

Yersinia pestis CO92 plasmid pPCP1

GenBank: AL109969.1

FASTA Graphics

Go to:

LOCUS AL109969 9612 bp DNA circular BCT 26-JUL-2016

DEFINITION Yersinia pestis CO92 plasmid pPCP1.

ACCESSION AL109969

VERSION AL109969.1

DBLINK BioProject: PRJNA34

BioSample: SAMEA1705942

KEYWORDS coagulase; fibrinolysin; IS100; ompT; pesticin; pesticin immunity; pim; pla; pst; rop; transcriptional regulator.

SOURCE Yersinia pestis CO92

ORGANISM Yersinia pestis CO92

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacterales; Yersiniaceae; Yersinia.

REFERENCE 1 (bases 1 to 9612)

AUTHORS Parkhill,J., Wren,B.W., Thomson,N.R., Titball,R.W., Holden,M.T., Prentice,M.B., Sebaihia,M., James,K.D., Churcher,C., Mungall,K.L., Baker,S., Basham,D., Bentley,S.D., Brooks,K., Cerdeno-Tarraga,A.M., Chillingworth,T., Cronin,A., Davies,R.M., Davis,P., Dougan,G., Feltwell,T., Hamlin,N., Holroyd,S., Jagels,K., Karlyshev,A.V., Leather,S., Moule,S., Oyston,P.C., Quail,M., Rutherford,K., Simmonds,M., Skelton,J., Stevens,K., Whitehead,S. and Barrell,B.G.

TITLE Genome sequence of Yersinia pestis, the causative agent of plague

JOURNAL Nature 413 (6855), 523-527 (2001)

PUBMED 11586360

REFERENCE 2 (bases 1 to 9612)

AUTHORS Parkhill,J., Barrell,B.G. and Rajandream,M.A.

TITLE Direct Submission

JOURNAL Submitted (19-AUG-1999) Yersinia pestis sequencing project, The Wellcome Trust Sanger Institute, Genome Campus, Hinxton, Cambridge CB10 1SA E-mail:parkhill@sanger.ac.uk DNA supplied by Dr. Andrey Karlyshev and Prof. Brendan Wren, [3]. Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT

COMMENT Notes:

Details of Y. pestis sequencing at the Sanger Centre are available on the World Wide Web.

(URL, http://www.sanger.ac.uk/Projects/Y_pestis/).

FEATURES Location/Qualifiers

source 1..9612

```

/organism="Yersinia pestis CO92"
/mol_type="genomic DNA"
/strain="CO-92"
/db_xref="taxon:214092"
/plasmid="pPCP1"
/note="biovar:Orientalis"
repeat_region 1..1954
/note="IS100"
gene 87..1109
/gene="YPPCP1.01"
CDS 87..1109
/gene="YPPCP1.01"
/note="YPPCP1.01, probable IS100 transposase, len: 340 aa;
identical to TR:P74993, and similar to e.g. TRA6_BACST
putative transposase for insertion sequence element IS5376
(400 aa), fasta scores; opt: 632 z-score: 1097.4 E(): 0,
38.1% identity in 328 aa overlap. Contains
helix-turn-helix motif from aa 19-40 (Score 2045, +6.15
SD)"
/codon_start=1
/transl_table=11
/product="putative transposase"
/protein_id="CAB53164.1"
/db_xref="GOA:Q7ARN5"
/db_xref="InterPro:IPR001584"
/db_xref="InterPro:IPR009057"
/db_xref="InterPro:IPR012337"
/db_xref="InterPro:IPR017894"
/db_xref="InterPro:IPR025246"
/db_xref="UniProtKB/TrEMBL:Q7ARN5"

/translation="MVTFETVMEIKILHKQGMSSRAIARELGISRNTVKRYLQAKSEP

PKYTTPRAVASLLDEYRDYIRQRIADAHYPKIPATVIAREIRDQGYRGGMTILRAFIR

SLSVPQEQEPAVRFETEPGRQMQVDWGTMRNGRSPLHVFVAVLGYSRMLYIEFTDNMR

YDTLETCHRNAFRFFGGVPREVLVDNMKTVVLQRDAYQTGQHRFHPSLWQFGKEMGFS

PRLCRPFRAQTKGKVERMVQYTRNSFYIPLMTRLRPMGITVDVETANRHGLRWLHDVA

NQRKHETIQARPCDRWLEEQQSMLALPPEKKEYDVHLDENLVNFDKHPLHHPLSIYDS
FCRGVA"
gene 1106..1888

```

```

CDS
    /gene="YPPCP1.02"
    1106..1888
    /gene="YPPCP1.02"
    /note="YPPCP1.02, IS100 ATP-binding protein, len: 260 aa;
    identical to TR:P74994 and similar to e.g. ISTB_ECOLI
    insertion sequence IS21 putative ATP-binding protein (265
    aa), fasta scores; opt: 730 z-score: 1078.9 E(): 0, 47.4%
    identity in 249 aa overlap. Contains PS00017
    ATP/GTP-binding site motif A (P-loop)"
    /codon_start=1
    /transl_table=11
    /product="putative ATP-binding protein"
    /protein_id="CAB53165.1"
    /db_xref="GOA:Q7ARK4"
    /db_xref="InterPro:IPR001270"
    /db_xref="InterPro:IPR002611"
    /db_xref="InterPro:IPR003593"
    /db_xref="InterPro:IPR027417"
    /db_xref="InterPro:IPR028350"
    /db_xref="UniProtKB/TrEMBL:Q7ARK4"

/translation="MMMELQHQRMLALAGQLQLESLISAAPALSQQAVDQEWSYMDFL
EHLHHEEKLARHQRKQAMYTRMAAFPAVKTFEEYDFTFATGAPQKQLQSLRSLSFIER
NENIVLLGPSGVGKTHLAIAMGYEAVRAGIKVRFTTAADLLLQLSTAQRQGRYKTTLQ
RGVMAPRLLIIDEIGYLPFSQEEAKLFFQVIKRYEKSAMILTSNLPFGQWDQTFAGD
AALTSAMLDRIHHSHVQIKGESYRLRQKRKAGVIAEANPE"
misc_feature    1433..1456
    /gene="YPPCP1.02"
    /note="PS00017 ATP/GTP-binding site motif A (P-loop)"
repeat_region   2265..2283
repeat_region   2321..2340
repeat_region   complement(2413..2432)
misc_feature    2924..4176
    /note="similar to ColE1 ori"
gene            2925..3119
    /gene="YPPCP1.03"
    /gene_synonym="rop"
CDS            2925..3119
    /gene="YPPCP1.03"
    /gene_synonym="rop"
    /note="YPPCP1.03, rop, plasmid replication regulatory

```

protein, len: 64 aa; highly similar to e.g. ROP_ECOLI regulatory protein ROP (RNA one modulator) from ColE1 (63 aa), fasta scores; opt: 241 z-score: 482.0 E(): 1.5e-19, 63.9% identity in 61 aa overlap."

/codon_start=1

/transl_table=11

/product="putative replication regulatory protein"

/protein_id="CAB53166.1"

/db_xref="InterPro:IPR000769"

/db_xref="UniProtKB/TrEMBL:Q9S1E1"

/translation="MNKQQQTALNMARFIRSQSLILLEKLDALDADEQAAMCERLHEL
AEELQNSIQARFEAESETGT"

gene 4355..4780

/gene="YPPCP1.04"

/gene_synonym="pim"

CDS 4355..4780

/gene="YPPCP1.04"

/gene_synonym="pim"

/note="YPPCP1.04, pim, pesticin immunity protein, len: 141
aa; identical to TR:Q57285 (EMBL:Z54145, X92856)"

/codon_start=1

/transl_table=11

/product="pesticin immunity protein"

/protein_id="CAB53167.1"

/db_xref="UniProtKB/TrEMBL:Q57285"

/translation="MISKLFCLALIFLSSSGLAEKNTYTAKDILQNLELNTFGNSLSH

GIYGKQTTFKQTEFTNIKSNTKKHIALINKDNSWMISLKILGKRDEYTVCFEDFSLI
RPPTYVAIHPLLIKVKSGNFIVVKEIKKSIPGCTVYYH"

gene complement(4815..5888)

/gene="YPPCP1.05c"

/gene_synonym="pst"

CDS complement(4815..5888)

/gene="YPPCP1.05c"

/gene_synonym="pst"

/note="YPPCP1.05c, pst, pesticin, len: 357 aa; identical
to TR:Q57159 (EMBL:X92856)"

/codon_start=1

/transl_table=11

/product="pesticin"

/protein_id="CAB53168.1"

/db_xref="GOA:Q57159"

/db_xref="InterPro:IPR023346"
/db_xref="InterPro:IPR023347"
/db_xref="PDB:4AQN"
/db_xref="PDB:4ARJ"
/db_xref="PDB:4ARL"
/db_xref="PDB:4ARM"
/db_xref="PDB:4ARP"
/db_xref="PDB:4ARQ"
/db_xref="PDB:4EPF"
/db_xref="PDB:4EPI"
/db_xref="PDB:4EXM"
/db_xref="UniProtKB/TrEMBL:Q57159"

/translation="MSDTMVNGSGGVPAFLFSGSTLSSYRPNFEANSITIALPHYVD

LPGRSNFKLMYIMGFPIDTEMEKDSEYSNKIRQESKISKTEGTVSYEQKITVETGQEK

DGVKVYRVMVLEGTIAESIEHLDKKENEDILNNNRNRIVLADNTVINFDNISQLKEFL

RRSVNIVDHDIFSSNGFEGFNPTSHFPSNPSSDYFNSTGVTFGSGVDLGQRSKQDLLN

DGVPQYIADRLDGYMMLRGKEAYDKVRTAPLTLSDNAHLLSNIIYIDKFSHKIEGLFN

DANIGLRFSDLPLRTRTALVSIGYQKGFKLSRTAPTWNKVIKDWNGLVNAFNIVD

GMSDRRKREGALVQKDIDSGLLK"

misc_feature 5910..5917

/note="difference: AF053945 has 7 A's this sequence has 8"

misc_feature 5927..5934

/note="difference: AF053945 has 9 A's this sequence has 8"

gene 6006..6422

/gene="YPPCP1.06"

CDS 6006..6422

/gene="YPPCP1.06"

/note="YPPCP1.06, unknown, len: 138 aa"

/codon_start=1

/transl_table=11

/product="hypothetical protein"

/protein_id="CAB53169.1"

/db_xref="UniProtKB/TrEMBL:Q9S1E0"

/translation="MKFHFCDLNHSYKNQEGKIRSRKTAPGNIRKKQKGDNVSKTKSG

RHRLSKTDKRLAALVVAGYEERTARDLIQKHVYTLTQADLRHLVSEISNGVGQSQAY

DAIYQARRIRLARKYLSGKKPEGVEPREGQEREDLP"

gene	6665..7603
	/gene="YPPCP1.07"
	/gene_synonym="pla"
CDS	6665..7603 /gene="YPPCP1.07" /gene_synonym="pla" /note="YPPCP1.07, pla, coagulase/fibrinolysin precursor (EC 3.4.21.-) (plasminogen activator), len: 312 aa; identical to COLY_YERPE. Contains PS00834 and PS00835 Serine proteases, omptin family signatures 1 and 2" /codon_start=1 /transl_table=11 /product="coagulase/fibrinolysin precursor" /protein_id="CAB53170.1" /db_xref="GOA:P17811" /db_xref="InterPro:IPR000036" /db_xref="InterPro:IPR020079" /db_xref="InterPro:IPR020080" /db_xref="InterPro:IPR023619" /db_xref="PDB:2X4M" /db_xref="PDB:2X55" /db_xref="PDB:2X56" /db_xref="PDB:4DCB"
	/translation="MKKSSIVATIITILSGSANAASSQLIPNISPDSTVAASTGMLS
	GKSHEMLYDAETGRKISQLDWKIKNVAILKGDISWDPYSFLTINARGWTSLASGSGNM
	DDYDWMNENQSEWTDHSSHPATNVNHANEYDLNVKGWLLQDENYKAGITAGYQETRFS
	WTATGGSYSYNNGAYTGNFPKGVRVIGYNQRFSPYIGLAGQYRINDFELNALFKFSD
	WVRAHDNDEHYMRDLTFREKTSGRYYGTVINAGYYVTPNAKVFAEFTYSKYDEGKGG
	TQTIDKNSGDSVSIIGDAAGISNKNYTVTAGLQYRF"
misc_feature	7007..7036 /gene="YPPCP1.07" /gene_synonym="pla" /note="PS00834 Serine proteases, omptin family signature 1"
misc_feature	7118..7168 /gene="YPPCP1.07" /gene_synonym="pla" /note="PS00835 Serine proteases, omptin family signature 2"

```

gene      complement(7790..8089)
          /gene="YPPCP1.08c"
CDS       complement(7790..8089)
          /gene="YPPCP1.08c"
          /note="YPPCP1.08c, probable transcriptional regulator,
          len: 99 aa; similar to several predicted transcriptional
          regulators e.g. TR:P95258 (EMBL:Z84498) Mycobacterium
          tuberculosis MTCY09F9.0 8C (Rv1956) (149 aa), fasta
          scores; opt: 190 z-score: 314.7 E(): 3.1e-10, 35.4%
          identity in 99 aa overlap. Contains helix-turn-helix motif
          from aa 42-63 (Score 2083, +6.28 SD)"
          /codon_start=1
          /transl_table=11
          /product="putative transcriptional regulator"
          /protein_id="CAB53171.1"
          /db_xref="GOA:Q9S1D9"
          /db_xref="InterPro:IPR001387"
          /db_xref="InterPro:IPR010982"
          /db_xref="UniProtKB/TrEMBL:Q9S1D9"

/translation="MRTLDEVIASRSPESQTRIKEMADEMILEVGLQMMREELQLSQK
QVAEAMGISQPAVTKLEQRGNDLKLATLKRYVEAMGGKSLDVELPTGRRVAFHV"
gene      complement(8089..8436)
          /gene="YPPCP1.09c"
CDS       complement(8089..8436)
          /gene="YPPCP1.09c"
          /note="YPPCP1.09c, unknown, len: 115 aa; similar to two
          hypothetical proteins from M. tuberculosis TR:O53468
          (EMBL:O53468) MTV018.09C (Rv2022c) (201 aa), fasta scores;
          opt: 271 z-score: 398.5 E(): 6.7e-15, 41.7% identity in
          108 aa overlap, and TR:O53332 (EMBL:AL021646) MTV014.26
          (Rv3182) (114 aa), fasta scores; opt: 214 z-score: 320.5
          E(): 1.5e-10, 39.4% identity in 104 aa overlap"
          /codon_start=1
          /transl_table=11
          /product="hypothetical protein"
          /protein_id="CAB53172.1"
          /db_xref="InterPro:IPR009241"
          /db_xref="UniProtKB/TrEMBL:Q9S1D8"

/translation="MWMVLFSQRFDDWLNEQEDALQEKVLADLKKLQVYGPELPRPYA
DTVKGSRYKNMKELRVQFSGRPIRAFYAFDPIRRAIVLCAGDKSNDKRFYEKLVRIAE

```

DEFTAHLENTLESK"

misc_feature 8528..8532

/note="difference: AF053945 has 3 T's this sequence has 5"

ORIGIN

1 tgaacgaac ggtgcaatag tgatccacac ccaacgcctg aaatcagatc cagggggtaa
61 tctgctctcc tgattcagga gagtttatgg tcacttttga gacagttatg gaaattaaaa
121 tctgcacaa gcaggggaatg agtagccggg cgattgccag agaactgggg atctcccgca
181 ataccgttaa acgttatttg caggcaaaat ctgagccgcc aaaatatacg ccgcgacctg
241 ctgttgcttc actcctggat gaataccggg attatattcg tcaacgcacg gccgatgctc
301 atccttacia aatcccggca acggtaatcg ctgcgcgagat cagagaccag ggaatcgtg
361 gcggaatgac cattctcagg gcattcattc gttctctctc gggtcctcag gagcaggagc
421 ctgccgttcg gttcgaaact gaacccggac gacagatgca gggtgactgg ggcactatgc
481 gtaatggtcg ctaccgctt cacgtgttcg ttgctgttct cgatacagc cgaatgctgt
541 acatcgaatt cactgacaat atgcgttatg acacgtgga gacctgcat cgtaatgcgt
601 tccgcttctt tgggtggtg cgcgcggaag tgttgatga caatatgaaa actgtggctc
661 tgcaacgtga cgcataatcag accggtcagc accggttcca tcctcgctg tggcagttcg
721 gcaaggagat gggcttctct ccccgactgt gtcgcccctt cagggcacag actaaaggta
781 aggtggaacg gatggtgcag tacacccgta acagtttta catccacta atgactcgc
841 tgcgcccgat ggggatcact gtcgatgttg aaacagccaa ccgccacggg ctgcgctggc
901 tgcacgatgt cgtaaccaa cgaaagcatg aaacaatcca ggcccgctcc tgcgatcgt
961 ggctcgaaga gcagcagtc atgctggcac tgctccgga gaaaaagag tatgacgtgc
1021 atcttgatga aaatctggtg aactcgaca aacacccct gcatcatcca ctctcatct
1081 acgactcatt ctgcagagga gtggcgtgat gatggaactg caacatcaac gactgatggc
1141 gctcgccggg cagttgcaac tggaaagcct tataagcgca gcgcctgcgc tgcacaaca
1201 ggcagtagac caggaatgga gttatatgga ctctcggag catctgctc atgaagaaaa
1261 actggcacgt catcaacgta aacaggcgat gtataccga atggcagcct tccggcggt
1321 gaaaacgttc gaagagtatg acttcacatt cgccaccgga gcaccgcaga agcaactcca
1381 gtcgttacgc tactcagct tcatagaacg taatgaaaat atcgtattac tggggccatc
1441 aggtgtgggg aaaaccatc tggcaatagc gatgggctat gaagcagtc gtgcaggtat
1501 caaagttcgc ttcacaacag cagcagatct gttactcag ttatctacgg cacaacgtca
1561 gggccgttat aaaacgacgc ttcagcgtgg agtaatggcc cccgcctgc tcatattga
1621 tgaaataggc tatctgccg tcaagcagga agaagcaaag ctgttctcc aggtcatcgc
1681 taaacgttac gaaaagagcg caatgatcct gacatccaat ctgccgttcg ggcagtgga
1741 tcaaacgttc gccggtgatg cagcactgac ctacgcgatg ctggaccgta tcttacacca
1801 ctacatgtc gttcaaatca aaggagaaag ctatcgactc agacagaaac gaaaggccgg
1861 ggttatagca gaagctaac ctgagtaaaa cggtgatca atattgggcc gttggtggag
1921 atataagtg atcactttc atccgtcgtt gacacctga tgaattcac gtgtcacgcc
1981 tgaataacaa gaatgccgga gatacgcagt catattttt acacaattct ctaatcccga
2041 caaggtcgta ggtcgttata ggaaaattct tagcaccatt ccggaacaat cagaacagca
2101 ggcatgaac gactgacaac attacgaata taaaaacgc acccgggcca gacattcccc
2161 ctactgatta aaccagccgg actgtccac ggaacggtct ttttaaccg acacacagtc
2221 tgagtacaga tacatgtcac gatgatgcag gattagcgga agagtgtgag cacgtttccg
2281 ggaactgtg tgaaccatag ctcaatattc gagtgagggc ataccgaaa cgcgctcaga
2341 ttcgttgtaa cgcgatttc cgtaccggg aatttttca gttgtttt cgtttcatgt

2401 cgtcagaaac gttctgagcg cgtttccggc atctgatgct acgcaaacca tccccatggt
2461 cagttgacag ccggaacac gcgggtgtcg ttttagcgta tcgacgggac ggcgtcgaga
2521 agcacaaaa acagatgttg tactcagtca gttgtttac agacagcact gcggcagatt
2581 gaaaaagtac cgtacttca ggaatgtcca gaaacatgt gtcagacttc gtttcccc
2641 ttccgggtga attttttgt catccgttca ggaatctct tataacgatt actccatttc
2701 aggatTTTT atgtggcgtt tactacaggc aggatattca aaggcaaaaa aatcccccg
2761 aacaggcgga acccgacag ggggagaacg aatcgctaaa taatttcgt agttgtattt
2821 cccatcgttg ctactgaac gggatgaatt tgccgcagtt tatcctgtaa aacaatcctg
2881 atttactcac actcacata tcactgacgg agcacaacgg aatagtgaac aaacaacaac
2941 aaactgcgct gaatatggcg cgattatca gaagccagag cctgatactg cttgaaaaac
3001 tggatgctct ggatgccgac gagcaggcgg ccatgtgtga acgactgcac gaactgcggg
3061 aagaactcca gaacagcatc caggctcgtt ttgaagccga aagtgaaca ggaacataac
3121 gaagtcctcg gagacgggtc cagcttgtct gtgaacggat gccgggagca gacaagcccc
3181 tcagggcgcg tcagcgggtt ttacgggtg tcggggcgca gccatgacct agtcacgtag
3241 cgatagcgga gtgtatactg gcttagtcat gcggcatcag tgcggattgt atgaaaagt
3301 caccatgtac ggtgtgaaat gccgcacaga tgcgtaagga gaacatgcag atgccgatgc
3361 tctccgctt cctcgtcac tgactcgtg cgctcggtcg ttccggtcgc gcgagcgggt
3421 tctgctact caaaagcggg gatactgta tccacacaat caggggataa cgccggaaag
3481 aacatgtgag caaaaaacga agacccaga aaaggccgcg ccggaggcgc ttttcata
3541 ggctccgcc ccctgacgag catcacaaaa atcgacgctc aagtcagagg tggcgaaacc
3601 cgacaggact taaagatacc agcgtttcc ccccggaagc tccctcgtgc gctctcgt
3661 tccgacctg ccgcttaccg gataccttc cgcttttc ccttcgggaa gcgtggcgt
3721 ttctcatagc tcacgtggtt ggtatctcag ttccgtgtag gtcgttcgct ccaagctggg
3781 ctgtgtgcac gaacccccg ttacgcccga cactgcgcc ttatccggt actatcgtct
3841 tgagtccaac ccgtaagac acgactttac gccactggca gcagccattg gtaactgaaa
3901 agtggattta gatacgaga actcttgaag ttgaagcctt atcgcggtc cactgaaagg
3961 acagcatttg gtatctgtgc tccacttaag ccagctacca caggttagaa agcctgagaa
4021 acttctaacc ttgaaagaa cccacgcctg agaacgtggg tttttcgtt tacaggcagc
4081 agattacgag cagaaaaaaa ggaatcaag aagatcctt gatctttct actgaattgc
4141 gctcccgatc agttcagcag aagattatga tggggttcta tgggtattgc tgcggtaaca
4201 cccatgttac ttgaggtgt atgtagtctg tgtagaatta tacacataag gcttaaactg
4261 ctctttttt tcaatatgca attggaagt cattgactac ataatagat tattccaaat
4321 aattatttta tgtaagaaca ggaatggagg ggaatgatc tcaaagtat ttgcttggc
4381 tctcatattt ttatcatcaa gtggcctgc agaaaaaac acatatacag caaaagacat
4441 cttgaaaaac ctagaattaa ataccttgg caattcattg tctcatggca tctatgggaa
4501 acagacaacc ttcaagcaaa ccgagtttac aaatattaaa agcaacacca aaaaacacat
4561 tgcacttatc aataaagaca actcatggat gatataccta aaaatactag gaattaagag
4621 agatgagtat actgtctgtt ttgaagatt ctctctaata agaccgcaa catatgtagc
4681 catacatcct ctactataa aaaaagtaaa atctggaac ttatagtag tgaaagaaat
4741 aaagaaatct atccctggtt gcactgtata ttatcattaa tagcaagccc ctattatta
4801 tgaggggctc atggttatt taacaatcca ctatcgatat cttttgcac cagagcgccc
4861 tctcgtttac gtctgcaga cattccatca acaatattat taaaagcatt tacaaggcca
4921 ttccagtctt ttgcgataac tttatccat actgtgggag cagttctgga taactaaac
4981 ccttttgat atccaataga caccagtgt gtacgggtc tcaacggtaa atcgctgaac

5041 cgaagaccga tattagcgtc attgaaaaga ccttcaatct tatgtgagaa ttatcaata
 5101 taaatattag ataagagatg agcttcatta tcagaaagcg tcagaggtgc tgttctcact
 5161 ttatcataag cctccttccc tcgaagcata taatacccat caagtctatc tgcaataac
 5221 tgagggacac cgtcattcaa taaatcctgt ttgcttcgct gaccaaggtc aaccccgaa
 5281 ccgaatgtaa caccgggtact gttaaaataa tcgctactag gattagacgg aaaatgactt
 5341 gtcggattaa acccttcaaa accattactg gagaaaatat cgtgggtcaac aatatttacc
 5401 gaacgacgta aaaattcctt cagttgacta atattgtcaa agttaatgac agtgttgctc
 5461 gctaggacga tgcgatttcg gttattatc agaattgtctt cgttctctt cttatcgaga
 5521 tgttcaatag attcggcaat cgttccctca agaaccatga caggtagac ttccacaccg
 5581 tcttttctt gacctgttc aacagttatt ttctgttcgt aagacacggc cccttcagtt
 5641 tttgaaattt tactttctg gcggatctta ttgaatatt cactgtctt ctccatctcc
 5701 gtatcaatcg gaaaccccat aatgtacatc agtttaaat tactccggcc aggcagatcc
 5761 acataatgtg gtaatgcaat tgtaatcgaa ttagcttcaa aatttggtct gtaactgctt
 5821 aatgtacttc cggaaaagag aaaagccgga acaccactg aaccattcac taccattgta
 5881 tctgacataa aaattcctt ttaacacata aaaaaaaca taagttaaa aaaatactgt
 5941 acataaaacc actgtttta tgtacagtaa taaaattacg ccgctttatt ttctctgtca
 6001 ataatatgaa atttcattt ttgatctga atcactcta taaaatcag gaagggaaga
 6061 ttgcgacgag aaaaacagca ccgggtaaca tcagaaaaaa acagaaagga gataacgtga
 6121 gcaaaacaaa atctggtcgc caccgactga gcaaaacaga caaacgctg ctgggtgcac
 6181 ttgtcgttc cggatagcaa gaacggacag cccgtgacct catccagaaa cagtttaca
 6241 cactgacaca ggccgacctg cgccatctgg tcagtgaat cagtaacggc gtgggacagt
 6301 cacaggccta cgatgcgatt taccaggcga gacgcattcg tctgcccgt aaatacctga
 6361 gcggaaaaa accggaaggg gtggaacccc gggaaggga ggaacggga gattaccat
 6421 aactccggtt atcagtacca tcggctcaac gtcggtgct ggatctgaaa aattcgctca
 6481 aaagatcata ttccctgga tttttccac cgttcttat gtgagcaaag tcacataatt
 6541 ctgtcagacg acgagaaaac ggatatcgat tattgtttaa ttttttaca ttataaaaa
 6601 tgaaattaga taatcagata caaataatat gtttcgttc atgcagagag attaagggtg
 6661 tctaatgaag aaaagttcta ttgtggcaac cattataact attctgtccg ggagtgtcaa
 6721 tgcagcatca tctcagttaa taccaatat atcccctgac agctttacag ttgcagcctc
 6781 caccgggatg ctgagtggaa agtctcatga aatgctttat gacgcagaaa caggaagaaa
 6841 gatcagccag ttagactgga agatcaaaaa tgtcgctatc ctgaaagggtg atatatcctg
 6901 ggatccatac tcatttctga cctgaatgc cagggggtgg acgtctctgg ctccgggtc
 6961 aggtaatatg gatgactacg actggatgaa tgaaaatcaa tctgagtga cagatcactc
 7021 atctcatcct gctacaaatg ttaatcatgc caatgaatat gacctcaatg tgaaaggctg
 7081 gttactccag gatgagaatt ataaagcagg tataacagca ggatatcagg aaacacgttt
 7141 cagttggaca gctacagggtg gttcatatag ttataataat ggagcttata ccggaaactt
 7201 ccgaaagga gtgcgggtaa taggttataa ccagcgctt tctatgcat atattggact
 7261 tgcaggccag tatgcatta atgattttga gttaaatgca ttatttaaat tcagcgactg
 7321 ggttcgggca catgataatg atgagcacta tatgagagat ctactttcc gtgagaagac
 7381 atccggctca cgttattatg gtaccgtaat taacgctgga tattatgtca cacctaatgc
 7441 caaagtcttt gcggaattta catacagtaa atatgatgag ggcaaaggag gtactcagac
 7501 cattgataag aatagtggag attctgtctc tattggcgga gatgctgccg gtatttcaa
 7561 taaaaattat actgtgacgg cgggtctgca atatcgcttc tgaataaac agatcatatc
 7621 tctctttca tctcccta gcggggagga tgtctgtgga aaggagggtg gtgttgacc

7681 aaccttcaga tgtgtgaaaa atcacctttt tcaccataat gacggggcgc tcattctgtt
7741 gttttgcctt gacattctcc acgtctttca gggcatggag aagggtcaa tagacatgga
7801 acgctactct ctttctgta ggaagctcaa catccaagct taatttgctt cccattgctt
7861 caacgtaacg cttaacgtc gccagcttta aatcatttcc gcgctgctcc agctttgtta
7921 ctgctggctg gcttataccc atcgccctag caacttgttt ttgtgataac tggagtctt
7981 cacgcatcat ctgcaagccg acctcaagaa tcattctatc tgccatttct ttaattcggt
8041 tctggctttc aggtgaacga ctggcaatca cctcatctaa tgttctcatt acttgctctc
8101 cagtgtgttc agatgtgctg taaattcatc ctgagctata cgcaccagtt ttcataaaa
8161 ccgcttatca ttacttttat ctctgcaca aagaacgata gcccgcgaa tcggatcgaa
8221 cgataaaaag gctcttatcg gacggccaga aaactgaacg cgaagctctt tcatttttt
8281 gtaccgagaa ctttcacgg tatcggcata tggcctgggt aactcaggtc cgtaaacctg
8341 tagctttttc aaatcagcca aaacctttc ctgaagagcg tcttctgct catttagcca
8401 gtcgtcaaat cgctggctaa aaagtacat ccacatgctc aacctataa cctgtagctt
8461 accccactaa caatataacc tacgagttat attttcaaga aaagctggct atttaacata
8521 acggcaattt ttgtacgcac cactgaaatg cgttcagcgc gatcacggca acagacaggc
8581 aaaaatagca acaaactcc cgaaaaaccg ccgcatcgc gcctgataaa ttttaacct
8641 atgcatatct atgcagccag gcgaatcac aacgaattgc ctgcctgatg taactgaaac
8701 ggggtgtttt tctgatttg gtgggcgtgg aagacggaac atgaacggga aaacagaatt
8761 catgccagat gagcgcatc tggcaattaa ggcaaacac agcaacaaag acacgccaga
8821 atcgcgcccgc gatatgtttt aacgcgattt tcagactcag acaaattcag cagaatgcta
8881 ctccattcac cgggctgatg gtgaatacat gcgtatccag gatgagtaca tttctggctc
8941 tgccacagct ctgtctgttg gcagctttcg cctgtccgga aacctgcta aaacgtccc
9001 gaaaggcctc tgaaccagaa agcaacaaaa cacaggccat taagtaaacc gcgttaaac
9061 acgtctgatg gattgctgca aaaaaagtc cctaaggag cagggactgt taaaccagt
9121 gaatagcgtc taaattaaag taagaatacg accaggtact cttcagaaaa gagattaatc
9181 caccgcacag aataatcaac agtaaaaaa aacaaccctg attttttatt tttcttttt
9241 tcgataaaaa caaaattaa gaaataatta atcagaacat tccttaactt cagggcattg
9301 cctgtgttcc attttgat tagtctgaaa ctccgaagg tggataacac ccggtattt
9361 tttgctaca taaagcccct cttcaggca gaggggctt ttttgcca ccacataaaa
9421 aaggccctca caggaggtgt tctgtaggg cgatgataa ggactgaatc gatggtaat
9481 atgtctagtc ctgacttttg catctccgaa tataaaacc tgtttaacgg catgcaaac
9541 caaaaaataa aaatgtgaca tcgcaatgcc agataatatt gacgcatgag ggaatgcgta
9601 ccccgacccc tg

//