

Dataset S2 (Construct VIII of Fig. 3A)

TITLE Polyprotein strategy for stoichiometric assembly of nitrogen fixation components for synthetic biology

AUTHORS Jianguo Yang, Xiaqing Xie, Nan Xiang, Zhe-Xian Tian, Ray Dixon, Yi-Ping Wang

FEATURES Location/Qualifiers

source 1..27022
 /organism="synthetic DNA construct"
 /lab_host="Escherichia coli"
 /mol_type="other DNA"

source 5339..5359
 /organism="synthetic DNA construct"
 /lab_host="Escherichia coli"
 /mol_type="other DNA"

promoter 75..107
 /note="tet promoter"
 /note="E. coli promoter for tetracycline efflux protein gene"

CDS 155..1345
 /codon_start=1
 /gene="tet"
 /product="tetracycline efflux protein"
 /note="Tc^R"
 /note="confers resistance to tetracycline"
 /translation="MKSNNALIVILGTVTLDAVGIGLVMPVLPGLLRDIVHSDSIASHY
 GVLLALYALMQFLCAPVLGALSDFRGRPVLLASLLGATIDYAIMATTPVLWILYAGRI
 VAGITGATGAVAGAYIADITDGEDRARHFGLMSACFGVGMVAGPVAGLLGAISLHAPF
 LAAAVLNGLNLLLGCFMQESHKGERRPMLRAFNPVSSFRWARGMTIVAALMTVFFIM
 QLVGQVPAALWVIFGEDRFRWSATMIGLSLAVFGILHALAQAFVTGPATKRFGEKQAI I
 AGMAADALGYVLLAFATRGWMAFPIMILLASGGIGMPALQAMLSRQVDDDHQQLQGS L
 AALTSLTSITGPLIVTAIYAASASTWNGLAWIVGAALYLVLCPALRRGAWSRATST"

CDS complement(2379..3038)
 /codon_start=1
 /gene="cat"
 /product="chloramphenicol acetyltransferase"
 /note="Cm^R"
 /note="confers resistance to chloramphenicol"
 /translation="MEKKITGYTTVDISQWHRKEHFQAFQVAQCTYNQTVQLDITAF L
 KTVKKNKHKFYPAFIHILARLMNAHPEFRMAMKDGELVIWDSVHPCYTVFHEQTETFSS
 LWSEYHDDFRQFLHIYSQDVACYGENLAYFPKGF IENMFFVSANPWVSFTSFDLNVANM
 DNFFAPVFTMGKYYTQGDKVLMLPLAIQVHHAVCDGFHVGRMLNELQQYCDEWQGGA"

promoter complement(3039..3141)
 /note="cat promoter"
 /note="promoter of the E. coli cat gene"

rep_origin complement(3667..4212)
 /direction=LEFT
 /note="p15A ori"

/note="Plasmids containing the medium-copy-number p15A origin of replication can be propagated in E. coli cells that contain a second plasmid with the ColE1 origin."
 terminator 4434..4449
 /note="Terminator-nifE"
 CDS 4496..5317
 /codon_start=1
 /note="nifU"
 /translation="MWNYSEKVKDHHFFNPRNARVVDNANAVGDVGSLSGCDALRLMLRV
 DPQSEIIEEAGFQTFGCGSAIASSSALTELIIGHTLAEAGQITNQQIADYLDGLPPEKM
 HCSVMGQEALRAAIANFRGESLEEEHDEGKLICKCFGVDEGHIRRAVQNNGLTTLAEVI
 NYTKAGGGCTSCHEKIELALAEILAQQPQTTPAVASGKDPHWQSVVDTIAELRPHIQAD
 GGDMLLSVTNHQVTVSLSGSCSGCMMTDMTLAWLQQKLMERTGCYMEVVA"
 CDS 5318..5338
 /codon_start=1
 /product="tobacco etch virus (TEV) protease recognition and cleavage site"
 /note="TEV site"
 /translation="ENLYFQS"
 CDS 5339..5359
 /codon_start=1
 /product="tobacco etch virus (TEV) protease recognition and cleavage site"
 /note="TEV site"
 /translation="ENLYFQS"
 CDS 5360..6559
 /codon_start=1
 /note="nifS"
 /translation="KQVYLDNNATTRLDPMVLEAMMPFLTDFYGNPSSIHDFGIPAQAA
 LERAHQQAAALLGAEYPSEIIFTSCATEATATAIASAIALLPERREIITSVVEHPATLA
 ACEHLERQGYRIHRIAVDSEGALDMAQFRAALSPRVALVSMWANNETGVLFPIGEMAE
 LAHEQGALFHCDAVQVVGKIPIAVGQTRIDMLSCSAHKFHGPKGVGCLYLRRGTRFRPL
 LRGGHQEYGRRAGTENICGIVGMGAACELANIHLPGMTHIGQLRNRLEHRLLASVPSVM
 VMGGQPRVPGTVNLAFEFIEGEAILLLLNQAGIAASSGSACTSGSLEPSHVMRAMNIP
 YTAAHGTRFSLSRYTREKEIDYVVATLPPIIDRLRALSPYWQNGKPRPADAVFTPVYG
 "
 terminator 6580..6602
 /note="Terminator-nifUF"
 CDS 6903..7781
 /codon_start=1
 /note="nifH"
 /translation="MTMRQCAIYGKGGIGKSTTTQNLVAALAEMGKVMIVGCDPKADS
 TRLILHAKAQNTIMEMAAEVGSVEDLELEDVLQIGYGDVRCAESGGPEPGVCAGRGI
 TAINFLEEEGAYEDDLDFVFDVLDVCGGFAMP IRENKAQE IYIVCSGEMMAMYAAN
 NISKGIVKYAKSGKVRLLGGLICNSRQTDREDELI IALAELKGTQMIHFVPRDNIVQRAE
 IRRMTVIEYDPACKQANEYRTLAQKIVNNTMKVVPTPCTMDELESLLMEFGIMEEEDTS
 IIGKTAEEENAA"
 CDS 7782..7802

```

/codon_start=1
/product="tobacco etch virus (TEV) protease recognition and
cleavage site"
/note="TEV site"
/translation="ENLYFQS"
CDS 7803..7823
/codon_start=1
/product="tobacco etch virus (TEV) protease recognition and
cleavage site"
/note="TEV site"
/translation="ENLYFQS"
CDS 7824..9272
/codon_start=1
/note="nifD"
/translation="MMTNATGERNLALIQEVLEVPETARKERRKHMVSDPEMESVGK
CIISNRKSQPGVMTVRGCAYAGSKGVVFGPIKDMAHISHGPGCGQYSRAGRNNYITGV
SGVDSFGTLNFTSDFQERDIVFGGDKKLSKLIEMELLFPLTKGITIQSECPVGLIGDD
ISAVANASSKALDKPVIPVRCEGFRGVSQSLGHHIANDVVRDWILNNREGQPFETPYD
VAIIGDYNIGGDAWASRILLEEMGLRVVAQWSGDGTLVEMENTPFVKLNLVHCYRSMNY
IARHMEEKHQIPWMEYNFFGPTKIAESLRKIADQFDDTIRANAEAVIARYEGQMAAIIA
KYRPRLEGRKVLVLYMGLRPRHVI GAYEDLGMEIIAAGYEFANDDYDRTL PDLKEGTL
LFDDASSYELEAFVKALKPDLIGSGIKEKYIFQKMGVPPFRQMHSWDYSGPYHG YDGF AI
FARDMDMTLNNPAWNELTAPWLKSA"
CDS 9273..9293
/codon_start=1
/product="tobacco etch virus (TEV) protease recognition and
cleavage site"
/note="TEV site"
/translation="ENLYFQS"
CDS 9294..9314
/codon_start=1
/product="tobacco etch virus (TEV) protease recognition and
cleavage site"
/note="TEV site"
/translation="ENLYFQS"
CDS 9315..10877
/codon_start=1
/note="nifK"
/translation="MSQTIDKINSCYPLFEQDEYQELFRNKRQLEEAHDAQRVQEVFAW
TTTAEYEALNFQREALTVDPKACQPLGAVLCSLGFANTLPYVHGSQGCVAYFRTYFNR
HFKEPIACVSDSMTEDAAVFGGNNMNLGLQNASALYKPEIIAVSTTCMAEVI GDDLQA
FIANAKKDGFDVSSIAVPHAHTPSFIGSHVTGWDNMFEGFAKTFTADYQQGPGKLPKLN
LVTGFETYLG NFRVLKRMMEQMAVPCSLSDPSEVLDT PADGHYRMYSGGTTQQEMKEA
PDAIDTLLLQPWQLLKS KKVQEMWNQPATEVAIPLGLAATDELLMTVSQ LSGKPIADA
LTLERGR LVDMM LDSHTWLHGK KFGLYGDPDFVMGLTRFLLEL GCEPTVILSHNANKRW
QKAMNKMLDASPYGRDSEVF INCDLWHFRSLMFTRQPDFMIGNSYKGF IQRDTLAKGKA
FEVPLIRLGFPLFDRHHLHRQTTWGYEGAMNIVTTLVNAVLEKLDSDTSQLGKTDYSFD
LVR"

```

terminator 10897..10915
 /note="Terminator1-nifH"
 terminator 11014..11040
 /note="Terminator2-nifH"
 terminator 11101..11131
 /note="Terminator-nifL"
 CDS complement (11154..12557)
 /codon_start=1
 /note="nifB"
 /translation="TSCSSFSGGKACRPADDSALTPLVADKAAAHPCYSRHHHRFARM
 HLPVAPACNLQCNYCNRFDCSNESRPGVSSSTLLTPEQAVVKVRQVAQAIPQLSVVGI
 A GPGDPLANIARTFRTLELIREQLPDLKLCSTNGLMLPDAVDRLLDVGVHDVTVTINTL
 DAETAAQIYAWLWLDGERYSGREAGEIL IARQLEGVRRLTAKGVLVKINSVLIPGINDS
 GMADVSRALRASGAFIHNIMPLIARPEHGTVFGLNGQPEPDAETLAATRSRCDEVMPQM
 THCHQCRADAIGMLGEDRSQQFTQLPAPESLPAWLPILHQAQLHASIATRGESEADDA
 CLVAVASSRGDVIDCHF GHADRFYIYLSAAGMVLVNERFTP KYCQGRDDCEPQDNAAR
 FAAILELLADV KAVFCVRIGHTPWQQLEQEGIEPCVDGAWRPVSEVLPAAWWQRRGSWP
 AALPHKGVA"
 CDS complement (12558..12602)
 /codon_start=1
 /note="GGGGS X 3"
 /translation="GGGSGGGSGGGGS"
 CDS complement (12603..13985)
 /codon_start=1
 /note="nifN"
 /translation="MADIFRTDKPLAVSPIKGTGQPLGAILASLGLIEHSIPLVHGAQGCS
 AFAKVFFIQHFHDPVPLQSTAMDPTSTIMGADGNIFTALDTLCQRNNPQAI VLLSTGLS
 EAQGS DISRVVRQFREEYPRHKGVA ILLTVNTPDFY GSMENGFSAVLESVIEQWVPPAPR
 PAQRNRRVNLVSHLCSPGDIEWLRRCVEAFGLQPIILPDLAQSMGHLAQGDFSPLTQ
 GGTPLRQIEQMGSLSCSFAIGVSLHRASSLLAPRCRGEVIALPHLMTLERCDAFIHQLA
 KISGRAVPEWLERQRGQLQDAMIDCHMWLQGQRMAIAAEGDLLAAWCDFANSQGMQPGP
 LVAPTGHPSLRQLPVERVVPGDLEDLQTLCAHPADLLVANSHARDLAEQFALPLVRAG
 FPLFDKLGEFRRVRQGYSGMRDTL FELANLIRERHHHLAHYRSPLRQNPESLSTGGAY
 AAD"
 CDS complement (13986..14006)
 /codon_start=1
 /product="tobacco etch virus (TEV) protease recognition and
 cleavage site"
 /note="TEV site"
 /translation="ENLYFQS"
 CDS complement (14007..14027)
 /codon_start=1
 /product="tobacco etch virus (TEV) protease recognition and
 cleavage site"
 /note="TEV site"
 /translation="ENLYFQS"
 CDS complement (14028..15398)
 /codon_start=1

```

/note="nifE"
/translation="MKGNEILALLDEPACEHNHKQKSGCSAPKPGATAGGCAFDGAQIT
LLPIADVAHLVHGPIGCAGSSWDNRGSASSGPTLNRLGFTTDLNEQDVIMGRGERLRFH
AVRHIVTRYHPAAVFIYNTCVPAMEGDDLEAVCQAAQTATGVPVIAIDAAGFYGSKNLG
NRLAGDVMVKRVIQREPAWPPESTLFAPEQRHDIGLIGEFNIAGEFWHIQPLDELGI
RVLGSLSGDGRFAEIQTMHRAQANMLVCSRALINVARALEQRYGTPWFEGSFYGIRATS
DALRQLAALLGDDDLRQRTEALIAREEQAAELALQPWREQLRGRKALLYTGGVKSWSVV
SALQDLGMTVVATGTRKSTEEDKQRIRELMGEEAVMLEEGNARTLLDVVYRYQADLMIA
GGRNMYTAYKARLPFLDINQEREHAFAGYQGIVTLARQLCQTINSP IWPQTHSRAPWR"
terminator 15558..15584
/note="Terminator2-nifH"
CDS complement(15632..17206)
/codon_start=1
/note="nifA"
/translation="MIHKSDSDTTVRRFDLSQQFTAMQRISVVLRSRTEASKTLQEVLS
VLHNDAFMQHGMI CLYDSQQEILSIEALQQTEDQTLPGSTQIRYRPGELVGTVLAQQG
SLVLPRVADDQRFLDRLSLYDYDLPFIAVPLMGPHSRPIGVLAHAMARQEERLPACTR
FLETVANLIAQTIRLMILPTSAAQAPQQSPRIERPRACTPSRGGFLENMVGKSPAMRQI
MDIIRQVSRWDTTVLVRGESGTGKELIANAIIHNSPRAAAAFVKFNCAALPDNLLESEL
FGHEKGAFTGAVRQRKGRFELADGGTLFLDEIGESSASFQAKLLRILQEGEMERVGGDE
TLRVNVRIIAATNRHLEEEVRLGHFREDLYRRLNVMPIALPLRERQEDIAELAHFLVR
KIAHSQGRTLRISDGAIRLLMEYSWPGNVRELENCLERSAVLSEGLIDRDVILFNHRD
NPPKALASSGPAEDGWLDNSLDERQRLIAALEKAGWVQAKAARLLGMTPRQVAYRIQIM
DITMPRL"
CDS complement(17203..18690)
/codon_start=1
/note="nifL"
/translation="MTLNMMLDNAVPEAIIAGALTQQHPGLFFTMVEQASVAISLTDARA
NITYANPAFCRQTGYSLAQLLNQNPRLASSQTPREIYQEMWQTLLQRQPWRGQLINQA
RDGGLYLVDIDITPVLNPQGELEHYLAMQRDISVSYTLEQRLRNHMTLMEAVLNNIPAA
VVVVDEQDRVMDNLAYKTFCADCGGKELLVELQVSPRKMGPQAEQILPVVVRGAVRWL
SVTCWALPGVSEEASRYFVDSAPARTLMVIADCTQQRQQQEQGRDLRLKQMTAGKLLA
AIRESLDAALIQLNCPINMLAAARRLNGEGSGNVALDAAWREGEEAMARLQRCRPSLEL
ESNAVWPLQPFDDLYALYRTRFDDRARLQVDMASPHLVGFGQRTQLLACLSLWLDRTL
ALAAELPSVPLEIELYAEEDGWLSTLYLNDNVPLLQVRYAHSPPDALNSPGKGMELRLIQ
TLVAYHRGAIELASRPQGGTSLVLRFPPLFNTLTGGEQ"
terminator 19080..19111
/note="Terminator-nifJ"
CDS complement(19161..19421)
/codon_start=1
/note="nifW"
/translation="MMEWFYQIPGVDELRSAESFFQFFAVPYQPELLGRCSLPVLATFH
RKLRAEVPLQNRLEDNDRAPWLLARRLLAESYQQQFQESGT"
CDS complement(19422..19442)
/codon_start=1
/product="tobacco etch virus (TEV) protease recognition and
cleavage site"
/note="TEV site"

```

CDS /translation="ENLYFQS"
complement (19443..19463)
/codon_start=1
/product="tobacco etch virus (TEV) protease recognition and
cleavage site"
/note="TEV site"

CDS /translation="ENLYFQS"
complement (19464..20603)
/codon_start=1
/note="nifV"
/translation="MERVLINDTTLRDGEQSPGVAFRTSEKVAIAEALYAAGITAMEVG
TPAMGDEEIARIQLVRRQLPDATLMTWCRMNALEIRQSADLGIDWVDISIPASDKLRQY
KLREPLAVLLERLAMFIHLAHTLGLKVCIGCEDASRASGQTLRAIAEVAQQCAAARLY
ADTVGLLDPFTTAAQISALRDVWSGEIEMHAHNDLGMATANTLAAVSAGATSVNTTVLG
LGERAGNAALETVALGLERCLGVETGVHFSALPALCQRVAEAAQRAIDPQQPLVGELVF
THESGVHVAALLRDSesyQSIAPSLMGRSRYLVLGKHSGRQAVNGVFDQMGYHLNAAQI
NQLLPAIRRFAENWKRS PKDYELVAIYDEL CGESALRARG"

CDS complement (20604..20624)
/codon_start=1
/product="tobacco etch virus (TEV) protease recognition and
cleavage site"
/note="TEV site"
/translation="ENLYFQS"

CDS complement (20625..20645)
/codon_start=1
/product="tobacco etch virus (TEV) protease recognition and
cleavage site"
/note="TEV site"
/translation="ENLYFQS"

CDS complement (20646..24158)
/codon_start=1
/note="nifJ"
/translation="MSGKMTMDGNAAA AWISYAFTEVA AIYPITPSTPMAENVDEWAA
QGKKNLFGQPVR LMEMQSEAGAAGAVH GALQAGAL TTTYTASQGLLLMIPNMYKIAGEL
LPGVFHVSARALATNSLNI FGDHQDVM AVRQTGCAMLAENNVQQVMDLSAV AHLAAIKG
RIPFVNFFDGFRTSHEIQKIEVLEYEQLATLLDRPALDSFRRNALHPDHPVIRGTAQNP
DIYFQEREAGNRFYQALPDIVESYMTQISALTGREYHLFNYTGAADAERV I IAMGSVCD
TVQEVDLNAAGEKVG LLSVHLFRPFS LAHFFAQLPKTVQRIAVLDRTKEPGAQAEPL
CLDVKNAFYHHDDAPLIVGGRYALGGKDVLPNDIAAVFDNLNKPLPMDGFTLGI VDDVT
FTSLPPRQQTLAVSHDGITACKFWGMGSDGTVGANKSAIKIIGDKTPLYAQAYFSYDSK
KSGGITVSHLRFGDRPINSPYLIHRADFISCSQQSYVERYDLLDGLKPGGTFLNCSWS
DAELEQHLPVGFKRYLARENIFHYTLNAVDIARELGLGGRFNMLMQAAFFKLA AIIDPQ
TAADYLKQAVEKSYGSKGAAVIEMNQRAIELGMASLHQVTIPAHWATLDEPAAQASAMM
PDFIRDILQPMNRQCGDQLPVSAFVGMEDGTFPSGTAAWEKRGIALEVPVWQPEGCTQC
NQCAFICPHAAIRPALLNGEEHDAAPVGLLSKPAQGAKEYHYHLAISPLDCSGCGNCVD
ICPARGKALKMQSLDSQRQMAPVWDYALALTPKSNPFRKTTVKGSQFETPLLEFSGACA
GCGETPYARLITQLFGDRMLIANATGCSSIWGASAPSI PYTTNHRGHGPAWANS LFEDN
AEFGLGMM LGGQAVRQQIADDMTAAALALPVSDELSDAMRQLAKQDEGE TRERADRLS

```

ERLAAEKEGVPLLEQLWQNRDYFVRRSQWIFGGDGWAYDIGFGGLDHVLASGEDVNILV
FDTEVYSNTGGQSSKSTPVAAIKFAAQGKRTRKKDLGMMMSYGVVYAQVAMGADKD
QTLRAIAEAEAWPGPSLVIAYAACINHGLKAGMRCSQREAKRAVEAGYWHLWRYHPQRE
AEGKTPFMLDSEEPESFRDFLLGEVRYASLHKTTPHLADALFSRTEEDARARFAQYRR
LAGEE"
terminator 24551..24570
            /note="Terminator-nifB"
CDS        complement(24604..25266)
            /codon_start=1
            /note="nifY"
            /translation="MSDNDTLFWRMLALFQSLPDLQPAQIVDWLAQESGETLTPERLAT
            LTQPQLAASFPSATAVMSPARWSRVMASLQGALPAHLRIVRPAQRTPQLLAAFCSQDGL
            VINGHFGQGRLFFIYAFDEQGGWLYDLRRYPSAPHQEQANEVRARLIEDCQLLFCQEIG
            GPAAARLIRHRIHPMKAQPGTTIQAQCEAINTLLAGRLPPWLAKRLNRDNPLEERVF"
CDS        complement(25267..25287)
            /codon_start=1
            /product="tobacco etch virus (TEV) protease recognition and
            cleavage site"
            /note="TEV site"
            /translation="ENLYFQS"
CDS        complement(25288..25308)
            /codon_start=1
            /product="tobacco etch virus (TEV) protease recognition and
            cleavage site"
            /note="TEV site"
            /translation="ENLYFQS"
CDS        complement(25309..26106)
            /codon_start=1
            /note="nifM"
            /translation="MNPWQRFARQRLARSRWNRDPAALDPADTPAFEQAWQRQCHMEQT
            IVARVPEGDIPAALLENIAASLAIWLDEGDFAPPERAAIVRHHARLELAFADIARQAPQ
            PDLSTVQAWYLRHQTFMRPEQRLTRHLLLTVDNDREAVHQRILGLYRQINASRDAFAP
            LAQRHSHCPSALEEGRLGWISRGLLYPQLETALFSLAENALSLPIASELGHLLWCEAI
            RPAAPMEPQQALESARDYLWQQSQQRHQRWLEQMISRQPGLCG"
CDS        complement(26107..26127)
            /codon_start=1
            /product="tobacco etch virus (TEV) protease recognition and
            cleavage site"
            /note="TEV site"
            /translation="ENLYFQS"
CDS        complement(26128..26148)
            /codon_start=1
            /product="tobacco etch virus (TEV) protease recognition and
            cleavage site"
            /note="TEV site"
            /translation="ENLYFQS"
CDS        complement(26149..26676)
            /codon_start=1

```

/note="nifF"

/translation="MANIGIFFGTDGKTRKIAKMIHKQLGELADAPVNINRTTLDDFM
AYPVLLLTPTLGDGQLPGLGCESESWSEFISGLDDASLKGKTVLFLGLDQRGYPD
NFVSGMRPLFDALSARGAQMIGSWPNEGYESASSALEGDRFVGLVLDQDNQFDQTEAR
LASWLEEIKRTVL"

ORIGIN

1 tttcagtgca atttatctct tcaaatgtag cacctctagt cagccccata cgatataagt
61 tgtaattctc atgtttgaca gcttatcgat cgaataagca cttaaatgcg gtagtttatac
121 acagttaaat tgctaacgca gtcaggcacc gtgtatgaaa tctaacaatg cgctcatcgt
181 catcctcggc accgtcacc tggatgctgt aggcataggc ttggttatgc cggtagctgc
241 gggcctcttg cgggatatcg tccattccga cagcatcgcc agtcactatg gcgtgctgct
301 ggcgctatat gcgttgatgc aatttctatg cgcacccggt ctccggagcac tgtccgaccg
361 ctttggccgc cgcccagtcc tgctcgcttc gctacttga gccactatcg actacgcgat
421 catggcgacc acacccgtcc tgtggatect ctacgccgga cgcacgtggt ccggcateac
481 cggcgccaca ggtgcggttg ctggcgcta tatcgccgac atcaccgatg gggaagatcg
541 ggctcgccac ttcgggctca tgagcgcttg ttcggcgtg ggtatggtg caggccccgt
601 ggccggggga ctgttgggcg ccatctcctt gcatgcacca ttccttgcgg cggcggtgct
661 caacggcctc aacctactac tgggctgctt cctaagcag gagtcgcata agggagagcg
721 tcgaccgatg cccttgagag ccttcaacc agtcagctcc ttccgggtggg cgcggggcat
781 gactatcgtc gccgcactta tgactgtctt cttatcatg caactcgtag gacaggtgcc
841 ggcagcgctc tgggtcattt tcggcgagga ccgcttctgc tggagcgca cgatgatcgg
901 cctgtcgtt gcggtattcg gaatcttga cgcctcget caagccttcg tcaactgtcc
961 cgccaccaa cgtttcggcg agaagcaggc cattatcgcc ggcatggcgg ccgacgcgt
1021 gggctacgtc ttgctggcgt tcgcgacgcg aggetggatg gcttccccca ttatgattct
1081 tctcgttcc ggccgcatcg ggatgcccgc gttgcaggcc atgctgtcca ggcaggtaga
1141 tgacgaccat cagggacagc ttcaaggatc gctcgcggt cttaccagcc taacttcgat
1201 cactggaccg ctgatcgtca cggcgattta tgccgctcg gcgagcaca ggaacgggtt
1261 ggcatggatt gtaggcgccg ccctatacct tgtctgcctc cccgcgttgc gtcgcggtgc
1321 atggagccgg gccacctga cctgaatgga agccggcggc acctcgctaa cggattcacc
1381 actccaagaa ttggagccaa tcaattcttg cggagaactg tgaatgcga aaccaacct
1441 tggcagaaca tatccatcgc gtccgccatc tccagcagcc gcacgcggcg catctcgggc
1501 agcgttgggt cctggccacg ggtgcgatg atcgtgctc tgtcgttag gaccggcta
1561 ggctggcggg gttgccttac tggtagcag aatgaatcac cgatacgca gcgaactga
1621 agcgaactgt gctgcaaac gtctgcgacc tgagcaaca catgaatggt ctccggttc
1681 cgtgtttcgt aaagtctgga aacgcggaag tcccctacgt gctgctgaag ttgcccga
1741 cagagagtgg aaccaaccg tgataccacg atactatgac tgagagtcaa cgccatgagc
1801 ggcctcattt cttattctga gttacaacag tccgcaccgc tgtccgtag ctcttcgg
1861 tgggcgcggg gcatgactat cgctgccgca cttatgactg tcttcttat catgcaactc
1921 gtaggacagg tgccggcagc gcccaacagt ccccgccca cggggcctgc caccatacc
1981 acgccgaac aagcgcctg caccattatg ttccgatct gcatcgagg atgctgctgg
2041 ctaccctgtg gaacacctac atctgtatta acgaagcgt aaccgtttt atcaggctct
2101 gggaggcaga ataatgatc atatcgtcaa ttattacct caggggaga gcctgagca
2161 actggcctca ggcatctgag aagcacacgg tcacactgct tccggtagtc aataaccgg
2221 taaaccagca atagacataa gcggctattt aacgacctg cctgaaccg acgaccgggt
2281 cgaatttgct ttcgaatttc tgccattcat ccgcttatta tcaattatc agcgtagca
2341 ccaggcgttt aaggccacca ataactgcct taaaaaatt acgccccgc ctgcccacta
2401 tcgcaatact gttgtaattc attaagcatt ctgccgacat ggaagccatc acagacggca
2461 tgatgaacct gaatcgccag cggcatcagc acctgtcgc cttgcgtata atatttgc

2521 atggtgaaaa cgggggcgaa gaagttgtcc atattggcca cgtttaaatc aaaactgggtg
2581 aaactcacc agggattggc tgagacgaaa aacatattct caataaaccc tttagggaaa
2641 taggccaggt tttcacgta acacgccaca tcttgcgaat atatgtgtag aaactgccgg
2701 aaatcgtcgt ggtattcact ccagagcgat gaaaacgttt cagtttgctc atggaaaacg
2761 gtgtaacaag ggtgaacact atcccatatc accagctcac cgtctttcat tgccatacgg
2821 aattccggat gagcattcat caggcgggca agaattgtaa taaaggccgg ataaaacttg
2881 tgcttatttt tctttacggc ctttaaaaag gccgtaatat ccagctgaac ggtctggtta
2941 taggtacatt gagcaactga ctgaaatgcc tcaaaatggt ctttacgatg ccattgggat
3001 atatcaacgg tggatatatc agtgattttt ttctccattt tagcttctt agctcctgaa
3061 aatctcgata actcaaaaaa tacgcccggg agtgatctta tttcattatg gtgaaagtgtg
3121 gaacctctta cgtgccgatc aacgtctcat tttcgccaaa agttggccca gggcttcccg
3181 gtatcaacag ggacaccagg atttatttat tctgcgaagt gatcttccgt cacaggtatt
3241 tattcggcgc aaagtgcgtc gggatgatgct gccaaactac tgatttagtg tatgatgggtg
3301 tttttgaggt gctccagtggt cttctgtttc tatcagctgt cctcctgtt cagctactga
3361 cggggtgggtg cgtaacggca aaagcaccgc cggacatcag cgcaagcggg gtgtatactg
3421 gcttactatg ttggcactga tgagggtgtc agtgaagtgc ttcattgtggc aggagaaaaa
3481 aggctgcacc ggtgcgtcag cagaatatgt gatacaggat atattccgct tctcgcctca
3541 ctgactcgtc acgctcggtc gttcgcactgc ggcgagcggg aatggcttac gaacggggcg
3601 gagatttcct ggaagatgcc aggaagatac ttaacaggga agtgagaggg ccgcggcaaa
3661 gccgtttttc cataggctcc gccccctga caagcatcac gaaatctgac gctcaaatca
3721 gtggtggcga aaccgcacag gactataaaag ataccaggcg tttccccctg gcggtccct
3781 cgtgcgtctc cctgttctcg cctttcgggt taccgggtgc attccgctgt tatggccgcg
3841 tttgtctcat tccacgcctg aactcagtt ccgggtaggc agttcgtcc aagctggact
3901 gtatgcacga accccccgtt cagtcagacc gctgcgcctt atccggtaac tategtcttg
3961 agtccaacce ggaaagacat gcaaaagcac cactggcagc agccactggt aattgattta
4021 gaggagttag tcttgaagtc atgcgccggg taaggctaaa ctgaaaggac aagttttggt
4081 gactgcgctc ctccaagcca gttacctcgg ttcaaagagt tggtagctca gagaacctc
4141 gaaaaaccgc cctgcaaggc ggttttttcg ttttcagagc aagagattac gcgcagacca
4201 aaacgatctc aagaagatca tcttattaat cagataaaat atttctagag tttaaacgca
4261 ggagcaggag tggcatgggg atccggacc cgcgccctag ccgtgtcgtt tctgtgacaa
4321 agcccacaaa acatcgcgac actgtaggac gaaccttgtc aggactaata cacaaccatt
4381 tgaaaaatat taattttatt ctctggtatc gcaattgcta gttcgttatc gccaccgcgc
4441 ttccgcgggtg aaccgcgccc cggcgttttc cgtcaacatc cctggagctg acagcatgtg
4501 gaattactcc gagaaagtaa aagaccattt ttttaacccc cgcaatgcgc gcgtggtgga
4561 caacgccaac gcggtaggcg acgtcggttc gtttaagctgc ggcgacgccc tgcgctgat
4621 gctgcgcgtc gaccgcgaaa gcgaaatcat tgaggaggcg ggcttccaga ccttcggctg
4681 cggcagcgcc atcgcctcct cctccgcgt gacggagctg attatcgcc ataccctcgc
4741 cgaagccggg cagataacca atcagcagat tgccgattat ctcgacggac tgccgccgga
4801 gaaaatgcac tgctcgggtg tgggccagga ggccctgcgc gcggccatcg ccaactttcg
4861 cggcgaaagc cttgaagagg agcacgacga gggcaagctg atctgcaaat gcttcggcgt
4921 cgatgaaggg catattcgcc gcgcggtaca gaacaacggg ctgaccacc ttgccgaggt
4981 gatcaactac accaaagcgg gcggcggctg cactcttgc cacgaaaaaa tcgagctggc
5041 cctggcggag atcctcgcgc agcagccgca gacgacgcca gccgtggcca gcggcaaaga
5101 tccgactgg cagagcgtcg tcgataacat cgcagaactg cggccgcata ttcaggccga
5161 cggcggcgat atggcgttac tcagcgtcac caaccaccag gtgaccgtca gcctctccgg
5221 cagctgtagc ggctgcatga tgaccgatat gaccctggcc tggtgcgac aaaaactgat
5281 ggaacgtacc ggctgttata tggaaagtgg ggccggccgaa aatctttatt ttcagtcgga
5341 aaatctttat tttcagtcga aacaggttta tctcgataac aacgccacca cccgtctgga

5401 cccgatggtc ctggaagcga tgatgccctt tttgaccgat ttttacggca acccctcgtc
5461 gatacacgat tttggcattc cggcccaggc ggctctggaa cgcgcgcate agcaggctgc
5521 ggcgctgctg ggcgcggagt atcccagcga gatcatcttt acctcctgcg ccaccegaagc
5581 caccgccacc gccatcgcct cggcgatcgc cctgctgcct gagcgtcgcg aatcatcac
5641 cagcgtggtc gaacatccgg cgacgctggc ggctgcgag cacctggagc gccagggcta
5701 ccggattcat cgcatcgcgg tggatagcga gggggcgcct gacatggcgc agttccgcgc
5761 ggcgcttagc ccgcgcgtcg cgttggtcag cgtgatgtgg gcgaataacg aaaccggggt
5821 gcttttcccg atcggcgaaa tggcggagct ggcccatgaa caaggggcgc tgtttcactg
5881 cgatgcggtg cagggtggcgc gaaaaatacc gatcgccgtg ggccagacc gcategatat
5941 gctctcctgc tcggcgcata agttccacgg gccaaaaggc gtaggctgtc tttatctgcg
6001 gcggggaacg cgctttcgcg cgctgctgcg cggcggtcac caggagtacg gtcggcgagc
6061 cgggacagaa aatatctgcg gaatcgtcgg catgggcgcg gcctgcgagc tggcgaatat
6121 tcacttgccg ggaatgacgc atatcggcca attgcgcaac aggctggagc atgcctgct
6181 ggccagcgtg ccgtcggcga tggatagcga cggcggccag ccgcgggtgc ccggcacggt
6241 gaatctggcc tttgagttaa ttgaaggcga agccattctg ctgctgttaa accaggccgg
6301 gatcgccgcc tccagcggca gcgcctgcac ctgaggctcg ctggaacct cccacgtgat
6361 gcgggcgatg aatatcccct acaccgccgc ccacggcacc atccgctttt ctctctcgcg
6421 ctacacccgg gagaaagaga tcgattacgt cgtcgccacg ctgccgccga ttategaccg
6481 gctgcgcgcg ctgtcgcctt actggcagaa cggcaagccg cccccggcgg acgccgtatt
6541 cacgccggtt tacggctaata cccgtcagcc gggactgtgc gggtagcctc ggcggctacc
6601 cgtaaacgcc tacagcacgg tggtttaaac actagatacg tacagcgcgc caagagattg
6661 cgtggaataa gacacagggg gcgacaagct gttgaacagg cgacaaagcg ccaccatggc
6721 cccggcaggc gcaattgttc tgtttcccac atttggcgcg cttattgtgc cgttttgttt
6781 tacgtcctgc gcggcgacaa ataactaact tcataaaaaat cataagaata cataaacagg
6841 cacggctggt atgttccctg cacttctctg ctggcaaaca ctcaacaaca ggagaagtca
6901 ccatgaccat cgtcgaatgc gctatttacg gtaaaggcgg tatcggtaaa tccaccacca
6961 cgcagaacct cgtcgcgcg ctggcggaga tgggtaagaa agtgatgatc gtcggctgcg
7021 atccgaaggc ggactccacc cgtctgattc tgcacgcaa agcacagaac accattatgg
7081 agatggccgc ggaagtcggc tcggctgagg acctcgaact cgaagacgtg ctgcaattg
7141 gctacggcga tgtgcgctgc gcggaatccg gcggcccgga gccaggcgtc ggctgcgcgg
7201 gacgcggcgt gatcacggcg atcaactttc ttgaagaaga aggcgcctac gaggacgatc
7261 tcgatttcgt gttctatgac gtgctcggcg acgtggtctg cggcggcttc gccatgccga
7321 tccgcgaaaa caaagcccag gagatctaca tcgtctgctc cggcgaaatg atggcgatgt
7381 acgcggccaa caatatctcc aaagggatcg ttaaatacgc caaatccggc aaggtgcgcc
7441 tcggcggcct gatctgtaac tcacgtcaga ccgaccgtga agacgaactg attattgccc
7501 tggcggaaaa gctcgggtacc cagatgatec actttgtgcc ccgcgacaac atcgtgcagc
7561 gcgcggagat ccgccgcgat acggttatcg agtacgacc cgctgtaaa caggccaacg
7621 aataccgcac cctggcgcag aagatcgtca acaacacat gaaagtgggt ccgacgccct
7681 gcaccatgga tgagctggaa tcgctgctga tggagtccgg catcatggaa gaggaagaca
7741 ccagcatcat tggcaaaacc gccgccgaag aaaacgcggc cgagaacctt tatttccaat
7801 cggagaatct ttacttccaa tcaatgatga ccaacgcaac gggcgaactg aatctggcgc
7861 tgatccagga agtcctggag gtgttcccgg aaaccgcgcg aaaagagcgc agaaagcaca
7921 tgatggctcag cgatccggaa atggagagcg tcggcaagtg cattatctct aaccgcaaat
7981 cacaaccggc cgtaatgacc gtacgcggct gcgcctacgc cggttccaaa ggggtggtat
8041 ttgggccgat taaggatatg gcccatattt cgcacggacc ggtcggctgc ggccagtatt
8101 cccgcgccgg acgacgcaac tactacaccg gactcagcgg cgtcgategc ttggcagcgc
8161 tgaacttcac ctctgatatt caggagcgcg acatcgtctt cggcggcgat aaaaagctca
8221 gcaagctgat tgaagagatg gagttgctgt tcccgtcac caaagggatc accattcagt

8281 cggaatgccc ggtggggctg atcggatgatg atatcagcgc ggtggccaac gccagcagca
8341 aggcgctgga taaaccggtg atccccgtac gctgcgaagg ctttcgcggc gtgtcgcagt
8401 ctctggggca ccataatgcc aacgacgtgg tgcgcgactg gatcctgaac aatcgcgaag
8461 gacagccggt tgaaccacc ccttacgatg tggcgcactc cggcgactac aacatcggcg
8521 gcgacgcctg ggccctcgcg attctgctgg aagagatggg gctacgggta gtcgcgcagt
8581 ggtccggcga cggcacgctg gtggagatgg agaatacccc attcgtcaag ctgaacctgg
8641 ttcactgcta ccgttcgatg aactatatcg cccgccatat ggaggagaaa catcagattc
8701 cgtggatgga gtacaacttc ttcgggccga ccaaaatcgc cgaatcgcct cgcaaaatcg
8761 ccgaccagtt cgacgatacc attcgcgcga acgccgaagc ggtgatcgcc cggtatgagg
8821 ggcagatggc ggcgattatc gccaaatata gcccgccct ggaggggct aaggtgctgc
8881 tctatatggg cggcctgcgg ccgcgccacg ttattggcgc ctatgaggat ctcgggatgg
8941 agatcatcgc cgcggctac gagtttgcg ataacgatga ttacaccgc accctgccgg
9001 atctgaaaga gggcacgctg ctgttcgatg acgccagcag ctacgagctg gaagcgttcg
9061 tcaaggcgct gaagcccgc cttatcggct ccggcatcaa ggaaaaatat atcttcaga
9121 aaatggcgct gccgttcgcg cagatgcact cgtgggacta ttccggcccg taccacggct
9181 acgatggttt cgccattttc gcccgcgata tggatatgac cctgaacaac ccggcgtgga
9241 acgaactgac cgctccgtgg ctgaagtctg cggaaaacct ctacttcag tcggagaatc
9301 tttactttca gtcaatgagc caaacgattg ataaaattaa tagctgttat ccgctattcg
9361 aacaggatga ataccaggag ctgttccgca ataagcggca gctggaagag gcgcacgatg
9421 cgcagcgcgt gcaggaggtc tttgcctgga ccaccaccgc cgagtatgaa gcgctgaatt
9481 tccagcgcga ggcgctgacc gttgaccgg cgaaaacctg ccagccgctt ggcgcggtgc
9541 tttgctcgtt gggatttgcc aacacctgc cgtatgtgca cggctctcag gggctcgtgg
9601 cctactttcg cacctatttt aaccgccatt tcaaagagcc gatcgctgc gtctccgact
9661 cgatgaccga agacgcggcg gtcttcggcg gcaacaacaa tatgaacctg ggctgcaga
9721 acgccagcgc gctgtacaaa ccggagatca ttgcggtgtc caccacctgc atggcggaag
9781 ttatcggcga tgacctgcag gcgtttatcg ccaacgctaa aaaagatggc ttcgtcgaca
9841 gcagcatcgc cgtgccccac gcccatacgc ccagctttat cggcagccac gtcaccggct
9901 gggataacat gtttgaaggc ttcgcaaaa cttcactgc ggactaccag gggcagccgg
9961 gcaaatgcc gaagctcaat ctggtgaccg gctttgaaac ctatctcggc aacttcgcg
10021 tattaagcg gatgatgga cagatggcgg tgccgtgcag cctgctctcc gatccgctgg
10081 aagttctcga cacgcccgc gacggccact atcggatgta ttccggcggc accacgcagc
10141 aggagatgaa agaggcccct gacgccatcg atacgctgct cctgcagccg tggcagctgc
10201 tgaagagcaa aaaagtgggt caggagatgt ggaaccagcc cgcaccgag gtcgccattc
10261 cgctggggct ggccgccacc gatgaactgc tgatgaccgt cagccagctt agcggcaagc
10321 cgattgccga cgcctcacc cttagcgcg gccggctggt tgacatgatg ctgactccc
10381 acacctggct gcacggcaag aagtttggc tgtacggcga tccggacttc gtgatggcc
10441 tcaccgctt cctgctggag ctgggctgcg agccaacggt gatcctgagc cataacgcca
10501 acaaacgctg gcaaaaagcg atgaacaaaa tgctcgcgac ctgcgctac gggcgcgata
10561 gcgaagtgtt tatcaactgc gatttgtggc acttccgttc gctgatgttc acccgtcagc
10621 cggactttat gatcggcaac tctacggca agtttatcca gcgcgatacc ctggcgaagg
10681 gtaaagcctt tgaagtggc cttatccgc tcggctttcc gctgttcgac cgcaccatc
10741 tgcaccgcca gacaacctgg ggttatgaag gggcgcgata cattgtgacg acgctggtga
10801 acgccgtgct ggagaaactg gatagcgata ccagccagct gggcaaac gattacagct
10861 tcgatctcgt ccgttaatcc ctgtttgtg cttgttgcg gctgacccc cgggcttttt
10921 ttcgcgtatg gacgctcttc cccacgttac gctcagggga atattccgtt cacggttgtt
10981 ccgggcttct tgatgcgcct aacccccctg ctgccagcct ttcataca aatagccatc
11041 ccagcgcgat aggtcatata cgtaactaga actagaaaagc ttggcaacc gcgtccctga
11101 attgtcgcaa ttccgccgcg ctggcgacaa tgtcctgaat ctacataag gcttcaggcg

11161 acccccttat gcggaacgc ggcaggccag ctccccgac gctgttgcca ccacgcgggc
11221 agcacttcgg agaccggccg ccacgcgccg tcaacgcagg gttcaatgcc ttectgttec
11281 agctgttgcc acggcgtatg gccgatacgc acgcagaata cggctttaac gtccgccage
11341 agttcgagga tcgccgcaaa cccggctgcg ttatcctgcg gctcgcagtc atcgcgcccc
11401 tgacaatatt tgggcgtaaa gcgctcgttg accagcacca taccggcggc cgagaggctg
11461 taaatgtaga accggtcggc gtgaccaaag tgacaatcaa tgacgtcccc gcggcttgac
11521 gccacggcga ccaggcaggc gtcacggct tcagattcgc cgcgggtcgc aatgctggcg
11581 tgcagctgcg cgcgctggtg gaggatcggc agccaggccg ggagactctc tggcgccgga
11641 agctgggtaa actgctggct gcggctcttcg ccgagcatcc caatggcgtc ggcgcgacac
11701 tggtagcagt gggatcactg cggcatcact tcgtcgcacc ggctgcgggt ggcggcgagc
11761 gtctcggcgt cggctccgg ctggccgttg aggccaaaca cctgcccgtg ctccggcctg
11821 gcgatcagcg gcataatatt atggataaac gcgccgtgg cccgcagcgc gcggctcacg
11881 tcggccatgc cgctatcgtt gataccgggg atcagcaccg aatttatctt caccagcacg
11941 cctttggcgg tcagcctgcg tacgccctca agctgacggg caatcaggat ctctcccgt
12001 tcgcgcccgc tgtagcgttc gccgtccagc catagccagg cgtagatttg cgcggcaate
12061 tccgcgtcga ggggtgtaat ggtgaccgtg acgtggtaa cgcggacatc cagcaggcgg
12121 tccaccgct caggcagcat cagtccgttg gtcgacaggc ataatttcag gtccggcagc
12181 tgttcgcgga tcagctccag ggtgcgaaag gtgcgggga tattggcgag cggatcgccg
12241 ggcccggcga tgcccaccac cgaaaagtgc gggatcgct cgcgcacctg acgactttc
12301 acgaccgct gttcaggcgt cagcagcgtt gacgataccc cggggcggga ctctgtgctg
12361 caatcgaatt tgcgattaca gtagttgcac tgcaaatgc agggggcgc gacgggcaga
12421 tgcacccgc cgaaacggtg atgcccattg cgagagtagc aggggtgcgc ggcagctta
12481 tcggccaca gcggcgtcaa tgcgctgtca tccgccggg gcgaggctt gccgccagaa
12541 aaagaggagc aggaagtaga acctccacct ccggaacct cacctccaga accacctca
12601 ccgtccgctg cgtacgcgcc acccgtggag agtgacgatt cggggttctg gcgcagcggc
12661 gatcggtagt gggcgagggt gtggtgacgc tcgcgtatca gtttgccag ctcaaacagc
12721 gtatcgcgca tcccgtata cccctgtcgc acccggcgga attcggcag cttgtcaaag
12781 agcggaaaac ccgcgcgcac cagcggcagc gcaaaactgt ccgccaggtc gcggcgctgc
12841 gagttcgcca ccagcaggtc ggcgggatgc gcgcacagca gggtttgag atctccaga
12901 tccccggca ccaccgctt caccggcagc tggcgcagc tgggatgacc ggtaggggc
12961 accagcgggc cgggctgcat cccctggctg ttggcgaaat cacaccagc cgcagcaaa
13021 tcgcttccg ccgctatcgc catgcgtg cctggagcc acatattgca gtcgatcctc
13081 gcacctgta gctggccgcg ctggcgttcc agccactcgg gaacggcgcg tccggaatt
13141 ttcgccagtt gatgaataaa ggcgtcgcag cgttcgaggg tcatcagggt cggcaggcgc
13201 ataacctgc cgcggcagc cggggccagc agcagatgagg cgcgatgaag ggagacgcca
13261 atggcgaagc tgcacaggct ttgccccatc tgctctatct ggcgcagcgg cgtcccgcc
13321 tgggtcagcg gcgagaaatc gccctgcgc aggtggcct ccatcgattg cgcaggtcc
13381 ggcaggatta tcggctgcag accaaaggct tcgacgcatc ggcgcagcca ctgatatcg
13441 cccggcgaac agagatggct gaccagcaga ttgaccggc gattgcgtg agccggcgc
13501 ggcgccggcg gcacccactg ctcaatgacg ctcttaaca ccgcgtgaa gccgttctc
13561 atggagccat aaaaatccgg cgtgttaacc gtcaatatcg ccaccctt atgccgggga
13621 tactcttcgc gaaactggcg aaccacgcgg gaaatatcgc tgccctgggc ctccgacagc
13681 ccggtgctga gcagtacgat agcctgcgga ttgttcgct ggcagagggt atccaggcgc
13741 gtaaaaatat tgccgtccgc gcccataatc gtcgacgtgg ggtccatcgc cgtcgactgc
13801 aggggaaccg ggtcgtggaa atgttgaata aaaaagactt tggcgaagc gctgcaccc
13861 tgcgcgccgt ggaccagagg gatgctgtgt tcgatcccga ggctggcgag gattgcgccg
13921 agcggctggc cggttttgat ggggctgacc gccagcggct tatcggctgc gaaaatgtct
13981 gccatcgatt ggaaatacag gttctccgac tgaataata agttctcgc ccacggggc

14041 cgagaatggg tttgcggcca gatggggctg ttgatggtct gacacagctg gcgggcgagg
14101 gtgacgatcc cctgatagcc agcgaaggcg tgttcgcgct cctgattgat atcgagaaac
14161 ggcagcctgg ctttatagcc ggtgtacata ttgcgtccgc cggcaatcat caggteggcc
14221 tgatagcgat agaccacatc cagcagcgtg cgggcgttgc cctcttccag cattaccgcc
14281 tcttcgccc tcagctcgcg gatccgctgt ttatcctctt cgggtgattt acgcgtgccg
14341 gttgccacca cggatcatgcc caaatcctgc agcgcggata ccaccgacca ggatttcacc
14401 ccgccggtat agagcagcgc tttgcggccg cgcagctggt cgcgccacgg ctgtagcgcc
14461 agttccgccg cctgttccctc ccgcgcaatc agcgttccgg tgcgctggcg aaggtcgtcg
14521 tcgccagca gcgccgccag ctggcgcagg gcgtcagagg tggcgcggat cccgtaaaag
14581 ctgccttcga accacggcgt gccgtagcgc tgctccaggg ctctggcgac gttaattaac
14641 gcccgcgagc agaccagcat attggcctgc gcccggtgca tggctctggat ctggcggaag
14701 cggccatcac cggagaggct gccgagcacg cggatcccca gttcgtcgag cagcggctga
14761 atatgccaga actcgcgggc aatattgaat tcgccaatca ggccaatata gtgacgctgc
14821 tccggggcaa agagcgtgct ctccggccag ggggcgggct cgcgctggcc gatgaccctg
14881 ttgacatga cgtcgcggc cagccggtta ccgagatttt tactgccgta gaaaccggcg
14941 gcgtcaatag cgataaccgg tacgccggtg gcggtctgcg cggcctggca taccgcttc
15001 aggtcatcgc cctccatggc cggtagcag gtgtttaga taaagaccgc cgcggatga
15061 tagcgggtga cgatatggcg cacggcgtga aacaagcggc gttcgcgcg gcccataatc
15121 acgtcctggt cgttgagatc ggtggtgaac ccgagccgat taagggtggg gccggagctg
15181 gcgctgccgc gtttatcca tgagcttccg gcgcagccga tggggccgtg gaccagatgc
15241 gccacgtcgg cgatgggcag cagggttatc tgcgcgccgt cgaacgcgca gccgccggcg
15301 gtggcgccgg gtttgggcgc gctgcagccg gatTTTTgtt tatggttggt ttcacaggcc
15361 ggttcatcca gcagcggcag aatttcattt ccctcatag tgacctggt gacggtgggg
15421 gatgaggggg aagatgcaat tcgcgctcca gaagcggagc cttgattttt aaagaaaaga
15481 ggttggttggc tttgcgacaa agccaacaat cggacaaggg atggcggcat gtgatgcttt
15541 atgacctatc gcgctgggat ggctatttgt tgatgaaagg ctggcagcga gggggttagg
15601 cgcatacaaga agcaagctta ctagtcttaa gtcacagtcg cggcatggtg atatccataa
15661 tctgaatgcg ataccgccacc tggcgcgggg tcatgccgag cagccgcgcc gctttggcct
15721 gcaccagcc cgctttttcc agggcggcga tcagccgctg gcgctcgtcg aggctgttat
15781 cgagccagcc gtctccgcc gggccgctgc tggcagcgc tttcggcggg ttatcgcgat
15841 ggttgaacag aatcacgtcc cggctatca ggccgcttc cgacagcacc gccgaacgtt
15901 cgagacagtt ttccagttcg cgcagtttc ccggccagct gtactccatc agcaggcga
15961 tcgccccatc gctgatgcgc agcgttcgcc cctggctgtg ggcgattttt cgcaccagaa
16021 agtgcgccag ctcggcgata tctcctggc gctcgcgcag cggcggcagc gcgataggca
16081 ttacgttcag gcggtagtat agatcctcgc ggaaatgacc cagccgcacc tctcttcca
16141 gatggcggtt ggtcggcg ataatagcgca cgttgaccgg cagggtttcg tcgcccgga
16201 cgcgctccat ctccccctct tgcagaatac gcagtagctt agcctgaaac gaggcgtgc
16261 tttcggcgat ctcatcgagg aataagggtgc cgccgtccgc cagctcaaag cggcctttcc
16321 gctggcgcac cgcgccgta aacgcgctt tctcatgacc aacagctcg ctctccagca
16381 ggttgtccgg cagcggcgc cagttaaatt tgacgaacgc cgcggcggcg cgcggagaat
16441 tatggtggat ggcgttggcg atgagctctt tcccgtgccc gctctcggc cgtaccagca
16501 ccgtggtatc ccagcgggaa acctgacgaa taatatccat aatctgccgc atcgccgggc
16561 tttaccgac catattttcc aggccgaaac cgcgcgaagg ggtacaggcg cgcgggcgt
16621 ctattctggg gctctgctgc ggcgctgcg cggcggaggt tggcaggatc atcaggcga
16681 tcgtctgggc gatcagattg gcgacggtt cgagaaagcg cgtgcaggcg ggcagccgt
16741 ctctctgacg cgccatcgcg tgcggcccca gtacgccgat gggccgggag tggggccca
16801 tcagcggaac ggcgataaac ggcaggtcat agtcgtacag gctcagacga tcgagaaaac
16861 gctggtcgtc ggcgacgcgc ggcagcacca gcgactggcc ctgcgccagc acggtaccga

16921 ctaatccttc ccccggccgg tagcgaattt gcgtactgcc gggcagcgtc tgatcttccg
16981 tttgctgcag cgcttcgatg ctcaggatct cctgctggct gtcgtacagg caaatcatcc
17041 cgtgctgcat aaaggcatcg ttatgtagca cgctcagaac ctctgcagg gttttgctcg
17101 ctccgggtggc gcgactcagg accacgctta tccgctgcat ggcggtaaac tgctgggaga
17161 gatcgaaacg tctgacggtg gtgtccgaat cggatttatg gatcattgct cacctccggt
17221 cagggtatta aagagcggga aacgcagaac caggctgggt cctccctgcg gtcgcgaagc
17281 cagttcaatc gcgcccggt gtaggcgac cagcgtttgg atcagccgca gctccatccc
17341 tttgccggga gagtttaggg catcggggga gtgggcgtag cgcacctgca gcagcgggac
17401 attgtcgttg agatacaaag agagccagcc ctctctctct tcggcgtaaa gctcgatctc
17461 cagcggtagc gagggcagct cggcggcgag ggccagcgtc cggtcgagcc ataaactcaa
17521 gcaggccagc agctgggtac gctggccgaa gccgaccaga tgcggcgatg ccatgtccac
17581 ctgcagccgc gcgcgatcgt caaagcgggt gcggtagagg gcgtacaggt cgtcaaaaaa
17641 gggctgaagc ggccagacgg cattgctttc cagctcaaga gaaggcggc agcgtgcag
17701 gcgcgccatg gcctcttcac ctccgcgcca cgcgcgtcc agcgcacgt tgccgctgcc
17761 ttcgccgttc agccggcggg ccgccgccag catattgatt gggcaattaa gctgaatcag
17821 cgccgcgtcc agcgactcgc gaatcgcggc cagcagctta ccggcggtea tttgctgttt
17881 cagacggtcg agccggccct gctcctgctg ctggcgtgc tgggtacagt cggcgatcac
17941 catcagcgtg cgcgccgggg cgctgtcgac gaagtagcgg ctggcttctt cactcacgcc
18001 gggcagcgc cagcaggtta ccgacagcca gcggaccgc ccgcaacca ccaccggcag
18061 gatttgctcc gcgccgggccc ccattttgcg cggggaaacc tggagctcga ccagcagctc
18121 tttcccggcg cagtccgcgc agaacgtttt gtaggcgaga ttatccatca ccaccgatc
18181 ctgctcatcg accacgacca cggcggcggg gatgttattg agcaccgctt ccattagcgt
18241 catatgattg cgcagccgct gttccagggt atagctgac ctgatatccc gctgcatcgc
18301 cagataatgc tccagctcgc cctgcggatt cagcaccggc gtgatatcga tatctaccag
18361 atacaggccg ccgtcgcggg cctgattaat tagctgaccg cgcacggct ggcgtggag
18421 cagggtttgc cacatctcct ggtagatctc gcgcggcgtc tggctgctgg ccagcaggcg
18481 cgggttttga ttgagcaatt gcgccagcga gtatccagtc tggcggcaaa acgccgggtt
18541 ggcgtaggta atattcggcc gggcatcggg gagggaaatc gctaccgatg cctgttcgac
18601 cattgtaaaa aacagccccg gatgttgttg agtcagcgc caaggcaatc cctcgggtac
18661 ggcgttatcg agcatcatat tcagggtcat cgcaaaactc tgtgtcgcag ttatgacaat
18721 ttgtccgaca ccgtgacagg gacgcccggg tgttttccga acggtgataa ccatgcaaac
18781 cgtgcgccct tatcggcgtg atgtgcagaa acattgatcc ggtaggggaa aagcaggcgc
18841 tttttgcatt ttaagcggcc cgccagcccc gccgcgcgc gcgggagagg gctgaagcga
18901 tcgttttggg gcagggatgg gccgcggtag tgcaaaagca cctggcacag ccttcgcaat
18961 acccctgcga gaacgcgtat tttcaacccc cttttccgac aggagcgcga cacttaagtc
19021 tagtgctagc tctagtagta ctcagcgcgc acaggggcgg gagaacttga tttcgtcgcg
19081 ccgggaagat ttagacgcgc gctttcccgg cctgctggat acgctgctta aggtcatgca
19141 gcaggagaac taaaggcccc ttatgttccg ctctcctgaa actgttgctg atagctctcc
19201 gcgagcagtc ttcgcgccag cagccagggc gcgcggctgt tatcctcgag ccggttttgc
19261 agcggcacct ccgcgcggag tttgcgatga aacgttgcca gcaccggcag gctgcagcgg
19321 ccaagcagct cgggctgata ggggacggcg aaaaactgaa aaaaagattc ggcggagcga
19381 agttcgtcca cgccgggaat ttgataaaac cactccatca ttgattggaa gtacaagttc
19441 tccgactgaa aatacagggt tccccctc gcccgagag cggattacc gcacagctcg
19501 tcgtagatag ccaccagctc gtaatctttc gggctgcgct tccagttctc ggcgaagcgg
19561 cggatggcgg gcagcagctg gttaatctgc gcggcgttga ggtgatagcc catctgtgca
19621 aaaacgccgt tgaccgcctg acgcccggag tgtttgccc gcaccagccg gtagctgcgg
19681 cccatcaggg aagggcgat ggactgtag ctctcgtgt cgcgcagcag cgcgccacg
19741 tggacacctg actcatgggt aaacaccagc tcgccacca gcggctgctg cgggtcgatg

19801 gcgcgctgcg cggcttccgc gacctctga cagagcgcgg gcagcgcga aaaatgcacg
19861 ccggtctcca cgccaggca gcgttcaagg cccagcgcga cggtttccag cgccgcgttg
19921 cccgcccgtc cgccgagacc gaggaccgtc gtattcacgc tggtagcccc ggcgcttacc
19981 gccgccagcg tattggcggg cgccataccc agatcgttat gggcatgcat ttcgatttcg
20041 ccggaccaga cgtcgcgcag ggccgagatt tgcgccgcgg tggtaaaagg gtcgagcagg
20101 ccgaccgtat cggcatagcg caggcgggcg gcggcgcatt gctgcgcgac ctgggcgata
20161 gcgcgcaggg tctggccgct ggcccgcgag gcgtcctcgc agccgatgca taccttcagg
20221 ccgaggggat gcgcaagatg gataaacatc gccagccgct ccagcagcac cgccagcggc
20281 tcgcgcagtt tgtactgccg cagcttatcc gaagccggaa tcgagataac caccagtcg
20341 atgcccagat cggcgcctcg gcggatctcc agcgcgttca tccgacacca ggtcatcagg
20401 gtcgcgctcg gcagctggcg acgcaccagc tggatccgcg cgatctctc gtcgccatc
20461 gccggggtgc cgacctccat cgccgttatt cctgcggcgt aaagcgcctc ggcaatggcg
20521 accttttcgc tggtagcga ggcgacgcc gggctctgct cgccgtcgcg cagggtggta
20581 tcgttaatca gcacgcgttc catcgattgg aaatacaggt tctccgactg aaaatataag
20641 ttctctctct cgccggccag ccgccgatac tgggcaaagc gggcccgcgc gtctctctcg
20701 gttcggctaa agagcgcac cgccagatgc ggcgtcgttt tgtgcagcga ggcgtagcgc
20761 acttcgcaa gcaaaaagtc gcggaagctc tctccggct cttcggaatc gagcataaac
20821 ggcgtcttac cttccgcttc ccgctgcgga tgatagcgc acaggtgcca gtatcccgcc
20881 tcaaccgcc gtttcgctc gcgctggctg cagcgcatac cggctttcag cccgtggtta
20941 atgcaggcgg cgtaggcaat caccagcgc ggtcccggcc aggttcggc ctgcccgatc
21001 gcccgtaggg tctgatcttt atcagcgcgc atcgcgacct gggccacgta cacattgccg
21061 tagctcatcg ccatcatgcc gagatcttt ttccgctgc gtttgccctg cgcggcaaac
21121 ttcgcgatgg ccgccaccgg ggtcgattta gacgactggc cgccggtatt ggagtaaacc
21181 tcggtgtcaa acaccagaat attgacgtct tccccgctc ccagcacgtg atcgagaccg
21241 ccgaagccga tctcgtaggc ccagccgctc ccgccgaaa tccactgca acgacgaaca
21301 aaatagtcgc ggttctgcca cagctgctcc aacagcggca cgccctctt ttccgccgc
21361 agccgttcgc tgagccggtc cgcgcgctc cgggtgccct cgccttcac ctgcttcgcc
21421 agccactggc gcattgcgtc gctaagtctc tcgctgacc gtagcgcag cgcgccggtc
21481 atatcatcgg cgatttgtt acgcaccgcc tggccgccga gcacatgcc gaggccaaac
21541 tccgattat cctcaaacag cgagttcgc catgccgggc catggccgc gtggttggtg
21601 gtatagggaa tcgacggcgc gctggctccc cagatagaag agcagccggt ggcgttagcg
21661 atcagcatcc ggtgcgcaaa cagctgggtt atcaggcggg cataaggcgt ttcaccgat
21721 cccgcgcagg cgccggaaaa ctccagcagc ggggtttcaa actggctgcc tttgaccgtc
21781 gtcttacgaa acggattgct ctccggcgtc agcgcagcg catagtcca gaccggcgc
21841 atctgacgct ggctatcgag agactgcatt ttaacgcct tgccgcgcgc gggacagata
21901 tccacgcagt tgccgcagcc ggaacaatcc agcggcgaga tagccagatg gtagtgatac
21961 tcttcgctc cctgcgcggg tttgctcagc agcccaaccg gcgcggcgtc atgtcttcg
22021 ccggttagca gcgccgggcg gatcgccgca tgcgggcaga taaaggcga ctggttacac
22081 tgcgtgcagc cctccggctg ccagaccggc acttcagcg cgatcccgcg tttctcccac
22141 gcggcgggtc ccgaaggaaa ggtcccgtc tccataccga cgaacgcgt caccggcagc
22201 tggtcgccgc actggcgggt catcggctgc agaatacgc ggtatgaaac cggcatcatg
22261 gctgatgctt gcgcccgggg ttcattcagc gtcgccagc gcgccggaat cgtcacctga
22321 tgcagcgagg ccatgcccag ctcgatcgc cgctggttca tctcaatcac cgcccccct
22381 ttgctgccgt agctttttc aaccgcctgc ttgaggtaac ccgcccggt ctgcccgtcg
22441 ataatgcccg ccagcttaaa gaacgccgc tgcatcagca tattaagcg cccgccagc
22501 ccgagctcgc gggcgatata cacggcgttc agggataaaa aatggatatt tcccgcgc
22561 agatagcgtt taaagccgac cggcagatgc tgctccagct ccgcatcga ccagctgcag
22621 ttgagtaaaa aggtcccgcc cggctttaat ccgtccagca gatcgtagc ctcaacgtag

22681 gactgctgcg aacaggagat aaaatcggcc cgatggatca ggtagggcga attgatcggc
22741 cggtcgccga agcgtaaatg tgaacggta atgccgccgg attttttcga gtcataagaa
22801 aagtaggcct gcgcgtagag cggcgtttta tcgccgataa ttttgatcgc gcttttattg
22861 gccccgacgg tgccgtccga gcccatgccc caaaatttac aggcggtgat gccgtcatgc
22921 gagaccgcca gcgtctgctg gcgcggcggt aacgaagtaa aggttacatc atcgacaate
22981 ccgagggtaa acccgtccat cggcagcggg ttattgaggt tatcaaagac ggccgcgata
23041 tcgttgggca gaacatcctt cccgccaaagc gcatagcggc cgccgacgat tagcggcgca
23101 tcgtcgtggg ggtagaaggc gtttttcaca tccaggcaca gcggttcage ctgagcggcg
23161 ggctctttgg tacggcaag gacggcaatc cgctgcacgg ttttcggcag ctgggcgaag
23221 aagtgggcca gcgaaaaagg gcgaaacaga tgcacgctga gcagcccgac cttctctccc
23281 gccgcgttca gcgtatccac cacttctcga acggtatcgc agaccgatcc cattgcgata
23341 atcaccggtt cggcatccgc cgcgccggta tagttaaaca gatgatactc ccggccgggtg
23401 agcgcgctga tttgcgtcat atagctttcg acaatgctgg gcagcgcctg ataaaaacgg
23461 ttgcccgcct cccgctcctg gaagtagata tccgggttct gcgccgttcc gcggatgacc
23521 ggatgatccg gatgcagcgc gttacggcgg aagctgtcga gcgcggggccg gtccagcagc
23581 gtcgccagct gctcatattc caacacctcg attttttgaa tttcgtcga ggtgcgaaaa
23641 ccgtcgaaga agttaacaaa cgggatgcgt cccttaatcg ccgccagatg ccccaccgcc
23701 gacaaatcca tcacctgctg cacgttggtc tccgccagca tcgcgcagcc ggtctggcgg
23761 accgccatca catcctgggtg atcgccaaaa atattcagcg aattggctgc cagcggccgg
23821 gcgctgacgt gaaagacgcc cggcagcagt tcaccggcga tttgtacat gttggggatc
23881 atcagcagca gcccctggga ggccgtatag gtgggtggtg gcgccccggc ctgcagcgcg
23941 ccgtggaccg cgcctgccgc gccggcctcc gactgcatct ccattaagcg caccggctgg
24001 ccaaaaagg tctttttccc ctgcgccgcc cactcgtcga cgttttcgc catcggcgtg
24061 gaggggggta tggggtaaat cgccgcgacc tcggtaaagg cataagatg ccaggccgcc
24121 gcggcggtgc catccattgt tttcattttt ccggacattg ttcaatctc gaaggtgaga
24181 ggcatcttcg ccgcctcaaa taagcggcaa acccagttgt tgctcaagc acagcctgtg
24241 ccagctcgcg gatgacagaa gagtagcgc gaattcaacg cgttatgaag agagtcgccg
24301 cgcagcgcgc caagagattg cgtggaataa gacacagggg gcgacaagct gttgaacagg
24361 cgacaaagcg ccaccatggc cccggcaggc gcaattgttc tgtttcccac atttggctgc
24421 cttattgtgc cgttttgttt tacgtcctgc gcggcgacaa ataactaact tcataaaaat
24481 cataagaata cataaacagg cacggcagta ctctagtat ttaaattaat cacatttgat
24541 gatgtgaaaa gccggaggga aacctccggc ttttttgttt gcagcgcgcc cttcgggaaa
24601 cggtcacaaa acgcgttctt ccagagggtt atccctgta agccgcttcg ccagccacgg
24661 cggcaaacgg ccggccagca gcgtattgat cgctcgcac tgtgctgaa tcgtcgtccc
24721 gggctgcgct ttcacgggtt ggatgcgatg gcggatcagc gcgcggcgcg cgggccccgc
24781 tatctcctgg caaacagca gctgacagtc ctcaataagc cgggcgcgca cttcgttggc
24841 ctctgctgg tggggggcgg agggatagcg gcgcagatcg tacaaccage cgcttgttc
24901 atcgaacgcg tagataaaaa acagtcgtcc ctggccgaaa tggccgtaa tcaccagccc
24961 atcctgggag caaaatgcc ccagcagctg cggcgtgcgc tgggcagggc gaacgatgcg
25021 taaatgggcg ggcagcgcgc cctgcaggct cgccatcacc cgcgaccage gagcggggga
25081 catcaccgcc gtcgcggagg gaaagctggc ggccagctgc ggctgggtca ggtcggccag
25141 acgctctggc gtcagcgtct cgccgctctc ctgcgccage cagtcgacga tttgcgccgg
25201 ctgtaggtcc ggcagagact gaaacagcgc cagcatacgc cagaataggg tatcgttgtc
25261 ggacattgat tgaaagtaca agttttccga ttggaaatag agattctccc cgcacagtcc
25321 cggctgacgg gaaatcatct gttccagcca ctggcgtgta tggcgtgct ggctctgctg
25381 ccaagataa tcgcgcgcgc tctccagcgc ctgctgcggc tccatgggcg cggcggggcg
25441 aatcgcttcg caccataaaa gatgccagcc cagttcgtg gcgatgggaa ggcttagcgc
25501 gttttccgcc agtgaacaca gcgcgggtctc gagctgcgga tagagcaggc cacggcta

25561 ccagcctaaa cgacctctt ccagcgcgct cgggcagtgg gaatggcgct gggccagcgg
25621 cgcgaaagcg tcccgcgagg cgttgatttg ccgatacagg ccgaggatcc gctggtgcac
25681 ggcttcgcgg tcgttatcga ccgtcagcag taaatggcgg gtcagacgct gttccgggcg
25741 cataaactgc gtctggtggc gcagatacca tgctgtacc gtggagagat ccggctgcgg
25801 cgcctggcgg gcgatatcgg cgaaggcgag ttccagccgg gcgtgatggc gcacgatggc
25861 agcgcgctcg ggcggcgcaa aatccccctc gtcgagccag atggcaaggg aggcagcgat
25921 attctccagc aacgccgccg gaatatcgcc ttcagggacc cgcgcgacga tcgtctgctc
25981 catatggcac tggcgttgcc aggctgttc aaaagccggc gtatcggccg gatccagggc
26041 cgccggatcg cgattccagc ggctgcgcgc cagccgctgc cgggcaaac gttgccacgg
26101 gttcattgat tggaaataca agttttctga ctgaaagtaa aggttttcca gcacggtgcg
26161 tttaatctcc tcaagccagc tcgccagacg cgcttcggtc tggtcgaact ggttatcctg
26221 atccagcacc agccaacaa agcggtcgcc ttccagcgcc gaggacgcgc tgaattcata
26281 accctcattt ggccagctgc caatcatctg cgcgccgcgc gcgctcaggg cgtcgaacag
26341 cgggcgcacg ccgctgacga agttgtccgg atagcctctc tgatcgcca ggccgaacag
26401 cgccacgggtt ttccctttca ggctggcgtc gtcgaggccg ctgataaatt cgctccatga
26461 ctcgctttcg catccggcct ccagccccgg cagctggccg tcgccgagcg tcggcgtgcc
26521 cagcagcagc accggatagg ccataaagtc gtccagcgtc gtgcggttaa tgttgaccgg
26581 ggcacccgcc agctcgccca gttgcttatg gatcattttc gcgattttgc gggttttacc
26641 ggtatcgggtg ccaaagaaaa taccaatgtt cgccatgttg cgtcctgtc ggaaaagggg
26701 gttgaaaata cgcgttctcg caggggtatt gcgaaggctg tgccaggttg ctttgacta
26761 ccgcggccca tccctgcccc aaaacgatcg cttcagccct ctcccgccgc gcgcggcggg
26821 gctggcgggg cgcttaaaat gcaaaaagcg cctgcttttc ccctaccgga tcaatgtttc
26881 tgcacatcac gccgataagg gcgcacgggt tgcatggta tcaccgttcg gaaaacaccg
26941 cggcgtccct gtcacgggtg cggacaaatt gtcataactg cgacacagga gtttgcgatg
27001 accctgaaat ttaaatacta ga

//