



Nucleotide

GenBank

## SARS coronavirus Urbani, complete genome

GenBank: AY278741.1

[FASTA](#) [Graphics](#)

Go to:

LOCUS AY278741 29727 bp RNA linear VRL 25-JUL-2016

DEFINITION SARS coronavirus Urbani, complete genome.

ACCESSION AY278741

VERSION AY278741.1

KEYWORDS .

SOURCE SARS coronavirus Urbani

ORGANISM [SARS coronavirus Urbani](#)

Viruses; Riboviria; Orthornavirae; Pisuviricota; Pisoniviricetes; Nidovirales; Cornidovirineae; Coronaviridae; Orthocoronavirinae; Betacoronavirus; Sarbecovirus.

REFERENCE 1 (bases 1 to 29727)

AUTHORS Rota,P.A., Oberste,M.S., Monroe,S.S., Nix,W.A., Campagnoli,R., Icenogle,J.P., Penaranda,S., Bankamp,B., Maher,K., Chen,M.H., Tong,S., Tamin,A., Lowe,L., Frace,M., DeRisi,J.L., Chen,Q., Wang,D., Erdman,D.D., Peret,T.C., Burns,C., Ksiazek,T.G., Rollin,P.E., Sanchez,A., Liffick,S., Holloway,B., Limor,J., McCaustland,K., Olsen-Rasmussen,M., Fouchier,R., Gunther,S., Osterhaus,A.D., Drosten,C., Pallansch,M.A., Anderson,L.J. and Bellini,W.J.

TITLE Characterization of a novel coronavirus associated with severe acute respiratory syndrome

JOURNAL Science 300 (5624), 1394-1399 (2003)

PUBMED [12730500](#)

REFERENCE 2 (bases 1 to 29727)

AUTHORS Monroe,S.S.

CONSRM CDC SARS Coronavirus Sequencing Team

TITLE SARS coronavirus (SARS-CoV), Urbani strain

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 29727)

AUTHORS Bellini,W.J., Campagnoli,R.P., Icenogle,J.P., Monroe,S.S., Nix,W.A., Oberste,M.S., Pallansch,M.A. and Rota,P.A.

TITLE Direct Submission

JOURNAL Submitted (17-APR-2003) Division of Viral and Rickettsial Diseases, Centers for Disease Control and Prevention, 1600 Clifton RD, NE, Atlanta, GA 30333, USA

FEATURES Location/Qualifiers

source 1..29727  
/organism="SARS coronavirus Urbani"  
/mol\_type="genomic RNA"  
/strain="Urbani"  
/db\_xref="taxon:[228330](#)"  
/cell\_line="Vero"

[5' UTR](#)  
[CDS](#) 1..264  
join(265..13398,13398..21485)  
/ribosomal\_slippage  
/note="ORF 1ab"  
/codon\_start=1  
/product="nonstructural polyprotein pp1ab"  
/protein\_id="AAP13442.1"

/translation="MESLVLGVNEKTHVQLSLPVLQVRDVLVRFVGFSDSVEEALSEARE  
HLKNGTCGLVELEKGVLPQLEQPYVFIKRS DALSTNHGHKVVVELVAEMDGIQYGRSGI  
TLGVLVPHVGETPIAYRNVLLRKNKNGKAGGHSYGD LKSYDLGDELGTDP IEDYEQN  
WNTKHGSGALRELTRELN GAVTRYVDN NFCGPDGYPLDCIKDFLARAGKSMCTLSEQ  
LDYIESKRGVYCCRRDHEHEIAWFTE RSDKSYEHQTPFEIKS AKKFDTFKGECPKFVFP

LNSKVKVIQPRVEKKKTEGFMGRIRSYPVSPQECNMHLSTLMKCNHCDEVSWQTC  
DFLTKATCEHCGETENLVEIEGPTTCGYLPTNAVVKMPCACQDPEIGPEHSVADYHNHNSN  
IETRLRKGGRTRCFGGCVFAYVGCYNKRAYWVPRASADIGSGHTGITGDNVETLNEDL  
LEILSRERVNINIVGDFHLENEVAIILASFASSTAFIDTIKSLDYKSFKTIVESCNGN  
YKVTGKGPVKGAWNIGQORSVLTPLCGFPSQAAGVIRSI FARTLDAANHSIPDLORAA  
VTILDGISEQSLRLVDAMVYTSDLLTNSVIMAYVTGGLVQQTQSWLSNLLGTTVEKL  
RPIFEWIEAKLSAGVEFLKDAWEILKFLITGVFDIVKGOIQVASDNKDCVKCFIDVV  
NKALEMCIDQVTTAGAKLRSLNLGEVFI AQSKGLYRQCIRGKEQLQLLMLPKAPKEVT  
FLEGSDHDTVLTSEEVVKNNGELEALETPVDSFTNGAIVGTPVCVNGMLLEIKDKEQ  
YCALSPGLLATNNVFRLLKGGAPIKGVTFGEDTVWEVQGYKNVRITFELDERVDKVLNE  
KCSVYTVESGTEVTEFACVVAEAVVKTLPVSDLLTNMGIDLDEWSVATFYLFDDAGE  
ENFSSRMYCSFYPPDEEEEDDAECEEEI D ETCHEHYGTEDDYQGLPLEFGASAETVR  
VEEEEEEEDWLDTTTEQSEIEPEPEPTPEEPVNFQFTGYLKLTDNVAIKCVDIVKEAQA  
NPMVIVNAANIHLKHGGVAGALNKATNGAMQKESDDYIKLNGPLTVGGSCLLSGHNL  
AKKCLHVVPNLNAGEDIQLLKAAYENFN SQDILLAPLLSAGIFGAKPLQSLQVCVQT  
VRTQVYIAVNDKALYEQVMDYLDNLKPRVEAPKQEEPPNTEDSKTEEKSVVQKPV DV  
KPKIKACIDEVTTTTLEETKFLTNKLLLFADINGKLYHDSQNMLRGEDMSFLKADAPYM  
VGDVITSGDITCVVIPSKKAGGTTEMLSRALKKVPVDEYITTYPGQGCAGYTL EEAKT  
ALKKCKSAFYVLPSEAPNAKEELGTVSWNLREMLAHAEERKLMPI CMDVRAIMATI  
QRKYKGIKI QEGIVDYGVRFFFYTSKEPVASIIITKLSLNEPLVTMPIGYVTHGFNLE  
EAARCMSRLKAPAVVSVSSPDVAVTTYNGYLTSSSKTSEEHFVETVSLRAGSYRDSYSG  
QRTELGVFELKRGDKIVYHTLESPVEFHLDGEVLSLDKLSLLSLREVKTIKVFTTVD  
NTNLHTQLVDMSTYGOQFGPTYLDGADVTKIKPHVNHEGKTFVFLPSDDTLRSEAFE  
YYHTLDESFLGRYMSALNHTKWKFPQVGLT SIKWADNNCYLSSVLLALQLEVKFN  
APALQEAAYRARAAGDAANFCALILAYSNKTVGELGVDRETMTHLLQHANLESAKRVLN  
VVCKHCQKQTTTTLGTVEAVMYMGLT SYDNLKTVSIPCVCGRDATQYLVQESSFVMM  
SAPPAEYKLGQGTFLCANEYTGNYQC GHYTHITAKETLYRIDGAHLTKMSEYKGPVTD  
VFYKETS YTTTTKPVSYKLDGVTYTEIEPKLDGYKKNAYYTEQPIDLVPTQLPNA  
SFDNFKLTC SNTKFADDLNQMTGFTK PASRELSVTFPDLNGDVVAIDYRHYSASFKK  
GAKLLHAKPIVWHINQATTKTTTFKPN TWCLRCLWSTKPVDTSNSEVLEHEDTQGMNLL  
ACESQOPTSEEVNPTIQKEVIECDVKTTEVVG NVILKPSDEGVKVTQELGHEDLMA  
AYVENTSITIKKPNELSLALGLKTIATHGIAAINSVPWSKILAYVKPFLGQAAITTSN  
CAKRLAQRVFNMYMPYVFTLLFQLCTFTKSTNSRIRASLPTTIAKNSVKSVAKLCLDA  
G1NYVKSPKF SKLFTIAMWLLLLSICLGLSICVTAAFVGLLSNFGAPSYCNGVRELYL  
NSSNVTMTDFCEGSFPCSI CLSGLDLSYPALETIQVTISSYKLDL TILGLAAEWL  
AYMLFTKFFYLLGLSAIMQVFFGYFASHFISNSWLMWFIISIVQMAPVSAMVRYIFF  
ASFYYIWKSYVHIMDGTSTSTCMCYKRN RATRVECTTIVNGMKRSFYVYANGGRGFC  
KTHNWNCLNCDTFTGSTFISDEVARDSLQFKRPI NPTDQSSYIVDSVAVKNGALHL  
YFDKAGQKTYERHPLSHFVNLNLRANNTKGS LPINVIVFDGKSKCEDASAKSASVYY  
SQLMCQPIILLDQVLVSDVGDSTEVSVKMF DAYVDTFSATFSVPMKELKALVATAHSE  
LAKGVALDGVSTFVSAARQGVVDVDTVDKDVIECLKLSHSDLEVTGDSCNNFMLTY  
NKVENMTPRDLGACIDCNARHINAQVAKSHNVSLIWNVKDYMSLSEQLRKQIRSAAKK  
NNIPFRLTCATTRQVVNVITTKISLGGKIVSTCFKLMKATLLCVLAALVCIYI VMPV  
HTLSIHGDYNEIIGYKAIQDGVTRDIISTDDCFANKHAGFDWFSQRGGSYKNDKSC  
PVVAAITREIGFIVPGLPGTVLRAINGDFLHFLPRVFSAVGNICYTPSKLIEYSDFA  
TSACVLA AECTIFKDAMGKPVPCYD TNLLEGSISYSELRPDTRYVLMGDSIIQFPNT  
YLEGSVRVVTTFDAEYCRHGT CERSEVGI CLSTSGRWLVNNEHYRALSGVFCGVDAMN  
LIANIFTPLVQPVGALDVSASV VAGGIIAILVTCAAYYFMKFRRVFGEYNHVVAANAL  
LFLMSFTILCLVPAYSFLPGVYSVFYLYLTFYFTNDVSFLAHLQWFAMFSPIVPFWIT  
AIYVFCISLKHCHWFNNYLKRKVMFNGVTFSTFEEAALCTFLLNKEMYLKLRSETLL  
PLTQYNRYLALYNKYK YFSGALDTSYREAAACHLAKALNDFSNSGADVLYQPQTSI  
TSAVLQSGFRKMAFP SGKVEGCMVQVTCGTTTTLNLGLWLD DTVYCPRHVIC TAEDMLNP  
NYEDLLIRKSNHSFLVQAGNVQLRVIGHSMQNC LLRLKVDTSNPKTPKYKFVRIQPGQ  
TFSVLACYN GSPSGVYQCAMPNHTIKGSFLNGSCGSVGFNIDYDCVSFCYMHMHEL P  
TGVHAGTDLEKFGYGFVDRQTAQAAGDTTITLNLVLAWL YAAVINGDRWFLNRFTTT  
LNDFNLVAMKYNYEPLTQDHVDILGPLSAQTGIAVLDMCAALKELLQNGMNTILGS  
TILEDEFTPFDDVVRQCSGVT FQGKFKKIVKGTTHHMLL TFLTSLLLILVQSTQWLSFFF  
VYENAFLPFTLGMATAACAMLLVKKHAF LCLFLLPSLATVAYFNMVYMPASWVMRI  
MTWLELADTSLSGYRLKDCVMYASALVLLILMTARTVYDDAARRVWTLMMNITLVYKV  
YYGNALDQAIMWALVISVTSNYSGVVTTIMFLARAI VFCVEYYP LFIITGNLQCI  
MLVYCF LGYCCCYFGLFCLLNRYFRLTLGVYDYL VSTQEFRYMNSQGLLPPKSSIDA  
FKLNKLLGIGGKPCIKVATVQSKMSDVKCTSVVLLSVLQQLRVESSK LWAQCQVLH  
NDILLAKD TTEAF EKMVSLLSVLLSMQGAVDINRLCEEMLDNRATLQAIASEFSSLPS  
YAAAYATAQEAYEQAVANGDSEVLLK LKLSLNVAKSEFDRDAAMQRKLEK MADQAMTQ  
MYKQARSEDKRAKVT SAMQTM LFTMLRKLNDALNNI INNARDGCVPLNI IPLTTAAK  
LMVVVDPYGYKNTCDGNFTFYASALWEIQVVDADSKIVQLSEINMDNSPNLAWPLI  
VTALRANS AVKLNNE LSPVALRQMSCAAGTTQTACTDDNALAYNNSKGGFRVLALL  
SDHQDLKWARFPKSDGTGTIYTELEPPCRFVTDTPKGPVKYLYFIKGLNNLNRGMVL  
GSLAATVRLQAGNATEVPANSTVLSFCAFAVDPAKAYKDYLASGGQPITNCVKMLC TH  
TGTGQAITVTP EANMDQESFGGASCCLYCRCHIDHPNPKGFCDLKGYVQIPTTCAND  
PVGFTLRNTVCTVCGMWKYGCS CDQLREPLMQSADASTFLNRVCGVSAARLT PCGTG  
TSTDVVYRAFDIYNEKVAGFAKFLKTNCCR FQEKDEEENLLDSYFVVKRHTMSNYQDE  
ETIYNLVKDCPAVAHVDFFKFRVDGDMVPHISRQRLTKYTMADLVYALRHDFEGNCDT  
LKEILVTYNCDDDYFNKKDWYDFVENPDILRVYANLGERVRQSLKTVQFC DAMRDA  
GIVGVLTLDNQDLNGNWDYDFGDFVQVAPGCGVP IVDSYSSLMPILTLTRALAAESHM  
DADLAKPLTKWDLKYDFTEERLCLFDRYFKYWDQTYHPNCINCLDDRCILHCAFNV  
LFTSVFPPTSFGPLVRKIFVDGVPFVSTGYHFRELGVVHNQD VNLHSSRLSFKELLV

YAADPAMHAASGNLLLDKRTTCFVSAALTNNVAFQTVKPGNFNKDFYDFAVSKGFFKE  
GSSVELKHFFDAQDGNAAISDYDYRYNLPMTCDIRQLLFFVEVVDKYDFCYDGGCIN  
ANQIVNMLDKSAGFPFNKWKARLYYDSMSYEDQDALFAYTKRNVIPITIQMNLKYA  
ISAKNRARTVAGVVICSTMTNRQFHQKLLKLSIAATRATVIGTSKFGYGGWHNMLKTV  
YSDVETPHLMGWDYPKCDRAMPNMLRIMASLVLARKHNTCCNLSHRFYRLANCAQVL  
SEMVMCGGSLYVKPGGTSDDATTAYANSVFNICQAVTANVALLSTDGNKIADKYVR  
NLQHRLYECLYRNRDVEHFVDEFYAYLRKHFSMMILSDDAVVCYNSNYAAQGLVASI  
KNFKAVLYYQNNVFMSEAKCWTETDLTKGPHEFCSQHTMLVKQGDYVYLPYDPSRI  
LGAGCFVDDIVKTDGTLMIERFVSLAIDAYPLTKHPNQEYADVFLHYLQYIRKLDHEL  
TGHMLDMYSVMLTNDNTSRYWEPEFYEAMYTPHTVLAQVACVLCNSQTSRLRCGACIR  
RPFLLCKCCYDHYVISTSHKLVLSVNPVYCNAPGCDVTDVTQLYLGGMSYYCKSHKPP  
SFPLCANGQVFLGYKNTCVGSDNVTDFNAIATCDWTNAGDYILANTCTERLKLFAAET  
LKATEETFKLKSYGIATVREVLSDRELHLSWEVKGPRPLNRNYVFTGYRVTKNSKVQI  
GEYTFEKGDYGDAVVYRGTTTYKLVNGDYFVLTSHYVMPLSAPTLVPEHYVITGLY  
PTLNIISDEFSSNVANYQKVGMMQYKSTLQGGPPTGKSHFAIGLALYPSARIVYACSH  
AAVDALCEKALYLPIDKCSRIIPARARVECFDKFKNSTLEQYVFCVNLPETTAD  
IVVFDEISMATNYDLVSVNARLRAKHVYVIGDPAQLPAPRTLLTKGTEPEYFNSVCR  
LMKTIGPDMFLGTCRRCPAEIVDVSALVYDNKLAHKDKSAQCCKMFYKGVITHDVS  
SAINRPQIGVVREFLTRNPAWRKAVFISPYNSQNAVASKILGLPTQTVDSSQGEYDY  
VIFTQTTETAHSCNVNRFNVAITRAKIGILCIMSDDRDLYDKLQFTSLEIPRRNVATLQ  
AENVTLGFKDKCSKIIITGLHPTQAPTHLSVDIKFKTEGLCVDIPGPKDMTYRRLISMM  
GFKMNYQVNGYPMNFITREEAIRHVRWIGFDVEGCHATRDAVGTNLPLQLGFSTGVN  
LVAVPTGYVDTENNETFRVNAKPPGQDFKHLIPLMYKGLPWNVVRIKIVQMLSDTL  
KGLSDRVVFLWAHGFEKSMKYFVKIGPERTCCLCDKRACTFSTSSDYACWNHVS  
FDYVYVNFHIDVQGWFTGNLQSNHDQHCQVHGNHVASCDAIMTRCLAVHECFVKRV  
DWSVEYPIIGDELRVNSACRQVHMVKSALLADKFPVLDHIGNPKAICKVPAEVEV  
KFYDAQPCSDKAYKIEELFYSYATHHDKFTDGVCLFWNCNVDRYPANAIVCRFDRVL  
SNLNLPGCDGGSLYVKNHAFHTPAFDKSAFTNLKQLPFFYYSDSPCESHGKQVSDID  
YVPLKSATCITRCNLGGAVCRHHANEYRQYLDAYNMMISAGFSLWIYKQFDYTNLWNT  
FTRLQSLENVAYNVVKNKGHFDGHAGEAPVSIINNAVYTKVDGIDVEIFENKTTLPVNV  
AFELWAKRNIKVPPEIKILNMLGVDIAANTVIWDYKREAPAHVSTIGVCTMTDIAKPP  
TESACSSLTVLFDGRVEGQVDFRNARNGVLITEGVSVKGLTPSKGPAQASVNGVTLIG  
ESVKTQFNFKKVDGIIQQLPETYFTQSRDLEDFKPRSQMETFLELAMDEFIORYKL  
EGYAFHEIYVYGDVSHGQLGGLHLMIGLAKRSQDSPLKLEDFIPMDSTVKNYFITDAQT  
GSSKCVCSVIDLDDFVEIISQDLSVISKVVKVTDYAEISFMLWCKDGHVETFPY  
KLQASQAWQPGVAMPNLKMQRMLEKCDLQNYGENAVIPKGINMNVAKYQLCQYLN  
TLTLAVPYNMRLIHFAGSDKGVAPGTAVLRQWLPTGTLVDSDLNDFVSDADSTLIG  
DCATVHTANKWDLIISDMYDPRTKHVTKEKESGFFTYLCEGFKQKALGGSIAVKI  
TEHSMADLYKLMGHFSWTAFTVTVNASSSEAFIIGANYLGKPKQIDGYTMHANYI  
FWRNTNPIQLSSYSLFDMSKFPLKLRGTAVMSLKENQINDMIYSLLEKGRLLIRENRR  
VVVSSDILVNN"

265..13413

/note="ORF 1a"

/codon\_start=1

/product="nonstructural polyprotein pp1a"

/protein\_id="AAP13439.1"

/translation="MESLVLGVNEKTHVQLSLPVLQVRDVLVVRGFGDSVEEALSEARE  
HLKNGTCGLVELEKGVLPQLEQPYVFIKRSDALSTNHGKVVVLAEMDGIQYGRSGI  
TLGVLVPHVGETPIAYRNVLLRKNNGKAGGHSYGDIDLKSVDLDELGTDPIEDYEON  
WNTKHGSGALRELTRELNGAVTRYVDNFCGPDGYPLDCKIDFLARAGKSMCTLSEQ  
LDYIESKRGVYCCRDHEHEIAWFTERSCKSYEHQTPFEIKSAKKFDTFKGCECPKVFVP  
LNSKVKVIQPRVEKKTEGFMGRIRSVYVPAQPCENMMHLSLTMKCNHCDEVSWQTC  
DFLKATCEHCGTENLVIEGPTTCGYLPTNAVVKMPCPACQDPEIGPEHSVADYHNHNS  
IETRLRKGGRTRCFGGCVFAYVGCYNKRAYWVPRASADIGSGHTGITGDNVETLNEDL  
LEILSRERNINIVGDFHLNNEVAIILASFSASTSAFIDTIKSLDYKSFKTIIVESCGN  
YKVTGKPKYKAWNIGQQRSVLTPLCGFPQAAGVIRSIFFARTLDAANHSIPDLQRAA  
VTILDGISEQSLRLVDAMVYTSDLLTNSVIIMAYVTGGLVQQTQSLNLLGTTVEKL  
RPIFEWIEAKLSAGVEFLKDAWEILKFLITGVFDIVKQIQVASDNIKDCVKCFIDVV  
NKALEMCIQVTTIAGAKLRSLNLGEVFIQAQSKGLYRQIRGKEQLQLMPLKAPKEVT  
FLEGDSDHTVLTSEEVVLKNGELEALETPVDSFTNGAIVGTPVCVNGMLLEIKDKEQ  
YCALSPGLLATNNVFRLLKGGAPIKGVTFGEDTVWEVQGYKNVRITFELDERVDKVLNE  
KCSVYTVESGTEVTEFACVVAEAVVKTLPVSDLLTNMGIDLDEWSVATFYLFDDAGE  
ENFSSRMYSFYPPDEEEEDDAECEEIEDETCEHEYGTEDDYQGLPEFGASAEVTR  
VEEEEEEEDWLDDETEQSEIEPEPEPTPEEPVQFTGYLKLTDNVAIKCVDIVKEAQS  
NPMVIVNAANIHLKHGGVAGALNKATNGAMQKESDDYIKLNGPLTVGGSCLLSGHNL  
AKKCLHVVGNLAGEDIQLLKAAYENFNSQDILLAPLLSAGIFGAKPLQSLQVCVQT  
VRTQVYIAVNDKALYEQVMDYLDNLKPRVEAPKQEEPPNTEDSKTEEKSVVQKPVVD  
KPKIKACIDEVTTTLEETKFLTNKLLLFADINGKLYHDSQNMRLRGEDMSFLEKADAPY  
VGDVITSGDITCVVIPSCKAGGTTEMLSRALKKVPVDEYITTYPGCAGYLEEAKT  
ALKKCKSAFYVLPSEAPNAKEEILGTVSWNLREMLAHAEETRKLMPICMDVRAIMATI  
QRKYGKIQIEGIVDYGVRFFYTSKEPVASIIITKLNLSNEPLVTMPIGYVTHGFNLE  
EAARCMRSLKAPAVVSVSPDAVTTYNGYLTSSSKTSEEHFVETVSLAGSYRDSYSG  
QRTELGVEFLKRGDKIVYHTLESPEVFHLDGEVLSLDKLSLLSREVKTIKVFVTTVD  
NTNLHTQLVDMSTYQQFGPTYLDGADVTKIKPHVNHGKTFVFLPSDDTLRSEAFE  
YYHTLDESFLGRYMSALNHTKWKFPQVGLTSLKADNCCYLSVLLALQLEVKFN  
APALQEAAYRARAGDAANFCALILAYSNKTVGELGDVRETMTLLQHANLESKRVLN  
VCKHCGQKTTTLTGVEAVMYMGLTSDNLKTGVSIPCVCGRDATQYLVQGESEFVMM

SAPPAEYKLGQGTFLCANEYTGNYQCGHYTHITAKETLYRIDGAHLTKMSEYKGPVTD  
VFYKETSYYTTIKPVSYKLDGVTYTEIEPKLDGYKKDNAYYTEQPIDLVPTQPLPNA  
SFDNFKLTCNSNFKADDLNQMTGFTKPARRELSVTFPPDLNGDVVAIDYRHYASAFKK  
GAKLLHKPIVWHINQATTKTTFKPNWCLRCLWSTKPVDTSNSEFEVLAVEDTQGMNDL  
ACESQOPTSEEVVENPTIQKEVIECDVKTTTEVVGNVILKPSDEGVKVTQELGHEDLMA  
AYVENTSITIKKPNELSLALGLKTIATHGIAAINSVPWSKILAYVKPFLGQAAITTSN  
CAKRLAQRVFNMYVFTLLFQLCTFTKSTNSRIRASLPTTIAKNSVKSVAKLCLDA  
GNYVKSXPKFSLFTIAMWLLLLSICLGLSICVTAAFVGLLSNFGAPSYCNGVRELYL  
NSSNVTMTDFCEGSPCSICLSGLDSDSYPALLETIQVTISSYKLDLTILGLAAEVL  
AYMLFTKFFYLLGLSAIMQVFFGYFASHFISNSWLMWFIISIVQMAPVSAMVMYIFF  
ASFYYIWKSYVHMDGCTSTCMCYKRNRRATRVECTTIVNGMKRSFYVYANGGRGFC  
KTHNWNCLNCDFTGTFISDEVARDSLQFKRPINPTDQSSYIVDSVAVKNGALHL  
YFDKAGQKTYERHPLSHFVNLNLRANNTKGSLPINVIIVFDGSKKCEDASASKASVYY  
SQLMCQPILLDDQVLVSDVGDSTEVSVKMFDAYVDTFSATFSVPMKELKALVATAHSE  
LAKGVALDGLVSTFVSAARQGVVDVDTKDVIECLKLSHSDLEVTGDCNNFMLTY  
NKVENMTPRDLGACIDCNARHINAQVAKSHNVSLIWNVKDYMSLSELRKQIRSAAKK  
NNIPFRLTCATTRQVNVNITTKISLGGKIVSTCFKMLKATLLCVLAALVCYIVMPV  
HTLSIHGDYNEIGYKAIQDGVTRDIISTDDCFANKHAGFDWFSQRGGSYKNDKSC  
PVVAAITREIGFIVPGLPGTVLRAINGDFLHFLPRVFSAVGNICYTPSKLIEYSDF  
TSACVLAECTIFKDMAGKPPVPCYDNTLLEGSISYSELRPDTRYVLMGDSIIQFPNT  
YLEGSVRVVTTFDAEYCRHGTCESEVIGLSTSGRWVLNNEHYRALSGVCGVDAMN  
LIANIFTPLVQPGALDVSASVAGGIIAILVTCAAYFMKFRVFGYENHVAANAL  
LFLMSFTILCLVPAYSFLPGVYSVFLYLTYFTNDVSFLAHLQWAMFSPVFPWIT  
AIYVFCISLKHCHWFFNNYLKRVFMFNGVTFSTFEEAALCTFLLNKEMYLKLRSETLL  
PLTQYNRYLALYNKYKYSFGALDTSYREAAACHLAKALNDFNSGADVLYQPPQTSI  
TSAVLQSGFRKMAFSPGKVEGCMVQVTCGTTTLNGLWLDDEVYCPRHVICTAEDMLNP  
NYEDLLIRKSNHSFLVQAGNVQLRVIGHSMQNCLLRLKVDTSNPKTPKYKFRVRIQPGQ  
TFSVLACYNGSPSGVYQCAMPNHTIKGSFLNGSCGSGVFNIDYDCVSFCYMHMELP  
TGVHAGTLEKGFYGFVDRQTAQAAGDTTITLNLVLAWYAAVINGDRWFLNRFTTT  
LNDFNLVAMKYNYPELTDHVDILGPLSAQGTIAVLDMAALKELLQNGMNGRTILGS  
TILEDEFTPFVDRQCSGVTFOGKFKKIVKGTTHHMLLTLTSLLLVQSTQWLSFFF  
VYENAFPLPTLGMATAAACAMLLVKHKHAFCLFLLPSLATVAYFNMVYMPASWVMRI  
MTWLELADTSLSGYRLKDCVYASALVLLILMTARTVYDAAARRVWTLMNVTLVYK  
YYGNALDQAI SMWALVISVTSNYSVGVVTTIMFLARAI VFCVEYYP LFLITGNTLQCI  
MLVYCFGLYCCCYFGLFCLLNRYFRLTLGVYDYLSTQEFYRMYNSQGLLPPKSSIDA  
FKLNKILLGIGGKPCIKVATVQSKMSDVKCTSVVLLSVLQQLRVESSEKLAQCQLH  
NDILLAKDTTEAFKEMVSLVLLSMQGAVDINRLCEEMLDNRATLQAIASEFSSLPS  
YAAAYATAQEAYEQAVANGDSEVLKLLKKS LNVAKSEFDRDAAMORKLEKMAQAMTQ  
MYKQARSEDKRAKVTSAMQTMFTMLRKLNDALNNIINNARDGCVPLNIIPLTTAAK  
LMVVVDPDYGYKNTCDGNTFTYASALWEIQVVDADSKIVQLEINMDNSPNLAWPLI  
VTALRANSVAVKLNQNELSPVALRQMSCAAGTTQTACTDDNALAYNNKSGGRFVLALL  
SDHQDLKWARFPKSDGTGTIYTELEPPCRFVTDTPKGPVKYLYFIKGLNNLRGMVL  
GSLAATVRLQAGNATEVPANSTVLSFCFAVDPKAYKYDYLASGGQPIITNCVKMLCTH  
TGTGQAITVTPCANMDQESFGGASCCLYCRCHIDHPNPKGFDLKGKVVQIPTTCAND  
PVGFTLRNTVCTVCGMWKGYGCSQDLREPLMQSADASTFLNGFAV"

<13398..21485

/note="ORF 1b; expressed via predicted -1 ribosomal frameshift"

/codon\_start=1

/product="nonstructural polyprotein"

/protein\_id="AAP13440.1"

/translation="RVCVSAARLTPCGTGTSTDVVYRAFDIYNEKAVGFAKFLKTCN  
CRFQEKDEEGLLDSYFVVKRHTMSNYQHEETIYNLVKDCPAVAVHDFKFRVDGDMV  
PHISRQLTKYTMADLVYALRHDFEGNCDLKEILVTYNCDDDYFNKWDYDFVENP  
DILRVYANLGERVRSLLKTVQFCAMRDAGIVGVLTDNQDLNGNWYDFGDFVQVAP  
GCGVPIVDYSYSLMPLILTLTRALAAESHMDADLAKPLIKWDLKDYDFTEERLCLFDR  
YFKYWDQTYHPNCINLDDRCILHCANFNVLVSTVFPPTSFGPLVRKIFVDGVPPFVVS  
TGYHFRELGVVHNDVNLHSSRLSFKELLVYAADPAMHAASGNLLLDKRTTCFVSAAL  
TNNVAFQTVKPGFNKDFYDFAVSKGFKEGSSVELKHFFFAQDGNAAISDYDYRYRN  
LPTMCDIRQLLFVVEVVDKYFDCYDGGCINANQVIVNNDKSAGFPFNKWKARLYDY  
SMSYEDQDALFAYTKRNVIPITITQMNLKYAISAKNRARTVAGVSICTMTNRRQFHQKL  
LKSIAATRGATVYIGTSKFGWNNMLKTVYSVDVETPHLMGWDPKCDRAMPNMLRIM  
ASLVLARKHNTCCNLSHRFYRLANEAQVLSMVMCGGSLYVKGPGTSSGDATTAYAN  
SVFNICQAVTANVALLSTDGKIAADKYVRNLQHRLEYECLYRNRVDHEFVDEFYAYL  
RKHFMSLSDDAVVCYNSNYAAQGLVASIKNFKAVLYYQNNVFMSEAKMWTETDLTK  
GPHEFCSQHTMLVKQDDYVYLPYPDPSRILGAGCFVDDIVKTDGTLMIERFVSLAID  
AYPLTKHPNQYADVHFLYLQYIRKLDHDELGHMLDMYSVMLTNDNTSRYWEPEFYEA  
MYTPHTVLQAVGACVLCNSQTSLRGACIRRPFLCCKCCYDHVISTSHKLVLSVNPYV  
CNAPGCDVTDVTLQYLGGMSYCKSHKPPISFPLCANGQVFLYKNTCVGSDNVDFN  
AIATCDWTNAGDYILANTCTERLKLFAAETLKATEETFKLSYGIATVREVLSDRELHL  
SWEVKGPRPPLNRNYVFTGYRVTKNSKVQIGEYTFEKGDYDVAVVYRGTTTYKLVNGD  
YFVLTSHTVMPLSAPTLVQEHYVRITGLYPTLNIISDESSNVANYQKVMQKYSTLQ  
GPPGTGKSHFAIGLALYPSARIVYACSHAAVDALCEKALKYLPIDKCSRIIPARAR  
VECFDKFVNSTLEQYVFCVNALPETTADIVVFEISMATNYDLVSVNARLRAKHYY  
YIGDPAQLPAPRTLLTKGTLEPEYFNSVCRMLKTIIGPDMFLGTCCRCPAEIVDVTYSAL  
VYDNKLAHKDKSAQCFKMFYKGVITHDVSAINRPQIGVREFLTRNPAWRKAVFIS  
PYNQNAVASKILGLPTQTVDSQGEYDYVIFTQTTETAHSCNVNRFNVAITRAKIG

CDS

ILCIMSDRDLYDKLQFTSLEIPRRNVATLQAEVNTGLFKDCSKIITGLHPTQAPTHLS  
VDIKFKTEGLCVDIPGPKDMTYRRLISMGMFKMNYQVNGYPNMFITREEAIRHVRAW  
IGFDVEGKREAPDVGNTLPLQLGFSTGVNLVAVPTGYVDTENNTEFTRVNAKPPPGD  
QFKHLIPLMYKGLPWNVVRKIVQMLSDTLKGLSDRVVFLWAHGFELTSMKYFVKIG  
PERTCCLCDKCRATCFSTSSDYACWNHNSVGFYVYVNFMDVQVQWGTGNLQSNHDQH  
CQVHGNNAHVASCDAIMTRCLAVHECFVKRVDWSVEYPIIGDELRVNSACRKYQHMVVK  
SALLADKFPVLDHIGNPKAIKCVPAEVEWKFYDAQPCSDKAYKIEELFYSYATHHDK  
FTDGVCLFVNCNVDRYPANAIIVCRFDTRVLSNLSLPGCDGGSLYVNHAFHTPAFDKS  
AFTNLKQLPFFYSDSPCESHGKQVVDIDYVPLKSATCITRNLGGAVCRHHANEYR  
QYLDAYNMIISAGFSLWIYKQFDYTNLWNTFTRLQSLENVAYNVVNGHFDGHAGEAP  
VSIINNAVYTKVDGIDVEIFENKTTLPVNVAFELWAKRNIKPVPEIKLNLDGVDIAA  
NTVIWDYKREAPAHVSTIGVCTMTDIAKKPTESACSSLTVLFDGRVEGQVDFRNARN  
GVLITEGSKGLTPSKGPAQASVNGVTLIGESVKTQFNFKVVDGIIQQLPETYFTQS  
RDLEDFKPRSQMETDFLELAMDEFIQRKLEGYAFEHIVYGDFFSHGQLGGLHLMIGLA  
KRSQDSPLKLEDFIPMDSTVKNYFITDAQTGSSKCVCSVIDLLDDFVEIISQDLSV  
ISKVVKVTIDYAEISFMLWCKDGHVETFPKQASQAWQPGVAMPNLKMQRMLEK  
DLQNYGENAIVPKGIMMNVAKYQQLCQYLNTLTLAVPYNMRVIFHGAGSDKGVAPGTA  
VLRQWLPPTGTLVDSLDLNFVSDADSTLIGDCATVHTANKWDLIISDMYDPRTKHVTK  
ENDSKEGFFTYLTCGFIKQKALALGGSIAVKITEHSWNADLYKLMGHFSWMTAFVTVNVA  
SSSEAFILGANYLGGKPEQIDGYTMHANYIFWRNTNPIQLSSYFLDMSKFPKLRGT  
AVMSLKENQINDMIYSLLEKGRLLIRENNRVVSSDILVNN"

CDS

21492..25259  
/note="surface spike glycoprotein"  
/codon\_start=1  
/product="S protein"  
/protein\_id="AAP13441.1"  
/translation="MFIFLLFLTLTSGSDLDRCTTFDDVQAPNYTQHTSSMRGVYYPD  
EIFRSDTLYLTQDLFLPFYSNVTGFHTINHTFGNPVFPKDGIFYFAATEKSNVVRGW  
FGSTMNKSQSVIIINNSTNVVIRACNFELCDNPFFAVSKPMGTQHTMIFDNFNCT  
FEYISDAFSLDVSEKSGNFKHLREFVFNKDKGFLYVYKGYQPIDVVRDLPSGFNTLKP  
IFKPLGINITNFRAILTAFAQDIWGTSAAYFVGYLKPTTFMLKYDENGITIDAV  
DCSQNPLAELKCSVKSFEIDKGIYQTSNFRVVPVSGDVVRFNITNLCPFGVEVFNATKF  
PSVYAWERKKISNCVADYSVLVNSTFFSTFKCYGVSATKLNLDLCSNVYADSFFVVKGD  
DVRQIAPGQTVIADYNYKLPDDFMGCVLAWNTRNIDATSTGNVYKYRILRHGKLRP  
FERDISNVFSPDGKPCPPALNRYWPLNDYGFYTTTIGIGYQYRNVVLSFELLNAPA  
TVCGPKLSTDLKINQCVNFNFLGTGTGVLTPSSKRFQPFQFGRDVSDFDTSVRDPK  
TSEILDISPCSFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIAHADQLTPAWRIYST  
GNVVFQTAQAGLIGAETHVDSYECDIPIGAGICASYHTVSLLRSTQSIVAYTMSLG  
ADSSIAYSNNTIAIPTNFSISITTEVMPVSMKTSVDCNMVICGDSTECANLLQYGS  
FCTQLNRALSGIAAEQDRNTREVFQVQKMYKPTLKYFGGFNSQILPDPLKPTKRS  
FIEDLLFNKVTLADAGFMKQYGECLGDINARDLCAQKFNGLTVLPLLLTDDMIAAYT  
AALVSGTATAGWTFGAGAALQIPFAMQMYRFGIGVTONVLYENQKQIANQFNKAIS  
QIQESLTTTSTALGKLQDVVNQNAQALNLTLVKQLSSNFGAISSVLDLILSRDLKVEAE  
VQIDRLITGRQLSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYH  
LMSFPQAAPHGVVFLHVTYVPSQERNFTTAPAICHEGKAYFPREGVVFVNGTSWFITQ  
RNFSPQIITDNTFVSGNCDVIGIINNTVYDPLQPELDSFKEELDKYFKNHTSPDV  
DLGDISGINASVNIQKEIDRLNEVAKNLNESLIDLQELGKYEYIKWPWYVWLFIA  
GLIAIVMVTILLCCMTCSCCLKAGACSCGSCCKFDEDDSEPVKGVKLYHT"

CDS

25268..26092  
/note="potential product, c-terminal similarity to porin"  
/codon\_start=1  
/product="protein X1"  
/protein\_id="AAP13446.1"  
/translation="MDLFMRFFTLGSIQAQPVKIDNASPASTVHATATIPLQASLPFG  
WLVIQVAFVAVFQATKIIALNKRWQLALYKGFQICNLLLLFVTIYSHLLLVAAGME  
AQFLYLIALIYFLQICINACRIIMRCWLCKWCKSKNPLLYDANYFVCWHTHNYDYCIPY  
NSVTDITVIVTEGDGISTPKLKEDYQIGGYSEDRHSGVKDYVVVHGYFTEVYQLESTQ  
ITTDGTIENATFFIFNKLKVDPPNVQIHTIDGSSGVANPAMDPIYDEPTTTTSVPL"

CDS

25689..26153  
/note="potential product"  
/codon\_start=1  
/product="protein X2"  
/protein\_id="AAP13447.1"  
/translation="MMPTTLFAGTHITMTTVYHITVSIQIQLSLLKVTAFQHNSKKT  
KLVVILRIGTQVLKTMSLYMAISPFTTSLSLHLLQTLVLMKMLHSSSLTSLKTHRM  
CKYTQSTALQELLIQWQIFMMSRRLLACLCKHKKVSTNLCTHSFRKKQVR"

CDS

26117..26347  
/note="envelope protein"  
/codon\_start=1  
/product="E protein"  
/protein\_id="AAP13443.1"  
/translation="MYSFVSEETGLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCC  
NIVNVSLVKPTVYVYSRVKNLNSSEGVPDLLV"

CDS

26398..27063  
/note="small membrane protein"  
/codon\_start=1  
/product="M protein"

/protein\_id="AAP13444.1"  
/translation="MADNGTITVEELKQLLEQWNLVIGFLFLAWIMLLQFAYSNRNRF  
LYIIKLVFLWLLWPVTLACFVLAAYRINWVTGGIAIAMACIVGLMWSYFVASFRLF  
ARTRSMWSFNPETNILLNVPLRGITVTRPLMESELVIGAVIIRGHLRMAGHPLGRCDI  
KDLPEKITVARTLSYYKLGASQVRGTDGFAAYNRYRIGNYKLNTHAGSNDNIAL  
LVQ"

CDS

27074..27265  
/note="potential product"  
/codon\_start=1  
/product="protein X3"  
/protein\_id="AAP13448.1"  
/translation="MFHLVDFQVTIAEILIIIMRTFRIAIWNLDVIISSIVRQLFKPL  
TKKNYSELDDDEEPMELDYP"

CDS

27273..27641  
/note="potential product"  
/codon\_start=1  
/product="protein X4"  
/protein\_id="AAP13449.1"  
/translation="MKIILFLTLIVFTSCELYHQECVVRGTTVLLKEPCPSGTYEENS  
PFHPLADNKFALCTCTSTHFACADGTRHTYQLRARSVSPKLFIRQEEVQQLYSPLF  
LIVAALVFLILCFIKRKTE"

CDS

27864..28118  
/note="potential product"  
/codon\_start=1  
/product="protein X5"  
/protein\_id="AAP13450.1"  
/translation="MCLKILVRYNTRGNTYSTAWL CALGKVL PFHRWHTMVQCTPNV  
TINCQDPAGGALIARCWYLHEGHQTAAFRDVVLVLRNKRTRN"

CDS

28120..29388  
/note="nucleocapsid protein"  
/codon\_start=1  
/product="N protein"  
/protein\_id="AAP13445.1"  
/translation="MSDNGPQSNQRSAPRITFGGPTDSTDNQNGGRNGARPKQRRPQ  
GLPNNTASWFTALTQHGKEELRFPRGQGVPIINTNSGPDQIGYYRRATRRVRGGDGKM  
KELSPRWYFYLLGTGPEASLPYGANKGIVVWATEGALNTPKDHIGTRNPNNAATVL  
QLPQGTTLKPGFYAEGSRGGSQASSRSSRSRGNRNSTPGSSRGNSPARMASGGGET  
ALALLLDRLNQLLESKVGKQQQGQTVTKKSAEASKKPROKRRTATKQYNVTOAFG  
RRGPEQTQGNFGDQLIRQGTDYKHWPIAQFAPSASAFFGMSRIGMEVTPSGTWLTY  
HGAIKLDDKDPQFKDNVILLNKHIDAYKTFPPTPEPKDKKKKTDEAQPLPQRQKKQPT  
VTLLPAADMDDFSRQLQNSMSGASADSTQA"  
29727

polyA\_site

ORIGIN

```
1 atattaggtt ttacctacc caggaaaagc caaccaacct cgatctcttg tagatctgtt
61 ctctaaacga actttaaaat ctgtgtagct gtcgctcggc tgcagtccta gtcacacctac
121 gcagtataaa caataataaa ttttactgtc gttgacaaga aacgagtaac tcgtccctct
181 tctgcagact gcttacgggt tcgtccgtgt tgcagtcgat catcagcata cctagggttc
241 gtcgggtgtg gaccgaaagg taagatggag agccttgttc ttggtgtcaa cgagaaaaaca
301 cacgtccaac tcagtttgc tgccttcag gttagagacg tgctagtgcg tggcttcggg
361 gactctgtgg aagagccctc atcggaggca cgtgaacacc tcaaaaatgg cacttgtggt
421 ctagtagagc tggaaaaagg cgtactgcc cagcttgaac agccctatgt gttcattaaa
481 cgttctgatg ccttaagcac caatcacggc cacaaggtcg ttgagctggt tgcagaaaatg
541 gacggcattc agtacggtc tagcgggata aactgggag tactcgtgcc acatgtgggc
601 gaaaccccaa ttgcataacc caatgttctt cttcgtgaaga acggaataaa gggagccggg
661 ggtcatagct atggcatcga tctaaagtct tatgacttag gtagcagact tggcactgat
721 ccattgaagc attatgaaca aaactggaac actaagcatg gcagtggtgc actccgtgaa
781 ctactcgtg agctcaatgg aggtgcagtc actcgtatg tgcacaaca tttctgtggc
841 ccagatgggt accctcttga ttgcatcaaa gatttctcgc cacgctcggg caagtcaatg
901 tgactctttt ccgaacaact tgattacatc gagtcaaga gaggtgtcta ctgctgccgt
961 gaccatgagc atgaatgtc ctggttact gagcgtctg ataagagcta cgagcaccag
1021 acacccttgc aaattaagag tgccaagaaa tttgacactt tcaaaaggga atgccccaaag
1081 tttgtgtttc ctcttaactc aaaagtcaaa gtcattcaac cacgtgttga aaagaaaaag
1141 actgaggggtt tcatggggcg tatacgtctc gtgtaccctg ttgcatctcc acaggaggtg
1201 aacaatatgc acttgtctac cttgatgaaa tgaatcatt gcgatgaagt ttcattggcag
1261 acgtgcgact tctgaagac cacttgtgaa cattgtggca ctgaaaattt agttattgaa
1321 ggacctacta catgtgggta cctacctact aatgctgtag tgaaaatgcc atgtcctgcc
1381 tgcacaagacc cagagattgg acctgagcat agtgttgtag attatcaaa ccaactcaaac
1441 attgaaactc gactccgcaa gggaggtagg actagatggt ttggaggctg tgtgtttgcc
1501 tatgttggtc gctataataa cgtgccttac tgggttctct gtagtagtgc tctattggc
1561 tcaggccata ctggcattac tgggtgacaat gtaggacct tgaatgagga tctccttagg
1621 aactgagtc gtgaacgtgt taacattaac attgttggcg attttctttt gaatgaagag
1681 gttgccatca ttttggcatc tttctctgct tctacaagtg cttttattga cactataaag
1741 agcttggatt acaagcttt caaaaccatt gttgagtcct gcgtaacta taaagtacc
1801 aagggaaagc ccgtaaaagg tgcttggaaac attggacaac agagatcagt ttttaaccca
1861 ctgtgtggtt tccctcaca ggctgctggt gttatcagat caatttttgc gcgcacactt
1921 gatgcagcaa accactcaat tcctgatttg caaagagcag ctgtcaccat acttgatggt
1981 atttctgaac agtcattacg tcttgtcgac gccatggttt atacttcaga cctgctcacc
```

2041 aacagtgtca ttattatgac atatgtaact ggtggtcttg tacaacagac ttctcagtg  
2101 ttgtctaate ttttgggac tactgttgaa aaactcaggc ctatctttga atggattgag  
2161 gcgaaactta gtgcaggatg tgaatttctc aaggatgctt gggagattct caaatttctc  
2221 attacaggtg tttttgacat cgtcaagggt caaatacagg ttgcttcaga taacatcaag  
2281 gattgtgtaa aatgcttcat tgatgtgtt aacaaggcac tcgaaatgag cattgatcaa  
2341 gtcaactatc ctggcgcaaa gttgcatca ctcaacttag gtgaagtctt catcgctcaa  
2401 agcaagggac tttaccgtca gtgtatacgt ggcaaggagc agctgcaact actcatgcct  
2461 cttaaaggac caaagaagt aacctttctt gaaggtgatt cacatgacac agtacttacc  
2521 tctgaggagg ttgttctcaa gaacggtgaa ctgcaagcac tcgagacgcc cgttgatagc  
2581 ttcaaaatg gagctatcgt tggcacacca gtctgtgtaa atggcctcat gctcttagag  
2641 attaaggaca aagaacaata ctgagcattg tctcttggt tactggctac aaacaatgtc  
2701 tttcgcttaa aagggggtgc accaatataa ggtgtaacct ttggagaaga tactgtttg  
2761 gaagttcaag gttacaagaa tgtgagaatc acatttgagc ttgatgaacg tttgacaaa  
2821 gtgcttaatg aaaagtgtc tgtctacact gttgaatccg gtaccgaagt tactgagttt  
2881 gcatgtgttg tagcagagtg tgtgtgaa gactttacaac cagtttctga tctccttacc  
2941 aacatgggta ttgatcttga tgagtggagt gtagctacat tctacttatt tgatgatgct  
3001 ggtgaagaaa acctttcatc acgtatgtat tgttctttt accctccaga tgaggaagaa  
3061 gaggacgatg cagagtgtga ggaagaagaa attgatgaaa cctgtgaaca tgagtacggt  
3121 acagaggatg attatcaagg tctccctctg gaatttggtg cctcagctga aacagttcga  
3181 gttgaggaag aagaagagga agactggctg gatgatacta ctgagcaatc agagattgag  
3241 ccagaaccag tggaaagacc taacacacc gttaatcagt ttactggtta tttaaaactt  
3301 actgacaatg ttgccattaa atgtgttgac atcgttaagg aggcacaaag tgctaactct  
3361 atggtgattg taaatgtctc taacatacac ctgaaacatg gtggtggtgt agcagggtca  
3421 ctcaacaagg caaccaatgg tgccatgcaa aaggagagtg atgattacat taagctaaat  
3481 ggccctctta cagttaggag gtctgtttg ctttctggac ataactctgc taagaagtgt  
3541 ctgcatgttg ttggaccctg cctaaatgca ggtgaggaca tccagcttct taaggcagca  
3601 tatgaaaatt tcaattcaca ggacatctta cttgacccat tgtgtcagc aggcataatt  
3661 ggtgctaaac cacttcagtc tttacaagtg tgcgtgcaga cggctcgtac acaggtttat  
3721 attgcagtca atgacaagg tctttatgag caggttgtca tggattatct tgataacctg  
3781 aagcctagag ttgaaagacc taaacaagag gagccaccaa acacagaaga ttccaaaact  
3841 gaggagaaat ctgtcgtaca gaagcctgtc gatgtgaagc caaaaatata ggcctgcatt  
3901 gatgagggta ccacaacact ggaagaaact aagtttctta ccaataagtt actctgtttt  
3961 gctgatatac atggtaagct ttaccatgat tctcagaaca tgcttagagg tgaagatag  
4021 tctttctctg agaaggatgc accttacctg gtagggtgag ttatcactag ttgtgatc  
4081 acttgtgttg taataccctc caaaaaggct ggtggcacta ctgagatgct ctcaagagct  
4141 ttgaagaaag tgccagttga tgagtatata accacgtacc ctggacaagg atgtgctggt  
4201 tatacacttg aggaagctaa gactgctctt aagaaatgca aatctgcatt ttatgtacta  
4261 ccttcagaag cacctaagtc taaggaagag attctaggaa ctgtatcctg gaatttgaga  
4321 gaaatgcttg ctcatgctga agagacaaaga aaattaatgc ctatatgcat ggtatgtaga  
4381 gccataatgg caaccatcca acgtaagtat aaaggaatta aaattcaaga gggcatcgtt  
4441 gactatgggt tccgatttct cttttatact agtaagagc ctgtagcttc tattattacg  
4501 aagctgaact ctctaataga gccgctgtc acaatgcaa ttggttatgt gacacatggt  
4561 tttaatcttg aagagctgc gcgctgtatg cgttctctta aagctcctgc gcagtgtca  
4621 gtatcatcac cagatgctgt tactacatat aatggatacc tcacttcgtc atcaagaca  
4681 tctgaggagc acctttaga aacagtttct ttggctggt cttacagaga ttggtcctat  
4741 tcaggacagc gtacagagtt aggtgttgaa tttcttaagc gtggtgacaa aattgtgtac  
4801 cacactctgg acagcccgt cgagtttcat cttgacggtg aggttcttct acttgacaaa  
4861 ctaaagagtc tcttatccct gcgggaggtt aagactataa aagtggtcac aactgtggac  
4921 aacactaatc tccacacaca gcttgggat atgtctatga catatggaca gcagtttgg  
4981 ccaacatact tggatgggtc tgatgttaca aaaatataac ctcatgtaa tcatgaggg  
5041 aagactttct ttgtactacc tagtgatgac acactacgta gtgaagcttt cgagtactac  
5101 cactactctg atgagaagtt tcttggtagg tacatgtctg ctttaaaaca cacaaagaaa  
5161 tggaaatttc ctcaagttgg tggtttaact tcaattaaat gggctgataa caattgttat  
5221 ttgtctagtg ttttatagc acttcaacag cttgaagca aattcaatgc accagcactt  
5281 caagaggctt attatagagc ccgtgctggt gatgctgcta acttttgtc actcatactc  
5341 gcttacagta ataaaactg ttggcagctt ggtgatgca gagaactat gaccactctt  
5401 ctacagcatg ctaatttggaa atctgcaaag cgagttctta atgtggtgtg taaacattgt  
5461 ggtcagaaaa ctactacctt aacgggtgta gaagctgtga tgtatatgg tactctatct  
5521 tatgataatc ttaagacagg tgtttccatt ccatgtgtgt gtggtcgtga tgtacacaa  
5581 tatctagtag aacaagagtc tcttttgtt atgatgctg caccacctgc ttagtataaa  
5641 ttacagcaag gtacattctt atgtgcgaat gagtacactg gtaactatca gttggtcat  
5701 tacactcata taactgctaa ggagaccctc tatcgtattg acggagctca cttacaaag  
5761 atgtcagagt acaaggacc agtgactgat gtttctaca aggaacatc ttacactaca  
5821 accatcaagc ctgtgtcgt taaactcgat ggagttactt acacagagat tgaacaaaa  
5881 ttggatgggt attataaaaa ggataatgct tactatacag agcagcctat agacttga  
5941 ccaactcaac cattacaaa tgcgagttt gataatttca aactcacatg ttctaacaca  
6001 aaatttgctg atgattataa tcaaatgaca ggcttcacaa agccagcttc acgagagcta  
6061 tctgtcacat tcttcccaga cttgaatggc gatgtagtgg ctattgacta tagacactat  
6121 tcagcgagtt tcaagaagg tgctaaatta ctgcataagc caattgtttg gcaattaac  
6181 caggctacaa ccaagacaac gttcaaacca aacacttgggt gtttactgtg tctttggagt  
6241 acaagccag tagatactt aaattcatt gaagttctgg cagtagaaga cacacaagga  
6301 atggacaatc ttgcttgtga aagtcaacaa cccacctctg aagaagtagt ggaatcct  
6361 accatacaga aggaagtcag agagtgtgac gtgaaacta ccgaagttgt agcagatgc  
6421 atacttaaac catcagatga aggtgttaaa gtaacacaag agttaggcca tgaggatctt  
6481 atggctgctt atgtgaaaa caaagcatt accatataa aacctaata gctttcacta  
6541 gccttaggtt taaaaacaat tgccactcat ggtattgctg caattaatag tgttcttgg  
6601 agtaaaattt tggcttatgt caaacattc ttaggacaag cagcaattac aacatcaaat

6661 tgcgctaaga gattagcaca acgtgtgttt aacaattata tgccttatgt gttacatta  
6721 ttgtccaat tgtgtacttt tactaaaagt accaattcta gaattagagc ttcactacct  
6781 acaactattg ctaaaaatgt tgtaagagt gttgtcaaat tatgtttgga tgcggcatt  
6841 aattatgtga agtaccacaa attttctaaa ttgttcacaa tgcgtatgtg gctattgttg  
6901 ttaagtattt gcttaggttc tctaactgtt gtaactgctg cttttggtgt actcttatct  
6961 aattttgggt ctccctctta ttgtaatggc gttagagaat tgmtatctaa ttcgtctaac  
7021 gttactacta tggatttctg tgaaggttct tttccttgca gcatttgttt aagtggatta  
7081 gactcccttg atcttatcc agctcttgaa accattcagg tgacgatttc atcgtacaag  
7141 ctgacttga caattttagg tctggccgct gagtgggttt tggcatatat gttgttcaca  
7201 aaattctttt atttattagg tctttcagct ataatgcagg tgmtctttgg ctattttgct  
7261 agtcatttca tcagcaattc ttggctcatg tggtttatca ttagtattgt acaaatggca  
7321 cccgtttctg caatggttag gatgtacatc tctttgctt ctttctacta catatggaag  
7381 agctatgttc atatactga tggttgcacc tcttcgactt gcatgatgtg ctataagcgc  
7441 aatcgtgcca caccgcttga gtgtacaact attgttaatg gcatgaagag atctttctat  
7501 gtctatgcaa atggaggccg tggcttctgc aagactcaca attggaattg tctcaattgt  
7561 gacacatttt gcactggtag tacattcatt agtgatgaag ttgctcgtga tttgtcactc  
7621 cagtttaaaa gaccaatcaa cctactgac cagtcacgt atattgttga tagtgttctg  
7681 gtgaaaaatg gcgcttca cctctacttt gacaaggctg gtcaaaagac ctatgagaga  
7741 catccgctct cccattttgt caatttagac aatttgagag ctaacaacac taaaggttca  
7801 ctgcctatta atgtcatagt tttgatggc aagtccaaat gcgacgagtc tgcctctaag  
7861 tctgcttctg tgtactacag tcagctgatg tgccaaccta tcttgttctg tgaccaagtt  
7921 cttgtatcag acgttgaga tagtactgaa gtttccgta agatgttga tgcctatgtc  
7981 gacacctttt cagcaacttt tagtgttctc atggaaaaac ttaaggcact tgttctaca  
8041 gctcacagcg agttagcaaa ggggttagct ttagatgggt tcctttctac attcgtgtca  
8101 gctgcccgc aaggtgttgt tgataccgat gttgacacaa aggatgttat tgaatgtctc  
8161 aaactttcac atctactga cttagaagtg acaggtgaca gttgtaacaa tttcatgctc  
8221 acctataata aggttgaaaa catgacgccc agagatcttg gcgcatgtat tgactgtaat  
8281 gcaaggcata tcaatgccc agtagcaaaa agtcacaatg tttcactcat ctggaatgta  
8341 aaagactaca tgtctttatc tgaacagctg cgtaaacaaa ttcgtagtgc tgccaagaag  
8401 aacaacatac cttttagact aacttgctct acaactagac aggttgtcaa tgtcataact  
8461 actaaaaatc cactcaaggg tggtaagatt gttagtactt gttttaaact tatgcttaag  
8521 gccacattat tgtgcgttct tgctgcattg gtttgttata tctgtatgcc agtacatata  
8581 ttgtcaatcc atgatggtta cacaaatgaa atcattggtt acaaaagccat tcaggatggt  
8641 gtcaactcgt acatcatttc tactgatgat tgttttgcaa ataaacatgc tggctttgac  
8701 gcatggttta gccagcgtgg tggttcatac aaaaatgaca aaagctgcc tgtagttagt  
8761 gctatcatta caagagagat tggttcata gtgcctggct taccgggtac tgtctgaga  
8821 gcaatcaatg gtgacttctt gcattttcta cctcgtgttt ttagtgtgtg tggcaacatt  
8881 tgctacacac ctccaaactc cattgagtat agtgattttg ctacctctgc ttcgcttctt  
8941 gctgctgagt ttacaatfff taaggatgct atgggcaaac ctgtgccata ttgttatgac  
9001 actaatttgc tagaggggtc tatttcttat agtgagctc gtccagacac tctgtatgtg  
9061 cttatggatg gttccatcat acagtttctc aacacttacc tggagggttc tgttagagta  
9121 gtaacaactt ttgatctgta gtactgtaga catggtacat gcgaaaggtc agaagtaggt  
9181 atttgccctat ttagcctggt tagatgggtt cttaaataag agcattacag agctctatca  
9241 ggagttttct gtggtgttga tgcgatgaat ctcatagcta acatctttac tcctctgtg  
9301 caactgtggt gtgctttaga tgtgtctgct tcagtagtgg ctggtggtat tattgccata  
9361 ttggtgactt gtgctgccta ctactttatg aaattcagac gtgtttttgg tgaatgacaac  
9421 catgttgttg ctgctaatgc acttttggtt ttgatgtctt tcaactactc tgttctgga  
9481 ccagcttaca gctttctgcc gggagtctac tcagcttttt acttgtactt gacattctat  
9541 ttaccaatg atgtttcatt ctggctcac cttcaatggt ttgccatggt ttctcctatt  
9601 gtgccttttt ggataacagc aatctatgta ttctgtattt ctctgaagca ctgccattgg  
9661 ttctttaaca actactctag gaaaagagtc atgtttaatg gatttaccatt tagtaccctc  
9721 gaggaggctg ctttgtgtac cttttgctc aacaaggaaa tgcacctaaa attgctgagc  
9781 gagacactgt tgccacttac acagtataac aggtatcttg ctctataata caagtacaag  
9841 tttttagctg gagccttaga tactaccagc fatcgtgaag cagcttctg ccacttagca  
9901 aaggctctaa atgacttttag caactcaggt gctgatgttc tctaccaacc accacagaca  
9961 tcaatcaact ctgctgttct gcagagtggt tttagaaaaa tggcattccc gtcaggcaaa  
10021 gttgaagggg gcatggtaca agtaacctgt ggaactacaa ctctaatagg attgtgggtg  
10081 gatgacacag tatakgtcc aagacatgct atttgcacag cagaagacat gcttaatcct  
10141 aactatgaag atctgctcat tgcgaaatcc aaccatagct tcttgttca ggtggcaat  
10201 gttcaacttc gtttattgg ccattctatg caaaattgct tgcctaggct taaagttgat  
10261 acttctaacc ctaagacacc caagtataaa tttgtccgta tccaacctgg tcaaacattt  
10321 tcagttctag catgctacaa tggttcacca tctgggtttt atcagtgctc catgagacct  
10381 aatcatacca ttaaagggtc tttccttaat ggatcatggt gtagtgtttg ttttaacatt  
10441 gattatgatt gcgtgtcttt ctgctatatg catcatatgg agcttccaac aggagtacac  
10501 gctggtagct acttagaagc taaattctat ggtccatttg ttgacagaca aactgacag  
10561 gctgcaggta cagacacaac cataacatta aatgttttgg catggctgta tgcgtctgtt  
10621 atcaatgggt ataggtggtt tcttaataga ttcaccacta ctttgaatga ctttaacctt  
10681 gtggcaatga agtacaacta tgaacctttg acacaagatc atgttgacat attgggacct  
10741 ctttctgctc aaacaggaat tgccccttta gatagtgtg ctgctttgaa agagctgctg  
10801 cagaatggta tgaatggctg tactatcctt ggtagcacta ttttagaaga tgaatttaca  
10861 ccattttagt ttgttagaca atgctctggt gttaccttc aaggttaagt caagaaaatt  
10921 gttaaagggca ctcatcattg gatgctttta actttcttga catcactatt gattcttgtt  
10981 caaagtacac agtggctact gttttcttt gtttacgaga atgctttctt gccatttact  
11041 cttggtatta tggcaattgc tgcctgtgct atgctgctg ttaagcataa gcacgcttc  
11101 ttgtgcttgt ttctgttacc ttctcttgca acagttgctt actttaatat ggtctacatg  
11161 cctgtagctg gggtagctg tatcatgaca tggcttgaat tggctgacac tagctgtctt  
11221 ggttataggc ttaaggattg tgttatgat gcttcagctt tagttttgct tattctcatg



11281 acagctcgca ctgtttatga tgatgctgct agacgtggtt ggacactgat gaatgtcatt  
11341 acacttggtt acaaagtcta ctatggtaat gcttttagatc aagctatttc catgtggggc  
11401 ttagttattt ctgtaacctc taactattct ggtgtcgta cgactatcat gtttttagct  
11461 agagctatag tgtttgtgtg ttgtgagtat taccattgt tatttattac tggcaacacc  
11521 ttacagtgtg tcatgcttgt ttattgtttc ttaggctatt gttgctgctg ctactttggc  
11581 cttttctggt tactcaaccg ttacttcagg cttactcttg gtgtttatga ctacttggtc  
11641 tctacacaag aatthagga tatgaactcc caggggcttt tgcctcctaa gagtagtatt  
11701 gatgctttca agcttaacat taagtgttg ggtattggag gtaaacatg atcraaggtt  
11761 gctactgtac agtctaaat gtctgacgta aagtgacat ctgtggtact gctctcggtt  
11821 cttcaacaac ttagagtaga gtcacttctt aaattgtggg cacaatgtgt acaactccac  
11881 aatgatattt ttcttgcaaa agacacaact gaagctttcg agaagatggt ttctcttttg  
11941 tctgttttgc tatccatgca ggtgctgta gacattaata ggttgtgca ggaatgctc  
12001 gataaccctg ctactcttca ggctattgct tcagaattta gttctttacc atcatatgcc  
12061 gcttatgcca ctgccagga ggctatgag caggctgtag ctaatggtga ttctgaagtc  
12121 gttctcaaaa agttaaagaa atctttgaat gtggctaaat ctgagtttga cctgtatgct  
12181 gccatgcaac gcaagtgtga aaagatggca gatcaggcta tgaccctaaat gtacaaacag  
12241 gcaagatctg aggcaagagta agcaaaagta actagtgtca tgcaaacatg gctcttctc  
12301 atgcttagga agcttgataa tgatgacctt aacaacatta tcaacaatgc gcgtgatggt  
12361 tgtgttccac tcaacatcat accattgact acagcagcca aactcatggt tgttgcctc  
12421 gattatggta cctacaagaa cacttgtgat ggtaacacct ttacatagc atctgactc  
12481 tgggaaatcc agcaagtgtg tgatgaggat agcaagattg ttcaacttag tgaattaac  
12541 atggacaatt caccaaatth ggcttggcct cttattgtta cagctctaag agccaactca  
12601 gctgttaaac tacagaataa tgaactgagt ccagtagcac tacgacagat gtcctgtgcg  
12661 gctggtacca cacaaacagc ttgtactgat gacaatgac ttgctacta taacaattcg  
12721 aaggaggta ggtttgtgct ggctacta tcagaccacc aagatctcaa atggctaga  
12781 tcccctaaga gttatggtat aggtacaatt tacacagAAC tggaaaccac ttgtaggtt  
12841 gttacagaca caccaaaagg gcctaaagt aaatacttgt acttcatcaa aggcctaac  
12901 aacctaaata gaggtatggt gctgggagc ttagtctgta cagtacgtct tcaggctgga  
12961 aatgtctacag aagtacctgc caattcaact gtgctttcct tctgtgcttt tgcagtagac  
13021 cctgctaagag catataagga ttacctagca agtggaggac aaccaatcac caactgtgtg  
13081 aagatgttgt gtacacacac tggtagcaga caggcaatta ctgtaacacc agaagctaac  
13141 atggaccaag agtcttttgg tgggtcttca tgttctctgt attgtatagc ccattgac  
13201 catccaaatc ctaaaaggatt ctgtgacttg aaaggtaagt acgtccaaat acctaccat  
13261 tgtgctaaty acccagtggg ttttactctt agaaacacag tctgtaccgt ctgcggaatg  
13321 tggaaagggt atggctgtag ttgtgaccaa ctccgcaac ccttgatgca gctgaggat  
13381 gcatcaacgt ttttaacagg gtttgcggtg taagtgcagc ccgtcttaca cctgaggca  
13441 caggcactag tactgatgtc gtctacaggg cttttgatat ttacaacgaa aaagtgtctg  
13501 gttttgcaaa gttcttaaaa actaatgtct gtcgcttcca ggagaaggat gaggaaggca  
13561 atttattaga ctcttacttt gtagttaaga ggcatactat gtctaacctac caacatgaag  
13621 agactattta taacttgggt aaagattgtc cagcggttgc tgtccatgac ttttcaagt  
13681 ttagagtaga tggtagcatg gtaccacata tatcacgtca gcgtctaact aaatacacia  
13741 tggctgattt agtctatgct ctacgtcatt ttgatgggg taattgtgat acattaaaag  
13801 aaatactcgt cacatacaat tgctgtgatg atgattattt caataagaag gattggtatg  
13861 acttctgtaga gaatcctgac atcttaccg tatarctaa cttagggtgag cgtgtacgcc  
13921 aatcattatt aaagactgta caattctgc atgctatgct tgatgagggc attgtaggcg  
13981 tactgacatt agataatcag gatcttaatg ggaactggta cgatttcggt gatttctgac  
14041 aagtagcacc aggtgagcgc ttaactattg tggattcata ttactcattg atgtggcca  
14101 tcctcaactt gactagggca ttggctgctg agtcccata ggtgctgat ctgcaaaac  
14161 cacttattaa gtgggatttg ctgaaatag attttacgga agagagactt tgtctcttcg  
14221 accgttattt taaatattgg gaccagacat accatcccaa ttgtattaac tgtttggatg  
14281 atagggtgat ccttcatgtg gcaaaactta atgtgttatt ttctactgtg tttccacct  
14341 caagttttgg accactagta agaaaaatat ttgtataggg tgttctttt gttgtttcaa  
14401 ctggatacca ttttctgtg ttaggagctg tacataatca ggtgtaaac ttacatagct  
14461 cgcgtctcag tttcaaggaa cttttagtgt atgctgctga tccagctatg catgcagctt  
14521 ctggcaattt attgctgat aaacgcacta catgctttc agtagctgca ctaacaaaca  
14581 atgttcttct tcaaacctgtc aaaccggta attttaataa agacttttat gactttgctg  
14641 tgtctaaagg tttctttaa gaaggaagt ctgttgaact aaaacacttc ttttctgctc  
14701 aggatggcaa cgtctctatc agtgattatg actattatcg ttataatctg ccaacaatgt  
14761 gtgatatcag acaactccta ttcgtagtgt aagttgttga taaatacttt gattgttacg  
14821 atgtggctg tattaatgct aaccaagtaa tctgtaacaa tctggataaa ttagctggtt  
14881 tcccatttaa taaatggggat aaggctagac tttattatga ctcaatgagt tatgaggatc  
14941 aagatgacct tttcgcgtat actaagcgtg atgtcatccc tactataact caaatgaatc  
15001 ttaagtatgc cattagtgtc aagaatagag ctgcaccgt agctgggtgc tctatctgta  
15061 gtactatgac aaatagacag tttcatcaga aattattgaa gtcaatagcc gccactagag  
15121 gagctactgt agtatttgg acaagcaagt tttacgggtg ctggcataat agtttaaaa  
15181 ctgtttacag tgatgtagaa actccacacc ttatgggttg ggattatcca aaatgtgaca  
15241 gagccatgcc taacatgctt aggataatgg cctctctgt tcttctcgc aaacataaca  
15301 cttgctgtaa cttatcacac cgtttctaca ggttagctaa cgagtgtgcg caagtattaa  
15361 gtgagatggt catgtgtgct ggctcaact atgttaaac aggtggaaaca tcatccggg  
15421 atgtacaac tgcttatgct aatagtgtct ttaacattg tcaagctgtt acagccaatg  
15481 taaatgacct tcttcaact gatgtaata agatagctga caagtatgct cgcaatctac  
15541 aacacaggct ctatgagtgt ctctatagaa ataggatgt tgatcatgaa ttcgtggatg  
15601 agttttacgc ttacctgctg aaacatttct ccatgatgat tctttctgat gatgacctg  
15661 tgtgtataaa cagtaactag cggctcaag gtttagtagc tagcattaa aactttaagg  
15721 cagttcttta ttatcaaaat aatgtgttca tgtctgaggc aaaatgttgg actgagactg  
15781 accttactaa aggacctcac gaatttctc cacagcatc aatgctagtt aaacaaggag  
15841 atgattacgt gtacctgctt taccagatc catcaagaat attaggcgca ggctgtttg

15901 tcgatgatat tgtcaaaaca gatggtacac ttatgattga aaggttcgtg tcaactggcta  
15961 ttgatgctta cccacttaca aaacatccta atcaggagta tgctgatgtc tttcacttgt  
16021 atttacaata cattaagaag ttacatgatg agcttactgg ccacatgttg gacatgtatt  
16081 ccgtaatgct aactaatgta aacacctcac ggtagctgga acctgagttt tatgaggcta  
16141 tgtacacacc acatacagtc ttgcaggctg taggtgcttg tgtattgtgc aattcacaga  
16201 cttcacttcg ttgcggtgcc tgtatttagga gaccattcct atgttgcaag tgctgctatg  
16261 accatgtcat ttcaacatca cacaaattag tgttgtctgt taatccctat gtttgcaatg  
16321 cccacgggtg tgaagtctact gatgtgacac aactgtatct aggaggtatg agctattatt  
16381 gcaagtcaca taagcctccc attagttttc cattatgtgc taatggtcag gtttttggtt  
16441 tatacaaaaa cacatgtgta ggcagtgaca atgtcactga cttcaatgag atagcaacat  
16501 gtgattggac taatgctggc gattacatac ttgccaacac ttgtactgag agactcaagc  
16561 ttttcgagc agaaacgctc aaagccactg aggaacatt taagctgtca tatggtattg  
16621 ctactgtacg cgaagtactc tctgacagag aattgcatct ttcattggag gttggaaaac  
16681 ctagaccacc attgaacaga aactatgtct ttactggtta ccgtgtaact aaaaatagta  
16741 aagtacagat tggagagta acctttgaaa aaggtgacta tgggtgatgct gttgtgtaca  
16801 gaggtactac gacatacaag ttgaatgttg gtgattactt tgggttgaca tctcacactg  
16861 taatgccact tgaagccact ctctagtc cacaagagca ctatgtgaga attactggct  
16921 tgtaccacaac actcaacatc tcagatgagt tttctagcaa tgttgcaaat tatcaaaagg  
16981 tcggcatgca aaagtactct acactccaag gaccacctgg tactggtaag agctattttg  
17041 ccatcggact tgctctctat taccatctg ctgcatagt gtataccgca tgctctcatg  
17101 cagctgttga tgaagcctat gaaaaggcat taaaatattt gcccatagat aaatgtagta  
17161 gaatcatacc tgcgctgctg cgcgtagagt gttttgataa attcaaatg aattcaacac  
17221 tagaacagta tgtttctctg actgtaaatg cattgccaga aacaactgct gacattgtag  
17281 tctttgatga aatctctatg gctactaatt atgacttgag tgttgtaaat gctagacttc  
17341 gtgcaaaaac ctactctat attggcagtc ctgctcaatt accagcccc cgcaacttc  
17401 tgactaaagg cacactagaa ccagaatatt ttaattcagt gtgcagact atgaaaaaca  
17461 taggtccaga catgttcctt ggaactgtc gccgtgtcc tgcgtgaaat gttgacactg  
17521 tgagtctttt agtttatgac aataagctaa aagcacacaa ggataaagca gctcaatgct  
17581 tcaaaatggt ctacaagggt gttattacac atgatgtttc atctgcaatc aacagacctc  
17641 aaataggcgt tgaagagaaa tttcttacac gcaatcctgc ttggagaaaa gctgttttta  
17701 tctcacctta taattcacag aacgctgtag cttcaaaaat cttaggattg cctacgcaga  
17761 ctggtgattc atcacagggt tctgaatag actatgtcat attcacacaa actactgaaa  
17821 cagcacactc ttgtaatgtc aaccgcttca atgtggctat cacaagggca aaaattggga  
17881 ttttgtgcat atgtctgat agagatcttt atgacaaaact gcaatttaca agtctagaaa  
17941 taccacgctg caatgtggct acattacaag cagaaaatgt aactggactt ttaaggact  
18001 gtagtaagat cattactggt ctctaccta cacaggcacc tacacacctc agcgttgata  
18061 taaagtcaa gactgaagta ttatgtgttg acataccagg cataccaag gacatgacct  
18121 accgtagact catctctatg atgggtttca aaatgaatta ccaagtcaat ggttacccta  
18181 atatgtttat caccgcgaa gaagctattc gtcacgttcg tgcgtggatt ggtttgatg  
18241 tagaggctg tcatgcaact agagatgctg tgggtactaa cctacctc cagctaggat  
18301 tttctacagg tgttaactta gtactgttac cgactggtta tgttgacact gaaaaataca  
18361 cagaattcac cagagttaat gcaaaacctc caccagggtg ccagtttaa catcttatac  
18421 cactcatgta taaaggctg ccttggatg tagtgctat taagatagta caaatctca  
18481 gtgatacact gaaaggattg tcagacagag tgcgttctg cctttggcg catggctttg  
18541 agcttacatc aatgaagtac tttgtcaaga ttggacctga aagaacctgt tgcctgtgtg  
18601 acaaacgtgc aacttgcttt tctacttcat cagatactta tgcctgtctg aatcattctg  
18661 tgggttttga ctatgtctat aaccattta tgattgatg tcagcagtg ggctttacg  
18721 gtaaccttca gagtaacct gaccaacatt gccagggtaca tggaaatgca catgtggcta  
18781 gttgtgatg tcatgact agatgtttag cagtcctga gtgctttgtt aagcgcgttg  
18841 attggtctgt tgaatacctt attataggag atgaactgag ggttaattct gcttgacaga  
18901 aagtacaaca ctgggtgtg aagctgcat tgcctgctga taagtttcca gttctctcatg  
18961 acattggaaa tcaaaagctc atcaagtggt tgcctcaggc tgaagtagaa tggaaagtct  
19021 acgatgctca gccatgtagt gacaaagctt acaaaataga ggagctcttc tattcttatg  
19081 ctacacatca cgataaattc actgatggtg tttgtttgtt ttggaattgt aacgttgatc  
19141 gttaccagc caatgcaatt gtgtgtaggt ttgacacaag agtctgtca aactgaaact  
19201 taccaggctg tgaatgtgt agtttgtatg tgaataagca tgcattcac actccagctt  
19261 tcgataaaaag tgcatttact aatttaaagc aattgccttt ctttactat tctgatagtc  
19321 cttgtgagtc tcatggcaaa caagttagt cggatattga ttatgttcca ctcaaatctg  
19381 ctactgtgat tacacgatg aatttaggtg gtgctgtttg cagacacat gcaaatgagt  
19441 accgacagta ctggatgca tataatatga tgatttctg tggatttag ctatggattt  
19501 acaacaatt tgaacttat aacctgtgga atacatttac caggttacag agtttagaaa  
19561 atgtggctta taatgttgtt aataaaggac actttgatgg acacgccgc gaagcacctg  
19621 tttccatcat taataatgct gtttacacaa aggtatagtg tattgatgag gagatctttg  
19681 aaaataagac aacacttctt gttaatgttg catttgagct tgggctaag cgtaacatta  
19741 aaccagtgc agagatgaag atactcaata atttgggtgt tgatctcgt gctaaactg  
19801 taactctgga ctacaaaaga gaagcccgag cacatgtatc tacaataggt gctgcacaa  
19861 tgaactgacat tgcgaagaaa cctactgaga gtgctgtttc ttcacttact gcttgtttg  
19921 atggttagat ggaaggacag gtagacctt ttagaaacgc ccgtaatggt gtttaataa  
19981 agaaagggtc agtcaaggt ctaaacctt caaagggacc agcaaacgt agcgtcaatg  
20041 gactcacatt aattggagaa tcagtaaaaa cacagtttaa ctactttaag aaagtagacg  
20101 gcattattca acagtgtcct gaaacctact ttactcagag cagagactta gaggatttta  
20161 agcccagatc acaaatggaa actgactttc tcgagctcgc tatggatgaa ttcatacagc  
20221 gatataagct cgagggtact gccttcaaac acatcgttta tggagatttc agtcatggac  
20281 aacttggcgg tcttcattha atgataggct tagccaagcg ctcaaacatg tcaccactta  
20341 aattagagga ttttatccct atggacagca cagtgaaaaa ttacttcata acagatgctc  
20401 aaacagggtc atcaaaatgt gtgtgtctg tgattgatct tttacttgat gactttgtcg  
20461 agataataaa gtcacaagat ttgtcagtga tttcaaaagt ggtcaaggtt acaattgact

20521 atgctgaaat ttcattcatg ctttgggtga aggatggaca tgttgaacc ttctaccaa  
20581 aactacaagc aagtcaagcg tggcaaccag gtgttgcgat gcctaacttg tacaagatgc  
20641 aaagaatgct tcttgaagc tgtgacctc agaattatgg tgaaatgct gttataccaa  
20701 aaggaataat gatcaatgtc gcaaagtata ctcaactgtg tcaactacta aatacactta  
20761 ctttagctgt accctacaac atgagagtta ttcactttgg tegtggctct gataaaggag  
20821 ttgcaccagg tacagctgtg ctcagacaat ggttgccaac tggcacacta cttgtcgatt  
20881 cagatcttaa tgacttcgtc tccgacgcag attctacttt aattggagac tgtgcaacag  
20941 tacatacgcg taataaagcg gaccttatta ttacgcatat gtatgacctt aggaccaaac  
21001 atgtgacaaa agagaatgac tctaaagaag ggtttttcac ttatctgtgt ggatttataa  
21061 agcaaaaact agccctgggt ggttctatag ctgtaaagat aacagagcat tcttggatg  
21121 ctgaccttta caagcttatg ggccatttct catggtggac agcttttgtt acaaatgtaa  
21181 atgcatcatc atcgaagca tttttaattg gggctaacta tcttggcaag ccgaggaac  
21241 aaattgatgg ctataccatg catgctaact acattttctg gaggaacaca aatcctatcc  
21301 agttgtcttc ctattcactc ttgacatga gcaaatctc tcttaatta agaggaactg  
21361 ctgtaatgtc tcttaaggag aatcaaatca atgatatgat ttatttctct ctggaaaaag  
21421 gtaggcttat cattagagaa aacaacagag ttgtggttc aagtgatatt cttgttaaca  
21481 actaaacgaa catgtttatt tcttattat tcttactct cactagtgtt agtgaccttg  
21541 accggtgcac cacttttgat gatgttcaag ctcctaatta cactcaacat acttcatcta  
21601 tgaggggggt ttactatcct gatgaaattt ttagatcaga cactctttat ttaactcagg  
21661 atttatttct tccattttat tctaagtta cagggttca tactattaat catacgtttg  
21721 gcaacctctg tacacttttt aaggatggta tttattttgc tgccacagag aatcaaatg  
21781 ttgtccgtgg ttgggttttt ggttctacca tgaacaaca gtcacagtcg gtgattatta  
21841 ttaacaattc tactaatgtt gttatacgag catgtaactt tgaattgtgt gacaacctt  
21901 tctttgctgt ttctaaccac atgggtacac agacacatac tatgatattc gataatgcat  
21961 ttaattgcac tttcagatc atactctgat cctttctgct tgaatgttca gaaagtctg  
22021 gtaattttaa acactttcga gagtttgtgt ttaaaaataa agatgggttt ctctatggtt  
22081 ataagggcta tcaacctata gatgtagttc gtgatctacc ttctggtttt aacctttga  
22141 aacctatttt taagttgcct ctgtgtatta acattacaaa ttttagagcc attcttacag  
22201 ccttttcacc tgctcaagac atttggggca cgtcagctgc agcctatttt gttggctatt  
22261 taaagccaac tacattttat ctcaagtatg atgaaaatgg tacaatcaca gatgctgttg  
22321 atgtttctca aatccactt gctgaactca aatgctctgt taagagcttt gagattgaca  
22381 aaggaattta ccagacctct aatttcaggg ttgttccctc agggatggtt gtgagattcc  
22441 ctaatattac aactttgtgt ccttttggag aggtttttaa tgctactaaa ttccttctg  
22501 tctatgcatg ggagagaaaa aaaatttcta attgtgttgc tgattactct gtgctctaca  
22561 actcaacatt ttttcaacc tttaaagtct atggcgtttc tgccactaag ttgaatgatc  
22621 tttgcttctc caatgtctat gcagattcct ttgtagtcaa gggagatgat gtaagacaaa  
22681 tagccccagg acaaacctgt gttattgctg attataatta taaattgcc a gatgatttca  
22741 tgggtgtgtg ccttctgttg aatactagga acattgatgc tacttcaact ggttaattata  
22801 attataaata taggtatcct agacatggca agcttaggcc ctttgagaga gacatatcta  
22861 atgtgccttt ctcccctgat ggcaaacctt gcacccacc tgctcttaat tgttattggc  
22921 cattaaatga ttatggtttt tacaccacta ctggcattgg ctaccaacct tacagagttg  
22981 tagtactttc ttttgaactt ttaaatgcac cggccacggg ttgtggacca aaattatcca  
23041 ctgaccttat taagaaccag tgtgtcaatt ttaattttaa tggactcact ggtactggtg  
23101 tgttaactcc ttcttcaag agatttcaac catttcaaca atttggccgt gatgtttctg  
23161 atttactgta ttccgtctga gatcctaaaa catctgaaat attagacatt tcacctgct  
23221 cttttggggg tgtaagtgtt attacacctg gaacaaatgc ttcatctgaa gttgctgttc  
23281 tatatcaaga ttttaactgc actgatgttt ctacagcaat tcatgcagat caactcacac  
23341 cagcttggcg catatattct actggaaaca atgtattcca gactcaagca ggtgtcttca  
23401 taggagctga gcatgtcgac acttcttatg agtgcgacat tcctattgga gctggcattt  
23461 gtgctagtta ccatacagtt tctttattac gtagtactag ccaaaaatct atttggctt  
23521 atactatgtc tttagggtct gatagttcaa ttgcttactc taataacacc attgtatata  
23581 ctactaactt tcaaatagc attactacag aagtaatgcc tgttctatg gctaaaaact  
23641 ccgtagattg taatatgtac atctgaggag atttactgta atgtgctaat ttgcttctcc  
23701 aatatggtag cttttgcaca caactaaatc gtgcaactct aggtattgct gctgaacagg  
23761 atcgcaacac acgtgaagtg ttcgctcaag tcaacaacaa gtacaaaacc ccaacttga  
23821 aatattttgg tggttttaa tttcacaaa tattacctga ccctcctaaag ccaactaaag  
23881 ggtcttttat tgaggacttg ctcttataa aggtgacact cgctgatgct ggcttcatga  
23941 agcaatattg cgaatgccta ggtgatatta atgctagaga tctcatttgt gcgagaagt  
24001 tcaatggact tacagtgttg ccacctctgc tcaactgatga tatgattgct gcctacactg  
24061 ctgctctagt tagtggtagt gccactgctg gatggacatt ttggtctggc gctgctctc  
24121 aaataccttt tgctatgcaa atggcatata ggttcaatgg cattggagtt acccaaatg  
24181 ttctctatga gaacaaaaa caaatcgcca accaatttaa caaggcgatt agtcaaatc  
24241 aagaatcact tacaacaaca tcaactgcat tgggcaagct gcaagagctt gttaacccaa  
24301 atgtcgaagc attaaacaca cttgttaaac aacttagctc taattttggt gcaatttcaa  
24361 gtgtgctaaa tgatctcctt tcgcaacttg ataaagtcca ggcggaggta caaattgaca  
24421 ggttaattac aggcagactt caaaccttc aaacctatgt aacacaaca ctaatcaggg  
24481 ctgctgaaat cagggcttct gctaactctg ctgctactaa aatgtctgag tgtgttctg  
24541 gacaatcaaa aagagttgac ttttgggaa agggctacca ccttatgtc ttcccacaag  
24601 cagccccgca tgggtgtgtc ttcctacatg tcacgtatgt gccatcccag gagaggagt  
24661 tcaccacagc gccagcaatt tgtcatgaag gcaaaacata cttccctcgt gaagggttt  
24721 ttgtgtttaa tggcacttct tggtttatta cacagaggaa cttcttttct ccacaaataa  
24781 ttactacaga caatacattt gtctcaggaa attgtgatgt cgttattggc atcataaaca  
24841 acacagttaa tgactctgtc caacctgagc tcgactcatt caaagaagag ctggacaagt  
24901 acttcaaaaa catacatca ccagatgttg atcttggcga catttcaggc attaacgctt  
24961 ctgtcgtcaa cattcaaaaa gaaattgacc gcctcaatga ggtcgtcaaa aatttaaatg  
25021 aatcactcat tgaccttcaa gaattgggaa aatagagca atatattaaa tggccttggg  
25081 atgtttggct cggcttcat gctggactaa ttgccatcgt catggttaca atcttgcctt

25141 gttgcatgac tagttgttgc agttgcctca aggggtgcatg ctcttgtggt tcttctgca  
25201 agttttagta ggatgactct gagccagttc tcaaggggtg caaattacat tacacataaa  
25261 cgaacttatg gatttgttta tgagattttt tactcttggg tcaattactg cacagccagt  
25321 aaaaattgac gaagcgttct ctgcaagtac tggttcatgct acagcaacga taccgtaca  
25381 agcctcactc ctttccggat ggcttggat tgccgttgca tttcttgcg ttttccagag  
25441 cgctaccaaa ataattgctc tcaataaaag atggcagcta gccctttata agggcttcca  
25501 gttcatttgc aatttactgc tgctatttgc taccatctat tcacatcttt tgcttgcgca  
25561 tgcaaggtatg gaggcgtcaat tttgtacct ctatgccttg atatattttc tacaatgcat  
25621 caacgcatgt agaattatta tgagatgttg gctttgttgg aagtgcaaat ccaagaacct  
25681 attactttat gatgccaaact actttgtttg ctggcacaca cataactatg actactgtat  
25741 accatataac agtgtcacag atacaattgt cgttactgaa ggtgacggca tttcaacacc  
25801 aaaactcaaa gaagactacc aaattggtgg ttattctgag gataggcact gaaggttaa  
25861 agactatgct gttgtacatg gctattttac cgaagtttac taccagcttg agtctacaca  
25921 aattactaca gacactggta ttgaaaatgc tacattcttc atctttaaca agcttgttaa  
25981 agaccaccg c aatgtgcaaa tacacacaat cgacggctct tcaggagtgg ctaatccagc  
26041 aatggatcca atttatgat agccgacgac gactactagc gtgccttctg aagcacaaga  
26101 aagtgagatc actttatgtg actcattctg ttcggaagaa acaggtacgt taatagttaa  
26161 tagcgtactt cttttcttg ctttctggtt attcttgcta gtcacactag ccatccttac  
26221 tgcgcttcga ttgtgtgcgt actgtgcaa tattgttaac gtgagtttag taaaaccaac  
26281 ggtttacgct tactcgcgtg ttaaaaatct gaactcttct gaaggagttc ctgatcttct  
26341 ggtcctaaac aactaactat tattattatt ctgtttggaa ctttaacatt gcttatcatg  
26401 gcagacaacg gtactattac cgttgaggag cttaacaac tcctggaaca atggaacctt  
26461 gtaataggtt tcctattctc agcctggatt atgttactac aatttgccta ttctaactcg  
26521 aacaggtttt tgtacataat aaagcttgtt ttctctggc tcttggggcc agtaacactt  
26581 gcttgttttg tgcttgcgct tgcctacaga attaatggg tgactggcgg gattgcgatt  
26641 gcaatggcct gattgttagg ctgtatgtgg cttagctact tcggttgcctt cttcaggctg  
26701 tttgctcgta cccgctcaat gtggtcattc aaccagaaa caaacattct tctcaatgtg  
26761 cctctccggg ggacaattgt gaccagaccg ctcatggaaa gtgaacttgt cattggtgct  
26821 gtgatcattc gtggtcactt gcgaatggc ggacacccc tagggcgtg tgacattaag  
26881 gacctgccaa aagagatcac tgtggctaca tcacgaacgc tttcttatta caaattagga  
26941 gcgtcgcagc gtgtaggcac tgattcagg tttgtgcat acaaccgcta ccgtattgga  
27001 aactataaat taaatacaga ccacgccggt agcaacgaca atattgcttt gctagtacag  
27061 taagtgaaca cagatgtttc atcttgttga ctccagggtt acaatagcag agatattgat  
27121 tatcattatg aggactttca ggattgctat ttggaatctt gacgttataa taagttcaat  
27181 agtgagacaa ttatttaagc ctctaactaa gaagaattat tcggagttag atgatgaaga  
27241 acctatggag ttagattatc cataaacga acatgaaat tattctcttc ctgacattga  
27301 ttgtattttac atcttgcgag ctatatcact atcaggagtg tgttagaggt acgactgtac  
27361 tactaaaaga accttgccca tcaggaacat acgagggcaa ttcaccattt caccctctg  
27421 ctgacaataa atttgcacta acttgcacta gcacacactt tgcttttgc tgtgctgacg  
27481 gtactcgaca tacctatcag ctgctgcaa gatcagtttc accaaaactt ttcacagac  
27541 aagaggaggt tcaacaagag ctctactcgc cactttttct cattgttgc tctctagtat  
27601 ttttaatact ttgcttcacc attaagagaa agacagaatg aatgagctca ctttaattga  
27661 cttctatttg tgctttttag ctttctgctc attccttgtt ttaataatgc tttattatatt  
27721 ttggttttca ctgaaatcc aggatctaga agaacttgt accaaagtct aaacgaacat  
27781 gaaacttctc atgtttttga ctgtatttc tctatgcagt tgcatatgca ctgtagtaca  
27841 gcgctgtgca tctaataaac ctcatgtgct tgaagatcct tgtaaggtag aacactaggg  
27901 gtaatactta tagcactgct tggctttgtg ctctaggaaa ggttttacct tttcatagat  
27961 ggcacactat gtttcaaa ca tgccacacta atgttactat caactgtcaa gatccagctg  
28021 gtggtgctct tatagctagg tgttggtagc ttcattgaagg tcaccaaact gctgcattta  
28081 gagagctact tgttgtttta aataaacgaa caaattaaaa tgcctgataa tggaccccaa  
28141 tcaaaccaac gtagtgcctc ccgcatcaca tttgggtggc ccacagattc aactgacaat  
28201 aaccagaatg gaggagcaca tggggcaagg ccaaaaacagc gccgacccca aggtttacc  
28261 aataatactg cgtcttgggt cacagctctc actcagcatg gcaaggagga acttagattc  
28321 cctcagagcc agggcgttcc aatcaacacc aatagtgtc cagatgacca aattggctac  
28381 taccgaagag ctaccgacg agttcgtggt ggtgacggca aatgaaaga gctcagcccc  
28441 agatggtact tctattacct ctattactc aggaactggc ccagaagctt cacttcccta cggcgctaac  
28501 aaagaaggca tcgtaggggt tgcaactgag ggagccttga atacaccaa agaccacatt  
28561 gccaccgca atcctaataa caatgctgcc accgtgctac aacttctca aggaacaaca  
28621 ttgcaaaaag gcttctacgc agagggaagc agaggcggca gtcaagcctc tttctgctcc  
28681 tcatcacgta tgcgcggtaa ttcaagaaat tcaactcctg gcagcagtag gggaaattct  
28741 cctgctcga tggctagcgg aggtggtgaa actgccctcg cgctattgct gctagacaga  
28801 ttgaaccagc ttgagagcaa agtttctggt aaaggccaac aacaacaagg ccaactgtc  
28861 actaagaaat ctgctgctga ggcattcaaa aagcctgcc aaaaacgtac tgcacaaaa  
28921 cagtacaacg tcaactaacg atttgggaga cgtggtccag aacaacccta aggaaattc  
28981 ggggaccaag acctaatcag acaaggaaact gattacaac attgggcga aattgcaca  
29041 tttgctccaa gtgcctctgc attctttgga atgtcacgca ttggcatgga agtcacacct  
29101 tcgggaacat ggctgactta tcatggagcc attaaattgg atgacaaaaga tccacaattc  
29161 aaagacaacg tcaactgct gaacaagcac attgacgat acaaaacatt cccacaaca  
29221 gagcctaaaa aggcaaaaaa gaaaaagact gatgaagctc agcctttgcc gcagagacaa  
29281 aagaagcagc ccaactgtgac tcttcttctc gcgctgaca tggatgattt ctccagacaa  
29341 cttcaaaatt ccatgagtg agcttctgct gattcaactc aggcataaac actcatgatg  
29401 accacacaag gcagatgggc tatgtaaacg ttttgcgaat tccgtttacg atacatagtc  
29461 tactcttgtg cagaatgaat tctcgtaact aaacagcaca agtaggttta gttacttta  
29521 atctcacata gcaactttta atcaatgtgt aacattaggg aggacttgaa agtagccaca  
29581 cattttcatc gaggccacgc ggagtacgat cgagggtaca gtgaataatg ctaggagag  
29641 ctgcctatat ggaagagccc taatgtgtaa aattaatttt agtagtgcta tccccatgtg  
29701 attttaatag cttcttagga gaatgac

