SEQUENCE LISTING

<110> FUNDACION PARA LA INVESTIGACION MEDICA APLICADA

<120> NUCLEIC ACID CONSTRUCTS AND GENE THERAPY VECTORS FOR USE IN THE TREATMENT OF WILSON DISEASE AND OTHER CONDITIONS

<130> 2014005EP00

<160> 21

<170> BiSSAP 1.3

<210> 1
<211> 5107
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleic acid construct of expression vector AAV2-AAT-wtATP7B

<220>
<221> repeat_region
<222> 1..141
<223> /note="5' ITR of adeno-associated virus serotype 2"

<220>
<221> promoter
<222> 156..460
<223> /note="alpha 1 antitrypsin"

<220>
<221> CDS
<222> 473..4870
<223> /note="Sequence encoding ATP7B (Copper transporting ATPase 2)"
/transl_table=1

<220>
<221> polyA_signal
<222> 4877..4932

<220>
<221> repeat_region
<222> 4968..5107
<223> /standard_name="3' ITR of adeno-associated virus serotype 2"

<400> 1
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60

gggcgacctt tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120

actccatcac taggggttcc tgcggccgca cgcgtcgcca ccccctccac cttggacaca 180

ggacgctgtg gtttctgagc caggtacaat gactcctttc ggtaagtgca gtggaagctg 240

tacactgccc aggcaaagcg tccgggcagc gtaggcgggc gactcagatc ccagccagtg 300

gacttagccc ctgtttgctc ctccgataac tggggtgacc ttggttaata ttcaccagca 360

gcctcccccg ttgcccctct ggatccactg cttaaatacg gacgaggaca gggccctgtc 420

tcctcagctt caggcaccac cactgacctg ggacagtgaa gcggccgcca cc atg cct 478
Met Pro
1

gag cag gag aga cag atc aca gcc aga gaa ggg gcc agt cgg aaa atc 526
Glu Gln Glu Arg Gln Ile Thr Ala Arg Glu Gly Ala Ser Arg Lys Ile
5 10 15

```
tta tct aag ctt tct ttg cct acc cgt gcc tgg gaa cca gca atg aag     574
Leu Ser Lys Leu Ser Leu Pro Thr Arg Ala Trp Glu Pro Ala Met Lys
    20                  25                  30

aag agt ttt gct ttt gac aat gtt ggc tat gaa ggt ggt ctg gat ggc     622
Lys Ser Phe Ala Phe Asp Asn Val Gly Tyr Glu Gly Gly Leu Asp Gly
35                  40                  45                  50

ctg ggc cct tct tct cag gtg gcc acc agc aca gtc agg atc ttg ggc     670
Leu Gly Pro Ser Ser Gln Val Ala Thr Ser Thr Val Arg Ile Leu Gly
                55                  60                  65

atg act tgc cag tca tgt gtg aag tcc att gag gac agg att tcc aat     718
Met Thr Cys Gln Ser Cys Val Lys Ser Ile Glu Asp Arg Ile Ser Asn
            70                  75                  80

ttg aaa ggc atc atc agc atg aag gtt tcc ctg gaa caa ggc agt gcc     766
Leu Lys Gly Ile Ile Ser Met Lys Val Ser Leu Glu Gln Gly Ser Ala
        85                  90                  95

act gtg aaa tat gtg cca tcg gtt gtg tgc ctg caa cag gtt tgc cat     814
Thr Val Lys Tyr Val Pro Ser Val Val Cys Leu Gln Gln Val Cys His
    100                 105                 110

caa att ggg gac atg ggc ttc gag gcc agc att gca gaa gga aag gca     862
Gln Ile Gly Asp Met Gly Phe Glu Ala Ser Ile Ala Glu Gly Lys Ala
115                 120                 125                 130

gcc tcc tgg ccc tca agg tcc ttg cct gcc cag gag gct gtg gtc aag     910
Ala Ser Trp Pro Ser Arg Ser Leu Pro Ala Gln Glu Ala Val Val Lys
                135                 140                 145

ctc cgg gtg gag ggc atg acc tgc cag tcc tgt gtc agc tcc att gaa     958
Leu Arg Val Glu Gly Met Thr Cys Gln Ser Cys Val Ser Ser Ile Glu
            150                 155                 160

ggc aag gtc cgg aaa ctg caa gga gta gtg aga gtc aaa gtc tca ctc    1006
Gly Lys Val Arg Lys Leu Gln Gly Val Val Arg Val Lys Val Ser Leu
        165                 170                 175

agc aac caa gag gcc gtc atc act tat cag cct tat ctc att cag ccc    1054
Ser Asn Gln Glu Ala Val Ile Thr Tyr Gln Pro Tyr Leu Ile Gln Pro
    180                 185                 190

gaa gac ctc agg gac cat gta aat gac atg gga ttt gaa gct gcc atc    1102
Glu Asp Leu Arg Asp His Val Asn Asp Met Gly Phe Glu Ala Ala Ile
195                 200                 205                 210

aag agc aaa gtg gct ccc tta agc ctg gga cca att gat att gag cgg    1150
Lys Ser Lys Val Ala Pro Leu Ser Leu Gly Pro Ile Asp Ile Glu Arg
                215                 220                 225

tta caa agc act aac cca aag aga cct tta tct tct gct aac cag aat    1198
Leu Gln Ser Thr Asn Pro Lys Arg Pro Leu Ser Ser Ala Asn Gln Asn
            230                 235                 240

ttt aat aat tct gag acc ttg ggg cac caa gga agc cat gtg gtc acc    1246
Phe Asn Asn Ser Glu Thr Leu Gly His Gln Gly Ser His Val Val Thr
        245                 250                 255

ctc caa ctg aga ata gat gga atg cat tgt aag tct tgc gtc ttg aat    1294
Leu Gln Leu Arg Ile Asp Gly Met His Cys Lys Ser Cys Val Leu Asn
    260                 265                 270

att gaa gaa aat att ggc cag ctc cta ggg gtt caa agt att caa gtg    1342
Ile Glu Glu Asn Ile Gly Gln Leu Leu Gly Val Gln Ser Ile Gln Val
275                 280                 285                 290
```

```
tcc ttg gag aac aaa act gcc caa gta aag tat gac cct tct tgt acc    1390
Ser Leu Glu Asn Lys Thr Ala Gln Val Lys Tyr Asp Pro Ser Cys Thr
                295                 300                 305

agc cca gtg gct ctg cag agg gct atc gag gca ctt cca cct ggg aat    1438
Ser Pro Val Ala Leu Gln Arg Ala Ile Glu Ala Leu Pro Pro Gly Asn
            310                 315                 320

ttt aaa gtt tct ctt cct gat gga gcc gaa ggg agt ggg aca gat cac    1486
Phe Lys Val Ser Leu Pro Asp Gly Ala Glu Gly Ser Gly Thr Asp His
        325                 330                 335

agg tct tcc agt tct cat tcc cct ggc tcc cca ccg aga aac cag gtc    1534
Arg Ser Ser Ser Ser His Ser Pro Gly Ser Pro Pro Arg Asn Gln Val
    340                 345                 350

cag ggc aca tgc agt acc act ctg att gcc att gcc ggc atg acc tgt    1582
Gln Gly Thr Cys Ser Thr Thr Leu Ile Ala Ile Ala Gly Met Thr Cys
355                 360                 365                 370

gca tcc tgt gtc cat tcc att gaa ggc atg atc tcc caa ctg gaa ggg    1630
Ala Ser Cys Val His Ser Ile Glu Gly Met Ile Ser Gln Leu Glu Gly
                375                 380                 385

gtg cag caa ata tcg gtg tct ttg gcc gaa ggg act gca aca gtt ctt    1678
Val Gln Gln Ile Ser Val Ser Leu Ala Glu Gly Thr Ala Thr Val Leu
            390                 395                 400

tat aat ccc tct gta att agc cca gaa gaa ctc aga gct gct ata gaa    1726
Tyr Asn Pro Ser Val Ile Ser Pro Glu Glu Leu Arg Ala Ala Ile Glu
        405                 410                 415

gac atg gga ttt gag gct tca gtc gtt tct gaa agc tgt tct act aac    1774
Asp Met Gly Phe Glu Ala Ser Val Val Ser Glu Ser Cys Ser Thr Asn
    420                 425                 430

cct ctt gga aac cac agt gct ggg aat tcc atg gtg caa act aca gat    1822
Pro Leu Gly Asn His Ser Ala Gly Asn Ser Met Val Gln Thr Thr Asp
435                 440                 445                 450

ggt aca cct aca tct gtg cag gaa gtg gct ccc cac act ggg agg ctc    1870
Gly Thr Pro Thr Ser Val Gln Glu Val Ala Pro His Thr Gly Arg Leu
                455                 460                 465

cct gca aac cat gcc ccg gac atc ttg gca aag tcc cca caa tca acc    1918
Pro Ala Asn His Ala Pro Asp Ile Leu Ala Lys Ser Pro Gln Ser Thr
            470                 475                 480

aga gca gtg gca ccg cag aag tgc ttc tta cag atc aaa ggc atg acc    1966
Arg Ala Val Ala Pro Gln Lys Cys Phe Leu Gln Ile Lys Gly Met Thr
        485                 490                 495

tgt gca tcc tgt gtg tct aac ata gaa agg aat ctg cag aaa gaa gct    2014
Cys Ala Ser Cys Val Ser Asn Ile Glu Arg Asn Leu Gln Lys Glu Ala
    500                 505                 510

ggt gtt ctc tcc gtg ttg gtt gcc ttg atg gca gga aag gca gag atc    2062
Gly Val Leu Ser Val Leu Val Ala Leu Met Ala Gly Lys Ala Glu Ile
515                 520                 525                 530

aag tat gac cca gag gtc atc cag ccc ctc gag ata gct cag ttc atc    2110
Lys Tyr Asp Pro Glu Val Ile Gln Pro Leu Glu Ile Ala Gln Phe Ile
                535                 540                 545

cag gac ctg ggt ttt gag gca gca gtc atg gag gac tac gca ggc tcc    2158
Gln Asp Leu Gly Phe Glu Ala Ala Val Met Glu Asp Tyr Ala Gly Ser
            550                 555                 560
```

```
gat ggc aac att gag ctg aca atc aca ggg atg acc tgc gcg tcc tgt    2206
Asp Gly Asn Ile Glu Leu Thr Ile Thr Gly Met Thr Cys Ala Ser Cys
        565                 570                 575

gtc cac aac ata gag tcc aaa ctc acg agg aca aat ggc atc act tat    2254
Val His Asn Ile Glu Ser Lys Leu Thr Arg Thr Asn Gly Ile Thr Tyr
    580                 585                 590

gcc tcc gtt gcc ctt gcc acc agc aaa gcc ctt gtt aag ttt gac ccg    2302
Ala Ser Val Ala Leu Ala Thr Ser Lys Ala Leu Val Lys Phe Asp Pro
595                 600                 605                 610

gaa att atc ggt cca cgg gat att atc aaa att att gag gaa att ggc    2350
Glu Ile Ile Gly Pro Arg Asp Ile Ile Lys Ile Ile Glu Glu Ile Gly
                615                 620                 625

ttt cat gct tcc ctg gcc cag aga aac ccc aac gct cat cac ttg gac    2398
Phe His Ala Ser Leu Ala Gln Arg Asn Pro Asn Ala His His Leu Asp
            630                 635                 640

cac aag atg gaa ata aag cag tgg aag aag tct ttc ctg tgc agc ctg    2446
His Lys Met Glu Ile Lys Gln Trp Lys Lys Ser Phe Leu Cys Ser Leu
        645                 650                 655

gtg ttt ggc atc cct gtc atg gcc tta atg atc tat atg ctg ata ccc    2494
Val Phe Gly Ile Pro Val Met Ala Leu Met Ile Tyr Met Leu Ile Pro
    660                 665                 670

agc aac gag ccc cac cag tcc atg gtc ctg gac cac aac atc att cca    2542
Ser Asn Glu Pro His Gln Ser Met Val Leu Asp His Asn Ile Ile Pro
675                 680                 685                 690

gga ctg tcc att cta aat ctc atc ttc ttt atc ttg tgt acc ttt gtc    2590
Gly Leu Ser Ile Leu Asn Leu Ile Phe Phe Ile Leu Cys Thr Phe Val
                695                 700                 705

cag ctc ctc ggt ggg tgg tac ttc tac gtt cag gcc tac aaa tct ctg    2638
Gln Leu Leu Gly Gly Trp Tyr Phe Tyr Val Gln Ala Tyr Lys Ser Leu
            710                 715                 720

aga cac agg tca gcc aac atg gac gtg ctc atc gtc ctg gcc aca agc    2686
Arg His Arg Ser Ala Asn Met Asp Val Leu Ile Val Leu Ala Thr Ser
        725                 730                 735

att gct tat gtt tat tct ctg gtc atc ctg gtg gtt gct gtg gct gag    2734
Ile Ala Tyr Val Tyr Ser Leu Val Ile Leu Val Val Ala Val Ala Glu
    740                 745                 750

aag gcg gag agg agc cct gtg aca ttc ttc gac acg ccc ccc atg ctc    2782
Lys Ala Glu Arg Ser Pro Val Thr Phe Phe Asp Thr Pro Pro Met Leu
755                 760                 765                 770

ttt gtg ttc att gcc ctg ggc cgg tgg ctg gaa cac ttg gca aag agc    2830
Phe Val Phe Ile Ala Leu Gly Arg Trp Leu Glu His Leu Ala Lys Ser
                775                 780                 785

aaa acc tca gaa gcc ctg gct aaa ctc atg tct ctc caa gcc aca gaa    2878
Lys Thr Ser Glu Ala Leu Ala Lys Leu Met Ser Leu Gln Ala Thr Glu
            790                 795                 800

gcc acc gtt gtg acc ctt ggt gag gac aat tta atc atc agg gag gag    2926
Ala Thr Val Val Thr Leu Gly Glu Asp Asn Leu Ile Ile Arg Glu Glu
        805                 810                 815

caa gtc ccc atg gag ctg gtg cag cgg ggc gat atc gtc aag gtg gtc    2974
Gln Val Pro Met Glu Leu Val Gln Arg Gly Asp Ile Val Lys Val Val
    820                 825                 830
```

```
cct ggg gga aag ttt cca gtg gat ggg aaa gtc ctg gaa ggc aat acc    3022
Pro Gly Gly Lys Phe Pro Val Asp Gly Lys Val Leu Glu Gly Asn Thr
835                 840                 845                 850

atg gct gat gag tcc ctc atc aca gga gaa gcc atg cca gtc act aag    3070
Met Ala Asp Glu Ser Leu Ile Thr Gly Glu Ala Met Pro Val Thr Lys
                855                 860                 865

aaa ccc gga agc act gta att gcg ggg tct ata aat gca cat ggc tct    3118
Lys Pro Gly Ser Thr Val Ile Ala Gly Ser Ile Asn Ala His Gly Ser
            870                 875                 880

gtg ctc att aaa gct acc cac gtg ggc aat gac acc act ttg gct cag    3166
Val Leu Ile Lys Ala Thr His Val Gly Asn Asp Thr Thr Leu Ala Gln
        885                 890                 895

att gtg aaa ctg gtg gaa gag gct cag atg tca aag gca ccc att cag    3214
Ile Val Lys Leu Val Glu Glu Ala Gln Met Ser Lys Ala Pro Ile Gln
    900                 905                 910

cag ctg gct gac cgg ttt agt gga tat ttt gtc cca ttt atc atc atc    3262
Gln Leu Ala Asp Arg Phe Ser Gly Tyr Phe Val Pro Phe Ile Ile Ile
915                 920                 925                 930

atg tca act ttg acg ttg gtg gta tgg att gta atc ggt ttt atc gat    3310
Met Ser Thr Leu Thr Leu Val Val Trp Ile Val Ile Gly Phe Ile Asp
                935                 940                 945

ttt ggt gtt gtt cag aga tac ttt cct aac ccc aac aag cac atc tcc    3358
Phe Gly Val Val Gln Arg Tyr Phe Pro Asn Pro Asn Lys His Ile Ser
            950                 955                 960

cag aca gag gtg atc atc cgg ttt gct ttc cag acg tcc atc acg gtg    3406
Gln Thr Glu Val Ile Ile Arg Phe Ala Phe Gln Thr Ser Ile Thr Val
        965                 970                 975

ctg tgc att gcc tgc ccc tgc tcc ctg ggg ctg gcc acg ccc acg gct    3454
Leu Cys Ile Ala Cys Pro Cys Ser Leu Gly Leu Ala Thr Pro Thr Ala
    980                 985                 990

gtc atg gtg ggc acc ggg gtg gcc gcg cag aac ggc atc ctc atc aag    3502
Val Met Val Gly Thr Gly Val Ala Ala Gln Asn Gly Ile Leu Ile Lys
995                 1000                1005                1010

gga ggc aag ccc ctg gag atg gcg cac aag ata aag act gtg atg ttt    3550
Gly Gly Lys Pro Leu Glu Met Ala His Lys Ile Lys Thr Val Met Phe
                1015                1020                1025

gac aag act ggc acc att acc cat ggc gtc ccc agg gtc atg cgg gtg    3598
Asp Lys Thr Gly Thr Ile Thr His Gly Val Pro Arg Val Met Arg Val
            1030                1035                1040

ctc ctg ctg ggg gat gtg gcc aca ctg ccc ctc agg aag gtt ctg gct    3646
Leu Leu Leu Gly Asp Val Ala Thr Leu Pro Leu Arg Lys Val Leu Ala
        1045                1050                1055

gtg gtg ggg act gcg gag gcc agc agt gaa cac ccc ttg ggc gtg gca    3694
Val Val Gly Thr Ala Glu Ala Ser Ser Glu His Pro Leu Gly Val Ala
    1060                1065                1070

gtc acc aaa tac tgt aaa gag gaa ctt gga aca gag acc ttg gga tac    3742
Val Thr Lys Tyr Cys Lys Glu Glu Leu Gly Thr Glu Thr Leu Gly Tyr
1075                1080                1085                1090

tgc acg gac ttc cag gca gtg cca ggc tgt gga att ggg tgc aaa gtc    3790
Cys Thr Asp Phe Gln Ala Val Pro Gly Cys Gly Ile Gly Cys Lys Val
                1095                1100                1105
```

```
agc aac gtg gaa ggc atc ctg gcc cac agt gag cgc cct ttg agt gca    3838
Ser Asn Val Glu Gly Ile Leu Ala His Ser Glu Arg Pro Leu Ser Ala
            1110                1115                1120

ccg gcc agt cac ctg aat gag gct ggc agc ctt ccc gca gaa aaa gat    3886
Pro Ala Ser His Leu Asn Glu Ala Gly Ser Leu Pro Ala Glu Lys Asp
        1125                1130                1135

gca gtc ccc cag acc ttc tct gtg ctg att gga aac cgt gag tgg ctg    3934
Ala Val Pro Gln Thr Phe Ser Val Leu Ile Gly Asn Arg Glu Trp Leu
    1140                1145                1150

agg cgc aac ggt tta acc att tct agc gat gtc agt gac gct atg aca    3982
Arg Arg Asn Gly Leu Thr Ile Ser Ser Asp Val Ser Asp Ala Met Thr
1155                1160                1165                1170

gac cac gag atg aaa gga cag aca gcc atc ctg gtg gct att gac ggt    4030
Asp His Glu Met Lys Gly Gln Thr Ala Ile Leu Val Ala Ile Asp Gly
                1175                1180                1185

gtg ctc tgt ggg atg atc gca atc gca gac gct gtc aag cag gag gct    4078
Val Leu Cys Gly Met Ile Ala Ile Ala Asp Ala Val Lys Gln Glu Ala
            1190                1195                1200

gcc ctg gct gtg cac acg ctg cag agc atg ggt gtg gac gtg gtt ctg    4126
Ala Leu Ala Val His Thr Leu Gln Ser Met Gly Val Asp Val Val Leu
        1205                1210                1215

atc acg ggg gac aac cgg aag aca gcc aga gct att gcc acc cag gtt    4174
Ile Thr Gly Asp Asn Arg Lys Thr Ala Arg Ala Ile Ala Thr Gln Val
    1220                1225                1230

ggc atc aac aaa gtc ttt gca gag gtg ctg cct tcg cac aag gtg gcc    4222
Gly Ile Asn Lys Val Phe Ala Glu Val Leu Pro Ser His Lys Val Ala
1235                1240                1245                1250

aag gtc cag gag ctc cag aat aaa ggg aag aaa gtc gcc atg gtg ggg    4270
Lys Val Gln Glu Leu Gln Asn Lys Gly Lys Lys Val Ala Met Val Gly
                1255                1260                1265

gat ggg gtc aat gac tcc ccg gcc ttg gcc cag gca gac atg ggt gtg    4318
Asp Gly Val Asn Asp Ser Pro Ala Leu Ala Gln Ala Asp Met Gly Val
            1270                1275                1280

gcc att ggc acc ggc acg gat gtg gcc atc gag gca gcc gac gtc gtc    4366
Ala Ile Gly Thr Gly Thr Asp Val Ala Ile Glu Ala Ala Asp Val Val
        1285                1290                1295

ctt atc aga aat gat ttg ctg gat gtg gtg gct agc att cac ctt tcc    4414
Leu Ile Arg Asn Asp Leu Leu Asp Val Val Ala Ser Ile His Leu Ser
    1300                1305                1310

aag agg act gtc cga agg ata cgc atc aac ctg gtc ctg gca ctg att    4462
Lys Arg Thr Val Arg Arg Ile Arg Ile Asn Leu Val Leu Ala Leu Ile
1315                1320                1325                1330

tat aac ctg gtt ggg ata ccc att gca gca ggt gtc ttc atg ccc atc    4510
Tyr Asn Leu Val Gly Ile Pro Ile Ala Ala Gly Val Phe Met Pro Ile
                1335                1340                1345

ggc att gtg ctg cag ccc tgg atg ggc tca gcg gcc atg gca gcc tcc    4558
Gly Ile Val Leu Gln Pro Trp Met Gly Ser Ala Ala Met Ala Ala Ser
            1350                1355                1360

tct gtg tct gtg gtg ctc tca tcc ctg cag ctc aag tgc tat aag aag    4606
Ser Val Ser Val Val Leu Ser Ser Leu Gln Leu Lys Cys Tyr Lys Lys
        1365                1370                1375
```

```
cct gac ctg gag agg tat gag gca cag gcg cat ggc cac atg aag ccc    4654
Pro Asp Leu Glu Arg Tyr Glu Ala Gln Ala His Gly His Met Lys Pro
    1380                1385                1390

ctg acg gca tcc cag gtc agt gtg cac ata ggc atg gat gac agg tgg    4702
Leu Thr Ala Ser Gln Val Ser Val His Ile Gly Met Asp Asp Arg Trp
1395                1400                1405                1410

cgg gac tcc ccc agg gcc aca cca tgg gac cag gtc agc tat gtc agc    4750
Arg Asp Ser Pro Arg Ala Thr Pro Trp Asp Gln Val Ser Tyr Val Ser
                1415                1420                1425

cag gtg tcg ctg tcc tcc ctg acg tcc gac aag cca tct cgg cac agc    4798
Gln Val Ser Leu Ser Ser Leu Thr Ser Asp Lys Pro Ser Arg His Ser
            1430                1435                1440

gct gca gca gac gat gat ggg gac aag tgg tct ctg ctc ctg aat ggc    4846
Ala Ala Ala Asp Asp Asp Gly Asp Lys Trp Ser Leu Leu Leu Asn Gly
        1445                1450                1455

agg gat gag gag cag tac atc tga ggtaccaata aagacctctt attttcattc   4900
Arg Asp Glu Glu Gln Tyr Ile
    1460                1465

atcaggtgtg gttggttttt ttgtgtgggg gcggatccat cggatcccgt gcggaccgag  4960

cggccgcagg aacccctagt gatggagttg gccactccct ctctgcgcgc tcgctcgctc  5020

actgaggccg ggcgaccaaa ggtcgcccga cgcccgggct ttgcccgggc ggcctcagtg  5080

agcgagcgag cgcgcagctg cctgcag                                      5107
```

<210> 2
<211> 1465
<212> PRT
<213> Artificial Sequence

<220>
<223> [CDS]:473..4870 from SEQ ID NO 1

<400> 2

```
Met Pro Glu Gln Glu Arg Gln Ile Thr Ala Arg Glu Gly Ala Ser Arg
1               5                  10                  15
Lys Ile Leu Ser Lys Leu Ser Leu Pro Thr Arg Ala Trp Glu Pro Ala
            20                  25                  30
Met Lys Lys Ser Phe Ala Phe Asp Asn Val Gly Tyr Glu Gly Gly Leu
        35                  40                  45
Asp Gly Leu Gly Pro Ser Ser Gln Val Ala Thr Ser Thr Val Arg Ile
    50                  55                  60
Leu Gly Met Thr Cys Gln Ser Cys Val Lys Ser Ile Glu Asp Arg Ile
65                  70                  75                  80
Ser Asn Leu Lys Gly Ile Ile Ser Met Lys Val Ser Leu Glu Gln Gly
                85                  90                  95
Ser Ala Thr Val Lys Tyr Val Pro Ser Val Val Cys Leu Gln Gln Val
            100                 105                 110
Cys His Gln Ile Gly Asp Met Gly Phe Glu Ala Ser Ile Ala Glu Gly
        115                 120                 125
Lys Ala Ala Ser Trp Pro Ser Arg Ser Leu Pro Ala Gln Glu Ala Val
    130                 135                 140
Val Lys Leu Arg Val Glu Gly Met Thr Cys Gln Ser Cys Val Ser Ser
145                 150                 155                 160
Ile Glu Gly Lys Val Arg Lys Leu Gln Gly Val Val Arg Val Lys Val
                165                 170                 175
Ser Leu Ser Asn Gln Glu Ala Val Ile Thr Tyr Gln Pro Tyr Leu Ile
            180                 185                 190
Gln Pro Glu Asp Leu Arg Asp His Val Asn Asp Met Gly Phe Glu Ala
        195                 200                 205
```

```
Ala Ile Lys Ser Lys Val Ala Pro Leu Ser Leu Gly Pro Ile Asp Ile
    210                 215                 220
Glu Arg Leu Gln Ser Thr Asn Pro Lys Arg Pro Leu Ser Ser Ala Asn
225                 230                 235                 240
Gln Asn Phe Asn Asn Ser Glu Thr Leu Gly His Gln Gly Ser His Val
                245                 250                 255
Val Thr Leu Gln Leu Arg Ile Asp Gly Met His Cys Lys Ser Cys Val
            260                 265                 270
Leu Asn Ile Glu Glu Asn Ile Gly Gln Leu Leu Gly Val Gln Ser Ile
        275                 280                 285
Gln Val Ser Leu Glu Asn Lys Thr Ala Gln Val Lys Tyr Asp Pro Ser
    290                 295                 300
Cys Thr Ser Pro Val Ala Leu Gln Arg Ala Ile Glu Ala Leu Pro Pro
305                 310                 315                 320
Gly Asn Phe Lys Val Ser Leu Pro Asp Gly Ala Glu Gly Ser Gly Thr
                325                 330                 335
Asp His Arg Ser Ser Ser Ser His Ser Pro Gly Ser Pro Pro Arg Asn
            340                 345                 350
Gln Val Gln Gly Thr Cys Ser Thr Thr Leu Ile Ala Ile Ala Gly Met
        355                 360                 365
Thr Cys Ala Ser Cys Val His Ser Ile Glu Gly Met Ile Ser Gln Leu
    370                 375                 380
Glu Gly Val Gln Gln Ile Ser Val Ser Leu Ala Glu Gly Thr Ala Thr
385                 390                 395                 400
Val Leu Tyr Asn Pro Ser Val Ile Ser Pro Glu Glu Leu Arg Ala Ala
                405                 410                 415
Ile Glu Asp Met Gly Phe Glu Ala Ser Val Val Ser Glu Ser Cys Ser
            420                 425                 430
Thr Asn Pro Leu Gly Asn His Ser Ala Gly Asn Ser Met Val Gln Thr
        435                 440                 445
Thr Asp Gly Thr Pro Thr Ser Val Gln Glu Val Ala Pro His Thr Gly
    450                 455                 460
Arg Leu Pro Ala Asn His Ala Pro Asp Ile Leu Ala Lys Ser Pro Gln
465                 470                 475                 480
Ser Thr Arg Ala Val Ala Pro Gln Lys Cys Phe Leu Gln Ile Lys Gly
                485                 490                 495
Met Thr Cys Ala Ser Cys Val Ser Asn Ile Glu Arg Asn Leu Gln Lys
            500                 505                 510
Glu Ala Gly Val Leu Ser Val Leu Val Ala Leu Met Ala Gly Lys Ala
        515                 520                 525
Glu Ile Lys Tyr Asp Pro Glu Val Ile Gln Pro Leu Glu Ile Ala Gln
    530                 535                 540
Phe Ile Gln Asp Leu Gly Phe Glu Ala Ala Val Met Glu Asp Tyr Ala
545                 550                 555                 560
Gly Ser Asp Gly Asn Ile Glu Leu Thr Ile Thr Gly Met Thr Cys Ala
                565                 570                 575
Ser Cys Val His Asn Ile Glu Ser Lys Leu Thr Arg Thr Asn Gly Ile
            580                 585                 590
Thr Tyr Ala Ser Val Ala Leu Ala Thr Ser Lys Ala Leu Val Lys Phe
        595                 600                 605
Asp Pro Glu Ile Ile Gly Pro Arg Asp Ile Ile Lys Ile Ile Glu Glu
    610                 615                 620
Ile Gly Phe His Ala Ser Leu Ala Gln Arg Asn Pro Asn Ala His His
625                 630                 635                 640
Leu Asp His Lys Met Glu Ile Lys Gln Trp Lys Lys Ser Phe Leu Cys
                645                 650                 655
Ser Leu Val Phe Gly Ile Pro Val Met Ala Leu Met Ile Tyr Met Leu
            660                 665                 670
Ile Pro Ser Asn Glu Pro His Gln Ser Met Val Leu Asp His Asn Ile
        675                 680                 685
Ile Pro Gly Leu Ser Ile Leu Asn Leu Ile Phe Phe Ile Leu Cys Thr
    690                 695                 700
Phe Val Gln Leu Leu Gly Gly Trp Tyr Phe Tyr Val Gln Ala Tyr Lys
705                 710                 715                 720
Ser Leu Arg His Arg Ser Ala Asn Met Asp Val Leu Ile Val Leu Ala
                725                 730                 735
Thr Ser Ile Ala Tyr Val Tyr Ser Leu Val Ile Leu Val Val Ala Val
            740                 745                 750
```

```
Ala Glu Lys Ala Glu Arg Ser Pro Val Thr Phe Phe Asp Thr Pro Pro
        755                 760                 765
Met Leu Phe Val Phe Ile Ala Leu Gly Arg Trp Leu Glu His Leu Ala
    770                 775                 780
Lys Ser Lys Thr Ser Glu Ala Leu Ala Lys Leu Met Ser Leu Gln Ala
785                 790                 795                 800
Thr Glu Ala Thr Val Val Thr Leu Gly Glu Asp Asn Leu Ile Ile Arg
                805                 810                 815
Glu Glu Gln Val Pro Met Glu Leu Val Gln Arg Gly Asp Ile Val Lys
            820                 825                 830
Val Val Pro Gly Gly Lys Phe Pro Val Asp Gly Lys Val Leu Glu Gly
        835                 840                 845
Asn Thr Met Ala Asp Glu Ser Leu Ile Thr Gly Glu Ala Met Pro Val
    850                 855                 860
Thr Lys Lys Pro Gly Ser Thr Val Ile Ala Gly Ser Ile Asn Ala His
865                 870                 875                 880
Gly Ser Val Leu Ile Lys Ala Thr His Val Gly Asn Asp Thr Thr Leu
                885                 890                 895
Ala Gln Ile Val Lys Leu Val Glu Glu Ala Gln Met Ser Lys Ala Pro
            900                 905                 910
Ile Gln Gln Leu Ala Asp Arg Phe Ser Gly Tyr Phe Val Pro Phe Ile
        915                 920                 925
Ile Ile Met Ser Thr Leu Thr Leu Val Val Trp Ile Val Ile Gly Phe
    930                 935                 940
Ile Asp Phe Gly Val Val Gln Arg Tyr Phe Pro Asn Pro Asn Lys His
945                 950                 955                 960
Ile Ser Gln Thr Glu Val Ile Ile Arg Phe Ala Phe Gln Thr Ser Ile
                965                 970                 975
Thr Val Leu Cys Ile Ala Cys Pro Cys Ser Leu Gly Leu Ala Thr Pro
            980                 985                 990
Thr Ala Val Met Val Gly Thr Gly Val Ala Ala Gln Asn Gly Ile Leu
        995                 1000                1005
Ile Lys Gly Gly Lys Pro Leu Glu Met Ala His Lys Ile Lys Thr Val
    1010                1015                1020
Met Phe Asp Lys Thr Gly Thr Ile Thr His Gly Val Pro Arg Val Met
1025                1030                1035                1040
Arg Val Leu Leu Leu Gly Asp Val Ala Thr Leu Pro Leu Arg Lys Val
                1045                1050                1055
Leu Ala Val Val Gly Thr Ala Glu Ala Ser Ser Glu His Pro Leu Gly
            1060                1065                1070
Val Ala Val Thr Lys Tyr Cys Lys Glu Glu Leu Gly Thr Glu Thr Leu
        1075                1080                1085
Gly Tyr Cys Thr Asp Phe Gln Ala Val Pro Gly Cys Gly Ile Gly Cys
    1090                1095                1100
Lys Val Ser Asn Val Glu Gly Ile Leu Ala His Ser Glu Arg Pro Leu
1105                1110                1115                1120
Ser Ala Pro Ala Ser His Leu Asn Glu Ala Gly Ser Leu Pro Ala Glu
                1125                1130                1135
Lys Asp Ala Val Pro Gln Thr Phe Ser Val Leu Ile Gly Asn Arg Glu
            1140                1145                1150
Trp Leu Arg Arg Asn Gly Leu Thr Ile Ser Ser Asp Val Ser Asp Ala
        1155                1160                1165
Met Thr Asp His Glu Met Lys Gly Gln Thr Ala Ile Leu Val Ala Ile
    1170                1175                1180
Asp Gly Val Leu Cys Gly Met Ile Ala Ile Ala Asp Ala Val Lys Gln
1185                1190                1195                1200
Glu Ala Ala Leu Ala Val His Thr Leu Gln Ser Met Gly Val Asp Val
                1205                1210                1215
Val Leu Ile Thr Gly Asp Asn Arg Lys Thr Ala Arg Ala Ile Ala Thr
            1220                1225                1230
Gln Val Gly Ile Asn Lys Val Phe Ala Glu Val Leu Pro Ser His Lys
        1235                1240                1245
Val Ala Lys Val Gln Glu Leu Gln Asn Lys Gly Lys Lys Val Ala Met
    1250                1255                1260
Val Gly Asp Gly Val Asn Asp Ser Pro Ala Leu Ala Gln Ala Asp Met
1265                1270                1275                1280
Gly Val Ala Ile Gly Thr Gly Thr Asp Val Ala Ile Glu Ala Ala Asp
                1285                1290                1295
```

Val Val Leu Ile Arg Asn Asp Leu Leu Asp Val Val Ala Ser Ile His
1300 1305 1310
Leu Ser Lys Arg Thr Val Arg Arg Ile Arg Ile Asn Leu Val Leu Ala
1315 1320 1325
Leu Ile Tyr Asn Leu Val Gly Ile Pro Ile Ala Ala Gly Val Phe Met
1330 1335 1340
Pro Ile Gly Ile Val Leu Gln Pro Trp Met Gly Ser Ala Ala Met Ala
1345 1350 1355 1360
Ala Ser Ser Val Ser Val Val Leu Ser Ser Leu Gln Leu Lys Cys Tyr
1365 1370 1375
Lys Lys Pro Asp Leu Glu Arg Tyr Glu Ala Gln Ala His Gly His Met
1380 1385 1390
Lys Pro Leu Thr Ala Ser Gln Val Ser Val His Ile Gly Met Asp Asp
1395 1400 1405
Arg Trp Arg Asp Ser Pro Arg Ala Thr Pro Trp Asp Gln Val Ser Tyr
1410 1415 1420
Val Ser Gln Val Ser Leu Ser Ser Leu Thr Ser Asp Lys Pro Ser Arg
1425 1430 1435 1440
His Ser Ala Ala Asp Asp Asp Gly Asp Lys Trp Ser Leu Leu Leu
1445 1450 1455
Asn Gly Arg Asp Glu Glu Gln Tyr Ile
1460 1465

<210> 3
<211> 5107
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleic acid construct of expression vector AAV2-AAT-coATP7B

<220>
<221> repeat_region
<222> 1..141
<223> /note="5' ITR of adeno-associated virus serotype 2"

<220>
<221> promoter
<222> 156..460
<223> /note="alpha 1 antitrypsin"

<220>
<221> CDS
<222> 473..4870
<223> /note="Codon optimized sequence encoding ATP7B"
/transl_table=1

<220>
<221> polyA_signal
<222> 4877..4932

<220>
<221> repeat_region
<222> 4968..5107
<223> /note="3' ITR of adeno-associated virus serotype 2"

<400> 3
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60

gggcgacctt tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120

actccatcac taggggttcc tgcggccgca cgcgtcgcca ccccctccac cttggacaca 180

ggacgctgtg gtttctgagc caggtacaat gactcctttc ggtaagtgca gtggaagctg 240

tacactgccc aggcaaagcg tccgggcagc gtaggcgggc gactcagatc ccagccagtg 300

gacttagccc ctgtttgctc ctccgataac tggggtgacc ttggttaata ttcaccagca 360

```
gcctcccccg ttgcccctct ggatccactg cttaaatacg gacgaggaca gggccctgtc   420

tcctcagctt caggcaccac cactgacctg ggacagtgaa gcggccgcca cc atg cca   478
                                                           Met Pro
                                                           1

gaa cag gaa cgc cag atc aca gca aga gag gga gca agt cgg aaa atc     526
Glu Gln Glu Arg Gln Ile Thr Ala Arg Glu Gly Ala Ser Arg Lys Ile
        5                   10                  15

ctg agc aaa ctg agc ctg cca acc aga gca tgg gaa ccc gca atg aag     574
Leu Ser Lys Leu Ser Leu Pro Thr Arg Ala Trp Glu Pro Ala Met Lys
    20                  25                  30

aaa agc ttc gcc ttt gac aac gtg gga tac gag gga ggg ctg gat gga     622
Lys Ser Phe Ala Phe Asp Asn Val Gly Tyr Glu Gly Gly Leu Asp Gly
35                  40                  45                  50

ctg gga cct agc tcc cag gtg gcc acc tct aca gtc cga atc ctg ggc     670
Leu Gly Pro Ser Ser Gln Val Ala Thr Ser Thr Val Arg Ile Leu Gly
                55                  60                  65

atg act tgc cag agt tgc gtg aaa tca att gaa gac cgg atc agt aat     718
Met Thr Cys Gln Ser Cys Val Lys Ser Ile Glu Asp Arg Ile Ser Asn
            70                  75                  80

ctg aag gga atc att agc atg aaa gtg tcc ctg gag cag ggc tca gcc     766
Leu Lys Gly Ile Ile Ser Met Lys Val Ser Leu Glu Gln Gly Ser Ala
        85                  90                  95

acc gtg aag tat gtc cct agc gtg gtc tgc ctg cag cag gtg tgc cac     814
Thr Val Lys Tyr Val Pro Ser Val Val Cys Leu Gln Gln Val Cys His
    100                 105                 110

cag atc ggc gat atg ggg ttc gag gcc tcc att gct gaa ggg aaa gcc     862
Gln Ile Gly Asp Met Gly Phe Glu Ala Ser Ile Ala Glu Gly Lys Ala
115                 120                 125                 130

gct tct tgg cct agc cgg tcc ctg cca gca cag gaa gca gtg gtc aag     910
Ala Ser Trp Pro Ser Arg Ser Leu Pro Ala Gln Glu Ala Val Val Lys
                135                 140                 145

ctg aga gtg gag gga atg aca tgc cag agc tgc gtg agc agt atc gaa     958
Leu Arg Val Glu Gly Met Thr Cys Gln Ser Cys Val Ser Ser Ile Glu
            150                 155                 160

gga aag gtc cga aaa ctg cag ggc gtg gtc cgg gtg aag gtc tct ctg    1006
Gly Lys Val Arg Lys Leu Gln Gly Val Val Arg Val Lys Val Ser Leu
        165                 170                 175

agt aac cag gag gcc gtg att acc tac cag ccc tat ctg atc cag cct    1054
Ser Asn Gln Glu Ala Val Ile Thr Tyr Gln Pro Tyr Leu Ile Gln Pro
    180                 185                 190

gaa gac ctg agg gat cac gtg aat gac atg ggc ttc gag gca gcc atc    1102
Glu Asp Leu Arg Asp His Val Asn Asp Met Gly Phe Glu Ala Ala Ile
195                 200                 205                 210

aag tcc aaa gtg gcc cca ctg tct ctg ggg ccc att gat atc gaa aga    1150
Lys Ser Lys Val Ala Pro Leu Ser Leu Gly Pro Ile Asp Ile Glu Arg
                215                 220                 225

ctg cag tcc acc aac cca aag agg ccc ctg tca agc gcc aac cag aac    1198
Leu Gln Ser Thr Asn Pro Lys Arg Pro Leu Ser Ser Ala Asn Gln Asn
            230                 235                 240

ttc aac aat agt gag acc ctg gga cac cag ggc tca cat gtg gtc aca    1246
```

```
Phe Asn Asn Ser Glu Thr Leu Gly His Gln Gly Ser His Val Val Thr
        245                 250                 255

ctg cag ctg agg att gac ggc atg cac tgc aag tct tgc gtg ctg aac    1294
Leu Gln Leu Arg Ile Asp Gly Met His Cys Lys Ser Cys Val Leu Asn
    260                 265                 270

att gag gaa aat atc ggc cag ctg ctg ggg gtg cag tct atc cag gtc    1342
Ile Glu Glu Asn Ile Gly Gln Leu Leu Gly Val Gln Ser Ile Gln Val
275                 280                 285                 290

agt ctg gag aac aag act gct cag gtg aaa tac gat cct tca tgc acc    1390
Ser Leu Glu Asn Lys Thr Ala Gln Val Lys Tyr Asp Pro Ser Cys Thr
                295                 300                 305

agc cca gtg gca ctg cag cgc gct atc gaa gca ctg ccc cct gga aat    1438
Ser Pro Val Ala Leu Gln Arg Ala Ile Glu Ala Leu Pro Pro Gly Asn
            310                 315                 320

ttc aag gtg agc ctg cct gac gga gca gag gga tcc gga acc gat cac    1486
Phe Lys Val Ser Leu Pro Asp Gly Ala Glu Gly Ser Gly Thr Asp His
        325                 330                 335

agg tcc tct agt tca cat tcc cca ggg tct cca cca cga aac cag gtg    1534
Arg Ser Ser Ser Ser His Ser Pro Gly Ser Pro Pro Arg Asn Gln Val
    340                 345                 350

cag gga aca tgt tcc acc aca ctg att gca atc gcc ggc atg act tgc    1582
Gln Gly Thr Cys Ser Thr Thr Leu Ile Ala Ile Ala Gly Met Thr Cys
355                 360                 365                 370

gcc tca tgc gtg cac agc att gaa ggg atg atc tct cag ctg gag gga    1630
Ala Ser Cys Val His Ser Ile Glu Gly Met Ile Ser Gln Leu Glu Gly
                375                 380                 385

gtg cag cag atc tca gtc agc ctg gcc gag ggc act gct acc gtg ctg    1678
Val Gln Gln Ile Ser Val Ser Leu Ala Glu Gly Thr Ala Thr Val Leu
            390                 395                 400

tac aat ccc agt gtc atc tca cct gag gaa ctg cgg gct gca att gag    1726
Tyr Asn Pro Ser Val Ile Ser Pro Glu Glu Leu Arg Ala Ala Ile Glu
        405                 410                 415

gac atg ggg ttc gaa gct tcc gtg gtc tcc gaa tct tgc agt acc aac    1774
Asp Met Gly Phe Glu Ala Ser Val Val Ser Glu Ser Cys Ser Thr Asn
    420                 425                 430

ccc ctg ggg aat cat tcc gcc gga aac tct atg gtg cag act acc gac    1822
Pro Leu Gly Asn His Ser Ala Gly Asn Ser Met Val Gln Thr Thr Asp
435                 440                 445                 450

ggg aca cct act tct gtg cag gag gtc gca cca cac aca gga cgc ctg    1870
Gly Thr Pro Thr Ser Val Gln Glu Val Ala Pro His Thr Gly Arg Leu
                455                 460                 465

cca gcc aat cat gct ccc gat atc ctg gcc aaa agc ccc cag tcc act    1918
Pro Ala Asn His Ala Pro Asp Ile Leu Ala Lys Ser Pro Gln Ser Thr
            470                 475                 480

cga gct gtg gca cct cag aag tgt ttt ctg cag atc aaa ggc atg acc    1966
Arg Ala Val Ala Pro Gln Lys Cys Phe Leu Gln Ile Lys Gly Met Thr
        485                 490                 495

tgc gcc tct tgc gtg agc aac att gag cgg aat ctg cag aag gaa gct    2014
Cys Ala Ser Cys Val Ser Asn Ile Glu Arg Asn Leu Gln Lys Glu Ala
    500                 505                 510

ggg gtg ctg agc gtg ctg gtc gca ctg atg gcc gga aag gct gag atc    2062
```

```
Gly Val Leu Ser Val Leu Val Ala Leu Met Ala Gly Lys Ala Glu Ile
515                 520                 525                 530

aag tac gac cct gaa gtg atc cag cca ctg gag att gcc cag ttc atc    2110
Lys Tyr Asp Pro Glu Val Ile Gln Pro Leu Glu Ile Ala Gln Phe Ile
                535                 540                 545

cag gat ctg ggc ttt gag gcc gct gtg atg gaa gac tat gct ggg agc    2158
Gln Asp Leu Gly Phe Glu Ala Ala Val Met Glu Asp Tyr Ala Gly Ser
            550                 555                 560

gat gga aac att gaa ctg acc atc acc gga atg act tgt gcc tct tgc    2206
Asp Gly Asn Ile Glu Leu Thr Ile Thr Gly Met Thr Cys Ala Ser Cys
        565                 570                 575

gtg cac aac atc gag agt aaa ctg act aga acc aat ggg att acc tac    2254
Val His Asn Ile Glu Ser Lys Leu Thr Arg Thr Asn Gly Ile Thr Tyr
    580                 585                 590

gcc agt gtg gcc ctg gct aca tca aag gct ctg gtg aaa ttc gac ccc    2302
Ala Ser Val Ala Leu Ala Thr Ser Lys Ala Leu Val Lys Phe Asp Pro
595                 600                 605                 610

gag atc att gga cct agg gat atc att aag atc att gag gaa atc ggc    2350
Glu Ile Ile Gly Pro Arg Asp Ile Ile Lys Ile Ile Glu Glu Ile Gly
                615                 620                 625

ttt cac gca agc ctg gcc cag cgc aac cca aat gcc cac cat ctg gac    2398
Phe His Ala Ser Leu Ala Gln Arg Asn Pro Asn Ala His His Leu Asp
            630                 635                 640

cat aag atg gag atc aag cag tgg aag aaa agt ttc ctg tgc tca ctg    2446
His Lys Met Glu Ile Lys Gln Trp Lys Lys Ser Phe Leu Cys Ser Leu
        645                 650                 655

gtg ttt gga atc ccc gtc atg gcc ctg atg atc tac atg ctg atc cct    2494
Val Phe Gly Ile Pro Val Met Ala Leu Met Ile Tyr Met Leu Ile Pro
    660                 665                 670

agc aac gag cca cac cag tcc atg gtg ctg gat cat aac atc att cct    2542
Ser Asn Glu Pro His Gln Ser Met Val Leu Asp His Asn Ile Ile Pro
675                 680                 685                 690

ggc ctg tcc atc ctg aat ctg att ttc ttt atc ctg tgc aca ttc gtg    2590
Gly Leu Ser Ile Leu Asn Leu Ile Phe Phe Ile Leu Cys Thr Phe Val
                695                 700                 705

cag ctg ctg gga ggc tgg tac ttt tat gtg cag gca tat aaa tca ctg    2638
Gln Leu Leu Gly Gly Trp Tyr Phe Tyr Val Gln Ala Tyr Lys Ser Leu
            710                 715                 720

cga cac cgg agc gcc aat atg gac gtg ctg att gtc ctg gca acc tct    2686
Arg His Arg Ser Ala Asn Met Asp Val Leu Ile Val Leu Ala Thr Ser
        725                 730                 735

atc gcc tac gtg tat agt ctg gtc atc ctg gtg gtc gca gtg gca gag    2734
Ile Ala Tyr Val Tyr Ser Leu Val Ile Leu Val Val Ala Val Ala Glu
    740                 745                 750

aag gca gaa cgg agc cca gtg act ttc ttt gat acc cct cca atg ctg    2782
Lys Ala Glu Arg Ser Pro Val Thr Phe Phe Asp Thr Pro Pro Met Leu
755                 760                 765                 770

ttc gtg ttt atc gct ctg ggc aga tgg ctg gaa cat ctg gca aag tca    2830
Phe Val Phe Ile Ala Leu Gly Arg Trp Leu Glu His Leu Ala Lys Ser
                775                 780                 785

aaa acc agc gag gct ctg gca aag ctg atg agc ctg cag gct acc gaa    2878
```

```
Lys Thr Ser Glu Ala Leu Ala Lys Leu Met Ser Leu Gln Ala Thr Glu
            790                 795                 800

gca aca gtg gtc act ctg gga gag gac aac ctg atc att cgc gag gaa    2926
Ala Thr Val Val Thr Leu Gly Glu Asp Asn Leu Ile Ile Arg Glu Glu
        805                 810                 815

cag gtg cct atg gaa ctg gtc cag cga ggc gat atc gtg aag gtg gtc    2974
Gln Val Pro Met Glu Leu Val Gln Arg Gly Asp Ile Val Lys Val Val
    820                 825                 830

cca ggg gga aaa ttc ccc gtg gac ggc aag gtc ctg gag ggg aat act    3022
Pro Gly Gly Lys Phe Pro Val Asp Gly Lys Val Leu Glu Gly Asn Thr
835                 840                 845                 850

atg gcc gat gaa tcc ctg atc acc ggc gag gct atg cct gtg aca aag    3070
Met Ala Asp Glu Ser Leu Ile Thr Gly Glu Ala Met Pro Val Thr Lys
                855                 860                 865

aaa cca gga tca act gtc att gct ggc agc atc aac gca cac ggg tcc    3118
Lys Pro Gly Ser Thr Val Ile Ala Gly Ser Ile Asn Ala His Gly Ser
            870                 875                 880

gtg ctg atc aag gcc aca cat gtc ggg aat gac aca act ctg gct cag    3166
Val Leu Ile Lys Ala Thr His Val Gly Asn Asp Thr Thr Leu Ala Gln
        885                 890                 895

att gtg aaa ctg gtc gag gaa gcc cag atg tcc aag gct cct atc cag    3214
Ile Val Lys Leu Val Glu Glu Ala Gln Met Ser Lys Ala Pro Ile Gln
    900                 905                 910

cag ctg gcc gat cgg ttc tcc ggc tac ttc gtg ccc ttc atc att atc    3262
Gln Leu Ala Asp Arg Phe Ser Gly Tyr Phe Val Pro Phe Ile Ile Ile
915                 920                 925                 930

atg tct aca ctg act ctg gtg gtc tgg att gtg atc gga ttc att gac    3310
Met Ser Thr Leu Thr Leu Val Val Trp Ile Val Ile Gly Phe Ile Asp
                935                 940                 945

ttt ggc gtg gtc cag aga tat ttt ccc aac cct aat aag cac atc agc    3358
Phe Gly Val Val Gln Arg Tyr Phe Pro Asn Pro Asn Lys His Ile Ser
            950                 955                 960

cag acc gaa gtg atc atc agg ttc gca ttt cag acc agt att aca gtg    3406
Gln Thr Glu Val Ile Ile Arg Phe Ala Phe Gln Thr Ser Ile Thr Val
        965                 970                 975

ctg tgc atc gcc tgc cca tgt tca ctg ggg ctg gct acc ccc aca gca    3454
Leu Cys Ile Ala Cys Pro Cys Ser Leu Gly Leu Ala Thr Pro Thr Ala
    980                 985                 990

gtg atg gtc gga aca gga gtg gca gca cag aac gga att ctg atc aag    3502
Val Met Val Gly Thr Gly Val Ala Ala Gln Asn Gly Ile Leu Ile Lys
995                 1000                1005                1010

ggc ggg aaa ccc ctg gag atg gcc cac aag atc aaa act gtg atg ttt    3550
Gly Gly Lys Pro Leu Glu Met Ala His Lys Ile Lys Thr Val Met Phe
                1015                1020                1025

gac aaa act ggg acc att aca cat gga gtg ccc cgc gtc atg cga gtg    3598
Asp Lys Thr Gly Thr Ile Thr His Gly Val Pro Arg Val Met Arg Val
            1030                1035                1040

ctg ctg ctg ggc gat gtg gca acc ctg cct ctg aga aag gtc ctg gca    3646
Leu Leu Leu Gly Asp Val Ala Thr Leu Pro Leu Arg Lys Val Leu Ala
        1045                1050                1055

gtg gtc gga aca gca gag gct agc tcc gaa cac cca ctg ggg gtg gcc    3694
```

```
Val Val Gly Thr Ala Glu Ala Ser Ser Glu His Pro Leu Gly Val Ala
    1060                1065                1070

gtc aca aag tac tgc aaa gag gaa ctg ggc act gag acc ctg ggg tat    3742
Val Thr Lys Tyr Cys Lys Glu Glu Leu Gly Thr Glu Thr Leu Gly Tyr
1075                1080                1085                1090

tgt act gac ttc cag gca gtg ccc gga tgc gga atc gga tgt aaa gtc    3790
Cys Thr Asp Phe Gln Ala Val Pro Gly Cys Gly Ile Gly Cys Lys Val
                1095                1100                1105

tct aac gtg gaa ggg att ctg gct cac agt gag cgg ccc ctg agc gca    3838
Ser Asn Val Glu Gly Ile Leu Ala His Ser Glu Arg Pro Leu Ser Ala
            1110                1115                1120

cct gca tcc cat ctg aat gaa gca gga agc ctg cca gca gag aag gac    3886
Pro Ala Ser His Leu Asn Glu Ala Gly Ser Leu Pro Ala Glu Lys Asp
        1125                1130                1135

gct gtg cct cag acc ttt tcc gtc ctg atc ggc aac aga gaa tgg ctg    3934
Ala Val Pro Gln Thr Phe Ser Val Leu Ile Gly Asn Arg Glu Trp Leu
    1140                1145                1150

cgg aga aat ggg ctg aca att tct agt gac gtg tcc gat gcc atg aca    3982
Arg Arg Asn Gly Leu Thr Ile Ser Ser Asp Val Ser Asp Ala Met Thr
1155                1160                1165                1170

gat cac gag atg aaa ggc cag act gca att ctg gtg gcc atc gac gga    4030
Asp His Glu Met Lys Gly Gln Thr Ala Ile Leu Val Ala Ile Asp Gly
                1175                1180                1185

gtc ctg tgc ggc atg att gct atc gca gat gcc gtg aag cag gag gct    4078
Val Leu Cys Gly Met Ile Ala Ile Ala Asp Ala Val Lys Gln Glu Ala
            1190                1195                1200

gca ctg gcc gtc cat acc ctg cag tct atg ggc gtg gac gtg gtc ctg    4126
Ala Leu Ala Val His Thr Leu Gln Ser Met Gly Val Asp Val Val Leu
        1205                1210                1215

atc acc ggg gat aac cgg aaa aca gct aga gca att gcc act caa gtg    4174
Ile Thr Gly Asp Asn Arg Lys Thr Ala Arg Ala Ile Ala Thr Gln Val
    1220                1225                1230

ggc atc aat aag gtg ttc gct gaa gtc ctg cct agc cac aag gtc gca    4222
Gly Ile Asn Lys Val Phe Ala Glu Val Leu Pro Ser His Lys Val Ala
1235                1240                1245                1250

aaa gtg cag gag ctg cag aac aag ggc aag aaa gtc gcc atg gtg gga    4270
Lys Val Gln Glu Leu Gln Asn Lys Gly Lys Lys Val Ala Met Val Gly
                1255                1260                1265

gac ggc gtg aat gat agc cca gct ctg gca cag gca gac atg gga gtc    4318
Asp Gly Val Asn Asp Ser Pro Ala Leu Ala Gln Ala Asp Met Gly Val
            1270                1275                1280

gct att ggg aca gga act gac gtg gca atc gag gcc gct gat gtg gtc    4366
Ala Ile Gly Thr Gly Thr Asp Val Ala Ile Glu Ala Ala Asp Val Val
        1285                1290                1295

ctg att agg aat gac ctg ctg gat gtg gtc gct tct att cat ctg agt    4414
Leu Ile Arg Asn Asp Leu Leu Asp Val Val Ala Ser Ile His Leu Ser
    1300                1305                1310

aag agg aca gtg agg cgc att cgc atc aac ctg gtg ctg gcc ctg atc    4462
Lys Arg Thr Val Arg Arg Ile Arg Ile Asn Leu Val Leu Ala Leu Ile
1315                1320                1325                1330

tac aat ctg gtg gga att cca atc gca gcc ggc gtg ttt atg cca att    4510
```

```
Tyr Asn Leu Val Gly Ile Pro Ile Ala Ala Gly Val Phe Met Pro Ile
                1335                1340                1345

ggg atc gtc ctg cag ccc tgg atg ggc tca gct gca atg gcc gct tca    4558
Gly Ile Val Leu Gln Pro Trp Met Gly Ser Ala Ala Met Ala Ala Ser
            1350                1355                1360

agc gtg agc gtg gtc ctg tcc tct ctg cag ctg aaa tgc tac aag aaa    4606
Ser Val Ser Val Val Leu Ser Ser Leu Gln Leu Lys Cys Tyr Lys Lys
        1365                1370                1375

cca gac ctg gag cgg tac gaa gct cag gca cac gga cat atg aag ccc    4654
Pro Asp Leu Glu Arg Tyr Glu Ala Gln Ala His Gly His Met Lys Pro
    1380                1385                1390

ctg acc gct tcc cag gtg tct gtc cac atc ggc atg gac gat aga tgg    4702
Leu Thr Ala Ser Gln Val Ser Val His Ile Gly Met Asp Asp Arg Trp
1395                1400                1405                1410

agg gac agc cca agg gcc act cca tgg gat cag gtc agt tac gtg agc    4750
Arg Asp Ser Pro Arg Ala Thr Pro Trp Asp Gln Val Ser Tyr Val Ser
                1415                1420                1425

cag gtc agc ctg agt tca ctg acc agc gac aag ccc tcc cgc cat tct    4798
Gln Val Ser Leu Ser Ser Leu Thr Ser Asp Lys Pro Ser Arg His Ser
            1430                1435                1440

gca gcc gct gat gac gac ggg gac aag tgg agc ctg ctg ctg aac gga    4846
Ala Ala Ala Asp Asp Asp Gly Asp Lys Trp Ser Leu Leu Leu Asn Gly
        1445                1450                1455

agg gac gaa gaa cag tat atc taa ggtaccaata aagacctctt attttcattc   4900
Arg Asp Glu Glu Gln Tyr Ile
    1460                1465

atcaggtgtg gttggttttt ttgtgtgggg gcggatccat cggatcccgt gcggaccgag  4960

cggccgcagg aacccctagt gatggagttg gccactccct ctctgcgcgc tcgctcgctc  5020

actgaggccg ggcgaccaaa ggtcgcccga cgcccgggct ttgcccgggc ggcctcagtg  5080

agcgagcgag cgcgcagctg cctgcag                                      5107

<210> 4
<211> 1465
<212> PRT
<213> Artificial Sequence

<220>
<223> [CDS]:473..4870 from SEQ ID NO 3

<400> 4
Met Pro Glu Gln Glu Arg Gln Ile Thr Ala Arg Glu Gly Ala Ser Arg
1               5                  10                  15
Lys Ile Leu Ser Lys Leu Ser Leu Pro Thr Arg Ala Trp Glu Pro Ala
            20                  25                  30
Met Lys Lys Ser Phe Ala Phe Asp Asn Val Gly Tyr Glu Gly Gly Leu
        35                  40                  45
Asp Gly Leu Gly Pro Ser Ser Gln Val Ala Thr Ser Thr Val Arg Ile
    50                  55                  60
Leu Gly Met Thr Cys Gln Ser Cys Val Lys Ser Ile Glu Asp Arg Ile
65                  70                  75                  80
Ser Asn Leu Lys Gly Ile Ile Ser Met Lys Val Ser Leu Glu Gln Gly
                85                  90                  95
Ser Ala Thr Val Lys Tyr Val Pro Ser Val Val Cys Leu Gln Gln Val
            100                 105                 110
Cys His Gln Ile Gly Asp Met Gly Phe Glu Ala Ser Ile Ala Glu Gly
        115                 120                 125
```

```
Lys Ala Ala Ser Trp Pro Ser Arg Ser Leu Pro Ala Gln Glu Ala Val
    130                 135                 140
Val Lys Leu Arg Val Glu Gly Met Thr Cys Gln Ser Cys Val Ser Ser
145                 150                 155                 160
Ile Glu Gly Lys Val Arg Lys Leu Gln Gly Val Val Arg Val Lys Val
                165                 170                 175
Ser Leu Ser Asn Gln Glu Ala Val Ile Thr Tyr Gln Pro Tyr Leu Ile
            180                 185                 190
Gln Pro Glu Asp Leu Arg Asp His Val Asn Asp Met Gly Phe Glu Ala
        195                 200                 205
Ala Ile Lys Ser Lys Val Ala Pro Leu Ser Leu Gly Pro Ile Asp Ile
    210                 215                 220
Glu Arg Leu Gln Ser Thr Asn Pro Lys Arg Pro Leu Ser Ser Ala Asn
225                 230                 235                 240
Gln Asn Phe Asn Asn Ser Glu Thr Leu Gly His Gln Gly Ser His Val
                245                 250                 255
Val Thr Leu Gln Leu Arg Ile Asp Gly Met His Cys Lys Ser Cys Val
            260                 265                 270
Leu Asn Ile Glu Glu Asn Ile Gly Gln Leu Leu Gly Val Gln Ser Ile
        275                 280                 285
Gln Val Ser Leu Glu Asn Lys Thr Ala Gln Val Lys Tyr Asp Pro Ser
    290                 295                 300
Cys Thr Ser Pro Val Ala Leu Gln Arg Ala Ile Glu Ala Leu Pro Pro
305                 310                 315                 320
Gly Asn Phe Lys Val Ser Leu Pro Asp Gly Ala Glu Gly Ser Gly Thr
                325                 330                 335
Asp His Arg Ser Ser Ser Ser His Ser Pro Gly Ser Pro Pro Arg Asn
            340                 345                 350
Gln Val Gln Gly Thr Cys Ser Thr Thr Leu Ile Ala Ile Ala Gly Met
        355                 360                 365
Thr Cys Ala Ser Cys Val His Ser Ile Glu Gly Met Ile Ser Gln Leu
    370                 375                 380
Glu Gly Val Gln Gln Ile Ser Val Ser Leu Ala Glu Gly Thr Ala Thr
385                 390                 395                 400
Val Leu Tyr Asn Pro Ser Val Ile Ser Pro Glu Glu Leu Arg Ala Ala
                405                 410                 415
Ile Glu Asp Met Gly Phe Glu Ala Ser Val Val Ser Glu Ser Cys Ser
            420                 425                 430
Thr Asn Pro Leu Gly Asn His Ser Ala Gly Asn Ser Met Val Gln Thr
        435                 440                 445
Thr Asp Gly Thr Pro Thr Ser Val Gln Glu Val Ala Pro His Thr Gly
    450                 455                 460
Arg Leu Pro Ala Asn His Ala Pro Asp Ile Leu Ala Lys Ser Pro Gln
465                 470                 475                 480
Ser Thr Arg Ala Val Ala Pro Gln Lys Cys Phe Leu Gln Ile Lys Gly
                485                 490                 495
Met Thr Cys Ala Ser Cys Val Ser Asn Ile Glu Arg Asn Leu Gln Lys
            500                 505                 510
Glu Ala Gly Val Leu Ser Val Leu Val Ala Leu Met Ala Gly Lys Ala
        515                 520                 525
Glu Ile Lys Tyr Asp Pro Glu Val Ile Gln Pro Leu Glu Ile Ala Gln
    530                 535                 540
Phe Ile Gln Asp Leu Gly Phe Glu Ala Ala Val Met Glu Asp Tyr Ala
545                 550                 555                 560
Gly Ser Asp Gly Asn Ile Glu Leu Thr Ile Thr Gly Met Thr Cys Ala
                565                 570                 575
Ser Cys Val His Asn Ile Glu Ser Lys Leu Thr Arg Thr Asn Gly Ile
            580                 585                 590
Thr Tyr Ala Ser Val Ala Leu Ala Thr Ser Lys Ala Leu Val Lys Phe
        595                 600                 605
Asp Pro Glu Ile Ile Gly Pro Arg Asp Ile Ile Lys Ile Ile Glu Glu
    610                 615                 620
Ile Gly Phe His Ala Ser Leu Ala Gln Arg Asn Pro Asn Ala His His
625                 630                 635                 640
Leu Asp His Lys Met Glu Ile Lys Gln Trp Lys Lys Ser Phe Leu Cys
                645                 650                 655
Ser Leu Val Phe Gly Ile Pro Val Met Ala Leu Met Ile Tyr Met Leu
            660                 665                 670
```

```
Ile Pro Ser Asn Glu Pro His Gln Ser Met Val Leu Asp His Asn Ile
        675                 680                 685
Ile Pro Gly Leu Ser Ile Leu Asn Leu Ile Phe Phe Ile Leu Cys Thr
    690                 695                 700
Phe Val Gln Leu Leu Gly Gly Trp Tyr Phe Tyr Val Gln Ala Tyr Lys
705                 710                 715                 720
Ser Leu Arg His Arg Ser Ala Asn Met Asp Val Leu Ile Val Leu Ala
                725                 730                 735
Thr Ser Ile Ala Tyr Val Tyr Ser Leu Val Ile Leu Val Val Ala Val
            740                 745                 750
Ala Glu Lys Ala Glu Arg Ser Pro Val Thr Phe Phe Asp Thr Pro Pro
        755                 760                 765
Met Leu Phe Val Phe Ile Ala Leu Gly Arg Trp Leu Glu His Leu Ala
    770                 775                 780
Lys Ser Lys Thr Ser Glu Ala Leu Ala Lys Leu Met Ser Leu Gln Ala
785                 790                 795                 800
Thr Glu Ala Thr Val Val Thr Leu Gly Glu Asp Asn Leu Ile Ile Arg
                805                 810                 815
Glu Glu Gln Val Pro Met Glu Leu Val Gln Arg Gly Asp Ile Val Lys
            820                 825                 830
Val Val Pro Gly Gly Lys Phe Pro Val Asp Gly Lys Val Leu Glu Gly
        835                 840                 845
Asn Thr Met Ala Asp Glu Ser Leu Ile Thr Gly Glu Ala Met Pro Val
    850                 855                 860
Thr Lys Lys Pro Gly Ser Thr Val Ile Ala Gly Ser Ile Asn Ala His
865                 870                 875                 880
Gly Ser Val Leu Ile Lys Ala Thr His Val Gly Asn Asp Thr Thr Leu
                885                 890                 895
Ala Gln Ile Val Lys Leu Val Glu Glu Ala Gln Met Ser Lys Ala Pro
            900                 905                 910
Ile Gln Gln Leu Ala Asp Arg Phe Ser Gly Tyr Phe Val Pro Phe Ile
        915                 920                 925
Ile Ile Met Ser Thr Leu Thr Leu Val Val Trp Ile Val Ile Gly Phe
    930                 935                 940
Ile Asp Phe Gly Val Val Gln Arg Tyr Phe Pro Asn Pro Asn Lys His
945                 950                 955                 960
Ile Ser Gln Thr Glu Val Ile Ile Arg Phe Ala Phe Gln Thr Ser Ile
                965                 970                 975
Thr Val Leu Cys Ile Ala Cys Pro Cys Ser Leu Gly Leu Ala Thr Pro
            980                 985                 990
Thr Ala Val Met Val Gly Thr Gly Val Ala Ala Gln Asn Gly Ile Leu
        995                 1000                1005
Ile Lys Gly Gly Lys Pro Leu Glu Met Ala His Lys Ile Lys Thr Val
    1010                1015                1020
Met Phe Asp Lys Thr Gly Thr Ile Thr His Gly Val Pro Arg Val Met
1025                1030                1035                1040
Arg Val Leu Leu Leu Gly Asp Val Ala Thr Leu Pro Leu Arg Lys Val
                1045                1050                1055
Leu Ala Val Val Gly Thr Ala Glu Ala Ser Ser Glu His Pro Leu Gly
            1060                1065                1070
Val Ala Val Thr Lys Tyr Cys Lys Glu Glu Leu Gly Thr Glu Thr Leu
        1075                1080                1085
Gly Tyr Cys Thr Asp Phe Gln Ala Val Pro Gly Cys Gly Ile Gly Cys
    1090                1095                1100
Lys Val Ser Asn Val Glu Gly Ile Leu Ala His Ser Glu Arg Pro Leu
1105                1110                1115                1120
Ser Ala Pro Ala Ser His Leu Asn Glu Ala Gly Ser Leu Pro Ala Glu
                1125                1130                1135
Lys Asp Ala Val Pro Gln Thr Phe Ser Val Leu Ile Gly Asn Arg Glu
            1140                1145                1150
Trp Leu Arg Arg Asn Gly Leu Thr Ile Ser Ser Asp Val Ser Asp Ala
        1155                1160                1165
Met Thr Asp His Glu Met Lys Gly Gln Thr Ala Ile Leu Val Ala Ile
    1170                1175                1180
Asp Gly Val Leu Cys Gly Met Ile Ala Ile Ala Asp Ala Val Lys Gln
1185                1190                1195                1200
Glu Ala Ala Leu Ala Val His Thr Leu Gln Ser Met Gly Val Asp Val
                1205                1210                1215
```

Val Leu Ile Thr Gly Asp Asn Arg Lys Thr Ala Arg Ala Ile Ala Thr
1220 1225 1230
Gln Val Gly Ile Asn Lys Val Phe Ala Glu Val Leu Pro Ser His Lys
1235 1240 1245
Val Ala Lys Val Gln Glu Leu Gln Asn Lys Gly Lys Lys Val Ala Met
1250 1255 1260
Val Gly Asp Gly Val Asn Asp Ser Pro Ala Leu Ala Gln Ala Asp Met
1265 1270 1275 1280
Gly Val Ala Ile Gly Thr Gly Thr Asp Val Ala Ile Glu Ala Ala Asp
1285 1290 1295
Val Val Leu Ile Arg Asn Asp Leu Leu Asp Val Val Ala Ser Ile His
1300 1305 1310
Leu Ser Lys Arg Thr Val Arg Arg Ile Arg Ile Asn Leu Val Leu Ala
1315 1320 1325
Leu Ile Tyr Asn Leu Val Gly Ile Pro Ile Ala Ala Gly Val Phe Met
1330 1335 1340
Pro Ile Gly Ile Val Leu Gln Pro Trp Met Gly Ser Ala Ala Met Ala
1345 1350 1355 1360
Ala Ser Ser Val Ser Val Val Leu Ser Ser Leu Gln Leu Lys Cys Tyr
1365 1370 1375
Lys Lys Pro Asp Leu Glu Arg Tyr Glu Ala Gln Ala His Gly His Met
1380 1385 1390
Lys Pro Leu Thr Ala Ser Gln Val Ser Val His Ile Gly Met Asp Asp
1395 1400 1405
Arg Trp Arg Asp Ser Pro Arg Ala Thr Pro Trp Asp Gln Val Ser Tyr
1410 1415 1420
Val Ser Gln Val Ser Leu Ser Ser Leu Thr Ser Asp Lys Pro Ser Arg
1425 1430 1435 1440
His Ser Ala Ala Ala Asp Asp Asp Gly Asp Lys Trp Ser Leu Leu Leu
1445 1450 1455
Asn Gly Arg Asp Glu Glu Gln Tyr Ile
1460 1465

<210> 5
<211> 695
<212> DNA
<213> Artificial Sequence

<220>
<223> Hybrid promoter EalbPa1AT

<220>
<221> enhancer
<222> 1..382
<223> /note="Albumin gene enhancer"

<220>
<221> promoter
<222> 391..695
<223> /note="Alpha 1 antitrypsin promoter"

<400> 5
ctcgaggttc ctagattaca ttacacattc tgcaagcata gcacagagca atgttctact 60

ttaattactt tcattttctt gtatcctcac agcctagaaa ataacctgcg ttacagcatc 120

cactcagtat cccttgagca tgaggtgaca ctacttaaca tagggacgag atggtacttt 180

gtgtctcctg ctctgtcagc agggcacagt acttgctgat accagggaat gtttgttctt 240

aaataccatc attccggacg tgtttgcctt ggccagtttt ccatgtacat gcagaaagaa 300

gtttggactg atcaatacag tcctctgcct ttaaagcaat aggaaaaggc caacttgtct 360

acgtttagta tgtggctgta gatctgtacc cgccaccccc tccaccttgg acacaggacg 420

ctgtggtttc tgagccaggt acaatgactc ctttcggtaa gtgcagtgga agctgtacac 480

tgcccaggca aagcgtccgg gcagcgtagg cgggcgactc agatcccagc cagtggactt 540

agcccctgtt tgctcctccg ataactgggg tgaccttggt taatattcac cagcagcctc 600

ccccgttgcc cctctggatc cactgcttaa atacggacga ggacagggcc ctgtctcctc 660

agcttcaggc accaccactg acctgggaca gtgaa 695

<210> 6
<211> 3817
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleic acid construct of expression vector
AAV2-AAT-ATP7B(d57-486)

<220>
<221> repeat_region
<222> 1..141
<223> /note="5' ITR of adeno-associated virus serotype 2"

<220>
<221> promoter
<222> 156..460
<223> /note="alpha 1 antitrysin"

<220>
<221> CDS
<222> 473..3580
<223> /note="Truncated ATP7B (Copper-transporting ATPase 2)"
/note="Sequence encoding a truncated ATP7B (Copper transporting ATPase 2), carrying a deletion of amino acids 57..486"
/transl_table=1

<220>
<221> polyA_signal
<222> 3587..3642

<220>
<221> repeat_region
<222> 3678..3817
<223> /note="3' ITR of adeno-associated virus serotype 2"

<400> 6
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60

gggcgacctt tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120

actccatcac taggggttcc tgcggccgca cgcgtcgcca ccccctccac cttggacaca 180

ggacgctgtg gtttctgagc caggtacaat gactcctttc ggtaagtgca gtggaagctg 240

tacactgccc aggcaaagcg tccgggcagc gtaggcgggc gactcagatc ccagccagtg 300

gacttagccc ctgtttgctc ctccgataac tggggtgacc ttggttaata ttcaccagca 360

gcctcccccg ttgcccctct ggatccactg cttaaatacg gacgaggaca gggccctgtc 420

tcctcagctt caggcaccac cactgacctg ggacagtgaa gcggccgcca cc atg cct 478
Met Pro
1

gag cag gag aga cag atc aca gcc aga gaa ggg gcc agt cgg aaa atc 526
Glu Gln Glu Arg Gln Ile Thr Ala Arg Glu Gly Ala Ser Arg Lys Ile
5 10 15

```
tta tct aag ctt tct ttg cct acc cgt gcc tgg gaa cca gca atg aag     574
Leu Ser Lys Leu Ser Leu Pro Thr Arg Ala Trp Glu Pro Ala Met Lys
    20                  25                  30

aag agt ttt gct ttt gac aat gtt ggc tat gaa ggt ggt ctg gat ggc     622
Lys Ser Phe Ala Phe Asp Asn Val Gly Tyr Glu Gly Gly Leu Asp Gly
35                  40                  45                  50

ctg ggc cct tct tct cag ccg cag aag tgc ttc tta cag atc aaa ggc     670
Leu Gly Pro Ser Ser Gln Pro Gln Lys Cys Phe Leu Gln Ile Lys Gly
                55                  60                  65

atg acc tgt gca tcc tgt gtg tct aac ata gaa agg aat ctg cag aaa     718
Met Thr Cys Ala Ser Cys Val Ser Asn Ile Glu Arg Asn Leu Gln Lys
            70                  75                  80

gaa gct ggt gtt ctc tcc gtg ttg gtt gcc ttg atg gca gga aag gca     766
Glu Ala Gly Val Leu Ser Val Leu Val Ala Leu Met Ala Gly Lys Ala
        85                  90                  95

gag atc aag tat gac cca gag gtc atc cag ccc ctc gag ata gct cag     814
Glu Ile Lys Tyr Asp Pro Glu Val Ile Gln Pro Leu Glu Ile Ala Gln
    100                 105                 110

ttc atc cag gac ctg ggt ttt gag gca gca gtc atg gag gac tac gca     862
Phe Ile Gln Asp Leu Gly Phe Glu Ala Ala Val Met Glu Asp Tyr Ala
115                 120                 125                 130

ggc tcc gat ggc aac att gag ctg aca atc aca ggg atg acc tgc gcg     910
Gly Ser Asp Gly Asn Ile Glu Leu Thr Ile Thr Gly Met Thr Cys Ala
                135                 140                 145

tcc tgt gtc cac aac ata gag tcc aaa ctc acg agg aca aat ggc atc     958
Ser Cys Val His Asn Ile Glu Ser Lys Leu Thr Arg Thr Asn Gly Ile
            150                 155                 160

act tat gcc tcc gtt gcc ctt gcc acc agc aaa gcc ctt gtt aag ttt    1006
Thr Tyr Ala Ser Val Ala Leu Ala Thr Ser Lys Ala Leu Val Lys Phe
        165                 170                 175

gac ccg gaa att atc ggt cca cgg gat att atc aaa att att gag gaa    1054
Asp Pro Glu Ile Ile Gly Pro Arg Asp Ile Ile Lys Ile Ile Glu Glu
    180                 185                 190

att ggc ttt cat gct tcc ctg gcc cag aga aac ccc aac gct cat cac    1102
Ile Gly Phe His Ala Ser Leu Ala Gln Arg Asn Pro Asn Ala His His
195                 200                 205                 210

ttg gac cac aag atg gaa ata aag cag tgg aag aag tct ttc ctg tgc    1150
Leu Asp His Lys Met Glu Ile Lys Gln Trp Lys Lys Ser Phe Leu Cys
                215                 220                 225

agc ctg gtg ttt ggc atc cct gtc atg gcc tta atg atc tat atg ctg    1198
Ser Leu Val Phe Gly Ile Pro Val Met Ala Leu Met Ile Tyr Met Leu
            230                 235                 240

ata ccc agc aac gag ccc cac cag tcc atg gtc ctg gac cac aac atc    1246
Ile Pro Ser Asn Glu Pro His Gln Ser Met Val Leu Asp His Asn Ile
        245                 250                 255

att cca gga ctg tcc att cta aat ctc atc ttc ttt atc ttg tgt acc    1294
Ile Pro Gly Leu Ser Ile Leu Asn Leu Ile Phe Phe Ile Leu Cys Thr
    260                 265                 270

ttt gtc cag ctc ctc ggt ggg tgg tac ttc tac gtt cag gcc tac aaa    1342
Phe Val Gln Leu Leu Gly Gly Trp Tyr Phe Tyr Val Gln Ala Tyr Lys
275                 280                 285                 290
```

```
tct ctg aga cac agg tca gcc aac atg gac gtg ctc atc gtc ctg gcc    1390
Ser Leu Arg His Arg Ser Ala Asn Met Asp Val Leu Ile Val Leu Ala
                295                 300                 305

aca agc att gct tat gtt tat tct ctg gtc atc ctg gtg gtt gct gtg    1438
Thr Ser Ile Ala Tyr Val Tyr Ser Leu Val Ile Leu Val Val Ala Val
            310                 315                 320

gct gag aag gcg gag agg agc cct gtg aca ttc ttc gac acg ccc ccc    1486
Ala Glu Lys Ala Glu Arg Ser Pro Val Thr Phe Phe Asp Thr Pro Pro
        325                 330                 335

atg ctc ttt gtg ttc att gcc ctg ggc cgg tgg ctg gaa cac ttg gca    1534
Met Leu Phe Val Phe Ile Ala Leu Gly Arg Trp Leu Glu His Leu Ala
    340                 345                 350

aag agc aaa acc tca gaa gcc ctg gct aaa ctc atg tct ctc caa gcc    1582
Lys Ser Lys Thr Ser Glu Ala Leu Ala Lys Leu Met Ser Leu Gln Ala
355                 360                 365                 370

aca gaa gcc acc gtt gtg acc ctt ggt gag gac aat tta atc atc agg    1630
Thr Glu Ala Thr Val Val Thr Leu Gly Glu Asp Asn Leu Ile Ile Arg
                375                 380                 385

gag gag caa gtc ccc atg gag ctg gtg cag cgg ggc gat atc gtc aag    1678
Glu Glu Gln Val Pro Met Glu Leu Val Gln Arg Gly Asp Ile Val Lys
            390                 395                 400

gtg gtc cct ggg gga aag ttt cca gtg gat ggg aaa gtc ctg gaa ggc    1726
Val Val Pro Gly Gly Lys Phe Pro Val Asp Gly Lys Val Leu Glu Gly
        405                 410                 415

aat acc atg gct gat gag tcc ctc atc aca gga gaa gcc atg cca gtc    1774
Asn Thr Met Ala Asp Glu Ser Leu Ile Thr Gly Glu Ala Met Pro Val
    420                 425                 430

act aag aaa ccc gga agc act gta att gcg ggg tct ata aat gca cat    1822
Thr Lys Lys Pro Gly Ser Thr Val Ile Ala Gly Ser Ile Asn Ala His
435                 440                 445                 450

ggc tct gtg ctc att aaa gct acc cac gtg ggc aat gac acc act ttg    1870
Gly Ser Val Leu Ile Lys Ala Thr His Val Gly Asn Asp Thr Thr Leu
                455                 460                 465

gct cag att gtg aaa ctg gtg gaa gag gct cag atg tca aag gca ccc    1918
Ala Gln Ile Val Lys Leu Val Glu Glu Ala Gln Met Ser Lys Ala Pro
            470                 475                 480

att cag cag ctg gct gac cgg ttt agt gga tat ttt gtc cca ttt atc    1966
Ile Gln Gln Leu Ala Asp Arg Phe Ser Gly Tyr Phe Val Pro Phe Ile
        485                 490                 495

atc atc atg tca act ttg acg ttg gtg gta tgg att gta atc ggt ttt    2014
Ile Ile Met Ser Thr Leu Thr Leu Val Val Trp Ile Val Ile Gly Phe
    500                 505                 510

atc gat ttt ggt gtt gtt cag aga tac ttt cct aac ccc aac aag cac    2062
Ile Asp Phe Gly Val Val Gln Arg Tyr Phe Pro Asn Pro Asn Lys His
515                 520                 525                 530

atc tcc cag aca gag gtg atc atc cgg ttt gct ttc cag acg tcc atc    2110
Ile Ser Gln Thr Glu Val Ile Ile Arg Phe Ala Phe Gln Thr Ser Ile
                535                 540                 545

acg gtg ctg tgc att gcc tgc ccc tgc tcc ctg ggg ctg gcc acg ccc    2158
Thr Val Leu Cys Ile Ala Cys Pro Cys Ser Leu Gly Leu Ala Thr Pro
            550                 555                 560
```

```
acg gct gtc atg gtg ggc acc ggg gtg gcc gcg cag aac ggc atc ctc    2206
Thr Ala Val Met Val Gly Thr Gly Val Ala Ala Gln Asn Gly Ile Leu
        565                 570                 575

atc aag gga ggc aag ccc ctg gag atg gcg cac aag ata aag act gtg    2254
Ile Lys Gly Gly Lys Pro Leu Glu Met Ala His Lys Ile Lys Thr Val
    580                 585                 590

atg ttt gac aag act ggc acc att acc cat ggc gtc ccc agg gtc atg    2302
Met Phe Asp Lys Thr Gly Thr Ile Thr His Gly Val Pro Arg Val Met
595                 600                 605                 610

cgg gtg ctc ctg ctg ggg gat gtg gcc aca ctg ccc ctc agg aag gtt    2350
Arg Val Leu Leu Leu Gly Asp Val Ala Thr Leu Pro Leu Arg Lys Val
                615                 620                 625

ctg gct gtg gtg ggg act gcg gag gcc agc agt gaa cac ccc ttg ggc    2398
Leu Ala Val Val Gly Thr Ala Glu Ala Ser Ser Glu His Pro Leu Gly
            630                 635                 640

gtg gca gtc acc aaa tac tgt aaa gag gaa ctt gga aca gag acc ttg    2446
Val Ala Val Thr Lys Tyr Cys Lys Glu Glu Leu Gly Thr Glu Thr Leu
        645                 650                 655

gga tac tgc acg gac ttc cag gca gtg cca ggc tgt gga att ggg tgc    2494
Gly Tyr Cys Thr Asp Phe Gln Ala Val Pro Gly Cys Gly Ile Gly Cys
    660                 665                 670

aaa gtc agc aac gtg gaa ggc atc ctg gcc cac agt gag cgc cct ttg    2542
Lys Val Ser Asn Val Glu Gly Ile Leu Ala His Ser Glu Arg Pro Leu
675                 680                 685                 690

agt gca ccg gcc agt cac ctg aat gag gct ggc agc ctt ccc gca gaa    2590
Ser Ala Pro Ala Ser His Leu Asn Glu Ala Gly Ser Leu Pro Ala Glu
                695                 700                 705

aaa gat gca gtc ccc cag acc ttc tct gtg ctg att gga aac cgt gag    2638
Lys Asp Ala Val Pro Gln Thr Phe Ser Val Leu Ile Gly Asn Arg Glu
            710                 715                 720

tgg ctg agg cgc aac ggt tta acc att tct agc gat gtc agt gac gct    2686
Trp Leu Arg Arg Asn Gly Leu Thr Ile Ser Ser Asp Val Ser Asp Ala
        725                 730                 735

atg aca gac cac gag atg aaa gga cag aca gcc atc ctg gtg gct att    2734
Met Thr Asp His Glu Met Lys Gly Gln Thr Ala Ile Leu Val Ala Ile
    740                 745                 750

gac ggt gtg ctc tgt ggg atg atc gca atc gca gac gct gtc aag cag    2782
Asp Gly Val Leu Cys Gly Met Ile Ala Ile Ala Asp Ala Val Lys Gln
755                 760                 765                 770

gag gct gcc ctg gct gtg cac acg ctg cag agc atg ggt gtg gac gtg    2830
Glu Ala Ala Leu Ala Val His Thr Leu Gln Ser Met Gly Val Asp Val
                775                 780                 785

gtt ctg atc acg ggg gac aac cgg aag aca gcc aga gct att gcc acc    2878
Val Leu Ile Thr Gly Asp Asn Arg Lys Thr Ala Arg Ala Ile Ala Thr
            790                 795                 800

cag gtt ggc atc aac aaa gtc ttt gca gag gtg ctg cct tcg cac aag    2926
Gln Val Gly Ile Asn Lys Val Phe Ala Glu Val Leu Pro Ser His Lys
        805                 810                 815

gtg gcc aag gtc cag gag ctc cag aat aaa ggg aag aaa gtc gcc atg    2974
Val Ala Lys Val Gln Glu Leu Gln Asn Lys Gly Lys Lys Val Ala Met
    820                 825                 830
```

```
gtg ggg gat ggg gtc aat gac tcc ccg gcc ttg gcc cag gca gac atg    3022
Val Gly Asp Gly Val Asn Asp Ser Pro Ala Leu Ala Gln Ala Asp Met
835                 840                 845                 850

ggt gtg gcc att ggc acc ggc acg gat gtg gcc atc gag gca gcc gac    3070
Gly Val Ala Ile Gly Thr Gly Thr Asp Val Ala Ile Glu Ala Ala Asp
                855                 860                 865

gtc gtc ctt atc aga aat gat ttg ctg gat gtg gtg gct agc att cac    3118
Val Val Leu Ile Arg Asn Asp Leu Leu Asp Val Val Ala Ser Ile His
            870                 875                 880

ctt tcc aag agg act gtc cga agg ata cgc atc aac ctg gtc ctg gca    3166
Leu Ser Lys Arg Thr Val Arg Arg Ile Arg Ile Asn Leu Val Leu Ala
        885                 890                 895

ctg att tat aac ctg gtt ggg ata ccc att gca gca ggt gtc ttc atg    3214
Leu Ile Tyr Asn Leu Val Gly Ile Pro Ile Ala Ala Gly Val Phe Met
    900                 905                 910

ccc atc ggc att gtg ctg cag ccc tgg atg ggc tca gcg gcc atg gca    3262
Pro Ile Gly Ile Val Leu Gln Pro Trp Met Gly Ser Ala Ala Met Ala
915                 920                 925                 930

gcc tcc tct gtg tct gtg gtg ctc tca tcc ctg cag ctc aag tgc tat    3310
Ala Ser Ser Val Ser Val Val Leu Ser Ser Leu Gln Leu Lys Cys Tyr
                935                 940                 945

aag aag cct gac ctg gag agg tat gag gca cag gcg cat ggc cac atg    3358
Lys Lys Pro Asp Leu Glu Arg Tyr Glu Ala Gln Ala His Gly His Met
            950                 955                 960

aag ccc ctg acg gca tcc cag gtc agt gtg cac ata ggc atg gat gac    3406
Lys Pro Leu Thr Ala Ser Gln Val Ser Val His Ile Gly Met Asp Asp
        965                 970                 975

agg tgg cgg gac tcc ccc agg gcc aca cca tgg gac cag gtc agc tat    3454
Arg Trp Arg Asp Ser Pro Arg Ala Thr Pro Trp Asp Gln Val Ser Tyr
    980                 985                 990

gtc agc cag gtg tcg ctg tcc tcc ctg acg tcc gac aag cca tct cgg    3502
Val Ser Gln Val Ser Leu Ser Ser Leu Thr Ser Asp Lys Pro Ser Arg
995                 1000                1005                1010

cac agc gct gca gca gac gat gat ggg gac aag tgg tct ctg ctc ctg    3550
His Ser Ala Ala Ala Asp Asp Asp Gly Asp Lys Trp Ser Leu Leu Leu
                1015                1020                1025

aat ggc agg gat gag gag cag tac atc tga ggtaccaata aagacctctt      3600
Asn Gly Arg Asp Glu Glu Gln Tyr Ile
            1030                1035

attttcattc atcaggtgtg gttggttttt ttgtgtgggg gcggatccat cggatcccgt  3660

gcggaccgag cggccgcagg aacccctagt gatggagttg gccactccct ctctgcgcgc  3720

tcgctcgctc actgaggccg ggcgaccaaa ggtcgcccga cgcccgggct ttgcccgggc  3780

ggcctcagtg agcgagcgag cgcgcagctg cctgcag                           3817
```

<210> 7
<211> 1035
<212> PRT
<213> Artificial Sequence

<220>
<223> [CDS]:473..3580 from SEQ ID NO 6

<400> 7

```
Met Pro Glu Gln Glu Arg Gln Ile Thr Ala Arg Glu Gly Ala Ser Arg
1               5                  10                  15
Lys Ile Leu Ser Lys Leu Ser Leu Pro Thr Arg Ala Trp Glu Pro Ala
            20                  25                  30
Met Lys Lys Ser Phe Ala Phe Asp Asn Val Gly Tyr Glu Gly Gly Leu
        35                  40                  45
Asp Gly Leu Gly Pro Ser Ser Gln Pro Gln Lys Cys Phe Leu Gln Ile
    50                  55                  60
Lys Gly Met Thr Cys Ala Ser Cys Val Ser Asn Ile Glu Arg Asn Leu
65                  70                  75                  80
Gln Lys Glu Ala Gly Val Leu Ser Val Leu Val Ala Leu Met Ala Gly
                85                  90                  95
Lys Ala Glu Ile Lys Tyr Asp Pro Glu Val Ile Gln Pro Leu Glu Ile
            100                 105                 110
Ala Gln Phe Ile Gln Asp Leu Gly Phe Glu Ala Ala Val Met Glu Asp
        115                 120                 125
Tyr Ala Gly Ser Asp Gly Asn Ile Glu Leu Thr Ile Thr Gly Met Thr
    130                 135                 140
Cys Ala Ser Cys Val His Asn Ile Glu Ser Lys Leu Thr Arg Thr Asn
145                 150                 155                 160
Gly Ile Thr Tyr Ala Ser Val Ala Leu Ala Thr Ser Lys Ala Leu Val
                165                 170                 175
Lys Phe Asp Pro Glu Ile Ile Gly Pro Arg Asp Ile Ile Lys Ile Ile
            180                 185                 190
Glu Glu Ile Gly Phe His Ala Ser Leu Ala Gln Arg Asn Pro Asn Ala
        195                 200                 205
His His Leu Asp His Lys Met Glu Ile Lys Gln Trp Lys Lys Ser Phe
    210                 215                 220
Leu Cys Ser Leu Val Phe Gly Ile Pro Val Met Ala Leu Met Ile Tyr
225                 230                 235                 240
Met Leu Ile Pro Ser Asn Glu Pro His Gln Ser Met Val Leu Asp His
                245                 250                 255
Asn Ile Ile Pro Gly Leu Ser Ile Leu Asn Leu Ile Phe Phe Ile Leu
            260                 265                 270
Cys Thr Phe Val Gln Leu Leu Gly Gly Trp Tyr Phe Tyr Val Gln Ala
        275                 280                 285
Tyr Lys Ser Leu Arg His Arg Ser Ala Asn Met Asp Val Leu Ile Val
    290                 295                 300
Leu Ala Thr Ser Ile Ala Tyr Val Tyr Ser Leu Val Ile Leu Val Val
305                 310                 315                 320
Ala Val Ala Glu Lys Ala Glu Arg Ser Pro Val Thr Phe Phe Asp Thr
                325                 330                 335
Pro Pro Met Leu Phe Val Phe Ile Ala Leu Gly Arg Trp Leu Glu His
            340                 345                 350
Leu Ala Lys Ser Lys Thr Ser Glu Ala Leu Ala Lys Leu Met Ser Leu
        355                 360                 365
Gln Ala Thr Glu Ala Thr Val Val Thr Leu Gly Glu Asp Asn Leu Ile
    370                 375                 380
Ile Arg Glu Glu Gln Val Pro Met Glu Leu Val Gln Arg Gly Asp Ile
385                 390                 395                 400
Val Lys Val Val Pro Gly Gly Lys Phe Pro Val Asp Gly Lys Val Leu
                405                 410                 415
Glu Gly Asn Thr Met Ala Asp Glu Ser Leu Ile Thr Gly Glu Ala Met
            420                 425                 430
Pro Val Thr Lys Lys Pro Gly Ser Thr Val Ile Ala Gly Ser Ile Asn
        435                 440                 445
Ala His Gly Ser Val Leu Ile Lys Ala Thr His Val Gly Asn Asp Thr
    450                 455                 460
Thr Leu Ala Gln Ile Val Lys Leu Val Glu Glu Ala Gln Met Ser Lys
465                 470                 475                 480
Ala Pro Ile Gln Gln Leu Ala Asp Arg Phe Ser Gly Tyr Phe Val Pro
                485                 490                 495
Phe Ile Ile Ile Met Ser Thr Leu Thr Leu Val Val Trp Ile Val Ile
            500                 505                 510
Gly Phe Ile Asp Phe Gly Val Val Gln Arg Tyr Phe Pro Asn Pro Asn
        515                 520                 525
Lys His Ile Ser Gln Thr Glu Val Ile Ile Arg Phe Ala Phe Gln Thr
```

530 535 540
Ser Ile Thr Val Leu Cys Ile Ala Cys Pro Cys Ser Leu Gly Leu Ala
545 550 555 560
Thr Pro Thr Ala Val Met Val Gly Thr Gly Val Ala Ala Gln Asn Gly
565 570 575
Ile Leu Ile Lys Gly Gly Lys Pro Leu Glu Met Ala His Lys Ile Lys
580 585 590
Thr Val Met Phe Asp Lys Thr Gly Thr Ile Thr His Gly Val Pro Arg
595 600 605
Val Met Arg Val Leu Leu Leu Gly Asp Val Ala Thr Leu Pro Leu Arg
610 615 620
Lys Val Leu Ala Val Val Gly Thr Ala Glu Ala Ser Ser Glu His Pro
625 630 635 640
Leu Gly Val Ala Val Thr Lys Tyr Cys Lys Glu Glu Leu Gly Thr Glu
645 650 655
Thr Leu Gly Tyr Cys Thr Asp Phe Gln Ala Val Pro Gly Cys Gly Ile
660 665 670
Gly Cys Lys Val Ser Asn Val Glu Gly Ile Leu Ala His Ser Glu Arg
675 680 685
Pro Leu Ser Ala Pro Ala Ser His Leu Asn Glu Ala Gly Ser Leu Pro
690 695 700
Ala Glu Lys Asp Ala Val Pro Gln Thr Phe Ser Val Leu Ile Gly Asn
705 710 715 720
Arg Glu Trp Leu Arg Arg Asn Gly Leu Thr Ile Ser Ser Asp Val Ser
725 730 735
Asp Ala Met Thr Asp His Glu Met Lys Gly Gln Thr Ala Ile Leu Val
740 745 750
Ala Ile Asp Gly Val Leu Cys Gly Met Ile Ala Ile Ala Asp Ala Val
755 760 765
Lys Gln Glu Ala Ala Leu Ala Val His Thr Leu Gln Ser Met Gly Val
770 775 780
Asp Val Val Leu Ile Thr Gly Asp Asn Arg Lys Thr Ala Arg Ala Ile
785 790 795 800
Ala Thr Gln Val Gly Ile Asn Lys Val Phe Ala Glu Val Leu Pro Ser
805 810 815
His Lys Val Ala Lys Val Gln Glu Leu Gln Asn Lys Gly Lys Lys Val
820 825 830
Ala Met Val Gly Asp Gly Val Asn Asp Ser Pro Ala Leu Ala Gln Ala
835 840 845
Asp Met Gly Val Ala Ile Gly Thr Gly Thr Asp Val Ala Ile Glu Ala
850 855 860
Ala Asp Val Val Leu Ile Arg Asn Asp Leu Leu Asp Val Val Ala Ser
865 870 875 880
Ile His Leu Ser Lys Arg Thr Val Arg Arg Ile Arg Ile Asn Leu Val
885 890 895
Leu Ala Leu Ile Tyr Asn Leu Val Gly Ile Pro Ile Ala Ala Gly Val
900 905 910
Phe Met Pro Ile Gly Ile Val Leu Gln Pro Trp Met Gly Ser Ala Ala
915 920 925
Met Ala Ala Ser Ser Val Ser Val Val Leu Ser Ser Leu Gln Leu Lys
930 935 940
Cys Tyr Lys Lys Pro Asp Leu Glu Arg Tyr Glu Ala Gln Ala His Gly
945 950 955 960
His Met Lys Pro Leu Thr Ala Ser Gln Val Ser Val His Ile Gly Met
965 970 975
Asp Asp Arg Trp Arg Asp Ser Pro Arg Ala Thr Pro Trp Asp Gln Val
980 985 990
Ser Tyr Val Ser Gln Val Ser Leu Ser Ser Leu Thr Ser Asp Lys Pro
995 1000 1005
Ser Arg His Ser Ala Ala Ala Asp Asp Asp Gly Asp Lys Trp Ser Leu
1010 1015 1020
Leu Leu Asn Gly Arg Asp Glu Glu Gln Tyr Ile
1025 1030 1035

<210> 8
<211> 3108
<212> DNA
<213> Artificial Sequence

<220>
<223> Codon optimized sequence coATP7B(d57-486) encoding truncated ATP7B, with deletion of amino acids 57-486

<220>
<221> CDS
<222> 1..3108
<223> /transl_table=1

<400> 8

```
atg cca gaa cag gaa cgc cag atc aca gca aga gag gga gca agt cgg      48
Met Pro Glu Gln Glu Arg Gln Ile Thr Ala Arg Glu Gly Ala Ser Arg
1               5                   10                  15

aaa atc ctg agc aaa ctg agc ctg cca acc aga gca tgg gaa ccc gca      96
Lys Ile Leu Ser Lys Leu Ser Leu Pro Thr Arg Ala Trp Glu Pro Ala
            20                  25                  30

atg aag aaa agc ttc gcc ttt gac aac gtg gga tac gag gga ggg ctg     144
Met Lys Lys Ser Phe Ala Phe Asp Asn Val Gly Tyr Glu Gly Gly Leu
        35                  40                  45

gat gga ctg gga cct agc tcc cag cct cag aag tgt ttt ctg cag atc     192
Asp Gly Leu Gly Pro Ser Ser Gln Pro Gln Lys Cys Phe Leu Gln Ile
    50                  55                  60

aaa ggc atg acc tgc gcc tct tgc gtg agc aac att gag cgg aat ctg     240
Lys Gly Met Thr Cys Ala Ser Cys Val Ser Asn Ile Glu Arg Asn Leu
65                  70                  75                  80

cag aag gaa gct ggg gtg ctg agc gtg ctg gtc gca ctg atg gcc gga     288
Gln Lys Glu Ala Gly Val Leu Ser Val Leu Val Ala Leu Met Ala Gly
                85                  90                  95

aag gct gag atc aag tac gac cct gaa gtg atc cag cca ctg gag att     336
Lys Ala Glu Ile Lys Tyr Asp Pro Glu Val Ile Gln Pro Leu Glu Ile
            100                 105                 110

gcc cag ttc atc cag gat ctg ggc ttt gag gcc gct gtg atg gaa gac     384
Ala Gln Phe Ile Gln Asp Leu Gly Phe Glu Ala Ala Val Met Glu Asp
        115                 120                 125

tat gct ggg agc gat gga aac att gaa ctg acc atc acc gga atg act     432
Tyr Ala Gly Ser Asp Gly Asn Ile Glu Leu Thr Ile Thr Gly Met Thr
    130                 135                 140

tgt gcc tct tgc gtg cac aac atc gag agt aaa ctg act aga acc aat     480
Cys Ala Ser Cys Val His Asn Ile Glu Ser Lys Leu Thr Arg Thr Asn
145                 150                 155                 160

ggg att acc tac gcc agt gtg gcc ctg gct aca tca aag gct ctg gtg     528
Gly Ile Thr Tyr Ala Ser Val Ala Leu Ala Thr Ser Lys Ala Leu Val
                165                 170                 175

aaa ttc gac ccc gag atc att gga cct agg gat atc att aag atc att     576
Lys Phe Asp Pro Glu Ile Ile Gly Pro Arg Asp Ile Ile Lys Ile Ile
            180                 185                 190

gag gaa atc ggc ttt cac gca agc ctg gcc cag cgc aac cca aat gcc     624
Glu Glu Ile Gly Phe His Ala Ser Leu Ala Gln Arg Asn Pro Asn Ala
        195                 200                 205

cac cat ctg gac cat aag atg gag atc aag cag tgg aag aaa agt ttc     672
His His Leu Asp His Lys Met Glu Ile Lys Gln Trp Lys Lys Ser Phe
    210                 215                 220

ctg tgc tca ctg gtg ttt gga atc ccc gtc atg gcc ctg atg atc tac     720
```

```
Leu Cys Ser Leu Val Phe Gly Ile Pro Val Met Ala Leu Met Ile Tyr
225                 230                 235                 240

atg ctg atc cct agc aac gag cca cac cag tcc atg gtg ctg gat cat     768
Met Leu Ile Pro Ser Asn Glu Pro His Gln Ser Met Val Leu Asp His
                245                 250                 255

aac atc att cct ggc ctg tcc atc ctg aat ctg att ttc ttt atc ctg     816
Asn Ile Ile Pro Gly Leu Ser Ile Leu Asn Leu Ile Phe Phe Ile Leu
            260                 265                 270

tgc aca ttc gtg cag ctg ctg gga ggc tgg tac ttt tat gtg cag gca     864
Cys Thr Phe Val Gln Leu Leu Gly Gly Trp Tyr Phe Tyr Val Gln Ala
        275                 280                 285

tat aaa tca ctg cga cac cgg agc gcc aat atg gac gtg ctg att gtc     912
Tyr Lys Ser Leu Arg His Arg Ser Ala Asn Met Asp Val Leu Ile Val
    290                 295                 300

ctg gca acc tct atc gcc tac gtg tat agt ctg gtc atc ctg gtg gtc     960
Leu Ala Thr Ser Ile Ala Tyr Val Tyr Ser Leu Val Ile Leu Val Val
305                 310                 315                 320

gca gtg gca gag aag gca gaa cgg agc cca gtg act ttc ttt gat acc    1008
Ala Val Ala Glu Lys Ala Glu Arg Ser Pro Val Thr Phe Phe Asp Thr
                325                 330                 335

cct cca atg ctg ttc gtg ttt atc gct ctg ggc aga tgg ctg gaa cat    1056
Pro Pro Met Leu Phe Val Phe Ile Ala Leu Gly Arg Trp Leu Glu His
            340                 345                 350

ctg gca aag tca aaa acc agc gag gct ctg gca aag ctg atg agc ctg    1104
Leu Ala Lys Ser Lys Thr Ser Glu Ala Leu Ala Lys Leu Met Ser Leu
        355                 360                 365

cag gct acc gaa gca aca gtg gtc act ctg gga gag gac aac ctg atc    1152
Gln Ala Thr Glu Ala Thr Val Val Thr Leu Gly Glu Asp Asn Leu Ile
    370                 375                 380

att cgc gag gaa cag gtg cct atg gaa ctg gtc cag cga ggc gat atc    1200
Ile Arg Glu Glu Gln Val Pro Met Glu Leu Val Gln Arg Gly Asp Ile
385                 390                 395                 400

gtg aag gtg gtc cca ggg gga aaa ttc ccc gtg gac ggc aag gtc ctg    1248
Val Lys Val Val Pro Gly Gly Lys Phe Pro Val Asp Gly Lys Val Leu
                405                 410                 415

gag ggg aat act atg gcc gat gaa tcc ctg atc acc ggc gag gct atg    1296
Glu Gly Asn Thr Met Ala Asp Glu Ser Leu Ile Thr Gly Glu Ala Met
            420                 425                 430

cct gtg aca aag aaa cca gga tca act gtc att gct ggc agc atc aac    1344
Pro Val Thr Lys Lys Pro Gly Ser Thr Val Ile Ala Gly Ser Ile Asn
        435                 440                 445

gca cac ggg tcc gtg ctg atc aag gcc aca cat gtc ggg aat gac aca    1392
Ala His Gly Ser Val Leu Ile Lys Ala Thr His Val Gly Asn Asp Thr
    450                 455                 460

act ctg gct cag att gtg aaa ctg gtc gag gaa gcc cag atg tcc aag    1440
Thr Leu Ala Gln Ile Val Lys Leu Val Glu Glu Ala Gln Met Ser Lys
465                 470                 475                 480

gct cct atc cag cag ctg gcc gat cgg ttc tcc ggc tac ttc gtg ccc    1488
Ala Pro Ile Gln Gln Leu Ala Asp Arg Phe Ser Gly Tyr Phe Val Pro
                485                 490                 495

ttc atc att atc atg tct aca ctg act ctg gtg gtc tgg att gtg atc    1536
```

```
Phe Ile Ile Ile Met Ser Thr Leu Thr Leu Val Val Trp Ile Val Ile
            500                 505                 510

gga ttc att gac ttt ggc gtg gtc cag aga tat ttt ccc aac cct aat    1584
Gly Phe Ile Asp Phe Gly Val Val Gln Arg Tyr Phe Pro Asn Pro Asn
        515                 520                 525

aag cac atc agc cag acc gaa gtg atc atc agg ttc gca ttt cag acc    1632
Lys His Ile Ser Gln Thr Glu Val Ile Ile Arg Phe Ala Phe Gln Thr
    530                 535                 540

agt att aca gtg ctg tgc atc gcc tgc cca tgt tca ctg ggg ctg gct    1680
Ser Ile Thr Val Leu Cys Ile Ala Cys Pro Cys Ser Leu Gly Leu Ala
545                 550                 555                 560

acc ccc aca gca gtg atg gtc gga aca gga gtg gca gca cag aac gga    1728
Thr Pro Thr Ala Val Met Val Gly Thr Gly Val Ala Ala Gln Asn Gly
                565                 570                 575

att ctg atc aag ggc ggg aaa ccc ctg gag atg gcc cac aag atc aaa    1776
Ile Leu Ile Lys Gly Gly Lys Pro Leu Glu Met Ala His Lys Ile Lys
            580                 585                 590

act gtg atg ttt gac aaa act ggg acc att aca cat gga gtg ccc cgc    1824
Thr Val Met Phe Asp Lys Thr Gly Thr Ile Thr His Gly Val Pro Arg
        595                 600                 605

gtc atg cga gtg ctg ctg ctg ggc gat gtg gca acc ctg cct ctg aga    1872
Val Met Arg Val Leu Leu Leu Gly Asp Val Ala Thr Leu Pro Leu Arg
    610                 615                 620

aag gtc ctg gca gtg gtc gga aca gca gag gct agc tcc gaa cac cca    1920
Lys Val Leu Ala Val Val Gly Thr Ala Glu Ala Ser Ser Glu His Pro
625                 630                 635                 640

ctg ggg gtg gcc gtc aca aag tac tgc aaa gag gaa ctg ggc act gag    1968
Leu Gly Val Ala Val Thr Lys Tyr Cys Lys Glu Glu Leu Gly Thr Glu
                645                 650                 655

acc ctg ggg tat tgt act gac ttc cag gca gtg ccc gga tgc gga atc    2016
Thr Leu Gly Tyr Cys Thr Asp Phe Gln Ala Val Pro Gly Cys Gly Ile
            660                 665                 670

gga tgt aaa gtc tct aac gtg gaa ggg att ctg gct cac agt gag cgg    2064
Gly Cys Lys Val Ser Asn Val Glu Gly Ile Leu Ala His Ser Glu Arg
        675                 680                 685

ccc ctg agc gca cct gca tcc cat ctg aat gaa gca gga agc ctg cca    2112
Pro Leu Ser Ala Pro Ala Ser His Leu Asn Glu Ala Gly Ser Leu Pro
    690                 695                 700

gca gag aag gac gct gtg cct cag acc ttt tcc gtc ctg atc ggc aac    2160
Ala Glu Lys Asp Ala Val Pro Gln Thr Phe Ser Val Leu Ile Gly Asn
705                 710                 715                 720

aga gaa tgg ctg cgg aga aat ggg ctg aca att tct agt gac gtg tcc    2208
Arg Glu Trp Leu Arg Arg Asn Gly Leu Thr Ile Ser Ser Asp Val Ser
                725                 730                 735

gat gcc atg aca gat cac gag atg aaa ggc cag act gca att ctg gtg    2256
Asp Ala Met Thr Asp His Glu Met Lys Gly Gln Thr Ala Ile Leu Val
            740                 745                 750

gcc atc gac gga gtc ctg tgc ggc atg att gct atc gca gat gcc gtg    2304
Ala Ile Asp Gly Val Leu Cys Gly Met Ile Ala Ile Ala Asp Ala Val
        755                 760                 765

aag cag gag gct gca ctg gcc gtc cat acc ctg cag tct atg ggc gtg    2352
```

```
Lys Gln Glu Ala Ala Leu Ala Val His Thr Leu Gln Ser Met Gly Val
    770                 775                 780

gac gtg gtc ctg atc acc ggg gat aac cgg aaa aca gct aga gca att    2400
Asp Val Val Leu Ile Thr Gly Asp Asn Arg Lys Thr Ala Arg Ala Ile
785                 790                 795                 800

gcc act caa gtg ggc atc aat aag gtg ttc gct gaa gtc ctg cct agc    2448
Ala Thr Gln Val Gly Ile Asn Lys Val Phe Ala Glu Val Leu Pro Ser
                805                 810                 815

cac aag gtc gca aaa gtg cag gag ctg cag aac aag ggc aag aaa gtc    2496
His Lys Val Ala Lys Val Gln Glu Leu Gln Asn Lys Gly Lys Lys Val
            820                 825                 830

gcc atg gtg gga gac ggc gtg aat gat agc cca gct ctg gca cag gca    2544
Ala Met Val Gly Asp Gly Val Asn Asp Ser Pro Ala Leu Ala Gln Ala
        835                 840                 845

gac atg gga gtc gct att ggg aca gga act gac gtg gca atc gag gcc    2592
Asp Met Gly Val Ala Ile Gly Thr Gly Thr Asp Val Ala Ile Glu Ala
    850                 855                 860

gct gat gtg gtc ctg att agg aat gac ctg ctg gat gtg gtc gct tct    2640
Ala Asp Val Val Leu Ile Arg Asn Asp Leu Leu Asp Val Val Ala Ser
865                 870                 875                 880

att cat ctg agt aag agg aca gtg agg cgc att cgc atc aac ctg gtg    2688
Ile His Leu Ser Lys Arg Thr Val Arg Arg Ile Arg Ile Asn Leu Val
                885                 890                 895

ctg gcc ctg atc tac aat ctg gtg gga att cca atc gca gcc ggc gtg    2736
Leu Ala Leu Ile Tyr Asn Leu Val Gly Ile Pro Ile Ala Ala Gly Val
            900                 905                 910

ttt atg cca att ggg atc gtc ctg cag ccc tgg atg ggc tca gct gca    2784
Phe Met Pro Ile Gly Ile Val Leu Gln Pro Trp Met Gly Ser Ala Ala
        915                 920                 925

atg gcc gct tca agc gtg agc gtg gtc ctg tcc tct ctg cag ctg aaa    2832
Met Ala Ala Ser Ser Val Ser Val Val Leu Ser Ser Leu Gln Leu Lys
    930                 935                 940

tgc tac aag aaa cca gac ctg gag cgg tac gaa gct cag gca cac gga    2880
Cys Tyr Lys Lys Pro Asp Leu Glu Arg Tyr Glu Ala Gln Ala His Gly
945                 950                 955                 960

cat atg aag ccc ctg acc gct tcc cag gtg tct gtc cac atc ggc atg    2928
His Met Lys Pro Leu Thr Ala Ser Gln Val Ser Val His Ile Gly Met
                965                 970                 975

gac gat aga tgg agg gac agc cca agg gcc act cca tgg gat cag gtc    2976
Asp Asp Arg Trp Arg Asp Ser Pro Arg Ala Thr Pro Trp Asp Gln Val
            980                 985                 990

agt tac gtg agc cag gtc agc ctg agt tca ctg acc agc gac aag ccc    3024
Ser Tyr Val Ser Gln Val Ser Leu Ser Ser Leu Thr Ser Asp Lys Pro
        995                 1000                1005

tcc cgc cat tct gca gcc gct gat gac gac ggg gac aag tgg agc ctg    3072
Ser Arg His Ser Ala Ala Ala Asp Asp Asp Gly Asp Lys Trp Ser Leu
    1010                1015                1020

ctg ctg aac gga agg gac gaa gaa cag tat atc taa                    3108
Leu Leu Asn Gly Arg Asp Glu Glu Gln Tyr Ile
1025                1030                1035
```

<210> 9

<211> 1035
<212> PRT
<213> Artificial Sequence

<220>
<223> [CDS]:1..3108 from SEQ ID NO 8

<400> 9

```
Met Pro Glu Gln Glu Arg Gln Ile Thr Ala Arg Glu Gly Ala Ser Arg
1               5                  10                  15
Lys Ile Leu Ser Lys Leu Ser Leu Pro Thr Arg Ala Trp Glu Pro Ala
            20                  25                  30
Met Lys Lys Ser Phe Ala Phe Asp Asn Val Gly Tyr Glu Gly Gly Leu
        35                  40                  45
Asp Gly Leu Gly Pro Ser Ser Gln Pro Gln Lys Cys Phe Leu Gln Ile
    50                  55                  60
Lys Gly Met Thr Cys Ala Ser Cys Val Ser Asn Ile Glu Arg Asn Leu
65                  70                  75                  80
Gln Lys Glu Ala Gly Val Leu Ser Val Leu Val Ala Leu Met Ala Gly
                85                  90                  95
Lys Ala Glu Ile Lys Tyr Asp Pro Glu Val Ile Gln Pro Leu Glu Ile
            100                 105                 110
Ala Gln Phe Ile Gln Asp Leu Gly Phe Glu Ala Ala Val Met Glu Asp
        115                 120                 125
Tyr Ala Gly Ser Asp Gly Asn Ile Glu Leu Thr Ile Thr Gly Met Thr
    130                 135                 140
Cys Ala Ser Cys Val His Asn Ile Glu Ser Lys Leu Thr Arg Thr Asn
145                 150                 155                 160
Gly Ile Thr Tyr Ala Ser Val Ala Leu Ala Thr Ser Lys Ala Leu Val
                165                 170                 175
Lys Phe Asp Pro Glu Ile Ile Gly Pro Arg Asp Ile Ile Lys Ile Ile
            180                 185                 190
Glu Glu Ile Gly Phe His Ala Ser Leu Ala Gln Arg Asn Pro Asn Ala
        195                 200                 205
His His Leu Asp His Lys Met Glu Ile Lys Gln Trp Lys Lys Ser Phe
    210                 215                 220
Leu Cys Ser Leu Val Phe Gly Ile Pro Val Met Ala Leu Met Ile Tyr
225                 230                 235                 240
Met Leu Ile Pro Ser Asn Glu Pro His Gln Ser Met Val Leu Asp His
                245                 250                 255
Asn Ile Ile Pro Gly Leu Ser Ile Leu Asn Leu Ile Phe Phe Ile Leu
            260                 265                 270
Cys Thr Phe Val Gln Leu Leu Gly Gly Trp Tyr Phe Tyr Val Gln Ala
        275                 280                 285
Tyr Lys Ser Leu Arg His Arg Ser Ala Asn Met Asp Val Leu Ile Val
    290                 295                 300
Leu Ala Thr Ser Ile Ala Tyr Val Tyr Ser Leu Val Ile Leu Val Val
305                 310                 315                 320
Ala Val Ala Glu Lys Ala Glu Arg Ser Pro Val Thr Phe Phe Asp Thr
                325                 330                 335
Pro Pro Met Leu Phe Val Phe Ile Ala Leu Gly Arg Trp Leu Glu His
            340                 345                 350
Leu Ala Lys Ser Lys Thr Ser Glu Ala Leu Ala Lys Leu Met Ser Leu
        355                 360                 365
Gln Ala Thr Glu Ala Thr Val Val Thr Leu Gly Glu Asp Asn Leu Ile
    370                 375                 380
Ile Arg Glu Glu Gln Val Pro Met Glu Leu Val Gln Arg Gly Asp Ile
385                 390                 395                 400
Val Lys Val Val Pro Gly Gly Lys Phe Pro Val Asp Gly Lys Val Leu
                405                 410                 415
Glu Gly Asn Thr Met Ala Asp Glu Ser Leu Ile Thr Gly Glu Ala Met
            420                 425                 430
Pro Val Thr Lys Lys Pro Gly Ser Thr Val Ile Ala Gly Ser Ile Asn
        435                 440                 445
Ala His Gly Ser Val Leu Ile Lys Ala Thr His Val Gly Asn Asp Thr
    450                 455                 460
Thr Leu Ala Gln Ile Val Lys Leu Val Glu Glu Ala Gln Met Ser Lys
465                 470                 475                 480
```

```
Ala Pro Ile Gln Gln Leu Ala Asp Arg Phe Ser Gly Tyr Phe Val Pro
                485                 490                 495
Phe Ile Ile Ile Met Ser Thr Leu Thr Leu Val Val Trp Ile Val Ile
            500                 505                 510
Gly Phe Ile Asp Phe Gly Val Val Gln Arg Tyr Phe Pro Asn Pro Asn
        515                 520                 525
Lys His Ile Ser Gln Thr Glu Val Ile Ile Arg Phe Ala Phe Gln Thr
    530                 535                 540
Ser Ile Thr Val Leu Cys Ile Ala Cys Pro Cys Ser Leu Gly Leu Ala
545                 550                 555                 560
Thr Pro Thr Ala Val Met Val Gly Thr Gly Val Ala Ala Gln Asn Gly
                565                 570                 575
Ile Leu Ile Lys Gly Gly Lys Pro Leu Glu Met Ala His Lys Ile Lys
            580                 585                 590
Thr Val Met Phe Asp Lys Thr Gly Thr Ile Thr His Gly Val Pro Arg
        595                 600                 605
Val Met Arg Val Leu Leu Leu Gly Asp Val Ala Thr Leu Pro Leu Arg
    610                 615                 620
Lys Val Leu Ala Val Val Gly Thr Ala Glu Ala Ser Ser Glu His Pro
625                 630                 635                 640
Leu Gly Val Ala Val Thr Lys Tyr Cys Lys Glu Glu Leu Gly Thr Glu
                645                 650                 655
Thr Leu Gly Tyr Cys Thr Asp Phe Gln Ala Val Pro Gly Cys Gly Ile
            660                 665                 670
Gly Cys Lys Val Ser Asn Val Glu Gly Ile Leu Ala His Ser Glu Arg
        675                 680                 685
Pro Leu Ser Ala Pro Ala Ser His Leu Asn Glu Ala Gly Ser Leu Pro
    690                 695                 700
Ala Glu Lys Asp Ala Val Pro Gln Thr Phe Ser Val Leu Ile Gly Asn
705                 710                 715                 720
Arg Glu Trp Leu Arg Arg Asn Gly Leu Thr Ile Ser Ser Asp Val Ser
                725                 730                 735
Asp Ala Met Thr Asp His Glu Met Lys Gly Gln Thr Ala Ile Leu Val
            740                 745                 750
Ala Ile Asp Gly Val Leu Cys Gly Met Ile Ala Ile Ala Asp Ala Val
        755                 760                 765
Lys Gln Glu Ala Ala Leu Ala Val His Thr Leu Gln Ser Met Gly Val
    770                 775                 780
Asp Val Val Leu Ile Thr Gly Asp Asn Arg Lys Thr Ala Arg Ala Ile
785                 790                 795                 800
Ala Thr Gln Val Gly Ile Asn Lys Val Phe Ala Glu Val Leu Pro Ser
                805                 810                 815
His Lys Val Ala Lys Val Gln Glu Leu Gln Asn Lys Gly Lys Lys Val
            820                 825                 830
Ala Met Val Gly Asp Gly Val Asn Asp Ser Pro Ala Leu Ala Gln Ala
        835                 840                 845
Asp Met Gly Val Ala Ile Gly Thr Gly Thr Asp Val Ala Ile Glu Ala
    850                 855                 860
Ala Asp Val Val Leu Ile Arg Asn Asp Leu Leu Asp Val Val Ala Ser
865                 870                 875                 880
Ile His Leu Ser Lys Arg Thr Val Arg Arg Ile Arg Ile Asn Leu Val
                885                 890                 895
Leu Ala Leu Ile Tyr Asn Leu Val Gly Ile Pro Ile Ala Ala Gly Val
            900                 905                 910
Phe Met Pro Ile Gly Ile Val Leu Gln Pro Trp Met Gly Ser Ala Ala
        915                 920                 925
Met Ala Ala Ser Ser Val Ser Val Val Leu Ser Ser Leu Gln Leu Lys
    930                 935                 940
Cys Tyr Lys Lys Pro Asp Leu Glu Arg Tyr Glu Ala Gln Ala His Gly
945                 950                 955                 960
His Met Lys Pro Leu Thr Ala Ser Gln Val Ser Val His Ile Gly Met
                965                 970                 975
Asp Asp Arg Trp Arg Asp Ser Pro Arg Ala Thr Pro Trp Asp Gln Val
            980                 985                 990
Ser Tyr Val Ser Gln Val Ser Leu Ser Ser Leu Thr Ser Asp Lys Pro
        995                1000                1005
Ser Arg His Ser Ala Ala Ala Asp Asp Asp Gly Asp Lys Trp Ser Leu
    1010                1015                1020
```

Leu Leu Asn Gly Arg Asp Glu Glu Gln Tyr Ile
1025 1030 1035

<210> 10
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer AAT-Forward

<400> 10
ctggtctaga acgcgtcgcc acccccctcca ccttgg 36

<210> 11
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer AAT-Reverse

<400> 11
atcatgatgc ggccgcttca ctgtcccagg tcagtg 36

<210> 12
<211> 3966
<212> DNA
<213> Artificial Sequence

<220>
<223> Sequence encoding truncated ATP7B(d223-366) that encodes ATP7B with a deletion of amino acids 223..366

<220>
<221> CDS
<222> 1..3966
<223> /transl_table=1

<400> 12
atg cct gag cag gag aga cag atc aca gcc aga gaa ggg gcc agt cgg 48
Met Pro Glu Gln Glu Arg Gln Ile Thr Ala Arg Glu Gly Ala Ser Arg
1 5 10 15

aaa atc tta tct aag ctt tct ttg cct acc cgt gcc tgg gaa cca gca 96
Lys Ile Leu Ser Lys Leu Ser Leu Pro Thr Arg Ala Trp Glu Pro Ala
20 25 30

atg aag aag agt ttt gct ttt gac aat gtt ggc tat gaa ggt ggt ctg 144
Met Lys Lys Ser Phe Ala Phe Asp Asn Val Gly Tyr Glu Gly Gly Leu
35 40 45

gat ggc ctg ggc cct tct tct cag gtg gcc acc agc aca gtc agg atc 192
Asp Gly Leu Gly Pro Ser Ser Gln Val Ala Thr Ser Thr Val Arg Ile
50 55 60

ttg ggc atg act tgc cag tca tgt gtg aag tcc att gag gac agg att 240
Leu Gly Met Thr Cys Gln Ser Cys Val Lys Ser Ile Glu Asp Arg Ile
65 70 75 80

tcc aat ttg aaa ggc atc atc agc atg aag gtt tcc ctg gaa caa ggc 288
Ser Asn Leu Lys Gly Ile Ile Ser Met Lys Val Ser Leu Glu Gln Gly
85 90 95

agt gcc act gtg aaa tat gtg cca tcg gtt gtg tgc ctg caa cag gtt 336
Ser Ala Thr Val Lys Tyr Val Pro Ser Val Val Cys Leu Gln Gln Val

```
                100                 105                 110

tgc cat caa att ggg gac atg ggc ttc gag gcc agc att gca gaa gga     384
Cys His Gln Ile Gly Asp Met Gly Phe Glu Ala Ser Ile Ala Glu Gly
        115                 120                 125

aag gca gcc tcc tgg ccc tca agg tcc ttg cct gcc cag gag gct gtg     432
Lys Ala Ala Ser Trp Pro Ser Arg Ser Leu Pro Ala Gln Glu Ala Val
    130                 135                 140

gtc aag ctc cgg gtg gag ggc atg acc tgc cag tcc tgt gtc agc tcc     480
Val Lys Leu Arg Val Glu Gly Met Thr Cys Gln Ser Cys Val Ser Ser
145                 150                 155                 160

att gaa ggc aag gtc cgg aaa ctg caa gga gta gtg aga gtc aaa gtc     528
Ile Glu Gly Lys Val Arg Lys Leu Gln Gly Val Val Arg Val Lys Val
                165                 170                 175

tca ctc agc aac caa gag gcc gtc atc act tat cag cct tat ctc att     576
Ser Leu Ser Asn Gln Glu Ala Val Ile Thr Tyr Gln Pro Tyr Leu Ile
            180                 185                 190

cag ccc gaa gac ctc agg gac cat gta aat gac atg gga ttt gaa gct     624
Gln Pro Glu Asp Leu Arg Asp His Val Asn Asp Met Gly Phe Glu Ala
        195                 200                 205

gcc atc aag agc aaa gtg gct ccc tta agc ctg gga cca att ggc atg     672
Ala Ile Lys Ser Lys Val Ala Pro Leu Ser Leu Gly Pro Ile Gly Met
    210                 215                 220

acc tgt gca tcc tgt gtc cat tcc att gaa ggc atg atc tcc caa ctg     720
Thr Cys Ala Ser Cys Val His Ser Ile Glu Gly Met Ile Ser Gln Leu
225                 230                 235                 240

gaa ggg gtg cag caa ata tcg gtg tct ttg gcc gaa ggg act gca aca     768
Glu Gly Val Gln Gln Ile Ser Val Ser Leu Ala Glu Gly Thr Ala Thr
                245                 250                 255

gtt ctt tat aat ccc tct gta att agc cca gaa gaa ctc aga gct gct     816
Val Leu Tyr Asn Pro Ser Val Ile Ser Pro Glu Glu Leu Arg Ala Ala
            260                 265                 270

ata gaa gac atg gga ttt gag gct tca gtc gtt tct gaa agc tgt tct     864
Ile Glu Asp Met Gly Phe Glu Ala Ser Val Val Ser Glu Ser Cys Ser
        275                 280                 285

act aac cct ctt gga aac cac agt gct ggg aat tcc atg gtg caa act     912
Thr Asn Pro Leu Gly Asn His Ser Ala Gly Asn Ser Met Val Gln Thr
    290                 295                 300

aca gat ggt aca cct aca tct gtg cag gaa gtg gct ccc cac act ggg     960
Thr Asp Gly Thr Pro Thr Ser Val Gln Glu Val Ala Pro His Thr Gly
305                 310                 315                 320

agg ctc cct gca aac cat gcc ccg gac atc ttg gca aag tcc cca caa    1008
Arg Leu Pro Ala Asn His Ala Pro Asp Ile Leu Ala Lys Ser Pro Gln
                325                 330                 335

tca acc aga gca gtg gca ccg cag aag tgc ttc tta cag atc aaa ggc    1056
Ser Thr Arg Ala Val Ala Pro Gln Lys Cys Phe Leu Gln Ile Lys Gly
            340                 345                 350

atg acc tgt gca tcc tgt gtg tct aac ata gaa agg aat ctg cag aaa    1104
Met Thr Cys Ala Ser Cys Val Ser Asn Ile Glu Arg Asn Leu Gln Lys
        355                 360                 365

gaa gct ggt gtt ctc tcc gtg ttg gtt gcc ttg atg gca gga aag gca    1152
Glu Ala Gly Val Leu Ser Val Leu Val Ala Leu Met Ala Gly Lys Ala
```

```
    370                 375                 380

gag atc aag tat gac cca gag gtc atc cag ccc ctc gag ata gct cag    1200
Glu Ile Lys Tyr Asp Pro Glu Val Ile Gln Pro Leu Glu Ile Ala Gln
385                 390                 395                 400

ttc atc cag gac ctg ggt ttt gag gca gca gtc atg gag gac tac gca    1248
Phe Ile Gln Asp Leu Gly Phe Glu Ala Ala Val Met Glu Asp Tyr Ala
                405                 410                 415

ggc tcc gat ggc aac att gag ctg aca atc aca ggg atg acc tgc gcg    1296
Gly Ser Asp Gly Asn Ile Glu Leu Thr Ile Thr Gly Met Thr Cys Ala
            420                 425                 430

tcc tgt gtc cac aac ata gag tcc aaa ctc acg agg aca aat ggc atc    1344
Ser Cys Val His Asn Ile Glu Ser Lys Leu Thr Arg Thr Asn Gly Ile
        435                 440                 445

act tat gcc tcc gtt gcc ctt gcc acc agc aaa gcc ctt gtt aag ttt    1392
Thr Tyr Ala Ser Val Ala Leu Ala Thr Ser Lys Ala Leu Val Lys Phe
    450                 455                 460

gac ccg gaa att atc ggt cca cgg gat att atc aaa att att gag gaa    1440
Asp Pro Glu Ile Ile Gly Pro Arg Asp Ile Ile Lys Ile Ile Glu Glu
465                 470                 475                 480

att ggc ttt cat gct tcc ctg gcc cag aga aac ccc aac gct cat cac    1488
Ile Gly Phe His Ala Ser Leu Ala Gln Arg Asn Pro Asn Ala His His
                485                 490                 495

ttg gac cac aag atg gaa ata aag cag tgg aag aag tct ttc ctg tgc    1536
Leu Asp His Lys Met Glu Ile Lys Gln Trp Lys Lys Ser Phe Leu Cys
            500                 505                 510

agc ctg gtg ttt ggc atc cct gtc atg gcc tta atg atc tat atg ctg    1584
Ser Leu Val Phe Gly Ile Pro Val Met Ala Leu Met Ile Tyr Met Leu
        515                 520                 525

ata ccc agc aac gag ccc cac cag tcc atg gtc ctg gac cac aac atc    1632
Ile Pro Ser Asn Glu Pro His Gln Ser Met Val Leu Asp His Asn Ile
    530                 535                 540

att cca gga ctg tcc att cta aat ctc atc ttc ttt atc ttg tgt acc    1680
Ile Pro Gly Leu Ser Ile Leu Asn Leu Ile Phe Phe Ile Leu Cys Thr
545                 550                 555                 560

ttt gtc cag ctc ctc ggt ggg tgg tac ttc tac gtt cag gcc tac aaa    1728
Phe Val Gln Leu Leu Gly Gly Trp Tyr Phe Tyr Val Gln Ala Tyr Lys
                565                 570                 575

tct ctg aga cac agg tca gcc aac atg gac gtg ctc atc gtc ctg gcc    1776
Ser Leu Arg His Arg Ser Ala Asn Met Asp Val Leu Ile Val Leu Ala
            580                 585                 590

aca agc att gct tat gtt tat tct ctg gtc atc ctg gtg gtt gct gtg    1824
Thr Ser Ile Ala Tyr Val Tyr Ser Leu Val Ile Leu Val Val Ala Val
        595                 600                 605

gct gag aag gcg gag agg agc cct gtg aca ttc ttc gac acg ccc ccc    1872
Ala Glu Lys Ala Glu Arg Ser Pro Val Thr Phe Phe Asp Thr Pro Pro
    610                 615                 620

atg ctc ttt gtg ttc att gcc ctg ggc cgg tgg ctg gaa cac ttg gca    1920
Met Leu Phe Val Phe Ile Ala Leu Gly Arg Trp Leu Glu His Leu Ala
625                 630                 635                 640

aag agc aaa acc tca gaa gcc ctg gct aaa ctc atg tct ctc caa gcc    1968
Lys Ser Lys Thr Ser Glu Ala Leu Ala Lys Leu Met Ser Leu Gln Ala
```

```
                645                 650                 655

aca gaa gcc acc gtt gtg acc ctt ggt gag gac aat tta atc atc agg    2016
Thr Glu Ala Thr Val Val Thr Leu Gly Glu Asp Asn Leu Ile Ile Arg
            660                 665                 670

gag gag caa gtc ccc atg gag ctg gtg cag cgg ggc gat atc gtc aag    2064
Glu Glu Gln Val Pro Met Glu Leu Val Gln Arg Gly Asp Ile Val Lys
        675                 680                 685

gtg gtc cct ggg gga aag ttt cca gtg gat ggg aaa gtc ctg gaa ggc    2112
Val Val Pro Gly Gly Lys Phe Pro Val Asp Gly Lys Val Leu Glu Gly
    690                 695                 700

aat acc atg gct gat gag tcc ctc atc aca gga gaa gcc atg cca gtc    2160
Asn Thr Met Ala Asp Glu Ser Leu Ile Thr Gly Glu Ala Met Pro Val
705                 710                 715                 720

act aag aaa ccc gga agc act gta att gcg ggg tct ata aat gca cat    2208
Thr Lys Lys Pro Gly Ser Thr Val Ile Ala Gly Ser Ile Asn Ala His
                725                 730                 735

ggc tct gtg ctc att aaa gct acc cac gtg ggc aat gac acc act ttg    2256
Gly Ser Val Leu Ile Lys Ala Thr His Val Gly Asn Asp Thr Thr Leu
            740                 745                 750

gct cag att gtg aaa ctg gtg gaa gag gct cag atg tca aag gca ccc    2304
Ala Gln Ile Val Lys Leu Val Glu Glu Ala Gln Met Ser Lys Ala Pro
        755                 760                 765

att cag cag ctg gct gac cgg ttt agt gga tat ttt gtc cca ttt atc    2352
Ile Gln Gln Leu Ala Asp Arg Phe Ser Gly Tyr Phe Val Pro Phe Ile
    770                 775                 780

atc atc atg tca act ttg acg ttg gtg gta tgg att gta atc ggt ttt    2400
Ile Ile Met Ser Thr Leu Thr Leu Val Val Trp Ile Val Ile Gly Phe
785                 790                 795                 800

atc gat ttt ggt gtt gtt cag aga tac ttt cct aac ccc aac aag cac    2448
Ile Asp Phe Gly Val Val Gln Arg Tyr Phe Pro Asn Pro Asn Lys His
                805                 810                 815

atc tcc cag aca gag gtg atc atc cgg ttt gct ttc cag acg tcc atc    2496
Ile Ser Gln Thr Glu Val Ile Ile Arg Phe Ala Phe Gln Thr Ser Ile
            820                 825                 830

acg gtg ctg tgc att gcc tgc ccc tgc tcc ctg ggg ctg gcc acg ccc    2544
Thr Val Leu Cys Ile Ala Cys Pro Cys Ser Leu Gly Leu Ala Thr Pro
        835                 840                 845

acg gct gtc atg gtg ggc acc ggg gtg gcc gcg cag aac ggc atc ctc    2592
Thr Ala Val Met Val Gly Thr Gly Val Ala Ala Gln Asn Gly Ile Leu
    850                 855                 860

atc aag gga ggc aag ccc ctg gag atg gcg cac aag ata aag act gtg    2640
Ile Lys Gly Gly Lys Pro Leu Glu Met Ala His Lys Ile Lys Thr Val
865                 870                 875                 880

atg ttt gac aag act ggc acc att acc cat ggc gtc ccc agg gtc atg    2688
Met Phe Asp Lys Thr Gly Thr Ile Thr His Gly Val Pro Arg Val Met
                885                 890                 895

cgg gtg ctc ctg ctg ggg gat gtg gcc aca ctg ccc ctc agg aag gtt    2736
Arg Val Leu Leu Leu Gly Asp Val Ala Thr Leu Pro Leu Arg Lys Val
            900                 905                 910

ctg gct gtg gtg ggg act gcg gag gcc agc agt gaa cac ccc ttg ggc    2784
Leu Ala Val Val Gly Thr Ala Glu Ala Ser Ser Glu His Pro Leu Gly
```

```
        915                 920                 925

gtg gca gtc acc aaa tac tgt aaa gag gaa ctt gga aca gag acc ttg    2832
Val Ala Val Thr Lys Tyr Cys Lys Glu Glu Leu Gly Thr Glu Thr Leu
    930                 935                 940

gga tac tgc acg gac ttc cag gca gtg cca ggc tgt gga att ggg tgc    2880
Gly Tyr Cys Thr Asp Phe Gln Ala Val Pro Gly Cys Gly Ile Gly Cys
945                 950                 955                 960

aaa gtc agc aac gtg gaa ggc atc ctg gcc cac agt gag cgc cct ttg    2928
Lys Val Ser Asn Val Glu Gly Ile Leu Ala His Ser Glu Arg Pro Leu
                965                 970                 975

agt gca ccg gcc agt cac ctg aat gag gct ggc agc ctt ccc gca gaa    2976
Ser Ala Pro Ala Ser His Leu Asn Glu Ala Gly Ser Leu Pro Ala Glu
            980                 985                 990

aaa gat gca gtc ccc cag acc ttc tct gtg ctg att gga aac cgt gag    3024
Lys Asp Ala Val Pro Gln Thr Phe Ser Val Leu Ile Gly Asn Arg Glu
        995                 1000                1005

tgg ctg agg cgc aac ggt tta acc att tct agc gat gtc agt gac gct    3072
Trp Leu Arg Arg Asn Gly Leu Thr Ile Ser Ser Asp Val Ser Asp Ala
    1010                1015                1020

atg aca gac cac gag atg aaa gga cag aca gcc atc ctg gtg gct att    3120
Met Thr Asp His Glu Met Lys Gly Gln Thr Ala Ile Leu Val Ala Ile
1025                1030                1035                1040

gac ggt gtg ctc tgt ggg atg atc gca atc gca gac gct gtc aag cag    3168
Asp Gly Val Leu Cys Gly Met Ile Ala Ile Ala Asp Ala Val Lys Gln
                1045                1050                1055

gag gct gcc ctg gct gtg cac acg ctg cag agc atg ggt gtg gac gtg    3216
Glu Ala Ala Leu Ala Val His Thr Leu Gln Ser Met Gly Val Asp Val
            1060                1065                1070

gtt ctg atc acg ggg gac aac cgg aag aca gcc aga gct att gcc acc    3264
Val Leu Ile Thr Gly Asp Asn Arg Lys Thr Ala Arg Ala Ile Ala Thr
        1075                1080                1085

cag gtt ggc atc aac aaa gtc ttt gca gag gtg ctg cct tcg cac aag    3312
Gln Val Gly Ile Asn Lys Val Phe Ala Glu Val Leu Pro Ser His Lys
    1090                1095                1100

gtg gcc aag gtc cag gag ctc cag aat aaa ggg aag aaa gtc gcc atg    3360
Val Ala Lys Val Gln Glu Leu Gln Asn Lys Gly Lys Lys Val Ala Met
1105                1110                1115                1120

gtg ggg gat ggg gtc aat gac tcc ccg gcc ttg gcc cag gca gac atg    3408
Val Gly Asp Gly Val Asn Asp Ser Pro Ala Leu Ala Gln Ala Asp Met
                1125                1130                1135

ggt gtg gcc att ggc acc ggc acg gat gtg gcc atc gag gca gcc gac    3456
Gly Val Ala Ile Gly Thr Gly Thr Asp Val Ala Ile Glu Ala Ala Asp
            1140                1145                1150

gtc gtc ctt atc aga aat gat ttg ctg gat gtg gtg gct agc att cac    3504
Val Val Leu Ile Arg Asn Asp Leu Leu Asp Val Val Ala Ser Ile His
        1155                1160                1165

ctt tcc aag agg act gtc cga agg ata cgc atc aac ctg gtc ctg gca    3552
Leu Ser Lys Arg Thr Val Arg Arg Ile Arg Ile Asn Leu Val Leu Ala
    1170                1175                1180

ctg att tat aac ctg gtt ggg ata ccc att gca gca ggt gtc ttc atg    3600
Leu Ile Tyr Asn Leu Val Gly Ile Pro Ile Ala Ala Gly Val Phe Met
```

```
1185                1190                1195                1200

ccc atc ggc att gtg ctg cag ccc tgg atg ggc tca gcg gcc atg gca    3648
Pro Ile Gly Ile Val Leu Gln Pro Trp Met Gly Ser Ala Ala Met Ala
                1205                1210                1215

gcc tcc tct gtg tct gtg gtg ctc tca tcc ctg cag ctc aag tgc tat    3696
Ala Ser Ser Val Ser Val Val Leu Ser Ser Leu Gln Leu Lys Cys Tyr
            1220                1225                1230

aag aag cct gac ctg gag agg tat gag gca cag gcg cat ggc cac atg    3744
Lys Lys Pro Asp Leu Glu Arg Tyr Glu Ala Gln Ala His Gly His Met
        1235                1240                1245

aag ccc ctg acg gca tcc cag gtc agt gtg cac ata ggc atg gat gac    3792
Lys Pro Leu Thr Ala Ser Gln Val Ser Val His Ile Gly Met Asp Asp
    1250                1255                1260

agg tgg cgg gac tcc ccc agg gcc aca cca tgg gac cag gtc agc tat    3840
Arg Trp Arg Asp Ser Pro Arg Ala Thr Pro Trp Asp Gln Val Ser Tyr
1265                1270                1275                1280

gtc agc cag gtg tcg ctg tcc tcc ctg acg tcc gac aag cca tct cgg    3888
Val Ser Gln Val Ser Leu Ser Ser Leu Thr Ser Asp Lys Pro Ser Arg
                1285                1290                1295

cac agc gct gca gca gac gat gat ggg gac aag tgg tct ctg ctc ctg    3936
His Ser Ala Ala Ala Asp Asp Asp Gly Asp Lys Trp Ser Leu Leu Leu
            1300                1305                1310

aat ggc agg gat gag gag cag tac atc tga                            3966
Asn Gly Arg Asp Glu Glu Gln Tyr Ile
        1315                1320
```

<210> 13
<211> 1321
<212> PRT
<213> Artificial Sequence

<220>
<223> [CDS]:1..3966 from SEQ ID NO 12

<400> 13

```
Met Pro Glu Gln Glu Arg Gln Ile Thr Ala Arg Glu Gly Ala Ser Arg
1               5                   10                  15
Lys Ile Leu Ser Lys Leu Ser Leu Pro Thr Arg Ala Trp Glu Pro Ala
            20                  25                  30
Met Lys Lys Ser Phe Ala Phe Asp Asn Val Gly Tyr Glu Gly Gly Leu
        35                  40                  45
Asp Gly Leu Gly Pro Ser Ser Gln Val Ala Thr Ser Thr Val Arg Ile
    50                  55                  60
Leu Gly Met Thr Cys Gln Ser Cys Val Lys Ser Ile Glu Asp Arg Ile
65                  70                  75                  80
Ser Asn Leu Lys Gly Ile Ile Ser Met Lys Val Ser Leu Glu Gln Gly
                85                  90                  95
Ser Ala Thr Val Lys Tyr Val Pro Ser Val Val Cys Leu Gln Gln Val
            100                 105                 110
Cys His Gln Ile Gly Asp Met Gly Phe Glu Ala Ser Ile Ala Glu Gly
        115                 120                 125
Lys Ala Ala Ser Trp Pro Ser Arg Ser Leu Pro Ala Gln Glu Ala Val
    130                 135                 140
Val Lys Leu Arg Val Glu Gly Met Thr Cys Gln Ser Cys Val Ser Ser
145                 150                 155                 160
Ile Glu Gly Lys Val Arg Lys Leu Gln Gly Val Val Arg Val Lys Val
                165                 170                 175
Ser Leu Ser Asn Gln Glu Ala Val Ile Thr Tyr Gln Pro Tyr Leu Ile
            180                 185                 190
Gln Pro Glu Asp Leu Arg Asp His Val Asn Asp Met Gly Phe Glu Ala
```

```
        195                 200                 205
Ala Ile Lys Ser Lys Val Ala Pro Leu Ser Leu Gly Pro Ile Gly Met
    210                 215                 220
Thr Cys Ala Ser Cys Val His Ser Ile Glu Gly Met Ile Ser Gln Leu
225                 230                 235                 240
Glu Gly Val Gln Gln Ile Ser Val Ser Leu Ala Glu Gly Thr Ala Thr
                245                 250                 255
Val Leu Tyr Asn Pro Ser Val Ile Ser Pro Glu Glu Leu Arg Ala Ala
            260                 265                 270
Ile Glu Asp Met Gly Phe Glu Ala Ser Val Val Ser Glu Ser Cys Ser
        275                 280                 285
Thr Asn Pro Leu Gly Asn His Ser Ala Gly Asn Ser Met Val Gln Thr
    290                 295                 300
Thr Asp Gly Thr Pro Thr Ser Val Gln Glu Val Ala Pro His Thr Gly
305                 310                 315                 320
Arg Leu Pro Ala Asn His Ala Pro Asp Ile Leu Ala Lys Ser Pro Gln
                325                 330                 335
Ser Thr Arg Ala Val Ala Pro Gln Lys Cys Phe Leu Gln Ile Lys Gly
            340                 345                 350
Met Thr Cys Ala Ser Cys Val Ser Asn Ile Glu Arg Asn Leu Gln Lys
        355                 360                 365
Glu Ala Gly Val Leu Ser Val Leu Val Ala Leu Met Ala Gly Lys Ala
    370                 375                 380
Glu Ile Lys Tyr Asp Pro Glu Val Ile Gln Pro Leu Glu Ile Ala Gln
385                 390                 395                 400
Phe Ile Gln Asp Leu Gly Phe Glu Ala Ala Val Met Glu Asp Tyr Ala
                405                 410                 415
Gly Ser Asp Gly Asn Ile Glu Leu Thr Ile Thr Gly Met Thr Cys Ala
            420                 425                 430
Ser Cys Val His Asn Ile Glu Ser Lys Leu Thr Arg Thr Asn Gly Ile
        435                 440                 445
Thr Tyr Ala Ser Val Ala Leu Ala Thr Ser Lys Ala Leu Val Lys Phe
    450                 455                 460
Asp Pro Glu Ile Ile Gly Pro Arg Asp Ile Ile Lys Ile Ile Glu Glu
465                 470                 475                 480
Ile Gly Phe His Ala Ser Leu Ala Gln Arg Asn Pro Asn Ala His His
                485                 490                 495
Leu Asp His Lys Met Glu Ile Lys Gln Trp Lys Lys Ser Phe Leu Cys
            500                 505                 510
Ser Leu Val Phe Gly Ile Pro Val Met Ala Leu Met Ile Tyr Met Leu
        515                 520                 525
Ile Pro Ser Asn Glu Pro His Gln Ser Met Val Leu Asp His Asn Ile
    530                 535                 540
Ile Pro Gly Leu Ser Ile Leu Asn Leu Ile Phe Phe Ile Leu Cys Thr
545                 550                 555                 560
Phe Val Gln Leu Leu Gly Gly Trp Tyr Phe Tyr Val Gln Ala Tyr Lys
                565                 570                 575
Ser Leu Arg His Arg Ser Ala Asn Met Asp Val Leu Ile Val Leu Ala
            580                 585                 590
Thr Ser Ile Ala Tyr Val Tyr Ser Leu Val Ile Leu Val Val Ala Val
        595                 600                 605
Ala Glu Lys Ala Glu Arg Ser Pro Val Thr Phe Phe Asp Thr Pro Pro
    610                 615                 620
Met Leu Phe Val Phe Ile Ala Leu Gly Arg Trp Leu Glu His Leu Ala
625                 630                 635                 640
Lys Ser Lys Thr Ser Glu Ala Leu Ala Lys Leu Met Ser Leu Gln Ala
                645                 650                 655
Thr Glu Ala Thr Val Val Thr Leu Gly Glu Asp Asn Leu Ile Ile Arg
            660                 665                 670
Glu Glu Gln Val Pro Met Glu Leu Val Gln Arg Gly Asp Ile Val Lys
        675                 680                 685
Val Val Pro Gly Gly Lys Phe Pro Val Asp Gly Lys Val Leu Glu Gly
    690                 695                 700
Asn Thr Met Ala Asp Glu Ser Leu Ile Thr Gly Glu Ala Met Pro Val
705                 710                 715                 720
Thr Lys Lys Pro Gly Ser Thr Val Ile Ala Gly Ser Ile Asn Ala His
                725                 730                 735
Gly Ser Val Leu Ile Lys Ala Thr His Val Gly Asn Asp Thr Thr Leu
```

```
            740                 745                 750
Ala Gln Ile Val Lys Leu Val Glu Glu Ala Gln Met Ser Lys Ala Pro
        755                 760                 765
Ile Gln Gln Leu Ala Asp Arg Phe Ser Gly Tyr Phe Val Pro Phe Ile
    770                 775                 780
Ile Ile Met Ser Thr Leu Thr Leu Val Val Trp Ile Val Ile Gly Phe
785                 790                 795                 800
Ile Asp Phe Gly Val Val Gln Arg Tyr Phe Pro Asn Pro Asn Lys His
                805                 810                 815
Ile Ser Gln Thr Glu Val Ile Ile Arg Phe Ala Phe Gln Thr Ser Ile
            820                 825                 830
Thr Val Leu Cys Ile Ala Cys Pro Cys Ser Leu Gly Leu Ala Thr Pro
        835                 840                 845
Thr Ala Val Met Val Gly Thr Gly Val Ala Ala Gln Asn Gly Ile Leu
    850                 855                 860
Ile Lys Gly Gly Lys Pro Leu Glu Met Ala His Lys Ile Lys Thr Val
865                 870                 875                 880
Met Phe Asp Lys Thr Gly Thr Ile Thr His Gly Val Pro Arg Val Met
                885                 890                 895
Arg Val Leu Leu Leu Gly Asp Val Ala Thr Leu Pro Leu Arg Lys Val
            900                 905                 910
Leu Ala Val Val Gly Thr Ala Glu Ala Ser Ser Glu His Pro Leu Gly
        915                 920                 925
Val Ala Val Thr Lys Tyr Cys Lys Glu Glu Leu Gly Thr Glu Thr Leu
    930                 935                 940
Gly Tyr Cys Thr Asp Phe Gln Ala Val Pro Gly Cys Gly Ile Gly Cys
945                 950                 955                 960
Lys Val Ser Asn Val Glu Gly Ile Leu Ala His Ser Glu Arg Pro Leu
                965                 970                 975
Ser Ala Pro Ala Ser His Leu Asn Glu Ala Gly Ser Leu Pro Ala Glu
            980                 985                 990
Lys Asp Ala Val Pro Gln Thr Phe Ser Val Leu Ile Gly Asn Arg Glu
        995                 1000                1005
Trp Leu Arg Arg Asn Gly Leu Thr Ile Ser Ser Asp Val Ser Asp Ala
    1010                1015                1020
Met Thr Asp His Glu Met Lys Gly Gln Thr Ala Ile Leu Val Ala Ile
1025                1030                1035                1040
Asp Gly Val Leu Cys Gly Met Ile Ala Ile Ala Asp Ala Val Lys Gln
                1045                1050                1055
Glu Ala Ala Leu Ala Val His Thr Leu Gln Ser Met Gly Val Asp Val
            1060                1065                1070
Val Leu Ile Thr Gly Asp Asn Arg Lys Thr Ala Arg Ala Ile Ala Thr
        1075                1080                1085
Gln Val Gly Ile Asn Lys Val Phe Ala Glu Val Leu Pro Ser His Lys
    1090                1095                1100
Val Ala Lys Val Gln Glu Leu Gln Asn Lys Gly Lys Lys Val Ala Met
1105                1110                1115                1120
Val Gly Asp Gly Val Asn Asp Ser Pro Ala Leu Ala Gln Ala Asp Met
                1125                1130                1135
Gly Val Ala Ile Gly Thr Gly Thr Asp Val Ala Ile Glu Ala Ala Asp
            1140                1145                1150
Val Val Leu Ile Arg Asn Asp Leu Leu Asp Val Val Ala Ser Ile His
        1155                1160                1165
Leu Ser Lys Arg Thr Val Arg Arg Ile Arg Ile Asn Leu Val Leu Ala
    1170                1175                1180
Leu Ile Tyr Asn Leu Val Gly Ile Pro Ile Ala Ala Gly Val Phe Met
1185                1190                1195                1200
Pro Ile Gly Ile Val Leu Gln Pro Trp Met Gly Ser Ala Ala Met Ala
                1205                1210                1215
Ala Ser Ser Val Ser Val Val Leu Ser Ser Leu Gln Leu Lys Cys Tyr
            1220                1225                1230
Lys Lys Pro Asp Leu Glu Arg Tyr Glu Ala Gln Ala His Gly His Met
        1235                1240                1245
Lys Pro Leu Thr Ala Ser Gln Val Ser Val His Ile Gly Met Asp Asp
    1250                1255                1260
Arg Trp Arg Asp Ser Pro Arg Ala Thr Pro Trp Asp Gln Val Ser Tyr
1265                1270                1275                1280
Val Ser Gln Val Ser Leu Ser Ser Leu Thr Ser Asp Lys Pro Ser Arg
```

eolf-seql

1285 1290 1295

His Ser Ala Ala Ala Asp Asp Asp Gly Asp Lys Trp Ser Leu Leu Leu

1300 1305 1310

Asn Gly Arg Asp Glu Glu Gln Tyr Ile

1315 1320

<210> 14
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer F1

<400> 14
ctagatgcgg ccgccaccat gcctg 25

<210> 15
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer R1

<400> 15
ctgagaagaa gggcccaggc c 21

<210> 16
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer F2

<400> 16
ggcccttctt ctcagccgca gaagtgcttc ttacag 36

<210> 17
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer R2

<400> 17
accaaaatcg ataaaaccga ttacaatcc 29

<210> 18
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer F3

<400> 18
acgcgtgcgg ccgccaccat gccag 25

<210> 19
<211> 21

<212> DNA
<213> Artificial Sequence

<220>
<223> Primer R3

<400> 19
ctgggagcta ggtcccagtc c 21

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

<220>
<223> Primer F4

<400> 20
ggacctagct cccagcctca gaagtgtttt ctgcag 36

<210> 21
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer R4

<400> 21
tgttcctcgc gaatgatcag gttgtcctc 29