Dengue virus 1, complete genome

NCBI Reference Sequence: NC_001477.1

FASTA Graphics

Go to:

LOCUS NC_001477 10735 bp ss-RNA linear VRL 03-MAY-2019

DEFINITION Dengue virus 1, complete genome.

ACCESSION NC_001477

VERSION NC_001477.1

DBLINK BioProject: PRJNA485481

KEYWORDS RefSeq.

SOURCE dengue virus type 1

ORGANISM dengue virus type 1

Viruses; Riboviria; Orthornavirae; Kitrinoviricota; Flasuviricetes; Amarillovirales; Flaviviridae; Orthoflavivirus; Euflavivirus; Orthoflavivirus denguei.

REFERENCE 1 (bases 1 to 10735)

AUTHORS Puri,B., Nelson,W.M., Henchal,E.A., Hoke,C.H., Eckels,K.H., Dubois,D.R., Porter,K.R. and Hayes,C.G.

TITLE Molecular analysis of dengue virus attenuation after serial passage in primary dog kidney cells

JOURNAL J. Gen. Virol. 78 (PT 9), 2287-2291 (1997)

PUBMED 9292016

REFERENCE 2 (bases 1 to 10735)

AUTHORS McKee,K.T. Jr., Bancroft,W.H., Eckels,K.H., Redfield,R.R., Summers,P.L. and Russell,P.K.

TITLE Lack of attenuation of a candidate dengue 1 vaccine (45AZ5) in human volunteers

JOURNAL Am. J. Trop. Med. Hyg. 36 (2), 435-442 (1987)

PUBMED 3826504

REFERENCE 3 (bases 1 to 10735)

CONSRTM NCBI Genome Project

TITLE Direct Submission

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

REFERENCE 4 (bases 1 to 10735)

AUTHORS Puri,B. and Nelson,W.M.

TITLE Direct Submission

JOURNAL Submitted (05-FEB-1997) Inf. Dis. Dept, Naval Medical Research Institute, 8901 Wisconsin Ave., Bethesda, MD 20889-5607, USA

COMMENT VALIDATED REFSEQ: This record has undergone validation or preliminary review. The reference sequence was derived from U88536.
COMPLETENESS: full length.

FEATURES Location/Qualifiers

source 1..10735

/organism="dengue virus type 1"
/mol_type="genomic RNA"
/db_xref="taxon:11053"
/clone="45AZ5"
/type="1"

5'UTR 1..94

stem_loop 2..69
/note="stem-loop A (SLA)"

regulatory 70..78
/regulatory_class="other"
/note="oligo U track spacer"

regulatory 79..94
/regulatory_class="promoter"
/note="5' upstream AUG region (UAR)"

stem_loop 79..93
/note="stem-loop B (SLB)"

gene 95..10273
/gene="POLY"
/locus_tag="DV1_gp1"
/gene_synonym="polyprotein gene"
/db_xref="GeneID:5075725"

CDS 95..10273
/gene="POLY"
/locus_tag="DV1_gp1"
/gene_synonym="polyprotein gene"
/codon_start=1
/product="polyprotein"
/protein_id="NP_059433.1"
/db_xref="GeneID:5075725"
/translation="MNNQRKKTGRPSFNMLKRARNRVSTVSQLAKRFSKGLLSGQGPM
KLVMAFIAFLRFLAIPPTAGILARWGSFKKNGAIKVLRGFKKEISNMLNIMNRRKRSV
TMLLMLLPTALAFHLTTRGGEPHMIVSKQERGKSLLFKTSAGVNMCTLIAMDLGELCE
DTMTYKCPRITETEPDDVDCWCNATETWVTYGTCSQTGEHRRDKRSVALAPHVGLGLE
TRTETWMSSEGAWKQIQKVETWALRHPGFTVIALFLAHAIGTSITQKGIIFILLMLVT
PSMAMRCVGIGNRDFVEGLSGATWVDVVLEHGSCVTTMAKDKPTLDIELLKTEVTNPA
VLRKLCIEAKISNTTTDSRCPTQGEATLVEEQDTNFVCRRTFVDRGWGNGCGLFGKGS
LITCAKFKCVTKLEGKIVQYENLKYSVIVTVHTGDQHQVGNETTEHGTTATITPQAPT
SEIQLTDYGALTLDCSPRTGLDFNEMVLLTMKKKSWLVHKQWFLDLPLPWTSGASTSQ
ETWNRQDLLVTFKTAHAKKQEVVVLGSQEGAMHTALTGATEIQTSGTTTIFAGHLKCR
LKMDKLILKGMSYVMCTGSFKLEKEVAETQHGTVLVQVKYEGTDAPCKIPFSSQDEKG
VTQNGRLITANPIVTDKEKPVNIEAEPPFGESYIVVGAGEKALKLSWFKKGSSIGKMF
EATARGARRMAILGDTAWDFGSIGGVFTSVGKLIHQIFGTAYGVLFSGVSWTMKIGIG
ILLTWLGLNSRSTSLSMTCIAVGMVTLYLGVMVQADSGCVINWKGRELKCGSGIFVTN
EVHTWTEQYKFQADSPKRLSAAIGKAWEEGVCGIRSATRLENIMWKQISNELNHILLE

NDMKFTVVVGDVSGILAQGKKMIRPQPMEHKYSWKSWGKAKIIGADVQNTTFIIDGPN
TPECPDNQRAWNIWEVEDYGFGIFTTNIWLKLRDSYTQVCDHRLMSAAIKDSKAVHAD
MGYWIESEKNETWKLARASFIEVKTCIWPKSHTLWSNGVLESEMIIPKIYGGPISQHN
YRPGYFTQTAGPWHLGKLELDFDLCEGTTVVVDEHCGNRGPSLRTTTVTGKTIHEWCC
RSCTLPPLRFKGEDGCWYGMEIRPVKEKEENLVKSMVSAGSGEVDSFSLGLLCISIMI
EEVMRSRWSRKMLMTGTLAVFLLLTMGQLTWNDLIRLCIMVGANASDKMGMGTTYLAL
MATFRMRPMFAVGLLFRRLTSREVLLLTVGLSLVASVELPNSLEELGDGLAMGIMMLK
LLTDFQSHQLWATLLSLTFVKTTFSLHYAWKTMAMILSIVSLFPLCLSTTSQKTTWLP
VLLGSLGCKPLTMFLITENKIWGRKSWPLNEGIMAVGIVSILLSSLLKNDVPLAGPLI
AGGMLIACYVISGSSADLSLEKAAEVSWEEEAEHSGASHNILVEVQDDGTMKIKDEER
DDTLTILLKATLLAISGVYPMSIPATLFVWYFWQKKKQRSGVLWDTPSPPEVERAVLD
DGIYRILQRGLLGRSQVGVGVFQEGVFHTMWHVTRGAVLMYQGKRLEPSWASVKKDLI
SYGGGWRFQGSWNAGEEVQVIAVEPGKNPKNVQTAPGTFKTPEGEVGAIALDFKPGTS
GSPIVNREGKIVGLYGNGVVTTSGTYVSAIAQAKASQEGPLPEIEDEVFRKRNLTIMD
LHPGSGKTRRYLPAIVREAIKRKLRTLVLAPTRVVASEMAEALKGMPIRYQTTAVKSE
HTGKEIVDLMCHATFTMRLLSPVRVPNYNMIIMDEAHFTDPASIAARGYISTRVGMGE
AAAIFMTATPPGSVEAFPQSNAVIQDEERDIPERSWNSGYDWITDFPGKTVWFVPSIK
SGNDIANCLRKNGKRVVQLSRKTFDTEYQKTKNNDWDYVVTTDISEMGANFRADRVID
PRRCLKPVILKDGPERVILAGPMPVTVASAAQRRGRIGRNQNKEGDQYIYMGQPLNND
EDHAHWTEAKMLLDNINTPEGIIPALFEPEREKSAAIDGEYRLRGEARKTFVELMRRG
DLPVWLSYKVASEGFQYSDRRWCFDGERNNQVLEENMDVEIWTKEGERKKLRPRWLDA
RTYSDPLALREFKEFAAGRRSVSGDLILEIGKLPQHLTQRAQNALDNLVMLHNSEQGG
KAYRHAMEELPDTIETLMLLALIAVLTGGVTLFFLSGRGLGKTSIGLLCVIASSALLW
MASVEPHWIAASIILEFFLMVLLIPEPDRQRTPQDNQLAYVVIGLLFMILTVAANEMG
LLETTKKDLGIGHAAAENHHHAAMLDVDLHPASAWTLYAVATTIITPMMRHTIENTTA
NISLTAIANQAAILMGLDKGWPISKMDIGVPLLALGCYSQVNPLTLTAAVLMLVAHYA
IIGPGLQAKATREAQKRTAAGIMKNPTVDGIVAIDLDPVVYDAKFEKQLGQIMLLILC
TSQILLMRTTWALCESITLATGPLTTLWEGSPGKFWNTTIAVSMANIFRGSYLAGAGL
AFSLMKSLGGGRRGTGAQGETLGEKWKRQLNQLSKSEFNTYKRSGIIEVDRSEAKEGL
KRGETTKHAVSRGTAKLRWFVERNLVKPEGKVIDLGCGRGGWSYYCAGLKKVTEVKGY
TKGGPGHEEPIPMATYGWNLVKLYSGKDVFFTPPEKCDTLLCDIGESSPNPTIEEGRT
LRVLKMVEPWLRGNQFCIKILNPYMPSVVETLEQMQRKHGGMLVRNPLSRNSTHEMYW
VSCGTGNIVSAVNMTSRMLLNRFTMAHRKPTYERDVDLGAGTRHVAVEPEVANLDIIG
QRIENIKNEHKSTWHYDEDNPYKTWAYHGSYEVKPSGSASSMVNGVVRLLTKPWDVIP
MVTQIAMTDTTPFGQQRVFKEKVDTRTPKAKRGTAQIMEVTARWLWGFLSRNKKPRIC
TREEFTRKVRSNAAIGAVFVDENQWNSAKEAVEDERFWDLVHRERELHKQGKCATCVY
NMMGKREKKLGEFGKAKGSRAIWYMWLGARFLEFEALGFMNEDHWFSRENSLSGVEGE
GLHKLGYILRDISKIPGGNMYADDTAGWDTRITEDDLQNEAKITDIMEPEHALLATSI
FKLTYQNKVVRVQRPAKNGTVMDVISRRDQRGSGQVGTYGLNTFTNMEAQLIRQMESE
GIFSPSELETPNLAERVLDWLKKHGTERLKRMAISGDDCVVKPIDDRFATALTALNDM
GKVRKDIPQWEPSKGWNDWQQVPFCSHHFHQLIMKDGREIVVPCRNQDELVGRARVSQ
GAGWSLRETACLGKSYAQMWQLMYFHRRDLRLAANAICSAVPVDWVPTSRTTWSIHAH
HQWMTTEDMLSVWNRVWIEENPWMEDKTHVSSWEDVPYLGKREDQWCGSLIGLTARAT
WATNIQVAINQVRRLIGNENYLDFMTSMKRFKNESDPEGALW"

mat_peptide 95..436
/gene="POLY"
/locus_tag="DV1_gp1"
/gene_synonym="polyprotein gene"
/product="anchored capsid protein ancC"
/protein_id="NP_722457.2"
/db_xref="VBRC:35735"
mat_peptide 95..394
/gene="POLY"
/locus_tag="DV1_gp1"
/gene_synonym="polyprotein gene"
/product="capsid protein C"
/note="added by NCBI staff following more recent
annotations of this virus sequence"
/protein_id="NP_722466.2"
/db_xref="VBRC:67793"
mat_peptide 437..934
/gene="POLY"
/locus_tag="DV1_gp1"
/gene_synonym="polyprotein gene"
/product="membrane glycoprotein precursor prM"
/protein_id="NP_733807.2"
mat_peptide 437..709
/gene="POLY"
/locus_tag="DV1_gp1"
/gene_synonym="polyprotein gene"
/product="protein pr"
/note="peptide pr"
/protein_id="YP_009164956.1"
mat_peptide 710..934
/gene="POLY"
/locus_tag="DV1_gp1"
/gene_synonym="polyprotein gene"
/product="membrane glycoprotein M"
/protein_id="NP_722459.2"
mat_peptide 935..2419
/gene="POLY"
/locus_tag="DV1_gp1"
/gene_synonym="polyprotein gene"
/product="envelope protein E"
/protein_id="NP_722460.2"
mat_peptide 2420..3475
/gene="POLY"
/locus_tag="DV1_gp1"

/gene_synonym="polyprotein gene"
/product="nonstructural protein NS1"
/protein_id="NP_722461.1"
/db_xref="VBRC:35739"
mat_peptide 3476..4129
/gene="POLY"
/locus_tag="DV1_gp1"
/gene_synonym="polyprotein gene"
/product="nonstructural protein NS2A"
/protein_id="NP_733808.1"
/db_xref="VBRC:35740"
mat_peptide 4130..4519
/gene="POLY"
/locus_tag="DV1_gp1"
/gene_synonym="polyprotein gene"
/product="nonstructural protein NS2B"
/protein_id="NP_733809.1"
/db_xref="VBRC:35741"
mat_peptide 4520..6376
/gene="POLY"
/locus_tag="DV1_gp1"
/gene_synonym="polyprotein gene"
/product="nonstructural protein NS3"
/protein_id="NP_722463.1"
/db_xref="VBRC:35742"
mat_peptide 6377..6757
/gene="POLY"
/locus_tag="DV1_gp1"
/gene_synonym="polyprotein gene"
/product="nonstructural protein NS4A"
/protein_id="NP_733810.1"
/db_xref="VBRC:35743"
mat_peptide 6758..6826
/gene="POLY"
/locus_tag="DV1_gp1"
/gene_synonym="polyprotein gene"
/product="protein 2K"
/note="added by NCBI staff following more recent
annotations of this virus sequence"
/protein_id="NP_722467.1"
/db_xref="VBRC:67794"
mat_peptide 6827..7573
/gene="POLY"
/locus_tag="DV1_gp1"

/gene_synonym="polyprotein gene"
/product="nonstructural protein NS4B"
/protein_id="NP_733811.1"
/db_xref="VBRC:35744"
mat_peptide 7574..10270
/gene="POLY"
/locus_tag="DV1_gp1"
/gene_synonym="polyprotein gene"
/product="RNA-dependent RNA polymerase NS5"
/protein_id="NP_722465.1"
/db_xref="VBRC:35745"
stem_loop 114..130
/note="capsid region hairpin (cHP)"
regulatory 132..142
/regulatory_class="other"
/note="5' conserved sequence (CS); also called cyclization sequence"
3'UTR 10274..10735
ncRNA 10320..10735
/ncRNA_class="lncRNA"
/product="sfRNA1"
/note="subgenomic flavivirus RNA"
stem_loop 10324..10389
/note="flaviviral nuclease-resistant RNA 1 (fNR1); also called stem-loop 1 or xrRNA1"
ncRNA 10393..10735
/ncRNA_class="lncRNA"
/product="sfRNA2"
/note="subgenomic flavivirus RNA"
stem_loop 10397..10458
/note="flaviviral nuclease-resistant RNA 2 (fNR2); also called stem-loop 2 or xrRNA2"
ncRNA 10463..10735
/ncRNA_class="lncRNA"
/product="sfRNA3"
/note="subgenomic flavivirus RNA"
stem_loop 10467..10549
/note="dumbbell 1 (DBI); also called xrRNA3"
ncRNA 10551..10735
/ncRNA_class="lncRNA"
/product="sfRNA4"
/note="subgenomic flavivirus RNA"
stem_loop 10551..10633
/note="dumbbell 2 (DBII); also called xrRNA4"

```
     regulatory      10630..10640
                     /regulatory_class="other"
                     /note="3' conserved sequence (CS); also called cyclization
                     sequence"
     stem_loop       10643..10656
                     /note="short hairpin (sHP)"
     regulatory      10654..10670
                     /regulatory_class="promoter"
                     /note="3' upstream AUG region (UAR)"
     stem_loop       10657..10735
                     /note="3' stem-loop (3'SL)"
ORIGIN
        1 agttgttagt ctacgtggac cgacaagaac agtttcgaat cggaagcttg cttaacgtag
       61 ttctaacagt tttttattag agagcagatc tctgatgaac aaccaacgga aaaagacggg
      121 tcgaccgtct ttcaatatgc tgaaacgcgc gagaaaccgc gtgtcaactg tttcacagtt
      181 ggcgaagaga ttctcaaaag gattgctttc aggccaagga cccatgaaat tggtgatggc
      241 ttttatagca ttcctaagat ttctagccat acctccaaca gcaggaattt tggctagatg
      301 gggctcattc aagaagaatg gagcgatcaa agtgttacgg ggtttcaaga aagaaatctc
      361 aaacatgttg aacataatga acaggaggaa aagatctgtg accatgctcc tcatgctgct
      421 gcccacagcc ctggcgttcc atctgaccac ccgaggggga gagccgcaca tgatagttag
      481 caagcaggaa agaggaaaat cacttttgtt taagacctct gcaggtgtca acatgtgcac
      541 ccttattgca atggatttgg gagagttatg tgaggacaca atgacctaca aatgcccccg
      601 gatcactgag acggaaccag atgacgttga ctgttggtgc aatgccacgg agacatgggt
      661 gacctatgga acatgttctc aaactggtga acaccgacga gacaaacgtt ccgtcgcact
      721 ggcaccacac gtagggcttg gtctagaaac aagaaccgaa acgtggatgt cctctgaagg
      781 cgcttggaaa caaatacaaa aagtggagac ctgggctctg agacacccag gattcacggt
      841 gatagccctt tttctagcac atgccatagg aacatccatc acccagaaag ggatcatttt
      901 tattttgctg atgctggtaa ctccatccat ggccatgcgg tgcgtgggaa taggcaacag
      961 agacttcgtg gaaggactgt caggagctac gtgggtggat gtggtactgg agcatggaag
     1021 ttgcgtcact accatggcaa aagacaaacc aacactggac attgaactct tgaagacgga
     1081 ggtcacaaac cctgccgtcc tgcgcaaact gtgcattgaa gctaaaatat caaacaccac
     1141 caccgattcg agatgtccaa cacaaggaga agccacgctg gtggaagaac aggacacgaa
     1201 ctttgtgtgt cgacgaacgt tcgtggacag aggctggggc aatggttgtg ggctattcgg
     1261 aaaaggtagc ttaataacgt gtgctaagtt taagtgtgtg acaaaactgg aaggaaagat
     1321 agtccaatat gaaaacttaa aatattcagt gatagtcacc gtacacactg gagaccagca
     1381 ccaagttgga aatgagacca cagaacatgg aacaactgca accataacac ctcaagctcc
     1441 cacgtcggaa atacagctga cagactacgg agctctaaca ttggattgtt cacctagaac
     1501 agggctagac tttaatgaga tggtgttgtt gacaatgaaa aaaaaatcat ggctcgtcca
     1561 caaacaatgg tttctagact taccactgcc ttggacctcg ggggcttcaa catcccaaga
     1621 gacttggaat agacaagact tgctggtcac atttaagaca gctcatgcaa aaaagcagga
     1681 agtagtcgta ctaggatcac aagaaggagc aatgcacact gcgttgactg gagcgacaga
     1741 aatccaaacg tctggaacga caacaatttt tgcaggacac ctgaaatgca gattaaaaat
     1801 ggataaactg attttaaaag ggatgtcata tgtaatgtgc acagggtcat tcaagttaga
     1861 gaaggaagtg gctgagaccc agcatggaac tgttctagtg caggttaaat acgaaggaac
```

1921 agatgcacca tgcaagatcc ccttctcgtc ccaagatgag aagggagtaa cccagaatgg
1981 gagattgata acagccaacc ccatagtcac tgacaaagaa aaaccagtca acattgaagc
2041 ggagccacct tttggtgaga gctacattgt ggtaggagca ggtgaaaaag ctttgaaact
2101 aagctggttc aagaagggaa gcagtatagg gaaaatgttt gaagcaactg cccgtggagc
2161 acgaaggatg gccatcctgg gagacactgc atgggacttc ggttctatag gaggggtgtt
2221 cacgtctgtg ggaaaactga tacaccagat ttttgggact gcgtatggag ttttgttcag
2281 cggtgtttct tggaccatga agataggaat agggattctg ctgacatggc taggattaaa
2341 ctcaaggagc acgtcccttt caatgacgtg tatcgcagtt ggcatggtca cactgtacct
2401 aggagtcatg gttcaggcgg actcgggatg tgtaatcaac tggaaaggca gagaactcaa
2461 atgtggaagc ggcatttttg tcaccaatga agtccacacc tggacagagc aatataaatt
2521 ccaggccgac tcccctaaga gactatcagc ggccattggg aaggcatggg aggagggtgt
2581 gtgtggaatt cgatcagcca ctcgtctcga gaacatcatg tggaagcaaa tatcaaatga
2641 attaaaccac atcttacttg aaaatgacat gaaatttaca gtggtcgtag gagacgttag
2701 tggaatcttg gcccaaggaa agaaaatgat taggccacaa cccatggaac acaaatactc
2761 gtggaaaagc tggggaaaag ccaaaatcat aggagcagat gtacagaata ccaccttcat
2821 catcgacggc ccaaacaccc cagaatgccc tgataaccaa agagcatgga acatttggga
2881 agttgaagac tatggatttg gaattttcac gacaaacata tggttgaaat tgcgtgactc
2941 ctacactcaa gtgtgtgacc accggctaat gtcagctgcc atcaaggata gcaaagcagt
3001 ccatgctgac atggggtact ggatagaaag tgaaaagaac gagacttgga agttggcaag
3061 agcctccttc atagaagtta agacatgcat ctggccaaaa tcccacactc tatggagcaa
3121 tggagtcctg gaaagtgaga tgataatccc aaagatatat ggaggaccaa tatctcagca
3181 caactacaga ccaggatatt tcacacaaac agcagggccg tggcacttgg gcaagttaga
3241 actagatttt gatttatgtg aaggtaccac tgttgttgtg gatgaacatt gtggaaatcg
3301 aggaccatct cttagaacca caacagtcac aggaaagaca atccatgaat ggtgctgtag
3361 atcttgcacg ttacccccc tacgtttcaa aggagaagac gggtgctggt acggcatgga
3421 aatcagacca gtcaaggaga aggaagagaa cctagttaag tcaatggtct ctgcagggtc
3481 aggagaagtg gacagttttt cactaggact gctatgcata tcaataatga tcgaagaggt
3541 aatgagatcc agatggagca gaaaaatgct gatgactgga acattggctg tgttcctcct
3601 tctcacaatg ggacaattga catggaatga tctgatcagg ctatgtatca tggttggagc
3661 caacgcttca gacaagatgg ggatgggaac aacgtaccta gctttgatgg ccactttcag
3721 aatgagacca atgttcgcag tcgggctact gtttcgcaga ttaacatcta gagaagttct
3781 tcttcttaca gttggattga gtctggtggc atctgtagaa ctaccaaatt ccttagagga
3841 gctaggggat ggacttgcaa tgggcatcat gatgttgaaa ttactgactg attttcagtc
3901 acatcagcta tgggctacct tgctgtcttt aacatttgtc aaaacaactt tttcattgca
3961 ctatgcatgg aagacaatgg ctatgatact gtcaattgta tctctcttcc ctttatgcct
4021 gtccacgact tctcaaaaaa caacatggct tccggtgttg ctgggatctc ttggatgcaa
4081 accactaacc atgtttctta taacagaaaa caaaatctgg ggaaggaaaa gctggcctct
4141 caatgaagga attatggctg ttggaatagt tagcattctt ctaagttcac ttctcaagaa
4201 tgatgtgcca ctagctggcc cactaatagc tggaggcatg ctaatagcat gttatgtcat
4261 atctggaagc tcggccgatt tatcactgga gaaagcggct gaggtctcct gggaagaaga
4321 agcagaacac tctggtgcct cacacaacat actagtggag gtccaagatg atggaaccat
4381 gaagataaag gatgaagaga gagatgacac actcaccatt ctcctcaaag caactctgct
4441 agcaatctca ggggtatacc caatgtcaat accggcgacc ctctttgtgt ggtatttttg
4501 gcagaaaaag aaacagagat caggagtgct atgggacaca cccagccctc cagaagtgga

4561 aagagcagtc cttgatgatg gcatttatag aattctccaa agaggattgt tgggcaggtc
4621 tcaagtagga gtaggagttt ttcaagaagg cgtgttccac acaatgtggc acgtcaccag
4681 gggagctgtc ctcatgtacc aagggaagag actggaacca agttgggcca gtgtcaaaaa
4741 agacttgatc tcatatggag gaggttggag gtttcaagga tcctggaacg cgggagaaga
4801 agtgcaggtg attgctgttg aaccggggaa gaaccccaaa aatgtacaga cagcgccggg
4861 taccttcaag acccctgaag gcgaagttgg agccatagct ctagacttta aacccggcac
4921 atctggatct cctatcgtga acagagaggg aaaaatagta ggtctttatg gaaatggagt
4981 ggtgacaaca agtggtacct acgtcagtgc catagctcaa gctaaagcat cacaagaagg
5041 gcctctacca gagattgagg acgaggtgtt taggaaaaga aacttaacaa taatggacct
5101 acatccagga tcgggaaaaa caagaagata ccttccagcc atagtccgtg aggccataaa
5161 aagaaagctg cgcacgctag tcttagctcc cacaagagtt gtcgcttctg aaatggcaga
5221 ggcgctcaag ggaatgccaa taaggtatca gacaacagca gtgaagagtg aacacacggg
5281 aaaggagata gttgacctta tgtgtcacgc cactttcact atgcgtctcc tgtctcctgt
5341 gagagttccc aattataata tgattatcat ggatgaagca cattttaccg atccagccag
5401 catagcagcc agagggtata tctcaacccg agtgggtatg ggtgaagcag ctgcgatttt
5461 catgacagcc actccccccg gatcggtgga ggcctttcca cagagcaatg cagttatcca
5521 agatgaggaa agagacattc ctgaaagatc atggaactca ggctatgact ggatcactga
5581 tttcccaggt aaaacagtct ggtttgttcc aagcatcaaa tcaggaaatg acattgccaa
5641 ctgtttaaga aagaatggga aacgggtggt ccaattgagc agaaaaactt ttgacactga
5701 gtaccagaaa acaaaaaata acgactggga ctatgttgtc acaacagaca tatccgaaat
5761 gggagcaaac ttccgagccg acagggtaat agacccgagg cggtgcctga aaccggtaat
5821 actaaaagat ggcccagagc gtgtcattct agccggaccg atgccagtga ctgtggctag
5881 cgccgcccag aggagaggaa gaattggaag gaaccaaaat aaggaaggcg atcagtatat
5941 ttacatggga cagcctctaa acaatgatga ggaccacgcc cattggacag aagcaaaaat
6001 gctccttgac aacataaaca caccagaagg gattatccca gccctctttg agccggagag
6061 agaaaagagt gcagcaatag acggggaata cagactacgg ggtgaagcga ggaaaacgtt
6121 cgtggagctc atgagaagag gagatctacc tgtctggcta tcctacaaag ttgcctcaga
6181 aggcttccag tactccgaca gaaggtggtg ctttgatggg gaaaggaaca accaggtgtt
6241 ggaggagaac atggacgtgg agatctggac aaaagaagga gaaagaaaga aactacgacc
6301 ccgctggctg gatgccagaa catactctga cccactggct ctgcgcgaat tcaaagagtt
6361 cgcagcagga agaagaagcg tctcaggtga cctaatatta gaaataggga aacttccaca
6421 acatttaacg caaagggccc agaacgcctt ggacaatctg gttatgttgc acaactctga
6481 acaaggagga aaagcctata gacacgccat ggaagaacta ccagacacca tagaaacgtt
6541 aatgctccta gctttgatag ctgtgctgac tggtggagtg acgttgttct tcctatcagg
6601 aaggggtcta ggaaaaacat ccattggcct actctgcgtg attgcctcaa gtgcactgtt
6661 atggatggcc agtgtggaac cccattggat agcggcctct atcatactgg agttctttct
6721 gatggtgttg cttattccag agccggacag acagcgcact ccacaagaca accagctagc
6781 atacgtggtg ataggtctgt tattcatgat attgacagtg gcagccaatg agatgggatt
6841 actggaaacc acaaagaagg acctggggat tggtcatgca gctgctgaaa accaccatca
6901 tgctgcaatg ctggacgtag acctacatcc agcttcagcc tggactctct atgcagtggc
6961 cacaacaatt atcactccca tgatgagaca cacaattgaa aacacaacgg caaatatttc
7021 cctgacagct attgcaaacc aggcagctat attgatggga cttgacaagg gatggccaat
7081 atcaaagatg gacataggag ttccacttct cgccttgggg tgctattctc aggtgaaccc
7141 gctgacgctg acagcggcgg tattgatgct agtggctcat tatgccataa ttggacccgg

7201 actgcaagca aaagctacta gagaagctca aaaaaggaca gcagccggaa taatgaaaaa
7261 cccaactgtc gacgggatcg ttgcaataga tttggaccct gtggtttacg atgcaaaatt
7321 tgaaaaacag ctaggccaaa taatgttgtt gatactttgc acatcacaga tcctcctgat
7381 gcggaccaca tgggccttgt gtgaatccat cacactagcc actggacctc tgactacgct
7441 ttgggaggga tctccaggaa aattctggaa caccacgata gcggtgtcca tggcaaacat
7501 ttttagggga agttatctag caggagcagg tctggccttt tcattaatga aatctctagg
7561 aggaggtagg agaggcacgg gagcccaagg ggaaacactg ggagaaaaat ggaaaagaca
7621 gctaaaccaa ttgagcaagt cagaattcaa cacttacaaa aggagtggga ttatagaggt
7681 ggatagatct gaagccaaag aggggttaaa aagaggagaa acgactaaac acgcagtgtc
7741 gagaggaacg gccaaactga ggtggtttgt ggagaggaac cttgtgaaac cagaagggaa
7801 agtcatagac ctcggttgtg gaagaggtgg ctggtcatat tattgcgctg ggctgaagaa
7861 agtcacagaa gtgaaaggat acacgaaagg aggacctgga catgaggaac caatcccaat
7921 ggcaacctat ggatggaacc tagtaaagct atactccggg aaagatgtat tctttacacc
7981 acctgagaaa tgtgacaccc tcttgtgtga tattggtgag tcctctccga acccaactat
8041 agaagaagga agaacgttac gtgttctaaa gatggtggaa ccatggctca gaggaaacca
8101 attttgcata aaaattctaa atccctatat gccgagtgtg gtagaaactt tggagcaaat
8161 gcaaagaaaa catggaggaa tgctagtgcg aaatccactc tcaagaaact ccactcatga
8221 aatgtactgg gtttcatgtg gaacaggaaa cattgtgtca gcagtaaaca tgacatctag
8281 aatgctgcta aatcgattca caatggctca caggaagcca acatatgaaa gagacgtgga
8341 cttaggcgct ggaacaagac atgtggcagt agaaccagag gtggccaacc tagatatcat
8401 tggccagagg atagagaata taaaaaatga acacaaatca acatggcatt atgatgagga
8461 caatccatac aaaacatggg cctatcatgg atcatatgag gtcaagccat caggatcagc
8521 ctcatccatg gtcaatggtg tggtgagact gctaaccaaa ccatgggatg tcattcccat
8581 ggtcacacaa atagccatga ctgacaccac acccttgga caacagaggg tgtttaaaga
8641 gaaagttgac acgcgtacac caaaagcgaa acgaggcaca gcacaaatta tggaggtgac
8701 agccaggtgg ttatggggtt ttctctctag aaacaaaaaa cccagaatct gcacaagaga
8761 ggagttcaca agaaaagtca ggtcaaacgc agctattgga gcagtgttcg ttgatgaaaa
8821 tcaatggaac tcagcaaaag aggcagtgga agatgaacgg ttctgggacc ttgtgcacag
8881 agagagggag cttcataaac aaggaaaatg tgccacgtgt gtctacaaca tgatgggaaa
8941 gagagagaaa aaattaggag agttcggaaa ggcaaaagga agtcgcgcaa tatggtacat
9001 gtggttggga gcgcgctttt tagagtttga agcccttggt ttcatgaatg aagatcactg
9061 gttcagcaga gagaattcac tcagtggagt ggaaggagaa ggactccaca aacttggata
9121 catactcaga gacatatcaa agattccagg gggaaatatg tatgcagatg acacagccgg
9181 atgggacaca agaataacag aggatgatct tcagaatgag gccaaaatca ctgacatcat
9241 ggaacctgaa catgccctat tggccacgtc aatctttaag ctaacctacc aaaacaaggt
9301 agtaagggtg cagagaccag cgaaaaatgg aaccgtgatg gatgtcatat ccagacgtga
9361 ccagagagga agtggacagg ttggaaccta tggcttaaac accttcacca acatggaggc
9421 ccaactaata agacaaatgg agtctgaggg aatcttttca cccagcgaat tggaaacccc
9481 aaatctagcc gaaagagtcc tcgactggtt gaaaaaacat ggcaccgaga ggctgaaaag
9541 aatggcaatc agtggagatg actgtgtggt gaaaccaatc gatgacagat ttgcaacagc
9601 cttaacagct ttgaatgaca tgggaaaggt aagaaaagac ataccgcaat gggaaccttc
9661 aaaaggatgg aatgattggc aacaagtgcc tttctgttca caccatttcc accagctgat
9721 tatgaaggat gggagggaga tagtggtgcc atgccgcaac caagatgaac ttgtaggtag
9781 ggccagagta tcacaaggcg ccggatggag cttgagagaa actgcatgcc taggcaagtc

```
 9841 atatgcacaa atgtggcagc tgatgtactt ccacaggaga gacttgagat tagcggctaa
 9901 tgctatctgt tcagccgttc cagttgattg ggtcccaacc agccgcacca cctggtcgat
 9961 ccatgcccac catcaatgga tgacaacaga agacatgttg tcagtgtgga atagggtttg
10021 gatagaggaa aacccatgga tggaggacaa gactcatgtg tccagttggg aagacgttcc
10081 atacctagga aaaagggaag atcaatggtg tggttcccta ataggcttaa cagcacgagc
10141 cacctgggcc accaacatac aagtggccat aaaccaagtg agaaggctca ttgggaatga
10201 gaattatcta gacttcatga catcaatgaa gagattcaaa aacgagagtg atcccgaagg
10261 ggcactctgg taagccaact cattcacaaa ataaaggaaa ataaaaaatc aaacaaggca
10321 agaagtcagg ccggattaag ccatagcacg gtaagagcta tgctgcctgt gagccccgtc
10381 caaggacgta aaatgaagtc aggccgaaag ccacggttcg agcaagccgt gctgcctgta
10441 gctccatcgt ggggatgtaa aaacccggga ggctgcaaac catggaagct gtacgcatgg
10501 ggtagcagac tagtggttag aggagacccc tcccaagaca caacgcagca gcggggccca
10561 acaccagggg aagctgtacc ctggtggtaa ggactagagg ttagaggaga ccccccgcac
10621 aacaacaaac agcatattga cgctgggaga gaccagagat cctgctgtct ctacagcatc
10681 attccaggca cagaacgcca aaaaatggaa tggtgctgtt gaatcaacag gttct
```