

# SEQUENCE LISTING

<110> Regeneron Pharmaceuticals, Inc.

<120> NON-HUMAN ANIMALS HAVING A DISRUPTION IN A C9ORF72 LOCUS

<130> 32698 (10152-US01)

<150> 62/168,171

<151> 2015-05-29

<150> 62/232,658

<151> 2015-09-25

<150> 62/245,382

<151> 2015-10-23

<160> 9

<170> PatentIn version 3.5

<210> 1

<211> 3198

<212> DNA

<213> Mus musculus

<400> 1

gtgtccgggg cggggcggtc ccggggcggg gcccggagcg ggctgcgggtt gcggtccttg 60

cgccggcgggt gaaggcgag cagcggcgag tggctattgc aagcggtcgg ataatgtgag 120

acctggaatg cagtgaagacc tgggatgcag ggatgtcgac tatctgcccc ccaccatctc 180

ctgctgttgc caagacagag attgctttta gtggtgaatc acccttggtg gcggctacct 240

ttgcttactg ggataatatt cttggtccta gagtaaggca tatttgggct ccaaagacag 300

accaagtgtc tctcagtgat ggagaaataa cttttcttgc caaccacact ctaaattggag 360

aaattcttcg aaatgcagag agtggggcta tagatgtaaa atttttgtc ttatctgaaa 420

aaggggtaat tattgtttca ttaatcttcg acggaaactg gaatggagat cggagcactt 480

atggactatc aattatactg ccgcagacag agctgagctt ctacctcca cttcacagag 540

tgtgtgttga caggctaaca cacattattc gaaaaggaag aatatggatg cataaggaaa 600

gacaagaaaa tgtccagaaa attgtcttgg aaggcacaga gaggatggaa gatcagggtc 660

agagtatcat tcccatgctt actggggaag tcattcctgt aatggagctg cttgcatcta 720

|                                                                    |      |
|--------------------------------------------------------------------|------|
| tgaaatccca cagtgttcct gaagacattg atatagctga tacagtgctc aatgatgatg  | 780  |
| acattggtga cagctgtcac gaaggctttc ttctcaatgc catcagctca cacctgcaga  | 840  |
| cctgtggctg ttccgttgta gttggcagca gtgcagagaa agtaaataag atagtaagaa  | 900  |
| cgctgtgcct ttttctgaca ccagcagaga ggaaatgctc caggctgtgt gaagcagaat  | 960  |
| cgctccttaa gtacgaatcg ggactctttg tgcaaggctt gctaaaggat gcaacaggca  | 1020 |
| gttttgtcct acccttccgg caagttatgt atgccccgta cccaccacg cacattgatg   | 1080 |
| tggatgtcaa cactgtcaag cagatgccac cgtgtcatga acatatttat aatcaacgca  | 1140 |
| gatacatgag gtcagagctg acagccttct ggagggaac ttcagaagag gacatggcgc   | 1200 |
| aggacaccat catctacaca gatgagagct tcactcctga tttgaatatt ttccaagatg  | 1260 |
| tcttacacag agacactcta gtgaaagcct tcctggatca ggtcttccat ttgaagcctg  | 1320 |
| gcctgtctct caggagtact ttccttgca agttcctcct cattcttcac agaaaagcct   | 1380 |
| tgacactaat caagtacatc gaggatgata cgcagaaggg gaaaaagccc tttaatgtctc | 1440 |
| ttcggaacct gaagatagat cttgatttaa cagcagaggg cgatcttaac ataataatgg  | 1500 |
| ctctagctga gaaaattaag ccaggcctac actctttcat ctttgggaga cttttctaca  | 1560 |
| ctagtgtaca agaacgtgat gttctaata ccttttgacc gtgtggtttg ctgtgtctgt   | 1620 |
| ctcttcacag tcacacctgc tgttacagtg tctcagcagt gtgtgggcac atccttcctc  | 1680 |
| ccgagtcctg ctgcaggaca gggtagacta cacttgtag tagaagtctg tacctgatgt   | 1740 |
| caggatgcac gttacagtga atgactcttc ctagaataga tgtactcttt tagggcctta  | 1800 |
| tgtttacaat tatcctaagt actattgctg tcttttaaag atatgaatga tggaatatac  | 1860 |
| acttgaccat aactgctgat tgggtttttg ttttgtttg tttgttttct tggaaactta   | 1920 |
| tgattcctgg ttacatgta ccacactgaa accctcgtta gctttacaga taaagtgtga   | 1980 |
| gttgacttcc tgcccctctg tgttctgtgg tatgtccgat tacttctgcc acagctaac   | 2040 |
| attagagcat ttaaagttag cagttcctca gaaaggaact tagtctgact acagattagt  | 2100 |
| tcttgagaga agacactgat agggcagagc tgtaggtgaa atcagttgtt agcccttcct  | 2160 |
| ttatagacgt agtccttcag attcgggtctg tacagaaatg ccgaggggtc atgcatgggc | 2220 |

|                                                                      |      |
|----------------------------------------------------------------------|------|
| cctgagtatc gtgacctgtg acaagttttt tgttggttta ttgtagttct gtcaaagaaa    | 2280 |
| gtggcatttg tttttataat tgttgccaac ttttaagggtt aattttcatt atttttgagc   | 2340 |
| cgaattaaaa tgcgcacctc ctgtgccttt cccaatcttg gaaaatataa tttcttggca    | 2400 |
| gagggtcaga tttcagggcc cagtcacttt catctgacca ccttttgcac ggctgccgtg    | 2460 |
| tgcttggtt agattagaag tccttggtta gtatgtcaga gtacattcgc tgataagatc     | 2520 |
| tttgaagagc aggggaagcgt cttgcctctt tcctttgggtt tctgcctgta ctctgggtgtt | 2580 |
| tcccggtgtc cctgcatcat aggaacagca gagaaatctg acccagtgct atttttctag    | 2640 |
| gtgctactat ggcaaactca agtgggtctgt ttctgttcct gtaacgttcg actatctcgc   | 2700 |
| tagctgtgaa gtactgatta gtggagttct gtgcaacagc agtgtaggag tatacacaaa    | 2760 |
| cacaaatatg tgtttctatt taaaactgtg gacttagcat aaaaaggag aatatattta     | 2820 |
| ttttttacaa aagggataaa aatgggcccc gttcctcacc caccagattt agcgagaaaa    | 2880 |
| agctttctat tctgaaaggt cacggtggct ttggcattac aaatcagaac aacacacact    | 2940 |
| gaccatgatg gcttgtgaac taactgcaag gcactccgtc atggtaagcg agtaggtccc    | 3000 |
| acctcctagt gtgccgctca ttgctttaca cagtagaatc ttatttgagt gctaattgtt    | 3060 |
| gtctttgctg ctttactgtg ttgttataga aaatgtaagc tgtacagtga ataagttatt    | 3120 |
| gaagcatgtg taaacactgt tatatatctt ttctcctaga tggggaattt tgaataaaat    | 3180 |
| acctttgaaa ttctgtgt                                                  | 3198 |

<210> 2  
 <211> 481  
 <212> PRT  
 <213> Mus musculus

<400> 2

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Ser | Thr | Ile | Cys | Pro | Pro | Pro | Ser | Pro | Ala | Val | Ala | Lys | Thr | Glu |
| 1   |     |     |     | 5   |     |     |     | 10  |     |     |     |     |     | 15  |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ile | Ala | Leu | Ser | Gly | Glu | Ser | Pro | Leu | Leu | Ala | Ala | Thr | Phe | Ala | Tyr |
|     |     |     | 20  |     |     |     | 25  |     |     |     |     |     | 30  |     |     |

Trp Asp Asn Ile Leu Gly Pro Arg Val Arg His Ile Trp Ala Pro Lys  
35 40 45

Thr Asp Gln Val Leu Leu Ser Asp Gly Glu Ile Thr Phe Leu Ala Asn  
50 55 60

His Thr Leu Asn Gly Glu Ile Leu Arg Asn Ala Glu Ser Gly Ala Ile  
65 70 75 80

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

Leu Ile Phe Asp Gly Asn Trp Asn Gly Asp Arg Ser Thr Tyr Gly Leu  
100 105 110

Ser Ile Ile Leu Pro Gln Thr Glu Leu Ser Phe Tyr Leu Pro Leu His  
115 120 125

Arg Val Cys Val Asp Arg Leu Thr His Ile Ile Arg Lys Gly Arg Ile  
130 135 140

Trp Met His Lys Glu Arg Gln Glu Asn Val Gln Lys Ile Val Leu Glu  
145 150 155 160

Gly Thr Glu Arg Met Glu Asp Gln Gly Gln Ser Ile Ile Pro Met Leu  
165 170 175

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

His Ser Val Pro Glu Asp Ile Asp Ile Ala Asp Thr Val Leu Asn Asp  
195 200 205

Asp Asp Ile Gly Asp Ser Cys His Glu Gly Phe Leu Leu Asn Ala Ile  
210 215 220

Ser Ser His Leu Gln Thr Cys Gly Cys Ser Val Val Val Gly Ser Ser

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 225 |     | 230 |     | 235 |     | 240 |     |     |     |     |     |     |     |     |     |
| Ala | Glu | Lys | Val | Asn | Lys | Ile | Val | Arg | Thr | Leu | Cys | Leu | Phe | Leu | Thr |
|     |     |     | 245 |     |     |     |     |     | 250 |     |     |     |     | 255 |     |
| Pro | Ala | Glu | Arg | Lys | Cys | Ser | Arg | Leu | Cys | Glu | Ala | Glu | Ser | Ser | Phe |
|     |     |     | 260 |     |     |     |     | 265 |     |     |     |     | 270 |     |     |
| Lys | Tyr | Glu | Ser | Gly | Leu | Phe | Val | Gln | Gly | Leu | Leu | Lys | Asp | Ala | Thr |
|     |     | 275 |     |     |     |     | 280 |     |     |     |     | 285 |     |     |     |
| Gly | Ser | Phe | Val | Leu | Pro | Phe | Arg | Gln | Val | Met | Tyr | Ala | Pro | Tyr | Pro |
|     | 290 |     |     |     |     |     | 295 |     |     |     | 300 |     |     |     |     |
| Thr | Thr | His | Ile | Asp | Val | Asp | Val | Asn | Thr | Val | Lys | Gln | Met | Pro | Pro |
| 305 |     |     |     |     | 310 |     |     |     |     | 315 |     |     |     |     | 320 |
| Cys | His | Glu | His | Ile | Tyr | Asn | Gln | Arg | Arg | Tyr | Met | Arg | Ser | Glu | Leu |
|     |     |     |     | 325 |     |     |     |     | 330 |     |     |     |     | 335 |     |
| Thr | Ala | Phe | Trp | Arg | Ala | Thr | Ser | Glu | Glu | Asp | Met | Ala | Gln | Asp | Thr |
|     |     |     | 340 |     |     |     |     |     | 345 |     |     |     |     | 350 |     |
| Ile | Ile | Tyr | Thr | Asp | Glu | Ser | Phe | Thr | Pro | Asp | Leu | Asn | Ile | Phe | Gln |
|     |     | 355 |     |     |     |     | 360 |     |     |     |     | 365 |     |     |     |
| Asp | Val | Leu | His | Arg | Asp | Thr | Leu | Val | Lys | Ala | Phe | Leu | Asp | Gln | Val |
|     | 370 |     |     |     |     | 375 |     |     |     |     | 380 |     |     |     |     |
| Phe | His | Leu | Lys | Pro | Gly | Leu | Ser | Leu | Arg | Ser | Thr | Phe | Leu | Ala | Gln |
| 385 |     |     |     |     | 390 |     |     |     |     | 395 |     |     |     |     | 400 |
| Phe | Leu | Leu | Ile | Leu | His | Arg | Lys | Ala | Leu | Thr | Leu | Ile | Lys | Tyr | Ile |
|     |     |     |     | 405 |     |     |     |     | 410 |     |     |     |     | 415 |     |
| Glu | Asp | Asp | Thr | Gln | Lys | Gly | Lys | Lys | Pro | Phe | Lys | Ser | Leu | Arg | Asn |
|     |     |     | 420 |     |     |     |     | 425 |     |     |     |     | 430 |     |     |

Leu Lys Ile Asp Leu Asp Leu Thr Ala Glu Gly Asp Leu Asn Ile Ile  
435 440 445

Met Ala Leu Ala Glu Lys Ile Lys Pro Gly Leu His Ser Phe Ile Phe  
450 455 460

Gly Arg Pro Phe Tyr Thr Ser Val Gln Glu Arg Asp Val Leu Met Thr  
465 470 475 480

Phe

<210> 3  
<211> 3435  
<212> DNA  
<213> Rattus norvegicus

<400> 3  
cgtttgtagt gtcagccatc ccaattgcct gttccttctc tgtgggagtg gtgtctagac 60  
agtccaggca gggatatgcta ggcagggtgcg ttttggttgc ctcagatcgc aacttgactc 120  
cataacggtg accaaagaca aaagaaggaa accagattaa aaagaaccgg acacagaccc 180  
ctgcagaatc tggagcggcc gtggttgggg gcggggctac gacggggcgg actcgggggc 240  
gtgggagggc ggggcccggg cggggcccgg agccggctgc ggttgcggtc cctgcgccgg 300  
cgggtgaaggc gcagcggcgg cgagtggcta ttgcaagcgt ttggataatg tgagacctgg 360  
gatgcaggga tgtcgactat ctgccccca ccatctcctg ctgttgccaa gacagagatt 420  
gctttaagtg gtgaatcacc cttgttggcg gctacctttg cttactggga taatattctt 480  
ggtcctagag taaggcacat ttgggctcca aagacagacc aagtactcct cagtgatgga 540  
gaaatcactt ttcttgccaa ccacactctg aatggagaaa ttcttcggaa tgcggagagt 600  
ggggcaatag atgtaaagtt ttttgtctta tctgaaaagg gcgtcattat tgtttcatta 660  
atcttcgacg ggaactggaa cggagatcgg agcacttacg gactatcaat tatactgccg 720  
cagacggagc tgagtttcta cctcccactg cacagagtgt gtgttgacag gctaacgcac 780  
atcattcgaa aaggaaggat atggatgcac aaggaaagac aagaaaatgt ccagaaaatt 840

|                                                                    |      |
|--------------------------------------------------------------------|------|
| gtcttggaag gcaccgagag gatggaagat cagggtcaga gtatcatccc tatgcttact  | 900  |
| ggggagggtca tccctgtgat ggagctgctt gcgtctatga gatcacacag tgttcctgaa | 960  |
| gacctcgata tagctgatac agtactcaat gatgatgaca ttggtgacag ctgtcatgaa  | 1020 |
| ggctttcttc tcaatgccat cagctcacat ctgcagacct gcggctgttc tgtgggtgga  | 1080 |
| ggcagcagtg cagagaaaagt aaataagata gtaagaacac tgtgcctttt tctgacacca | 1140 |
| gcagagagga agtgctccag gctgtgtgaa gccgaatcgt cctttaaata cgaatctgga  | 1200 |
| ctctttgtac aaggcttgct aaaggatgcg actggcagtt ttgtactacc tttccggcaa  | 1260 |
| gttatgtatg ccccttatcc caccacacac atcgatgtgg atgtcaacac tgtcaagcag  | 1320 |
| atgccaccgt gtcatgaaca tatttataat caacgcagat acatgagggtc agagctgaca | 1380 |
| gccttctgga gggcaacttc agaagaggac atggctcagg acaccatcat ctacacagat  | 1440 |
| gagagcttca ctctgatatt gaatatatttc caagatgtct tacacagaga cactctagt  | 1500 |
| aaagcctttc tggatcaggt cttccatttg aagcctggcc tgtctctcag gagtactttc  | 1560 |
| cttgcacagt tcctcctcat tcttcacaga aaagccttga cactaatcaa gtacatagag  | 1620 |
| gatgacacgc agaaggggaa aaagcccttt aagtctcttc ggaacctgaa gatagatctt  | 1680 |
| gatttaacag cagagggcga ccttaacata ataatggctc tagctgagaa aattaagcca  | 1740 |
| ggcctacact ctttcatctt cgggagacct ttctacacta gtgtccaaga acgtgatgtt  | 1800 |
| ctaatgactt tttaaacatg tggtttgctc cgtgtgtctc atgacagtca cacttgctgt  | 1860 |
| tacagtgtct cagcgctttg gacacatcct tcctccaggg tcctgccgca ggacacgtta  | 1920 |
| cactacactt gtcagtagag gtctgtacca gatgtcaggt acatcgttgt agtgaatgtc  | 1980 |
| tcttttccta gactagatgt accctcgtag ggacttatgt ttacaaccct cctaagtact  | 2040 |
| agtgctgtct tgtaaggata cgaatgaagg gatgtaaact tcaccacaac tgctggttgg  | 2100 |
| ttttgttggt tttgtttttt gaaacttata attcatgggt tacatgcac acactgaaac   | 2160 |
| cctagttagc tttttacagg taagctgtga gttgactgcc tgtccctgtg ttctctggcc  | 2220 |
| tgtacgatct gtggcgtgta ggatcacttt tgcaacaact aaaaactaaa gcactttgtt  | 2280 |
| tgcagttcta cagaaagcaa cttagtctgt ctgcagattc gtttttgaaa gaagacatga  | 2340 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| gaaagcggag ttttaggtga agtcagttgt tggatcttcc tttatagact tagtccttta  | 2400 |
| gatgtggtct gtatagacat gccaaccat catgcatggg cactgaatat cgtgaactgt   | 2460 |
| ggtatgcttt ttgttggttt attgtacttc tgtcaaagaa agtggcattg gtttttataa  | 2520 |
| ttgttgccaa gttttaaggt taattttcat tatttttgag ccaaattaaa atgtgcacct  | 2580 |
| cctgtgcctt tccaatctt ggaaaatata atttcttggc agaaggtcag atttcagggc   | 2640 |
| ccagtcactt tcgtctgact tccctttgca cagtccgcca tgggcctggc ttagaagttc  | 2700 |
| ttgtaaacta tgccagagag tacattcgct gataaaatct tctttgcaga gcaggagagc  | 2760 |
| ttcttgccctc tttcctttca tttctgcctg gactttggtg ttctccacgt tccctgcatc | 2820 |
| ctaaggacag caggagaact ctgaccccag tgctatttct ctaggtgcta ttgtggcaaa  | 2880 |
| ctcaagcggg ccgtctctgt ccctgtaacg ttcgtacctt gctggctgtg aagtactgac  | 2940 |
| tggtaaagct ccgtgctaca gcagtgtagg gtatacacia acacaagtaa gtgttttatt  | 3000 |
| taaaactgtg gacttagcat aaaaaggag actatattta ttttttacia aagggataaa   | 3060 |
| aatggaaccc tttcctcacc caccagattt agtcagaaaa aaacattcta ttctgaaagg  | 3120 |
| tcacagtggg ttgacatga cacatcagaa caacgcacac tgtccatgat ggcttatgaa   | 3180 |
| ctccaagtca ctccatcatg gtaaattggg agatccctcc ttctagtgtg ccacaccatt  | 3240 |
| gcttcccaca gtagaatctt atttaagtgc taagtgttgt ctctgctggg ttactctgtt  | 3300 |
| gttttagaga atgtaagttg tatagtgaat aagttattga agcatgtgta aacactgtta  | 3360 |
| tacatctttt ctcttagatg gggaatttgg aataaaatac ctttaaaatt caaaaaaaaa  | 3420 |
| aaaaaaaaaa aaaaa                                                   | 3435 |

<210> 4  
 <211> 481  
 <212> PRT  
 <213> Rattus norvegicus

<400> 4

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Ser | Thr | Ile | Cys | Pro | Pro | Pro | Ser | Pro | Ala | Val | Ala | Lys | Thr | Glu |
| 1   |     |     |     | 5   |     |     |     | 10  |     |     |     |     |     | 15  |     |

Ile Ala Leu Ser Gly Glu Ser Pro Leu Leu Ala Ala Thr Phe Ala Tyr  
20 25 30

Trp Asp Asn Ile Leu Gly Pro Arg Val Arg His Ile Trp Ala Pro Lys  
35 40 45

Thr Asp Gln Val Leu Leu Ser Asp Gly Glu Ile Thr Phe Leu Ala Asn  
50 55 60

His Thr Leu Asn Gly Glu Ile Leu Arg Asn Ala Glu Ser Gly Ala Ile  
65 70 75 80

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

Leu Ile Phe Asp Gly Asn Trp Asn Gly Asp Arg Ser Thr Tyr Gly Leu  
100 105 110

Ser Ile Ile Leu Pro Gln Thr Glu Leu Ser Phe Tyr Leu Pro Leu His  
115 120 125

Arg Val Cys Val Asp Arg Leu Thr His Ile Ile Arg Lys Gly Arg Ile  
130 135 140

Trp Met His Lys Glu Arg Gln Glu Asn Val Gln Lys Ile Val Leu Glu  
145 150 155 160

Gly Thr Glu Arg Met Glu Asp Gln Gly Gln Ser Ile Ile Pro Met Leu  
165 170 175

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

His Ser Val Pro Glu Asp Leu Asp Ile Ala Asp Thr Val Leu Asn Asp  
195 200 205

Asp Asp Ile Gly Asp Ser Cys His Glu Gly Phe Leu Leu Asn Ala Ile

210

215

220

Ser Ser His Leu Gln Thr Cys Gly Cys Ser Val Val Val Gly Ser Ser  
 225 230 235 240

Ala Glu Lys Val Asn Lys Ile Val Arg Thr Leu Cys Leu Phe Leu Thr  
 245 250 255

Pro Ala Glu Arg Lys Cys Ser Arg Leu Cys Glu Ala Glu Ser Ser Phe  
 260 265 270

Lys Tyr Glu Ser Gly Leu Phe Val Gln Gly Leu Leu Lys Asp Ala Thr  
 275 280 285

Gly Ser Phe Val Leu Pro Phe Arg Gln Val Met Tyr Ala Pro Tyr Pro  
 290 295 300

Thr Thr His Ile Asp Val Asp Val Asn Thr Val Lys Gln Met Pro Pro  
 305 310 315 320

Cys His Glu His Ile Tyr Asn Gln Arg Arg Tyr Met Arg Ser Glu Leu  
 325 330 335

Thr Ala Phe Trp Arg Ala Thr Ser Glu Glu Asp Met Ala Gln Asp Thr  
 340 345 350

Ile Ile Tyr Thr Asp Glu Ser Phe Thr Pro Asp Leu Asn Ile Phe Gln  
 355 360 365

Asp Val Leu His Arg Asp Thr Leu Val Lys Ala Phe Leu Asp Gln Val  
 370 375 380

Phe His Leu Lys Pro Gly Leu Ser Leu Arg Ser Thr Phe Leu Ala Gln  
 385 390 395 400

Phe Leu Leu Ile Leu His Arg Lys Ala Leu Thr Leu Ile Lys Tyr Ile  
 405 410 415

Glu Asp Asp Thr Gln Lys Gly Lys Lys Pro Phe Lys Ser Leu Arg Asn  
420 425 430

Leu Lys Ile Asp Leu Asp Leu Thr Ala Glu Gly Asp Leu Asn Ile Ile  
435 440 445

Met Ala Leu Ala Glu Lys Ile Lys Pro Gly Leu His Ser Phe Ile Phe  
450 455 460

Gly Arg Pro Phe Tyr Thr Ser Val Gln Glu Arg Asp Val Leu Met Thr  
465 470 475 480

Phe

<210> 5  
<211> 680  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Synthetic polynucleotide-prot promoter

<400> 5  
ccagtagcag cacccacgtc caccttctgt ctagtaatgt ccaacacctc cctcagtcca 60  
aacactgctc tgcattccatg tggctcccat ttatacctga agcacttgat ggggcctcaa 120  
tggttttacta gagcccaccc ccctgcaact ctgagaccct ctggatttgt ctgtcagtgc 180  
ctcactgggg cgttggataa tttcttaaaa ggtcaagttc cctcagcagc attctctgag 240  
cagtctgaag atgtgtgctt ttcacagttc aaatccatgt ggctgtttca cccacctgcc 300  
tggccttggg ttatctatca ggacctagcc tagaagcagg tgtgtggcac ttaacaccta 360  
agctgagtga ctaactgaac actcaagtgg atgccatctt tgtcacttct tgactgtgac 420  
acaagcaact cctgatgcca aagccctgcc caccctctc atgcccataat ttggacatgg 480  
tacaggctcct cactggccat ggtctgtgag gtcctgggtcc tctttgactt cataattcct 540  
agggggccact agtatctata agaggaagag ggtgctggct cccaggccac agcccacaaa 600

|                                                                     |     |
|---------------------------------------------------------------------|-----|
| attccacctg ctcacagggtt ggctggctcg acccagggtgg tgtcccctgc tctgagccag | 660 |
| ctcccggcca agccagcacc                                               | 680 |

<210> 6  
 <211> 1052  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic polynucleotide-Blimp1 promoter 1kb

|                                                                    |      |
|--------------------------------------------------------------------|------|
| <400> 6                                                            |      |
| tgccatcatc acaggatgtc cttccttctc cagaagacag actggggctg aaggaaaagc  | 60   |
| cggccagggt cagaacgagc ccctaataatt actgcctcca acagctttcc actcactgcc | 120  |
| cccagcccaa catccccctt ttaactggga agcattccta ctctccattg tacgcacacg  | 180  |
| ctcgggaagcc tggctgtggg tttgggcatg agaggcaggg acaacaaaac cagtatatat | 240  |
| gattataact ttttcctgtt tccctatttc caaatggctg aaaggaggaa gttaggtcta  | 300  |
| cctaagctga atgtattcag ttagcaggag aaatgaaatc ctatacgttt aatactagag  | 360  |
| gagaaccgcc ttagaatatt tatttcattg gcaatgactc caggactaca cagcgaaatt  | 420  |
| gtattgcatg tgctgcaaaa atacttttagc tctttccttc gaagtacgtc ggatcctgta | 480  |
| attgagacac cgagtttagg tgactagggt tttcttttga ggaggagtcc cccaccccg   | 540  |
| cccgtctgtc cgcgacagga agctagcgat ccggaggact tagaatacaa tcgtagtgtg  | 600  |
| ggtaaacaatg gagggcaagc gcctgcaaag ggaagtaaga agattcccag tccttggtga | 660  |
| aatccatttg caaacagagg aagctgccgc gggtcgcagt cggtagggggg aagccctgaa | 720  |
| ccccacgctg cacggctggg ctggccagggt gcggccacgc ccccatcgcg gcggctggta | 780  |
| ggagtgaatc agaccgtcag tattggtaaa gaagtctgcg gcagggcagg gagggggaag  | 840  |
| agtagtcagt cgctcgctca ctgctcgct cgcacagaca ctgctgcagt gacactcggc   | 900  |
| cctccagtgt cgcgagagac caagagcagc gcgcagcacc tgtccgcccg gagcgagccc  | 960  |
| ggcccgcggc cgtagaaaag gagggaccgc cgaggtgcgc gtcagtactg ctcagcccgg  | 1020 |
| cagggacgcg ggaggatgtg gactgggtgg ac                                | 1052 |

<210> 7  
<211> 2008  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Synthetic polynucleotide-Blimp1 promoter 2kb

<400> 7  
gtgggtgctga ctcagcatcg gttaataaac cctctgcagg aggctggatt tcttttgttt 60  
aattatcact tggacctttc tgagaactct taagaattgt tcattcgggt ttttttgttt 120  
tgttttgggt tggttttttt gggttttttt tttttttttt tttttggttt ttggagacag 180  
ggtttctctg tatatagccc tggcacaaga gcaagctaac agcctgtttc ttcttgggtgc 240  
tagcgcccc tctggcagaa aatgaaataa cagggtggacc tacaaccccc cccccccccc 300  
ccagtgtatt ctactcttgt ccccggtata aatttgattg ttccgaacta cataaattgt 360  
agaaggattt tttagatgca catatcattt tctgtgatac cttccacaca cccctcccc 420  
ccaaaaaaat ttttctggga aagtttcttg aaaggaaaac agaagaacaa gcctgtcttt 480  
atgattgagt tgggcttttg ttttgctgtg tttcatttct tcctgtaaac aaataactcaa 540  
atgtccactt cattgtatga ctaagttggg atcattaggt tgggtctggg tgtgtgaatg 600  
tgggtgtgga tctggatgtg ggtgggtgtg tatgccccgt gtgttttagaa tactagaaaa 660  
gataccacat cgtaaaacttt tgggagagat gattttttaa aatgggggtg ggggtgaggg 720  
gaacctgcga tgaggcaagc aagataaggg gaagacttga gtttctgtga tctaaaaagt 780  
cgctgtgatg ggatgctggc tataaatggg cccttagcag cattgtttct gtgaattgga 840  
ggatccctgc tgaaggcaaa agaccattga aggaagtacc gcatctgggt tgttttgtaa 900  
tgagaagcag gaatgcaagg tccacgctct taataataaa caaacaggac attgtatgcc 960  
atcatcacag gatgtccttc cttctccaga agacagactg gggctgaagg aaaagccggc 1020  
caggctcaga acgagcccca ctaattactg cctccaacag ctttccactc actgccccca 1080  
gccaacatc ccctttttta ctgggaagca ttcctactct ccattgtacg cacacgctcg 1140  
gaagcctggc tgtgggtttg ggcatgagag gcagggacaa caaaaccagt atatatgatt 1200

|                                                                     |      |
|---------------------------------------------------------------------|------|
| ataacttttt cctgtttccc tattttccaaa tgggtcgaaag gaggaagtta ggtctaccta | 1260 |
| agctgaatgt attcagtttag caggagaaaat gaaatcctat acgtttaata ctagaggaga | 1320 |
| accgccttag aatattttatt tcattggcaa tgactccagg actacacagc gaaattgtat  | 1380 |
| tgcatgtgct gccaaaatac tttagctctt tccttcgaag tacgtcggat cctgtaattg   | 1440 |
| agacaccgag tttaggtgac tagggttttc ttttgaggag gagtccccca ccccgccccg   | 1500 |
| ctctgccgcg acaggaagct agcgatccgg aggacttaga atacaatcgt agtgtgggta   | 1560 |
| aacatggagg gcaagcgcct gcaaagggaa gtaagaagat tcccagtcct tgttgaaatc   | 1620 |
| catttgcaaa cagaggaagc tgccgcgggt cgagtcggt ggggggaagc cctgaacccc    | 1680 |
| acgctgcacg gctgggctgg ccaggtgcgg ccacgcccc atcgcggcgg ctggtaggag    | 1740 |
| tgaatcagac cgtcagtatt ggtaaagaag tctgcggcag ggcagggagg gggaagagta   | 1800 |
| gtcagtcgct cgctcactcg ctgcctcgca cagacactgc tgcagtgaca ctcgccctc    | 1860 |
| cagtgtcgcg gagacgcaag agcagcgcgc agcacctgtc cgcccggagc gagcccggcc   | 1920 |
| cgcgggccgta gaaaaggagg gaccgccgag gtgcgcgtca gtactgctca gcccggcagg  | 1980 |
| gacgcgggag gatgtggact gggtggac                                      | 2008 |

<210> 8

<211> 8388

<212> DNA

<213> Artificial Sequence

<220>

<223> Polynucleotide sequence of a modified C9ORF72 locus from lacZ through the 3' loxP site

<400> 8

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| ggtaccgatt taaatgatcc agtggtcctg cagaggagag attgggagaa tcccgggtgtg | 60  |
| acacagctga acagactagc cgcccaccct ccctttgctt cttggagaaa cagtgaggaa  | 120 |
| gctaggacag acagaccaag ccagcaactc agatctttga acggggagtg gagatttgcc  | 180 |
| tggtttccgg caccagaagc ggtgccggaa agctggctgg agtgcgatct tcctgaggcc  | 240 |
| gatactgtcg tcgtccccctc aaactggcag atgcacggtt acgatgcgcc catctacacc | 300 |
| aacgtgacct atcccattac ggtcaatccg ccgtttgttc ccacggagaa tccgacgggt  | 360 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| tgttactcgc tcacatttaa tgttgatgaa agctggctac aggaaggcca gacgcgaatt  | 420  |
| atTTTTgatg gCGTtaactc ggcgtttcat ctgtggtgca acgggcgctg ggtcggttac  | 480  |
| ggccaggaca gtcgTTtgcc gtctgaattt gacctgagcg cattttttacg cgccggagaa | 540  |
| aaccgcctcg cggTgatggt gctgcgctgg agtgacggca gttatctgga agatcaggat  | 600  |
| atgtggcgga tgagcggcatt tttccgtgac gtctcgTTgc tgcataaacc gactacacaa | 660  |
| atcagcgatt tccatgtTgc cactcgcttt aatgatgatt tcagccgcgc tgtactggag  | 720  |
| gctgaagtTc agatgtgcgg cgagttgcgt gactacctac gggtaacagt ttctttatgg  | 780  |
| cagggtgaaa cgcaggtcgc cagcggcacc gcgcTTTTcg gcggtgaaat tatcgatgag  | 840  |
| cgtggtggTt atgccgatcg cgtcacacta cgtctgaacg tcgaaaaccc gaaactgtgg  | 900  |
| agcgccgaaa tcccgaatct ctatcgTgcg gtggttgaac tgcacaccgc cgacggcacg  | 960  |
| ctgattgaag cagaagcctg cgatgtcggT ttcgcgagg tgccgattga aaatggtctg   | 1020 |
| ctgctgctga acggcaagcc gttgctgatt cgaggcgTta accgtcacga gcatcatcct  | 1080 |
| ctgcatggTc aggtcatgga tgagcagacg atggTgcagg atatcctgct gatgaagcag  | 1140 |
| aacaacttta acgccgtgcg ctgttcgcat tatccgaacc atccgctgtg gtacacgctg  | 1200 |
| tgcgaccgct acggcctgta tgtggtggat gaagccaata ttgaaacca cggcattggtg  | 1260 |
| ccaatgaatc gtctgaccga tgatccgcgc tggctaccgg cgatgagcga acgcgtaacg  | 1320 |
| cgaatggtgc agcgcgatcg taatcacccg agtgtgatca tctggTcgct ggggaatgaa  | 1380 |
| tcaggccacg gcgctaata cgcgcgctg tatcgctgga tcaaattctgt cgatccttcc   | 1440 |
| cgcgggtgc agtatgaagg cggcggagcc gacaccacgg ccaccgatat tatttgcccg   | 1500 |
| atgtacgcgc gcgtggatga agaccagccc ttcccggtg tgccgaaatg gtccatcaaa   | 1560 |
| aaatggcttt cgctacctgg agagacgcgc ccgctgatcc ttTgcgaata cggccacgcg  | 1620 |
| atgggtaaca gtcttggcgg tttcgctaaa tactggcagg cgTttcgtca gtatccccgt  | 1680 |
| ttacagggcg gcttcgtctg ggactgggtg gatcagtcgc tgattaaata tgatgaaaac  | 1740 |
| ggcaaccctg ggtcggctta cggcggTgat tttggcgata cgccgaacga tcgccagtTc  | 1800 |
| tgtatgaacg gtctggtctt tgccgaccgc acgccgcatc cagcgctgac ggaagcaaaa  | 1860 |

|                                                                     |      |
|---------------------------------------------------------------------|------|
| caccagcagc agtttttcca gttccgttta tccgggcaaa ccatcgaagt gaccagcgaa   | 1920 |
| tacctgttcc gtcatagcga taacgagctc ctgcactgga tggtagcgct ggatggtaag   | 1980 |
| ccgctggcaa gcggtgaagt gcctctggat gtcgctccac aaggtaaaca gttgattgaa   | 2040 |
| ctgcctgaac taccgcagcc ggagagcgcc gggcaactct ggctcacagt acgcgtagtg   | 2100 |
| caaccgaacg cgaccgcatg gtcagaagcc gggcacatca gcgcctggca gcagtggcgt   | 2160 |
| ctggcggaaa acctcagtgt gacgctcccc gccgcgtccc acgccatccc gcatctgacc   | 2220 |
| accagcgaiaa tggattttttg catcgagctg ggtaataagc gttggcaatt taaccgccag | 2280 |
| tcaggctttc tttcacagat gtggattggc gataaaaaac aactgctgac gccgctgcgc   | 2340 |
| gatcagttca cccgtgcacc gctggataac gacattggcg taagtgaagc gacccgcatt   | 2400 |
| gaccctaacg cctgggtcga acgctggaag gcggcgggcc attaccaggc cgaagcagcg   | 2460 |
| ttgttgcaat gcacggcaga tacacttgct gatgcggtgc tgattacgac cgctcacgcg   | 2520 |
| tggcagcatc aggggaaaaac cttatttatc agccggaaaa cctaccggat tgatggtagt  | 2580 |
| ggtcaaatgg cgattaccgt tgatgttgaa gtggcgagcg atacaccgca tccggcgcgg   | 2640 |
| attggcctga actgccagct ggcgcaggta gcagagcggg taaactggct cggattaggg   | 2700 |
| ccgcaagaaa actatcccga ccgccttact gccgcctgtt ttgaccgctg ggatctgcca   | 2760 |
| ttgtcagaca tgtatacccc gtacgtcttc ccgagcgaiaa acggtctgcg ctgcgggacg  | 2820 |
| cgcaattga attatggccc acaccagtgg cgcggcgact tccagttcaa catcagccgc    | 2880 |
| tacagtcaac agcaactgat ggaaaccagc catcgccatc tgctgcacgc ggaagaaggc   | 2940 |
| acatggctga atatcgacgg tttccatatg gggattgggt gcgacgactc ctggagccccg  | 3000 |
| tcagtatcgg cggaattcca gctgagcgcc ggtcgctacc attaccagtt ggtctggtgt   | 3060 |
| caaaaataat aataaccggg caggggggat ctaagctcta gataagtaat gatcataatc   | 3120 |
| agccatatca catctgtaga ggttttactt gctttaaaaa acctcccaca cctccccctg   | 3180 |
| aacctgaaac ataaaaatgaa tgcaattgtt gttgttaact tgtttattgc agcttataat  | 3240 |
| ggttacaaat aaagcaatag catcacaaat ttcacaaata aagcattttt ttactgcat    | 3300 |
| tctagtgtgt gtttgtccaa actcatcaat gtatcttatc atgtctggat cccccggcta   | 3360 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| gagtttaaac actagaacta gtggatcccc gggctcgata actataacgg tcctaaggta  | 3420 |
| gcgactcgag ataacttcgt ataatgtatg ctatacgaag ttatatgcat ggccctccgcg | 3480 |
| ccgggtttttg gcgcctcccc cgggcgcccc cctcctcacg gcgagcgctg ccacgtcaga | 3540 |
| cgaagggcgc agcgagcgtc ctgatccttc cgcccgagc ctcaggacag cggcccgcgtg  | 3600 |
| ctcataagac tcggccttag aaccccagta tcagcagaag gacatttttag gacgggactt | 3660 |
| gggtgactct agggcactgg ttttctttcc agagagcgga acaggcgagg aaaagtagtc  | 3720 |
| ccttctcggc gattctgcgg agggatctcc gtggggcggt gaacgccgat gattatataa  | 3780 |
| ggacgcgccg ggtgtggcac agctagtctc gtcgcagccg ggatttgggt cgcggttctt  | 3840 |
| gtttgtggat cgctgtgac gtcacttgggt gagtagcggg ctgctgggct ggccggggct  | 3900 |
| ttcgtggccg ccgggcccgt cgggtgggacg gaagcgtgtg gagagaccgc caagggctgt | 3960 |
| agtctgggtc cgcgagcaag gttgccctga actggggggt ggggggagcg cagcaaaatg  | 4020 |
| gcggctgttc ccgagtcttg aatggaagac gcttgtgagg cgggctgtga ggtcgttgaa  | 4080 |
| acaaggtggg gggcatgggt ggccggcaaga acccaaggtc ttgaggcctt cgctaatacg | 4140 |
| ggaaagctct tattcgggtg agatgggctg gggcaccatc tggggaccct gacgtgaagt  | 4200 |
| ttgtcactga ctggagaact cggtttgtcg tctgttgcgg gggcggcagt tatggcggtg  | 4260 |
| ccgttgggca gtgcacccgt acctttggga gcgcgcgcc tcgtcgtgtc gtgacgtcac   | 4320 |
| ccgttctgtt ggcttataat gcagggtggg gccacctgcc ggtaggtgtg cggtaggctt  | 4380 |
| ttctccgtcg caggacgcag ggttcgggcc tagggtaggc tctcctgaat cgacaggcgc  | 4440 |
| cggacctctg gtgaggggag ggataagtga ggcgtcagtt tctttggtcg gttttatgta  | 4500 |
| cctatcttct taagtagctg aagctccggt tttgaactat gcgctcgggg ttggcgagt   | 4560 |
| tgttttgtga agtttttttag gcaccttttg aaatgtaatc atttgggtca atatgtaatt | 4620 |
| ttcagtgtta gactagtaaa ttgtccgcta aattctggcc gtttttggct tttttgtag   | 4680 |
| acgtgttgac aattaatcat cggcatagta tatcggcata gtataatacg acaaggtag   | 4740 |
| gaactaaacc atgggatcgg ccattgaaca agatggattg cacgcagggt ctccggccgc  | 4800 |
| ttgggtggag aggctattcg gctatgactg ggcacaacag acaatcggct gctctgatgc  | 4860 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| cgccgtgttc cggctgtcag cgcaggggcg cccggttctt tttgtcaaga ccgacctgtc  | 4920 |
| cggtgccctg aatgaactgc aggacgaggc agcgcggtta tcgtggctgg ccacgacggg  | 4980 |
| cgttccttgc gcagctgtgc tcgacgttgt cactgaagcg ggaagggact ggctgctatt  | 5040 |
| gggcgaagtg ccggggcagg atctcctgtc atctcacctt gctcctgccg agaaagtatc  | 5100 |
| catcatggct gatgcaatgc ggcggctgca tacgcttgat ccggctacct gccattcga   | 5160 |
| ccaccaagcg aaacatcgca tcgagcgagc acgtactcgg atggaagccg gtcttgtcga  | 5220 |
| tcaggatgat ctggacgaag agcatcaggg gctcgcgcca gccgaactgt tcgccaggct  | 5280 |
| caaggcgcgc atgcccgcgc gcgatgatct cgtcgtgacc catggcgatg cctgcttgcc  | 5340 |
| gaatatcatg gtggaaaatg gccgcttttc tggattcatc gactgtggcc ggctgggtgt  | 5400 |
| ggcggaccgc tatcaggaca tagcgttggc taccctgat attgctgaag agcttggcgg   | 5460 |
| cgaatgggct gaccgcttcc tcgtgcttta cggatatgcc gctcccgatt cgcagcgcat  | 5520 |
| cgccttctat cgccttcttg acgagttctt ctgaggggat ccgctgtaag tctgcagaaa  | 5580 |
| ttgatgatct attaaacaat aaagatgtcc actaaaatgg aagtttttcc tgtcatactt  | 5640 |
| tgtaagaag ggtgagaaca gagtacctac attttgaatg gaaggattgg agctacgggg   | 5700 |
| gtgggggtgg ggtgggatta gataaatgcc tgctctttac tgaaggctct ttactattgc  | 5760 |
| tttatgataa tgtttcatag ttggatatca taatttaaac aagcaaaacc aaattaaggg  | 5820 |
| ccagctcatt cctcccactc atgatctata gatctataga tctctcgtgg gatcattggt  | 5880 |
| tttctcttga ttcccacttt gtggttctaa gtactgtgggt ttccaaatgt gtcagtttca | 5940 |
| tagcctgaag aacgagatca gcagcctctg ttccacatac acttcattct cagtattggt  | 6000 |
| ttgccaagtt ctaattccat cagacctcga cctgcagccc ctagtcgagc gccagtagca  | 6060 |
| gcaccacgt ccaccttctg tctagtaatg tccaacacct ccctcagtcc aaacactgct   | 6120 |
| ctgcatccat gtggctccca ttatatacctg aagcacttga tggggcctca atgttttact | 6180 |
| agagcccacc cccctgcaac tctgagacct tctggatttg tctgtcagtg cctcactggg  | 6240 |
| gcgttggata atttcttaaa aggtcaagtt ccctcagcag cattctctga gcagtctgaa  | 6300 |
| gatgtgtgct tttcacagtt caaatccatg tggctgtttc acccacctgc ctggccttgg  | 6360 |

|                                                                     |      |
|---------------------------------------------------------------------|------|
| gttatctatc aggacctagc ctagaagcag gtgtgtggca cttaacacct aagctgagtg   | 6420 |
| actaactgaa cactcaagtg gatgccatct ttgtcacttc ttgactgtga cacaagcaac   | 6480 |
| tcctgatgcc aaagccctgc ccacccctct catgcccata tttggacatg gtacagggtcc  | 6540 |
| tcactggcca tggctctgtga ggtcctggtc ctctttgact tcataattcc taggggccac  | 6600 |
| tagtatctat aagaggaaga gggtgctggc tcccaggcca cagcccacaa aattccacct   | 6660 |
| gctcacagggt tggctggctc gacccagggtg gtgtcccctg ctctgagcca gctcccggcc | 6720 |
| aagccagcac catgggaacc cccaagaaga agaggaagggt gcgtaccgat ttaaattcca  | 6780 |
| atttactgac cgtacaccaa aatttgcctg cattaccggt cgatgcaacg agtgatgagg   | 6840 |
| ttcgcaagaa cctgatggac atgttcaggg atcgccaggc gttttctgag catacctgga   | 6900 |
| aaatgcttct gtccgtttgc cggtcgtggg cggcatgggt caagttgaat aaccggaaat   | 6960 |
| ggtttccgc agaacctgaa gatgttcgcg attatcttct atatcttcag gcgcgcggtc    | 7020 |
| tggcagtaaa aactatccag caacatttgg gccagctaaa catgcttcat cgtcgggtccg  | 7080 |
| ggctgccacg accaagtgac agcaatgctg tttcactggt tatgcggcgg atccgaaaag   | 7140 |
| aaaacgttga tgccggtgaa cgtgcaaaac aggctctagc gttcgaacgc actgatttcg   | 7200 |
| accaggttcg ttcactcatg gaaaatagcg atcgctgcca ggatatacgt aatctggcat   | 7260 |
| ttctggggat tgcttataac accctgttac gtatagccga aattgccagg atcagggtta   | 7320 |
| aagatatctc acgtactgac ggtgggagaa tgттаатсса tattggcaga acgaaaacgc   | 7380 |
| tggttagcac cgcagggtga gagaaggcac ttagcctggg ggtaactaaa ctggtcgagc   | 7440 |
| gatggatttc cgtctctggt gtagctgatg atccgaataa ctacctgttt tgccgggtca   | 7500 |
| gaaaaaatgg tgttgccgcg ccatctgcca ccagccagct atcaactcg gccttgaag     | 7560 |
| ggatttttga agcaactcat cgattgattt acggcgctaa ggtaaataa aaatttttaa    | 7620 |
| gtgtataatg tgttaaacta ctgattctaa ttgtttgtgt attttaggat gactctggtc   | 7680 |
| agagatacct ggcttggctt ggacacagtg cccgtgtcgg agccgcgcga gatatggccc   | 7740 |
| gcgctggagt ttcaataccg gagatcatgc aagctgggtg ctggaccaat gtaaatattg   | 7800 |
| tcatgaacta tatccgtaac ctggatagtg aaacaggggc aatggtgcgc ctgctggaag   | 7860 |

|                                                                   |      |
|-------------------------------------------------------------------|------|
| atggcgattg atctagataa gtaatgatca taatcagcca tatcacatct gtagaggttt | 7920 |
| tacttgcttt aaaaaacctc ccacacctcc ccctgaacct gaaacataaa atgaatgcaa | 7980 |
| ttgttgttgt taaacctgcc ctagttgcgg ccaattccag ctgagcgtga gctcaccatt | 8040 |
| accagttggt ctggtgtcaa aaataataat aaccgggcag gggggatcta agctctagat | 8100 |
| aagtaatgat cataatcagc catatcacat ctgtagaggt ttacttgct ttaaaaaacc  | 8160 |
| tcccacacct cccctgaac ctgaaacata aaatgaatgc aattgttgtt gttaacttgt  | 8220 |
| ttattgcagc ttataatggt tacaataaaa gcaatagcat cacaaatttc acaaataaag | 8280 |
| catttttttc actgcattct agttgtggtt tgtccaaact catcaatgta tcttatcatg | 8340 |
| tctggatcct cgacataact tcgtataatg tatgctatac gaagttat              | 8388 |

<210> 9

<211> 17210

<212> DNA

<213> Artificial Sequence

<220>

<223> Polynucleotide sequence of a modified C9ORF72 locus from Exon 1a through 3' UTR

<400> 9

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| cgttttagt gacagccatc ccaattgccc tttccttcta ggtggaaagt ggtgtctaga   | 60  |
| cagtccaggg aggggtgtcg agggaggtgc gttttggtt cctcagctcg caacttaact   | 120 |
| ccacaacggt gaccaaggac aaaagaagga aacaagactg cagagatccg caccggggag  | 180 |
| ccctgcagat tctgggtctg ctgtggactg ggggcgggac tgcgactggg cgggcctggg  | 240 |
| ggcgtgtccg gggcggggcg gtcccggggc ggggcccgga gcgggctgcg gttgcgggtcc | 300 |
| ctgcgccggc ggtgaaggcg cagcagcggc gagtgggtga gtgagacgcg cgggcggagg  | 360 |
| ggggctgctg ccacggtcgg ctgcggggcc ggccggctcc gggtaggcgc ggggtttttt  | 420 |
| tctccttcga ggtgaactcc tccctgtccc ccgggcgaaa gagcccttgg ccttgcagga  | 480 |
| gttgccgggg ccgcggcggt gcggagggga tggggatggg cctcatcttt gctgtccgcc  | 540 |
| cgcgctcccc gatcccgacc cggagcgtct cccgggccct tgagggaacc ctccgggagt  | 600 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| acggcgagcg cggccccac cgccacaagc ctgggccccca ggggcctggc ccggcgacag  | 660  |
| ctggtgggtc ctgcgacca gtcaggtctc ccgaggggtcc ccgcccggga ggagaaagcg  | 720  |
| ccggtgggat ggagtaagga cggacagaac aacacgcagg caggatttcg cagaagtttg  | 780  |
| caaggagtgc ggatgcccac ttacatgggc tgctactctt accaggttgt tccccagttc  | 840  |
| tgtgggacgt gacctggttg cctcacagct ccgcggttgt acagacttat taaaggaagt  | 900  |
| gaccattgtg acttgggcat cacttgactg atggtaatca gttgcagaga gagaagtgca  | 960  |
| ctgattaagt ctgtccacac agggctctgtc tggccaggag tgcatttgcc tgggagggat | 1020 |
| tggttgcgct ttctggtgtg gggactatta ggctcttgta gagttttgtc ccggcagatg  | 1080 |
| gataaatttc ttgttacct gttcccgttc gtcaccagtt gagaaaaacg ggtacacagt   | 1140 |
| ctgtctcagt agtactttta ctttatatta agggcccaaa agggactgga aaatacttta  | 1200 |
| agatagaatc gttagtccac ttggaaaact taaaatatga gagagagagg ggggggggag  | 1260 |
| agagagagag agagagagag agaaaggaag gaagaaggag gaagaggagg aggaaagaga  | 1320 |
| ttgagattat gttaataata tggaatcaga atatttgaaa tatagtaagc gtccccctcag | 1380 |
| ttaaagagga cattccagga ggccccagt atagcctgaa atctcaggaa acgcctacat   | 1440 |
| acacccatcg tgtggatata ggtgttttcc cttcattaca tttcatacac agatgttaaa  | 1500 |
| gtttagaaaag taggcacaat aagagattac aaataactga taataaagtc gagccattgc | 1560 |
| agctgctctg taaaagtcct gtgaatgtga tcgctttgtg tttcaaagta acttactgta  | 1620 |
| cttcacccct gttaagcaaa acaagattca cctgaacgca ggcaccttgg taccttggca  | 1680 |
| gacaccagat ctgataacca agaggatgga gaagtagtgg cagacagtgt ggagagcatg  | 1740 |
| aatatgctag acaaaagggt gaatcataac ctaggagcag aaagcaggta tttcatcatc  | 1800 |
| ctccacagta aaaacctatg tcacgtaaaa aacctacaag tagtttttct tttactcttt  | 1860 |
| ttgaatgaaa gcttgctaca ggcaactgaaa gttaaaataa tctgtggatc aggaggaaca | 1920 |
| ggggttttct gtctgagtca ctgctgacta gcacctcagt gaccattggc actgtgggaa  | 1980 |
| accccagagt cagttggaaa cttcgaaact aaaggtgacg gtgttcttat ttcatagaac  | 2040 |
| acaaaaaata agaggggtta cagcctgcgc tgcagactgg acattcaaca agcatttaaa  | 2100 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| tttctgggag acaaagttaa atataacttt aaaagttggt aaaatactct gtttggctat  | 2160 |
| gttggccatc caatgtttgc ttttagaaaa tgactgaatg gataaaacgt ctatcttttg  | 2220 |
| agcctgccct agacccccat gttgagtga tactgtccaa gtgttaggtt agccggcctg   | 2280 |
| agaaacttgg atctaggcaa gatggcacag tcctggtgtc atgagtatgc atgtgagttt  | 2340 |
| tggctgaaat tgaacatttg tagagaatga caaaggctgg tctggcaagt agtcactgt   | 2400 |
| ctttacagtg gtcttggtta gttcctgttt ggctgagagg gctggttgat ggctgtcctg  | 2460 |
| cccctcttcc cacaagtgga agccttatgg tataattctt gatcacagta gcagtaggca  | 2520 |
| aatgaacttc ctcaaagcag cctggaaagc tgatTTTTTTT ttctttcttt ctcttttttt | 2580 |
| tttttttttc acaaggtaa agaaaaaaca aagggttca aatgtgccag tctgctaaca    | 2640 |
| gtgttaacat gtttattaac ataaataaac tttattagtt tttggaagta ttggttaagc  | 2700 |
| cctcgtgacc cctgaactcg gtttatagag tgatgagtcg tagcctcact ctggtttgga  | 2760 |
| ctctggcttc tctcagaaga ctctgtggct aatgttaacc ttctgaagta gccagaaaac  | 2820 |
| atataagcaa aagtctgtga ggttgaaatg aatTTTTTg ccacatttgt atatgggttc   | 2880 |
| ccaccaatgc taacttcagg tgttagtaat atcagactca cagcttccct gattacactt  | 2940 |
| cgctataaga ctttattttt taggtcatag gaatttcccc tttttcatga ttcttaaadc  | 3000 |
| atgaaataac atagtctaaa aatacggat tcctgaaata aacaatttct aagttttaag   | 3060 |
| ctgcgtgcta ttctgaacag tctgatgccc tctttagct tttacttgtt cctaccccgg   | 3120 |
| gcatgggtga ttctttgtc caaacatctg tctgttgat ccacactgga ttgcaccacc    | 3180 |
| tgcgtgctag tcagtcactc agacatttta gttataaggt agcttatatt tactccttat  | 3240 |
| tttatttaat aatggcctca tagcaaggcg gtaatgatac tggtaatttg ggtttgctta  | 3300 |
| agaggagcca tgaagtagtt ttaaataaaa aggtgaaaat tcccactata gtttggaggg  | 3360 |
| ggaggctata ctggtactac tacgattcac ggtaagacta aatcttctgt gaaattatga  | 3420 |
| aggagaaaaa gttacactgg tctggtcttg ctgttggtt aattttatag ttataaccac   | 3480 |
| tgtacatgat aaataaccct aaaacaatga atttgtaggt ggatggcata atctgaaaac  | 3540 |
| catgttctga gcagttgatg gcagcaggct gtgctggaag tgttaggcat atttatagat  | 3600 |

|                                                                     |      |
|---------------------------------------------------------------------|------|
| ttcagcccaa gttctgaaga ggctggagag atggctcagt ggtaaagagt gcttgctatt   | 3660 |
| gcagaggacc taggttcctc tacaggcacc aggcaagcgt gggacacact gagatacata   | 3720 |
| cagacaaaac ataaaaattaa ataaattgtg cataataata ctagtaatata atgagtaaaa | 3780 |
| taaggataaa tacacatcat aattaaataa ataaattgta aagttcccta gaagtgaggg   | 3840 |
| tcaccaagcc attcacaaga tggctgcgct gatgcaggga tatatgtgaa ctagaaaaag   | 3900 |
| gtcaaaactta acagagaagt tccaaggcat gctactgcag gcttggctag catgcttgac  | 3960 |
| ctgcagaaat gctgacggcc actgggagggt tttcacaaat gaggaattag aagaactttt  | 4020 |
| tttactaatc tccagaaaaa aaaaaggga gaagaaactg aagcagcctg tgatgtggac    | 4080 |
| cagaaacgca gtgacagtaa catgtgtgac attgcaaagg catgaaagga cagagctgtg   | 4140 |
| gaatacagac ctcagggtgga gctcagcata gagtcattcg gggattatgc ctgctgcagc  | 4200 |
| aacaaaagga tgagctcaaa agagacaccg acttctgaat gcagtgggtg tttgttttgt   | 4260 |
| tttgtttcaa atgaattggg cagaaaactt tccagctgtg gaagcttctg aaccgtccct   | 4320 |
| tgctgctgac atctaagcgt ccgctgtgtc ccagctcagt gatctagggt cttccaaaca   | 4380 |
| gatggtccgg tgctgagcac tttgaatctc aatcctgagt ttctaccacg cttttggcca   | 4440 |
| tttaattccc agataaaaaga cacatacaac ctttatatatt ataataaacc ttagtcagca | 4500 |
| caagagctga gcaaatatct gtcctctatg ctattatatc tattaccag ccaataaccc    | 4560 |
| cattctataa tttgctgtgc ttcattctggg ctgctcttaa cttcagtcag ccagcccacg  | 4620 |
| tggccattat ttttaagattt ttttacccca tagtgtcttc tcactttact ttacattttt  | 4680 |
| ctctctctcc tcatggttct cctctgacct caagcctagg aaccctaaac cccacccatg   | 4740 |
| tctcttctgc ccatctattg gctgtaggca tctttattca ccaatcagga taacttggag   | 4800 |
| gcaagggttaa gtagtctcct gggcttaggt gctgtctctg ggagcaacca gtatttagca  | 4860 |
| tagcaaaaaga ccagacctcc acaatgatca ctctgacct cggggcagaa ggcacctact   | 4920 |
| agcctgtgcc actcacctca ctttgttgaa tcacatctta tcctgtagtg tgtatcactg   | 4980 |
| cctgttatca caggaaaaag tgagtcccat caaataagat gtttcagaaa gagaccatgt   | 5040 |
| tcatataatt atcattctgg taagctttta atggttatat tttgttatta atctctttgt   | 5100 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| tcctattttg caaattatac cttacagtaa atatatatgc atccaatggg gtctttgaat  | 5160 |
| tcctccccgg ggagtaggag gactctttga ggatgggctg catttaaagc taaacaacgc  | 5220 |
| aacatgacct ttagtcctta tagatagcct agagatgaga ctaaataaaa gaaatggtat  | 5280 |
| ataatgcttt aagtttccca atcagcttaa aagcttttcc tataaatcct taagattatg  | 5340 |
| ctctggggct caatactgct tcaagaaggg cttttctttt gtatttagaa ttattcacct  | 5400 |
| ttttaaacia aaggagaaaa tggaatagaa atatgtttgc aacataattt tatgactatg  | 5460 |
| tgttttatttc gcgtgttctg tgggcctgca gtttgctgct gttaatgagg acaacagtgg | 5520 |
| caccaatata gtttccactc agattacatt ctctgttccc tttctgaaag ctgccctctc  | 5580 |
| cactgggccc aaaagagtca gtatcttaaa caagctgtac aacttagata accatggctt  | 5640 |
| cttcagacta gttaattgac atatattaaa aagtaaatag taccaaagtg aatttctgaa  | 5700 |
| attaaaaatg aacatttaaa aactctaggt aaactattcc ttagagttaa gtgttttgcc  | 5760 |
| aagttctgta atcataatat gatagaaacg ctactcagc attctaaata tagaagttac   | 5820 |
| tccttcgcat gacactctaa ttcttgataa ggtggagaaa gagagagaga gagggggaga  | 5880 |
| gacagaaaat atgggtggttc aaggaccatt tgagggaatt agttatgttc ttccgtcctc | 5940 |
| tgtggatctt aggggttgaa tacagtcatt gagctcggtg gatggctgtc ctgttgaaag  | 6000 |
| gtctgcccag cagagcaaata agactttttt atttacctgg acatccgttt gtgactaatc | 6060 |
| taatgttcac tcccaaagta atcacacaga cagagaggta gcttccttca gtactcttac  | 6120 |
| cttacctgaa tcctaccatt ttgttatttt ttttccactt taaatctttg attatgtgtt  | 6180 |
| tttaattaga aaatttgcac acaaatttcc atacagtatg tagaattgac tgtgtttgaa  | 6240 |
| tgggtgaaga tccacatgtg taaccctagc tctggactgg ctctgagctt gtttgctctt  | 6300 |
| ctctttttgtg ttctgagtaa ctgaaactct ttcatcttag cagcttagta tgcgcccttc | 6360 |
| acattgctgt gctgcctgct gcactaacat tactcctttg cttatgttcc ccttcctgat  | 6420 |
| tcagtgtcat ttttaagcagt agtactggac ctcatgacct tagccggagc tcatgaggt  | 6480 |
| gacagggtg aggctctgct gctgtctttt gagcttacct ctttttaatg ttttatggta   | 6540 |
| tttctgctgc caggtttggg ggttttggtt tgttttgttt tttgtttttt gtttttttta  | 6600 |

|                                                                       |      |
|-----------------------------------------------------------------------|------|
| atthttctagg aacacctaga aaacacaaaac taggaaactt aaaagagcag cgtcttggttc  | 6660 |
| cctgcgttct agaaaagtcca agcctaattgc cagtgtcatg gttgtcagga acatgagcct   | 6720 |
| ctgaaggctt cttgggaaaac ctttcttggtc tcaacacctc tgggtggcaag cagtagtcca  | 6780 |
| tgggtactctc tctgtccacg gtcagcatcc cagtccttgc cttttatctt tgtgcagccg    | 6840 |
| accagctttg ctttagtctg tctccttctc aggtctcctt ccccgctcct ctttaagcaca    | 6900 |
| gcagtcattg gattagagcc catccttccc tcggatggcc catttgacct aatthttacgt    | 6960 |
| atthtgtaact aagggtcccat ttactttacac agggccctcc ctttcctggtt ttgttcttta | 7020 |
| gctgaaatgg tttggagacc aaatatccaa tcattacaat tgtgcacaag ctatgttcat     | 7080 |
| ttggaggtaa taaaggctca ttcttttgctt ctattgggtat gtgacattht tctaagtcac   | 7140 |
| ttgggggtttg atagatatct ttaaatggct gaacctgatc actgttcttt tgtatgtccc    | 7200 |
| tgttttagcta ttgcaagcgt tcggataatg tgagacctgg aatgcagtga gacctgggat    | 7260 |
| gcagggatgg gtaccgattt aaatgatcca gtggtcctgc agaggagaga ttgggagaat     | 7320 |
| cccggtgtga cacagctgaa cagactagcc gcccaccctc cttttgcttc ttggagaaaac    | 7380 |
| agtgagggaag ctaggacaga cagaccaagc cagcaactca gatctttgaa cggggagtgg    | 7440 |
| agatthtgcct ggthttccggc accagaagcg gtgccggaaa gctggctgga gtgcgatctt   | 7500 |
| cctgaggccg atactgtcgt cgtccctca aactggcaga tgcacggtta cgatgcgccc      | 7560 |
| atctacacca acgtgacctt tcccattacg gtcaatccgc cgtttgttcc cacggagaat     | 7620 |
| ccgacgggtt gttactcgct cacatthtaat gttgatgaaa gctggctaca ggaaggccag    | 7680 |
| acgcaatta thtttgatgg cgttaactcg gcgtthcatc tgtggtgcaa cgggcgctgg      | 7740 |
| gtcggttacg gccaggacag tcgtttgccg tctgaatthg acctgagcg c atthttacgc    | 7800 |
| gccggagaaa accgcctcgc ggtgatgggtg ctgcgctgga gtgacggcag ttatctggaa    | 7860 |
| gatcaggata tgtggcggat gagcggcatt ttccgtgacg tctcgttgct gcataaaccg     | 7920 |
| actacacaaa tcagcgattt ccatgttgcc actcgttht atgatgattt cagccgcgct      | 7980 |
| gtactggagg ctgaagttca gatgtgcggc gagttgcgtg actacctacg ggtaacagtt     | 8040 |
| tcthttatggc aggggtgaaa gcaggtcgcc agcggcaccg cgcctthtcgg cgggtgaaatt  | 8100 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| atc gatgagc gtggtggtta tgccgatcgc gtcacactac gtctgaacgt cgaaaacccg | 8160 |
| aaactgtgga ggcgccgaaat cccgaatctc tatcgtgcgg tggttgaact gcacaccgcc | 8220 |
| gacggcacgc tgattgaagc agaagcctgc gatgtcggtt tccgcgaggt gcggattgaa  | 8280 |
| aatggtctgc tgctgctgaa cggcaagccg ttgctgattc gaggcgttaa ccgtcacgag  | 8340 |
| catcatcctc tgcatggtca ggtcatggat gagcagacga tggcgcagga tatcctgctg  | 8400 |
| atgaagcaga acaactttta cgccgtgcgc tgttcgcatt atccgaacca tccgctgtgg  | 8460 |
| tacacgctgt gcgaccgcta cggcctgtat gtggtggatg aagccaatat tgaaaccac   | 8520 |
| ggcatggtgc caatgaatcg tctgaccgat gatccgcgct ggctaccggc gatgagcgaa  | 8580 |
| cgcgtaacgc gaatggtgca gcgcgatcgt aatcacccga gtgtgatcat ctggtcgtg   | 8640 |
| gggaatgaat caggccacgg cgctaatac gacgcgctgt atcgttgat caaatctgtc    | 8700 |
| gatccttccc gcccggtgca gtatgaaggc ggcgagccg acaccacggc caccgatatt   | 8760 |
| atgtgcccga tgtacgcgcg cgtggatgaa gaccagccct tcccggctgt gccgaaatgg  | 8820 |
| tccatcaaaa aatggctttc gctacctgga gagacgcgcc cgctgaccta ttgcgaatac  | 8880 |
| gcccacgcga tgggtaacag tcttggcggc ttcgctaaat actggcaggc gtttcgtcag  | 8940 |
| tatccccgtt tacagggcgg cttcgtctgg gactgggtgg atcagtcgct gattaaatat  | 9000 |
| gatgaaaacg gcaacccgtg gtcggcttac ggcggtgatt ttggcgatac gccgaacgat  | 9060 |
| cgccagttct gtatgaacgg tctggtcttt gccgaccgca cgccgcatcc agcgctgacg  | 9120 |
| gaagcaaaac accagcagca gtttttccag ttccgtttat ccgggcaaac catcgaagtg  | 9180 |
| accagcgaat acctgttccg tcatagcgat aacgagctcc tgcactggat ggtggcgctg  | 9240 |
| gatggtaagc cgctggcaag cggatgaagt cctctggatg tcgctccaca aggtaaacag  | 9300 |
| ttgattgaac tgcctgaact accgcagccg gagagcgccg ggcaactctg gctcacagta  | 9360 |
| cgcgtagtgc aaccgaacgc gaccgcatgg tcagaagccg ggacacatcag cgcctggcag | 9420 |
| cagtggcgctc tggcggaaaa cctcagtgtg acgctccccg ccgcgtccca cgccatcccc | 9480 |
| catctgacca ccagcgaaat ggatttttgc atcgagctgg gtaataagcg ttggcaattt  | 9540 |
| aaccgccagt caggctttct ttcacagatg tggattggcg ataaaaaaca actgctgacg  | 9600 |

|                                                                    |       |
|--------------------------------------------------------------------|-------|
| ccgctgcgcg atcagttcac ccgtgcaccg ctggataacg acattggcgt aagtgaagcg  | 9660  |
| acccgcattg accctaacgc ctgggtcgaa cgctggaagg cggcgggcca ttaccaggcc  | 9720  |
| gaagcagcgt tgttgcagtg cacggcagat acacttgctg atgcggtgct gattacgacc  | 9780  |
| gctcacgcgt ggcagcatca ggggaaaacc ttatttatca gccggaaaac ctaccggatt  | 9840  |
| gatggtagt gtcaaatggc gattaccgtt gatgttgaag tggcgagcga tacaccgcat   | 9900  |
| ccggcgcgga ttggcctgaa ctgccagctg gcgcaggtag cagagcgggt aaactggctc  | 9960  |
| ggattagggc cgcaagaaaa ctatcccgac cgccttactg ccgcctgttt tgaccgctgg  | 10020 |
| gatctgccat tgtcagacat gtataccccg tacgtcttcc cgagcgaaaa cggctctgcgc | 10080 |
| tgcgggacgc gcgaattgaa ttatggccca caccagtggc gcggcgactt ccagttcaac  | 10140 |
| atcagccgct acagtcaaca gcaactgatg gaaaccagcc atcgccatct gctgcacgcg  | 10200 |
| gaagaaggca catggctgaa tatcgacggt ttccatatgg ggattggtgg cgacgactcc  | 10260 |
| tggagcccgt cagtatcggc ggaattccag ctgagcgccg gtcgctacca ttaccagttg  | 10320 |
| gtctggtgtc aaaaataata ataaccgggc aggggggatc taagctctag ataagtaatg  | 10380 |
| atcataatca gccatatcac atctgtagag gttttacttg ctttaaaaaa cctccccacac | 10440 |
| ctccccctga acctgaaaca taaaatgaat gcaattgttg ttgttaactt gtttattgca  | 10500 |
| gcttataatg gttacaaata aagcaatagc atcacaaatt tcacaaataa agcatttttt  | 10560 |
| tcactgcatt ctagttgtgg tttgtccaaa ctcatcaatg tatcttatca tgtctggatc  | 10620 |
| ccccggctag agtttaaaca ctagaactag tggatccccg ggctcgataa ctataacggt  | 10680 |
| cctaaggtag cgactcgaga taacttcgta taatgtatgc tatacgaagt tatatgcatg  | 10740 |
| gcctccgcgc cgggttttgg cgcctccgc gggcgcccc ctcctcacgg cgagcgctgc    | 10800 |
| cacgtcagac gaagggcgca gcgagcgtcc tgatccttcc gcccggacgc tcaggacagc  | 10860 |
| ggcccgtgc tcataagact cggccttaga accccagtat cagcagaagg acattttagg   | 10920 |
| acgggacttg ggtgactcta gggcactggt tttctttcca gagagcggaa caggcgagga  | 10980 |
| aaagtagtcc cttctcggcg attctgcgga gggatctccg tggggcggtg aacgccgatg  | 11040 |
| attatataag gacgcgccgg gtgtggcaca gctagttccg tcgcagccgg gatttgggtc  | 11100 |

|                                                                    |       |
|--------------------------------------------------------------------|-------|
| gcggttcttg tttgtggatc gctgtgatcg tcacttgggtg agtagcgggc tgctgggctg | 11160 |
| gccggggctt tcgtggccgc cgggccgctc ggtgggacgg aagcgtgtgg agagaccgcc  | 11220 |
| aagggctgta gtctgggtcc gcgagcaagg ttgccctgaa ctgggggttg gggggagcgc  | 11280 |
| agcaaaatgg cggctgttcc cgagtcttga atggaagacg cttgtgaggc gggctgtgag  | 11340 |
| gtcgttgaaa caaggtgggg ggcatgggtg gcggcaagaa cccaaggtct tgaggccttc  | 11400 |
| gctaatacgg gaaagctctt attcgggtga gatgggctgg ggcaccatct ggggaccctg  | 11460 |
| acgtgaagtt tgtcactgac tggagaactc ggtttgtcgt ctgttgccgg ggcggcagtt  | 11520 |
| atggcgggtc cgttgggcag tgcacccgta ctttggggag cgcgcgccct cgtcgtgtcg  | 11580 |
| tgacgtcacc cgttctgttg gcttataatg cagggtgggg ccacctgccg gtaggtgtgc  | 11640 |
| ggtaggcttt tctccgtcgc aggacgcagg gttcgggcct agggtaggct ctcctgaatc  | 11700 |
| gacaggcgcc ggacctctgg tgaggggagg gataagttag gcgtcagttt ctttggtcgg  | 11760 |
| ttttatgtac ctatcttctt aagtagctga agctccggtt ttgaactatg cgctcggggt  | 11820 |
| tggcgagtgt gttttgtgaa gttttttagg caccttttga aatgtaatca tttgggtcaa  | 11880 |
| tatgtaattt tcagtgttag actagtaaata tgtccgctaa attctggccg tttttggctt | 11940 |
| ttttgttaga cgtgttgaca attaatacgc ggcatagtat atcggcatag tataatacga  | 12000 |
| caaggtgagg aactaaacca tgggatcggc cattgaacaa gatggattgc acgcaggttc  | 12060 |
| tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga caatcggtg   | 12120 |
| ctctgatgcc gccgtgttcc ggctgtcagc gcaggggcgc ccggttcttt ttgtcaagac  | 12180 |
| cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat cgtggctggc  | 12240 |
| cacgacgggc gttccttgcg cagctgtgct cgacgttgct actgaagcgg gaagggactg  | 12300 |
| gctgctattg ggcaagtgc cggggcagga tctcctgtca tctcaccttg ctcctgccga   | 12360 |
| gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc cggctacctg  | 12420 |
| cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga tggaagccgg  | 12480 |
| tcttgtcgat caggatgatc tggacgaaga gcatcagggg ctcgcgccag ccgaactgtt  | 12540 |
| cgccaggctc aaggcgcgca tgcccagcgg cgatgatctc gtcgtgaccc atggcgatgc  | 12600 |

|                                                                    |       |
|--------------------------------------------------------------------|-------|
| ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg actgtggccg  | 12660 |
| gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata ttgctgaaga  | 12720 |
| gcttggcggc gaatgggctg accgcttcct cgtgctttac ggtatcgccg ctcccgattc  | 12780 |
| gcagcgcac gccttctatc gccttcttga cgagtcttc tgaggggatc cgctgtaagt    | 12840 |
| ctgcagaaat tgatgatcta ttaaacaata aagatgtcca ctaaaatgga agtttttcct  | 12900 |
| gtcatacttt gttaagaagg gtgagaacag agtacctaca ttttgaatgg aaggattgga  | 12960 |
| gctacggggg tgggggtggg gtgggattag ataaatgcct gctctttact gaaggctctt  | 13020 |
| tactattgct ttatgataat gtttcatagt tggatatcat aatttaaaca agcaaaacca  | 13080 |
| aattaagggc cagctcattc ctcccactca tgatctatag atctatagat ctctcgtagg  | 13140 |
| atcattgttt ttctcttgat tcccactttg tggttctaag tactgtgggt tccaaatgtg  | 13200 |
| tcagtttcat agcctgaaga acgagatcag cagcctctgt tccacataca cttcattctc  | 13260 |
| agtattgttt tgccaagttc taattccatc agacctcgac ctgcagcccc tagtcgagcg  | 13320 |
| ccagtagcag caccacgtc caccttctgt ctagtaatgt ccaacacctc cctcagtcca   | 13380 |
| aacactgctc tgcatccatg tggctcccat ttataacctga agcacttgat ggggcctcaa | 13440 |
| tgttttacta gagcccaccc ccctgcaact ctgagaccct ctggatttgt ctgtcagtgc  | 13500 |
| ctcactgggg cgttggataa tttcttaaaa ggtcaagttc cctcagcagc atttctgag   | 13560 |
| cagtctgaag atgtgtgctt ttcacagttc aaatccatgt ggctgtttca cccacctgcc  | 13620 |
| tggccttggg ttatctatca ggacctagcc tagaagcagg tgtgtggcac ttaacaccta  | 13680 |
| agctgagtga ctaactgaac actcaagtgg atgccatctt tgtcacttct tgactgtgac  | 13740 |
| acaagcaact cctgatgcca aagccctgcc caccctctc atgcccataat ttggacatgg  | 13800 |
| tacaggctct cactggccat ggtctgtgag gtcctgggtcc tctttgactt cataattcct | 13860 |
| aggggccact agtatctata agaggaagag ggtgctggct cccaggccac agcccacaaa  | 13920 |
| attccacctg ctcacagggt ggctggctcg acccagggtg tgtcccctgc tctgagccag  | 13980 |
| ctcccgcca agccagcacc atgggaaccc ccaagaagaa gaggaagggt cgtagcgatt   | 14040 |
| taaattccaa ttactgacc gtacaccaaa atttgcctgc attaccgggt gatgcaacga   | 14100 |

|             |             |             |             |             |             |       |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| gtgatgaggt  | tcgcaagaac  | ctgatggaca  | tgttcaggga  | tcgccaggcg  | ttttctgagc  | 14160 |
| atacctggaa  | aatgcttctg  | tccgtttgcc  | ggtcgtgggc  | ggcatggtgc  | aagttgaata  | 14220 |
| accggaaatg  | gtttcccgca  | gaacctgaag  | atgttcgcga  | ttatcttcta  | tatcttcagg  | 14280 |
| cgcgcggtct  | ggcagtaaaa  | actatccagc  | aacatttggg  | ccagctaaac  | atgcttcatc  | 14340 |
| gtcgggccgg  | gctgccacga  | ccaagtgaac  | gcaatgctgt  | ttcactgggt  | atgcggcgga  | 14400 |
| tccgaaaaga  | aaacgttgat  | gccggtgaac  | gtgcaaaaaca | ggctctagcg  | ttcgaacgca  | 14460 |
| ctgatttcga  | ccaggttcgt  | tcactcatgg  | aaaatagcga  | tcgctgccag  | gatatacgta  | 14520 |
| atctggcatt  | tctggggatt  | gcttataaca  | ccctgttacg  | tatagccgaa  | attgccagga  | 14580 |
| tcagggttaa  | agatatctca  | cgtactgacg  | gtgggagaat  | gttaatccat  | attggcagaa  | 14640 |
| cgaaaacgct  | ggtttagcacc | gcagggtgtag | agaaggcact  | tagcctgggg  | gtaactaaac  | 14700 |
| tggtcgagcg  | atggatttcc  | gtctctgggtg | tagctgatga  | tccgaataac  | tacctgtttt  | 14760 |
| gccgggtcag  | aaaaaatggt  | gttgccgcgc  | catctgccac  | cagccagcta  | tcaactcgcg  | 14820 |
| ccctggaagg  | gattttttgaa | gcaactcatc  | gattgattta  | cggcgctaag  | gtaaatataa  | 14880 |
| aatttttaag  | tgtataatgt  | gttaaactac  | tgattctaata | tgtttgtgta  | tttttaggatg | 14940 |
| actctggtca  | gagataacctg | gcctgggtctg | gacacagtgc  | ccgtgtcgga  | gccgcgcgag  | 15000 |
| atatggccccg | cgctggagtt  | tcaataccgg  | agatcatgca  | agctgggtggc | tggaccaatg  | 15060 |
| taaatattgt  | catgaactat  | atccgtaacc  | tggatagtga  | aacaggggca  | atggtgcgcc  | 15120 |
| tgctggaaga  | tggcgattga  | tctagataag  | taatgatcat  | aatcagccat  | atcacatctg  | 15180 |
| tagaggtttt  | acttgcttta  | aaaaacctcc  | cacacctccc  | cctgaacctg  | aaacataaaa  | 15240 |
| tgaatgcaat  | tgttgttggt  | aaacctgccc  | tagttgcggc  | caattccagc  | tgagcgtgag  | 15300 |
| ctcaccatta  | ccagttggtc  | tggtgtcaaa  | aataataata  | accgggcagg  | ggggatctaa  | 15360 |
| gctctagata  | agtaatgatc  | ataatcagcc  | atatcacatc  | tgtagagggt  | ttacttgctt  | 15420 |
| taaaaaacct  | cccacacctc  | cccctgaacc  | tgaaacataa  | aatgaatgca  | attgttggtg  | 15480 |
| ttaacttggt  | tattgcagct  | tataatgggt  | acaaataaag  | caatagcatc  | acaaatttca  | 15540 |
| caaataaagc  | atttttttca  | ctgcattcta  | gttgtgggtt  | gtccaaactc  | atcaatgtat  | 15600 |

|                                                                    |       |
|--------------------------------------------------------------------|-------|
| cttatcatgt ctggatcctc gacataactt cgtataatgt atgctatacg aagttattgc  | 15660 |
| tagcgtgtct cagcagtggtg tgggcacatc cttcctcccg agtcctgctg caggacaggg | 15720 |
| tacactacac ttgtcagtag aagtctgtac ctgatgtcag gtgcatcggtt acagtgaatg | 15780 |
| actcttccta gaatagatgt actcttttag ggccttatgt ttacaattat cctaagtact  | 15840 |
| attgctgtct tttaaagata tgaatgatgg aatatacact tgaccataac tgctgattgg  | 15900 |
| ttttttgttt tgttttgttt gttttcttgg aaacttatga ttcctggttt acatgtacca  | 15960 |
| cactgaaacc ctcgtagct ttacagataa agtgtgagtt gacttcctgc ccctctgtgt   | 16020 |
| tctgtggat gtccgattac ttctgccaca gctaaacatt agagcattta aagtttgcag   | 16080 |
| ttcctcagaa aggaacttag tctgactaca gattagttct tgagagaaga cactgatagg  | 16140 |
| gcagagctgt aggtgaaatc agttgttagc ctttccttta tagacgtagt ctttcagatt  | 16200 |
| cggctctgtac agaaatgccg aggggtcatg catgggccct gagtatcgtg acctgtgaca | 16260 |
| agttttttgt tggtttattg tagttctgtc aaagaaagtg gcatttgttt ttataattgt  | 16320 |
| tgccaacttt taaggttaat tttcattatt tttgagccga attaaaatgc gcacctctg   | 16380 |
| tgcccttccc aatcttgga aatataattt cttggcagag ggtcagattt cagggccag    | 16440 |
| tcactttcat ctgaccaccc tttgcacggc tgccgtgtgc ctggcttaga ttagaagtcc  | 16500 |
| ttgttaagta tgtcagagta cattcgtga taagatcttt gaagagcagg gaagcgtctt   | 16560 |
| gcctctttcc tttggtttct gcctgtactc tgggtgtttcc cgtgtcacct gcatcatagg | 16620 |
| aacagcagag aaatctgacc cagtgcattt tttctaggtg ctactatggc aaactcaagt  | 16680 |
| ggtctgtttc tgttcctgta acgttcgact atctcgctag ctgtgaagta ctgattagtg  | 16740 |
| gagttctgtg caacagcagt gtaggagtat acacaaacac aaatatgtgt ttctatttaa  | 16800 |
| aactgtggac ttagcataaa aaggggagaat atattttattt ttacaaaag ggataaaaat | 16860 |
| gggccccgtt cctcaccac cagatttagc gagaaaaagc tttctattct gaaaggtcac   | 16920 |
| ggtggctttg gcattacaaa tcagaacaac acacactgac catgatggct tgtgaactaa  | 16980 |
| ctgcaaggca ctccgtcatg gtaagcgagt aggtcccacc tcctagtgtg ccgctcattg  | 17040 |
| ctttacacag tagaatctta tttgagtgtc aattgttgtc tttgctgctt tactgtgttg  | 17100 |

ttatagaaaa tgtaagctgt acagtgaata agttattgaa gcatgtgtaa acactgttat 17160

atatcttttc tcctagatgg ggaattttga ataaaatacc ttgaaattc 17210