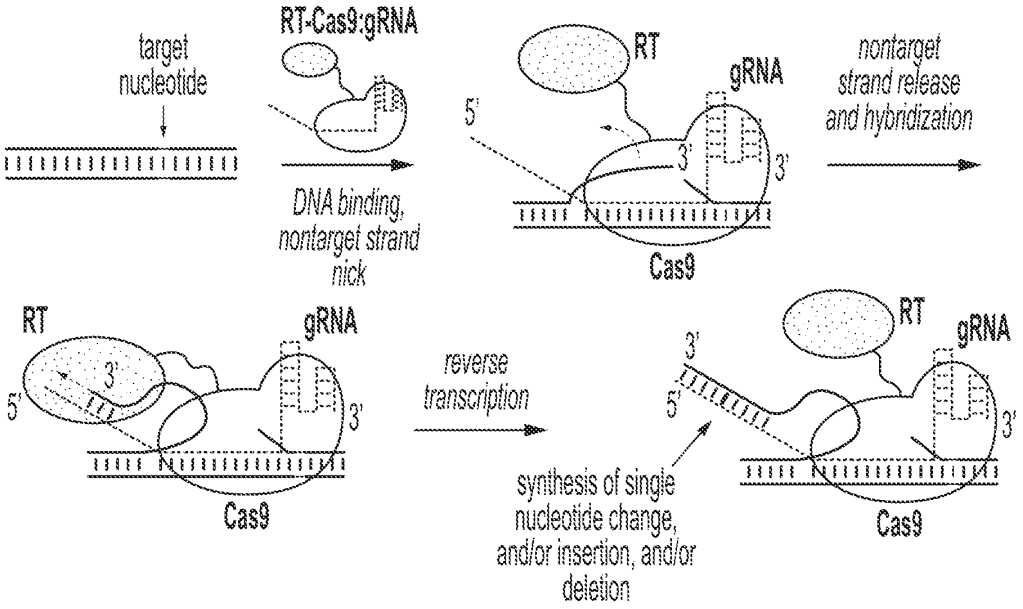

target nucleotide
RT-Cas9:gRNA
DNA binding, nontarget strand nick
RT
gRNA
5'
3'
3'
Cas9
nontarget strand release and hybridization
RT
gRNA
5'
3'
3'
Cas9
reverse transcription
3'
5'
RT
gRNA
3'
Cas9
synthesis of single nucleotide change, and/or insertion, and/or deletion

FIG. 1A

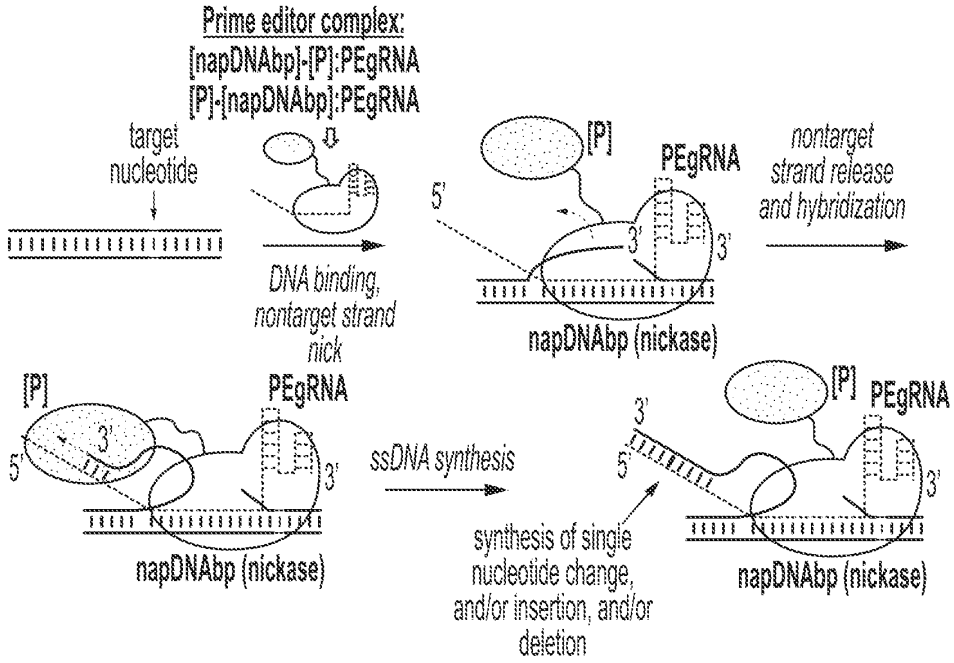

Prime editor complex:
[napDNAbp]-[P]:PEgRNA
[P]-[napDNAbp]:PEgRNA
target nucleotide
DNA binding, nontarget strand nick
[P]
PEgRNA
5'
3'
3'
napDNAbp (nickase)
nontarget strand release and hybridization
[P]
PEgRNA
5'
3'
3'
napDNAbp (nickase)
ssDNA synthesis
3'
5'
[P]
PEgRNA
3'
napDNAbp (nickase)
synthesis of single nucleotide change, and/or insertion, and/or deletion

FIG. 1B

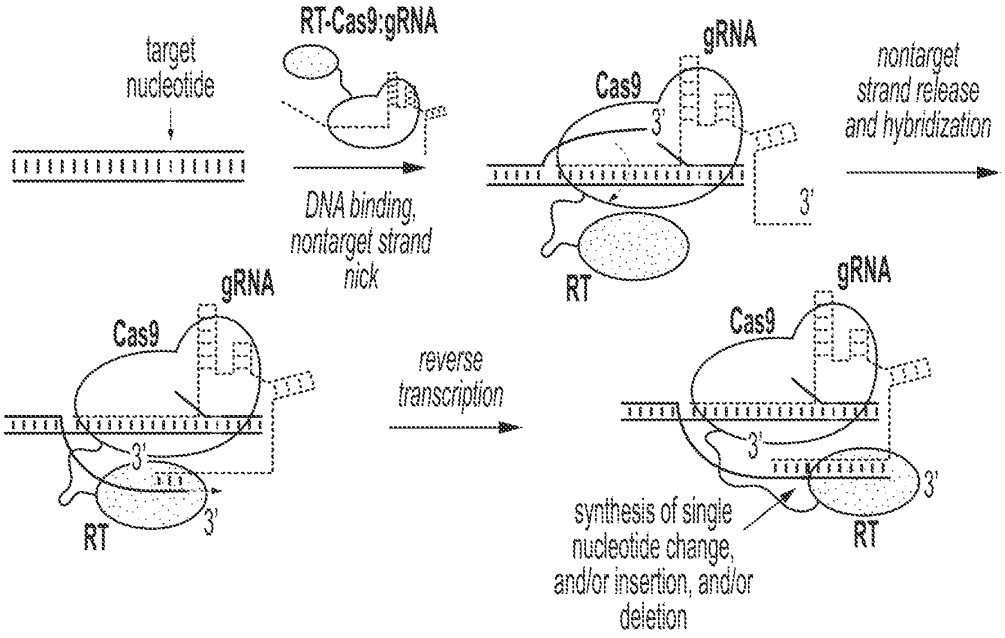

target nucleotide
RT-Cas9:gRNA
DNA binding, nontarget strand nick
gRNA
Cas9
3'
RT
nontarget strand release and hybridization
reverse transcription
synthesis of single nucleotide change, and/or insertion, and/or deletion

FIG. 1C

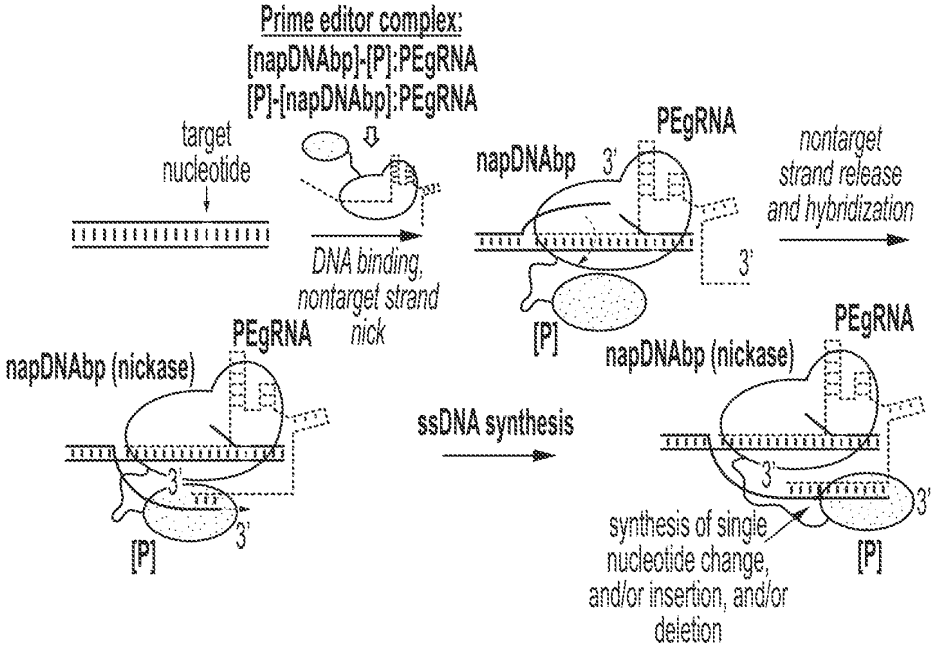

Prime editor complex:
[napDNAbp]-[P]:PEgRNA
[P]-[napDNAbp]:PEgRNA
target nucleotide
DNA binding, nontarget strand nick
PEgRNA
napDNAbp
3'
[P]
nontarget strand release and hybridization
napDNAbp (nickase)
ssDNA synthesis
synthesis of single nucleotide change, and/or insertion, and/or deletion

FIG. 1D

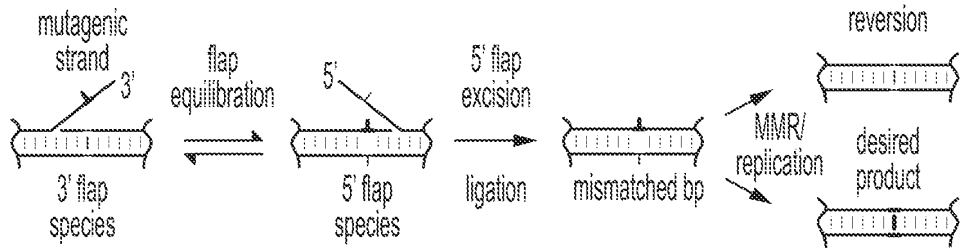

mutagenic strand
3'
3' flap species
flap equilibration
5'
5' flap species
5' flap excision
ligation
mismatched bp
MMR/ replication
reversion
desired product

FIG. 1E

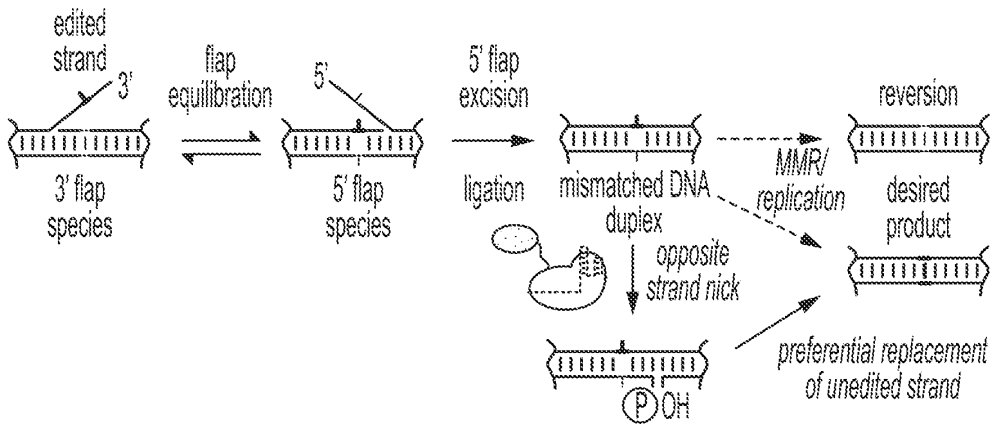

edited strand
3'
3' flap species
flap equilibration
5'
5' flap species
5' flap excision
ligation
mismatched DNA duplex
opposite strand nick
MMR/ replication
reversion
desired product
preferential replacement of unedited strand
P
OH

FIG. 1F

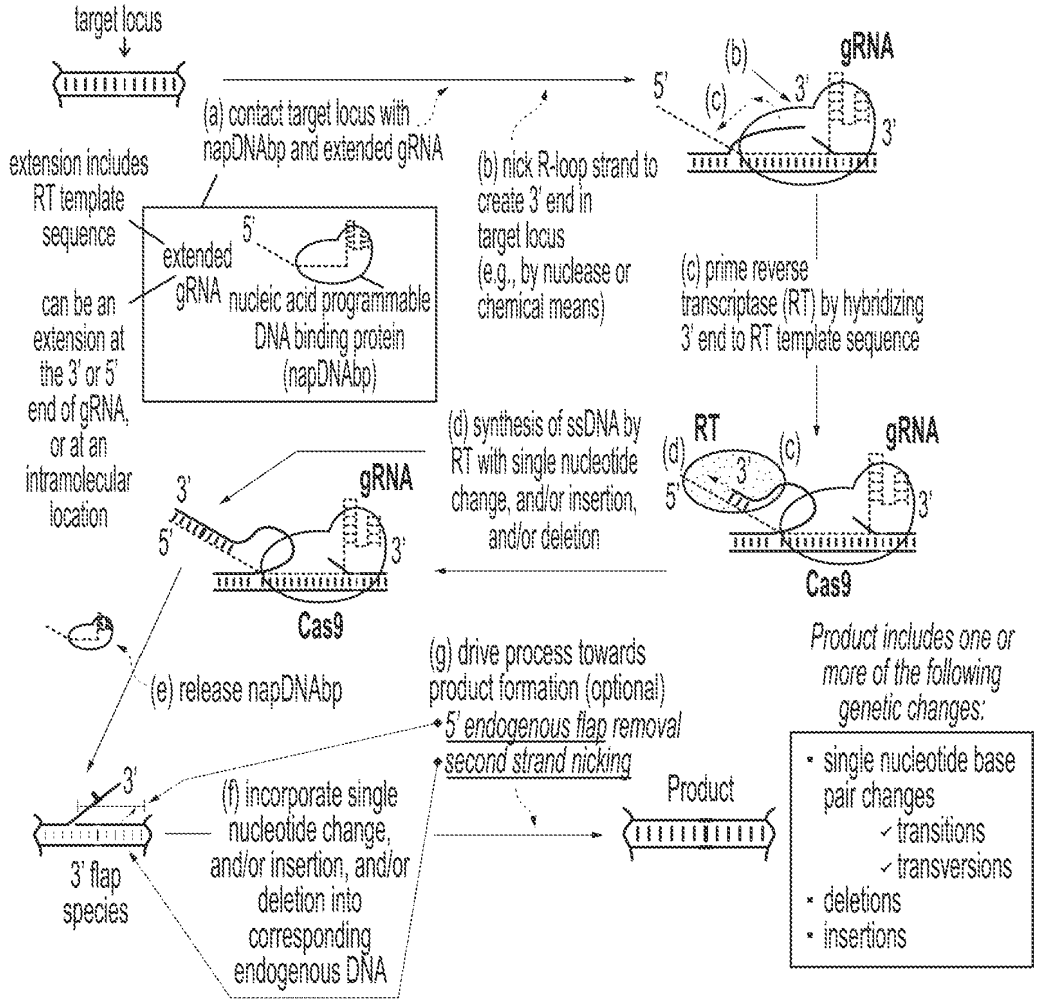

target locus
(a) contact target locus with napDNAbp and extended gRNA
extension includes RT template sequence
can be an extension at the 3' or 5' end of gRNA, or at an intramolecular location
extended gRNA
nucleic acid programmable DNA binding protein (napDNAbp)
(b) nick R-loop strand to create 3' end in target locus (e.g., by nuclease or chemical means)
gRNA
(c) prime reverse transcriptase (RT) by hybridizing 3' end to RT template sequence
RT
(d) synthesis of ssDNA by RT with single nucleotide change, and/or insertion, and/or deletion
Cas9
(e) release napDNAbp
3' flap species
(f) incorporate single nucleotide change, and/or insertion, and/or deletion into corresponding endogenous DNA
(g) drive process towards product formation (optional)
5' endogenous flap removal
second strand nicking
Product
Product includes one or more of the following genetic changes:
single nucleotide base pair changes
transitions
transversions
deletions
insertions

FIG. 1G

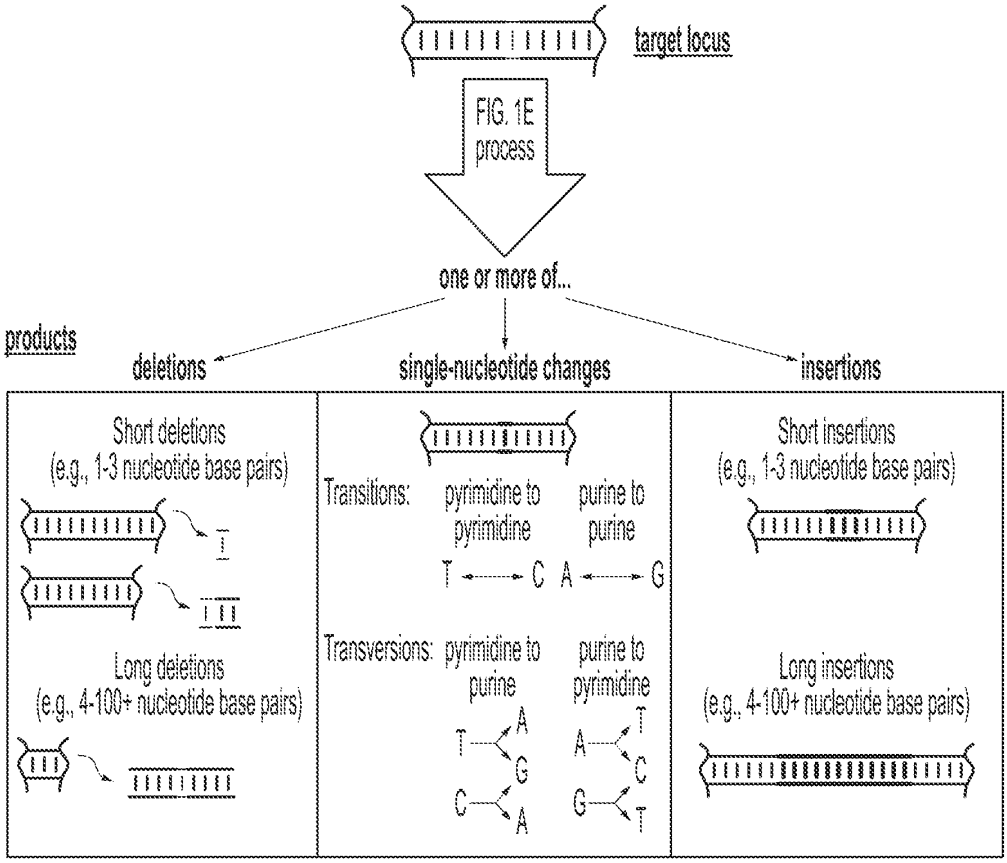

target locus
FIG. 1E process
one or more of...
products
deletions
single-nucleotide changes
insertions
Short deletions
(e.g., 1-3 nucleotide base pairs)
Long deletions
(e.g., 4-100+ nucleotide base pairs)
Transitions:
pyrimidine to pyrimidine
purine to purine
T ⟷ C A ⟷ G
Transversions:
pyrimidine to purine
purine to pyrimidine
T A G C A
A T C G T
Short insertions
(e.g., 1-3 nucleotide base pairs)
Long insertions
(e.g., 4-100+ nucleotide base pairs)

FIG. 1H

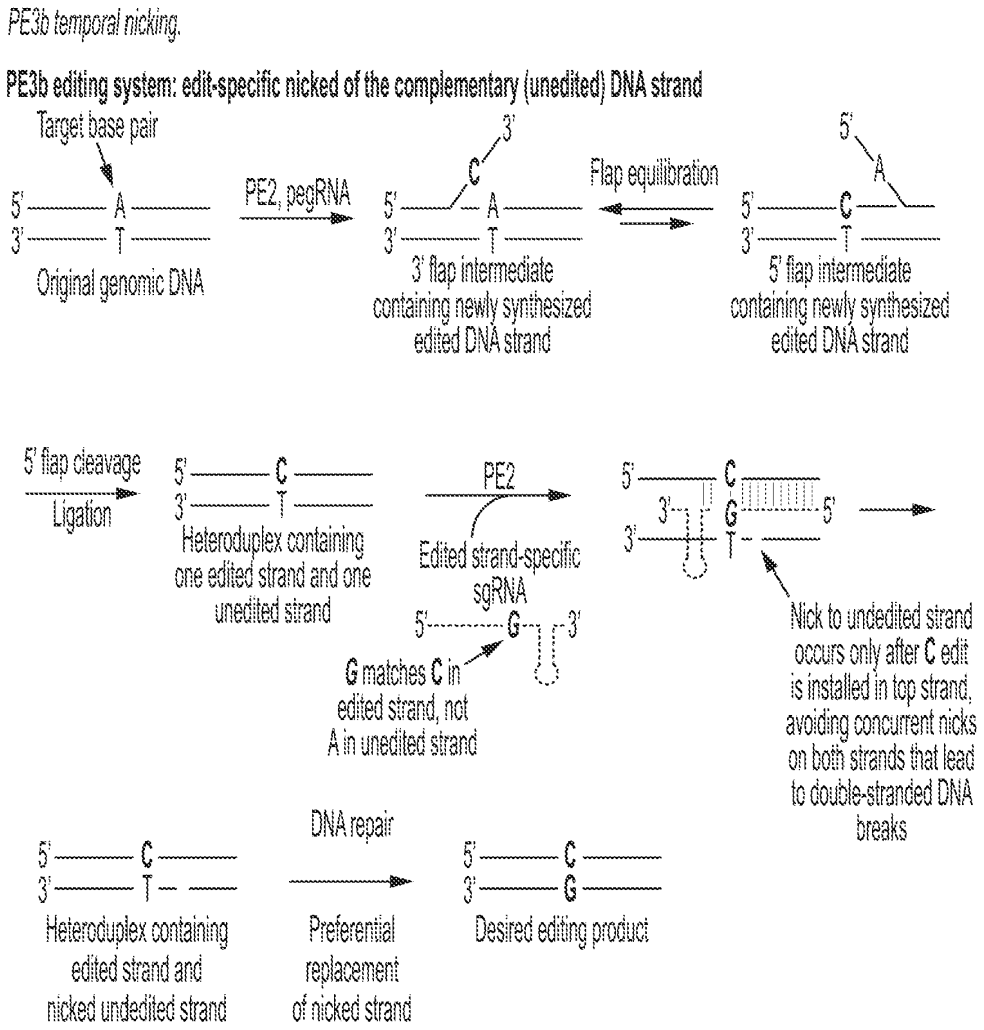

PE3b temporal nicking.
PE3b editing system: edit-specific nicked of the complementary (unedited) DNA strand
Target base pair
Original genomic DNA
PE2, pegRNA
3' flap intermediate containing newly synthesized edited DNA strand
Flap equilibration
5' flap intermediate containing newly synthesized edited DNA strand
5' flap cleavage
Ligation
Heteroduplex containing one edited strand and one unedited strand
PE2
Edited strand-specific sgRNA
G matches C in edited strand, not A in unedited strand
Nick to unedited strand occurs only after C edit is installed in top strand, avoiding concurrent nicks on both strands that lead to double-stranded DNA breaks
DNA repair
Heteroduplex containing edited strand and nicked unedited strand
Preferential replacement of nicked strand
Desired editing product

FIG. 1I

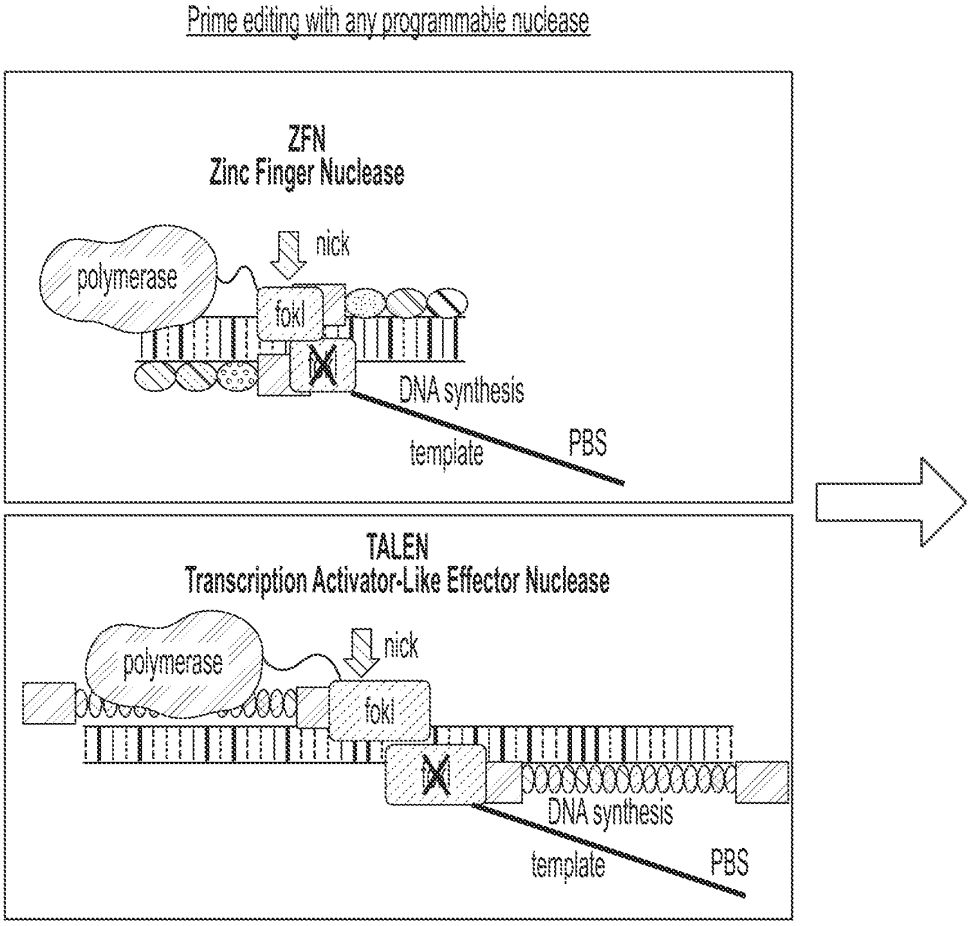

Prime editing with any programmable nuclease
ZFN
Zinc Finger Nuclease
polymerase
nick
fokI
DNA synthesis
template
PBS
TALEN
Transcription Activator-Like Effector Nuclease
polymerase
nick
fokI
DNA synthesis
template
PBS

FIG. 1J

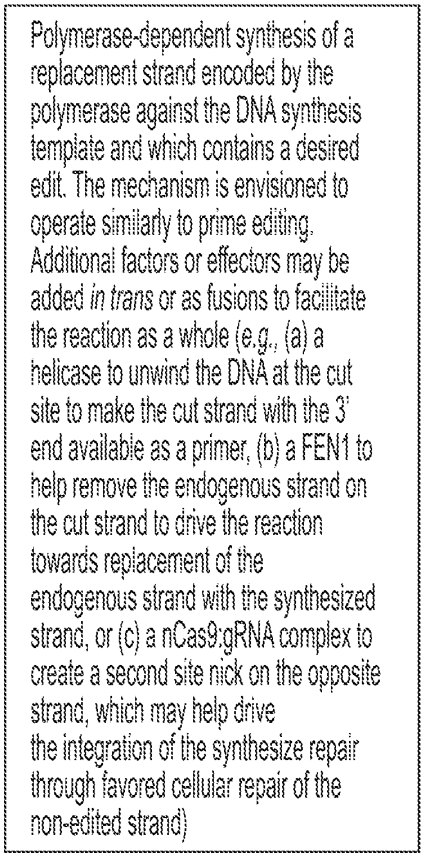

Polymerase-dependent synthesis of a replacement strand encoded by the polymerase against the DNA synthesis template and which contains a desired edit. The mechanism is envisioned to operate similarly to prime editing. Additional factors or effectors may be added *in trans* or as fusions to facilitate the reaction as a whole (*e.g.*, (a) a helicase to unwind the DNA at the cut site to make the cut strand with the 3' end available as a primer, (b) a FEN1 to help remove the endogenous strand on the cut strand to drive the reaction towards replacement of the endogenous strand with the synthesized strand, or (c) a nCas9:gRNA complex to create a second site nick on the opposite strand, which may help drive the integration of the synthesize repair through favored cellular repair of the non-edited strand)

FIG. 1K

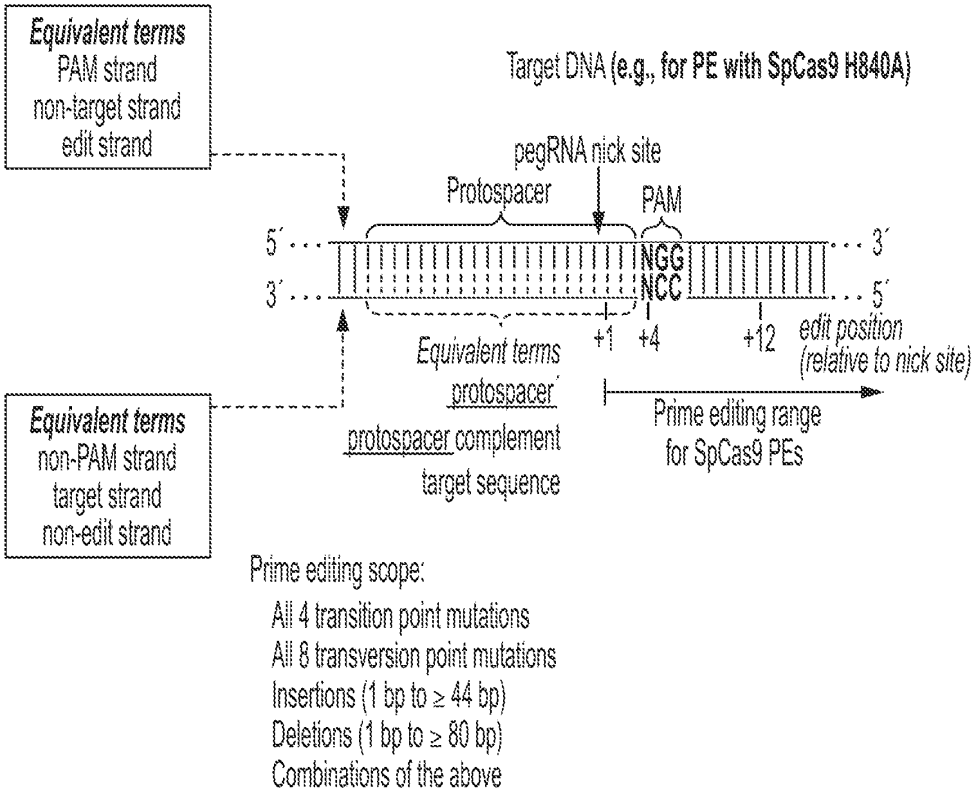

Equivalent terms
PAM strand
non-target strand
edit strand
Target DNA (e.g., for PE with SpCas9 H840A)
pegRNA nick site
Protospacer
PAM
5´
3´
NGG
NCC
3´
5´
+1
+4
+12
edit position
(relative to nick site)
Equivalent terms
protospacer
protospacer complement
target sequence
Prime editing range
for SpCas9 PEs
Equivalent terms
non-PAM strand
target strand
non-edit strand
Prime editing scope:
All 4 transition point mutations
All 8 transversion point mutations
Insertions (1 bp to ≥ 44 bp)
Deletions (1 bp to ≥ 80 bp)
Combinations of the above

FIG. 1L

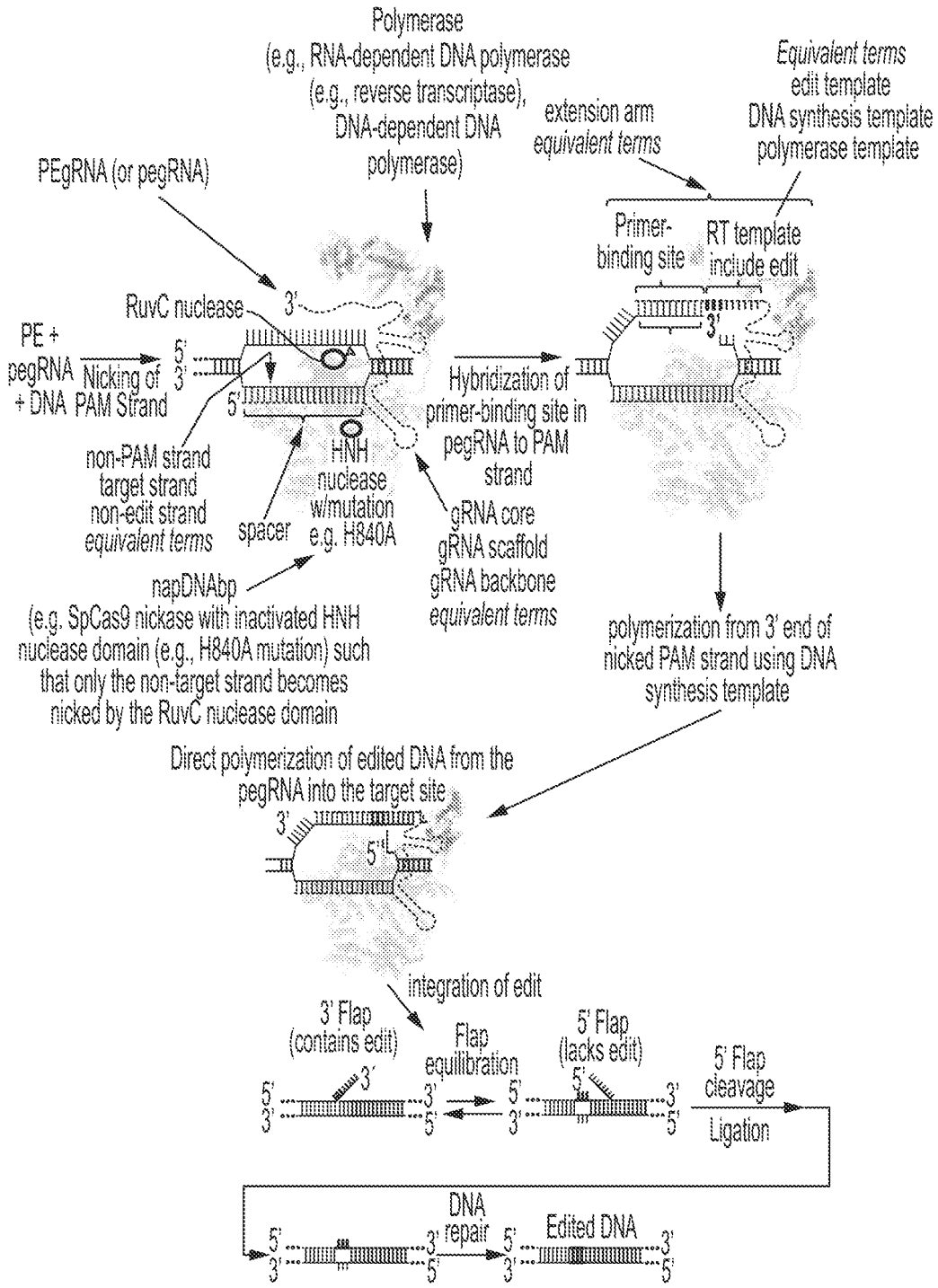

Polymerase
(e.g., RNA-dependent DNA polymerase
(e.g., reverse transcriptase),
DNA-dependent DNA
polymerase)
extension arm
equivalent terms
Equivalent terms
edit template
DNA synthesis template
polymerase template
PEgRNA (or pegRNA)
Primer-
binding site
RT template
include edit
RuvC nuclease
3'
PE +
pegRNA
+ DNA
Nicking of
PAM Strand
5'
3'
5'
Hybridization of
primer-binding site in
pegRNA to PAM
strand
non-PAM strand
target strand
non-edit strand
equivalent terms
spacer
HNH
nuclease
w/mutation
e.g. H840A
gRNA core
gRNA scaffold
gRNA backbone
equivalent terms
napDNAbp
(e.g. SpCas9 nickase with inactivated HNH
nuclease domain (e.g., H840A mutation) such
that only the non-target strand becomes
nicked by the RuvC nuclease domain
polymerization from 3' end of
nicked PAM strand using DNA
synthesis template
Direct polymerization of edited DNA from the
pegRNA into the target site
integration of edit
3' Flap
(contains edit)
Flap
equilibration
5' Flap
(lacks edit)
5' Flap
cleavage
Ligation
DNA
repair
Edited DNA

FIG. 1M

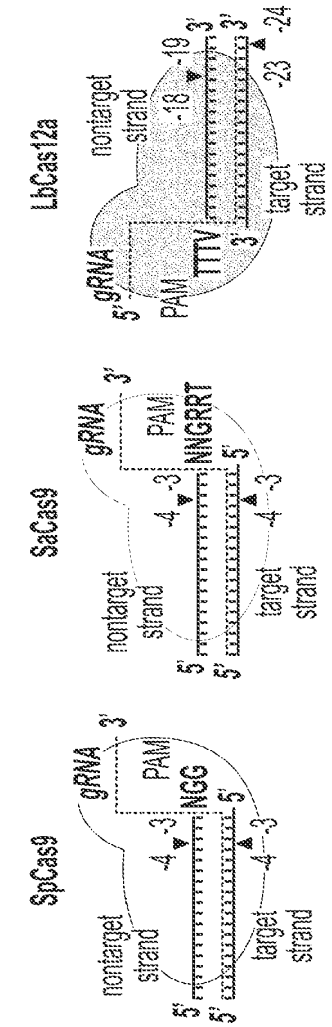

SpCas9
gRNA
3'
PAM
NGG
5'
-4
-3
nontarget strand
target strand
SaCas9
gRNA
3'
PAM
NNGRRT
LbCas12a
nontarget strand
-18
-19
3'
-23
-24
gRNA
5'
PAM
TTTV
target strand

FIG. 2

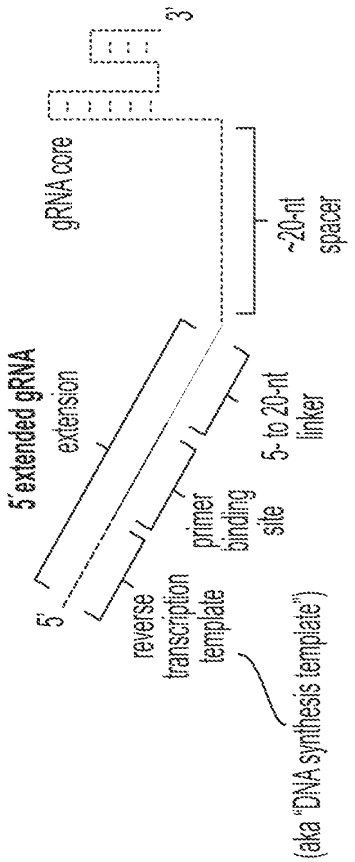

5' extended gRNA
extension
gRNA core
3'
5'
reverse transcription template
primer binding site
5- to 20-nt linker
~20-nt spacer
(aka "DNA synthesis template")

FIG. 3A

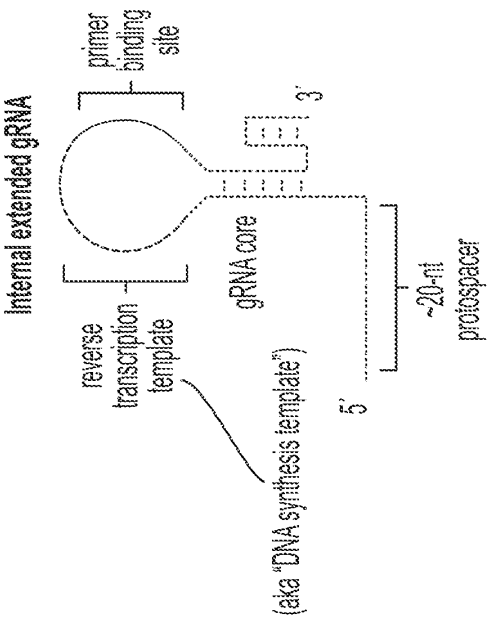

Internal extended gRNA
primer binding site
3'
gRNA core
reverse transcription template
(aka "DNA synthesis template")
5'
~20-nt protospacer

FIG. 3C

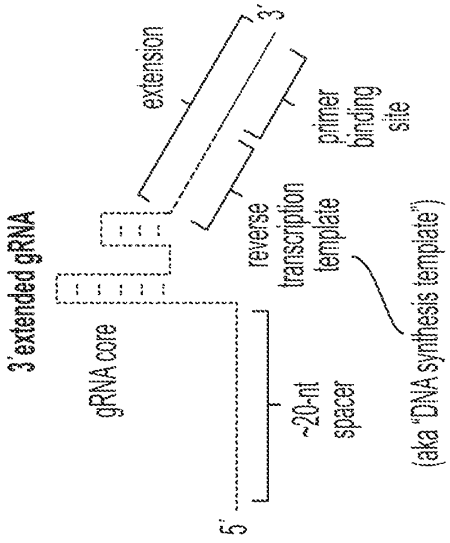

3' extended gRNA
extension
3'
gRNA core
reverse transcription template
primer binding site
5'
~20-nt spacer
(aka "DNA synthesis template")

FIG. 3B

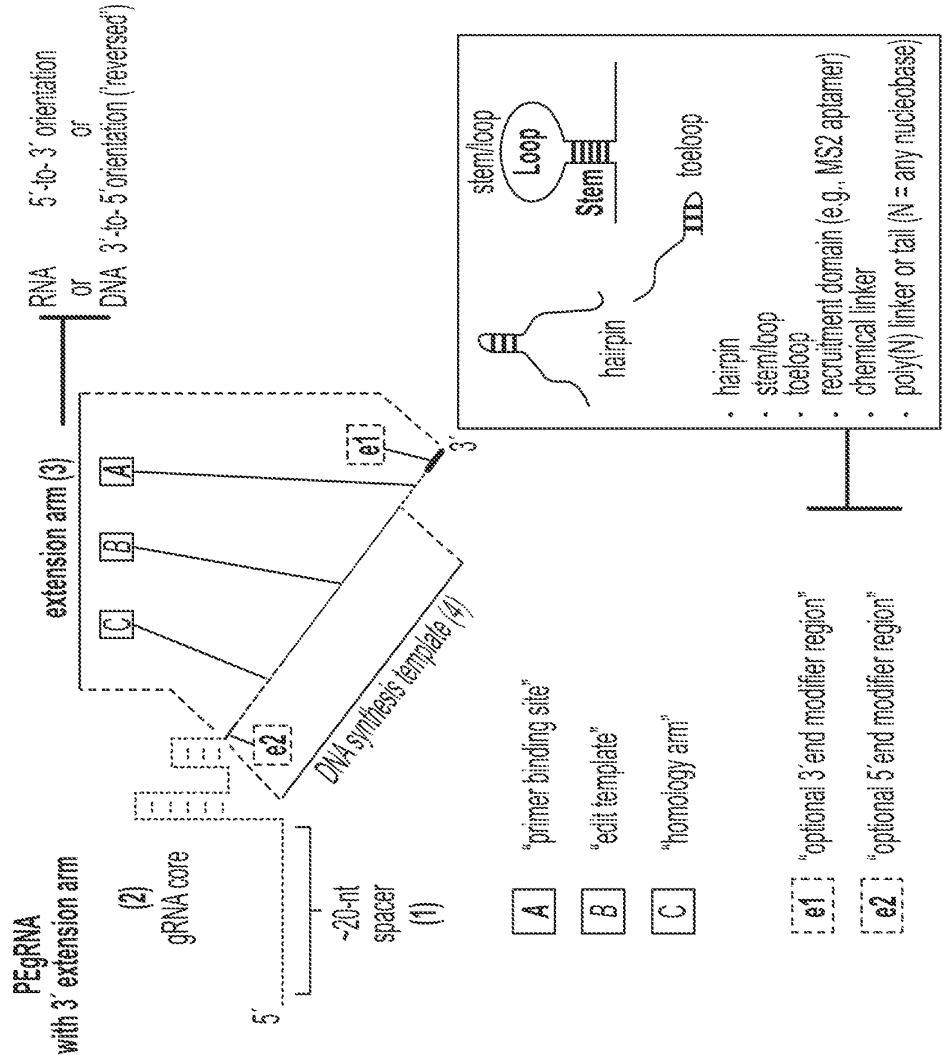

PEgRNA
with 3′ extension arm
(2)
gRNA core
~20-nt
spacer
(1)
5′
e2
DNA synthesis template (4)
extension arm (3)
A
B
C
e1
3′
RNA
or
DNA
5′-to-3′ orientation
or
3′-to-5′ orientation ("reversed")
stem/loop
Loop
Stem
hairpin
toeloop
• hairpin
• stem/loop
• toeloop
• recruitment domain (e.g., MS2 aptamer)
• chemical linker
• poly(N) linker or tail (N = any nucleobase)
A "primer binding site"
B "edit template"
C "homology arm"
e1 "optional 3′ end modifier region"
e2 "optional 5′ end modifier region"

FIG. 3D

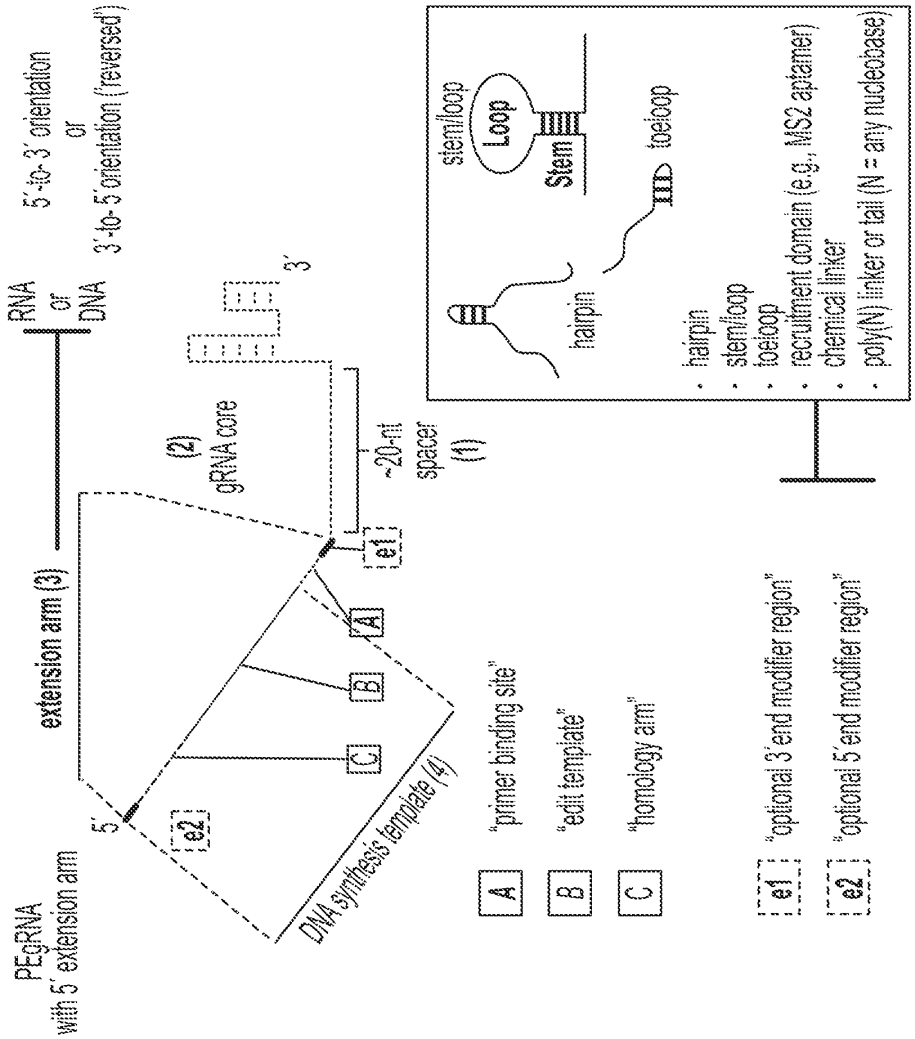

PEgRNA
with 5' extension arm
5'
e2
extension arm (3)
RNA
or
DNA
5'-to-3' orientation
or
3'-to-5' orientation ("reversed")
(2)
gRNA core
3'
~20-nt
spacer
(1)
e1
A
B
C
DNA synthesis template (4)
stem/loop
Loop
Stem
hairpin
toeloop
· hairpin
· stem/loop
· toeloop
· recruitment domain (e.g., MS2 aptamer)
· chemical linker
· poly(N) linker or tail (N = any nucleobase)
A "primer binding site"
B "edit template"
C "homology arm"
e1 "optional 3' end modifier region"
e2 "optional 5' end modifier region"

FIG. 3E

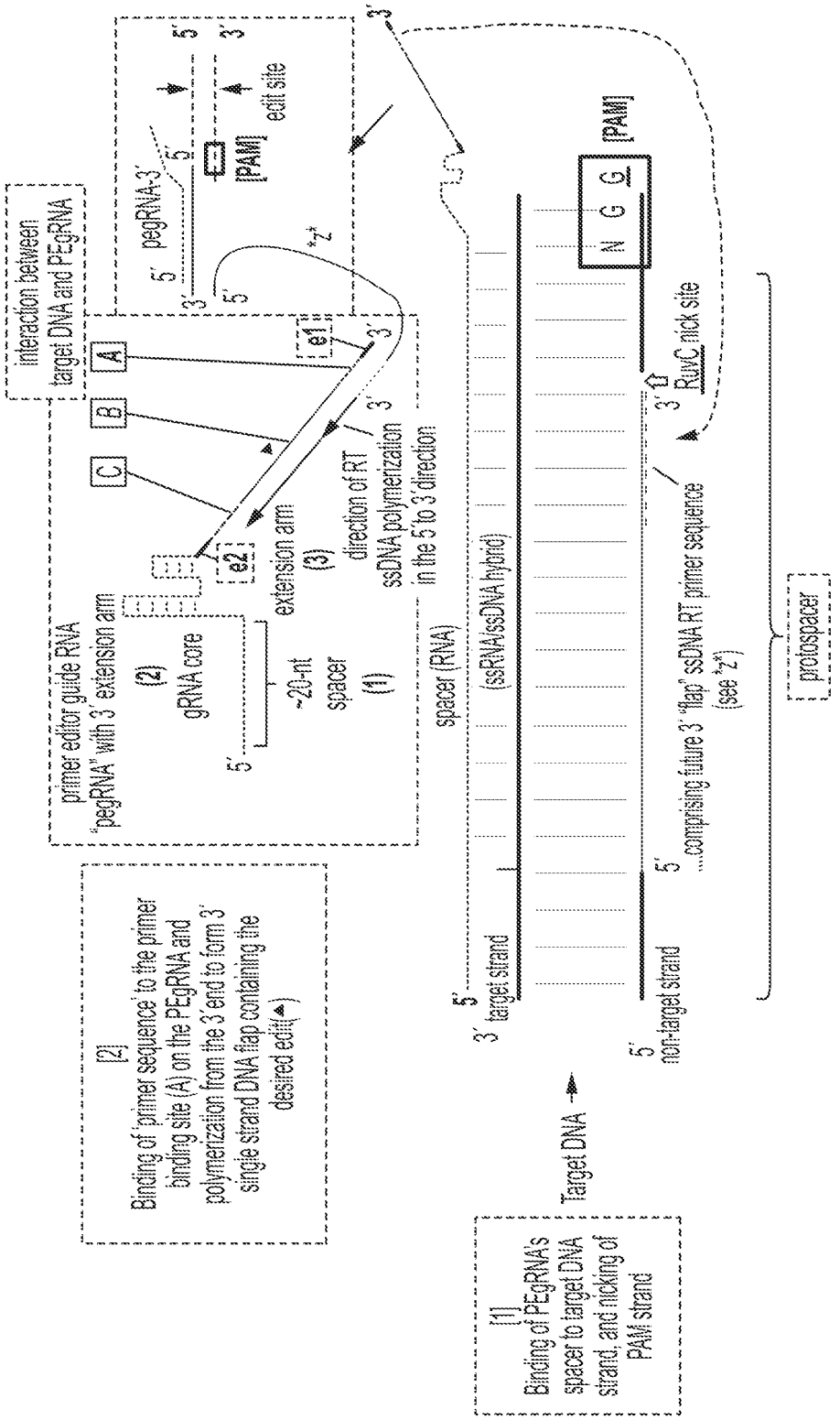

interaction between target DNA and PEgRNA
primer editor guide RNA "pegRNA" with 3' extension arm
(2) gRNA core
~20-nt spacer (1)
extension arm (3)
direction of RT ssDNA polymerization in the 5' to 3' direction
spacer (RNA)
(ssRNA/ssDNA hybrid)
3' target strand
5'
5' non-target strand
3'
...comprising future 3' "flap" ssDNA RT primer sequence (see "z")
RuvC nick site
[PAM]
N G G
edit site
pegRNA-3'
"z"
e1
e2
A
B
C
protospacer
[2] Binding of 'primer sequence' to the primer binding site (A) on the PEgRNA and polymerization from the 3' end to form 3' single strand DNA flap containing the desired edit(★)
[1] Binding of PEgRNA's spacer to target DNA strand, and nicking of PAM strand
Target DNA

FIG. 3F

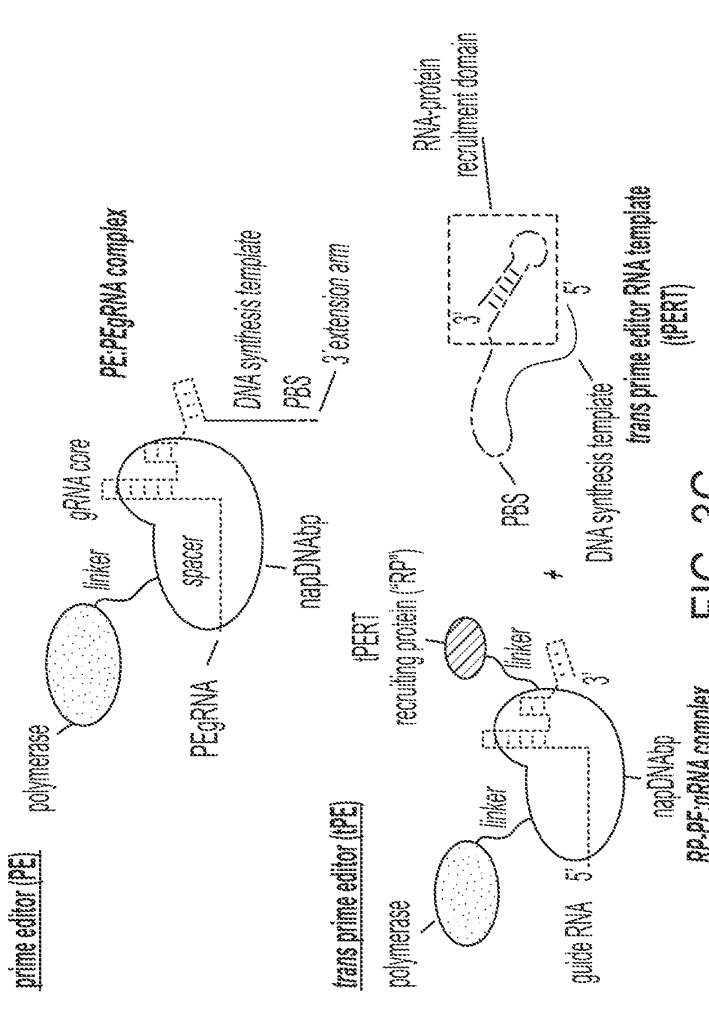

prime editor (PE)
polymerase
linker
gRNA core
spacer
napDNAbp
PEgRNA
PE:PEgRNA complex
DNA synthesis template
PBS
3' extension arm
trans prime editor (tPE)
polymerase
linker
guide RNA
5'
3'
napDNAbp
RP:PEgRNA complex
tPERT
recruiting protein ("RP")
linker
RNA-protein recruitment domain
3'
5'
PBS
DNA synthesis template
trans prime editor RNA template (tPERT)

FIG. 3G

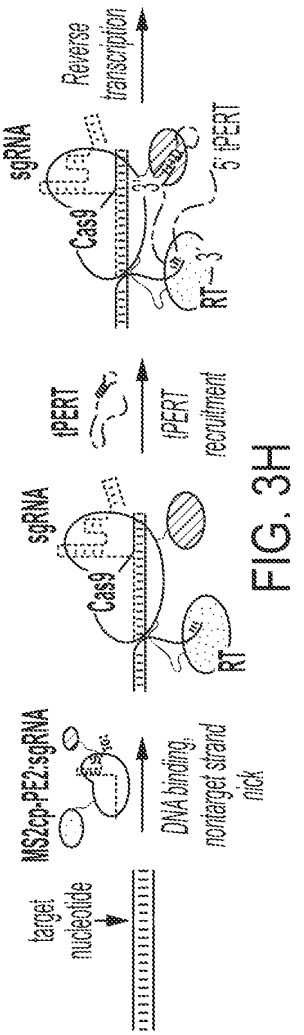

target nucleotide
MS2cp-PE2:sgRNA
DNA binding, nontarget strand nick
sgRNA
Cas9
RT
tPERT
tPERT recruitment
sgRNA
Cas9
Reverse transcription
RT
3'
5' tPERT

FIG. 3H

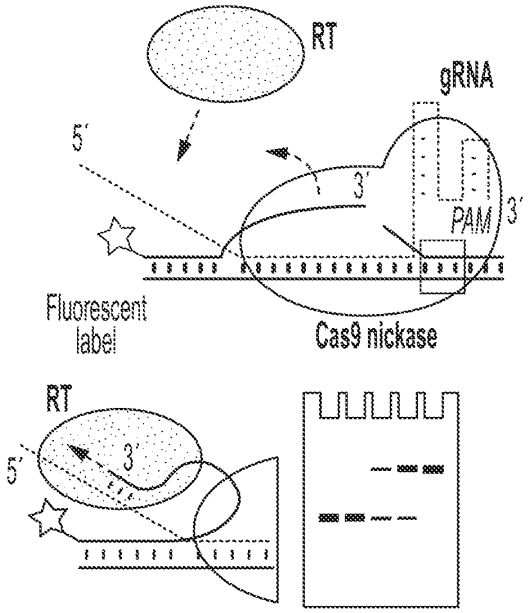


FIG. 4A

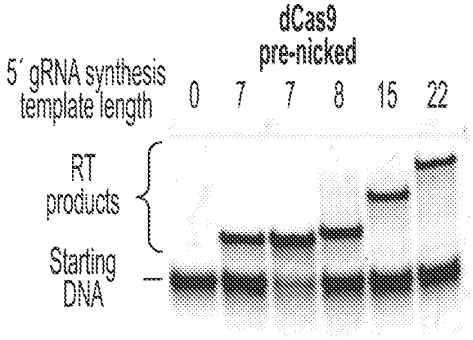


FIG. 4B

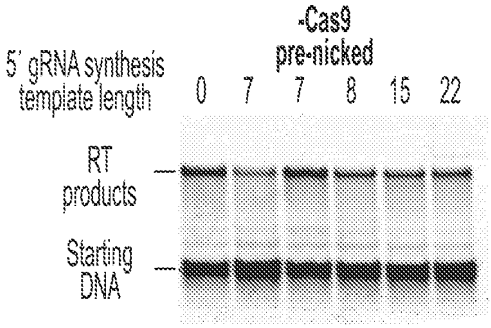


FIG. 4C

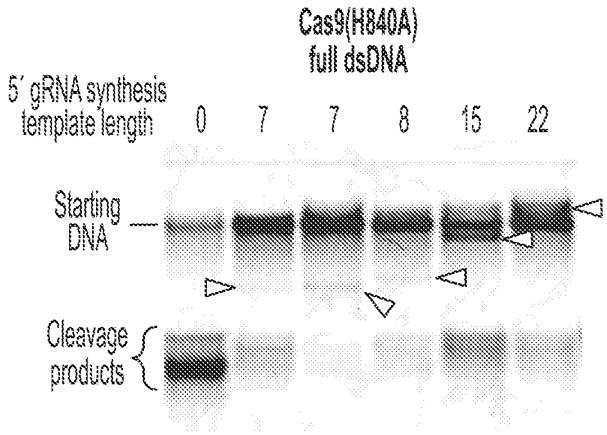

Cas9(H840A)
full dsDNA
5′ gRNA synthesis template length
0
7
7
8
15
22
Starting DNA
Cleavage products

FIG. 4D

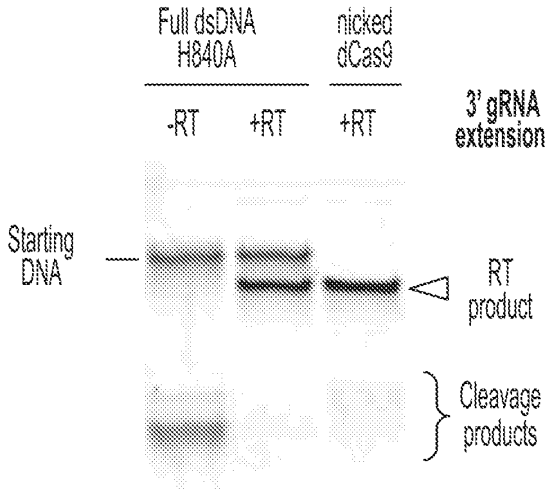

Full dsDNA H840A
nicked dCas9
-RT
+RT
+RT
3′ gRNA extension
Starting DNA
RT product
Cleavage products

FIG. 4E

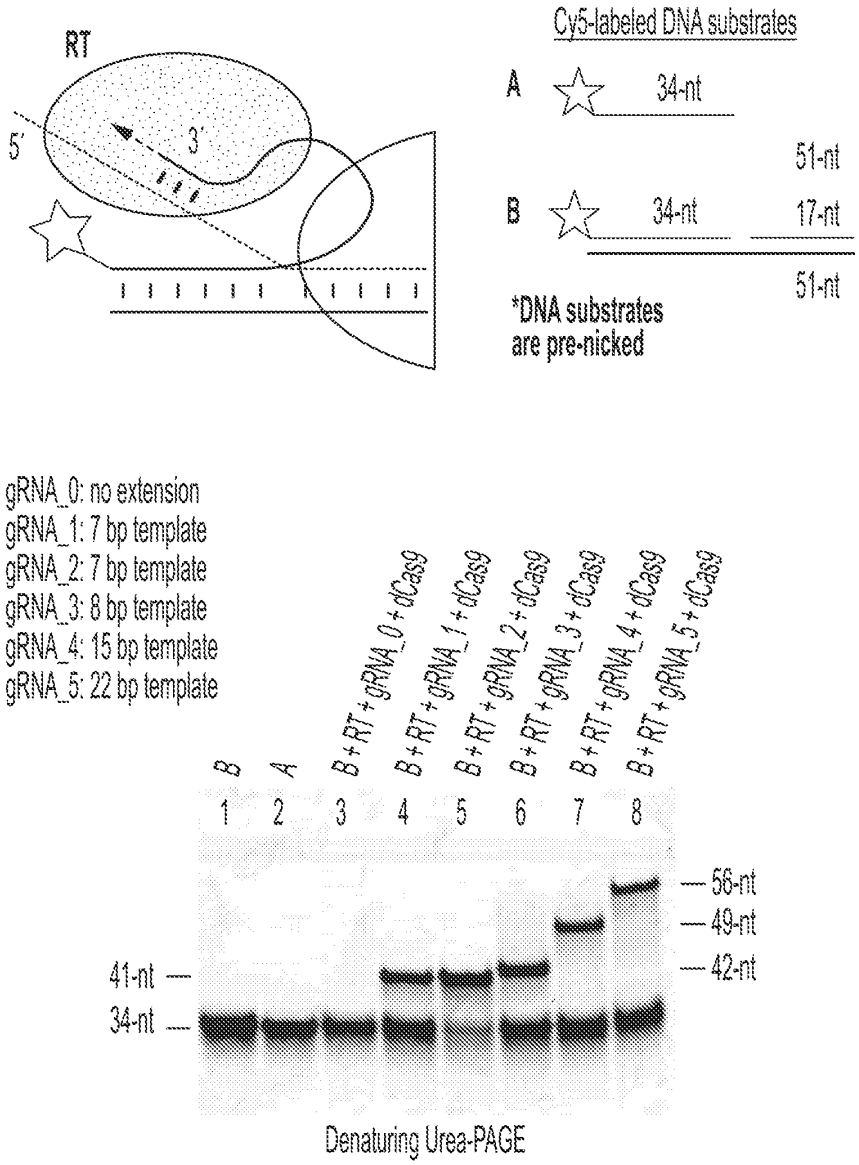

RT
5'
3'
Cy5-labeled DNA substrates
A
34-nt
51-nt
B
34-nt
17-nt
51-nt
*DNA substrates
are pre-nicked
gRNA_0: no extension
gRNA_1: 7 bp template
gRNA_2: 7 bp template
gRNA_3: 8 bp template
gRNA_4: 15 bp template
gRNA_5: 22 bp template
B
A
B+RT+gRNA_0+dCas9
B+RT+gRNA_1+dCas9
B+RT+gRNA_2+dCas9
B+RT+gRNA_3+dCas9
B+RT+gRNA_4+dCas9
B+RT+gRNA_5+dCas9
1 2 3 4 5 6 7 8
56-nt
49-nt
42-nt
41-nt
34-nt
Denaturing Urea-PAGE

FIG. 5

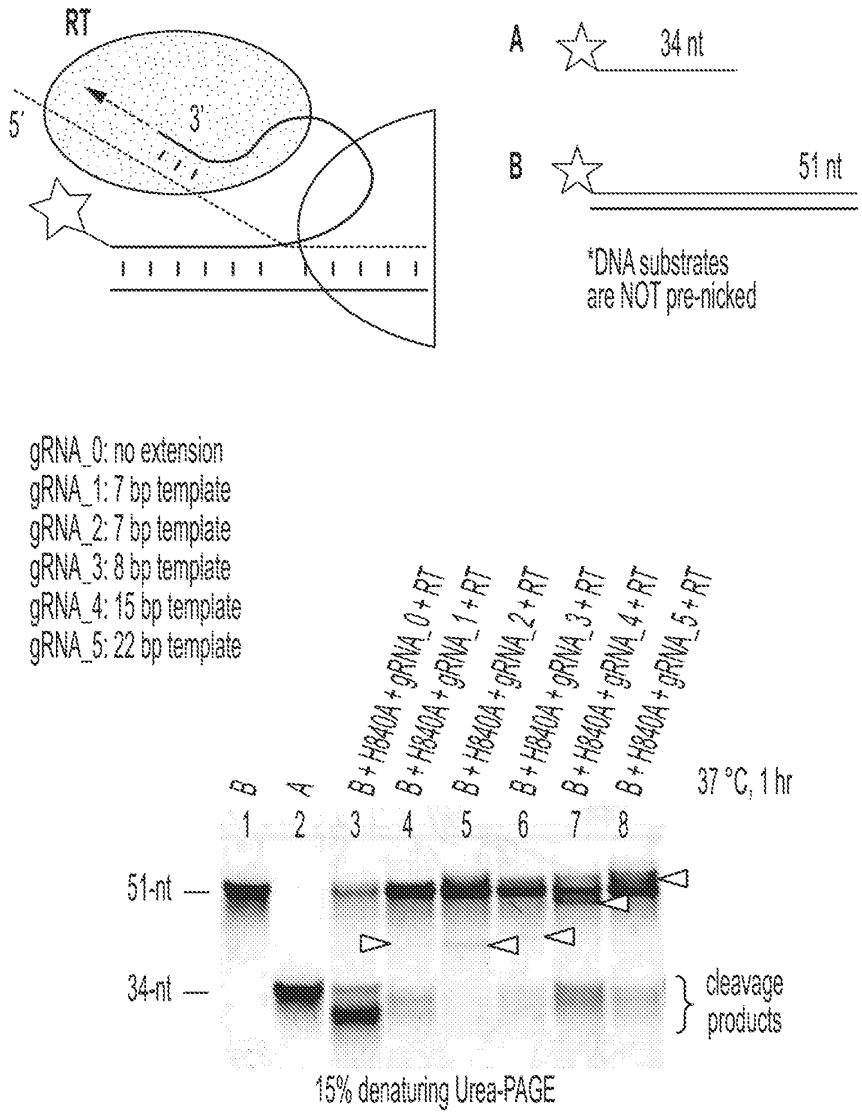

RT
5'
3'
A
34 nt
B
51 nt
*DNA substrates
are NOT pre-nicked
gRNA_0: no extension
gRNA_1: 7 bp template
gRNA_2: 7 bp template
gRNA_3: 8 bp template
gRNA_4: 15 bp template
gRNA_5: 22 bp template
B
A
B + H840A + gRNA_0 + RT
B + H840A + gRNA_1 + RT
B + H840A + gRNA_2 + RT
B + H840A + gRNA_3 + RT
B + H840A + gRNA_4 + RT
B + H840A + gRNA_5 + RT
37 °C, 1 hr
1 2 3 4 5 6 7 8
51-nt
34-nt
cleavage products
15% denaturing Urea-PAGE

FIG. 6

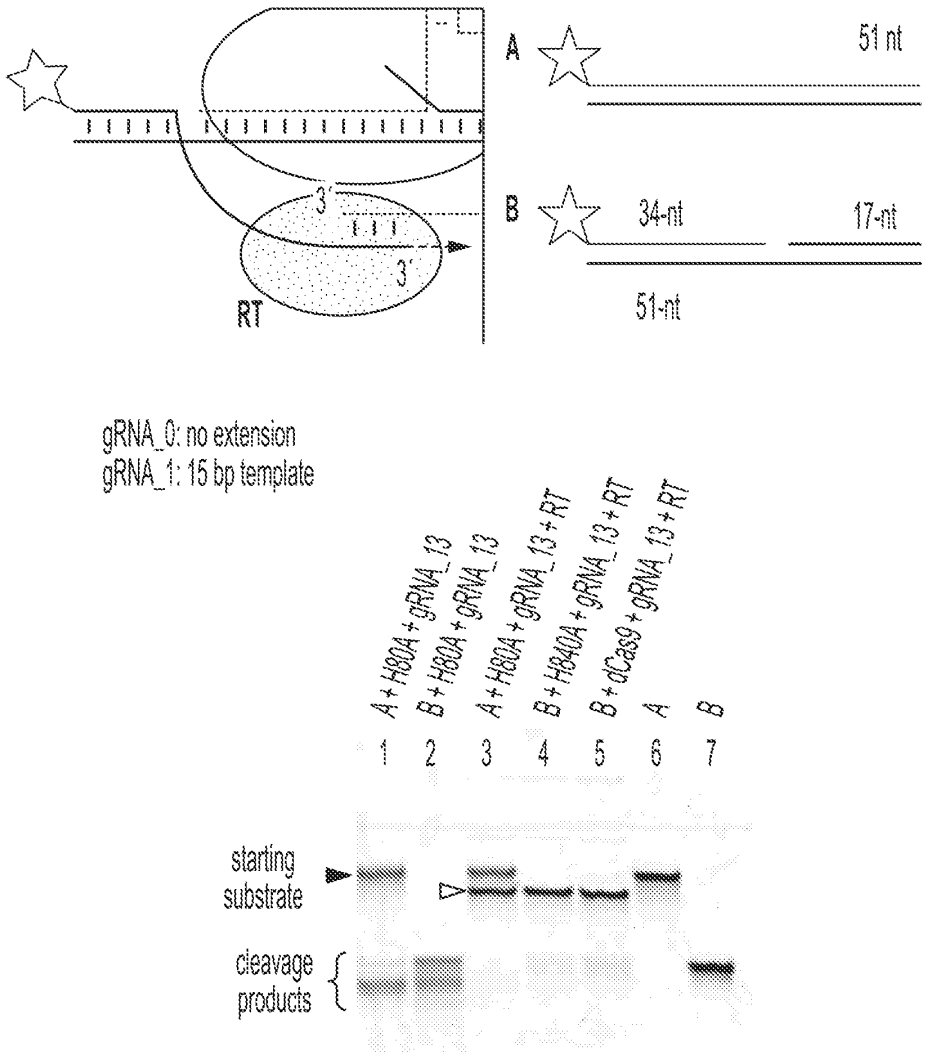

A
51 nt
B
34-nt
17-nt
51-nt
3'
3'
RT
gRNA_0: no extension
gRNA_1: 15 bp template
A + H80A + gRNA_13
B + H80A + gRNA_13
A + H80A + gRNA_13 + RT
B + H840A + gRNA_13 + RT
B + dCas9 + gRNA_13 + RT
A
B
1 2 3 4 5 6 7
starting substrate
cleavage products

FIG. 7

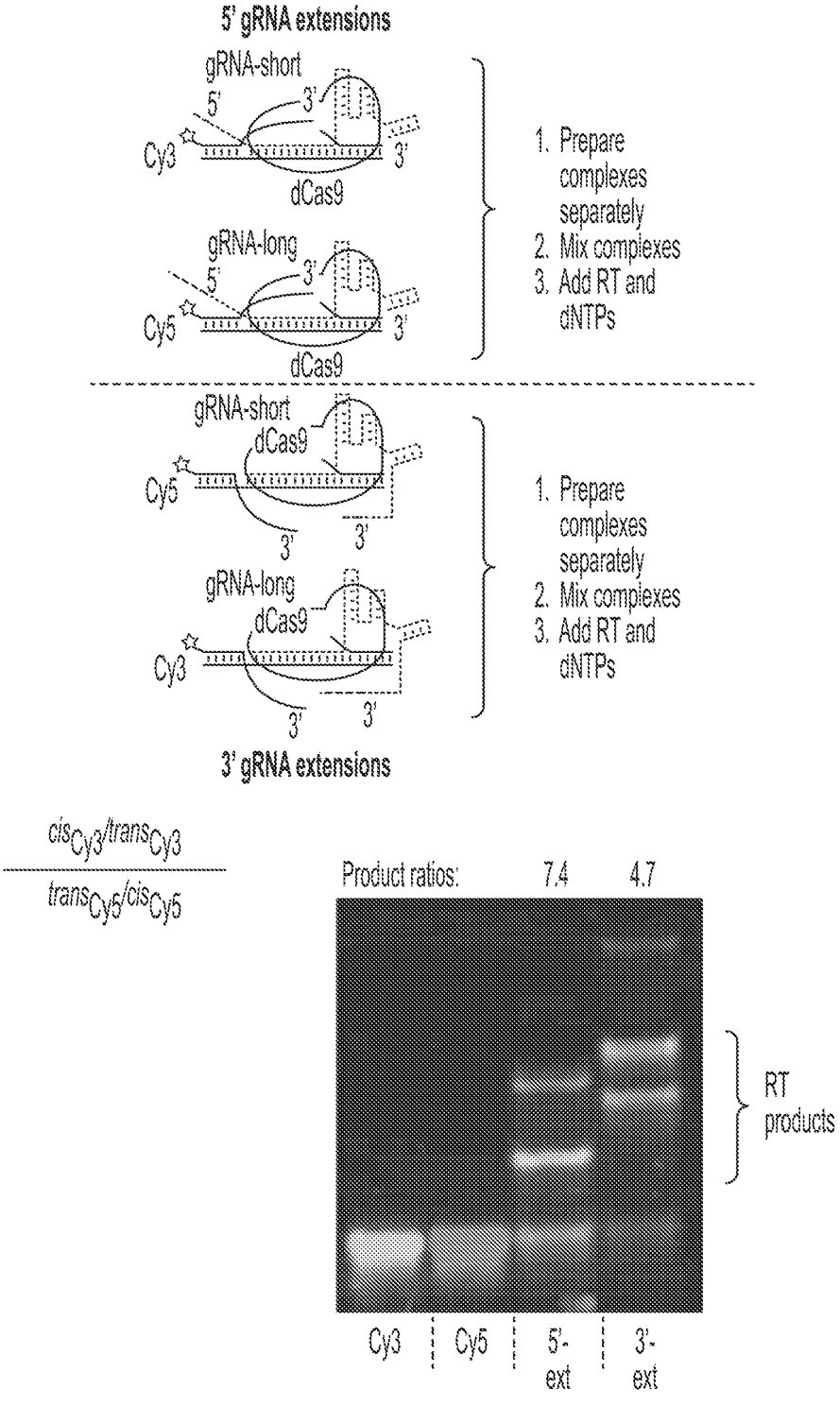

5' gRNA extensions
gRNA-short
5'
3'
Cy3
3'
dCas9
gRNA-long
5'
3'
Cy5
3'
dCas9
1. Prepare complexes separately
2. Mix complexes
3. Add RT and dNTPs
gRNA-short
dCas9
Cy5
3'
3'
gRNA-long
dCas9
Cy3
3'
3'
1. Prepare complexes separately
2. Mix complexes
3. Add RT and dNTPs
3' gRNA extensions
cisCy3/transCy3
transCy5/cisCy5
Product ratios:
7.4
4.7
RT products
Cy3
Cy5
5'-ext
3'-ext

FIG. 8

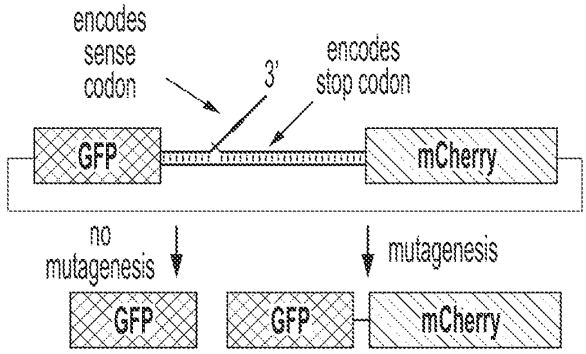

encodes
sense
codon
3'
encodes
stop codon
GFP
mCherry
no
mutagenesis
mutagenesis
GFP
GFP
mCherry

FIG. 9A

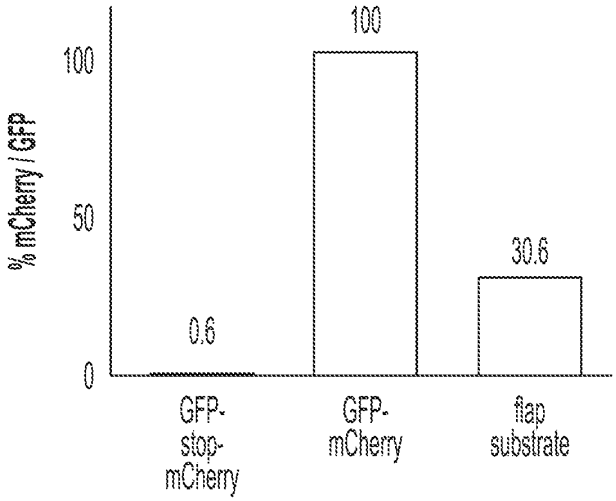

% mCherry / GFP
100
50
0
100
0.6
30.6
GFP-
stop-
mCherry
GFP-
mCherry
flap
substrate

FIG. 9B

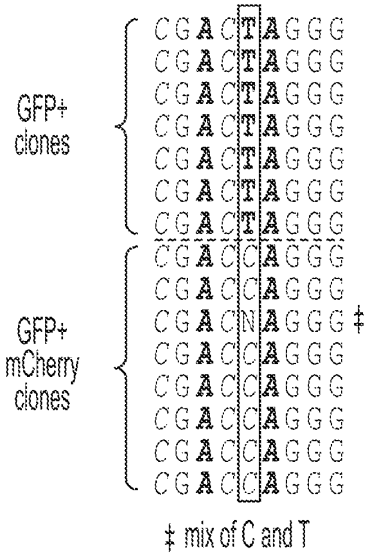

GFP+
clones
CGACTAGGG
CGACTAGGG
CGACTAGGG
CGACTAGGG
CGACTAGGG
CGACTAGGG
CGACTAGGG
GFP+
mCherry
clones
CGACCAGGG
CGACCAGGG
CGACNAGGG ‡
CGACCAGGG
CGACCAGGG
CGACCAGGG
CGACCAGGG
CGACCAGGG
‡ mix of C and T

FIG. 9C

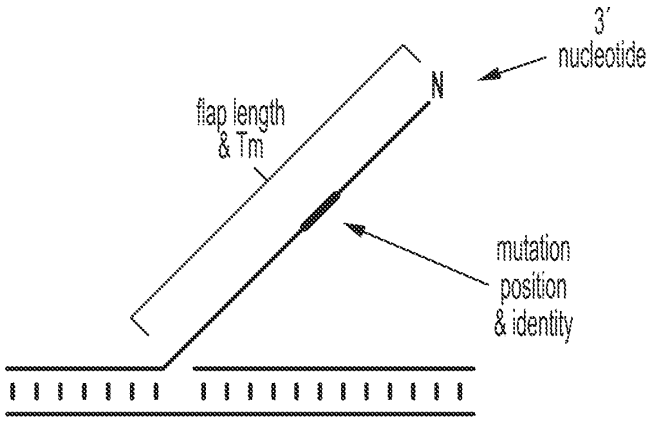

3'
nucleotide
N
flap length
& Tm
mutation
position
& identity

FIG. 9D

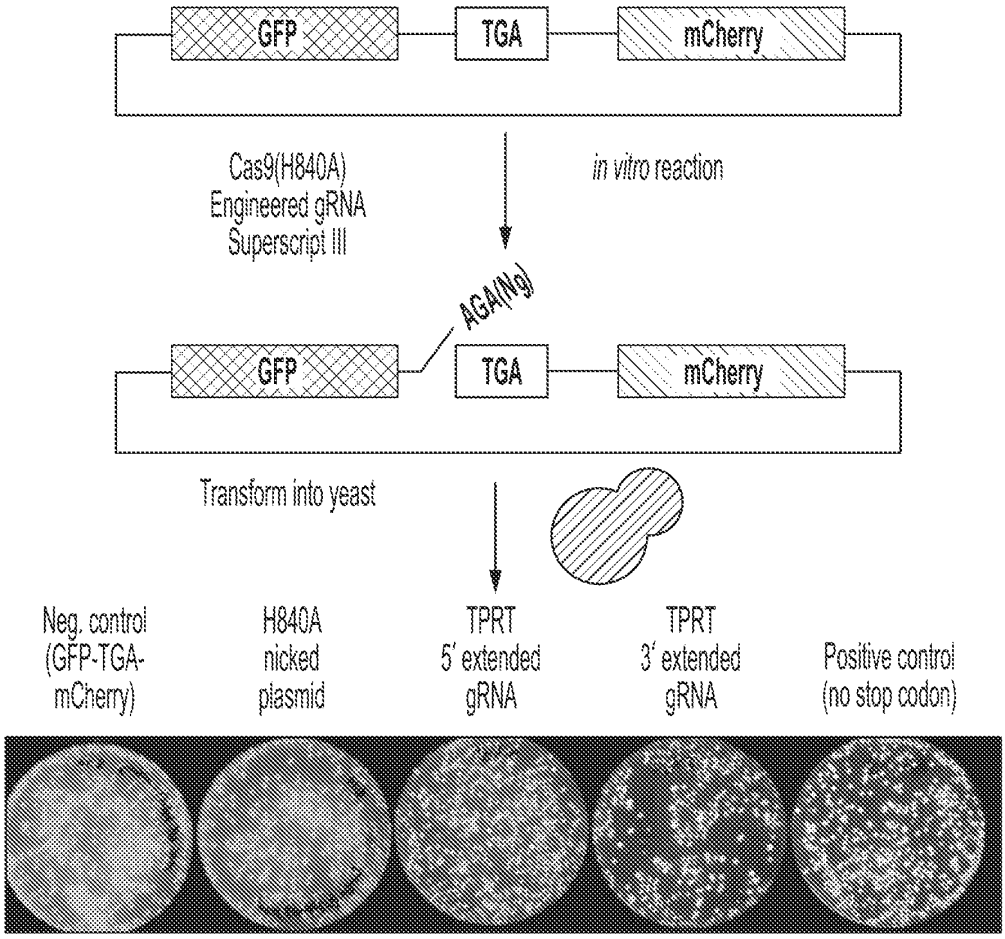

GFP
TGA
mCherry
Cas9(H840A)
Engineered gRNA
Superscript III
in vitro reaction
AGA(Ng)
GFP
TGA
mCherry
Transform into yeast
Neg. control (GFP-TGA-mCherry)
H840A nicked plasmid
TPRT 5′ extended gRNA
TPRT 3′ extended gRNA
Positive control (no stop codon)

FIG. 10

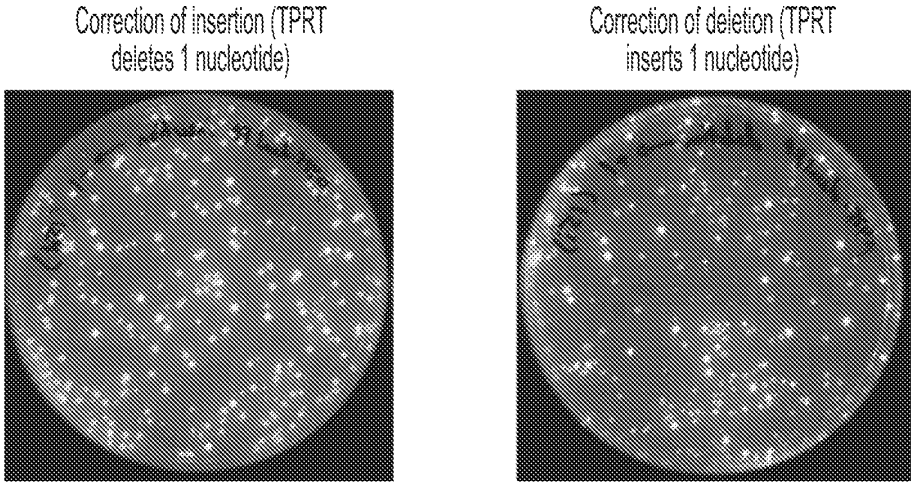

Correction of insertion (TPRT deletes 1 nucleotide)
Correction of deletion (TPRT inserts 1 nucleotide)

FIG. 11

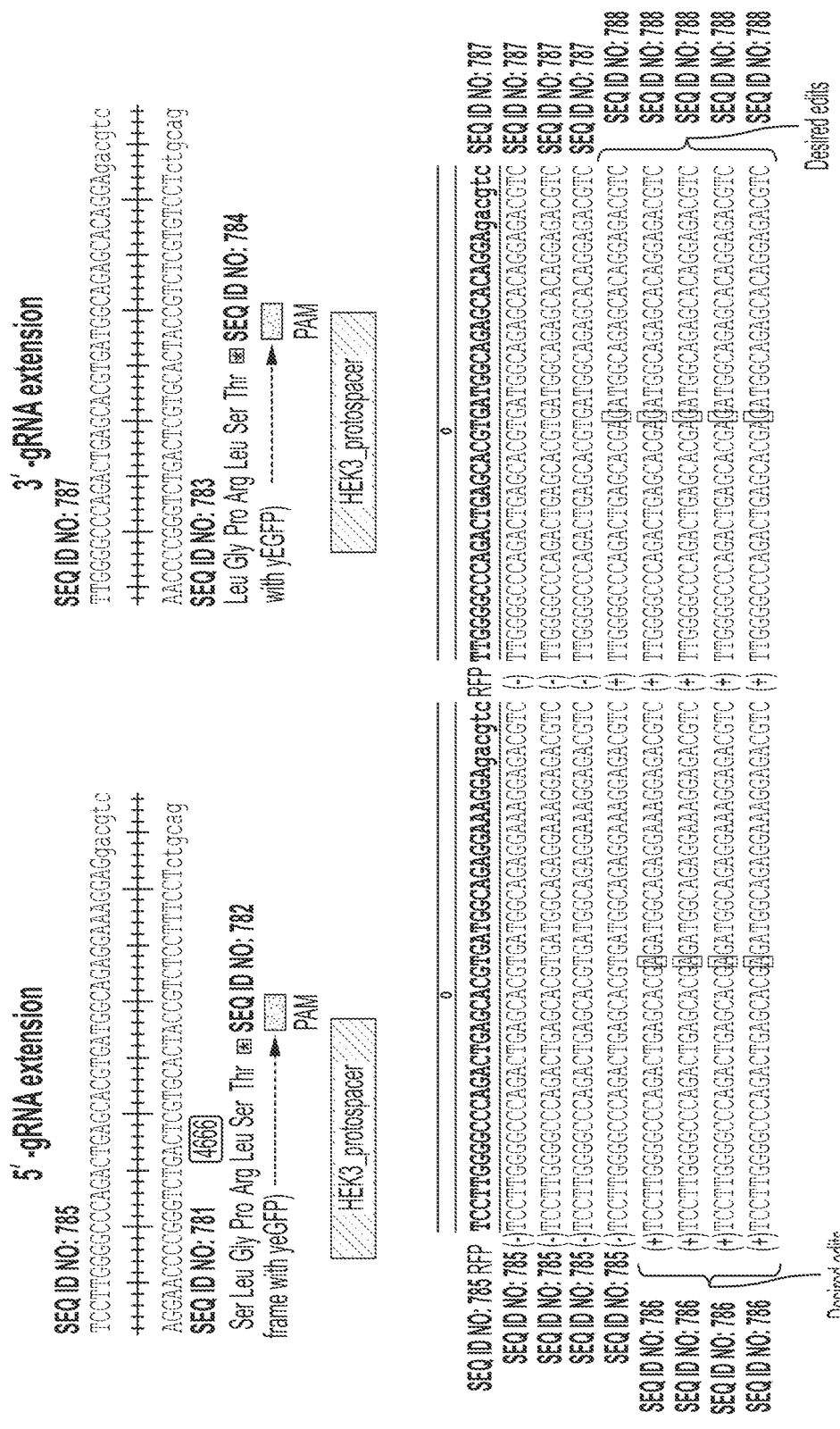

5' -gRNA extension
SEQ ID NO: 785
TCCTTGGGCCCAGACTGAGCACGTGATGGCAGGAAAGGACGAGgacgtc
SEQ ID NO: 781
4666
Ser Leu Gly Pro Arg Leu Ser Thr
frame with yeGFP)
PAM
HEK3_protospacer
3' -gRNA extension
SEQ ID NO: 787
TTGGGCCCAGACTGAGCACGTGATGGCAGGAAAGGACGAGgacgtc
AAGCCCGGGTCTGACTCGTGCACTACCGTTTCCTTTGCCTCctgcag
SEQ ID NO: 783
Leu Gly Pro Arg Leu Ser Thr
SEQ ID NO: 784
with yEGFP)
PAM
HEK3_protospacer
Desired edits
Desired edits

FIG. 12

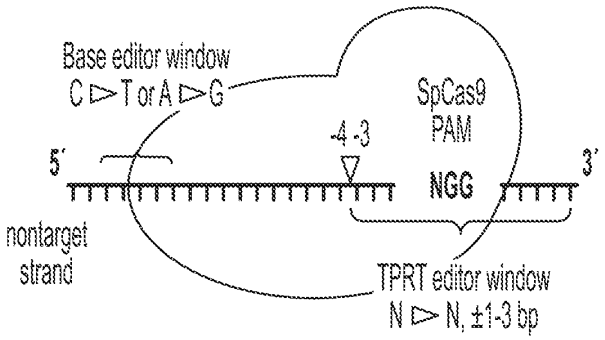

Base editor window
C ⊳ T or A ⊳ G
SpCas9
PAM
-4 -3
5′
3′
NGG
nontarget
strand
TPRT editor window
N ⊳ N, ±1-3 bp

FIG. 13

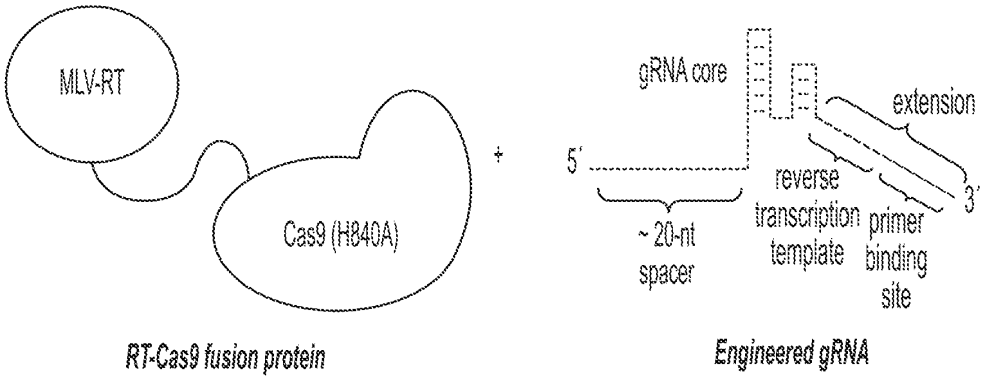

MLV-RT
Cas9 (H840A)
+
gRNA core
extension
5′
3′
reverse
transcription
template
primer
binding
site
~ 20-nt
spacer
RT-Cas9 fusion protein
Engineered gRNA

FIG. 14

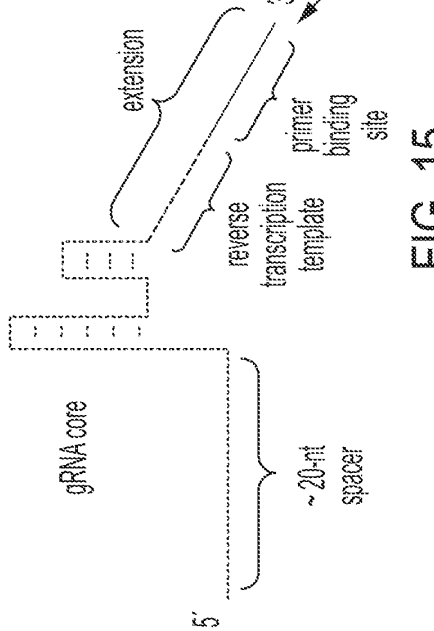

gRNA core
~ 20-nt spacer
reverse transcription template
primer binding site
extension
7, 8, 9, 10, 11 or 12 nucleotides long
5'
3'

FIG. 15

HEK3 locus
SEQ ID NO: 789
SEQ ID NO: 790
spacer_3
spacer_3
spacer_2
editing target
PAM
A
B
spacer_4_rev
spacer_5_rev
spacer_6
C
D
E
297

FIG. 16

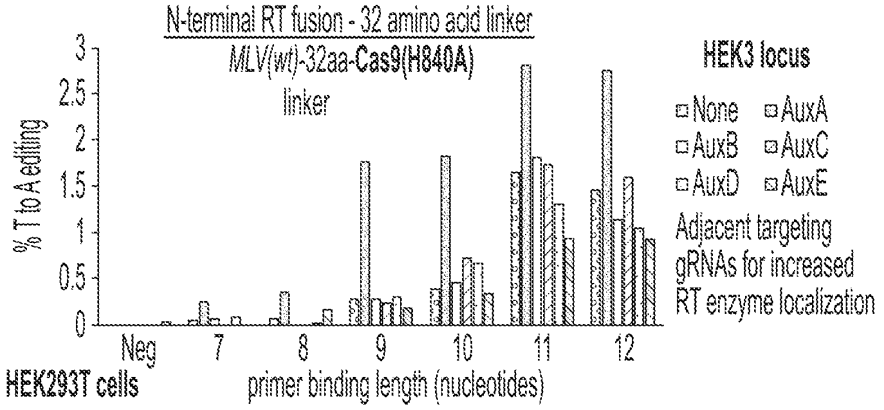


FIG. 17A

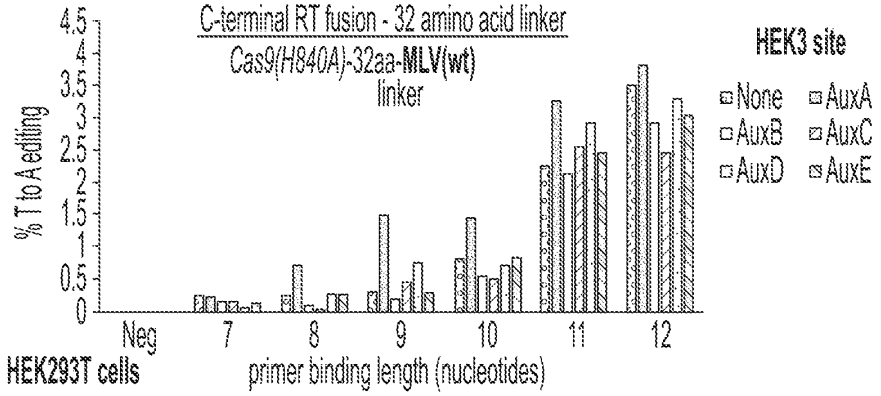


FIG. 17B

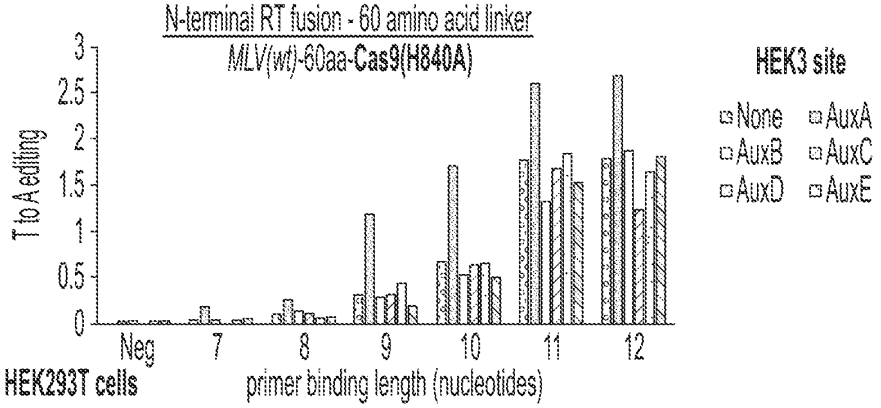


FIG. 17C

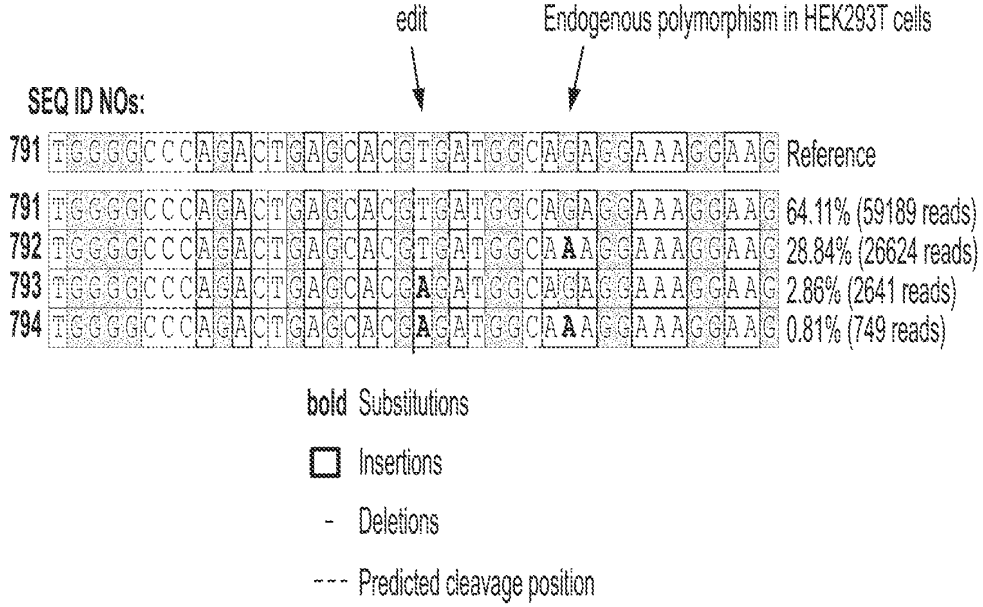

edit
Endogenous polymorphism in HEK293T cells
SEQ ID NOs:
791 TGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAG Reference
791 TGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAG 64.11% (59189 reads)
792 TGGGGCCCAGACTGAGCACGTGATGGCAAAGGAAAGGAAG 28.84% (26624 reads)
793 TGGGGCCCAGACTGAGCACGAGATGGCAGAGGAAAGGAAG 2.86% (2641 reads)
794 TGGGGCCCAGACTGAGCACGAGATGGCAAAGGAAAGGAAG 0.81% (749 reads)
bold Substitutions
Insertions
- Deletions
--- Predicted cleavage position

FIG. 18

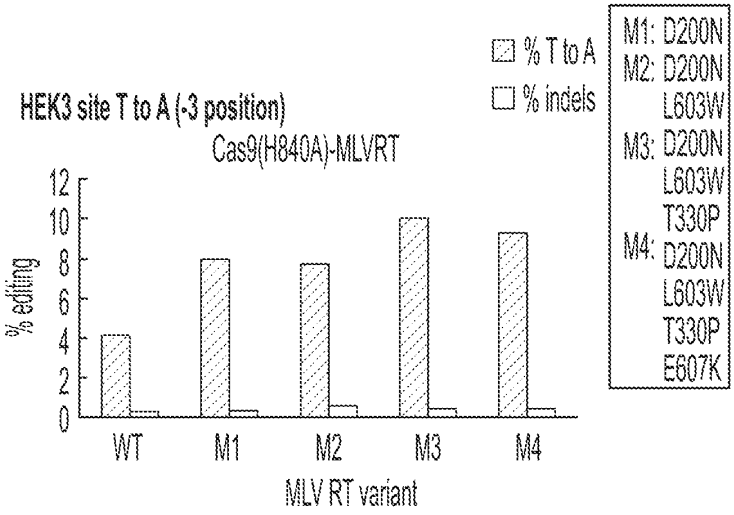

HEK3 site T to A (-3 position)
Cas9(H840A)-MLVRT
% T to A
% indels
% editing
0
2
4
6
8
10
12
WT
M1
M2
M3
M4
MLV RT variant
M1: D200N
M2: D200N L603W
M3: D200N L603W T330P
M4: D200N L603W T330P E607K

FIG. 19

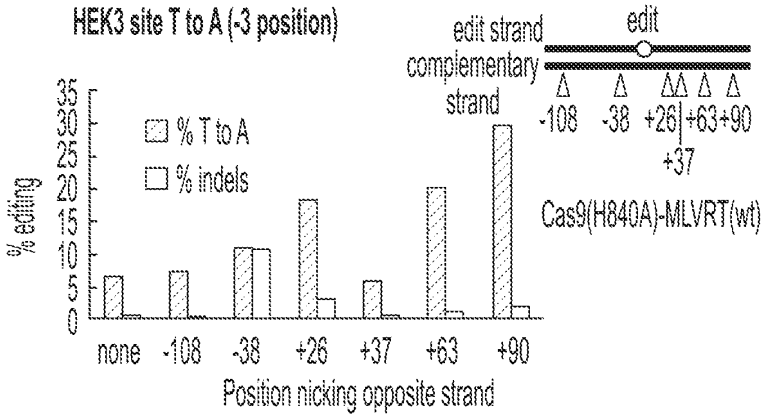

HEK3 site T to A (-3 position)
% T to A
% indels
% editing
0
5
10
15
20
25
30
35
none
-108
-38
+26
+37
+63
+90
Position nicking opposite strand
edit strand
complementary strand
edit
-108
-38
+26
+37
+63+90
Cas9(H840A)-MLVRT(wt)

FIG. 20

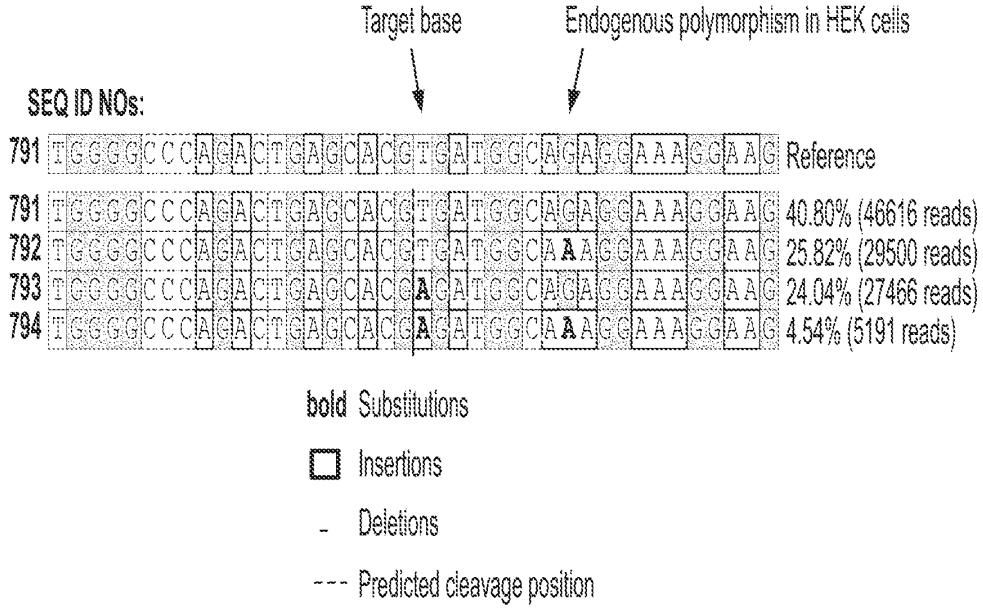

Target base
Endogenous polymorphism in HEK cells
SEQ ID NOs:
791 TGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAG Reference
791 TGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAG 40.80% (46616 reads)
792 TGGGGCCCAGACTGAGCACGTGATGGCAAAGGAAAGGAAG 25.82% (29500 reads)
793 TGGGGCCCAGACTGAGCACGAGATGGCAGAGGAAAGGAAG 24.04% (27466 reads)
794 TGGGGCCCAGACTGAGCACGAGATGGCAAAGGAAAGGAAG 4.54% (5191 reads)
bold Substitutions
Insertions
- Deletions
--- Predicted cleavage position

FIG. 21

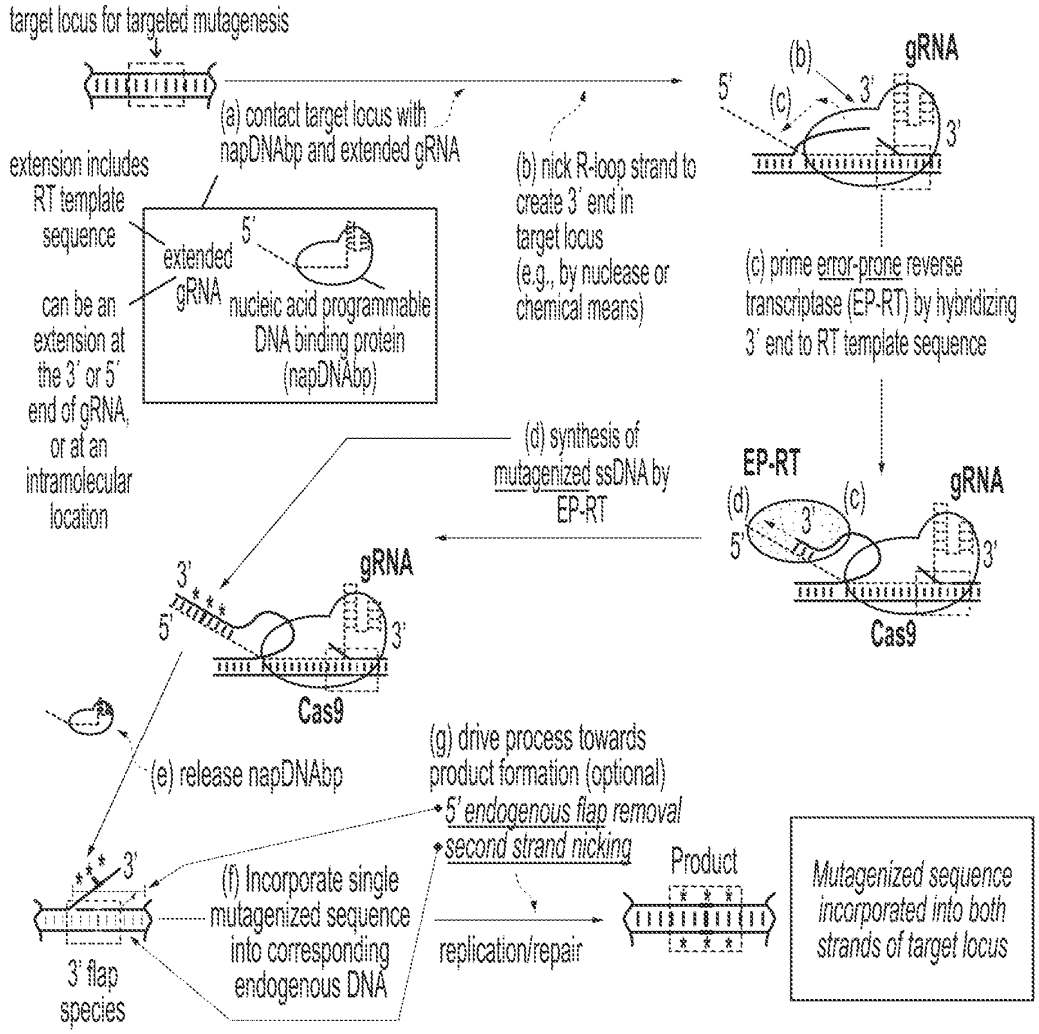

target locus for targeted mutagenesis
(a) contact target locus with napDNAbp and extended gRNA
extension includes RT template sequence
can be an extension at the 3′ or 5′ end of gRNA, or at an intramolecular location
extended gRNA
5′
nucleic acid programmable DNA binding protein (napDNAbp)
(b) nick R-loop strand to create 3′ end in target locus (e.g., by nuclease or chemical means)
(b)
gRNA
5′
(c)
3′
3′
(c) prime error-prone reverse transcriptase (EP-RT) by hybridizing 3′ end to RT template sequence
EP-RT
(d)
(c)
gRNA
5′
3′
3′
Cas9
(d) synthesis of mutagenized ssDNA by EP-RT
3′
5′
gRNA
3′
Cas9
(e) release napDNAbp
(g) drive process towards product formation (optional)
5′ endogenous flap removal
second strand nicking
3′
3′ flap species
(f) Incorporate single mutagenized sequence into corresponding endogenous DNA
replication/repair
Product
Mutagenized sequence incorporated into both strands of target locus

FIG. 22

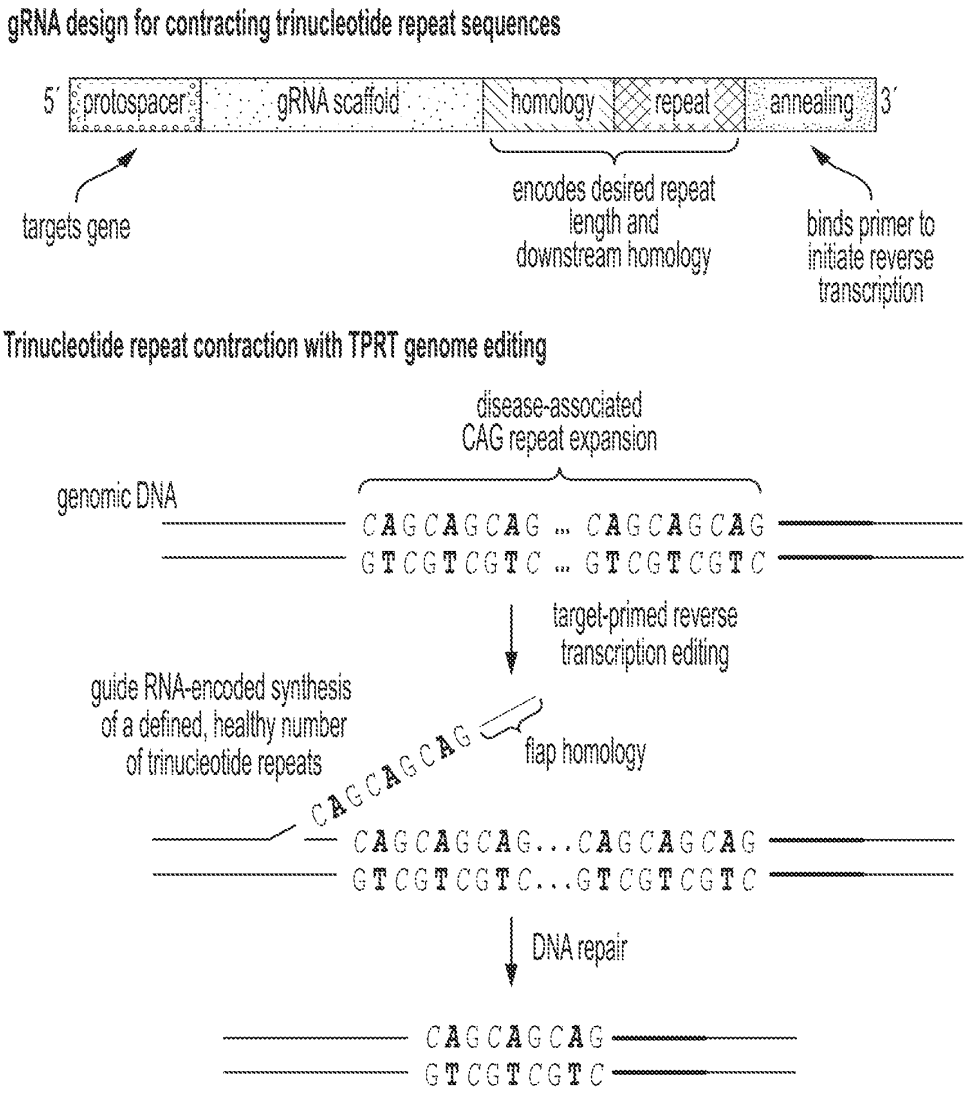

gRNA design for contracting trinucleotide repeat sequences
5′
protospacer
gRNA scaffold
homology
repeat
annealing
3′
targets gene
encodes desired repeat length and downstream homology
binds primer to initiate reverse transcription
Trinucleotide repeat contraction with TPRT genome editing
disease-associated CAG repeat expansion
genomic DNA
CAGCAGCAG … CAGCAGCAG
GTCGTCGTC … GTCGTCGTC
target-primed reverse transcription editing
guide RNA-encoded synthesis of a defined, healthy number of trinucleotide repeats
CAGCAGCAG
flap homology
CAGCAGCAG...CAGCAGCAG
GTCGTCGTC...GTCGTCGTC
DNA repair
CAGCAGCAG
GTCGTCGTC

FIG. 23

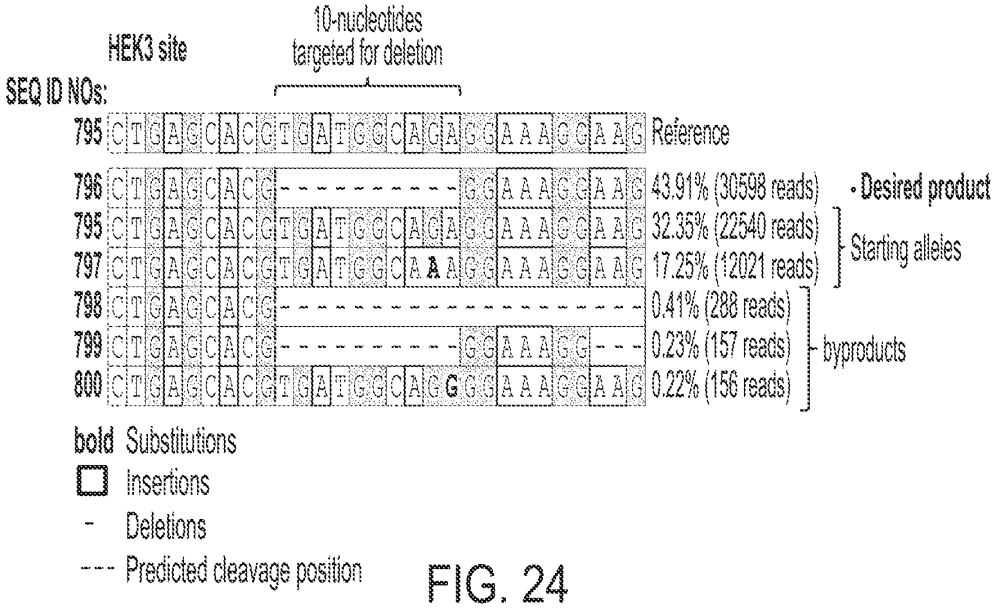

HEK3 site
10-nucleotides targeted for deletion
SEQ ID NOs:
795 CTGAGCACGTGATGGCAGAGGAAAGGAAG Reference
796 CTGAGCACG---------GGAAAGGAAG 43.91% (30598 reads) • Desired product
795 CTGAGCACGTGATGGCAGAGGAAAGGAAG 32.35% (22540 reads) Starting alleles
797 CTGAGCACGTGATGGCAAAGGAAAGGAAG 17.25% (12021 reads) Starting alleles
798 CTGAGCACG---------- 0.41% (288 reads) byproducts
799 CTGAGCACG---------GGAAAGG--- 0.23% (157 reads) byproducts
800 CTGAGCACGTGATGGCAGGGAAAGGAAG 0.22% (156 reads) byproducts
bold Substitutions
Insertions
- Deletions
--- Predicted cleavage position

FIG. 24

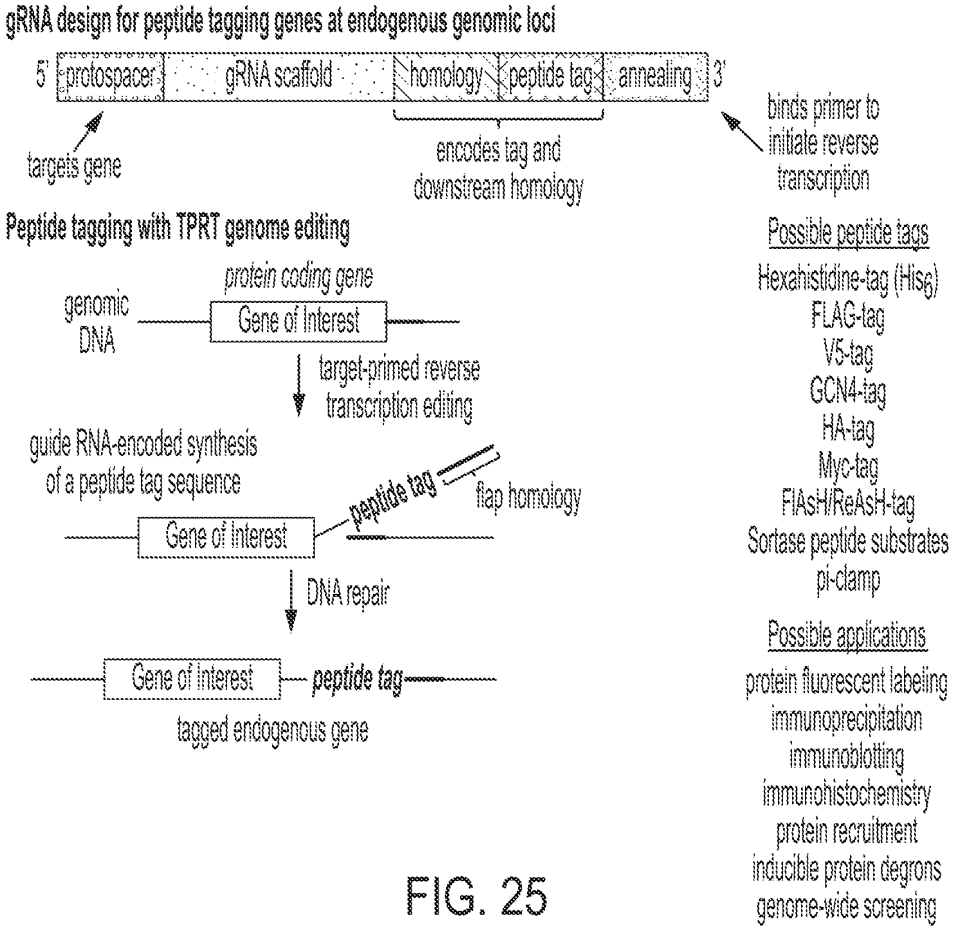

gRNA design for peptide tagging genes at endogenous genomic loci
5' protospacer gRNA scaffold homology peptide tag annealing 3'
targets gene
encodes tag and downstream homology
binds primer to initiate reverse transcription
Peptide tagging with TPRT genome editing
protein coding gene
genomic DNA
Gene of Interest
target-primed reverse transcription editing
guide RNA-encoded synthesis of a peptide tag sequence
peptide tag
flap homology
Gene of Interest
DNA repair
Gene of Interest peptide tag
tagged endogenous gene
Possible peptide tags
Hexahistidine-tag (His6)
FLAG-tag
V5-tag
GCN4-tag
HA-tag
Myc-tag
FlAsH/ReAsH-tag
Sortase peptide substrates
pi-clamp
Possible applications
protein fluorescent labeling
immunoprecipitation
immunoblotting
immunohistochemistry
protein recruitment
inducible protein degrons
genome-wide screening

FIG. 25

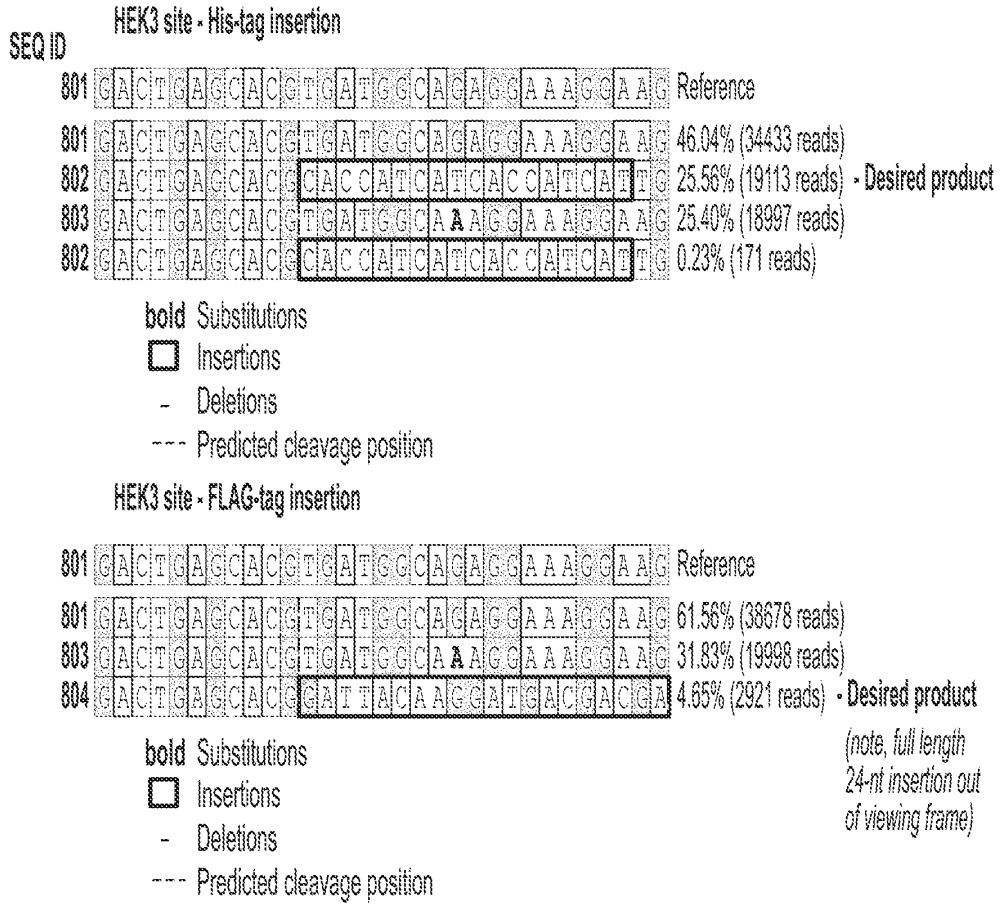


FIG. 26A

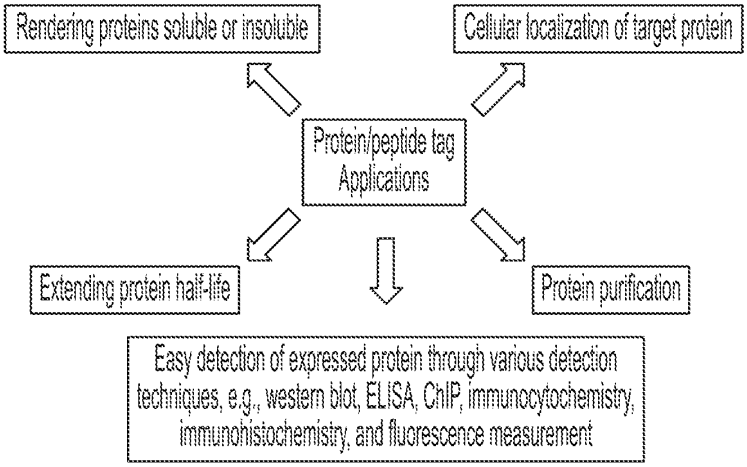


FIG. 26B

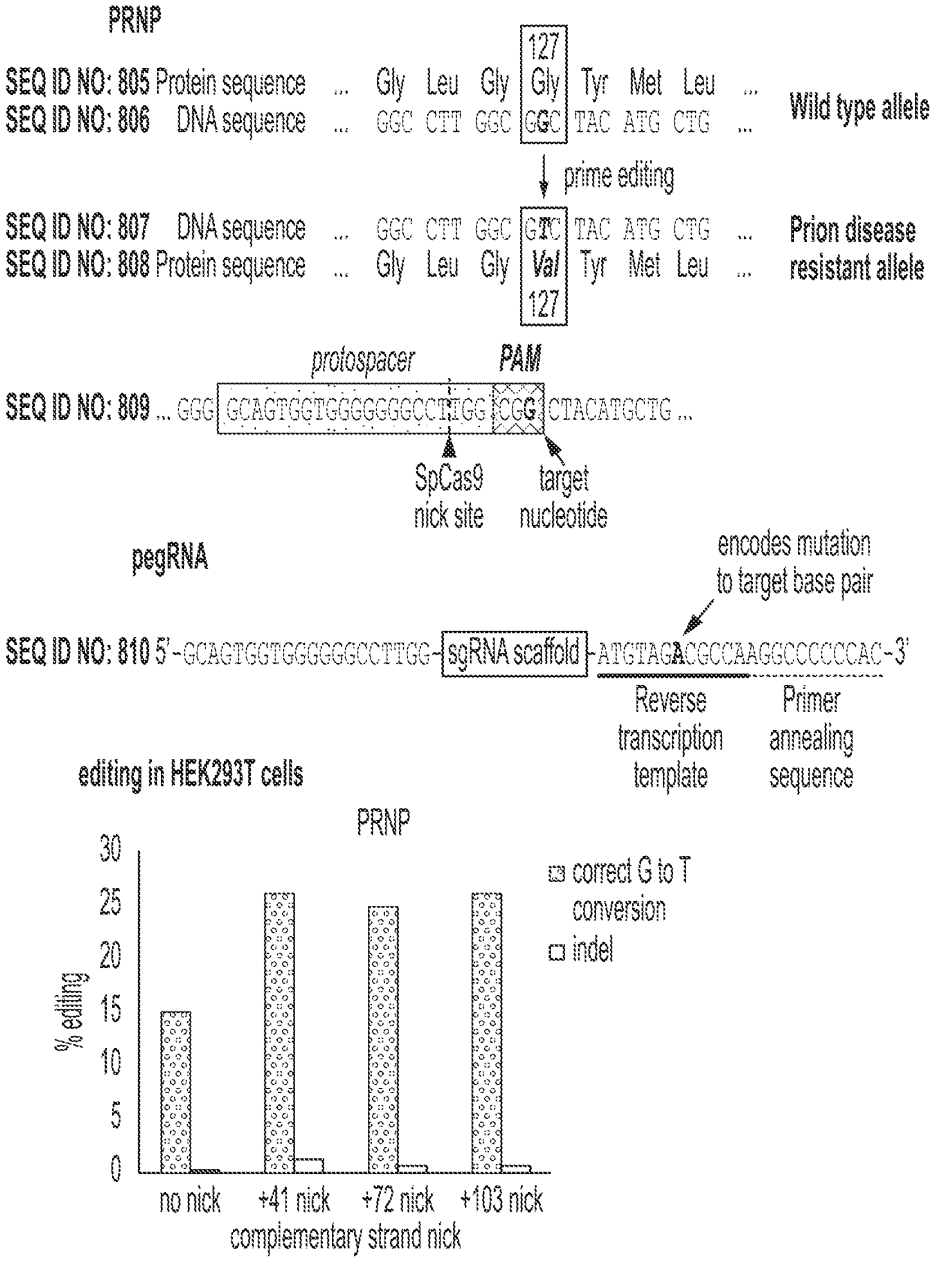

PRNP
SEQ ID NO: 805 Protein sequence ... Gly Leu Gly Gly Tyr Met Leu ...
127
SEQ ID NO: 806 DNA sequence ... GGC CTT GGC GGC TAC ATG CTG ...
Wild type allele
prime editing
SEQ ID NO: 807 DNA sequence ... GGC CTT GGC GTC TAC ATG CTG ...
Prion disease
SEQ ID NO: 808 Protein sequence ... Gly Leu Gly Val Tyr Met Leu ...
127
resistant allele
protospacer
PAM
SEQ ID NO: 809 ... GGG GCAGTGGTGGGGGGGCCTTGG CGG CTACATGCTG ...
SpCas9 nick site
target nucleotide
pegRNA
encodes mutation to target base pair
SEQ ID NO: 810 5'-GCAGTGGTGGGGGGGCCTTGG-sgRNA scaffold-ATGTAGACGCCAAGGCCCCCCAC-3'
Reverse transcription template
Primer annealing sequence
editing in HEK293T cells
PRNP
% editing
30
25
20
15
10
5
0
correct G to T conversion
indel
no nick
+41 nick
+72 nick
+103 nick
complementary strand nick

FIG. 27

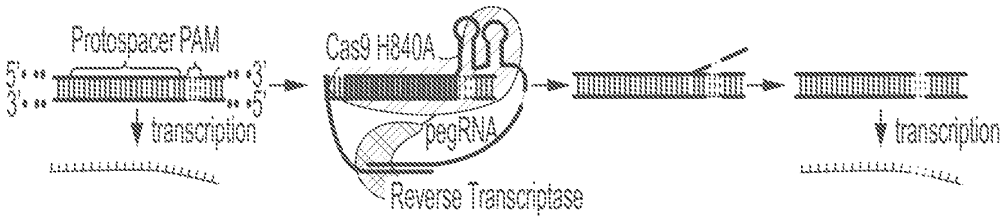


FIG. 28A

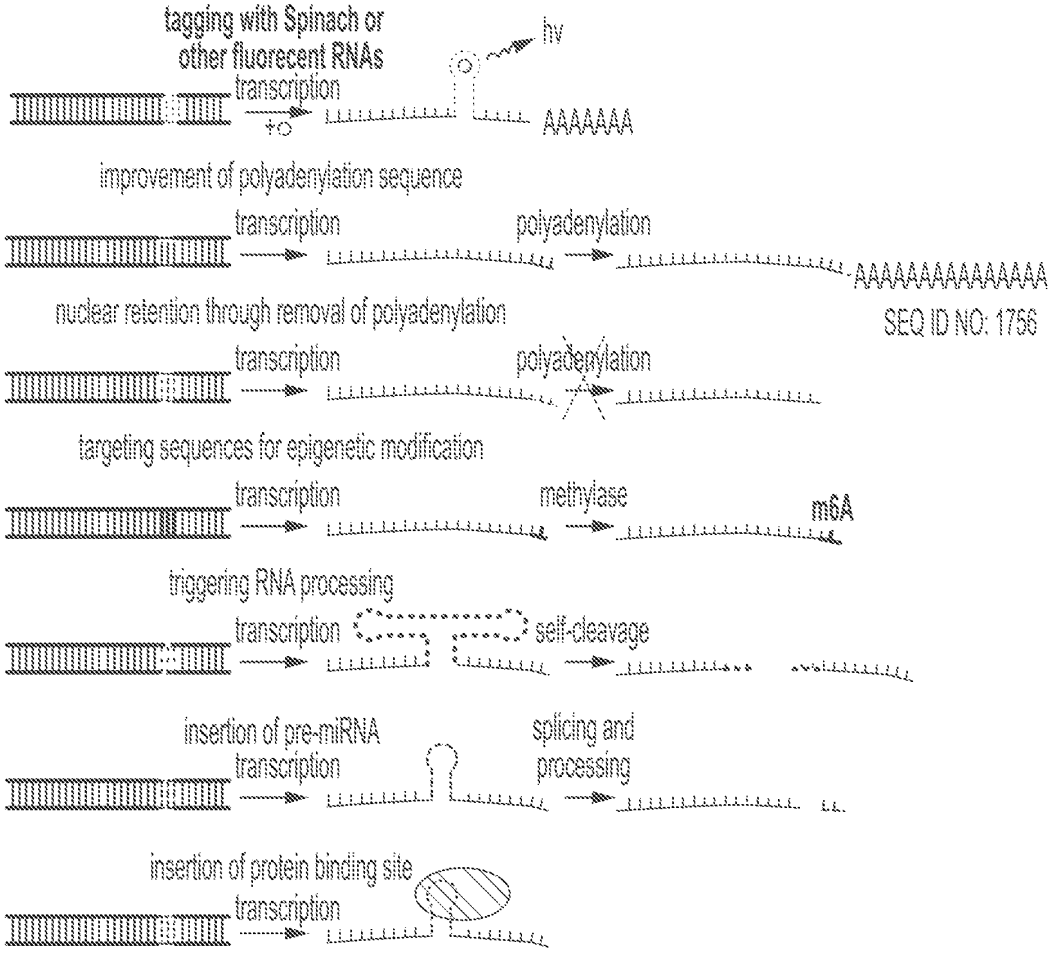


FIG. 28B

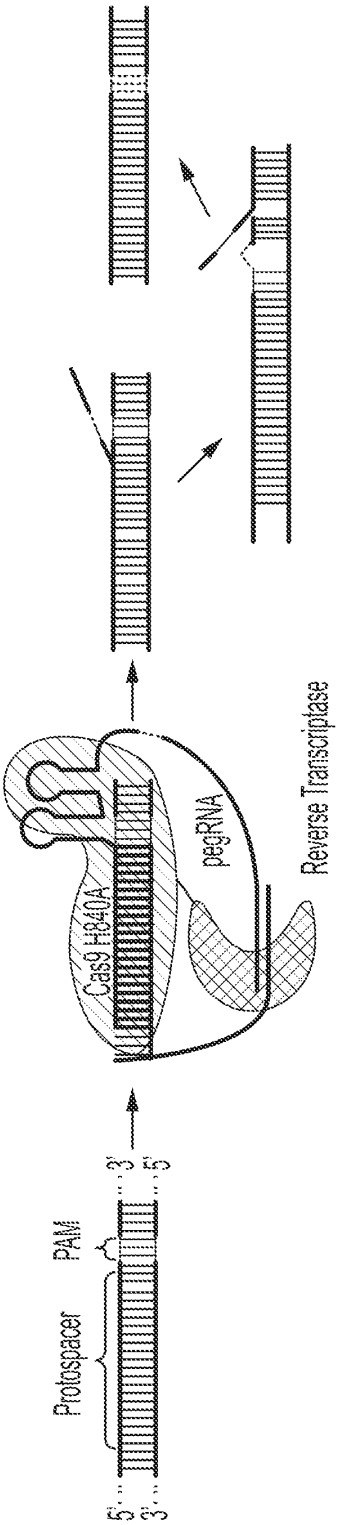

Protospacer
PAM
3'
5'
Cas9 H840A
pegRNA
Reverse Transcriptase

FIG. 29A

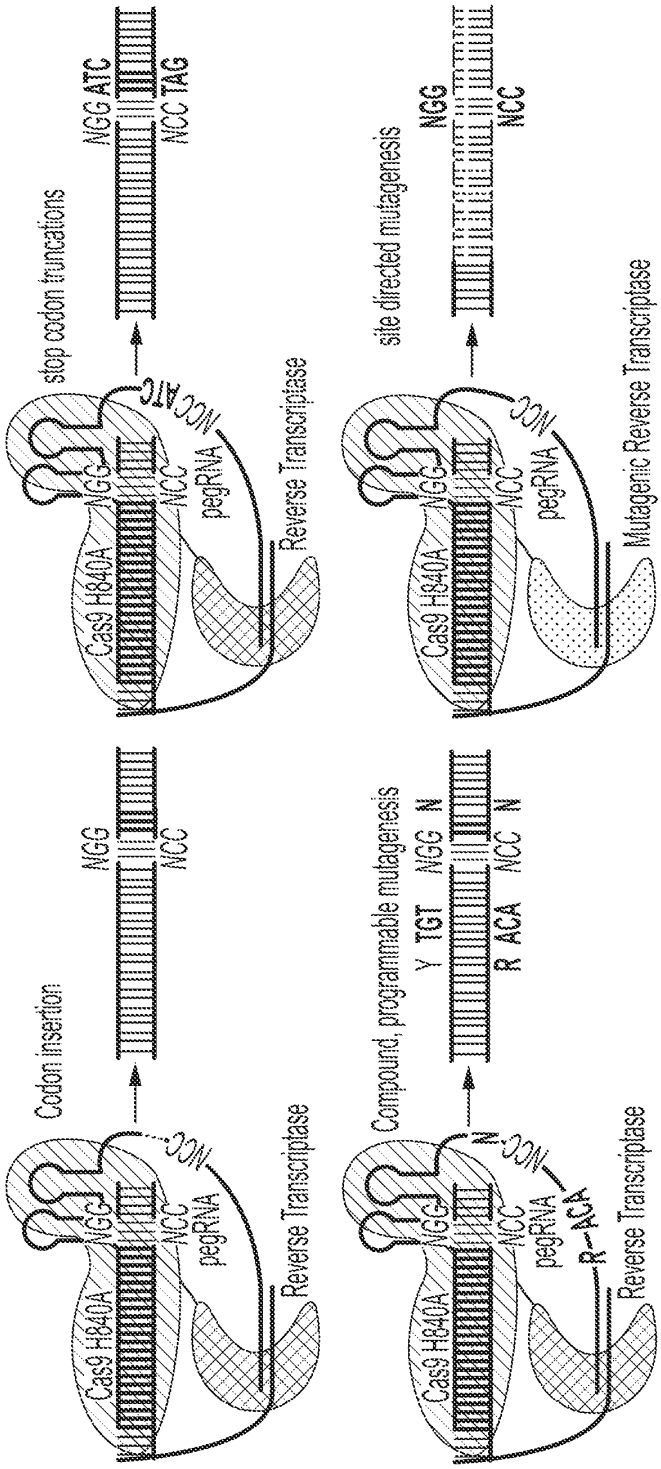

Codon insertion
NGG
NCC
Cas9 H840A
NGG
NCC
pegRNA
NCC
Reverse Transcriptase
stop codon truncations
NGG ATC
NCC TAG
NCC ATC
Compound, programmable mutagenesis
Y TGT NGG N
R ACA NCC N
N
R-ACA
site directed mutagenesis
Mutagenic Reverse Transcriptase

FIG. 29B

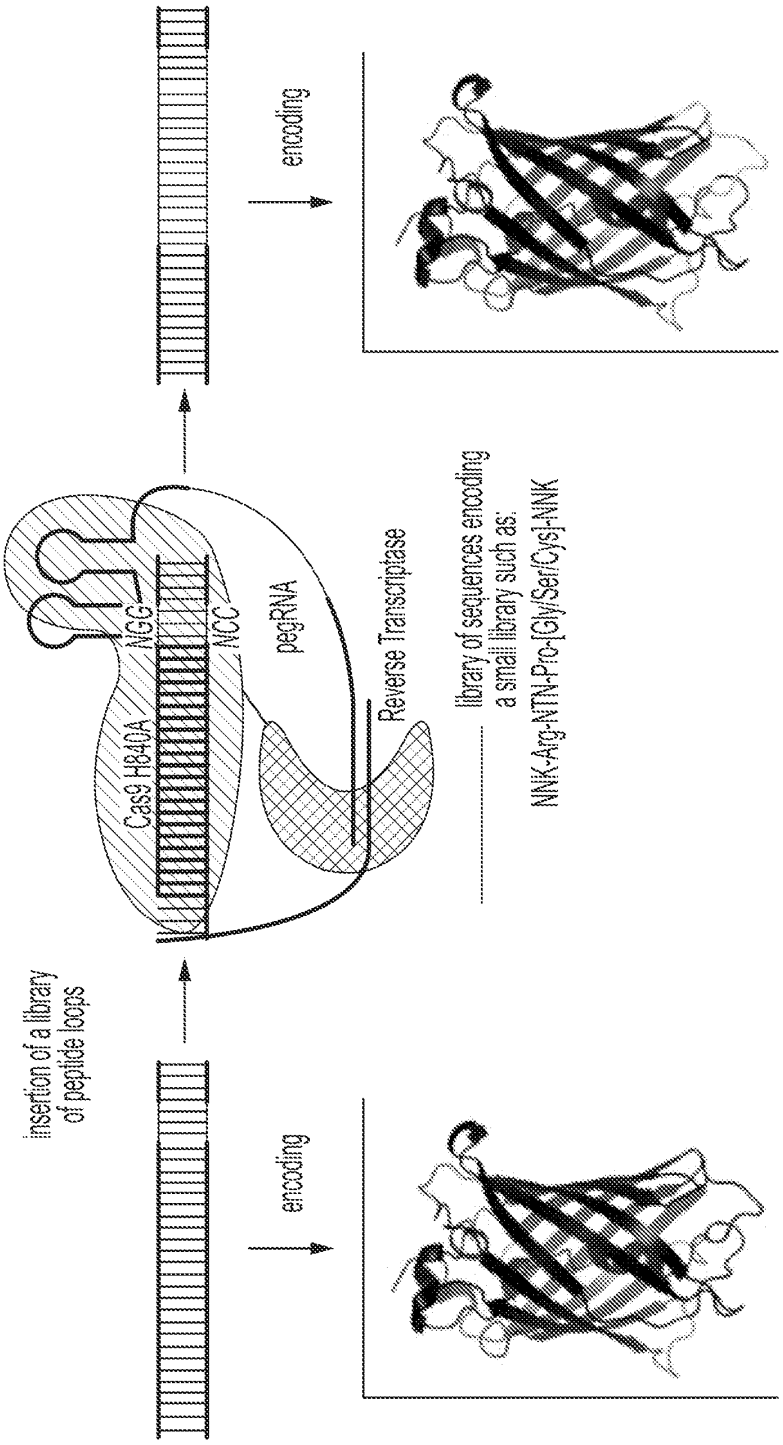

insertion of a library of peptide loops
encoding
encoding
Cas9 H840A
NGG
NCC
pegRNA
Reverse Transcriptase
library of sequences encoding a small library such as:
NNK-Arg-NTN-Pro-[Gly/Ser/Cys]-NNK

FIG. 29C

programmable codon deletion and truncations

SEQ ID NO:811	**GFP**	NFKIRHNIEDGSVQLADHYQQNTPI
SEQ ID NO:812	del5	NFKI-HNIEDGSVQLADHYQQNTPI
SEQ ID NO:813	del6	NFKIR-NIEDGSVQLADHYQQNTPI
SEQ ID NO:814	del8	NFKIRHN-EDGSVQLADHYQQNTPI
SEQ ID NO:815	del 8-13	NFKIRHN-----QLADHYQQNTPI
SEQ ID NO:816	del 1-5	-----HNIEDGSVQLADHYQQNTPI
SEQ ID NO:817	del 15-25	NFKIRHNIEDGSVQ----------

FIG. 29D

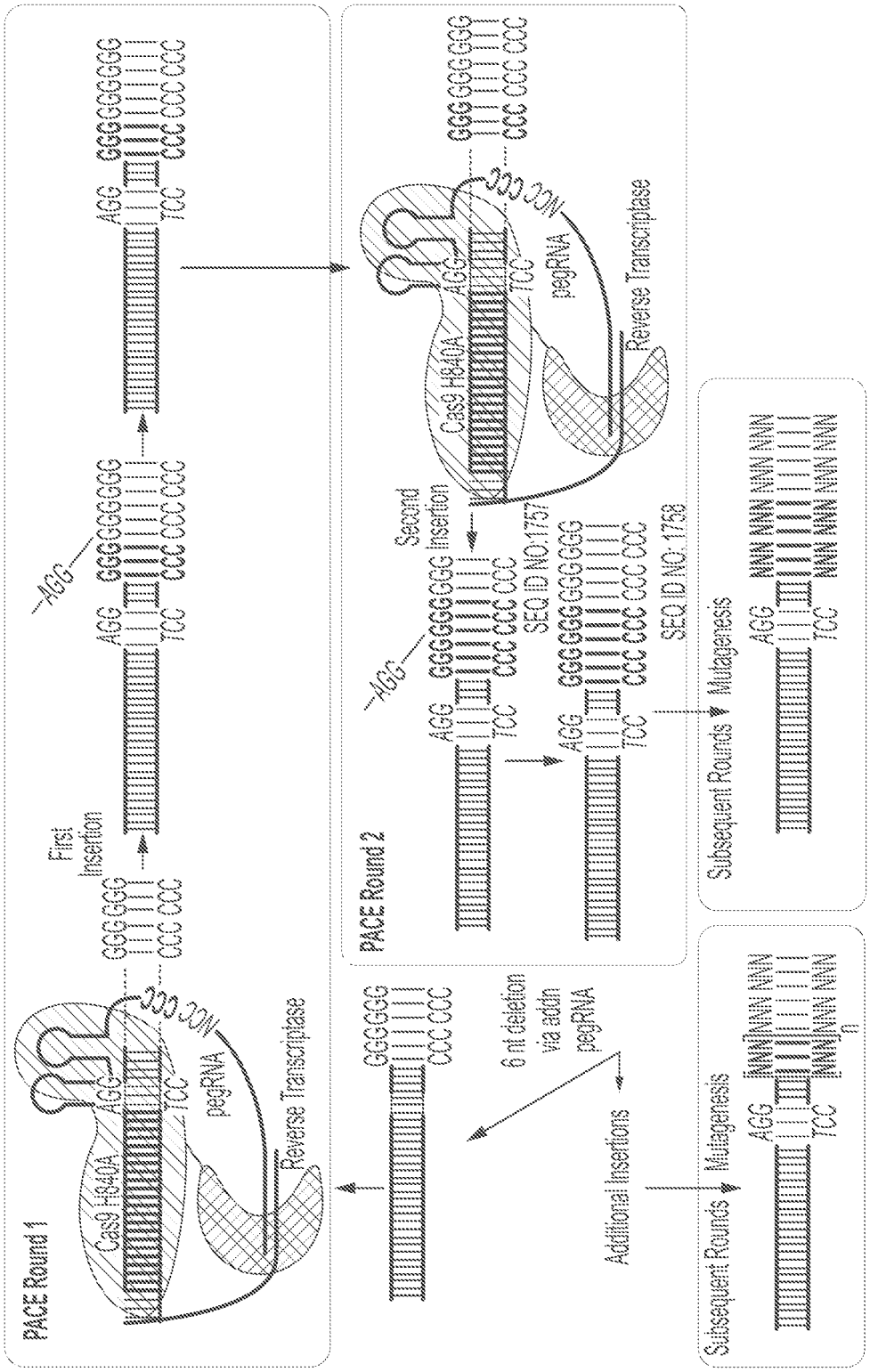

PACE Round 1
Cas9 H840A
pegRNA
Reverse Transcriptase
First Insertion
PACE Round 2
Second Insertion
SEQ ID NO:1757
SEQ ID NO:1758
6 nt deletion via addn pegRNA
Additional Insertions
Subsequent Rounds
Mutagenesis

FIG. 30

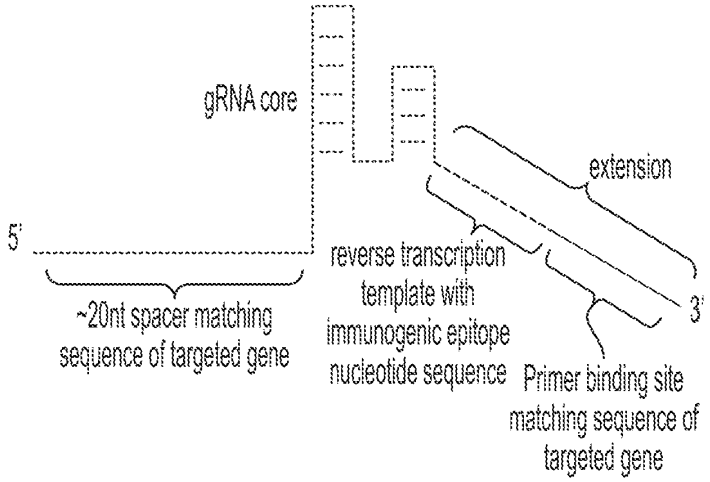

gRNA core
extension
5'
reverse transcription template with immunogenic epitope nucleotide sequence
3'
~20nt spacer matching sequence of targeted gene
Primer binding site matching sequence of targeted gene

FIG. 31

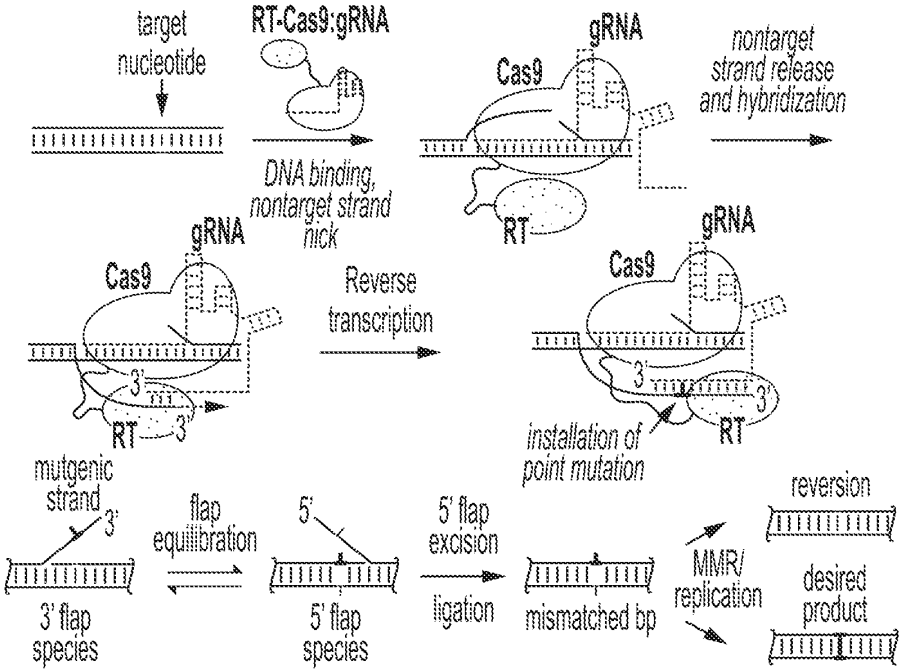

target nucleotide
RT-Cas9:gRNA
DNA binding, nontarget strand nick
gRNA
Cas9
RT
nontarget strand release and hybridization
Reverse transcription
installation of point mutation
mutgenic strand
3'
3' flap species
flap equilibration
5'
5' flap species
5' flap excision
ligation
mismatched bp
MMR/ replication
reversion
desired product

FIG. 32

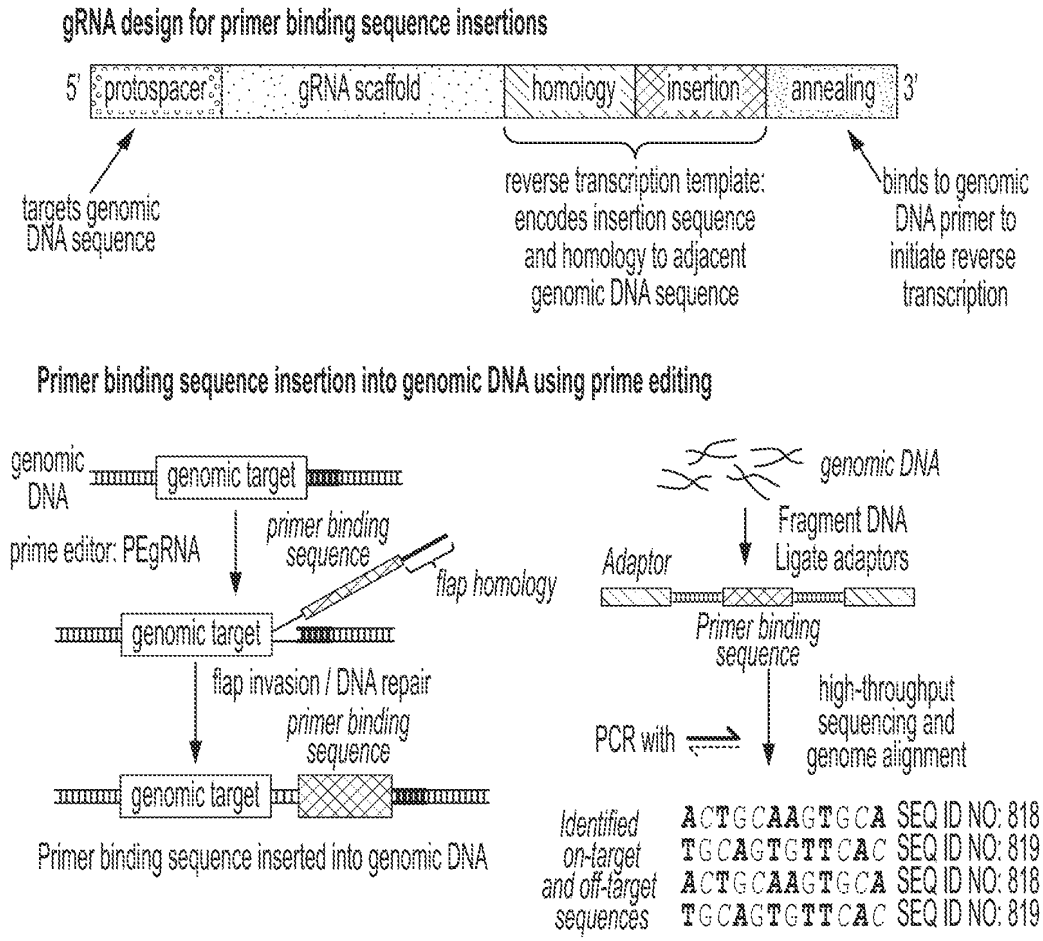

gRNA design for primer binding sequence insertions
5'
protospacer
gRNA scaffold
homology
insertion
annealing
3'
targets genomic DNA sequence
reverse transcription template: encodes insertion sequence and homology to adjacent genomic DNA sequence
binds to genomic DNA primer to initiate reverse transcription
Primer binding sequence insertion into genomic DNA using prime editing
genomic DNA
genomic target
prime editor: PEgRNA
primer binding sequence
flap homology
genomic target
flap invasion / DNA repair
primer binding sequence
genomic target
Primer binding sequence inserted into genomic DNA
genomic DNA
Fragment DNA
Ligate adaptors
Adaptor
Primer binding sequence
PCR with
high-throughput sequencing and genome alignment
Identified on-target and off-target sequences
ACTGCAAGTGCA SEQ ID NO: 818
TGCAGTGTTCAC SEQ ID NO: 819
ACTGCAAGTGCA SEQ ID NO: 818
TGCAGTGTTCAC SEQ ID NO: 819

FIG. 33

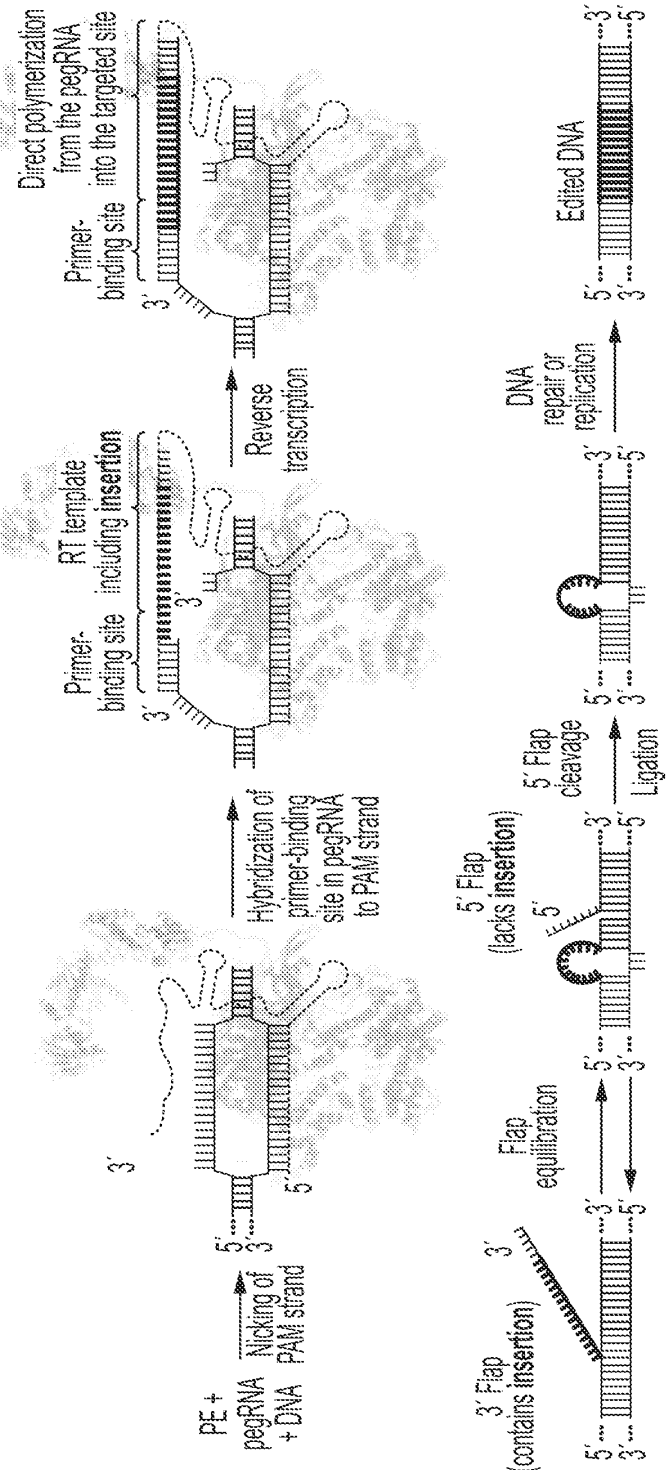

PE + pegRNA + DNA
Nicking of PAM strand
Hybridization of primer-binding site in pegRNA to PAM strand
Primer-binding site
RT template including insertion
Reverse transcription
Primer-binding site
Direct polymerization from the pegRNA into the targeted site
3' Flap (contains insertion)
Flap equilibration
5' Flap (lacks insertion)
5' Flap cleavage
Ligation
DNA repair or replication
Edited DNA

FIG. 34

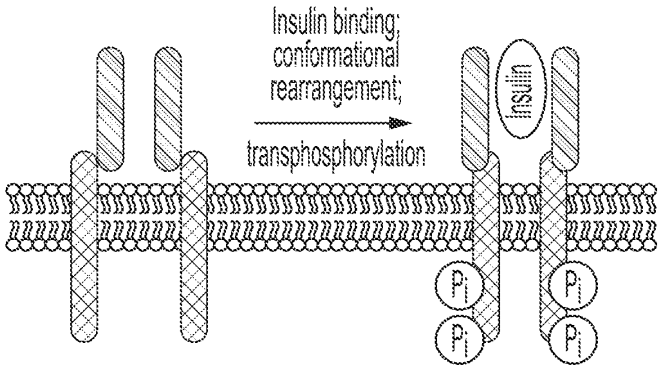


FIG. 35A

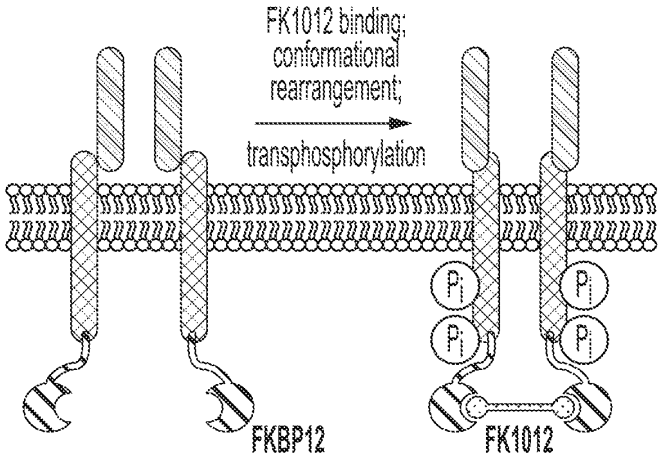


FIG. 35B

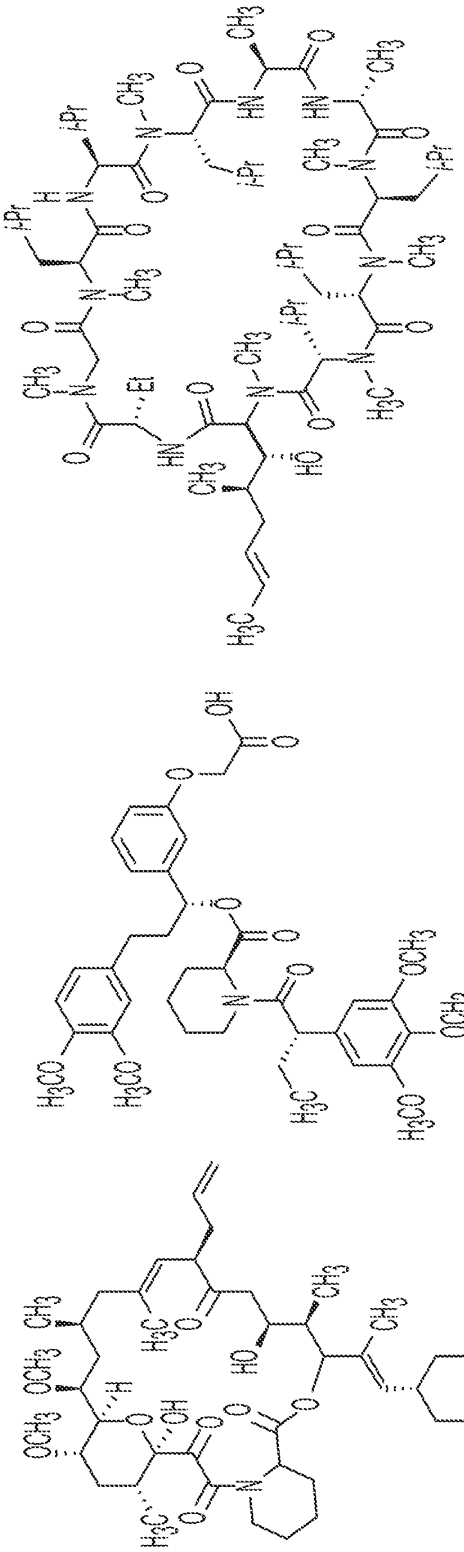

1
FK506
target: FKBP12
2
bumped FK506 mimic
target: FKBP12 F36V
3
cyclosporin A
target: cyclophilin

FIG. 36

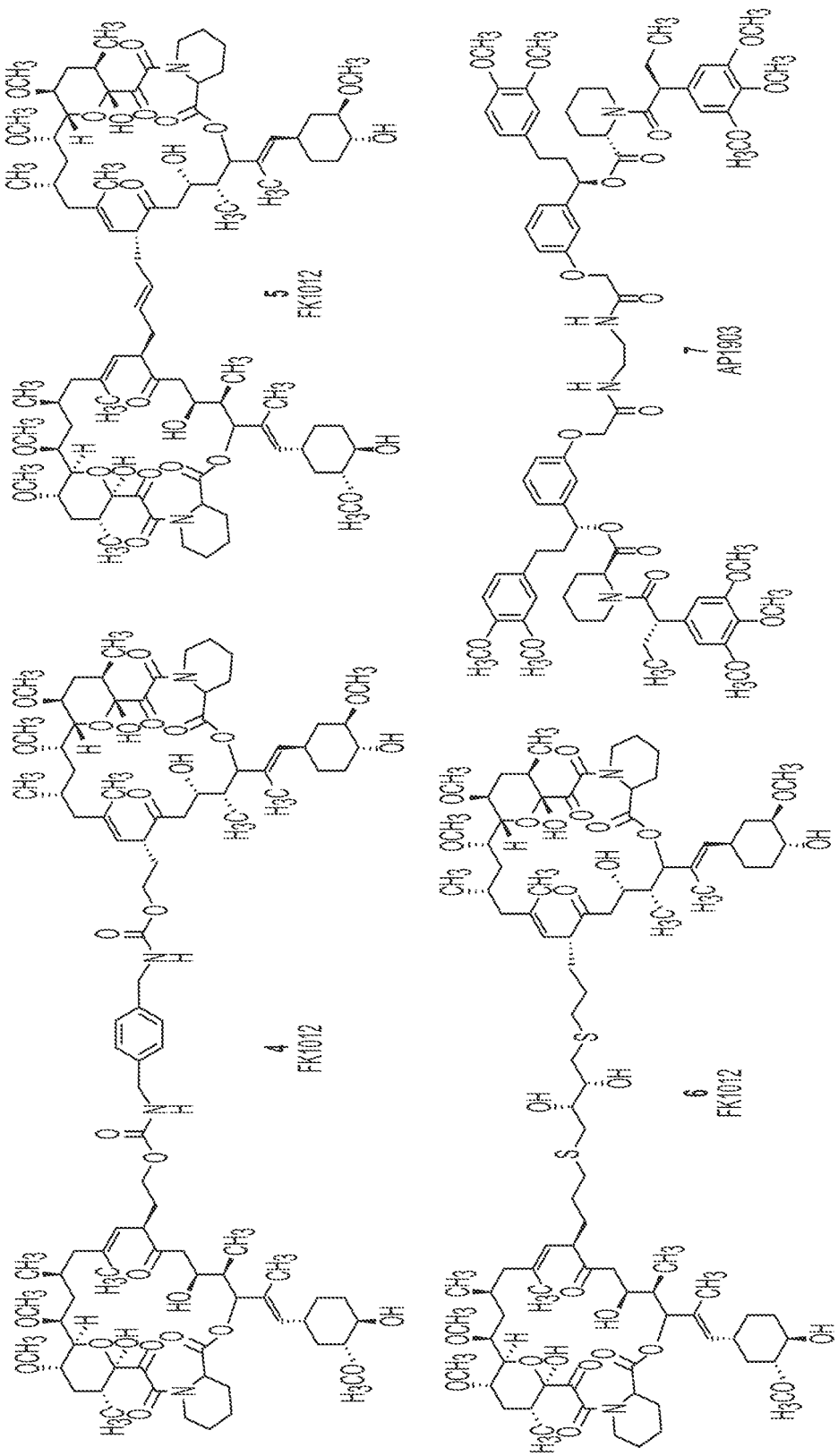


FIG. 37A

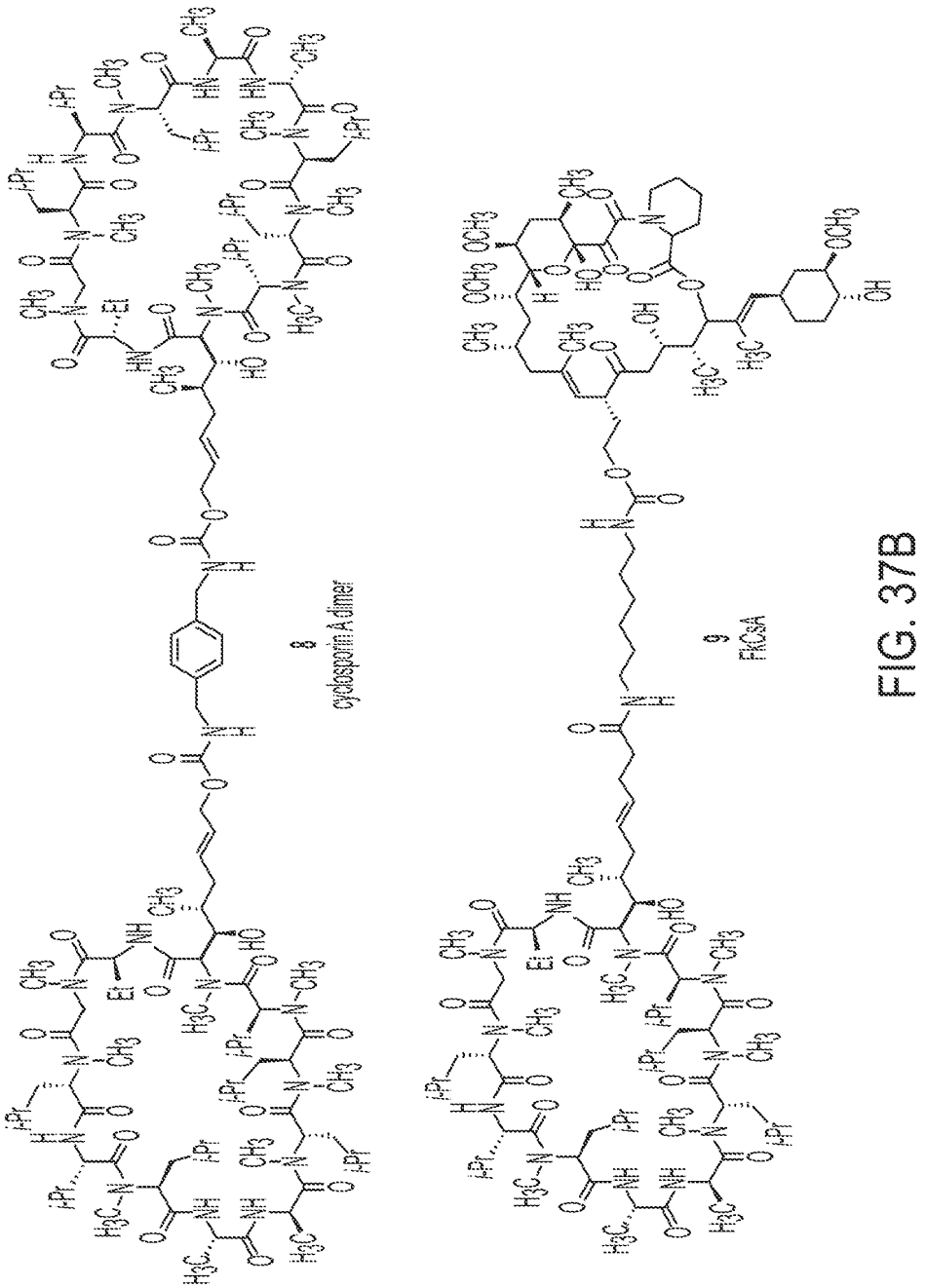


FIG. 37B

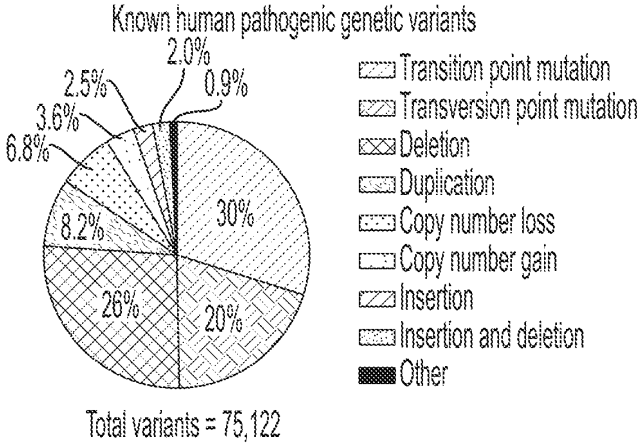


FIG. 38A

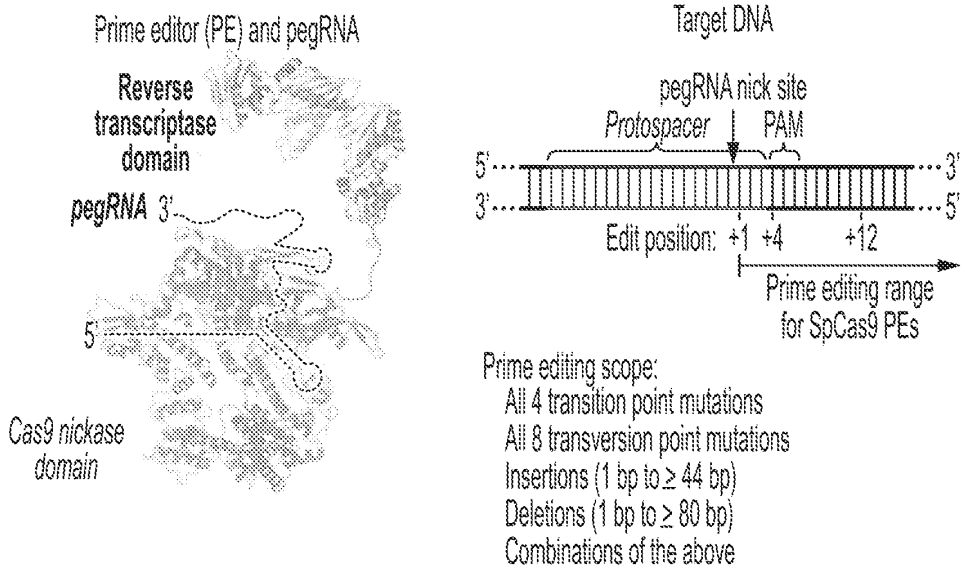


FIG. 38B

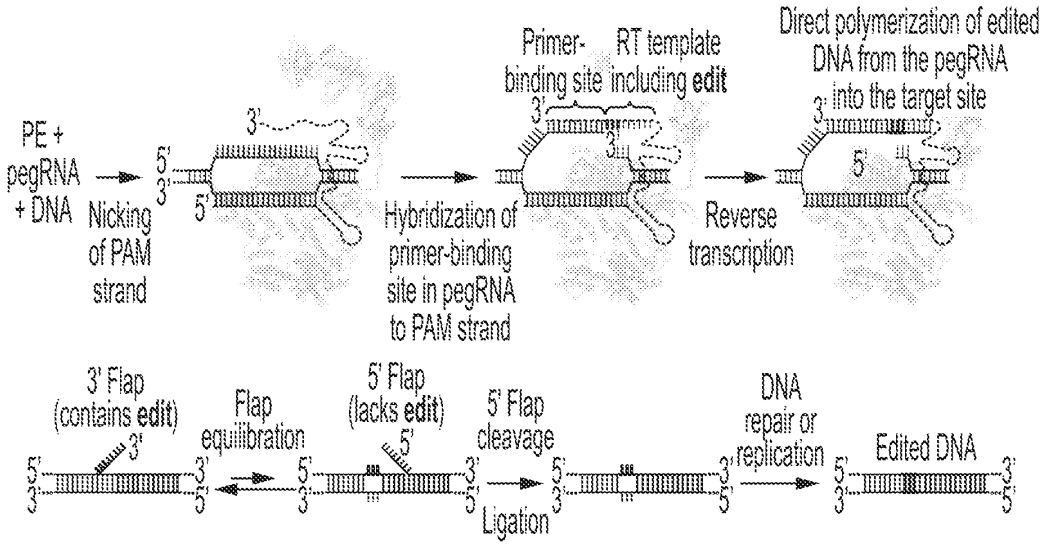


FIG. 38C

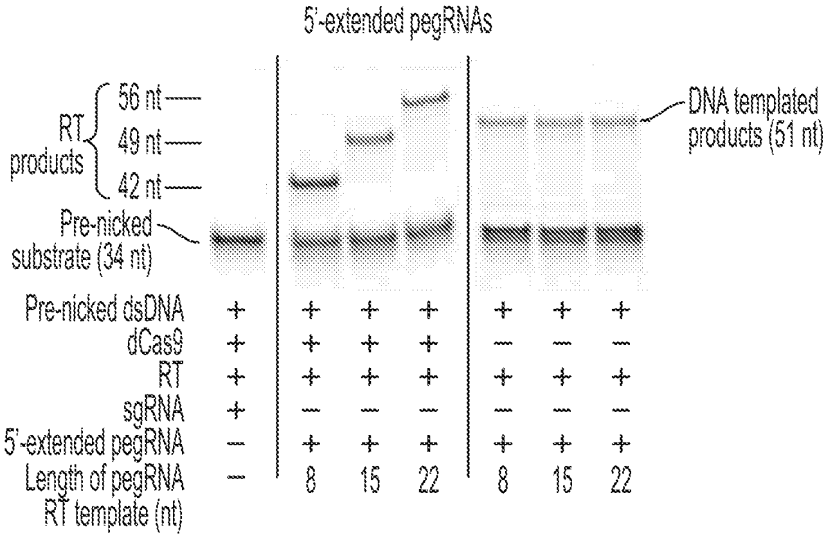


FIG. 38D

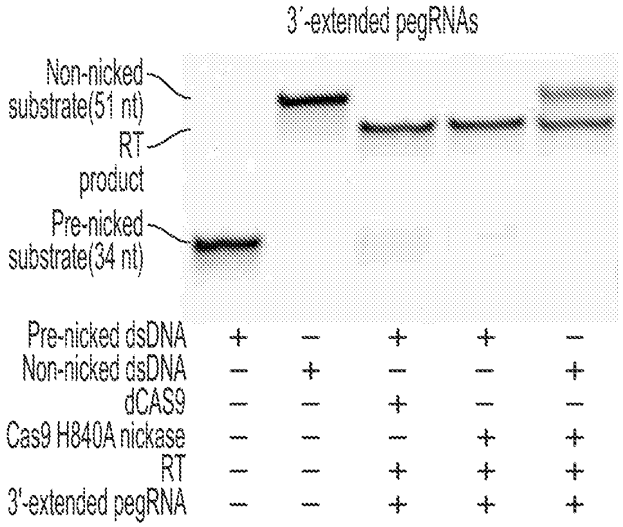


FIG. 38E

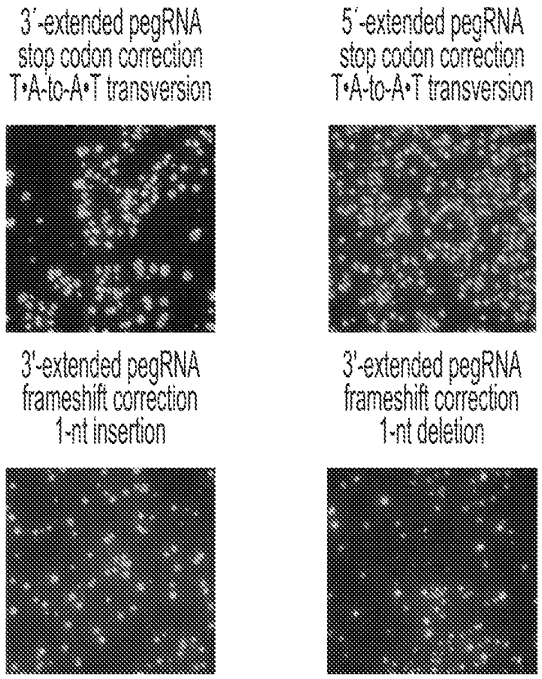


FIG. 38F

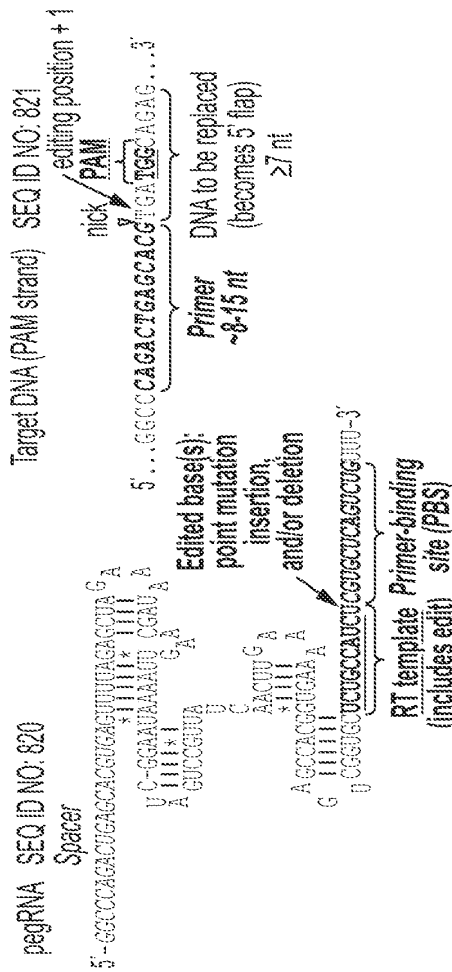


FIG. 39A

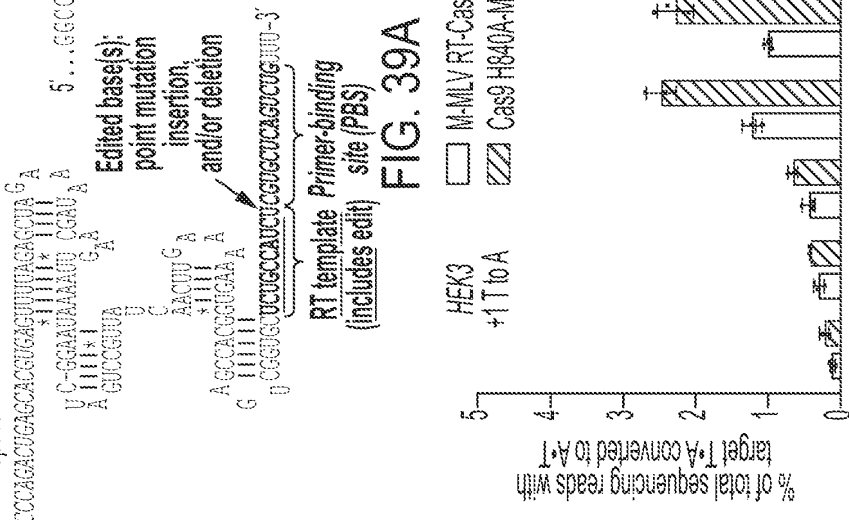


FIG. 39B

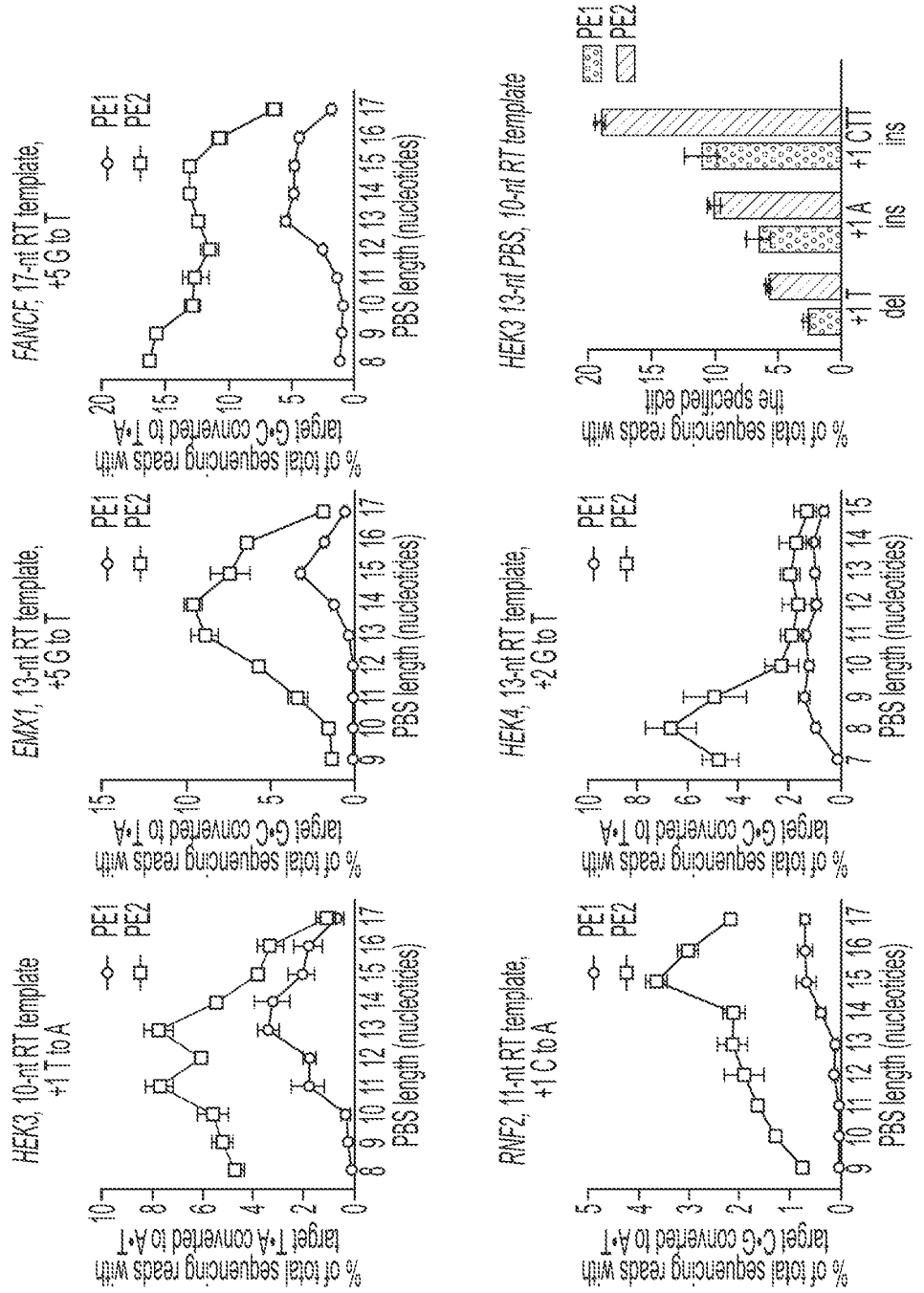

HEK3, 10-nt RT template, +1 T to A
EMX1, 13-nt RT template, +5 G to T
FANCF, 17-nt RT template, +5 G to T
RNF2, 11-nt RT template, +1 C to A
HEK4, 13-nt RT template, +2 G to T
HEK3 13-nt PBS, 10-nt RT template
PE1
PE2
% of total sequencing reads with target T•A converted to A•T
% of total sequencing reads with target G•C converted to T•A
% of total sequencing reads with target C•G converted to A•T
% of total sequencing reads with the specified edit
PBS length (nucleotides)
del
ins
+1T
+1A
+1CTT

FIG. 39C

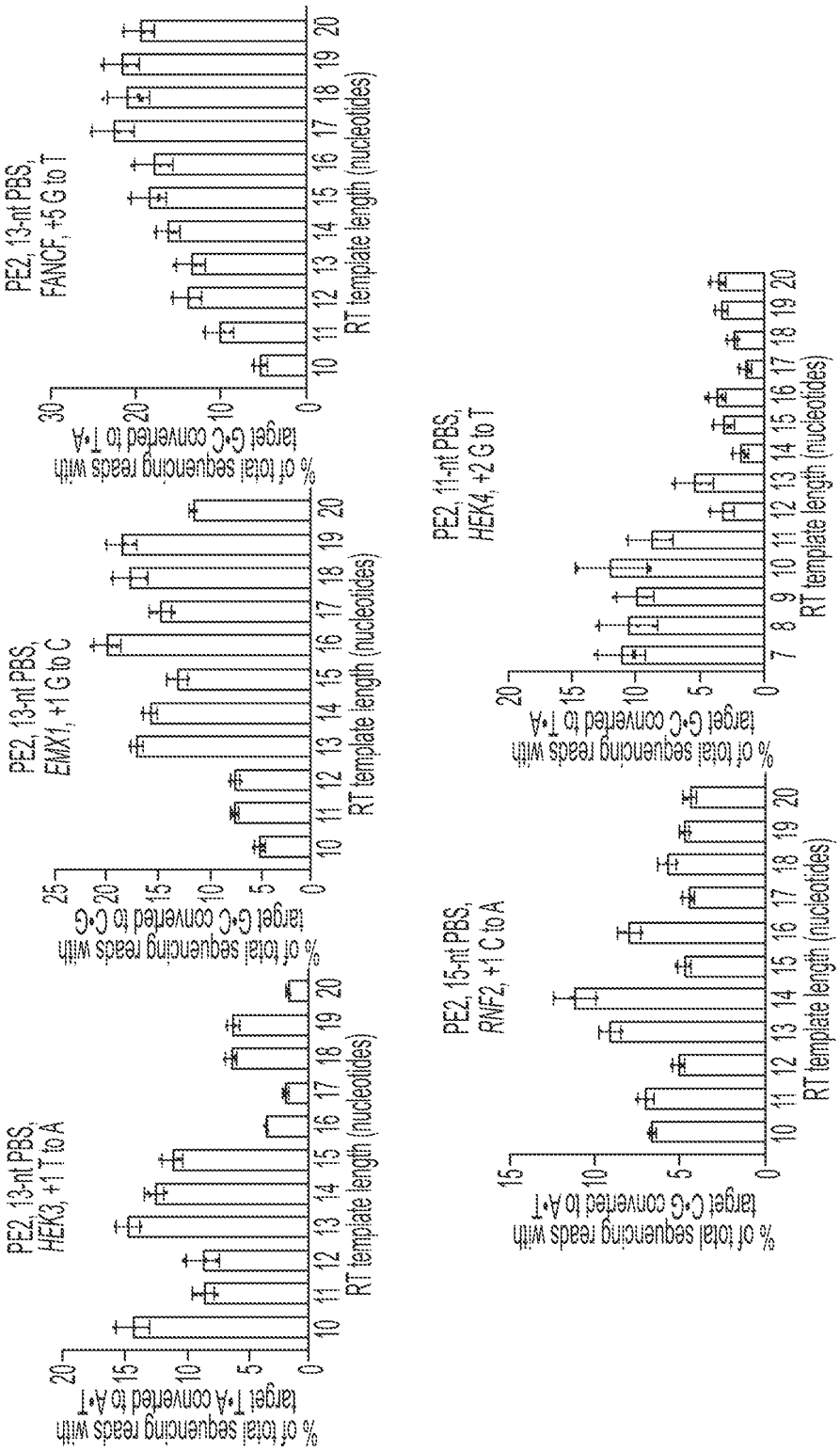

PE2, 13-nt PBS,
HEK3, +1 T to A
% of total sequencing reads with target T•A converted to A•T
RT template length (nucleotides)
PE2, 13-nt PBS,
EMX1, +1 G to C
% of total sequencing reads with target G•C converted to C•G
RT template length (nucleotides)
PE2, 13-nt PBS,
FANCF, +5 G to T
% of total sequencing reads with target G•C converted to T•A
RT template length (nucleotides)
PE2, 15-nt PBS,
RNF2, +1 C to A
% of total sequencing reads with target C•G converted to A•T
RT template length (nucleotides)
PE2, 11-nt PBS,
HEK4, +2 G to T
% of total sequencing reads with target G•C converted to T•A
RT template length (nucleotides)

FIG. 39D

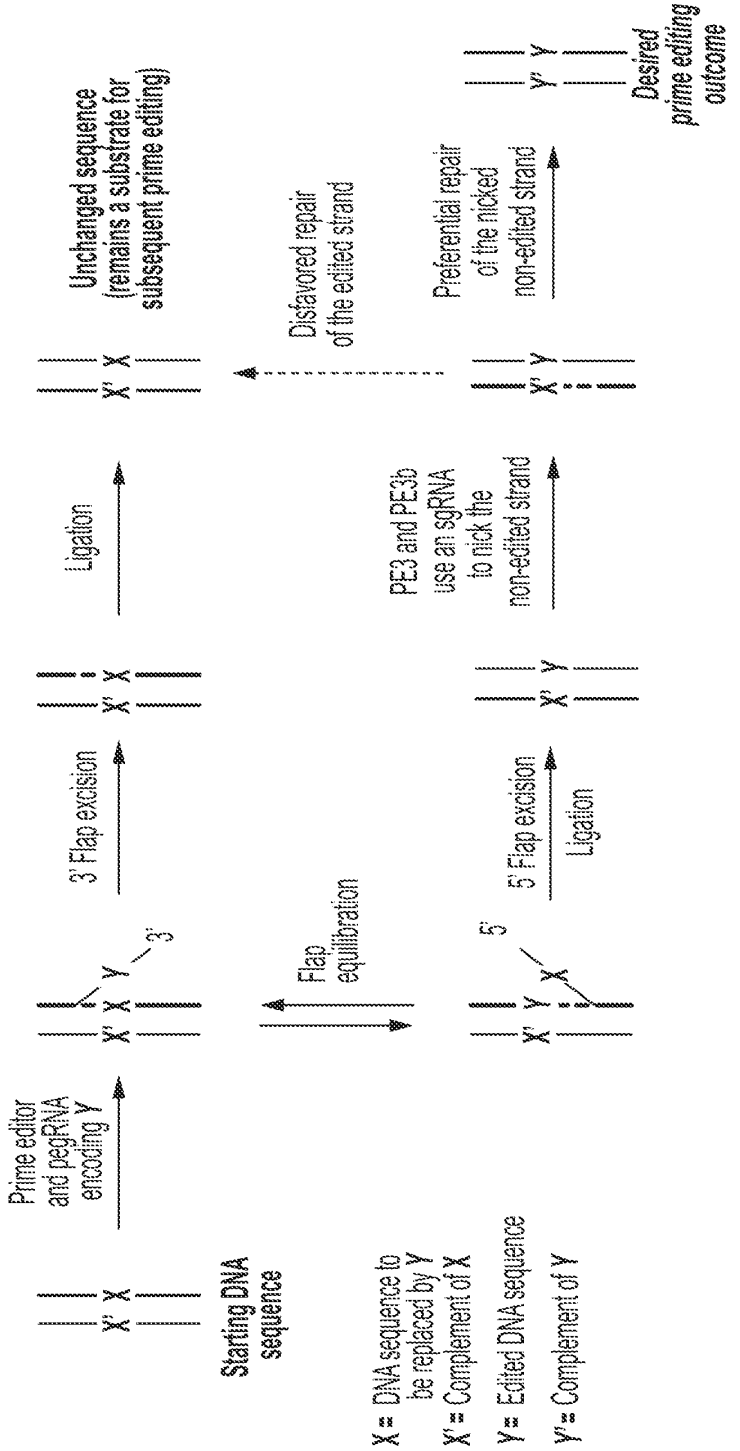

Starting DNA sequence
Prime editor and pegRNA encoding Y
3' Flap excision
Ligation
Unchanged sequence (remains a substrate for subsequent prime editing)
Disfavored repair of the edited strand
Flap equilibration
PE3 and PE3b use an sgRNA to nick the non-edited strand
Preferential repair of the nicked non-edited strand
Desired prime editing outcome
5' Flap excision
Ligation
X = DNA sequence to be replaced by Y
X' = Complement of X
Y = Edited DNA sequence
Y' = Complement of Y

FIG. 40A

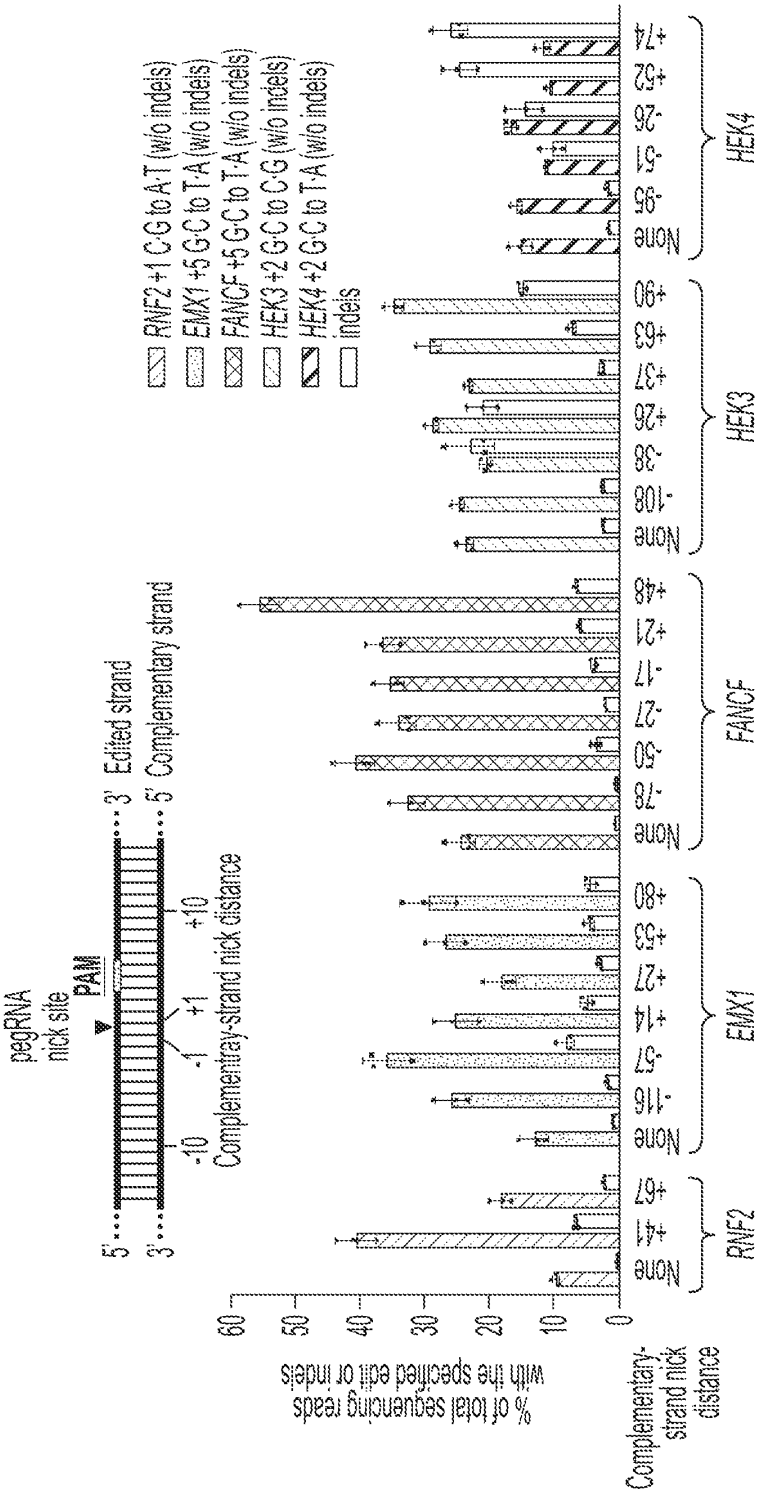

pegRNA
nick site
PAM
5'
3'
3' Edited strand
5' Complementary strand
-10
-1
+1
+10
Complementary-strand nick distance
RNF2 +1 C·G to A·T (w/o indels)
EMX1 +5 G·C to T·A (w/o indels)
FANCF +5 G·C to T·A (w/o indels)
HEK3 +2 G·C to C·G (w/o indels)
HEK4 +2 G·C to T·A (w/o indels)
indels
% of total sequencing reads with the specified edit or indels
0
10
20
30
40
50
60
Complementary-strand nick distance
None
+41
+67
RNF2
None
-116
-57
+14
+27
+53
+80
EMX1
None
-78
-50
-27
-17
+21
+48
FANCF
None
-108
-38
+26
+37
+63
+90
HEK3
None
-95
-51
-28
+52
+74
HEK4

FIG. 40B

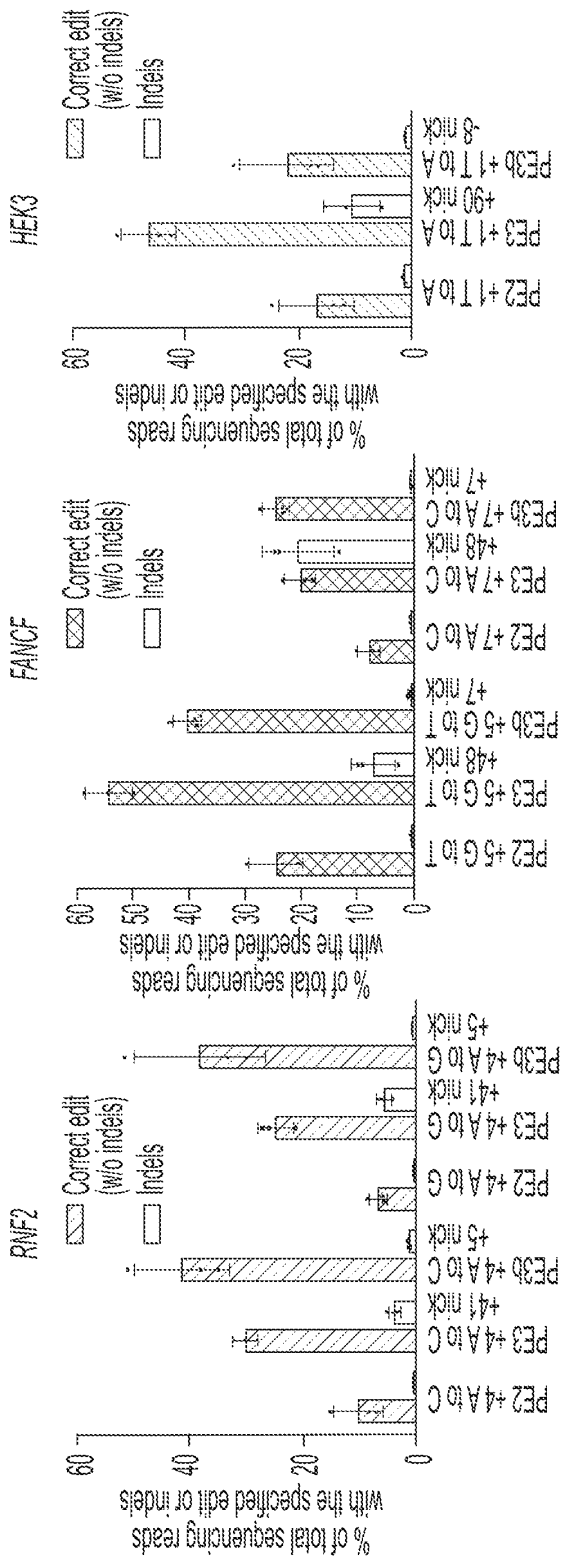

RNF2
FANCF
HEK3
Correct edit (w/o indels)
Indels
% of total sequencing reads with the specified edit or indels
PE2 +4 A to C
PE3 +4 A to C
+41 nick
PE3b +4 A to C
+5 nick
PE2 +4 A to G
PE3 +4 A to G
+41 nick
PE3b +4 A to G
+5 nick
PE2 +5 G to T
PE3 +5 G to T
+48 nick
PE3b +5 G to T
+7 nick
PE2 +7 A to C
PE3 +7 A to C
+48 nick
PE3b +7 A to C
+7 nick
PE2 +1 T to A
PE3 +1 T to A
+90 nick
PE3b +1 T to A
-8 nick

FIG. 40C

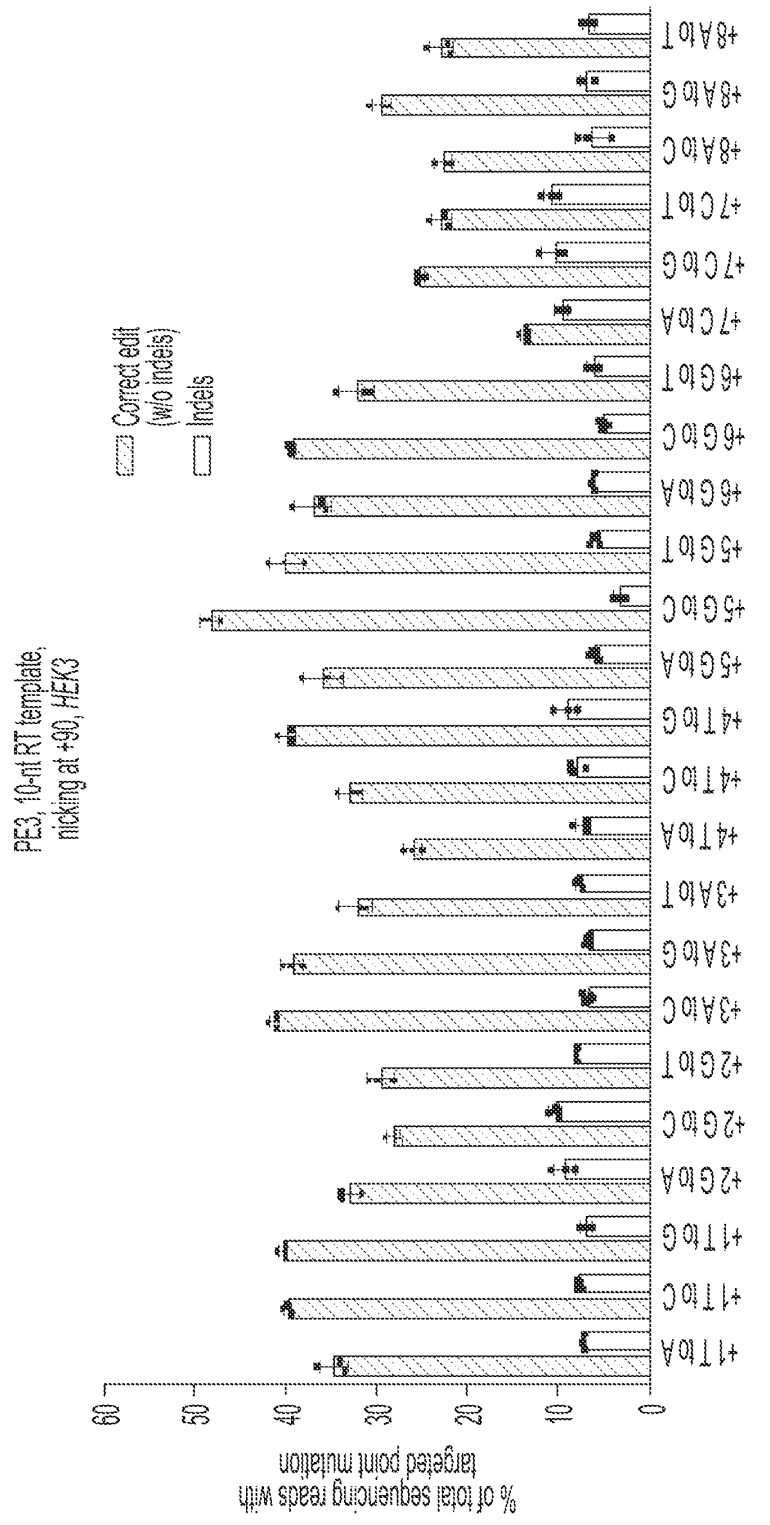

PE3, 10-nt RT template, nicking at +90, HEK3
Correct edit (w/o indels)
Indels
% of total sequencing reads with targeted point mutation
0
10
20
30
40
50
60
+1T to A
+1T to C
+1T to G
+2G to A
+2G to C
+2G to T
+3A to C
+3A to G
+3A to T
+4T to A
+4T to C
+4T to G
+5G to A
+5G to C
+5G to T
+6G to A
+6G to C
+6G to T
+7C to A
+7C to G
+7C to T
+8A to C
+8A to G
+8A to T

FIG. 41A

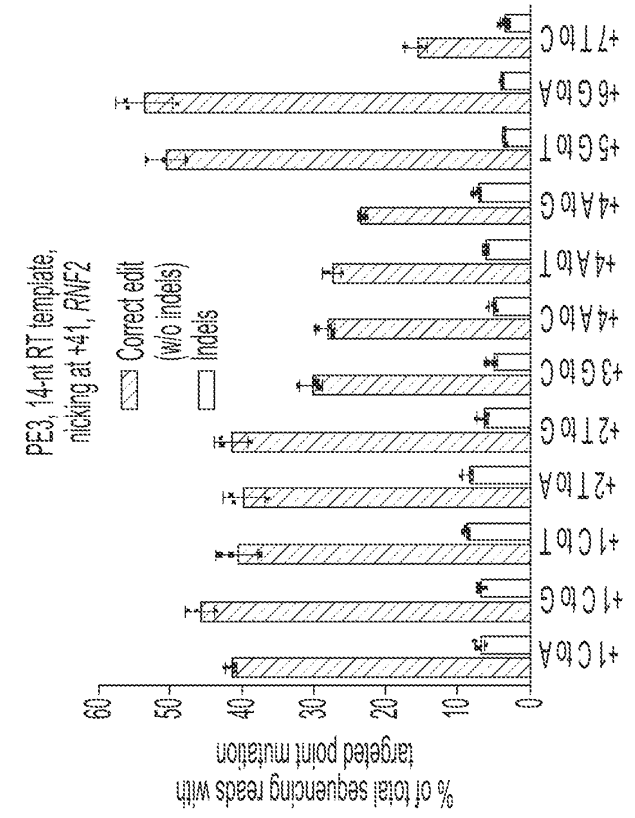


FIG. 41C

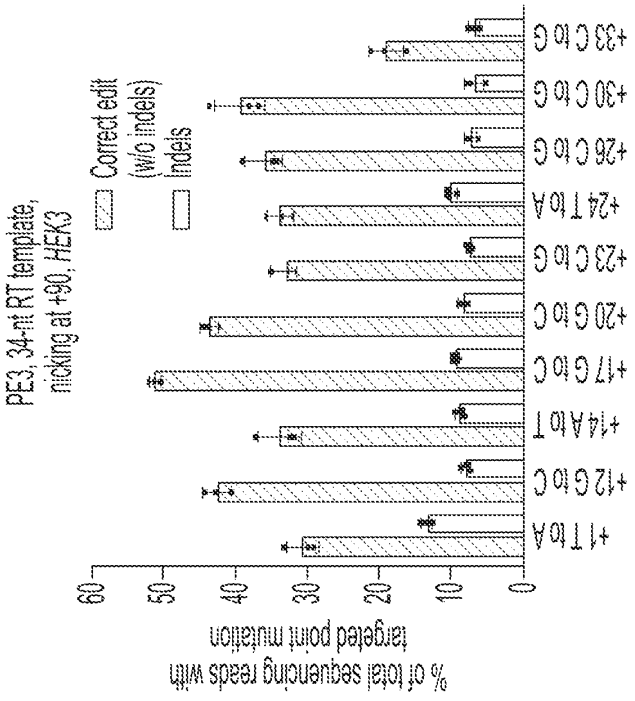


FIG. 41B

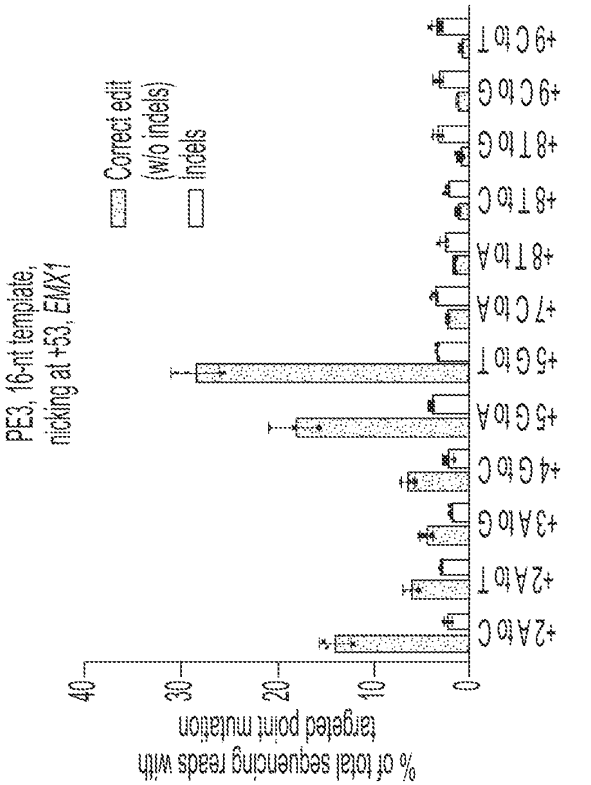


FIG. 41E

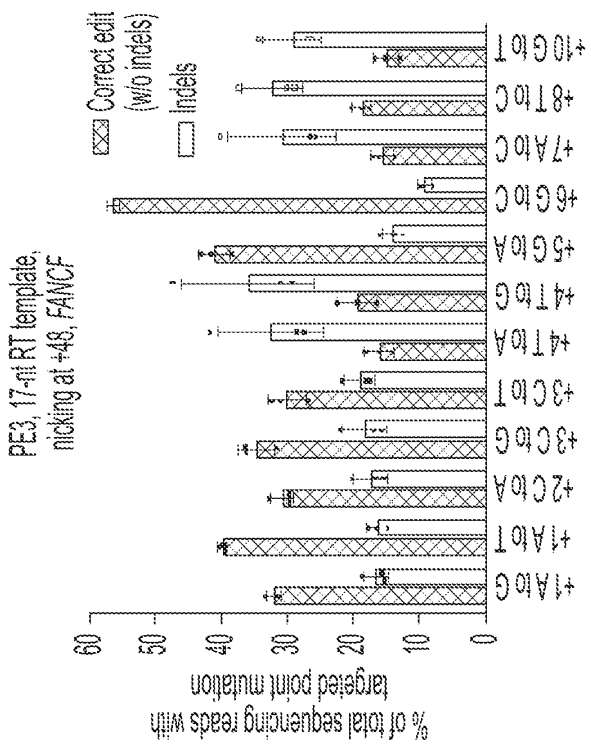


FIG. 41D

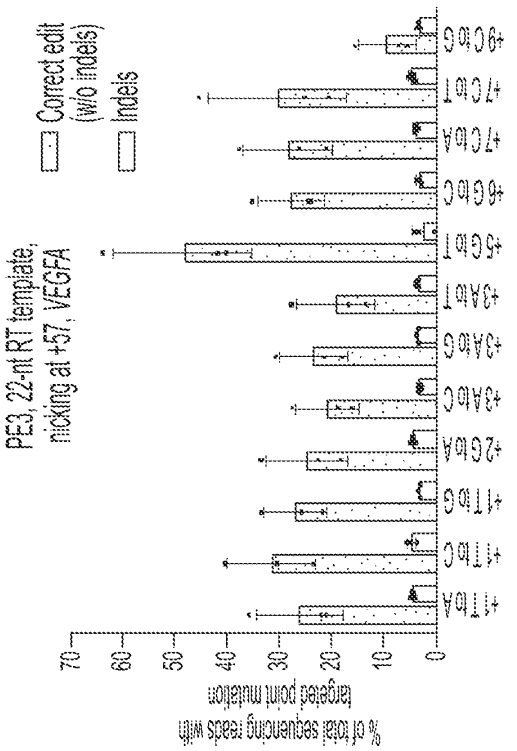


FIG. 41G

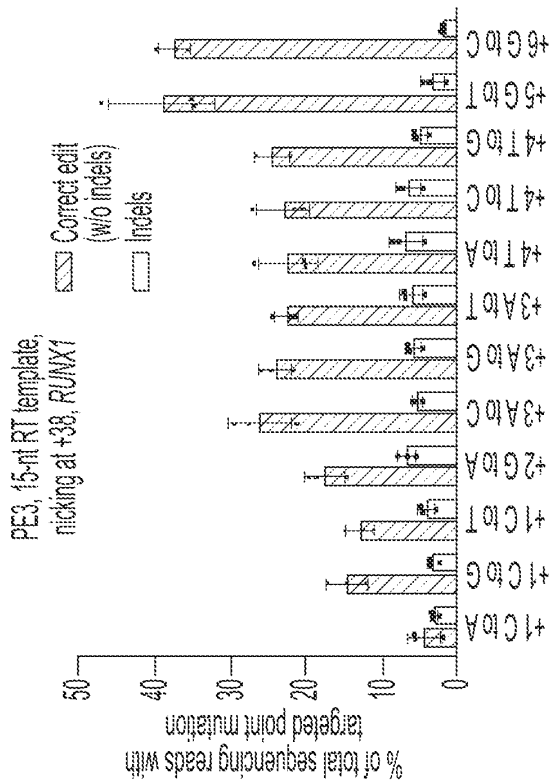


FIG. 41F

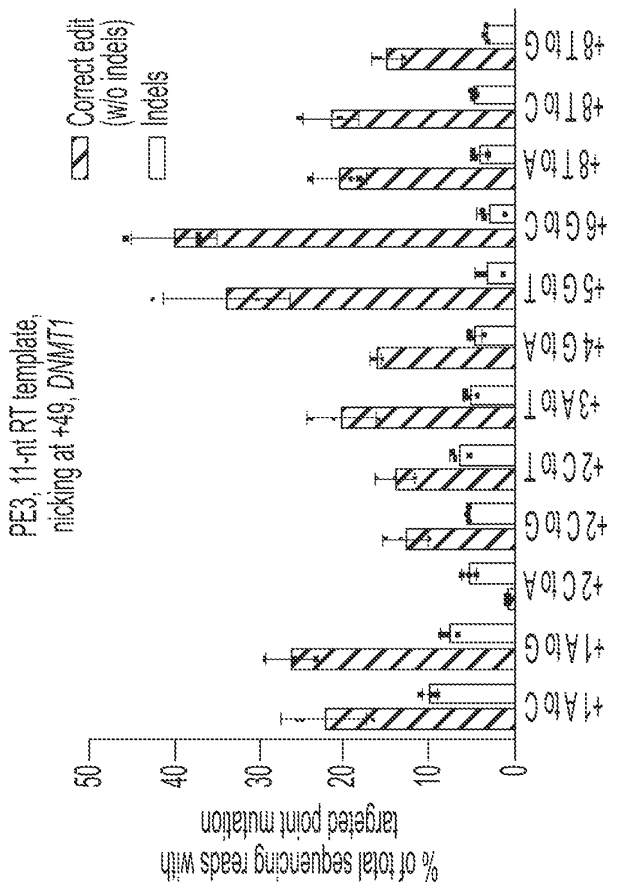

PE3, 11-nt RT template,
nicking at +49, DNMT1
Correct edit
(w/o indels)
Indels
% of total sequencing reads with
targeted point mutation
0
10
20
30
40
50
+1A to C
+1A to G
+2C to A
+2C to G
+2C to T
+3A to T
+4G to A
+5G to T
+6G to C
+8T to A
+8T to C
+8T to G

FIG. 41H

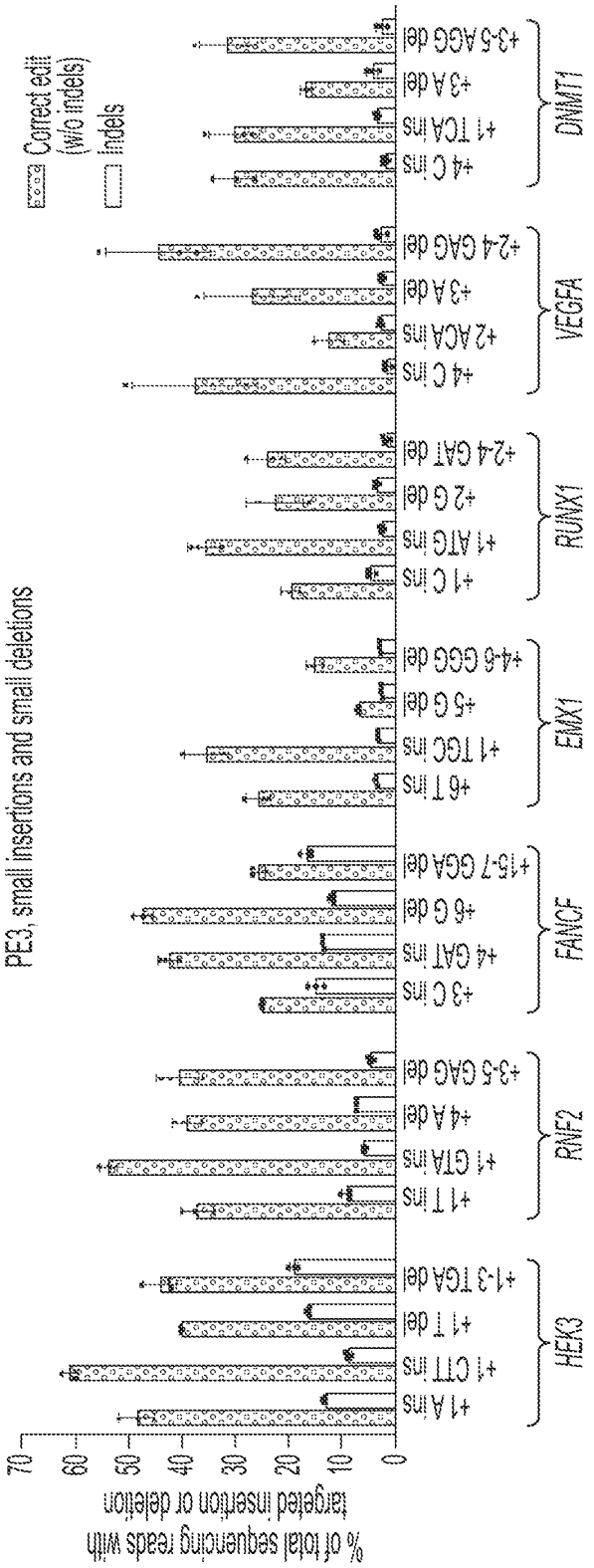

PE3, small insertions and small deletions
% of total sequencing reads with targeted insertion or deletion
0
10
20
30
40
50
60
70
Correct edit (w/o indels)
Indels
+1 A ins
+1 CTT ins
+1 T del
+1-3 TGA del
HEK3
+1 T ins
+1 GTA ins
+4 A del
+3-5 GAG del
RNF2
+3 C ins
+4 GAT ins
+6 G del
+15-7 GGA del
FANCF
+6 T ins
+1 TGC ins
+5 G del
+4-6 GGG del
EMX1
+1 C ins
+1 ATG ins
+2 G del
+2-4 GAT del
RUNX1
+4 C ins
+2 ACA ins
+3 A del
+2-4 GAG del
VEGFA
+4 C ins
+1 TCA ins
+3 A del
+3-5 AGG del
DNMT1

FIG. 411

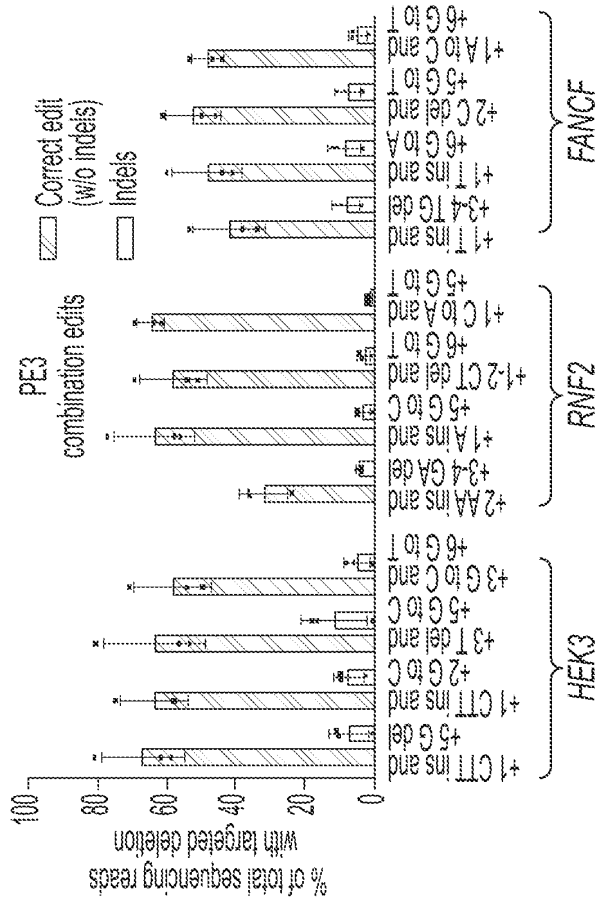


FIG. 41J

FIG. 41K

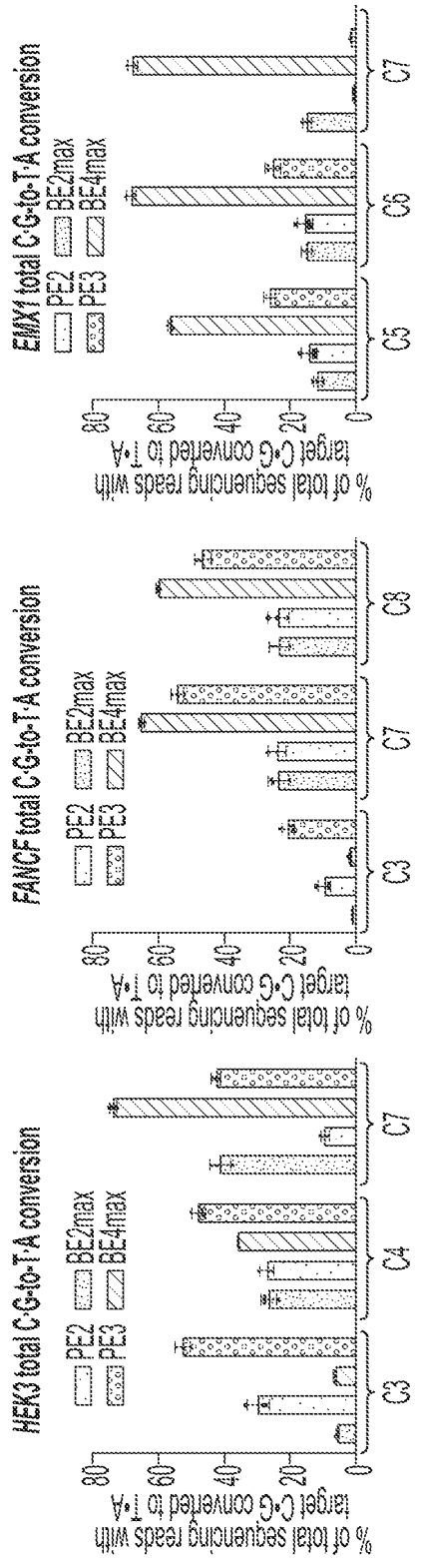


FIG. 42A

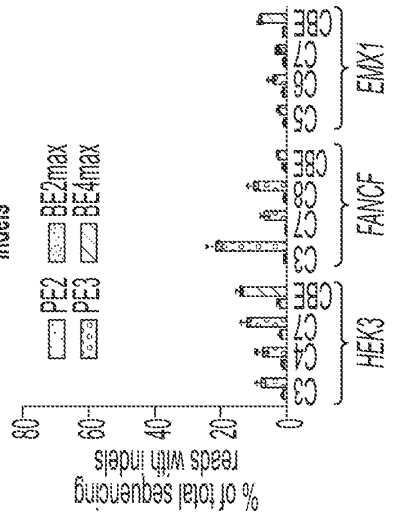


FIG. 42B

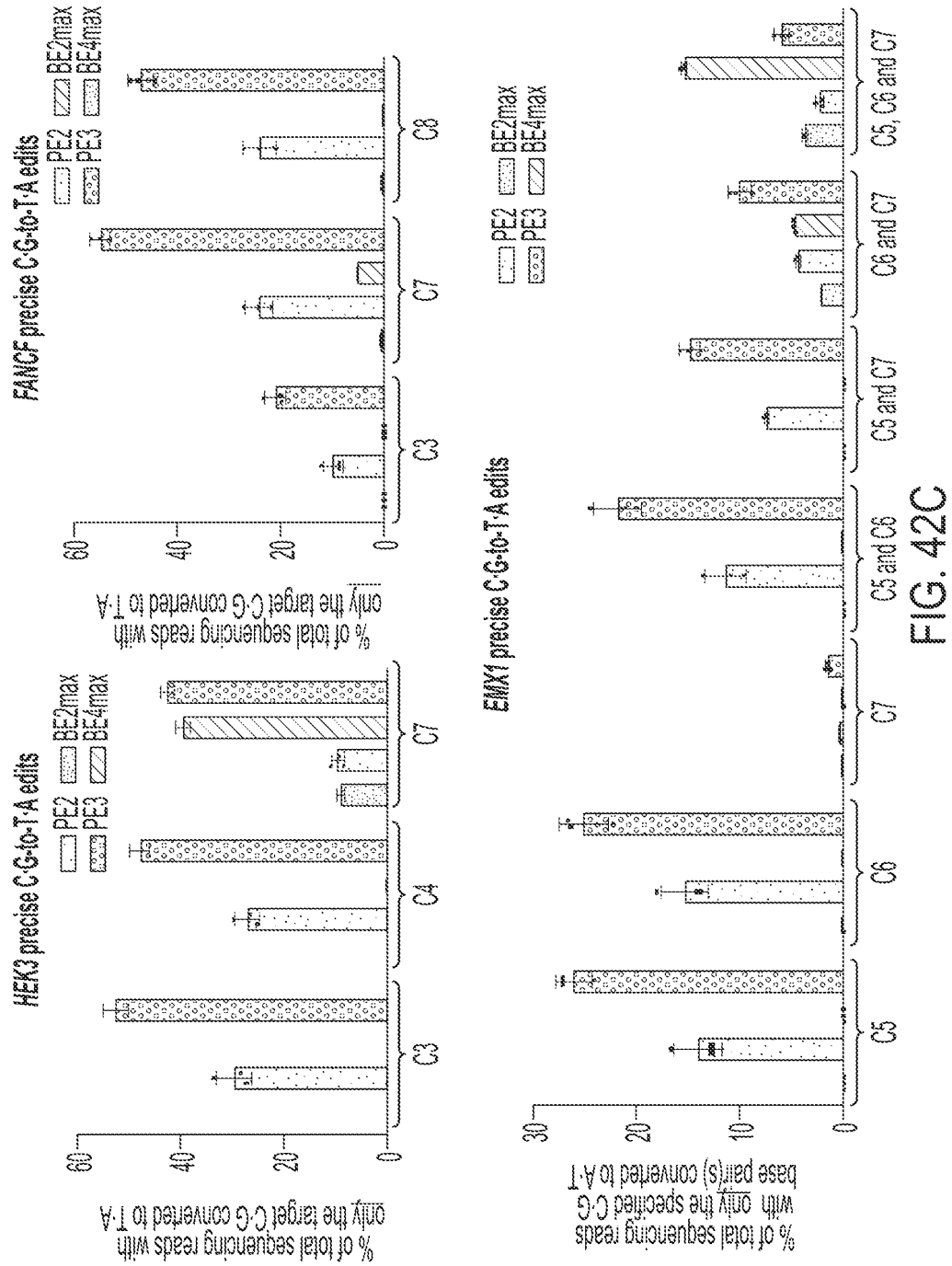

HEK3 precise C·G-to-T·A edits
PE2
PE3
BE2max
BE4max
% of total sequencing reads with only the target C·G converted to T·A
0
20
40
60
C3
C4
C7
FANCF precise C·G-to-T·A edits
C8
EMX1 precise C·G-to-T·A edits
% of total sequencing reads with only the specified C·G base pair(s) converted to A·T
0
10
20
30
C5
C6
C5 and C6
C5 and C7
C6 and C7
C5, C6 and C7

FIG. 42C

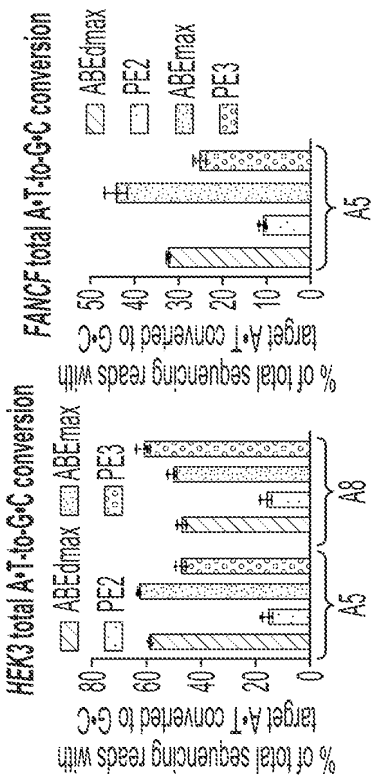


FIG. 42D

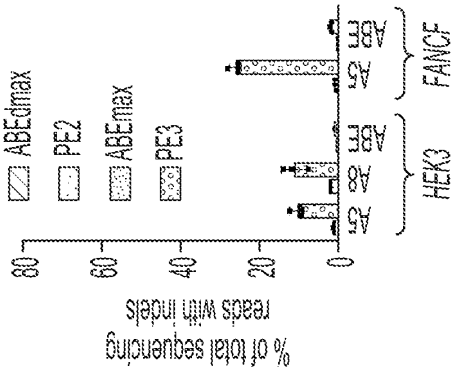


FIG. 42F

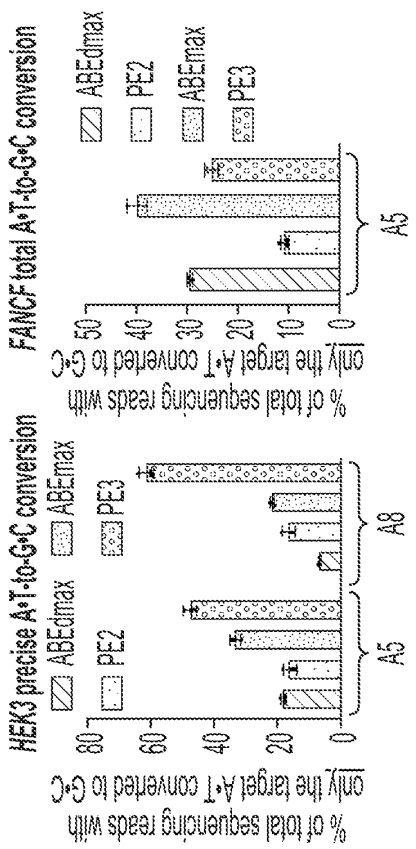


FIG. 42E

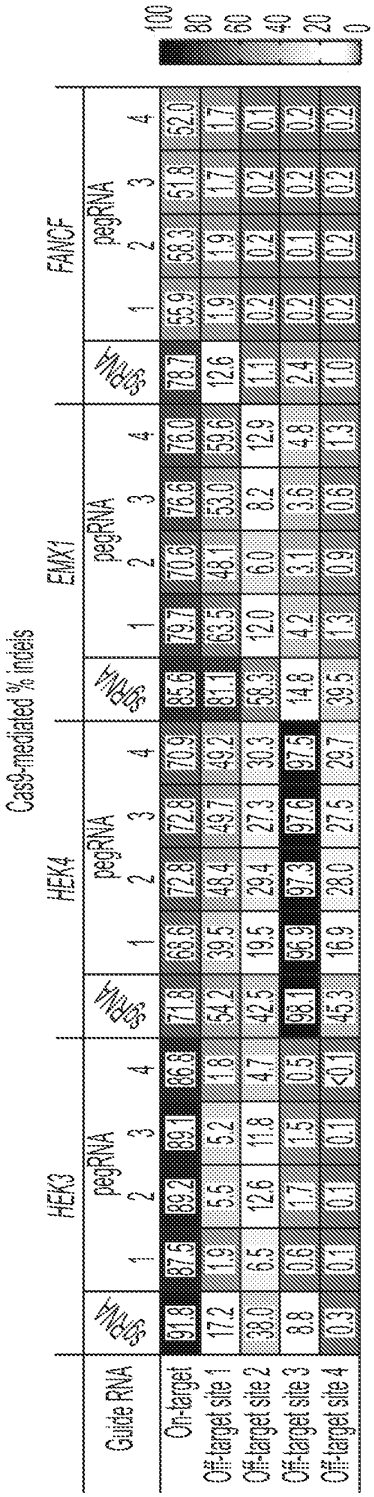

Cas9-mediated % indels
HEK3
HEK4
EMX1
FANCF
Guide RNA
sgRNA
pegRNA
1
2
3
4
On-target
Off-target site 1
Off-target site 2
Off-target site 3
Off-target site 4
91.8
87.5
89.2
89.1
66.6
71.8
69.6
72.8
72.8
70.9
85.6
79.7
70.6
76.6
76.0
78.7
55.9
58.3
61.8
52.0
17.2
1.9
5.5
5.2
1.8
54.2
39.5
48.4
49.7
49.2
81.1
63.5
48.1
53.0
59.6
12.6
1.9
1.9
1.7
1.7
30.0
6.5
12.6
11.8
4.7
42.5
19.5
29.4
27.3
30.3
58.3
12.0
6.0
8.2
12.9
1.1
0.2
0.2
0.2
0.1
8.8
0.6
1.7
1.5
0.5
98.1
96.9
97.3
97.6
97.5
14.8
4.2
3.1
3.6
4.8
2.4
0.2
0.1
0.2
0.2
0.3
0.1
0.1
0.1
<0.1
45.3
16.9
28.0
27.6
29.7
39.5
1.3
0.9
0.6
1.3
1.0
0.2
0.2
0.2
0.2
100
80
60
40
20
0

FIG. 42G

PE-mediated % editing and % indels (in parentheses)

Guide RNA	*HEK3* (PE3) pegRNA 1	2	3	4	*HEK4* (PE2) pegRNA 1	2	3	4	*EMX1* (PE3) pegRNA 1	2	3	4	*FANCF* (PE3) pegRNA 1	2	3	4
On-target	44.2 (11.9)	61.2 (8.8)	40.4 (16.5)	48.4 (3.3)	18.2 (0.9)	14.4 (1.8)	9.8 (2.0)	7.9 (2.2)	28.6 (3.5)	14.1 (2.4)	35.7 (3.3)	15.4 (2.9)	56.9 (9.3)	32.4 (16.7)	42.8 (13.6)	47.6 (12.0)
Off-target site 1	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	0.4 (<0.1)	<0.1 (<0.1)	0.4 (<0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (0.1)	0.7 (0.1)	<0.1 (0.1)	<0.1 (0.1)
Off-target site 2	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (<0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (0.1)	<0.1 (<0.1)
Off-target site 3	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	0.2 (0.5)	6.8 (1.9)	19.2 (0.5)	7.9 (3.5)	<0.1 (0.3)	<0.1 (0.3)	<0.1 (0.3)	<0.1 (0.3)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)
Off-target site 4	<0.1 (0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (0.1)	<0.1 (0.2)	<0.1 (<0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (0.2)	<0.1 (0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)

Color scale: 0, 20, 40, 60, 80, 100

FIG. 42H

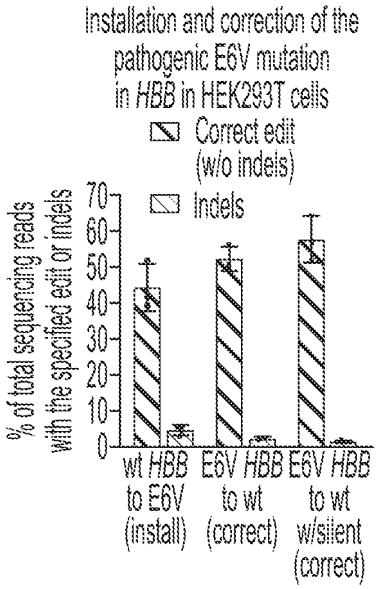


FIG. 43A

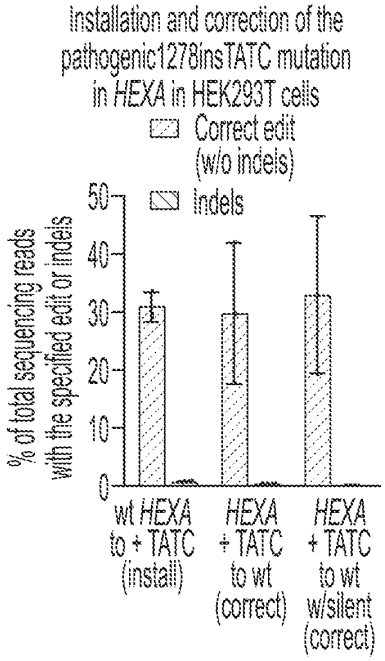


FIG. 43B

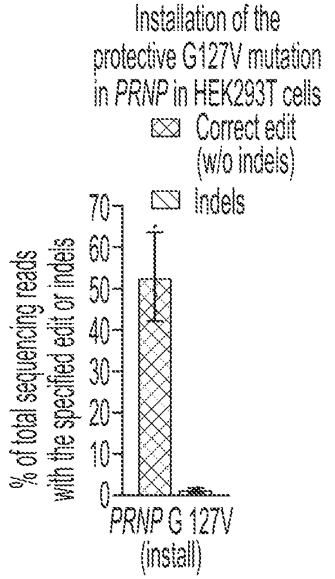


FIG. 43C

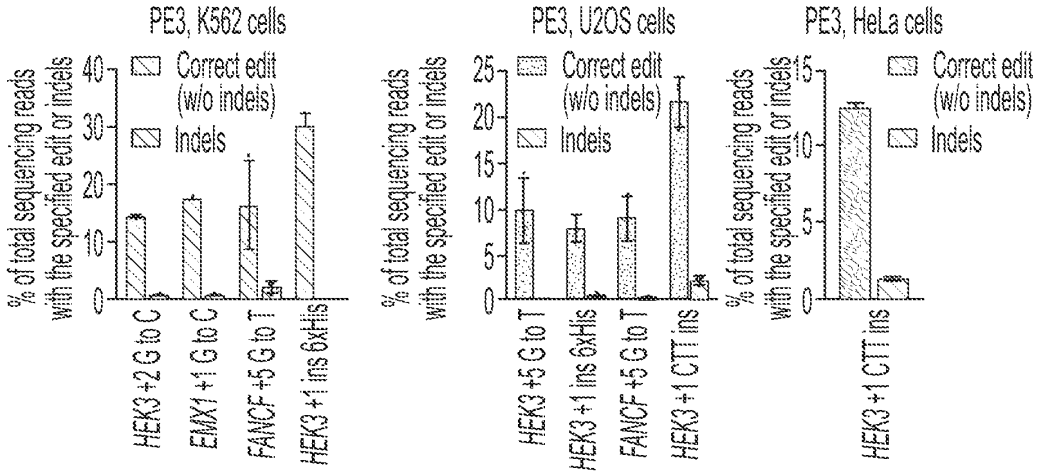


FIG. 43D

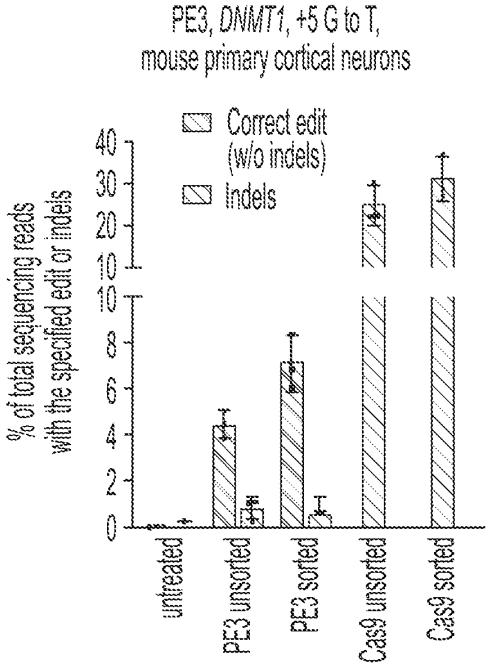


FIG. 43E

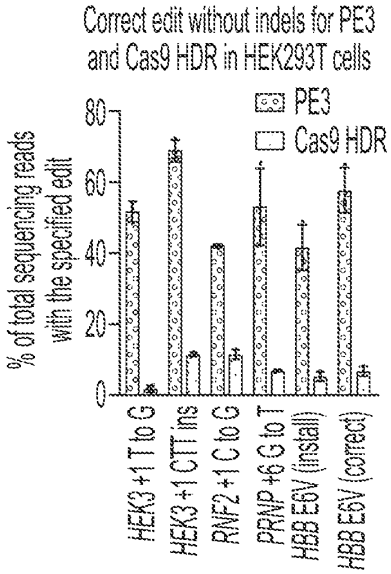


FIG. 43F

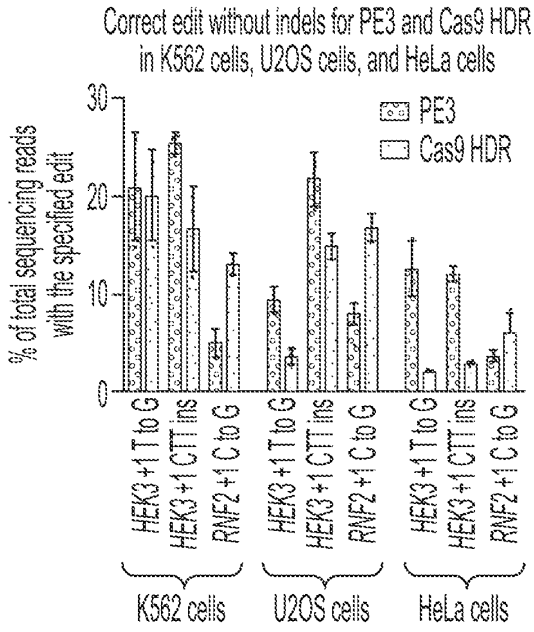


FIG. 43G

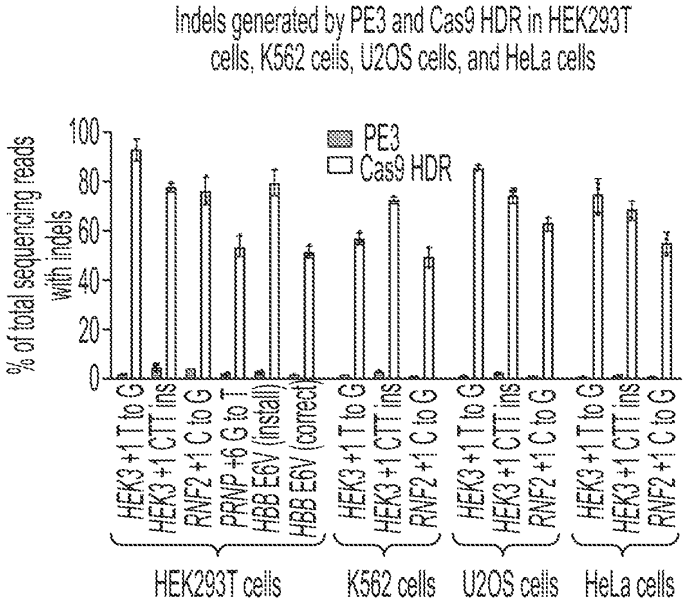


FIG. 43H

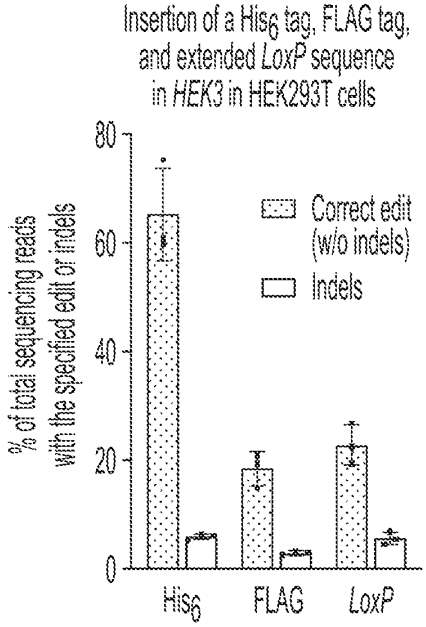


FIG. 43I

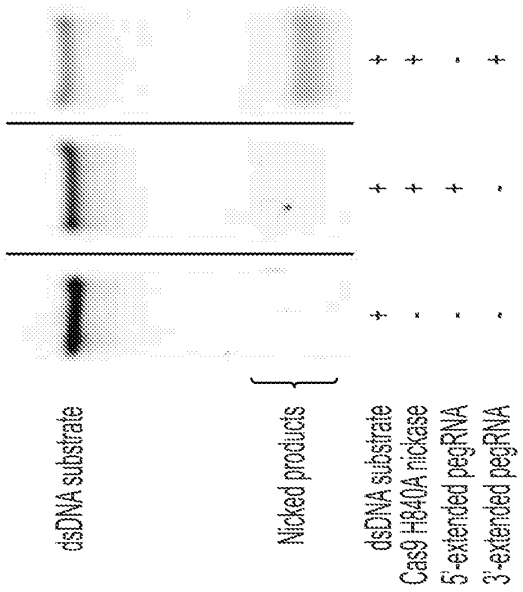


FIG. 44B

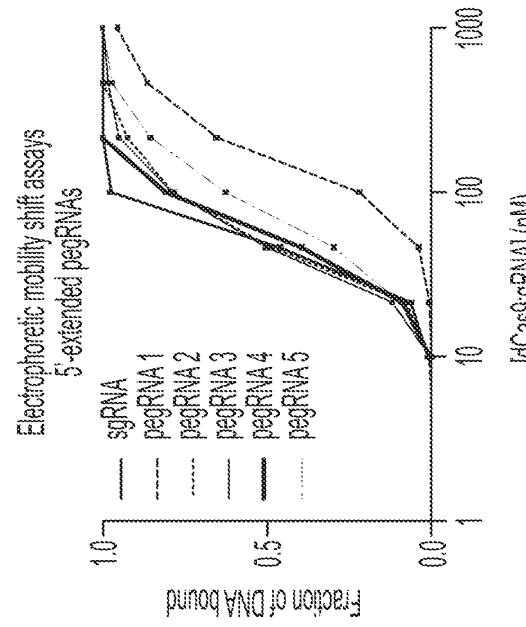


FIG. 44A

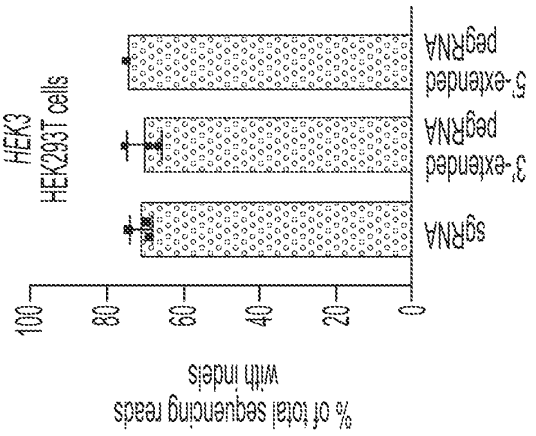

% of total sequencing reads with indels
100
80
60
40
20
0
HEK3
HEK293T cells
sgRNA
3'-extended pegRNA
5'-extended pegRNA

FIG. 44C

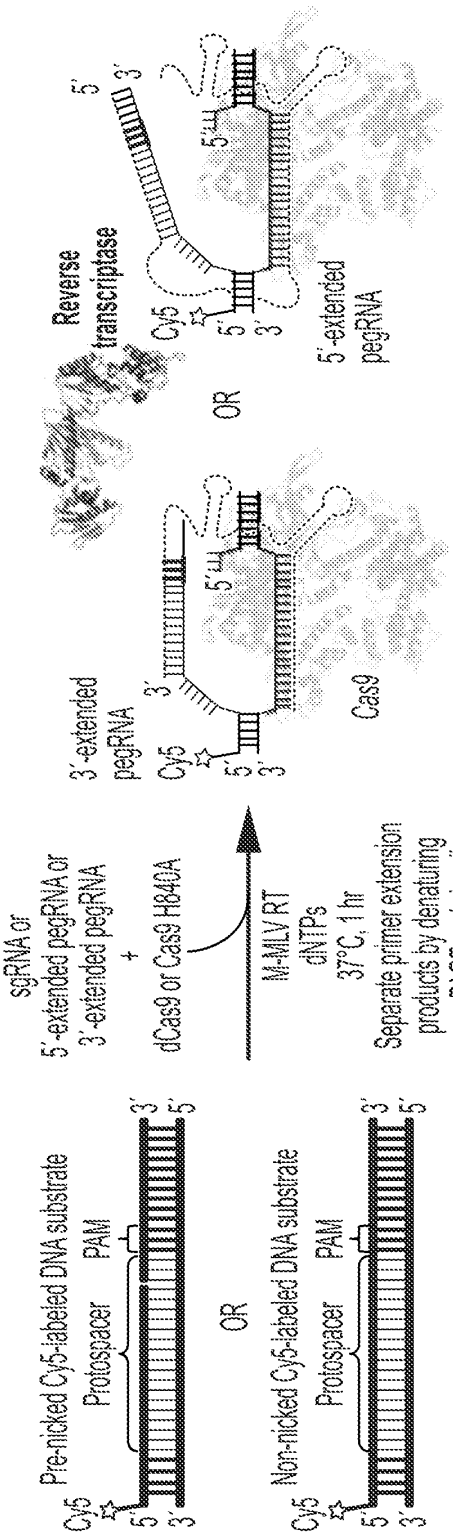

Pre-nicked Cy5-labeled DNA substrate
Cy5
Protospacer
PAM
5'
3'
OR
Non-nicked Cy5-labeled DNA substrate
sgRNA or
5'-extended pegRNA or
3'-extended pegRNA
+
dCas9 or Cas9 H840A
M-MLV RT
dNTPs
37°C, 1 hr
Separate primer extension products by denaturing urea PAGE and visualize Cy5 fluorescence
3'-extended pegRNA
Cas9
Reverse transcriptase
5'-extended pegRNA

FIG. 44D

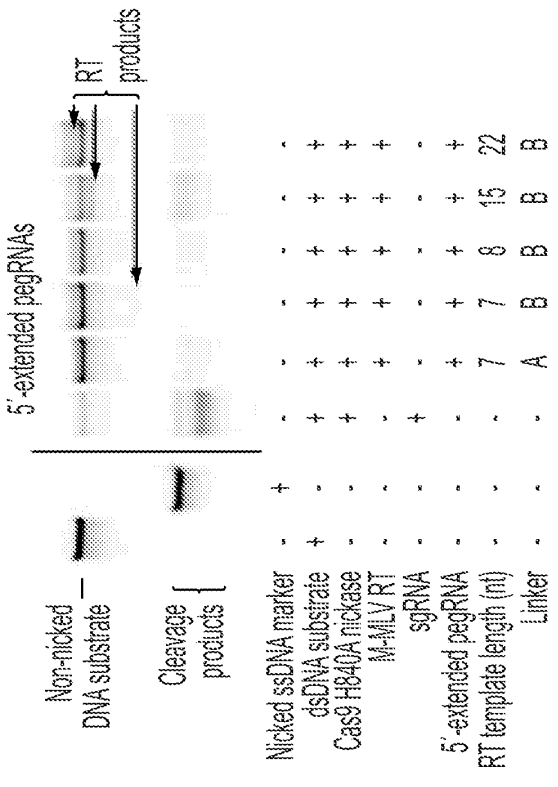


FIG. 44E

FIG. 44F

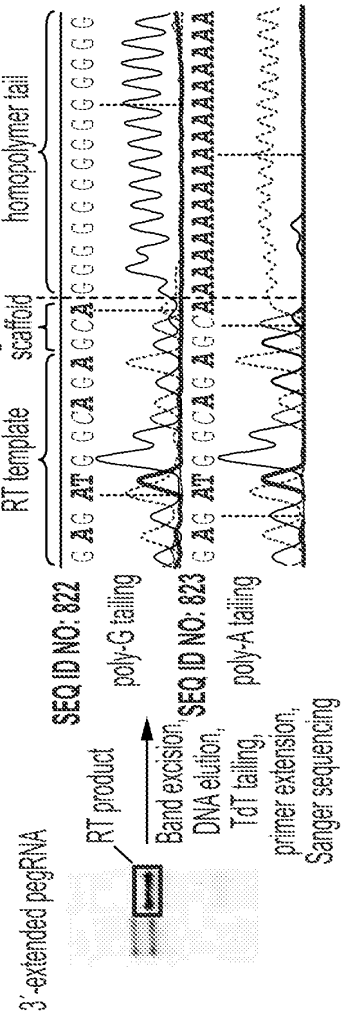


FIG. 44G

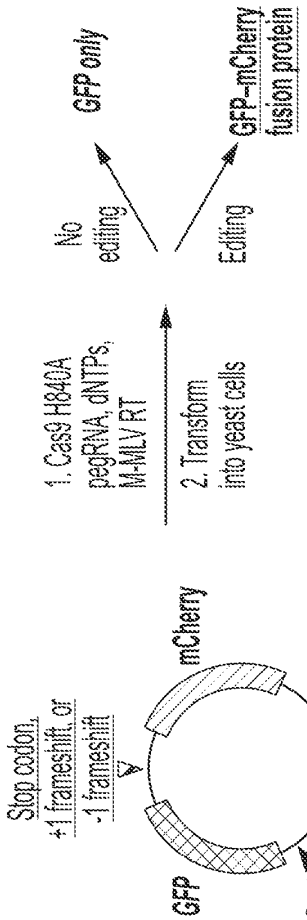


FIG. 45A

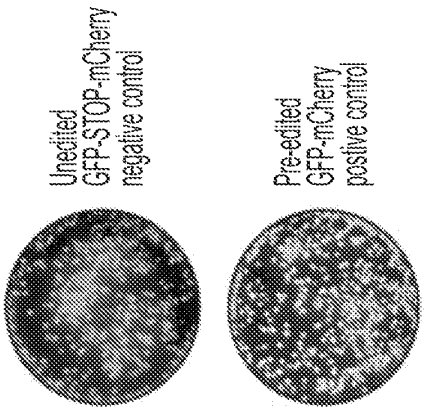

Unedited
GFP-STOP-mCherry
negative control
Pre-edited
GFP-mCherry
positive control

FIG. 45B

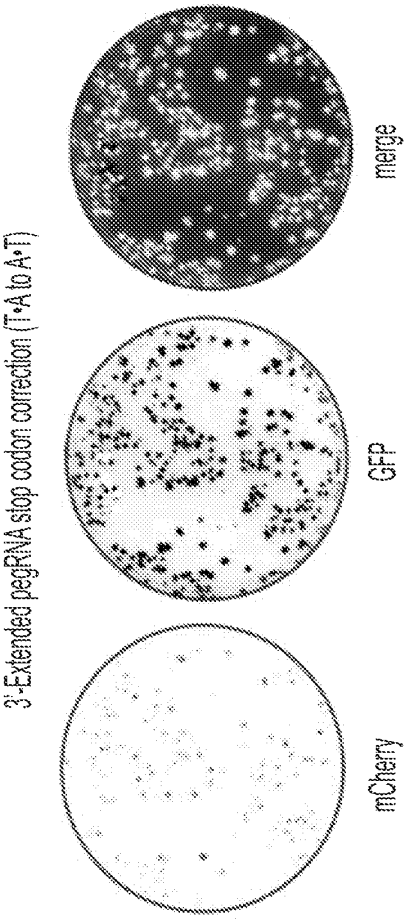


FIG. 45C

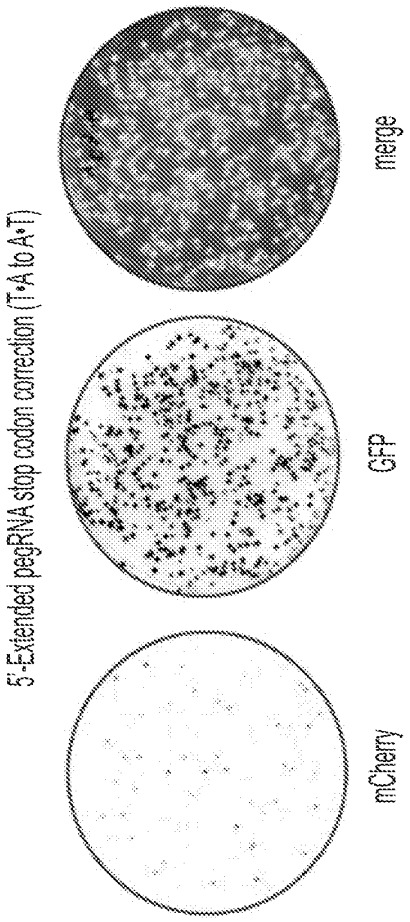


FIG. 45D

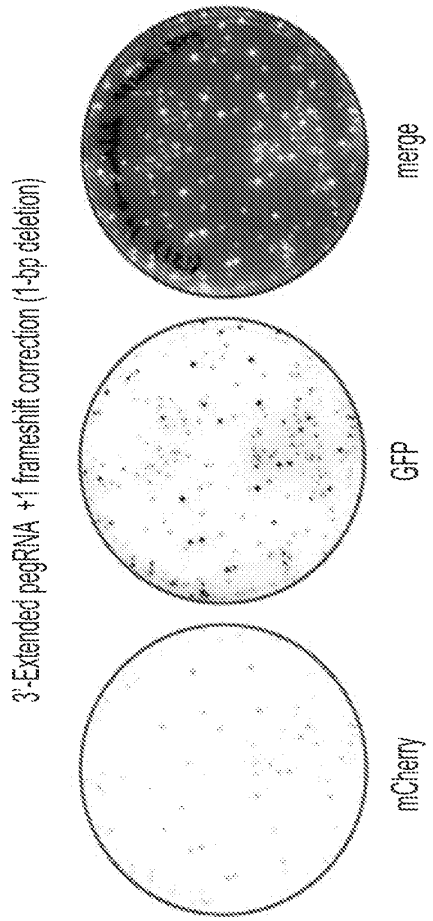


FIG. 45E

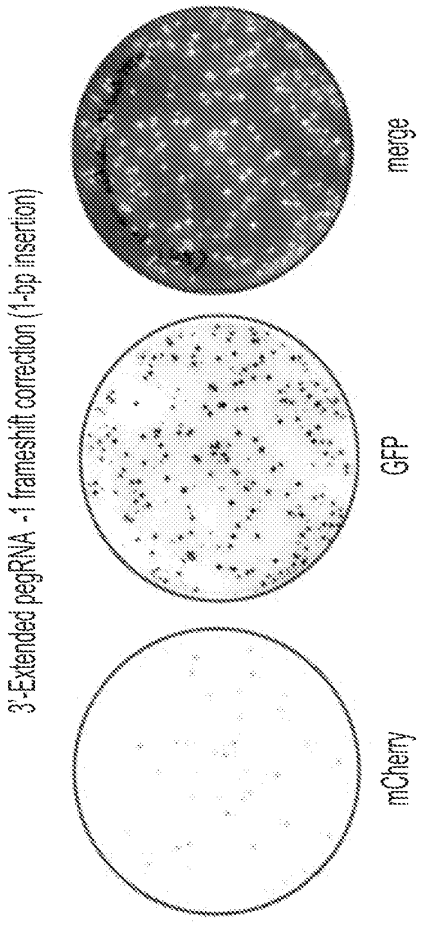


FIG. 45F

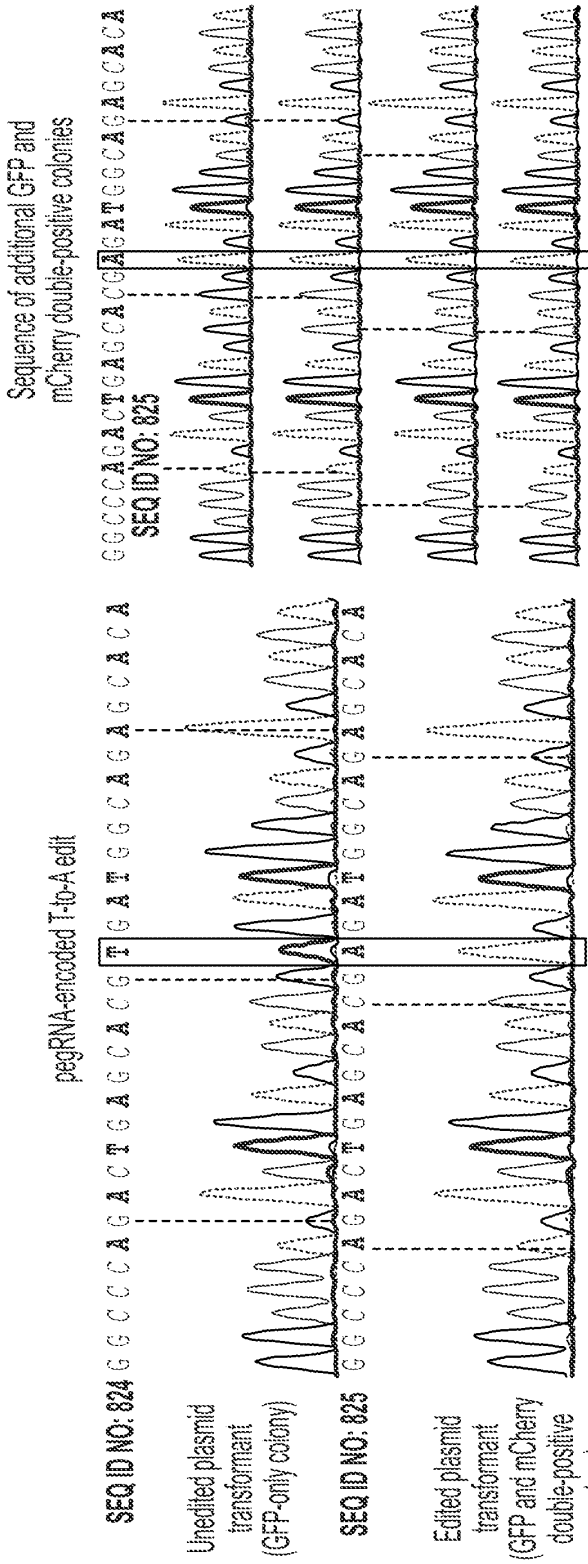

Sequence of additional GFP and mCherry double-positive colonies
SEQ ID NO: 825
GGCCCAGACTGAGCACGCATGGCAGACA
pegRNA-encoded T-to-A edit
SEQ ID NO: 824 GGCCCAGACTGAGCACGTGATGGCAGACA
Unedited plasmid transformant (GFP-only colony)
SEQ ID NO: 825 GGCCCAGACTGAGCACGATGGCAGACA
Edited plasmid transformant (GFP and mCherry double-positive colony)

FIG. 45G

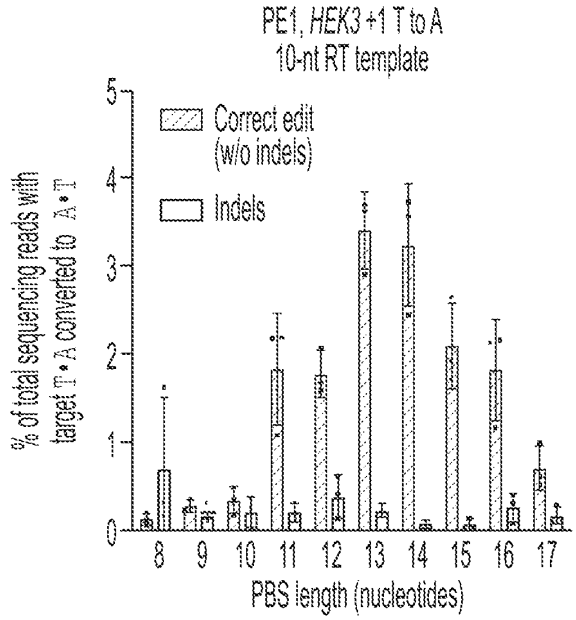

PE1, HEK3 +1 T to A
10-nt RT template
Correct edit
(w/o indels)
Indels
% of total sequencing reads with
target T•A converted to A•T
0
1
2
3
4
5
8
9
10
11
12
13
14
15
16
17
PBS length (nucleotides)

FIG. 46A

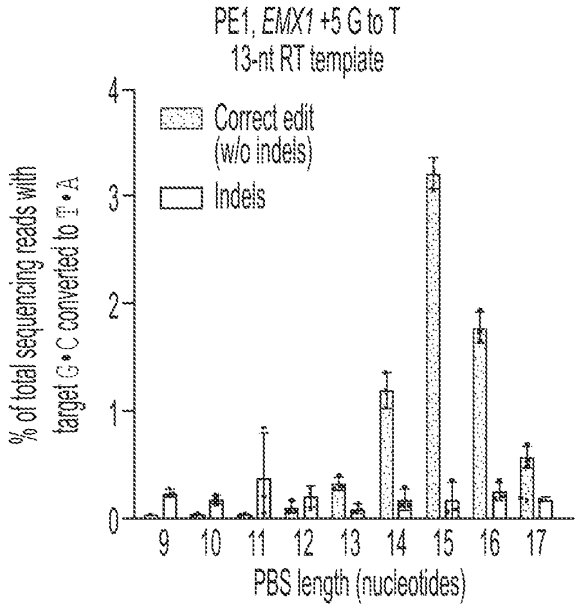


FIG. 46B

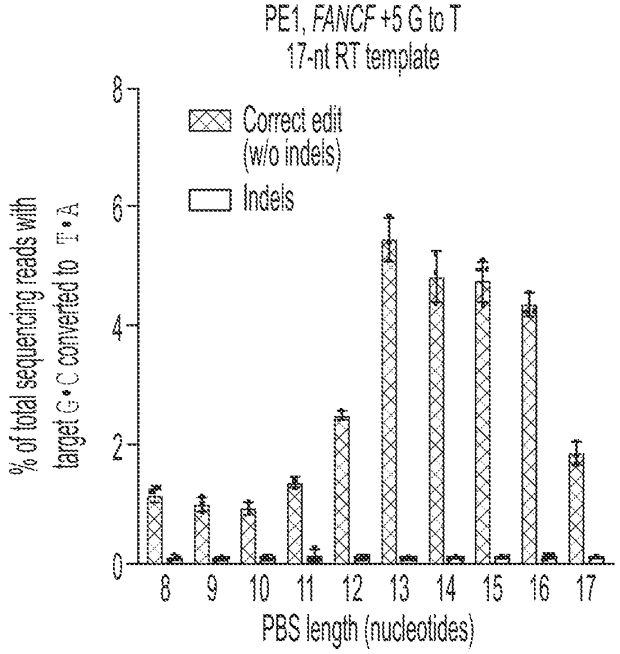


FIG. 46C

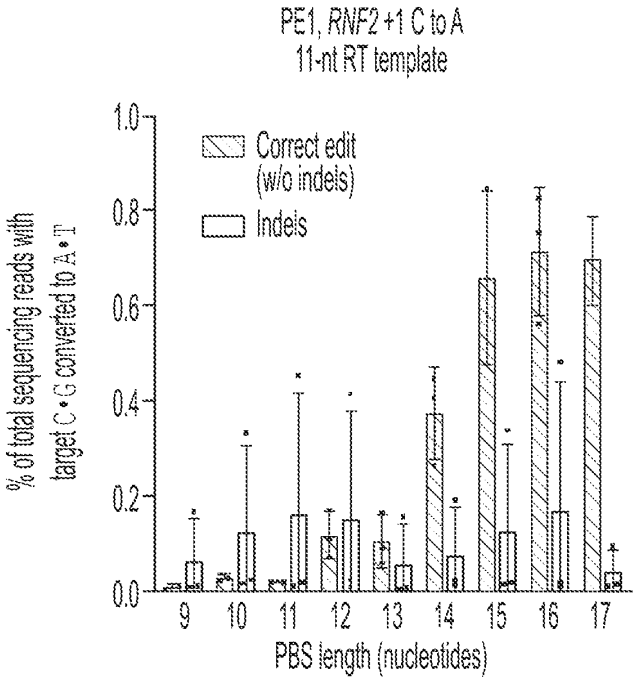


FIG. 46D

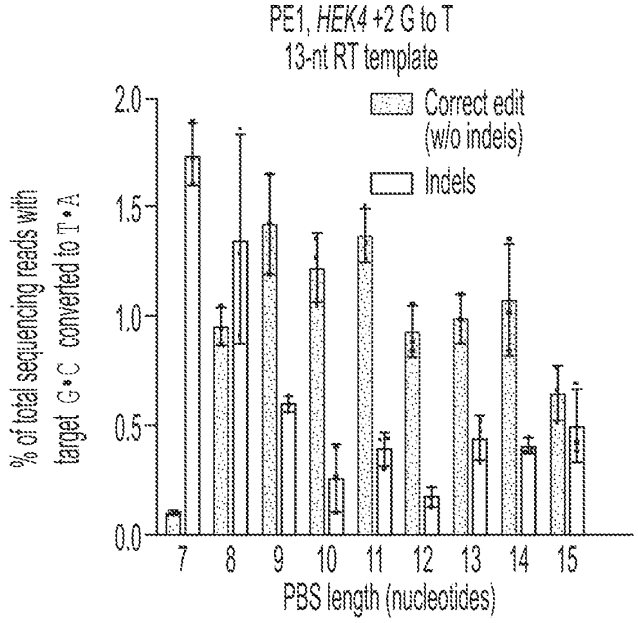


FIG. 46E

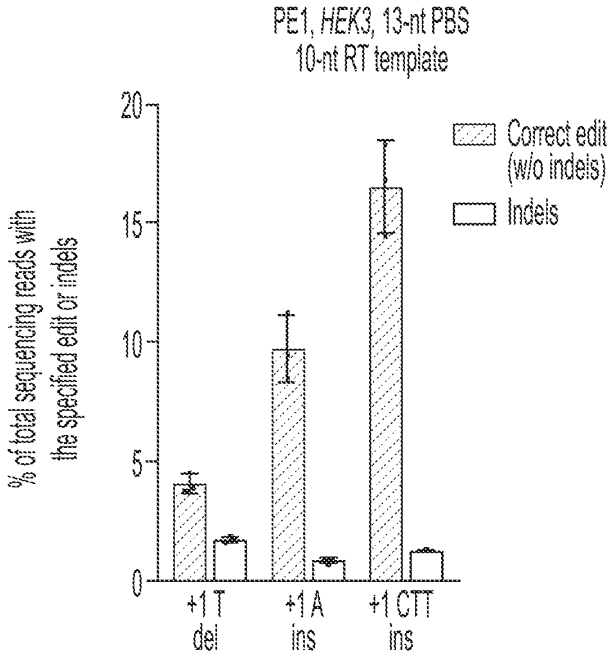


FIG. 46F

PE1	Cas9 H840A-M-MLV RT (wt)
PE1-M1	Cas9 H840A-M-MLV RT (D200N)
PE1-M2	Cas9 H840A-M-MLV RT (D200N+L603W)
PE1-M3	Cas9 H840A-M-MLV RT (D200N+L603W+T330P)
PE1-M6	Cas9 H840A-M-MLV RT (D200N+L603W+T330P+D524G+E562Q+D583N)
PE1-M15	Cas9 H840A-M-MLV RT (P51L+S67R+E67K+T197A+H204R+E302K+F309N+W1331F+T330P+L345G+ N454K+D524G+D583N+H594Q+D653N)
PE1-M3inv	M-MLV RT (D200N+L603W+T330P)-Cas9 H840A
PE2	Cas9 H840A-M-MLV RT (D200N+L603W+T330P+T306K+W313F)

FIG. 47A

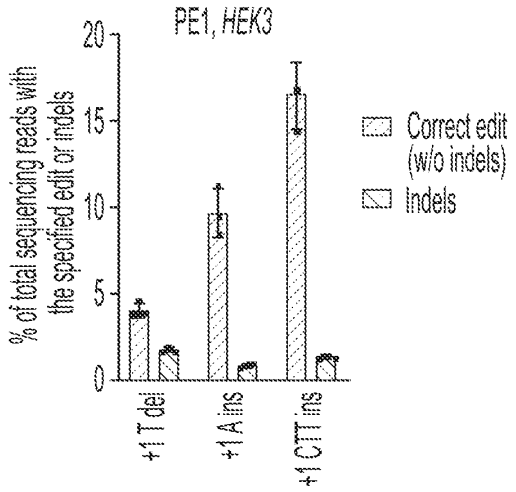


FIG. 47B

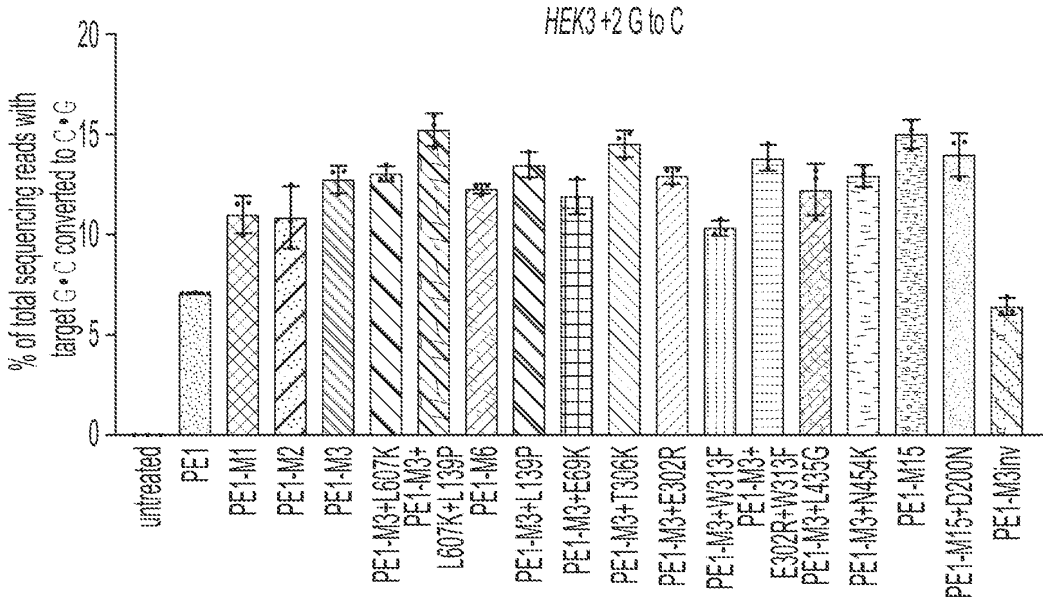


FIG. 47C

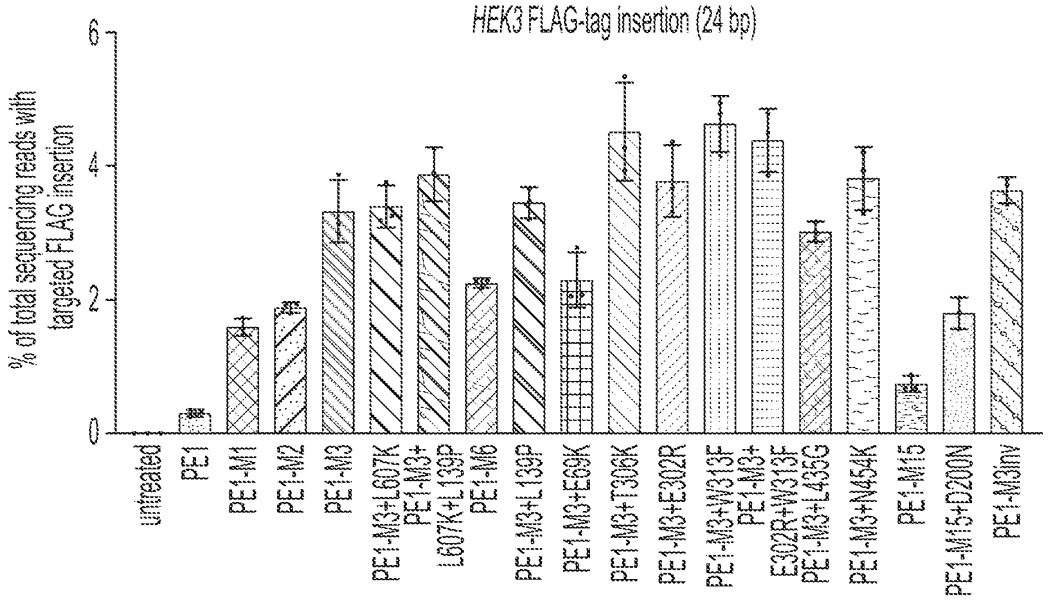


FIG. 47D

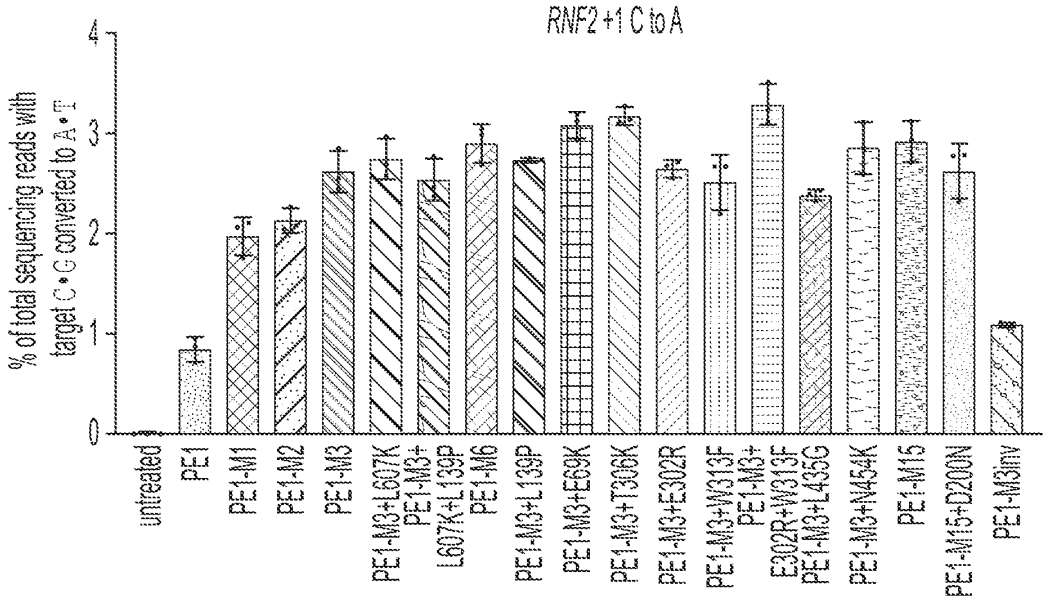


FIG. 47E

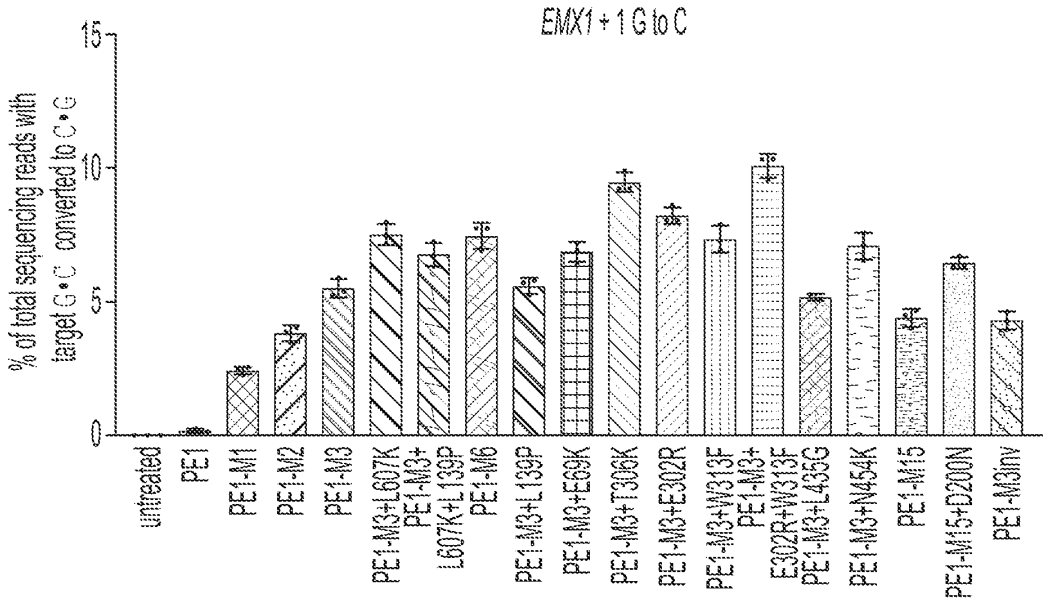


FIG. 47F

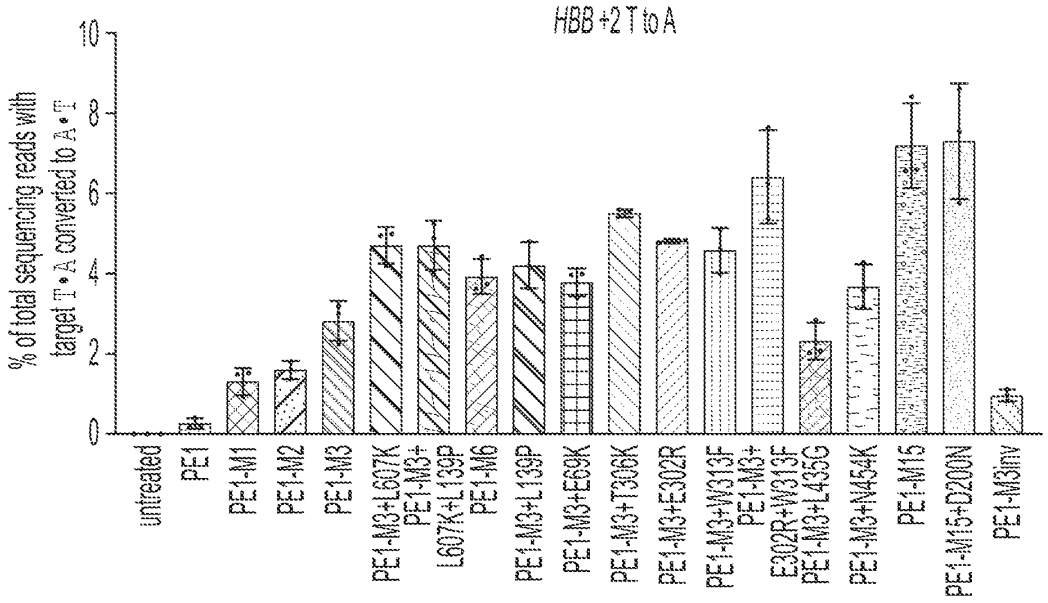


FIG. 47G

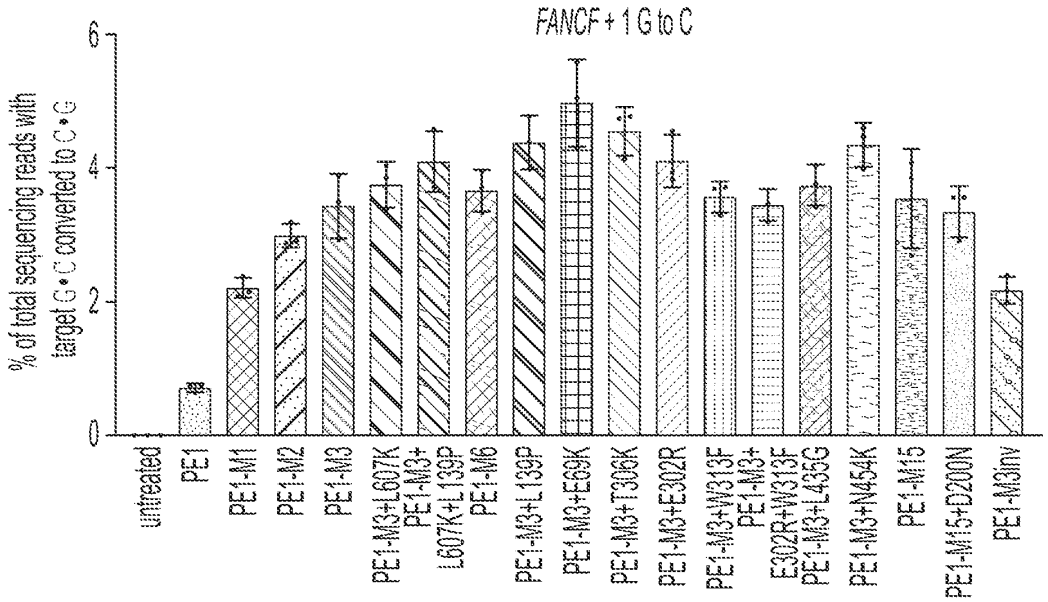


FIG. 47H

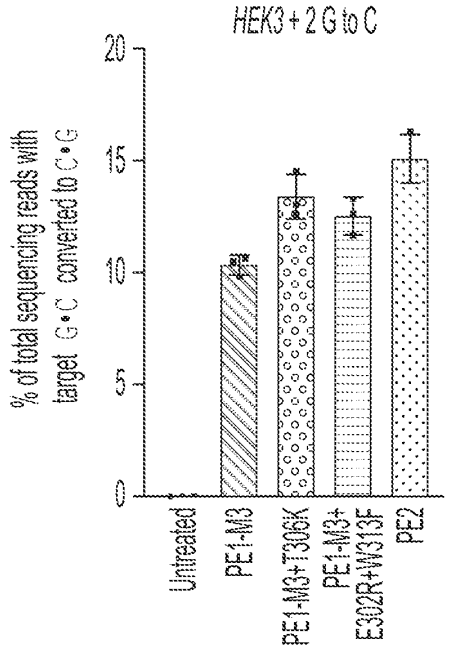


FIG. 47I

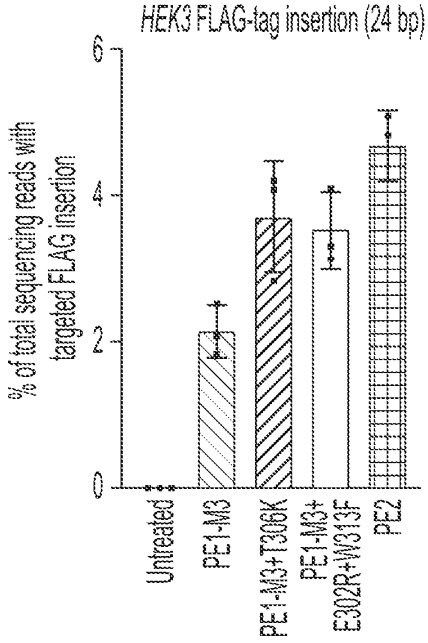


FIG. 47J

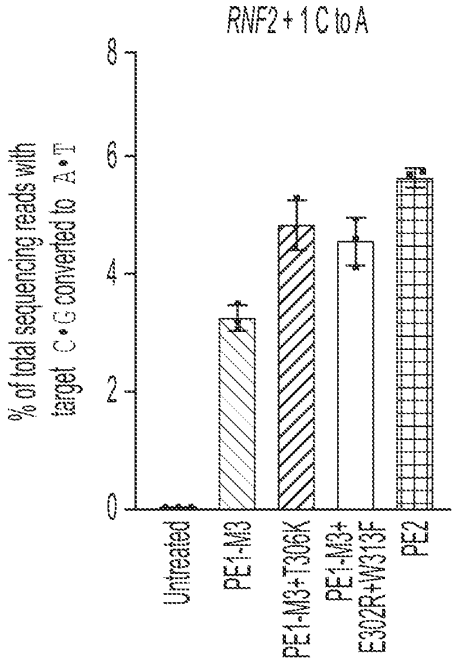


FIG. 47K

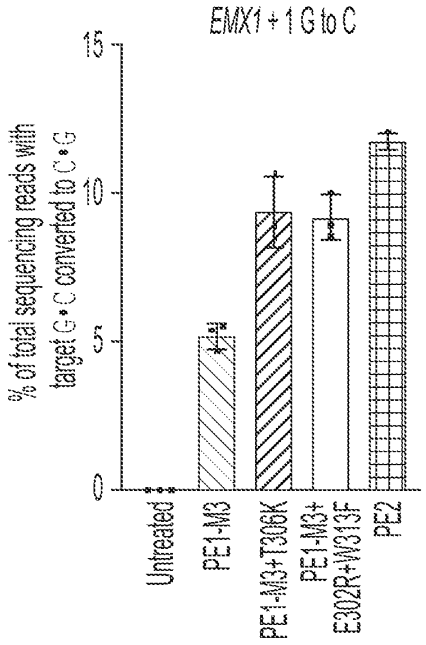


FIG. 47L

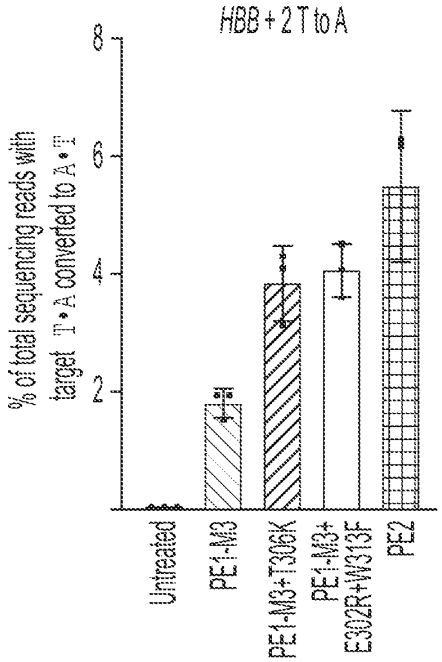


FIG. 47M

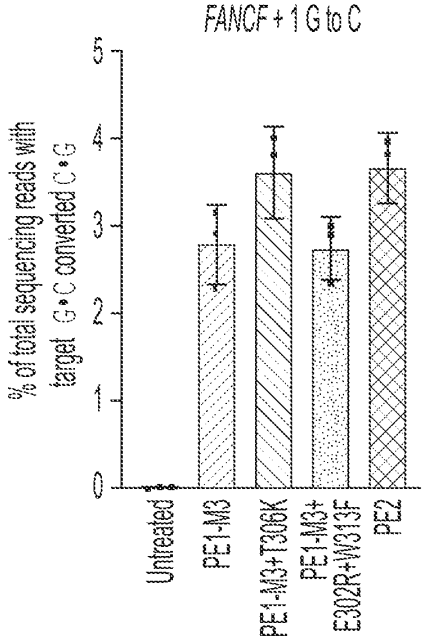


FIG. 47N

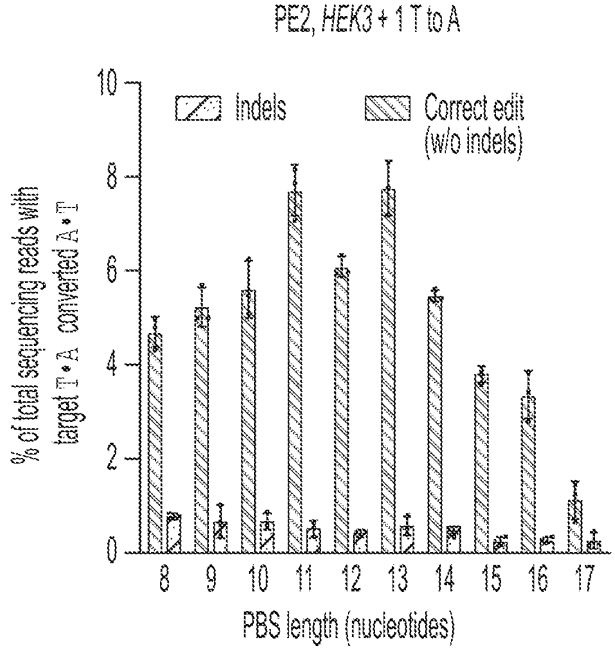


FIG. 47O

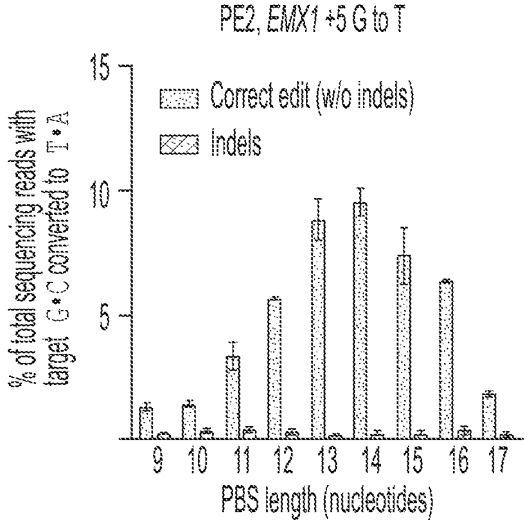


FIG. 47P

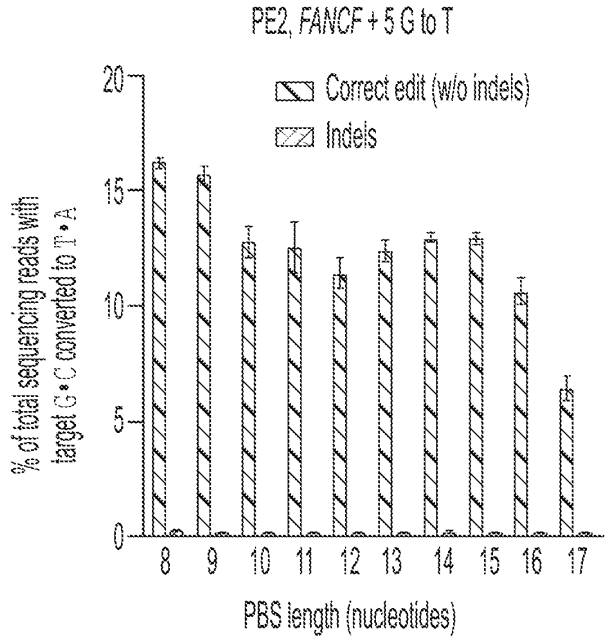


FIG. 47Q

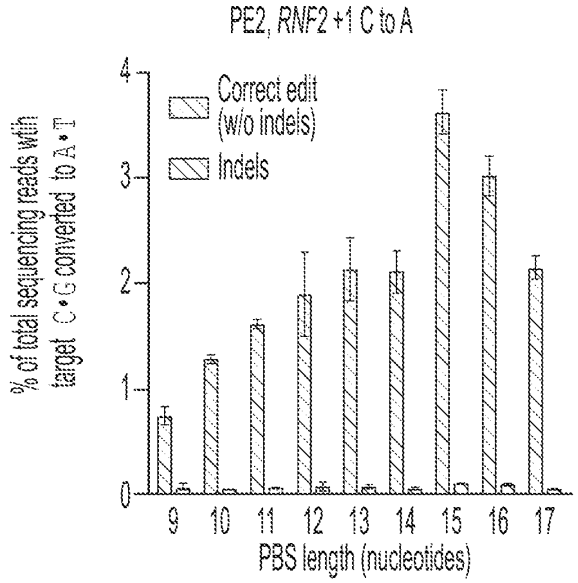


FIG. 47R

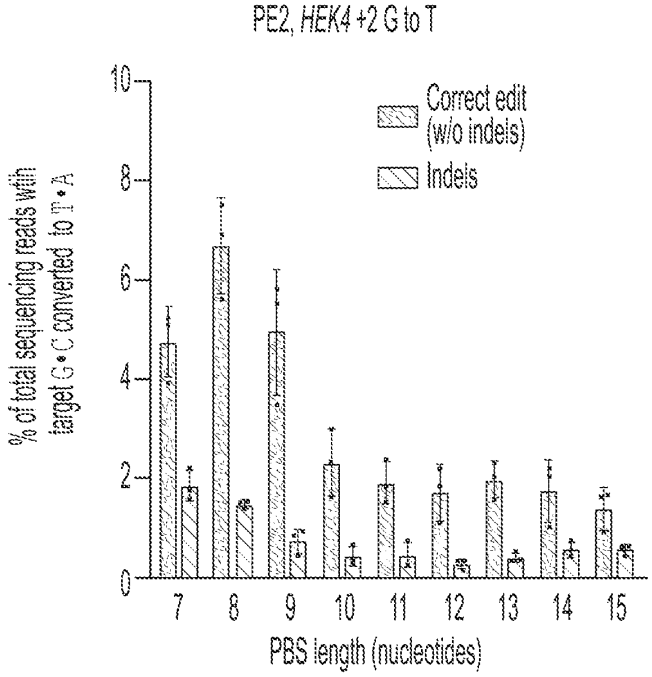


FIG. 47S

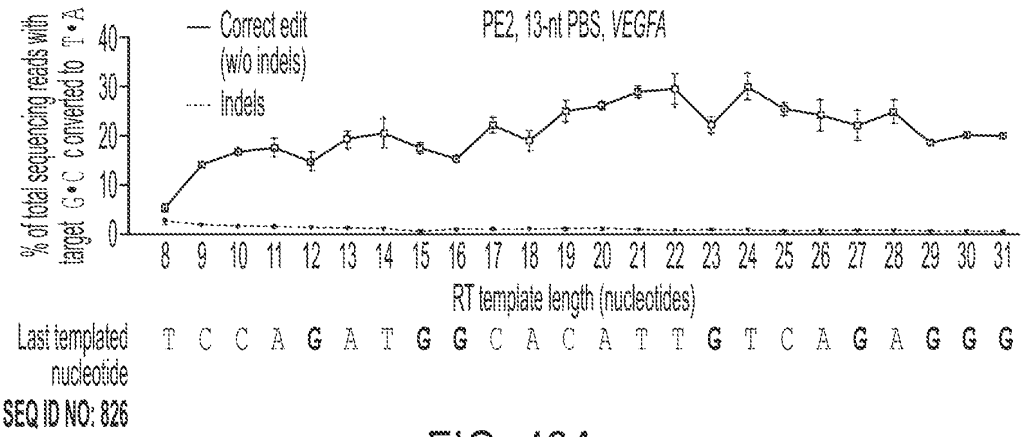


FIG. 48A

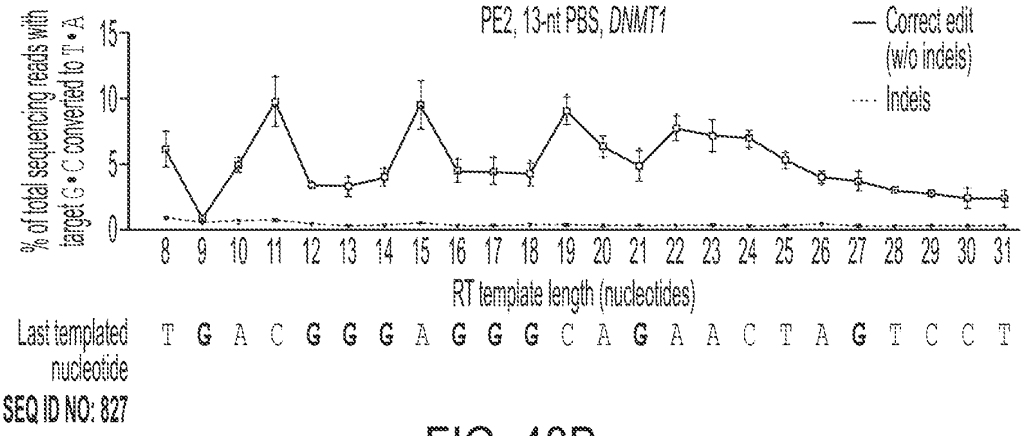


FIG. 48B

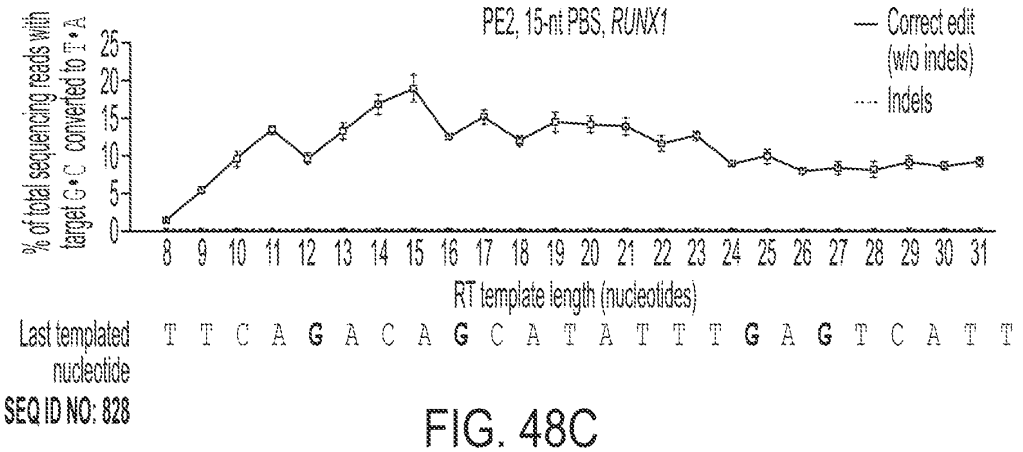


FIG. 48C

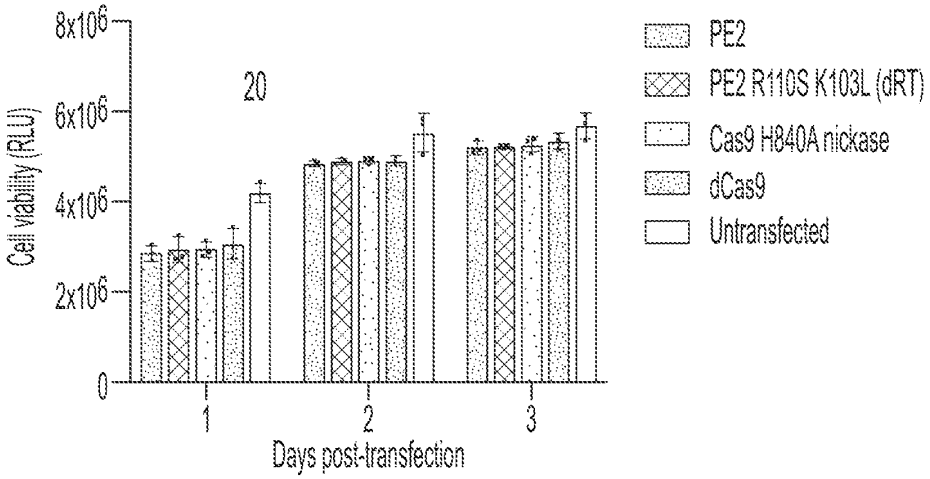


FIG. 49A

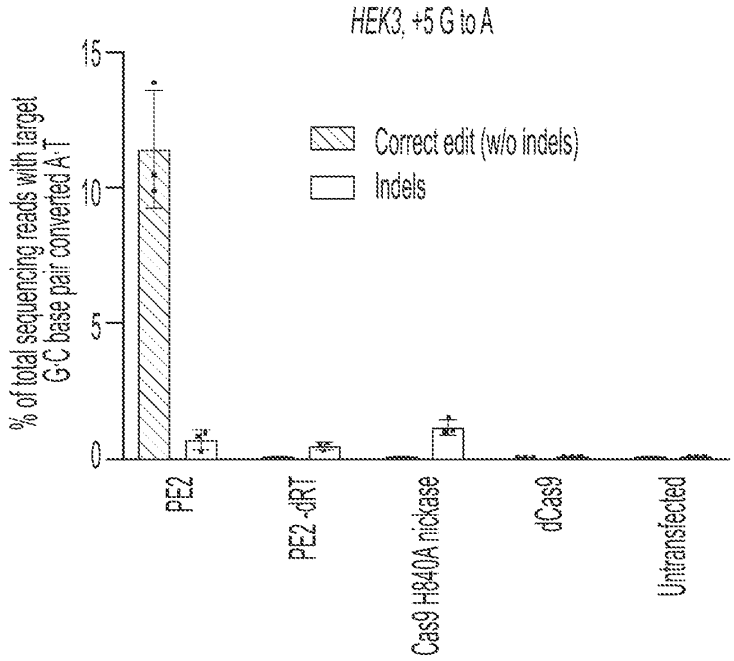


FIG. 49B

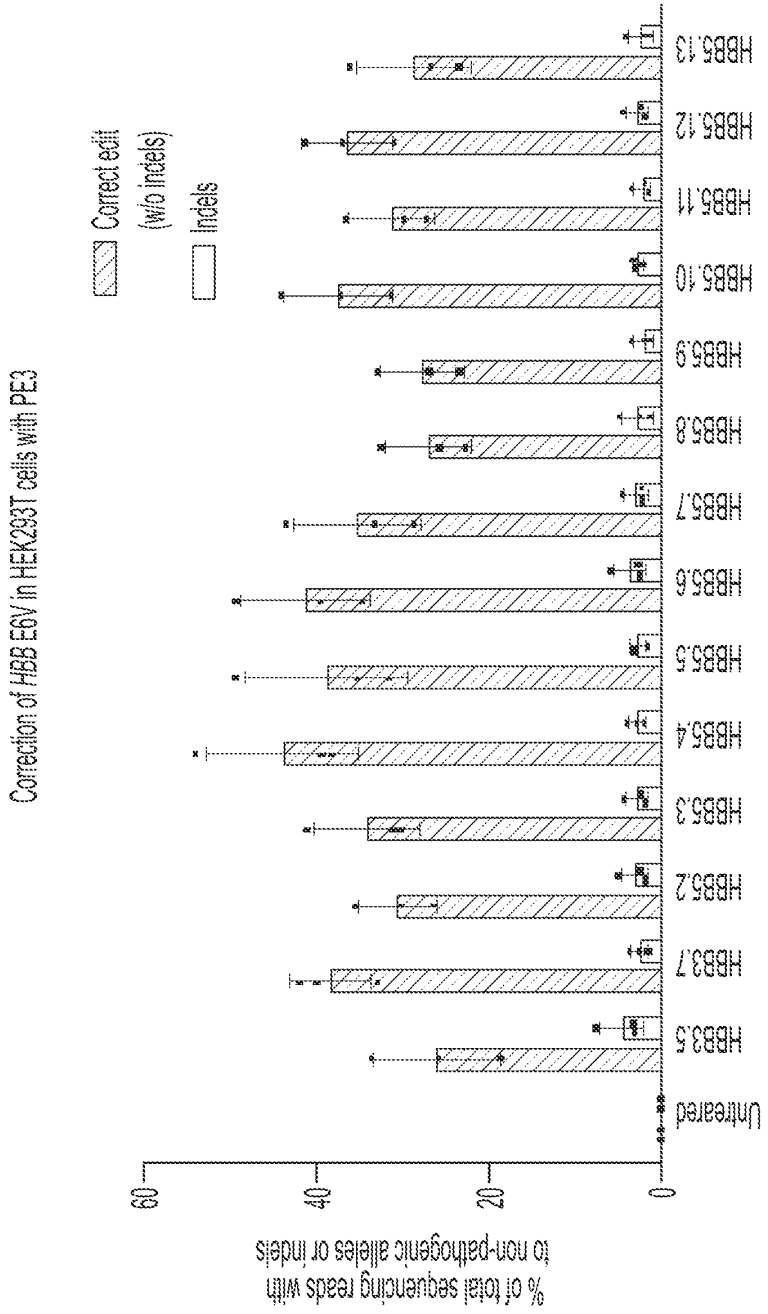

Correction of HBB E6V in HEK293T cells with PE3
% of total sequencing reads with to non-pathogenic alleles or indels
0
20
40
60
Correct edit (w/o indels)
Indels
Unlreated
HBB3.5
HBB3.7
HBB5.2
HBB5.3
HBB5.4
HBB5.5
HBB5.6
HBB5.7
HBB5.8
HBB5.9
HBB5.10
HBB5.11
HBB5.12
HBB5.13
pegRNA

FIG. 50A

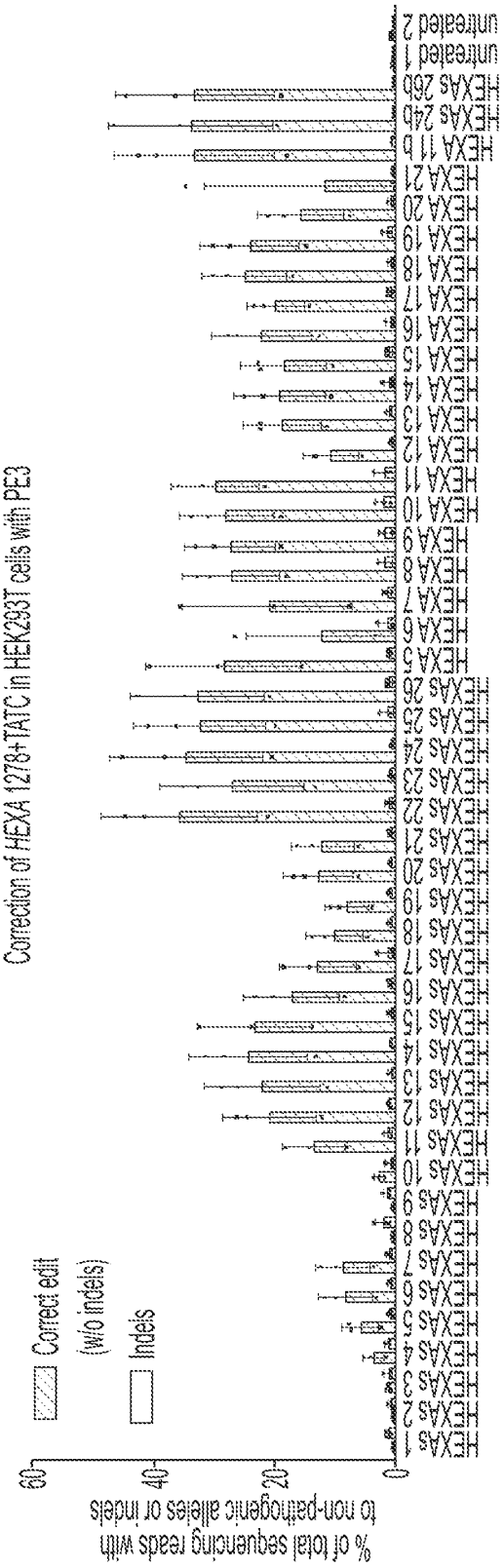

Correction of HEXA 1278+TATC in HEK293T cells with PE3
% of total sequencing reads with to non-pathogenic alleles or indels
0
20
40
60
Indels
Correct edit (w/o indels)
pegRNA
HEXAs 1
HEXAs 2
HEXAs 3
HEXAs 4
HEXAs 5
HEXAs 6
HEXAs 7
HEXAs 8
HEXAs 9
HEXAs 10
HEXAs 11
HEXAs 12
HEXAs 13
HEXAs 14
HEXAs 15
HEXAs 16
HEXAs 17
HEXAs 18
HEXAs 19
HEXAs 20
HEXAs 21
HEXAs 22
HEXAs 23
HEXAs 24
HEXAs 25
HEXAs 26
HEXA 1
HEXA 2
HEXA 3
HEXA 4
HEXA 5
HEXA 6
HEXA 7
HEXA 8
HEXA 9
HEXA 10
HEXA 11
HEXA 12
HEXA 13
HEXA 14
HEXA 15
HEXA 16
HEXA 17
HEXA 18
HEXA 19
HEXA 20
HEXA 21
HEXA 11b
HEXAs 24b
HEXAs 26b
untreated 1
untreated 2

FIG. 50B

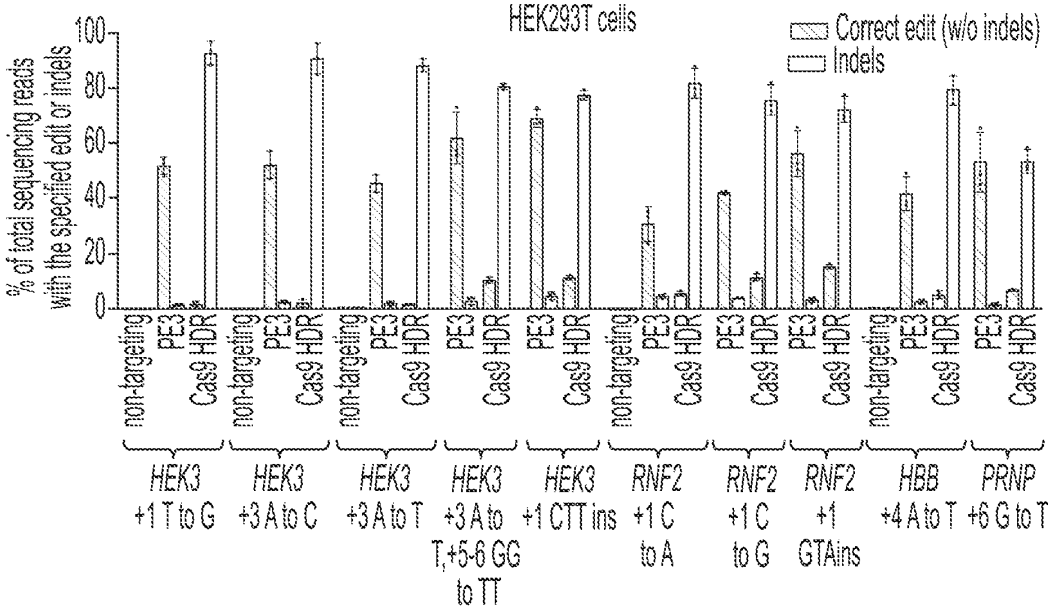


FIG. 51A

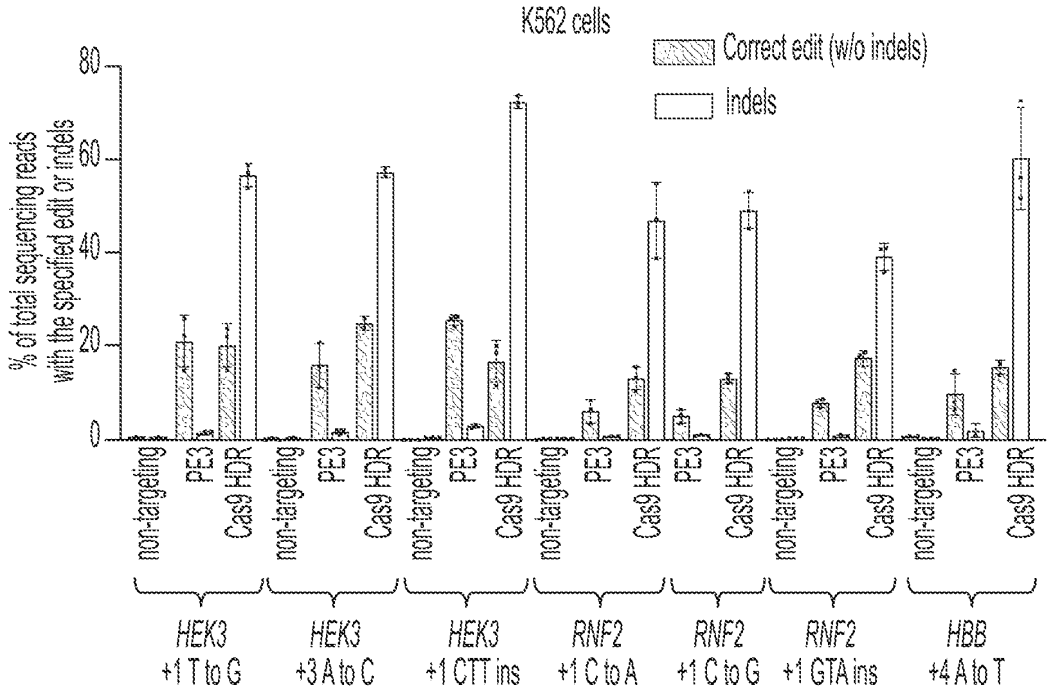


FIG. 51B

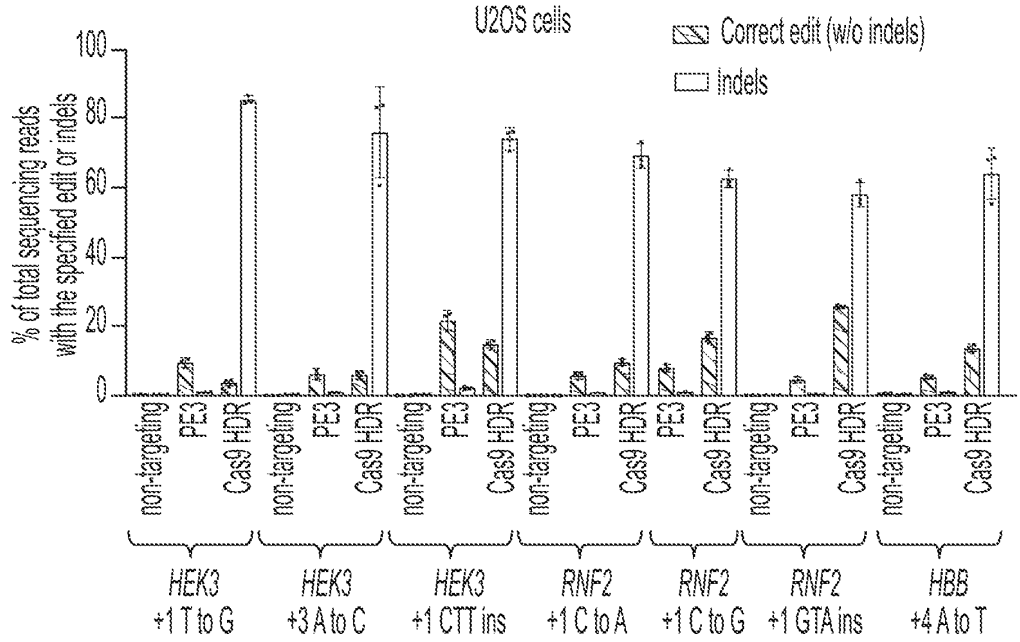


FIG. 51C

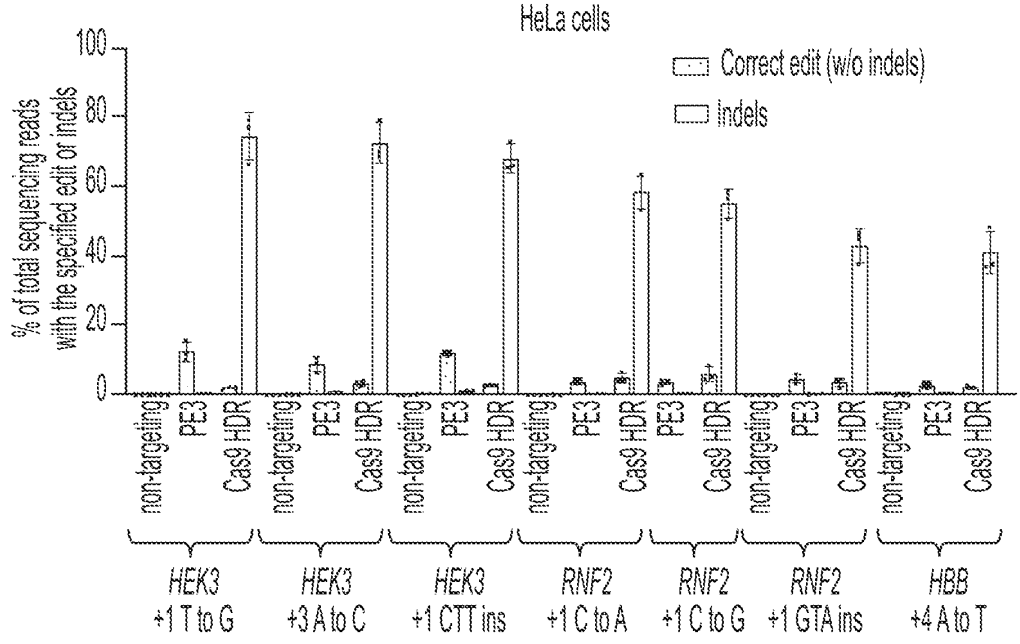


FIG. 51D

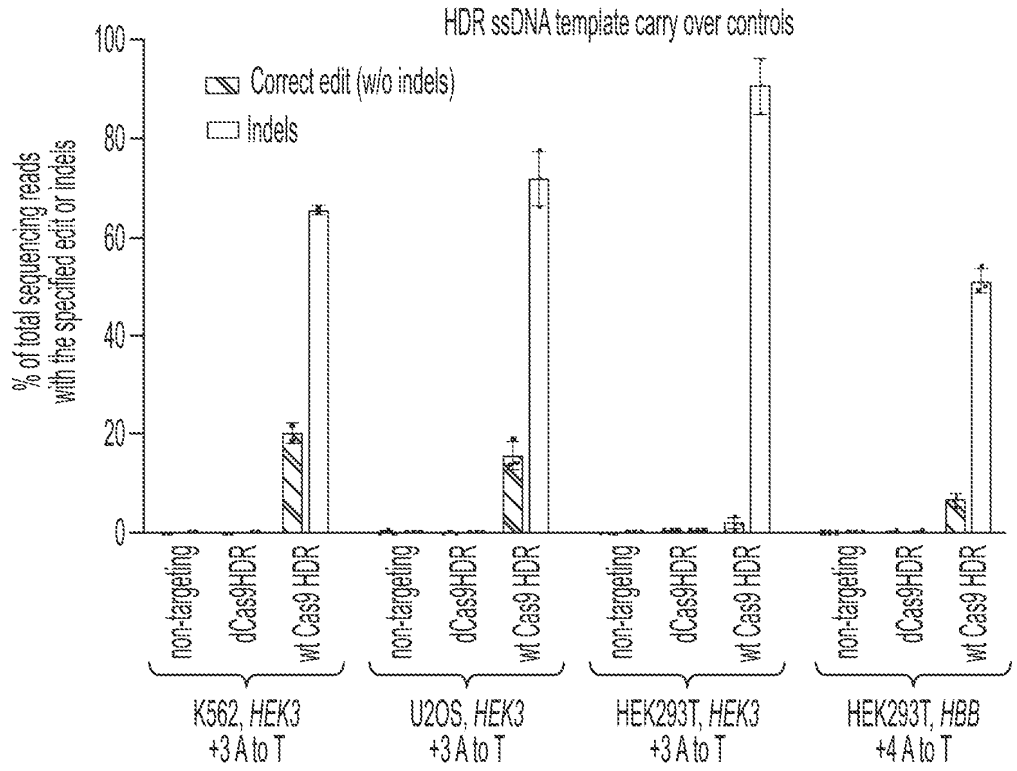

HDR ssDNA template carry over controls
Correct edit (w/o indels)
Indels
% of total sequencing reads with the specified edit or indels
100
80
60
40
20
0
non-targeting
dCas9HDR
wt Cas9 HDR
K562, HEK3 +3 A to T
U2OS, HEK3 +3 A to T
HEK293T, HEK3 +3 A to T
HEK293T, HBB +4 A to T

FIG. 51E

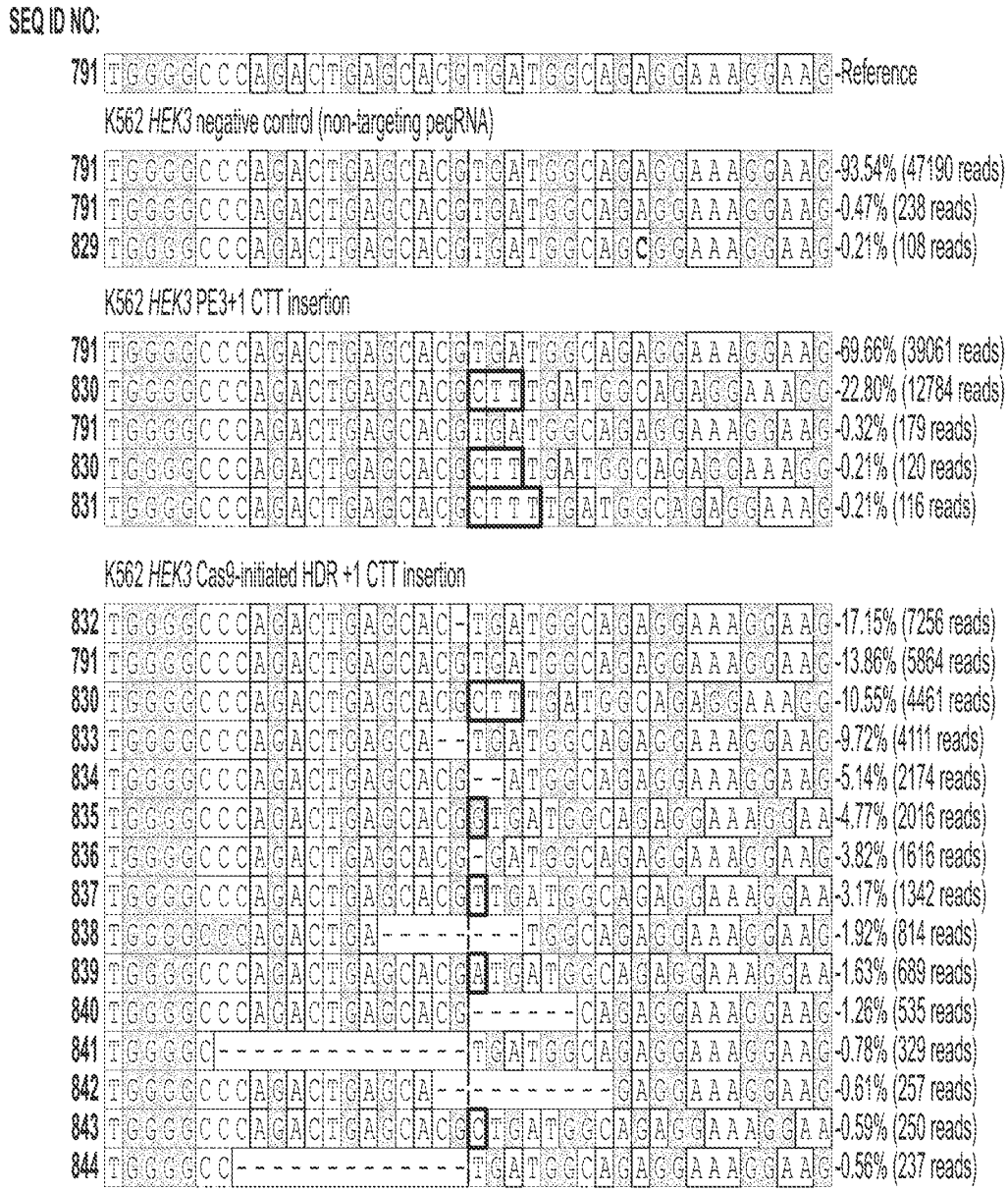

SEQ ID NO:
791 TGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAG -Reference
K562 HEK3 negative control (non-targeting pegRNA)
791 TGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAG -93.54% (47190 reads)
791 TGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAG -0.47% (238 reads)
829 TGGGGCCCAGACTGAGCACGTGATGGCAGCGGAAAGGAAG -0.21% (108 reads)
K562 HEK3 PE3+1 CTT insertion
791 TGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAG -69.66% (39061 reads)
830 TGGGGCCCAGACTGAGCACGCTTTGATGGCAGAGGAAAGG -22.80% (12784 reads)
791 TGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAG -0.32% (179 reads)
830 TGGGGCCCAGACTGAGCACGCTTTGATGGCAGAGGAAAGG -0.21% (120 reads)
831 TGGGGCCCAGACTGAGCACGCTTTTGATGGCAGAGGAAAG -0.21% (116 reads)
K562 HEK3 Cas9-initiated HDR +1 CTT insertion
832 TGGGGCCCAGACTGAGCAC-TGATGGCAGAGGAAAGGAAG -17.15% (7256 reads)
791 TGGGGCCCAGACTGAGCACGTGATGGCAGAGGAAAGGAAG -13.86% (5864 reads)
830 TGGGGCCCAGACTGAGCACGCTTTGATGGCAGAGGAAAGG -10.55% (4461 reads)
833 TGGGGCCCAGACTGAGCA--TGATGGCAGAGGAAAGGAAG -9.72% (4111 reads)
834 TGGGGCCCAGACTGAGCAC--ATGGCAGAGGAAAGGAAG -5.14% (2174 reads)
835 TGGGGCCCAGACTGAGCACGTTGATGGCAGAGGAAAGGAA -4.77% (2016 reads)
836 TGGGGCCCAGACTGAGCAC-GATGGCAGAGGAAAGGAAG -3.82% (1616 reads)
837 TGGGGCCCAGACTGAGCACGTTGATGGCAGAGGAAAGGAA -3.17% (1342 reads)
838 TGGGGCCCAGACTGA-----TGGCAGAGGAAAGGAAG -1.92% (814 reads)
839 TGGGGCCCAGACTGAGCACGATGATGGCAGAGGAAAGGAA -1.63% (689 reads)
840 TGGGGCCCAGACTGAGCAC------CAGAGGAAAGGAAG -1.26% (535 reads)
841 TGGGGC----------TGATGGCAGAGGAAAGGAAG -0.78% (329 reads)
842 TGGGGCCCAGACTGAGCA--------GAGGAAAGGAAG -0.61% (257 reads)
843 TGGGGCCCAGACTGAGCACGTTGATGGCAGAGGAAAGGAA -0.59% (250 reads)
844 TGGGGCC----------TGATGGCAGAGGAAAGGAAG -0.56% (237 reads)

FIG. 51F

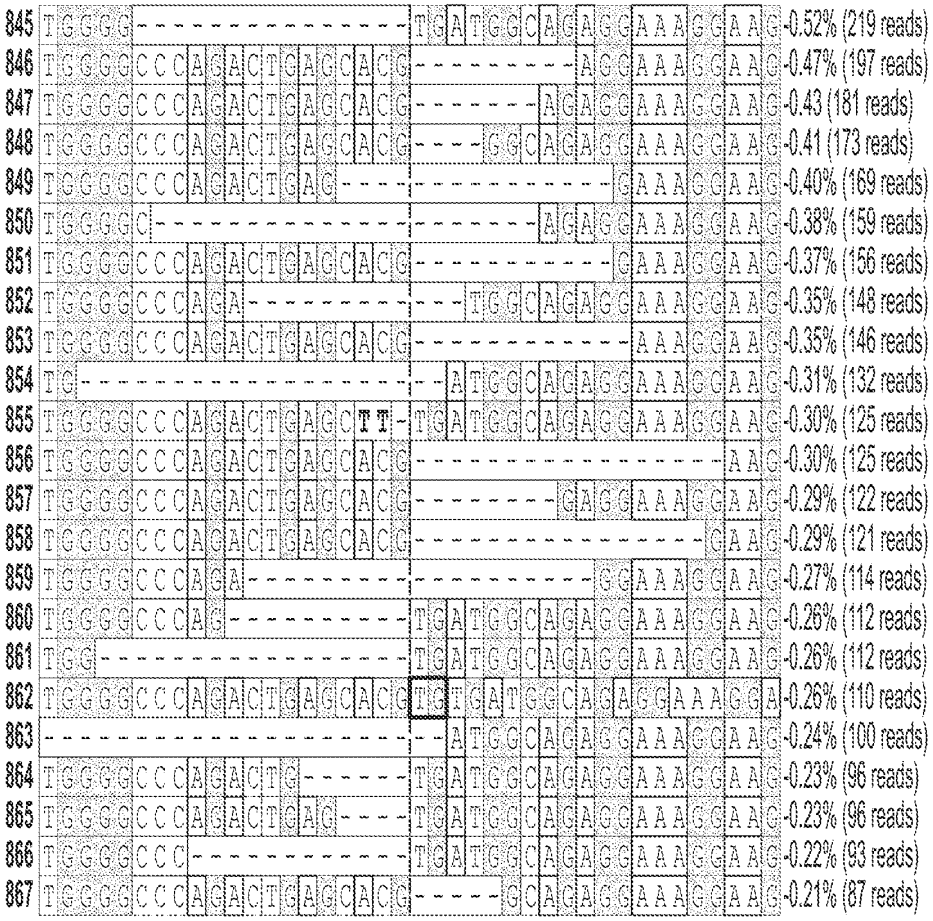

845 -0.52% (219 reads)
846 -0.47% (197 reads)
847 -0.43 (181 reads)
848 -0.41 (173 reads)
849 -0.40% (169 reads)
850 -0.38% (159 reads)
851 -0.37% (156 reads)
852 -0.35% (148 reads)
853 -0.35% (146 reads)
854 -0.31% (132 reads)
855 -0.30% (125 reads)
856 -0.30% (125 reads)
857 -0.29% (122 reads)
858 -0.29% (121 reads)
859 -0.27% (114 reads)
860 -0.26% (112 reads)
861 -0.26% (112 reads)
862 -0.26% (110 reads)
863 -0.24% (100 reads)
864 -0.23% (96 reads)
865 -0.23% (96 reads)
866 -0.22% (93 reads)
867 -0.21% (87 reads)
bold Substitutions
Insertions
- Deletions
---- Predicted
cleavage position

FIG. 51G

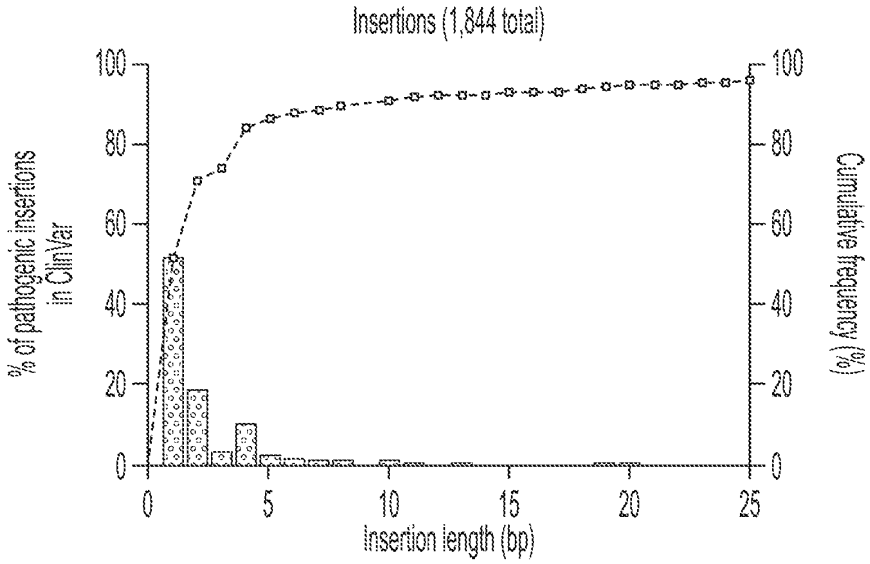


FIG. 52A

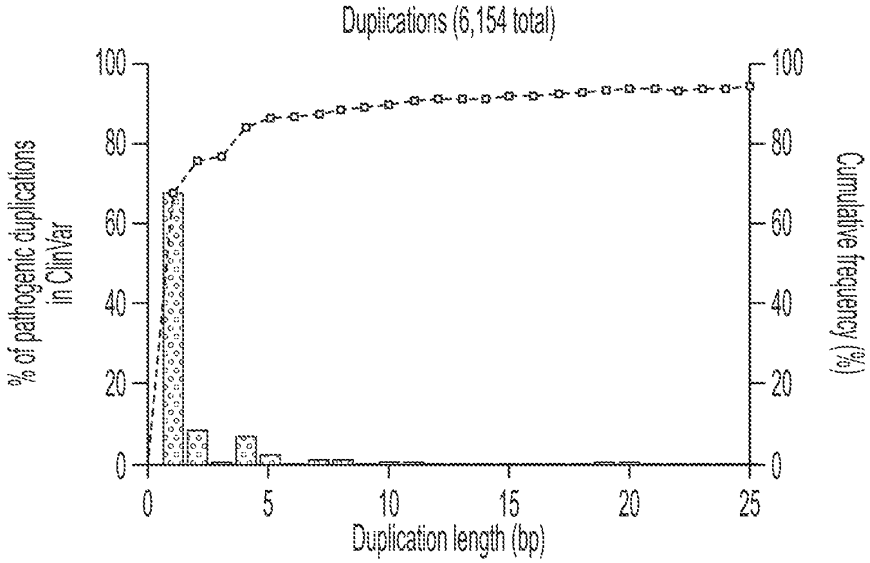


FIG. 52B

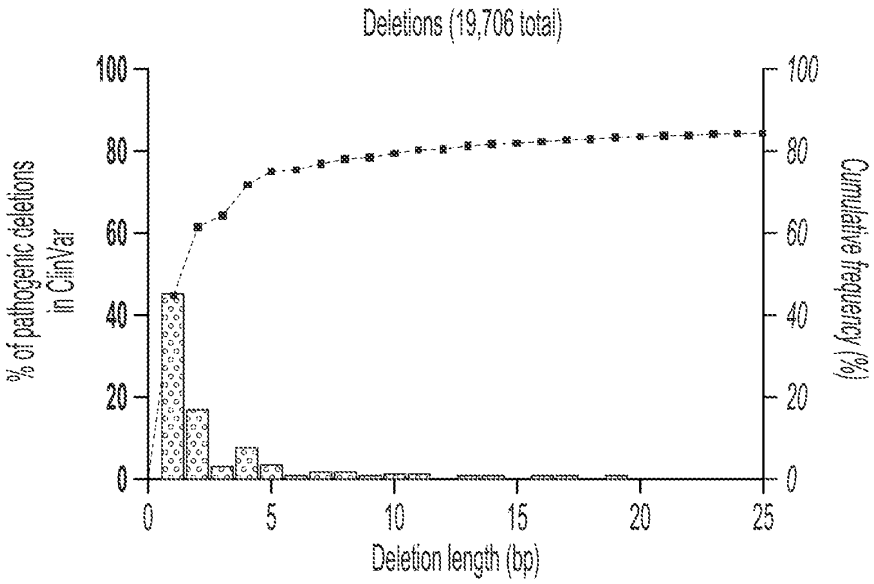


FIG. 52C

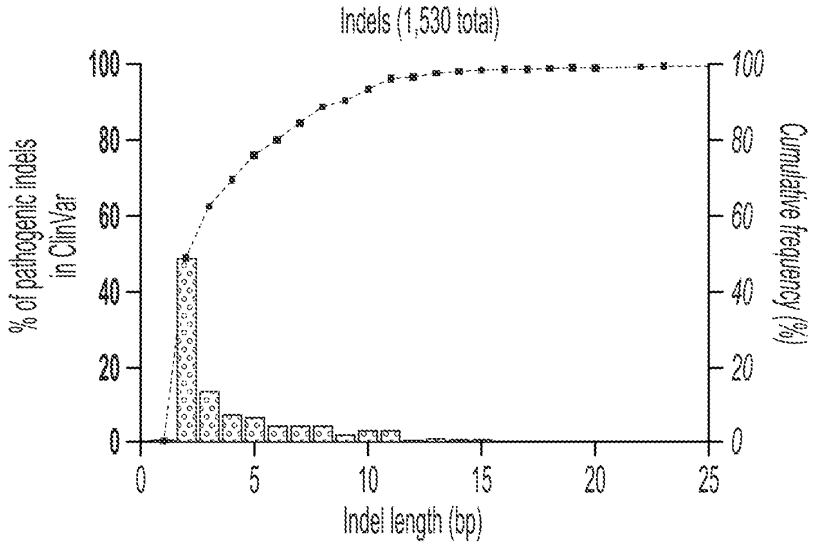


FIG. 52D

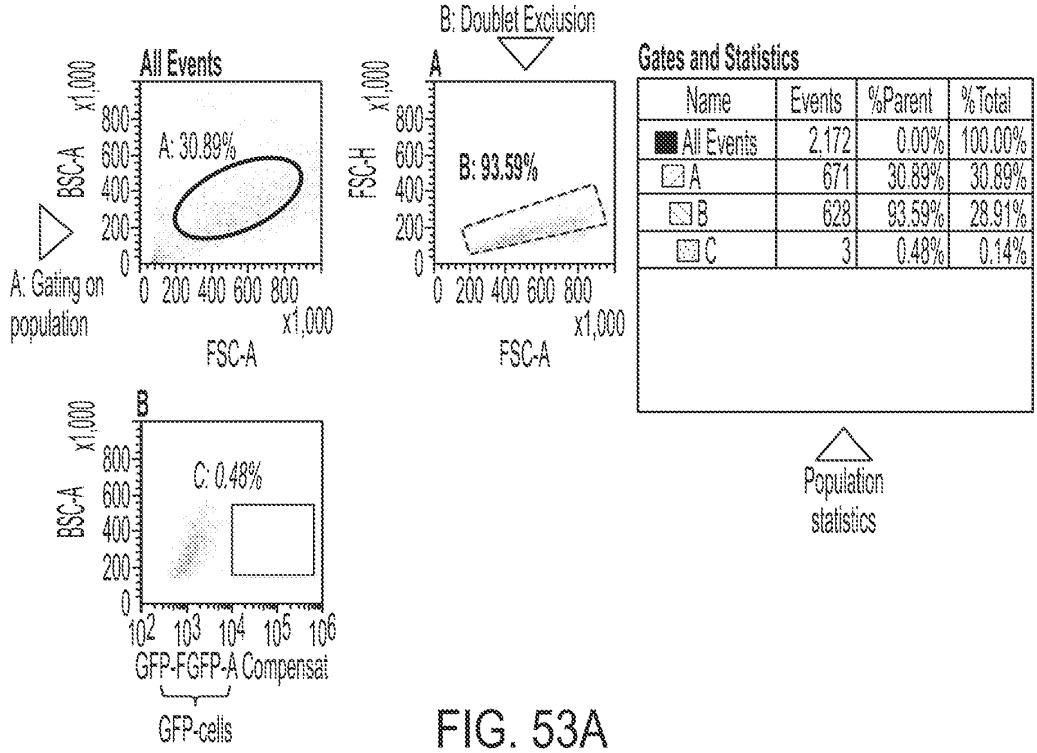


FIG. 53A

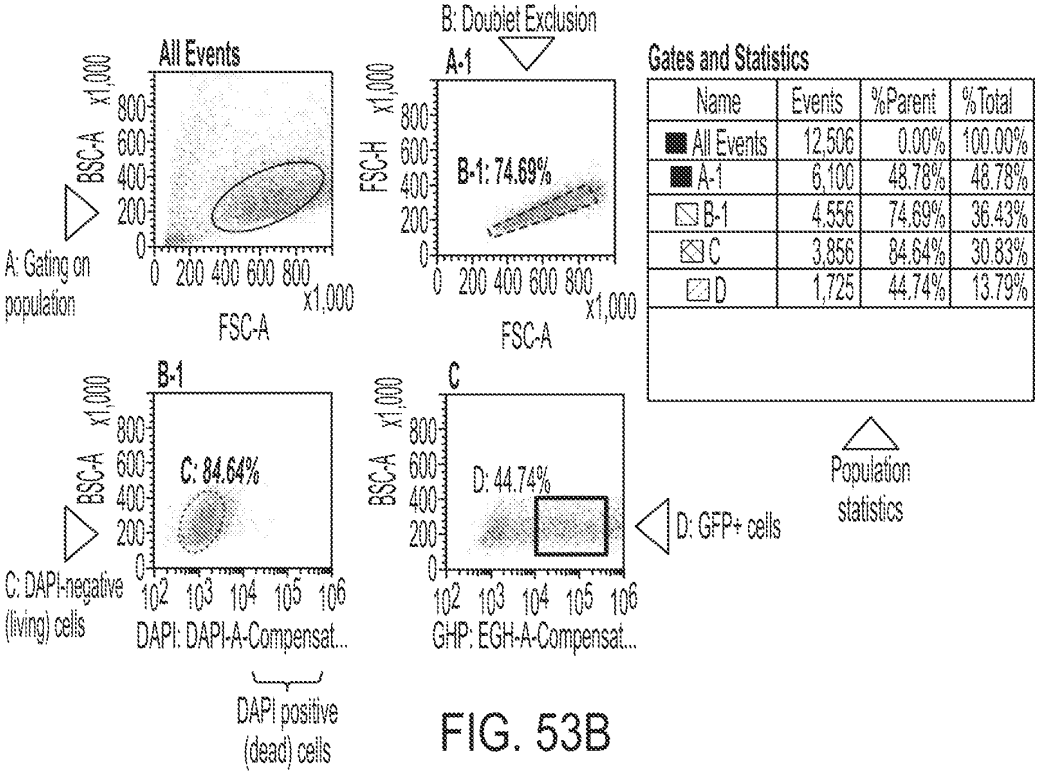


FIG. 53B

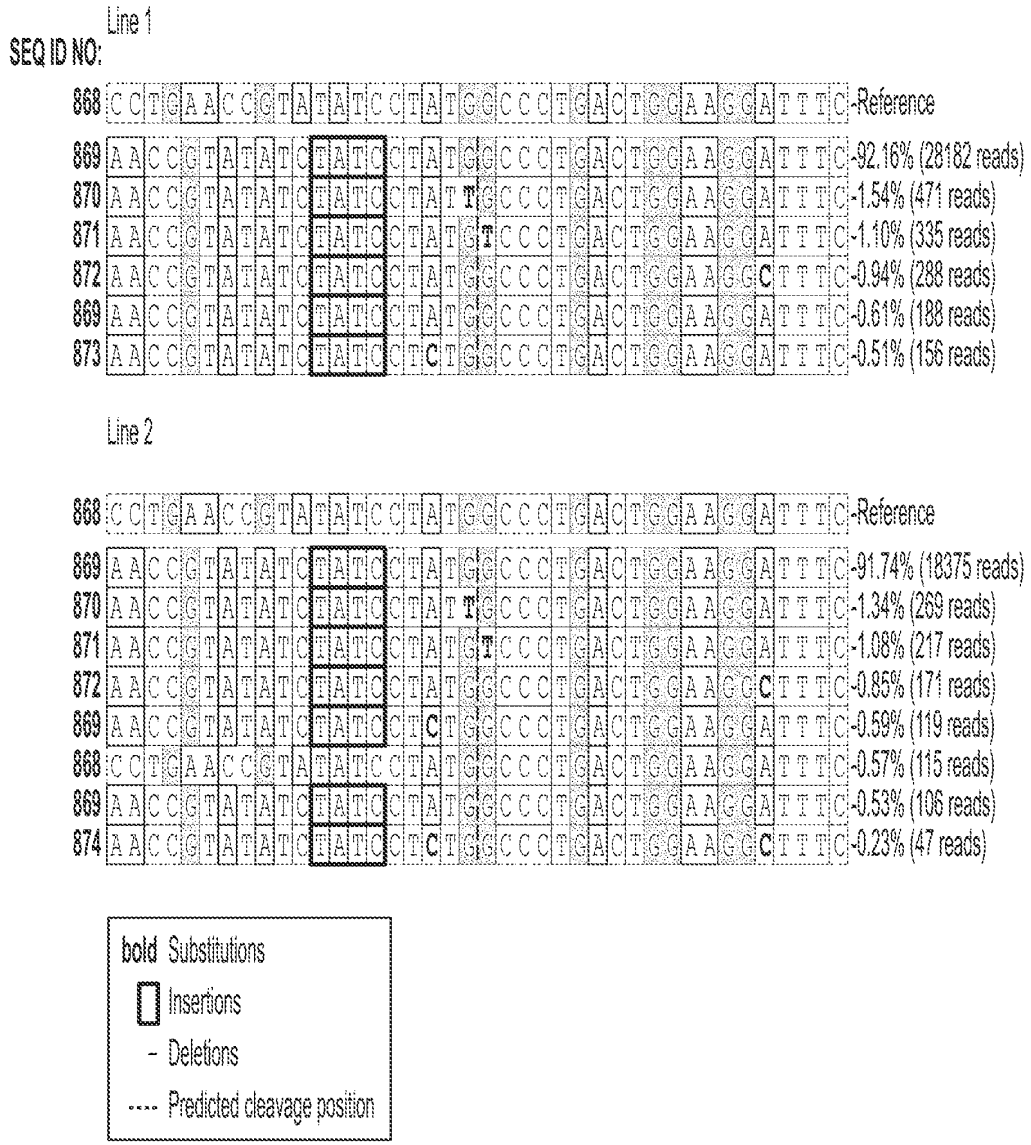

SEQ ID NO:
Line 1
868 CCTGAACGTATATCCTATGGCCCTGACTGGAAGATTTC -Reference
869 AACCGTATATCTATCCTATGGCCCTGACTGGAAGATTTC -92.16% (28162 reads)
870 AACCGTATATCTATCCTATTGCCCTGACTGGAAGATTTC -1.54% (471 reads)
871 AACCGTATATCTATCCTATGTCCCTGACTGGAAGATTTC -1.10% (335 reads)
872 AACCGTATATCTATCCTATGGCCCTGACTGGAAGCTTTC -0.94% (288 reads)
869 AACCGTATATCTATCCTATGGCCCTGACTGGAAGATTTC -0.61% (188 reads)
873 AACCGTATATCTATCCTCTGGCCCTGACTGGAAGATTTC -0.51% (156 reads)
Line 2
868 CCTGAACGTATATCCTATGGCCCTGACTGGAAGATTTC -Reference
869 AACCGTATATCTATCCTATGGCCCTGACTGGAAGATTTC -91.74% (18375 reads)
870 AACCGTATATCTATCCTATTGCCCTGACTGGAAGATTTC -1.34% (269 reads)
871 AACCGTATATCTATCCTATGTCCCTGACTGGAAGATTTC -1.08% (217 reads)
872 AACCGTATATCTATCCTATGGCCCTGACTGGAAGCTTTC -0.85% (171 reads)
869 AACCGTATATCTATCCTCTGGCCCTGACTGGAAGATTTC -0.59% (119 reads)
868 CCTGAACGTATATCCTATGGCCCTGACTGGAAGATTTC -0.57% (115 reads)
869 AACCGTATATCTATCCTATGGCCCTGACTGGAAGATTTC -0.53% (106 reads)
874 AACCGTATATCTATCCTCTGGCCCTGACTGGAAGCTTTC -0.23% (47 reads)
bold Substitutions
Insertions
- Deletions
---- Predicted cleavage position

FIG. 53C

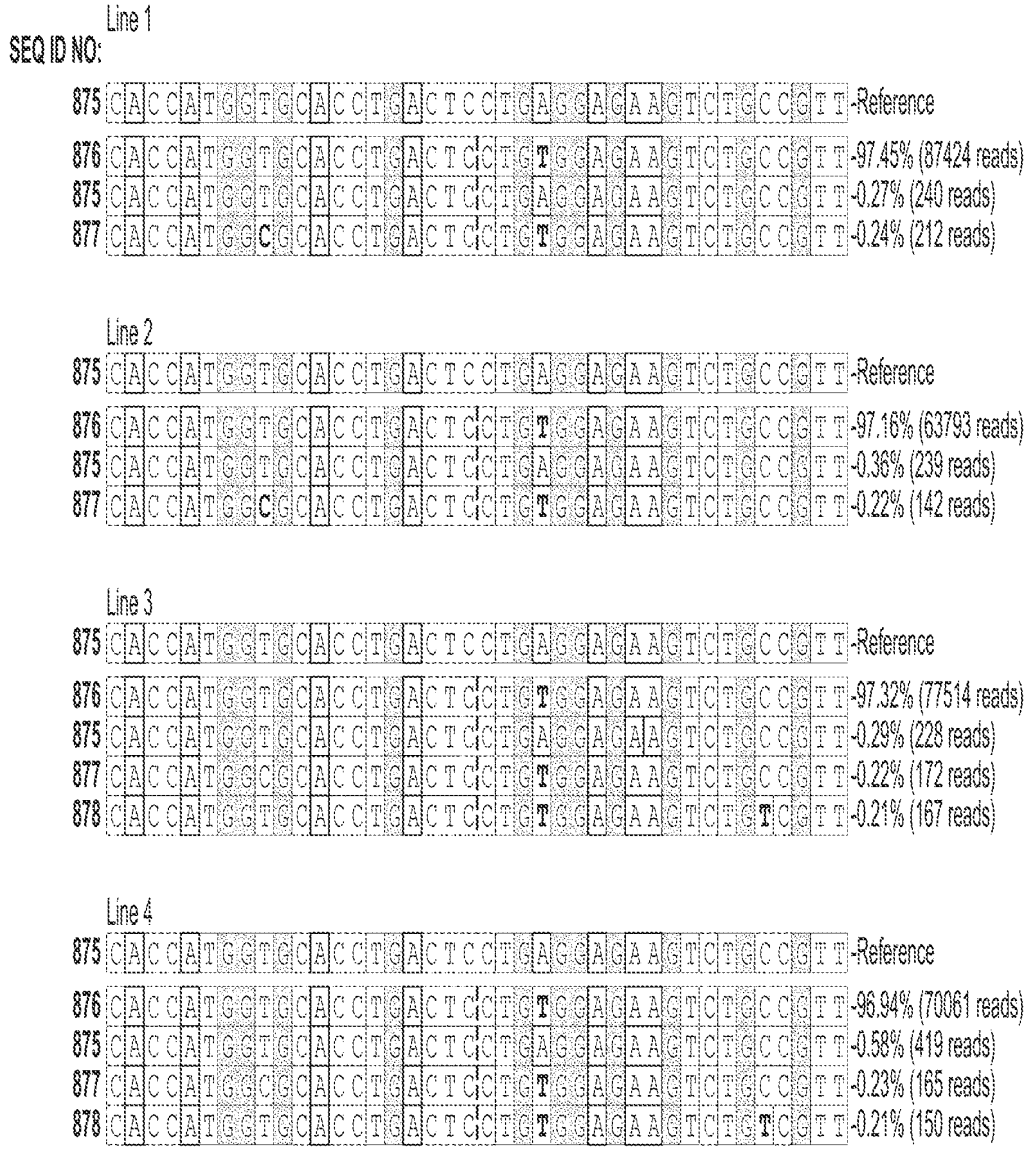

SEQ ID NO:
Line 1
875 CACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT -Reference
876 CACCATGGTGCACCTGACTCCTGTGGAGAAGTCTGCCGTT -97.45% (87424 reads)
875 CACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT -0.27% (240 reads)
877 CACCATGGCGCACCTGACTCCTGTGGAGAAGTCTGCCGTT -0.24% (212 reads)
Line 2
875 CACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT -Reference
876 CACCATGGTGCACCTGACTCCTGTGGAGAAGTCTGCCGTT -97.16% (63793 reads)
875 CACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT -0.36% (239 reads)
877 CACCATGGCGCACCTGACTCCTGTGGAGAAGTCTGCCGTT -0.22% (142 reads)
Line 3
875 CACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT -Reference
876 CACCATGGTGCACCTGACTCCTGTGGAGAAGTCTGCCGTT -97.32% (77514 reads)
875 CACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT -0.29% (228 reads)
877 CACCATGGCGCACCTGACTCCTGTGGAGAAGTCTGCCGTT -0.22% (172 reads)
878 CACCATGGTGCACCTGACTCCTGTGGAGAAGTCTGTCGTT -0.21% (167 reads)
Line 4
875 CACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT -Reference
876 CACCATGGTGCACCTGACTCCTGTGGAGAAGTCTGCCGTT -96.94% (70061 reads)
875 CACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT -0.58% (419 reads)
877 CACCATGGCGCACCTGACTCCTGTGGAGAAGTCTGCCGTT -0.23% (165 reads)
878 CACCATGGTGCACCTGACTCCTGTGGAGAAGTCTGTCGTT -0.21% (150 reads)

FIG. 53D

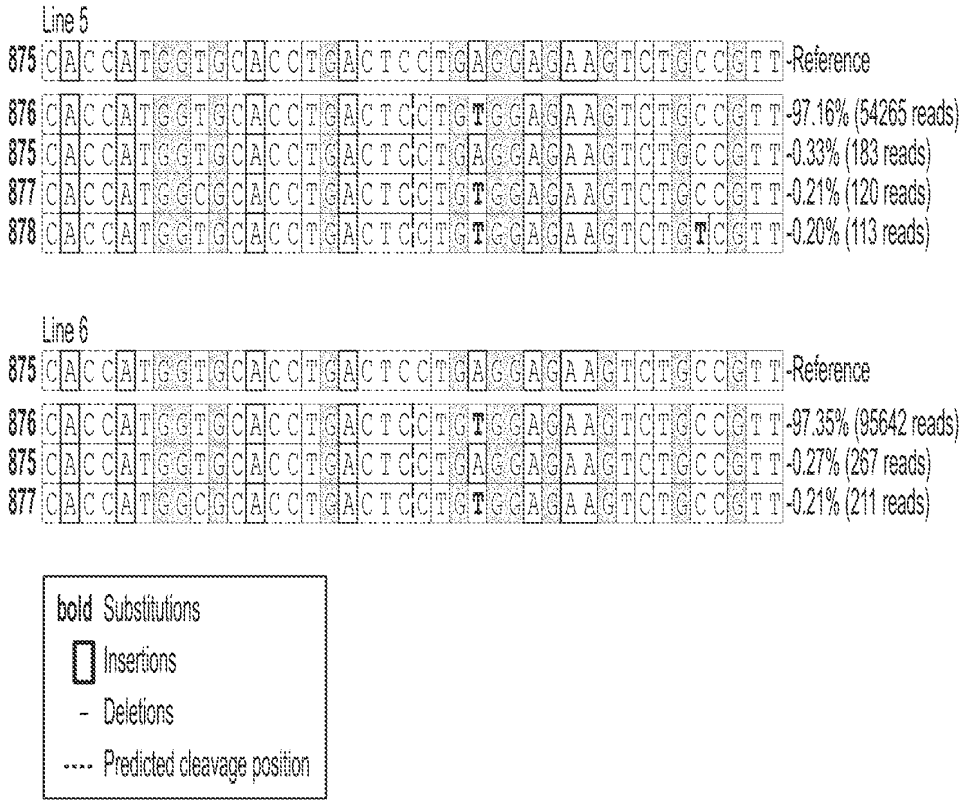
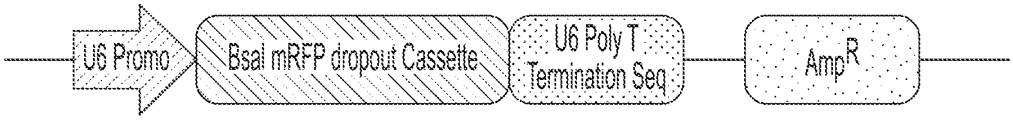

Line 5
875 CACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT -Reference
876 CACCATGGTGCACCTGACTCCTGTGGAGAAGTCTGCCGTT -97.16% (54265 reads)
875 CACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT -0.33% (183 reads)
877 CACCATGGCGCACCTGACTCCTGTGGAGAAGTCTGCCGTT -0.21% (120 reads)
878 CACCATGGTGCACCTGACTCCTGTGGAGAAGTCTGTCGTT -0.20% (113 reads)
Line 6
875 CACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT -Reference
876 CACCATGGTGCACCTGACTCCTGTGGAGAAGTCTGCCGTT -97.35% (95642 reads)
875 CACCATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTT -0.27% (267 reads)
877 CACCATGGCGCACCTGACTCCTGTGGAGAAGTCTGCCGTT -0.21% (211 reads)
bold Substitutions
Insertions
- Deletions
---- Predicted cleavage position

FIG. 53E

Protocol to simultaneously clone a spacer sequence (Target Site) and pegRNA 3' extension containing a PBS and RT template into a human U6 promotor pegRNA expression vector:

[Component 1] Undigested pU6-pegRNA-GG-Vector Schematic:



Plasmid	Description	Features	Cut with
pU6-pegRNA-GG-Vector	Human pegRNA U6 expression vector	U6_BsaI-mRFP-cassette_U6-Term	BsaI-HFV2

[Component 2] Ordering oligonucleotides to generate the spacer (target) sequence for human pegRNA vectors:

Top and bottom oligos to be annealed and then assembled into digested pU6-pegRNA-GG-Vector:

SEQ ID NO: 879 5'-CACCNNNNNNNNNNNNNNNNNNNNGTTTT - 3'

SEQ ID NO: 880 3' - NNNNNNNNNNNNNNNNNNNNCAAAATCTC - 5'

(The N20 mock protospacer sequence shown can be of variable length, as needed)

[Component 3] Ordering oligonucleotides to generate the SpCas9 sgRNA scaffold for human pegRNA vectors:

Top and bottom oligos to be annealed and then assembled into digested pU6-pegRNA-GG-Vector

(Phos) 5' - AGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCG - 3' SEQ ID NO: 881

SEQ ID NO: 882 3' - GATCTTTATCGTTCAATTTTATTCCGATCAGGCAATAGTTGAACTTTTTCACCGTGGCTCAGCCACG - 5' **(Phos)**

(Note that these oligos **must be 5' phosphorylated** to permit pegRNA vector assembly)

[Component 4] Ordering oligonucleotides to generate the pegRNA 3' extension containing a PBS and RT template sequence for human pegRNA vectors:

Top and bottom oligos to be annealed and then assembled into digested pU6-pegRNA-GG-Vector:

SEQ ID NO: 883 5' -GTGCNNNNNNNNNNNNNNNNNNNNNNNNNNNN - 3'

SEQ ID NO: 884 3' - NNNNNNNNNNNNNNNNNNNNNNNNNNNNAAA - 5'

(The N30 mock reverse transcriptase template sequence shown can be of variable length, as needed)

Schematic of final assembled U6 pegRNA expression vector from components 1, 2, 3, and 4:

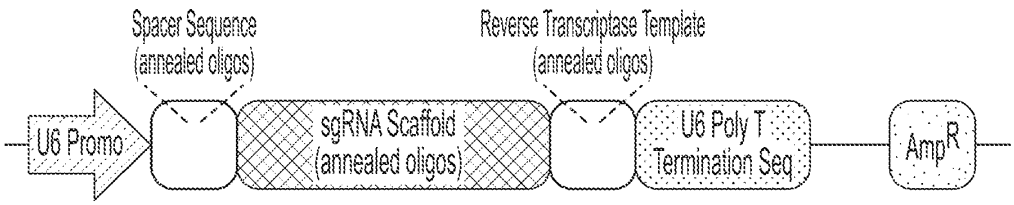


FIG. 54

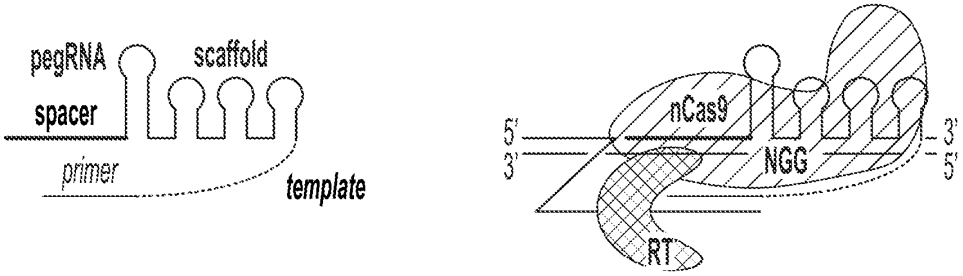


FIG. 55A

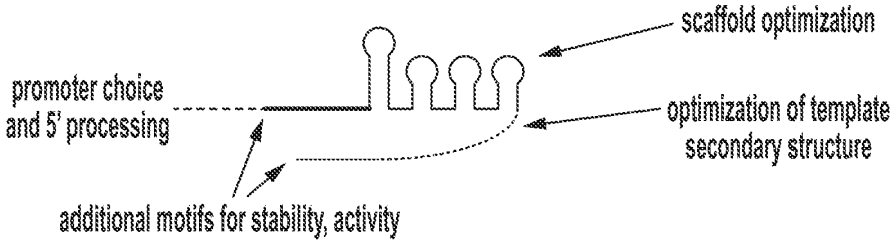


FIG. 55B

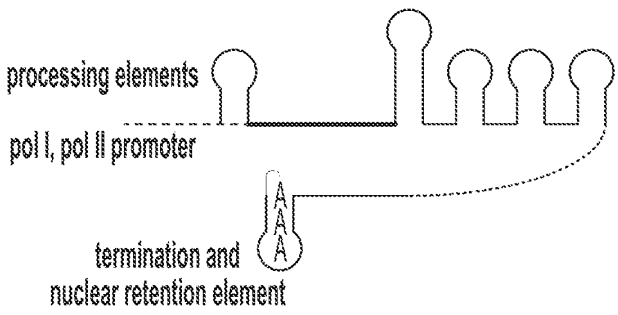


FIG. 55C

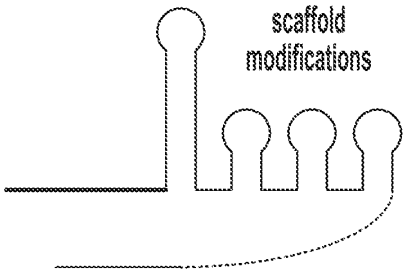


FIG. 55D

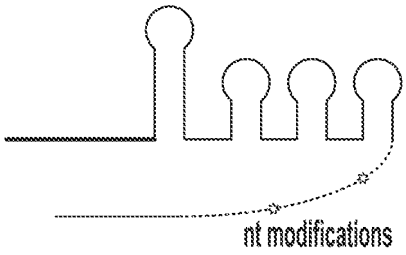


FIG. 55E

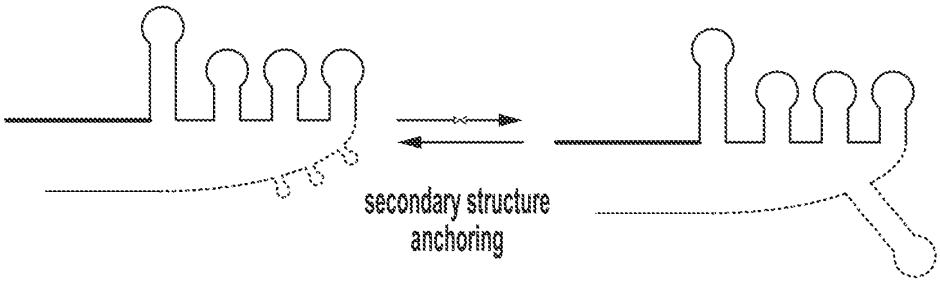


FIG. 55F

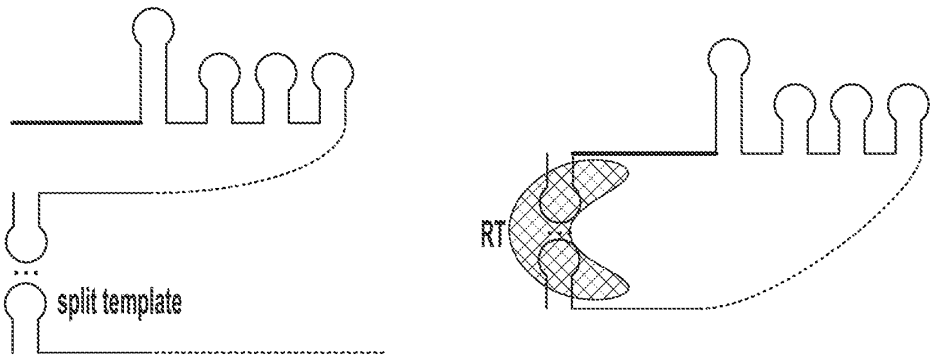


FIG. 55G

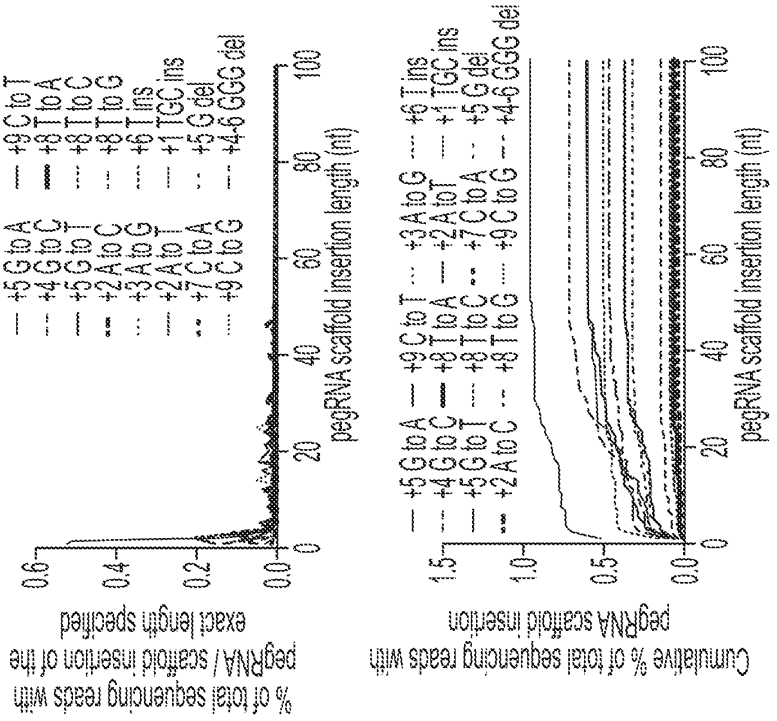
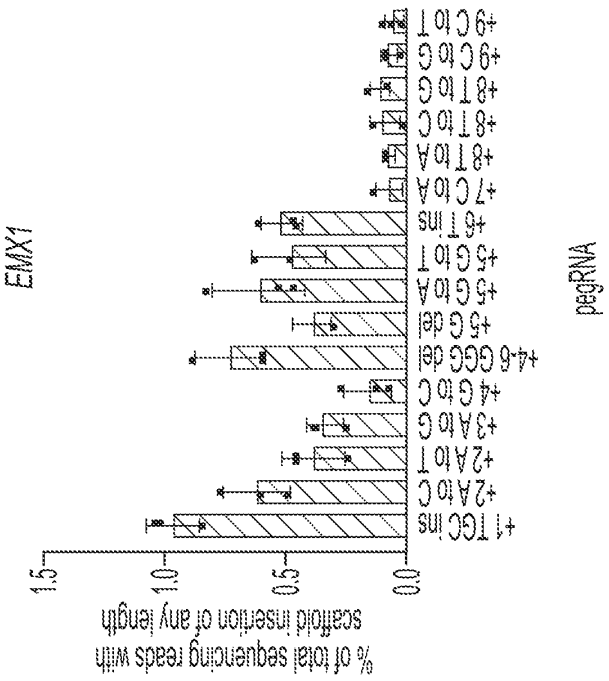

% of total sequencing reads with pegRNA / scaffold insertion of the exact length specified
pegRNA scaffold insertion length (nt)
+5 G to A
+4 G to C
+5 G to T
+2 A to C
+3 A to G
+2 A to T
+7 C to A
+9 C to G
+9 C to T
+8 T to A
+8 T to C
+8 T to G
+6 T ins
+1 TGC ins
+5 G del
+4-6 GGG del
Cumulative % of total sequencing reads with pegRNA scaffold insertion
pegRNA scaffold insertion length (nt)

FIG. 56A


EMX1
% of total sequencing reads with scaffold insertion of any length
pegRNA
+1 TGC ins
+2 A to C
+2 A to T
+3 A to G
+4 G to C
+4-6 GGG del
+5 G del
+5 G to A
+5 G to T
+6 T ins
+7 C to A
+8 T to A
+8 T to C
+8 T to G
+9 C to G
+9 C to T

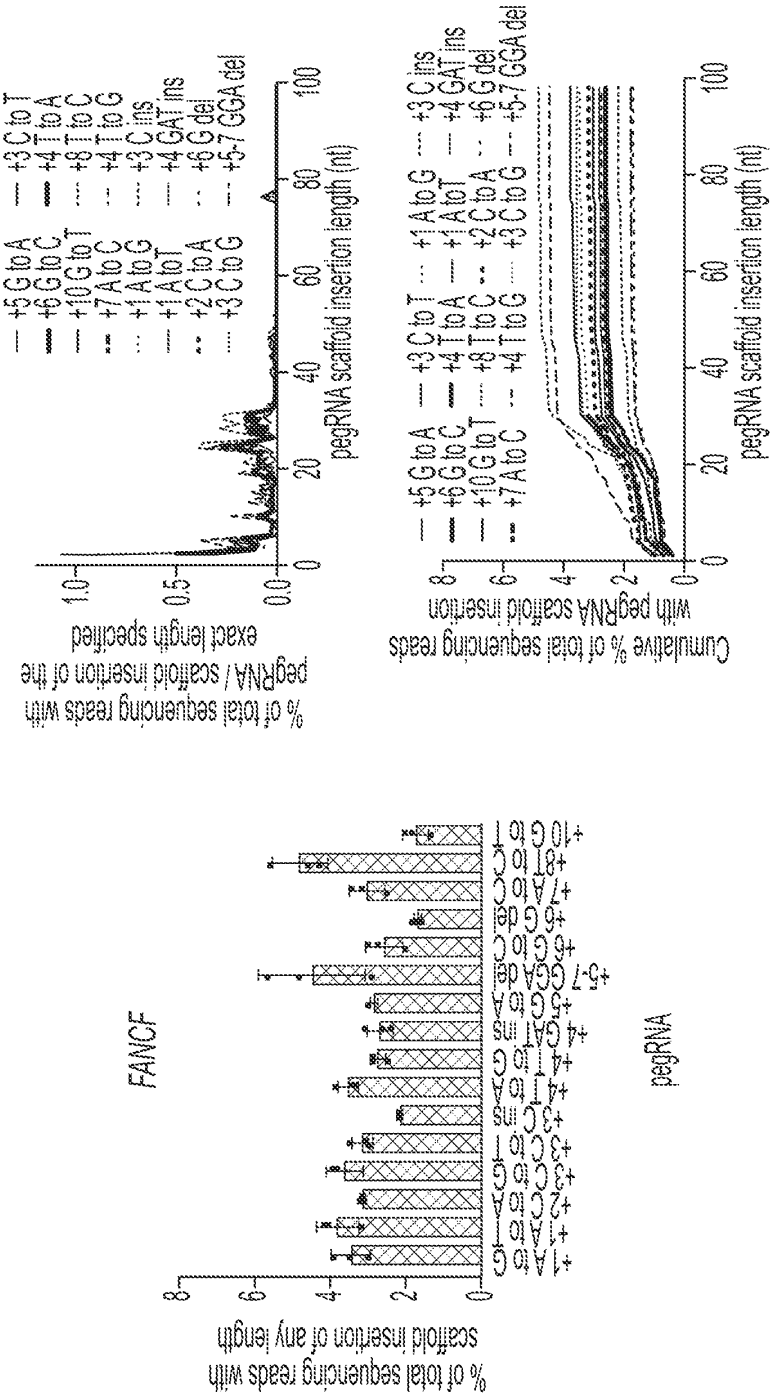

% of total sequencing reads with pegRNA / scaffold insertion of the exact length specified
pegRNA scaffold insertion length (nt)
+5 G to A
+6 G to C
+10 G to T
+7 A to C
+1 A to G
+1 A to T
+2 C to A
+3 C to G
+3 C to T
+4 T to A
+8 T to C
+4 T to G
+3 C ins
+4 GAT ins
+6 G del
+5-7 GGA del
Cumulative % of total sequencing reads with pegRNA scaffold insertion
pegRNA scaffold insertion length (nt)
+5 G to A
+6 G to C
+10 G to T
+7 A to C
+3 C to T
+1 A to T
+4 T to A
+8 T to C
+4 T to G
+2 C to A
+3 C to G
+3 C ins
+4 GAT ins
+6 G del
+5-7 GGA del
% of total sequencing reads with scaffold insertion of any length
FANCF
pegRNA
+1 A to G
+1 A to T
+2 C to A
+3 C to G
+3 C to T
+3 C ins
+4 T to A
+4 T to G
+4 GAT ins
+5 G to A
+5-7 GGA del
+6 G to C
+6 G del
+7 A to C
+8 T to C
+10 G to T

FIG. 56B

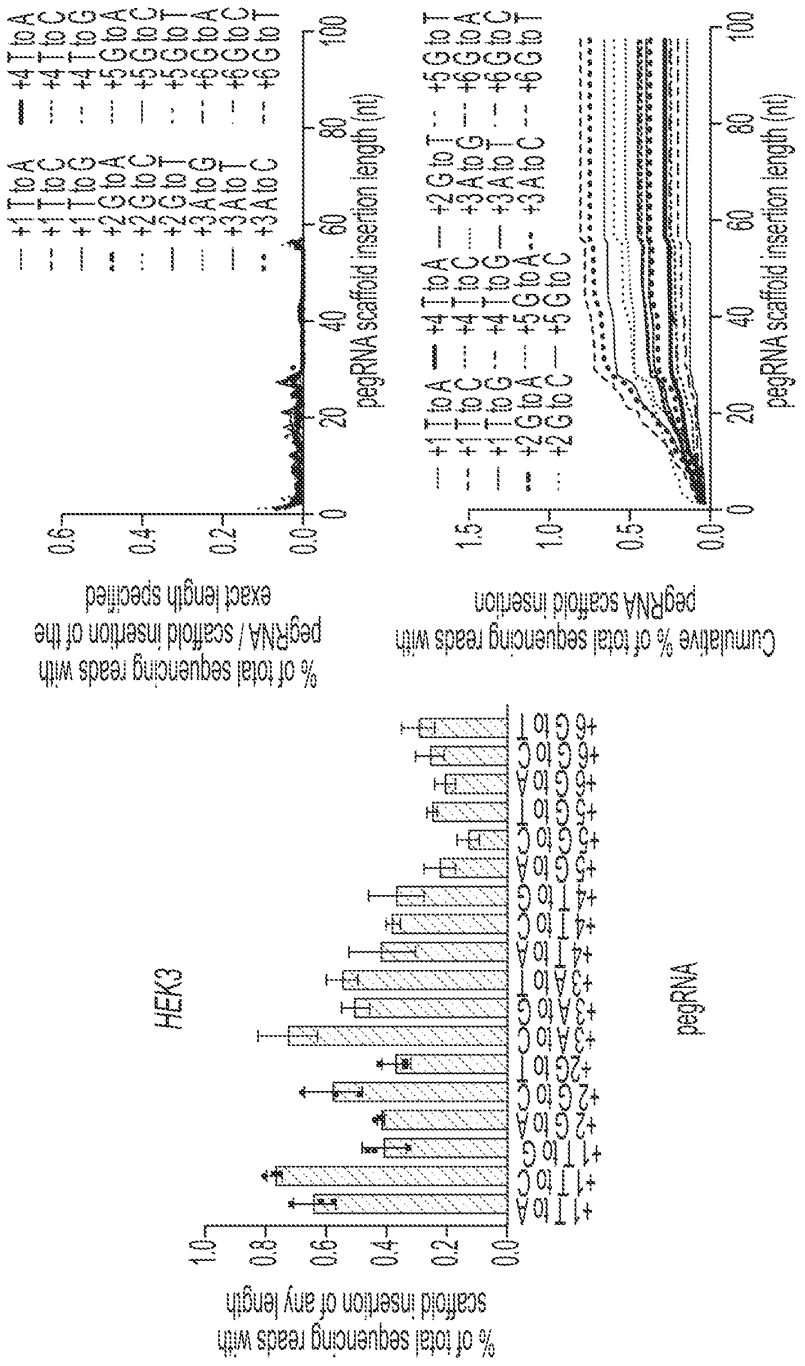

HEK3
% of total sequencing reads with scaffold insertion of any length
pegRNA
+1T to A
+1T to C
+1T to G
+2G to A
+2G to C
+2G to T
+3A to C
+3A to G
+3A to T
+4T to A
+4T to C
+4T to G
+5G to A
+5G to C
+5G to T
+6G to A
+6G to C
+6G to T
% of total sequencing reads with pegRNA / scaffold insertion of the exact length specified
pegRNA scaffold insertion length (nt)
Cumulative % of total sequencing reads with pegRNA scaffold insertion
pegRNA scaffold insertion length (nt)

FIG. 56C

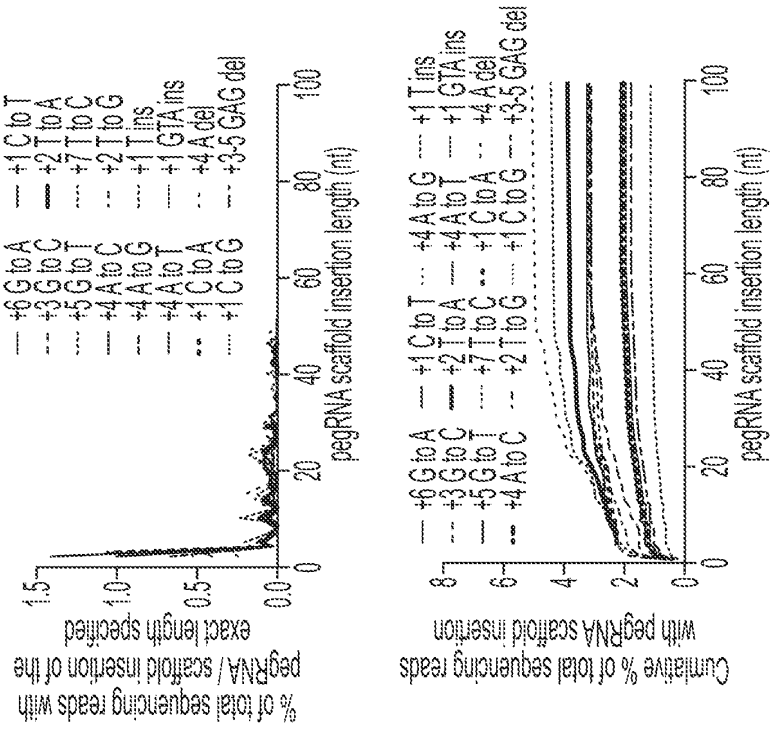
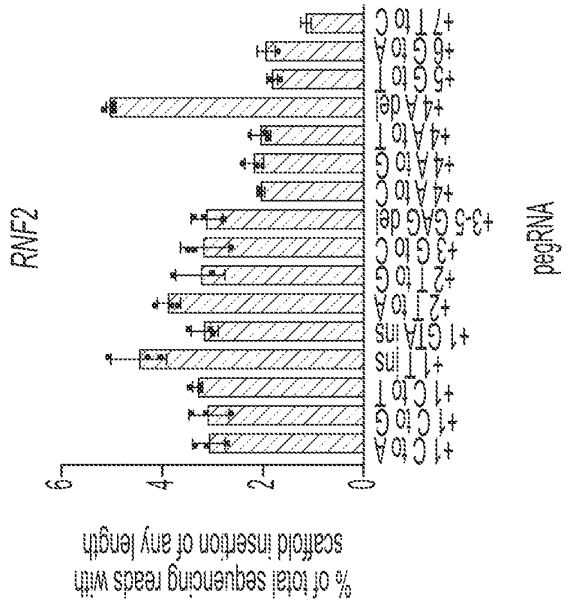

% of total sequencing reads with pegRNA / scaffold insertion of the exact length specified
pegRNA scaffold insertion length (nt)
— +6 G to A
-- +3 G to C
··· +5 G to T
— +4 A to C
-- +4 A to G
··· +4 A to T
— +1 C to A
-- +1 C to G
··· +1 C to T
-- +2 T to A
··· +7 T to C
-- +2 T to G
··· +1 T ins
— +1 GTA ins
-- +4 A del
··· +3-5 GAG del
Cumulative % of total sequencing reads with pegRNA scaffold insertion
pegRNA scaffold insertion length (nt)
— +6 G to A
-- +3 G to C
··· +5 G to T
— +4 A to C
-- +4 A to G
··· +4 A to T
— +1 C to A
-- +1 C to G
··· +1 C to T
-- +2 T to A
··· +7 T to C
-- +2 T to G
··· +1 T ins
— +1 GTA ins
-- +4 A del
··· +3-5 GAG del

FIG. 56D


% of total sequencing reads with scaffold insertion of any length
RNF2
pegRNA
+1C to G
+1C to A
+1C to T
+1T ins
+1GTA ins
+2T to A
+2T to G
+3G to T
+3-5 GAG del
+4A to C
+4A to G
+4A to T
+4A del
+5G to T
+6G to A
+7T to C

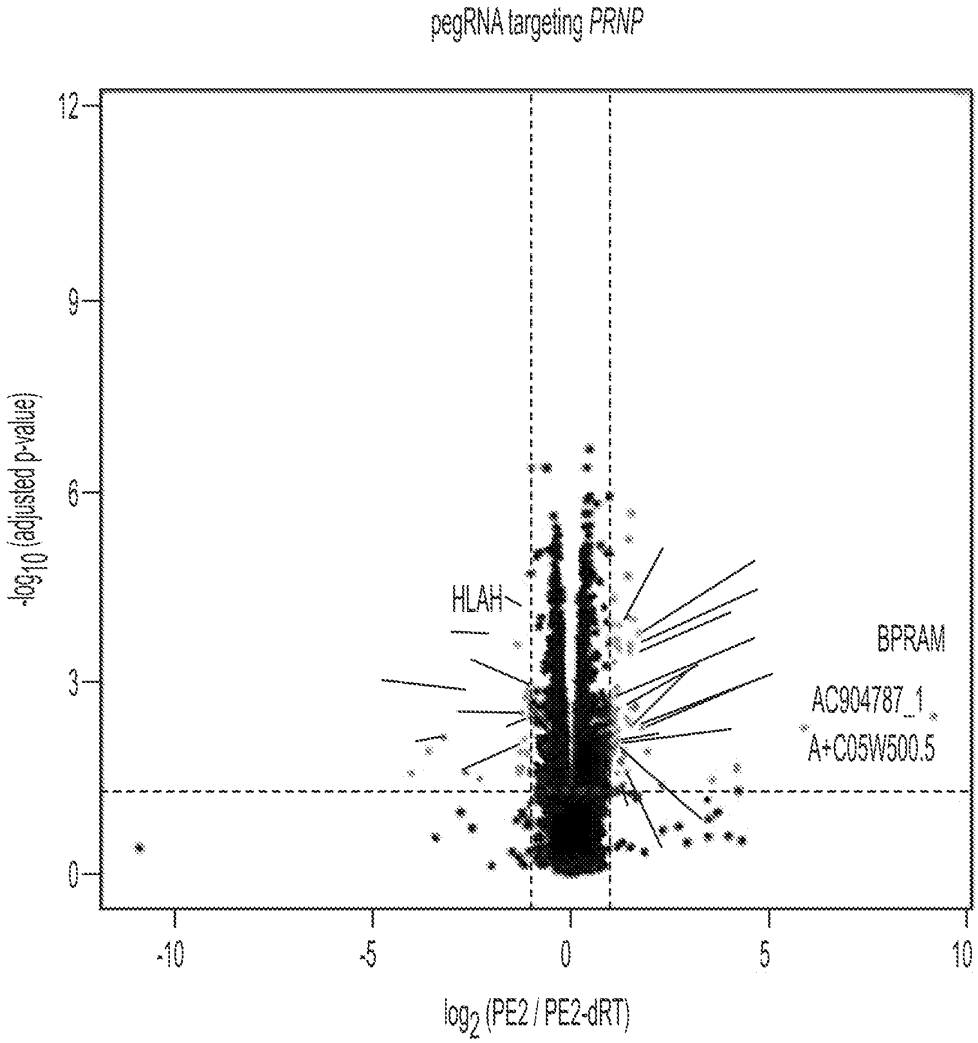

pegRNA targeting PRNP
12
9
6
3
0
-log10 (adjusted p-value)
HLAH
BPRAM
AC904787_1
A+C05W500.5
-10
-5
0
5
10
log2 (PE2 / PE2-dRT)

FIG. 57A

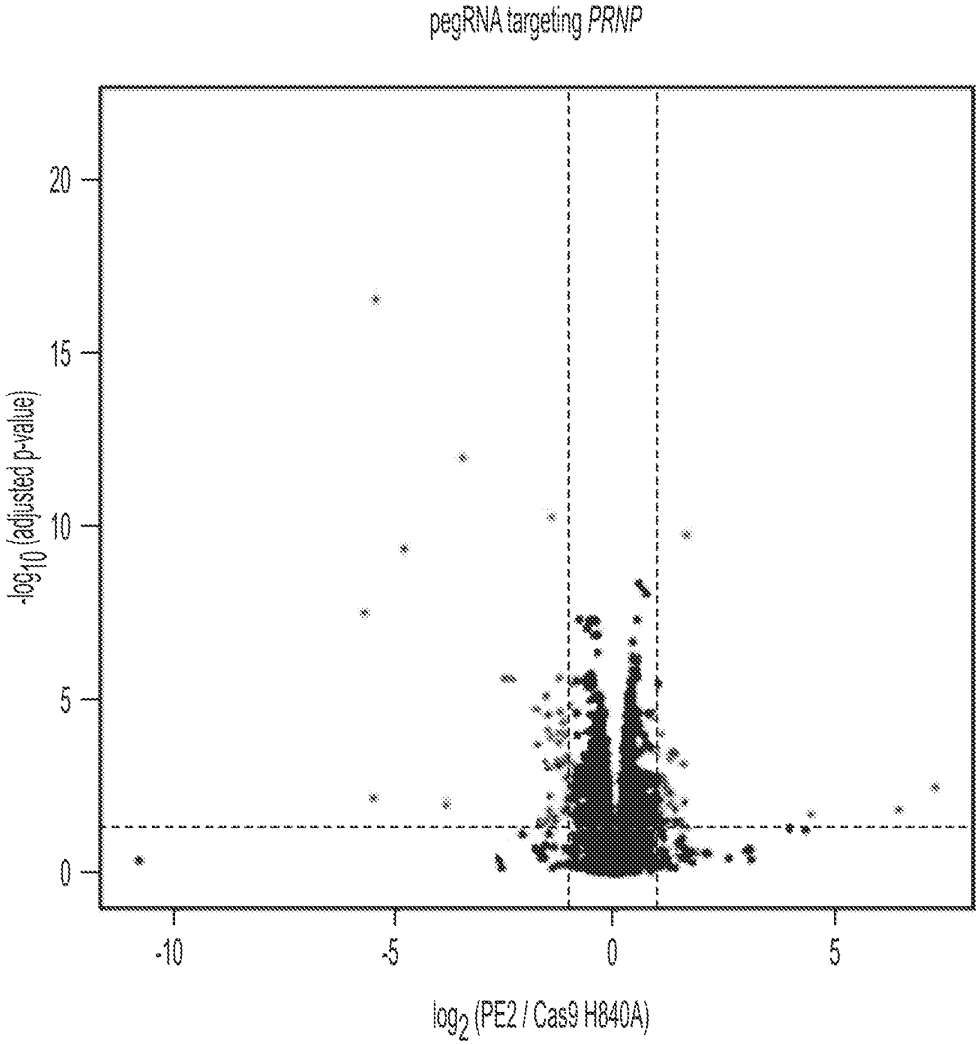

pegRNA targeting PRNP
20
15
10
5
0
-log10 (adjusted p-value)
-10
-5
0
5
log2 (PE2 / Cas9 H840A)

FIG. 57B

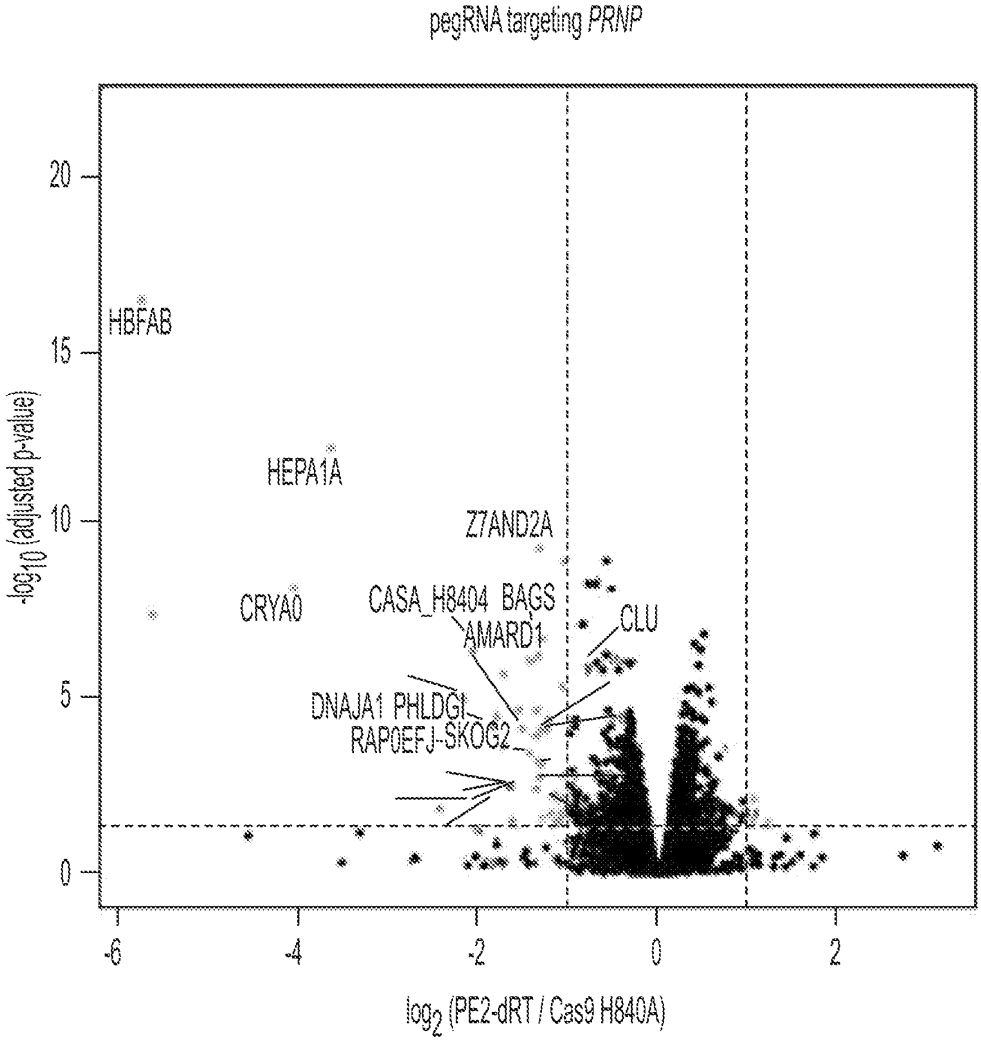

pegRNA targeting PRNP
20
15
10
5
0
-log10 (adjusted p-value)
HBFAB
HEPA1A
Z7AND2A
CRYAB
CASA_H8404
BAGS
AMARD1
CLU
DNAJA1
PHLDG1
RAP0EFJ-SKOG2
-6
-4
-2
0
2
log2 (PE2-dRT / Cas9 H840A)

FIG. 57C

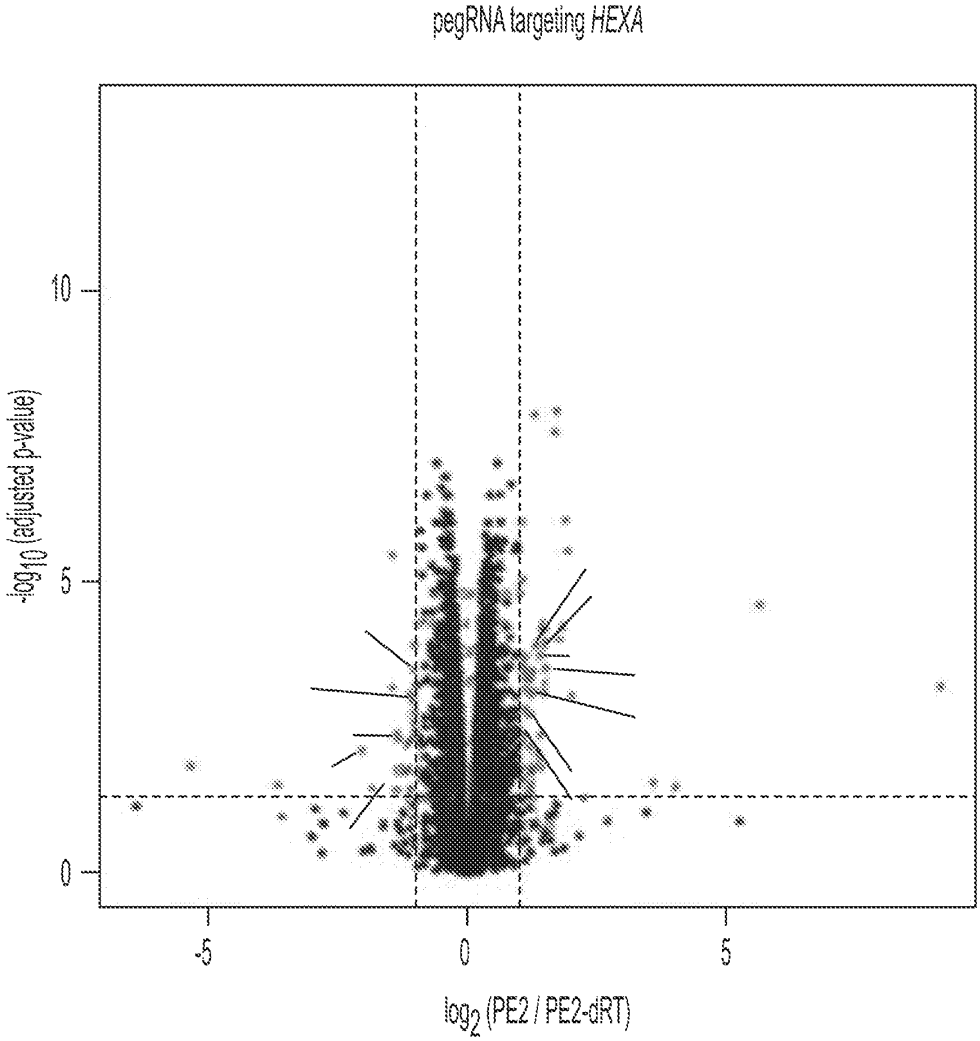

pegRNA targeting HEXA
10
5
0
-log10 (adjusted p-value)
-5
0
5
log2 (PE2 / PE2-dRT)

FIG. 57D

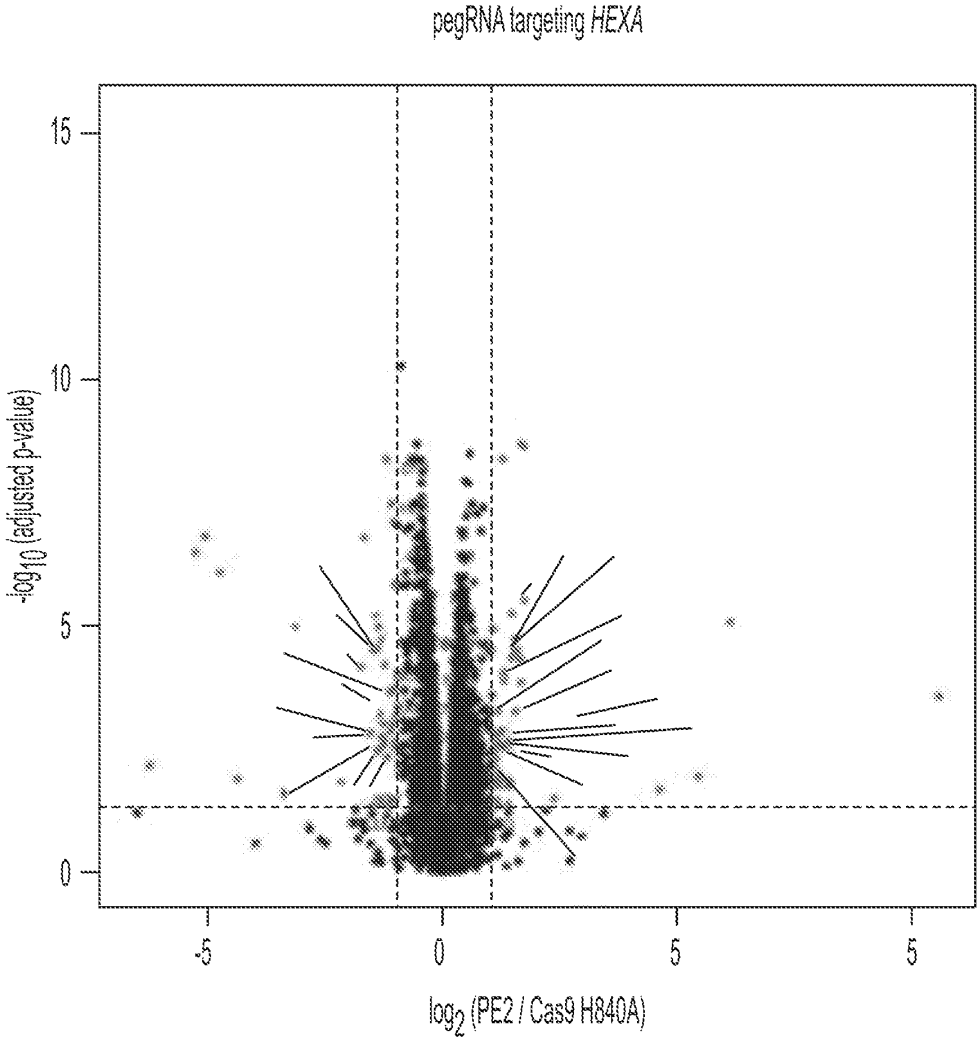

pegRNA targeting HEXA
15
10
5
0
-log10 (adjusted p-value)
-5
0
5
5
log2 (PE2 / Cas9 H840A)

FIG. 57E

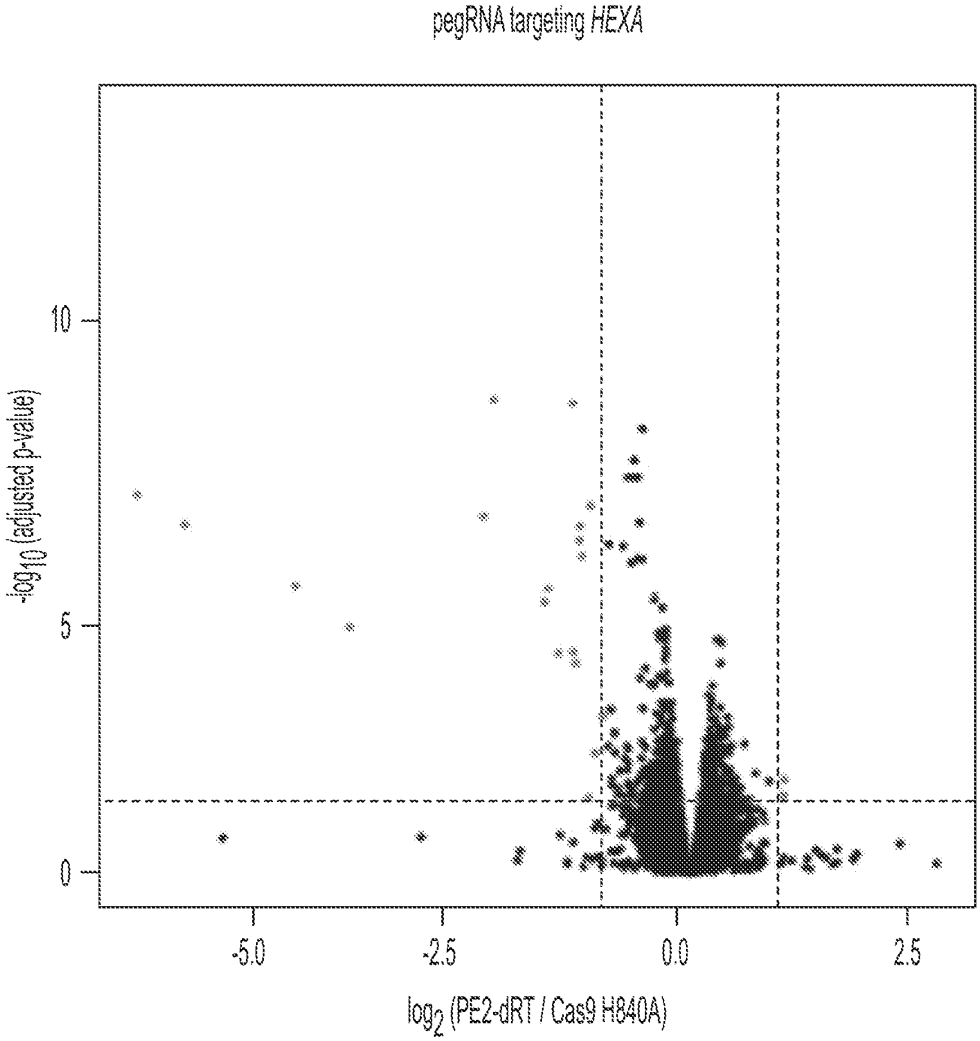

pegRNA targeting HEXA
10
5
0
-log10 (adjusted p-value)
-5.0
-2.5
0.0
2.5
log2 (PE2-dRT / Cas9 H840A)

FIG. 57F

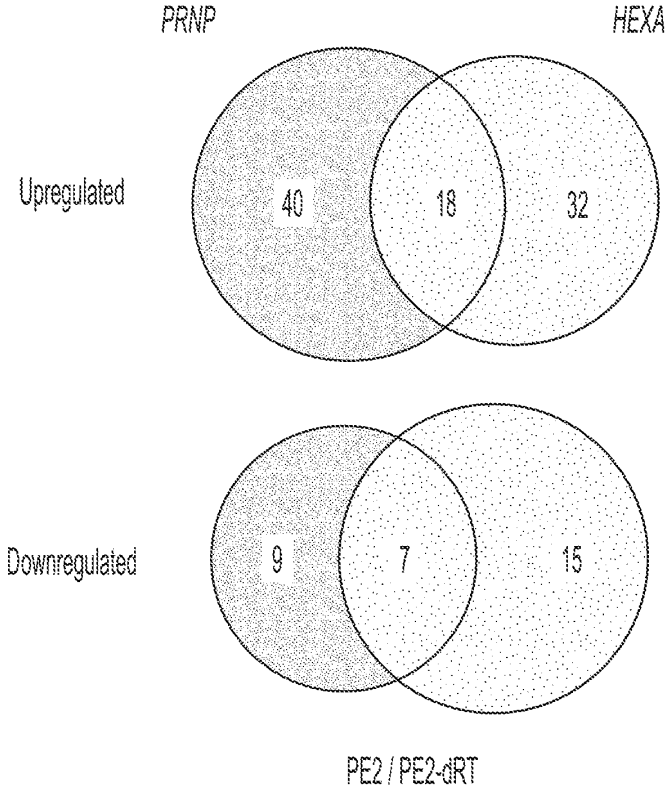

PRNP
HEXA
Upregulated
40
18
32
Downregulated
9
7
15
PE2 / PE2-dRT

FIG. 57G

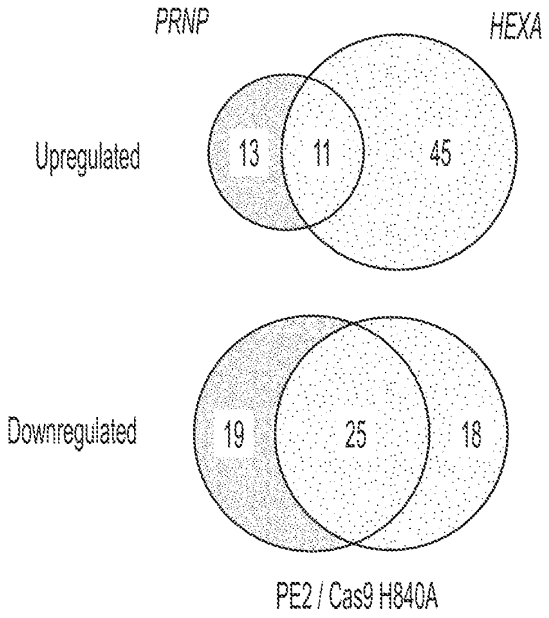


FIG. 57H

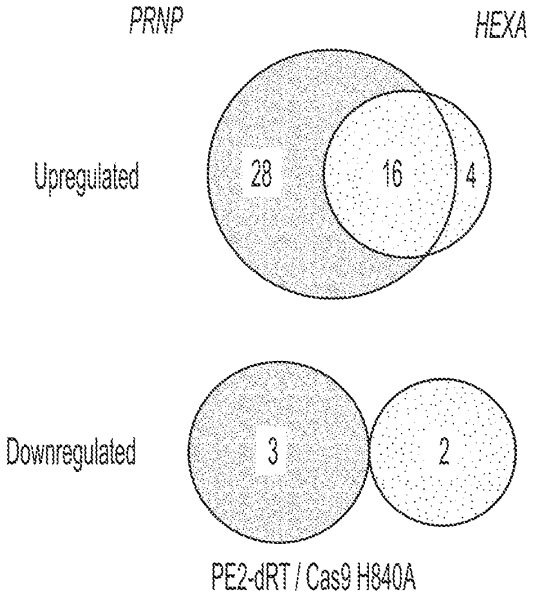


FIG. 57I

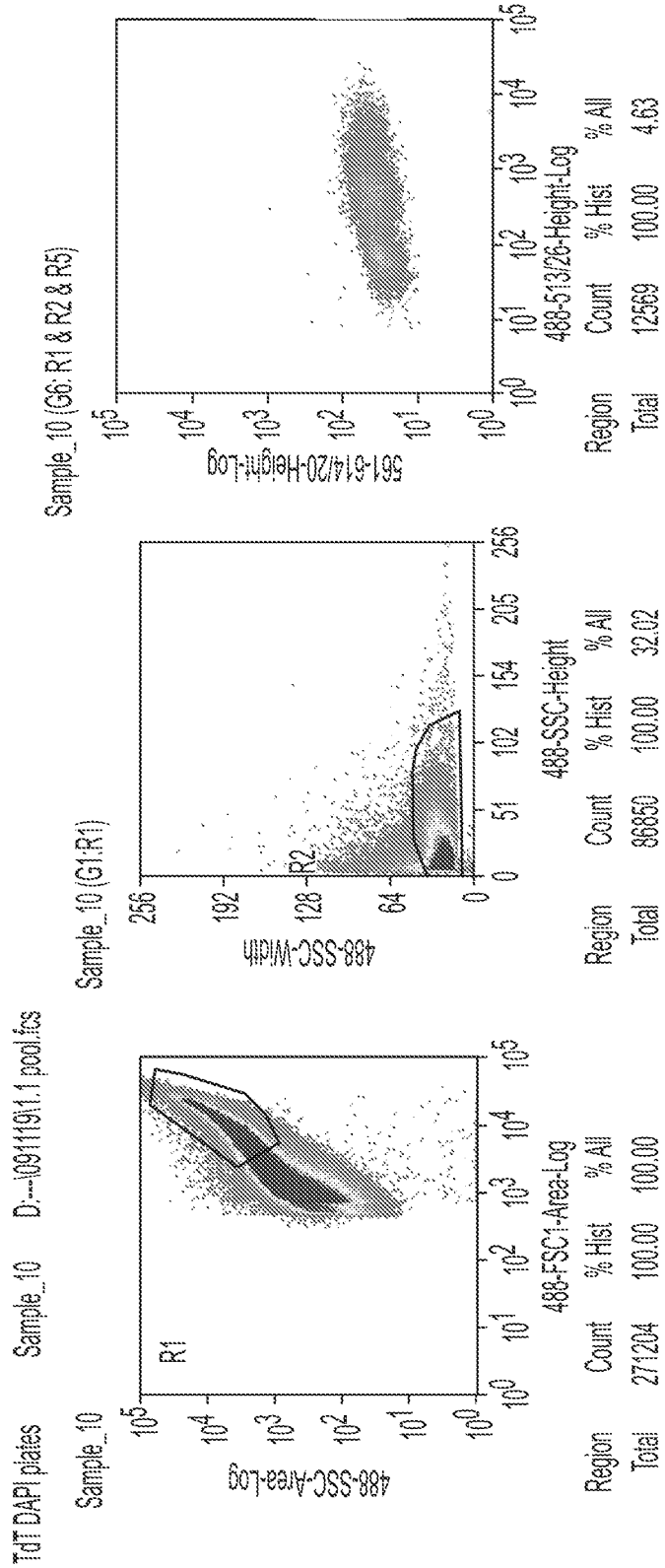

TdT DAPI plates
Sample_10
D:\---\09111911.1 pool.fcs
Sample_10
488-SSC-Area-Log
488-FSC1-Area-Log
R1
Region Count % Hist % All
Total 271204 100.00 100.00
R1 66850 32.02 32.02
Sample_10 (G1:R1)
488-SSC-Width
488-SSC-Height
R2
Region Count % Hist % All
Total 86850 100.00 32.02
R2 81289 93.60 29.97
Sample_10 (G6: R1 & R2 & R5)
561-614/20-Height-Log
488-513/26-Height-Log
Region Count % Hist % All
Total 12569 100.00 4.63

FIG. 58A

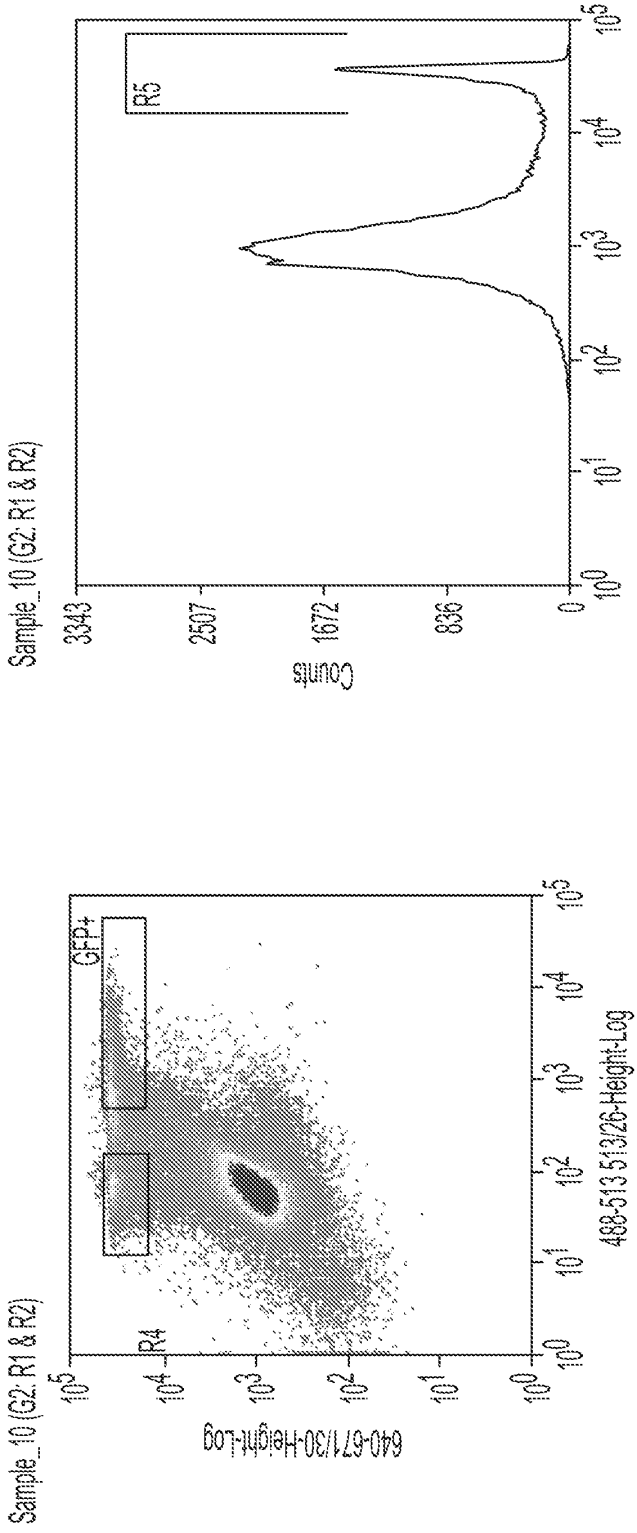

Sample_10 (G2: R1 & R2)
640-671/30-Height-Log
488-513/26-Height-Log
GFP+
R4
Sample_10 (G2: R1 & R2)
Counts
640-671/30-Height-Log
R5
Region Count % Hist Mean CV
Total 81289 100.00 6521.68 179.64
R5 12569 15.46 32516.62 24.27
Region Count % Hist % All Bounds Mode Count
Total 81289 100.00 29.97 (1.00,1000000.04) (100000.04,1... 128
GFP + 4418 5.43 1.63 (495.60,48559.55) (58170.94,1... 40
R4 4709 5.79 1.74 (11.98,46415.91) (150.13,1501... 34
zniziolek Wednesday, September 11, 2019 01:36:4 1PM Page 1

FIG. 58B

Cloning overview

1. **Digest pU6-pegRNA-GG-Vector plasmid (component 1)** with *Bsa*I and isolate the plasmid fragment (~2.2kb) containing the origin of replication, U6 promoter, U6 poly-T termination sequence, and Amp^R gene
2. **Order oligonucleotides for:**
   a. The desired spacer (target) sequence flanked by indicated overhangs (**component 2**)
      i. Use the desired target's 5'-3' sequence for the top strand oligonucleotide (including the 5' CACC and 3' GTTTT overhangs) and use the reverse complement of the target sequence for the bottom strand oligonucleotide (including the 5' CTCTAAAAC overhang). Spacer sequences must begin with a G nucleotide for efficient transcription initiation.
   b. The desired pegRNA 3' extension template flanked by the indicated overhangs (**component 3**)
      i. Use the RNA sense sequence as the top strand oligonucleotide (featuring the 5' GTGC overhang) and use the reverse complement of this sequence for the bottom strand oligonucleotide (featuring the 5' AAAA overhang).
   c. SpCas9 sgRNA scaffold sequence featuring compatible golden gate overhangs (**component 4**)
      i. These oligonucleotides are not the complete scaffold sequence, as overhangs from the remaining components contribute several missing nucleotides
      ii. Note: these oligonucleotides must be 5' phosphorylated. Oligonucleotides can be 5' phosphorylated by the manufacturer or 5' phosphorylated enzymatically using T4 PNK (see protocol below)
3. **Anneal top and bottom oligonucleotides** for components 2, 3, and 4 in separate annealing reactions according to the protocol below. If the SpCas9 sgRNA scaffold sequence (component 4) was not phosphorylated, phosphorylate with T4 PNK.
4. **Golden Gate assembly of isolated 2.2-kb fragment from component 1 with components 2, 3, and 4**
5. **Transform the ligation product into *E. coli*.** The antibiotic resistance conferred by component 1 from the pU6- pegRNA-GG-vector plasmid is ampicillin and carbenicillin resistance.
6. **Isolate and sequence** plasmids from the resulting clonal transformants

FIG. 59A

pegRNA cloning protocol

Step 1: Digest pU6-pegRNA-GG-Vector plasmid (component 1)

Combine the following in a PCR tube:

2000 ng pU6-pegRNA-GG-Vector (component 1)	X µL
Bsa1-HFv2 (NEB)	1.0 µL
10x Cutsmart Buffer	3.0 µL
H_2O	to 30.0 µL
Total reaction volume	30.0 µL

Incubate at 37 °C for 4-16 hours
Isolate ~2.2-kb fragment from cut plasmid.

FIG. 59B

Steps 2 and 3: Order and anneal oligonucleotide parts (components 2, 3, and 4)

Materials

Annealing buffer: H_2O supplemented with 10 mM Tris-Cl pH 8.5 and 50 mM NaCl
Complementary oligonucleotide pairs

Protocol

Combine the following in a PCR tube:

Top oligonucleotide, 100 µM	1.0 µL
Bottom oligonucleotide, 100 µM	1.0 µL
Annealing buffer (components 2, 3, and 4)	23.0 µL
Total reaction volume	25.0 µL

In thermocycler, heat at 95 °C for 3 minutes, then cool gradually (0.1 °C/s) to 22 °C

Dilute annealed oligonucleotides 1:4 by adding 75 µL H_2O. The final concentration of each oligonucleotide will be 1 µM after this dilution. Do not dilute the sgRNA scaffold (component 4) if phosphorylating by PNK in step 2.5.

FIG. 59C

Step 2.b.ii.: sgRNA scaffold phosphorylation (unnecessary if oligonucleotides were purchased phosphorylated)

Protocol

Combine the following in a PCR tube:

4 µM oligonucleotide duplex from step 1	6.25 µL
10x T4 DNA ligase buffer (NEB)	2.50 µL
T4 PNK (NEB)	0.50 µL
H_2O	15.75 µL
Total reaction volume	20.0 µL

In thermocycler, incubate at 37 °C for 60 minutes
Following this phosphorylation, annealed scaffold oligonucleotides are now at a concentration of 1 µM.
Proceed to step 3.

FIG. 59D

Step 4: peg RNA assembly

pegRNA, Golden Gate assembly reaction

Digested pU6-pegRNA-GG plasmid-vector - Pre-cut, isolated 2.2-kb fragment	1.00 µL @ 30 ng/µL
Annealed protospacer oligonucleotides (component 2)	1.00 µL @ 1 µM
Annealed pegRNA 3'-extension oligonucleotides (component 4)	1.00 µL @ 1 µM
sgRNA scaffold annealed oligonucleotides (component 3) - Oligonucleotides *must be phosphorylated*	1.00 µL @ 1 µM
*Bsa*I-HFv2 (NEB)	0.25 µL
T4 DNA ligase (NEB)	0.50 µL
10x T4 DNA ligase buffer (NEB)	1.00 µL
H_2O	4.25 µL
Total reaction volume:	10.0 µL

Incubate at room temperature for 10 min
Following incubation, perform the following program in a thermocycler:
5 min at 37 °C then 5 min at 80 °C then hold at 12 °C

FIG. 59E

Steps 5 and 6: Transformation of assembled plasmids

Transform 1 µL of the 10 µL assembly reaction into 10 µL of competent cells. The desired transformants will be resistant to ampicillin and carbenicillin.

The following diagram summarizes the pegRNA cloning protocol.

FIG. 59F

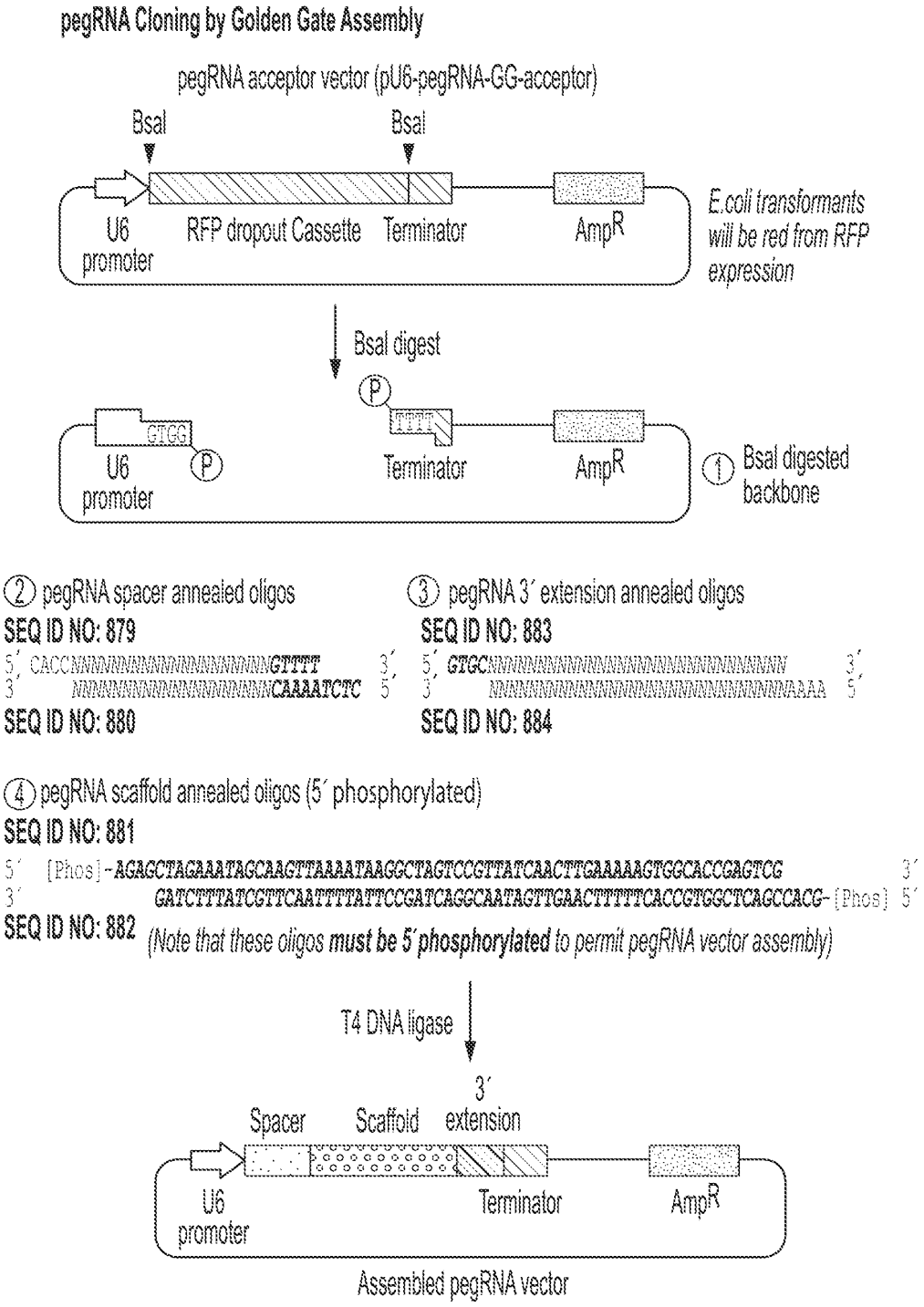

pegRNA Cloning by Golden Gate Assembly
pegRNA acceptor vector (pU6-pegRNA-GG-acceptor)
BsaI
BsaI
U6 promoter
RFP dropout Cassette
Terminator
AmpR
E.coli transformants will be red from RFP expression
BsaI digest
GTGG
P
TTTT
U6 promoter
Terminator
AmpR
1 BsaI digested backbone
2 pegRNA spacer annealed oligos
SEQ ID NO: 879
5′ CACCNNNNNNNNNNNNNNNNNNNGTTTT 3′
3′ NNNNNNNNNNNNNNNNNNNCAAAATCTC 5′
SEQ ID NO: 880
3 pegRNA 3′ extension annealed oligos
SEQ ID NO: 883
5′ GTGCNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3′
3′ NNNNNNNNNNNNNNNNNNNNNNNNNNNNAAA 5′
SEQ ID NO: 884
4 pegRNA scaffold annealed oligos (5′ phosphorylated)
SEQ ID NO: 881
5′ [Phos]-AGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAGTGGCACCGAGTCG 3′
3′ GATCTTTATCGTTCAATTTTATTCCGATCAGGCAATAGTTGAACTTTTTCACCGTGGCTCAGCCACG-[Phos] 5′
SEQ ID NO: 882
(Note that these oligos must be 5′ phosphorylated to permit pegRNA vector assembly)
T4 DNA ligase
3′ extension
Spacer
Scaffold
U6 promoter
Terminator
AmpR
Assembled pegRNA vector

FIG. 59G

```
## sgRNA scaffold sequence search ##

import pandas as pd
import Bio as bio
from Bio import SeqIO
import glob

#generates list of fastq files to analyze
sources = glob.glob('*.fastq')

#reads the fastq files into a dictionary with the file names as keys
fastqdict = {}
for i in range(len(sources)):
        temp = list(SeqIO.parse(sources[i],"fastq"))
        fastqdict[sources[i]]= [str(temp[k].seq) for k in range(len(temp))]

#the referenced sequence to be searched for is entered into the following dictionary with
#an appropriate key
scaffdict =
{'HEK3':'CAGAGGACCGACTCGGTCCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTT
GCTATTTCTAGCTCTAAAACTCACGTGCTCAGTCTGGGCCGGTG', (SEQ ID NO: 885)
'EMX1':'ATCACGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACT
TGCTATTTCTAGCTCTAAAACTTCTTCTTCTGCTCGGACTCGGTG', (SEQ ID NO: 886)
'FANCF':'TTTCCGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAAC
TTGCTATTTCTAGCTCTAAAACGGTGCTGCAGAAGGGATTCCGGTG', (SEQ ID NO: 887)
'RNF2':'TCGTTGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTT
GCTATTTCTAGCTCTAAAACCAGGTAATGACTAAGATGACGGTG'} (SEQ ID NO: 888)

#matches and counts iterative slices of the reference string to the appropriate fastq files
#reference key must be contained in the name of the fastq file
#generated values represent cumulative counts for a minimum degree of sgRNA integration
#i.e. a given value x means x reads contain y or more bases of the scaffold
resultdict = dict.fromkeys(sources)
for key in fastqdict:
   for scaffold in scaffdict:
      if scaffold in str(key):
         resultlist = []
         for j in range(len(scaffdict[scaffold])):
```

FIG. 60A

```
                extent = scaffdict[scaffold][0:(j+1)]
                counter = 0
                for i in range(len(fastqdict[key])):
                    if extent in fastqdict[key][i]:
                        counter = counter + 1
                resultlist.append(counter)
            resultdict[key]=resultlist

#writes the results into a dataframe indexed from 1
resultdf = pd.DataFrame.from_dict(resultdict)
resultdf = resultdf.reindex(sorted(resultdf.columns), axis=1)
resultdf.index = range(1,len(resultdf)+1)

#converts the cumulative count values into specific counts
#i.e. a given value x means x reads contain exactly y bases of the scaffold
resultdf2=resultdf.copy()
for entry in resultdf:
    for i in range(1,len(resultdf[entry])+1):
        try:
            resultdf2[entry][i] = resultdf[entry][i]-resultdf[entry][i+1]
        except:
            resultdf2[entry][i] = resultdf[entry][i]

#converts the specific counts values into frequencies
resultdf3=resultdf2.copy()
for entry in resultdf3:
    resultdf3[entry]=resultdf2[entry].div(resultdf[entry][1])*100

#reads the results into excel files
resultdf.to_excel('cumulativecounts.xlsx')
resultdf2.to_excel('specificcounts.xlsx')
resultdf3.to_excel('specificfrequencies.xlsx')
```

FIG. 60B

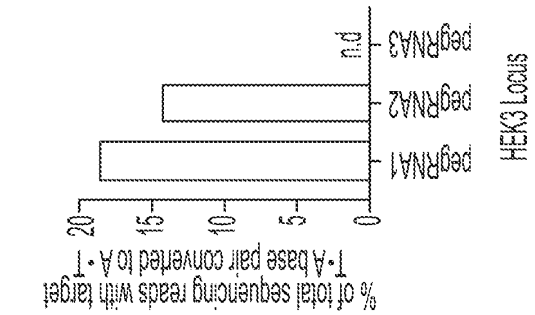


FIG. 62A

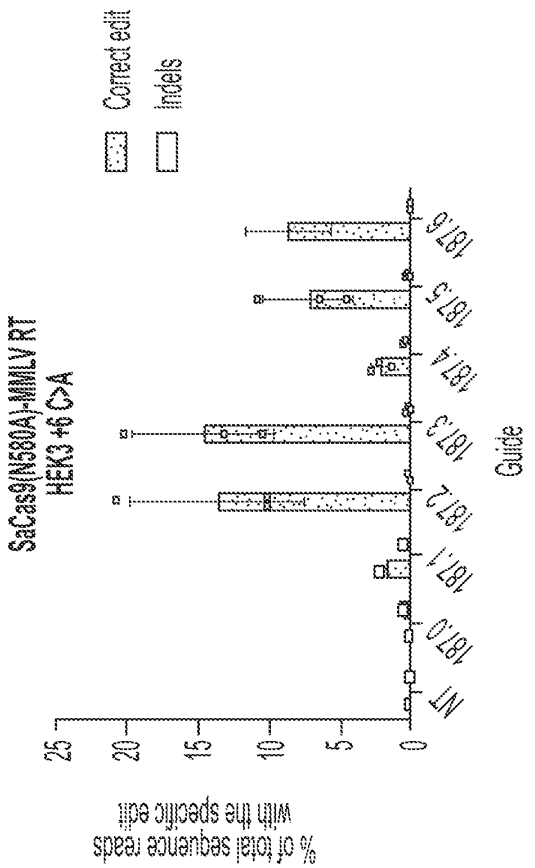


FIG. 61

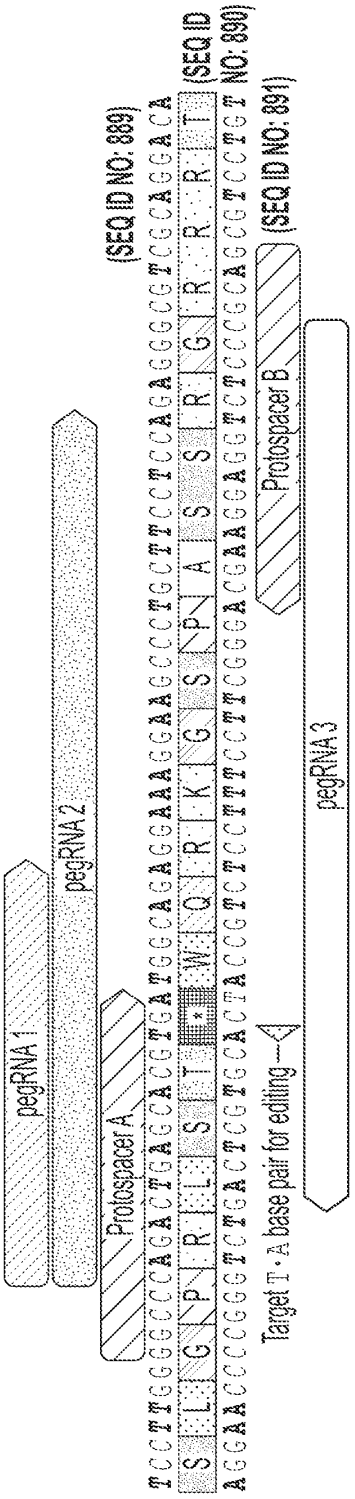

pegRNA 1
pegRNA 2
Protospacer A
pegRNA 3
Protospacer B
Target T • A base pair for editing
(SEQ ID NO: 889)
(SEQ ID NO: 890)
(SEQ ID NO: 891)

FIG. 62B

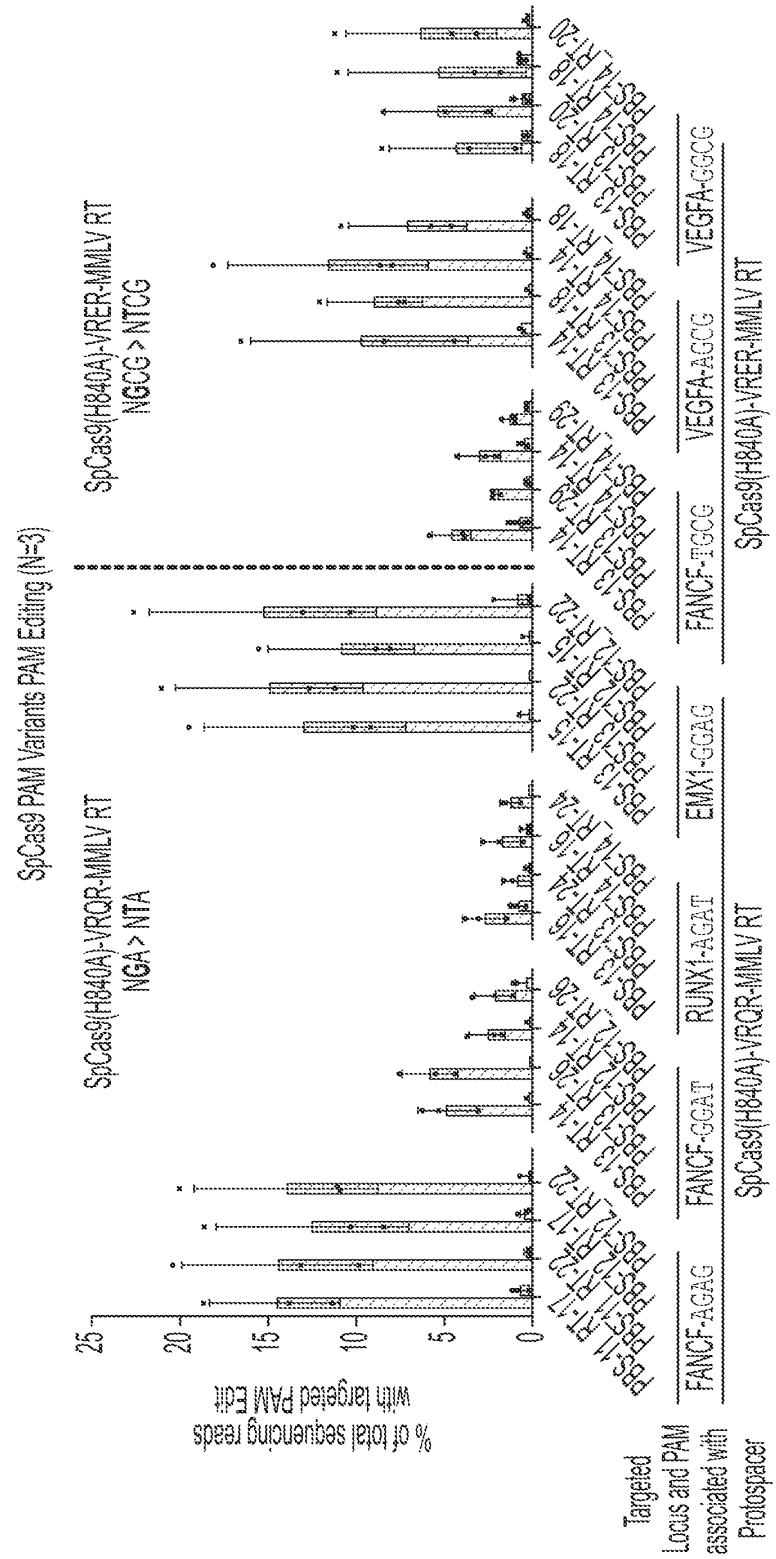

SpCas9 PAM Variants PAM Editing (N=3)
SpCas9(H840A)-VRQR-MMLV RT
NGA > NTA
SpCas9(H840A)-VRER-MMLV RT
NGCG > NTCG
% of total sequencing reads with targeted PAM Edit
25
20
15
10
5
0
Targeted Locus and PAM associated with Protospacer
FANCF-AGAG
FANCF-GGAT
RUNX1-AGAT
EMX1-GGAG
FANCF-TGCG
VEGFA-AGCG
VEGFA-GGCG
SpCas9(H840A)-VRQR-MMLV RT
SpCas9(H840A)-VRER-MMLV RT
pegRNA Primer Binding Site (PBS) Length and RT Template (RT) Length and Prime Editing System Used
Percent G>T Editing From Total Reads
Indels from total reads

FIG. 63

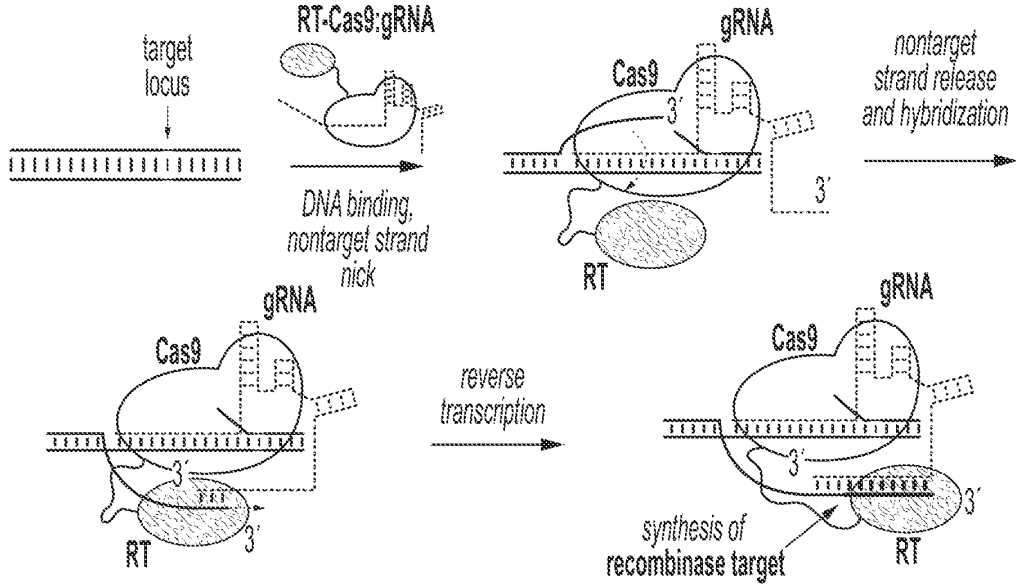


FIG. 64A

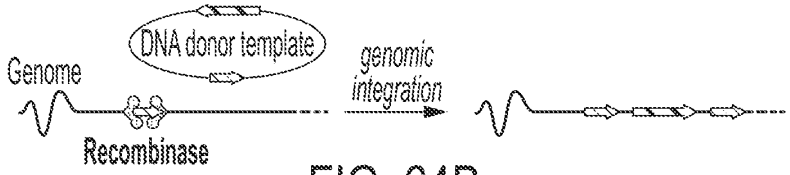


FIG. 64B



FIG. 64C



FIG. 64D

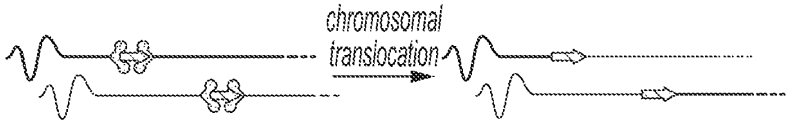


FIG. 64E

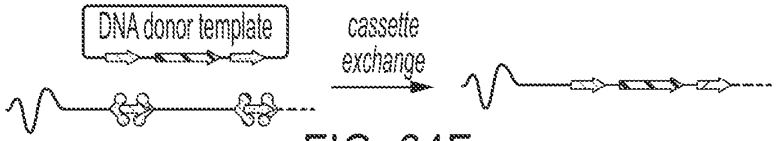


FIG. 64F

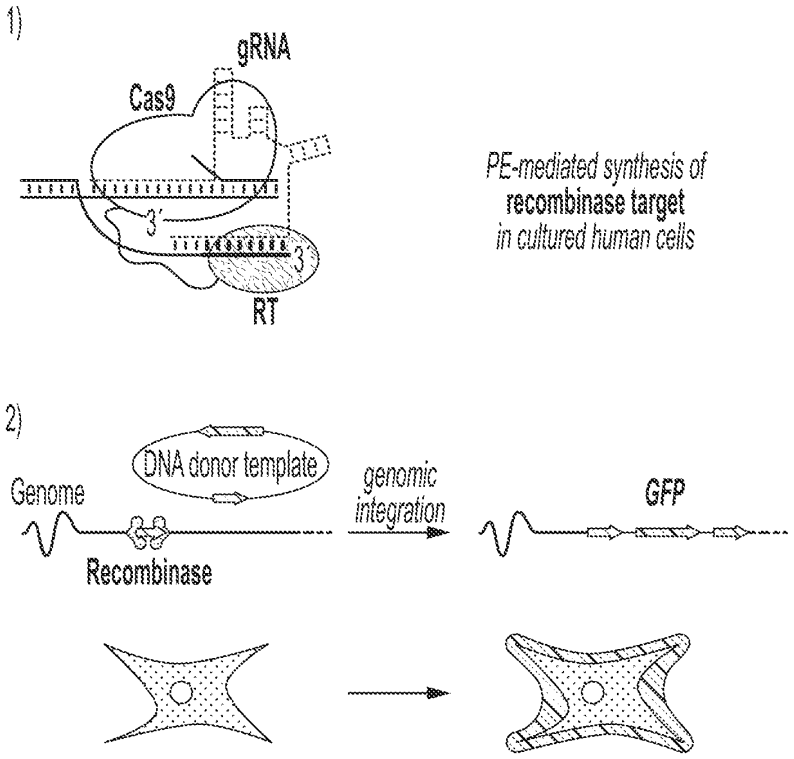

1)
gRNA
Cas9
3'
3
RT
PE-mediated synthesis of
recombinase target
in cultured human cells
2)
DNA donor template
Genome
Recombinase
genomic
integration
GFP

FIG. 65

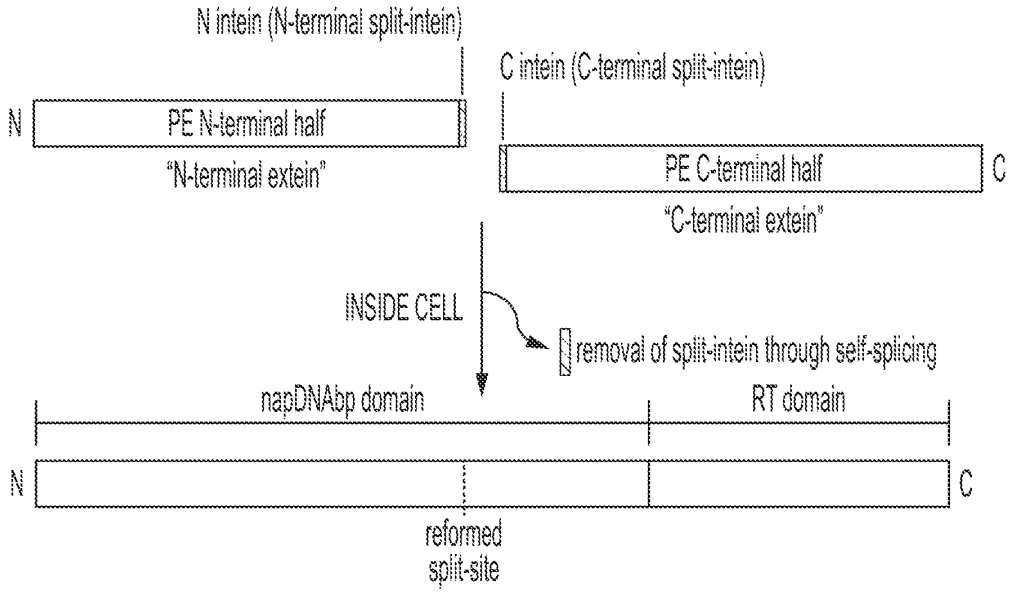

N intein (N-terminal split-intein)
C intein (C-terminal split-intein)
N
PE N-terminal half
"N-terminal extein"
PE C-terminal half
C
"C-terminal extein"
INSIDE CELL
removal of split-intein through self-splicing
napDNAbp domain
RT domain
N
C
reformed
split-site
Prime Editor

FIG. 66

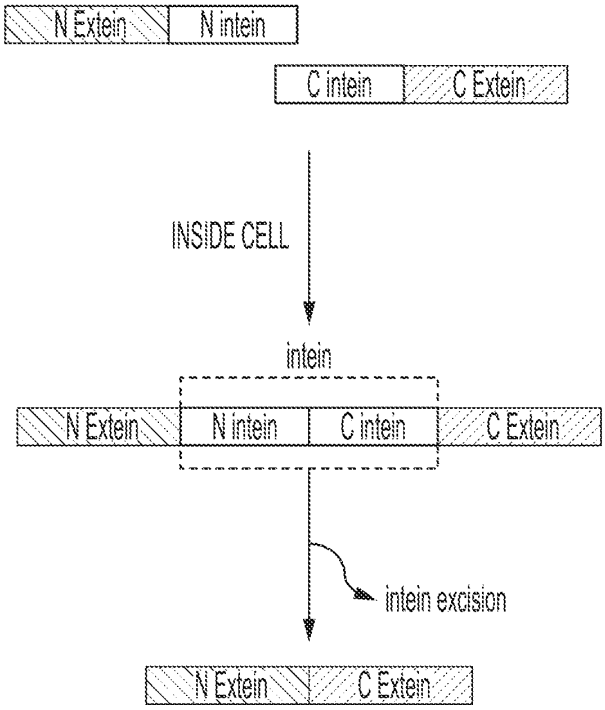

N Extein
N intein
C intein
C Extein
INSIDE CELL
intein
N Extein
N intein
C intein
C Extein
intein excision
N Extein
C Extein

FIG. 67A

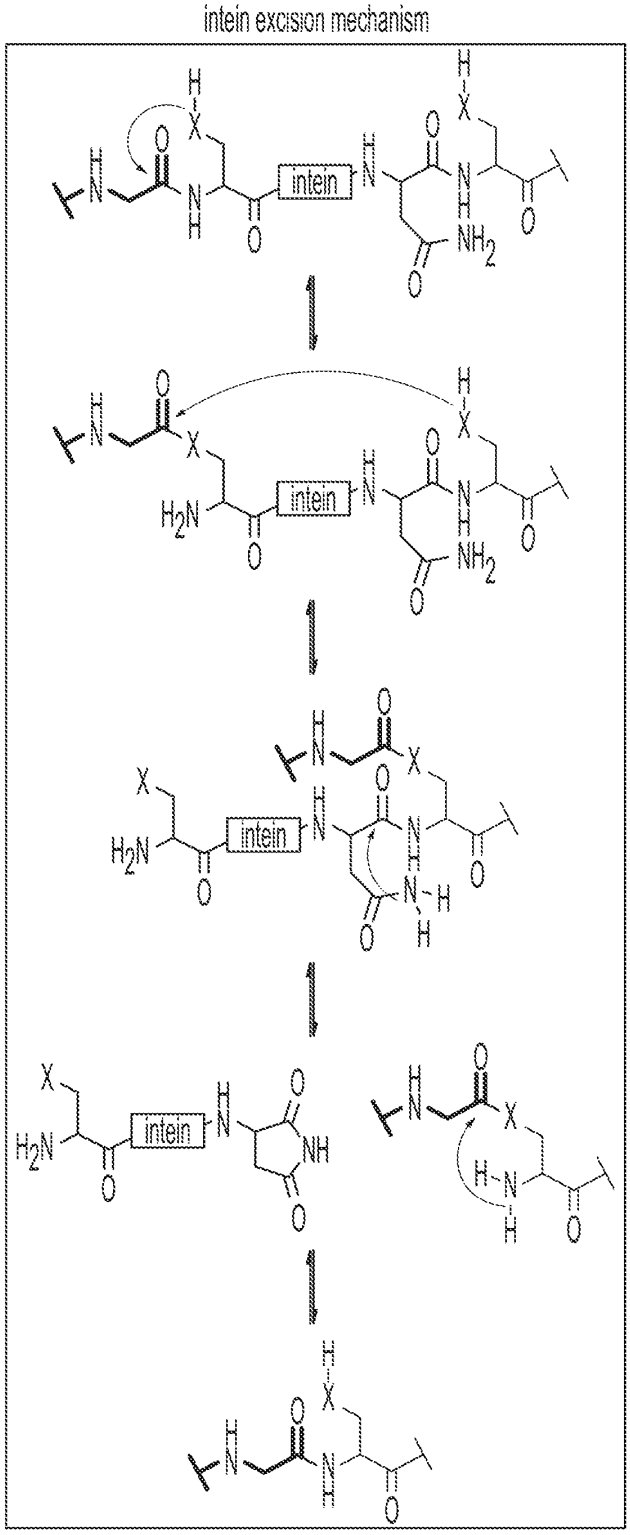


FIG. 67B

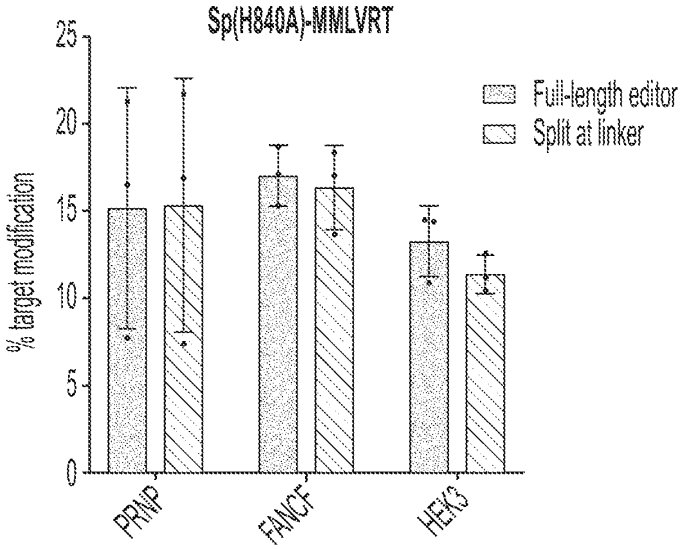


FIG. 68A

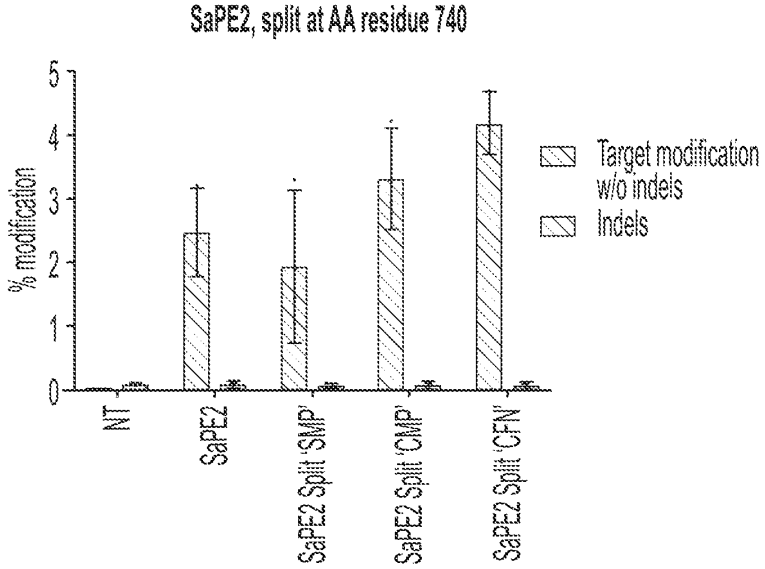


FIG. 68B

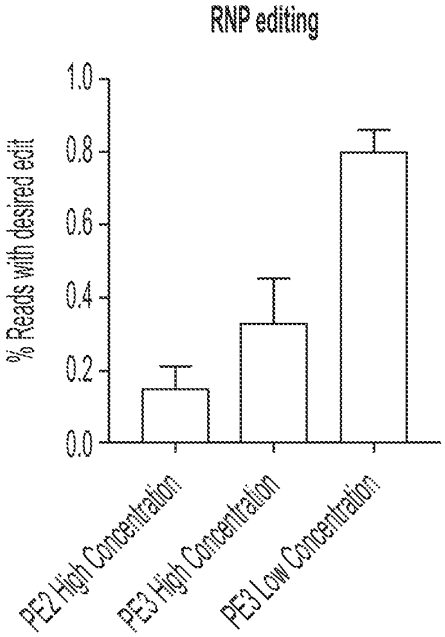

RNP editing
% Reads with desired edit
1.0
0.8
0.6
0.4
0.2
0.0
PE2 High Concentration
PE3 High Concentration
PE3 Low Concentration

FIG. 68C

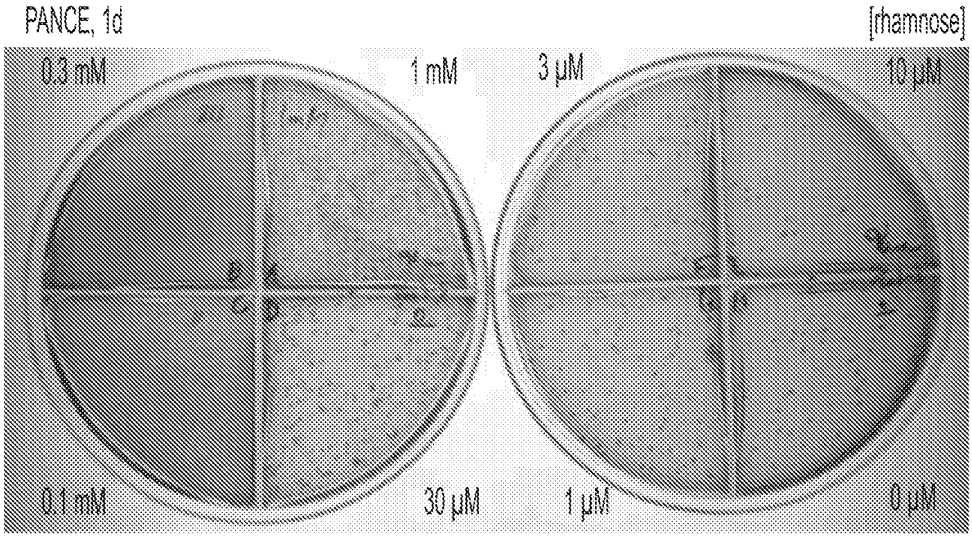


FIG. 69

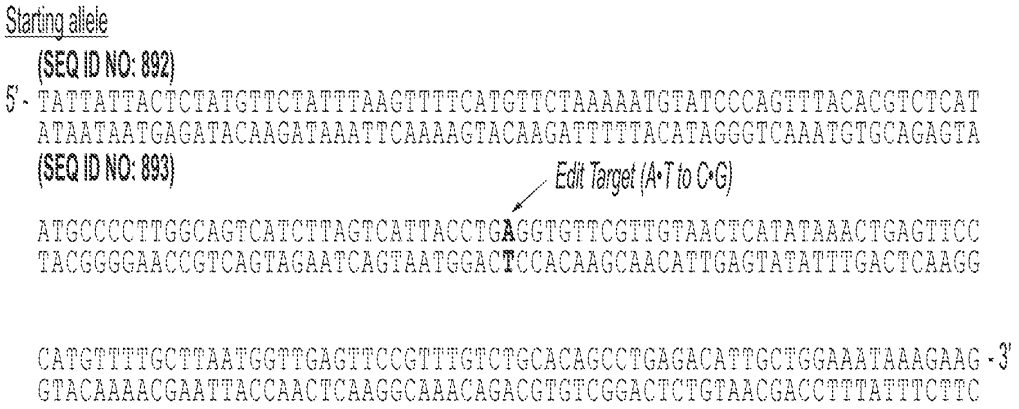


FIG. 70A

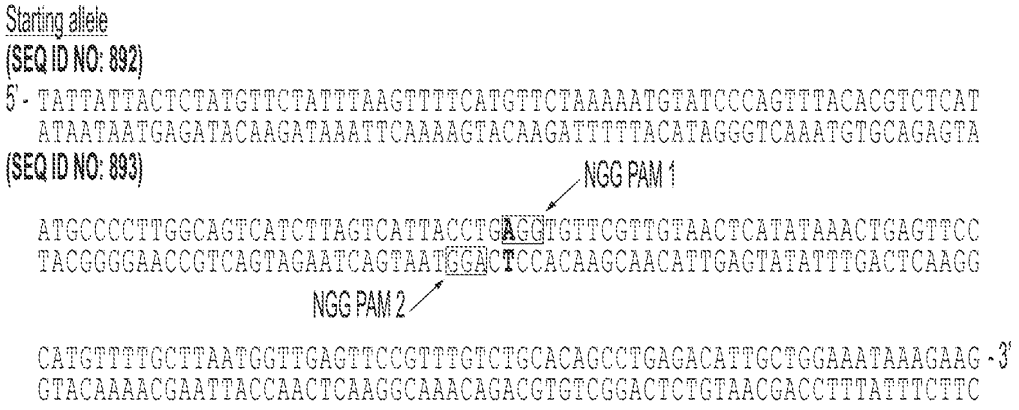


FIG. 70B

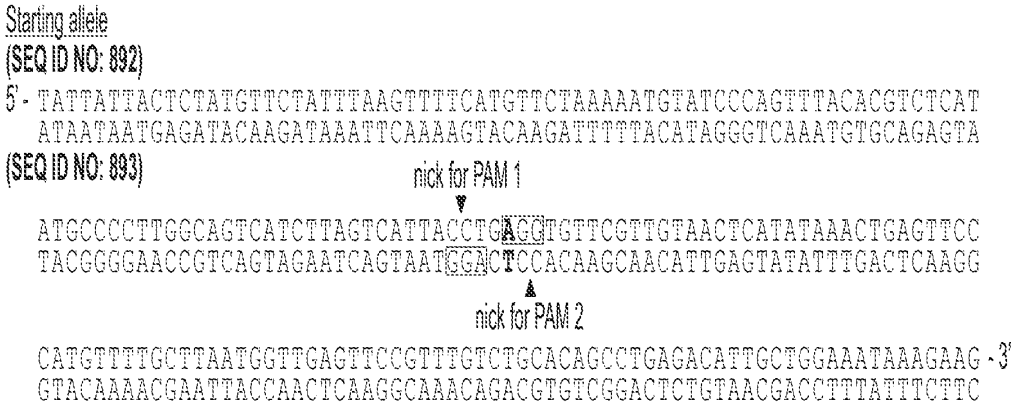


FIG. 70C



FIG. 70D



FIG. 70E

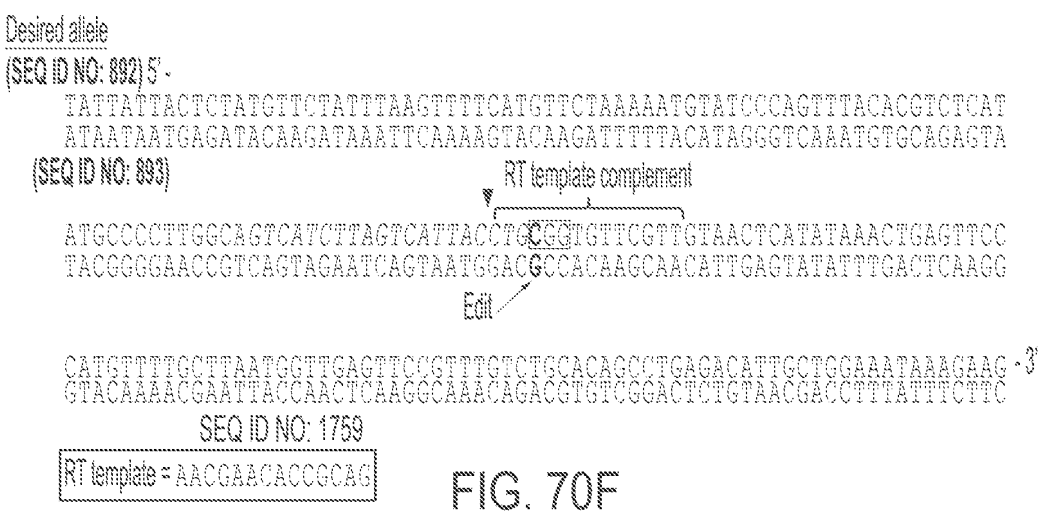


FIG. 70F

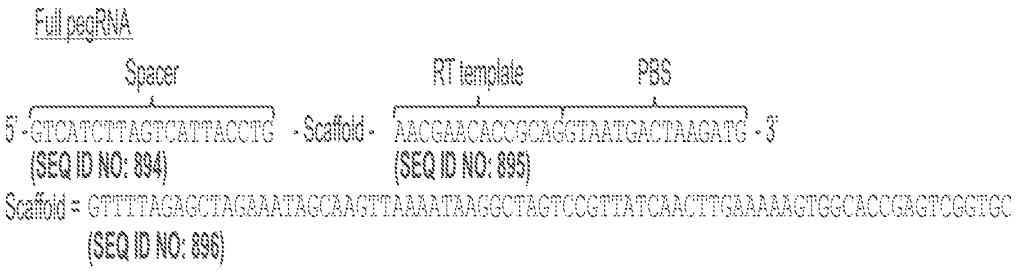


FIG. 70G

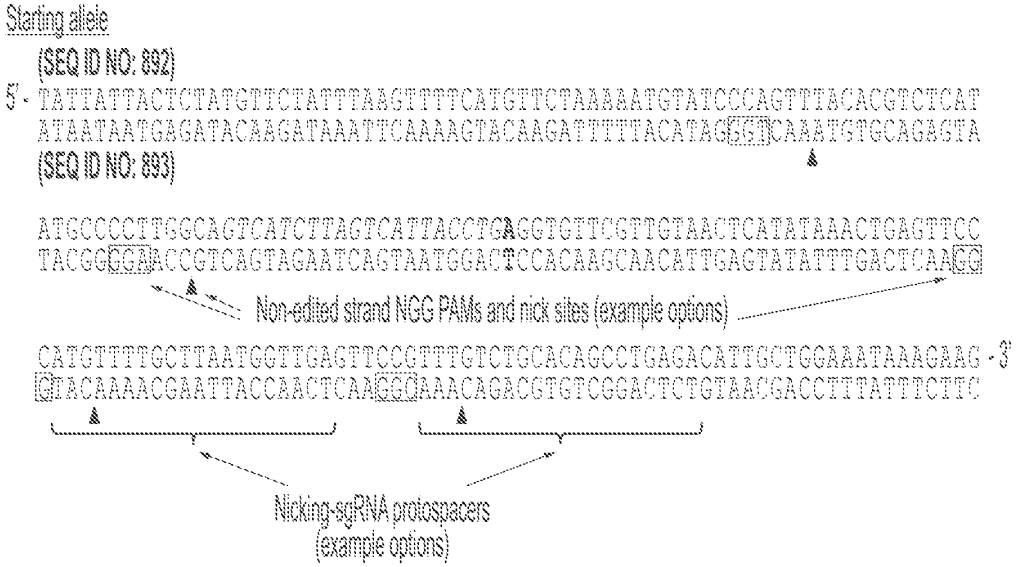


FIG. 70H



FIG. 70I

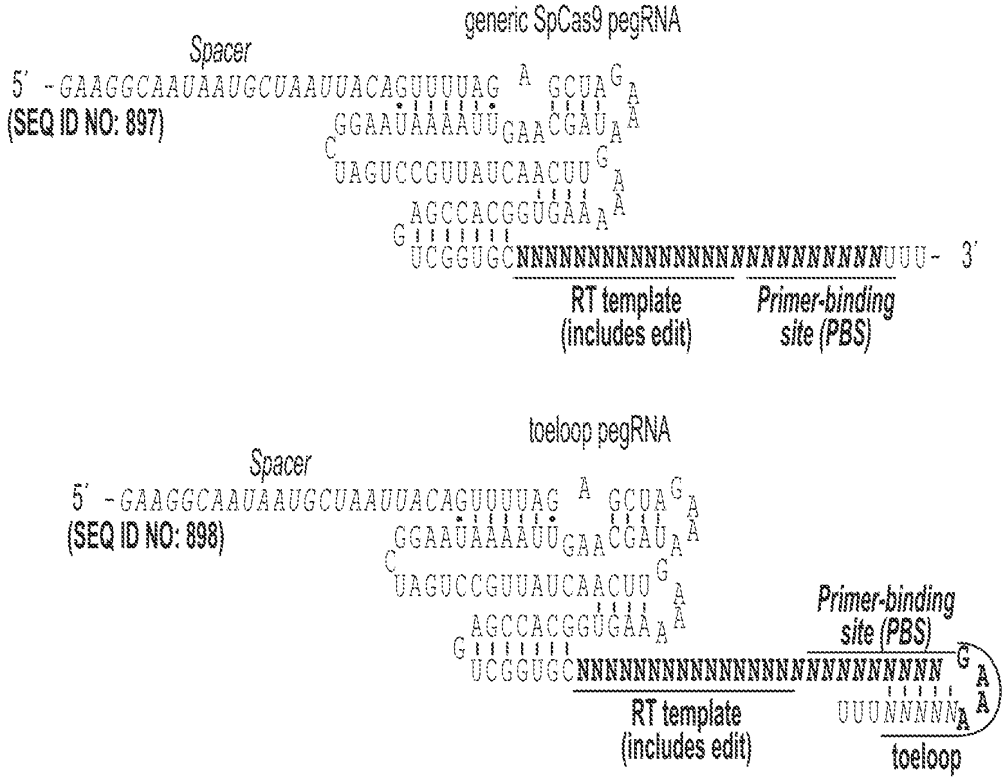


FIG. 71A

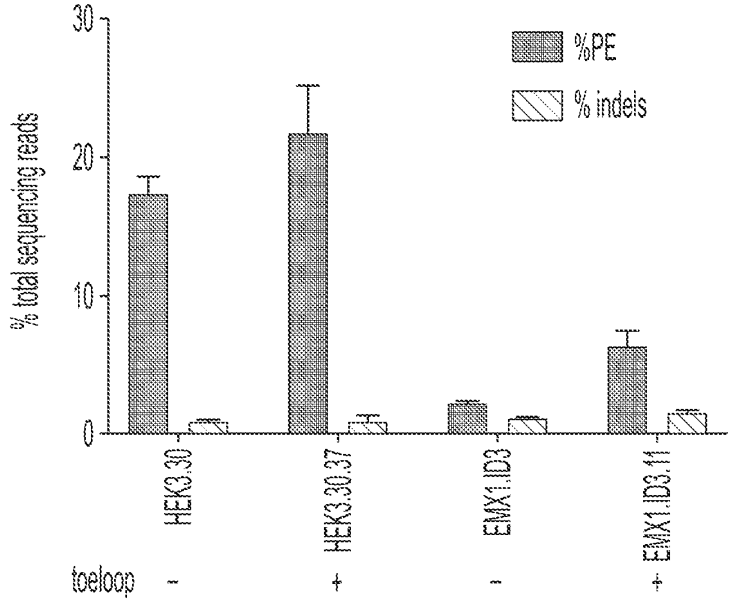


FIG. 71B

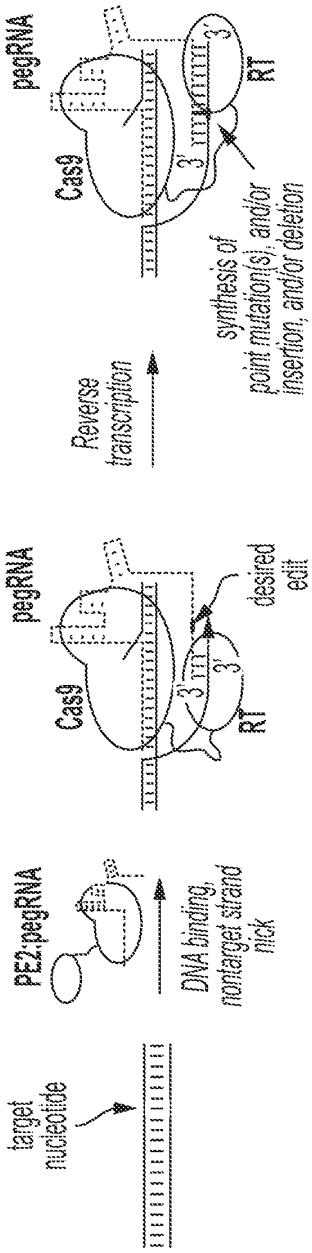


FIG. 72A

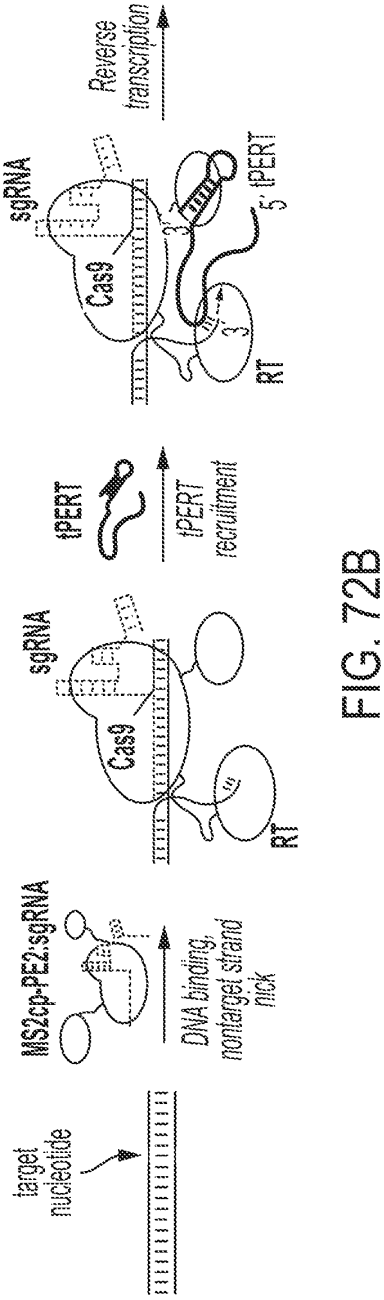


FIG. 72B

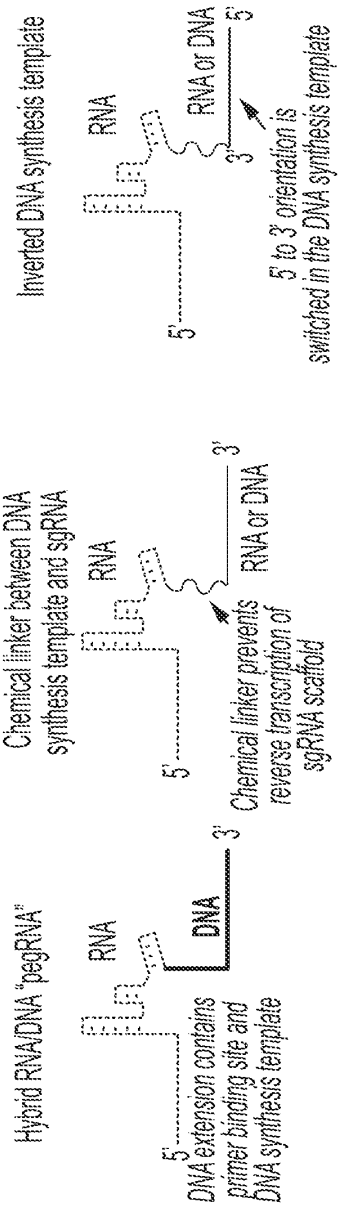

Hybrid RNA/DNA "pegRNA"
RNA
DNA
5'
3'
DNA extension contains primer binding site and DNA synthesis template
Chemical linker between DNA synthesis template and sgRNA
RNA
5'
Chemical linker prevents reverse transcription of sgRNA scaffold
RNA or DNA
3'
DNA-dependent DNA polymerase can be used in place of reverse transcriptase
Inverted DNA synthesis template
RNA
5'
RNA or DNA
3'
5'
5' to 3' orientation is switched in the DNA synthesis template

FIG. 72C

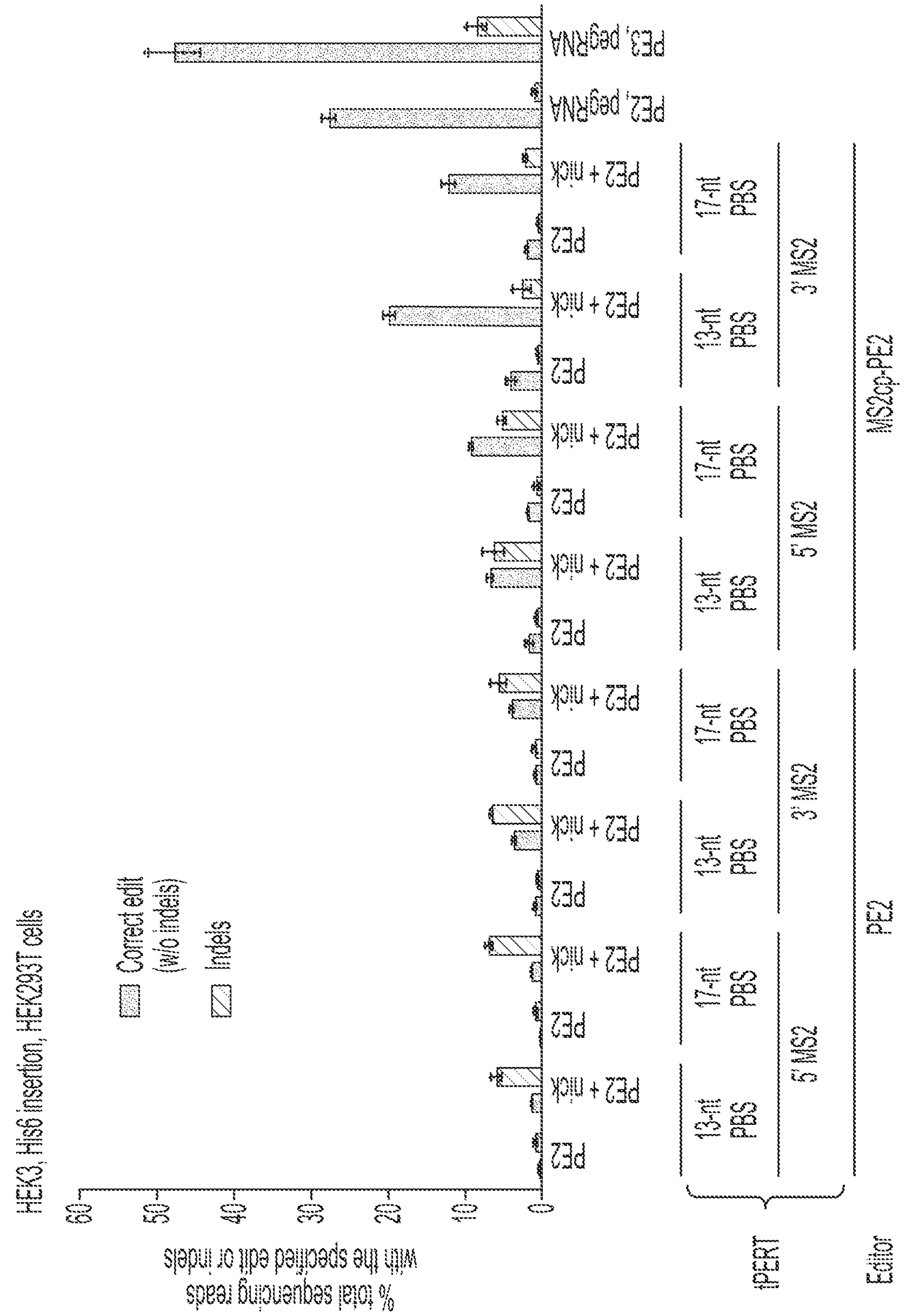

HEK3, His6 insertion, HEK293T cells
Correct edit (w/o indels)
Indels
% total sequencing reads with the specified edit or indels
60
50
40
30
20
10
0
PE2
PE2 + nick
PE2, pegRNA
PE3, pegRNA
13-nt PBS
17-nt PBS
5' MS2
3' MS2
tPERT
PE2
MS2cp-PE2
Editor

FIG. 73

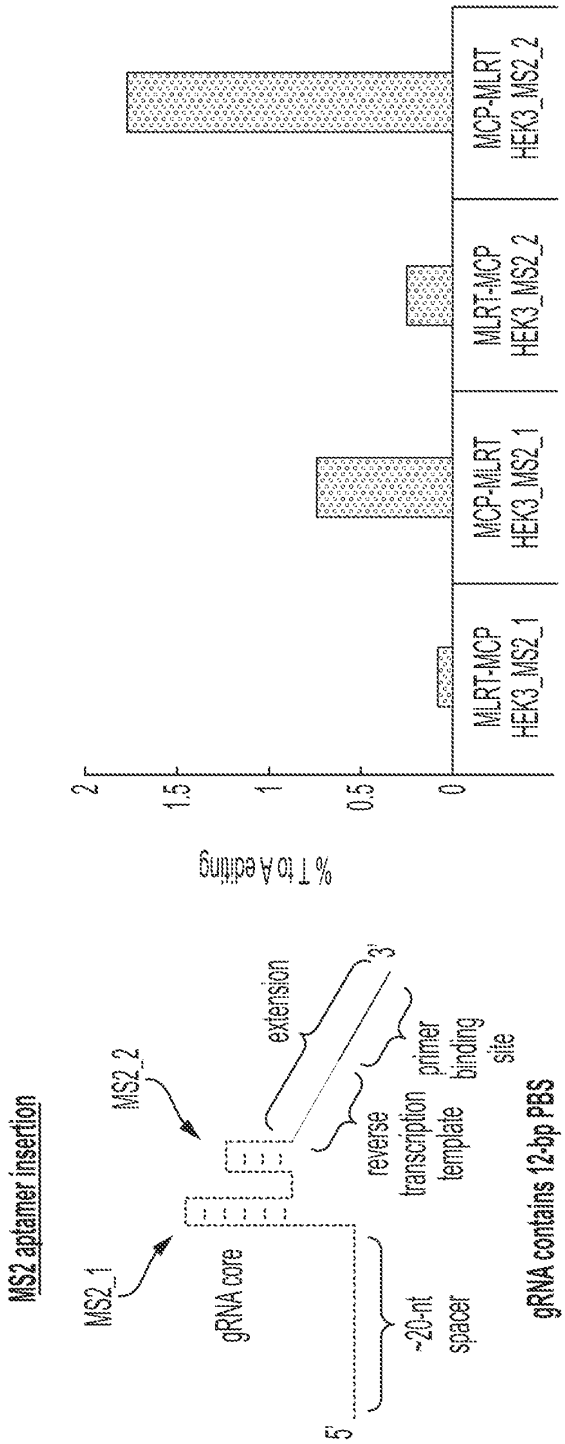

% T to A editing
2
1.5
1
0.5
0
MLRT-MCP HEK3_MS2_1
MCP-MLRT HEK3_MS2_1
MLRT-MCP HEK3_MS2_2
MCP-MLRT HEK3_MS2_2
MS2 aptamer insertion
MS2_1
MS2_2
gRNA core
~20-nt spacer
5'
3'
extension
reverse transcription template
primer binding site
gRNA contains 12-bp PBS

FIG. 74

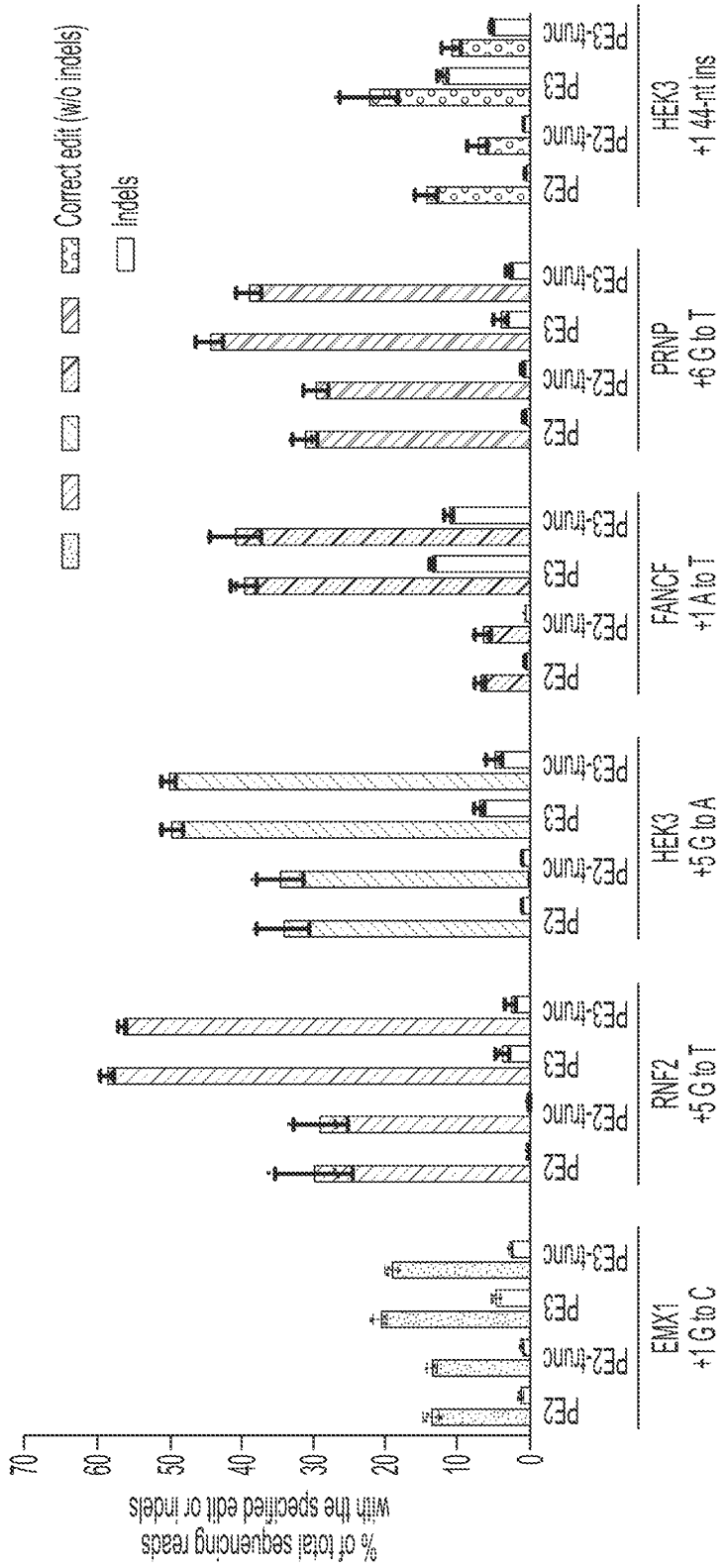

% of total sequencing reads with the specified edit or indels
0
10
20
30
40
50
60
70
Indels
Correct edit (w/o indels)
EMX1 +1 G to C
RNF2 +5 G to T
HEK3 +5 G to A
FANCF +1 A to T
PRNP +6 G to T
HEK3 +144-nt ins
PE2
PE2-trunc
PE3
PE3-trunc

FIG. 75

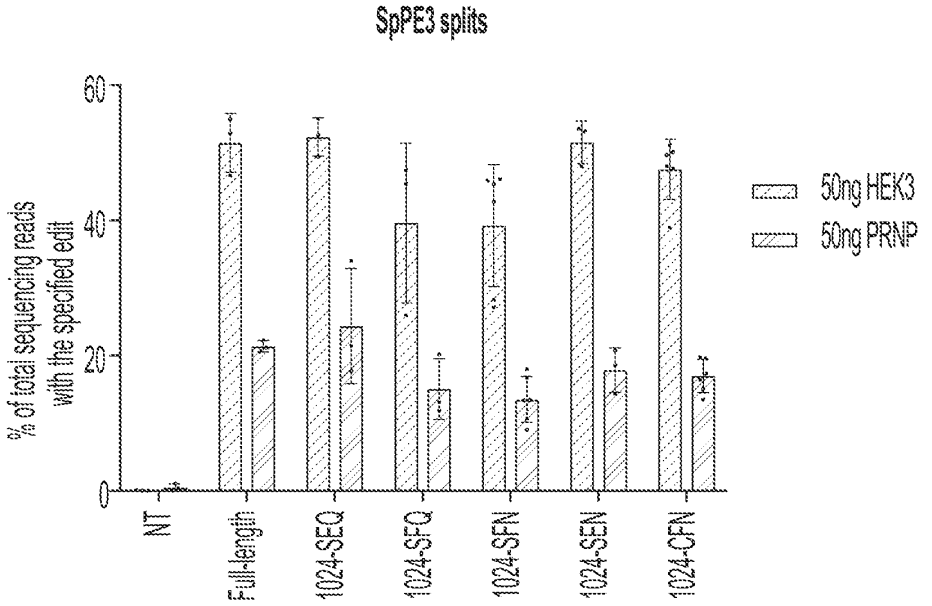

SpPE3 splits
% of total sequencing reads with the specified edit
60
40
20
0
NT
Full-length
1024-SEQ
1024-SFQ
1024-SFN
1024-SEN
1024-CFN
50ng HEK3
50ng PRNP

FIG. 76

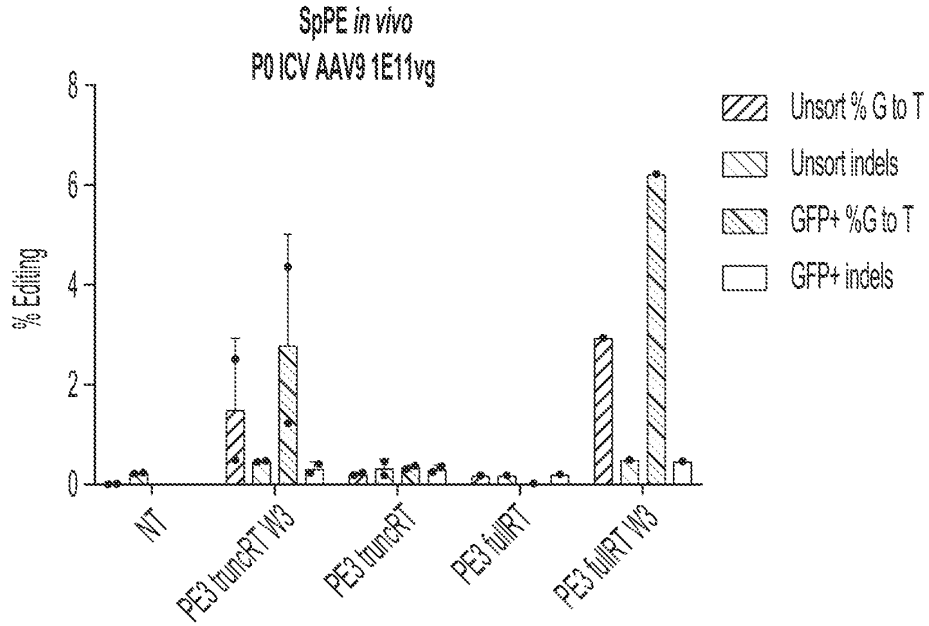

SpPE in vivo
P0 ICV AAV9 1E11vg
% Editing
8
6
4
2
0
NT
PE3 truncRT W3
PE3 truncRT
PE3 fullRT
PE3 fullRT W3
Unsort % G to T
Unsort indels
GFP+ %G to T
GFP+ indels

FIG. 77

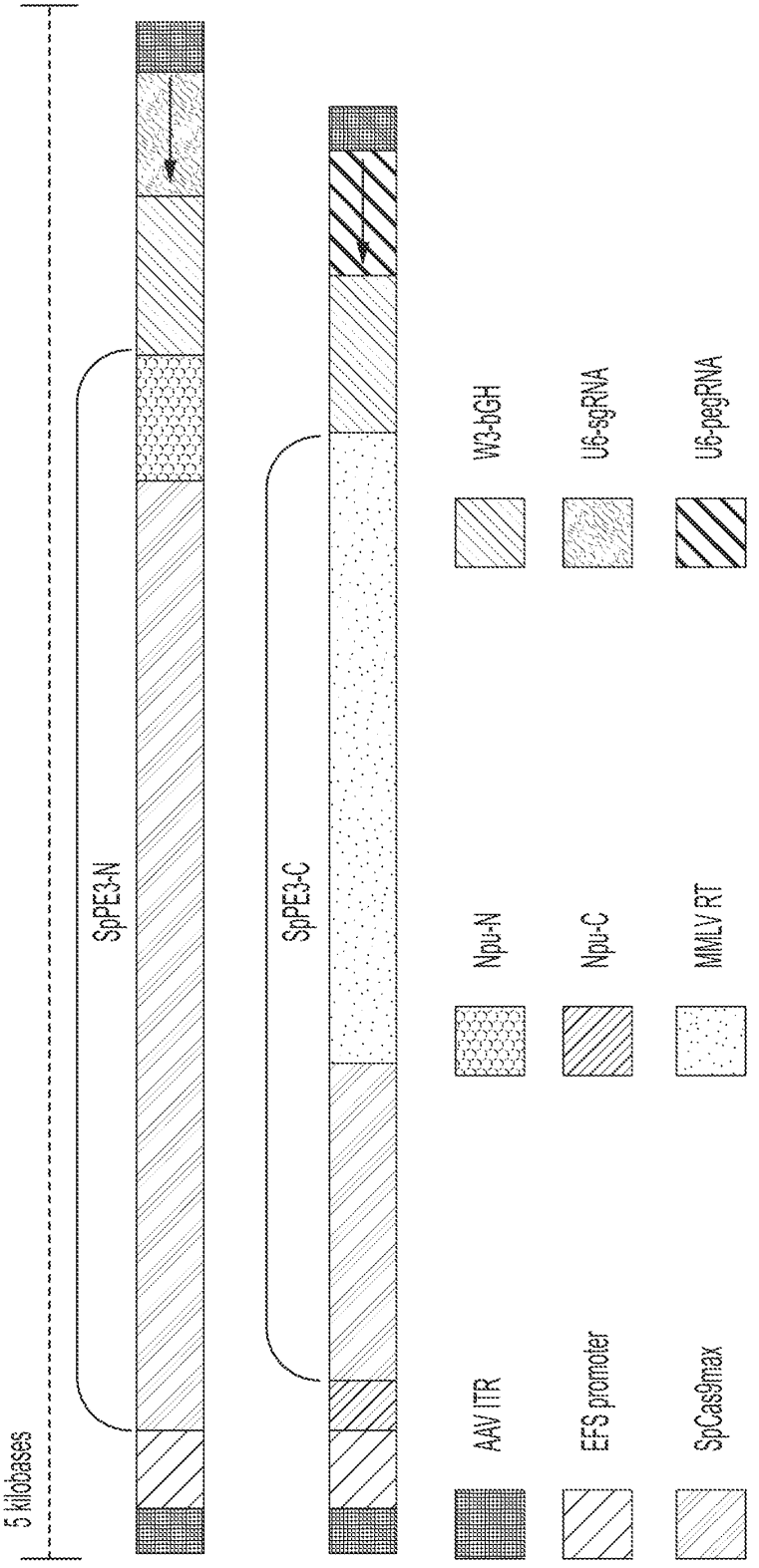

5 kilobases
SpPE3-N
SpPE3-C
AAV ITR
EFS promoter
SpCas9max
Npu-N
Npu-C
MMLV RT
W3-bGH
U6-sgRNA
U6-pegRNA

FIG. 78

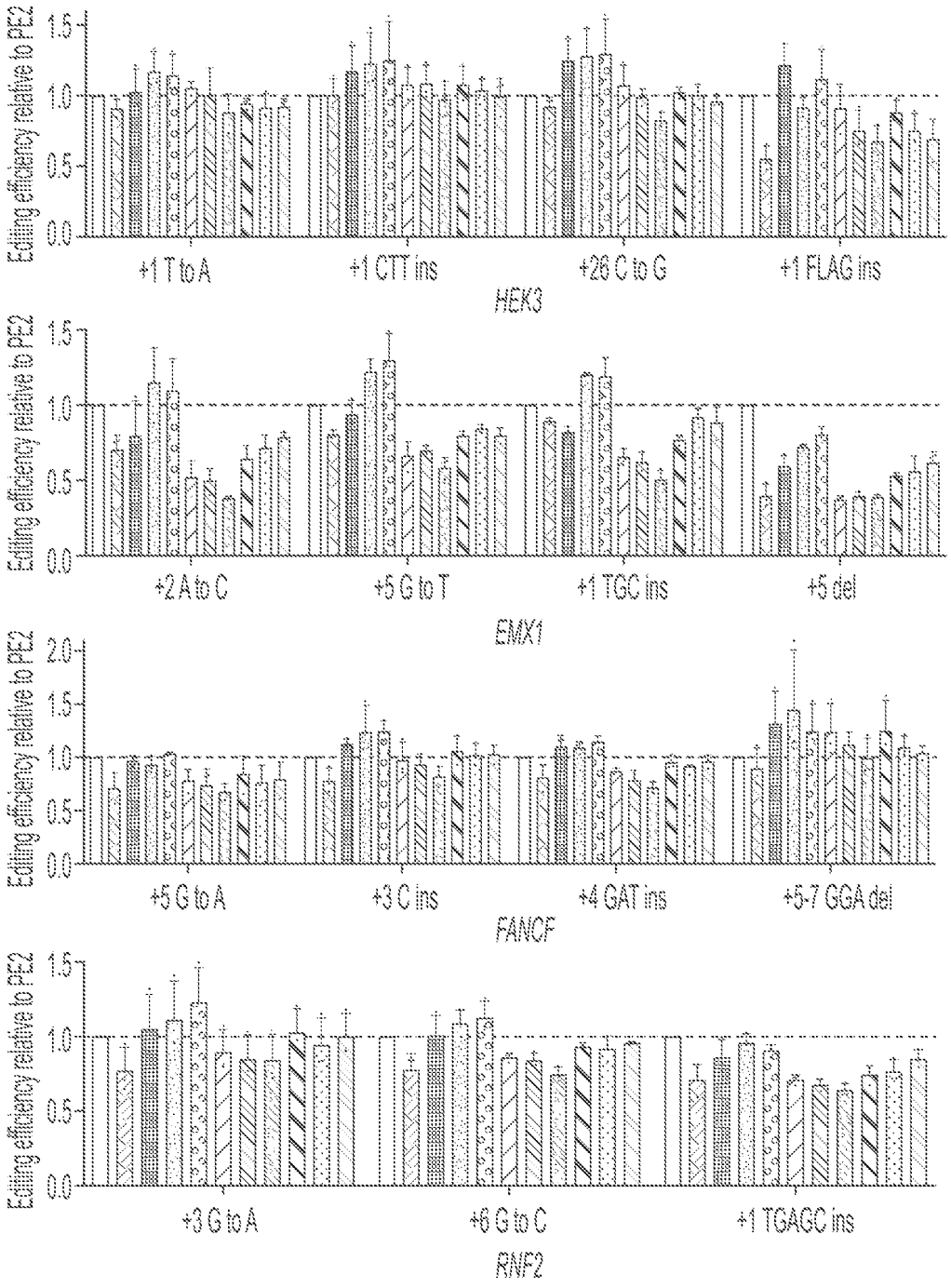

Editing efficiency relative to PE2
0.0
0.5
1.0
1.5
+1 T to A
+1 CTT ins
+26 C to G
+1 FLAG ins
HEK3
Editing efficiency relative to PE2
0.0
0.5
1.0
1.5
+2 A to C
+5 G to T
+1 TGC ins
+5 del
EMX1
Editing efficiency relative to PE2
0.0
0.5
1.0
1.5
2.0
+5 G to A
+3 C ins
+4 GAT ins
+5-7 GGA del
FANCF
Editing efficiency relative to PE2
0.0
0.5
1.0
1.5
+3 G to A
+6 G to C
+1 TGAGC ins
RNF2
PE2
1x SGGS (SEQ ID NO: 8)
2x SGGS (SEQ ID NO: 9)
3x SGGS (SEQ ID NO: 10)
1x XTEN (SEQ ID NO: 11)
No linker
1x Gly
1x Pro
1x EAAAK (SEQ ID NO: 12)
2x EAAAK (SEQ ID NO: 13)
3x EAAAK (SEQ ID NO: 14)

FIG. 79

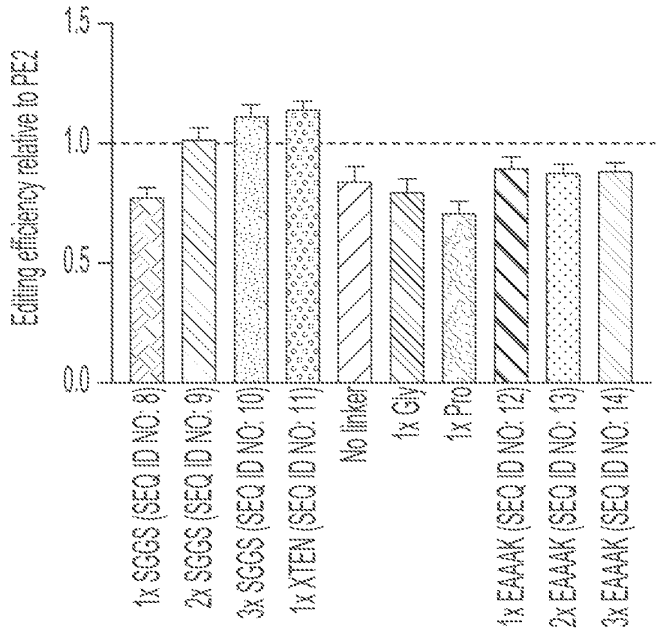

Editing efficiency relative to PE2
1.5
1.0
0.5
0.0
1x SGGS (SEQ ID NO: 8)
2x SGGS (SEQ ID NO: 9)
3x SGGS (SEQ ID NO: 10)
1x XTEN (SEQ ID NO: 11)
No linker
1x Gly
1x Pro
1x EAAK (SEQ ID NO: 12)
2x EAAK (SEQ ID NO: 13)
3x EAAK (SEQ ID NO: 14)

FIG. 80

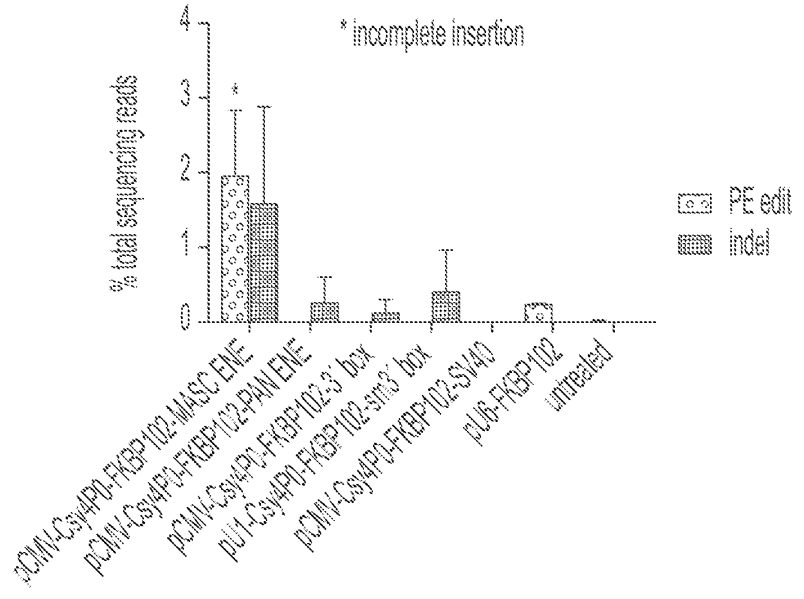

* incomplete insertion
% total sequencing reads
4
3
2
1
0
*
PE edit
indel
pCMV-Csy4P0-FKBP102-MASC ENE
pCMV-Csy4P0-FKBP102-PAN ENE
pCMV-Csy4P0-FKBP102-3' box
pU1-Csy4P0-FKBP102-sm3' box
pCMV-Csy4P0-FKBP102-SV40
pU6-FKBP102
untreated

FIG. 81

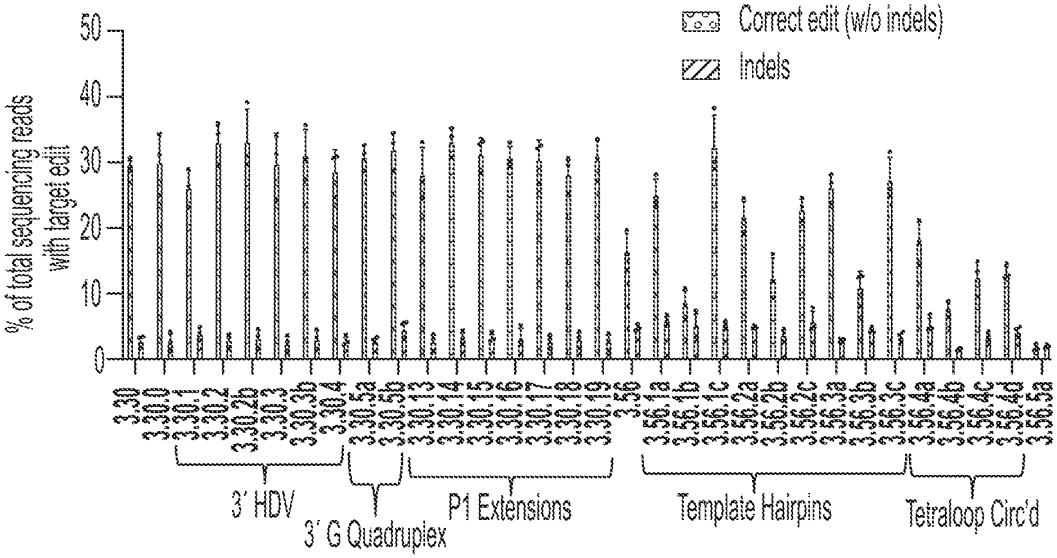

Correct edit (w/o indels)
Indels
% of total sequencing reads with target edit
50
40
30
20
10
0
3.30
3.30.0
3.30.1
3.30.2
3.30.2b
3.30.3
3.30.3b
3.30.4
3.30.5a
3.30.5b
3.30.13
3.30.14
3.30.15
3.30.16
3.30.17
3.30.18
3.30.19
3.56
3.56.1a
3.56.1b
3.56.1c
3.56.2a
3.56.2b
3.56.2c
3.56.3a
3.56.3b
3.56.3c
3.56.4a
3.56.4b
3.56.4c
3.56.4d
3.56.5a
3' HDV
3' G Quadruplex
P1 Extensions
Template Hairpins
Tetraloop Circ'd

FIG. 82

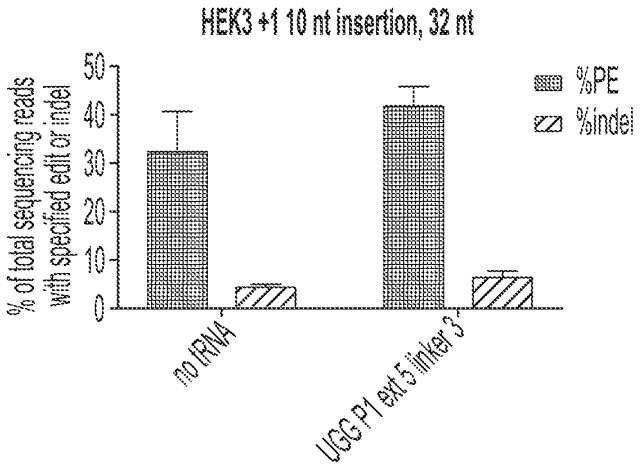

HEK3 +1 10 nt insertion, 32 nt
%PE
%indel
% of total sequencing reads with specified edit or indel
50
40
30
20
10
0
no tRNA
UGC P1 ext 5' linker 3'

FIG. 83

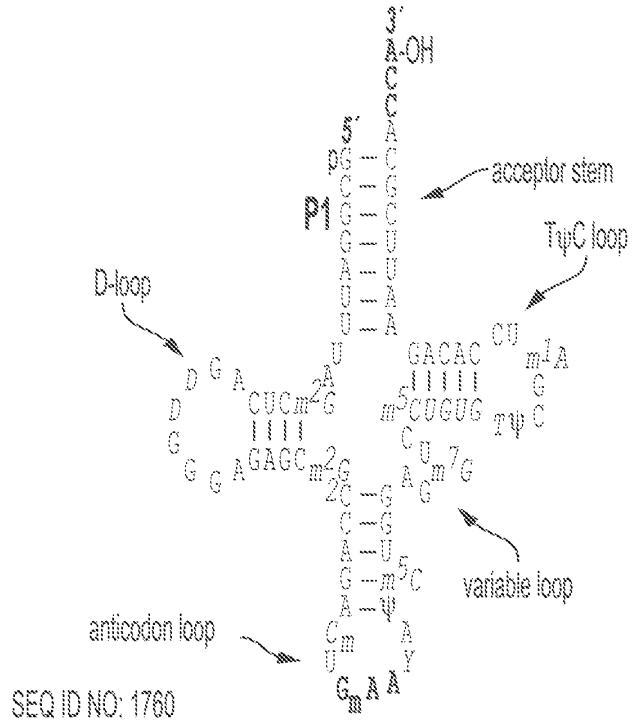


FIG. 84A

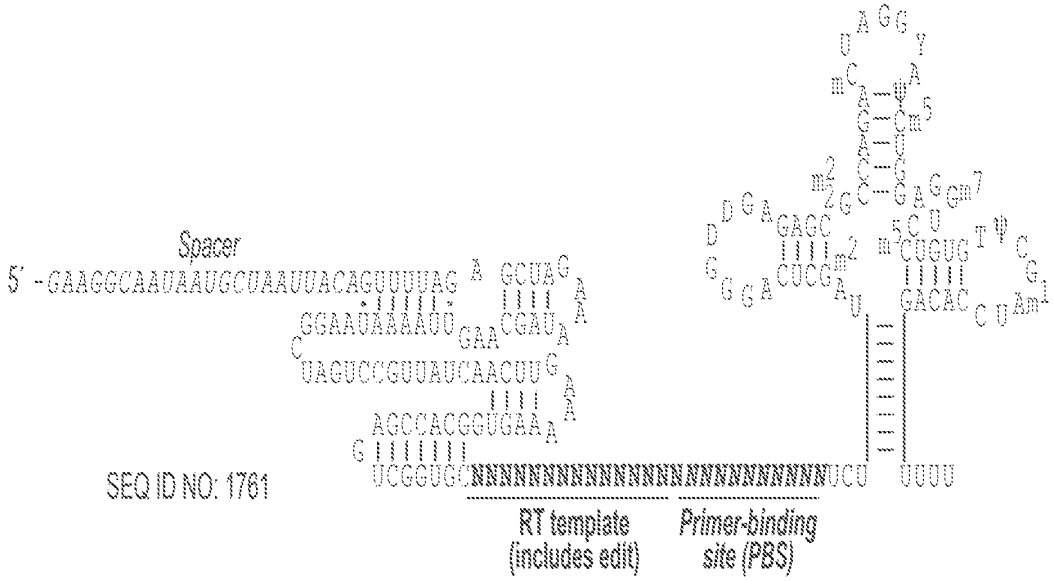


FIG. 84B

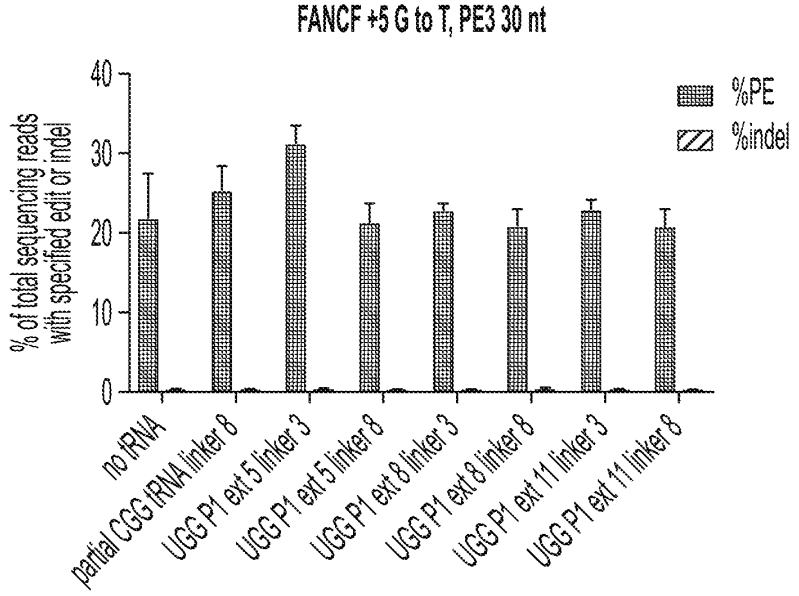

FANCF +5 G to T, PE3 30 nt
%PE
%indel
% of total sequencing reads with specified edit or indel
40
30
20
10
0
no tRNA
partial CGG tRNA linker 8
UGG P1 ext 5 linker 3
UGG P1 ext 5 linker 8
UGG P1 ext 8 linker 3
UGG P1 ext 8 linker 8
UGG P1 ext 11 linker 3
UGG P1 ext 11 linker 8

FIG. 85

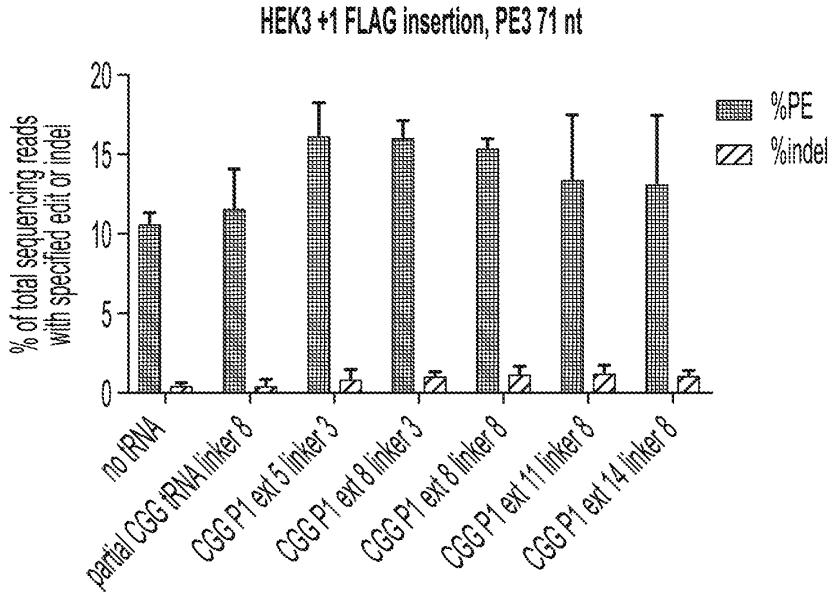

HEK3 +1 FLAG insertion, PE3 71 nt
%PE
%indel
% of total sequencing reads with specified edit or indel
20
15
10
5
0
no tRNA
partial CGG tRNA linker 8
CGG P1 ext 5 linker 3
CGG P1 ext 8 linker 3
CGG P1 ext 8 linker 8
CGG P1 ext 11 linker 8
CGG P1 ext 14 linker 8

FIG. 86

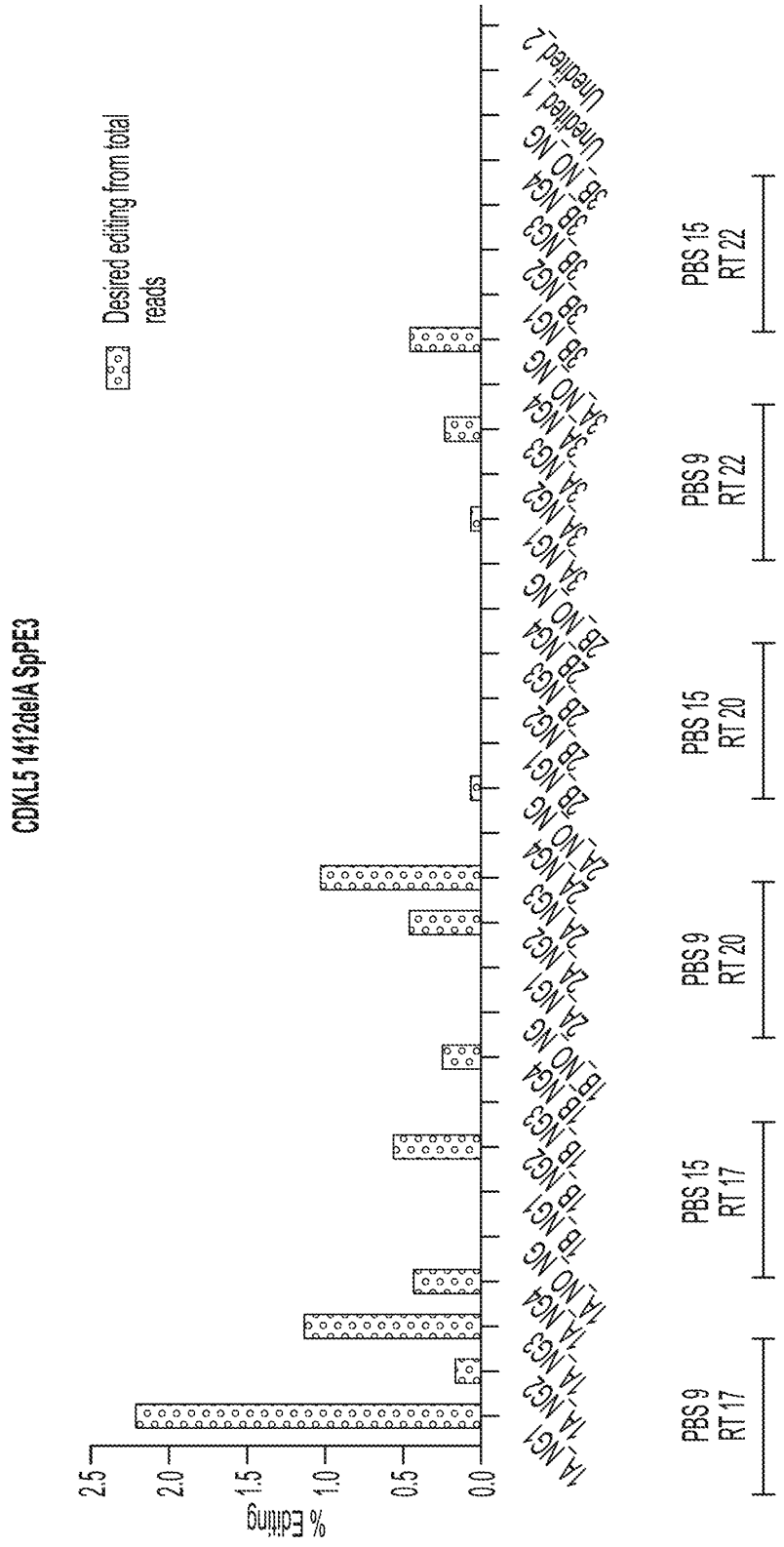

CDKL5 1412delA SpPE3
Desired editing from total reads
% Editing
2.5
2.0
1.5
1.0
0.5
0.0
14_NG1
14_NG2
14_NG3
14_NG4
14_NO_NG
18_NG1
18_NG2
18_NG3
18_NG4
18_NO_NG
24_NG1
24_NG2
24_NG3
24_NG4
24_NO_NG
28_NG1
28_NG2
28_NG3
28_NG4
28_NO_NG
34_NG1
34_NG2
34_NG3
34_NG4
34_NO_NG
38_NG1
38_NG2
38_NG3
38_NG4
38_NO_NG
Unedited_1
Unedited_2
PBS 9 RT 17
PBS 15 RT 17
PBS 9 RT 20
PBS 15 RT 20
PBS 9 RT 22
PBS 15 RT 22

FIG. 87

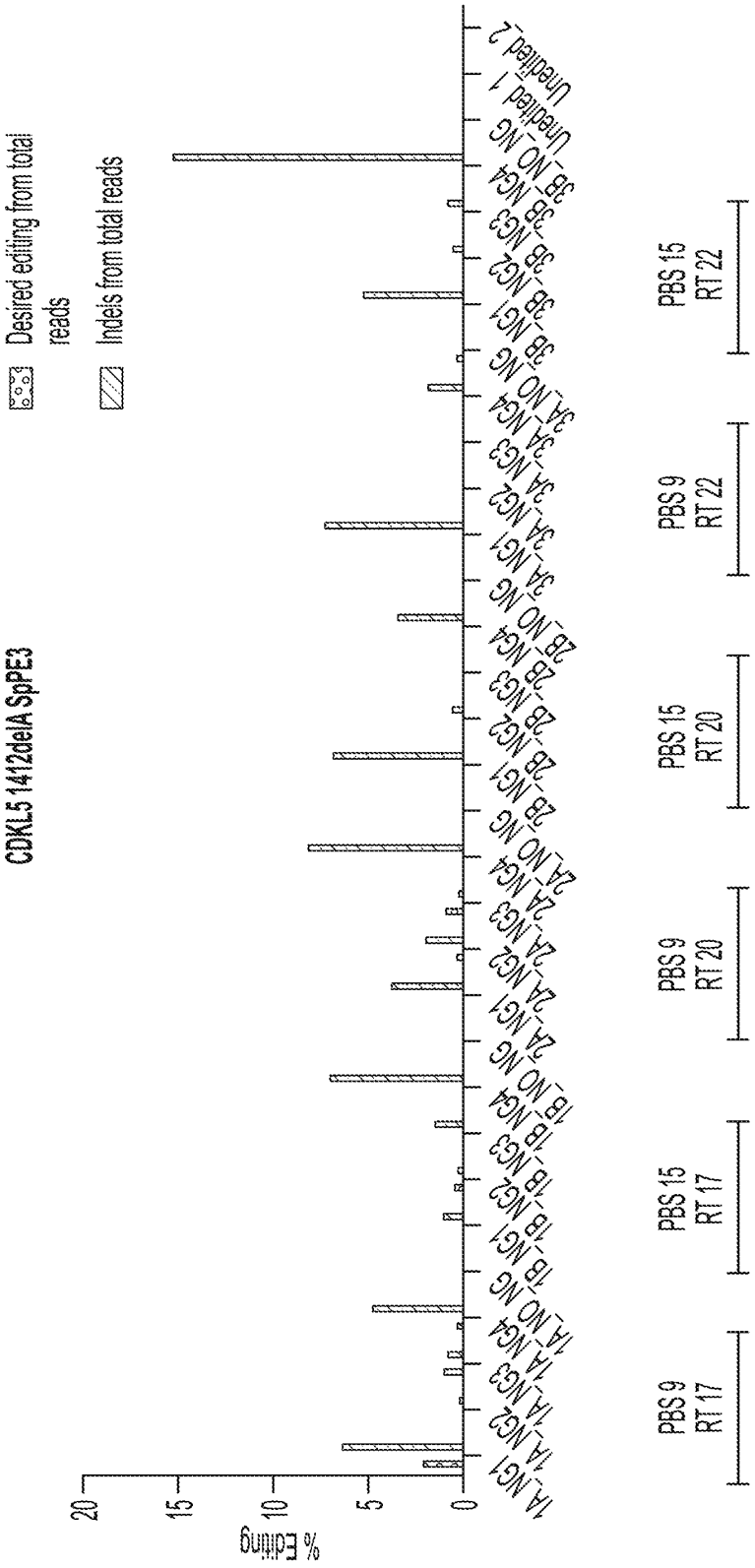

CDKL5 1412delA SpPE3
Desired editing from total reads
Indels from total reads
% Editing
0
5
10
15
20
1A_NG1
1A_1A_NG2
1A_1A_NG3
1A_1A_NG4
1A_NO_NG
1B_1B_NG1
1B_1B_NG2
1B_1B_NG3
1B_1B_NG4
1B_NO_NG
2A_2A_NG1
2A_2A_NG2
2A_2A_NG3
2A_2A_NG4
2A_NO_NG
2B_2B_NG1
2B_2B_NG2
2B_2B_NG3
2B_2B_NG4
2B_NO_NG
3A_3A_NG1
3A_3A_NG2
3A_3A_NG3
3A_3A_NG4
3A_NO_NG
3B_3B_NG1
3B_3B_NG2
3B_3B_NG3
3B_3B_NG4
3B_NO_NG
Unedited_1
Unedited_2
PBS 9
RT 17
PBS 15
RT 17
PBS 9
RT 20
PBS 15
RT 20
PBS 9
RT 22
PBS 15
RT 22

FIG. 88

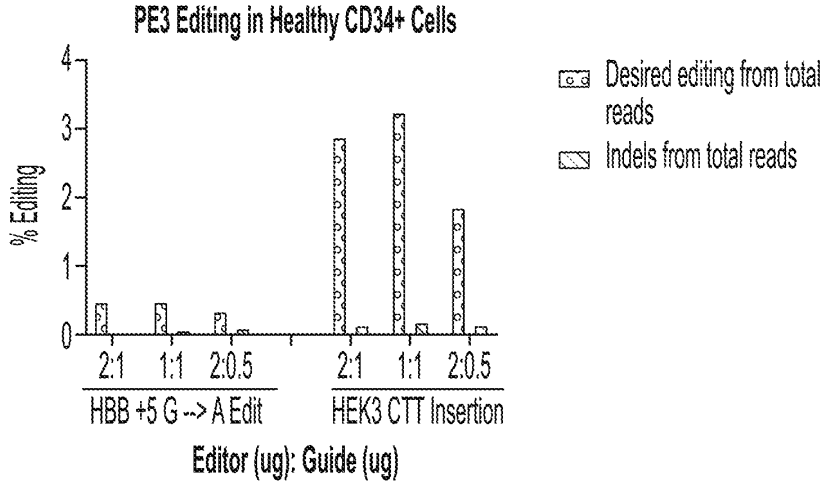

PE3 Editing in Healthy CD34+ Cells
% Editing
4
3
2
1
0
2:1
1:1
2:0.5
2:1
1:1
2:0.5
HBB +5 G --> A Edit
HEK3 CTT Insertion
Editor (ug): Guide (ug)
Desired editing from total reads
Indels from total reads

FIG. 89

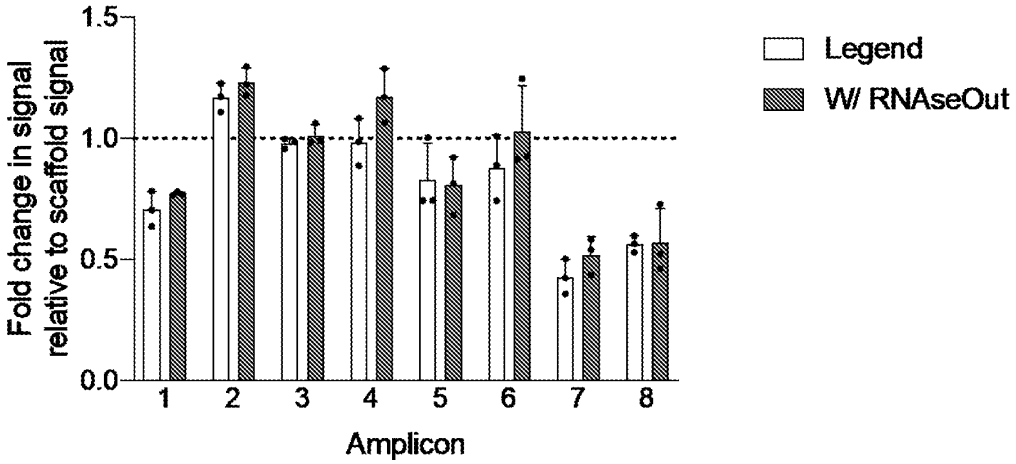

1.5
1.0
0.5
0.0
Fold change in signal relative to scaffold signal
1
2
3
4
5
6
7
8
Amplicon
Legend
W/ RNAseOut

FIG. 90A

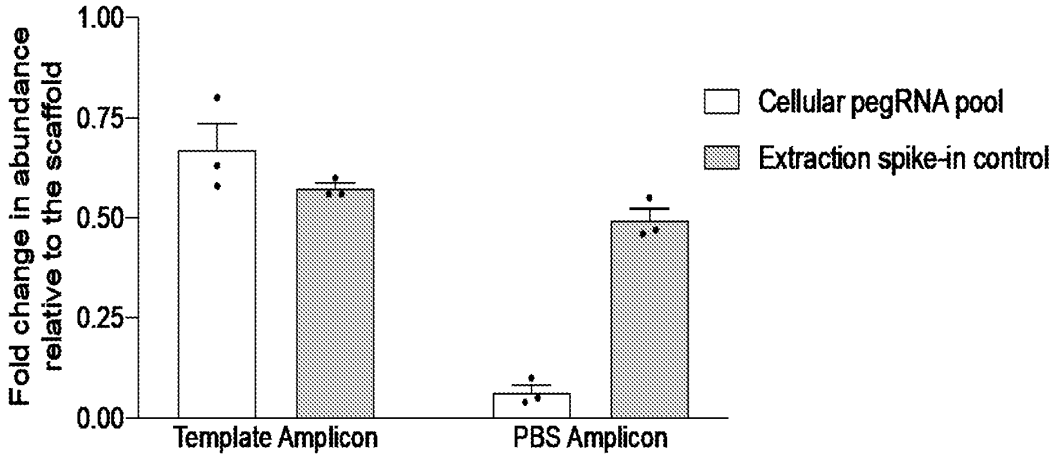

Fold change in abundance relative to the scaffold
1.00
0.75
0.50
0.25
0.00
Template Amplicon
PBS Amplicon
Cellular pegRNA pool
Extraction spike-in control

FIG. 90B

pegRNA sequence: SEQ ID NO: 1762
GGCCCAGACTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGA
AAAGTGGCACCGAGTCGGTGCTGGAGGAAGCAGGGCTTCCTTTCCTCTGCCATCACTTATCGTCGTCATCCTTGT
ATCCGTGCTCAGTCTG

Template amplicon sequence: SEQ ID NO: 1763
AAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTGGAGGAAGCAGGGCTTCCTTTC
TCTGCCATCACTTATCGTC

PBS amplicon sequence: SEQ ID NO: 1764
TGAAAAAGTGGCACCGAGTCGGTGCTGGAGGAAGCAGGGCTTCCTTTCCTCTGCCATCACTTATCGTCGTCATCC
TGTAATCCGTGCTCAGTCTGT

FIG. 90C

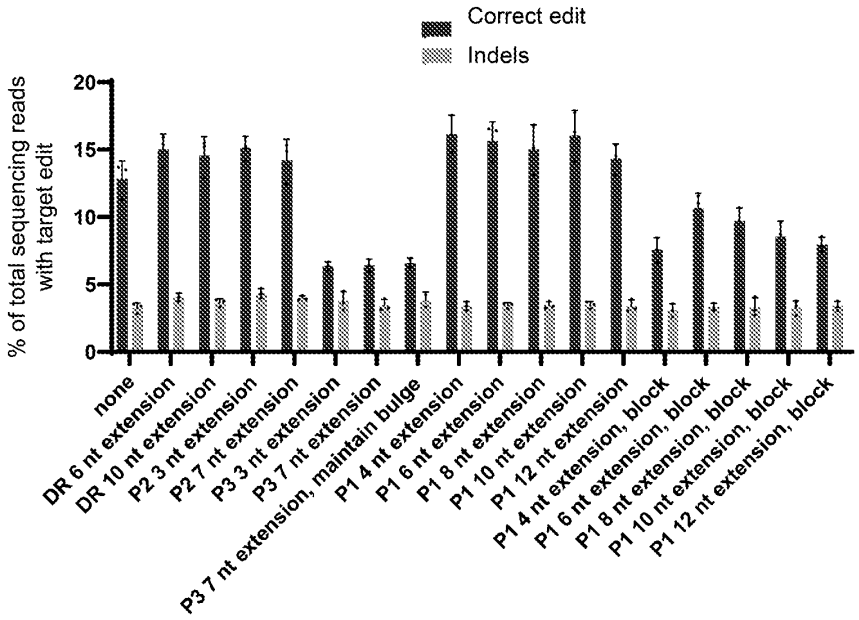


FIG. 91A

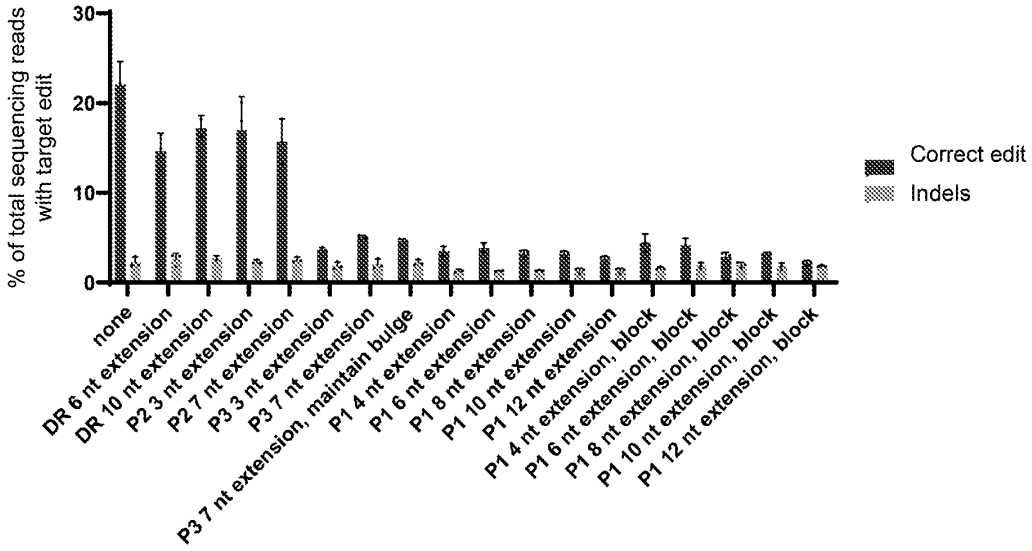


FIG. 91B

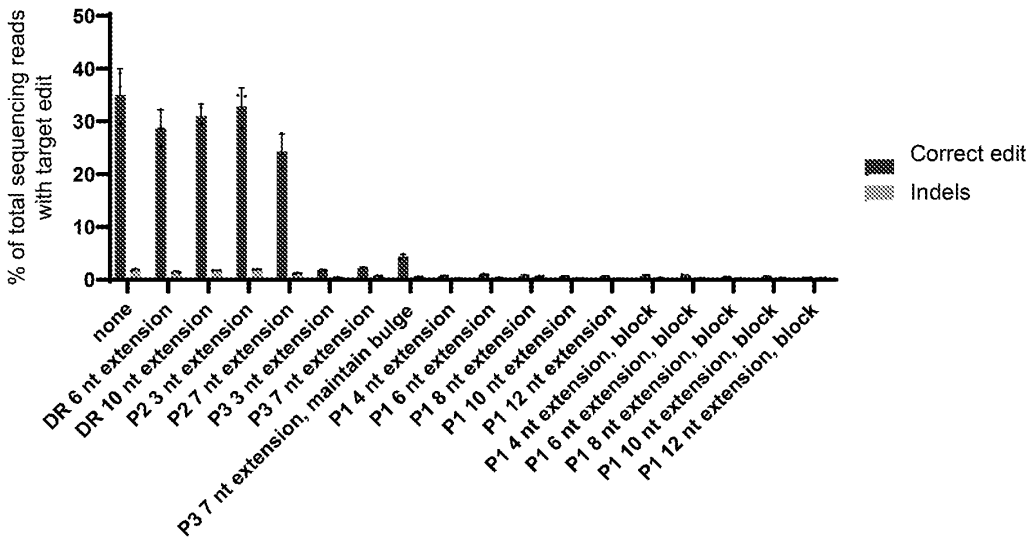


FIG. 91C

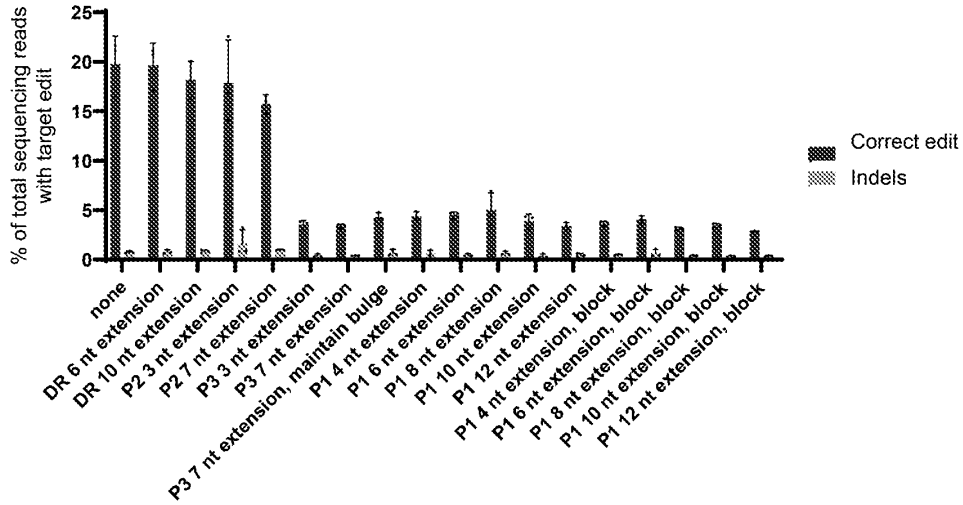


FIG. 91D

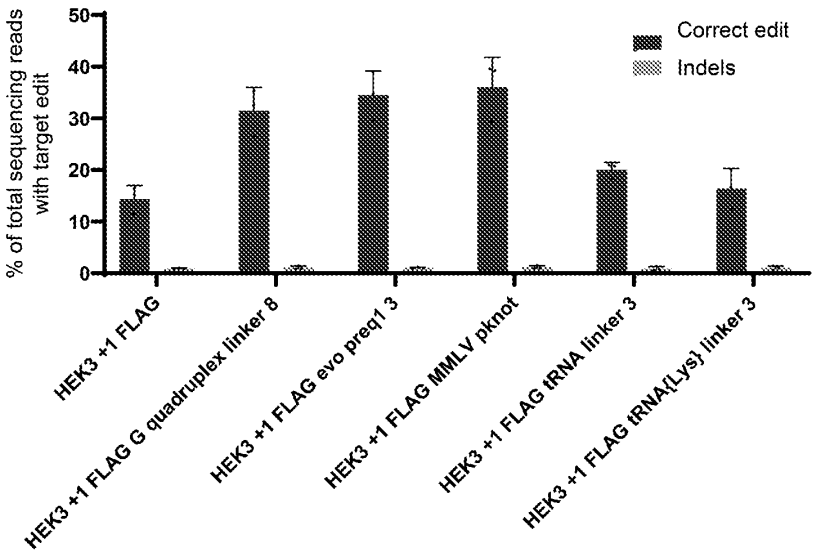


FIG. 92A

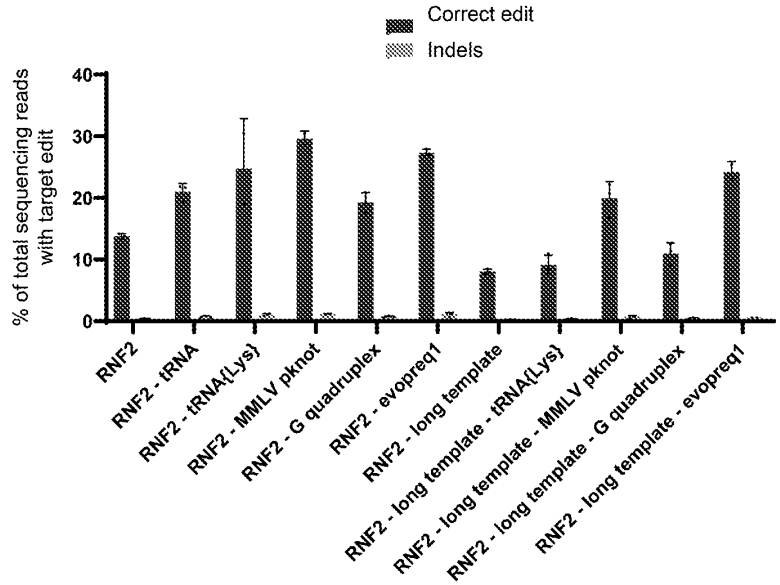


FIG. 92B

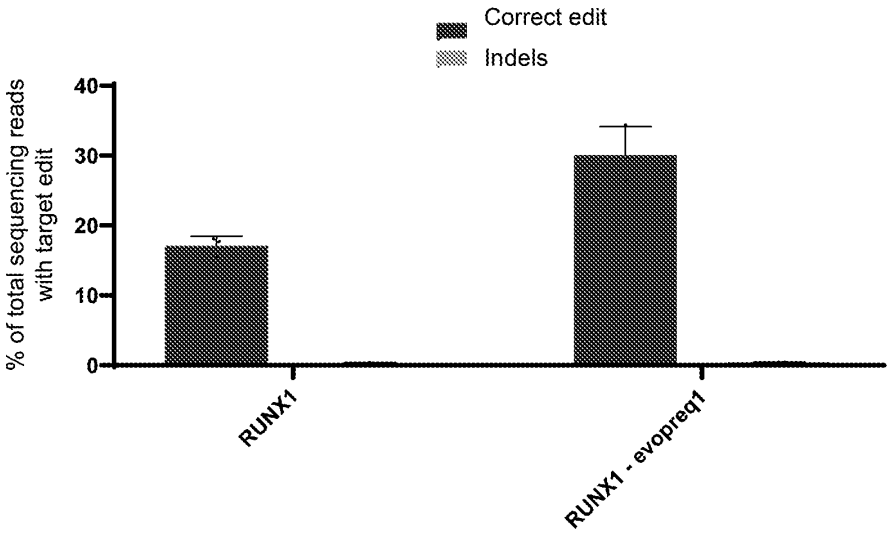

Correct edit
Indels
% of total sequencing reads with target edit
40
30
20
10
0
RUNX1
RUNX1 - evopreq1

FIG. 92C

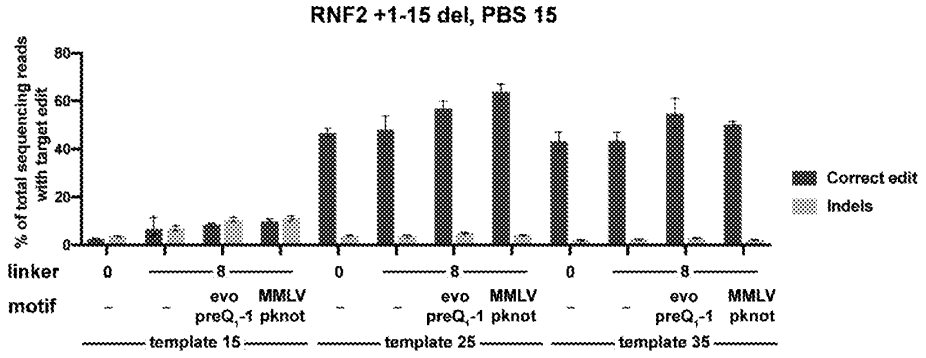


FIG. 93A

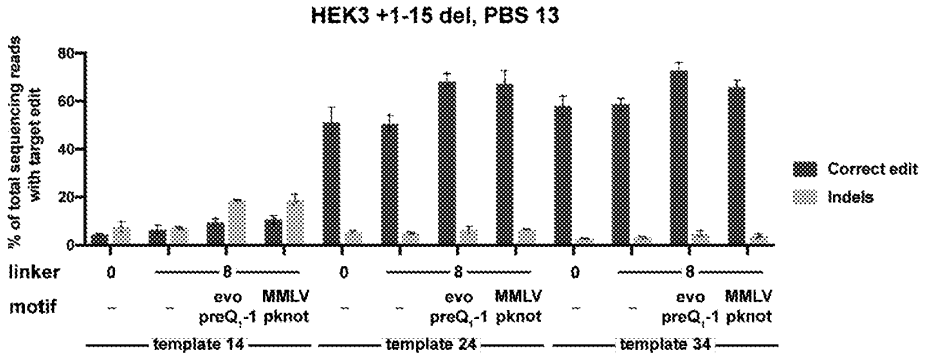


FIG. 93B

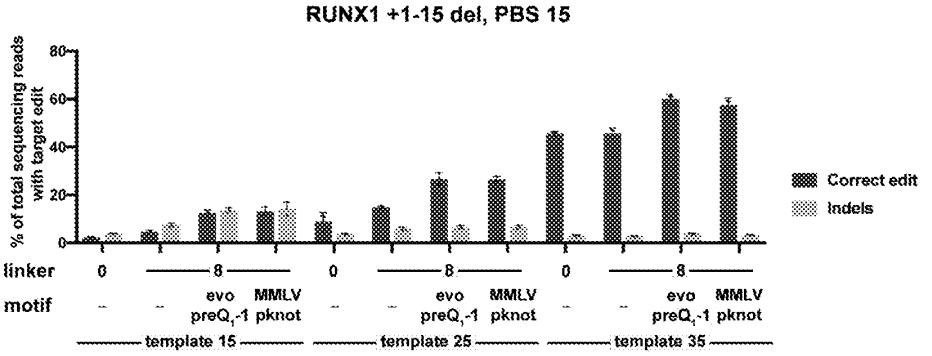


FIG. 93C

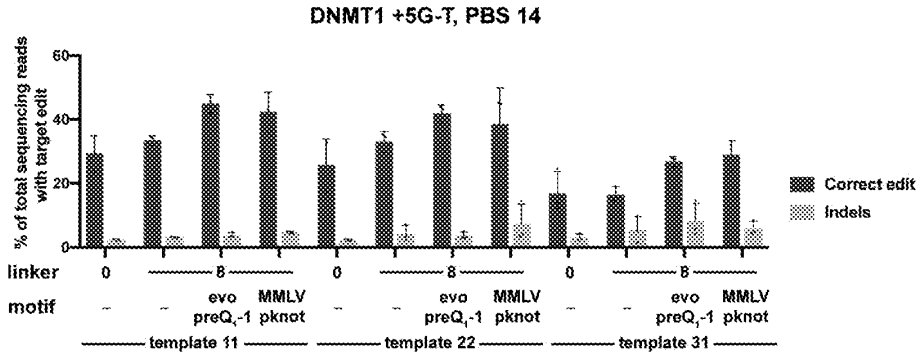


FIG. 93D

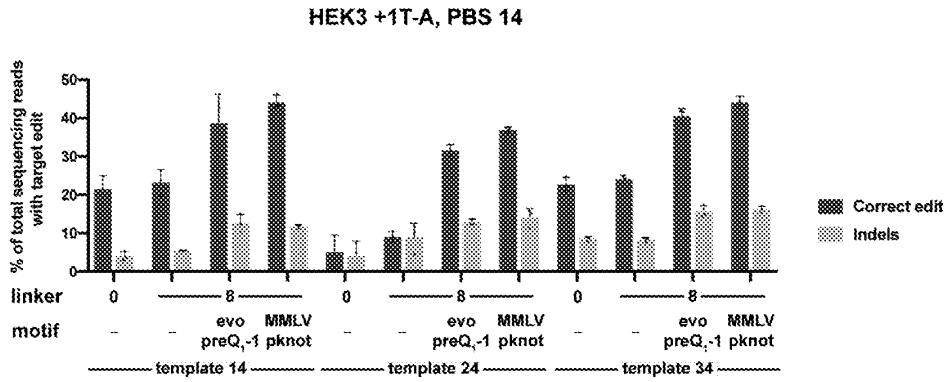


FIG. 93E

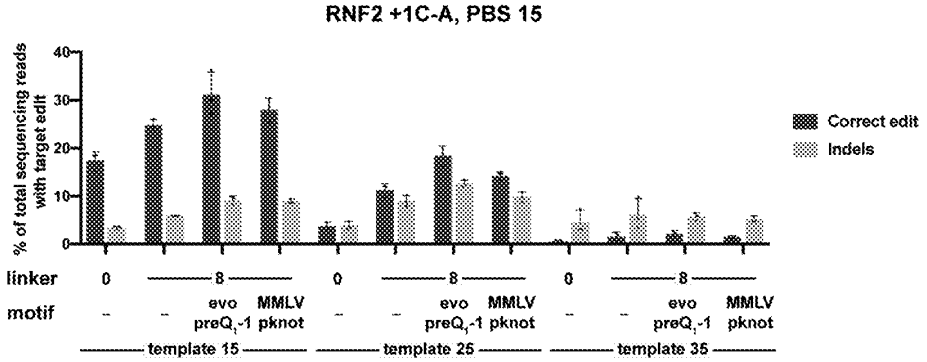


FIG. 93F

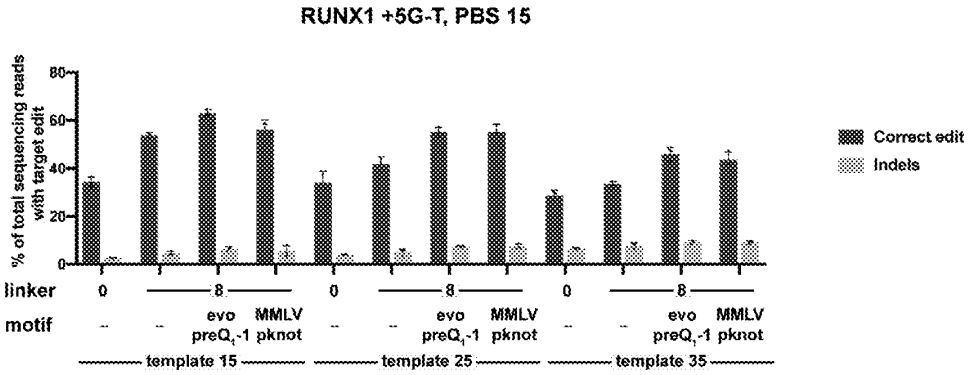


FIG. 93G

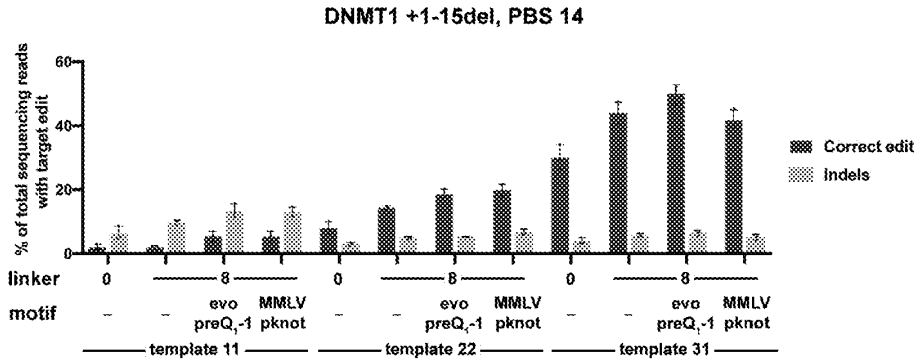


FIG. 93H

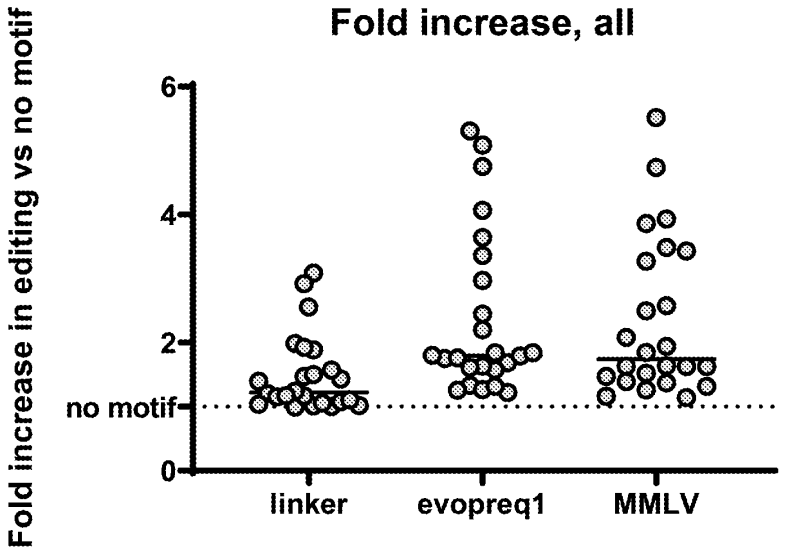

Fold increase, all
Fold increase in editing vs no motif
6
4
2
no motif
0
linker
evopreq1
MMLV

FIG. 94

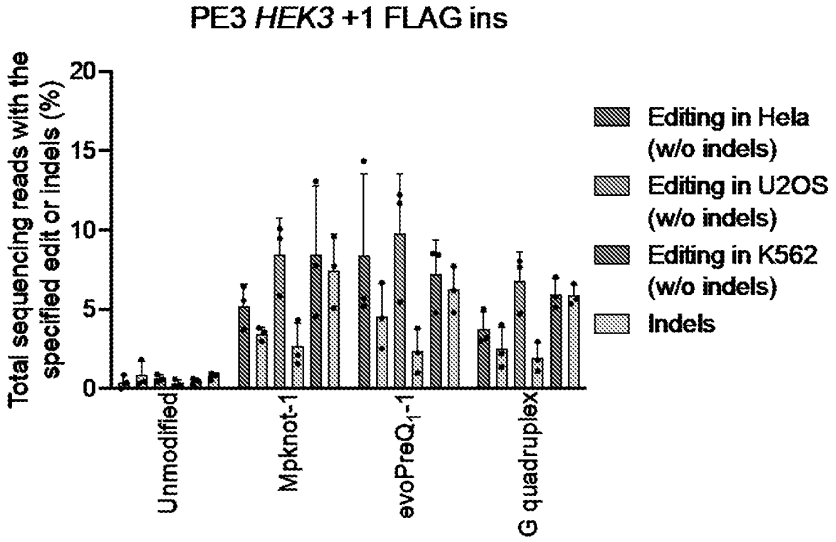


FIG. 95A

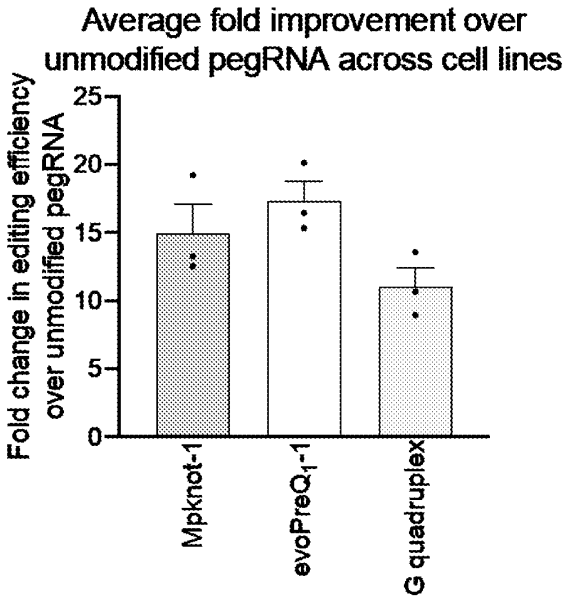


FIG. 95B

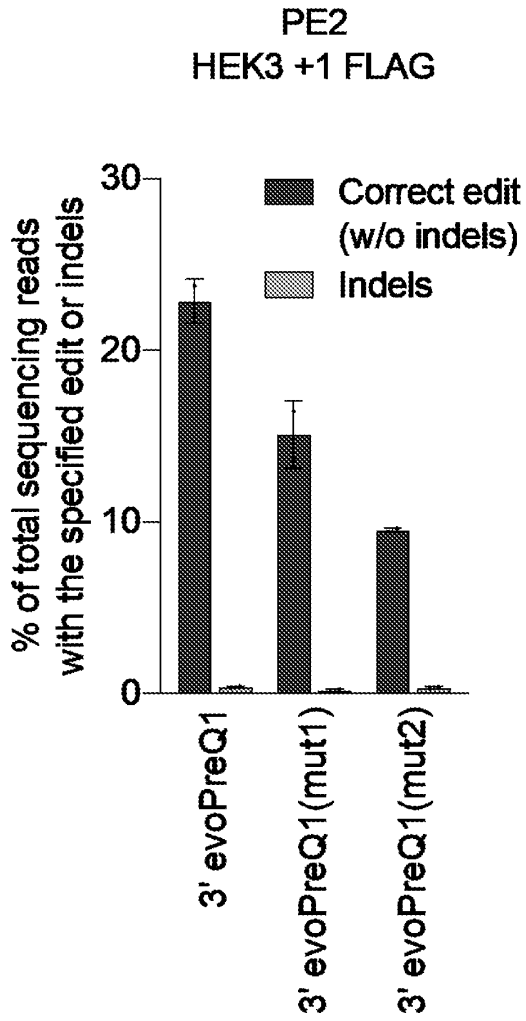

PE2
HEK3 +1 FLAG
% of total sequencing reads with the specified edit or indels
30
20
10
0
Correct edit (w/o indels)
Indels
3' evoPreQ1
3' evoPreQ1(mut1)
3' evoPreQ1(mut2)

FIG. 96

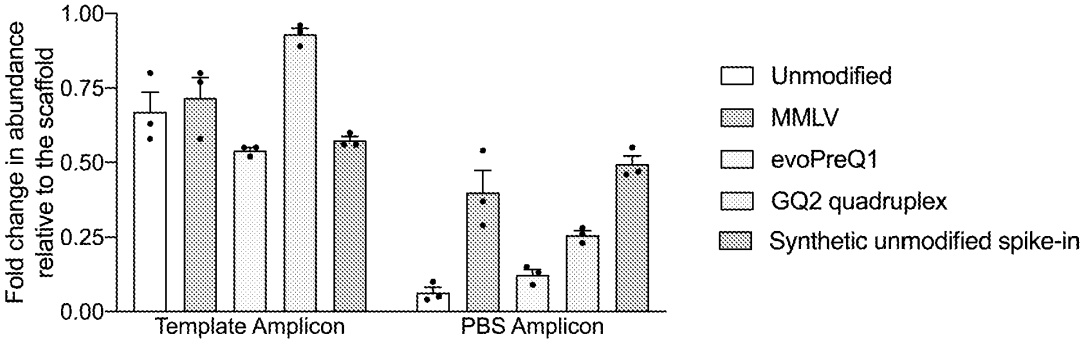

Fold change in abundance relative to the scaffold
1.00
0.75
0.50
0.25
0.00
Template Amplicon
PBS Amplicon
Unmodified
MMLV
evoPreQ1
GQ2 quadruplex
Synthetic unmodified spike-in

FIG. 97

pegRNA appended at 3' end with nucleic acid moiety

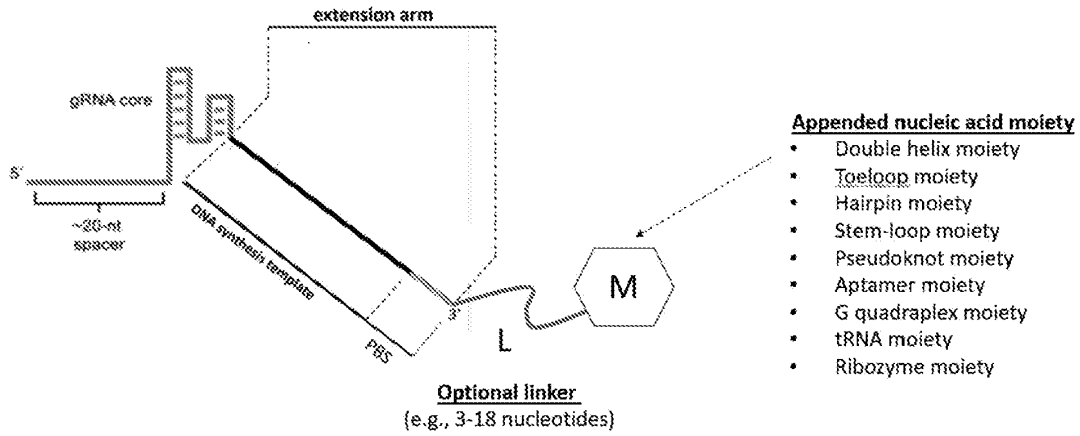


FIG. 98

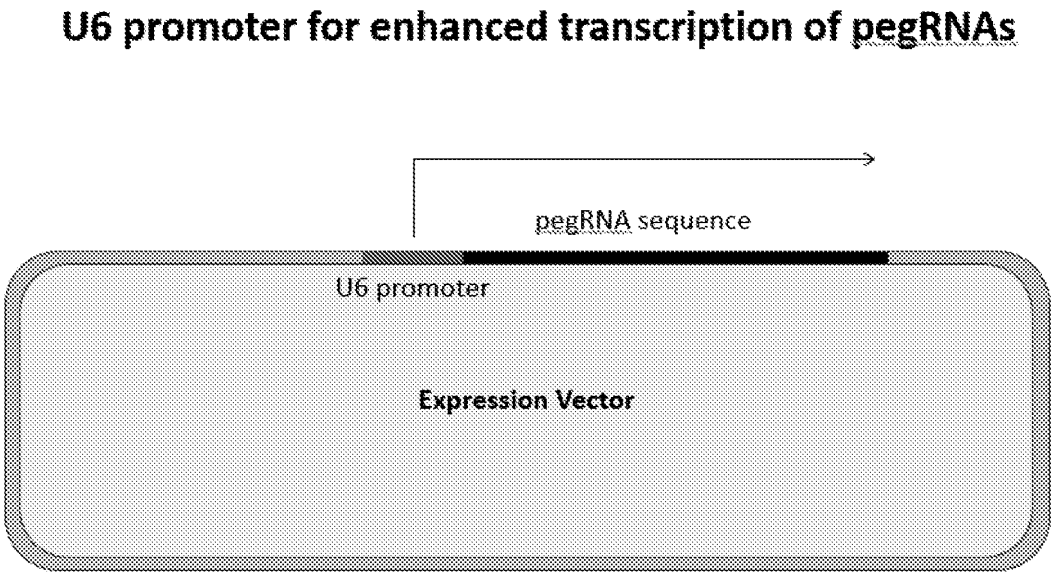

U6 promoter for enhanced transcription of pegRNAs
pegRNA sequence
U6 promoter
Expression Vector

FIG. 99

HEK3 +1FLAG

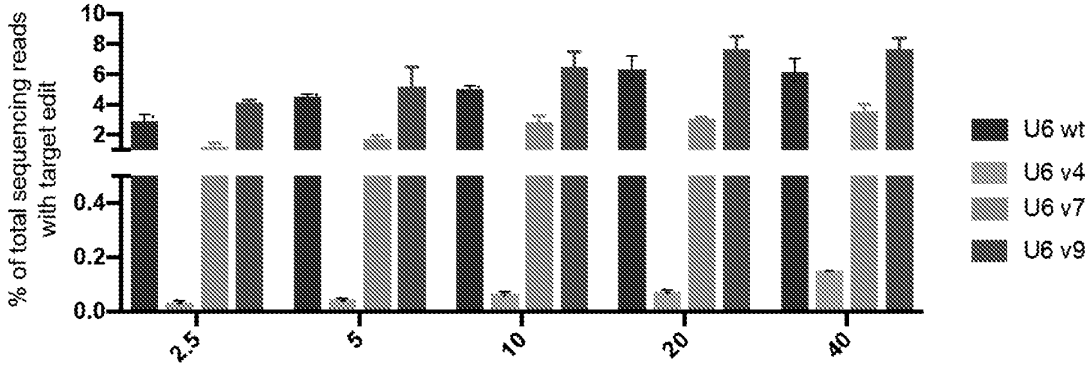


FIG. 100

HEK3 +1CTT

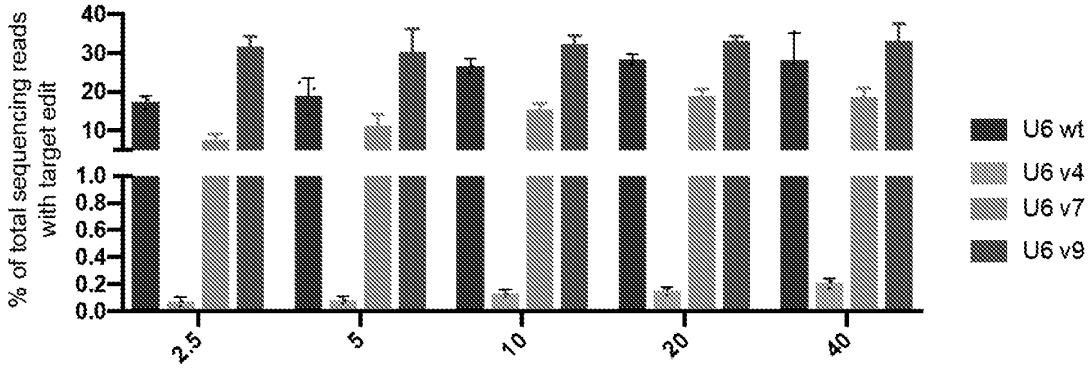


FIG. 100B

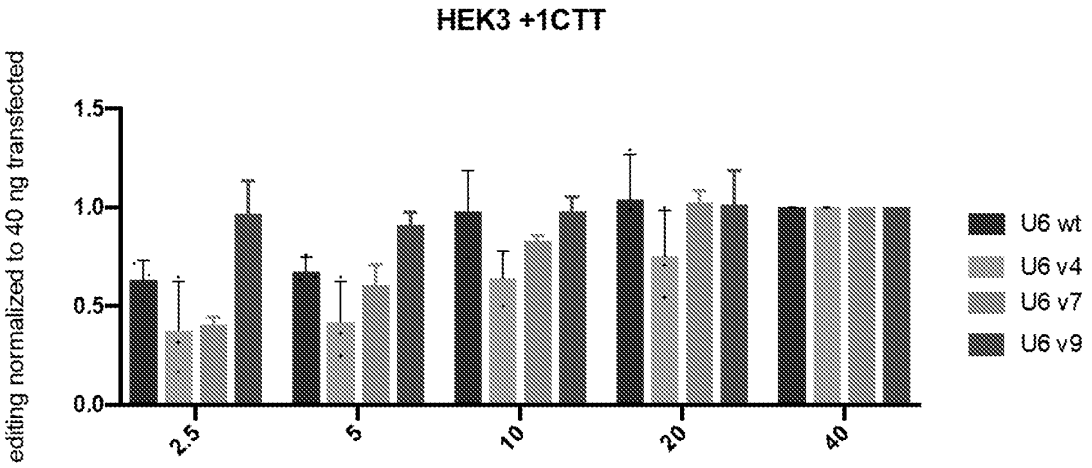


FIG. 100C

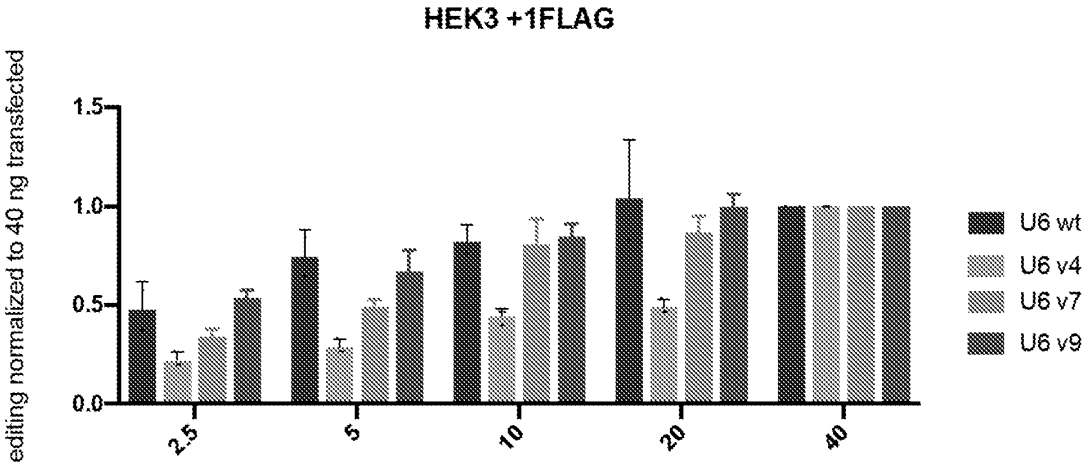


FIG. 100D

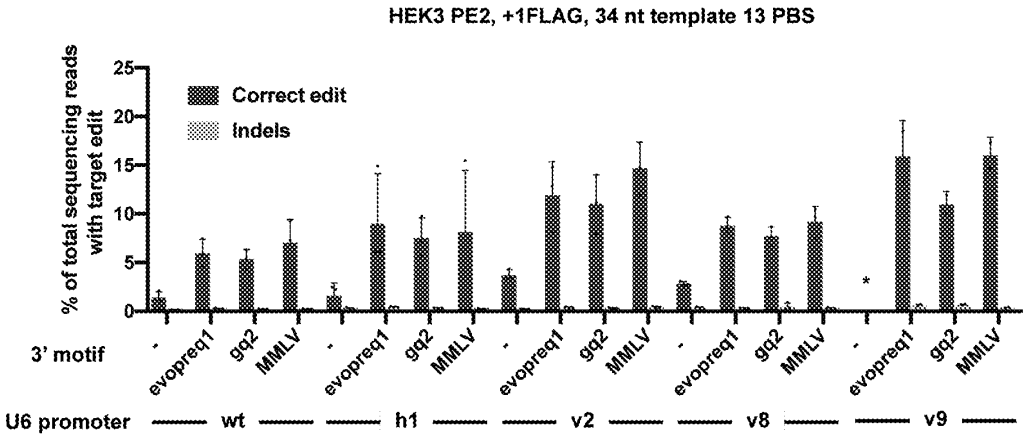
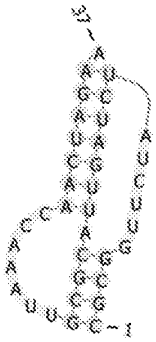

HEK3 PE2, +1FLAG, 34 nt template 13 PBS
% of total sequencing reads with target edit
25
20
15
10
5
0
Correct edit
Indels
3' motif
-
evopreq1
gq2
MMLV
U6 promoter
wt
h1
v2
v8
v9

FIG. 100E

FS-2

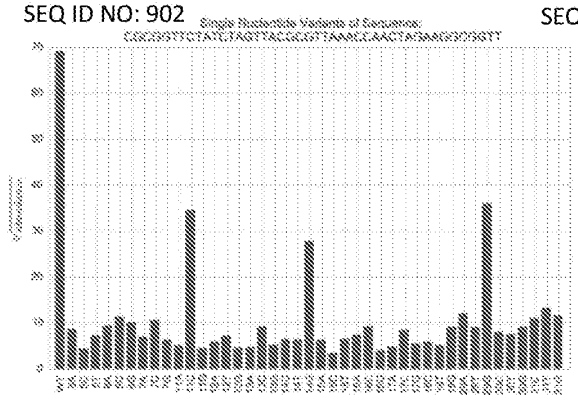
Motif 103



No. supporting sequences	23
Mean abundance	20,074
Stand. dev. abundance	69,219
Mode neighborhood occ.	0.055
Mode neighborhood entropy	0.348



SEQ ID NO: 902



SEQ ID NO: 902

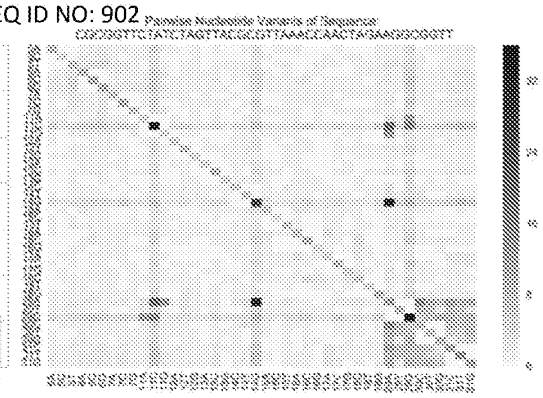
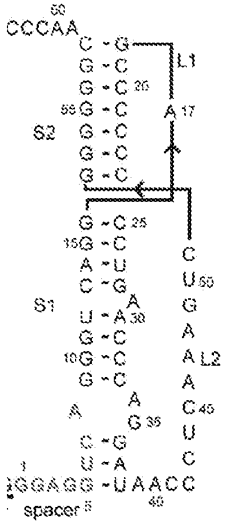


FIG. 101

SEQ ID NO: 195



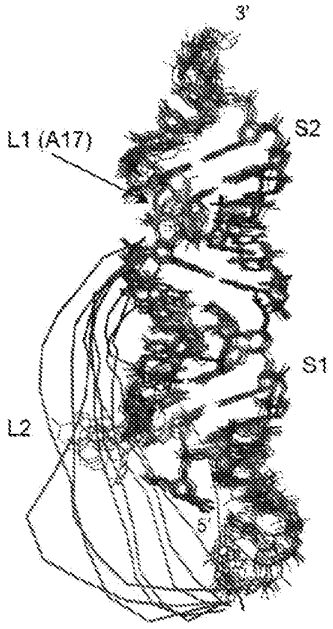


FIG. 102

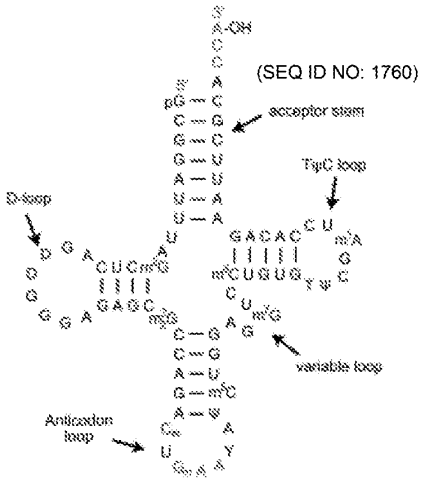


FIG. 103A

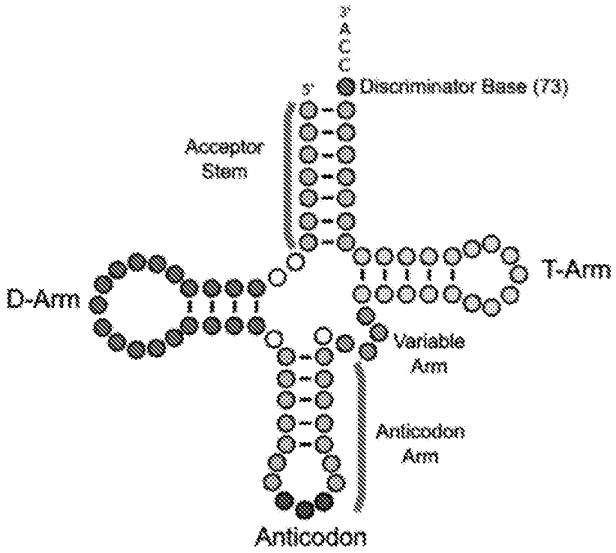


FIG. 103B

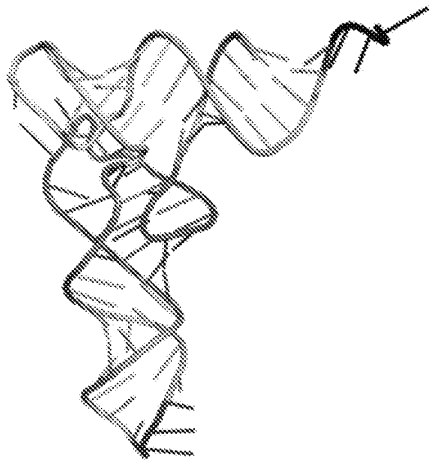


FIG. 103C

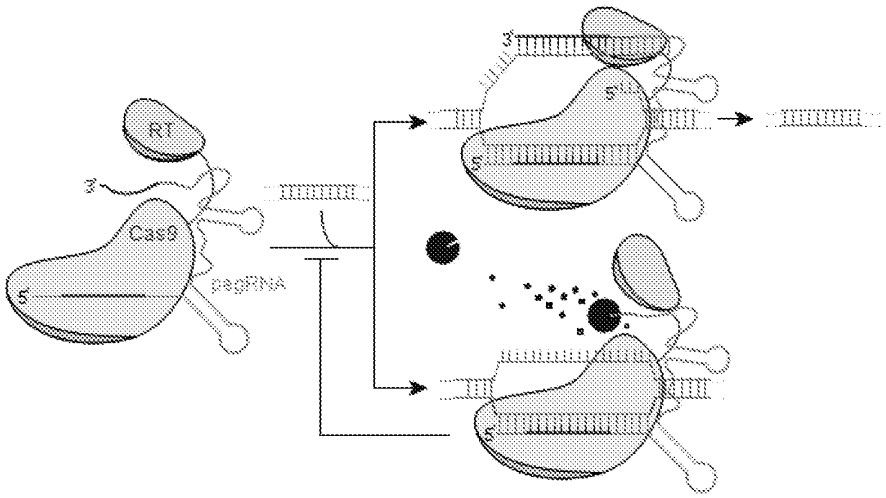


FIG. 104A

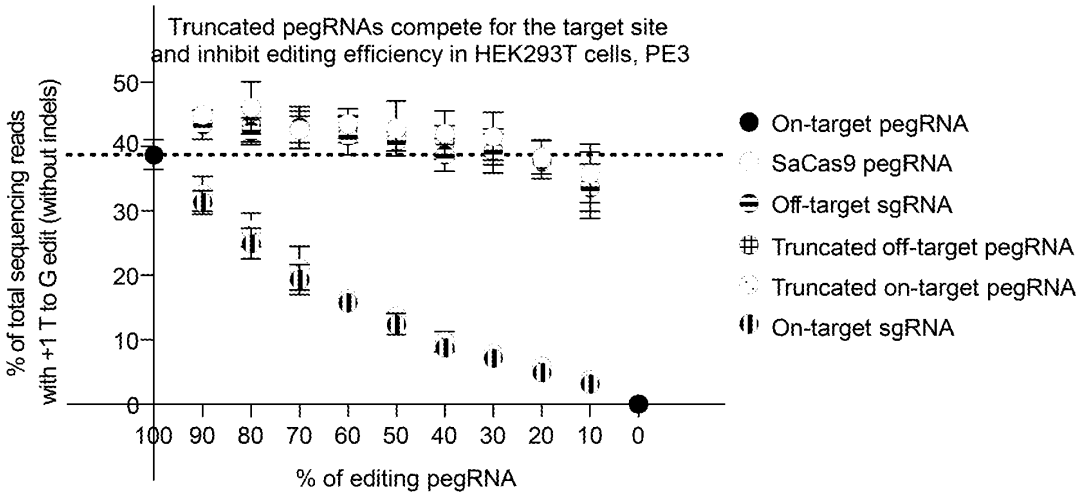


FIG. 104B

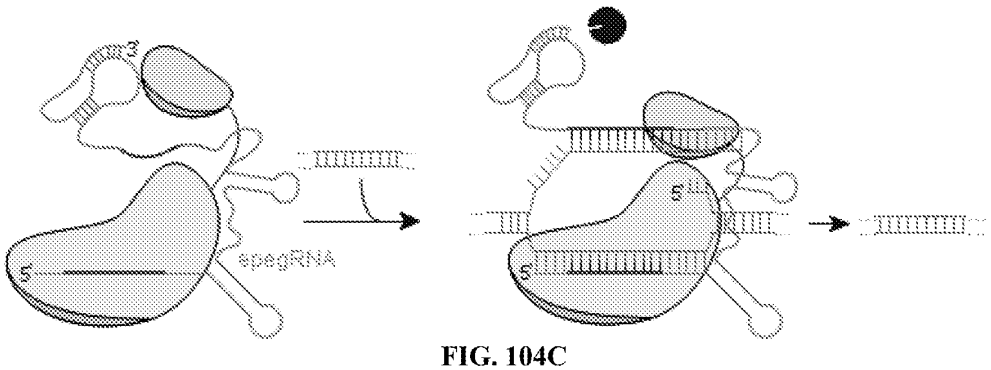


FIG. 104C

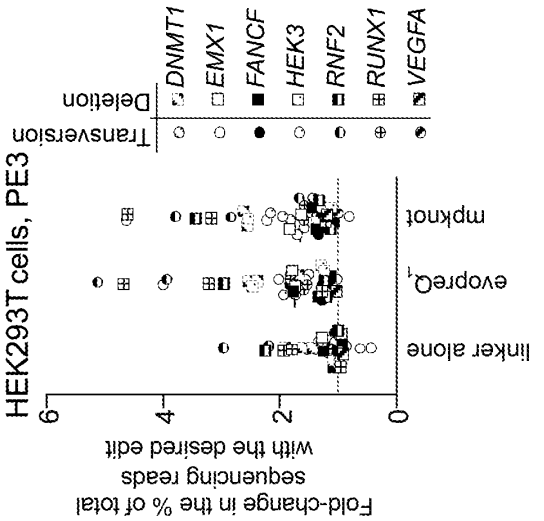


FIG. 105B

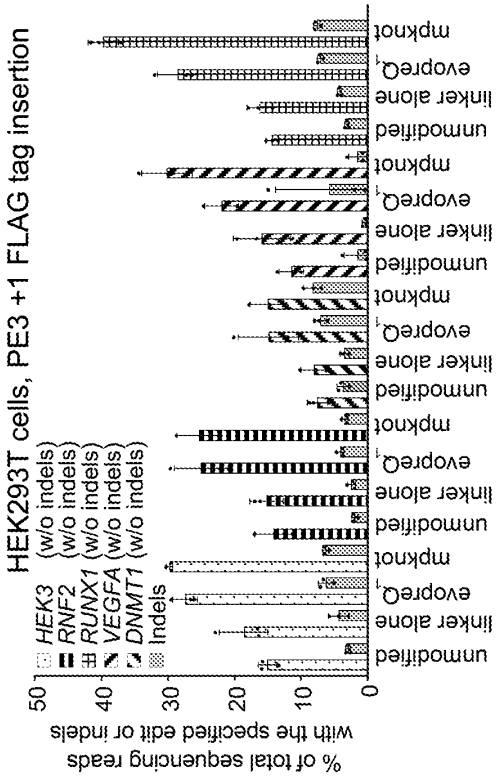


FIG. 105A

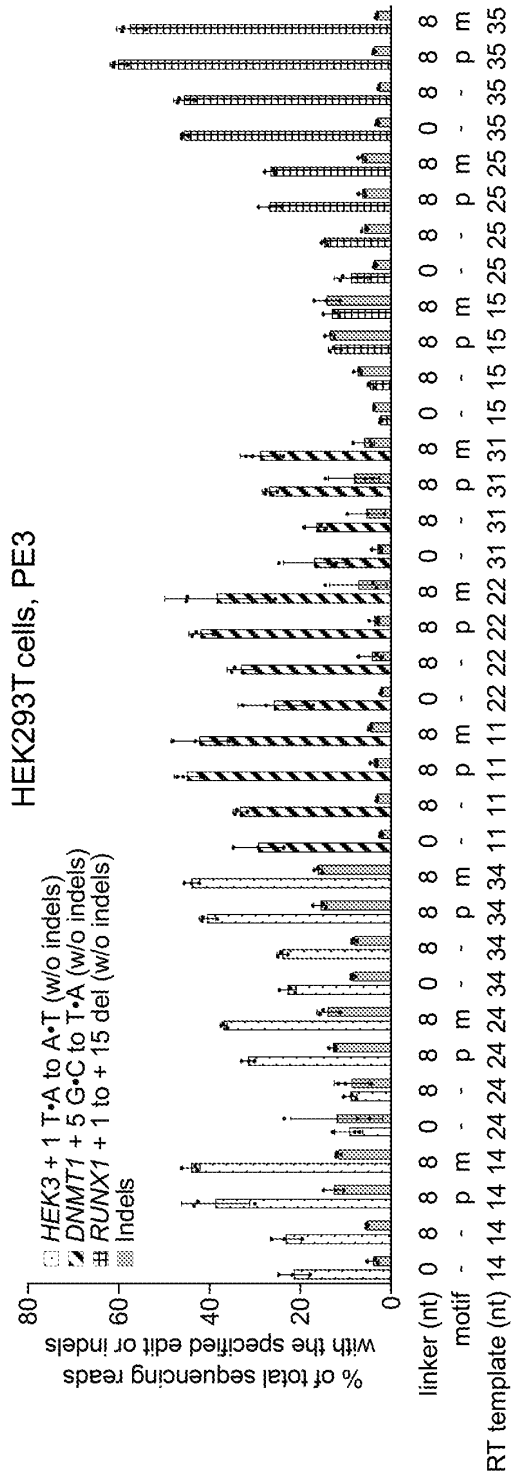

HEK293T cells, PE3
HEK3 + 1 T•A to A•T (w/o indels)
DNMT1 + 5 G•C to T•A (w/o indels)
RUNX1 + 1 to + 15 del (w/o indels)
Indels
% of total sequencing reads with the specified edit or indels
0
20
40
60
80
linker (nt) 0 8 8 8 0 8 8 8 8 0 8 8 8 8 0 8 8 8 8 0 8 8 8 8 0 8 8 8 8 0 8 8 8 8 0 8 8 8 8 0 8 8 8 8
motif - - p m - - p m - - p m - - p m - - p m - - p m - - p m - - p m - - p m - - p m - - p m
RT template (nt) 14 14 14 24 24 24 24 34 34 34 34 11 11 11 11 22 22 22 22 31 31 31 31 15 15 15 15 25 25 25 25 35 35 35 35

FIG. 105C

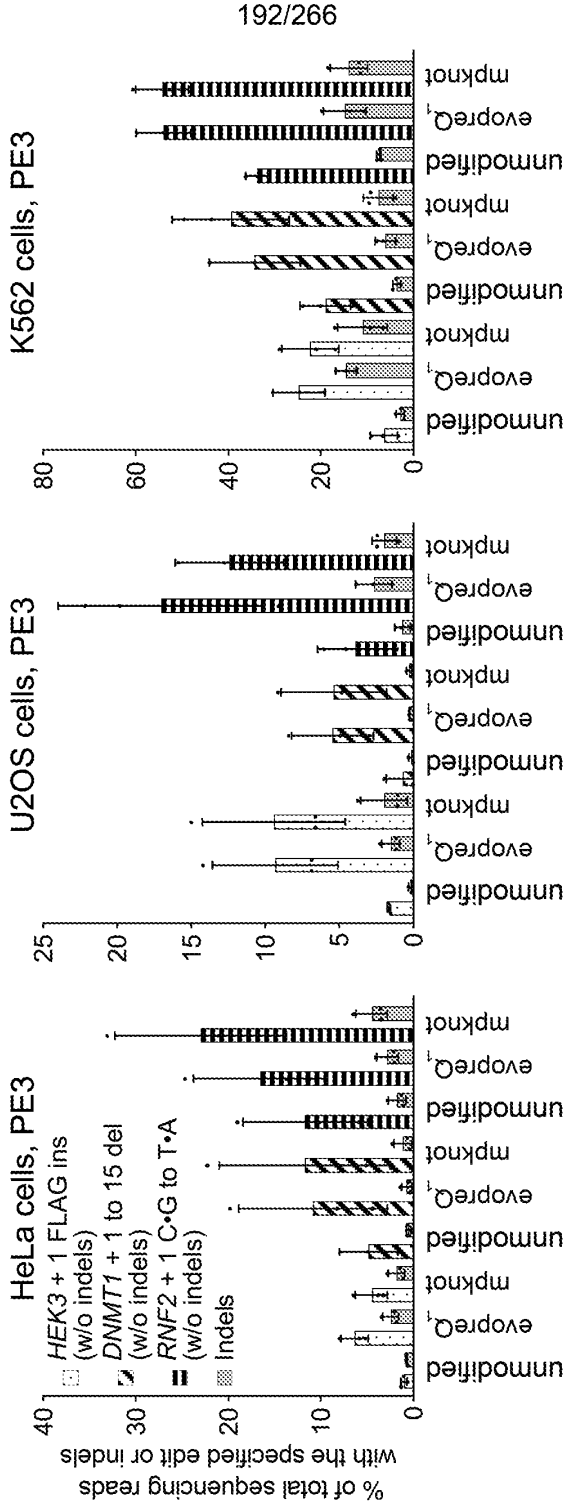

HeLa cells, PE3
U2OS cells, PE3
K562 cells, PE3
% of total sequencing reads with the specified edit or indels
HEK3 + 1 FLAG ins (w/o indels)
DNMT1 + 1 to 15 del (w/o indels)
RNF2 + 1 C•G to T•A (w/o indels)
Indels
unmodified
evopreQ1
mpknot

FIG. 105D

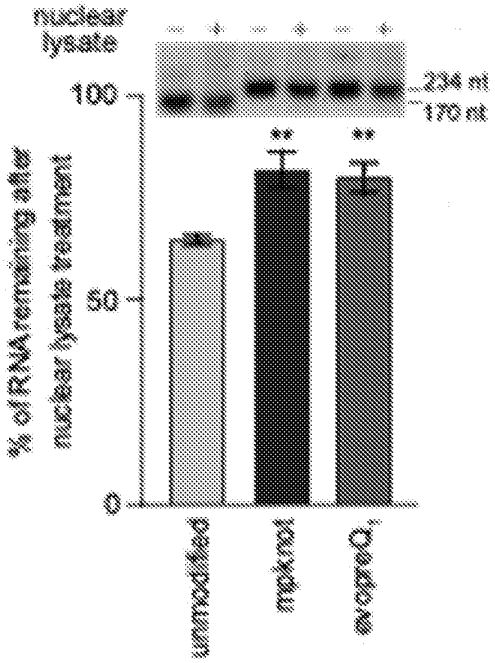


FIG. 106A

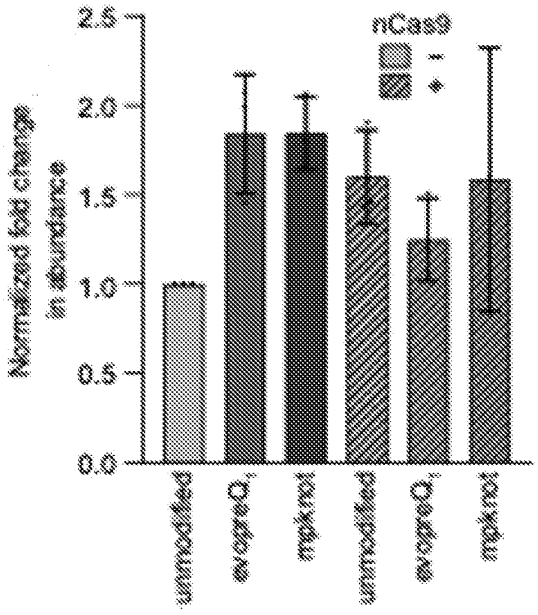


FIG. 106B

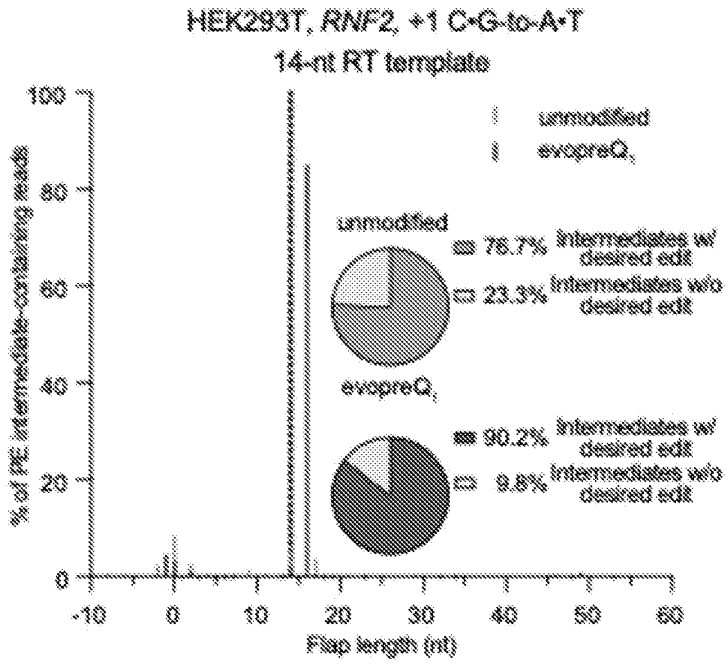


FIG. 106C

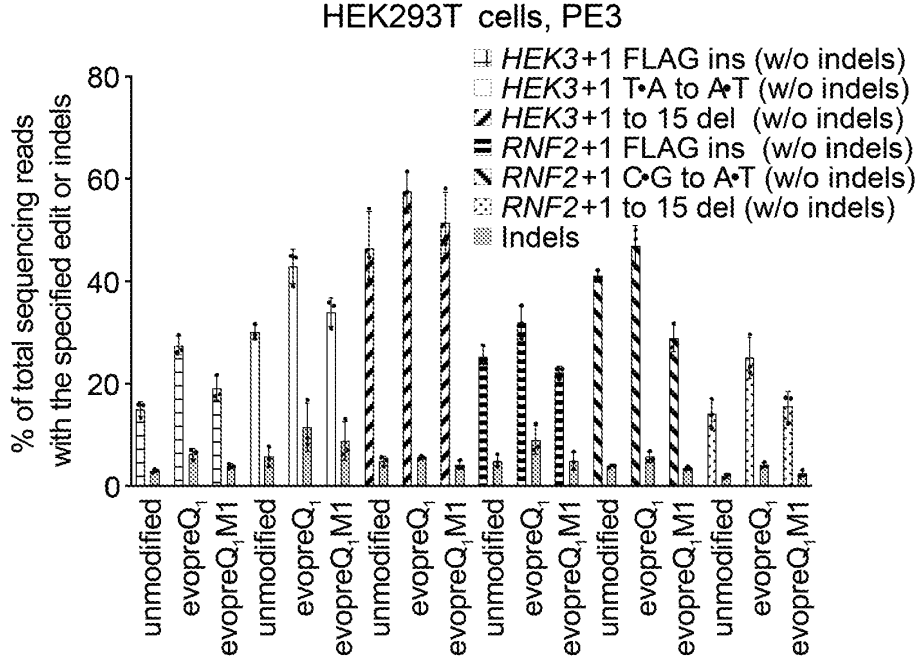


FIG. 106D

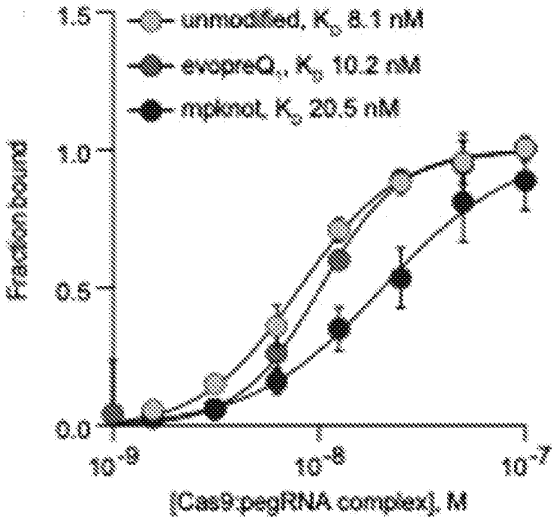


FIG. 106E

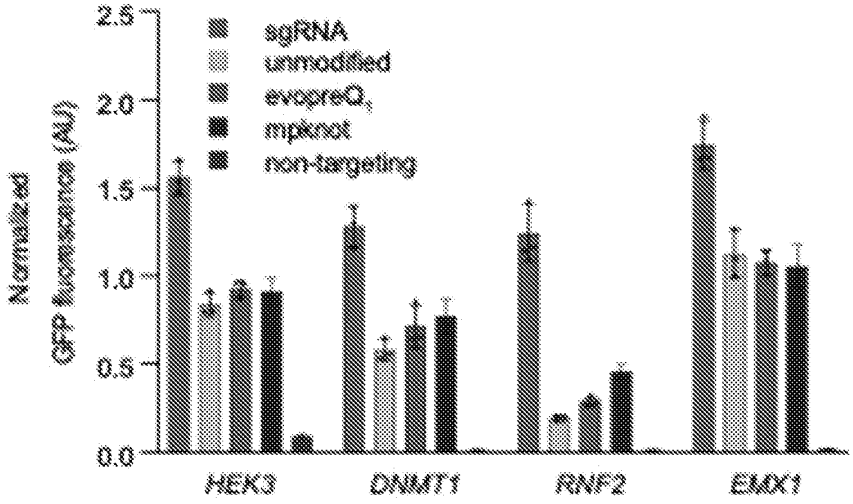


FIG. 106F

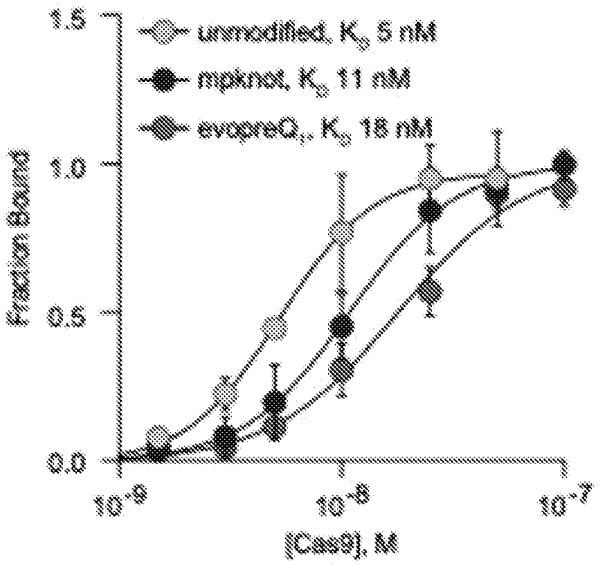


FIG. 106G

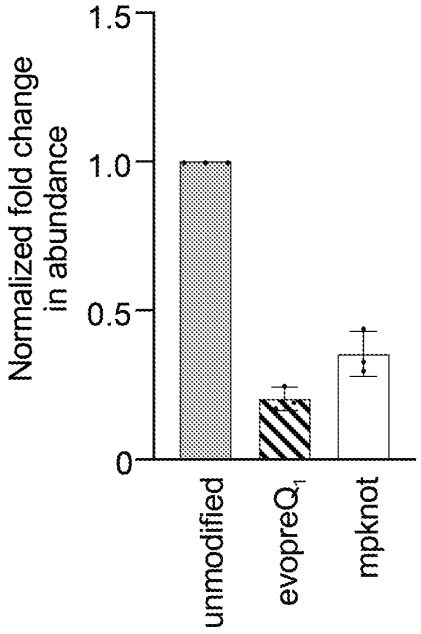


FIG. 106H

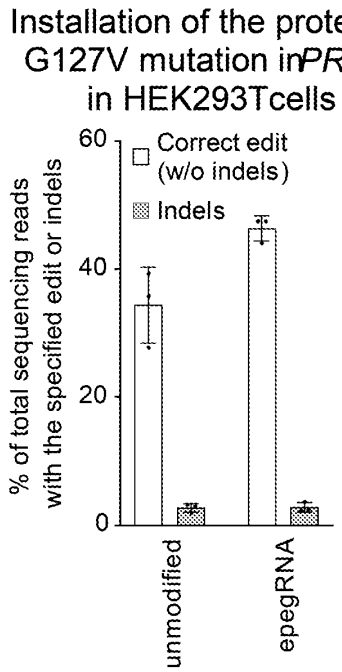


FIG. 107A

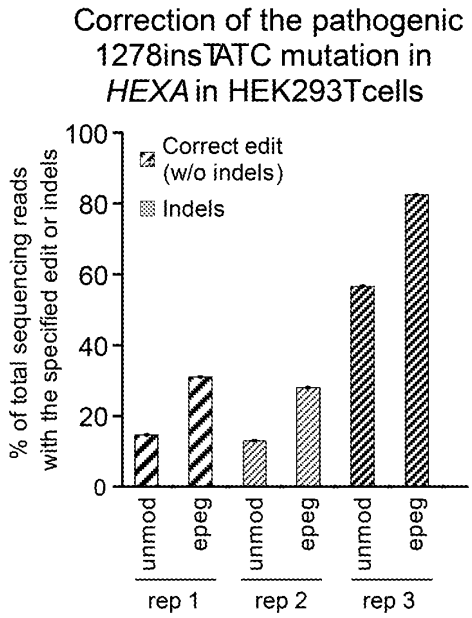


FIG. 107B

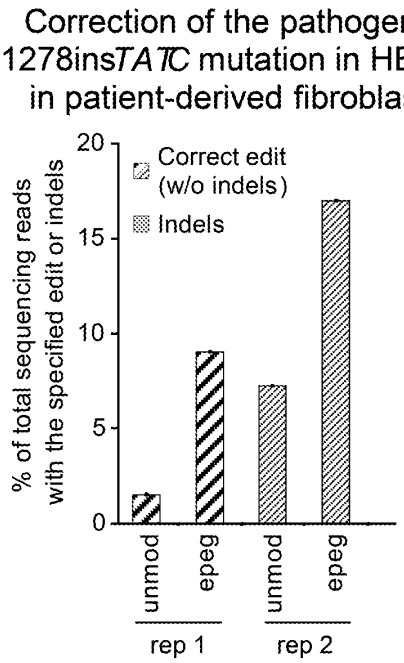


FIG. 107C

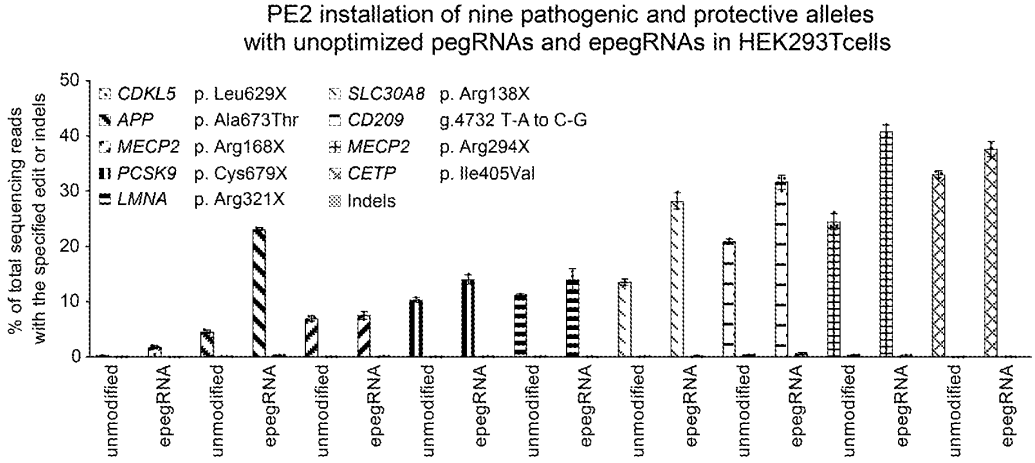


FIG. 107D

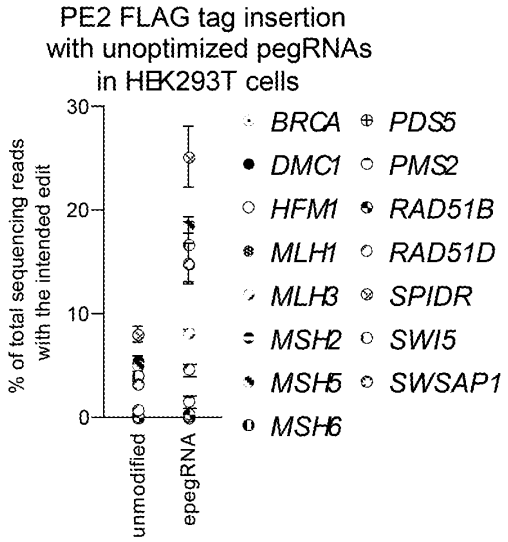

PE2 FLAG tag insertion
with unoptimized pegRNAs
in HEK293T cells
% of total sequencing reads
with the intended edit
30
20
10
0
unmodified
epegRNA
BRCA
DMC1
HFM1
MLH1
MLH3
MSH2
MSH5
MSH6
PDS5
PMS2
RAD51B
RAD51D
SPIDR
SWI5
SWSAP1

FIG. 107E

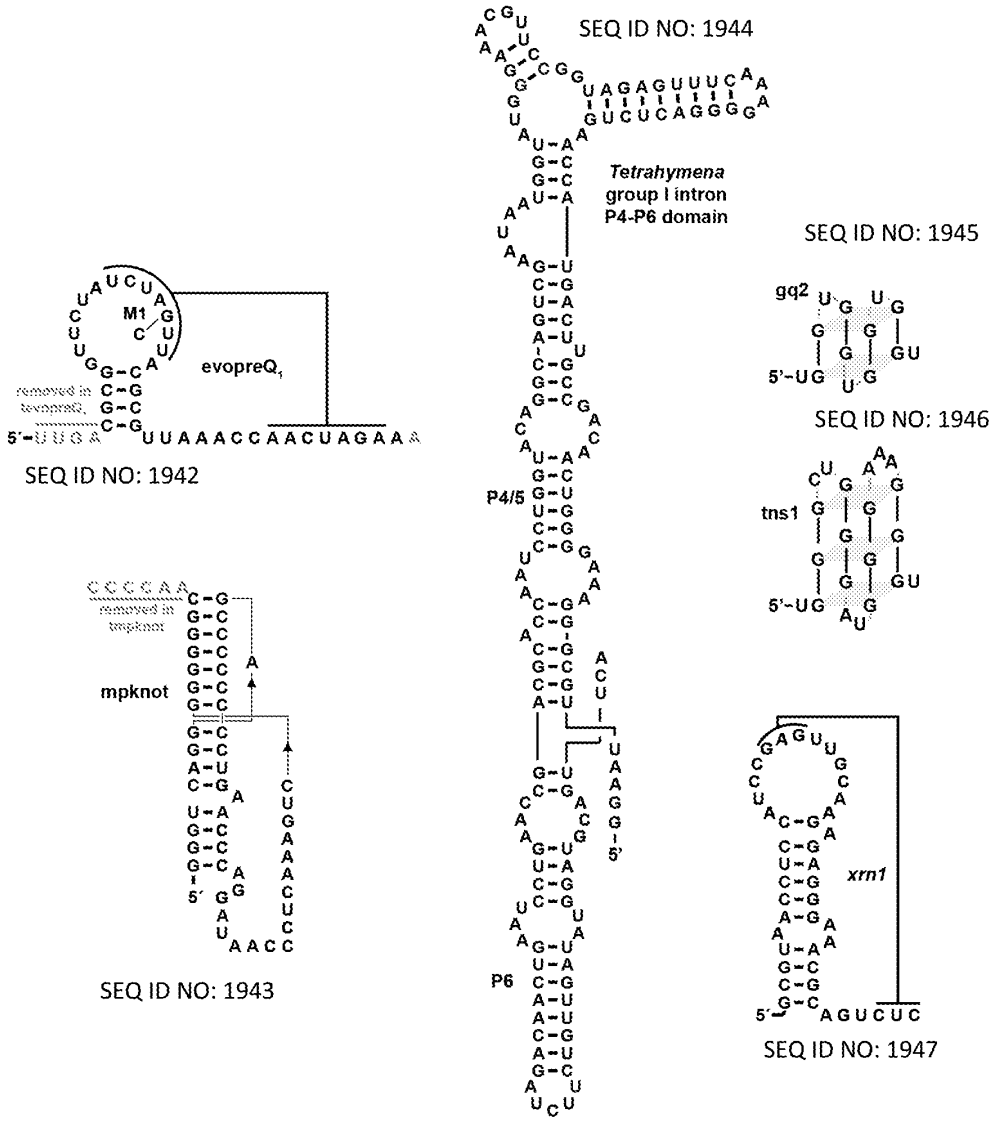

SEQ ID NO: 1944
Tetrahymena
group I intron
P4-P6 domain
P4/5
P6
SEQ ID NO: 1945
gq2
SEQ ID NO: 1946
tns1
SEQ ID NO: 1947
xrn1
M1
evopreQ1
SEQ ID NO: 1942
mpknot
SEQ ID NO: 1943

FIG. 108

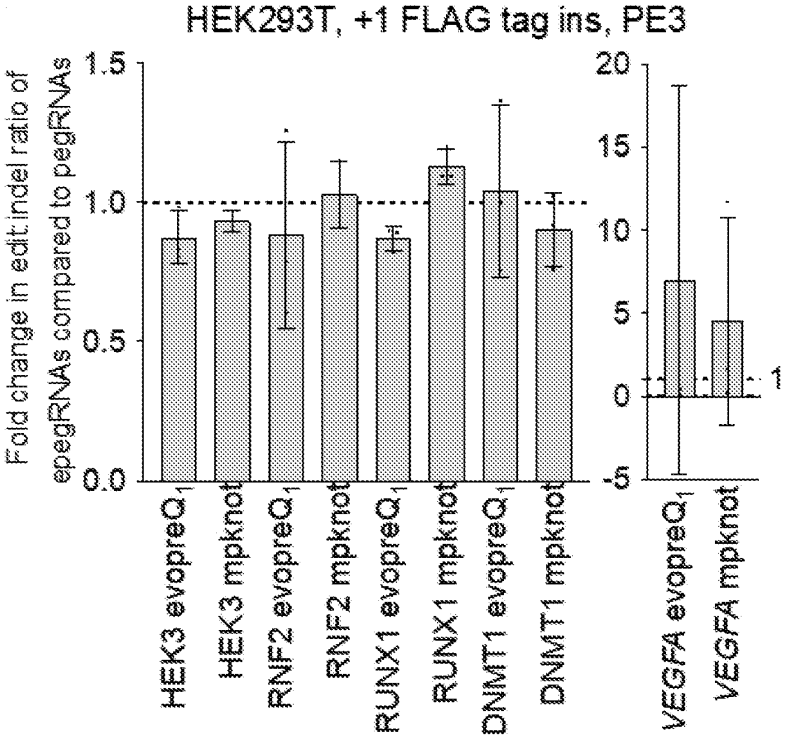


FIG. 109A

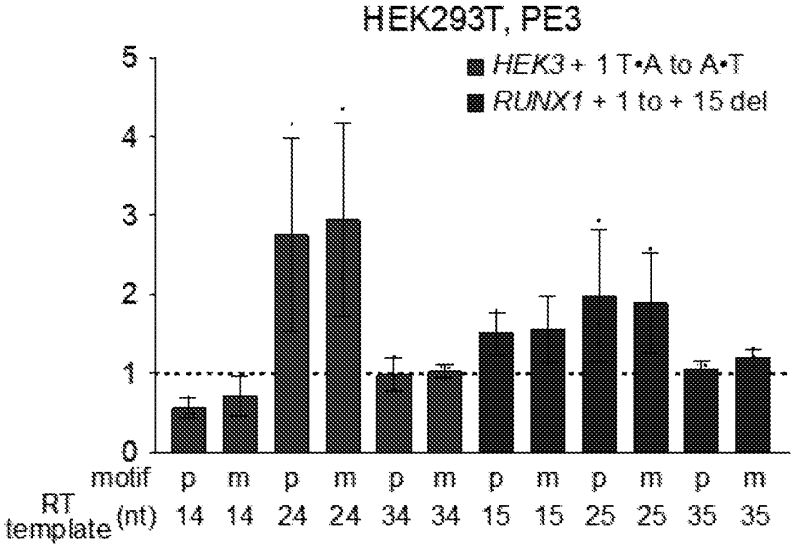


FIG. 109B

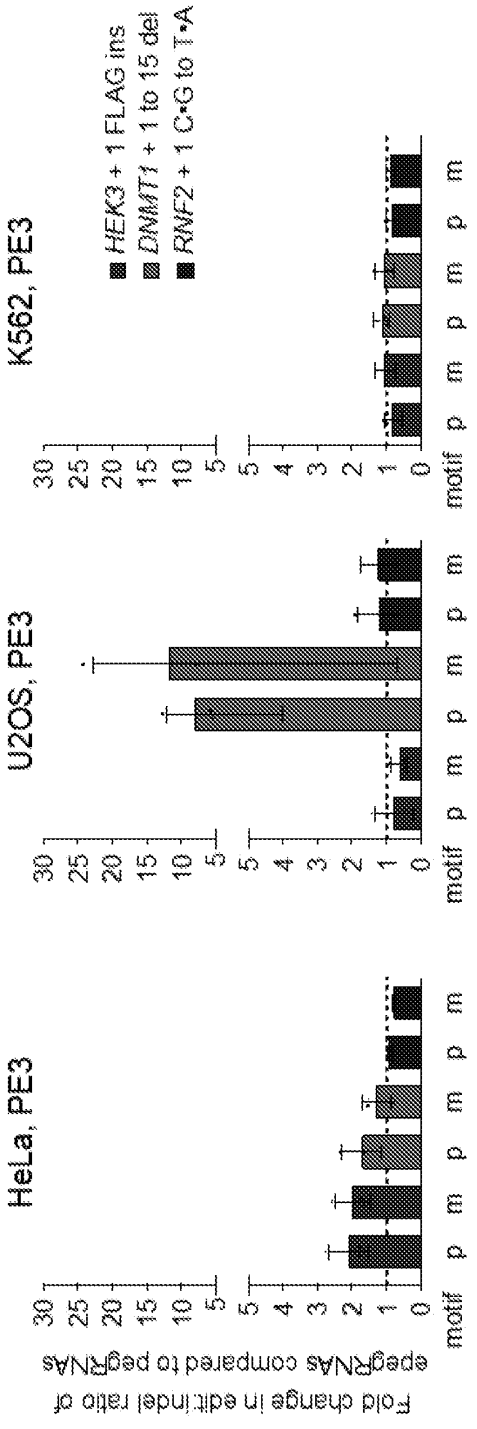

HeLa, PE3
U2OS, PE3
K562, PE3
HEK3 + 1 FLAG ins
DNMT1 + 1 to 15 del
RMF2 + 1 C•G to T•A
Fold change in edit:indel ratio of epegRNAs compared to pegRNAs
motif p m p m p m

FIG. 109C

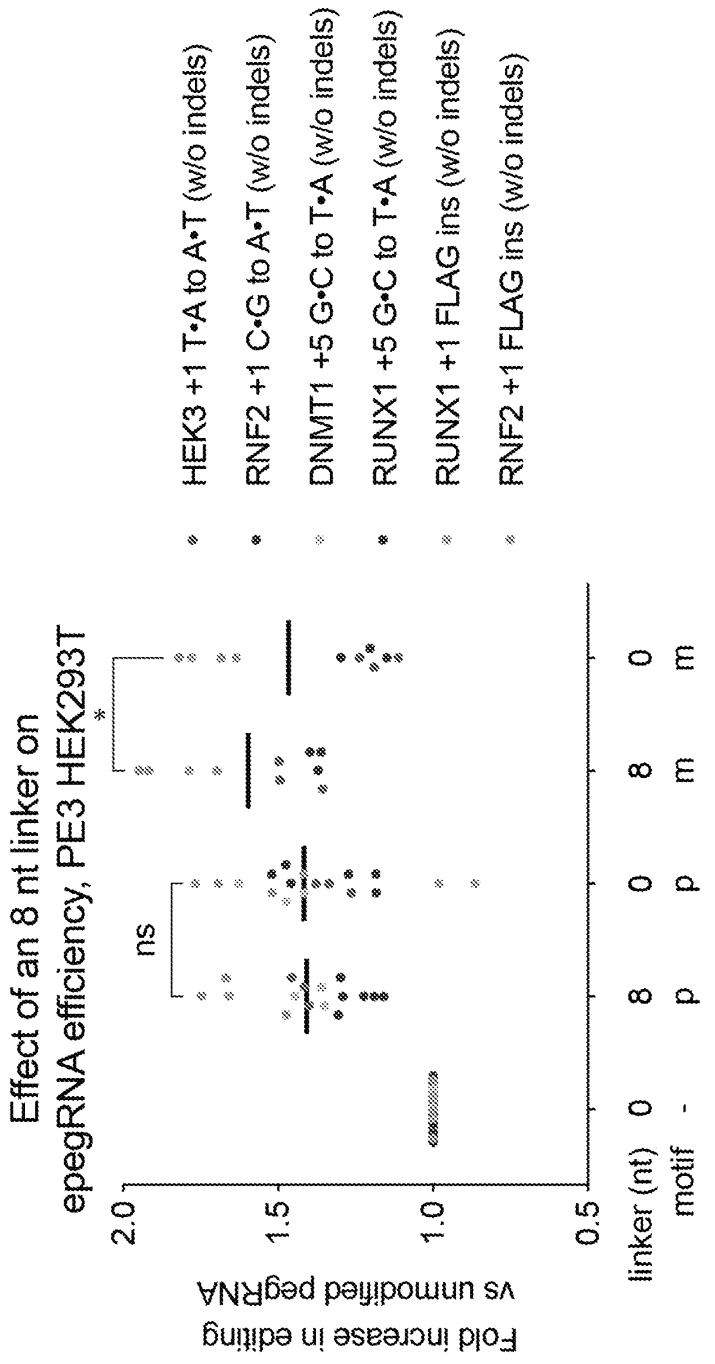

Effect of an 8 nt linker on
epegRNA efficiency, PE3 HEK293T
Fold increase in editing
vs unmodified pegRNA
2.0
1.5
1.0
0.5
ns
*
linker (nt) 0 8 0 8 0
motif - p p m m
HEK3 +1 T•A to A•T (w/o indels)
RNF2 +1 C•G to A•T (w/o indels)
DNMT1 +5 G•C to T•A (w/o indels)
RUNX1 +5 G•C to T•A (w/o indels)
RUNX1 +1 FLAG ins (w/o indels)
RNF2 +1 FLAG ins (w/o indels)

FIG. 110

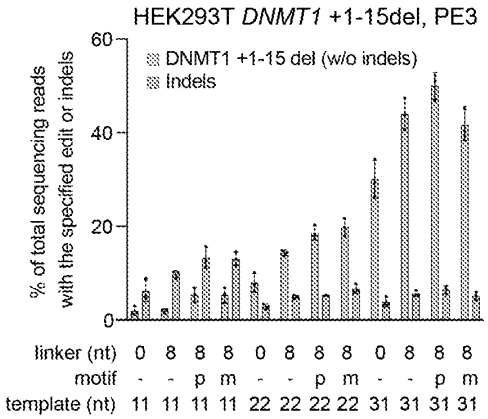


FIG. 111A

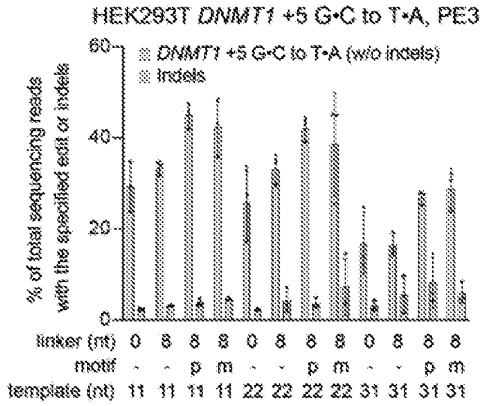


FIG. 111B

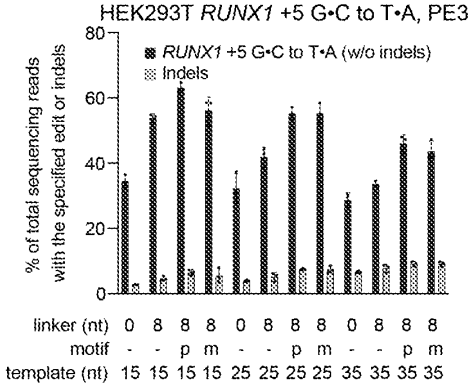


FIG. 111C

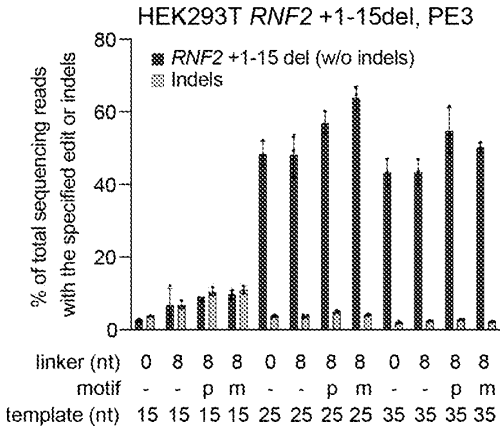


FIG. 111D

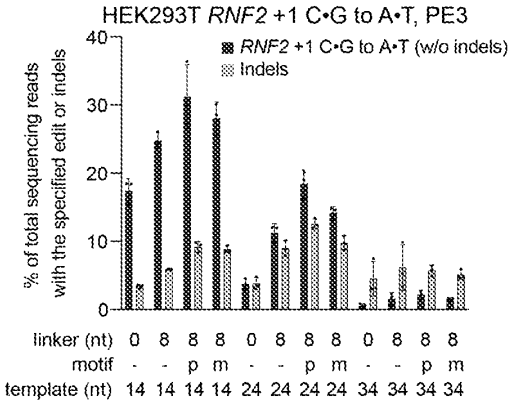


FIG. 111E

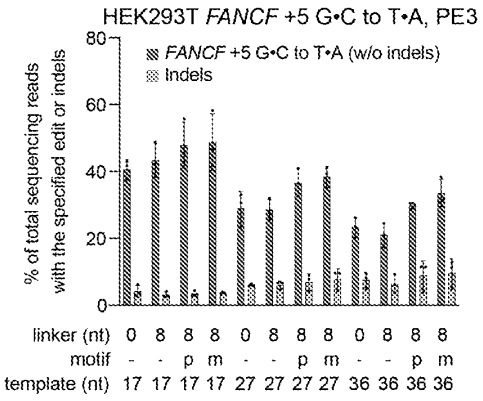


FIG. 111F

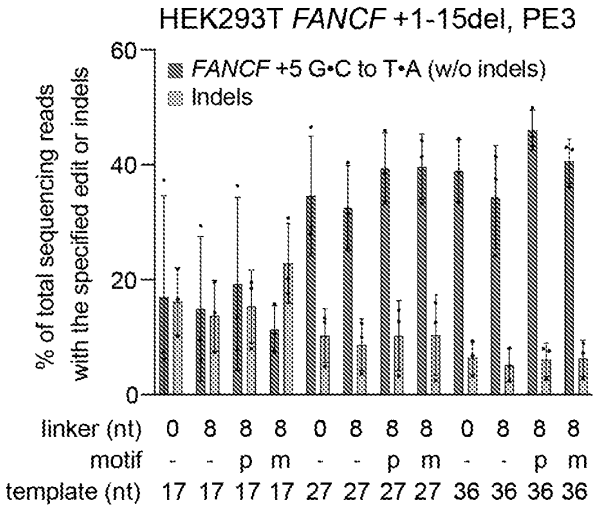


FIG. 111G

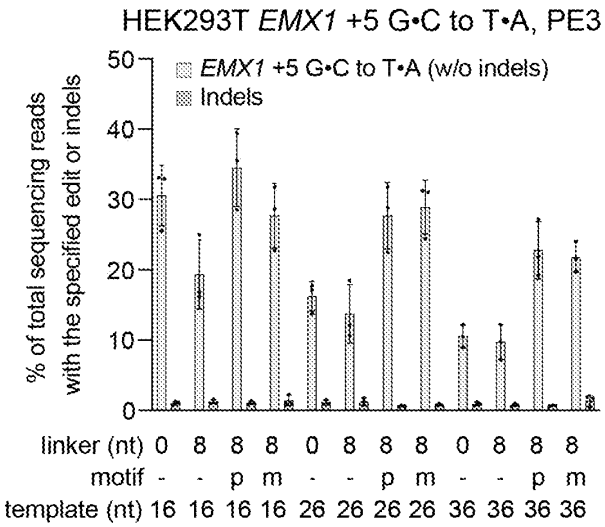


FIG. 111H

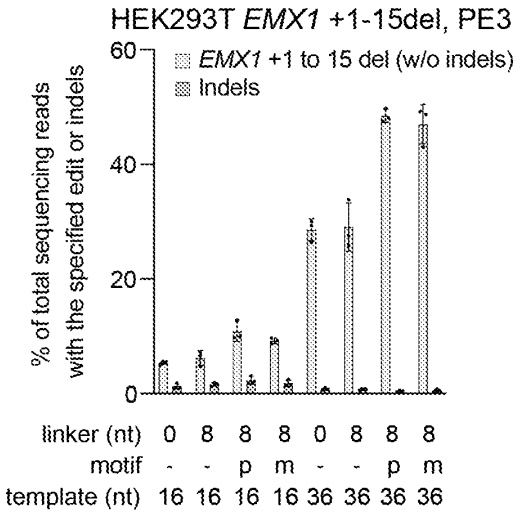


FIG. 111I

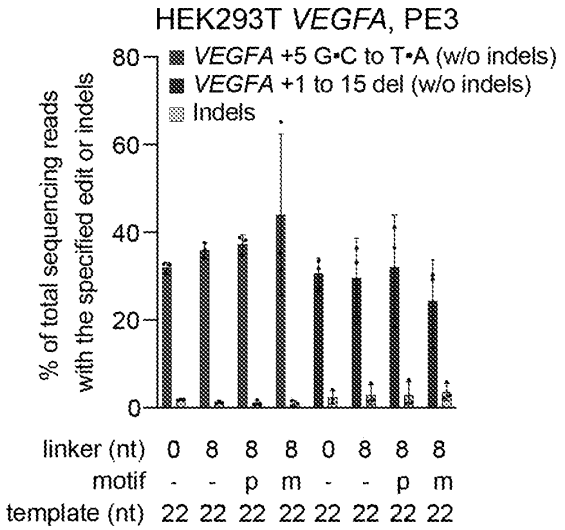


FIG. 111J

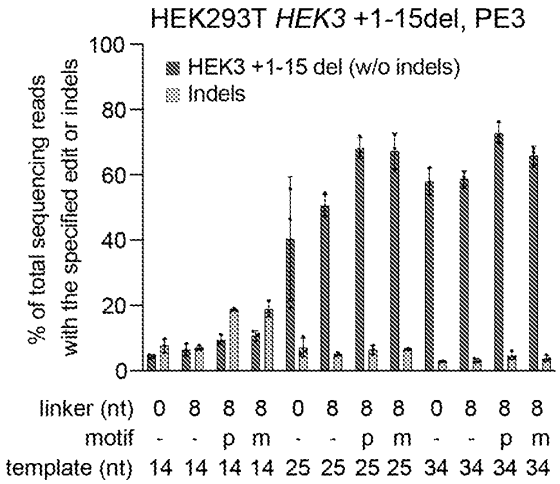


FIG. 111K

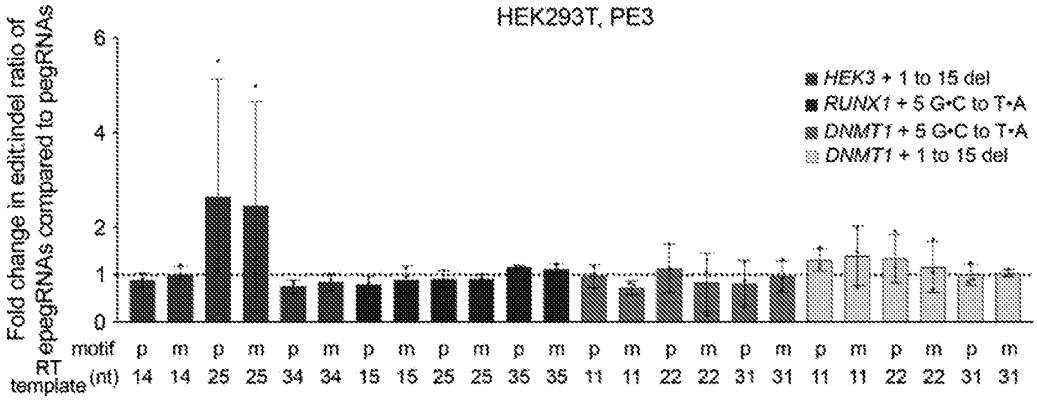


FIG. 112A

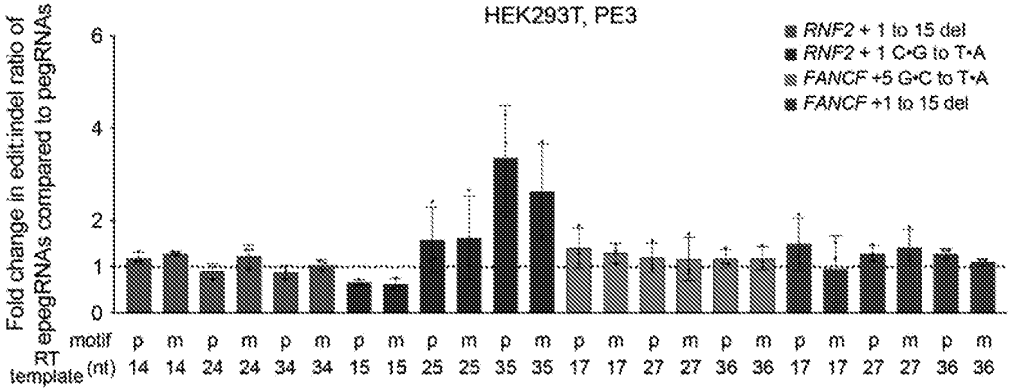


FIG. 112B

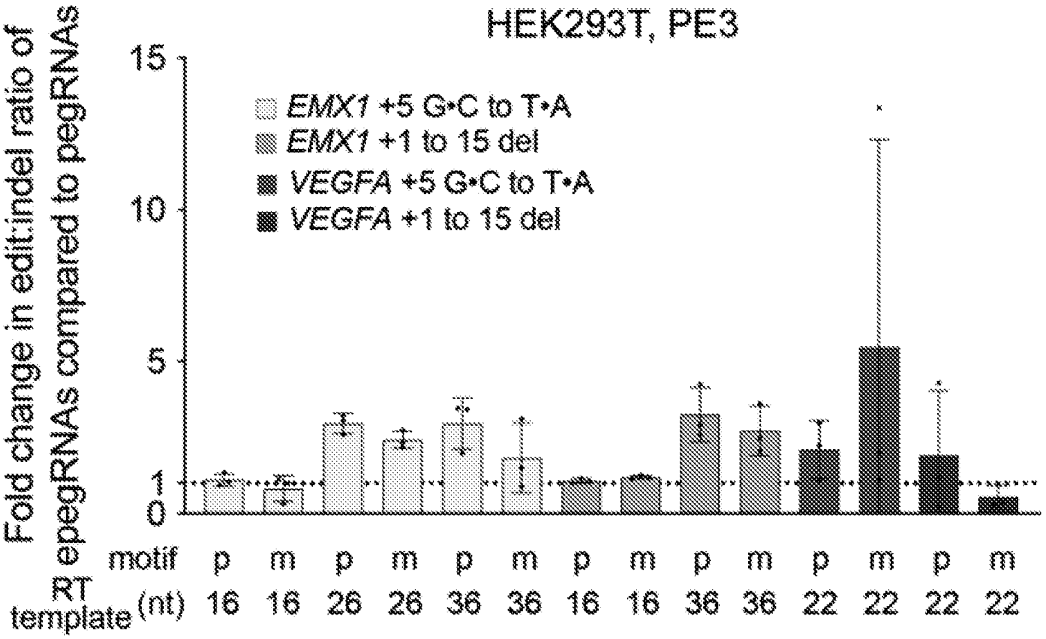

HEK293T, PE3
Fold change in edit:indel ratio of epegRNAs compared to pegRNAs
15
10
5
1
0
EMX1 +5 G•C to T•A
EMX1 +1 to 15 del
VEGFA +5 G•C to T•A
VEGFA +1 to 15 del
motif p m p m p m p m p m p m p m p m
RT template (nt) 16 16 26 26 36 36 16 16 36 36 22 22 22 22

FIG. 112C

site \ pegRNA	...	*HEK3* (PE3)				...	*EMX1* (PE3)				...	*FANCF* (PE3)			
		pt mtn, -	pt mtn, m	del, -	del, p		pt mtn, -	pt mtn, p	del, -	del, m		pt mtn, -	pt mtn, m	del, -	del, m
On-target		21.5 (4.0)	44.1 (11.7)	57.9 (2.8)	72.7 (4.7)		30.5 (1.1)	34.5 (1.1)	28.6 (0.8)	47 (0.5)		40.7 (4.2)	48.8 (3.7)	38.9 (6.4)	40.6 (6.2)
Off-target 1		<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)		<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)		<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)
Off-target 2		<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)		<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)		<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)
Off-target 3		<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)		<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)		<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)
Off-target 4		<0.1 (<0.1)	<0.1 (0.1)	<0.1 (0.1)	<0.1 (0.1)		<0.1 (0.1)	<0.1 (0.1)	<0.1 (<0.1)	<0.1 (0.1)		<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)	<0.1 (<0.1)

FIG. 113

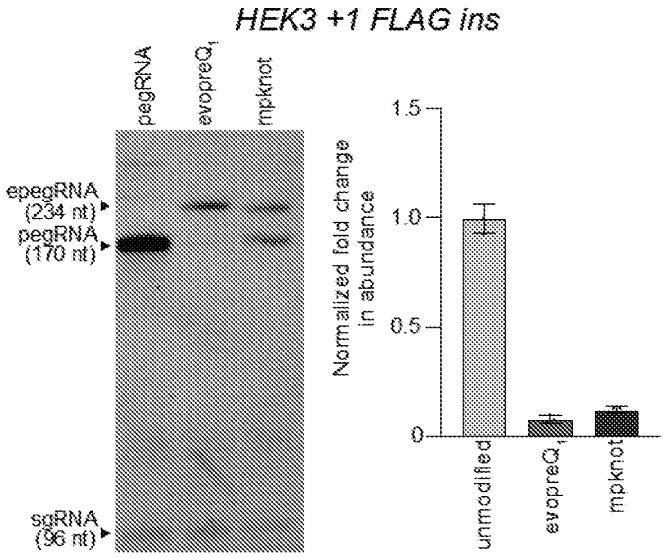


FIG. 114A

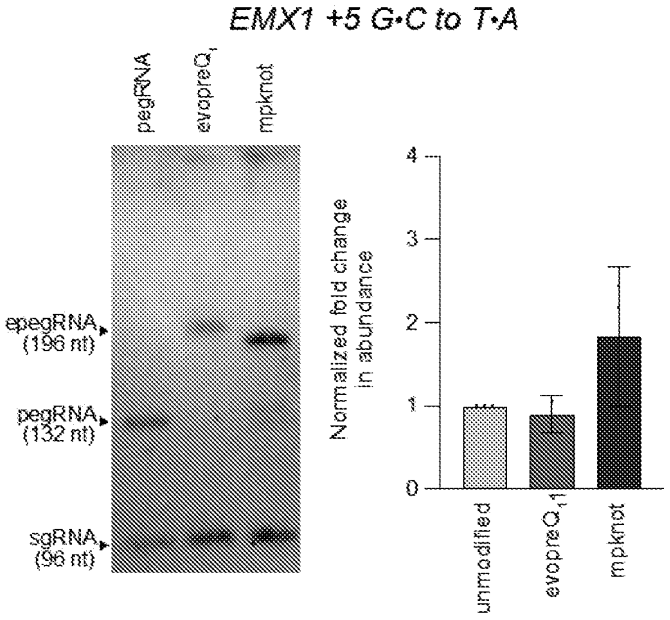


FIG. 114B

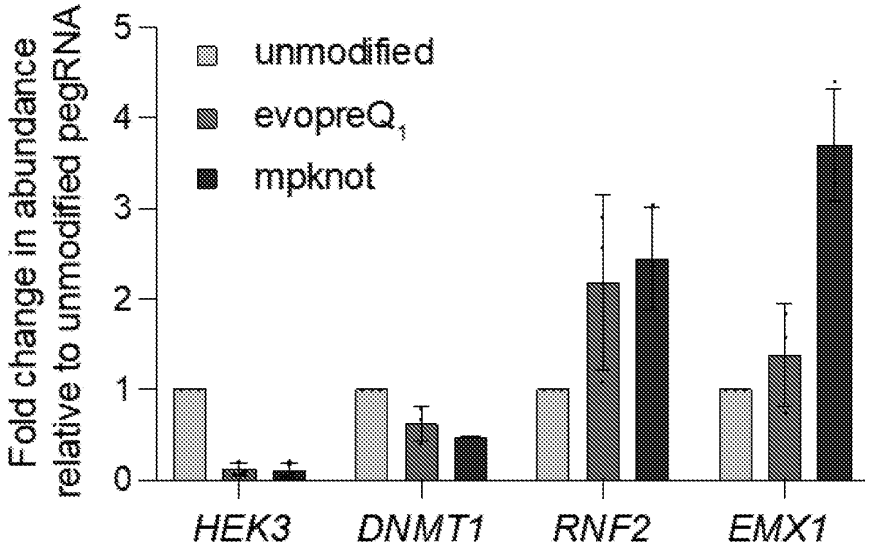


FIG. 114C

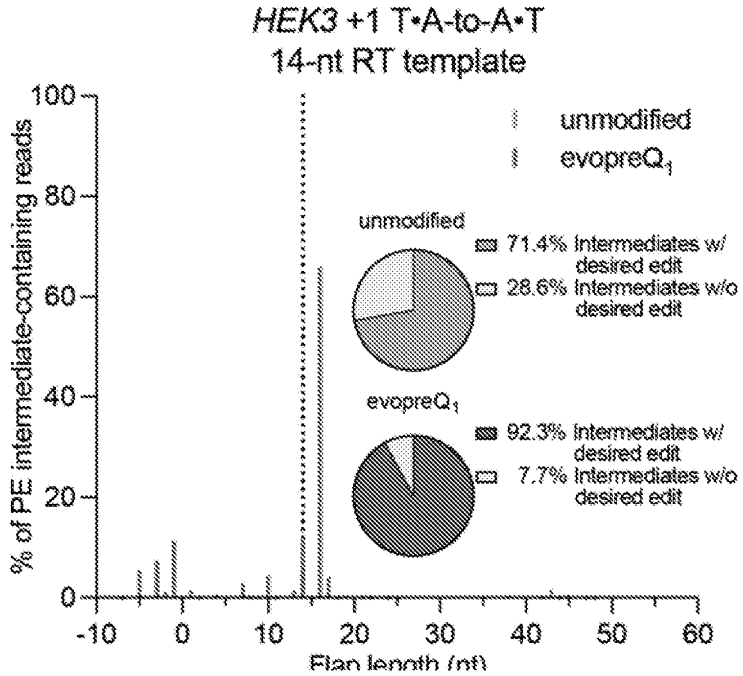


FIG. 115A

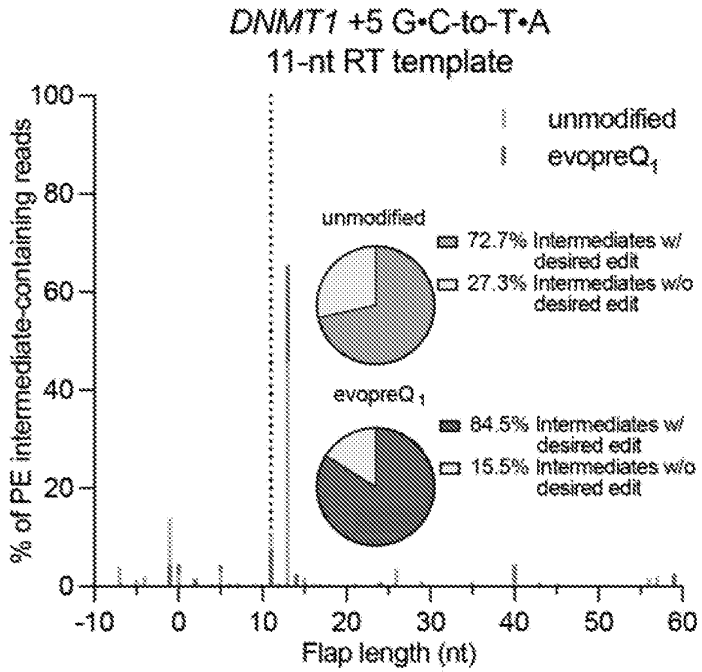


FIG. 115B

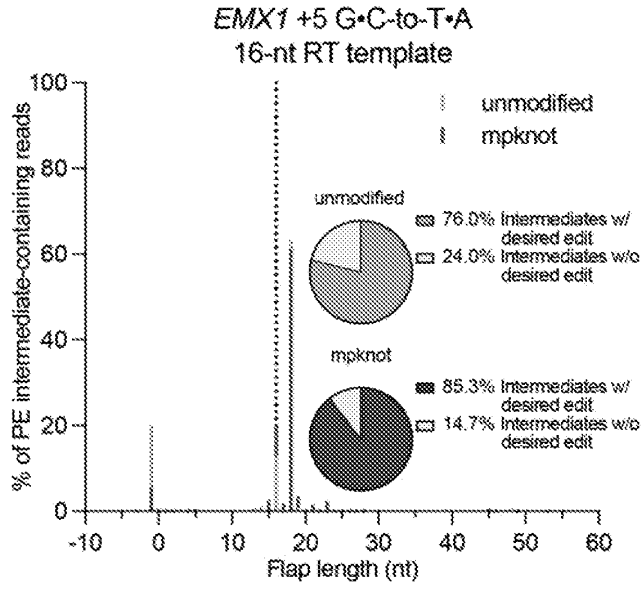


FIG. 115C

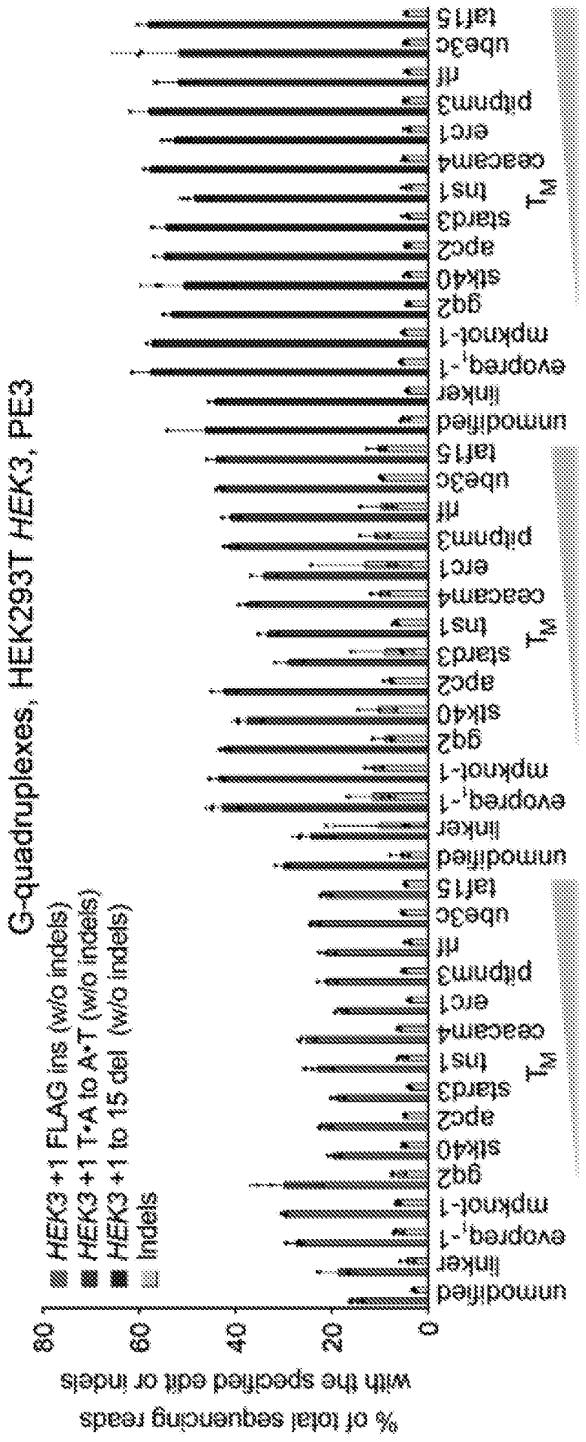

G-quadruplexes, HEK293T HEK3, PE3
HEK3 +1 FLAG ins (w/o indels)
HEK3 +1 T•A to A•T (w/o indels)
HEK3 +1 to 15 del (w/o indels)
Indels
% of total sequencing reads with the specified edit or indels
0
20
40
60
80
unmodified
linker
evopreq1-1
mpknot-1
gq2
stk40
apc2
stard3
tns1
ceacam4
erc1
pitpnm3
rtf
ube3c
tarl5
T_M

FIG. 116A

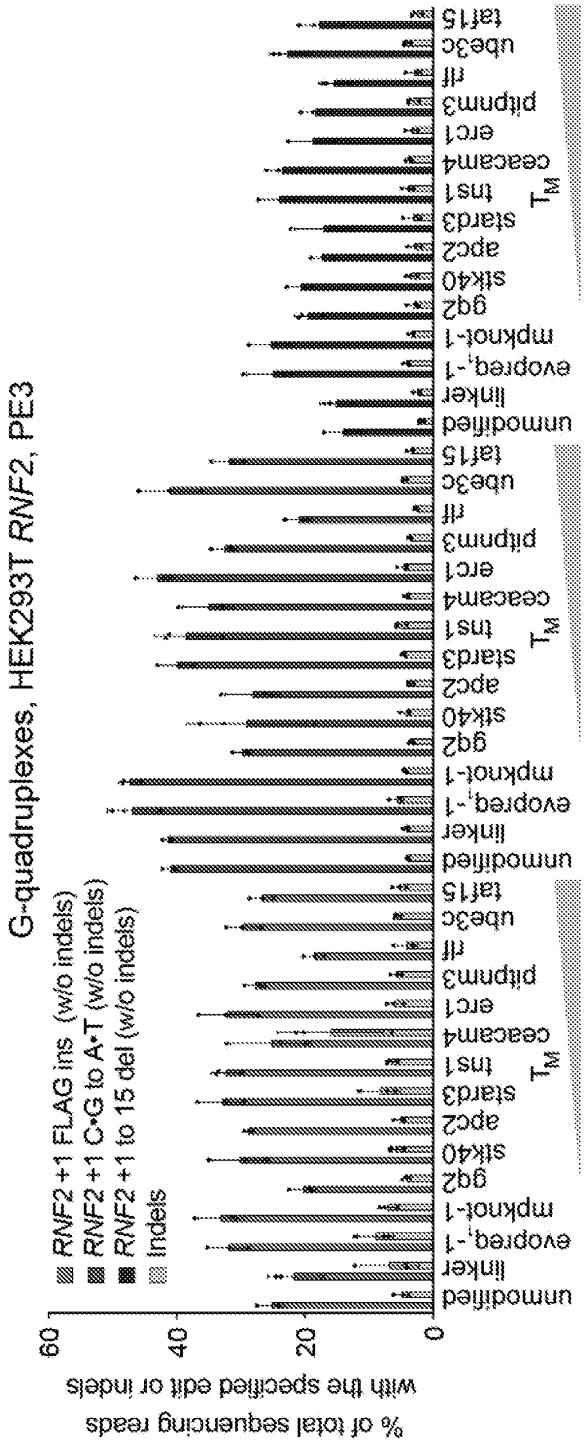

G-quadruplexes, HEK293T RNF2, PE3
RNF2 +1 FLAG ins (w/o indels)
RNF2 +1 C•G to A•T (w/o indels)
RNF2 +1 to 15 del (w/o indels)
Indels
% of total sequencing reads with the specified edit or indels
0
20
40
60
unmodified
linker
evopreq1-1
mpknot-1
gq2
stk40
apc2
stard3
tns1
ceacam4
erc1
pitpnm3
rff
ube3c
taf15
TM

FIG. 116B

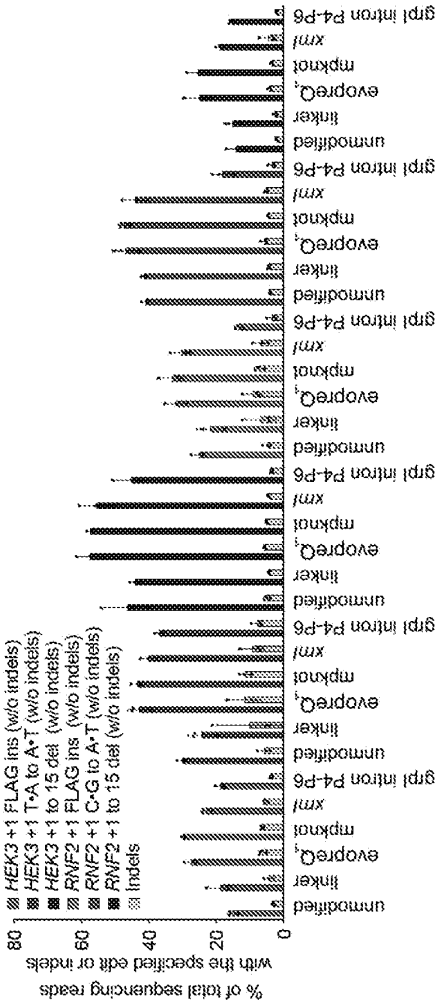


FIG. 116D

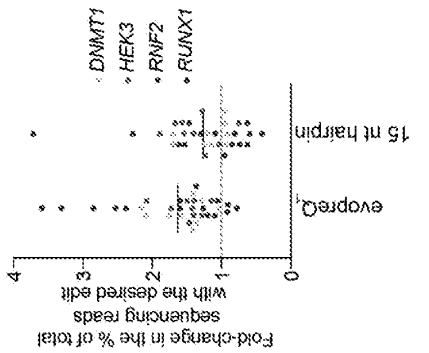


FIG. 116C

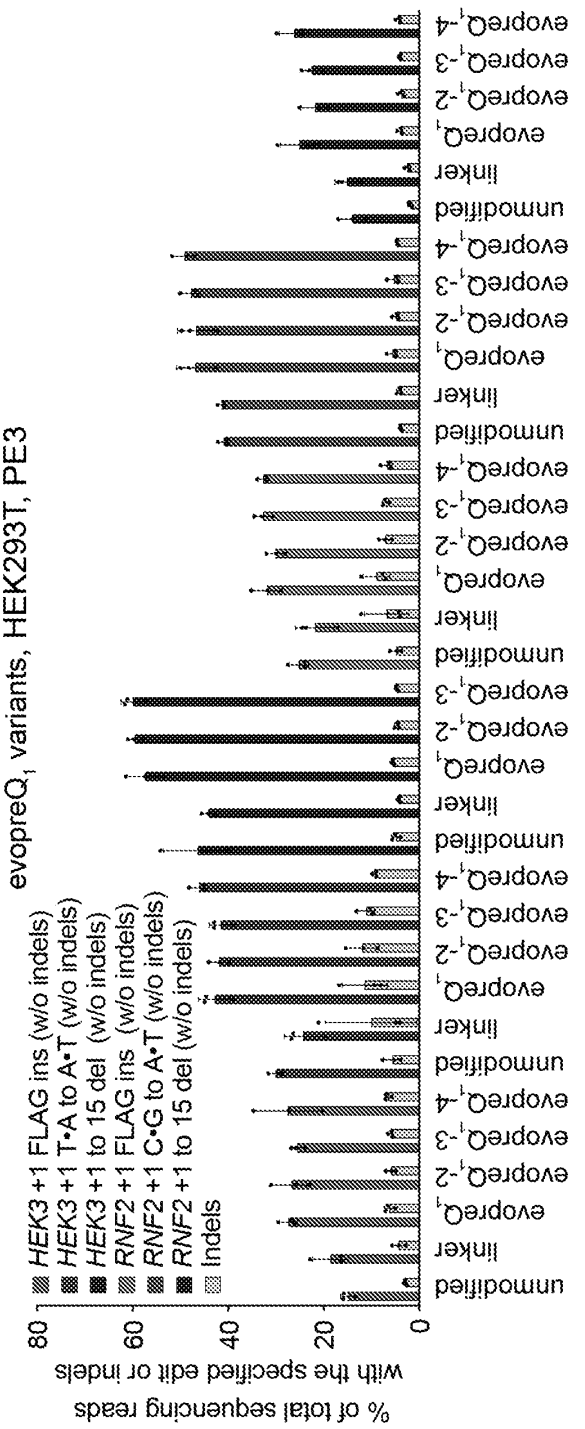

% of total sequencing reads with the specified edit or indels
0
20
40
60
80
HEK3 +1 FLAG ins (w/o indels)
HEK3 +1 T•A to A•T (w/o indels)
HEK3 +1 to 15 del (w/o indels)
RNF2 +1 FLAG ins (w/o indels)
RNF2 +1 C•G to A•T (w/o indels)
RNF2 +1 to 15 del (w/o indels)
Indels
evopreQ1 variants, HEK293T, PE3
unmodified
linker
evopreQ1
evopreQ1-2
evopreQ1-3
evopreQ1-4
unmodified
linker
evopreQ1
evopreQ1-2
evopreQ1-3
evopreQ1-4
unmodified
linker
evopreQ1
evopreQ1-2
evopreQ1-3
evopreQ1-4
unmodified
linker
evopreQ1
evopreQ1-2
evopreQ1-3
evopreQ1-4
unmodified
linker
evopreQ1
evopreQ1-2
evopreQ1-3
evopreQ1-4

FIG. 117A

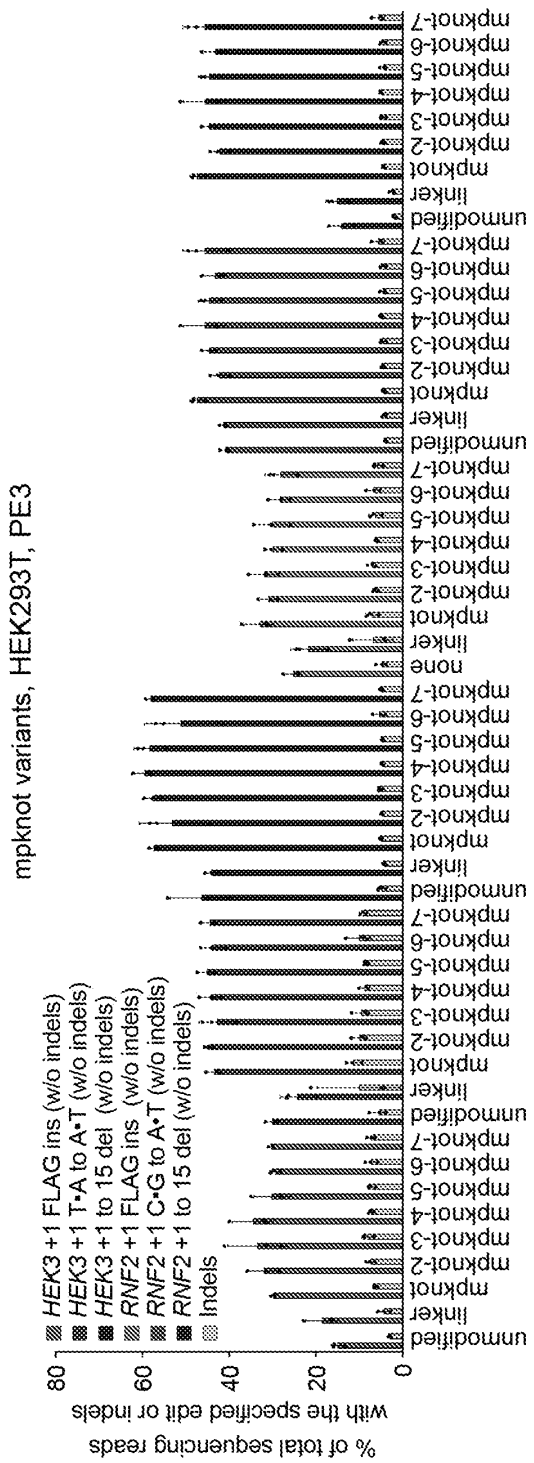

mpknot variants, HEK293T, PE3
% of total sequencing reads with the specified edit or indels
0
20
40
60
80
HEK3 +1 FLAG ins (w/o indels)
HEK3 +1 T•A to A•T (w/o indels)
HEK3 +1 to 15 del (w/o indels)
RNF2 +1 FLAG ins (w/o indels)
RNF2 +1 C•G to A•T (w/o indels)
RNF2 +1 to 15 del (w/o indels)
Indels
unmodified
linker
mpknot
mpknot-2
mpknot-3
mpknot-4
mpknot-5
mpknot-6
mpknot-7
none

FIG. 117B

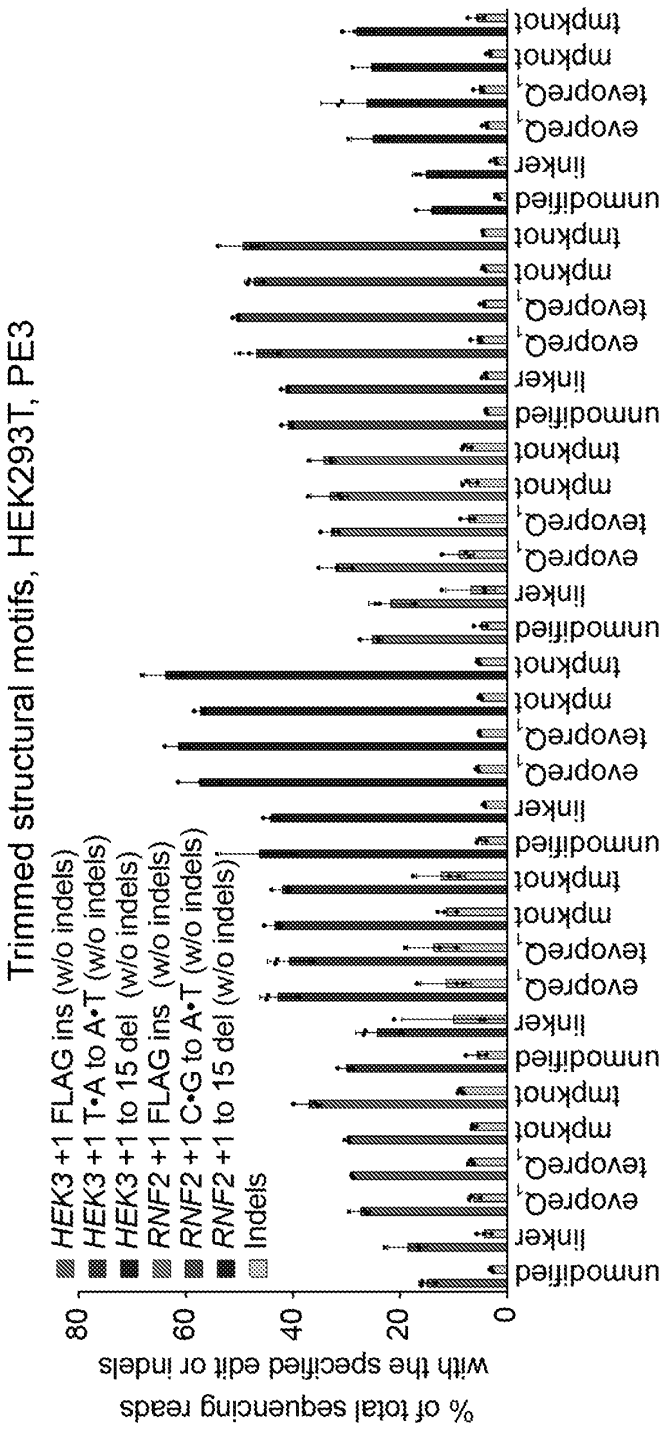

Trimmed structural motifs, HEK293T, PE3
% of total sequencing reads with the specified edit or indels
0
20
40
60
80
HEK3 +1 FLAG ins (w/o indels)
HEK3 +1 T•A to A•T (w/o indels)
HEK3 +1 to 15 del (w/o indels)
RNF2 +1 FLAG ins (w/o indels)
RNF2 +1 C•G to A•T (w/o indels)
RNF2 +1 to 15 del (w/o indels)
Indels
unmodified
linker
evopreQ1
tevopreQ1
mpknot
tmpknot
unmodified
linker
evopreQ1
tevopreQ1
mpknot
tmpknot
unmodified
linker
evopreQ1
tevopreQ1
mpknot
tmpknot
unmodified
linker
evopreQ1
tevopreQ1
mpknot
tmpknot
unmodified
linker
evopreQ1
tevopreQ1
mpknot
tmpknot

FIG. 117C

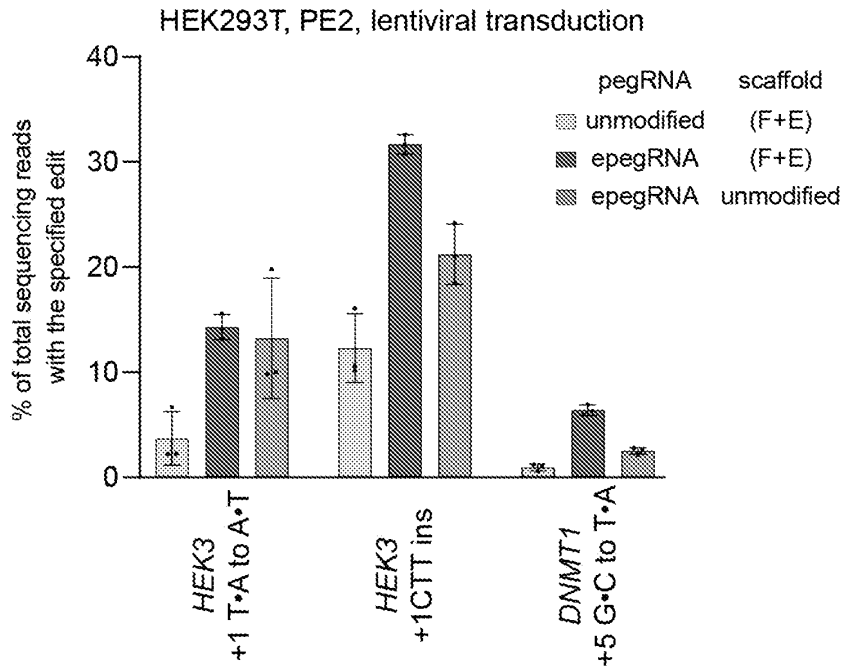

HEK293T, PE2, lentiviral transduction
% of total sequencing reads with the specified edit
40
30
20
10
0
pegRNA scaffold
unmodified (F+E)
epegRNA (F+E)
epegRNA unmodified
HEK3 +1 T•A to A•T
HEK3 +1CTT ins
DNMT1 +5 G•C to T•A

FIG. 118

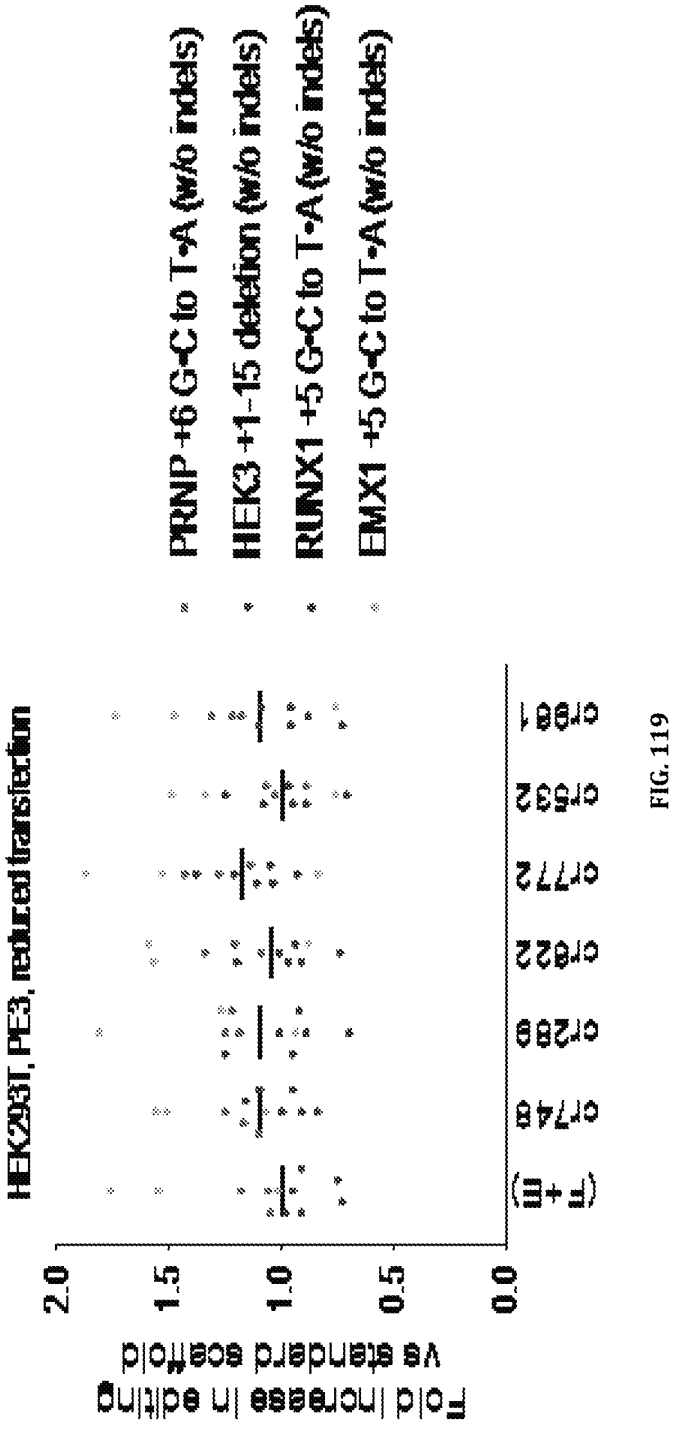

PRNP +6 G•C to T•A (w/o indels)
HEK3 +1-15 deletion (w/o indels)
RUNX1 +5 G•C to T•A (w/o indels)
EMX1 +5 G•C to T•A (w/o indels)
HEK293T, PE3, reduced transfection
Fold increase in editing vs standard scaffold
2.0
1.5
1.0
0.5
0.0
(F+E)
cr748
cr288
cr822
cr772
cr532
cr981

FIG. 119

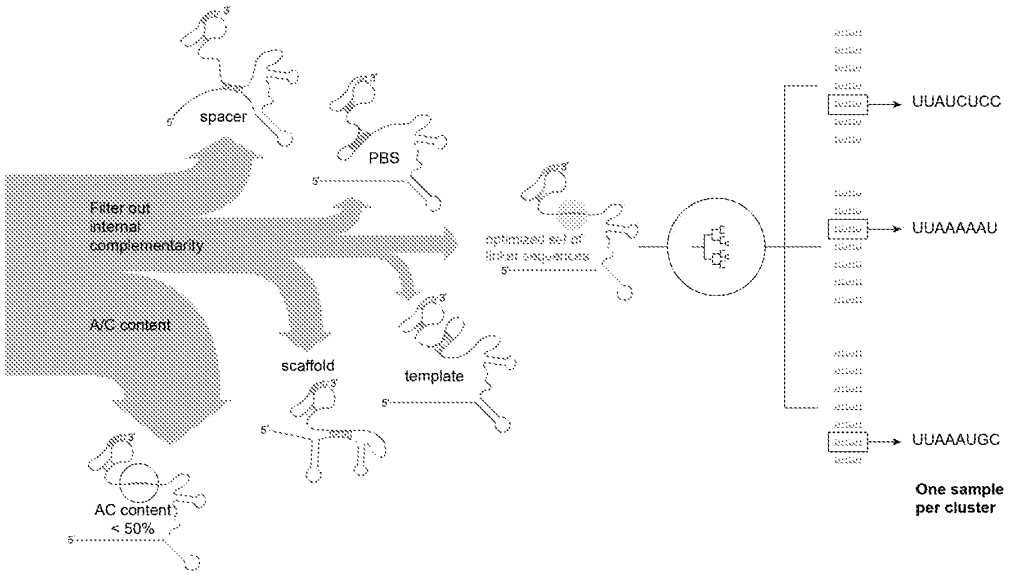


FIG. 120A

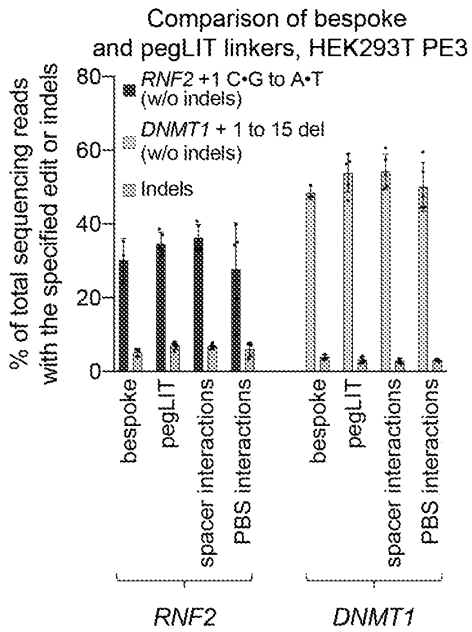


FIG. 120B

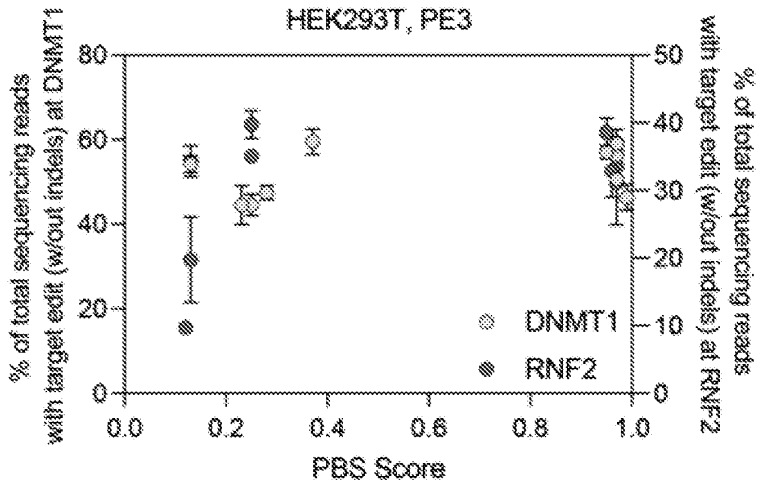


FIG. 120C

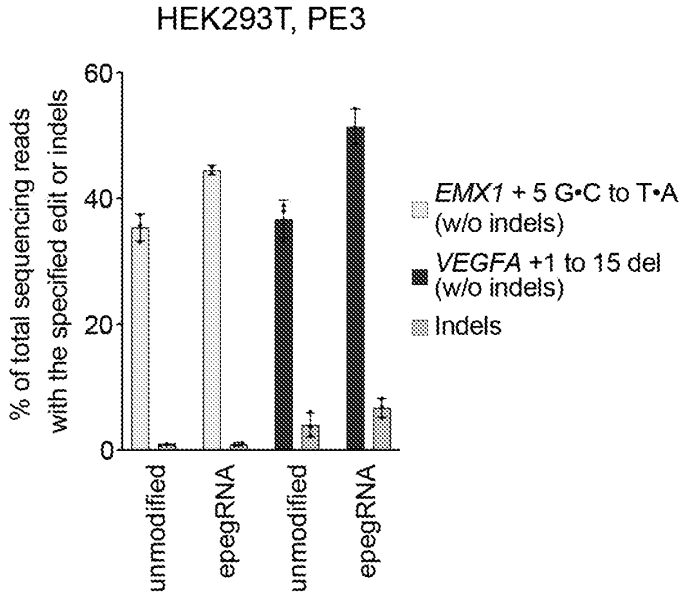


FIG. 120D

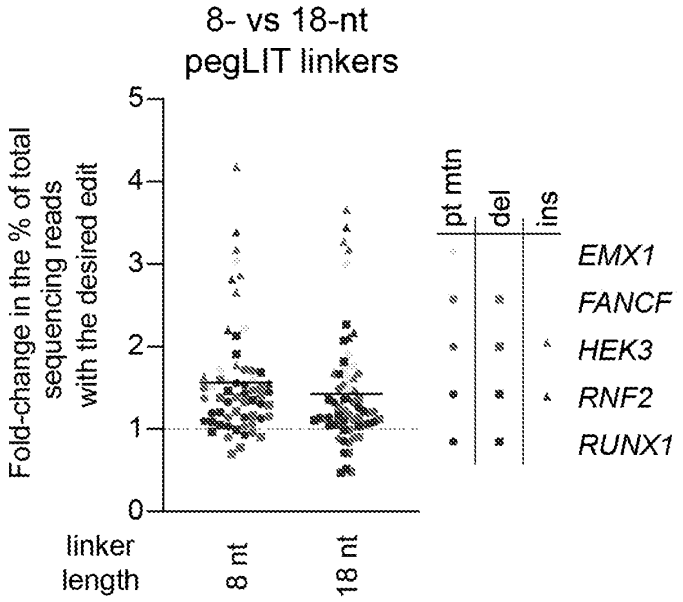

8- vs 18-nt
pegLIT linkers
Fold-change in the % of total sequencing reads with the desired edit
5
4
3
2
1
0
linker length
8 nt
18 nt
pt mtn
del
ins
EMX1
FANCF
HEK3
RNF2
RUNX1

FIG. 120E

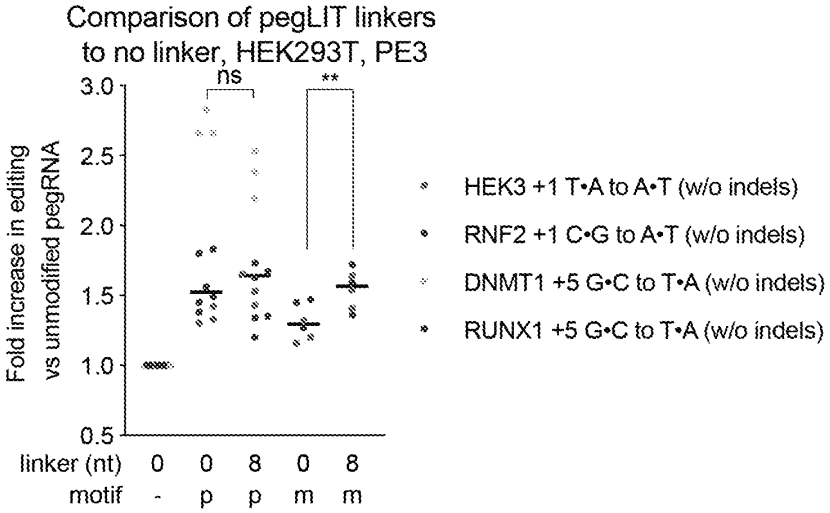


FIG. 120F

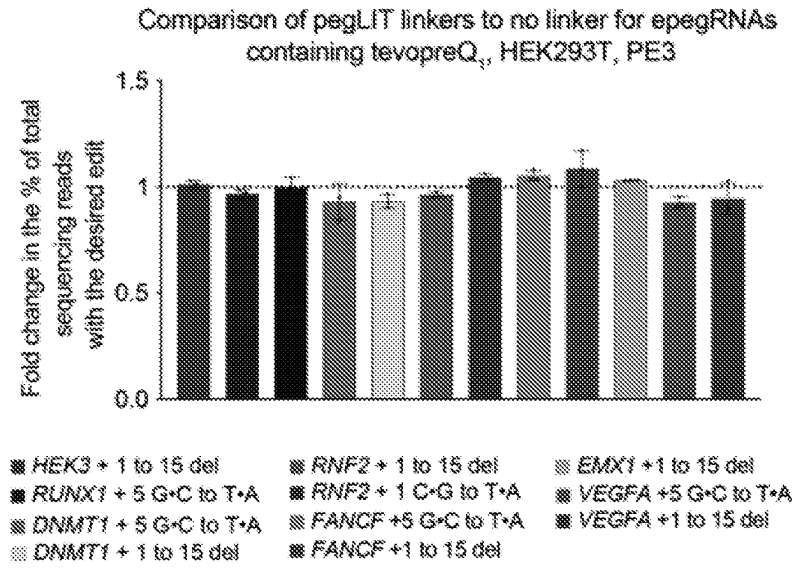


FIG. 120G

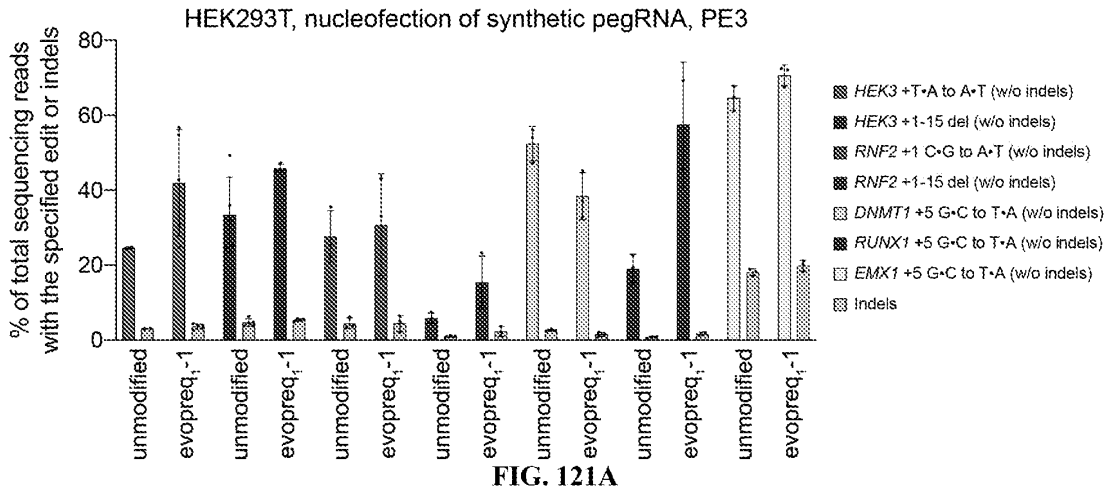


FIG. 121A

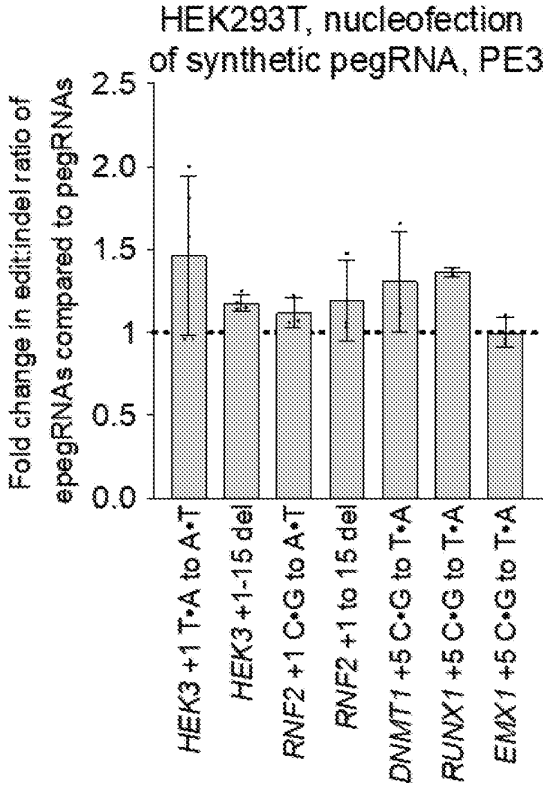


FIG. 121B

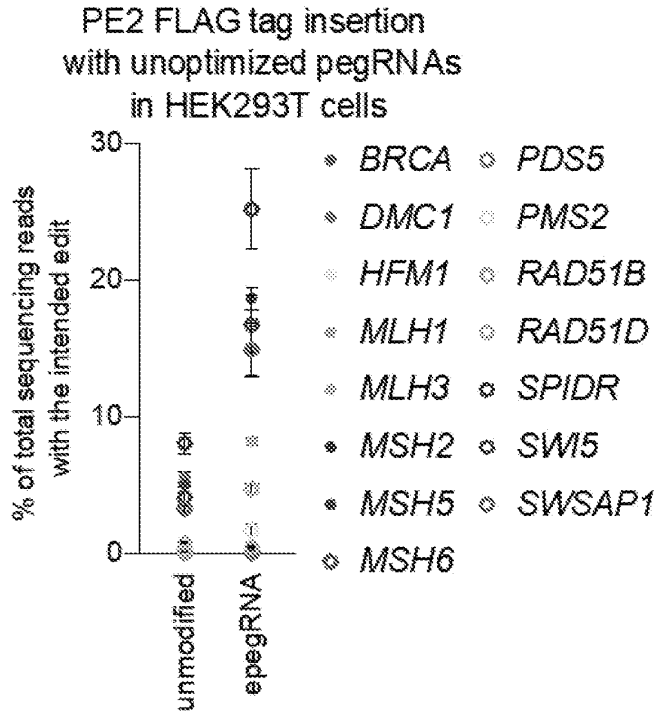


FIG. 122A

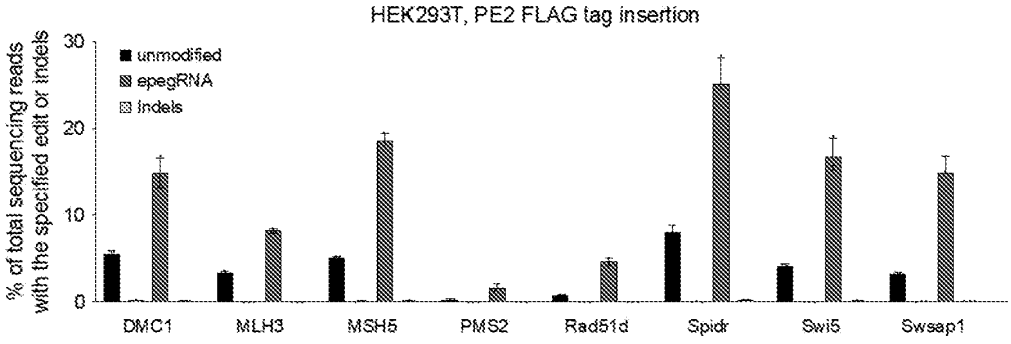


FIG. 122B

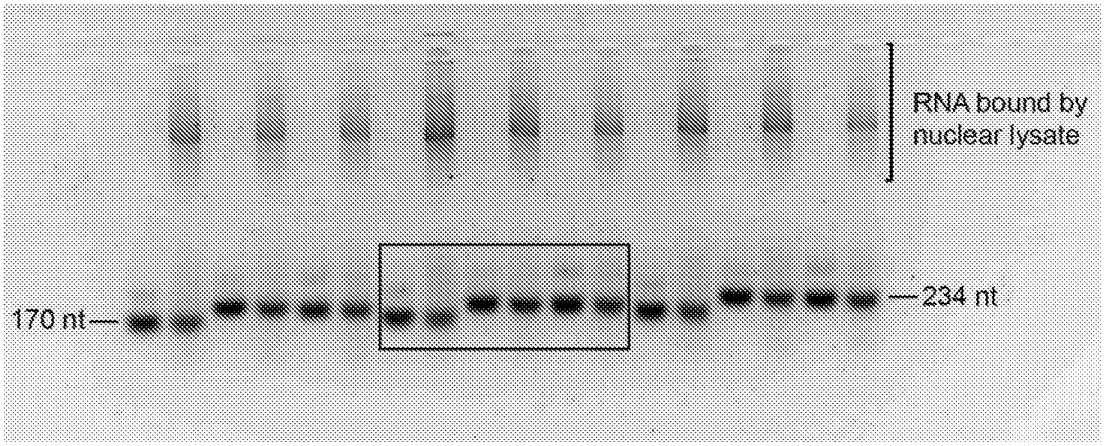


FIG. 123

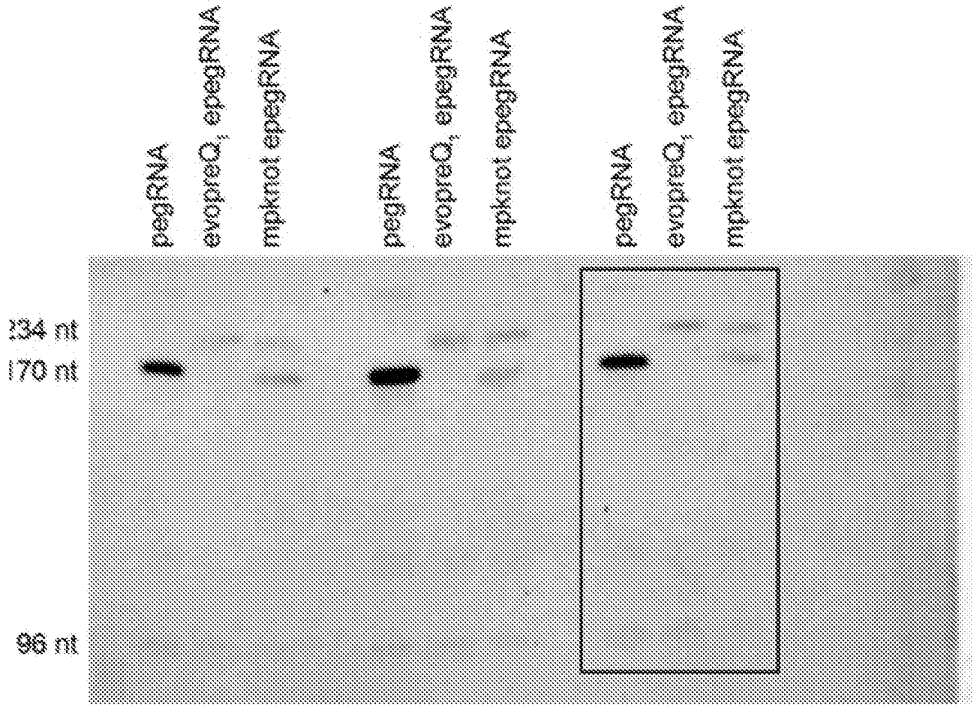


FIG. 124A

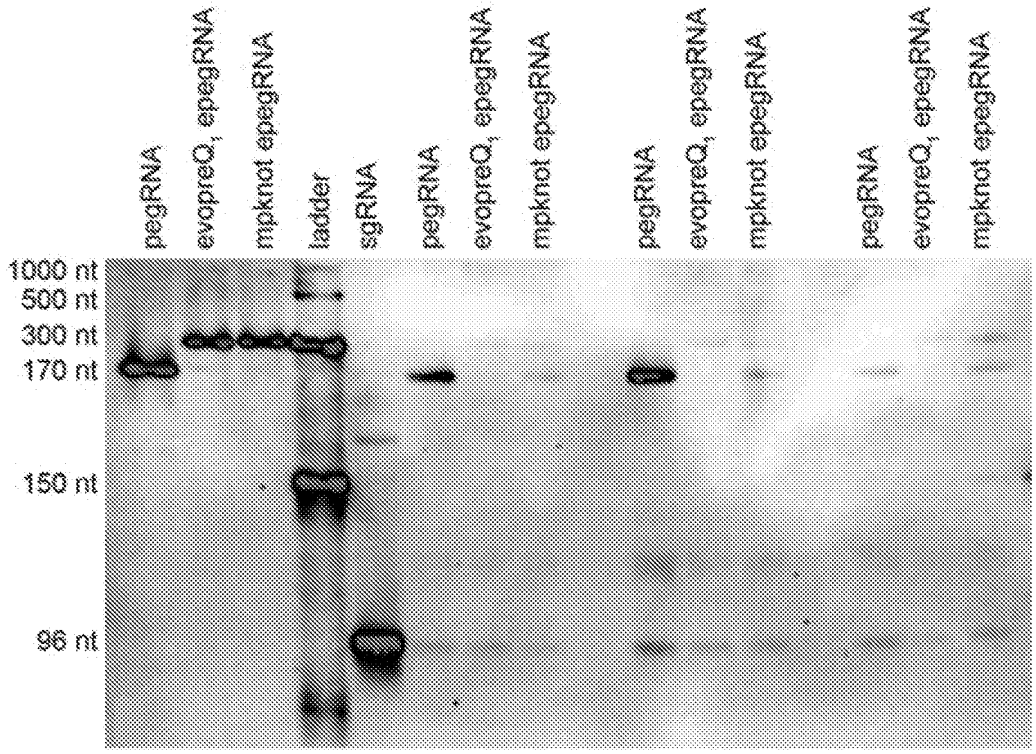

pegRNA
evopreQ1 epegRNA
mpknot epegRNA
ladder
sgRNA
pegRNA
evopreQ1 epegRNA
mpknot epegRNA
pegRNA
evopreQ1 epegRNA
mpknot epegRNA
pegRNA
evopreQ1 epegRNA
mpknot epegRNA
1000 nt
500 nt
300 nt
170 nt
150 nt
96 nt

FIG. 124B

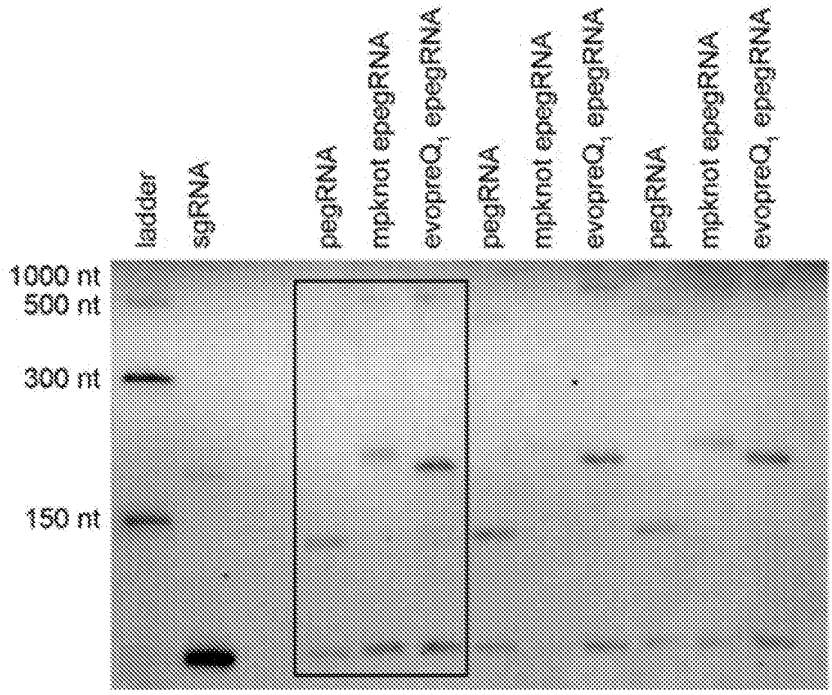

ladder
sgRNA
pegRNA
mpknot epegRNA
evopreQ1 epegRNA
pegRNA
mpknot epegRNA
evopreQ1 epegRNA
pegRNA
mpknot epegRNA
evopreQ1 epegRNA
1000 nt
500 nt
300 nt
150 nt

FIG. 124C

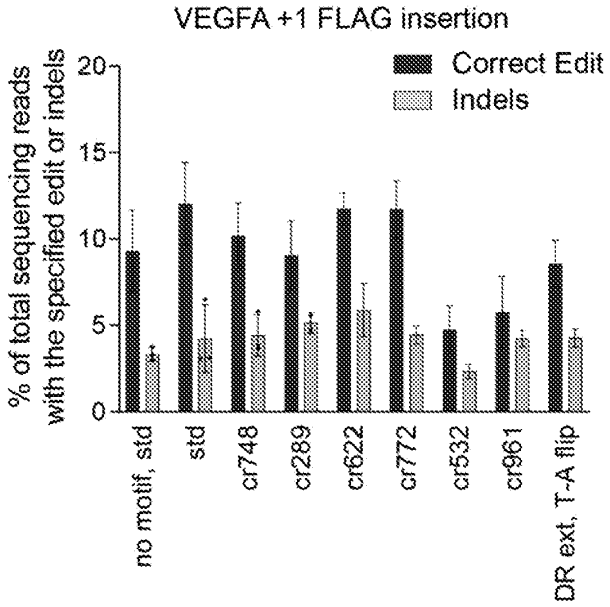


FIG. 125A

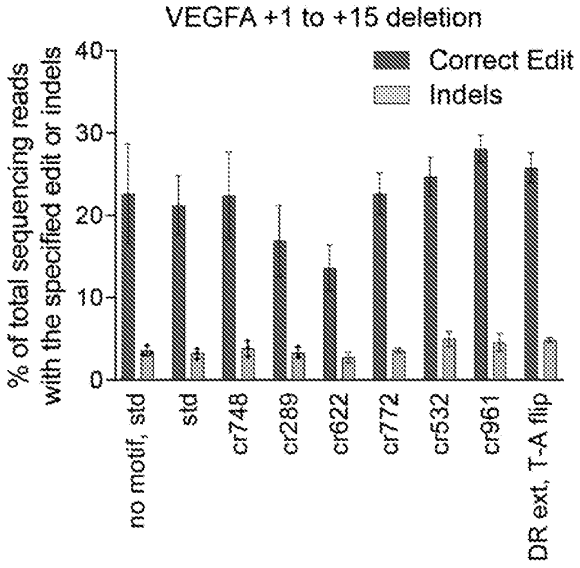


FIG. 125B

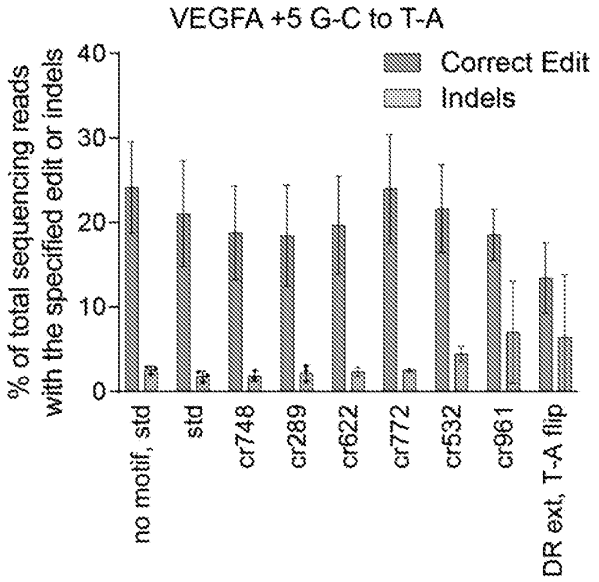


FIG. 125C

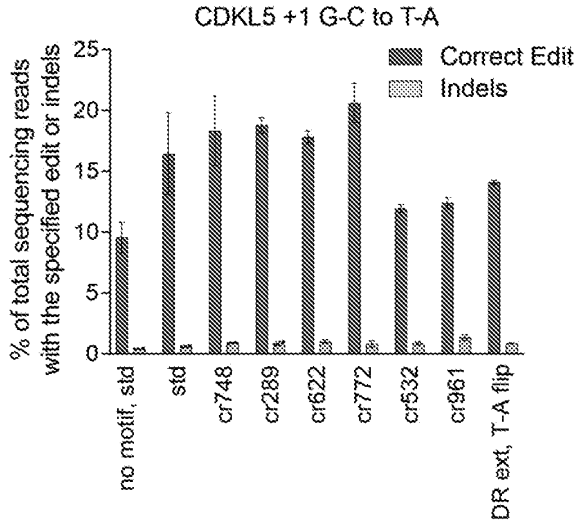


FIG. 125D

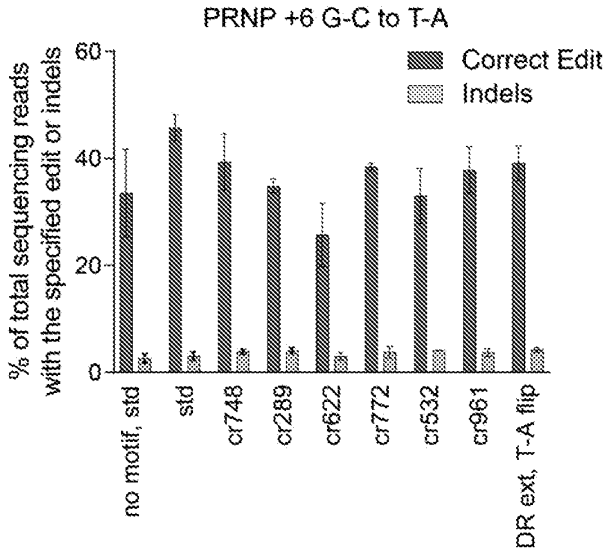

PRNP +6 G-C to T-A
Correct Edit
Indels
% of total sequencing reads with the specified edit or indels
60
40
20
0
no motif, std
std
cr748
cr289
cr622
cr772
cr532
cr961
DR ext, T-A flip

FIG. 125E

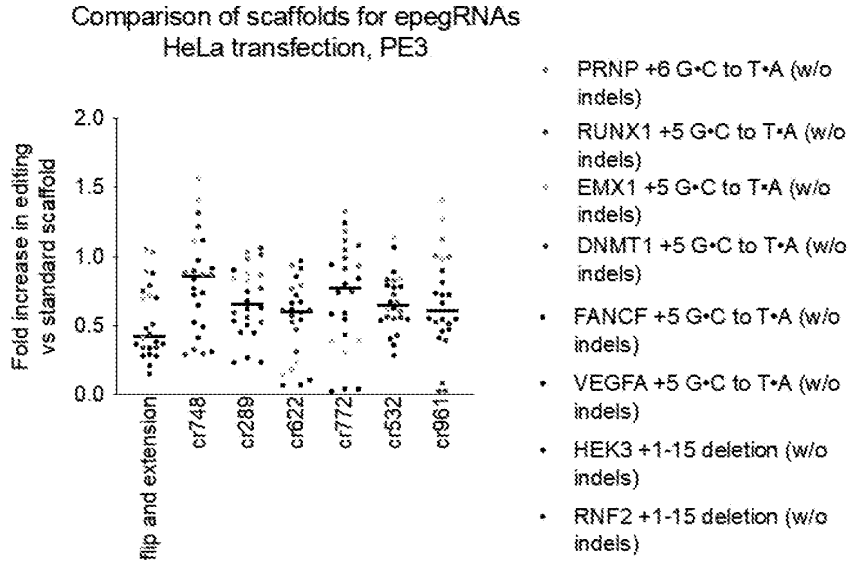


FIG. 126A

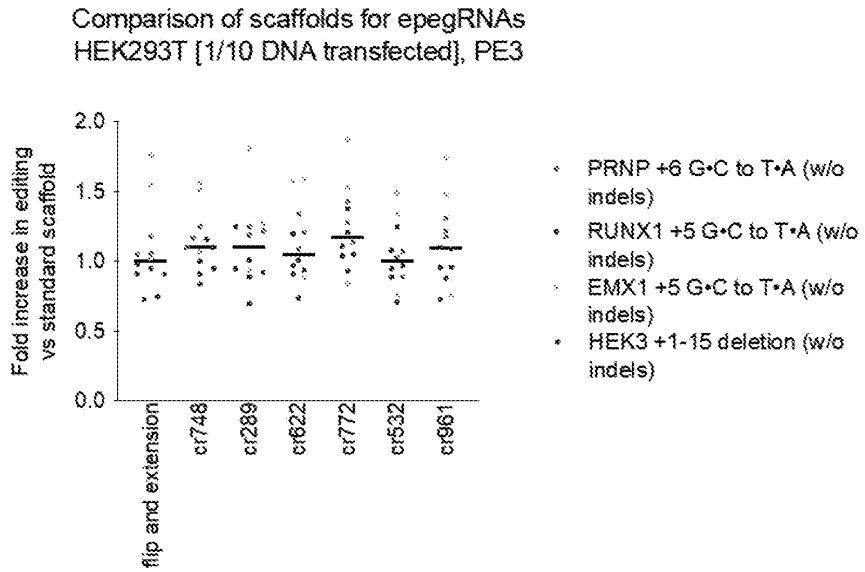


FIG. 126B

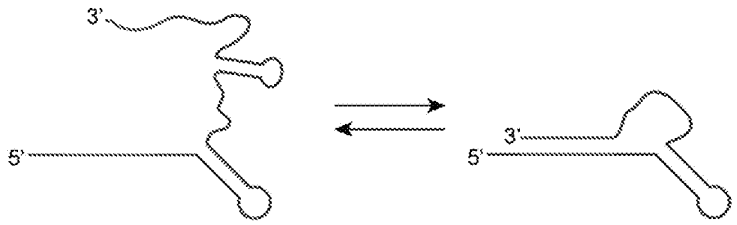


FIG. 127A

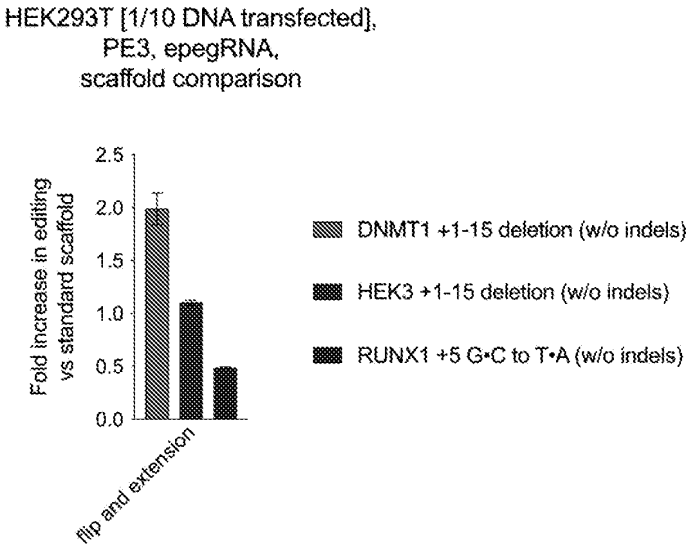


FIG. 127B

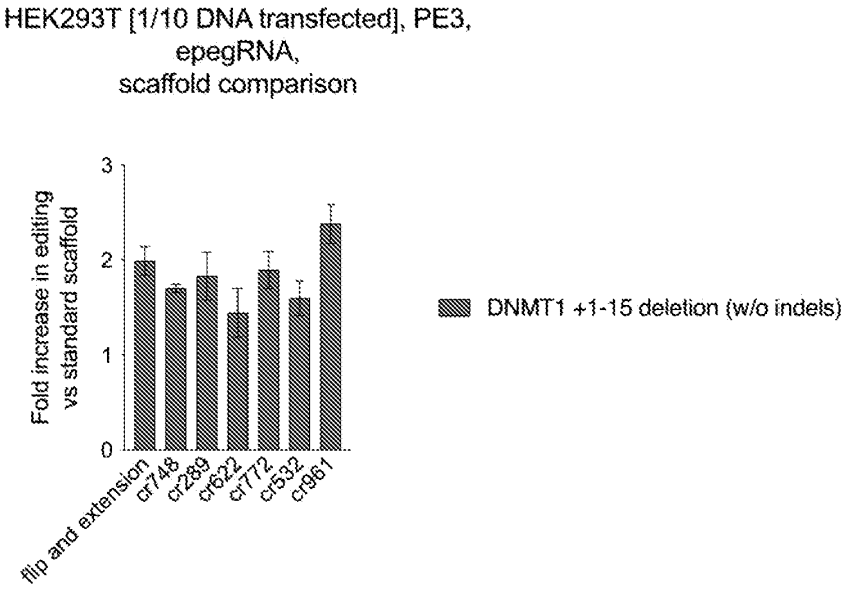

HEK293T [1/10 DNA transfected], PE3,
epegRNA,
scaffold comparison
Fold increase in editing
vs standard scaffold
3
2
1
0
flip and extension
cr748
cr289
cr622
cr772
cr532
cr961
DNMT1 +1-15 deletion (w/o indels)

FIG. 127C

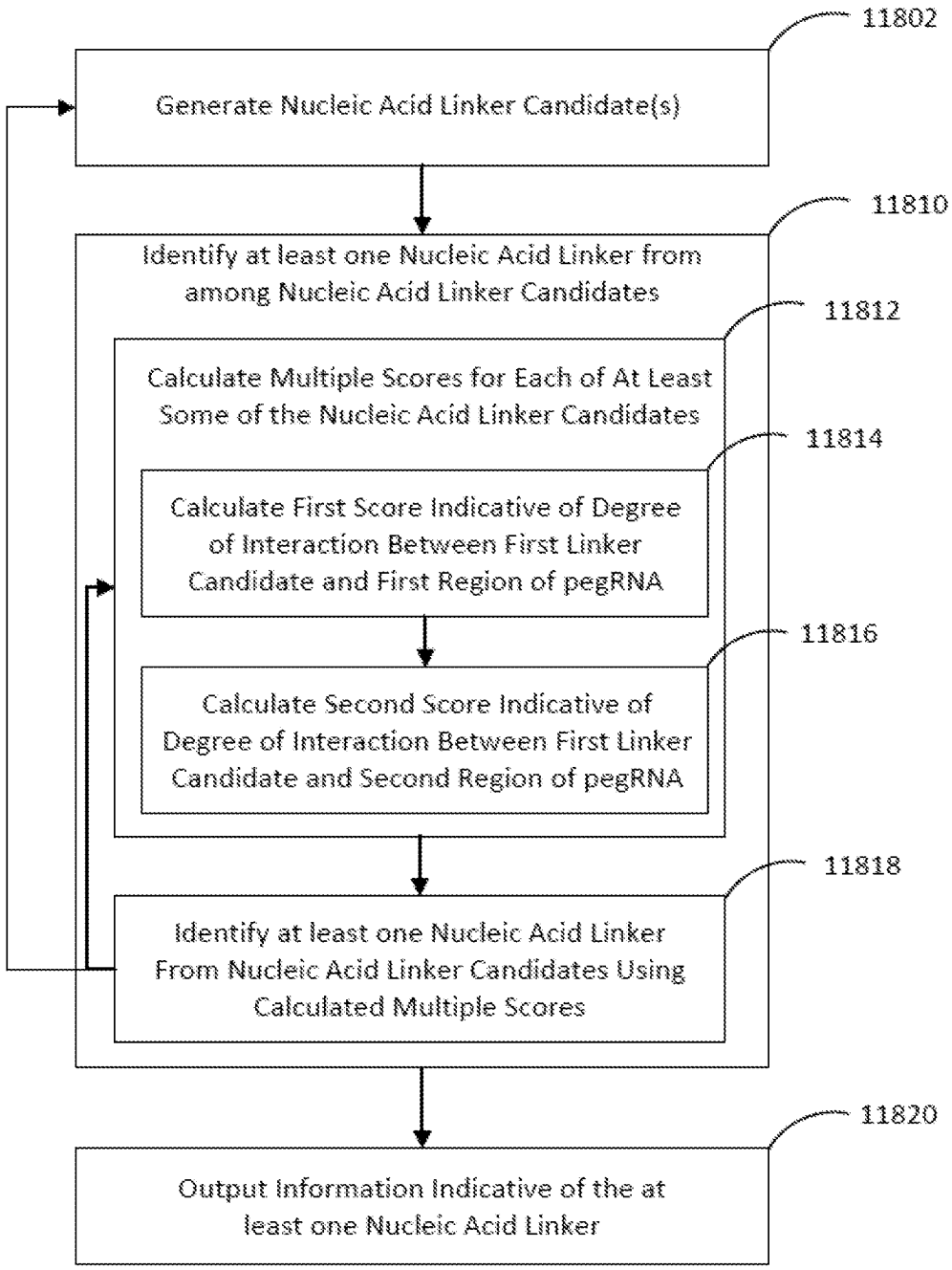

11802
Generate Nucleic Acid Linker Candidate(s)
11810
Identify at least one Nucleic Acid Linker from among Nucleic Acid Linker Candidates
11812
Calculate Multiple Scores for Each of At Least Some of the Nucleic Acid Linker Candidates
11814
Calculate First Score Indicative of Degree of Interaction Between First Linker Candidate and First Region of pegRNA
11816
Calculate Second Score Indicative of Degree of Interaction Between First Linker Candidate and Second Region of pegRNA
11818
Identify at least one Nucleic Acid Linker From Nucleic Acid Linker Candidates Using Calculated Multiple Scores
11820
Output Information Indicative of the at least one Nucleic Acid Linker

FIG. 128

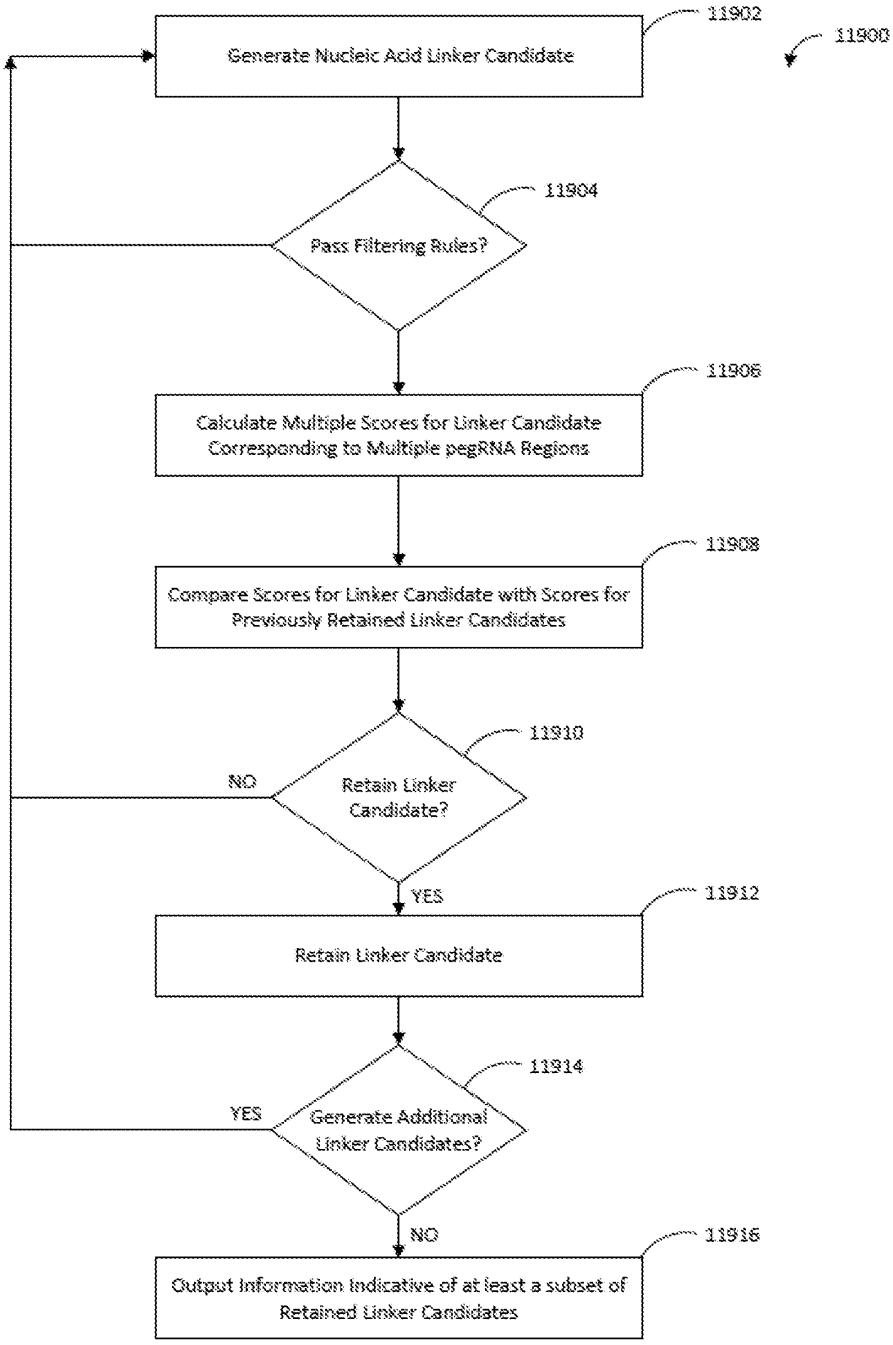

11900
11902
Generate Nucleic Acid Linker Candidate
11904
Pass Filtering Rules?
11906
Calculate Multiple Scores for Linker Candidate Corresponding to Multiple pegRNA Regions
11908
Compare Scores for Linker Candidate with Scores for Previously Retained Linker Candidates
11910
Retain Linker Candidate?
NO
YES
11912
Retain Linker Candidate
11914
Generate Additional Linker Candidates?
YES
NO
11916
Output Information Indicative of at least a subset of Retained Linker Candidates

FIG. 129

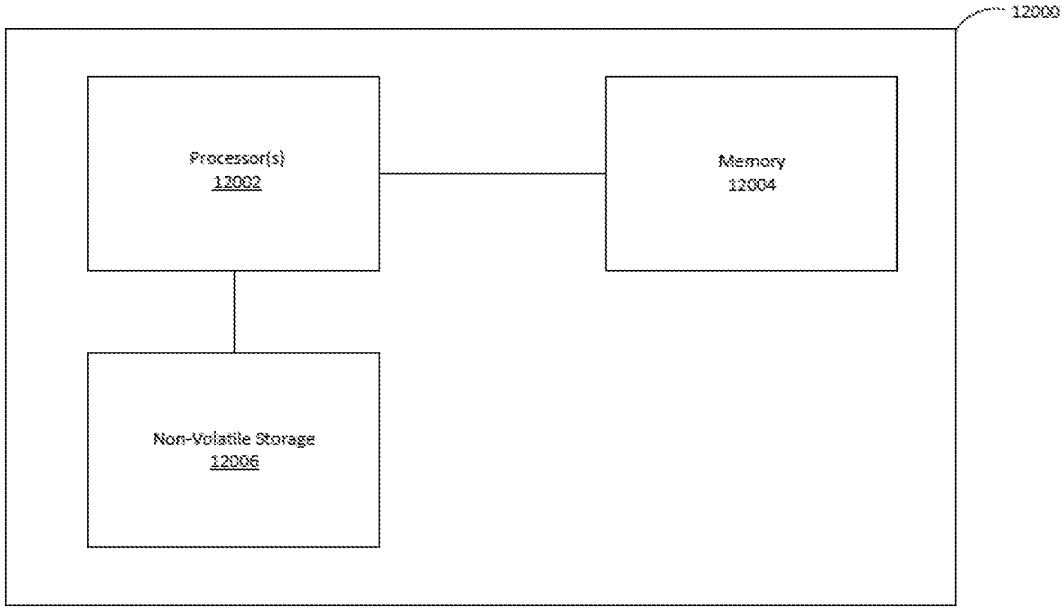

12000
Processor(s)
12002
Memory
12004
Non-Volatile Storage
12006

FIG. 130

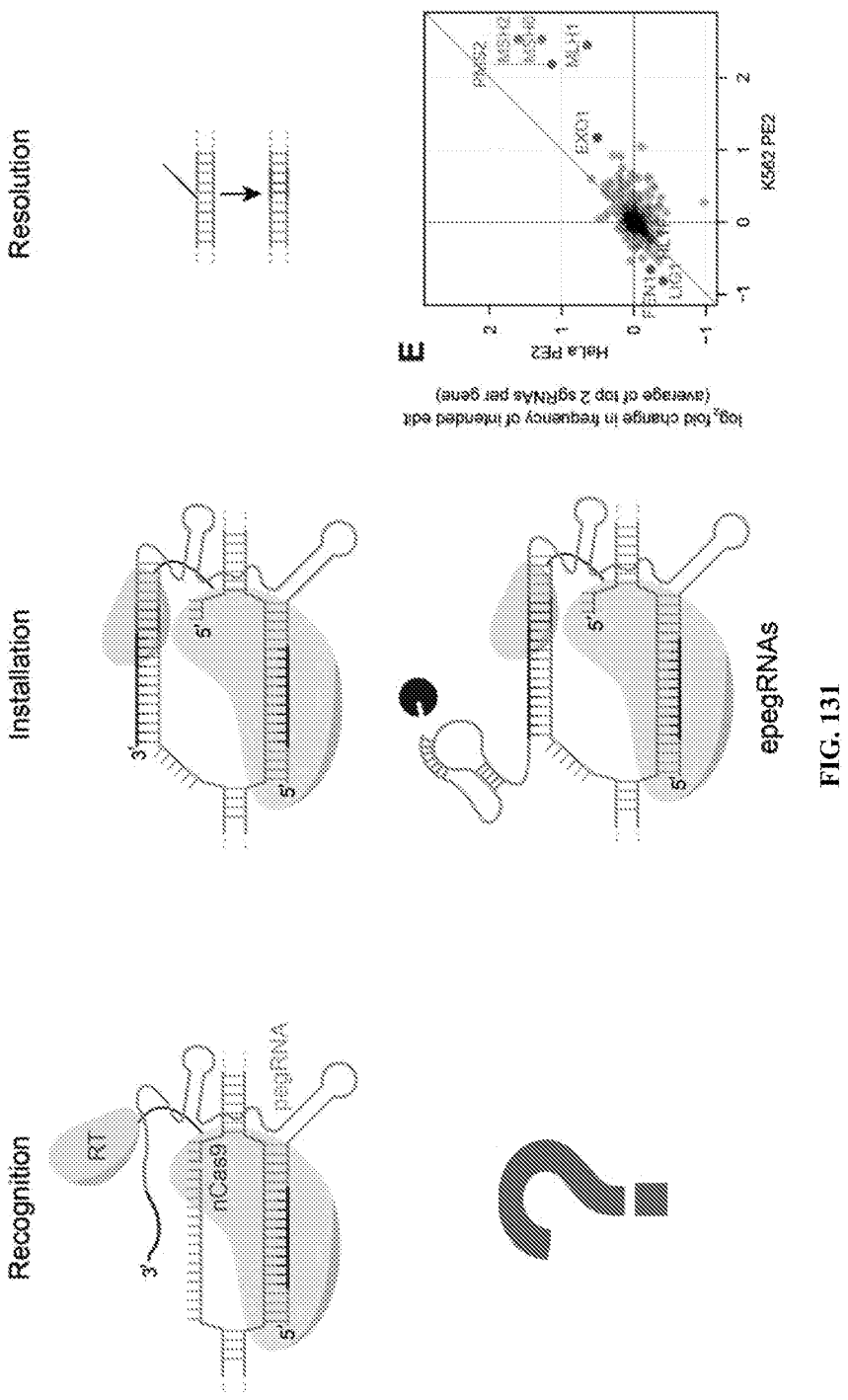

Resolution
Installation
Recognition
RT
nCas9
pegRNA
3'
5'
epegRNAs
E
HeLa PE2
K562 PE2
log2 fold change in frequency of intended edit
(average of top 2 sgRNAs per gene)

FIG. 131

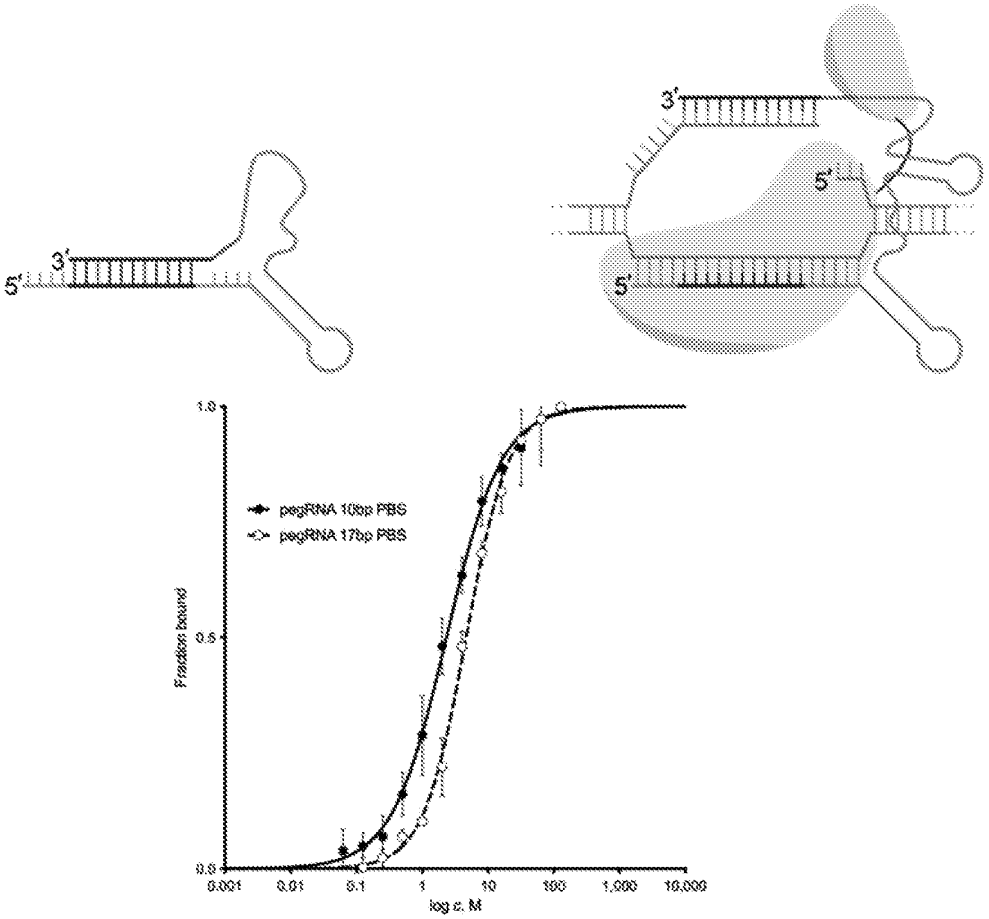

3'
5'
3'
5'
5'
1.0
0.5
0.0
Fraction bound
pegRNA 10bp PBS
pegRNA 17bp PBS
0.001
0.01
0.1
1
10
100
1,000
10,000
log c, M

FIG. 132

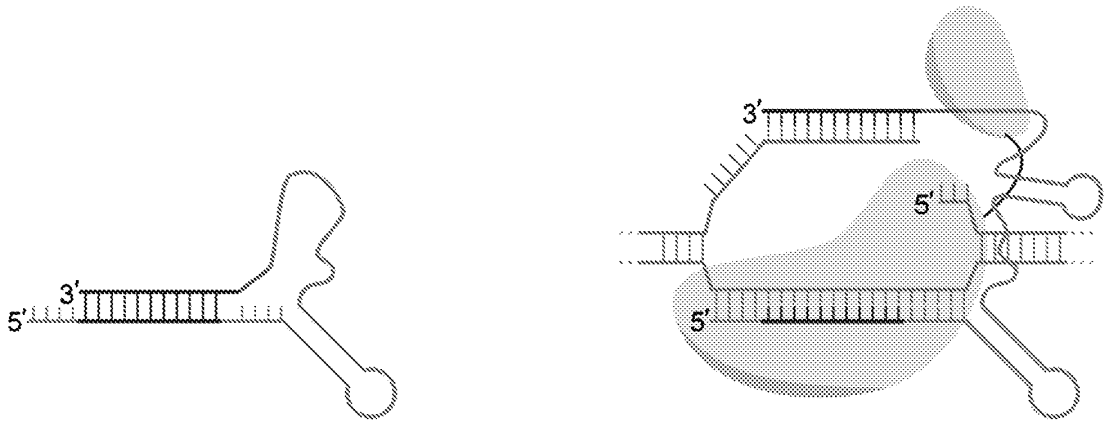


FIG. 133

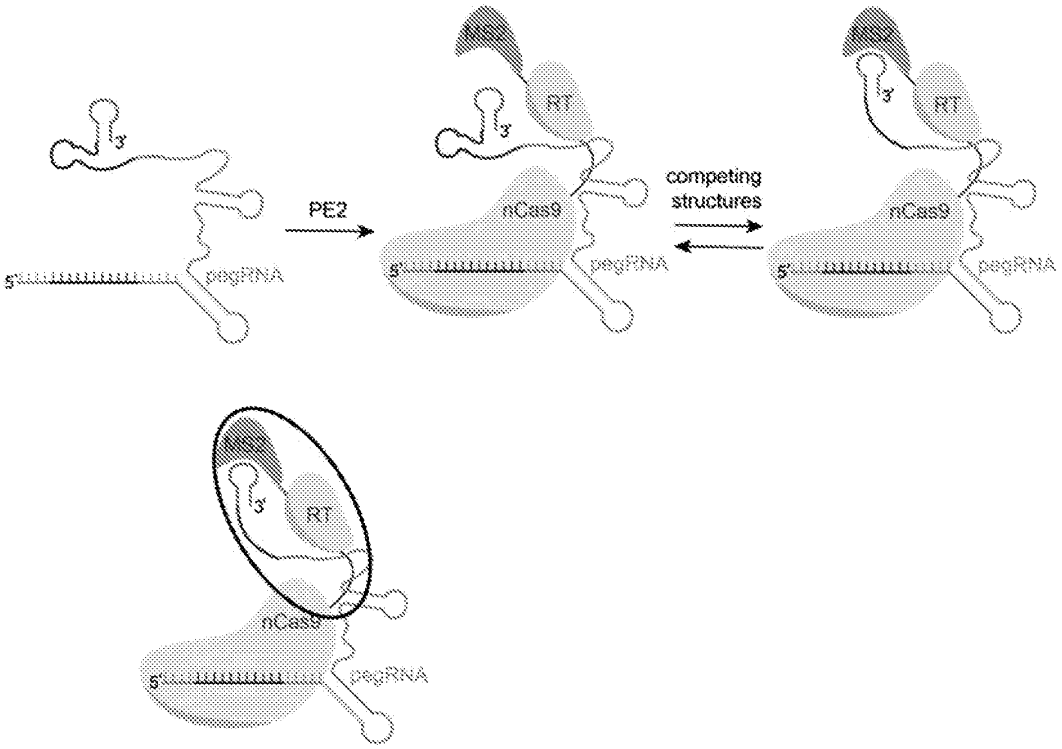
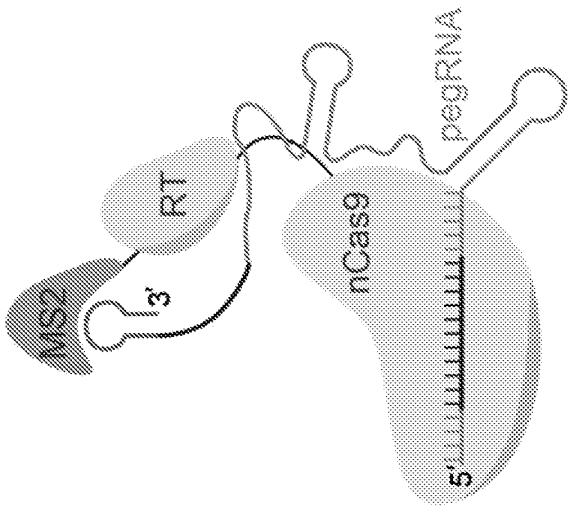
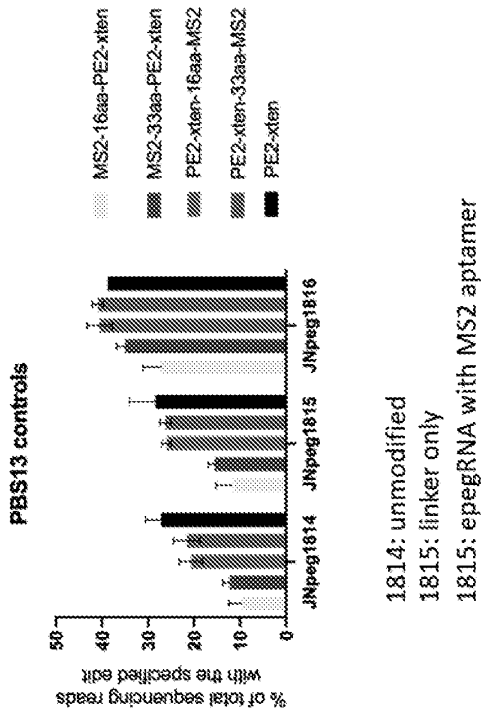


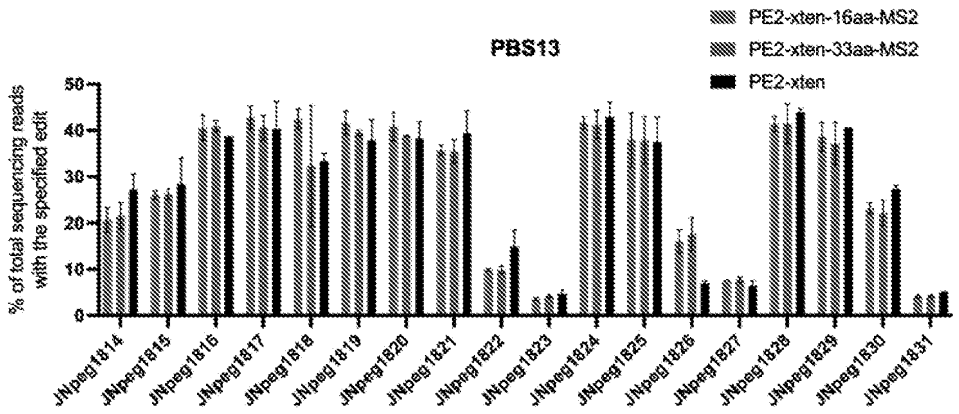
FIG. 134





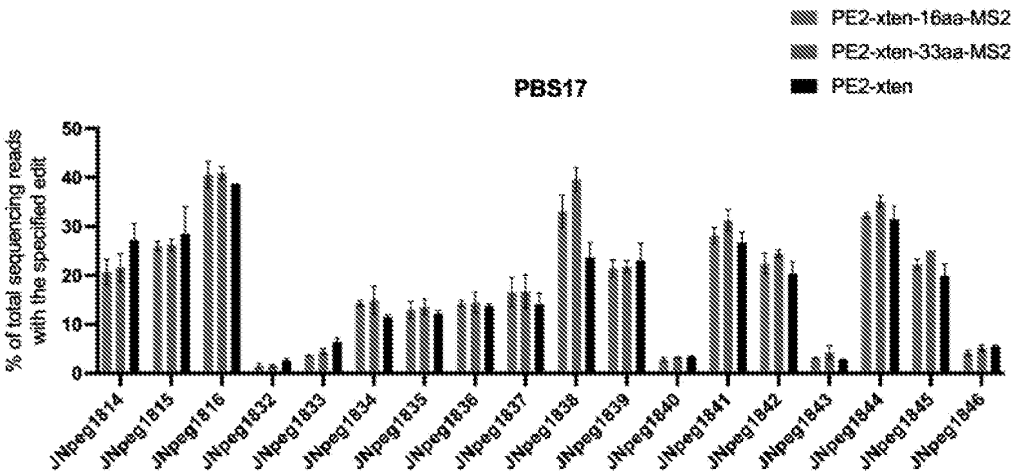
1814: unmodified
1815: linker only
1815: epegRNA with MS2 aptamer

FIG. 135



PBS = 13 all
1814: unmodified
1815: linker only
1816: epegRNA with MS2 aptamer
1817-1819: no toehold, MS2 aptamer
1820-1831: toeholds of various strengths, MS2 aptamer

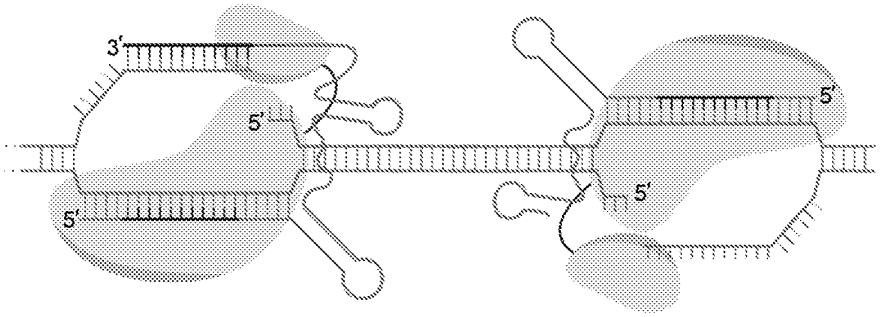
FIG. 136



1814: unmodified PBS 13
1815: linker only PBS 13
1816: epegRNA with MS2 aptamer PBS 13
PBS 17 all
1832-1833: unmodified; linker alone
1834-1837: no toehold, MS2 aptamer
1838-1846: toeholds of various strengths, MS2 aptamer

FIG. 137

PE3, pegRNA, ngRNA



PE3, sgRNA, pegRNA

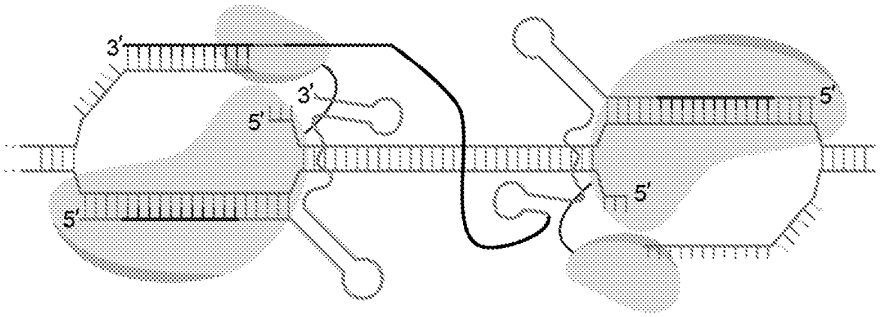


FIG. 138

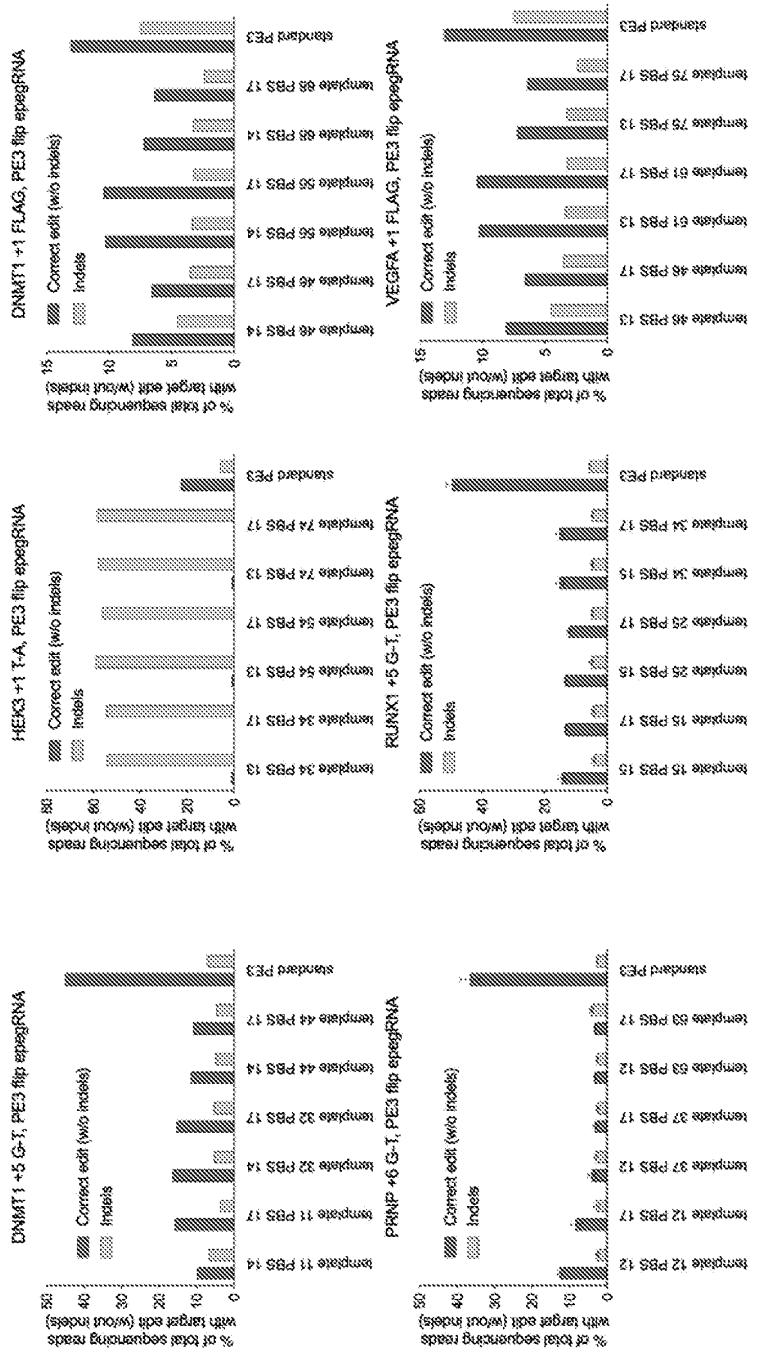

DNMT1 +1 FLAG, PE3 flip epegRNA
VEGFA +1 FLAG, PE3 flip epegRNA
HEK3 +1 T-A, PE3 flip epegRNA
RUNX1 +5 G-T, PE3 flip epegRNA
DNMT1 +5 G-T, PE3 flip epegRNA
PRNP +6 G-T, PE3 flip epegRNA
% of total sequencing reads with target edit (w/out indels)
Correct edit (w/o indels)
Indels
template 46 PBS 14
template 46 PBS 17
template 56 PBS 14
template 56 PBS 17
template 68 PBS 14
template 68 PBS 17
standard PE3
template 46 PBS 13
template 46 PBS 17
template 61 PBS 13
template 61 PBS 17
template 76 PBS 13
template 76 PBS 17
template 34 PBS 13
template 34 PBS 17
template 54 PBS 13
template 54 PBS 17
template 74 PBS 13
template 74 PBS 17
template 15 PBS 13
template 15 PBS 17
template 25 PBS 15
template 25 PBS 17
template 34 PBS 15
template 34 PBS 17
template 11 PBS 14
template 11 PBS 17
template 32 PBS 14
template 32 PBS 17
template 44 PBS 14
template 44 PBS 17
template 12 PBS 12
template 12 PBS 17
template 37 PBS 12
template 37 PBS 17
template 53 PBS 12
template 53 PBS 17

FIG. 139

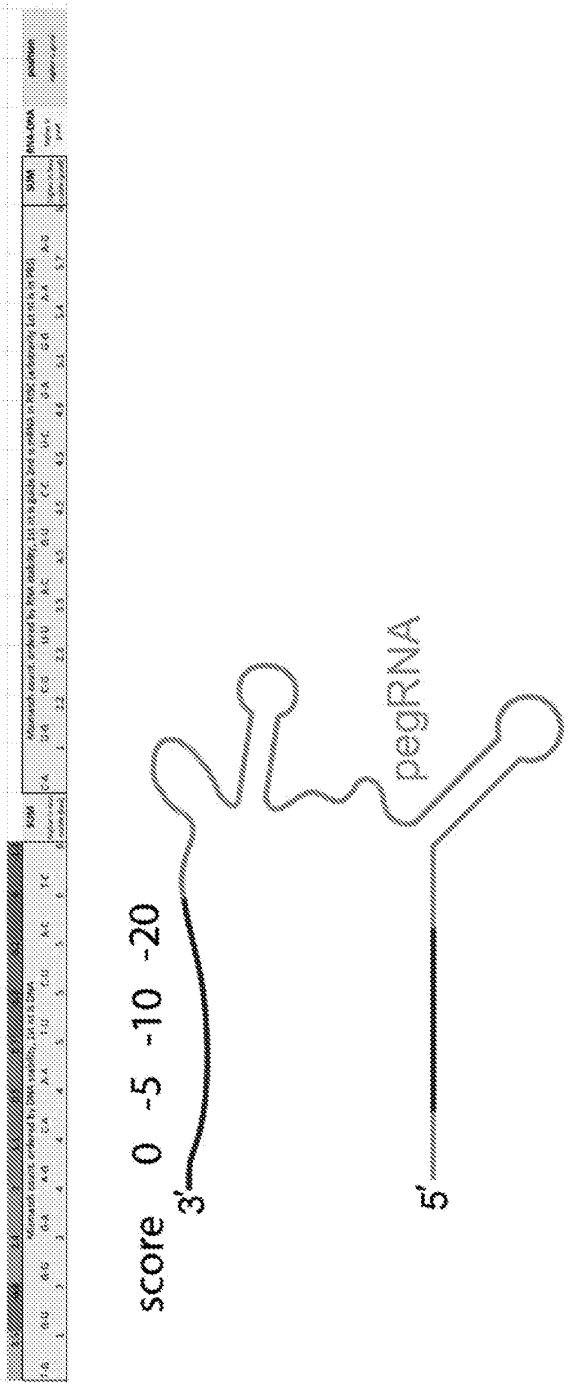


FIG. 140

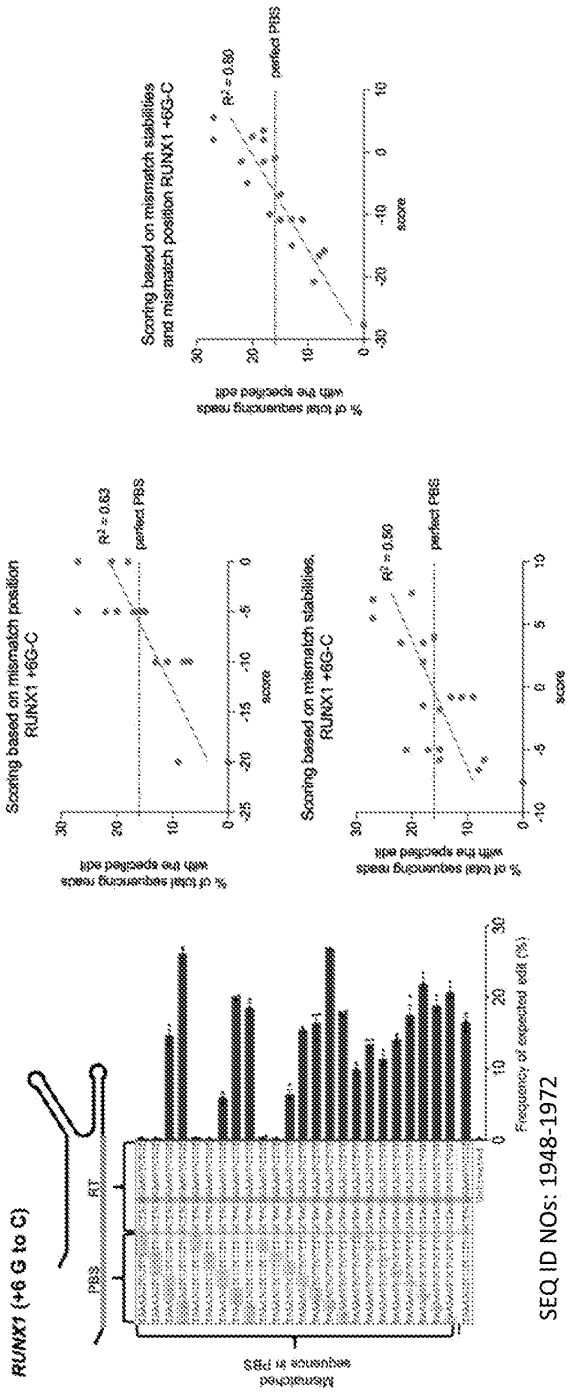

RUNX1 (+6 G to C)
PBS
RT
Mismatched sequence in PBS
Frequency of expected edit (%)
SEQ ID NOs: 1948-1972
Scoring based on mismatch position RUNX1 +6G-C
% of total sequencing reads with the specified edit
score
R² = 0.63
perfect PBS
Scoring based on mismatch stabilities. RUNX1 +6G-C
R² = 0.50
Scoring based on mismatch stabilities and mismatch position RUNX1 +6G-C
R² = 0.80

FIG. 141

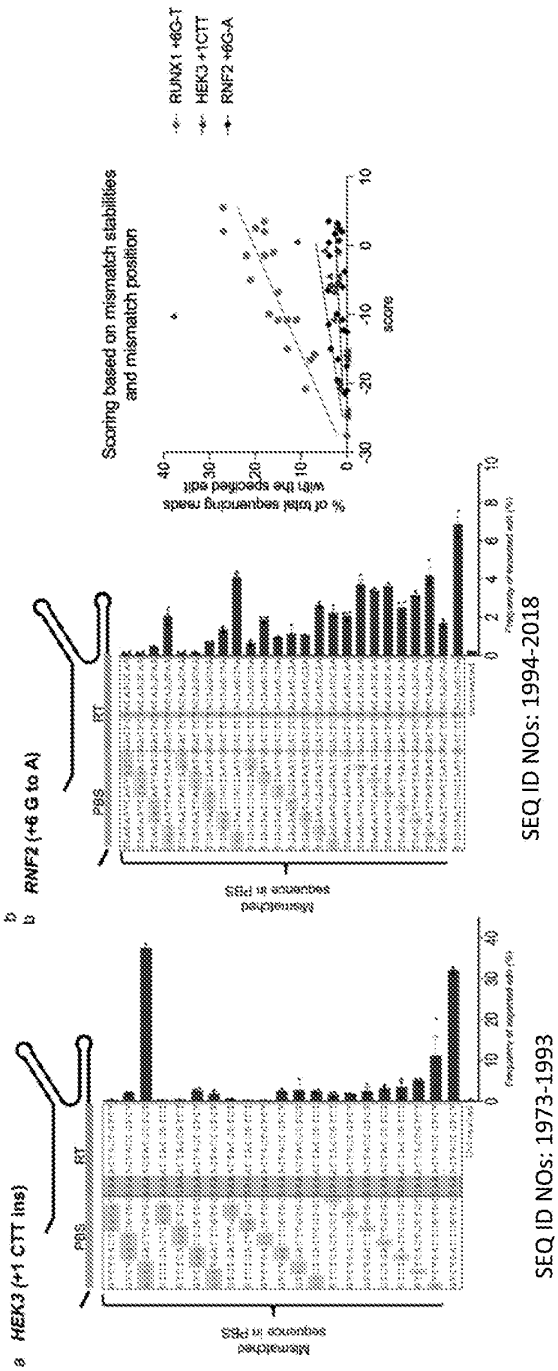

a HEK3 (+1 CTT ins)
b RNF2 (+6 G to A)
PBS
RT
Mismatches in PBS
% of total sequencing reads with the specified edit
Scoring based on mismatch stabilities and mismatch position
score
RUNX1 +6G-T
HEK3 +1CTT
RNF2 +6G-A
SEQ ID NOs: 1973-1993
SEQ ID NOs: 1994-2018

FIG. 142

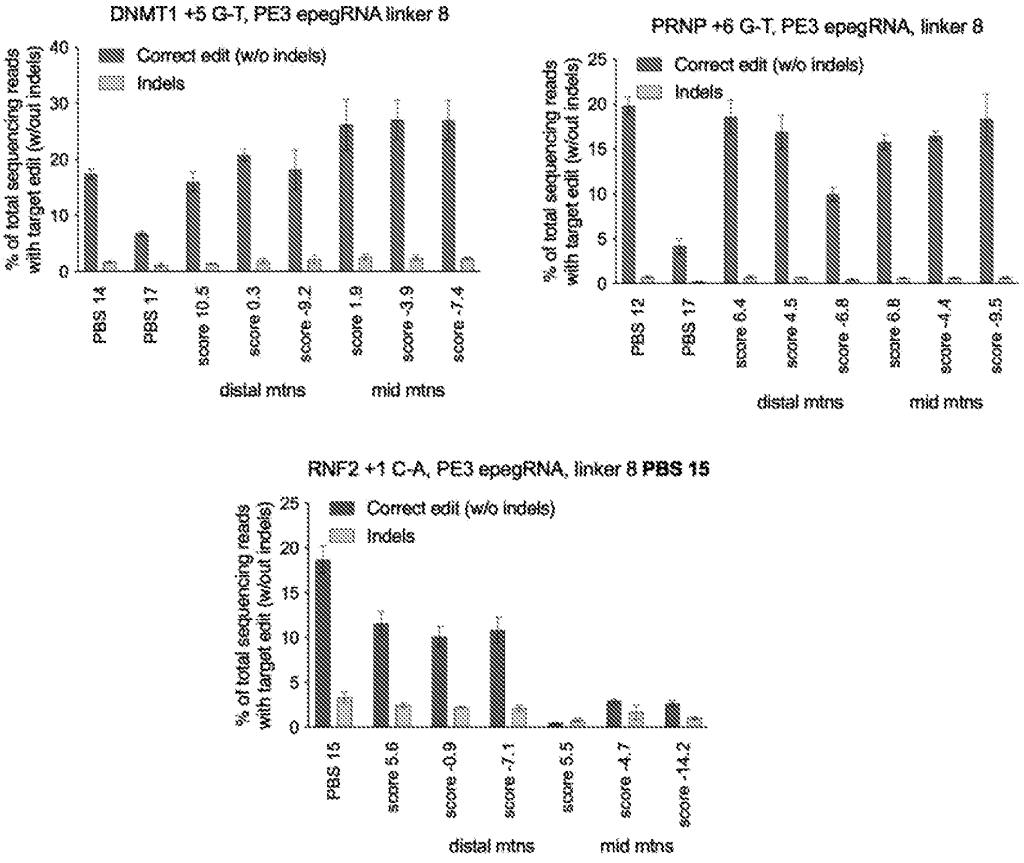

DNMT1 +5 G-T, PE3 epegRNA linker 8
% of total sequencing reads with target edit (w/out indels)
Correct edit (w/o indels)
Indels
PBS 14
PBS 17
score 10.5
score 0.3
score -9.2
score 1.9
score -3.9
score -7.4
distal mtns
mid mtns
PRNP +6 G-T, PE3 epegRNA, linker 8
PBS 12
PBS 17
score 6.4
score 4.5
score -6.8
score 6.8
score -4.4
score -9.5
RNF2 +1 C-A, PE3 epegRNA, linker 8 PBS 15
PBS 15
score 5.6
score -0.9
score -7.1
score 5.5
score -4.7
score -14.2

FIG. 143

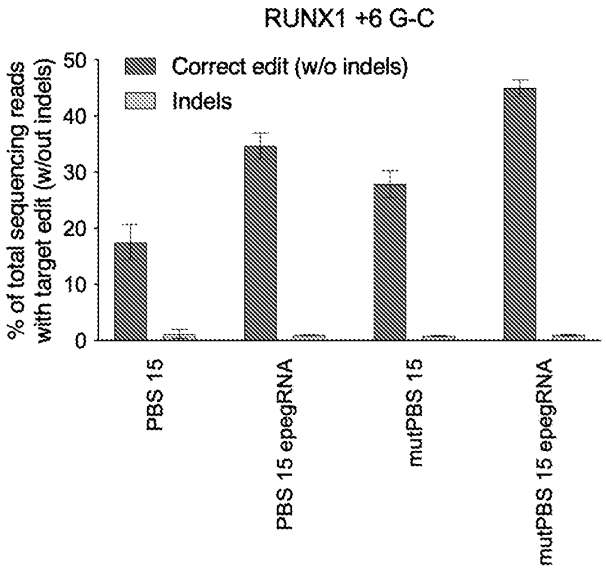

RUNX1 +6 G-C
% of total sequencing reads with target edit (w/out indels)
50
40
30
20
10
0
Correct edit (w/o indels)
Indels
PBS 15
PBS 15 epegRNA
mutPBS 15
mutPBS 15 epegRNA

FIG. 144

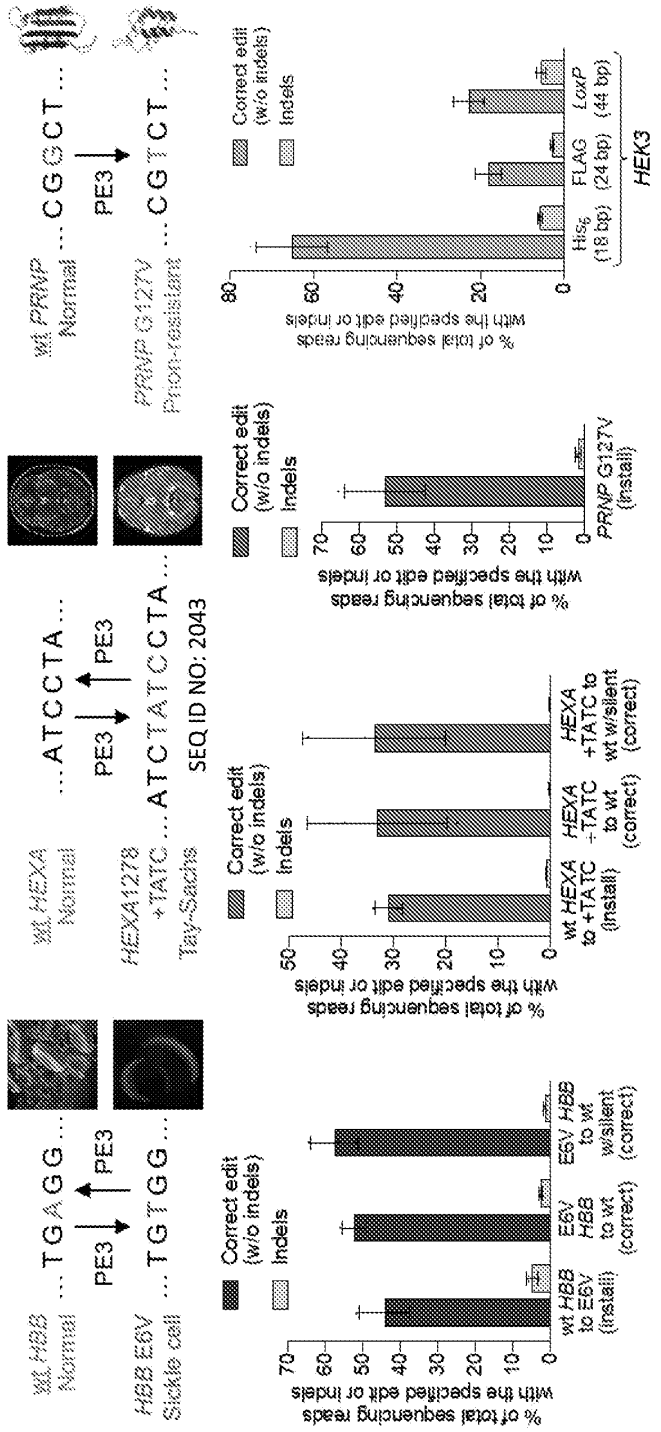

wt HBB
Normal
...TGAGG...
PE3
PE3
HBB E6V
Sickle cell
...TGTGG...
wt HEXA
Normal
...ATCCTA...
PE3
PE3
HEXA1278
+TATC
Tay-Sachs
...ATCTATCCTA...
SEQ ID NO: 2043
wt PRNP
Normal
...CGCCT...
PE3
PRNP G127V
Prion-resistant
...CGTCT...
% of total sequencing reads with the specified edit or indels
Correct edit (w/o indels)
Indels
wt HBB to E6V (install)
E6V HBB to wt (correct)
E6V HBB to wt w/silent (correct)
wt HEXA to +TATC (install)
HEXA +TATC to wt (correct)
HEXA +TATC to wt w/silent (correct)
PRNP G127V (install)
His6 (18 bp)
FLAG (24 bp)
LoxP (44 bp)
HEK3

FIG. 145

SEQ ID NO: 2019
CDKL5 5'··· TGGGGAATTGTGACAATATAGCTAT ··· 3'
CDKL5 1412delA 5'··· TGGGGAATTGTG–CAATATAGCTAT ··· 3'
SEQ ID NO: 2020
1412delA
Prime editing in principle can directly install or correct mutations that cause CDD
Cas9-induced nick
PAM
1412delA
SEQ ID NO: 2022
5'···GAGGGACTCCTAGAGGACTGGGGAATTGTGCAATATAGCTAT ··· 3'
Spacer
5'-GAGGGACUCCUAGAGGACUGGGUUUUAGAGCUA
Primer ~13 nt
DNA to be replaced by prime editing
RT template that inserts the missing A in CDKL5
CDKL5 c.1412delA locus
One of many possible pegRNAs that programs a prime editor to insert the missing A at the target site
SEQ ID NO: 2021
RT template (includes edit)
Primer-binding site (PBS)

FIG. 146

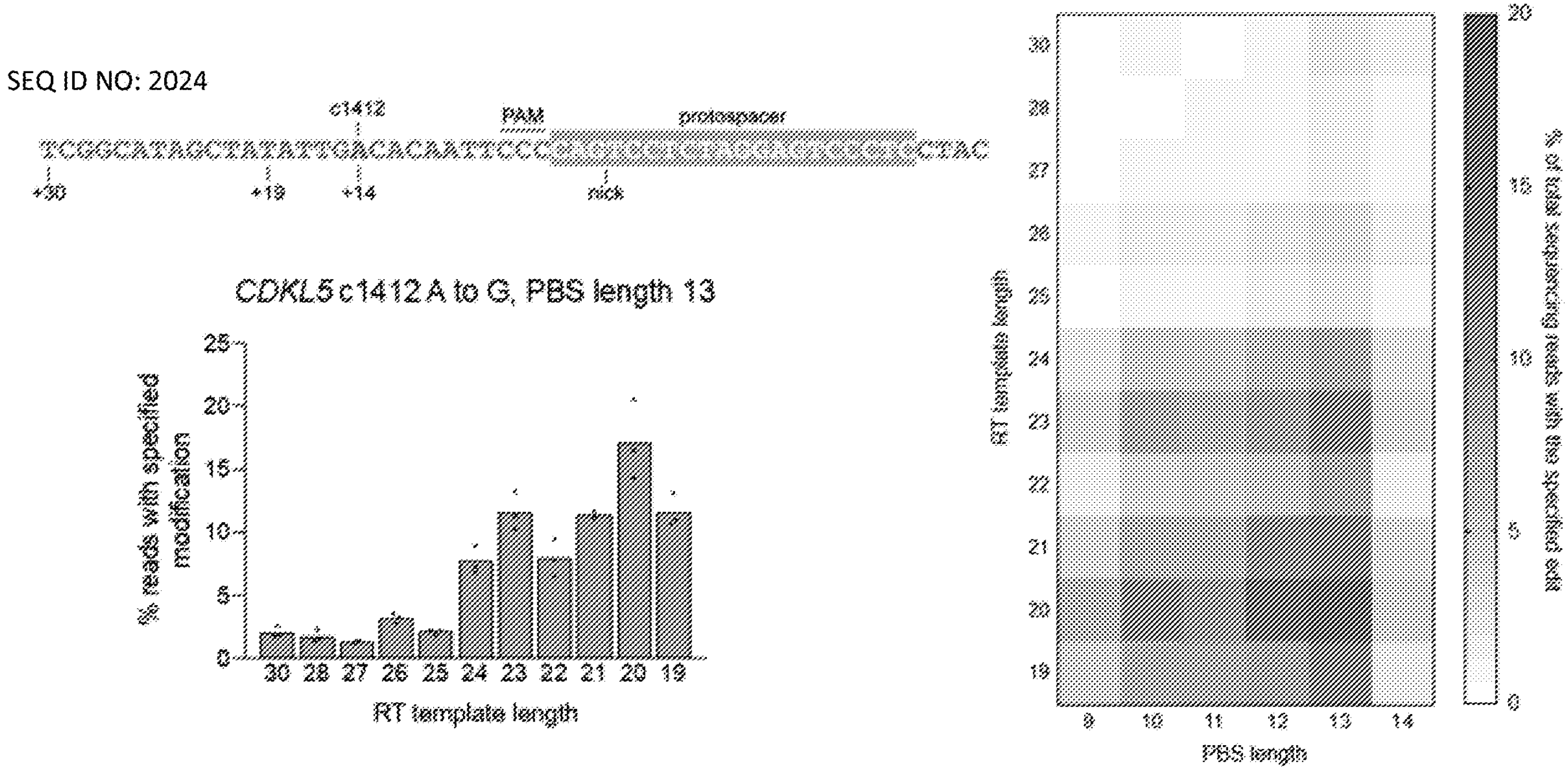

SEQ ID NO: 2024
c1412
PAM
protospacer
TCGGCATAGCTATATTGACACAATTCCC
CTAC
+30
+19
+14
nick
CDKL5 c1412 A to G, PBS length 13
% reads with specified modification
RT template length
% of total sequencing reads with the specified edit
PBS length

FIG. 147

human CDKL5 exon 8
SEQ ID NO: 2024
SEQ ID NO: 2025
SEQ ID NO: 2026
CDKL5 pathogenic missense mutations clustered in catalytic domain
Spacer
SEQ ID NO: 2027
Cas9-induced nick
SEQ ID NO: 2028
PAM
Primer
DNA to be replaced by prime editing
RT template
(programs a PE to restore any of the above mutations to wild-type)
Primer-binding site (PBS)

FIG. 148

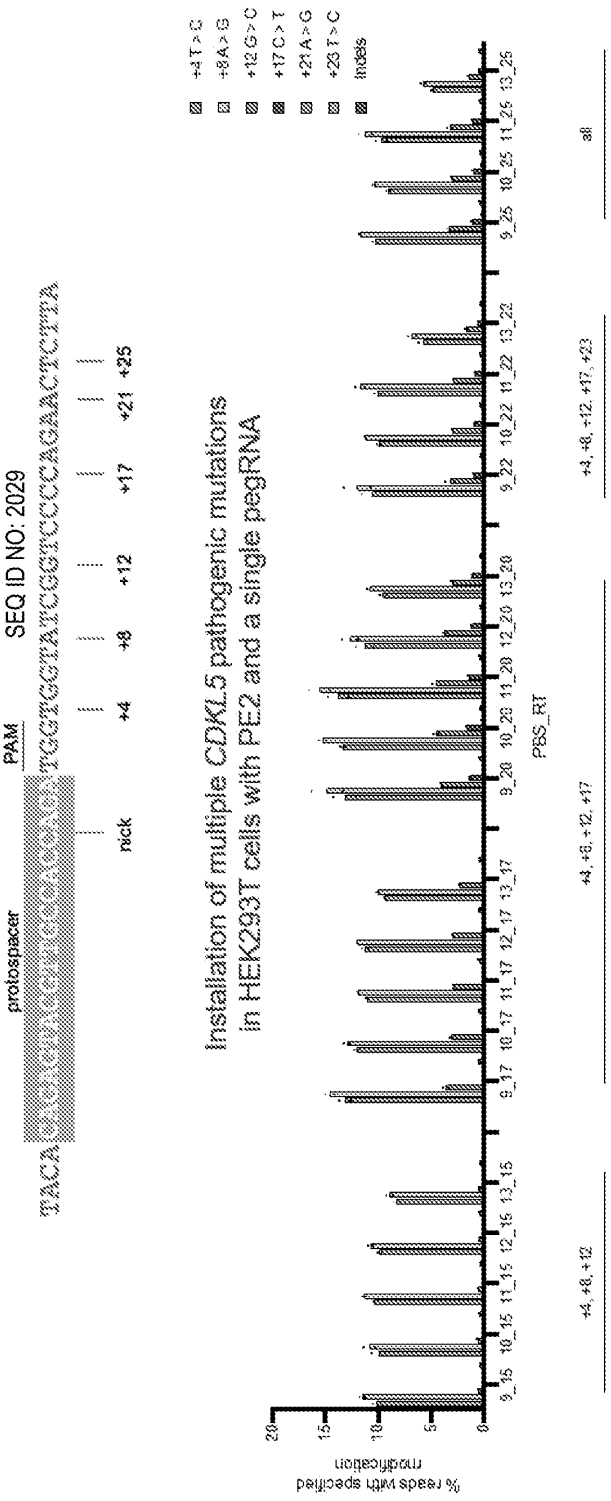

protospacer
PAM
SEQ ID NO: 2029
TACA
TGGTGGTATCGGTCCCCAGAACTCTTA
nick
+4
+8
+12
+17
+21
+25
Installation of multiple CDKL5 pathogenic mutations
in HEK293T cells with PE2 and a single pegRNA
+4 T > C
+8 A > G
+12 G > C
+17 C > T
+21 A > G
+23 T > C
indels
% reads with specified modification
PBS_RT
+4, +8, +12
+4, +8, +12, +17
+4, +8, +12, +17, +23
all

FIG. 149

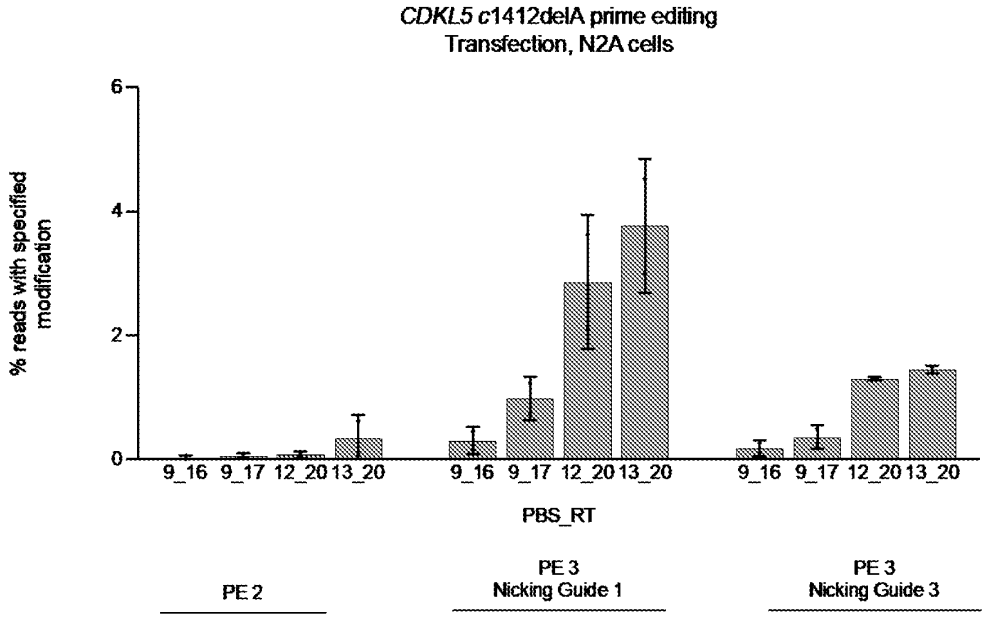


FIG. 150

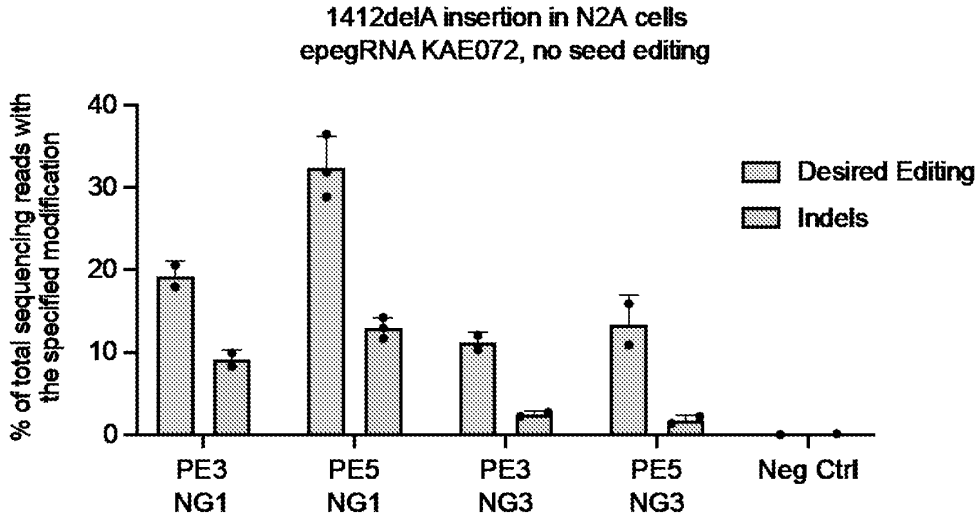


FIG. 151

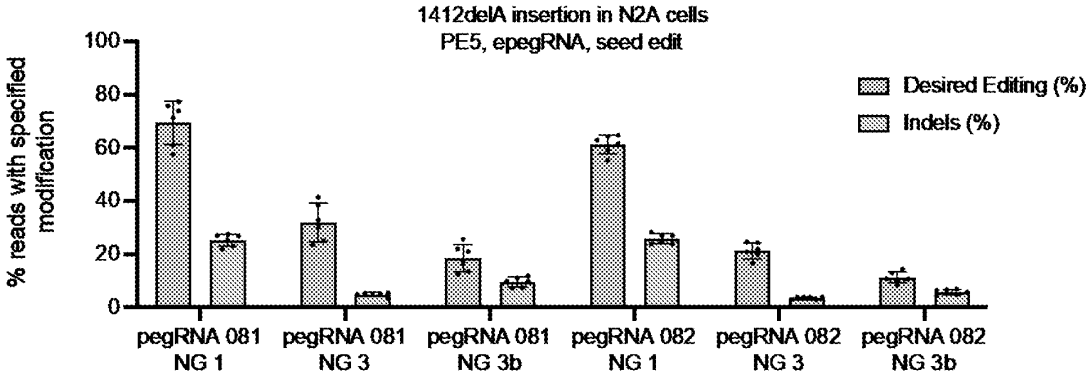


FIG. 152

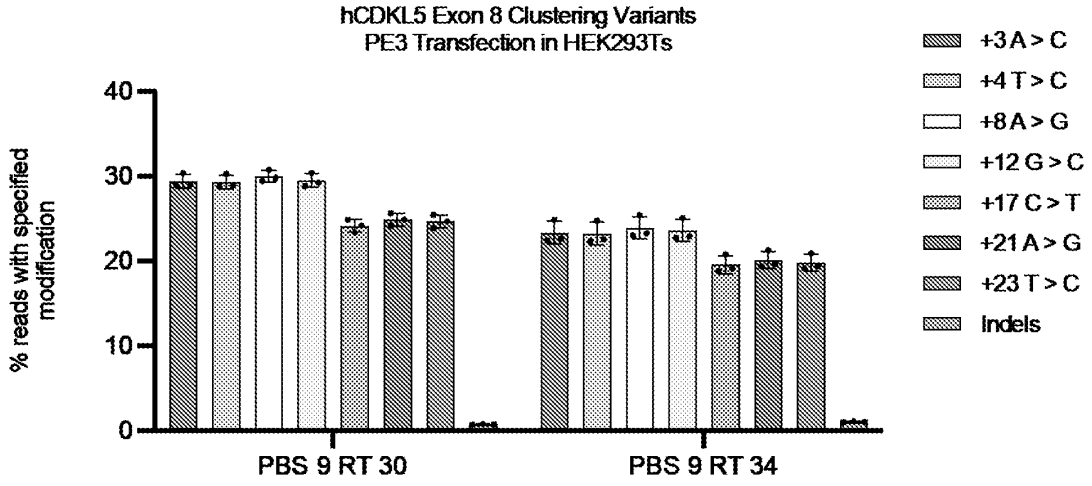


FIG. 153

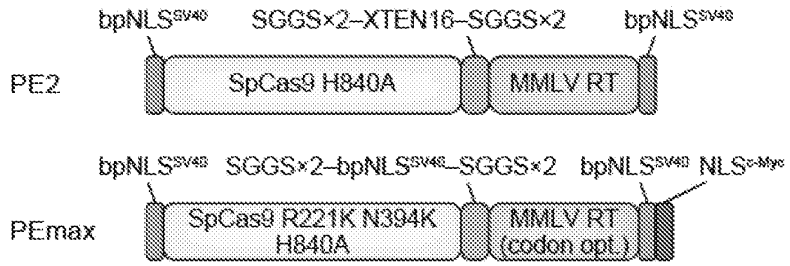


FIG. 154

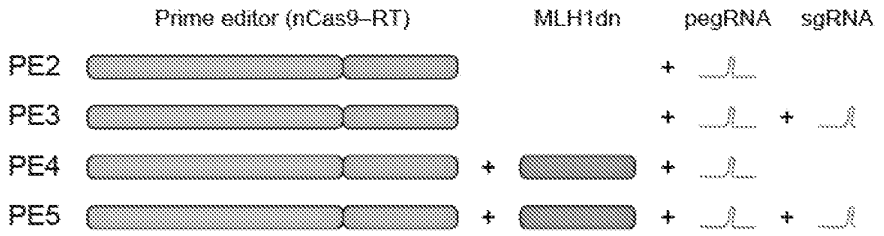


FIG. 155

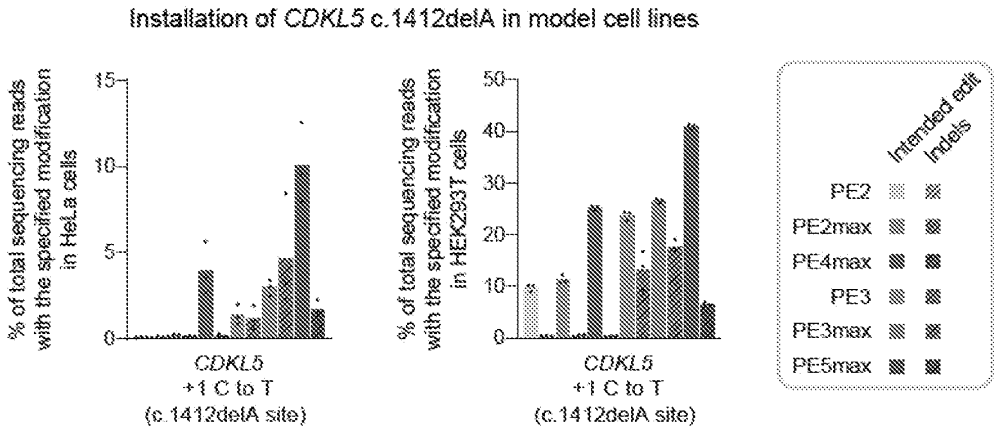


FIG. 156

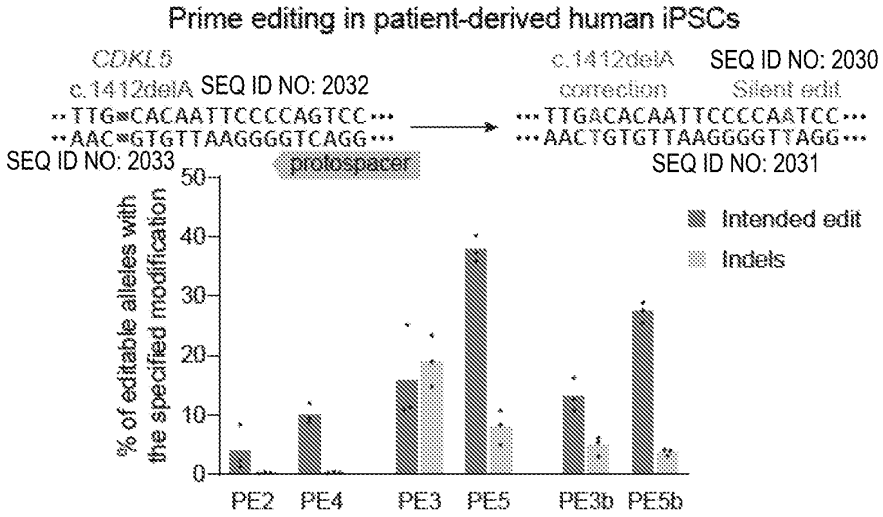


FIG. 157

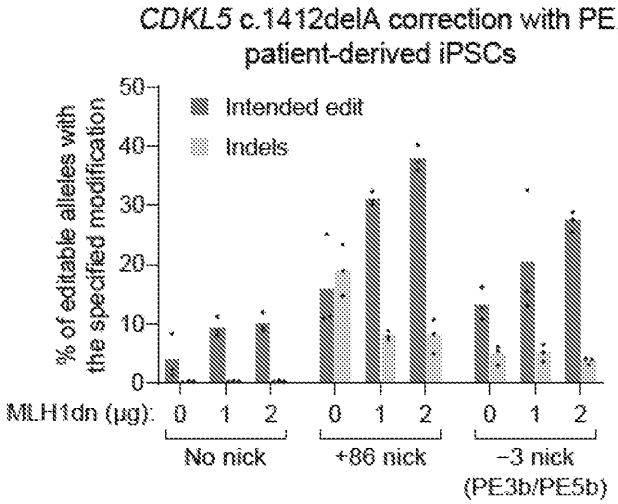


FIG. 158

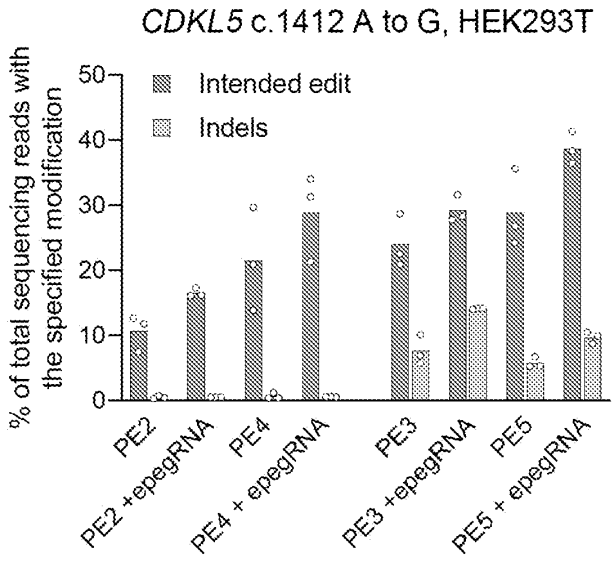


FIG. 159

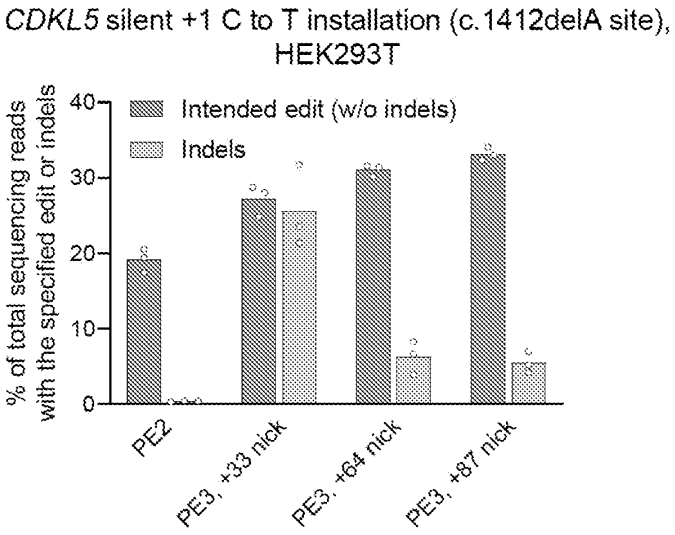


FIG. 160

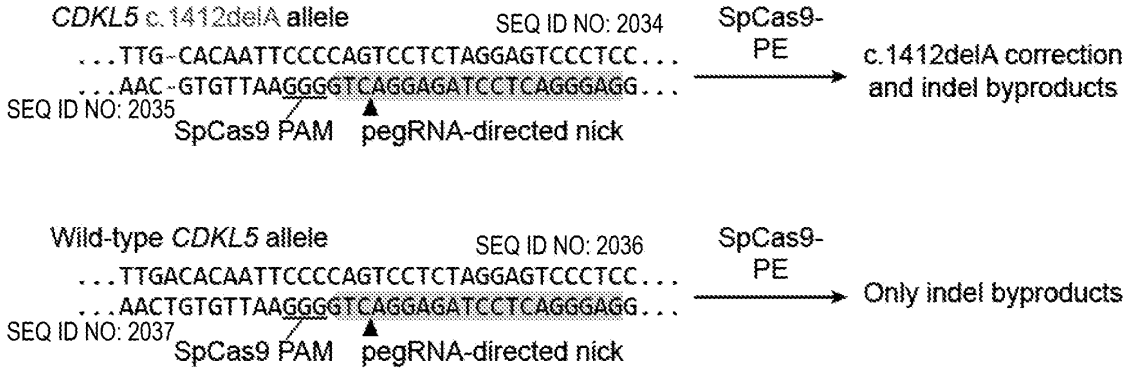


FIG. 161

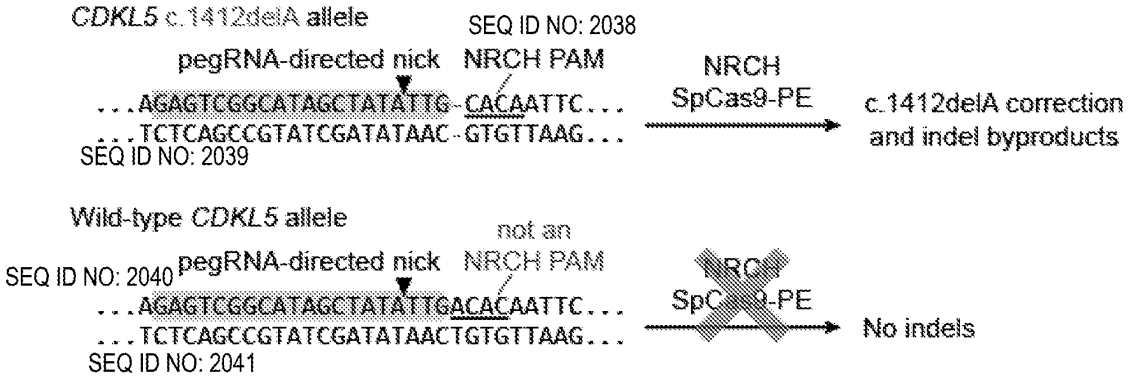


FIG. 162

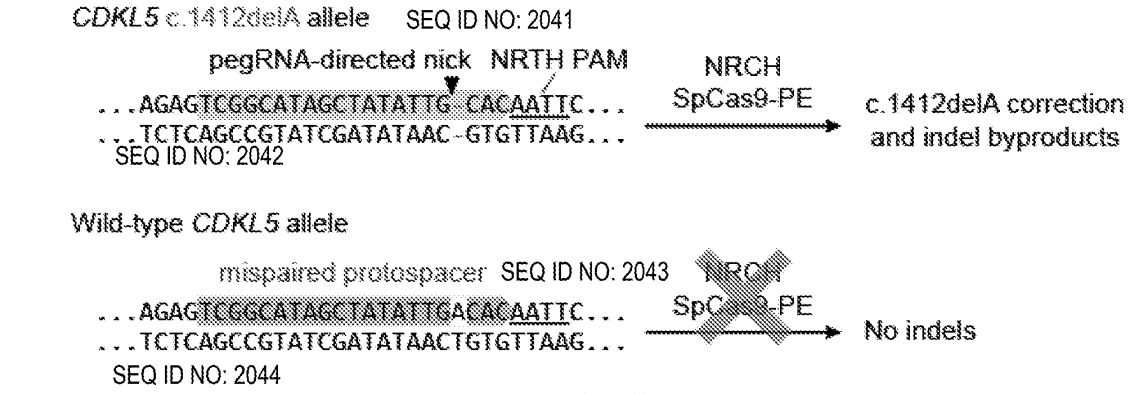


FIG. 163

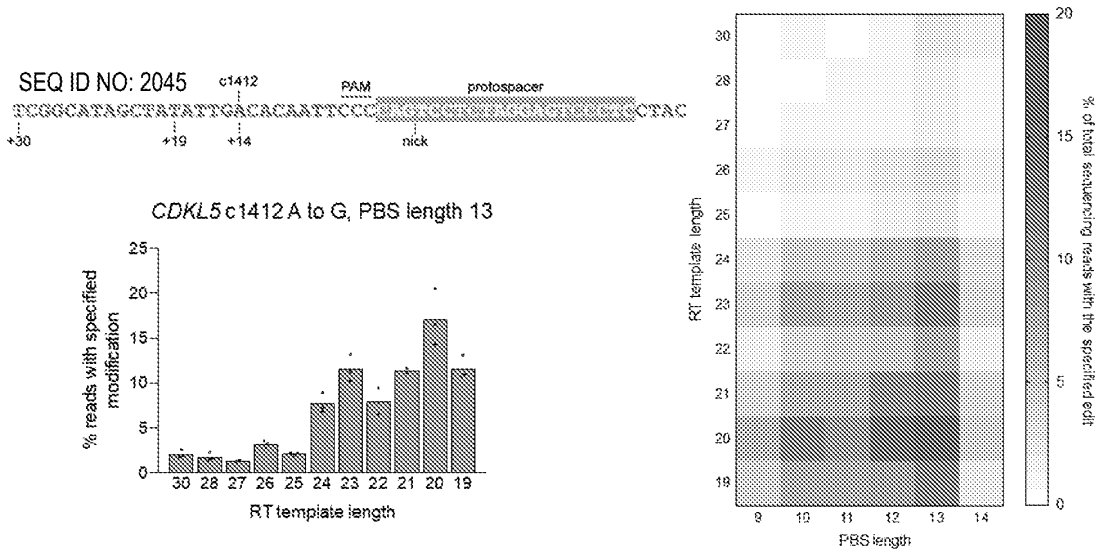

SEQ ID NO: 2045
c1412
PAM
protospacer
+30
+19
+14
nick
CDKL5 c1412 A to G, PBS length 13
% reads with specified modification
RT template length
RT template length
PBS length
% of total sequencing reads with the specified edit

FIG. 164

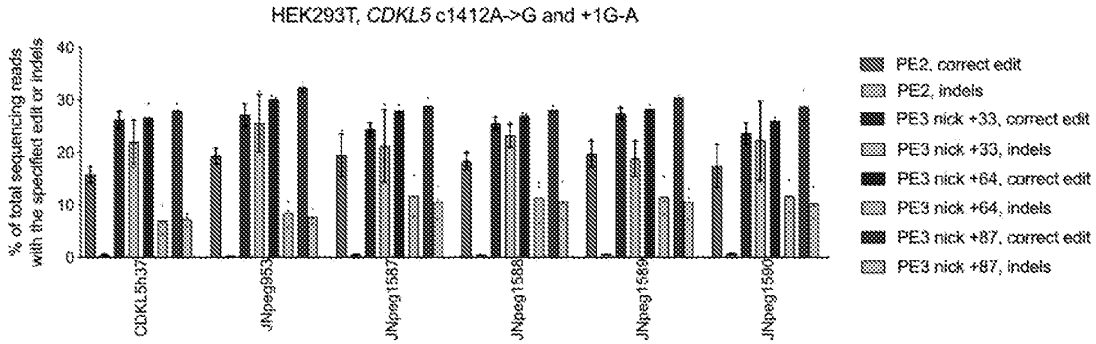

HEK293T, CDKL5 c1412A->G and +1G-A
% of total sequencing reads with the specified edit or indels
40
30
20
10
0
CDKL5h37
JNpeg955
JNpeg1587
JNpeg1588
JNpeg1589
JNpeg1590
PE2, correct edit
PE2, indels
PE3 nick +33, correct edit
PE3 nick +33, indels
PE3 nick +64, correct edit
PE3 nick +64, indels
PE3 nick +87, correct edit
PE3 nick +87, indels

FIG. 165