIDKVLNEK
CSAYTVELGTEVNEFACVVADAVIKTLQPVSELLTPLGIDLDEWSMATYYLFDSEGEF
KLASHMYCSFYPPDEDEEEGDCEEEEFEPSTQY EYGTEDDYQGKPLEFGATSAALQPE
EEQEEDWLDDDSQQT VGGQDGEDNQT TTIQTIVEVQPQLEMELTPVVQTIEVNSFSG
YLKLTDNVYIKNADIVEEAKVKPTVVVNAANVYLKHGGGVAGALNKATNNAMQVESD
DYIATNGPLKVGGS CVLSGHNLAKHCLHVVGPVNVKGEDIQLLKSAYENFNQHEVLLA
PLLSAGIFGADPIHSLRVCVDTVRTNVYLA VFDKNLYDKLVSSFLEMKSEKQVEQKIA

EIPKEEVKPFITESKPSVEQRKQDDKKIKACVEEVTTTTLEETKFLTENLLLYIDINGN
LHPDSATLVSDIDITFLKKDAPYIVGDVVQEGVLTAVVIPTKKAGGTTEMLAKALRKV
PTDNYITTYPGQGLNGYTVVEEAKTVLKKCKSAFYILPSIISNEKQEILGTVSWNLREM
LAHAEETRKLMPVCVETKAIVSTIQRKYKGIKIQEGVVDYGARFYFYTSKTTVASLIN
TLNDLNETLVTMPLGYVTHGLNLEEAARYMRSCLKVPATVSVSSPDAVTAYNGYLTSSS
KTPEEHFIETISLAGSYKDWSYSGQSTQLGIEFLKRGDKSVYYTSNPTTFHLDGEVIT
FDNLKTLTSLREVRTIKVFTTVDNINLHTQVVDMSMTYGQQFGPTYLDGADVTKIKPH
NSHEGKTFYVLPNDTLRVEAFEYHYHTTDPNFLGRYMSALNHTKKWKYPQVNGLTSIK
WADNNCYLATALLTLOQIELKFNPPALQDAYYRARAGEAANFCALILAYCNKTVGELG
DVRETMSYLFQHANLDSCKRVLNVVCKTCGQQQTTLKGVEAVMYMGTLSEYQFKKGVQ
IPCTCGKQATKYLVOQESPFVMMSSAPPAQYELKHGTFTCASEYTGNYQCGHYKHITSK
ETLYCIDGALLTKSSEYKGPITDVFYKENSYTTTTIKPVTYKLDGVVCTEIDPKLDNYY
KKDNSYFTEQPIDLVPNQYPYNASFDNFKFVCDNIKFADDLNQLTGYKKPASRELKVT
FFPDLNGDVVAIDYKHYTPSFKKGAKLLHKPIVWHVNNATNKATYKPNWTCIRCLWST
KPVETSNSFDVLKSEDAQGMNDLACEDLKPVSEEVVENPTIQKDVLECNVKTTEVVGD
IILKPANNSLKITEEVGHGHTDLMAAYVDNSSLTIKKPNELSRVLGLKTLATHGLAAVNS
VPWDTIANYAKPFLNKVVSTTTNIVTRCLNRVCTNYMPYFFTTTTLLQLCTFTRSTNSRI
KASMPPTIAKNTVKSVMGKFCLEASFNYLKS PNF SKLINIIIFLLLLSVCLGSLIYSTA
ALGVLMSNLGMPSYCTGYREGYLNSTNVTIATYCTGSI PCSVCLSGLDSDLTYP SLET
IQITISSFKWDLTAFGLVAEWFLAYILFTRFFYVLGLAAIMQLFFSYFAVHFISNSWL
MWLIINLVQMAPISAMVRMYIFFASFYVWKS YVHVVDGCNSSTCMMCYKRN RATRVE
CTTIVNGVRRSFYVYANGGKGFCKLHNWNCVNCDTFCAGSTFISDEVARDLSLQFKRP
INPTDQSSYIVDSVTVKNGSIHLYFDKAGQKTYERHSLSHFVNLDNLRANNTKGS LPI
NVIVFDGKSKCEE SAKSASVYYSQLMCQPILLLDQALVSDVGDSA EVAVKMF DAYVN
TFSSTFNVPMEK LKTLVATAEAE LAKNVSLDNVLSTFISAARQGFVDS DVETKD VVEC
LKLSHQSDIEVTGDSCNNYMLTYNKVENMTPRDLGACIDCSARHINAQVAKSHNIALI
WNVKDFMSLSEQLRKQIRSAAKNNLPFKLTCATTRQVVNVVTTKIALKGGKIVNNWL
KQLIKVTLVFLFVA AIFYLITPVHVMSKHTDFSSEIIGYKAIDGGVTRDIASDTDCFA

NKHADFDTWFSQRGGSYTNDKACPLIAAVITREVGFFVPGPLPGTILRRTTNGDFLHFLP
RVFSAVGNICYTPSKLIEYTDFAVSACVLAECTIFKDasGKVPYCYDTNVLEGSVA
YESLRPDTRYVLMdGSIIQFPNTYLEGSVRVVTTFDSEYCRHGTCERSEAGVCVSTSG
RWVLNNDYYRSLPGVFCGVDAVNLLTNMFTPLIQPIGALDISASIVAGGIVAIIVTCL
AYYFMRFRRAFGEYSHVAFNTLLFLMSFTVLCCLTPVYSFLPGVYSVIYLYLTFYLTN
DVSFLAHIQWMMVMTPLVPFWITIAIYIICISTKHfYWFFSNYLKRRVVFNGVSFSTFE
EAALCTFLLNKEMYLKLRSDVLLPLTQYNRYLALYNKYKYFSGAMDTTSYREAACCHL
AKALNDFSNSGSDVLYQPPQTSITSAVLQSGFRKMAFSPGKVEGCMVQVTCGTTTLNG
LWLDDVVYCPRHVICTSEDMLNPNYEDLLIRKSNHNFLVQAGNVQLRVIGHSMQNCVL
KLKVDtanPKTPKYKFVRIQPGQTFsvLACYNGSPSGVYQCAMRPNFTIKGSFLNGSC
GSVGFNIDYDCVSFCYMHMELPTGVHAGTDLEGNFYGPFVDRQTAQAAGTDTTITVN
VLAWLYAAVINGDRWFLNRFTTTLNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAV
LDMCASLKELLQNGMNGRTILGSALLEDEFTPFdvVRQCSGVTfQSAVKRTIKGTHHW
LLLTILTSLLVLVQSTQWSLFFFLYENAFLPFAMGI IAMSafAMMFVKHKHAFCLCLFL
LPSLATVAYFNMVYMPASWVMRIMTWLDMVDTSLSGFKLKDCVMYASAVVLLILMTAR
TVYDDGARRVWTLMNVLTLVYKVYYGNALDQAISMWALI ISVTSNYSGVVTVMFLAR
GIVFMCVEYCPiFFITGNTLQCIMLVYCFLGyFCTCYFGLFCLLNRYFRLTLGVYDYL
VSTQEFryMNSQGLLPKNSIDAFKLNiKLLGVGGKPCIKVATVQSKMSDVKCTSVVL
LSVLQQLRVESSSKLWAQCVQLHNDILLAKDTTEAFEKMVSLLSVLLSMQGAVDINKL
CEEMLDNRATLQAIASEFSSLPsYAAFATAQEAYEQAVANGDSEVVlKklKksLNVAK
SEFDRDAAMQRKLEKMAQAMTQMYKQARSEDKRAKVTsAMQTMLFTMLRKLNDALN
NIINNARDGCvPLNIIPLTTAAKLMVVI PDYNTYKNTCDGTTFTYASALWEIQQVVDA
DSKIVQLSEISMDNSPNLAWPLIVTALRANSaVKLQNNELSPVALRQMSCAAGTTQTA
CTDDNALAYYNTTKGRFVLALLSDLQDLKWARFPKSDGTGTIYTELEPPCRFVTDTP
KGPVKYLYFIKGLNlnRGMVLGSLAATVRLQAGNATEVPANSTVLSFCAFAVDAAK
AYKDYLASGGQPItNCVkmLCTHTGTGQAITVTPEANMDQESFGGASCCLYCRCHIDH
PNPKGFCDLKGKYVQIPttCANDPVGFtLKNTVCTVCGMWKGYGCSCDQLREPMLQSA

[mat_peptide](#)

DAQSFLNGFAV"
266..805
/gene="ORF1ab"

[mat peptide](#) /locus_tag="GU280_gp01"
 /product="leader protein"
 /note="nsp1; produced by both ppla and pplab"
 /protein_id="[YP_009742608.1](#)"
 806..2719
 /gene="ORF1ab"
 /locus_tag="GU280_gp01"
 /product="nsp2"
 /note="produced by both ppla and pplab"
 /protein_id="[YP_009742609.1](#)"
 2720..8554
 /gene="ORF1ab"
 /locus_tag="GU280_gp01"
 /product="nsp3"
 /note="former nsp1; conserved domains are: N-terminal acidic (Ac), predicted phosphoesterase, papain-like proteinase, Y-domain, transmembrane domain 1 (TM1), adenosine diphosphate-ribose 1''-phosphatase (ADRP); produced by both ppla and pplab"
 /protein_id="[YP_009742610.1](#)"
 8555..10054
 /gene="ORF1ab"
 /locus_tag="GU280_gp01"
 /product="nsp4"
 /note="nsp4B_TM; contains transmembrane domain 2 (TM2); produced by both ppla and pplab"
 /protein_id="[YP_009742611.1](#)"
 10055..10972
 /gene="ORF1ab"
 /locus_tag="GU280_gp01"
 /product="3C-like proteinase"
 /note="nsp5A_3CLpro and nsp5B_3CLpro; main proteinase (Mpro); mediates cleavages downstream of nsp4. 3D structure of the SARSr-CoV homolog has been determined (Yang et al., 2003); produced by both ppla and pplab"
 /protein_id="[YP_009742612.1](#)"
 10973..11842
 /gene="ORF1ab"
 /locus_tag="GU280_gp01"
 /product="nsp6"
 /note="nsp6_TM; putative transmembrane domain; produced
 by both ppla and pplab"
 /protein_id="[YP_009742613.1](#)"
 11843..12091
 /gene="ORF1ab"
 /locus_tag="GU280_gp01"
 /product="nsp7"
 /note="produced by both ppla and pplab"
 /protein_id="[YP_009742614.1](#)"
 12092..12685
 /gene="ORF1ab"
 /locus_tag="GU280_gp01"
 /product="nsp8"
 /note="produced by both ppla and pplab"
 /protein_id="[YP_009742615.1](#)"
 12686..13024
[mat peptide](#)

```

/gene="ORF1ab"
/locus_tag="GU280_gp01"
/product="nsp9"
/note="ssRNA-binding protein; produced by both ppla and
pplab"
/protein_id="YP\_009742616.1"
13025..13441
mat peptide /gene="ORF1ab"
/locus_tag="GU280_gp01"
/product="nsp10"
/note="nsp10_CysHis; formerly known as growth-factor-
like protein (GFL); produced by both ppla and pplab"
/protein_id="YP\_009742617.1"
13442..13480
mat peptide /gene="ORF1ab"
/locus_tag="GU280_gp01"
/product="nsp11"
/note="produced by ppla only"
/protein_id="YP\_009725312.1"
13476..13503
stem loop /gene="ORF1ab"
/locus_tag="GU280_gp01"
/inference="COORDINATES:
profile:Rfam-release-14.1:RF00507,Infernal:1.1.2"
/function="Coronavirus frameshifting stimulation element
stem-loop 1"
13488..13542
stem loop /gene="ORF1ab"
/locus_tag="GU280_gp01"
/inference="COORDINATES:
profile:Rfam-release-14.1:RF00507,Infernal:1.1.2"
/function="Coronavirus frameshifting stimulation element
stem-loop 2"
21563..25384
gene /gene="S"
/locus_tag="GU280_gp02"
/gene_synonym="spike glycoprotein"
/db_xref="GeneID:43740568"
21563..25384
CDS /gene="S"
/locus_tag="GU280_gp02"
/gene_synonym="spike glycoprotein"
/note="structural protein; spike protein"
/codon_start=1
/product="surface glycoprotein"
/protein_id="YP\_009724390.1"
/db_xref="GeneID:43740568"

```

```

/translation="MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFR
SSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIR
GWIFGTTLDSKTQSLILVNNATNVVIVKCEVFQFCNDPFLGVYYHKNNKSWMESEFRVY
SSANNCTFEYVSQPFLMDLEGGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPO

```

GFSALEPLVDLPIGINITRFQTLALHRSYLTPGDSSSGWTAGAAAYVGYLQPRFTL
LKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITN
LCPFGEVFNATRFASVYAWNKRKISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCF
TNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYN
YLYRLFRKSNLKPFERDISTEIQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPY
RVVVLSEFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFG
RDIADTTDAVRDPQTLLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAI
HADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPR
RARSVASQSI IAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTM
YICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFG
GFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFN
GLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQN
VLYENQKLIANQFNSAIGKIQDSLSTASALGKLQDVVNQNAQALNTLVKQLSSNFGA
ISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMS
ECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTVPAQEKNFTTAPAICHGKAH
FPREGVFSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELD
SFKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELG
KYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCCLKGCCSCGSCCKFDEDDSE

PVLKGVKLHYT"
gene 25393..26220
/gene="ORF3a"
/locus_tag="GU280_gp03"
/db_xref="GeneID:43740569"
CDS 25393..26220
/gene="ORF3a"
/locus_tag="GU280_gp03"
/codon_start=1
/product="ORF3a protein"
/protein_id="YP_009724391.1"
/db_xref="GeneID:43740569"

/translation="MDLFMRIFTIGTVTLKQGEIKDATPSDFVRATATIPIQASLPFG
WLIVGVALLAVFQSASKIITLKKRWQLALSKGVHVCNLLLLLVTVYSHLLLVAAGLE
APFLYLYALVYFLQSINFVRIIMRLWLCWKCRSKNPLLYDANYFLCWHTNCYDYCIPY
NSVTSSIVITSGDGTTSPISEHDYQIGGYTEKWESGVKDCVVLHSYFTSDYYQLYSTQ

LSTDTGVEHVTFFFIYNKIVDEPEEHVQIHTIDGSSGVVNPVMEPIYDEPTTTTTSVPL"

gene 26245..26472
/gene="E"
/locus_tag="GU280_gp04"
/db_xref="GeneID:[43740570](#)"
CDS 26245..26472
/gene="E"
/locus_tag="GU280_gp04"
/note="ORF4; structural protein; E protein"
/codon_start=1
/product="envelope protein"
/protein_id="[YP_009724392.1](#)"
/db_xref="GeneID:[43740570](#)"

/translation="MYSFVSEETGTLIVNSVLLFLAFVVFLVTLAILTALRLCAYCC
NIVNVSLVKPSFYVYSRVKKNLSSSRVPDLLV"

gene 26523..27191
/gene="M"
/locus_tag="GU280_gp05"
/db_xref="GeneID:[43740571](#)"
CDS 26523..27191
/gene="M"
/locus_tag="GU280_gp05"
/note="ORF5; structural protein"
/codon_start=1
/product="membrane glycoprotein"
/protein_id="[YP_009724393.1](#)"
/db_xref="GeneID:[43740571](#)"

/translation="MADSNGTITVEELKKLLEQWNLVIGFLFTWICLLQFAYANRNR
FLYIIKLIIFLWLLWPVTLACFVLAAYRINWITGGIAIAMACLVGLMMLSIFYFIASFRL
FARTSRMWSFNPETNILLNVPLHGHTILTRPLLESELVIGAVILRGHLRIAGHHLGRCD
IKDLPKEITVATSRTLSSYYKLGASQRVAGDSGFAAYSRYRIGNYKLNTDHSSSSDNIA
LLVQ"

gene 27202..27387
/gene="ORF6"
/locus_tag="GU280_gp06"
/db_xref="GeneID:[43740572](#)"
CDS 27202..27387
/gene="ORF6"
/locus_tag="GU280_gp06"
/codon_start=1
/product="ORF6 protein"
/protein_id="[YP_009724394.1](#)"
/db_xref="GeneID:[43740572](#)"

/translation="MFHLVDFQVTIAEILLIIMRTFKVSIWNLDYIINLIKKNLSKSL
TENKYSQLDEEQPMEID"

gene 27394..27759
/gene="ORF7a"
/locus_tag="GU280_gp07"
/db_xref="GeneID:[43740573](#)"
CDS 27394..27759

/gene="ORF7a"
/locus_tag="GU280_gp07"
/codon_start=1
/product="ORF7a protein"
/protein_id="[YP_009724395.1](#)"
/db_xref="GeneID:[43740573](#)"

/translation="MKIILFLALITLATCELYHYQECVRGTTVLLKEPCSSGTYESGNS

PFHPLADNKFALTCFSTQFAFACPDGVKHVYQLRARSVSPKLFIRQEEVQELYSPIFL
IVAAIVFITLCFTLKRKTE"

[gene](#)

27756..27887
/gene="ORF7b"
/locus_tag="GU280_gp08"
/db_xref="GeneID:[43740574](#)"

[CDS](#)

27756..27887
/gene="ORF7b"
/locus_tag="GU280_gp08"
/codon_start=1
/product="ORF7b"
/protein_id="[YP_009725318.1](#)"
/db_xref="GeneID:[43740574](#)"

/translation="MIELSLIDFYLCFLAFLLLFLVLIIMLIIFWFSLELQDHNETCHA"

[gene](#)

27894..28259
/gene="ORF8"
/locus_tag="GU280_gp09"
/db_xref="GeneID:[43740577](#)"

[CDS](#)

27894..28259
/gene="ORF8"
/locus_tag="GU280_gp09"
/codon_start=1
/product="ORF8 protein"
/protein_id="[YP_009724396.1](#)"
/db_xref="GeneID:[43740577](#)"

/translation="MKFLVFLGIITTVAAFHQECSLQSQCTQHQPVVDDPCPIHFYSK

WYIRVGARKSAPLIELCVDEAGSKSPIQYIDIGNYTVSCLPFTINCQEPKLGSLVVR
SFYEDFLEYHDVRVVLDFI"

[gene](#)

28274..29533
/gene="N"
/locus_tag="GU280_gp10"
/db_xref="GeneID:[43740575](#)"

[CDS](#)

28274..29533
/gene="N"
/locus_tag="GU280_gp10"
/note="ORF9; structural protein"
/codon_start=1
/product="nucleocapsid phosphoprotein"
/protein_id="[YP_009724397.2](#)"
/db_xref="GeneID:[43740575](#)"

/translation="MSDNGPQNQRNAPRITFGGSPDSTGNSQNGERSGARSKQRRPQG

LPNNTASWFTALTQHGKEDLKFPRGQGVPIINTNSSPDDQIGYYRRATRRIRGGDGKMK

DLSPRWYFYLLGTGPEAGLPYGANKDGI IIVVATEGALNTPKDHIGTRNPANNAIIVLQ
 LPQGTTLPKGFYAEGSRGGSQASSRSSSRSRNSSRNSTPGSSRGTS PARMAGNGGDAA
 LALLLLDRLNQLESKMSGKGQQQQGQTVTKKSAAEASKKPRQKRTATKAYNVTQAFGR
 RGPEQTQGNFGDQELIRQGTDYKHWPOIAQFAPSASAFFGMSRIGMEVTPSGTWLTYT
 GAIKLDDKDPNFKDQVILLNKHIDAYKTFPPTEPKKDKKKKADETOALPQRQKKQOTV

TLLPAADLDDFSKQLQSSMSSADSTQA"
 gene 29558..29674
 /gene="ORF10"
 /locus_tag="GU280_gp11"
 /db_xref="GeneID:43740576"
 CDS 29558..29674
 /gene="ORF10"
 /locus_tag="GU280_gp11"
 /codon_start=1
 /product="ORF10 protein"
 /protein_id="YP_009725255.1"
 /db_xref="GeneID:43740576"
 /translation="MGYINVFAFPFTIYSLLLCRMNSRNYIAQVDVVNFNLT"
 stem loop 29609..29644
 /gene="ORF10"
 /locus_tag="GU280_gp11"
 /inference="COORDINATES:
 profile::Rfam-release-14.1:RF00165,Infernal:1.1.2"
 /function="Coronavirus 3' UTR pseudoknot stem-loop 1"
 stem loop 29629..29657
 /gene="ORF10"
 /locus_tag="GU280_gp11"
 /inference="COORDINATES:
 profile::Rfam-release-14.1:RF00165,Infernal:1.1.2"
 /function="Coronavirus 3' UTR pseudoknot stem-loop 2"
 3'UTR 29675..29903
 stem loop 29728..29768
 /inference="COORDINATES:
 profile:Rfam-release-14.1:RF00164,Infernal:1.1.2"
 /note="basepair exception: alignment to the Rfam model
 implies coordinates 29740:29758 form a noncanonical C:T
 basepair, but the homologous positions form a highly
 conserved C:G basepair in other viruses, including SARS
 (NC_004718.3)"
 /function="Coronavirus 3' stem-loop II-like motif (s2m)"

ORIGIN

```

1 attaaagggtt tataccttcc caggtaacaa accaaccaac tttcgatctc ttgtagatct
61 gttctctaaa cgaactttaa aatctgtgtg gctgtcactc ggctgcatgc ttagtgcact
121 cacgcagtat aattaataac taattactgt cgttgacagg acacgagtaa ctcgtctatc
181 ttctgcaggc tgcttacggt ttcgtccgtg ttgcagccga tcatcagcac atctagggtt
241 cgtccgggtg tgaccgaaag gtaagatgga gaggccttgtc cctggtttca acgagaaaac
301 acacgtccaa ctcagtttgc ctgttttaca ggttcgcgac gtgctcgtac gtggctttgg
361 agactccgtg gaggaggtct tatcagaggc acgtcaacat cttaaagatg gcacttgtgg
421 cttagtagaa gttgaaaaag gcgttttgcc tcaacttgaa cagccctatg tgttcatcaa
481 acgttcggat gctcgaactg cacctcatgg tcatgttatg gttgagctgg tagcagaact
541 cgaaggcatt cagtacggtc gtagtggtga gacacttggg gtccttgtcc ctcatgtggg
601 cgaaatacca gtggcttacc gcaaggttct tcttcgtaag aacggtaata aaggagctgg

```

661 tggccatagt tacggcgccg atctaaagtc atttgactta ggcgacgagc ttggcactga
721 tccttatgaa gattttcaag aaaactggaa cactaaacat agcagtggtg ttaccctgta
781 actcatgcbt gagcttaacg gaggggcata cactcgctat gtcgataaca acttctgtgg
841 ccctgatggc taccctcttg agtgcattaa agaccttcta gcacgtgctg gtaaagcttc
901 atgcactttg tccgaacaac tggactttat tgacactaag aggggtgtat actgctgccg
961 tgaacatgag catgaaattg cttggtacac ggaacgttct gaaaagagct atgaattgca
1021 gacacctttt gaaattaaat tggca**aagaa** atttgacacc ttcaatgggg aatgtccaaa
1081 ttttgtatth cccttaaatt ccataatcaa gactattcaa ccaagggttg **aaaagaa**aaa
1141 gcttgatggc tttatgggta gaattcgatc tgtctatcca gttgcbgtcac caaatgaatg
1201 caaccaaagtg tgcctttcaa ctctcatgaa gtgtgatcat tgtggtgaaa cttcatggca
1261 gacgggcbgat tttgttaaag ccacttgcbga attttgtggc actgagaatt tgactaaaga
1321 aggtgccact acttgtggtt acttacccca aaatgctggt gttaaaaatt attgtccagc
1381 atgtcacaat tcagaagtag gacctgagca tagtcttgcc gaataccata atgaatctgg
1441 cttgaaaacc attcttcgta agggtggtcg cactattgcc tttggaggct gtgtgttctc
1501 ttatgttggg tgccataaca agtgtgccta ttgggttcca cbgtgtagcg ctaacatagg
1561 ttgtaaccat acaggtggtg ttggagaagg ttccgaaggc cttaatgaca accttcttga
1621 aatactccaa aaagagaaaag tcaacatcaa tattgttggg gactttaaac ttaatgaaga
1681 gatcbgccatt attttggcat ctttttctgc ttccacaagt gcttttgtgg aaactgtgaa
1741 aggtttggat tataaagcat tcaaacaaat tgttgaatcc tgtggtaatt ttaaagttac
1801 aaaaggaaaa gctaaaaaag gtgcctggaa tattggtgaa cagaaatcaa tactgagtc
1861 tctttatgca tttgcatcag aggtgctcg tgtgtacga tcaattttct ccbgactct
1921 tgaaactgct caaaattctg tgcgtgtttt acagaaggcc gctataacaa tactagatgg
1981 aatttcacag tattcactga gactcattga tgctatgatg ttcacatctg atttggctac
2041 taacaatcta gttgtaatgg cctacattac aggtgggtgt gttcagttga cttcgcagtg
2101 gctaaactaac atctttggca ctgtttatga aaaactcaaa ccbgtccttg attggettga
2161 agagaagttt aaggaagggt tagagtttct tagagacggg tgggaaattg ttaaatttat
2221 ctcaacctgt gcttgtgaaa ttgtcgggtg acaaattgtc acctgtgcaa aggaaattaa
2281 ggagagtgtt cagacattct ttaagcttgt aaataaattt ttggctttgt gtgctgactc
2341 tatcattatt ggtggagcta aacttaaagc cttgaattta ggtgaaacat ttgtcagca
2401 ctcaaaagga ttgtacagaa agtgtgttaa atccagagaa gaaactggcc tactcatgcc
2461 tctaaaagcc ccaa**aagaa** ttatcttctt agaggagaa acacttccca cagaagtgtt
2521 aacagaggaa gttgtcttga aaactgggtg tttacaacca ttagaacaac ctactagtga
2581 agctgtttaa gctccattgg ttggtacacc agtttgtatt aacgggctta tgttgctcga
2641 aatcaaagac acagaaaagt actgtgcctt tgcacctaat atgatggtaa caaacaatac
2701 cttcacactc aaaggcgggt caccaacaaa ggttactttt ggtgatgaca ctgtgataga
2761 agtgcaaggc tacaagagtg tgaatatcac ttttgaactt gatgaaagga ttgataaagt
2821 acttaatgag aagtgtctct cctatacagt tgaactcggg acagaagtaa atgagttcgc
2881 ctgtgttgtg gcagatgctg tcataaaaaac tttgcaacca gtatctgaat tacttacacc
2941 actgggcatt gatttagatg agtggagtat ggctacatac tacttatttg atgagttcgg
3001 tgagtttaaa ttggcttcac atatgtattg ttctttctac cctccagatg aggatgaaga
3061 agaaggatgat tgtg**aagaa** aagagtttga gccatcaact caatatgagt atgggtactga
3121 agatgattac caaggtaaac ctttggattt tgggtgccact tctgctgctc ttcaacctga
3181 agaagagcaa **gaaga**agatt ggttagatga tgatagtcaa caaactgttg gtcaacaaga
3241 cggcagtgag gacaatcaga caactactat tcaaacaatt gttgaggttc aacctcaatt
3301 agagatggaa cttacaccag ttgttcagac tattgaagtg aatagtttta gtggttattt
3361 aaaacttact gacaatgtat acattaaaaa tgcagacatt gtgga**aagaa** ctaaaaaggt
3421 aaaaccaaca gtggttggtta atgcagccaa tgtttacctt aaacatggag gagggttgc
3481 aggaacctta aataaggcta ctaacaatgc catgcaagtt gaatctgatg attcatagc
3541 tactaagga ccacttaaag tgggtggtag ttgtgtttta agcbgacaca atcttcttaa
3601 aactgtctt catgttgtcg gcccaaatgt taacaaaggc gaagacattc aacttcttaa
3661 gagtgcctt gaaaatttta atcagcacga agttctactt gcaccattat tatcagctgg
3721 tatttttggg gctgacccta tacattcttt aagagtttgt gtagatactg ttcgcacaaa
3781 tgtctactta gctgtctttg ataaaaatct ctatgacaaa cttgtttcaa gctttttgga
3841 aatgaagagt gaaaagcaag ttgaacaaaa gatcbgtgag attcctaaag aggaagttaa
3901 gccatttata actgaaagta aaccttcagt tgaacagaga aaacaagatg at**aagaa**aat
3961 caaagcttgt gttg**aagaa** ttacaacaac tctgga**aagaa** actaagttcc tcacagaaaa
4021 cttgttactt tatattgaca ttaatggcaa tcttcatcca gattctgcca ctcttgttag

4081 tgacattgac atcactttct ta**aagaa**aga tgctccatat atagtgggtg atgttgttca
4141 agaggggtgtt ttaactgctg tggttatacc tactaaaaag gctgggtggca ctactgaaat
4201 gctagcgaag gctttgagaa aagtgccaac agacaattat ataaccactt acccgggtca
4261 ggggtttaaag ggttacactg tagaggaggc aaagacagtg cttaaaaagt gtaaaagtgc
4321 cttttacatt ctaccatcta ttatctctaa tgagaagcaa gaaattcttg gaactgtttc
4381 ttggaatttg cgagaaatgc ttgcacatgc ag**aagaa**aca cgcaaattaa tgctgtctg
4441 tgtggaaact aaagccatag tttcaactat acagcgtaaa tataagggta ttaaaataca
4501 agaggggtgtg gttgattatg gtgctagatt ttacttttac accagtaaaa caactgtagc
4561 gtcacttatac aacacactta acgatctaaa tgaaactctt gttacaatgc cacttgggta
4621 tgtaacacat ggcttaaatt tgg**aagaa**gc tgctcgggat atgagatctc tcaaagtgcc
4681 agctacagtt tctgtttctt cacctgatgc tgttacagcg tataatgggt atcttacttc
4741 ttcttctaaa acacctgaag aacattttat tgaaacctac tcacttgctg gttcctataa
4801 agattgggtcc tattctggac aatctacaca actagggtata gaatttctta agagaggtga
4861 taaaagtgtg tattacacta gtaatcctac cacattccac ctagatgggt aagttatcac
4921 ctttgacaat cttaagacac ttctttcttt gagagaagtg aggactatta aggtgtttac
4981 aacagtagac aacattaacc tccacacgca agttgtggac atgtcaatga catatggaca
5041 acagtttggt ccaacttatt tggatggagc tgatgttact aaaataaac ctcataatc
5101 acatgaaggt aaaacatfff atgttttacc taatgatgac actctacgtg ttgaggcttt
5161 tgagtactac cacacaactg atcctagttt tctgggtagg tacatgtcag cattaatca
5221 cactaaaaag tggaaatacc cacaagttaa tgggttact tctattaaat gggcagataa
5281 caactgttat cttgccactg cattgttaac actccaaca atagagttga agtttaatcc
5341 acctgctcta caagatgctt attacagagc aagggctggg gaagctgcta acttttgtgc
5401 acttatctta gcctactgta ataagacagt aggtgagtta ggtgatgta gagaacaat
5461 gagttacttg tttcaacatg ccaatttaga ttcttgcaaa agagtcttga acgtgggtg
5521 taaaacttgt ggacaacagc agacaaccct taaggggtgta gaagctgta tgtacatggg
5581 cacactttct tatgaacaat tt**aagaa**agg tgttcagata ccttgtagct gtggtaaaca
5641 agctacaaaa tatctagtac aacaggagtc accttttgtt atgatgtcag caccacctgc
5701 tcagtatgaa cttaagcatg gtacatttac ttgtgctagt gactacactg gtaattacca
5761 gtgtgggtcac tataaacata taacttctaa agaaactttg tattgcatag acgggtgctt
5821 acttacaag tcctcagaat acaaaggtcc tattacggat gttttctaca **aagaa**acag
5881 ttacacaaca accataaaac cagttactta taaattggat ggtgttggtt gtacagaat
5941 tgaccctaag ttggacaatt attat**aagaa** agacaattct tatttcacag agcaaccaat
6001 tgatcttgta ccaaaccaac catatccaaa cgcaagcttc gataatttta agtttgtatg
6061 tgataatata aaatttgctg atgattttaa ccagttactt ggttataaga aacctgcttc
6121 aagagagctt aaagttacat ttttccctga cttaaattgg gatgtgggtg ctattgatta
6181 taaacactac acacctctt tt**aagaa**agg agctaaattg ttacataaac ctattgtttg
6241 gcatgttaac aatgcaacta ataaagccac gtataaacca aatacctggg gtatacgttg
6301 tctttggagc acaaaaccag ttgaaacatc aaattcgttt gatgtactga agtcagagga
6361 cgcgagggga atggataatc ttgcctgcca agatctaaaa ccagtctctg **aagaa**gtagt
6421 ggaaaatcct accatacaga aagacgttct tgagtgtaat gtgaaaacta ccgaagttgt
6481 aggagacatt atacttaaac cagcaaataa tagtttataa attacagaag aggttggcca
6541 cacagatcta atggctgctt atgtagacaa ttctagtctt actattaaga aacctaatga
6601 attatctaga gtattaggtt tgaaaaccct tgctactcat ggttttagct ctgttaatag
6661 tgtcccttgg gatactatag ctaattatgc taagcctttt cttacaaaag ttgttagtac
6721 aactactaac atagttacac ggtgttttaa ccgtgtttgt actaattata tgccttattt
6781 ctttacttta ttgctacaat tgtgtacttt tactagaagt acaaattcta gaattaaagc
6841 atctatgccg actactatag ca**aagaa**tac tgttaaagat gtcggtaaat tttgtctaga
6901 ggcttcattt aattatttga agtcaccta tttttctaaa ctgataaata ttataatttg
6961 ttttttacta ttaagtgttt gcctaggttc tttaatctac tcaaccgctg ctttaggtgt
7021 tttaatgtct aatttaggca tgccttctta ctgtagtggg tacagagaag gctatttgaa
7081 ctactactaat gtcactattg caacctactg tactggttct ataccttga gtgtttgtct
7141 tagtgggtta gattctttag acacctatcc ttcttttaga actatacaaa ttaccatttc
7201 atctttttaa tgggatttaa ctgcttttgg cttagttgca gagtgggttt tggcatatat
7261 tcttttctact aggtttttct atgtacttgg attggctgca atcatgcaat tgtttttcag
7321 ctatttttga gtacatttta ttagtaattc ttggcttatg tggttaataa ttaatcttgt
7381 acaaatggcc ccgatttcag ctatgggttag aatgtacatc ttctttgcat cattttatta
7441 tgtatggaaa agttatgtgc atgtttaga cggttgta tcatcaactt gtatgatgtg

7501 ttacaaacgt aatagagcaa caagagtcga atgtacaact attgttaatg gtgttagaag
7561 gtcctttttat gtctatgcta atggaggtaa aggccttttgc aaactacaca attggaattg
7621 tgttaattgt gatacattct gtgctggtag tacatttatt agtgatgaag ttgcgagaga
7681 cttgtcacta cagtttaaaa gaccaataaa tcctactgac cagtcttctt acatcgttga
7741 tagtgttaca gtg**aagaa**tg gttccatcca tctttacttt gataaagctg gtcaaaagac
7801 ttatgaaaga cattctctct ctcattttgt taacttagac aacctgagag ctaataacac
7861 taaaggttca ttgcctatta atgttatagt ttttgatggt aatcaaaat gtg**aagaa**tc
7921 atctgcaaaa tcagcgtctg tttactacag tcagcttatg tgtcaaccta tactgttact
7981 agatcaggca ttagtgtctg atgttgggta tagtgcgga gttgcagtta aaatgtttga
8041 tgcttacggt aatacgtttt catcaacttt taacgtacca atggaaaaac tcaaaacact
8101 agttgcaact gcagaagctg aacttgcaaa gaatgtgtcc ttagacaatg tcttatctac
8161 ttttatttca gcagctcggc aagggtttgt tgattcagat gtagaacta aagatgttgt
8221 tgaatgtcct aaattgtcac atcaatctga catagaagtt actggcgata gttgtaataa
8281 ctatatgctc acctataaca aagttgaaaa catgacacc cgtgaccttg gtgcttztat
8341 tgactgtagt gcgcgtcata ttaatgcgca ggtagcaaaa agtcacaaca ttgctttgat
8401 atggaacggt aaagatttca tgtcattgtc tgaacaacta cgaaaacaaa tacgtagtgc
8461 tgctaaaaag aataaacttac cttttaagtt gacatgtgca actactagac aagttgttaa
8521 tgttgtaaca acaaagatag cacttaaggg tggtaaaaatt gtttaataatt ggttgaagca
8581 gtttaattaaa gttacacttg tgttctttt tgttgctgct attttctatt taataacacc
8641 tgttcatgtc atgtctaaac atactgactt ttcaagtga atcataggat acaaggctat
8701 tgatgggtgg gtcactcgtg acatagcatc tacagatact tgttttgcta acaaacatgc
8761 tgattttgac acatggttta gccagcgtgg tggtagttat actaatgaca aagcttgccc
8821 attgattgct gcagtcataa caagagaagt gggttttgtc gtgcctgggt tgcctggcac
8881 gatattacgc acaactaatg gtgacttttt gcatttctta cctagagttt ttagtgcagt
8941 tggtaacatc tgttacacac catcaaaact tatagagtac actgactttg caacatcagc
9001 ttgtgttttg gctgctgaat gtacaatttt taaagatgct tctggtaagc cagtaccata
9061 ttgttatgat accaatgtac tagaaggttc tgttgcttat gaaagtttac gccctgacac
9121 acgttatgtg ctcatggatg gctctattat tcaatttctt aacacctacc ttgaaggttc
9181 tgttagagtg gtaacaactt ttgattctga gtactgtagg cacggcactt gtgaaagatc
9241 agaagctggc gtttgtgtat ctactagtga tagatgggta cttacaatg attattacag
9301 atctttacca ggagttttct gtggtgtaga tgcgtgaaat ttacttacta atatttttac
9361 accactaatt caacctattg gtgctttgga catatcagca tctatagtag ctggtgggat
9421 tgtagctatc gtagtaacat gccttgccca ctattttatg aggtttagaa gagcttttgg
9481 tgaatacagt catgtagtgt cctttaatac tttactatc cttatgtcat tcaactgact
9541 ctgttttaaca ccagtttact cattcttacc tgggtgtttat tctgttattt acttgtactt
9601 gacattttat cttactaatg atgtttcttt tttagcacat attcagtgga tggttatggt
9661 cacaccttta gtacctttct ggataacaat tgcttatatc atttgtattt ccacaaagca
9721 tttctattgg ttcttttagta attacctaaa gagacgtgta gtccttfaat gtgtttcctt
9781 tagtactttt **aagaa**gctg cgctgtgcac ctttttgta aata**aagaa** tgtatctaaa
9841 gttgcgtagt gatgtgctat tacctcttac gcaatataat agatacttag ctctttataa
9901 taagtacaag tatttttagt gagcaatgga tacaactagc tacagagaag ctgcttgggt
9961 tcatctcgca aaggctctca atgacttcag taactcaggt tctgatgttc tttaccaacc
10021 accacaaacc tctatcacct cagctgtttt gcagagtggg tttagaaaaa tggcattccc
10081 atctggtaaa gttgaggggt gtatggtaca agtaacttgt ggtacaacta cacttaacgg
10141 tctttggcct gatgacgtag tttactgtcc aagacatgtg atctgcacct ctgaagacat
10201 gcttaaccct aattatgaag atttactcat tcgtaagtct aatcataatt tcttggta
10261 ggctggtaat gttcaactca gggttattgg acattctatg caaaattgtg tacttaagct
10321 taaggttgat acagccaatc ctaagacacc taagtataag tttgttcgca tttcaaccag
10381 acagactttt tcagtgttag cttgttacia tggttcacca tctgggtgtt accaatgtgc
10441 tatgaggccc aatttcaact ttaaggggtc attccttaat ggttcatgtg gtatggtgg
10501 ttttaacata gattatgact gtgtctcttt ttgttacatg caccatattg aattaccaac
10561 tggagttcat gctggcacag acttagaagg taacttttat ggaccttttg ttgacaggca
10621 aacagcacia gcagctggta cggacacaaac tattacagtt aatgttttag cttgggtgta
10681 cgctgctggt ataaatggag acaggtgggt tctcaatcga tttaccacia ctcttaatga
10741 ctttaacctt gtggctatga agtacaatta tgaacctcta acacaagacc atgttgacat
10801 actaggacct ctttctgctc aaactggaat tgccgtttta gatatgtgtg cttcattaaa
10861 agaattactg caaaatggta tgaatggacg taccatattg ggtagtgtct tattagaaga

10921 tgaatttaca ccttttgatg ttgttagaca atgctcaggt gttactttcc aaagtgcagt
10981 gaa**aagaa**ca atcaagggta cacaccactg gttgttactc acaattttga cttcactttt
11041 agtttttagtc cagagtactc aatgggtcttt gttctttttt ttgtatgaaa atgccttttt
11101 accttttgct atgggtatta ttgctatgtc tgcttttgca atgatgtttg tcaaacataa
11161 gcatgcattt ctctgtttgt ttttgttacc ttctcttgcc actgtagctt attttaatat
11221 ggtctatatg cctgctagtt ggggtgatgcg tattatgaca tggttggata tggttgatac
11281 tagtttgtct ggttttaagc taaaagactg tgttatgtat gcatcagctg tagtgttact
11341 aatccttatg acagc**aagaa** ctggtgatga tgatgggtgct aggagagtgt ggacacttat
11401 gaatgtcttg acactcgttt ataaagtta ttatggtaat gcttttagatc aagccatttc
11461 catgtgggct cttataatct ctgttacttc taactactca ggtgtagtta caactgtcat
11521 gtttttggcc agaggtattg tttttatgtg tgttgagtat tgccctattt tcttcataac
11581 tggaatata cttcagtgta taatgctagt ttattgtttc ttaggctatt tttgtacttg
11641 ttactttggc ctcttttggt tactcaaccg ctactttaga ctgactcttg gtgtttatga
11701 ttacttagtt tctacacagg agtttagata tatgaattca cagggactac tcccacccaa
11761 gaatagcata gatgccttca aactcaacat taaattggtg ggtgttggtg gcaaaccctg
11821 tatcaaagta gccactgtac agtctaaaat gtcagatgta aagtgcacat cagtagtctt
11881 actctcagtt ttgcaacaac tcagagtaga atcatcatct aaattgtggg ctcaatgtgt
11941 ccagttacac aatgacattc tcttagctaa agatactact gaagcctttg aaaaaatggt
12001 ttcactactt tctgttttgc tttccatgca ggggtgctgta gacataaaca agctttgtga
12061 agaaatgctg gacaacaggg caaccttaca agctatagcc tcagagtta gttcccttcc
12121 atcatatgca gcttttgcta ctgctcaaga agcttatgag caggctgttg ctaatggtga
12181 ttctgaagtt gttcttaaaa agttg**aagaa** gtctttgaat gtggctaagt ctgaatttga
12241 ccgtgatgca gccatgcaac gtaagttgga aaagatggct gatcaagcta tgacccaaat
12301 gtataaacag gctagatctg aggacaagag ggcaaaagtt actagtgcta tgcagacaat
12361 gcttttccact atgcttagaa agttggataa tgatgcactc aacaacatta tcaacaatgc
12421 aagagatggg tgtgttccct tgaacataat acctcttaca acagcagcca aactaatggt
12481 tgtcatacca gactataaca catataaaaa tacgtgtgat ggtacaacat ttacttatgc
12541 atcagcattg tgggaaatcc aacaggttgt agatgcagat agtaaaattg ttcaacttag
12601 tgaaattagt atggacaatt cacctaattt agcatggcct cttattgtaa cagctttaag
12661 ggccaattct gctgtcaaat tacagaataa tgagcttagt cctgttgcac tacgacagat
12721 gtcttgtgct gccggtaacta cacaaactgc ttgcaactgac gacaatgcgt tagcttacta
12781 caacacaaca aagggaggta ggtttgtact tgcactgta tccgatttac aggatttgaa
12841 atgggctaga ttccctaaga gtgatggaac tggtaactatc tatacagaac tggaaaccacc
12901 ttgtaggttt gttacagaca cacctaaagg tccctaaagt aagtatttat actttattaa
12961 aggattaaac aacctaaata gaggtatggt acttggtagt ttagctgcca cagtagtctt
13021 acaagctggg aatgcaacag aagtgcctgc caattcaact gtattatctt tctgtgcttt
13081 tgctgtagat gctgctaaag cttacaaaga ttatctagct agtgggggac aaccaatcac
13141 taattgtggt aagatgttgt gtacacacac tggtaactggt caggcaataa cagttacacc
13201 ggaagccaat atggatcaag aatcctttgg tgggtgcatcg tgttgtctgt actgccgttg
13261 ccacatagat catccaaatc ctaaaggatt ttgtgactta aaaggtaagt atgtacaaat
13321 acctacaact tgtgctaatt accctgtggg ttttacactt aaaaacacag tctgtaccgt
13381 ctgcggtatg tggaaagggt atggctgtag ttgtgatcaa ctccgcgaac ccatgcttca
13441 gtcagctgat gcacaatcgt ttttaaaccg gtttgcggtg taagtgcagc ccgtcttaca
13501 ccgtgcgga caggcactag tactgatgtc gtatacaggg cttttgacat ctacaatgat
13561 aaagtagctg gttttgctaa attcctaaaa actaattggt gtcgcttcca agaaaaggac
13621 gaagatgaca atttaattga ttcttacttt gtagttaaga gacacacttt ctctaactac
13681 caacatgaag aaacaattta taatttactt aaggattgtc cagctggtgc taaacatgac
13741 ttctttaagt ttagaataga cggtgacatg gtaccacata tatcagtca acgtcttact
13801 aaatacaciaa tggcagacct cgtctatgct ttaaggcatt ttgatgaagg taattgtgac
13861 acattaaaag aaatacttgt cacatacaat tgtttgtgat atgattatth caataaaaag
13921 gactggtagt atttttaga aaaccagat atattacgcg tatacgcaa cttagggtgaa
13981 cgtgtacgcc aagctttggt aaaaacagta caattctgtg atgceatgcg aaatgctggt
14041 attgttggtg tactgacatt agataatcaa gatctcaatg gtaactggta tgatttcggt
14101 gatttcatac aaaccacgcc aggtagtgga gttcctgttg tagattctta ttattcattg
14161 ttaatgccta tattaacctt gaccagggct ttaactgcag agtcacatgt tgacactgac
14221 ttaacaaagc cttacattaa gtgggatttg ttaaaatatg acttcacgga agagagggtta
14281 aaactctttg accgttattt taaatattgg gatcagacat accacccaaa ttgtgttaac

14341 tgttttggatg acagatgcat tctgcattgt gcaaacttta atgtttttatt ctctacagtg
14401 ttcccaccta caagtttttgg accactagtg agaaaaatat ttgtttgatgg tgtttccattt
14461 gtagtttcaa ctggatacca cttcagagag ctaggtgttg tacataatca ggatgtaaac
14521 ttacatagct ctagacttag ttttaaggaa ttacttgtgt atgctgctga ccctgctatg
14581 cacgctgctt ctggtaatct attactagat aaacgcacta cgtgcttttc agtagctgca
14641 cttactaaca atgtttgcttt tcaaactgtc aaacccggtt attttaacaa agacttctat
14701 gactttgctg tgtctaaggg tttctttaag gaaggaagtt ctgttgaatt aaaacacttc
14761 ttctttgctc aggatggtaa tgcctgtatc agcgattatg actactatcg ttataatcta
14821 ccaacaatgt gtgatatcag acaactacta tttgtagtgg aagttgttga taagtacttt
14881 gattgttacg atgggtggctg tattaatgct aaccaagtca tcgtcaacaa cctagacaaa
14941 tcagctgggtt ttccatttaa taaatggggat aaggctagac tttattatga ttcaatgagt
15001 tatgaggatc aagatgcact tttcgcatac acaaaacgta atgtcatccc tactataact
15061 caaatgaatc ttaagtatgc cattaagtga **aagaa**tagag ctgcaccgtg agctgggtgc
15121 tctatctgta gtactatgac caatagacag tttcatcaaa aattattgaa atcaatagcc
15181 gccactagag gagctactgt agtaattgga acaagcaaat tctatgggtgg ttggcacaac
15241 atgttaaaaa ctgtttatag tgatgtagaa aaccctcacc ttatggggttg ggattatcct
15301 aaatgtgata gagccatgcc taacatgctt agaattatgg cctcacttgt tcttgctcgc
15361 aacatacaa cgtgtttgtag cttgtcacac cgtttctata gattagctaa tgagtgtgct
15421 caagtattga gtgaaatggg catgtgtggc ggttcactat atgttaaacc aggtggaacc
15481 tcatcaggag atgccacaac tgcttatgct aatagtgttt ttaacatttg tcaagctgct
15541 acggccaatg ttaatgcact tttatctact gatggtaaca aaattgccga taagtatgct
15601 cgcaatttac aacacagact ttatgagtgt ctctatagaa atagagatgt tgacacagac
15661 tttgtgaatg agttttacgc atatttgcgt aaacatttct caatgatgat actctctgac
15721 gatgctggtg tgtgtttcaa tagcacttat gcatctcaag gtctagtggc tagcataaag
15781 aactttaagt cagttcttta ttatcaaaac aatgttttta tgcctgaagc aaaatggttg
15841 actgagactg accttactaa aggacctcat gaattttgct ctcaacatac aatgctagtt
15901 aaacaggggtg atgattatgt gtaccttctc taccagatc catc**aagaa**t cctagggggcc
15961 ggctgtttttg tagatgatat cgtaaaaaca gatggtacac ttatgattga acggttcgctg
16021 tcttttagcta tagatgctta cccacttact aaacatccta atcaggagta tgctgatgct
16081 tttctattgt acttacaata cat**aagaa**ag ctacatgatg agttaacagg acacatggtta
16141 gatcatgtatt ctggtatgct tactaactgat aacacttcaa ggtattggga acctgagttt
16201 tatgaggcta tgtacacacc gcatacagtc ttacaggctg ttggggccttg tgttctttgc
16261 aattcacaga cttcattaag atgtgggtgct tgcatacgtg gaccattctt atgttgtaaa
16321 tgcctgttacg accatgtcat atcaacatca cataaattag tcttgtctgt taatccgtat
16381 gtttgcaatg ctccagggtg tgatgtcaca gatgtgactc aactttactt aggaggatg
16441 agctattatt gtaaatcaca taaaccacc attagttttc cattgtgtgc taatggacaa
16501 gtttttgggtt tatataaaaa tacatgtgtt ggtagcgata atgttactga ctttaatgca
16561 attgcaacat gtgactggac aaatgctggg gattacattt tagctaacac ctgtactgaa
16621 agactcaagc tttttgcagc agaaacgctc aaagctactg aggagacatt taaactgtct
16681 tatgggtattg ctactgtacg tgaagtgtcg tctgacagag aattacatct ttcattggaa
16741 gttggtaaac ctagaccacc acttaaccga aattatgtct ttactgggta tcgtgtaact
16801 aaaaacagta aagtacaaat aggagagtac acctttgaaa aaggtgacta tgggtgatgct
16861 gttgtttacc gaggtacaac aacttacaac ttaaatgttg gtgattattt tgcctgaca
16921 tcacatacag taatgccatt aagtgcacct acactagtgc cacaagagca ctatgttaga
16981 attactggct tatacccaac actcaatatc tcagatgagt tttctagcaa tgggtgcaat
17041 tatcaaaagg ttggtatgca aaagtattct acactccagg gaccacctgg tactggtaag
17101 agtcattttg ctattggcct agctctctac taccctctcg ctgcatagt gtatacagct
17161 tgcctctcatg ccgctgttga tgcactatgt gagaaggcat taaaatattt gcctatagat
17221 aaatgtagta gaattatacc tgcacgtgct cgtgtagagt gttttgataa attcaagtg
17281 aattcaacat tagaacagta tgtcttttgt actgtaaag cattgcctga gacgacagca
17341 gatatagttg tctttgatga aatttcaatg gccacaaatt atgatttgag tgtgtcaat
17401 gccagattac gtgctaagca ctatgtgtac attggcgacc ctgctcaatt acctgcacca
17461 cgcacattgc taactaaggg cacactagaa ccagaatatt tcaattcagt gtgtagactt
17521 atgaaaacta taggtccaga catgttctctc ggaacttgtc ggcgttgtcc tgcgtgaaat
17581 gttgacactg tgagtgtttt ggtttatgat aataagctta aagcacataa agacaaatca
17641 gctcaatgct ttaaaatggt ttataagggt gttatcacgc atgatgtttc atctgcaatt
17701 aacaggccac aaataggcgt ggttaagagaa ttccttacac gtaaccctgc ttggagaaaa

17761 gctgtcttta tttcacctta taattcacag aatgctgtag cctcaaagat tttgggacta
17821 ccaactcaaa ctggttgatc atcacagggc tcagaatatg actatgtcat attcactcaa
17881 accactgaaa cagctcactc ttgtaatgta aacagattta atggttgctat taccagagca
17941 aaagtaggca tactttgcat aatgtctgat agagaccttt atgacaagtt gcaatttaca
18001 agtcttgaaa ttccacgtag gaatgtggca actttacaag ctgaaaatgt aacaggactc
18061 tttaaagatt gtagtaaggt aatcactggg ttacatccta cacaggcacc tacacacctc
18121 agtgttgaca ctaaattcaa aactgaaggt ttatgtggtg acatacctgg catacctaag
18181 gacatgacct atagaagact catctctatg atgggtttta aatgaatta tcaagttaat
18241 ggttacccta acatgtttat caccgcgaa gaagctataa gacatgtacg tgcattggatt
18301 ggcttcgatg tcgaggggtg tcatgctact agagaagctg ttggtacca tttaccttta
18361 cagctaggtt tttctacagg tgtaaaccta gttgctgtac ctacaggtta tgttgataca
18421 cctaataata cagatttttc cagagttagt gctaaccac cgctggaga tcaatttaaa
18481 cacctcatac cacttatgta caaaggactt ccttggatg tagtgcgtat aaagattgta
18541 caaatgttaa gtgacacact taaaaatctc tctgacagag tcgtatttgt cttatgggca
18601 catggctttg agttgacatc tatgaagtat tttgtgaaaa taggacctga gcgcacctgt
18661 tgtctatgtg atagacgtgc cacatgcttt tccactgctt cagacactta tgctgtttgg
18721 catcattcta ttggatttga ttacgtctat aatccgttta tgattgatgt tcaacaatgg
18781 ggtttttacag gtaacctaca aagcaaccat gatctgtatt gtcaagtcca tggtaatgca
18841 catgtagcta gttgtgatgc aatcatgact aggtgtctag ctgtccacga gtgctttggt
18901 aagcgtgttg actggactat tgaatatcct ataattgggtg atgaactgaa gattaatgcg
18961 gctttagtaa aggttcaaca catggttgtt aaagctgcat tattagcaga caaattccca
19021 gttcttcacg acattggtaa ccctaaagct attaatgtgt tacctcaagc tgatgtagaa
19081 tgggaagttct atgatgcaca gccttgtagt gacaaagctt ataaaataga agaattattc
19141 tattcttatg ccacacattc tgacaaattc acagatgggtg tatgcctatt ttggaattgc
19201 aatgtcgata gatatcctgc taattccatt gtttgtagat ttgacactag agtgcctatc
19261 aaccttaact tgctgggtg tgatgggtggc agtttgtatg taaataaaca tgcattccac
19321 acaccagctt ttgataaaaag tgcttttgtt aatttaaaac aattaccatt tttctattac
19381 tctgacagtc catgtgagtc tcatggaaaa caagtagtgt cagatataga ttatgtacca
19441 ctaaagtctg ctacgtgtat aacacgttgc aatttaggtg gtgctgtctg tagacatcat
19501 gctaaatgag acagattgta tctcagtct tataacatga tgatctcagc tggctttagc
19561 ttgtgggttt acaacaatt tctacttat aacctctgga acacttttac aagacttcag
19621 agtttagaaa atgtggcttt taatgttga aataagggac actttgatgg acaacaggtg
19681 gaagtaccag tttctatcat taataacact gtttacacaa aagttgatgg tgttgatgta
19741 gaattgtttg aaaataaaac aacattacct gttaatgtag catttgagct ttgggctaag
19801 cgcaacatta aaccagtacc agaggtgaaa atactcaata atttgggtgt ggacattgct
19861 gctaatactg tgatctggga ctacaaaaga gatgctccag cacatatac tactattggt
19921 gtttgttcta tgactgacat agcc**aagaa** ccaactgaaa cgatttgtgc accactcact
19981 gtcttttttg atggtagagt tgatgggtcaa gtagacttat ttagaaatgc ccgtaatggt
20041 gttcttatta cagaaggtag tgtaaaaggt ttacaacat ctgtaggtcc caaacaagct
20101 agtcttaatg gagtcacatt aattggagaa gccgtaaaaa cacagttcaa ttattataag
20161 aaagttgatg gtgttgcaca acaattacct gaaacttact ttactcagag tagaaattta
20221 **caagaa**ttta aaccaggag tcaaattgaa attgatttct tagaattagc tatggatgaa
20281 ttcattgaac ggtataaatt agaaggtat gccttcgaac atatcgttta tggagatttt
20341 agtcatagtc agttaggtgg tttacatcta ctgattggac tagctaaacg ttttaaggaa
20401 tcaccttttg **aattagaaga** **ttttattcct** **atggacagta** **cagttaaaa** **ctatttcata**
20461 **acagatgcgc** **aaacaggttc** **atctaagtg** **gtgtgtctg** **ttattgatt** **attacttgat**
20521 **gattttgttg** **aaataataaa** **atccaagat** **ttatctgtag** **tttctaaggt** **tgtcaaagt**
20581 **actattgact** **atacagaaat** **ttcatttatg** **cttttggtgta** **aagatggcca** **tgtagaaca**
20641 **ttttaccaa** **aattacaatc** **tagtcaagcg** **tggaaccgg** **gtgttgctat** **gcctaattc**
20701 **tacaaaatgc** **aaagaatgct** **attagaaaag** **tgtgacctc** **aaaattatgg** **tgatagtgca**
20761 **acattacct** **aaggcataat** gatgaatgtc gcaaaatata ctcaactgtg tcaatattta
20821 aacacattaa cattagctgt accctataat atgagagtta tacattttgg tgcgtgttct
20881 gataaaggag ttgcaccagg tacagctgtt ttaagacagt ggttgectac ggttacgctg
20941 cttgtcgatt cagatcttaa tgactttgtc tctgatgcag attcaacttt gattggtgat
21001 tgtgcaactg tacatacagc taataaatgg gatctcatta ttagtgatat gtacgacctc
21061 aagactaaaa atgttacaaa agaaaatgac tctaaagagg gt**ttttcac** **ttacatttgt**
21121 **gg**gtttatac aacaaaagct agctcttggg ggttccgtgg ctataaagat aacagaacat

21181 tcttggaatg ctgatcttta taagctcatg ggacacttgc catgggtggac agcctttggt
21241 actaatgtga atgcgcatc atctgaagca tttttaattg gatgtaatta tcttgccaaa
21301 ccacgcgaac aaatagatgg ttatgtcatg catgcaaatt acatattttg gaggaataca
21361 aatccaattc agttgtcttc ctattcttta **tttgacatga gtaaatttcc ccttaaatta**
21421 **aggggtactg ctgttatgtc tttaaaagaa ggtcaaatca atgatatgat tttatctctt**
21481 **ccttagtaaag gttagacttat aattagaga aacaaacagag ttgttatttc tagtgatggt**
21541 **cctgttaaca actaaacgaa caatgtttgt tttcttggtt ttattgccac tagtctctag**
21601 tcagtgtggt aatcttacia ccagaactca attaccctct gcatacacta attctttcac
21661 acgtgggtgt tattaccctg acaaagtttt cagatcctca gttttacatt caactcagga
21721 cttgtttctta cctttctttt ccaatgttac ttggttccat gctatacatg tctctgggac
21781 caatgggtact aagagggttg ataaccctgt cctaccattt aatgatgggtg tttattttgc
21841 ttccactgag aagtctaaca taataagagc ctggattttt ggtactactt tagattcgaa
21901 gcccagttcc ctacttattg ttaataacgc tactaatggt gttattaaag tctgtgaatt
21961 tcaattttgt aatgatccat ttttgggtgt ttattaccac aaaaacaaca aaagttggat
22021 ggaaagtgag ttcagagttt attctagtgc gaataattgc acttttgaat atgtctctca
22081 gccttttctt atggacctg aaggaaaaca gggtaatttc aaaaatctta gggaaattgt
22141 gtttaagaaat attgatgggt attttaaaat atattctaag cacacgccta ttaatttagt
22201 gcgtgatctc cctcaggggt tttcggcttt agaaccattg gtagatttgc caataggat
22261 taacatcact aggtttcaaa ctttacttgc tttacataga agttatttga ctctgggtga
22321 ttctttctca ggttggacag ctgggtgctgc agcttattat gttgggttctc tccaacctag
22381 gacttttcta ttaaaatata atgaaaatgg aaccattaca gatgctgtag actgtgcaact
22441 tgaccctctc tcagaaacaa agtgtacggt gaaatccttc actgtagaaa aaggaatcta
22501 tcaaacttct aacttttagag tccaaccaac agaactctatt gttagatttc ctaatattac
22561 aaacttgtgc ccttttgggt aagtttttaa cgccaccaga tttgcatctg tttatgcttg
22621 gaacaggaag agaactcagca actgtgttgc tgattattct gtcctatata attccgcctc
22681 attttccact ttttaagtgt atggagtgtc tctactaaa ttaaagatgc tctgctttac
22741 taatgtctat gcagattcat ttgtaattag aggtgatgaa gtcagacaaa tcgctccagg
22801 gcaaactgga aagattgctg attataatta taaattacca gatgatttca cagggtgctg
22861 tatagctttg aattctaaca atcttgattc taaggttgggt ggaattata attacctgta
22921 tagattggtt aggaagtcta atctcaaacc ttttgagaga gatatttcaa ctgaaatcta
22981 tcagccgggt agcacacct gtaatgggtg tgaagggttt aattgttact ttcctttaca
23041 atcatatggt ttccaacca ctaatgggtg tggttaccaa ccatacagag tagtagtact
23101 ttcttttgaa cttctacatg caccagcaac tgtttgtgga cctaaaaagt ctactaattt
23161 ggttaaaaaa aatgtgtca atttcaact caatgggtta acaggcacag gtgttcttac
23221 tgagtctaac aaaaagtttc tgcctttcca acaatttggc agagacattg ctgacactac
23281 tgatgctgtc cgtgatccac agacacttga gattcttgac attacaccat gttcttttgg
23341 tgggtgtcagt gttataaac caggaacaaa tacttctaac cagggtgctg tcttttatca
23401 ggatggtaac tgcacagaag tccctgttgc tattcatgca gatcaactta ctctacttg
23461 gcgtgtttat tctacaggtt ctaatgtttt tcaaacacgt gcaggctggt taataggggc
23521 tgaacatgtc aacaactcat atgagtgtga cataccatt ggtgcaggta tatgcgctag
23581 ttatcagact cagactaatt ctcctcggcg ggcacgtagt gtagctagtc aatccatcat
23641 tgcctacact atgtcacttg gtgcagaaaa ttcagttgct tactctaata actctattgc
23701 catacccaca aattttacta ttagtgttac cacagaaatt ctaccagtgt ctatgaccaa
23761 gacatcagta gattgtacaa tgtacatttg tgggtgattca actgaaatgca gcaatctttt
23821 gttgcaatat ggcagttttt gtacacaatt aaaccgtgct ttaactggaa tagctgttga
23881 acaagacaaa aacacccaag aagtttttgc acaagtcaaa caaatttaca aaacaccacc
23941 aattaaagat tttgggtggt ttaatttttc acaaataa cagatccat caaaaccaag
24001 caagaaatca tttattgaag atctactttt caacaaagtg acactgtcag atgtgtgctt
24061 catcaaaca tatgggtgatt gccttgggtg tattgctgct agagacctca tttgtgcaca
24121 aaagtttaac ggccttactg ttttgcacc tttgctcaca gatgaaatga ttgctcaata
24181 cacttctgca ctgttagcgg gtacaatcac ttctggttgg accttgggtg cagggtgctgc
24241 attacaaata ccatttgcta tgcaaatggc ttataggttt aatgggtattg gagttacaca
24301 gaatgttctc tatgagaacc aaaaattgat tgccaaccaa tttaatagtg ctattggcaa
24361 aattcaagac tcactttctt ccacagcaag tgcacttggg aaacttcaag atgtgggtcaa
24421 ccaaaatgca caagctttaa acacgcttgt taacaactt agctccaatt ttgggtgcaat
24481 ttcaagtgtt ttaaatgata tcttttcacg tcttgacaaa gttgaggctg aagtgcaaat
24541 tgatagggtg atcacaggca gacttcaag tttgcagaca tatgtgactc aacaattaat

24601 tagagctgca gaaatcagag cttctgctaa tcttgcctgct actaaaatgt cagagtgtgt
24661 acttggacaa tcaaaaagag ttgatttttg tggaaagggc tatcatctta tgtccttccc
24721 tcagtcagca cctcatgggtg tagtcttctt gcattgtgact tatgtccctg cacaagaaaa
24781 gaacttcaca actgctcctg ccatttgtca tgatggaaaa gcacactttc ctctggaagg
24841 tgtctttggt tcaaatggca cacactgggt tgtaacacaa aggaattttt atgaaccaca
24901 aatcattact acagacaaca catttgtgtc tggtaactgt gatgttgtaa taggaattgt
24961 caacaacaca gtttatgatc ctttgcaacc tgaattagac tcattcaagg aggagttaga
25021 taaatatttt aagaatcata catcaccaga tgttgattta ggtgacatct ctggcattaa
25081 tgcttcagtt gtaaaccattc aaaaagaaat tgaccgcctc aatgaggttg ccaagaattt
25141 aaatgaatct ctcatcgatc tccaagaact tggaaagtat gagcagtata taaaatggcc
25201 atgggtacatt tggctagggt ttatagctgg cttgattgcc atagtaattg tgacaattat
25261 gctttgctgt atgaccagtt gctgtagttg tctcaagggc tgttgttctt gtggatcctg
25321 ctgcaaattt gatgaagacg actctgagcc agtgcctcaa ggagtcaaat tacattacac
25381 ataaacgaac ttatggattt gtttatgaga atcttcacaa ttggaactgt aactttgaag
25441 caaggtgaaa tcaaggatgc tactccttca gattttgttc gcgctactgc aacgataccg
25501 atacaagcct cactcccttt cggatggctt attgttggcg ttgcacttct tgetgttttt
25561 cagagcgctt ccaaaatcat aaccctcaa aagagatggc aactagcact ctccaaggtt
25621 gttcactttg tttgcaactt gctgttggtt tttgtaacag tttactcaca cttttgtctc
25681 gttgctgctg gccttgaagc cccttttctc tatctttatg ctttagtcta cttcttgcag
25741 agtataaact ttgtagaat aataatgagg ctttggcttt gctggaaatg ccgttccaaa
25801 aaccattac tttatgatgc caactatttt ctttgcctggc atactaattg ttacgactat
25861 tgtatacctt acaatagtgt aacttcttca attgtcatta cttcaggtga tggcacaaca
25921 agtcctatth ctgaacatga ctaccagatt ggtggttata ctgaaaaatg ggaatctgga
25981 gtaaaagact gtgttgattt acacagttac ttcacttcag actattacca gctgtactca
26041 actcaattga gtacagacac tgggtgttgaa catgttacct tcttcateta caataaaatt
26101 gttgatgagc ctgagaaca tgtccaaatt cacacaatcg acggttcatc cggagtgtgt
26161 aatccagtaa tggaaaccaat ttatgatgaa ccgacgacga ctactagect gcctttgtaa
26221 gcacaagctg atgagtacga acttatgtac tcattcgttt cggaaagagac aggtacgtta
26281 atagtttaata gcgtacttct ttttcttctg ttcgtgggtat tcttgcctagt tacactagcc
26341 atccttactg cgcttctgatt gtgtgcgtac tgctgcaata ttgttaactg gactcttcta
26401 aaaccttctt tttacgttta ctctcgtgtt aaaaatctga attcttctag atttctctgat
26461 cttctggctt aaacgaacta aatattatat tagtttttct gtttggact ttaattttag
26521 ccatggcaga ttccaacggg actattaccg ttgaagagct taaaaagctc cttgaacaat
26581 ggaacctagt aataggtttc ctattcctta catggatttg tcttctacaa tttgectatg
26641 ccaacaggaa taggtttttg tatataatta agttaatttt cctctggctg ttatggccag
26701 taacttttagc ttgtttttgt cttgctgctg tttacagaat aaattggatc accggtggaa
26761 ttgctatcgc aatggcttgt cttgtaggct tgatgtggct cagctacttc attgcttctt
26821 tcagactggt tgcgcgtacg cgttccatgt ggtcattcaa tccagaaact aacattcttc
26881 tcaacgtgcc actccatggc actattctga ccagaccgct tctagaaagt gaactcgtaa
26941 tcggagctgt gatccttctg ggacatcttc gtattgctgg acaccatcta ggacgctgtg
27001 acatcaagga cctgcctaaa gaaatcactg ttgctacatc acgaacgctt tcttattaca
27061 aattgggagc ttgcgagcgt gtagcaggtg actcaggttt tgctgcatac agtcgctaca
27121 ggattggcaa ctataaatta aacacagacc attccagtag cagtgacaat attgctttgc
27181 ttgtacagta agtgacaaca gatgtttcat ctctgtgact ttcaggttac tatagcagag
27241 atattactaa ttattatgag gactttttaa gtttccattt ggaatcttga ttacatcata
27301 aacctcataa ttaaaaattt atctaagtca ctaactgaga ataaatattc tcaattagat
27361 gaagagcaac caatggagat tgattaaacg aacatgaaaa ttattctttt cttggcactg
27421 ataacactcg ctacttgtga gctttatcac taccaagagt gtgttagagg tacaacagta
27481 cttttaaag aaccttgctc ttctggaaca tacgagggca attcaccatt tcatcctcta
27541 gctgataaca aatttgcact gacttgcctt agcactcaat ttgcttttgc ttgtcctgac
27601 ggcgtaaaac acgtctatca gttacgtgcc agatcagttt cacctaaact gttcatcaga
27661 caagaggaag ttcagaact ttactctcca atttttctta ttgttgcggc aatagtgttt
27721 ataacacttt gcttcacact caaagaaag acagaatgat tgaactttca ttaattgact
27781 tctatttgtg ctttttagcc tttctgctat tcttgtttt aattatgctt attatctttt
27841 ggttctcact tgaactgcaa gatcataatg aaacttgtca cgcctaaacg aacatgaaat
27901 ttcttgtttt cttaggaatc atcacaactg tagctgcatt tcaccaagaa tgtagtttac
27961 agtcatgtac tcaacatcaa ccatatgtag ttgatgacc gtgtcctatt cacttctatt

28021 ctaaattggta tattagagta ggagctagaa aatcagcacc ttttaattgaa ttgtgcgtgg
28081 atgaggctgg ttctaaatca cccattcagt acatcgatat cggtaattat acagtttccct
28141 gtttaccttt tacaattaat tgccaggaac ctaaattggg tagtcttgta gtgcgttggt
28201 cgttctatga agacttttta gagtatcatg acgttcgtgt tgttttagat ttcattctaaa
28261 cgaacaaact aaaatgtctg ataattggacc caaaaatcag cgaaatgcac cccgcattac
28321 gtttggtgga ccctcagatt caactggcag taaccagaat ggagaacgca gtggggcgcg
28381 atcaaaaaca cgtcggcccc aaggtttacc caataatact gcgtcttggt tcaccgctct
28441 cactcaacat ggcaaggaag accttaaatt ccctcgagga caaggcggtc caattaacac
28501 caatagcagt ccagatgacc aaattggcta ctaccgaaga gctaccagac gaattcgtgg
28561 tgggtgacggg aaaatgaaag atctcagtc aagatgggtat ttctactacc taggaactgg
28621 gccagaagct ggacttccct atgggtgctaa caaagacggc atcatatggg ttgcaactga
28681 gggagccttg aatacaccaa aagatcacat tggcaccgc aatcctgcta acaatgtgc
28741 aatcgtgcta caacttctc aaggaacaac attgccaaaa ggcttctacg cagaagggag
28801 cagaggcggc agtcaagcct cttctcgttc ctcattcacgt agtcgcaaca gttc**aagaa**
28861 ttcaactcca ggcagcagta ggggaacttc tctgctaga atggctggca atggcggta
28921 tgctgctctt gctttgctgc tgcctgacag attgaaccag cttgagagca aaatgtctgg
28981 taaaggccaa caacaacaag gccaaactgt cact**aagaa** tctgctgctg aggttctaa
29041 gaagcctcgg caaaaacgta ctgccactaa agcatacaat gtaacacaag ctttcggcag
29101 acgtggtcca gaacaaacc aaggaaattt tggggaccag gaactaatca gacaaggaac
29161 tgattacaaa cattggccgc aaattgcaca atttgcccc agcgcttcag cgttcttcgg
29221 aatgtcgcgc attggcatgg aagtcacacc ttcgggaacg tggttgacct acacaggtgc
29281 catcaaattg gatgacaaag atccaaattt caaagatcaa gtcattttgc tgaataagca
29341 tattgacgca taaaaacat tcccaccaac agagcctaaa aaggacaaaa ag**aagaa**ggc
29401 tgatgaaact caagccttac cgcagagaca **aagaa**acag caaactgtga ctcttcttcc
29461 tgctgcagat ttggatgatt tctccaaaca attgcaaca tccatgagca gtgctgactc
29521 aactcaggcc taaactcatg cagaccacac aaggcagatg ggctatataa acgttttcgc
29581 ttttcggttt acgatataa gtctactctt gtgcagaatg aattctcgta actacatagc
29641 acaagtagat gtagttaact ttaatctcac atagcaatct ttaatcagtg tgtaacatta
29701 gggaggactt gaaagagcca ccacattttc accgaggcca cgcggagtac gatcgagtgt
29761 acagtgaaca atgctagggg gagctgccta tatggaagag ccctaattgtg taaaattaat
29821 tttagtagtg ctatccccat gtgattttaa tagcttctta ggagaatgac aaaaaaaaaa
29881 aaaaaaaaaa aaaaaaaaaa aaa

//