

Rous sarcoma virus, complete genome

NCBI Reference Sequence: NC_001407.1

[FASTA Graphics](#)

[Go to:](#)

LOCUS NC_001407 9392 bp ss-RNA linear VRL 13-AUG-2018

DEFINITION Rous sarcoma virus, complete genome.

ACCESSION NC_001407

VERSION NC_001407.1

DBLINK BioProject: [PRJNA485481](#)

KEYWORDS RefSeq.

SOURCE Rous sarcoma virus

ORGANISM [Rous sarcoma virus](#)
Viruses; Riboviria; Pararnavirae; Artverviricota; Revtraviricetes;
Ortervirales; Retroviridae; Orthoretrovirinae; Alpharetrovirus.

REFERENCE 1 (bases 1 to 9392)

AUTHORS Petropoulos,C.J.

TITLE Appendix 2: Retroviral taxonomy, protein structure, sequences, and genetic maps

JOURNAL (in) Coffin,J.M. (Ed.);
RETROVIRUSES: 757;
Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, NY, USA (1997)

REFERENCE 2 (bases 1 to 9392)

CONSRMT NCBI Genome Project

TITLE Direct Submission

JOURNAL Submitted (03-SEP-2004) National Center for Biotechnology Information, NIH, Bethesda, MD 20894, USA

REFERENCE 3 (bases 1 to 9392)

AUTHORS Chappey,C.

TITLE Direct Submission

JOURNAL Submitted (12-NOV-1997) NIH, NLM, Rockville Pike, Bethesda, MD 20894, USA

COMMENT PROVISIONAL [REFSEQ](#): This record has not yet been subject to final NCBI review. The reference sequence was derived from [AF033808](#).
COMPLETENESS: full length.

FEATURES Location/Qualifiers

source 1..9392
/organism="Rous sarcoma virus"
/mol_type="genomic RNA"
/db_xref="taxon:[11886](#)"

[gene](#) 1..9392
/locus_tag="RSVgpl"
/db_xref="GeneID:[2193432](#)"

[mRNA](#) 1..9392

```

/ gene="gag-pol"
mRNA      join(1..397,5078..9392)
/ gene="env"
/ product="gPr95 polyprotein precursor"
5'UTR     22..102
CDS       join(380..397,5078..6865)
/ gene="env"
/ codon_start=1
/ product="gPr95 polyprotein precursor"
/ protein_id="NP_056885.1"
/ translation="MEAVIKAFLTGYPGKTSKKDSKEKPLATSKKDPEKTPLLPTRVN
YILIIGVLVLC EVTGV RADVHLL EQPNLWITWANRTGQTD FCLSTQSATSPFQTCLI
GIPSPISEGDFKGYVSDTNCSTVGTDRVLVLSASITGGPDNSTTLTYRKVSCLLLKLN
VSMWDEPPELQLLGSQSLPNVTNITQVSGVAGGC VYFAPRATGLFLGWSKQGLSRFLLR
HPFTSTSNSTEPFTVVTADRHNLFMGSEYCGAYGYRFEIYNCSQTRNTYRCGDVGGT
GLPETWCRGKGGIWVNQSKEINETEPFSFTANCTGSNLGNVSGCCGE PITILPLGAWI
DSTQGSFTKPKALPPAIFLICDRAWQGIPSRVGGPCYLGKLTMLAPNHTDILKILA
NSSRTGIRRKRSVSHLDDTCSDEVQLWGPTARIFASILAPGVAAAQALREIERLACWS
VKQANLTSSLGDLDDVTSIRHAVLQNRAAIDFLLAHGHGCE DVAGMCCFNLSDHS
ESIQKKFQLMKKHVNKIGVSDPIGSWLRGIFGGIGEWAVHLLKGLLLGLVVILLLV
CLPCLLQFVSSSIRKMINSSINYHTEYRKMGGAV"
mat_peptide 5249..6268
/ gene="env"
/ product="gp85 SU"
/ protein_id="NP_955614.1"
mat_peptide 6269..6862
/ gene="env"
/ product="gp37 TM"
/ protein_id="NP_955615.1"
CDS       join(380..2482,2482..5190)
/ gene="gag-pol"
/ ribosomal_slippage
/ codon_start=1
/ product="Pr180 polyprotein precursor"
/ protein_id="NP_056886.1"
/ translation="MEAVIKVISSACKTYCGKTSPSKKEIGAMLSLLQKEGLLMSPSD
LYSPGSWDPI TAALSQRAMILGKSGELKTWGLVLGALKAAAREEQVTSEQAKFWLGLGG
GRVSPPGPECIEKPATERRIDKGEEVGETTVQRDAKMAPEETATPKTVGTSCYHCGTA
IGCNCATASAPPPPYVGSGLYPSLAGVGEQQGQGGDTPPGAEQSRAEPGHAGQAPGPA
LTDWARVREELASTGPPVVAMPVVIKTEGPAWTPLEPKLITRLADTVRTKGLRSPITM
AEVEALMSSPLLPHDVTNLMRVILGPAPYALWMDAWGVQLQTVIAAATRDPRHPANGQ
GRGERTNLNRLKGLADGMVGNPQGQAALLRPGELVAITASALQAFREVARLAEPAGPW
ADIMQGPSESFVDFANRLIKAVEGSDLPPSARAPVIIDCFRQKSQPDIIQLIRTAPST
LTPGEIIKYVLDRQKTAPLTDQGIAAAMSSAIQPLIMAVVNRERDGGQTSGSGRARGL

```

CYTCGSPGHYQAQCPKKRKSNGSRERCQLCNGMGHNAKQCRKRDGNQGQRPKGKLSSG
PWPGEPPAVSLAMTMEHKDRPLVRVILTNTGSHPVKQRSVYITALD SGADITIISE
EDWPTDWPVMEANPQIHGIGGGIPMRKSRDMIELGVINRDGSLERPLLLFPVAMVR
GSILGRDCLQGLGLRLTNLIGRATVLTVALHLAIPKWKPDHTPVWIDQWFLPEGKLV
ALTQLVEKELQLGHIEPSLSCWNTPVFVIRKASGSYRLLHDLRAVNAKLVPF GAVQQG
APVLSALPRGWPLMVLDLKDCFFSIPLAEQDREAFATLPSVNNQAPARRFQWKVLPQ
GMTCSPTICQLVVGQVLEPLRLKHPSLCMLHYMDDL LLAASSHDGLEAAGEEVISTLE
RAGFTISPDKVQREPGVQYLGYKLGSTYVAPVGLVAEPRIATLWDVQKLVGSLQWLRP
ALGIPPRLMGPFYEQLRGSDPNEAREWN LDMKMAWREIVRLSTTAALERWDPALPLEG
AVARCEQGAIGVLGQGLSTHPRCLWLFSTQPTKAFTAWLEVLTLLITKLRA SAVRTF
GKEVDILLLPACFREDLPLPEGILLALKGFAGKIRSSDTPSIFDIARPLHVS LKVRVT
DHPVPGPTVFTDASSSTHKGVVVWREGPRWEIKEIADLGASVQQLEARAVAMALLWP
TTPTNVVTDSAFVAKMLLKMQEGVPSTAAAFILEDALSQRSAMAAVLHVRSHSEVPG
FFTEGNDVADSQATFQAYPLREAKDLHTALHIGPRALSKACNISMQQAREVVQTC PHC
NSAPALEAGVNPRGLGPLQIWQTDFTLEPRMAPRSWLAVTVDTASSAIVVTQHGRVTS
VAVQHHWATAIAVLGRP KAIKTDNGSCFTSKSTREWLARWGIAHTTGIPGNSQGGAMV
ERANRLKDRIRVLAEGDGFMKRIPTSKQGELLAKAMYALNHFERGENTKTPIQKHWR
PTVLTEGPPVKIRIETGEWEKGWNVLVWGRGYAAVKNRD TDKVIWVPSRKVKPDITQK
DEVTKKDEASPLFAGISDWIPWEDEQEG LQGETASNKQERPGEDTLAANES"

mat_peptide join(2102..2482,2482..2502)

/gene="gag-pol"

/product="p15 PR"

/protein_id="[NP_955613.1](#)"

mat_peptide 2503..4218

/gene="gag-pol"

/product="p68a RT"

/protein_id="[NP_955611.1](#)"

mat_peptide 4219..5187

/gene="gag-pol"

/product="p32 IN"

/protein_id="[NP_955612.1](#)"

gene 380..2485

/gene="gag"

/locus_tag="RSVgp2"

/db_xref="GeneID:[1491923](#)"

CDS 380..2485

/gene="gag"

/locus_tag="RSVgp2"

/codon_start=1

/product="Pr76 polyprotein precursor"

/protein_id="[NP_056887.1](#)"

/db_xref="GeneID:[1491923](#)"

/translation="MEAVIKVISSACKTYCGKTSPSKKEIGAMLSLLQKEGLLMSPSD
LYSPGSWDPITAALSQRAMILGKSGELKTWGLVLGALKAAAREEQVTSEQAKFWLGLGG

GRVSPPGPECIEKPATERRIDKGEEVGETTVQRDAKMAPEETATPKTVGTSCYHCGTA
IGCNCATASAPPPPYVGSGLYPSLAGVGEQQGQGGDTPPGAEQSRAPGHAGQAPGPA
LTDWARVREELASTGPPVVAMPVVIKTEGPAWTPLEPKLITRLADTVRTKGLRSPITM
AEVEALMSSPLLPHDVTNLMRVILGPAPYALWMDAWGVQLQTVIAATRDPHRHPANGQ
GRGERTNLNRLKGLADGMVGNPQGQAALLRPGELVAITASALQAFREVARLAEPAGPW
ADIMQGPSESFVDFANRLIKAVEGSDLPPSARAPVIIDCFRQKSQPDIIQLIRTAPST
LTPGEI IKYVLDRQKTAPLTDQGIAAAMSSAIQPLIMAVVNRERDGGQTGSGGRARGL
CYTCGSPGHYQAQCPKKRKSGNSRERCQLCNGMGHNAKQCRKRDGNQGRPGKGLSSG
PWPGEPPAVSLAMTMEHKDRPLVRVILTNTGSHPVKQRSVYITALDLSGADITIISE
EDWPTDWPVMEANPQIHGIGGGIPMRKSRDMIELGVINRDGSLERPLLLFPAVAMVR
GSILGRDCLQGLGLRLTNL"

mat_peptide 380..844
/gene="gag"
/locus_tag="RSVgp2"
/product="p19"
/protein_id="[NP_955607.1](#)"

mat_peptide 911..1096
/gene="gag"
/locus_tag="RSVgp2"
/product="pp10"
/protein_id="[NP_955608.1](#)"

mat_peptide 1097..1816
/gene="gag"
/locus_tag="RSVgp2"
/product="p27 CA"
/protein_id="[NP_955609.1](#)"

mat_peptide 1844..2110
/gene="gag"
/locus_tag="RSVgp2"
/product="p12 NC"
/protein_id="[NP_955610.1](#)"

gene 2482..5190
/gene="pol"
/locus_tag="RSVgp3"
/db_xref="GeneID:[1491924](#)"

gene 7129..8709
/gene="src"
/locus_tag="RSVgp4"
/db_xref="GeneID:[1491925](#)"

mRNA 7129..8709
/gene="src"
/locus_tag="RSVgp4"
/product="p60 src"
/db_xref="GeneID:[1491925](#)"

[CDS](#) 7129..8709
 /gene="src"
 /locus_tag="RSVgp4"
 /note="p60-SRC phosphoprotein"
 /codon_start=1
 /product="p60 src"
 /protein_id="[NP_056888.1](#)"
 /db_xref="GeneID:[1491925](#)"
 /translation="MGSSKSKPKDPSQRRHSLEPPDSTHHGGFPASQTPDETAAPDAH
 RNPSRSFGTVATEPKLFWGFNTSDTVTSPQRAGALAGGVTTTFVALYDYESWTETDLSF
 KKGERLQIVNNTGDDWLAHSLTTGQTGYIPSNYVAPSDSIQAEWYFGKITRRESER
 LLLNPENPRGTFLVRKSETAKGAYCLSVSDFDNAKGPNVKHYKIYKLYSGGFYITSRT
 QFGSLQQLVAYYSKHADGLCHRLANVCPTSKPQTQGLAKDAWEIPRESLRLEAKLGQG
 CFGEVVMGTWNDTTRVAIKTLKPGTMSPEAFLEAQVMKKLRHEKLVQLYAVVSEEP
 YIVIEYMSKGSLLDFLKGMGKYLRLPQLVDMAAQIASGMAYVERMNYVHRDLRAANI
 LVGENLVCKVADFGRLARLIEDNEYTARQGAKFPIKWTAPAAALYGRFTIKSDVWSFGI
 LLTELTTKGRVPYPGMVNREVLQDQVERGYRMPCCPECPESLHDLMCQCWRKDPEERPT
 FKYLQAQLLPACVLEVAE"

[mat_peptide](#) 7132..8706
 /gene="src"
 /locus_tag="RSVgp4"
 /product="pp60 SRC"
 /protein_id="[NP_955616.1](#)"

[3'UTR](#) 9058..9292
 /locus_tag="RSVgp1"

ORIGIN

```

1 gccattttac cattcaccac attggtgtgc acctgggttg atggccggac cgtcgattcc
61 ctaacgattg cgaacacctg aatgaagcag aaggcttcac ttggtgaccc cgacgtgata
121 gttagggaa agtggtcggc cacagacggc gtggcgatcc tgccctcatc cgtctcgctt
181 attcggggag cggacgatga ccctagtaga gggggctgcg gcttaggagg gcagaagctg
241 agtggcgctg gagggagctc tactgcaggg agcccagata ccctaccgag aactcagaga
301 gtcgttgtaa gacgggaagg aagcccgacg actgagcagt ccacccagg cgtgattctg
361 gtcgcccggg ggatcaagca tggaagccgt cataaagggtg atttcgtccg cgtgtaaaac
421 ctattgcggg aaaacctctc cttctaagaa ggaaataggg gccatgttgt ccctcttaca
481 aaaggaaggg ttgcttatgt ctccctcaga cttatatcc cgggggtcct gggatcccat
541 taccgcggcg ctatcccagc gggctatgat acttgggaaa tcgggagagt taaaaacctg
601 gggattggtt ttgggggcat tgaaggcggc tcgagaggaa caggttacat ctgagcaagc
661 aaagtthtgg ttgggattag ggggagggag ggtctctccc ccagggtccg agtgcacga
721 gaaaccagca acggagcggc gaatcgacaa aggggaggaa gtgggagaaa caactgtgca
781 gcgagatgag aagatggcgc cggaggaaac ggccacacct aaaaccgttg gcacatcctg
841 ctatcattgc ggaacagcta ttggctgtaa ttgcgccaca gcctcggtc ctcctcctcc
901 ttatgtgggg agtggthtgt atccttcctt ggcgggggtg ggagagcagc agggccaggg
961 gggtagacac cctccggggg cggaacagtc aagggcggag ccagggcgat cgggtcaggc
1021 tcctgggccg gccctgactg actgggcaag ggtcaggagg gagcttgca gtactgttcc

```

1081 gcccggtggtg gccatgcctg tagtgattaa gacagaggga cccgcttgga cccctctgga
1141 gccaaaattg atcacaagac tggctgatac ggtcaggacc aagggcttac gatccccgat
1201 tactatggca gaagtggaag cgcttatgtc ctccccgctg ctgccgcatg acgtcacgaa
1261 tctaattgaga gttatTTTTtag ggcttgcccc atatgcctta tggatggacg cttgggggagt
1321 ccaactccag acagttatag cggcagccac tcgcgacccc cgacaccag cgaacggtca
1381 agggcggggg gaacggacta atttgaatcg cttaaagggc ttagctgatg ggatggtggg
1441 caaccacag ggtcaggccg cattattaag accgggggaa ttggttgcta ttacggcgtc
1501 ggctctccag gcgttttagag aggttgcccc gctggcgga cctgcaggtc catgggcgga
1561 catcatgcag ggaccatctg agtcctttgt tgattttgcc aatcggtta taaaggcgtt
1621 tgaggggtca gatctcccg cttccgcgcg ggctccggtg atcattgact gcttttagga
1681 gaagtcacag ccagatatc agcagcttat acggacagca ccctccacgc tgaccacccc
1741 aggagagata attaaatatg tgctagacag gcagaagact gccctctta cggatcaagg
1801 catagccgcg gccatgtcgt ctgctatcca gcccttaatt atggcagtag tcaatagaga
1861 gagggatgga caaactgggt cgggtggtcg tgcccgaggg ctctgctaca cttgtggatc
1921 cccgggacat tatcaggcgc agtgcccga aaaacggaag tcaggaaaca gccgtgagcg
1981 atgtcagttg tgtaacggga tgggacacaa cgctaaacag tgtaggaagc gggatggcaa
2041 ccagggccaa cgcccaggaa aaggtctctc ttcggggccg tggcccgcc ctgagccacc
2101 tgccgtctcg ttagcgatga caatggaaca taaagatcg cccttggtta gggtcattct
2161 gactaacact gggagtcac cgggtcaaaca gcgttcggtg tatatcaccg cgctgttgga
2221 ctctggagcg gacatcacta ttatttcaga ggaggattg cccaccgatt ggccagtgat
2281 ggaggccgcg aaccgcaga tccatgggat aggaggggga attcccatgc gaaaatctcg
2341 tgacatgata gagttggggg ttattaaccg agacgggtct ttggagcgac ccctgctcct
2401 cttccccgca gtagctatgg ttagaggag tatcctagga agagattgtc tgcagggcct
2461 agggctccgc ttgacaaatt tatagggag gccactgttc tcaactgttc gctacatctg
2521 gctattccgc tcaaatggaa gccagaccac acgcctgtgt ggattgacca gtggcccctc
2581 cctgaaggta aacttgtagc gctaacgcaa ttagtgaaa aagaattaca gttaggacat
2641 atagaacctt cacttagttg ttggaacaca cctgtcttcg tgatccgga ggttccggg
2701 tcttaccgct tactgcatga tttgcgcgt gttaacgcca agcttgttcc ttttggggcc
2761 gtccaacagg gggcgccagt tctctccgcg ctcccgctg gctggcccct gatggtctta
2821 gacctcaagg attgcttctt ttctatccct cttgcggaac aagatcgca agcttttgca
2881 tttacgctcc cctctgtgaa taaccaggcc cccgctcgaa gattccaatg gaaggtcttg
2941 cccaaggga tgacctgttc tccactatc tgtcagttgg tagtgggtca ggtacttgag
3001 cccttgcgac tcaagcacc atctctgtgc atgttgcat atatggatga tcttttgcta
3061 gccgcctcaa gtcacgatg gttggaagcg gcaggggag aggttatcag tacattggaa
3121 agagccgggt tcaactattc gcctgataag gtccagagg agcccgagt acaatatctt
3181 gggtaacaagt taggcagtac gtatgtagca cccgtaggcc tggtagcaga acccaggata
3241 gccacctgtt gggatgttca aaagctgggt gggtcacttc agtggcttcg ccagcgta
3301 ggaatccgc cagactgat gggcccttc tatgagcagt tacgagggtc agatcctaac
3361 gaggcgagg aatggaatct agacatgaaa atggcctgga gagagatcgt acggcttagc
3421 accactgct ccttggaac atgggacct gccctgcctc tggaaggagc ggtcgctaga
3481 tgtgaacagg gggcaatagg ggttttggga cagggactgt ccacacaccc aaggccatgc
3541 ttgtggttat tctccacca accaccaag gcgtttactg cttgggttaga agtgctcacc
3601 cttttgatta ctaagctacg tgcttcggca gtgcgaacct ttggcaagg ggtcgatatc
3661 ctctgtttgc ctgcatgct tcgggaggac cttccgctcc cagaggggat cctgttagcc

3721 ctttaaggggt ttgcaggaaa aatcaggagt agtgacacgc catctatctt tgacattgctg
3781 cgtccactgc atgtttctct gaaagtgagg gttaccgacc accctgtgcc gggaccact
3841 gtctttactg acgcctcctc aagcacccat aaggggtg tagtctggag ggagggccca
3901 aggtgggaga taaaagaaat agctgatttg ggggcaagtg tacaacaact ggaagcacgc
3961 gctgtggcca tggcacttct gctgtggccg acaacgcca ctaatgtagt gactgactcc
4021 gcgtttgttg cgaaaatgtt actcaagatg ggacaggagg gagtccgctc tacagcggcg
4081 gcttttattt tagaggatgc gttaagccaa aggtcagcca tggccgccgt tctccacgtg
4141 cggagtcatc ctgaagtgcc agggtttttc acagaaggaa atgacgtggc agatagccaa
4201 gccaccttcc aagcgtatcc cttgagagag gctaaagatc ttcataccgc tctccatatt
4261 ggaccccgcg cgctatccaa agcgtgtaat atatctatgc agcaggctag ggaggttgtt
4321 cagacctgcc cgcattgtaa ttcagccctc gcgttgaggc ccggagtaaa ccctaggggt
4381 ttgggacccc tacagatatg gcagacagac tttacgcttg agcctagaat ggcccccgct
4441 tcctggctcg ctgttactgt ggataccgcc tcatcagcga tagtcgtaac tcagcatggc
4501 cgtgtcacat cggttgctgt acaacatcat tgggccacgc ctatcgccgt tttgggaaga
4561 ccaaaggcca taaaaacaga taatgggtcc tgcttcacgt ctaaaccac gcgagagtgg
4621 ctgcgagat ggggatagc acacaccacc gggattccgc gtaattccca gggatcaagc
4681 atggtagagc gggccaaccg gctcctgaaa gataggatcc gtgtgcttgc ggagggggac
4741 ggctttatga aaagaatccc caccagcaaa cagggggaac tattagccaa ggcaatgtat
4801 gccctcaatc actttgagcg tggtgaaaac acgaaaacac cgatacaaaa aactggaga
4861 cctaccgttc ttacagaagg acccccggtt aaaatacgaa tagagacagg ggagtgggaa
4921 aaaggatgga acgtgctggt ctggggacga ggttatgccg ctgtgaaaaa cagggacact
4981 gataaggtta tttgggtacc ctctcgaaaa gttaaaccgc acatcaccca aaaggatgag
5041 gtgactaaga aagatgaggc gagccctctt tttgcaggca tttctgactg gataccctgg
5101 gaagacgagc aagaaggact ccaaggagaa accgctagca acaagcaaga aagaccgga
5161 gaagacaccc ttgctgcca cgagagttaa ttatattctc attattggtg tcctggtctt
5221 gtgtgaggtt acgggggtaa gagctgatgt tcaattactc gagcagccag ggaacctttg
5281 gattacatgg gccaacgta caggccaaac ggatttctgc ctctctacac agtcagccac
5341 ctcccctttt caaacatgtt tgataggtat cccgtctcct atttccgaag gtgattttaa
5401 gggatatgtt tctgatacaa attgctccac tgtgggaact gaccggttag tcttgtcagc
5461 cagcattacc ggcgccctg acaacagcac caccctact tatcgaaag tttcatgcct
5521 gctgttaaag ctgaacgtct ccatgtggga tgagccacct gaactgcagc tgctaggttc
5581 ccagtctctc cctaacgtta ctaacattac tcaggtctct ggcgtggccg ggggatgtgt
5641 atatttcgcc ccaagggcc ctggcctgtt tttaggttg tctaaacaag gtctctcgcg
5701 gttcctcctc cgtcacccct ttacctccac ctctaactcc acggaaccgt tcacggtggt
5761 gacagcggat agacacaatc tttttatggg gagtgagtag tgtggtgcat atggctacag
5821 attttgggaa atatataact gctcacagac taggaatact taccgctgtg gagacgtggg
5881 aggtactggc ctccctgaaa cctggtgcag aggaaaagga ggtatatggg ttaatcaatc
5941 aaaggaaatt aatgagacag agccgttcag ttttactgcg aactgtactg gcagtaattt
6001 gggtaatgtc agcggatgtt gcggagaacc aatcacgatt ctcccactag gggcatggat
6061 cgacagtacg caaggtagtt tactaaacc aaaagcgcta ccaccgcaa ttttcctcat
6121 ttgtggggat cgcgcagtc aaggaattcc cagtcgtccg gtagggggcc cctgctattt
6181 aggcaagctt accatgttag caccaccca tacagatatt ctcaaaatac ttgctaattc
6241 gtgcgggaca ggtataagac gtaaacgaag cgtctcacac ctggatgata catgctcaga
6301 tgaagtacag ctttggggtc ctacagcaag aatctttgca tctatcttag ccccggggt

6361 agcagctgcg caagccttaa gagaaattga gagactagcc tgttggtccg ttaaacaggc
6421 taacttgaca acatcactcc tcggggactt attggatgat gtcacgagta ttcgacacgc
6481 ggtcctgcag aaccgagcgg ctattgactt cttgcttcta gtcacaggcc atggctgtga
6541 ggacgttgcc ggaatgtgtt gtttcaatct gagtgatcac agtgaatcta tacagaagaa
6601 gttccagcta atgaagaaac atgtcaataa gatcggcgtg gacagcgacc caatcggaa
6661 ttggctgcga gggatattcg ggggaatagg ggaatgggcc gttcatctgc taaaaggact
6721 gcttttgggg cttgtagtta ttttattgct actggtgtgc ctgccttgcc ttttacaatt
6781 tgtgtctagt agtattcgaa agatgattaa tagttcaatc aactatcata ctgaatacag
6841 gaagatgcag ggcggagcag tctagagctc agttataata atcctgcgaa tcgggctgta
6901 acggggcaag gcttgaccga ggggactata acatgtatag gcgaaaagcg gggctctcgt
6961 tgtaacgcgc ttaggaagtc ccctcgaggt atggcagata tgctcttgca tagggggaaa
7021 aaatgtagtc ttaatatgtt ctgtgtgctg caggagctaa gctgactctg ctgggtgcct
7081 cgcgtagcac tgtggccagg cggtagctgg gacgtgcagc cgaccaccat ggggagcagc
7141 aagagcaagc ctaaggaccc cagccagcgc cggcacagcc tggagccacc cgacagcacc
7201 caccacgggg gattcccagc ctgcgagacc cccgacgaga cagcagcccc cgacgcacac
7261 cgcaacccca gcgcctcctt cgggaccgtg gccaccgagc ccaagctctt ctggggcttc
7321 aacacttctg acaccgtcac gtcgccgagc cgtgccgggg cactggctgg cggcgtcacc
7381 actttcgtgg ctctctacga ctacgagtcc tggactgaaa cggacttgct cttcaagaaa
7441 ggagaacgcc tgcagattgt caacaacacg gaaggtgact ggtggctggc tcattccctc
7501 actacaggac agacgggcta catccccagt aactatgtcg cgccctcaga ctccatccag
7561 gctgaagagt ggtacttttg gaagatcact cgtcgggagt ccgagcggct gctgcttaac
7621 cccgaaaacc cccggggaac cttcttggtc cgggaagagc agacggcaaa ggggtgcctat
7681 tgccctctccg tttctgactt tgacaacgcc aaggggcca atgtgaagca ctacaagatc
7741 tacaagctgt acagcggcgg cttctacatc acctcacgca cacagttcgg cagcctacag
7801 cagctggtgg cctactactc caaacatgct gatggcttgt gccaccgctt ggccaacgct
7861 tgccccacgt ccaagcccca gaccagggga ctgcgcaagg acgcgtggga aatcccccg
7921 gagtgcgtac ggctggaggc gaagctgggg cagggctgct ttggagaggt ctggatgggg
7981 acctggaacg acaccaccag agtggccata aagactctga agcccgccac catgtccccg
8041 gaggccttcc tgcaggaagc ccaagtgatg aagaagctcc ggcattgagaa gctgggttcag
8101 ctgtacgcag tgggtgtcga agagcccatc tacatcgtca ttgagtacat gagcaagggg
8161 agcctcctgg atttctgaa gggagagatg ggcaagtacc tgcggtgcc acagctcgtc
8221 gatatggctg ctgagattgc atccggcatg gcctatgtgg agagaatgaa ctacgtgcac
8281 cgagacctgc gggcgccaa catcctggtg ggggagaacc tgggtgtgaa ggtggctgac
8341 ttcgggctg cagcctcat cgaggacaac gagtacacag cagggcaagg tgccaagttc
8401 cccatcaagt ggacagcccc cgaggcagcc ctctatggcc ggttcacat caagtcggat
8461 gtctggctct tcggcatcct gctgactgag ctgaccacca agggccgggt gccataccca
8521 gggatggtca acagggaggt gctggaccag gtggagaggg gctaccgcat gccctgccg
8581 ccgagtgcc ccgagtcgt gcatgacctc atgtgccagt gctggcgga ggaccctgag
8641 gagcgccca cctttaagta cctgcaggcc cagctgctcc ctgcttgtgt gttggaggtc
8701 gctgagtaag tacgaggcgt gacctacaat tgctcaaata atgcttctgt agaaattgtt
8761 tagcattagc cgtcctgcgt tgctccgcga tgtacgggtc aggtataatg tgcagtttga
8821 ctgaggggac catgatgtgt ataggcgtca agcggggctt cggttgtacg cggataggaa
8881 tccccacag acaattctgc ttggaatatg atggcgtctt ccctgttttg cccttagact
8941 attcgagttg cctctgtgga ttagggctgg aggcagcacg gatagtctga tggccaaata

9001 aggcaggcaa gacagctatt tgtaactgcg aaatacgctt ttgcataggg agggggaaat
9061 gtagtcttat gcaatactcc ttagtcttg caacatgctt atgtaacgat gagttagcaa
9121 tatgccttac aaggaaagaa aaggcacctg gcatgccgat tggtagtagt aaggtagtac
9181 gatcgtgcct tattaggaag gtatcagacg ggtctaacaat ggattggacg aacctgaa
9241 ttccgcatcg cagagatatt gtatttaagt gcctagctcg atacaataaa cgccatttta
9301 ccattcacca cattggtgtg cacctgggtt gatggccgga ccgtcgattc cctaacgatt
9361 gcgaacacct gaatgaagca gaaggcttca tt

//