

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

NCBI Reference Sequence: NC_045512.2

[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; 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="ORFlab polyprotein"
/protein_id="[YP_009724389.1](#)"
/db_xref="GeneID:43740578"

/translation="MESLVPGFNEKTHVQLSLPVLQVRDVLVRGFGDSVVEEVLSEARQ
HLKDGTCGLVEVEKGVLPQLEQPYVFIKRS DARTAPHGHVMVELVAELEGIQYGRSGE
TLGVLVPHVGEIPVAYRKVLLRKNNGKAGGHSYGADLKSFDLGDELGTDPYEDFQEN
WNTKHSSGVTRELMRELNGGAYTRYVDNNFCGPDGYPLECIKDLLARAGKASCTLSEQ
LDFIDTKRGVYCCREHEHEIAWYTERSEKSYELQTPFEIKLAKKFDTFNGECPNFVFP
LNSIIKTIQPRVEKKKLDGFMGRIRSVPVSPNECNQMCLSTLMKCDHCGETSWQTG
DFVKATCEFCGTENLTKEGATTCGYLPQNAVVKIYCPACHNSEVGPESHSLAEYHNESG
LKTILRKGGR TIAFGGCVFSYVGCHNK CAYWVPRASANIGCNHTGVVGESEGLNDNL
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
ITILDGISQYSLRLIDAMMFTSDLATNNLVVMAYITGGVVQLTSQWLTNIFGTVYEKL
KPVLDWLEEKFKEGVEFLRDGWEIVKFI STCACEIVGGQIVTCAKEIKESVQTFFKLV
NKFLALCADS I I I G G A K L K A L N L G E T F V T H S K G L Y R K C V K S R E E T G L L M P L K A P K E I I
FLEGETLPTEVLTEEVLKTDGLQPLEQPTSEAVEAPLVGTPVCINGLMLLEIKDTEK
YCALAPNMMVTNNTFTLKG G A P T 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
CSAYTVELGTEVNEFACVVADAVIKTLQPVSELLTPLGIDLDEWSMATYYLFDSEGEF
KLASHMYCSFYPPDEDEEEEGDCEEEEFEPSTQY E Y G T E D D Y Q G K P L E F G A T S A A L Q P E
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
YLKLTDNVYIKNADIVEEAKKVKPTVVVNAANVYLKHGGGVAGALNKATNNAMQVESD
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
PLLSAGIFGADPIHSLRVCVDTVRTNVYLAVFDKNLYDKLVSSFLEMKSEKQVEQKIA
EIPKEEVKPFITESKPSVEQRKQDDKKIKACVEEVTTTTLEETKFLTENLLLYIDINGN
LHPDSATLVSDIDITFLK KDAPYIVGDVVQEGVLTAVVIPTKKAGGTTEMLAKALRKV
PTDNYITTYPGQLNGYTVVEEAKTVLKKCKSAFYILPSIISNEKQEILGTVSWNLREM
LAHAETRKLMPVCVETKAI V S T I Q R K Y K G I K I 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

KTPEEHFIETISLAGSYKDWSYSGQSTQLGIEFLKRGDKSVYYTNSNPTTFHLDGEVIT
FDNLKTLTLLSLREVRTIKVFTTVDNINLHTQVVDMSMTYGQQFGPTYLDGADVTKIKPH
NSHEGKTFYVLPNDTLRVEAFEYHYHTDPSFLGRYMSALNHTKKWKYPQVNGLTSIK
WADNNCYLATALLTLOQIELKFNPALQDAYYRARAGEAANFCALILAYCNKTVGELG
DVRETMSYLFQHANLDSCKRVLNVVCKTCGQQQTTLKGVEAVMYMGTLSEYQFKKGVQ
IPCTCGKQATKYLVOQESPFVMSAPPAQYELKHGTFTCASEYTGNYQCGHYKHITSK
ETLYCIDGALLTKSSEYKGPITDVFYKENSYTTTTIKPVTYKLDGVVCTEIDPKLDNYY
KKDNSYFTEQPIDLVPNQYPNASFDNFKFVCDNIKFADDLNQLTGYYKPPASRELKVT
FFPDLNGDVVAIDYKHYTPSFKKGAKLLHKPIVWHVNNATNKATYKPNWTWCIRCLWST
KPVETSNSFDVLKSEDAQGMNDLACEDLKPVSEEVVENPTIQKDVLECNVKTTEVVGD
IILKPANNSLKITEEVGHTDLMAAYVDNSSLTIKKPNELSRVLGLKTLATHGLAAVNS
VPWDTIANYAKPFLNKVVSTTTNIVTRCLNRVCTNYMPYFFTLQLCTFTRSTNSRI
KASMPPTIAKNTVKSVMGKFCLEASFNLYLKSPNFSKLINIIIFLILLSVCLGSLIYSTA
ALGVLMNSNLGMPYSYCTGYREGYLNSTNVTIATYCTGSI PCSVCLSGLDSDLTYPSET
IQITISSFKWDLTAFGLVAEWFLAYILFTRFFYVLGLAAIMQLFFSYFAVHFISNSWL
MWLIINLVQMAPI SAMVRMYIFFASFYVWKSIVHVVDGCNSSTCMCYKRNRRATRVE
CTTIVNGVRRSFYVYANGGKGFCKLHNWNCVNCDFCAGSTFISDEVARDLSLQFKRP
INPTDQSSYIVDSVTVKNGSIHLYFDKAGQKTYERHSLSHFVNLDNLRANNTKGSPLI
NVIVFDGKSKCEESSAKSASVYYSQLMCQPIILLDQALVSDVGDSEAVAVKMFDAYVN
TFSSTFNVPMEKLTLVATAEAEALAKNVSLDNVLSFTFISAARQGFVDSVETKDVVEC
LKLSHQSDIEVTGDSCNNYMLTYNKVENMTPRDLGACIDCSARHINAQVAKSHNIALI
WNVKDFMSLSEQLRKQIRSAAKNNLPFKLTCATTRQVVNVVTTKIALKGGKIVNNWL
KQLIKVTLVFLFVAAI FYLITPVHVMSKHTDFSSEIIGYKAIDGGVTRDIASDTDCFA
NKHADFDTWFSQRGGSYTNKACPLIAAVITREVGFFVPLPGTILRTTNGDFLHFLP
RVFSAVGNICYTPSKLIEYTD FATSACVLAAECTIFKDASGKVPYCYDTNVLEGSVA
YESLRPDTRYVLMDSIIQFPNTYLEGSRVVVTTFDSEYCRHGT CERSEAGVCVSTSG
RWVLNNDYYRSLPGVFCGVDVAVNLLTNMFTPLIQPIGALDISASIVAGGIVAIIVTCL
AYYFMRFRRAFGEYSHVVAFNLLFLMSFTVLCLTPVYSFLPGVYSVIYLYLTFYLTN

DVSFLAHIQWMVMFTPLVPFWITIAIYIICISTKHFYWFFSNYLKRRVVFNGVSFSTFE
EAALCTFLLNKEMYLKLRSVLLPLTQYNRYLALYNKYKYFSGAMDTTSYREAACCHL
AKALNDFSNSGSDVLYQPPQTSITSAVLQSGFRKMAFSPGKVEGCMVQVTCGTTTLNG
LWLDDVVYCPRHVICTSEDMLNPNYEDLLIRKSNHNFLVQAGNVQLRVIGHSMQNCVL
KLKVDTANPKTPKYKFVRIQPGQTFSVLACYNGSPSGVYQCAMRPNFTIKGSFLNGSC
GSVGFNIDYDCVSFCYMHMELPTGVHAGTDLEGNFYGPFVDRQTAQAAGTDTTITVN
VLAWLYAAVINGDRWFLNRFTTTLNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAV
LDMCASLKELLQNGMNGRTILGSALLEDEFTPFDDVVRQCSGVTFQSAVKRTIKGTHHW
LLLTIILTSLLVLVQSTQWSLFFFLYENAFPLPFAMGIIAMSAFAMMFVKHKAFLCLFL
LPSLATVAYFNMVYMPASWVMRIMTWLDMVDTSLSGFKLKDCVMYASAVVLLILMTAR
TVYDDGARRVWTLMNVLTLVYKVYYGNALDQAI SMWALI I SVTSNYSGVTTVMFLAR
GIVFMCVEYCPDIFFITGNLQCI MLVYCF LGYFCTCYFGLFCLLNRYFRLTLGVYDYL
VSTQEFRYMNSQGLLPKNSIDAFKLNKLLGVGGKPCIKVATVQSKMSDVKCTSVVL
LSVLQQLRVESSSKLWAQCVQLHNDILLAKDTTEAFEKMSVLLSVLLSMQGAVDINKL
CEEMLDNRATLQAIASEFSSLPSYAAFATAQEAYEQAVANGDSEVVLLKCLKKSLNVAK
SEFDRDAAMQRKLEK MADQAMTQMYKQARSEDKRAKVT SAMQTMLFTMLRKL DNDALN
NIINNARDGCVPLNIIPLTTAAKLMVVI PDYNTYKNTCDGTTFTYASALWEIQQV VDA
DSKIVQLSEI SMDNSPNLAWPLIVTALRANS AVKLQNNELSPVALRQM SCAAGTTQTA
CTDDNALAYYNTTKGGRFVLALLSDLQDLKWARFPKSDGTGTIYTELEPPCRFVTDTP
KGPVKYLYFIKGLNNLNRGMVLGSLAATVRLQAGNATEVPANSTVLSFCAFAVDAAK
AYKDYLASGGQPITNCVKMLCTHTGTGQAITVTPEANMDQESFGGASCCLYCRCHIDH
PNPKGFCDLKGKYVQIPTTCANDPVGFTLKN TVCTVCGMWKGYGCSCDQLREPMLQSA
DAQSFLNRVCGVSAARLTPCGTGTSTDVYRAFDIYNDKVAGFAKFLKTNCCRFQEKD
EDDNLIDSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHDFFKFRIDGDMVPHISRQR
LTKYTMADLVYALRHFDEGNCDTLKEILVTYNCCDDDYFNKKDWYDFVENPDILRVYA
NLGERVRQALLKTVQFCDAMRNAGIVGVLTLDNQLNGNWDYDFGDFIQTTPGSGVPVV
DSYSSLMPILTLTRALTAESHVDTDLTKPYIKWDLKDYDFTEERLKLFD RYFKYWDQ
TYHPNCVNCLDDRCILHCANFNVL FSTVFPPTSFGPLVRKIFVDGVPFVVSTGYHFRE

LGVVHNQDVNLHSSRSLFKELLVYAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQ
TVKPGNFNKDFYDFAVSKGFFKEGSSVELKHFFFAQDGNAAISDYDYRYNLPTMCDI
RQLLFVVEVVDKYFDCYDGGCINANQVIVNNLDKSAGFPFNKKGKARLYYDSMSYEDQ
DALFAYTKRNVIPTITQMNLYAISAKNRARTVAGVSICTMTNRQFHQKLLKSIAT
RGATVVIIGTSKFGYGGWHNMLKTVYSDVENPHLMGWDYPKCDRAMPNMLRIMASLVLAR
KHTTCCSLSHRFYRLANCAQVLSSEMVMCGGSLYVKPGGTSSGDATTAYANSVFNICQ
AVTANVNALLSTDGNKIADKYVRNLQHRLYECLYRNRDVTDFVNEFYAYLRKHFSMM
ILSDDAVVCFNSTYASQGLVASIKNFKSVLYYQNNVFMSEAKCWTETDLTKGPHEFCS
QHTMLVKQGDDYVYLPYDPDSRILGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKH
PNQEYADVFLYLQYIRKLDDELTDGHMLDMYSVMLTNDNTSRYWEPEFYEAMYTPHTV
LQAVGACVLCNSQTSLRGACIRRPFLCCKCCYDHVISTSHKLVLSVNPYVCNAPGCD
VTDVTQLYLGMSYYCKSHKPPISFPLCANGQVFGLYKNTCVGSDNVTFNAIATCDW
TNAGDYILANTCTERLKLFAAETLKATEETFKLSYGIATVREVLSRELHLSWEVGKP
RPPLNRNYVFTGYRVTKNSKVQIGEYTFEKGDYGDVVYRGTTTTYKLVNGDYFVLTSH
TVMPLSAPTLVLPQEHYVRITGLYPTLNI SDEFSSNVANYQKVGMQKYSTLQGPPGTGK
SHFAIGLALYYP SARIVYTACSHA AVDALCEKALKYLPIDKCSRIIPARARVECFDKF
KVNSTLEQYVFC TVNALPETTADIVVFEI SMATNYDLSVVNARLR AKHYVYIGDPAQ
LPAPRTLLTKGTLEPEYFNSVCRLMKTIGPDMFLGTCRRCPAEIVDTV SALVYDNKLLK
AHKDKSAQC FKM FYKGVITHDVSSAINRPQIGVVREFLTRNPAWRKAVFISPYN SQNA
VASKILGLPTQTVDSSQ GSEYDYVIFTQT TETAHSCNVNRFNVAITRAKVGILCIMS D
RDLYDKLQFTSLEIPRRNVATLQAENV TGLFKDCSKVITGLHPTQAPTHLSVDTKFKT
EGLCVDIPGIPKDMTYRRLISMMGFKMNYQVNGYPNMFITREEAIRHVRAWIGFDVEG
CHATREAVGTNLPLQLGFSTGVNLVAVPTGYVDT PNNTDFSRVSAKPPPGDQFKHLIP
LMYKGLPWNVVR IKIVQMLSDTLKNLSDRVVFVLWAHGFELTSMKYFVKIGPERTCCL
CDRRATCFSTASDTYACWHHSIGFDYVYNPFMIDVQQWGFTGNLQSNHDLYCQVHGNA
HVASCDAIMTRCLAVHECFVKRVDWTIEYPIIGDELKINAACR KVQHMVVK AALLADK
FPVLHDIGNPKAIKCVPAQADVEWKFYDAQPCSDKAYKIEELFYSYATHSDKFTDGVCL
FWNCNVDRYPANSIVCRFDTRVLSNLNLP GCDGGSLYVNKHAFHTPAFDKSAFVNLKQ

LPFFYYSDSPCESHGKQVVSDIDYVPLKSATCITRCNLGGAVCRHHANEYRLYLDAYN
MMISAGFSLWVYKQFDTYNLWNTFTRLQSLNVAFNVVNKGHFDGQQGEVPSIINNT
VYTKVDGVDVELFENKTTLPVNVAFELWAKRNIKPVEVKILNNLGVDIAANTVIWDY
KRDAPAHISTIGVCSMTDIAKKPTETICAPLTVFFDGRVDGQVDLFRNARNGVLITEG
SVKGLQPSVGPQASLNGVTLIGEAVKTQFNYYKKVDGVVQQLPETYFTQSRNLQEFK
PRSQMEIDFLELAMDEFIERKLEGYAFEHIVYGDFSHSQLGGLHLLIGLAKRFKESP
FELEDFIPMDSTVKNYFITDAQTGSSKCVCSVIDLLLDDFVEI IKSQDLSVVSQVVKV
TIDYTEISFMLWCKDGHVETFPYKQSSQAWQPGVAMPNLYKMQRMLLEKCDLQNYGD
SATLPKGIMMNVAKYTQLCQYLNTLTLAVPYNMRVIHFGAGSDKGVAPGTAVLRQWLP
TGTLVDSLDLNDVSDADSTLIGDCATVHTANKWDLIISDMYDPKTKNVTKENDSKEG
FFTYICGFIQQKLALGGSVAIKITEHSWNADLYKLMGHFAWWTAFVTNVNASSSEAFI

IGCNYLKGPREQIDGYVMHANYIFWRNTNPIQLSSYSLFDMSKFPLKLRGTAVMSLKE
GQINDMILSLLSKGRLIIRENNRVVISSDVLVNN"

[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 ppl1a and ppl1ab"
[mat peptide](#) /protein_id="YP_009725301.1"
 10973..11842
 /gene="ORF1ab"
 /locus_tag="GU280_gp01"
 /product="nsp6"
 /note="nsp6_TM; putative transmembrane domain; produced

by

both ppl1a and ppl1ab"
[mat peptide](#) /protein_id="YP_009725302.1"
 11843..12091
 /gene="ORF1ab"
 /locus_tag="GU280_gp01"
 /product="nsp7"
 /note="produced by both ppl1a and ppl1ab"
[mat peptide](#) /protein_id="YP_009725303.1"
 12092..12685
 /gene="ORF1ab"
 /locus_tag="GU280_gp01"
 /product="nsp8"
 /note="produced by both ppl1a and ppl1ab"
[mat peptide](#) /protein_id="YP_009725304.1"
 12686..13024
 /gene="ORF1ab"
 /locus_tag="GU280_gp01"
 /product="nsp9"
 /note="ssRNA-binding protein; produced by both ppl1a and ppl1ab"
[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 ppl1a and ppl1ab"
[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 ppl1ab 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 ppl1ab 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"
[CDS](#) /protein_id="[YP_009725311.1](#)"
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
TLGLVLPVPHVGEIPVAYRKVLLRKNNGKAGGHSYGADLKSFDLGDELGTDPTYEDFQEN
WNTKHSSGVTRELMRELNGGAYTRYVDNNFCGPDGYPLECIKDLLARAGKASCTLSEQ
LDFIDTKRGVYCCREHEHEIAWYTERSEKSYELQTPFEIKLAKKFDTFNGECPNFVFP
LNSIIKTIQPRVEKKKLDGFMGRIRSVYPVASPNECNQMCLSTLMKCDHCGETSWQTG
DFVKATCEFCGTENLTKEGATTCGYLPQNAVVKIYCPACHNSEVGPEHSLAEYHNESG
LKTILRKGGR TIAFGGCVFSYVGCHNKAYWVPRASANIGCNHTGVVGESEGLNDNL
LEILQKEKVNINIVGDFKLNEEIAIILASFSASTSAFVETVKGLDYKAFKQIVESC GN
FKVTKGKAKKGAWNIGE QKSILSPLYAFASEAARVRSIFSRTLETAQNSVRVLQKAA
ITILDGISQYSLRLIDAMMFTSDLATNNLVVMAYITGGVVQLTSQWLTNIFGTVYEKL
KPVLDWLEEKFKEGVEFLRDGWEIVKFISTCACEIVGGQIVTCAKEIKESVQTFFKLV
NKFLALCADSIIIGGAKLKALNLGETFVTHSKGLYRKCVKSREETGLLMPLKAPKEII
FLEGETLPTEVLTEE VVLKTGDLQPLEQPTSEAVEAPLVGTPVCINGLM LLEIKDTEK
YCALAPNMMVTNNTFTLKG GAPT KVTFGDDTVIEVQGYKSVNITFELDERIDKVLNEK

CSAYTVELGTEVNEFACVVADAVIKTLQPVSELLTPLGIDLDEWSMATYYLFDSEGEF
KLASHMYCSFYPPDEDEEEGDCEEEEFEPSTQYEYGTEDDYQGKPLEFGATSAALQPE
EEQEEDWLDDDSQQTIVGQQDGSQDQTTTIQTIQVIVEVQPQLEMELTPVVQTIIEVNSFSG
YLKLTDNVYIKNADIVEEAKKVKPTVVVNAANVYLKHGGGVAGALNKATNNAMQVESD
DYIATNGPLKVGGSVLSGHNLAKHCLHVVGPNVKNKEDIQLLKSAYENFNQHEVLLA
PLLSAGIFGADPIHSLRVCVDTVRTNVYLAVFDKNLYDKLVSSFLEMKSEKQVEQKIA
EIPKEEVKPFITESKPSVEQRKQDDKIKACVEEVTTTTLEETKFLTENLLLYIDINGN
LHPDSATLVSDIDITFLKKDAPYIVGDVVQEGVLTAVVIPTKKAGGTTEMLAKALRKV
PTDNYITTYPGQGLNGYTVEEAKTVLKKCKSAFYILPSIISNEKQEILGTVSWNLREM
LAHAETRKLMPVCVETKAIVSTIQRKYKGIKIQEGVVDYGARFYFYTSKTTVASLIN
TLNDLNETLVTMPLGYVTHGLNLEEAARYMRSKVPATVSVSSPDAVTAYNGYLTSSS
KTPEEHFIETISLAGSYKDWSYSGQSTQLGIEFLKRGDKSVYYTSNPTTFHLDGEVIT
FDNLKTLTSLREVRTIKVFTTVDNINLHTQVVDMSTYQQQFGPTYLDGADVTKIKPH
NSHEGKTFYVLPNDTLRVEAFEYHYHTTDPFLGRYMSALNHTKKWKYPQVNGLTSIK
WADNNCYLATALLTQQIELKFNPPALQDAYRARAGEAANFCALILAYCNKTVGELG
DVRETMSYLFQHANLDSCKRVLNVVCKTCGQQQTTLKGVEAVMYMGTLSEYQFKKGVO
IPCTCGKQATKYLVOQESPFVMMSSAPPAQYELKHGTFTCASEYTGNYQCGHYKHITSK
ETLYCIDGALLTKSSEYKGPITDVFYKENSYTTTTIKPVTYKLDGVVCTEIDPKLDNYY
KKDNSYFTEQPIDLVPNQYPNASFDNFKFVCDNIKFADDLNQLTGKYPASRELKVT
FFPDLNGDVVAIDYKHYTPSFKKGAKLLHKPIVWHVNNATNKATYKPNWTCIRCLWST
KPVETSNSFDVLKSEDAQGMDNLACEDLKPVSEEVVENPTIQKDVLECNVKTTEVVGD
IILKPANNSLKITEEVGHTDLMAAYVDNSSLTIKKPNELSRVLGLKTLATHGLAAVNS
VPWDTIANYAKPFLNKVVSTTTNIVTRCLNRVCTNYMPYFFTTTTLLQLCTFTRSTNSRI
KASMPPTIAKNTVKSVMGKFCLEASFNYLKS PNF SKLINIIWFLLLSVCLGSLIYSTA
ALGVLMNSNLGMPYCTGYREGYLNSTNVTIATYCTGSI PCSVCLSGLDLDTYPSLET
IQITISSFKWDLTAFGLVAEWFILFTRFFYVLGLAAIMQLFFSYFAVHFISNSWL
MWLIINLVQMAPISAMVRMYIFFASFYVWKS YVHVVDGCNSSTCMMCYKRNRRATVE
CTTIVNGVRRSFYVYANGGKGFCKLHNWNCVNCDTFCAGSTFISDEVARDLSLQFKRP

INPTDQSSYIVDSVTVKNGSIHLYFDKAGQKTYERHSLSHFVNLDNLRANNTKGS LPI
NVIVFDGKSKCEESSAKSASVYYSQLMCQPILLLDQALVSDVGDSA EVAVKMF DAYVN
TFSSTFNVPMEK LKTLVATAEAE LAKNVSLDNV LSTFISAARQGFVDS DVETKDVVEC
LKL SHQSDIEVTGDSCNNYMLTYNKVENMTPRDLGACIDCSARHINAQVAKSHNIALI
WNVKDFMSLSEQLRKQIRSAAKNNLPFKLTCATTRQVVNVVTTKIALKGGKIVNNWL
KQLIKVTLVFLFVA AIFYLITPVHVMSKHTDFSSEIIGYKAIDGGVTRDIAS TDT CFA
NKHADFDTWFSQRGGSYTNDKACPLIAAVITREVG FVVPGLPGTILRTTNGDFLHFLP
RVFSAVGNICYTPSKLIEYTD FATSACV LAAECTIFKDASGKPV PYCYDTNVLEGSVA
YESLRPDTRYV LMDGSI IQFPNTYLEG SVRVVTTFDSEYCRHGT CERSEAGVCVSTSG
RWVLNNDY YRSLPGVFCGVD AVNLLTNMFTPLIQPIGALDISASIVAGGIVAIVVTCL
A YFMRFRRAFGEYSHV VAFNTLLFLMSFTVLCLTPVYSFLPGVYSVIYLYLTFYLTN
DVSFLAHIQWMMVFTPLVPFWITIA YIICISTKH FYWFFSNYLKRRVVFNGVSFSTFE
EAALCTFLLNKEMYLKLRSDVLLPLTQYNRYLALYNKYKYFSGAMD TTSYRE AACCHL
AKALNDFSNSGSDVLYQPPQTSITSAVLQSGFRKMAFP SGKVEGCMVQVTCGTTTLNG
LWLDDVVYCPRHVIC TSEDMLNPNYEDLLIRKSNHNFLVQAGNVQLRVIGHSMQNCVL
KLKVD TANPKTPKYKFVRIQPGQTF SVLACYNGSPSGVYQCAMRPNFTIKGSFLNGSC
GSVGFNIDYDCVSFCYMHMELPTGVHAGTDLEGNFYGPFVDRQTAQAAGTDTTITVN
VLAWLYAAVINGDRWFLNRFTTTLNDFNLVAMKYN YEPLTQDHVDILGPLSAQTGI AV
LDMCASLKELLQNGMNGRTILGSALLEDEFTPF DVVRQCSGVTFQSAVKRTIKGTHHW
LLLTI L TSLLVLVQSTQW S LFFFLYENAF L P F AMGI I AMSAFAMMFVKHKHAF LCLFL
LPSLATVAYFNMVYMPASWVMRIMTWLDMVDTSLSGFKLKDCVMYASAVVLLILMTAR
TVYDDGARRVWTL MNVLT LVYKVYYGNALDQAISMWALIISVTSNYSGVVTTVMFLAR
GIVFMCVEYCP IFFITGN TLQCI MLVYCF LGYFCTCYFGLFCLLNRYFRLTLGVYDYL
VSTQEF RYMNSQGLLPKNSIDAFKLN I KLLGVGGKPCIKVATVQSKMSDVKCTSVVL
LSVLQQLRVES SSKLWAQC VQLHNDILLAKDTTEAF EKMSVLLSVLLSMQGAVDINKL
CEEMLDNRATLQAIASEFSSLPSYAAFATAQEAYEQAVANGDSEVVLK LK LKSLNVAK
SEFDRDAAMQRKLEK MADQAMTQMYKQARSEDKRAKVT SAMQTMLFTMLRKL DNDALN
NIINNARDGCVPLNIIPLTTAAKLMVVI PDYNTYKNTCDGTTFTYASALWEIQQV VDA

DSKIVQLSEIISMDNSPNLAWPLIVTALRANSVAVKLVQNNELSPVALRQMSCAAGTTQTA
CTDDNALAYYNTTKGGRFVLLALLSDLQDLKWARFPKSDGTGTIYTELEPPCRFVTDTP
KGPVKVLYLFIKGLNNLNLRGMVLGSLAATVRLQAGNATEVPANSTVLSFCAFAVDAAK
AYKDYLASGGQPITNCVKMLCTHTGTGQAITVTPEANMDQESFGGASCCLYCRCHIDH
PNPKGFCDLKGGYVQIPTTCANDPVGFTLKNTVCTVCGMWKGYGCSCDQLREPMLQSA

[mat peptide](#) DAQSFLNGFAV"
266..805
/gene="ORF1ab"
/locus_tag="GU280_gp01"
/product="leader protein"
/note="nsp1; produced by both ppla and pplab"
/protein_id="YP_009742608.1"

[mat peptide](#) 806..2719
/gene="ORF1ab"
/locus_tag="GU280_gp01"
/product="nsp2"
/note="produced by both ppla and pplab"
/protein_id="YP_009742609.1"

[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_009742610.1"

[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_009742611.1"

[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_009742612.1"

[mat peptide](#) 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"

[mat peptide](#) 11843..12091
 /gene="ORF1ab"
 /locus_tag="GU280_gp01"
 /product="nsp7"
 /note="produced by both ppla and pplab"
 /protein_id="[YP_009742614.1](#)"

[mat peptide](#) 12092..12685
 /gene="ORF1ab"
 /locus_tag="GU280_gp01"
 /product="nsp8"
 /note="produced by both ppla and pplab"
 /protein_id="[YP_009742615.1](#)"

[mat peptide](#) 12686..13024
 /gene="ORF1ab"
 /locus_tag="GU280_gp01"
 /product="nsp9"
 /note="ssRNA-binding protein; produced by both ppla and pplab"
 /protein_id="[YP_009742616.1](#)"

[mat peptide](#) 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"
 /protein_id="[YP_009742617.1](#)"

[mat peptide](#) 13442..13480
 /gene="ORF1ab"
 /locus_tag="GU280_gp01"
 /product="nsp11"
 /note="produced by ppla only"
 /protein_id="[YP_009725312.1](#)"

[stem loop](#) 13476..13503
 /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"

[stem loop](#) 13488..13542
 /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"

[gene](#) 21563..25384
 /gene="S"
 /locus_tag="GU280_gp02"
 /gene_synonym="spike glycoprotein"
 /db_xref="GeneID:[43740568](#)"

[CDS](#) 21563..25384
 /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
GWIFGTTLDSTQSLIVNNAITNVVIKVEFQFCNDPFLGVYYHKNNKSWMESEFRVY
SSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPO
GFSALEPLVDLPIGINITRFQTLALHRSYLT PGDSSSGWTAGAAAYVGYLQPRFTL
LKYNENGTITDAVDCALDPLSETKCTLKSFVTEKGIYQTSNFRVQPTESIVRFPNITN
LCPFGEVFNATRFASVYAWNKRKISNCVADYSVLYNSASFSTFKCYGVSPTKLNLDLFC
TNVYADSFVIRGDEVQRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGNYN
YLYRLFRRKSNLKPFERDISTEIIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPY
RVVVLSEFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFG
RDIADTTDAVRDPQTLIELDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAI
HADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPR
RARSVASQSIIAYTMSLGAENSVAYSNNIAIPTNFTISVTTEILPVSMTKTSVDCTM
YICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFG
GFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFN
GLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQN
VLYENQKLIANQFNSAIGKIQDSLSTASALGKLQDVVNQNAQALNTLVKQLSSNFGA
ISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMS
ECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVITYVPAQEKNFTTAPAICHGKAH
FPREGVFSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELD
SFKEELDKYFKNHTSPDVLDGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELG
KYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSE

[gene](#) PVLKGVKLYHT"
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
WLIVGVALLAVFQSASKIITLKKRWQLALSkgVHFVCNLLLLFVTVYSHLLLVAAGLE
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="MYSFVSEETGTLIVNSVLLFLAFVVLLVTLAILTALRLCAYCC
NIVNVSLVKPSFYVYSRVKLNLSRVPDLLV"

[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
FLYIIKLIFLWLLWPVTLACFVLAAYRINWITGGIAIAMACLVGLMWLSYFIASFRL
FATRSMWSFNPETNILLNVPLHG TILTRPLLESELVIGAVILRGHLRIAGHHLGRCD
IKDLPKEITVATSRTL SYYKLGASQRVAGDSGFAAYSRYRIGNYKLNTDHSSSSDNIA
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="MFHLVDFQVTIAEILLIIMRTFKVSIWNLDYIINLIKLNLSKSL
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="MKIILFLALITLATCELYHYQECVRGTTVLLKEPCSSGTYEGNS

PFHPLADNKFALTCFSTQFAFACPDGVKHVYQLRARSVSPKLFIRQEEVQELYSPIFL
IVAAIVFITLCTLKRKTE"

[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="MIELSLIDFYLCFLAFLFLVLIIMLIIFWFSLELQDHNETCHA"

[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="MKFLVFLGIITTVAAFHQECSLQSQCTQHQPYPVDDPCPIHFYSK

WYIRVGARKSAPLIELCVDEAGSKSPIQYIDIGNYTVSCLPFTINCQEPKLGSLVVR
SFYEDFLEYHDVRVLDLFI"

[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="MSDNGPQNQRNAPRITFGGPSDSTGSNQNGERSGARSKQRRPQG
LPNNTASWFTALTQHGKEDLKFPRGQGVPIINTNSSPDDQIGYYRRATRIRIRGGDGKMK
DLSRWYFYFYLGTGPEAGLPYGANKDGIWVATEGALNTPKDHIGTRNPANNAIVLQ
LPQGTTLPKGFYAEGSRGGSQASSRSSRSRNSSRNSTPGSSRGTS PARMAGNGGDA
LALLLLDRLNQLESKMSGKGGQQQQGQTVTKKSAAEASKKPRQKRTATKAYNVTQAFGR
RGPEQTQGNFGDQELIRQGTDYKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTYT
GAIKLDDKDPNFKDQVILLNKHIDAYKTFPPTEPKKDKKKKADETQALPQRQKKQQTV
TLLPAADLDDFSKQLQQSMSSADSTQA"

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

stem loop

29675..29903
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 attaaaggtt tataccttcc caggaacaa accaaccaac tttcgatctc ttgtagatct

61 gttctctaaa cgaactttaa aatctgtgtg gctgtcactc ggctgcatgc ttagtgcact
121 cacgcagtat aattaataac taattactgt cgttgacagg acacgagtaa ctctgtctatc
181 ttctgcaggc tgcttacggg ttcgtccgtg ttgcagccga tcatcagcac atctagggtt
241 cgtccgggtg tgaccgaaag gtaagatgga gagccttgtc cctgggttca acgagaaaac
301 acacgtccaa ctgagtttgc ctgttttaca ggttcgcgac gtgctcgtac gtggccttgg
361 agactccgtg gaggagggtc tatcagaggc acgtcaacat cttaaagatg gcacttgtgg
421 cttagtagaa gttgaaaaag gcgttttggc tcaacttgaa cagccctatg tgttcatcaa
481 acgttcggat gctcgaactg cacctcatgg tcatgttatg gttgagctgg tagcagaact
541 cgaaggcatt cagtacggtc gtagtgggtg gacacttggg gtccttgtcc ctcatgtggg
601 cgaataacca gtggccttacc gcaaggttct tcttcgtaag aacggtaata aaggagctgg
661 tggccatagt tacggcgccg atctaaagtc atttgactta ggcgacgagc ttggcactga
721 tccttatgaa gattttcaag aaaactggaa cactaaacat agcagtgggt ttaccctgtg
781 actcatgctg 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 tggcaaagaa atttgacacc ttcaatgggg aatgtccaaa
1081 ttttgtatth cccttaaatt ccataatcaa gactattcaa ccaaggggtg aaaagaaaa
1141 gcttgatggc tttatgggta gaattcgatc tgtctatcca gttgcgtcac caaatgaatg
1201 caaccaaagtg tgcctttcaa ctctcatgaa gtgtgatcat tgtgggtaaa ctcatggca
1261 gacgggcatg tttgttaaag ccacttgcca attttgtggc actgagaatt tgactaaaga
1321 aggtgccact acttgtgggtt acttacccca aaatgctggt gttaaaattt atgtccagc
1381 atgtcacaat tcagaagtag gacctgagca tagtcttggc gaataccata atgaatctgg
1441 cttgaaaacc attcttcgta aggggtggctg cactattgcc tttggaggct gtgtgttctc
1501 ttatgttggg tgccataaca agtgtgcta ttgggttcca cgtgctagcg ctaacatagg
1561 ttgtaaccat acagggtgtg ttggagaagg ttccgaaggc cttaatgaca accttcttga
1621 aatactccaa aaagagaaaag tcaacatcaa tattgttggg gactttaaac ttaatgaaga
1681 gatcgccatt attttggcat ctttttctgc ttccacaagt gcttttgtgg aaactgtgaa
1741 aggtttggat tataaagcat tcaaacaaat tgttgaatcc tgtggtaatt ttaaagttac
1801 aaaagaaaaa gctaaaaaag gtgcctggaa tattgggtgaa cagaaatcaa tactgagctc
1861 tctttatgca tttgcatcag aggtgctcg tgtgtacga tcaattttct cccgactct
1921 tgaactgct caaaattctg tgcgtgtttt acagaaggcc gctataacaa tactagatgg
1981 aatttcacag tattcactga gactcattga tgctatgatg ttcacatctg atttggctac
2041 taacaatcta gttgtaattg cctacattac aggtgggtgtt gttcagttga cttcgcagtg
2101 gctaactaac atctttggca ctgtttatga aaaactcaaa cccgtccttg atttggcttga
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 ttgtcacgca
2401 ctcaaaggga ttgtacagaa agtgtgttaa atccagagaa gaaactggcc tactcatgcc
2461 tctaaaagcc caaaagaaa ttatcttctt agagggagaa acacttccca cagaagtgtt
2521 aacagaggaa gttgtcttga aaactgggtg tttacaacca ttagaacaac ctactagtga
2581 agctgttgaa gctccattgg ttggtacacc agtttgtatt aacgggctta tgttgcctga
2641 aatcaaagac acagaaaagt actgtgccct tgcacctaat atgatggtaa caaacaatac
2701 cttcacactc aaaggcgggt caccaacaaa ggttactttt ggtgatgaca ctgtgataga
2761 agtgcaaggg tacaagagtg tgaatatcac ttttgaactt gatgaaagga ttgataaagt
2821 acttaattgag aagtgtctct cctatacagt tgaactcggg acagaagtaa atgagttcgc
2881 ctgtgttgtg gcagatgctg tcataaaaac tttgcaacca gtatctgaat tacttacacc
2941 actgggcatt gatttagatg agtggagtat ggctacatac tacttatttg atgagtctgg
3001 tgagttttaa ttggcttcac atatgtattg ttctttctac cctccagatg aggatgaaga
3061 agaaggatgat tgtgaagaag aagagtttga gccatcaact caatatgagt atgggtactga
3121 agatgattac caaggtaaac ctttggattt tgggtgccact tctgctgctc ttcaacctga
3181 agaagagcaa gaagaagatt ggttagatga tgatagtcaa caaactgttg gtcaacaaga
3241 cggcagtgag gacaatcaga caactactat tcaacaattt gttgaggttc aacctcaatt
3301 agagatggaa cttacaccag ttgttcagac tattgaagtg aatagtttta gtggttatth
3361 aaaacttact gacaatgtat acattaaaaa tgcagacatt gtggaagaag ctaaaagggt
3421 aaaaccaaca gtggttgtta atgcagccaa tgtttacctt aaacatggag gaggtgttgc

3481 aggagcctta aataaggcta ctaacaatgc catgcaagtt gaatctgatg attacatagc
3541 tactaatgga ccacttaaag tgggtggttag ttgtgtttta agcggacaca atcttgctaa
3601 aactgtctt catgttgtcg gcccaaagt taacaaaggt gaagacattc aacttcttaa
3661 gagtgcttat gaaaatttta atcagcacga agtttactt gcaccattat tatcagctgg
3721 ttttttggg gctgacccta tacattcttt aagagtttgt gtagatactg ttcgcacaaa
3781 tgttacttta gctgtctttg ataaaaatct ctatgacaaa cttgtttcaa gctttttgga
3841 aatgaagagt gaaaagcaag ttgaacaaaa gatcgctgag attcctaaag aggaagttaa
3901 gccatttata actgaaagta aaccttcagt tgaacagaga aaacaagatg ataagaaaat
3961 caaagcttgt gttgaagaag ttacaacaac tctggaagaa actaagttcc tcacagaaaa
4021 cttgttactt tatattgaca ttaatggcaa tcttcatcca gattctgcca ctcttgttag
4081 tgacattgac atcactttct taaagaaaga tgctccatat atagtgggtg atgttgttca
4141 agagggtggt ttaactgctg tggttatacc tactaaaaag gctggtggca ctactgaaat
4201 gctagcgaag gctttgagaa aagtgcacac agacaattat ataaccactt acccgggtca
4261 gggttttaat ggttacactg tagaggaggc aaagacagtg cttaaaaagt gtaaaagtgc
4321 cttttacatt ctaccatcta ttatctctaa tgagaagcaa gaaattcttg gaactgtttc
4381 ttggaatttg cgagaaatgc ttgcacatgc agaagaaaca cgcaaattaa tgctgtctg
4441 tgtggaaact aaagccatag tttcaactat acagcgtaaa tataagggtg ttaaaataca
4501 agagggtgtg gttgattatg gtgctagatt ttacttttac accagtaaaa caactgtagc
4561 gtcacttatc aacacactta acgatctaaa tgaaactctt gttacaatgc cacttggcta
4621 tgtaacacat ggcttaaatt tggaagaagc tgctcgggat atgagatctc tcaaagtgcc
4681 agctacagtt tctgtttctt cacctgatgc tgttacagcg tataatgggt atcttacttc
4741 ttcttctaaa acacctgaag aacattttat tgaaaccatc tcacttgctg gttcctataa
4801 agattgggtcc tattctggac aatctacaca actaggtata gaatttctta agagaggtga
4861 taaaagtgta tattacacta gtaatcctac cacattccac ctagatgggt aagttatcac
4921 ctttgacaat cttaagacac ttctttcttt gagagaagtg aggactatta aggtgtttac
4981 aacagtagac aacattaacc tccacacgca agttgtggac atgtcaatga catatggaca
5041 acagtttggg ccaacttatt tggatggagc tgatgttact aaaataaaaac ctcataattc
5101 acatgaaggt aaaacatttt atgttttacc taatgatgac actctacgtg ttgaggcttt
5161 tgagtactac cacacaactg atcctagttt tctgggtagg tacatgtcag cattaatatca
5221 cactaaaaag tggaaatacc cacaagttaa tggtttaact tctattaaat gggcagataa
5281 caactgttat cttgccactg cattgttaac actccaacaa atagagtuga agtttaatcc
5341 acctgtctca caagatgctt attacagagc aagggtggt gaagctgcta acttttgtgc
5401 acttatctta gcctactgta ataagacagt aggtgagtta ggtgatgta gagaaacaat
5461 gagttacttg tttcaacatg ccaatttaga ttcttgcaaa agagtcttga acgtggtgtg
5521 taaaacttgt ggacaacagc agacaaccct taagggtgta gaagctgta tgtacatggg
5581 cacactttct tatgaacaat ttaagaaagg tgttcagata ccttgtacgt gtggtaaaaca
5641 agctacaaaa tatctagtac aacaggagtc accttttgtt atgatgtcag caccacctgc
5701 tcagtatgaa cttaagcatg gtacatttac ttgtgctagt gagtacactg gtaattacca
5761 gtgtgggtcac tataaacata taacttctaa agaaactttg tattgcatag acgggtgctt
5821 acttacaag tcctcagaat acaaaggtcc tattacggat gttttctaca aagaaaacag
5881 ttacacaaca accataaaac cagttactta taaattggat ggtgttgttt gtacagaaat
5941 tgaccctaag ttggacaatt attataagaa agacaattct tatttcacag agcaaccaat
6001 tgatcttgta ccaaaccaac catatccaaa cgcaagcttc gataatttta agtttgtatg
6061 tgataatata aaatttgctg atgattttaa ccagtttaact ggttataaga aacctgcttc
6121 aagagagctt aaagttacat ttttccctga cttaaaggt gatgtgggtg ctattgatta
6181 taaacactac acaccctctt ttaagaaagg agctaaattg ttacataaac ctattgtttg
6241 gcatgttaac aatgcaacta ataaagccac gtataaacca aatacctggt gtatacgttg
6301 tctttggagc acaaaaccag ttgaaacatc aaattcgttt gatgtactga agtcagagga
6361 cgcgcaggga atggataatc ttgcctgcga agatctaaaa ccagtctctg aagaagtagt
6421 ggaaaatcct accatacaga aagacgttct tgagtgtaat gtgaaaacta ccgaagttgt
6481 aggagacatt atacttaaac cagcaataaa tagtttaaaa attacagaag aggttggcca
6541 cacagatcta atggctgctt atgtagacaa ttctagtctt actattaaga aacctaataga
6601 attatctaga gtattaggtt tgaaaaccct tgctactcat ggttttagct ctgttaatatg
6661 tgtcccttgg gatactatag ctaattatgc taagcctttt cttaacaaaag ttgttagtac
6721 aactactaac atagttacac ggtgtttaa cctgttttgt actaattata tgccttattt
6781 ctttacttta ttgctacaat tgtgtacttt tactagaagt acaaattcta gaattaaagc
6841 atctatgccg actactatag caaagaatac tgttaagagt gtcggtaaat tttgtctaga

6901 ggcttcattt aattatattga agtcaccta tttttctaaa ctgataaata ttataatttg
6961 gtttttacta ttaagtgttt gcctagggtc tttaatctac tcaaccgctg ctttaggtgt
7021 tttaatgtct aatttaggca tgccttctta ctgtactggt tacagagaag gctatttgaa
7081 ctctactaat gtcactattg caacctactg tactggttct ataccttgta gtgtttgtct
7141 tagtggttta gattcttttag acacctatcc ttcttttagaa actatacaaa ttaccatttc
7201 atctttttaa tgggatttaa ctgcttttgg cttagttgca gagtgggttt tggcatatat
7261 tcttttctact aggtttttct atgtacttgg attggctgca atcatgcaat tgttttctag
7321 ctattttgca gtacatttta ttagtaattc ttggcttatg tggttaataa ttaatcttgt
7381 acaaatggcc ccgatttcag ctatggtttag aatgtacatc ttctttgcat cattttatta
7441 tgtatggaaa agttatgtgc atgtttaga cggttgtaat tcatcaactt gtatgatgtg
7501 ttacaaaact aatagagcaa caagagtcca atgtacaact attgttaatg gtgttagaag
7561 gtccttttat gtctatgcta atggaggtaa aggccttttgc aaactacaca attggaattg
7621 tgttaattgt gatacattct gtgctggtag tacatttatt agtcatgcaag ttgctgagaga
7681 cttgtcacta cagtttaaaa gaccaataaa tctactgac cagtcttctt acatcgttga
7741 tagtgttaca gtgaagaatg gttccatcca tctttacttt gataaagctg gtcaaaagac
7801 ttatgaaaga cattctctct ctcattttgt taacttagac aacctgagag ctaataacac
7861 taaaggttca ttgcctatta atgttatagt ttttgatggt aaatcaaaat gtgaagaatc
7921 atctgcaaaa tcagcgtctg tttactacag tcagcttatg tgtcaaccta tactgttact
7981 agatcaggca ttagtgtctg atgttgggtga tagtgcggaa gttgcagtta aaatgtttga
8041 tgcttacggt aatacgtttt catcaacttt taacgtacca atggaaaaac tcaaaacact
8101 agttgcaact gcagaagctg aacttgcaaa gaatgtgtcc ttagacaatg tcttatctac
8161 ttttatttca gcagctcggc aagggtttgt tgattcagat gtagaaacta aagatgttgt
8221 tgaatgtctt aaattgtcac atcaatctga catagaagtt actggcgata gttgtaataa
8281 ctatatgctc acctataaca aagttgaaaa catgacaccc cgtgaccttg gtgcttgtat
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 tggtaaaatt gttaataatt ggttgaagca
8581 gttaattaaa gttacacttg tgttctttt tgttctgct attttctatt taataacacc
8641 tgttcatgtc atgtctaaac atactgactt ttcaagtga atcataggt 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 tagaagggtc tgttgcttat gaaagtttac gccctgacac
9121 acgttatgtg ctcatggatg gctctattat tcaatttctt aacacctacc ttgaagggtc
9181 tgttagagtg gtaacaactt ttgattctga gtactgtagg cacggcactt gtgaaagatc
9241 agaagctggg gtttgtgtat ctactagtgg tagatgggta cttacaatg attattacag
9301 atctttacca ggagttttct gtggtgtaga tgctgtaaat ttacttacta atatgtttac
9361 accactaatt caacctattg gtgctttgga catatcagca tctatagtag ctgggtggtat
9421 tgtagctatc gtagtaacat gccttgctta ctattttatg aggttttagaa gagcttttgg
9481 tgaatacagt catgtagttg ctttaatac tttactattc cttatgtcat tcaactgact
9541 ctgtttaaca ccagtttact cattcttacc tgggtgtttat tctgttattt acttgtactt
9601 gacattttat cttactaatg atgtttcttt tttagcacat attcagtgga tggttatggt
9661 cacaccttta gtaccttctt ggatacaaat tgcttatatc atttgtattt ccacaagca
9721 ttctatttgg ttcttttagta attacctaaa gagacgtgta gtctttaatg tgttttctt
9781 tagtactttt gaagaagctg cgctgtgcac ctttttggta aataaagaaa tgtatctaaa
9841 gttgcgtagt gatgtgctat tacctcttac gcaatataat agatacttag ctctttataa
9901 taagtacaag tatttttagt gagcaatgga tacaactagc tacagagaag ctgcttgttg
9961 tcatctcgca aaggetctca atgacttcag taactcaggt tctgatgttc tttaccaacc
10021 accacaaacc tctatcacct cagctgtttt gcagagtggg tttagaaaaa tggcattccc
10081 atctggtaaa gttgaggggt gtatggtaca agtaacttgt ggtacaacta cacttaacgg
10141 tctttggctt gatgacgtag tttactgtcc aagacatgtg atctgcacct ctgaagacat
10201 gcttaaccct aattatgaag atttactcat tcgtaagtct aatcataatt tcttgggtaca
10261 ggctggtaat gttcaactca gggttattgg acattctatg caaaattgtg tacttaagct

10321 taaggttgat acagccaatc ctaagacacc taagtataag tttgttcgca ttcaaccagg
10381 acagactttt tcagtgttag cttgttacaa tgggtcacca tctgggtggtt accaatgtgc
10441 tatgaggccc aatttcacta ttaaggggtc attccttaat ggttcatgtg gtagtgttgg
10501 ttttaacata gattatgact gtgtctcttt ttgttacatg caccatattg aattaccaac
10561 tggagttcat gctggcacag acttagaagg taacttttat ggaccttttg ttgacaggca
10621 aacagcacia gcagctggta cggacacaac tattacagtt aatgttttag cttgggtgta
10681 cgctgctgtt ataaatggag acaggtgggt tctcaatcga tttaccacia ctcttaatga
10741 ctttaacctt gtggctatga agtacaatta tgaacctcta acacaagacc atgttgacat
10801 actaggacct ctttctgctc aaactggaat tgccgtttta gatatgtgtg cttcattaaa
10861 agaattactg caaaatggta tgaatggacg taccatattg ggtagtgtt tattagaaga
10921 tgaatttaca ctttttgatg ttgttagaca atgtcaggt gttactttcc aaagtgcagt
10981 gaaaagaaca atcaagggta cacaccactg gttgttactc acaattttga cttcactttt
11041 agtttttagt cagagtactc aatggctctt gttctttttt ttgtatgaaa atgccttttt
11101 accttttgct atgggtatta ttgctatgtc tgcttttgca atgatgtttg tcaaacataa
11161 gcatgcattt ctctgtttgt tttgtttacc ttctcttgcc actgtagctt attttaatat
11221 ggtctatatg cctgctagtt ggggtgatgcg tattatgaca tgggtggata tgggtgatac
11281 tagtttgtct ggttttaagc taaaagactg tgttatgtat gcatcagctg tagtgttact
11341 aatccttatg acagcaagaa ctgtgtatga tgatgggtgct aggagagtgt ggacacttat
11401 gaatgtcttg aactcgttt ataaagttta ttatggtaat gcttttagatc aagccatttc
11461 catgtgggct cttataatct ctgttacttc taactactca ggtgtagtta caactgtcat
11521 gtttttgccc agaggtattg tttttatgtg tgttgagtat tgccctattt tcttcataac
11581 tggtaataca cttcagtgta taatgctagt ttattgtttc ttaggctatt tttgtacttg
11641 ttactttggc ctcttttggt tactcaaccg ctactttaga ctgactcttg gtgtttatga
11701 ttacttagtt tctacacagg agtttagata tatgaattca cagggactac tcccacccaa
11761 gaatagcata gatgccttca aactcaacat taaattgttg ggtgttggty gcaaaccttg
11821 tatcaaagta gccactgtac agtctaaaat gtcagatgta aagtgcacat cagtagtctt
11881 actctcagtt ttgcaacaac tcagagtata atcatcatct aaattgtggg ctcaatgtgt
11941 ccagttacac aatgacattc tcttagctaa agatactact gaagcctttg aaaaaatggt
12001 ttcaactact tctgttttgc tttccatgca ggggtgctgta gacataaaca agctttgtga
12061 atcaatgctg gacaacaggg caaccttaca agctatagcc tcagagttta gttcccttcc
12121 atcatatgca gcttttgcta ctgctcaaga agcttatgag caggctgttg ctaattgtga
12181 ttctgaagtt gttcttaaaa agttgaagaa gtctttgaat gtggctaaat ctgaatttga
12241 ccgtgatgca gccatgcaac gtaagttgga aaagatggct gatcaagcta tgacccaaat
12301 gtataaacag gctagatctg aggacaagag ggcaaaagt actagtgcta tgcagacaat
12361 gcttttctact 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 tgaatttagt atggacaatt cacctaattt agcatggcct cttattgtaa cagctttaag
12661 ggccaattct gctgtcaaat tacagaataa tgagcttagt cctgttgcac tacgacagat
12721 gtcttgtgct gccggtacta cacaaactgc ttgactgat gacaatgctg tagcttacta
12781 caacacaaca aagggaggta ggtttgtact tgcactgta tccgatttac aggatttgaa
12841 atgggctaga ttccctaaga gtgatggaac tggactatc tatacagaac tggaccacc
12901 ttgtaggttt gttacagaca cacctaaagg tccctaaagt aagtatttat actttattaa
12961 aggattaaac aacctaaata gaggtatggt acttggtagt ttactgtcca cagtacgtct
13021 acaagctggg aatgcaacag aagtgcctgc caattcaact gtattatctt tctgtgcttt
13081 tgctgtatag gctgctaaag cttacaaga ttatctagct agtgggggac aaccaatcac
13141 taattgtgtt aagatgttgt gtacacacac tggtagtggg caggcaataa cagttacacc
13201 ggaagccaat atggatcaag aatcctttgg tgggtgcatc tgttgtctgt actgccgttg
13261 ccacatagat catccaaatc ctaaaggatt ttgtgactta aaaggtaagt atgtacaaat
13321 acctacaact tgtgctaatt accctgtggg ttttactact aaaaacacag tctgtaccgt
13381 ctgctgtatg tggaaagggt atggctgtag ttgtgatcaa ctccgcgaac ccatgcttca
13441 gtcagctgat gcacaatcgt ttttaaacgg gtttgccgtg taagtgcagc ccgtcttaca
13501 ccgtgctgca caggcactag tactgatgtc gtatacaggg cttttgacat ctacaatgat
13561 aaagtagctg gttttgctaa attcctaaaa actaattgtt gtcgcttcca agaaaaggac
13621 gaagatgaca atttaattga ttcttacttt gtagttaaga gacacacttt ctctaactac
13681 caacatgaag aaacaattta taatttactt aaggattgtc cagctgttgc taacatgac

13741 ttctttaagt ttagaataga cgggtgacatg gtaccacata tatcacgtca acgtcttact
13801 aaatacacaa tggcagacct cgtctatgct ttaaggcatt ttgatgaagg taattgtgac
13861 acattaaaag aaatacttgt cacatacaat tgttgtgatg atgattatth caataaaaag
13921 gactgggatg atttttaga aaaccagat atattacgag tatacgcaa cttaggtgaa
13981 cgtgtacgcc aagctttgtt aaaaacagta caattctgtg atgccatgag aaatgctggt
14041 attgttggtg tactgacatt agataatcaa gatctcaatg gtaactggta tgatttcggt
14101 gatttcatac aaaccagcc aggtagtgga gttcctgttg tagattctta ttattcattg
14161 ttaatgccta tattaacct gaccagggt ttaactgcag agtcacatgt tgacactgac
14221 ttaacaaagc cttacattaa gtgggatttg ttaaaatag acttcacgga agagagggta
14281 aaactctttg accgttattt taaatattgg gatcagacat accacccaaa ttgtgttaac
14341 tgtttggatg acagatgcat tctgcattgt gcaaacttta atgttttatt ctctacagtg
14401 tccccacct caagttttgg accactagt agaaaaatat ttgttgatgg tgttccattt
14461 gtagtttcaa ctggatacca cttcagagag ctagggtgtg tacataatca ggatgtaaac
14521 ttacatagct ctgacttag ttttaaggaa ttacttgtgt atgctgctga ccctgctatg
14581 cacgctgctt ctggtaatct attactagat aaacgcacta cgtgcttttc agtagctgca
14641 cttactaaca atgttgcttt tcaaactgtc aaaccggta attttaaca agacttctat
14701 gactttgctg tgtctaaggg tttctttaag gaaggaagtt ctggtgaatt aaaacacttc
14761 ttctttgctc aggatggtaa tgcctgctatc agcgattatg actactatcg ttataatcta
14821 ccaacaatgt gtgatatcag acaactacta tttgtagtgg aagttgttga taagtacttt
14881 gattgttacg atgggtggctg tattaatgct aaccaagtca tcgtcaacaa cctagacaaa
14941 tcagctgggt ttccatttaa taaatggggg aaggctagac tttattatga ttcaatgagt
15001 tatgaggatc aagatgcact tttcgcatat acaaaacgta atgtcatccc tactataact
15061 caaatgaatc ttaagtatgc cattagtgc aagaatagag ctgcaccgt agctgggtgc
15121 tctatctgta gtactatgac caatagacag tttcatcaaa aattattgaa atcaatagcc
15181 gccactagag gagctactgt agtaattgga acaagcaaat tctatgggtg ttggcacaac
15241 atgttaaaaa ctgtttatag tgatgtagaa aaccctcacc ttatgggttg ggattatcct
15301 aaatgtgata gagccatgcc taacatgctt agaattatgg cctcacttgt tcttgctcgc
15361 aaacatacaa cgtgtttagt cttgtcacac cgtttctata gattagctaa tgagtgtgct
15421 caagtattga gtgaaatggg catgtgtggc ggttcactat atgttaaac aggtggaacc
15481 tctcaggag atgccacaac tgcttatgct aatagtgtt ttaacatttg tcaagctgtc
15541 acggccaatg ttaatgcact tttatctact gatggtaaca aaattgccga taagtgtgc
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 tgtctgaagc aaaatgttgg
15841 actgagactg accttactaa aggacctcat gaattttgct ctcaacatac aatgctagtt
15901 aaacaggggt atgattatgt gtaccttct taccagatc catcaagaat cctaggggccc
15961 ggctgttttg tagatgatat cgtaaaaaca gatggtacac ttatgattga acgggtcgtg
16021 tcttttagcta tagatgctta cccacttact aaacatccta atcaggagta tgctgatgtc
16081 tttcatttgt acttacaata cataagaaag ctacatgatg agttaacagg acacatgtta
16141 gacatgtatt ctgttatgct tactaatgat aacacttcaa ggtattggga acctgagttt
16201 tatgaggcta tgtacacacc gcatacagtc ttacaggctg ttggggcttg tgttctttgc
16261 aattcacaga cttcattaag atgtggtgct tgcatacgt gaccattctt atgttgtaaa
16321 tgcgtttacg accatgtcat atcaacatca cataaattag tcttgtctgt taatccgtat
16381 gtttgcaatg ctccagggtg tgatgtcaca gatgtgactc aactttactt aggaggatg
16441 agctattatt gtaaatacaca taaaccacc attagtthtc cattgtgtgc taatggacaa
16501 gtttttgggt tatataaaaa taactgtggt ggtagcgata atgttactga ctttaatgca
16561 attgcaacat gtgactggac aaatgctggg gattacatt tagctaacac ctgtactgaa
16621 agactcaagc tttttgcagc agaaacgctc aaagctactg aggagacatt taaactgtct
16681 tatggatttg ctactgtacg tgaagtgtg tctgacagag aattacatct tcatgggaa
16741 gttggtaaac ctgaccacc acttaaccga aattatgtct ttactgggta tctgtgtaact
16801 aaaaacagta aagtacaaat aggagagtac acctttgaaa aagggtgacta tgggtgatgct
16861 gttgtttacc gaggtacaac aacttacaac ttaaatgttg gtgattatth tgtgctgaca
16921 tcacatacag taatgccatt aagtgcacct acactagtgc cacaagagca ctatgttaga
16981 attactggct tatacccaac actcaatatc tcagatgagt tttctagcaa tgttgcaaat
17041 tatcaaaagg ttggtatgca aaagtattct acactccagg gaccacctgg tactggtaag
17101 agtcattttg ctattggcct agctctctac tacccttctg ctgcgatagt gtatacagct

17161 tgctctcatg ccgctgttga tgcactatgt gagaaggcat taaaatattt gcctatagat
17221 aatgtagta gaattatacc tgcacgtgct cgtgtagagt gttttgataa attcaaagtg
17281 aattcaacat tagaacagta tgtcttttgt actgtaaagt cattgcctga gacgacagca
17341 gatatagttg tctttgatga aatttcaatg gccacaaatt atgatttgag tgttgtcaat
17401 gccagattac gtgctaagca ctatgtgtac attggcgacc ctgctcaatt acctgcacca
17461 cgcacattgc taactaaggg cacactagaa ccagaatatt tcaattcagt gtgtagactt
17521 atgaaaacta taggtccaga catgttcctc ggaacttgtc ggcgttgtcc tgcetgaaatt
17581 gttgacactg tgagtgtctt ggtttatgat aataagctta aagcacataa agacaaatca
17641 gctcaatgct ttaaaatggt ttataagggg gttatcacgc atgatgtttc atctgcaatt
17701 aacaggccac aaatagggct ggtaagagaa ttccttacac gtaaccctgc ttggagaaaa
17761 gctgtcttta tttcacctta taattcacag aatgctgtag cctcaaagat tttgggacta
17821 ccaactcaaa ctggttgattc atcacagggc tcagaatatg actatgtcat attcactcaa
17881 accactgaaa cagctcactc ttgtaatgta aacagattta atggtgctat taccagagca
17941 aaagtaggca tactttgcat aatgtctgat agagaccttt atgacaagtt gcaatttaca
18001 agtcttgaiaa ttccacgtag gaatgtggca actttacaag ctgaaaatgt aacaggactc
18061 tttaaagatt gtagtaaggt aatcactggg ttacatccta cacaggcacc tacacacctc
18121 agtgttgaca ctaaattcaa aactgaaggt ttatgtgttg acatacctgg catacctaag
18181 gacatgacct atagaagact catctctatg atgggtttta aaatgaatta tcaagttaat
18241 ggttacccta acatgtttat cccccgcaa gaagctataa gacatgtacg tgcattggatt
18301 ggcttcgatg tgcaggggtg tcatgctact agagaagctg ttggtacca tttaccttta
18361 cagctagggt tttctacagg tgttaaccta gttgctgtac ctacaggtta tgttgataca
18421 cctaataata cagatttttc cagagttagt gctaaaccac cgcctggaga tcaatttaaa
18481 cacctcatac cacttatgta caaaggactt ctttggaatg tagtgcgtat aaagattgta
18541 caaatgttaa gtgacacact taaaaatctc tctgacagag tctgatttgt cttatgggca
18601 catggctttg agttgacatc tatgaagtat tttgtgaaaa taggacctga gcgcacctgt
18661 tgtctatgtg atagacgtgc cacatgcttt tccactgctt cagacactta tgctgtttgg
18721 catcattcta ttggatttga ttacgtctat aatccgttta tgattgatgt tcaacaatgg
18781 ggttttacag gtaacctaca aagcaacctat gatctgtatt gtcaagtcca tggtaatgca
18841 catgtagcta gttgtgatgc aatcatgact aggtgtctag ctgtccacga gtgctttggt
18901 aagcgtgttg actggactat tgaatatcct ataattggtg atgaactgaa gattaatgcg
18961 gcttgtagaa aggttcaaca catggttgtt aaagctgcat tattagcaga caaattccca
19021 gttcttcacg acattggtaa ccctaaagct attaagtgtg tacctcaagc tgatgtagaa
19081 tgggaagttct atgatgcaca gccttgtagt gacaaagctt ataaaataga agaattatc
19141 tattcttatg ccacacatte tgacaaatc acagatggtg tatgcctatt ttggaattgc
19201 aatgtcgata gatatactgc taattccatt gtttgtatag ttgacactag agtgcctatc
19261 aaccttaact tgctgggttg tgatgggtggc agtttgtatg taaataaaca tgcattccac
19321 acaccagctt ttgataaaag tgcttttgtt aattttaaac aattaccatt tttctattac
19381 tctgacagtc catgtgagtc tcatggaaaa caagtagtgt cagatataga ttatgtacca
19441 ctaaagtctg ctacgtgtat aacacgttgc aatttaggtg gtgctgtctg tagacatcat
19501 gctaattgagt acagattgta tctcgatgct tataacatga tgatctcagc tggctttagc
19561 ttgtgggttt acaacaatt tgatacttat aacctctgga acacttttac aagacttcag
19621 agtttagaaa atgtggcttt taatgttga aataagggac actttgatgg acaacagggg
19681 gaagtaccag tttctatcat taataacact gtttacacaa aagttgatgg tgttgatgta
19741 gaattgtttg aaaataaaac aacattacct gttaatgtag ctttgagct ttgggctaag
19801 cgcaacatta aaccagtacc agaggtgaaa atactcaata atttgggtgt ggacattgct
19861 gctaatactg tgatctggga ctacaaaaga gatgctccag cacatatac tactattggt
19921 gtttgttcta tgactgacat agccaagaaa ccaactgaaa cgatttgtgc accactcact
19981 gtcttttttg atggtagagt tgatggtcaa gttagactat ttagaaatgc ccgtaatggt
20041 gttcttatta cagaaggtag tgttaaaggt ttacaacct ctgtaggtcc caacaagct
20101 agtcttaatg gagtcacatt aattggagaa gccgtaaaaa cacagttcaa ttattataag
20161 aaagttgatg gtgttgtcca acaattacct gaaacttact ttactcagag tagaaattta
20221 caagaattta aaccaggag tcaaatggaa attgatttct tagaattagc tatggatgaa
20281 ttcattgaac ggtataaatt agaaggctat gccttcgaac atatcgttta tggagatttt
20341 agtcatagtc agttagggtg tttacatcta ctgattggac tagctaaacg ttttaaggaa
20401 tcaccttttg aattagaaga ttttattcct atggacagta cagttaaaaa ctatttcata
20461 acagatgctc aaacagggtc atctaagtgt gtgtgttctg ttattgattt attacttgat
20521 gattttgttg aaataataaa atcccaagat ttatctgtag tttctaaggt tgtcaaagtg

20581 actattgact atacagaaat ttcatttatg ctttggtgta aagatggcca ttagaaca
20641 ttttacccea aattacaatc tagtcaagcg tggcaaccgg gtgttgctat gcctaactct
20701 tacaaaatgc aaagaatgct attagaaaag tgtgacctc aaaattatgg tgatagtgca
20761 acattaccta aaggcataat gatgaatgtc gcaaaaatata ctcaactgtg tcaatattta
20821 aacacattaa cattagctgt accctataat atgagagtta tacatatttg tgctggttct
20881 gataaaggag ttgcaccagg tacagctggt ttaagacagt ggttgcctac gggtagcgtg
20941 cttgtcgatt cagatcttaa tgactttgtc tctgatgcag attcaacttt gattggtgat
21001 tgtgcaactg tacatacagc taataaatgg gatctcatta ttagtgatat gtacgacct
21061 aagactaaaa atgttacaaa agaaaatgac tctaaagagg gttttttcac ttacatttgt
21121 gggtttatac aacaaaagct agctcttggg ggttccgtgg ctataaagat aacagaacat
21181 tcttggaaatg ctgatcttta taagctcatg ggacacttcg catggtggac agcctttggt
21241 actaatgtga atgctcatc atctgaagca tttttaattg gatgtaatta tcttggcaaa
21301 ccacgcgaac aaatagatgg ttatgtcatg catgcaaatt acatattttg gaggaataca
21361 aatccaattc agttgtcttc ctattcttta tttgacatga gtaaatttcc ccttaaatta
21421 aggggtactg ctgttatgtc tttaaaagaa ggtcaaatac atgatatgat tttatctctt
21481 cttagtaaag gtagacttat aattagagaa aacaacagag ttgttatttc tagtgatggt
21541 cttgttaaca actaaacgaa caatgtttgt tttcttgtt ttattgccac tagtctctag
21601 tcagtgtggt aatcttaca ccagaactca attaccctc gcatacacta attctttcac
21661 acgtgggtgt tattaccctg acaaagtttt cagatcctca gttttacatt caactcagga
21721 cttgttctta ctttctttt ccaatgttac ttggttccat gctatacatg tctctgggac
21781 caatggtact aagaggtttg ataaccctgt cctaccattt aatgatggtg tttattttgc
21841 ttccactgag aagtctaaca taataagagg ctggattttt ggtactactt tagattcgaa
21901 gaccagtcct ctacttattg ttaataacgc tactaatgtt gttattaag tctgtgaatt
21961 tcaattttgt aatgatccat ttttgggtgt ttattaccac aaaaacaaca aaagttggat
22021 ggaaagtgag ttcagagttt attctagtgc gaataattgc acttttgaat atgtctctca
22081 gccttttctt atggaccttg aaggaaaaca gggtaatttc aaaaatctta gggaaattgt
22141 gtttaagaat attgatggtt attttaaaat atattctaag cacacgccta ttaatttagt
22201 gcgtgatctc cctcaggggt tttcggcttt agaaccattg gtagatttgc caataggat
22261 taacatcact aggtttcaaa ctttacttgc tttacataga agttatttga ctctgggtga
22321 ttcttcttca ggttggacag ctggtgctgc agcttattat gtgggttat tccaacctag
22381 gacttttcta ttaaaatata atgaaaatgg aaccattaca gatgctgtag actgtcact
22441 tgaccctctc tcagaaacaa agtgtacgtt gaaatccttc actgtagaaa aaggaatcta
22501 tcaaacttct aactttagag tccaaccaac agaactctatt gttagatttc ctaatattac
22561 aaacttgtgc ctttttgggtg aagtttttaa cgccaccaga tttgcatctg tttatgcttg
22621 gaacaggaag agaatcagca actgtgttgc tgattattct gtcctatata attccgcctc
22681 attttccact tttaagtgtt atggagtgtc tctactaaa ttaaagtatc tctgctttac
22741 taatgtctat gcagattcat ttgtaattag aggtgatgaa gtcagacaaa tcgctccagg
22801 gcaaactgga aagattgctg attataatta taaattacca gatgatttta caggctgcgt
22861 tatagcttgg aattctaaca atcttgattc taaggttggt ggttaattata attacctgta
22921 tagattgttt aggaagtcta atctcaaacc ttttgagaga gatatttcaa ctgaaatcta
22981 tcaggccggt agcacacctt gtaatggtgt tgaaggtttt aattgttact ttcctttaca
23041 atcatatggt ttccaacca ctaatggtgt tggttaccaa ccatacagag tagtagtact
23101 ttcttttgaa cttctacatg caccagcaac tgtttgtgga cctaaaaagt ctactaattt
23161 ggttaaaaaa aaatgtgtca atttcaactt caatggttta acaggcacag gtgttcttac
23221 tgagtctaac aaaaagtttc tgctttcca acaatttggc agagacattg ctgacactac
23281 tgatgctgtc cgtgatccac agacacttga gattcttgac attacaccat gttcttttgg
23341 tgggtgctagt gttataacac caggaacaaa tacttctaac caggttgcgt tctttatca
23401 ggatgttaac tgcacagaag tccctgttgc tattcatgca gatcaactta ctctacttg
23461 gcgtgtttat tctacaggtt ctaagttttt tcaaacacgt gcaggctgtt taataggggc
23521 tgaacatgtc aacaactcat atgagtgtga cataccatt ggtgcaggta tatgcgctag
23581 ttatcagact cagactaatt ctctcggcg ggcacgtagt gtagctagtc aatccatcat
23641 tgctacact atgtcacttg gtgcagaaaa ttcagttgct tactetaata actctattgc
23701 catacccaca aattttacta ttagtgttac cacagaaatt ctaccagtgt ctatgaccaa
23761 gacatcagta gattgtacaa tgtacatttg tggtgattca actgaatgca gcaatctttt
23821 gttgcaatat ggcagttttt gtacacaatt aaaccgtgct ttaactggaa tagctgttga
23881 acaagacaaa aacacccaag aagtttttgc acaagtcaaa caaatttaca aaacaccacc
23941 aattaaagat tttggtggtt ttaatttttc acaaatatta ccagatccat caaaaccaag

24001 caagagggtca tttattgaag atctactttt caacaaagtg acacttgcag atgctggcctt
24061 catcaaaciaa tatgggtgatt gccttgggtga tattgctgct agagacctca tttgtgcaca
24121 aaagtttaac ggccttactg ttttgcacc tttgctcaca gatgaaatga ttgctcaata
24181 cacttctgca ctgttagcgg gtacaatcac ttctgggttg acccttgggtg caggtgctgc
24241 attacaaata ccatttgcta tgcaaatggc ttataggttt aatgggtattg gagttacaca
24301 gaatgttctc tatgagaacc aaaaattgat tgccaaccaa tttaatagtg ctattggcaa
24361 aattcaagac tcactttctt ccacagcaag tgcacttggg aaacttcaag atgtgggtcaa
24421 ccaaaatgca caagctttta acacgcttgt taacaactt agctccaatt ttggtgcaat
24481 ttcaagtgtt ttaaattgata tcctttcacg tcttgacaaa gttgaggctg aagtgcaaat
24541 tgatagggtt atcacaggca gacttcaaag tttgcagaca tatgtgactc aacaattaat
24601 tagagctgca gaaatcagag cttctgctaa tcttgctgct actaaaatgt cagagtgtgt
24661 acttggacaa tcaaaaagag ttgattttttg tggaaagggc tatcatctta tgccttccc
24721 tcaagtgcga cctcatgggt tagtcttctt gcatgtgact tatgtccctg cacaagaaaa
24781 gaacttcaca actgctcctg ccatttgtca tgatggaaaa gcacactttc ctcgtgaagg
24841 tgtctttgtt 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 aatgagggtt ccaagaattt
25141 aatgaatct ctcatcgatc tccaagaact tggaaagtat gagcagtata taaaatggcc
25201 atggtacatt tggctagggt ttatagctgg cttgattgcc atagtaatgg tgacaattat
25261 gctttgctgt atgaccagtt gctgtagttg tctcaagggc tgttgttctt gtggatcctg
25321 ctgcaaattt gatgaagacg actctgagcc agtgctcaaa ggagtcaaat tacattacac
25381 ataaacgaac ttatggattt gtttatgaga atcttcacaa ttggaactgt aactttgaag
25441 caaggtgaaa tcaaggatgc tactccttca gattttgttc gcgctactgc aacgataccg
25501 atacaagcct cactcccttt cggatggcct attgttggcg ttgcaacttct tgctgttttt
25561 cagagcgctt ccaaaatcat aaccctcaa aagagatggc aactagcact ctccaaggg
25621 gttcactttg tttgcaactt gctgttgttg tttgtaacag tttactcaca ccttttgctc
25681 gttgtctgct gccttgaagc ccttttctc tatctttatg ctttagtcta cttcttgcaag
25741 agtataaact ttgtaagaat aataattgag ctttggcctt gctggaaatg ccgttccaaa
25801 aaccattac tttatgatgc ttaactattt ctttgcctggc atactaattg ttacgactat
25861 tgtatacctt acaatagtg aacttcttca attgtcatta cttcaggtga tggcacaaca
25921 agtcctatct ctgaacatga ctaccagatt ggtgggttata ctgaaaaatg ggaatctgga
25981 gtaaaagact gtgttgattt acacagttac ttcacttcag actattacca gctgtactca
26041 actcaattga gtacagacac tgggtgttga catgttacct tcttcatcta caataaaatt
26101 gttgatgagc ctgaagaaca tgtccaaatt cacacaatcg acggttcac cggagtgtgt
26161 aatccagtaa tggaaaccaat ttatgatgaa cgcagcagca ctactagcgt gcctttgtaa
26221 gcacaagctg atgagtacga acttatgtac tcattcgttt cgggaagagac aggtacgtta
26281 atagttaata gcgtacttct ttttcttctg ttctgtggtat tcttgctagt tacactagcc
26341 atccttactg cgcttcgatt gtgtgctgac tgctgcaata ttgttaacgt gactcttgta
26401 aaaccttctt tttacgttta ctctcgtgtt aaaaatctga attcttctag agttcctgat
26461 cttctggctt aaacgaacta aatattatat tagtttttct gtttggaaact ttaattttag
26521 ccattggcaga ttccaacggt actattaccg ttgaagagct taaaaagctc cttgaacaat
26581 ggaacctagt aataggtttc ctattcctta catggatttg tcttctacaa tttgcctatg
26641 ccaacaggaa taggtttttg tatataatta agttaatttt cctctggctg ttatggccag
26701 taacttttagc ttgttttgtg cttgctgctg tttacagaat aaattggatc accggtggaa
26761 ttgctatcgc aatggcttgt cttgtaggct tgatgtggct cagctacttc attgcttctt
26821 tcgactgtt tgccgctacg cgttccatgt ggtcattcaa tccagaactc aacattcttc
26881 tcaacgtgcc actccatggc actattctga ccagaccgct tctagaaagt gaactcgtaa
26941 tccgagctgt gatccttctg ggacatcttc gtattgctgg acaccatcta ggacgctgtg
27001 acatcaagga cctgcctaaa gaaatcactg ttgctacatc acgaacgctt tcttattaca
27061 aattgggagc ttgcgagcgt gtagcagggt actcaggttt tgcctgcatac agtcgctaca
27121 ggattggcaa ctataaatta aacacagacc attccagtag cagtgacaat attgctttgc
27181 ttgtacagta agtgacaaca gatgtttcat ctcgttgact 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 cttttaaaag aaccttgctc ttctggaaca tacgagggca attcaccatt tcatcctcta
27541 gctgataaca aatttgcact gacttgcttt agcactcaat ttgcttttgc ttgtcctgac
27601 ggcgtaaaac acgtctatca gttacgtgcc agatcagttt cacctaaact gttcatcaga
27661 caagaggaag ttcaagaact ttactctcca atttttctta ttgttgcggc aatagtgttt
27721 ataacacttt gcttcacact caaaagaaag acagaatgat tgaactttca ttaattgact
27781 tctatattgtg ctttttagcc tttctgctat tccttgtttt aattatgctt attatctttt
27841 ggttctcact tgaactgcaa gatcataatg aaacttgta cgcctaaacg aacatgaaat
27901 ttcttgtttt cttaggaatc atcacaactg tagctgcatt tcaccaagaa tgtagtttac
27961 agtcatgtac tcaacatcaa ccatatgtag ttgatgacct gtgtcctatt cacttctatt
28021 ctaaagtgta tattagagta ggagctagaa aatcagcacc ttttaattgaa ttgtgcgtgg
28081 atgaggctgg ttctaataca cccattcagt acatcgatat cggtaattat acagtttcct
28141 gtttaccttt tacaattaat tgccaggaac ctaaattggg tagtcttgta gtgcgttggt
28201 cgttctatga agacttttta gagtatcatg acgttcgtgt tgttttagat ttcactctaaa
28261 cgaacaaact aaaatgtctg ataatggacc caaaaatcag cgaaatgcac cccgcattac
28321 gtttggtgga ccctcagatt caactggcag taaccagaat ggagaacgca gtggggcgcg
28381 atcaaaaaca cgtcggcccc aaggtttacc caataatact gcgtcttggt tcaccgctct
28441 cactcaacat ggcaaggaag accttaatt ccctcgagga caaggcgctc caattaacac
28501 caatagcagt ccagatgacc aaattggcta ctaccgaaga gctaccagac gaattcgtgg
28561 tgggtgacggg aaaatgaaag atctcagtc aagatgggat ttctactacc taggaactgg
28621 gccagaagct ggacttcctt atgggtgctaa caaagacggc atcatatggg ttgcaactga
28681 gggagccttg aatacaccaa aagatcacat tggcaccgc aatcctgcta acaatgctgc
28741 aatcgtgcta caacttcctc aaggaacaac attgccaaaa ggcttctacg cagaagggag
28801 cagaggcggc agtcaagcct cttctcgctt ctcacacgt agtcgcaaca gttcaagaaa
28861 ttcaactcca ggcagcagta ggggaacttc tcctgctaga atggctggca atggcggtga
28921 tgctgctctt gctttgctgc tgcttgacag attgaaccag cttgagagca aaatgtctgg
28981 taaaggccaa caacaacaag gccaaactgt cactaagaaa tctgctgctg aggcttctaa
29041 gaagcctcgg caaaaacgta ctgccactaa agcatacaat gtaacacaag ctttcggcag
29101 acgtggtcca gaacaaacc aaggaaattt tggggaccag gaactaatca gacaaggaac
29161 tgattacaaa cattggccgc aaattgcaca atttgcccc agcgttcag cgttcttcgg
29221 aatgtcgcgc attggcatgg aagtccacc ttcgggaacg tggttgacct acacaggtgc
29281 catcaaattg gatgacaaag atccaaattt caaagatcaa gtcattttgc tgaataagca
29341 tattgacgca taaaaacat tcccaccaac agagcctaaa aaggacaaaa agaagaaggc
29401 tgatgaaact caagccttac cgcagagaca gaagaaacag caaactgtga ctcttcttcc
29461 tgctgcagat ttggatgatt tctccaaaca attgcaaca tccatgagca gtgctgactc
29521 aactcaggcc taaactcatg cagaccacac aaggcagatg ggctatataa acgttttcgc
29581 ttttcctgtt acgatataa gtctactctt gtgcagaatg aattctcgta actacatagc
29641 acaagtagat gtagttaact ttaatctcac atagcaatct ttaatcagtg tghtaacatta
29701 gggaggactt gaaagagcca ccacattttc accgaggcca cgcggagtac gatcgagtgt
29761 acagtgaaca atgctagggg gagctgccta tatggaagag ccctaattgtg taaaattaat
29821 tttagtagtg ctatccccat gtgattttta tagcttctta ggagaatgac aaaaaaaaaa
29881 aaaaaaaaaa aaaaaaaaaa aaa

//