



## General Description

DNA pHAGE2-TetOminiCV-Oct4F2AKlf4-IRES-Sox2E2AcMyc-W  
Entire molecule length: 11745 bp

## Standard Fields

## User Fields

## Author(s)

### Author

Gustavo Mostoslavsky

### Original author

Gustavo Mostoslavsky

## References

## Comments

## Feature Map

### CDS (7 total)

#### HIV gag

Start: 790 End: 1144

HIV gag sequence from 336bp to 688bp (Clone to PvuII site in HIV and filled in ClaI digest to

#### env

Start: 1156 End: 2013

HIV env sequence from bp7168 to 8025 (contains RRE) Cloned from BglII to BamHI out of HIV

#### Oct4

Start: 2685 End: 3740

#### Klf4

Start: 3816 End: 5240

#### Sox2

Start: 5828 End: 6784

#### cMyc

Start: 6854 End: 8173

#### AMPr

Start: 9359 End: 10198

Ampicillin Resistance Gene

Original Location Description:

6153..6992

Qualifiers:

/gene="AMPr"

/product="beta-lactamase (mature form)"

### LTR (2 total)

#### 5' LTR

Start: 1 End: 636

Long Terminal Repeat

Original Location Description:

1..636  
Qualifiers:  
/gene="LTR"

### 3' LTR

Start: 8782 End: 9103

### Misc. Feature (26 total)

#### HIV U3

Start: 1 End: 454  
Full Length HIV U3

#### HIV R

Start: 455 End: 551  
HIV Repeat Region

#### HIV U5

Start: 552 End: 636  
HIV U5

#### PBS-K

Start: 636 End: 658  
tRNA binding site for Lysine tRNA

#### PSI

Start: 637 End: 1155  
PSI Packaging Sequence  
Original Location Description:  
637..1155  
Qualifiers:  
/gene="psi"  
/product="pbs-gag"

#### HIV-PSI

Start: 697 End: 806

#### HR-ePSI

Start: 807 End: 1144  
HIV gag sequence 351bp to 688bp

#### Stop Codon

Start: 853 End: 855  
Stop codon in gag put in frame by ClaI fill in

#### RRE

Start: 1245 End: 1593  
Rev Responsive Element according to Belasco in Molecular Cell, Vol. 7, 603-614, March, 2001

#### RRE

Start: 1303 End: 1536  
Rev Responsive Element  
Original Location Description:  
1303..1536  
Qualifiers:  
/gene="rre"  
/product="minimal RRE"

#### Junk

Start: 2014 End: 2027  
Linker Junk Sequence probably left over from cloning

#### HIV cpPu (Trip)

Start: 2034 End: 2211  
Central Poly Purine Track  
Original Location Description:

2034..2211  
Qualifiers:  
/gene="cppt"

TetO

Start: 2238 End: 2555  
Tetracyclin responsive element (TRE)

GSG bridge

Start: 3741 End: 3749

F2A

Start: 3750 End: 3815

IRES

Start: 5247 End: 5824

GSG bridge

Start: 6785 End: 6793

E2A

Start: 6794 End: 6853

WPRE

Start: 8186 End: 8777  
Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element  
Feature extends beyond the specified end point  
Original Location Description:  
3563..4154  
Qualifiers:  
/gene="WPRE"

HIV nef

Start: 8782 End: 8883  
HIV nef

HIV PPT

Start: 8832 End: 8846  
PolyPurine Tract

dU3

Start: 8849 End: 8885  
1st Half of deleted U3

dU3

Start: 8905 End: 8923  
2nd Half of deleted U3

HIV R

Start: 8924 End: 9020

HIV U5

Start: 9021 End: 9103  
HIV U5

SV40 ori

Start: 11197 End: 11320  
Original Location Description:  
7991..8114  
Qualifiers:  
/gene="SV40 ori"

PolyA Signal (1 total)

G-T cluster

Start: 570 End: 577

GT Cluster shown to be necessary for Poly-A site activity by Cullen in J Virol. 1989 January;

PolyA Site (1 total)

pA Site

Start: 550 End: 550  
Exact base to which Poly-A is added

Primer Binding Site (27 total)

pHAGE-A0050-8050

Start: 158 End: 179 (Complementary)

pHAGE-S501-1000

Start: 401 End: 421

pHAGE-A0550-0050

Start: 667 End: 690 (Complementary)

pHAGE-S1001-1500

Start: 898 End: 919

pHAGE-A1050-0550

Start: 1166 End: 1189 (Complementary)

pHAGE-S1501-2000

Start: 1391 End: 1417

pHAGE-A1550-1050

Start: 1650 End: 1678 (Complementary)

pHAGE-S2001-2500

Start: 1898 End: 1922

pHAGE-A2050-1550

Start: 2084 End: 2112 (Complementary)

pHAGE 5' IRES

Start: 5263 End: 5283

pHAGE 3' CDS 1-1 (Binds IRES)

Start: 5271 End: 5291 (Complementary)

pHAGE 5' CDS 2-1 Start (Binds IRES)

Start: 5527 End: 5545

pHAGE 3' CDS 1-2 (Binds IRES)

Start: 5527 End: 5545 (Complementary)

pHAGE 3' CDS

Start: 8214 End: 8248 (Complementary)

pHAGE-A3550-3050

Start: 8292 End: 8315 (Complementary)

pHAGE-S4001-4500

Start: 8445 End: 8465

pHAGE-S6001-6500

Start: 9135 End: 9157

pHAGE-A6050-5550

Start: 9408 End: 9431 (Complementary)

pHAGE-S6501-7000

Start: 9608 End: 9632

pHAGE-A6550-6050

Start: 9877 End: 9900 (Complementary)

pHAGE-S7001-7500

Start: 10128 End: 10152

pHAGE-A7050-6550

Start: 10388 End: 10409 (Complementary)

pHAGE-S7501-8000

Start: 10632 End: 10656

pHAGE-A7550-7050

Start: 10899 End: 10920 (Complementary)

pHAGE-S8001-8500

Start: 11135 End: 11154

pHAGE-A8050-7550

Start: 11375 End: 11401 (Complementary)

pHAGE-S1-500

Start: 11640 End: 11668

#### Promoter Eukaryotic (1 total)

miniCMV

Start: 2556 End: 2675

#### Promoter Prokaryotic (1 total)

P-Bla

Start: 9306 End: 9340

Beta Lactamase Promoter

#### Replication Origin (1 total)

ORI

Start: 10427 End: 10960

Bacterial Origin of Replication

Original Location Description:

7221..7754

Qualifiers:

/gene="ORI"

/product="ColE1 origin of replication"

#### Splicing Signal (3 total)

5' SD 289

Start: 744 End: 745

Major 5' Splice Donor 289

SA 7925

Start: 1911 End: 1912

SA 4459

Start: 2188 End: 2189

#### Stem Loop (4 total)

SL-1

Start: 697 End: 731

PSI Stem loop 1

### SL-2

Start: 736 End: 754  
PSI Stem loop 2

### SL-3

Start: 766 End: 779  
PSI Stem Loop 3

### SL-4

Start: 793 End: 806  
PSI Stem Loop 4

### 5' UTR (1 total)

#### 5' UTR

Start: 551 End: 635  
HIV 5' UTR

### Mutation (2 total)

#### Fill-in

Start: 833 End: 834  
Filled in Cla Site from original HIV genome to disrupt gag polyprotein ORF

#### RRE Mutation

Start: 1357 End: 1357  
This G is a C in the WT HIV-1 genome

### Restriction/Methylation Map

Enzyme	# of cuts	Positions
Acc65I	3	2541 5680 8779
AfIII	3	518 3244 8987
ApaLI	4	5716 8397 9457 10703
Ascl	1	8899
Aval	13	296 1928 2238 2545 3061 4257 4365 4762 5989 6383 7040 7705 9105
BamHI	1	2772
BclI	1	6652
BglII	3	474 8815 8943
Clal	1	8176
DraI	7	2062 5239 8832 9551 10243 10262 11703
EagI	7	1145 1149 2678 3963 4393 6096 6165
EcoRI	1	2029
EcoRV	4	36 115 334 6998
EgeI	6	640 3508 4571 6272 6622 6674
FseI	1	1151
FspI	1	9906
HindIII	7	532 1088 1676 2790 5463 5875 9001
KpnI	3	2545 5684 8783
MscI	2	3994 5817
NaeI	3	1149 6275 8702
NcoI	3	2684 3815 6550

NdeI	1	5824
NotI	1	2678
PmlI	2	291 5555
PstI	6	3092 4225 6451 6976 7393 8892
PvuI	1	9760
PvuII	5	436 6365 7151 7815 8810
SacI	7	492 683 2539 2593 4388 4904 8961
SacII	3	6171 7355 8691
SbfI	2	7393 8892
SdaI	2	7393 8892
SfiI	1	11249
SmaI	5	2547 4259 5991 6385 9107
SpeI	1	2213
StuI	4	2576 3229 3691 11295
XcmI	3	4666 6557 6590
XhoI	1	2238
XmaI	5	2545 4257 5989 6383 9105

No cuts: BsrGI, HpaI, MluI, NheI, PaeI, PmeI, Sall, SrfI, SwaI, XbaI

### Sequence

```

1  tggaaagggct aattcactcc caaagaagac aagatatcct tgatctgtgg atctaccaca
cacaaggcta cttccctgat tagcagaact acacaccagg
101 gccaggggtc agatatccac tgaccttgg atggtgctac aagctagtag cagttgagcc
agataaggta gaagaggcca ataaaggaga gaacaccagc
201 ttgttacacc ctgtgagcct gcatgggatg gatgaccgag agagagaagt gttagagtgg
aggtttgaca gccgcctagc atttcatcac gtggcccag
301 agctgcatcc ggagtacttc aagaactgct gatatcgagc ttgctacaag ggactttccg
ctggggactt tccagggagg cgtggcctgg gcgggactgg
401 ggagtggcga gccctcagat cctgcatata agcagctgct ttttgctgt actgggtctc
tctggttaga ccagatctga gcctgggagc tctctggcta
501 actaggggaa cactgctta agcctcaata agcttgctt tgagtgttc aagtagtgtg
tgcccgtctg ttgtgtgact ctggtaacta gagatccctc
601 agacctttt agtcagtgtg gaaaatctct agcagtggcg cccgaacagg gacttgaaag
cgaaagggaa accagaggag ctctctcgac gcaggactcg
701 gcttgctgaa gcgcgcacgg caagaggcga gggcgggcga ctggtgagta cgccaaaaat
ttgactagc ggaggctaga aggagagaga tgggtgagc
801 agcgtcagta ttaagcgggg gagaattaga tcgcatggg aaaaaattcg gttaggcca
gggggaaaga aaaaataaa attaaaacat atagtatggg
901 caagcagggg gctagaacga ttcgcagtta atcctggcct gttagaaca tcagaaggct
gtagacaaat actgggacag ctacaacat cccttcagac
1001 aggatcagaa gaacttagat cattatataa tacagtagca accctctatt gtgtgcatca
aaggatagag ataaaagaca ccaaggaagc tttagacaag
1101 atagaggaag agcaaaaca aagtaagacc accgcacagc aagcggccgg ccgctgatct
tcagacctgg aggaggagat atgagggaca attggagaag
1201 tgaattatat aatataaag tagtaaaaat tgaaccatta ggagtagcac ccaccaaggc
aaagagaaga gtggtgcaga gagaaaaaag agcagtggga
1301 ataggagctt tgttccttgg gttcttggga gcagcaggaa gcactatggg cgcagcgtca
atgacgctga cggtacaggc cagacaatta ttgtctggta
1401 tagtgcagca gcagaacaat ttgctgaggg ctattgaggg gcaacagcat ctgttgcaac
tcacagtctg gggcatcaag cagctccagg caagaatcct
1501 ggctgtggaa agatacctaa aggatcaaca gtcctggggg atttgggggt gctctggaaa
actcatttgc accactgctg tgccctggaa tgctagtgtg
1601 agtaataaat ctctggaaca gatttggaa cacacgacct ggatggagtg ggacagagaa
attaacaatt acacaagctt aatacactcc ttaattgaag
1701 aatcgcaaaa ccagcaagaa aagaatgaac agaattatt ggaattagat aaatgggcaa
gtttgtggaa ttggtttaac ataacaaatt ggctgtggta

```



1801 tataaaatta ttcataatga tagtaggagg cttgtaggt ttaagaatag tttttgctgt  
actttctata gtgaatagag ttaggcaggg atattcacca  
1901 ttatcgtttc agaccacact cccaaccccg aggggaccg acaggcccg aggaatagaa  
gaagaaggtg gagagagaga cagagacaga tccattcgat  
2001 tagtgaacgg atctcgacgg tatcgccgaa ttcacaaatg gcagtattca tccacaattt  
taaaagaaaa ggggggattg gggggtacag tgcaggggaa  
2101 agaatagtag acataatagc aacagacata caaactaaag aattacaaaa acaaattaca  
aaaattcaaa attttcgggt ttattacagg gacagcagag  
2201 atccagtttg gactagtcca caccctaact gacacactcg agtttaccac tcctatcag  
tgatagagaa aagtgaaagt cgagtttacc actccctatc  
2301 agtgatagag aaaagtgaaa gtcgagttta cactcccta tcagtgatag agaaaagtga  
aagtcgagtt taccactccc tatcagtgat agagaaaagt  
2401 gaaagtcgag ttaccactc cctatcagtg atagagaaaa gtgaaagtcg agtttaccac  
tcctatcag tgatagagaa aagtgaaagt cgagtttacc  
2501 actccctatc agtgatagag aaaagtgaaa gtcgagctcg gtacccgggt cgaggtaggg  
gtgtacgggt ggaggcctat ataagcagag ctcgtttagt  
2601 gaaccgtcag atcgccctgga gacgccatcc acgctgtttt gacctccata gaagacaccg  
ggaccgatcc agcctgcggc cgccatggct ggacacctgg  
2701 cttcagactt cgccttctca cccccaccg tgggggtga tgggtcagca gggctggagc  
cgggctgggt ggatcctcga acctggctaa gttccaagg  
2801 gctccaggt gggcctgga tcggaccagg ctgagaggta ttggggatct ccccatgtcc  
gcccgcatac gagttctgcg gagggatggc atactgtgga  
2901 cctcaggttg gactgggct agtccccca gttggcgtgg agactttgca gcctgagggc  
caggcaggag cacgagtgga aagcaactca gagggaacct  
3001 cetctgagcc ctgtgcccgc cgcaccaatg cegtgaagt ggagaaggtg gaaccaactc  
ccgaggatc ccaggacatg aaagccctgc agaaggagct  
3101 agaacagttt gccaagctgc tgaagcagaa gaggatcacc ttgggggtaca cccaggccga  
cgtggggctc acctgggcg ttctcttg aaaggtgttc  
3201 agccagacca catctgctc cttcagggcc ttgcagctca gccttaagaa catgtgtaag  
ctgcccggcc tgctggagaa gtgggtggag gaagccgaca  
3301 acaatgagaa ccttcaggag atatgcaaat cggagaccct ggtgcaggcc cggagagaa  
agcgaactg cattgagaac cgtgtgaggt ggagtctgga  
3401 gaccatgttt ctgaagtgc cgaagccctc cctacagcag atcactcaca tcgccaatca  
gcttgggcta gagaaggatg tggttcagat atggttctgt  
3501 aaccggcgc agaagggcaa aagatcaagt attgagtatt cccaacgaga agagtatgag  
gctacagggg cacctttccc agggggggct gtatccttc  
3601 ctctgcccc aggtccccac ttggcacc caggctatgg aagccccac ttcaccacac  
tctactcagt ccttttctc gaggcgagg cctttccctc  
3701 tgttcccgtc actgctctgg gctctcccat gcattcaaac ggaagtggcg tgaacagac  
tttgaatttt gaccttctca agttggcggg agacgtggag  
3801 tccaaccag ggcccatggc tgcagcgc gctctgctcc cgtccttctc cagttcgcg  
tccggcccgg cgggaaggga gaagacactg cgtccagcag  
3901 gtgccccgac taaccgttgg cgtgaggaac tctctacat gaagcgactt cccccacttc  
cggcccgc ctacgacctg gcggcagcg tggccacaga  
4001 cctggagagt ggcggagctg gtgcagctt cagcagtaac aaccggccc tcttagcccg  
gagggagacc gaggagttca acgacctcct ggacctagac  
4101 ttatccttt ccaactcgtc aaccaccag gaatcgggtg ccgccaccgt gaccacctcg  
gcgtcagctt catcctcgtc ttcccagcg agcagcggcc  
4201 ctgcccagc gcctccacc tgcagcttca gctatccgat ccgggcccgg ggtgaccggg  
gcgtggctgc cagcaacaca ggtggagggc tctctacag  
4301 ccgagaatct gcgccacctc ccacggccc cttcaacctg gcggacatca atgacgtgag  
cccctcgggc ggcttctggt ctgagctcct gcggccggag  
4401 ttggaccag tataattcc gccacagcag cctcagccgc caggtggcgg gctgatgggg  
aagtttgtg tgaaggcgtc tctgaccacc cctggcagcg  
4501 agtacagcag cccttcggtc atcagtggtta gcaaaggaag ccagacggc agccaccccg  
tggtagtggc gccctacagc ggtggcccgc cgcgcatgtg  
4601 cccaagatt aagcaagagg cgttccgctc ctgcaggtc agccggctcc tagaggccca  
tttgagcgt ggaccccagc tcagcaacgg ccaccggccc  
4701 aacacacag acttccccct gggggggcag ctccccacca ggactacccc tacttgagt  
cccaggaac tgctgaacag cagggactgt caccctggcc  
4801 tgctcttcc ccaggatc catccccatc cggggcccaa ctaccctcct ttctgcccag  
accagatgca gtcacaagtc cctctctcc attatcaaga  
4901 gctcatgcca ccgggttct gctgcccaga ggagcccaag ccaagaggg gaagaaggtc  
gtggcccgg aaaagaacag ccaccacac ttgtgactat

5001 gcaggctgtg gcaaaaccta taccaagagt tctcatctca aggcacacct gcgaactcac  
acaggcgaga aaccttacca ctgtgactgg gacggctgtg  
5101 ggtggaaatt cgcccgtcc gatgaactga ccaggcacta ccgcaaacac acagggcacc  
ggccctttca gtgccagaag tgtgacaggg cttttccag  
5201 gtcggaccac cttgccttac acatgaagag gcacttttaa agatccctcc ccccccccta  
acgttactgg ccgaagccgc ttggaataag gccgggtgtgc  
5301 gttgtctat atgttatttt ccaccatatt gccgtctttt ggcaatgtga gggcccggaa  
acctggccct gtctttttga cgagcattcc taggggtctt  
5401 tcccctctcg ccaaaggaat gcaaggtctg ttgaatgtcg tgaaggaagc agttcctctg  
gaagcttctt gaagacaaac aacgtctgta gcgacccttt  
5501 gcaggcagcg gaacccccca cctggcgaca ggtgcctctg cggccaaaag ccacgtgtat  
aagatacacc tgcaaagggc gcacaacccc agtgccacgt  
5601 tgtgagttgg atagttgtgg aaagagtcaa atggctctcc tcaagcgtat tcaacaaggg  
gctgaaggat gcccagaagg taccatttgg tatgggatct  
5701 gatctggggc ctcgggtgcac atgctttaca tgtgtttagt cgagggttaa aaaacgtcta  
ggccccccga accacgggga cgtggttttc ctttgaaaa  
5801 cacgatgata atatggccac acatgatgat tataacatga tggagacgga gctgaagccg  
ccgggcccgc agcaagcttc gggggggcgc ggcggaggag  
5901 gcaacgccac ggcggcggc accggcgga accagaagaa cagcccggac cgcgtcaaga  
ggcccatgaa cgccttcatg gtatggctcc gggggcagc  
6001 gcgtaagatg gcccaggaga accccaagat gcacaactcg gagatcagca agcgcctggg  
cgcggagtgg aaacttttgt ccgagaccga gaagcggccg  
6101 ttcctcgacg aggccaagcg gctgcgcgct ctgcacatga aggagcacc ggattataaa  
taccggcccgc ggcggaaaaac caagacgctc atgaagaagg  
6201 ataagtacac gcttcccga ggttgcctgg ccccggcgg gaacagcatg gcgagcgggg  
ttggggtggg cgcgggctg ggtgcggcgc tgaaccagcg  
6301 catggacagc tacgcgcaca tgaacggctg gagcaacggc agctacagca tgatgcagga  
gcagctgggc taccgcagc acccggcct caacgctcac  
6401 ggcgcggcac agatgcaacc gatgcaccgc tacgacgtca gcgccctgca gtacaactcc  
atgaccagct cgcagacct catgaacggc tcgcccact  
6501 acagcatgct ctactcgag cagggcacc cggtatggc gctgggctcc atgggctctg  
tggtaagtc cgaggccagc tccagcccc cgtgggttac  
6601 ctctctctcc cactccaggg cgcctgcca ggcgggggac ctccgggaca tgatcagcat  
gtacctccc ggcgcggagg tgccggagcc cgtgcgccc  
6701 agtagactgc acatggcca gcaactacc agcggcccgg tgcccggcac ggccattaac  
ggcacactgc cctgtgcga catgggtagt gggcaatgta  
6801 ctaactacgc tttgttgaac ctgcgtggcg atgttgaag taaccccggt cctatgcccc  
tcaacgtgaa cttcaccia aggaactatg acctcgacta  
6901 cgactccgta cagccctatt tcatctgcga cgaggaagag aatttctatc accagcaaca  
gcagagcgag ctgcagccgc ccgcgccag tgaggatatac  
7001 tggaagaaat tcgagctgct tcccacccc ccctgtccc cgagccgccc ctccgggctc  
tgctctccat cctatgttgc ggtcgtacg tcttctccc  
7101 caaggaaga cgatgacggc ggcgggtggca acttctccac cgcgatcag ctggagatga  
tgaccgatt acttgagga gatgtgtga accagactt  
7201 catctgcgat cctgacgagc agaccttcat caagaacatc atcatccagg actgtatgtg  
gagcggtttc tcagccgctg ccaagctggt ctcgagaaag  
7301 ctggcctcct accaggctgc gcgcaaagac agcaccagcc tgagccccgc ccgcccggc  
agcgtctgct ccacctccag cctgtacctg caggacctca  
7401 ccgcccgcgc gtcggagtgc attgacctc cagtgtctt tccctaccgc ctcaacgaca  
gcagctgcc caaatcctgt acctcgtcc attccacggc  
7501 cttctctcct tctcggact cgtgctgtc ctccgagtcc tcccacggg ccagccctga  
gcccctagtg ctgcatgagg agacaccgcc caccaccagc  
7601 agcactctg aagaagagca agaagatgag gaagaaattg atgtgggtgc tgtggagaag  
aggcaaacc ctgccaagag gtcggagtgc ggtcatctc  
7701 catcccggagg ccacagcaaa cctccgcaca gccactggt cctcaagagg tgccacgtct  
cactcaca gcacaactac gccgacccc cctccacaag  
7801 gaagactat ccagctgcca agagggcca gttggacagt ggcagggtcc tgaagcagat  
cagcaacaac cgcaagtgtc ccagccccag gtcctcagac  
7901 acggaggaaa acgacaagag gcggacacac aacgtcttgg aacgtcagag gaggaacgag  
ctgaagcga gcttttttgc cctgctgac cagatccctg  
8001 aattggaaaa caacgaaaag gcccccaag tagtgatect caaaaaagcc accgcctaca  
tctgtctcat tcaagcagac gagcacaagc taacctctga  
8101 aaaggactta ttgagaaac gacgagaaca gttgaaacac aaactcgaac agcttcgaaa  
ctctggtgca taaatcgata gatcctaate aacctctgga

8201 ttacaaaatt tgtgaaagat tgactggat tcttaactat gttgctcctt ttacgctatg  
tggatacgt gctttaatgc ctttgtatca tgctattgct  
8301 tcccgtatgg ctttcatttt ctctccttg tataaatcct ggttgctgct tctttatgag  
gagttgtggc cggttgtcag gcaactggc gtgggtgca  
8401 ctgtgtttgc tgacgcaacc cccactgggt ggggcattgc caccacctgt cagctccttt  
cggggacttt cgctttcccc ctccctattg ccacggcgga  
8501 actcatcgcc gectgccttg cccgctgctg gacaggggct cggctggttg gcaactgaca  
ttccgtgggtg ttgtcgggga aatcatcgct ctttccttgg  
8601 ctgctcgctt gtgttgccac ctggattctg cgcgggacgt ccttctgcta cgtcccttgc  
gccctcaatc cagcggacct tccttccgc ggctgctgc  
8701 cggtctgctg gectcttccg cgtcttccgc ttccctca gacgagtcgg atctcccttt  
gggcccctc cccgctgggt accttaaga ccaatgactt  
8801 acaaggcagc tgtagatcct agccactttt taaaagaaaa ggggggactg gaagggctaa  
ttcactccca acgaagaca gatcacctgc aggacaggcg  
8901 cgcctgctt tttgcttga ctgggtctct ctggtagac cagatctgag cctgggagct  
ctctggctaa ctagggaacc cactgctaa gcctcaataa  
9001 agcttgcctt gagtgcttca agtagtggt gcccgtctgt tgtgtgactc tggtaactag  
agatccctca gaccctttta gtcagtgtgg aaaatcteta  
9101 gcaccgggc gattaaggaa agggctagat cattcttga gacgaaagg cctcgtgata  
cgcctatttt tataggtaa tgcctgata ataattggtt  
9201 cttagacgtc aggtggcact tttcgggaa atgtgcgctg aaccctatt tgtttatttt  
tctaaataca tcaaatatg tatccgctca tgagacaata  
9301 accctgataa atgcttcaat aatattgaaa aggaagagt atgagtattc aacatttccg  
tgtgcctt attccctttt ttgcggcatt ttgccttct  
9401 gttttgctc accagaaa gctgggtaa gtaaaagatg ctgaagatca gttgggtgca  
cgagtgggtt acatcgaact ggatctcaac agcggtaaga  
9501 tccttgagag ttttcgccc gaagaacgtt ttcaatgat gagcactttt aaagttctgc  
tatgtggcgc ggtattatcc cgtgttgacg cgggcaaga  
9601 gcaactcggg cgcgcatac actattctca gaatgacttg gttgagtact caccagtcac  
agaaaagcat cttacggatg gcatgacagt aagagaatta  
9701 tgcagtgtg ccataacct gagtgataac actgcggcca acttacttct gacaacgatc  
ggaggaccga aggagctaac cgctttttg cacaacatgg  
9801 gggatcatgt aactgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg  
acgagcgtga caccagatg cctgtagca tggcaacaac  
9901 gttgcgaaa ctattaactg gcgaactact tactctagct tcccggcaac aattaataga  
ctggatggag gcgataaag ttgcaggacc acttctgcgc  
10001 tcggcccttc cggctggctg gtttattgct gataaatctg gagccggatg gcgtgggtct  
cgcggtatca ttgcagcact gggccagat ggtaagcct  
10101 cccgtatcgt agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac  
agatcgtgta gataggtgcc tcaactgatta agcattggta  
10201 actgtcagac caagtttact catatatact ttagattgat taaaacttc atttttaatt  
taaaaggatc taggtgaaga tcctttttga taatctcatg  
10301 accaaaatcc cttaactgta gttttcgtc cactgagcgt cagacccgt agaaaagatc  
aaaggatctt cttagatcc tttttctg cgcgtaactt  
10401 gctgcttga acaaaaaaaaa ccaccgctac cagcgggtgg ttgtttgccg gatcaagagc  
taccaactct ttttcgaag gtaactggct tcagcagagc  
10501 gcagatacca aatactgtt tctagtgtg gccgtagtta ggccaccact tcaagaactc  
tgtagcaccg cctacatacc tcgctctgct aatcctgtta  
10601 ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg ggttgactc aagacgatag  
ttaccggata aggcgcagcg gtcgggctga acggggggtt  
10701 cgtgcacaca gccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg  
agctatgaga aagcgcacg ctcccgaag ggagaaaggc  
10801 ggacaggtat cggtaagcg gcagggtcgg aacaggagag cgcacgagg agcttccagg  
gggaaacgcc tggatcttt atagctctgt cgggttctgc  
10901 cacctctgac ttgagcgtc atttttgtg tgctcgtcag gggggcggag cctatggaaa  
aacgccaga acgggcctt ttacggctc ctggcctttt  
11001 cgtgccttt tgctcacatg ttcttctg cgttatcccc tgattctgtg gataaccgta  
ttaccgctt tgagttagct gataccgctc gccgagcgc  
11101 aacgaccgag cgcagcagc cagttagcga ggaagcggaa gagcgccaa tacgcaaacc  
gcctctccc gcgcgttggc cgattcatta atgcagcaag  
11201 ctcatggctg actaattttt tttatttatg cagaggccga ggccgctcgc gcctctgagc  
tattccagaa gtagttagga gcttttttg gaggcctagg  
11301 cttttgaaa aagctcccc tggcagcaca gtttccccga ctggaaagcg ggcagtgagc  
gcaacgcaat taatgtgagt tagctcactc attaggcacc

11401 ccaggcttta cactttatgc ttccggctcg tatgttggtg ggaattgtga gcggataaca  
atcacacaca ggaaacagct atgacatgat tacgaatttc  
11501 acaaataaag ctttttttc actgcattct agttgtggtt tgtccaaact catcaatgta  
tcttatcatg tctggatcaa ctggataact caagctaacc  
11601 aaaatcatcc caaacttccc accccatacc ctattaccac tgccaattac ctgtggtttc  
atttactcta aacctgtgat tcctctgaat tttttcatt  
11701 ttaaagaaat tgtatttggt aaatatgtac taaaactta gtagt

## Component Fragments