

SEQUENCE LISTING

<110> TRANSLATE BIO INC.

<120> GENERATION OF OPTIMIZED NUCLEOTIDE SEQUENCES

<130> MRT-2131W0

<141> 2021-05-07

<150> US 62/978,180

<151> 2020-02-18

<150> US 63/021,345

<151> 2020-05-07

<160> 32

<170> SeqWin2010, version 1.0

<210> 1

<211> 874

<212> PRT

<213> Bacteriophage SP6

<400> 1

Met Gln Asp Leu His Ala Ile Gln Leu Gln Leu Glu Glu Glu Met Phe
1 5 10 15

Asn Gly Gly Ile Arg Arg Phe Glu Ala Asp Gln Gln Arg Gln Ile Ala
20 25 30

Ala Gly Ser Glu Ser Asp Thr Ala Trp Asn Arg Arg Leu Leu Ser Glu
35 40 45

Leu Ile Ala Pro Met Ala Glu Gly Ile Gln Ala Tyr Lys Glu Glu Tyr
50 55 60

Glu Gly Lys Lys Gly Arg Ala Pro Arg Ala Leu Ala Phe Leu Gln Cys
65 70 75 80

Val Glu Asn Glu Val Ala Ala Tyr Ile Thr Met Lys Val Val Met Asp
85 90 95

Met Leu Asn Thr Asp Ala Thr Leu Gln Ala Ile Ala Met Ser Val Ala
100 105 110

Glu Arg Ile Glu Asp Gln Val Arg Phe Ser Lys Leu Glu Gly His Ala
115 120 125

Ala Lys Tyr Phe Glu Lys Val Lys Lys Ser Leu Lys Ala Ser Arg Thr
130 135 140

Lys Ser Tyr Arg His Ala His Asn Val Ala Val Val Ala Glu Lys Ser

145					150					155				160
Val	Ala	Glu	Lys	Asp	Ala	Asp	Phe	Asp	Arg	Trp	Glu	Ala	Trp	Pro Lys
				165					170					175
Glu	Thr	Gln	Leu	Gln	Ile	Gly	Thr	Thr	Leu	Leu	Glu	Ile	Leu	Glu Gly
			180					185					190	
Ser	Val	Phe	Tyr	Asn	Gly	Glu	Pro	Val	Phe	Met	Arg	Ala	Met	Arg Thr
		195					200					205		
Tyr	Gly	Gly	Lys	Thr	Ile	Tyr	Tyr	Leu	Gln	Thr	Ser	Glu	Ser	Val Gly
	210					215					220			
Gln	Trp	Ile	Ser	Ala	Phe	Lys	Glu	His	Val	Ala	Gln	Leu	Ser	Pro Ala
225					230					235				240
Tyr	Ala	Pro	Cys	Val	Ile	Pro	Pro	Arg	Pro	Trp	Arg	Thr	Pro	Phe Asn
				245					250					255
Gly	Gly	Phe	His	Thr	Glu	Lys	Val	Ala	Ser	Arg	Ile	Arg	Leu	Val Lys
			260					265					270	
Gly	Asn	Arg	Glu	His	Val	Arg	Lys	Leu	Thr	Gln	Lys	Gln	Met	Pro Lys
		275					280					285		
Val	Tyr	Lys	Ala	Ile	Asn	Ala	Leu	Gln	Asn	Thr	Gln	Trp	Gln	Ile Asn
	290					295					300			
Lys	Asp	Val	Leu	Ala	Val	Ile	Glu	Glu	Val	Ile	Arg	Leu	Asp	Leu Gly
305					310					315				320
Tyr	Gly	Val	Pro	Ser	Phe	Lys	Pro	Leu	Ile	Asp	Lys	Glu	Asn	Lys Pro
				325					330					335
Ala	Asn	Pro	Val	Pro	Val	Glu	Phe	Gln	His	Leu	Arg	Gly	Arg	Glu Leu
			340					345					350	
Lys	Glu	Met	Leu	Ser	Pro	Glu	Gln	Trp	Gln	Gln	Phe	Ile	Asn	Trp Lys
		355					360					365		
Gly	Glu	Cys	Ala	Arg	Leu	Tyr	Thr	Ala	Glu	Thr	Lys	Arg	Gly	Ser Lys
	370					375					380			
Ser	Ala	Ala	Val	Val	Arg	Met	Val	Gly	Gln	Ala	Arg	Lys	Tyr	Ser Ala
385					390					395				400
Phe	Glu	Ser	Ile	Tyr	Phe	Val	Tyr	Ala	Met	Asp	Ser	Arg	Ser	Arg Val
				405					410					415
Tyr	Val	Gln	Ser	Ser	Thr	Leu	Ser	Pro	Gln	Ser	Asn	Asp	Leu	Gly Lys
			420					425					430	
Ala	Leu	Leu	Arg	Phe	Thr	Glu	Gly	Arg	Pro	Val	Asn	Gly	Val	Glu Ala

435					440					445					
Leu	Lys	Trp	Phe	Cys	Ile	Asn	Gly	Ala	Asn	Leu	Trp	Gly	Trp	Asp	Lys
	450					455					460				
Lys	Thr	Phe	Asp	Val	Arg	Val	Ser	Asn	Val	Leu	Asp	Glu	Glu	Phe	Gln
465					470					475					480
Asp	Met	Cys	Arg	Asp	Ile	Ala	Ala	Asp	Pro	Leu	Thr	Phe	Thr	Gln	Trp
				485					490					495	
Ala	Lys	Ala	Asp	Ala	Pro	Tyr	Glu	Phe	Leu	Ala	Trp	Cys	Phe	Glu	Tyr
			500					505					510		
Ala	Gln	Tyr	Leu	Asp	Leu	Val	Asp	Glu	Gly	Arg	Ala	Asp	Glu	Phe	Arg
		515					520					525			
Thr	His	Leu	Pro	Val	His	Gln	Asp	Gly	Ser	Cys	Ser	Gly	Ile	Gln	His
	530					535					540				
Tyr	Ser	Ala	Met	Leu	Arg	Asp	Glu	Val	Gly	Ala	Lys	Ala	Val	Asn	Leu
545					550					555					560
Lys	Pro	Ser	Asp	Ala	Pro	Gln	Asp	Ile	Tyr	Gly	Ala	Val	Ala	Gln	Val
				565					570					575	
Val	Ile	Lys	Lys	Asn	Ala	Leu	Tyr	Met	Asp	Ala	Asp	Asp	Ala	Thr	Thr
			580					585					590		
Phe	Thr	Ser	Gly	Ser	Val	Thr	Leu	Ser	Gly	Thr	Glu	Leu	Arg	Ala	Met
		595					600					605			
Ala	Ser	Ala	Trp	Asp	Ser	Ile	Gly	Ile	Thr	Arg	Ser	Leu	Thr	Lys	Lys
		610				615					620				
Pro	Val	Met	Thr	Leu	Pro	Tyr	Gly	Ser	Thr	Arg	Leu	Thr	Cys	Arg	Glu
625					630					635					640
Ser	Val	Ile	Asp	Tyr	Ile	Val	Asp	Leu	Glu	Glu	Lys	Glu	Ala	Gln	Lys
				645					650					655	
Ala	Val	Ala	Glu	Gly	Arg	Thr	Ala	Asn	Lys	Val	His	Pro	Phe	Glu	Asp
			660					665					670		
Asp	Arg	Gln	Asp	Tyr	Leu	Thr	Pro	Gly	Ala	Ala	Tyr	Asn	Tyr	Met	Thr
		675					680					685			
Ala	Leu	Ile	Trp	Pro	Ser	Ile	Ser	Glu	Val	Val	Lys	Ala	Pro	Ile	Val
	690					695					700				
Ala	Met	Lys	Met	Ile	Arg	Gln	Leu	Ala	Arg	Phe	Ala	Ala	Lys	Arg	Asn
705					710					715					720
Glu	Gly	Leu	Met	Tyr	Thr	Leu	Pro	Thr	Gly	Phe	Ile	Leu	Glu	Gln	Lys

	725		730		735	
Ile Met Ala Thr Glu Met Leu Arg Val Arg Thr Cys Leu Met Gly Asp						
	740		745		750	
Ile Lys Met Ser Leu Gln Val Glu Thr Asp Ile Val Asp Glu Ala Ala						
	755		760		765	
Met Met Gly Ala Ala Ala Pro Asn Phe Val His Gly His Asp Ala Ser						
	770		775		780	
His Leu Ile Leu Thr Val Cys Glu Leu Val Asp Lys Gly Val Thr Ser						
785		790		795		800
Ile Ala Val Ile His Asp Ser Phe Gly Thr His Ala Asp Asn Thr Leu						
	805		810		815	
Thr Leu Arg Val Ala Leu Lys Gly Gln Met Val Ala Met Tyr Ile Asp						
	820		825		830	
Gly Asn Ala Leu Gln Lys Leu Leu Glu Glu His Glu Val Arg Trp Met						
	835		840		845	
Val Asp Thr Gly Ile Glu Val Pro Glu Gln Gly Glu Phe Asp Leu Asn						
	850		855		860	
Glu Ile Met Asp Ser Glu Tyr Val Phe Ala						
865		870				

<210> 2
 <211> 2625
 <212> DNA
 <213> Bacteriophage SP6

<400> 2

atgcaagatt	tacacgctat	ccagcttcaa	ttagaagaag	agatgtttta	tggtggcatt	60
cgtcgcttcg	aagcagatca	acaacgccag	attgcagcag	gtagcgagag	cgacacagca	120
tggaaccgcc	gcctgttgct	agaacttatt	gcacctatgg	ctgaaggcat	tcaggcttat	180
aaagaagagt	acgaaggtaa	gaaaggtcgt	gcacctcgcg	cattggcttt	cttacaatgt	240
gtagaaaatg	aagttgcagc	atacatcact	atgaaagttg	ttatggatat	gctgaatacg	300
gatgctaccc	ttcaggctat	tgcaatgagt	gtagcagaac	gcattgaaga	ccaagtgcgc	360
ttttctaagc	tagaagggtca	cgccgctaaa	tactttgaga	aggttaagaa	gtcactcaag	420
gctagccgta	ctaagtcata	tcgtcacgct	cataacgtag	ctgtagttgc	tgaaaaatca	480
gttgcaaaaa	aggacgcgga	ctttgaccgt	tgggaggcgt	ggccaaaaga	aactcaattg	540
cagattggta	ctaccttgct	tgaaatctta	gaaggtagcg	ttttctataa	tggtgaacct	600
gtatttatgc	gtgctatgcg	cacttatggc	ggaaagacta	tttactactt	acaaacttct	660
gaaagtgtag	gccagtggat	tagcgcattc	aaagagcacg	tagcgcaatt	aagcccagct	720
tatgcccctt	gcgtaatccc	tcctcgtcct	tggagaactc	catttaattg	agggttccat	780
actgagaagg	tagctagccg	tatccgtctt	gtaaaaggta	accgtgagca	tgtacgcaag	840
ttgactcaaa	agcaaatgcc	aaaggtttat	aaggctatca	acgcattaca	aaatacacaa	900
tggcaaatca	acaaggatgt	attagcagtt	attgaagaag	taatccgctt	agaccttggt	960
tatggtgtac	cttccttcaa	gccactgatt	gacaaggaga	acaagccagc	taacccggta	1020
cctgttgaat	tccaacacct	gcgcggtcgt	gaactgaaag	agatgctatc	acctgagcag	1080
tggcaacaat	tcattaactg	gaaaggcgaa	tgcgcgcgcc	tatataccgc	agaaactaag	1140

cgcggttcaa	agtccgccgc	cgttgttcgc	atggtaggac	aggcccgtaa	atatagcgcc	1200
tttgaatcca	tttacttcgt	gtacgcaatg	gatagccgca	gccgtgtcta	tgtgcaatct	1260
agcacgctct	ctccgcagtc	taacgactta	ggtaaggcat	tactccgctt	taccgaggga	1320
cgccctgtga	atggcgtaga	agcgcttaaa	tggttctgca	tcaatgggtgc	taacctttgg	1380
ggatgggaca	agaaaacttt	tgatgtgctc	gtgtctaacg	tattagatga	ggaattccaa	1440
gatatgtgtc	gagacatcgc	cgagaccctc	ctcacattca	cccaatgggc	taaagctgat	1500
gcaccttatg	aattcctcgc	ttggtgcttt	gagtatgctc	aataccttga	tttgggtggat	1560
gaaggaaggg	ccgacgaatt	ccgcactcac	ctaccagtac	atcaggacgg	gtcttgttca	1620
ggcattcagc	actatagtgc	tatgcttcgc	gacgaagtag	gggcaaagc	tgtaaacctg	1680
aaaccctccg	atgcaccgca	ggatatctat	ggggcggtgg	cgcaagtggg	tatcaagaag	1740
aatgcgctat	atatgggatgc	ggacgatgca	accacgttta	cttctggtag	cgtcacgctg	1800
tccggtacag	aactgcgagc	aatggctagc	gcatgggata	gtattgggat	taccgtagc	1860
ttaacaaaaa	agcccgtgat	gaccttgcca	tatggttcta	ctcgcttaac	ttgccgtgaa	1920
tctgtgattg	attacatcgt	agacttagag	gaaaaagagg	cgcagaaggc	agtagcagaa	1980
gggcggacgg	caaacaaggt	acatcctttt	gaagacgac	gtcaagatta	cttgactccg	2040
ggcgagcgtt	acaactacat	gacggcacta	atctggcctt	ctatttctga	agtagttaag	2100
gcaccgatag	tagctatgaa	gatgatacgc	cagcttgcac	gctttgcagc	gaaacgtaat	2160
gaaggcctga	tgtacacctt	gcctactggc	ttcatcttag	aacagaagat	catggcaacc	2220
gagatgctac	gcgtgcgtac	ctgtctgatg	ggtgatatca	agatgtccct	tcaggttgaa	2280
acggatatcg	tagatgaagc	cgctatgatg	ggagcagcag	cacctaattt	cgtaacgggt	2340
catgacgcaa	gtcaccttat	ccttaccgta	tgtgaattgg	tagacaaggg	cgtaactagt	2400
atcgtgtaa	tccacgactc	ttttggtact	catgcagaca	acaccctcac	tcttagagtg	2460
gcacttaaag	ggcagatggg	tgcaatgtat	attgatggta	atgcgcttca	gaaactactg	2520
gaggagcatg	aagtgcgctg	gatggttgat	acaggtatcg	aagtacctga	gcaaggggag	2580
ttcgacctta	acgaaatcat	ggattctgaa	tacgtatttg	cctaa		2625

<210> 3
 <211> 18
 <212> DNA
 <213> Bacteriophage SP6

<400> 3
 atttaggtga cactatag 18

<210> 4
 <211> 23
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic oligonucleotide

<400> 4
 atttagggga cactatagaa gag 23

<210> 5
 <211> 22
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic oligonucleotide

<400> 5

atttagggga cactatagaa gg	22
<210> 6	
<211> 23	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Synthetic oligonucleotide	
<400> 6	
atttagggga cactatagaa ggg	23
<210> 7	
<211> 20	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Synthetic oligonucleotide	
<400> 7	
atttaggtga cactatagaa	20
<210> 8	
<211> 22	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Synthetic oligonucleotide	
<400> 8	
atttaggtga cactatagaa ga	22
<210> 9	
<211> 23	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Synthetic oligonucleotide	
<400> 9	
atttaggtga cactatagaa gag	23
<210> 10	
<211> 22	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Synthetic oligonucleotide	

<400>	10	
atttaggtga cactatagaa gg		22
<210>	11	
<211>	23	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Synthetic oligonucleotide	
<400>	11	
atttaggtga cactatagaa ggg		23
<210>	12	
<211>	23	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Synthetic oligonucleotide	
<220>		
<221>	misc_feature	
<222>	(22)	
<223>	n is a, c, t or g	
<400>	12	
atttaggtga cactatagaa gng		23
<210>	13	
<211>	24	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Synthetic oligonucleotide	
<400>	13	
catacgattt aggtgacact atag		24
<210>	14	
<211>	18	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Bacteriophage T7	
<400>	14	
taatacgact cactatag		18
<210>	15	
<211>	1480	

<212> PRT
<213> Artificial Sequence

<220>
<223> Homo sapiens

<400> 15
Met Gln Arg Ser Pro Leu Glu Lys Ala Ser Val Val Ser Lys Leu Phe
1 5 10 15
Phe Ser Trp Thr Arg Pro Ile Leu Arg Lys Gly Tyr Arg Gln Arg Leu
20 25 30
Glu Leu Ser Asp Ile Tyr Gln Ile Pro Ser Val Asp Ser Ala Asp Asn
35 40 45
Leu Ser Glu Lys Leu Glu Arg Glu Trp Asp Arg Glu Leu Ala Ser Lys
50 55 60
Lys Asn Pro Lys Leu Ile Asn Ala Leu Arg Arg Cys Phe Phe Trp Arg
65 70 75 80
Phe Met Phe Tyr Gly Ile Phe Leu Tyr Leu Gly Glu Val Thr Lys Ala
85 90 95
Val Gln Pro Leu Leu Leu Gly Arg Ile Ile Ala Ser Tyr Asp Pro Asp
100 105 110
Asn Lys Glu Glu Arg Ser Ile Ala Ile Tyr Leu Gly Ile Gly Leu Cys
115 120 125
Leu Leu Phe Ile Val Arg Thr Leu Leu Leu His Pro Ala Ile Phe Gly
130 135 140
Leu His His Ile Gly Met Gln Met Arg Ile Ala Met Phe Ser Leu Ile
145 150 155 160
Tyr Lys Lys Thr Leu Lys Leu Ser Ser Arg Val Leu Asp Lys Ile Ser
165 170 175
Ile Gly Gln Leu Val Ser Leu Leu Ser Asn Asn Leu Asn Lys Phe Asp
180 185 190
Glu Gly Leu Ala Leu Ala His Phe Val Trp Ile Ala Pro Leu Gln Val
195 200 205
Ala Leu Leu Met Gly Leu Ile Trp Glu Leu Leu Gln Ala Ser Ala Phe
210 215 220
Cys Gly Leu Gly Phe Leu Ile Val Leu Ala Leu Phe Gln Ala Gly Leu
225 230 235 240
Gly Arg Met Met Met Lys Tyr Arg Asp Gln Arg Ala Gly Lys Ile Ser
245 250 255

Glu	Arg	Leu	Val	Ile	Thr	Ser	Glu	Met	Ile	Glu	Asn	Ile	Gln	Ser	Val	260	265	270
Lys	Ala	Tyr	Cys	Trp	Glu	Glu	Ala	Met	Glu	Lys	Met	Ile	Glu	Asn	Leu	275	280	285
Arg	Gln	Thr	Glu	Leu	Lys	Leu	Thr	Arg	Lys	Ala	Ala	Tyr	Val	Arg	Tyr	290	295	300
Phe	Asn	Ser	Ser	Ala	Phe	Phe	Phe	Ser	Gly	Phe	Phe	Val	Val	Phe	Leu	305	310	315
Ser	Val	Leu	Pro	Tyr	Ala	Leu	Ile	Lys	Gly	Ile	Ile	Leu	Arg	Lys	Ile	325	330	335
Phe	Thr	Thr	Ile	Ser	Phe	Cys	Ile	Val	Leu	Arg	Met	Ala	Val	Thr	Arg	340	345	350
Gln	Phe	Pro	Trp	Ala	Val	Gln	Thr	Trp	Tyr	Asp	Ser	Leu	Gly	Ala	Ile	355	360	365
Asn	Lys	Ile	Gln	Asp	Phe	Leu	Gln	Lys	Gln	Glu	Tyr	Lys	Thr	Leu	Glu	370	375	380
Tyr	Asn	Leu	Thr	Thr	Thr	Glu	Val	Val	Met	Glu	Asn	Val	Thr	Ala	Phe	385	390	395
Trp	Glu	Glu	Gly	Phe	Gly	Glu	Leu	Phe	Glu	Lys	Ala	Lys	Gln	Asn	Asn	405	410	415
Asn	Asn	Arg	Lys	Thr	Ser	Asn	Gly	Asp	Asp	Ser	Leu	Phe	Phe	Ser	Asn	420	425	430
Phe	Ser	Leu	Leu	Gly	Thr	Pro	Val	Leu	Lys	Asp	Ile	Asn	Phe	Lys	Ile	435	440	445
Glu	Arg	Gly	Gln	Leu	Leu	Ala	Val	Ala	Gly	Ser	Thr	Gly	Ala	Gly	Lys	450	455	460
Thr	Ser	Leu	Leu	Met	Val	Ile	Met	Gly	Glu	Leu	Glu	Pro	Ser	Glu	Gly	465	470	475
Lys	Ile	Lys	His	Ser	Gly	Arg	Ile	Ser	Phe	Cys	Ser	Gln	Phe	Ser	Trp	485	490	495
Ile	Met	Pro	Gly	Thr	Ile	Lys	Glu	Asn	Ile	Ile	Phe	Gly	Val	Ser	Tyr	500	505	510
Asp	Glu	Tyr	Arg	Tyr	Arg	Ser	Val	Ile	Lys	Ala	Cys	Gln	Leu	Glu	Glu	515	520	525
Asp	Ile	Ser	Lys	Phe	Ala	Glu	Lys	Asp	Asn	Ile	Val	Leu	Gly	Glu	Gly	530	535	540

Gly	Ile	Thr	Leu	Ser	Gly	Gly	Gln	Arg	Ala	Arg	Ile	Ser	Leu	Ala	Arg	545	550	555	560
Ala	Val	Tyr	Lys	Asp	Ala	Asp	Leu	Tyr	Leu	Leu	Asp	Ser	Pro	Phe	Gly		565	570	575
Tyr	Leu	Asp	Val	Leu	Thr	Glu	Lys	Glu	Ile	Phe	Glu	Ser	Cys	Val	Cys		580	585	590
Lys	Leu	Met	Ala	Asn	Lys	Thr	Arg	Ile	Leu	Val	Thr	Ser	Lys	Met	Glu		595	600	605
His	Leu	Lys	Lys	Ala	Asp	Lys	Ile	Leu	Ile	Leu	His	Glu	Gly	Ser	Ser	610	615	620	
Tyr	Phe	Tyr	Gly	Thr	Phe	Ser	Glu	Leu	Gln	Asn	Leu	Gln	Pro	Asp	Phe	625	630	635	640
Ser	Ser	Lys	Leu	Met	Gly	Cys	Asp	Ser	Phe	Asp	Gln	Phe	Ser	Ala	Glu		645	650	655
Arg	Arg	Asn	Ser	Ile	Leu	Thr	Glu	Thr	Leu	His	Arg	Phe	Ser	Leu	Glu		660	665	670
Gly	Asp	Ala	Pro	Val	Ser	Trp	Thr	Glu	Thr	Lys	Lys	Gln	Ser	Phe	Lys		675	680	685
Gln	Thr	Gly	Glu	Phe	Gly	Glu	Lys	Arg	Lys	Asn	Ser	Ile	Leu	Asn	Pro	690	695	700	
Ile	Asn	Ser	Ile	Arg	Lys	Phe	Ser	Ile	Val	Gln	Lys	Thr	Pro	Leu	Gln	705	710	715	720
Met	Asn	Gly	Ile	Glu	Glu	Asp	Ser	Asp	Glu	Pro	Leu	Glu	Arg	Arg	Leu		725	730	735
Ser	Leu	Val	Pro	Asp	Ser	Glu	Gln	Gly	Glu	Ala	Ile	Leu	Pro	Arg	Ile		740	745	750
Ser	Val	Ile	Ser	Thr	Gly	Pro	Thr	Leu	Gln	Ala	Arg	Arg	Arg	Gln	Ser		755	760	765
Val	Leu	Asn	Leu	Met	Thr	His	Ser	Val	Asn	Gln	Gly	Gln	Asn	Ile	His	770	775	780	
Arg	Lys	Thr	Thr	Ala	Ser	Thr	Arg	Lys	Val	Ser	Leu	Ala	Pro	Gln	Ala	785	790	795	800
Asn	Leu	Thr	Glu	Leu	Asp	Ile	Tyr	Ser	Arg	Arg	Leu	Ser	Gln	Glu	Thr		805	810	815
Gly	Leu	Glu	Ile	Ser	Glu	Glu	Ile	Asn	Glu	Glu	Asp	Leu	Lys	Glu	Cys		820	825	830

Phe	Phe	Asp	Asp	Met	Glu	Ser	Ile	Pro	Ala	Val	Thr	Thr	Trp	Asn	Thr
835		840						845							
Tyr	Leu	Arg	Tyr	Ile	Thr	Val	His	Lys	Ser	Leu	Ile	Phe	Val	Leu	Ile
850		855						860							
Trp	Cys	Leu	Val	Ile	Phe	Leu	Ala	Glu	Val	Ala	Ala	Ser	Leu	Val	Val
865		870						875							
Leu	Trp	Leu	Leu	Gly	Asn	Thr	Pro	Leu	Gln	Asp	Lys	Gly	Asn	Ser	Thr
		885						890							
His	Ser	Arg	Asn	Asn	Ser	Tyr	Ala	Val	Ile	Ile	Thr	Ser	Thr	Ser	Ser
		900						905							
Tyr	Tyr	Val	Phe	Tyr	Ile	Tyr	Val	Gly	Val	Ala	Asp	Thr	Leu	Leu	Ala
915		920						925							
Met	Gly	Phe	Phe	Arg	Gly	Leu	Pro	Leu	Val	His	Thr	Leu	Ile	Thr	Val
930		935						940							
Ser	Lys	Ile	Leu	His	His	Lys	Met	Leu	His	Ser	Val	Leu	Gln	Ala	Pro
945		950						955							
Met	Ser	Thr	Leu	Asn	Thr	Leu	Lys	Ala	Gly	Gly	Ile	Leu	Asn	Arg	Phe
		965						970							
Ser	Lys	Asp	Ile	Ala	Ile	Leu	Asp	Asp	Leu	Leu	Pro	Leu	Thr	Ile	Phe
		980						985							
Asp	Phe	Ile	Gln	Leu	Leu	Leu	Ile	Val	Ile	Gly	Ala	Ile	Ala	Val	Val
995		1000						1005							
Ala	Val	Leu	Gln	Pro	Tyr	Ile	Phe	Val	Ala	Thr	Val	Pro	Val	Ile	Val
1010		1015						1020							
Ala	Phe	Ile	Met	Leu	Arg	Ala	Tyr	Phe	Leu	Gln	Thr	Ser	Gln	Gln	Leu
1025		1030						1035							
Lys	Gln	Leu	Glu	Ser	Glu	Gly	Arg	Ser	Pro	Ile	Phe	Thr	His	Leu	Val
		1045						1050							
Thr	Ser	Leu	Lys	Gly	Leu	Trp	Thr	Leu	Arg	Ala	Phe	Gly	Arg	Gln	Pro
		1060						1065							
Tyr	Phe	Glu	Thr	Leu	Phe	His	Lys	Ala	Leu	Asn	Leu	His	Thr	Ala	Asn
1075		1080						1085							
Trp	Phe	Leu	Tyr	Leu	Ser	Thr	Leu	Arg	Trp	Phe	Gln	Met	Arg	Ile	Glu
1090		1095						1100							
Met	Ile	Phe	Val	Ile	Phe	Phe	Ile	Ala	Val	Thr	Phe	Ile	Ser	Ile	Leu
1105		1110						1115							

Thr	Thr	Gly	Glu	Gly	Glu	Gly	Arg	Val	Gly	Ile	Ile	Leu	Thr	Leu	Ala			
				1125					1130					1135				
Met	Asn	Ile	Met	Ser	Thr	Leu	Gln	Trp	Ala	Val	Asn	Ser	Ser	Ile	Asp			
			1140					1145					1150					
Val	Asp	Ser	Leu	Met	Arg	Ser	Val	Ser	Arg	Val	Phe	Lys	Phe	Ile	Asp			
		1155					1160					1165						
Met	Pro	Thr	Glu	Gly	Lys	Pro	Thr	Lys	Ser	Thr	Lys	Pro	Tyr	Lys	Asn			
	1170						1175				1180							
Gly	Gln	Leu	Ser	Lys	Val	Met	Ile	Ile	Glu	Asn	Ser	His	Val	Lys	Lys			
1185					1190					1195					1200			
Asp	Asp	Ile	Trp	Pro	Ser	Gly	Gly	Gln	Met	Thr	Val	Lys	Asp	Leu	Thr			
				1205					1210					1215				
Ala	Lys	Tyr	Thr	Glu	Gly	Gly	Asn	Ala	Ile	Leu	Glu	Asn	Ile	Ser	Phe			
			1220					1225					1230					
Ser	Ile	Ser	Pro	Gly	Gln	Arg	Val	Gly	Leu	Leu	Gly	Arg	Thr	Gly	Ser			
		1235					1240					1245						
Gly	Lys	Ser	Thr	Leu	Leu	Ser	Ala	Phe	Leu	Arg	Leu	Leu	Asn	Thr	Glu			
	1250					1255					1260							
Gly	Glu	Ile	Gln	Ile	Asp	Gly	Val	Ser	Trp	Asp	Ser	Ile	Thr	Leu	Gln			
1265					1270					1275					1280			
Gln	Trp	Arg	Lys	Ala	Phe	Gly	Val	Ile	Pro	Gln	Lys	Val	Phe	Ile	Phe			
				1285					1290					1295				
Ser	Gly	Thr	Phe	Arg	Lys	Asn	Leu	Asp	Pro	Tyr	Glu	Gln	Trp	Ser	Asp			
			1300					1305					1310					
Gln	Glu	Ile	Trp	Lys	Val	Ala	Asp	Glu	Val	Gly	Leu	Arg	Ser	Val	Ile			
		1315					1320					1325						
Glu	Gln	Phe	Pro	Gly	Lys	Leu	Asp	Phe	Val	Leu	Val	Asp	Gly	Gly	Cys			
	1330					1335					1340							
Val	Leu	Ser	His	Gly	His	Lys	Gln	Leu	Met	Cys	Leu	Ala	Arg	Ser	Val			
1345					1350					1355					1360			
Leu	Ser	Lys	Ala	Lys	Ile	Leu	Leu	Leu	Asp	Glu	Pro	Ser	Ala	His	Leu			
				1365					1370					1375				
Asp	Pro	Val	Thr	Tyr	Gln	Ile	Ile	Arg	Arg	Thr	Leu	Lys	Gln	Ala	Phe			
			1380					1385					1390					
Ala	Asp	Cys	Thr	Val	Ile	Leu	Cys	Glu	His	Arg	Ile	Glu	Ala	Met	Leu			
	1395						1400					1405						

Glu Cys Gln Gln Phe Leu Val Ile Glu Glu Asn Lys Val Arg Gln Tyr
1410 1415 1420

Asp Ser Ile Gln Lys Leu Leu Asn Glu Arg Ser Leu Phe Arg Gln Ala
1425 1430 1435 1440

Ile Ser Pro Ser Asp Arg Val Lys Leu Phe Pro His Arg Asn Ser Ser
1445 1450 1455

Lys Cys Lys Ser Lys Pro Gln Ile Ala Ala Leu Lys Glu Glu Thr Glu
1460 1465 1470

Glu Glu Val Gln Asp Thr Arg Leu
1475 1480

<210> 16
<211> 140
<212> RNA
<213> Artificial Sequence

<220>
<223> 5' UTR sequence

<400> 16
ggacagaucg ccuggagacg ccauccacgc uguuuugacc uccauagaag acaccgggac 60
cgauccagcc uccgcggccg ggaacggugc auuggaacgc ggauuccccg ugccaagagu 120
gacucaccgu ccuugacacg 140

<210> 17
<211> 105
<212> RNA
<213> Artificial Sequence

<220>
<223> 3' UTR sequence

<400> 17
cggguggcau cccugugacc ccuccccagu gccucuccug gccuggaag uugccacucc 60
agugcccacc agccuugucc uaauaaaauu aaguugcauc aagcu 105

<210> 18
<211> 105
<212> RNA
<213> Artificial Sequence

<220>
<223> 3' UTR sequence

<400> 18
ggguggcauc ccugugaccc cucccccagug ccucuccugg cccuggaagu ugccacucca 60
gugccacca gccuuguccu aaauaaaaua aguugcauca aagcu 105

<210> 19

<211> 582
<212> DNA
<213> Artificial Sequence

<220>

<223> Homo sapiens EPO sequence, codon optimized, reference

<400> 19

```
atgggtgtgc acgaatgtcc tgcttggtg tggctccttc tctccctgct gtccctgcct 60
cttggactcc cgggtgcttg agcaccctcg agactgatct gcgacagcag ggtgctcgag 120
cgctacctcc tggaagccaa ggaagccgaa aacatcacta ctggctgcgc cgaacactgc 180
tccctgaacg agaacatcac cgtgccggac accaaggtca acttctacgc gtggaagaga 240
atggaggctc gacagcaagc cgtggaagtg tggcagggac ttgcgctcct gtcggaagcc 300
gtgctgaggg gacaagccct gctcgtgaac agctcacagc cttgggagcc cctgcagctg 360
catgtcgaca aggccgtgtc cggactgcgc tctactacca ctctgctgag ggccttgggt 420
gcccagaaag aggctatttc cccaccggat gcagcctcgg cagctcctct gcggaccatt 480
acggcggaca ctttcggaa gctgttccgc gtctacagca atttcctccg ggggaagttg 540
aaactgtata ccggcgaagc ctgtcggact ggcgatcgct ga 582
```

<210> 20
<211> 582
<212> DNA
<213> Artificial Sequence

<220>

<223> Homo sapiens EPO sequence, codon optimized, #1

<400> 20

```
atgggggttc atgagtgtcc agcttggtt tggctcctgc tcagcttgct tagtctccct 60
ttgggcctgc ccgtgctggg cggccctcca cgcttgatct gtgacagcag ggtcttggaa 120
cggatatttc ttgaagctaa agaagctgag aacataacaa cgggatgtgc tgaacattgc 180
tccttgaacg aaaacatcac agttcccgac acaaaagtca atttttacgc atggaagcgg 240
atggaggttg gccagcaagc tgtggaggtc tggcaagggc tggctcttct cagtgaagcc 300
gtgctgcgcg gacaagcact cttggtgaac tccagccagc cctgggagcc cttcagctc 360
catgtcgata aagcagttag cggcctccga tcattgacta ccctccttag ggctttgggt 420
gcacaaaaag aggccatttc accaccggac gcggcaagtg ctgctccgtt gcgaactata 480
actgtcgaca cttccggaa actttttcgg gtatattcca actttctcag ggggaaactc 540
aagctctaca ccggcgaggc gtgccgaact ggagaccgct ga 582
```

<210> 21
<211> 582
<212> DNA
<213> Artificial Sequence

<220>

<223> Homo sapiens EPO sequence, codon optimized, #2

<400> 21

```
atgggcgtac atgaatgtcc ggcatggctt tggctgctgc tgtccctgct gagtttgccg 60
ctgggcctcc ccgtcctcgg cgctcccccg agactcattt gcgactctag ggtcctcgaa 120
cgctatctgc tggaagcaaa agaagctgag aacataacta caggatgcgc tgagcactgt 180
tccttgaatg agaatatcac agtacctgac actaaggtga atttttacgc atggaaacgc 240
atggaagtgg gtcagcaggc cgtggaagtg tggcagggcc tggcgctgct gtccgaggct 300
gttcttagag gccaaagcct gttggtcaat tcctctcaac cctgggagcc cctccagctg 360
```

catgttgata	aagccgtctc	tggtctccgg	tcccttacca	ccctgctcag	ggcacttggc	420
gcacagaagg	aagctatctc	ccccccagac	gctgccagtg	ccgccccct	ccggactatt	480
accgccgata	ctttcaggaa	actgtttcga	gtctatagca	attttctccg	cgggaaactg	540
aagctgtata	caggtgaggc	ctgcaggaca	ggagatcgct	ga		582

<210> 22
 <211> 582
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Homo sapiens EPO sequence, codon optimized, #3

<400> 22						
atgggctg	acgaatgtcc	tgcttggtg	tggtctgctg	tgagtctgct	gtctctgcct	60
ctgggactg	ctgttcttgg	agccccctc	agactgatct	gcgacagcag	agtgctggaa	120
agatacctg	tggaagccaa	agaggccgag	aacatcacaa	caggctgtgc	cgagcactgc	180
agcctgaacg	agaatatcac	cgtgcctgac	accaaagtga	acttctacgc	ctggaagcgg	240
atggaagtgg	gacagcaggc	tgtggaagtt	tggcaaggac	tggccctgct	gtctgaagct	300
gttctgagag	gacaggctct	gctgggtcaat	agctctcagc	cttggaacc	tctccagctg	360
catgtggata	aggccgtgtc	tggcctgaga	agcctgacaa	cactgctgag	agccctggga	420
gccagaaaag	aggccatttc	tccacctgat	gctgccagcg	ctgccccctc	gagaacaatc	480
accgccgaca	ccttcagaaa	gctgttccgg	gtgtacagca	acttcctgcg	gggcaagctg	540
aaactgtaca	ccggcgaagc	ctgcagaacc	ggcgatagat	aa		582

<210> 23
 <211> 582
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Homo sapiens EPO sequence, codon optimized, #4

<400> 23						
atgggggtg	acgagtgtcc	tgcttggtg	tggttgctg	tgtccctgct	gtctctgcca	60
ctgggactg	cagtgtctgg	agctccacct	aggctgatct	gcgacagccg	ggctctggag	120
aggtacctg	tcgaggccaa	ggaggccgag	aacattacca	caggctgcgc	cgagcactgc	180
agcctgaacg	agaacattac	agtgtccgat	acaaaggtga	acttctacgc	ctggaagagg	240
atggaggtgg	gccagcaggc	cgtggaggtg	tggcaggggc	tggccctgct	gagcagggcc	300
gtgctgaggg	gccaagccct	gctgggtcaac	agcagccagc	cttgggagcc	cctgcagctc	360
cacgtggaca	aggctgtgtc	tggcttgagg	tctctcacia	cattgctgag	ggccctgggc	420
gcacagaaag	aagctatcag	cccacctgat	gccgctagt	ccgctccact	gcggacaatt	480
accgccgata	cctttagaaa	attgttcagg	gtctactcca	actttttgcg	cgggaagctg	540
aagctctata	ccggcgaggc	ctgccggaca	ggggacagat	ga		582

<210> 24
 <211> 582
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Homo sapiens EPO sequence, codon optimized, #5

<400> 24

atgggagtgc	acgaatgtcc	tgcattggctc	tggctcctgc	tgtctctcct	gagcctgcca	60
ctgggactcc	cagtgtctggg	agcaccctcc	aggctgatct	gcgattctcg	ggtgctggag	120
cgctacctgc	tcgaggctaa	ggaggccgag	aatatcacta	ctgggtgtgc	cgaacactgt	180
agcctcaatg	aaaacattac	agtcccagat	accaaggtga	acttttatgc	atggaagagg	240
atggagggtcg	ggcagcaggc	agtggagggtg	tggcagggac	tggctctgct	gtccgaagcc	300
gtgctcagag	gtcaggccct	gctgggttaat	tccagccagc	cttgggaacc	tctgcagctg	360
catgtggaca	aggcagtgtc	tggcctgaga	tcctttacta	cactgctgag	agcactgggg	420
gctcagaaaag	aagctatttc	cccaccagac	gccgcctcag	cagcacctct	ccggaccatc	480
actgctgaca	ccttccgcaa	gctcttttagg	gtgtactcca	acttcctgcg	cgggaagctc	540
aagctgtaca	ccggcgaagc	ctgcaggacc	ggggatcgct	ga		582

<210> 25
 <211> 4443
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Homo sapiens CFTR sequence, codon optimized, reference

<400>	25					
atgcaacgct	ctcctcttga	aaaggcctcg	gtgggtgtcca	agctcttctt	ctcgtggact	60
agaccatcc	tgagaaaggg	gtacagacag	cgcttggagc	tgtccgatat	ctatcaaadc	120
ccttccgtgg	actccgcgga	caacctgtcc	gagaagctcg	agagagaatg	ggacagagaa	180
ctcgcctcaa	agaagaaccc	gaagctgatt	aatgcgctta	ggcgggtgctt	tttctggcgg	240
ttcatgttct	acggcatctt	cctctacctg	ggagagggtca	ccaaggccgt	gcagcccctg	300
ttgctgggac	ggattattgc	ctcctacgac	cccgacaaca	aggaagaaaag	aagcatcgct	360
atctacttgg	gcatcgggtc	gtgcctgctt	ttcatcgctc	ggaccctctt	gttgcacatc	420
gctatttttcg	gcctgcatca	cattggcatg	cagatgagaa	ttgccatgtt	ttccctgatc	480
tacaagaaaa	ctctgaagct	ctcgagccgc	gtgcttgaca	agattttccat	cggccagctc	540
gtgtccctgc	tctccaacaa	tctgaacaag	ttcgacgagg	gcctcgccct	ggcccacttc	600
gtgtggatcg	cccctctgca	agtggcgctt	ctgatgggccc	tgatctggga	gctgctgcaa	660
gcctcggcat	tctgtgggct	tggattcctg	atcgtgctgg	cactgttcca	ggccggactg	720
gggcggatga	tgatgaagta	cagggaccag	agagccggaa	agattttccga	acggctgggtg	780
atcacttcgg	aaatgatcga	aaacatccag	tcagtgaagg	cctactgctg	ggaagaggcc	840
atggaaaaga	tgattgaaaa	cctccggcaa	accgagctga	agctgaccgc	caaggccgct	900
tacgtgcgct	atttcaactc	gtccgctttc	ttcttctccg	ggttcttctg	ggtgtttctc	960
tccgtgctcc	cctacgccct	gattaaggga	atcatcctca	ggaagatctt	caccaccatt	1020
tccttctgta	tcgtgctccg	catggccgtg	accggcagct	tcccatgggc	cgtgcagact	1080
tggtacgact	ccctgggagc	cattaacaag	atccaggact	tccttcaaaa	gcaggagtac	1140
aagaccctcg	agtacaacct	gactactacc	gaggtcgtga	tggaaaacgt	caccgccttt	1200
tgggaggagg	gatttggcga	actgttcgag	aaggccaagc	agaacaacaa	caaccgcaag	1260
acctcgaacg	gtgacgactc	cctcttcttt	tcaaaactta	gcctgctcgg	gacgcccgtg	1320
ctgaaggaca	ttaacttcaa	gatcgaaaaga	ggacagctcc	tggcgggtggc	cggatcgacc	1380
ggagccggaa	agacttccct	gctgatgggtg	atcatgggag	agcttgaacc	tagcgaggga	1440
aagatcaagc	actccggccg	catcagcttc	tgtagccagt	tttcttggtg	catgcccggg	1500
accattaagg	aaaacatcat	cttcggcgctg	tcctacgatg	aataccgcta	ccggtccgtg	1560
atcaaagcct	gccagctgga	agaggatatt	tcaaagttcg	cggagaaaaga	taacatcgtg	1620
ctgggcgaag	gggggtattac	cttgtcgggg	ggccagcggg	ctagaatctc	gctggccaga	1680
gccgtgtata	aggacgccga	cctgtatctc	ctggactccc	ccttcggata	cctggacgtc	1740
ctgaccgaaa	aggagatctt	cgaatcgtgc	gtgtgcaagc	tgatggctaa	caagactcgc	1800
atcctcgtga	cctccaaaat	ggagcacctg	aagaaggcag	acaagattct	gattctgcat	1860
gaggggtcct	cctactttta	cggcaccttc	tcggagttgc	agaacttgca	gcccgaacttc	1920
tcatcgaagc	tgatgggttg	cgacagcttc	gaccagttct	ccgccgaaag	aaggaaactcg	1980
atcctgacgg	aaaccttgca	ccgcttctct	ttggaaggcg	acgcccctgt	gtcatggacc	2040

gagactaaga	agcagagctt	caagcagacc	ggggaattcg	gcgaaaagag	gaagaacagc	2100
atcttgaacc	ccattaactc	catccgcaag	ttctcaatcg	tgcaaaagac	gccactgcag	2160
atgaacggca	ttgaggagga	ctccgacgaa	ccccttgaga	ggcgcctgtc	cctggtgccg	2220
gacagcgagc	agggagaagc	catcctgcct	cggatttccg	tgatctccac	tggtccgacg	2280
ctccaagccc	ggcggcggca	gtccgtgctg	aacctgatga	cccacagcgt	gaaccagggc	2340
caaaacattc	accgcaagac	taccgcatcc	acccggaaag	tgtccctggc	acctcaagcg	2400
aatcttaccg	agctcgacat	ctactcccgg	agactgtcgc	aggaaaccgg	gctcgaaatt	2460
tccgaagaaa	tcaacgagga	ggatctgaaa	gagtgttctt	tcgacgatat	ggagtcgata	2520
cccgccgtga	cgacttggaa	cacttatctg	cggtacatca	ctgtgcacaa	gtcattgata	2580
ttcgtgctga	tttggcgcct	ggtgattttc	ctggccgagg	tcgcggcctc	actggtggtg	2640
ctctggctgt	tgggaaacac	gcctctgcaa	gacaagggaa	actccacgca	ctcgagaaac	2700
aacagctatg	ccgtgattat	cacttccacc	tcctcttatt	acgtgttcta	catctacgtc	2760
ggagtggcgg	ataccctgct	cgcgatgggt	ttcttcagag	gactgccgct	ggtccacacc	2820
ttgatcaccg	tcagcaagat	tcttcaccac	aagatgttgc	atagcgtgct	gcaggccccc	2880
atgtccaccc	tcaacactct	gaaggccgga	ggcattctga	acagattctc	caaggacatc	2940
gctatcctgg	acgatctcct	gccgcttacc	atctttgact	tcattccagct	gctgctgata	3000
gtgattggag	caatcgcagt	ggtggcgggt	ctgcagcctt	acattttcgt	ggccactgtg	3060
ccggtcattg	tggcgttcat	catgctgcgg	gcctacttcc	tcctaaaccag	ccagcagctg	3120
aagcaactgg	aatccgaggg	acgatccccc	atcttcactc	accttgtgac	gtcgttgaag	3180
ggactgtgga	ccctccgggc	tttcggacgg	cagccctact	tcgaaaccct	cttcacaaag	3240
gccctgaacc	tccacaccgc	caattgggtt	ctgtacctgt	ccaccctgcg	gtggttccag	3300
atgcgcatcg	agatgatttt	cgtcatcttc	ttcatcgcgg	tcacattcat	cagcatcctg	3360
actaccggag	agggagaggg	acgggtcggg	ataatcctga	ccctcgccat	gaacattatg	3420
agcaccctgc	agtgggcagt	gaacagctcg	atcgacgtgg	acagcctgat	gcgaagcgtc	3480
agccgcgtgt	tcaagttcat	cgacatgcct	actgagggaa	aaccactaa	gtccactaag	3540
ccctacaaaa	atggccagct	gagcaaggct	atgatcatcg	aaaactccca	cgtgaagaag	3600
gacgatattt	ggccctccgg	aggtcaaatg	accgtgaagg	acctgaccgc	aaagtacacc	3660
gagggaggaa	acgccattct	cgaaaacatc	agcttctcca	tttcgccggg	acagcgggtc	3720
ggccttctcg	ggcggaccgg	ttccgggaag	tcaactctgc	tgtcggcttt	cctccggctg	3780
ctgaataacc	agggggaaat	ccaaattgac	ggcgtgtctt	gggattccat	tactctgcag	3840
cagtggcggg	aggccttcgg	cgtgatcccc	cagaagggtg	tcattcttct	gggtaccttc	3900
cggagaagcc	tggatcctta	cgagcagtgg	agcgaccaag	aaatctggaa	ggtcgccgac	3960
gaggtcggcc	tgcgctccgt	gattgaacaa	tttcctggaa	agctggactt	cgtgctcgtc	4020
gacgggggat	gtgtcctgtc	gcacggacat	aagcagctca	tgtgcctcgc	acggtccgtg	4080
ctctccaagg	ccaagattct	gctgctggac	gaaccttcgg	cccacctgga	tccggtcacc	4140
taccagatca	tcaggaggac	cctgaagcag	gcctttgccg	attgcaccgt	gattctctgc	4200
gagcaccgca	tcgaggccat	gctggagtgc	cagcagttcc	tggctatcga	ggagaacaag	4260
gtccgccaat	acgactccat	tcaaaagctc	ctcaacgagc	ggtcgtgttt	cagacaagct	4320
atttcaccgt	ccgatagagt	gaagctcttc	ccgcacgcga	acagctcaaa	gtgcaaatcg	4380
aagccgcaga	tcgcagcctt	gaaggaagag	actgaggaag	aggtgcagga	caccggcctt	4440
taa						4443

<210> 26
 <211> 4443
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Homo sapiens CFTR sequence, codon optimized, hCFTR #1

<400>	26					
atgcagcggg	ccccgctcga	aaaggccagt	gtcgtgtcca	aactcttctt	ctcatggact	60
cgccctatcc	ttagaaaggg	gtatcggcag	aggcttgagt	tgtctgacat	ctaccagatc	120
ccctcggtag	attcggcgga	taacctctcg	gagaagctcg	aacgggaatg	ggaccgcgaa	180

ctcgcgtcta	agaaaaaccc	gaagctcatc	aacgcactga	gaaggtgctt	cttctggcgg	240
ttcatgttct	acggtatctt	cttgtatctc	ggggagggtca	caaaagcagt	ccaacccctg	300
ttgttgggtc	gcattatcgc	ctcgtacgac	cccgataaca	aagaagaacg	gagcatcgcg	360
atctacctcg	ggatcggact	gtgtttgctt	ttcatcgtca	gaacactttt	gttgcattcca	420
gcaatcttcg	gcctccatca	catcggtatg	cagatgcgaa	tcgctatgtt	tagcttgatc	480
tacaaaaaga	cactgaaact	ctcgtcgcgg	gtgttggata	agatttccat	cggtcagttg	540
gtgtccctgc	ttagtaataa	cctcaacaaa	ttcgatgagg	gactggcgct	ggcacatttc	600
gtgtggattg	ccccgttgca	agtcgccctt	ttgatgggcc	ttatttggga	actcttgcat	660
gcatctgcct	tttgtggcct	gggatttctg	attgtgttgg	cattgtttca	ggctgggctt	720
gggcggatga	tgatgaagta	tcgcgaccag	agagcgggta	aaatctcgga	aagactcgtc	780
atcacttcgg	aaatgatcga	aaacatccag	tcgggtcaaag	cctattgctg	ggaagaagct	840
atggagaaga	tgattgaaaa	cctccgccaa	actgagctga	aactgaccgc	caaggcggcg	900
tatgtccggg	atttcaattc	gtcagcgttc	ttcttttccg	ggttcttctg	tgtctttctc	960
tcggttttgc	cttatgcctt	gattaagggg	attatcctcc	gcaagatttt	caccacgatt	1020
tcgttctgca	ttgtattgca	catggcagtg	acacggcaat	ttccgtgggc	cgtgcagaca	1080
tggtatgact	cgcttggagc	gatcaacaaa	atccaagact	tcttgcaaaa	gcaagagtac	1140
aagaccctgg	agtacaatct	tactactacg	gaggtagtaa	tggagaatgt	gacggctttt	1200
tgggaagagg	gttttggaga	gctcttcgag	aaagcaaagc	agaataacaa	caaccgcaag	1260
acctcaaatg	gggacgattc	cctgtttttc	tcgaacttct	ccctgctcgg	aacaccctg	1320
ttgaaggaca	tcaatttcaa	gattgagagg	ggacagcttc	tcgcggtagc	gggaagcact	1380
gggtgcgggaa	aaactagcct	cttgatgggtg	attatggggg	agcttgagcc	cagcgagggg	1440
aagattaaac	actccgggcg	tatctcattc	tgtagccagt	tttcatggat	catgcccgga	1500
accattaaag	agaacatcat	tttcggagta	tcctatgatg	agtaccgata	cagatcggtc	1560
attaaggcgt	gccagttgga	agaggacatt	tctaagttcg	ccgagaagga	taacatcgtc	1620
ttgggagaag	gggggtattac	attgtcggga	gggcagcgag	cgcggtatcag	cctcgcgaga	1680
gcggtataca	aagatgcaga	tttgtacctg	ctcgattcac	cgtttggata	cctcgacgta	1740
ttgacagaaa	aagaaatctt	cgagtcgtgc	gtgtgtaaac	ttatggctaa	taagacgaga	1800
atcctgggtga	catcaaaaat	ggaacacctt	aagaaggcgg	acaagatcct	gatcctccac	1860
gaaggatcgt	cctactttta	cggcactttc	tcagagttgc	aaaacttgca	gccggacttc	1920
tcaagcaaac	tcatgggggtg	tgactcattc	gaccagttca	gcgcggaacg	gcggaactcg	1980
atcttgacgg	aaacgctgca	ccgattctcg	cttgagggtg	atgccccggt	atcgtggacc	2040
gagacaaaga	agcagtcgtt	taagcagaca	ggagaatttg	gtgagaaaag	aaagaacagt	2100
atcttgaatc	ctattaactc	aattcgcaag	ttctcaatcg	tccagaaaaac	tccactgcag	2160
atgaatggaa	ttgaagagga	ttcggacgaa	cccctggagc	gcaggcttag	cctcgtgccg	2220
gattcagagc	aaggggaggc	cattcttccc	cggatttcgg	tgatttcaac	cggacctaca	2280
cttcaggcga	ggcgaaggca	atccgtgctc	aacctcatga	cgcattcggg	aaaccagggg	2340
caaaacattc	accgcaaaac	gacggcctca	acgagaaaag	tgtcacttgc	accccaggcg	2400
aatttgactg	aactcgacat	ctacagccgt	aggctttcgc	aagaaaccgg	acttgagatc	2460
agcgaagaaa	tcaatgaaga	agatttgaaa	gagtgtttct	ttgatgacat	ggaatcaatc	2520
ccagcgggtga	caacgtggaa	cacatacttg	cgttacatca	cgggtgcacaa	gtccttgatt	2580
ttcgtcctca	tttggtgcct	cgtgatcttt	ctcgtgagg	tcgcagcgtc	acttgtggtc	2640
ctctggctgc	ttggtaatac	gcccttgcaa	gacaaaaggca	attctacaca	ctcaagaaac	2700
aattcctatg	ccgtgattat	cacttctaca	agctcgtatt	acgtgtttta	catctacgta	2760
ggagtggccg	acactctgct	cgcgatgggt	ttcttccgag	gactcccact	cgttcacacg	2820
cttatcactg	tctccaagat	tctccaccat	aagatgcttc	atagcgtact	gcaggctccc	2880
atgtccacct	tgaatacgct	caaggcggga	ggtattttga	atcgcttctc	aaaagatatt	2940
gcaatttttg	atgaccttct	gcccctgacg	atcttcgact	tcattccagtt	gttgcctgatc	3000
gtgattgggg	ctattgcagt	agtcgctgtc	ctccagcctt	acatttttgt	cgcgaccggt	3060
ccggtgatcg	tggcgtttat	catgctgcgg	gcctattttc	tgcagacgtc	acagcagctt	3120
aagcaactgg	agtctgaagg	gaggctgcct	atcttttacgc	atcttgtgac	cagtttgaag	3180
ggattgtgga	cgttgcgcgc	ctttggcagg	cagccctact	ttgaaacact	gttccacaaa	3240
gcgctgaatc	tccatacggc	aaattgggtt	ttgtatttga	gtaccctccg	atggtttcag	3300
atgcgcattg	agatgatttt	tgtgatcttc	tttatcgcgg	tgacttttat	ctccatcttg	3360
accacgggag	agggcgaggg	acgggtcggg	attatcctga	cactcgccat	gaacattatg	3420

agcactttgc	agtgggcagt	gaacagctcg	attgatgtgg	atagcctgat	gaggtccgtt	3480
tcgagggtct	ttaagttcat	cgacatgccg	acggaggga	agcccacaaa	aagtacgaaa	3540
ccctataaga	atgggcaatt	gagtaaggta	atgatcatcg	agaacagtca	cgtgaagaag	3600
gatgacatct	ggcctagcgg	gggtcagatg	accgtgaagg	acctgacggc	aaaatacacc	3660
gagggaggga	acgcaatcct	tgaaaacatc	tcgttcagca	ttagccccgg	tcagcgtgtg	3720
gggttgctcg	ggaggaccgg	gtcaggaaaa	tcgacgttgc	tgtcggcctt	cttgagactt	3780
ctgaatacag	agggtgagat	ccagatcgac	ggcgtttcgt	gggatagcat	caccttgacg	3840
cagtggcggg	aagcgtttgg	agtaatcccc	caaaaggctt	ttatcttttag	cggaaccttc	3900
cgaagaatc	tcgatcctta	tgaacagtgg	tcagatcaag	agatttgga	agtcgcggac	3960
gaggttggcc	ttcggagtgt	aatcgagcag	tttccgggaa	aactcgactt	tgtccttgta	4020
gatgggggat	gcgtcctgtc	gcatgggcac	aagcagctca	tgtgcctggc	gcgatccgtc	4080
ctctctaaag	cgaaaattct	tctcttggat	gaaccttcgg	cccatctgga	cccggtaacg	4140
tatcagatca	tcagaaggac	acttaagcag	gcgtttgccg	actgcacggg	gattctctgt	4200
gagcatcgta	tcgaggccat	gctcgaatgc	cagcaatttc	ttgtcatcga	agagaataag	4260
gtccgccagt	acgactccat	ccagaagctg	cttaatgaga	gatcattgtt	ccggcaggcg	4320
atttcaccat	ccgatagggt	gaaacttttt	ccacacagaa	attcgtcgaa	gtgcaagtcc	4380
aaaccgcaga	tcgcggcctt	gaaagaagag	actgaagaag	aagttcaaga	cacgcgtctt	4440
taa						4443

<210> 27
 <211> 4443
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Homo sapiens CFTR sequence, codon optimized, hCFTR #2

<400>	27					
atgcagcggt	ctcccctgga	gaaggcttct	gtggtgagta	aacttttttt	ctcctggacc	60
agacctatcc	tgaggaaagg	ctacaggcag	agactggagc	tctctgacat	ataccagata	120
ccttcagtcg	atagcgccga	caacctgagc	gagaagctgg	aacgcgagtg	ggacagagag	180
ctggcaagca	agaagaacct	aaagctgatt	aatgccctga	gaagggtgtt	cttctggaga	240
ttcatgttct	acggaatctt	tctgtatctg	ggggagggtta	caaaggctgt	gcaacccttg	300
ctgctcggca	gaatcatcgc	ctcatacgat	ccagacaaca	aggaagaaag	aagcatcgcc	360
atctacctgg	gcattggcct	ctgcctcctg	tttattgtgc	ggactctgct	gctgcaccca	420
gcaattttcg	ggttgcatca	tattggcatg	cagatgcgca	ttgctatgtt	ttccctcatc	480
tacaaaaaga	cactgaaact	cagctcccgg	gtgctggaca	agatctccat	cggccaactg	540
gtgtctctcc	tgagcaataa	cttgaataag	ttcgacgaag	ggctggccct	ggcacacttc	600
gtgtggattg	ccccctgca	ggtggccctg	ctgatgggac	tgatttggga	actgctgcag	660
gctagcgctt	tctgcggcct	ggggttcctg	atcgtgctgg	cactgtttca	ggcaggcctg	720
ggccgtatga	tgatgaagta	cagagaccag	agggccggga	agatctccga	acggctcggt	780
attacctctg	agatgatcga	gaacattcag	tctgtgaaag	cctactgctg	ggaggaggct	840
atggagaaga	tgatcgagaa	tctgagacag	accgagctga	agctgaccag	aaaggccgcc	900
tacgtgaggt	acttcaacag	cagtgccttc	ttcttctctg	ggttcttcgt	tgtgtttctg	960
agcgtgctgc	catacgtctt	catcaaaggc	atcatcctgc	ggaagatctt	caccaccatc	1020
agcttttgca	tcgtgcttag	aatggccgtg	acacggcagt	tcccatgggc	cgttcaaact	1080
tggtatgatt	ccctgggcgc	catcaaaaa	atccaggatt	tcctgcagaa	gcaggaatac	1140
aagacactcg	aatataacct	cacaactact	gaggtgggtta	tggagaacgt	gactgccttc	1200
tgggaggagg	ggttcggaga	gctttttgag	aaggccaaac	agaataataa	taaccgcaaa	1260
accagcaacg	gcgacgacag	cctgttcttc	tccaattttt	ctctcctggg	aacaccctgc	1320
ctcaaagaca	tcaactttaa	gatcgagagg	ggccagctgc	tcgccgtcgc	cggatccaca	1380
ggcgccggca	agacctctct	gctgatgggt	atcatgggcg	aactggagcc	ctccgagggc	1440
aagattaagc	actcaggaag	aatctccttt	tgtagccagt	tcagttggat	tatgcccggc	1500
actattaagc	agaatatcat	ttttgggggtg	agctatgatg	agtatcggtg	tcggagcggt	1560

atcaaagcct	gtcagctgga	ggaggatatc	agcaagttcg	cagagaagga	taatattgtg	1620
ctgggagagg	gaggaatcac	cctgagcgga	ggccagagag	ccagaatctc	actggcccgg	1680
gccgtctaca	aggacgccga	cctttacctt	ctggacagtc	cctttggata	tctggatgtg	1740
ctgactgaaa	aggagatctt	cgagtcttgt	gtgtgcaagc	tgatggctaa	caagaccggg	1800
atcctagtga	ctagtaagat	ggagcacctg	aagaaggcag	acaagatctt	gattctgcac	1860
gagggatcct	cttactttta	cggcaccttt	agcgagctgc	agaacctcca	gcccgatctt	1920
tcatctaagc	tgatgggctg	tgatagcttc	gaccagttct	ctgccgagcg	cagaaacagc	1980
atcctgacag	agacactgca	ccggttttca	ctggagggcg	acgcccctgt	cagctggacc	2040
gagacaaaaa	agcagtcttt	caagcagaca	ggcgagtctg	gcgagaagcg	caaaaacagc	2100
atcctgaatc	caatcaactc	tataaggaag	tttagcatcg	tgcagaagac	accctccag	2160
atgaacggca	tcgaagagga	cagtgcagag	cccctggagc	ggcgcctgag	cctcgtgcct	2220
gacagcgaac	agggcgaggc	catcctgcct	aggatcagcg	tgatttcaac	cgggccaaca	2280
ctgcaggcta	ggagaagaca	gtcagtgcct	aacctgatga	cacatagcgt	gaatcaggga	2340
cagaacatcc	atcgaaaaac	cacagcctct	actcgaaaag	tgctactggc	tcctcaggct	2400
aatctgacag	agctggacat	ctatagcagg	aggctgagcc	aggagacagg	cctggagatc	2460
agtgaggaga	tcaacgaaga	ggacctgaag	gagtgccttt	tcgatgacat	ggagagtatc	2520
cccgccgtca	ccacctggaa	tacctacctc	cggtacatca	cagtgcacaa	gtccctcatc	2580
tttgtgctga	tttggctgcct	cgtgatcttt	ctcgcagaag	tggccgcctc	cctgggtggtg	2640
ctgtggctgt	tggggaatac	tccactgcag	gacaaaaggca	attctacaca	cagcaggaat	2700
aattcctatg	ccgtgattat	caccagcaca	tcctcttact	acgtgttcta	catctacgtg	2760
ggagtggcag	atactctgct	tgcaatgggc	ttcttcaggg	ggctgcccct	ggtgcacaca	2820
ctgatcacag	tgtccaagat	cctccaccat	aaaatgctcc	acagcgtgct	gcaggcacc	2880
atgagcacc	tgaacacact	gaaggccggc	ggcatcctga	atcgcttttc	caaagacatc	2940
gccatcctcg	acgatctcct	gccactgacc	atcttcgatt	ttatccagct	gctgctgac	3000
gtgatcgggg	ccatcgccgt	ggtggccgtg	ctgcagccat	acattttcgt	ggctacagtg	3060
cccgtgatcg	ttgcctttat	catgctgaga	gcctacttcc	tgacagcttc	tcagcagctg	3120
aagcagctgg	agagcgaagg	gagaagcccc	atcttcactc	acctggtgac	aagcctgaag	3180
ggactctgga	ccctgagagc	cttcggccgg	cagccctatt	tcgagaccct	gtttcacaa	3240
gccctcaacc	tgacacacagc	caactggttc	ctctacctgt	ccaccctgag	gtggttcag	3300
atgaggattg	aaatgatctt	cgtgattttt	ttcatcgccg	tgacattcat	tagcattctg	3360
accaccggcg	agggggaggg	gagagtgggc	atcatcctga	cccttgccat	gaacattatg	3420
agcacactgc	agtgggcccgt	gaatagtagt	atcgacgtgg	acagtctgat	gaggtccgtg	3480
agccgggtgt	tcaagttcat	tgacatgccc	acagaaggga	aaccaccaa	aagcaccaag	3540
ccctacaaga	acgggcagct	gtccaagggt	atgatcatcg	agaactctca	cgtgaagaag	3600
gacgacattt	ggcccagcgg	cggccagatg	acagtgaag	atctgaccgc	caaatacacc	3660
gagggaggca	acgccatcct	cgaaaacatt	agcttctcta	tcagccctgg	acagagggtg	3720
ggcctgctgg	gccggacagg	ctcagggaa	agtactctgc	tgctagcatt	cctgaggctc	3780
ctgaacacag	agggcgagat	ccagattgac	ggcgtgtcct	gggactccat	cacctgcag	3840
cagtggcgga	aggctttcgg	ggtgatcccc	cagaagggtg	tcattctttag	cggcactttc	3900
agaaagaatc	tggaccctta	tgagcagtgg	agtaccagg	agatctggaa	agtggccgat	3960
gaggtcggac	tgaggagcgt	gatcgagcag	tttccaggga	agctggactt	tgtgctggtg	4020
gatggcggat	gcgtgctgtc	tcacggccat	aaacagctga	tgtgtctggc	ccggtccgtg	4080
ctgtctaagg	ccaagatcct	gctgctggac	gaaccctccg	cccacctgga	ccccgtgaca	4140
taccagatca	tcaggagaac	tctcaagcag	gccttcgccg	actgtaccgt	gattctgtgc	4200
gagcaccgca	ttgaagctat	gctggagtgt	cagcagttcc	tggtgatcga	ggaaaataag	4260
gtgaggcagt	acgacagcat	ccagaagctg	ctgaacgagc	gctccctgtt	ccgccaggct	4320
atctcccat	cagaccgggt	gaagctcttc	ccccacagaa	actcctcaaa	gtgcaagtc	4380
aagccccaga	tcgccgccct	gaaggaggag	accgaggagg	aggtgcagga	caccaggctg	4440
tga						4443

<210> 28

<211> 4443

<212> DNA

<213> Homo sapiens CFTR sequence, codon optimized, hCFTR #3

<400> 28

atgcagcgct	cgcctctgga	aaaggcgagc	gtcgtgtcaa	agctattctt	ttcttggacc	60
cggccccattc	tcaggaaggg	ctacaggcag	aggctggagt	tgagcgacat	ctatcagatt	120
ccttccgtgg	acagcgccga	caacctgagc	gagaagctgg	aaagggagt	ggaccgcgaa	180
ctggcaagca	aaaagaaccc	caagctgata	aatgccctga	gaagggtgtt	cttttggaga	240
ttcatgttct	acgggatctt	tctgtatctg	ggcgagggtta	caaaggctgt	gcagcccctg	300
ctgctcggca	gaatcatcgc	ctcatatgat	ccagacaaca	aggaagaaag	aagcatcgcc	360
atctacctgg	gcattggcct	ctgcctcctg	tttattgtgc	ggactctgct	gctgcaccca	420
gcaatttttcg	ggttgcatca	tattggcatg	cagatgcgca	ttgctatgtt	ttccctcatc	480
tacaaaaaga	cactgaaact	cagctcccgg	gtgctggaca	agatctccat	cggccaactg	540
gtgtctctcc	tgagcaataa	cttgaataag	ttcgacgaag	ggctggccct	ggcacacttc	600
gtgtggattg	ccccctgca	ggtggccctg	ctgatgggac	tgatttggga	actgctgcag	660
gctagcgctt	tctgcggcct	ggggttcctg	atcgtgctgg	cactgtttca	ggcaggcctg	720
ggccgtatga	tgatgaagta	cagagaccag	agggccggga	agatctccga	acggctcgtt	780
attacctctg	agatgatcga	gaacattcag	tctgtgaaag	cctactgctg	ggaggaggct	840
atggagaaga	tgatcgagaa	tctgagacag	accgagctga	agctgaccag	aaaggccgcc	900
tacgtgaggt	acttcaacag	cagtgccttc	ttcttctctg	gcttcttctg	tgtgtttctg	960
agcgtgctgc	catagcctct	catcaaaggc	atcatcctgc	ggaagatctt	caccaccatc	1020
agcttttgca	tcgtgcttag	aatggccgtg	acccggcagt	tcccatgggc	cgtgcaaact	1080
tggtatgatt	ccctgggcgc	catcaacaaa	atccaggatt	tcctgcagaa	gcaggaatac	1140
aagacactcg	aatataatct	cacaactact	gaggtgggtta	tggaagaacgt	gactgccttc	1200
tgggaggagg	ggttcggaga	gctttttgag	aaggcaaaac	agaataacaa	caaccgcaaa	1260
accagcaacg	gcgacgacag	cctgttcttc	tccaattttt	ctctcctggg	aacaccgcgc	1320
ctcaaagaca	tcaactttaa	gatcgagagg	ggacagctgc	tcgcagtcgc	cggatccaca	1380
ggcgcgggca	agacctctct	gctgatgggt	atcatgggag	aactggagcc	atccgagggc	1440
aagattaagc	acagtggaag	aatctccttt	tgtagccagt	tcagttggat	tatgcccggc	1500
actattaagg	agaatatcat	ttttgggggtg	agctatgatg	agtatcggta	tcggagcggt	1560
atcaaagcct	gtcagctgga	ggaggatatc	agcaaattcg	cagagaagga	taatatcgtg	1620
ctggggggagg	ggggaatcac	cctgagcgga	ggccagagag	ccagaatctc	actggcccgg	1680
gccgtctaca	aggacgccga	cctttacctt	ctggacagtc	cctttggata	tctggatgtg	1740
ctgactgaaa	aggagatctt	cgagtcttgt	gtgtgcaagc	tgatggctaa	taagaccggg	1800
atcctagtga	ccagtaagat	ggagcacctg	aagaaggcag	acaagatctt	gattctgcac	1860
gagggatcct	cttactttta	cggcaccttt	agcgagctgc	agaatctcca	gcccgatctt	1920
tcacttaagc	tgatgggctg	tgatagcttc	gaccagttct	ctgccgagcg	cagaaacagc	1980
atcctgacag	agacactgca	ccggttttca	ctggagggcg	acgcccctgt	cagctggacc	2040
gagacaaaaa	agcagtcttt	caagcagaca	ggcgagttcg	gcgagaagcg	caaaaacagc	2100
atcctgaatc	caatcaatc	tataaggaag	tttagcatcg	tcagaagac	acccctccag	2160
atgaacggca	tcgaagagga	cagtgcagag	cccctggagc	ggcgcctgag	cctcgtgcct	2220
gacagcgaac	agggcgaggc	catcctgcct	aggatcagcg	tgatttcaac	cgggccaaca	2280
ctgcaggcta	ggagaagaca	gtcagtgcct	aacctgatga	cacatagcgt	gaatcaggga	2340
cagaacatcc	atcgaaaaac	cacagcctct	actcgaaaag	tgtcactggc	tcctcagggt	2400
aatctgacag	agctggacat	ctatagcagg	aggctgagcc	aggagacagg	cctggagatc	2460
agtgaggaga	tcaacgaaga	ggacctgaag	gagtgttttt	tcgatgacat	ggagagtatc	2520
cccgccgtca	ccacctggaa	tacctacctc	cggtacatca	cagtgcacaa	gtccctcatc	2580
tttgtgctga	tttgggtgcct	cgtgatcttt	ctcgcagaag	tggccgcctc	cctggtgggtg	2640
ctgtggctgt	tggggaatac	tccactgcag	gacaaaaggca	attctacaca	cagcaggaat	2700
aattcctatg	ccgtgattat	caccagcaca	tcctcttact	acgtgtttta	catctacgtg	2760
ggagtggcag	atactctgct	tgcaatgggc	ttcttcaggg	ggctgcccct	ggtgcacaca	2820
ctgatcacag	tgtccaagat	cctccaccat	aaaatgtctc	acagcgtgct	gcaggcacc	2880
atgagcacc	tgaacacact	gaaggccggc	ggcatcctga	atcgcttttc	caaagacatc	2940
gccatcctcg	acgatctcct	gccactgacc	atcttcgatt	ttatccagct	gctgctgac	3000
gtgatcgggg	ccatcgccgt	ggtggccgtg	ctgcagccat	acattttcgt	ggctacagtg	3060
cccgtgatcg	ttgcctttat	catgctgaga	gcctacttcc	tcgagacttc	tcagcagctg	3120

aagcagctgg	agagcgaagg	gagaagcccc	atcttcactc	acctggtgac	aagcctgaag	3180
ggactctgga	ccctgagagc	cttcggccgg	cagccctatt	tcgagaccct	gtttcacaag	3240
gccctcaacc	tgcacacagc	caactggttt	ctctacctgt	ccaccctgag	gtggttccag	3300
atgaggattg	aaatgatctt	cgtgattttt	ttcatcgccg	tgacattcat	tagcattctg	3360
accaccggcg	agggggaggg	gagagtgggc	atcatcctga	cccttgccat	gaacattatg	3420
tccacactgc	agtgggccgt	gaatagttca	atcgacgtgg	acagtctgat	gagggtccgtg	3480
agccgggtgt	tcaagttcat	tgacatgccc	acagagggga	aaccaccaa	aagcaccaag	3540
ccctacaaga	acgggcagct	gtccaagggt	atgatcatcg	agaactctca	cgtgaagaag	3600
gacgacattt	ggcccagcgg	cggccagatg	acagtgaaag	atctgaccgc	caaatacacc	3660
gagggaggca	acgccatcct	cgaaaacatt	agcttctcta	tcagccctgg	acagagggtg	3720
ggcctgctgg	gccggacagg	ctcaggggaag	agtactctgc	tgtcagcatt	cctgaggctc	3780
ctgaacacag	agggcgagat	ccagattgac	ggcgtgtcct	gggactccat	caccctgcag	3840
cagtggcgga	aggctttcgg	ggtgatcccc	cagaaggtgt	tcattctttag	cggcactttc	3900
agaaagaatc	tggaccctta	tgagcagtgg	agtgaccagg	agatctggaa	agtggccgat	3960
gaggtcggac	tgaggagcgt	gatcgagcag	tttccaggga	agctggactt	tgtgctggtg	4020
gatggcggat	gcgtgctgtc	tcacggccat	aaacagctga	tgtgtctggc	ccggtccgtg	4080
ctgtctaagg	ccaagatcct	gctgctggac	gaaccctccg	cccacctgga	ccccgtgaca	4140
taccagatca	tcaggagaac	tctcaagcag	gccttcgccg	actgtaccgt	gattctgtgc	4200
gagcaccgca	ttgaagctat	gctggagtgt	cagcagttcc	tggtgatcga	ggaaaataag	4260
gtgaggcagt	acgacagcat	ccagaagctg	ctgaacgagc	gctccctgtt	ccgccaggct	4320
atctcccat	cagaccgggt	gaagctcttc	ccccacagaa	actcctcaaa	gtgcaagtcc	4380
aagccccaga	tcgccgccct	gaaggaggag	accgaggagg	aggtgcagga	caccaggctg	4440
tga						4443

<210> 29
 <211> 2100
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Homo sapiens DNAI1 sequence, codon optimized, DNAI1 #1

<400>	29					
atgatcccag	cttctgccaa	ggccccacac	aagcagccac	acaaacagag	catttccatt	60
gggcgcggca	caagggaagag	agacgaggac	tcaggcacag	aggtgggcga	aggaaccgac	120
gagtgggctc	agagcaaagc	cacagtgagg	ccccagatc	agctggagct	gacagacgcc	180
gagctgaagg	aggagtttac	ccgcatictg	actgccaata	accacacgc	accccagaac	240
atcgtgcgct	attcttttaa	ggaaggaacc	tataagccaa	tcggctttgt	caatcagctg	300
gctgtgcact	acaccaggt	tgggaacctg	atccccaagg	atagcgacga	gggcaggaga	360
cagcattata	gagacgagct	cgtcgccgga	agccaggagt	ctgtcaaagt	gatcagcgaa	420
acaggaaacc	tggaggagga	tgaggagccc	aaggaactgg	aaaccgagcc	tggcagccag	480
acagatgtgc	cagccgcagg	agccgcagag	aaggtgacag	aagaggagct	catgaccccc	540
aaacagccaa	aggagcggaa	actgacaaac	cagttcaact	tcagcgaaag	agccagccag	600
acctacaata	accccgtgcg	ggacagagaa	tgccagacag	agcctccacc	acgcaccaac	660
ttctccgcaa	cagctaacca	gtgggagatc	tatgatgcct	acgtggagga	gctggaaaag	720
caggagaaga	ccaaagaaaa	ggagaaagcc	aagacccttg	tcgccaagaa	gtccggcaaa	780
atggctatga	gaaagctgac	atctatggaa	tcccagactg	atgacctgat	caagctgtct	840
caggcagcca	agattatgga	aagaatggtg	aatcagaaca	cctatgacga	catcgcccag	900
gatttttaagt	actatgatga	cgctgcagac	gagtatagag	atcaggtggg	gaccctgctg	960
ccactgtgga	agttccagaa	tgacaaggct	aagcgcctgt	ccgtgacagc	tctgtgctgg	1020
aatccaaaat	atagggacct	cttcgccgtg	ggctacggct	cttatgactt	catgaagcag	1080
tcacgcggga	tgctgctgct	gtacagcctg	aaaaatccct	cctttcccga	gtacatgttc	1140
agctctaact	ccggggtcct	gtgtctggat	attcatgtgg	accatccata	cctggtggct	1200
gtcgggcact	acgatggaaa	cgtggctatc	tacaatctga	agaagccaca	ctcccagccc	1260

tccttttgct	cctccgcaa	gtccggcaag	cactccgacc	ctgtgtggca	ggtcaagtgg	1320
cagaaggacg	acatggacca	gaacctgaac	ttcttttctg	tgtctagcga	tggcaggatc	1380
gtgtcctgga	ccctgggtgaa	gagaaaactg	gtgcacatcg	atgttatcaa	gctcaaagtc	1440
gaggggaagca	ccaccgaggt	tcctgagggc	ctgcagctgc	accagtgagg	ctgcggcaca	1500
gccttcgact	ttcataaaga	gattgactac	atgttcctgg	tgggcacaga	ggaggggaag	1560
atctacaagt	gctccaaatc	ctactccagc	cagtttctgg	acacttacga	cgctcataat	1620
atgagcgtgg	acaccgtgtc	ctggaaccct	taccacacaa	aggtgttcat	gagctgcagc	1680
agcgactgga	ctgtgaagat	ttgggaccat	actatcaaaa	ccccaatgtt	tatctatgat	1740
ctcaattctg	ccgtgggcga	cgtggcttgg	gccccctatt	cctccacagt	gttcgcagcc	1800
gtgactaccg	acggaaaagc	ccacattttc	gacctcgcta	ttaacaagta	tgaggccatt	1860
tgttaaccagc	cagtggctgc	caagaagaac	cgctgaccc	acgtgcagtt	caacctgatt	1920
cacccaatta	tcattgtggg	ggacgacaga	ggacacatta	tctcactgaa	gctgtctcct	1980
aatctgagaa	agatgcctaa	ggagaagaaa	ggacaggagg	tgcagaaggg	ccctgccgtg	2040
gaaattgcca	aactcgacaa	gctgctgaac	ctggtgaggg	aggtgaagat	caagacatga	2100

<210> 30
 <211> 2100
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Homo sapiens DNAI1 sequence, codon optimized, DNAI1 #2

<400>	30	
atgatccccg	catccgcaa	agcccctcat
ggacggggga	cccggaaaag	ggatgaggac
gaatgggcac	agagtaaggc	taccgtgaga
gaactgaagg	aggagtttac	taggatcctg
atcgtcagat	atagcttcaa	agagggcaca
gctgtgcatt	acacacaggt	ggggaacctt
cagcattata	gagatgaact	ggttgacagga
accggcaacc	tggaagaaga	tgaggaacca
acagacgtgc	cagcagctgg	cgctgccgag
aaacagccaa	aagagaggaa	gctgacaaac
acctacaaca	accagtgctg	cgaccgggag
ttttctgcta	ctgcaaatca	gtgggagatc
caggaaaaga	ctaaggagaa	ggagaaggca
atggcaatgc	ggaagctgac	ttctatggaa
caggcagcta	agattatgga	aaggatggtc
gactttaagt	attatgatga	tgccgctgac
ccactgtgga	agtttcaaaa	cgacaaggct
aatcccaagt	accgggacct	ctttgccgtg
tccagaggca	tgctgctgct	gtacagcttg
agctctaatt	ctggagtgat	gtgcctggac
gttgacact	atgacggcaa	tgtggccatc
tccttttgta	gctctgcaaa	gtccggaaaag
cagaaagacg	acatggacca	gaatctgaac
gtctcatgga	ctctgggtcaa	acggaagctg
gaaggatcga	ctactgaggt	gccagaagga
gcatttgatt	tccataaaga	aatcgactac
atctataagt	gtagcaaatc	ctattctagc
atgtccgtgg	acactgtaag	ctggaacccc
tccgattgga	ctgttaagat	ttgggatcac
ctgaactctg	ccgtggggga	tgtggcctgg
aaacagcccc	acaaacagtc	catctccatt
tctgggacgg	aagttggaga	aggcactgac
cctcccgacc	agctggagct	cactgacgca
acagcaaata	acccccacgc	cccacagaat
tacaagccta	ttgggttcgt	gaaccagctg
attcctaaag	actctgatga	aggccgcaga
ctgtgaaagt	gatttagcgag	
aaagaactgg	agacagagcc	tgggtctcag
aaagtgcag	aggaggagct	gatgacacct
caattcaatt	tttccgaacg	ggcatcacag
aacctcctcc	tagaacaac	
acgtggagga	gctggagaag	
tggccaaaaa	atccggcaaa	
agccagactg	atgacctgat	caaactgtcc
catatgacga	cattgtctcag	
accaagtggg	gacactgctg	
ccgtgacagc	actctgctgg	
cttacgactt	catgaaacag	
cctttcccga	gtacatgttc	
atccacgtgg	atcaccctta	cctcgtggcc
tacaacctga	aaaaaccaca	ctctcagcct
ccgtgtggca	agtgaatgg	
tctcttcaga	cggcagaatc	
acgtgatcaa	actcaaggctc	
accagtgagg	atgtggaaact	
tgggaactga	agaggggaag	
atacatacga	cgctcacaac	
aggtgttcat	gtcctgcagc	
cccctatgtt	tatctacgat	
gctccacagt	cttcgcagct	

gtcactaccg	atggaaaggc	ccacattttt	gacctggcta	tcaacaaata	cgaggccatc	1860
tgcaatcagc	ctgtggcagc	aaagaagaac	cgcttgactc	acgtgcaatt	caacctgatt	1920
caccctatca	tcattgtttg	ggatgatagg	ggccacatta	tttctctaaa	gctgtcccca	1980
aatctgcgga	aaatgcccaa	ggagaagaaa	ggccaggagg	tgcagaaagg	cccagccgtt	2040
gaaatcgcaa	agctggacaa	gctgctcaac	ctcgtccggg	aggttaaaat	caaaacctga	2100

<210> 31
 <211> 2100
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Homo sapiens DNAI1 sequence, codon optimized, DNAI1 #3

<400>	31	
atgatcccag	caagcgccaa	ggccccacac
ggcaggggga	ctcgcaagag	agacgaggac
gaatggggcc	agtctaaggc	cactgttcgc
gagctgaagg	aagagttcac	caggattctg
attgtgagat	attcttttaa	ggagggcact
gcagtgcatt	acactcaagt	gggcaacctg
cagcactata	gggacgagct	ggtcgctggg
actggcaacc	tggaagagga	tgaggagcca
acagatgtgc	ccgccgctgg	agctgctgag
aaacagccta	aggaaacgga	gctcaccaac
acttataata	accctgtgcg	cgaccgggag
ttctccgcca	ctgccaaacca	gtgggaaatc
caggagaaaa	ctaaggagaa	agaaaaggcc
atggccatga	gaaaactgac	ctccatggag
caggcagcca	agatcatgga	gaggatggtg
gactttaaat	actacgatga	tgccgctgac
cctctgtgga	aattccagaa	tgataaggct
aaccctaagt	accgcgacct	ctttgctgtg
tccagaggca	tgctgctgct	gtacagcttg
agctctaatt	ctggagtgat	gtgcctggac
gttggaact	atgacggcaa	tgtggccatc
tccttttgta	gctctgcaaa	gtccggaaaag
cagaaagacg	acatggacca	gaatctgaac
gtctcatgga	ctctggtcaa	acggaagctg
gaaggatcga	ctactgaggt	gccagaagga
gcatttgatt	tccataaaga	aatcgactac
atctataagt	gtagcaaata	ctattctagc
atgtccgtgg	acactgtaag	ctggaacccc
tccgattgga	ctgttaagat	ttgggatcac
ctgaactctg	ccgtggggga	tgtggcctgg
gtcactaccg	atggaaaggc	ccacattttt
tgcaatcagc	ctgtggcagc	aaagaagaac
caccctatca	tcattgtttg	ggatgatagg
aatctgcgga	aaatgcccaa	ggagaagaaa
gaaatcgcaa	agctggacaa	gctgctcaac
		ctcgtccggg
		aggttaaaat
		caaaacctga
		2100

<210> 32
 <211> 2100
 <212> DNA

<213> Artificial Sequence

<220>

<223> Homo sapiens DNAI1 sequence, codon optimized, DNAI1 #4

<400> 32

atgatccccg	cctccgccaa	agccccctcac	aagcaaccgc	acaagcaaag	cattagcatt	60
gggcggggta	ctcggaagcg	cgacgaggac	tcgggaactg	aagtcggaga	ggggaccgac	120
gaatggggcg	agtcaaagcg	caccgtgctg	ccaccggacc	agctcgagct	gaccgatgct	180
gagctgaagg	aggagtttac	ccggatcctg	acagccaaca	acccacatgc	accgcagaac	240
atcgtgcggt	acagcttcaa	agagggaact	tataagccca	ttggcttcgt	gaaccaactc	300
gcggtgcatt	acaccaagt	cggaaacctt	attccgaagg	actcggacga	aggcagacgc	360
cagcactacc	gggacgagct	cgtggcagga	tcccaggaag	gcgtcaaggt	catttccgag	420
actggcaacc	tcgaggagga	cgaagaacct	aaggagctgg	aaaccgaacc	cggatcccag	480
accgacgtgc	cggccgctgg	ggctgccgag	aaagtcactg	aagaggaact	catgaccccc	540
aagcagccga	aagagagaaa	gctcaccaac	caattcaact	tcagcgagcg	cggcagccaa	600
acctacaaca	accagtcag	ggatcgggaa	tgtcagaccg	aaccgcctcc	gagaacgaac	660
ttctcggcga	ccgcgaacca	atgggagatc	tacgacgcct	acgtggaaga	actggaaaag	720
caggaanaaa	ctaaggaaaa	ggaaaaggcc	aagactcccg	tcgccaagaa	gtcgggcaaa	780
atggccatgc	ggaagctcac	ctccatggaa	tcacagactg	acgacttgat	caagttgagc	840
caggccgcaa	agatcatgga	gcgcatggtc	aaccaaata	cttacgacga	tatcgcccaa	900
gacttcaagt	actacgacga	cgctgccgat	gaataccgag	atcaagtcgg	caccctactg	960
ccgctttgga	agttccagaa	tgacaaggcc	aagaggctga	gcgtgaccgc	gctgtgctgg	1020
aaccccaaat	accgcgacct	cttcgccgtg	ggatacggct	cctacgattt	catgaagcag	1080
agccggggaa	tgttgctcct	ttactccctg	aagaaccctt	ccttccctga	gtacatgttc	1140
agctcaaaca	gcggcgtgat	gtgcctcgac	attcacgtgg	accaccctta	cctcgtggcc	1200
gtgggtcact	acgacggcaa	cgtcgcgatc	tacaacttga	agaagccgca	ttcacagccc	1260
tcgttttgct	cctcggccaa	gtccggcaaa	cattcggacc	cagtgtggca	agtcaagtgg	1320
cagaaagatg	acatggacca	aaacttgaac	ttcttcagcg	tgtcctccga	cggacggatc	1380
gtgtcctgga	ccctcgtgaa	gcggaagttg	gtgcatatcg	acgtgatcaa	attgaaggtc	1440
gagggttcga	ccaccgaagt	gcctgaaggc	ctgcagcttc	accccgctgg	atgcggcact	1500
gccttcgact	tccacaagga	gatcgactac	atgttcctcg	tgggaaccga	ggaagggaag	1560
atctacaaat	gcagcaagtc	ctactcatca	caattcctgg	atacctacga	tgccacaac	1620
atgagcgtgg	ataccgtgtc	gtggaacccc	tatcacacca	aggtattcat	gtcctgctcc	1680
tccgactgga	ccgtcaagat	ttgggaccac	accatcaaga	cccccatgtt	catctacgac	1740
ctgaactccg	ccgtggggga	tgtggcctgg	gccccctact	cgtcgaccgt	gtttgccgcg	1800
gtcaccacgg	acggaaagcg	acacattttc	gaccttgcca	ttaacaaata	cgaggcgatt	1860
tgcaaccagc	ccgtggccgc	caaaaagaac	cgcctgacct	acgttcaatt	caacttaatc	1920
cacccaatca	tcacgtcggg	cgatgacaga	ggacacatta	ttagcctgaa	acttagcccc	1980
aacctccgca	agatgcccaa	ggagaagaag	ggacaggaag	tccagaaggg	ccctgccgtg	2040
gagattgcaa	agctcgataa	gctcctgaac	ttagtcgggg	aagtgaagat	caagacttaa	2100