

SEQUENCE LISTING

<110> ULTRAGENYX PHARMACEUTICAL INC.

<120> GENE THERAPY CONSTRUCTS FOR TREATING WILSON DISEASE

<130> ULP-003WO

<150> 62/788,324

<151> 2019-01-04

<150> 62/834,830

<151> 2019-04-16

<160> 15

<170> PatentIn version 3.5

<210> 1

<211> 3603

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of artificial sequence: Del 1-3 ATP7B, native

<220>

<221> misc_feature

<222> (223)..(225)

<223> Nucleotides 223-225 encode for serine residue S340 (numbering based on wild-type protein sequence)

<220>

<221> misc_feature

<222> (226)..(228)

<223> Nucleotides 226-228 encode for serine residue S341 (numbering based on wild-type protein sequence)

<400> 1

atgcctgagc aggagagaca gatcacagcc agagaagggg ccagtcggaa aatcttatct 60

aagctttctt tgcctacccg tgcctgggaa ccagcaatga agaagagttt tgcttttgac 120

aatgttggt atgaaggtgg tctggatggc ctgggccctt cttctcaggt ggccaccagc 180

acagtcaggg atggagccga agggagtggg acagatcaca ggtcttccag ttctcattcc 240

cctgggtccc caccgagaaa ccaggtccag ggcacatgca gtaccactct gattgccatt 300

gccggcatga cctgtgcatc ctgtgtccat tccattgaag gcatgatctc ccaactggaa 360

ggggtgcagc aaatatcggg gtctttggcc gaagggactg caacagttct ttataatccc 420

tctgtaatta gccagaaga actcagagct gctatagaag acatgggatt tgaggcttca	480
gtcgttttctg aaagctgttc tactaaccct cttggaaacc acagtgtctgg gaattccatg	540
gtgcaaacta cagatgggtac acctacatct gtgcaggaag tggctcccca cactgggagg	600
ctccctgcaa accatgcccc ggacatcttg gcaaagtccc cacaatcaac cagagcagtg	660
gcaccgcaga agtgcttctt acagatcaaa ggcatgacct gtgcatcctg tgtgtctaac	720
atagaaagga atctgcagaa agaagctggg gttctctccg tgttggttgc cttgatggca	780
ggaaaggcag agatcaagta tgaccagag gtcattcagc ccctcgagat agctcagttc	840
atccaggacc tgggttttga ggacagcagtc atggaggact acgcaggctc cgatggcaac	900
attgagctga caatcacagg gatgacctgc gcgtcctgtg tccacaacat agagtccaaa	960
ctcacgagga caaatggcat cacttatgcc tccgttgccc ttgccaccag caaagccctt	1020
gttaagtttg acccggaat tatcggtcca cgggatatta tcaaaattat tgaggaaatt	1080
ggctttcatg cttccctggc ccagagaaac cccaacgctc atcacttggg ccacaagatg	1140
gaaataaagc agtggagaa gtctttcctg tgcagcctgg tgtttggcat ccctgtcatg	1200
gccttaatga tctatatgct gataccagc aacgagcccc accagtccat ggtcctggac	1260
cacaacatca ttccaggact gtccattcta aatctcatct tctttatctt gtgtaccttt	1320
gtccagctcc tcggtgggtg gtacttctac gttcaggcct acaaattctt gagacacagg	1380
tcagccaaca tggacgtgct catcgtcctg gccacaagca ttgcttatgt ttattctctg	1440
gtcatcctgg tggttgctgt ggctgagaag gcggagagga gccctgtgac attcttcgac	1500
acgccccca tgctctttgt gttcattgcc ctgggccggt ggctggaaca cttggcaaag	1560
agcaaacct cagaagccct ggctaaactc atgtctctcc aagccacaga agccaccgtt	1620
gtgacccttg gtgaggacaa tttaatcatc agggaggagc aagtcccat ggagctgggtg	1680
cagcggggcg atatcgtcaa ggtgggtccct gggggaaagt ttccagtgga tgggaaagtc	1740
ctggaaggca ataccatggc tgatgagtcc ctcatcacag gagaagccat gccagtcact	1800
aagaaacccg gaagcactgt aattgcgggg tctataaatg cacatggctc tgtgtcatt	1860
aaagctaccc acgtgggcaa tgacaccact ttggctcaga ttgtgaaact ggtggaagag	1920
gctcagatgt caaaggcacc cattcagcag ctggctgacc ggtttagtgg atattttgtc	1980

ccatttatca tcatcatgtc aactttgacg ttggtggtat ggattgtaat cggttttatc	2040
gatttttggtg ttgttcagag atactttcct aacccaaca agcacatctc ccagacagag	2100
gtgatcatcc ggtttgcttt ccagacgtcc atcacggtgc tgtgcattgc ctgcccctgc	2160
tccctggggc tggccacgcc cacggctgtc atggtgggca ccgggggtggc cgcgcagaac	2220
ggcatcctca tcaagggagg caagcccctg gagatggcgc acaagataaa gactgtgatg	2280
tttgacaaga ctggcaccat tacccatggc gtccccaggg tcatgcggtt gctcctgctg	2340
ggggatgtgg ccacactgcc cctcaggaag gttctggctg tgggtggggac tgcggaggcc	2400
agcagtgaac accccttggg cgtggcagtc accaaatact gtaaagagga acttggaaca	2460
gagaccttgg gatactgcac ggacttccag gcagtgccag gctgtggaat tgggtgcaaa	2520
gtcagcaacg tggaaggcat cctggcccac agtgagcgcc ctttgagtgc accggccagt	2580
cacctgaatg aggctggcag cttcccgca gaaaaagatg cagtcccca gaccttctct	2640
gtgctgattg gaaaccgtga gtggctgagg cgcaacggtt taaccatttc tagcgatgtc	2700
agtgacgcta tgacagacca cgagatgaaa ggacagacag ccaccttggg ggctattgac	2760
ggtgtgctct gtgggatgat cgcaatcgca gacgctgtca agcaggaggc tgccctggct	2820
gtgcacacgc tgacagagcat ggggtgtggac gtggttctga tcacggggga caaccggaag	2880
acagccagag ctattgccac ccagggttggc atcaacaaag tctttgcaga ggtgctgcct	2940
tcgcacaagg tggccaaggt ccaggagctc cagaataaag ggaagaaagt cgccatgggtg	3000
ggggatgggg tcaatgactc cccggccttg gccaggcag acatgggtgt ggccattggc	3060
accggcacgg atgtggccat cgaggcagcc gacgtcgtcc ttatcagaaa tgatttgctg	3120
gatgtgggtg ctagcattca cttttccaag aggactgtcc gaaggatacg catcaacctg	3180
gtcctggcac tgatttataa cctggttggg ataccattg cagcaggtgt cttcatgccc	3240
atcggcattg tgctgcagcc ctggatgggc tcagcggcca tggcagcctc ctctgtgtct	3300
gtggtgctct catccctgca gctcaagtgc tataagaagc ctgacctgga gaggtatgag	3360
gcacaggcgc atggccacat gaagcccctg acggcatccc aggtcagtgt gcacataggc	3420
atggatgaca ggtggcggga ctccccagg gccacacat gggaccaggc cagctatgtc	3480
agccagggtg cgctgtcctc cctgacgtcc gacaagccat ctcggcacag cgctgcagca	3540
gacgatgatg gggacaagtg gtctctgtc ctgaatggca gggatgagga gcagtacatc	3600

tga 3603

<210> 2
<211> 145
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence: AAV2 ITR

<400> 2
ttggccactc cctctctgcg cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc 60
cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgag agagggagtg 120
gccaactcca tcactagggg ttcct 145

<210> 3
<211> 100
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence: enTTR Enhancer sequence

<400> 3
ctacctgtg atgcccggc ccctgttcaa acatgtccta atactctgtc tctgcaaggg 60
tcatacgtag ttttccatct tactcaacat cctcccagtg 100

<210> 4
<211> 95
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence: SV40 Small T Intron

<400> 4
gctctaaggt aaatataaaa tttttaagtg tataatgtgt taaactactg attctaattg 60
tttctctctt ttagattcca acctttggaa ctgat 95

<210> 5
<211> 315
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of artificial sequence: rHBB Intron

<400> 5

gtgagcgggc gggacggccc ttctcctccg ggctgtaatt agcgcttggt ttaatgacgg 60

cttgtttctt ttctgtggct gcgtgaaagc cttgaggggc tccgggaggg ccctttgtgc 120

ggggggagcg gctcggggct gtccgcgggg ggacggctgc cttcgggggg gacggggcag 180

ggcgggggttc ggcttctggc gtgtgaccgg cggctcaaga gcctctgcta accatgttca 240

tgcttcttc ttttcttac agctcctggg caacgtgctg gttattgtgc tgtctcatca 300

ttttggcaaa gaatt 315

<210> 6

<211> 215

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of artificial sequence: BGH polyadenylation signal sequence

<400> 6

gcctcgactg tgccttctag ttgccagcca tctgttgttt gcccctcccc cgtgccttcc 60

ttgaccctgg aaggtgccac tcccactgtc ctttcctaataaaaatgagga aattgcatcg 120

cattgtctga gtaggtgtca ttctattctg ggggggtgggg tggggcagga cagcaagggg 180

gaggattggg aagacaatag caggcatgct gggga 215

<210> 7

<211> 198

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of artificial sequence: SV40 polyadenylation signal sequence

<400> 7

gatccagaca tgataagata cattgatgag tttggacaaa ccacaactag aatgcagtga 60

aaaaaatgct ttatttgtga aatttgtgat gctattgctt tatttgtaac cattataagc 120

tgcaataaac aagttaacaa caacaattgc attcatttta tgtttcaggt tcagggggag 180

gtgtgggagg ttttttag 198

<210> 8
 <211> 1200
 <212> PRT
 <213> Artificial Sequence

 <220>
 <223> Description of artificial sequence: DEL 1-3 ATP7B Native, protein

<220>
 <221> MISC_FEATURE
 <222> (75)..(75)
 <223> Amino acid residue at position 75 corresponds to Serine S340
 (numbering based on wild-type protein sequence)

<220>
 <221> MISC_FEATURE
 <222> (76)..(76)
 <223> Amino acid residue at position 76 corresponds to Serine S341
 (numbering based on wild-type protein sequence)

<400> 8

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 Asp
 50 55 60

Gly Ala Glu Gly Ser Gly Thr Asp His Arg Ser Ser Ser Ser His Ser
 65 70 75 80

Pro Gly Ser Pro Pro Arg Asn Gln Val Gln Gly Thr Cys Ser Thr Thr
 85 90 95

Leu Ile Ala Ile Ala Gly Met Thr Cys Ala Ser Cys Val His Ser Ile
 100 105 110

Glu Gly Met Ile Ser Gln Leu Glu Gly Val Gln Gln Ile Ser Val Ser
115 120 125

Leu Ala Glu Gly Thr Ala Thr Val Leu Tyr Asn Pro Ser Val Ile Ser
130 135 140

Pro Glu Glu Leu Arg Ala Ala Ile Glu Asp Met Gly Phe Glu Ala Ser
145 150 155 160

Val Val Ser Glu Ser Cys Ser Thr Asn Pro Leu Gly Asn His Ser Ala
165 170 175

Gly Asn Ser Met Val Gln Thr Thr Asp Gly Thr Pro Thr Ser Val Gln
180 185 190

Glu Val Ala Pro His Thr Gly Arg Leu Pro Ala Asn His Ala Pro Asp
195 200 205

Ile Leu Ala Lys Ser Pro Gln Ser Thr Arg Ala Val Ala Pro Gln Lys
210 215 220

Cys Phe Leu Gln Ile Lys Gly Met Thr Cys Ala Ser Cys Val Ser Asn
225 230 235 240

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

Ala Leu Met Ala Gly Lys Ala Glu Ile Lys Tyr Asp Pro Glu Val Ile
260 265 270

Gln Pro Leu Glu Ile Ala Gln Phe Ile Gln Asp Leu Gly Phe Glu Ala
275 280 285

Ala Val Met Glu Asp Tyr Ala Gly Ser Asp Gly Asn Ile Glu Leu Thr
290 295 300

Ile Thr Gly Met Thr Cys Ala Ser Cys Val His Asn Ile Glu Ser Lys
305 310 315 320

Leu Thr Arg Thr Asn Gly Ile Thr Tyr Ala Ser Val Ala Leu Ala Thr

Glu Asp Asn Leu Ile Ile Arg Glu Glu Gln Val Pro Met Glu Leu Val
545 550 555 560

Gln Arg Gly Asp Ile Val Lys Val Val Pro Gly Gly Lys Phe Pro Val
565 570 575

Asp Gly Lys Val Leu Glu Gly Asn Thr Met Ala Asp Glu Ser Leu Ile
580 585 590

Thr Gly Glu Ala Met Pro Val Thr Lys Lys Pro Gly Ser Thr Val Ile
595 600 605

Ala Gly Ser Ile Asn Ala His Gly Ser Val Leu Ile Lys Ala Thr His
610 615 620

Val Gly Asn Asp Thr Thr Leu Ala Gln Ile Val Lys Leu Val Glu Glu
625 630 635 640

Ala Gln Met Ser Lys Ala Pro Ile Gln Gln Leu Ala Asp Arg Phe Ser
645 650 655

Gly Tyr Phe Val Pro Phe Ile Ile Ile Met Ser Thr Leu Thr Leu Val
660 665 670

Val Trp Ile Val Ile Gly Phe Ile Asp Phe Gly Val Val Gln Arg Tyr
675 680 685

Phe Pro Asn Pro Asn Lys His Ile Ser Gln Thr Glu Val Ile Ile Arg
690 695 700

Phe Ala Phe Gln Thr Ser Ile Thr Val Leu Cys Ile Ala Cys Pro Cys
705 710 715 720

Ser Leu Gly Leu Ala Thr Pro Thr Ala Val Met Val Gly Thr Gly Val
725 730 735

Ala Ala Gln Asn Gly Ile Leu Ile Lys Gly Gly Lys Pro Leu Glu Met
740 745 750

Ala His Lys Ile Lys Thr Val Met Phe Asp Lys Thr Gly Thr Ile Thr
755 760 765

His Gly Val Pro Arg Val Met Arg Val Leu Leu Leu Gly Asp Val Ala
770 775 780

Thr Leu Pro Leu Arg Lys Val Leu Ala Val Val Gly Thr Ala Glu Ala
785 790 795 800

Ser Ser Glu His Pro Leu Gly Val Ala Val Thr Lys Tyr Cys Lys Glu
805 810 815

Glu Leu Gly Thr Glu Thr Leu Gly Tyr Cys Thr Asp Phe Gln Ala Val
820 825 830

Pro Gly Cys Gly Ile Gly Cys Lys Val Ser Asn Val Glu Gly Ile Leu
835 840 845

Ala His Ser Glu Arg Pro Leu Ser Ala Pro Ala Ser His Leu Asn Glu
850 855 860

Ala Gly Ser Leu Pro Ala Glu Lys Asp Ala Val Pro Gln Thr Phe Ser
865 870 875 880

Val Leu Ile Gly Asn Arg Glu Trp Leu Arg Arg Asn Gly Leu Thr Ile
885 890 895

Ser Ser Asp Val Ser Asp Ala Met Thr Asp His Glu Met Lys Gly Gln
900 905 910

Thr Ala Ile Leu Val Ala Ile Asp Gly Val Leu Cys Gly Met Ile Ala
915 920 925

Ile Ala Asp Ala Val Lys Gln Glu Ala Ala Leu Ala Val His Thr Leu
930 935 940

Gln Ser Met Gly Val Asp Val Val Leu Ile Thr Gly Asp Asn Arg Lys
945 950 955 960

Thr Ala Arg Ala Ile Ala Thr Gln Val Gly Ile Asn Lys Val Phe Ala
965 970 975

Glu Val Leu Pro Ser His Lys Val Ala Lys Val Gln Glu Leu Gln Asn
980 985 990

Lys Gly Lys Lys Val Ala Met Val Gly Asp Gly Val Asn Asp Ser Pro
995 1000 1005

Ala Leu Ala Gln Ala Asp Met Gly Val Ala Ile Gly Thr Gly Thr
1010 1015 1020

Asp Val Ala Ile Glu Ala Ala Asp Val Val Leu Ile Arg Asn Asp
1025 1030 1035

Leu Leu Asp Val Val Ala Ser Ile His Leu Ser Lys Arg Thr Val
1040 1045 1050

Arg Arg Ile Arg Ile Asn Leu Val Leu Ala Leu Ile Tyr Asn Leu
1055 1060 1065

Val Gly Ile Pro Ile Ala Ala Gly Val Phe Met Pro Ile Gly Ile
1070 1075 1080

Val Leu Gln Pro Trp Met Gly Ser Ala Ala Met Ala Ala Ser Ser
1085 1090 1095

Val Ser Val Val Leu Ser Ser Leu Gln Leu Lys Cys Tyr Lys Lys
1100 1105 1110

Pro Asp Leu Glu Arg Tyr Glu Ala Gln Ala His Gly His Met Lys
1115 1120 1125

Pro Leu Thr Ala Ser Gln Val Ser Val His Ile Gly Met Asp Asp
1130 1135 1140

Arg Trp Arg Asp Ser Pro Arg Ala Thr Pro Trp Asp Gln Val Ser
1145 1150 1155

Tyr Val Ser Gln Val Ser Leu Ser Ser Leu Thr Ser Asp Lys Pro

1160		1165		1170
Ser Arg His Ser Ala Ala Ala Asp Asp Asp Gly Asp Lys Trp Ser				
1175		1180		1185
Leu Leu Leu Asn Gly Arg Asp Glu Glu Gln Tyr Ile				
1190		1195		1200

<210> 9
 <211> 2211
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of artificial sequence: Nucleic acid sequence
 encoding AAV9 capsid protein VP1

<400> 9

atggctgccg atggttatct tccagattgg ctcgaggaca accttagtga aggaattcgc	60
gagtgggtggg ctttgaaacc tggagcccct caaccaagg caaatcaaca acatcaagac	120
aacgctcgag gtcttgtgct tccgggttac aaataccttg gacccggcaa cggactcgac	180
aagggggagc cggatcaacgc agcagacgcg gcggccctcg agcacgacaa ggcctacgac	240
cagcagctca aggccggaga caacccttac ctcaagtaca accacgccga cgccgagttc	300
caggagcggc tcaaagaaga tacgtctttt gggggcaacc tcgggcgagc agtcttccag	360
gccaaaaaga ggcttcttga acctcttggt ctggttgagg aagcggctaa gacggctcct	420
ggaaagaaga ggcctgtaga gcagtctcct caggaaccgg actcctccgc gggatttggc	480
aaatcgggtg cacagccgc taaaaagaga ctcaatttcg gtcagactgg cgacacagag	540
tcagtcccag acctcaacc aatcggagaa cctcccgag cccctcagg tgtgggatct	600
cttacaatgg cttcaggtgg tggcgcacca gtggcagaca ataacgaagg tgccgatgga	660
gtgggtagtt cctcgggaaa ttggcattgc gattcccaat ggctggggga cagagtcac	720
accaccagca cccgaacctg ggccctgccc acctacaaca atcacctcta caagcaaadc	780
tccaacagca catctggagg atcttcaaata gacaacgcct acttcggcta cagcaccacc	840
tgggggtatt ttgacttcaa cagattccac tgccacttct caccacgtga ctggcagcga	900
ctcatcaaca acaactgggg attccggcct aagcgactca acttcaagct cttcaacatt	960

cagggtcaaag aggttacgga caacaatgga gtcaagacca tcgccaataa ccttaccagc	1020
acgggtccagg tcttcacgga ctacagactat cagctcccgt acgtgctcgg gtcggctcac	1080
gagggctgcc tcccgccgtt cccagcggac gttttcatga ttcctcagta cgggtatctg	1140
acgcttaatg atggaagcca ggccgtgggt cgttcgtcct ttactgcct ggaatatttc	1200
ccgtcgcaaa tgctaagaac gggtacaac ttccagttca gctacgagtt tgagaacgta	1260
cctttccata gcagctacgc tcacagccaa agcctggacc gactaatgaa tccactcatc	1320
gaccaatact tgtactatct ctcaaagact attaacgggt ctggacagaa tcaacaaacg	1380
ctaaaattca gtgtggccgg acccagcaac atggctgtcc aggaagaaa ctacatacct	1440
ggaccagct accgacaaca acgtgtctca accactgtga ctcaaaacaa caacagcgaa	1500
tttgcttggc ctggagcttc ttcttgggct ctcaatggac gtaatagctt gatgaatcct	1560
ggacctgcta tggccagcca caaagaagga gaggaccgtt tctttccttt gtctggatct	1620
ttaatTTTTG gcaacaagg aactggaaga gacaacgtgg atgcggacaa agtcatgata	1680
accaacgaag aagaaattaa aactactaac ccggtagcaa cggagtccta tggacaagtg	1740
gccacaaacc accagagtgc ccaagcacag gcgcagaccg gctgggttca aaaccaagga	1800
atacttccgg gtatggtttg gcaggacaga gatgtgtacc tgcaaggacc catttgggcc	1860
aaaattcctc acacggacgg caactttcac ctttctccgc tgatgggagg gtttggaatg	1920
aagcaccgc ctctcagat cctcatcaaa aacacacctg tacctgcgga tcctccaacg	1980
gccttcaaca aggacaagct gaactctttc atcaccagct attctactgg ccaagtcagc	2040
gtggagatcg agtgggagct gcagaaggaa aacagcaagc gctggaaccc ggagatccag	2100
tacacttcca actattacaa gtctaataat gttgaatttg ctgttaatac tgaaggtgta	2160
tatagtgaac cccgccccat tggcaccaga tacctgactc gtaatctgta a	2211

<210> 10

<211> 736

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of artificial sequence: Amino acid sequence encoding AAV9 capsid protein VP1

<400> 10

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
 1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro
 20 25 30

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

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
 50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
 65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
 85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
 100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro
 115 120 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
 130 135 140

Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
 145 150 155 160

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

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
 180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
 195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400

Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu
405 410 415

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu

420	425	430													
Asp	Arg	Leu	Met	Asn	Pro	Leu	Ile	Asp	Gln	Tyr	Leu	Tyr	Tyr	Leu	Ser
		435					440					445			
Lys	Thr	Ile	Asn	Gly	Ser	Gly	Gln	Asn	Gln	Gln	Thr	Leu	Lys	Phe	Ser
	450					455					460				
Val	Ala	Gly	Pro	Ser	Asn	Met	Ala	Val	Gln	Gly	Arg	Asn	Tyr	Ile	Pro
465					470					475					480
Gly	Pro	Ser	Tyr	Arg	Gln	Gln	Arg	Val	Ser	Thr	Thr	Val	Thr	Gln	Asn
				485					490					495	
Asn	Asn	Ser	Glu	Phe	Ala	Trp	Pro	Gly	Ala	Ser	Ser	Trp	Ala	Leu	Asn
			500					505					510		
Gly	Arg	Asn	Ser	Leu	Met	Asn	Pro	Gly	Pro	Ala	Met	Ala	Ser	His	Lys
		515					520					525			
Glu	Gly	Glu	Asp	Arg	Phe	Phe	Pro	Leu	Ser	Gly	Ser	Leu	Ile	Phe	Gly
	530					535					540				
Lys	Gln	Gly	Thr	Gly	Arg	Asp	Asn	Val	Asp	Ala	Asp	Lys	Val	Met	Ile
545					550					555					560
Thr	Asn	Glu	Glu	Glu	Ile	Lys	Thr	Thr	Asn	Pro	Val	Ala	Thr	Glu	Ser
				565					570					575	
Tyr	Gly	Gln	Val	Ala	Thr	Asn	His	Gln	Ser	Ala	Gln	Ala	Gln	Ala	Gln
			580					585					590		
Thr	Gly	Trp	Val	Gln	Asn	Gln	Gly	Ile	Leu	Pro	Gly	Met	Val	Trp	Gln
	595						600					605			
Asp	Arg	Asp	Val	Tyr	Leu	Gln	Gly	Pro	Ile	Trp	Ala	Lys	Ile	Pro	His
	610					615					620				
Thr	Asp	Gly	Asn	Phe	His	Pro	Ser	Pro	Leu	Met	Gly	Gly	Phe	Gly	Met
625					630					635					640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln
675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 11
<211> 6
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence: Consensus Kozak Sequence

<400> 11
gccgcc 6

<210> 12
<211> 190
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of artificial sequence: TTR promoter Sequence

<400> 12
atttcataga acgaatgttc cgatgctcta atctctctag acaaggttca tatttgtatg 60
ggttacttat tctctctttg ttgactaagt caataatcag aatcagcagg tttgcagtca 120
gattggcagg gataagcagc ctagctcagg agaagtgagt ataaaagccc caggctggga 180

gcagccatca

190

<210> 13

<211> 1120

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of artificial sequence: DEL 1-4 ATP7B Native, protein

<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 Ser
50 55 60

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

Val Val Ser Glu Ser Cys Ser Thr Asn Pro Leu Gly Asn His Ser Ala
85 90 95

Gly Asn Ser Met Val Gln Thr Thr Asp Gly Thr Pro Thr Ser Val Gln
100 105 110

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

Ile Leu Ala Lys Ser Pro Gln Ser Thr Arg Ala Val Ala Pro Gln Lys
130 135 140

Cys Phe Leu Gln Ile Lys Gly Met Thr Cys Ala Ser Cys Val Ser Asn
145 150 155 160

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

Ala Leu Met Ala Gly Lys Ala Glu Ile Lys Tyr Asp Pro Glu Val Ile
180 185 190

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

Ala Val Met Glu Asp Tyr Ala Gly Ser Asp Gly Asn Ile Glu Leu Thr
210 215 220

Ile Thr Gly Met Thr Cys Ala Ser Cys Val His Asn Ile Glu Ser Lys
225 230 235 240

Leu Thr Arg Thr Asn Gly Ile Thr Tyr Ala Ser Val Ala Leu Ala Thr
245 250 255

Ser Lys Ala Leu Val Lys Phe Asp Pro Glu Ile Ile Gly Pro Arg Asp
260 265 270

Ile Ile Lys Ile Ile Glu Glu Ile Gly Phe His Ala Ser Leu Ala Gln
275 280 285

Arg Asn Pro Asn Ala His His Leu Asp His Lys Met Glu Ile Lys Gln
290 295 300

Trp Lys Lys Ser Phe Leu Cys Ser Leu Val Phe Gly Ile Pro Val Met
305 310 315 320

Ala Leu Met Ile Tyr Met Leu Ile Pro Ser Asn Glu Pro His Gln Ser
325 330 335

Met Val Leu Asp His Asn Ile Ile Pro Gly Leu Ser Ile Leu Asn Leu
340 345 350

Ile Phe Phe Ile Leu Cys Thr Phe Val Gln Leu Leu Gly Gly Trp Tyr
355 360 365

Phe Tyr Val Gln Ala Tyr Lys Ser Leu Arg His Arg Ser Ala Asn Met

370

375

380

Asp Val Leu Ile Val Leu Ala Thr Ser Ile Ala Tyr Val Tyr Ser Leu
 385 390 395 400

Val Ile Leu Val Val Ala Val Ala Glu Lys Ala Glu Arg Ser Pro Val
 405 410 415

Thr Phe Phe Asp Thr Pro Pro Met Leu Phe Val Phe Ile Ala Leu Gly
 420 425 430

Arg Trp Leu Glu His Leu Ala Lys Ser Lys Thr Ser Glu Ala Leu Ala
 435 440 445

Lys Leu Met Ser Leu Gln Ala Thr Glu Ala Thr Val Val Thr Leu Gly
 450 455 460

Glu Asp Asn Leu Ile Ile Arg Glu Glu Gln Val Pro Met Glu Leu Val
 465 470 475 480

Gln Arg Gly Asp Ile Val Lys Val Val Pro Gly Gly Lys Phe Pro Val
 485 490 495

Asp Gly Lys Val Leu Glu Gly Asn Thr Met Ala Asp Glu Ser Leu Ile
 500 505 510

Thr Gly Glu Ala Met Pro Val Thr Lys Lys Pro Gly Ser Thr Val Ile
 515 520 525

Ala Gly Ser Ile Asn Ala His Gly Ser Val Leu Ile Lys Ala Thr His
 530 535 540

Val Gly Asn Asp Thr Thr Leu Ala Gln Ile Val Lys Leu Val Glu Glu
 545 550 555 560

Ala Gln Met Ser Lys Ala Pro Ile Gln Gln Leu Ala Asp Arg Phe Ser
 565 570 575

Gly Tyr Phe Val Pro Phe Ile Ile Ile Met Ser Thr Leu Thr Leu Val
 580 585 590

Val Trp Ile Val Ile Gly Phe Ile Asp Phe Gly Val Val Gln Arg Tyr
 595 600 605

Phe Pro Asn Pro Asn Lys His Ile Ser Gln Thr Glu Val Ile Ile Arg
 610 615 620

Phe Ala Phe Gln Thr Ser Ile Thr Val Leu Cys Ile Ala Cys Pro Cys
 625 630 635 640

Ser Leu Gly Leu Ala Thr Pro Thr Ala Val Met Val Gly Thr Gly Val
 645 650 655

Ala Ala Gln Asn Gly Ile Leu Ile Lys Gly Gly Lys Pro Leu Glu Met
 660 665 670

Ala His Lys Ile Lys Thr Val Met Phe Asp Lys Thr Gly Thr Ile Thr
 675 680 685

His Gly Val Pro Arg Val Met Arg Val Leu Leu Leu Gly Asp Val Ala
 690 695 700

Thr Leu Pro Leu Arg Lys Val Leu Ala Val Val Gly Thr Ala Glu Ala
 705 710 715 720

Ser Ser Glu His Pro Leu Gly Val Ala Val Thr Lys Tyr Cys Lys Glu
 725 730 735

Glu Leu Gly Thr Glu Thr Leu Gly Tyr Cys Thr Asp Phe Gln Ala Val
 740 745 750

Pro Gly Cys Gly Ile Gly Cys Lys Val Ser Asn Val Glu Gly Ile Leu
 755 760 765

Ala His Ser Glu Arg Pro Leu Ser Ala Pro Ala Ser His Leu Asn Glu
 770 775 780

Ala Gly Ser Leu Pro Ala Glu Lys Asp Ala Val Pro Gln Thr Phe Ser
 785 790 795 800

Val Leu Ile Gly Asn Arg Glu Trp Leu Arg Arg Asn Gly Leu Thr Ile
805 810 815

Ser Ser Asp Val Ser Asp Ala Met Thr Asp His Glu Met Lys Gly Gln
820 825 830

Thr Ala Ile Leu Val Ala Ile Asp Gly Val Leu Cys Gly Met Ile Ala
835 840 845

Ile Ala Asp Ala Val Lys Gln Glu Ala Ala Leu Ala Val His Thr Leu
850 855 860

Gln Ser Met Gly Val Asp Val Val Leu Ile Thr Gly Asp Asn Arg Lys
865 870 875 880

Thr Ala Arg Ala Ile Ala Thr Gln Val Gly Ile Asn Lys Val Phe Ala
885 890 895

Glu Val Leu Pro Ser His Lys Val Ala Lys Val Gln Glu Leu Gln Asn
900 905 910

Lys Gly Lys Lys Val Ala Met Val Gly Asp Gly Val Asn Asp Ser Pro
915 920 925

Ala Leu Ala Gln Ala Asp Met Gly Val Ala Ile Gly Thr Gly Thr Asp
930 935 940

Val Ala Ile Glu Ala Ala Asp Val Val Leu Ile Arg Asn Asp Leu Leu
945 950 955 960

Asp Val Val Ala Ser Ile His Leu Ser Lys Arg Thr Val Arg Arg Ile
965 970 975

Arg Ile Asn Leu Val Leu Ala Leu Ile Tyr Asn Leu Val Gly Ile Pro
980 985 990

Ile Ala Ala Gly Val Phe Met Pro Ile Gly Ile Val Leu Gln Pro Trp
995 1000 1005

Met Gly Ser Ala Ala Met Ala Ala Ser Ser Val Ser Val Val Leu
 1010 1015 1020

Ser Ser Leu Gln Leu Lys Cys Tyr Lys Lys Pro Asp Leu Glu Arg
 1025 1030 1035

Tyr Glu Ala Gln Ala His Gly His Met Lys Pro Leu Thr Ala Ser
 1040 1045 1050

Gln Val Ser Val His Ile Gly Met Asp Asp Arg Trp Arg Asp Ser
 1055 1060 1065

Pro Arg Ala Thr Pro Trp Asp Gln Val Ser Tyr Val Ser Gln Val
 1070 1075 1080

Ser Leu Ser Ser Leu Thr Ser Asp Lys Pro Ser Arg His Ser Ala
 1085 1090 1095

Ala Ala Asp Asp Asp Gly Asp Lys Trp Ser Leu Leu Leu Asn Gly
 1100 1105 1110

Arg Asp Glu Glu Gln Tyr Ile
 1115 1120

<210> 14
 <211> 4485
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of artificial sequence: Expression cassette of
 exemplary vector DTC319

<400> 14
 ttggccactc cctctctgcg cgctcgctcg ctactgagg ccgggcgacc aaaggtcgcc 60
 cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgag agagggagtg 120
 gccaactcca tcactagggg ttctctacc tcgtgatcgc ccggcccctg ttcaaactg 180
 tcctaatact ctgtctctgc aagggtcatc agtagttttc catcttactc aacatcctcc 240
 cagtgatctc atagaacgaa tggtccgatg ctctaatttc tctagacaag gttcatatct 300
 gtatgggtta cttattctct ctttgttgac taagtcaata atcagaatca gcaggtttgc 360

agtcagattg gcagggataa gcagcctagc tcaggagaag tgagtataaa agccccaggc	420
tgggagcagc catcagctct aaggtaaata taaaatTTTT aagtgtataa tgtgttaaac	480
tactgattct aattgtttct ctcttttaga ttccaacctt tggaactgat gccgccacca	540
tgcttgagca ggagagacag atcacagcca gagaaggggc cagtcggaaa atcttatcta	600
agctttcttt gcctacccgt gcctgggaac cagcaatgaa gaagagtttt gcttttgaca	660
atgttggcta tgaaggtggc ctggatggcc tgggcccttc ttctcaggtg gccaccagca	720
cagtcaggga tggagccgaa gggagtggga cagatcacag gtcttccagt tctcattccc	780
ctggctcccc accgagaaac caggtccagg gcacatgcag taccactctg attgccattg	840
ccggcatgac ctgtgcatcc tgtgtccatt ccattgaagg catgatctcc caactggaag	900
gggtgcagca aatatcgggt tctttggccg aagggactgc aacagttctt tataatccct	960
ctgtaattag cccagaagaa ctcagagctg ctatagaaga catgggattt gaggcttcag	1020
tcgtttctga aagctgttct actaaccctc ttggaaacca cagtgtctggg aattccatgg	1080
tgcaaaactac agatggtaca cctacatctg tgcaggaagt ggctccccac actgggaggc	1140
tccctgcaaa ccatgccccg gacatcttgg caaagtcccc acaatcaacc agagcagtgg	1200
caccgcagaa gtgcttctta cagatcaaag gcatgacctg tgcatcctgt gtgtctaaca	1260
tagaaaggaa tctgcagaaa gaagctggtg ttctctccgt gttggttgcc ttgatggcag	1320
gaaaggcaga gatcaagtat gaccagagg tcatccagcc cctcgagata gctcagttca	1380
tccaggacct gggttttgag gcagcagtca tggaggacta cgcaggctcc gatggcaaca	1440
ttgagctgac aatcacaggg atgacctgcg cgtcctgtgt ccacaacata gagtccaaac	1500
tcacgaggac aaatggcatc acttatgcct ccgttgccct tgccaccagc aaagcccttg	1560
ttaagtttga cccggaaatt atcgggtccac gggatattat caaaattatt gaggaaattg	1620
gctttcatgc ttccctggcc cagagaaacc ccaacgtca tcacttggac cacaagatgg	1680
aaataaagca gtggaagaag tctttcctgt gcagcctggc gtttggcatc cctgtcatgg	1740
ccttaatgat ctatatgctg ataccagca acgagcccca ccagtccatg gtcctggacc	1800
acaacatcat tccaggactg tccattctaa atctcatctt ctttatcttg tgtaccttg	1860
tccagctcct cggtgggtgg tacttctacg ttcaggccta caaatctctg agacacaggt	1920

cagccaacat ggacgtgctc atcgtcctgg ccacaagcat tgcttatgtt tattctctgg	1980
tcatcctggt ggttgctgtg gctgagaagg cggagaggag ccctgtgaca ttcttcgaca	2040
cgccccccat gctctttgtg ttcattgccc tgggccggtg gctggaacac ttggcaaaga	2100
gcaaaacctc agaagccctg gctaaactca tgtctctcca agccacagaa gccaccgttg	2160
tgacccttgg tgaggacaat ttaatcatca gggaggagca agtccccatg gagctggtgc	2220
agcggggcga tatcgtcaag gtggtccctg ggggaaagt tccagtggat gggaaagtcc	2280
tggaaggcaa taccatggct gatgagtccc tcatcacagg agaagccatg ccagtcacta	2340
agaaacccgg aagcactgta attgcggggt ctataaatgc acatggctct gtgctcatta	2400
aagctacca cgtgggcaat gacaccactt tggctcagat tgtgaaactg gtggaagagg	2460
ctcagatgtc aaaggcaccc attcagcagc tggctgaccg gtttagtgga tattttgtcc	2520
catttatcat catcatgtca actttgacgt tgggtggtatg gattgtaatc ggttttatcg	2580
attttggtgt gtgtcagaga tactttccta accccaacaa gcacatctcc cagacagagg	2640
tgatcatccg gtttgctttc cagacgtcca tcacggtgct gtgcattgcc tgcccctgct	2700
ccctggggct ggccacgccc acggctgtca tgggtgggcac cggggtggcc gcgcagaacg	2760
gcatcctcat caagggaggc aagcccctgg agatggcgca caagataaag actgtgatgt	2820
ttgacaagac tggcaccatt acccatggcg tccccagggt catgcgggtg ctctgtctgg	2880
gggatgtggc cacttgccc ctgaggaagg ttctggctgt ggtggggact gcggaggcca	2940
gcagtgaaca ccccttgggc gtggcagtca ccaaatactg taaagaggaa cttggaacag	3000
agaccttggg atactgcacg gacttccagg cagtgccagg ctgtggaatt gggtgcaaag	3060
tcagcaacgt ggaaggcatc ctggcccaca gtgagcgccc tttgagtga cggccagtc	3120
acctgaatga ggctggcagc cttcccgcag aaaaagatgc agtccccag accttctctg	3180
tgctgattgg aaaccgtgag tggctgaggc gcaacggttt aaccatttct agcgatgtca	3240
gtgacgctat gacagaccac gagatgaaag gacagacagc catcctggtg gctattgacg	3300
gtgtgctctg tgggatgatc gcaatcgag acgctgtcaa gcaggaggct gccctggctg	3360
tgcacacgct gcagagcatg ggtgtggacg tggttctgat cacgggggac aaccggaaga	3420
cagccagagc tattgccacc caggttggca tcaacaaagt ctttgcagag gtgctgcctt	3480
cgcacaaagt ggccaagtc caggagctcc agaataaagg gaagaaagtc gccatggtgg	3540

gggatgggggt caatgactcc ccggccttgg cccaggcaga catgggtgtg gccattggca	3600
ccggcacgga tgtggccatc gaggcagccg acgtcgtcct tatcagaaat gatttgctgg	3660
atgtgggtggc tagcattcac ctttccaaga ggactgtccg aaggatacgc atcaacctgg	3720
tcctggcact gatttataac ctggttggga taccattgc agcaggtgtc ttcattgcca	3780
tcggcattgt gctgcagccc tggatgggct cagcggccat ggagcctcc tctgtgtctg	3840
tggtgctctc atccctgcag ctcaagtgtg ataagaagcc tgacctggag aggtatgagg	3900
cacaggcgca tggccacatg aagcccctga cggcatcca ggtagtgtg cacataggca	3960
tggatgacag gtggcgggac tccccaggc ccacaccatg ggaccaggtc agctatgtca	4020
gccaggtgtc gctgtcctcc ctgacgtccg acaagccatc tcggcacagc gctgcagcag	4080
acgatgatgg ggacaagtgg tctctgctcc tgaatggcag ggatgaggag cagtacatct	4140
gagatccaga catgataaga tacattgatg agtttggaca aaccacaact agaatgcagt	4200
gaaaaaaaaatg ctttatttgt gaaatttgtg atgctattgc tttatttcta accattataa	4260
gctgcaataa acaagttaac aacaacaatt gcattcattt tatgtttcag gttcaggggg	4320
aggtgtggga ggttttttag aggaaccct agtgatggag ttggccactc cctctctgcg	4380
cgctcgctcg ctactgagg ccgcccgggc aaagcccggg cgtcgggcga cttttggtcg	4440
cccggcctca gtgagcgagc gagcgcgag agagggtgtg gccaa	4485

<210> 15

<211> 3603

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of artificial sequence: Del 1-3 ATP7B,
codon-optimized

<400> 15

atgcccgagc aggaaagaca gatcaccgcc agagagggcg ccagccggaa gatcctgagc	60
aagctgagcc tgcccaccag agcctgggag cccgccatga agaagtcctt cgccttcgac	120
aacgtgggct acgagggcgg cctggacgga ctgggacat cttctcaggt ggccacaagc	180
accgtgcgcg acggagccga gggctccgga acagatcaca gaagcagcag cagccacagc	240
cctggcagcc cccctagaaa tcaggtgcag ggcacctgta gcaccaccct gatcgccatt	300

gccggcatga catgcgccag ctgcgtgcac tctattgagg gcatgatctc ccagctggaa	360
ggcgtgcagc agatcagtgt gtctctggcc gagggcacccg ccacagtgtg gtacaaccct	420
agcgtgatca gccccgaaga actgagagcc gccattgagg acatgggatt cgaagccagc	480
gtggtgtccg agagctgtc caccaaccct ctgggcaatc acagcgccgg caacagcatg	540
gtgcagacca ccgacggcac ccccaccagc gtgcaggaag tggccccaca tacaggcaga	600
ctgcccgcc atcacgcccc cgatatcctg gccaaagacc cccagagtac aagagccgtg	660
ggccccaga agtgcttcct gcagatcaag ggcatgactt gtgcctcttg tgtgtccaat	720
atcgagcgga acctgcagaa agaggccggc gtgctgtctg tgctggtggc tctgatggcc	780
ggcaaggccg agatcaaata cgaccccgaa gtgattcagc ccctggaaat cgcccagttt	840
atccaggacc tgggctttga agccgccgtg atggaagatt acgccggctc cgacggcaac	900
atcgagctga ccatcaccgg aatgacctgc gcctcctgtg tgcacaacat tgagtccaag	960
ctgaccggga ccaacggcat cacctacgcc tctgtggctc tggccacctc caaggccctc	1020
gtgaagtctg atcccagat catcggcccc agggacatca tcaagatcat cgaagagatc	1080
ggcttcacg ccagcctggc ccagaggaac cctaacgccc accacctgga ccacaagatg	1140
gaaatcaagc agtgaagaa aagcttcctg tgcagcctgg tgttcggcat ccccgatg	1200
gccctgatga tctacatgt gatccccagc aacgagcccc accagtccat ggtgctggat	1260
cacaacatca tccccggcct gtctatcctg aacctgatct tcttcacct gtgcaccttc	1320
gtgcagctgc tgggcggctg gtacttctac gtgcaggcct acaagtcct gcggcacaga	1380
tccgccaaca tggacgtgt gatcgtgtg gccacatcta tcgcctacgt gtactccctc	1440
gtgatcctgg tggtagccgt ggccgagaaa gccgagagaa gccctgtgac cttcttcgac	1500
accccccta tgctgttcgt gtttatcgcc ctgggcccgt ggctggaaca cctggccaaa	1560
agcaagacca gcgaggccct ggctaagctg atgagtctgc aggccaccga ggccacagtc	1620
gtgaccctgg gcgaggacaa cctgatcatc cgcgaggaac aggtgccaat ggaactggtg	1680
cagcggggcg acatcgtgaa ggtggtgcct ggccggcaagt tccccgtgga cggaaaagt	1740
ctggaaggga ataccatggc cgacgagagc ctgatcacag gcgaggccat gcccgtagcc	1800
aagaaacctg gcagcacagt gatcgccggc agcatcaatg cccacggcag cgtgctgatt	1860

aaggccacac acgtgggcaa cgataccacc ctggctcaga ttgtgaagct ggtggaagag	1920
gcccagatga gcaaggcccc cattcagcag ctggctgacc gggttcagcgg ctacttcgtg	1980
ccctttatca tcatcatgag caccctgaca ctggctgtgt ggatcgtgat cggctttatc	2040
gacttcggag tgggtgcagag atacttcccc aaccctaaca agcacatcag ccagacagaa	2100
gtgatcatca gattcgcctt tcagaccagc atcaccgtgc tgtgtatcgc ctgcccctgt	2160
agcctgggac tggccacacc taccgctgtg atgggtgggaa caggcgtggc cgctcagaac	2220
ggcatcctga tcaagggggg caagcctctg gaaatggctc acaagatcaa gaccgtgatg	2280
ttcgacaaga ccggcaccat caccacggc gtgcccagag tgatgagagt gctgctgctg	2340
ggggatgtgg ccaccctgcc tctgagaaag gtgctggctg tcgtgggcac agccgaggct	2400
agctctgaac acccactggg agtggccgtg acaaagtact gcaaagagga actgggcacc	2460
gaaaccctgg gctactgcac cgactttcag gccgtgcctg gctgtggcat cggctgcaag	2520
gtgtccaacg tggaaggcat cctggcccac agcgagaggc cactgtctgc ccctgccagc	2580
cacctgaacg aggccggatc tctgcccgcc gaaaaggacg ctgtgcccc aaccttctct	2640
gtgctgattg gcaacagaga gtggctgcgg cggaacggcc tgaccatctc ctccgatgtg	2700
tccgacgcca tgaccgacca cgagatgaag ggccagaccg ccattctggt ggccattgac	2760
ggggtgctgt gcggcatgat cgcaatcgcc gatgccgtga aacaggaagc agcactggcc	2820
gtgcacaccc tgcagtctat gggagtggat gtgggtgctga tcaccggcga caacagaaag	2880
accgccaggg ccattgccac ccagggtggc atcaacaagg tgttcgccga ggtgctgccc	2940
agccacaaaag tggccaaggt gcaggaactg cagaacaaag gcaaaaaggt ggccatggtg	3000
ggagatggcg tgaacgactc tcctgctctg gcccaggcag atatgggcgt ggccatcggc	3060
acaggcaccg acgtggcaat tgaggctgct gacgtggtgc tgattcggaa cgacctgctg	3120
gacgtggtgg cctccatcca cctgtccaag agaaccgtgc ggcggatcag aatcaacctg	3180
gtgctggcac tgatctataa cctcgtgggc atccctatcg ccgctggcgt gttcatgcct	3240
atcggaatcg tgctgcagcc ctggatgggc tctgccgcca tggctgcaag ctccgtgtct	3300
gtggtgctgt ccagcctgca gctgaagtgc tacaagaagc ccgacctgga aagatacagag	3360
gcccaggccc acggacacat gaagcctctg acagcctccc aggtgtccgt gcacatcggc	3420
atggacgaca gatggcggga cagccctaga gccaccctt gggatcaggt gtcatacgtg	3480

tcacaggtgt ccctgagcag cctgaccagc gacaagccca gcagacatag cgccgctgcc	3540
gacgacgatg gggacaagtg gtccttgctg ctgaacggcc gggatgagga acagtacatc	3600
tga	3603