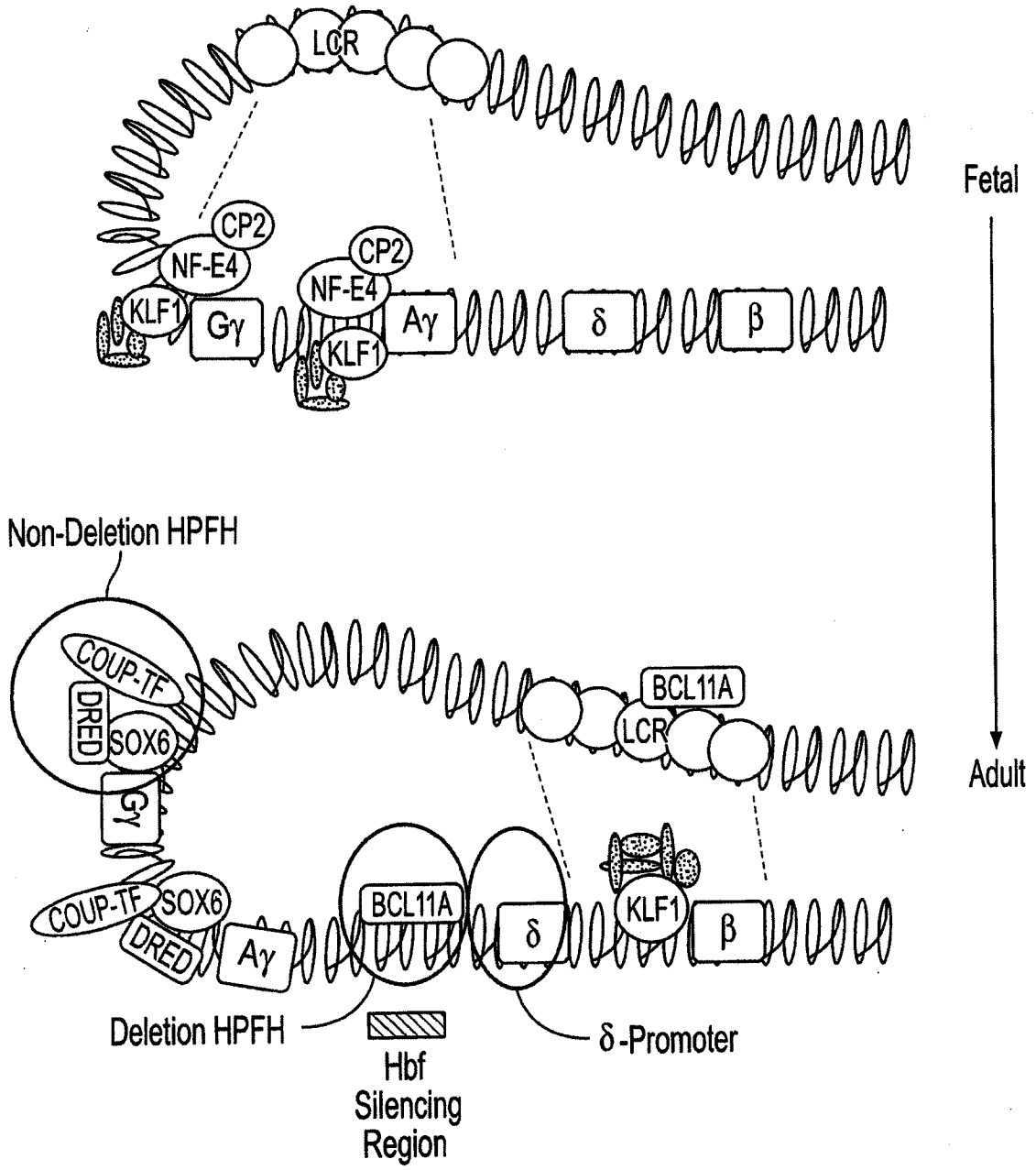

LCR
CP2
NF-E4
KLF1
Gγ
CP2
NF-E4
KLF1
Aγ
δ
β
Fetal
Non-Deletion HPFH
COUP-TF
DRED
SOX6
Gγ
BCL11A
LCR
Adult
COUP-TF
SOX6
DRED
Aγ
BCL11A
δ
KLF1
β
Deletion HPFH
Hbf
Silencing
Region
δ-Promoter

FIG. 1

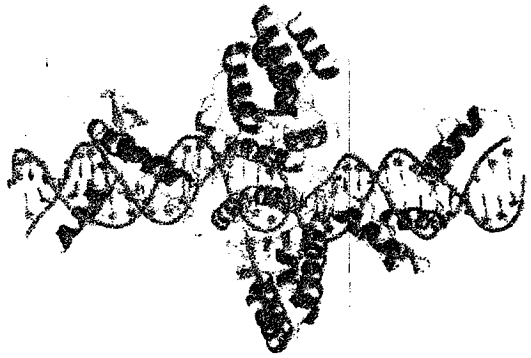
"TALEN"

TAL Effector Nuclease
Two chains
5400 bp total CDS
5' overhang
Easy retargeting
Reasonably high specificity
Difficult vectorization/delivery



"ZFN"

Zinc Finger Nuclease
Two chains
2400 bp total CDS
5' overhang
Moderate retargeting
Reasonably high specificity
Moderate vectorization/delivery



"HE"

Homing Endonuclease
Single chain
800 bp CDS
3' overhang
Difficult retargeting
High specificity
Easiest vectorization/delivery

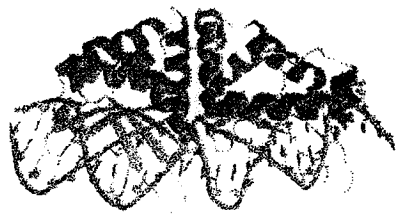


FIG. 2

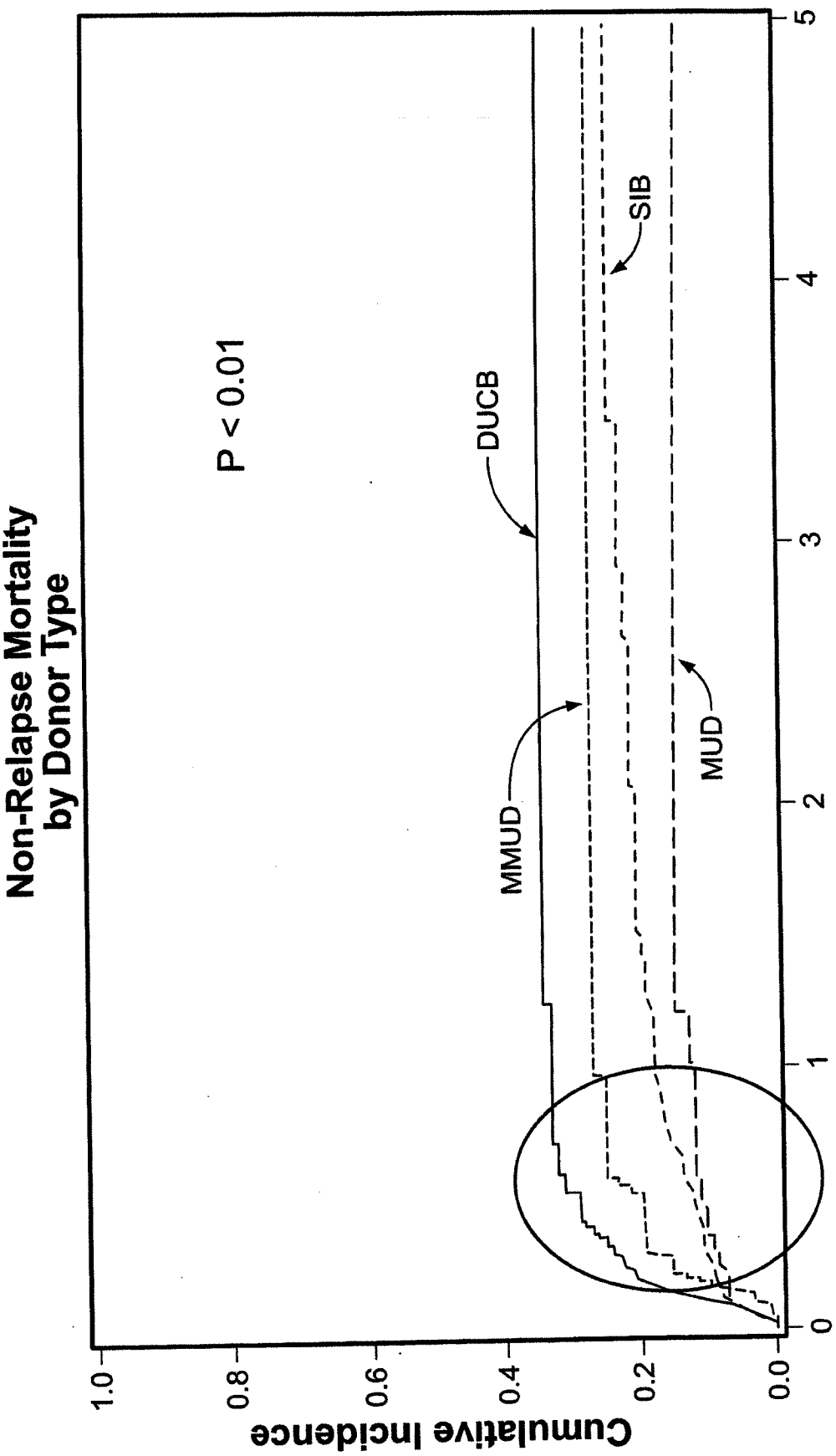

Non-Relapse Mortality
by Donor Type
P < 0.01
DUCB
SIB
MMUD
MUD
Cumulative Incidence
0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3
4
5
Years Post-Transplantation

FIG. 3

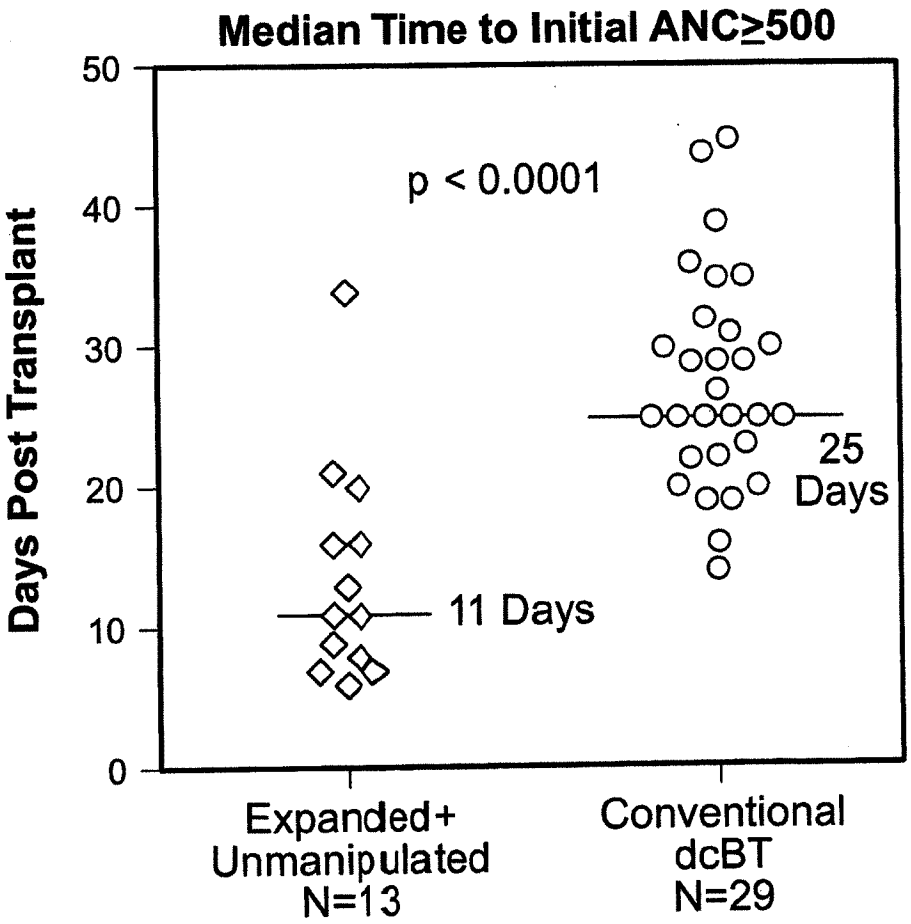

Median Time to Initial ANC≥500
Days Post Transplant
50
40
30
20
10
0
p < 0.0001
11 Days
25
Days
Expanded+
Unmanipulated
N=13
Conventional
dcBT
N=29

FIG. 4

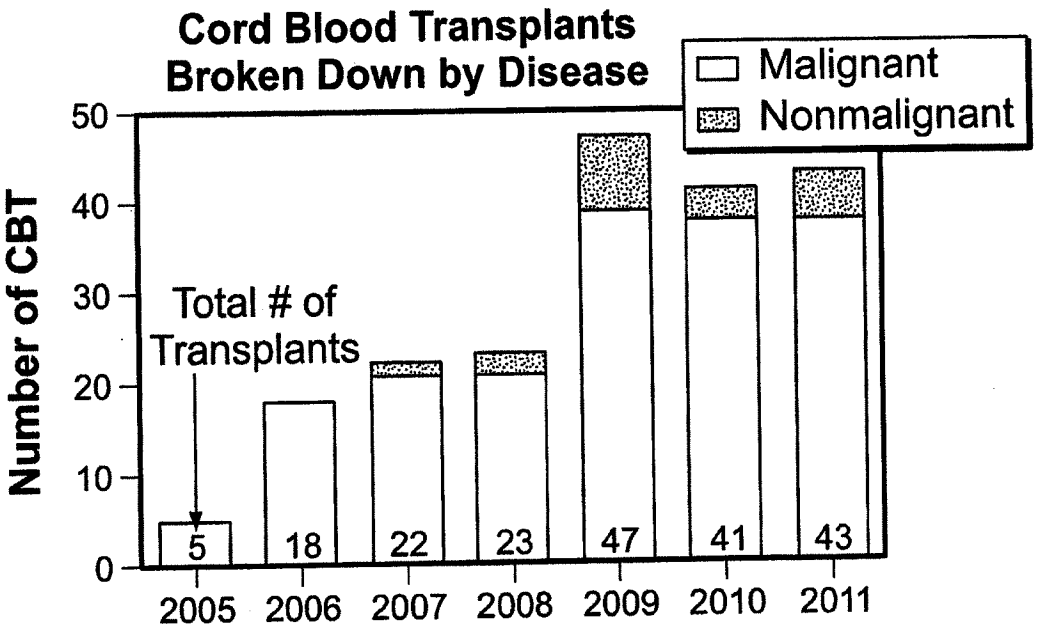

Cord Blood Transplants
Broken Down by Disease
Malignant
Nonmalignant
Number of CBT
50
40
30
20
10
0
Total # of
Transplants
5
18
22
23
47
41
43
2005
2006
2007
2008
2009
2010
2011

FIG. 5

(SEQ ID NO: 1)

Nucleotide Sequence of a Human Fetal Hemoglobin (HbF) Silencing Region

(chr11:5212342-5215944 in HG18)

```
   1  CCAGTGAGCAGGTTGGTTTAAGATAAGCAGGGTTTCATTAGTTTGTGAGA   50
  51  ATGAAAAATGAACCTTCATTCCACTATTCCCTTAACTTGCCCTGAGATTG  100
 101  GCTGTTCTGTCATGTGTGTCTTGACTCAGAAACCCTGTTCTCCTCTACAT  150
 151  ATCTCCCCACCGCATCTCTTTCAGCAGTTGTTTCTAAAAATATCCTCCTA  200
 201  GTTTCATTTTTGCAGAAGTGTTTTAGGCTAATATAGTGGAATGTATCTTA  250
 251  GAGTTTAACTTATTTGTTTCTGTCACTTTATACTAAGAAAACTTATCTAA  300
 301  AAGCAGATGTTTTAACAAGTTGACTCAATATAAAGTTCTTCTTTGCCTCT  350
 351  AGAGATTTTTGTCTCCAAGGGAATTTTGAGAGGTTGGAATGGACAAATCT  400
 401  ATTGCTGCAGTTTAAACTTGCTTGCTTCCTCCTTCTTTTGGTAAATTCTT  450
 451  CCTATAATAAAACTCtaatttttattatattgaaataaatatccattaa  500
 501  aagaatatttaaaaaatGAATAGTGTTTATTTACCAGTTATTGAAATAGG  550
 551  TTCTGGAAACATGAATTTTAAGGTTAACATTTTAATGACAGATAAAATCA  600
 601  AATATTATATACAAATATTTTGAATGTTTAAAATTATGGTATGACTAAAG  650
 651  AAAGAATGCAAAGTGAAAAGTAGATTTACCATATTCAGCCAGATTAAATT  700
 701  TAACGAAGTTCCTGGGAATATGCTAGTACAGAACATTTTTACAGATGTGT  750
 751  TCTTAAAAAAAAATGTGGAATTAGACCCAGGAATGAAGATCCCAGTAGTT  800
 801  TTTCACTCTTTTCTGAATTCAAATAATGCCACAATGGCAGACAAATACAC  850
 851  ACCCATGAGCATATCCAAaaggaaggattgaaggaaagaggaggaagaaa  900
 901  tggagaaaggaaggaaggaagaggggaagagagaggatggaagggatgga  950
 951  ggagaagaaggaaAAATAAATAATGgagaggagaggagaaaaaaggaggg 1000
1001  gagaggagaggagaagggatagggaagagaaagagaaagggaagggaaga 1050
1051  gaggaaagaagagaagaggagagaaaagaaacgaagagaggggaagggaa 1100
1101  ggaaaaaaaagaggaaaaaagagacaagagaagagaTaagactgacagtt 1150
1151  caaattttggtggtgatatggatcaatagaaactcaaactctgttggtga 1200
1201  cactgtacaatagtataacccctttggaaaacctttaatagtatccacaa 1250
1251  atgctggatgcttgataagtctattacctagcaattacatttttagatat 1300
```

FIG. 6

1301 tcagaaacacatgcatgtgtgtatccaaagacatgtatagaaatgcttat 1350
1351 gacagcaataatcataaaaacctcaaaccggtagccacttaaatgcttac 1400
1401 caacagtagaattgataaattacggtatagtcaaagaatagaatattaca 1450
1451 cagaaatgaaaagaatcaactactgcttaacacgtagcgatacaaatgca 1500
1501 ttttacagcatttggttgattaaaagtaaccagaggtgagttcaaactat 1550
1551 atgactttatttgtatatagAAAGATGGATGATGTGCCTGAGATTCTGAT 1600
1601 CACAAGGGGAAATGTTATAAAATAGGGTAGAGAGGAGCCATGAATGACCT 1650
1651 TTAAACTTTGTTACAAGTTATTTTTCTGTAACCTggaagccaacgaaaga 1700
1701 tattgaataattcaagaaaggtggtggcatggtttgatttgtgtctttaa 1750
1751 aagattattctcacttagtgaagaaatgtattttagaagtagagaaaatg 1800
1801 ggagacaaatagctgggcttctgttgcagtaGGGAAGAAAGTGACAATGC 1850
1851 CATTTCTATTATCAGACTTGGACCATGACGGTGATGTCAGTCGTGAACAC 1900
1901 AAGAATAGGGCCACATTTGTGAGTTTAGTGGTACGATaaaatcagaaata 1950
1951 cagtcttggatacattgtattgtatgcactcttgtaaaatgcaaaaagat 2000
2001 gtacttagatatgtggatctggagctcagAAAGAATACAACCAGGTCAAG 2050
2051 AATACAGAATGGAACAGAACATACAAGAACAGATCATAATGTGCTGTGTG 2100
2101 AATCACTACCACTACCTGTTAAAAATGACAGATGATGTACTTCATcaata 2150
2151 tctccttaaaatcttagaatgtgtttgtgagggaggaattatgtttccaa 2200
2201 ttcatatataagaaaattgattctaaaaaaaatgttaggtaaattcttaa 2250
2251 ggccaTGAGGACTGTTATTTGATCTTTGTCTGTTAATTCCAAAGACTTGG 2300
2301 CTTTTCACTTTAATTCTGTTCTACCTGAAATGATTTTACACATTGGGAGA 2350
2351 TCTGGTTACATGTTTATTCTATATGGATTGCATTGAGAGGATTTGTATAA 2400
2401 CAGAATAAGGTCttttttttctttctcttctgagatggagtttcatccct 2450
2451 attgcccaagctagagtgcaatggtgcaatctaggctcaccgcaacctct 2500
2501 gcctcctgggttcaagcaattctcctgcctcagccacctgaatagctggg 2550
2551 actgcaggcatgcaccacacgcccggctgattttgtatttttagtagaga 2600
2601 tggggtttcaccatgttggtcaggctggtcttgaactcctgacctcaagt 2650
2651 gatctgcctgccttggcctcccaaagtgctgggtttacaagcctgagcca 2700
2701 ccgcatccagccAGGATAAGGTCTAAAAGTGGAAAGAATAGCATCTACTC 2750

FIG. 6 (Cont.)

6

2751	TTGTTCAGGAAACAATGAGGACCTGACTGGGCAGTAAGAGTGGTGATTAA	2800
2801	TAGATAGGGACAAATTGAAGCAGAATCGAACTGTTGATTAgaggtaggga	2850
2851	aatgattttaatctgtgaccttggtgaatgggcaagtagctatctaatga	2900
2901	ctaaaatggaaaacactggaagagaaacagttttagtataacaagtgaaa	2950
2951	tacccatgctgagtctgaggtgcctataggacatctatataaataagccc	3000
3001	agtacattgtttgatatatgggtttggcactgaggttggaggtcagaggt	3050
3051	tagaaatcagagttgggaattgggattatacaggctgtatttaagagttt	3100
3101	agatataactgtgaatccaagagtgtgaTGAATACAAAGTTAAATGAAGG	3150
3151	ACCTTTAATGAACACCAACATTTAATGTGAAATCTCAAGGAAGTATGAAG	3200
3201	TAAGACATAGTCCCCAAAATCCCCGATGATTTTAGAACTCAGTATCGATT	3250
3251	TTAATTAGTGTAATGCCAATGTGGGTTAGAATGGAAGTCAACTTGCTGTT	3300
3301	GGTTTCAGAGCAGGTAGGAGATAAGGTTCTAGATTTTGACACAGTGAAAA	3350
3351	GCTGAAACAAAAAGGAAAAGGTAGGGTGAAAGATGGGAAATGTATGTAAG	3400
3401	GAGGATGAGCCACATGGTATGGGAGGTATACTAAGGACTCTAGGGTCAGA	3450
3451	GAAATATGGGTTATATCCTTCTACAAAATTCACATTCTTggctgggtgtg	3500
3501	gtggctcacgcctgtgatcccagcactttcagaggccgaggagggtggat	3550
3551	cacctgatgttaggagttcgagatcagcctgaccaacatggtgaaacccc	3600
3601	cta	3603

FIG. 6 (Cont.)

(SEQ ID NO: 2)

Nucleotide Sequence of a 350 bp Portion of a Human HbF Silencing Region (Including Sites for Bcl11a Occupancy)

1	AAAGATGGATGATGTGCCTGAGATTCTGATCACAAGGGGAAATGTTATAA	50
51	AATAGGGTAGAGAGGAGCCATGAATGACCTTTAAACTTTGTTACAAGTTA	100
101	TTTTTCTGTAACCTggaagccaacgaaagatattgaataattcaagaaag	150
151	gtggtggcatggtttgatttgtgtctttaaaagattattctcacttagtg	200
201	aagaaatgtattttagaagtagagaaaatgggagacaaatagctgggctt	250
251	ctgttgcagtaGGGAAGAAAGTGACAATGCCATTTCTATTATCAGACTTG	300
301	GACCATGACGGTGATGTCAGTCGTGAACACAAGAATAGGGCCACATTTGT	350

FIG. 7

(SEQ ID NO: 13)

Human b-globin Gene from 1 kb Upstream of the Cap Site through the PolyA Site

1	GCAATGAAAATAAATGTTTTTTATTAGGCAGAATCCAGATGCTCAAGGCC	50
51	CTTCATAATATCCCCCAGTTTAGTAGTTGGACTTAGGGAACAAAGGAACC	100
101	TTTAATAGAAATTGGACAGCAAGAAAGCGAGCTTAGTGATACTTGTGGGC	150
151	CAGGGCATTAGCCACACCAGCCACCACTTTCTGATAGGCAGCCTGCACTG	200
201	GTGGGGTGAATTCTTTGCCAAAGTGATGGGCCAGCACACAGACCAGCACG	250
251	TTGCCCAGGAGCTGTGGGAGGAAGATAAGAGGTATGAACATGATTAGCAA	300
301	AAGGGCCTAGCTTGGACTCAGAATAATCCAGCCTTATCCCAACCATAAAA	350
351	TAAAAGCAGAATGGTAGCTGGATTGTAGCTGCTATTAGCAATATGAAACC	400
401	TCTTACATCAGTTACAATTTATATGCAGAAATATTTATATGCAGAGATAT	450
451	TGCTATTGCCTTAACCCAGAAATTATCACTGTTATTCTTTAGAATGGTGC	500
501	AAAGAGGCATGATACATTGTATCATTATTGCCCTGAAAGAAAGAGATTAG	550
551	GGAAAGTATTAGAAATAAGATAAACAAAAAAGTATATTAAAAGAAGAAAG	600
601	CATTTTTTAAAATTACAAATGCAAaattaccctgatttggtcaatatgtg	650
651	tacacatattaaaacattacactttaacccataaatatgtataatgatta	700
701	tgtatcaattaaaaataaaaGAAAATAAAGTAGGGAGATTATGAATATGC	750
751	AAATAAGCACACATATATTCCAAATAGTAATGTACTAGGCAGACTGTGTA	800
801	AAGTTTTTTTTTTAAGTTACTTAATGTATCTCAGAGATATTTCCTTTTGTT	850
851	ATACACAATGTTAAGGCATTAAGTATAATAGTAAAAATTGCGgagaagaa	900
901	aaaaaaagaaagcaagaattaaacaaaagaaaacaattgttatgaacagc	950
951	aaataaaagaaactaaaaCGATCCTGAGACTTCCACACTGATGCAATCAT	1000
1001	TCGTCTGTTTCCCATTCTAAACTGTACCCTGTTACTTATCCCCTTCCTAT	1050
1051	GACATGAACTTAACCATAGAAAAGAAGGGGAAAGAAAACATCAAGCGTCC	1100
1101	CATAGACTCACCCTGAAGTTCTCAGGATCCACGTGCAGCTTGTCACAGTG	1150
1151	CAGCTCACTCAGTGTGGCAAAGGTGCCCTTGAGGTTGTCCAGGTGAGCCA	1200
1201	GGCCATCACTAAAGGCACCGAGCACTTTCTTGCCATGAGCCTTCACCTTA	1250
1251	GGGTTGCCCATAACAGCATCAGGAGTGGACAGATCCCCAAAGGACTCAAA	1300

FIG. 8

1301	GAACCTCTGGGTCCAAGGGTAGACCACCAGCAGCCTAAGGGTGGGAAAAT	1350
1351	AGACCAATAGGCAGAGAGAGTCAGTGCCTATCAGAAACCCAAGAGTCTTC	1400
1401	TCTGTCTCCACATGCCCAGTTTCTATTGGTCTCCTTAAACCTGTCTTGTA	1450
1451	ACCTTGATACCAACCTGCCCAGGGCCTCACCACCAACTTCATCCACGTTC	1500
1501	ACCTTGCCCCACAGGGCAGTAACGGCAGACTTCTCCTCAGGAGTCAGATG	1550
1551	CACCATGGTGTCTGTTTGAGGTTGCTAGTGAACACAGTTGTGTCAGAAGC	1600
1601	AAATGTAAGCAATAGATGGCTCTGCCCTGACTTTTATGCCCAGCCCTGGC	1650
1651	TCCTGCCCTCCCTGCTCCTGGGAGTAGATTGGCCAACCCTAGGGTGTGGC	1700
1701	TCCACAGGGTGAGGTCTAAGTGATGACAGCCGTACCTGTCCTTGGCTCTT	1750
1751	CTGGCACTGGCTTAGGAGTTGGACTTCAAACCCTCAGCCCTCCCTCTAAG	1800
1801	ATATATCTCTTGGCCCCATACCATCAGTACAAATTGCTACTAAAAACATC	1850
1851	CTCCTTTGCAAGTGTATTTACGTAATATTTGGAAtcacagcttggtaagc	1900
1901	atattgaagatcgttttcccaattttcttattacacaaataagaagttga	1950
1951	tgcactaaaagtggaagagttttgtctaCCATAATTCAGCTTTGGGATAT	2000
2001	GTAGATGGATCTCTTCCTGCGTCTCCAGAATATGCAAAATACTTACAGGA	2050
2051	CAGAATGGATGAAAACTCTACCTCGGTTCTAAGCATATCTTCTCCTTATT	2100
2101	TGGATTAAAACCTTCTGGTAAGAAAAGAAAAAatatatatatatatgtgt	2150
2151	gtatatatacacacatacatatacatatatatGCATTCATTTGTTGTTGT	2200
2201	TTTTCTTAATTTGCTCATGCATGCTAATAAATTATGTCTAAAAATAGAAT	2250
2251	AAATACAAATCAATGTGCTCTGTGCATTAGTTACTTATTAGGTTTTGGGA	2300
2301	AACAAGAGATAAAAAACTAGAGACCTCTTAATGCAGTCAAAAATACAAAT	2350
2351	AAATAAAAAGTCACTTACAACCCAAAGTGTGACTATCAATGGGGTAATCA	2400
2401	GTGGTGTCAAATAGGAGGTTAACTGGGGACATCTAACTGTTTCTGCCTGG	2450
2451	ACTAATCTGCAAGAGTGTCTGGGGGAACAAAAAGCCTCTGTGACTTAGAA	2500
2501	AGTAGGGGTAGGAGGGGAAAAGGTCTTCTACTTGGCTCAGATTATTTTTT	2550
2551	TCCTCTAGTCCACTAAGAATACTGCGTTTTAAAATCATTTCCTTGATTCA	2600
2601	AGTTCC	2606

FIG. 8 (Cont.)

(SEQ ID NO: 14)

606 bp Region of Human b-globin Spanning from the Promoter into Intron 2
(Mutations within this Region Result in Thalassemias and Sickle Cell Disease)

```
  1  CACCCTGAAGTTCTCAGGATCCACGTGCAGCTTGTCACAGTGCAGCTCAC  50
 51  TCAGTGTGGCAAAGGTGCCCTTGAGGTTGTCCAGGTGAGCCAGGCCATCA 100
101  CTAAAGGCACCGAGCACTTTCTTGCCATGAGCCTTCACCTTAGGGTTGCC 150
151  CATAACAGCATCAGGAGTGGACAGATCCCCAAAGGACTCAAAGAACCTCT 200
201  GGGTCCAAGGGTAGACCACCAGCAGCCTAAGGGTGGGAAAATAGACCAAT 250
251  AGGCAGAGAGAGTCAGTGCCTATCAGAAACCCAAGAGTCTTCTCTGTCTC 300
301  CACATGCCCAGTTTCTATTGGTCTCCTTAAACCTGTCTTGTAACCTTGAT 350
351  ACCAACCTGCCCAGGGCCTCACCACCAACTTCATCCACGTTCACCTTGCC 400
401  CCACAGGGCAGTAACGGCAGACTTCTCCTCAGGAGTCAGATGCACCATGG 450
451  TGTCTGTTTGAGGTTGCTAGTGAACACAGTTGTGTCAGAAGCAAATGTAA 500
501  GCAATAGATGGCTCTGCCCTGACTTTTATGCCCAGCCCTGGCTCCTGCCC 550
551  TCCCTGCTCCTGGGAGTAGATTGGCCAACCCTAGGGTGTGGCTCCACAGG 600
601  GTGAGG                                              606
```

FIG. 9

(SEQ ID NO: 24)

Nucleotide Sequence of Human Bcl11a cDNA (CCDS1862.1)

1	ATGTCTCGCCGCAAGCAAGGCAAACCCCAGCACTTAAGCAAACGGGAATT	50
51	CTCGCCCGAGCCTCTTGAAGCCATTCTTACAGATGATGAACCAGACCACG	100
101	GCCCGTTGGGAGCTCCAGAAGGGGATCATGACCTCCTCACCTGTGGGCAG	150
151	TGCCAGATGAACTTCCCATTGGGGGACATTCTTATTTTTATCGAGCACAA	200
201	ACGGAAACAATGCAATGGCAGCCTCTGCTTAGAAAAAGCTGTGGATAAGC	250
251	CACCTTCCCCTTCACCAATCGAGATGAAAAAAGCATCCAATCCCGTGGAG	300
301	GTTGGCATCCAGGTCACGCCAGAGGATGACGATTGTTTATCAACGTCATC	350
351	TAGAGGAATTTGCCCCAAACAGGAACACATAGCAGATAAACTTCTGCACT	400
401	GGAGGGGCCTCTCCTCCCCTCGTTCTGCACATGGAGCTCTAATCCCCACG	450
451	CCTGGGATGAGTGCAGAATATGCCCCGCAGGGTATTTGTAAAGATGAGCC	500
501	CAGCAGCTACACATGTACAACTTGCAAACAGCCATTCACCAGTGCATGGT	550
551	TTCTCTTGCAACACGCACAGAACACTCATGGATTAAGAATCTACTTAGAA	600
601	AGCGAACACGGAAGTCCCCTGACCCCGCGGGTTGGTATCCCTTCAGGACT	650
651	AGGTGCAGAATGTCCTTCCCAGCCACCTCTCCATGGGATTCATATTGCAG	700
701	ACAATAACCCCTTTAACCTGCTAAGAATACCAGGATCAGTATCGAGAGAG	750
751	GCTTCCGGCCTGGCAGAAGGGCGCTTTCCACCCACTCCCCCCCTGTTTAG	800
801	TCCACCACCGAGACATCACTTGGACCCCCACCGCATAGAGCGCCTGGGGG	850
851	CGGAAGAGATGGCCCTGGCCACCCATCACCCGAGTGCCTTTGACAGGGTG	900
901	CTGCGGTTGAATCCAATGGCTATGGAGCCTCCCGCCATGGATTTCTCTAG	950
951	GAGACTTAGAGAGCTGGCAGGGAACACGTCTAGCCCACCGCTGTCCCCAG	1000
1001	GCCGGCCCAGCCCTATGCAAAGGTTACTGCAACCATTCCAGCCAGGTAGC	1050
1051	AAGCCGCCCTTCCTGGCGACGCCCCCCCTCCCTCCTCTGCAATCCGCCCC	1100
1101	TCCTCCCTCCCAGCCCCCGGTCAAGTCCAAGTCATGCGAGTTCTGCGGCA	1150
1151	AGACGTTCAAATTTCAGAGCAACCTGGTGGTGCACCGGCGCAGCCACACG	1200
1201	GGCGAGAAGCCCTACAAGTGCAACCTGTGCGACCACGCGTGCACCCAGGC	1250
1251	CAGCAAGCTGAAGCGCCACATGAAGACGCACATGCACAAATCGTCCCCCA	1300
1301	TGACGGTCAAGTCCGACGACGGTCTCTCCACCGCCAGCTCCCCGGAACCC	1350

FIG. 10

1351	GGCACCAGCGACTTGGTGGGCAGCGCCAGCAGCGCGCTCAAGTCCGTGGT	1400
1401	GGCCAAGTTCAAGAGCGAGAACGACCCCAACCTGATCCCGGAGAACGGGG	1450
1451	ACGAGGAGGAAGAGGAGGACGACGAGGAAGAGGAAGAAGAGGAGGAAGAG	1500
1501	GAGGAGGAGGAGCTGACGGAGAGCGAGAGGGTGGACTACGGCTTCGGGCT	1550
1551	GAGCCTGGAGGCGGCGCGCCACCACGAGAACAGCTCGCGGGGCGCGGTCG	1600
1601	TGGGCGTGGGCGACGAGAGCCGCGCCCTGCCCGACGTCATGCAGGGCATG	1650
1651	GTGCTCAGCTCCATGCAGCACTTCAGCGAGGCCTTCCACCAGGTCCTGGG	1700
1701	CGAGAAGCATAAGCGCGGCCACCTGGCCGAGGCCGAGGGCCACAGGGACA	1750
1751	CTTGCGACGAAGACTCGGTGGCCGGCGAGTCGGACCGCATAGACGATGGC	1800
1801	ACTGTTAATGGCCGCGGCTGCTCCCCGGGCGAGTCGGCCTCGGGGGGCCT	1850
1851	GTCCAAAAAGCTGCTGCTGGGCAGCCCCAGCTCGCTGAGCCCCTTCTCTA	1900
1901	AGCGCATCAAGCTCGAGAAGGAGTTCGACCTGCCCCCGGCCGCGATGCCC	1950
1951	AACACGGAGAACGTGTACTCGCAGTGGCTCGCCGGCTACGCGGCCTCCAG	2000
2001	GCAGCTCAAAGATCCCTTCCTTAGCTTCGGAGACTCCAGACAATCGCCTT	2050
2051	TTGCCTCCTCGTCGGAGCACTCCTCGGAGAACGGGAGTTTGCGCTTCTCC	2100
2101	ACACCGCCCGGGGAGCTGGACGGAGGGATCTCGGGGCGCAGCGGCACGGG	2150
2151	AAGTGGAGGGAGCACGCCCCATATTAGTGGTCCGGGCCCGGGCAGGCCCA	2200
2201	GCTCAAAAGAGGGCAGACGCAGCGACACTTGTGAGTACTGTGGGAAAGTC	2250
2251	TTCAAGAACTGTAGCAATCTCACTGTCCACAGGAGAAGCCACACGGGCGA	2300
2301	AAGGCCTTATAAATGCGAGCTGTGCAACTATGCCTGTGCCCAGAGTAGCA	2350
2351	AGCTCACCAGGCACATGAAAACGCATGGCCAGGTGGGGAAGGACGTTTAC	2400
2401	AAATGTGAAATTTGTAAGATGCCTTTTAGCGTGTACAGTACCCTGGAGAA	2450
2451	ACACATGAAAAAATGGCACAGTGATCGAGTGTTGAATAATGATATAAAAA	2500
2501	CTGAATAG	2508

FIG. 10 (Cont.)

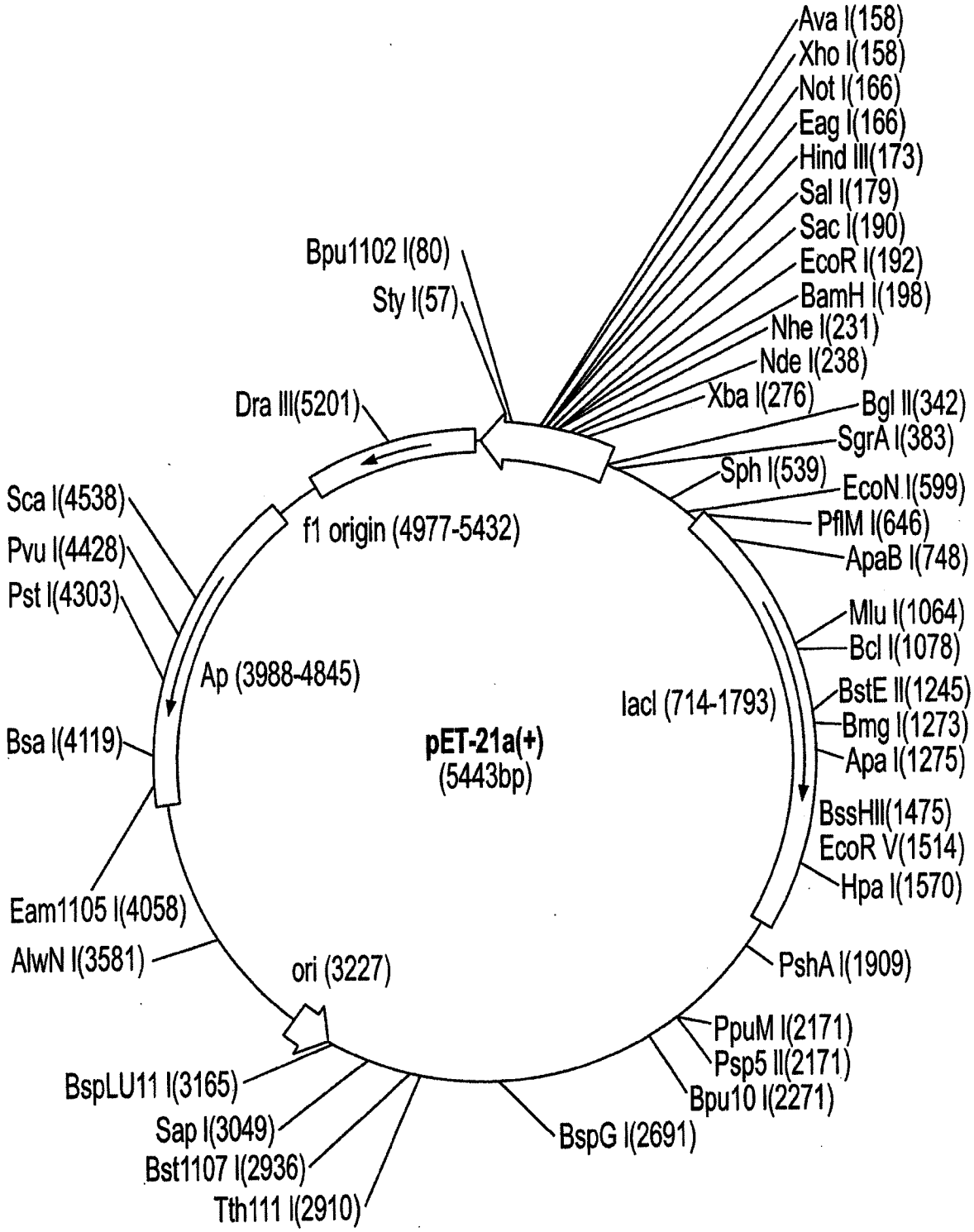

Ava I(158)
Xho I(158)
Not I(166)
Eag I(166)
Hind III(173)
Sal I(179)
Sac I(190)
EcoR I(192)
BamH I(198)
Nhe I(231)
Nde I(238)
Xba I(276)
Bgl II(342)
SgrA I(383)
Sph I(539)
EcoN I(599)
PflM I(646)
ApaB I(748)
Mlu I(1064)
Bcl I(1078)
BstE II(1245)
Bmg I(1273)
Apa I(1275)
BssHII(1475)
EcoR V(1514)
Hpa I(1570)
PshA I(1909)
PpuM I(2171)
Psp5 II(2171)
Bpu10 I(2271)
BspG I(2691)
Tth111 I(2910)
Bst1107 I(2936)
Sap I(3049)
BspLU11 I(3165)
ori (3227)
AlwN I(3581)
Eam1105 I(4058)
Bsa I(4119)
Pst I(4303)
Pvu I(4428)
Sca I(4538)
Dra III(5201)
Bpu1102 I(80)
Sty I(57)
f1 origin (4977-5432)
Ap (3988-4845)
lacI (714-1793)
pET-21a(+)
(5443bp)

FIG. 11

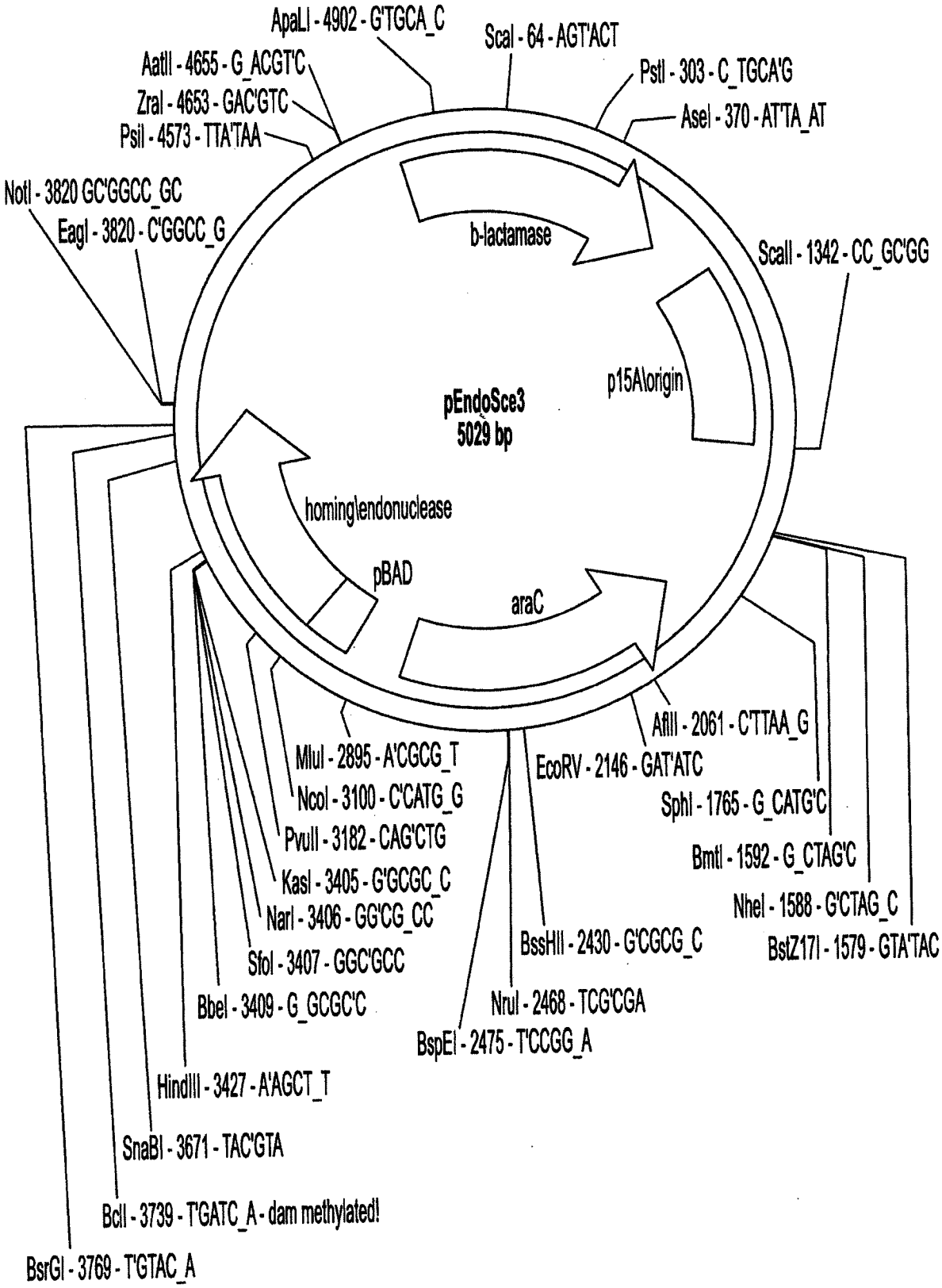

ApaLI - 4902 - G'TGCA_C
ScaI - 64 - AGT'ACT
AatII - 4655 - G_ACGT'C
PstI - 303 - C_TGCA'G
ZraI - 4653 - GAC'GTC
AseI - 370 - AT'TA_AT
PsiI - 4573 - TTA'TAA
NotI - 3820 GC'GGCC_GC
EagI - 3820 - C'GGCC_G
b-lactamase
ScaII - 1342 - CC_GC'GG
p15A\origin
pEndoSce3
5029 bp
homing\endonuclease
pBAD
araC
AflII - 2061 - C'TTAA_G
MluI - 2895 - A'CGCG_T
EcoRV - 2146 - GAT'ATC
NcoI - 3100 - C'CATG_G
SphI - 1765 - G_CATG'C
PvuII - 3182 - CAG'CTG
BmtI - 1592 - G_CTAG'C
KasI - 3405 - G'GCGC_C
NheI - 1588 - G'CTAG_C
NarI - 3406 - GG'CG_CC
BstZ17I - 1579 - GTA'TAC
SfoI - 3407 - GGC'GCC
BssHII - 2430 - G'CGCG_C
BbeI - 3409 - G_GCGC'C
NruI - 2468 - TCG'CGA
BspEI - 2475 - T'CCGG_A
HindIII - 3427 - A'AGCT_T
SnaBI - 3671 - TAC'GTA
BclI - 3739 - T'GATC_A - dam methylated!
BsrGI - 3769 - T'GTAC_A

FIG. 12

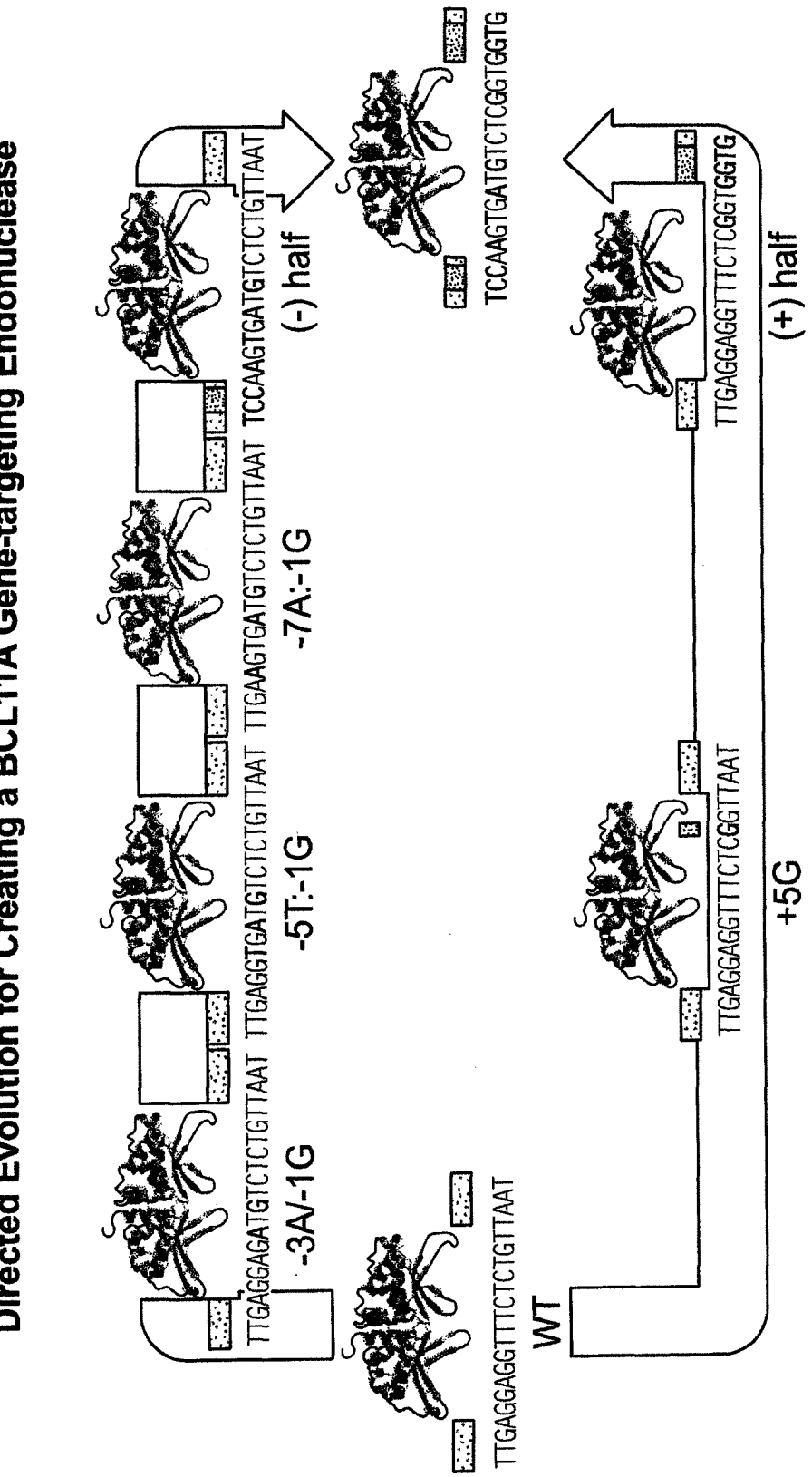

Directed Evolution for Creating a BCL11A Gene-targeting Endonuclease
TTGAGGAGATGTCTCTGTTAAT
-3A/-1G
TTGAGGAGGTTTCTCTGTTAAT
WT
TTGAGGAGATGTCTCTGTTAAT TTGAAGTGATGTCTCTGTTAAT TCCAAGTGATGTCTCTGTTAAT
-5T:-1G
-7A:-1G
(-) half
TCCAAGTGATGTCTCGGTGGTG
TTGAGGAGGTTTCTCGGTTAAT
+5G
TTGAGGAGGTTTCTCGGTGGTG
(+) half

FIG. 13

(SEQ ID NO: 28)

Nucleotide Sequence of I-HjeMI, Codon Optimized for Expression in E.coli

1	ATGGGATCCCACATGGACCTGACCTACGCTTACCTGGTTGGTCTGTTCGA	50
51	AGGTGACGGTTACTTCTCTATCACCAAAAAGGGTAAATACCTGACCTACG	100
101	AACTGGGTATCGAACTGTCTATCAAAGACGTTCAGCTGATCTACAAAATC	150
151	AAAGACATCCTGGGTGTTGGTAAAGTTTCTTTCCGTAAACGTAACGAAAT	200
201	CGAAATGGTTTCTCTGCGTATCCGTGACAAGAATCACCTGAAAAACTTCA	250
251	TCCTGCCGATCTTCGACAAATACCCGATGCTGTCTAACAAGCAGTACGAC	300
301	TACCTGCGTTTCAAAGACGCTCTCCTGTCTAACATTATCTACTCTGACGA	350
351	TCTGCCGGAATACGCTCGTTCTAACGAATCTATCAACTCTGTTGACTCTA	400
401	TTATCAACACCTCTTACTTCTCTGCTTGGCTGGTTGGTTTCATCGAAGCT	450
451	GAAGGTTGCTTCTCTACCTACAAACTGAACAAAGATGACGATTACCTGAT	500
501	CGCTTCTTTCGACATCGCTCAGAAAGACGGTGACATCCTGATCTCTGCTA	550
551	TCCACAAATACCTGTCTTTCACCACGAAAATCTACCTGGACAAAACCAAC	600
601	TGCTCTCGTCTGAAAGTGACCGGTGTACGTTCTGTTAAAAACGTGGTTAA	650
651	ATTCATCCAGGGTGCTCCGGTTAAACTGCTCGGTAACAAGAAACTGCAGT	700
701	ACAAACTGTGGATCAAACAGCTGCGTAAAATCTCTCGTTACTCTGAAAAA	750
751	ATCCAGCTGCCGTCTAACTAC	771

FIG. 14

(SEQ ID NO: 29)

Nucleotide Sequence of I-HjeMI, Codon Optimized for Expression in Mammals

1	ATGGGCAGCCACATGGACCTGACCTACGCCTATCTGGTCGGCCTGTTCGA	50
51	GGGCGACGGCTATTTTAGCATAACCAAGAAGGGCAAGTATCTGACGTATG	100
101	AACTGGGCATCGAGCTCTCCATCAAGGACGTGCAGCTCATCTACAAGATC	150
151	AAGGACATCCTCGGCGTGGGCAAAGTGTCTTTTAGGAAGAGGAACGAGAT	200
201	CGAGATGGTCAGCCTGCGAATCAGGGACAAAAACCACCTGAAGAACTTCA	250
251	TCCTGCCCATCTTCGACAAGTACCCCATGCTGAGCAACAAGCAGTACGAC	300
301	TATCTCCGATTCAAGGATGCCCTCCTGTCCAACATCATCTATAGCGACGA	350
351	CCTGCCCGAGTACGCCAGGAGCAACGAGTCAATCAATAGCGTGGACAGCA	400
401	TCATCAACACCTCATACTTCAGCGCCTGGCTGGTTGGCTTCATCGAGGCC	450
451	GAGGGCTGCTTCAGCACCTACAAGCTCAACAAGGACGACGATTATTTGAT	500
501	CGCGAGCTTCGATATAGCCCAGAAGGACGGCGACATTCTCATCTCCGCGA	550
551	TCCACAAATACCTGAGCTTCACGACCAAAATCTACCTGGACAAGACCAAC	600
601	TGTAGCAGGCTCAAGGTCACCGGCGTGAGGAGCGTCAAGAACGTGGTTAA	650
651	GTTCATCCAGGGTGCGCCGGTCAAGTTGCTGGGTAACAAGAAGCTGCAGT	700
701	ACAAACTTTGGATAAAGCAGCTGCGCAAGATCTCCCGATACAGCGAGAAA	750
751	ATCCAGCTGCCCAGTAACTAC	771

FIG. 15

(SEQ ID NO: 30)

Amino Acid Sequence of the Homing Endonuclease I-HjeMI

1	MGDLTYAYLVGLFEGDGYFSITKKGKYLTYELGIELSIKDVQLIYKIKDI	50
51	LGVGKVSFRKRNEIEMVSLRIRDKNHLKNFILPIFDKYPMLSNKQYDYLR	100
101	FKDALLSNIIYSDDLPEYARSNESINSVDSIINTSYFSAWLVGFIEAEGC	150
151	FSTYKLNKDDDYLIASFDIAQKDGDILISAIHKYLSFTTKIYLDKTNCSR	200
201	LKVTGVRSVKNVVKFIQGAPVKLLGNKKLQYKLWIKQLRKISRYSEKIQL	250
251	PSNY	254

FIG. 16

(SEQ ID NO: 31)

Nucleotide Sequence for a BCL11A Gene Targeting Nuclease Based on the Homing Endonuclease I-HjeMI

(Codon Optimized for Expression in E.coli and Obtained through Directed Evolution in IVC and in Bacteria)

1	ATGGGATCCCACATGGACCTGACCTACGCTTACCTGGTTGGTCTGTTCGA	50
51	AGGTGACGGTTACTTCACTATCGCTAAAGCGGGTAAGTACCTGAATTACG	100
101	AGCTGGGTATCACGCTGTCTATCAAAGACGCTCAGCTGATCTACAAAATC	150
151	AAAGACATCCTGGGTGTTGGTAATGTTTATTTCCGGAAATATAGGCAGCA	200
201	TGAAATGGTTTCTCTGCGGATCCAGGACAAGAATCACCTGAAAAACTTCA	250
251	TCCTGCCGATCTTCGACAAATACCCGATGCTGTCTAACAAGCAGTACGAC	300
301	TACCTGCGTTTCAAAGACGCTCTCCTGTCTAACATTATCTACTCTGACGA	350
351	TCTGCCGGAATACGCTCGTTCTAACGAATCTATCAACTCTGTTGACTCTA	400
401	TTATCAACACCTCTTACTTCTCTGCTTGGCTGGTTGGTTTCATCGAAGCT	450
451	GAAGGTTGCTTCACGACCTACAAAGCGAGTAAAGATAAGTACCTGACGGC	500
501	TGGGTTCAGTATCGCTCAGAAAGACGGTGACATCCTGATCTCTGCTATCC	550
551	ACAAATACCTGTCTTTCACCACGAAACCGTACAAGGACAAAACCAACTGC	600
601	TCTCATCTGAAGGTGACCGGTGTACGTTCTGTTAACAACGTGGTTAAATT	650
651	CATCCAGGGTGCTCCGGTTAAACTGCTCGGTAACAAGAAACTGCAGTACA	700
701	AACTGTGGATCAAACAGCTGCGTAAAATCTCTCGTTACTCTGAAAAAATC	750
751	CAGCTGCCGTCTAACTAC	768

FIG. 17

(SEQ ID NO: 32)

Nucleotide Sequence of a BCL11A Gene Targeting Nuclease Based on the Homing Endonuclease I-HjeMI

(Codon Optimized for Expression in Mammals and Obtained through Directed Evolution in IVC and in Bacteria)

```
   1 ATGGGCAGCCACATGGACCTGACCTACGCCTATCTGGTCGGCCTGTTCGA  50
  51 GGGCGACGGCTATTTTACaATAGCtAAGgccGGCAAGTATCTGAacTACG 100
 101 AGCTGGGCATCacaCTCTCCATCAAGGACGCtCAGCTCATCTACAAGATC 150
 151 AAGGACATCCTCGGCGTGGGCAAcGTGtaCTTTAGGAAGtacAggcAaca 200
 201 tGAGATGGTCAGCCTGCGAATCcaGGACAAAAACCACCTGAAGAACTTCA 250
 251 TCCTGCCCATCTTCGACAAGTACCCCATGCTGAGCAACAAGCAGTACGAC 300
 301 TACCTGCGATTCAAGGATGCCCTCCTGTCCAACATCATCTATAGCGACGA 350
 351 CCTGCCCGAGTACGCCAGGAGCAACGAGTCAATCAATAGCGTGGACAGCA 400
 401 TCATCAACACCTCATACTTCAGCGCCTGGCTGGTTGGCTTCATCGAGGCC 450
 451 GAGGGCTGCTTCAcCACCTACAAGgcatcCAAGGACaAgTACCTGAcaGC 500
 501 GgGCTTCtccATAGCCCAGAAGGACGGCGACATTCTCATCTCCGCGATCC 550
 551 ACAAATACCTGAGCTTCACGACCAAAccCTACaaaGACAAGACCAACTGT 600
 601 AGCcacCTCAAGGTCACCGGCGTGAGGAGCGTCAAtAACGTGGTTAAGTT 650
 651 CATCCAGGGTGCGCCGGTCAAGCTGCTGGGTAACAAGAAGCTGCAGTACA 700
 701 AACTTTGGATAAAGCAGCTGCGCAAGATCTCCCGATACAGCGAGAAAATC 750
 751 CAGCTGCCCAGTAACTAC                                  768
```

FIG. 18

(SEQ ID NO: 33)

Amino Acid Sequence of a BCL11A Gene Targeting Nuclease Based on the Homing Endonuclease I-HjeMI

```
  1 MGSHMDLTYAYLVGLFEGDGYFTIAKAGKYLNYELGITLSIKDAQLIYKI  50
 51 KDILGVGNVYFRKYRQHEMVSLRIQDKNHLKNFILPIFDKYPMLSNKQYD 100
101 YLRFKDALLSNIIYSDDLPEYARSNESINSVDSIINTSYFSAWLVGFIEA 150
151 EGCFTTYKASKDKYLTAGFSIAQKDGDILISAIHKYLSFTTKPYKDKTNC 200
201 SHLKVTGVRSVNNVVKFIQGAPVKLLGNKKLQYKLWIKQLRKISRYSEKI 250
251 QLPSNY                                              256
```

FIG. 19

Distribution of Amino-acid Residues Different between the BCL11A Gene-targeting Endonuclease and its Parental LHE I-HjeMI

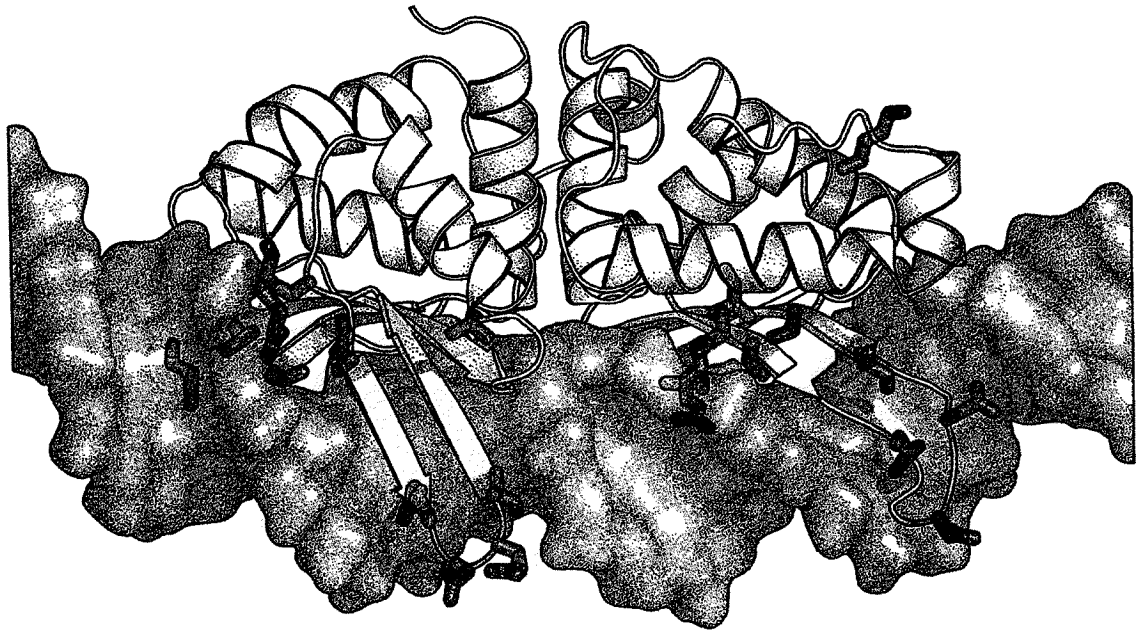


FIG. 20

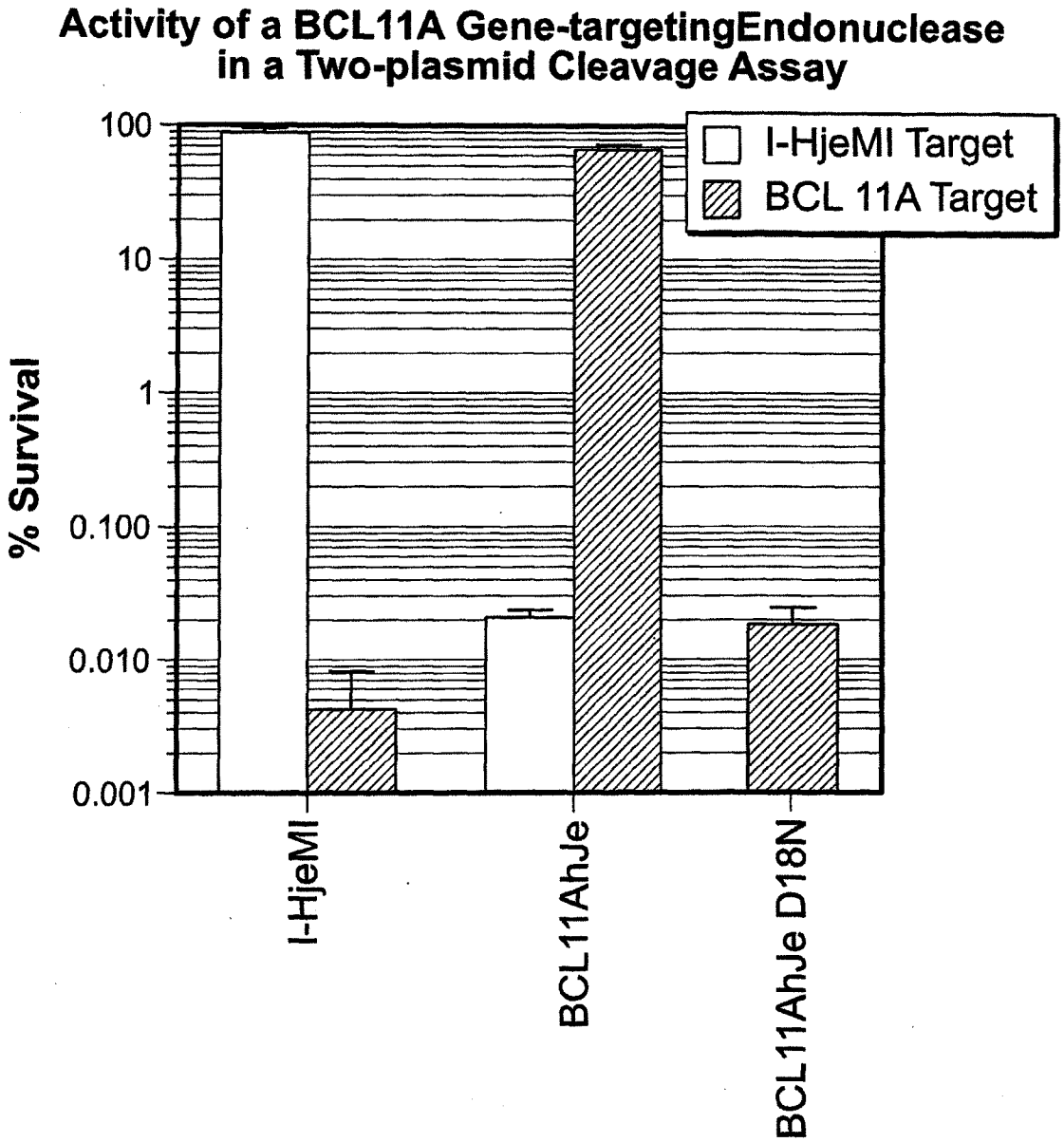

Activity of a BCL11A Gene-targetingEndonuclease
in a Two-plasmid Cleavage Assay
% Survival
100
10
1
0.100
0.010
0.001
I-HjeMI Target
BCL 11A Target
I-HjeMI
BCL11AhJe
BCL11AhJe D18N

FIG. 21

(SEQ ID NO: 34)
Nucleotide Sequence of I-OnuI, Codon Optimized for Expression in *E. coli*
(parental enzyme for HEs targeting the HbF silencing)

```
   1 ATGTCCGCCTACATGTCCCGTCGCGAGTCCATTAACCCGTGGATTCTCAC  50
  51 CGGTTTCGCCGACGCGGAAGGCTCCTTTTTGCTGCGCATCCGCAACAACA 100
 101 ACAAGTCCAGCGTCGGCTACTCCACTGAGCTCGGCTTCCAAATTACACTT 150
 151 CATAACAAGGACAAGAGCATTCTTGAGAACATCCAGTCAACATGGAAGGT 200
 201 GGGCGTGATCGCCAACAGCGGTGACAACGCCGTGTCGCTGAAGGTCACGC 250
 251 GTTTTGAGGACCTGAAGGTCATTATCGACCATTTTGAAAAATACCCACTG 300
 301 ATTACGCAGAAGCTCGGTGACTACATGCTGTTTAAGCAGGCGTTTTGCGT 350
 351 CATGGAGAACAAGGAGCATTTGAAGATTAATGGTATCAAGGAGCTGGTGC 400
 401 GCATTAAGGCAAAGCTCAATTGGGGTCTGACGGATGAGCTGAAGAAGGCC 450
 451 TTTCCGGAGATCATCTCGAAGGAGCGCTCCCTCATCAACAAGAACATCCC 500
 501 TAATTTCAAGTGGCTGGCGGGTTTTACCTCGGGCGAGGGTTGCTTCTTTG 550
 551 TTAACCTGATCAAGTCAAAGTCGAAGCTAGGTGTCCAGGTGCAGCTGGTG 600
 601 TTCAGCATTACCCAACACATCAAGGATAAGAACCTCATGAACTCTCTGAT 650
 651 TACCTACTTGGGCTGCGGCTACATTAAGGAGAAAAACAAGAGTGAGTTCT 700
 701 CCTGGCTTGACTTCGTCGTCACGAAATTCTCCGACATCAACGACAAGATC 750
 751 ATTCCGGTCTTTCAGGAAAACACGCTCATCGGCGTGAAGCTCGAGGACTT 800
 801 CGAGGATTGGTGTAAGGTCGCTAAGCTGATCGAGGAGAAAAAGCACCTGA 850
 851 CAGAAAGTGGCCTGGACGAGATCAAGAAGATTAAGCTGAACATGAACAAG 900
 901 GGCAGAGTATTC                                        912
```

FIG. 22A

(SEQ ID NO: 15)

Amino Acid Sequence of I-Onulhoming endonuclease.

```
SAYMSRRESI NPWILTGFAD AEGSFLLRIR NNNKSSVGYS TELGFQITLH NKDKSILENI  60
QSTWKVGVIA NSGDNAVSLK VTRFEDLKVI IDHFEKYPLI TQKLGDYMLF KQAFCVMENK 120
EHLKINGIKE LVRIKAKLNW GLTDELKKAF PEIISKERSL INKNIPNFKW LAGFTSGEGC 180
FFVNLIKSKS KLGVQVQLVF SITQHIKDKN LMNSLITYLG CGYIKEKNKS EFSWLDFVVT 240
KFSDINDKII PVFQENTLIG VKLEDFEDWC KVAKLIEEKK HLTESGLDEI KKIKLNMNKG 300
RVF                                                               303
```

FIG. 22B

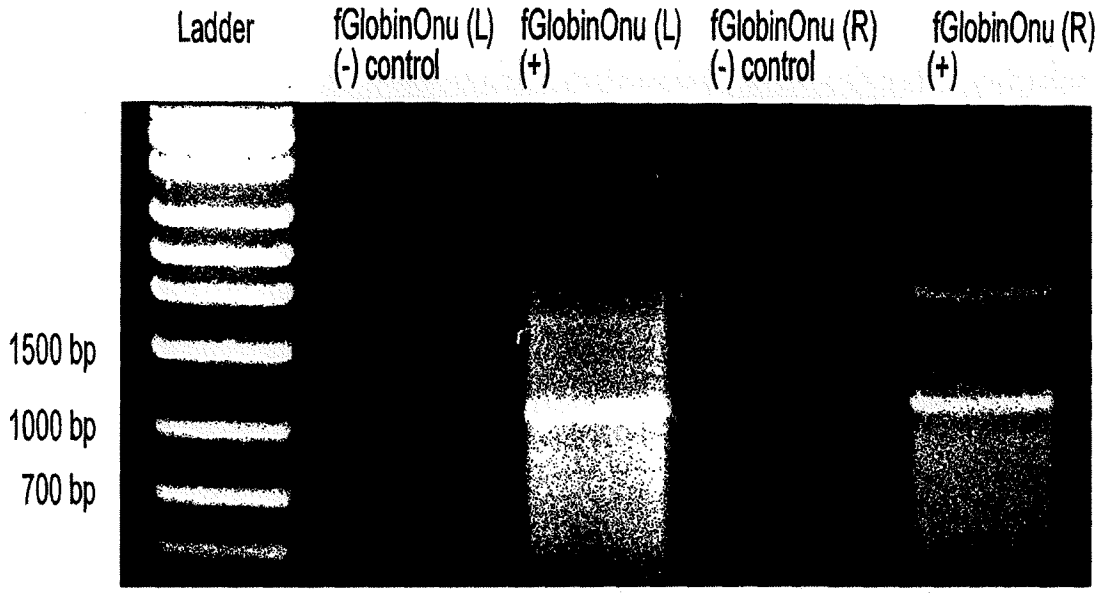

Ladder
fGlobinOnu (L)
(-) control
fGlobinOnu (L)
(+)
fGlobinOnu (R)
(-) control
fGlobinOnu (R)
(+)
1500 bp
1000 bp
700 bp

FIG. 23

(SEQ ID NO: 35)

Nucleotide Sequence of MegaTAL:5.5 RVD + Y2 I-AniI

1	gtggatctacgcacgctcggctacagtcagcagcagcaagagaagatcaa	50
51	accgaaggtgcgttcgacagtggcgcagcaccacgaggcactggtgggcc	100
101	atgggtttacacacgcgcacatcgttgcgctcagccaacacccggcagcg	150
151	ttagggaccgtcgctgtcacgtatcagcacataatcacggcgttgccaga	200
201	ggcgacacacgaagacatcgttggcgtcggcaaacagtggtccggcgcac	250
251	gcgccctggaggccttgctcacggatgcgggggagttgagaggtccgccg	300
301	ttacagttggacacaggccaacttgtgaagattgcaaaacgtggcggcgt	350
351	gaccgcaatggaggcagtgcatgcatcgcgcaatgcactgacgggtgccc	400
401	ccctgaacctgaccccggaccaagtggtggctatcgccagcaacaatggc	450
451	ggcaagcaagcgctcgaaacggtgcagcggctgttgccggtgctgtgcca	500
501	ggaccatggcctgactccggaccaagtggtggctatcgccagccacgatg	550
551	gcggcaagcaagcgctcgaaacggtgcagcggctgttgccggtgctgtgc	600
601	caggaccatggcctgaccccggaccaagtggtggctatcgccagcaacat	650
651	tggcggcaagcaagcgctcgaaacggtgcagcggctgttgccggtgctgt	700
701	gccaggaccatggcctgaccccggaccaagtggtggctatcgccagcaac	750
751	aatggcggcaagcaagcgctcgaaacggtgcagcggctgttgccggtgct	800
801	gtgccaggaccatggcctgactccggaccaagtggtggctatcgccagcc	850
851	acgatggcggcaagcaagcgctcgaaacggtgcagcggctgttgccggtg	900

FIG. 24

901 ctgtgccaggaccatggcctgaccccggaccaagtggtggctatcgccag 950
951 caacggtggcggcaagcaagcgctcgaaagcattgtggcccagctgagcc 1000
1001 ggcctgatccggcgttggccgcgttgaccaacgaccacctcgtcgccttg 1050
1051 gcctgcctcggcggacgtcctgccatggatgcagtgaaaaagggattgcc 1100
1101 gcacgcgccggaattgatcagaagagtcaatcgccgtattggcgaacgca 1150
1151 cgtcccatcgcgttgcgatatctagagtgggaggaagcgatcttacgtac 1200
1201 gcgtatttagttggtctctacgaaggggatggatactttagtatcaccaa 1250
1251 gaaaggcaagtacttgacttatgaattaggtattgagctgagcatcaaag 1300
1301 acgtccaattgatttacaagatcaagaaaatcctaggtattggcatcgta 1350
1351 agcttcaggaagagaaacgagattgaaatggttgcattgaggatccgtga 1400
1401 taagaatcatctaaaatctaagatattgcctatatttgagaagtatccaa 1450
1451 tgttttccaacaaacagtacgactatttaagattcaggaatgcattgtta 1500
1501 tctggcattatatacctagaagacttgcctgattacactagaagtgacga 1550
1551 accattgaattctatagaatccattatcaacacatcatacttctccgcct 1600
1601 ggctagttggatttatagaagctgagggctgtttcagtgtgtacaagctg 1650
1651 aacaaagacgatgactacttgattgcttcattcgacattgcccaaagaga 1700
1701 tggtgatatcttgatttcagcaattaggaagtacttaagtttcactacta 1750
1751 aggtttacctagacaagactaattgtagcaaattgaaggtcactagtgtt 1800
1801 agatccgtcgagaacatcattaagtttctgcagaatgctcctgtcaaatt 1850
1851 gttaggcaacaagaaactgcaatacaagttgtggttgaaacaactaagga 1900
1901 agatttctaggtattccgagaagatcaagattccatcaaactacgtcgac 1950
1951 cgagcatcttaccgccatttatacccatatttgttctgtttttcttga 1998

FIG. 24 (Cont.)

(SEQ ID NO: 36)

Amino Acid Sequence of MegaTAL:5.5 RVD + Y2 I-AniI

```
   1 VDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAA  50
  51 LGTVAVTYQHIITALPEATHEDIVGVGKQWSGARALEALLTDAGELRGPP 100
 101 LQLDTGQLVKIAKRGGVTAMEAVHASRNALTGAPLNLTPDQVVAIASNNG 150
 151 GKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLC 200
 201 QDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASN 250
 251 NGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPV 300
 301 LCQDHGLTPDQVVAIASNGGGKQALESIVAQLSRPDPALAALTNDHLVAL 350
 351 ACLGGRPAMDAVKKGLPHAPELIRRVNRRIGERTSHRVAISRVGGSDLTY 400
 401 AYLVGLYEGDGYFSITKKGKYLTYELGIELSIKDVQLIYKIKKILGIGIV 450
 451 SFRKRNEIEMVALRIRDKNHLKSKILPIFEKYPMFSNKQYDYLRFRNALL 500
 501 SGIIYLEDLPDYTRSDEPLNSIESIINTSYFSAWLVGFIEAEGCFSVYKL 550
 551 NKDDDYLIASFDIAQRDGDILISAIRKYLSFTTKVYLDKTNCSKLKVTSV 600
 601 RSVENIIKFLQNAPVKLLGNKKLQYKLWLKQLRKISRYSEKIKIPSNYVD 650
 651 RASYRHLYPYLFCFS                                    665
```

FIG. 25

(SEQ ID NO: 37)

Nucleotide Sequence of Cas9 Endonuclease

(from Mali *et al.*, *Science* (2013))

1	gccacc**ATG**GACAAGAAGTACTCC**ATTGGGCTCGATATCGGCACAAACAG**	50
51	**CGTCGGCTGGGCCGTCATT**ACGGACGAGTACAAGGTGCCGAGCAAAAAAT	100
101	TCAAAGTTCTGGGCAATACCGATCGCCACAGCATAAAGAAGAACCTCATT	150
151	GGCGCCCTCCTGTTCGACTCCGGGGAGACGGCCGAAGCCACGCGGCTCAA	200
201	AAGAACAGCACGGCGCAGATATACCCGCAGAAAGAATCGGATCTGCTACC	250
251	TGCAGGAGATCTTTAGTAATGAGATGGCTAAGGTGGATGACTCTTTCTTC	300
301	CATAGGCTGGAGGAGTCCTTTTTGGTGGAGGAGGATAAAAAGCACGAGCG	350
351	CCACCCAATCTTTGGCAATATCGTGGACGAGGTGGCGTACCATGAAAAGT	400
401	ACCCAACCATATATCATCTGAGGAAGAAGCTTGTAGACAGTACTGATAAG	450
451	GCTGACTTGCGGTTGATCTATCTCGCGCTGGCGCATATGATCAAATTTCG	500
501	GGGACACTTCCTCATCGAGGGGGACCTGAACCCAGACAACAGCGATGTCG	550
551	ACAAACTCTTTATCCAACTGGTTCAGACTTACAATCAGCTTTTCGAAGAG	600
601	AACCCGATCAACGCATCCGGAGTTGACGCCAAAGCAATCCTGAGCGCTAG	650
651	GCTGTCCAAATCCCGGCGGCTCGAAAACCTCATCGCACAGCTCCCTGGGG	700
701	AGAAGAAGAACGGCCTGTTTGGTAATCTTATCGCCCTGTCACTCGGGCTG	750
751	ACCCCCAACTTTAAATCTAACTTCGACCTGGCCGAAGATGCCAAGCTTCA	800
801	ACTGAGCAAAGACACCTACGATGATGATCTCGACAATCTGCTGGCCCAGA	850
851	TCGGCGACCAGTACGCAGACCTTTTTTTGGCGGCAAAGAACCTGTCAGAC	900
901	GCCATTCTGCTGAGTGATATTCTGCGAGTGAACACGGAGATCACCAAAGC	950
951	TCCGCTGAGCGCTAGTATGATCAAGCGCTATGATGAGCACCACCAAGACT	1000
1001	TGACTTTGCTGAAGGCCCTTGTCAGACAGCAACTGCCTGAGAAGTACAAG	1050
1051	GAAATTTTCTTCGATCAGTCTAAAAATGGCTACGCCGGATACATTGACGG	1100
1101	CGGAGCAAGCCAGGAGGAATTTTACAAATTTATTAAGCCCATCTTGGAAA	1150
1151	AAATGGACGGCACCGAGGAGCTGCTGGTAAAGCTTAACAGAGAAGATCTG	1200
1201	TTGCGCAAACAGCGCACTTTCGACAATGGAAGCATCCCCCACCAGATTCA	1250
1251	CCTGGGCGAACTGCACGCTATCCTCAGGCGGCAAGAGGATTTCTACCCCT	1300

FIG. 26

1301 TTTTGAAAGATAACAGGGAAAAGATTGAGAAAATCCTCACATTTCGGATA 1350
1351 CCCTACTATGTAGGCCCCCTCGCCCGGGGAAATTCCAGATTCGCGTGGAT 1400
1401 GACTCGCAAATCAGAAGAGACCATCACTCCCTGGAACTTCGAGGAAGTCG 1450
1451 TGGATAAGGGGGCCTCTGCCCAGTCCTTCATCGAAAGGATGACTAACTTT 1500
1501 GATAAAAATCTGCCTAACGAAAAGGTGCTTCCTAAACACTCTCTGCTGTA 1550
1551 CGAGTACTTCACAGTTTATAACGAGCTCACCAAGGTCAAATACGTCACAG 1600
1601 AAGGGATGAGAAAGCCAGCATTCCTGTCTGGAGAGCAGAAGAAAGCTATC 1650
1651 GTGGACCTCCTCTTCAAGACGAACCGGAAAGTTACCGTGAAACAGCTCAA 1700
1701 AGAAGACTATTTCAAAAAGATTGAATGTTTCGACTCTGTTGAAATCAGCG 1750
1751 GAGTGGAGGATCGCTTCAACGCATCCCTGGGAACGTATCACGATCTCCTG 1800
1801 AAAATCATTAAAGACAAGGACTTCCTGGACAATGAGGAGAACGAGGACAT 1850
1851 TCTTGAGGACATTGTCCTCACCCTTACGTTGTTTGAAGATAGGGAGATGA 1900
1901 TTGAAGAACGCTTGAAAACTTACGCTCATCTCTTCGACGACAAAGTCATG 1950
1951 AAACAGCTCAAGAGGCGCCGATATACAGGATGGGGGCGGCTGTCAAGAAA 2000
2001 ACTGATCAATGGGATCCGAGACAAGCAGAGTGGAAAGACAATCCTGGATT 2050
2051 TTCTTAAGTCCGATGGATTTGCCAACCGGAACTTCATGCAGTTGATCCAT 2100
2101 GATGACTCTCTCACCTTTAAGGAGGACATCCAGAAAGCACAAGTTTCTGG 2150
2151 CCAGGGGGACAGTCTTCACGAGCACATCGCTAATCTTGCAGGTAGCCCAG 2200
2201 CTATCAAAAAGGGAATACTGCAGACCGTTAAGGTCGTGGATGAACTCGTC 2250
2251 AAAGTAATGGGAAGGCATAAGCCCGAGAATATCGTTATCGAGATGGCCCG 2300
2301 AGAGAACCAAACTACCCAGAAGGGACAGAAGAACAGTAGGGAAAGGATGA 2350
2351 AGAGGATTGAAGAGGGTATAAAAGAACTGGGGTCCCAAATCCTTAAGGAA 2400
2401 CACCCAGTTGAAAACACCCAGCTTCAGAATGAGAAGCTCTACCTGTACTA 2450
2451 CCTGCAGAACGGCAGGGACATGTACGTGGATCAGGAACTGGACATCAATC 2500
2501 GGCTCTCC**GACTACGACGTGGATCATATCGTGCCCCAGTCTTTTCTCAAA** **2550**
2551 **GATGATTCTATTGATAATAAAGTGTTGACAAGATCCGATAAAAAT**AGAGG 2600
2601 GAAGAGTGATAACGTCCCCTCAGAAGAAGTTGTCAAGAAAATGAAAAATT 2650
2651 ATTGGCGGCAGCTGCTGAACGCCAAACTGATCACACAACGGAAGTTCGAT 2700
2701 AATCTGACTAAGGCTGAACGAGGTGGCCTGTCTGAGTTGGATAAAGCCGG 2750
2751 CTTCATCAAAAGGCAGCTTGTTGAGACACGCCAGATCACCAAGCACGTGG 2800
2801 CCCAAATTCTCGATTCACGCATGAACACCAAGTACGATGAAAATGACAAA 2850
2851 CTGATTCGAGAGGTGAAAGTTATTACTCTGAAGTCTAAGCTGGTCTCAGA 2900

FIG. 26 (Cont.)

2901	TTTCAGAAAGGACTTTCAGTTTTATAAGGTGAGAGAGATCAACAATTACC	2950
2951	ACCATGCGCATGATGCCTACCTGAATGCAGTGGTAGGCACTGCACTTATC	3000
3001	AAAAAATATCCCAAGCTTGAATCTGAATTTGTTTACGGAGACTATAAAGT	3050
3051	GTACGATGTTAGGAAAATGATCGCAAAGTCTGAGCAGGAAATAGGCAAGG	3100
3101	CCACCGCTAAGTACTTCTTTTACAGCAATATTATGAATTTTTTCAAGACC	3150
3151	GAGATTACACTGGCCAATGGAGAGATTCGGAAGCGACCACTTATCGAAAC	3200
3201	AAACGGAGAAACAGGAGAAATCGTGTGGGACAAGGGTAGGGATTTCGCGA	3250
3251	CAGTCCGGAAGGTCCTGTCCATGCCGCAGGTGAACATCGTTAAAAAGACC	3300
3301	GAAGTACAGACCGGAGGCTTCTCCAAGGAAAGTATCCTCCCGAAAAGGAA	3350
3351	CAGCGACAAGCTGATCGCACGCAAAAAAGATTGGGACCCCAAGAAATACG	3400
3401	GCGGATTCGATTCTCCTACAGTCGCTTACAGTGTACTGGTTGTGGCCAAA	3450
3451	GTGGAGAAAGGGAAGTCTAAAAAACTCAAAAGCGTCAAGGAACTGCTGGG	3500
3501	CATCACAATCATGGAGCGATCAAGCTTCGAAAAAAACCCCATCGACTTTC	3550
3551	TCGAGGCGAAAGGATATAAAGAGGTCAAAAAAGACCTCATCATTAAGCTT	3600
3601	CCCAAGTACTCTCTCTTTGAGCTTGAAAACGGCCGGAAACGAATGCTCGC	3650
3651	TAGTGCGGGCGAGCTGCAGAAAGGTAACGAGCTGGCACTGCCCTCTAAAT	3700
3701	ACGTTAATTTCTTGTATCTGGCCAGCCACTATGAAAAGCTCAAAGGGTCT	3750
3751	CCCGAAGATAATGAGCAGAAGCAGCTGTTCGTGGAACAACACAAACACTA	3800
3801	CCTTGATGAGATCATCGAGCAAATAAGCGAATTCTCCAAAAGAGTGATCC	3850
3851	TCGCCGACGCTAACCTCGATAAGGTGCTTTCTGCTTACAATAAGCACAGG	3900
3901	GATAAGCCCATCAGGGAGCAGGCAGAAAACATTATCCACTTGTTTACTCT	3950
3951	GACCAACTTGGGCGCGCCTGCAGCCTTCAAGTACTTCGACACCACCATAG	4000
4001	ACAGAAAGCGGTACACCTCTACAAAGGAGGTCCTGGACGCCACACTGATT	4050
4051	CATCAGTCAATTACGGGGCTCTATGAAACAAGAATCGACCTCTCTCAGCT	4100
4101	CGGTGGAGACAGCAGGGCTGACCCCAAGAAGAAGAGGAAGGTGTGA	4046

FIG. 26 (Cont.)

(SEQ ID NO: 38)

Nucleotide Sequence of RNA Guide Strand for use with Cas9 Endonuclease (from Mali *et al.*, *Science* (2013))

```
   1 TGTACAAAAAAGCAGGCTTTAAAGGAACCAATTCAGTCGACTGGATCCGG  50
  51 TACCAAGGTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATT 100
 101 TGCATATACGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGACT 150
 151 GTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATT 200
 201 TCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTATCATA 250
 251 TGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTG 300
 301 TGGAAAGGACGAAACACCGNNNNNNNNNNNNNNNNNNGTTTTAGAGCTA 350
 351 GAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGG 400
 401 CACCGAGTCGGTGCTTTTTTTTCTAGACCCAGCTTTCTTGTACAAAGTTGG 450
 451 CATTA                                              455
```

FIG. 27

(SEQ ID NO: 62)

Nucleotide Sequence of a I-CpaMI homing endonuclease

(ORF, codon optimized for mammalian expression)

ATGAACACCA GCTCTAGCTT CAATCCCTGG TTCCTGACCG GCTTTAGCGA TGCAGAGTGC 60
TCTTTCAGCA TCCTGATACA GGCCAACAGC AAGTACTCCA CCGGTTGGAG GATCAAGCCC 120
GTGTTCGCCA TCGGCTTGCA CAAGAAGGAC CTGGAGCTTC TGAAGAGAAT CCAGAGCTAT 180
CTGGGCGTGG GCAAGATACA CATTCACGGC AAAGACAGCA TTCAGTTCAG GATTGACAGC 240
CCCAAGGAGC TGGAGGTGAT CATCAACCAC TTTGAGAACT ACCCCCTGGT AACCGCCAAG 300
TGGGCCGACT ACACCCTCTT TAAGAAGGCC CTGGACGTAA TTCTGTTGAA GGAGCACCTG 360
AGCCAGAAGG GCCTGCTTAA ACTGGTAGGC ATTAAGGCGA GCCTGAATCT CGGGTTGAAC 420
GGCAGCCTCA AGGAGGCGTT CCCGAACTGG GAAGAACTGC AGATCGACAG GCCGAGCTAC 480
GTGTTCAAGG GCATCCCCGA CCCCAACTGG ATCAGCGGCT TCGCGTCAGG CGATAGCAGC 540
TTTAATGTGA AAATCAGCAA CTCCCCCACG TCACTGCTCA ATAAAAGGGT GCAGCTGAGG 600
TTCGGCATCG GACTGAACAT CAGAGAGAAA GCCCTTATCC AATACCTGGT GGCCTACTTT 660
GACCTGTCAG ACAACCTGAA GAACATCTAC TTCGACCTGA ACAGCGCACG GTTCGAGGTG 720
GTGAAGTTCA GCGACATCAC CGACAAGATC ATCCCCTTCT TCGACAAGTA CAGCATACAA 780
GGCAAGAAGA GCCTGGACTA CATCAACTTC AAGGAAGTGG CCGACATTAT CAAGAGCAAG 840
AACCATCTTA CTAGCGAGGG CTTCCAGGAA ATCTTGGACA TCAAAGCCAG TATGAACAAG 900

FIG. 28

(SEQ ID NO: 63)

Amino Acid Sequence of I-CpaMI homing endonuclease

```
MNTSSSFNPW FLTGFSDAEC SFSILIQANS KYSTGWRIKP VFAIGLHKKD LELLKRIQSY  60
LGVGKIHIHG KDSIQFRIDS PKELEVIINH FENYPLVTAK WADYTLFKKA LDVILLKEHL 120
SQKGLLKLVG IKASLNLGLN GSLKEAFPNW EELQIDRPSY VFKGIPDPNW ISGFASGDSS 180
FNVKISNSPT SLLNKRVQLR FGIGLNIREK ALIQYLVAYF DLSDNLKNIY FDLNSARFEV 240
VKFSDITDKI IPFFDKYSIQ GKKSLDYINF KEVADIIKSK NHLTSEGFQE ILDIKASMNK 300
```

FIG. 29

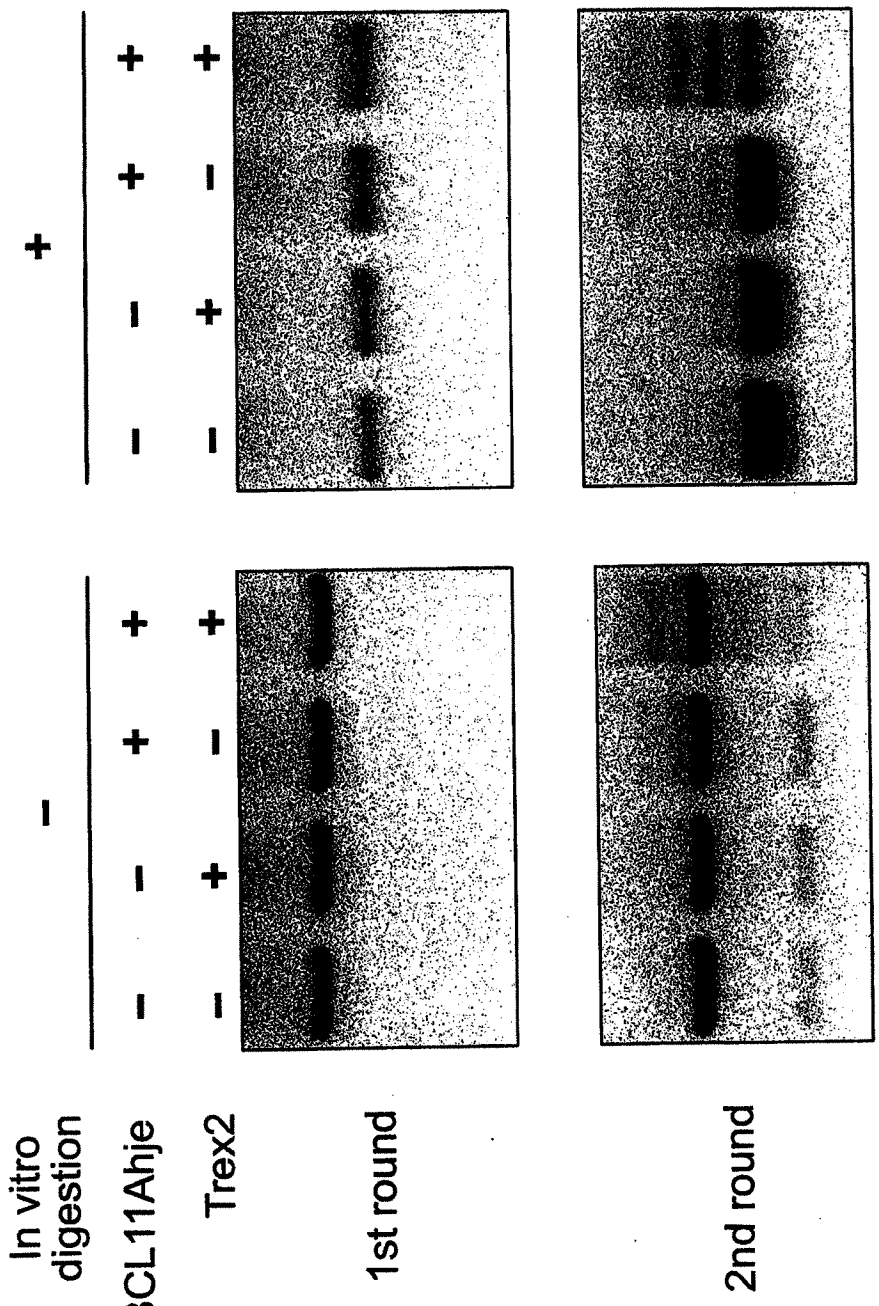


FIG. 30

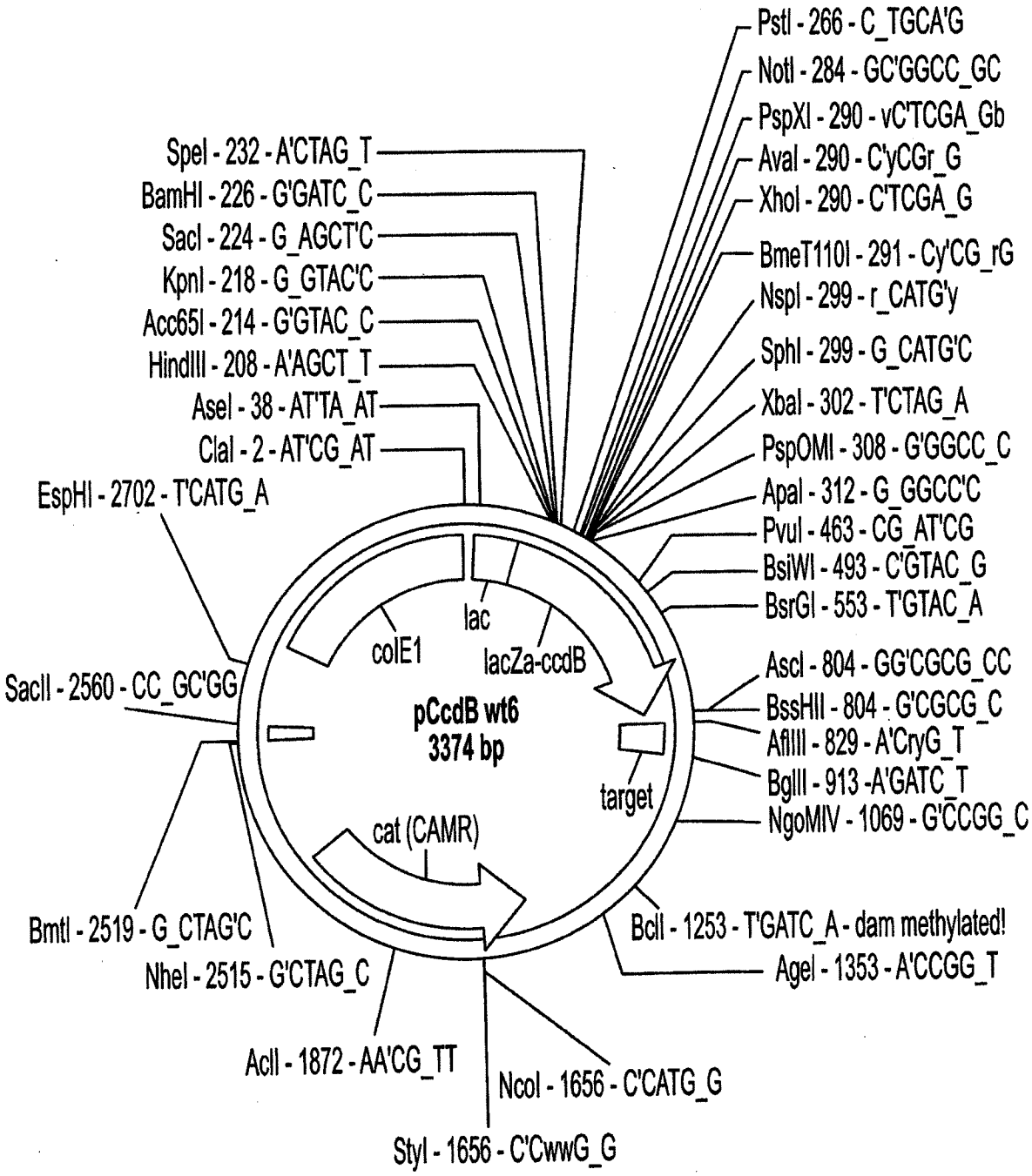

PstI - 266 - C_TGCA'G
NotI - 284 - GC'GGCC_GC
PspXI - 290 - vC'TCGA_Gb
AvaI - 290 - C'yCGr_G
XhoI - 290 - C'TCGA_G
BmeT110I - 291 - Cy'CG_rG
NspI - 299 - r_CATG'y
SphI - 299 - G_CATG'C
XbaI - 302 - T'CTAG_A
PspOMI - 308 - G'GGCC_C
ApaI - 312 - G_GGCC'C
PvuI - 463 - CG_AT'CG
BsiWI - 493 - C'GTAC_G
BsrGI - 553 - T'GTAC_A
AscI - 804 - GG'CGCG_CC
BssHII - 804 - G'CGCG_C
AflIII - 829 - A'CryG_T
BglII - 913 -A'GATC_T
NgoMIV - 1069 - G'CCGG_C
BclI - 1253 - T'GATC_A - dam methylated!
AgeI - 1353 - A'CCGG_T
NcoI - 1656 - C'CATG_G
StyI - 1656 - C'CwwG_G
AclI - 1872 - AA'CG_TT
NheI - 2515 - G'CTAG_C
BmtI - 2519 - G_CTAG'C
SacII - 2560 - CC_GC'GG
EspHI - 2702 - T'CATG_A
ClaI - 2 - AT'CG_AT
AseI - 38 - AT'TA_AT
HindIII - 208 - A'AGCT_T
Acc65I - 214 - G'GTAC_C
KpnI - 218 - G_GTAC'C
SacI - 224 - G_AGCT'C
BamHI - 226 - G'GATC_C
SpeI - 232 - A'CTAG_T
lac
colE1
lacZa-ccdB
pCcdB wt6
3374 bp
target
cat (CAMR)

FIG. 31

(SEQ ID NO: 64)

Nucleotide sequence of BCL11A gene targeting nuclease-encoding plasmid (pExodusBCL11Ahje)

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTAT
CTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGA
CAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATT
GATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAA
CTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGT
AACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGT
ATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTA
TGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAA
TGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC
AAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAG
GTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAG
GGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTTGGTACCGAGCTCGGATCCACTAGTCCAGTGTGGTGGAATTCTGC
AGGTACGTTGACGCCGCCACCATGGGATATCCATACGATGTCCCAGATTATGCGCCACCTAAGAAGAAACGCAAAGTCCC
CGGGGGCAGCCACATGGACCTGACCTACGCCTATCTGGTCGGCCTGTTCGAGGGCGACGGCTATTTTAcaATAgCtAAGg
ccGGCAAGTATCTGAacTACGAGCTGGGCATCacaCTCTCCATCAAGGACGctCAGCTCATCTACAAGATCAAGGACATC
CTCGGCGTGGGCAAcGTGtaCTTTAGGAAGtacAggcAacatGAGATGGTCAGCCTGCGAATCcaGGACAAAAACCACCT
GAAGAACTTCATCCTGCCCATCTTCGACAAGTACCCCATGCTGAGCAACAAGCAGTACGACTACCTGCGATTCAAGGATG
CCCTCCTGTCCAACATCATCTATAGCGACGACCTGCCCGAGTACGCCAGGAGCAACGAGTCAATCAATAGCGTGGACAGC
ATCATCAACACCTCATACTTCAGCGCCTGGCTGGTTGGCTTCATCGAGGCCGAGGGCTGCTTCAcCACCTACAAGgcatc
CAAGGACaAgTACCTGAcaGCGgGCTTCtccATAGCCCAGAAGGACGGCGACATTCTCATCTCCGCGATCCACAAATACC
TGAGCTTCACGACCAAAccCTACaaaGACAAGACCAACTGTAGCcacCTCAAGGTCACCGGCGTGAGGAGCGTCAAtAAC
GTGGTTAAGTTCATCCAGGGTGCGCCGGTCAAGCTGCTGGGTAACAAGAAGCTGCAGTACAAACTTTGGATAAAGCAGCT
GCGCAAGATCTCCCGATACAGCGAGAAAATCCAGCTGCCCAGTAACTACTAATCTAGAGGGCCCGTTTAAACCCGCTGAT
CAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCC
ACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGG
GGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTG
AGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTG
GTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCAC
GTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACC
CCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAG
TCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATA
AGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTAATTCTGTGGAA
TGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCCTATCAGGACATAGCGTTG
GCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGA
TTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAAGC
GACGCCCAACCTGCCATCACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGG
ACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTAT
AATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTC

FIG. 32

CAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGT
TTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCC
TAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCA
TTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGC
GCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAAC
GCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAG
GCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACC
AGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTC
CCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGG
CTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGAC
ACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTG
AAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAA
AAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCA
GAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGG
ATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAG
TATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTT
CATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATG
ATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGG
TCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTT
TGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCC
CAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAG
AAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGAT
GCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCG
TCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACT
CTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTT
TCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGA
ATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATG
TATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC

FIG. 32 (Cont.)

(SEQ ID NO: 65)

Nucleotide sequence of TREX2-encoding plasmid (pExodus CMV.Trex2)

gacggatcgggagatctcccgatcccctatggtgcactctcagtacaatctgctctgatgccgcatagtt
aagccagtatctgctccctgcttgtgtgttggaggtcgctgagtagtgcgcgagcaaaatttaagctaca
acaaggcaaggcttgaccgacaattgcatgaagaatctgcttagggttaggcgttttgcgctgcttcgcg
atgtacgggccagatatacgcgttgacattgattattgactagttattaatagtaatcaattacggggtc
attagttcatagcccatatatggagttccgcgttacataacttacggtaaatggcccgcctggctgaccg
cccaacgacccccgcccattgacgtcaataatgacgtatgttcccatagtaacgccaatagggactttcc
attgacgtcaatgggtggagtatttacggtaaactgcccacttggcagtacatcaagtgtatcatatgcc
aagtacgccccctattgacgtcaatgacggtaaatggcccgcctggcattatgcccagtacatgacctta
tgggactttcctacttggcagtacatctacgtattagtcatcgctattaccatggtgatgcggttttggc
agtacatcaatgggcgtggatagcggtttgactcacggggatttccaagtctccaccccattgacgtcaa
tgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgtaacaactccgccccattgacg
caaatgggcggtaggcgtgtacggtgggaggtctatataagcagagctctctggctaactagagaaccca
ctgcttactggcttatcgaaattaatacgactcactatagggagacccaagctggctagcatgtctgagc
cacctcgggctgagacctttgtattcctggacctagaagccactgggctcccaaacatggaccctgagat
tgcagagatatccctttttgctgttcaccgctcttccctggagaacccagaacgggatgattctggttcc
ttggtgctgccccgtgttctggacaagctcacactgtgcatgtgcccggagcgcccctttactgccaagg
ccagtgagattactggtttgagcagcgaaagcctgatgcactgcgggaaggctggtttcaatggcgctgt
ggtaaggacactgcagggcttcctaagccgccaggagggccccatctgccttgtggcccacaatggcttc
gattatgacttcccactgctgtgcacggagctacaacgtctgggtgcccatctgccccaagacactgtct
gcctggacacactgcctgcattgcggggcctggaccgtgctcacagccacggcaccagggctcaaggccg
caaaagctacagcctggccagtctcttccaccgctacttccaggctgaacccagtgctgcccattcagca
gaaggtgatgtgcacaccctgcttctgatcttcctgcatcgtgctcctgagctgctcgcctgggcagatg

FIG. 33

agcaggcccgcagctgggctcatattgagcccatgtacgtgccacctgatggtccaagcctcgaagcctg
aattctgcagatatccagcacagtggcggccgctcgagtctagagggcccgtttaaacccgctgatcagc
ctcgactgtgccttctagttgccagccatctgttgtttgcccctcccccgtgccttccttgaccctggaa
ggtgccactcccactgtcctttcctaataaaatgaggaaattgcatcgcattgtctgagtaggtgtcatt
ctattctgggggtggggtggggcaggacagcaaggggaggattgggaagacaatagcaggcatgctgg
ggatgcggtgggctctatggcttctgaggcggaaagaaccagctggggctctagggggtatccccacgcg
ccctgtagcggcgcattaagcgcggcgggtgtggtggttacgcgcagcgtgaccgctacacttgccagcg
ccctagcgcccgctcctttcgctttcttcccttcctttctcgccacgttcgccggctttccccgtcaagc
tctaaatcgggggctcccttagggttccgatttagtgctttacggcacctcgaccccaaaaaacttgat
tagggtgatggttcacgtagtgggccatcgccctgatagacggtttttcgcccttgacgttggagtcca
cgttctttaatagtggactcttgttccaaactggaacaacactcaaccctatctcggtctattcttttga
tttataagggattttgccgatttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcg
aattaattctgtggaatgtgtgtcagttagggtgtggaaagtccccaggctccccagcaggcagaagtat
gcaaagcctatcaggacatagcgttggctacccgtgatattgctgaagagcttggcggcgaatgggctga
ccgcttcctcgtgctttacggtatcgccgctcccgattcgcagcgcatcgccttctatcgccttcttgac
gagttcttctgagcgggactctggggttcgaaatgaccgaccaagcgacgcccaacctgccatcacgaga
tttcgattccaccgccgccttctatgaaaggttgggcttcggaatcgttttccgggacgccggctggatg
atcctccagcgcggggatctcatgctggagttcttcgcccaccccaacttgtttattgcagcttataatg
gttacaaataaagcaatagcatcacaaatttcacaaataaagcatttttttcactgcattctagttgtgg
tttgtccaaactcatcaatgtatcttatcatgtctgtataccgtcgacctctagctagagcttggcgtaa
tcatggtcatagctgtttcctgtgtgaaattgttatccgctcacaattccacacaacatacgagccggaa
gcataaagtgtaaagcctggggtgcctaatgagtgagctaactcacattaattgcgttgcgctcactgcc
cgctttccagtcgggaaacctgtcgtgccagctgcattaatgaatcggccaacgcgcggggagaggcggt
ttgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgag
cggtatcagctcactcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacat
gtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgttttccataggctc
cgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaa

FIG. 33 (Cont.)

gataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggata
cctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcg
gtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccgaccgctgcgccttat
ccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaa
caggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctac
actagaagaacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagct
cttgatccggcaaacaaaccaccgctggtagcggtttttttgtttgcaagcagcagattacgcgcagaaa
aaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgt
taagggattttggtcatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagtt
ttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaatgcttaatcagtgaggcacc
tatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgata
cgggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcaccggctccagatt
tatcagcaataaaccagccagccggaagggccgagcgcagaagtggtcctgcaactttatccgcctccat
ccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttgtt
gccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaac
gatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgt
tgtcagaagtaagttggccgcagtgttatcactcatggttatggcagcactgcataattctcttactgtc
atgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgagaatagtgtatgc
ggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagt
gctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcg
atgtaacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaa
aaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcatactctt
cctttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatt
tagaaaaataaacaaataggggttccgcgcacatttccccgaaaagtgccacctgacgtc

FIG. 33 (Cont.)

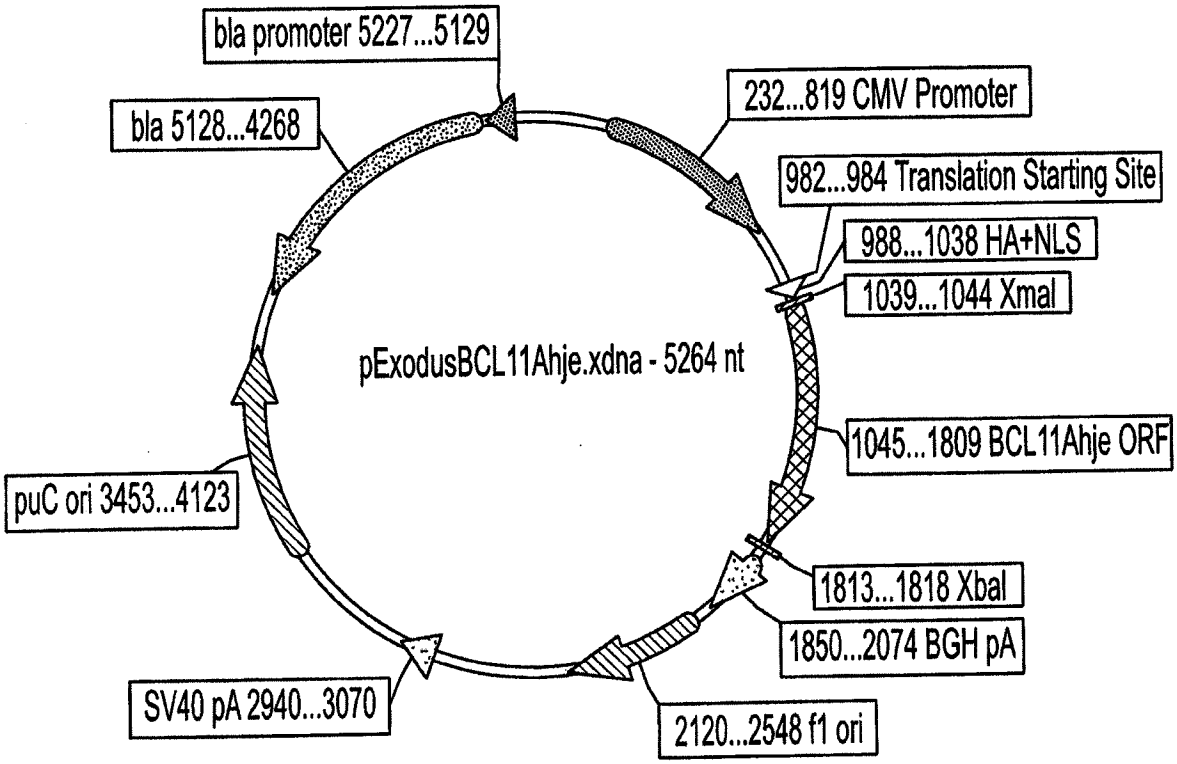

bla promoter 5227...5129
232...819 CMV Promoter
bla 5128...4268
982...984 Translation Starting Site
988...1038 HA+NLS
1039...1044 XmaI
pExodusBCL11Ahje.xdna - 5264 nt
1045...1809 BCL11Ahje ORF
puC ori 3453...4123
1813...1818 XbaI
1850...2074 BGH pA
SV40 pA 2940...3070
2120...2548 f1 ori

FIG. 34

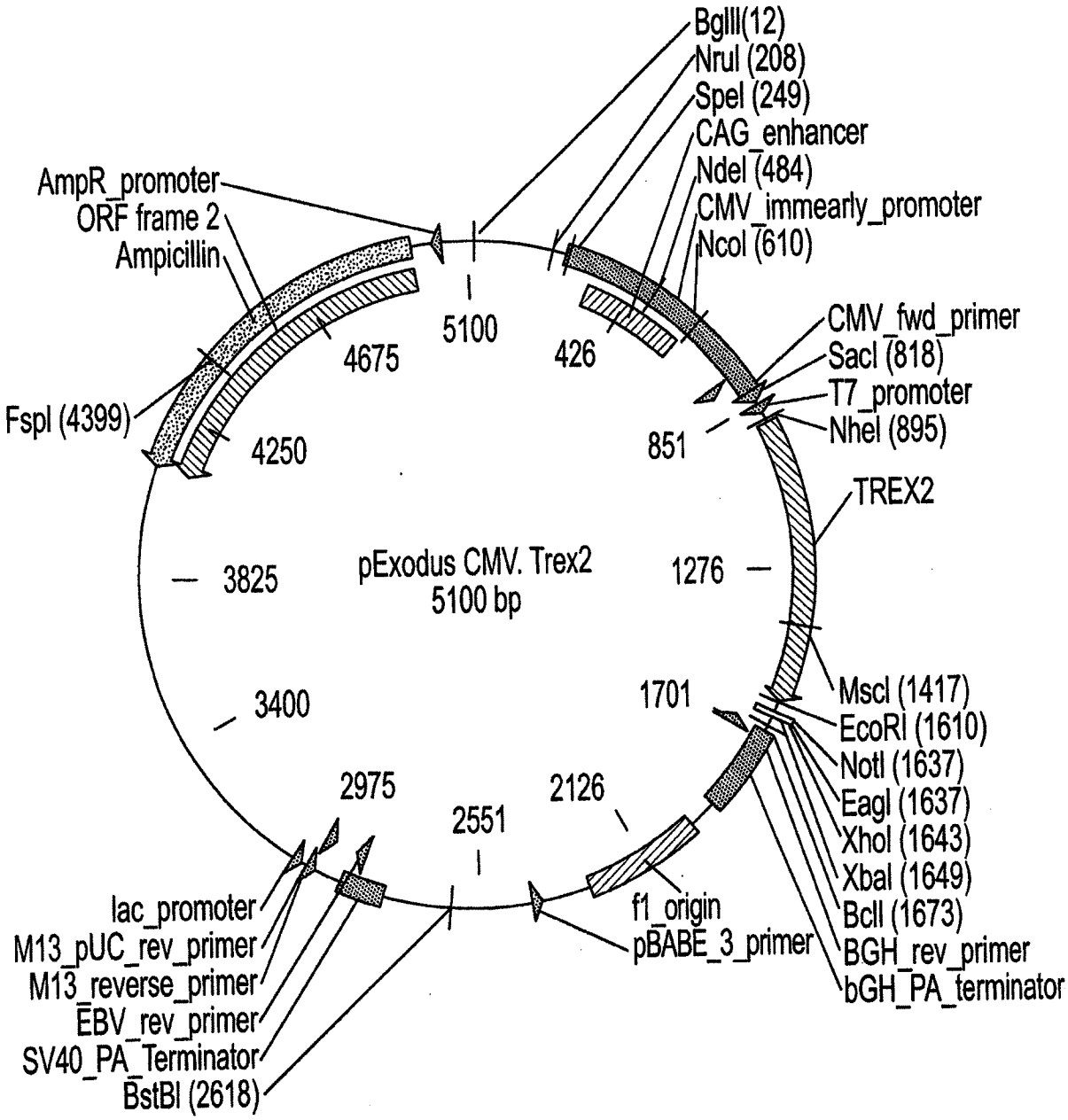

BglII(12)
NruI (208)
SpeI (249)
CAG_enhancer
NdeI (484)
CMV_immearly_promoter
NcoI (610)
CMV_fwd_primer
SacI (818)
T7_promoter
NheI (895)
TREX2
MscI (1417)
EcoRI (1610)
NotI (1637)
EagI (1637)
XhoI (1643)
XbaI (1649)
BclI (1673)
BGH_rev_primer
bGH_PA_terminator
f1_origin
pBABE_3_primer
lac_promoter
M13_pUC_rev_primer
M13_reverse_primer
EBV_rev_primer
SV40_PA_Terminator
BstBI (2618)
AmpR_promoter
ORF frame 2
Ampicillin
FspI (4399)
5100
426
851
1276
1701
2126
2551
2975
3400
3825
4250
4675
pExodus CMV. Trex2
5100 bp

FIG. 35

SEQUENCE LISTING

<110> Fred Hutchinson Cancer Research Center

<120> COMPOSITIONS AND METHODS FOR THE TREATMENT OF HEMOGLOBINOPATHIES

<130> SG4919A

<140> TBD
<141> 2013-02-22

<150> US 61/603,231
<151> 2012-02-24

<160> 69

<170> PatentIn version 3.5

<210> 1
<211> 3603
<212> DNA
<213> Homo sapiens

<400> 1
ccagtgagca ggttggttta agataagcag ggtttcatta gtttgtgaga atgaaaaatg 60

aaccttcatt ccactattcc cttaacttgc cctgagattg gctgttctgt catgtgtgtc 120

ttgactcaga aaccctgttc tcctctacat atctccccac cgcatctctt tcagcagttg 180

tttctaaaaa tatcctccta gtttcatttt tgcagaagtg ttttaggcta atatagtgga 240

atgtatctta gagtttaact tatttgtttc tgtcacttta tactaagaaa acttatctaa 300

aagcagatgt tttaacaagt tgactcaata taaagttctt ctttgcctct agagattttt 360

gtctccaagg gaattttgag aggttggaat ggacaaatct attgctgcag tttaaacttg 420

cttgcttcct ccttcttttg gtaaattctt cctataataa aactctaatt ttttattata 480

ttgaaataaa tatccattaa aagaatattt aaaaaatgaa tagtgtttat ttaccagtta 540

ttgaaatagg ttctggaaac atgaatttta aggttaacat tttaatgaca gataaaatca 600

aatattatat acaaatattt tgaatgttta aaattatggt atgactaaag aaagaatgca 660

aagtgaaaag tagatttacc atattcagcc agattaaatt taacgaagtt cctgggaata 720

tgctagtaca gaacattttt acagatgtgt tcttaaaaaa aaatgtggaa ttagacccag 780

gaatgaagat cccagtagtt tttcactctt ttctgaattc aaataatgcc acaatggcag 840

acaaatacac acccatgagc atatccaaaa ggaaggattg aaggaaagag gaggaagaaa 900

tggagaaagg aaggaaggaa gaggggaaga gagaggatgg aagggatgga ggagaagaag 960

gaaaaataaa taatggagag gagaggagaa aaaaggaggg gagaggagag gagaagggat 1020

agggaagaga aagagaaagg gaagggaaga gaggaaagaa gagaagagga gagaaaagaa 1080

acgaagagag gggaagggaa ggaaaaaaaa gaggaaaaaa gagacaagag aagagataag 1140

actgacagtt caaattttgg tggtgatatg gatcaataga aactcaaact ctgttggtga 1200

cactgtacaa tagtataacc cctttggaaa acctttaata gtatccacaa atgctggatg 1260

cttgataagt ctattaccta gcaattacat ttttagatat tcagaaacac atgcatgtgt 1320
gtatccaaag acatgtatag aaatgcttat gacagcaata atcataaaaa cctcaaaccg 1380
gtagccactt aaatgcttac caacagtaga attgataaat tacggtatag tcaaagaata 1440
gaatattaca cagaaatgaa aagaatcaac tactgcttaa cacgtagcga tacaaatgca 1500
ttttacagca tttggttgat taaaagtaac cagaggtgag ttcaaactat atgactttat 1560
ttgtatatag aaagatggat gatgtgcctg agattctgat cacaagggga aatgttataa 1620
aatagggtag agaggagcca tgaatgacct ttaaactttg ttacaagtta tttttctgta 1680
acctggaagc caacgaaaga tattgaataa ttcaagaaag gtggtggcat ggtttgattt 1740
gtgtctttaa aagattattc tcacttagtg aagaaatgta ttttagaagt agagaaaatg 1800
ggagacaaat agctgggctt ctgttgcagt agggaagaaa gtgacaatgc catttctatt 1860
atcagacttg gaccatgacg gtgatgtcag tcgtgaacac aagaataggg ccacatttgt 1920
gagtttagtg gtacgataaa atcagaaata cagtcttgga tacattgtat tgtatgcact 1980
cttgtaaaat gcaaaaagat gtacttagat atgtggatct ggagctcaga aagaatacaa 2040
ccaggtcaag aatacagaat ggaacagaac atacaagaac agatcataat gtgctgtgtg 2100
aatcactacc actacctgtt aaaaatgaca gatgatgtac ttcatcaata tctccttaaa 2160
atcttagaat gtgtttgtga gggaggaatt atgtttccaa ttcatatata agaaaattga 2220
ttctaaaaaa aatgttaggt aaattcttaa ggccatgagg actgttattt gatctttgtc 2280
tgttaattcc aaagacttgg cttttcactt taattctgtt ctacctgaaa tgattttaca 2340
cattgggaga tctggttaca tgtttattct atatggattg cattgagagg atttgtataa 2400
cagaataagg tctttttttc ttttctcttc tgagatggag tttcatccct attgcccaag 2460
ctagagtgca atggtgcaat ctaggctcac cgcaacctct gcctcctggg ttcaagcaat 2520
tctcctgcct cagccacctg aatagctggg actgcaggca tgcaccacac gcccggctga 2580
ttttgtattt ttagtagaga tggggtttca ccatgttggt caggctggtc ttgaactcct 2640
gacctcaagt gatctgcctg ccttggcctc ccaaagtgct gggtttacaa gcctgagcca 2700
ccgcatccag ccaggataag gtctaaaagt ggaaagaata gcatctactc ttgttcagga 2760
aacaatgagg acctgactgg gcagtaagag tggtgattaa tagataggga caaattgaag 2820
cagaatcgaa ctgttgatta gaggtaggga aatgatttta atctgtgacc ttggtgaatg 2880
ggcaagtagc tatctaatga ctaaaatgga aaacactgga agagaaacag ttttagtata 2940
acaagtgaaa tacccatgct gagtctgagg tgcctatagg acatctatat aaataagccc 3000
agtacattgt ttgatatatg ggtttggcac tgaggttgga ggtcagaggt tagaaatcag 3060
agttgggaat tgggattata caggctgtat ttaagagttt agatataact gtgaatccaa 3120
gagtgtgatg aatacaaagt taaatgaagg acctttaatg aacaccaaca tttaatgtga 3180
aatctcaagg aagtatgaag taagacatag tccccaaaat ccccgatgat tttagaactc 3240
agtatcgatt ttaattagtg taatgccaat gtgggttaga atggaagtca acttgctgtt 3300

ggtttcagag caggtaggag ataaggttct agattttgac acagtgaaaa gctgaaacaa 3360

aaaggaaaag gtagggtgaa agatgggaaa tgtatgtaag gaggatgagc cacatggtat 3420

gggaggtata ctaaggactc tagggtcaga gaaatatggg ttatatcctt ctacaaaatt 3480

cacattcttg gctgggtgtg gtggctcacg cctgtgatcc cagcactttc agaggccgag 3540

gagggtggat cacctgatgt taggagttcg agatcagcct gaccaacatg gtgaaacccc 3600

cta 3603

<210> 2
<211> 350
<212> DNA
<213> Homo sapiens

<400> 2
aaagatggat gatgtgcctg agattctgat cacaagggga aatgttataa aatagggtag 60

agaggagcca tgaatgacct ttaaactttg ttacaagtta tttttctgta acctggaagc 120

caacgaaaga tattgaataa ttcaagaaag gtggtggcat ggtttgattt gtgtctttaa 180

aagattattc tcacttagtg aagaaatgta ttttagaagt agagaaaatg ggagacaaat 240

agctgggctt ctgttgcagt agggaagaaa gtgacaatgc catttctatt atcagacttg 300

gaccatgacg gtgatgtcag tcgtgaacac aagaataggg ccacatttgt 350

<210> 3
<211> 21
<212> DNA
<213> Homo sapiens

<400> 3
gtgcacctca ctccagagga g 21

<210> 4
<211> 21
<212> DNA
<213> Homo sapiens

<400> 4
gtgcacctca ctccagtgga g 21

<210> 5
<211> 22
<212> DNA
<213> Homo sapiens

<400> 5
tttccactta ttcaaccttt ta 22

<210> 6
<211> 22
<212> DNA
<213> Homo sapiens

<400> 6
tgtggcccta ttcttgtgtt ca 22

<210> 7

<211> 22
<212> DNA
<213> Homo sapiens

<400> 7
cattgtcact ttcttcccta ct 22

<210> 8
<211> 22
<212> DNA
<213> Homo sapiens

<400> 8
taaaatacat ttcttcacta ag 22

<210> 9
<211> 22
<212> DNA
<213> Homo sapiens

<400> 9
actaagtgag aataatcttt ta 22

<210> 10
<211> 22
<212> DNA
<213> Homo sapiens

<400> 10
gccaccacct ttcttgaatt at 22

<210> 11
<211> 22
<212> DNA
<213> Homo sapiens

<400> 11
tcttgaatta ttcaatatct tt 22

<210> 12
<211> 22
<212> DNA
<213> Homo sapiens

<400> 12
ttaaaggtca ttcatggctc ct 22

<210> 13
<211> 2606
<212> DNA
<213> Homo sapiens

<400> 13
gcaatgaaaa taaatgtttt ttattaggca gaatccagat gctcaaggcc cttcataata 60

tcccccagtt tagtagttgg acttagggaa caaaggaacc tttaatagaa attggacagc 120

aagaaagcga gcttagtgat acttgtgggc cagggcatta gccacaccag ccaccacttt 180

ctgataggca gcctgcactg gtggggtgaa ttctttgcca aagtgatggg ccagcacaca 240

gaccagcacg ttgcccagga gctgtgggag gaagataaga ggtatgaaca tgattagcaa 300

aagggcctag cttggactca gaataatcca gccttatccc aaccataaaa taaaagcaga 360
atggtagctg gattgtagct gctattagca atatgaaacc tcttacatca gttacaattt 420
atatgcagaa atatttatat gcagagatat tgctattgcc ttaacccaga aattatcact 480
gttattcttt agaatggtgc aaagaggcat gatacattgt atcattattg ccctgaaaga 540
aagagattag ggaaagtatt agaaataaga taaacaaaaa agtatattaa aagaagaaag 600
catttttaa aattacaaat gcaaaattac cctgatttgg tcaatatgtg tacacatatt 660
aaaacattac actttaaccc ataaatatgt ataatgatta tgtatcaatt aaaaataaaa 720
gaaaataaag tagggagatt atgaatatgc aaataagcac acatatattc caaatagtaa 780
tgtactaggc agactgtgta aagtttttt ttaagttact taatgtatct cagagatatt 840
tccttttgtt atacacaatg ttaaggcatt aagtataata gtaaaaattg cggagaagaa 900
aaaaaaagaa agcaagaatt aaacaaaaga aaacaattgt tatgaacagc aaataaaaga 960
aactaaaacg atcctgagac ttccacactg atgcaatcat tcgtctgttt cccattctaa 1020
actgtaccct gttacttatc cccttcctat gacatgaact taaccataga aaagaagggg 1080
aaagaaaaca tcaagcgtcc catagactca ccctgaagtt ctcaggatcc acgtgcagct 1140
tgtcacagtg cagctcactc agtgtggcaa aggtgccctt gaggttgtcc aggtgagcca 1200
ggccatcact aaaggcaccg agcactttct tgccatgagc cttcacctta gggttgccca 1260
taacagcatc aggagtggac agatccccaa aggactcaaa gaacctctgg gtccaagggt 1320
agaccaccag cagcctaagg gtgggaaaat agaccaatag gcagagagag tcagtgccta 1380
tcagaaaccc aagagtcttc tctgtctcca catgcccagt ttctattggt ctccttaaac 1440
ctgtcttgta accttgatac caacctgccc agggcctcac caccaacttc atccacgttc 1500
accttgcccc acagggcagt aacggcagac ttctcctcag gagtcagatg caccatggtg 1560
tctgtttgag gttgctagtg aacacagttg tgtcagaagc aaatgtaagc aatagatggc 1620
tctgccctga cttttatgcc cagccctggc tcctgccctc cctgctcctg ggagtagatt 1680
ggccaaccct agggtgtggc tccacagggt gaggtctaag tgatgacagc cgtacctgtc 1740
cttggctctt ctggcactgg cttaggagtt ggacttcaaa ccctcagccc tccctctaag 1800
atatatctct tggccccata ccatcagtac aaattgctac taaaaacatc ctcctttgca 1860
agtgtattta cgtaatattt ggaatcacag cttggtaagc atattgaaga tcgttttccc 1920
aattttctta ttacacaaat aagaagttga tgcactaaaa gtggaagagt tttgtctacc 1980
ataattcagc tttgggatat gtagatggat ctcttcctgc gtctccagaa tatgcaaaat 2040
acttacagga cagaatggat gaaaactcta cctcggttct aagcatatct tctccttatt 2100
tggattaaaa ccttctggta agaaaagaaa aaatatatat atatatgtgt gtatatatac 2160
acacatacat atacatatat atgcattcat ttgttgttgt ttttcttaat ttgctcatgc 2220
atgctaataa attatgtcta aaaatagaat aaatacaaat caatgtgctc tgtgcattag 2280
ttacttatta ggttttggga aacaagagat aaaaaactag agacctctta atgcagtcaa 2340
aaatacaaat aaataaaaag tcacttacaa cccaaagtgt gactatcaat ggggtaatca 2400

gtggtgtcaa ataggaggtt aactggggac atctaactgt ttctgcctgg actaatctgc 2460

aagagtgtct gggggaacaa aaagcctctg tgacttagaa agtaggggta ggaggggaaa 2520

aggtcttcta cttggctcag attatttttt tcctctagtc cactaagaat actgcgtttt 2580

aaaatcattt ccttgattca agttcc 2606

<210> 14
<211> 606
<212> DNA
<213> Homo sapiens

<400> 14
caccctgaag ttctcaggat ccacgtgcag cttgtcacag tgcagctcac tcagtgtggc 60

aaaggtgccc ttgaggttgt ccaggtgagc caggccatca ctaaaggcac cgagcacttt 120

cttgccatga gccttcacct tagggttgcc cataacagca tcaggagtgg acagatcccc 180

aaaggactca aagaacctct gggtccaagg gtagaccacc agcagcctaa gggtgggaaa 240

atagaccaat aggcagagag agtcagtgcc tatcagaaac ccaagagtct tctctgtctc 300

cacatgccca gtttctattg gtctccttaa acctgtcttg taaccttgat accaacctgc 360

ccagggcctc accaccaact tcatccacgt tcaccttgcc ccacagggca gtaacggcag 420

acttctcctc aggagtcaga tgcaccatgg tgtctgtttg aggttgctag tgaacacagt 480

tgtgtcagaa gcaaatgtaa gcaatagatg gctctgccct gacttttatg cccagccctg 540

gctcctgccc tccctgctcc tgggagtaga ttggccaacc ctagggtgtg gctccacagg 600

gtgagg 606

<210> 15
<211> 303
<212> PRT
<213> Ophiostoma novo-ulmi

<400> 15

Ser Ala Tyr Met Ser Arg Arg Glu Ser Ile Asn Pro Trp Ile Leu Thr
1 5 10 15

Gly Phe Ala Asp Ala Glu Gly Ser Phe Leu Leu Arg Ile Arg Asn Asn
20 25 30

Asn Lys Ser Ser Val Gly Tyr Ser Thr Glu Leu Gly Phe Gln Ile Thr
35 40 45

Leu His Asn Lys Asp Lys Ser Ile Leu Glu Asn Ile Gln Ser Thr Trp
50 55 60

Lys Val Gly Val Ile Ala Asn Ser Gly Asp Asn Ala Val Ser Leu Lys
65 70 75 80

Val Thr Arg Phe Glu Asp Leu Lys Val Ile Ile Asp His Phe Glu Lys
85 90 95

```
Tyr Pro Leu Ile Thr Gln Lys Leu Gly Asp Tyr Met Leu Phe Lys Gln
            100                 105                 110

Ala Phe Cys Val Met Glu Asn Lys Glu His Leu Lys Ile Asn Gly Ile
        115                 120                 125

Lys Glu Leu Val Arg Ile Lys Ala Lys Leu Asn Trp Gly Leu Thr Asp
    130                 135                 140

Glu Leu Lys Lys Ala Phe Pro Glu Ile Ile Ser Lys Glu Arg Ser Leu
145                 150                 155                 160

Ile Asn Lys Asn Ile Pro Asn Phe Lys Trp Leu Ala Gly Phe Thr Ser
                165                 170                 175

Gly Glu Gly Cys Phe Phe Val Asn Leu Ile Lys Ser Lys Ser Lys Leu
            180                 185                 190

Gly Val Gln Val Gln Leu Val Phe Ser Ile Thr Gln His Ile Lys Asp
        195                 200                 205

Lys Asn Leu Met Asn Ser Leu Ile Thr Tyr Leu Gly Cys Gly Tyr Ile
    210                 215                 220

Lys Glu Lys Asn Lys Ser Glu Phe Ser Trp Leu Asp Phe Val Val Thr
225                 230                 235                 240

Lys Phe Ser Asp Ile Asn Asp Lys Ile Ile Pro Val Phe Gln Glu Asn
                245                 250                 255

Thr Leu Ile Gly Val Lys Leu Glu Asp Phe Glu Asp Trp Cys Lys Val
            260                 265                 270

Ala Lys Leu Ile Glu Glu Lys Lys His Leu Thr Glu Ser Gly Leu Asp
        275                 280                 285

Glu Ile Lys Lys Ile Lys Leu Asn Met Asn Lys Gly Arg Val Phe
    290                 295                 300
```

<210> 16
<211> 110
<212> DNA
<213> Homo sapiens

<400> 16
tgggggcccc ttccccacac tatctcaatg caaatatctg tctgaaacgg tccctggcta 60

aactccaccc atgggttggc cagccttgcc ttgaccaata gccttgacaa 110

<210> 17
<211> 110
<212> DNA
<213> Homo sapiens

<400> 17
tgggggcgcc ttccccacac tatctcaatg caaatatctg tctgaaacgg tccctggcta 60

aactccaccc atgggttggc cagccttgcc ttgaccaata gccttgacaa 110

<210> 18
<211> 110
<212> DNA
<213> Homo sapiens

<400> 18
tgggggcccc ttccccacac tatctcaatg caaacatctg tctgaaacgg tccctggcta 60

aactccaccc atgggttggc cagccttgcc ttgaccaata gccttgacaa 110

<210> 19
<211> 110
<212> DNA
<213> Homo sapiens

<400> 19
tgggggcccc ttccccacac tatctcaatg caaatatctg tctgaaacgg tccctggcta 60

aactccaccc atgggttggc cagccttgcc ttgactaata gccttgacaa 110

<210> 20
<211> 110
<212> DNA
<213> Homo sapiens

<400> 20
tgggggcccc ttctccacac tatctcaatg caaatatctg tctgaaacgg tccctggcta 60

aactccaccc atgggttggc cagccttgcc ttgaccaata gccttgacaa 110

<210> 21
<211> 110
<212> DNA
<213> Homo sapiens

<400> 21
tgggggcccc ttccccacac tatctcaatg caaacatctg tctgaaacgg tccctggcta 60

aactccaccc atgggttggc cagccttgcc ttgaccaata gccttgacaa 110

<210> 22
<211> 110
<212> DNA
<213> Homo sapiens

<400> 22
tgggggcccc ttccccacac tatctcaatg caaatatctg tctgaaacgg tccctggcta 60

aactccaccc atgggttggc cagccttgcc ttaaccaata gccttgacaa 110

<210> 23
<211> 110
<212> DNA
<213> Homo sapiens

<400> 23
tgggggcccc ttccccacac tatctcaatg caaatatctg tctgaaacgg tccctggcta 60

aactccaccc atgggttggc cagccttgcc ttgaccaata gccttgacaa 110

<210> 24
<211> 2508
<212> DNA
<213> Homo sapiens

<400> 24
atgtctcgcc gcaagcaagg caaaccccag cacttaagca aacgggaatt ctcgcccgag 60

cctcttgaag ccattcttac agatgatgaa ccagaccacg gcccgttggg agctccagaa 120

ggggatcatg acctcctcac ctgtgggcag tgccagatga acttcccatt gggggacatt 180

cttattttta tcgagcacaa acggaaacaa tgcaatggca gcctctgctt agaaaaagct 240

gtggataagc caccttcccc ttcaccaatc gagatgaaaa aagcatccaa tcccgtggag 300

gttggcatcc aggtcacgcc agaggatgac gattgtttat caacgtcatc tagaggaatt 360

tgccccaaac aggaacacat agcagataaa cttctgcact ggaggggcct ctcctcccct 420

cgttctgcac atggagctct aatccccacg cctgggatga gtgcagaata tgccccgcag 480

ggtatttgta aagatgagcc cagcagctac acatgtacaa cttgcaaaca gccattcacc 540

agtgcatggt ttctcttgca acacgcacag aacactcatg gattaagaat ctacttagaa 600

agcgaacacg gaagtcccct gaccccgcgg gttggtatcc cttcaggact aggtgcagaa 660

tgtccttccc agccacctct ccatgggatt catattgcag acaataaccc ctttaacctg 720

ctaagaatac caggatcagt atcgagagag gcttccggcc tggcagaagg gcgctttcca 780

cccactcccc ccctgtttag tccaccaccg agacatcact tggaccccca ccgcatagag 840

cgcctggggg cggaagagat ggccctggcc acccatcacc cgagtgcctt tgacagggtg 900

ctgcggttga atccaatggc tatggagcct cccgccatgg atttctctag gagacttaga 960

gagctggcag ggaacacgtc tagcccaccg ctgtccccag gccggcccag ccctatgcaa 1020

aggttactgc aaccattcca gccaggtagc aagccgccct tcctggcgac gccccccctc 1080

cctcctctgc aatccgcccc tcctccctcc cagcccccgg tcaagtccaa gtcatgcgag 1140

ttctgcggca agacgttcaa atttcagagc aacctggtgg tgcaccggcg cagccacacg 1200

ggcgagaagc cctacaagtg caacctgtgc gaccacgcgt gcacccaggc cagcaagctg 1260

aagcgccaca tgaagacgca catgcacaaa tcgtccccca tgacggtcaa gtccgacgac 1320

ggtctctcca ccgccagctc cccggaaccc ggcaccagcg acttggtggg cagcgccagc 1380

agcgcgctca agtccgtggt ggccaagttc aagagcgaga acgaccccaa cctgatcccg 1440

gagaacgggg acgaggagga agaggaggac gacgaggaag aggaagaaga ggaggaagag 1500

gaggaggagg agctgacgga gagcgagagg gtggactacg gcttcgggct gagcctggag 1560

gcggcgcgcc accacgagaa cagctcgcgg ggcgcggtcg tgggcgtggg cgacgagagc 1620

cgcgccctgc ccgacgtcat gcagggcatg gtgctcagct ccatgcagca cttcagcgag 1680

gccttccacc aggtcctggg cgagaagcat aagcgcggcc acctggccga ggccgagggc 1740

cacagggaca cttgcgacga agactcggtg gccggcgagt cggaccgcat agacgatggc 1800

actgttaatg gccgcggctg ctccccgggc gagtcggcct cgggggcct gtccaaaaag 1860

ctgctgctgg gcagccccag ctcgctgagc cccttctcta agcgcatcaa gctcgagaag 1920

gagttcgacc tgcccccggc cgcgatgccc aacacggaga acgtgtactc gcagtggctc 1980

gccggctacg cggcctccag gcagctcaaa gatcccttcc ttagcttcgg agactccaga 2040

caatcgcctt ttgcctcctc gtcggagcac tcctcggaga acgggagttt gcgcttctcc 2100

acaccgcccg gggagctgga cggagggatc tcggggcgca gcggcacggg aagtggaggg 2160

agcacgcccc atattagtgg tccgggcccg ggcaggccca gctcaaaaga gggcagacgc 2220

agcgacactt gtgagtactg tgggaaagtc ttcaagaact gtagcaatct cactgtccac 2280

aggagaagcc acacgggcga aaggccttat aaatgcgagc tgtgcaacta tgcctgtgcc 2340

cagagtagca agctcaccag gcacatgaaa acgcatggcc aggtggggaa ggacgtttac 2400

aaatgtgaaa tttgtaagat gccttttagc gtgtacagta ccctggagaa acacatgaaa 2460

aaatggcaca gtgatcgagt gttgaataat gatataaaaa ctgaatag 2508

<210> 25
<211> 22
<212> DNA
<213> Homo sapiens

<400> 25
atgggattca tattgcagac aa 22

<210> 26
<211> 22
<212> DNA
<213> Homo sapiens

<400> 26
agccattgga ttcaaccgca gc 22

<210> 27
<211> 22
<212> DNA
<213> Homo sapiens

<400> 27
caaacagcca ttcaccagtg ca 22

<210> 28
<211> 771
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide Sequence of I-HjeMI, Codon Optimized for Expression in E. coli

<400> 28
atgggatccc acatggacct gacctacgct tacctggttg gtctgttcga aggtgacggt 60

tacttctcta tcaccaaaaa gggtaaatac ctgacctacg aactgggtat cgaactgtct 120

atcaaagacg ttcagctgat ctacaaaatc aaagacatcc tgggtgttgg taaagtttct 180

ttccgtaaac gtaacgaaat cgaaatggtt tctctgcgta tccgtgacaa gaatcacctg 240

aaaaacttca tcctgccgat cttcgacaaa tacccgatgc tgtctaacaa gcagtacgac 300

tacctgcgtt tcaaagacgc tctcctgtct aacattatct actctgacga tctgccggaa 360

tacgctcgtt ctaacgaatc tatcaactct gttgactcta ttatcaacac ctcttacttc 420

tctgcttggc tggttggttt catcgaagct gaaggttgct tctctaccta caaactgaac 480

aaagatgacg attacctgat cgcttctttc gacatcgctc agaaagacgg tgacatcctg 540

atctctgcta tccacaaata cctgtctttc accacgaaaa tctacctgga caaaaccaac 600

tgctctcgtc tgaaagtgac cggtgtacgt tctgttaaaa acgtggttaa attcatccag 660

ggtgctccgg ttaaactgct cggtaacaag aaactgcagt acaaactgtg gatcaaacag 720

ctgcgtaaaa tctctcgtta ctctgaaaaa atccagctgc cgtctaacta c 771

<210> 29
<211> 771
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide Sequence of I-HjeMI, Codon Optimized for Expression in Mammals

<400> 29
atgggcagcc acatggacct gacctacgcc tatctggtcg gcctgttcga gggcgacggc 60

tattttagca taaccaagaa gggcaagtat ctgacgtatg aactgggcat cgagctctcc 120

atcaaggacg tgcagctcat ctacaagatc aaggacatcc tcggcgtggg caaagtgtct 180

tttaggaaga ggaacgagat cgagatggtc agcctgcgaa tcagggacaa aaaccacctg 240

aagaacttca tcctgcccat cttcgacaag taccccatgc tgagcaacaa gcagtacgac 300

tatctccgat tcaaggatgc cctcctgtcc aacatcatct atagcgacga cctgcccgag 360

tacgccagga gcaacgagtc aatcaatagc gtggacagca tcatcaacac ctcatacttc 420

agcgcctggc tggttggctt catcgaggcc gagggctgct tcagcaccta caagctcaac 480

aaggacgacg attatttgat cgcgagcttc gatatagccc agaaggacgg cgacattctc 540

atctccgcga tccacaaata cctgagcttc acgaccaaaa tctacctgga caagaccaac 600

tgtagcaggc tcaaggtcac cggcgtgagg agcgtcaaga acgtggttaa gttcatccag 660

ggtgcgccgg tcaagttgct gggtaacaag aagctgcagt acaaactttg gataaagcag 720

ctgcgcaaga tctcccgata cagcgagaaa atccagctgc ccagtaacta c 771

<210> 30
<211> 254
<212> PRT
<213> Trichoderma reesei

<400> 30

Met Gly Asp Leu Thr Tyr Ala Tyr Leu Val Gly Leu Phe Glu Gly Asp
1 5 10 15

Gly Tyr Phe Ser Ile Thr Lys Lys Gly Lys Tyr Leu Thr Tyr Glu Leu

20 25 30

Gly Ile Glu Leu Ser Ile Lys Asp Val Gln Leu Ile Tyr Lys Ile Lys
35 40 45

Asp Ile Leu Gly Val Gly Lys Val Ser Phe Arg Lys Arg Asn Glu Ile
50 55 60

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

Ile Leu Pro Ile Phe Asp Lys Tyr Pro Met Leu Ser Asn Lys Gln Tyr
85 90 95

Asp Tyr Leu Arg Phe Lys Asp Ala Leu Leu Ser Asn Ile Ile Tyr Ser
100 105 110

Asp Asp Leu Pro Glu Tyr Ala Arg Ser Asn Glu Ser Ile Asn Ser Val
115 120 125

Asp Ser Ile Ile Asn Thr Ser Tyr Phe Ser Ala Trp Leu Val Gly Phe
130 135 140

Ile Glu Ala Glu Gly Cys Phe Ser Thr Tyr Lys Leu Asn Lys Asp Asp
145 150 155 160

Asp Tyr Leu Ile Ala Ser Phe Asp Ile Ala Gln Lys Asp Gly Asp Ile
165 170 175

Leu Ile Ser Ala Ile His Lys Tyr Leu Ser Phe Thr Thr Lys Ile Tyr
180 185 190

Leu Asp Lys Thr Asn Cys Ser Arg Leu Lys Val Thr Gly Val Arg Ser
195 200 205

Val Lys Asn Val Val Lys Phe Ile Gln Gly Ala Pro Val Lys Leu Leu
210 215 220

Gly Asn Lys Lys Leu Gln Tyr Lys Leu Trp Ile Lys Gln Leu Arg Lys
225 230 235 240

Ile Ser Arg Tyr Ser Glu Lys Ile Gln Leu Pro Ser Asn Tyr
245 250

<210> 31
<211> 768
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide Sequence for a BCL11A Gene Targeting Nuclease Based on the Homing Endonuclease I-HjeMI; Codon Optimized for Expression in E. coli and Obtained through Directed Evolution in IVC and in Bacteria

<400> 31
atgggatccc acatggacct gacctacgct tacctggttg gtctgttcga aggtgacggt 60

tacttcacta tcgctaaagc gggtaagtac ctgaattacg agctgggtat cacgctgtct 120

atcaaagacg ctcagctgat ctacaaaatc aaagacatcc tgggtgttgg taatgtttat 180

ttccggaaat ataggcagca tgaaatggtt tctctgcgga tccaggacaa gaatcacctg 240

aaaaacttca tcctgccgat cttcgacaaa tacccgatgc tgtctaacaa gcagtacgac 300

tacctgcgtt tcaaagacgc tctcctgtct aacattatct actctgacga tctgccggaa 360

tacgctcgtt ctaacgaatc tatcaactct gttgactcta ttatcaacac ctcttacttc 420

tctgcttggc tggttggttt catcgaagct gaaggttgct tcacgaccta caaagcgagt 480

aaagataagt acctgacggc tgggttcagt atcgctcaga aagacggtga catcctgatc 540

tctgctatcc acaaatacct gtctttcacc acgaaaccgt acaaggacaa aaccaactgc 600

tctcatctga aggtgaccgg tgtacgttct gttaacaacg tggttaaatt catccagggt 660

gctccggtta aactgctcgg taacaagaaa ctgcagtaca aactgtggat caaacagctg 720

cgtaaaatct ctcgttactc tgaaaaaatc cagctgccgt ctaactac 768

<210> 32
<211> 768
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide Sequence of a BCL11A Gene Targeting Nuclease Based on the Homing Endonuclease I-HjeMI; Codon Optimized for Expression in Mammals and Obtained through Directed Evolution in IVC and in Bacteria

<400> 32
atgggcagcc acatggacct gacctacgcc tatctggtcg gcctgttcga gggcgacggc 60

tattttacaa tagctaaggc cggcaagtat ctgaactacg agctgggcat cacactctcc 120

atcaaggacg ctcagctcat ctacaagatc aaggacatcc tcggcgtggg caacgtgtac 180

tttaggaagt acaggcaaca tgagatggtc agcctgcgaa tccaggacaa aaaccacctg 240

aagaacttca tcctgcccat cttcgacaag taccccatgc tgagcaacaa gcagtacgac 300

tacctgcgat tcaaggatgc cctcctgtcc aacatcatct atagcgacga cctgcccgag 360

tacgccagga gcaacgagtc aatcaatagc gtggacagca tcatcaacac ctcatacttc 420

agcgcctggc tggttggctt catcgaggcc gagggctgct tcaccaccta caaggcatcc 480

aaggacaagt acctgacagc gggcttctcc atagcccaga aggacggcga cattctcatc 540

tccgcgatcc acaaatacct gagcttcacg accaaaccct acaaagacaa gaccaactgt 600

agccacctca aggtcaccgg cgtgaggagc gtcaataacg tggttaagtt catccagggt 660

gcgccggtca agctgctggg taacaagaag ctgcagtaca aactttggat aaagcagctg 720

cgcaagatct cccgatacag cgagaaaatc agctgccca gtaactac 768

<210> 33

<211> 256
<212> PRT
<213> Artificial Sequence

<220>
<223> Amino Acid Sequence of a BCL11A Gene Targeting Nuclease Based on the Homing Endonuclease I-HjeMI

<400> 33

Met Gly Ser His Met Asp Leu Thr Tyr Ala Tyr Leu Val Gly Leu Phe
1 5 10 15

Glu Gly Asp Gly Tyr Phe Thr Ile Ala Lys Ala Gly Lys Tyr Leu Asn
20 25 30

Tyr Glu Leu Gly Ile Thr Leu Ser Ile Lys Asp Ala Gln Leu Ile Tyr
35 40 45

Lys Ile Lys Asp Ile Leu Gly Val Gly Asn Val Tyr Phe Arg Lys Tyr
50 55 60

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

Lys Asn Phe Ile Leu Pro Ile Phe Asp Lys Tyr Pro Met Leu Ser Asn
85 90 95

Lys Gln Tyr Asp Tyr Leu Arg Phe Lys Asp Ala Leu Leu Ser Asn Ile
100 105 110

Ile Tyr Ser Asp Asp Leu Pro Glu Tyr Ala Arg Ser Asn Glu Ser Ile
115 120 125

Asn Ser Val Asp Ser Ile Ile Asn Thr Ser Tyr Phe Ser Ala Trp Leu
130 135 140

Val Gly Phe Ile Glu Ala Glu Gly Cys Phe Thr Thr Tyr Lys Ala Ser
145 150 155 160

Lys Asp Lys Tyr Leu Thr Ala Gly Phe Ser Ile Ala Gln Lys Asp Gly
165 170 175

Asp Ile Leu Ile Ser Ala Ile His Lys Tyr Leu Ser Phe Thr Thr Lys
180 185 190

Pro Tyr Lys Asp Lys Thr Asn Cys Ser His Leu Lys Val Thr Gly Val
195 200 205

Arg Ser Val Asn Asn Val Val Lys Phe Ile Gln Gly Ala Pro Val Lys
210 215 220

Leu Leu Gly Asn Lys Lys Leu Gln Tyr Lys Leu Trp Ile Lys Gln Leu
225 230 235 240

Arg Lys Ile Ser Arg Tyr Ser Glu Lys Ile Gln Leu Pro Ser Asn Tyr
245 250 255

<210> 34
<211> 912
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide Sequence of I-OnuI, Codon Optimized for Expression in E. coli

<400> 34
atgtccgcct acatgtcccg tcgcgagtcc attaacccgt ggattctcac cggtttcgcc 60

gacgcggaag gctccttttt gctgcgcatc cgcaacaaca acaagtccag cgtcggctac 120

tccactgagc tcggcttcca aattacactt cataacaagg acaagagcat tcttgagaac 180

atccagtcaa catggaaggt gggcgtgatc gccaacagcg gtgacaacgc cgtgtcgctg 240

aaggtcacgc gttttgagga cctgaaggtc attatcgacc attttgaaaa atacccactg 300

attacgcaga agctcggtga ctacatgctg tttaagcagg cgttttgcgt catggagaac 360

aaggagcatt tgaagattaa tggtatcaag gagctggtgc gcattaaggc aaagctcaat 420

tggggtctga cggatgagct gaagaaggcc tttccggaga tcatctcgaa ggagcgctcc 480

ctcatcaaca agaacatccc taatttcaag tggctggcgg gttttacctc gggcgagggt 540

tgcttctttg ttaacctgat caagtcaaag tcgaagctag gtgtccaggt gcagctggtg 600

ttcagcatta cccaacacat caaggataag aacctcatga actctctgat tacctacttg 660

ggctgcggct acattaagga gaaaaacaag agtgagttct cctggcttga cttcgtcgtc 720

acgaaattct ccgacatcaa cgacaagatc attccggtct ttcaggaaaa cacgctcatc 780

ggcgtgaagc tcgaggactt cgaggattgg tgtaaggtcg ctaagctgat cgaggagaaa 840

aagcacctga cagaaagtgg cctggacgag atcaagaaga ttaagctgaa catgaacaag 900

ggcagagtat tc 912

<210> 35
<211> 1998
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide Sequence of MegaTAL:5.5 RVD + Y2 I-AniI

<400> 35
gtggatctac gcacgctcgg ctacagtcag cagcagcaag agaagatcaa accgaaggtg 60

cgttcgacag tggcgcagca ccacgaggca ctggtgggcc atgggtttac acacgcgcac 120

atcgttgcgc tcagccaaca cccggcagcg ttagggaccg tcgctgtcac gtatcagcac 180

ataatcacgg cgttgccaga ggcgacacac gaagacatcg ttggcgtcgg caaacagtgg 240

tccggcgcac gcgccctgga ggccttgctc acggatgcgg gggagttgag aggtccgccg 300

ttacagttgg acacaggcca acttgtgaag attgcaaaac gtggcggcgt gaccgcaatg 360

gaggcagtgc atgcatcgcg caatgcactg acgggtgccc ccctgaacct gaccccggac 420
caagtggtgg ctatcgccag caacaatggc ggcaagcaag cgctcgaaac ggtgcagcgg 480
ctgttgccgg tgctgtgcca ggaccatggc ctgactccgg accaagtggt ggctatcgcc 540
agccacgatg gcggcaagca agcgctcgaa acggtgcagc ggctgttgcc ggtgctgtgc 600
caggaccatg gcctgacccc ggaccaagtg gtggctatcg ccagcaacat tggcggcaag 660
caagcgctcg aaacggtgca gcggctgttg ccggtgctgt gccaggacca tggcctgacc 720
ccggaccaag tggtggctat cgccagcaac aatggcggca agcaagcgct cgaaacggtg 780
cagcggctgt tgccggtgct gtgccaggac catggcctga ctccggacca agtggtggct 840
atcgccagcc acgatggcgg caagcaagcg ctcgaaacgg tgcagcggct gttgccggtg 900
ctgtgccagg accatggcct gaccccggac caagtggtgg ctatcgccag caacggtggc 960
ggcaagcaag cgctcgaaag cattgtggcc cagctgagcc ggcctgatcc ggcgttggcc 1020
gcgttgacca acgaccacct cgtcgccttg gcctgcctcg gcggacgtcc tgccatggat 1080
gcagtgaaaa agggattgcc gcacgcgccg gaattgatca gaagagtcaa tcgccgtatt 1140
ggcgaacgca cgtcccatcg cgttgcgata tctagagtgg gaggaagcga tcttacgtac 1200
gcgtatttag ttggtctcta cgaaggggat ggatacttta gtatcaccaa gaaaggcaag 1260
tacttgactt atgaattagg tattgagctg agcatcaaag acgtccaatt gatttacaag 1320
atcaagaaaa tcctaggtat tggcatcgta agcttcagga agagaaacga gattgaaatg 1380
gttgcattga ggatccgtga taagaatcat ctaaaatcta agatattgcc tatatttgag 1440
aagtatccaa tgttttccaa caaacagtac gactatttaa gattcaggaa tgcattgtta 1500
tctggcatta tatacctaga agacttgcct gattacacta gaagtgacga accattgaat 1560
tctatagaat ccattatcaa cacatcatac ttctccgcct ggctagttgg atttatagaa 1620
gctgagggct gtttcagtgt gtacaagctg aacaaagacg atgactactt gattgcttca 1680
ttcgacattg cccaaagaga tggtgatatc ttgatttcag caattaggaa gtacttaagt 1740
ttcactacta aggtttacct agacaagact aattgtagca aattgaaggt cactagtgtt 1800
agatccgtcg agaacatcat taagtttctg cagaatgctc ctgtcaaatt gttaggcaac 1860
aagaaactgc aatacaagtt gtggttgaaa caactaagga agatttctag gtattccgag 1920
aagatcaaga ttccatcaaa ctacgtcgac cgagcatctt accgccattt atacccatat 1980
ttgttctgtt tttcttga 1998

<210> 36
<211> 665
<212> PRT
<213> Artificial Sequence

<220>
<223> Amino Acid Sequence of MegaTAL:5.5 RVD + Y2 I-AniI

<400> 36

Val Asp Leu Arg Thr Leu Gly Tyr Ser Gln Gln Gln Gln Glu Lys Ile
1 5 10 15

```
Lys Pro Lys Val Arg Ser Thr Val Ala Gln His His Glu Ala Leu Val
            20                  25                  30

Gly His Gly Phe Thr His Ala His Ile Val Ala Leu Ser Gln His Pro
        35                  40                  45

Ala Ala Leu Gly Thr Val Ala Val Thr Tyr Gln His Ile Ile Thr Ala
    50                  55                  60

Leu Pro Glu Ala Thr His Glu Asp Ile Val Gly Val Gly Lys Gln Trp
65                  70                  75                  80

Ser Gly Ala Arg Ala Leu Glu Ala Leu Leu Thr Asp Ala Gly Glu Leu
                85                  90                  95

Arg Gly Pro Pro Leu Gln Leu Asp Thr Gly Gln Leu Val Lys Ile Ala
            100                 105                 110

Lys Arg Gly Gly Val Thr Ala Met Glu Ala Val His Ala Ser Arg Asn
        115                 120                 125

Ala Leu Thr Gly Ala Pro Leu Asn Leu Thr Pro Asp Gln Val Val Ala
    130                 135                 140

Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala Leu Glu Thr Val Gln Arg
145                 150                 155                 160

Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp Gln Val
                165                 170                 175

Val Ala Ile Ala Ser His Asp Gly Gly Lys Gln Ala Leu Glu Thr Val
            180                 185                 190

Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr Pro Asp
        195                 200                 205

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

Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly Leu Thr
225                 230                 235                 240

Pro Asp Gln Val Val Ala Ile Ala Ser Asn Asn Gly Gly Lys Gln Ala
                245                 250                 255

Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp His Gly
            260                 265                 270

Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser His Asp Gly Gly Lys
        275                 280                 285
```

```
Gln Ala Leu Glu Thr Val Gln Arg Leu Leu Pro Val Leu Cys Gln Asp
    290                 295                 300

His Gly Leu Thr Pro Asp Gln Val Val Ala Ile Ala Ser Asn Gly Gly
305                 310                 315                 320

Gly Lys Gln Ala Leu Glu Ser Ile Val Ala Gln Leu Ser Arg Pro Asp
                325                 330                 335

Pro Ala Leu Ala Ala Leu Thr Asn Asp His Leu Val Ala Leu Ala Cys
            340                 345                 350

Leu Gly Gly Arg Pro Ala Met Asp Ala Val Lys Lys Gly Leu Pro His
        355                 360                 365

Ala Pro Glu Leu Ile Arg Arg Val Asn Arg Arg Ile Gly Glu Arg Thr
    370                 375                 380

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

Ala Tyr Leu Val Gly Leu Tyr Glu Gly Asp Gly Tyr Phe Ser Ile Thr
                405                 410                 415

Lys Lys Gly Lys Tyr Leu Thr Tyr Glu Leu Gly Ile Glu Leu Ser Ile
            420                 425                 430

Lys Asp Val Gln Leu Ile Tyr Lys Ile Lys Lys Ile Leu Gly Ile Gly
        435                 440                 445

Ile Val Ser Phe Arg Lys Arg Asn Glu Ile Glu Met Val Ala Leu Arg
    450                 455                 460

Ile Arg Asp Lys Asn His Leu Lys Ser Lys Ile Leu Pro Ile Phe Glu
465                 470                 475                 480

Lys Tyr Pro Met Phe Ser Asn Lys Gln Tyr Asp Tyr Leu Arg Phe Arg
                485                 490                 495

Asn Ala Leu Leu Ser Gly Ile Ile Tyr Leu Glu Asp Leu Pro Asp Tyr
            500                 505                 510

Thr Arg Ser Asp Glu Pro Leu Asn Ser Ile Glu Ser Ile Ile Asn Thr
        515                 520                 525

Ser Tyr Phe Ser Ala Trp Leu Val Gly Phe Ile Glu Ala Glu Gly Cys
    530                 535                 540

Phe Ser Val Tyr Lys Leu Asn Lys Asp Asp Asp Tyr Leu Ile Ala Ser
545                 550                 555                 560
```

```
Phe Asp Ile Ala Gln Arg Asp Gly Asp Ile Leu Ile Ser Ala Ile Arg
                565                 570                 575

Lys Tyr Leu Ser Phe Thr Thr Lys Val Tyr Leu Asp Lys Thr Asn Cys
            580                 585                 590

Ser Lys Leu Lys Val Thr Ser Val Arg Ser Val Glu Asn Ile Ile Lys
        595                 600                 605

Phe Leu Gln Asn Ala Pro Val Lys Leu Leu Gly Asn Lys Lys Leu Gln
    610                 615                 620

Tyr Lys Leu Trp Leu Lys Gln Leu Arg Lys Ile Ser Arg Tyr Ser Glu
625                 630                 635                 640

Lys Ile Lys Ile Pro Ser Asn Tyr Val Asp Arg Ala Ser Tyr Arg His
                645                 650                 655

Leu Tyr Pro Tyr Leu Phe Cys Phe Ser
            660                 665
```

<210> 37
<211> 4146
<212> DNA
<213> Artificial Sequence

<220>
<223> Engineered type II CRISPR system for human cells; expression format and full sequence of cas9 gene insert of Mali et al.

<400> 37

```
gccaccatgg acaagaagta ctccattggg ctcgatatcg gcacaaacag cgtcggctgg       60

gccgtcatta cggacgagta caaggtgccg agcaaaaaat tcaaagttct gggcaatacc      120

gatcgccaca gcataaagaa gaacctcatt ggcgccctcc tgttcgactc cggggagacg      180

gccgaagcca cgcggctcaa aagaacagca cggcgcagat atacccgcag aaagaatcgg      240

atctgctacc tgcaggagat ctttagtaat gagatggcta aggtggatga ctctttcttc      300

cataggctgg aggagtcctt tttggtggag gaggataaaa agcacgagcg ccacccaatc      360

tttggcaata tcgtggacga ggtggcgtac catgaaaagt acccaaccat atatcatctg      420

aggaagaagc ttgtagacag tactgataag gctgacttgc ggttgatcta tctcgcgctg      480

gcgcatatga tcaaatttcg gggacacttc ctcatcgagg gggacctgaa cccagacaac      540

agcgatgtcg acaaactctt tatccaactg gttcagactt acaatcagct tttcgaagag      600

aacccgatca acgcatccgg agttgacgcc aaagcaatcc tgagcgctag gctgtccaaa      660

tcccggcggc tcgaaaacct catcgcacag ctccctgggg agaagaagaa cggcctgttt      720

ggtaatctta tcgccctgtc actcgggctg acccccaact ttaaatctaa cttcgacctg      780

gccgaagatg ccaagcttca actgagcaaa gacacctacg atgatgatct cgacaatctg      840

ctggcccaga tcggcgacca gtacgcagac cttttttgg cggcaaagaa cctgtcagac      900

gccattctgc tgagtgatat tctgcgagtg aacacggaga tcaccaaagc tccgctgagc      960
```

gctagtatga tcaagcgcta tgatgagcac caccaagact tgactttgct gaaggccctt 1020
gtcagacagc aactgcctga gaagtacaag gaaattttct tcgatcagtc taaaaatggc 1080
tacgccggat acattgacgg cggagcaagc caggaggaat tttacaaatt tattaagccc 1140
atcttggaaa aaatggacgg caccgaggag ctgctggtaa agcttaacag agaagatctg 1200
ttgcgcaaac agcgcacttt cgacaatgga agcatccccc accagattca cctgggcgaa 1260
ctgcacgcta tcctcaggcg gcaagaggat ttctaccct ttttgaaaga taacagggaa 1320
aagattgaga aaatcctcac atttcggata ccctactatg taggcccct cgcccgggga 1380
aattccagat tcgcgtggat gactcgcaaa tcagaagaga ccatcactcc ctggaacttc 1440
gaggaagtcg tggataaggg ggcctctgcc cagtccttca tcgaaaggat gactaacttt 1500
gataaaaatc tgcctaacga aaaggtgctt cctaaacact ctctgctgta cgagtacttc 1560
acagtttata acgagctcac caaggtcaaa tacgtcacag aagggatgag aaagccagca 1620
ttcctgtctg gagagcagaa gaaagctatc gtggacctcc tcttcaagac gaaccggaaa 1680
gttaccgtga aacagctcaa agaagactat ttcaaaaaga ttgaatgttt cgactctgtt 1740
gaaatcagcg gagtggagga tcgcttcaac gcatccctgg gaacgtatca cgatctcctg 1800
aaaatcatta aagacaagga cttcctggac aatgaggaga acgaggacat tcttgaggac 1860
attgtcctca cccttacgtt gtttgaagat agggagatga ttgaagaacg cttgaaaact 1920
tacgctcatc tcttcgacga caaagtcatg aaacagctca agaggcgccg atatacagga 1980
tgggggcggc tgtcaagaaa actgatcaat gggatccgag acaagcagag tggaaagaca 2040
atcctggatt ttcttaagtc cgatggattt gccaaccgga acttcatgca gttgatccat 2100
gatgactctc tcacctttaa ggaggacatc cagaaagcac aagtttctgg ccaggggac 2160
agtcttcacg agcacatcgc taatcttgca ggtagcccag ctatcaaaaa gggaatactg 2220
cagaccgtta aggtcgtgga tgaactcgtc aaagtaatgg gaaggcataa gcccgagaat 2280
atcgttatcg agatggcccg agagaaccaa actacccaga agggacagaa gaacagtagg 2340
gaaaggatga agaggattga agagggtata aaagaactgg ggtcccaaat ccttaaggaa 2400
cacccagttg aaaacaccca gcttcagaat gagaagctct acctgtacta cctgcagaac 2460
ggcagggaca tgtacgtgga tcaggaactg gacatcaatc ggctctccga ctacgacgtg 2520
gatcatatcg tgccccagtc ttttctcaaa gatgattcta ttgataataa agtgttgaca 2580
agatccgata aaaatagagg gaagagtgat aacgtcccct cagaagaagt tgtcaagaaa 2640
atgaaaaatt attggcggca gctgctgaac gccaaactga tcacacaacg gaagttcgat 2700
aatctgacta aggctgaacg aggtggcctg tctgagttgg ataaagccgg cttcatcaaa 2760
aggcagcttg ttgagacacg ccagatcacc aagcacgtgg cccaaattct cgattcacgc 2820
atgaacacca agtacgatga aaatgacaaa ctgattcgag aggtgaaagt tattactctg 2880
aagtctaagc tggtctcaga tttcagaaag gactttcagt tttataaggt gagagagatc 2940
aacaattacc accatgcgca tgatgcctac ctgaatgcag tggtaggcac tgcacttatc 3000

aaaaaatatc ccaagcttga atctgaattt gtttacggag actataaagt gtacgatgtt 3060

aggaaaatga tcgcaaagtc tgagcaggaa ataggcaagg ccaccgctaa gtacttcttt 3120

tacagcaata ttatgaattt tttcaagacc gagattacac tggccaatgg agagattcgg 3180

aagcgaccac ttatcgaaac aaacggagaa acaggagaaa tcgtgtggga caagggtagg 3240

gatttcgcga cagtccggaa ggtcctgtcc atgccgcagg tgaacatcgt taaaaagacc 3300

gaagtacaga ccggaggctt ctccaaggaa agtatcctcc cgaaaaggaa cagcgacaag 3360

ctgatcgcac gcaaaaaaga ttgggacccc aagaaatacg gcggattcga ttctcctaca 3420

gtcgcttaca gtgtactggt tgtggccaaa gtggagaaag ggaagtctaa aaaactcaaa 3480

agcgtcaagg aactgctggg catcacaatc atggagcgat caagcttcga aaaaaacccc 3540

atcgactttc tcgaggcgaa aggatataaa gaggtcaaaa aagacctcat cattaagctt 3600

cccaagtact ctctctttga gcttgaaaac ggccggaaac gaatgctcgc tagtgcgggc 3660

gagctgcaga aaggtaacga gctggcactg ccctctaaat acgttaattt cttgtatctg 3720

gccagccact atgaaaagct caaagggtct cccgaagata atgagcagaa gcagctgttc 3780

gtggaacaac acaaacacta ccttgatgag atcatcgagc aaataagcga attctccaaa 3840

agagtgatcc tcgccgacgc taacctcgat aaggtgcttt ctgcttacaa taagcacagg 3900

gataagccca tcagggagca ggcagaaaac attatccact tgtttactct gaccaacttg 3960

ggcgcgcctg cagccttcaa gtacttcgac accaccatag acagaaagcg gtacacctct 4020

acaaaggagg tcctggacgc cacactgatt catcagtcaa ttacggggct ctatgaaaca 4080

agaatcgacc tctctcagct cggtggagac agcagggctg accccaagaa gaagaggaag 4140

gtgtga 4146

<210> 38
<211> 455
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide Sequence of RNA Guide Strand for use with Cas9 Endonuclease

<220>
<221> misc_feature
<222> (320)..(338)
<223> n is a, c, g, or t

<400> 38
tgtacaaaaa agcaggcttt aaaggaacca attcagtcga ctggatccgg taccaaggtc 60

gggcaggaag agggcctatt tcccatgatt ccttcatatt tgcatatacg atacaaggct 120

gttagagaga taattagaat taatttgact gtaaacacaa agatattagt acaaaatacg 180

tgacgtagaa agtaataatt tcttgggtag tttgcagttt taaaattatg ttttaaaatg 240

gactatcata tgcttaccgt aacttgaaag tatttcgatt tcttggcttt atatatcttg 300

tggaaaggac gaaacaccgn nnnnnnnnnn nnnnnnnngt tttagagcta gaaatagcaa 360

gttaaaataa ggctagtccg ttatcaactt gaaaaagtgg caccgagtcg gtgctttttt 420

tctagaccca gctttcttgt acaaagttgg catta 455

<210> 39
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> BCL11A gene target site

<400> 39
tccaagtgat gtctcggtgg tg 22

<210> 40
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Targeted Homing Endonuclease Target Site

<400> 40
tttccaatta ttcaaccttt ta 22

<210> 41
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Targeted Homing Endonuclease Target Site

<400> 41
tcttgaatta ttcaaccttt ta 22

<210> 42
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Targeted Homing Endonuclease Target Site

<400> 42
tttccattta ttcaatattt ta 22

<210> 43
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Targeted Homing Endonuclease Target Site

<400> 43
tttccattta ttcaatatct tt 22

<210> 44
<211> 12
<212> DNA
<213> Artificial Sequence

<220>
<223> U6 Promoter Sequence of a Generic Cas9 Guide RNA

<400> 44
ggacgaaaca cc 12

<210> 45
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Generic Target-specific Sequence of a Generic Cas9 Guide RNA

<220>
<221> misc_feature
<222> (2)..(20)
<223> n is a, c, g, or t

<400> 45
gnnnnnnnnn nnnnnnnnnn 20

<210> 46
<211> 77
<212> DNA
<213> Artificial Sequence

<220>
<223> Guide RNA Scaffold Sequence of a Generic Cas9 Guide RNA

<400> 46
gttttagagc tagaaatagc aagttaaaat aaggctagtc cgttatcaac ttgaaaaagt 60

ggcaccgagt cggtgct 77

<210> 47
<211> 10
<212> DNA
<213> Artificial Sequence

<220>
<223> Poly T Tail of a Generic Cas9 Guide RNA

<220>
<221> misc_feature
<222> (1)..(4)
<223> n is a, c, g, or t

<400> 47
nnnnttttt 10

<210> 48
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Target-specific Sequence for Cas9 Guide RNA GGN20GG-B

<400> 48
gccatttcta ttatcagact tgg 23

<210> 49
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Target-specific Sequence for Cas9 Guide RNA GGN20GG-C

<400> 49
gctgggcttc tgttgcagta ggg 23

<210> 50
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Target-specific Sequence for Cas9 Guide RNA GGN20GG-D

<400> 50
gaaaatggga gacaaatagc tgg 23

<210> 51
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Target-specific Sequence for Cas9 Guide RNA GGN20GG-E

<400> 51
gaataattca agaaaggtgg tgg 23

<210> 52
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Target-specific Sequence for Cas9 Guide RNA GGN20GG-F

<400> 52
gatattgaat aattcaagaa agg 23

<210> 53
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Target-specific Sequence for Cas9 Guide RNA GGN20GG-G

<400> 53
gcctgagatt ctgatcacaa ggg 23

<210> 54
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Target-specific Sequence for Cas9 Guide RNA GGN20GG-H

<400> 54
ggtaaattct taaggccatg agg 23

<210> 55
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Engineered target site

<400> 55
ttgaggagat gtctctgtta at 22

<210> 56
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Engineered target site

<400> 56
ttgaggtgat gtctctgtta at 22

<210> 57
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Engineered target site

<400> 57
ttgaagtgat gtctctgtta at 22

<210> 58
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Engineered target site

<400> 58
tccaagtgat gtctctgtta at 22

<210> 59
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Engineered target site

<400> 59
tccaagtgat gtctctgtta at 22

<210> 60
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Engineered target site

<400> 60
ttgaggaggt ttctgtgtta at 22

<210> 61
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Engineered target site

<400> 61
ttgaggaggt ttctcggtgg tg 22

<210> 62
<211> 900
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide Sequence of a I-CpaMI homing endonuclease; ORF, codon optimized for mammalian expression

<400> 62
atgaacacca gctctagctt caatccctgg ttcctgaccg gctttagcga tgcagagtgc 60

tctttcagca tcctgataca ggccaacagc aagtactcca ccggttggag gatcaagccc 120

gtgttcgcca tcggcttgca caagaaggac ctggagcttc tgaagagaat ccagagctat 180

ctgggcgtgg gcaagataca cattcacggc aaagacagca ttcagttcag gattgacagc 240

cccaaggagc tggaggtgat catcaaccac tttgagaact acccccctggt aaccgccaag 300

tgggccgact acaccctctt taagaaggcc ctggacgtaa ttctgttgaa ggagcacctg 360

agccagaagg gcctgcttaa actggtaggc attaaggcga gcctgaatct cgggttgaac 420

ggcagcctca aggaggcgtt cccgaactgg gaagaactgc agatcgacag gccgagctac 480

gtgttcaagg gcatccccga ccccaactgg atcagcggct tcgcgtcagg cgatagcagc 540

tttaatgtga aaatcagcaa ctcccccacg tcactgctca ataaaagggt gcagctgagg 600

ttcggcatcg gactgaacat cagagagaaa gcccttatcc aatacctggt ggcctacttt 660

gacctgtcag acaacctgaa gaacatctac ttcgacctga acagcgcacg gttcgaggtg 720

gtgaagttca gcgacatcac cgacaagatc atccccttct tcgacaagta cagcatacaa 780

ggcaagaaga gcctggacta catcaacttc aaggaagtgg ccgacattat caagagcaag 840

aaccatctta ctagcgaggg cttccaggaa atcttggaca tcaaagccag tatgaacaag 900

<210> 63
<211> 300
<212> PRT
<213> Cryphonectria parasitica

<400> 63

Met Asn Thr Ser Ser Ser Phe Asn Pro Trp Phe Leu Thr Gly Phe Ser
1 5 10 15

```
Asp Ala Glu Cys Ser Phe Ser Ile Leu Ile Gln Ala Asn Ser Lys Tyr
            20                  25                  30

Ser Thr Gly Trp Arg Ile Lys Pro Val Phe Ala Ile Gly Leu His Lys
        35                  40                  45

Lys Asp Leu Glu Leu Leu Lys Arg Ile Gln Ser Tyr Leu Gly Val Gly
    50                  55                  60

Lys Ile His Ile His Gly Lys Asp Ser Ile Gln Phe Arg Ile Asp Ser
65                  70                  75                  80

Pro Lys Glu Leu Glu Val Ile Ile Asn His Phe Glu Asn Tyr Pro Leu
                85                  90                  95

Val Thr Ala Lys Trp Ala Asp Tyr Thr Leu Phe Lys Lys Ala Leu Asp
            100                 105                 110

Val Ile Leu Leu Lys Glu His Leu Ser Gln Lys Gly Leu Leu Lys Leu
        115                 120                 125

Val Gly Ile Lys Ala Ser Leu Asn Leu Gly Leu Asn Gly Ser Leu Lys
    130                 135                 140

Glu Ala Phe Pro Asn Trp Glu Glu Leu Gln Ile Asp Arg Pro Ser Tyr
145                 150                 155                 160

Val Phe Lys Gly Ile Pro Asp Pro Asn Trp Ile Ser Gly Phe Ala Ser
                165                 170                 175

Gly Asp Ser Ser Phe Asn Val Lys Ile Ser Asn Ser Pro Thr Ser Leu
            180                 185                 190

Leu Asn Lys Arg Val Gln Leu Arg Phe Gly Ile Gly Leu Asn Ile Arg
        195                 200                 205

Glu Lys Ala Leu Ile Gln Tyr Leu Val Ala Tyr Phe Asp Leu Ser Asp
    210                 215                 220

Asn Leu Lys Asn Ile Tyr Phe Asp Leu Asn Ser Ala Arg Phe Glu Val
225                 230                 235                 240

Val Lys Phe Ser Asp Ile Thr Asp Lys Ile Ile Pro Phe Phe Asp Lys
                245                 250                 255

Tyr Ser Ile Gln Gly Lys Lys Ser Leu Asp Tyr Ile Asn Phe Lys Glu
            260                 265                 270

Val Ala Asp Ile Ile Lys Ser Lys Asn His Leu Thr Ser Glu Gly Phe
        275                 280                 285
```

Gln Glu Ile Leu Asp Ile Lys Ala Ser Met Asn Lys
290 295 300

<210> 64
<211> 5264
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of BCL11A gene targeting nuclease-encoding plasmid pExodusBCL11Ahje

<400> 64
gacggatcgg gagatctccc gatcccctat ggtgcactct cagtacaatc tgctctgatg 60

ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120

cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180

ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240

gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata 300

tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 360

cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420

attgacgtca atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt 480

atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540

atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 600

tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660

actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 720

aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg 780

gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca 840

ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gctggctagc 900

gtttaaactt aagcttggta ccgagctcgg atccactagt ccagtgtggt ggaattctgc 960

aggtacgttg acgccgccac catgggatat ccatacgatg tcccagatta tgcgccacct 1020

aagaagaaac gcaaagtccc cgggggcagc cacatggacc tgacctacgc ctatctggtc 1080

ggcctgttcg agggcgacgg ctattttaca atagctaagg ccggcaagta tctgaactac 1140

gagctgggca tcacactctc catcaaggac gctcagctca tctacaagat caaggacatc 1200

ctcggcgtgg gcaacgtgta ctttaggaag tacaggcaac atgagatggt cagcctgcga 1260

atccaggaca aaaaccacct gaagaacttc atcctgccca tcttcgacaa gtaccccatg 1320

ctgagcaaca agcagtacga ctacctgcga ttcaaggatg ccctcctgtc caacatcatc 1380

tatagcgacg acctgcccga gtacgccagg agcaacgagt caatcaatag cgtggacagc 1440

atcatcaaca cctcatactt cagcgcctgg ctggttggct tcatcgaggc cgagggctgc 1500

ttcaccacct acaaggcatc caaggacaag tacctgacag cgggcttctc catagcccag 1560

aaggacggcg acattctcat ctccgcgatc cacaaatacc tgagcttcac gaccaaaccc 1620

tacaaagaca agaccaactg tagccacctc aaggtcaccg gcgtgaggag cgtcaataac 1680
gtggttaagt tcatccaggg tgcgccggtc aagctgctgg gtaacaagaa gctgcagtac 1740
aaactttgga taaagcagct gcgcaagatc tcccgataca gcgagaaaat ccagctgccc 1800
agtaactact aatctagagg gcccgtttaa acccgctgat cagcctcgac tgtgccttct 1860
agttgccagc catctgttgt ttgcccctcc cccgtgcctt ccttgaccct ggaaggtgcc 1920
actcccactg tcctttccta ataaaatgag gaaattgcat cgcattgtct gagtaggtgt 1980
cattctattc tgggggggtgg ggtggggcag gacagcaagg gggaggattg ggaagacaat 2040
agcaggcatg ctggggatgc ggtgggctct atggcttctg aggcggaaag aaccagctgg 2100
ggctctaggg ggtatcccca cgcgccctgt agcggcgcat taagcgcggc gggtgtggtg 2160
gttacgcgca gcgtgaccgc tacacttgcc agcgccctag cgcccgctcc tttcgctttc 2220
ttcccttcct ttctcgccac gttcgccggc tttccccgtc aagctctaaa tcgggggctc 2280
cctttagggt tccgatttag tgctttacgg cacctcgacc ccaaaaaact tgattagggt 2340
gatggttcac gtagtgggcc atcgccctga tagacggttt ttcgcccttt gacgttggag 2400
tccacgttct ttaatagtgg actcttgttc caaactggaa caacactcaa ccctatctcg 2460
gtctattctt ttgatttata agggattttg ccgatttcgg cctattggtt aaaaaatgag 2520
ctgatttaac aaaaatttaa cgcgaattaa ttctgtggaa tgtgtgtcag ttagggtgtg 2580
gaaagtcccc aggctcccca gcaggcagaa gtatgcaaag cctatcagga catagcgttg 2640
gctacccgtg atattgctga agagcttggc ggcgaatggg ctgaccgctt cctcgtgctt 2700
tacggtatcg ccgctcccga ttcgcagcgc atcgccttct atcgccttct tgacgagttc 2760
ttctgagcgg gactctgggg ttcgaaatga ccgaccaagc gacgcccaac ctgccatcac 2820
gagatttcga ttccaccgcc gccttctatg aaaggttggg cttcggaatc gttttccggg 2880
acgccggctg gatgatcctc cagcgcgggg atctcatgct ggagttcttc gcccacccca 2940
acttgtttat tgcagcttat aatggttaca aataaagcaa tagcatcaca aatttcacaa 3000
ataaagcatt tttttcactg cattctagtt gtggtttgtc caaactcatc aatgtatctt 3060
atcatgtctg tataccgtcg acctctagct agagcttggc gtaatcatgg tcatagctgt 3120
ttcctgtgtg aaattgttat ccgctcacaa ttccacacaa catacgagcc ggaagcataa 3180
agtgtaaagc ctggggtgcc taatgagtga gctaactcac attaattgcg ttgcgctcac 3240
tgcccgcttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg 3300
cggggagagg cggtttgcgt attgggcgct cttccgcttc ctcgctcact gactcgctgc 3360
gctcggtcgt tcggctgcgg cgagcggtat cagctcactc aaaggcggta atacggttat 3420
ccacagaatc aggggataac gcaggaaaga acatgtgagc aaaaggccag caaaaggcca 3480
ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag gctccgcccc cctgacgagc 3540
atcacaaaaa tcgacgctca agtcagaggt ggcgaaaccc gacaggacta taaagatacc 3600
aggcgtttcc ccctggaagc tccctcgtgc gctctcctgt tccgaccctg ccgcttaccg 3660
gatacctgtc cgcctttctc ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta 3720

ggtatctcag ttcggtgtag gtcgttcgct ccaagctggg ctgtgtgcac gaacccccg 3780

ttcagcccga ccgctgcgcc ttatccggta actatcgtct tgagtccaac ccggtaagac 3840

acgacttatc gccactggca gcagccactg gtaacaggat tagcagagcg aggtatgtag 3900

gcggtgctac agagttcttg aagtggtggc ctaactacgg ctacactaga agaacagtat 3960

ttggtatctg cgctctgctg aagccagtta ccttcggaaa aagagttggt agctcttgat 4020

ccggcaaaca aaccaccgct ggtagcggtt tttttgtttg caagcagcag attacgcgca 4080

gaaaaaaagg atctcaagaa gatcctttga tcttttctac ggggtctgac gctcagtgga 4140

acgaaaactc acgttaaggg attttggtca tgagattatc aaaaaggatc ttcacctaga 4200

tccttttaaa ttaaaaatga agttttaaat caatctaaag tatatatgag taaacttggt 4260

ctgacagtta ccaatgctta atcagtgagg cacctatctc agcgatctgt ctatttcgtt 4320

catccatagt tgcctgactc cccgtcgtgt agataactac gatacgggag ggcttaccat 4380

ctggccccag tgctgcaatg ataccgcgag acccacgctc accggctcca gatttatcag 4440

caataaacca gccagccgga agggccgagc gcagaagtgg tcctgcaact ttatccgcct 4500

ccatccagtc tattaattgt tgccgggaag ctagagtaag tagttcgcca gttaatagtt 4560

tgcgcaacgt tgttgccatt gctacaggca tcgtggtgtc acgctcgtcg tttggtatgg 4620

cttcattcag ctccggttcc caacgatcaa ggcgagttac atgatccccc atgttgtgca 4680

aaaaagcggt tagctccttc ggtcctccga tcgttgtcag aagtaagttg gccgcagtgt 4740

tatcactcat ggttatggca gcactgcata attctcttac tgtcatgcca tccgtaagat 4800

gcttttctgt gactggtgag tactcaacca agtcattctg agaatagtgt atgcggcgac 4860

cgagttgctc ttgcccggcg tcaatacggg ataataccgc gccacatagc agaactttaa 4920

aagtgctcat cattggaaaa cgttcttcgg ggcgaaaact ctcaaggatc ttaccgctgt 4980

tgagatccag ttcgatgtaa cccactcgtg cacccaactg atcttcagca tcttttactt 5040

tcaccagcgt ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa 5100

gggcgacacg gaaatgttga atactcatac tcttcctttt tcaatattat tgaagcattt 5160

atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa aataaacaaa 5220

taggggttcc gcgcacattt ccccgaaaag tgccacctga cgtc 5264

<210> 65
<211> 5100
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of TREX2-encoding plasmid pExodus CMV.Trex2

<400> 65
gacggatcgg gagatctccc gatcccctat ggtgcactct cagtacaatc tgctctgatg 60

ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120

cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180

ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata 300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc 360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc 420
attgacgtca atgggtggag tatttacggt aaactgccca cttggcagta catcaagtgt 480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca 600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg 780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca 840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gctggctagc 900
atgtctgagc cacctcgggc tgagaccttt gtattcctgg acctagaagc cactgggctc 960
ccaaacatgg accctgagat tgcagagata tccctttttg ctgttcaccg ctcttccctg 1020
gagaacccag aacgggatga ttctggttcc ttggtgctgc cccgtgttct ggacaagctc 1080
acactgtgca tgtgcccgga gcgccccttt actgccaagg ccagtgagat tactggtttg 1140
agcagcgaaa gcctgatgca ctgcgggaag gctggtttca atggcgctgt ggtaaggaca 1200
ctgcagggct tcctaagccg ccaggagggc cccatctgcc ttgtggccca caatggcttc 1260
gattatgact tcccactgct gtgcacggag ctacaacgtc tgggtgccca tctgccccaa 1320
gacactgtct gcctggacac actgcctgca ttgcggggcc tggaccgtgc tcacagccac 1380
ggcaccaggg ctcaaggccg caaaagctac agcctggcca gtctcttcca ccgctacttc 1440
caggctgaac ccagtgctgc ccattcagca gaaggtgatg tgcacaccct gcttctgatc 1500
ttcctgcatc gtgctcctga gctgctcgcc tgggcagatg agcaggcccg cagctgggct 1560
catattgagc ccatgtacgt gccacctgat ggtccaagcc tcgaagcctg aattctgcag 1620
atatccagca cagtggcggc cgctcgagtc tagagggccc gtttaaaccc gctgatcagc 1680
ctcgactgtg ccttctagtt gccagccatc tgttgtttgc ccctcccccg tgccttcctt 1740
gaccctggaa ggtgccactc ccactgtcct ttcctaataa aatgaggaaa ttgcatcgca 1800
ttgtctgagt aggtgtcatt ctattctggg gggtggggtg gggcaggaca gcaaggggga 1860
ggattgggaa gacaatagca ggcatgctgg ggatgcggtg ggctctatgg cttctgaggc 1920
ggaaagaacc agctggggct ctaggggtga tccccacgcg ccctgtagcg gcgcattaag 1980
cgcggcgggt gtggtggtta cgcgcagcgt gaccgctaca cttgccagcg ccctagcgcc 2040
cgctcctttc gctttcttcc cttcctttct cgccacgttc gccggctttc cccgtcaagc 2100
tctaaatcgg gggctccctt tagggttccg atttagtgct ttacggcacc tcgaccccaa 2160
aaaacttgat tagggtgatg gttcacgtag tgggccatcg ccctgataga cggtttttcg 2220
ccctttgacg ttggagtcca cgttctttaa tagtggactc ttgttccaaa ctggaacaac 2280

actcaaccct atctcggtct attcttttga tttataaggg attttgccga tttcggccta 2340
ttggttaaaa aatgagctga tttaacaaaa atttaacgcg aattaattct gtggaatgtg 2400
tgtcagttag ggtgtggaaa gtccccaggc tccccagcag gcagaagtat gcaaagccta 2460
tcaggacata gcgttggcta cccgtgatat tgctgaagag cttggcggcg aatgggctga 2520
ccgcttcctc gtgctttacg gtatcgccgc tcccgattcg cagcgcatcg ccttctatcg 2580
ccttcttgac gagttcttct gagcgggact ctggggttcg aaatgaccga ccaagcgacg 2640
cccaacctgc catcacgaga tttcgattcc accgccgcct tctatgaaag gttgggcttc 2700
ggaatcgttt tccgggacgc cggctggatg atcctccagc gcggggatct catgctggag 2760
ttcttcgccc accccaactt gtttattgca gcttataatg gttacaaata aagcaatagc 2820
atcacaaatt tcacaaataa agcatttttt tcactgcatt ctagttgtgg tttgtccaaa 2880
ctcatcaatg tatcttatca tgtctgtata ccgtcgacct ctagctagag cttggcgtaa 2940
tcatggtcat agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc acacaacata 3000
cgagccggaa gcataaagtg taaagcctgg ggtgcctaat gagtgagcta actcacatta 3060
attgcgttgc gctcactgcc cgctttccag tcgggaaacc tgtcgtgcca gctgcattaa 3120
tgaatcggcc aacgcgcggg gagaggcggt ttgcgtattg ggcgctcttc cgcttcctcg 3180
ctcactgact cgctgcgctc ggtcgttcgg ctgcggcgag cggtatcagc tcactcaaag 3240
gcggtaatac ggttatccac agaatcaggg gataacgcag gaaagaacat gtgagcaaaa 3300
ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc 3360
cgccccctg acgagcatca caaaaatcga cgctcaagtc agaggtggcg aaacccgaca 3420
ggactataaa gataccaggc gtttcccccct ggaagctccc tcgtgcgctc tcctgttccg 3480
accctgccgc ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgctttct 3540
catagctcac gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt 3600
gtgcacgaac cccccgttca gcccgaccgc tgcgccttat ccggtaacta tcgtcttgag 3660
tccaacccgg taagacacga cttatcgcca ctggcagcag ccactggtaa caggattagc 3720
agagcgaggt atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac 3780
actagaagaa cagtatttgg tatctgcgct ctgctgaagc cagttacctt cggaaaaaga 3840
gttggtagct cttgatccgg caaacaaacc accgctggta gcggtttttt tgtttgcaag 3900
cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg 3960
tctgacgctc agtggaacga aaactcacgt taagggattt tggtcatgag attatcaaaa 4020
aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata 4080
tatgagtaaa cttggtctga cagttaccaa tgcttaatca gtgaggcacc tatctcagcg 4140
atctgtctat ttcgttcatc catagttgcc tgactccccg tcgtgtagat aactacgata 4200
cgggagggct taccatctgg ccccagtgct gcaatgatac cgcgagaccc acgctcaccg 4260
gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgcag aagtggtcct 4320

gcaactttat ccgcctccat ccagtctatt aattgttgcc gggaagctag agtaagtagt 4380

tcgccagtta atagtttgcg caacgttgtt gccattgcta caggcatcgt ggtgtcacgc 4440

tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga 4500

tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt tgtcagaagt 4560

aagttggccg cagtgttatc actcatggtt atggcagcac tgcataattc tcttactgtc 4620

atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa 4680

tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa taccgcgcca 4740

catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg aaaactctca 4800

aggatcttac cgctgttgag atccagttcg atgtaaccca ctcgtgcacc caactgatct 4860

tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc 4920

gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatactctt cctttttcaa 4980

tattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt 5040

tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc 5100

<210> 66
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 66
gctggaatgg ttgcagtaac 20

<210> 67
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 67
caaacagcca ttcaccagtg 20

<210> 68
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR Primer

<400> 68
ctgccagctc tctaagtctc c 21

<210> 69
<211> 21
<212> DNA
<213> Artificial Sequence

<220>

<223> PCR Primer

<400> 69
tgcaacacgc acagaacact c 21