**Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/USA/CT-CDC-LC0566727/2022 ORF1ab polyprotein (ORF1ab) and ORF1a polyprotein (ORF1ab) genes, partial cds; surface glycoprotein (S) and ORF3a protein (ORF3a) genes, complete cds; envelope p...**

GenBank: ON151370.1

[FASTA](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?report=fasta) [Graphics](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?report=graph)

[Go to:](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1" \l "goto2216763641_0)

LOCUS ON151370 28514 bp RNA linear VRL 05-APR-2022

DEFINITION Severe acute respiratory syndrome coronavirus 2 isolate

SARS-CoV-2/human/USA/CT-CDC-LC0566727/2022 ORF1ab polyprotein

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

glycoprotein (S) and ORF3a protein (ORF3a) genes, complete cds;

envelope protein (E) and membrane glycoprotein (M) genes, partial

cds; ORF6 protein (ORF6), ORF7a protein (ORF7a), ORF7b (ORF7b), and

ORF8 protein (ORF8) genes, complete cds; and nucleocapsid

phosphoprotein (N) gene, partial cds.

ACCESSION ON151370

VERSION ON151370.1

DBLINK BioProject: [PRJNA716984](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA716984)

BioSample: [SAMN27304078](https://www.ncbi.nlm.nih.gov/biosample/SAMN27304078)

KEYWORDS purposeofsampling:baselinesurveillance.

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 28514)

AUTHORS Howard,D., Batra,D., Cook,P.W., Caravas,J., Rambo-Martin,B.,

Sammons,S., Unoarumhi,Y., Schmerer,M., Lacek,K.A., Kendall,T.,

Caban Figueroa,V., Morrison,S., Gulvick,C., Agarwal,M., Almasri,E.,

Boles,D., Burns,A., Charoensri,N., Cohen,O., Countryman,S.,

Cristobal,M.A., Croy,B., Dale,S., Deshmukh,H., Douglas,A.,

Drouillon,V., Eisenberg,M., Engler,H., Ghatti,R., Gupta,P.,

Hicks,S., Humphrey,J., Iyer,L., Pfefferle,L., Jain,M., Robinson,M.,

Kolli,M., Krueger,B., Kuphal,T., Letovsky,S., Levandoski,M.,

Lukasik,C., Meltzer,J., Norvell,B., Nye,M., Parker,S.,

Petropoulos,C., Pruitt,J., Ragan,S., Ryan,S., Sapeta,M.,

Schroth,J., Selvaraju,S.B., Stevovic,G., Suchanek,A., Throop,A.,

Tilson,L., Urban,T., Voshell,J., Wagner,K., Williams,J.,

Williamson,M., Zeng,Q., Zwiefelhofer,T., Paden,C.R. and

MacCannell,D.

TITLE CDC Sars CoV2 Sequencing Baseline Constellation

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 28514)

AUTHORS Howard,D., Batra,D., Cook,P.W., Caravas,J., Rambo-Martin,B.,

Sammons,S., Unoarumhi,Y., Schmerer,M., Lacek,K.A., Kendall,T.,

Caban Figueroa,V., Morrison,S., Gulvick,C., Sula,E., Agarwal,M.,

Almasri,E., Boles,D., Burns,A., Charoensri,N., Cohen,O.,

Countryman,S., Cristobal,M.A., Croy,B., Dale,S., Deshmukh,H.,

Douglas,A., Drouillon,V., Eisenberg,M., Engler,H., Ghatti,R.,

Gupta,P., Hicks,S., Humphrey,J., Iyer,L., Pfefferle,L., Jain,M.,

Robinson,M., Kolli,M., Krueger,B., Kuphal,T., Letovsky,S.,

Levandoski,M., Lukasik,C., Meltzer,J., Norvell,B., Nye,M.,

Parker,S., Petropoulos,C., Pruitt,J., Ragan,S., Ryan,S., Sapeta,M.,

Schroth,J., Selvaraju,S.B., Stevovic,G., Suchanek,A., Throop,A.,

Tilson,L., Urban,T., Voshell,J., Wagner,K., Williams,J.,

Williamson,M., Zeng,Q., Zwiefelhofer,T., Paden,C.R. and

MacCannell,D.

TITLE Direct Submission

JOURNAL Submitted (05-APR-2022) Respiratory Viruses Branch, Division of

Viral Diseases, Centers for Disease Control and Prevention, 1600

Clifton Rd, Atlanta, GA 30329, USA

COMMENT ##Assembly-Data-START##

Assembly Method :: CLC Genomics

Sequencing Technology :: PacBio Sequel II

##Assembly-Data-END##

FEATURES Location/Qualifiers

source 1..28514

/organism="Severe acute respiratory syndrome coronavirus

2"

/mol\_type="genomic RNA"

/isolate="SARS-CoV-2/human/USA/CT-CDC-LC0566727/2022"

/isolation\_source="Nasal Swabs"

/host="Homo sapiens"

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

/country="USA: Connecticut"

/collection\_date="2022-03-22"

[gene](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?from=1&to=20931) <1..20931

/gene="ORF1ab"

[CDS](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?location=1:12844,12844:20931) join(<1..12844,12844..20931)

/gene="ORF1ab"

/ribosomal\_slippage

/codon\_start=2

/product="ORF1ab polyprotein"

/protein\_id="[UOF55137.1](https://www.ncbi.nlm.nih.gov/protein/2216763642)"

/translation="YRKVLLRKNGNKGAGGHRYGADLKSFDLGDELGTDPYEDFQENW

NTKHSSGVTRELMRELNGGAYTRYVDNNFCGPDGYPLECIKDLLARAGKASCTLSEQL

DFIDTKRGVYCCREHEHEIAWYTERSEKSYELQTPFEIKLAKKFDTFNGECPNFVFPL

NSIIKTIQPRVEKKKLDGFMGRIRSVYPVASPNECNQMCLSTLMKCDHCGETSWQTGD

FVKATCEFCGTENLTKEGATTCGYLPQNAVVKIYCPACHNSEVGPEHSLAEYHNESGL

KTILRKGGRTIAFGGCVFSYVGCHNKCAYWVPRASANIGCNHTGVVGEGSEGLNDNLL

EILQKEKVNINIVGDFKLNEEIAIILASFSASTSAFVETVKGLDYKAFKQIVESCGNF

KVTKGKAKKGAWNIGEQKSILSPLYAFASEAARVVRSIFSRTLETAQNSVRVLQKAAI

TILDGISQYSLRLIDAMMFTSDLATNNLVVMAYITGGVVQLTSQWLTNIFGTVYEKLK

PVLDWLEEKFKEGVEFLRDGWEIVKFISTCACEIVGGQIVTCAKEIKESVQTFFKLVN

KFLALCADSIIIGGAKLKALNLGETFVTHSKGLYRKCVKSREETGLLMPLKAPKEIIF

LEGETLPTEVLTEEVVLKTGDLQPLEQPTSEAVEAPLVGTPVCINGLMLLEIKDTEKY

CALAPNMMVTNNTFTLKGGAPTKVTFGDDTVIEVQGYKSVNIIFELDERIDKVLNEKC

SAYTVELGTEVNEFACVVADAVIKTLQPVSELLTPLGIDLDEWSMATYYLFDESGEFK

LASHMYCSFYPPDEDEEEGDCEEEEFEPSTQYEYGTEDDYQGKPLEFGATSAALQPEE

EQEEDWLDDDSQQTVGQQDGSEDNQTTTIQTIVEVQPQLEMELTPVVQTIEVNSFSGY

LKLTDNVYIKNADIVEEAKKVKPTVVVNAANVYLKHGGGVAGALNKATNNAMQVESDD

YIATNGPLKVGGSCVLSGHNLAKHCLHVVGPNVNKGEDIQLLKSAYENFNQHEVLLAP

LLSAGIFGADPIHSLRVCVDTVRTNVYLAVFDKNLYDKLVSSFLEMKSEKQVEQKIAE

IPKEEVKPFITESKPSVEQRKQDDKKIKACVEEVTTTLEETKFLTENLLLYIDINGNL

HPDSATLVSDIDITFLKKDAPYIVGDVVQEGVLTAVVIPTKKASGTTEMLAKALRKVP

TDNYITTYPGQGLNGYTVEEAKTVLKKCKSAFYILPSIISNEKQEILGTVSWNLREML

AHAEETRKLMPVCVETKAIVSTIQRKYKGIKIQEGVVDYGARFYFYTSKTTVASLINT

LNDLNETLVTMPLGYVTHGLNLEEAARYMRSLKVPATVSVSSPDAVTAYNGYLTSSSK

TPEEHFIETISLAGSYKDWSYSGQSTQLGIEFLKRGDKSVYYTSNPTTFHLDGEVITF

DNLKTLLSLREVRTIKVFTTVDNINLHTQVVDMSMTYGQQFGPTYLDGADVTKIKPHN

SHEGKTFYVLPNDDTLRVEAFEYYHTTDPSFLGRYMSALNHTKKWKYPQVNGLTSIKW

ADNNCYLATALLTLQQIELKFNPPALQDAYYRARAGEAANFCALILAYCNKTVGELGD

VRETMSYLFQHANLDSCKRVLNVVCKTCGQQQTTLKGVEAVMYMGTLSYEQFKKGVQI

PCTCGKQATKYLVQQESPFVMMSAPPAQYELKHGTFTCASEYTGNYQCGHYKHITSKE

TLYCIDGALLTKSSEYKGPITDVFYKENSYTTTIKPVTYKLDGVVCTEIDPKLDNYYK

KDNSYFTEQPIDLVPNQPYPNASFDNFKFVCDNIKFADDLNQLTGYKKPASRELKVTF

FPDLNGDVVAIDYKHYTPSFKKGAKLLHKPIVWHVNNATNKATYKPNTWCIRCLWSTK

PVETSNSFDVLKSEDAQGMDNLACEDLKPVSEEVVENPTIQKDVLECNVKTTEVVGDI

ILKPANNSLKITEEVGHTDLMAAYVDNSSLTIKKPNELSRVLGLKTLATHGLAAVNSV

PWDTIANYAKPFLNKVVSTTTNIVTRCLNRVCTNYMPYFFTLLLQLCTFTRSTNSRIK

ASMPTTIAKNTVKSVGKFCLEASFNYLKSPNFSKLINIIIWFLLLSVCLGSLIYSTAA

LGVLMSNLGMPSYCTGYREGYLNSTNVTIATYCTGSIPCSVCLSGLDSLDTYPSLETI

QITISSFKWDLTAFGLVAEWFLAYILFTRFFYVLGLAAIMQLFFSYFAVHFISNSWLM

WLIINLVQMAPISAMVRMYIFFASFYYVWKSYVHVVDGCNSSTCMMCYKRNRATRVEC

TTIVNGVRRSFYVYANGGKGFCKLHNWNCVNCDTFCAGSTFISDEVARDLSLQFKRPI

NPTDQSSYIVDSVTVKNGSIHLYFDKAGQKTYERHSLSHFVNLDNLRANNTKGSLPIN

VIVFDGKSKCEESSAKSASVYYSQLMCQPILLLDQALVSDVGDSAEVAVKMFDAYVNT

FSSTFNVPMEKLKTLVATAEAELAKNVSLDNVLSTFISAARQGFVDSDVETKDVVECL

KLSHQSDIEVTGDSCNNYMLTYNKVENMTPRDLGACIDCSARHINAQVAKSHNIALIW

NVKDFMSLSEQLRKQIRSAAKKNNLPFKLTCATTRQVVNVVTTKIALKGGKIVNNWLK

QLIKVTLVFLFVAAIFYLITPVHVMSKHTDFSSEIIGYKAIDGGVTRDIASTDTCFAN

KHADFDTWFSQRGGSYTNDKACPLIAAVITREVGFVVPGLPGTILRTTNGDFLHFLPR

VFSAVGNICYTPSKLIEYTDFATSACVLAAECTIFKDASGKPVPYCYDTNVLEGSVAY

ESLRPDTRYVLMDGSIIQFPNTYLEGSVRVVTTFDSEYCRHGTCERSEAGVCVSTSGR

WVLNNDYYRSLPGVFCGVDAVNLFTNMFTPLIQPIGALDISASIVAGGIVAIVVTCLA

YYFMRFRRAFGEYSHVVAFNTLLFLMSFIVLCLTPVYSFLPGVYSVIYLYLTFYLTND

VSFLAHIQWMVMFTPLVPFWITIAYIICISTKHFYWFFSNYLKRRVVFNGVSFSTFEE

AALCTFLLNKEMYLKLRSDVLLPFTQYNRYLALYNKYKYFSGAMDTTSYREAACCHLA

KALNDFSNSGSDVLYQPPQISITSAVLQSGFRKMAFPSGKVEGCMVQVTCGTTTLNGL

WLDDVVYCPRHVICTSEDMLNPNYEDLLIRKSNHNFLVQAGNVQLRVIGHSMQNCVLK

LKVDTANPKTPKYKFVRIQPGQTFSVLACYNGSPSGVYQCAMRHNFTIKGSFLNGSCG

SVGFNIDYDCVSFCYMHHMELPTGVHAGTDLEGNFYGPFVDRQTAQAAGTDTTITVNV

LAWLYAAVINGDRWFLNRFTTTLNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAVL

DMCASLKELLQNGMNGRTILGSALLEDEFTPFDVVRQCSGVTFQSAVKRTIKGTHHWL

LLTILTSLLVLVQSTQWSLFFFLYENAFLPFAMGIIAMSAFAMMFVKHKHAFLCLFLL

PSLATVAYFNMVYMPASWVMRIMTWLDMVDTSLKLKDCVMYASAVVLLILMTARTVYD

DGARRVWTLMNVLTLVYKVYYGNALDQAISMWALIISVTSNYSGVVTTVMFLARGIVF

MCVEYCPIFFITGNTLQCIMLVYCFLGYFCTCYFGLFCLLNRYFRLTLGVYDYLVSTQ

EFRYMNSQGLLPPKNSIDAFKLNIKLLGVGGKPCIKVATVQSKMSDVKCTSVVLLSVL

QQLRVESSSKLWAQCVQLHNDILLAKDTTEAFEKMVSLLSVLLSMQGAVDINKLCEEM

LDNRATLQAIASEFSSLPSYAAFATAQEAYEQAVANGDSEVVLKKLKKSLNVAKSEFD

RDAAMQRKLEKMADQAMTQMYKQARSEDKRAKVTSAMQTMLFTMLRKLDNDALNNIIN

NARDGCVPLNIIPLTTAAKLMVVIPDYNTYKNTCDGTTFTYASALWEIQQVVDADSKI

VQLSEISMDNSPNLAWPLIVTALRANSAVKLQNNELSPVALRQMSCAAGTTQTACTDD

NALAYYNTTKGGRFVLALLSDLQDLKWARFPKSDGTGTIYTELEPPCRFVTDTPKGPK

VKYLYFIKGLNNLNRGMVLGSLAATVRLQAGNATEVPANSTVLSFCAFAVDAAKAYKD

YLASGGQPITNCVKMLCTHTGTGQAITVTPEANMDQESFGGASCCLYCRCHIDHPNPK

GFCDLKGKYVQIPTTCANDPVGFTLKNTVCTVCGMWKGYGCSCDQLREPMLQSADAQS

FLNRVCGVSAARLTPCGTGTSTDVVYRAFDIYNDKVAGFAKFLKTNCCRFQEKDEDDN

LIDSYFVVKRHTFSNYQHEETIYNLLKDCPAVAKHDFFKFRIDGDMVPHISRQRLTKY

TMADLVYALRHFDEGNCDTLKEILVTYNCCDDDYFNKKDWYDFVENPDILRVYANLGE

RVRQALLKTVQFCDAMRNAGIVGVLTLDNQDLNGNWYDFGDFIQTTPGSGVPVVDSYY

SLLMPILTLTRALTAESHVDTDLTKPYIKWDLLKYDFTEERLKLFDRYFKYWDQTYHP

NCVNCLDDRCILHCANFNVLFSTVFPLTSFGPLVRKIFVDGVPFVVSTGYHFRELGVV

HNQDVNLHSSRLSFKELLVYAADPAMHAASGNLLLDKRTTCFSVAALTNNVAFQTVKP

GNFNKDFYDFAVSKGFFKEGSSVELKHFFFAQDGNAAISDYDYYRYNLPTMCDIRQLL

FVVEVVDKYFDCYDGGCINANQVIVNNLDKSAGFPFNKWGKARLYYDSMSYEDQDALF

AYTKRNVIPTITQMNLKYAISAKNRARTVAGVSICSTMTNRQFHQKLLKSIAATRGAT

VVIGTSKFYGGWHNMLKTVYSDVENPHLMGWDYPKCDRAMPNMLRIMASLVLARKHTT

CCSLSHRFYRLANECAQVLSEMVMCGGSLYVKPGGTSSGDATTAYANSVFNICQAVTA

NVNALLSTDGNKIADKYVRNLQHRLYECLYRNRDVDTDFVNEFYAYLRKHFSMMILSD

DAVVCFNSTYASQGLVASIKNFKSVLYYQNNVFMSEAKCWTETDLTKGPHEFCSQHTM

LVKQGDDYVYLPYPDPSRILGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKHPNQE

YADVFHLYLQYIRKLHDELTGHMLDMYSVMLTNDNTSRYWEPEFYEAMYTPHTVLQAV

GACVLCNSQTSLRCGACIRRPFLCCKCCYDHVISTSHKLVLSVNPYVCNAPGCDVTDV

TQLYLGGMSYYCKSHKPPISFPLCANGQVFGLYKNTCVGSDNVTDFNAIATCDWTNAG

DYILANTCTERLKLFAAETLKATEETFKLSYGIATVREVLSDRELHLSWEVGKPRPPL

NRNYVFTGYRVTKNSKVQIGEYTFEKGDYGDAVVYRGTTTYKLNVGDYFVLTSHTVMP

LSAPTLVPQEHYVRITGLYPTLNISDEFSSNVANYQKVGMQKYSTLQGPPGTGKSHFA

IGLALYYPSARIVYTACSHAAVDALCEKALKYLPIDKCSRIIPARARVECFDKFKVNS

TLEQYVFCTVNALPETTADIVVFDEISMATNYDLSVVNARLCAKHYVYIGDPAQLPAP

RTLLTKGTLEPEYFNSVCRLMKTIGPDMFLGTCRRCPAEIVDTVSALVYDNKLKAHKD

KSAQCFKMFYKGVITHDVSSAINRPQIGVVREFLTRNPAWRKAVFISPYNSQNAVASK

ILGLPTQTVDSSQGSEYDYVIFTQTTETAHSCNVNRFNVAITRAKVGILCIMSDRDLY

DKLQFTSLEIPRRNVATLQAENVTGLFKDCSKVITGLHPTQAPTHLSVDTKFKTEGLC

VDVPGIPKDMTYRRLISMMGFKMNYQVNGYPNMFITREEAIRHVRAWIGFDVEGCHAT

REAVGTNLPLQLGFSTGVNLVAVPTGYVDTPNNTDFSRVSAKPPPGDQFKHLIPLMYK

GLPWNVVRIKIVQMLSDTLKNLSDRVVFVLWAHGFELTSMKYFVKIGPERTCCLCDRR

ATCFSTASDTYACWHHSIGFDYVYNPFMIDVQQWGFTGNLQSNHDLYCQVHGNAHVAS

CDAIMTRCLAVHECFVKRVDWTIEYPIIGDELKINAACRKVQHMVVKAALLADKFPVL

HDIGNPKAIKCVPQADVEWKFYDAQPCSDKAYKIEELFYSYATHSDKFTDGVCLFWNC

NVDRYPANSIVCRFDTRVLSNLNLPGCDGGSLYVNKHAFHTPAFDKSAFVNLKQLPFF

YYSDSPCESHGKQVVSDIDYVPLKSATCITRCNLGGAVCRHHANEYRLYLDAYNMMIS

AGFSLWVYKQFDTYNLWNTFTRLQSLENVAFNVVNKGHFDGQQGEVPVSIINNTVYTK

VDGVDVELFENKTTLPVNVAFELWAKRNIKPVPEVKILNNLGVDIAANTVIWDYKRDA

PAHISTIGVCSMTDIAKKPIETICAPLTVFFDGRVDGQVDLFRNARNGVLITEGSVKG

LQPSVGPKQASLNGVTLIGEAVKTQFNYYKKVDGVVQQLPETYFTQSRNLQEFKPRSQ

MEIDFLELAMDEFIERYKLEGYAFEHIVYGDFSHSQLGGLHLLIGLAKRFKESPFELE

DFIPMDSTVKNYFITDAQTGSSKCVCSVIDLLLDDFVEIIKSQDLSVVSKVVKVTIDY

TEISFMLWCKDGHVETFYPKLQSSQAWQPGVAMPNLYKMQRMLLEKCDLQNYGDSATL

PKGIMMNVAKYTQLCQYLNTLTLAVPYNMRVIHFGAGSDKGVAPGTAVLRQWLPTGTL

LVDSDLNDFVSDADSTLIGDCATVHTANKWDLIISDMYDPKTKNVTKENDSKEGFFTY

ICGFIQQKLALGGSVAIKITEHSWNADLYKLMGHFAWWTAFVTNVNASSSEAFLIGCN

YLGKPREQIDGYVMHANYIFWRNTNPIQLSSYSLFDMSKFPLKLRGTAVMSLKEGQIN

DMILSLLSKGRLIIRENNRVVISSDVLVNN"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55137.1?from=1&to=63) <1..190

/gene="ORF1ab"

/product="leader protein"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55137.1?from=64&to=701) 191..2104

/gene="ORF1ab"

/product="nsp2"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55137.1?from=702&to=2646) 2105..7939

/gene="ORF1ab"

/product="nsp3"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55137.1?from=2647&to=3146) 7940..9439

/gene="ORF1ab"

/product="nsp4"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55137.1?from=3147&to=3452) 9440..10357

/gene="ORF1ab"

/product="3C-like proteinase"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55137.1?from=3453&to=3739) 10358..11218

/gene="ORF1ab"

/product="nsp6"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55137.1?from=3740&to=3822) 11219..11467

/gene="ORF1ab"

/product="nsp7"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55137.1?from=3823&to=4020) 11468..12061

/gene="ORF1ab"

/product="nsp8"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55137.1?from=4021&to=4133) 12062..12400

/gene="ORF1ab"

/product="nsp9"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55137.1?from=4134&to=4272) 12401..12817

/gene="ORF1ab"

/product="nsp10"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55137.1?from=4273&to=5204) join(12818..12844,12844..15612)

/gene="ORF1ab"

/product="RNA-dependent RNA polymerase"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55137.1?from=5205&to=5805) 15613..17415

/gene="ORF1ab"

/product="helicase"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55137.1?from=5806&to=6332) 17416..18996

/gene="ORF1ab"

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

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55137.1?from=6333&to=6678) 18997..20034

/gene="ORF1ab"

/product="endoRNAse"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55137.1?from=6679&to=6976) 20035..20928

/gene="ORF1ab"

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

[CDS](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?from=1&to=12859) <1..12859

/gene="ORF1ab"

/codon\_start=2

/product="ORF1a polyprotein"

/protein\_id="[UOF55138.1](https://www.ncbi.nlm.nih.gov/protein/2216763643)"

/translation="YRKVLLRKNGNKGAGGHRYGADLKSFDLGDELGTDPYEDFQENW

NTKHSSGVTRELMRELNGGAYTRYVDNNFCGPDGYPLECIKDLLARAGKASCTLSEQL

DFIDTKRGVYCCREHEHEIAWYTERSEKSYELQTPFEIKLAKKFDTFNGECPNFVFPL

NSIIKTIQPRVEKKKLDGFMGRIRSVYPVASPNECNQMCLSTLMKCDHCGETSWQTGD

FVKATCEFCGTENLTKEGATTCGYLPQNAVVKIYCPACHNSEVGPEHSLAEYHNESGL

KTILRKGGRTIAFGGCVFSYVGCHNKCAYWVPRASANIGCNHTGVVGEGSEGLNDNLL

EILQKEKVNINIVGDFKLNEEIAIILASFSASTSAFVETVKGLDYKAFKQIVESCGNF

KVTKGKAKKGAWNIGEQKSILSPLYAFASEAARVVRSIFSRTLETAQNSVRVLQKAAI

TILDGISQYSLRLIDAMMFTSDLATNNLVVMAYITGGVVQLTSQWLTNIFGTVYEKLK

PVLDWLEEKFKEGVEFLRDGWEIVKFISTCACEIVGGQIVTCAKEIKESVQTFFKLVN

KFLALCADSIIIGGAKLKALNLGETFVTHSKGLYRKCVKSREETGLLMPLKAPKEIIF

LEGETLPTEVLTEEVVLKTGDLQPLEQPTSEAVEAPLVGTPVCINGLMLLEIKDTEKY

CALAPNMMVTNNTFTLKGGAPTKVTFGDDTVIEVQGYKSVNIIFELDERIDKVLNEKC

SAYTVELGTEVNEFACVVADAVIKTLQPVSELLTPLGIDLDEWSMATYYLFDESGEFK

LASHMYCSFYPPDEDEEEGDCEEEEFEPSTQYEYGTEDDYQGKPLEFGATSAALQPEE

EQEEDWLDDDSQQTVGQQDGSEDNQTTTIQTIVEVQPQLEMELTPVVQTIEVNSFSGY

LKLTDNVYIKNADIVEEAKKVKPTVVVNAANVYLKHGGGVAGALNKATNNAMQVESDD

YIATNGPLKVGGSCVLSGHNLAKHCLHVVGPNVNKGEDIQLLKSAYENFNQHEVLLAP

LLSAGIFGADPIHSLRVCVDTVRTNVYLAVFDKNLYDKLVSSFLEMKSEKQVEQKIAE

IPKEEVKPFITESKPSVEQRKQDDKKIKACVEEVTTTLEETKFLTENLLLYIDINGNL

HPDSATLVSDIDITFLKKDAPYIVGDVVQEGVLTAVVIPTKKASGTTEMLAKALRKVP

TDNYITTYPGQGLNGYTVEEAKTVLKKCKSAFYILPSIISNEKQEILGTVSWNLREML

AHAEETRKLMPVCVETKAIVSTIQRKYKGIKIQEGVVDYGARFYFYTSKTTVASLINT

LNDLNETLVTMPLGYVTHGLNLEEAARYMRSLKVPATVSVSSPDAVTAYNGYLTSSSK

TPEEHFIETISLAGSYKDWSYSGQSTQLGIEFLKRGDKSVYYTSNPTTFHLDGEVITF

DNLKTLLSLREVRTIKVFTTVDNINLHTQVVDMSMTYGQQFGPTYLDGADVTKIKPHN

SHEGKTFYVLPNDDTLRVEAFEYYHTTDPSFLGRYMSALNHTKKWKYPQVNGLTSIKW

ADNNCYLATALLTLQQIELKFNPPALQDAYYRARAGEAANFCALILAYCNKTVGELGD

VRETMSYLFQHANLDSCKRVLNVVCKTCGQQQTTLKGVEAVMYMGTLSYEQFKKGVQI

PCTCGKQATKYLVQQESPFVMMSAPPAQYELKHGTFTCASEYTGNYQCGHYKHITSKE

TLYCIDGALLTKSSEYKGPITDVFYKENSYTTTIKPVTYKLDGVVCTEIDPKLDNYYK

KDNSYFTEQPIDLVPNQPYPNASFDNFKFVCDNIKFADDLNQLTGYKKPASRELKVTF

FPDLNGDVVAIDYKHYTPSFKKGAKLLHKPIVWHVNNATNKATYKPNTWCIRCLWSTK

PVETSNSFDVLKSEDAQGMDNLACEDLKPVSEEVVENPTIQKDVLECNVKTTEVVGDI

ILKPANNSLKITEEVGHTDLMAAYVDNSSLTIKKPNELSRVLGLKTLATHGLAAVNSV

PWDTIANYAKPFLNKVVSTTTNIVTRCLNRVCTNYMPYFFTLLLQLCTFTRSTNSRIK

ASMPTTIAKNTVKSVGKFCLEASFNYLKSPNFSKLINIIIWFLLLSVCLGSLIYSTAA

LGVLMSNLGMPSYCTGYREGYLNSTNVTIATYCTGSIPCSVCLSGLDSLDTYPSLETI

QITISSFKWDLTAFGLVAEWFLAYILFTRFFYVLGLAAIMQLFFSYFAVHFISNSWLM

WLIINLVQMAPISAMVRMYIFFASFYYVWKSYVHVVDGCNSSTCMMCYKRNRATRVEC

TTIVNGVRRSFYVYANGGKGFCKLHNWNCVNCDTFCAGSTFISDEVARDLSLQFKRPI

NPTDQSSYIVDSVTVKNGSIHLYFDKAGQKTYERHSLSHFVNLDNLRANNTKGSLPIN

VIVFDGKSKCEESSAKSASVYYSQLMCQPILLLDQALVSDVGDSAEVAVKMFDAYVNT

FSSTFNVPMEKLKTLVATAEAELAKNVSLDNVLSTFISAARQGFVDSDVETKDVVECL

KLSHQSDIEVTGDSCNNYMLTYNKVENMTPRDLGACIDCSARHINAQVAKSHNIALIW

NVKDFMSLSEQLRKQIRSAAKKNNLPFKLTCATTRQVVNVVTTKIALKGGKIVNNWLK

QLIKVTLVFLFVAAIFYLITPVHVMSKHTDFSSEIIGYKAIDGGVTRDIASTDTCFAN

KHADFDTWFSQRGGSYTNDKACPLIAAVITREVGFVVPGLPGTILRTTNGDFLHFLPR

VFSAVGNICYTPSKLIEYTDFATSACVLAAECTIFKDASGKPVPYCYDTNVLEGSVAY

ESLRPDTRYVLMDGSIIQFPNTYLEGSVRVVTTFDSEYCRHGTCERSEAGVCVSTSGR

WVLNNDYYRSLPGVFCGVDAVNLFTNMFTPLIQPIGALDISASIVAGGIVAIVVTCLA

YYFMRFRRAFGEYSHVVAFNTLLFLMSFIVLCLTPVYSFLPGVYSVIYLYLTFYLTND

VSFLAHIQWMVMFTPLVPFWITIAYIICISTKHFYWFFSNYLKRRVVFNGVSFSTFEE

AALCTFLLNKEMYLKLRSDVLLPFTQYNRYLALYNKYKYFSGAMDTTSYREAACCHLA

KALNDFSNSGSDVLYQPPQISITSAVLQSGFRKMAFPSGKVEGCMVQVTCGTTTLNGL

WLDDVVYCPRHVICTSEDMLNPNYEDLLIRKSNHNFLVQAGNVQLRVIGHSMQNCVLK

LKVDTANPKTPKYKFVRIQPGQTFSVLACYNGSPSGVYQCAMRHNFTIKGSFLNGSCG

SVGFNIDYDCVSFCYMHHMELPTGVHAGTDLEGNFYGPFVDRQTAQAAGTDTTITVNV

LAWLYAAVINGDRWFLNRFTTTLNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAVL

DMCASLKELLQNGMNGRTILGSALLEDEFTPFDVVRQCSGVTFQSAVKRTIKGTHHWL

LLTILTSLLVLVQSTQWSLFFFLYENAFLPFAMGIIAMSAFAMMFVKHKHAFLCLFLL

PSLATVAYFNMVYMPASWVMRIMTWLDMVDTSLKLKDCVMYASAVVLLILMTARTVYD

DGARRVWTLMNVLTLVYKVYYGNALDQAISMWALIISVTSNYSGVVTTVMFLARGIVF

MCVEYCPIFFITGNTLQCIMLVYCFLGYFCTCYFGLFCLLNRYFRLTLGVYDYLVSTQ

EFRYMNSQGLLPPKNSIDAFKLNIKLLGVGGKPCIKVATVQSKMSDVKCTSVVLLSVL

QQLRVESSSKLWAQCVQLHNDILLAKDTTEAFEKMVSLLSVLLSMQGAVDINKLCEEM

LDNRATLQAIASEFSSLPSYAAFATAQEAYEQAVANGDSEVVLKKLKKSLNVAKSEFD

RDAAMQRKLEKMADQAMTQMYKQARSEDKRAKVTSAMQTMLFTMLRKLDNDALNNIIN

NARDGCVPLNIIPLTTAAKLMVVIPDYNTYKNTCDGTTFTYASALWEIQQVVDADSKI

VQLSEISMDNSPNLAWPLIVTALRANSAVKLQNNELSPVALRQMSCAAGTTQTACTDD

NALAYYNTTKGGRFVLALLSDLQDLKWARFPKSDGTGTIYTELEPPCRFVTDTPKGPK

VKYLYFIKGLNNLNRGMVLGSLAATVRLQAGNATEVPANSTVLSFCAFAVDAAKAYKD

YLASGGQPITNCVKMLCTHTGTGQAITVTPEANMDQESFGGASCCLYCRCHIDHPNPK

GFCDLKGKYVQIPTTCANDPVGFTLKNTVCTVCGMWKGYGCSCDQLREPMLQSADAQS

FLNGFAV"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55138.1?from=1&to=63) <1..190

/gene="ORF1ab"

/product="leader protein"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55138.1?from=64&to=701) 191..2104

/gene="ORF1ab"

/product="nsp2"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55138.1?from=702&to=2646) 2105..7939

/gene="ORF1ab"

/product="nsp3"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55138.1?from=2647&to=3146) 7940..9439

/gene="ORF1ab"

/product="nsp4"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55138.1?from=3147&to=3452) 9440..10357

/gene="ORF1ab"

/product="3C-like proteinase"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55138.1?from=3453&to=3739) 10358..11218

/gene="ORF1ab"

/product="nsp6"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55138.1?from=3740&to=3822) 11219..11467

/gene="ORF1ab"

/product="nsp7"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55138.1?from=3823&to=4020) 11468..12061

/gene="ORF1ab"

/product="nsp8"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55138.1?from=4021&to=4133) 12062..12400

/gene="ORF1ab"

/product="nsp9"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55138.1?from=4134&to=4272) 12401..12817

/gene="ORF1ab"

/product="nsp10"

[mat\_peptide](https://www.ncbi.nlm.nih.gov/protein/UOF55138.1?from=4273&to=4285) 12818..12856

/gene="ORF1ab"

/product="nsp11"

[stem\_loop](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?from=12852&to=12879) 12852..12879

/gene="ORF1ab"

/note="Coronavirus frameshifting stimulation element

stem-loop 1"

[stem\_loop](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?from=12864&to=12918) 12864..12918

/gene="ORF1ab"

/note="Coronavirus frameshifting stimulation element

stem-loop 2"

[gene](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?from=20939&to=24751) 20939..24751

/gene="S"

[CDS](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?from=20939&to=24751) 20939..24751

/gene="S"

/codon\_start=1

/product="surface glycoprotein"

/protein\_id="[UOF55139.1](https://www.ncbi.nlm.nih.gov/protein/2216763644)"

/translation="MFVFLVLLPLVSSQCVNLITRTQSYTNSFTRGVYYPDKVFRSSV

LHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWI

FGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLDVYYHKNNKSWMESEFRVYSSA

NNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLGRDLPQGFS

ALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKY

NENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCP

FDEVFNATRFASVYAWNRKRISNCVADYSVLYNFAPFFAFKCYGVSPTKLNDLCFTNV

YADSFVIRGNEVSQIAPGQTGNIADYNYKLPDDFTGCVIAWNSNKLDSKVGGNYNYLY

RLFRKSNLKPFERDISTEIYQAGNKPCNGVAGFNCYFPLRSYGFRPTYGVGHQPYRVV

VLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDI

ADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHAD

QLTPTWRVYSTGSNVFQTRAGCLIGAEYVNNSYECDIPIGAGICASYQTQTKSHRRAR

SVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYIC

GDSTECSNLLLQYGSFCTQLKRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKYFGGFN

FSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLT

VLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLY

ENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNHNAQALNTLVKQLSSKFGAISS

VLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECV

LGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPR

EGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFK

EELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYE

QYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVL

KGVKLHYT"

[gene](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?from=24760&to=25587) 24760..25587

/gene="ORF3a"

[CDS](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?from=24760&to=25587) 24760..25587

/gene="ORF3a"

/codon\_start=1

/product="ORF3a protein"

/protein\_id="[UOF55140.1](https://www.ncbi.nlm.nih.gov/protein/2216763645)"

/translation="MDLFMRIFTIGTVTLKQGEIKDATPSDFVRATATIPIQASLPFG

WLIVGVALLAVFQSASKIITLKKRWQLALSKGVHFVCNLLLLFVTVYSHLLLVAAGLE

APFLYLYALVYFLQSINFVRIIMRLWLCWKCRSKNPLLYDANYFLCWHTNCYDYCIPY

NSVTSSIVITSGDGTTSPISEHDYQIGGYTEKWESGVKDCVVLHSYFTSDYYQLYSTQ

LSTDIGVEHVTFFIYNKIVDEPEEHVQIHTIDGSSGVVNPVMEPIYDEPTTTTSVPL"

[gene](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?from=25612&to=25821) 25612..>25821

/gene="E"

[CDS](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?from=25612&to=25821) 25612..>25821

/gene="E"

/codon\_start=1

/product="envelope protein"

/protein\_id="[UOF55141.1](https://www.ncbi.nlm.nih.gov/protein/2216763646)"

/translation="MYSFVSEEIGTLIVNSVLLFLAFVVFLLVTLAILTALRLCXXXX

XXXXXXXXXXXFYVYSRVKNLNSSRV"

gap 25822..25928

/estimated\_length=107

[gene](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?from=25929&to=26558) <25929..26558

/gene="M"

[CDS](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?from=25929&to=26558) <25929..26558

/gene="M"

/codon\_start=1

/product="membrane glycoprotein"

/protein\_id="[UOF55142.1](https://www.ncbi.nlm.nih.gov/protein/2216763647)"

/translation="KKLLEEWNLVIGFLFLTWICLLQFXXXXXXXXXXXXXXXXXXXX

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXFIASFRLFARTRSMWSFNPE

TNILLNVPLHGTILTRPLLESELVIGAVILRGHLRIAGHHLGRCDIKDLPKEITVATS

RTLSYYKLGASQRVAGDSGFAAYSRYRIGNYKLNTDHSSSSDNIALLVQ"

gap 26002..26174

/estimated\_length=173

[gene](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?from=26569&to=26754) 26569..26754

/gene="ORF6"

[CDS](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?from=26569&to=26754) 26569..26754

/gene="ORF6"

/codon\_start=1

/product="ORF6 protein"

/protein\_id="[UOF55143.1](https://www.ncbi.nlm.nih.gov/protein/2216763648)"

/translation="MFHLVDFQVTIAEILLIIMRTFKVSIWNLDYIINLIIKNLSKSL

TENKYSQLDEEQPMEIL"

[gene](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?from=26761&to=27126) 26761..27126

/gene="ORF7a"

[CDS](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?from=26761&to=27126) 26761..27126

/gene="ORF7a"

/codon\_start=1

/product="ORF7a protein"

/protein\_id="[UOF55144.1](https://www.ncbi.nlm.nih.gov/protein/2216763649)"

/translation="MKIILFLALITLATCELYHYQECVRGTTVLLKEPCSSGTYEGNS

PFHPLADNKFALTCFSTQFAFACPDGVKHVYQLRARSVSPKLFIRQEEVQELYSPIFL

IVAAIVFITLCFTLKRKTE"

[gene](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?from=27123&to=27254) 27123..27254

/gene="ORF7b"

[CDS](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?from=27123&to=27254) 27123..27254

/gene="ORF7b"

/codon\_start=1

/product="ORF7b"

/protein\_id="[UOF55145.1](https://www.ncbi.nlm.nih.gov/protein/2216763650)"

/translation="MIELSLIDFYLCFLAFLLFLVLIMLIIFWFSLELQDHNETCHA"

[gene](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?from=27261&to=27626) 27261..27626

/gene="ORF8"

[CDS](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?from=27261&to=27626) 27261..27626

/gene="ORF8"

/codon\_start=1

/product="ORF8 protein"

/protein\_id="[UOF55146.1](https://www.ncbi.nlm.nih.gov/protein/2216763651)"

/translation="MKFLVFLGIITTVAAFHQECSLQSCTQHQPYVVDDPCPIHFYSK

WYIRVGARKSAPLIELCVDEAGSKSPIQYIDIGNYTVSCLPFTINCQEPKLGSLVVRC

SFYEDFLEYHDVRVVLDFI"

[gene](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?from=27641&to=28514) 27641..>28514

/gene="N"

[CDS](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?from=27641&to=28514) 27641..>28514

/gene="N"

/codon\_start=1

/product="nucleocapsid phosphoprotein"

/protein\_id="[UOF55147.1](https://www.ncbi.nlm.nih.gov/protein/2216763652)"

/translation="MSDNGPQNQRNALRITFGGPSDSTGSNQNGGARSKQRRPQGLPN

NTASWFTALTQHGKEDLKFPRGQGVPINTNSSPDDQIGYYRRATRRIRGGDGKMKXXX

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXLPQ

GTTLPKGFYAEGSRGGSQASSRSSSRSRNSSRNSTPGSSKRTSPARMAGNGGDAALAL

LLLDRLNQLESKMSGKGQQQQGQTVTKKSAAEASKKPRQKRTATKAYNVTQAFGRRGP

EQTQGNFGDQELIRQ"

gap 27938..28111

/estimated\_length=174

ORIGIN

1 ttaccgcaag gttcttcttc gtaagaacgg taataaagga gctggtggcc ataggtacgg

61 cgccgatcta aagtcatttg acttaggcga cgagcttggc actgatcctt atgaagattt

121 tcaagaaaac tggaacacta aacatagcag tggtgttacc cgtgaactca tgcgtgagct

181 taacggaggg gcatacactc gctatgtcga taacaacttc tgtggccctg atggctaccc

241 tcttgagtgc attaaagacc ttctagcacg tgctggtaaa gcttcatgca ctttgtccga

301 acaactggac tttattgaca ctaagagggg tgtatactgc tgccgtgaac atgagcatga

361 aattgcttgg tacacggaac gttctgaaaa gagctatgaa ttgcagacac cttttgaaat

421 taaattggca aagaaatttg acaccttcaa tggggaatgt ccaaattttg tatttccctt

481 aaattccata atcaagacta ttcaaccaag ggttgaaaag aaaaagcttg atggctttat

541 gggtagaatt cgatctgtct atccagttgc gtcaccaaat gaatgcaacc aaatgtgcct

601 ttcaactctc atgaagtgtg atcattgtgg tgaaacttca tggcagacgg gcgattttgt

661 taaagccact tgcgaatttt gtggcactga gaatttgact aaagaaggtg ccactacttg

721 tggttactta ccccaaaatg ctgttgttaa aatttattgt ccagcatgtc acaattcaga

781 agtaggacct gagcatagtc ttgccgaata ccataatgaa tctggcttga aaaccattct

841 tcgtaagggt ggtcgcacta ttgcctttgg aggctgtgtg ttctcttatg ttggttgcca

901 taacaagtgt gcctattggg ttccacgtgc tagcgctaac ataggttgta accatacagg

961 tgttgttgga gaaggttccg aaggtcttaa tgacaacctt cttgaaatac tccaaaaaga

1021 gaaagtcaac atcaatattg ttggtgactt taaacttaat gaagagatcg ccattatttt

1081 ggcatctttt tctgcttcca caagtgcttt tgtggaaact gtgaaaggtt tggattataa

1141 agcattcaaa caaattgttg aatcctgtgg taattttaaa gttacaaaag gaaaagctaa

1201 aaaaggtgcc tggaatattg gtgaacagaa atcaatactg agtcctcttt atgcatttgc

1261 atcagaggct gctcgtgttg tacgatcaat tttctcccgc actcttgaaa ctgctcaaaa

1321 ttctgtgcgt gttttacaga aggccgctat aacaatacta gatggaattt cacagtattc

1381 actgagactc attgatgcta tgatgttcac atctgatttg gctactaaca atctagttgt

1441 aatggcctac attacaggtg gtgttgttca gttgacttcg cagtggctaa ctaacatctt

1501 tggcactgtt tatgaaaaac tcaaacccgt ccttgattgg cttgaagaga agtttaagga

1561 aggtgtagag tttcttagag acggttggga aattgttaaa tttatctcaa cctgtgcttg

1621 tgaaattgtc ggtggacaaa ttgtcacctg tgcaaaggaa attaaggaga gtgttcagac

1681 attctttaag cttgtaaata aatttttggc tttgtgtgct gactctatca ttattggtgg

1741 agctaaactt aaagccttga atttaggtga aacatttgtc acgcactcaa agggattgta

1801 cagaaagtgt gttaaatcca gagaagaaac tggcctactc atgcctctaa aagccccaaa

1861 agaaattatc ttcttagagg gagaaacact tcccacagaa gtgttaacag aggaagttgt

1921 cttgaaaact ggtgatttac aaccattaga acaacctact agtgaagctg ttgaagctcc

1981 attggttggt acaccagttt gtattaacgg gcttatgttg ctcgaaatca aagacacaga

2041 aaagtactgt gcccttgcac ctaatatgat ggtaacaaac aataccttca cactcaaagg

2101 cggtgcacca acaaaggtta cttttggtga tgacactgtg atagaagtgc aaggttacaa

2161 gagtgtgaat atcatttttg aacttgatga aaggattgat aaagtactta atgagaagtg

2221 ctctgcctat acagttgaac tcggtacaga agtaaatgag ttcgcctgtg ttgtggcaga

2281 tgctgtcata aaaactttgc aaccagtatc tgaattactt acaccactgg gcattgattt

2341 agatgagtgg agtatggcta catactactt atttgatgag tctggtgagt ttaaattggc

2401 ttcacatatg tattgttctt tttaccctcc agatgaggat gaagaagaag gtgattgtga

2461 agaagaagag tttgagccat caactcaata tgagtatggt actgaagatg attaccaagg

2521 taaacctttg gaatttggtg ccacttctgc tgctcttcaa cctgaagaag agcaagaaga

2581 agattggtta gatgatgata gtcaacaaac tgttggtcaa caagacggca gtgaggacaa

2641 tcagacaact actattcaaa caattgttga ggttcaacct caattagaga tggaacttac

2701 accagttgtt cagactattg aagtgaatag ttttagtggt tatttaaaac ttactgacaa

2761 tgtatacatt aaaaatgcag acattgtgga agaagctaaa aaggtaaaac caacagtggt

2821 tgttaatgca gccaatgttt accttaaaca tggaggaggt gttgcaggag ccttaaataa

2881 ggctactaac aatgccatgc aagttgaatc tgatgattac atagctacta atggaccact

2941 taaagtgggt ggtagttgtg ttttaagcgg acacaatctt gctaaacact gtcttcatgt

3001 tgtcggccca aatgttaaca aaggtgaaga cattcaactt cttaagagtg cttatgaaaa

3061 ttttaatcag cacgaagttc tacttgcacc attattatca gctggtattt ttggtgctga

3121 ccctatacat tctttaagag tttgtgtaga tactgttcgc acaaatgtct acttagctgt

3181 ctttgataaa aatctctatg acaaacttgt ttcaagcttt ttggaaatga agagtgaaaa

3241 gcaagttgaa caaaagatcg ctgagattcc taaagaggaa gttaagccat ttataactga

3301 aagtaaacct tcagttgaac agagaaaaca agatgataag aaaatcaaag cttgtgttga

3361 agaagttaca acaactctgg aagaaactaa gttcctcaca gaaaacttgt tactttatat

3421 tgacattaat ggcaatcttc atccagattc tgccactctt gttagtgaca ttgacatcac

3481 tttcttaaag aaagatgctc catatatagt gggtgatgtt gttcaagagg gtgttttaac

3541 tgctgtggtt atacctacta aaaaggctag tggcactact gaaatgctag cgaaagcttt

3601 gagaaaagtg ccaacagaca attatataac cacttacccg ggtcagggtt taaatggtta

3661 cactgtagag gaggcaaaga cagtgcttaa aaagtgtaaa agtgcttttt acattctacc

3721 atctattatc tctaatgaga agcaagaaat tcttggaact gtttcttgga atttgcgaga

3781 aatgcttgca catgcagaag aaacacgcaa attaatgcct gtctgtgtgg aaactaaagc

3841 catagtttca actatacagc gtaaatataa gggtattaaa atacaagagg gtgtggttga

3901 ttatggtgct agattttact tttacaccag taaaacaact gtagcgtcac ttatcaacac

3961 acttaacgat ctaaatgaaa ctcttgttac aatgccactt ggctatgtaa cacatggctt

4021 aaatttggaa gaagctgctc ggtatatgag atctctcaaa gtgccagcta cagtttctgt

4081 ttcttcacct gatgctgtta cagcgtataa tggttatctt acttcttctt ctaaaacacc

4141 tgaagaacat tttattgaaa ccatctcact tgctggttcc tataaagatt ggtcctattc

4201 tggacaatct acacaactag gtatagaatt tcttaagaga ggtgataaaa gtgtatatta

4261 cactagtaat cctaccacat tccacctaga tggtgaagtt atcacctttg acaatcttaa

4321 gacacttctt tctttgagag aagtgaggac tattaaggtg tttacaacag tagacaacat

4381 taacctccac acgcaagttg tggacatgtc aatgacatat ggacaacagt ttggtccaac

4441 ttatttggat ggagctgatg ttactaaaat aaaacctcat aattcacatg aaggtaaaac

4501 attttatgtt ttacctaatg atgacactct acgtgttgag gcttttgagt actaccacac

4561 aactgatcct agttttctgg gtaggtacat gtcagcatta aatcacacta aaaagtggaa

4621 atacccacaa gttaatggtt taacttctat taaatgggca gataacaact gttatcttgc

4681 cactgcattg ttaacactcc aacaaataga gttgaagttt aatccacctg ctctacaaga

4741 tgcttattac agagcaaggg ctggtgaagc tgctaacttt tgtgcactta tcttagccta

4801 ctgtaataag acagtaggtg agttaggtga tgttagagaa acaatgagtt acttgtttca

4861 acatgccaat ttagattctt gcaaaagagt cttgaacgtg gtgtgtaaaa cttgtggaca

4921 acagcagaca acccttaagg gtgtagaagc tgttatgtac atgggcacac tttcttatga

4981 acaatttaag aaaggtgttc agataccttg tacgtgtggt aaacaagcta caaaatatct

5041 agtacaacag gagtcacctt ttgttatgat gtcagcacca cctgctcagt atgaacttaa

5101 gcatggtaca tttacttgtg ctagtgagta cactggtaat taccagtgtg gtcactataa

5161 acatataact tctaaagaaa ctttgtattg catagacggt gctttactta caaagtcctc

5221 agaatacaaa ggtcctatta cggatgtttt ctacaaagaa aacagttaca caacaaccat

5281 aaaaccagtt acttataaat tggatggtgt tgtttgtaca gaaattgacc ctaagttgga

5341 caattattat aagaaagaca attcttattt cacagagcaa ccaattgatc ttgtaccaaa

5401 ccaaccatat ccaaacgcaa gcttcgataa ttttaagttt gtatgtgata atatcaaatt

5461 tgctgatgat ttaaaccagt taactggtta taagaaacct gcttcaagag agcttaaagt

5521 tacatttttc cctgacttaa atggtgatgt ggtggctatt gattataaac actacacacc

5581 ctcttttaag aaaggagcta aattgttaca taaacctatt gtttggcatg ttaacaatgc

5641 aactaataaa gccacgtata aaccaaatac ctggtgtata cgttgtcttt ggagcacaaa

5701 accagttgaa acatcaaatt cgtttgatgt actgaagtca gaggacgcgc agggaatgga

5761 taatcttgcc tgcgaagatc taaaaccagt ctctgaagaa gtagtggaaa atcctaccat

5821 acagaaagac gttcttgagt gtaatgtgaa aactaccgaa gttgtaggag acattatact

5881 taaaccagca aataatagtt taaaaattac agaagaggtt ggccacacag atctaatggc

5941 tgcttatgta gacaattcta gtcttactat taagaaacct aatgaattat ctagagtatt

6001 aggtttgaaa acccttgcta ctcatggttt agctgctgtt aatagtgtcc cttgggatac

6061 tatagctaat tatgctaagc cttttcttaa caaagttgtt agtacaacta ctaacatagt

6121 tacacggtgt ttaaaccgtg tttgtactaa ttatatgcct tatttcttta ctttattgct

6181 acaattgtgt acttttacta gaagtacaaa ttctagaatt aaagcatcta tgccgactac

6241 tatagcaaag aatactgtta agagtgtcgg taaattttgt ctagaggctt catttaatta

6301 tttgaagtca cctaattttt ctaaactgat aaatattata atttggtttt tactattaag

6361 tgtttgccta ggttctttaa tctactcaac cgctgcttta ggtgttttaa tgtctaattt

6421 aggcatgcct tcttactgta ctggttacag agaaggctat ttgaactcta ctaatgtcac

6481 tattgcaacc tactgtactg gttctatacc ttgtagtgtt tgtcttagtg gtttagattc

6541 tttagacacc tatccttctt tagaaactat acaaattacc atttcatctt ttaaatggga

6601 tttaactgct tttggcttag ttgcagagtg gtttttggca tatattcttt tcactaggtt

6661 tttctatgta cttggattgg ctgcaatcat gcaattgttt ttcagctatt ttgcagtaca

6721 ttttattagt aattcttggc ttatgtggtt aataattaat cttgtacaaa tggccccgat

6781 ttcagctatg gttagaatgt acatcttctt tgcatcattt tattatgtat ggaaaagtta

6841 tgtgcatgtt gtagacggtt gtaattcatc aacttgtatg atgtgttaca aacgtaatag

6901 agcaacaaga gtcgaatgta caactattgt taatggtgtt agaaggtcct tttatgtcta

6961 tgctaatgga ggtaaaggct tttgcaaact acacaattgg aattgtgtta attgtgatac

7021 attctgtgct ggtagtacat ttattagtga tgaagttgcg agagacttgt cactacagtt

7081 taaaagacca ataaatccta ctgaccagtc ttcttacatc gttgatagtg ttacagtgaa

7141 gaatggttcc atccatcttt actttgataa agctggtcaa aagacttatg aaagacattc

7201 tctctctcat tttgttaact tagacaacct gagagctaat aacactaaag gttcattgcc

7261 tattaatgtt atagtttttg atggtaaatc aaaatgtgaa gaatcatctg caaaatcagc

7321 gtctgtttac tacagtcagc ttatgtgtca acctatactg ttactagatc aggcattagt

7381 gtctgatgtt ggtgatagtg cggaagttgc agttaaaatg tttgatgctt acgttaatac

7441 gttttcatca acttttaacg taccaatgga aaaactcaaa acactagttg caactgcaga

7501 agctgaactt gcaaagaatg tgtccttaga caatgtctta tctactttta tttcagcagc

7561 tcggcaaggg tttgttgatt cagatgtaga aactaaagat gttgttgaat gtcttaaatt

7621 gtcacatcaa tctgacatag aagttactgg cgatagttgt aataactata tgctcaccta

7681 taacaaagtt gaaaacatga caccccgtga ccttggtgct tgtattgact gtagtgcgcg

7741 tcatattaat gcgcaggtag caaaaagtca caacattgct ttgatatgga acgttaaaga

7801 tttcatgtca ttgtctgaac aactacgaaa acaaatacgt agtgctgcta aaaagaataa

7861 cttacctttt aagttgacat gtgcaactac tagacaagtt gttaatgttg taacaacaaa

7921 gatagcactt aagggtggta aaattgttaa taattggttg aagcagttaa ttaaagttac

7981 acttgtgttc ctttttgttg ctgctatttt ctatttaata acacctgttc atgtcatgtc

8041 taaacatact gacttttcaa gtgaaatcat aggatacaag gctattgatg gtggtgtcac

8101 tcgtgacata gcatctacag atacttgttt tgctaacaaa catgctgatt ttgacacatg

8161 gtttagccag cgtggtggta gttatactaa tgacaaagct tgcccattga ttgctgcagt

8221 cataacaaga gaagtgggtt ttgtcgtgcc tggtttgcct ggcacgatat tacgcacaac

8281 taatggtgac tttttgcatt tcttacctag agtttttagt gcagttggta acatctgtta

8341 cacaccatca aaacttatag agtacactga ctttgcaaca tcagcttgtg ttttggctgc

8401 tgaatgtaca atttttaaag atgcttctgg taagccagta ccatattgtt atgataccaa

8461 tgtactagaa ggttctgttg cttatgaaag tttacgccct gacacacgtt atgtgctcat

8521 ggatggctct attattcaat ttcctaacac ctaccttgaa ggttctgtta gagtggtaac

8581 aacttttgat tctgagtact gtaggcacgg cacttgtgaa agatcagaag ctggtgtttg

8641 tgtatctact agtggtagat gggtacttaa caatgattat tacagatctt taccaggagt

8701 tttctgtggt gtagatgctg taaatttatt tactaatatg tttacaccac taattcaacc

8761 tattggtgct ttggacatat cagcatctat agtagctggt ggtattgtgg ctatcgtagt

8821 aacatgcctt gcctactatt ttatgaggtt tagaagagct tttggtgaat acagtcatgt

8881 agttgccttt aatactttac tattccttat gtcattcatt gtactctgtt taacaccagt

8941 ttactcattc ttacctggtg tttattctgt tatttacttg tacttgacat tttatcttac

9001 taatgatgtt tcttttttag cacatattca gtggatggtt atgttcacac ctttagtacc

9061 tttctggata acaattgctt atatcatttg tatttccaca aagcatttct attggttctt

9121 tagtaattac ctaaagagac gtgtagtctt taatggtgtt tcctttagta cttttgaaga

9181 agctgcgctg tgcacctttt tgttaaataa agaaatgtat ctaaagttgc gtagtgatgt

9241 gctattacct tttacgcaat ataatagata cttagctctt tataataagt acaagtattt

9301 tagtggagca atggatacaa ctagctacag agaagctgct tgttgtcatc tcgcaaaggc

9361 tctcaatgac ttcagtaact caggttctga tgttctttac caaccaccac aaatctctat

9421 cacctcagct gttttgcaga gtggttttag aaaaatggca ttcccatctg gtaaagttga

9481 gggttgtatg gtacaagtaa cttgtggtac aactacactt aacggtcttt ggcttgatga

9541 cgtagtttac tgtccaagac atgtgatctg cacctctgaa gatatgctta accctaatta

9601 tgaagattta ctcattcgta agtctaatca taatttcttg gtacaggctg gtaatgttca

9661 actcagggtt attggacatt ctatgcaaaa ttgtgtactt aagcttaagg ttgatacagc

9721 caatcctaag acacctaagt ataagtttgt tcgcattcaa ccaggacaga ctttttcagt

9781 gttagcttgt tacaatggtt caccatctgg tgtttaccaa tgtgctatga gacacaattt

9841 cactattaag ggttcattcc ttaatggttc atgtggtagt gttggtttta acatagatta

9901 tgactgtgtc tctttttgtt acatgcacca tatggaatta ccaactggag ttcatgctgg

9961 cacagactta gaaggtaact tttatggacc ttttgttgac aggcaaacag cacaagcagc

10021 tggtacggac acaactatta cagttaatgt tttagcttgg ttgtacgctg ctgttataaa

10081 tggagacagg tggtttctca atcgatttac cacaactctt aatgacttta accttgtggc

10141 tatgaagtac aattatgaac ctctaacaca agaccatgtt gacatactag gacctctttc

10201 tgctcaaact ggaattgccg ttttagatat gtgtgcttca ttaaaagaat tactgcaaaa

10261 tggtatgaat ggacgtacca tattgggtag tgctttatta gaagatgaat ttacaccttt

10321 tgatgttgtt agacaatgct caggtgttac tttccaaagt gcagtgaaaa gaacaatcaa

10381 gggtacacac cactggttgt tactcacaat tttgacttca cttttagttt tagtccagag

10441 tactcaatgg tctttgttct tttttttgta tgaaaatgcc tttttacctt ttgctatggg

10501 tattattgct atgtctgctt ttgcaatgat gtttgtcaaa cataagcatg catttctctg

10561 tttgtttttg ttaccttctc ttgccactgt agcttatttt aatatggtct atatgcctgc

10621 tagttgggtg atgcgtatta tgacatggtt ggatatggtt gatactagtt tgaagctaaa

10681 agactgtgtt atgtatgcat cagctgtagt gttactaatc cttatgacag caagaactgt

10741 gtatgatgat ggtgctagga gagtgtggac acttatgaat gtcttgacac tcgtttataa

10801 agtttattat ggtaatgctt tagatcaagc catttccatg tgggctctta taatctctgt

10861 tacttctaac tactcaggtg tagttacaac tgtcatgttt ttggccagag gtattgtttt

10921 tatgtgtgtt gagtattgcc ctattttctt cataactggt aatacacttc agtgtataat

10981 gctagtttat tgtttcttag gctatttttg tacttgttac tttggcctct tttgtttact

11041 caaccgctac tttagactga ctcttggtgt ttatgattac ttagtttcta cacaggagtt

11101 tagatatatg aattcacagg gactactccc acccaagaat agcatagatg ccttcaaact

11161 caacattaaa ttgttgggtg ttggtggcaa accttgtatc aaagtagcca ctgtacagtc

11221 taaaatgtca gatgtaaagt gcacatcagt agtcttactc tcagttttgc aacaactcag

11281 agtagaatca tcatctaaat tgtgggctca atgtgtccag ttacacaatg acattctctt

11341 agctaaagat actactgaag cctttgaaaa aatggtttca ctactttctg ttttgctttc

11401 catgcagggt gctgtagaca taaacaagct ttgtgaagaa atgctggaca acagggcaac

11461 cttacaagct atagcctcag agtttagttc ccttccatca tatgcagctt ttgctactgc

11521 tcaagaagct tatgagcagg ctgttgctaa tggtgattct gaagttgttc ttaaaaagtt

11581 gaagaagtct ttgaatgtgg ctaaatctga atttgaccgt gatgcagcca tgcaacgtaa

11641 gttggaaaag atggctgatc aagctatgac ccaaatgtat aaacaggcta gatctgagga

11701 caagagggca aaagttacta gtgctatgca gacaatgctt ttcactatgc ttagaaagtt

11761 ggataatgat gcactcaaca acattatcaa caatgcaaga gatggttgtg ttcccttgaa

11821 cataatacct cttacaacag cagccaaact aatggttgtc ataccagact ataacacata

11881 taaaaatacg tgtgatggta caacatttac ttatgcatca gcattgtggg aaatccaaca

11941 ggttgtagat gcagatagta aaattgttca acttagtgaa attagtatgg acaattcacc

12001 taatttagca tggcctctta ttgtaacagc tttaagggcc aattctgctg tcaaattaca

12061 gaataatgag cttagtcctg ttgcactacg acagatgtct tgtgctgccg gtactacaca

12121 aactgcttgc actgatgaca atgcgttagc ttactacaac acaacaaagg gaggtaggtt

12181 tgtacttgca ctgttatccg atttacagga tttgaaatgg gctagattcc ctaagagtga

12241 tggaactggt actatttata cagaactgga accaccttgt aggtttgtta cagacacacc

12301 taaaggtcct aaagtgaagt atttatactt tattaaagga ttaaacaacc taaatagagg

12361 tatggtactt ggtagtttag ctgccacagt acgtctacaa gctggtaatg caacagaagt

12421 gcctgccaat tcaactgtat tatctttctg tgcttttgct gtagatgctg ctaaagctta

12481 caaagattat ctagctagtg ggggacaacc aatcactaat tgtgttaaga tgttgtgtac

12541 acacactggt actggtcagg caataacagt tacaccggaa gccaatatgg atcaagaatc

12601 ctttggtggt gcatcgtgtt gtctgtactg ccgttgccac atagatcatc caaatcctaa

12661 aggattttgt gacttaaaag gtaagtatgt acaaatacct acaacttgtg ctaatgaccc

12721 tgtgggtttt acacttaaaa acacagtctg taccgtctgc ggtatgtgga aaggttatgg

12781 ctgtagttgt gatcaactcc gcgaacccat gcttcagtca gctgatgcac aatcgttttt

12841 aaacgggttt gcggtgtaag tgcagcccgt cttacaccgt gcggcacagg cactagtact

12901 gatgtcgtat acagggcttt tgacatctac aatgataaag tagctggttt tgctaaattc

12961 ctaaaaacta attgttgtcg cttccaagaa aaggacgaag atgacaattt aattgattct

13021 tactttgtag ttaagagaca cactttctct aactaccaac atgaagaaac aatttataat

13081 ttacttaagg attgtccagc tgttgctaaa catgacttct ttaagtttag aatagacggt

13141 gacatggtac cacatatatc acgtcaacgt cttactaaat acacaatggc agacctcgtc

13201 tatgctttaa ggcattttga tgaaggtaat tgtgacacat taaaagaaat acttgtcaca

13261 tacaattgtt gtgatgatga ttatttcaat aaaaaggact ggtatgattt tgtagaaaac

13321 ccagatatat tacgcgtata cgccaactta ggtgaacgtg tacgccaagc tttgttaaaa

13381 acagtacaat tctgtgatgc catgcgaaat gctggtattg ttggtgtact gacattagat

13441 aatcaagatc tcaatggtaa ctggtatgat ttcggtgatt tcatacaaac cacgccaggt

13501 agtggagttc ctgttgtaga ttcttattat tcattgttaa tgcctatatt aaccttgacc

13561 agggctttaa ctgcagagtc acatgttgac actgacttaa caaagcctta cattaagtgg

13621 gatttgttaa aatatgactt cacggaagag aggttaaaac tctttgaccg ttattttaaa

13681 tattgggatc agacatacca cccaaattgt gttaactgtt tggatgacag atgcattctg

13741 cattgtgcaa actttaatgt tttattctct acagtgttcc cacttacaag ttttggacca

13801 ctagtgagaa aaatatttgt tgatggtgtt ccatttgtag tttcaactgg ataccacttc

13861 agagagctag gtgttgtaca taatcaggat gtaaacttac atagctctag acttagtttt

13921 aaggaattac ttgtgtatgc tgctgaccct gctatgcacg ctgcttctgg taatctatta

13981 ctagataaac gcactacgtg cttttcagta gctgcactta ctaacaatgt tgcttttcaa

14041 actgtcaaac ccggtaattt taacaaagac ttctatgact ttgctgtgtc taagggtttc

14101 tttaaggaag gaagttctgt tgaattaaaa cacttcttct ttgctcagga tggtaatgct

14161 gctatcagcg attatgacta ctatcgttat aatctaccaa caatgtgtga tatcagacaa

14221 ctactatttg tagttgaagt tgttgataag tactttgatt gttacgatgg tggctgtatt

14281 aatgctaacc aagtcatcgt caacaaccta gacaaatcag ctggttttcc atttaataaa

14341 tggggtaagg ctagacttta ttatgattca atgagttatg aggatcaaga tgcacttttc

14401 gcatatacaa aacgtaatgt catccctact ataactcaaa tgaatcttaa gtatgccatt

14461 agtgcaaaga atagagctcg caccgtagct ggtgtctcta tctgtagtac tatgaccaat

14521 agacagtttc atcaaaaatt attgaaatca atagccgcca ctagaggagc tactgtagta

14581 attggaacaa gcaaattcta tggtggttgg cacaacatgt taaaaactgt ttatagtgat

14641 gtagaaaacc ctcaccttat gggttgggat tatcctaaat gtgatagagc catgcctaac

14701 atgcttagaa ttatggcctc acttgttctt gctcgcaaac atacaacgtg ttgtagcttg

14761 tcacaccgtt tctatagatt agctaatgag tgtgctcaag tattgagtga aatggtcatg

14821 tgtggcggtt cactatatgt taaaccaggt ggaacctcat caggagatgc cacaactgct

14881 tatgctaata gtgtttttaa catttgtcaa gctgtcacgg ccaatgttaa tgcactttta

14941 tctactgatg gtaacaaaat tgccgataag tatgtccgca atttacaaca cagactttat

15001 gagtgtctct atagaaatag agatgttgac acagactttg tgaatgagtt ttacgcatat

15061 ttgcgtaaac atttctcaat gatgatactt tctgacgatg ctgttgtgtg tttcaatagc

15121 acttatgcat ctcaaggtct agtggctagc ataaagaact ttaagtcagt tctttattat

15181 caaaacaatg tttttatgtc tgaagcaaaa tgttggactg agactgacct tactaaagga

15241 cctcatgaat tttgctctca acatacaatg ctagttaaac agggtgatga ttatgtgtac

15301 cttccttacc cagatccatc aagaatccta ggggccggct gttttgtaga tgatatcgta

15361 aaaacagatg gtacacttat gattgaacgg ttcgtgtctt tagctataga tgcttaccca

15421 cttactaaac atcctaatca ggagtatgct gatgtctttc atttgtactt acaatacata

15481 agaaagctac atgatgagtt aacaggacac atgttagaca tgtattctgt tatgcttact

15541 aatgataaca cttcaaggta ttgggaacct gagttttatg aggctatgta cacaccgcat

15601 acagtcttac aggctgttgg ggcttgtgtt ctttgcaatt cacagacttc attaagatgt

15661 ggtgcttgca tacgtagacc attcttatgt tgtaaatgct gttacgacca tgtcatatca

15721 acatcacata aattagtctt gtctgttaat ccgtatgttt gcaatgctcc aggttgtgat

15781 gtcacagatg tgactcaact ttacttagga ggtatgagct attattgtaa atcacataaa

15841 ccacccatta gttttccatt gtgtgctaat ggacaagttt ttggtttata taaaaataca

15901 tgtgttggta gcgataatgt tactgacttt aatgcaattg caacatgtga ctggacaaat

15961 gctggtgatt acattttagc taacacctgt actgaaagac tcaagctttt tgcagcagaa

16021 acgctcaaag ctactgagga gacatttaaa ctgtcttatg gtattgctac tgtacgtgaa

16081 gtgctgtctg acagagaatt acatctttca tgggaagttg gtaaacctag accaccactt

16141 aaccgaaatt atgtctttac tggttatcgt gtaactaaaa acagtaaagt acaaatagga

16201 gagtacacct ttgaaaaagg tgactatggt gatgctgttg tttaccgagg tacaacaact

16261 tacaaattaa atgttggtga ttattttgtg ctgacatcac atacagtaat gccattaagt

16321 gcacctacac tagtgccaca agagcactat gttagaatta ctggcttata cccaacactc

16381 aatatctcag atgagttttc tagcaatgtt gcaaattatc aaaaggttgg tatgcaaaag

16441 tattctacac tccagggacc acctggtact ggtaagagtc attttgctat tggcctagct

16501 ctctactacc cttctgctcg catagtgtat acagcttgct ctcatgccgc tgttgatgca

16561 ctatgtgaga aggcattaaa atatttgcct atagataaat gtagtagaat tatacctgca

16621 cgtgctcgtg tagagtgttt tgataaattc aaagtgaatt caacattaga acagtatgtc

16681 ttttgtactg taaatgcatt gcctgagacg acagcagata tagttgtctt tgatgaaatt

16741 tcaatggcca caaattatga tttgagtgtt gtcaatgcca gattatgtgc taagcactat

16801 gtgtacattg gcgaccctgc tcaattacct gcaccacgca cattgctaac taagggcaca

16861 ctagaaccag aatatttcaa ttcagtgtgt agacttatga aaactatagg tccagacatg

16921 ttcctcggaa cttgtcggcg ttgtcctgct gaaattgttg acactgtgag tgctttggtt

16981 tatgataata agcttaaagc acataaagac aaatcagctc aatgctttaa aatgttttat

17041 aagggtgtta tcacgcatga tgtttcatct gcaattaaca ggccacaaat aggcgtggta

17101 agagaattcc ttacacgtaa ccctgcttgg agaaaagctg tctttatttc accttataat

17161 tcacagaatg ctgtagcctc aaagattttg ggactaccaa ctcaaactgt tgattcatca

17221 cagggctcag aatatgacta tgtcatattc actcaaacca ctgaaacagc tcactcttgt

17281 aatgtaaaca gatttaatgt tgctattacc agagcaaaag taggcatact ttgcataatg

17341 tctgatagag acctttatga caagttgcaa tttacaagtc ttgaaattcc acgtaggaat

17401 gtggcaactt tacaagctga aaatgtaaca ggactcttta aagattgtag taaggtaatc

17461 actgggttac atcctacaca ggcacctaca cacctcagtg ttgacactaa attcaaaact

17521 gaaggtttat gtgttgacgt acctggcata cctaaggaca tgacctatag aagactcatc

17581 tctatgatgg gttttaaaat gaattatcaa gttaatggtt accctaacat gtttatcacc

17641 cgcgaagaag ctataagaca tgtacgtgca tggattggct tcgatgtcga ggggtgtcat

17701 gctactagag aagctgttgg taccaattta cctttacagc taggtttttc tacaggtgtt

17761 aacctagttg ctgtacctac aggttatgtt gatacaccta ataatacaga tttttccaga

17821 gttagtgcta aaccaccgcc tggagatcaa tttaaacacc tcataccact tatgtacaaa

17881 ggacttcctt ggaatgtagt gcgtataaag attgtacaaa tgttaagtga cacacttaaa

17941 aatctctctg acagagtcgt atttgtctta tgggcacatg gctttgagtt gacatctatg

18001 aagtattttg tgaaaatagg acctgagcgc acctgttgtc tatgtgatag acgtgccaca

18061 tgcttttcca ctgcttcaga cacttatgcc tgttggcatc attctattgg atttgattac

18121 gtctataatc cgtttatgat tgatgttcaa caatggggtt ttacaggtaa cctacaaagc

18181 aaccatgatc tgtattgtca agtccatggt aatgcacatg tagctagttg tgatgcaatc

18241 atgactaggt gtctagctgt ccacgagtgc tttgttaagc gtgttgactg gactattgaa

18301 tatcctataa ttggtgatga actgaagatt aatgcggctt gtagaaaggt tcaacacatg

18361 gttgttaaag ctgcattatt agcagacaaa ttcccagttc ttcacgacat tggtaaccct

18421 aaagctatta agtgtgtacc tcaagctgat gtagaatgga agttctatga tgcacagcct

18481 tgtagtgaca aagcttataa aatagaagaa ttattctatt cttatgccac acattctgac

18541 aaattcacag atggtgtatg cctattttgg aattgcaatg tcgatagata tcctgctaat

18601 tccattgttt gtagatttga cactagagtg ctatctaacc ttaacttgcc tggttgtgat

18661 ggtggcagtt tgtatgtaaa taaacatgca ttccacacac cagcttttga taaaagtgct

18721 tttgttaatt taaaacaatt accatttttc tattactctg acagtccatg tgagtctcat

18781 ggaaaacaag tagtgtcaga tatagattat gtaccactaa agtctgctac gtgtataaca

18841 cgttgcaatt taggtggtgc tgtctgtaga catcatgcta atgagtacag attgtatctc

18901 gatgcttata acatgatgat ctcagctggc tttagcttgt gggtttacaa acaatttgat

18961 acttataacc tctggaacac ttttacaaga cttcagagtt tagaaaatgt ggcttttaat

19021 gttgtaaata agggacactt tgatggacaa cagggtgaag taccagtttc tatcattaat

19081 aacactgttt acacaaaagt tgatggtgtt gatgtagaat tgtttgaaaa taaaacaaca

19141 ttacctgtta atgtagcatt tgagctttgg gctaagcgca acattaaacc agtaccagag

19201 gtgaaaatac tcaataattt gggtgtggac attgctgcta atactgtgat ctgggactac

19261 aaaagagatg ctccagcaca tatatctact attggtgttt gttctatgac tgacatagcc

19321 aagaaaccaa ttgaaacgat ttgtgcacca ctcactgtct tttttgatgg tagagttgat

19381 ggtcaagtag acttatttag aaatgcccgt aatggtgttc ttattacaga gggtagtgtt

19441 aaaggtttac aaccatctgt aggtcccaaa caagctagtc ttaatggagt cacattaatt

19501 ggagaagccg taaaaacaca gttcaattat tataagaaag ttgatggtgt tgtccaacaa

19561 ttacctgaaa cttactttac tcagagtaga aatttacaag aatttaaacc caggagtcaa

19621 atggaaattg atttcttaga attagctatg gatgaattca ttgaacggta taaattagaa

19681 ggctatgcct tcgaacatat cgtttatgga gattttagtc atagtcagtt aggtggttta

19741 catctactga ttggactagc taaacgtttt aaggaatcac cttttgaatt agaagatttt

19801 attcctatgg acagtacagt taaaaactat ttcataacag atgcgcaaac aggttcatct

19861 aagtgtgtgt gttctgttat tgatttatta cttgatgatt ttgttgaaat aataaaatcc

19921 caagatttat ctgtagtttc taaggttgtc aaagtgacta ttgactatac agaaatttca

19981 tttatgcttt ggtgtaaaga tggccatgta gaaacatttt acccaaaatt acaatctagt

20041 caagcgtggc aaccgggtgt tgctatgcct aatctttaca aaatgcaaag aatgctatta

20101 gaaaagtgtg accttcaaaa ttatggtgat agtgcaacat tacctaaagg cataatgatg

20161 aatgtcgcaa aatatactca actgtgtcaa tatttaaaca cattaacatt agctgtaccc

20221 tataatatga gagttataca ttttggtgct ggttctgata aaggagttgc accaggtaca

20281 gctgttttaa gacagtggtt gcctacgggt acgctgcttg tcgattcaga tcttaatgac

20341 tttgtctctg atgcagattc aactttgatt ggtgattgtg caactgtaca tacagctaat

20401 aaatgggatc tcattattag tgatatgtac gaccctaaga ctaaaaatgt tacaaaagaa

20461 aatgactcta aagagggttt tttcacttac atttgtgggt ttatacaaca aaagctagct

20521 cttggaggtt ccgtggctat aaagataaca gaacattctt ggaatgctga tctttataag

20581 ctcatgggac acttcgcatg gtggacagcc tttgttacta atgtgaatgc gtcatcatct

20641 gaagcatttt taattggatg taattatctt ggcaaaccac gcgaacaaat agatggttat

20701 gtcatgcatg caaattacat attttggagg aatacaaatc caattcagtt gtcttcctat

20761 tctttatttg acatgagtaa atttcccctt aaattaaggg gtactgctgt tatgtcttta

20821 aaagaaggtc aaatcaatga tatgatttta tctcttctta gtaaaggtag acttataatt

20881 agagaaaaca acagagttgt tatttctagt gatgttcttg ttaacaacta aacgaacaat

20941 gtttgttttt cttgttttat tgccactagt ctctagtcag tgtgttaatc ttataaccag

21001 aactcaatca tacactaatt ctttcacacg tggtgtttat taccctgaca aagttttcag

21061 atcctcagtt ttacattcaa ctcaggactt gttcttacct ttcttttcca atgttacttg

21121 gttccatgct atacatgtct ctgggaccaa tggtactaag aggtttgata accctgtcct

21181 accatttaat gatggtgttt attttgcttc cactgagaag tctaacataa taagaggctg

21241 gatttttggt actactttag attcgaagac ccagtcccta cttattgtta ataacgctac

21301 taatgttgtt attaaagtct gtgaatttca attttgtaat gatccatttt tggatgttta

21361 ttaccacaaa aacaacaaaa gttggatgga aagtgagttc agagtttatt ctagtgcgaa

21421 taattgcact tttgaatatg tctctcagcc ttttcttatg gaccttgaag gaaaacaggg

21481 taatttcaaa aatcttaggg aatttgtgtt taagaatatt gatggttatt ttaaaatata

21541 ttctaagcac acgcctatta atttagggcg tgatctccct cagggttttt cggctttaga

21601 accattggta gatttgccaa taggtattaa catcactagg tttcaaactt tacttgcttt

21661 acatagaagt tatttgactc ctggtgattc ttcttcaggt tggacagctg gtgctgcagc

21721 ttattatgtg ggttatcttc aacctaggac ttttctatta aaatataatg aaaatggaac

21781 cattacagat gctgtagact gtgcacttga ccctctctca gaaacaaagt gtacgttgaa

21841 atccttcact gtagaaaaag gaatctatca aacttctaac tttagagtcc aaccaacaga

21901 atctattgtt agatttccta atattacaaa cttgtgccct tttgatgaag tttttaacgc

21961 caccagattt gcatctgttt atgcttggaa caggaagaga atcagcaact gtgttgctga

22021 ttattctgtc ctatataatt tcgcaccatt tttcgctttt aagtgttatg gagtgtctcc

22081 tactaaatta aatgatctct gctttactaa tgtctatgca gattcatttg taattagagg

22141 taatgaagtc agccaaatcg ctccagggca aactggaaat attgctgatt ataattataa

22201 attaccagat gattttacag gctgcgttat agcttggaat tctaacaagc ttgattctaa

22261 ggttggtggt aattataatt acctgtatag attgtttagg aagtctaatc tcaaaccttt

22321 tgagagagat atttcaactg aaatctatca ggccggtaac aaaccttgta atggtgttgc

22381 aggttttaat tgttactttc ctttacgatc atatggtttc cgacccactt atggtgttgg

22441 tcaccaacca tacagagtag tagtactttc ttttgaactt ctacatgcac cagcaactgt

22501 ttgtggacct aaaaagtcta ctaatttggt taaaaacaaa tgtgtcaatt tcaacttcaa

22561 tggtttaaca ggcacaggtg ttcttactga gtctaacaaa aagtttctgc ctttccaaca

22621 atttggcaga gacattgctg acactactga tgctgtccgt gatccacaga cacttgagat

22681 tcttgacatt acaccatgtt cttttggtgg tgtcagtgtt ataacaccag gaacaaatac

22741 ttctaaccag gttgctgttc tttatcagga tgttaactgc acagaagtcc ctgttgctat

22801 tcatgcagat caacttactc ctacttggcg tgtttattct acaggttcta atgtttttca

22861 aacacgtgca ggctgtttaa taggggctga atatgtcaac aactcatatg agtgtgacat

22921 acccattggt gcaggtatat gcgctagtta tcagactcag actaagtctc atcggcgggc

22981 acgtagtgta gctagtcaat ccatcattgc ctacactatg tcacttggtg cagaaaattc

23041 agttgcttac tctaataact ctattgccat acccacaaat tttactatta gtgttaccac

23101 agaaattcta ccagtgtcta tgaccaagac atcagtagat tgtacaatgt acatttgtgg

23161 tgattcaact gaatgcagca atcttttgtt gcaatatggc agtttttgta cacaattaaa

23221 acgtgcttta actggaatag ctgttgaaca agacaaaaac acccaagaag tttttgcaca

23281 agtcaaacaa atttacaaaa caccaccaat taaatatttt ggtggtttta atttttcaca

23341 aatattacca gatccatcaa aaccaagcaa gaggtcattt attgaagatc tacttttcaa

23401 caaagtgaca cttgcagatg ctggcttcat caaacaatat ggtgattgcc ttggtgatat

23461 tgctgctaga gacctcattt gtgcacaaaa gtttaacggc cttactgttt tgccaccttt

23521 gctcacagat gaaatgattg ctcaatacac ttctgcactg ttagcgggta caatcacttc

23581 tggttggacc tttggtgcag gtgctgcatt acaaatacca tttgctatgc aaatggctta

23641 taggtttaat ggtattggag ttacacagaa tgttctctat gagaaccaaa aattgattgc

23701 caaccaattt aatagtgcta ttggcaaaat tcaagactca ctttcttcca cagcaagtgc

23761 acttggaaaa cttcaagatg tggtcaacca taatgcacaa gctttaaaca cgcttgttaa

23821 acaacttagc tccaaatttg gtgcaatttc aagtgtttta aatgatatcc tttcacgtct

23881 tgacaaagtt gaggctgaag tgcaaattga taggttgatc acaggcagac ttcaaagttt

23941 gcagacatat gtgactcaac aattaattag agctgcagaa atcagagctt ctgctaatct

24001 tgctgctact aaaatgtcag agtgtgtact tggacaatca aaaagagttg atttttgtgg

24061 aaagggctat catcttatgt ccttccctca gtcagcacct catggtgtag tcttcttgca

24121 tgtgacttat gtccctgcac aagaaaagaa cttcacaact gctcctgcca tttgtcatga

24181 tggaaaagca cactttcctc gtgaaggtgt ctttgtttca aatggcacac actggtttgt

24241 aacacaaagg aatttttatg aaccacaaat cattactaca gacaacacat ttgtgtctgg

24301 taactgtgat gttgtaatag gaattgtcaa caacacagtt tatgatcctt tgcaacctga

24361 attagattca ttcaaggagg agttagataa atattttaag aatcatacat caccagatgt

24421 tgatttaggt gacatctctg gcattaatgc ttcagttgta aacattcaaa aagaaattga

24481 ccgcctcaat gaggttgcca agaatttaaa tgaatctctc atcgatctcc aagaacttgg

24541 aaagtatgag cagtatataa aatggccatg gtacatttgg ctaggtttta tagctggctt

24601 gattgccata gtaatggtga caattatgct ttgctgtatg accagttgct gtagttgtct

24661 caagggctgt tgttcttgtg gatcctgctg caaatttgat gaagacgact ctgagccagt

24721 gctcaaagga gtcaaattac attacacata aacgaactta tggatttgtt tatgagaatc

24781 ttcacaattg gaactgtaac tttgaagcaa ggtgaaatca aggatgctac tccttcagat

24841 tttgttcgcg ctactgcaac gataccgata caagcctcac tccctttcgg atggcttatt

24901 gttggcgttg cacttcttgc tgtttttcag agcgcttcca aaatcataac tctcaaaaag

24961 agatggcaac tagcactctc caagggtgtt cactttgttt gcaacttgct gttgttgttt

25021 gtaacagttt actcacacct tttgctcgtt gctgctggcc ttgaagcccc ttttctctat

25081 ctttatgctt tagtctactt cttgcagagt ataaactttg taagaataat aatgaggctt

25141 tggctttgct ggaaatgccg ttccaaaaac ccattacttt atgatgccaa ctattttctt

25201 tgctggcata ctaattgtta cgactattgt ataccttaca atagtgtaac ttcttcaatt

25261 gtcattactt caggtgatgg cacaacaagt cctatttctg aacatgacta ccagattggt

25321 ggttatactg aaaaatggga atctggagta aaagactgtg ttgtattaca cagttacttc

25381 acttcagact attaccagct gtactcaact caattgagta cagacattgg tgttgaacat

25441 gttaccttct tcatctacaa taaaattgtt gatgagcctg aagaacatgt ccaaattcac

25501 acaatcgacg gttcatccgg agttgttaat ccagtaatgg aaccaattta tgatgaaccg

25561 acgacgacta ctagcgtgcc tttgtaagca caagctgatg agtacgaact tatgtactca

25621 ttcgtttcgg aagagatagg tacgttaata gttaatagcg tacttctttt tcttgctttc

25681 gtggtattct tgctagttac actagccatc cttactgcgc ttcgattgtg tnnnnnnnnn

25741 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnttttt acgtttactc tcgtgttaaa

25801 aatctgaatt cttctagagt t

[gap 107 bp] [Expand Ns](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?expand-gaps=on)

25929 aa aaagctcctt gaagaatgga acctagtaat aggtttccta ttccttacat

25981 ggatttgtct tctacaattt g

[gap 173 bp] [Expand Ns](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?expand-gaps=on)

26175 ttcatt gcttctttca gactgtttgc gcgtacgcgt tccatgtggt

26221 catttaatcc agaaactaac attcttctca acgtgccact ccatggcact attctgacca

26281 gaccgcttct agaaagtgaa ctcgtaatcg gagctgtgat ccttcgtgga catcttcgta

26341 ttgctggaca ccatctagga cgctgtgaca tcaaggacct gcctaaagaa atcactgttg

26401 ctacatcacg aacgctttct tattacaaat tgggagcttc gcagcgtgta gcaggtgact

26461 caggttttgc tgcatacagt cgctacagga ttggcaacta taaattaaac acagaccatt

26521 ccagtagcag tgacaatatt gctttgcttg tacagtaagt gacaacagat gtttcatctc

26581 gttgactttc aggttactat agcagagata ttactaatta ttatgcggac ttttaaagtt

26641 tccatttgga atcttgatta catcataaac ctcataatta aaaatttatc taagtcacta

26701 actgagaata aatattctca attagatgaa gagcaaccaa tggagattct ctaaacgaac

26761 atgaaaatta ttcttttctt ggcactgata acactcgcta cttgtgagct ttatcactac

26821 caagagtgtg ttagaggtac aacagtactt ttaaaagaac cttgctcttc tggaacatac

26881 gagggcaatt caccatttca tcctctagct gataacaaat ttgcactgac ttgctttagc

26941 actcaatttg cttttgcttg tcctgacggc gtaaaacacg tctatcagtt acgtgccaga

27001 tcagtttcac ctaaactgtt catcagacaa gaggaagttc aagaacttta ctctccaatt

27061 tttcttattg ttgcggcaat agtgtttata acactttgct tcacactcaa aagaaagaca

27121 gaatgattga actttcatta attgacttct atttgtgctt tttagccttt ctgttattcc

27181 ttgttttaat tatgcttatt atcttttggt tctcacttga actgcaagat cataatgaaa

27241 cttgtcacgc ctaaacgaac atgaaatttc ttgttttctt aggaatcatc acaactgtag

27301 ctgcatttca ccaagaatgt agtttacagt catgtactca acatcaacca tatgtagttg

27361 atgacccgtg tcctattcac ttctattcta aatggtatat tagagtagga gctagaaaat

27421 cagcaccttt aattgaattg tgcgtggatg aggctggttc taaatcaccc attcagtaca

27481 tcgatatcgg taattataca gtttcctgtt taccttttac aattaattgc caggaaccta

27541 aattgggtag tcttgtagtg cgttgttcgt tctatgaaga ctttttagag tatcatgacg

27601 ttcgtgttgt tttagatttc atctaaacga acaaacttaa atgtctgata atggacccca

27661 aaatcagcga aatgcactcc gcattacgtt tggtggaccc tcagattcaa ctggcagtaa

27721 ccagaatggt ggggcgcgat caaaacaacg tcggccccaa ggtttaccca ataatactgc

27781 gtcttggttc accgctctca ctcaacatgg caaggaagac cttaaattcc ctcgaggaca

27841 aggcgttcca attaacacca atagcagtcc agatgaccaa attggctact accgaagagc

27901 taccagacga attcgtggtg gtgacggtaa aatgaaa

[gap 174 bp] [Expand Ns](https://www.ncbi.nlm.nih.gov/nuccore/ON151370.1?expand-gaps=on)

28112 cttcctcaa ggaacaacat tgccaaaagg

28141 cttctacgca gaagggagca gaggcggcag tcaagcctct tctcgttcct catcacgtag

28201 tcgcaacagt tcaagaaatt caactccagg cagcagtaaa cgaacttctc ctgctagaat

28261 ggctggcaat ggcggtgatg ctgctcttgc tttgctgctg cttgacagat tgaaccagct

28321 tgagagcaaa atgtctggta aaggccaaca acaacaaggc caaactgtca ctaagaaatc

28381 tgctgctgag gcttctaaga agcctcggca aaaacgtact gccactaaag catacaatgt

28441 aacacaagct ttcggcagac gtggtccaga acaaacccaa ggaaattttg gggaccagga

28501 actaatcaga caag

//