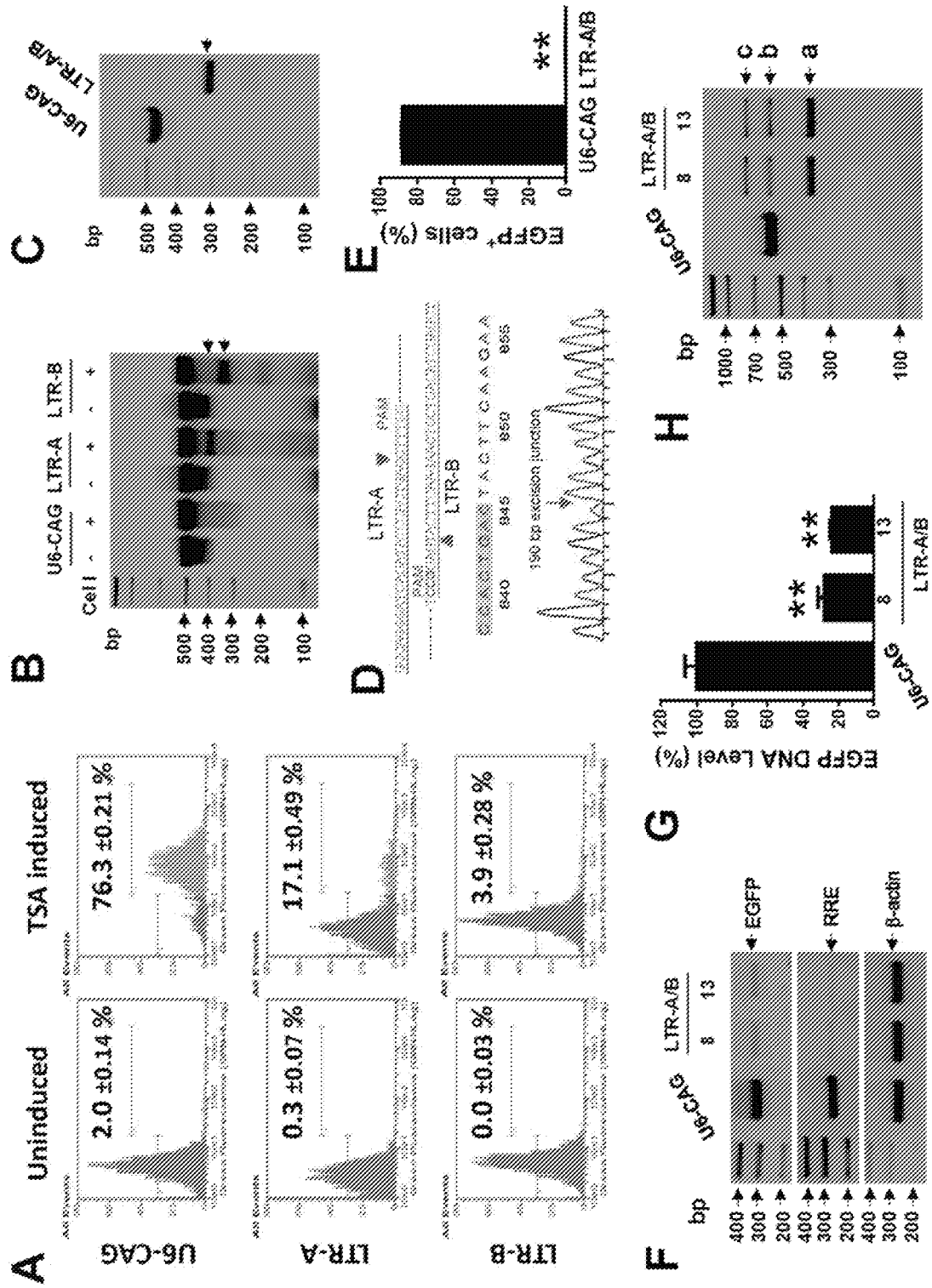

A
Uninduced
TSA induced
U6-CAG
LTR-A
LTR-B
2.0 ±0.14 %
76.3 ±0.21 %
0.3 ±0.07 %
17.1 ±0.49 %
0.0 ±0.03 %
3.9 ±0.28 %
B
U6-CAG LTR-A LTR-B
Cell
bp
500
400
300
200
100
C
U6-CAG LTR-A/B
bp
500
400
300
200
100
D
LTR-A
PAM
LTR-B
840 845 850 855
190 bp excision junction
E
EGFP+ cells (%)
100
80
60
40
20
0
**
U6-CAG LTR-A/B
F
U6-CAG LTR-A/B
8 13
bp
EGFP
RRE
β-actin
400
300
200
G
EGFP DNA Level (%)
120
100
80
60
40
20
0
**
**
U6-CAG
8 13
LTR-A/B
H
U6-CAG LTR-A/B
8 13
bp
1000
700
500
300
100
c
b
a

FIGURE 1

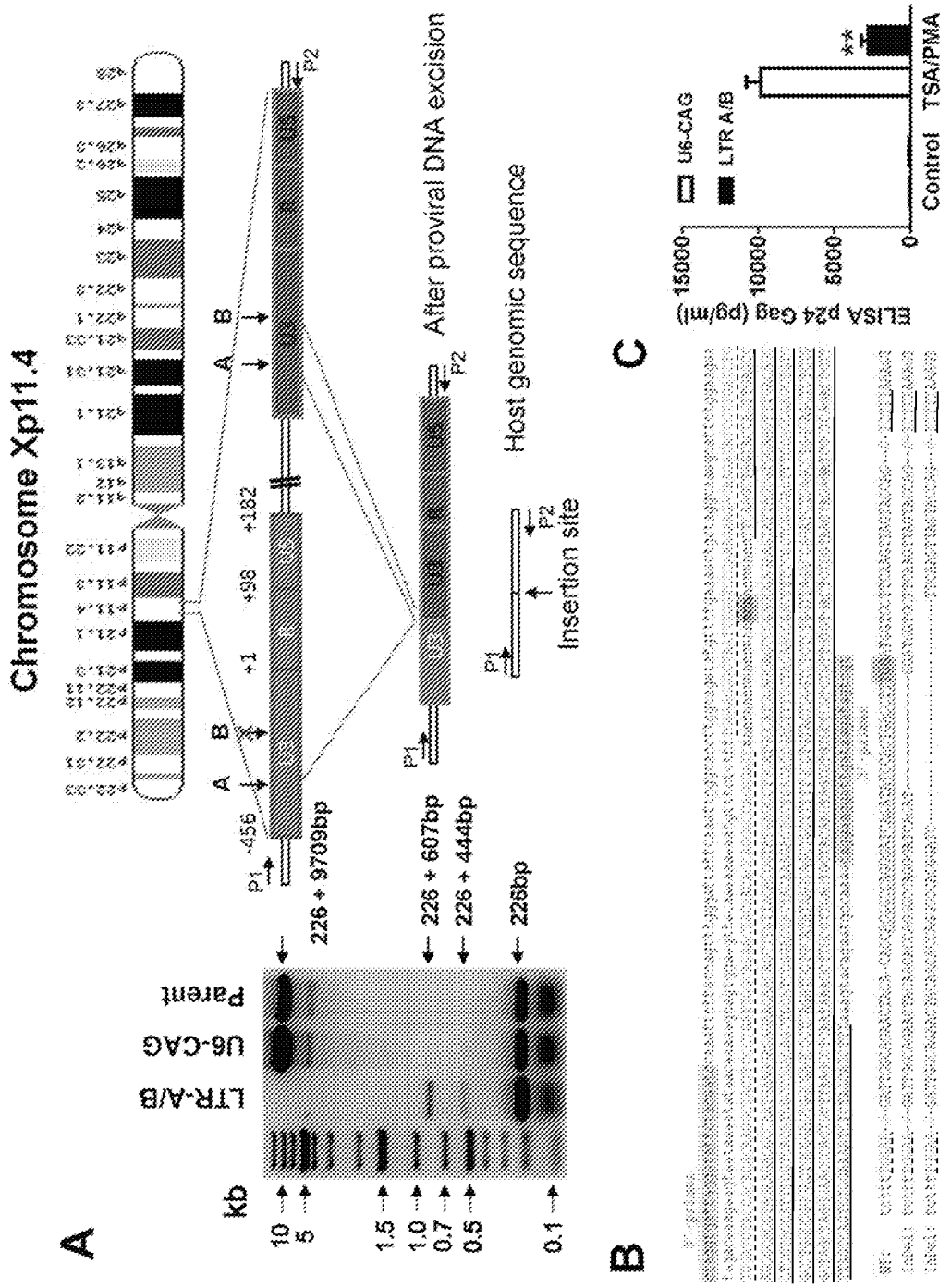

A
Chromosome Xp11.4
P1
P2
-456
+1
+98
+182
A B
226 + 9709bp
After proviral DNA excision
Host genomic sequence
Insertion site
226 + 607bp
226 + 444bp
226bp
kb
10
5
1.5
1.0
0.7
0.5
0.1
LTR-A/B
U6-CAG
Parent
B
C
ELISA p24 Gag (pg/ml)
15000
10000
5000
0
U6-CAG
LTR A/B
**
Control
TSA/PMA

FIGURE 2

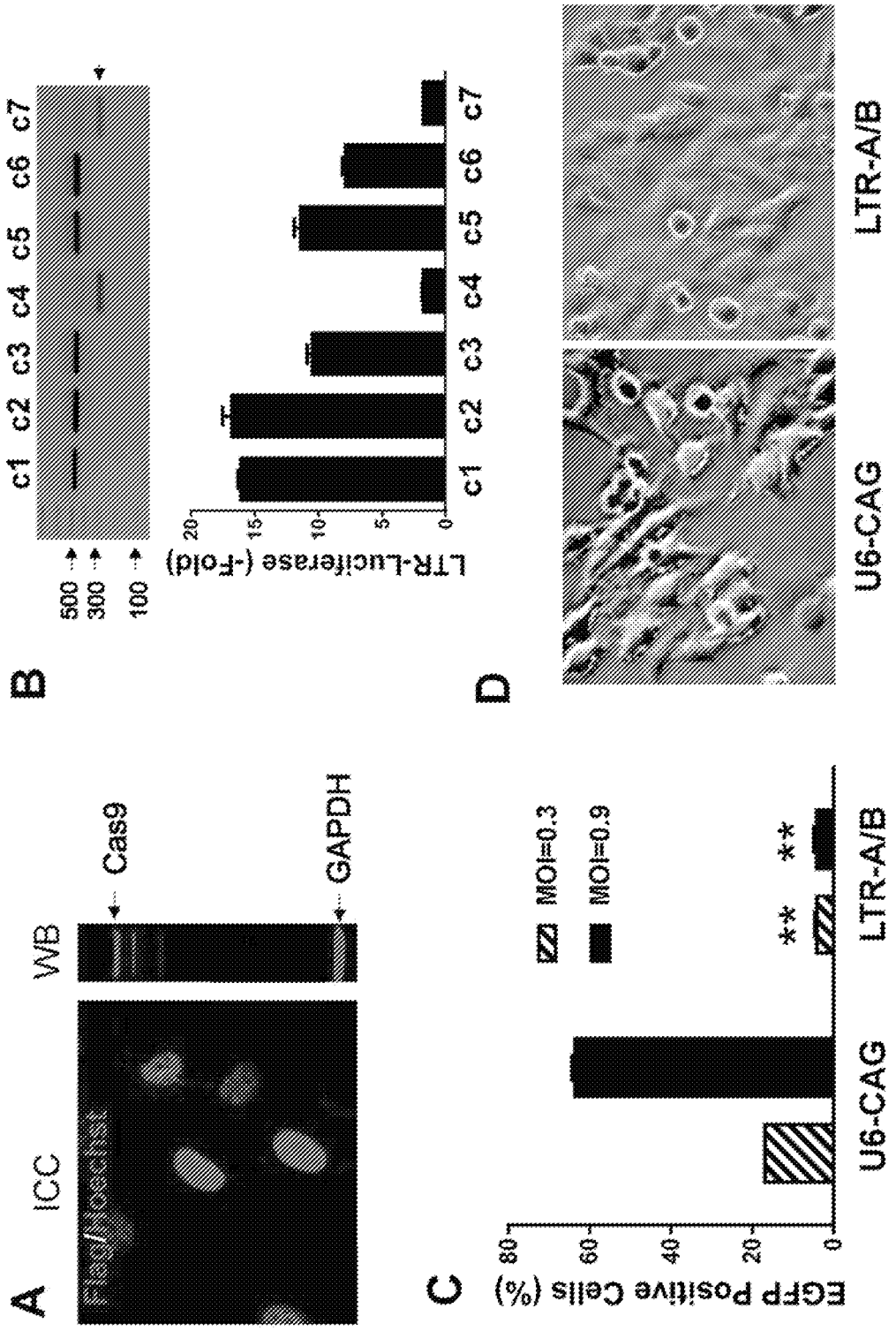
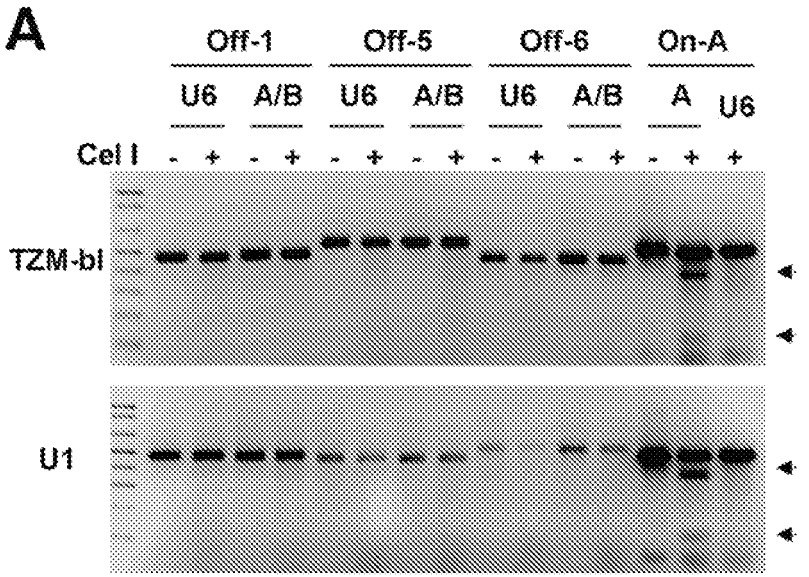
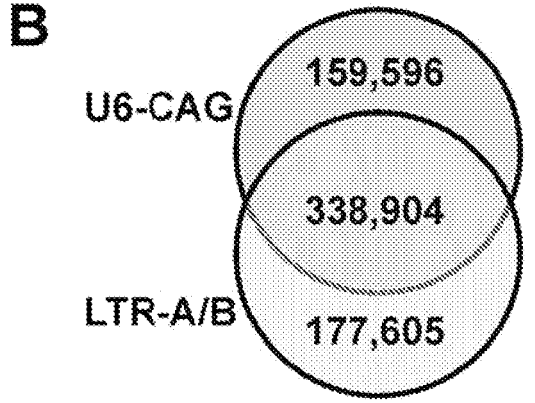
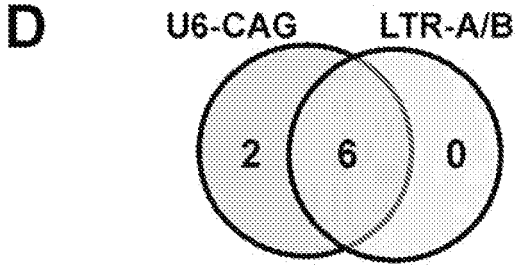

A
ICC
Flag/Hoechst
WB
Cas9
GAPDH
B
c1 c2 c3 c4 c5 c6 c7
500
300
100
LTR-Luciferase (-Fold)
20
15
10
5
0
c1 c2 c3 c4 c5 c6 c7
C
EGFP Positive Cells (%)
80
60
40
20
0
MOI=0.3
MOI=0.9
**
**
U6-CAG
LTR-A/B
D
U6-CAG
LTR-A/B

FIGURE 3







C

Name	gRNA Target Site				Indels							Genotype	
	genome	start	end	strand	genome	start	end	reference	quality	filter	mutation	U6-CAG	LTR-A/B
LTR4	chr11	103632094	103632108	+	chr11	1.04E+08	103632209	ATC	272.76	PASS	A	1/1	1/1
LTR4	chr14	24193190	24193204	-	chr14	24192852	24192875	GAGAGCTGTC TCAAAAAAAG	96.74	PASS	G	0/1	0/0
LTR6	chr6	98501053	98501068	+	chr6	98501264	98501267	CAGA	323.75	PASS	C	1/1	1/1
LTR6	chr6	98501053	98501068	+	chr6	98501280	98501282	GCA	208.27	PASS	G	1/1	1/1
LTR7	chr7	72210628	72210642	-	chr7	72210827	72210828	GT	111.24	LowQual	G	0/1	0/1
LTR7	chr7	72210628	72210642	-	chr7	72210828	72210832	GTGT	25.74	LowQual	G	0/1	0/0
LTR9	chr9	95871690	95871704	+	chr9	95871594	95871596	GCA	456.73	PASS	C	0/1	0/1
LTR6	HIV	162	184	+	HIV	371	375	CTCA	14.87	LowQual	C	0/0	0/1
LTR8	chr15	82132445	82132459	+	chr15	82132764	82132764	T	109.76	PASS	TTGTG,TTGTGTGTG	1/1	1/1
LTR6	HIV	9281	9313	-	HIV	9816	9816	A	174.66	PASS	ATGTGTCCCCTGTGTTCTT	0/0	0/1

FIGURE 4

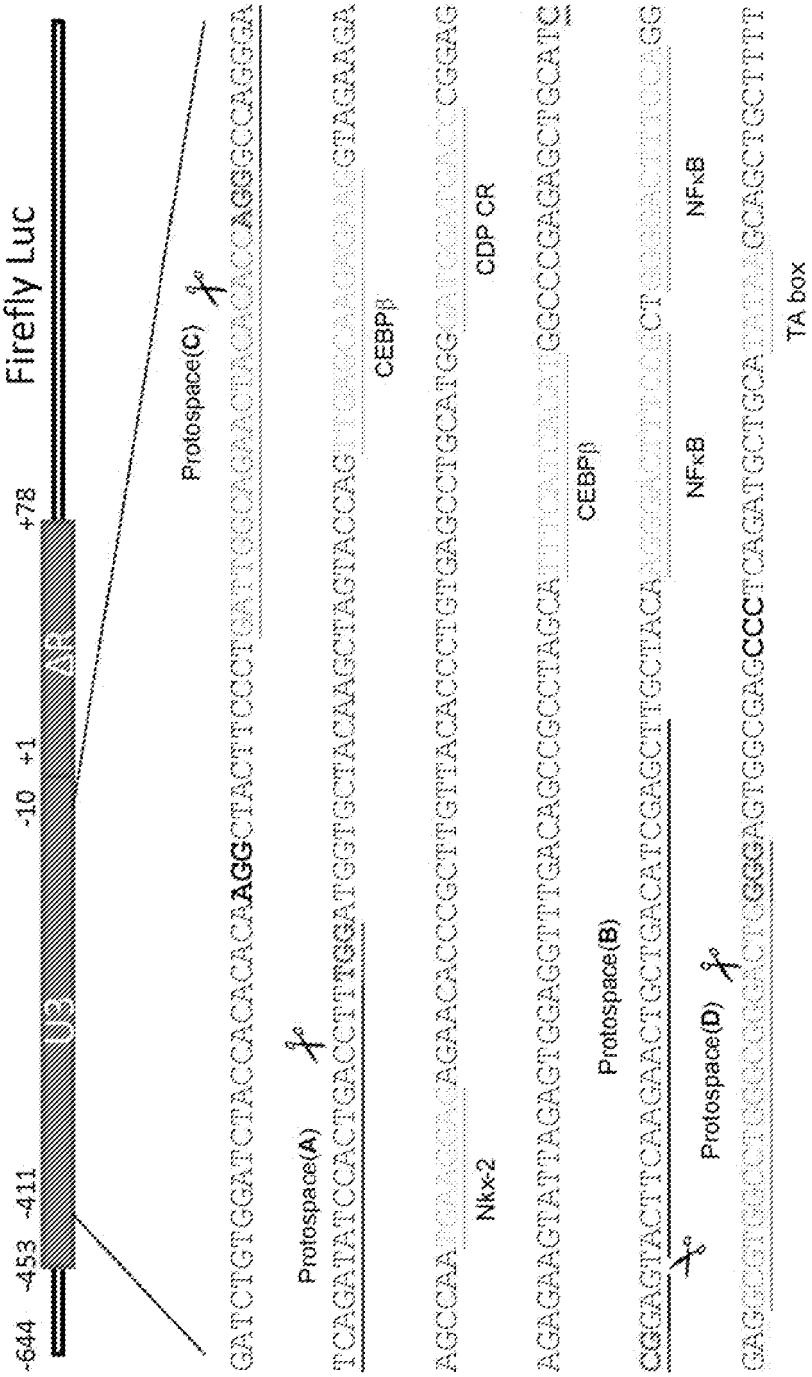

Firefly Luc
-644
-453
-411
U3
-10
+1
AS
+78
Protospace(C)
GATCTGTGGATCTACCACACACAAGGCTACTTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGA
Protospace(A)
TCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAG
CEBPβ
AGCCAA
AGAACACCCGCTTGTTACACCCTGTGAGCCTGCATGG
Nkx-2
CDP CR
AGAGAAGTATTAGAGTGGAGGTTTGACAGCCGCCTAGCA
CEBPβ
Protospace(B)
CGGAGTACTTCAAGAACTGCTGACATCGAGCTTGCTACA
NFκB
Protospace(D)
NFκB
GAGCGCTTGCCTTGCCCGCTCGGGAGTGGCGAGCCCTCAGATGCTGCATA
TA box

FIGURE 5

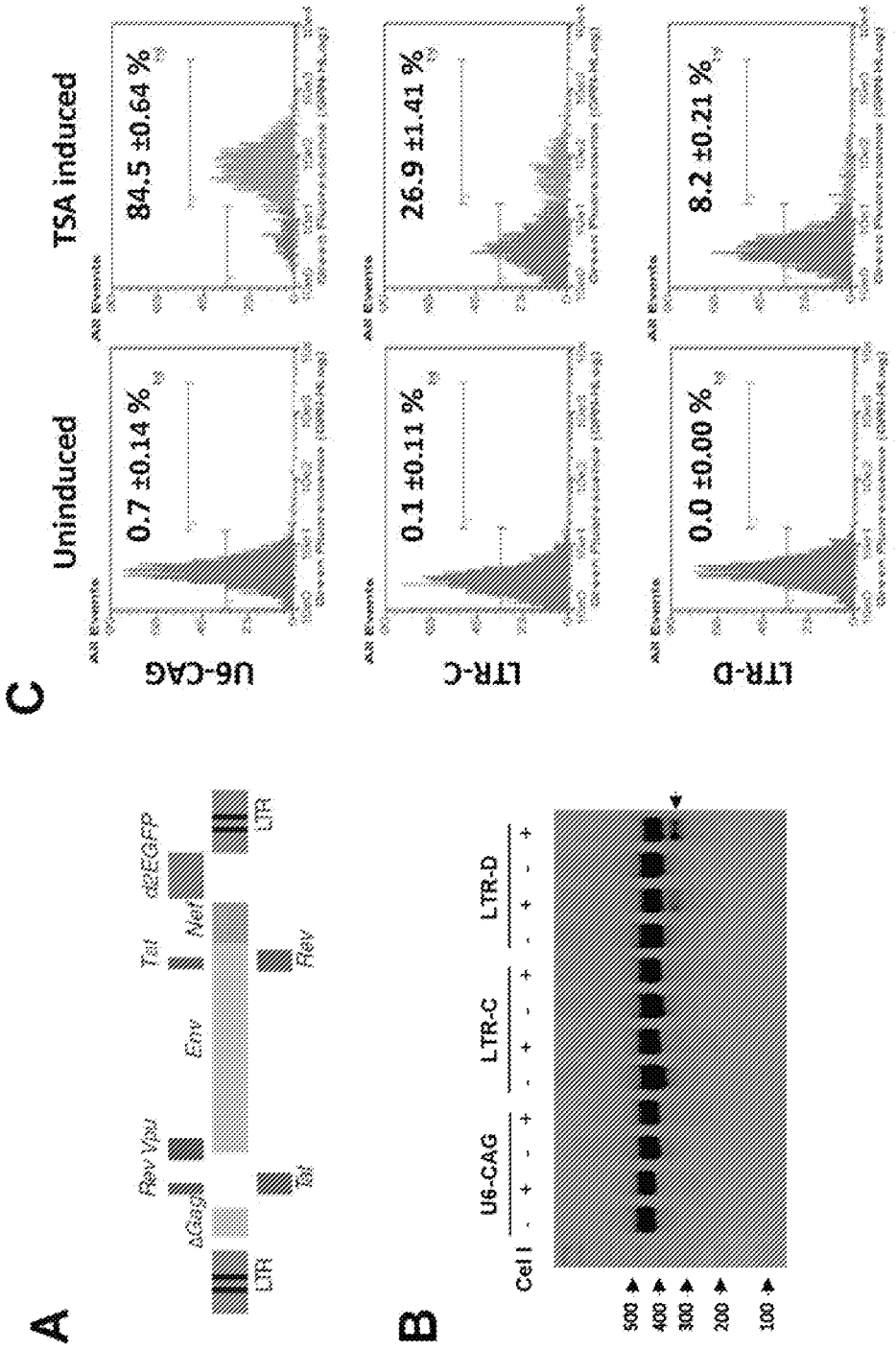

A
LTR
ΔGag
Rev Vpu
Tat
Rev
Env
Tat
d2EGFP
Nef
LTR
B
U6-CAG
LTR-C
LTR-D
Cell I - + - + - + - + - +
500
400
300
200
100
C
Uninduced
TSA induced
U6-CAG
LTR-C
LTR-D
0.7 ±0.14 %
84.5 ±0.64 %
0.1 ±0.11 %
26.9 ±1.41 %
0.0 ±0.00 %
8.2 ±0.21 %

FIGURE 6

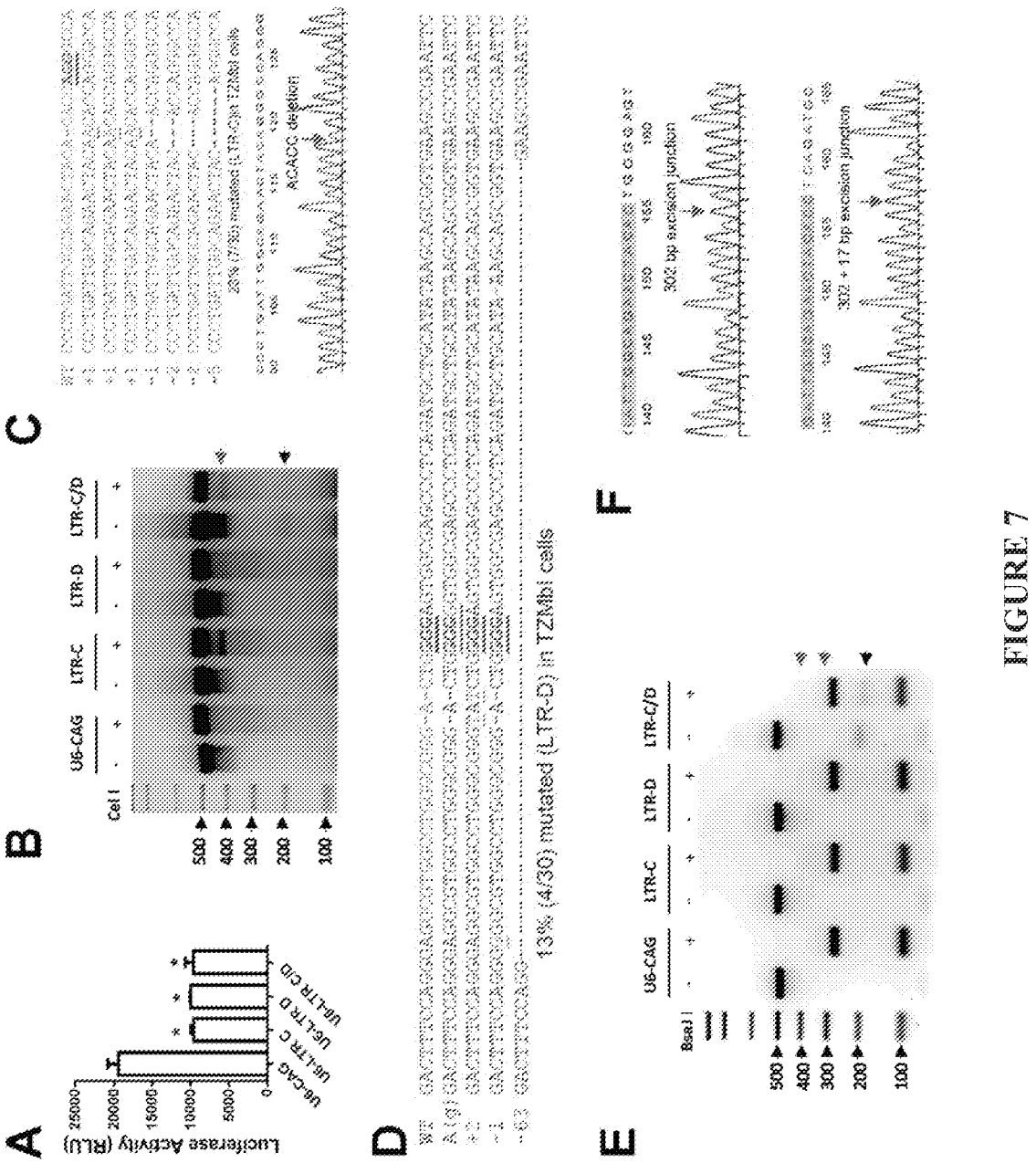

A
Luciferase Activity (RLU)
25000
20000
15000
10000
5000
0
U6-CAG
U6-LTR C
U6-LTR D
U6-LTR C/D
B
U6-CAG
LTR-C
LTR-D
LTR-C/D
Cell
500
400
300
200
100
C
25% (7/30) mutated (LTR-C) in TZMbl cells
ACAGC deletion
D
WT
A (9)
+2
-1
-63
13% (4/30) mutated (LTR-D) in TZMbl cells
E
U6-CAG
LTR-C
LTR-D
LTR-C/D
BssaI
500
400
300
200
100
F
302 bp excision junction
302 + 17 bp excision junction

FIGURE 7

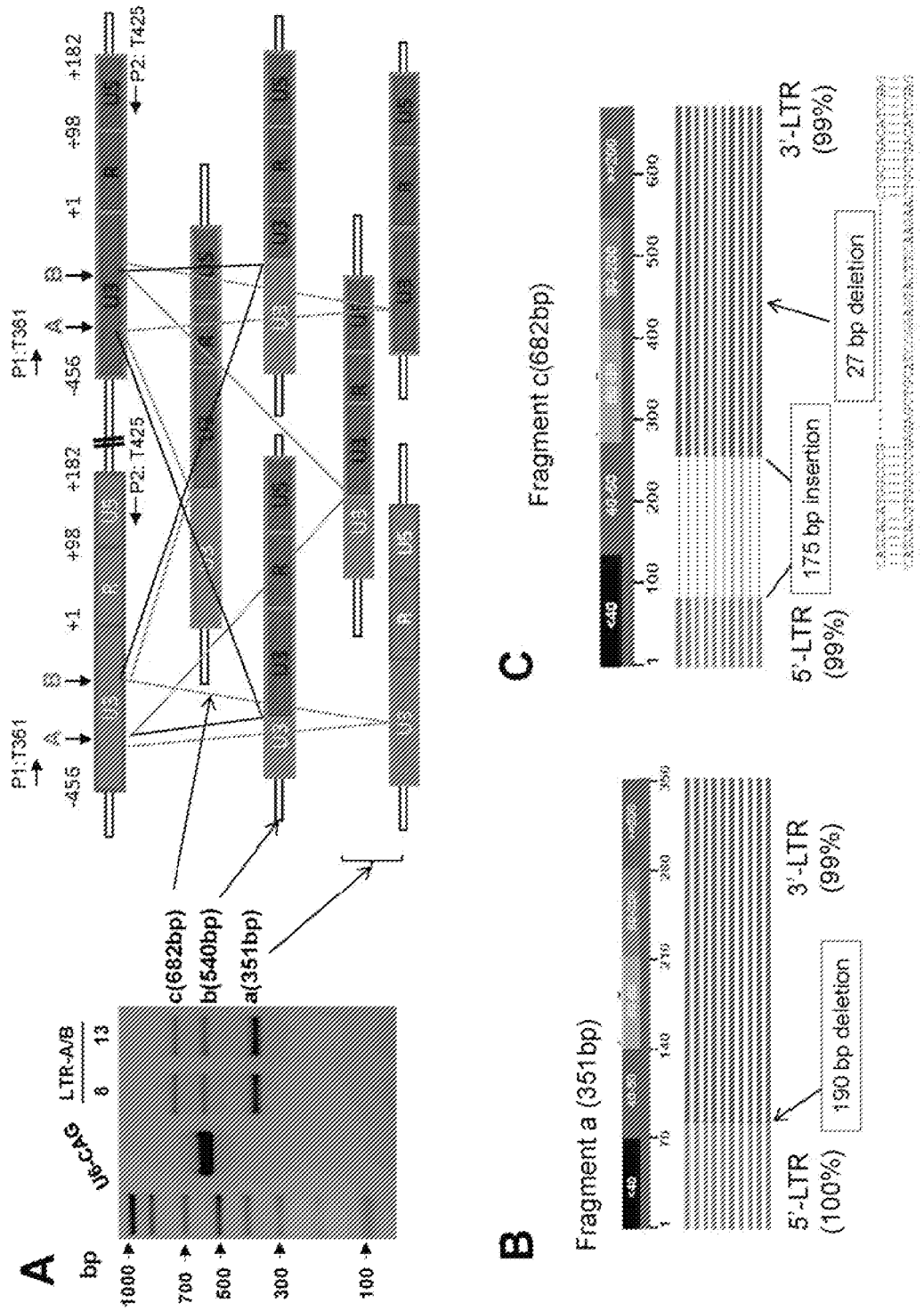

A
bp
1000
700
500
300
100
U6-CAG
LTR-A/B
8
13
c(682bp)
b(540bp)
a(351bp)
P1:T361
P2: T425
-456
+1
+98
+182
B
Fragment a (351bp)
5'-LTR
(100%)
190 bp deletion
3'-LTR
(99%)
C
Fragment c(682bp)
5'-LTR
(99%)
175 bp insertion
27 bp deletion
3'-LTR
(99%)

FIGURE 8

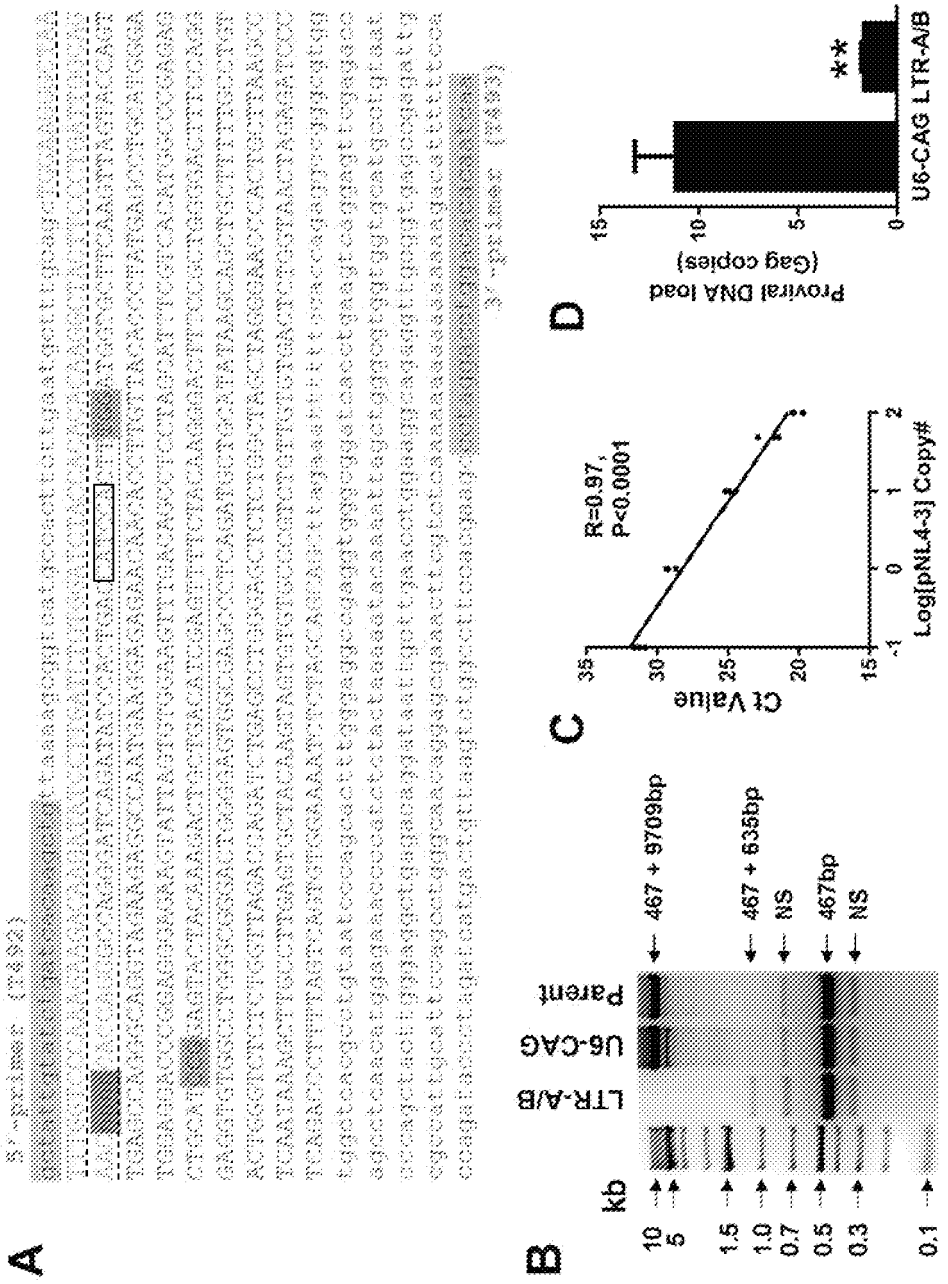

A
B
kb
LTR-A/B
U6-CAG
Parent
467 + 9709bp
467 + 835bp
NS
467bp
NS
C
R=0.97,
P<0.0001
Ct Value
Log[pNL4-3] Copy#
D
Provirial DNA load
(Gag copies)
U6-CAG LTR-A/B

FIGURE 9

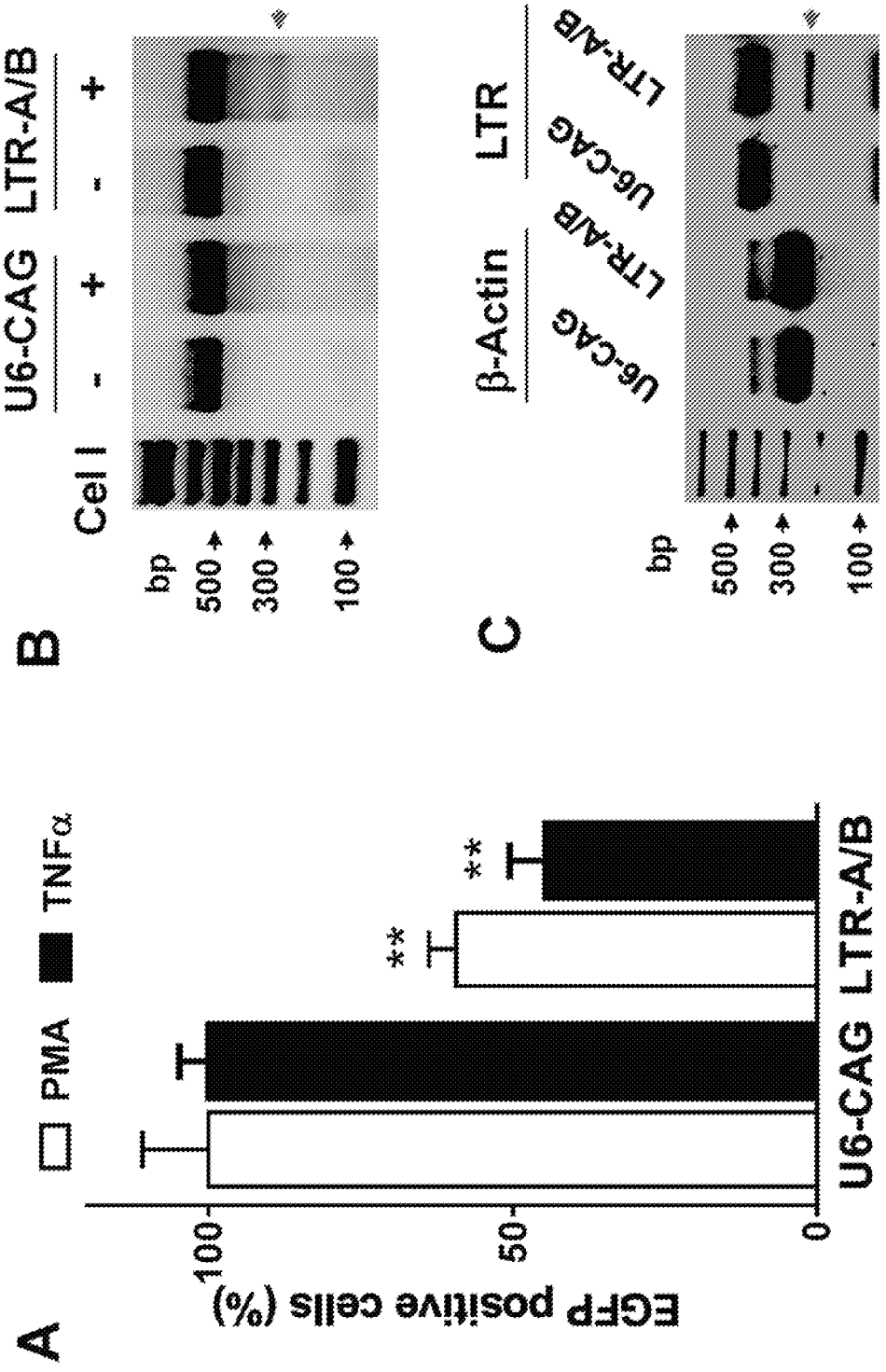

A
PMA
TNFα
EGFP positive cells (%)
100
50
0
**
**
U6-CAG
LTR-A/B
B
U6-CAG
LTR-A/B
Cel I
-
+
-
+
bp
500
300
100
C
β-Actin
U6-CAG
LTR-A/B
LTR
U6-CAG
LTR-A/B
bp
500
300
100

FIGURE 10

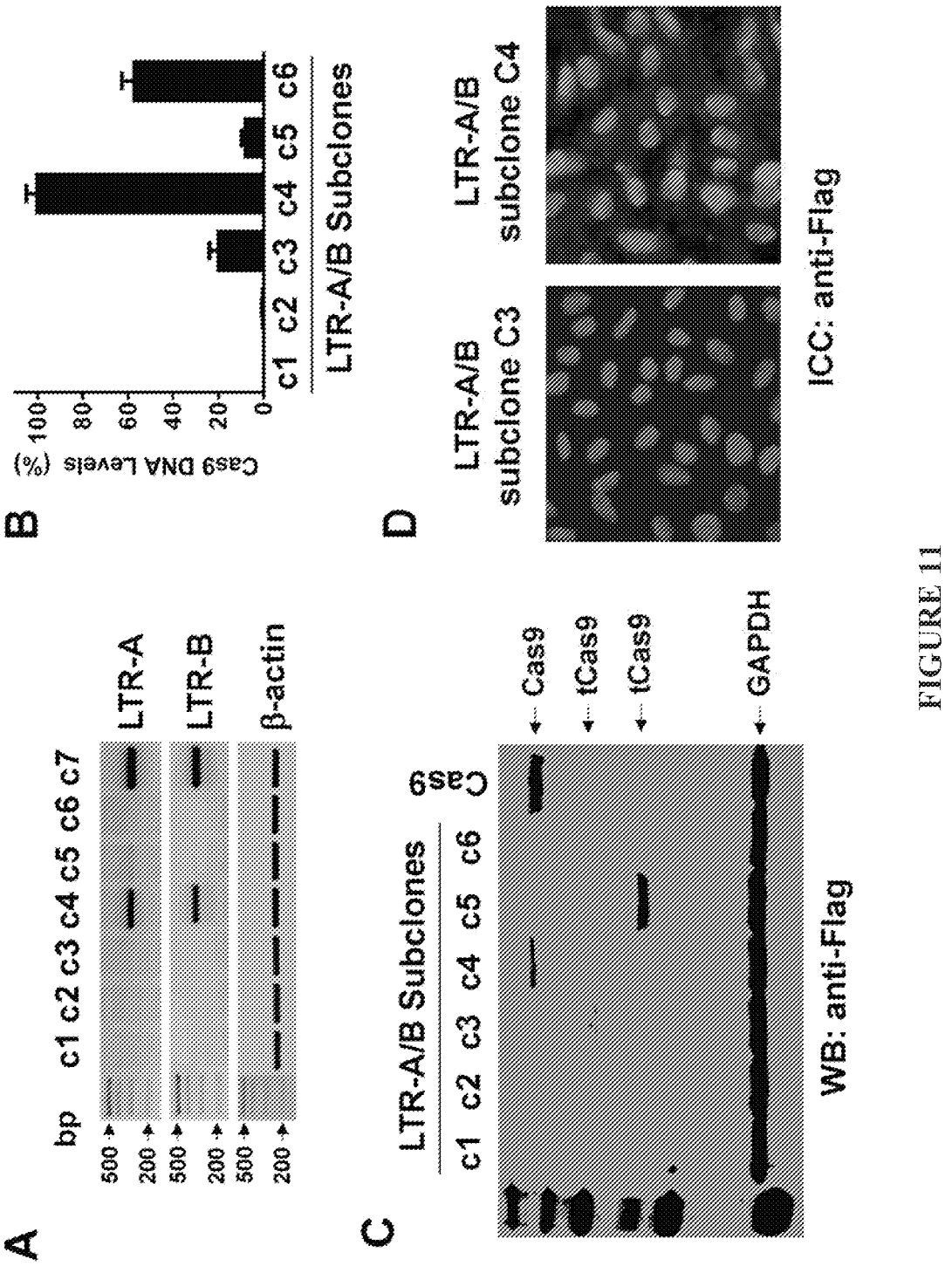

A
bp c1 c2 c3 c4 c5 c6 c7
500
200
LTR-A
LTR-B
β-actin
B
Cas9 DNA Levels (%)
100
80
60
40
20
0
c1 c2 c3 c4 c5 c6
LTR-A/B Subclones
C
LTR-A/B Subclones
c1 c2 c3 c4 c5 c6 Cas9
Cas9
tCas9
tCas9
GAPDH
WB: anti-Flag
D
LTR-A/B subclone C3
LTR-A/B subclone C4
ICC: anti-Flag

FIGURE 11

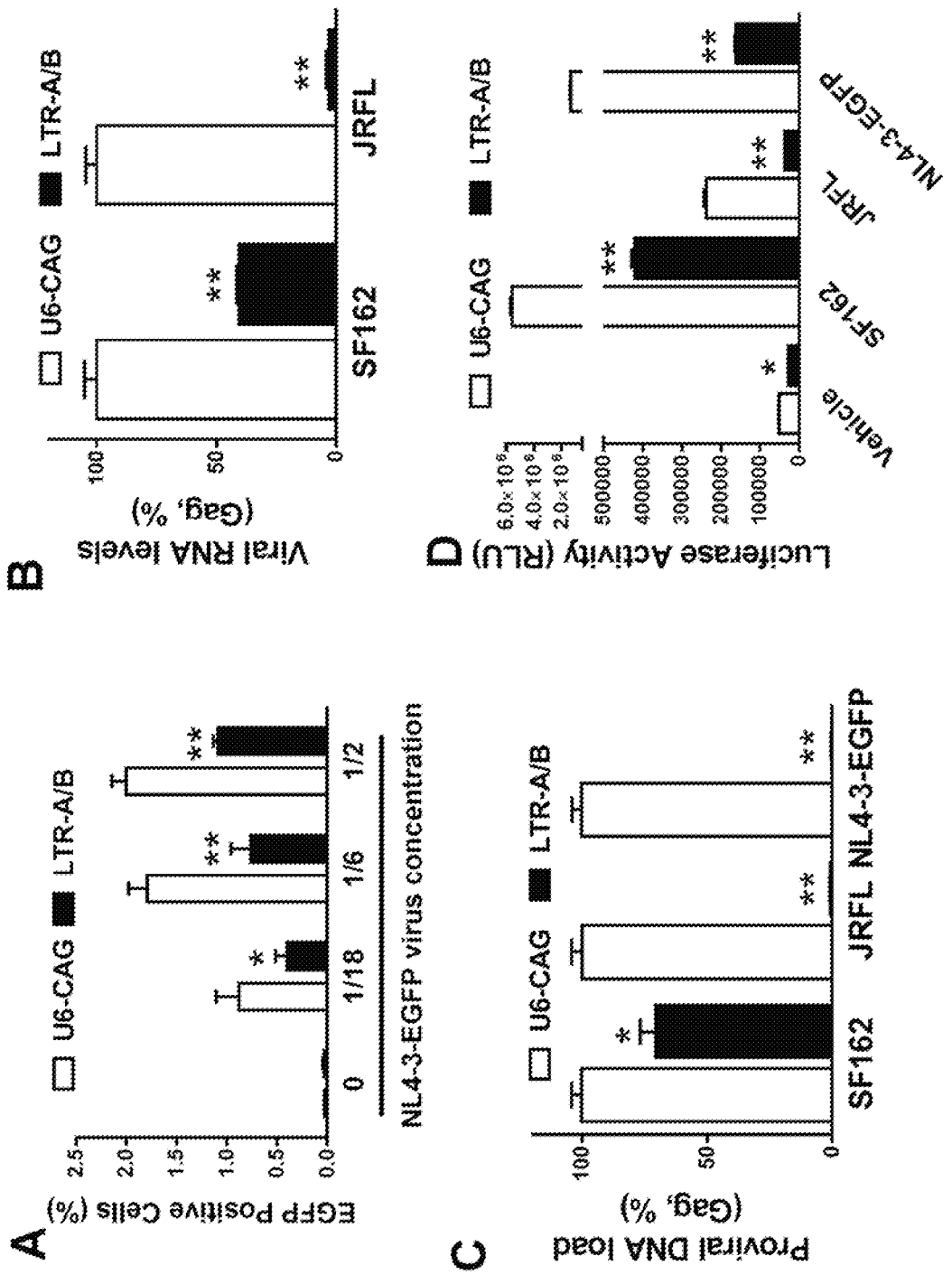

A
EGFP Positive Cells (%)
U6-CAG
LTR-A/B
NL4-3-EGFP virus concentration
0
1/18
1/6
1/2
B
Viral RNA levels (Gag, %)
SF162
JRFL
C
Proviral DNA load (Gag, %)
SF162
JRFL NL4-3-EGFP
D
Luciferase Activity (RLU)
Vehicle
SF162
JRFL
NL4-3-EGFP

FIGURE 12

Table S1. Predicted LTR gRNAs and their off-target numbers (100% match)

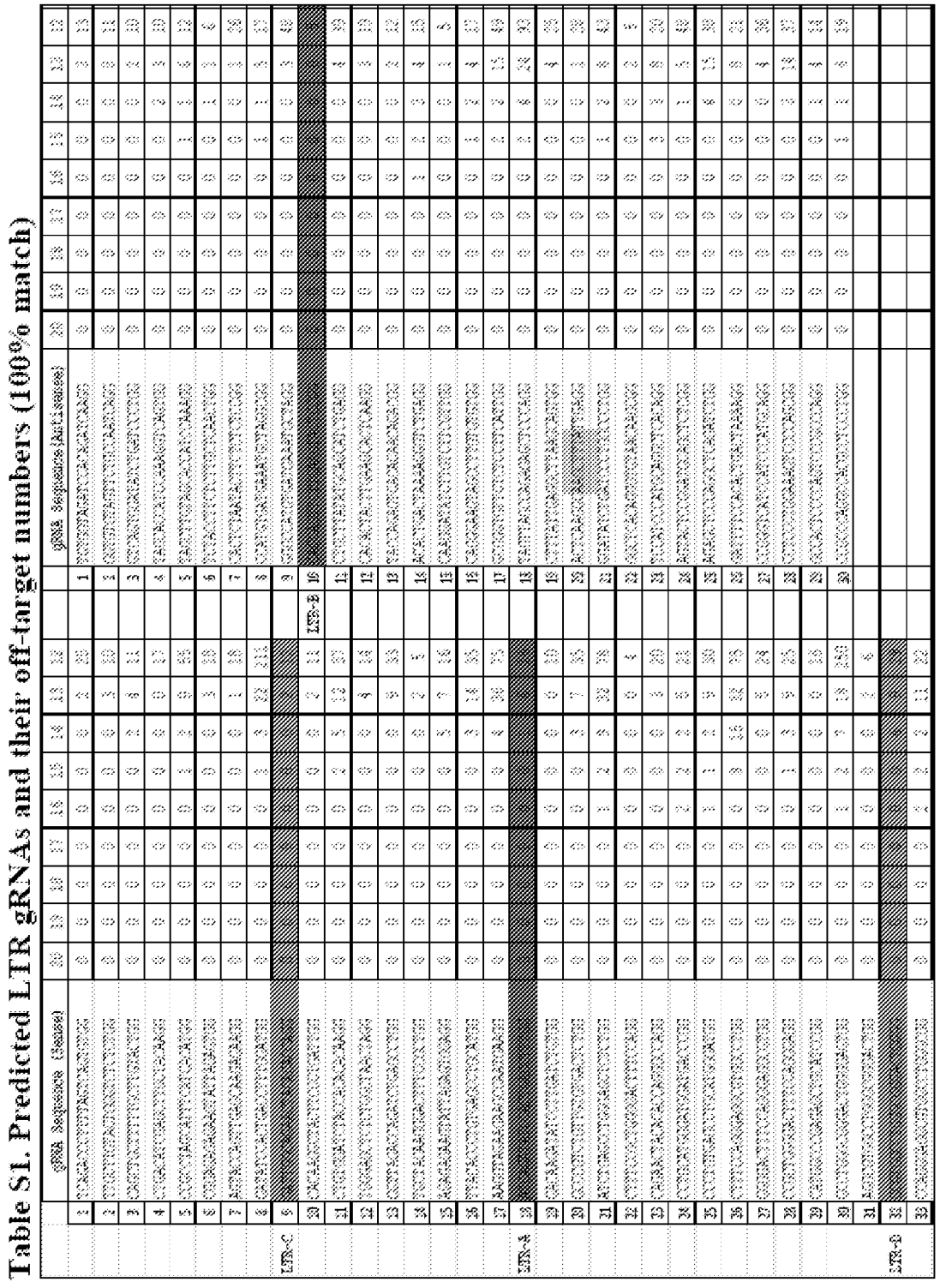


FIGURE 13

Table S2. Oligonucleotides for gRNA targeting sites and primers used for PCR and sequencing

Target name	Direction	Sequences (5' to 3')
LTR-A	T353: Sense	aaacAGGGCCAGGGATCAGATATCCACTGACCTTgt
	T354: Antisense	taaacAAGGTCAGTGGATATCTGATCCCTGGCCCT
LTR-B	T355: Anti-sense	aaacAGCTCGATGTCAGCAGTTCTTGAAGTACTCgt
	T356: Sense	taaacGAGTACTTCAAGAACTGCTGACATCGAGCT
LTR-C	T357: Sense	caccGATTGGCAGAACTACACACC
	T358: Antisense	aaacGGTGTGTAGTTCTGCCAATC
LTR-D	T359: Sense	caccGCGTGGCCTGGGCGGGACTG
	T360: Antisense	aaacCAGTCCCGCCCAGGCCACGC
PCR primer		
LTR -453/F	Sense	TGGAAGGGCTAATTCACTCCCAAC
LTR +43/R	Anti-sense	CCGAGAGCTCCCAGGCTCAGATCT
LTR -411/F	T361: Sense	cacaGATCTGTGGATCTACCACACACA
LTR +129/R	T363: Anti-sense	aaacGAGTCACACAACAGACGGGC
Cas-hU6/5'/XhoBm	T351: Sense	cgcctcgaggatccGAGGGCCTATTTCCCATGATTCC
Cas-CAG/3'/EcoR	T352: Anti-sense	tgtgaattcAGGCGGGCCATTTACCGTAAGTTATG
U1-Chromosome X	T485: Sense	ACGACTATCTTATCAATCCTTCCTG
	T486: Anti-sense	CTAGGTGATTAGGATATTCTACAATC
U1-Chromosome 2	T492: Sense	GCTATTGTATCTGATCACAAGCTG
	T493: Anti-sense	TTGATTGTGTGTCCAGGTCCTAGG
d2EGFP	T494: Sense	GCAAGGGCGAGGAGCTGTTCACC
	T495: Anti-sense	TTGTAGTTGCCGTCGTCCTTGAAG
Gag	T457: Sense	AATGGTACATCAGGCCATATCAC
	T458: Anti-sense	CCCACTGTGTTTAGCATGGTATT
Cas9	T477: Sense	CACAGCATCAAGAAGAACCTGAT
	T491: Anti-sense	TCTTCCGTCTGGTGTATCTTCTTC
RRE	Sense	CGCCAAGCTTGAATAGGAGCTTTGTTCC
	Antisense	CTAGGATCCAGGAGCTGTTGATCCTTTAGG
Off-Target (OT)		
LTR-A-OT-1	T465: Sense	GTGGACTTTGGATGGTGAGATAG
	T466: Anti-sense	GCCTGGCAAGAGTGAACTGAGTC
LTR-A-OT-2	T467: Sense	AAGATAATGAGTTGTGGCAGAGC
	T468: Anti-sense	TCTACCTGGTAATCCAGCATCTGG
LTR-A-OT-3	T469: Sense	ATAGGAGGAAGGCACCAAGAGGG
	T470: Anti-sense	AATGATGCTTTGGTCCTACTCCT
LTR-A-OT-4	T471: Sense	TGCTCTTGCTACTCTGGCATGTAC
	T472: Anti-sense	AATCTACCTCTGAGAGCTGCAGG
LTR-A-OT-5	T473: Sense	TCAGACACAGCTGAAGCAGAGGC
	T474: Anti-sense	ATGCCAGTGTCAGTAGATGTCAG
LTR-A-OT-6	T475: Sense	TCAAGATCAGCCAGAGTGCACATG
	T476: Anti-sense	TGCTCTTCCGAGCCTCTCTGGAG
Others		
hU6-sequence	T428: Sense	ATGGACTATCATATGCTTACCG
LSP1	Sense	GCTTCAGCAAGCCGAGTCCTGCGTCGAG
LSP2	Anti-sense	GCTCCTCTGGTTTCCCTTTCGCTTTCAA
AP1	Sense	GTAATACGACTCACTATAGGGC
AP2	Anti-sense	ACTATAGGGCACGCGTGGT

FIGURE 14

Table S3. Locations of predicted gRNA targeting sites of LTR-A and LTR-B

name	query Seq	subject Id	Identity (%)	E-Value	start	end	strand	ref Seq	mismatch (12bp seed)
LTRA	ATCAGATATCCACTGACCTTTGG	HIV	100	7.00E-04	162	184	+	ATCAGATATCCACTGACCTTTGG	0
LTRA	TCAGATATCCACTGACCTTTGG	HIV	100	0.003	163	184	+	TCAGATATCCACTGACCTTTGG	0
LTRA	TCAGATATCCACTGACCTTTGG	HIV	100	0.003	9091	9113	+	TCAGATATCCACTGACCTTTGG	0
LTRA	CAGATATCCACTGACCTTTGG	HIV	100	0.009	164	184	+	CAGATATCCACTGACCTTTGG	0
LTRA	CAGATATCCACTGACCTTTGG	HIV	100	0.009	9092	9112	+	CAGATATCCACTGACCTTTGG	0
LTRA	AGATATCCACTGACCTTTGG	HIV	100	0.033	165	184	+	AGATATCCACTGACCTTTGG	0
LTRA	AGATATCCACTGACCTTTGG	HIV	100	0.033	9093	9112	+	AGATATCCACTGACCTTTGG	0
LTRA	GATATCCACTGACCTTTGG	HIV	100	0.12	166	184	+	GATATCCACTGACCTTTGG	0
LTRA	GATATCCACTGACCTTTGG	HIV	100	0.12	9094	9112	+	GATATCCACTGACCTTTGG	0
LTRA	ATATCCACTGACCTTTGG	HIV	100	0.42	167	184	+	ATATCCACTGACCTTTGG	0
LTRA	ATATCCACTGACCTTTGG	HIV	100	0.42	9095	9112	+	ATATCCACTGACCTTTGG	0
LTRA	TATCCACTGACCTTGGG	chr5	100	1.5	21926317	21926333	+	TATCCACTGACCTTGGG	0
LTRA	TATCCACTGACCTTTGG	HIV	100	1.5	168	184	+	TATCCACTGACCTTTGG	0
LTRA	TATCCACTGACCTTTGG	HIV	100	1.5	9096	9112	+	TATCCACTGACCTTTGG	0
LTRA	TATCCACTGACCTTAAG	chr3	100	1.5	116712577	116712593	+	TATCCACTGACCTTAAG	0
LTRA	TATCCACTGACCTTGAG	chr6	100	1.5	32460607	32460623	+	TATCCACTGACCTTGAG	0
LTRA	ATCCACTGACCTTAGG	chr3	100	5.4	2669092	2669107	+	ATCCACTGACCTTAGG	0
LTRA	ATCCACTGACCTTAGG	chr3	100	5.4	158293369	158293384	+	ATCCACTGACCTTAGG	0
LTRA	ATCCACTGACCTTGGG	chr20	100	5.4	46918344	46918359	+	ATCCACTGACCTTGGG	0
LTRA	ATCCACTGACCTTGGG	chr14	100	5.4	86310067	86310082	-	ATCCACTGACCTTGGG	0
LTRA	ATCCACTGACCTTGGG	chr5	100	5.4	21926318	21926333	+	ATCCACTGACCTTGGG	0
LTRA	ATCCACTGACCTTGGG	chr4	100	5.4	95491921	95491936	-	ATCCACTGACCTTGGG	0
LTRA	ATCCACTGACCTTTGG	HIV	100	5.4	169	184	+	ATCCACTGACCTTTGG	0
LTRA	ATCCACTGACCTTTGG	HIV	100	5.4	9097	9112	+	ATCCACTGACCTTTGG	0
LTRA	ATCCACTGACCTTTGG	chr6	100	5.4	98901053	98901068	+	ATCCACTGACCTTTGG	0
LTRA	ATCCACTGACCTTAAG	chr7	100	5.4	155511293	155511308	-	ATCCACTGACCTTAAG	0
LTRA	ATCCACTGACCTTAAG	chr3	100	5.4	116712578	116712593	+	ATCCACTGACCTTAAG	0
LTRA	ATCCACTGACCTTCAG	chr5	100	5.4	152371289	152371304	+	ATCCACTGACCTTCAG	0
LTRA	ATCCACTGACCTTCAG	chr4	100	5.4	110823169	110823184	-	ATCCACTGACCTTCAG	0
LTRA	ATCCACTGACCTTGAG	chrX	100	5.4	74260260	74260275	+	ATCCACTGACCTTGAG	0
LTRA	ATCCACTGACCTTGAG	chr6	100	5.4	32460608	32460623	+	ATCCACTGACCTTGAG	0
LTRA	TCCACTGACCTTAGG	chr12	100	20	14485012	14485026	-	TCCACTGACCTTAGG	0
LTRA	TCCACTGACCTTAGG	chr7	100	20	72210628	72210642	-	TCCACTGACCTTAGG	0
LTRA	TCCACTGACCTTAGG	chr6	100	20	160845640	160845654	+	TCCACTGACCTTAGG	0
LTRA	TCCACTGACCTTAGG	chr3	100	20	2669093	2669107	+	TCCACTGACCTTAGG	0
LTRA	TCCACTGACCTTAGG	chr3	100	20	158293370	158293384	+	TCCACTGACCTTAGG	0
LTRA	TCCACTGACCTTAGG	chr2	100	20	237551230	237551244	-	TCCACTGACCTTAGG	0
LTRA	TCCACTGACCTTGGG	chr20	100	20	46918345	46918359	+	TCCACTGACCTTGGG	0
LTRA	TCCACTGACCTTGGG	chr14	100	20	86310067	86310081	-	TCCACTGACCTTGGG	0
LTRA	TCCACTGACCTTGGG	chr12	100	20	116054688	116054702	+	TCCACTGACCTTGGG	0
LTRA	TCCACTGACCTTGGG	chr11	100	20	103532094	103532108	+	TCCACTGACCTTGGG	0
LTRA	TCCACTGACCTTGGG	chr10	100	20	132186431	132186445	-	TCCACTGACCTTGGG	0
LTRA	TCCACTGACCTTGGG	chr8	100	20	144600475	144600489	-	TCCACTGACCTTGGG	0
LTRA	TCCACTGACCTTGGG	chr5	100	20	21926319	21926333	+	TCCACTGACCTTGGG	0
LTRA	TCCACTGACCTTGGG	chr4	100	20	95491921	95491935	-	TCCACTGACCTTGGG	0
LTRA	TCCACTGACCTTTGG	HIV	100	20	170	184	+	TCCACTGACCTTTGG	0
LTRA	TCCACTGACCTTTGG	HIV	100	20	9098	9112	+	TCCACTGACCTTTGG	0
LTRA	TCCACTGACCTTTGG	chr16	100	20	86962569	86962583	+	TCCACTGACCTTTGG	0
LTRA	TCCACTGACCTTTGG	chr11	100	20	68156214	68156228	+	TCCACTGACCTTTGG	0
LTRA	TCCACTGACCTTTGG	chr6	100	20	98901054	98901068	+	TCCACTGACCTTTGG	0
LTRA	TCCACTGACCTTTGG	chr5	100	20	72600080	72600094	+	TCCACTGACCTTTGG	0
LTRA	TCCACTGACCTTTGG	chr5	100	20	136458169	136458183	+	TCCACTGACCTTTGG	0
LTRA	TCCACTGACCTTTGG	chr4	100	20	25353030	25353044	+	TCCACTGACCTTTGG	0
LTRA	TCCACTGACCTTTGG	chr2	100	20	207833373	207833387	+	TCCACTGACCTTTGG	0
LTRA	TCCACTGACCTTAAG	chr15	100	20	67850506	67850520	-	TCCACTGACCTTAAG	0
LTRA	TCCACTGACCTTAAG	chr7	100	20	155511293	155511307	-	TCCACTGACCTTAAG	0
LTRA	TCCACTGACCTTAAG	chr5	100	20	25142541	25142555	-	TCCACTGACCTTAAG	0
LTRA	TCCACTGACCTTAAG	chr3	100	20	116712579	116712593	+	TCCACTGACCTTAAG	0
LTRA	TCCACTGACCTTAAG	chr1	100	20	163298514	163298528	+	TCCACTGACCTTAAG	0
LTRA	TCCACTGACCTTCAG	chr20	100	20	22136764	22136778	-	TCCACTGACCTTCAG	0
LTRA	TCCACTGACCTTCAG	chr19	100	20	50519462	50519476	-	TCCACTGACCTTCAG	0
LTRA	TCCACTGACCTTCAG	chr18	100	20	74623621	74623635	-	TCCACTGACCTTCAG	0
LTRA	TCCACTGACCTTCAG	chr16	100	20	71402733	71402747	+	TCCACTGACCTTCAG	0
LTRA	TCCACTGACCTTCAG	chr14	100	20	24193180	24193194	-	TCCACTGACCTTCAG	0
LTRA	TCCACTGACCTTCAG	chr11	100	20	133664063	133664077	+	TCCACTGACCTTCAG	0
LTRA	TCCACTGACCTTCAG	chr9	100	20	140394271	140394285	+	TCCACTGACCTTCAG	0
LTRA	TCCACTGACCTTCAG	chr6	100	20	47685256	47685270	-	TCCACTGACCTTCAG	0
LTRA	TCCACTGACCTTCAG	chr5	100	20	152371290	152371304	+	TCCACTGACCTTCAG	0
LTRA	TCCACTGACCTTCAG	chr4	100	20	110823169	110823183	-	TCCACTGACCTTCAG	0
LTRA	TCCACTGACCTTCAG	chr3	100	20	46255327	46255341	+	TCCACTGACCTTCAG	0
LTRA	TCCACTGACCTTCAG	chr3	100	20	186757301	186757315	+	TCCACTGACCTTCAG	0
LTRA	TCCACTGACCTTGAG	chrX	100	20	74260261	74260275	+	TCCACTGACCTTGAG	0
LTRA	TCCACTGACCTTGAG	chr11	100	20	76052171	76052185	-	TCCACTGACCTTGAG	0
LTRA	TCCACTGACCTTGAG	chr9	100	20	33927660	33927674	-	TCCACTGACCTTGAG	0
LTRA	TCCACTGACCTTGAG	chr9	100	20	71035331	71035345	+	TCCACTGACCTTGAG	0
LTRA	TCCACTGACCTTGAG	chr9	100	20	95871690	95871704	+	TCCACTGACCTTGAG	0
LTRA	TCCACTGACCTTGAG	chr7	100	20	137681847	137681861	+	TCCACTGACCTTGAG	0
LTRA	TCCACTGACCTTGAG	chr6	100	20	32460609	32460623	+	TCCACTGACCTTGAG	0
LTRA	TCCACTGACCTTGAG	chr3	100	20	42344237	42344251	+	TCCACTGACCTTGAG	0
LTRA	TCCACTGACCTTGAG	chr2	100	20	64643586	64643600	+	TCCACTGACCTTGAG	0
LTRA	TCCACTGACCTTTAG	chr16	100	20	55133552	55133566	-	TCCACTGACCTTTAG	0
LTRA	TCCACTGACCTTTAG	chr15	100	20	90072212	90072226	-	TCCACTGACCTTTAG	0
LTRA	TCCACTGACCTTTAG	chr12	100	20	69006300	69006314	+	TCCACTGACCTTTAG	0
LTRA	TCCACTGACCTTTAG	chr3	100	20	170680338	170680352	-	TCCACTGACCTTTAG	0
LTRA	TCCACTGACCTTTAG	chr2	100	20	215414950	215414964	-	TCCACTGACCTTTAG	0

FIGURE 15

name	query Seq	subject Id	Identity (%)	E-Value	start	end	strand	ref Seq	mismatch (12bp seed)
LTRB	CAGCAGTTCTTGAAGTACTCCGG	HIV	100	7.00E-04	9291	9313	-	CAGCAGTTCTTGAAGTACTCCGG	0
LTRB	AGCAGTTCTTGAAGTACTCCGG	HIV	100	0.003	9291	9312	-	AGCAGTTCTTGAAGTACTCCGG	0
LTRB	GCAGTTCTTGAAGTACTCCGG	HIV	100	0.009	9291	9311	-	GCAGTTCTTGAAGTACTCCGG	0
LTRB	CAGTTCTTGAAGTACTCCGG	HIV	100	0.033	9291	9310	-	CAGTTCTTGAAGTACTCCGG	0
LTRB	AGTTCTTGAAGTACTCCGG	HIV	100	0.12	9291	9309	-	AGTTCTTGAAGTACTCCGG	0
LTRB	GTTCTTGAAGTACTCCGG	HIV	100	0.42	9291	9308	-	GTTCTTGAAGTACTCCGG	0
LTRB	TTCTTGAAGTACTCCGG	HIV	100	1.5	9291	9307	-	TTCTTGAAGTACTCCGG	0
LTRB	TCTTGAAGTACTCCGG	HIV	100	5.4	9291	9306	-	TCTTGAAGTACTCCGG	0
LTRB	TCTTGAAGTACTCTAG	chr11	100	5.4	91845834	91845849	-	TCTTGAAGTACTCTAG	0
LTRB	CTTGAAGTACTCAGG	chr19	100	20	45672789	45672803	-	CTTGAAGTACTCAGG	0
LTRB	CTTGAAGTACTCAGG	chr15	100	20	82132445	82132459	+	CTTGAAGTACTCAGG	0
LTRB	CTTGAAGTACTCAGG	chr11	100	20	94282411	94282425	+	CTTGAAGTACTCAGG	0
LTRB	CTTGAAGTACTCAGG	chr2	100	20	193312744	193312758	-	CTTGAAGTACTCAGG	0
LTRB	CTTGAAGTACTCCGG	HIV	100	20	9291	9305	-	CTTGAAGTACTCCGG	0
LTRB	CTTGAAGTACTCTGG	chr15	100	20	61274973	61274987	-	CTTGAAGTACTCTGG	0
LTRB	CTTGAAGTACTCAAG	chrX	100	20	36051764	36051778	-	CTTGAAGTACTCAAG	0
LTRB	CTTGAAGTACTCAAG	chr15	100	20	31316465	31316479	-	CTTGAAGTACTCAAG	0
LTRB	CTTGAAGTACTCAAG	chr13	100	20	23054474	23054488	-	CTTGAAGTACTCAAG	0
LTRB	CTTGAAGTACTCAAG	chr9	100	20	83208046	83208060	+	CTTGAAGTACTCAAG	0
LTRB	CTTGAAGTACTCAAG	chr8	100	20	13956368	13956382	+	CTTGAAGTACTCAAG	0
LTRB	CTTGAAGTACTCCAG	chr16	100	20	57449025	57449039	-	CTTGAAGTACTCCAG	0
LTRB	CTTGAAGTACTCCAG	chr15	100	20	41397831	41397845	-	CTTGAAGTACTCCAG	0
LTRB	CTTGAAGTACTCCAG	chr11	100	20	70255488	70255502	-	CTTGAAGTACTCCAG	0
LTRB	CTTGAAGTACTCCAG	chr3	100	20	134149643	134149657	+	CTTGAAGTACTCCAG	0
LTRB	CTTGAAGTACTCTAG	chr11	100	20	91845834	91845848	-	CTTGAAGTACTCTAG	0
LTRB	CTTGAAGTACTCTAG	chr1	100	20	224520600	224520614	+	CTTGAAGTACTCTAG	0

FIGURE 15 continued

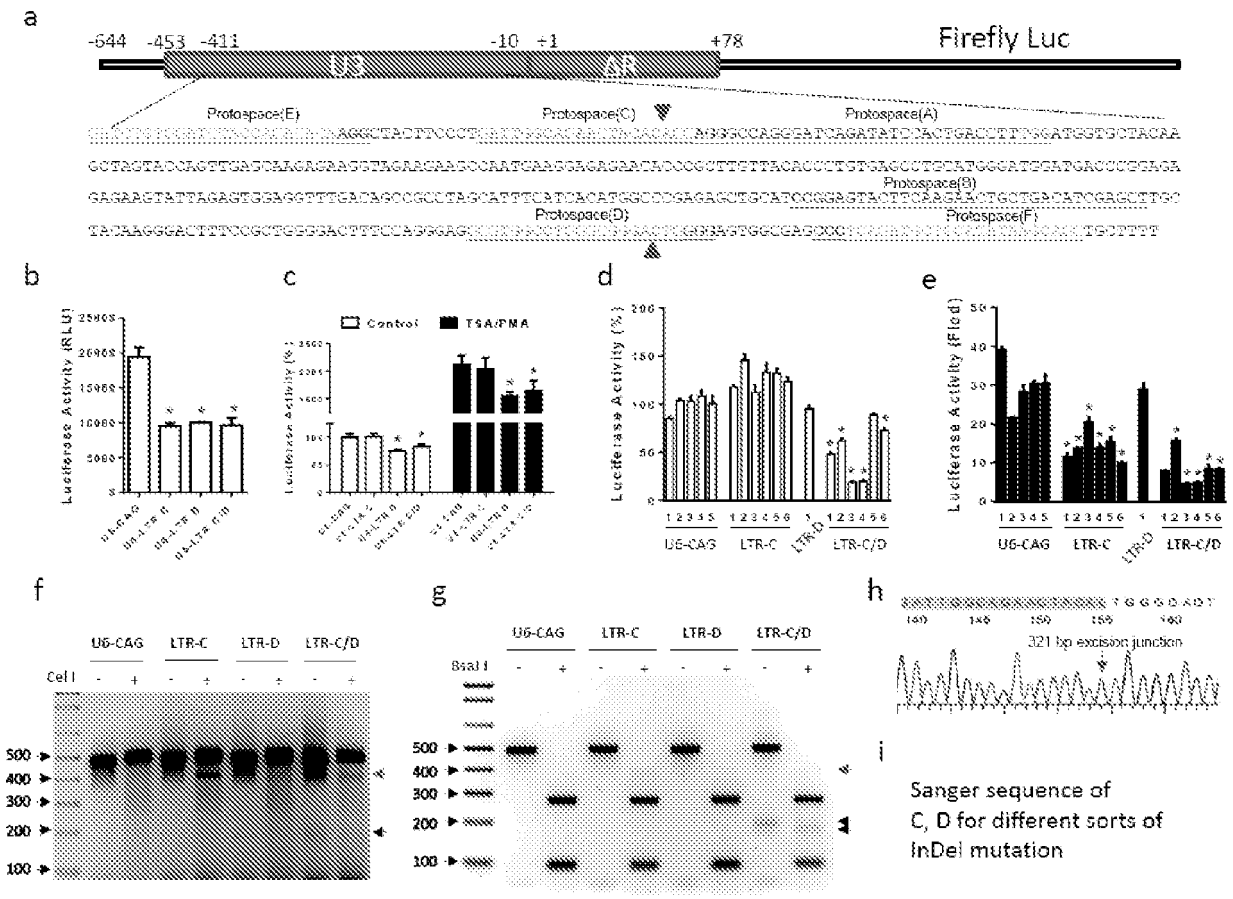

a
-644 -453 -411
-10 +1
+78
Firefly Luc
Protospace(E)
Protospace(C)
Protospace(A)
Protospace(B)
Protospace(D)
Protospace(F)
b
c
Control
TSA/5AZA
d
e
U6-CAG
LTR-C
LTR-D
LTR-C/D
f
g
Cell
BssH I
500
400
300
200
100
h
321 bp excision junction
i
Sanger sequence of
C, D for different sorts of
InDel mutation

FIGURE 16

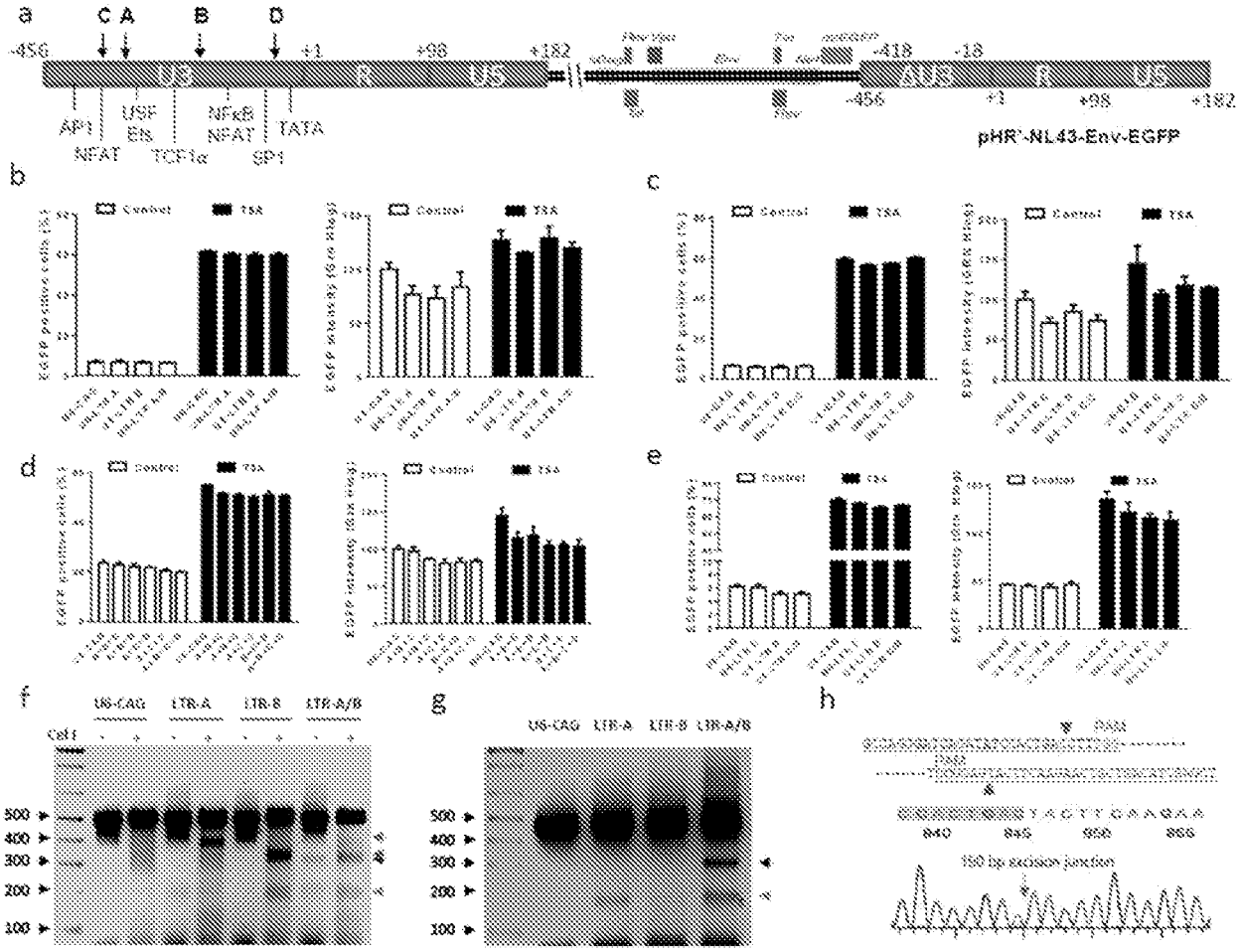

a
c A
B
D
-456
+1
+98
+182
-418
-18
U3
R
U5
AUG
AP1
NFAT
USF
Ets
TCF1α
NFκB
NFAT
SP1
TATA
pHR'-NL43-Env-EGFP
b
c
d
e
Control
TSA
f
US-CAG
LTR-A
LTR-B
LTR-A/B
Cell
500
400
300
200
100
g
US-CAG
LTR-A
LTR-B
LTR-A/B
h
DAS
DAS
840
845
850
855
150 bp excision junction

FIGURE 17

HIV-1 LTR (644 bp)

Human immunodeficiency virus type 1, NY5/BRU (LAV-1) recombinant clone pNL4-3.
ACCESSION number: M19921

TGGAAGGGCTAATTCACTCCCAACGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAAGGCTACT

TCCCTGATTGGCAGAACTACACACCAGGGCCAGGGATCAGATATCCACTGACCTTTGGATGGTGCTACAAG

U3 region

CTAGTACCAGTTGAGCAAGAGAAGGTAGAAGAAGCCAATGAAGGAGAGAACACCCGCTTGTTACACCCTG

TGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTATTAGAGTGGAGGTTTGACAGCCGCCTAGCATT

TCATCACATGGCCCGAGAGCTGCATCCGGAGTACTTCAAGAACTGCTGACATCGAGCTTGCTACAAGGGAC

TTTCCGCTGGGGACTTTCCAGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATGCT

R region

GCATATAAGCAGCTGCTTTTTGCTTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCT

GGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCG

U5 region

TCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCA

FIGURE 18

SIVmm239-LTR(818bp)

Simian (macaque) immunodeficiency virus, isolate 239
Genebank number: M33262

TGGAAGGGATTTATTACAGTGCAAGAAGACATAGAATCTTAGACATATACTTAGAAAAGGAAGAAGGCATCAT

ACCAGATTGGCAGGATTACACCTCAGGACCAGGAATTAGATACCCAAAGACATTTGGCTGGCTATGGAAATTA

U3 region

GTCCCTGTAAATGTATCAGATGAGGCACAGGAGGATGAGGAGCATTATTTAATGCATCCAGCTCAAACTTCCC

AGTGGGATGACCCTTGGGGAGAGGTTCTAGCATGGAAGTTTGATCCAACTCTGGCCTACACTTATGAGGCATA

TGTTAGATACCCAGAAGAGTTTGGAAGCAAGTCAGGCCTGTCAGAGGAAGAGGTTAGAAGAAGGCTAACCGCA

AGAGGCCTTCTTAACATGGCTGACAAGAAGGAAACTCGCTGAAACAGCAGGGACTTTCCACAAGGGGATGTTA

CGGGGAGGTACTGGGGAGGAGCCGGTCGGGAACGCCCACTTTCTTGATGTATAAATATCACTGCATTTCGCTC

TGTATTCAGTCGCTCTGCGGAGAGGCTGGCAGATTGAGCCCTGGGAGGTTCTCTCCAGCACTAGCAGGTAGAG

R region

CCTGGGTGTTCCCTGCTAGACTCTCACCAGCACTTGGCCGGTGCTGGGCAGAGTGACTCCACGCTTGCTTGCT

TAAAGCCCTCTTCAATAAAGCTGCCATTTTAGAAGTAAGCTAGTGTGTGTTCCCATCTCTCCTAGCCGCCGCC

U5 region

TGGTCAACTCGGTACTCAATAATAAGAAGACCCTGGTCTGTTAGGACCCTTTCTGCTTTGGGAAACCGAAGCA

GGAAAATCCCTAGCA

FIGURE 19