

```
{\rtf0\ansi{\fonttbl\fl\modern Courier;\f0\swiss Helvetica;}
\paperw9840
\paperh8400
\margl120
\margr120
\pard\tx200\tx1160\tx2120\tx3080\tx4040\tx5000\tx5960\tx6920\tx7880\tx8840\t
x980
0\tx10760\tx11720\tx12680\tx13640\tx14600\tx15560\tx16520\tx17480\tx18440\
fl\b0\
i0\ulnone\fs20\li200\fc0\cf0 \
LOCUS   ALSFV42S 11442 bp RNA      VRL    14-OCT-1993\
DEFINITION  Semliki forest virus 42S RNA genome.\
ACCESSION  X04129 J02361 J02362 L00018 V01399 V01400 V01401\
KEYWORDS   capsid protein; genome; nonstructural protein; polyprotein; \
           structural protein.\
SOURCE     Semliki Forest virus.\
ORGANISM   Semliki Forest virus\
           Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;\
           Alphavirus.\
REFERENCE  1 (bases 1 to 7520)\
AUTHORS    Takkinen,K.\
TITLE      Complete nucleotide sequence of the nonstructural protein genes of\
           Semliki Forest virus\
JOURNAL    Nucleic Acids Res. 14 (14), 5667-5682 (1986)\
MEDLINE    86286581\
REFERENCE  2 (bases 7121 to 7480)\
AUTHORS    Riedel,H., Lehrach,H. and Garoff,H.\
TITLE      Nucleotide sequence at the junction between the nonstructural and\
           the structural genes of the semliki forest virus genome\
JOURNAL    J. Virol. 42 (2), 725-729 (1982)\
MEDLINE    82217052\
REFERENCE  3 (bases 7387 to 8335)\
AUTHORS    Garoff,H., Frischauf,A.M., Simons,K., Lehrach,H. and Delius,H.\
TITLE      The capsid protein of Semliki Forest virus has clusters of basic\
           amino acids and prolines in its amino-terminal region\
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 77 (11), 6376-6380 (1980)\
MEDLINE    81101055\
REFERENCE  4 (bases 8212 to 11442)\
AUTHORS    Garoff,H., Frischauf,A.M., Simons,K., Lehrach,H. and Delius,H.\
TITLE      Nucleotide sequence of cdna coding for Semliki Forest virus\
           membrane glycoproteins\
JOURNAL    Nature 288 (5788), 236-241 (1980)\
MEDLINE    81052444\
COMMENT\
           Data kindly reviewed (23-MAR-1987) by K. Takkinen.\
           \
           NCBI gi: 58614\
FEATURES   Location/Qualifiers\
```

source 1..11442\
/organism="Semliki Forest virus"\
CDS 86..1696\
/note="NCBI gi: 58616"\
/codon_start=1\
/product="nonstructural protein nsP1"\

/translation="MAAKVHVDIEADSPFIKSLQKAFPSFEVESLQVTPNDHANAR
AF\

SHLATKLTIEQETDKDTLILDIGSAPSRMMSTHKYHCVCPMRS AEDPERL
DSYAKKLA\

AASGKVLVDREIAGKITDLQTVMATPDAESPTFCLHTDVT CRTAAEVAVY
QDVYAVHAP\

TSLYHQAMKGVRTAYWIGFDTPPFMFDALAGAYPTYATNWADEQVLQ
ARNIGLCAASL\

TEGRLGKLSILRKKQLKPCDTVMFVSGSTLYTESRKL LRSWHLPSVFHLK
GKQSFTCR\

CDTIVSCEGYVVKKITMCPGLYGKTVGYAVTYHAEGFLVCKTTDTVKGE
RVSFVCTY\

VPSTICDQMTGILATDVTPEDAQLLVGLNQRI VVNGRTQRNTNTMKNY
LLPIVAVAF\

SKWAREYKADLDDEKPLGVRERSLTCCCLWAFKTRKMHTMYKKPDTQ
TIVKVPSEFNS\

FVIPSLWSTGLAIPVRSRIKMLLAKKTKRELIPVLDASSARDAEQEEKERL
EAELTRE\

ALPPLVPIAPAETGVVDVDVEELEYHAGA"\
CDS 86..7381\
/note="(aa 1-2431); NCBI gi: 58615"\
/codon_start=1\
/product="polyprotein of non-structural proteins"\

/translation="MAAKVHVDIEADSPFIKSLQKAFPSFEVESLQVTPNDHANAR
AF\

SHLATKLTIEQETDKDTLILDIGSAPSRMMSTHKYHCVCPMRS AEDPERL
DSYAKKLA\

AASGKVLVDREIAGKITDLQTVMATPDAESPTFCLHTDVT CRTAAEVAVY
QDVYAVHAP\

TSLYHQAMKGVRTAYWIGFDTPFMFDALAGAYPTYATNWADEQVLQ
ARNIGLCAASL\

TEGRLGKLSILRKKQLKPCDTVMFSGSTLYTESRKLRSWHLPSVFHLK
GKQSFTCR\

CDTIVSCEGYVVKKITMCPGLYGKTVGYAVTYHAEGFLVCKTTDTVKGE
RVSFPVCTY\

VPSTICDQMTGILATDVTPEDAQKLLVGLNQRIVVNGRTQRNTNTMKNY
LLPIVAVAF\

SKWAREYKADLDDEKPLGVRERSLTCCLWAFKTRKMHTMYKKPDTQ
TIVKVPSEFNS\

FVIPSLWSTGLAIPVRSRIKMLLAKKTKRELIPVLDASSARDAEQEERL
EAELTRE\

ALPPLVPIAPAETGVVDVDVEELEYHAGAGVVETPRSALKVTAQPNDVL
LGNYVVLSP\

QTVLKSSKLAPVHPLAEQVKIITHNGRAGGYQVDGYDGRVLLPCGSAIPV
PEFQALSE\

SATMVYNEREFVNRKLYHIAVHGPSLNTDEENYEKVRAERTDAEYVFDV
DKKCCVKRE\

EASGLVLVGELTNPPFHEFAYEGLKIRPSAPYKTTVVGVFGVPGSGKSII
KSLVTKH\

DLVTSGKKENCQEIVNDVKKHRGKGTSRENSDSILLNGCRRAVDILYVD
EAFACHSGT\

LLALIALVKPRSKVVLCDPKQCGFFNMMQLKVNFNHNICTEVCHKISISR
RCTRPVTA\

IVSTLHYGGKMRTTNPCKPIIIDTTGQTKPKPGDIVLTCFRGWAKQLQL
DYRGHEVM\

TAAASQGLTRKGVYAVRQKVNENPLYAPASEHVNVLTRTEDRLVWKT
LAGDPWIKVL\

SNIPQGNFTATLEEWQEEHDKIMKVIEGPAAPVDAFQNKANVCWAKSLV
PVLDTAGIR\

LTAEEWSTIITAFKEDRAYSPVVALNEICTKYYGVDLDSGLFSAPKVS LY
YENNHWDN\

RPGGRMYGFNAATAARLEARHTFLKGQWHTGKQAVIAERKIQPLSVLD
NVIPINRRLP\

HALVAEYKTVKGSRVEWLVNKNRGYHVLLVSEYNLALPRRRVTWLSPL
NVTGADRCYD\

LSLGLPADAGRFDLVFVNIHTEFRIHHYQQCVDHAMKLQMLGGDALRLL
KPGGILMRA\

YGYADKISEAVVSSLSRKFSARVLRPDCVTSNTEVFLLFSNFDNGKRPST
LHQMNTK\

LSAVYAGEAMHTAGCAPSYRVKRADIATCTEAAVVNAANARGTVGDG
VCRAVAKKWPS\

AFKGAATPVGTIKTVMCGSYPVIHAVAPNFSATTEAEGDRELAAYRAV
AAEVNRLSL\

SSVAIPLLSTGVFSGGRDRLQQSLNHLFTAMDATDADVTIYCRDKSWEK
KIQEAIDMR\

TAVELLNDDVELTTDLVRVHPDSSLVGRKGYSTTDGSLYSYFEGTKFNQ
AAIDMAEIL\

TLWPRLQEANERICLYALGETMDNIGSKCPVNDSSTPPRTVPCLCRYA
MTAERAR\

LRSHQVKSMVVCSSFPLPKYHVDGVQKVKCEKVLLFDPTVPSVVSPRKY
AASDDHSD\

RSLRGFDLDWTTDSSSTASDTMSLPSLQSCDIDSIYEPMAPIVVTADVHPE
PAGIADL\

AADVHPEPADHVDLENPIPPRPKRAAYLASRAAERPVPAPRKPTPAPRT
AFRNKLPL\

TFGDFDEHEVDALASGITFGDFDDVLRRLGRAGAYIFSSDTGSGHLQQKSV
RQHNLQCA\

QLDAVQEEKMYPPKLDTEREKLLLLKMQMHPSEANKSRYQSRKVENMK
ATVVDRLTSG\

ARLYTGADVGRIPTYAVRYPRPVYSPTVIERFSSPDVAIAACNEYLSRNYP
TVASYQI\

TDEYDAYLDMVDGSDSCLDRATFCPAKLRCYPKHHAYHQPTVRSVAVPSP
FQNTLQNVL\

AAATKRNCNVTQMRELPTMDSA VFNVECFKRYACSGEYWEEYAKQPIRI
TTENITTYV\

TKLKGPKAAALFAKTHNLVPLQEVPMDRFTVDMKRDKVTPGKHTTE
RPKVQVIQAA\

EPLATAYLCGIHRELVRRNLNAVLRPNVHTLFDMSAEDFDIASHFHPGD
PVLETDIA\

SFDKSQDDSLALTGLMILEDLGVDQYLLDLIEAAFGEISSCHLPTGTRFKF
GAMMKSG\

MFLTFLINTVLNITIASRVLEQRLTDSACAAFIGDDNIVHGVISDKLMAER
CASWVNM\

EVKIIDAVMGEKPPYFCGGFIVFDSVTQTACRVSDPLKRLFGLGKPLTAE
DKQDEDRR\

RALSDEVSKWFRTGLGAELEVALTSRYEVEGCKSILIAMTTTLARDIKAFK
KLRGPVIH\

LYGGPRLVR"\

CDS 1697..4090\
/product="nonstructural protein nsP2"\

CDS 4091..5536\
/product="nonstructural protein nsP3"\

misc_feature 5225..5248\
/note="imperfect repeat 1"\

misc_feature 5267..5290\
/note="imperfect repeat 1"\

misc_feature 5378..5389\
/note="imperfect repeat 2"\

misc_feature 5399..5410\
/note="imperfect repeat 2"\

misc_feature 5438..5455\
/note="imperfect repeat 3"\

misc_feature 5489..5506\
/note="imperfect repeat 3"\

CDS 5537..7378\
/product="nonstructural protein nsP4"\

misc_feature 7369..11442\
/note="26S RNA"\

CDS 7420..>11442\
/product="capsid protein"\

BASE COUNT 3051 a 3002 c 3088 g 2301 t\
ORIGIN \

1 atggcggatg tgtgacatac acgacgcca aagattttgt tccagctcct gccacctcgg\
61 ctacgcgaga gattaaccac ccacgatggc cgccaaagtg catgttgata ttgaggctga\
121 cagccattc atcaagtctt tgcagaaggc atttccgtcg ttcgaggtgg agtcattgca\
181 gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt\
241 gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt\
301 gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt\
361 gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt\
421 gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt\
481 gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt\
541 gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt\
601 gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt\
661 gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt\
721 gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt\
781 gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt\
841 gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt\
901 gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt\
961 gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt\
1021 gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt\
1081 gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt\
1141 gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt\
1201 gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt gctgctgctt

181 ggacacacca aatgaccatg caaatgccag agcattttcg cacctggcta ccaaattgat\
241 cgagcaggag actgacaaa acacactcat cttggatc ggcagtgcgc cttccaggag\
301 aatgatgtct acgcacaaat accactgcgt atgccctatg cgcagcgcag aagaccccga\
361 aaggctcgat agctacgcaa agaaactggc agcggcctcc gggaaaggtgc tggatagaga\
421 gatcgcagga aaaatcaccg acctgcagac cgtcatggct acgccagacg ctgaatctcc\
481 taccttttgc ctgcatacag acgtcacgtg tcgtacggca gccgaagtgg ccgtatacca\
541 ggacgtgtat gctgtacatg caccaacatc gctgtacat caggcgatga aaggtgtcag\
601 aacggcgtat tggattgggt ttgacaccac cccgtttatg tttgacgcgc tagcaggcgc\
661 gtatccaacc tacgccaaa actgggcccga cgagcaggtg ttacaggcca ggaacatagg\
721 actgtgtgca gcatccttga ctgagggag actcggcaaa ctgtccattc tccgaagaa\
781 gcaattgaaa cttgacgaca cagtcatgtt ctcggtagga tctacattgt aactgagag\
841 cagaagcta ctgaggagct ggcacttacc ctccgtattc cacctgaaag gtaacaatc\
901 ctttacctgt aggtgcgata ccatcgtatc atgtgaaggg tacgtagtta agaaatcac\
961 tatgtcccc ggctgtacg gtaaacgggt agggtagcc gtgacgtatc acgcggaggg\
1021 attectagtg tgcaagacca cagacactgt caaaggagaa agagtctcat tcctgtatg\
1081 cacctacgtc cctcaacca tctgtgatca aatgactggc atactagcga ccgacgtcac\
1141 accggaggac gcacagaagt tgttagtggg attgaatcag aggatagttg tgaacggag\
1201 aacacagcga aactaaca cgatgaagaa ctatctgctt ccgattgtgg ccgtcgcatt\
1261 tagcaagtgg gcgaggggaat acaaggcaga cttgatgat gaaaaacctc tgggtgtccg\
1321 agagaggtca ctacttctg gctgcttggg ggcattttaa acgaggaaga tgcacaccat\
1381 gtacaagaaa ccagacacc agacaatagt gaaggtgect tcagagtta actcgtctg\
1441 catcccagc ctatggctc caggcctgc aatcccagtc agatcacgca ttaagatgt\
1501 tttggccaag aagaccaagc gagagttaat acctgttctc gacgcgtcgt cagccaggga\
1561 tgctgaacaa gaggagaagg agaggttggg gcccagctg actagagaag ccttaccacc\
1621 cctcgtcccc atcgcgccgg cggagacggg agtcgtcgcac gtcgacgtg aagaactaga\
1681 gtatcacgca ggtgcagggg tcgtggaac acctcgcagc gcgttgaag tcaccgcaca\
1741 gccgaacgac gtactactag gaaattactg agttctgtcc ccgacagacc tgctcaagag\
1801 ctccaagttg gccccgtgc accctctagc agagcaggtg aaaataata cacataacgg\
1861 gagggccggc ggttaccagg tcgacggata tgacggcagg gtctactac catgtggatc\
1921 ggccattccg gtcctctgagt ttcaagctt gagcagagc gccactatgg tgtacaacga\
1981 aaggagttc gtcaacagga aactatacca tattgccgtt cacggaccgt cgctgaacac\
2041 cgacgaggag aactacgaga aagtcagagc tgaagaact gacgccagc acgtgttcga\
2101 cgtagataaa aatgctgctc tcaagagaga ggaagcgtc ggtttggtg tgggtggaga\
2161 gtaaccaac cccccgttc atgaattcgc ctacgaaggg ctgaagatca ggccgtcggc\
2221 accatataag actacagtag taggagtctt tggggttccg ggatcaggca agtctctat\
2281 tattaagagc ctctgacca aacacgatc gtcaccagc ggcaagaagg agaactgcca\
2341 ggaaatagtt aacgacgtga agaagcaccg cgggaagggg acaagtaggg aaaacagtga\
2401 ctccatctg ctaaacgggt gtcgtcgtc cgtggacac ctatatgtgg acgaggcttt\
2461 cgcttgccat tccgtactc tgctggcctt aattgctctt gttaacctc ggagcaaagt\
2521 ggtgttatgc ggagacccca agcaatcggg attcttcaat atgatgcagc ttaaggtgaa\
2581 ctcaaccac aacatctgca ctgaagtatg tcataaaagt atatccagac gttgcacgcg\
2641 tcagtcacg gccatcgtgt ctacgttga ctacggaggc aagatgcgca cgaccaacc\
2701 gtgcaacaaa cccataatca tagacaccac aggacagacc aagcccaagc caggagacat\
2761 cgtgttaaca tgcttccgag gctgggcaaa gcagctgcag ttgactacc gtggacacga\
2821 agtcatgaca gcagcagcat ctacggcct caccgcaaa ggggtatag ccgtaaggca\
2881 gaaggtgaat gaaaacctt tttatgccc tcgctcggag cacgtgaatg tactgctgac\
2941 gcgactgag gataggctgg tgtggaaaac gctggccggc gatccctgga ttaaggtcct\
3001 atcaaacatt ccacagggtg actttacggc cacattggaa gaatggcaag aagaacacga

3061 caaaataatg aaggtgattg aaggaccggc tgcgcctgtg gacgcgttcc agaacaagc\
3121 gaacgtgtgt tgggcgaaaa gcttggtgcc tgcctggac actgccgga tcagattgac\
3181 agcagaggag tggagcacca taattacagc atttaaggag gacagagctt actctccagt\
3241 ggtggccttg aatgaaattt gcaccaagta ctatggagtt gacctggaca gtggcctgtt\
3301 ttctgccccg aaggtgtccc tgtattacga gaacaaccac tgggataaca gacctggtgg\
3361 aaggatgtat ggattcaatg ccgaacagc tgccaggctg gaagctagac ataccttct\
3421 gaaggggcag tggcatacgg gcaagcaggc agttatcgca gaaagaaaaa tccaaccgct\
3481 ttctgtctg gacaatgtaa ttctatcaa ccgaggctg ccgacgccc tgggtgctga\
3541 gtacaagacg gtaaaaggca gtaggggtga gtggctggc aataaagtaa gagggtagca\
3601 cgtcctgtg gtgagtgagt acaacctggc tttgctcga cgcagggtca cttggtgtc\
3661 accgctgaat gtcacaggcg ccgataggct ctacgaccta agttaggac tgccggctga\
3721 cgccggcagg ttcgactgg tcttttgaa cattcacacg gaattcagaa tccaccacta\
3781 ccagcagtgt gtcgaccagc ccatgaagct gcagatgctt gggggagatg cgctacgact\
3841 gctaaaaccg ggcggcatct tgatgagagc ttacggatac gccgataaaa tcagcgaagc\
3901 cgtgtttcc tcttaagca gaaagttctc gtctgcaaga gtgtgccc cggattgtgt\
3961 caccagcaat acagaagtgt tctgtctt ctccaactt gacaacggaa agagacctc\
4021 tacgtacac cagatgaata ccaagctgag tgccgtgtat gccggagaag ccatgcacac\
4081 ggccgggtgt gcaccatct acagagttaa gagagcagac atagccactg gcacagaagc\
4141 ggctgtggtt aacgcagcta acgcccgtgg aactgtaggg gatggcgtat gcagggcctg\
4201 gggaagaaa tggccgtcag ctttaaggg agcagcaaca ccagtgggca caattaaac\
4261 agtcatgtgc ggctctacc ccgtcatcca cgtgtagcg cctaattct ctgccacgac\
4321 tgaagcggaa ggggaccgag aattggccc tctctaccgg gcagtggccg ccgaagtaaa\
4381 cagactgtca ctgagcagcg tagccatccc gctgctgtcc acaggagtgt tcagcggcgg\
4441 aagagatagg ctgcagcaat cctcaacca tctattaca gcaatggacg ccacggacgc\
4501 tgactgacc atctactgca gagacaaaag ttgggagaag aaaatccagg aagccattga\
4561 catgaggacg gctgtggagt tgctcaatga tgactggag ctgaccacag acttggtagt\
4621 agtgcaccg gacagcagcc tgggtggctg taagggtac agtaccactg acgggtcgtt\
4681 gtactctac tttgaaggta cgaaattcaa ccaggctgct attgatattg cagagatact\
4741 gacgttggg cccagactgc aggaggcaaa cgaacggata tgcctatac cgctgggcga\
4801 aacaatggac aacatcggat ccaaatgtcc ggtgaacgat tccgattcat caacacctc\
4861 caggacagt cctgcctgt gccgctacgc aatgacagca gaacggatcg cccgcctag\
4921 gtcacacaa gtaaaagca tgggtggttg ctatctttt cccctcccga aataccatgt\
4981 agatggggtg cagaaggtaa agtgcgagaa gtttctctg ttcgaccga cggtagctt\
5041 agtggttagt ccgcggaagt atgccgcatc tacgacggac cactcagatc ggtcgttagt\
5101 agggttgac ttgactgga ccaccgactc gtcttccact gccagcgata ccatgtcgtt\
5161 acccagttg cagtcgtgtg acatcgactc gatctacgag ccaatggctc ccatagtagt\
5221 gacggctgac gtacacctg aaccgcagg catcgcggac ctggcggcag atgtgacccc\
5281 tgaaccgca gaccatgtg acctcgagaa cccgattcct ccaccgccc cgaagagagc\
5341 tgcatactt gcctcccgc cggcggagcg accggtgccg gcgcccagaa agccgacgcc\
5401 tgcccgaagg actgcgttta ggaacaagct gcctttgacg ttcggcact ttgacgagca\
5461 cgaggtcgtat gcgttgccct ccgggattac ttcggagac ttcgacgacg tctgctgact\
5521 aggcgcgcg ggtgcatata ttttctctc ggacactggc agcggacatt tacaacaaaa\
5581 atccgttagg cagcacaatc tccagtgcg acaactggat gcggtccagg aggagaaaat\
5641 gtaccgcca aaattgata ctgagaggga gaagctgtt ctgctgaaaa tgcagatgca\
5701 cccatcggag gctaataaga gtcgatacca gtctcgaaa gtggagaaca tgaagccac\
5761 ggtggtggac aggtcacat cgggggccc agttgacacg ggagcggacg taggcccgat\
5821 accaacatac gcggttcgtt acccccgcc cgtgtactcc cctaccgtga tcgaaagatt\
5881 ctaagcccc gatgtagcaa tcgacgctg caacgaatac ctatccagaa attaccaac\

5941 agtggcgtcg taccagataa cagatgaata cgacgcatac ttggacatgg ttgacgggtc\
6001 ggatagttgc ttggacagag cgacattctg cccggcgaag ctccgggtgct acccgaaaca\
6061 tcatgctac caccagccga ctgtacgcag tgccgtcccg tcacctttc agaacacact\
6121 acagaacgtg ctageggccg ccaccaagag aaactgcaac gtcacgcaaa tgcgagaact\
6181 acccaccatg gactcggcag tgttaacgt ggagtgttc aagcgctatg cctgctccgg\
6241 agaattattg gaagaatatg ctaaacaacc tatccggata accactgaga acatcactac\
6301 ctatgtgacc aaattgaaag gcccgaagc tgctgcctg ttcgtaaga ccacaactt\
6361 ggttccgctg caggagggtc ccatggacag attcacggtc gacatgaaac gagatgtcaa\
6421 agtactcca gggacgaaac acacagagga aagacccaaa gtccaggtaa tcaagcagc\
6481 ggagccattg gcgaccgctt acctgtgccc catccacagg gaattagtaa ggagactaaa\
6541 tgctgtgta cgcctaacg tgcacacatt gttgatatg tcggccgaag actttgacgc\
6601 gatcatgcc tctacttcc acccaggaga cccggttcta gagacggaca ttgatcatt\
6661 cgacaaaagc caggacgact ccttgctct tacaggttta atgatctcg aagatctagg\
6721 ggtgatcag tacctgctgg acttgatcga ggcagcctt ggggaaatat ccagctgtca\
6781 cctaccaact ggcacgcgt tcaagtccg agctatgatg aaatcgggca tgttctgac\
6841 tttgttatt aacctgtt tgaacatcac catagcaagc agggacttg agcagagact\
6901 cactgactcc gctgtgccc ccttcatcgg cgacgacaac atcgttcag gagtgatctc\
6961 cgacaagctg atggcggaga ggtgcgcgtc gtgggtcaac atggaggta agatcattga\
7021 cgctgtcatg ggcgaaaaac cccatattt ttgtggggga tcatagttt ttgacagcgt\
7081 cacacagacc gctgcccgtg ttccagacc acttaagcgc ctgtcaagt tgggtaagcc\
7141 gtaacagct gaagacaagc aggacgaaga caggcgacga gactgagtg acgaggttag\
7201 caagtgggtc cggacaggct tgggggcccga actggagggtg gactaacat ctaggtatga\
7261 ggtagagggc tgcaaaagta tctcatagc catgaccacc ttggcgaggg acattaaggc\
7321 gttaagaaa ttgagaggac ctgtataca cctctacggc ggtctagat tgggtgctta\
7381 atacacagaa ttctgattat agcgcactat tatagcacca tgaattacat ccttacgcaa\
7441 acgttttacg gccgccggtg gcgcccgcgc ccggcggccc gtccttggcc gttgcaggcc\
7501 actccggtg ctcccgtct ccccacttc caggcccagc agatgcagca actcatcagc\
7561 gccgtaaatg cgctgacaat gagacagaac gcaattgctc ctgctaggcc tcccaaacca\
7621 aagaagaaga agacaacca accaaagccg aaaacgcagc ccaagaagat caacggaaaa\
7681 acgcagcagc aaaagaaga agacaagcaa gccgacaaga agaagaaga acccgaaaa\
7741 agagaagaa tgtgcatgaa gattgaaat gactgtatct tcgaagtcaa acacgaagga\
7801 aaggtcactg ggtacgcctg cctggtgggc gacaaagtca tgaacctgc ccactgaaa\
7861 ggagtcactg acaacgcgga cctggcaaag ctagcttca agaatcgag caagtatgac\
7921 cttgagtgtg ccagatacc agttcatatg aggtcggatg cctcaaagta cacgcatgag\
7981 aagcccaggg gacactataa ctggcaccac ggggctgttc agtacagcgg aggtaggttc\
8041 actatacca caggagcggg caaacgggga gacagtggcc ggccatctt tgacaacaag\
8101 gggagggtag tcgtatcgt cctgggcccg gccaacgagg gtcacgcac agcactgtcg\
8161 gtggtcact ggaacaaga tatggtgact agagtgacc cggaggggtc cgaagagtgg\
8221 tccgcccgc tgattactgc catgtgtgct ctgccaatg ctacctccc gtgctccag\
8281 ccccgtgtg tacctgtg ctatgaaaac aacgcagagg ccactactg gatgctcgag\
8341 gataacgtg ataggccagg gtactacgac ctcttcagg cagcctgac gtgccgaaac\
8401 ggaacaagc accggcgcag cgtgtcga cactcaacg tgtataaggc tacagccct\
8461 tacatcgct actgcgccga ctgcccagca gggcactcgt gcatagccc cgtagcaatt\
8521 gaagcggta ggtccgaagc taccgacggg atgctgaaga ttcagttct gcacaaatt\
8581 ggcatagata agagtgaca tcatgactac acgaagataa ggtacgcaga cgggcacgcc\
8641 attgagaatg ccgtccggtc atcttgaag gtagccacct ccggagactg ttcgtccat\
8701 ggcacaatg gacatttcat actggcaaag tcccaccgg gtgaattct gcaggtctcg\
8761 atccaggaca ccagaacgc ggtccgtgc tcagaatac aatatcatca tgacctcaa\

8821 ccggtgggta gagaaaaatt tacaattaga ccacactatg gaaaagagat cccttgacc\
8881 acttatcaac agaccacagc ggagaccgtg gaggaaatcg acatgcatat gccgccagat\
8941 acgccggaca ggacgttgc atcacagcaa tctggcaatg taaagatcac agtcggagga\
9001 aagaaggtga aatacaactg cacctgtgga accggaaacg ttggcactac taattcggac\
9061 atgacgatca acacgtgtct aatagagcag tggcactgtc cagtgcagga ccataagaaa\
9121 tggcagttca actcaccttt cgtcccgaga gccgacgaac cggctagaaa aggcaaagtc\
9181 catatcccat tcccgttgga caacatcaca tgcagagttc caatggcgcg cgaaccaacc\
9241 gtatccacg gcaaaagaga agtgacactg caccttcacc cagatcatcc cagctcttt\
9301 tctaccgca cactgggtga ggaccgcag taccagagg aatgggtgac agcggcggtg\
9361 gaacggacca taccgtacc agtggacggg atggagtacc actggggaaa caacgacca\
9421 gtgaggcttt ggtctcaact caccactgaa gggaaaccgc acggctggcc gcatcagatc\
9481 gtacagtact actatgggtt taccggcc gctacagtat ccgcggtcgt cgggatgagc\
9541 ttactggcgt tgatctgat ctctcgtcgt tgctacatgc tggttgcggc ccgcagtaag\
9601 tgcttgacc cttatgcttt aacaccagga gctgcagttc cgtggacgct ggggatactc\
9661 tgctgcgcc cgcgggcgca cgcagctagt tggcagaga ctatggccta cttgtgggac\
9721 caaaaccaag cgtgttctg gttggagttt gggccctg ttgctgcat cctcatcctc\
9781 acgtattgcc tcagaaactg gctgtgtgc tgaagagcc ttcttttt agtgctactg\
9841 agcctcgggg caaccgccag agcttacgaa cattcgacag taatgccgaa cgtggtgggg\
9901 tccccata aggtcacat tgaaggcca ggatagacc cctcacttt gcagatgcag\
9961 gttgtgaaa ccagcctga accaaccctt aatttggat acataacctg tgagtacaag\
10021 acggtcgtcc cgtcggcgtg cgtgaagtgc tgcggcgct cagagtgtc cactaaagag\
10081 aagcctgact accaatgcaa ggttacaca ggcgtgtacc cgttcatgt gggaggggca\
10141 tattgctct gcgactcaga aaacacgcaa ctacagagg cgtacgtcga tcgatcggac\
10201 gtatgcaggc atgatcagc atctgcttac aaagccata cagatcgtc gaaggccaaa\
10261 gtgagggtta tttacggcaa cgtaaaccag actgtggatg tttacgtgaa cggagaccat\
10321 gccgtcaga taggggttac tcaattcata ttggggcgc tgcacggc ctggaccccg\
10381 ttgacaaca agatagctgt gtacaaagac gaagtgtca atcaggactt cccgccgtac\
10441 ggatctgggc aaccagggcg ctccggcgac atccaaagca gaacagtgga gagtaacgac\
10501 ctgtacgca acacggcact gaagctggca cgccttcac ccggcatggt ccatgtaccg\
10561 tacacacaga cacctcagg gtccaatat ttgctaaagg aaaaaggac agccctaat\
10621 acgaaggetc cttttgctg ccaaatcaaa acgaaccctg tcagggccat gaactgcgcc\
10681 gtgggaaaca tcctgtctc catgaattg cctgacagcg ctttaccg cattgtcag\
10741 gcgccacca tcattgacct gacttcaca tggctacct gtacgcactc ctggatttc\
10801 ggcggcgtct tgacactgac gtacaagacc aacaagaacg gggactgtc tttactcgt\
10861 cacttaacg tagtactct acaggaggcc acagcaaaag tgaagacagc aggtaaggtg\
10921 accttacct tctccagcg aagcgcata cctctttt gggtgtcgt atgcagtct\
10981 agggccacct gttcagcgtc gttgagccc ccgaaagacc acatagccc atatcggct\
11041 agccacagta acgtagtgt tccagacatg tgggcaccg cactatcat ggtgcagaaa\
11101 atctcgggtg gtctgggggc ctctgcaatc ggcctatcc tgggtcgtgt tttgtcact\
11161 tgattgggc tccgagata agttagggtg ggcaatggca ttgatatagc aagaaaattg\
11221 aaaacagaaa aagttaggtg aagcaatggc atataacct aactgtataa cttgtaacaa\
11281 agcgcaaaa gacctgcga attggcccc tggctcgcct cacggaaact cggggcaact\
11341 catattgaca cattaattg caataattg aagcttacc aagcttaatt cgacgaataa\
11401 ttgattttt attttttt gcaattggtt ttaatttt cc

}