SEQUENCE LISTING

<110> SIGMA-ALDRICH CO. LLC
CHEN, Fuqiang
DAVIS, Gregory D.
KANG, Qiaohua
KNIGHT, Scott W.

<120> CRISPR-BASED GENOME MODIFICATION AND REGULATION

<130> 047497-465606

<150> US 61/734,256
<151> 2012-12-06

<150> US 61/758,624
<151> 2013-01-30

<150> US 61/761,046
<151> 2013-02-05

<150> US 61/794,422
<151> 2013-03-15

<160> 27

<170> PatentIn version 3.5

<210> 1
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> SYNTHESIZED

<400> 1

Pro Lys Lys Lys Arg Lys Val
1 5

<210> 2
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> SYNTHESIZED

<400> 2

Pro Lys Lys Lys Arg Arg Val
1 5

<210> 3
<211> 16
<212> PRT
<213> Artificial Sequence

<220>
<223> SYNTHESIZED

<400> 3

Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys
1 5 10 15

<210> 4
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> SYNTHESIZED

<400> 4

Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Pro Pro Gln Pro Lys Lys
1 5 10 15

Lys Arg Lys Val
20

<210> 5
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> SYNTHESIZED

<400> 5

Pro Leu Ser Ser Ile Phe Ser Arg Ile Gly Asp Pro Pro Lys Lys Lys
1 5 10 15

Arg Lys Val

<210> 6
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> SYNTHESIZED

<400> 6

Gly Ala Leu Phe Leu Gly Trp Leu Gly Ala Ala Gly Ser Thr Met Gly
1 5 10 15

Ala Pro Lys Lys Lys Arg Lys Val
20

<210> 7
<211> 27
<212> PRT
<213> Artificial Sequence

<220>
<223> SYNTHESIZED

<400> 7

Gly Ala Leu Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly
1 5 10 15

Ala Trp Ser Gln Pro Lys Lys Lys Arg Lys Val
20 25

<210> 8
<211> 21
<212> PRT
<213> Artificial Sequence

<220>
<223> SYNTHESIZED

<400> 8

Lys Glu Thr Trp Trp Glu Thr Trp Trp Thr Glu Trp Ser Gln Pro Lys
1 5 10 15

Lys Lys Arg Lys Val
20

<210> 9
<211> 1374
<212> PRT
<213> Artificial Sequence

<220>
<223> SYNTHESIZED

<400> 9

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

Gly Trp Ala Val Ile Thr Asp Asp 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 Gly 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 Ala 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 Ile Tyr
            180                 185                 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Arg 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 Arg 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 Ser 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

Ala 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 Gly 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 His Ser Leu
705                 710                 715                 720

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Val Lys  Lys Asp Leu Ile Ile  Lys Leu Pro Lys Tyr  Ser Leu Phe
    1190                1195                1200

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

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

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

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

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

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

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

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

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

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

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

Lys Lys Lys Arg Lys Val
1370

<210> 10
<211> 4122
<212> DNA
<213> Artificial Sequence

<220>
<223> SYNTHESIZED

<400> 10
atggacaaga agtacagcat cggcctggac atcggcacca actctgtggg ctgggccgtg 60

atcaccgacg actacaaggt gcccagcaag aaattcaagg tgctgggcaa caccgaccgg 120

cacagcatca agaagaacct gatcggcgcc ctgctgttcg gctctggcga aacagccgag 180

gccacccggc tgaagagaac cgccagaaga agatacacca gacggaagaa ccggatctgc 240

tatctgcaag agatcttcag caacgagatg gccaaggtgg acgacagctt cttccacaga 300

ctggaagagt ccttcctggt ggaagaggat aagaagcacg agcggcaccc catcttcggc 360

aacatcgtgg acgaggtggc ctaccacgag aagtacccca ccatctacca cctgagaaag 420

aagctggccg acagcaccga caaggccgac ctgagactga tctacctggc cctggcccac 480

atgatcaagt tccggggcca cttcctgatc gagggcgacc tgaaccccga caacagcgac 540

gtggacaagc tgttcatcca gctggtgcag atctacaatc agctgttcga ggaaaacccc 600

atcaacgcca gcagagtgga cgccaaggcc atcctgagcg ccagactgag caagagcaga 660

cggctggaaa atctgatcgc ccagctgccc ggcgagaagc ggaatggcct gttcggcaac 720

ctgattgccc tgagcctggg cctgaccccc aacttcaaga gcaacttcga cctggccgag 780

gatgccaaac tgcagctgag caaggacacc tacgacgacg acctggacaa cctgctggcc 840

cagatcggcg accagtacgc cgacctgttt ctggccgcca agaacctgtc cgacgccatc 900

ctgctgagcg acatcctgag agtgaacagc gagatcacca aggcccccct gtccgcctct 960

atgatcaaga gatacgacga gcaccaccag gacctgaccc tgctgaaagc tctcgtgcgg 1020

cagcagctgc ctgagaagta caaagagatt ttcttcgacc agagcaagaa cggctacgcc 1080

ggctacatcg atggcggagc cagccaggaa gagttctaca agttcatcaa gcccatcctg 1140

gaaaagatgg acggcaccga ggaactgctc gtgaagctga acagagagga cctgctgcgg 1200

aagcagcgga ccttcgacaa cggcagcatc ccccaccaga tccacctggg agagctgcac 1260

gccattctgc ggcggcagga agatttttac ccattcctga aggacaaccg ggaaaagatc 1320

gagaagatcc tgaccttcag aatcccctac tacgtgggcc ctctggccag gggaaacagc 1380

agattcgcct ggatgaccag aaagagcgag gaaaccatca ccccctggaa cttcgaggaa 1440

gtggtggaca agggcgccag cgcccagagc ttcatcgagc ggatgaccaa cttcgataag 1500

aacctgccca acgagaaggt gctgcccaag cacagcctgc tgtacgagta cttcaccgtg 1560

tacaacgagc tgaccaaagt gaaatacgtg accgagggaa tgcggaagcc cgcctttctg 1620

agcggcgagc agaaaaaggc catcgtggac ctgctgttca agaccaaccg gaaagtgacc 1680

gtgaagcagc tgaaagagga ctacttcaag aaaatcgagt gcttcgacag cgtggaaatc 1740

agcggcgtgg aagatcggtt caacgcctcc ctgggcgcct atcacgatct gctgaaaatt 1800

atcaaggaca aggacttcct ggacaatgag gaaaacgagg acattctgga agatatcgtg 1860

ctgaccctga cactgtttga ggaccggggc atgatcgagg aacggctgaa aacctatgcc 1920

cacctgttcg acgacaaagt gatgaagcag ctgaagcggc ggagatacac cggctggggc 1980

aggctgagcc ggaagctgat caacggcatc cgggacaagc agtccggcaa gacaatcctg 2040

gatttcctga agtccgacgg cttcgccaac agaaacttca tgcagctgat ccacgacgac 2100

agcctgacct ttaaagagga catccagaaa gcccaggtgt ccggccaggg acactctctg 2160

cacgagcaga tcgccaatct ggccggatcc cccgccatta agaagggcat cctgcagaca 2220

gtgaagattg tggacgagct cgtgaaagtg atgggccaca agcccgagaa catcgtgatc 2280

gaaatggcca gagagaacca gaccacccag aagggacaga agaacagccg cgagagaatg 2340

aagcggatcg aagagggcat caaagagctg ggcagccaga tcctgaaaga acaccccgtg 2400

gaaaacaccc agctgcagaa cgagaagctg tacctgtact acctgcagaa tgggcgggat 2460

atgtacgtgg accaggaact ggacatcaac cggctgtccg actacgatgt ggaccacatt 2520

gtgccccagt ccttcatcaa ggacgactcc atcgataaca aagtgctgac tcggagcgac 2580

aagaaccggg gcaagagcga caacgtgccc tccgaagagg tcgtgaagaa gatgaagaac 2640

tactggcgcc agctgctgaa tgccaagctg attacccaga ggaagttcga caatctgacc 2700

aaggccgaga gaggcggcct gagcgaactg gataaggccg gcttcattaa gcggcagctg 2760

gtggaaaccc ggcagatcac aaagcacgtg gcacagatcc tggactcccg gatgaacact 2820

aagtacgacg agaacgacaa actgatccgg gaagtgaaag tgatcaccct gaagtccaag 2880

ctggtgtccg acttcagaaa ggatttccag ttttacaaag tgcgcgagat caacaactac 2940

caccacgccc acgacgccta cctgaacgcc gtcgtgggaa ccgccctgat caaaaagtac 3000

cctaagctgg aaagcgagtt cgtgtacggc gattacaagg tgtacgacgt gcggaagatg 3060

atcgccaaga gcgagcagga aatcggcaag gctaccgcca agtacttctt ctacagcaac 3120

atcatgaact ttttcaagac cgagatcaca ctggccaacg gcgagatcag aaagcggcct 3180

ctgatcgaga caaacggcga aaccggggag atcgtgtggg ataagggccg ggattttgcc 3240

acagtgcgga aagtgctgtc catgccccaa gtgaatatcg tgaaaaagac cgaggtgcag 3300

accggcggct tcagcaaaga gtctatcctg cccaagagga actccgacaa gctgatcgcc 3360

agaaagaagg attgggaccc taagaagtac ggcggctttg acagccccac cgtggcctac 3420

tctgtgctgg tggtggccaa agtggaaaag ggcaagtcca agaaactgaa gagtgtgaaa 3480

gagctgctgg ggatcaccat catggaaaga agcagcttcg agaagaatcc catcgacttt 3540

ctggaagcca agggctacaa agaagtgaaa aaggacctga tcatcaagct gcctaagtac 3600

tccctgttcg agctggaaaa cggccggaag cggatgctgg cttctgccgg cgaactgcag 3660

aagggaaacg agctggccct gccctccaaa tatgtgaact tcctgtacct ggccagccac 3720

tatgagaagc tgaagggctc ccccgaggat aatgagcaga aacagctgtt tgtggaacag 3780

cacaagcact acctggacga gatcatcgag cagattagcg agttctccaa gcgcgtgatc 3840

ctggccgatg ccaacctgga caaggtgctg agcgcctaca acaagcaccg ggataagccc 3900

atcagagagc aggccgagaa tatcatccac ctgtttaccc tgaccaacct gggagcccct 3960

gccgccttca agtactttga caccaccatc gaccggaaga ggtacaccag caccaaagag 4020

gtgctggacg ccaccctgat ccaccagagc atcaccggcc tgtacgagac acggatcgac 4080

ctgtctcagc tgggaggcga ccccaagaaa aagcgcaaag tg 4122

<210> 11
<211> 4764
<212> DNA
<213> Homo sapiens

<400> 11
gcgggcgggc ggtgcgatgt ccggagagga tggcccggcg gctggcccgg gggcggcggc 60

ggcggctgcc cgggagcggc gacgggagca gctgcggcag tgggggggcgc gggcgggcgc 120

cgagcctggc cccggagagc gccgcgcccg caccgtccgc ttcgagcgcg ccgccgagtt 180

cctggcggcc tgtgcgggcg gcgacctgga cgaggcgcgt ctgatgctgc gcgccgccga 240

ccctggcccc ggcgccgagc tcgaccccgc cgcgccgccg cccgcccgcg ccgtgctgga 300

ctccaccaac gccgacggta tcagcgccct gcaccaggtc agcgcccccc gcccggcgtc 360

tcccggggcc aggtccaccc tctgctgcgc cacctggggc atcctccttc cccgttgcca 420

gtctcgatcc gccccgtcgt tcctggccct gggctttgcc accctatgct gacaccccgt 480

cccagtcccc cttaccattc cccttcgacc accccacttc cgaattggag ccgcttcaac 540

tggccctggg cttagccact ctgtgctgac cactctgccc caggcctcct taccattccc 600

cttcgaccta ctctcttccg cattggagtc gctttaactg gccctggctt tggcagcctg 660

tgctgaccca tgcagtcctc cttaccatcc ctccctcgac ttcccctctt ccgatgttga 720

gcccctccag ccggtcctgg actttgtctc cttccctgcc ctgccctctc ctgaacctga 780

gccagctccc atagctcagt ctggtctatc tgcctggccc tggccattgt cactttgcgc 840

tgccctcctc tcgcccccga gtgcccttgc tgtgccgccg gaactctgcc ctctaacgct 900

gccgtctctc tcctgagtcc ggaccacttt gagctctact ggcttctgcg ccgcctctgg 960

cccactgttt ccccttccca ggcaggtcct gctttctctg acctgcattc tctccctgg 1020

gcctgtgccg ctttctgtct gcagcttgtg gcctgggtca cctctacggc tggcccagat 1080

ccttccctgc cgcctccttc aggttccgtc ttcctccact ccctcttccc cttgctctct 1140

gctgtgttgc tgcccaagga tgctctttcc ggagcacttc cttctcggcg ctgcaccacg 1200

tgatgtcctc tgagcggatc ctccccgtgt ctgggtcctc tccgggcatc tctcctccct 1260

cacccaaccc catgccgtct tcactcgctg ggttcccttt tccttctcct tctggggcct 1320

gtgccatctc tcgtttctta ggatggcctt ctccgacgga tgtctccctt gcgtcccgcc 1380

tccccttctt gtaggcctgc atcatcaccg tttttctgga caaccccaaa gtaccccgtc 1440

tccctggctt tagccacctc tccatcctct tgctttcttt gcctggacac cccgttctcc 1500

tgtggattcg ggtcacctct cactcctttc atttgggcag ctcccctacc ccccttacct 1560

ctctagtctg tgctagctct tccagccccc tgtcatggca tcttccaggg gtccgagagc 1620

tcagctagtc ttcttcctcc aacccgggcc cctatgtcca cttcaggaca gcatgtttgc 1680

tgcctccagg gatcctgtgt ccccgagctg ggaccacctt atattcccag ggccggttaa 1740

tgtggctctg gttctgggta cttttatctg tcccctccac cccacagtgg ggccactagg 1800

gacaggattg gtgacagaaa agccccatcc ttaggcctcc tccttcctag tctcctgata 1860

ttgggtctaa cccccacctc ctgttaggca gattccttat ctggtgacac acccccattt 1920

cctggagcca tctctctcct tgccagaacc tctaaggttt gcttacgatg gagccagaga 1980

ggatcctggg agggagagct tggcaggggg tgggagggaa ggggggggatg cgtgacctgc 2040

ccggttctca gtggccaccc tgcgctaccc tctcccagaa cctgagctgc tctgacgcgg 2100

ccgtctggtg cgtttcactg atcctggtgc tgcagcttcc ttacacttcc caagaggaga 2160

agcagtttgg aaaaacaaaa tcagaataag ttggtcctga gttctaactt tggctcttca 2220

cctttctagt ccccaattta tattgttcct ccgtgcgtca gttttacctg tgagataagg 2280

ccagtagcca gccccgtcct ggcagggctg tggtgaggag gggggtgtcc gtgtggaaaa 2340

ctccctttgt gagaatggtg cgtcctaggt gttcaccagg tcgtggccgc ctctactccc 2400

tttctctttc tccatccttc tttccttaaa gagtccccag tgctatctgg gacatattcc 2460

tccgcccaga gcagggtccc gcttccctaa ggccctgctc tgggcttctg ggtttgagtc 2520

cttggcaagc ccaggagagg cgctcaggct tccctgtccc ccttcctcgt ccaccatctc 2580

atgcccctgg ctctcctgcc ccttccctac aggggttcct ggctctgctc ttcagactga 2640

gccccgttcc cctgcatccc cgttcccctg catcccccct cccctgcatc ccccagaggc 2700

cccaggccac ctacttggcc tggaccccac gagaggccac cccagccctg tctaccaggc 2760

tgccttttgg gtggattctc ctccaactgt ggggtgactg cttggcaaac tcactcttcg 2820

gggtatccca ggaggcctgg agcattgggg tgggctgggg ttcagagagg agggattccc 2880

ttctcaggtt acgtggccaa gaagcagggg agctgggttt gggtcaggtc tgggtgtggg 2940

gtgaccagct tatgctgttt gcccaggaca gcctagtttt agcactgaaa ccctcagtcc 3000

taggaaaaca gggatggttg gtcactgtct ctgggtgact cttgattccc ggccagtttc 3060

tccacctggg gctgtgtttc tcgtcctgca tccttctcca ggcaggtccc caagcatcgc 3120

ccccctgctg tggctgttcc caagttctta gggtacccca cgtgggttta tcaaccactt 3180

ggtgaggctg gtaccctgcc cccattcctg caccccaatt gccttagtgg ctaggggttt 3240

gggggctaga gtaggagggg ctggagccag gattcttagg gctgaacaga gaagagctgg 3300

gggcctgggc tcctgggttt gagagaggag gggctggggc ctggactcct gggtccgagg 3360

gaggaggggc tggggcctgg actcctgggt ctgagggtgg agggactggg ggcctggact 3420

cctgggtccg agggaggagg ggctggggcc tggactcgtg ggtctgaggg aggaggggct 3480

gggggcctgg acttctgggt cttagggagg cggggctggg cctggacccc tgggtctgaa 3540

tggggagagg ctgggggcct ggactccttc atctgagggc ggaagggctg gggcctggcc 3600

tcctgggttg aatggggagg ggttgggcct ggactctgga gtccctggtg cccaggcctc 3660

aggcatcttt cacagggatg cctgtactgg gcaggtcctt gaaagggaaa ggcccattgc 3720

tctccttgcc cccctcccct atcgccatga caactgggtg gaaataaacg agccgagttc 3780

atcccgttcc cagggcacgt gcggcccctt cacagcccga gtttccatga cctcatgctc 3840

ttggccctcg tagctccctc ccgcctcctc cagatgggca gctttggaga ggtgagggac 3900

ttgggggta atttatcccg tggatctagg agtttagctt cactccttcc tcagctccag 3960

ttcaggtccc ggagcccacc cagtgtccac aaggcctggg gcaagtccct cctccgaccc 4020

cctggacttc ggcttttgtc cccccaagtt ttggacccct aagggaagaa tgagaaacgg 4080

tggcccgtgt cagcccctgg ctgcagggcc ccgtgcagag ggggcctcag tgaactggag 4140

tgtgacagcc tggggcccag gcacacaggt gtgcagctgt ctcacccctc tgggagtccc 4200

gcccaggccc ctgagtctgt cccagcacag ggtggccttc ctccaccctg catagccctg 4260

ggcccacggc ttcgttcctg cagagtatct gctggggtgg tttccgagct tgacccttgg 4320

aaggacctgg ctgggtttaa ggcaggaggg gctgggggcc aggactcctg gctctgaagg 4380

aggaggggct ggaacctctt ccctagtctg agcactggaa gcgccacctg tgggtggtga 4440

cgggggtttt gccgtgtcta acaggtacca tgtggggttc ccgcacccag atgagaagcc 4500

ccctcccttc cccgttcact tcctgtttgc agatagccag gagtcctttc gtggtttcca 4560

ctgagcactg aaggcctggc cggcctgacc actgggcaac caggcgtatc ttaaacagcc 4620

agtggccaga ggctgttggg tcattttccc cactgtccta gcaccgtgtc cctggatctg 4680

ttttcgtggc tccctctgga gtcccgactt gctgggacac cgtggctggg gtaggtgcgg 4740

ctgacggctg tttcccaccc ccag 4764

<210> 12
<211> 42
<212> RNA
<213> Artificial Sequence

<220>
<223> SYNTHESIZED

<400> 12
accccacagu ggggccacua guuuuagagc uaugcuguuu ug 42

<210> 13
<211> 86
<212> RNA
<213> Artificial Sequence

<220>
<223> SYNTHESIZED

<400> 13
ggaaccauuc aaaacagcau agcaaguuaa aauaaggcua guccguuauc aacuugaaaa 60

aguggcaccg agucggugcu uuuuuu 86

<210> 14
<211> 62
<212> RNA
<213> Artificial Sequence

<220>
<223> SYNTHESIZED

<400> 14
accccacagu ggggccacua guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc 60

cg 62

<210> 15
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> SYNTHESIZED

<400> 15
ctgacctctt ctcttcctcc cacag 25

<210> 16
<211> 1009
<212> DNA
<213> Artificial Sequence

<220>
<223> SYNTHESIZED

<400> 16
gccaccatgg actacaaaga cgatgacgac aaggtcgact ctagagctgc agagagcgac 60

gagagcggcc tgcccgccat ggagatcgag tgccgcatca ccggcaccct gaacggcgtg 120

gagttcgagc tggtgggcgg cggagagggc acccccgagc agggccgcat gaccaacaag 180

atgaagagca ccaaaggcgc cctgaccttc agcccctacc tgctgagcca cgtgatgggc 240

tacggcttct accacttcgg cacctacccc agcggctacg agaacccctt cctgcacgcc 300

atcaacaacg gcggctacac caacacccgc atcgagaagt acgaggacgg cggcgtgctg 360

cacgtgagct tcagctaccg ctacgaggcc ggccgcgtga tcggcgactt caaggtgatg 420

ggcaccggct tccccgagga cagcgtgatc ttcaccgaca agatcgtccg cagcaacgcc 480

accgtggagc acctgcaccc catgggcgat aacgatctgg atggcagctt cacccgcacc 540

ttcagcctgc gcgacggcgg ctactacagc tccgtggtgg acagccacat gcacttcaag 600

agcgccatcc accccagcat cctgcagaac gggggcccca tgttcgcctt ccgccgcgtg 660

gaggaggatc acagcaacac cgagctgggc atcgtggagt accagcacgc cttcaagacc 720

ccggatgcag atgccggtga agaatgaaga tctctgtgcc ttctagttgc cagccatctg 780

ttgtttgccc ctccccgtg ccttccttga ccctggaagg tgccactccc actgtccttt 840

cctaataaaa tgaggaaatt gcatcgcatt gtctgagtag gtgtcattct attctggggg 900

gtggggtggg gcaggacagc aaggggagg attgggaaga caatagcagg catgctgggg 960

atgcggtggg ctctatggac tcgaggttta aacgtcgacg cggccgcgt 1009

<210> 17
<211> 355
<212> PRT
<213> Artificial Sequence

<220>
<223> SYNTHESIZED

<400> 17

Met Ser Gly Glu Asp Gly Pro Ala Ala Gly Pro Gly Ala Ala Ala Ala
1 5 10 15

Ala Ala Arg Glu Arg Arg Arg Glu Gln Leu Arg Gln Trp Gly Ala Arg
20 25 30

Ala Gly Ala Glu Pro Gly Pro Gly Glu Arg Arg Ala Arg Thr Val Arg

```
        35                  40                  45

Phe Glu Arg Ala Ala Glu Phe Leu Ala Ala Cys Ala Gly Gly Asp Leu
    50                  55                  60

Asp Glu Ala Arg Leu Met Leu Arg Ala Ala Asp Pro Gly Pro Gly Ala
65                  70                  75                  80

Glu Leu Asp Pro Ala Ala Pro Pro Pro Ala Arg Ala Val Leu Asp Ser
                85                  90                  95

Thr Asn Ala Asp Gly Ile Ser Ala Leu His Gln Ala Thr Met Asp Tyr
            100                 105                 110

Lys Asp Asp Asp Asp Lys Val Asp Ser Arg Ala Ala Glu Ser Asp Glu
        115                 120                 125

Ser Gly Leu Pro Ala Met Glu Ile Glu Cys Arg Ile Thr Gly Thr Leu
    130                 135                 140

Asn Gly Val Glu Phe Glu Leu Val Gly Gly Gly Glu Gly Thr Pro Glu
145                 150                 155                 160

Gln Gly Arg Met Thr Asn Lys Met Lys Ser Thr Lys Gly Ala Leu Thr
                165                 170                 175

Phe Ser Pro Tyr Leu Leu Ser His Val Met Gly Tyr Gly Phe Tyr His
            180                 185                 190

Phe Gly Thr Tyr Pro Ser Gly Tyr Glu Asn Pro Phe Leu His Ala Ile
        195                 200                 205

Asn Asn Gly Gly Tyr Thr Asn Thr Arg Ile Glu Lys Tyr Glu Asp Gly
    210                 215                 220

Gly Val Leu His Val Ser Phe Ser Tyr Arg Tyr Glu Ala Gly Arg Val
225                 230                 235                 240

Ile Gly Asp Phe Lys Val Met Gly Thr Gly Phe Pro Glu Asp Ser Val
                245                 250                 255

Ile Phe Thr Asp Lys Ile Val Arg Ser Asn Ala Thr Val Glu His Leu
            260                 265                 270

His Pro Met Gly Asp Asn Asp Leu Asp Gly Ser Phe Thr Arg Thr Phe
        275                 280                 285

Ser Leu Arg Asp Gly Gly Tyr Tyr Ser Ser Val Val Asp Ser His Met
    290                 295                 300

His Phe Lys Ser Ala Ile His Pro Ser Ile Leu Gln Asn Gly Gly Pro
```

305 310 315 320

Met Phe Ala Phe Arg Arg Val Glu Glu Asp His Ser Asn Thr Glu Leu
325 330 335

Gly Ile Val Glu Tyr Gln His Ala Phe Lys Thr Pro Asp Ala Asp Ala
340 345 350

Gly Glu Glu
355

<210> 18
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> SYNTHESIZED

<400> 18
ccactctgtg ctgaccactc t 21

<210> 19
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> SYNTHESIZED

<400> 19
gcggcactcg atctcca 17

<210> 20
<211> 711
<212> DNA
<213> Mus musculus

<400> 20
gagcggctgc ggggcgggtg caagcacgtt tccgacttga gttgcctcaa gaggggcgtg 60

ctgagccaga cctccatcgc gcactccggg gagtggaggg aaggagcgag ggctcagttg 120

ggctgttttg gaggcaggaa gcacttgctc tcccaaagtc gctctgagtt gttatcagta 180

agggagctgc agtggagtag gcggggagaa ggccgcaccc ttctccggag gggggagggg 240

agtgttgcaa tacctttctg ggagttctct gctgcctcct ggcttctgag gaccgccctg 300

ggcctgggag aatcccttcc ccctcttccc tcgtgatctg caactccagt ctttctagaa 360

gatgggcggg agtcttctgg gcaggcttaa aggctaacct ggtgtgtggg cgttgtcctg 420

caggggaatt gaacaggtgt aaaattggag ggacaagact tcccacagat tttcggtttt 480

gtcgggaagt tttttaatag gggcaaataa ggaaaatggg aggataggta gtcatctggg 540

gttttatgca gcaaaactac aggttattat tgcttgtgat ccgcctcgga gtattttcca 600

tcgaggtaga ttaaagacat gctcacccga gttttatact ctcctgcttg agatccttac 660

tacagtatga aattacagtg tcgcgagtta gactatgtaa gcagaatttt a 711

<210> 21
<211> 42
<212> RNA
<213> Artificial Sequence

<220>
<223> SYNTHESIZED

<400> 21
cuccagucuu ucuagaagau guuuuagagc uaugcuguuu ug 42

<210> 22
<211> 42
<212> RNA
<213> Artificial Sequence

<220>
<223> SYNTHESIZED

<400> 22
ugaacaggug uaaaauugga guuuuagagc uaugcuguuu ug 42

<210> 23
<211> 42
<212> RNA
<213> Artificial Sequence

<220>
<223> SYNTHESIZED

<400> 23
ugucgggaag uuuuuuaaua guuuuagagc uaugcuguuu ug 42

<210> 24
<211> 642
<212> DNA
<213> Rattus rattus

<400> 24
gggattcctc cttgagttgt ggcactgagg aacgtgctga acaagaccta cattgcactc 60

cagggagtgg atgaaggagt tggggctcag tcgggttgta ttggagacaa gaagcacttg 120

ctctccaaaa gtcggtttga gttatcatta agggagctgc agtggagtag gcggagaaaa 180

ggccgcaccc ttctcaggac gggggagggg agtgttgcaa tacctttctg ggagttctct 240

gctgcctcct gtcttctgag gaccgccctg ggcctggaag attcccttcc cccttcttcc 300

ctcgtgatct gcaactggag tctttctgga agataggcgg gagtcttctg ggcaggctta 360

aaggctaacc tggtgcgtgg ggcgttgtcc tgcagaggaa ttgaacaggt gtaaaattgg 420

aggggcaaga cttcccacag attttcgatt gtgttgttaa gtattgtaat aggggcaaat 480

aagggaaata gactaggcac tcacctgggg ttttatgcag caaaactaca ggttattatt 540

gcttgtgatc cgccctggag aatttttcac cgaggtagat tgaagacatg cccacccaaa 600

ttttaatatt cttccacttg cgatccttgc tacagtatga aa 642

<210> 25

<211> 42
<212> RNA
<213> Artificial Sequence

<220>
<223> SYNTHESIZED

<400> 25
aggggaagg gaaucuucca guuuuagagc uaugcuguuu ug 42

<210> 26
<211> 42
<212> RNA
<213> Artificial Sequence

<220>
<223> SYNTHESIZED

<400> 26
ucugcaacug gagucuuucu guuuuagagc uaugcuguuu ug 42

<210> 27
<211> 42
<212> RNA
<213> Artificial Sequence

<220>
<223> SYNTHESIZED

<400> 27
aggcgggagu cuucugggca guuuuagagc uaugcuguuu ug 42