

Vector construction with pRosa26-DEST

Peter Hohenstein

This overview is based on my experience with the GateWay system and the pRosa26-DEST vector in particular. Everything (vector and information) comes 'as it is' without any warranties. Feel free to contact me with questions, tips, problems, ideas or if you want to be kept informed on further improvements of the vector at peter.hohenstein@hgu.mrc.ac.uk

pRosa26-DEST was made by inserting a GateWay conversion cassette into the *Xho*I site of pBigT, followed by transfer of *Pac*I / *Asc*I fragment of the resulting vector to the pRosa26-PA vector (see <http://www.biomedcentral.com/1471-213X/1/4> for details on these vectors). The resulting vector resembles the construct that was used for the R26R Cre reporter mouse made by Phil Soriano in which the expression of lacZ is coupled to the endogenous Rosa26 locus via a splice acceptor site, but only after removal of the lox-STOP-lox cassette (which is actually the neo-R cassette used during the targeting). The only thing changed from the pBigT vectors and the R26R reporter is the cloning strategy, so the resulting knock-in should be as good as these systems, and as good (or as bad) as the use of the Rosa26 locus in any situation.

Generating Entry vectors for pRosa26 knock-in vectors

More detailed information on this is available in the GateWay manual (<http://www.invitrogen.com/content/sfs/manuals/gatewayman.pdf>). Please also check this for general information on the GateWay system.

Briefly, the options are as follows:

1. Use conventional cloning to insert your cDNA of interest in one of the pENTR vectors (http://www.invitrogen.com/content/sfs/manuals/pentr_man.pdf). This can be straightforward when you're lucky with restriction sites. However, it is important to check reading frame and unexpected start / stop codons.
2. Generate a PCR fragment with primers extended with *attB* sites and recombine into a pDONR vector (http://www.invitrogen.com/content/sfs/manuals/pdonr_man.pdf) using a BP reaction. This is my favourite method, since it gives complete control over start and stop signals and it allows extra sequences like epitope or purification tags to be included. Disadvantage is the sequencing of several clones that will be needed to check for PCR errors.
3. TOPO clone a PCR fragment without *att* sites added to the primers into an appropriate vector (http://www.invitrogen.com/downloads/B-12645_Gateway_TOPO.pdf). Will still need to be sequenced.
4. If you have a GateWay expression construct with your cDNA of interest, you can perform a BP reaction on this vector with a pDONR vector and grow bacteria on kanamycin selection. The resulting clones will be Entry clones with your cDNA.
5. You can buy sequence verified GateWay Entry clones from Invitrogen and several not-for-profit clone distribution sources.

You can check reading frame etc. of your construct at the design stage using Vector NTI, which is these days freely available for not-for-profit researchers (<https://catalog.invitrogen.com/index.cfm?fuseaction=userGroup.home>). With this program you can generate your Entry vector *in silico*, perform an *in silico* LR reaction with a DEST vector that has a CMV promoter and analyse the ORF in the resulting expression vector before proceeding to spending bench time and money. However, you can also do this with old-fashioned pen, paper and common sense.

Generating the Rosa26 knock-in vectors for cDNA expression

pRosa26-DEST contains the lethal *ccdB* counter selection gene used in the GateWay system. Therefore it can only be maintained in bacteria that are insensitive to this gene, like DB3.1 cells. The supplied bacteria are from this strain and carry the vector. If you need to retransform the vector, you will need to use DB3.1 cells.

The original pBigT vector system (<http://www.srinivas.org/>) on which pRosa26-DEST is based is notorious for its instability. In pRosa26-DEST this problem is partially but not completely solved. My experience so far with generating 10 different constructs shows that at least 3 / 16 *E.coli* clones tested are correct. For this I standard use Stb13 cells, which carry an extra mutation that makes instable constructs like lentivirus vectors more stable. I also grow the cells at 30° C to be on the safe side. I never tried standard DH5 α cells at 37° C.

I use the pRosa26-DEST with maxi-prep quality; my Entry vectors are usually mini-prep, in some cases I used mini-prep from an automated robotics system with some of the magnetic beads still in it and it worked fine, so apparently quality of this DNA is not the most critical factor.

1. Perform a standard 10 μ l LR reaction:
 - 150 ng pRosa26-DEST
 - 150 ng Entry vector
 - TE to 8 μ l
 - 2 μ l LR clonase II
2. Incubate 4-6 h at rt.
3. Add 1 μ l ProtK (2 μ g/ μ l, supplied with LR enzyme mix) and incubate 10' at 37° C. I often actually take 5 μ l in a fresh tube and treat this with ProtK, while incubating the remainder o/n at rt (just in case, I never actually needed this).
4. Transform 2-3 μ l into One-Shot Stb13 cells according to the supplied protocol. I let them recover shaking for 1 ½ h rather than 1 h at 30° C.
5. Plate on Amp plates and grow o/n at 30° C.
6. The next day pick 16 colonies and grow o/n at 30° for miniprep
7. On day 3 I first test all minipreps by digesting with *KpnI* (or it isoschizomer *Asp718*), which will linearize the vector, any unwanted recombination will give a clearly shorter fragment than you would expect. The clones that look correct on this digest I test further with *PacI/AscI* double digest (NEB buffer 4 + BSA) which will release the fragment as it was originally derived from pBigT, which will be easier to check the length after inserting your cDNA (remember that the GateWay reaction not only removes the sequence between the *att* sites, but that the *attR* sites from pRosa26-DEST will be replaced by *attB* sites which are only 25 nt each instead of 125, which will reduce the length of the fragment with 200 nt). I run the samples on 0.6-0.7% TAE gels at a relatively low voltage (max. 80 V).
8. Grow a 400 ml maxiprep of a correct clone o/n at 30° C.
9. Maxiprep the vector, take up DNA in 150 μ l TE and measure concentration.
10. Test 1 μ l DNA again with *KpnI* and *PacI/AscI* and if correct I digest 100 μ g vector with *KpnI* for electroporation (if your cDNA contains a *KpnI* site, you should also be able to use *XhoI*, *SciI*, *Acc65I*, *BcgI* or *AloI* for linearization, though I never had to test this).
11. After digestions and testing 1 μ l on gel for proper digestion, precipitate the DNA and keep it o/n at -20° C. The tube is ready to take into the ES lab.

Generating the Rosa26 knock-in vectors for Cre regulated RNAi

The idea of using the Rosa26 system for Cre-regulated RNAi is based on the fact that endogenous miRNAs are processed out of bigger RNA polymerase II transcribed transcripts (from introns or UTRs) by the Drosha complex, before further processing by Dicer and RISC into functional siRNAs. Since the Rosa26 knock-in systems will generate a fusion transcript between exon 1 of the Rosa26 transcript and the inserted sequence, it would recapitulate a normal situation for the expression of miRNA molecules. The work of the labs of Brian Cullen, Greg Hannon and others has shown that replacing the target sequence of an endogenous miRNA by a target sequence against your favourite gene will result in efficient knockdown of this gene (better than the comparable shRNA, especially at low copy numbers, likely due to better processing by Drosha and Dicer) from RNA polymerase II driven constructs and with better target sequence design rules.

The protocol described here is based in the Gateway-compatible miRNA vectors from Invitrogen pcDNA6.2-GW/miR (http://www.invitrogen.com/content/sfs/vectors/pcdna6_2gw_miR_map.pdf) and pcDNA6.2-GW/EmGFP-miR (http://www.invitrogen.com/content/sfs/vectors/pcdna6_2gw_miR_map.pdf). These vectors are based on the mouse miR-155 miRNA and target sequences are easily cloned into the system as dsDNA 64-mer oligos. I have generated another Gateway-compatible miRNA vector based on the human miR-30 backbone as used in the Hannon shRNA^{mir} library and using their cloning method. Though I'm sure their shotgun cloning and sequence-till-you-drop method is great for their whole genome libraries, I didn't find the method efficient enough (at least in my hands) for just few specific constructs at a time. Two other advantages of the Invitrogen vectors are the possibilities of linking GFP expression to the expression of the miRNA and linking several miRNA sequences in the same construct (either against the same gene to increase knockdown further if needed or against different genes to knockdown more than one gene with one construct). So far I haven't seen the GFP expression when targets were knocked into the Rosa locus (though I haven't looked too hard yet, in overexpression situations it's working great). My guess is that at low copy numbers (as in the single copy Rosa26 knock-in) all expressed transcripts are processed by Drosha, leading to loss of the GFP mRNA. In contrast, when doing overexpression experiments you're saturating the miRNA processing machinery, so the GFP you'll see is basically the leftovers. I've no hard data to support this, but in overexpression experiment we've seen GFP-negative cells that have great knockdown (I guess these are cells with only a few copies of the vector) but never GFP-positive cells that didn't show the knockdown. I haven't tried the linking of target sequence yet, but I cannot think of a reason why this wouldn't work when knocked into the Rosa26 locus.

Biggest disadvantage of the Invitrogen vectors is the fact that they're being sold as linearized vectors that you cannot maintain yourself. Personally I think that the speed of making constructs with these vectors more than compensates for this, but if people want to try my miR-30 based vector just contact me. Note that I haven't used these in the Rosa locus yet.

1. Design the 2 64-mer oligos using the design tool on the Invitrogen website (<https://rnaidesigner.invitrogen.com/rnaiexpress/>). After hybridising them the dsDNA oligo will have the correct overhang to insert into the Invitrogen vectors.
2. Order the cheapest quality oligos that you think you can get away with, in my experience this means the cheapest oligos from either Invitrogen or Sigma are fine, don't bother even trying cheap Eurogentec oligos.
3. Dilute and hybridise the oligos as described in the manual that comes with the vector.
4. After hybridising I continue directly with the ligation as described in the manual, I never test the oligos on gel. I usually ligate for 1h but I have done 5 min, which gave fewer colonies but still more than enough.
5. After ligation proceed immediately with transformation, when plating don't forget the vectors are Spec resistant, not Amp or Kan.
6. I sometimes get a mixture of small and large colonies. If this is the case, pick the large ones and grow o/n for miniprep. I usually grow 4 colonies / construct
7. Have the minipreps sequenced to check for mutations in the oligos. Sequencing the hairpins can be tricky, in our hands we get good results using 30 pmol of the forward sequencing

primer as described in the vector manual per reaction with the addition of 5% DMSO to the sequencing reaction.

8. When correct clones are identified, they can directly be used for overexpression experiments (for instance for testing). At this stage the GFP cassette can be removed if needed, or different target sequences can be linked, for all this see the vector manual. Alternatively, the miRNA (with or without GFP) can be shuttled into pRosa26-DEST in a combined LR/BP reaction.
9. Perform a BP reaction:
 - 150 ng pDONR221 (circular)
 - 150 ng miRNA vector (circular)
 - TE to 8 μ l
 - 2 μ l BP clonase II
10. Incubate 6 h at rt
11. Perform a LR reaction:
 - 3 μ l BP reaction
 - 150 ng pRosa26-DEST (circular)
 - TE to 8 μ l
 - 2 μ l LR clonase II
12. Incubate for o/n at rt
13. Add 1 μ l ProtK and incubate 10' at 37° C
14. Briefly cool reaction on ice
15. Transform 3 μ l to Stbl3 cells
16. Plate on Amp-R plates and grow o/n at 30° C
17. Check colonies via miniprep and digestion as described for the cDNA constructs.

Targeting with pRosa26-DEST based vectors

All our experiments so far have been with E14-IV cells, which are 129Ola. If you're using cells from another cell line or genetic background I'd be interested in your results.

We routinely electroporate 1×10^7 cells with 100 μ g linearized vector. Cells are plated on 10 x 100 mm dishes and grown o/n without selection. The next day we put our cells on neomycin selection. We pick colonies after 10 days selection. The vector targets with approximately 25% efficiency, so on average we pick 48-60 clones. This way we can always choose between sufficient numbers of good looking clones.

Clones are tested via Southern blot. The probe can be isolated from pRosa-5' (see <http://www.fhrc.org/science/labs/soriano/vectors/pROSA26-5prime.html>) with *EcoRI/HindIII*). The probe gives a wt band of 11 kb and a correctly targeted band of 3.8 kb on a *EcoRV* digest of genomic DNA. It doesn't give nice blots, but good enough for scoring.

We're currently using a first PCR screen to test colonies for correct homologues recombination. Based on these results we select a few clones to expand, freeze and confirm on Southern.

The PCR is based on the protocol as found on Phil Soriano's website. The most important part of the protocol is the PCR buffer. I've spend months and months being too lazy to make this buffer and just trying new primer sets with normal PCR buffer, without any success. With this buffer the PCR works like a treat. The primer sequences given here are different from the primers on Phil Soriano's website. They were designed and tested by Wenhao Xu (University of Virginia) and work better in our hands than the other primers. Finally, a hot start is important, but since we're doing our PCRs in 96 well plates with adhesive lids I don't find a normal hot start very convenient. The Invitrogen Platinum taq is an automatic hot start taq that works great in this PCR without any hassle.

Rosa26 PCR screen of ES clones

1. After picking, ES clones are grown in 96 well plates
2. Continue growing until medium in 80% of the wells turn orange-yellow in one day (change medium on all wells daily once the first wells turn yellow)
3. Split plate in 2 new gelatinised plates: plate A gets 20% of the cells, plate B gets 80%.
4. Next day wash cells in plate B with PBS and put 200 μ l lysis buffer on cells
5. Grow o/n at 37° C
6. Next day use 1 μ l lysate in a 25 μ l PCR reaction and analyse on 1.2% agarose gel. I usually continue with 4 PCR-positive clones.
7. Continue growing cells in plate A until wells are 70-80% confluent
8. Trypsinize selected wells and transfer complete well to 24 wells well
9. When 70-80% confluent passage complete well to 5 new wells on 24 wells plate
10. Grow to 80% confluency, pool cells from 4 wells and freeze in 8 vials for injection and *in vitro* activation of the construct, grow 5th well completely confluent for DNA isolation for Southern blot confirmation

Lysis buffer (12 ml):

10 x GB buffer:	1.2 ml
10% Triton X-100:	0.6 ml
milliQ H ₂ O:	10.2 ml
store at 4° C	

before use add 0.4 μ l ProtK (20 mg/ml) / 200 μ l lysis buffer; I dissolve ProtK in 50% glycerol for storage at -20° C to prevent subsequent freezing/thawing cycles)

10 x GB buffer (10 ml):

2M TRIS pH 8.8:	3.35 ml
1 M (NH ₄) ₂ SO ₄ :	1.66 ml
0.5 M MgCl ₂ :	1.34 ml
milliQ H ₂ O:	3.65 ml

store at -20° C

PCR mix (1 sample):

10 x MGB buffer:	2.5 µl
10 x dNTP (2 mM each):	2.5 µl
DMSO:	2.5 µl
F primer (12,5 µM):	0.2 µl
R primer (12,5 µM):	0.2 µl
Invitrogen Platinum Taq:	0.2 µl
β-mercapto-ethanol:	0.125 µl
milliQ H ₂ O:	11.775 µl

10 x MGB buffer (10 ml):

2M TRIS pH 8.8:	3.35 ml
1 M (NH ₄) ₂ SO ₄ :	1.66 ml
0.5 M MgCl ₂ :	1.30 ml
0.5% gelatine:	2.00 ml
milliQ H ₂ O:	1.69 ml

PCR conditions:

3' 95° C
45 X 20" 94° C
30" 62° C
2'30" 68° C
7' 72° C
10' 4° C
end

primer sequences:

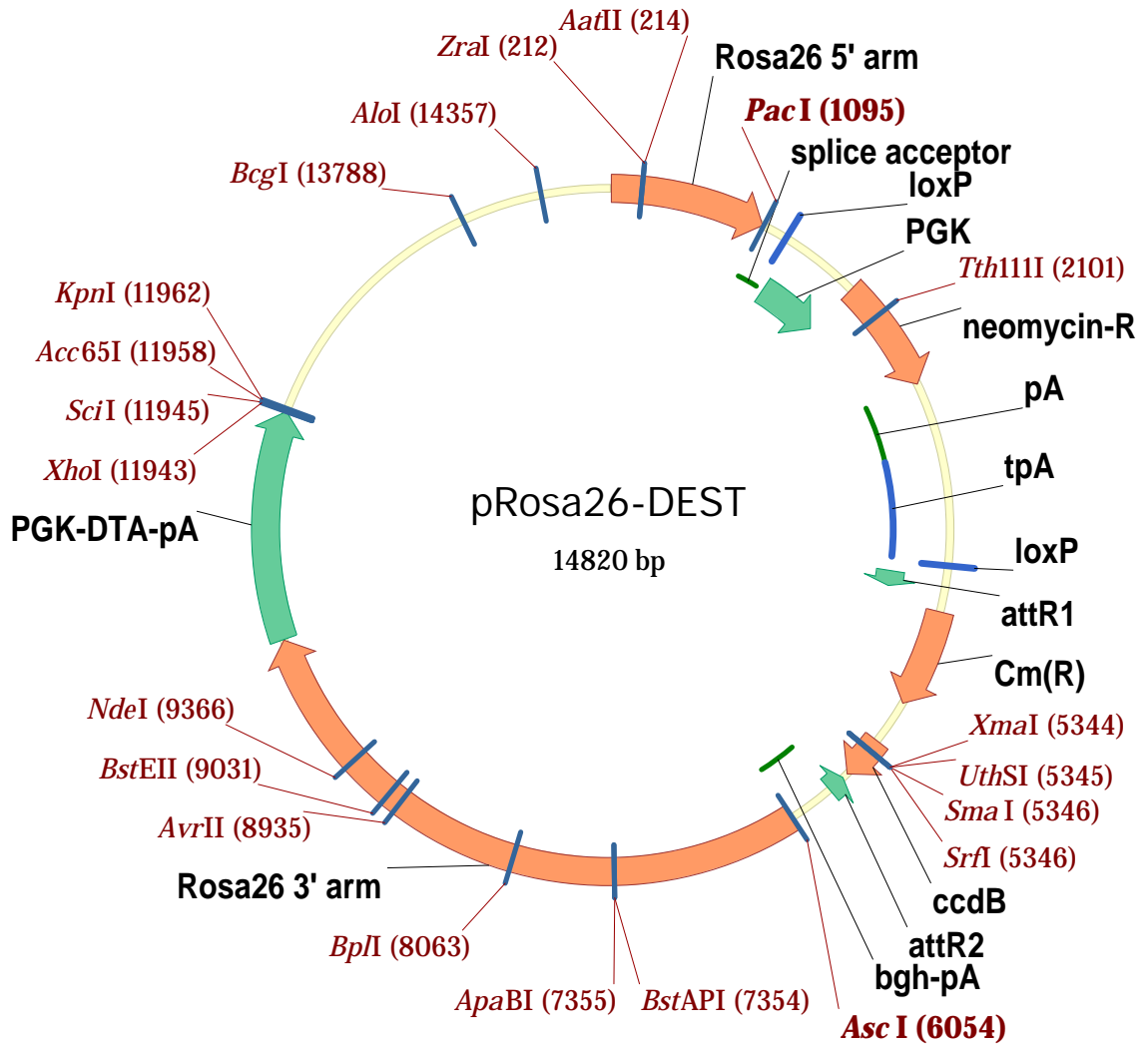
F: GGCGGACTGGCGGGACTA

R: GGGACAGGATAAGTATGACATCATCAAGG

Fragment length approx. 1.5 kb

Some handy cat. no. for Invitrogen stuff you might need...

BP clonase II: 11789020
LR Clonase II: 11791020
Stb13 cells: C7373-03
DB3.1 cells: 11782-018



LOCUS pRosa26-DEST 14820 bp DNA circular 9-MAR-2006
DEFINITION GateWay destination vector for knock-in in Rosa26 locus.
SOURCE

ORGANISM

COMMENT This file is created by Vector NTI

<http://www.invitrogen.com/>

COMMENT VNTDATE|404447100|

COMMENT VNTDBDATE|404447684|

COMMENT LSOWNER|

COMMENT VNTNAME|pRosa26-DEST|

COMMENT VNTAUTHORNAME|Peter Hohenstein|

COMMENT VNTREPLTYPE|Plasmid

COMMENT VNTTEXTCHREPL|Bacteria

FEATURES Location/Qualifiers

loci 1..1084
/vntifkey="18"
/label=Rosa26\5\arm

loci 6065..10335
/vntifkey="18"
/label=Rosa26\3\arm

misc_feature 10337..11945
/vntifkey="21"
/label=PGK-DTA-pA

polyA_signal 5738..6051
/vntifkey="25"
/label=bgh-pA

splicing_signal 1098..1271
/vntifkey="38"
/label=splice\acceptor

misc_feature 1272..1305
/vntifkey="21"
/label=loxP

promoter 1329..1840
/vntifkey="29"
/label=PGK

exon 1845..2648
/vntifkey="61"
/label=neomycin-R

polyA_signal 2649..3123
/vntifkey="25"
/label=pA

terminator 3124..3909
/vntifkey="43"
/label=tpA

misc_feature 3938..3971
/vntifkey="21"
/label=loxP

CDS 5262..5567
/vntifkey="4"
/label=ccdB

misc_recomb complement(5608..5732)
/vntifkey="86"
/label=attR2

CDS 4262..4942
/vntifkey="4"
/label=Cm(R)

misc_recomb 4029..4153
/vntifkey="86"
/label=attR1

BASE COUNT 3648 a 3385 c 3681 g 4088 t 18 others

ORIGIN

1 ccccgcgcca ggcctccga gcgtggtgga gccgttctgt gagacagccg ggtacgagtc
 61 gtgacgctgg aaggggcaag cgggtggtgg gcaggaatgc ggtccgacct gcagcaaccg
 121 gagggggagg gagaaggagg cggaaaagt cccaccggac gcggccatgg ctcggggggg
 181 ggggggagc ggaggascgc tccggccga cgtctcgtcg ctgattggct tyttttctc
 241 ccgccgtgtg tgaaaacaca aatggcgtgt tttggtggc gtaaggcgcc tgtcagttaa
 301 cggcagccgg agtgcgcagc cggccgagc ctgcctctgc ccactgggtg gggcgggagg
 361 taggtgggtg gaggcgagct gnacgtgcgg gcgcggtcgg cctctggcgg ggcggggagg
 421 gggaggagg gtcagcgaaa gtagctcgcg cgcgagcggc cggccacct ccccttctc
 481 tgggggagtc gttttaccg ccgcccggc ggcctcgtcg tctgattggc tctcggggcc
 541 cagaaaactg gcccttgcca ttggctcgtg ttcgtgcaag ttgagttcat ccgcccggca
 601 accggccctg cgaggaggcg ctcccaggtt ccggccctcc cctcggcccc gcgcgcgaga
 661 gtctggcgc gcgcccctgc gcaacgtggc aggaagcgcg cgtcgggggc ggggacgggc
 721 agtagggctg agcggctcgc gggcgggtgc aagcacgttt ccgacttgag ttgcctcaag
 781 aggggcgtgc tgagccagac ctccatcgcg cactccgggg agtggaggga aggagcagg
 841 gctcagttgg gctgttttgg aggcaggaa cacttgcctt cccaaagtcc ctctgagttg
 901 ttatcagtaa gggagctgca gtggagtagg cgggggagaag gccgcacct tctccggagg
 961 ggggagggga gtgttgcaat acctttctg gagttctctg ctgcctcctg gcttctgagg
 1021 accggccctg gcctgggaga atcccttccc ccctctccc ctctgactt gcaactccag
 1081 tctttctagt taattaagg atctgtagg cgcagtagtc cagggttccc ttgatgatgt
 1141 catacttacc ctgtcccttt tttttccaca gctcgcggtt gaggacaaac tcttcgcggt
 1201 ctttccagtg gggatcgacg gtatcgtaga gtcgaggccg ctctagaact agtggatccg
 1261 gaacccttaa tataacttcg tataatgtat gctatacgaa gttattagg cctcgcacct
 1321 gcaggaattc taccgggtag gggaggcgt tttcccagg cagtctggag catgcgcttt
 1381 agcagccccg ctggcacttg gcgctacaca agtggcctct ggcctcgcac acattccaca
 1441 tccaccggta gcgccaaccg gctccgttct ttgggtggccc cttcgcgcca ccttctactc
 1501 cccccctagt caggaagttc cccccgccc cgcagctcgc gtcgtgcagg acgtgacaaa
 1561 tggaagtagc acgtctcact agtctcgtgc agatggacag caccgctgag caatggaagc
 1621 gggtaggcct ttggggcagc ggccaatagc agctttgctc cttcgccttc tgggctcaga
 1681 ggctgggaa ggggtgggtc gggggcgggc tcaggggcgg gctcaggggc ggggcggggc
 1741 cgaaggtcct cccgaggccc ggcattctcg cacgcttcaa aagcgcacgt ctgcgcgct
 1801 gttctcctct tctcctctc cggcctttc gacctgcagc caatatggga tccggctatg
 1861 aacaagatgg attgcacgca ggttctccgg ccgcttgggt ggagaggcta ttcggctatg
 1921 actgggcaca acagacaatc ggctgctctg atgcgcctg gttccggctg tcagcgcagg
 1981 ggcgcccggg tctttttgtc aagaccgacc tgtccggtgc cctgaatgaa ctgcaggacg
 2041 aggcagcgcg gctatcgtgg ctggccaaga cgggcgttcc ttgcgcagct gtgctcgacg
 2101 ttgtcactga agcgggaagg gactggctgc tattggggca agtgcggggg caggatctcc
 2161 tgtcactca ccttgcctct gccgagaag tatccatcat ggctgatgca atgcggcggc
 2221 tgcatacgct tggatccgct acctgcctat tgcaccacca agcgaacat cgcactgcagc
 2281 gagcacgtac tgggatggaa gccggctctg tgcactcagg tgatctggac gaagagcatc
 2341 aggggctcgc gccagccgaa ctgctcgcga ggctcaaggc gcgcatgcc gacggcgatg
 2401 atctcgtcgt gaccatggc gatgcctgct tgccgaatat catggtggaa aatggcgcct
 2461 tttctggatt catcagctgt ggcggctggt gtgtggcggg ccgctatcag gacatagcgt
 2521 tggctaccgg tgatattgct gaagagcttg gcggcgaatg ggctgaccgc ttctcgtgc
 2581 tttacgggat cgccgctccc gattcgcagc gcatcgcctt ctatcgcctt cttgagagt
 2641 tcttctgagg ggatccgctg taagtctgca gaaattgatg atctattaaa caataaagat
 2701 gtccactaaa atggaagttt ttctgtcat actttgttaa gaagggtgag aacagagtac
 2761 ctacattttg aatggaagga ttggagctac ggggggtggg gtgggggtggg attagataaa
 2821 tgcctgctct ttactgaagg ctcttacta ttgctttatg ataagtttc atagttggat
 2881 atcataattt aaacaagcaa aaccaaatta agggccagct cattcctccc actcatgatc
 2941 tatagatcta tagatctctc gtgggatcat tgtttttctc ttgattccca ctttgggtt
 3001 ctaagtagct tggtttccaa atgtgtcagt ttcatagcct gaagaacgag atcagcagcc
 3061 tctgttccac atacacttca ttctcagtat tgttttgcca agttctaatt ccatcagaag
 3121 cttgcagatc tgcgactcta gaggatctgc gactctagag gatcataatc agccatacca
 3181 cattttaga ggttttactt gctttaaaaa acctcccaca cctcccctg aacctgaaac
 3241 ataaaatgaa tgcaattgtt gttgttaact tgtttattgc agcttataat ggttcaaaa
 3301 aaagcaatag catcacaatc ttcacaaata aagcattttt ttcactgcat tctagttgtg
 3361 gtttgcctca actcatcaat gtatctatc atgtctggat ctgcgactct agagactcat
 3421 aatcagccat accacatttg tagaggtttt acttgcttta aaaaacctcc cacacctccc
 3481 cctgaacctg aacataaaaa tgaatgcaat tgttgtgtgt aacttgttta ttgcagctta
 3541 taatggttac aaataaagca atagcatcac aaatttcaca aataaagcat ttttttact
 3601 gcattctagt tgtggtttgt ccaaactcat caatgtatct tatcatgtct ggatctgcga
 3661 ctctagaggga tcataatcag ccataaccaca tttgtagagg ttttacttgc tttaaaaaac
 3721 tccccacacc tccccctgaa cctgaaacat aaaatgaatg caattgttgt tgttaacttg
 3781 tttattgcag cttataatgg ttacaaataa agcaatagca tcacaaattt cacaaataaa
 3841 gcattttttt cactgcattc tagttgtggg ttgtccaaac tcatcaatgt atcttatcat
 3901 gtctggatcc ccatcaagct gatccggaac ctttaataa acttctgata atgtatgcta
 3961 tacgaagtta ttaggtccct cgacctgcag cccaagctag cttatcgata ccgtcagctt

4021 cgaatcaaac aagtttgtac aaaaaagctg aacgagaaac gtaaaatgat ataaatatca
4081 atatatataa ttagattttg cataaaaaac agactacata atactgtaaa acacaacata
4141 tccagtccta tgggcggccg cattaggcac cccaggcttt acactttatg cttccggctc
4201 gtataatgtg tggattttga gttaggatcc gtcgagattt tcaggagcta aggaagctaa
4261 aatggagaaa aaaatcactg gatataccac cgttgatata tcccaatggc atcgtaaaga
4321 acattttgag gcatttcagt cagttgctca atgtacctat aaccagaccg ttcagctgga
4381 tattacggcc tttttaaaga ccgtaaagaa aaataagcac aagttttatc cggcctttat
4441 tcacattctt gcccgcctga tgaatgctca tccggaattc cgtatggcaa tgaagacgg
4501 tgagctgggtg atatgggata gtgttcaccc ttgttacacc gttttccatg agcaactga
4561 aacgttttca tcgctctgga gtgaatacca cgacgatttc cggcagtttc tacacatata
4621 ttcgacaagat gtggctgtgt acggtgaaaa cctggcctat tccctaaag ggtttatga
4681 gaatatgttt ttcgtctcag ccaatccctg ggtgagtttc accagttttg atttaaacgt
4741 ggccaatatg gacaacttct tcgccccctg tttcacctat ggcaaatatt atacgcaagg
4801 cgacaagggtg ctgatgccgc tggcgattca ggttcatcat gccgtttgtg atggcttcca
4861 tgtcggcaga atgcttaatg aattacaaca gtactgcgat gagtggcagg cggggcgtaa
4921 tctagaggat ccggcttact aaaagccaga taacagtatg cgtattttgag cgtgatttt
4981 tgcggtataa gaatataatc tgatatgtat acccgaagta tgtcaaaaaa aggtatgcta
5041 atatggccag tattacagtg acagttgaca gcgacagcta tcagttgctc aaggcatata
5101 tgatgtcaat atctccggtc tggtaagcac aacctgcag aatgaagccc gtcgtctgag
5161 tgccgaacgc tggaaagcgg aaaatcagga agggatggct gaggtcgccc ggtttattga
5221 aatgaacggc tcttttctg acgagaacag gggctggtga aatgcagttt aaggtttaca
5281 cctataaaaag agagagccgt tatcgtctgt ttgtggatgt acagagtgat attattgaca
5341 cgccccggcg acggatgggt atccccctgg ccagtgcacg tctgctgtca gataaagtcc
5401 cccgtgaact ttaccgggtg gtgcataatc gggatgaaag ctggcgcag atgaccaccg
5461 atatggccag tgtgcccgtc tccgttatcg gggagaagt ggctgatctc agccaccgag
5521 aaaatgacat caaaaaagcc attaacctga tgttctgggg aatataaatg tcaggctccc
5581 ttatacacag ccagctctga ggtcgaccat agtgactgga tatgttgtgt tttacagtat
5641 tatgtagtct gttttttatg caaaatctaa tttaatatat tgatatttat atcattttac
5701 gtttctctgt cagctttctt gtacaaaagt gttcgattcg aggggcccc gggggcgccc
5761 gcgagctcgc tgatcagcct cgactgtgcc ttctagttag cagccatctg ttgtttgccc
5821 ctccccctg ccttccttga cctgcctcgg tgcactccc actgtccttt cctaataaaa
5881 tgaggaatt gcctgcatt gtctgagtag gtgtcattct attctggggg gtggggtggg
5941 gcaggacagc aagggggagg attgggaaga caatagcagg catgctgggg atgcggtggg
6001 ctctatggct tctgaggcgg aaagaaccag ctggggctcg atcctctagt tggcgcgccc
6061 tagaagatgg gcgggagtct tctggcagg cttaaaaggct aacctggtgt gtggcggtt
6121 tctgacagg gaattgaaca ggtgtaaaat tggagggaca agacttccc cagattttcg
6181 gttttgtcgg gaagttttt aataggggca aatagaaaaa tggagatag gactcatctg
6241 gggtttatgc agcaaaaacta caggtatatt gcttgtatcc gcctcggaga tttccatgag
6301 gagataaaga catgtcacc cagtttatac tctcctgctt agatcctact acagtatgaa
6361 atacagtgtg gcgaggtaga ctatgtaagc agatttaatc attttaaaaga gccagtact
6421 tcatatccat ttctcccgt ccttctgcag ccttatcaaa aggtatttag aacctcatt
6481 ttagcccat tttcatttat tatactggct tatccaacc ctgacagag cattggcatt
6541 ttccctttcc tgatcttaga agtctgatga ctcatgaaac cagacagatt agttacatac
6601 accacaaatc gaggctgtag ctggggcctc aacctgcag ttctttata actcctagt
6661 acactttttg ttgatccttt gccttgatcc ttaattttca gtgtctatca cctctccgct
6721 caggtggtgt tccacatttg ggcctattct cagtccagg agttttacaa caatagatgt
6781 attgagaatc caacctaaag cttactttc cactcccatg aatgcctctc tctttttct
6841 ccattataac tgagctatwa ccattaatgg tttcaggtgg atgtctctc ccccaatata
6901 cctgatgtat ctacatattg ccaggctgat attttaagac atwaaaggta tatttcatta
6961 ttgagccaca tggatttgat tactgtact aaaatttgt cattgtacac atctgtaaaa
7021 ggtggttctc tttggaatgc aaagttcagg tgtttgtgt ctttctgac ctaaggtctt
7081 gtgagcttgt attttttcta tttagcagt gctttctctt ggactggctt gactcatggc
7141 attctacacg ttattgctgg tctaaatgtg attttgcaa gcttcttcag gacctataat
7201 tttgcttgac ttgtagcaa acacaagtaa aatgattaag caacaaatgt atttgtgaag
7261 cttggttttt aggttgttgt gttgtgtgtg cttgtgctct ataataatac tatccagggg
7321 ctggagaggt ggctcggagt tcaagagcac agactgctct tccagaagtc ctgagttcaa
7381 ttcccagcaa ccacatgggt gctcacaacc atctgtaatg ggatctgatg cctctctctg
7441 gtgtgtctga agaccacaag tgtattcaca ttaataaaat aatcctcctt ctctctctt
7501 tttttttttt aaagagaatw ctgtctccag tagaattact gaagtaatga aatactttgt
7561 gtttgttcca atatggwagc caataatcaa atactcttwa gcactggaaa tgtaccaagg
7621 aactatttta ttttaagtwa ctgtggacag aggagccata actgcagact tgtgggatac
7681 agaagaccaa tgcagactta atgtctttc tcttacta agcaataaag aaataaaaa
7741 tgaacttcta gtatctatt tgttaactg ctactttac taacttttgt gcttcatcta
7801 tacaaagctg aaagctaagt ctcagacat tactaaacat gaaagcaagt aatgataat
7861 ttggatttca aaaatgtagg gccaggttt agccagccag tgggtgtgct tgctttatg
7921 ccttaatccc agcactctgg aggcagagac aggcagatct ctgagtttga gccagcctg
7981 gtctacacat caagttctat ctaggatagc caggaataca cacagaacc ctgttgggga
8041 ggggggctct gagatttcat aaaattataa ttgaagcatt ccctaatgag ccactatgga

8101 tgtggctaaa tccgtctacc tttctgatga gatttgggta ttatTTTTTc tgtctctgct
8161 gttggttggg tcttttgaca ctgtggcctt tcttaaagcc tccttcctcg ccatgtggctc
8221 tcttgtttgc tactaacttc ccatggctta aatggcatgg ctttttgctt tctaagggca
8281 gctgctgagw tttgcagcct gatttccagg gtgggggttg gaaatctttc aaacataaa
8341 attgctcttt aatTTTTTTt taaaaaatgg gttatataat aaacctcata aaatagtat
8401 gaggagtgag gtggactaat attaagtgtt ccctcccta taaaagagct attaaggctt
8461 tttgtcttat actaactttt tttttaaatg tggatctttt agaaccaagg gtcttagagt
8521 tttagtatac agaaactgtt gcatcgctta atcagatTTT ctagtttcaa atccagagaa
8581 tccaaattct tcacagccaa agtcaaatta agaatttctg actttaatgt tatttgctac
8641 tgtgaatata aaatgatagc ttttcttgag gcagggtctc actatgtatc tctgcctgat
8701 ctgcaacaag atatgtagac taaagtctcg cctgcttttg tctcctgaat actaaggTTa
8761 aaatgtagta atacttttgg aacttgacag tcagattcct ttatagggga cacactaagg
8821 gagcttgggt gatagtTggT aaatgtgttt aagtgatgaa aacttgaatt attatcaccg
8881 caacctactt tttaaaaaaa aaagccaggc ctgtagagc atgctaaggg atccctagga
8941 cttgctgagc acacaagagt agtactTggc aggctcctgg tgagagcata tttcaaaaa
9001 caaggcagac aaccaagaaa ctacagtaag gttacctgtc ttaaccatc tgcataatac
9061 cagggatatt aaaatattcc aaataatatt tcattcaagt tttccccat caaattggga
9121 catgcaattc tccggtgaat aggcagagtT ggaactaaa caaatgtTgg tttgtgatt
9181 tgtgaaattg ttttcaagtg atagtTaaag cccatgagat acagaacaaa gctgctattt
9241 cgaggctctc tggttatact cagaagcact tctttgggtt tcctgcact atcctgatca
9301 tgtgctaggc ctwccttagg ctgattgttg ttcaataac ttaagtTcc tgtcaggTga
9361 tgtcatatga tttcatatat caaggcaaaa catgttatat atgtTaaaca tttgkactta
9421 atgtgaaagt taggtctttg tgggttttga ttttaatttc aaaacctgag ctaaataagt
9481 cattttacat gtcttacatt tggTgaattg tatattTgg tttgcaggca agactctctg
9541 acctagtaac cctcctatag agcactTtgc tgggtcacia gtctaggagt caagcatttc
9601 accttgaagt tgagacgttt tgttagtgta tactagtTat atgtTggagg acatgtttat
9661 ccagaagata ttcaggacta tttttgactg ggctaaggaa ttgattctga ttagcactgt
9721 tagtgagcat tgagtggcct ttaggctTga attggagtca cttgtatate tcaataaatg
9781 ctggcctttt ttwaaaagcc cttgtctttt atcaccctgt tttctacata atttttgttc
9841 aaagaatac ttgtttggat ctctttttga caacaatagc atgttttcaa gccataattt
9901 ttttcctttt tttttttttt tttggttttt cgagacaggg tttctctgta tagcctggc
9961 tgtcctggaa ctcaactTgt agaccaggct ggcctcgaac tcagaaatcc gctgctctc
10021 gcctcctgag tgccgggatt aaaggcgtgc accaccacgc ctggctaagt tggatatttt
10081 gtatataact ataaccaata ctaactccac tgggtggatt ttaattcag tcagtagTct
10141 taagtgtctt ttattggccc ttattaaat ctactgtTca ctctaacaga ggctgtTgga
10201 ctagtgsac taagcaactt cctacggata tactagcaga taagggtcag ggatagaaac
10261 tagtctagcg ttttgtatac ctaccagctt atactacctt gttctgatag aaatatttag
10321 gacatctagc ttatcgatcc ttcgacgTga tcgataagct tgatatcgaa ttctaccggg
10381 taggggaggc gcttttccaa ggcagctTga gcatgcgctt agcagccccg ctggcacttg
10441 gcgctacaca agtggcctyt ggcctcgcac acattccaca tccaccggta ggcgccaacc
10501 ggctccgttc tttggtggcc ccttcgccc accttctwct cctcccctag tcaggaaagt
10561 ccccccgcc ccgcagctcg cgtcgtTsaag acgtgacaaa tggaaagtag acgtctcact
10621 agtctcgtca gatggacagc accgctgagc aatggaagcg ggtaggcctt tggggcagcg
10681 gccaatagca gctttgtccc ttcgctttct gggctcagag gctgggaag ggtgggtccg
10741 ggggcggctc caggggcggg ctacggggcg gggcggggcg ccgaaggtcc tccggaggcc
10801 cggcattctg cacgctTcaa aagcgcacgt ctgcccgcgt gttctctctc tctcatctc
10861 cgggcctttc gacctgcagg tctctgccaat ggatcctgat gatgtTgTta tcttctaat
10921 cttttgtatg gaaaactttt cttcgtacca cgggactaaa cctggttatg tagattccat
10981 tcaaaaaggt atacaaaagc caaaatctgg tacacaagga aattatgacg atgattggaa
11041 agggTTTTat agtaccgaca ataaatacga cgctgcggga tactctgtag ataatgaaaa
11101 cccgctctct ggaaaagctg gaggcgtTgt caaagtgacg tatccaggac tgacgaaggt
11161 tctcgacta aaagtggata atgccgaaac tattaagaaa gagttagggt taagtctcac
11221 tgaaccgttg atggagcaag tcggaacgga agagtTtatc aaaaggttcg gtgatggTgc
11281 ttcgctgTga gtgctcagcc ttcctctcgc tgaggggagt tctagcgttg aatatattaa
11341 taactgggaa caggcgaag cgTtaagcgt agaactTgag attaattttg aaacctgTgg
11401 aaaacgtggc caagatgcga tgtatgagta tatggctcaa gcctgtgcag gaaactgTgt
11461 cagcagatct ctttTgTgaag gaacctact tctgtTgtgt gacataattg gacaaactac
11521 ctacagagat ttaaagctct aaggTaaata taaaattttt aagtTataaa tTgtTaaac
11581 tactgattct aattgtTttgt gtattttTaga ttccaacctc tggaaactgat gaatgggagc
11641 agtggTggaa tgcagatcct agagctcgtc gatcagcctc gactgtgcct tctagtTgcc
11701 agccatctgt tgtttgccc tccccgTgc cttcctTgac cctggaaggt gccactcca
11761 ctgtcctttc ctaataaaaat gaggaaattg catcgcatTg tctgagtagg tgtcatttca
11821 ttctgggggg tggggtgggg caggacagca agggggagga tgggaaagc aatagcaggc
11881 atgctgggga tgcggtgggc tctatTgctt ctgagggcga aagaaccagc tgggctcga
11941 cctcgagggg gggcccggta cccagctttt gttcccttta gtgagggTta attgcccgtc
12001 tggcgtaatc atggTcatag ctgtttcctg tgtgaaattg ttatccgctc acaattccac
12061 acaacatacg agccggaagc ataaagtTga aagcctgggg tgccaatga gtgagctaac
12121 tcacattaat tgcgtTgcgc tcactgcccg ctttccagtc gggaaacctg tctgTccagc

12181 tgcattaatg aatcggccaa cgcgcgggga gaggcggttt gcgtattggg cgctcttccg
12241 cttcctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc
12301 actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt
12361 gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcggtgctg gcggttttcc
12421 ataggctccg cccccctgac gagcatcaca aaaatcgacg ctcaagtcaag aggtgcgcaa
12481 acccgacagg actataaaga taccaggcgt tccccctggg aagctccctc gtgcgctctc
12541 ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcc ggaagcgtgg
12601 cgctttctca tagctcacgc tgtaggtatc tcagttcggg gtaggtcgtt cgctccaagc
12661 tgggctgtgt gcacgaacc cccgttcagc ccgaccgctg cgccctatcc ggtaactatc
12721 gtcttgagtc caaccggta agacacgact tatcgccact ggcagcagcc actggtaaca
12781 ggattagcag agcaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact
12841 acggctacac tagaaggaca gtatttggta tctgcgctct gctgaagcca gttacctccg
12901 gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggtttt
12961 ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct
13021 tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga
13081 gattatcaaa aaggatctt acctagatcc ttttaaatca aaaatgaagt tttaaatcaa
13141 tctaaagtat atatgagtaa acttggctcg acagttacca atgcttaatc agtgaggcac
13201 ctatctcagc gatctgtcta tttcgttcac ccatagttgc ctgactccc gtcgctaga
13261 taactacgat acggggggc ttaccatctg gccccagtgc tgcaatgata cgcgagacc
13321 cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg gccgagcga
13381 gaagtggctc tgcaacttta tccgcctcca tccagtctat taattgttgc cgggaagcta
13441 gagtaagtag ttcgccagt aatagtttgc gcaacgttgt tgccattgct acaggcatcg
13501 tgggtgtcacg ctgctcgtt ggtatggctt cattcagctc cggttcccaa cgatcaaggc
13561 gagttacatg atccccatg ttgtgcaaaa aagcggtag ctccctcggg cctccgactc
13621 ttgtcagaag taagttggcc gcagtgttat cactcatggt tatggcagca ctgcataatt
13681 ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtagtac tcaaccaagt
13741 cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgtca atacgggata
13801 ataccgcgcc acatagcaga actttaaag tgctcatcat tggaaaacgt tcttcggggc
13861 gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaacc actcgtgcac
13921 ccaactgatc ttcagcatct tttactttca ccagcgttc tgggtgagca aaaacaggaa
13981 ggcaaaatgc cgcaaaaag ggaataaggc cgacacggaa atgttgaata ctcatactct
14041 tcctttttca atattattga agcatttacc agggttattg tctcatgagc ggatacatat
14101 ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc
14161 cacctaaatt gtaagcgtta atattttgtt aaaattcggg ttaaattttt gttaaatcag
14221 ctcatTTTTT aaccaatagg ccgaaatcgg caaaatccct tataaatcaa aagaatagac
14281 cgagataggg ttgagtgttg ttccagtttg gaacaagagt ccactattaa agaacgtgga
14341 ctccaacgct aaagggcgaa aaaccgtcta tcagggcgat ggcccactac gtgaaccatc
14401 accctaatca agtttttgg ggtcgaggtg ccgtaaagca ctaaactcgg accctaaagg
14461 gagccccga tttagagctt gacggggaaa gccggcgaac gtggcgagaa aggaagggaa
14521 gaaagcgaag ggagcgggag ctagggcgct ggcaagtgtg gcggtcacgc tgcgcgtaac
14581 caccacacc gccgcgctta atgcgcccgt acagggcgcg tcccatcgc cattcaggct
14641 gcgcaactgt tgggaagggc gatcgggtcg ggcctcttcc ctattacgcc agctggcgaa
14701 agggggatgt gctgcaaggc gattaagttg ggtaacgcca gggttttccc agtcacgacg
14761 ttgtaaaacg acggccagtg agcgcgcgta atacgactca ctataggcgg aattggagct