

US 20230235325A1

(19) **United States**

(12) **Patent Application Publication**
Davidson et al.

(10) **Pub. No.: US 2023/0235325 A1**

(43) **Pub. Date: Jul. 27, 2023**

(54) **COMPOSITIONS AND METHODS FOR
TREATING HUNTINGTON'S DISEASE AND
RELATED DISORDERS**

Publication Classification

(71) Applicant: **The Children's Hospital of
Philadelphia, Philadelphia, PA (US)**

(72) Inventors: **Beverly L. Davidson, North Liberty, IA
(US); Alejandro Mas Monteys,
Philadelphia, PA (US); Shauna Ebanks,
Claymont, DE (US)**

(21) Appl. No.: **17/938,857**

(22) Filed: **Oct. 7, 2022**

Related U.S. Application Data

(63) Continuation of application No. 15/766,549, filed on
Apr. 6, 2018, now Pat. No. 11,473,084, filed as
application No. PCT/US2016/056417 on Oct. 11,
2016.

(60) Provisional application No. 62/239,714, filed on Oct.
9, 2015, provisional application No. 62/352,788, filed
on Jun. 21, 2016.

(51) **Int. Cl.**

C12N 15/113 (2006.01)

C12N 15/861 (2006.01)

C12N 15/88 (2006.01)

C12N 15/66 (2006.01)

C12N 15/10 (2006.01)

A61K 48/00 (2006.01)

A61P 25/00 (2006.01)

C12N 9/22 (2006.01)

(52) **U.S. Cl.**

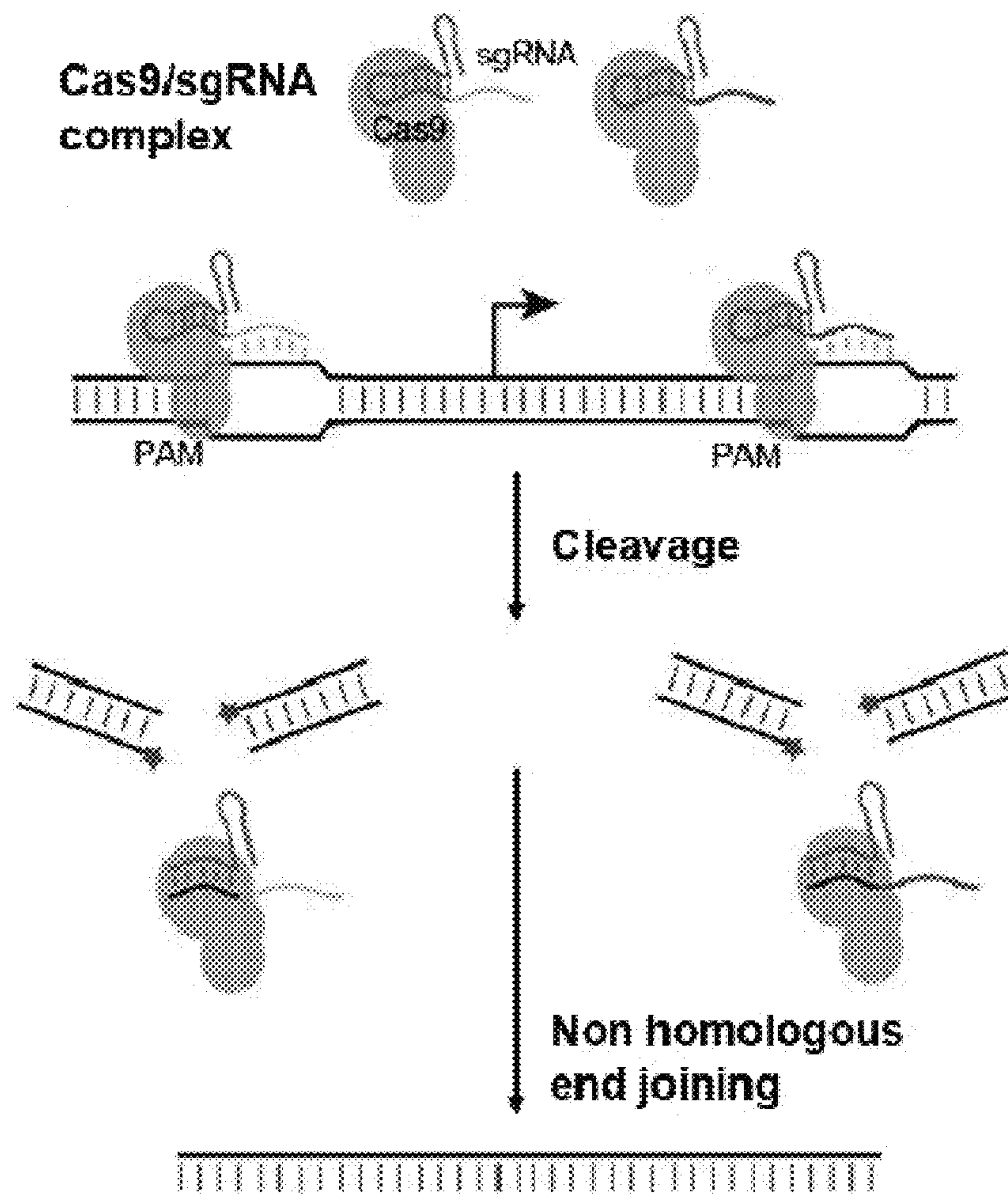
CPC **C12N 15/113** (2013.01); **C12N 15/861**
(2013.01); **C12N 15/88** (2013.01); **C12N**
15/66 (2013.01); **C12N 15/102** (2013.01);
A61K 48/005 (2013.01); **A61P 25/00**
(2018.01); **C12N 9/22** (2013.01); **C12N**
2310/20 (2017.05); **C12N 2330/51** (2013.01);
C12N 2800/80 (2013.01)

(57)

ABSTRACT

Compositions and methods are provided for the inhibition,
treatment and/or prevention of Huntington's disease and
related disorders.

Specification includes a Sequence Listing.



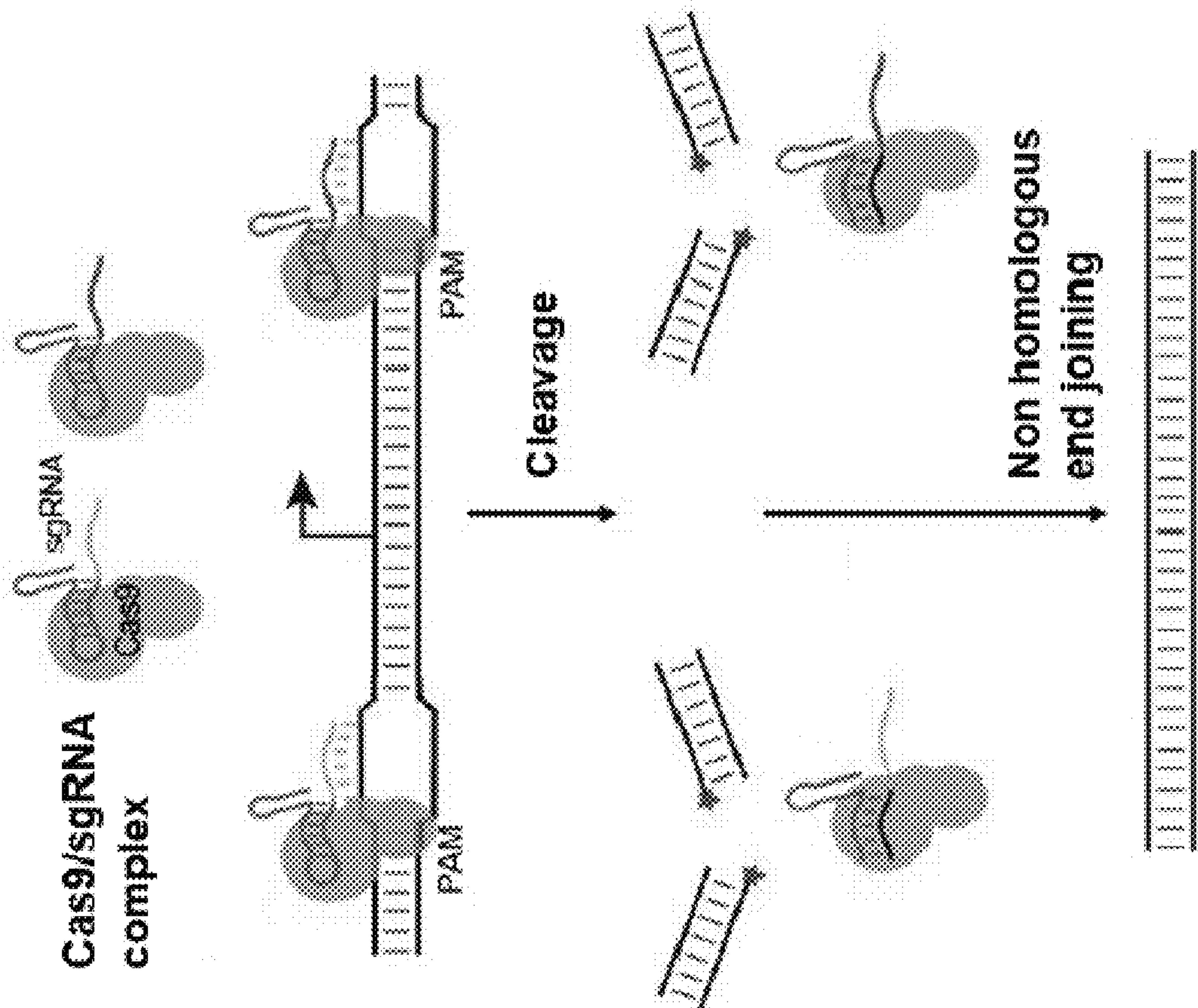


Figure 1A

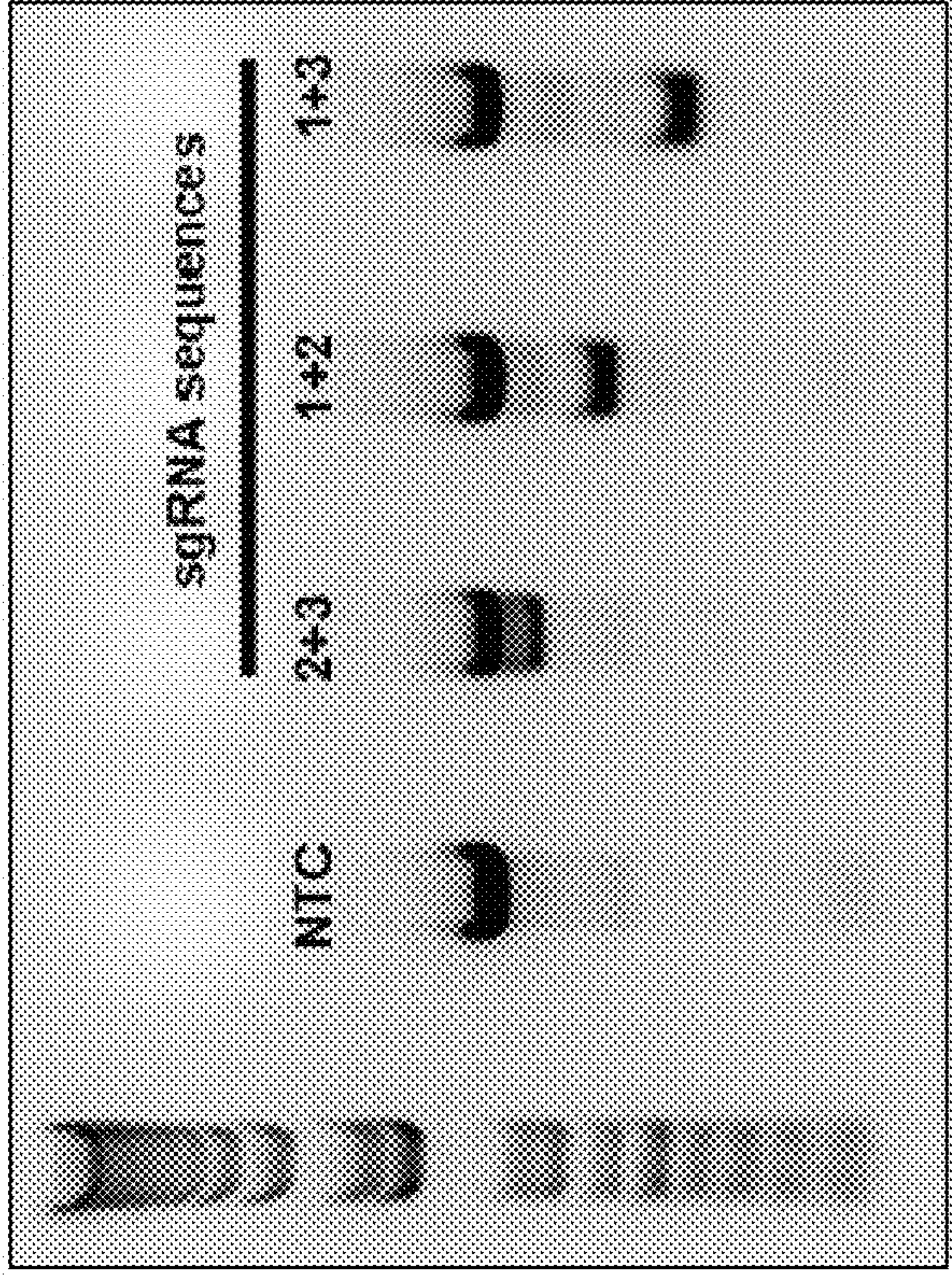
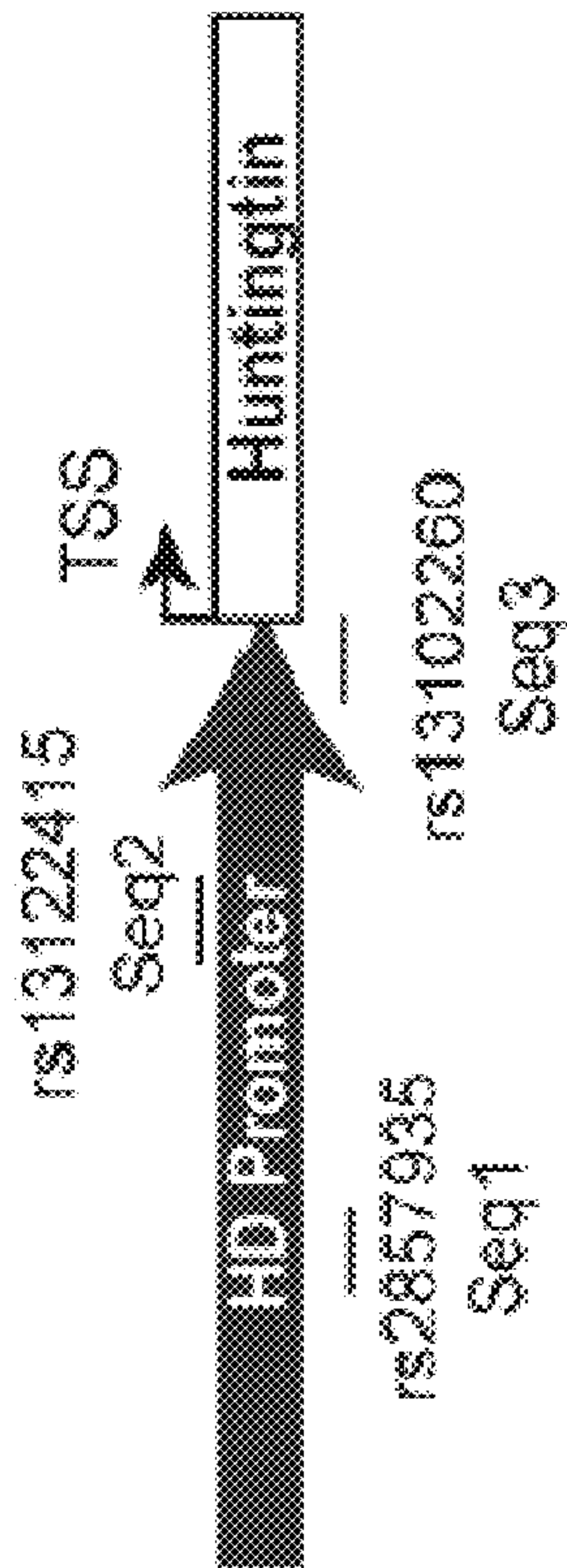


Figure 1B

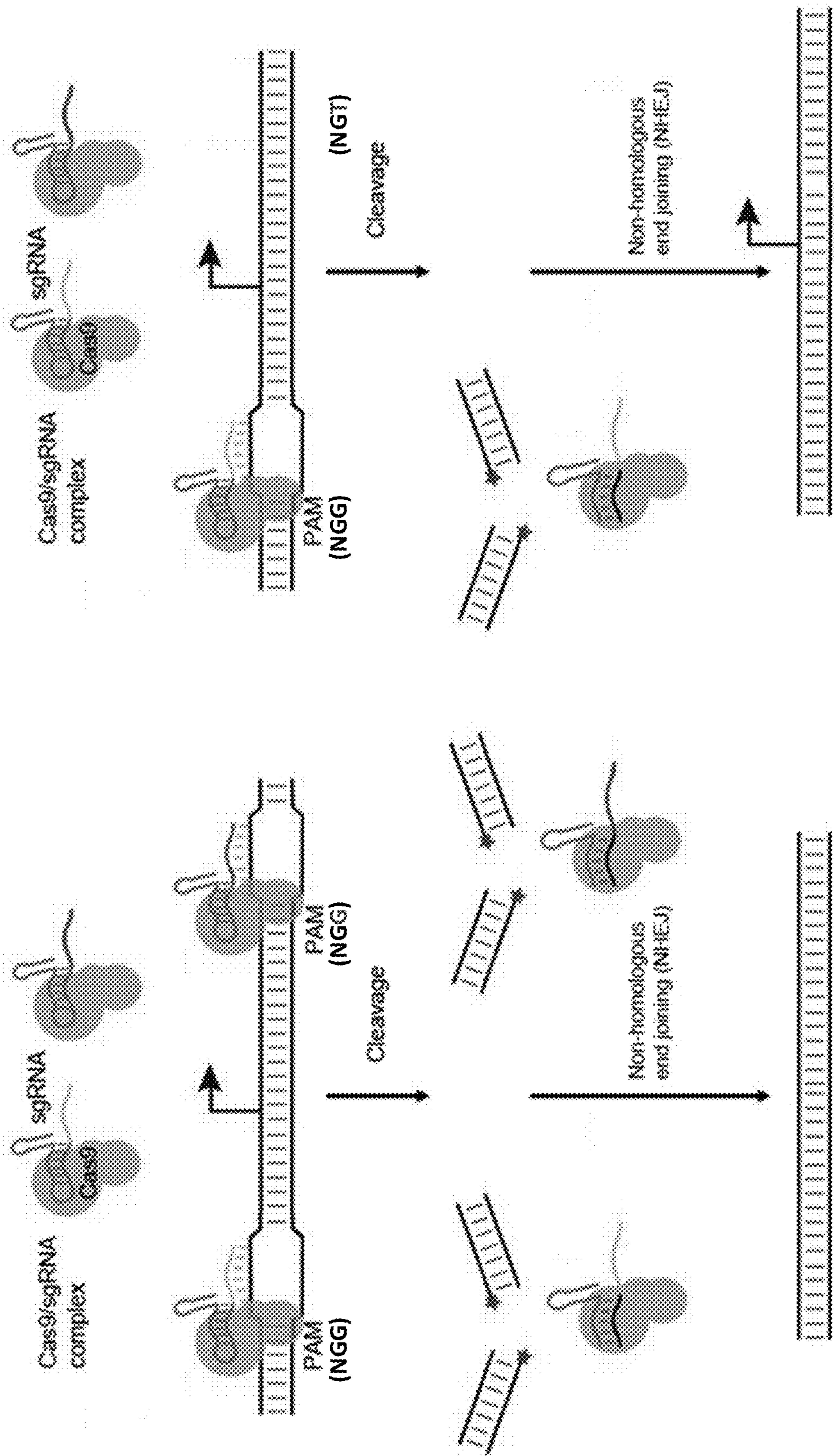


Figure 1C

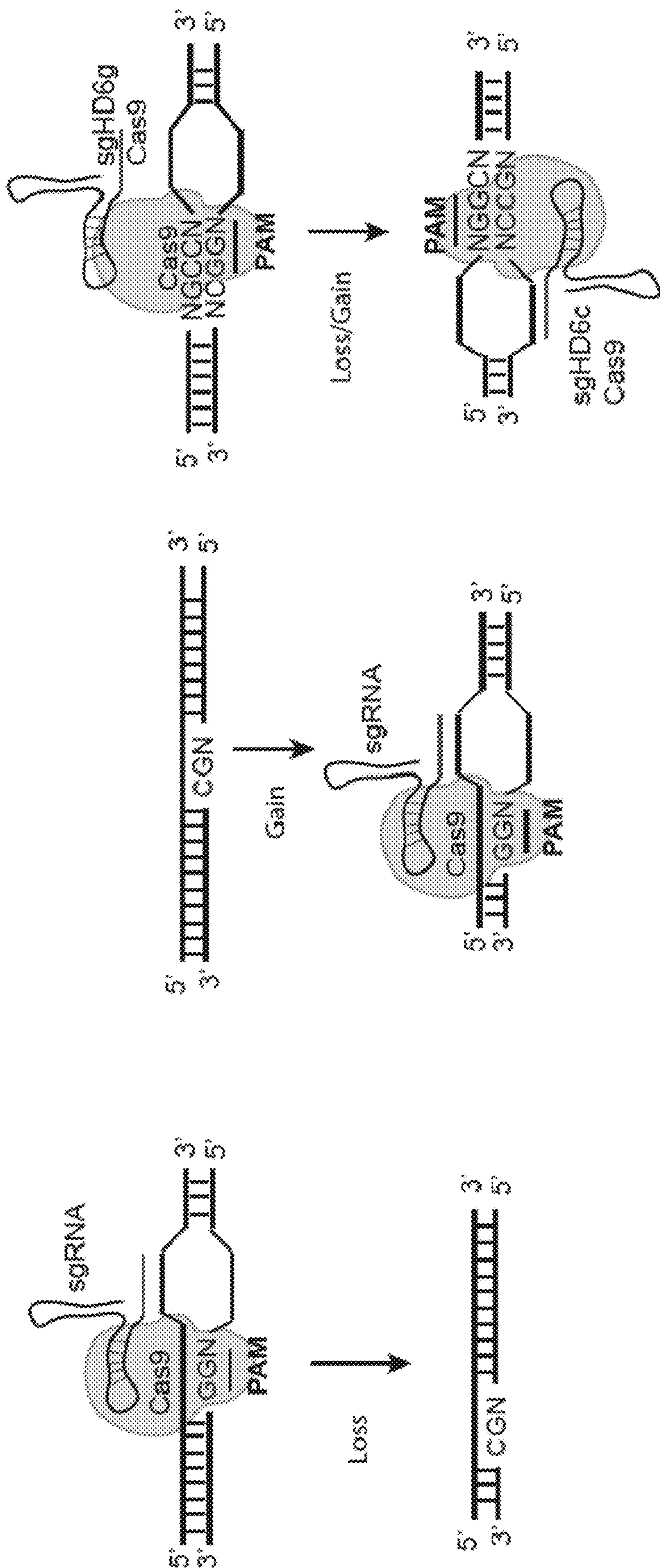
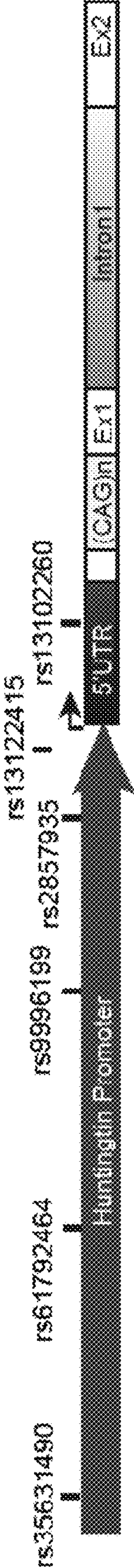
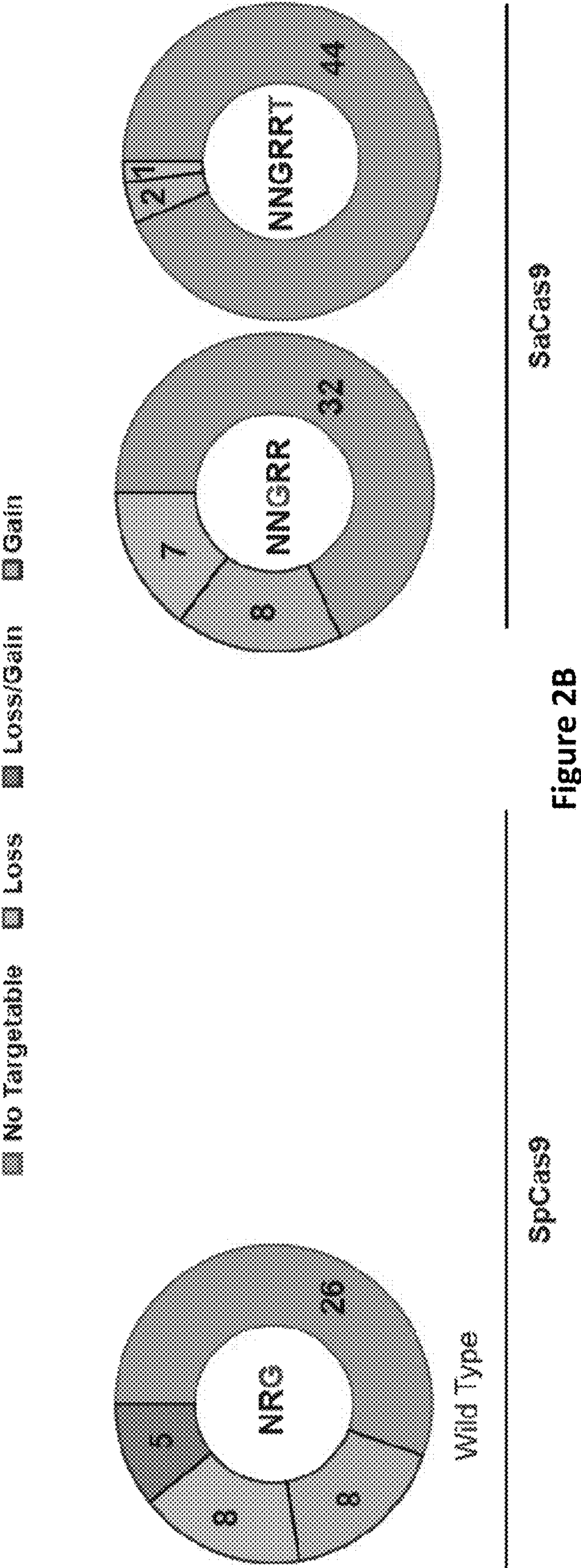


Figure 2A



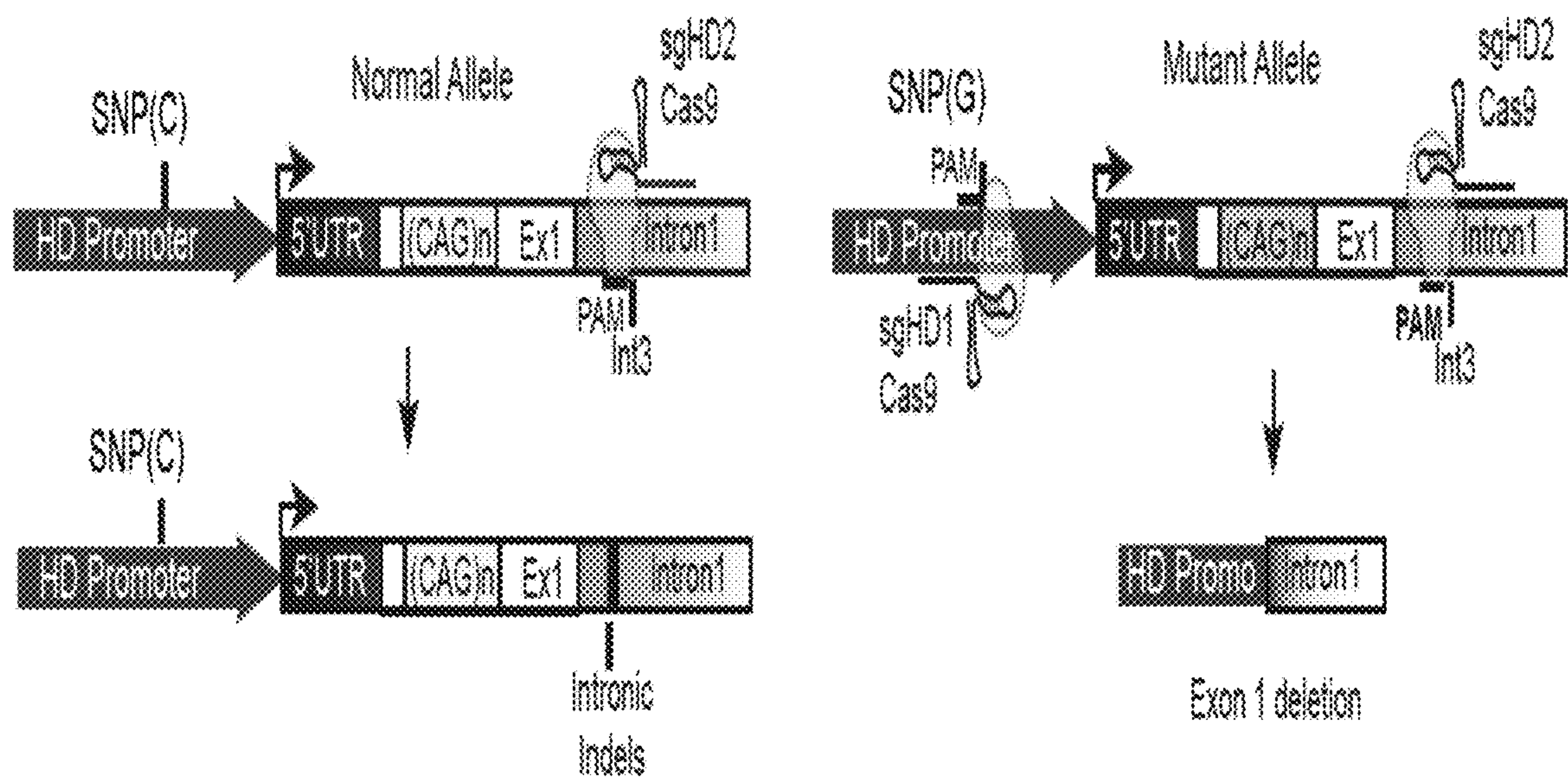


Figure 3A

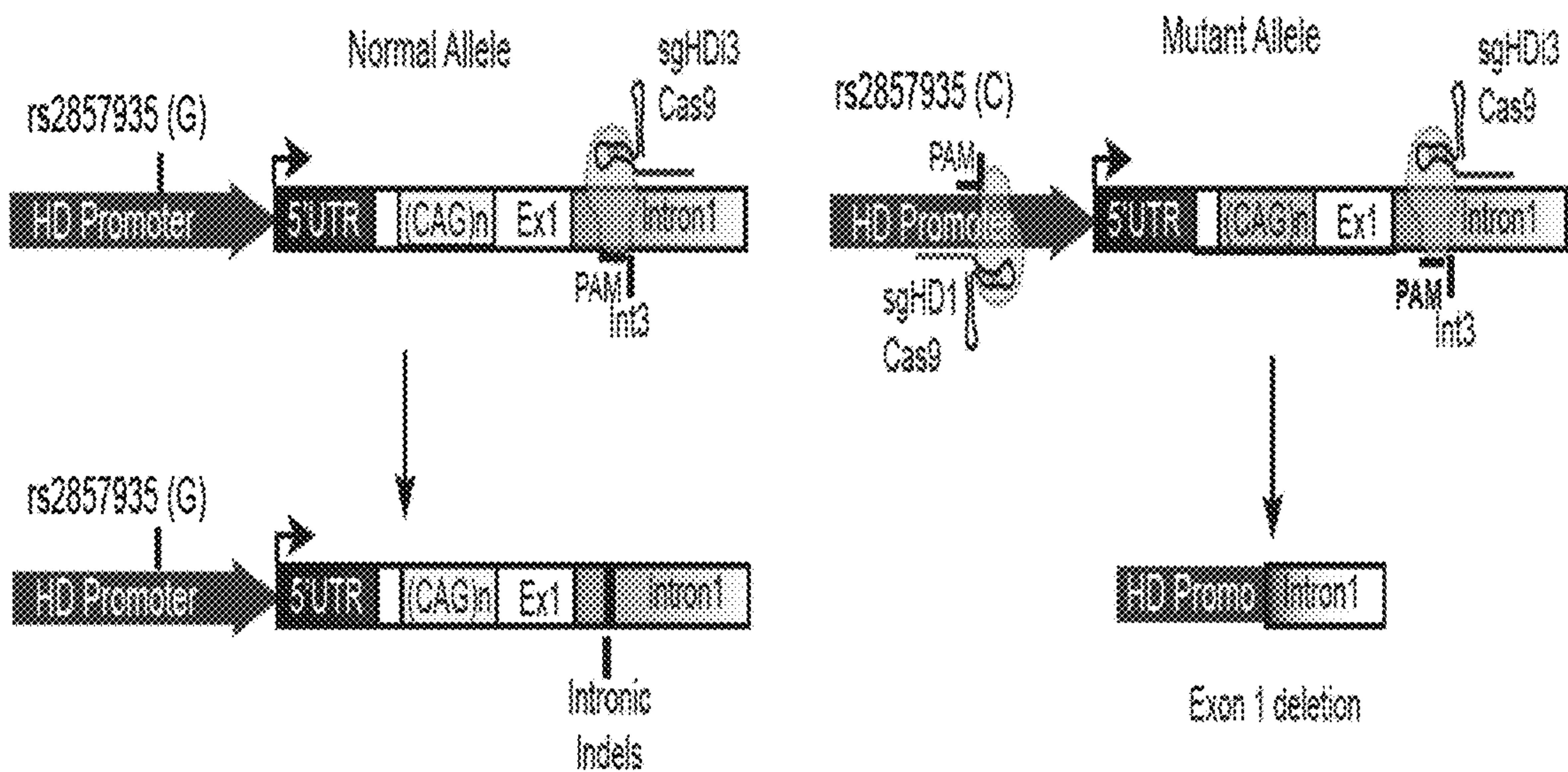


Figure 3B

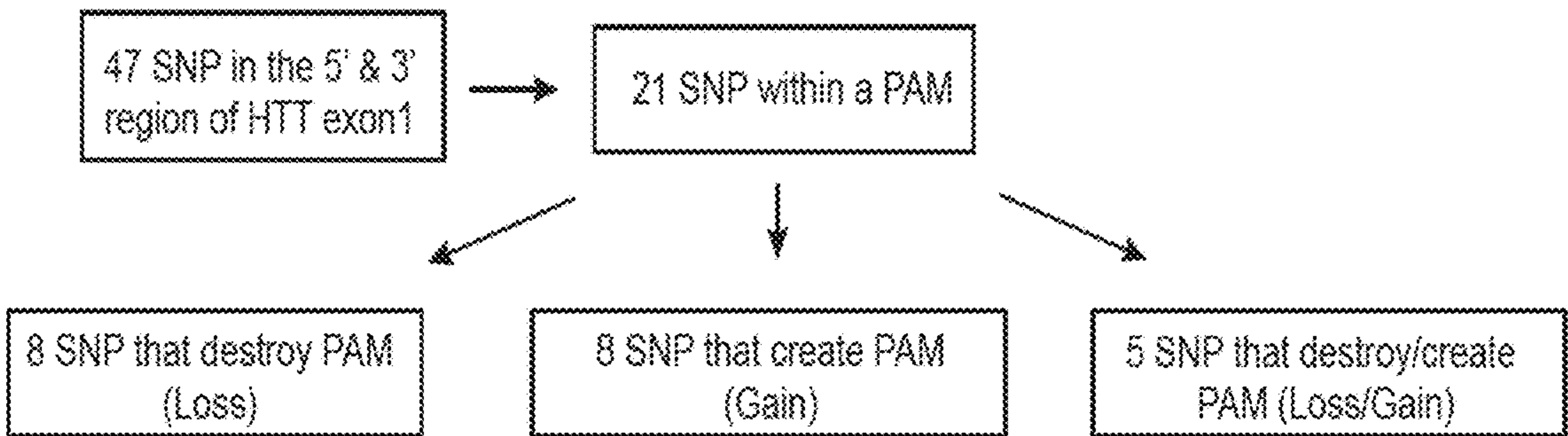


Figure 3C

# Variant ID	Location	Sequence Variation
rs35631490	Promoter	GTCTGCGTCAGGGTTTCCTTCTTTT [C/G] CAGCCCCACCCCGCGTGCATCCCCAC
rs73086139	Promoter	TCAAGGCCTCTTCTCTCTTCTCGGC [A/G] GGACAGGCACAGGCAGGTGGCCAGG
rs73086140	Promoter	GCCAGGTGTCATGCTTAGCTCCCCG [C/G] CCAGTGAGATTCTTTCATTTAACAA
rs113541600	Promoter	TAGGAACCTTATTTCTCTCTCGCTC [-/T] TTTTTTTTTTTTGAGACAGAGTCT
rs61791259	Promoter	GATTACAGGCACCTGCCACCATGCC [C/T] GGCTAATTTTTGTATTTTGTAGTTGA
rs61792460	Promoter	GAGGGTTTCATCTTGTGGTCAGGC [A/G] GACTTGAACCTCCTGACCTCAGGTGA
rs73086144	Promoter	GCCACTGCGCCTTCATCTCTCTTCT [A/G] TGTATGTGTACGCTGTTTTTCTTT
rs73086145	Promoter	GCTGTTTTTCTTTAGAAATGGGGGA [C/G] GTTATCAGGCTCTACATGGTGTGTA
rs61792461	Promoter	TCTACATGGTGTGTAGTCCGCTAGC [A/G] TGTGTAAAGCCTTTCCTGTGTAC
rs61792462	Promoter	CTGTGTACAAGTGCTCATCTGGAA [C/G] AGGATTCTAATGACTGCCTGTGGCT
rs61090955	Promoter	TCATTTTATGTGATTCCCTTCTAGA [A/G] GTACTACTCATTACTTCTGCTTGCA
rs77384845	Promoter	TAGCTGAAGGAAGGACAGGGACTGT [C/T] ATACACTAGCTAAGAGGCAAACTGC
rs10011412	Promoter	agctgaaggaaggacagggaactgtc [A/G] TACACTAGCTAAGAGGCAAACTGCT
rs61792464	Promoter	TCCCTCATTCAAGTTGATGTCTTAA [C/G] CCCAGAACCTCAGAATGGGATTGT
rs112435590	Promoter	TCCATGCCAAGAAGGCAACAGAGAG [G/T] GCCAGGGAGACTGAAGTCATACCCT
rs111670395	Promoter	CCCAGGTTCAAGCAATTCTGCCTCA [A/G] CCTCCGGAATAGCTGGGACTACAGG
rs762855	Promoter	TTGAGAAGGACAGCAGAGAAACAGC [C/T] GTTAGTtccagttcttgggaggct
rs9996199	Promoter	TGTGGCCTGGCTAAAGTAGGCTTTA [C/G] TGGGCTCCTCTCTGCCTGCATCACC
rs143861513	Promoter	TCCGGGGCGCTGCGCTGGGACCGAT [-/G] GGGGGGCGCCAGGCCTGTGGACACC
rs28431418	Promoter	CTCCCCGCAGGGCTGTCCGGGTGAG [C/T] ATGGCTCTGGCCACGGGCCAGTGTG
rs2857935	Promoter	GGGCGCAGGCCCATGCGGAAAGGAT [A/C/G] CCCCGCCGACGCCTGGAGCGGGGCG
rs28616835	Promoter	GCGCCGCGCTCGGCGCCCTCCA [C/T] GGCCCCGCCCGTCCATGGCCCCGT
rs149624523	Promoter	CCCCGCCCCGGCCTCGCCACGCCCC [C/T] ACCTCACCACGCCCCCGCATCGCC
rs13122415	Promoter	attacagtctcaccacgccccgtcc [C/G] CTCTCCGTTGAGCCCCGCGCCTTCG
rs113331544	5'UTR	GGCCTTGCTGTGTGAGGCAGAACCT [-/GCGGGG] GCGGGGCGAGGGCGGGCTGGTTCC
rs13132932	5'UTR	TGTGTGAGGCAGAACCTGCGGGGGC [A/G] GGGGCGGGCTGGTTCCCTGGCCAGC
rs13102260	5'UTR	AGCGTCTGGGACGCAAGGCGCCGTG [A/G] GGGCTGCCGGGACGGGTCCAAGATG
rs10009935	Intron 1	CTCACTTGGGTCTTCCCTTGTCTC [C/T] CGCGAGGGGAGGCAGAGCCTTGTGTG
rs58870770	Intron 1	TGAATGAGTTGTGGTTGCCAAGTAA [-/A] GTGGTGAACCTACGTGGTGATTAAT
rs34045730	Intron 1	GAGGTGTACATTTTACCAGTATTCC [A/T] GTCAGGCTTGCCAGAATACGGGGGG
rs28656215	Intron 1	GGAAGTCTGTGTGTGAGTGTACAG [C/T] AGGAGTTAGGAAGTACTCTGGTGCA
rs28571971	Intron 1	ACAGTAGGAGTTAGGAAGTACTCTG [C/G] TGCAGTTCAGGCCTTCTCTTACCT
rs28583447	Intron 1	CAGTAGGAGTTAGGAAGTACTCTGG [C/T] GCAGTTCAGGCCTTCTCTTACCTC
rs28468636	Intron 1	TCTCTTACCTCTCAGTATTCTATTT [C/G] CGATCTGGATGTGTCCCAGATGGCA
rs28564368	Intron 1	TATTCTATTTCCGATCTGGATGTGT [A/C] CCAGATGGCATTGTGGTAAGAATATC
rs28485764	Intron 1	GATGGCATTGTGGTAAGAATATCTCT [A/G] TTAAGACTGATTAATTTTGTAGTAAT
rs77173925	Intron 1	GAATAAATTATTCTAAAGGATGGAA [A/G] AACTTTTTGGATATTTGGAGAAATT
rs3905238	Intron 1	TTGTATCATGTCAATGTATTACTTA [C/T] GCAAAAATAATACATTAAAAAAAT
rs33950430	Intron 1	AGGTATTCACATAATTTTGAGTAACA [-/AACA] CTGCTCACAAAGTTTGGATTTTGGC
rs28377140	Intron 1	AGGCAATTAATACTTGCTTCTGGCA [C/G] TTTCTTATTCTCCTTCAGATTCTTA
rs3856973	Intron 1	ttaaaaaataaaaaataaGTTAACACT [C/T] GATTAACCCTGACATTTCCCTATCC
rs4498089	Intron 1	AGTGTTAACCTATTTTATTTTAA [A/G] AAAATTGTTAAGGGCTTTCAGCAA
rs112353753	Intron 1	GACATGCACTGCCATGCCTGGGTAA [-/T] TTTTTTTTTTCCCCGAGACGGAG
rs57666989	Intron 1	CTGCAAGCTCCGCTTCCCGAGTTCA [C/T] GCCATTCTCCTGCCTCAGTCTCCCA
rs10006129	Intron 1	tcttgatctcctgacctcgctcatcc [C/G] ccgaccttgatgacccacctcg
rs28696693	Intron 1	GGTAATTTTGTATTTTGTAGAG [A/G] TGGGGTTTGGCATGATGAGCAGGC
rs28393280	Intron 1	GGATTTTGAATGCGGAACCAACTGC [A/G] CTTGTTGAACTCTGCTAAGTATAAC

Figure 3D

1000 Genome DATA			Reference Allele			Minor allele		
Ref.Alele	Min. Allele	MAF >0.05	PAM Motif	Strand	PAM Ref>Min	PAM Motif	Strand	PAM Ref>Min
C	G	0.1074	TGG	Negative	Loss	-	-	-
A	G	0.0659	-	-	-	-	-	-
C	G	0.0962	GGG	Negative	Loss	CGG	Positive	Gain
T	delT	0.2091	-	-	-	-	-	-
C	T	0.0661	-	-	-	-	-	-
G	A	0.0966	-	-	-	-	-	-
G	A	0.0899	-	-	-	-	-	-
C	G	0.0899	-	-	-	GAG	Positive	Gain
A	G	0.0901	-	-	-	-	-	-
C	G	0.0579	-	-	-	AAG	Positive	Gain
A	G	0.0865	-	-	-	GAG	Positive	Gain
C	T	0.0984	-	-	-	-	-	-
A	G	0.1158	-	-	-	-	-	-
G	C	0.1372	GGG	Negative	Loss	AAG	Positive	Gain
G	T	0.0881	AGG	Positive	Loss	-	-	-
G	A	0.0663	CAG	Positive	Loss	-	-	-
G	A (Rev strand)	0.4828	-	-	-	-	-	-
C	G	0.1575	CAG	Negative	Loss	TAG	Positive	Gain
-	insG	0.0599	-	-	-	-	-	-
T	C	0.1438	-	-	-	-	-	-
G	C/T (Rev Strand)	0.226	GGG	Negative	Loss	-	-	-
C	T	0.0839	-	-	-	-	-	-
T	C	0.0927	-	-	-	-	-	-
C	G	0.1082	AGG	Negative	Loss	-	-	-
-	insGCGGGG	0.0851	-	-	-	-	-	-
A	G	0.0727	-	-	-	-	-	-
G	A	0.1581	TGG	Positive	Loss	-	-	-
T	C	0.0877	-	-	-	CGG	Negative	Gain
A	delA	0.0845	-	-	-	-	-	-
A	T	0.0589	-	-	-	-	-	-
T	C	0.1544	-	-	-	-	-	-
G	C	0.0839	TGG	Positive	Loss	CAG	Negative	Gain
T	C	0.0839	-	-	-	-	-	-
C	G	0.0839	CGG	Negative	Loss	-	-	-
C	A	0.0839	GGG	Negative	Loss	-	-	-
G	A	0.0845	-	-	-	-	-	-
A	G	0.0843	-	-	-	AAG	Positive	Gain
A	G (Rev Strand)	0.4443	-	-	-	-	-	-
AACA	delAACA	0.0835	-	-	-	-	-	-
G	C	0.1064	CAG	Positive	Loss	AAG	Negative	Gain
G	A (Rev strand)	0.4081	-	-	-	-	-	-
A	G	0.2917	-	-	-	AAG	Positive	Gain
T	delT	0.1154	-	-	-	-	-	-
C	T	0.1424	-	-	-	-	-	-
G	C	0.0845	-	-	-	GGG	Negative	Gain
A	G	0.0839	-	-	-	AGG	Positive	Gain
A	G	0.0717	-	-	-	-	-	-

Figure 3E

SpCas9_VQR/EQR (NGAN_NGNG/NGAG)					
Reference Allele			Minor allele		
PAM Motif	Strand	PAM Ref>Min	PAM Motif	Strand	PAM Ref>Min
GGAA	Negative	Loss	-	-	-
-	-	-	GGCG	Positive	Gain
GGCG	Negative	Loss	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-
GGCG	Positive	Loss	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-
-	-	-	AGCG	Positive	Gain
-	-	-	AGAG	Positive	Gain
-	-	-	-	-	-
TGAC	Negative	Loss	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-
GGAT	Negative	Loss	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-
GGGG	Positive	Loss	-	-	-
-	-	-	GGCG	Positive	Gain
-	-	-	GGCG	Positive	Gain
-	-	-	TGAG	Positive	Gain
CGAG	Negative	Loss	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-
-	-	-	TGCG	Negative	Gain
-	-	-	TGCG	Positive	Gain
GGAC/TGGG	Neg/Neg	Loss	-	-	-
-	-	-	-	-	-
-	-	-	AGAA	Positive	Gain
-	-	-	TGCG	Negative	Gain
-	-	-	-	-	-
-	-	-	-	-	-
CGAG	Negative	Loss	-	-	-
-	-	-	AGAA	Positive	Gain
-	-	-	-	-	-
GGCG	Negative	Loss	-	-	-
-	-	-	CGGG	Negative	Gain
AGAT	Positive	Loss	-	-	-
-	-	-	TGCG	Positive	Gain

Figure 3F

[illegible]

Figure 3G

SaCas9 (NNGRRT/NNGRR)					
Reference Allele			Minor allele		
PAM Motif	Strand	PAM Ref>Min	PAM Motif	Strand	PAM Ref>Min
TGGAA	Negative	Loss	-	-	-
-	-	-	GCGGG	Positive	Gain
-	-	-	-	-	-
-	-	-	-	-	-
CCGGG	Negative	Loss	-	-	-
GCGGA	Positive	Loss	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-
-	-	-	AAGAG	Positive	Gain
-	-	-	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-
GAGAG	Positive	-	GAGAGT	Positive	Gain
-	-	-	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-
GTGAGT	Positive	Loss	GTGAG	Positive	-
CGGGG	Negative	-	CGGGGT	Negative	Gain
-	-	-	-	-	-
-	-	-	GTGGG	Negative	Gain
AGGGG	Negative	Loss	-	-	-
-	-	-	-	-	-
-	-	-	GCGGG	Positive	Gain
TGGGG	Positive	Loss	-	-	-
-	-	-	CGGGA	Negative	Gain
-	-	-	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-
CGGAA	Negative	Loss	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-
-	-	-	AAGAA	Positive	Gain
-	-	-	-	-	-
-	-	-	-	-	-
-	-	-	-	-	-
TCGAGT	Negative	Loss	-	-	-
-	-	-	AAGAA	Positive	Gain
-	-	-	-	-	-
-	-	-	-	-	-
-	-	-	GGGGG	Negative	Gain
-	-	-	-	-	-
-	-	-	-	-	-

Figure 3H

LbCpf1 (TTTN/TTCN/CTTN/TCTN/ATTN/TCCN/TTGN/GTTN/CCCN/CCTN...)

Reference Allele

Minor allele

[illegible]

Figure 3I

[illegible]

Figure 3J

SNP ID: rs9996199 (C/G) PAM: Gain/Loss

Major allele (C): PAM motif on the negative strand (NAG)

(70) 5' GTGGCCTGGCTAAAGTAGGCTTTA 3' CCGAGGAGAGACGGACG 5' (72)
 (151) 3' CACCGGACCGATTTCATCCGAAATGACCCGAGGAGAGACGGACGTAGTG 5'
 PAM

Minor Allele (G): PAM motif on the Positive strand (NAG)

PAM
 (71) 5' GTGGCCTGGCTAAAGTAGGCTTTA 3' TGGGCTCCTCTCTGCCTGCATCAC 3'
 (152) 3' CACCGGACCGATTTCATCCGAAATCACCCGAGGAGAGACGGACGTAGTG 5'
 5' GCCTGGCTAAAGTAGGCTT 3' (73)

SNP ID: rs2857935 (C/G/T) PAM: Loss

Major Allele (C): PAM motif on the Positive strand (NGG)

PAM
 (66) 5' GTCGCCCCGCTCCAGGCGTCGGCGGGGGATCCTTTCCGCATGGGCCTGC 3'
 (153) 3' CAGCGGGGCGAGGTCCGCAGCCGCCCCCTAGGAAAGGCGTACCCGGACG 5'
 5' GCTCCAGGCGTCGGCGG 3' (69)

Minor Allele (G): Disruption of the PAM motive.

(67) 5' GTCGCCCCGCTCCAGGCGTCGGCGGGGGATCCTTTCCGCATGGGCCTGC 3'
 (154) 3' CAGCGGGGCGAGGTCCGCAGCCGCCCCCTAGGAAAGGCGTACCCGGACG 5'

Minor Allele (T): Disruption of the PAM motive.

(68) 5' GTCGCCCCGCTCCAGGCGTCGGCGGGGGAATCCTTTCCGCATGGGCCTGC 3'
 (155) 3' CAGCGGGGCGAGGTCCGCAGCCGCCCCCTAGGAAAGGCGTACCCGGACG 5'

Figure 3K

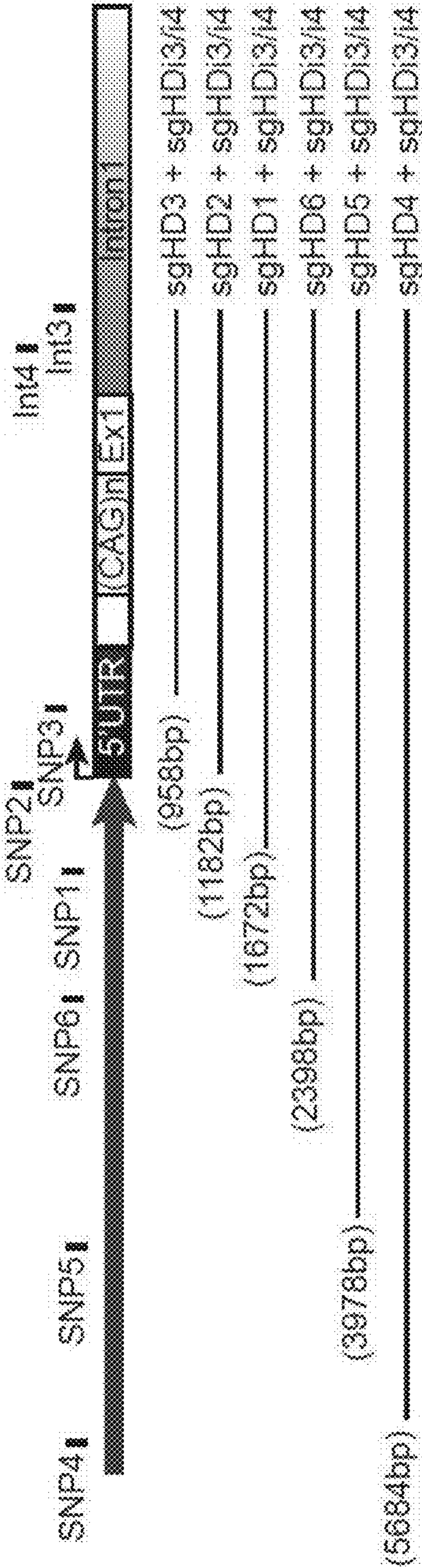


Figure 4A

rs2857935	rs13122415	rs13102260	rs35631490	rs61792464	rs9996199
SNP1 (C/G/T)	SNP2 (C/G)	SNP3 (G/A)	SNP4 (C/G)	SNP5 (C/G)	SNP6 (C/G)
C	C	G	C	G	C
Present	Present	Present	Present	Present	Present
sgHD1	sgHD2	sgHD3	sgHD4	sgHD5G	sgHD6C

Figure 4B

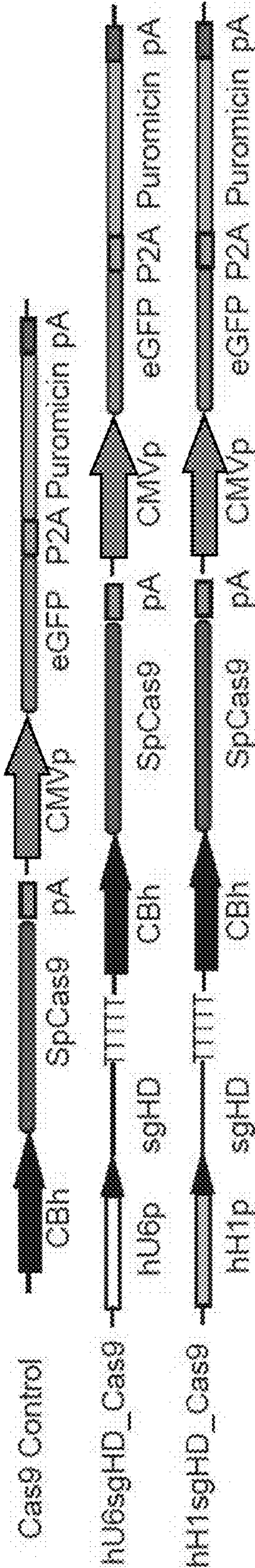


Figure 4C

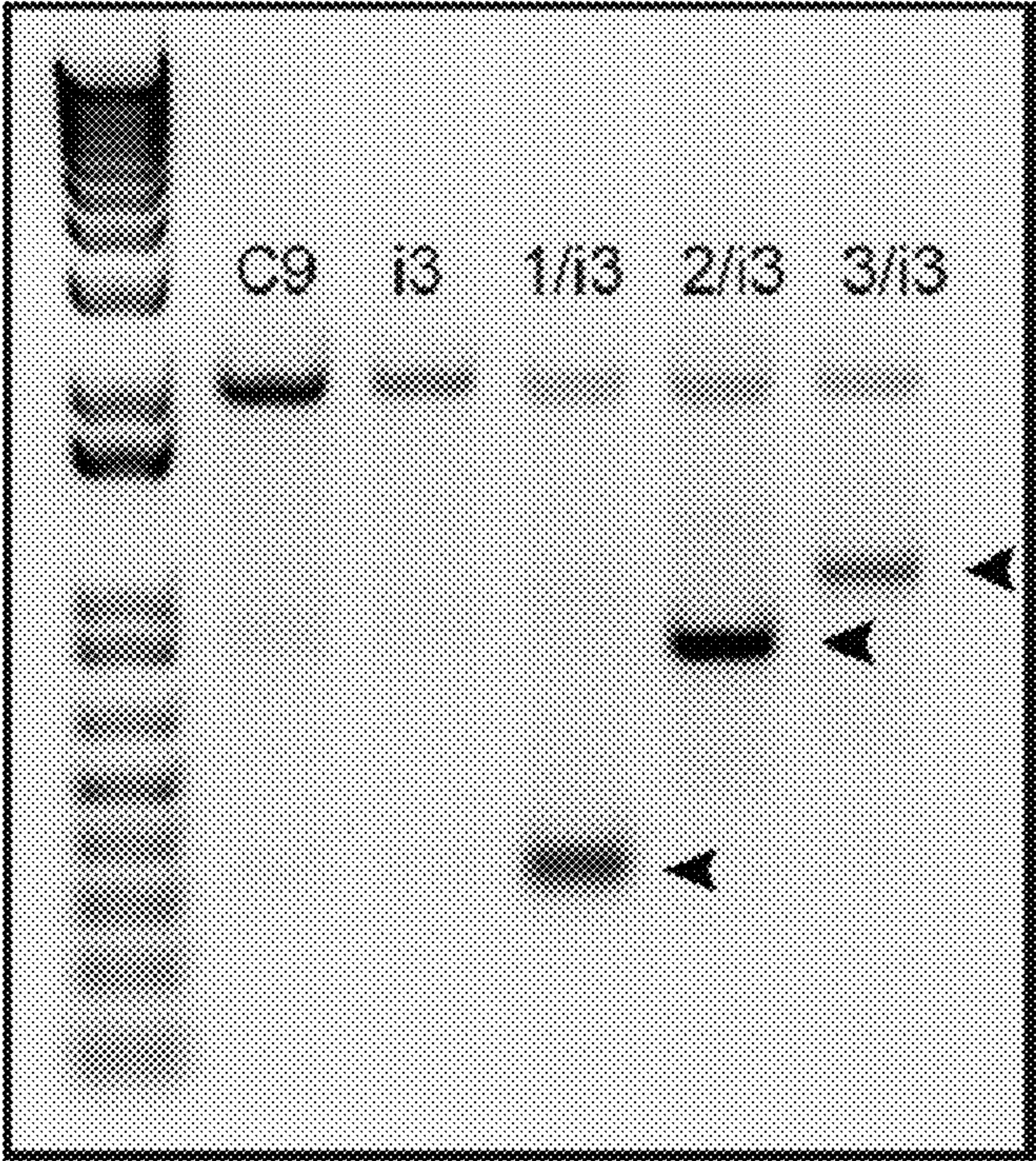


Figure 4D

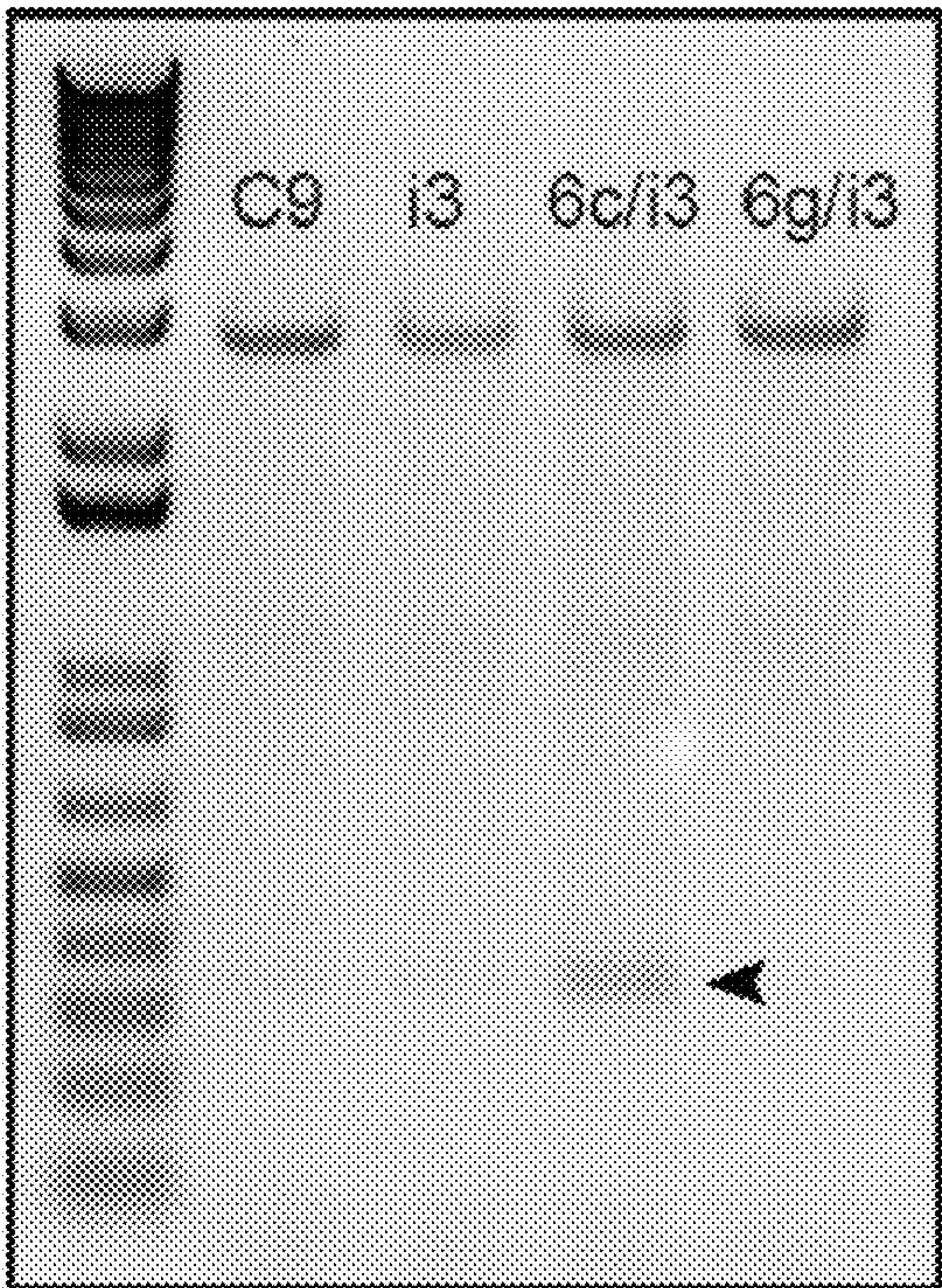


Figure 4E

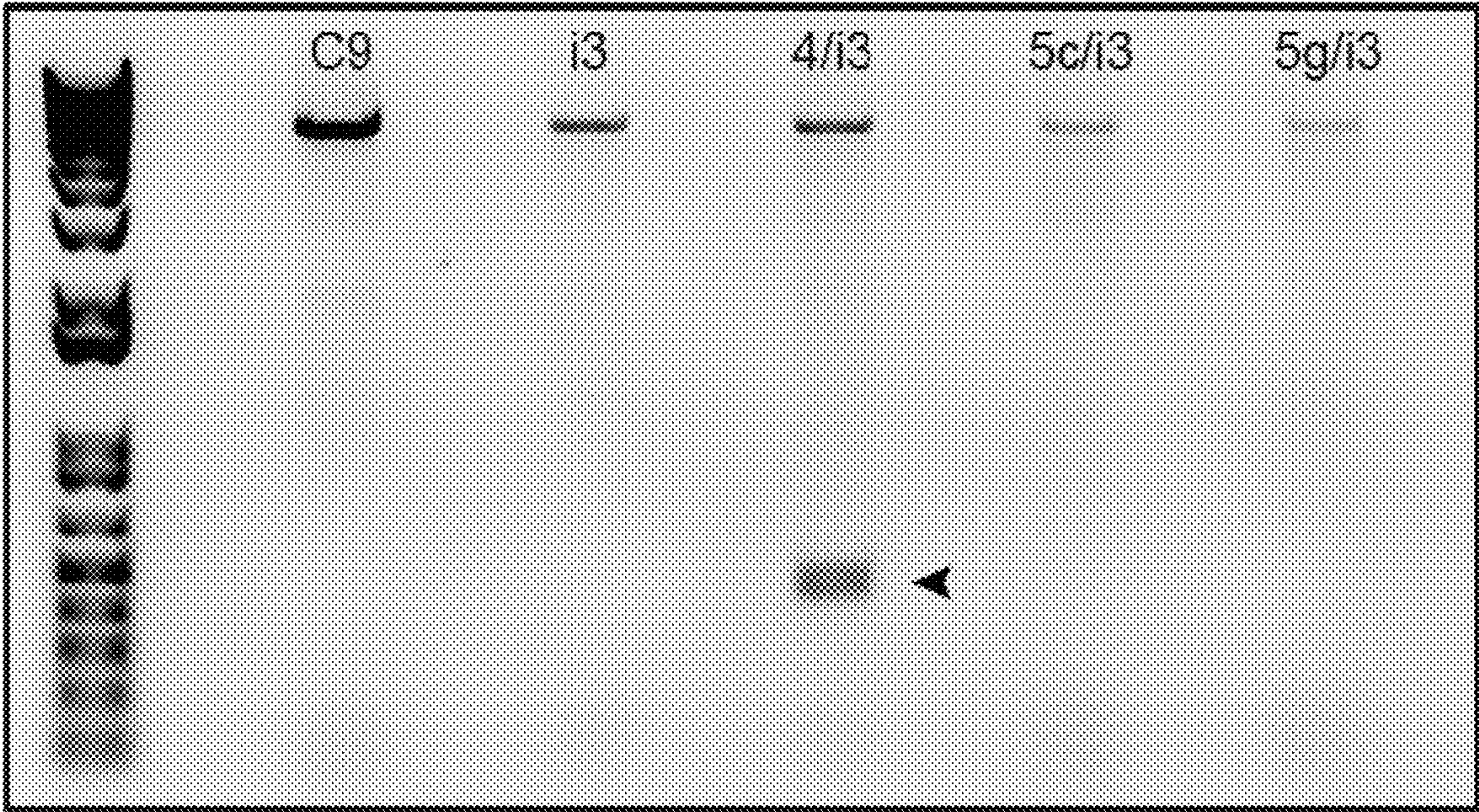


Figure 4F

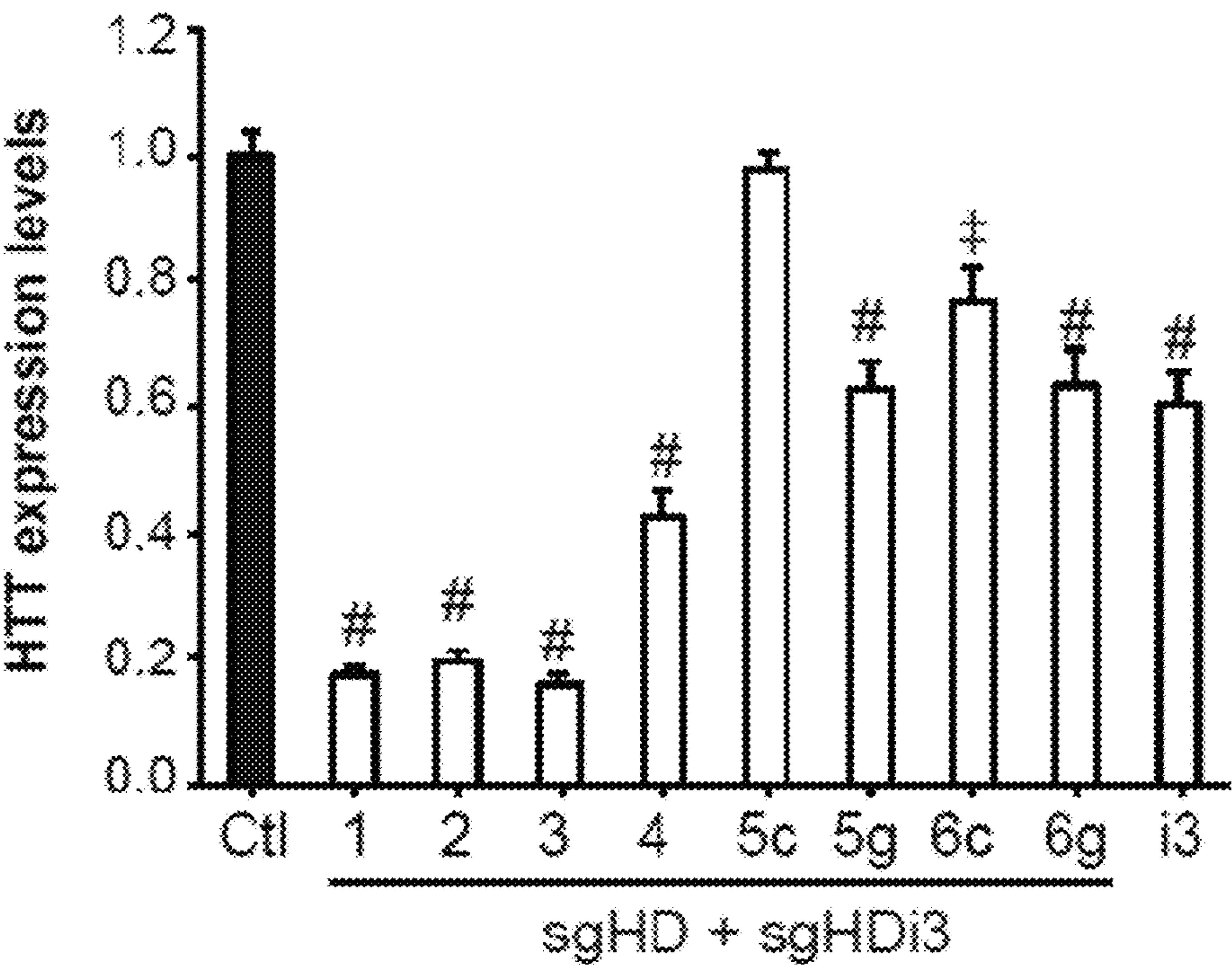


Figure 4G

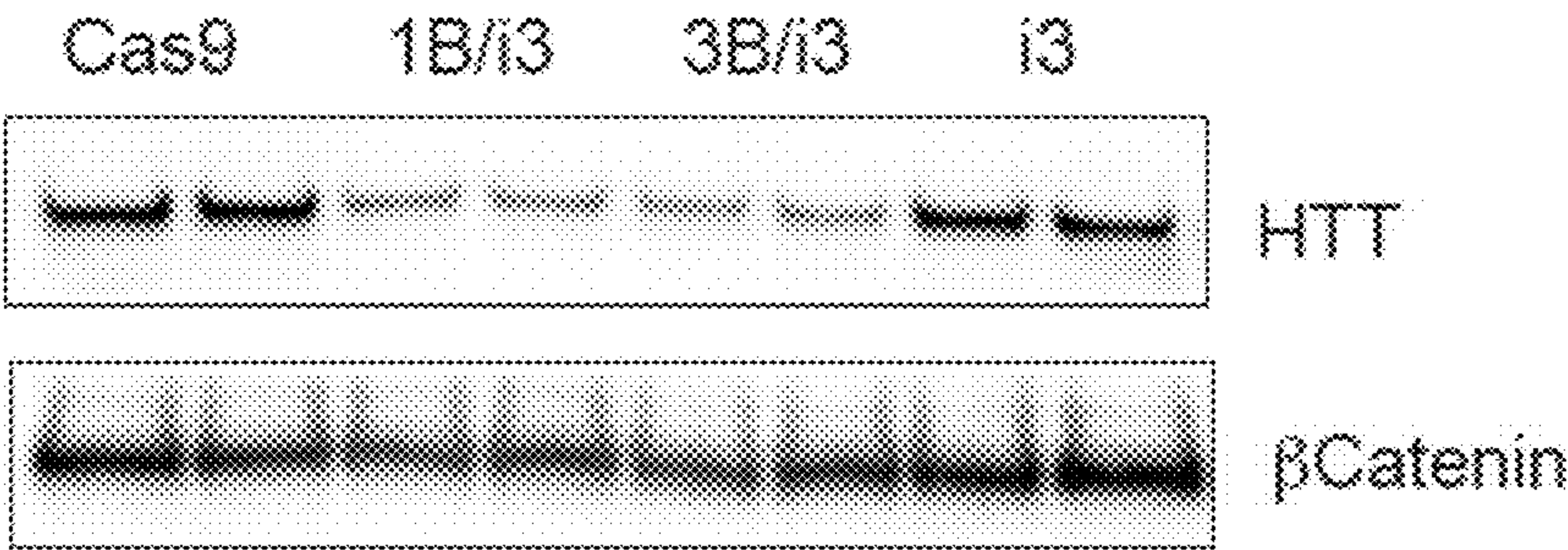


Figure 4H

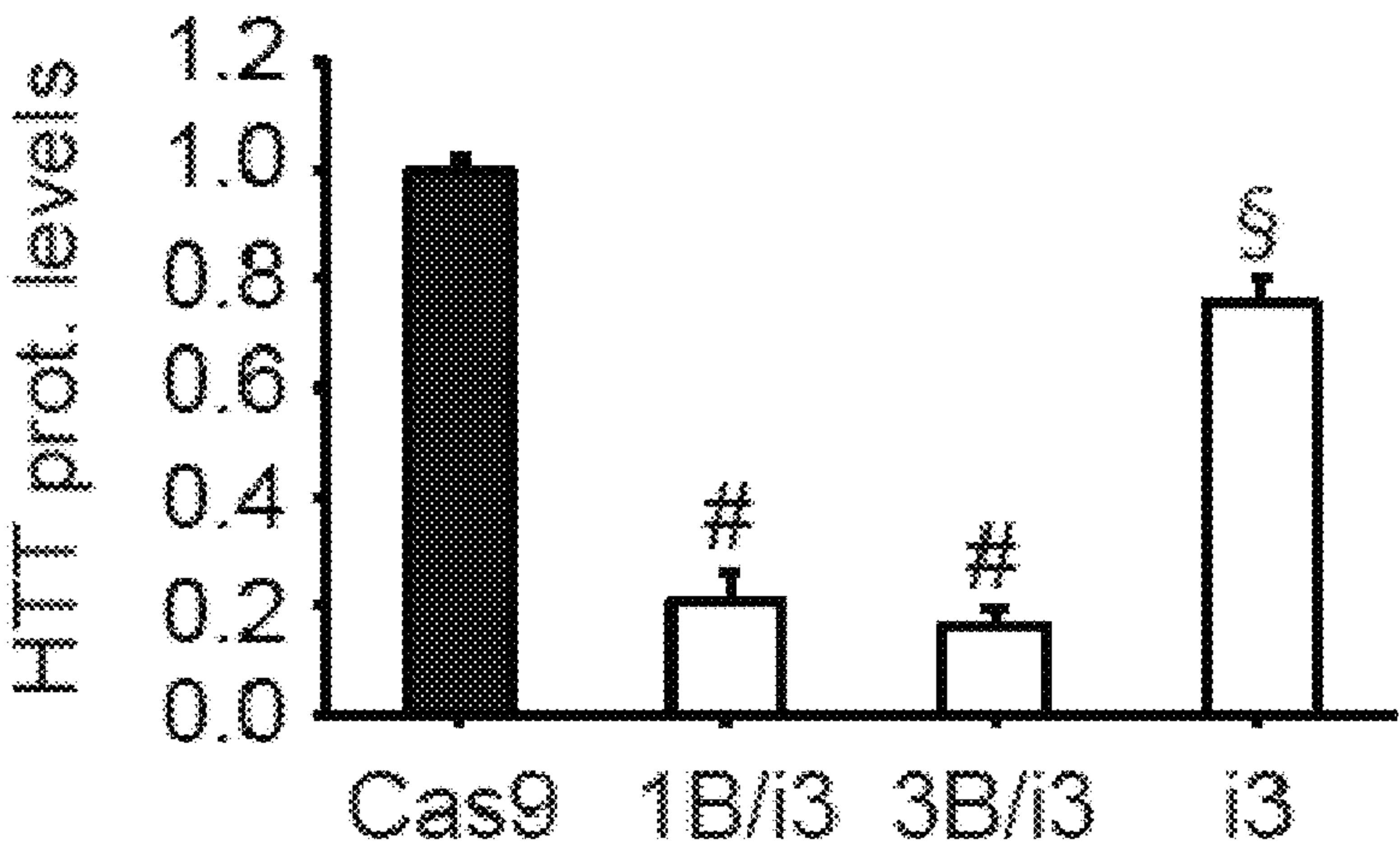


Figure 4I

sgHD1B/i3

(121) CCGCTCAGAGTCCACCGCCGCGCTGTCCGCCCTCCAGGCGCTCCGGGGGCG--//--ATGCTTTTAGGACGCTCCGCGGAGTGGCGCGCGCGCGCGGAGTGAG (122)
 CCGCTCAGAGTCCACCGCCGCGCTGTCCGCCCTCCAGGCGCTCC--CGCGCGGAGTGGCGCGCGCGCGCGCGGAGTGAG (123)
 CCGCTCAGAGTCCACCGCCGCGCTGTCCGCCCTCCAGGCGCTCC--CGCGCGGAGTGGCGCGCGCGCGCGGAGTGAG (124)
 CCGCTCAGAGTCCACCGCGCGC-----GGAGTGGCGCGCGCGCGCGGAGTGAG (125)
 CCGCTCAGAGTCCACCGCGCGCTGTCCGCCCTCCAGGCGCTCC--CGCGCGGAGTGGCGCGCGCGCGGAGTGAG (126)

HD promoter sequence upstream sgHD1B

HD intron 1 sequence downstream sgHD13

sgHD2B/i3

(127) CCGATTACAGTCTCACCACGCGCCGCTGCTCCGCTCCGCTGAGCCCGCGCC--//--ATGCTTTTAGGACGCTCCGCGGAGTGGCGCGCGCGCGGAGTGAG (128)
 CCGATTACAGTCTCACCACGCGCCGCTGCTCCGCTGCTCC-----CGCGCGGAGTGGCGCGCGCGCGCGGAGTGAG (129)
 CCGATTACAGTCTCACCACGCGCCGCTGCTCCGCTGCTCC--TCGCGCGGAGTGGCGCGCGCGCGCGGAGTGAG (130)
 CCGATTACAGTCTCACCACGCGCCGCTGCTCCGCTGCTCC-----CGCGCGGAGTGGCGCGCGCGCGGAGTGAG (131)
 CCGATTACAGTCTCACCACGCGCCGCTGCTCCGCTGCTCC--CGCGCGGAGTGGCGCGCGCGCGGAGTGAG (132)

HD promoter sequence upstream sgHD2B

HD intron 1 sequence downstream sgHD13

sgHD3B/i3

(133) CCGCGCCCGCGCTCCCGCGCGCGCGAGCGCTGCTGCGCGCGCGCGCGCGGAGTGAG (134)
 CCGCGCCCGCGCTCCCGCGCGCGCGAGCGCTGCTGCGCGCGCGCGCGGAGTGAG (135)
 CCGCGCCCGCGCTCCCGCGCGCGCGAGCGCTGCGCGCGCGCGCGGAGTGAG (136)
 CCGCGCCCGCGCTCCCGCGCGCGCGAGCGCTGCGCGCGCGCGCGGAGTGAG (137)
 CCGCGCCCGCGCTCCCGCGCGCGCGAGCGCTGCGCGCGCGCGCGGAGTGAG (138)

HD promoter sequence upstream sgHD3B

HD intron 1 sequence downstream sgHD13

sgHD4/i3

(139) TGGGGTCTGGCTCAGGGTTCTCTCTTTTTCAGGCGCGCGCGCGCGCGGAGTGAG (140)
 TGGGGTCTGGCTCAGGGTTCTCTCTTTTTCAGGCGCGCGCGCGCGGAGTGAG (141)
 TGGGGTCTGGCTCAGGGTTCTCTCTTTTTCAGGCGCGCGCGCGCGGAGTGAG (142)
 TGGGGTCTGGCTCAGGGTTCTCTCTTTTTCAGGCGCGCGCGCGCGGAGTGAG (143)
 TGGGGTCTGGCTCAGGGTTCTCTCTTTTTCAGGCGCGCGCGCGCGGAGTGAG (144)

HD promoter sequence upstream sgHD4

HD intron 1 sequence downstream sgHD13

sgHD6C/i3

(145) CAGGTGTGGCTGGCTAAAGTAGGCTTTTACGCGCTCTCTGCGCTGCTC--//--ATGCTTTTAGGACGCTCCGCGGAGTGGCGCGCGCGCGGAGTGAG (146)
 CAGGTGTGGCTGGCTAAAGTAGGCTTTTACGCGCTCTCTGCGCTGCTC--CGCGCGGAGTGGCGCGCGCGCGCGGAGTGAG (147)
 CAGGTGTGGCTGGCTAAAGTAGGCTTTTACGCGCTCTCTGCGCTGCTC--TCGCGCGGAGTGGCGCGCGCGCGCGGAGTGAG (148)
 CAGGTGTGGCTGGCTAAAGTAGGCTTTTACGCGCTCTCTGCGCTGCTC--CGCGCGGAGTGGCGCGCGCGCGCGGAGTGAG (149)
 CAGGTGTGGCTGGCTAAAGTAGGCTTTTACGCGCTCTCTGCGCTGCTC--CGCGCGGAGTGGCGCGCGCGCGCGGAGTGAG (150)

HD promoter sequence upstream sgHD6C

HD intron 1 sequence downstream sgHD13

FIG. 4J

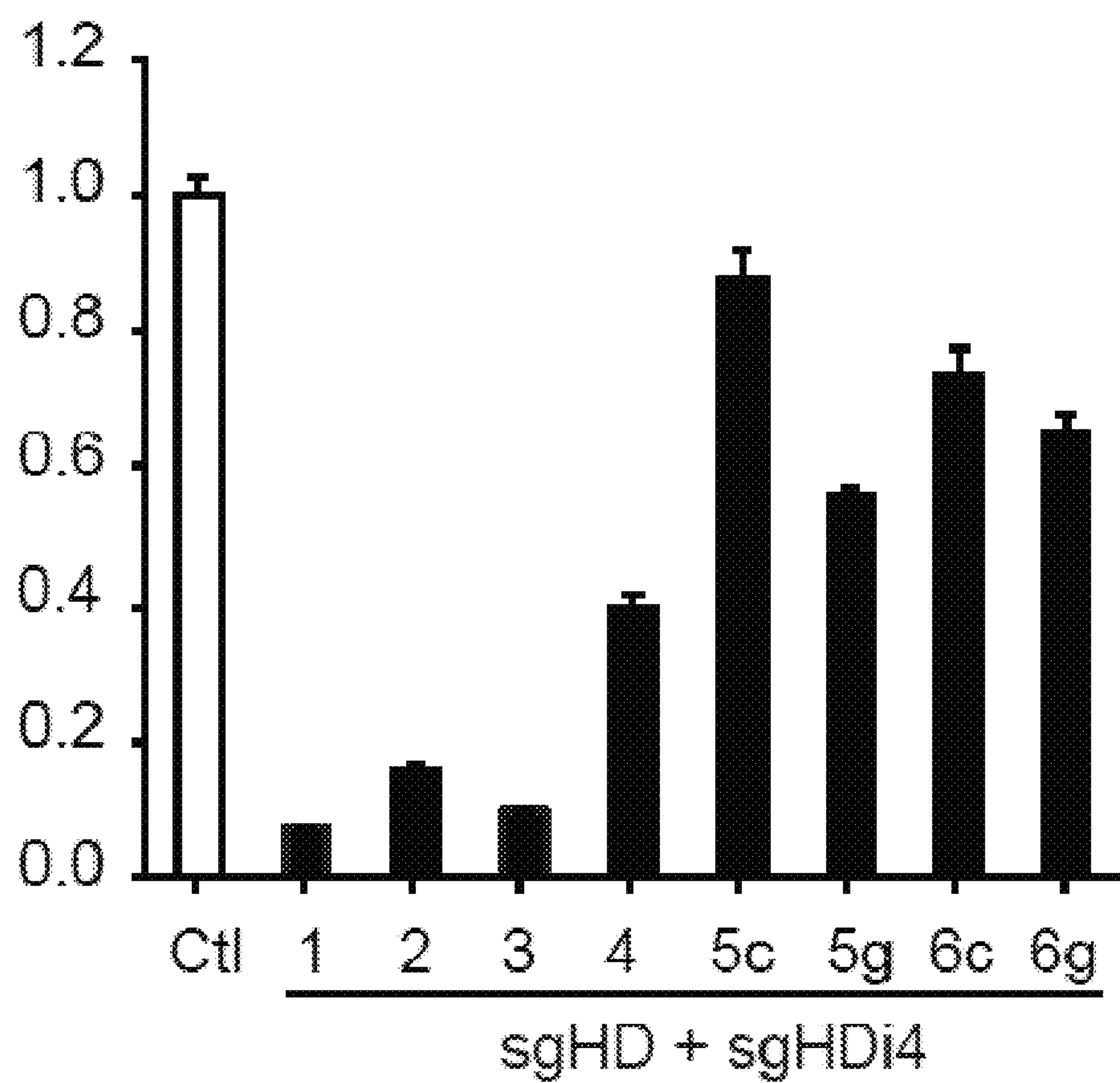
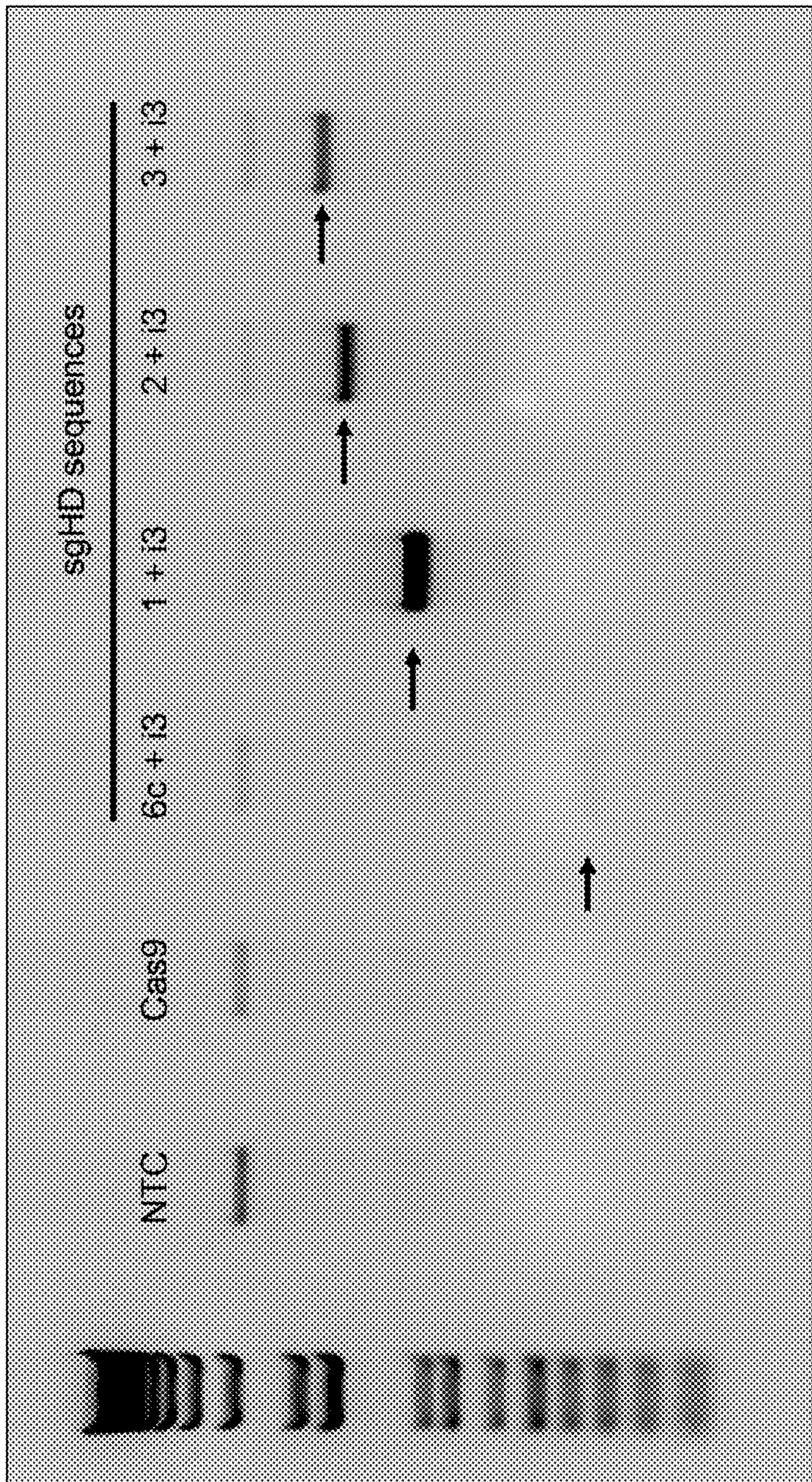
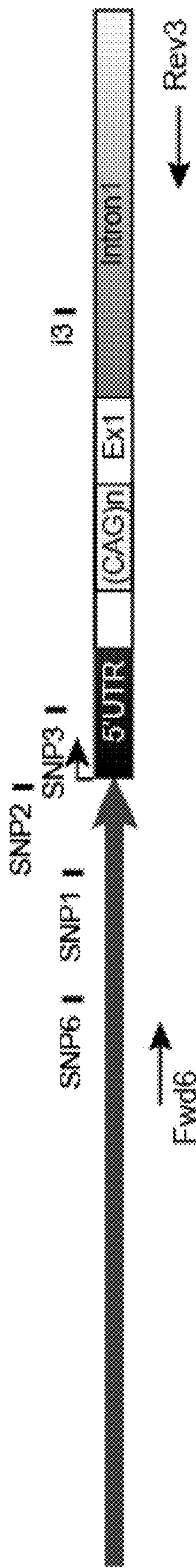
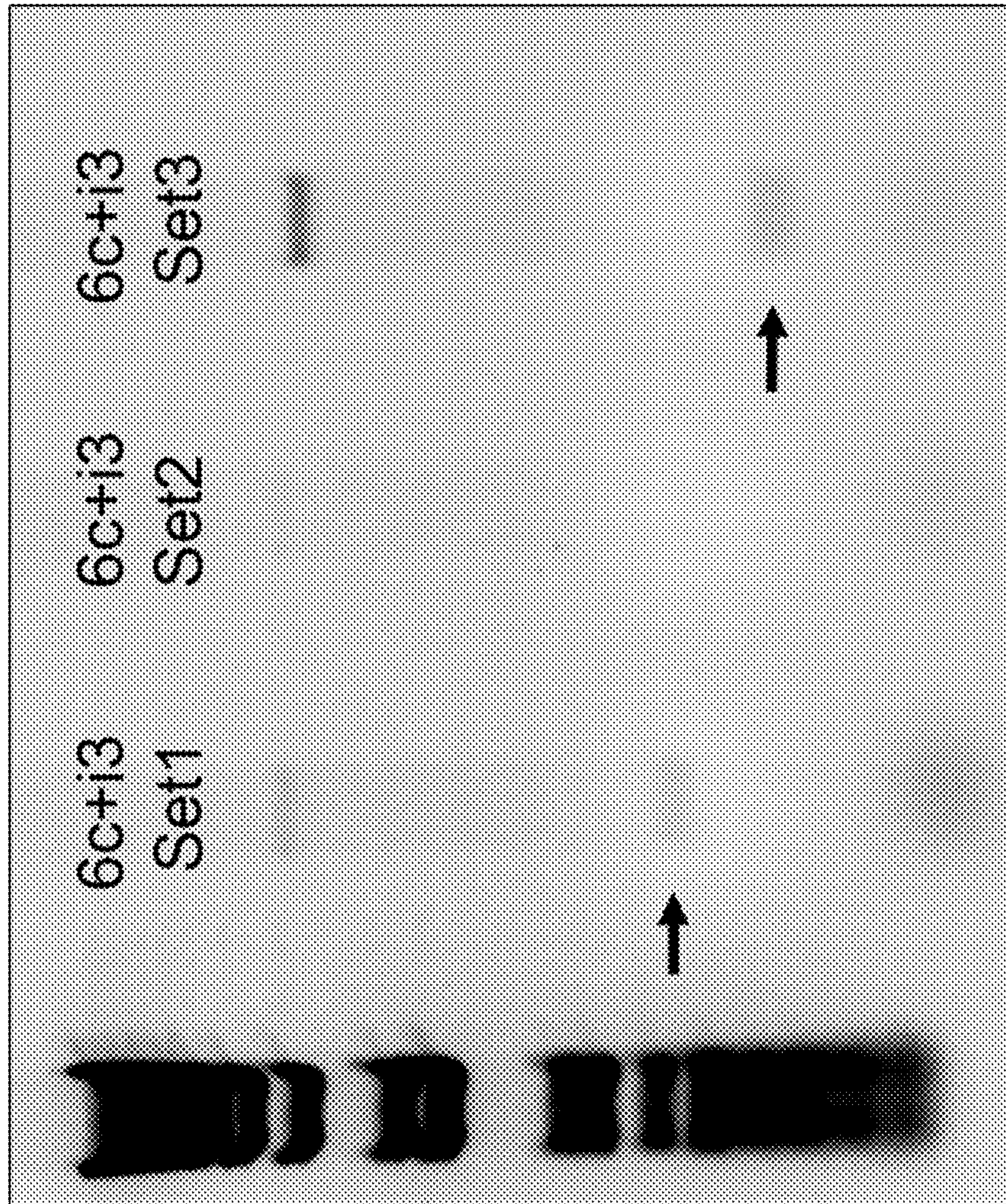
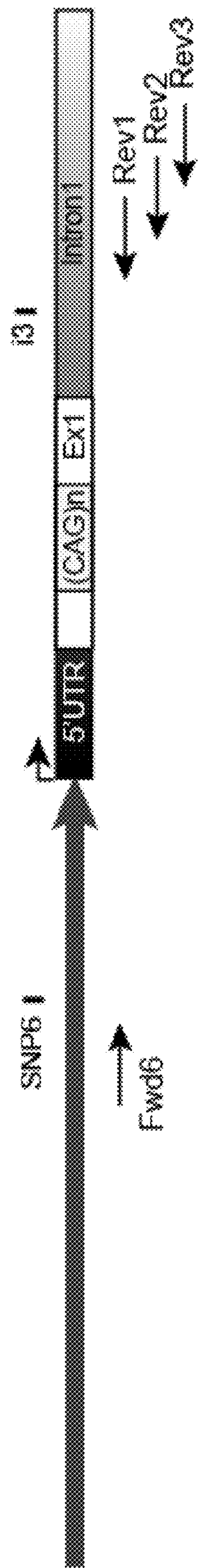


Figure 4K



Expected Bands:
Cas9 Ctl: 3207 bp
6c + i3: 333bp (NHEJ)
1 + i3: 1059bp (NHEJ)
2 + i3: 1549bp (NHEJ)
3 + i3: 1773bp (NHEJ)

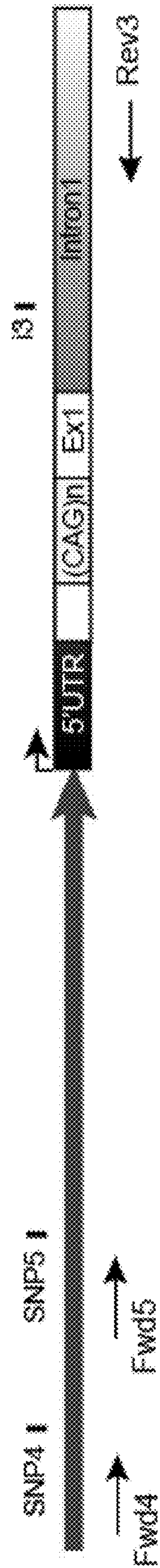
Figure 4L



Expected Bands:

- Cas9 Ctl: 3207 bp
- 6c + i3: 593 bp (NHEJ) Set 1
- 6c + i3: 563 bp (NHEJ) Set 2
- 6c + i3: 333 bp (NHEJ) Set 3

Figure 4M



Expected Bands:

Cas9 Ctl: 6105 bp	Cas9 Ctl: 5050 bp
4 + i3: 421 bp (NHEJ)	5 + i3: 596 bp (NHEJ)

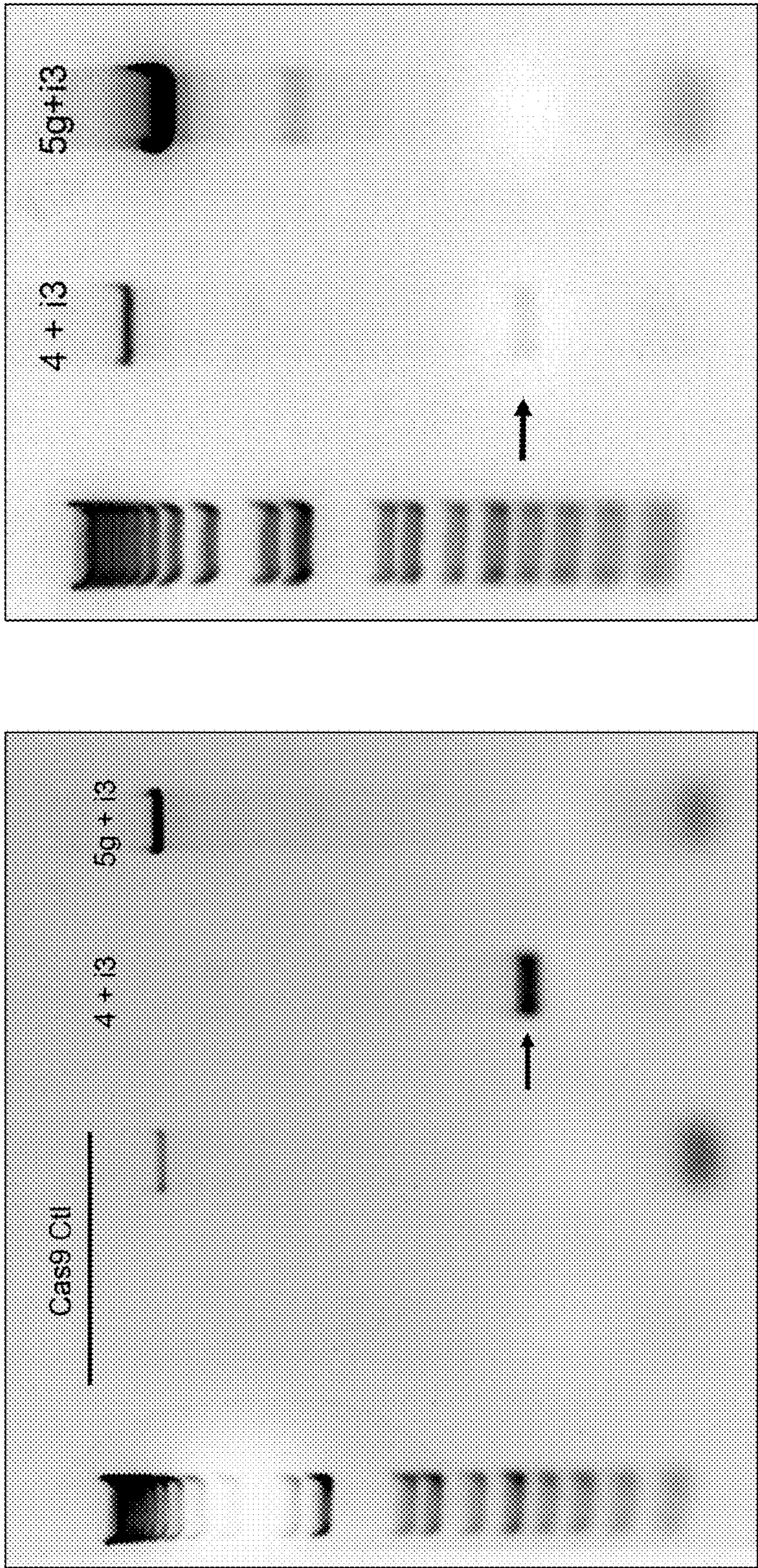


Figure 4N

Minor Allele:

GAGGGTTTCATCTTTGTTGGTCAGGCGAGACTTGAACCTCTGACCTCAGGTGA (247)

CTCCCAAAGTAGAACAACCAGTCCCTGTAACCTGAGGACTGGAGTCCACT (248)

SpCas9_VRER (PAM Motif: NGCG)

Ref. Allele:

PAM

GAGGGTTTCATCTTTGTTGGTCAGGCGAGACTTGAACCTCTGACCTCAGGTGA (245)

CTCCCAAAGTAGAACAACCAGTCCCTGTAACCTGAGGACTGGAGTCCACT (246)

GGTTTCATCTTTGTTGGTCAG (164)

Minor Allele:

GAGGGTTTCATCTTTGTTGGTCAGGCGAGACTTGAACCTCTGACCTCAGGTGA (247)

CTCCCAAAGTAGAACAACCAGTCCCTGTAACCTGAGGACTGGAGTCCACT (248)

SaCas9 (NNGRRR/NNGRR)

Ref. Allele:

PAM

GAGGGTTTCATCTTTGTTGGTCAGGCGAGACTTGAACCTCTGACCTCAGGTGA (245)

CTCCCAAAGTAGAACAACCAGTCCCTGTAACCTGAGGACTGGAGTCCACT (246)

GGTTTCATCTTTGTTGGTCAG (165)

Minor Allele:

GAGGGTTTCATCTTTGTTGGTCAGGCGAGACTTGAACCTCTGACCTCAGGTGA (247)

CTCCCAAAGTAGAACAACCAGTCCCTGTAACCTGAGGACTGGAGTCCACT (248)

# Variant ID	Sequence Variation	Ref. Allele	Min. allele
rs73085145	GCTGTTTTTCTTTAGAAATGGGGGA[C/G]GTTATCAGGCTCTACATGGTGTGTA (81)	C	G

SpCas9_WT (PAM Motif: NRG)

Ref. Allele:

GCTGTTTTTCTTTAGAAATGGGGGAAGTTATCAGGCTCTACATGGTGTGTA (249)

CGACAAAAAAGAAATCTTACCCCTGCAATAGTCCGAGATGTACACACAT (250)

Minor Allele

PAM

GCTGTTTTTCTTTAGAAATGGGGGAAGTTATCAGGCTCTACATGGTGTGTA (251)

CGACAAAAAAGAAATCTTACCCCTGCAATAGTCCGAGATGTACACACAT (252)

GTTTTTCTTTAGAAATGGGG (166)

# Variant ID	Sequence Variation	Ref. Allele	Min. allele
rs61792461	TCTACATGGTGTGTAGTGGGCTAGC[A/G]TGTGTAAAGCCTTTCCCTGTGTAC (82)	A	G

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)

Ref. Allele:

TCTACATGGTGTGTAGTGGGCTAGCATGTTGTAAAGCCTTTCCCTGTGTAC (253)

AGATGTACACACATCAGCCGATCGTACAACATTCCGAAAGGGACACAGTG (254)

Minor Allele

PAM

TCTACATGGTGTGTAGTGGGCTAGCATGTTGTAAAGCCTTTCCCTGTGTAC (255)

AGATGTACACACATCAGCCGATCGTACAACATTCCGAAAGGGACACAGTG (256)

ACATGGTGTGTAGTGGGCTA (167)

SpCas9_VRER (PAM Motif: NGCG)

Ref. Allele:

TCTACATGGTGTGTAGTGGGCTAGCATGTTGTAAAGCCTTTCCCTGTGTAC (253)

AGATGTACACACATCAGCCGATCGTACAACATTCCGAAAGGGACACAGTG (254)

Minor Allele

PAM

TCTACATGGTGTGTAGTGGGCTAGCATGTTGTAAAGCCTTTCCCTGTGTAC (255)

AGATGTACACACATCAGCCGATCGTACAACATTCCGAAAGGGACACAGTG (256)

ACATGGTGTGTAGTGGGCTA (167)

Figure 5C

# Variant ID	Sequence Variation	Ref. Allele	Min. allele
rs61792462	CTGTGTCACAAGTGCCTCATCTCGAAGAGGATTCTAATGACTGCCTGTGGCT (81) C	C	G

SpCas9_WT (PAM Motif: NRG)

Ref. Allele:

CTGTGTCACAAGTGCCTCATCTCGAAGAGGATTCTAATGACTGCCTGTGGCT (257)

GACACAGTGTTCACGAGTAGACCTTCTCCTAAGATTACTGACGGACACCGA (258)

Minor Allele

PAM

CTGTGTCACAAGTGCCTCATCTCGAAGAGGATTCTAATGACTGCCTGTGGCT (259)

GACACAGTGTTCACGAGTAGACCTTCTCCTAAGATTACTGACGGACACCGA (260)

GTCACAAGTGCCTCATCTCG (168)

SpCas9_VQR/QR (PAM Motif: NGRN-NGRN/NGAG)

Ref. Allele:

CTGTGTCACAAGTGCCTCATCTCGAAGAGGATTCTAATGACTGCCTGTGGCT (257)

GACACAGTGTTCACGAGTAGACCTTCTCCTAAGATTACTGACGGACACCGA (258)

Minor Allele

PAM

CTGTGTCACAAGTGCCTCATCTCGAAGAGGATTCTAATGACTGCCTGTGGCT (259)

GACACAGTGTTCACGAGTAGACCTTCTCCTAAGATTACTGACGGACACCGA (260)

GTCACAAGTGCCTCATCTCGA (169)

SpCas9 (NNGRRR/NNGRR)

Ref. Allele:

CTGTGTCACAAGTGCCTCATCTCGAAGAGGATTCTAATGACTGCCTGTGGCT (257)

GACACAGTGTTCACGAGTAGACCTTCTCCTAAGATTACTGACGGACACCGA (258)

Minor Allele

PAM

CTGTGTCACAAGTGCCTCATCTCGAAGAGGATTCTAATGACTGCCTGTGGCT (259)

GACACAGTGTTCACGAGTAGACCTTCTCCTAAGATTACTGACGGACACCGA (260)

GTCACAAGTGCCTCATCTCG (168)

LbCpf1 (TTN/TTCH/CTN/TCN/ATTN/TCCH/...)

Ref. Allele:

CTGTGTCACAAGTGCCTCATCTCGAAGAGGATTCTAATGACTGCCTGTGGCT (257)

GACACAGTGTTCACGAGTAGACCTTCTCCTAAGATTACTGACGGACACCGA (258)

Minor Allele

GACACAGTGTTCACGAGTAGAC (170)

CTGTGTCACAAGTGCCTCATCTCGAAGAGGATTCTAATGACTGCCTGTGGCT (259)

GACACAGTGTTCACGAGTAGACCTTCTCCTAAGATTACTGACGGACACCGA (260)

PAM

# Variant ID	Sequence Variation	Ref. Allele	Min. allele
rs61698955	TCATTTTATGTGATTCCTTTCTAGA[A/G]GTACTACTCACTACTTCTGCTTCCA (84)	A	G

SpCas9_WT (PAM Motif: NRG)

Ref. Allele:

TCATTTTATGTGATTCCTTTCTAGAAGTACTACTCACTACTTCTGCTTCCA (261)

AGTAAATACACTAAGGAAAGATCTTCATGATGATTAATGAAGACCAACGT (262)

Minor Allele

PAM

TCATTTTATGTGATTCCTTTCTAGAAGTACTACTCACTACTTCTGCTTCCA (263)

AGTAAATACACTAAGGAAAGATCTTCATGATGATTAATGAAGACCAACGT (264)

TTTTATGTGATTCCTTTCTA (171)

LbCpf1 (TTN/TTCH/CTN/TCN/ATTN/TCCH/...)

Ref. Allele:

GTAAATACACTAAGGAAAG (172)

TCATTTTATGTGATTCCTTTCTAGAAGTACTACTCACTACTTCTGCTTCCA (261)

AGTAAATACACTAAGGAAAGATCTTCATGATGATTAATGAAGACCAACGT (262)

PAM

Minor Allele

TCATTTTATGTGATTCCTTTCTAGAAGTACTACTCACTACTTCTGCTTCCA (263)

AGTAAATACACTAAGGAAAGATCTTCATGATGATTAATGAAGACCAACGT (264)

Figure 5D

Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs77384845	TAGCTGAAAGGACAGGCGACTGTTC/TATACACTAGCTAAGACGGCAACTTC (85)	C	T

SpCas9_VQR/EQR (FAM NotIc: NGAN-NGNG/NGAG)
Ref. Allele:

REF. A11818:
ATGTTAATCCGATTCTCTGG (173)
TAGCTGAAGGAAGGCACGGACTGTGATACACTAGCTAAGAGGCCAAACTGG (265)
ATCGACTTCTCTCTCTGTAATTCAGCAGTATGTGATGATTCCTCTCTTTTGAGG (266)
PAM
Minor Allele
TAGCTGAAGGAAGGCACGGACTGTGATACACTAGCTAAGAGGCCAAACTGG (267)
ATCGACTTCTCTCTCTGTAATTCAGCAGTATGTGATGATTCCTCTCTTTTGAGG (268)

Variant ID	Sequence Variation	Ref. Allele	Min. allele
rs61792464	TCCCTCATTCAGGTGATGTCTTAA(C/G)CCCCCGAACCTCAGAAATGGGATTGTT (87)	G	C

SpCas2_WT (PAM Motif: NRG)
Ref. Allele:

```

ref. Allele:
GGTUTTGGAGGCTTAACCTA (174)
TCCCTCATTTCAGGTTGATGTCCTTAACCCCCAGAACCTCAGAAATGGGATTGT (269)
ACGGAGTAAGTTCACACTACAGGAATGGGGGGCTCTGGAGATCTTTACCTTAACA (270)
      PAM
Minor Allele
TCCCTCATTTCAGGTTGATGTCCTTAACCCCCAGAACCTCAGAAATGGGATTGT (271)
ACCGACTAACTCCAACTACACCATTCCCGGCTCTTCCACTCTTAAACCTAACA (272)
TCATTTCAGGTTGATGTTCTT (175)

```

LbCnF1 (TTTN/TTCN/CTTN/TCFN/ATFN/TCCN/L)

Ref. Alleles:
TCCCTCAFTTCAGGTTGATGTCCTAAACCCAGAACTCAGAATGGGATGTT (269)
AAGGAGTAACGTCCTAACTACAGGATTTGAGGTTCTCTGGAGTCCTTAACCTAACA (270)

Minor Allele
GGAGTAAGTCCAACTACAGG (176)
TCCCTCATTCAGGTTGATPTCCCTAAACCCAGAACCTCAGAAATGGGATGT (271)
AGGGATAAATCCAACTACAGGATTCGGGGCTCTGGAGATCTTACCCCTAAAC (272)
FAM

# Variant ID	Sequence Variation	Ref.Allele	Min. allele
rs112435598	TCCATGCCGAGAGAGGGGCACAGAGAG G/T GCCAGGGCAGACTCGAAGTTCATACCTT (88)	G	T

SpCas9_WT (FAM Motif: NRG)
Ref. Allele:

TCCATGCCCAAGAAAGGCACACAGACAGGGCCCAAGGGAGACATGAAGTCTATACCTT (273)
 AGCTACGCTTCTTCCGTTGCTCTCTCCCGGCTCCCTCTGACTTCAGTATGCGGA (274)
 ATGCCCAACAAGGCACACAG (177)

Minor Allele
TCCATGCCCAACAAAGGCACACAGAGTCCCAAGGAGACCTGAAGTCTATACCGT (275)
AGCTACGGTTCCTCCGTTGTCTCTCAGCGGTCCCGTCTGACTTCAGTATCGGA (276)

SaCas9 (NNGRR/NNGRR)
Ref. Alleles:

PAM

TCGATGCCAAGAAGGCCAACAGACAGAGCGCCAGGGACACTGAAGTCTATACCCCT (273)
AGGTACGGTTCCTCCCGTGTCTCTCCCGGTCCTCTGACTTCAGTATGGGA (274)
TCCATGCCAAGAAGGCCAACA (178)

Minor Allele

PAM

TCCATGCCCAACAAGGCCAACAGACAGAGTCCAGGGAGACTGAAAGTCATACCCCT (275)

AGGTACGGTCTCTCCGTTTCTCTCTCAGGGTCCCTCTGACTTCAGTATGGGA (276)

TCCATGCCCAACAAGGCCAACA (178)

# Variant ID	Sequence Variation	Ref. Allele	Min. allele
rs111670395	CCTAGGTTCAAGCAATTTTCTTCA1A/G1CCTGTTTAAATAGCTGGGACTACAG (89) G	A	

SpCas9_NT (PAM Notif: NRG)
Ref. Allele:

PAM
 CCGAGGTTCAAGCAATTCCTGCGCTCAGCCTCCGGAATAGCTGGGAATACAGG (277)
 GGGTCCAGGTTGTTAAGACGGAGTTCGGAGGCTTATGACACCTHATSPCC (278)
 GGTTCAAGCAATTCCTGCGCT (179)

Figure 5E

Minor Allele

CCCAGGTTCAAACAATTCTGCTCAACCTCCGGAATAGCTGGGACTACAGG (279)
GGGTCCCAAGTTCTGTTAAGACGGAGTTGGAGGCTTATCGACCTGATGTCC (280)

=====

# Variant ID	Sequence Variation	Ref. Allele	Min. allele
rs2836199	TGTGGGCTGGCTAAAGTAGGCTTAA[C/G]TGGGCTCTCTCTGCTGCTATCACC (91)	C	G

SpCas9_NT (PAM Motif: NRG)

Ref. Allele:

CGGAGGAGAGACGGAG (180)
TGTGGGCTGGCTAAAGTAGGCTTAACTGGGCTCTCTCTGCTGCTATCACC (281)
ACACCGGACCGATTTCATCCCAATGAGCCGAGGACAGACGGAGCTAGTGG (282)
PAM

Minor Allele

PAM
TGTGGGCTGGCTAAAGTAGGCTTAACTGGGCTCTCTCTGCTGCTATCACC (283)
ACACCGGACCGATTTCATCCCAATCAGCCGAGGACAGACGGAGCTAGTGG (284)
GGGCTGGCTAAAGTAGGCTT (181)

=====

# Variant ID	Sequence Variation	Ref. Allele	Min. allele
rs28431418	CTCCCCCAGGGCTGTCCGCTGAG[C/T]ATGGCTCTGCCACGGGCTATGTC (93) T	C	

SaCas9 (NNGRRT/NNGRR)

Ref. Allele:

PAM
CTCCCCCAGGGCTGTCCGCTGAGTATGGCTCTGCCACGGGCTATGTC (285)
GAGGGGCTTCCGACAGGCTTCTATACCGAGACGGCTGCGGCTCAGAC (286)
TCCCCCAGGGCTGTCCG (182)

Minor Allele

PAM
CTCCCCCAGGGCTGTCCGCTGAGTATGGCTCTGCCACGGGCTATGTC (287)
GAGGGGCTTCCGACAGGCTTCTATACCGAGACGGCTGCGGCTCAGAC (288)
TCCCCCAGGGCTGTCCG (182)

=====

# Variant ID	Sequence Variation	Ref. Allele	Min. allele
rs2857935	GGGCGCAGGCCCATGCGGAAGGAT[A/C/G]CCCCCGGAGGCTGGAGCGGGGG (94)	G	C/T (Rev Strand)

SpCas9_NT (PAM Motif: NRG)

Ref. Allele:

GGGCGCTCGGGAUCTCG (183)
GGGCGCAGGCCCATGCGGAAGGATCCCCCGGAGGCTGGAGCGGGGG (289)
CCCCGCTCGGCTAGGCTTTCTTAGGGGGGGGCTGGGACCTGCCCCGC (290)
PAM

Minor Allele

GGGCGCAGGCCCATGCGGAAGGATCCCCCGGAGGCTGGAGCGGGGG (291)
CCCCGCTCGGCTAGGCTTTCTTAGGGGGGGGCTGGGACCTGCCCCGC (292)

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)

Ref. Allele:

GGGCGGCTCGGGAUCTCG (183)
GGGCGCAGGCCCATGCGGAAGGATCCCCCGGAGGCTGGAGCGGGGG (289)
CCCCGCTCGGCTAGGCTTTCTTAGGGGGGGGCTGGGACCTGCCCCGC (290)
PAM

Minor Allele

GGGCGCAGGCCCATGCGGAAGGATCCCCCGGAGGCTGGAGCGGGGG (291)
CCCCGCTCGGCTAGGCTTTCTTAGGGGGGGGCTGGGACCTGCCCCGC (292)

SaCas9 (NNGRRT/NNGRR)

Ref. Allele:

GGCTCGGGAUCTGCCCCG (184)
GGGCGCAGGCCCATGCGGAAGGATCCCCCGGAGGCTGGAGCGGGGG (289)
CCCCGCTCGGCTAGGCTTTCTTAGGGGGGGGCTGGGACCTGCCCCGC (290)
PAM

Minor Allele

GGCTCGGGAUCTGCCCCG (184)
GGGCGCAGGCCCATGCGGAAGGATCCCCCGGAGGCTGGAGCGGGGG (293)
CCCCGCTCGGCTAGGCTTTCTTAGGGGGGGGCTGGGACCTGCCCCGC (294)
PAM

Figure 5F

Variant ID Sequence Variation Ref. Allele Min. allele
rs113331544 GGCCTTGTCTGTGTGAGGCAGAACCT[~/GCGGGG]GCGGGGGCAGGGGGCGGGCTGGTTCC (98) - inaGCGGGG

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)
Ref. Allele:

PAM

GGCCTTGTCTGTGTGAGGCAGAACCTGCGGGGGCAGGGGGCGGGCTGGTTCC (303)
CCGGAACGACACACTCCCTCTTGGACGCCCCCGCTCCCCCGGACCAAGG (304)
GCTGTGTGAGGCAGAACCT (190)
Minor Allele

PAM

GGCCTTGTCTGTGTGAGGCAGAACCTGCGGGGGCAGGGGGCGGGCTGGTTCC (305)
CCGGAACGACACACTCCCTCTTGGACGCCCCCGCTCCCCCGGACCAAGG (306)
GTGAGGCAGAACCTGCGGGG (191)

SpCas9_VRRR (PAM Motif: NGCG)
Ref. Allele:

PAM

GGCCTTGTCTGTGTGAGGCAGAACCTGCGGGGGCAGGGGGCGGGCTGGTTCC (303)
CCGGAACGACACACTCCCTCTTGGACGCCCCCGCTCCCCCGGACCAAGG (304)
GCTGTGTGAGGCAGAACCT (190)
Minor Allele

PAM

GGCCTTGTCTGTGTGAGGCAGAACCTGCGGGGGCAGGGGGCGGGCTGGTTCC (305)
CCGGAACGACACACTCCCTCTTGGACGCCCCCGCTCCCCCGGACCAAGG (306)
GTGAGGCAGAACCTGCGGGG (191)

Variant ID Sequence Variation Ref. Allele Min. allele
rs13132932 TGTGTGAGGCAGAACCTGCGGGGGC[A/G]GCGGGGGCTGGTTCCCTGGCCAGC (99) A G

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)
Ref. Allele:

TGTGTGAGGCAGAACCTGCGGGGGCAGGGGGCGGGCTGGTTCCCTGGCCAGC (307)
ACACACTCCCTCTTGGACGCCCCCGCTCCCCCGGACCAAGGGACCGGTCC (308)
Minor Allele

PAM

TGTGTGAGGCAGAACCTGCGGGGGCAGGGGGCGGGCTGGTTCCCTGGCCAGC (309)
ACACACTCCCTCTTGGACGCCCCCGCTCCCCCGGACCAAGGGACCGGTCC (310)
GAGGCAGAACCTGCGGGG (192)

SpCas9_VRRR (PAM Motif: NGCG)
Ref. Allele:

TGTGTGAGGCAGAACCTGCGGGGGCAGGGGGCGGGCTGGTTCCCTGGCCAGC (307)
ACACACTCCCTCTTGGACGCCCCCGCTCCCCCGGACCAAGGGACCGGTCC (308)
Minor Allele

PAM

TGTGTGAGGCAGAACCTGCGGGGGCAGGGGGCGGGCTGGTTCCCTGGCCAGC (309)
ACACACTCCCTCTTGGACGCCCCCGCTCCCCCGGACCAAGGGACCGGTCC (310)
GAGGCAGAACCTGCGGGG (192)

SaCas9 (NNSRRR/NNRRR)
Ref. Allele:

TGTGTGAGGCAGAACCTGCGGGGGCAGGGGGCGGGCTGGTTCCCTGGCCAGC (307)
ACACACTCCCTCTTGGACGCCCCCGCTCCCCCGGACCAAGGGACCGGTCC (308)
Minor Allele

PAM

TGTGTGAGGCAGAACCTGCGGGGGCAGGGGGCGGGCTGGTTCCCTGGCCAGC (309)
ACACACTCCCTCTTGGACGCCCCCGCTCCCCCGGACCAAGGGACCGGTCC (310)
GAGGCAGAACCTGCGGGG (192)

Variant ID Sequence Variation Ref. Allele Min. allele
rs13102260 AGCGTCTGGGACGCAGGGGGCGGTG[A/G]GGGCTGCGGGGACGGGTCCAGATG G A
(100)

SpCas9_WT (PAM Motif: NRG)
Ref. Allele:

PAM

AGCGTCTGGGACGCAGGGGGCGGTGAGGGGGACGGGTCCAGATG (311)
TGGCAGACCTGCGGTTCGGGGGAGGGGGACGGGTCCAGATG (312)
GTCTGGGACCGAAGGCGCG (193)

Figure 5H

Figure 51

Minor Allele
TGAATGAGTTGTGGTTGCCAAGTAAGTGGTGAACCTTACCTGGTGATTAA (321)
ACTTACTCAGACACCAACGGTTTATTCACCACTTGAATGCCACCACTAATTA (322)

AsCpf1(TTN)
Ref. Allele:
CTTACTCAGACACCAACGGTTC (197)
TGAATGAGTTGTGGTTGCCAAGTAAGTGGTGAACCTTACCTGGTGATTAA (319)
ACTTACTCAGACACCAACGGTTTATTCACCACTTGAATGCCACCACTAATTA (320)
PAM

Minor Allele
TGAATGAGTTGTGGTTGCCAAGTAAGTGGTGAACCTTACCTGGTGATTAA (321)
ACTTACTCAGACACCAACGGTTCATTCACCACTTGAATGCCACCACTAATTA (322)

=====

# Variant ID	Sequence Variation	Ref. Allele	Min. allele
rs28571971	ACAGTAGGAGTTAGGAAGTACTCTG(C/G)TGAATTCAGGCTTTCTCTTACCT (105)	C	

SpCas9_WT (PAM Motif: NRG)
Ref. Allele:
ACAGTAGGAGTTAGGAAGTACTCTGCTCCAGTTCAGGCTTTCTCTTACCT (323)
TGTCACTCCTCAATCCTTCATGAGAGGACCTTCAAGTCCGGAAAGAGAAATGA (324)
GGAGTTAGGAAGTACTC (198)
Minor Allele:
GTCAAGTCCGGAAAGAGAAATGA (199)
ACAGTAGGAGTTAGGAAGTACTCTGCTCCAGTTCAGGCTTTCTCTTACCT (325)
TGTCACTCCTCAATCCTTCATGAGAGGACCTTCAAGTCCGGAAAGAGAAATGA (326)
PAM

=====

# Variant ID	Sequence Variation	Ref. Allele	Min. allele
rs28583447	CAGTAGGAGTTAGGAAGTACTCTGG(C/T)TGAATTCAGGCTTTCTCTTACCT (106)	T	C

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)
Ref. Allele:
CAGTAGGAGTTAGGAAGTACTCTGGTCCAGTTCAGGCTTTCTCTTACCTC (327)
GTCATCCTCAATCCTTCATGAGAGGACCTCAAGTCCGGAAAGAGAAATGGAG (328)
Minor Allele:
TCAAGTCCGGAAAGAGAAATGA (200)
CAGTAGGAGTTAGGAAGTACTCTGGTCCAGTTCAGGCTTTCTCTTACCTC (329)
GTCATCCTCAATCCTTCATGAGAGGACCTCAAGTCCGGAAAGAGAAATGGAG (330)
PAM

SpCas9_VRR (PAM Motif: NGCG)
Ref. Allele:
CAGTAGGAGTTAGGAAGTACTCTGGTCCAGTTCAGGCTTTCTCTTACCTC (327)
GTCATCCTCAATCCTTCATGAGAGGACCTCAAGTCCGGAAAGAGAAATGGAG (328)
Minor Allele:
PAM
CAGTAGGAGTTAGGAAGTACTCTGGTCCAGTTCAGGCTTTCTCTTACCTC (329)
GTCATCCTCAATCCTTCATGAGAGGACCTCAAGTCCGGAAAGAGAAATGGAG (330)
GGAGTTAGGAAGTACTCTG (201)
Minor Allele:
TCAAGTCCGGAAAGAGAAATGA (200)
CAGTAGGAGTTAGGAAGTACTCTGGTCCAGTTCAGGCTTTCTCTTACCTC (329)
GTCATCCTCAATCCTTCATGAGAGGACCTCAAGTCCGGAAAGAGAAATGGAG (330)
PAM

=====

# Variant ID	Sequence Variation	Ref. Allele	Min. allele
rs28468636	TCTCTTACCTCTCAGTATTCTATTT(C/G)CGATCTGGATCTGTCCAGATGGCA (107)	C	G

SpCas9_WT (PAM Motif: NRG)
Ref. Allele:
TAGACCTACACAGGGTCTA (202)
TCTCTTACCTCTCAGTATTCTATTTCCGATCTGGATCTGTCCAGATGGCA (331)
AGAGAAATGGAGGTCATAAGATAAAGGCTAGACCTACACAGGGTCTACCGT (332)
PAM

Minor Allele
TCTCTTACCTCTCAGTATTCTATTTCCGATCTGGATCTGTCCAGATGGCA (333)
AGAGAAATGGAGGTCATAAGATAAAGGCTAGACCTACACAGGGTCTACCGT (334)

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)
Ref. Allele:

Figure 5J

TCTCTTACCTCTCAGTATTCTATTTCCGATCTGGATGTGTCCCAGATGGCA (331)
AGAGAATGGAGAGTCATAAGATAAAGGCTAGACCTACACAGGGTCTACCGT (332)

Minor Allele PAM
TCTCTTACCTCTCAGTATTCTATTTCCGATCTGGATGTGTCCCAGATGGCA (333)
AGAGAATGGAGAGTCATAAGATAAAGGCTAGACCTACACAGGGTCTACCGT (334)
ACCTCTCAGTATTCTATTT (203)

SpCas9_VRER (PAM Motif: NGCG)
Ref. Allele:
TCTCTTACCTCTCAGTATTCTATTTCCGATCTGGATGTGTCCCAGATGGCA (331)
AGAGAATGGAGAGTCATAAGATAAAGGCTAGACCTACACAGGGTCTACCGT (332)

Minor Allele PAM
TCTCTTACCTCTCAGTATTCTATTTCCGATCTGGATGTGTCCCAGATGGCA (333)
AGAGAATGGAGAGTCATAAGATAAAGGCTAGACCTACACAGGGTCTACCGT (334)
ACCTCTCAGTATTCTATTT (203)

SpCas9 (NNGRRRT/NNGRR):
Ref. Allele: TAGACCTACACAGGGTCTA (202)
TCTCTTACCTCTCAGTATTCTATTTCCGATCTGGATGTGTCCCAGATGGCA (331)
AGAGAATGGAGAGTCATAAGATAAAGGCTAGACCTACACAGGGTCTACCGT (332)
PAM

Minor Allele
TCTCTTACCTCTCAGTATTCTATTTCCGATCTGGATGTGTCCCAGATGGCA (333)
AGAGAATGGAGAGTCATAAGATAAAGGCTAGACCTACACAGGGTCTACCGT (334)

LbCpf1 (TTTN/TTCH/CTTN/TCIN/ATTN/TCCN/...):

Ref. Allele: PAM
TCTCTTACCTCTCAGTATTCTATTTCCGATCTGGATGTGTCCCAGATGGCA (331)
AGAGAATGGAGAGTCATAAGATAAAGGCTAGACCTACACAGGGTCTACCGT (332)
ATCTGGATGTGTCCCAGATG (204)

Minor Allele
TCTCTTACCTCTCAGTATTCTATTTCCGATCTGGATGTGTCCCAGATGGCA (333)
AGAGAATGGAGAGTCATAAGATAAAGGCTAGACCTACACAGGGTCTACCGT (334)

Variant ID Sequence Variation Ref. Allele Min. allele
rs28564368 TATTCTATTTCCGATCTGGATGTGT{A/C}CCAGATGGCATTTGGTAAGAATATC C A
(108)

SpCas9_WT (PAM Motif: NRS)
Ref. Allele: TCTACCGTAAACCATTCCTA (205)
TATTCTATTTCCGATCTGGATGTGTCCCAGATGGCATTTGGTAAGAATATC (335)
ATAAGATAAAGGCTAGACCTACACAGGGTCTACCGTAAACCATTCCTATAG (336)
PAM

Minor Allele
TATTCTATTTCCGATCTGGATGTGTACCAGATGGCATTTGGTAAGAATATC (337)
ATAAGATAAAGGCTAGACCTACACAGGGTCTACCGTAAACCATTCCTATAG (338)

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)
Ref. Allele: GTCTACCGTAAACCATTCCT (206)
TATTCTATTTCCGATCTGGATGTGTCCCAGATGGCATTTGGTAAGAATATC (335)
ATAAGATAAAGGCTAGACCTACACAGGGTCTACCGTAAACCATTCCTATAG (336)
PAM

Minor Allele
TATTCTATTTCCGATCTGGATGTGTACCAGATGGCATTTGGTAAGAATATC (337)
ATAAGATAAAGGCTAGACCTACACAGGGTCTACCGTAAACCATTCCTATAG (338)

Ref. Allele: TCTACCGTAAACCATTCCTA (205)
TATTCTATTTCCGATCTGGATGTGTCCCAGATGGCATTTGGTAAGAATATC (335)
ATAAGATAAAGGCTAGACCTACACAGGGTCTACCGTAAACCATTCCTATAG (336)
PAM

Minor Allele
TATTCTATTTCCGATCTGGATGTGTACCAGATGGCATTTGGTAAGAATATC (337)
ATAAGATAAAGGCTAGACCTACACAGGGTCTACCGTAAACCATTCCTATAG (338)

LbCpf1 (TTTN/TTCH/CTTN/TCIN/ATTN/TCCN/...):

Ref. Allele: PAM
TATTCTATTTCCGATCTGGATGTGTCCCAGATGGCATTTGGTAAGAATATC (335)
ATAAGATAAAGGCTAGACCTACACAGGGTCTACCGTAAACCATTCCTATAG (336)
AGATGGCATTTGGTAAGAATAT (207)

Minor Allele
TATTCTATTTCCGATCTGGATGTGTACCAGATGGCATTTGGTAAGAATATC (337)
ATAAGATAAAGGCTAGACCTACACAGGGTCTACCGTAAACCATTCCTATAG (338)

Figure 5K

=====			
Variant ID	Sequence Variation	Ref. Allele	Min. allele
rs77173925	GAATAAATTATTCTAAAGGATGGAA{A/G}AACTTTTGGATATTTGGAGAAATT	A	G
	(110)		
SpCas9_WT (PAM Motif: <u>NRG</u>)			
Ref. Allele:			
GAATAAATTATTCTAAAGGATGGAAAAACCTTTTGGATATTTGGAGAAATT (339)			
CTTATTTAATAAGATTTCCTACCTTTTGA AAAACCTATAAACCTCTTTAA (340)			
Minor Allele			
PAM			
GAATAAATTATTCTAAAGGATGGAA <u>AA</u> AACTTTTGGATATTTGGAGAAATT (341)			
CTTATTTAATAAGATTTCCTACCTTTTGA AAAACCTATAAACCTCTTTAA (342)			
AAATTATTCTAAAGGATGG (208)			
SpCas9_VQR/EQR (PAM Motif: <u>NGAN-NGNG/NGAG</u>)			
Ref. Allele:			
GAATAAATTATTCTAAAGGATGGAAAAACCTTTTGGATATTTGGAGAAATT (339)			
CTTATTTAATAAGATTTCCTACCTTTTGA AAAACCTATAAACCTCTTTAA (340)			
Minor Allele			
PAM			
GAATAAATTATTCTAAAGGATGGAA <u>AA</u> AACTTTTGGATATTTGGAGAAATT (341)			
CTTATTTAATAAGATTTCCTACCTTTTGA AAAACCTATAAACCTCTTTAA (342)			
ATTATTCTAAAGGATGGAA (209)			
SaCas9 (NNGRRR/ <u>NNGRR</u>):			
Ref. Allele:			
GAATAAATTATTCTAAAGGATGGAAAAACCTTTTGGATATTTGGAGAAATT (339)			
CTTATTTAATAAGATTTCCTACCTTTTGA AAAACCTATAAACCTCTTTAA (340)			
Minor Allele			
PAM			
GAATAAATTATTCTAAAGGATGGAA <u>AA</u> AACTTTTGGATATTTGGAGAAATT (341)			
CTTATTTAATAAGATTTCCTACCTTTTGA AAAACCTATAAACCTCTTTAA (342)			
AATTATTCTAAAGGATGG (210)			
AsCpf1 (TTTN)			
Ref. Allele:			
ATTTAATAAGATTTCCTAC (211)			
GAATAAATTATTCTAAAGGATGGAAAAACCTTTTGGATATTTGGAGAAATT (339)			
CTTATTTAATAAGATTTCCTACCTTTTGA AAAACCTATAAACCTCTTTAA (340)			
PAM			
Minor Allele			
GAATAAATTATTCTAAAGGATGGAA <u>AA</u> AACTTTTGGATATTTGGAGAAATT (341)			
CTTATTTAATAAGATTTCCTACCTTTTGA AAAACCTATAAACCTCTTTAA (342)			
=====			
Variant ID	Sequence Variation	Ref. Allele	Min. allele
rs3965238	TTGTATCATGTCAATGTATTACTTA{C/T}GCAAAAATAATACATTAAAAAAAT	A	G (Rev Strand)
	(111)		
SpCas9_VQR/EQR (PAM Motif: <u>NGAN-NGNG/NGAG</u>)			
Ref. Allele:			
TTGTATCATGTCAATGTATTACTTATGCAAAAATAATACATTAAAAAAAT (343)			
AACATAGTACAGTTACATAATGAATACGTTTTTATTATGTAATTTTTTTTA (344)			
Minor Allele			
TTTTATTATGTAATTTTTTT (212)			
TTGTATCATGTCAATGTATTACTTATGCAAAAATAATACATTAAAAAAAT (345)			
AACATAGTACAGTTACATAATGAATGCGTTTTTATTATGTAATTTTTTTTA (346)			
PAM			
SpCas9_VRER (PAM Motif: <u>NGCG</u>)			
Ref. Allele:			
TTGTATCATGTCAATGTATTACTTATGCAAAAATAATACATTAAAAAAAT (343)			
AACATAGTACAGTTACATAATGAATACGTTTTTATTATGTAATTTTTTTTA (344)			
Minor Allele			
TTTTATTATGTAATTTTTTT (212)			
TTGTATCATGTCAATGTATTACTTATGCAAAAATAATACATTAAAAAAAT (345)			
AACATAGTACAGTTACATAATGAATGCGTTTTTATTATGTAATTTTTTTTA (346)			
PAM			

Figure 5L

```

# Variant ID      Sequence Variation      Ref. Allele      Min. allele
rs33950438  AGGTATTCACATAATTTTGAAGTAAACA[-/AACA]CTGCTCACAAGTTTGGATTTTGGC  AACA  delAACA
                                                    (112)

LbCpf1(TTIN/TTDN/CTTN/TCIN/ATTN/TCEN/...):
Ref. Allele:
ATAAGTGATTAAAACTCATT (213)
AGGTATTCACATAATTTTGAAGTAAACAAGTCTGCTCACAAGTTTGGATTTTGGC (347)
TCCATAAGTGATTAAAACTCATTGTTTGTGACGAGTGTTCCTCAACCTAAAACCG (348)
PAM

Minor Allele
AGGTATTCACATAATTTTGAAGTAAACAAGTCTGCTCACAAGTTTGGATTTTGGC (349)
TCCATAAGTGATTAAAACTCATTGTTTGTGACGAGTGTTCCTCAACCTAAAACCG (350)

AsCpf1(TTTN)
Ref. Allele:
ATAAGTGATTAAAACTCATT (213)
AGGTATTCACATAATTTTGAAGTAAACAAGTCTGCTCACAAGTTTGGATTTTGGC (347)
TCCATAAGTGATTAAAACTCATTGTTTGTGACGAGTGTTCCTCAACCTAAAACCG (348)
PAM

Minor Allele
AGGTATTCACATAATTTTGAAGTAAACAAGTCTGCTCACAAGTTTGGATTTTGGC (349)
TCCATAAGTGATTAAAACTCATTGTTTGTGACGAGTGTTCCTCAACCTAAAACCG (350)

=====

# Variant ID      Sequence Variation      Ref. Allele      Min. allele
rs28377140  AGGCAATTAACTCTGCTCTGCGCA[C/G]TTTCTTATTCTCCTTCAGATTCTTA  G  C
                                                    (113)

SpCas9_WT (PAM Motif: NRG)
Ref. Allele:      PAM
AGGCAATTAACTCTGCTCTGCGCACTTTCTTATTCTCCTTCAGATTCTTA (351)
TCCGTTAATTATGAAAGGAGACCGTCAAGCAATAAGAGGAAGTCTAAGGAT (352)
ATTAACTCTGCTCTGCG (214)

Minor Allele
(215) AGAATAAGAGGAAGTCTAAGG
AGGCAATTAACTCTGCTCTGCGCACTTTCTTATTCTCCTTCAGATTCTTA (353)
TCCGTTAATTATGAAAGGAGACCGTCAAGCAATAAGAGGAAGTCTAAGGAT (354)
PAM

=====

# Variant ID      Sequence Variation      Ref. Allele      Min. allele
rs3856973  ttacaaataaaaaataaGTTAACT[C/T]GATTAACCTGACATTTCCCTATCC  G  A (Rev strand)
                                                    (114)

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)

Ref. Allele:      CTAATGCGGACTGTAAAGGG (216)
TTAAAAATAAAAAATAAGTTAACTCTGATTAACCTGACATTTCCCTATCC (355)
AATTTTATTTTATTTCAATTGTTGAGCTAATTGGGACTGTAAAGGGATAGG (356)
PAM

Minor Allele
TTAAAAATAAAAAATAAGTTAACTCTGATTAACCTGACATTTCCCTATCC (357)
AATTTTATTTTATTTCAATTGTTGAGCTAATTGGGACTGTAAAGGGATAGG (358)

BsCas9(HNGRRT/HNGRR):
Ref. Allele:      AATTTGGGACTGTAAAGGG (217)
TTAAAAATAAAAAATAAGTTAACTCTGATTAACCTGACATTTCCCTATCC (355)
AATTTTATTTTATTTCAATTGTTGAGCTAATTGGGACTGTAAAGGGATAGG (356)
PAM

Minor Allele
TTAAAAATAAAAAATAAGTTAACTCTGATTAACCTGACATTTCCCTATCC (357)
AATTTTATTTTATTTCAATTGTTGAGCTAATTGGGACTGTAAAGGGATAGG (358)

LbCpf1(TTIN/TTDN/CTTN/TCIN/ATTN/TCEN/...):
Ref. Allele:
TTAAAAATAAAAAATAAGTTAACTCTGATTAACCTGACATTTCCCTATCC (355)
AATTTTATTTTATTTCAATTGTTGAGCTAATTGGGACTGTAAAGGGATAGG (356)

Minor Allele
PAM
TTAAAAATAAAAAATAAGTTAACTCTGATTAACCTGACATTTCCCTATCC (357)
AATTTTATTTTATTTCAATTGTTGAGCTAATTGGGACTGTAAAGGGATAGG (358)
(218) AATTAACCTGACATTTCCCTA

=====

# Variant ID      Sequence Variation      Ref. Allele      Min. allele
rs4498089  AGTGTAACTTATTTTATTTTAA[A/G]AAAATTTTAAAGGGCTTTTCAGCA  A  G
                                                    (115)

SpCas9_WT (PAM Motif: NRG)
Ref. Allele:

```

Figure 5M

AGTGTAACTTATTTTATTTTAAAAAATTGTTAAGGGCTTTCCAGCAA (359)
TCACAATTGAATAAAAAATAAAAAATTTTAAACAATTCCCGAAAGGTGGT (360)

Minor Allele PAM
AGTGTAACTTATTTTATTTTAAAGAAATTTGTTAAGGGCTTTCCAGCAA (361)
TCACAATTGAATAAAAAATAAAAAATTTTAAACAATTCCCGAAAGGTGGT (362)
GTTAACTTATTTTATTTT (219)

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)
Ref. Allele:
AGTGTAACTTATTTTATTTTAAAAAATTGTTAAGGGCTTTCCAGCAA (359)
TCACAATTGAATAAAAAATAAAAAATTTTAAACAATTCCCGAAAGGTGGT (360)

Minor Allele PAM
AGTGTAACTTATTTTATTTTAAAGAAATTTGTTAAGGGCTTTCCAGCAA (361)
TCACAATTGAATAAAAAATAAAAAATTTTAAACAATTCCCGAAAGGTGGT (362)
GTTAACTTATTTTATTTT (220)

SaCas9 (NNGRRT/NNGRR):
Ref. Allele:
AGTGTAACTTATTTTATTTTAAAAAATTGTTAAGGGCTTTCCAGCAA (359)
TCACAATTGAATAAAAAATAAAAAATTTTAAACAATTCCCGAAAGGTGGT (360)

Minor Allele PAM
AGTGTAACTTATTTTATTTTAAAGAAATTTGTTAAGGGCTTTCCAGCAA (361)
TCACAATTGAATAAAAAATAAAAAATTTTAAACAATTCCCGAAAGGTGGT (362)
GTTAACTTATTTTATTTT (219)

AaCpf1 (TTN):
Ref. Allele:
CACAAATTGAATAAAAAATAAA (221)
AGTGTAACTTATTTTATTTTAAAAAATTGTTAAGGGCTTTCCAGCAA (359)
TCACAATTGAATAAAAAATAAAAAATTTTAAACAATTCCCGAAAGGTGGT (360)
PAM

Minor Allele
AGTGTAACTTATTTTATTTTAAAGAAATTTGTTAAGGGCTTTCCAGCAA (361)
TCACAATTGAATAAAAAATAAAAAATTTTAAACAATTCCCGAAAGGTGGT (362)

§ Variant ID Sequence Variation Ref. Allele Min. allele
rs57666989 CTGCAAGCTCCGCTCCCGAGTTCA[C/T]GCCATTCTCCTGCTCAGTCTCCCA C T
(117)

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)
Ref. Allele: TAAGAGGACGGAGTCAGAGG (222)
CTGCAAGCTCCGCTCCCGAGTTCAAGCCATTCTCCTGCTCAGTCTCCCA (363)
GACGTTCCAGGCGAAGCGCTCAAGTCCCGTAAGAGGACGGAGTCAGAGG (364)
PAM

Minor Allele
CTGCAAGCTCCGCTCCCGAGTTCAAGCCATTCTCCTGCTCAGTCTCCCA (365)
GACGTTCCAGGCGAAGCGCTCAAGTCCCGTAAGAGGACGGAGTCAGAGG (366)

SpCas9_VRRR (PAM Motif: NGCG)
Ref. Allele: TAAGACACCCACTCAGAG (223)
CTGCAAGCTCCGCTCCCGAGTTCAAGCCATTCTCCTGCTCAGTCTCCCA (363)
GACGTTCCAGGCGAAGCGCTCAAGTCCCGTAAGAGGACGGAGTCAGAGG (364)
PAM

Minor Allele
CTGCAAGCTCCGCTCCCGAGTTCAAGCCATTCTCCTGCTCAGTCTCCCA (365)
GACGTTCCAGGCGAAGCGCTCAAGTCCCGTAAGAGGACGGAGTCAGAGG (366)

§ Variant ID Sequence Variation Ref. Allele Min. allele
rs10006128 tcttgatctcctgacctcgtcatcc[C/G]cagacattgtgatacggccacctcg G C
(118)

SpCas9_WT (PAM Motif: NRG)
Ref. Allele:
TCTTGATCTCCTGACCTCCTCATCCCGGACCTTGTGATCCGCCCACCTCG (367)
AGAACTAGAGGACTCGAGCACTAGGCGGCTCGAACAATAGCGCGGTGGAGC (368)

Minor Allele CTGGAACACTAGGCGGTTGG (224)
TCTTGATCTCCTGACCTCCTCATCCCGGACCTTGTGATCCGCCCACCTCG (369)
AGAACTAGAGGACTCGAGCACTAGGCGGCTCGAACAATAGGCGGTTGGAGC (370)

Figure 5N

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)
 Ref. Allele:
 TCTTGATCTCCTGACCTCGTCATCCCGGACCTTGTGATCCGCCCACTCG (367)
 AGAAGTAGAGGACTGGAGCAGTAGGGGGCTGGAACTAGGGGGGTGGAGC (368)

Minor Allele (224) CTGGAACACTAGGGGGGTGG
 TCTTGATCTCCTGACCTCGTCATCCCGGACCTTGTGATCCGCCCACTCG (369)
 AGAAGTAGAGGACTGGAGCAGTAGGGGGCTGGAACTAGGGGGGTGGAGC (370)
 PAM

SpCas9_VRRR (PAM Motif: NGCG)
 Ref. Allele:
 TCTTGATCTCCTGACCTCGTCATCCCGGACCTTGTGATCCGCCCACTCG (367)
 AGAAGTAGAGGACTGGAGCAGTAGGGGGCTGGAACTAGGGGGGTGGAGC (368)
 PAM

Minor Allele
 TCTTGATCTCCTGACCTCGTCATCCCGGACCTTGTGATCCGCCCACTCG (369)
 AGAAGTAGAGGACTGGAGCAGTAGGGGGCTGGAACTAGGGGGGTGGAGC (370)

SpCas9 (NNGRRR/NNGRR):
 Ref. Allele:
 TCTTGATCTCCTGACCTCGTCATCCCGGACCTTGTGATCCGCCCACTCG (367)
 AGAAGTAGAGGACTGGAGCAGTAGGGGGCTGGAACTAGGGGGGTGGAGC (368)

Minor Allele (224) CTGGAACACTAGGGGGGTGG
 TCTTGATCTCCTGACCTCGTCATCCCGGACCTTGTGATCCGCCCACTCG (369)
 AGAAGTAGAGGACTGGAGCAGTAGGGGGCTGGAACTAGGGGGGTGGAGC (370)
 PAM

```

=====
# Variant ID      Sequence Variation                                     Ref. Allele      Min. allele
rs28696683      CCTAATTTTGTATTTTACTAGAG[A/G]TCCCGCTTTCCCATCATCACCAGCC      A                C
(119)
  
```

SpCas9 (PAM Motif: NRG)
 Ref. Allele:
 CCTAATTTTGTATTTTACTAGAGATGGGGTTTCCCATCATCAGCAGGC (371)
 CCTAATTTTGTATTTTACTAGAGATGGGGTTTCCCATCATCAGCAGGC (372)

Minor Allele PAM
 CCTAATTTTGTATTTTACTAGAGATGGGGTTTCCCATCATCAGCAGGC (373)
 CCTAATTTTGTATTTTACTAGAGATGGGGTTTCCCATCATCAGCAGGC (374)
 ATTTTGTATTTTACTAG (225)

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)
 Ref. Allele: PAM
 CCTAATTTTGTATTTTACTAGAGATGGGGTTTCCCATCATCAGCAGGC (371)
 CCTAATTTTGTATTTTACTAGAGATGGGGTTTCCCATCATCAGCAGGC (372)
 TTTTGTATTTTACTAG (226)

Minor Allele
 CCTAATTTTGTATTTTACTAGAGATGGGGTTTCCCATCATCAGCAGGC (373)
 CCTAATTTTGTATTTTACTAGAGATGGGGTTTCCCATCATCAGCAGGC (374)

```

=====
# Variant ID      Sequence Variation                                     Ref. Allele      Min. allele
rs28393289      GGATTTTGAATGCGGAACCAACTGCGCTTGTGAACTCTGCTAAGTATAAC      A                G
(120)
  
```

SpCas9_VQR/EQR (PAM Motif: NGAN-NGNG/NGAG)
 Ref. Allele:
 GGATTTTGAATGCGGAACCAACTGCGCTTGTGAACTCTGCTAAGTATAAC (375)
 CCTAAACTTACGCGCTTGTGAACTCTGCTAAGTATAAC (376)

Minor Allele: PAM
 GGATTTTGAATGCGGAACCAACTGCGCTTGTGAACTCTGCTAAGTATAAC (377)
 CCTAAACTTACGCGCTTGTGAACTCTGCTAAGTATAAC (378)
 ATTTTGAATGCGGAACCAAC (227)

SpCas9_VRRR (PAM Motif: NGCG)
 Ref. Allele:
 GGATTTTGAATGCGGAACCAACTGCGCTTGTGAACTCTGCTAAGTATAAC (375)
 CCTAAACTTACGCGCTTGTGAACTCTGCTAAGTATAAC (376)

Minor Allele: PAM
 GGATTTTGAATGCGGAACCAACTGCGCTTGTGAACTCTGCTAAGTATAAC (377)
 CCTAAACTTACGCGCTTGTGAACTCTGCTAAGTATAAC (378)
 ATTTTGAATGCGGAACCAAC (227)

Minor Allele: ACAACTTGAGACGATTCT (228)
 GGATTTTGAATGCGGAACCAACTGCGCTTGTGAACTCTGCTAAGTATAAC (377)
 CCTAAACTTACGCGCTTGTGAACTCTGCTAAGTATAAC (378)
 PAM

Figure 50

Sequence sgHD1

Complementary guide: 5'GCTCCAGGCGTCGGCGG 3' (sgHD1) (n=17 nt)(379)
sgRNA expression cassette: hU6 promoter

Major Allele (C): Have a PAM motif on the Positive strand (NGG)
GTCGCCCCGCTCCAGGCGTCGGCGGGGATCCTTTCCGCATGGGCCTGCGCC (389)
CAGCGGGGCGAGGTCCGCAGCCGCCCCCTAGGAAAGGCGTACCCGGACGCGG (390)
GCTCCAGGCGTCGGCGG (379)

Minor Allele (G): Disruption of the PAM motive.
GTCGCCCCGCTCCAGGCGTCGGCGGGGATCCTTTCCGCATGGGCCTGCGCC (391)
CAGCGGGGCGAGGTCCGCAGCCGCCCCCTAGGAAAGGCGTACCCGGACGCGG (392)

Sequence sgHD2

Complementary guide: 5'GGCGCGGGCTCAACGGAG 3' (sgHD2) (n=19 nt)(380)
sgRNA expression cassette: hU6 promoter

Major Allele (C): Has a PAM motif on the Negative strand (NGG)
GAGGCAACTCGGGGCGCGG (380)
TTACAGTCTCACCACGCCCCGTCCCTCTCCGTTGAGCCCCGCGCCTTC (393)
AATGTCAGAGTGCTCGGGGCGAGGGAGAGGCAACTCGGGGCGCGGAAG (394)

Minor Allele (G): Disruption of the PAM motive
TTACAGTCTCACCACGCCCCGTCCCTCTCCGTTGAGCCCCGCGCCTTC (395)
AATGTCAGAGTGCTCGGGGCGAGGGAGAGGCAACTCGGGGCGCGGAAG (396)

Sequence sgHD3

Complementary guide: 5'GTCTGGGACGCAAGGCGCCG3' (sgHD3) (n=20 nt)(381)
sgRNA expression cassette: hU6 promoter

Major Allele (G): Has a PAM motif on the Positive strand (NGG) Use U6 promoter
GCGTCTGGGACGCAAGGCGCCGTGAGGGCTGCCGGGACGGGTCCAAGAT (397)
CGCAGACCCCTGCGTTCCGCGGCACCCCCAGGGCCCTGCCAGGTTCTA (398)
GTCTGGGACGCAAGGCGCCG (381)

Minor Allele (A): Loss PAM motif on the Positive strand (NGG) Use U6 promoter
GCGTCTGGGACGCAAGGCGCCGTGAGGGCTGCCGGGACGGGTCCAAGAT (399)
CGCAGACCCCTGCGTTCCGCGGCACCCCCAGGGCCCTGCCAGGTTCTA (400)

Sequence sgHD4

Complementary guide: 5' GATGCACCGGGCTGGGGC 3' (n=19 nt)(382)
sgRNA expression cassette: hU6 promoter

Major Allele: Has a PAM motif on the negative strand (NGG)
3' CGGGGTGGGGCGCACGTAG 5' (382)
5' TCTGCGTCAGGGTTTCCTTCTTTTCAGCCCCACCCCGCGTGCATCCCA 3' (401)
3' AGACGCAGTCCCAAAGGAAGAAAAGTTCGGGGTGGGGCGCACGTAGGGT 5' (402)

Minor Allele: SNP produces a Loss of the PAM motif.
5' TCTGCGTCAGGGTTTCCTTCTTTTCAGCCCCACCCCGCGTGCATCCCA 3' (403)
3' AGACGCAGTCCCAAAGGAAGAAAAGTTCGGGGTGGGGCGCACGTAGGGT (404)

Figure 6A

Sequence sgHD5g and sgHD5c

Complementary guide: 5'ATTCAGGTTGATGTCCT 3' (sgHD5g) (n=17 nt)(383)
 Complementary guide: 5'ATCCCATTTCTGAGGTTCTGG 3' (sgHD5c) (n=20 nt)(384)
 sgRNA expression cassette: hH1 promoter

Major Allele: Has a PAM motif on the Positive strand (NAG)

5'CCCTCATTTCAGGTTGATGTCCTAAGCCCCAGAACCTCAGAATGGGATTG 3' (405)
 3' GGGAGTAAGTCCAACTACAGGATTGCGGGTCTTGGAGTCTTACCCTAAC 5' (406)
 ATTCAGGTTGATGTCCT (383)

Minor Allele: generates a PAM motif on the Negative strand (NAG)

GGTCTTGGAGTCTTACCCTA (384)
 5'CCCTCATTTCAGGTTGATGTCCTAAGCCCCAGAACCTCAGAATGGGATTG 3' (407)
 3'GGGAGTAAGTCCAACTACAGGATTGCGGGTCTTGGAGTCTTACCCTAAC 5' (408)

Sequence sgHD6c and sgHD6g

Complementary guide: 5'GCAGGCAGAGAGGAGCC 3' (sgHD6c) (n=17 nt)(385)
 Complementary guide: 5'GCCTGGCTAAAGTAGGCTT 3' (sgHD6g) (n=19 nt)(386)
 sgRNA expression cassette: hU6 promoter

Major Allele (C): Have a PAM motif on the negative strand (NAG)

CCGAGGAGAGACGGACG (385)
 5' GTGGCCTGGCTAAAGTAGGCTTTACTGGGCTCCTCTCTGCCTGCATCAC 3' (409)
 3' CACCGGACCGATTTCATCCGAAATGACCCGAGGAGAGACGGACGTAGTG 5' (410)

Minor Allele (G): Generates a PAM motif on the Positive strand (NAG)

5' GTGGCCTGGCTAAAGTAGGCTTTACTGGGCTCCTCTCTGCCTGCATCAC 3' (411)
 3' CACCGGACCGATTTCATCCGAAATCACCCGAGGAGAGACGGACGTAGTG 5' (412)
 GCCTGGCTAAAGTAGGCTT (386)

Sequence sgHDi3

Complementary guide: 5'GCTTTTAGGACGCCTCGG 3' (sgHDi3) (n=18 nt)
 sgRNA expression cassette: hU6 promoter (387)

5' AATGCTTTTAGGACGCCTCGGCGGGAGTGGCGGGGCAGGGGGGGGGCG 3' (413)
 3' TTACGAAAATCCTGCGGAGCCGCCCTCACCGCCCCGTCCCCCCCCCGC 5' (414)
 GCTTTTAGGACGCCTCGG (387)

Sequence sgHDi4

Complementary guide: 5' GCGGGACACTTCGAGAGG 3' (sgHDi4) (n=18 nt)
 sgRNA expression cassette: hU6 promoter (388)

5' GCGCGGGACACTTCGAGAGGAGGCGGGGTTTGGAGCTGGAGAGATGT 3' (415)
 3' CCGCGCCCTGTGAAGCTCTCCTCCGCCCAACCTCGACCTCTCTACA 5' (416)
 GCGGGACACTTCGAGAGG (388)

Figure 6B

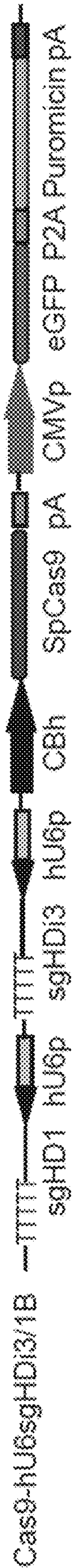


Figure 7A

ND31551		ND33392		
HD Allele	Normal	Mutant	Normal	Mutant
CAG repeat	18	39	17	57
rs2857935	C	G	G	C
PAM motif	Present	Absent	Absent	Present

Figure 7B

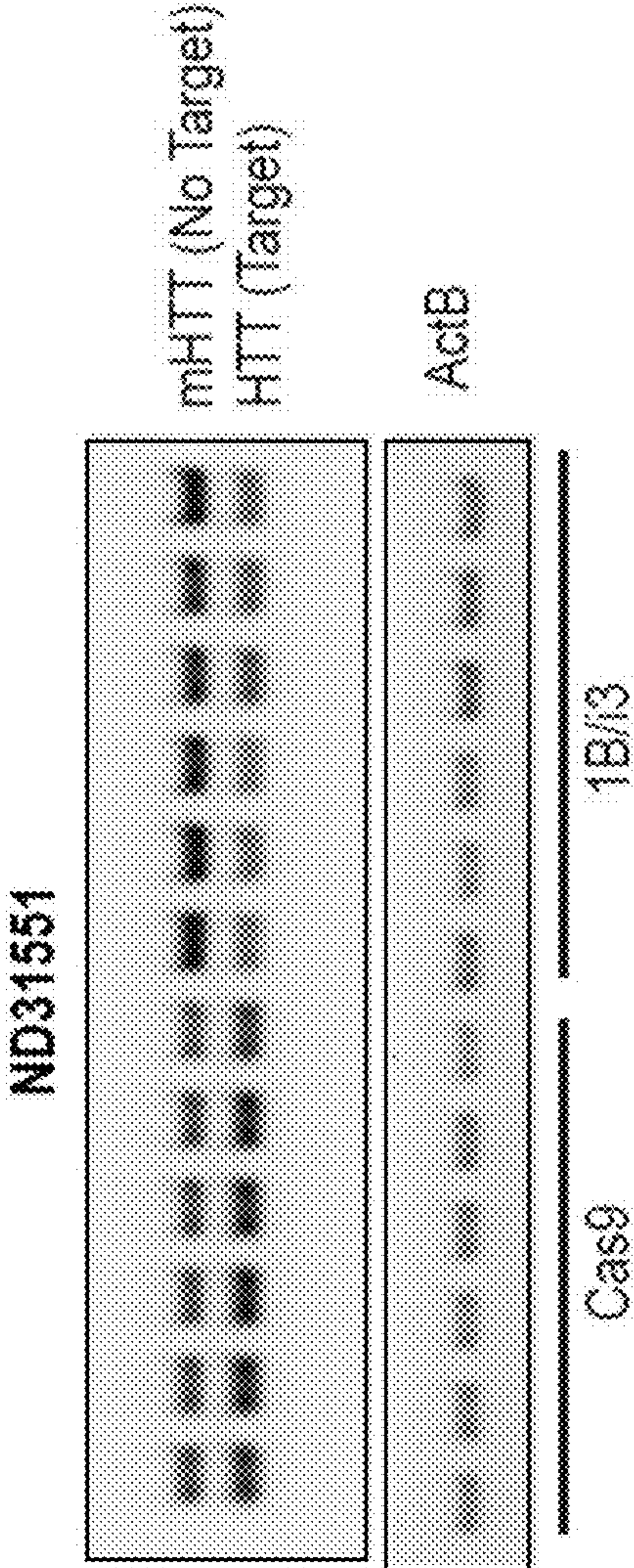


Figure 7D

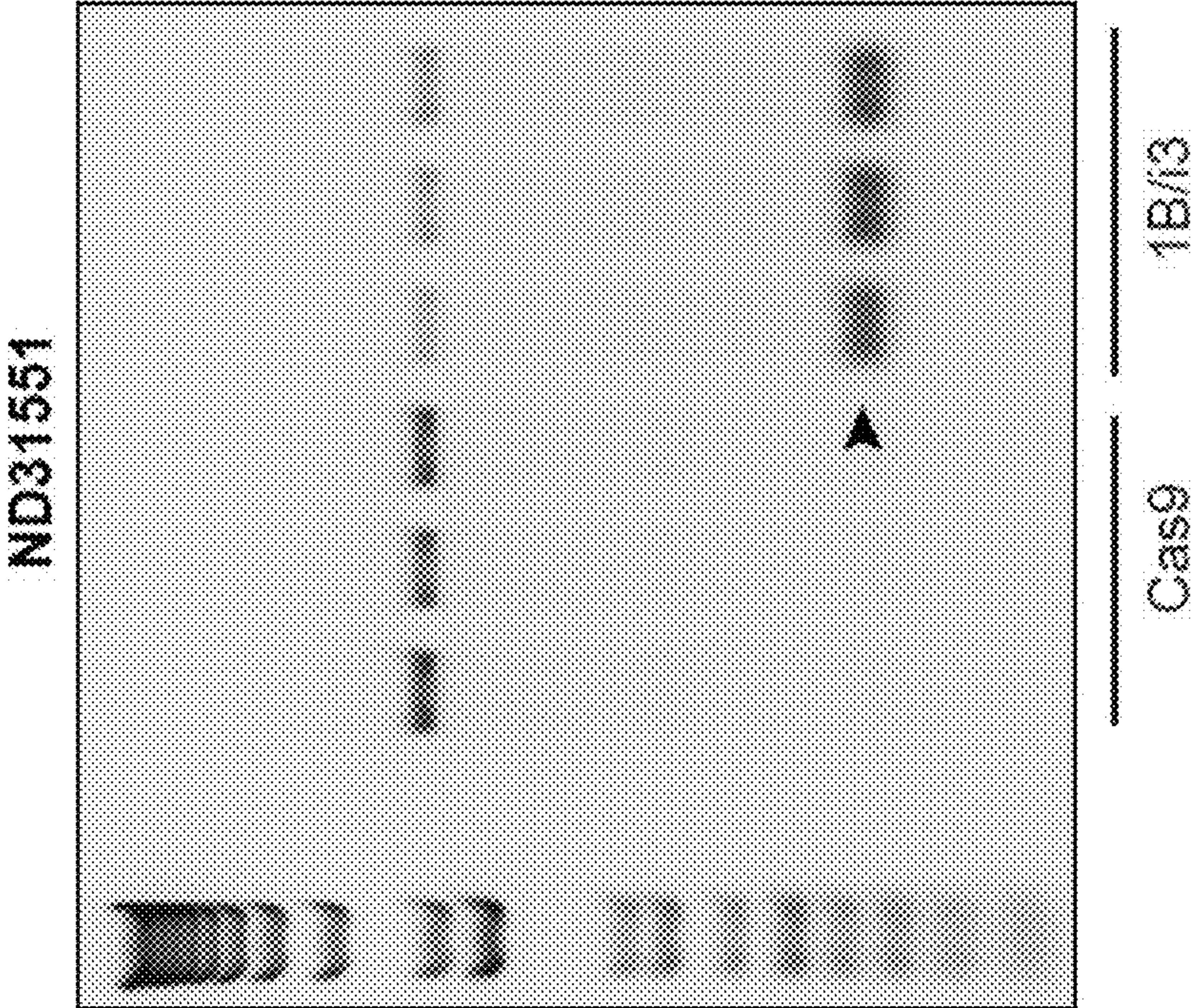


Figure 7C

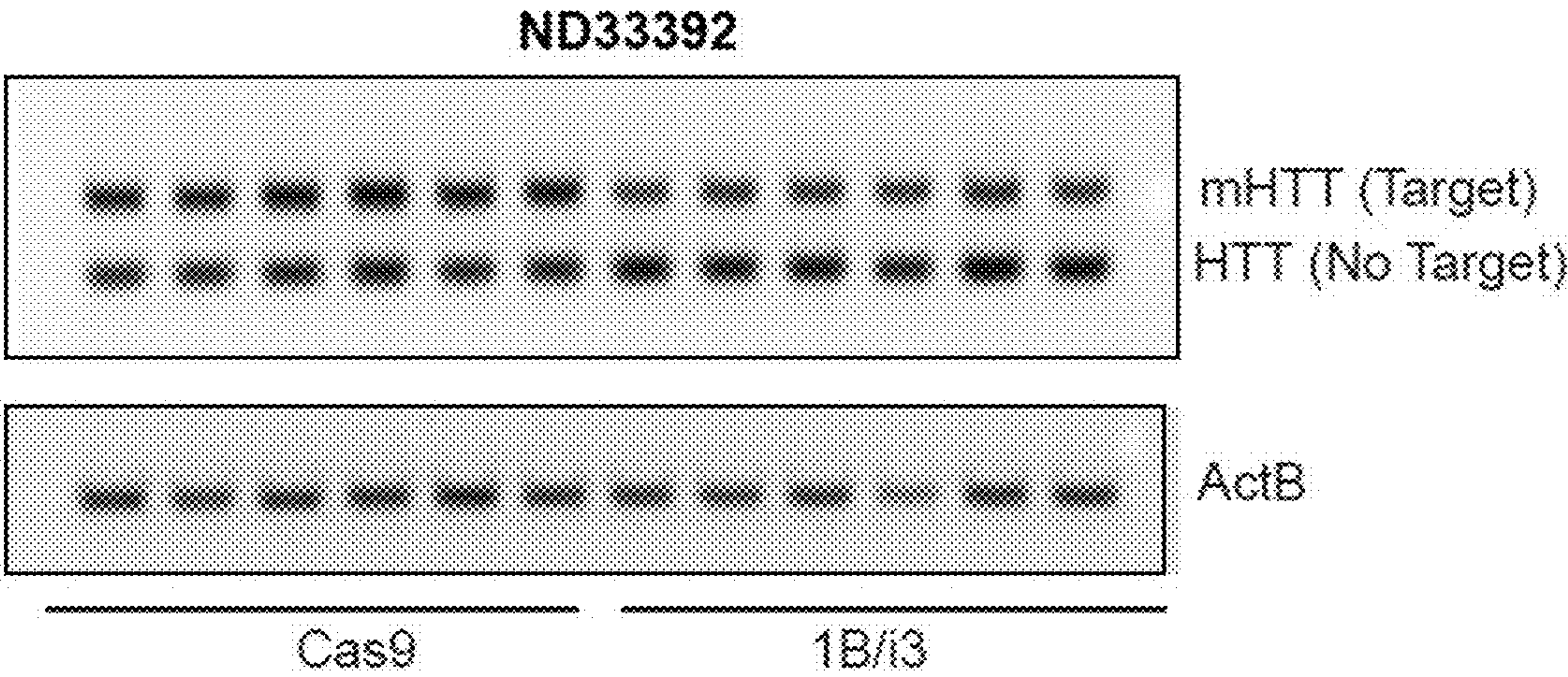


Figure 7E

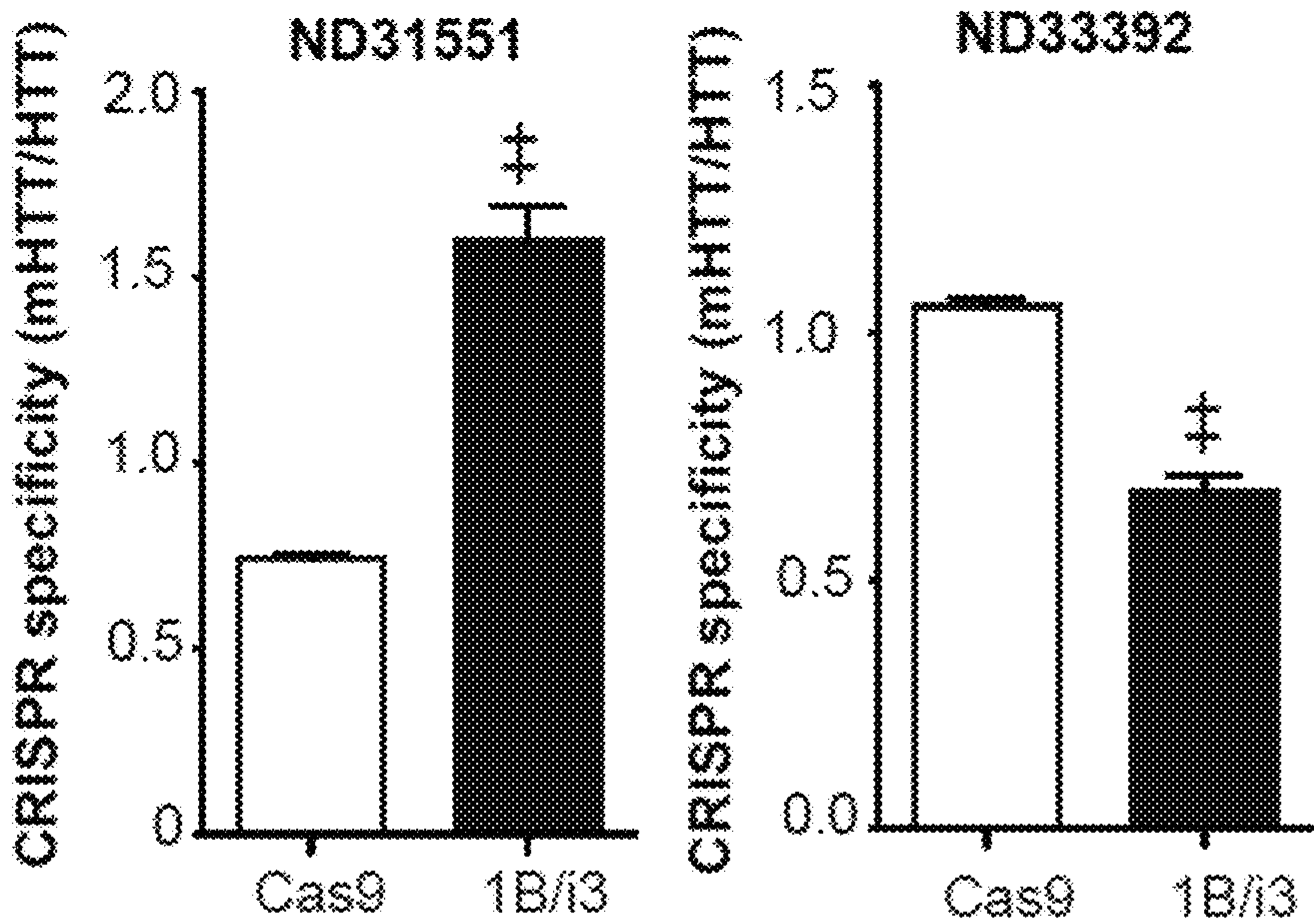


Figure 7F

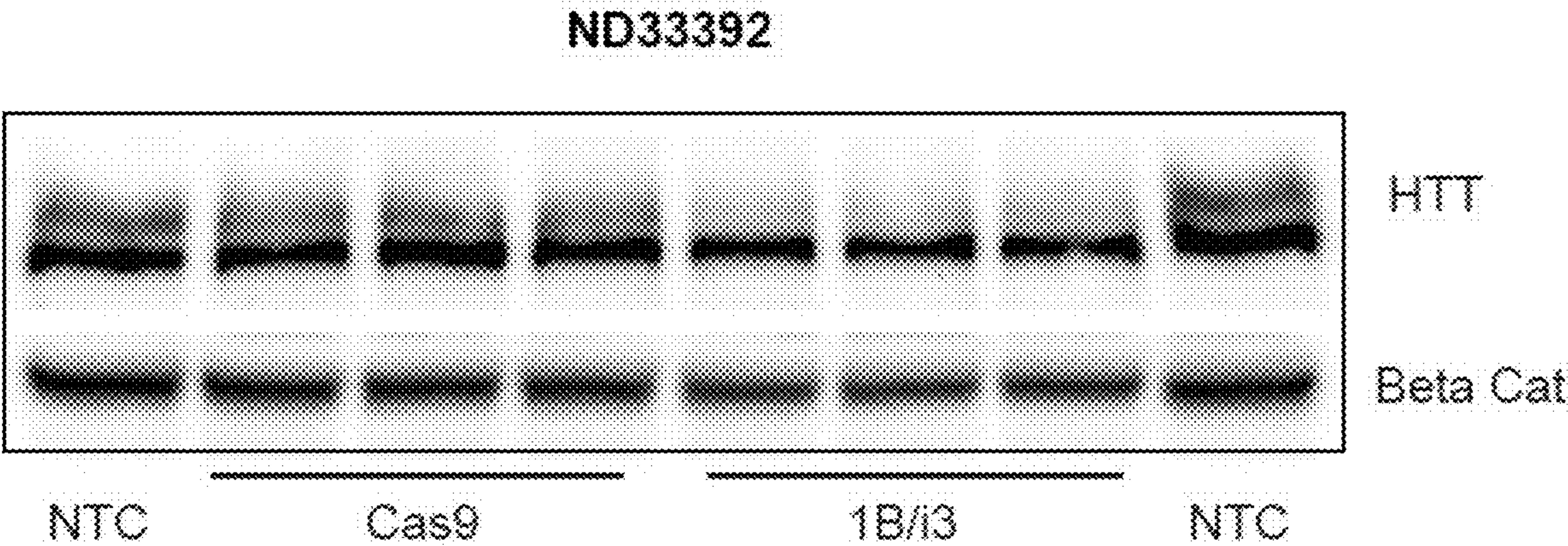


Figure 7G

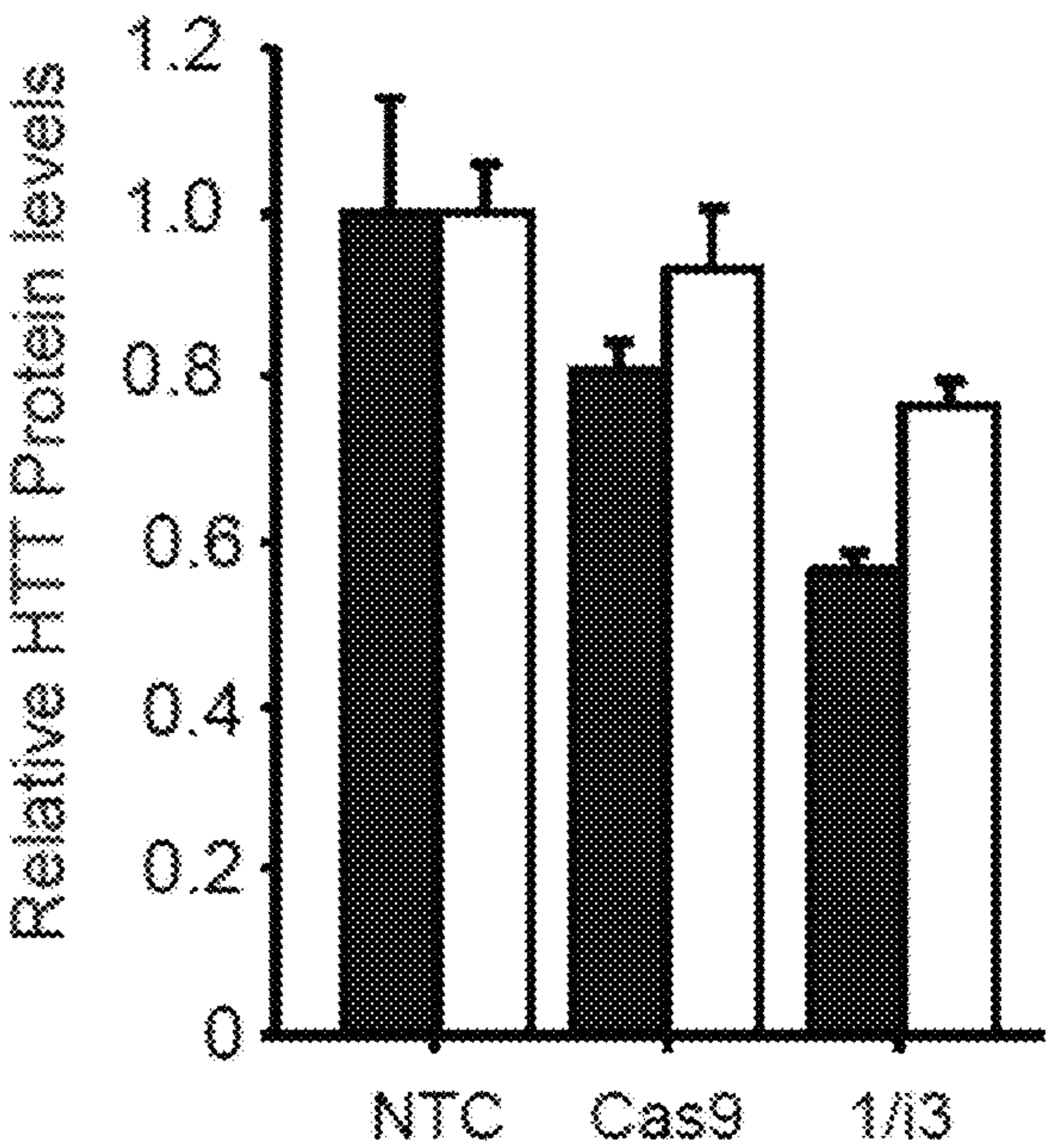


Figure 7H

	sgHD1	sgHD2	sgHD3	sgHDi3
Comp. (nt)	17	19	20	18
0 mm	1	1	1	1
1 mm	2	0	0	0
2 mm	32	1	0	2
3 mm	382	38	8	18
Total	416	40	7	21

Figure 8A

	sgHD1	sgHD2	sgHD3	sgHDi3
Intergenic	71	11	2	6
Promoter	58	4	0	1
5'UTR	64	3	0	0
Intron	143	20	5	13
Exon	75	1	0	1
3'UTR	5	1	0	0
Total	416	40	7	21

Figure 8B

5' GTTTAATTGAGTTGTCATATGTTAATAACGGTAT 3'
 5' ATACCGTTATTAACATATGACAACTCAATTAAAC 3' (Rev Complement)

TGTCGCCCCGCTCCAGGCGTCGGTC-----GGCGGGAGTGGCGGGGCAGGGGGGG
 TGTCGCCCCGCTCCAGGCGTCGG--ATACCGTTATTAACATATGACAACTCAATTAAAC--CGGCGGGAGTGGCGGGGCAGGGGGGG
 TGTCGCCCCGCTCCAGGCGTCG-----GTTATTAACATATGACAACTCAATTAAAC-----CGGCGGGAGTGGCGGGGCAGGGGGGG
 TGTCGCCCCGCTCCAGGCGTCGG-GTTTAATTGAGTTGTCATATGTTAATAACGGTATCTCGGCGGGAGTGGCGGGGCAGGGGGGG
 TCGCCCCGCTCCAGGCGTCGG-----TATTAACATATGACAACTCAATTAAAC-----CGGCGGGAGTGGCGGGGCAGGGGGGG
 TGTCGCCCCGCTCCAGGCGTCGG-----ATTAACATATGACAACTCAATTAAAC-----CGGCGAGTGGCGGGGCAGGGGGGG

HTT Promoter ODN HTT Intron I

Figure 8C

Gene	Chromosome	Strand	Mismatches	Mismatch position	Location	Guide	Indels
SLC45A4	chr8	-	2	GCTCCAtGCGgCGGCGG	Exon	sgHD1	Undetected
CHRNA2	chr8	+	3	GCTtCAGGCGTaGGgGG	Exon	sgHD1	Undetected
TRIB1	chr8	+	3	GCTCCgGGgGcCGGCGG	Exon	sgHD1	Undetected
SMARCD1	chr12	+	3	GCTCCAaGCGgCGGCGc	Exon	sgHD1	Undetected
COX11	chr17	+	3	GCTcTTAGGcCGCCgCGG	Exon/5'UTR	sgHDi3	Undetected
STT3B	chr3	+	2	GCaCCAGGCGgCGGCGG	5'UTR	sgHD1	Undetected
NUP210	chr3	-	3	GCTTTaAGGAtGCCTtGG	Intron	sgHDi3	Undetected
DNAJC16	chr1	+	3	aCTTTTAGGAgGCCTaGG	Intron	sgHDi3	Undetected
LEPR	chr1	+	3	GCaTTTAGGAgGCCTaGG	Intron	sgHDi3	Undetected
CBFA2T3	chr16	+	1	GCTCCAGGCGTgGGCGG	Intron	sgHD1	Undetected
NAV2	chr11	+	2	GCTTTTgGGACGCCTCaG	Intron	sgHDi3	Undetected

Figure 8D

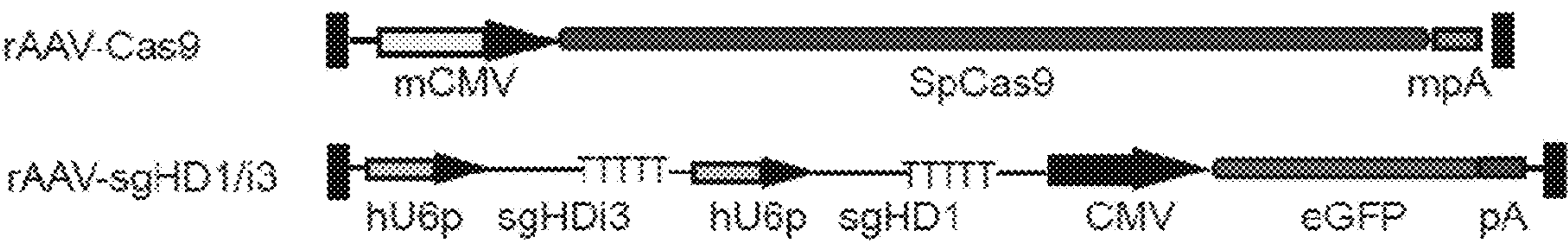


Figure 9A

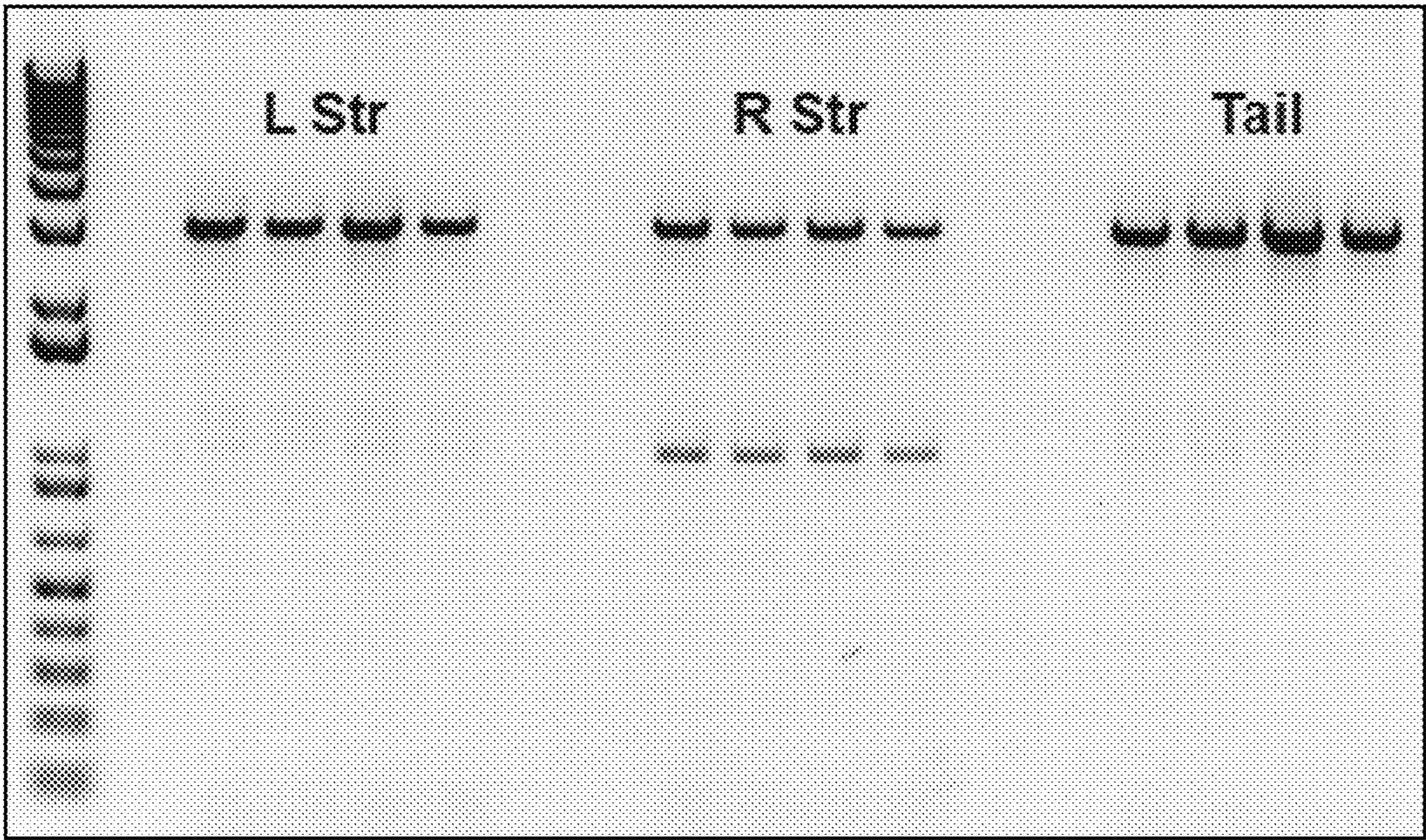


Figure 9B

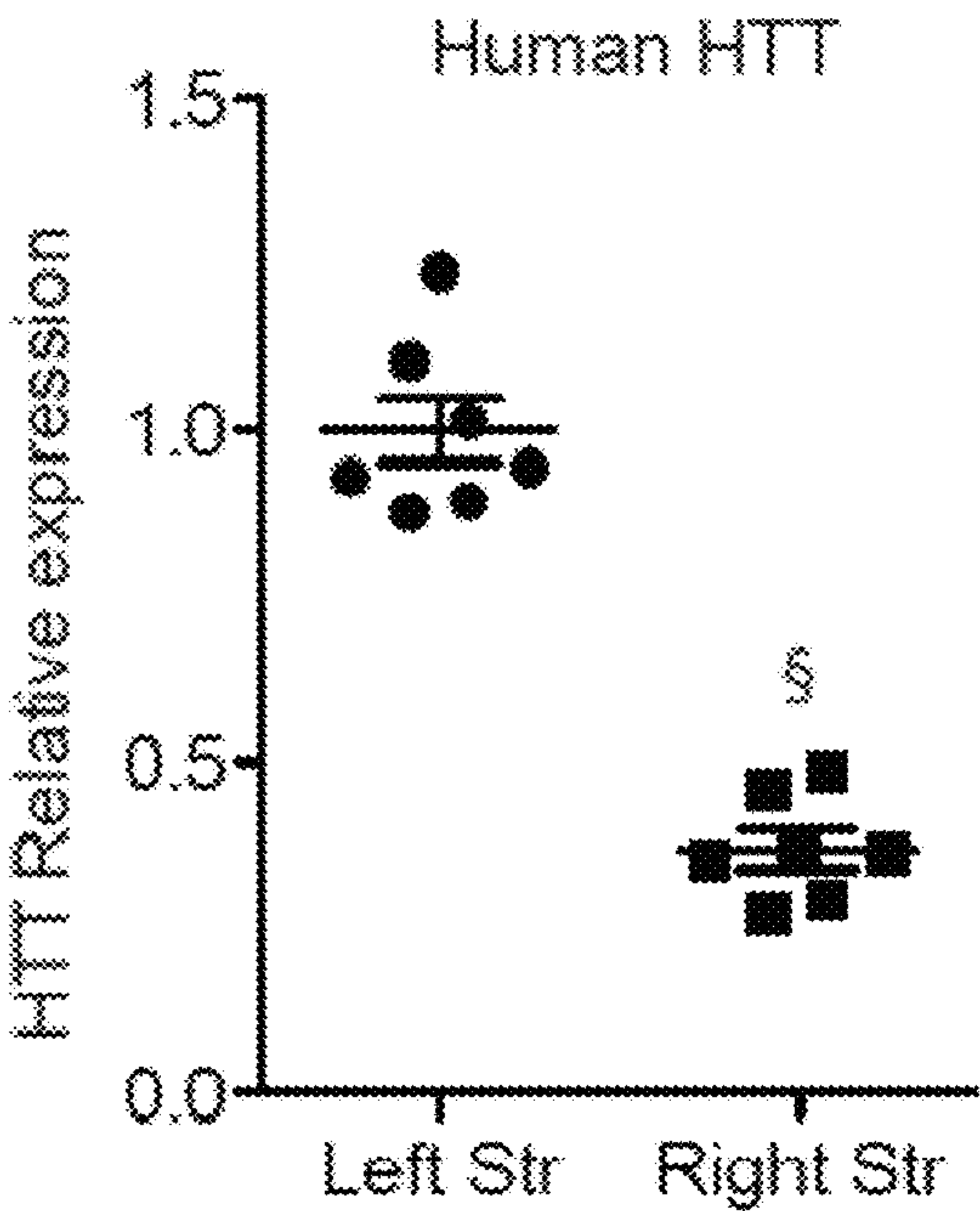


Figure 9C

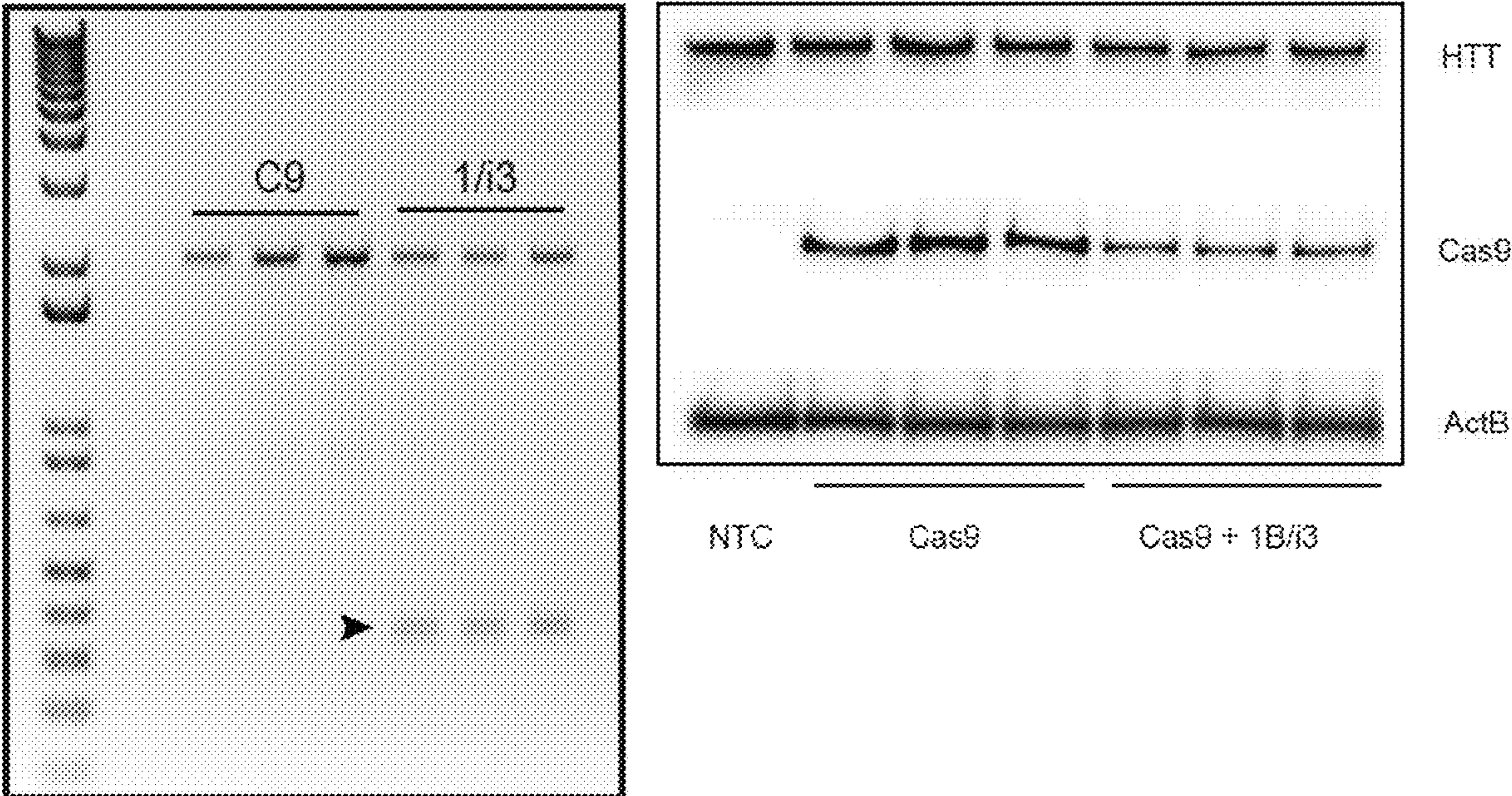


Figure 9D

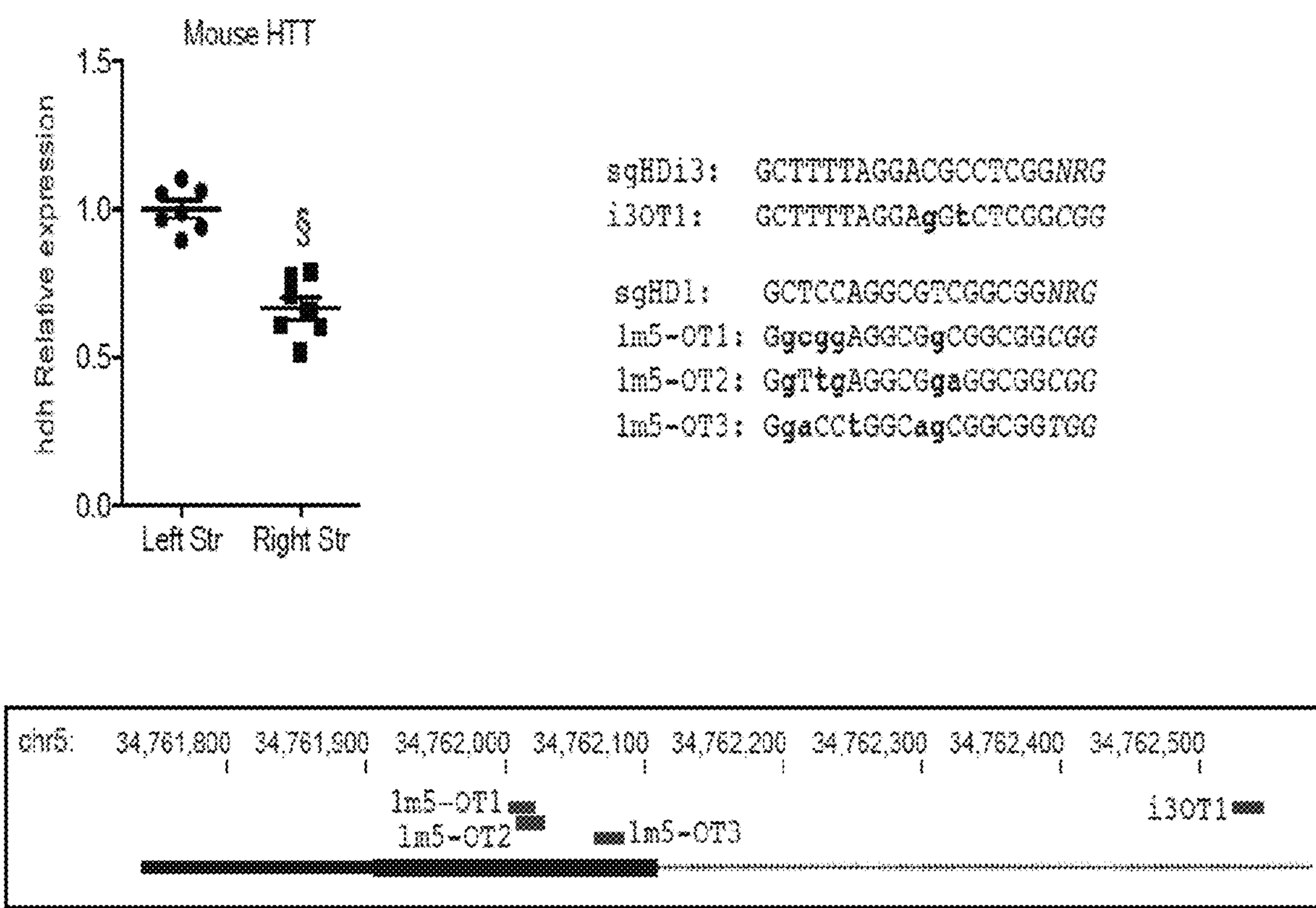


Figure 9E

COMPOSITIONS AND METHODS FOR TREATING HUNTINGTON'S DISEASE AND RELATED DISORDERS

[0001] This application is a continuation of U.S. patent application Ser. No. 15/766,549, filed on Apr. 6, 2018, which is a § 371 application of PCT/US2016/056417, filed Oct. 11, 2016, which claims priority under 35 U.S.C. § 119(e) to U.S. Provisional Patent Application No. 62/352,788, filed Jun. 21, 2016 and U.S. Provisional Patent Application No. 62/239,714, filed Oct. 9, 2015. The foregoing applications are incorporated by reference herein.

[0002] Incorporated herein by reference in its entirety is the Sequence Listing being concurrently submitted via EFS-Web as a XML file named SeqList, created Mar. 6, 2023, and having a size of 587,341 bytes.

[0003] This invention was made with government support under grant number NS084475 awarded by the National Institutes of Health. The government has certain rights in the invention.

FIELD OF THE INVENTION

[0004] The present invention relates to polyglutamine (polyQ) disorders, particularly Huntington's disease. Specifically, the instant invention provides compositions and methods for the treatment and/or prevention of Huntington's disease and related disorders.

BACKGROUND OF THE INVENTION

[0005] Huntington disease (HD) is an autosomal dominant neurodegenerative disease that manifest in adults (adult-onset HD) or children (Juvenile HD). HD is part of the family of polyglutamine (polyQ) disorders comprising at least nine different neurodegenerative diseases that result from the expansion of a triplet CAG repeat in specific genes (Walker, F. O. (2007) *Lancet*, 369:218-228). In HD, the disease causing mutation is found in the first exon of the huntingtin gene, and although mutant huntingtin is ubiquitously expressed, the brain, and particularly the striatum and motor cortex are the earliest and most affected (Walker, F. O. (2007) *Lancet*, 369:218-228; The Huntington's Disease Collaborative Research Group (1993) *Cell*, 72, 971-983). Patients with HD develop progressive neurodegeneration leading to death, generally within 20 years of onset.

[0006] There is no cure for HD, and treatments are focused on managing its symptoms (Johnson et al. (2010) *Hum. Mol. Genet.*, 19:R98-R102). Earlier studies using genetically modified mouse models showed that HD-like disease phenotypes can be resolved if mutant huntingtin expression is eliminated, even at advanced disease stages (Yamamoto et al. (2000) *Cell*, 101:57-66; Diaz-Hernandez et al. (2005) *J. Neurosci.*, 25:9773-9781). RNA interference (RNAi), a method of reducing gene expression, has emerged as a leading therapeutic option. RNAi does not eliminate all mutant huntingtin, however, and therefore it remains expressed at low levels. For this reason, there is a need for effective therapeutics for HD.

SUMMARY OF THE INVENTION

[0007] In accordance with one aspect of the instant invention, methods and compositions for inhibiting, treating, and/or preventing a polyglutamine disorder (e.g., Huntington's disease) in a subject are provided. In accordance with

another aspect of the instant invention, methods for reducing the expression of a mutant protein (e.g., mutant huntingtin) encoded by an allele of a gene associated with a polyglutamine disorder (e.g., Huntington's disease) in a cell are provided. In a particular embodiment, the methods of the instant invention comprise administering to the subject or cell a nucleic acid molecule encoding Cas9 and at least one guide RNA (e.g., sgRNA). In a particular embodiment, the guide RNA are administered to the subject or cell as a nucleic acid molecule (e.g., an expression vector or viral vector) encoding the guide RNA. In a particular embodiment, the methods of the instant invention comprise administering two guide RNAs to the subject or cell. In a particular embodiment, one guide RNA targets a sequence 5' of exon 1 (e.g., within the 5' untranslated region, within the promoter, or within the first 2 kb 5' of the transcription start site) and one guide RNA targets a sequence within intron 1. In a particular embodiment, at least one guide RNA administered to the subject or cell targets a sequence adjacent to a PAM present on only one allele of the gene (e.g., the mutant allele). In a particular embodiment, at least one of the guide RNAs targets a sequence specifically set forth herein.

[0008] The instant invention also encompasses guide RNAs, nucleic acid molecules (e.g., an expression vector or viral vector) encoding the guide RNAs, and compositions comprising the guide RNAs and/or nucleic acid molecules (e.g., an expression vector or viral vector) encoding the guide RNAs. In a particular embodiment, the composition and nucleic acid molecules (e.g., an expression vector or viral vector) encoding the guide RNAs contain or encode more than one guide RNA.

BRIEF DESCRIPTIONS OF THE DRAWING

[0009] FIG. 1A provides a schematic depicting the genomic deletion mechanism by the sgRNA/Cas9 complex. FIG. 1B shows the targeting sequences of the HTT promoter (top) and PCR results demonstrating genomic deletion in transfected cells (bottom). FIG. 1C provides a schematic depicting genomic deletion mechanism by sgRNA/Cas9 complex (left) and strategy adopted to achieve allele specific gene editing based on location of single nucleotide polymorphisms at the third nucleotide sequence of the PAM motif (right).

[0010] FIG. 2A provides a schematic of the targeting of Cas9 to allele specific PAM motifs by sgRNA. The loss of a PAM (left), gain of a PAM (center), and a loss/gain (right) are depicted. FIG. 2B provides examples of allele specific PAM motifs based on *Streptococcus pyogenes* (SpCas9; PAM—NRG) and *Staphylococcus aureus* Cas9 (SaCas9; PAM—NNGR, NNRR). The nucleotide variation of a SNP within a PAM alters Cas9 recognition resulting in the loss (Left), the gain (Middle), or the simultaneous loss of a PAM in one DNA strand and the gain of a PAM on the opposite strand (Right). FIG. 2C shows the locations of examples (Table 1) of the SNP ID's and genomic localization of the nucleotide variants targeted by SpCas9/sgRNA complex in the promoter and first intron sequence of the human huntingtin genomic locus.

[0011] FIGS. 3A-3C: SNP dependent editing for Huntington disease therapy. FIG. 3A: Cartoon depicting the allele specific editing strategy to abrogate mutant HTT expression. SNPs within PAM sequences upstream of HTT exon1 permit specific targeted deletions of the mutant allele when present in heterozygosity. After DNA repair mutant HTT exon1 is

deleted by a pair of sgRNA/Cas9 complexes binding upstream and downstream of exon1 (Right), whereas intronic indels could be generated by a single dsDNA break in the normal allele (Left). FIG. 3B provides a schematic of a normal allele and a mutant allele in the presence of sgHD1 and sgHDi3. The presence of both sgRNAs causes the cleavage of only the mutant allele. FIG. 3C: 21 out of 47 prevalent SNPs flanking HTT exon1 are located within predicted critical positions of a PAM sequence for the CRISPR/SpCas9 system analyzed. The minor frequency allele either mediates the loss (8 SNPs), again (8 SNPs) or a loss/gain (5 SNPs) of a PAM motif. FIGS. 3D-3J: List of the 36 prevalent SNPs upstream/downstream of HTT exon1 that fall within the nucleotide positions of the different CRISPR/Cas9 and CRISPR/Cpf1 systems analyzed. Sequences in FIG. 3D are SEQ ID NOs: 74-120, from top to bottom. FIG. 3K shows the flanking sequence of the indicated SNPs, the location of the PAMs, and an example of complementary sequence to target the site for cleavage. Nucleotide variations at specific genomic DNA sequences results in the loss or the loss/gain of SNP-dependent PAM motifs which generate or eliminate the recruitment of SpCas9/sgRNA to the huntingtin genomic locus. SEQ ID NOs are provided in parentheses.

[0012] FIGS. 4A-4I: Cleavage of SNP-dependent sgHD/SpCas9 complexes in HEK293 cells. FIG. 4A: Cartoon depicting the relative position of the 6 prevalent SNP-dependent PAMs upstream of HTT exon1, and 2 common PAMs within HTT intron 1. The estimated size of the targeted deleted sequence is indicated. FIG. 4B: Genotype of the prevalent SNPs within the HTT promoter in HEK293 cells. All SNPs were homozygous for the nucleotide variation and the PAM motif was present for the sgRNA indicated. FIG. 4C: Diagram of the CRISPR expression systems transfected into HEK293 cells. FIGS. 4D-4F: Genomic PCR showing HTT exon1-targeted deletion by sgRNA/SpCas9 pair complexes binding upstream and downstream of the target sequence. FIG. 4G: RT-qPCR analysis of HTT mRNA levels in HEK293 cells transfected with sgHD/SpCas9 expression cassettes targeting upstream promoter SNPs and the common intronic sgHDi3 sequence. All samples are normalized to human GAPDH and results are the mean $\pm$ SEM relative to cells transfected with plasmids containing the SpCas9 only control. (n=64; $P<0.001$, $\#P<0.0001$, One way ANOVA followed by a Bonferroni's post-hoc). FIG. 4H: sgHD1/i3/SpCas9 and sgHD3/i3/SpCas9 and sgHDi3/SpCas9 expression cassettes were transfected into HEK293 cells, and endogenous HTT protein levels were determined after puromycin selection and expansion. Cells transfected with Cas9 only was used as a control and betacatenin served as a loading control. FIG. 4I: Quantification of HTT protein levels after treatment with sgHD/SpCas9 complexes. Data are the mean $\pm$ SEM relative to cells transfected with plasmids containing SpCas9 only control. (n=6; $\#P<0.0001$, $\S P<0.001$, one way ANOVA followed by Bonferroni's post-hoc). FIG. 4J: Sanger sequencing of PCR amplified products after of HTT exon1 editing with sgHD/Cas9 expression vectors targeting upstream promoter SNPs and the intronic i3 PAM. SEQ ID NOs are provided in parentheses. FIG. 4K: RT-qPCR analysis of HTT mRNA levels in HEK293 cells transfected with sgHD/SpCas9 expression cassettes targeting upstream promoter SNPs and common intronic sgHDi4 sequences. All samples are normalized to human GAPDH and results are mean $\pm$ SEM

relative to cells transfected with plasmids containing SpCas9 only control. (n=4). FIG. 4L provides a schematic depicting the strategy to delete 1st exon sequence of mutant huntingtin gene using the CRISPR/Cas9 system targeting the SNP-derived PAM motifs at positions SNP1, SNP2, SNP3 and SNP6 in combination with the intronic i3 PAM motif (top). A representative experiment is also provided (bottom) showing deletion of HTT genomic fragment after coexpression of SNP and i3 targeting CRISPR/Cas9 expression systems. Arrow indicates expected DNA bands derived from cleavage and NHEJ recombination. FIG. 4M provides a schematic depicting the strategy to delete 1st exon sequence of mutant huntingtin gene using the CRISPR/Cas9 system targeting the SNP6 in combination with i3 PAM motif (top). A representative experiment is also provided (bottom) showing deletion of HTT genomic fragment after coexpression of SNP6 and i3 targeting CRISPR/Cas9 expression systems. Arrow indicates expected DNA bands derived from cleavage and NHEJ recombination using 3 different primer sets. FIG. 4N provides a schematic depicting the strategy to delete 1st exon sequence of mutant huntingtin gene using the CRISPR/Cas9 system targeting the SNP4, SNP5 derived PAM motifs in combination with i3 PAM motif (top). Representative experiments are provided (bottom) showing deletion of HTT genomic fragment after coexpression of SNP targeting CRISPR/Cas9 systems. Arrow indicates expected DNA bands derived from cleavage and NHEJ recombination.

[0013] FIGS. 5A-5O: List of SNP-dependent PAMs targeted by different CRISPR systems. SEQ ID NOs are provided in parentheses.

[0014] FIGS. 6A-6B: List of sgRNA sequence designed for targeting prevalent SNP-dependent PAM motifs present upstream 5' of HTT exon1, and common sgHD sequences within HTT intron 1. SEQ ID NOs are provided in parentheses.

[0015] FIGS. 7A-7H: Assessment of allele specific cleavage in human HD fibroblasts. FIG. 7A: Cartoon depicting the CRISPR expression plasmid use to co-express sgHD1 and sgHDi3 expression cassettes. SpCas9 and the selective reporter eGFP/puromycin expression cassettes present in the same plasmid are also shown. FIG. 7B: ND31551 and ND33392 HD fibroblasts lines were selected to determine allele specific target deletions of HTT. CAG repeat length, nucleotide variation and the allele location of the PAM motif are indicated. FIG. 7C: Representative genomic PCR showing HTT exon1 deletion of DNA harvested from the electroporated ND31551 HD fibroblast cell line. Arrow indicates the expected PCR amplification product resulting from allele specific deletion. FIGS. 7D and 7E: Semi-quantitative PCR reaction showing reduction of the targeted allele containing the conserved PAM sequence. For ND31551 fibroblasts the PAM sequence is conserved in the normal allele, while for ND33392 fibroblasts the PAM sequence is in the mutant allele. Expression levels are reduced only on the PAM-containing allele. FIG. 7F: Quantification of mRNA reduction in treated HD fibroblasts. Data show the ratio between mRNA levels of the mutant with respect to the normal allele, relative to cells electroporated with vectors expressing only the Cas9 control. The results are mean $\pm$ SEM relative to cells transfected with plasmids containing SpCas9 only control. (n=6; $P<0.01$, Mann Whitney T-test). FIG. 7G: Representative western blot showing allele specific depletion (upper band of HTT doublet consisting of normal (lower) and expanded polyQ-containing proteins) after electroporation

of HD fibroblasts with sgHD/SpCas9 expression vectors. FIG. 7H: Western blot quantification of HD fibroblasts electroporated with Cas9-hU6sgHD1/i3 expression vector.

[0016] FIGS. 8A-8D: Assessing off target activity of sgHD1 and sgHDi3/Cas9. FIG. 8A: Table depicting the number of off-target sites for the most active sequences predicted to bind with 1, 2 or 3 mismatches. The nucleotide length of the complementary guide sequence is also indicated. FIG. 8B: Table highlighting the number of off-target sites binding at different genomic regions using the UCSC genome browser. FIG. 8C: HD fibroblasts were electroporated with plasmids expressing sgHD1/i3 and SpCas9 along with an ODN sequence (SEQ ID NO: 417 and SEQ ID NO: 418 (reverse complement)). Sanger sequencing results showed the incorporation of the ODN sequence at the DNA cleavage site. HTT promoter sequence and HTT intron sequence outside the ODN sequence are also depicted (SEQ ID NOs: 419-424, from top to bottom). FIG. 8D: Sanger sequencing results from 11 predicted off target sites. Gene name, chromosome position, DNA strand, number of mismatches and position within the guide, gene location, sgRNA sequence and indel presence or absence are indicated. Sequences are SEQ ID NOs: 425-435, from top to bottom.

[0017] FIGS. 9A-9C: In vivo gene editing of the mutant HTT allele. FIG. 9A: Cartoon depicting rAAV shuttle vectors containing the SpCas9 and sgHD1/i3 expression cassettes. mCMV, minimal CMV promoter; mpA, minimal polyA. hU6p, human U6 promoter; pA, SV40 polyA. FIG. 9B: PCR of isolated genomic DNA showing human HTT exon1 targeted deletion after injection of vectors expressing SpCas9 and sgRNA sequences. LStr, left Striatum; RStr, Right Striatum. FIG. 9C: RT-QPCR analysis of HTT mRNA levels in striatum samples harvested 3 weeks after SpCas9 and sgHD1/i3 delivery. All samples were normalized to beta actin and results are mean $\pm$ SEM relative to uninjected striatal samples (n=7 animals per group, § P<0.001, Mann Whitney test). FIG. 9D: Genomic PCR products of DNA harvested after HTT exon1 editing in HEK 293 cells with rAAV2/1 SpCas9 and rAAV2/1 vectors (Left) and representative western blot showing reduced human HTT protein levels in HEK293 cells after editing with rAAV2/1 SpCas9 and rAAV2/1 vectors (Right). FIG. 9E: RT-QPCR analysis of mouse Htt mRNA levels in striatum samples harvested 3 weeks after rAAV. SpCas9 and rAAV.sgHDi3/1 injection. All samples were normalized to beta actin and results are mean $\pm$ SEM relative to uninjected striatal samples (n=7 animals per group, § P<0.001, Mann Whitney test). Pairing between sgHD1 and sgHDi3 to mouse HTT exon1 sequences is also shown. One off-target site was identified for sgHDi3 in mouse Htt that contained 2 mismatches. Three off-targets were identified in the case of sgHD1, all with 5 mismatches between the mouse genomic sequence and the small guide RNA sequence. Sequences are SEQ ID NOs: 436-441, from top to bottom.

DETAILED DESCRIPTION OF THE INVENTION

[0018] Herein, therapeutics tools to inhibit mutant protein expression using genome-editing strategies based on the CRISPR/Cas9 technology are provided. Targeted gene deletions can be introduced when guide RNAs (e.g., two guide RNAs) complex with Cas9 and mediate dsDNA breaks followed by DNA repair. Given the potency of the CRISPR/

Cas9 technology for targeting both alleles, and the fact that huntingtin is an important protein for cell viability, a concern of this method is that one will eliminate all huntingtin in the cell—good and bad. Thus, approaches that selectively target the expression of the mutant huntingtin allele are desirable. Here, a gene editing approach for specific targeting of the mutant huntingtin allele is provided. The strategy takes advantage of single nucleotide polymorphisms prevalent in the population for which a targeting sequence (PAM motif) is generated depending on the nucleotide variation. gRNA sequences have been developed that recruit CRISPR/Cas9 complex to this allele-specific PAM motif. The safety and allele specificity may be examined in vitro using human cell lines (e.g., HEK293, NT2, HELA, neuronal precursor cell lines, HD fibroblasts (e.g., derived from human patients) and in vivo with a new transgenic mouse expressing the nucleotide variants within the HTT locus.

[0019] As explained above, the therapeutic benefit of RNA interference (RNAi) to reduce mutant huntingtin expression in different mouse models has been shown (Harper et al. (2005) Proc. Natl. Acad. Sci., 102:5820-5825; Boudreau et al. (2009) Mol. Ther., 17:1053-1063; Drouet et al. (2009) Annals of Neurol., 65:276-285). However, RNAi treatment does not completely eliminate mutant huntingtin expression, and the mutant protein remains present at low levels. This remaining protein may mitigate positive effects of RNAi therapy in HD patients. Thus, efforts to eliminate expression of all mutant HTT protein have been sought. Genome editing nucleases present such an opportunity.

[0020] Gene editing based on bacterial endonucleases such as CRISPR-associated protein-9 (Cas9) from *Streptococcus pyogenes* has revolutionized the field (Cong et al. (2013) Science 339:819-823; Ran et al. (2013) Nature Protocols 8:2281-2308; Mali et al. (2013) Science 339:823-826; Jinek et al. (2012) Science 337:816-821). The RNA-guided CRISPR/Cas9 system involves expressing Cas9 along with a guide RNA molecule (gRNA). When coexpressed, gRNAs bind and recruit Cas9 to a specific genomic target sequence where it mediates a double strand DNA (dsDNA) break and activates the dsDNA break repair machinery. Specific DNA fragments can be deleted when two gRNA/Cas9 complexes generate dsDNA breaks at relative proximity, and the genomic DNA (genDNA) is repaired by nonhomologous end joining (FIG. 1A). gRNAs were originally designed for recruiting the CRISPR/Cas9 complex to the HTT promoter to delete regulatory sequences and inhibit HTT expression (FIG. 1B). However, the complete elimination of the expression of both the normal and mutant alleles is likely not tolerable in human adults with HD. As such, an allele specific genome editing approach is provided herein where the CRISPR/Cas9 complex is targeted to the mutant HTT allele.

[0021] The binding specificity of the CRISPR/Cas9 complex depends on two different elements. First, the binding complementarity between the targeted genDNA sequence and the complementary recognition sequence of the gRNA. Second, the presence of a protospacer-adjacent motif (PAM) juxtaposed to the genDNA/gRNA complementary region. Whereas single point mutations in the complementary recognition sequence permit Cas9-mediated DNA cleavage, the preservation of an intact PAM motif is critical (Jinek et al. (2012) Science 337:816-821; Hsu et al. (2013) Nat. Biotech., 31:827-832; Sternberg et al. (2014) Nature 507:62-67). The PAM motif for *S. Pyogenes* Cas9 has been fully

characterized, and is NGG or NAG (Jinek et al. (2012) Science 337:816-821; Hsu et al. (2013) Nat. Biotech., 31:827-832). While any nucleotide type can be found at the first position of the PAM motif, a C/T nucleotide at position 2 and/or a C/T/A nucleotide at position 3 can disrupt the PAM motif and subsequently inhibit Cas9-mediated dsDNA cleavage. Thus, PAM motifs containing single nucleotide polymorphisms (SNP) at positions two or three will confer allele cleavage selectivity when targeted with CRISPR/Cas9 complexes.

[0022] The CAG (encoding glutamine) disease expansion in HTT is located within the 1st exon of the HTT gene (The Huntington's Disease Collaborative Research Group (1993) Cell 72:971-983). A short exon 1 HTT polyadenylated mRNA resulting from aberrant splicing of the mutant allele is reported to be translated into a pathogenic exon 1 HTT protein and contribute to disease progression (Sathasivam et al. (2013) Proc. Natl. Acad. Sci., 110:2366-2370; Gipson et al. (2013) RNA Biol., 10:1647-1652). The main regulatory regions of the HTT promoter are located within the first 2 Kb upstream of the transcription start site (Holzmann et al. (2001) Brain Res., 92:85-97). The 1000 Genome database has reported the location and allele frequency of prevalent SNPs located in the promoter and within the first intronic sequence of the HTT gene (Abecasis et al. (2010) Nature 467:1061-1073). Table 1 provides selected SNPs for allele-specific gene editing on the promoter of the HTT gene. Two examples of SNPs in the first intron of the HTT gene are rs28377140 (3,079,906; G/C; + strand; gain/loss) and rs4498089 (3,080,199; A/G; + strand; gain).

destruction of the target DNA (Jinek et al. (2012) Science, 337:816-821; Sapranaukas et al. (2011) Nucleic Acids Res. 39:9275-9282). The two nucleases generate double-stranded breaks. The double-stranded endonuclease activity of Cas9 requires a target sequence (e.g., ~20 nucleotides) and a short conserved sequence (~2-5 nucleotides; e.g., 3 nucleotides) known as protospacer-associated motif (PAM), which follows immediately 3'—of the CRISPR RNA (crRNA) complementary sequence (Jinek et al. (2012) Science, 337: 816-821; Nishimasu et al. (2014) Cell 156(5):935-49; Swarts et al. (2012) PLoS One, 7:e35888; Sternberg et al. (2014) Nature 507(7490):62-7). Guidelines and computer-assisted methods for generating gRNAs are available (see, e.g, CRISPR Design Tool (crispr.mit.edu/); Hsu et al. (2013) Nat. Biotechnol. 31:827-832; www.addgene.org/CRISPR; and CRISPR gRNA Design tool—DNA2.0 (www.dna20.com/eCommerce/startCas9)). Typically, the PAM sequence is present in the DNA target sequence but not in the gRNA sequence.

[0025] As stated above, wild-type Cas9 creates a site-specific double-stranded DNA break. The double strand break can be repaired by non-homologous end joining (NHEJ) pathway yielding an insertion and/or deletion or, in the presence of a donor template, by homology-directed repair (HDR) pathway for replacement mutations (Overballe-Petersen et al. (2013) Proc. Natl. Acad. Sci. U.S.A. 110:19860-19865; Gong et al. (2005) Nat. Struct. Mol. Biol. 12:304-312). A Cas9 mutant may also be used in the instant invention (e.g., a mutant with an inactivated HNH and/or RuvC nuclease). In a particular embodiment, the mutant is

TABLE 1

Prevalence of the targeted SNPs in the human huntingtin locus. From left to right the table information is: genomic location of the SNP (promoter, 5'UTR), variant ID (SNP identification), location (genomic position), SNP (nucleotide variation), allele frequency based on information obtained from 1000 Genomes data base (Reference nucleotide vs variant nucleotide polymorphism), strand (positive or negative genomic DNA strand), PAM (single nucleotide variation produces either the loss or the loss/gain of a PAM motif at the opposite genomic DNA strand).							
Allele frequency							
Location	Variant ID	Location	SNP	Reference	1000G MAF	Strand	PAM
Promoter	rs35631490	3,071,679	C/G	C = 0.8926	G = 0.1074	+	Loss
Promoter	rs61792464	3,073,385	G/C	G = 0.8628	C = 0.1372	+	Gain/Loss
Promoter	rs9996199	3,074,965	C/G	C = 0.8425	G = 0.1575	+	Gain/Loss
Promoter	rs2857935	3,075,691	C/G/T	C = 0.7710	G = 0.2260	–	Loss
Promoter	rs13122415	3,076,181	C/G	C = 0.8918	G = 0.1082	+	Loss
5'UTR	rs13102260	3,076,405	G/A	G = 0.8419	A = 0.1581	+	Loss

[0023] FIG. 2A provides a schematic of the targeting of Cas9 to allele specific PAM motifs by gRNA. The loss of a PAM (left), gain of a PAM (center), and a loss/gain (right) are depicted. FIG. 2B provides examples of allele specific PAM motifs based on *Streptococcus pyogenes* (SpCas9; PAM—NRG) and *Staphylococcus aureus* Cas9 (SaCas9; PAM—NNGRR, NNGRRT; Ann Ran et al. (2015) Nature 520:186-191). FIG. 2C shows the locations of the alleles listed in Table 1.

[0024] Clustered, regularly interspaced, short palindromic repeat (CRISPR)/Cas9 technology is well known in the art (see, e.g., Sander et al. (2014) Nature Biotech., 32:347-355; Jinek et al. (2012) Science, 337:816-821). Cas9 possesses two nuclease domains, a RuvC-like nuclease domain and a HNH-like nuclease domain, and is responsible for the

Cas9 D10A. Cas9 D10A nicks single-strand DNA rather than generate a double strand break (Cong et al. (2013) Science, 339:819-823; Davis et al. (2014) Proc. Natl. Acad. Sci., 111:E924-932). The nicks are repaired by HDR pathway. Two gRNAs can be used to generate a staggered double strand break with Cas9 D10A.

[0026] In accordance with the instant invention, methods of treating, inhibiting, and/or preventing a polyglutamine disorder (e.g., Huntington's disease) are provided. In accordance with another aspect of the instant invention, methods for reducing the expression of a mutant protein (e.g., mutant huntingtin) encoded by an allele of a gene associated with a polyglutamine disorder (e.g., Huntington's disease) in a cell are provided. Polyglutamine (polyQ) disorder are generally neurodegenerative disorders which are caused by expanded

cytosine-adenine-guanine (CAG) repeats (e.g., greater than about 36 repeats) encoding a long polyQ tract in the respective proteins. Polyglutamine (polyQ) disorders include, without limitation, spinocerebellar ataxia (SCA; types 1, 2, 3, 6, 7, 17), Machado-Joseph disease (MJD/SCA3), Huntington's disease (HD), dentatorubral pallidoluysian atrophy (DRPLA), and spinal and bulbar muscular atrophy, X-linked 1 (SMA1/SBMA). In a particular embodiment, the methods of the instant invention comprise inhibiting, reducing, or eliminating mutant protein (e.g., HTT) expression. In a particular embodiment, the method comprises inactivating (e.g., cleaving and/or deleting (at least partially (e.g., deleting the first exon))) mutant alleles (e.g., HTT alleles) using CRISPR/Cas9 technology. In a particular embodiment, the method comprises administering at least one Cas9 (e.g., a nucleic acid molecule encoding Cas9) and at least one gRNA (e.g., a nucleic acid molecule encoding said gRNA) to said subject. Examples of Cas9 include, without limitation, *Streptococcus pyogenes* Cas9, Cas9 D10A, high fidelity Cas9 (Kleinstiver et al. (2016) Nature, 529:490-495; Slaymaker et al. (2016) Science, 351:84-88), Cas9 nickase (Ran et al. (2013) Cell, 154:1380-1389), *Streptococcus pyogenes* Cas9 with altered PAM specificities (e.g., SpCas9_VQR, SpCas9_EQR, and SpCas9_VRER; Kleinstiver et al. (2015) Nature, 523:481-485), *Staphylococcus aureus* Cas9, the CRISPR/Cpf1 system of *Acidaminococcus*, and the CRISPR/Cpf1 system of Lachnospiraceae. In a particular embodiment, the Cas9 is *S. pyogenes* Cas9. In a particular embodiment, the Cas9 has an inactivated HNH and/or RuvC nuclease, particularly Cas9 D10A. In a particular embodiment, at least two gRNAs are delivered. In a particular embodiment, at least one gRNA targets a region adjacent to a PAM present in only one allele (e.g., in the mutant allele). In a particular embodiment, the targeted PAM is in the 5'UTR, promoter, or first intron. In a particular embodiment, a second gRNA which is not targeted to the allele specific PAM is provided. In a particular embodiment, the second gRNA targets anywhere from the 5'UTR to the 3'UTR of the gene (e.g., HTT gene), particularly within the first intron. In a particular embodiment, the method further comprises the administration of a donor nucleic acid molecule (e.g., DNA; e.g., a nucleic acid molecule encoding the desired sequence). The donor DNA may be a replacement sequence (e.g., wild-type) for the sequence excised from the mutant. The nucleic acids of the instant invention may be administered consecutively (before or after) and/or at the same time (concurrently). The nucleic acid molecules may be administered in the same composition or in separate compositions. In a particular embodiment, the nucleic acid molecules are delivered in a single vector (e.g., a viral vector).

[0027] The methods of the instant invention may also comprise the administration of an additional therapeutic for Huntington's disease or the related disorder. Other therapeutics include, without limitation: haloperidol, tetrabenazine, amantadine, huntingtin antisense, huntingtin siRNA, antidepressants, and antianxiety medications. The nucleic acids of the instant invention and the other therapeutics may be administered consecutively (before and/or after) and/or at the same time (concurrently). The other therapeutics may be administered in the same composition or in separate compositions as the nucleic acid molecules of the instant invention.

[0028] In a particular embodiment, the nucleic acid molecules of the instant invention are delivered (e.g., via infec-

tion, transfection, electroporation, etc.) and expressed in cells via a vector (e.g., a plasmid), particularly a viral vector. The expression vectors of the instant invention may employ a strong promoter, a constitutive promoter, and/or a regulated promoter. In a particular embodiment, the nucleic acid molecules are expressed transiently. Examples of promoters are well known in the art and include, but are not limited to, RNA polymerase II promoters, the T7 RNA polymerase promoter, and RNA polymerase III promoters (e.g., U6 and H1; see, e.g., Myslinski et al. (2001) Nucl. Acids Res., 29:2502-09). Examples of expression vectors for expressing the molecules of the invention include, without limitation, plasmids and viral vectors (e.g., adeno-associated viruses (AAVs), adenoviruses, retroviruses, and lentiviruses).

[0029] In a particular embodiment, the guide RNA of the instant invention may comprise separate nucleic acid molecules. For example, one RNA specifically hybridizes to a target sequence (crRNA) and another RNA (trans-activating crRNA (tracrRNA)) which specifically hybridizes with the crRNA. In a particular embodiment, the guide RNA is a single molecule (sgRNA) which comprises a sequence which specifically hybridizes with a target sequence (crRNA; complementary sequence) and a sequence recognized by Cas9 (e.g., a tracrRNA sequence; scaffold sequence). Examples of gRNA scaffold sequences are well known in the art (e.g., 5'-GUUUUAGAGC UAGAAUAGC AAGUUAUAAU AAGGCUAGUC CGUUAUCAAC UUGAAAAAGU GGCACCGAGU CGGUGCUUUU; SEQ ID NO: 442). As used herein, the term "specifically hybridizes" does not mean that the nucleic acid molecule needs to be 100% complementary to the target sequence. Rather, the sequence may be at least 80%, 85%, 90%, 95%, 97%, 99%, or 100% complementary to the target sequences. The greater the complementarity reduces the likelihood of undesired cleavage events at other sites of the genome. In a particular embodiment, the region of complementarity (e.g., between a guide RNA and a target sequence) is at least about 10, at least about 12, at least about 15, at least about 17, at least about 20, at least about 25, at least about 30, at least about 35, or more nucleotides. In a particular embodiment, the region of complementarity (e.g., between a guide RNA and a target sequence) is about 15 to about 25 nucleotides, about 15 to about 23 nucleotides, about 16 to about 23 nucleotides, about 17 to about 21 nucleotide, or about 20 nucleotides. In a particular embodiment, the guide RNA targets a sequence or comprises a sequence (inclusive of RNA version of DNA molecules) as set forth in the Example or Figures provided herein (see, e.g., the guide or target sequences provided in FIGS. 5 and 6). In a particular embodiment, the guide RNA targets a sequence or comprises a sequence which has at least 80%, 85%, 90%, 95%, 97%, 99%, or 100% homology or identity to a sequence set forth in the Examples and/or Figures (e.g., FIGS. 5 and 6) provided herein. In a particular embodiment, the guide RNA targets a sequence selected from the group consisting of SEQ ID NOs: 156-228 and 379-388. In a particular embodiment, the guide RNA comprises a sequence selected from the group consisting of SEQ ID NOs: 156-228 and 379-388 (e.g., in RNA form). The sequences may be extended or shortened by 1, 2, 3, 4, or 5 nucleotides at the end of the sequence opposite from the PAM (i.e., the 5' end). When the sequence is extended the added nucleotides should correspond to the HTT sequence. In a particular embodiment, one gRNA comprises sgHD1,

sgHD2, sgHD3, sgHD4, sgHD5c, sgHD6c, or sgHD6g. In a particular embodiment, the second gRNA comprises sgHDi3 or sgHDi4. In a particular embodiment, one gRNA comprises sgHD1 and one gRNA comprises sgHDi3.

[0030] As stated hereinabove, the instant invention provides nucleic acid molecules, vectors, and compositions and methods for the inhibition, treatment, and/or prevention of Huntington's disease and related disorders. Compositions comprising at least one nucleic acid described herein are also encompassed by the instant invention. In a particular embodiment, the composition comprises at least one, particularly at least two, guide RNA (e.g., a nucleic acid molecule encoding the guide RNA (e.g., an expression vector)) and at least one pharmaceutically acceptable carrier. The composition may further comprise at least one Cas9 (e.g., a nucleic acid molecule encoding Cas9) and/or at least one donor nucleic acid molecule. The composition may further comprise at least one additional therapeutic (as described above). In a particular embodiment, all of the nucleic acid molecules are encoded within a single expression vector (e.g., viral vector (e.g., AAV)). Alternatively, the other nucleic acid molecules may be contained within a separate composition(s) with at least one pharmaceutically acceptable carrier. The present invention also encompasses kits comprising a first composition comprising at least one guide RNA (e.g., a nucleic acid molecule encoding the guide RNA (e.g., an expression vector)) and a second composition comprising at least one Cas9 (e.g., a nucleic acid molecule encoding Cas9) and/or at least one donor nucleic acid molecule. The first and second compositions may further comprise at least one pharmaceutically acceptable carrier. In a particular embodiment, the kits of the instant invention comprise a first composition comprising at least one guide RNA (e.g., a nucleic acid molecule encoding the guide RNA (e.g., an expression vector)), at least one Cas9 (e.g., a nucleic acid molecule encoding Cas9), and/or at least one donor nucleic acid molecule (optionally all within a single vector) and a second composition comprising at least one additional therapeutic. The first and second compositions may further comprise at least one pharmaceutically acceptable carrier.

[0031] As explained hereinabove, the compositions of the instant invention are useful for treating Huntington's disease and related disorders. A therapeutically effective amount of the composition may be administered to a subject in need thereof. The dosages, methods, and times of administration are readily determinable by persons skilled in the art, given the teachings provided herein.

[0032] The components as described herein will generally be administered to a patient as a pharmaceutical preparation. The term "patient" or "subject" as used herein refers to human or animal subjects. The components of the instant invention may be employed therapeutically, under the guidance of a physician for the treatment of the indicated disease or disorder.

[0033] The pharmaceutical preparation comprising the components of the invention may be conveniently formulated for administration with an acceptable medium (e.g., pharmaceutically acceptable carrier) such as water, buffered saline, ethanol, polyol (for example, glycerol, propylene glycol, liquid polyethylene glycol and the like), dimethyl sulfoxide (DMSO), oils, detergents, suspending agents or suitable mixtures thereof. The concentration of the agents in the chosen medium may be varied and the medium may be

chosen based on the desired route of administration of the pharmaceutical preparation. Except insofar as any conventional media or agent is incompatible with the agents to be administered, its use in the pharmaceutical preparation is contemplated.

[0034] Selection of a suitable pharmaceutical preparation depends upon the method of administration chosen. For example, the components of the invention may be administered by direct injection into any desired tissue (e.g., brain) or into the surrounding area. In this instance, a pharmaceutical preparation comprises the components dispersed in a medium that is compatible with blood or the target tissue.

[0035] The therapy may be, for example, administered parenterally, by injection into the blood stream (e.g., intravenous), or by subcutaneous, intramuscular or intraperitoneal injection. Pharmaceutical preparations for injection are known in the art. If injection is selected as a method for administering the therapy, steps must be taken to ensure that sufficient amounts of the molecules reach their target cells to exert a biological effect.

[0036] Pharmaceutical compositions containing a compound of the present invention as the active ingredient in intimate admixture with a pharmaceutical carrier can be prepared according to conventional pharmaceutical compounding techniques. The carrier may take a wide variety of forms depending on the form of preparation desired for administration, e.g., intravenous, oral or parenteral. In preparing the antibody in oral dosage form, any of the usual pharmaceutical media may be employed, such as, for example, water, glycols, oils, alcohols, flavoring agents, preservatives, coloring agents and the like in the case of oral liquid preparations (such as, for example, suspensions, elixirs and solutions); or carriers such as starches, sugars, diluents, granulating agents, lubricants, binders, disintegrating agents and the like in the case of oral solid preparations (such as, for example, powders, capsules and tablets). Injectable suspensions may be prepared, in which case appropriate liquid carriers, suspending agents and the like may be employed.

[0037] A pharmaceutical preparation of the invention may be formulated in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form, as used herein, refers to a physically discrete unit of the pharmaceutical preparation appropriate for the patient undergoing treatment. Each dosage should contain a quantity of active ingredient calculated to produce the desired effect in association with the selected pharmaceutical carrier. Procedures for determining the appropriate dosage unit are well known to those skilled in the art. Dosage units may be proportionately increased or decreased based on the weight of the patient. Appropriate concentrations for alleviation of a particular pathological condition may be determined by dosage concentration curve calculations, as known in the art.

[0038] The methods of the instant invention may further comprise monitoring the disease or disorder in the subject after administration of the composition(s) of the instant invention to monitor the efficacy of the method. For example, the subject may be monitored for characteristics of Huntington's disease and related disorders.

Definitions

[0039] The singular forms "a," "an," and "the" include plural referents unless the context clearly dictates otherwise.

[0040] The terms “isolated” is not meant to exclude artificial or synthetic mixtures with other compounds or materials, or the presence of impurities that do not interfere with the fundamental activity, and that may be present, for example, due to incomplete purification, or the addition of stabilizers.

[0041] “Pharmaceutically acceptable” indicates approval by a regulatory agency of the Federal or a state government or listed in the U.S. Pharmacopeia or other generally recognized pharmacopeia for use in animals, and more particularly in humans.

[0042] A “carrier” refers to, for example, a diluent, preservative, antioxidant, solubilizer, emulsifier, adjuvant, excipient, bulking substances, auxiliary agent or vehicle with which an active agent of the present invention is administered. Pharmaceutically acceptable carriers can be sterile liquids, such as water and oils, including those of petroleum, animal, vegetable or synthetic origin, such as peanut oil, soybean oil, mineral oil, sesame oil and the like. Water or aqueous saline solutions and aqueous dextrose and glycerol solutions are preferably employed as carriers, particularly for injectable solutions. Suitable pharmaceutical carriers are described, for example, in “Remington’s Pharmaceutical Sciences” by E. W. Martin.

[0043] The term “treat” as used herein refers to any type of treatment that imparts a benefit to a patient suffering from an injury, including improvement in the condition of the patient (e.g., in one or more symptoms), delay in the progression of the condition, etc.

[0044] As used herein, the term “prevent” refers to the prophylactic treatment of a subject who is at risk of developing a condition and/or sustaining an injury, resulting in a decrease in the probability that the subject will develop conditions associated with the injury.

[0045] A “therapeutically effective amount” of a compound or a pharmaceutical composition refers to an amount effective to prevent, inhibit, or treat a particular injury and/or the symptoms thereof. For example, “therapeutically effective amount” may refer to an amount sufficient to modulate the pathology associated traumatic brain injury in a patient.

[0046] As used herein, the term “subject” refers to an animal, particularly a mammal, particularly a human.

[0047] A “vector” is a genetic element, such as a plasmid, cosmid, bacmid, phage or virus, to which another genetic sequence or element (either DNA or RNA) may be attached so as to bring about the replication and/or expression of the attached sequence or element. A vector may be either RNA or DNA and may be single or double stranded. A vector may comprise expression operons or elements such as, without limitation, transcriptional and translational control sequences, such as promoters, enhancers, translational start signals, polyadenylation signals, terminators, and the like, and which facilitate the expression of a polynucleotide or a polypeptide coding sequence in a host cell or organism.

[0048] The following example describes illustrative methods of practicing the instant invention and is not intended to limit the scope of the invention in any way.

Example

[0049] Huntington disease (HD) is a fatal dominantly inherited neurodegenerative disorder caused by CAG repeat expansion (~>36 repeats) within the first exon of the huntingtin gene. Although mutant huntingtin (mHTT) is ubiquitously expressed, the brain shows robust and early degen-

eration. Current RNA interference-based approaches for lowering mHTT expression have been efficacious in mouse models, but basal mutant protein levels are still detected. To fully mitigate expression from the mutant allele, allele specific genome editing can occur via prevalent promoter-resident single nucleotide polymorphisms (SNP) in heterozygosity with the mutant allele. Here, SNPs are identified that either cause or destroy PAM motifs critical for CRISPR selective editing of one allele vs. the other in cell from HD patients, and in a transgenic HD model harboring the human allele.

[0050] Briefly, SNPs within the promoter or intron 1 with an allele frequency >10% in the population were identified using the NCBI website and the 1000 Genome database. 8 SNPs in which the nucleotide change either disrupt and/or generate a PAM motif to allow target specifically of the mutant HTT allele were identified (Table 1 provides certain examples). gRNA sequences targeting selected allele specific PAM motifs were designed and cloned into a CRISPR/Cas9 expression cassette containing a puromycin selection marker.

[0051] Standard human laboratory cell lines (e.g., HEK293, HELA, and NT2), patient fibroblasts, and iPS cells may be genotyped for the selected SNPs. Different cell lines containing the selected nucleotide variants can be used to determine allele selectivity and safety of the approach. For HD fibroblasts containing heterozygous SNPs, direct sequencing of PCR amplified genome sequences containing CAG repeat or SNP linkage by circularization (SLiC) can be used to identify the linkage between CAG repeat length and the nucleotide identity (Liu et al. (2008) Nature Methods 5:951-953). After transfection, genomic DNA can be isolated and CRISPR/Cas9 mediated deletions determined by PCR. The effects on huntingtin expression can be determined by measuring RNA and protein levels by Q-PCR and Western blot, respectively. Unintended cleavage of selected sequences can be determined by using GUIDE-seq method (Tsai et al. (2015) Nature Biotechnol., 33:187-197). Knock in transgenic mice can also be generated where the promoter, the first exon and/or the first intronic sequence of the mouse Htt gene are replaced by the orthologous human HTT sequence containing the different SNP-nucleotide variants on each allele. The therapeutic efficacy of the selected sgRNA sequences (e.g., delivered by an AAV viral vector) can be studied in vivo using current mouse models which exhibit disease phenotypes or the knock in mice.

[0052] As explained above, Huntington Disease (HD) is a fatal neurodegenerative disorder due to polyglutamine (polyQ) disorder caused by triplet CAG repeat expansion in the huntingtin (HTT) gene. Although huntingtin is ubiquitously expressed, the neuropathology of HD is characterized by early striatal atrophy followed by volume loss in other brain areas (Walker, F. O. (2007) Seminars Neurol., 27:143-150; The Huntington’s Disease Collaborative Research Group (1993) Cell, 72:971-983). There is no cure for HD and treatments are focused on symptom management (Johnson et al. (2010) Hum. Mol. Genet., 19:R98-R102). Earlier studies using genetically modified mouse models showed that HD-like phenotypes can be resolved if mutant huntingtin expression is eliminated, even at advanced disease stages (Yamamoto et al. (2000) Cell, 101:57-66; Diaz-Hernandez et al. (2005) J. Neurosci., 25:9773-9781), suggesting that therapeutic strategies focused on eliminating mutant huntingtin expression will be highly beneficial. As examples,

knock down strategies using RNA interference (RNAi) or antisense oligonucleotides, which reduce mutant huntingtin expression either alone or together with the normal huntingtin, are beneficial in various mouse models (Harper et al. (2005) Proc. Natl. Acad. Sci., 102:5820-5825; Boudreau et al. (2009) Mol. Ther., 17:1053-1063; Drouet et al. (2009) Ann. Neurol., 65:276-285; Kordasiewicz et al. (2012) Neuron, 74:1031-1044). Other strategies, such as genome editing with zinc finger nucleases targeted to the CAG-repeat expansion region, have also been tried (Garriga-Canut et al. (2012) Proc. Natl. Acad. Sci., 109:E3136-3145).

[0053] Genome editing with the recently discovered CRISPR/Cas9 system represents an exciting alternative for tackling dominantly inherited genetic disorders such as HD (Jinek et al. (2012) Science, 337:816-821; Mali et al. (2013) Science, 339:823-826; Cong et al. (2013) Science, 339:819-823). The most recent system advancements involves expressing Cas9 along with a guide RNA such as a single guide RNA molecule. When co-expressed, gRNAs bind and recruit Cas9 to a specific genomic target sequence where it mediates a double strand DNA (dsDNA) break, activating the dsDNA break repair machinery. Targeted gene deletions by non-homologous end joining (NHEJ) can be made when a pair of gRNA/Cas9 complexes bind in proximity and produce dsDNA breaks (Cong et al. (2013) Science, 339:819-823; Ran et al. (2013) Nat. Protoc., 8:2281-2308; Jinek et al. (2013) eLife, 2:e00471).

[0054] Given the potency and sequence specificity of the CRISPR/Cas9 targeting, and the fact that huntingtin is an important protein for cell viability (Dragatsis et al. (2000) Nat. Genet., 26:300-306), the use of CRISPR/Cas9 to direct allele specific genome editing is an attractive alternative to the partial reduction approach using ASOs or RNAi methods. Targeting specificity of the CRISPR/Cas9 complex is regulated by two different elements, first, the binding complementarity between the targeted genomic DNA sequence (genDNA) and the ~20 nt-guiding sequence of the gRNA, and secondly, the presence of a protospacer-adjacent motif (PAM) juxtaposed to the genDNA/gRNA complementary region (Jinek et al. (2012) Science, 337:816-821; Cong et al. (2013) Science, 339:819-823; Sternberg et al. (2014) Nature, 507:62-67). While previous studies have shown that nucleotide mismatches at positions 1-10 on the sgRNA-target site interface are not well tolerated for cleavage, sequence context at this region is crucial to determine which nucleotide positions are more effective to influence cleavage (Jinek et al. (2012) Science, 337:816-821; Ran et al. (2013) Nat. Protoc., 8:2281-2308; Sternberg et al. (2014) Nature, 507:62-67; Fu et al. (2013) Nat. Biotech., 31:822-826; Kescu et al. (2014) Nat. Biotech., 32:677-683). However, the preservation of an intact PAM motif appears to be critical and genome wide studies searching for Cas9 off-target cleavage events demonstrate that mutations on the PAM motif result on an important reduction of cleavage efficacy (Anders et al. (2014) Nature, 513:569-573; Kleinstiver et al. (2015) Nature, 523:481-485; Tsai et al. et al. (2015) Nat. Biotech., 33:187-197; Zetsche et al. (2015) Cell, 163:759-771; Ran et al. (2015) Nature, 520:186-191). Therefore, allele specific gene editing can be achieved by taking advantage of prevalent single nucleotide polymorphisms (SNPs) that either eliminate or create a PAM sequence. In HD, polyglutamine repeat expansion occurs within exon 1 (Walker, F. O. (2007) Seminars Neurol., 27:143-150). Because the main regulatory elements for HTT expression

reside within the first 2 Kb 5' of the transcription start site (Coles et al. (1998) Hum. Mol. Genet., 7:791-800), SNP-dependent PAMs in heterozygosity with the mutation are natural CRISPR/Cas9 targets for allele specific editing. Genomic regions adjacent to HTT exon-1 were screened to identify SNPs that were prevalent, and were within the critical position for CRISPR/Cas9- or CRISPR/Cpf1-directed editing. Their utility was tested for allele-specific editing in HD patient cell lines and a mouse model expressing full length mutant human HTT.

Methods

Prediction of SNP-Dependent PAM Motifs:

[0055] SNPs with a prevalence of $\geq 5\%$ located upstream (6.5 Kb) and downstream (Intron 1) HTT exon1 were obtained from the 1000 Genomes database using the NCBI variation viewer website (www.ncbi.nlm.nih.gov/variation/view/?q=HTT&filters=source:dbsnp&asm=GCF_000001405.25). To predict SNP-dependent PAM motifs, SNPs were screened against the consensus PAM sequences of *Streptococcus Pyogenes* (SpCas9, NGG or NAG) and *Staphylococcus aureus* (SaCas9 NNGRRT), or the CRISPR/Cpf1 systems of *Acidaminococcus* (AsCpf1, NTTT) and *Lachnospiraceae* (LbCpf1, heterogenous PAMs). Only those SNPs positioned in a conserved nucleotide PAM position in which the nucleotide variation disrupted the consensus PAM were predicted as SNP-dependent PAM motifs.

Cell Culture and Transfection:

[0056] Human embryonic kidney (HEK293) cells were maintained in DMEM media containing 10% Fetal Bovine Serum (FBS), 1% L-Glutamine and 1% penicillin/streptomycin at 37° C. with 5% CO₂. Cells were cultured in 24 well plates and transfected at 80-90% confluence using Lipofectamine® 2000 transfection reagent, according to the manufacturer's protocol. Human HD patient fibroblasts were maintained on MEM media supplemented with 15% Fetal Bovine Serum (FBS), 1% MEM non-essential amino acids, 1% penicillin/streptomycin and 1% L-Glutamine at 37° C. with 5% CO₂. DNA transfection was done by electroporation using Invitrogen Neon® transfection reagent using the electroporation conditions (ND31551: 1650V, 10 ms, 3 pulses; ND33392: 1450V, 20 ms, 2 pulses), and following the guidelines provided by manufacturer.

sgRNA and Cas9 Plasmid Construction:

[0057] The plasmid pX330 containing the SpCas9 and sgRNA expression cassettes was used as a template for further modifications. To determine transfection efficacy and for selecting positive transfected cells a CMV reporter cassette expressing eGFP/P2A/puromycin fusion protein was cloned downstream of the SpCas9 expression cassette. For all sgRNAs the guide complementary sequences were cloned using a single cloning step with a pair of partially complementary oligonucleotides. The oligo pairs encoding the genomic complementary guide sequences were annealed and ligated into the BbsI cloning site upstream and in frame with the invariant scaffold of the sgRNA sequence.

Genomic DNA Extraction, SNP Genotyping and Genome Editing Analysis:

[0058] Genomic DNA from HD fibroblast and HEK293 cell lines was extracted using DNeasy® Blood & Tissue kit

(Qiagen) according to manufacturer’s instructions. SNPs were genotyped by direct Sanger sequencing of PCR amplified products containing the SNPs and using the primers listed on Table 2. To determine which nucleotide variation of SNP1 (rs2857935) was linked to the normal or the mutant allele the genomic sequence containing SNP1 and the CAG repeat was amplified by PCR and cloned into TOPO® plasmids using the TOPO TA cloning kit, and subsequently

transformed into DH5alpha competent cells. Individual colonies were analyzed using Sanger sequencing to determine which nucleotide variant is associated with the normal or mutant allele. Deletions of HTT exon 1 were confirmed on genomic DNA samples by PCR, using primers binding outside the intervening segment cleaved by the sgRNA/ SpCas9 complex pair (Table 2).

TABLE 2

List of primers and oligonucleotides.	
Oligos to generate guide sequences:	
Name :	Sequence (SEQ ID NOs: 1-20)
PosCRPAS1	caccGCTCCAGGCGTCGGCGG
NegCRPAS1	aaacCCGCCGACGCCTGGAGC
PosCRPAS2	caccGGCGCGGGGCTCAACGGAG
NegCRPAS2	aaacCTCCGTTGAGCCCCGCGCC
PosCRPAS3	caccGTCTGGGACGCAAGGCGCCG
NegCRPAS3	aaacCGGCGCCTTGCGTCCCAGAC
PosCRPAS4	caccGATGCACGCGGGGTGGGGC
NegCRPAS4	aaacGCCCCACCCCGCGTGCATC
PosCRPAS5G	tcccATTAGGTTGATGTCCT
NegCRPAS5G	aaacAGGACATCAACCTGAAT
PosCRPAS5C	tcccATCCCATTCTGAGGTTCTGG
NegCRPAS5C	aaacCCAGAACCTCAGAATGGGAT
PosCRPAS6C	caccGCAGGCAGAGAGGAGCC
NegCRPAS6C	aaacGGCTCCTCTCTGCCTGC
PosCRPAS6G	caccGCCTGGCTAAAGTAGGCTT
NegCRPAS6G	aaacAAGCCTACTTTAGCCAGGC
PosCRI3	caccGCTTTTAGGACGCCICGG
NegCRI3	aaacCCGAGGCGTCCTAAAAGC
PosCRI4	caccGCGGGACACTTCGAGAGG
NegCRI4	aaacCCTCTCGAAGTGTCCCGC
Primers to assess cleavage	
Name :	Sequence (SEQ ID NOs: 21-25)
Fwd1SNP	5'-GAC CAC GCG CAT TCT CT-3'
Fwd4SNP	5'-GGA AAC AGG ACA GAT GAA GGAG-3'
Fwd5SNP	5'-CAG CTC AGA CGG AAG TGT ATT T-3'
Fwd6SNP	5'-CTC CCA AGA ACT GGG AAC TAA C 3'
Rev3Cleavage	5'-ACC ACC GTG ATC ATG AAC TAA A-3'

TABLE 2-continued	
List of primers and oligonucleotides.	
Primers for genotyping	
Name:Sequence (SEQ ID NOs in parentheses)	
Fwd1SNP	5'-GAC CAC GCG CAT TCT CT-3' (21)
Rev1SNP	5'-CGG GAC TGC ATG GTA AGG-3' (26)
Seq1SNP	5'-GCT GTC CGG GTG AGT ATG-3' (27)
FwdSNP2	5'-CCC ACC TCT CAC CTT CCT-3' (28)
RevSNP2	5'-CAG CAT GAT TGA CAG CCC TA-3' (29)
Seq2SNP	5'-CCG CGA CAC TTC ACA CA-3' (30)
FwdSNP3	5'-CCG CGA CAC TTC ACA CA-3' (31)
RevSNP3	5'-TGC TGC TGG AAG GAC TTG 3' (32)
Seq3SNP	5'-TAG GGC TGT CAA TCA TGC TG 3' (33)
Fwd4SNP	5'-GGA AAC AGG ACA GAT GAA GGAG-3' (22)
Rev4SNP	5'-GGG AAT TGA GGG CGG TTT AT-3' (34)
Seq4SNP	5'-TTT ACC AGC TCC TGG CTT TC-3' (35)
Fwd5SNP	5'-CAG CTC AGA CGG AAG TGT ATT T-3' (23)
Rev5SNP	5' GAG CAT GTC CGT GTC CTA ATC-3' (36)
Seq5SNP	5'-TCC CTG GCT AGC ACT TAC TT-3' (37)
Fwd6SNP	5'-CTC CCA AGA ACT GGG AAC TAA C 3' (24)
Rev6SNP	5'-TGT GAT TAG TGC AGC GAG AAG-3' (38)
Seq6SNP	5'-CTG TTT CTC TGC TGT CCT TCT C-3' (39)
Primers for SQ-PCR reaction	
Name:Sequence (SEQ ID NOs: 40-43)	
FwdHTT	TCGGTGCAGCGGCTCCTC
Rev HTT	ATGGCGACCCCTGGAAAAGCTG
FwdActB	TTCGCGGGCGACGATGC
RevActB	CGTACATGGCTGGGGTGTG
Primers to determine off target indels	
Name:Sequence (SEQ ID NOs: 44-65)	
CBFA2T3_Fwd	TCTGTGGTTCAGCCGACTTC
CBFA2T3_Rev	ACACAATACCGTGGCAGAGG
SLC45A_Fwd	GACCCAAGCTTGCCGTAGTA
SLC45A_Rev	ACCTGTTCAGCATCGACGAG
STT3B_Fwd	CCTAACGGACCTGTCGCTTT
STT3B_Rev	TGAGGGACGACTTGTGCTTG
NAV2_Fwd	CCACGAGTGCACACAGTTTG
NAV2_Rev	CTCAAGGACTGCTGGCTCAA
NUP210_Fwd	AGCTGCGTGATCTTGACCAA

TABLE 2-continued
List of primers and oligonucleotides.

NUP210_Rev	GGTGGTTCAGGCTCTTTCCA
DNAJC16_Fwd	TTCATGCACCTCCTCCCAT
DNAJC16_Rev	TGAGTGCAGCGACATGATCA
LEPR_Fwd	TGAGATGTGCCTCCCTCAGA
LEPR_Rev	AACTAGTGGCATGCGTTTGC
CHRNA2_Fwd	CCTTCTGCATGTGGGGTGAT
CHRNA2_Rev	TGAGATCATCCCGTCCACCT
TRIB1_Fwd	TCCCGGGACTTAAAAAGCCG
TRIB1_Rev	ACCTGGTCAAATGGCGTCTT
SMARCD1_Fwd	TATGGTTTTCCCTCCCGGAC
SMARCD1_Rev	AGCAGGTGTGTAAGTGCCTC
COX11_Fwd	GTTAGAGGCTGCGGACCTTT
COX11_Rev	GCCGTTTCTTAGGCCAGAGT

RNA Extraction, RT-QPCR and SQ-PCR of HTT Expression Levels:

[0059] Total RNA was extracted using TRIzol® (Life Technologies, Grand Island, N.Y., USA) according to the manufacturer’s protocol, with the exception of 1 µl Glycoblue™ (Life Technologies, Grand Island, N.Y.) in addition to the aqueous phase on the isopropanol precipitation step and a single wash with cold 70% ethanol. RNA samples were quantified by spectrophotometry and subsequently cDNAs were generated from 1 µg of total RNA with random hexamers (TaqMan® RT reagents, Applied Biosystems). To determine human HTT expression levels in HD fibroblasts and HEK293 cells, TaqMan® probes for human HTT and glyceraldehyde 3-phosphate dehydrogenase (GAPDH) mRNAs obtained from Applied Biosystems were used. For determining human and mouse HTT expression levels in BACHD mice samples TaqMan® Probes for human HTT mRNA, mouse Htt mRNA and mouse beta actin mRNA obtained from Applied Biosystems were used. Relative HTT gene expression was determined using the ddCt method. Allele specific editing was determined by a Semi Quantitative PCR amplification of the CAG repeat within HTT exon 1. RT-PCR experiments were carried out with cDNAs generated from 1 µg of total RNA and using 80 ng for PCR amplification. The RT-exponential phase was determined on 25-30 cycles to allow semi-quantitative (SQ) comparison of cDNAs developed from identical reactions with Biolase™ Taq Polymerase (Bioline Inc, MA). A SQ-PCR reaction for Actin B (20 cycles) was used as a reference gene to determine loading differences between samples. The primers are shown in Table 2. Briefly, the high capacity cDNA kit from Applied Biosystems was used for the Rt-reaction with random hexamers.

Huntingtin Western Blots:

[0060] HEK293 cells were transfected with sgRNA/Sp-Cas9 expression cassettes, selected for 2 days with puromy-

cin (3 µM) and expanded until cells reached confluence. Then, cells were rinsed once with PBS and lysed with passive lysis buffer (PBL, Promega). Protein concentrations were determined using the DC protein assay (BioRad) and 15 µg of protein loaded on a 3-8% NuPAGE® Tris-Acetate gel (Novex Life Technologies). HD fibroblast cells were electroporated with sgRNA/SpCas9 expression cassettes, selected for 2 days with puromycin (204) and expanded until cells reached confluence. Cells were then rinsed with iced-cold PBS, de-attached, pelleted, snap froze, and lysed with SDP lysis buffer (50 mM Tris pH8.0, 150 mM NaCl, 1% NP40, 1× complete protease inhibitors, 1×phosphatase inhibitors, 100 mM PMSF) followed by incubation on ice for 20 minutes with occasional vortexing. Debris was removed by centrifugation (15 min, 20,000 g 4° C.) and the supernatant retained. Protein concentrations were determined using the DC protein assay (BioRad). Samples (25 µg) were prepared for immunoblotting by denaturing the lysates in LDS sample buffer (Invitrogen) with 2×reducing agent (100 mM DTT, Invitrogen) and heating to 70° C. for 10 minutes. Samples were resolved on a 10% low-bis acrylamide gels (200:1 acrilamide:Bis) with Tris-glycine running buffer (25 mM Tris, 190 mM Glycine, 0.1% SDS) containing 10.7 mM Beta mercaptoethanol. Gels were run on ice for 40 minutes at 90V through the stack, then at 190 V. Proteins were transferred overnight at 30V and 4° C. onto polyvinylidene fluoride (PVDF) membranes with NuPage® transfer buffer (Invitrogen: 25 mM Bicine, 25 mM Bis-Tris, 1.025 mM EDTA, 5% MeOH, pH7.2). Membranes were blocked with 5% milk in PBS-T and then blotted with a Human anti-HTT (1:5000, Millipore, CA, USA), or rabbit anti beta-actin (1:40000, Sigma) antibodies followed by horseradish peroxidase-coupled antibodies (1:10,000, mouse; or 1:50,000, Rabbit; Jackson ImmunoResearch, West Grove, Pa.). Blots were developed with ECL Plus reagents (Amersham Pharmacia). HTT reduction was determined by densitometry (n=3 independent experiments) of

protein levels relative to beta catenin with the VersaDoc™ Imaging System (Biorad) and Quantity One analysis software.

rAAV Vector Design and Production:

[0061] For in vivo studies, two different rAAV vectors were generated. One expressed SpCas9 and one the sgRNAs expression cassettes. SpCas9 was expressed under the control of a minimal cytomegalovirus immediate-early gene enhancer/promoter region (CMV promoter) and cloned upstream of a minimal poly A sequence (FBAAV-Cas9). The sgRNA expression cassettes were moved into an AAV shuttle plasmid containing an eGFP gene under the control of the CMV promoter and upstream of a SV40 pA signal. All rAAV plasmid shuttles have AAV2 inverted terminal repeat sequences. RAAV vectors were produced by standard calcium phosphate transfection method in HEK293 cells by using the Ad Helper, AAV1 transpackaging and AAV shuttle plasmids. Vector titers were determined by RT PCR and were 1×10^{13} vg/ml. Vector purity was also tested by silver stain.

Off-Target Analysis:

[0062] Potential off target loci for sgHD guide sequences in the human genome were determined using the Cas9-Off finder algorithm (Bae et al. (2014) Bioinform., 30:1473-1475). Genomic DNA was extracted from HD human fibroblasts electroporated with sgHD1 and sgHDi3 and amplicons generated with Phusion® polymerase using PCR primers flanking the potential site. Amplicons were subjected to Sanger sequencing to determine mutations in the cleavage site using specific primers, as well as cloned into TOPO-cloning system for sequence confirmation using 3-4 colonies/site.

Mouse Studies:

[0063] BACHD mice were obtained from Jackson Laboratories (Bar Harbor, Me.). Mice were housed in a temperature-controlled environment on a 12-hour light/dark cycle. Food and water were provided ad libitum. Mice were injected with a combination 1:1 of rAAV2/1-SpCas9 virus and rAAV-hU6sgRNA/eGFP virus. For rAAV injections, mice were anesthetized with isoflurane, and 5 µl of rAAV mixture injected unilaterally into the right striata at 0.2 µl/min (coordinates: +0.86 mm rostral to Bregma, +/-1.8 mm lateral to medial, ~2.5 mm ventral from brain surface). After 3 weeks, mice were anesthetized with a Ketamine and Xylazine mix and perfused with 18 ml of 0.9% cold saline mixed with 2-ml RNAlater® (Ambion) solution. Brains were removed, blocked and cut into 1-mm-thick coronal slices. Tissue punches from striata were taken using a tissue corer (1.4-mm in diameter; Zivic Instruments, Pittsburgh, Pa.). All tissue punches were flash frozen in liquid nitrogen and stored at -80° C. until use.

Statistical Analysis

[0064] All statistical analyses were performed using Graph-Pad Prism v5.0 software. All data was analyzed using one-way ANOVA followed by a Bonferroni's post-hoc, or a Mann Whitney test as indicated. Statistical significance was considered * P<0.05, P<0.01, § P<0.001, #P<0.0001.

Results

Screening SNP-Derived PAM Motifs in the HTT Genomic Locus

[0065] A goal was to delete the mutant HTT allele using SNP-dependent PAMs flanking HTT exon 1 that when present in heterozygosity would tether the Cas9 protein to the mutant, but not the normal allele (FIG. 3A). The CRISPR/SpCas9 system from *Streptococcus pyogenes* is the most widely used, and its PAM sequence (NRG, where N represents any nucleotide, R a purine and the conservation of a guanine) have been characterized (Jinek et al. (2012) Science, 337:816-821; Cong et al. (2013) Science, 339:819-823). SNPs present at specific PAM positions could generate, remove, or simultaneously do both in a strand specific way (FIG. 3B). Using the NCBI website and the 1000 Genomes database, the upstream (3 Kb, Promoter/5'UTR) and downstream (6.5 Kb, Intron 1) genomic sequences of HTT exon-1 were screened and 47 SNPs with a prevalence of more than 5% were identified. Of these, 21 were present within the conserved third nucleotide of the PAM sequence of SpCas9 (FIG. 3C, FIG. 5A-5O, FIGS. 3D-3J). Overall, the nucleotide variation caused the loss (8 SNPs), gain (8 SNPs), or simultaneously the loss in one DNA strand and the gain on the opposite strand (5 SNPs, Loss/Gain) (FIG. 1C, FIGS. 3D-3J). FIG. 3K shows the flanking sequence of the indicated SNPs, the location of the PAMs, and an example of complementary sequence to target the site for cleavage.

Experimental Validation of HTT Promoter SNP-Dependent PAM Motifs

[0066] Small guide RNAs (sgRNAs; single guide RNAs) were generated that bind adjacent to PAM sequences representing 5 of the identified SNPs in the upstream ~6 Kb (SNPs 1, 2, 4, 5, and 6) and to a SNP-dependent PAM near the transcription start site (SNP3) (FIGS. 4A, 4B, 6A, and 6B) to test as candidates for CRISPR/Cas9 cleavage in HEK 293 cells. These SNPs have an allele frequency of >10% in the general population, and the nucleotide variations cause the Loss or a Loss/Gain of the PAM motif (Table 3). Common sgRNAs were also designed for the first HTT intron (sgHDi3 and sgHDi4; FIG. 4A and FIGS. 6A and 6B). The sgRNAs were cloned downstream of the hU6 or hH1 promoter, along with other elements as depicted (FIG. 2C). HEK 293 cells, which are homozygous for the targeting SNPs (FIG. 4B) were transfected with SpCas9 and sgRNA expression plasmids and genomic deletion assessed. DNA products of the anticipated size were amplified in most of the sgRNA/SpCas9 pair complexes tested (FIG. 4D, 4E, 4F). As expected, HTT remained intact in cells expressing SpCas9 or sgHDi3 only, or co-expressing sgHDi3 with a sgRNA sequence for which a PAM sequence is absent in the HTT promoter (sgHD5c/i3 and sgHD6g/i3). HTT exon1 cleavage in cells was not detected transfected with sgHD5g/i3, in spite of the presence of the PAM. Both sgHD1 and sgHD5g have a 17 nt complementary sequence, yet sgHD1/i3 eliminated HTT exon 1 while sgHD5/i3 did not. Interestingly, sgHD1 has 8 guanines, 6 cytosines and 1 adenosine whereas sgHD5g has 4 guanines, 3 cytosines and 3 adenosines. This is consistent with work showing a direct correlation between the sequence composition of the sgRNA complementary region to sgRNA activity, with the most active sequences enriched for guanine and cytosine and depleted of adenosine

(Moreno-Mateos et al. (2015) Nat. Meth., 12:982-988). Sanger sequencing of the small-amplified PCR products confirmed HTT exon1 deletion and dsDNA repair (FIG. 4J).

TABLE 3

List of prevalent SNPs located upstream of human HTT exon1. SNP ID, location, nucleotide variation, allele nucleotide frequency, strand and the effect of the nucleotide variation on the PAM sequence are indicated.							
Allele frequency							
SNP ID	Variant ID	Location	SNP	Reference	1000G MAF	Strand	PAM
SNP4	rs35631490	3,071,679	C/G	C = 0.8926	G = 0.1074	+	Loss
SNP5	rs61792464	3,073,385	G/C	G = 0.8628	C = 0.1372	+	Gain/Loss
SNP6	rs9996199	3,074,965	C/G	C = 0.8425	G = 0.1575	+	Gain/Loss
SNP1	rs2857935	3,075,691	C/G/T	C = 0.7710	G = 0.2260	–	Loss
SNP2	rs13122415	3,076,181	C/G	C = 0.8918	G = 0.1082	+	Loss
SNP3	rs13102260	3,076,405	G/A	G = 0.8419	A = 0.1581	+	Loss

[0067] HTT mRNA and protein levels were reduced in cells following editing, as determined by Q-PCR and western blot, respectively (FIGS. 4G, 4H, 4I, and 4K). Reduction of HTT mRNA levels was greater in cells expressing sgRNA/SpCas9 complex pairs that generated small targeted deletions, indicating that HTT exon 1 removal efficacy may be influenced by the distance between the two dsDNA breaks (compare sgHD1, 2 and 3 versus sgHD4 and 6) (FIG. 4G). Also, the results indicate a preference of SpCas9 for NGG over NAG PAM sequences (compare sgHD1, 2, 3, 4 (NGG) vs sgHD6c (NAG)) (FIG. 4G). Interestingly, cells expressing sgHDi3 alone, or in combination with sgHD6g or sgHD5g, also showed reduced HTT mRNA and protein levels albeit not to as great an extent as those where HTT exon 1 was removed. This indicates that elements within the first intron, disrupted after DNA repair, could affect expression of the normal HTT allele (FIGS. 4G, 4H, 4I).

[0068] FIG. 4L provides a schematic depicting the strategy to delete 1st exon sequence of mutant huntingtin gene using the CRISPR/Cas9 system targeting the SNP-derived PAM motifs at positions SNP1, SNP2, SNP3 and SNP6 in combination with the intronic i3 PAM motif. FIG. 4M provides a schematic depicting the strategy to delete 1st exon sequence of mutant huntingtin gene using the CRISPR/Cas9 system targeting the SNP6 in combination with i3 PAM

motif. FIG. 4N provides a schematic depicting the strategy to delete 1st exon sequence of mutant huntingtin gene using the CRISPR/Cas9 system targeting the SNP4, SNP5 derived PAM motifs in combination with i3 PAM motif.

Assessment of Allele Specificity of Editing in HD Human Fibroblasts.

[0069] Expression vectors for sgHD1/i3 and SpCas9, or SpCas9 only, were generated (FIG. 7A) for screening in HD fibroblast cell lines. Twenty-three lines were screened for SNP heterozygosity using direct Sanger sequencing of PCR amplified products. Eleven lines were heterozygous for SNP1, 1 line was heterozygous for SNP2, SNP4 and SNP6, and 2 lines were heterozygous for SNP3 and SNP5 (Table 4). Two lines, ND31551 and ND33392, which are heterozygous for the SNP1 on opposite alleles, were chosen for specificity testing (FIG. 7B, Table 5). PCR of genomic DNA showed target cleavage in cells transfected with plasmids expressing sgHD1/i3 and SpCas9 relative to those lacking sgRNAs (FIG. 7C). Semi-quantitative PCR for the normal and mutant HTT mRNAs showed target mRNA knockdown (FIGS. 7D, 7E, 7F), which for ND31551 is the normal allele, and for ND33392 is the mutant allele. Western blot for protein confirmed allele specific reduction of the target allele (FIGS. 7G, 7H).

TABLE 4

List of Huntington's disease (HD) fibroblast genotypes for the various SNPs. HD fibroblast ID, CAG repeat length of the normal and mutant allele, and presence or absence of the SNP heterozygosity are indicated.							
HD Fibroblast	CAG repeat	SNP1	SNP2	SNP3	SNP4	SNP5	SNP6
		rs2857935 (C > G/T)	rs13122415 (C > G)	rs13102260 (G > A)	rs35631490 (C > G)	rs61792464 (C > G)	rs9996199 (C > G)
		0.226	0.1082	0.1581	0.1074	0.1372	0.1575
GM04723	CAG:72	C/G	C	G	C	G	C
GM04869	CAG:50	C	C	G	C	G	C
GM04767	CAG:47	C/G	C	G	C	G	C
GM04867	ND	C	C	G	C	G	C
GM04849	ND	C	C	G	C	G	C
GM04689	CAG:46	C	C	G	C	G	C
ND29801	CAG:40	C	C	G	C	G	C
ND29970	CAG:40	C	C	G/A	C	G/C	G/C
ND30013	CAG:43	C	C	G	C	G	C
ND30015	CAG:41	C/G	C	G	C	G	C
ND30016	CAG:41	C/G	C	G	C	G	C
ND30047	CAG:41	C/G	C	G	C	G	C
ND30259	CAG:38	C	C	G	C	G	C

TABLE 4-continued

List of Huntington's disease (HD) fibroblast genotypes for the various SNPs. HD fibroblast ID, CAG repeat length of the normal and mutant allele, and presence or absence of the SNP heterozygosity are indicated.							
HD Fibroblast	CAG repeat	SNP1	SNP2	SNP3	SNP4	SNP5	SNP6
		rs2857935 (C > G/T)	rs13122415 (C > G)	rs13102260 (G > A)	rs35631490 (C > G)	rs61792464 (C > G)	rs9996199 (C > G)
		0.226	0.1082	0.1581	0.1074	0.1372	0.1575
ND30422	CAG:40	C/G	C	G	C	G	C
ND30626	CAG:41	C/G	C	G	C	G	C
ND30967	CAG:43	C	C	G	C	G	C
ND31038	CAG:44	C/G	C	G	C	G	C
ND31551	CAG:39	C/G	C/G	G/A	C/G	G	G/C
ND31846	CAG:40	C	C	G	C	G	C
ND33392	CAG:57	C/G	C	G	C	G	C
ND33947	CAG:40	C/G	C	G	C	G	C
ND40536	CAG:66	C	C	G	G	G	C
ND40534	CAG:46/26	C	C	G	C	G	C

TABLE 5

Genotypes of 11 HD fibroblast lines heterozygous for SNP1. Fibroblast ID, CAG repeat length, nucleotide variation for normal and mutant allele, allele that contains the PAM motif and family ID of the HD fibroblast line are indicated.						
HD Fibroblast	CAG repeat	SNP	Huntingtin Allele			Family
			Normal	Mutant	Targeted	
GM04723	CAG:72/17	C/G	G	C	Mutant	691
GM04767	CAG:47/18	C/G	G	C	Mutant	691
ND30015	CAG:41/20	C/G	G	C	Mutant	NINDS3749
ND30016	CAG:41/21	C/G	G	C	Mutant	NINDS3749
ND30047	CAG:41/18	C/G	G	C	Mutant	NINDS3753
ND30422	CAG:40/18	C/G	G	C	Mutant	NINDS3751
ND30626	CAG:41/21	C/G	G	C	Mutant	NINDS3752
ND31038	CAG:44/19	C/G	C	G	Normal	NINDS3752
ND31551	CAG:39/18	C/G	C	G	Normal	Unknown
ND33392	CAG:57/17	C/G	G	C	Mutant	NINDS4250
ND33947	GAG:40/18	C/G	G	C	Mutant	Unknown

Assessment of Off-Target Cleavage Sites in HD Human Edited Fibroblasts.

[0070] Although truncated sgRNA sequences (~<20 nt) are reported to have higher selectivity for the on-target site, any sgRNA/Cas9 complex can also generate unwanted dsDNA breaks at off-target sites that resemble the on-target sequence (Fu et al. (2013) Nat. Biotech., 31:822-826; Kleinstiver et al. (2015) Nature, 523:481-485; Hsu et al. (2013) Nat. Biotech., 31:827-832; Cho et al. (2013) Nat. Biotech., 31:230-232). The Cas9-Off finder algorithm was used to predict the number of potential off-target sites for the most effective sgRNAs (sgHD1, sgHD2, sgHD3 and sgHDi3), and the UCSC genome browser for mapping their location in the human genome (Bae et al. (2014) Bioinform., 30:1473-1475). The screen identified 416 sites for sgHD1, whereas 40, 21 and 7 off-targets are predicted for sgHD2, sgHDi3 and sgHD3, respectively (FIG. 8A). Of note, sgHD1 has the shortest complementary sequence (17 nt), which could explain its higher frequency for genomic off-targets. Importantly, all guides showed full complementary only to HTT, and more than 90% of the off-targets have 3 mismatches. As expected, they occur in the promoter, 5'UTR, exons, introns, 3' UTR, and intergenic regions. The highest number was predicted within introns (FIG. 8B). HD fibroblasts were electroporated with vectors expressing SpCas9 and sgHD1/i3, along with a short ODN sequence for map-

ping off-target dsDNA breaks (Tsai et al. (2015) Nat. Biotech., 33:187-197). As expected, the ODN was incorporated within the HTT gene locus (FIG. 8C) but it was not detected in any of the 11 top off-target sites tested (FIG. 8D).

Allele-Specific Editing In Vivo.

[0071] BachHD mice are transgenic for a modified human HD allele (Gray et al. (2008) J. Neurosci., 28:6182-6195), which fortuitously contains SNPs 1, 2 and 3. These mice were used to first evaluate the efficacy of mutant HTT editing in vivo at the genomic level. For this, recombinant AAVs (rAAVs) expressing either SpCas9 (rAAV.SpCas9) or the sgRNAs (rAAV.sgHD1/i3) were generated, which effectively delete human HTT exon 1 in vitro (FIG. 9A, 9D). Mice were injected on the right hemisphere with rAAV.SpCas9 plus rAAV.sgHD1B/i3. The left hemisphere was used as a control and left uninjected. Brains were harvested 3 weeks later and genomic DNA isolated. PCR amplification of genomic DNA demonstrates cleavage in the setting of Cas9 and sgRNA expressing AAVs only (FIG. 9B). Accordingly, HTT mRNA levels reduced on the right, but not the left hemisphere, in concordance with DNA cleavage (FIG. 9C). Interestingly, mouse Httm RNA levels were also reduced on the injected hemisphere, although to a lesser degree than the human HTT allele. Several binding sites for sgHD1 and sgHDi3 were identified within the mouse Htt

locus. Three binding sites were identified for sgHD1, all within exon 1 and containing 5 mismatches. In contrast, a single binding site with 2 mismatches within the first intron was predicted for sgHDi3 (FIG. 9E). Indels within the intron caused by sgHDi3 may disrupt transcription factor binding sites, and similar to HEK293 cells expressing sgHDi3 alone (FIG. 4G), reduce mouse HTT expression.

Other Cas9 Systems for Silencing the HTT Allele.

[0072] It was then screened which of the 47 SNPs flanking HTT exon1 were contained within their conserved PAM nucleotide positions. Engineered SpCas9 variants from *Streptococcus Pyogenes* with altered PAM specificities have been generated (SpCas9 VQR, SpCas9 EQR, and SpCas9 VRER) (Kleinstiver et al. (2015) Nature, 523:481-485). The SpCas9 VQR variant strongly recognizes sequences bearing the NGAN PAM and with lower efficiency those sites with a NGNG motif. SpCas9_EQR is more specific for an NGAG PAM. In contrast, SpCas9_VRER has a strong selectivity for a NGCG PAM sequence with no cleavage activity when this is varied. For SpCas9 VQR, the SNP could be positioned at the 2nd or the 3rd nucleotide of the NGAN PAM, as well as at the 2nd and 4th nucleotide of the NGNG PAM sequence. In contrast, because of the selectivity of the SpCas9 EQR for NGAG and SpCas9 VRER for NGCG sequences, the SNP could be permitted at any position of their PAM (FIG. 5A-5O, FIGS. 3D-3J). The discovery of SaCas9 from *Staphylococcus aureus* has extended the number of CRISPR/Cas9 systems, with the advantage that a SaCas9-encoding transgene can be easily package into AAV viral vectors (Ran et al. (2015) Nature, 520:186-191). SaCas9 primarily recognizes a NNGRRT PAM, although dsDNA breaks are also observed at DNA targets adjacent to NNGRR motifs. For SaCas9 only those SNPs positioned at the 3rd nucleotide of the PAM would allow for allele specificity (FIG. 5A-5O, FIGS. 3D-3J). A new Class 2 CRISPR system was recently identified that contains Cpf1 as effector protein to mediate dsDNA breaks (Zetsche et al. (2015) Cell, 163:759-771). Unlike Cas9 that recognizes a G-rich PAM motif, the Cpf1 PAM motif is T-rich. Currently, 16 Cpf1-family proteins have been characterized, but only the Cpf1 proteins from *Acidaminococcus* (AsCpf1) and *Lachnospiraceae* (LbCpf1) have shown robust DNA interference activity when expressed in mammalian cells. AsCpf1 has strong selectivity for a TTTN PAM and does not recognize any sequence variants. Therefore, SNP presents at any position of the TTTN PAM could disrupt AsCpf1 recognition. In contrast, LbCpf1 recognizes multiple T-rich PAMs, albeit with different cleavage activity. Thus, for LbCpf1, only those SNPs where the variant nucleotide did not generate any other PAM that could be recognized above LbCpf1 cleavage threshold activity could be considered for allele discrimination (Zetsche et al. (2015) Cell, 163:759-771) (FIGS. 5A-5O, FIGS. 3D-3J).

[0073] Overall, 36 SNPs located within the specific PAM positions described above were identified. Again, instances where the nucleotide variation caused the loss (12 SNPs), gain (11 SNPs), or a simultaneous loss in one DNA strand and a gain on the opposite strand (13 SNPs) were identified (FIGS. 5A-5O, FIGS. 3D-3J). Of special interests are the SNPs that generate a Loss/Gain, since CRISPR complexes could be designed for any of the two possible nucleotides linked to the mutant allele. Of note, instances were found where the same Cas9 protein could target each nucleotide

variation using a different sgRNA sequence, or alternatively, a different CRISPR effector protein could be used to target each nucleotide variant. Two interesting observations also arose from the screen. One, in the rs113331544 SNP, for which the minor allele contains a six-nucleotide insertion, the same PAM sequence is present on both alleles, but a different sgRNA sequence could be designed to tether SpCas9 to the mutant allele depending on the nucleotide variation. Two, for the rs28393280 and the rs28583447 SNPs, the nucleotide variation causes the gain of 2 PAM motifs on the same allele, one on the positive and the other on the negative DNA strand. Those SNPs could be appropriate for targeting with a nickase effector protein, which would efficiently generate on-target dsDNA breaks without detectable damage at potential off-target sites (Ran et al. (2013) Cell, 154:1380-1389).

[0074] Currently, reduction of HTT mRNA levels with RNAi and ASOs are the leading therapeutic options for HD (Kordasiewicz et al. (2012) Neuron, 74:1031-1044; McBride et al. (2011) Mol. Ther., 19:2152-2162). However, it is unknown whether these treatments will be beneficial in HD patients, since the mutant protein is not completely eliminated. Additionally, the normal allele is reduced relative to normal levels as a consequence of the non-allele specific gene silencing approach.

[0075] Targeted gene deletions can be generated when two sgRNA/Cas9 complexes cause dsDNA breaks followed by DNA repair (Cong et al. (2013) Science, 339:819-823; Jinek et al. (2013) eLife, 2:e00471). Given the potency of CRISPR/Cas9 and the high likelihood of cleaving both HTT alleles, the role of HTT protein on important cellular functions, and the fact that a complete loss of the huntingtin gene in adult mice causes progressive neurodegeneration (Dragatsis et al. (2000) Nat. Genet., 26:300-306), allele specificity for editing is imperative. Earlier work demonstrated that Cas9 causes dsDNA breaks when mismatches are present between the guide and the targeted sequence, but only if a PAM motif is near the target sequence (Fu et al. (2013) Nat. Biotech., 31:822-826; Hsu et al. (2013) Nat. Biotech., 31:827-832). Genome wide studies and in vitro library screens have provided information regarding the conservation for each nucleotide within a PAM sequence for several of the available Cas9 proteins (Kleinstiver et al. (2015) Nature, 523:481-485; Tsai et al. (2015) Nat. Biotech., 33:187-197; Zetsche et al. (2015) Cell, 163:759-771; Ran et al. (2015) Nature, 520:186-191). Cas9 PAM recognition could be disrupted on a single allele if a SNP located at these conserved nucleotides were present in heterozygosity. Thus, single allele targeted deletions could be generated to mitigate the expression of the mutant, but not the normal allele.

[0076] Guide RNAs were designed that bind and tether SpCas9 to six prevalent SNPs located 5' of HTT exon1, which in combination with a guide binding within the first HTT intron effectively eliminate expression of the HTT protein. The distance between upstream and downstream guides influenced editing efficacy, as well as confirmed the SpCas9 preference in HD cell lines. The studies also indicate that intronic transcription binding sites may effect HTT gene expression, since indels generated by SpCas9 within the HTT intron reduced gene expression. This is important when designing intronic guide sequences, since expression of the normal allele could also be affected. The allele specificity of the instant approach was demonstrated using human fibroblast cell lines for which these SNPs are present in het-

erozygosity. HTT exon 1 excision was observed only on the alleles where the nucleotide variation did not disrupt the PAM motif.

[0077] Interestingly, SNP1 (rs2857935) has a prevalence of 22% among the human population. In the HD fibroblast lines, 9 out of 11 were heterozygous for the SNP and the PAM was linked to the mutant allele. This raises the exciting possibility that this SNP is in linkage disequilibrium with the mutant allele in the general HD population.

[0078] The importance of on-target selectivity is crucial when using gene-editing approaches. In the instant strategy truncated sgRNA guides were used, which have been shown to minimize unintended dsDNA breaks (Fu et al. (2014) Nat. Biotech., 32:279-284). Potential off-targets from the guides were screened for using an in silico approach, and most of the off-target binding sites contained 3 mismatches within intronic regions. Notably, additional tools with significant on-target selectivity such as the High fidelity Cas9 proteins and the Cas9 nickases can be used in the instant methods (Ran et al. (2013) Cell, 154:1380-1389; Kleinstiver et al. (2016) Nature, 529:490-495; Slaymaker et al. (2016) Science, 351:84-88).

[0079] The approach was also demonstrated in vivo using an HD mouse model. rAAV delivery of the sgRNA/SpCas9

complexes reduced human mutant HTT expression to 40% in treated hemispheres, a level of reduction known to provide benefit by RNAi or ASOs (Harper et al. (2005) Proc. Natl. Acad. Sci., 102:5820-5825; Boudreau et al. (2009) Mol. Ther., 17:1053-1063; Kordasiewicz et al. (2012) Neuron, 74:1031-1044). Notably, Cas9 and/or the sgRNAs may be transiently expressed in the instant methods (Hendel et al. (2015) Nat. Biotech., 33:985-989; Randar et al. (2015) Proc. Natl. Acad. Sci., 112:E7110-7117).

[0080] Thus, a strategy for allele specific genome-editing of mutant HTT based on CRISPR/Cas9 technology has been developed that takes advantage of highly prevalent SNPs at the HTT locus for guiding mutant allele specific cleavage.

[0081] A number of publications and patent documents are cited throughout the foregoing specification in order to describe the state of the art to which this invention pertains. The entire disclosure of each of these citations is incorporated by reference herein.

[0082] While certain of the preferred embodiments of the present invention have been described and specifically exemplified above, it is not intended that the invention be limited to such embodiments. Various modifications may be made thereto without departing from the scope and spirit of the present invention, as set forth in the following claims.

SEQUENCE LISTING

Sequence total quantity: 442		
SEQ ID NO: 1	moltype = DNA length = 21	
FEATURE	Location/Qualifiers	
misc_feature	1..21	
	note = PosCRPAS1	
source	1..21	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 1		
caccgctcca ggcgtcggcg g		21
SEQ ID NO: 2	moltype = DNA length = 21	
FEATURE	Location/Qualifiers	
misc_feature	1..21	
	note = NegCRPAS1	
source	1..21	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 2		
aaacccgccg acgctggag c		21
SEQ ID NO: 3	moltype = DNA length = 23	
FEATURE	Location/Qualifiers	
misc_feature	1..23	
	note = PosCRPAS2	
source	1..23	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 3		
caccggcgcg gggctcaacg gag		23
SEQ ID NO: 4	moltype = DNA length = 23	
FEATURE	Location/Qualifiers	
misc_feature	1..23	
	note = NegCRPAS2	
source	1..23	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 4		
aaacctccgt tgagccccgc gcc		23
SEQ ID NO: 5	moltype = DNA length = 24	
FEATURE	Location/Qualifiers	
misc_feature	1..24	

-continued

source	note = PosCRPAS3 1..24 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 5		
caccgtctgg gacgcaaggc gccg		24
SEQ ID NO: 6	moltype = DNA length = 24	
FEATURE	Location/Qualifiers	
misc_feature	1..24	
	note = NegCRPAS3	
source	1..24 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 6		
aaaccggcgc cttgcgtccc agac		24
SEQ ID NO: 7	moltype = DNA length = 23	
FEATURE	Location/Qualifiers	
misc_feature	1..23	
	note = PosCRPAS4	
source	1..23 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 7		
caccgatgca cgcggggtgg ggc		23
SEQ ID NO: 8	moltype = DNA length = 23	
FEATURE	Location/Qualifiers	
misc_feature	1..23	
	note = NegCRPAS4	
source	1..23 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 8		
aaacgccccca ccccgctgc atc		23
SEQ ID NO: 9	moltype = DNA length = 21	
FEATURE	Location/Qualifiers	
misc_feature	1..21	
	note = PosCRPAS5G	
source	1..21 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 9		
tcccattcag gttgatgtcc t		21
SEQ ID NO: 10	moltype = DNA length = 21	
FEATURE	Location/Qualifiers	
misc_feature	1..21	
	note = NegCRPAS5G	
source	1..21 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 10		
aaacaggaca tcaacctgaa t		21
SEQ ID NO: 11	moltype = DNA length = 24	
FEATURE	Location/Qualifiers	
misc_feature	1..24	
	note = PosCRPAS5C	
source	1..24 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 11		
tcccatccca ttctgaggtt ctgg		24
SEQ ID NO: 12	moltype = DNA length = 24	
FEATURE	Location/Qualifiers	
misc_feature	1..24	
	note = NegCRPAS5C	
source	1..24 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 12		
aaaccagaa cctcagaatg ggat		24

SEQ ID NO: 13	moltype = DNA	length = 21	
FEATURE	Location/Qualifiers		
misc_feature	1..21		
	note = PosCRPAS6C		
source	1..21		
	mol_type = other DNA		
	organism = synthetic construct		
SEQUENCE: 13			
caccgcaggc agagaggagc c			21
SEQ ID NO: 14	moltype = DNA	length = 21	
FEATURE	Location/Qualifiers		
misc_feature	1..21		
	note = NegCRPAS6C		
source	1..21		
	mol_type = other DNA		
	organism = synthetic construct		
SEQUENCE: 14			
aaacggctcc tctctgctg c			21
SEQ ID NO: 15	moltype = DNA	length = 23	
FEATURE	Location/Qualifiers		
misc_feature	1..23		
	note = PosCRPAS6G		
source	1..23		
	mol_type = other DNA		
	organism = synthetic construct		
SEQUENCE: 15			
caccgcctgg ctaaagtagg ctt			23
SEQ ID NO: 16	moltype = DNA	length = 23	
FEATURE	Location/Qualifiers		
misc_feature	1..23		
	note = NegCRPAS6G		
source	1..23		
	mol_type = other DNA		
	organism = synthetic construct		
SEQUENCE: 16			
aaacaagcct actttagcca ggc			23
SEQ ID NO: 17	moltype = DNA	length = 22	
FEATURE	Location/Qualifiers		
misc_feature	1..22		
	note = PosCRi3		
source	1..22		
	mol_type = other DNA		
	organism = synthetic construct		
SEQUENCE: 17			
caccgctttt aggacgcttc gg			22
SEQ ID NO: 18	moltype = DNA	length = 22	
FEATURE	Location/Qualifiers		
misc_feature	1..22		
	note = NegCRi3		
source	1..22		
	mol_type = other DNA		
	organism = synthetic construct		
SEQUENCE: 18			
aaaccgagg cgtcctaata gc			22
SEQ ID NO: 19	moltype = DNA	length = 22	
FEATURE	Location/Qualifiers		
misc_feature	1..22		
	note = PosCRi4		
source	1..22		
	mol_type = other DNA		
	organism = synthetic construct		
SEQUENCE: 19			
caccgcggga cacttcgaga gg			22
SEQ ID NO: 20	moltype = DNA	length = 22	
FEATURE	Location/Qualifiers		
misc_feature	1..22		
	note = NegCRi4		
source	1..22		

-continued

	mol_type = other DNA organism = synthetic construct	
SEQUENCE: 20		
aaaccctctc gaagtgtccc gc		22
SEQ ID NO: 21	moltype = DNA length = 17	
FEATURE	Location/Qualifiers	
misc_feature	1..17	
	note = Fwd1SNP	
source	1..17	
	mol_type = other DNA organism = synthetic construct	
SEQUENCE: 21		
gaccacgcgc attctct		17
SEQ ID NO: 22	moltype = DNA length = 22	
FEATURE	Location/Qualifiers	
misc_feature	1..22	
	note = Fwd4SNP	
source	1..22	
	mol_type = other DNA organism = synthetic construct	
SEQUENCE: 22		
ggaaacagga cagatgaagg ag		22
SEQ ID NO: 23	moltype = DNA length = 22	
FEATURE	Location/Qualifiers	
misc_feature	1..22	
	note = Fwd5SNP	
source	1..22	
	mol_type = other DNA organism = synthetic construct	
SEQUENCE: 23		
cagctcagac ggaagtgtat tt		22
SEQ ID NO: 24	moltype = DNA length = 22	
FEATURE	Location/Qualifiers	
misc_feature	1..22	
	note = Fwd6SNP	
source	1..22	
	mol_type = other DNA organism = synthetic construct	
SEQUENCE: 24		
ctcccaagaa ctgggaacta ac		22
SEQ ID NO: 25	moltype = DNA length = 22	
FEATURE	Location/Qualifiers	
misc_feature	1..22	
	note = Rev3Cleavage	
source	1..22	
	mol_type = other DNA organism = synthetic construct	
SEQUENCE: 25		
accaccgtga tcatgaacta aa		22
SEQ ID NO: 26	moltype = DNA length = 18	
FEATURE	Location/Qualifiers	
misc_feature	1..18	
	note = Rev1SNP	
source	1..18	
	mol_type = other DNA organism = synthetic construct	
SEQUENCE: 26		
cgggactgca tggtaagg		18
SEQ ID NO: 27	moltype = DNA length = 18	
FEATURE	Location/Qualifiers	
misc_feature	1..18	
	note = Sqe1SNP	
source	1..18	
	mol_type = other DNA organism = synthetic construct	
SEQUENCE: 27		
gctgtccggg tgagtatg		18
SEQ ID NO: 28	moltype = DNA length = 18	

-continued

FEATURE	Location/Qualifiers	
misc_feature	1..18	
	note = FwdSNP2	
source	1..18	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 28		
cccacctctc accttcct		18
SEQ ID NO: 29	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = RevSNP2	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 29		
cagcatgatt gacagcccta		20
SEQ ID NO: 30	moltype = DNA length = 17	
FEATURE	Location/Qualifiers	
misc_feature	1..17	
	note = Seq2SNP	
source	1..17	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 30		
ccgcgacact tcacaca		17
SEQ ID NO: 31	moltype = DNA length = 17	
FEATURE	Location/Qualifiers	
misc_feature	1..17	
	note = FwdSNP3	
source	1..17	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 31		
ccgcgacact tcacaca		17
SEQ ID NO: 32	moltype = DNA length = 18	
FEATURE	Location/Qualifiers	
misc_feature	1..18	
	note = RevSNP3	
source	1..18	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 32		
tgctgctgga aggacttg		18
SEQ ID NO: 33	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = Seq3SNP	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 33		
tagggctgtc aatcatgctg		20
SEQ ID NO: 34	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = Rev4SNP	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 34		
gggaattgag ggcggtttat		20
SEQ ID NO: 35	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = Seq4SNP	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	

-continued

SEQUENCE: 35
 tttaccagct cctggctttc 20

SEQ ID NO: 36 moltype = DNA length = 21
 FEATURE Location/Qualifiers
 misc_feature 1..21
 note = Rev5SNP
 source 1..21
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 36
 gagcatgtcc gtgtcctaata c 21

SEQ ID NO: 37 moltype = DNA length = 20
 FEATURE Location/Qualifiers
 misc_feature 1..20
 note = Seq5SNP
 source 1..20
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 37
 tccctggcta gcacttactt 20

SEQ ID NO: 38 moltype = DNA length = 21
 FEATURE Location/Qualifiers
 misc_feature 1..21
 note = Rev6SNP
 source 1..21
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 38
 tgtgattagt gcagcgagaa g 21

SEQ ID NO: 39 moltype = DNA length = 22
 FEATURE Location/Qualifiers
 misc_feature 1..22
 note = Seq6SNP
 source 1..22
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 39
 ctgtttctct gctgtccttc tc 22

SEQ ID NO: 40 moltype = DNA length = 18
 FEATURE Location/Qualifiers
 misc_feature 1..18
 note = Fwd HTT
 source 1..18
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 40
 tcggtgcagc ggctcctc 18

SEQ ID NO: 41 moltype = DNA length = 21
 FEATURE Location/Qualifiers
 misc_feature 1..21
 note = Rev HTT
 source 1..21
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 41
 atggcgaccc tggaaaagct g 21

SEQ ID NO: 42 moltype = DNA length = 17
 FEATURE Location/Qualifiers
 misc_feature 1..17
 note = FwdActB
 source 1..17
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 42
 ttcgcgggcg acgatgc 17

SEQ ID NO: 43 moltype = DNA length = 20
 FEATURE Location/Qualifiers
 misc_feature 1..20

-continued

source	note = RevActB 1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 43		
cgtacatggc tggggtgttg		20
SEQ ID NO: 44	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = CBFA2T3_Fwd	
source	1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 44		
tctgtgggttc agccgacttc		20
SEQ ID NO: 45	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = CBFA2T3_Rev	
source	1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 45		
acacaataacc gtggcagagg		20
SEQ ID NO: 46	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = SLC45A_Fwd	
source	1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 46		
gacccaagct tgccgtagta		20
SEQ ID NO: 47	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = SLC45A_Rev	
source	1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 47		
acctgttcag catcgacgag		20
SEQ ID NO: 48	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = STT3B_Fwd	
source	1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 48		
cctaacggac ctgtcgcttt		20
SEQ ID NO: 49	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = STT3B_Rev	
source	1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 49		
tgagggacga cttgtgcttg		20
SEQ ID NO: 50	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = NAV2_Fwd	
source	1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 50		
ccacgagtgc acacagtttg		20

-continued

SEQ ID NO: 51	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = NAV2_Rev	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 51		
ctcaaggact gctggctcaa		20
SEQ ID NO: 52	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = NUP210_Fwd	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 52		
agctgcgtga tcttgaccaa		20
SEQ ID NO: 53	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = NUP210_Rev	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 53		
ggtgggttcag gctctttcca		20
SEQ ID NO: 54	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = DNAJC16_Fwd	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 54		
tctcatgcac ctctcccat		20
SEQ ID NO: 55	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = DNAJC16_Rev	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 55		
tgagtgcagc gacatgatca		20
SEQ ID NO: 56	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = LEPR_Fwd	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 56		
tgagatgtgc ctcctcaga		20
SEQ ID NO: 57	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = LEPR_Rev	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 57		
aactagtggc atgcgtttgc		20
SEQ ID NO: 58	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = CHRNA2_Fwd	
source	1..20	

-continued

	mol_type = other DNA organism = synthetic construct	
SEQUENCE: 58 ccttctgcat gtggggtgat		20
SEQ ID NO: 59 FEATURE misc_feature source	moltype = DNA length = 20 Location/Qualifiers 1..20 note = CHRNA2_Rev 1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 59 tgagatcatc ccgtccacct		20
SEQ ID NO: 60 FEATURE misc_feature source	moltype = DNA length = 20 Location/Qualifiers 1..20 note = TRIB1_Fwd 1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 60 tcccgggact taaaaagccg		20
SEQ ID NO: 61 FEATURE misc_feature source	moltype = DNA length = 20 Location/Qualifiers 1..20 note = TRIB1_Rev 1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 61 acctggtcaa atggcgtctt		20
SEQ ID NO: 62 FEATURE misc_feature source	moltype = DNA length = 20 Location/Qualifiers 1..20 note = SMARCD1_Fwd 1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 62 tatggttttc cctcccggac		20
SEQ ID NO: 63 FEATURE misc_feature source	moltype = DNA length = 20 Location/Qualifiers 1..20 note = SMARCD1_Rev 1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 63 agcaggtgtg taactgcctc		20
SEQ ID NO: 64 FEATURE misc_feature source	moltype = DNA length = 20 Location/Qualifiers 1..20 note = COX11_Fwd 1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 64 gttagaggct gcggaccttt		20
SEQ ID NO: 65 FEATURE misc_feature source	moltype = DNA length = 20 Location/Qualifiers 1..20 note = COX11_Rev 1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 65 gccgtttctt aggccagagt		20
SEQ ID NO: 66	moltype = DNA length = 49	

-continued

FEATURE	Location/Qualifiers	
misc_feature	1..49	
	note = SNP1 major allele	
source	1..49	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 66		
gtcgccccgc tccaggcgtc ggcgggggat cctttccgca tgggcctgc		49
SEQ ID NO: 67	moltype = DNA length = 49	
FEATURE	Location/Qualifiers	
misc_feature	1..49	
	note = SNP1 minor allele (G)	
source	1..49	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 67		
gtcgccccgc tccaggcgtc ggcgggggcat cctttccgca tgggcctgc		49
SEQ ID NO: 68	moltype = DNA length = 49	
FEATURE	Location/Qualifiers	
misc_feature	1..49	
	note = SNP1 minor allele (T)	
source	1..49	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 68		
gtcgccccgc tccaggcgtc ggcgggggaat cctttccgca tgggcctgc		49
SEQ ID NO: 69	moltype = DNA length = 17	
FEATURE	Location/Qualifiers	
misc_feature	1..17	
	note = sgHD1	
source	1..17	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 69		
gctccaggcg tcggcgg		17
SEQ ID NO: 70	moltype = DNA length = 49	
FEATURE	Location/Qualifiers	
misc_feature	1..49	
	note = rs9996199 major allele	
source	1..49	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 70		
gtggcctggc taaagtaggc ttactgggc tcctctctgc ctgcatcac		49
SEQ ID NO: 71	moltype = DNA length = 49	
FEATURE	Location/Qualifiers	
misc_feature	1..49	
	note = rs9996199 minor allele	
source	1..49	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 71		
gtggcctggc taaagtaggc tttagtgggc tcctctctgc ctgcatcac		49
SEQ ID NO: 72	moltype = DNA length = 17	
FEATURE	Location/Qualifiers	
misc_feature	1..17	
	note = rs9996199 major allele target sequence	
source	1..17	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 72		
gcaggcagag aggagcc		17
SEQ ID NO: 73	moltype = DNA length = 19	
FEATURE	Location/Qualifiers	
misc_feature	1..19	
	note = rs9996199 minor allele target sequence	
source	1..19	
	mol_type = other DNA	
	organism = synthetic construct	

-continued

SEQUENCE: 73
gcctggctaa agtaggctt 19

SEQ ID NO: 74 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
 note = rs35631490
misc_difference 26
 note = n = c or g
source 1..51
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 74
gtctgcgtca gggtttcctt cttttncagc cccaccccgc gtgcatccca c 51

SEQ ID NO: 75 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
 note = rs73086139
misc_difference 26
 note = n = a or g
source 1..51
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 75
tcaaggcctc ttctctcttc tcggcnggac aggcacaggc aggtggccag g 51

SEQ ID NO: 76 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
 note = rs73086140
misc_difference 26
 note = n = c or g
source 1..51
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 76
gccaggtgtc atgcttagct ccccgncag tgagattctt tcatttaaca a 51

SEQ ID NO: 77 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
 note = rs113541600
misc_difference 26
 note = n = t or absent
source 1..51
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 77
taggaacctt atttctctct cgctcntttt tttttttttg agacagagtc t 51

SEQ ID NO: 78 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
 note = rs61791259
misc_difference 26
 note = n = c or t
source 1..51
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 78
gattacaggc acctgccacc atgccnggct aatttttgta ttttagttg a 51

SEQ ID NO: 79 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
 note = rs61792460
misc_difference 26
 note = n = a or g
source 1..51
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 79
gagggtttca tcttggttgt caggcngact tgaactcctg acctcaggtg a 51

SEQ ID NO: 80 moltype = DNA length = 51

-continued

FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = rs73086144	
misc_difference	26	
	note = n = a or g	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 80		
gccactgcgc cttcatctct cttctntgta tgtgtacgct gttttttctt t		51
SEQ ID NO: 81	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = rs73086145	
misc_difference	26	
	note = n = c or g	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 81		
gctgtttttt ctttagaatg ggggangtta tcaggctcta catgggtgtg a		51
SEQ ID NO: 82	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = rs61792461	
misc_difference	26	
	note = n = a or g	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 82		
tctacatggt gtgtagtcgg ctagcntgtt gtaagccttt ccctgtgtca c		51
SEQ ID NO: 83	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = rs61792462	
misc_difference	26	
	note = n = c or g	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 83		
ctgtgtcaca agtgctcatc tggaanagga ttctaatagac tgctgtggc t		51
SEQ ID NO: 84	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = rs61090955	
misc_difference	26	
	note = n = a or g	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 84		
tcattttatg tgattccttt ctagangtac tactcattac ttctgcttgc a		51
SEQ ID NO: 85	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = rs77384845	
misc_difference	26	
	note = n = c or t	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 85		
tagctgaagg aaggacagg actgtnatac actagctaag aggcaaactg c		51
SEQ ID NO: 86	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = rs10011412	
misc_difference	26	

-continued

source	note = n = a or g 1..51 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 86		
agctgaagga aggacaggga ctgtcntaca ctagctaaga ggcaaactgc t		51
SEQ ID NO: 87	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51 note = rs61792464	
misc_difference	26 note = n = c or g	
source	1..51 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 87		
tccttcattc aggttgatgt cctaancccc agaacctcag aatgggattg t		51
SEQ ID NO: 88	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51 note = rs112435590	
misc_difference	26 note = n = g or t	
source	1..51 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 88		
tccatgccaa gaaggcaaca gagagngcca gggagactga agtcataccc t		51
SEQ ID NO: 89	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51 note = rs111670395	
misc_difference	26 note = n = a or g	
source	1..51 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 89		
cccaggttca agcaattctg cctcancctc cggaatagct gggactacag g		51
SEQ ID NO: 90	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51 note = rs762855	
misc_difference	26 note = n = c or t	
source	1..51 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 90		
ttgagaagga cagcagagaa acagcngtta gttcccagtt cttgggaggc t		51
SEQ ID NO: 91	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51 note = rs9996199	
misc_difference	26 note = n = c or g	
source	1..51 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 91		
tgtggcctgg ctaaagtagg ctttantggg ctctctctg cctgcatcac c		51
SEQ ID NO: 92	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51 note = rs143861513	
misc_difference	26 note = n = g or absent	
source	1..51 mol_type = other DNA organism = synthetic construct	

-continued

SEQUENCE: 92
tccggggcgc tgcgctggga ccgatngggg ggcgccaggc ctgtggacac c 51

SEQ ID NO: 93 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
 note = rs28431418
misc_difference 26
 note = n = c or t
source 1..51
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 93
ctccccgcag ggctgtccgg gtgagnatgg ctctggccac gggccagtgt g 51

SEQ ID NO: 94 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
 note = rs2857935
misc_difference 26
 note = n = a, c, or g
source 1..51
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 94
gggcgcaggc ccatgcggaa aggatncccc gccgacgcct ggagcggggc g 51

SEQ ID NO: 95 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
 note = rs28616835
misc_difference 26
 note = n = c or t
source 1..51
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 95
gcgcccgcgc tcggcgcccc ctccanggcc ccgcccgc tc catggccccg t 51

SEQ ID NO: 96 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
 note = rs149624523
misc_difference 26
 note = n = c or t
source 1..51
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 96
ccccgccccg gcctcgccac gcccenacct caccacgccc cccgcatcgc c 51

SEQ ID NO: 97 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
 note = rs13122415
misc_difference 26
 note = n = c or g
source 1..51
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 97
attacagtct caccacgccc cgtcnctct ccgttgagcc ccgcgccttc g 51

SEQ ID NO: 98 moltype = DNA length = 56
FEATURE Location/Qualifiers
misc_feature 1..56
 note = rs113331544
misc_difference 26..31
 note = sequence is gcgggg or absent
source 1..56
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 98
ggccttgctg tgtgaggcag aacctnnnnn ngcgggggca ggggcgggct gggtcc 56

SEQ ID NO: 99 moltype = DNA length = 51

-continued

FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = rs13132932	
misc_difference	26	
	note = n = a or g	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 99		
tgtgtgagggc agaacctgcg ggggcngggg cgggctgggt ccttgccag c		51
SEQ ID NO: 100	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = rs13102260	
misc_difference	26	
	note = n = a or g	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 100		
agcgtctggg acgcaaggcg ccgtgngggc tgccgggacg ggtccaagat g		51
SEQ ID NO: 101	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = rs10009935	
misc_difference	26	
	note = n = c or t	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 101		
ctcacttggg tcttccttg tcctcncgcg aggggaggca ggccttggt g		51
SEQ ID NO: 102	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = rs58870770	
misc_difference	26	
	note = n = a or absent	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 102		
tgaatgagtt gtggttgcca agtaangtgg tgaacttacg tggtgattaa t		51
SEQ ID NO: 103	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = rs34045730	
misc_difference	26	
	note = n = a or t	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 103		
gaggtgtaca tttaccagt attcngtca ggcttgccag aatacggggg g		51
SEQ ID NO: 104	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = rs28656215	
misc_difference	26	
	note = n = c or t	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 104		
ggaagtctgt gtgtcgagtg tacagnagga gttaggaagt actctggtgc a		51
SEQ ID NO: 105	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = rs28571971	
misc_difference	26	

-continued

source	note = n = c or g 1..51 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 105		
acagtaggag ttaggaagta ctctgntgca gttcaggcct ttctcttacc t		51
SEQ ID NO: 106	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51 note = rs28583447	
misc_difference	26 note = n = c or t	
source	1..51 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 106		
cagtaggagt taggaagtac tctggngcag ttcaggcctt tctcttacct c		51
SEQ ID NO: 107	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51 note = rs28468636	
misc_difference	26 note = n = c or g	
source	1..51 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 107		
tctcttacct ctcagtattc tatttncgat ctggatgtgt cccagatggc a		51
SEQ ID NO: 108	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51 note = rs28564368	
misc_difference	26 note = n = a or c	
source	1..51 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 108		
tattctatatt ccgatctgga tgtgtncag atggcatttg gtaagaatat c		51
SEQ ID NO: 109	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51 note = rs28485764	
misc_difference	26 note = n = a or g	
source	1..51 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 109		
gatggcattt ggtaagaata tctctnttaa gactgattaa tttttagtaa t		51
SEQ ID NO: 110	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51 note = rs77173925	
misc_difference	26 note = n = a or g	
source	1..51 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 110		
gaataaatta ttctaaagga tggaanaact ttttggatat ttggagaaat t		51
SEQ ID NO: 111	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51 note = rs3905238	
misc_difference	26 note = n = c or t	
source	1..51 mol_type = other DNA organism = synthetic construct	

-continued

SEQUENCE: 111
ttgtatcatg tcaatgtatt acttangcaa aaataataca ttaaaaaaaaa t 51

SEQ ID NO: 112 moltype = DNA length = 54
FEATURE Location/Qualifiers
misc_feature 1..54
 note = rs33950430
misc_difference 26..29
 note = n = sequence is aaca or absent
source 1..54
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 112
aggatttcac taattttgag taacannnnc tgctcacaaa gtttggattt tggc 54

SEQ ID NO: 113 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
 note = rs28377140
misc_difference 26
 note = n = c or g
source 1..51
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 113
aggcaattaa tacttgcttc tggcantttc ttattctcct tcagattcct a 51

SEQ ID NO: 114 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
 note = rs3856973
misc_difference 26
 note = n = c or t
source 1..51
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 114
ttaaaaataa aaataagtta aactngatt aaccctgaca tttccctatc c 51

SEQ ID NO: 115 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
 note = rs4498089
misc_difference 26
 note = n = a or g
source 1..51
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 115
agtgttaact tatttttatt tttaanaaaa ttgttaaggc ctttccagca a 51

SEQ ID NO: 116 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
 note = rs112353753
misc_difference 26
 note = n = t or absent
source 1..51
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 116
gacatgcact gccatgcctg ggtaantttt tttttttccc ccgagacgga g 51

SEQ ID NO: 117 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
 note = rs57666989
misc_difference 26
 note = n = c or t
source 1..51
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 117
ctgcaagctc cgcttcccga gttcangcca ttctcctgcc tcagttctccc a 51

SEQ ID NO: 118 moltype = DNA length = 51

-continued

FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = rs10006129	
misc_difference	26	
	note = n = c or g	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 118		
tcttgatctc ctgacctcgt c	catcncgga ccttgatgac cgccacctc g	51
SEQ ID NO: 119	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = rs28696693	
misc_difference	26	
	note = n = a or g	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 119		
ggtaattttt gtatttttag	tagagntggg gttttgccat gatgagcagg c	51
SEQ ID NO: 120	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = rs28393280	
misc_difference	26	
	note = n = a or g	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 120		
ggattttgaa tgcggaacca	actgcncctg ttgaactctg ctaagtataa c	51
SEQ ID NO: 121	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 121		
cggtcagag tccacggccg	gctgtcgccc cgctccaggc gtcggcgggg g	51
SEQ ID NO: 122	moltype = DNA length = 55	
FEATURE	Location/Qualifiers	
misc_feature	1..55	
	note = Synthetic sequence	
source	1..55	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 122		
atgcttttag gacgcctcgg	cgggagtggc ggggcggggg ggggcgggga gtgag	55
SEQ ID NO: 123	moltype = DNA length = 82	
FEATURE	Location/Qualifiers	
misc_feature	1..82	
	note = Synthetic sequence	
source	1..82	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 123		
cggtcagag tccacggccg	gctgtcgccc cgctccaggc gtcgcggcgg gagtggcggg	60
gcgggggggg gcggggagtg	ag	82
SEQ ID NO: 124	moltype = DNA length = 82	
FEATURE	Location/Qualifiers	
misc_feature	1..82	
	note = Synthetic sequence	
source	1..82	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 124		
cggtcagag tccacggccg	gctgtcgccc cgccccaggc gtcgcggcgg gagtggcggg	60
gcgggggggg gcggggagtg	ag	82

-continued

SEQ ID NO: 125 moltype = DNA length = 56
FEATURE Location/Qualifiers
misc_feature 1..56
 note = Synthetic sequence
source 1..56
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 125
cggctcagag tccacggccg gcgggagtg cggggcgggg gggggcgggg agtgag 56

SEQ ID NO: 126 moltype = DNA length = 82
FEATURE Location/Qualifiers
misc_feature 1..82
 note = Synthetic sequence
source 1..82
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 126
cggctcagag tccacggccg gctgtcgccc cgctccaggc gtcggggcgg gagtggcggg 60
gcgggggggg gcggggagtg ag 82

SEQ ID NO: 127 moltype = DNA length = 50
FEATURE Location/Qualifiers
misc_feature 1..50
 note = Synthetic sequence
source 1..50
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 127
cccattacag tctcaccacg ccccgctccc tctccgttga gccccgcgcc 50

SEQ ID NO: 128 moltype = DNA length = 55
FEATURE Location/Qualifiers
misc_feature 1..55
 note = Synthetic sequence
source 1..55
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 128
atgcttttag gacgcctcgg cgggagtggc ggggcggggg ggggcggggg gtgag 55

SEQ ID NO: 129 moltype = DNA length = 73
FEATURE Location/Qualifiers
misc_feature 1..73
 note = Synthetic sequence
source 1..73
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 129
cccattacag tctcaccacg ccccgctccc tctcccggcg ggagtgcggg gcgggggggg 60
gcggggagtg agg 73

SEQ ID NO: 130 moltype = DNA length = 74
FEATURE Location/Qualifiers
misc_feature 1..74
 note = Synthetic sequence
source 1..74
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 130
cccattacag tctcaccacg ccccgctccc tctcctcggc gggagtgcgg ggcggggggg 60
ggcggggagtg gagg 74

SEQ ID NO: 131 moltype = DNA length = 73
FEATURE Location/Qualifiers
misc_feature 1..73
 note = Synthetic sequence
source 1..73
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 131
cccattacag tctcaccacg ccccgctccc tctcccggcg ggagtggcgg ggcggggggg 60
gggcggggag tga 73

SEQ ID NO: 132 moltype = DNA length = 73

-continued

FEATURE	Location/Qualifiers	
misc_feature	1..73	
	note = Synthetic sequence	
source	1..73	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 132		
ccattacag tctcaccacg ccccggtcccc tctcgcgggc ggagtggcgg ggcggggggg		60
ggcggggagt gag		73
SEQ ID NO: 133	moltype = DNA length = 49	
FEATURE	Location/Qualifiers	
misc_feature	1..49	
	note = Synthetic sequence	
source	1..49	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 133		
cgcggccccg cctccgccgg cgcagcgtct gggacgcaag gcgccgtg		49
SEQ ID NO: 134	moltype = DNA length = 55	
FEATURE	Location/Qualifiers	
misc_feature	1..55	
	note = Synthetic sequence	
source	1..55	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 134		
atgcttttag gacgcctcgg cgggagtggc gggcgggggg gggcgggga gtgag		55
SEQ ID NO: 135	moltype = DNA length = 71	
FEATURE	Location/Qualifiers	
misc_feature	1..71	
	note = Synthetic sequence	
source	1..71	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 135		
cgcggccccg cctccgccgg cgcagcgtct gggcggcggg agtggcgggg cggggggggg		60
cggggagtga g		71
SEQ ID NO: 136	moltype = DNA length = 82	
FEATURE	Location/Qualifiers	
misc_feature	1..82	
	note = Synthetic sequence	
source	1..82	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 136		
cgcggccccg cctccgccgg cgcagcgtct gggacgcaag gcgtcgccgg gagtggcggg		60
gcgggggggg gcggggagtg ag		82
SEQ ID NO: 137	moltype = DNA length = 83	
FEATURE	Location/Qualifiers	
misc_feature	1..83	
	note = Synthetic sequence	
source	1..83	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 137		
cgcggccccg cctccgccgg cgcagcgtct gggacgcaag gcgatcgccg ggagtggcgg		60
ggcggggggg ggcggggagt gag		83
SEQ ID NO: 138	moltype = DNA length = 81	
FEATURE	Location/Qualifiers	
misc_feature	1..81	
	note = Synthetic sequence	
source	1..81	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 138		
tccggccccg cctccgccgg cgcagcgtct gggacgcaag gcgcggcggg agtggcgggg		60
cggggggggg cggggagtga g		81
SEQ ID NO: 139	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	

-continued

misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 139		
tggggtctgc gtcagggttt ccttcttttc cagccccacc ccgcgtgcat c		51
SEQ ID NO: 140	moltype = DNA length = 55	
FEATURE	Location/Qualifiers	
misc_feature	1..55	
	note = Synthetic sequence	
source	1..55	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 140		
atgcttttag gacgcctcgg cgggagtggc ggggcggggg ggggcgggga gtgag		55
SEQ ID NO: 141	moltype = DNA length = 69	
FEATURE	Location/Qualifiers	
misc_feature	1..69	
	note = Synthetic sequence	
source	1..69	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 141		
tggggtctgc gtcagggttt ccttcttttc cagccgggag tggcggggcg ggggggggcg		60
gggagtgag		69
SEQ ID NO: 142	moltype = DNA length = 73	
FEATURE	Location/Qualifiers	
misc_feature	1..73	
	note = Synthetic sequence	
source	1..73	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 142		
tggggtctgc gtcagggttt ccttcttttc cagccgggag ggagtggcgg ggcggggggg		60
ggcggggagt gag		73
SEQ ID NO: 143	moltype = DNA length = 73	
FEATURE	Location/Qualifiers	
misc_feature	1..73	
	note = Synthetic sequence	
source	1..73	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 143		
tggggtctgc gtcagggttt ccttcttttc cagccgggag ggagtggcgg ggcggggggg		60
ggcggggagt gag		73
SEQ ID NO: 144	moltype = DNA length = 67	
FEATURE	Location/Qualifiers	
misc_feature	1..67	
	note = Synthetic sequence	
source	1..67	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 144		
tggggtctgc gtcagggttt ccttcttttc ggcgggagtg gcggggcggg ggggggaggg		60
gagtgag		67
SEQ ID NO: 145	moltype = DNA length = 52	
FEATURE	Location/Qualifiers	
misc_feature	1..52	
	note = Synthetic sequence	
source	1..52	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 145		
cagggtgtggc ctggctaaag taggctttac tgggctcttc tctgctgca tc		52
SEQ ID NO: 146	moltype = DNA length = 55	
FEATURE	Location/Qualifiers	
misc_feature	1..55	
	note = Synthetic sequence	

-continued

```

source                1..55
                      mol_type = other DNA
                      organism = synthetic construct

SEQUENCE: 146
atgcttttag gacgcctcgg cgggagtggc ggggcggggg ggggcgggga gtgag      55

SEQ ID NO: 147        moltype = DNA   length = 72
FEATURE              Location/Qualifiers
misc_feature          1..72
                      note = Synthetic sequence
source                1..72
                      mol_type = other DNA
                      organism = synthetic construct

SEQUENCE: 147
caggtgtggc ctggctaaag taggctttac tgggcggcgg gagtggcggg gcgggggggg  60
gcggggagtg ag                                           72

SEQ ID NO: 148        moltype = DNA   length = 74
FEATURE              Location/Qualifiers
misc_feature          1..74
                      note = Synthetic sequence
source                1..74
                      mol_type = other DNA
                      organism = synthetic construct

SEQUENCE: 148
caggtgtggc ctggctaaag taggctttac tgggctcggc gggagtggcg gggcgggggg  60
gggcggggag tgag                                           74

SEQ ID NO: 149        moltype = DNA   length = 73
FEATURE              Location/Qualifiers
misc_feature          1..73
                      note = Synthetic sequence
source                1..73
                      mol_type = other DNA
                      organism = synthetic construct

SEQUENCE: 149
caggtgtggc ctggctaaag taggctttac tgggcggcgg ggagtggcgg ggcggggggg  60
ggcggggagt gag                                           73

SEQ ID NO: 150        moltype = DNA   length = 73
FEATURE              Location/Qualifiers
misc_feature          1..73
                      note = Synthetic sequence
source                1..73
                      mol_type = other DNA
                      organism = synthetic construct

SEQUENCE: 150
caggtgtggc ctggctaaag taggctttac tgggcggcgg ggagtggcgg ggcggggggg  60
ggcggggagt gag                                           73

SEQ ID NO: 151        moltype = DNA   length = 49
FEATURE              Location/Qualifiers
misc_feature          1..49
                      note = Synthetic sequence
source                1..49
                      mol_type = other DNA
                      organism = synthetic construct

SEQUENCE: 151
gtgatgcagg cagagaggag cccagtaaag cctacttttag ccaggccac      49

SEQ ID NO: 152        moltype = DNA   length = 49
FEATURE              Location/Qualifiers
misc_feature          1..49
                      note = Synthetic sequence
source                1..49
                      mol_type = other DNA
                      organism = synthetic construct

SEQUENCE: 152
gtgatgcagg cagagaggag cccactaaag cctacttttag ccaggccac      49

SEQ ID NO: 153        moltype = DNA   length = 49
FEATURE              Location/Qualifiers
misc_feature          1..49
                      note = Synthetic sequence
source                1..49
                      mol_type = other DNA

```

-continued

```

                                organism = synthetic construct
SEQUENCE: 153
gcaggcccat gcggaaagga tccccgccg acgcctggag cggggcgac          49

SEQ ID NO: 154                moltype = DNA  length = 49
FEATURE                       Location/Qualifiers
misc_feature                  1..49
                                note = Synthetic sequence
source                        1..49
                                mol_type = other DNA
                                organism = synthetic construct

SEQUENCE: 154
gcaggcccat gcggaaagga tgccccgccg acgcctggag cggggcgac          49

SEQ ID NO: 155                moltype = DNA  length = 49
FEATURE                       Location/Qualifiers
misc_feature                  1..49
                                note = Synthetic sequence
source                        1..49
                                mol_type = other DNA
                                organism = synthetic construct

SEQUENCE: 155
gcaggcccat gcggaaagga ttccccgccg acgcctggag cggggcgac          49

SEQ ID NO: 156                moltype = DNA  length = 20
FEATURE                       Location/Qualifiers
misc_feature                  1..20
                                note = Synthetic sequence
source                        1..20
                                mol_type = other DNA
                                organism = synthetic construct

SEQUENCE: 156
ggatgcacgc ggggtggggc          20

SEQ ID NO: 157                moltype = DNA  length = 20
FEATURE                       Location/Qualifiers
misc_feature                  1..20
                                note = Synthetic sequence
source                        1..20
                                mol_type = other DNA
                                organism = synthetic construct

SEQUENCE: 157
gatgcacgcg ggggtggggc          20

SEQ ID NO: 158                moltype = DNA  length = 20
FEATURE                       Location/Qualifiers
misc_feature                  1..20
                                note = Synthetic sequence
source                        1..20
                                mol_type = other DNA
                                organism = synthetic construct

SEQUENCE: 158
gccccacccc gcgtgcatcc          20

SEQ ID NO: 159                moltype = DNA  length = 19
FEATURE                       Location/Qualifiers
misc_feature                  1..19
                                note = Synthetic sequence
source                        1..19
                                mol_type = other DNA
                                organism = synthetic construct

SEQUENCE: 159
ggcctcttct ctcttctcg          19

SEQ ID NO: 160                moltype = DNA  length = 20
FEATURE                       Location/Qualifiers
misc_feature                  1..20
                                note = Synthetic sequence
source                        1..20
                                mol_type = other DNA
                                organism = synthetic construct

SEQUENCE: 160
ttaaatgaaa gaatctcact          20

SEQ ID NO: 161                moltype = DNA  length = 20
FEATURE                       Location/Qualifiers

```

-continued

misc_feature	1..20	
	note = Synthetic sequence	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 161		
aggtgtcatg cttagctccc		20
SEQ ID NO: 162	moltype = DNA length = 22	
FEATURE	Location/Qualifiers	
misc_feature	1..22	
	note = Synthetic sequence	
source	1..22	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 162		
ttaaatgaaa gaatctcact gg		22
SEQ ID NO: 163	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = Synthetic sequence	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 163		
actaaaaata caaaaattag		20
SEQ ID NO: 164	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = Synthetic sequence	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 164		
ggtttcatct tggtggtcag		20
SEQ ID NO: 165	moltype = DNA length = 21	
FEATURE	Location/Qualifiers	
misc_feature	1..21	
	note = Synthetic sequence	
source	1..21	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 165		
gggtttcatc ttgttggtca g		21
SEQ ID NO: 166	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = Synthetic sequence	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 166		
gttttttctt tagaatgggg		20
SEQ ID NO: 167	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = Synthetic sequence	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 167		
acatggtgtg tagtcggcta		20
SEQ ID NO: 168	moltype = DNA length = 19	
FEATURE	Location/Qualifiers	
misc_feature	1..19	
	note = Synthetic sequence	
source	1..19	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 168		

-continued

gtcacaagtg ctcactctgg	19
SEQ ID NO: 169 moltype = DNA length = 20 FEATURE Location/Qualifiers misc_feature 1..20 note = Synthetic sequence source 1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 169 gtcacaagtg ctcactctgga	20
SEQ ID NO: 170 moltype = DNA length = 23 FEATURE Location/Qualifiers misc_feature 1..23 note = Synthetic sequence source 1..23 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 170 ccagatgagc acttgtaga cag	23
SEQ ID NO: 171 moltype = DNA length = 20 FEATURE Location/Qualifiers misc_feature 1..20 note = Synthetic sequence source 1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 171 ttttatgtga ttcctttcta	20
SEQ ID NO: 172 moltype = DNA length = 21 FEATURE Location/Qualifiers misc_feature 1..21 note = Synthetic sequence source 1..21 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 172 agaaaaggaat cacataaaat g	21
SEQ ID NO: 173 moltype = DNA length = 17 FEATURE Location/Qualifiers misc_feature 1..17 note = Synthetic sequence source 1..17 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 173 gcctcttagc tagtgta	17
SEQ ID NO: 174 moltype = DNA length = 20 FEATURE Location/Qualifiers misc_feature 1..20 note = Synthetic sequence source 1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 174 atcccattct gaggttctgg	20
SEQ ID NO: 175 moltype = DNA length = 19 FEATURE Location/Qualifiers misc_feature 1..19 note = Synthetic sequence source 1..19 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 175 tcattcaggt tgatgtcct	19
SEQ ID NO: 176 moltype = DNA length = 20 FEATURE Location/Qualifiers misc_feature 1..20 note = Synthetic sequence	

-continued

source	1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 176		
ggacatcaac ctgaatgagg		20
SEQ ID NO: 177	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20 note = Synthetic sequence	
source	1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 177		
atgccaagaa ggcaacagag		20
SEQ ID NO: 178	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20 note = Synthetic sequence	
source	1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 178		
tccatgccaa gaaggcaaca		20
SEQ ID NO: 179	moltype = DNA length = 19	
FEATURE	Location/Qualifiers	
misc_feature	1..19 note = Synthetic sequence	
source	1..19 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 179		
ggttcaagca attctgcct		19
SEQ ID NO: 180	moltype = DNA length = 17	
FEATURE	Location/Qualifiers	
misc_feature	1..17 note = Synthetic sequence	
source	1..17 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 180		
gcaggcagag aggagcc		17
SEQ ID NO: 181	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20 note = Synthetic sequence	
source	1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 181		
ggcctggcta aagtaggctt		20
SEQ ID NO: 182	moltype = DNA length = 19	
FEATURE	Location/Qualifiers	
misc_feature	1..19 note = Synthetic sequence	
source	1..19 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 182		
tccccgcagg gctgtccgg		19
SEQ ID NO: 183	moltype = DNA length = 17	
FEATURE	Location/Qualifiers	
misc_feature	1..17 note = Synthetic sequence	
source	1..17 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 183		
gctccaggcg tcggcgg		17

-continued

SEQ ID NO: 184	moltype = DNA length = 19	
FEATURE	Location/Qualifiers	
misc_feature	1..19	
	note = Synthetic sequence	
source	1..19	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 184		
gccccgctcc aggcgtcgg		19
SEQ ID NO: 185	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = Synthetic sequence	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 185		
ccgccgacgc ctggagcggg		20
SEQ ID NO: 186	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = Synthetic sequence	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 186		
gatgcggggg gcgtggtgag		20
SEQ ID NO: 187	moltype = DNA length = 19	
FEATURE	Location/Qualifiers	
misc_feature	1..19	
	note = Synthetic sequence	
source	1..19	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 187		
ggcgcggggc tcaacggag		19
SEQ ID NO: 188	moltype = DNA length = 17	
FEATURE	Location/Qualifiers	
misc_feature	1..17	
	note = Synthetic sequence	
source	1..17	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 188		
gcggggctca acggaga		17
SEQ ID NO: 189	moltype = DNA length = 16	
FEATURE	Location/Qualifiers	
misc_feature	1..16	
	note = Synthetic sequence	
source	1..16	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 189		
gcggggctca acggag		16
SEQ ID NO: 190	moltype = DNA length = 19	
FEATURE	Location/Qualifiers	
misc_feature	1..19	
	note = Synthetic sequence	
source	1..19	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 190		
gctgtgtgag gcagaacct		19
SEQ ID NO: 191	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = Synthetic sequence	
source	1..20	
	mol_type = other DNA	

-continued

SEQUENCE: 191	organism = synthetic construct	
gtgaggcaga acctgcgggg		20
SEQ ID NO: 192	moltype = DNA length = 18	
FEATURE	Location/Qualifiers	
misc_feature	1..18	
	note = Synthetic sequence	
source	1..18	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 192		
gaggcagaac ctgcgggg		18
SEQ ID NO: 193	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = Synthetic sequence	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 193		
gtctgggacg caaggcgccg		20
SEQ ID NO: 194	moltype = DNA length = 22	
FEATURE	Location/Qualifiers	
misc_feature	1..22	
	note = Synthetic sequence	
source	1..22	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 194		
gcgtctggga cgcaaggcgc cg		22
SEQ ID NO: 195	moltype = DNA length = 17	
FEATURE	Location/Qualifiers	
misc_feature	1..17	
	note = Synthetic sequence	
source	1..17	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 195		
ggctctgcct cccctcg		17
SEQ ID NO: 196	moltype = DNA length = 18	
FEATURE	Location/Qualifiers	
misc_feature	1..18	
	note = Synthetic sequence	
source	1..18	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 196		
ggctctgcct cccctcgc		18
SEQ ID NO: 197	moltype = DNA length = 21	
FEATURE	Location/Qualifiers	
misc_feature	1..21	
	note = Synthetic sequence	
source	1..21	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 197		
cttggaacc acaactcatt c		21
SEQ ID NO: 198	moltype = DNA length = 17	
FEATURE	Location/Qualifiers	
misc_feature	1..17	
	note = Synthetic sequence	
source	1..17	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 198		
ggagttaggga agtactc		17
SEQ ID NO: 199	moltype = DNA length = 22	
FEATURE	Location/Qualifiers	

-continued

misc_feature	1..22	
	note = Synthetic sequence	
source	1..22	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 199		
ggtaagagaa aggcctgaac tg		22
SEQ ID NO: 200	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = Synthetic sequence	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 200		
gtaagagaaa ggcctgaact		20
SEQ ID NO: 201	moltype = DNA length = 19	
FEATURE	Location/Qualifiers	
misc_feature	1..19	
	note = Synthetic sequence	
source	1..19	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 201		
ggagttagga agtactctg		19
SEQ ID NO: 202	moltype = DNA length = 19	
FEATURE	Location/Qualifiers	
misc_feature	1..19	
	note = Synthetic sequence	
source	1..19	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 202		
atctgggaca catccagat		19
SEQ ID NO: 203	moltype = DNA length = 19	
FEATURE	Location/Qualifiers	
misc_feature	1..19	
	note = Synthetic sequence	
source	1..19	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 203		
acctctcagt attctattt		19
SEQ ID NO: 204	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = Synthetic sequence	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 204		
atctgggaca catccagatg		20
SEQ ID NO: 205	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = Synthetic sequence	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 205		
attcttacca aatgccatct		20
SEQ ID NO: 206	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = Synthetic sequence	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 206		

-continued

ttctttaccaa atgccatctg	20
SEQ ID NO: 207 moltype = DNA length = 22 FEATURE Location/Qualifiers misc_feature 1..22 note = Synthetic sequence source 1..22 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 207 agatggcatt tggtaagaat at	22
SEQ ID NO: 208 moltype = DNA length = 19 FEATURE Location/Qualifiers misc_feature 1..19 note = Synthetic sequence source 1..19 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 208 aaattattct aaaggatgg	19
SEQ ID NO: 209 moltype = DNA length = 19 FEATURE Location/Qualifiers misc_feature 1..19 note = Synthetic sequence source 1..19 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 209 attattctaa aggatggaa	19
SEQ ID NO: 210 moltype = DNA length = 18 FEATURE Location/Qualifiers misc_feature 1..18 note = Synthetic sequence source 1..18 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 210 aattattcta aaggatgg	18
SEQ ID NO: 211 moltype = DNA length = 20 FEATURE Location/Qualifiers misc_feature 1..20 note = Synthetic sequence source 1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 211 ccatccttta gaataattta	20
SEQ ID NO: 212 moltype = DNA length = 19 FEATURE Location/Qualifiers misc_feature 1..19 note = Synthetic sequence source 1..19 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 212 ttttttaatg tattatttt	19
SEQ ID NO: 213 moltype = DNA length = 20 FEATURE Location/Qualifiers misc_feature 1..20 note = Synthetic sequence source 1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 213 ttactcaaaa ttagtgaata	20
SEQ ID NO: 214 moltype = DNA length = 18 FEATURE Location/Qualifiers misc_feature 1..18 note = Synthetic sequence	

-continued

source	1..18 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 214		
attaatactt gcttctgg		18
SEQ ID NO: 215	moltype = DNA length = 21	
FEATURE	Location/Qualifiers	
misc_feature	1..21 note = Synthetic sequence	
source	1..21 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 215		
ggaatctgaa ggagaataag a		21
SEQ ID NO: 216	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20 note = Synthetic sequence	
source	1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 216		
gggaaatgtc agggttaatc		20
SEQ ID NO: 217	moltype = DNA length = 18	
FEATURE	Location/Qualifiers	
misc_feature	1..18 note = Synthetic sequence	
source	1..18 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 217		
gggaaatgtc agggttaa		18
SEQ ID NO: 218	moltype = DNA length = 21	
FEATURE	Location/Qualifiers	
misc_feature	1..21 note = Synthetic sequence	
source	1..21 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 218		
attaaccctg acatttcct a		21
SEQ ID NO: 219	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20 note = Synthetic sequence	
source	1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 219		
gttaacttat ttttattttt		20
SEQ ID NO: 220	moltype = DNA length = 21	
FEATURE	Location/Qualifiers	
misc_feature	1..21 note = Synthetic sequence	
source	1..21 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 220		
gttaacttat ttttattttt a		21
SEQ ID NO: 221	moltype = DNA length = 21	
FEATURE	Location/Qualifiers	
misc_feature	1..21 note = Synthetic sequence	
source	1..21 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 221		
aaaataaaaa taagttaaca c		21

-continued

SEQ ID NO: 222	moltype = DNA length = 21	
FEATURE	Location/Qualifiers	
misc_feature	1..21	
	note = Synthetic sequence	
source	1..21	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 222		
gggagactga ggcaggagaa t		21
SEQ ID NO: 223	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = Synthetic sequence	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 223		
ggagactgag gcaggagaat		20
SEQ ID NO: 224	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = Synthetic sequence	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 224		
ggtgggcgga tcacaaggtc		20
SEQ ID NO: 225	moltype = DNA length = 19	
FEATURE	Location/Qualifiers	
misc_feature	1..19	
	note = Synthetic sequence	
source	1..19	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 225		
atTTTTgtat ttttagtag		19
SEQ ID NO: 226	moltype = DNA length = 18	
FEATURE	Location/Qualifiers	
misc_feature	1..18	
	note = Synthetic sequence	
source	1..18	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 226		
tttttgtatt ttttagtag		18
SEQ ID NO: 227	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = Synthetic sequence	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 227		
atTTTgaatg cggaaccaac		20
SEQ ID NO: 228	moltype = DNA length = 19	
FEATURE	Location/Qualifiers	
misc_feature	1..19	
	note = Synthetic sequence	
source	1..19	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 228		
tacttagcag agttcaaca		19
SEQ ID NO: 229	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	

-continued

organism = synthetic construct

SEQUENCE: 229
gtctgcgtca gggtttcctt cttttccagc cccaccccg cgtgcatccca c 51

SEQ ID NO: 230 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
note = Synthetic sequence
source 1..51
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 230
gtgggatgca cgcggggtgg ggctggaaaa gaaggaaacc ctgacgcaga c 51

SEQ ID NO: 231 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
note = Synthetic sequence
source 1..51
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 231
gtctgcgtca gggtttcctt cttttgcagc cccaccccg cgtgcatccca c 51

SEQ ID NO: 232 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
note = Synthetic sequence
source 1..51
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 232
gtgggatgca cgcggggtgg ggctgcaaaa gaaggaaacc ctgacgcaga c 51

SEQ ID NO: 233 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
note = Synthetic sequence
source 1..51
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 233
tcaaggcctc ttctctcttc tcggcaggac aggcacaggc aggtggccag g 51

SEQ ID NO: 234 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
note = Synthetic sequence
source 1..51
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 234
cctggccacc tgctgtgcc tgcctgccg agaagagaga agaggccttg a 51

SEQ ID NO: 235 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
note = Synthetic sequence
source 1..51
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 235
tcaaggcctc ttctctcttc tcggcgggac aggcacaggc aggtggccag g 51

SEQ ID NO: 236 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
note = Synthetic sequence
source 1..51
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 236
cctggccacc tgctgtgcc tgcctgccg agaagagaga agaggccttg a 51

SEQ ID NO: 237 moltype = DNA length = 51
FEATURE Location/Qualifiers

-continued

misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 237		
gccaggtgtc atgcttagct ccccgcccag tgagattctt tcatttaaca a		51
SEQ ID NO: 238	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 238		
ttgttaaagt aaagaatctc actgggcccgg gagctaagca tgacacctgg c		51
SEQ ID NO: 239	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 239		
gccaggtgtc atgcttagct ccccgcccag tgagattctt tcatttaaca a		51
SEQ ID NO: 240	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 240		
ttgttaaagt aaagaatctc actggcccgg gagctaagca tgacacctgg c		51
SEQ ID NO: 241	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 241		
gattacaggc acctgccacc atgcccggct aatttttgta ttttagttg a		51
SEQ ID NO: 242	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 242		
tcaactaaaa atacaaaaat tagccgggca tgggtggcagg tgcctgtaat c		51
SEQ ID NO: 243	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 243		
gattacaggc acctgccacc atgcctggct aatttttgta ttttagttg a		51
SEQ ID NO: 244	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 244		

-continued

tcaactaaaa atacaaaaat tagccaggca tgggtggcagg tgccctgtaat c	51
SEQ ID NO: 245	moltype = DNA length = 51
FEATURE	Location/Qualifiers
misc_feature	1..51
	note = Synthetic sequence
source	1..51
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 245	
gagggtttca tcttgttggt caggcggact tgaactcctg acctcagggtg a	51
SEQ ID NO: 246	moltype = DNA length = 51
FEATURE	Location/Qualifiers
misc_feature	1..51
	note = Synthetic sequence
source	1..51
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 246	
tcacctgagg tcaggagttc aagtccgcct gaccaacaag atgaaaccct c	51
SEQ ID NO: 247	moltype = DNA length = 51
FEATURE	Location/Qualifiers
misc_feature	1..51
	note = Synthetic sequence
source	1..51
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 247	
gagggtttca tcttgttggt caggcagact tgaactcctg acctcagggtg a	51
SEQ ID NO: 248	moltype = DNA length = 51
FEATURE	Location/Qualifiers
misc_feature	1..51
	note = Synthetic sequence
source	1..51
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 248	
tcacctgagg tcaggagttc aagtctgcct gaccaacaag atgaaaccct c	51
SEQ ID NO: 249	moltype = DNA length = 51
FEATURE	Location/Qualifiers
misc_feature	1..51
	note = Synthetic sequence
source	1..51
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 249	
gctgtttttt cttagaatg ggggacgtta tcaggctcta catggtgtgt a	51
SEQ ID NO: 250	moltype = DNA length = 51
FEATURE	Location/Qualifiers
misc_feature	1..51
	note = Synthetic sequence
source	1..51
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 250	
tacacacat gtagagcctg ataacgtccc ccattctaaa gaaaaaacag c	51
SEQ ID NO: 251	moltype = DNA length = 51
FEATURE	Location/Qualifiers
misc_feature	1..51
	note = Synthetic sequence
source	1..51
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 251	
gctgtttttt cttagaatg ggggagggtta tcaggctcta catggtgtgt a	51
SEQ ID NO: 252	moltype = DNA length = 51
FEATURE	Location/Qualifiers
misc_feature	1..51
	note = Synthetic sequence

-continued

source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 252		
tacacacccat gtagagcctg ataacctccc ccattctaaa gaaaaaacag c		51
SEQ ID NO: 253	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 253		
tctacatggt gtgtagtcgg ctagcatggt gtaagccttt ccctgtgtca c		51
SEQ ID NO: 254	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 254		
gtgacacagg gaaaggctta caacatgcta gccgactaca caccatgtag a		51
SEQ ID NO: 255	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 255		
tctacatggt gtgtagtcgg ctagcgtggt gtaagccttt ccctgtgtca c		51
SEQ ID NO: 256	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 256		
gtgacacagg gaaaggctta caacacgcta gccgactaca caccatgtag a		51
SEQ ID NO: 257	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 257		
ctgtgtcaca agtgctcatc tggaacagga ttctaatacgc tgctgtggc t		51
SEQ ID NO: 258	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 258		
agccacaggc agtcattaga atcctgttcc agatgagcac ttgtgacaca g		51
SEQ ID NO: 259	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 259		
ctgtgtcaca agtgctcatc tggaagagga ttctaatacgc tgctgtggc t		51

-continued

SEQ ID NO: 260	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51		
	note = Synthetic sequence		
source	1..51		
	mol_type = other DNA		
	organism = synthetic construct		
SEQUENCE: 260			
agccacaggc agtcattaga atcctcttcc agatgagcac ttgtgacaca g			51
SEQ ID NO: 261	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51		
	note = Synthetic sequence		
source	1..51		
	mol_type = other DNA		
	organism = synthetic construct		
SEQUENCE: 261			
tcattttatg tgattccttt ctagaagtac tactcattac ttctgcttgc a			51
SEQ ID NO: 262	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51		
	note = Synthetic sequence		
source	1..51		
	mol_type = other DNA		
	organism = synthetic construct		
SEQUENCE: 262			
tgcaagcaga agtaatgagt agtacttcta gaaaggaatc acataaaatg a			51
SEQ ID NO: 263	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51		
	note = Synthetic sequence		
source	1..51		
	mol_type = other DNA		
	organism = synthetic construct		
SEQUENCE: 263			
tcattttatg tgattccttt ctagaggtag tactcattac ttctgcttgc a			51
SEQ ID NO: 264	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51		
	note = Synthetic sequence		
source	1..51		
	mol_type = other DNA		
	organism = synthetic construct		
SEQUENCE: 264			
tgcaagcaga agtaatgagt agtacctcta gaaaggaatc acataaaatg a			51
SEQ ID NO: 265	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51		
	note = Synthetic sequence		
source	1..51		
	mol_type = other DNA		
	organism = synthetic construct		
SEQUENCE: 265			
tagctgaagg aaggacaggg actgtcatatc actagctaag aggcaaactg c			51
SEQ ID NO: 266	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51		
	note = Synthetic sequence		
source	1..51		
	mol_type = other DNA		
	organism = synthetic construct		
SEQUENCE: 266			
gcagtttgcc tcttagctag tgtatgacag tcctgtcct tccttcagct a			51
SEQ ID NO: 267	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51		
	note = Synthetic sequence		
source	1..51		
	mol_type = other DNA		

-continued

	organism = synthetic construct	
SEQUENCE: 267		
tagctgaagg aaggacaggg actgttatac actagctaag aggcaaactg c		51
SEQ ID NO: 268	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 268		
gcagtttgcc tcttagctag tgtataacag tcctgtcct tccttcagct a		51
SEQ ID NO: 269	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 269		
tcctcattc aggttgatgt cctaaccccc agaacctcag aatgggattg t		51
SEQ ID NO: 270	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 270		
acaatcccat tctgaggttc tgggggcttag gacatcaacc tgaatgaggg a		51
SEQ ID NO: 271	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 271		
tcctcattc aggttgatgt cctaagcccc agaacctcag aatgggattg t		51
SEQ ID NO: 272	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 272		
acaatcccat tctgaggttc tggggcttag gacatcaacc tgaatgaggg a		51
SEQ ID NO: 273	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 273		
tccatgcaa gaaggcaaca gagaggcca gggagactga agtcataccc t		51
SEQ ID NO: 274	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 274		
agggtatgac ttcagtcctc ctggccctct ctgttgctt cttggcatgg a		51
SEQ ID NO: 275	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	

-continued

misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 275		
tccatgccaa gaaggcaaca gagagtgcca gggagactga agtcataccc t		51
SEQ ID NO: 276	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 276		
agggtatgac ttcagtctcc ctggcactct ctgttgccctt cttggcatgg a		51
SEQ ID NO: 277	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 277		
cccaggttca agcaattctg cctcagcctc cggaatagct gggactacag g		51
SEQ ID NO: 278	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 278		
cctgtagtcc cagctattcc ggaggctgag gcagaattgc ttgaacctgg g		51
SEQ ID NO: 279	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 279		
cccaggttca agcaattctg cctcaacctc cggaatagct gggactacag g		51
SEQ ID NO: 280	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 280		
cctgtagtcc cagctattcc ggagggtgag gcagaattgc ttgaacctgg g		51
SEQ ID NO: 281	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 281		
tgtggcctgg ctaaagtagg ctttactggg ctctctctg cctgcatcac c		51
SEQ ID NO: 282	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 282		

-continued

ggtgatgcag gcagagagga gcccagtaaa gcctacttta gccaggccac a 51

SEQ ID NO: 283 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
note = Synthetic sequence
source 1..51
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 283
tgtggcctgg ctaaagtagg ctttagtggg ctctctctg cctgcatcac c 51

SEQ ID NO: 284 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
note = Synthetic sequence
source 1..51
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 284
ggtgatgcag gcagagagga gcccactaaa gcctacttta gccaggccac a 51

SEQ ID NO: 285 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
note = Synthetic sequence
source 1..51
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 285
ctccccgcag ggctgtccgg gtgagtatgg ctctggccac gggccagtgt g 51

SEQ ID NO: 286 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
note = Synthetic sequence
source 1..51
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 286
cacactggcc cgtggccaga gccataactca cccggacagc cctgcgggga g 51

SEQ ID NO: 287 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
note = Synthetic sequence
source 1..51
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 287
ctccccgcag ggctgtccgg gtgagcatgg ctctggccac gggccagtgt g 51

SEQ ID NO: 288 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
note = Synthetic sequence
source 1..51
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 288
cacactggcc cgtggccaga gccatgctca cccggacagc cctgcgggga g 51

SEQ ID NO: 289 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
note = Synthetic sequence
source 1..51
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 289
gggcgcaggc ccatgcggaa aggatcccc gccgacgcct ggagcggggc g 51

SEQ ID NO: 290 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
note = Synthetic sequence

-continued

source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 290		
cgccccgctc caggcgtcgg cgggggatacc tttccgcatg ggctcgccc c		51
SEQ ID NO: 291	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 291		
gggcgcaggc ccatgcggaa aggatgcccc gccgacgcct ggagcggggc g		51
SEQ ID NO: 292	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 292		
cgccccgctc caggcgtcgg cggggcatcc tttccgcatg ggctcgccc c		51
SEQ ID NO: 293	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 293		
gggcgcaggc ccatgcggaa aggatacccc gccgacgcct ggagcggggc g		51
SEQ ID NO: 294	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 294		
cgccccgctc caggcgtcgg cggggatatcc tttccgcatg ggctcgccc c		51
SEQ ID NO: 295	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 295		
ccccgccccg gcctcgccac gcccctacct caccacgccc cccgcatcgc c		51
SEQ ID NO: 296	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 296		
ggcgatgcgg ggggcgtggt gaggtagggg cgtggcgagg cgggggcggg g		51
SEQ ID NO: 297	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 297		
ccccgccccg gcctcgccac gccccacct caccacgccc cccgcatcgc c		51

-continued

SEQ ID NO: 298	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51	note = Synthetic sequence	
source	1..51	mol_type = other DNA	
		organism = synthetic construct	
SEQUENCE: 298			
ggcgatgcgg	ggggcggtgt	gaggtggggg	cgtggcgagg ccggggcggg g 51
SEQ ID NO: 299	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51	note = Synthetic sequence	
source	1..51	mol_type = other DNA	
		organism = synthetic construct	
SEQUENCE: 299			
attacagtct	caccacgccc	cgtccctct	ccgttgagcc ccgcgccttc g 51
SEQ ID NO: 300	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51	note = Synthetic sequence	
source	1..51	mol_type = other DNA	
		organism = synthetic construct	
SEQUENCE: 300			
cgaaggcgcg	gggctcaacg	gagaggggac	ggggcggtgt gagactgtaa t 51
SEQ ID NO: 301	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51	note = Synthetic sequence	
source	1..51	mol_type = other DNA	
		organism = synthetic construct	
SEQUENCE: 301			
attacagtct	caccacgccc	cgtccgctct	ccgttgagcc ccgcgccttc g 51
SEQ ID NO: 302	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51	note = Synthetic sequence	
source	1..51	mol_type = other DNA	
		organism = synthetic construct	
SEQUENCE: 302			
cgaaggcgcg	gggctcaacg	gagagcggac	ggggcggtgt gagactgtaa t 51
SEQ ID NO: 303	moltype = DNA	length = 50	
FEATURE	Location/Qualifiers		
misc_feature	1..50	note = Synthetic sequence	
source	1..50	mol_type = other DNA	
		organism = synthetic construct	
SEQUENCE: 303			
ggccttgctg	tgtgaggcag	aacctgcggg	ggcaggggcg ggctgggtcc 50
SEQ ID NO: 304	moltype = DNA	length = 50	
FEATURE	Location/Qualifiers		
misc_feature	1..50	note = Synthetic sequence	
source	1..50	mol_type = other DNA	
		organism = synthetic construct	
SEQUENCE: 304			
ggaaccagcc	cgccctgcc	cccgcaggtt	ctgcctcaca cagcaaggcc 50
SEQ ID NO: 305	moltype = DNA	length = 56	
FEATURE	Location/Qualifiers		
misc_feature	1..56	note = Synthetic sequence	
source	1..56	mol_type = other DNA	

-continued

	organism = synthetic construct	
SEQUENCE: 305		
ggccttgctg tgtgaggcag aacctgcggg ggcgggggca ggggcgggct ggttcc		56
SEQ ID NO: 306	moltype = DNA length = 56	
FEATURE	Location/Qualifiers	
misc_feature	1..56	
	note = Synthetic sequence	
source	1..56	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 306		
ggaaccagcc cgccctgcc cccgcccccg caggttctgc ctcacacagc aaggcc		56
SEQ ID NO: 307	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 307		
tgtgtgaggc agaacctgcg ggggcagggg cgggctgggt ccctggccag c		51
SEQ ID NO: 308	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 308		
gctggccagg gaaccagccc gcccctgccc ccgcagggtc tgctcacac a		51
SEQ ID NO: 309	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 309		
tgtgtgaggc agaacctgcg ggggcggggg cgggctgggt ccctggccag c		51
SEQ ID NO: 310	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 310		
gctggccagg gaaccagccc gccccgccc ccgcagggtc tgctcacac a		51
SEQ ID NO: 311	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 311		
agcgtctggg acgcaaggcg ccgtgggggc tgccgggacg ggtccaagat g		51
SEQ ID NO: 312	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 312		
catcttgac ccgtcccgcc agccccacg gcgccttgcg tcccagacgc t		51
SEQ ID NO: 313	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	

-continued

misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 313		
agcgtctggg acgcaaggcg ccgtgagggc tgccgggacg ggtccaagat g		51
SEQ ID NO: 314	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 314		
catcttggac ccgtcccggc agccctcacg gcgccttgcg tcccagacgc t		51
SEQ ID NO: 315	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 315		
ctcacttggg tcttcccttg tcctctcgcg aggggaggca gagecctgtt g		51
SEQ ID NO: 316	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 316		
caacaaggct ctgcctcccc tcgcgagagg acaagggaag acccaagtga g		51
SEQ ID NO: 317	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 317		
ctcacttggg tcttcccttg tcctcccgcg aggggaggca gagecctgtt g		51
SEQ ID NO: 318	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 318		
caacaaggct ctgcctcccc tcgcgaggagg acaagggaag acccaagtga g		51
SEQ ID NO: 319	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 319		
tgaatgagtt gtggttgcca agtaaagtgg tgaacttacg tggtgattaa t		51
SEQ ID NO: 320	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 320		

-continued

attaatcacc acgtaagtgc accactttac ttggcaacca caactcattc a	51
SEQ ID NO: 321 moltype = DNA length = 50	
FEATURE Location/Qualifiers	
misc_feature	1..50
note = Synthetic sequence	
source	1..50
mol_type = other DNA	
organism = synthetic construct	
SEQUENCE: 321	
tgaatgagtt gtggttgcca agtaagtggg gaacttacgt ggtgattaat	50
SEQ ID NO: 322 moltype = DNA length = 50	
FEATURE Location/Qualifiers	
misc_feature	1..50
note = Synthetic sequence	
source	1..50
mol_type = other DNA	
organism = synthetic construct	
SEQUENCE: 322	
attaatcacc acgtaagtgc accacttact tggcaaccac aactcattca	50
SEQ ID NO: 323 moltype = DNA length = 51	
FEATURE Location/Qualifiers	
misc_feature	1..51
note = Synthetic sequence	
source	1..51
mol_type = other DNA	
organism = synthetic construct	
SEQUENCE: 323	
acagtaggag ttaggaagta ctctgggtgca gttcaggcct ttctcttacc t	51
SEQ ID NO: 324 moltype = DNA length = 51	
FEATURE Location/Qualifiers	
misc_feature	1..51
note = Synthetic sequence	
source	1..51
mol_type = other DNA	
organism = synthetic construct	
SEQUENCE: 324	
aggtaagaga aaggcctgaa ctgcaccaga gtacttccta actcctactg t	51
SEQ ID NO: 325 moltype = DNA length = 51	
FEATURE Location/Qualifiers	
misc_feature	1..51
note = Synthetic sequence	
source	1..51
mol_type = other DNA	
organism = synthetic construct	
SEQUENCE: 325	
acagtaggag ttaggaagta ctctgctgca gttcaggcct ttctcttacc t	51
SEQ ID NO: 326 moltype = DNA length = 51	
FEATURE Location/Qualifiers	
misc_feature	1..51
note = Synthetic sequence	
source	1..51
mol_type = other DNA	
organism = synthetic construct	
SEQUENCE: 326	
aggtaagaga aaggcctgaa ctgcagcaga gtacttccta actcctactg t	51
SEQ ID NO: 327 moltype = DNA length = 51	
FEATURE Location/Qualifiers	
misc_feature	1..51
note = Synthetic sequence	
source	1..51
mol_type = other DNA	
organism = synthetic construct	
SEQUENCE: 327	
cagtaggagt taggaagtac tctgggtgcag ttcaggcctt tctcttacct c	51
SEQ ID NO: 328 moltype = DNA length = 51	
FEATURE Location/Qualifiers	
misc_feature	1..51
note = Synthetic sequence	

-continued

source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 328		
gaggtaagag aaaggcctga actgcaccag agtacttcct aactcctact g		51
SEQ ID NO: 329	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 329		
cagtaggagt taggaagtac tctggcgcag ttcaggcctt tctcttacct c		51
SEQ ID NO: 330	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 330		
gaggtaagag aaaggcctga actgcgccag agtacttcct aactcctact g		51
SEQ ID NO: 331	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 331		
tctcttacct ctcagtattc tatttccgat ctggatgtgt cccagatggc a		51
SEQ ID NO: 332	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 332		
tgccatctgg gacacatcca gatcggaat agaatactga gaggtaagag a		51
SEQ ID NO: 333	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 333		
tctcttacct ctcagtattc tatttgcgat ctggatgtgt cccagatggc a		51
SEQ ID NO: 334	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 334		
tgccatctgg gacacatcca gatcgcaaat agaatactga gaggtaagag a		51
SEQ ID NO: 335	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 335		
tattctatctt ccgatctgga tgtgtcccag atggcatttg gtaagaatat c		51

-continued

SEQ ID NO: 336	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51	note = Synthetic sequence	
source	1..51	mol_type = other DNA	
		organism = synthetic construct	
SEQUENCE: 336			
gatattctta ccaaagcca tctgggacac atccagatcg gaaatagaat a			51
SEQ ID NO: 337	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51	note = Synthetic sequence	
source	1..51	mol_type = other DNA	
		organism = synthetic construct	
SEQUENCE: 337			
tattctatctt ccgatctgga tgtgtaccag atggcatttg gtaagaatat c			51
SEQ ID NO: 338	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51	note = Synthetic sequence	
source	1..51	mol_type = other DNA	
		organism = synthetic construct	
SEQUENCE: 338			
gatattctta ccaaagcca tctgggtacac atccagatcg gaaatagaat a			51
SEQ ID NO: 339	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51	note = Synthetic sequence	
source	1..51	mol_type = other DNA	
		organism = synthetic construct	
SEQUENCE: 339			
gaataaatta ttctaaagga tggaaaaact ttttgatata ttggagaaat t			51
SEQ ID NO: 340	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51	note = Synthetic sequence	
source	1..51	mol_type = other DNA	
		organism = synthetic construct	
SEQUENCE: 340			
aatttctcca aatatccaaa aagtttttcc atcctttaga ataatttatt c			51
SEQ ID NO: 341	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51	note = Synthetic sequence	
source	1..51	mol_type = other DNA	
		organism = synthetic construct	
SEQUENCE: 341			
gaataaatta ttctaaagga tggagaact ttttgatata ttggagaaat t			51
SEQ ID NO: 342	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51	note = Synthetic sequence	
source	1..51	mol_type = other DNA	
		organism = synthetic construct	
SEQUENCE: 342			
aatttctcca aatatccaaa aagttcttcc atcctttaga ataatttatt c			51
SEQ ID NO: 343	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51	note = Synthetic sequence	
source	1..51	mol_type = other DNA	

-continued

SEQUENCE: 343	organism = synthetic construct	
ttgtatcatg tcaatgtatt acttatgcaa aaataataca ttaaaaaaaaa t		51
SEQ ID NO: 344	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 344		
atTTTTTTta atgtattatt ttTgcataag taatacattg acatgataca a		51
SEQ ID NO: 345	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 345		
ttgtatcatg tcaatgtatt acttacgcaa aaataataca ttaaaaaaaaa t		51
SEQ ID NO: 346	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 346		
atTTTTTTta atgtattatt ttTgcgtaag taatacattg acatgataca a		51
SEQ ID NO: 347	moltype = DNA length = 54	
FEATURE	Location/Qualifiers	
misc_feature	1..54	
	note = Synthetic sequence	
source	1..54	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 347		
aggtattcac taattttgag taacaaacac tgctcacaaa gtttgattt tggc		54
SEQ ID NO: 348	moltype = DNA length = 54	
FEATURE	Location/Qualifiers	
misc_feature	1..54	
	note = Synthetic sequence	
source	1..54	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 348		
gccaaaatcc aaactttgtg agcagtgtt gttactcaaa attagtgaat acct		54
SEQ ID NO: 349	moltype = DNA length = 50	
FEATURE	Location/Qualifiers	
misc_feature	1..50	
	note = Synthetic sequence	
source	1..50	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 349		
aggtattcac taattttgag taacactgct cacaaagttt ggattttggc		50
SEQ ID NO: 350	moltype = DNA length = 50	
FEATURE	Location/Qualifiers	
misc_feature	1..50	
	note = Synthetic sequence	
source	1..50	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 350		
gccaaaatcc aaactttgtg agcagtgtta ctcaaaatta gtgaatacct		50
SEQ ID NO: 351	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	

-continued

misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 351		
aggcaattaa tacttgcttc tggcagtttc ttattctcct tcagattcct a		51
SEQ ID NO: 352	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 352		
taggaatctg aaggagaata agaaactgcc agaagcaagt attaattgcc t		51
SEQ ID NO: 353	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 353		
aggcaattaa tacttgcttc tggcactttc ttattctcct tcagattcct a		51
SEQ ID NO: 354	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 354		
taggaatctg aaggagaata agaaagtgcc agaagcaagt attaattgcc t		51
SEQ ID NO: 355	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 355		
ttaaaaataa aaataagtta acactcgatt aaccctgaca tttccctatc c		51
SEQ ID NO: 356	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 356		
ggatagggaa atgtcagggt taatcgagtg ttaacttatt tttattttta a		51
SEQ ID NO: 357	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 357		
ttaaaaataa aaataagtta acacttgatt aaccctgaca tttccctatc c		51
SEQ ID NO: 358	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 358		

-continued

ggataggggaa atgtcagggt taatcaagtg ttaacttatt tttattttta a 51

SEQ ID NO: 359 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
note = Synthetic sequence
source 1..51
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 359
agtgttaact tatttttatt tttaaaaaaa ttgttaaggg ctttccagca a 51

SEQ ID NO: 360 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
note = Synthetic sequence
source 1..51
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 360
ttgctggaaa gcccttaaca atttttttaa aaataaaaat aagttaacac t 51

SEQ ID NO: 361 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
note = Synthetic sequence
source 1..51
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 361
agtgttaact tatttttatt tttaagaaaa ttgttaaggg ctttccagca a 51

SEQ ID NO: 362 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
note = Synthetic sequence
source 1..51
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 362
ttgctggaaa gcccttaaca attttcttaa aaataaaaat aagttaacac t 51

SEQ ID NO: 363 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
note = Synthetic sequence
source 1..51
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 363
ctgcaagctc cgcttcccga gttcacgcca ttctcctgcc tcagtcctcc a 51

SEQ ID NO: 364 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
note = Synthetic sequence
source 1..51
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 364
tgggagactg aggcaggaga atggcgtgaa ctcgggaagc ggagcttgca g 51

SEQ ID NO: 365 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
note = Synthetic sequence
source 1..51
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 365
ctgcaagctc cgcttcccga gttcatgcca ttctcctgcc tcagtcctcc a 51

SEQ ID NO: 366 moltype = DNA length = 51
FEATURE Location/Qualifiers
misc_feature 1..51
note = Synthetic sequence

-continued

source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 366		
tgggagactg aggcaggaga atggcatgaa ctcgggaagc ggagcttgca g		51
SEQ ID NO: 367	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 367		
tcttgatctc ctgacctcgt catccgccga ccttgatgac cgcccacctc g		51
SEQ ID NO: 368	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 368		
cgaggtgggc ggatcacaag gtcggcggat gacgaggtca ggagatcaag a		51
SEQ ID NO: 369	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 369		
tcttgatctc ctgacctcgt catccccga ccttgatgac cgcccacctc g		51
SEQ ID NO: 370	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 370		
cgaggtgggc ggatcacaag gtcgggggat gacgaggtca ggagatcaag a		51
SEQ ID NO: 371	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 371		
ggtaattttt gtatttttag tagagatggg gttttgccat gatgagcagg c		51
SEQ ID NO: 372	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 372		
gcctgctcat catggcaaaa ccccatctct actaaaaata caaaaattac c		51
SEQ ID NO: 373	moltype = DNA length = 51	
FEATURE	Location/Qualifiers	
misc_feature	1..51	
	note = Synthetic sequence	
source	1..51	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 373		
ggtaattttt gtatttttag tagaggtggg gttttgccat gatgagcagg c		51

-continued

SEQ ID NO: 374	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51		
	note = Synthetic sequence		
source	1..51		
	mol_type = other DNA		
	organism = synthetic construct		
SEQUENCE: 374			
gcctgctcat catggcaaaa cccacacctct actaaaaata caaaaattac c			51
SEQ ID NO: 375	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51		
	note = Synthetic sequence		
source	1..51		
	mol_type = other DNA		
	organism = synthetic construct		
SEQUENCE: 375			
ggattttgaa tgcggaacca actgcacttg ttgaactctg ctaagtataa c			51
SEQ ID NO: 376	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51		
	note = Synthetic sequence		
source	1..51		
	mol_type = other DNA		
	organism = synthetic construct		
SEQUENCE: 376			
gttatactta gcagagttca acaagtgacag ttggttccgc attcaaaatc c			51
SEQ ID NO: 377	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51		
	note = Synthetic sequence		
source	1..51		
	mol_type = other DNA		
	organism = synthetic construct		
SEQUENCE: 377			
ggattttgaa tgcggaacca actgcgcttg ttgaactctg ctaagtataa c			51
SEQ ID NO: 378	moltype = DNA	length = 51	
FEATURE	Location/Qualifiers		
misc_feature	1..51		
	note = Synthetic sequence		
source	1..51		
	mol_type = other DNA		
	organism = synthetic construct		
SEQUENCE: 378			
gttatactta gcagagttca acaagcgacag ttggttccgc attcaaaatc c			51
SEQ ID NO: 379	moltype = DNA	length = 17	
FEATURE	Location/Qualifiers		
misc_feature	1..17		
	note = Synthetic sequence		
source	1..17		
	mol_type = other DNA		
	organism = synthetic construct		
SEQUENCE: 379			
gctccaggcg tcggcgg			17
SEQ ID NO: 380	moltype = DNA	length = 19	
FEATURE	Location/Qualifiers		
misc_feature	1..19		
	note = Synthetic sequence		
source	1..19		
	mol_type = other DNA		
	organism = synthetic construct		
SEQUENCE: 380			
ggcgcggggc tcaacggag			19
SEQ ID NO: 381	moltype = DNA	length = 20	
FEATURE	Location/Qualifiers		
misc_feature	1..20		
	note = Synthetic sequence		
source	1..20		
	mol_type = other DNA		

-continued

SEQUENCE: 381	organism = synthetic construct	
gtctgggacg caaggcgccg		20
SEQ ID NO: 382	moltype = DNA length = 19	
FEATURE	Location/Qualifiers	
misc_feature	1..19	
	note = Synthetic sequence	
source	1..19	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 382		
gatgcacgcg ggggtggggc		19
SEQ ID NO: 383	moltype = DNA length = 17	
FEATURE	Location/Qualifiers	
misc_feature	1..17	
	note = Synthetic sequence	
source	1..17	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 383		
attcaggttg atgtcct		17
SEQ ID NO: 384	moltype = DNA length = 20	
FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = Synthetic sequence	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 384		
atcccattct gaggttctgg		20
SEQ ID NO: 385	moltype = DNA length = 17	
FEATURE	Location/Qualifiers	
misc_feature	1..17	
	note = Synthetic sequence	
source	1..17	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 385		
gcaggcagag aggagcc		17
SEQ ID NO: 386	moltype = DNA length = 19	
FEATURE	Location/Qualifiers	
misc_feature	1..19	
	note = Synthetic sequence	
source	1..19	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 386		
gcctggctaa agtaggctt		19
SEQ ID NO: 387	moltype = DNA length = 18	
FEATURE	Location/Qualifiers	
misc_feature	1..18	
	note = Synthetic sequence	
source	1..18	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 387		
gcttttagga cgctcgg		18
SEQ ID NO: 388	moltype = DNA length = 18	
FEATURE	Location/Qualifiers	
misc_feature	1..18	
	note = Synthetic sequence	
source	1..18	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 388		
gcgggacact tcgagagg		18
SEQ ID NO: 389	moltype = DNA length = 52	
FEATURE	Location/Qualifiers	

-continued

misc_feature	1..52	
	note = Synthetic sequence	
source	1..52	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 389		
gtcgccccgc tccaggcgtc ggcggggggat cctttccgca tgggcctgcg cc		52
SEQ ID NO: 390	moltype = DNA length = 52	
FEATURE	Location/Qualifiers	
misc_feature	1..52	
	note = Synthetic sequence	
source	1..52	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 390		
ggcgcaggcc catgcggaaa ggatcccccg ccgacgcctg gagcggggcg ac		52
SEQ ID NO: 391	moltype = DNA length = 52	
FEATURE	Location/Qualifiers	
misc_feature	1..52	
	note = Synthetic sequence	
source	1..52	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 391		
gtcgccccgc tccaggcgtc ggcgggggcat cctttccgca tgggcctgcg cc		52
SEQ ID NO: 392	moltype = DNA length = 52	
FEATURE	Location/Qualifiers	
misc_feature	1..52	
	note = Synthetic sequence	
source	1..52	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 392		
ggcgcaggcc catgcggaaa ggatgccccg ccgacgcctg gagcggggcg ac		52
SEQ ID NO: 393	moltype = DNA length = 49	
FEATURE	Location/Qualifiers	
misc_feature	1..49	
	note = Synthetic sequence	
source	1..49	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 393		
ttacagtctc accacgcccc gtcccctctc cggtgagccc cgcgccttc		49
SEQ ID NO: 394	moltype = DNA length = 49	
FEATURE	Location/Qualifiers	
misc_feature	1..49	
	note = Synthetic sequence	
source	1..49	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 394		
gaaggcgcgg ggctcaacgg agagggggacg gggcgtggtg agactgtaa		49
SEQ ID NO: 395	moltype = DNA length = 49	
FEATURE	Location/Qualifiers	
misc_feature	1..49	
	note = Synthetic sequence	
source	1..49	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 395		
ttacagtctc accacgcccc gtccgctctc cggtgagccc cgcgccttc		49
SEQ ID NO: 396	moltype = DNA length = 49	
FEATURE	Location/Qualifiers	
misc_feature	1..49	
	note = Synthetic sequence	
source	1..49	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 396		

-continued

gaaggcgagg ggctcaacgg agagcggacg gggcgtgggtg agactgtaa	49
SEQ ID NO: 397	moltype = DNA length = 49
FEATURE	Location/Qualifiers
misc_feature	1..49
	note = Synthetic sequence
source	1..49
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 397	
gcgtctggga cgcaaggcgc cgtgggggct gccgggacgg gtccaagat	49
SEQ ID NO: 398	moltype = DNA length = 49
FEATURE	Location/Qualifiers
misc_feature	1..49
	note = Synthetic sequence
source	1..49
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 398	
atcttgacc cgtcccggca gccccacgg cgccttgcgt cccagacgc	49
SEQ ID NO: 399	moltype = DNA length = 49
FEATURE	Location/Qualifiers
misc_feature	1..49
	note = Synthetic sequence
source	1..49
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 399	
gcgtctggga cgcaaggcgc cgtgagggt gccgggacgg gtccaagat	49
SEQ ID NO: 400	moltype = DNA length = 49
FEATURE	Location/Qualifiers
misc_feature	1..49
	note = Synthetic sequence
source	1..49
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 400	
atcttgacc cgtcccggca gccctcacgg cgccttgcgt cccagacgc	49
SEQ ID NO: 401	moltype = DNA length = 49
FEATURE	Location/Qualifiers
misc_feature	1..49
	note = Synthetic sequence
source	1..49
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 401	
tctgcgtcag ggtttccttc ttttcagcc ccaccccgcg tgcattcca	49
SEQ ID NO: 402	moltype = DNA length = 49
FEATURE	Location/Qualifiers
misc_feature	1..49
	note = Synthetic sequence
source	1..49
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 402	
tgggatgcac gcggggtggg gctggaaaag aaggaaaccc tgacgcaga	49
SEQ ID NO: 403	moltype = DNA length = 49
FEATURE	Location/Qualifiers
misc_feature	1..49
	note = Synthetic sequence
source	1..49
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 403	
tctgcgtcag ggtttccttc ttttcagcc ccaccccgcg tgcattcca	49
SEQ ID NO: 404	moltype = DNA length = 49
FEATURE	Location/Qualifiers
misc_feature	1..49
	note = Synthetic sequence

-continued

```

source                1..49
                      mol_type = other DNA
                      organism = synthetic construct

SEQUENCE: 404
tgggatgcac gcggggtggg gctgcaaaag aaggaaaccc tgacgcaga          49

SEQ ID NO: 405        moltype = DNA  length = 49
FEATURE              Location/Qualifiers
misc_feature          1..49
                      note = Synthetic sequence
source                1..49
                      mol_type = other DNA
                      organism = synthetic construct

SEQUENCE: 405
ccctcattca ggttgatgtc ctaagcccca gaacctcaga atgggattg          49

SEQ ID NO: 406        moltype = DNA  length = 49
FEATURE              Location/Qualifiers
misc_feature          1..49
                      note = Synthetic sequence
source                1..49
                      mol_type = other DNA
                      organism = synthetic construct

SEQUENCE: 406
caatcccatt ctgaggttct ggggcttagg acatcaacct gaatgaggg          49

SEQ ID NO: 407        moltype = DNA  length = 49
FEATURE              Location/Qualifiers
misc_feature          1..49
                      note = Synthetic sequence
source                1..49
                      mol_type = other DNA
                      organism = synthetic construct

SEQUENCE: 407
ccctcattca ggttgatgtc ctaaccccca gaacctcaga atgggattg          49

SEQ ID NO: 408        moltype = DNA  length = 49
FEATURE              Location/Qualifiers
misc_feature          1..49
                      note = Synthetic sequence
source                1..49
                      mol_type = other DNA
                      organism = synthetic construct

SEQUENCE: 408
caatcccatt ctgaggttct ggggggtagg acatcaacct gaatgaggg          49

SEQ ID NO: 409        moltype = DNA  length = 49
FEATURE              Location/Qualifiers
misc_feature          1..49
                      note = Synthetic sequence
source                1..49
                      mol_type = other DNA
                      organism = synthetic construct

SEQUENCE: 409
gtggcctggc taaagtaggc tttactgggc tcctctctgc ctgcatcac          49

SEQ ID NO: 410        moltype = DNA  length = 49
FEATURE              Location/Qualifiers
misc_feature          1..49
                      note = Synthetic sequence
source                1..49
                      mol_type = other DNA
                      organism = synthetic construct

SEQUENCE: 410
gtgatgcagg cagagaggag cccagtaaag cctactttag ccaggccac          49

SEQ ID NO: 411        moltype = DNA  length = 49
FEATURE              Location/Qualifiers
misc_feature          1..49
                      note = Synthetic sequence
source                1..49
                      mol_type = other DNA
                      organism = synthetic construct

SEQUENCE: 411
gtggcctggc taaagtaggc tttagtgggc tcctctctgc ctgcatcac          49

```

-continued

SEQ ID NO: 412	moltype = DNA	length = 49	
FEATURE	Location/Qualifiers		
misc_feature	1..49	note = Synthetic sequence	
source	1..49	mol_type = other DNA	
		organism = synthetic construct	
SEQUENCE: 412			
gtgatgcagg cagagaggag cccactaaag cctacttttag ccaggccac			49
SEQ ID NO: 413	moltype = DNA	length = 48	
FEATURE	Location/Qualifiers		
misc_feature	1..48	note = Synthetic sequence	
source	1..48	mol_type = other DNA	
		organism = synthetic construct	
SEQUENCE: 413			
aatgcttttta ggacgcctcg gcgggagtgg cggggcaggg gggggggcg			48
SEQ ID NO: 414	moltype = DNA	length = 48	
FEATURE	Location/Qualifiers		
misc_feature	1..48	note = Synthetic sequence	
source	1..48	mol_type = other DNA	
		organism = synthetic construct	
SEQUENCE: 414			
cgcccccccc ctgccccgcc actcccgccg aggcgtccta aaagcatt			48
SEQ ID NO: 415	moltype = DNA	length = 48	
FEATURE	Location/Qualifiers		
misc_feature	1..48	note = Synthetic sequence	
source	1..48	mol_type = other DNA	
		organism = synthetic construct	
SEQUENCE: 415			
ggcgcgggac acttcgagag gaggcgggggt ttggagctgg agagatgt			48
SEQ ID NO: 416	moltype = DNA	length = 48	
FEATURE	Location/Qualifiers		
misc_feature	1..48	note = Synthetic sequence	
source	1..48	mol_type = other DNA	
		organism = synthetic construct	
SEQUENCE: 416			
acatctctcc agctccaaac cccgcctcct ctcgaagtgt cccgcgcc			48
SEQ ID NO: 417	moltype = DNA	length = 34	
FEATURE	Location/Qualifiers		
misc_feature	1..34	note = Synthetic sequence	
source	1..34	mol_type = other DNA	
		organism = synthetic construct	
SEQUENCE: 417			
gtttaattga gttgtcatat gttaataacg gtat			34
SEQ ID NO: 418	moltype = DNA	length = 34	
FEATURE	Location/Qualifiers		
misc_feature	1..34	note = Synthetic sequence	
source	1..34	mol_type = other DNA	
		organism = synthetic construct	
SEQUENCE: 418			
ataccgttat taacatatga caactcaatt aaac			34
SEQ ID NO: 419	moltype = DNA	length = 50	
FEATURE	Location/Qualifiers		
misc_feature	1..50	note = Synthetic sequence	
source	1..50	mol_type = other DNA	

organism = synthetic construct	
SEQUENCE: 419	
tgtcgccccg ctccaggcgt cggtcggcgg gagtggcggg gcaggggggg	50
SEQ ID NO: 420	moltype = DNA length = 83
FEATURE	Location/Qualifiers
misc_feature	1..83
	note = Synthetic sequence
source	1..83
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 420	
tgtcgccccg ctccaggcgt cggataccgt tattaacata tgacaactca attaaaccgg	60
cgggagtggc ggggcagggg ggg	83
SEQ ID NO: 421	moltype = DNA length = 77
FEATURE	Location/Qualifiers
misc_feature	1..77
	note = Synthetic sequence
source	1..77
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 421	
tgtcgccccg ctccaggcgt cggttattaa catatgacaa ctcaattaaa ccggcggggag	60
tggcggggca gggggggg	77
SEQ ID NO: 422	moltype = DNA length = 85
FEATURE	Location/Qualifiers
misc_feature	1..85
	note = Synthetic sequence
source	1..85
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 422	
tgtcgccccg ctccaggcgt cgggtttaat tgagttgtca tatgttaata acggtatctc	60
ggcgggagtg gcggggcagg gggggg	85
SEQ ID NO: 423	moltype = DNA length = 74
FEATURE	Location/Qualifiers
misc_feature	1..74
	note = Synthetic sequence
source	1..74
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 423	
tcgccccgct ccaggcgctg gtattaacat atgacaactc aattaaaccg gcgggagtg	60
cggggcaggg gggg	74
SEQ ID NO: 424	moltype = DNA length = 72
FEATURE	Location/Qualifiers
misc_feature	1..72
	note = Synthetic sequence
source	1..72
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 424	
tgtcgccccg ctccaggcgt cggattaaca tatgacaact caattaaacc gggagtggcg	60
gggcaggggg gg	72
SEQ ID NO: 425	moltype = DNA length = 17
FEATURE	Location/Qualifiers
misc_feature	1..17
	note = Synthetic sequence
source	1..17
	mol_type = other DNA
	organism = synthetic construct
SEQUENCE: 425	
gctccatgcg gcggcgg	17
SEQ ID NO: 426	moltype = DNA length = 17
FEATURE	Location/Qualifiers
misc_feature	1..17
	note = Synthetic sequence
source	1..17
	mol_type = other DNA
	organism = synthetic construct

-continued

SEQUENCE: 426		
gcttcaggcg taggggg		17
SEQ ID NO: 427	moltype = DNA length = 17	
FEATURE	Location/Qualifiers	
misc_feature	1..17	
	note = Synthetic sequence	
source	1..17	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 427		
gctccggggg ccggcgg		17
SEQ ID NO: 428	moltype = DNA length = 17	
FEATURE	Location/Qualifiers	
misc_feature	1..17	
	note = Synthetic sequence	
source	1..17	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 428		
gctccaagcg gcggcgc		17
SEQ ID NO: 429	moltype = DNA length = 18	
FEATURE	Location/Qualifiers	
misc_feature	1..18	
	note = Synthetic sequence	
source	1..18	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 429		
gctcttaggc cgccgcg		18
SEQ ID NO: 430	moltype = DNA length = 17	
FEATURE	Location/Qualifiers	
misc_feature	1..17	
	note = Synthetic sequence	
source	1..17	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 430		
gcaccaggcg gcggcgg		17
SEQ ID NO: 431	moltype = DNA length = 18	
FEATURE	Location/Qualifiers	
misc_feature	1..18	
	note = Synthetic sequence	
source	1..18	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 431		
gctttaagga tgccttgg		18
SEQ ID NO: 432	moltype = DNA length = 18	
FEATURE	Location/Qualifiers	
misc_feature	1..18	
	note = Synthetic sequence	
source	1..18	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 432		
acttttagga ggcctagg		18
SEQ ID NO: 433	moltype = DNA length = 18	
FEATURE	Location/Qualifiers	
misc_feature	1..18	
	note = Synthetic sequence	
source	1..18	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 433		
gcatttagga ggcctagg		18
SEQ ID NO: 434	moltype = DNA length = 17	
FEATURE	Location/Qualifiers	
misc_feature	1..17	

-continued

source	note = Synthetic sequence 1..17 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 434 gctccaggcg tgggcgg		17
SEQ ID NO: 435 FEATURE misc_feature	moltype = DNA length = 18 Location/Qualifiers 1..18 note = Synthetic sequence	
source	1..18 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 435 gcttttggga cgcctcag		18
SEQ ID NO: 436 FEATURE misc_feature	moltype = DNA length = 21 Location/Qualifiers 1..21 note = Synthetic sequence	
misc_difference	19 note = n = a, c, t, or g	
misc_difference	20 note = r = a or g	
source	1..21 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 436 gcttttagga cgcctcggnr g		21
SEQ ID NO: 437 FEATURE misc_feature	moltype = DNA length = 21 Location/Qualifiers 1..21 note = Synthetic sequence	
source	1..21 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 437 gcttttagga ggtctcggcg g		21
SEQ ID NO: 438 FEATURE misc_feature	moltype = DNA length = 20 Location/Qualifiers 1..20 note = Synthetic sequence	
misc_difference	19 note = n = a, t, c, or g	
misc_difference	20 note = r = a or g	
source	1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 438 gctccaggcg tcggcggnrg		20
SEQ ID NO: 439 FEATURE misc_feature	moltype = DNA length = 20 Location/Qualifiers 1..20 note = Synthetic sequence	
source	1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 439 ggcggaggcg gcggcggcgg		20
SEQ ID NO: 440 FEATURE misc_feature	moltype = DNA length = 20 Location/Qualifiers 1..20 note = Synthetic sequence	
source	1..20 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 440 ggttgaggcg gaggcggcgg		20
SEQ ID NO: 441	moltype = DNA length = 20	

-continued

FEATURE	Location/Qualifiers	
misc_feature	1..20	
	note = Synthetic sequence	
source	1..20	
	mol_type = other DNA	
	organism = synthetic construct	
SEQUENCE: 441		
ggacctggca gcggcggtgg		20
SEQ ID NO: 442	moltype = RNA length = 80	
FEATURE	Location/Qualifiers	
misc_feature	1..80	
	note = Synthetic sequence	
source	1..80	
	mol_type = other RNA	
	organism = synthetic construct	
SEQUENCE: 442		
gttttagagc tagaaatagc aagttaaaat aaggctagtc cgttatcaac ttgaaaaagt		60
ggcaccgagt cggtgctttt		80

- What is claimed is:
1. A method for inhibiting, treating, and/or preventing a polyglutamine disorder in a subject in need thereof, said method comprising reducing the expression of a mutant protein encoded by an allele of a gene associated with the polyglutamine disorder in said subject, wherein said polyglutamine disorder is characterized by an abnormally high number of cytosine-adenine-guanine (CAG) repeats resulting in an extended polyglutamine tract in the encoded for mutant protein, and wherein said method comprising administering to the subject a nucleic acid molecule encoding Cas9 and at least one guide RNA.

2. A method for reducing the expression of a mutant protein encoded by an allele of a gene associated with a polyglutamine disorder in a cell, said method comprising delivering to the cell a nucleic acid molecule encoding Cas9 and at least one guide RNA, wherein said polyglutamine disorder is characterized by an abnormally high number of cytosine-adenine-guanine (CAG) repeats resulting in an extended polyglutamine tract in the encoded for mutant protein.

3. The method of claim 1, wherein said polyglutamine disorder is Huntington's disorder and said gene is the huntingtin (HTT) gene.

4. The method of claim 1, wherein said method comprises administering to the subject a nucleic acid molecule encoding Cas9, at least one guide RNA, and, optionally at least one donor DNA.

5. The method of claim 1, wherein said guide RNA are administered as a nucleic acid molecule encoding said guide RNA.

6. The method of claim 5, wherein said nucleic acid molecules are administered in an expression vector.

7. The method of claim 6, wherein said expression vector is a viral vector.

8. The method of claim 1, wherein at least one guide RNA targets a sequence adjacent to a PAM present on only one allele of the gene.

9. The method of claim 8, wherein said PAM is present only on the mutant allele.

10. The method of claim 1, wherein two guide RNAs are administered.

11. The method of claim 10, wherein at least one of the guide RNA targets a sequence adjacent to a PAM present on only one allele of the gene.

12. The method of claim 10, wherein a first guide RNA targets a sequence within the promoter or 5' untranslated region of the gene and a second guide RNA targets a sequence within the first intron of the gene.

13. The method of claim 10, wherein the first guide RNA comprises sgHD1 and the second guide RNA comprises sgHDi3.

14. An isolated guide RNA, wherein said guide RNA targets a sequence adjacent to a PAM present on only one allele of a gene associated with a polyglutamine disorder.

15. A nucleic acid molecule encoding at least one of the guide RNA of claim 14.

16. A vector comprising the nucleic acid molecule of claim 15.

17. The vector of claim 16, which is a viral vector.

18. The vector of claim 16, further comprising a nucleic acid molecule encoding Cas9.

19. A composition comprising at least one nucleic acid molecule of claim 15 and a carrier.

20. A composition comprising at least one guide RNA of claim 14 and a carrier.
- * * * * *