SEQUENCE LISTING

<110> AGENCY FOR SCIENCE TECHNOLOGY AND RESEARCH
NG, HUCK-HUI
HENG, Dominic JC

<120> A Nuclear receptor and the use of the same in the Reprogramming of cells

<130> AY/YF/ak/2010.5274

<140> PCT/SG2010/000372
<141> 2010-09-30

<150> SG 200906546-7
<151> 2009-09-30

<150> SG 201000140-2
<151> 2010-01-09

<160> 11

<170> PatentIn version 3.5

<210> 1
<211> 2762
<212> DNA
<213> Mus musculus

<400> 1
gctgtaagcc aaaggactgc caataatttc gctaagaatg tctgctagtt tggatactgg 60

agattttcaa gaatttctta agcatggact tacagctatt gcgtctgcac cagggtcaga 120

gactcgccac tcccccaaac gtgaggaaca actccgggaa aaacgtgctg ggcttccgga 180

ccgacaccga cgccccattc ccgcccgcag ccgccttgtc atgctgccca aagtggagac 240

ggaagcccca ggactggtcc gatcgcatgg ggaacagggg cagatgccag aaaacatgca 300

agtgtctcaa tttaaaatgg tgaattactc ctatgatgaa gatctggaag agctatgtcc 360

tgtgtgtggc gataaagtgt ctgggtacca ttacggtctc ctcacgtgcg aaagctgcaa 420

gggtttttt aagcgaactg tccaaaacca aaaaaggtac acgtgcatag agaaccagaa 480

ttgccaaatt gacaaaacgc agagaaaacg atgtccctac tgtcgattca aaaaatgtat 540

cgatgttggg atgaagctgg aagccgtaag agccgaccgc atgcgagggg gcagaaataa 600

gtttgggcca atgtacaaga gagacagggc tttgaagcag cagaagaaag ccctcattcg 660

agccaatgga cttaagctgg aagccatgtc tcaggtgatc caagcaatgc cctcagacct 720

gacctctgca attcagaaca ttcattccgc ctccaaaggc ctacctctga gccatgtagc 780

cttgcctccg acagactatg acagaagtcc ctttgtcaca tctcccatta gcatgacaat 840

gccacctcac agcagcctgc atggttacca accctatggt cactttccta gtcgggccat 900

caagtctgag tacccagacc cctactccag ctcacctgag tcaatgatgg gttactccta 960

catggatggt taccagacaa actccccggc cagcatccca cacctgatac tggaactttt 1020

gaagtgtgaa ccagatgagc ctcaagttca agcgaagatc atggcttacc tccagcaaga 1080

gcagagtaac cgaaacaggc aagaaaagct gagcgcattt gggcttttat gcaaaatggc 1140

ggaccagacc ctgttctcca ttgttgagtg ggccaggagt agtatcttct tcagggaact 1200

gaaggttgat gaccaaatga agctgcttca aaactgctgg agtgagctct tgattctcga 1260

tcacatttac cgacaagtgg cgcatgggaa ggaagggaca atcttcctgg ttactggaga 1320

acacgtggac tactccacca tcatctcaca cacagaagtc gcgttcaaca acctcctgag 1380

tctcgcacag gagctggtgg tgaggctccg ttcccttcag ttcgatcagc gggagtttgt 1440

atgtctcaag ttcctggtgc tgttcagctc agatgtgaag aacctggaga acctgcagct 1500

ggtggaaggt gtccaagagc aggtgaatgc cgccctgctg gactacacgg tttgcaacta 1560

cccacaacag actgagaaat tcggacagct acttcttcgg ctacccgaga tccgggcaat 1620

cagcaagcag gcagaagact acctgtacta taagcacgtg aacggggatg tgccctataa 1680

taacctcctc attgagatgc tgcatgccaa aagagcctaa gtccccaccc ctggaagctt 1740

gctctaggaa cacagactgg aaggagaaga ggaggacgat gacagaaaca caatactctg 1800

aactgctcca agcaatgcta attataaact tggtttaaag acactgaatt ttaaaagcat 1860

aataattaaa tacctaatag caaataaatg atatatcagg gtatttgtac tgcaaactgt 1920

gaatcaaagg ctgtatgaat caaaggattc atatgaaaga cattgtaatg gggtggattg 1980

aacttacaga tggagaccaa taccacagca gaataaaaat ggacagaaca atccttgtat 2040

atttaaacta atctgctatt aagaaattca gaagttgatc tctgttatta attggatttg 2100

tcctgaatta ctccgtggtg acgctgaaca actcaagaat acatgggctg tgccttggca 2160

gcccctcccc atcctcccca ccaccaccac ccccaccccc acaaggccct ataccttctg 2220

acctgtgagc cctgaagcta ttttaaggac ttctgttcag ccatacccag tagtagctcc 2280

actaaaccat gatttctgga tgtctgtgtc ttagacctgc caacagctaa taagaacaat 2340

gtataaatat gtcagcttgc attttaaata tgtgctgaag tttgttttgt cgtgtgttcg 2400

taattaaaaa gaaaacgggc agtaaccctc ttctatataa gcattagtta atattaaggg 2460

aaatcaaaca aatctaagcc aatactccca acaagcaagt tagatcttac ttctgctgct 2520

gttgctgaaa tgtggctttg gcatggttgg gtttcataaa acttttggc caagaggctt 2580

gttagtatac atccatctgt ttagtcatca aggtttgtag ttcacttaaa aaaaaataaa 2640

ccactagaca tcttttgctg aatgtcaaat agtcacagtc taagtagcca aaaagtcaaa 2700

gcgtgttaaa cattgccaaa tgaaggaaag ggtgagctgc aaaggggatg gttcgaggtt 2760

ca 2762

<210> 2
<211> 560
<212> PRT
<213> Mus musculus

<400> 2

Met Ser Ala Ser Leu Asp Thr Gly Asp Phe Gln Glu Phe Leu Lys His
1 5 10 15

Gly Leu Thr Ala Ile Ala Ser Ala Pro Gly Ser Glu Thr Arg His Ser
20 25 30

Pro Lys Arg Glu Glu Gln Leu Arg Glu Lys Arg Ala Gly Leu Pro Asp
35 40 45

Arg His Arg Arg Pro Ile Pro Ala Arg Ser Arg Leu Val Met Leu Pro
50 55 60

Lys Val Glu Thr Glu Ala Pro Gly Leu Val Arg Ser His Gly Glu Gln
65 70 75 80

Gly Gln Met Pro Glu Asn Met Gln Val Ser Gln Phe Lys Met Val Asn
85 90 95

Tyr Ser Tyr Asp Glu Asp Leu Glu Glu Leu Cys Pro Val Cys Gly Asp
100 105 110

Lys Val Ser Gly Tyr His Tyr Gly Leu Leu Thr Cys Glu Ser Cys Lys
115 120 125

Gly Phe Phe Lys Arg Thr Val Gln Asn Gln Lys Arg Tyr Thr Cys Ile
130 135 140

Glu Asn Gln Asn Cys Gln Ile Asp Lys Thr Gln Arg Lys Arg Cys Pro
145 150 155 160

Tyr Cys Arg Phe Lys Lys Cys Ile Asp Val Gly Met Lys Leu Glu Ala
165 170 175

Val Arg Ala Asp Arg Met Arg Gly Gly Arg Asn Lys Phe Gly Pro Met
180 185 190

Tyr Lys Arg Asp Arg Ala Leu Lys Gln Gln Lys Lys Ala Leu Ile Arg
195 200 205

Ala Asn Gly Leu Lys Leu Glu Ala Met Ser Gln Val Ile Gln Ala Met
210 215 220

Pro Ser Asp Leu Thr Ser Ala Ile Gln Asn Ile His Ser Ala Ser Lys
225 230 235 240

Gly Leu Pro Leu Ser His Val Ala Leu Pro Pro Thr Asp Tyr Asp Arg
245 250 255

Ser Pro Phe Val Thr Ser Pro Ile Ser Met Thr Met Pro Pro His Ser
260 265 270

Ser Leu His Gly Tyr Gln Pro Tyr Gly His Phe Pro Ser Arg Ala Ile
275 280 285

Lys Ser Glu Tyr Pro Asp Pro Tyr Ser Ser Ser Pro Glu Ser Met Met
290 295 300

Gly Tyr Ser Tyr Met Asp Gly Tyr Gln Thr Asn Ser Pro Ala Ser Ile
305 310 315 320

Pro His Leu Ile Leu Glu Leu Leu Lys Cys Glu Pro Asp Glu Pro Gln
325 330 335

Val Gln Ala Lys Ile Met Ala Tyr Leu Gln Gln Glu Gln Ser Asn Arg
340 345 350

Asn Arg Gln Glu Lys Leu Ser Ala Phe Gly Leu Leu Cys Lys Met Ala
355 360 365

Asp Gln Thr Leu Phe Ser Ile Val Glu Trp Ala Arg Ser Ser Ile Phe
370 375 380

Phe Arg Glu Leu Lys Val Asp Asp Gln Met Lys Leu Leu Gln Asn Cys
385 390 395 400

Trp Ser Glu Leu Leu Ile Leu Asp His Ile Tyr Arg Gln Val Ala His
405 410 415

Gly Lys Glu Gly Thr Ile Phe Leu Val Thr Gly Glu His Val Asp Tyr
420 425 430

Ser Thr Ile Ile Ser His Thr Glu Val Ala Phe Asn Asn Leu Leu Ser
435 440 445

Leu Ala Gln Glu Leu Val Val Arg Leu Arg Ser Leu Gln Phe Asp Gln
450 455 460

Arg Glu Phe Val Cys Leu Lys Phe Leu Val Leu Phe Ser Ser Asp Val
465 470 475 480

Lys Asn Leu Glu Asn Leu Gln Leu Val Glu Gly Val Gln Glu Gln Val
485 490 495

Asn Ala Ala Leu Leu Asp Tyr Thr Val Cys Asn Tyr Pro Gln Gln Thr
500 505 510

Glu Lys Phe Gly Gln Leu Leu Leu Arg Leu Pro Glu Ile Arg Ala Ile
515 520 525

Ser Lys Gln Ala Glu Asp Tyr Leu Tyr Tyr Lys His Val Asn Gly Asp
530 535 540

Val Pro Tyr Asn Asn Leu Leu Ile Glu Met Leu His Ala Lys Arg Ala
545 550 555 560

<210> 3
<211> 560
<212> PRT
<213> Mus musculus

<400> 3

Met Ser Ala Ser Leu Asp Thr Gly Asp Phe Gln Glu Phe Leu Lys His
1 5 10 15

Gly Leu Thr Ala Ile Ala Ser Ala Pro Gly Ser Glu Thr Arg His Ser
20 25 30

Pro Lys Arg Glu Glu Gln Leu Arg Glu Lys Arg Ala Gly Leu Pro Asp
35 40 45

Arg His Arg Arg Pro Ile Pro Ala Arg Ser Arg Leu Val Met Leu Pro
50 55 60

Lys Val Glu Thr Glu Ala Pro Gly Leu Val Arg Ser His Gly Glu Gln
65 70 75 80

Gly Gln Met Pro Glu Asn Met Gln Val Ser Gln Phe Lys Met Val Asn
85 90 95

Tyr Ser Tyr Asp Glu Asp Leu Glu Glu Leu Cys Pro Val Cys Gly Asp
100 105 110

Lys Val Ser Gly Tyr His Tyr Gly Leu Leu Thr Cys Glu Ser Cys Lys
115 120 125

Gly Phe Phe Lys Arg Thr Val Gln Asn Gln Lys Arg Tyr Thr Cys Ile
130 135 140

Glu Asn Gln Asn Cys Gln Ile Asp Lys Thr Gln Arg Lys Arg Cys Pro
145 150 155 160

Tyr Cys Arg Phe Lys Lys Cys Ile Asp Val Gly Met Lys Leu Glu Ala
165 170 175

Val Arg Ala Asp Arg Met Arg Gly Gly Arg Asn Lys Phe Gly Pro Met
180 185 190

Tyr Lys Arg Asp Arg Ala Leu Lys Gln Gln Lys Lys Ala Leu Ile Arg
195 200 205

Ala Asn Gly Leu Lys Leu Glu Ala Met Ser Gln Val Ile Gln Ala Met
210 215 220

Pro Ser Asp Leu Thr Ser Ala Ile Gln Asn Ile His Ser Ala Ser Lys
225 230 235 240

Gly Leu Pro Leu Ser His Val Ala Leu Pro Pro Thr Asp Tyr Asp Arg

245 250 255

Ser Pro Phe Val Thr Ser Pro Ile Ser Met Thr Met Pro Pro His Ser
260 265 270

Ser Leu His Gly Tyr Gln Pro Tyr Gly His Phe Pro Ser Arg Ala Ile
275 280 285

Lys Ser Glu Tyr Pro Asp Pro Tyr Ser Ser Ser Pro Glu Ser Met Met
290 295 300

Gly Tyr Ser Tyr Met Asp Gly Tyr Gln Thr Asn Ser Pro Ala Ser Ile
305 310 315 320

Pro His Leu Ile Leu Glu Leu Leu Lys Cys Glu Pro Asp Glu Pro Gln
325 330 335

Val Gln Ala Lys Ile Met Ala Tyr Leu Gln Gln Glu Gln Ser Asn Arg
340 345 350

Asn Arg Gln Glu Lys Leu Ser Ala Phe Gly Leu Leu Cys Lys Met Ala
355 360 365

Asp Gln Thr Leu Phe Ser Ile Val Glu Trp Ala Arg Ser Ser Ile Phe
370 375 380

Phe Arg Glu Leu Lys Val Asp Asp Gln Met Lys Leu Leu Gln Asn Cys
385 390 395 400

Trp Ser Glu Leu Leu Ile Leu Asp His Ile Tyr Arg Gln Val Ala His
405 410 415

Gly Lys Glu Gly Thr Ile Phe Leu Val Thr Gly Glu His Val Asp Tyr
420 425 430

Ser Thr Ile Ile Ser His Thr Glu Val Ala Phe Asn Asn Leu Leu Ser
435 440 445

Leu Ala Gln Glu Leu Val Val Arg Leu Arg Ser Leu Gln Phe Asp Gln
450 455 460

Arg Glu Phe Val Cys Leu Lys Phe Leu Val Leu Phe Ser Ser Asp Val
465 470 475 480

Lys Asn Leu Glu Asn Leu Gln Leu Val Glu Gly Val Gln Glu Gln Val
485 490 495

Asn Ala Ala Leu Leu Asp Tyr Thr Val Cys Asn Tyr Pro Gln Gln Thr
500 505 510

Glu Lys Phe Gly Gln Leu Leu Leu Arg Leu Pro Glu Ile Arg Ala Ile

515 520 525

Ser Lys Gln Ala Glu Asp Tyr Leu Tyr Tyr Lys His Val Asn Gly Asp
530 535 540

Val Pro Tyr Asn Asn Leu Leu Ile Glu Met Leu His Ala Lys Arg Ala
545 550 555 560

<210> 4
<211> 2947
<212> DNA
<213> Mus musculus

<400> 4
gaggggagga ggaaaggacg atcggacagg gccagtttcc agtccgccgc tgcccgcccg 60
ctgctgggtg aagaagtttc tgagagcccg ctagccactg ccctacctga ggcctgggag 120
cctccccacc aggaccctgg tgtccagtgt ccacccttat ccggctgaga attctccttc 180
cgttcagcgg acgccgcggg catggactat tcgtacgacg aggacctgga cgagctgtgt 240
ccagtgtgtg gtgacaaggt gtcgggctac cactacgggc tgctcacgtg cgagagctgc 300
aagggcttct tcaagcgcac agtccagaac aacaagcatt acacgtgcac cgagagtcag 360
agctgcaaaa tcgacaagac gcagcgtaag cgctgtccct tctgccgctt ccagaagtgc 420
ctgacggtgg gcatgcgcct ggaagctgtg cgtgctgatc gaatgcgggg tggccggaac 480
aagtttgggc ccatgtacaa gagagaccgg gccttgaagc agcagaagaa agcacagatt 540
cgggccaatg gcttcaagct ggagaccgga ccaccgatgg gggtgccccc gccacccccc 600
cccccaccgg actacatgtt acccccctagc ctgcacgcac cggagcccaa ggccctggtc 660
tctggcccac ccagtgggcc gctgggtgac tttggagccc catctctacc catggctgtg 720
cctggtcccc acggacctct ggctggctac ctctatcctg ccttctctaa ccgcaccatc 780
aagtctgagt atccagagcc ctatgccagc cccccacaac agccagggcc accctacagc 840
tatccagagc ccttctcagg agggcccaat gtaccagagc tcatattgca gctgctgcaa 900
ctagagccag aggaggacca ggtgcgcgct cgcatcgtgg gctgtctgca ggagccagcc 960
aaaagccgct ctgaccagcc agcgcccttc agcctcctct gcagaatggc cgaccagacc 1020
tttatctcca ttgtcgactg ggcacgaagg tgcatggtct ttaaggagct ggaggtggct 1080
gaccagatga cactgctgca gaactgttgg agcgagctgc tggtgttgga ccacatctac 1140
cgccaggtcc agtacggcaa ggaagacagc atcctgctgg ttactggaca ggaggtggag 1200
ctgagcacag tggctgtgca ggctggctcc ctgctgcaca gcctggtgct gcgggcccaa 1260
gagttagtgc tccagttgca tgcactgcag ctggaccgcc aggagttcgt ctgtctcaag 1320
ttcctcatcc tcttcagcct cgatgtgaaa ttcctgaaca accacagcct cgtaaaggac 1380
gcccaggaaa aggccaacgc tgccctgttg gattacacct tgtgtcacta cccacactgc 1440
ggggacaaat tccagcagtt gctattgtgc ctggtggagg tgcgggccct gagcatgcag 1500
gccaaggagt acctgtacca caagcatttg ggcaacgaga tgccccgcaa caaccttctc 1560

attgagatgc tgcaggccaa gcagacttga gcctgggtgc caggcagcgg gcaataggca 1620

gggatgccac tgcctccaaa agactccttg cattaggtga tccaggagcc ctgtcactaa 1680

gcccctgccc ctgagctcca gagctgtgtg tttgggcaag gatgggcggg gattggccgg 1740

ggcaggttgc ctttactagc cattggcctg tgtccgccac ttggagtgcc ccaaaggggg 1800

cttctaacca ttccttcctc catcagcccc cagcttttt tcctggtatc tgaggtccca 1860

ggaggaggct caggattccc tggtgggtct ggatgtccct tgggtcagag gtcatccttt 1920

ccctctctcc tgttatcaga ggcaaaggaa ggtctacagg catcaatgag ggcaaaggag 1980

ggggtctcca gactccactg aagcaggaag tccactgttg taaactgagt ttgctaaatt 2040

gggtccccag aggataccat gagagtgggt agggcaaaaa gagccctttc cgccctctac 2100

ccatctaatt ctgatcctct acctgtagga ggactttggt gtgatcatcc ttctcccagg 2160

gcccggctac ccagggagga ggagtctggt gtagccaaca ttcctgccct aaccctgccc 2220

atcaccagct ggctgggctg gtatttatct gcaaggttga agtcactggg attcttttcc 2280

tttcacctag atagtccttg gaaagtgtgt gagagagaag tgggcaggag acagactggg 2340

gactgagctg ggatatgggg actagcatca aagctttctc ctgacatctc tttccaagag 2400

tcggggtggc atctgtaccc cacctcaccc ccgagaagtg ctattgcttg ccctctgcct 2460

cagccccact aggggaacaa caggaggcct gctggggctt agagtccgtg caggtgggga 2520

tatgggtaaa tctaggagaa ctcacagatc tttatatgag gacagtgctg aggactttct 2580

catggctcca tccttttggt ccctcgccac taccccttgaa gctggcttca gttccctggc 2640

tgctgctttg cctcctgaaa gccactctgt aggaccaagc actcggggga gaggcctaag 2700

ccatcctctg ttccagactg gacatccact gtctttcctg ctttcgcgtc agatttacag 2760

cttatgctag gcccacccaa ctggacaagg ctgtctcctg tcttctacta ccctggctca 2820

gcccccacct ctgcccctga aatgcgtgct cccaccaagg ccagagaccc acagccccaa 2880

gacaagaagt gcccttataa acccctgcag ccctgcagcc ctgaaataaa ttttgcaatt 2940

agtttcc 2947

<210> 5
<211> 462
<212> PRT
<213> Mus musculus

<400> 5

Met Asp Tyr Ser Tyr Asp Glu Asp Leu Asp Glu Leu Cys Pro Val Cys
1 5 10 15

Gly Asp Lys Val Ser Gly Tyr His Tyr Gly Leu Leu Thr Cys Glu Ser
20 25 30

Cys Lys Gly Phe Phe Lys Arg Thr Val Gln Asn Asn Lys His Tyr Thr
35 40 45

Cys Thr Glu Ser Gln Ser Cys Lys Ile Asp Lys Thr Gln Arg Lys Arg
50 55 60

Cys Pro Phe Cys Arg Phe Gln Lys Cys Leu Thr Val Gly Met Arg Leu
65 70 75 80

Glu Ala Val Arg Ala Asp Arg Met Arg Gly Gly Arg Asn Lys Phe Gly
85 90 95

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

Ile Arg Ala Asn Gly Phe Lys Leu Glu Thr Gly Pro Pro Met Gly Val
115 120 125

Pro Pro Pro Pro Pro Pro Pro Pro Asp Tyr Met Leu Pro Pro Ser Leu
130 135 140

His Ala Pro Glu Pro Lys Ala Leu Val Ser Gly Pro Pro Ser Gly Pro
145 150 155 160

Leu Gly Asp Phe Gly Ala Pro Ser Leu Pro Met Ala Val Pro Gly Pro
165 170 175

His Gly Pro Leu Ala Gly Tyr Leu Tyr Pro Ala Phe Ser Asn Arg Thr
180 185 190

Ile Lys Ser Glu Tyr Pro Glu Pro Tyr Ala Ser Pro Pro Gln Gln Pro
195 200 205

Gly Pro Pro Tyr Ser Tyr Pro Glu Pro Phe Ser Gly Gly Pro Asn Val
210 215 220

Pro Glu Leu Ile Leu Gln Leu Leu Gln Leu Glu Pro Glu Glu Asp Gln
225 230 235 240

Val Arg Ala Arg Ile Val Gly Cys Leu Gln Glu Pro Ala Lys Ser Arg
245 250 255

Ser Asp Gln Pro Ala Pro Phe Ser Leu Leu Cys Arg Met Ala Asp Gln
260 265 270

Thr Phe Ile Ser Ile Val Asp Trp Ala Arg Arg Cys Met Val Phe Lys
275 280 285

Glu Leu Glu Val Ala Asp Gln Met Thr Leu Leu Gln Asn Cys Trp Ser
290 295 300

Glu Leu Leu Val Leu Asp His Ile Tyr Arg Gln Val Gln Tyr Gly Lys
305 310 315 320

Glu Asp Ser Ile Leu Leu Val Thr Gly Gln Glu Val Glu Leu Ser Thr
325 330 335

Val Ala Val Gln Ala Gly Ser Leu Leu His Ser Leu Val Leu Arg Ala
340 345 350

Gln Glu Leu Val Leu Gln Leu His Ala Leu Gln Leu Asp Arg Gln Glu
355 360 365

Phe Val Cys Leu Lys Phe Leu Ile Leu Phe Ser Leu Asp Val Lys Phe
370 375 380

Leu Asn Asn His Ser Leu Val Lys Asp Ala Gln Glu Lys Ala Asn Ala
385 390 395 400

Ala Leu Leu Asp Tyr Thr Leu Cys His Tyr Pro His Cys Gly Asp Lys
405 410 415

Phe Gln Gln Leu Leu Leu Cys Leu Val Glu Val Arg Ala Leu Ser Met
420 425 430

Gln Ala Lys Glu Tyr Leu Tyr His Lys His Leu Gly Asn Glu Met Pro
435 440 445

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

<210> 6
<211> 88
<212> PRT
<213> Artificial Sequence

<220>
<223> from Nuclear receptor subfamily 5 sequences

<400> 6

Asp Glu Asp Leu Glu Glu Leu Cys Pro Val Cys Gly Asp Lys Val Ser
1 5 10 15

Gly Tyr His Tyr Gly Leu Leu Thr Cys Glu Ser Cys Lys Gly Phe Phe
20 25 30

Lys Arg Thr Val Gln Asn Gln Lys Arg Tyr Thr Cys Ile Glu Asn Gln
35 40 45

Asn Cys Gln Ile Asp Lys Thr Gln Arg Lys Arg Cys Pro Tyr Cys Arg
50 55 60

Phe Lys Lys Cys Ile Asp Val Gly Met Lys Leu Glu Ala Val Arg Ala
65 70 75 80

Asp Arg Met Arg Gly Gly Arg Asn

<210> 7
<211> 244
<212> PRT
<213> Artificial Sequence

<220>
<223> from Nuclear receptor subfamily 5 proteins

<400> 7

Pro Ala Ser Ile Pro His Leu Ile Leu Glu Leu Leu Lys Cys Glu Pro
1 5 10 15

Asp Glu Pro Gln Val Gln Ala Lys Ile Met Ala Tyr Leu Gln Gln Glu
20 25 30

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

Cys Lys Met Ala Asp Gln Thr Leu Phe Ser Ile Val Glu Trp Ala Arg
50 55 60

Ser Ser Ile Phe Phe Arg Glu Leu Lys Val Asp Asp Gln Met Lys Leu
65 70 75 80

Leu Gln Asn Cys Trp Ser Glu Leu Leu Ile Leu Asp His Ile Tyr Arg
85 90 95

Gln Val Ala His Gly Lys Glu Gly Thr Ile Phe Leu Val Thr Gly Glu
100 105 110

His Val Asp Tyr Ser Thr Ile Ile Ser His Thr Glu Val Ala Phe Asn
115 120 125

Asn Leu Leu Ser Leu Ala Gln Glu Leu Val Val Arg Leu Arg Ser Leu
130 135 140

Gln Phe Asp Gln Arg Glu Phe Val Cys Leu Lys Phe Leu Val Leu Phe
145 150 155 160

Ser Ser Asp Val Lys Asn Leu Glu Asn Leu Gln Leu Val Glu Gly Val
165 170 175

Gln Glu Gln Val Asn Ala Ala Leu Leu Asp Tyr Thr Val Cys Asn Tyr
180 185 190

Pro Gln Gln Thr Glu Lys Phe Gly Gln Leu Leu Leu Arg Leu Pro Glu
195 200 205

Ile Arg Ala Ile Ser Lys Gln Ala Glu Asp Tyr Leu Tyr Tyr Lys His
210 215 220

Val Asn Gly Asp Val Pro Tyr Asn Asn Leu Leu Ile Glu Met Leu His
225 230 235 240

Ala Lys Arg Ala

<210> 8
<211> 1683
<212> DNA
<213> Artificial Sequence

<220>
<223> sumoylated mutant of nuclear receptor construct with two lysine residues mutated

<400> 8
atgtctgcta gtttggatac tggagatttt caagaatttc ttaagcatgg acttacagct 60

attgcgtctg caccagggtc agagactcgc cactccccca aacgtgagga acaactccgg 120

gaaaaacgtg ctgggcttcc ggaccgacac cgacgcccca ttcccgcccg cagccgcctt 180

gtcatgctgc ccaaagtgga gacggaagcc ccaggactgg tccgatcgca tggggaacag 240

gggcagatgc cagaaaacat gcaagtgtct caatttaaaa tggtgaatta ctcctatgat 300

gaagatctgg aagagctatg tcctgtgtgt ggcgataaag tgtctgggta ccattacggt 360

ctcctcacgt gcgaaagctg caagggtttt tttaagcgaa ctgtccaaaa ccaaaaaagg 420

tacacgtgca tagagaacca gaattgccaa attgacaaaa cgcagagaaa acgatgtccc 480

tactgtcgat tcaaaaaatg tatcgatgtt gggatgaggc tggaagccgt aagagccgac 540

cgcatgcgag ggggcagaaa taagtttggg ccaatgtaca agagagacag ggctttgaag 600

cagcagaaga aagccctcat tcgagccaat ggacttaagc tggaagccat gtctcaggtg 660

atccaagcaa tgccctcaga cctgacctct gcaattcaga acattcattc cgcctccaaa 720

ggcctacctc tgagccatgt agccttgcct ccgacagact atgacagaag tccctttgtc 780

acatctccca ttagcatgac aatgccacct cacagcagcc tgcatggtta ccaaccctat 840

ggtcactttc ctagtcgggc catcaggtct gagtacccag acccctactc cagctcacct 900

gagtcaatga tgggttactc ctacatggat ggttaccaga caaactcccc ggccagcatc 960

ccacacctga tactggaact tttgaagtgt gaaccagatg agcctcaagt tcaagcgaag 1020

atcatggctt acctccagca agagcagagt aaccgaaaca ggcaagaaaa gctgagcgca 1080

tttgggcttt tatgcaaaat ggcggaccag accctgttct ccattgttga gtgggccagg 1140

agtagtatct tcttcaggga actgaaggtt gatgaccaaa tgaagctgct tcaaaactgc 1200

tggagtgagc tcttgattct cgatcacatt taccgacaag tggcgcatgg gaaggaaggg 1260

acaatcttcc tggttactgg agaacacgtg gactactcca ccatcatctc acacacagaa 1320

gtcgcgttca acaacctcct gagtctcgca caggagctgg tggtgaggct ccgttccctt 1380

cagttcgatc agcgggagtt tgtatgtctc aagttcctgg tgctgttcag ctcagatgtg 1440

aagaacctgg agaacctgca gctggtggaa ggtgtccaag agcaggtgaa tgccgccctg 1500

ctggactaca cggtttgcaa ctacccacaa cagactgaga aattcggaca gctacttctt 1560

cggctacccg agatccgggc aatcagcaag caggcagaag actacctgta ctataagcac 1620

gtgaacgggg atgtgcccta taataacctc ctcattgaga tgctgcatgc caaaagagcc 1680

taa 1683

<210> 9
<211> 560
<212> PRT
<213> Artificial Sequence

<220>
<223> A sumoylated mutant of Nr5a2 expression product with two lysine residues mutated

<400> 9

Met Ser Ala Ser Leu Asp Thr Gly Asp Phe Gln Glu Phe Leu Lys His
1 5 10 15

Gly Leu Thr Ala Ile Ala Ser Ala Pro Gly Ser Glu Thr Arg His Ser
20 25 30

Pro Lys Arg Glu Glu Gln Leu Arg Glu Lys Arg Ala Gly Leu Pro Asp
35 40 45

Arg His Arg Arg Pro Ile Pro Ala Arg Ser Arg Leu Val Met Leu Pro
50 55 60

Lys Val Glu Thr Glu Ala Pro Gly Leu Val Arg Ser His Gly Glu Gln
65 70 75 80

Gly Gln Met Pro Glu Asn Met Gln Val Ser Gln Phe Lys Met Val Asn
85 90 95

Tyr Ser Tyr Asp Glu Asp Leu Glu Glu Leu Cys Pro Val Cys Gly Asp
100 105 110

Lys Val Ser Gly Tyr His Tyr Gly Leu Leu Thr Cys Glu Ser Cys Lys
115 120 125

Gly Phe Phe Lys Arg Thr Val Gln Asn Gln Lys Arg Tyr Thr Cys Ile
130 135 140

Glu Asn Gln Asn Cys Gln Ile Asp Lys Thr Gln Arg Lys Arg Cys Pro
145 150 155 160

Tyr Cys Arg Phe Lys Lys Cys Ile Asp Val Gly Met Arg Leu Glu Ala
165 170 175

Val Arg Ala Asp Arg Met Arg Gly Gly Arg Asn Lys Phe Gly Pro Met
180 185 190

Tyr Lys Arg Asp Arg Ala Leu Lys Gln Gln Lys Lys Ala Leu Ile Arg
195 200 205

Ala Asn Gly Leu Lys Leu Glu Ala Met Ser Gln Val Ile Gln Ala Met
210 215 220

Pro Ser Asp Leu Thr Ser Ala Ile Gln Asn Ile His Ser Ala Ser Lys
225 230 235 240

Gly Leu Pro Leu Ser His Val Ala Leu Pro Pro Thr Asp Tyr Asp Arg
245 250 255

Ser Pro Phe Val Thr Ser Pro Ile Ser Met Thr Met Pro Pro His Ser
260 265 270

Ser Leu His Gly Tyr Gln Pro Tyr Gly His Phe Pro Ser Arg Ala Ile
275 280 285

Arg Ser Glu Tyr Pro Asp Pro Tyr Ser Ser Ser Pro Glu Ser Met Met
290 295 300

Gly Tyr Ser Tyr Met Asp Gly Tyr Gln Thr Asn Ser Pro Ala Ser Ile
305 310 315 320

Pro His Leu Ile Leu Glu Leu Leu Lys Cys Glu Pro Asp Glu Pro Gln
325 330 335

Val Gln Ala Lys Ile Met Ala Tyr Leu Gln Gln Glu Gln Ser Asn Arg
340 345 350

Asn Arg Gln Glu Lys Leu Ser Ala Phe Gly Leu Leu Cys Lys Met Ala
355 360 365

Asp Gln Thr Leu Phe Ser Ile Val Glu Trp Ala Arg Ser Ser Ile Phe
370 375 380

Phe Arg Glu Leu Lys Val Asp Asp Gln Met Lys Leu Leu Gln Asn Cys
385 390 395 400

Trp Ser Glu Leu Leu Ile Leu Asp His Ile Tyr Arg Gln Val Ala His
405 410 415

Gly Lys Glu Gly Thr Ile Phe Leu Val Thr Gly Glu His Val Asp Tyr
420 425 430

Ser Thr Ile Ile Ser His Thr Glu Val Ala Phe Asn Asn Leu Leu Ser
435 440 445

Leu Ala Gln Glu Leu Val Val Arg Leu Arg Ser Leu Gln Phe Asp Gln
450 455 460

Arg Glu Phe Val Cys Leu Lys Phe Leu Val Leu Phe Ser Ser Asp Val
465 470 475 480

Lys Asn Leu Glu Asn Leu Gln Leu Val Glu Gly Val Gln Glu Gln Val
485 490 495

Asn Ala Ala Leu Leu Asp Tyr Thr Val Cys Asn Tyr Pro Gln Gln Thr
500 505 510

Glu Lys Phe Gly Gln Leu Leu Leu Arg Leu Pro Glu Ile Arg Ala Ile
515 520 525

Ser Lys Gln Ala Glu Asp Tyr Leu Tyr Tyr Lys His Val Asn Gly Asp
530 535 540

Val Pro Tyr Asn Asn Leu Leu Ile Glu Met Leu His Ala Lys Arg Ala
545 550 555 560

<210> 10
<211> 1683
<212> DNA
<213> Artificial Sequence

<220>
<223> A sumoylation construct of Nr5a2 with five lysine residues mutated

<400> 10
atgtctgcta gtttggatac tggagatttt caagaatttc ttaagcatgg acttacagct 60

attgcgtctg caccagggtc agagactcgc cactccccca aacgtgagga acaactccgg 120

gaaaaacgtg ctgggcttcc ggaccgacac cgacgcccca ttcccgcccg cagccgcctt 180

gtcatgctgc ccaaagtgga gacggaagcc ccaggactgg tccgatcgca tggggaacag 240

gggcagatgc cagaaaacat gcaagtgtct caatttaaaa tggtgaatta ctcctatgat 300

gaagatctgg aagagctatg tcctgtgtgt ggcgataaag tgtctgggta ccattacggt 360

ctcctcacgt gcgaaagctg caagggtttt tttaagcgaa ctgtccaaaa ccaaaaaagg 420

tacacgtgca tagagaacca gaattgccaa attgacaaaa cgcagagaaa acgatgtccc 480

tactgtcgat tcaaaaaatg tatcgatgtt gggatgaggc tggaagccgt aagagccgac 540

cgcatgcgag ggggcagaaa taagtttggg ccaatgtaca agagagacag ggctttgaag 600

cagcagaaga aagccctcat tcgagccaat ggacttaggc tggaagccat gtctcaggtg 660

atccaagcaa tgccctcaga cctgacctct gcaattcaga acattcattc cgcctccaaa 720

ggcctacctc tgagccatgt agccttgcct ccgacagact atgacagaag tccctttgtc 780

acatctccca ttagcatgac aatgccacct cacagcagcc tgcatggtta ccaaccctat 840

ggtcactttc ctagtcgggc catcaggtct gagtacccag acccctactc cagctcacct 900

gagtcaatga tgggttactc ctacatggat ggttaccaga caaactcccc ggccagcatc 960

ccacacctga tactggaact tttgaggtgt gaaccagatg agcctcaagt tcaagcgaag 1020

atcatggctt acctccagca agagcagagt aaccgaaaca ggcaagaaaa gctgagcgca 1080

tttgggcttt tatgcaaaat ggcggaccag accctgttct ccattgttga gtgggccagg 1140

agtagtatct tcttcaggga actgagggtt gatgaccaaa tgaagctgct tcaaaactgc 1200

tggagtgagc tcttgattct cgatcacatt taccgacaag tggcgcatgg gaaggaaggg 1260

acaatcttcc tggttactgg agaacacgtg gactactcca ccatcatctc acacacagaa 1320

gtcgcgttca acaacctcct gagtctcgca caggagctgg tggtgaggct ccgttccctt 1380

cagttcgatc agcgggagtt tgtatgtctc aagttcctgg tgctgttcag ctcagatgtg 1440

aagaacctgg agaacctgca gctggtggaa ggtgtccaag agcaggtgaa tgccgccctg 1500

ctggactaca cggtttgcaa ctacccacaa cagactgaga aattcggaca gctacttctt 1560

cggctacccg agatccgggc aatcagcaag caggcagaag actacctgta ctataagcac 1620

gtgaacgggg atgtgcccta taataacctc ctcattgaga tgctgcatgc caaaagagcc 1680

taa 1683

<210> 11
<211> 560
<212> PRT
<213> Artificial Sequence

<220>
<223> An expression product of a sumoylation mutant of Nr5a2 with 5 lysine residues mutated

<400> 11

Met Ser Ala Ser Leu Asp Thr Gly Asp Phe Gln Glu Phe Leu Lys His
1 5 10 15

Gly Leu Thr Ala Ile Ala Ser Ala Pro Gly Ser Glu Thr Arg His Ser
20 25 30

Pro Lys Arg Glu Glu Gln Leu Arg Glu Lys Arg Ala Gly Leu Pro Asp
35 40 45

Arg His Arg Arg Pro Ile Pro Ala Arg Ser Arg Leu Val Met Leu Pro
50 55 60

Lys Val Glu Thr Glu Ala Pro Gly Leu Val Arg Ser His Gly Glu Gln
65 70 75 80

Gly Gln Met Pro Glu Asn Met Gln Val Ser Gln Phe Lys Met Val Asn
85 90 95

Tyr Ser Tyr Asp Glu Asp Leu Glu Glu Leu Cys Pro Val Cys Gly Asp
100 105 110

Lys Val Ser Gly Tyr His Tyr Gly Leu Leu Thr Cys Glu Ser Cys Lys
115 120 125

Gly Phe Phe Lys Arg Thr Val Gln Asn Gln Lys Arg Tyr Thr Cys Ile
130 135 140

Glu Asn Gln Asn Cys Gln Ile Asp Lys Thr Gln Arg Lys Arg Cys Pro
145 150 155 160

Tyr Cys Arg Phe Lys Lys Cys Ile Asp Val Gly Met Arg Leu Glu Ala
165 170 175

Val Arg Ala Asp Arg Met Arg Gly Gly Arg Asn Lys Phe Gly Pro Met
180 185 190

Tyr Lys Arg Asp Arg Ala Leu Lys Gln Gln Lys Lys Ala Leu Ile Arg
195 200 205

Ala Asn Gly Leu Arg Leu Glu Ala Met Ser Gln Val Ile Gln Ala Met
210 215 220

Pro Ser Asp Leu Thr Ser Ala Ile Gln Asn Ile His Ser Ala Ser Lys
225 230 235 240

Gly Leu Pro Leu Ser His Val Ala Leu Pro Pro Thr Asp Tyr Asp Arg
245 250 255

Ser Pro Phe Val Thr Ser Pro Ile Ser Met Thr Met Pro Pro His Ser
260 265 270

Ser Leu His Gly Tyr Gln Pro Tyr Gly His Phe Pro Ser Arg Ala Ile
275 280 285

Arg Ser Glu Tyr Pro Asp Pro Tyr Ser Ser Ser Pro Glu Ser Met Met
290 295 300

Gly Tyr Ser Tyr Met Asp Gly Tyr Gln Thr Asn Ser Pro Ala Ser Ile
305 310 315 320

Pro His Leu Ile Leu Glu Leu Leu Arg Cys Glu Pro Asp Glu Pro Gln
325 330 335

Val Gln Ala Lys Ile Met Ala Tyr Leu Gln Gln Glu Gln Ser Asn Arg
340 345 350

Asn Arg Gln Glu Lys Leu Ser Ala Phe Gly Leu Leu Cys Lys Met Ala
355 360 365

Asp Gln Thr Leu Phe Ser Ile Val Glu Trp Ala Arg Ser Ser Ile Phe
370 375 380

Phe Arg Glu Leu Arg Val Asp Asp Gln Met Lys Leu Leu Gln Asn Cys
385 390 395 400

Trp Ser Glu Leu Leu Ile Leu Asp His Ile Tyr Arg Gln Val Ala His
405 410 415

Gly Lys Glu Gly Thr Ile Phe Leu Val Thr Gly Glu His Val Asp Tyr
420 425 430

Ser Thr Ile Ile Ser His Thr Glu Val Ala Phe Asn Asn Leu Leu Ser
435 440 445

Leu Ala Gln Glu Leu Val Val Arg Leu Arg Ser Leu Gln Phe Asp Gln
450 455 460

Arg Glu Phe Val Cys Leu Lys Phe Leu Val Leu Phe Ser Ser Asp Val
465 470 475 480

Lys Asn Leu Glu Asn Leu Gln Leu Val Glu Gly Val Gln Glu Gln Val
485 490 495

Asn Ala Ala Leu Leu Asp Tyr Thr Val Cys Asn Tyr Pro Gln Gln Thr
500 505 510

Glu Lys Phe Gly Gln Leu Leu Leu Arg Leu Pro Glu Ile Arg Ala Ile
515 520 525

Ser Lys Gln Ala Glu Asp Tyr Leu Tyr Tyr Lys His Val Asn Gly Asp
530 535 540

Val Pro Tyr Asn Asn Leu Leu Ile Glu Met Leu His Ala Lys Arg Ala
545 550 555 560