Venezuelan equine encephalitis virus, complete genome

NCBI Reference Sequence: NC_001449.1

FASTA Graphics

Go to:

LOCUS NC_001449 11444 bp ss-RNA linear VRL 13-AUG-2018

DEFINITION Venezuelan equine encephalitis virus, complete genome.

ACCESSION NC_001449

VERSION NC_001449.1

DBLINK BioProject: PRJNA485481

KEYWORDS RefSeq.

SOURCE Venezuelan equine encephalitis virus (VEEV)

ORGANISM Venezuelan equine encephalitis virus

Viruses; Riboviria; Orthornavirae; Kitrinoviricota; Alsuviricetes; Martellivirales; Togaviridae; Alphavirus.

REFERENCE 1 (bases 1 to 11444)

AUTHORS Kinney,R.M., Tsuchiya,K.R., Sneider,J.M. and Trent,D.W.

TITLE Genetic evidence that epizootic Venezuelan equine encephalitis (VEE) viruses may have evolved from enzootic VEE subtype I-D virus

JOURNAL Virology 191 (2), 569-580 (1992)

PUBMED 1448915

REFERENCE 2 (bases 1 to 11444)

CONSRTM NCBI Genome Project

TITLE Direct Submission

JOURNAL Submitted (01-AUG-2000) National Center for Biotechnology Information, NIH, Bethesda, MD 20894, USA

REFERENCE 3 (bases 1 to 11444)

AUTHORS Kinney,R.

TITLE Direct Submission

JOURNAL Submitted (11-JUL-1993) Molecular Virology, Centers for Disease Control and Prevention, PO Box 2087, Fort Collins, CO 80522, USA

COMMENT VALIDATED REFSEQ: This record has undergone validation or preliminary review. The reference sequence was derived from L04653.

COMPLETENESS: full length.FEATURES Location/Qualifiers

source 1..11444

/organism="Venezuelan equine encephalitis virus"

/mol_type="genomic RNA"

/db_xref="taxon:11036"

/tissue_lib="pUC18 and M13 of Kinney et al." gene

1..11444

/locus_tag="VEEVgp1"

/db_xref="GeneID:2652925" mRNA 1..11444

/locus_tag="VEEVgp1"

/function="genomic mRNA of VEE virus as determined from

strain P676"

/db_xref="GeneID:2652925" gene 1..7526

/gene="NS"

/locus_tag="VEEVgp2"

/db_xref="GeneID:2652923" 5'UTR 1..44

/gene="NS"

/locus_tag="VEEVgp2" CDS 45..7526

/gene="NS"

/locus_tag="VEEVgp2"

/note="possible incorporation of arginine, cysteine or

tryptophan at read through of UGA codon"

/codon_start=1

/transl_except=(pos:5682..5684,aa:Arg)

/product="non-structural polyprotein precursor P1234"

/protein_id="NP_040822.1"

/db_xref="GeneID:2652923"

/translation="MEKVHVDIEEDSPFLRALQRSFPQFEVEAKQVTDNDHANARAFS

HLASKLIETEVDPSDTILDIGSAPARRMYSKHKYHCICPMRCAEDPDRLYKYATKLKK

NCKEITDKELDKKMKELAAVMSDPDLETETMCLHDDESCRYEGQVAVYQDVYAVDGPT

SLYHQANKGVRVAYWIGFDTTPFMFKNLAGAYPSYSTNWADETVLTARNIGLCSSDVM

ERSRRGMSILRKKYLKPSNNVLFSVGSTIYHEKRDLLRSWHLPSVFHLRGKQNYTCRC

ETIVSCDGYVVKRIAISPGLYGKPSGYAATMHREGFLCCKVTDTLNGERVSFPVCTYV

PATLCDQMTGILATDVSADDAQKLLVGLNQRIVVNGRTQRNTNTMKNYLLPVVAQAFA

RWAKEYKEDQEDERPLGLRDRQLVMGCCWAFRRHKITSIYKRPDTQTIIKVNSDFHSF

VLPRIGSNTLEIGLRTRIRKMLEEHKEPSPLITAEDIQEAKCAADEAKEVREAEELRA

ALPPLAADFEEPTLEADVDLMLQEAGAGSVETPRGLIKVTSYAGEDKIGSYAVLSPQA

VLKSEKLSCIHPLAEQVIVITHSGRKGRYAVEPYHGKVVVPEGHAIPVQDFQALSESA

TIVYNEREFVNRYLHHIATHGGALNTDEEYYKTVKPSEHDGEYLYDIDRKQCVKKELV

TGLGLTGELVDPPFHEFAYESLRTRPAAPYQVPTIGVYGVPGSGKSGIIKSAVTKKDL

VVSAKKENCAEIIRDVKKMKGLDVNARTVDSVLLNGCKHPVETLYIDEAFACHAGTLR

ALIAIIRPKKAVLCGDPKQCGFFNMMCLKVHFNHEICTQVFHKSISRRCTKSVTSVVS

TLFYDKRMRTTNPKETKIVIDTTGSTKPKQDDLILTCFRGWVKQLQIDYKGNEIMTAA

ASQGLTRKGVYAVRYKVNENPLYAPTSEHVNVLLTRTEDRIVWKTLAGDPWIKILTAK

YPGNFTATIEEWQAEHDAIMRHILERPDPTDVFQNKANVCWAKALVPVLKTAGIDMTT

EQWNTVDYFETDKAHSAEIVLNQLCVRFFGLDLDSGLFSAPTVPLSIRNNHWDNSPSP

NMYGLNKEVVRQLSRRYPQLPRAVATGRVYDMNTGTLRNYDPRINLVPVNRRLPHALV

LHHNEHPQSDFSSFVSKLKGRTVLVVGEKLSVPGKKVDWLSDQPEATFRARLDLGIPG

DVPKYDIVFINVRTPYKYHHYQQCEDHAIKLSMLTKKACLHLNPGGTCVSIGYGYADR

ASESIIGAIARQFKFSRVCKPKSSHEETEVLFVFIGYDRKARTHNPYKLSSTLTNIYT

GSRLHEAGCAPSYHVVRGDIATATEGVIINAANSKGQPGGGVCGALYKKFPESFDLQP

IEVGKARLVKGAAKHIIHAVGPNFNKVSEVEGDKQLAEAYESIAKIVNDNNYKSVAIP

LLSTGIFSGNKDRLTQSLNHLLTALDTTDADVAIYCRDKKWEMTLKEAVARREAVEEI

CISDDSSVTEPDAELVRVHPKSSLAGRKGYSTSDGKTFSYLEGTKFHQAAKDIAEINA

MWPVATEANEQVCMYILGESMSSIRSKCPVEESEASTPPSTLPCLCIHAMTPERVQRL

KASRPEQITVCSSFPLPKYRITGVQKIQCSQPILFSPKVPAYIHPRKYLVETPPVEET

PESPAENQSTEGTPEQPALVNVDATRTRMPEPIIIEEEEEDSISLLSDGPTHQVLQVE

ADIHGSPSVSSSSWSIPHASDFDVDSLSILDTLDGASVTSGAVSAETNSYFARSMEFR

ARPVPAPRTVFRNPPHPAPRTRTPPLAHSRASSRTSLVSTPPGVNRVITREELEALTP

SRAPSRSASRTSLVSNPPGVNRVITREEFEAFVAQQQRRFDAGAYIFSSDTGQGHLQQ

KSVRQTVLSEVVLERTELEISYAPRLDQEKEELLRKKLQLNPTPANRSRYQSRRVENM

KAITARRILQGLGHYLKAEGKVECYRTLHPVPLYSSSVNRAFSSPKVAVEACNAMLKE

NFPTVASYCIIPEYDAYLDMVDGASCCLDTASFCPAKLRSFPKKHSYLEPTIRSAVPS

AIQNTLQNVLAAATKRNCNVTQMRELPVLDSAAFNVECFKKYACNNEYWETFKENPIR

LTEENVVNYITKLKGPKAAALFAKTHNLNMLQDIPMDRFVMDLKRDVKVTPGTKHTEE

RPKVQVIQAADPLATADLCGIHRELVRRLNAVLLPNIHTLFDMSAEDFDAIIAEHFQP

GDCVLETDIASFDKSEDDAMALTALMILEDLGVDAELLTLIEAAFGEISSIHLPTKTK

FKFGAMMKSGMFLTLFVNTVINIVIASRVLRERLTGSPCAAFIGDDNIVKGVKSDKLM

ADRCATWLNMEVKIIDAVVGEKAPYFCGGFILCDSVTGTACRVADPLKRLFKLGKPLA

VDDEHDDDRRRALHEESTRWNRVGILPELCKAVESRYETVGTSIIVMAMTTLASSVKS

FSYLRGAPITLYG" mat_peptide 45..1649

/gene="NS"

/locus_tag="VEEVgp2"

/product="mRNA-capping enzyme nsP1"

/function="minus strand RNA synthesis; methyltransferase;

guanyltransferase"

/protein_id="NP_740696.1" mat_peptide 1650..4031

/gene="NS"

/locus_tag="VEEVgp2"

/product="protease nsP2"

/function="RNA helicase; nonstructural proteinase;

necessary for subgenomic 26S mRNA synthesis"

/function="replication"

/protein_id="NP_740697.1" mat_peptide 4032..5702

/gene="NS"

/locus_tag="VEEVgp2"

/product="non-structural protein nsp3"

/function="replication"

/protein_id="NP_740698.1" mat_peptide 5703..7520

/gene="NS"

/locus_tag="VEEVgp2"

/product="RNA-directed RNA polymerase nsP4"

/function="RNA polymerase"

/function="replication"

/protein_id="NP_740699.1" CDS 45..5684

/gene="NS"

/locus_tag="VEEVgp2"

/codon_start=1

/product="non-structural polyprotein precursor P123"

/protein_id="NP_040823.1"

/db_xref="GeneID:2652923"

/translation="MEKVHVDIEEDSPFLRALQRSFPQFEVEAKQVTDNDHANARAFS
HLASKLIETEVDPSDTILDIGSAPARRMYSKHKYHCICPMRCAEDPDRLYKYATKLKK
NCKEITDKELDKKMKELAAVMSDPDLETETMCLHDDESCRYEGQVAVYQDVYAVDGPT
SLYHQANKGVRVAYWIGFDTTPFMFKNLAGAYPSYSTNWADETVLTARNIGLCSSDVM
ERSRRGMSILRKKYLKPSNNVLFSVGSTIYHEKRDLLRSWHLPSVFHLRGKQNYTCRC
ETIVSCDGYVVKRIAISPGLYGKPSGYAATMHREGFLCCKVTDTLNGERVSFPVCTYV
PATLCDQMTGILATDVSADDAQKLLVGLNQRIVVNGRTQRNTNTMKNYLLPVVAQAFA
RWAKEYKEDQEDERPLGLRDRQLVMGCCWAFRRHKITSIYKRPDTQTIIKVNSDFHSF
VLPRIGSNTLEIGLRTRIRKMLEEHKEPSPLITAEDIQEAKCAADEAKEVREAEELRA
ALPPLAADFEEPTLEADVDLMLQEAGAGSVETPRGLIKVTSYAGEDKIGSYAVLSPQA
VLKSEKLSCIHPLAEQVIVITHSGRKGRYAVEPYHGKVVVPEGHAIPVQDFQALSESA
TIVYNEREFVNRYLHHIATHGGALNTDEEYYKTVKPSEHDGEYLYDIDRKQCVKKELV
TGLGLTGELVDPPFHEFAYESLRTRPAAPYQVPTIGVYGVPGSGKSGIIKSAVTKKDL
VVSAKKENCAEIIRDVKKMKGLDVNARTVDSVLLNGCKHPVETLYIDEAFACHAGTLR
ALIAIIRPKKAVLCGDPKQCGFFNMMCLKVHFNHEICTQVFHKSISRRCTKSVTSVVS
TLFYDKRMRTTNPKETKIVIDTTGSTKPKQDDLILTCFRGWVKQLQIDYKGNEIMTAA
ASQGLTRKGVYAVRYKVNENPLYAPTSEHVNVLLTRTEDRIVWKTLAGDPWIKILTAK
YPGNFTATIEEWQAEHDAIMRHILERPDPTDVFQNKANVCWAKALVPVLKTAGIDMTT
EQWNTVDYFETDKAHSAEIVLNQLCVRFFGLDLDSGLFSAPTVPLSIRNNHWDNSPSP
NMYGLNKEVVRQLSRRYPQLPRAVATGRVYDMNTGTLRNYDPRINLVPVNRRLPHALV
LHHNEHPQSDFSSFVSKLKGRTVLVVGEKLSVPGKKVDWLSDQPEATFRARLDLGIPG
DVPKYDIVFINVRTPYKYHHYQQCEDHAIKLSMLTKKACLHLNPGGTCVSIGYGYADR
ASESIIGAIARQFKFSRVCKPKSSHEETEVLFVFIGYDRKARTHNPYKLSSTLTNIYT
GSRLHEAGCAPSYHVVRGDIATATEGVIINAANSKGQPGGGVCGALYKKFPESFDLQP
IEVGKARLVKGAAKHIIHAVGPNFNKVSEVEGDKQLAEAYESIAKIVNDNNYKSVAIP
LLSTGIFSGNKDRLTQSLNHLLTALDTTDADVAIYCRDKKWEMTLKEAVARREAVEEI
CISDDSSVTEPDAELVRVHPKSSLAGRKGYSTSDGKTFSYLEGTKFHQAAKDIAEINA
MWPVATEANEQVCMYILGESMSSIRSKCPVEESEASTPPSTLPCLCIHAMTPERVQRL
KASRPEQITVCSSFPLPKYRITGVQKIQCSQPILFSPKVPAYIHPRKYLVETPPVEET
PESPAENQSTEGTPEQPALVNVDATRTRMPEPIIIEEEEEDSISLLSDGPTHQVLQVE
ADIHGSPSVSSSSWSIPHASDFDVDSLSILDTLDGASVTSGAVSAETNSYFARSMEFR
ARPVPAPRTVFRNPPHPAPRTRTPPLAHSRASSRTSLVSTPPGVNRVITREELEALTP

SRAPSRSASRTSLVSNPPGVNRVITREEFEAFVAQQQ" misc_feature 7524..7561

/locus_tag="VEEVgp1"

/note="putative"

/function="noncoding segment between nonstructural and structural genes" gene 7532..11444

/gene="26S mRNA"

/locus_tag="VEEVgp3"

/db_xref="GeneID:2652924" mRNA 7532..11444

/gene="26S mRNA"

/locus_tag="VEEVgp3"

/product="26S mRNA region"

/note="The structural proteins of the virus are translated as apolyprotein precursor molecule from an intracellular, 26SmRNA species that is identical to the 3'-one-third portion of the viral genomic mRNA; putative"

/function="intracellular, subgenomic viral mRNA species"

/db_xref="GeneID:2652924" CDS 7562..11329

/gene="26S mRNA"

/locus_tag="VEEVgp3"

/codon_start=1

/product="structural polyprotein precursor"

/protein_id="NP_040824.1"

/db_xref="GeneID:2652924"

/translation="MFPFQPMYPMQPMPYRNPFAAPRRPWFPRTDPFLAMQVQELTRS
MANLTFKQRRDAPPEGPPAKKPKREAPQKQKGGGQGKKKKNQGKKKAKTGPPNPKAQS
GNKKKPNKKPGKRQRMVMKLESDKTFPIMLEGKINGYACVVGGKLFRPMHVEGKIDND
VLAALKTKKASKYDLEYADVPQNMRADTFKYTHEKPQGYYSWHHGAVQYENGRFTVPK
GVGAKGDSGRPILDNQGRVVAIVLGGVNEGSRTALSVVMWNEKGVTVKYTPENCEQWS
LVTTMCLLANVTFPCAEPPICYDRKPAETLAMLSVNVDNPGYDELLEAAVKCPGRKRR
STEELFKEYKLTRPYMARCIRCAVGSCHSPIAIEAVKSDGHDGYVRLQTSSQYGLDSS
GNLKGRTMRYDMHGTIEEIPLHQVSLHTSRPCHIVDGHGYFLLARCPAGDSITMEFKK
GSVTHSCSVPYEVKFNPVGRELYTHPPEHGAEQACQVYAHDAQNRGAYVEMHLPGSEV
DSSLISLSGSSVTVTPPVGTSALVKCKCGGTKISETINKAKQFSQCTKKEQCRAYRLQ

NDKWVYNSDKLPKAAGATLKGKLHVPFLLADGKCTVPLAPEPMITFGFRSVSLKLHPK
NPTYLTTRQLADEPHYTHELISEPAVRNFTVTEKGWEFVWGNHPPKRFWAQETAPGNP
HGLPHEVITHYYHRYPMSTILGLSICAAIVTVSVAASTWLFCKSRVSCLTPYRLTPNA
RMPLCLAVLCCARTARAETTWESLDHLWNNNQQMFWIQLLIPLAALIVVTRLLKCVCC
VVPFLVVAGAAGAGAYEHATTMPSQAGISYNTIVNRAGYAPLPISITPTKIKLIPTVN
LEYVTCHYKTGMDSPAIKCCGSQECTPTNRPDEQCKVFTGVYPFMWGGAYCFCDTENT
QVSKAYVMKSDDCLADHAEAYKAHTASVQAFLNITVGEHSIVTTVYVNGETPVNFNGV
KLTAGPLSTAWTPFDRKIVQYAGEIYNYDFPEYGAGQPGAFGDIQSRTVSSSDLYANT
NLVLQRPKAGAIHVPYTQAPSGFEQWKKDKAPSLKFTAPFGCEIYTNPIRAENCAVGS
IPLAFDIPDALFTRVSETPTLSAAECTLNECVYSSDFGGIATVKYSASKSGKCAVHVP
SGTATLKEAAVELTEQGSATIHFSTANIHPEFRLQICTSYVTCKGDCHPPKDHIVTHP
QYHAQTFTAAVSKTAWTWLTSLLGGSAVIIIIGLVLATIVAMYVLTNQKHN"

mat_peptide 7562..8386

/gene="26S mRNA"

/locus_tag="VEEVgp3"

/product="core protein"

/function="encloses the genomic mRNA molecule"

/standard_name="capsid protein"

/note="putative"

/protein_id="NP_740700.1" mat_peptide 8387..8563

/gene="26S mRNA"

/locus_tag="VEEVgp3"

/product="E3 envelope protein"

/standard_name="E3 protein"

/note="putative"

/protein_id="NP_741965.1" mat_peptide 8564..9832

/gene="26S mRNA"

/locus_tag="VEEVgp3"

/product="E2 envelope glycoprotein"

/function="cell attachment, elicits virus neutralizing
antibodies"

/standard_name="E2 envelope glycoprotein"

/note="putative"

/protein_id="NP_741966.1" mat_peptide 9833..10000

/gene="26S mRNA"

/locus_tag="VEEVgp3"

/product="6K membrane protein"

/protein_id="NP_818994.1" mat_peptide 10001..11326

/gene="26S mRNA"

/locus_tag="VEEVgp3"

/product="E1 envelope glycoprotein"

/function="viral fusion, immunogenecity"

/standard_name="E1 envelope glycoprotein"

/note="putative"

/protein_id="NP_741967.1" gene 7562..10047

/locus_tag="VEEVgp4"

/db_xref="GeneID:13165423" CDS
join(7562..9970,9970..10047)

/locus_tag="VEEVgp4"

/note="Truncated verstion of structural polyprotein that
will be produced when frameshifting occurs at nt 9970;
Truncated version of structural polyprotein and transframe
fusion protein have been added by NCBI Staff with the kind
help of Dr. David Karlin (Department of Zoology,
University of Oxford, UK) and Dr. Andrew Firth (Department
of Pathology, University of Cambridge, UK)"

/codon_start=1

/product="truncated polyprotein"

/protein_id="YP_006491235.1"

/db_xref="GeneID:13165423"

/translation="MFPFQPMYPMQPMPYRNPFAAPRRPWFPRTDPFLAMQVQELTRS
MANLTFKQRRDAPPEGPPAKKPKREAPQKQKGGGQGKKKKNQGKKKAKTGPPNPKAQS
GNKKKPNKKPGKRQRMVMKLESDKTFPIMLEGKINGYACVVGGKLFRPMHVEGKIDND
VLAALKTKKASKYDLEYADVPQNMRADTFKYTHEKPQGYYSWHHGAVQYENGRFTVPK
GVGAKGDSGRPILDNQGRVVAIVLGGVNEGSRTALSVVMWNEKGVTVKYTPENCEQWS
LVTTMCLLANVTFPCAEPPICYDRKPAETLAMLSVNVDNPGYDELLEAAVKCPGRKRR
STEELFKEYKLTRPYMARCIRCAVGSCHSPIAIEAVKSDGHDGYVRLQTSSQYGLDSS
GNLKGRTMRYDMHGTIEEIPLHQVSLHTSRPCHIVDGHGYFLLARCPAGDSITMEFKK

GSVTHSCSVPYEVKFNPVGRELYTHPPEHGAEQACQVYAHDAQNRGAYVEMHLPGSEV

DSSLISLSGSSVTVTPPVGTSALVKCKCGGTKISETINKAKQFSQCTKKEQCRAYRLQ

NDKWVYNSDKLPKAAGATLKGKLHVPFLLADGKCTVPLAPEPMITFGFRSVSLKLHPK

NPTYLTTRQLADEPHYTHELISEPAVRNFTVTEKGWEFVWGNHPPKRFWAQETAPGNP

HGLPHEVITHYYHRYPMSTILGLSICAAIVTVSVAASTWLFCKSRVSCLTPYRLTPNA

RMPLCLAVLCCARTARAETTWESLDHLWNNNQQMFWIQLLIPLAALIVVTRLLKCVCC

VVPFLSRGRRRRRRRLRARDHDAEPSGNLV" mat_peptide
join(9833..9970,9970..10044)

/locus_tag="VEEVgp4"

/product="transframe fusion protein"

/note="Transframe fusion (TF) protein expressed via

programmed ribosomal frameshifting. TF protein presumably

plays a stabilizing role in the virion structure."

/protein_id="YP_006491236.1" 3'UTR 11327..11444

/gene="26S mRNA"

/locus_tag="VEEVgp3" polyA_site 11444

/gene="26S mRNA"

/locus_tag="VEEVgp3"ORIGIN

```
      1 atgggcggcg caagagagaa gcccaaacca ttacctacc caaaatggag aaagttcacg
     61 ttgacatcga ggaagacagc ccattcctca gagctttaca acggagcttc ccgcagtttg
    121 aggtagaagc caagcaggtc actgataatg accatgctaa tgccagagcg ttttcgcatc
    181 tggcttcaaa actgatcgaa acggaggtgg acccatccga cacgatcctt gacattggaa
    241 gtgcgcccgc ccgcagaatg tattctaagc ataagtatca ttgcatctgt ccgatgagat
    301 gtgcggaaga tccggacaga ttgtacaagt atgcaactaa gctgaagaaa aattgcaagg
    361 aaataactga caaggaattg gacaagaaaa tgaaggagct cgccgccgtc atgagcgacc
    421 ctgacctgga aactgagact atgtgcctcc acgacgatga gtcatgtcgc tacgaggggc
    481 aagtcgctgt ttaccaggat gtatacgcag ttgacggacc gacaagtctc tatcaccaag
    541 ccaacaaggg agttagagtc gcctactgga taggctttga caccacccct tttatgttta
    601 agaacttggc tggagcatat ccatcatact ctaccaactg ggccgacgaa accgtgttaa
    661 cggctcgtaa cataggccta tgcagctccg acgtcatgga gcggtcacgt agagggatgt
    721 ccattcttag gaagaagtat ttgaaaccat ccaataatgt cctattctct gttggctcga
    781 ccatctacca cgagaagagg gacttactga ggagctggca cctgccgtct gtatttcact
    841 tacgtggcaa gcaaaattac acatgtcggt gtgagactat agttagttgc gacgggtacg
    901 tcgttaaaag aatagctatc agtccaggcc tgtatgggaa gccttcaggc tatgctgcta
```

961 cgatgcaccg cgagggattc ttgtgctgca aagtgacaga cacattgaac ggggagaggg

1021 tctcttttcc cgtgtgcacg tatgtgccag ctacattgtg tgaccaaatg actggcatac

1081 tggcaacaga tgtcagtgcg gacgacgcgc aaaaactgct ggttgggctc aaccagcgca

1141 tagtcgtcaa cggtcgcacc caaagaaaca ccaataccat gaagaattat cttttgcccg

1201 tagtggccca ggcatttgct aggtgggcaa aggaatataa ggaagatcaa gaagatgaga

1261 ggccactagg actacgagat agacagttag tcatggggtg ctgctgggct tttagaaggc

1321 acaagataac atctatttat aagcgcccag atacccaaac catcatcaaa gtgaacagcg

1381 atttccactc attcgtgctg cccaggatag gcagtaacac actggagatc gggctgagaa

1441 cgagaatcag gaaaatgcta gaagagcaca aggagccgtc acctctcatt actgccgagg

1501 acatacaaga ggctaagtgc gcagccgatg aggctaagga agtgcgtgaa gccgaggagc

1561 tgcgcgctgc tctaccacct ttggcagctg attttgagga gcccactctg gaagccgatg

1621 tcgacttgat gttacaagag gctggggccg gctcagtgga gacacctcgt ggcttgataa

1681 aggttaccag ctatgccggc gaggacaaga tcggctctta cgcagtgctt tctccacagg

1741 ctgtactcaa gagtgagaaa ctatcttgca ttcaccctct cgctgaacaa gtcatagtga

1801 taacacactc tggccgaaaa gggcgttatg ccgtggaacc ctaccatgga aaagtagtgg

1861 tgccagaggg acatgcaata cccgtccagg actttcaagc tctgagtgaa agtgccacca

1921 tcgtgtacaa cgaacgagag ttcgtaaaca ggtacctgca ccatattgcc acacatggag

1981 gagcgctgaa cacagatgaa gaatattaca aaactgtcaa gcccagcgag cacgacggcg

2041 aatacctgta cgacatcgac aggaaacaat gcgtcaagaa agaattagtc actgggctag

2101 ggcttacagg cgagctggtg gatcctccct tccatgaatt tgcctacgag agtctgagaa

2161 cacgtccggc cgctccttac caagtaccaa ccataggggt gtatggcgtg ccggggtcag

2221 gcaagtctgg catcattaaa agcgcagtca ccaaaaaaga tctggtggtg agcgccaaga

2281 aagaaaactg cgcagaaata ataagggacg tcaagaaaat gaaagggctg gacgtcaatg

2341 ccagaactgt ggactcagtg ctcttgaatg gatgcaaaca ccccgtagag accctgtata

2401 ttgacgaagc ttttgcttgt catgcaggca ctctcagagc gctcatagcc atcataagac

2461 ctaaaaaggc agtgctctgc ggggatccaa aacagtgtgg ctttttcaat atgatgtgcc

2521 tgaaagtgca ttttaaccac gagatttgca cgcaggtctt ccacaaaagc atctctcgcc

2581 gttgcactaa atccgtgact tcggtcgtct caaccttgtt ttacgacaaa aggatgagaa

2641 cgacgaaccc gaaagagact aagattgtga ttgacactac tggcagtacc aaaccgaagc

2701 aggacgatct cattctcact tgtttcagag ggtgggtgaa gcagttgcaa atagattaca

2761 aaggcaacga aataatgacg gcagctgcct ctcaagggct gacccgtaaa ggcgtgtatg

2821 ccgttcggta caaggtgaat gaaaatcccc tgtacgcacc cacctcagaa catgtgaacg

2881 tcctactgac ccgcacggag gaccgtatcg tgtggaaaac actagccggt gatccatgga

2941 taaaaatact gacggccaag tatcctggga acttcactgc cacgatagag gaatggcaag

3001 cagagcatga tgccatcatg aggcacatct tggagagacc ggaccctacc gacgttttcc

3061 aaaataaggc gaacgtgtgt tgggccaagg ctttggtgcc ggtactgaag actgcaggca

3121 tagacatgac cactgaacaa tggaacactg tggattactt cgaaacggac aaagctcact

3181 cagcagagat agtattgaac caactatgcg tgaggttctt tggactcgac ctggactccg

3241 gtctattttc tgcacccact gttccgttat ccattaggaa taatcactgg gataattccc

3301 cgtcgcctaa catgtacggg ttgaataaag aagtggtccg ccagctctcc cgcaggtacc

3361 cacaactgcc tcgagcagtt gccaccggaa gagtctatga catgaacact ggcacgctgc

3421 gcaattatga tccgcgcata aatctagtac ctgtgaacag aagactgcct catgctttag

3481 tcctccacca taatgaacac ccacagagtg acttttcttc attcgtcagc aaactgaagg

3541 gcagaactgt cttggtggtc ggggagaagt tgtccgtccc aggcaaaaag gtcgactggt

3601 tgtcagacca gcctgaggct acctttagag ctcggctgga tttaggtatc ccaggtgacg

3661 tgcccaaata cgacattgta tttattaacg tgaggactcc atataaatac catcattatc

3721 agcagtgtga agaccacgcc attaagctta gtatgttgac caagaaagct tgtctgcatt

3781 tgaatcccgg cggaacctgc gtcagcatag gttatggtta cgctgacagg gccagcgaga

3841 gcatcattgg tgctatagcg cggcagttca agttctcccg ggtatgcaaa ccgaaatcct

3901 cacatgaaga gacagaagta ctgtttgtat tcattgggta cgatcgcaag gcccgtacgc

3961 acaatcctta caagctttca tctaccttga ccaacatcta tacaggttcc agactccacg

4021 aagccggatg cgcaccctca tatcatgtgg tgcgagggga tattgccacg gccaccgaag

4081 gagtgatcat aaatgctgct aacagcaaag gacaacctgg cggaggggtg tgcggagcgc

4141 tgtataagaa attcccggaa agcttcgatt tacagccgat cgaagtagga aaagcgcgac

4201 tggtcaaagg tgcagctaaa catatcattc atgccgtagg accaaacttc aacaaagttt

4261 cggaagttga aggggacaaa cagttggcag aggcttatga gtccatcgct aaaattgtca

4321 acgataacaa ttacaagtca gtagcgattc cactgttgtc caccggcatc ttttccggga

4381 acaaagatcg actaacccaa tcattgaacc atttgctgac agctttagac accactgatg

4441 cagatgtagc catatactgc agggacaaga aatgggaaat gactctcaag gaagcagtgg

4501 ctaggagaga agcagtggag gagatatgca tatcagacga ctcttcggtg acagaaccgg

4561 atgcagagct ggtgagggta catccgaaga gttctttggc tggaaggaag ggctacagca

4621 caagtgatgg caagactttc tcatatttgg aagggaccaa atttcaccag gcggccaagg

4681 atatagcaga aattaatgcc atgtggccag ttgcaacgga ggccaatgag caagtatgca

4741 tgtatatcct cggtgaaagc atgagcagca ttaggtcgaa atgccccgtc gaggagtcgg

4801 aagcctccac accacctagc acgctgcctt gcttgtgcat ccatgctatg actccagaaa

4861 gagtacaacg cctaaaagcc tcacgtccag aacaaattac tgtgtgctca tcctttccat

4921 tgccgaagta tagaatcact ggtgtgcaga agatccagtg ctcccagcct atactgttct

4981 caccgaaggt gcctgcgtac attcatccac ggaagtacct cgtggaaaca ccaccggtag

5041 aagagactcc ggagtcgccg gcagagaacc aatccacaga ggggacacct gaacaaccag

5101 cacttgtaaa cgtggatgca accaggacta gaatgcctga accgatcatc attgaagagg

5161 aagaagagga tagtataagt ttgctgtcag acggcccgac ccaccaggtg ctgcaagtcg

5221 aggcagacat tcacgggtcg ccttctgtat ccagctcatc ctggtccatt cctcatgcat

5281 ccgactttga tgtggacagc ttatccatcc ttgacaccct ggatggagct agcgtgacca

5341 gcggggcagt gtcagccgag actaactcct acttcgcaag gagcatggag tttcgggcgc

5401 gaccggtgcc tgcgcctcga accgtattca ggaaccctcc acatcccgca ccgcgcacaa

5461 gaacaccgcc acttgcacac agcagggcca gctcgagaac tagcctagtt tccacccccgc

5521 caggcgtgaa tagggtgatt actagagagg agctcgaggc gcttaccccg tcccgcgctc

5581 ctagcaggtc ggcctcaaga actagcctgg tctctaaccc gccaggcgta aatagggtga

5641 ttacaagaga ggagtttgag gcgttcgtag cacaacaaca atgacggttt gacgcgggtg

5701 catacatctt ttcctccgat accggtcaag ggcatttaca acaaaaatca gtaaggcaaa

5761 cggtgttatc cgaagtggtg ttggagagga ccgaattgga gatttcgtat gccccgcgcc

5821 tcgaccagga aaaagaagaa ctactacgca agaaattaca gctgaatccc acacctgcta

5881 acagaagcag ataccagtcc aggagggtgg agaatatgaa agccataaca gctagacgta

5941 ttctgcaagg cctagggcat tatttgaagg cagaaggaaa agtggagtgc tatcgaaccc

6001 tgcatcctgt tcctttgtat tcatctagtg tgaatcgtgc tttttcaagc cccaaggtcg

6061 cagtggaagc ctgcaatgcc atgctgaaag aaaattttcc gactgtagct tcctactgta

6121 ttattccaga gtacgatgcc tatctggaca tggttgacgg cgcttcttgt tgcttagaca

6181 ctgccagttt ttgccctgcg aagctgcgca gctttccaaa gaaacactcc tatttggaac

6241 ccacaatacg gtcggcagtg ccatcagcga ttcagaacac gctccagaac gtcctggcag

6301 ctgccacaaa aagaaattgc aacgtcacgc aaatgagaga attgcccgta ttggattcgg

6361 ctgcctttaa tgtggaatgc ttcaagaaat atgcgtgcaa taatgaatat tgggaaacgt

6421 ttaaagaaaa ccccatcagg cttactgaag aaaatgtggt aaattacatt actaaattaa

6481 aaggaccaaa agctgctgct ctttttgcga agacacataa tttgaatatg ttacaggaca

6541 taccaatgga caggtttgta atggacttaa agagggacgt gaaagtgact ccaggaacaa

6601 aacatactga agaacggccc aaggtacagg tgattcaggc tgccgatcca ctagcgacag

6661 cggatctgtg cggaatccac cgggagttgg ttaggagatt aaatgctgtc ctgcttccga

6721 acatccatac actgtttgac atgtcggctg aagactttga cgctattatt gccgagcatt

6781 tccagcctgg ggactgtgta ctggaaactg acattgcgtc gtttgataaa agtgaggacg

6841 acgccatggc tctgaccgcg ttaatgattc tggaagacct aggagtggac gcagagctgt

6901 tgacgctgat tgaggcggct ttcggcgaaa tatcatcaat acatttgccc accaaaacta

6961 aatttaaatt cggagccatg atgaaatccg gaatgttcct cacactgttt gtgaacacag

7021 tcatcaacat cgtaatcgca agcagagtgt taagagagcg gctaaccgga tcaccatgtg

7081 cagcattcat tggagatgac aatatcgtga aaggagtcaa atctgacaaa ttaatggcag

7141 acaggtgcgc cacttggttg aacatggaag tcaagatcat agacgccgtg gtgggcgaga

7201 aagcgcccta tttttgtgga gggtttatct tgtgtgactc cgtgaccggc acagcgtgcc

7261 gtgtggcaga ccccctaaaa aggctgttta agcttggcaa acccctggca gtagacgatg

7321 aacatgacga tgacaggaga agggcattac acgaagagtc aacacgctgg aatcgagtgg

7381 gaattcttcc agagctgtgt aaggcagtag aatcaaggta tgaaaccgta ggaacttcca

7441 tcatagttat ggccatgact actctagcta gcagtgttaa atcattcagc tacctgagag

7501 gggcccctat aactctctac ggctaacctg aatggactac gacatagtct agtccgccaa

7561 gatgttcccg ttccaaccaa tgtatccgat gcagccaatg ccctatcgta acccgttcgc

7621 ggccccgcgc aggccctggt tccccagaac cgacccttttt ctggcgatgc aggtgcagga

7681 attaacccgc tcgatggcta acctgacgtt caagcaacgc cgggacgcgc cacctgaggg

7741 gccacctgct aagaaaccta agagggaggc cccgcaaaag caaaaagggg gaggccaagg

7801 gaagaagaag aagaaccagg ggaagaagaa ggccaagacg gggccgccta atccgaaggc

7861 acagagtgga aacaagaaga agcccaacaa gaaaccaggc aagagacagc gcatggtcat

7921 gaaattggaa tctgacaaga cattcccaat tatgctggaa gggaagatta acggctacgc

7981 ttgcgtggtc ggagggaagt tattcaggcc gatgcacgtg gaaggcaaga tcgacaacga

8041 cgttctggcc gcacttaaga cgaagaaagc atccaaatat gatcttgagt atgcagatgt

8101 gccacagaac atgcgggccg atacattcaa gtacacccat gagaagcccc aaggctatta

8161 cagctggcat catggagcag tccaatatga aaatgggcgt ttcacggtgc caaaaggagt

8221 tggggccaag ggagacagcg gaagacccat tctggataat cagggacggg tggtcgctat

8281 tgtgctggga ggtgtgaatg aaggatctag gacagccctt tcagtcgtca tgtggaacga

8341 gaagggagta actgtgaagt atactccgga gaactgcgag caatggtcac tagtgaccac

8401 tatgtgcctg ctcgccaatg tgacgttccc atgtgccgaa ccaccaattt gctacgacag

8461 aaaaccagca gagactttgg ccatgctcag cgttaacgtt gacaacccgg gctacgatga

8521 gctgctggaa gcagctgtta agtgccccgg aagaaaaagg agatctaccg aggagctgtt

8581 taaggagtat aagctaacgc gcccttacat ggccagatgc atcagatgtg ccgttgggag

8641 ctgccatagt ccaatagcaa ttgaggcagt gaagagcgac gggcacgacg gctatgttag

8701 acttcagact tcctcgcagt atggcctgga ttcctctggc aacttaaagg gaaggactat

8761 gcggtatgat atgcacggga ccattgaaga gataccacta catcaagtgt cactccacac

8821 atctcgcccg tgtcacattg tggatgggca tggttatttt ctgcttgcta ggtgcccggc

8881 aggggactcc atcaccatgg aatttaagaa aggttcagtc acacactcct gctcagtgcc

8941 gtatgaagtg aaatttaatc ctgtaggcag agaactctac actcatccac cagaacacgg

9001 agcagagcaa gcgtgccaag tctacgcgca cgatgcacag aacagaggag cttatgtcga

9061 gatgcacctc ccgggctcag aagtggacag cagtttgatt tccttgagcg gcagttcagt

9121 caccgtgaca cctcctgtcg ggactagcgc cttggtgaaa tgcaagtgcg gcggcacaaa

9181 gatctccgaa accatcaaca aggcaaaaca gttcagccag tgcacaaaga aggagcagtg

9241 cagagcatat cgactgcaga atgacaagtg ggtgtataat tctgacaaac tgcccaaagc

9301 agcgggagcc accctaaaag gaaaactaca cgtcccgttc ttgctggcag acggcaaatg

9361 caccgtgcct ctagcaccgg aacctatgat aaccttcggt ttccgatcag tgtcactgaa

9421 actgcaccct aagaatccca catatctgac cactcgccaa cttgctgatg agcctcatta

9481 cacgcacgag ctcatatctg aaccagctgt taggaatttt accgtcactg aaaaggggtg

9541 ggagtttgta tggggaaacc atccgccgaa aaggttttgg gcacaggaaa cagcacccgg

9601 aaatccacat gggctgccac atgaggtgat aactcattat taccacagat accctatgtc

9661 caccatcctg ggtttgtcaa tttgcgccgc cattgtaacc gtttccgttg cagcgtccac

9721 ctggctgttt tgcaaatcca gagtttcgtg cctaactcct taccggctaa cacctaacgc

9781 caggatgccg ctttgcctgg ccgtgctttg ctgcgcccgc actgcccggg ccgagaccac

9841 ctgggagtcc ttggatcacc tatggaacaa taaccaacag atgttctgga ttcaattgct

9901 gatccctctg gccgccttga ttgtagtgac tcgcctgctc aagtgcgtgt gctgtgtagt

9961 gccttttta gtcgtggccg gcgccgcagg cgccggcgcc tacgagcacg cgaccacgat

10021 gccgagccaa gcgggaatct cgtataacac catagtcaac agagcaggct acgcgccact

10081 ccctatcagc ataacaccaa caaagatcaa gctgataccc acagtgaact tggagtacgt

10141 cacctgccac tacaaaacag gaatggattc accagccatc aaatgctgcg gatctcagga

10201 atgtactcca actaacaggc ctgatgaaca gtgcaaagtc ttcacagggg tttacccgtt

10261 catgtgggga ggtgcatatt gcttttgcga cactgagaat actcaggtca gcaaggccta

10321 cgtaatgaaa tctgacgact gccttgcgga tcatgctgaa gcatacaaag cgcacacagc

10381 ctcagtgcag gcgttcctca acatcacagt gggggaacac tctattgtga ccaccgtgta

10441 tgtgaatgga gaaactcctg tgaacttcaa tggggtcaaa ctaactgcag gtccactttc

10501 cacagcttgg acacccttg acagaaaaat cgtgcagtat gccggggaga tctataatta

10561 cgattttcct gagtatgggg caggacaacc aggagcattt ggagacatac aatccagaac

10621 agtctcaagc tcagatctgt atgccaatac caacctagtg ctgcagagac ccaaagcagg

10681 agcgatccat gtgccataca ctcaggcacc atcgggtttt gagcaatgga agaaagataa

10741 agctccgtca ttgaaattca ccgccccttt cggatgcgaa atatatacaa accccattcg

10801 cgccgaaaat tgtgctgtag ggtcaattcc attagccttt gacattcccg acgccttgtt

```
10861 caccagggtg tcagaaacac cgacactttc agcggccgaa tgcactctta acgagtgcgt
10921 gtattcatcc gactttggcg ggatcgccac ggtcaagtat tcggccagca agtcaggcaa
10981 gtgcgcagtc catgtgccat cagggactgc taccctaaaa gaagcagcag tcgagctaac
11041 cgagcaaggg tcggcgacca ttcatttctc gaccgcaaat atccacccgg agttcaggct
11101 ccaaatatgc acatcatatg tcacgtgcaa aggtgattgt cacccccccga aagaccacat
11161 tgtgacacac ccccagtatc acgcccaaac atttacagcc gcggtgtcaa aaaccgcgtg
11221 gacgtggtta acatccctgc tgggaggatc ggccgtaatt attataattg gcttagtgct
11281 ggctactatt gtggccatgt acgtgctgac caaccagaaa cataattgaa catagcagca
11341 attggcaagc tgcttatata gaacttgcgg cgattggcat gccgctttaa aattttattt
11401 tattttcttt tcttttccga atcggatttt gtttttaata tttc
//
```