SEQUENCE LISTING

<110> Poseida Therapeutics
OSTERTAG, Eric
SHEDLOCK, Devon

<120> MODIFIED STEM CELL MEMORY T CELLS, METHODS OF MAKING AND METHODS OF USING SAME

<130> POTH-012/01WO

<150> 62/402,707
<151> 2016-09-30

<150> 62/553,058
<151> 2017-08-31

<150> 62/556,309
<151> 2017-09-08

<150> 62/502,508
<151> 2017-05-05

<160> 42

<170> PatentIn version 3.5

<210> 1
<211> 89
<212> PRT
<213> Artificial Sequence

<220>
<223> FN3 domain consensus sequence

<400> 1

Leu Pro Ala Pro Lys Asn Leu Val Val Ser Glu Val Thr Glu Asp Ser
1 5 10 15

Leu Arg Leu Ser Trp Thr Ala Pro Asp Ala Ala Phe Asp Ser Phe Leu
20 25 30

Ile Gln Tyr Gln Glu Ser Glu Lys Val Gly Glu Ala Ile Asn Leu Thr
35 40 45

Val Pro Gly Ser Glu Arg Ser Tyr Asp Leu Thr Gly Leu Lys Pro Gly
50 55 60

Thr Glu Tyr Thr Val Ser Ile Tyr Gly Val Lys Gly Gly His Arg Ser

65 70 75 80

Asn Pro Leu Ser Ala Glu Phe Thr Thr
85

<210> 2
<211> 90
<212> PRT
<213> Artificial Sequence

<220>
<223> FN3 consensus sequence

<400> 2

Met Leu Pro Ala Pro Lys Asn Leu Val Val Ser Glu Val Thr Glu Asp
1 5 10 15

Ser Leu Arg Leu Ser Trp Thr Ala Pro Asp Ala Ala Phe Asp Ser Phe
20 25 30

Leu Ile Gln Tyr Gln Glu Ser Glu Lys Val Gly Glu Ala Ile Asn Leu
35 40 45

Thr Val Pro Gly Ser Glu Arg Ser Tyr Asp Leu Thr Gly Leu Lys Pro
50 55 60

Gly Thr Glu Tyr Thr Val Ser Ile Tyr Gly Val Lys Gly Gly His Arg
65 70 75 80

Ser Asn Pro Leu Ser Ala Glu Phe Thr Thr
85 90

<210> 3
<211> 270
<212> DNA
<213> Artificial Sequence

<220>
<223> FN3 consensus sequence

<400> 3
atgctgcctg caccaaagaa cctggtggtg tctcatgtga cagaggatag tgccagactg 60

tcatggactg ctcccgacgc agccttcgat agttttatca tcgtgtaccg ggagaacatc 120

gaaaccggcg aggccattgt cctgacagtg ccagggtccg aacgctctta tgacctgaca 180

gatctgaagc ccggaactga gtactatgtg cagatcgccg gcgtcaaagg aggcaatatc 240

agcttccctc tgtccgcaat cttcaccaca 270

<210> 4
<211> 594
<212> PRT
<213> Trichoplusia ni

<400> 4

Met Gly Ser Ser Leu Asp Asp Glu His Ile Leu Ser Ala Leu Leu Gln
1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Ile Ser Asp
20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu
50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn
65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His
85 90 95

Cys Trp Ser Thr Ser Lys Ser Thr Arg Arg Ser Arg Val Ser Ala Leu
100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile
115 120 125

Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile
130 135 140

Ser Glu Ile Val Lys Trp Thr Asn Ala Glu Ile Ser Leu Lys Arg Arg
145 150 155 160

```
Glu Ser Met Thr Gly Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile
                165                 170                 175

Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn
            180                 185                 190

His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
        195                 200                 205

Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu
    210                 215                 220

Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val
225                 230                 235                 240

Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile
                245                 250                 255

Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu
            260                 265                 270

Gly Phe Arg Gly Arg Cys Pro Phe Arg Met Tyr Ile Pro Asn Lys Pro
        275                 280                 285

Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Tyr Lys
    290                 295                 300

Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn
305                 310                 315                 320

Gly Val Pro Leu Gly Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val
                325                 330                 335

His Gly Ser Cys Arg Asn Ile Thr Cys Asp Asn Trp Phe Thr Ser Ile
            340                 345                 350

Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val
        355                 360                 365

Gly Thr Val Arg Ser Asn Lys Arg Glu Ile Pro Glu Val Leu Lys Asn
    370                 375                 380
```

```
Ser Arg Ser Arg Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro
385                 390                 395                 400

Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
                405                 410                 415

Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
            420                 425                 430

Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
        435                 440                 445

Leu Asp Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg
    450                 455                 460

Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn
465                 470                 475                 480

Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
                485                 490                 495

Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Ser Leu Thr Ser
            500                 505                 510

Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
        515                 520                 525

Arg Asp Asn Ile Ser Asn Ile Leu Pro Asn Glu Val Pro Gly Thr Ser
    530                 535                 540

Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr
545                 550                 555                 560

Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Asn Ala Ser Cys Lys Lys
                565                 570                 575

Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
            580                 585                 590
```

Cys Phe

<210> 5
<211> 594
<212> PRT
<213> Artificial Sequence

<220>
<223> Super Piggybac Transposase

<400> 5

Met Gly Ser Ser Leu Asp Asp Glu His Ile Leu Ser Ala Leu Leu Gln
1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp
20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu
50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn
65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His
85 90 95

Cys Trp Ser Thr Ser Lys Ser Thr Arg Arg Ser Arg Val Ser Ala Leu
100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile
115 120 125

Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile
130 135 140

Ser Glu Ile Val Lys Trp Thr Asn Ala Glu Ile Ser Leu Lys Arg Arg
145 150 155 160

```
Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile
                165                 170                 175

Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn
            180                 185                 190

His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
        195                 200                 205

Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu
    210                 215                 220

Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val
225                 230                 235                 240

Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile
                245                 250                 255

Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu
            260                 265                 270

Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro
        275                 280                 285

Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys
    290                 295                 300

Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn
305                 310                 315                 320

Gly Val Pro Leu Gly Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val
                325                 330                 335

His Gly Ser Cys Arg Asn Ile Thr Cys Asp Asn Trp Phe Thr Ser Ile
            340                 345                 350

Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val
        355                 360                 365

Gly Thr Val Arg Ser Asn Lys Arg Glu Ile Pro Glu Val Leu Lys Asn
    370                 375                 380
```

```
Ser Arg Ser Arg Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro
385                 390                 395                 400

Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
                405                 410                 415

Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
            420                 425                 430

Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
        435                 440                 445

Leu Asp Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg
    450                 455                 460

Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn
465                 470                 475                 480

Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
                485                 490                 495

Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Ser Leu Thr Ser
            500                 505                 510

Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
        515                 520                 525

Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser
    530                 535                 540

Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr
545                 550                 555                 560

Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Asn Ala Ser Cys Lys Lys
                565                 570                 575

Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
            580                 585                 590
```

Cys Phe

<210> 6
<211> 340
<212> PRT
<213> Artificial Sequence

<220>
<223> Sleeping Beauty Transposase

<400> 6

Met Gly Lys Ser Lys Glu Ile Ser Gln Asp Leu Arg Lys Lys Ile Val
1 5 10 15

Asp Leu His Lys Ser Gly Ser Ser Leu Gly Ala Ile Ser Lys Arg Leu
20 25 30

Lys Val Pro Arg Ser Ser Val Gln Thr Ile Val Arg Lys Tyr Lys His
35 40 45

His Gly Thr Thr Gln Pro Ser Tyr Arg Ser Gly Arg Arg Arg Tyr Leu
50 55 60

Ser Pro Arg Asp Glu Arg Thr Leu Val Arg Lys Val Gln Ile Asn Pro
65 70 75 80

Arg Thr Thr Ala Lys Asp Leu Val Lys Met Leu Glu Glu Thr Gly Thr
85 90 95

Lys Val Ser Ile Ser Thr Val Lys Arg Val Leu Tyr Arg His Asn Leu
100 105 110

Lys Gly Arg Ser Ala Arg Lys Lys Pro Leu Leu Gln Asn Arg His Lys
115 120 125

Lys Ala Arg Leu Arg Phe Ala Thr Ala His Gly Asp Lys Asp Arg Thr
130 135 140

Phe Trp Arg Asn Val Leu Trp Ser Asp Glu Thr Lys Ile Glu Leu Phe
145 150 155 160

```
Gly His Asn Asp His Arg Tyr Val Trp Arg Lys Lys Gly Glu Ala Cys
                165                 170                 175

Lys Pro Lys Asn Thr Ile Pro Thr Val Lys His Gly Gly Gly Ser Ile
            180                 185                 190

Met Leu Trp Gly Cys Phe Ala Ala Gly Gly Thr Gly Ala Leu His Lys
        195                 200                 205

Ile Asp Gly Ile Met Arg Lys Glu Asn Tyr Val Asp Ile Leu Lys Gln
    210                 215                 220

His Leu Lys Thr Ser Val Arg Lys Leu Lys Leu Gly Arg Lys Trp Val
225                 230                 235                 240

Phe Gln Met Asp Asn Asp Pro Lys His Thr Ser Lys Val Val Ala Lys
                245                 250                 255

Trp Leu Lys Asp Asn Lys Val Lys Val Leu Glu Trp Pro Ser Gln Ser
            260                 265                 270

Pro Asp Leu Asn Pro Ile Glu Asn Leu Trp Ala Glu Leu Lys Lys Arg
        275                 280                 285

Val Arg Ala Arg Arg Pro Thr Asn Leu Thr Gln Leu His Gln Leu Cys
    290                 295                 300

Gln Glu Glu Trp Ala Lys Ile His Pro Thr Tyr Cys Gly Lys Leu Val
305                 310                 315                 320

Glu Gly Tyr Pro Lys Arg Leu Thr Gln Val Lys Gln Phe Lys Gly Asn
                325                 330                 335

Ala Thr Lys Tyr
            340
```

<210> 7
<211> 340
<212> PRT
<213> Artificial Sequence

<220>

<223> hyperactive sleeping beauty transposase

<400> 7

```
Met Gly Lys Ser Lys Glu Ile Ser Gln Asp Leu Arg Lys Arg Ile Val
1               5                  10                  15

Asp Leu His Lys Ser Gly Ser Ser Leu Gly Ala Ile Ser Lys Arg Leu
            20                  25                  30

Ala Val Pro Arg Ser Ser Val Gln Thr Ile Val Arg Lys Tyr Lys His
        35                  40                  45

His Gly Thr Thr Gln Pro Ser Tyr Arg Ser Gly Arg Arg Arg Tyr Leu
    50                  55                  60

Ser Pro Arg Asp Glu Arg Thr Leu Val Arg Lys Val Gln Ile Asn Pro
65                  70                  75                  80

Arg Thr Thr Ala Lys Asp Leu Val Lys Met Leu Glu Glu Thr Gly Thr
                85                  90                  95

Lys Val Ser Ile Ser Thr Val Lys Arg Val Leu Tyr Arg His Asn Leu
            100                 105                 110

Lys Gly His Ser Ala Arg Lys Lys Pro Leu Leu Gln Asn Arg His Lys
        115                 120                 125

Lys Ala Arg Leu Arg Phe Ala Thr Ala His Gly Asp Lys Asp Arg Thr
    130                 135                 140

Phe Trp Arg Asn Val Leu Trp Ser Asp Glu Thr Lys Ile Glu Leu Phe
145                 150                 155                 160

Gly His Asn Asp His Arg Tyr Val Trp Arg Lys Lys Gly Glu Ala Cys
                165                 170                 175

Lys Pro Lys Asn Thr Ile Pro Thr Val Lys His Gly Gly Gly Ser Ile
            180                 185                 190

Met Leu Trp Gly Cys Phe Ala Ala Gly Gly Thr Gly Ala Leu His Lys
        195                 200                 205
```

```
Ile Asp Gly Ile Met Asp Ala Val Gln Tyr Val Asp Ile Leu Lys Gln
    210                 215                 220

His Leu Lys Thr Ser Val Arg Lys Leu Lys Leu Gly Arg Lys Trp Val
225                 230                 235                 240

Phe Gln His Asp Asn Asp Pro Lys His Thr Ser Lys Val Val Ala Lys
                245                 250                 255

Trp Leu Lys Asp Asn Lys Val Lys Val Leu Glu Trp Pro Ser Gln Ser
            260                 265                 270

Pro Asp Leu Asn Pro Ile Glu Asn Leu Trp Ala Glu Leu Lys Lys Arg
        275                 280                 285

Val Arg Ala Arg Arg Pro Thr Asn Leu Thr Gln Leu His Gln Leu Cys
    290                 295                 300

Gln Glu Glu Trp Ala Lys Ile His Pro Asn Tyr Cys Gly Lys Leu Val
305                 310                 315                 320

Glu Gly Tyr Pro Lys Arg Leu Thr Gln Val Lys Gln Phe Lys Gly Asn
                325                 330                 335

Ala Thr Lys Tyr
            340
```

<210> 8
<211> 21
<212> PRT
<213> Homo sapiens

<400> 8

```
Met Ala Leu Pro Val Thr Ala Leu Leu Leu Pro Leu Ala Leu Leu Leu
1               5                   10                  15

His Ala Ala Arg Pro
            20
```

<210> 9

<211> 63
<212> DNA
<213> Artificial Sequence

<220>
<223> nucleotide sequence for human CD8alpha signal peptide

<400> 9
atggcactgc cagtcaccgc cctgctgctg cctctggctc tgctgctgca cgcagctaga 60

cca 63

<210> 10
<211> 24
<212> PRT
<213> Homo sapiens

<400> 10

Ile Tyr Ile Trp Ala Pro Leu Ala Gly Thr Cys Gly Val Leu Leu Leu
1 5 10 15

Ser Leu Val Ile Thr Leu Tyr Cys
20

<210> 11
<211> 72
<212> DNA
<213> Artificial Sequence

<220>
<223> nucleotide sequence coding for human CD8alpha transmembrane domain

<400> 11
atctacattt gggcaccact ggccgggacc tgtggagtgc tgctgctgag cctggtcatc 60

acactgtact gc 72

<210> 12
<211> 112
<212> PRT
<213> Homo sapiens

<400> 12

Arg Val Lys Phe Ser Arg Ser Ala Asp Ala Pro Ala Tyr Lys Gln Gly
1 5 10 15

```
Gln Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr
            20                  25                  30

Asp Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys
        35                  40                  45

Pro Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys
    50                  55                  60

Asp Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg
65                  70                  75                  80

Arg Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala
                85                  90                  95

Thr Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg
            100                 105                 110
```

<210> 13
<211> 336
<212> DNA
<213> Artificial Sequenec

<220>
<223> nucleotide sequence encoding the CD28 costimulatory domain

<400> 13

```
cgcgtgaagt ttagtcgatc agcagatgcc ccagcttaca aacagggaca gaaccagctg       60

tataacgagc tgaatctggg ccgccgagag gaatatgacg tgctggataa gcggagagga      120

cgcgaccccg aaatgggagg caagcccagg cgcaaaaacc ctcaggaagg cctgtataac      180

gagctgcaga aggacaaaat ggcagaagcc tattctgaga tcggcatgaa gggggagcga      240

cggagaggca aagggcacga tgggctgtac cagggactga gcaccgccac aaaggacacc      300

tatgatgctc tgcatatgca ggcactgcct ccaagg                                336
```

<210> 14
<211> 42
<212> PRT
<213> Homo sapiens

<400> 14

```
Lys Arg Gly Arg Lys Lys Leu Leu Tyr Ile Phe Lys Gln Pro Phe Met
```

1 5 10 15

Arg Pro Val Gln Thr Thr Gln Glu Glu Asp Gly Cys Ser Cys Arg Phe
20 25 30

Pro Glu Glu Glu Glu Gly Gly Cys Glu Leu
35 40

<210> 15
<211> 126
<212> DNA
<213> Artificial Sequence

<220>
<223> nucleotide sequence encoding the 4-1BB costimulatory domain

<400> 15
aagagaggca ggaagaaact gctgtatatt ttcaaacagc ccttcatgcg ccccgtgcag 60

actacccagg aggaagacgg gtgctcctgt cgattccctg aggaagagga aggcgggtgt 120

gagctg 126

<210> 16
<211> 45
<212> PRT
<213> Homo sapiens

<400> 16

Thr Thr Thr Pro Ala Pro Arg Pro Pro Thr Pro Ala Pro Thr Ile Ala
1 5 10 15

Ser Gln Pro Leu Ser Leu Arg Pro Glu Ala Cys Arg Pro Ala Ala Gly
20 25 30

Gly Ala Val His Thr Arg Gly Leu Asp Phe Ala Cys Asp
35 40 45

<210> 17
<211> 135
<212> DNA
<213> Artificial Sequence

<220>
<223> nucleotide sequence encoding human CD8alpha hinge

<400> 17
actaccacac cagcacctag accaccaact ccagctccaa ccatcgcgag tcagcccctg 60

agtctgagac ctgaggcctg caggccagct gcaggaggag ctgtgcacac caggggcctg 120

gacttcgcct gcgac 135

<210> 18
<211> 18
<212> PRT
<213> Thosea asigna

<400> 18

Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn Pro
1 5 10 15

Gly Pro

<210> 19
<211> 21
<212> PRT
<213> Artificial Sequence

<220>
<223> GSG-T2A

<400> 19

Gly Ser Gly Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu
1 5 10 15

Glu Asn Pro Gly Pro
20

<210> 20
<211> 63
<212> DNA
<213> Artificial Sequence

<220>
<223> GSG-T2A

<400> 20
ggatctggag agggaagggg aagcctgctg acctgtggag acgtggagga aaacccagga 60

cca 63

<210> 21
<211> 20
<212> PRT
<213> Equine rhinitis A

<400> 21

Gln Cys Thr Asn Tyr Ala Leu Leu Lys Leu Ala Gly Asp Val Glu Ser
1   5   10   15

Asn Pro Gly Pro
            20

<210> 22
<211> 23
<212> PRT
<213> Artificial Sequence

<220>
<223> GSG-E2A peptide

<400> 22

Gly Ser Gly Gln Cys Thr Asn Tyr Ala Leu Leu Lys Leu Ala Gly Asp
1   5   10   15

Val Glu Ser Asn Pro Gly Pro
            20

<210> 23
<211> 22
<212> PRT
<213> Foot and nout disease virus type O

<400> 23

Val Lys Gln Thr Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val
1   5   10   15

Glu Ser Asn Pro Gly Pro
            20

<210> 24
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> GSG-F2A peptide

<400> 24

Gly Ser Gly Val Lys Gln Thr Leu Asn Phe Asp Leu Leu Lys Leu Ala
1   5   10   15

Gly Asp Val Glu Ser Asn Pro Gly Pro
            20   25

<210> 25
<211> 19
<212> PRT
<213> Porcine teschovirus-1

<400> 25

Ala Thr Asn Phe Ser Leu Leu Lys Gln Ala Gly Asp Val Glu Glu Asn
1   5   10   15

Pro Gly Pro

<210> 26
<211> 22
<212> PRT
<213> Artificial Sequence

<220>
<223> GSG-P2A peptide

<400> 26

Gly Ser Gly Ala Thr Asn Phe Ser Leu Leu Lys Gln Ala Gly Asp Val
1   5   10   15

Glu Glu Asn Pro Gly Pro
            20

<210> 27
<211> 5296
<212> DNA
<213> Artificial Sequence

<220>
<223> Helraiser transposon

<400> 27
tcctatataa taaaagagaa acatgcaaat tgaccatccc tccgctacgc tcaagccacg 60

cccaccagcc aatcagaagt gactatgcaa attaacccaa caaagatggc agttaaattt 120

gcatacgcag gtgtcaagcg ccccaggagg caacggcggc cgcgggctcc caggaccttc 180

gctggccccg ggaggcgagg ccggccgcgc ctagccacac ccgcgggctc ccgggacctt 240

cgccagcaga gagcagagcg ggagagcggg cggagagcgg gaggtttgga ggacttggca 300

gagcaggagg ccgctggaca tagagcagag cgagagagag ggtggcttgg agggcgtggc 360

tccctctgtc accccagctt cctcatcaca gctgtggaaa ctgacagcag ggaggaggaa 420

gtcccacccc cacagaatca gccagaatca gccgttggtc agacagctct cagcggcctg 480

acagccagga ctctcattca cctgcatctc agaccgtgac agtagagagg tgggactatg 540

tctaaagaac aactgttgat acaacgtagc tctgcagccg aaagatgccg gcgttatcga 600

cagaaaatgt ctgcagagca acgtgcgtct gatcttgaaa gaaggcggcg cctgcaacag 660

aatgtatctg aagagcagct actggaaaaa cgtcgctctg aagccgaaaa acagcggcgt 720

catcgacaga aaatgtctaa agaccaacgt gcctttgaag ttgaaagaag gcggtggcga 780

cgacagaata tgtctagaga acagtcatca acaagtacta ccaataccgg taggaactgc 840

cttctcagca aaaatggagt acatgaggat gcaattctcg aacatagttg tggtggaatg 900

actgttcgat gtgaattttg cctatcacta aatttctctg atgaaaaacc atccgatggg 960

aaatttactc gatgttgtag caaagggaaa gtctgtccaa atgatataca ttttccagat 1020

tacccggcat atttaaaaag attaatgaca aacgaagatt ctgacagtaa aaatttcatg 1080

gaaaatattc gttccataaa tagttctttt gcttttgctt ccatgggtgc aaatattgca 1140

tcgccatcag gatatgggcc atactgtttt agaatacacg gacaagttta tcaccgtact 1200

ggaactttac atccttcgga tggtgtttct cggaagtttg ctcaactcta tattttggat 1260

acagccgaag ctacaagtaa aagattagca atgccagaaa accagggctg ctcagaaaga 1320

ctcatgatca acatcaacaa cctcatgcat gaaataaatg aattaacaaa atcgtacaag 1380

atgctacatg aggtagaaaa ggaagcccaa tctgaagcag cagcaaaagg tattgctccc 1440

acagaagtaa caatggcgat taaatacgat cgtaacagtg acccaggtag atataattct 1500

ccccgtgtaa ccgaggttgc tgtcatattc agaaacgaag atggagaacc tccttttgaa 1560

agggacttgc tcattcattg taaaccagat cccaataatc caaatgccac taaaatgaaa 1620

caaatcagta tcctgtttcc tacattagat gcaatgacat atcctattct ttttccacat 1680

ggtgaaaaag gctggggaac agatattgca ttaagactca gagacaacag tgtaatcgac 1740

aataatacta gacaaaatgt aaggacacga gtcacacaaa tgcagtatta tggatttcat 1800

ctctctgtgc gggacacgtt caatcctatt ttaaatgcag gaaaattaac tcaacagttt 1860

attgtggatt catattcaaa aatggaggcc aatcggataa atttcatcaa agcaaaccaa 1920

tctaagttga gagttgaaaa atatagtggt ttgatggatt atctcaaatc tagatctgaa 1980

aatgacaatg tgccgattgg taaaatgata atacttccat catcttttga gggtagtccc 2040

agaaatatgc agcagcgata tcaggatgct atggcaattg taacgaagta tggcaagccc 2100

gatttattca taaccatgac atgcaacccc aaatgggcag atattacaaa caatttacaa 2160

cgctggcaaa aagttgaaaa cagacctgac ttggtagcca gagtttttaa tattaagctg 2220

aatgctcttt taaatgatat atgtaaattc catttatttg gcaaagtaat agctaaaatt 2280

catgtcattg aatttcagaa acgcggactg cctcacgctc acatattatt gatattagat 2340

agtgagtcca aattacgttc agaagatgac attgaccgta tagttaaggc agaaattcca 2400

gatgaagacc agtgtcctcg actttttcaa attgtaaaat caaatatggt acatggacca 2460

tgtggaatac aaaatccaaa tagtccatgt atggaaaatg gaaaatgttc aaagggatat 2520

ccaaaagaat ttcaaaatgc gaccattgga aatattgatg gatatcccaa atacaaacga 2580

agatctggta gcaccatgtc tattggaaat aaagttgtcg ataacacttg gattgtccct 2640

tataacccgt atttgtgcct taaatataac tgtcatataa atgttgaagt ctgtgcatca 2700

attaaaagtg tcaaatattt atttaaatac atctataaag ggcacgattg tgcaaatatt 2760

caaatttctg aaaaaaatat tatcaatcat gacgaagtac aggacttcat tgactccagg 2820

tatgtgagcg ctcctgaggc tgtttggaga ctttttgcaa tgcgaatgca tgaccaatct 2880

catgcaatca caagattagc tattcatttg ccaaatgatc agaatttgta ttttcatacc 2940

gatgattttg ctgaagtttt agatagggct aaaaggcata actcgacttt gatggcttgg 3000

ttcttattga atagagaaga ttctgatgca cgtaattatt attattggga gattccacag 3060

cattatgtgt ttaataattc tttgtggaca aaacgccgaa agggtgggaa taaagtatta 3120

ggtagactgt tcactgtgag ctttagagaa ccagaacgat attaccttag acttttgctt 3180

ctgcatgtaa aaggtgcgat aagttttgag gatctgcgaa ctgtaggagg tgtaacttat 3240

gatacatttc atgaagctgc taaacaccga ggattattac ttgatgacac tatctggaaa 3300

gatacgattg acgatgcaat catccttaat atgcccaaac aactacggca acttttgca 3360

tatatatgtg tgtttggatg tccttctgct gcagacaaat tatgggatga gaataaatct 3420

cattttattg aagatttctg ttggaaatta caccgaagag aaggtgcctg tgtgaactgt 3480

gaaatgcatg cccttaacga aattcaggag gtattcacat tgcatggaat gaaatgttca 3540

catttcaaac ttccggacta tcctttatta atgaatgcaa atacatgtga tcaattgtac 3600

gagcaacaac aggcagaggt tttgataaat tctctgaatg atgaacagtt ggcagccttt 3660

cagactataa cttcagccat cgaagatcaa actgtacacc ccaaatgctt tttcttggat 3720

ggtccaggtg gtagtggaaa aacatatctg tataaagttt taacacatta tattagaggt 3780

cgtggtggta ctgttttacc cacagcatct acaggaattg ctgcaaattt acttcttggt 3840

ggaagaacct ttcattccca atataaatta ccaattccat taaatgaaac ttcaatttct 3900

agactcgata taaagagtga agttgctaaa accattaaaa aggcccaact tctcattatt 3960

gatgaatgca ccatggcatc cagtcatgct ataaacgcca tagatagatt actaagagaa 4020

attatgaatt tgaatgttgc atttggtggg aaagttctcc ttctcggagg ggattttcga 4080

caatgtctca gtattgtacc acatgctatg cgatcggcca tagtacaaac gagtttaaag 4140

tactgtaatg tttggggatg tttcagaaag ttgtctctta aaacaaatat gagatcagag 4200

gattctgctt atagtgaatg gttagtaaaa cttggagatg gcaaacttga tagcagtttt 4260

catttaggaa tggatattat tgaaatcccc catgaaatga tttgtaacgg atctattatt 4320

gaagctacct ttggaaatag tatatctata gataatatta aaaatatatc taaacgtgca 4380

attctttgtc caaaaaatga gcatgttcaa aaattaaatg aagaaatttt ggatatactt 4440

gatggagatt ttcacacata tttgagtgat gattccattg attcaacaga tgatgctgaa 4500

aaggaaaatt ttcccatcga atttcttaat agtattactc cttcgggaat gccgtgtcat 4560

aaattaaaat tgaaagtggg tgcaatcatc atgctattga gaaatcttaa tagtaaatgg 4620

ggtctttgta atggtactag atttattatc aaaagattac gacctaacat tatcgaagct 4680

gaagtattaa caggatctgc agagggagag gttgttctga ttccaagaat tgatttgtcc 4740

ccatctgaca ctggcctccc atttaaatta attcgaagac agtttcccgt gatgccagca 4800

tttgcgatga ctattaataa atcacaagga caaactctag acagagtagg aatattccta 4860

cctgaacccg ttttcgcaca tggtcagtta tatgttgctt tctctcgagt tcgaagagca 4920

tgtgacgtta aagttaaagt tgtaaatact tcatcacaag ggaaattagt caagcactct 4980

gaaagtgttt ttactcttaa tgtggtatac agggagatat tagaataagt ttaatcactt 5040

tatcagtcat tgtttgcatc aatgttgttt ttatatcatg tttttgttgt ttttatatca 5100

tgtctttgtt gttgttatat catgttgtta ttgtttattt attaataaat ttatgtatta 5160

ttttcatata cattttactc atttcctttc atctctcaca cttctattat agagaaaggg 5220

caaatagcaa tattaaaata tttcctctaa ttaattccct ttcaatgtgc acgaatttcg 5280

tgcaccgggc cactag 5296

<210> 28
<211> 1496
<212> PRT
<213> Artificial Sequence

<220>
<223> Helitron transposase

<400> 28

Met Ser Lys Glu Gln Leu Leu Ile Gln Arg Ser Ser Ala Ala Glu Arg
1 5 10 15

Cys Arg Arg Tyr Arg Gln Lys Met Ser Ala Glu Gln Arg Ala Ser Asp
20 25 30

Leu Glu Arg Arg Arg Arg Leu Gln Gln Asn Val Ser Glu Glu Gln Leu
35 40 45

Leu Glu Lys Arg Arg Ser Glu Ala Glu Lys Gln Arg Arg His Arg Gln
50 55 60

Lys Met Ser Lys Asp Gln Arg Ala Phe Glu Val Glu Arg Arg Arg Trp
65 70 75 80

Arg Arg Gln Asn Met Ser Arg Glu Gln Ser Ser Thr Ser Thr Thr Asn
85 90 95

Thr Gly Arg Asn Cys Leu Leu Ser Lys Asn Gly Val His Glu Asp Ala
100 105 110

```
Ile Leu Glu His Ser Cys Gly Gly Met Thr Val Arg Cys Glu Phe Cys
        115                 120                 125

Leu Ser Leu Asn Phe Ser Asp Glu Lys Pro Ser Asp Gly Lys Phe Thr
    130                 135                 140

Arg Cys Cys Ser Lys Gly Lys Val Cys Pro Asn Asp Ile His Phe Pro
145                 150                 155                 160

Asp Tyr Pro Ala Tyr Leu Lys Arg Leu Met Thr Asn Glu Asp Ser Asp
                165                 170                 175

Ser Lys Asn Phe Met Glu Asn Ile Arg Ser Ile Asn Ser Ser Phe Ala
            180                 185                 190

Phe Ala Ser Met Gly Ala Asn Ile Ala Ser Pro Ser Gly Tyr Gly Pro
        195                 200                 205

Tyr Cys Phe Arg Ile His Gly Gln Val Tyr His Arg Thr Gly Thr Leu
    210                 215                 220

His Pro Ser Asp Gly Val Ser Arg Lys Phe Ala Gln Leu Tyr Ile Leu
225                 230                 235                 240

Asp Thr Ala Glu Ala Thr Ser Lys Arg Leu Ala Met Pro Glu Asn Gln
                245                 250                 255

Gly Cys Ser Glu Arg Leu Met Ile Asn Ile Asn Asn Leu Met His Glu
            260                 265                 270

Ile Asn Glu Leu Thr Lys Ser Tyr Lys Met Leu His Glu Val Glu Lys
        275                 280                 285

Glu Ala Gln Ser Glu Ala Ala Ala Lys Gly Ile Ala Pro Thr Glu Val
    290                 295                 300

Thr Met Ala Ile Lys Tyr Asp Arg Asn Ser Asp Pro Gly Arg Tyr Asn
305                 310                 315                 320
```

```
Ser Pro Arg Val Thr Glu Val Ala Val Ile Phe Arg Asn Glu Asp Gly
                325                 330                 335

Glu Pro Pro Phe Glu Arg Asp Leu Leu Ile His Cys Lys Pro Asp Pro
            340                 345                 350

Asn Asn Pro Asn Ala Thr Lys Met Lys Gln Ile Ser Ile Leu Phe Pro
        355                 360                 365

Thr Leu Asp Ala Met Thr Tyr Pro Ile Leu Phe Pro His Gly Glu Lys
    370                 375                 380

Gly Trp Gly Thr Asp Ile Ala Leu Arg Leu Arg Asp Asn Ser Val Ile
385                 390                 395                 400

Asp Asn Asn Thr Arg Gln Asn Val Arg Thr Arg Val Thr Gln Met Gln
                405                 410                 415

Tyr Tyr Gly Phe His Leu Ser Val Arg Asp Thr Phe Asn Pro Ile Leu
            420                 425                 430

Asn Ala Gly Lys Leu Thr Gln Gln Phe Ile Val Asp Ser Tyr Ser Lys
        435                 440                 445

Met Glu Ala Asn Arg Ile Asn Phe Ile Lys Ala Asn Gln Ser Lys Leu
    450                 455                 460

Arg Val Glu Lys Tyr Ser Gly Leu Met Asp Tyr Leu Lys Ser Arg Ser
465                 470                 475                 480

Glu Asn Asp Asn Val Pro Ile Gly Lys Met Ile Ile Leu Pro Ser Ser
                485                 490                 495

Phe Glu Gly Ser Pro Arg Asn Met Gln Gln Arg Tyr Gln Asp Ala Met
            500                 505                 510

Ala Ile Val Thr Lys Tyr Gly Lys Pro Asp Leu Phe Ile Thr Met Thr
        515                 520                 525

Cys Asn Pro Lys Trp Ala Asp Ile Thr Asn Asn Leu Gln Arg Trp Gln
    530                 535                 540
```

```
Lys Val Glu Asn Arg Pro Asp Leu Val Ala Arg Val Phe Asn Ile Lys
545                 550                 555                 560

Leu Asn Ala Leu Leu Asn Asp Ile Cys Lys Phe His Leu Phe Gly Lys
                565                 570                 575

Val Ile Ala Lys Ile His Val Ile Glu Phe Gln Lys Arg Gly Leu Pro
            580                 585                 590

His Ala His Ile Leu Leu Ile Leu Asp Ser Glu Ser Lys Leu Arg Ser
        595                 600                 605

Glu Asp Asp Ile Asp Arg Ile Val Lys Ala Glu Ile Pro Asp Glu Asp
    610                 615                 620

Gln Cys Pro Arg Leu Phe Gln Ile Val Lys Ser Asn Met Val His Gly
625                 630                 635                 640

Pro Cys Gly Ile Gln Asn Pro Asn Ser Pro Cys Met Glu Asn Gly Lys
                645                 650                 655

Cys Ser Lys Gly Tyr Pro Lys Glu Phe Gln Asn Ala Thr Ile Gly Asn
            660                 665                 670

Ile Asp Gly Tyr Pro Lys Tyr Lys Arg Arg Ser Gly Ser Thr Met Ser
        675                 680                 685

Ile Gly Asn Lys Val Val Asp Asn Thr Trp Ile Val Pro Tyr Asn Pro
    690                 695                 700

Tyr Leu Cys Leu Lys Tyr Asn Cys His Ile Asn Val Glu Val Cys Ala
705                 710                 715                 720

Ser Ile Lys Ser Val Lys Tyr Leu Phe Lys Tyr Ile Tyr Lys Gly His
                725                 730                 735

Asp Cys Ala Asn Ile Gln Ile Ser Glu Lys Asn Ile Ile Asn His Asp
            740                 745                 750
```

```
Glu Val Gln Asp Phe Ile Asp Ser Arg Tyr Val Ser Ala Pro Glu Ala
        755                 760                 765

Val Trp Arg Leu Phe Ala Met Arg Met His Asp Gln Ser His Ala Ile
    770                 775                 780

Thr Arg Leu Ala Ile His Leu Pro Asn Asp Gln Asn Leu Tyr Phe His
785                 790                 795                 800

Thr Asp Asp Phe Ala Glu Val Leu Asp Arg Ala Lys Arg His Asn Ser
                805                 810                 815

Thr Leu Met Ala Trp Phe Leu Leu Asn Arg Glu Asp Ser Asp Ala Arg
            820                 825                 830

Asn Tyr Tyr Tyr Trp Glu Ile Pro Gln His Tyr Val Phe Asn Asn Ser
        835                 840                 845

Leu Trp Thr Lys Arg Arg Lys Gly Gly Asn Lys Val Leu Gly Arg Leu
    850                 855                 860

Phe Thr Val Ser Phe Arg Glu Pro Glu Arg Tyr Tyr Leu Arg Leu Leu
865                 870                 875                 880

Leu Leu His Val Lys Gly Ala Ile Ser Phe Glu Asp Leu Arg Thr Val
                885                 890                 895

Gly Gly Val Thr Tyr Asp Thr Phe His Glu Ala Ala Lys His Arg Gly
            900                 905                 910

Leu Leu Leu Asp Asp Thr Ile Trp Lys Asp Thr Ile Asp Asp Ala Ile
        915                 920                 925

Ile Leu Asn Met Pro Lys Gln Leu Arg Gln Leu Phe Ala Tyr Ile Cys
    930                 935                 940

Val Phe Gly Cys Pro Ser Ala Ala Asp Lys Leu Trp Asp Glu Asn Lys
945                 950                 955                 960

Ser His Phe Ile Glu Asp Phe Cys Trp Lys Leu His Arg Arg Glu Gly
                965                 970                 975
```

```
Ala Cys Val Asn Cys Glu Met His Ala Leu Asn Glu Ile Gln Glu Val
            980                 985                 990


Phe Thr Leu His Gly Met Lys Cys  Ser His Phe Lys Leu  Pro Asp Tyr
        995                 1000                1005


Pro Leu  Leu Met Asn Ala Asn  Thr Cys Asp Gln Leu  Tyr Glu Gln
    1010                1015                1020


Gln Gln  Ala Glu Val Leu Ile  Asn Ser Leu Asn Asp  Glu Gln Leu
    1025                1030                1035


Ala Ala  Phe Gln Thr Ile Thr  Ser Ala Ile Glu Asp  Gln Thr Val
    1040                1045                1050


His Pro  Lys Cys Phe Phe Leu  Asp Gly Pro Gly Gly  Ser Gly Lys
    1055                1060                1065


Thr Tyr  Leu Tyr Lys Val Leu  Thr His Tyr Ile Arg  Gly Arg Gly
    1070                1075                1080


Gly Thr  Val Leu Pro Thr Ala  Ser Thr Gly Ile Ala  Ala Asn Leu
    1085                1090                1095


Leu Leu  Gly Gly Arg Thr Phe  His Ser Gln Tyr Lys  Leu Pro Ile
    1100                1105                1110


Pro Leu  Asn Glu Thr Ser Ile  Ser Arg Leu Asp Ile  Lys Ser Glu
    1115                1120                1125


Val Ala  Lys Thr Ile Lys Lys  Ala Gln Leu Leu Ile  Ile Asp Glu
    1130                1135                1140


Cys Thr  Met Ala Ser Ser His  Ala Ile Asn Ala Ile  Asp Arg Leu
    1145                1150                1155


Leu Arg  Glu Ile Met Asn Leu  Asn Val Ala Phe Gly  Gly Lys Val
    1160                1165                1170
```

```
Leu Leu  Leu Gly Gly Asp Phe  Arg Gln Cys Leu Ser  Ile Val Pro
    1175                1180                1185

His Ala  Met Arg Ser Ala Ile  Val Gln Thr Ser Leu  Lys Tyr Cys
    1190                1195                1200

Asn Val  Trp Gly Cys Phe Arg  Lys Leu Ser Leu Lys  Thr Asn Met
    1205                1210                1215

Arg Ser  Glu Asp Ser Ala Tyr  Ser Glu Trp Leu Val  Lys Leu Gly
    1220                1225                1230

Asp Gly  Lys Leu Asp Ser Ser  Phe His Leu Gly Met  Asp Ile Ile
    1235                1240                1245

Glu Ile  Pro His Glu Met Ile  Cys Asn Gly Ser Ile  Ile Glu Ala
    1250                1255                1260

Thr Phe  Gly Asn Ser Ile Ser  Ile Asp Asn Ile Lys  Asn Ile Ser
    1265                1270                1275

Lys Arg  Ala Ile Leu Cys Pro  Lys Asn Glu His Val  Gln Lys Leu
    1280                1285                1290

Asn Glu  Glu Ile Leu Asp Ile  Leu Asp Gly Asp Phe  His Thr Tyr
    1295                1300                1305

Leu Ser  Asp Asp Ser Ile Asp  Ser Thr Asp Asp Ala  Glu Lys Glu
    1310                1315                1320

Asn Phe  Pro Ile Glu Phe Leu  Asn Ser Ile Thr Pro  Ser Gly Met
    1325                1330                1335

Pro Cys  His Lys Leu Lys Leu  Lys Val Gly Ala Ile  Ile Met Leu
    1340                1345                1350

Leu Arg  Asn Leu Asn Ser Lys  Trp Gly Leu Cys Asn  Gly Thr Arg
    1355                1360                1365

Phe Ile  Ile Lys Arg Leu Arg  Pro Asn Ile Ile Glu  Ala Glu Val
    1370                1375                1380
```

Leu Thr Gly Ser Ala Glu Gly Glu Val Val Leu Ile Pro Arg Ile
1385 1390 1395

Asp Leu Ser Pro Ser Asp Thr Gly Leu Pro Phe Lys Leu Ile Arg
1400 1405 1410

Arg Gln Phe Pro Val Met Pro Ala Phe Ala Met Thr Ile Asn Lys
1415 1420 1425

Ser Gln Gly Gln Thr Leu Asp Arg Val Gly Ile Phe Leu Pro Glu
1430 1435 1440

Pro Val Phe Ala His Gly Gln Leu Tyr Val Ala Phe Ser Arg Val
1445 1450 1455

Arg Arg Ala Cys Asp Val Lys Val Lys Val Val Asn Thr Ser Ser
1460 1465 1470

Gln Gly Lys Leu Val Lys His Ser Glu Ser Val Phe Thr Leu Asn
1475 1480 1485

Val Val Tyr Arg Glu Ile Leu Glu
1490 1495

<210> 29
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Helraiser palindromic sequence

<400> 29
gtgcacgaat ttcgtgcacc gggccactag 30

<210> 30
<211> 649
<212> PRT
<213> Oryzias latipes

<400> 30

Met Glu Glu Val Cys Asp Ser Ser Ala Ala Ala Ser Ser Thr Val Gln

```
1               5                   10                  15

Asn Gln Pro Gln Asp Gln Glu His Pro Trp Pro Tyr Leu Arg Glu Phe
            20                  25                  30

Phe Ser Leu Ser Gly Val Asn Lys Asp Ser Phe Lys Met Lys Cys Val
        35                  40                  45

Leu Cys Leu Pro Leu Asn Lys Glu Ile Ser Ala Phe Lys Ser Ser Pro
    50                  55                  60

Ser Asn Leu Arg Lys His Ile Glu Arg Met His Pro Asn Tyr Leu Lys
65                  70                  75                  80

Asn Tyr Ser Lys Leu Thr Ala Gln Lys Arg Lys Ile Gly Thr Ser Thr
                85                  90                  95

His Ala Ser Ser Ser Lys Gln Leu Lys Val Asp Ser Val Phe Pro Val
            100                 105                 110

Lys His Val Ser Pro Val Thr Val Asn Lys Ala Ile Leu Arg Tyr Ile
        115                 120                 125

Ile Gln Gly Leu His Pro Phe Ser Thr Val Asp Leu Pro Ser Phe Lys
    130                 135                 140

Glu Leu Ile Ser Thr Leu Gln Pro Gly Ile Ser Val Ile Thr Arg Pro
145                 150                 155                 160

Thr Leu Arg Ser Lys Ile Ala Glu Ala Ala Leu Ile Met Lys Gln Lys
                165                 170                 175

Val Thr Ala Ala Met Ser Glu Val Glu Trp Ile Ala Thr Thr Thr Asp
            180                 185                 190

Cys Trp Thr Ala Arg Arg Lys Ser Phe Ile Gly Val Thr Ala His Trp
        195                 200                 205

Ile Asn Pro Gly Ser Leu Glu Arg His Ser Ala Ala Leu Ala Cys Lys
    210                 215                 220
```

```
Arg Leu Met Gly Ser His Thr Phe Glu Val Leu Ala Ser Ala Met Asn
225                 230                 235                 240

Asp Ile His Ser Glu Tyr Glu Ile Arg Asp Lys Val Val Cys Thr Thr
                245                 250                 255

Thr Asp Ser Gly Ser Asn Phe Met Lys Ala Phe Arg Val Phe Gly Val
            260                 265                 270

Glu Asn Asn Asp Ile Glu Thr Glu Ala Arg Arg Cys Glu Ser Asp Asp
        275                 280                 285

Thr Asp Ser Glu Gly Cys Gly Glu Gly Ser Asp Gly Val Glu Phe Gln
    290                 295                 300

Asp Ala Ser Arg Val Leu Asp Gln Asp Asp Gly Phe Glu Phe Gln Leu
305                 310                 315                 320

Pro Lys His Gln Lys Cys Ala Cys His Leu Leu Asn Leu Val Ser Ser
                325                 330                 335

Val Asp Ala Gln Lys Ala Leu Ser Asn Glu His Tyr Lys Lys Leu Tyr
            340                 345                 350

Arg Ser Val Phe Gly Lys Cys Gln Ala Leu Trp Asn Lys Ser Ser Arg
        355                 360                 365

Ser Ala Leu Ala Ala Glu Ala Val Glu Ser Glu Ser Arg Leu Gln Leu
    370                 375                 380

Leu Arg Pro Asn Gln Thr Arg Trp Asn Ser Thr Phe Met Ala Val Asp
385                 390                 395                 400

Arg Ile Leu Gln Ile Cys Lys Glu Ala Gly Glu Gly Ala Leu Arg Asn
                405                 410                 415

Ile Cys Thr Ser Leu Glu Val Pro Met Phe Asn Pro Ala Glu Met Leu
            420                 425                 430

Phe Leu Thr Glu Trp Ala Asn Thr Met Arg Pro Val Ala Lys Val Leu
```

```
        435                 440                 445

Asp Ile Leu Gln Ala Glu Thr Asn Thr Gln Leu Gly Trp Leu Leu Pro
    450                 455                 460

Ser Val His Gln Leu Ser Leu Lys Leu Gln Arg Leu His His Ser Leu
465                 470                 475                 480

Arg Tyr Cys Asp Pro Leu Val Asp Ala Leu Gln Gln Gly Ile Gln Thr
                485                 490                 495

Arg Phe Lys His Met Phe Glu Asp Pro Glu Ile Ile Ala Ala Ala Ile
            500                 505                 510

Leu Leu Pro Lys Phe Arg Thr Ser Trp Thr Asn Asp Glu Thr Ile Ile
        515                 520                 525

Lys Arg Gly Met Asp Tyr Ile Arg Val His Leu Glu Pro Leu Asp His
    530                 535                 540

Lys Lys Glu Leu Ala Asn Ser Ser Ser Asp Asp Glu Asp Phe Phe Ala
545                 550                 555                 560

Ser Leu Lys Pro Thr Thr His Glu Ala Ser Lys Glu Leu Asp Gly Tyr
                565                 570                 575

Leu Ala Cys Val Ser Asp Thr Arg Glu Ser Leu Leu Thr Phe Pro Ala
            580                 585                 590

Ile Cys Ser Leu Ser Ile Lys Thr Asn Thr Pro Leu Pro Ala Ser Ala
        595                 600                 605

Ala Cys Glu Arg Leu Phe Ser Thr Ala Gly Leu Leu Phe Ser Pro Lys
    610                 615                 620

Arg Ala Arg Leu Asp Thr Asn Asn Phe Glu Asn Gln Leu Leu Leu Lys
625                 630                 635                 640

Leu Asn Leu Arg Phe Tyr Asn Phe Glu
                645
```

<210> 31
<211> 4682
<212> DNA
<213> Oryzias latipes

<400> 31
cagaggtgta aagtacttga gtaattttac ttgattactg tacttaagta ttatttttgg 60

ggatttttac tttacttgag tacaattaaa aatcaatact tttactttta cttaattaca 120

ttttttaga aaaaaaagta ctttttactc cttacaattt tatttacagt caaaaagtac 180

ttatttttg gagatcactt cattctattt tcccttgcta ttaccaaacc aattgaattg 240

cgctgatgcc cagtttaatt taaatgttat ttattctgcc tatgaaaatc gttttcacat 300

tatatgaaat tggtcagaca tgttcattgg tcctttggaa gtgacgtcat gtcacatcta 360

ttaccacaat gcacagcacc ttgacctgga aattagggaa attataacag tcaatcagtg 420

gaagaaaatg gaggaagtat gtgattcatc agcagctgcg agcagcacag tccaaaatca 480

gccacaggat caagagcacc cgtggccgta tcttcgcgaa ttcttttctt taagtggtgt 540

aaataaagat tcattcaaga tgaaatgtgt cctctgtctc ccgcttaata aagaaatatc 600

ggccttcaaa agttcgccat caaacctaag gaagcatatt gaggtaagta cattaagtat 660

tttgttttac tgatagtttt ttttttttt ttttttttt tttttgggtg tgcatgtttt 720

gacgttgatg gcgcgccttt tatatgtgta gtaggcctat tttcactaat gcatgcgatt 780

gacaatataa ggctcacgta ataaaatgct aaaatgcatt tgtaattggt aacgttaggt 840

ccacgggaaa tttggcgcct attgcagctt tgaataatca ttatcattcc gtgctctcat 900

tgtgtttgaa ttcatgcaaa acacaagaaa accaagcgag aaatttttt ccaaacatgt 960

tgtattgtca aaacggtaac actttacaat gaggttgatt agttcatgta ttaactaaca 1020

ttaaataacc atgagcaata catttgttac tgtatctgtt aatctttgtt aacgttagtt 1080

aatagaaata cagatgttca ttgtttgttc atgttagttc acagtgcatt aactaatgtt 1140

aacaagatat aaagtattag taaatgttga aattaacatg tatacgtgca gttcattatt 1200

agttcatgtt aactaatgta gttaactaac gaaccttatt gtaaaagtgt taccatcaaa 1260

actaatgtaa tgaaatcaat tcaccctgtc atgtcagcct tacagtcctg tgtttttgtc 1320

aatataatca gaaataaaat taatgtttga ttgtcactaa atgctactgt atttctaaaa 1380

tcaacaagta tttaacatta taaagtgtgc aattggctgc aaatgtcagt tttattaaag 1440

ggttagttca cccaaaaatg aaaataatgt cattaatgac tcgccctcat gtcgttccaa 1500

gcccgtaaga cctccgttca tcttcagaac acagtttaag atattttaga tttagtccga 1560

gagctttctg tgcctccatt gagaatgtat gtacggtata ctgtccatgt ccagaaaggt 1620

aataaaaaca tcaaagtagt ccatgtgaca tcagtgggtt agttagaatt ttttgaagca 1680

tcgaatacat tttggtccaa aaataacaaa acctacgact ttattcggca ttgtattctc 1740

ttccgggtct gttgtcaatc cgcgttcacg acttcgcagt gacgctacaa tgctgaataa 1800

agtcgtaggt tttgttattt ttggaccaaa atgtattttc gatgcttcaa ataattctac 1860

ctaacccact gatgtcacat ggactacttt gatgttttta ttacctttct ggacatggac 1920

agtataccgt acatacattt tcagtggagg gacagaaagc tctcggacta aatctaaaat 1980

atcttaaact gtgttccgaa gatgaacgga ggtgttacgg gcttggaacg acatgagggt 2040

gagtcattaa tgacatcttt tcatttttgg gtgaactaac cctttaatgc tgtaatcaga 2100

gagtgtatgt gtaattgtta catttattgc atacaatata aatatttatt tgttgttttt 2160

acagagaatg cacccaaatt acctcaaaaa ctactctaaa ttgacagcac agaagagaaa 2220

gatcgggacc tccacccatg cttccagcag taagcaactg aaagttgact cagttttccc 2280

agtcaaacat gtgtctccag tcactgtgaa caaagctata ttaaggtaca tcattcaagg 2340

acttcatcct ttcagcactg ttgatctgcc atcatttaaa gagctgatta gtacactgca 2400

gcctggcatt tctgtcatta caaggcctac tttacgctcc aagatagctg aagctgctct 2460

gatcatgaaa cagaaagtga ctgctgccat gagtgaagtt gaatggattg caaccacaac 2520

ggattgttgg actgcacgta gaaagtcatt cattggtgta actgctcact ggatcaaccc 2580

tggaagtctt gaaagacatt ccgctgcact tgcctgcaaa agattaatgg gctctcatac 2640

ttttgaggta ctggccagtg ccatgaatga tatccactca gagtatgaaa tacgtgacaa 2700

ggttgtttgc acaaccacag acagtggttc caactttatg aaggctttca gagtttttgg 2760

tgtggaaaac aatgatatcg agactgaggc aagaaggtgt gaaagtgatg acactgattc 2820

tgaaggctgt ggtgagggaa gtgatggtgt ggaattccaa gatgcctcac gagtcctgga 2880

ccaagacgat ggcttcgaat tccagctacc aaaacatcaa aagtgtgcct gtcacttact 2940

taacctagtc tcaagcgttg atgcccaaaa agctctctca aatgaacact acaagaaact 3000

ctacagatct gtctttggca aatgccaagc tttatggaat aaaagcagcc gatcggctct 3060

agcagctgaa gctgttgaat cagaaagccg gcttcagctt ttaaggccaa accaaacgcg 3120

gtggaattca acttttatgg ctgttgacag aattcttcaa atttgcaaag aagcaggaga 3180

aggcgcactt cggaatatat gcacctctct tgaggttcca atgtaagtgt ttttcccctc 3240

tatcgatgta aacaaatgtg ggttgttttt gtttaatact ctttgattat gctgatttct 3300

cctgtaggtt taatccagca gaaatgctgt tcttgacaga gtgggccaac acaatgcgtc 3360

cagttgcaaa agtactcgac atcttgcaag cggaaacgaa tacacagctg gggtggctgc 3420

tgcctagtgt ccatcagtta agcttgaaac ttcagcgact ccaccattct ctcaggtact 3480

gtgacccact tgtggatgcc ctacaacaag gaatccaaac acgattcaag catatgtttg 3540

aagatcctga gatcatagca gctgccatcc ttctccctaa atttcggacc tcttggacaa 3600

atgatgaaac catcataaaa cgaggtaaat gaatgcaagc aacatacact tgacgaattc 3660

taatctgggc aacctttgag ccataccaaa attattcttt tatttattta tttttgcact 3720

ttttaggaat gttatatccc atctttggct gtgatctcaa tatgaatatt gatgtaaagt 3780

attcttgcag caggttgtag ttatccctca gtgtttcttg aaaccaaact catatgtatc 3840

atatgtggtt tggaaatgca gttagatttt atgctaaaat aagggatttg catgatttta 3900

gatgtagatg actgcacgta aatgtagtta atgacaaaat ccataaaatt tgttcccagt 3960

cagaagcccc tcaaccaaac ttttctttgt gtctgctcac tgtgcttgta ggcatggact 4020

acatcagagt gcatctggag cctttggacc acaagaagga attggccaac agttcatctg 4080

atgatgaaga ttttttcgct tctttgaaac cgacaacaca tgaagccagc aaagagttgg 4140

atggatatct ggcctgtgtt tcagacacca gggagtctct gctcacgttt cctgctattt 4200

gcagcctctc tatcaagact aatacacctc ttcccgcatc ggctgcctgt gagaggcttt 4260

tcagcactgc aggattgctt ttcagcccca aaagagctag gcttgacact aacaattttg 4320

agaatcagct tctactgaag ttaaatctga ggttttacaa ctttgagtag cgtgtactgg 4380

cattagattg tctgtcttat agtttgataa ttaaatacaa acagttctaa agcaggataa 4440

aaccttgtat gcatttcatt taatgttttt tgagattaaa agcttaaaca agaatctcta 4500

gttttctttc ttgcttttac ttttacttcc ttaatactca agtacaattt taatggagta 4560

ctttttttact tttactcaag taagattcta gccagatact tttactttta attgagtaaa 4620

attttcccta agtacttgta ctttcacttg agtaaaattt ttgagtactt tttacacctc 4680

tg 4682

<210> 32
<211> 1053
<212> PRT
<213> Artificial Sequence

<220>
<223> dSaCas9

<400> 32

Met Lys Arg Asn Tyr Ile Leu Gly Leu Ala Ile Gly Ile Thr Ser Val
1 5 10 15

Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly
20 25 30

Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg
35 40 45

Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile
50 55 60

Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His
65 70 75 80

Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu
85 90 95

Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu
100 105 110

Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr
115 120 125

Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala
130 135 140

Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys
145 150 155 160

```
Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr
                165                 170                 175

Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln
            180                 185                 190

Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg
        195                 200                 205

Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys
    210                 215                 220

Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe
225                 230                 235                 240

Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr
                245                 250                 255

Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn
            260                 265                 270

Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe
        275                 280                 285

Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu
    290                 295                 300

Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys
305                 310                 315                 320

Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr
                325                 330                 335

Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala
            340                 345                 350

Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu
        355                 360                 365

Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser
    370                 375                 380
```

```
Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile
385                 390                 395                 400

Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala
                405                 410                 415

Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln
            420                 425                 430

Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro
        435                 440                 445

Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile
    450                 455                 460

Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg
465                 470                 475                 480

Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys
                485                 490                 495

Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr
            500                 505                 510

Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp
        515                 520                 525

Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu
    530                 535                 540

Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro
545                 550                 555                 560

Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys
                565                 570                 575

Gln Glu Glu Ala Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu
            580                 585                 590
```

```
Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile
        595                 600                 605

Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu
    610                 615                 620

Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp
625                 630                 635                 640

Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu
                645                 650                 655

Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys
            660                 665                 670

Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp
        675                 680                 685

Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp
    690                 695                 700

Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys
705                 710                 715                 720

Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys
                725                 730                 735

Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu
            740                 745                 750

Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp
        755                 760                 765

Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile
    770                 775                 780

Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu
785                 790                 795                 800

Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu
                805                 810                 815
```

```
Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His
            820                 825                 830

Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly
        835                 840                 845

Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr
    850                 855                 860

Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile
865                 870                 875                 880

Lys Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp
                885                 890                 895

Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr
            900                 905                 910

Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val
        915                 920                 925

Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser
    930                 935                 940

Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala
945                 950                 955                 960

Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly
                965                 970                 975

Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile
            980                 985                 990

Glu Val Asn Met Ile Asp Ile Thr  Tyr Arg Glu Tyr Leu  Glu Asn Met
        995                 1000                1005

Asn Asp  Lys Arg Pro Pro Arg  Ile Ile Lys Thr Ile  Ala Ser Lys
    1010                1015                1020
```

Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu
    1025   1030   1035

Tyr Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys Gly
    1040   1045   1050

<210> 33
<211> 1368
<212> PRT
<213> Artificial Sequence

<220>
<223> dCas9 with X at position 1, X can be any amino acid

<220>
<221> misc_feature
<222> (1)..(1)
<223> Xaa can be any naturally occurring amino acid

<400> 33

Xaa Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val
1   5   10   15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
            20   25   30

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
        35   40   45

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
    50   55   60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
65   70   75   80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
                85   90   95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
            100   105   110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
        115   120   125

```
His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
    130                 135                 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
145                 150                 155                 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
                165                 170                 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
            180                 185                 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
        195                 200                 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
    210                 215                 220

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
225                 230                 235                 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
                245                 250                 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
            260                 265                 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
        275                 280                 285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
    290                 295                 300

Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305                 310                 315                 320

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
                325                 330                 335
```

```
Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
            340                 345                 350

Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
        355                 360                 365

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
    370                 375                 380

Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
385                 390                 395                 400

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
                405                 410                 415

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
            420                 425                 430

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
        435                 440                 445

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
    450                 455                 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
465                 470                 475                 480

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
                485                 490                 495

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
            500                 505                 510

Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
        515                 520                 525

Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
    530                 535                 540

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
545                 550                 555                 560
```

```
Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
                565                 570                 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
            580                 585                 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
        595                 600                 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
    610                 615                 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625                 630                 635                 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
                645                 650                 655

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
            660                 665                 670

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
        675                 680                 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
    690                 695                 700

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
705                 710                 715                 720

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
                725                 730                 735

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
            740                 745                 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
        755                 760                 765
```

```
Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
    770                 775                 780

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785                 790                 795                 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
                805                 810                 815

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
            820                 825                 830

Leu Ser Asp Tyr Asp Val Asp Ala Ile Val Pro Gln Ser Phe Leu Lys
        835                 840                 845

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
    850                 855                 860

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
865                 870                 875                 880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
                885                 890                 895

Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
            900                 905                 910

Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
        915                 920                 925

Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
    930                 935                 940

Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
945                 950                 955                 960

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
                965                 970                 975

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
            980                 985                 990
```

```
Val Gly Thr Ala Leu Ile Lys Lys  Tyr Pro Lys Leu Glu  Ser Glu Phe
        995                 1000                1005


Val Tyr  Gly Asp Tyr Lys Val  Tyr Asp Val Arg Lys  Met Ile Ala
    1010                1015                1020


Lys Ser  Glu Gln Glu Ile Gly  Lys Ala Thr Ala Lys  Tyr Phe Phe
    1025                1030                1035


Tyr Ser  Asn Ile Met Asn Phe  Phe Lys Thr Glu Ile  Thr Leu Ala
    1040                1045                1050


Asn Gly  Glu Ile Arg Lys Arg  Pro Leu Ile Glu Thr  Asn Gly Glu
    1055                1060                1065


Thr Gly  Glu Ile Val Trp Asp  Lys Gly Arg Asp Phe  Ala Thr Val
    1070                1075                1080


Arg Lys  Val Leu Ser Met Pro  Gln Val Asn Ile Val  Lys Lys Thr
    1085                1090                1095


Glu Val  Gln Thr Gly Gly Phe  Ser Lys Glu Ser Ile  Leu Pro Lys
    1100                1105                1110


Arg Asn  Ser Asp Lys Leu Ile  Ala Arg Lys Lys Asp  Trp Asp Pro
    1115                1120                1125


Lys Lys  Tyr Gly Gly Phe Asp  Ser Pro Thr Val Ala  Tyr Ser Val
    1130                1135                1140


Leu Val  Val Ala Lys Val Glu  Lys Gly Lys Ser Lys  Lys Leu Lys
    1145                1150                1155


Ser Val  Lys Glu Leu Leu Gly  Ile Thr Ile Met Glu  Arg Ser Ser
    1160                1165                1170


Phe Glu  Lys Asn Pro Ile Asp  Phe Leu Glu Ala Lys  Gly Tyr Lys
    1175                1180                1185
```

```
Glu Val  Lys Lys Asp Leu Ile  Ile Lys Leu Pro Lys  Tyr Ser Leu
    1190                1195                1200

Phe Glu  Leu Glu Asn Gly Arg  Lys Arg Met Leu Ala  Ser Ala Gly
    1205                1210                1215

Glu Leu  Gln Lys Gly Asn Glu  Leu Ala Leu Pro Ser  Lys Tyr Val
    1220                1225                1230

Asn Phe  Leu Tyr Leu Ala Ser  His Tyr Glu Lys Leu  Lys Gly Ser
    1235                1240                1245

Pro Glu  Asp Asn Glu Gln Lys  Gln Leu Phe Val Glu  Gln His Lys
    1250                1255                1260

His Tyr  Leu Asp Glu Ile Ile  Glu Gln Ile Ser Glu  Phe Ser Lys
    1265                1270                1275

Arg Val  Ile Leu Ala Asp Ala  Asn Leu Asp Lys Val  Leu Ser Ala
    1280                1285                1290

Tyr Asn  Lys His Arg Asp Lys  Pro Ile Arg Glu Gln  Ala Glu Asn
    1295                1300                1305

Ile Ile  His Leu Phe Thr Leu  Thr Asn Leu Gly Ala  Pro Ala Ala
    1310                1315                1320

Phe Lys  Tyr Phe Asp Thr Thr  Ile Asp Arg Lys Arg  Tyr Thr Ser
    1325                1330                1335

Thr Lys  Glu Val Leu Asp Ala  Thr Leu Ile His Gln  Ser Ile Thr
    1340                1345                1350

Gly Leu  Tyr Glu Thr Arg Ile  Asp Leu Ser Gln Leu  Gly Gly Asp
    1355                1360                1365
```

<210> 34
<211> 199
<212> PRT
<213> Clostridium sp. 7_2_43FAA

<400> 34

```
Glu Gly Ile Lys Ser Asn Ile Ser Leu Leu Lys Asp Glu Leu Arg Gly
1               5                   10                  15

Gln Ile Ser His Ile Ser His Glu Tyr Leu Ser Leu Ile Asp Leu Ala
            20                  25                  30

Phe Asp Ser Lys Gln Asn Arg Leu Phe Glu Met Lys Val Leu Glu Leu
        35                  40                  45

Leu Val Asn Glu Tyr Gly Phe Lys Gly Arg His Leu Gly Gly Ser Arg
    50                  55                  60

Lys Pro Asp Gly Ile Val Tyr Ser Thr Thr Leu Glu Asp Asn Phe Gly
65                  70                  75                  80

Ile Ile Val Asp Thr Lys Ala Tyr Ser Glu Gly Tyr Ser Leu Pro Ile
                85                  90                  95

Ser Gln Ala Asp Glu Met Glu Arg Tyr Val Arg Glu Asn Ser Asn Arg
            100                 105                 110

Asp Glu Glu Val Asn Pro Asn Lys Trp Trp Glu Asn Phe Ser Glu Glu
        115                 120                 125

Val Lys Lys Tyr Tyr Phe Val Phe Ile Ser Gly Ser Phe Lys Gly Lys
    130                 135                 140

Phe Glu Glu Gln Leu Arg Arg Leu Ser Met Thr Thr Gly Val Asn Gly
145                 150                 155                 160

Ser Ala Val Asn Val Val Asn Leu Leu Leu Gly Ala Glu Lys Ile Arg
                165                 170                 175

Ser Gly Glu Met Thr Ile Glu Glu Leu Glu Arg Ala Met Phe Asn Asn
            180                 185                 190

Ser Glu Phe Ile Leu Lys Tyr
        195
```

<210> 35

<211> 4897
<212> DNA
<213> Artificial Sequence

<220>
<223> PB-EF1a vector

<400> 35
tgtacataga ttaaccctag aaagataatc atattgtgac gtacgttaaa gataatcatg 60

cgtaaaattg acgcatgtgt tttatcggtc tgtatatcga ggtttattta ttaatttgaa 120

tagatattaa gttttattat atttacactt acatactaat aataaattca acaaacaatt 180

tatttatgtt tatttattta ttaaaaaaaa acaaaaactc aaaatttctt ctataaagta 240

acaaaacttt tatcgaatac ctgcagcccg ggggatgcag agggacagcc ccccccaaa 300

gcccccaggg atgtaattac gtccctcccc cgctaggggg cagcagcgag ccgcccgggg 360

ctccgctccg gtccggcgct ccccccgcat ccccgagccg gcagcgtgcg gggacagccc 420

gggcacgggg aaggtggcac gggatcgctt tcctctgaac gcttctcgct gctctttgag 480

cctgcagaca cctgggggga tacggggaaa agttgactgt gcctttcgat cgaaccatgg 540

acagttagct ttgcaaagat ggataaagtt ttaaacagag aggaatcttt gcagctaatg 600

gaccttctag gtcttgaaag gagtgggaat tggctccggt gcccgtcagt gggcagagcg 660

cacatcgccc acagtccccg agaagttggg gggaggggtc ggcaattgaa ccggtgccta 720

gagaaggtgg cgcggggtaa actgggaaag tgatgtcgtg tactggctcc gccttttcc 780

cgagggtggg ggagaaccgt atataagtgc agtagtcgcc gtgaacgttc tttttcgcaa 840

cgggtttgcc gccagaacac aggtaagtgc cgtgtgtggt tcccgcgggc ctggcctctt 900

tacgggttat ggcccttgcg tgccttgaat tacttccacc tggctgcagt acgtgattct 960

tgatcccgag cttcgggttg gaagtgggtg ggagagttcg aggccttgcg cttaaggagc 1020

cccttcgcct cgtgcttgag ttgaggcctg gcctgggcgc tggggccgcc gcgtgcgaat 1080

ctggtggcac cttcgcgcct gtctcgctgc tttcgataag tctctagcca tttaaaattt 1140

ttgatgacct gctgcgacgc tttttttctg gcaagatagt cttgtaaatg cgggccaaga 1200

tctgcacact ggtatttcgg tttttggggc cgcgggcggc gacggggccc gtgcgtccca 1260

gcgcacatgt tcggcgaggc ggggcctgcg agcgcggcca ccgagaatcg gacgggggta 1320

gtctcaagct ggccggcctg ctctggtgcc tggcctcgcg ccgccgtgta tcgccccgcc 1380

ctgggcggca aggctggccc ggtcggcacc agttgcgtga gcggaaagat ggccgcttcc 1440

cggccctgct gcagggagct caaaatggag gacgcggcgc tcgggagagc gggcgggtga 1500

gtcacccaca caaaggaaaa gggcctttcc gtcctcagcc gtcgcttcat gtgactccac 1560

ggagtaccgg gcgccgtcca ggcacctcga ttagttctcg agcttttgga gtacgtcgtc 1620

tttaggttgg ggggaggggt tttatgcgat ggagtttccc cacactgagt gggtggagac 1680

tgaagttagg ccagcttggc acttgatgta attctccttg gaatttgccc tttttgagtt 1740

tggatcttgg ttcattctca agcctcagac agtggttcaa agtttttttc ttccatttca 1800

ggtgtcgtga gaattctaat acgactcact atagggtgtg ctgtctcatc attttggcaa 1860

agattggcca ccaagcttgt cctgcaggag ggtcgacgcc tctagacggg cggccgctcc 1920

ggatccacgg gtaccgatca catatgcctt taattaaaca ctagttctat agtgtcacct 1980

aaattccctt tagtgagggt taatggccgt aggccgccag aattgggtcc agacatgata 2040

agatacattg atgagtttgg acaaaccaca actagaatgc agtgaaaaaa atgctttatt 2100

tgtgaaattt gtgatgctat tgctttattt gtaaccatta taagctgcaa taaacaagtt 2160

aacaacaaca attgcattca ttttatgttt caggttcagg gggaggtgtg ggaggttttt 2220

tcggactcta ggacctgcgc atgcgcttgg cgtaatcatg gtcatagctg tttcctgttt 2280

tccccgtatc cccccaggtg tctgcaggct caaagagcag cgagaagcgt tcagaggaaa 2340

gcgatcccgt gccaccttcc ccgtgcccgg gctgtccccg cacgctgccg gctcggggat 2400

gcgggggggag cgccggaccg gagcggagcc ccgggcggct cgctgctgcc ccctagcggg 2460

ggagggacgt aattacatcc ctgggggctt tgggggggggg ctgtccctct caccgcggtg 2520

gagctccagc ttttgttcga attggggccc cccctcgagg gtatcgatga tatctataac 2580

aagaaaatat atatataata agttatcacg taagtagaac atgaaataac aatataatta 2640

tcgtatgagt taaatcttaa aagtcacgta aaagataatc atgcgtcatt ttgactcacg 2700

cggtcgttat agttcaaaat cagtgacact taccgcattg acaagcacgc ctcacgggag 2760

ctccaagcgg cgactgagat gtcctaaatg cacagcgacg gattcgcgct atttagaaag 2820

agagagcaat atttcaagaa tgcatgcgtc aattttacgc agactatctt tctagggtta 2880

atctagctag ccttaagggc gcctattgcg ttgcgctcac tgcccgcttt ccagtcggga 2940

aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt 3000

attgggcgct cttccgcttc ctcgctcact gactcgctgc gctcggtcgt tcggctgcgg 3060

cgagcggtat cagctcactc aaaggcggta atacggttat ccacagaatc aggggataac 3120

gcaggaaaga acatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac 3180

cccgtagaaa agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc 3240

ttgcaaacaa aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca 3300

actcttttc cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgttcttcta 3360

gtgtagccgt agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct 3420

ctgctaatcc tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg 3480

gactcaagac gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc 3540

acacagccca gcttggagcg aacgacctac accgaactga gatacctaca gcgtgagcta 3600

tgagaaagcg ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg 3660

gtcggaacag gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt 3720

cctgtcgggt ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc gtcaggggg 3780

cggagcctat ggaaaaacgc cagcaacgcg gccttttac ggttcctggc cttttgctgg 3840

ccttttgctc acatgagatt atcaaaaagg atcttcacct agatcctttt aaattaaaaa 3900

tgaagtttta aatcaatcta aagtatatat gagtaaactt ggtctgacag tcagaagaac 3960

tcgtcaagaa ggcgatagaa ggcgatgcgc tgcgaatcgg gagcggcgat accgtaaagc 4020

acgaggaagc ggtcagccca ttcgccgcca agctcttcag caatatcacg ggtagccaac 4080

gctatgtcct gatagcggtc cgccacaccc agccggccac agtcgatgaa tccagaaaag 4140

cggccatttt ccaccatgat attcggcaag caggcatcgc catgggtcac gacgagatcc 4200

tcgccgtcgg gcatgctcgc cttgagcctg gcgaacagtt cggctggcgc gagcccctga 4260

tgctcttcgt ccagatcatc ctgatcgaca agaccggctt ccatccgagt acgtgctcgc 4320

tcgatgcgat gtttcgcttg gtggtcgaat gggcaggtag ccggatcaag cgtatgcagc 4380

cgccgcattg catcagccat gatggatact ttctcggcag gagcaaggtg agatgacagg 4440

agatcctgcc ccggcacttc gcccaatagc agccagtccc ttcccgcttc agtgacaacg 4500

tcgagcacag ctgcgcaagg aacgcccgtc gtggccagcc acgatagccg cgctgcctcg 4560

tcttgcagtt cattcagggc accggacagg tcggtcttga caaaaagaac cgggcgcccc 4620

tgcgctgaca gccggaacac ggcggcatca gagcagccga ttgtctgttg tgcccagtca 4680

tagccgaata gcctctccac ccaagcggcc ggagaacctg cgtgcaatcc atcttgttca 4740

atcataatat tattgaagca tttatcaggg ttcgtctcgt cccggtctcc tcccaatgca 4800

tgtcaatatt ggccattagc catattattc attggttata tagcataaat caatattggc 4860

tattggccat tgcatacgtt gtatctatat cataata 4897

<210> 36
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> EF1a-2r primer

<400> 36
caccggagcc aattcccact 20

<210> 37
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> AAVS-3r primer

<400> 37
ctgcaccacg tgatgtcctc 20

<210> 38
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> SV40pA-1r primer

<400> 38
gtaaccatta taagctgcaa taaacaag 28

<210> 39
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> AAVS-2f primer

<400> 39
ctggggactc tttaaggaaa gaag 24

<210> 40
<211> 1591
<212> PRT
<213> Artificial Sequence

<220>
<223> dCas9-Clo051

<400> 40

Met Ala Pro Lys Lys Lys Arg Lys Val Glu Gly Ile Lys Ser Asn Ile
1 5 10 15

Ser Leu Leu Lys Asp Glu Leu Arg Gly Gln Ile Ser His Ile Ser His
20 25 30

Glu Tyr Leu Ser Leu Ile Asp Leu Ala Phe Asp Ser Lys Gln Asn Arg
35 40 45

Leu Phe Glu Met Lys Val Leu Glu Leu Leu Val Asn Glu Tyr Gly Phe
50 55 60

Lys Gly Arg His Leu Gly Gly Ser Arg Lys Pro Asp Gly Ile Val Tyr
65 70 75 80

Ser Thr Thr Leu Glu Asp Asn Phe Gly Ile Ile Val Asp Thr Lys Ala
85 90 95

Tyr Ser Glu Gly Tyr Ser Leu Pro Ile Ser Gln Ala Asp Glu Met Glu
100 105 110

Arg Tyr Val Arg Glu Asn Ser Asn Arg Asp Glu Glu Val Asn Pro Asn
115 120 125

Lys Trp Trp Glu Asn Phe Ser Glu Glu Val Lys Lys Tyr Tyr Phe Val
130 135 140

Phe Ile Ser Gly Ser Phe Lys Gly Lys Phe Glu Glu Gln Leu Arg Arg
145 150 155 160

```
Leu Ser Met Thr Thr Gly Val Asn Gly Ser Ala Val Asn Val Val Asn
                165                 170                 175

Leu Leu Leu Gly Ala Glu Lys Ile Arg Ser Gly Glu Met Thr Ile Glu
            180                 185                 190

Glu Leu Glu Arg Ala Met Phe Asn Asn Ser Glu Phe Ile Leu Lys Tyr
        195                 200                 205

Gly Gly Gly Gly Ser Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly
    210                 215                 220

Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro
225                 230                 235                 240

Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys
                245                 250                 255

Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu
            260                 265                 270

Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys
        275                 280                 285

Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys
    290                 295                 300

Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu
305                 310                 315                 320

Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp
                325                 330                 335

Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys
            340                 345                 350

Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu
        355                 360                 365

Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly
```

```
    370                 375                 380

Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu
385                 390                 395                 400

Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser
                405                 410                 415

Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg
            420                 425                 430

Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly
        435                 440                 445

Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe
    450                 455                 460

Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys
465                 470                 475                 480

Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp
                485                 490                 495

Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile
            500                 505                 510

Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro
        515                 520                 525

Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu
    530                 535                 540

Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys
545                 550                 555                 560

Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp
                565                 570                 575

Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu
            580                 585                 590
```

```
Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu
        595                 600                 605

Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His
    610                 615                 620

Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp
625                 630                 635                 640

Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu
                645                 650                 655

Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser
            660                 665                 670

Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp
        675                 680                 685

Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile
    690                 695                 700

Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu
705                 710                 715                 720

Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu
                725                 730                 735

Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu
            740                 745                 750

Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn
        755                 760                 765

Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile
    770                 775                 780

Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn
785                 790                 795                 800

Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys
```

```
                805                 810                 815

Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val
            820                 825                 830

Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu
        835                 840                 845

Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys
    850                 855                 860

Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn
865                 870                 875                 880

Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys
                885                 890                 895

Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp
            900                 905                 910

Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln
        915                 920                 925

Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala
    930                 935                 940

Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val
945                 950                 955                 960

Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala
                965                 970                 975

Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg
            980                 985                 990

Met Lys Arg Ile Glu Glu Gly Ile  Lys Glu Leu Gly Ser  Gln Ile Leu
        995                 1000                1005

Lys Glu  His Pro Val Glu Asn  Thr Gln Leu Gln Asn  Glu Lys Leu
    1010                1015                1020
```

```
Tyr Leu  Tyr Tyr Leu Gln Asn  Gly Arg Asp Met Tyr  Val Asp Gln
    1025                1030                1035

Glu Leu  Asp Ile Asn Arg Leu  Ser Asp Tyr Asp Val  Asp Ala Ile
    1040                1045                1050

Val Pro  Gln Ser Phe Leu Lys  Asp Asp Ser Ile Asp  Asn Lys Val
    1055                1060                1065

Leu Thr  Arg Ser Asp Lys Asn  Arg Gly Lys Ser Asp  Asn Val Pro
    1070                1075                1080

Ser Glu  Glu Val Val Lys Lys  Met Lys Asn Tyr Trp  Arg Gln Leu
    1085                1090                1095

Leu Asn  Ala Lys Leu Ile Thr  Gln Arg Lys Phe Asp  Asn Leu Thr
    1100                1105                1110

Lys Ala  Glu Arg Gly Gly Leu  Ser Glu Leu Asp Lys  Ala Gly Phe
    1115                1120                1125

Ile Lys  Arg Gln Leu Val Glu  Thr Arg Gln Ile Thr  Lys His Val
    1130                1135                1140

Ala Gln  Ile Leu Asp Ser Arg  Met Asn Thr Lys Tyr  Asp Glu Asn
    1145                1150                1155

Asp Lys  Leu Ile Arg Glu Val  Lys Val Ile Thr Leu  Lys Ser Lys
    1160                1165                1170

Leu Val  Ser Asp Phe Arg Lys  Asp Phe Gln Phe Tyr  Lys Val Arg
    1175                1180                1185

Glu Ile  Asn Asn Tyr His His  Ala His Asp Ala Tyr  Leu Asn Ala
    1190                1195                1200

Val Val  Gly Thr Ala Leu Ile  Lys Lys Tyr Pro Lys  Leu Glu Ser
    1205                1210                1215

Glu Phe  Val Tyr Gly Asp Tyr  Lys Val Tyr Asp Val  Arg Lys Met
```

```
    1220                1225                1230

Ile Ala  Lys Ser Glu Gln Glu  Ile Gly Lys Ala Thr  Ala Lys Tyr
    1235                1240                1245

Phe Phe  Tyr Ser Asn Ile Met  Asn Phe Phe Lys Thr  Glu Ile Thr
    1250                1255                1260

Leu Ala  Asn Gly Glu Ile Arg  Lys Arg Pro Leu Ile  Glu Thr Asn
    1265                1270                1275

Gly Glu  Thr Gly Glu Ile Val  Trp Asp Lys Gly Arg  Asp Phe Ala
    1280                1285                1290

Thr Val  Arg Lys Val Leu Ser  Met Pro Gln Val Asn  Ile Val Lys
    1295                1300                1305

Lys Thr  Glu Val Gln Thr Gly  Gly Phe Ser Lys Glu  Ser Ile Leu
    1310                1315                1320

Pro Lys  Arg Asn Ser Asp Lys  Leu Ile Ala Arg Lys  Lys Asp Trp
    1325                1330                1335

Asp Pro  Lys Lys Tyr Gly Gly  Phe Asp Ser Pro Thr  Val Ala Tyr
    1340                1345                1350

Ser Val  Leu Val Val Ala Lys  Val Glu Lys Gly Lys  Ser Lys Lys
    1355                1360                1365

Leu Lys  Ser Val Lys Glu Leu  Leu Gly Ile Thr Ile  Met Glu Arg
    1370                1375                1380

Ser Ser  Phe Glu Lys Asn Pro  Ile Asp Phe Leu Glu  Ala Lys Gly
    1385                1390                1395

Tyr Lys  Glu Val Lys Lys Asp  Leu Ile Ile Lys Leu  Pro Lys Tyr
    1400                1405                1410

Ser Leu  Phe Glu Leu Glu Asn  Gly Arg Lys Arg Met  Leu Ala Ser
    1415                1420                1425
```

Ala Gly Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys
    1430   1435   1440

Tyr Val Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys
    1445   1450   1455

Gly Ser Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln
    1460   1465   1470

His Lys His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe
    1475   1480   1485

Ser Lys Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu
    1490   1495   1500

Ser Ala Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala
    1505   1510   1515

Glu Asn Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro
    1520   1525   1530

Ala Ala Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr
    1535   1540   1545

Thr Ser Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser
    1550   1555   1560

Ile Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly
    1565   1570   1575

Gly Asp Gly Ser Pro Lys Lys Lys Arg Lys Val Ser Ser
    1580   1585   1590

<210> 41
<211> 4073
<212> DNA
<213> Artificial Sequence

<220>
<223> pHR-GFP- selection gene plasmid

<400> 41

ggggccacta gggacaggat cggcgtcttc actcgctggg ttccctttc cttctccttc 60

tggggcctgt gccatctctc gtttcttagg atggccttct ccgacggatg tctcccttgc 120

gtcccgcctc cccttcttgt aggcctgcat catcaccgtt tttctggaca accccaaagt 180

accccgtctc cctggcttta gccacctctc catcctcttg ctttctttgc ctggacaccc 240

cgttctcctg tggattcggg tcacctctca ctcctttcat ttgggcagct cccctacccc 300

ccttacctct ctagtctgtg ctagctcttc cagcccccctg tcatggcatc ttccaggggt 360

ccgagagctc agctagtctt cttcctccaa cccgggcccc tatgtccact tcaggacagc 420

atgtttgctg cctccaggga tcctgtgtcc ccgagctggg accaccttat attcccaggg 480

ccggttaatg tggctctggt tctgggtact tttatctgtc ccgacgtccg atcgaaccat 540

ggacagttag ctttgcaaag atggataaag ttttaaacag agaggaatct ttgcagctaa 600

tggaccttct aggtcttgaa aggagtggga attggctccg gtgcccgtca gtgggcagag 660

cgcacatcgc ccacagtccc cgagaagttg gggggagggg tcggcaattg aaccggtgcc 720

tagagaaggt ggcgcggggt aaactgggaa agtgatgtcg tgtactggct ccgccttttt 780

cccgagggtg ggggagaacc gtatataagt gcagtagtcg ccgtgaacgt tctttttcgc 840

aacgggtttg ccgccagaac acaggtaagt gccgtgtgtg gttcccgcgg gcctggcctc 900

tttacgggtt atggcccttg cgtgccttga attacttcca cctggctgca gtacgtgatt 960

cttgatcccg agcttcgggt tggaagtggg tgggagagtt cgaggccttg cgcttaagga 1020

gcccccttcgc ctcgtgcttg agttgaggcc tggcctgggc gctggggccg ccgcgtgcga 1080

atctggtggc accttcgcgc ctgtctcgct gctttcgata agtctctagc catttaaaat 1140

ttttgatgac ctgctgcgac gctttttttc tggcaagata gtcttgtaaa tgcgggccaa 1200

gatctgcaca ctggtatttc ggtttttggg gccgcgggcg gcgacggggc ccgtgcgtcc 1260

cagcgcacat gttcggcgag gcggggcctg cgagcgcggc caccgagaat cggacggggg 1320

tagtctcaag ctggccggcc tgctctggtg cctggcctcg cgccgccgtg tatcgccccg 1380

ccctgggcgg caaggctggc ccggtcggca ccagttgcgt gagcggaaag atggccgctt 1440

cccggccctg ctgcagggag ctcaaaatgg aggacgcggc gctcgggaga gcgggcgggt 1500

gagtcaccca cacaaaggaa aagggccttt ccgtcctcag ccgtcgcttc atgtgactcc 1560

acggagtacc gggcgccgtc caggcacctc gattagttct cgagcttttg gagtacgtcg 1620

tctttaggtt ggggggaggg gttttatgcg atggagtttc cccacactga gtgggtggag 1680

actgaagtta ggccagcttg gcacttgatg taattctcct tggaatttgc ccttttgag 1740

tttggatctt ggttcattct caagcctcag acagtggttc aaagtttttt tcttccattt 1800

caggtgtcgt gagaattcta atacgactca ctatagggtg tgctgtctca tcattttggc 1860

aaagattggc caccaagctt accgccatgg tgagcaaggg cgaggagctg ttcaccgggg 1920

tggtgcccat cctggtcgag ctggacggcg acgtaaacgg ccacaagttc agcgtgtccg 1980

gcgagggcga gggcgatgcc acctacggca agctgaccct gaagttcatc tgcaccaccg 2040

gcaagctgcc cgtgccctgg cccaccctcg tgaccaccct gacctacggc gtgcagtgct 2100

tcagccgcta ccccgaccac atgaagcagc acgacttctt caagtccgcc atgcccgaag 2160

gctacgtcca ggagcgcacc atcttcttca aggacgacgg caactacaag acccgcgccg 2220

aggtgaagtt cgagggcgac accctggtga accgcatcga gctgaagggc atcgacttca 2280

aggaggacgg caacatcctg gggcacaagc tggagtacaa ctacaacagc cacaacgtct 2340

atatcatggc cgacaagcag aagaacggca tcaaggtgaa cttcaagatc cgccacaaca 2400

tcgaggacgg cagcgtgcag ctcgccgacc actaccagca gaacaccccc atcggcgacg 2460

gccccgtgct gctgcccgac aaccactacc tgagcaccca gtccgccctg agcaaagacc 2520

ccaacgagaa gcgtgatcac atggtcctgc tggagttcgt gaccgccgcc gggatcactc 2580

tcggcatgga cgagctgtac aaggaaggaa gaggcagcct gctgacatgt ggcgacgtgg 2640

aggagaaccc tggcccaatg gtgggcagcc tgaattgtat cgtggccgtg tcccagaaca 2700

tgggcatcgg caagaatggc gattttcctt ggccccctct gagaaatgag tccagatact 2760

ttcagaggat gaccacaacc agctccgtgg agggcaagca gaacctggtc atcatgggca 2820

agaagacatg gttctctatc ccagagaaga accgccccct gaagggccgg atcaatctgg 2880

tgctgagcag ggagctgaag gagccacccc agggagcaca ctttctgtcc aggtctctgg 2940

acgatgccct gaagctgacc gagcagcctg agctggccaa caaggtggac atggtgtgga 3000

tcgtgggcgg ctctagcgtg tataaggagg ccatgaatca ccctggccac ctgaagctgt 3060

tcgtgacacg gatcatgcag gactttgagt ccgatacctt ctttccagag atcgacctgg 3120

agaagtacaa gctgctgccc gagtatcctg gcgtgctgtc tgatgtgcag gaggagaagg 3180

gcatcaagta caagttcgag gtgtatgaga agaacgattg ataacatatg cctttaatta 3240

aacactagtt ctatagtgtc acctaaattc cctttagtga gggttaatgg ccgtaggccg 3300

ccagaattgg gtccagacat gataagatac attgatgagt ttggacaaac cacaactaga 3360

atgcagtgaa aaaaatgctt tatttgtgaa atttgtgatg ctattgcttt atttgtaacc 3420

attataagct gcaataaaca agttaacaac aacaattgca ttcattttat gtttcaggtt 3480

cagggggagg tgtgggaggt tttttcggac tctaggacct gcgcatgcgc ttggggtacc 3540

taggatatcg acagaaaagc cccatcctta ggcctcctcc ttcctagtct cctgatattg 3600

ggtctaaccc ccacctcctg ttaggcagat tccttatctg gtgacacacc cccatttcct 3660

ggagccatct ctctccttgc cagaacctct aaggtttgct tacgatggag ccagagagga 3720

tcctgggagg gagagcttgg cagggggtgg gagggaaggg ggggatgcgt gacctgcccg 3780

gttctcagtg gccaccctgc gctaccctct cccagaacct gagctgctct gacgcggccg 3840

tctggtgcgt ttcactgatc ctggtgctgc agcttcctta cacttcccaa gaggagaagc 3900

agtttggaaa aacaaaatca gaataagttg gtcctgagtt ctaactttgg ctcttcacct 3960

ttctagtccc caatttatat tgttcctccg tgcgtcagtt ttacctgtga gataaggcca 4020

gtagccagcc ccgtcctggc agggctgtgc ctctagggac aggattggtg acg 4073

<210> 42
<211> 3089
<212> DNA
<213> Artificial Sequence

<220>
<223> pMMEJ-GFP-selection gene plasmid

<400> 42
ccaatcctgt ccctagtggc cccactgtgg ggacgtccga tcgaaccatg gacagttagc 60

tttgcaaaga tggataaagt tttaaacaga gaggaatctt tgcagctaat ggaccttcta 120

ggtcttgaaa ggagtgggaa ttggctccgg tgcccgtcag tgggcagagc gcacatcgcc 180

cacagtcccc gagaagttgg ggggaggggt cggcaattga accggtgcct agagaaggtg 240

gcgcggggta aactgggaaa gtgatgtcgt gtactggctc cgccttttтc ccgagggtgg 300

gggagaaccg tatataagtg cagtagtcgc cgtgaacgtt ctttttcgca acgggtttgc 360

cgccagaaca caggtaagtg ccgtgtgtgg ttcccgcggg cctggcctct ttacgggtta 420

tggcccttgc gtgccttgaa ttacttccac ctggctgcag tacgtgattc ttgatcccga 480

gcttcgggtt ggaagtgggt gggagagttc gaggccttgc gcttaaggag ccccttcgcc 540

tcgtgcttga gttgaggcct ggcctgggcg ctggggccgc cgcgtgcgaa tctggtggca 600

ccttcgcgcc tgtctcgctg ctttcgataa gtctctagcc atttaaaatt tttgatgacc 660

tgctgcgacg ctttttttct ggcaagatag tcttgtaaat gcgggccaag atctgcacac 720

tggtatttcg gtttttgggg ccgcgggcgg cgacggggcc cgtgcgtccc agcgcacatg 780

ttcggcgagg cggggcctgc gagcgcggcc accgagaatc ggacgggggt agtctcaagc 840

tggccggcct gctctggtgc ctggcctcgc gccgccgtgt atcgccccgc cctgggcggc 900

aaggctggcc cggtcggcac cagttgcgtg agcggaaaga tggccgcttc ccggccctgc 960

tgcagggagc tcaaaatgga ggacgcggcg ctcgggagag cgggcgggtg agtcacccac 1020

acaaaggaaa agggcctttc cgtcctcagc cgtcgcttca tgtgactcca cggagtaccg 1080

ggcgccgtcc aggcacctcg attagttctc gagcttttgg agtacgtcgt ctttaggttg 1140

gggggagggg ttttatgcga tggagtttcc ccacactgag tgggtggaga ctgaagttag 1200

gccagcttgg cacttgatgt aattctcctt ggaatttgcc ctttttgagt ttggatcttg 1260

gttcattctc aagcctcaga cagtggttca aagtttttt cttccatttc aggtgtcgtg 1320

agaattctaa tacgactcac tatagggtgt gctgtctcat cattttggca aagattggcc 1380

accaagctta ccgccatggt gagcaagggc gaggagctgt tcaccggggt ggtgcccatc 1440

ctggtcgagc tggacggcga cgtaaacggc cacaagttca gcgtgtccgg cgagggcgag 1500

ggcgatgcca cctacggcaa gctgaccctg aagttcatct gcaccaccgg caagctgccc 1560

gtgccctggc ccaccctcgt gaccaccctg acctacggcg tgcagtgctt cagccgctac 1620

cccgaccaca tgaagcagca cgacttcttc aagtccgcca tgcccgaagg ctacgtccag 1680

gagcgcacca tcttcttcaa ggacgacggc aactacaaga cccgcgccga ggtgaagttc 1740

gagggcgaca ccctggtgaa ccgcatcgag ctgaagggca tcgacttcaa ggaggacggc 1800

aacatcctgg ggcacaagct ggagtacaac tacaacagcc acaacgtcta tatcatggcc 1860

gacaagcaga agaacggcat caaggtgaac ttcaagatcc gccacaacat cgaggacggc 1920

agcgtgcagc tcgccgacca ctaccagcag aacaccccca tcggcgacgg ccccgtgctg 1980

ctgcccgaca accactacct gagcacccag tccgccctga gcaaagaccc caacgagaag 2040

cgtgatcaca tggtcctgct ggagttcgtg accgccgccg ggatcactct cggcatggac 2100

gagctgtaca aggaaggaag aggcagcctg ctgacatgtg gcgacgtgga ggagaaccct 2160

ggcccaatgg tgggcagcct gaattgtatc gtggccgtgt cccagaacat gggcatcggc 2220

aagaatggcg attttccttg gccccctctg agaaatgagt ccagatactt tcagaggatg 2280

accacaacca gctccgtgga gggcaagcag aacctggtca tcatgggcaa gaagacatgg 2340

ttctctatcc cagagaagaa ccgccccctg aagggccgga tcaatctggt gctgagcagg 2400

gagctgaagg agccacccca gggagcacac tttctgtcca ggtctctgga cgatgccctg 2460

aagctgaccg agcagcctga gctggccaac aaggtggaca tggtgtggat cgtgggcggc 2520

tctagcgtgt ataaggaggc catgaatcac cctggccacc tgaagctgtt cgtgacacgg 2580

atcatgcagg actttgagtc cgataccttc tttccagaga tcgacctgga gaagtacaag 2640

ctgctgcccg agtatcctgg cgtgctgtct gatgtgcagg aggagaaggg catcaagtac 2700

aagttcgagg tgtatgagaa gaacgattga taacatatgc ctttaattaa acactagttc 2760

tatagtgtca cctaaattcc ctttagtgag ggttaatggc cgtaggccgc cagaattggg 2820

tccagacatg ataagataca ttgatgagtt tggacaaacc acaactagaa tgcagtgaaa 2880

aaaatgcttt atttgtgaaa tttgtgatgc tattgcttta tttgtaacca ttataagctg 2940

caataaacaa gttaacaaca acaattgcat tcattttatg tttcaggttc agggggaggt 3000

gtgggaggtt ttttcggact ctaggacctg cgcatgcgct tggggtacct aggatatcgg 3060

atggggcttt tctgtcacca atcctgagg 3089