**Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/USA/CA-LACPHL-AF08976/2022 ORF1ab polyprotein (ORF1ab) gene, partial cds; and ORF1a polyprotein (ORF1ab), surface glycoprotein (S), ORF3a protein (ORF3a), envelope protein (E), membrane g...**

GenBank: ON421981.1

[FASTA](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?report=fasta) [Graphics](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?report=graph)

[Go to:](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1" \l "goto2234631602_0)

LOCUS ON421981 29651 bp RNA linear VRL 06-MAY-2022

DEFINITION Severe acute respiratory syndrome coronavirus 2 isolate

SARS-CoV-2/human/USA/CA-LACPHL-AF08976/2022 ORF1ab polyprotein

(ORF1ab) gene, partial cds; and ORF1a polyprotein (ORF1ab), surface

glycoprotein (S), ORF3a protein (ORF3a), envelope protein (E),

membrane glycoprotein (M), ORF6 protein (ORF6), ORF7a protein

(ORF7a), ORF7b (ORF7b), ORF8 protein (ORF8), nucleocapsid

phosphoprotein (N), and ORF10 protein (ORF10) genes, complete cds.

ACCESSION ON421981

VERSION ON421981.1

DBLINK BioProject: [PRJNA746690](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA746690)

BioSample: [SAMN28104190](https://www.ncbi.nlm.nih.gov/biosample/SAMN28104190)

KEYWORDS .

SOURCE Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)

ORGANISM [Severe acute respiratory syndrome coronavirus 2](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=2697049)

Viruses; Riboviria; Orthornavirae; Pisuviricota; Pisoniviricetes;

Nidovirales; Cornidovirineae; Coronaviridae; Orthocoronavirinae;

Betacoronavirus; Sarbecovirus.

REFERENCE 1 (bases 1 to 29651)

AUTHORS Garrigues,J.M., Hemarajata,P. and Green,N.M.

TITLE Direct Submission

JOURNAL Submitted (05-MAY-2022) Los Angeles County Public Health

Laboratories, Los Angeles County Public Health Lab microbial

pathogen submission group, 12750 Erickson Avenue, Downey, CA 90242,

USA

COMMENT ##Assembly-Data-START##

Assembly Method :: Clear Labs BIP-Wv7 v. Clear Labs BIP-Wv7

Sequencing Technology :: ONT

##Assembly-Data-END##

FEATURES Location/Qualifiers

source 1..29651

/organism="Severe acute respiratory syndrome coronavirus

2"

/mol\_type="genomic RNA"

/isolate="SARS-CoV-2/human/USA/CA-LACPHL-AF08976/2022"

/isolation\_source="clinical"

/host="Homo sapiens"

/db\_xref="taxon:[2697049](https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=2697049)"

/country="USA"

/collection\_date="2022-04-20"

[gene](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=212&to=20613) 212..>20613

/gene="ORF1ab"

[CDS](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?location=212:13405,13405:20613) join(212..13405,13405..>20613)

/gene="ORF1ab"

/ribosomal\_slippage

/codon\_start=1

/product="ORF1ab polyprotein"

/protein\_id="[UPX94056.1](https://www.ncbi.nlm.nih.gov/protein/2234631604)"

/translation="MESLVPGFNEKTHVQLSLPVLQVRDVLVRGFGDSVEEVLSEARQ

HLKDGTCGLVEVEKGVLPQLEQPYVFIKRSDARTAPHGHVMVELVAELEGIQYGRSGE

TLGVLVPHVGEIPVAYRKVLLRKNGNKGAGGHRYGADLKSFDLGDELGTDPYEDFQEN

WNTKHSSGVTRELMRELNGGAYTRYVDNNFCGPDGYPLECIKDLLARAGKASCTLSEQ

LDFIDTKRGVYCCREHEHEIAWYTERSEKSYELQTPFEIKLAKKFDTFNGECPNFVFP

LNSIIKTIQPRVEKKKLDGFMGRIRSVYPVASPNECNQMCLSTLMKCDHCGETSWQTG

DFVKATCEFCGTENLTKEGATTCGYLPQNAVVKIYCPACHNSEVGPEHSLAEYHNESG

LKTILRKGGRTIAFGGCVFSYVGCHNKCAYWVPRASANIGCNHTGVVGEGSEGLNDNL

LEILQKEKVNINIVGDFKLNEEIAIILASFSASTSAFVETVKGLDYKAFKQIVESCGN

FKVTKGKAKKGAWNIGEQKSILSPLYAFASEAARVVRSIFSRTLETAQNSVRVLQKAA

ITILDGISQYSLRLIDAMMFTSDLATNNLVVMAYITGGVVQLTSQWLTNIFGTVYEKL

KPVLDWLEEKFKEGVEFLRDGWEIVKFISTCACEIVGGQIVTCAKEIKESVQTFFKLV

NKFLALCADSIIIGGAKLKALNLGETFVTHSKGLYRKCVKSREETGLLMPLKAPKEII

FLEGETLPTEVLTEEVVLKTGDLQPLEQPTSEAVEAPLVGTPVCINGLMLLEIKDTEK

YCALAPNMMVTNNTFTLKGGAPTKVTFGDDTVIEVQGYKSVNIIFELDERIDKVLNEK

CSAYTVELGTEVNEFACVVADAVIKTLQPVSELLTPLGIDLDEWSMATYYLFDESGEF

KLASHMYCSFYPPDEDEEEGDCEEEEFEPSTQYEYGTEDDYQGKPLEFGATSAALQPE

EEQEEDWLDDDSQQTVGQQDGSEDNQTTTIQTIVEVQPQLEMELTPVVQTIEVNSFSG

YLKLTDNVYIKNADIVEEAKKVKPTVVVNAANVYLKHGGGVAGALNKATNNAMQVESD

DYIATNGPLKVGGSCVLSGHNLAKHCLHVVGPNVNKGEDIQLLKSAYENFNQHEVLLA

PLLSAGIFGADPIHSLRVCVDTVRTNVYLAVFDKNLYDKLVSSFLEMKSEKQVEQKIA

EIPKEEVKPFITESKPSVEQRKQDDKKIKACVEEVTTTLEETKFLTENLLLYIDINGN

LHPDSATLVSDIDITFLKKDAPYIVGDVVQEGVLTAVVIPTKKASGTTEMLAKALRKV

PTDNYITTYPGQGLNGYTVEEAKTVLKKCKSAFYILPSIISNEKQEILGTVSWNLREM

LAHAEETRKLMPVCVETKAIVSTIQRKYKGIKIQEGVVDYGARFYFYTSKTTVASLIN

TLNDLNETLVTMPLGYVTHGLNLEEAARYMRSLKVPATVSVSSPDAVTAYNGYLTSSS

KTPEEHFIETISLAGSYKDWSYSGQSTQLGIEFLKRGDKSVYYTSNPTTFHLDGEVIT

FDNLKTLLSLREVRTIKVFTTVDNINLHTQVVDMSMTYGQQFGPTYLDGADVTKIKPH

NSHEGKTFYVLPNDDTLRVEAFEYYHTTDPSFLGRYMSALNHTKKWKYPQVNGLTSIK

WADNNCYLATALLTLQQIELKFNPPALQDAYYRARAGEAANFCALILAYCNKTVGELG

DVRETMSYLFQHANLDSCKRVLNVVCKTCGQQQTTLKGVEAVMYMGTLSYEQFKKGVQ

IPCTCGKQATKYLVQQESPFVMMSAPPAQYELKHGTFTCASEYTGNYQCGHYKHITSK

ETLYCIDGALLTKSSEYKGPITDVFYKENSYTTTIKPVTYKLDGVVCTEIDPKLDNYY

KKDNSYFTEQPIDLVPNQPYPNASFDNFKFVCDNIKFADDLNQLTGYKKPASRELKVT

FFPDLNGDVVAIDYKHYTPSFKKGAKLLHKPIVWHVNNATNKATYKPNTWCIRCLWST

KPVETSNSFDVLKSEDAQGMDNLACEDLKPVSEEVVENPTIQKDVLECNVKTTEVVGD

IILKPANNSLKITEEVGHTDLMAAYVDNSSLTIKKPNELSRVLGLKTLATHGLAAVNS

VPWDTIANYAKPFLNKVVSTTTNIVTRCLNRVCTNYMPYFFTLLLQLCTFTRSTNSRI

KASMPTTIAKNTVKSVGKFCLEASFNYLKSPNFSKLINIIIWFLLLSVCLGSLIYSTA

ALGVLMSNLGMPSYCTGYREGYLNSTNVTIATYCTGSIPCSVCLSGLDSLDTYPSLET

IQITISSFKWDLTAFGLVAEWFLAYILFTRFFYVLGLAAIMQLFFSYFAVHFISNSWL

MWLIINLVQMAPISAMVRMYIFFASFYYVWKSYVHVVDGCNSSTCMMCYKRNRATRVE

CTTIVNGVRRSFYVYANGGKGFCKLHNWNCVNCDTFCAGSTFISDEVARDLSLQFKRP

INPTDQSSYIVDSVTVKNGSIHLYFDKAGQKTYERHSLSHFVNLDNLRANNTKGSLPI

NVIVFDGKSKCEESSAKSASVYYSQLMCQPILLLDQALVSDVGDSAEVAVKMFDAYVN

TFSSTFNVPMEKLKTLVATAEAELAKNVSLDNVLSTFISAARQGFVDSDVETKDVVEC

LKLSHQSDIEVTGDSCNNYMLTYNKVENMTPRDLGACIDCSARHINAQVAKSHNIALI

WNVKDFMSLSEQLRKQIRSAAKKNNLPFKLTCATTRQVVNVVTTKIALKGGKIVNNWL

KQLIKVTLVFLFVAAIFYLITPVHVMSKHTDFSSEIIGYKAIDGGVTRDIASTDTCFA

NKHADFDTWFSQRGGSYTNDKACPLIAAVITREVGFVVPGLPGTILRTTNGDFLHFLP

RVFSAVGNICYTPSKLIEYTDFATSACVLAAECTIFKDASGKPVPYCYDTNVLEGSVA

YESLRPDTRYVLMDGSIIQFPNTYLEGSVRVVTTFDSEYCRHGTCERSEAGVCVSTSG

RWVLNNDYYRSLPGVFCGVDAVNLFTNMFTPLIQPIGALDISASIVAGGIVAIVVTCL

AYYFMRFRRAFGEYSHVVAFNTLLFLMSFIVLCLTPVYSFLPGVYSVIYLYLTFYLTN

DVSFLAHIQWMVMFTPLVPFWITIAYIICISTKHFYWFFSNYLKRRVVFNGVSFSTFE

EAALCTFLLNKEMYLKLRSDVLLPFTQYNRYLALYNKYKYFSGAMDTTSYREAACCHL

AKALNDFSNSGSDVLYQPPQISITSAVLQSGFRKMAFPSGKVEGCMVQVTCGTTTLNG

LWLDDVVYCPRHVICTSEDMLNPNYEDLLIRKSNHNFLVQAGNVQLRVIGHSMQNCVL

KLKVDTANPKTPKYKFVRIQPGQTFSVLACYNGSPSGVYQCAMRHNFTIKGSFLNGSC

GSVGFNIDYDCVSFCYMHHMELPTGVHAGTDLEGNFYGPFVDRQTAQAAGTDTTITVN

VLAWLYAAVINGDRWFLNRFTTTLNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAV

LDMCASLKELLQNGMNGRTILGSALLEDEFTPFDVVRQCSGVTFQSAVKRTIKGTHHW

LLLTILTSLLVLVQSTQWSLFFFLYENAFLPFAMGIIAMSAFAMMFVKHKHAFLCLFL

LPSLATVAYFNMVYMPASWVMRIMTWLDMVDTSLKLKDCVMYASAVVLLILMTARTVY

DDGARRVWTLMNVLTLVYKVYYGNALDQAISMWALIISVTSNYSGVVTIVMFLARGIV

FMCVEYCPIFFITGNTLQCIMLVYCFLGYFCTCYFGLFCLLNRYFRLTLGVYDYLVST

QEFRYMNSQGLLPPKNSIDAFKLNIKLLGVGGKPCIKVATVQSKMSDVKCTSVVLLSV

LQQLRVESSSKLWAQCVQLHNDILLAKDTTEAFEKMVSLLSVLLSMQGAVDINKLCEE

MLDNRATLQAIASEFSSLPSYAAFATAQEAYEQAVANGDSEVVLKKLKKSLNVAKSEF

DRDAAMQRKLEKMADQAMTQMYKQARSEDKRAKVTSAMQTMLFTMLRKLDNDALNNII

NNARDGCVPLNIIPLTTAAKLMVVIPDYNTYKNTCDGTTFTYASALWEIQQVVDADSK

IVQLSEISMDNSPNLAWPLIVTALRANSAVKLQNNELSPVALRQMSCAAGTTQTACTD

DNALAYYNTTKGGRFVLALLSDLQDLKWARFPKSDGTGTIYTELEPPCRFVTDTPKGP

KVKYLYFIKGLNNLNRGMVLGSLAATVRLQAGNATEVPANSTVLSFCAFAVDAAKAYK

DYLASGGQPITNCVKMLCTHTGTGQAITVTPEANMDQESFGGASCCLYCRCHIDHPNP

KGFCDLKGKYVQIPTTCANDPVGFTLKNTVCTVCGMWKGYGCSCDQLREPMLQSADAQ

SFLNRVCGVSAARLTPCGTGTSTDVVYRAFDIYNDKVAGFAKFLKTNCCRFQEKDEDD

NLIDSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHDFFKFRIDGDMVPHISRQRLTK

YTMADLVYALRHFDEGNCDTLKEILVTYNCCDDDYFNKKDWYDFVENPDILRVYANLG

ERVRQALLKTVQFCDAMRNAGIVGVLTLDNQDLNGNWYDFGDFIQTTPGSGVPVVDSY

YSLLMPILTLTRALTAESHVDTDLTKPYIKWDLLKYDFTEERLKLFDRYFKYWDQTYH

PNCVNCLDDRCILHCANFNVLFSTVFPLTSFGPLVRKIFVDGVPFVVSTGYHFRELGV

VHNQDVNLHSSRLSFKELLVYAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQTVK

PGNFNKDFYDFAVSKGFFKEGSSVELKHFFFAQDGNAAISDYDYYRYNLPTMCDIRQL

LFVVEVVDKYFDCYDGGCINANQVIVNNLDKSAGFPFNKWGKARLYYDSMSYEDQDAL

FAYTKRNVIPTITQMNLKYAISAKNRARTVAGVSICSTMTNRQFHQKLLKSIAATRGA

TVVIGTSKFYGGWHNMLKTVYSDVENPHLMGWDYPKCDRAMPNMLRIMASLVLARKHT

TCCSLSHRFYRLANECAQVLSEMVMCGGSLYVKPGGTSSGDATTAYANSVFNICQAVT

ANVNALLSTDGNKIADKYVRNLQHRLYECLYRNRDVDTDFVNEFYAYLRKHFSMMILS

DDAVVCFNSTYASQGLVASIKNFKSVLYYQNNVFMSEAKCWTETDLTKGPHEFCSQHT

MLVKQGDDYVYLPYPDPSRILGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKHPNQ

EYADVFHLYLQYIRKLHDELTGHMLDMYSVMLTNDNTSRYWEPEFYEAMYTPHTVLQA

VGACVLCNSQTSLRCGACIRRPFLCCKCCYDHVISTSHKLVLSVNPYVCNAPGCDVTD

VTQLYLGGMSYYCKSHKPPISFPLCANGQVFGLYKNTCVGSDNVTDFNAIATCDWTNA

GDYILANTCTERLKLFAAETLKATEETFKLSYGIATVREVLSDRELHLSWEVGKPRPP

LNRNYVFTGYRVTKNSKVQIGEYTFEKGDYGDAVVYRGTTTYKLNVGDYFVLTSHTVM

PLSAPTLVPQEHYVRITGLYPTLNISDEFSSNVANYQKVGMQKYSTLQGPPGTGKSHF

AIGLALYYPSARIVYTACSHAAVDALCEKALKYLPIDKCSRIIPARARVECFDKFKVN

STLEQYVFCTVNALPETTADIVVFDEISMATNYDLSVVNARLCAKHYVYIGDPAQLPA

PRTLLTKGTLEPEYFNSVCRLMKTIGPDMFLGTCRRCPAEIVDTVSALVYDNKLKAHK

DKSAQCFKMFYKGVITHDVSSAINRPQIGVVREFLTRNPAWRKAVFISPYNSQNAVAS

KILGLPTQTVDSSQGSEYDYVIFTQTTETAHSCNVNRFNVAITRAKVGILCIMSDRDL

YDKLQFTSLEIPRRNVATLQAENVTGLFKDCSKVITGLHPTQAPTHLSVDTKFKTEGL

CVDVPGIPKDMTYRRLISMMGFKMNYQVNGYPNMFITREEAIRHVRAWIGFDVEGCHA

TREAVGTNLPLQLGFSTGVNLVAVPTGYVDTPNNTDFSRVSAKPPPGDQFKHLIPLMY

KGLPWNVVRIKIVQMLSDTLKNLSDRVVFVLWAHGFELTSMKYFVKIGPERTCCLCDR

RATCFSTASDTYACWHHSIGFDYVYNPFMIDVQQWGFTGNLQSNHDLYCQVHGNAHVA

SCDAIMTRCLAVHECFVKRVDWTIEYPIIGDELKINAACRKVQHMVVKAALLADKFPV

LHDIGNPKAIKCVPQADVEWKFYDAQPCSDKAYKIEELFYSYATHSDKFTDGVCLFWN

CNVDRYPANSIVCRFDTRVLSNLNLPGCDGGSLYVNKHAFHTPAFDKSAFVNLKQLPF

FYYSDSPCESHGKQVVSDIDYVPLKSATCITRCNLGGAVCRHHANEYRLYLDAYNMMI

SAGFSLWVYKQFDTYNLWNTFTRLQSLENVAFNVVNKGHFDGQQGEVPVSIINNTVYT

KVDGVDVELFENKTTLPVNVAFELWAKRNIKPVPEVKILNNLGVDIAANTVIWDYKRD

APAHISTIGVCSMTDIAKKPIETICAPLTVFFDGRVDGQVDLFRNARNGVLITEGSVK

GLQPSVGPKQASLNGVTLIGEAVKTQFNYYKKVDGVVQQLPETYFTQSRNLQEFKPRS

QMEIDFLELAMDEFIERYKLEGYAFEHIVYGDFSHSQLGGLHLLIGLAKRFKESPFEL

EDFIPMDSTVKNYFITDAQTGSSKCVCSVIDLLLDDFVEIIKSQDLSVVSKVVKVTID

YTEISFMLWCKDGHVETFYPKLQSSQAWQ"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94056.1?from=1&to=180) 212..751

/gene="ORF1ab"

/product="leader protein"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94056.1?from=181&to=818) 752..2665

/gene="ORF1ab"

/product="nsp2"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94056.1?from=819&to=2763) 2666..8500

/gene="ORF1ab"

/product="nsp3"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94056.1?from=2764&to=3263) 8501..10000

/gene="ORF1ab"

/product="nsp4"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94056.1?from=3264&to=3569) 10001..10918

/gene="ORF1ab"

/product="3C-like proteinase"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94056.1?from=3570&to=3856) 10919..11779

/gene="ORF1ab"

/product="nsp6"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94056.1?from=3857&to=3939) 11780..12028

/gene="ORF1ab"

/product="nsp7"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94056.1?from=3940&to=4137) 12029..12622

/gene="ORF1ab"

/product="nsp8"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94056.1?from=4138&to=4250) 12623..12961

/gene="ORF1ab"

/product="nsp9"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94056.1?from=4251&to=4389) 12962..13378

/gene="ORF1ab"

/product="nsp10"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94056.1?from=4390&to=5321) join(13379..13405,13405..16173)

/gene="ORF1ab"

/product="RNA-dependent RNA polymerase"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94056.1?from=5322&to=5922) 16174..17976

/gene="ORF1ab"

/product="helicase"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94056.1?from=5923&to=6449) 17977..19557

/gene="ORF1ab"

/product="3'-to-5' exonuclease"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94056.1?from=6450&to=6795) 19558..20595

/gene="ORF1ab"

/product="endoRNAse"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94056.1?from=6796&to=6801) 20596..>20613

/gene="ORF1ab"

/product="2'-O-ribose methyltransferase"

[CDS](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=212&to=13420) 212..13420

/gene="ORF1ab"

/codon\_start=1

/product="ORF1a polyprotein"

/protein\_id="[UPX94055.1](https://www.ncbi.nlm.nih.gov/protein/2234631603)"

/translation="MESLVPGFNEKTHVQLSLPVLQVRDVLVRGFGDSVEEVLSEARQ

HLKDGTCGLVEVEKGVLPQLEQPYVFIKRSDARTAPHGHVMVELVAELEGIQYGRSGE

TLGVLVPHVGEIPVAYRKVLLRKNGNKGAGGHRYGADLKSFDLGDELGTDPYEDFQEN

WNTKHSSGVTRELMRELNGGAYTRYVDNNFCGPDGYPLECIKDLLARAGKASCTLSEQ

LDFIDTKRGVYCCREHEHEIAWYTERSEKSYELQTPFEIKLAKKFDTFNGECPNFVFP

LNSIIKTIQPRVEKKKLDGFMGRIRSVYPVASPNECNQMCLSTLMKCDHCGETSWQTG

DFVKATCEFCGTENLTKEGATTCGYLPQNAVVKIYCPACHNSEVGPEHSLAEYHNESG

LKTILRKGGRTIAFGGCVFSYVGCHNKCAYWVPRASANIGCNHTGVVGEGSEGLNDNL

LEILQKEKVNINIVGDFKLNEEIAIILASFSASTSAFVETVKGLDYKAFKQIVESCGN

FKVTKGKAKKGAWNIGEQKSILSPLYAFASEAARVVRSIFSRTLETAQNSVRVLQKAA

ITILDGISQYSLRLIDAMMFTSDLATNNLVVMAYITGGVVQLTSQWLTNIFGTVYEKL

KPVLDWLEEKFKEGVEFLRDGWEIVKFISTCACEIVGGQIVTCAKEIKESVQTFFKLV

NKFLALCADSIIIGGAKLKALNLGETFVTHSKGLYRKCVKSREETGLLMPLKAPKEII

FLEGETLPTEVLTEEVVLKTGDLQPLEQPTSEAVEAPLVGTPVCINGLMLLEIKDTEK

YCALAPNMMVTNNTFTLKGGAPTKVTFGDDTVIEVQGYKSVNIIFELDERIDKVLNEK

CSAYTVELGTEVNEFACVVADAVIKTLQPVSELLTPLGIDLDEWSMATYYLFDESGEF

KLASHMYCSFYPPDEDEEEGDCEEEEFEPSTQYEYGTEDDYQGKPLEFGATSAALQPE

EEQEEDWLDDDSQQTVGQQDGSEDNQTTTIQTIVEVQPQLEMELTPVVQTIEVNSFSG

YLKLTDNVYIKNADIVEEAKKVKPTVVVNAANVYLKHGGGVAGALNKATNNAMQVESD

DYIATNGPLKVGGSCVLSGHNLAKHCLHVVGPNVNKGEDIQLLKSAYENFNQHEVLLA

PLLSAGIFGADPIHSLRVCVDTVRTNVYLAVFDKNLYDKLVSSFLEMKSEKQVEQKIA

EIPKEEVKPFITESKPSVEQRKQDDKKIKACVEEVTTTLEETKFLTENLLLYIDINGN

LHPDSATLVSDIDITFLKKDAPYIVGDVVQEGVLTAVVIPTKKASGTTEMLAKALRKV

PTDNYITTYPGQGLNGYTVEEAKTVLKKCKSAFYILPSIISNEKQEILGTVSWNLREM

LAHAEETRKLMPVCVETKAIVSTIQRKYKGIKIQEGVVDYGARFYFYTSKTTVASLIN

TLNDLNETLVTMPLGYVTHGLNLEEAARYMRSLKVPATVSVSSPDAVTAYNGYLTSSS

KTPEEHFIETISLAGSYKDWSYSGQSTQLGIEFLKRGDKSVYYTSNPTTFHLDGEVIT

FDNLKTLLSLREVRTIKVFTTVDNINLHTQVVDMSMTYGQQFGPTYLDGADVTKIKPH

NSHEGKTFYVLPNDDTLRVEAFEYYHTTDPSFLGRYMSALNHTKKWKYPQVNGLTSIK

WADNNCYLATALLTLQQIELKFNPPALQDAYYRARAGEAANFCALILAYCNKTVGELG

DVRETMSYLFQHANLDSCKRVLNVVCKTCGQQQTTLKGVEAVMYMGTLSYEQFKKGVQ

IPCTCGKQATKYLVQQESPFVMMSAPPAQYELKHGTFTCASEYTGNYQCGHYKHITSK

ETLYCIDGALLTKSSEYKGPITDVFYKENSYTTTIKPVTYKLDGVVCTEIDPKLDNYY

KKDNSYFTEQPIDLVPNQPYPNASFDNFKFVCDNIKFADDLNQLTGYKKPASRELKVT

FFPDLNGDVVAIDYKHYTPSFKKGAKLLHKPIVWHVNNATNKATYKPNTWCIRCLWST

KPVETSNSFDVLKSEDAQGMDNLACEDLKPVSEEVVENPTIQKDVLECNVKTTEVVGD

IILKPANNSLKITEEVGHTDLMAAYVDNSSLTIKKPNELSRVLGLKTLATHGLAAVNS

VPWDTIANYAKPFLNKVVSTTTNIVTRCLNRVCTNYMPYFFTLLLQLCTFTRSTNSRI

KASMPTTIAKNTVKSVGKFCLEASFNYLKSPNFSKLINIIIWFLLLSVCLGSLIYSTA

ALGVLMSNLGMPSYCTGYREGYLNSTNVTIATYCTGSIPCSVCLSGLDSLDTYPSLET

IQITISSFKWDLTAFGLVAEWFLAYILFTRFFYVLGLAAIMQLFFSYFAVHFISNSWL

MWLIINLVQMAPISAMVRMYIFFASFYYVWKSYVHVVDGCNSSTCMMCYKRNRATRVE

CTTIVNGVRRSFYVYANGGKGFCKLHNWNCVNCDTFCAGSTFISDEVARDLSLQFKRP

INPTDQSSYIVDSVTVKNGSIHLYFDKAGQKTYERHSLSHFVNLDNLRANNTKGSLPI

NVIVFDGKSKCEESSAKSASVYYSQLMCQPILLLDQALVSDVGDSAEVAVKMFDAYVN

TFSSTFNVPMEKLKTLVATAEAELAKNVSLDNVLSTFISAARQGFVDSDVETKDVVEC

LKLSHQSDIEVTGDSCNNYMLTYNKVENMTPRDLGACIDCSARHINAQVAKSHNIALI

WNVKDFMSLSEQLRKQIRSAAKKNNLPFKLTCATTRQVVNVVTTKIALKGGKIVNNWL

KQLIKVTLVFLFVAAIFYLITPVHVMSKHTDFSSEIIGYKAIDGGVTRDIASTDTCFA

NKHADFDTWFSQRGGSYTNDKACPLIAAVITREVGFVVPGLPGTILRTTNGDFLHFLP

RVFSAVGNICYTPSKLIEYTDFATSACVLAAECTIFKDASGKPVPYCYDTNVLEGSVA

YESLRPDTRYVLMDGSIIQFPNTYLEGSVRVVTTFDSEYCRHGTCERSEAGVCVSTSG

RWVLNNDYYRSLPGVFCGVDAVNLFTNMFTPLIQPIGALDISASIVAGGIVAIVVTCL

AYYFMRFRRAFGEYSHVVAFNTLLFLMSFIVLCLTPVYSFLPGVYSVIYLYLTFYLTN

DVSFLAHIQWMVMFTPLVPFWITIAYIICISTKHFYWFFSNYLKRRVVFNGVSFSTFE

EAALCTFLLNKEMYLKLRSDVLLPFTQYNRYLALYNKYKYFSGAMDTTSYREAACCHL

AKALNDFSNSGSDVLYQPPQISITSAVLQSGFRKMAFPSGKVEGCMVQVTCGTTTLNG

LWLDDVVYCPRHVICTSEDMLNPNYEDLLIRKSNHNFLVQAGNVQLRVIGHSMQNCVL

KLKVDTANPKTPKYKFVRIQPGQTFSVLACYNGSPSGVYQCAMRHNFTIKGSFLNGSC

GSVGFNIDYDCVSFCYMHHMELPTGVHAGTDLEGNFYGPFVDRQTAQAAGTDTTITVN

VLAWLYAAVINGDRWFLNRFTTTLNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAV

LDMCASLKELLQNGMNGRTILGSALLEDEFTPFDVVRQCSGVTFQSAVKRTIKGTHHW

LLLTILTSLLVLVQSTQWSLFFFLYENAFLPFAMGIIAMSAFAMMFVKHKHAFLCLFL

LPSLATVAYFNMVYMPASWVMRIMTWLDMVDTSLKLKDCVMYASAVVLLILMTARTVY

DDGARRVWTLMNVLTLVYKVYYGNALDQAISMWALIISVTSNYSGVVTIVMFLARGIV

FMCVEYCPIFFITGNTLQCIMLVYCFLGYFCTCYFGLFCLLNRYFRLTLGVYDYLVST

QEFRYMNSQGLLPPKNSIDAFKLNIKLLGVGGKPCIKVATVQSKMSDVKCTSVVLLSV

LQQLRVESSSKLWAQCVQLHNDILLAKDTTEAFEKMVSLLSVLLSMQGAVDINKLCEE

MLDNRATLQAIASEFSSLPSYAAFATAQEAYEQAVANGDSEVVLKKLKKSLNVAKSEF

DRDAAMQRKLEKMADQAMTQMYKQARSEDKRAKVTSAMQTMLFTMLRKLDNDALNNII

NNARDGCVPLNIIPLTTAAKLMVVIPDYNTYKNTCDGTTFTYASALWEIQQVVDADSK

IVQLSEISMDNSPNLAWPLIVTALRANSAVKLQNNELSPVALRQMSCAAGTTQTACTD

DNALAYYNTTKGGRFVLALLSDLQDLKWARFPKSDGTGTIYTELEPPCRFVTDTPKGP

KVKYLYFIKGLNNLNRGMVLGSLAATVRLQAGNATEVPANSTVLSFCAFAVDAAKAYK

DYLASGGQPITNCVKMLCTHTGTGQAITVTPEANMDQESFGGASCCLYCRCHIDHPNP

KGFCDLKGKYVQIPTTCANDPVGFTLKNTVCTVCGMWKGYGCSCDQLREPMLQSADAQ

SFLNGFAV"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94055.1?from=1&to=180) 212..751

/gene="ORF1ab"

/product="leader protein"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94055.1?from=181&to=818) 752..2665

/gene="ORF1ab"

/product="nsp2"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94055.1?from=819&to=2763) 2666..8500

/gene="ORF1ab"

/product="nsp3"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94055.1?from=2764&to=3263) 8501..10000

/gene="ORF1ab"

/product="nsp4"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94055.1?from=3264&to=3569) 10001..10918

/gene="ORF1ab"

/product="3C-like proteinase"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94055.1?from=3570&to=3856) 10919..11779

/gene="ORF1ab"

/product="nsp6"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94055.1?from=3857&to=3939) 11780..12028

/gene="ORF1ab"

/product="nsp7"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94055.1?from=3940&to=4137) 12029..12622

/gene="ORF1ab"

/product="nsp8"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94055.1?from=4138&to=4250) 12623..12961

/gene="ORF1ab"

/product="nsp9"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94055.1?from=4251&to=4389) 12962..13378

/gene="ORF1ab"

/product="nsp10"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UPX94055.1?from=4390&to=4402) 13379..13417

/gene="ORF1ab"

/product="nsp11"

[stem\_loop](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=13413&to=13440) 13413..13440

/gene="ORF1ab"

/note="Coronavirus frameshifting stimulation element

stem-loop 1"

[stem\_loop](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=13425&to=13479) 13425..13479

/gene="ORF1ab"

/note="Coronavirus frameshifting stimulation element

stem-loop 2"

gap 20614..21499

/estimated\_length=886

[gene](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=21500&to=25312) 21500..25312

/gene="S"

[CDS](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=21500&to=25312) 21500..25312

/gene="S"

/codon\_start=1

/product="surface glycoprotein"

/protein\_id="[UPX94057.1](https://www.ncbi.nlm.nih.gov/protein/2234631605)"

/translation="MFVFLVLLPLVSSQCVNLTTRTQSYTNSFTRGVYYPDKVFRSSV

LHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWI

FGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLDVYYHKNNKSWMESEFRVYSSA

NNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLGRDLPQGFS

ALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKY

NENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCP

FDEVFNATRFASVYAWNRKRISNCVADYSVLYNFAPFFAFKCYGVSPTKLNDLCFTNV

YADSFVIRGNEVSQIAPGQTGNIADYNYKLPDDFTGCVIAWNSNKLDSKVGGNYNYLY

RLFRKSNLKPFERDISTEIYQAGNKPCNGVAGFNCYFPLRSYGFRPTYGVGHQPYRVV

VLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDI

ADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHAD

QLTPTWRVYSTGSNVFQTRAGCLIGAEYVNNSYECDIPIGAGICASYQTQTKSHRRAR

SVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYIC

GDSTECSNLLLQYGSFCTQLKRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKYFGGFN

FSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLT

VLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLY

ENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNHNAQALNTLVKQLSSKFGAISS

VLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECV

LGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPR

EGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFK

EELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYE

QYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVL

KGVKLHYT"

[gene](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=25321&to=26148) 25321..26148

/gene="ORF3a"

[CDS](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=25321&to=26148) 25321..26148

/gene="ORF3a"

/codon\_start=1

/product="ORF3a protein"

/protein\_id="[UPX94058.1](https://www.ncbi.nlm.nih.gov/protein/2234631606)"

/translation="MDLFMRIFTIGTVTLKQGEIKDATPSDFVRATATIPIQASLPFG

WLIVGVALLAVFQSASKIITLKKRWQLALSKGVHFVCNLLLLFVTVYSHLLLVAAGLE

APFLHLYALVYFLQSINFVRIIMRLWLCWKCRSKNPLLYDANYFLCWHTNCYDYCIPY

NSVTSSIVITSGDGTTSPISEHDYQIGGYTEKWESGVKDCVVLHSYFTSDYYQLYSTQ

LSTDIGVEHVTFFIYNKIVDEPEEHVQIHTIDGSSGVVNPVMEPIYDEPTTTTSVPL"

[gene](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=26173&to=26400) 26173..26400

/gene="E"

[CDS](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=26173&to=26400) 26173..26400

/gene="E"

/codon\_start=1

/product="envelope protein"

/protein\_id="[UPX94059.1](https://www.ncbi.nlm.nih.gov/protein/2234631607)"

/translation="MYSFVSEEIGTLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCC

NIVNVSLVKPSFYVYSRVKNLNSSRVPDLLV"

[gene](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=26451&to=27119) 26451..27119

/gene="M"

[CDS](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=26451&to=27119) 26451..27119

/gene="M"

/codon\_start=1

/product="membrane glycoprotein"

/protein\_id="[UPX94060.1](https://www.ncbi.nlm.nih.gov/protein/2234631608)"

/translation="MADSNGTITVEELKKLLEEWNLVIGFLFLTWICLLQFAYANRNR

FLYIIKLIFLWLLWPVTLTCFVLAAVYRINWITGGIAIAMACLVGLMWLSYFIASFRL

FARTRSMWSFNPETNILLNVPLHGTILTRPLLESELVIGAVILRGHLRIAGHHLGRCD

IKDLPKEITVATSRTLSYYKLGASQRVAGDSGFAAYSRYRIGNYKLNTDHSSSSDNIA

LLVQ"

[gene](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=27130&to=27315) 27130..27315

/gene="ORF6"

[CDS](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=27130&to=27315) 27130..27315

/gene="ORF6"

/codon\_start=1

/product="ORF6 protein"

/protein\_id="[UPX94061.1](https://www.ncbi.nlm.nih.gov/protein/2234631609)"

/translation="MFHLVDFQVTIAEILLIIMRTFKVSIWNLDYIINLIIKNLSKSL

TENKYSQLDEEQPMEIL"

[gene](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=27322&to=27687) 27322..27687

/gene="ORF7a"

[CDS](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=27322&to=27687) 27322..27687

/gene="ORF7a"

/codon\_start=1

/product="ORF7a protein"

/protein\_id="[UPX94062.1](https://www.ncbi.nlm.nih.gov/protein/2234631610)"

/translation="MKIILFLALITLATCELYHYQECVRGTTVLLKEPCSSGTYEGNS

PFHPLADNKFALTCFSTQFAFACPDGVKHVYQLRARSVSPKLFIRQEEVQELYSPIFL

IVAAIVFITLCFTLKRKTE"

[gene](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=27684&to=27815) 27684..27815

/gene="ORF7b"

[CDS](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=27684&to=27815) 27684..27815

/gene="ORF7b"

/codon\_start=1

/product="ORF7b"

/protein\_id="[UPX94063.1](https://www.ncbi.nlm.nih.gov/protein/2234631611)"

/translation="MIELSLIDFYLCFLAFLLFLVLIMLIIFWFSLELQDHNETCHA"

[gene](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=27822&to=28187) 27822..28187

/gene="ORF8"

[CDS](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=27822&to=28187) 27822..28187

/gene="ORF8"

/codon\_start=1

/product="ORF8 protein"

/protein\_id="[UPX94064.1](https://www.ncbi.nlm.nih.gov/protein/2234631612)"

/translation="MKFLVFLGIITTVAAFHQECSLQSCTQHQPYVVDDPCPIHFYSK

WYIRVGARKSAPLIELCVDEAGSKSPIQYIDIGNYTVSCLPFTINCQEPKLGSLVVRC

SFYEDFLEYHDVRVVLDFI"

[gene](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=28202&to=29452) 28202..29452

/gene="N"

[CDS](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=28202&to=29452) 28202..29452

/gene="N"

/codon\_start=1

/product="nucleocapsid phosphoprotein"

/protein\_id="[UPX94065.1](https://www.ncbi.nlm.nih.gov/protein/2234631613)"

/translation="MSDNGPQNQRNALRITFGGPSDSTGSNQNGGARSKQRRPQGLPN

NTASWFTALTQHGKEDLKFPRGQGVPINTNSSPDDQIGYYRRATRRIRGGDGKMKDLS

PRWYFYYLGTGPEAGLPYGANKDGIIWVATEGALNTPKDHIGTRNPANNAAIVLQLPQ

GTTLPKGFYAEGSRGGSQASSRSSSRSRNSSRNSTPGSSKRTSPARMAGNGGDAALAL

LLLDRLNQLESKMSGKGQQQQGQTVTKKSAAEASKKPRQKRTATKAYNVTQAFGRRGP

EQTQGNFGDQELIRQGTDYKHWPQIAQFAPSASAFFGMSRIGMEVTPSGTWLTYTGAI

KLDDKDPNFKDQVILLNKHIDAYKTFPPTEPKKDKKKKADETQALPQRQKKQQTVTLL

PAADLDDFSKQLQQSMSRADSTQA"

[gene](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=29477&to=29593) 29477..29593

/gene="ORF10"

[CDS](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=29477&to=29593) 29477..29593

/gene="ORF10"

/codon\_start=1

/product="ORF10 protein"

/protein\_id="[UPX94066.1](https://www.ncbi.nlm.nih.gov/protein/2234631614)"

/translation="MGYINVFAFPFTIYSLLLCRMNSRNYIAQVDVVNFNLT"

[stem\_loop](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=29528&to=29563) 29528..29563

/gene="ORF10"

/note="Coronavirus 3' UTR pseudoknot stem-loop 1"

[stem\_loop](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=29548&to=29576) 29548..29576

/gene="ORF10"

/note="Coronavirus 3' UTR pseudoknot stem-loop 2"

[stem\_loop](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?from=29647&to=29651) 29647..>29651

/note="Coronavirus 3' stem-loop II-like motif (s2m)"

ORIGIN

1 agatctgttc tctaaacgaa ctttaaaatc tgtgtggctg tcactcggct gcatgcttag

61 tgcactcacg cagtataatt aataactaat tactgtcgtt gacaggacac gagtaactcg

121 tctatcttct gcaggctgct tacggtttcg tccgtgttgc agccgatcat cagcacatct

181 aggttttgtc cgggtgtgac cgaaaggtaa gatggagagc cttgtccctg gtttcaacga

241 gaaaacacac gtccaactca gtttgcctgt tttacaggtt cgcgacgtgc tcgtacgtgg

301 ctttggagac tccgtggagg aggtcttatc agaggcacgt caacatctta aagatggcac

361 ttgtggctta gtagaagttg aaaaaggcgt tttgcctcaa cttgaacagc cctatgtgtt

421 catcaaacgt tcggatgctc gaactgcacc tcatggtcat gttatggttg agctggtagc

481 agaactcgaa ggcattcagt acggtcgtag tggtgagaca cttggtgtcc ttgtccctca

541 tgtgggcgaa ataccagtgg cttaccgcaa ggttcttctt cgtaagaacg gtaataaagg

601 agctggtggc cataggtacg gcgccgatct aaagtcattt gacttaggcg acgagcttgg

661 cactgatcct tatgaagatt ttcaagaaaa ctggaacact aaacatagca gtggtgttac

721 ccgtgaactc atgcgtgagc ttaacggagg ggcatacact cgctatgtcg ataacaactt

781 ctgtggccct gatggctacc ctcttgagtg cattaaagac cttctagcac gtgctggtaa

841 agcttcatgc actttgtccg aacaactgga ctttattgac actaagaggg gtgtatactg

901 ctgccgtgaa catgagcatg aaattgcttg gtacacggaa cgttctgaaa agagctatga

961 attgcagaca ccttttgaaa ttaaattggc aaagaaattt gacaccttca atggggaatg

1021 tccaaatttt gtatttccct taaattccat aatcaagact attcaaccaa gggttgaaaa

1081 gaaaaagctt gatggcttta tgggtagaat tcgatctgtc tatccagttg cgtcaccaaa

1141 tgaatgcaac caaatgtgcc tttcaactct catgaagtgt gatcattgtg gtgaaacttc

1201 atggcagacg ggcgattttg ttaaagccac ttgcgaattt tgtggcactg agaatttgac

1261 taaagaaggt gccactactt gtggttactt accccaaaat gctgttgtta aaatttattg

1321 tccagcatgt cacaattcag aagtaggacc tgagcatagt cttgccgaat accataatga

1381 atctggcttg aaaaccattc ttcgtaaggg tggtcgcact attgcctttg gaggctgtgt

1441 gttctcttat gttggttgcc ataacaagtg tgcctattgg gttccacgtg ctagcgctaa

1501 cataggttgt aaccatacag gtgttgttgg agaaggttcc gaaggtctta atgacaacct

1561 tcttgaaata ctccaaaaag agaaagtcaa catcaatatt gttggtgact ttaaacttaa

1621 tgaagagatc gccattattt tggcatcttt ttctgcttcc acaagtgctt ttgtggaaac

1681 tgtgaaaggt ttggattata aagcattcaa acaaattgtt gaatcctgtg gtaattttaa

1741 agttacaaaa ggaaaagcta aaaaaggtgc ctggaatatt ggtgaacaga aatcaatact

1801 gagtcctctt tatgcatttg catcagaggc tgctcgtgtt gtacgatcaa ttttctcccg

1861 cactcttgaa actgctcaaa attctgtgcg tgttttacag aaggccgcta taacaatact

1921 agatggaatt tcacagtatt cactgagact cattgatgct atgatgttca catctgattt

1981 ggctactaac aatctagttg taatggccta cattacaggt ggtgttgttc agttgacttc

2041 gcagtggcta actaacatct ttggcactgt ttatgaaaaa ctcaaacccg tccttgattg

2101 gcttgaagag aagtttaagg aaggtgtaga gtttcttaga gacggttggg aaattgttaa

2161 atttatctca acctgtgctt gtgaaattgt cggtggacaa attgtcacct gtgcaaagga

2221 aattaaggag agtgttcaga cattctttaa gcttgtaaat aaatttttgg ctttgtgtgc

2281 tgactctatc attattggtg gagctaaact taaagccttg aatttaggtg aaacatttgt

2341 cacgcactca aagggattgt acagaaagtg tgttaaatcc agagaagaaa ctggcctact

2401 catgcctcta aaagccccaa aagaaattat cttcttagag ggagaaacac ttcccacaga

2461 agtgttaaca gaggaagttg tcttgaaaac tggtgattta caaccattag aacaacctac

2521 tagtgaagct gttgaagctc cattggttgg tacaccagtt tgtattaacg ggcttatgtt

2581 gctcgaaatc aaagacacag aaaagtactg tgcccttgca cctaatatga tggtaacaaa

2641 caataccttc acactcaaag gcggtgcacc aacaaaggtt acttttggtg atgacactgt

2701 gatagaagtg caaggttaca agagtgtgaa tatcattttt gaacttgatg aaaggattga

2761 taaagtactt aatgagaagt gctctgccta tacagttgaa ctcggtacag aagtaaatga

2821 gttcgcctgt gttgtggcag atgctgtcat aaaaactttg caaccagtat ctgaattact

2881 tacaccactg ggcattgatt tagatgagtg gagtatggct acatactact tatttgatga

2941 gtctggtgag tttaaattgg cttcacatat gtattgttct ttttaccctc cagatgagga

3001 tgaagaagaa ggtgattgtg aagaagaaga gtttgagcca tcaactcaat atgagtatgg

3061 tactgaagat gattaccaag gtaaaccttt ggaatttggt gccacttctg ctgctcttca

3121 acctgaagaa gagcaagaag aagattggtt agatgatgat agtcaacaaa ctgttggtca

3181 acaagacggc agtgaggaca atcagacaac tactattcaa acaattgttg aggttcaacc

3241 tcaattagag atggaactta caccagttgt tcagactatt gaagtgaata gttttagtgg

3301 ttatttaaaa cttactgaca atgtatacat taaaaatgca gacattgtgg aagaagctaa

3361 aaaggtaaaa ccaacagtgg ttgttaatgc agccaatgtt taccttaaac atggaggagg

3421 tgttgcagga gccttaaata aggctactaa caatgccatg caagttgaat ctgatgatta

3481 catagctact aatggaccac ttaaagtggg tggtagttgt gttttaagcg gacacaatct

3541 tgctaaacac tgtcttcatg ttgtcggccc aaatgttaac aaaggtgaag acattcaact

3601 tcttaagagt gcttatgaaa attttaatca gcacgaagtt ctacttgcac cattattatc

3661 agctggtatt tttggtgctg accctataca ttctttaaga gtttgtgtag atactgttcg

3721 cacaaatgtc tacttagctg tctttgataa aaatctctat gacaaacttg tttcaagctt

3781 tttggaaatg aagagtgaaa agcaagttga acaaaagatc gctgagattc ctaaagagga

3841 agttaagcca tttataactg aaagtaaacc ttcagttgaa cagagaaaac aagatgataa

3901 gaaaatcaaa gcttgtgttg aagaagttac aacaactctg gaagaaacta agttcctcac

3961 agaaaacttg ttactttata ttgacattaa tggcaatctt catccagatt ctgccactct

4021 tgttagtgac attgacatca ctttcttaaa gaaagatgct ccatatatag tgggtgatgt

4081 tgttcaagag ggtgttttaa ctgctgtggt tatacctact aaaaaggcta gtggcactac

4141 tgaaatgcta gcgaaagctt tgagaaaagt gccaacagac aattatataa ccacttaccc

4201 gggtcagggt ttaaatggtt acactgtaga ggaggcaaag acagtgctta aaaagtgtaa

4261 aagtgctttt tacattctac catctattat ctctaatgag aagcaagaaa ttcttggaac

4321 tgtttcttgg aatttgcgag aaatgcttgc acatgcagaa gaaacacgca aattaatgcc

4381 tgtctgtgtg gaaactaaag ccatagtttc aactatacag cgtaaatata agggtattaa

4441 aatacaagag ggtgtggttg attatggtgc tagattttac ttttacacca gtaaaacaac

4501 tgtagcgtca cttatcaaca cacttaacga tctaaatgaa actcttgtta caatgccact

4561 tggctatgta acacatggct taaatttgga agaagctgct cggtatatga gatctctcaa

4621 agtgccagct acagtttctg tttcttcacc tgatgctgtt acagcgtata atggttatct

4681 tacttcttct tctaaaacac ctgaagaaca ttttattgaa accatctcac ttgctggttc

4741 ctataaagat tggtcctatt ctggacaatc tacacaacta ggtatagaat ttcttaagag

4801 aggtgataaa agtgtatatt acactagtaa tcctaccaca ttccacctag atggtgaagt

4861 tatcaccttt gacaatctta agacacttct ttctttgaga gaagtgagga ctattaaggt

4921 gtttacaaca gtagacaaca ttaacctcca cacgcaagtt gtggacatgt caatgacata

4981 tggacaacag tttggtccaa cttatttgga tggagctgat gttactaaaa taaaacctca

5041 taattcacat gaaggtaaaa cattttatgt tttacctaat gatgacactc tacgtgttga

5101 ggcttttgag tactaccaca caactgatcc tagttttctg ggtaggtaca tgtcagcatt

5161 aaatcacact aaaaagtgga aatacccaca agttaatggt ttaacttcta ttaaatgggc

5221 agataacaac tgttatcttg ccactgcatt gttaacactc caacaaatag agttgaagtt

5281 taatccacct gctctacaag atgcttatta cagagcaagg gctggtgaag ctgctaactt

5341 ttgtgcactt atcttagcct actgtaataa gacagtaggt gagttaggtg atgttagaga

5401 aacaatgagt tacttgtttc aacatgccaa tttagattct tgcaaaagag tcttgaacgt

5461 ggtgtgtaaa acttgtggac aacagcagac aacccttaag ggtgtagaag ctgttatgta

5521 catgggcaca ctttcttatg aacaatttaa gaaaggtgtt cagatacctt gtacgtgtgg

5581 taaacaagct acaaaatatc tagtacaaca ggagtcacct tttgttatga tgtcagcacc

5641 acctgctcag tatgaactta agcatggtac atttacttgt gctagtgagt acactggtaa

5701 ttaccagtgt ggtcactata aacatataac ttctaaagaa actttgtatt gcatagacgg

5761 tgctttactt acaaagtcct cagaatacaa aggtcctatt acggatgttt tctacaaaga

5821 aaacagttac acaacaacca taaaaccagt tacttataaa ttggatggtg ttgtttgtac

5881 agaaattgac cctaagttgg acaattatta taagaaagac aattcttatt tcacagagca

5941 accaattgat cttgtaccaa accaaccata tccaaacgca agcttcgata attttaagtt

6001 tgtatgtgat aatatcaaat ttgctgatga tttaaaccag ttaactggtt ataagaaacc

6061 tgcttcaaga gagcttaaag ttacattttt ccctgactta aatggtgatg tggtggctat

6121 tgattataaa cactacacac cctcttttaa gaaaggagct aaattgttac ataaacctat

6181 tgtttggcat gttaacaatg caactaataa agccacgtat aaaccaaata cctggtgtat

6241 acgttgtctt tggagcacaa aaccagttga aacatcaaat tcgtttgatg tactgaagtc

6301 agaggacgcg cagggaatgg ataatcttgc ctgcgaagat ctaaaaccag tctctgaaga

6361 agtagtggaa aatcctacca tacagaaaga cgttcttgag tgtaatgtga aaactaccga

6421 agttgtagga gacattatac ttaaaccagc aaataatagt ttaaaaatta cagaagaggt

6481 tggccacaca gatctaatgg ctgcttatgt agacaattct agtcttacta ttaagaaacc

6541 taatgaatta tctagagtat taggtttgaa aacccttgct actcatggtt tagctgctgt

6601 taatagtgtc ccttgggata ctatagctaa ttatgctaag ccttttctta acaaagttgt

6661 tagtacaact actaacatag ttacacggtg tttaaaccgt gtttgtacta attatatgcc

6721 ttatttcttt actttattgc tacaattgtg tacttttact agaagtacaa attctagaat

6781 taaagcatct atgccgacta ctatagcaaa gaatactgtt aagagtgtcg gtaaattttg

6841 tctagaggct tcatttaatt atttgaagtc acctaatttt tctaaactga taaatattat

6901 aatttggttt ttactattaa gtgtttgcct aggttcttta atctactcaa ccgctgcttt

6961 aggtgtttta atgtctaatt taggcatgcc ttcttactgt actggttaca gagaaggcta

7021 tttgaactct actaatgtca ctattgcaac ctactgtact ggttctatac cttgtagtgt

7081 ttgtcttagt ggtttagatt ctttagacac ctatccttct ttagaaacta tacaaattac

7141 catttcatct tttaaatggg atttaactgc ttttggctta gttgcagagt ggtttttggc

7201 atatattctt ttcactaggt ttttctatgt acttggattg gctgcaatca tgcaattgtt

7261 tttcagctat tttgcagtac attttattag taattcttgg cttatgtggt taataattaa

7321 tcttgtacaa atggccccga tttcagctat ggttagaatg tacatcttct ttgcatcatt

7381 ttattatgta tggaaaagtt atgtgcatgt tgtagacggt tgtaattcat caacttgtat

7441 gatgtgttac aaacgtaata gagcaacaag agtcgaatgt acaactattg ttaatggtgt

7501 tagaaggtcc ttttatgttt atgctaatgg aggtaaaggc ttttgcaaac tacacaattg

7561 gaattgtgtt aattgtgata cattctgtgc tggtagtaca tttattagtg atgaagttgc

7621 gagagacttg tcactacagt ttaaaagacc aataaatcct actgaccagt cttcttacat

7681 cgttgatagt gttacagtga agaatggttc catccatctt tactttgata aagctggtca

7741 aaagacttat gaaagacatt ctctctctca ttttgttaac ttagacaacc tgagagctaa

7801 taacactaaa ggttcattgc ctattaatgt tatagttttt gatggtaaat caaaatgtga

7861 agaatcatct gcaaaatcag cgtctgttta ctacagtcag cttatgtgtc aacctatact

7921 gttactagat caggcattag tgtctgatgt tggtgatagt gcggaagttg cagttaaaat

7981 gtttgatgct tacgttaata cgttttcatc aacttttaac gtaccaatgg aaaaactcaa

8041 aacactagtt gcaactgcag aagctgaact tgcaaagaat gtgtccttag acaatgtctt

8101 atctactttt atttcagcag ctcggcaagg gtttgttgat tcagatgtag aaactaaaga

8161 tgttgttgaa tgtcttaaat tgtcacatca atctgacata gaagttactg gcgatagttg

8221 taataactat atgctcacct ataacaaagt tgaaaacatg acaccccgtg accttggtgc

8281 ttgtattgac tgtagtgcgc gtcatattaa tgcgcaggta gcaaaaagtc acaacattgc

8341 tttgatatgg aacgttaaag atttcatgtc attgtctgaa caactacgaa aacaaatacg

8401 tagtgctgct aaaaagaata acttaccttt taagttgaca tgtgcaacta ctagacaagt

8461 tgttaatgtt gtaacaacaa agatagcact taagggtggt aaaattgtta ataattggtt

8521 gaagcagtta attaaagtta cacttgtgtt cctttttgtt gctgctattt tctatttaat

8581 aacacctgtt catgtcatgt ctaaacatac tgacttttca agtgaaatca taggatacaa

8641 ggctattgat ggtggtgtca ctcgtgacat agcatctaca gatacttgtt ttgctaacaa

8701 acatgctgat tttgacacat ggtttagcca gcgtggtggt agttatacta atgacaaagc

8761 ttgcccattg attgctgcag tcataacaag agaagtgggt tttgtcgtgc ctggtttgcc

8821 tggcacgata ttacgcacaa ctaatggtga ctttttgcat ttcttaccta gagtttttag

8881 tgcagttggt aacatctgtt acacaccatc aaaacttata gagtacactg actttgcaac

8941 atcagcttgt gttttggctg ctgaatgtac aatttttaaa gatgcttctg gtaagccagt

9001 accatattgt tatgatacca atgtactaga aggttctgtt gcttatgaaa gtttacgccc

9061 tgacacacgt tatgtgctca tggatggctc tattattcaa tttcctaaca cctaccttga

9121 aggttctgtt agagtggtaa caacttttga ttctgagtac tgtaggcacg gcacttgtga

9181 aagatcagaa gctggtgttt gtgtatctac tagtggtaga tgggtactta acaatgatta

9241 ttacagatct ttaccaggag ttttctgtgg tgtagatgct gtaaatttat ttactaatat

9301 gtttacacca ctaattcaac ctattggtgc tttggacata tcagcatcta tagtagctgg

9361 tggtattgtg gctatcgtag taacatgcct tgcctactat tttatgaggt ttagaagagc

9421 ttttggtgaa tacagtcatg tagttgcctt taatacttta ctattcctta tgtcattcat

9481 tgtactctgt ttaacaccag tttactcatt cttacctggt gtttattctg ttatttactt

9541 gtacttgaca ttttatctta ctaatgatgt ttctttttta gcacatattc agtggatggt

9601 tatgttcaca cctttagtac ctttctggat aacaattgct tatatcattt gtatttccac

9661 aaagcatttc tattggttct ttagtaatta cctaaagaga cgtgtagtct ttaatggtgt

9721 ttcctttagt acttttgaag aagctgcgct gtgcaccttt ttgttaaata aagaaatgta

9781 tctaaagttg cgtagtgatg tgctattacc ttttacgcaa tataatagat acttagctct

9841 ttataataag tacaagtatt ttagtggagc aatggataca actagctaca gagaagctgc

9901 ttgttgtcat ctcgcaaagg ctctcaatga cttcagtaac tcaggttctg atgttcttta

9961 ccaaccacca caaatctcta tcacctcagc tgttttgcag agtggtttta gaaaaatggc

10021 attcccatct ggtaaagttg agggttgtat ggtacaagta acttgtggta caactacact

10081 taacggtctt tggcttgatg acgtagttta ctgtccaaga catgtgatct gcacctctga

10141 agatatgctt aaccctaatt atgaagattt actcatccgt aagtctaatc ataatttctt

10201 ggtacaggct ggtaatgttc aactcagggt tattggacat tctatgcaaa attgtgtact

10261 taagcttaag gttgatacag ccaatcctaa gacacctaag tataagtttg ttcgcattca

10321 accaggacag actttttcag tgttagcttg ttacaatggt tcaccatctg gtgtttacca

10381 atgtgctatg agacacaatt tcactattaa gggttcattc cttaatggtt catgtggtag

10441 tgttggtttt aacatagatt atgactgtgt ctctttttgt tacatgcacc atatggaatt

10501 accaactgga gttcatgctg gcacagactt agaaggtaac ttttatggac cttttgttga

10561 caggcaaaca gcacaagcag ctggtacgga cacaactatt acagttaatg ttttagcttg

10621 gttgtacgct gctgttataa atggagacag gtggtttctc aatcgattta ccacaactct

10681 taatgacttt aaccttgtgg ctatgaagta caattatgaa cctctaacac aagaccatgt

10741 tgacatacta ggacctcttt ctgctcaaac tggaattgcc gttttagata tgtgtgcttc

10801 attaaaagaa ttactgcaaa atggtatgaa tggacgtacc atattgggta gtgctttatt

10861 agaagatgaa tttacacctt ttgatgttgt tagacaatgc tcaggtgtta ctttccaaag

10921 tgcagtgaaa agaacaatca agggtacaca ccactggttg ttactcacaa ttttgacttc

10981 acttttagtt ttagtccaga gtactcaatg gtctttgttc ttttttttgt atgaaaatgc

11041 ctttttacct tttgctatgg gtattattgc tatgtctgct tttgcaatga tgtttgtcaa

11101 acataagcat gcatttctct gtttgttttt gttaccttct cttgccactg tagcttattt

11161 taatatggtc tatatgcctg ctagttgggt gatgcgtatt atgacatggt tggatatggt

11221 tgatactagt ttgaagctaa aagactgtgt tatgtatgca tcagctgtag tgttactaat

11281 ccttatgaca gcaagaactg tgtatgatga tggtgctagg agagtgtgga cacttatgaa

11341 tgtcttgaca ctcgtttata aagtttatta tggtaatgct ttagatcaag ccatttccat

11401 gtgggctctt ataatctctg ttacttctaa ctactcaggt gtagttacaa ttgtcatgtt

11461 tttggccaga ggtattgttt ttatgtgtgt tgagtattgc cctattttct tcataactgg

11521 taatacactt cagtgtataa tgctagttta ttgtttctta ggctattttt gtacttgtta

11581 ctttggcctc ttttgtttac tcaaccgcta ctttagactg actcttggtg tttatgatta

11641 cttagtttct acacaggagt ttagatatat gaattcacag ggactactcc cacccaagaa

11701 tagcatagat gccttcaaac tcaacattaa attgttgggt gttggtggca aaccttgtat

11761 caaagtagcc actgtacagt ctaaaatgtc agatgtaaag tgcacatcag tagtcttact

11821 ctcagttttg caacaactca gagtagaatc atcatctaaa ttgtgggctc aatgtgtcca

11881 gttacacaat gacattctct tagctaaaga tactactgaa gcctttgaaa aaatggtttc

11941 actactttct gttttgcttt ccatgcaggg tgctgtagac ataaacaagc tttgtgaaga

12001 aatgctggac aacagggcaa ccttacaagc tatagcctca gagtttagtt cccttccatc

12061 atatgcagct tttgctactg ctcaagaagc ttatgagcag gctgttgcta atggtgattc

12121 tgaagttgtt cttaaaaagt tgaagaagtc tttgaatgtg gctaaatctg aatttgaccg

12181 tgatgcagcc atgcaacgta agttggaaaa gatggctgat caagctatga cccaaatgta

12241 taaacaggct agatctgagg acaagagggc aaaagttact agtgctatgc agacaatgct

12301 tttcactatg cttagaaagt tggataatga tgcactcaac aacattatca acaatgcaag

12361 agatggttgt gttcccttga acataatacc tcttacaaca gcagccaaac taatggttgt

12421 cataccagac tataacacat ataaaaatac gtgtgatggt acaacattta cttatgcatc

12481 agcattgtgg gaaatccaac aggttgtaga tgcagatagt aaaattgttc aacttagtga

12541 aattagtatg gacaattcac ctaatttagc atggcctctt attgtaacag ctttaagggc

12601 caattctgct gtcaaattac agaataatga gcttagtcct gttgcactac gacagatgtc

12661 ttgtgctgcc ggtactacac aaactgcttg cactgatgac aatgcgttag cttactacaa

12721 cacaacaaag ggaggtaggt ttgtacttgc actgttatcc gatttacagg atttgaaatg

12781 ggctagattc cctaagagtg atggaactgg tactatttat acagaactgg aaccaccttg

12841 taggtttgtt acagacacac ctaaaggtcc taaagtgaag tatttatact ttattaaagg

12901 attaaacaac ctaaatagag gtatggtact tggtagttta gctgccacag tacgtctaca

12961 agctggtaat gcaacagaag tgcctgccaa ttcaactgta ttatctttct gtgcttttgc

13021 tgtagatgct gctaaagctt acaaagatta tctagctagt gggggacaac caatcactaa

13081 ttgtgttaag atgttgtgta cacacactgg tactggtcag gcaataacag ttacaccgga

13141 agccaatatg gatcaagaat cctttggtgg tgcatcgtgt tgtctgtact gccgttgcca

13201 catagatcat ccaaatccta aaggattttg tgacttaaaa ggtaagtatg tacaaatacc

13261 tacaacttgt gctaatgacc ctgtgggttt tacacttaaa aacacagtct gtaccgtctg

13321 cggtatgtgg aaaggttatg gctgtagttg tgatcaactc cgcgaaccca tgcttcagtc

13381 agctgatgca caatcgtttt taaacgggtt tgcggtgtaa gtgcagcccg tcttacaccg

13441 tgcggcacag gcactagtac tgatgtcgta tacagggctt ttgacatcta caatgataaa

13501 gtagctggtt ttgctaaatt cctaaaaact aattgttgtc gcttccaaga aaaggacgaa

13561 gatgacaatt taattgattc ttactttgta gttaagagac acactttctc taactaccaa

13621 catgaagaaa caatttataa tttacttaag gattgtccag ctgttgctaa acatgacttc

13681 tttaagttta gaatagacgg tgacatggta ccacatatat cacgtcaacg tcttactaaa

13741 tacacaatgg cagacctcgt ctatgcttta aggcattttg atgaaggtaa ttgtgacaca

13801 ttaaaagaaa tacttgtcac atacaattgt tgtgatgatg attatttcaa taaaaaggac

13861 tggtatgatt ttgtagaaaa cccagatata ttacgcgtat acgccaactt aggtgaacgt

13921 gtacgccaag ctttgttaaa aacagtacaa ttctgtgatg ccatgcgaaa tgctggtatt

13981 gttggtgtac tgacattaga taatcaagat ctcaatggta actggtatga tttcggtgat

14041 ttcatacaaa ccacgccagg tagtggagtt cctgttgtag attcttatta ttcattgtta

14101 atgcctatat taaccttgac cagggcttta actgcagagt cacatgttga cactgactta

14161 acaaagcctt acattaagtg ggatttgtta aaatatgact tcacggaaga gaggttaaaa

14221 ctctttgacc gttattttaa atattgggat cagacatacc acccaaattg tgttaactgt

14281 ttggatgaca gatgcattct gcattgtgca aactttaatg ttttattctc tacagtgttc

14341 ccacttacaa gttttggacc actagtgaga aaaatatttg ttgatggtgt tccatttgta

14401 gtttcaactg gataccactt cagagagcta ggtgttgtac ataatcagga tgtaaactta

14461 catagctcta gacttagttt taaggaatta cttgtgtatg ctgctgaccc tgctatgcac

14521 gctgcttctg gtaatctatt actagataaa cgcactacgt gcttttcagt agctgcactt

14581 actaacaatg ttgcttttca aactgtcaaa cccggtaatt ttaacaaaga cttctatgac

14641 tttgctgtgt ctaagggttt ctttaaggaa ggaagttctg ttgaattaaa acacttcttc

14701 tttgctcagg atggtaatgc tgctatcagc gattatgact actatcgtta taatctacca

14761 acaatgtgtg atatcagaca actactattt gtagttgaag ttgttgataa gtactttgat

14821 tgttacgatg gtggctgtat taatgctaac caagtcatcg tcaacaacct agacaaatca

14881 gctggttttc catttaataa atggggtaag gctagacttt attatgattc aatgagttat

14941 gaggatcaag atgcactttt cgcatataca aaacgtaatg tcatccctac tataactcaa

15001 atgaatctta agtatgccat tagtgcaaag aatagagctc gcaccgtagc tggtgtctct

15061 atctgtagta ctatgaccaa tagacagttt catcaaaaat tattgaaatc aatagccgcc

15121 actagaggag ctactgtagt aattggaaca agcaaattct atggtggttg gcacaacatg

15181 ttaaaaactg tttatagtga tgtagaaaac cctcacctta tgggttggga ttatcctaaa

15241 tgtgatagag ccatgcctaa catgcttaga attatggcct cacttgttct tgctcgcaaa

15301 catacaacgt gttgtagctt gtcacaccgt ttctatagat tagctaatga gtgtgctcaa

15361 gtattgagtg aaatggtcat gtgtggcggt tcactatatg ttaaaccagg tggaacctca

15421 tcaggagatg ccacaactgc ttatgctaat agtgttttta acatttgtca agctgtcacg

15481 gccaatgtta atgcactttt atctactgat ggtaacaaaa ttgccgataa gtatgtccgc

15541 aatttacaac acagacttta tgagtgtctc tatagaaata gagatgttga cacagacttt

15601 gtgaatgagt tttacgcata tttgcgtaaa catttctcaa tgatgatact ttctgacgat

15661 gctgttgtgt gtttcaatag cacttatgca tctcaaggtc tagtggctag cataaagaac

15721 tttaagtcag ttctttatta tcaaaacaat gtttttatgt ctgaagcaaa atgttggact

15781 gagactgacc ttactaaagg acctcatgaa ttttgctctc aacatacaat gctagttaaa

15841 cagggtgatg attatgtgta ccttccttac ccagatccat caagaatcct aggggccggc

15901 tgttttgtag atgatatcgt aaaaacagat ggtacactta tgattgaacg gttcgtgtct

15961 ttagctatag atgcttaccc acttactaaa catcctaatc aggagtatgc tgatgtcttt

16021 catttgtact tacaatacat aagaaagcta catgatgagt taacaggaca catgttagac

16081 atgtattctg ttatgcttac taatgataac acttcaaggt attgggaacc tgagttttat

16141 gaggctatgt acacaccgca tacagtctta caggctgttg gggcttgtgt tctttgcaat

16201 tcacagactt cattaagatg tggtgcttgc atacgtagac cattcttatg ttgtaaatgc

16261 tgttacgacc atgtcatatc aacatcacat aaattagtct tgtctgttaa tccgtatgtt

16321 tgcaatgctc caggttgtga tgtcacagat gtgactcaac tttacttagg aggtatgagc

16381 tattattgta aatcacataa accacccatt agttttccat tgtgtgctaa tggacaagtt

16441 tttggtttat ataaaaatac atgtgttggt agcgataatg ttactgactt taatgcaatt

16501 gcaacatgtg actggacaaa tgctggtgat tacattttag ctaacacctg tactgaaaga

16561 ctcaagcttt ttgcagcaga aacgctcaaa gctactgagg agacatttaa actgtcttat

16621 ggtattgcta ctgtacgtga agtgctgtct gacagagaat tacatctttc atgggaagtt

16681 ggtaaaccta gaccaccact taaccgaaat tatgtcttta ctggttatcg tgtaactaaa

16741 aacagtaaag tacaaatagg agagtacacc tttgaaaaag gtgactatgg tgatgctgtt

16801 gtttaccgag gtacaacaac ttacaaatta aatgttggtg attattttgt gctgacatca

16861 catacagtaa tgccattaag tgcacctaca ctagtgccac aagagcacta tgttagaatt

16921 actggcttat acccaacact caatatctca gatgagtttt ctagcaatgt tgcaaattat

16981 caaaaggttg gtatgcaaaa gtattctaca ctccagggac cacctggtac tggtaagagt

17041 cattttgcta ttggcctagc tctctactac ccttctgctc gcatagtgta tacagcttgc

17101 tctcatgccg ctgttgatgc actatgtgag aaggcattaa aatatttgcc tatagataaa

17161 tgtagtagaa ttatacctgc acgtgctcgt gtagagtgtt ttgataaatt caaagtgaat

17221 tcaacattag aacagtatgt cttttgtact gtaaatgcat tgcctgagac gacagcagat

17281 atagttgtct ttgatgaaat ttcaatggcc acaaattatg atttgagtgt tgtcaatgcc

17341 agattatgtg ctaagcacta tgtgtacatt ggcgaccctg ctcaattacc tgcaccacgc

17401 acattgctaa ctaagggcac actagaacca gaatatttca attcagtgtg tagacttatg

17461 aaaactatag gtccagacat gttcctcgga acttgtcggc gttgtcctgc tgaaattgtt

17521 gacactgtga gtgctttggt ttatgataat aagcttaaag cacataaaga caaatcagct

17581 caatgcttta aaatgtttta taagggtgtt atcacgcatg atgtttcatc tgcaattaac

17641 aggccacaaa taggcgtggt aagagaattc cttacacgta accctgcttg gagaaaagct

17701 gtctttattt caccttataa ttcacagaat gctgtagcct caaagatttt gggactacca

17761 actcaaactg ttgattcatc acagggctca gaatatgact atgtcatatt cactcaaacc

17821 actgaaacag ctcactcttg taatgtaaac agatttaatg ttgctattac cagagcaaaa

17881 gtaggcatac tttgcataat gtctgataga gacctttatg acaagttgca atttacaagt

17941 cttgaaattc cacgtaggaa tgtggcaact ttacaagctg aaaatgtaac aggactcttt

18001 aaagattgta gtaaggtaat cactgggtta catcctacac aggcacctac acacctcagt

18061 gttgacacta aattcaaaac tgaaggttta tgtgttgacg tacctggcat acctaaggac

18121 atgacctata gaagactcat ctctatgatg ggttttaaaa tgaattatca agttaatggt

18181 taccctaaca tgtttatcac ccgcgaagaa gctataagac atgtacgtgc atggattggc

18241 ttcgatgtcg aggggtgtca tgctactaga gaagctgttg gtaccaattt acctttacag

18301 ctaggttttt ctacaggtgt taacctagtt gctgtaccta caggttatgt tgatacacct

18361 aataatacag atttttccag agttagtgct aaaccaccgc ctggagatca atttaaacac

18421 ctcataccac ttatgtacaa aggacttcct tggaatgtag tgcgtataaa gattgtacaa

18481 atgttaagtg acacacttaa aaatctctct gacagagtcg tatttgtctt atgggcacat

18541 ggctttgagt tgacatctat gaagtatttt gtgaaaatag gacctgagcg cacctgttgt

18601 ctatgtgata gacgtgccac atgcttttcc actgcttcag acacttatgc ctgttggcat

18661 cattctattg gatttgatta cgtctataat ccgtttatga ttgatgttca acaatggggt

18721 tttacaggta acctacaaag caaccatgat ctgtattgtc aagtccatgg taatgcacat

18781 gtagctagtt gtgatgcaat catgactagg tgtctagctg tccacgagtg ctttgttaag

18841 cgtgttgact ggactattga atatcctata attggtgatg aactgaagat taatgcggct

18901 tgtagaaagg ttcaacacat ggttgttaaa gctgcattat tagcagacaa attcccagtt

18961 cttcacgaca ttggtaaccc taaagctatt aagtgtgtac ctcaagctga tgtagaatgg

19021 aagttctatg atgcacagcc ttgtagtgac aaagcttata aaatagaaga attattctat

19081 tcttatgcca cacattctga caaattcaca gatggtgtat gcctattttg gaattgcaat

19141 gtcgatagat atcctgctaa ttccattgtt tgtagatttg acactagagt gctatctaac

19201 cttaacttgc ctggttgtga tggtggcagt ttgtatgtaa ataaacatgc attccacaca

19261 ccagcttttg ataaaagtgc ttttgttaat ttaaaacaat taccattttt ctattactct

19321 gacagtccat gtgagtctca tggaaaacaa gtagtgtcag atatagatta tgtaccacta

19381 aagtctgcta cgtgtataac acgttgcaat ttaggtggtg ctgtctgtag acatcatgct

19441 aatgagtaca gattgtatct cgatgcttat aacatgatga tctcagctgg ctttagcttg

19501 tgggtttaca aacaatttga tacttataac ctctggaaca cttttacaag acttcagagt

19561 ttagaaaatg tggcttttaa tgttgtaaat aagggacact ttgatggaca acagggtgaa

19621 gtaccagttt ctatcattaa taacactgtt tacacaaaag ttgatggtgt tgatgtagaa

19681 ttgtttgaaa ataaaacaac attacctgtt aatgtagcat ttgagctttg ggctaagcgc

19741 aacattaaac cagtaccaga ggtgaaaata ctcaataatt tgggtgtgga cattgctgct

19801 aatactgtga tctgggacta caaaagagat gctccagcac atatatctac tattggtgtt

19861 tgttctatga ctgacatagc caagaaacca attgaaacga tttgtgcacc actcactgtc

19921 ttttttgatg gtagagttga tggtcaagta gacttattta gaaatgcccg taatggtgtt

19981 cttattacag agggtagtgt taaaggttta caaccatctg taggtcccaa acaagctagt

20041 cttaatggag tcacattaat tggagaagcc gtaaaaacac agttcaatta ttataagaaa

20101 gttgatggtg ttgtccaaca attacctgaa acttacttta ctcagagtag aaatttacaa

20161 gaatttaaac ccaggagtca aatggaaatt gatttcttag aattagctat ggatgaattc

20221 attgaacggt ataaattaga aggctatgcc ttcgaacata tcgtttatgg agattttagt

20281 catagtcagt taggtggttt acatctactg attggactag ctaaacgttt taaggaatca

20341 ccttttgaat tagaagattt tattcctatg gacagtacag ttaaaaacta tttcataaca

20401 gatgcgcaaa caggttcatc taagtgtgtg tgttctgtta ttgatttatt acttgatgat

20461 tttgttgaaa taataaaatc ccaagattta tctgtagttt ctaaggttgt caaagtgact

20521 attgactata cagaaatttc atttatgctt tggtgtaaag atggccatgt agaaacattt

20581 tacccaaaat tacaatctag tcaagcgtgg caa

[gap 886 bp] [Expand Ns](https://www.ncbi.nlm.nih.gov/nuccore/ON421981.1?expand-gaps=on)

21500 a tgtttgtttt tcttgtttta ttgccactag tctctagtca

21541 gtgtgttaat cttacaacca gaactcaatc atacactaat tctttcacac gtggtgttta

21601 ttaccctgac aaagttttca gatcctcagt tttacattca actcaggact tgttcttacc

21661 tttcttttcc aatgttactt ggttccatgc tatacatgtc tctgggacca atggtactaa

21721 gaggtttgat aaccctgtcc taccatttaa tgatggtgtt tattttgctt ccactgagaa

21781 gtctaacata ataagaggct ggatttttgg tactacttta gattcgaaga cccagtccct

21841 acttattgtt aataacgcta ctaatgttgt tattaaagtc tgtgaatttc aattttgtaa

21901 tgatccattt ttggatgttt attaccacaa aaacaacaaa agttggatgg aaagtgagtt

21961 cagagtttat tctagtgcga ataattgcac ttttgaatat gtctctcagc cttttcttat

22021 ggaccttgaa ggaaaacagg gtaatttcaa aaatcttagg gaatttgtgt ttaagaatat

22081 tgatggttat tttaaaatat attctaagca cacgcctatt aatttagggc gtgatctccc

22141 tcagggtttt tcggctttag aaccattggt agatttgcca ataggtatta acatcactag

22201 gtttcaaact ttacttgctt tacatagaag ttatttgact cctggtgatt cttcttcagg

22261 ttggacagct ggtgctgcag cttattatgt gggttatctt caacctagga cttttctatt

22321 aaaatataat gaaaatggaa ccattacaga tgctgtagac tgtgcacttg accctctctc

22381 agaaacaaag tgtacgttga aatccttcac tgtagaaaaa ggaatctatc aaacttctaa

22441 ctttagagtc caaccaacag aatctattgt tagatttcct aatattacaa acttgtgccc

22501 ttttgatgaa gtttttaacg ccaccagatt tgcatctgtt tatgcttgga acaggaagag

22561 aatcagcaac tgtgttgctg attattctgt cctatataat ttcgcaccat ttttcgcttt

22621 taagtgttat ggagtgtctc ctactaaatt aaatgatctc tgctttacta atgtctatgc

22681 agattcattt gtaattagag gtaatgaagt cagccaaatc gctccagggc aaactggaaa

22741 tattgctgat tataattata aattaccaga tgattttaca ggctgcgtta tagcttggaa

22801 ttctaacaag cttgattcta aggttggtgg taattataat tacctgtata gattgtttag

22861 gaagtctaat ctcaaacctt ttgagagaga tatttcaact gaaatctatc aggccggtaa

22921 caaaccttgt aatggtgttg caggttttaa ttgttacttt cctttacgat catatggttt

22981 ccgacccact tatggtgttg gtcaccaacc atacagagta gtagtacttt cttttgaact

23041 tctacatgca ccagcaactg tttgtggacc taaaaagtct actaatttgg ttaaaaacaa

23101 atgtgtcaat ttcaacttca atggtttaac aggcacaggt gttcttactg agtctaacaa

23161 aaagtttctg cctttccaac aatttggcag agacattgct gacactactg atgctgtccg

23221 tgatccacag acacttgaga ttcttgacat tacaccatgt tcttttggtg gtgtcagtgt

23281 tataacacca ggaacaaata cttctaacca ggttgctgtt ctttatcagg gtgttaactg

23341 cacagaagtc cctgttgcta ttcatgcaga tcaacttact cctacttggc gtgtttattc

23401 tacaggttct aatgtttttc aaacacgtgc aggctgttta ataggggctg aatatgtcaa

23461 caactcatat gagtgtgaca tacccattgg tgcaggtata tgcgctagtt atcagactca

23521 gactaagtct catcggcggg cacgtagtgt agctagtcaa tccatcattg cctacactat

23581 gtcacttggt gcagaaaatt cagttgctta ctctaataac tctattgcca tacccacaaa

23641 ttttactatt agtgttacca cagaaattct accagtgtct atgaccaaga catcagtaga

23701 ttgtacaatg tacatttgtg gtgattcaac tgaatgcagc aatcttttgt tgcaatatgg

23761 cagtttttgt acacaattaa aacgtgcttt aactggaata gctgttgaac aagacaaaaa

23821 cacccaagaa gtttttgcac aagtcaaaca aatttacaaa acaccaccaa ttaaatattt

23881 tggtggtttt aatttttcac aaatattacc agatccatca aaaccaagca agaggtcatt

23941 tattgaagat ctacttttca acaaagtgac acttgcagat gctggcttca tcaaacaata

24001 tggtgattgc cttggtgata ttgctgctag agacctcatt tgtgcacaaa agtttaacgg

24061 ccttactgtt ttgccacctt tgctcacaga tgaaatgatt gctcaataca cttctgcact

24121 gttagcgggt acaatcactt ctggttggac ctttggtgca ggtgctgcat tacaaatacc

24181 atttgctatg caaatggctt ataggtttaa tggtattgga gttacacaga atgttctcta

24241 tgagaaccaa aaattgattg ccaaccaatt taatagtgct attggcaaaa ttcaagactc

24301 actttcttcc acagcaagtg cacttggaaa acttcaagat gtggtcaacc ataatgcaca

24361 agctttaaac acgcttgtta aacaacttag ctccaaattt ggtgcaattt caagtgtttt

24421 aaatgatatc ctttcacgtc ttgacaaagt tgaggctgaa gtgcaaattg ataggttgat

24481 cacaggcaga cttcaaagtt tgcagacata tgtgactcaa caattaatta gagctgcaga

24541 aatcagagct tctgctaatc ttgctgctac taaaatgtca gagtgtgtac ttggacaatc

24601 aaaaagagtt gatttttgtg gaaagggcta tcatcttatg tccttccctc agtcagcacc

24661 tcatggtgta gtcttcttgc atgtgactta tgtccctgca caagaaaaga acttcacaac

24721 tgctcctgcc atttgtcatg atggaaaagc acactttcct cgtgaaggtg tctttgtttc

24781 aaatggcaca cactggtttg taacacaaag gaatttttat gaaccacaaa tcattactac

24841 agacaacaca tttgtgtctg gtaactgtga tgttgtaata ggaattgtca acaacacagt

24901 ttatgatcct ttgcaacctg aattagattc attcaaggag gagttagata aatattttaa

24961 gaatcataca tcaccagatg ttgatttagg tgacatctct ggcattaatg cttcagttgt

25021 aaacattcaa aaagaaattg accgcctcaa tgaggttgcc aagaatttaa atgaatctct

25081 catcgatctc caagaacttg gaaagtatga gcagtatata aaatggccat ggtacatttg

25141 gctaggtttt atagctggct tgattgccat agtaatggtg acaattatgc tttgctgtat

25201 gaccagttgc tgtagttgtc tcaagggctg ttgttcttgt ggatcctgct gcaaatttga

25261 tgaagacgac tctgagccag tgctcaaagg agtcaaatta cattacacat aaacgaactt

25321 atggatttgt ttatgagaat ctttacaatt ggaactgtaa ctttgaagca aggtgaaatc

25381 aaggatgcta ctccttcaga ttttgttcgc gctactgcaa cgataccgat acaagcctca

25441 ctccctttcg gatggcttat tgttggcgtt gcacttcttg ctgtttttca gagcgcttcc

25501 aaaatcataa ctctcaaaaa gagatggcaa ctagcactct ccaagggtgt tcactttgtt

25561 tgcaacttgc tgttgttgtt tgtaacagtt tactcacacc ttttgctcgt tgctgctggc

25621 cttgaagccc cttttctcca tctttatgct ttagtctact tcttgcagag tataaacttt

25681 gtaagaataa taatgaggct ttggctttgc tggaaatgcc gttccaaaaa cccattactt

25741 tatgatgcca actattttct ttgctggcat actaattgtt acgactattg tataccttac

25801 aatagtgtaa cttcttcaat tgtcattact tcaggtgatg gcacaacaag tcctatttct

25861 gaacatgact accagattgg tggttatact gaaaaatggg aatctggagt aaaagactgt

25921 gttgtattac acagttactt cacttcagac tattaccagc tgtactcaac tcaattgagt

25981 acagacattg gtgttgaaca tgttaccttc ttcatctaca ataaaattgt tgatgagcct

26041 gaagaacatg tccaaattca cacaatcgac ggttcatccg gagttgttaa tccagtaatg

26101 gaaccaattt atgatgaacc gacgacgact actagcgtgc ctttgtaagc acaagctgat

26161 gagtacgaac ttatgtactc attcgtttcg gaagagatag gtacgttaat agttaatagc

26221 gtacttcttt ttcttgcttt cgtggtattc ttgctagtta cactagccat ccttactgcg

26281 cttcgattgt gtgcgtactg ctgcaatatt gttaacgtga gtcttgtaaa accttctttt

26341 tacgtttact ctcgtgttaa aaatctgaat tcttctagag ttcctgatct tctggtctaa

26401 acgaactaaa tattatatta gtttttctgt ttggaacttt aattttagcc atggcagatt

26461 ccaacggtac tattaccgtt gaagagctta aaaagctcct tgaagaatgg aacctagtaa

26521 taggtttcct attccttaca tggatttgtc ttctacaatt tgcctatgcc aacaggaata

26581 ggtttttgta tataattaag ttaattttcc tctggctgtt atggccagta actttaactt

26641 gttttgtgct tgctgctgtt tacagaataa attggatcac cggtggaatt gctatcgcaa

26701 tggcttgtct tgtaggcttg atgtggctca gctacttcat tgcttctttc agactgtttg

26761 cgcgtacgcg ttccatgtgg tcatttaatc cagaaactaa cattcttctc aacgtgccac

26821 tccatggcac tattctgacc agaccgcttc tagaaagtga actcgtaatc ggagctgtga

26881 tccttcgtgg acatcttcgt attgctggac accatctagg acgctgtgac atcaaggacc

26941 tgcctaaaga aatcactgtt gctacatcac gaacgctttc ttattacaaa ttgggagctt

27001 cgcagcgtgt agcaggtgac tcaggttttg ctgcatacag tcgctacagg attggcaact

27061 ataaattaaa cacagaccat tccagtagca gtgacaatat tgctttgctt gtacagtaag

27121 tgacaacaga tgtttcatct cgttgacttt caggttacta tagcagagat attactaatt

27181 attatgcgga cttttaaagt ttccatttgg aatcttgatt acatcataaa cctcataatt

27241 aaaaatttat ctaagtcact aactgagaat aaatattctc aattagatga agagcaacca

27301 atggagattc tctaaacgaa catgaaaatt attcttttct tggcactgat aacactcgct

27361 acttgtgagc tttatcacta ccaagagtgt gttagaggta caacagtact tttaaaagaa

27421 ccttgctctt ctggaacata cgagggcaat tcaccatttc atcctctagc tgataacaaa

27481 tttgcactga cttgctttag cactcaattt gcttttgctt gtcctgacgg cgtaaaacac

27541 gtctatcagt tacgtgccag atcagtttca cctaaactgt tcatcagaca agaggaagtt

27601 caagaacttt actctccaat ttttcttatt gttgcggcaa tagtgtttat aacactttgc

27661 ttcacactca aaagaaagac agaatgattg aactttcatt aattgacttc tatttgtgct

27721 ttttagcctt tctgttattc cttgttttaa ttatgcttat tatcttttgg ttctcacttg

27781 aactgcaaga tcataatgaa acttgtcacg cctaaacgaa catgaaattt cttgttttct

27841 taggaatcat cacaactgta gctgcatttc accaagaatg tagtttacag tcatgtactc

27901 aacatcaacc atatgtagtt gatgacccgt gtcctattca cttctattct aaatggtata

27961 ttagagtagg agctagaaaa tcagcacctt taattgaatt gtgcgtggat gaggctggtt

28021 ctaaatcacc cattcagtac atcgatatcg gtaattatac agtttcctgt ttacctttta

28081 caattaattg ccaggaacct aaattgggta gtcttgtagt gcgttgttcg ttctatgaag

28141 actttttaga gtatcatgac gttcgtgttg ttttagattt catctaaacg aacaaactta

28201 aatgtctgat aatggacccc aaaatcagcg aaatgcactc cgcattacgt ttggtggacc

28261 ctcagattca actggcagta accagaatgg tggggcgcga tcaaaacaac gtcggcccca

28321 aggtttaccc aataatactg cgtcttggtt caccgctctc actcaacatg gcaaggaaga

28381 ccttaaattc cctcgaggac aaggcgttcc aattaacacc aatagcagtc cagatgacca

28441 aattggctac taccgaagag ctaccagacg aattcgtggt ggtgacggta aaatgaaaga

28501 tctcagtcca agatggtatt tctactacct aggaactggg ccagaagctg gacttcccta

28561 tggtgctaac aaagacggca tcatatgggt tgcaactgag ggagccttga atacaccaaa

28621 agatcacatt ggcacccgca atcctgctaa caatgctgca atcgtgctac aacttcctca

28681 aggaacaaca ttgccaaaag gcttctacgc agaagggagc agaggcggca gtcaagcctc

28741 ttctcgttcc tcatcacgta gtcgcaacag ttcaagaaat tcaactccag gcagcagtaa

28801 acgaacttct cctgctagaa tggctggcaa tggcggtgat gctgctcttg ctttgctgct

28861 gcttgacaga ttgaaccagc ttgagagcaa aatgtctggt aaaggccaac aacaacaagg

28921 ccaaactgtc actaagaaat ctgctgctga ggcttctaag aagcctcggc aaaaacgtac

28981 tgccactaaa gcatacaatg taacacaagc tttcggcaga cgtggtccag aacaaaccca

29041 aggaaatttt ggggaccagg aactaatcag acaaggaact gattacaaac attggccgca

29101 aattgcacaa tttgccccca gcgcttcagc gttcttcgga atgtcgcgca ttggcatgga

29161 agtcacacct tcgggaacgt ggttgaccta cacaggtgcc atcaaattgg atgacaaaga

29221 tccaaatttc aaagatcaag tcattttgct gaataagcat attgacgcat acaaaacatt

29281 cccaccaaca gagcctaaaa aggacaaaaa gaagaaggct gatgaaactc aagccttacc

29341 gcagagacag aagaaacagc aaactgtgac tcttcttcct gctgcagatt tggatgattt

29401 ctccaaacaa ttgcaacaat ccatgagccg tgctgactca actcaggcct aaactcatgc

29461 agaccacaca aggcagatgg gctatataaa cgttttcgct tttccgttta cgatatatag

29521 tctactcttg tgcagaatga attctcgtaa ctacatagca caagtagatg tagttaactt

29581 taatctcaca tagcaatctt taatcagtgt gtaacattag ggaggacttg aaagagccac

29641 tacattttca c

//