SEQUENCE LISTING

<110> Rocket Pharmaceuticals, LTD
Beard, Brian
Ricks, David
Prabhakar, Raj

<120> GENE THERAPY VECTORS FOR INFANTILE MALIGNANT OSTEOPETROSIS

<130> ROPA-003/01WO 329592-2098

<150> US 62/852,216
<151> 2019-05-23

<160> 39

<170> PatentIn version 3.5

<210> 1
<211> 3384
<212> DNA
<213> Artificial Sequence

<220>
<223> Made in Lab - Expression cassette

<400> 1
ggctccggtg cccgtcagtg ggcagagcgc acatcgccca cagtccccga gaagttgggg 60

ggaggggtcg gcaattgaac cggtgcctag agaaggtggc gcggggtaaa ctgggaaagt 120

gatgtcgtgt actggctccg cctttttccc gagggtgggg gagaaccgta tataagtgca 180

gtagtcgccg tgaacgttct ttttcgcaac gggtttgccg ccagaacaca ggtgtcgtga 240

cgcgggatcc gccaccatgg gctccatgtt tcggagcgag gaggtggccc tggtccagct 300

ctttctgccc acagcggctg cctacacctg cgtgagtcgg ctgggcgagc tgggcctcgt 360

ggagttcaga gacctcaacg cctcggtgag cgccttccag agacgctttg tggttgatgt 420

tcggcgctgt gaggagctgg agaagacctt caccttcctg caggaggagg tgcggcgggc 480

tgggctggtc ctgcccccgc caaaggggag gctgccggca cccccacccc gggacctgct 540

gcgcatccag gaggagacgg agcgcctggc ccaggagctg cgggatgtgc ggggcaacca 600

gcaggccctg cgggcccagc tgcaccagct gcagctccac gccgccgtgc tacgccaggg 660

ccatgaacct cagctggcag ccgcccacac agatggggcc tcagagagga cgcccctgct 720

ccaggccccc ggggggccgc accaggacct gagggtcaac tttgtggcag gtgccgtgga 780

gccccacaag gcccctgccc tagagcgcct gctctggagg gcctgcagag gcttcctcat 840

tgccagcttc agggagctgg agcagccgct ggagcacccc gtgacgggcg agccagccac 900

gtggatgacc ttcctcatct cctactgggg tgagcagatc ggacagaaga tccgcaagat 960

cacggactgc ttccactgcc acgtcttccc gtttctgcag caggaggagg cccgcctcgg 1020

ggccctgcag cagctgcaac agcagagcca ggagctgcag gaggtcctcg gggagacaga 1080

gcggttcctg agccaggtgc taggccgggt gctgcagctg ctgccgccag ggcaggtgca 1140

ggtccacaag atgaaggccg tgtacctggc cctgaaccag tgcagcgtga gcaccacgca 1200

caagtgcctc attgccgagg cctggtgctc tgtgcgagac ctgcccgccc tgcaggaggc 1260

cctgcgggac agctcgatgg aggagggagt gagtgccgtg gctcaccgca tcccctgccg 1320

ggacatgccc cccacactca tccgcaccaa ccgcttcacg gccagcttcc agggcatcgt 1380

ggatgcctac ggcgtgggcc gctaccagga ggtcaacccc gctccctaca ccatcatcac 1440

cttccccttc ctgtttgctg tgatgttcgg ggatgtgggc cacgggctgc tcatgttcct 1500

cttcgccctg gccatggtcc ttgcggagaa ccgaccggct gtgaaggccg cgcagaacga 1560

gatctggcag actttcttca ggggccgcta cctgctcctg cttatgggcc tgttctccat 1620

ctacaccggc ttcatctaca acgagtgctt cagtcgcgcc accagcatct tcccctcggg 1680

ctggagtgtg gccgccatgg ccaaccagtc tggctggagt gatgcattcc tggcccagca 1740

cacgatgctt accctggacc ccaacgtcac cggtgtcttc ctgggaccct acccctttgg 1800

catcgatcct atttggagcc tggctgccaa ccacttgagc ttcctcaact ccttcaagat 1860

gaagatgtcc gtcatcctgg gcgtcgtgca catggccttt ggggtggtcc tcggagtctt 1920

caaccacgtg cactttggcc agaggcaccg gctgctgctg gagacgctgc cggagctcac 1980

cttcctgctg ggactcttcg gttacctcgt gttcctagtc atctacaagt ggctgtgtgt 2040

ctgggctgcc agggccgcct cggcccccag catcctcatc cacttcatca acatgttcct 2100

cttctcccac agccccagca acaggctgct ctacccccgg caggaggtgg tccaggccac 2160

gctggtggtc ctggccttgg ccatggtgcc catcctgctg cttggcacac ccctgcacct 2220

gctgcaccgc caccgccgcc gcctgcggag gaggcccgct gaccgacagg aggaaaacaa 2280

ggccgggttg ctggacctgc ctgacgcatc tgtgaatggc tggagctccg atgaggaaaa 2340

ggcaggggggc ctggatgatg aagaggaggc cgagctcgtc ccctccgagg tgctcatgca 2400

ccaggccatc cacaccatcg agttctgcct gggctgcgtc tccaacaccg cctcctacct 2460

gcgcctgtgg gccctgagcc tggcccacgc ccagctgtcc gaggttctgt gggccatggt 2520

gatgcgcata ggcctgggcc tgggccggga ggtgggcgtg gcggctgtgg tgctggtccc 2580

catctttgcc gcctttgccg tgatgaccgt ggctatcctg ctggtgatgg agggactctc 2640

agccttcctg cacgccctgc ggctgcactg ggtggaattc cagaacaagt tctactcagg 2700

cacgggctac aagctgagtc ccttcacctt cgctgccaca gatgactagt aagtcgacgg 2760

atcccccggg ctgcaggaat tcgagcatct taccgccatt tatacccata tttgttctgt 2820

ttttcttgat ttgggtatac atttaaatgt taataaaaca aaatggtggg gcaatcattt 2880

acatttttag ggatatgtaa ttactagttc aggtgtattg ccacaagaca aacatgttaa 2940

gaaactttcc cgttatttac gctctgttcc tgttaatcaa cctctggatt acaaaatttg 3000

tgaaagattg actgatattc ttaactatgt tgctcctttt acgctgtgtg gatatgctgc 3060

tttaatgcct ctgtatcatg ctattgcttc ccgtacggct ttcgttttct cctccttgta 3120

taaatcctgg ttgctgtctc tttatgagga gttgtggccc gttgtccgtc aacgtggcgt 3180

ggtgtgctct gtgtttgctg acgcaacccc cactggctgg ggcattgcca ccacctgtca 3240

actcctttct gggactttcg ctttcccccc cccgatcgcc acggcagaac tcatcgccgc 3300

ctgccttgcc cgctgctgga caggggctag gttgctgggc actgataatt ccgtggtgtt 3360

gtcggggaag ctgacgtcct ttcg 3384

<210> 2
<211> 243
<212> DNA
<213> Homo sapiens

<400> 2
ggctccggtg cccgtcagtg ggcagagcgc acatcgccca cagtccccga gaagttgggg 60

ggaggggtcg gcaattgaac cggtgcctag agaaggtggc gcggggtaaa ctgggaaagt 120

gatgtcgtgt actggctccg cctttttccc gagggtgggg gagaaccgta tataagtgca 180

gtagtcgccg tgaacgttct ttttcgcaac gggtttgccg ccagaacaca ggtgtcgtga 240

cgc 243

<210> 3
<211> 2493
<212> DNA

<213> Homo sapiens

<400> 3

atgggctcca tgtttcggag cgaggaggtg gccctggtcc agctctttct gcccacagcg 60

gctgcctaca cctgcgtgag tcggctgggc gagctgggcc tcgtggagtt cagagacctc 120

aacgcctcgg tgagcgcctt ccagagacgc tttgtggttg atgttcggcg ctgtgaggag 180

ctggagaaga ccttcacctt cctgcaggag gaggtgcggc gggctgggct ggtcctgccc 240

ccgccaaagg ggaggctgcc ggcaccccca ccccgggacc tgctgcgcat ccaggaggag 300

acggagcgcc tggcccagga gctgcgggat gtgcggggca accagcaggc cctgcgggcc 360

cagctgcacc agctgcagct ccacgccgcc gtgctacgcc agggccatga acctcagctg 420

gcagccgccc acacagatgg ggcctcagag aggacgcccc tgctccaggc ccccgggggg 480

ccgcaccagg acctgagggt caactttgtg gcaggtgccg tggagcccca caaggcccct 540

gccctagagc gcctgctctg gagggcctgc agaggcttcc tcattgccag cttcagggag 600

ctggagcagc cgctggagca ccccgtgacg ggcgagccag ccacgtggat gaccttcctc 660

atctcctact ggggtgagca gatcggacag aagatccgca agatcacgga ctgcttccac 720

tgccacgtct tcccgtttct gcagcaggag gaggcccgcc tcggggccct gcagcagctg 780

caacagcaga gccaggagct gcaggaggtc ctcggggaga cagagcggtt cctgagccag 840

gtgctaggcc gggtgctgca gctgctgccg ccagggcagg tgcaggtcca caagatgaag 900

gccgtgtacc tggccctgaa ccagtgcagc gtgagcacca cgcacaagtg cctcattgcc 960

gaggcctggt gctctgtgcg agacctgccc gccctgcagg aggccctgcg ggacagctcg 1020

atggaggagg gagtgagtgc cgtggctcac cgcatcccct gccgggacat gccccccaca 1080

ctcatccgca ccaaccgctt cacggccagc ttccagggca tcgtggatgc ctacggcgtg 1140

ggccgctacc aggaggtcaa ccccgctccc tacaccatca tcaccttccc cttcctgttt 1200

gctgtgatgt tcggggatgt gggccacggg ctgctcatgt tcctcttcgc cctggccatg 1260

gtccttgcgg agaaccgacc ggctgtgaag gccgcgcaga acgagatctg gcagactttc 1320

ttcaggggcc gctacctgct cctgcttatg ggcctgttct ccatctacac cggcttcatc 1380

tacaacgagt gcttcagtcg cgccaccagc atcttcccct cgggctggag tgtggccgcc 1440

atggccaacc agtctggctg gagtgatgca ttcctggccc agcacacgat gcttaccctg 1500

gaccccaacg tcaccggtgt cttcctggga ccctaccct ttggcatcga tcctatttgg 1560

agcctggctg ccaaccactt gagcttcctc aactccttca agatgaagat gtccgtcatc 1620

ctgggcgtcg tgcacatggc ctttggggtg gtcctcggag tcttcaacca cgtgcacttt 1680

ggccagaggc accggctgct gctggagacg ctgccggagc tcaccttcct gctgggactc 1740

ttcggttacc tcgtgttcct agtcatctac aagtggctgt gtgtctgggc tgccagggcc 1800

gcctcggccc ccagcatcct catccacttc atcaacatgt tcctcttctc ccacagcccc 1860

agcaacaggc tgctctaccc ccggcaggag gtggtccagg ccacgctggt ggtcctggcc 1920

ttggccatgg tgcccatcct gctgcttggc acacccctgc acctgctgca ccgccaccgc 1980

cgccgcctgc ggaggaggcc cgctgaccga caggaggaaa acaaggccgg gttgctggac 2040

ctgcctgacg catctgtgaa tggctggagc tccgatgagg aaaaggcagg gggcctggat 2100

gatgaagagg aggccgagct cgtcccctcc gaggtgctca tgcaccaggc catccacacc 2160

atcgagttct gcctgggctg cgtctccaac accgcctcct acctgcgcct gtgggccctg 2220

agcctggccc acgcccagct gtccgaggtt ctgtgggcca tggtgatgcg cataggcctg 2280

ggcctgggcc gggaggtggg cgtggcggct gtggtgctgg tccccatctt tgccgccttt 2340

gccgtgatga ccgtggctat cctgctggtg atggagggac tctcagcctt cctgcacgcc 2400

ctgcggctgc actgggtgga attccagaac aagttctact caggcacggg ctacaagctg 2460

agtcccttca ccttcgctgc cacagatgac tag 2493

<210> 4
<211> 606
<212> DNA
<213> Woodchuck hepatitis virus

<400> 4
attcgagcat cttaccgcca tttataccca tatttgttct gtttttcttg atttgggtat 60

acatttaaat gttaataaaa caaaatggtg gggcaatcat ttacattttt agggatatgt 120

aattactagt tcaggtgtat tgccacaaga caaacatgtt aagaaacttt cccgttattt 180

acgctctgtt cctgttaatc aacctctgga ttacaaaatt tgtgaaagat tgactgatat 240

tcttaactat gttgctcctt ttacgctgtg tggatatgct gctttaatgc ctctgtatca 300

tgctattgct tcccgtacgg ctttcgtttt ctcctccttg tataaatcct ggttgctgtc 360

tctttatgag gagttgtggc ccgttgtccg tcaacgtggc gtggtgtgct ctgtgtttgc 420

tgacgcaacc cccactggct ggggcattgc caccacctgt caactccttt ctgggacttt 480

cgctttcccc ctcccgatcg ccacggcaga actcatcgcc gcctgccttg cccgctgctg 540

gacaggggct aggttgctgg gcactgataa ttccgtggtg ttgtcgggga agctgacgtc 600

ctttcg 606

<210> 5
<211> 830
<212> PRT
<213> Homo sapiens

<400> 5

Met Gly Ser Met Phe Arg Ser Glu Glu Val Ala Leu Val Gln Leu Phe
1 5 10 15

Leu Pro Thr Ala Ala Ala Tyr Thr Cys Val Ser Arg Leu Gly Glu Leu
20 25 30

Gly Leu Val Glu Phe Arg Asp Leu Asn Ala Ser Val Ser Ala Phe Gln
35 40 45

Arg Arg Phe Val Val Asp Val Arg Arg Cys Glu Glu Leu Glu Lys Thr
50 55 60

Phe Thr Phe Leu Gln Glu Glu Val Arg Arg Ala Gly Leu Val Leu Pro
65 70 75 80

Pro Pro Lys Gly Arg Leu Pro Ala Pro Pro Pro Arg Asp Leu Leu Arg
85 90 95

Ile Gln Glu Glu Thr Glu Arg Leu Ala Gln Glu Leu Arg Asp Val Arg
100 105 110

Gly Asn Gln Gln Ala Leu Arg Ala Gln Leu His Gln Leu Gln Leu His
115 120 125

Ala Ala Val Leu Arg Gln Gly His Glu Pro Gln Leu Ala Ala Ala His
130 135 140

Thr Asp Gly Ala Ser Glu Arg Thr Pro Leu Leu Gln Ala Pro Gly Gly
145 150 155 160

```
Pro His Gln Asp Leu Arg Val Asn Phe Val Ala Gly Ala Val Glu Pro
               165                 170                 175

His Lys Ala Pro Ala Leu Glu Arg Leu Leu Trp Arg Ala Cys Arg Gly
            180                 185                 190

Phe Leu Ile Ala Ser Phe Arg Glu Leu Glu Gln Pro Leu Glu His Pro
        195                 200                 205

Val Thr Gly Glu Pro Ala Thr Trp Met Thr Phe Leu Ile Ser Tyr Trp
    210                 215                 220

Gly Glu Gln Ile Gly Gln Lys Ile Arg Lys Ile Thr Asp Cys Phe His
225                 230                 235                 240

Cys His Val Phe Pro Phe Leu Gln Gln Glu Glu Ala Arg Leu Gly Ala
               245                 250                 255

Leu Gln Gln Leu Gln Gln Gln Ser Gln Glu Leu Gln Glu Val Leu Gly
            260                 265                 270

Glu Thr Glu Arg Phe Leu Ser Gln Val Leu Gly Arg Val Leu Gln Leu
        275                 280                 285

Leu Pro Pro Gly Gln Val Gln Val His Lys Met Lys Ala Val Tyr Leu
    290                 295                 300

Ala Leu Asn Gln Cys Ser Val Ser Thr Thr His Lys Cys Leu Ile Ala
305                 310                 315                 320

Glu Ala Trp Cys Ser Val Arg Asp Leu Pro Ala Leu Gln Glu Ala Leu
               325                 330                 335

Arg Asp Ser Ser Met Glu Glu Gly Val Ser Ala Val Ala His Arg Ile
            340                 345                 350

Pro Cys Arg Asp Met Pro Pro Thr Leu Ile Arg Thr Asn Arg Phe Thr
        355                 360                 365
```

```
Ala Ser Phe Gln Gly Ile Val Asp Ala Tyr Gly Val Gly Arg Tyr Gln
    370                 375                 380

Glu Val Asn Pro Ala Pro Tyr Thr Ile Ile Thr Phe Pro Phe Leu Phe
385                 390                 395                 400

Ala Val Met Phe Gly Asp Val Gly His Gly Leu Leu Met Phe Leu Phe
                405                 410                 415

Ala Leu Ala Met Val Leu Ala Glu Asn Arg Pro Ala Val Lys Ala Ala
            420                 425                 430

Gln Asn Glu Ile Trp Gln Thr Phe Phe Arg Gly Arg Tyr Leu Leu Leu
        435                 440                 445

Leu Met Gly Leu Phe Ser Ile Tyr Thr Gly Phe Ile Tyr Asn Glu Cys
    450                 455                 460

Phe Ser Arg Ala Thr Ser Ile Phe Pro Ser Gly Trp Ser Val Ala Ala
465                 470                 475                 480

Met Ala Asn Gln Ser Gly Trp Ser Asp Ala Phe Leu Ala Gln His Thr
                485                 490                 495

Met Leu Thr Leu Asp Pro Asn Val Thr Gly Val Phe Leu Gly Pro Tyr
            500                 505                 510

Pro Phe Gly Ile Asp Pro Ile Trp Ser Leu Ala Ala Asn His Leu Ser
        515                 520                 525

Phe Leu Asn Ser Phe Lys Met Lys Met Ser Val Ile Leu Gly Val Val
    530                 535                 540

His Met Ala Phe Gly Val Val Leu Gly Val Phe Asn His Val His Phe
545                 550                 555                 560

Gly Gln Arg His Arg Leu Leu Leu Glu Thr Leu Pro Glu Leu Thr Phe
                565                 570                 575

Leu Leu Gly Leu Phe Gly Tyr Leu Val Phe Leu Val Ile Tyr Lys Trp
            580                 585                 590
```

```
Leu Cys Val Trp Ala Ala Arg Ala Ala Ser Ala Pro Ser Ile Leu Ile
        595                 600                 605

His Phe Ile Asn Met Phe Leu Phe Ser His Ser Pro Ser Asn Arg Leu
    610                 615                 620

Leu Tyr Pro Arg Gln Glu Val Val Gln Ala Thr Leu Val Val Leu Ala
625                 630                 635                 640

Leu Ala Met Val Pro Ile Leu Leu Leu Gly Thr Pro Leu His Leu Leu
                645                 650                 655

His Arg His Arg Arg Arg Leu Arg Arg Arg Pro Ala Asp Arg Gln Glu
            660                 665                 670

Glu Asn Lys Ala Gly Leu Leu Asp Leu Pro Asp Ala Ser Val Asn Gly
        675                 680                 685

Trp Ser Ser Asp Glu Glu Lys Ala Gly Gly Leu Asp Asp Glu Glu Glu
    690                 695                 700

Ala Glu Leu Val Pro Ser Glu Val Leu Met His Gln Ala Ile His Thr
705                 710                 715                 720

Ile Glu Phe Cys Leu Gly Cys Val Ser Asn Thr Ala Ser Tyr Leu Arg
                725                 730                 735

Leu Trp Ala Leu Ser Leu Ala His Ala Gln Leu Ser Glu Val Leu Trp
            740                 745                 750

Ala Met Val Met Arg Ile Gly Leu Gly Leu Gly Arg Glu Val Gly Val
        755                 760                 765

Ala Ala Val Val Leu Val Pro Ile Phe Ala Ala Phe Ala Val Met Thr
    770                 775                 780

Val Ala Ile Leu Leu Val Met Glu Gly Leu Ser Ala Phe Leu His Ala
785                 790                 795                 800
```

Leu Arg Leu His Trp Val Glu Phe Gln Asn Lys Phe Tyr Ser Gly Thr
805 810 815

Gly Tyr Lys Leu Ser Pro Phe Thr Phe Ala Ala Thr Asp Asp
820 825 830

<210> 6
<211> 13
<212> DNA
<213> Artificial Sequence

<220>
<223> Consensus Kozak sequence

<400> 6
gccgccacca tgg 13

<210> 7
<211> 387
<212> DNA
<213> Mus musculus

<400> 7
tggctaataa aggaaattta ttttcattgc aatagtgtgt tggaattttt tgtgtctctc 60

actcggaagg acatatggga gggcaaatca tttaaaacat cagaatgagt atttggttta 120

gagtttggca acatatgccc atatgctggc tgccatgaac aaaggttggc tataaagagg 180

tcatcagtat atgaaacagc cccctgctgt ccattcctta ttccatagaa aagccttgac 240

ttgaggttag attttttta tattttgttt tgtgttattt ttttctttaa catccctaaa 300

attttcctta catgttttac tagccagatt tttcctcctc tcctgactac tcccagtcat 360

agctgtccct cttctcttat ggagatc 387

<210> 8

<400> 8
000

<210> 9

<400> 9
000

<210> 10

<400> 10

000

<210> 11

<400> 11
000

<210> 12

<400> 12
000

<210> 13

<400> 13
000

<210> 14
<211> 13
<212> RNA
<213> Artificial Sequence

<220>
<223> Consensus Kozak sequence

<400> 14
gccgccrcca ugg 13

<210> 15
<211> 8
<212> RNA
<213> Artificial Sequence

<220>
<223> Kozak sequence

<220>
<221> misc_feature
<222> (2)..(3)
<223> n is A, C, T, G or U

<220>
<221> misc_feature
<222> (4)..(4)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (8)..(8)
<223> n is A, C, T, G or U

<400> 15

agnnaugn 8

<210> 16
<211> 7
<212> RNA
<213> Artificial Sequence

<220>
<223> Kozak sequence

<220>
<221> misc_feature
<222> (2)..(3)
<223> n is A, C, T, G or U

<400> 16
annaugg 7

<210> 17
<211> 7
<212> RNA
<213> Artificial Sequence

<220>
<223> Kozak sequence

<400> 17
accaugg 7

<210> 18
<211> 10
<212> RNA
<213> Artificial Sequence

<220>
<223> n is A, C, T, G or U

<400> 18
gacaccaugg 10

<210> 19
<211> 235
<212> DNA
<213> Bos taurus

<400> 19
tcgactgtgc cttctagttg ccagccatct gttgtttgcc cctcccccgt gccttccttg 60

accctggaag gtgccactcc cactgtcctt tcctaataaa atgaggaaat tgcatcgcat 120

tgtctgagta ggtgtcattc tattctgggg ggtggggtgg ggcaggacag caagggggag 180

gattgggagg acaatagcag gcatgctggg gatgcggtgg gctctatggc ttctg 235

<210> 20
<211> 222
<212> DNA
<213> simian virus 40

<400> 20
cagacatgat aagatacatt gatgagtttg gacaaaccac aactagaatg cagtgaaaaa 60

aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt ataagctgca 120

ataaacaagt taacaacaac aattgcattc attttatgtt tcaggttcag ggggagatgt 180

gggaggtttt ttaaagcaag taaaacctct acaaatgtgg ta 222

<210> 21
<211> 202
<212> DNA
<213> Homo sapiens

<400> 21
ctgcccgggt ggcatccctg tgacccctcc ccagtgcctc tcctggccct ggaagttgcc 60

actccagtgc ccaccagcct tgtcctaata aaattaagtt gcatcatttt gtctgactag 120

gtgtccttct ataatattat ggggtggagg ggggtggtat ggagcaaggg gcccaagttg 180

ggaagaaacc tgtagggcct gc 202

<210> 22
<211> 141
<212> DNA
<213> Escherichia coli

<400> 22
gtagaattgg taaagagagt cgtgtaaaat atcgagttcg cacatcttgt tgtctgatta 60

ttgatttttg gcgaaaccat ttgatcatat gacaagatgt gtatctacct taacttaatg 120

attttgataa aaatcattag g 141

<210> 23
<211> 7660
<212> DNA
<213> Artificial Sequence

<220>
<223> Made in Lab - plasmid construct

<400> 23
acattgatta ttgactagtt attaatagta atcaattacg gggtcattag ttcatagccc 60

atatatggag ttccgcgtta cataacttac ggtaaatggc ccgcctggct gaccgcccaa 120

cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc caatagggac 180

tttccattga cgtcaatggg tggagtattt acggtaaact gcccacttgg cagtacatca 240

agtgtatcat atgccaagta cgccccctat tgacgtcaat gacggtaaat ggcccgcctg 300

gcattatgcc cagtacatga ccttatggga ctttcctact tggcagtaca tctacgtatt 360

agtcatcgct attaccatgg tgatgcggtt ttggcagtac atcaatgggc gtggatagcg 420

gtttgactca cggggatttc caagtctcca ccccattgac gtcaatggga gtttgttttg 480

gcaccaaaat caacgggact ttccaaaatg tcgtaacaac tccgccccat tgacgcaaat 540

gggcggtagg cgtgtacggt gggaggtcta tataagcaga gctcgtttag tgaaccgggg 600

tctctctggt tagaccagat ctgagcctgg gagctctctg gctaactagg gaacccactg 660

cttaagcctc aataaagctt gccttgagtg cttcaagtag tgtgtgcccg tctgttgtgt 720

gactctggta actagagatc cctcagaccc ttttagtcag tgtggaaaat ctctagcagt 780

ggcgcccgaa cagggacttg aaagcgaaag ggaaaccaga ggagctctct cgacgcagga 840

ctcggcttgc tgaagcgcgc acggcaagag gcgaggggcg gcgactggtg agtacgccaa 900

aaattttgac tagcggaggc tagaaggaga gagatgggtg cgagagcgtc agtattaagc 960

gggggagaat tagatcgcga tgggaaaaaa ttcggttaag gccaggggga aagaaaaaat 1020

ataaattaaa acatatagta tgggcaagca gggagctaga acgattcgca gttaatcctg 1080

gcctgttaga aacatcagaa ggctgtagac aaatactggg acagctacaa ccatcccttc 1140

agacaggatc agaagaactt agatcattat ataatacagt agcaaccctc tattgtgtgc 1200

atcaaaggat agagataaaa gacaccaagg aagctttaga caagatagag gaagagcaaa 1260

acaaaagtaa gaccaccgca cagcaagcgg ccgctgatct tcagacctgg aggaggagat 1320

atgagggaca attggagaag tgaattatat aaatataaag tagtaaaaat tgaaccatta 1380

ggagtagcac ccaccaaggc aaagagaaga gtggtgcaga gagaaaaaag agcagtggga 1440

ataggagctt tgttccttgg gttcttggga gcagcaggaa gcactatggg cgcagcgtca 1500

atgacgctga cggtacaggc cagacaatta ttgtctggta tagtgcagca gcagaacaat 1560

ttgctgaggg ctattgaggc gcaacagcat ctgttgcaac tcacagtctg gggcatcaag 1620

cagctccagg caagaatcct ggctgtggaa agatacctaa aggatcaaca gctcctgggg 1680

atttggggtt gctctggaaa actcatttgc accactgctg tgccttggaa tgctagttgg 1740

agtaataaat ctctggaaca gatttggaat cacacgacct ggatggagtg ggacagagaa 1800

attaacaatt acacaagctt aatacactcc ttaattgaag aatcgcaaaa ccagcaagaa 1860

aagaatgaac aagaattatt ggaattagat aaatgggcaa gtttgtggaa ttggtttaac 1920

ataacaaatt ggctgtggta tataaaatta ttcataatga tagtaggagg cttggtaggt 1980

ttaagaatag tttttgctgt actttctata gtgaatagag ttaggcaggg atattcacca 2040

ttatcgtttc agacccacct cccaaccccg aggggacccg acaggcccga aggaatagaa 2100

gaagaaggtg gagagagaga cagagacaga tccattcgat tagtgaacgg atctcgacgg 2160

tatcggttaa cttttaaaag aaaaggggggg attgggggggt acagtgcagg ggaaagaata 2220

gtagacataa tagcaacaga catacaaact aaagaattac aaaaacaaat tacaaaaatt 2280

caaaatttta tcgatcacga gactagcctc gagaagcttg atcgattggc tccggtgccc 2340

gtcagtgggc agagcgcaca tcgcccacag tccccgagaa gttgggggga ggggtcggca 2400

attgaaccgg tgcctagaga aggtggcgcg gggtaaactg ggaaagtgat gtcgtgtact 2460

ggctccgcct ttttcccgag ggtgggggag aaccgtatat aagtgcagta gtcgccgtga 2520

acgttctttt tcgcaacggg tttgccgcca gaacacaggt gtcgtgacgc gggatccgcc 2580

accatgggct ccatgtttcg gagcgaggag gtggccctgg tccagctctt tctgcccaca 2640

gcggctgcct acacctgcgt gagtcggctg ggcgagctgg gcctcgtgga gttcagagac 2700

ctcaacgcct cggtgagcgc cttccagaga cgctttgtgg ttgatgttcg gcgctgtgag 2760

gagctggaga agaccttcac cttcctgcag gaggaggtgc ggcgggctgg gctggtcctg 2820

cccccgccaa aggggaggct gccggcaccc ccaccccggg acctgctgcg catccaggag 2880

gagacggagc gcctggccca ggagctgcgg gatgtgcggg gcaaccagca ggccctgcgg 2940

gcccagctgc accagctgca gctccacgcc gccgtgctac gccagggcca tgaacctcag 3000

ctggcagccg cccacacaga tggggcctca gagaggacgc ccctgctcca ggccccgggg 3060

gggccgcacc aggacctgag ggtcaacttt gtggcaggtg ccgtggagcc ccacaaggcc 3120

cctgccctag agcgcctgct ctggagggcc tgcagaggct tcctcattgc cagcttcagg 3180

gagctggagc agccgctgga gcaccccgtg acgggcgagc cagccacgtg gatgaccttc 3240

ctcatctcct actggggtga gcagatcgga cagaagatcc gcaagatcac ggactgcttc 3300

cactgccacg tcttcccgtt tctgcagcag gaggaggccc gcctcggggc cctgcagcag 3360

ctgcaacagc agagccagga gctgcaggag gtcctcgggg agacagagcg gttcctgagc 3420

caggtgctag gccgggtgct gcagctgctg ccgccagggc aggtgcaggt ccacaagatg 3480

aaggccgtgt acctggccct gaaccagtgc agcgtgagca ccacgcacaa gtgcctcatt 3540

gccgaggcct ggtgctctgt gcgagacctg cccgccctgc aggaggccct gcgggacagc 3600

tcgatggagg agggagtgag tgccgtggct caccgcatcc cctgccggga catgcccccc 3660

acactcatcc gcaccaaccg cttcacggcc agcttccagg gcatcgtgga tgcctacggc 3720

gtgggccgct accaggaggt caaccccgct ccctacacca tcatcacctt ccccttcctg 3780

tttgctgtga tgttcgggga tgtgggccac gggctgctca tgttcctctt cgccctggcc 3840

atggtccttg cggagaaccg accggctgtg aaggccgcgc agaacgagat ctggcagact 3900

ttcttcaggg gccgctacct gctcctgctt atgggcctgt tctccatcta caccggcttc 3960

atctacaacg agtgcttcag tcgcgccacc agcatcttcc cctcgggctg gagtgtggcc 4020

gccatggcca accagtctgg ctggagtgat gcattcctgg cccagcacac gatgcttacc 4080

ctggacccca acgtcaccgg tgtcttcctg ggaccctacc cctttggcat cgatcctatt 4140

tggagcctgg ctgccaacca cttgagcttc ctcaactcct tcaagatgaa gatgtccgtc 4200

atcctgggcg tcgtgcacat ggcctttggg gtggtcctcg gagtcttcaa ccacgtgcac 4260

tttggccaga ggcaccggct gctgctggag acgctgccgg agctcacctt cctgctggga 4320

ctcttcggtt acctcgtgtt cctagtcatc tacaagtggc tgtgtgtctg ggctgccagg 4380

gccgcctcgg cccccagcat cctcatccac ttcatcaaca tgttcctctt ctcccacagc 4440

cccagcaaca ggctgctcta cccccggcag gaggtggtcc aggccacgct ggtggtcctg 4500

gccttggcca tggtgcccat cctgctgctt ggcacacccc tgcacctgct gcaccgccac 4560

cgccgccgcc tgcggaggag gcccgctgac cgacaggagg aaaacaaggc cgggttgctg 4620

gacctgcctg acgcatctgt gaatggctgg agctccgatg aggaaaaggc agggggcctg 4680

gatgatgaag aggaggccga gctcgtcccc tccgaggtgc tcatgcacca ggccatccac 4740

accatcgagt tctgcctggg ctgcgtctcc aacaccgcct cctacctgcg cctgtgggcc 4800

ctgagcctgg cccacgccca gctgtccgag gttctgtggg ccatggtgat gcgcataggc 4860

ctgggcctgg gccgggaggt gggcgtggcg gctgtggtgc tggtccccat ctttgccgcc 4920

tttgccgtga tgaccgtggc tatcctgctg gtgatggagg gactctcagc cttcctgcac 4980

gccctgcggc tgcactgggt ggaattccag aacaagttct actcaggcac gggctacaag 5040

ctgagtccct tcaccttcgc tgccacagat gactagtaag tcgacggatc ccccgggctg 5100

caggaattcg agcatcttac cgccatttat acccatattt gttctgtttt tcttgatttg 5160

ggtatacatt taaatgttaa taaaacaaaa tggtggggca atcatttaca tttttaggga 5220

tatgtaatta ctagttcagg tgtattgcca caagacaaac atgttaagaa actttcccgt 5280

tatttacgct ctgttcctgt taatcaacct ctggattaca aaatttgtga aagattgact 5340

gatattctta actatgttgc tccttttacg ctgtgtggat atgctgcttt aatgcctctg 5400

tatcatgcta ttgcttcccg tacggctttc gttttctcct ccttgtataa atcctggttg 5460

ctgtctcttt atgaggagtt gtggcccgtt gtccgtcaac gtggcgtggt gtgctctgtg 5520

tttgctgacg caacccccac tggctggggc attgccacca cctgtcaact cctttctggg 5580

actttcgctt tccccctccc gatcgccacg gcagaactca tcgccgcctg ccttgcccgc 5640

tgctggacag gggctaggtt gctgggcact gataattccg tggtgttgtc ggggaagctg 5700

acgtcctttc gaattcgata tcaagctgta cctttaagac caatgactta caaggcagct 5760

gtagatctta gccacttttt aaaagaaaag gggggactgg aagggctaat tcactcccaa 5820

cgaagacaag atctgctttt tgcttgtact gggtctctct ggttagacca gatctgagcc 5880

tgggagctct ctggctaact agggaaccca ctgcttaagc ctcaataaag cttgccttga 5940

gtgcttcaag tagtgtgtgc ccgtctgttg tgtgactctg gtaactagag atccctcaga 6000

cccttttagt cagtgtggaa aatctctagc agtagtagtt catgtcatct tattattcag 6060

tatttataac ttgcaaagaa atgaatatca gagagtgaga ggaacttgtt tattgcagct 6120

tataatggtt acaaataaag caatagcatc acaaatttca caaataaagc attttttttca 6180

ctgcattcta gttgtggttt gtccaaactc atcaatgtat cttatcatgt ctggctctag 6240

ctatcccgcc cctaactccg cccatcccgc ccctaactcc gcccagttcc gcccattctc 6300

cgccccatgg ctgactaatt ttttttattt atgcagaggc cgaggccgcc tcggcctctg 6360

agctattcca gaagtagtga ggaggctttt ttggaggcct aggtagcccg cctaatgagc 6420

gggctttttt ttcttaggcc ttcttccgct tcctcgctca ctgactcgct gcgctcggtc 6480

gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg taatacggtt atccacagaa 6540

tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt 6600

aaaaaggccg cgttgctggc gtttttccat aggctccgcc cccctgacga gcatcacaaa 6660

aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt 6720

cccccctggaa gctccctcgt gcgctctcct gttccgaccc tgccgcttac cggatacctg 6780

tccgcctttc tcccttcggg aagcgtggcg ctttctcata gctcacgctg taggtatctc 6840

agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc 6900

gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtaag acacgactta 6960

tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt aggcggtgct 7020

acagagttct tgaagtggtg gcctaactac ggctacacta gaagaacagt atttggtatc 7080

tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa 7140

caaaccaccg ctggtagcgg tggtttttt gtttgcaagc agcagattac gcgcagaaaa 7200

aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa 7260

aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt 7320

ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac 7380

agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcatcc 7440

atagttgcct gactcctgca aaccacgttg tggtagaatt ggtaaagaga gtcgtgtaaa 7500

atatcgagtt cgcacatctt gttgtctgat tattgatttt tggcgaaacc atttgatcat 7560

atgacaagat gtgtatctac cttaacttaa tgattttgat aaaaatcatt aggtacctgt 7620

acatttatat tggctcatgt ccaacattac cgccatgttg 7660

<210> 24
<211> 511
<212> DNA
<213> Homo sapiens

<400> 24
ggggttgggg ttgcgccttt tccaaggcag ccctgggttt gcgcagggac gcggctgctc 60

tgggcgtggt tccgggaaac gcagcggcgc cgaccctggg tctcgcacat tcttcacgtc 120

cgttcgcagc gtcacccgga tcttcgccgc tacccttgtg ggcccccggg cgacgcttcc 180

tgctccgccc ctaagtcggg aaggttcctt gcggttcgcg gcgtgccgga cgtgacaaac 240

ggaagccgca cgtctcacta gtaccctcgc agacggacag cgccagggag caatggcagc 300

gcgccgaccg cgatgggctg tggccaatag cggctgctca gcagggcgcg ccgagagcag 360

cggccgggaa ggggcggtgc gggaggcggg gtgtggggcg gtagtgtggg ccctgttcct 420

gcccgcgcgg tgttccgcat tctgcaagcc tccggagcgc acgtcggcag tcggctccct 480

cgttgaccga atcaccgacc tctctcccca g 511

<210> 25
<211> 7613
<212> DNA
<213> Artificial Sequence

<220>
<223> Made in Lab - plasmid construct

<400> 25
acattgatta ttgactagtt attaatagta atcaattacg gggtcattag ttcatagccc 60

atatatggag ttccgcgtta cataacttac ggtaaatggc ccgcctggct gaccgcccaa 120

cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc caatagggac 180

tttccattga cgtcaatggg tggagtattt acggtaaact gcccacttgg cagtacatca 240

agtgtatcat atgccaagta cgccccctat tgacgtcaat gacggtaaat ggcccgcctg 300

gcattatgcc cagtacatga ccttatggga ctttcctact tggcagtaca tctacgtatt 360

agtcatcgct attaccatgg tgatgcggtt ttggcagtac atcaatgggc gtggatagcg 420

gtttgactca cggggatttc caagtctcca ccccattgac gtcaatggga gtttgttttg 480

gcaccaaaat caacgggact ttccaaaatg tcgtaacaac tccgccccat tgacgcaaat 540

gggcggtagg cgtgtacggt gggaggtcta tataagcaga gctcgtttag tgaaccgggg 600

tctctctggt tagaccagat ctgagcctgg gagctctctg gctaactagg gaacccactg 660

cttaagcctc aataaagctt gccttgagtg cttcaagtag tgtgtgcccg tctgttgtgt 720

gactctggta actagagatc cctcagaccc ttttagtcag tgtggaaaat ctctagcagt 780

ggcgcccgaa cagggacttg aaagcgaaag ggaaaccaga ggagctctct cgacgcagga 840

ctcggcttgc tgaagcgcgc acggcaagag gcgaggggcg gcgactggtg agtacgccaa 900

aaattttgac tagcggaggc tagaaggaga gagatgggtg cgagagcgtc agtattaagc 960

gggggagaat tagatcgcga tgggaaaaaa ttcggttaag gccaggggga aagaaaaaat 1020

ataaattaaa acatatagta tgggcaagca gggagctaga acgattcgca gttaatcctg 1080

gcctgttaga aacatcagaa ggctgtagac aaatactggg acagctacaa ccatcccttc 1140

agacaggatc agaagaactt agatcattat ataatacagt agcaaccctc tattgtgtgc 1200

atcaaaggat agagataaaa gacaccaagg aagctttaga caagatagag gaagagcaaa 1260

acaaaagtaa gaccaccgca cagcaagcgg ccgctgatct tcagacctgg aggaggagat 1320

atgagggaca attggagaag tgaattatat aaatataaag tagtaaaaat tgaaccatta 1380

ggagtagcac ccaccaaggc aaagagaaga gtggtgcaga gagaaaaaag agcagtggga 1440

ataggagctt tgttccttgg gttcttggga gcagcaggaa gcactatggg cgcagcgtca 1500

atgacgctga cggtacaggc cagacaatta ttgtctggta tagtgcagca gcagaacaat 1560

ttgctgaggg ctattgaggc gcaacagcat ctgttgcaac tcacagtctg gggcatcaag 1620

cagctccagg caagaatcct ggctgtggaa agatacctaa aggatcaaca gctcctgggg 1680

atttggggtt gctctggaaa actcatttgc accactgctg tgccttggaa tgctagttgg 1740

agtaataaat ctctggaaca gatttggaat cacacgacct ggatggagtg ggacagagaa 1800

attaacaatt acacaagctt aatacactcc ttaattgaag aatcgcaaaa ccagcaagaa 1860

aagaatgaac aagaattatt ggaattagat aaatgggcaa gtttgtggaa ttggtttaac 1920

ataacaaatt ggctgtggta tataaaatta ttcataatga tagtaggagg cttggtaggt 1980

ttaagaatag tttttgctgt actttctata gtgaatagag ttaggcaggg atattcacca 2040

ttatcgtttc agacccacct cccaaccccg aggggacccg acaggcccga aggaatagaa 2100

gaagaaggtg gagagagaga cagagacaga tccattcgat tagtgaacgg atctcgacgg 2160

tatcggttaa cttttaaaag aaaaggggggg attgggggggt acagtgcagg ggaaagaata 2220

gtagacataa tagcaacaga catacaaact aaagaattac aaaaacaaat tacaaaaatt 2280

caaaatttta tccgtcagtg ggcagagcgc acatcgccca cagtccccga gaagttgggg 2340

ggaggggtcg gcaattgaac cggtgcctag agaaggtggc gcggggtaaa ctgggaaagt 2400

gatgtcgtgt actggctccg cctttttccc gagggtgggg gagaaccgta tataagtgca 2460

gtagtcgccg tgaacgttct ttttcgcaac gggtttgccg ccagaacaca ggtgtcgtga 2520

cgcgggatcc gccaccatgg gctccatgtt tcggagcgag gaggtggccc tggtccagct 2580

ctttctgccc acagcggctg cctacacctg cgtgagtcgg ctgggcgagc tgggcctcgt 2640

ggagttcaga gacctcaacg cctcggtgag cgccttccag agacgctttg tggttgatgt 2700

tcggcgctgt gaggagctgg agaagacctt caccttcctg caggaggagg tgcggcgggc 2760

tgggctggtc ctgcccccgc caaaggggag gctgccggca cccccacccc gggacctgct 2820

gcgcatccag gaggagacgg agcgcctggc ccaggagctg cgggatgtgc ggggcaacca 2880

gcaggccctg cgggcccagc tgcaccagct gcagctccac gccgccgtgc tacgccaggg 2940

ccatgaacct cagctggcag ccgcccacac agatggggcc tcagagagga cgcccctgct 3000

ccaggccccc ggggggccgc accaggacct gagggtcaac tttgtggcag gtgccgtgga 3060

gccccacaag gcccctgccc tagagcgcct gctctggagg gcctgcagag gcttcctcat 3120

tgccagcttc agggagctgg agcagccgct ggagcacccc gtgacgggcg agccagccac 3180

gtggatgacc ttcctcatct cctactgggg tgagcagatc ggacagaaga tccgcaagat 3240

cacggactgc ttccactgcc acgtcttccc gtttctgcag caggaggagg cccgcctcgg 3300

ggccctgcag cagctgcaac agcagagcca ggagctgcag gaggtcctcg gggagacaga 3360

gcggttcctg agccaggtgc taggccgggt gctgcagctg ctgccgccag ggcaggtgca 3420

ggtccacaag atgaaggccg tgtacctggc cctgaaccag tgcagcgtga gcaccacgca 3480

caagtgcctc attgccgagg cctggtgctc tgtgcgagac ctgcccgccc tgcaggaggc 3540

cctgcgggac agctcgatgg aggagggagt gagtgccgtg gctcaccgca tcccctgccg 3600

ggacatgccc cccacactca tccgcaccaa ccgcttcacg gccagcttcc agggcatcgt 3660

ggatgcctac ggcgtgggcc gctaccagga ggtcaacccc gctccctaca ccatcatcac 3720

cttccccttc ctgtttgctg tgatgttcgg ggatgtgggc cacgggctgc tcatgttcct 3780

cttcgccctg gccatggtcc ttgcggagaa ccgaccggct gtgaaggccg cgcagaacga 3840

gatctggcag actttcttca ggggccgcta cctgctcctg cttatgggcc tgttctccat 3900

ctacaccggc ttcatctaca acgagtgctt cagtcgcgcc accagcatct tcccctcggg 3960

ctggagtgtg gccgccatgg ccaaccagtc tggctggagt gatgcattcc tggcccagca 4020

cacgatgctt accctggacc ccaacgtcac cggtgtcttc ctgggaccct acccctttgg 4080

catcgatcct atttggagcc tggctgccaa ccacttgagc ttcctcaact ccttcaagat 4140

gaagatgtcc gtcatcctgg gcgtcgtgca catggccttt ggggtggtcc tcggagtctt 4200

caaccacgtg cactttggcc agaggcaccg gctgctgctg gagacgctgc cggagctcac 4260

cttcctgctg ggactcttcg gttacctcgt gttcctagtc atctacaagt ggctgtgtgt 4320

ctgggctgcc agggccgcct cggcccccag catcctcatc cacttcatca acatgttcct 4380

cttctcccac agccccagca acaggctgct ctacccccgg caggaggtgg tccaggccac 4440

gctggtggtc ctggccttgg ccatggtgcc catcctgctg cttggcacac ccctgcacct 4500

gctgcaccgc caccgccgcc gcctgcggag gaggcccgct gaccgacagg aggaaaacaa 4560

ggccgggttg ctggacctgc ctgacgcatc tgtgaatggc tggagctccg atgaggaaaa 4620

ggcaggggggc ctggatgatg aagaggaggc cgagctcgtc ccctccgagg tgctcatgca 4680

ccaggccatc cacaccatcg agttctgcct gggctgcgtc tccaacaccg cctcctacct 4740

gcgcctgtgg gccctgagcc tggcccacgc ccagctgtcc gaggttctgt gggccatggt 4800

gatgcgcata ggcctgggcc tgggccggga ggtgggcgtg gcggctgtgg tgctggtccc 4860

catctttgcc gcctttgccg tgatgaccgt ggctatcctg ctggtgatgg agggactctc 4920

agccttcctg cacgccctgc ggctgcactg ggtggaattc cagaacaagt tctactcagg 4980

cacgggctac aagctgagtc ccttcacctt cgctgccaca gatgactagt aagtcgacgg 5040

atcccccggg ctgcaggaat tcgagcatct taccgccatt tatacccata tttgttctgt 5100

ttttcttgat ttgggtatac atttaaatgt taataaaaca aaatggtggg gcaatcattt 5160

acatttttag ggatatgtaa ttactagttc aggtgtattg ccacaagaca aacatgttaa 5220

gaaactttcc cgttatttac gctctgttcc tgttaatcaa cctctggatt acaaaatttg 5280

tgaaagattg actgatattc ttaactatgt tgctcctttt acgctgtgtg gatatgctgc 5340

tttaatgcct ctgtatcatg ctattgcttc ccgtacggct ttcgttttct cctccttgta 5400

taaatcctgg ttgctgtctc tttatgagga gttgtggccc gttgtccgtc aacgtggcgt 5460

ggtgtgctct gtgtttgctg acgcaacccc cactggctgg ggcattgcca ccacctgtca 5520

actcctttct gggactttcg ctttcccccc cccgatcgcc acggcagaac tcatcgccgc 5580

ctgccttgcc cgctgctgga caggggctag gttgctgggc actgataatt ccgtggtgtt 5640

gtcggggaag ctgacgtcct ttcgaattcg atatcaagct gtacctttaa gaccaatgac 5700

ttacaaggca gctgtagatc ttagccactt tttaaaagaa aaggggggac tggaagggct 5760

aattcactcc caacgaagac aagatctgct ttttgcttgt actgggtctc tctggttaga 5820

ccagatctga gcctgggagc tctctggcta actagggaac ccactgctta agcctcaata 5880

aagcttgcct tgagtgcttc aagtagtgtg tgcccgtctg ttgtgtgact ctggtaacta 5940

gagatccctc agacccttttt agtcagtgtg gaaaatctct agcagtagta gttcatgtca 6000

tcttattatt cagtatttat aacttgcaaa gaaatgaata tcagagagtg agaggaactt 6060

gtttattgca gcttataatg gttacaaata aagcaatagc atcacaaatt tcacaaataa 6120

agcatttttt tcactgcatt ctagttgtgg tttgtccaaa ctcatcaatg tatcttatca 6180

tgtctggctc tagctatccc gcccctaact ccgcccatcc cgcccctaac tccgcccagt 6240

tccgcccatt ctccgcccca tggctgacta attttttttа tttatgcaga ggccgaggcc 6300

gcctcggcct ctgagctatt ccagaagtag tgaggaggct tttttggagg cctaggtagc 6360

ccgcctaatg agcgggcttt ttttttcttag gccttcttcc gcttcctcgc tcactgactc 6420

gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg 6480

gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa 6540

ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccccctga 6600

cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag 6660

ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct 6720

taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc atagctcacg 6780

ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc 6840

ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt ccaacccggt 6900

aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta 6960

tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaagaac 7020

agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc 7080

ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat 7140

tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc 7200

tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt 7260

cacctagatc cttttaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta 7320

aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct 7380

atttcgttca tccatagttg cctgactcct gcaaaccacg ttgtggtaga attggtaaag 7440

agagtcgtgt aaaatatcga gttcgcacat cttgttgtct gattattgat ttttggcgaa 7500

accatttgat catatgacaa gatgtgtatc taccttaact taatgatttt gataaaaatc 7560

attaggtacc tgtacattta tattggctca tgtccaacat taccgccatg ttg 7613

<210> 26
<211> 47
<212> DNA
<213> Artificial Sequence

<220>
<223> Made in Lab - section of plasmid construct

<400> 26
gatcacgaga ctagcctcga gaagcttgat cgattggctc cggtgcc 47

<210> 27
<211> 7646
<212> DNA
<213> Artificial Sequence

<220>
<223> Made in lab - plasmid construct

<400> 27
ggctccggtg cccgtcagtg ggcagagcgc acatcgccca cagtccccga gaagttgggg 60

ggaggggtcg gcaattgaac cggtgcctag agaaggtggc gcggggtaaa ctgggaaagt 120

gatgtcgtgt actggctccg cctttttccc gagggtgggg gagaaccgta tataagtgca 180

gtagtcgccg tgaacgttct ttttcgcaac gggtttgccg ccagaacaca ggtgtcgtga 240

cgcgggatcc gccaccatgg gctccatgtt tcggagcgag gaggtggccc tggtccagct 300

ctttctgccc acagcggctg cctacacctg cgtgagtcgg ctgggcgagc tgggcctcgt 360

ggagttcaga gacctcaacg cctcggtgag cgccttccag agacgctttg tggttgatgt 420

tcggcgctgt gaggagctgg agaagacctt caccttcctg caggaggagg tgcggcgggc 480

tgggctggtc ctgcccccgc caaaggggag gctgccggca cccccacccc gggacctgct 540

gcgcatccag gaggagacgg agcgcctggc ccaggagctg cgggatgtgc ggggcaacca 600

gcaggccctg cgggcccagc tgcaccagct gcagctccac gccgccgtgc tacgccaggg 660

ccatgaacct cagctggcag ccgcccacac agatggggcc tcagagagga cgcccctgct 720

ccaggccccc ggggggccgc accaggacct gagggtcaac tttgtggcag gtgccgtgga 780

gccccacaag gcccctgccc tagagcgcct gctctggagg gcctgcagag gcttcctcat 840

tgccagcttc agggagctgg agcagccgct ggagcacccc gtgacgggcg agccagccac 900

gtggatgacc ttcctcatct cctactgggg tgagcagatc ggacagaaga tccgcaagat 960

cacggactgc ttccactgcc acgtcttccc gtttctgcag caggaggagg cccgcctcgg 1020

ggccctgcag cagctgcaac agcagagcca ggagctgcag gaggtcctcg gggagacaga 1080

gcggttcctg agccaggtgc taggccgggt gctgcagctg ctgccgccag ggcaggtgca 1140

ggtccacaag atgaaggccg tgtacctggc cctgaaccag tgcagcgtga gcaccacgca 1200

caagtgcctc attgccgagg cctggtgctc tgtgcgagac ctgcccgccc tgcaggaggc 1260

cctgcgggac agctcgatgg aggagggagt gagtgccgtg gctcaccgca tcccctgccg 1320

ggacatgccc cccacactca tccgcaccaa ccgcttcacg gccagcttcc agggcatcgt 1380

ggatgcctac ggcgtgggcc gctaccagga ggtcaacccc gctccctaca ccatcatcac 1440

cttccccttc ctgtttgctg tgatgttcgg ggatgtgggc cacgggctgc tcatgttcct 1500

cttcgccctg gccatggtcc ttgcggagaa ccgaccggct gtgaaggccg cgcagaacga 1560

gatctggcag actttcttca ggggccgcta cctgctcctg cttatgggcc tgttctccat 1620

ctacaccggc ttcatctaca acgagtgctt cagtcgcgcc accagcatct tcccctcggg 1680

ctggagtgtg gccgccatgg ccaaccagtc tggctggagt gatgcattcc tggcccagca 1740

cacgatgctt accctggacc ccaacgtcac cggtgtcttc ctgggaccct acccctttgg 1800

catcgatcct atttggagcc tggctgccaa ccacttgagc ttcctcaact ccttcaagat 1860

gaagatgtcc gtcatcctgg gcgtcgtgca catggccttt ggggtggtcc tcggagtctt 1920

caaccacgtg cactttggcc agaggcaccg gctgctgctg gagacgctgc cggagctcac 1980

cttcctgctg ggactcttcg gttacctcgt gttcctagtc atctacaagt ggctgtgtgt 2040

ctgggctgcc agggccgcct cggcccccag catcctcatc cacttcatca acatgttcct 2100

cttctcccac agccccagca acaggctgct ctacccccgg caggaggtgg tccaggccac 2160

gctggtggtc ctggccttgg ccatggtgcc catcctgctg cttggcacac ccctgcacct 2220

gctgcaccgc caccgccgcc gcctgcggag gaggcccgct gaccgacagg aggaaaacaa 2280

ggccgggttg ctggacctgc ctgacgcatc tgtgaatggc tggagctccg atgaggaaaa 2340

ggcaggggc ctggatgatg aagaggaggc cgagctcgtc ccctccgagg tgctcatgca 2400

ccaggccatc cacaccatcg agttctgcct gggctgcgtc tccaacaccg cctcctacct 2460

gcgcctgtgg gccctgagcc tggcccacgc ccagctgtcc gaggttctgt gggccatggt 2520

gatgcgcata ggcctgggcc tgggccggga ggtgggcgtg gcggctgtgg tgctggtccc 2580

catctttgcc gcctttgccg tgatgaccgt ggctatcctg ctggtgatgg agggactctc 2640

agccttcctg cacgccctgc ggctgcactg ggtggaattc cagaacaagt tctactcagg 2700

cacgggctac aagctgagtc ccttcacctt cgctgccaca gatgactagt aagtcgacgg 2760

atcccccggg ctgcaggaat tcgagcatct taccgccatt tatacccata tttgttctgt 2820

ttttcttgat ttgggtatac atttaaatgt taataaaaca aaatggtggg gcaatcattt 2880

acatttttag ggatatgtaa ttactagttc aggtgtattg ccacaagaca aacatgttaa 2940

gaaactttcc cgttatttac gctctgttcc tgttaatcaa cctctggatt acaaaatttg 3000

tgaaagattg actgatattc ttaactatgt tgctcctttt acgctgtgtg gatatgctgc 3060

tttaatgcct ctgtatcatg ctattgcttc ccgtacggct ttcgttttct cctccttgta 3120

taaatcctgg ttgctgtctc tttatgagga gttgtggccc gttgtccgtc aacgtggcgt 3180

ggtgtgctct gtgtttgctg acgcaacccc cactggctgg ggcattgcca ccacctgtca 3240

actcctttct gggactttcg ctttcccccc cccgatcgcc acggcagaac tcatcgccgc 3300

ctgccttgcc cgctgctgga caggggctag gttgctgggc actgataatt ccgtggtgtt 3360

gtcggggaag ctgacgtcct ttcgaattcg atatcaagct gtacctttaa gaccaatgac 3420

ttacaaggca gctgtagatc ttagccactt tttaaaagaa aaggggggac tggaagggct 3480

aattcactcc caacgaagac aagatctgct ttttgcttgt actgggtctc tctggttaga 3540

ccagatctga gcctgggagc tctctggcta actagggaac ccactgctta agcctcaata 3600

aagcttgcct tgagtgcttc aagtagtgtg tgcccgtctg ttgtgtgact ctggtaacta 3660

gagatccctc agacccttttt agtcagtgtg gaaaatctct agcagtagta gttcatgtca 3720

tcttattatt cagtatttat aacttgcaaa gaaatgaata tcagagagtg agaggaactt 3780

gtttattgca gcttataatg gttacaaata aagcaatagc atcacaaatt tcacaaataa 3840

agcatttttt tcactgcatt ctagttgtgg tttgtccaaa ctcatcaatg tatcttatca 3900

tgtctggctc tagctatccc gcccctaact ccgcccatcc cgcccctaac tccgcccagt 3960

tccgcccatt ctccgcccca tggctgacta atttttttta tttatgcaga ggccgaggcc 4020

gcctcggcct ctgagctatt ccagaagtag tgaggaggct tttttggagg cctaggtagc 4080

ccgcctaatg agcgggcttt tttttcttag gccttcttcc gcttcctcgc tcactgactc 4140

gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg 4200

gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa 4260

ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccctga 4320

cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag 4380

ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct 4440

taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc atagctcacg 4500

ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc 4560

ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt ccaacccggt 4620

aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta 4680

tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaagaac 4740

agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc 4800

ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat 4860

tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc 4920

tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt 4980

cacctagatc cttttaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta 5040

aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct 5100

atttcgttca tccatagttg cctgactcct gcaaaccacg ttgtggtaga attggtaaag 5160

agagtcgtgt aaaatatcga gttcgcacat cttgttgtct gattattgat ttttggcgaa 5220

accatttgat catatgacaa gatgtgtatc taccttaact taatgatttt gataaaaatc 5280

attaggtacc tgtacattta tattggctca tgtccaacat taccgccatg ttgacattga 5340

ttattgacta gttattaata gtaatcaatt acggggtcat tagttcatag cccatatatg 5400

gagttccgcg ttacataact tacggtaaat ggcccgcctg gctgaccgcc caacgacccc 5460

cgcccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat 5520

tgacgtcaat gggtggagta tttacggtaa actgcccact tggcagtaca tcaagtgtat 5580

catatgccaa gtacgccccc tattgacgtc aatgacggta aatggcccgc ctggcattat 5640

gcccagtaca tgaccttatg ggactttcct acttggcagt acatctacgt attagtcatc 5700

gctattacca tggtgatgcg gttttggcag tacatcaatg ggcgtggata gcggtttgac 5760

tcacggggat ttccaagtct ccaccccatt gacgtcaatg ggagtttgtt ttggcaccaa 5820

aatcaacggg actttccaaa atgtcgtaac aactccgccc cattgacgca aatgggcggt 5880

aggcgtgtac ggtgggaggt ctatataagc agagctcgtt tagtgaaccg gggtctctct 5940

ggttagacca gatctgagcc tgggagctct ctggctaact agggaaccca ctgcttaagc 6000

ctcaataaag cttgccttga gtgcttcaag tagtgtgtgc ccgtctgttg tgtgactctg 6060

gtaactagag atccctcaga ccctttttagt cagtgtggaa aatctctagc agtggcgccc 6120

gaacagggac ttgaaagcga aagggaaacc agaggagctc tctcgacgca ggactcggct 6180

tgctgaagcg cgcacggcaa gaggcgaggg gcggcgactg gtgagtacgc caaaaatttt 6240

gactagcgga ggctagaagg agagagatgg gtgcgagagc gtcagtatta agcgggggag 6300

aattagatcg cgatgggaaa aaattcggtt aaggccaggg ggaaagaaaa aatataaatt 6360

aaaacatata gtatgggcaa gcagggagct agaacgattc gcagttaatc ctggcctgtt 6420

agaaacatca gaaggctgta gacaaatact gggacagcta caaccatccc ttcagacagg 6480

atcagaagaa cttagatcat tatataatac agtagcaacc ctctattgtg tgcatcaaag 6540

gatagagata aaagacacca aggaagcttt agacaagata gaggaagagc aaaacaaaag 6600

taagaccacc gcacagcaag cggccgctga tcttcagacc tggaggagga gatatgaggg 6660

acaattggag aagtgaatta tataaatata aagtagtaaa aattgaacca ttaggagtag 6720

cacccaccaa ggcaaagaga agagtggtgc agagagaaaa aagagcagtg ggaataggag 6780

ctttgttcct tgggttcttg ggagcagcag gaagcactat gggcgcagcg tcaatgacgc 6840

tgacggtaca ggccagacaa ttattgtctg gtatagtgca gcagcagaac aatttgctga 6900

gggctattga ggcgcaacag catctgttgc aactcacagt ctggggcatc aagcagctcc 6960

aggcaagaat cctggctgtg gaaagatacc taaaggatca acagctcctg gggatttggg 7020

gttgctctgg aaaactcatt tgcaccactg ctgtgccttg gaatgctagt tggagtaata 7080

aatctctgga acagatttgg aatcacacga cctggatgga gtgggacaga gaaattaaca 7140

attacacaag cttaatacac tccttaattg aagaatcgca aaaccagcaa gaaaagaatg 7200

aacaagaatt attggaatta gataaatggg caagtttgtg gaattggttt aacataacaa 7260

attggctgtg gtatataaaa ttattcataa tgatagtagg aggcttggta ggtttaagaa 7320

tagtttttgc tgtactttct atagtgaata gagttaggca gggatattca ccattatcgt 7380

ttcagaccca cctcccaacc ccgaggggac ccgacaggcc cgaaggaata gaagaagaag 7440

gtggagagag agacagagac agatccattc gattagtgaa cggatctcga cggtatcggt 7500

taacttttaa aagaaaaggg gggattgggg ggtacagtgc aggggaaaga atagtagaca 7560

taatagcaac agacatacaa actaaagaat tacaaaaaca aattacaaaa attcaaaatt 7620

ttatcgatca cgagactagc ctcgag 7646

<210> 28
<211> 234
<212> DNA
<213> human immunodeficiency virus

<400> 28
tggaagggct aattcactcc caacgaagac aagatctgct ttttgcttgt actgggtctc 60

tctggttaga ccagatctga gcctgggagc tctctggcta actagggaac ccactgctta 120

agcctcaata aagcttgcct tgagtgcttc aagtagtgtg tgcccgtctg ttgtgtgact 180

ctggtaacta gagatccctc agacccttttt agtcagtgtg gaaaatctct agca 234

<210> 29
<211> 132
<212> DNA
<213> Simian virus 40

<400> 29
aacttgttta ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca 60

aataaagcat ttttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct 120

tatcatgtct gg 132

<210> 30
<211> 160
<212> DNA
<213> Simian virus 40

<400> 30
tcccgcccct aactccgccc atcccgcccc taactccgcc cagttccgcc cattctccgc 60

cccatggctg actaattttt tttatttatg cagaggccga ggccgcctcg gcctctgagc 120

tattccagaa gtagtgagga ggctttttg gaggcctagg 160

<210> 31
<211> 1015
<212> DNA
<213> Artificial Sequence

<220>
<223> Made in Lab

<400> 31
tcttccgctt cctcgctcac tgactcgctg cgctcggtcg ttcggctgcg gcgagcggta 60

tcagctcact caaaggcggt aatacggtta tccacagaat caggggataa cgcaggaaag 120

aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg 180

tttttccata ggctccgccc ccctgacgag catcacaaaa atcgacgctc aagtcagagg 240

tggcgaaacc cgacaggact ataaagatac caggcgtttc cccctggaag ctccctcgtg 300

cgctctcctg ttccgaccct gccgcttacc ggatacctgt ccgcctttct cccttcggga 360

agcgtggcgc tttctcatag ctcacgctgt aggtatctca gttcggtgta ggtcgttcgc 420

tccaagctgg gctgtgtgca cgaaccccc gttcagcccg accgctgcgc cttatccggt 480

aactatcgtc ttgagtccaa cccggtaaga cacgacttat cgccactggc agcagccact 540

ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagttctt gaagtggtgg 600

cctaactacg gctacactag aagaacagta tttggtatct gcgctctgct gaagccagtt 660

accttcggaa aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcggt 720

ggtttttttg tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct 780

ttgatctttt ctacggggtc tgacgctcag tggaacgaaa actcacgtta agggattttg 840

gtcatgagat tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt 900

aaatcaatct aaagtatata tgagtaaact tggtctgaca gttaccaatg cttaatcagt 960

gaggcaccta tctcagcgat ctgtctattt cgttcatcca tagttgcctg actcc 1015

<210> 32
<211> 139
<212> DNA
<213> Escherichia coli

<400> 32
gtagaattgg taaagagagt cgtgtaaaat atcgagttcg cacatcttgt tgtctgatta 60

ttgatttttg gcgaaaccat ttgatcatat gacaagatgt gtatctacct taacttaatg 120

attttgataa aaatcatta 139

<210> 33
<211> 577
<212> DNA
<213> Human betaherpesvirus 5

<400> 33
acattgatta ttgactagtt attaatagta atcaattacg gggtcattag ttcatagccc 60

atatatggag ttccgcgtta cataacttac ggtaaatggc ccgcctggct gaccgcccaa 120

cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc caatagggac 180

tttccattga cgtcaatggg tggagtattt acggtaaact gcccacttgg cagtacatca 240

agtgtatcat atgccaagta cgcccccctat tgacgtcaat gacggtaaat ggcccgcctg 300

gcattatgcc cagtacatga ccttatggga ctttcctact tggcagtaca tctacgtatt 360

agtcatcgct attaccatgg tgatgcggtt ttggcagtac atcaatgggc gtggatagcg 420

gtttgactca cggggatttc caagtctcca ccccattgac gtcaatggga gtttgttttg 480

gcaccaaaat caacgggact ttccaaaatg tcgtaacaac tccgccccat tgacgcaaat 540

gggcggtagg cgtgtacggt gggaggtcta tataagc 577

<210> 34
<211> 188
<212> DNA
<213> human immunodeficiency virus

<400> 34
gtctctctgg ttagaccaga tctgagcctg ggagctctct ggctaactag ggaacccact 60

gcttaagcct caataaagct tgccttgagt gcttcaagta gtgtgtgccc gtctgttgtg 120

tgactctggt aactagagat ccctcagacc cttttagtca gtgtggaaaa tctctagcag 180

tggcgccc 188

<210> 35
<211> 45
<212> DNA
<213> Human immunodeficiency virus 1

<400> 35
tgagtacgcc aaaaattttg actagcggag gctagaagga gagag 45

<210> 36
<211> 362
<212> DNA
<213> human immunodeficiency virus

<400> 36
atgggtgcga gagcgtcagt attaagcggg ggagaattag atcgcgatgg gaaaaaattc 60

ggttaaggcc agggggaaag aaaaaatata aattaaaaca tatagtatgg gcaagcaggg 120

agctagaacg attcgcagtt aatcctggcc tgttagaaac atcagaaggc tgtagacaaa 180

tactgggaca gctacaacca tcccttcaga caggatcaga agaacttaga tcattatata 240

atacagtagc aaccctctat tgtgtgcatc aaaggataga gataaaagac accaaggaag 300

ctttagacaa gatagaggaa gagcaaaaca aaagtaagac caccgcacag caagcggccg 360

ct 362

<210> 37
<211> 858
<212> DNA
<213> Artificial Sequence

<220>
<223> Made in Lab - plasmid element

<400> 37
gatcttcaga cctggaggag gagatatgag ggacaattgg agaagtgaat tatataaata 60

taaagtagta aaaattgaac cattaggagt agcacccacc aaggcaaaga gaagagtggt 120

gcagagagaa aaaagagcag tgggaatagg agctttgttc cttgggttct tgggagcagc 180

aggaagcact atgggcgcag cgtcaatgac gctgacggta caggccagac aattattgtc 240

tggtatagtg cagcagcaga acaatttgct gagggctatt gaggcgcaac agcatctgtt 300

gcaactcaca gtctggggca tcaagcagct ccaggcaaga atcctggctg tggaaagata 360

cctaaaggat caacagctcc tggggatttg gggttgctct ggaaaactca tttgcaccac 420

tgctgtgcct tggaatgcta gttggagtaa taaatctctg gaacagattt ggaatcacac 480

gacctggatg gagtgggaca gagaaattaa caattacaca agcttaatac actccttaat 540

tgaagaatcg caaaaccagc aagaaaagaa tgaacaagaa ttattggaat tagataaatg 600

ggcaagtttg tggaattggt ttaacataac aaattggctg tggtatataa aattattcat 660

aatgatagta ggaggcttgg taggtttaag aatagttttt gctgtacttt ctatagtgaa 720

tagagttagg cagggatatt caccattatc gtttcagacc cacctcccaa ccccgagggg 780

acccgacagg cccgaaggaa tagaagaaga aggtggagag agagacagag acagatccat 840

tcgattagtg aacggatc 858

<210> 38
<211> 118
<212> DNA
<213> human immunodeficiency virus

<400> 38
ttttaaaaga aaaggggggа ttgggggggta cagtgcaggg gaaagaatag tagacataat 60

agcaacagac atacaaacta aagaattaca aaaacaaatt acaaaaattc aaaatttt 118

<210> 39
<211> 3847
<212> DNA
<213> Artificial Sequence

<220>
<223> Made in lab - plasmid backbone construct

<400> 39
aacttgttta ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca 60

aataaagcat ttttttcact gcattctagt tgtggtttgt ccaaactcat caatgtatct 120

tatcatgtct ggctctagct atcccgcccc taactccgcc catcccgccc ctaactccgc 180

ccagttccgc ccattctccg ccccatggct gactaatttt ttttatttat gcagaggccg 240

aggccgcctc ggcctctgag ctattccaga agtagtgagg aggctttttt ggaggcctag 300

gtagcccgcc taatgagcgg gctttttttt cttaggcctt cttccgcttc ctcgctcact 360

gactcgctgc gctcggtcgt tcggctgcgg cgagcggtat cagctcactc aaaggcggta 420

atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc aaaaggccag 480

caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag gctccgcccc 540

cctgacgagc atcacaaaaa tcgacgctca agtcagaggt ggcgaaaccc gacaggacta 600

taaagatacc aggcgtttcc ccctggaagc tccctcgtgc gctctcctgt tccgaccctg 660

ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gcgtggcgct ttctcatagc 720

tcacgctgta ggtatctcag ttcggtgtag gtcgttcgct ccaagctggg ctgtgtgcac 780

gaaccccccg ttcagcccga ccgctgcgcc ttatccggta actatcgtct tgagtccaac 840

ccggtaagac acgacttatc gccactggca gcagccactg gtaacaggat tagcagagcg 900

aggtatgtag gcggtgctac agagttcttg aagtggtggc ctaactacgg ctacactaga 960

agaacagtat ttggtatctg cgctctgctg aagccagtta ccttcggaaa aagagttggt 1020

agctcttgat ccggcaaaca aaccaccgct ggtagcggtg gtttttttgt ttgcaagcag 1080

cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc tacggggtct 1140

gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt atcaaaaagg 1200

atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta aagtatatat 1260

gagtaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat ctcagcgatc 1320

tgtctatttc gttcatccat agttgcctga ctcctgcaaa ccacgttgtg gtagaattgg 1380

taaagagagt cgtgtaaaat atcgagttcg cacatcttgt tgtctgatta ttgatttttg 1440

gcgaaaccat ttgatcatat gacaagatgt gtatctacct taacttaatg attttgataa 1500

aaatcattag gtacctgtac atttatattg gctcatgtcc aacattaccg ccatgttgac 1560

attgattatt gactagttat taatagtaat caattacggg gtcattagtt catagcccat 1620

atatggagtt ccgcgttaca taacttacgg taaatggccc gcctggctga ccgcccaacg 1680

acccccgccc attgacgtca ataatgacgt atgttcccat agtaacgcca atagggactt 1740

tccattgacg tcaatgggtg gagtatttac ggtaaactgc ccacttggca gtacatcaag 1800

tgtatcatat gccaagtacg ccccctattg acgtcaatga cggtaaatgg cccgcctggc 1860

attatgccca gtacatgacc ttatgggact ttcctacttg gcagtacatc tacgtattag 1920

tcatcgctat taccatggtg atgcggtttt ggcagtacat caatgggcgt ggatagcggt 1980

ttgactcacg gggatttcca agtctccacc ccattgacgt caatgggagt ttgttttggc 2040

accaaaatca acgggacttt ccaaaatgtc gtaacaactc cgccccattg acgcaaatgg 2100

gcggtaggcg tgtacggtgg gaggtctata taagcagagc tcgtttagtg aaccggggtc 2160

tctctggtta gaccagatct gagcctggga gctctctggc taactaggga acccactgct 2220

taagcctcaa taaagcttgc cttgagtgct tcaagtagtg tgtgcccgtc tgttgtgtga 2280

ctctggtaac tagagatccc tcagaccctt ttagtcagtg tggaaaatct ctagcagtgg 2340

cgcccgaaca gggacttgaa agcgaaaggg aaaccagagg agctctctcg acgcaggact 2400

cggcttgctg aagcgcgcac ggcaagaggc gaggggcggc gactggtgag tacgccaaaa 2460

attttgacta gcggaggcta gaaggagaga gatgggtgcg agagcgtcag tattaagcgg 2520

gggagaatta gatcgcgatg ggaaaaaatt cggttaaggc caggggaaa gaaaaaatat 2580

aaattaaaac atatagtatg ggcaagcagg gagctagaac gattcgcagt taatcctggc 2640

ctgttagaaa catcagaagg ctgtagacaa atactgggac agctacaacc atcccttcag 2700

acaggatcag aagaacttag atcattatat aatacagtag caaccctcta ttgtgtgcat 2760

caaaggatag agataaaaga caccaaggaa gctttagaca agatagagga agagcaaaac 2820

aaaagtaaga ccaccgcaca gcaagcggcc gctgatcttc agacctggag gaggagatat 2880

gagggacaat tggagaagtg aattatataa atataaagta gtaaaaattg aaccattagg 2940

agtagcaccc accaaggcaa agagaagagt ggtgcagaga gaaaaaagag cagtgggaat 3000

aggagctttg ttccttgggt tcttgggagc agcaggaagc actatgggcg cagcgtcaat 3060

gacgctgacg gtacaggcca gacaattatt gtctggtata gtgcagcagc agaacaattt 3120

gctgagggct attgaggcgc aacagcatct gttgcaactc acagtctggg gcatcaagca 3180

gctccaggca agaatcctgg ctgtggaaag atacctaaag gatcaacagc tcctggggat 3240

ttggggttgc tctggaaaac tcatttgcac cactgctgtg ccttggaatg ctagttggag 3300

taataaatct ctggaacaga tttggaatca cacgacctgg atggagtggg acagagaaat 3360

taacaattac acaagcttaa tacactcctt aattgaagaa tcgcaaaacc agcaagaaaa 3420

gaatgaacaa gaattattgg aattagataa atgggcaagt ttgtggaatt ggtttaacat 3480

aacaaattgg ctgtggtata taaaattatt cataatgata gtaggaggct tggtaggttt 3540

aagaatagtt tttgctgtac tttctatagt gaatagagtt aggcagggat attcaccatt 3600

atcgtttcag acccacctcc caaccccgag gggacccgac aggcccgaag gaatagaaga 3660

agaaggtgga gagagagaca gagacagatc cattcgatta gtgaacggat ctcgacggta 3720

tcggttaact tttaaaagaa aaggggggat tgggggggtac agtgcagggg aaagaatagt 3780

agacataata gcaacagaca tacaaactaa agaattacaa aaacaaatta caaaaattca 3840

aaatttt 3847