M-MLV reverse transcriptase
pegRNA
Cas9 nickase
Nicking of PAM strand
Priming of pegRNA extension with PAM strand
Primer binding site
RT template including edit
Reverse transcription
Polymerization of edited DNA from the pegRNA into the target site
Flap equilibration
5' Flap (contains edit)
5' Flap (lacks edit)
5' Flap cleavage
Ligation
DNA repair
Edited DNA
Nicking of non-edited strand
Preferential repair of nicked strand
Edited DNA
What cellular factors repair these DNA prime editing intermediates?

FIG. 1

PE target site design criteria:
1. Compatible with SaCas9-based PEs
2. Editing efficiency above background (>5%)
CRISPRi sgRNA
PE target site
Lentivirus packaging
Lentiviral sgRNA library
Transduction MOI=0.1
SpCas9-KRAB HeLa cells
Transfect SaCas9 PE and pegRNA
Sequence sgRNAs and PE targets
editing
10%
15%
5%
Correlate PE outcomes with gene interference

FIG. 2

HeLa expressing
dSpCas9-KRAB
Transduce sgRNA
with PE target site
non-targeting
CRISPRi sgRNA
PE target
site
Transfect SaCas9 PE
and pegRNA
Sequence PE
target
editing
10%

FIG. 3A

Target site editing, HeLa
+6 G to C
Indels
% reads with specified modification
0
5
10
15
20
PBS length: 8 9 8 9 8 9 8 9 8 9 8 9 8 9 8 9 8 9 8 9 8 9 8 9 8 9 8 9 8 9 8 9
RT template length: 13 13 14 14 15 15 16 16 13 13 14 14 15 15 16 16 13 13 14 14 15 15 16 16 13 13 14 14 15 15 16 16
SaPE2
SaPE3 +50 nick
SaPE3 -50 nick

FIG. 3B

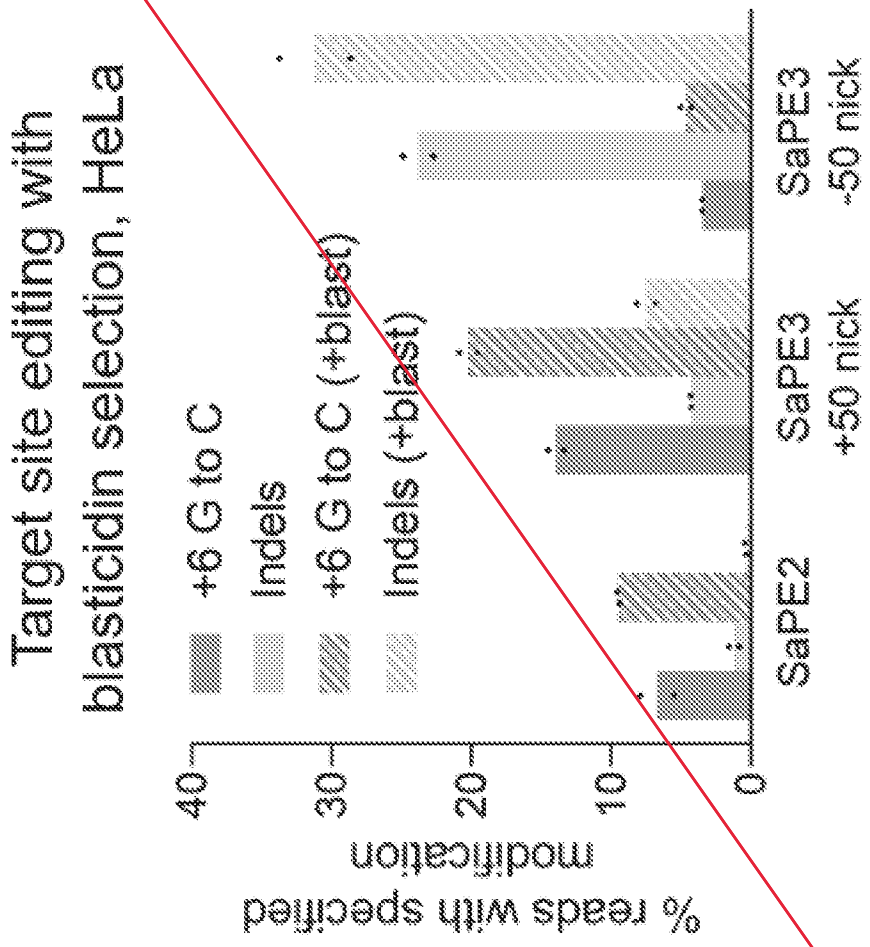

Target site editing with
blasticidin selection, HeLa
+6 G to C
Indels
+6 G to C (+blast)
Indels (+blast)
% reads with specified
modification
40
30
20
10
0
SaPE2
SaPE3
+50 nick
SaPE3
-50 nick

FIG. 3C

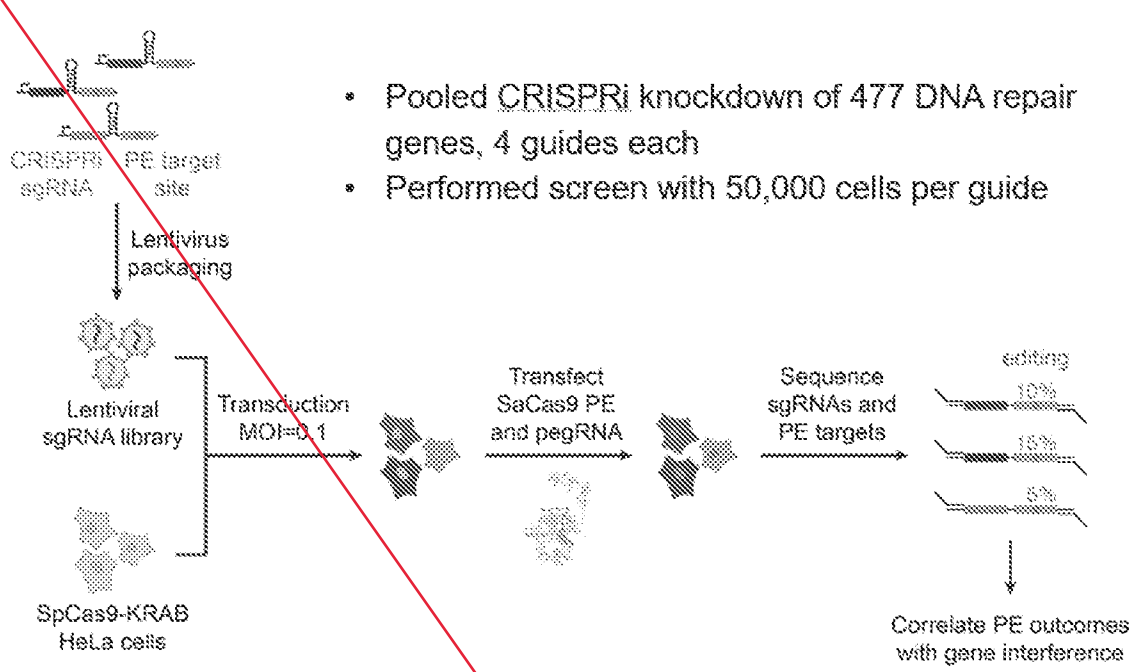

CRISPRi sgRNA
PE target site
Lentivirus packaging
Lentiviral sgRNA library
SpCas9-KRAB HeLa cells
Transduction MOI=0.1
Transfect SaCas9 PE and pegRNA
Sequence sgRNAs and PE targets
editing
10%
15%
5%
Correlate PE outcomes with gene interference
• Pooled CRISPRi knockdown of 477 DNA repair genes, 4 guides each
• Performed screen with 50,000 cells per guide

FIG. 4A

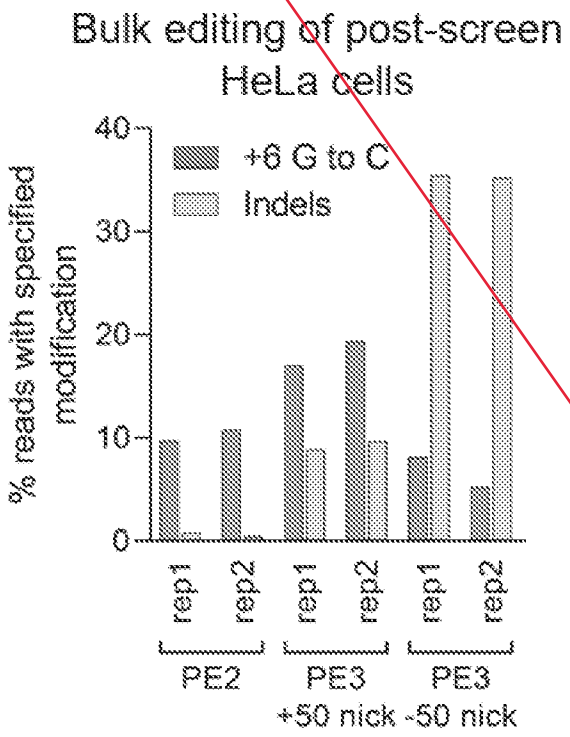

Bulk editing of post-screen HeLa cells
% reads with specified modification
40
30
20
10
0
+6 G to C
Indels
rep1
rep2
rep1
rep2
rep1
rep2
PE2
PE3
+50 nick
PE3
-50 nick

FIG. 4B

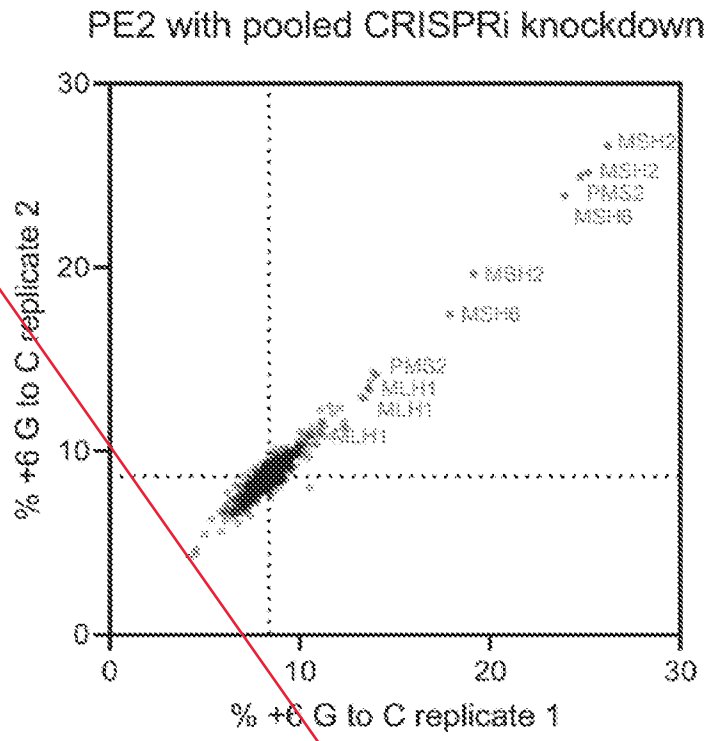

PE2 with pooled CRISPRi knockdown
% +6 G to C replicate 2
% +6 G to C replicate 1
0
10
20
30
MSH2
MSH2
PMS2
MSH6
MSH2
MSH6
PMS2
MLH1
MLH1
MLH1

FIG. 5A

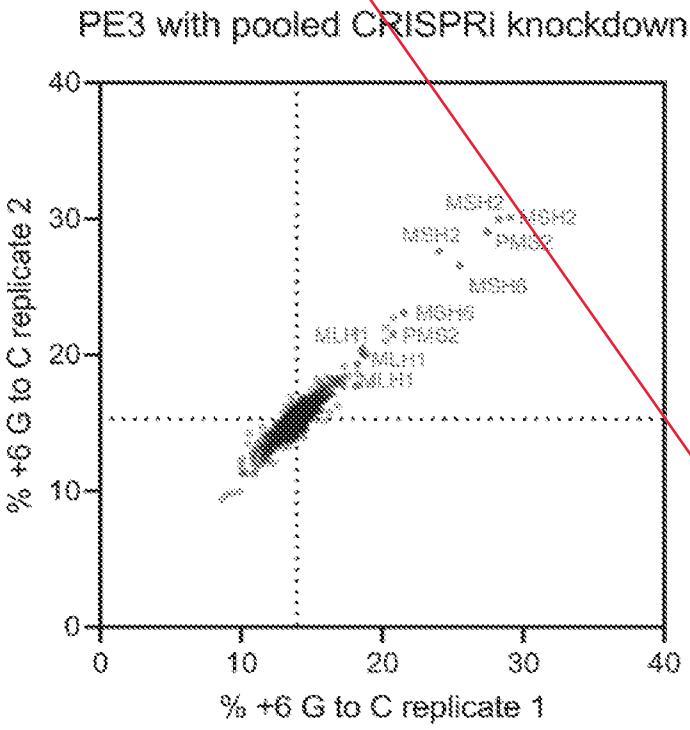

PE3 with pooled CRISPRi knockdown
% +6 G to C replicate 2
% +6 G to C replicate 1
0
10
20
30
40
MSH2
MSH2
MSH2
PMS2
MSH6
MSH6
MLH1
PMS2
MLH1
MLH1

FIG. 5B

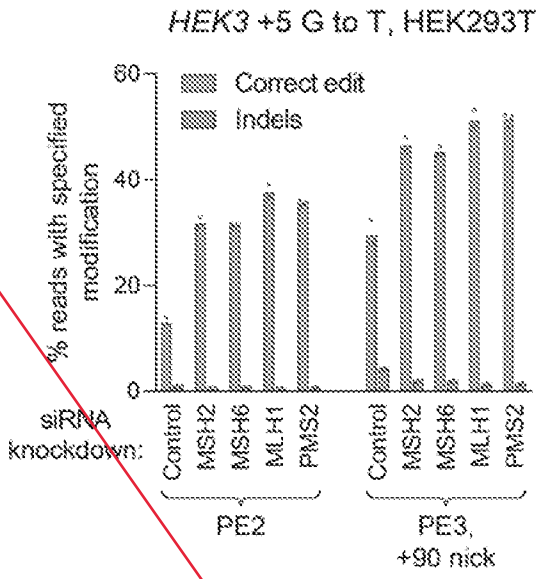


FIG. 6A

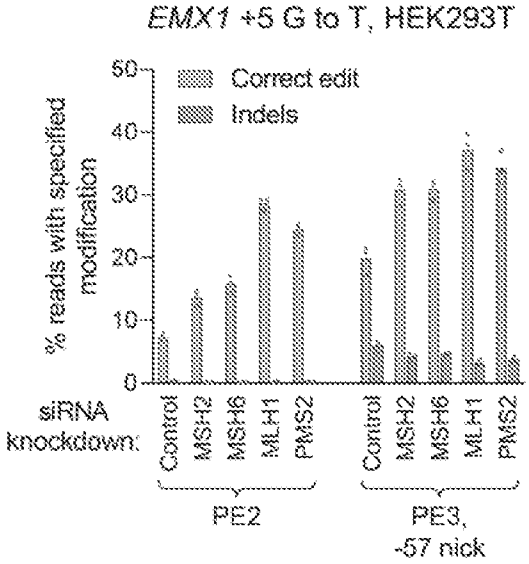


FIG. 6B

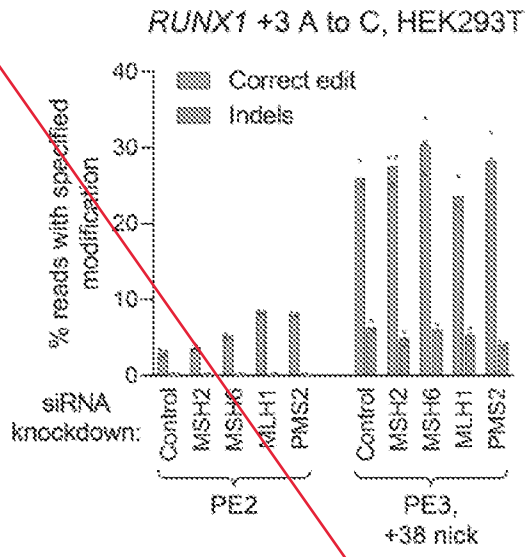


FIG. 6C

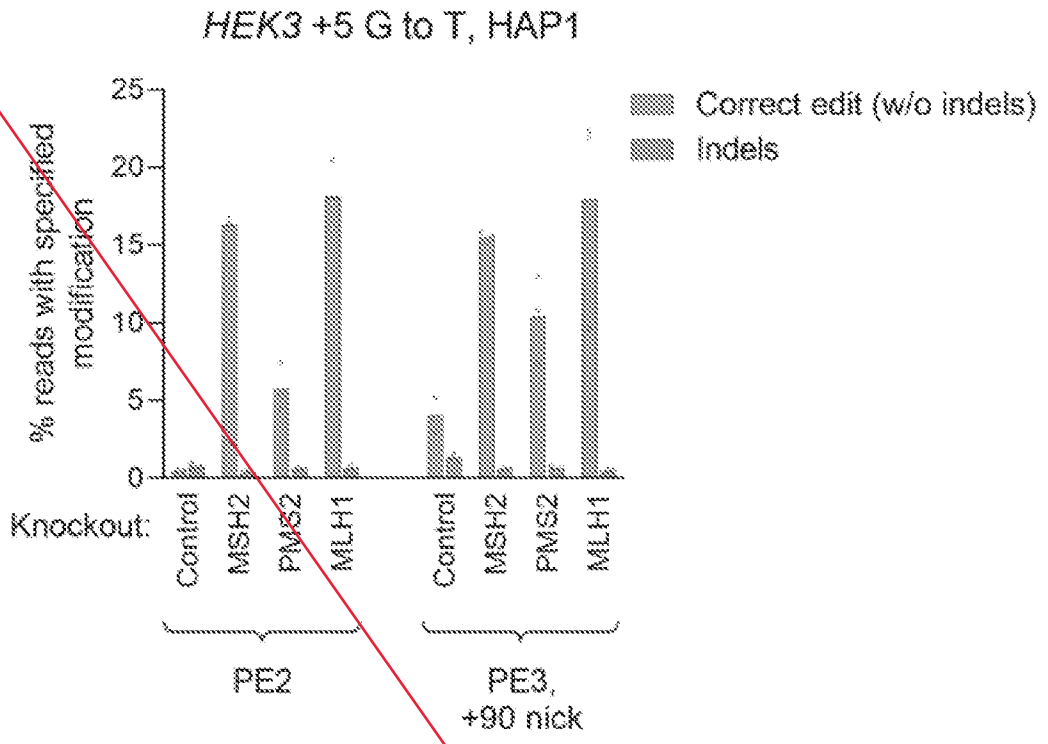

HEK3 +5 G to T, HAP1
Correct edit (w/o indels)
Indels
% reads with specified modification
25
20
15
10
5
0
Knockout:
Control
MSH2
PMS2
MLH1
Control
MSH2
PMS2
MLH1
PE2
PE3,
+90 nick

FIG. 7A

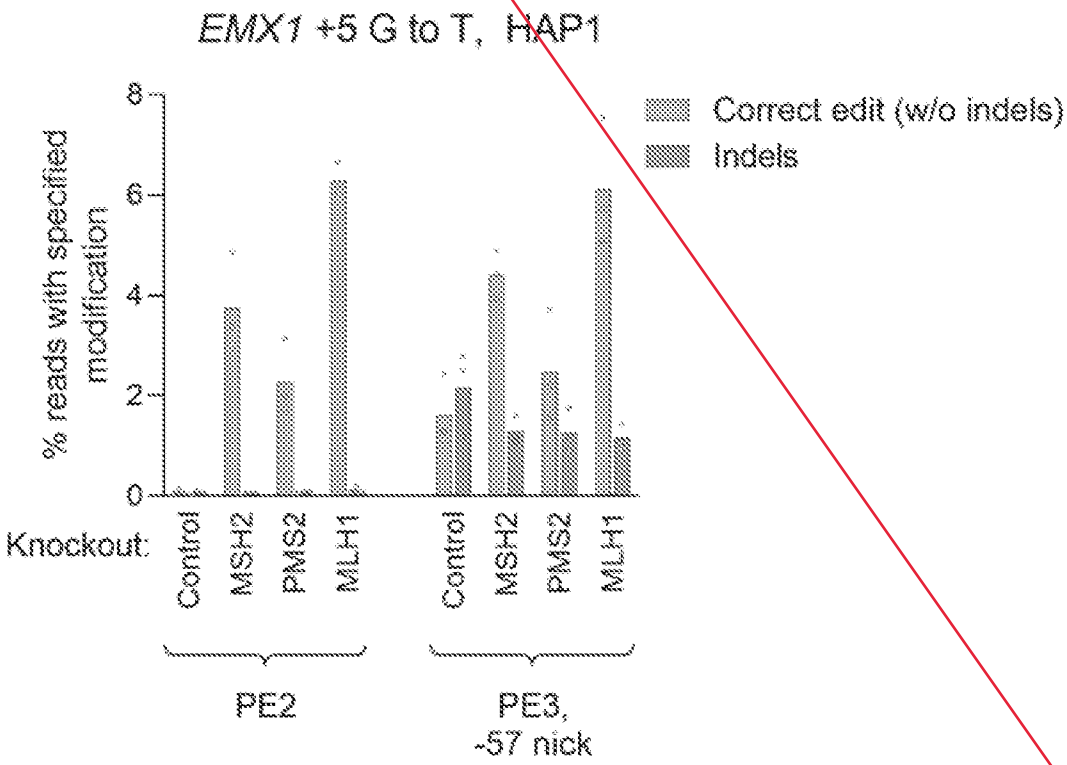

EMX1 +5 G to T, HAP1
Correct edit (w/o indels)
Indels
% reads with specified modification
8
6
4
2
0
Knockout:
Control
MSH2
PMS2
MLH1
Control
MSH2
PMS2
MLH1
PE2
PE3,
-57 nick

FIG. 7B

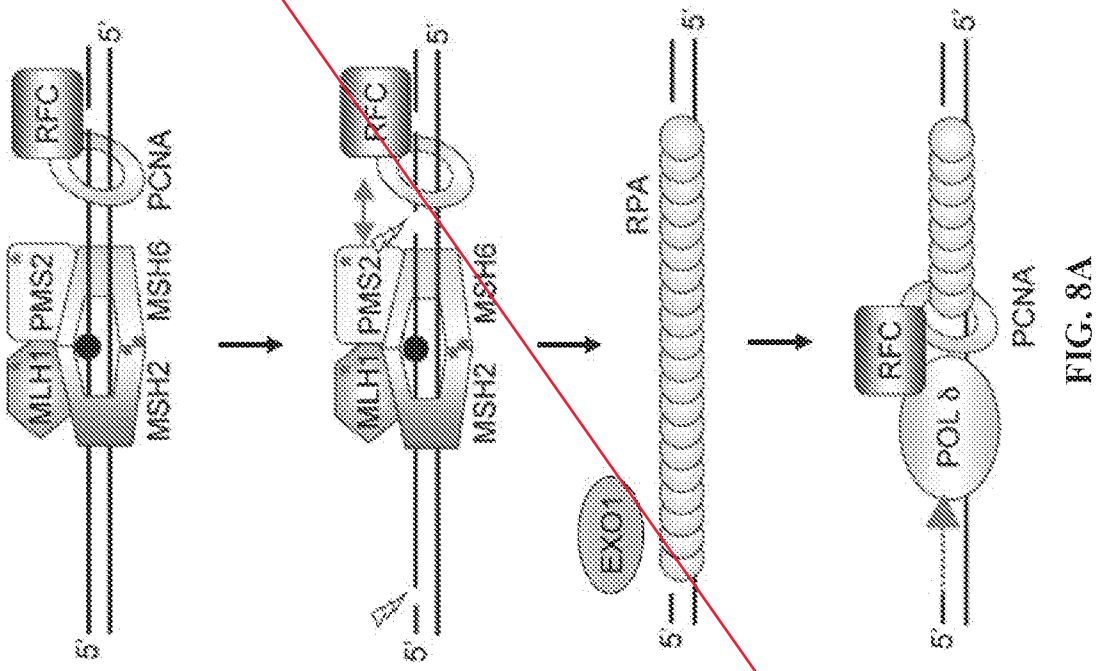


FIG. 8A

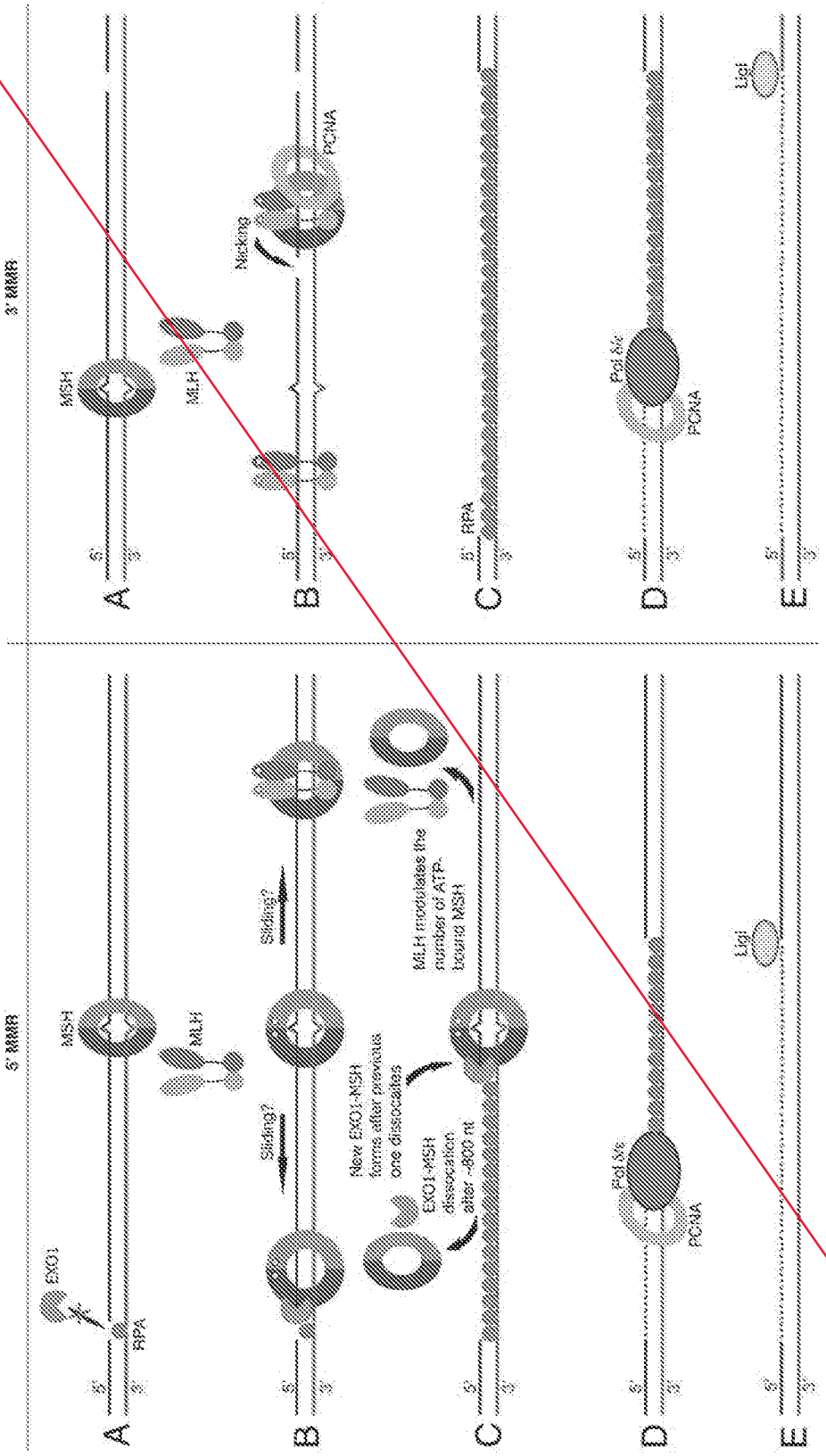

5' MMR
3' MMR
EXO1
MSH
MLH
RPA
PCNA
Sliding?
New EXO1-MSH forms after previous one dissociates
EXO1-MSH dissociation after ~800 nt
MLH modulates the number of ATP-bound MSH
Nicking
Pol δ/ε
Ligi
A
B
C
D
E
5'
3'

FIG. 8B

Flap equilibration
5' Flap cleavage
Ligation
MMR, with strand discrimination
MMR, no strand discrimination
Unedited
Edited

FIG. 9A

Flap equilibration
5' Flap cleavage
Ligation
Inhibited MMR
Inhibited MMR
Unedited
Edited

FIG. 9B

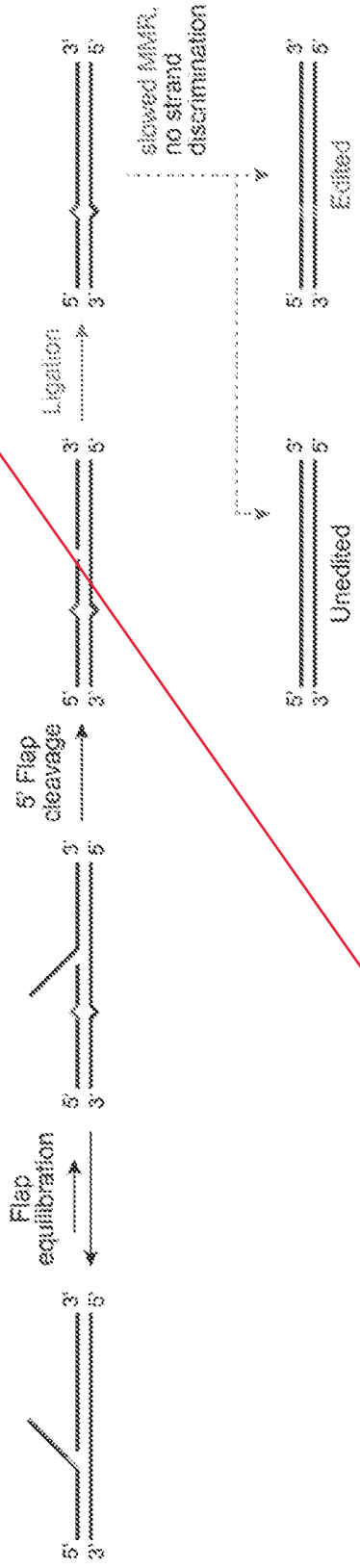

Flap equilibration
5' Flap cleavage
Ligation
slowed MMR, no strand discrimination
Unedited
Edited

FIG. 9C

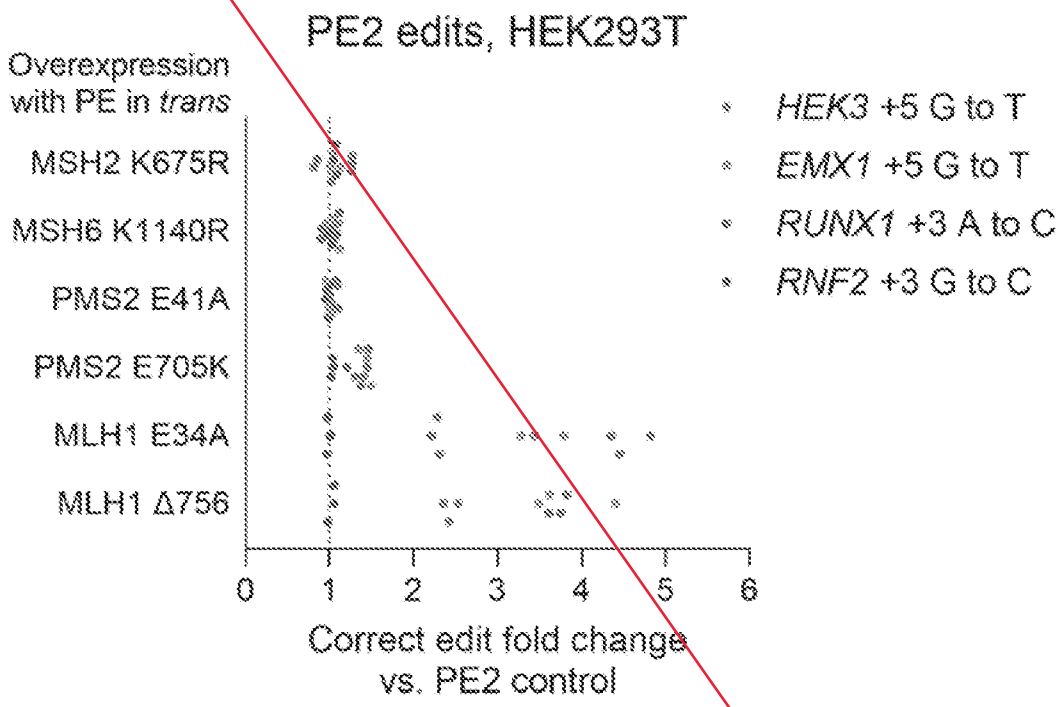

PE2 edits, HEK293T
Overexpression with PE in trans
MSH2 K675R
MSH6 K1140R
PMS2 E41A
PMS2 E705K
MLH1 E34A
MLH1 Δ756
HEK3 +5 G to T
EMX1 +5 G to T
RUNX1 +3 A to C
RNF2 +3 G to C
0
1
2
3
4
5
6
Correct edit fold change vs. PE2 control

FIG. 10

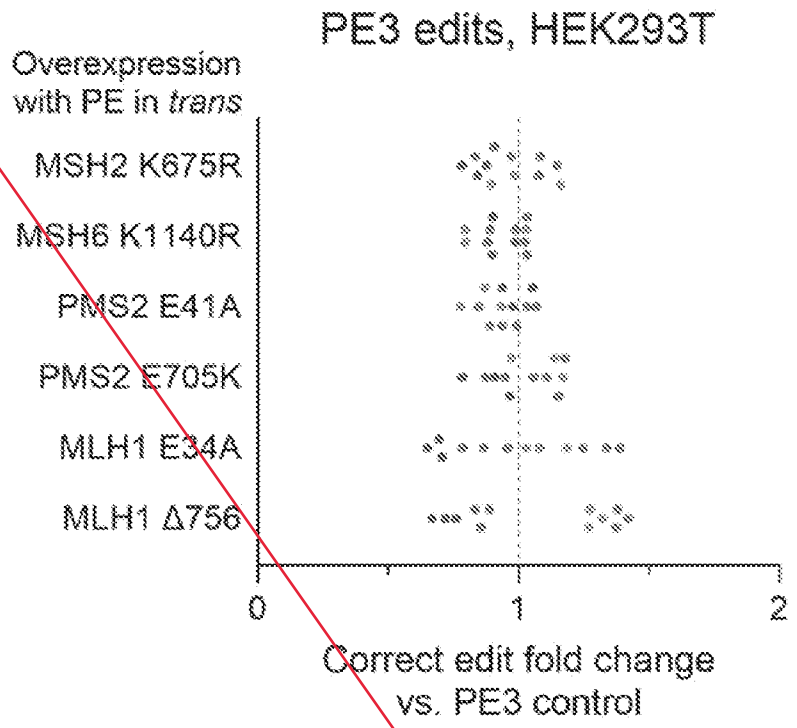


FIG. 11A

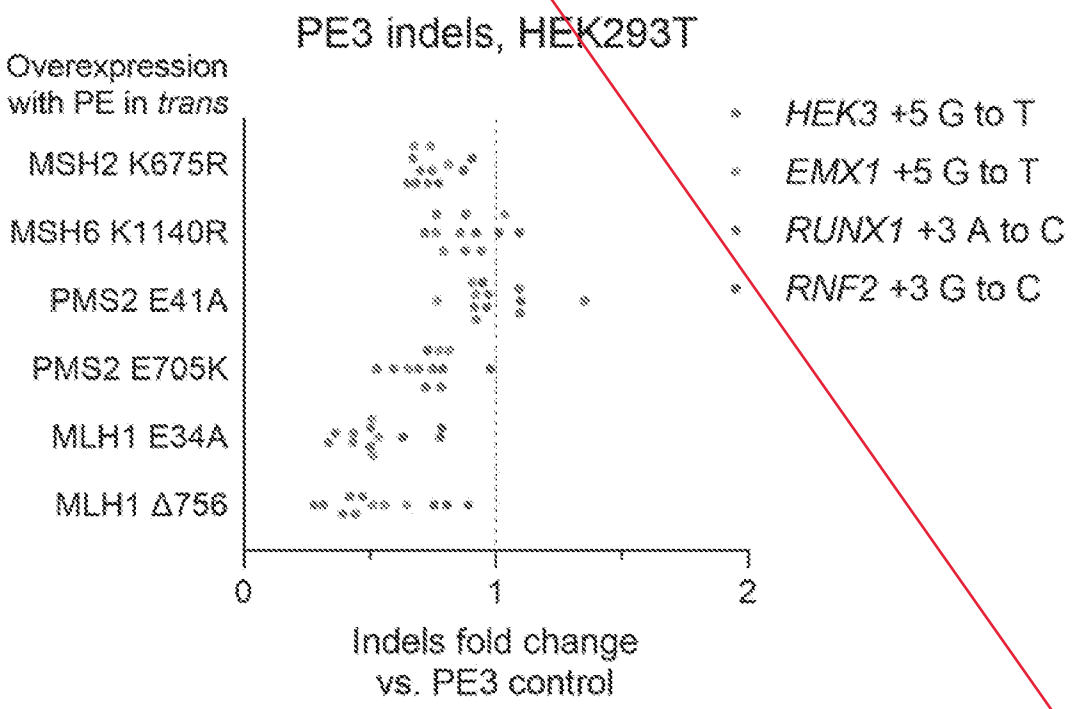


FIG. 11B

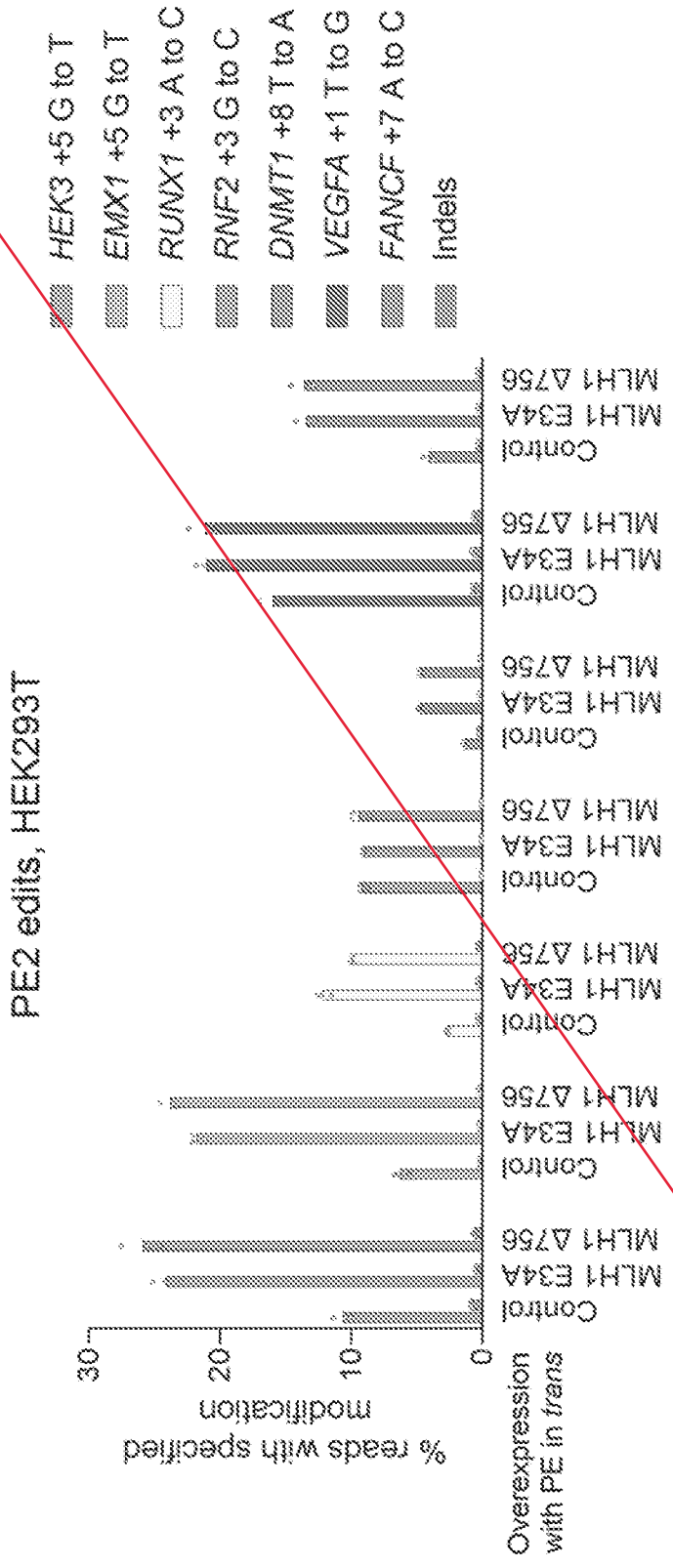

PE2 edits, HEK293T
% reads with specified modification
30
20
10
0
Overexpression with PE in trans
Control
MLH1 E34A
MLH1 Δ756
HEK3 +5 G to T
EMX1 +5 G to T
RUNX1 +3 A to C
RNF2 +3 G to C
DNMT1 +8 T to A
VEGFA +1 T to G
FANCF +7 A to C
Indels

FIG. 12A

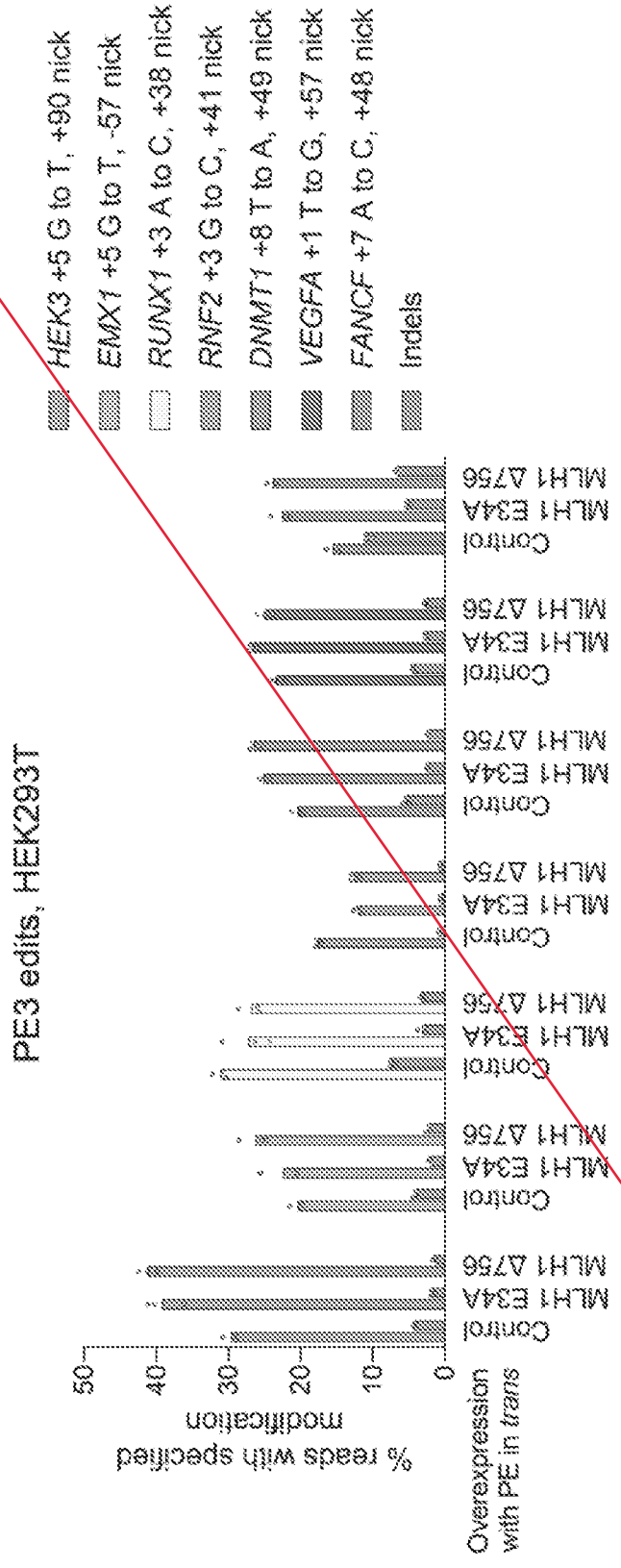

PE3 edits, HEK293T
% reads with specified modification
0
10
20
30
40
50
Overexpression with PE in trans
Control
MLH1 E34A
MLH1 Δ756
HEK3 +5 G to T, +90 nick
EMX1 +5 G to T, -57 nick
RUNX1 +3 A to C, +38 nick
RNF2 +3 G to C, +41 nick
DNMT1 +8 T to A, +49 nick
VEGFA +1 T to G, +57 nick
FANCF +7 A to C, +48 nick
Indels

FIG. 12B

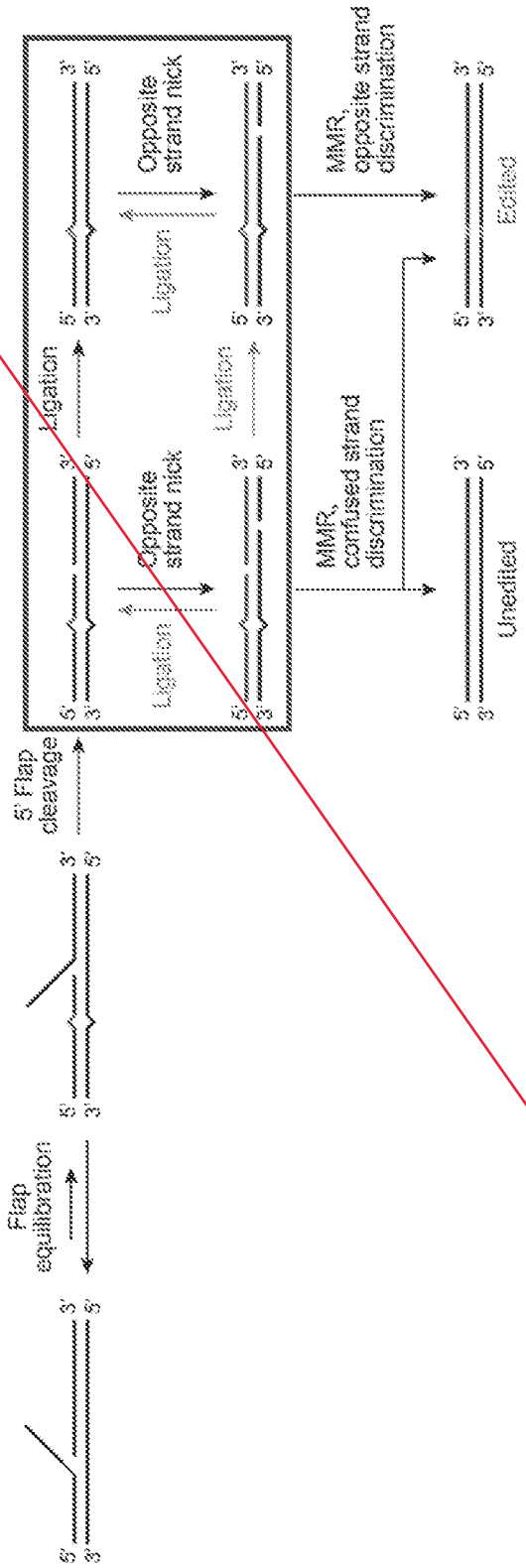

Flap equilibration
5' Flap cleavage
Ligation
Opposite strand nick
Ligation
Ligation
Opposite strand nick
Ligation
MMR, confused strand discrimination
MMR, opposite strand discrimination
Unedited
Edited

FIG. 13

Unedited:edited mixture
Edited favored
Opposite strand nick
Ligation
Ligation
Opposite strand nick
Ligation
Ligation
Unedited favored
Unedited:edited mixture
5'
3'
3'
5'
3'
6'
5'
3'
3'
5'
5'
3'
3'
5'
5'
3'

FIG. 14

MLH1
1
335
461
501
756
ATPase
MutS interaction
NLS
PMS2 dimerization
Endonuclease
PE2 editing fold-change, HEK293T
Overexpression with PE in trans
MLH1
MLH1 E34A
MLH1 Δ756
MLH1dn = MLH1 Δ754-756
MLH1 E34A Δ754-756
MLH1 1-335
MLH1 1-335 E34A
MLH1dn^NTD = MLH1 1-335 NLS^SV40
MLH1 501-756
MLH1 501-753
MLH1 461-753
NLS^SV40 MLH1 501-753
* HEK3 +5 G to T
* EMX1 +5 G to T
* RUNX1 +3 A to C
0
1
2
3
4
5
Correct edit fold change vs. PE2 control

FIG. 15A

MLH1
1
335
460
500
756
ATPase
MutS interaction
NLS
PMS2 dimerization
Endonuclease
PE2 edits, HEK293T
Overexpression with PE in trans
Control
MLH1 Δ754-756 (MLH1neg)
MLH1 1-335 NLS
MLH1 1-335 NLS genscript
MLH1 1-400 genscript
MLH1 1-460 genscript
MLH1 1-485 genscript
MLH1 1-501 genscript
MLH1 1-400 NLS genscript
MLH1 1-460 NLS genscript
MLH1 1-485 NLS genscript
MLH1 1-501 NLS genscript
HEK3 +5 G to T
EMX1 +5 G to T
0
10
20
30
40
50
% reads with specified modification

FIG. 15B

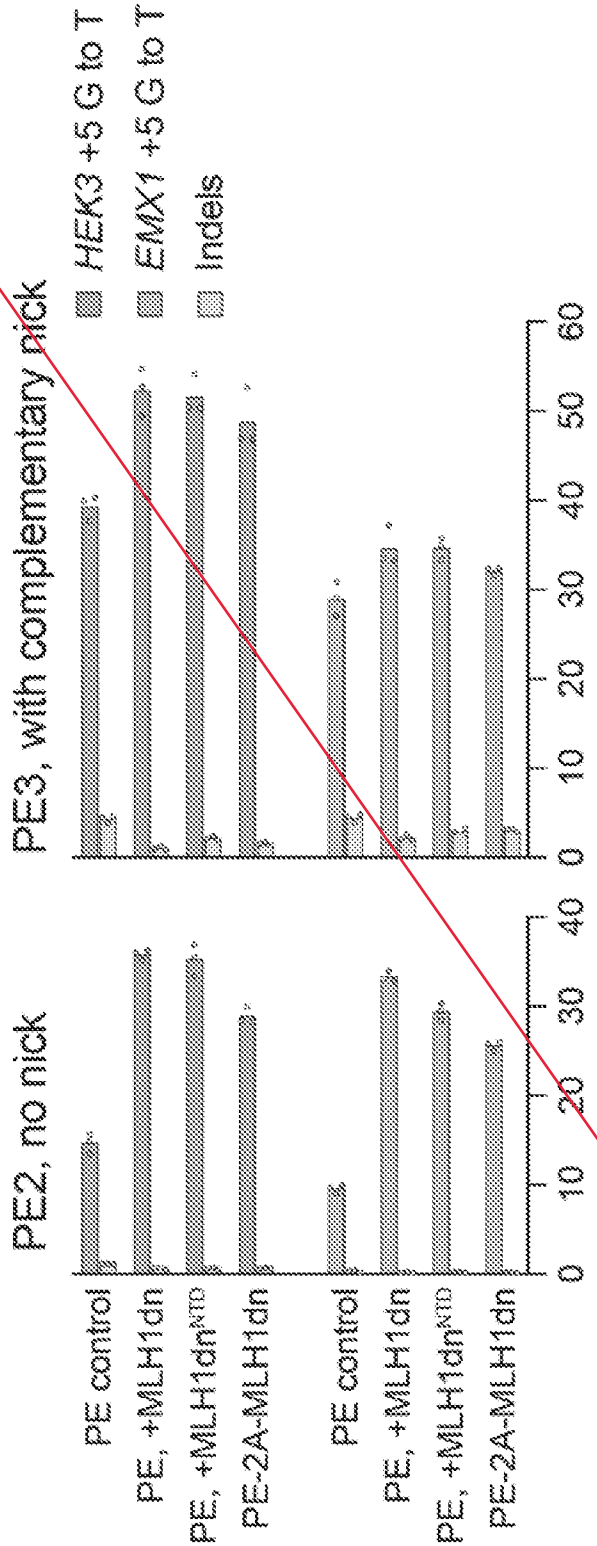

PE2, no nick
PE3, with complementary nick
HEK3 +5 G to T
EMX1 +5 G to T
Indels
PE control
PE, +MLH1dn
PE, +MLH1dn^NTD
PE-2A-MLH1dn
PE control
PE, +MLH1dn
PE, +MLH1dn^NTD
PE-2A-MLH1dn
0
10
20
30
40
0
10
20
30
40
50
60
% reads with specified modification

FIG. 15C

FIG. 15D

FIG. 15E

FIG. 15F

FIG. 15G

FIG. 15H

PE3 indels, HEK293T
% reads with indels
0
10
20
30
HEK3 +1 T to A
+90 nick
EMX1 +5 G to T
-57 nick
DNMT1 +1 A to C
+49 nick
FANCF +3 C to T
+48 nick
Untreated
PE
PE, +MLH1dn
nCas9-dead RT
nCas9-dead RT, +MLH1dn
nCas9
nCas9, +MLH1dn
PE, silent pegRNA
PE, silent pegRNA, +MLH1dn

FIG. 16

HEK3 +1 T to A, +90 nick
DNMT1 +1 A to C, +49 nick
PE
PE, +
dn
nCas9-dead RT
nCas9-dead RT, +MLH1dn
nCas9
nCas9, +MLH1dn
PE, silent pegRNA
PE, silent pegRNA, +MLH1dn
position in target locus

FIG. 17

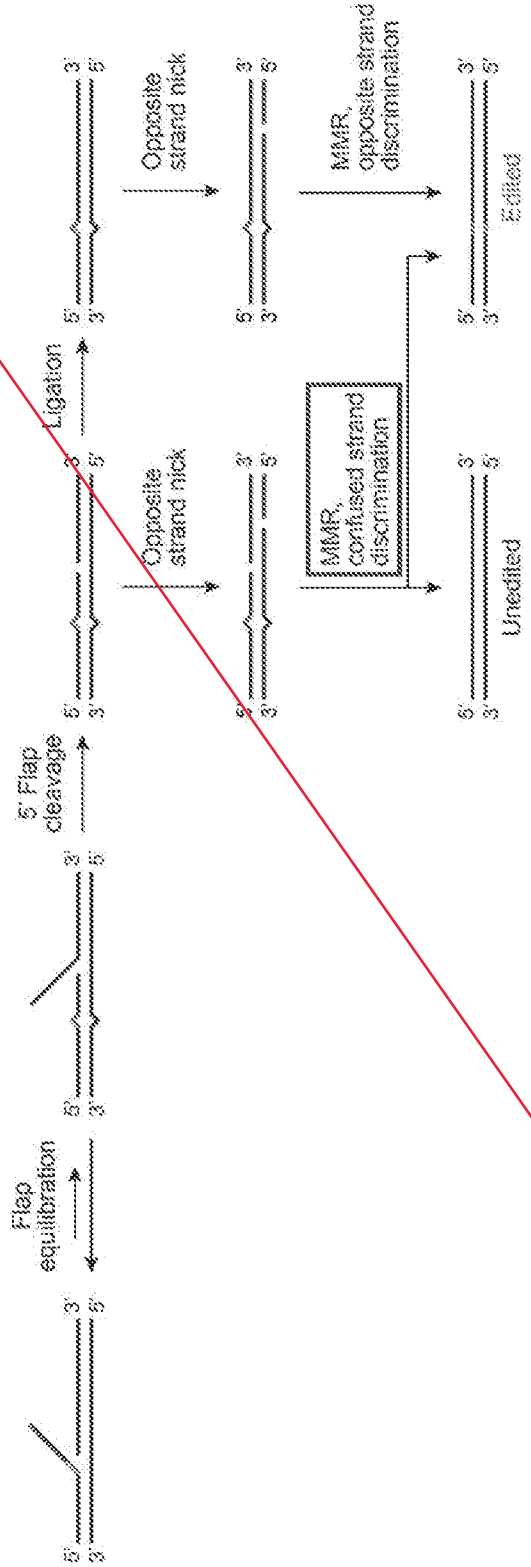

Flap equilibration
5' Flap cleavage
Ligation
Opposite strand nick
Opposite strand nick
MMR, confused strand discrimination
MMR, opposite strand discrimination
Edited
Unedited

FIG. 18

Flap equilibration
5' Flap cleavage
Opposite strand nick
MutS recognition, MutL incision
5' excision by ExoI
DSB repair
Unintended insertions and deletions
5'
3'

FIG. 19A

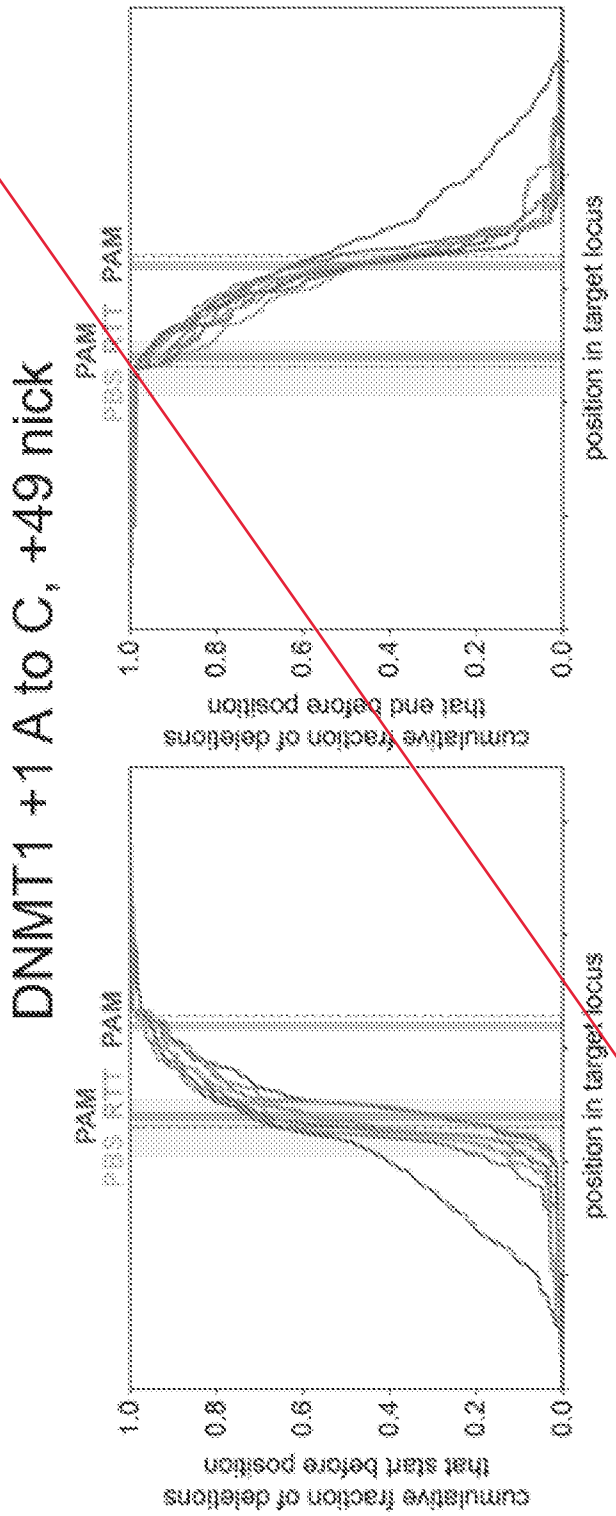

DNMT1 +1 A to C, +49 nick
cumulative fraction of deletions that start before position
cumulative fraction of deletions that end before position
1.0
0.8
0.6
0.4
0.2
0.0
PAM
RBS
NTT
PAM
position in target locus
PE3
PE3, +MLH1neg

FIG. 19B

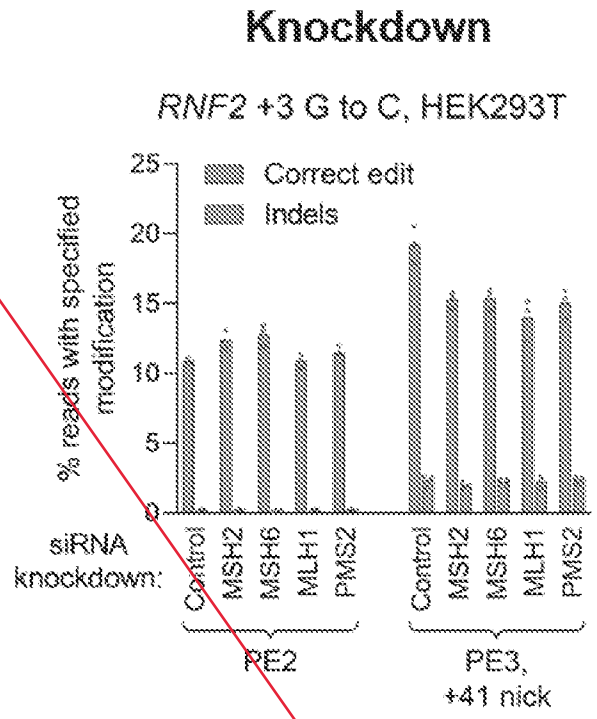


FIG. 20A

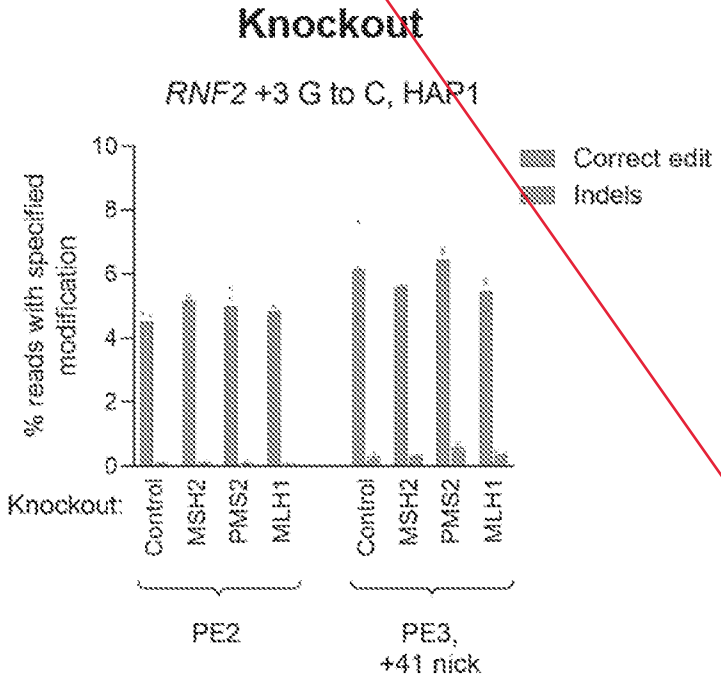


FIG. 20B

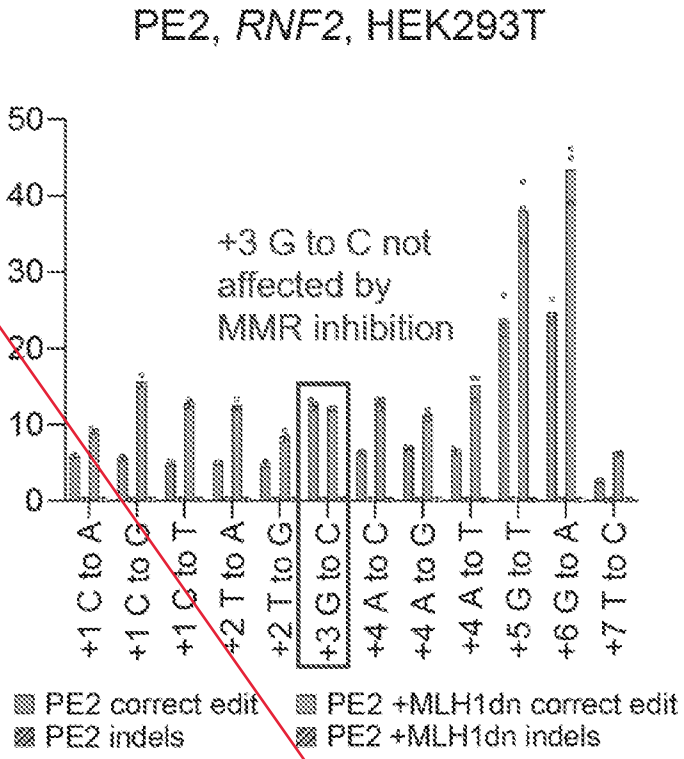


FIG. 21A

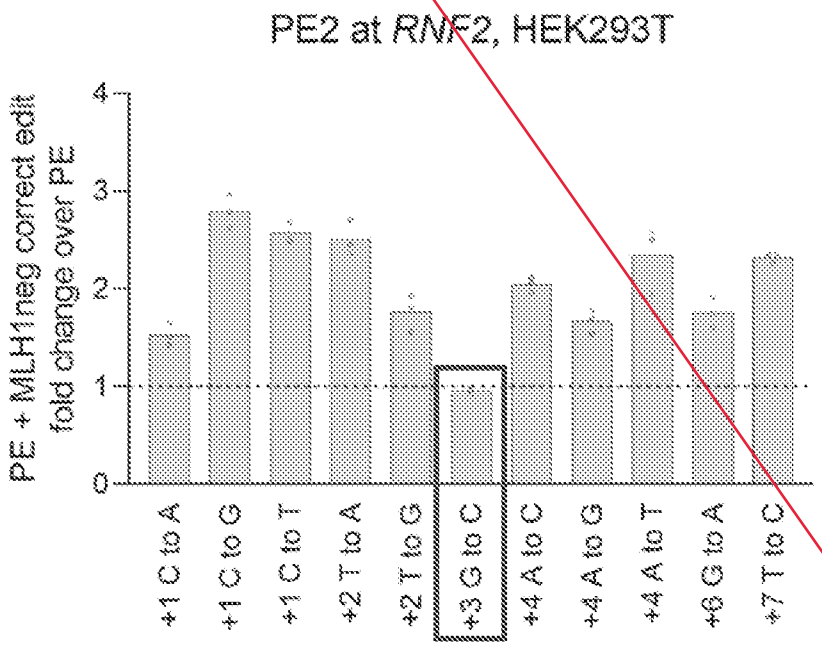


FIG. 21B

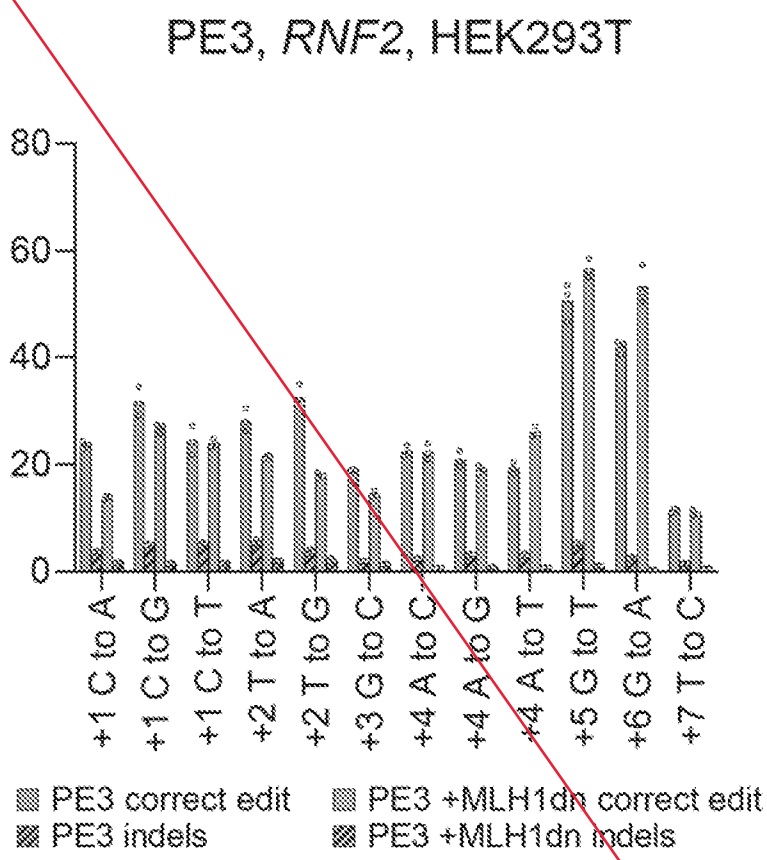

PE3, RNF2, HEK293T
80
60
40
20
0
+1 C to A
+1 C to G
+1 C to T
+2 T to A
+2 T to G
+3 G to C
+4 A to C
+4 A to G
+4 A to T
+5 G to T
+6 G to A
+7 T to C
PE3 correct edit
PE3 indels
PE3 +MLH1dn correct edit
PE3 +MLH1dn indels

FIG. 21C

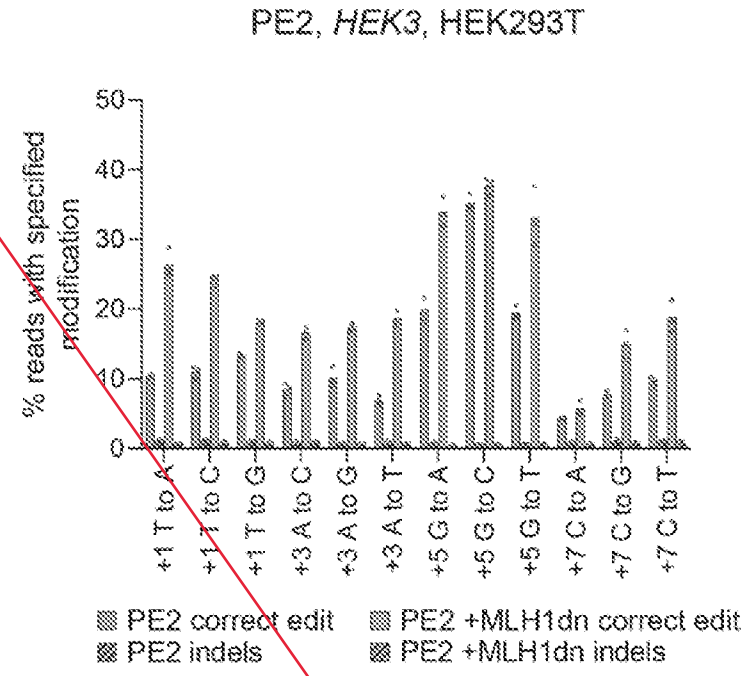


FIG. 22A

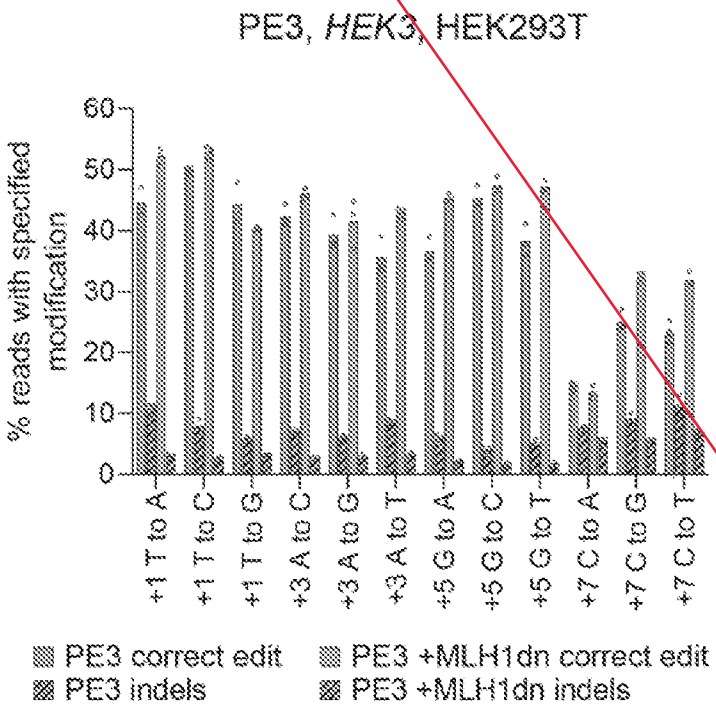


FIG. 22B

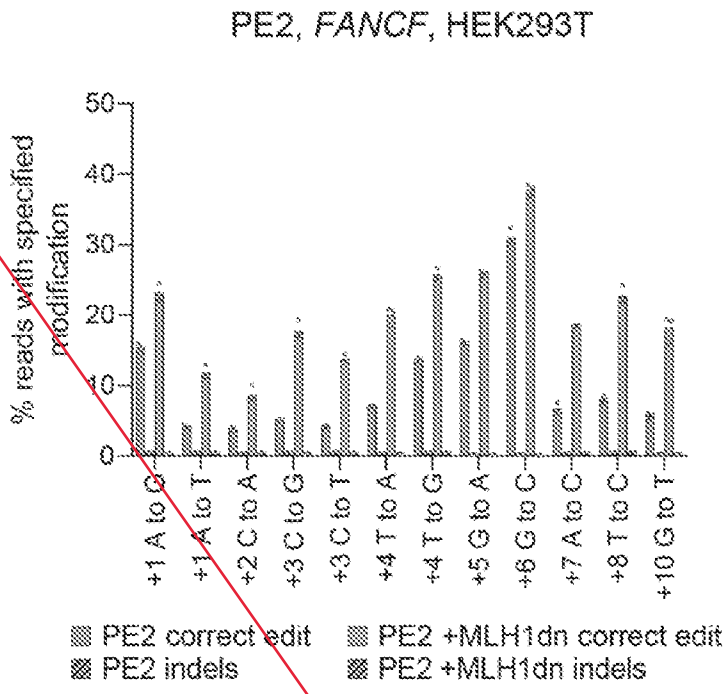


FIG. 23A

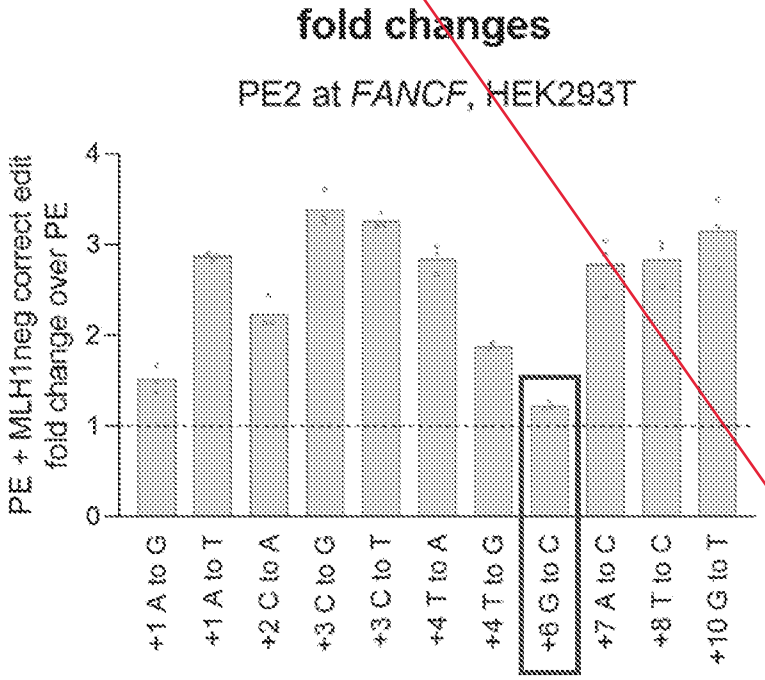


FIG. 23B

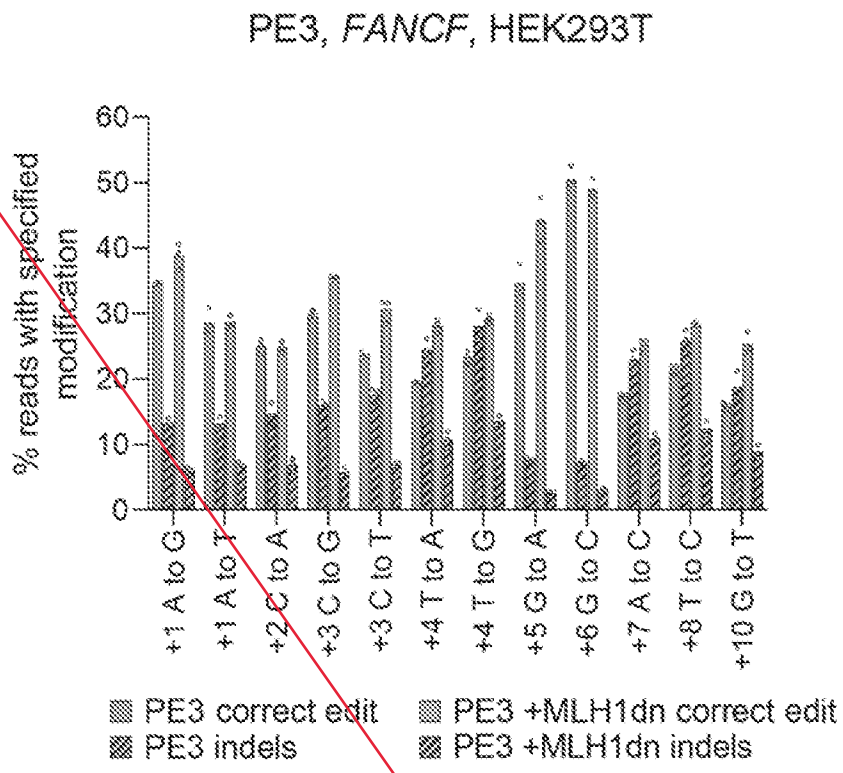


FIG. 23C

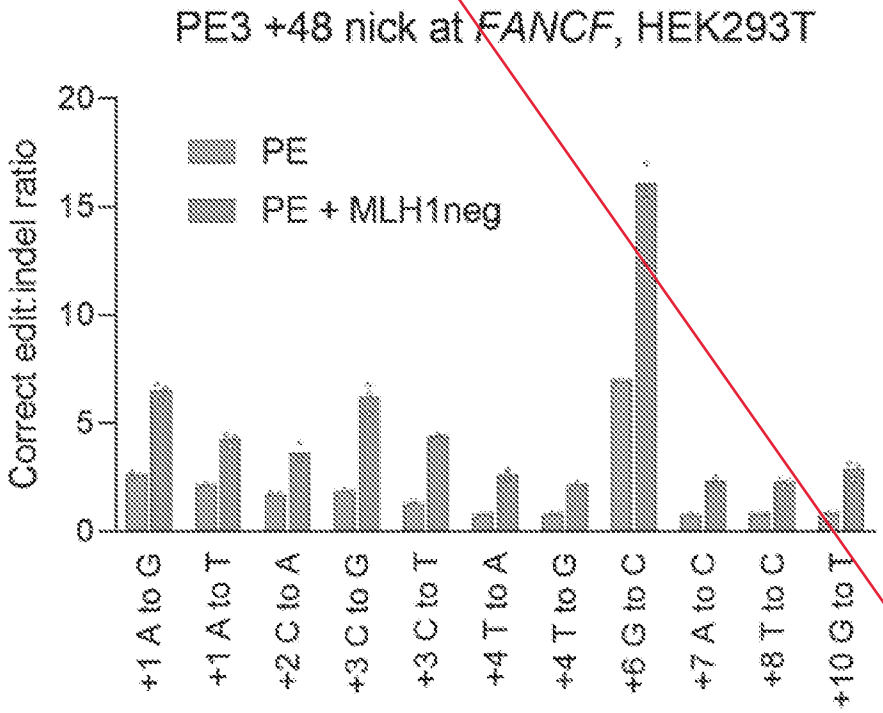


FIG. 23D

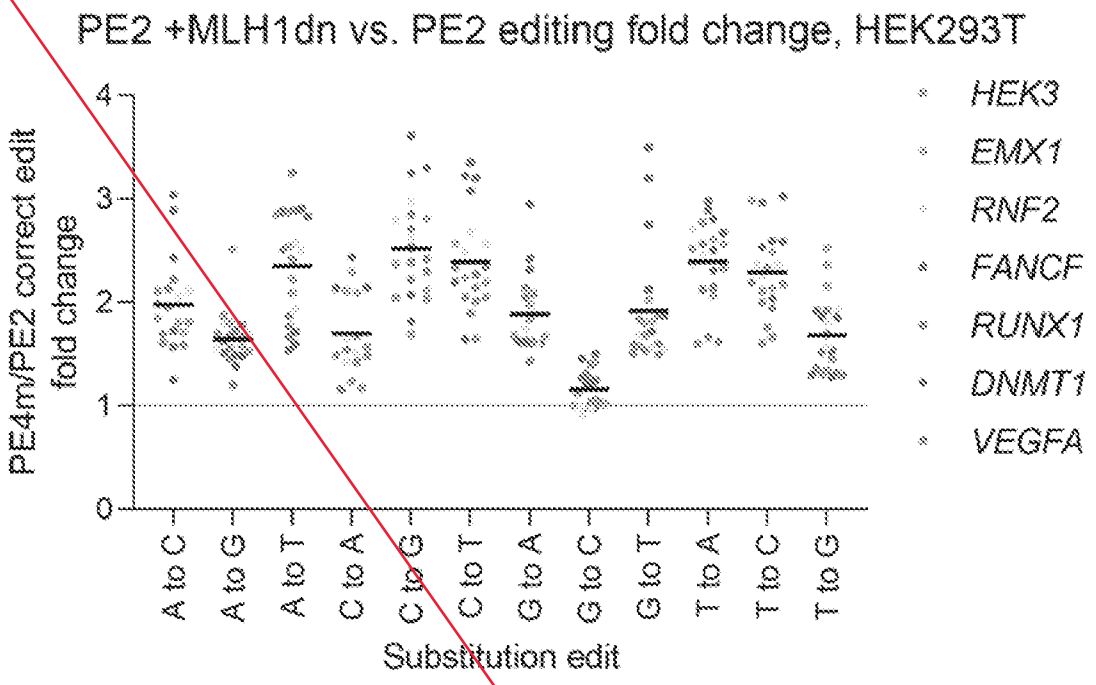


FIG. 24A

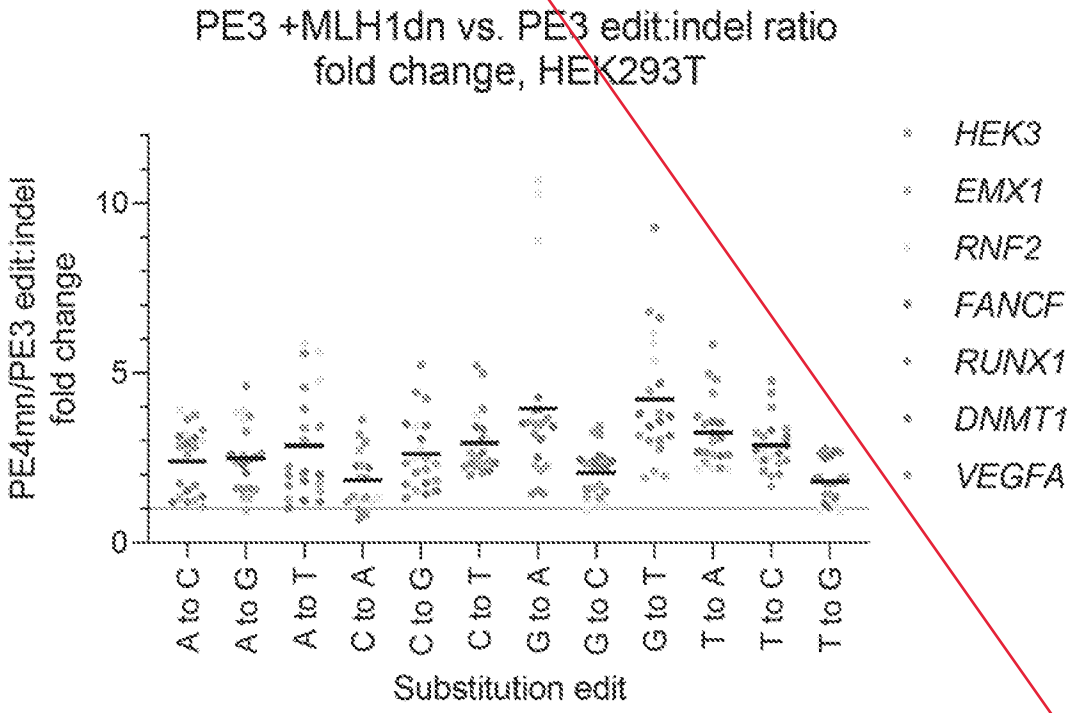


FIG. 24B

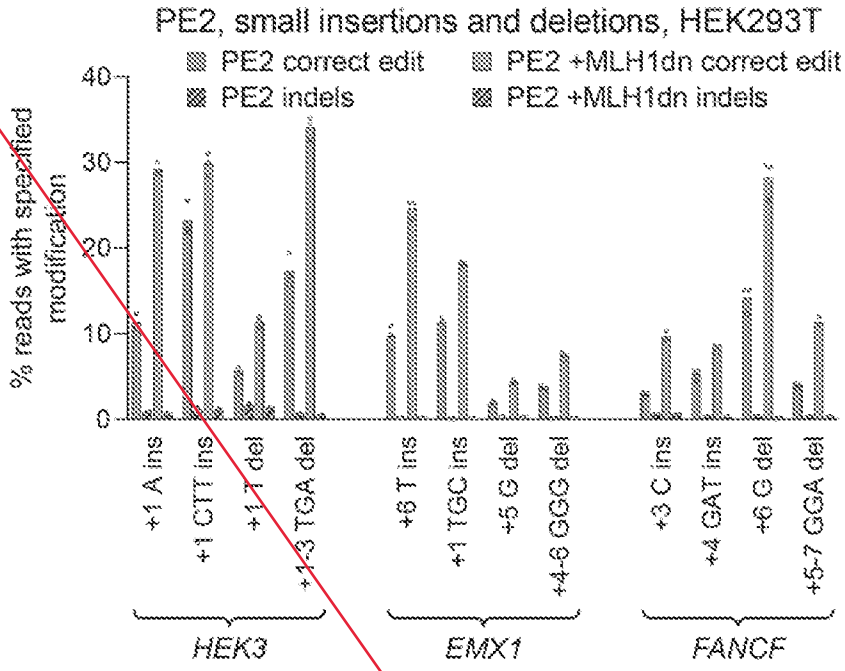


FIG. 25A

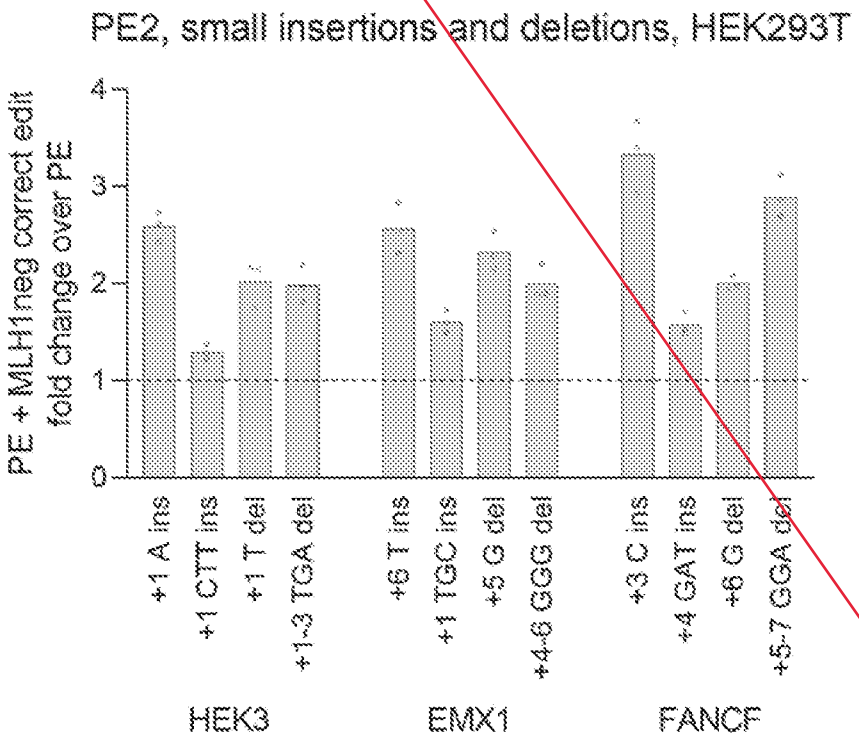


FIG. 25B

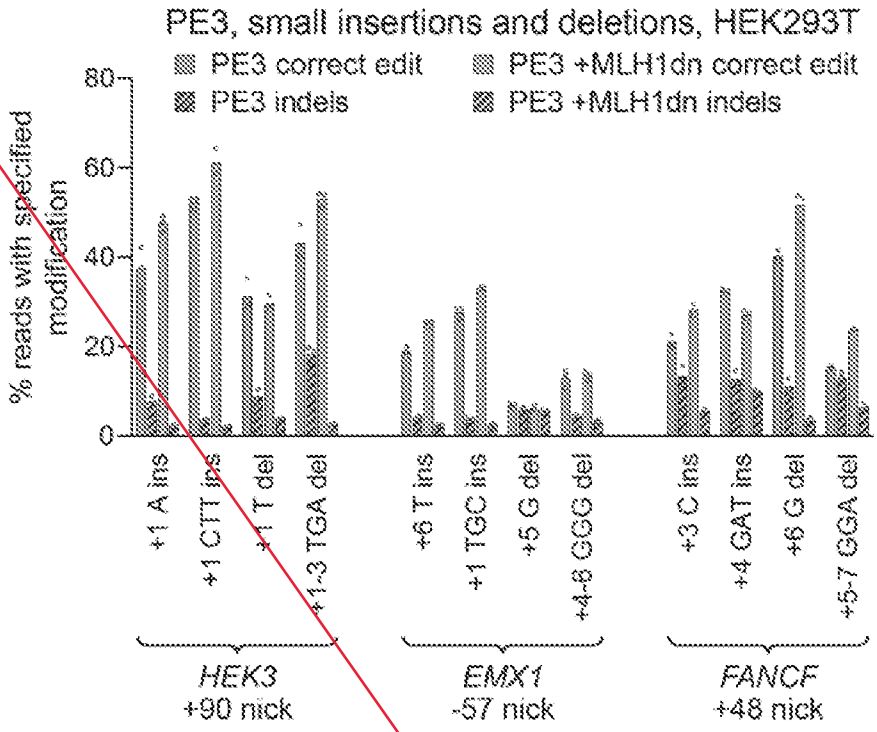


FIG. 25C

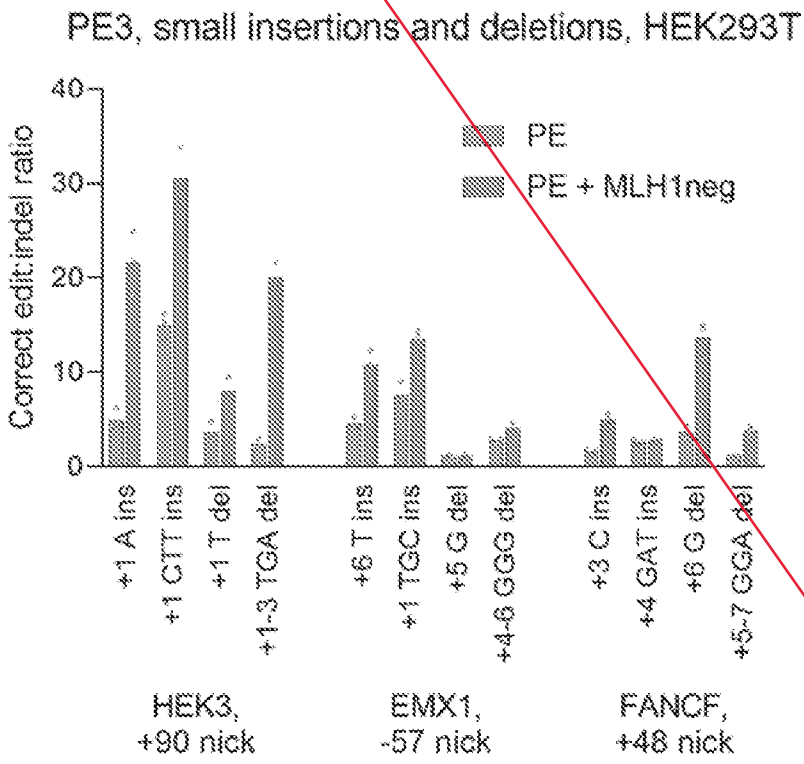


FIG. 25D

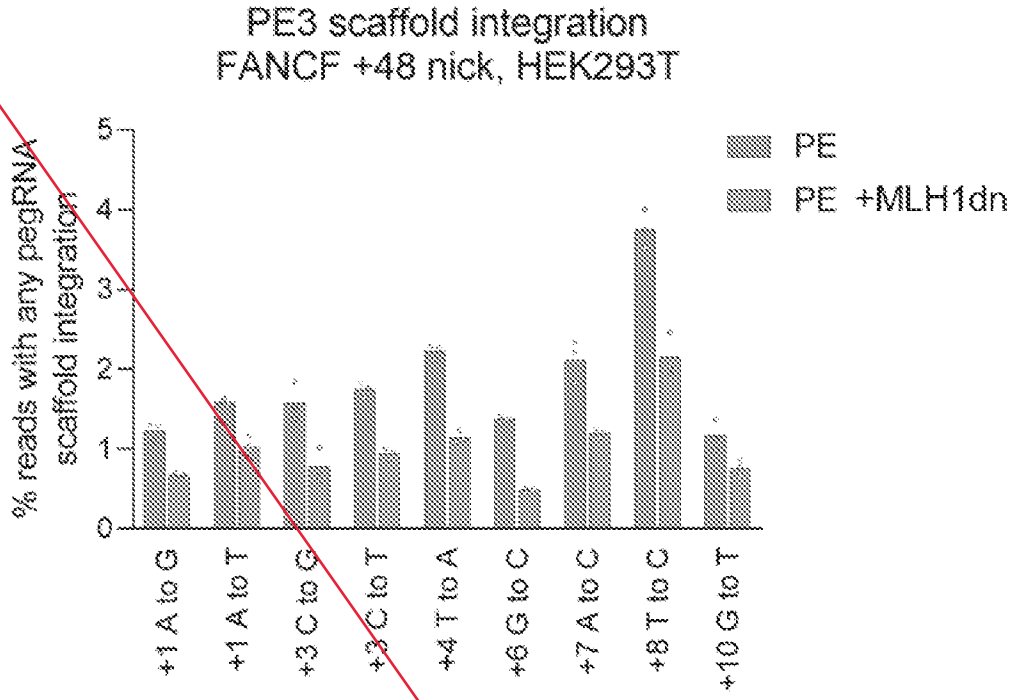


FIG. 26A

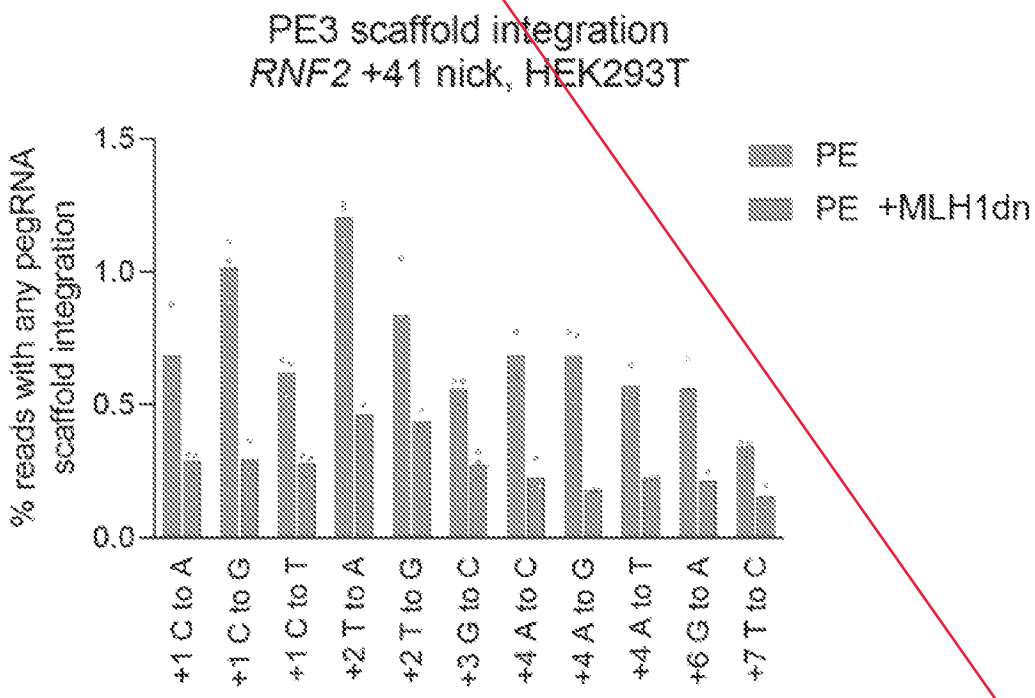


FIG. 26B

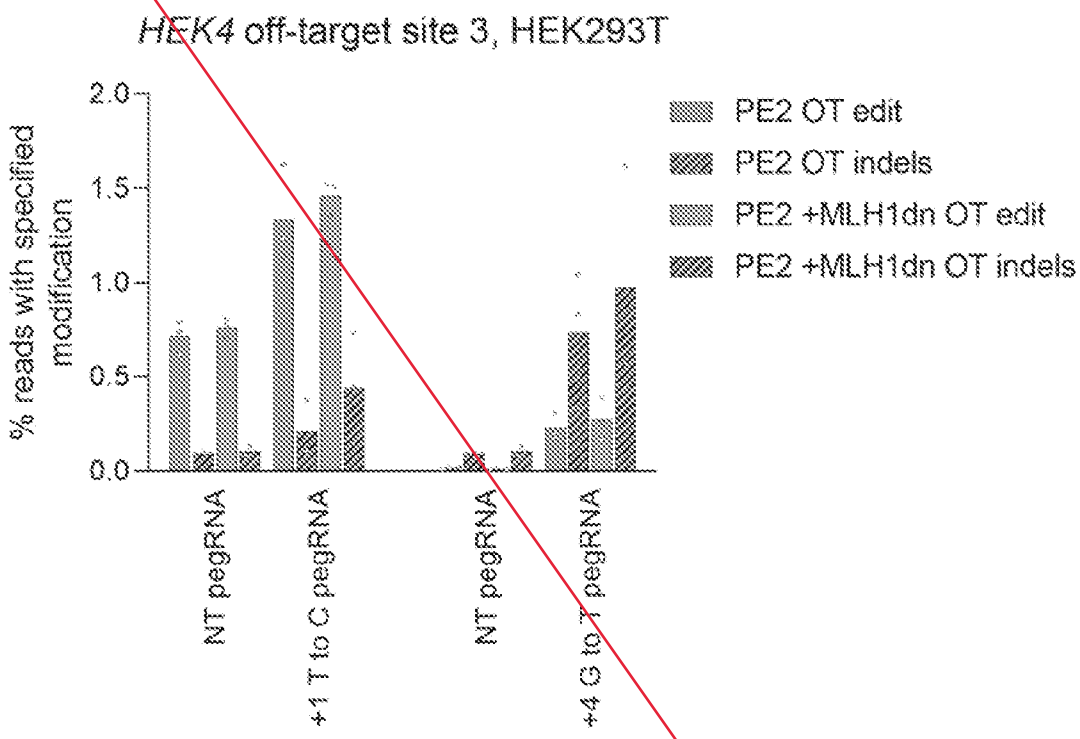

HEK4 off-target site 3, HEK293T
% reads with specified modification
2.0
1.5
1.0
0.5
0.0
PE2 OT edit
PE2 OT indels
PE2 +MLH1dn OT edit
PE2 +MLH1dn OT indels
NT pegRNA
+1 T to C pegRNA
NT pegRNA
+4 G to T pegRNA

FIG. 27

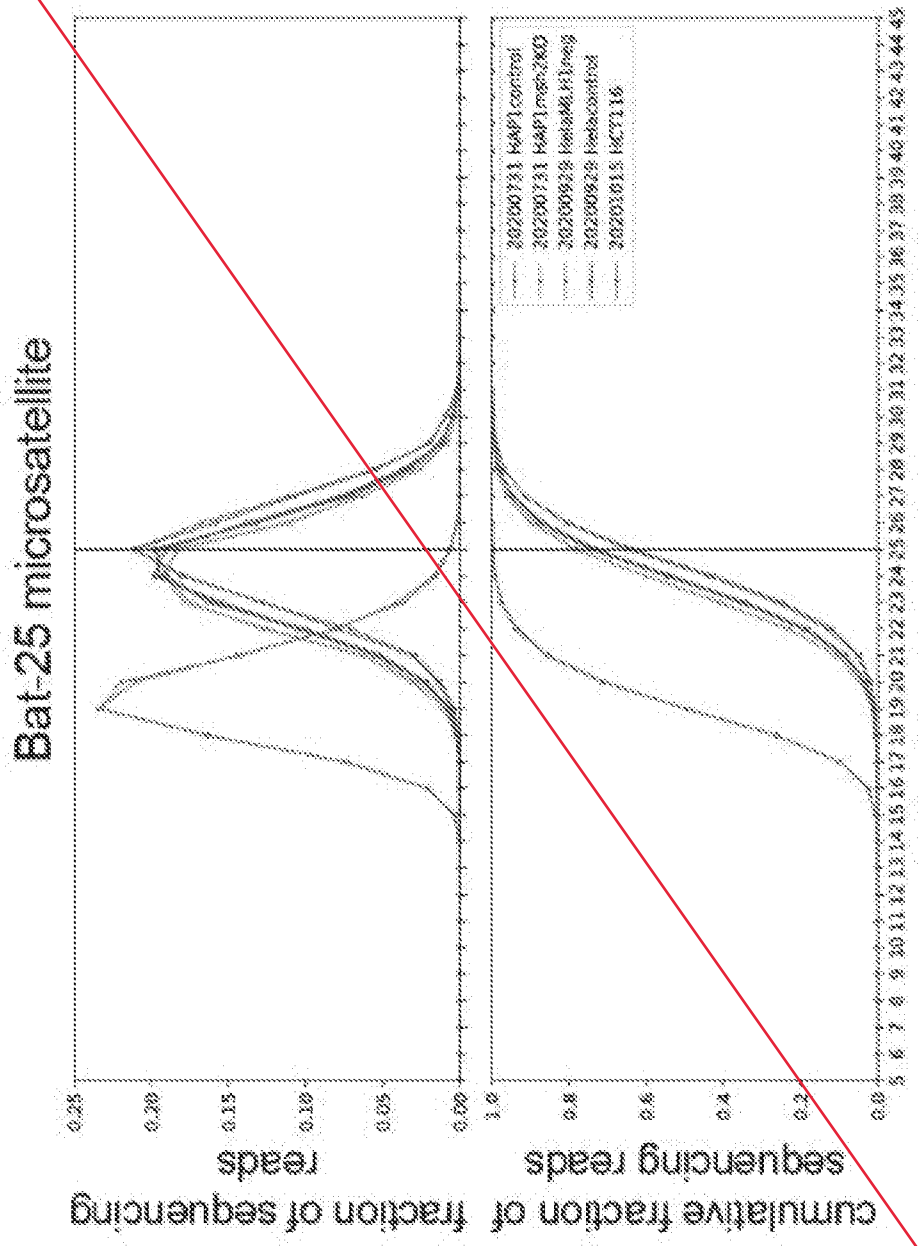

Bat-25 microsatellite
fraction of sequencing reads
cumulative fraction of sequencing reads
distribution of microsatellite lengths (nt)
20200731 HAP1control
20200731 HAP1msh2KO
20200828 Hela4h Rong
20200828 Hela control
20201015 HCT116

FIG. 28A

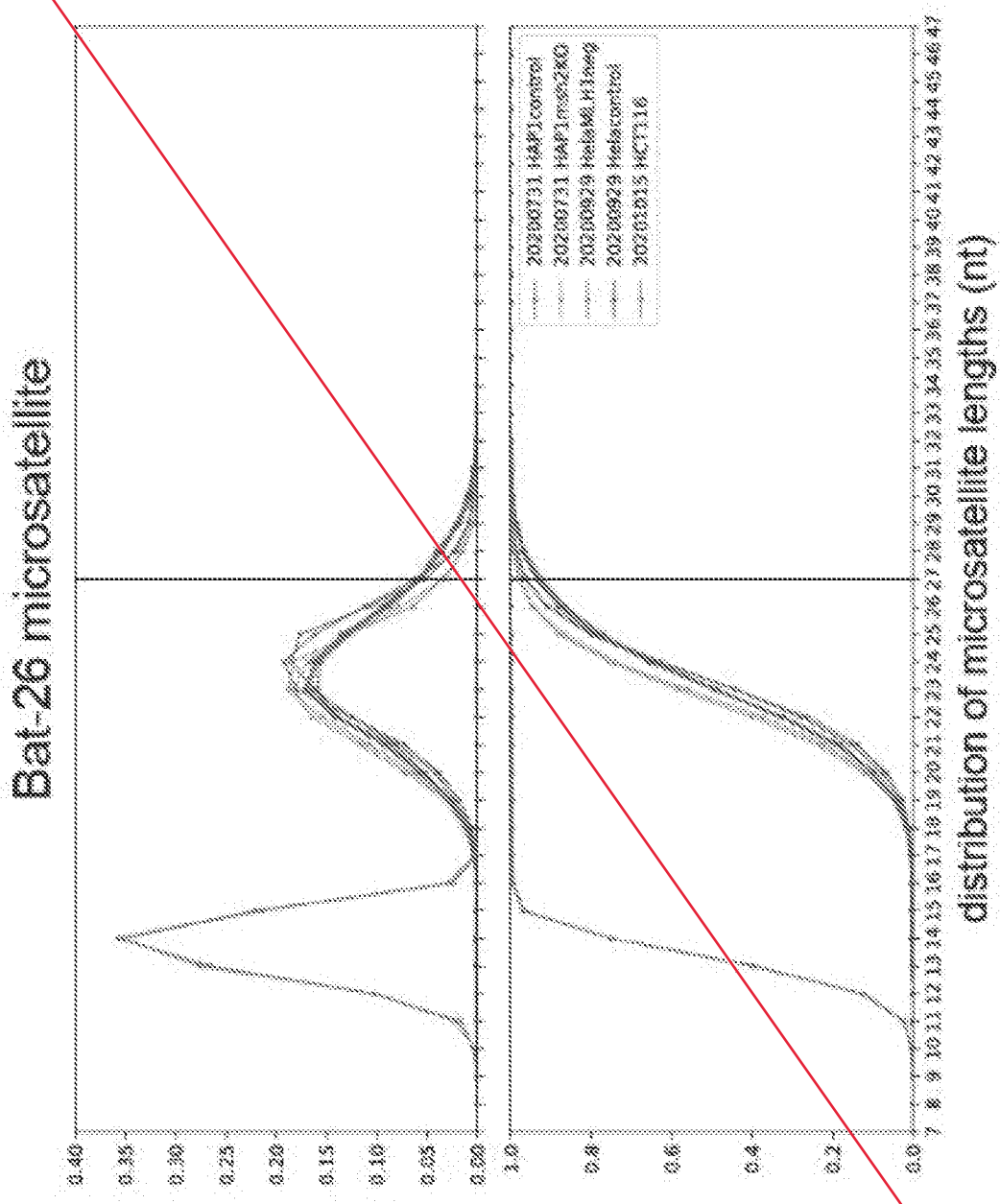

Bat-26 microsatellite
distribution of microsatellite lengths (nt)
20200731 HAP1control
20200731 HAP1msi2KO
20200929 HeLaMLH1wt
20200929 HeLacontrol
20201015 HCT116

FIG. 28B

HEK4 +2 G to T, HEK293T
PE correct edit
PE indels
PE +MLH1dn correct edit
PE +MLH1dn indels
% reads with specified modification
0
5
10
15
20
Nick position:
none
-52
-26
-19
+52
+74
+131
+169

FIG. 29

Prime editing at disease sites, HEK293T
% reads with specified modification
0
20
40
60
80
PE correct edit
PE indels
PE + MLH1dn correct edit
PE + MLH1dn indels
PE2 PE3
CFTR F508Δ install
PE2 PE3
CFTR W1282X install
PE2 PE3
APOE R136S install
PE2 PE3
HBB E6V install
PE2 PE3
CDKL5 c1412 A to G

FIG. 30

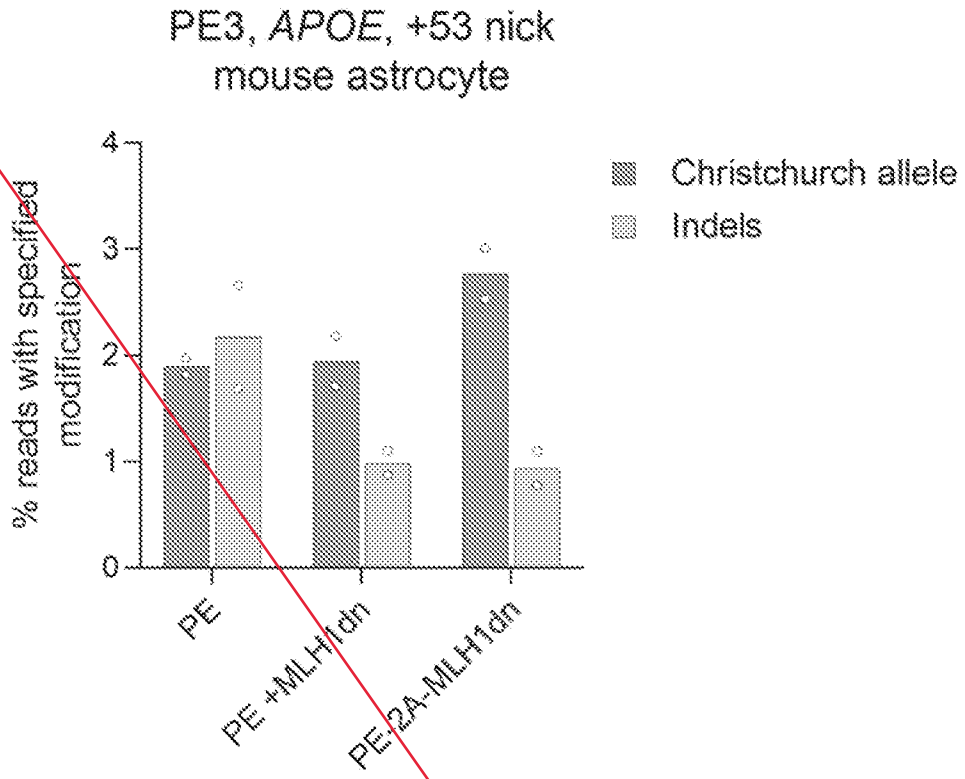

PE3, APOE, +53 nick
mouse astrocyte
% reads with specified modification
0
1
2
3
4
Christchurch allele
Indels
PE
PE +MLH1dn
PE-2A-MLH1dn

FIG. 31

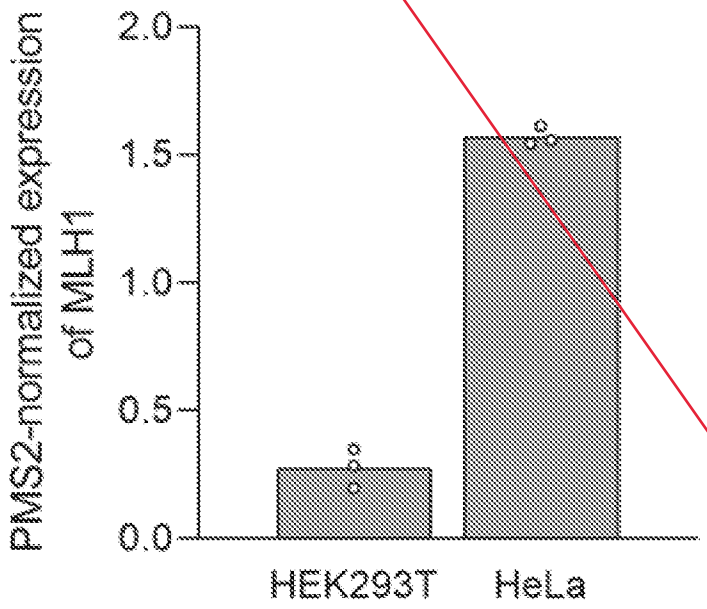

PMS2-normalized expression of MLH1
0.0
0.5
1.0
1.5
2.0
HEK293T
HeLa

FIG. 32

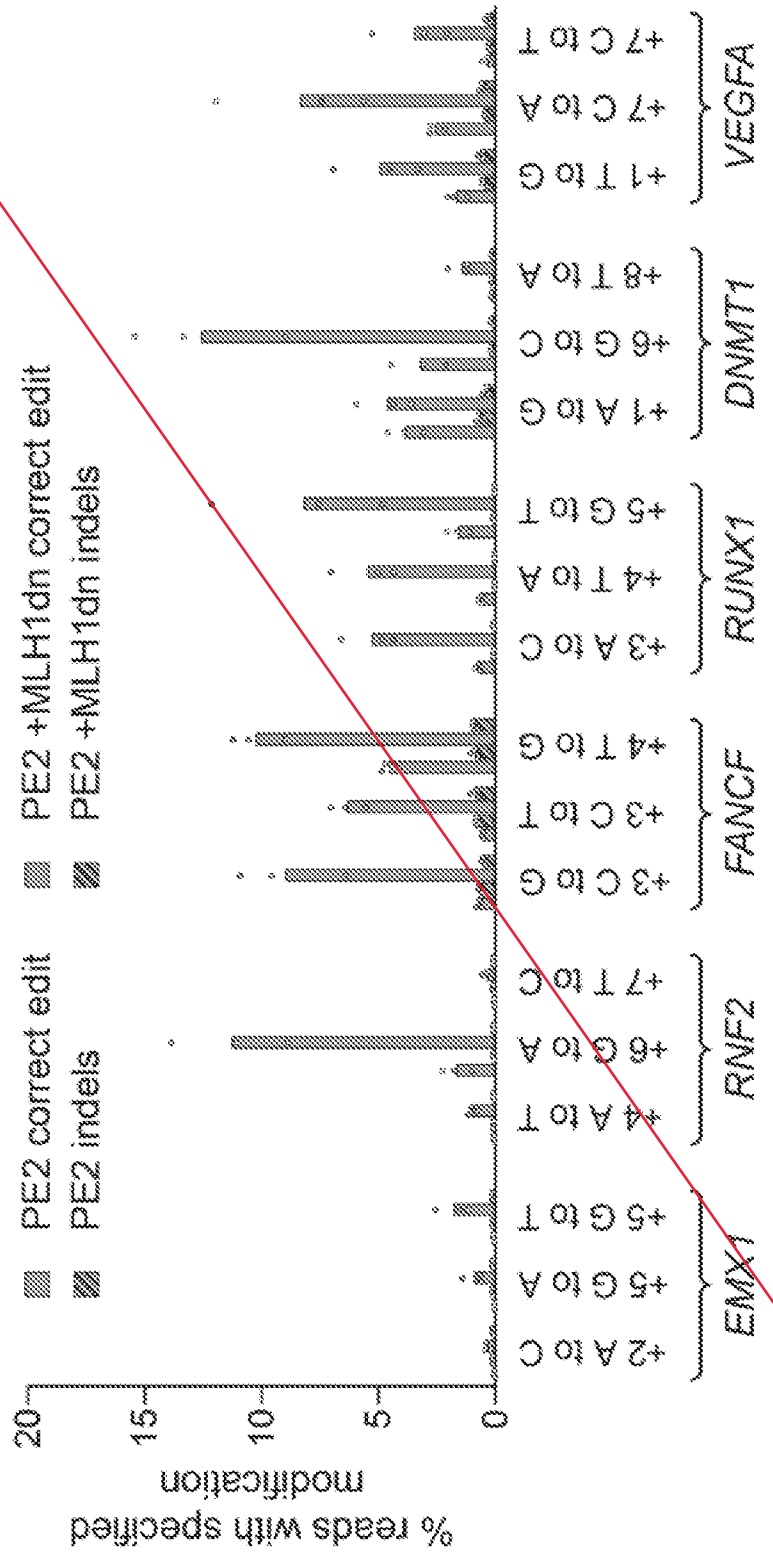

PE2, HeLa
PE2 correct edit
PE2 indels
PE2 +MLH1dn correct edit
PE2 +MLH1dn indels
% reads with specified modification
0
5
10
15
20
+2 A to C
+5 G to A
+5 G to T
+4 A to T
+6 G to A
+7 T to C
+3 C to G
+3 C to T
+4 T to G
+3 A to C
+4 T to A
+5 G to T
+1 A to G
+6 G to C
+8 T to A
+1 T to G
+7 C to A
+7 C to T
EMX1
RNF2
FANCF
RUNX1
DNMT1
VEGFA

FIG. 33A

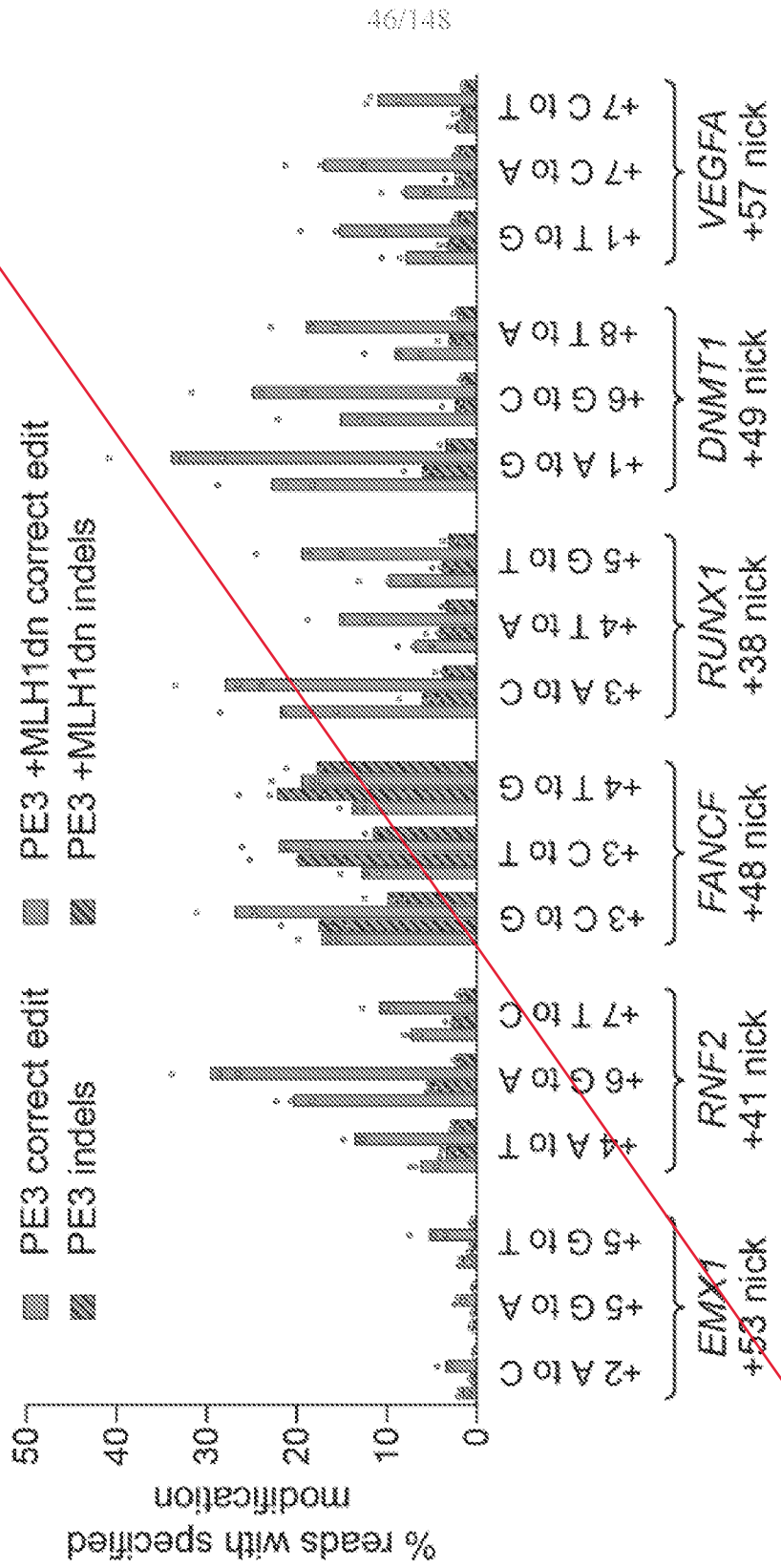

PE3, HeLa
PE3 correct edit
PE3 indels
PE3 +MLH1dn correct edit
PE3 +MLH1dn indels
% reads with specified modification
0
10
20
30
40
50
+2 A to C
+5 G to A
+5 G to T
+4 A to T
+6 G to A
+7 T to C
+3 C to G
+3 C to T
+4 T to G
+3 A to C
+4 T to A
+5 G to T
+1 A to G
+6 G to C
+8 T to A
+1 T to G
+7 C to A
+7 C to T
EMX1 +53 nick
RNF2 +41 nick
FANCF +48 nick
RUNX1 +38 nick
DNMT1 +49 nick
VEGFA +57 nick

FIG. 33B

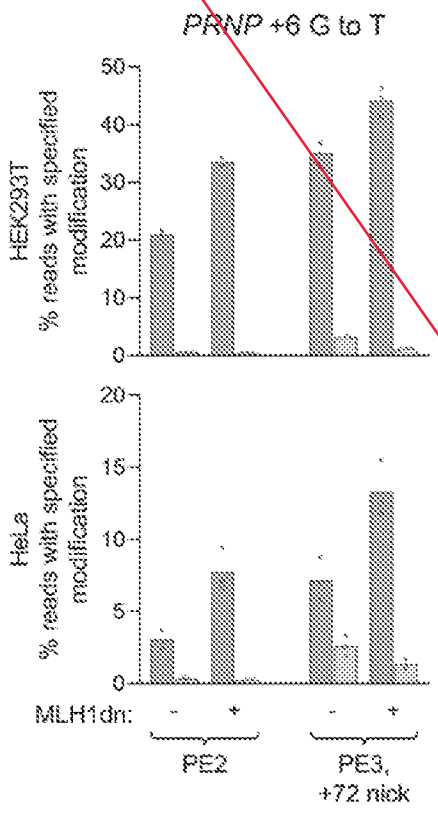


FIG. 34A

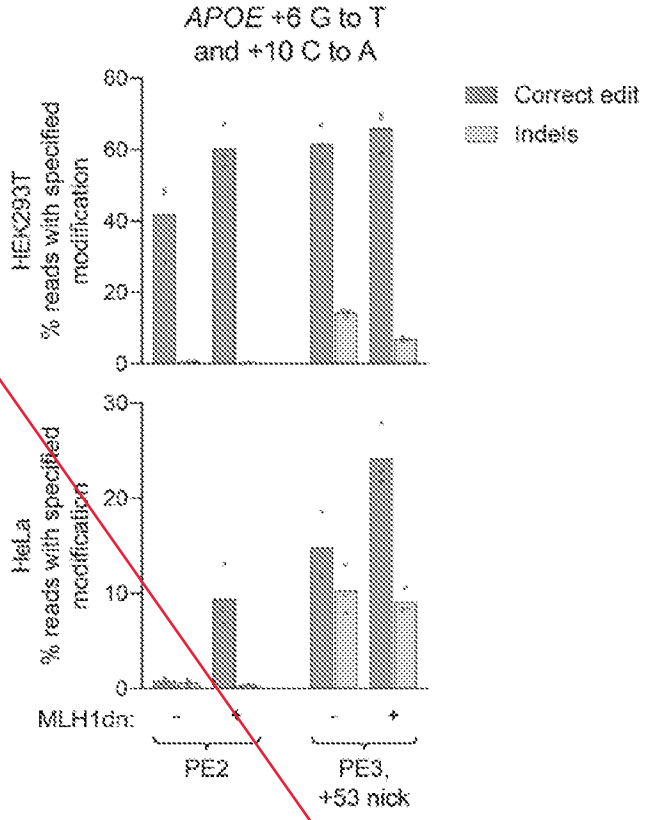


FIG. 34B

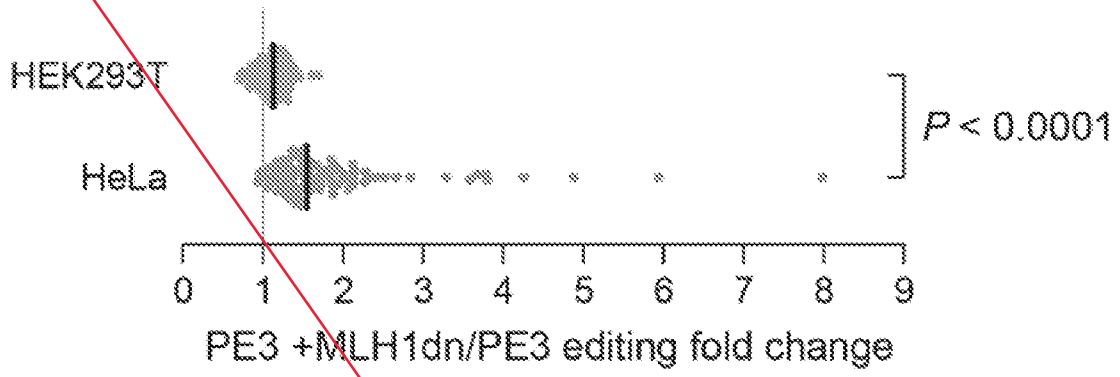


FIG. 35A

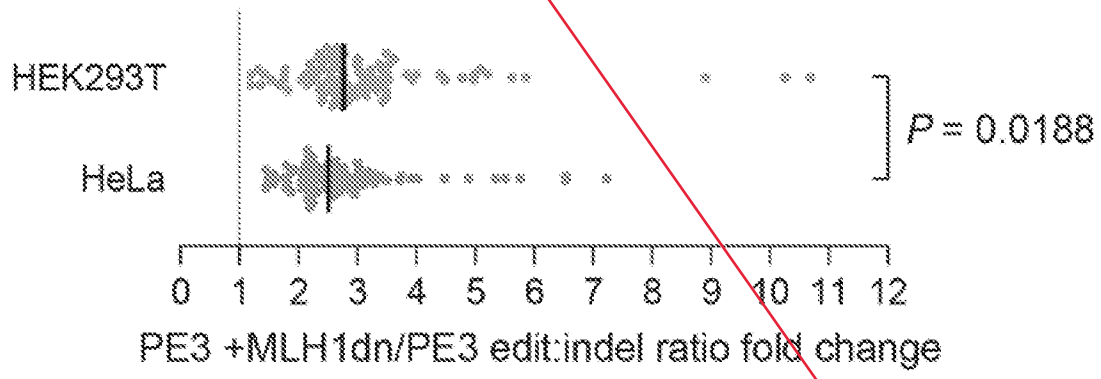


FIG. 35B

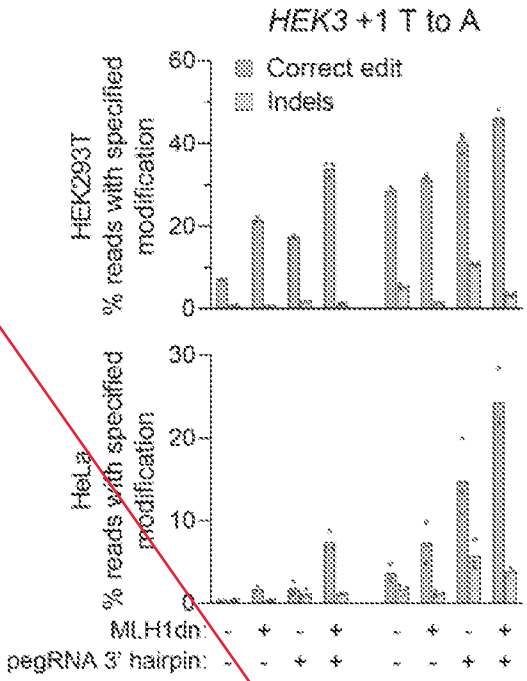


FIG. 36A

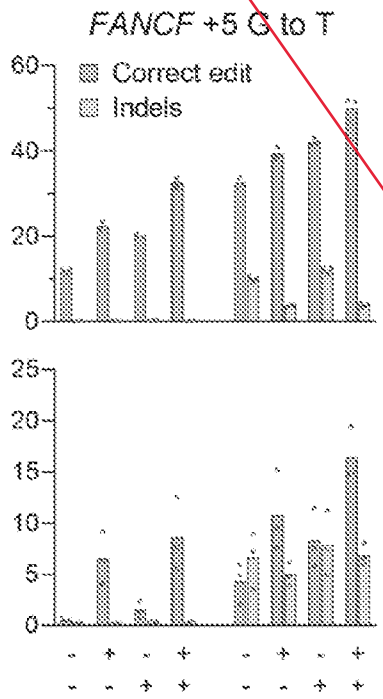


FIG. 36B

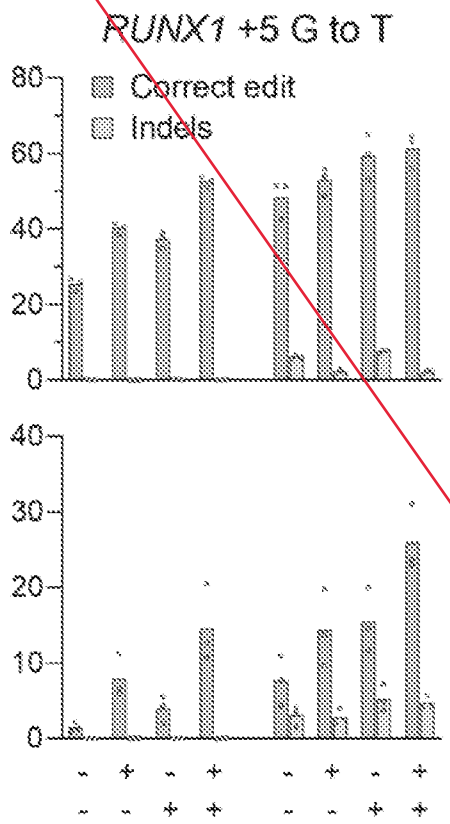


FIG. 36C

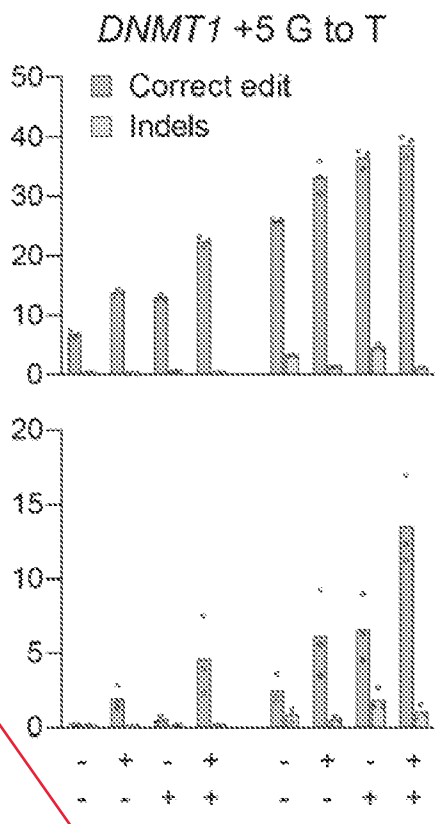


FIG. 36D

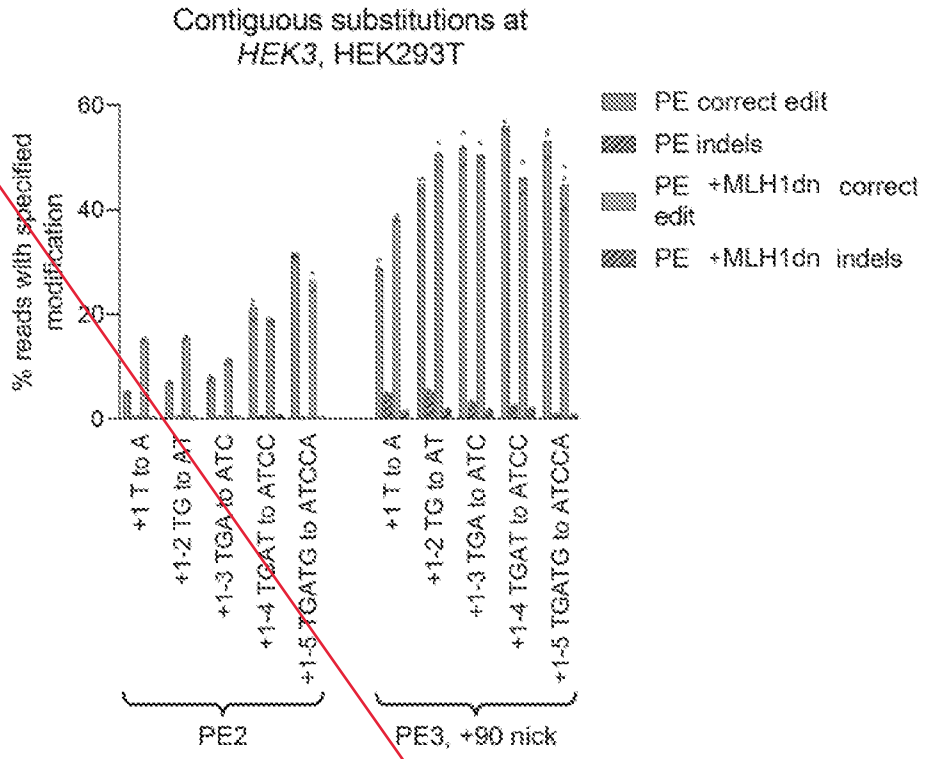


FIG. 37A

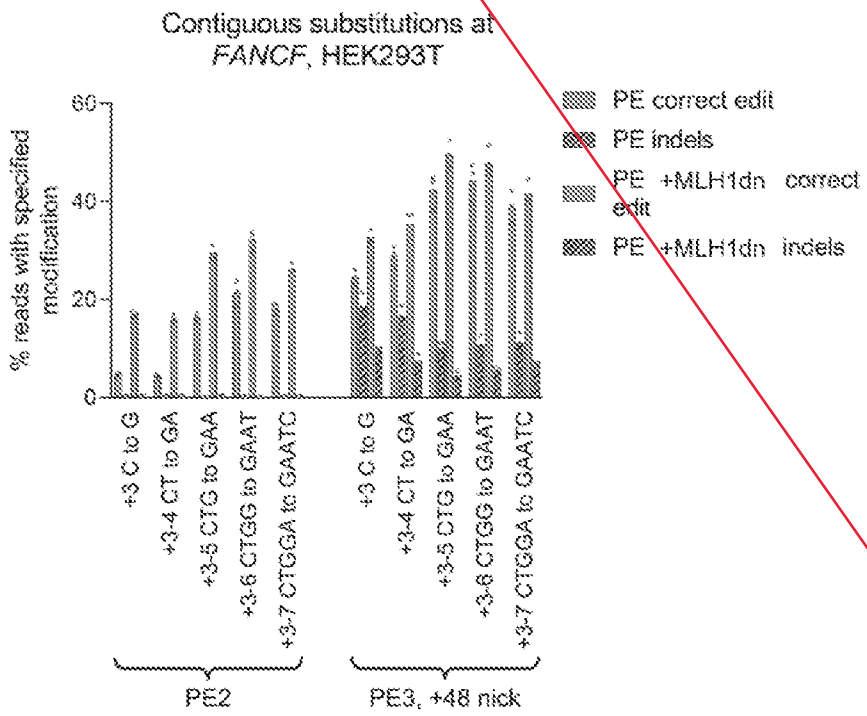


FIG. 37B

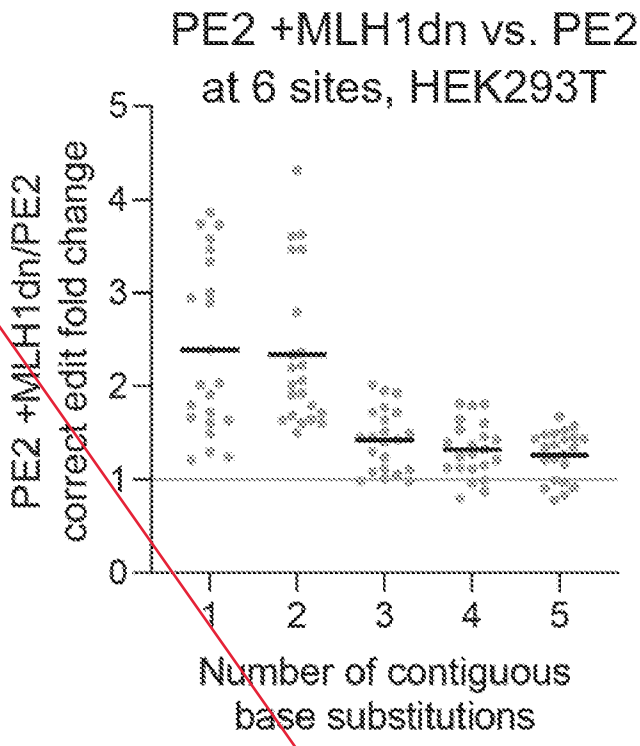


FIG. 38

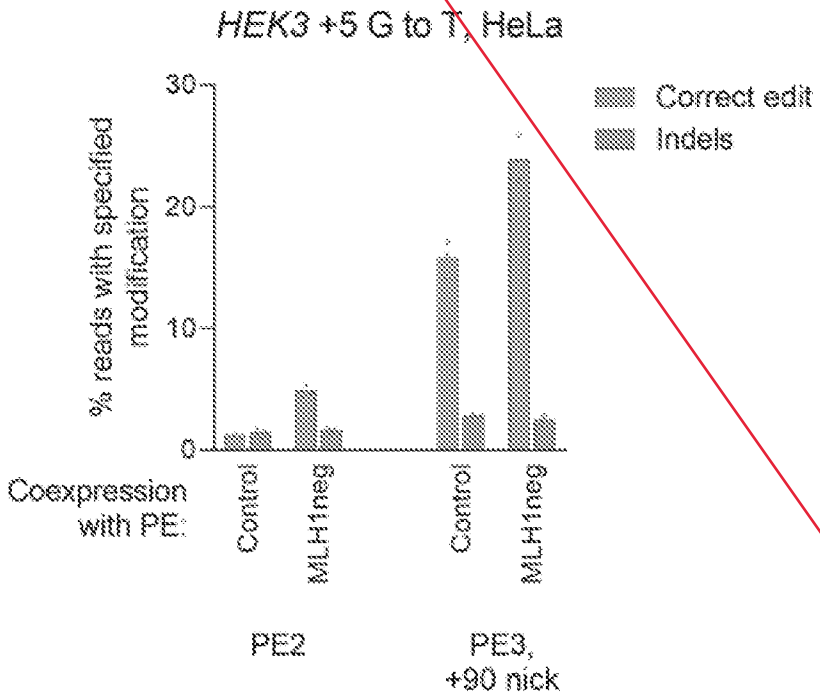


FIG. 39A

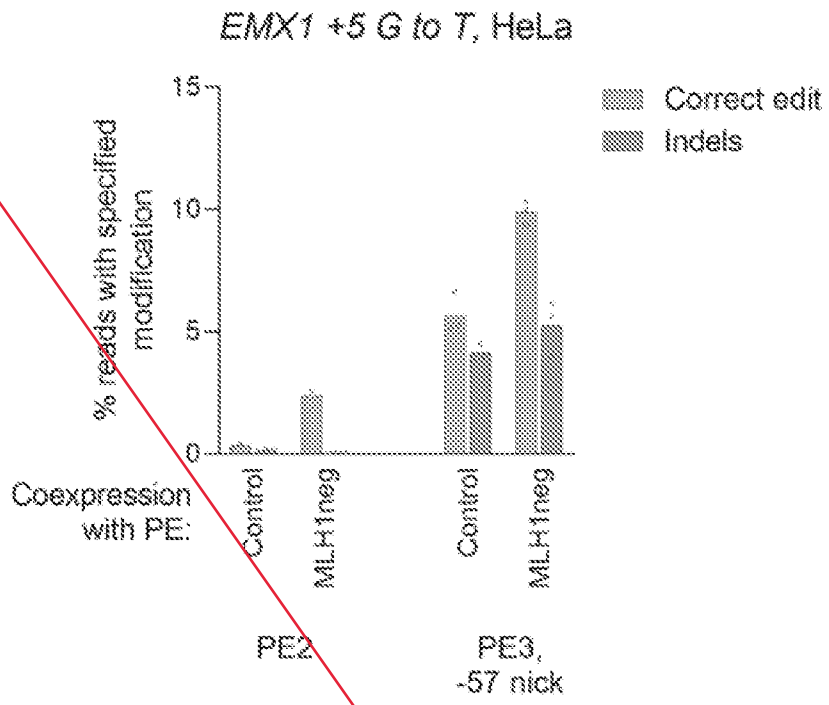


FIG. 39B

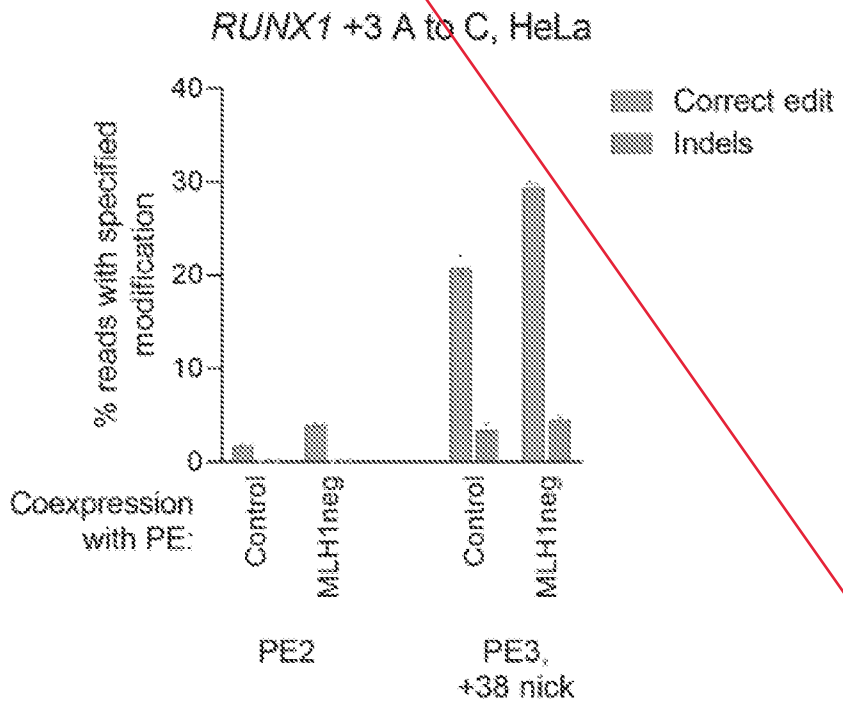


FIG. 39C

FIG. 40A

FIG. 40B

FIG. 40C

FIG. 40D

FIG. 40E

FIG. 40F

FIG. 40G

FIG. 41A

FIG. 41B

Installation of additional edits from nearly-matched scaffold sequence
PBS
RT template
scaffold
protospacer
pegRNA
Alignment to pegRNA
Editing outcome
Target site in genome
Edit position
(SEQ ID NO: 368)
Mismatch
(SEQ ID NO: 369)
Mismatch
1-nt insertion
(SEQ ID NO: 370)

FIG. 41C

Joining of reverse transcribed
sequence at unintended location

FIG. 41D

FIG. 41E

FIG. 41F

FIG. 41G

FIG. 41H

FIG. 41I

FIG. 41J

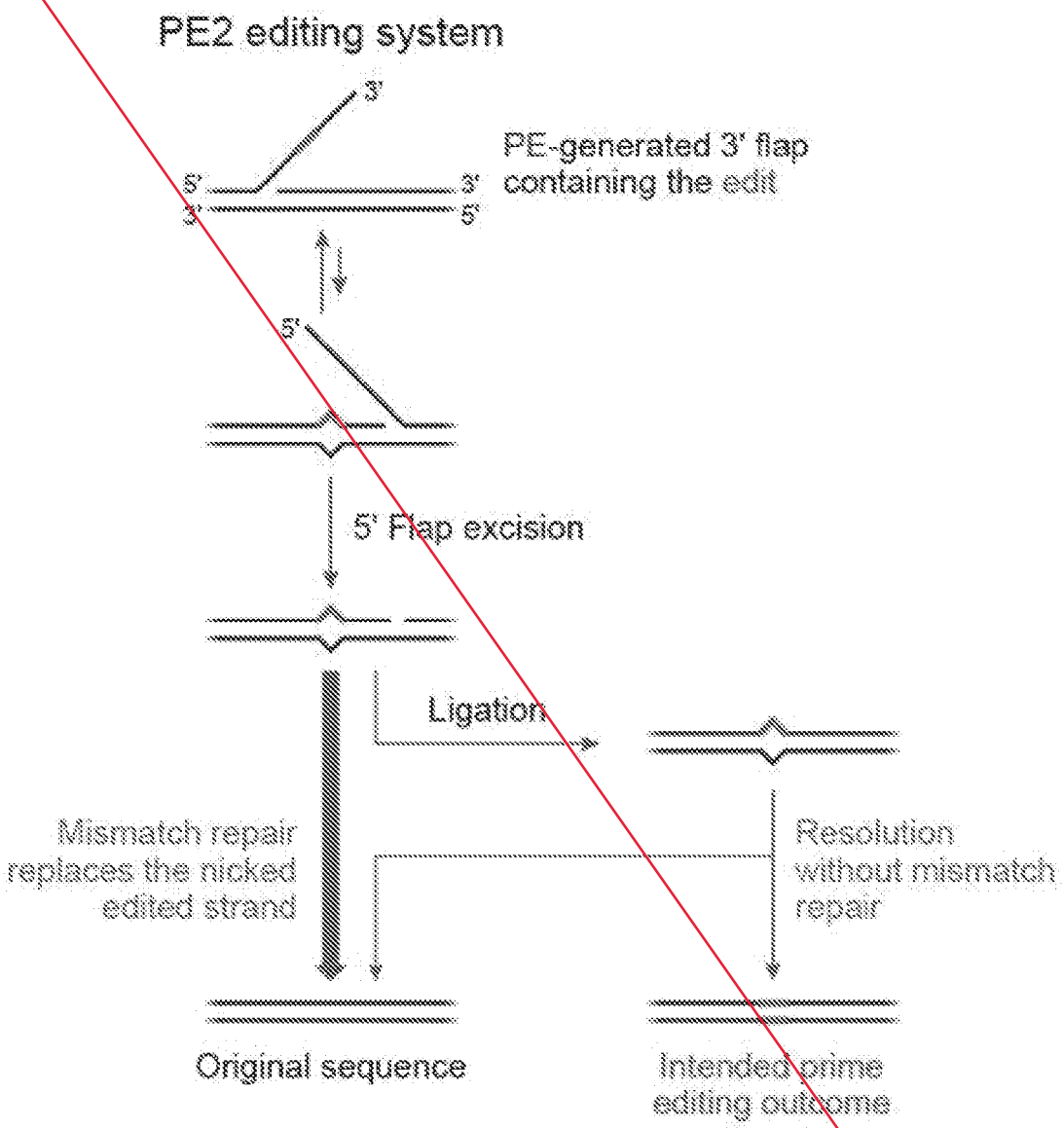

PE2 editing system
3'
5'
3'
3'
5'
PE-generated 3' flap containing the edit
5'
5' Flap excision
Ligation
Mismatch repair replaces the nicked edited strand
Resolution without mismatch repair
Original sequence
Intended prime editing outcome

FIG. 42A

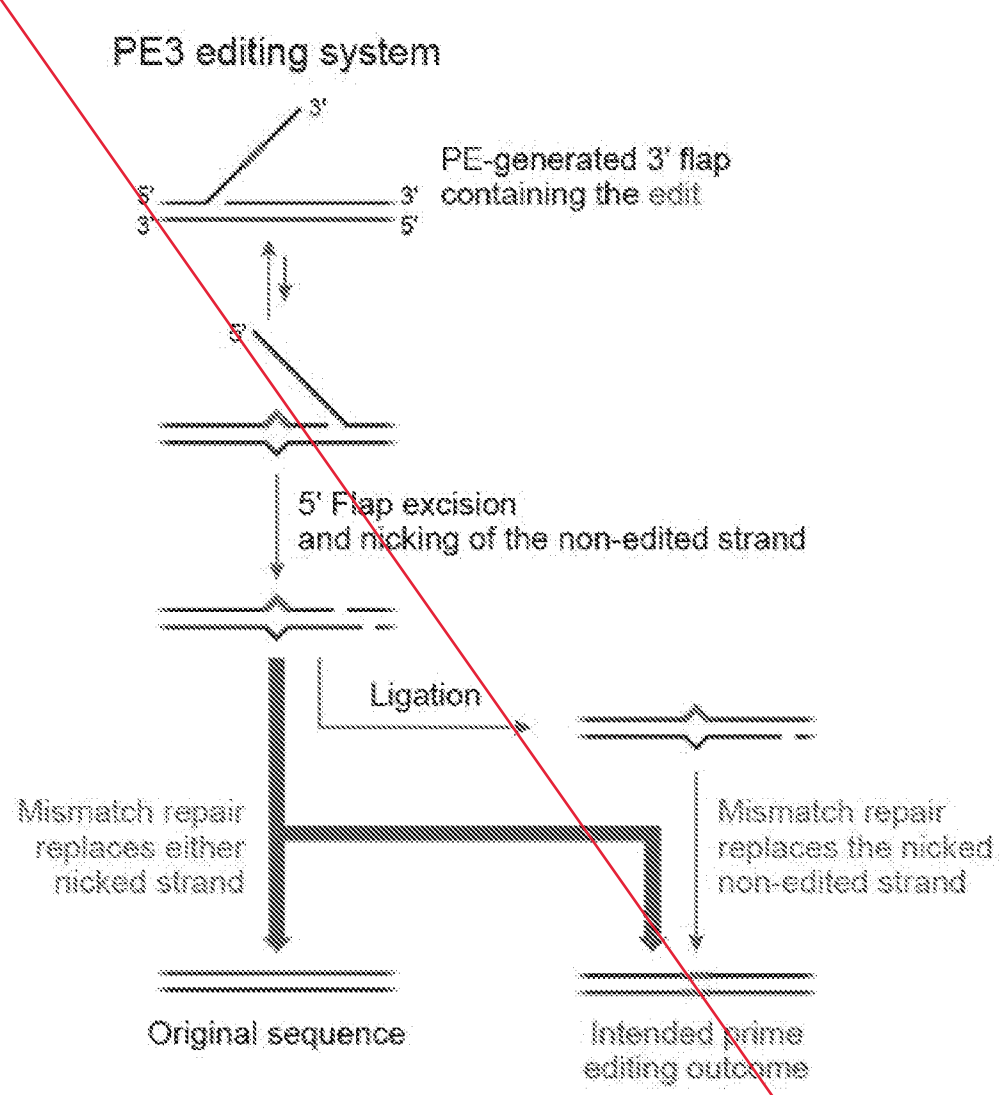

PE3 editing system
3'
5'
3'
3'
5'
PE-generated 3' flap
containing the edit
5'
5' Flap excision
and nicking of the non-edited strand
Ligation
Mismatch repair
replaces either
nicked strand
Mismatch repair
replaces the nicked
non-edited strand
Original sequence
Intended prime
editing outcome

FIG. 42B

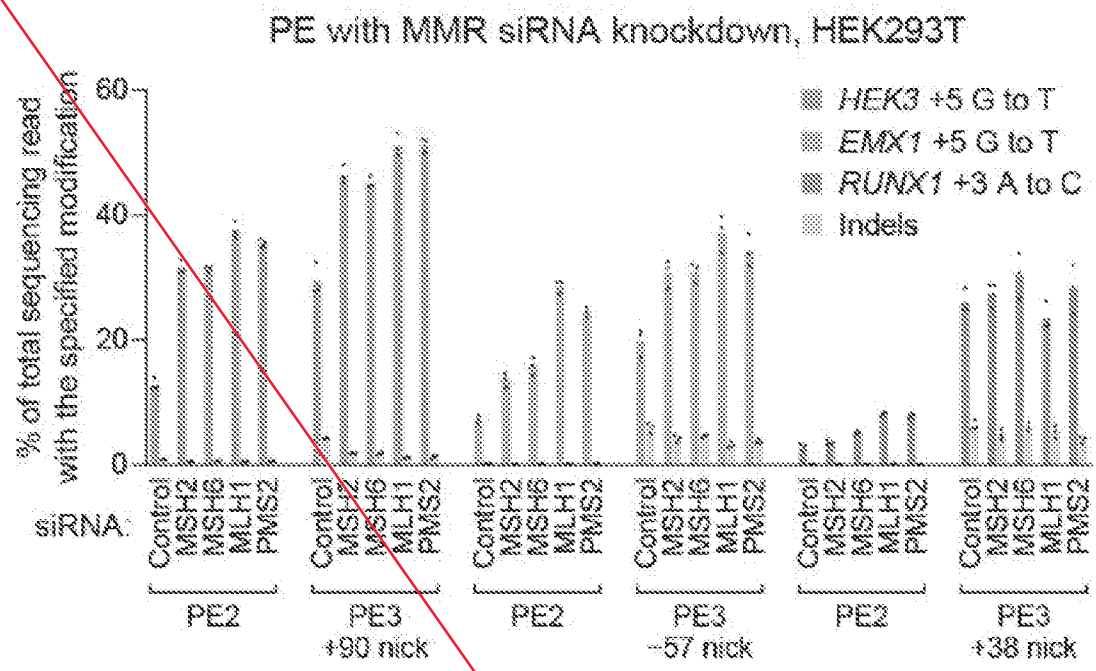


FIG. 42C

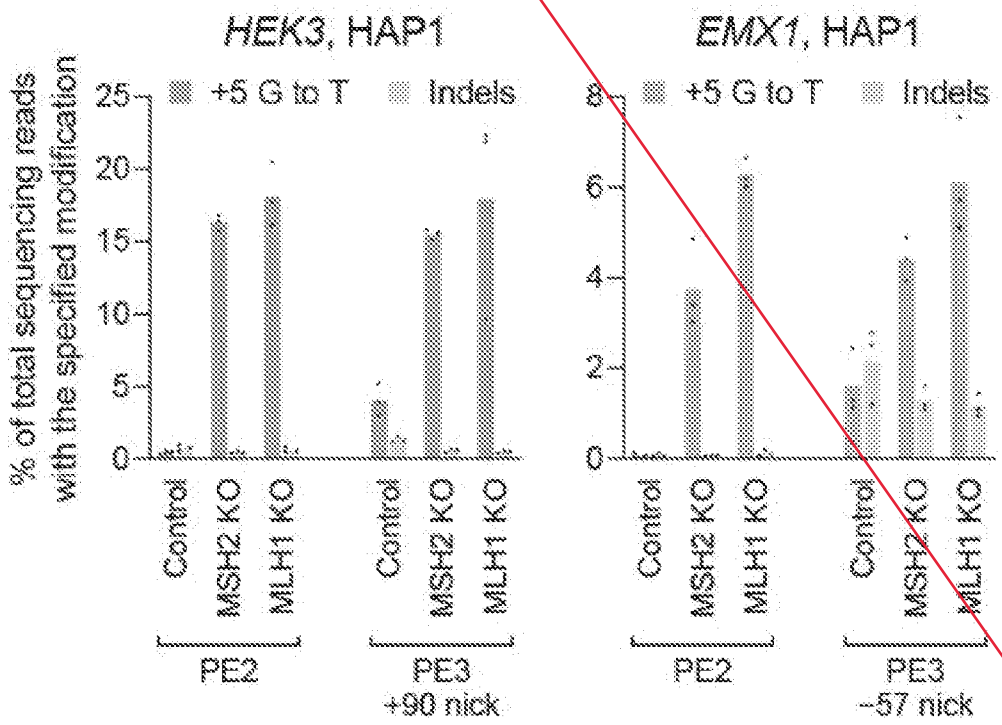


FIG. 42D

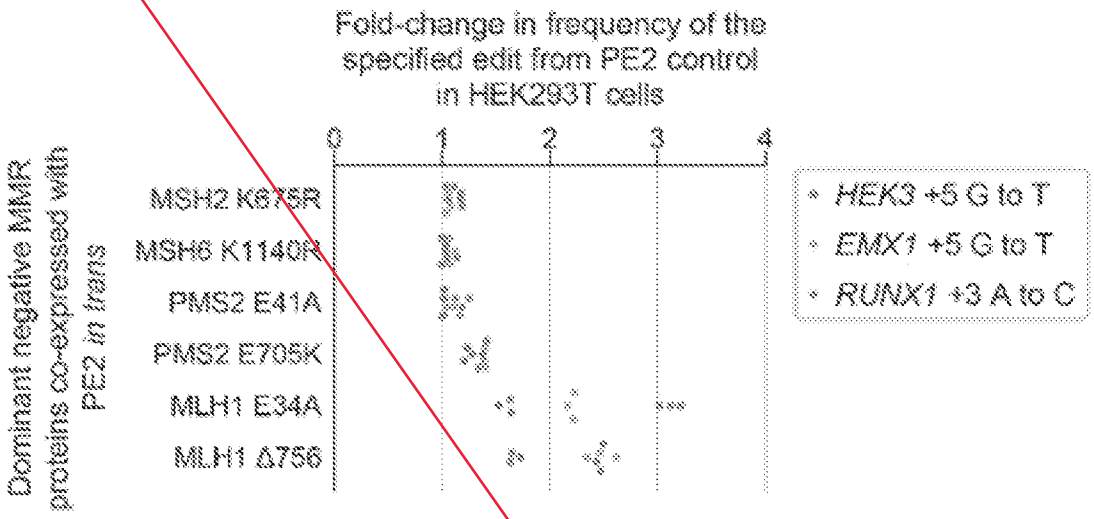


FIG. 43A

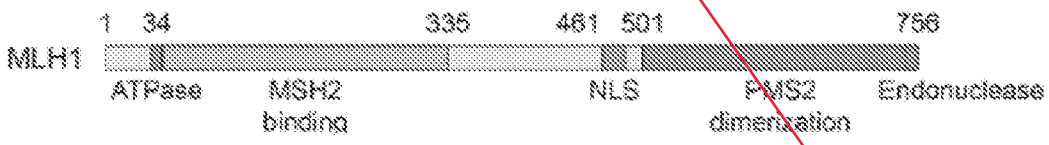


FIG. 43B

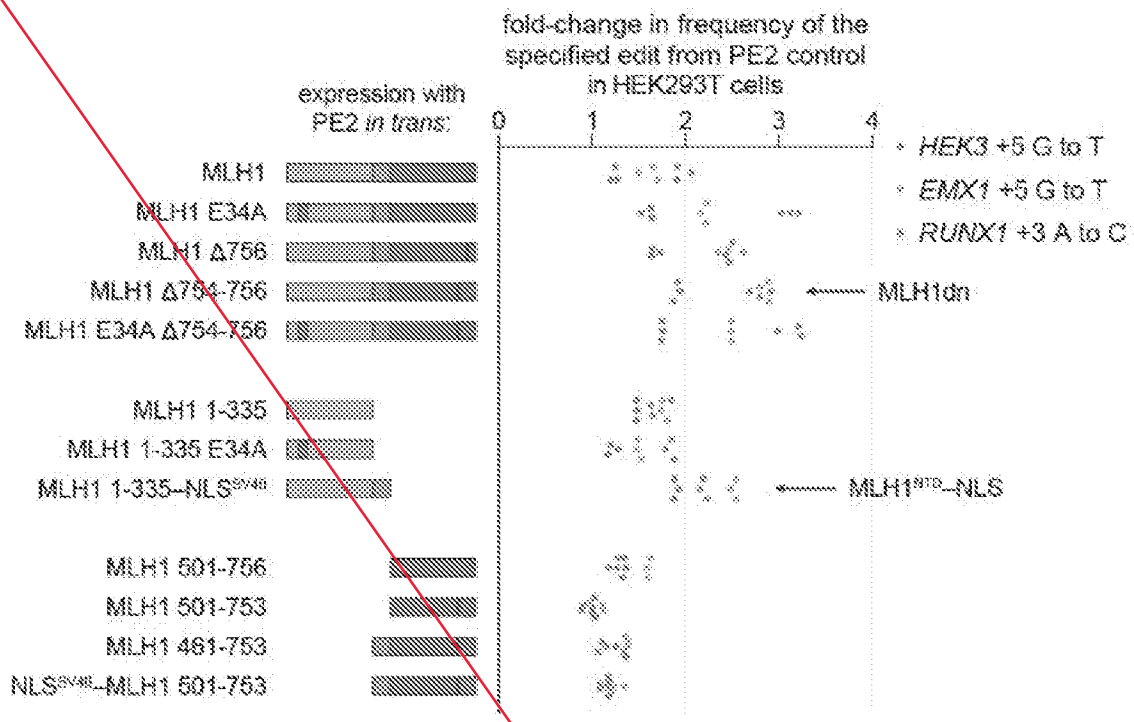


FIG. 43C

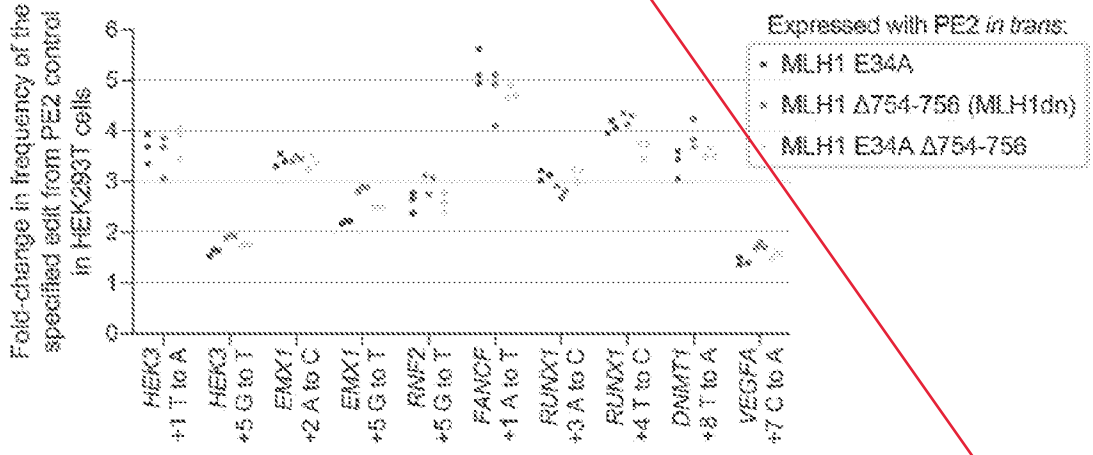


FIG. 43D

HEK3 +5 G to T
EMX1 +5 G to T
RUNX1 +3 A to C
Indels
PE2
PE2 +MLH1dn
PE2 +MLH1mut–NLS
PE2–P2A–MLH1dn
PE2
PE2 +MLH1dn
PE2 +MLH1mut–NLS
PE2–P2A–MLH1dn
PE2
PE2 +MLH1dn
PE2 +MLH1mut–NLS
PE2–P2A–MLH1dn
0 10 20 30 40
Without additional nick
0 10 20 30 40 50 60
With complementary-strand nick
% of total sequencing reads with the specified modification in HEK293T cells

FIG. 43E

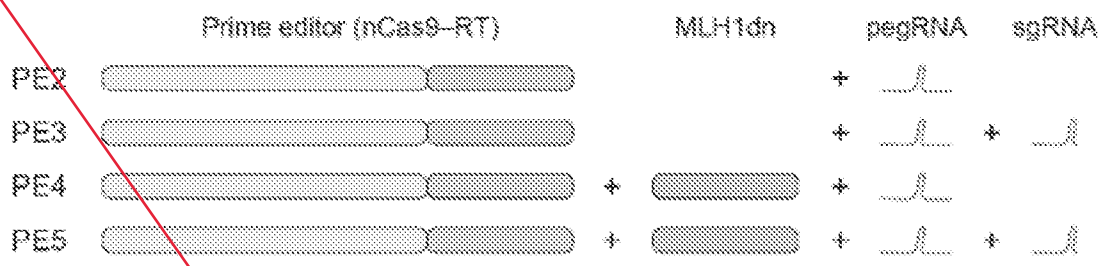


FIG. 43F

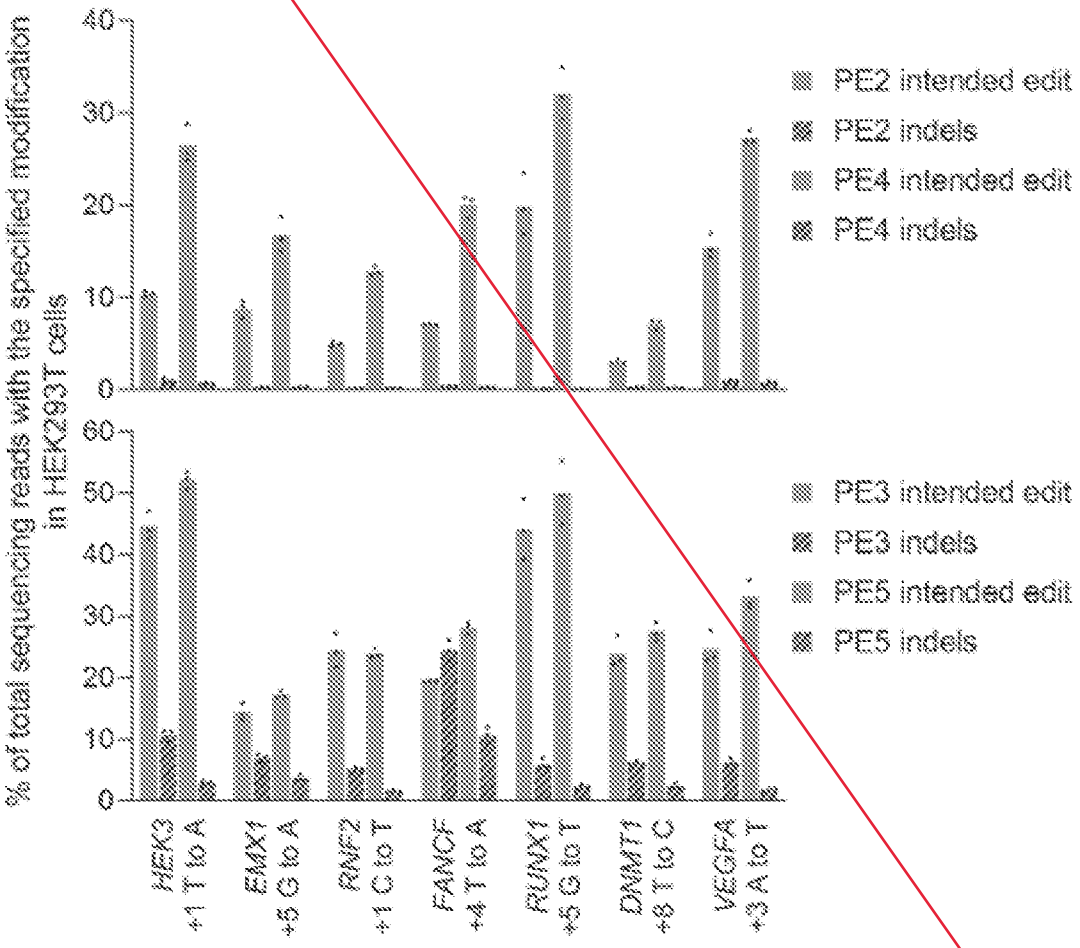


FIG. 43G

Substitution prime edit:
Mismatch created (edited strand-unedited strand):
C to G
G-G
T to A
A-A
C to T
T-G
A to T
T-T
T to C
C-A
A to C
C-T
G to T
T-C
G to A
A-C
C to A
A-G
T to G
G-A
A to G
G-T
G to C
C-C
PE4/PE2 fold change in frequency of the intended edit in HEK293T cells
0
1
2
3
4
PE5/PE3 fold change in intended edit:indel ratio in HEK293T cells
0
2
4
6
8
10
12
Seed edits (+1-3)
PAM edits (+5-6)
Other edits (+4,7-10)

FIG. 44A

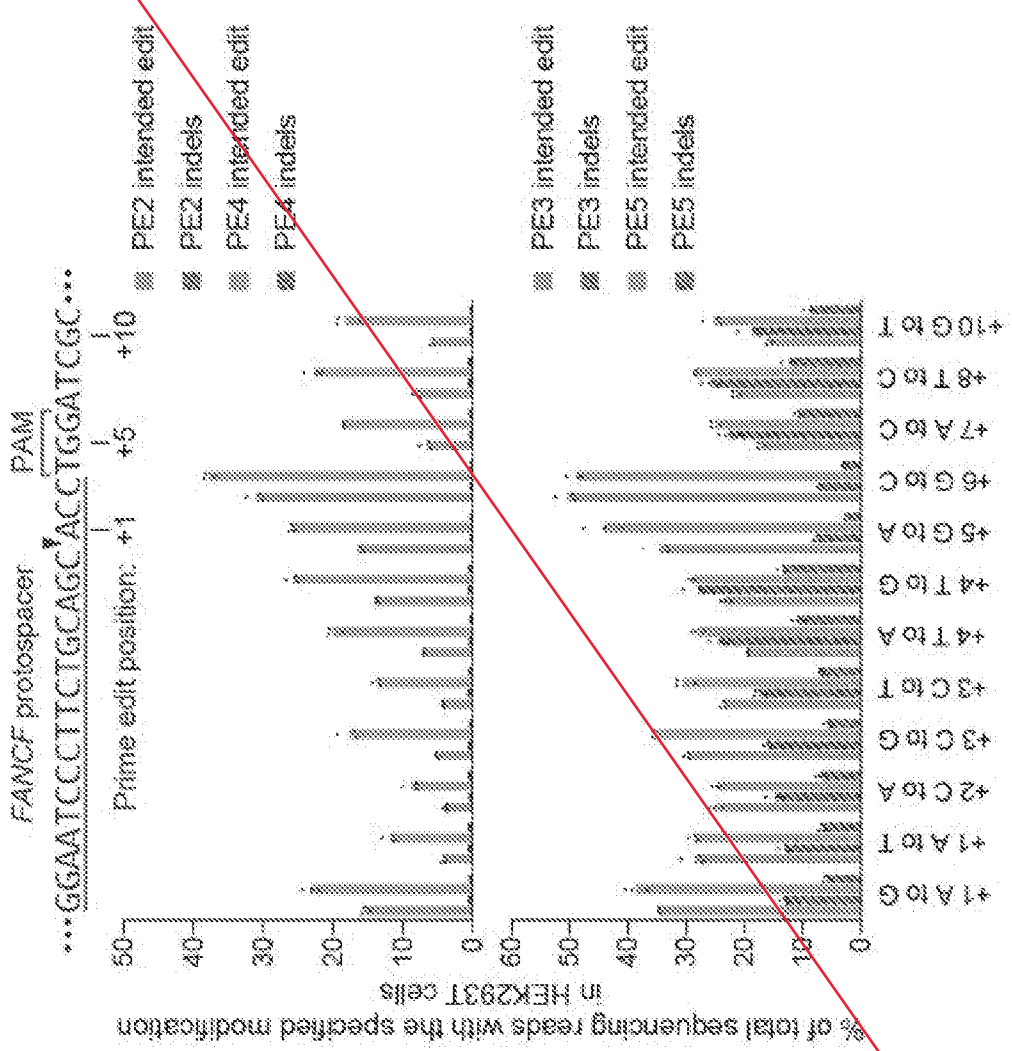

FANCF protospacer
PAM
...GGAATCCCCTTCTGCAGCACCTGGATCGC...
Prime edit position: +1 +5 +10
PE2 intended edit
PE2 indels
PE4 intended edit
PE4 indels
PE3 intended edit
PE3 indels
PE5 intended edit
PE5 indels
% of total sequencing reads with the specified modification in HEK293T cells
+1 A to G
+1 A to T
+2 C to A
+3 C to G
+3 C to T
+4 T to A
+4 T to G
+5 G to A
+6 G to C
+7 A to C
+8 T to C
+10 G to T

FIG. 44B

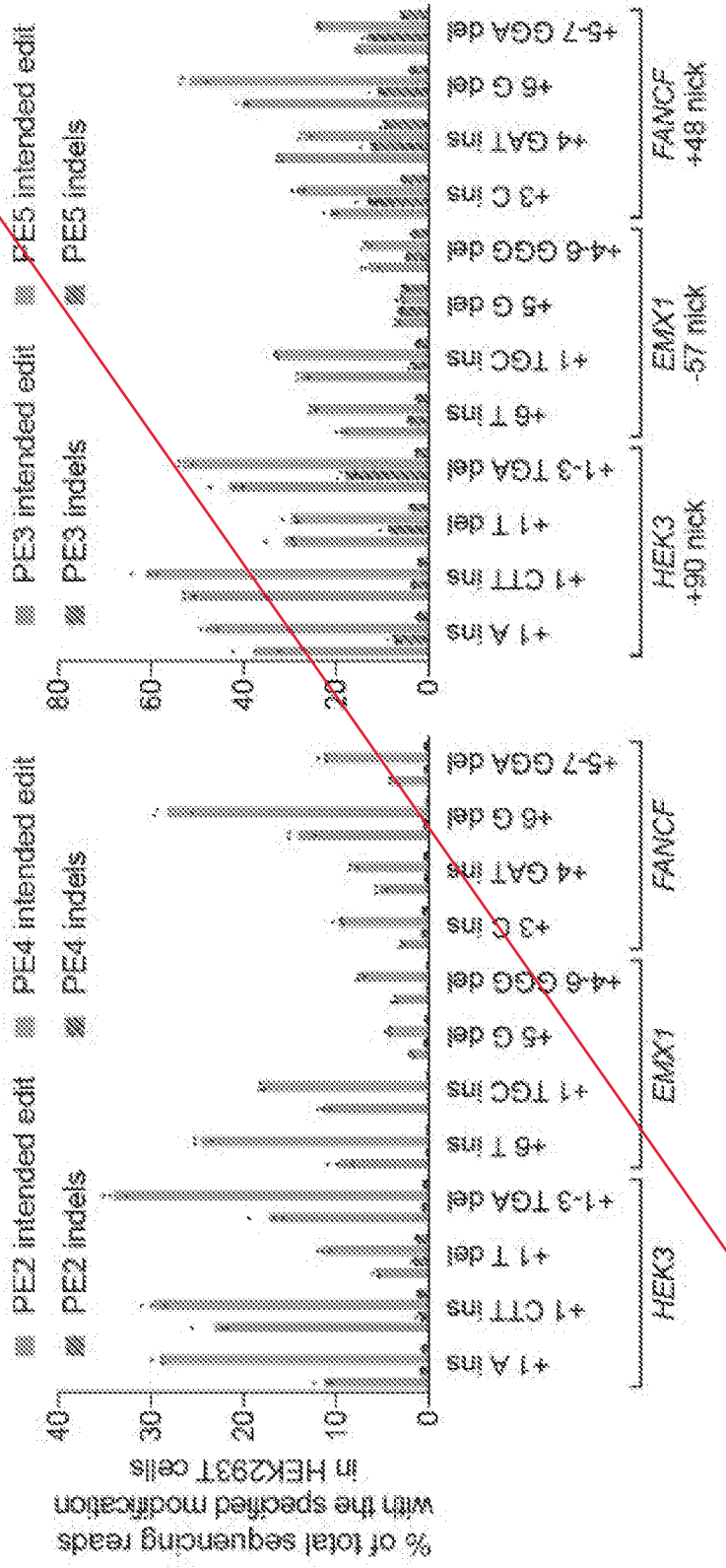

% of total sequencing reads with the specified modification in HEK293T cells
PE2 intended edit
PE2 indels
PE4 intended edit
PE4 indels
PE3 intended edit
PE3 indels
PE5 intended edit
PE5 indels
+1 A ins
+1 CTT ins
+1 T del
+1-3 TGA del
+6 T ins
+1 TGC ins
+5 G del
+4-6 GGG del
+3 C ins
+4 GAT ins
+6 G del
+5-7 GGA del
HEK3 +90 nick
EMX1 -57 nick
FANCF +48 nick

FIG. 44C

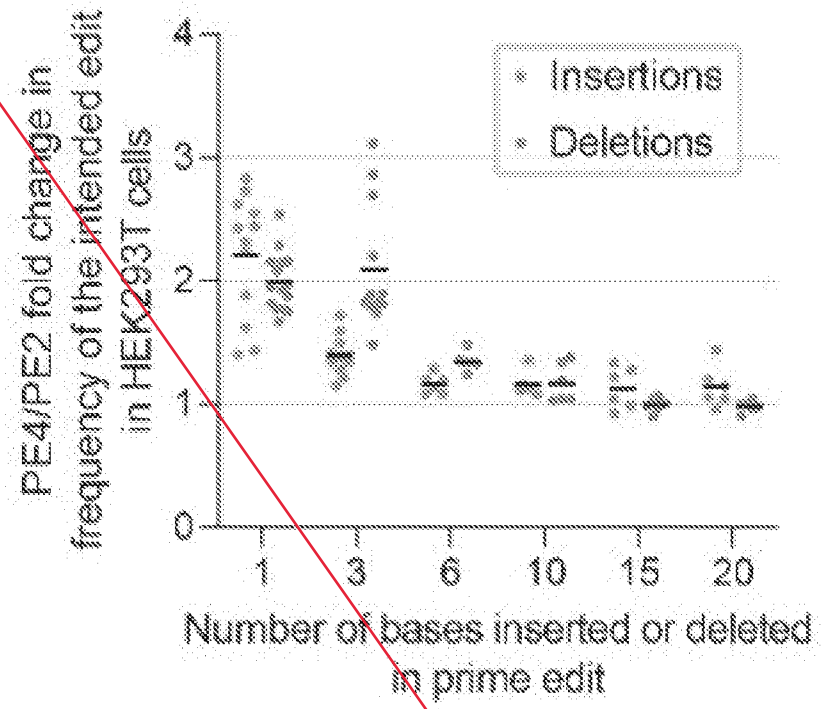

PE4/PE2 fold change in frequency of the intended edit in HEK293T cells
0
1
2
3
4
Insertions
Deletions
1
3
6
10
15
20
Number of bases inserted or deleted in prime edit

FIG. 44D

FIG. 44F

FIG. 44E

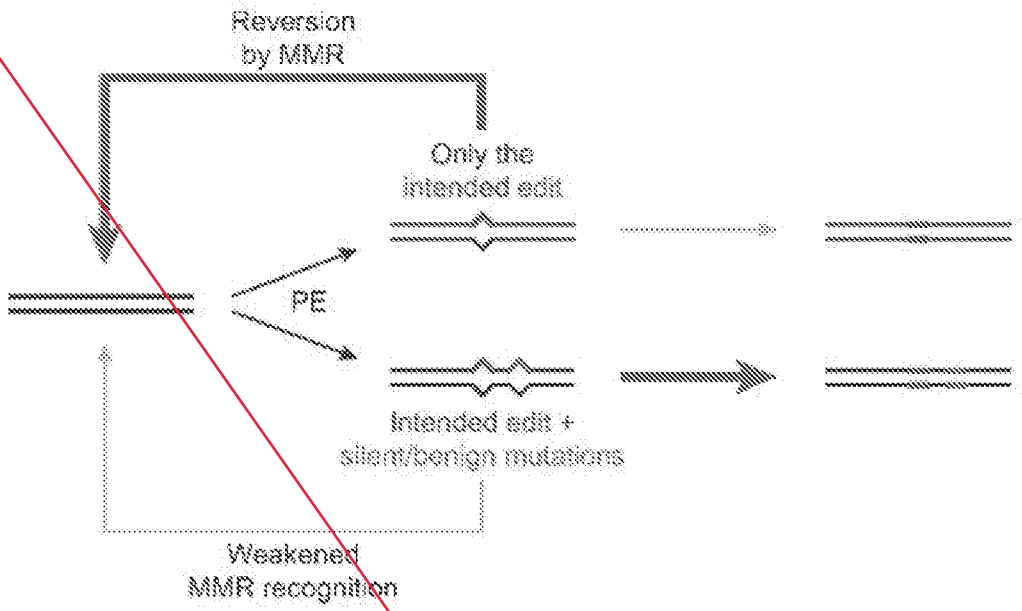


FIG. 44G

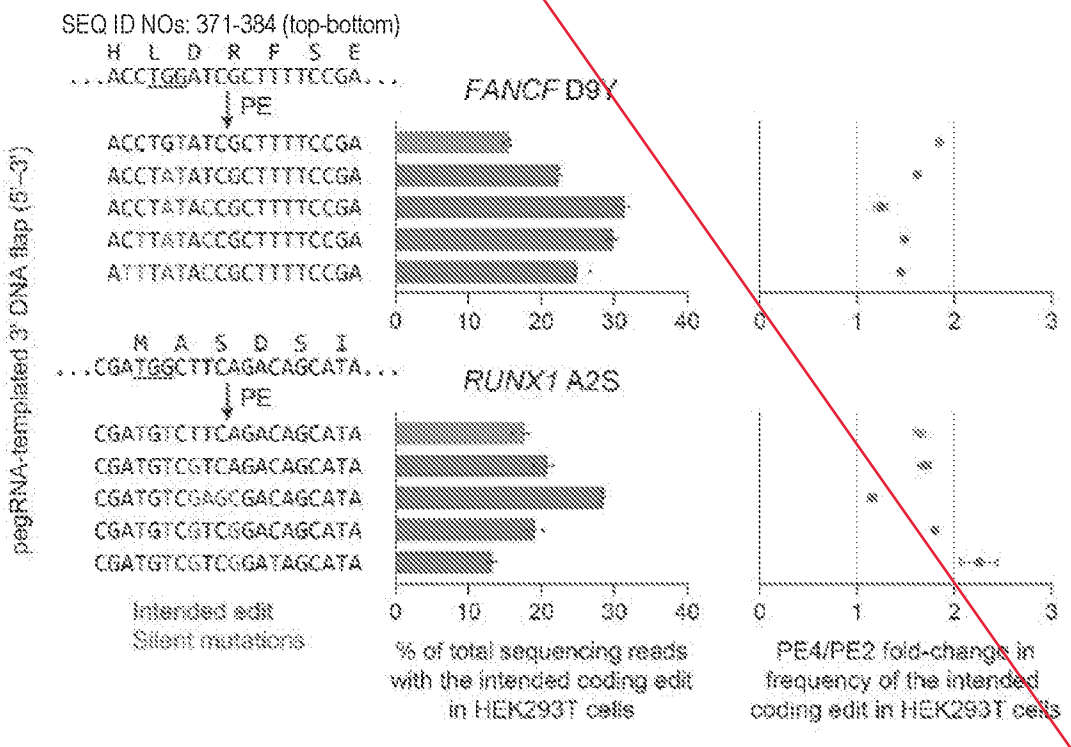


FIG. 44H

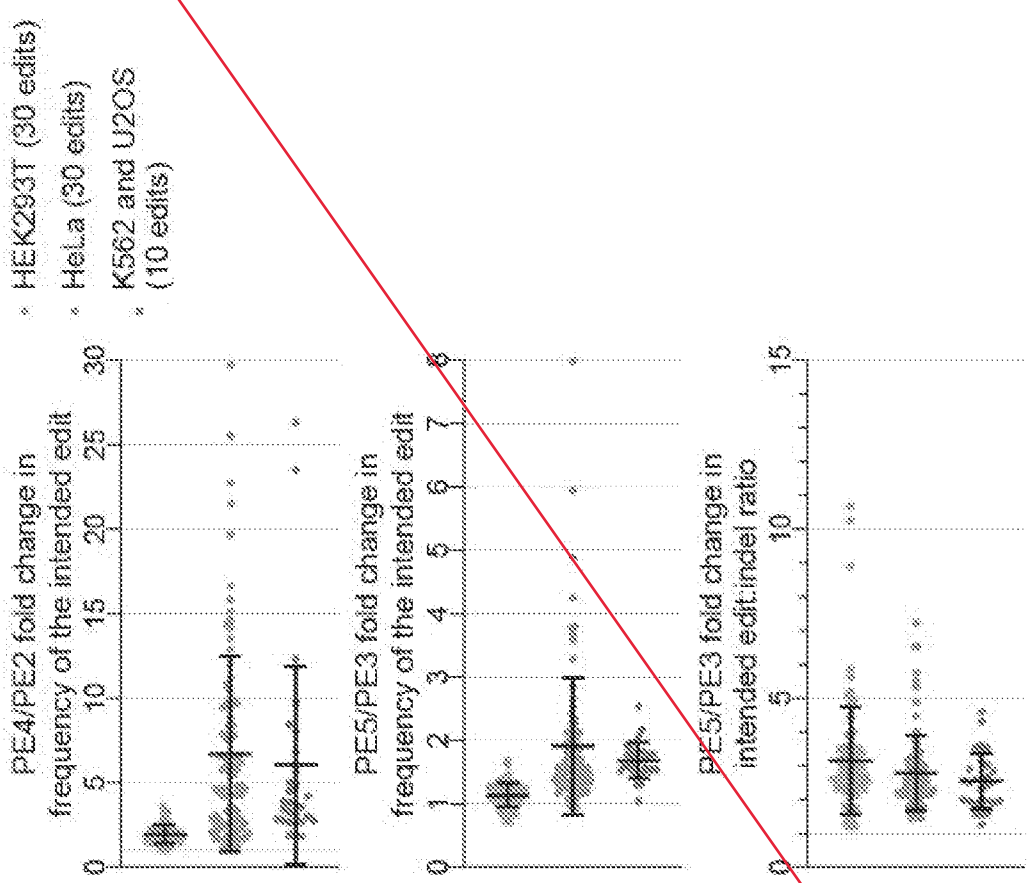

HEK293T (30 edits)
HeLa (30 edits)
K562 and U2OS (10 edits)
PE4/PE2 fold change in frequency of the intended edit
PE5/PE3 fold change in frequency of the intended edit
PE5/PE3 fold change in intended edit:indel ratio

FIG. 44I

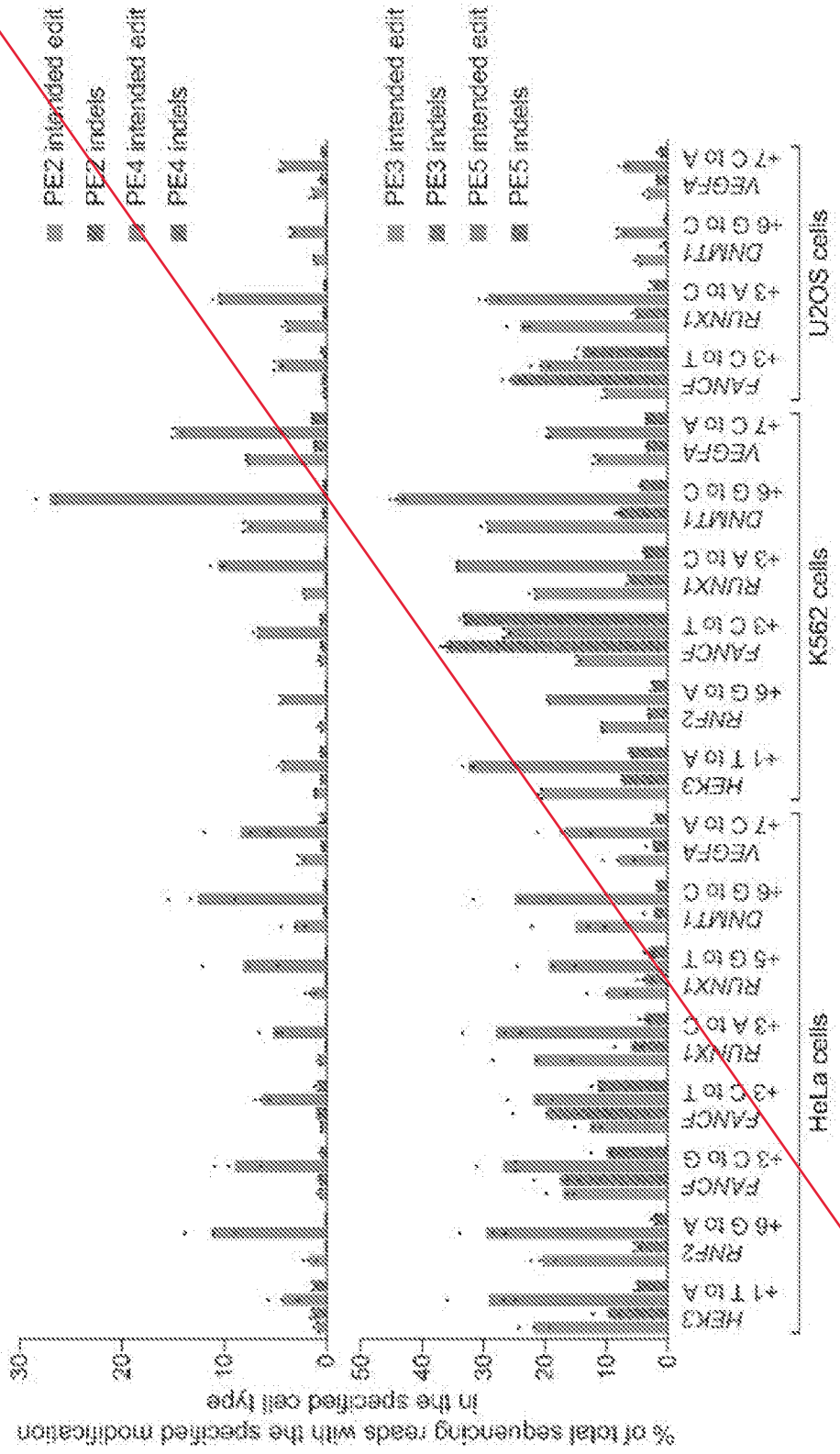

% of total sequencing reads with the specified modification in the specified cell type
PE2 intended edit
PE2 indels
PE4 intended edit
PE4 indels
PE3 intended edit
PE3 indels
PE5 intended edit
PE5 indels
HEK3 +1 T to A
RNF2 +6 G to A
FANCF +3 C to G
RUNX1 +3 A to C
FANCF +3 C to T
RUNX1 +5 G to T
DNMT1 +6 G to C
VEGFA +7 C to A
HeLa cells
HEK3 +1 T to A
RNF2 +6 G to A
FANCF +3 C to T
RUNX1 +3 A to C
DNMT1 +6 G to C
VEGFA +7 C to A
K562 cells
FANCF +3 C to T
RUNX1 +3 A to C
DNMT1 +6 G to C
VEGFA +7 C to A
U2OS cells

FIG. 44J

FIG. 45A

FIG. 45B

FIG. 45C

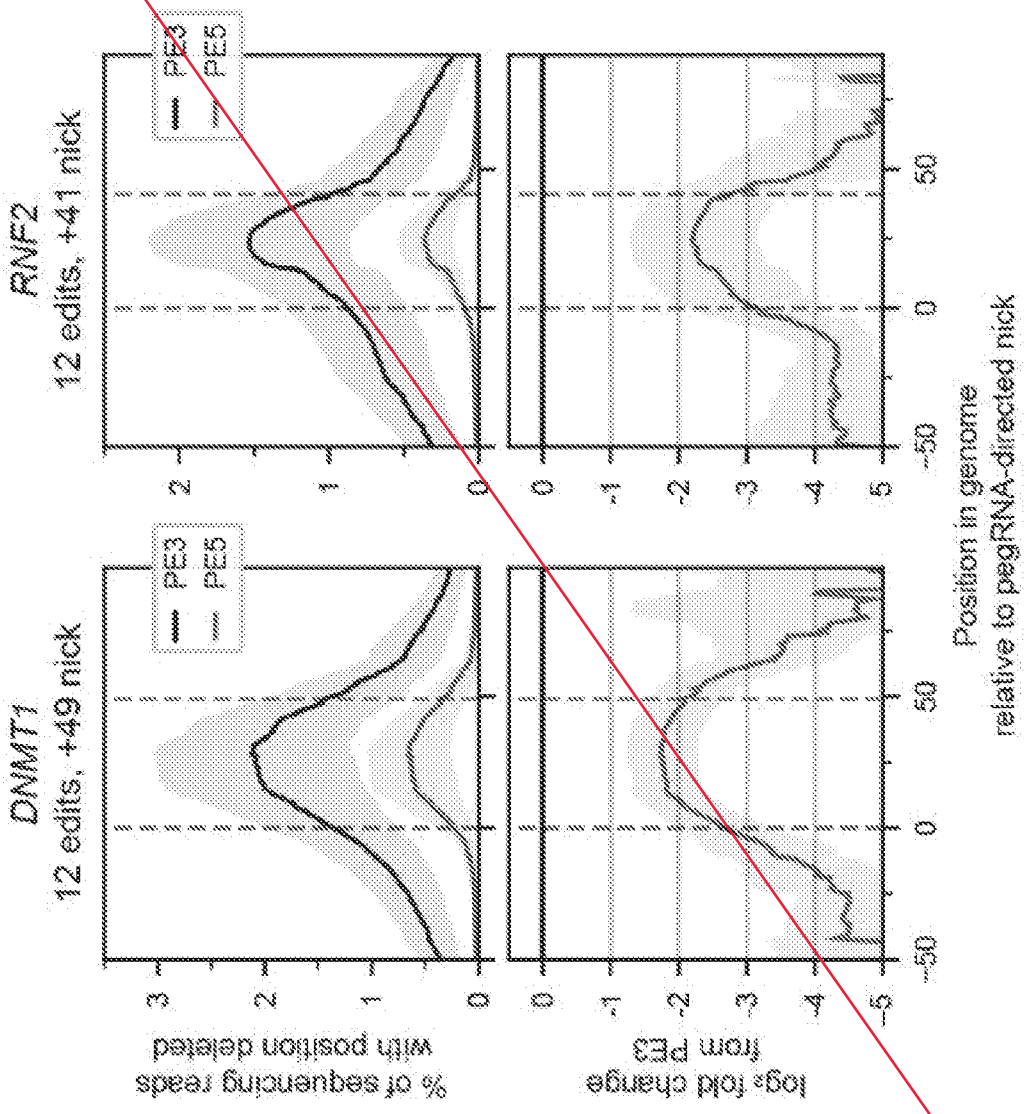

DNMT1
12 edits, +49 nick
RNF2
12 edits, +41 nick
PE3
PE5
% of sequencing reads with position deleted
log2 fold change from PE3
Position in genome relative to pegRNA-directed nick

FIG. 45D

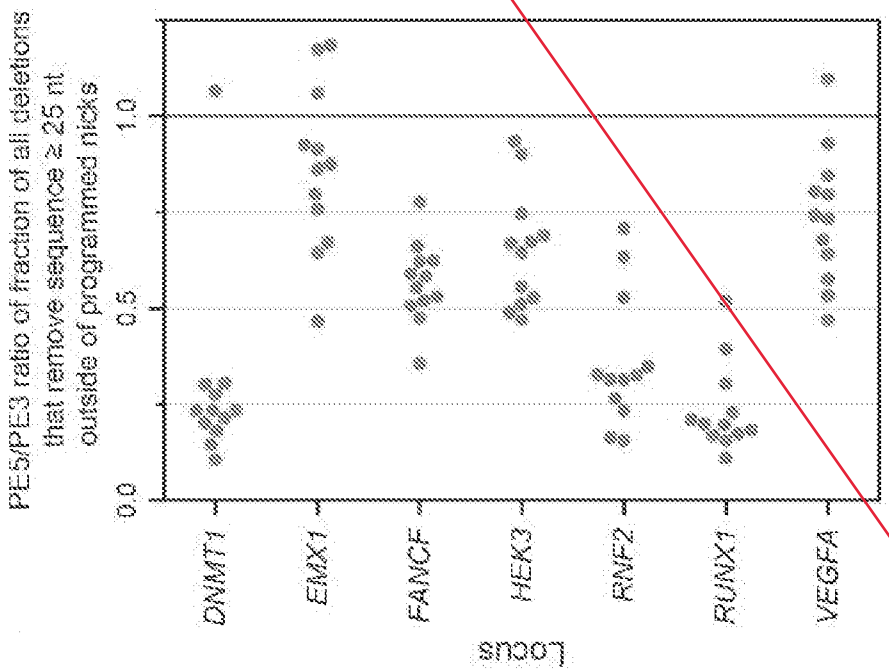


FIG. 45E

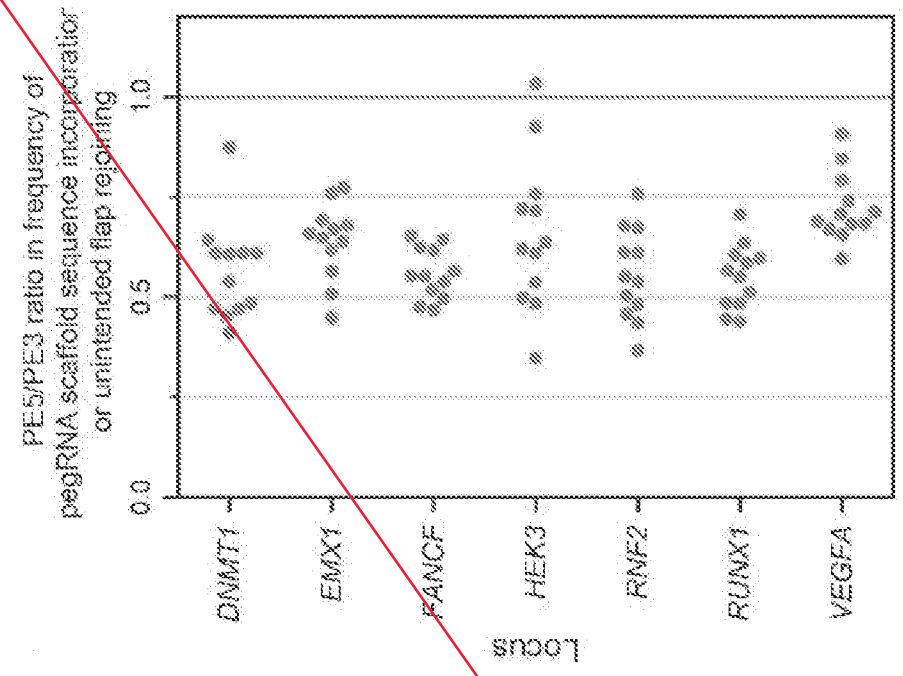


FIG. 45F

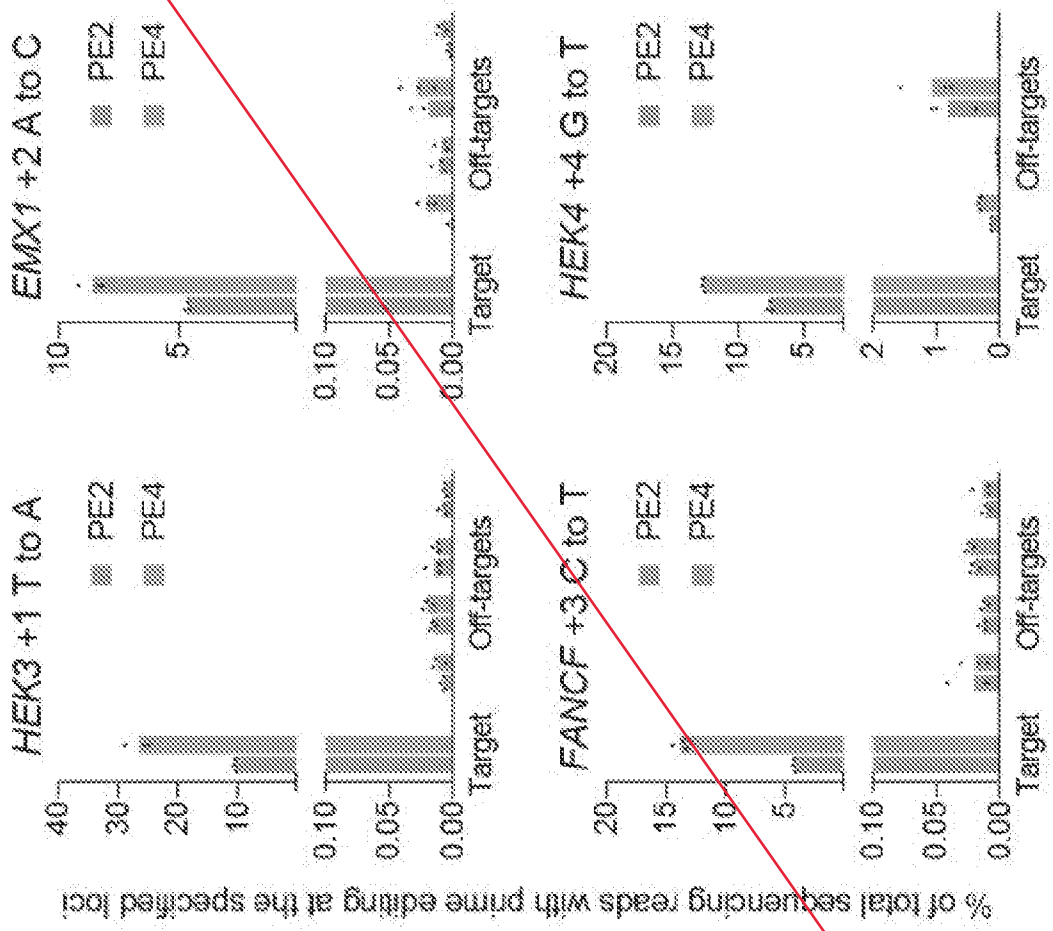

HEK3 +1 T to A
EMX1 +2 A to C
FANCF +3 C to T
HEK4 +4 G to T
PE2
PE4
Target
Off-targets
% of total sequencing reads with prime editing at the specified loci

FIG. 45G

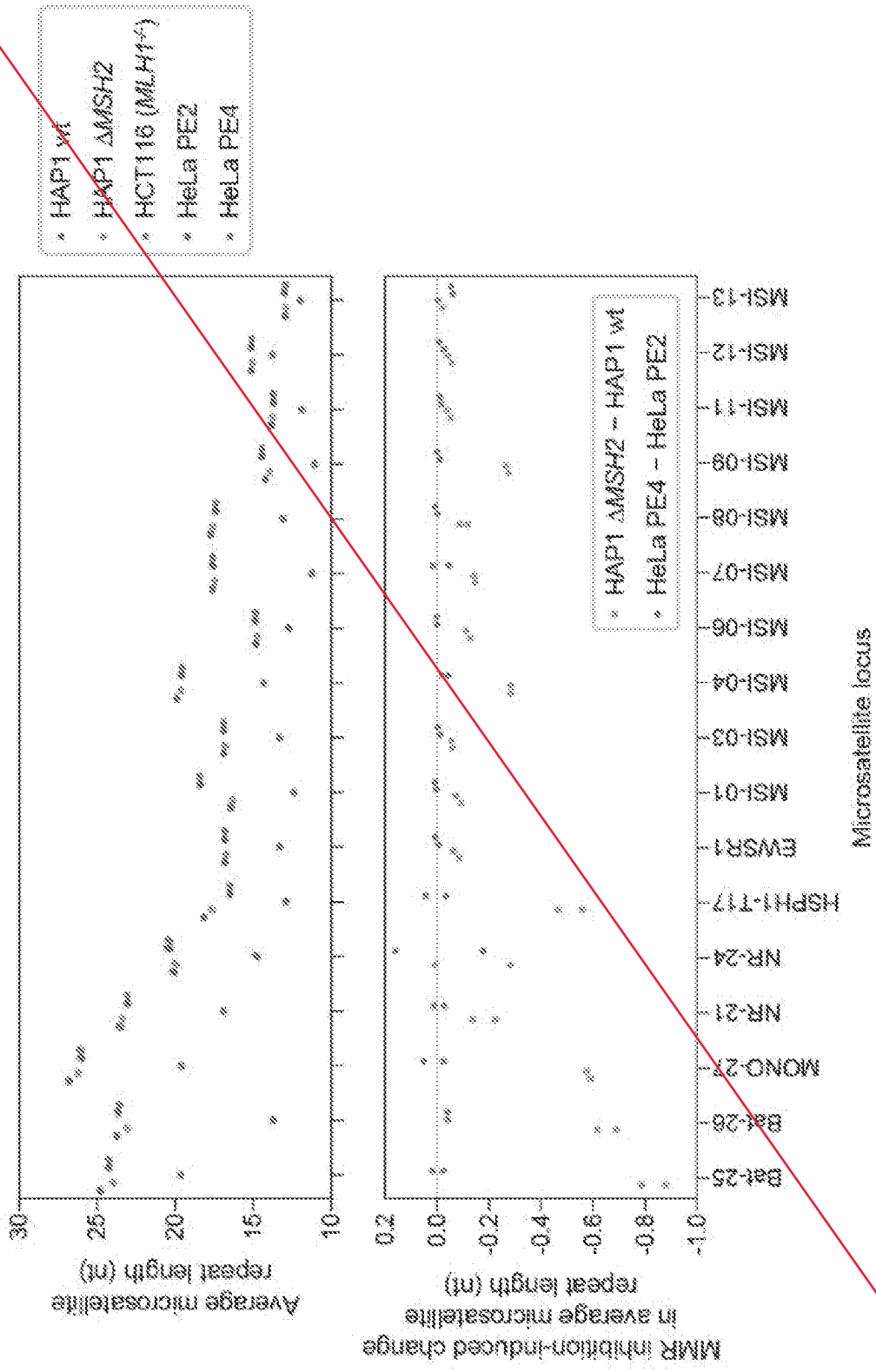

HAP1 wt
HAP1 ΔMSH2
HCT116 (MLH1-/-)
HeLa PE2
HeLa PE4
Average microsatellite repeat length (nt)
30
25
20
15
10
MMR inhibition-induced change in average microsatellite repeat length (nt)
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
HAP1 ΔMSH2 – HAP1 wt
HeLa PE4 – HeLa PE2
Bat-25
Bat-26
MONO-27
NR-21
NR-24
HSPH1-T17
EWSR1
MSI-01
MSI-03
MSI-04
MSI-06
MSI-07
MSI-08
MSI-08
MSI-09
MSI-11
MSI-12
MSI-13
Microsatellite locus

FIG. 45H

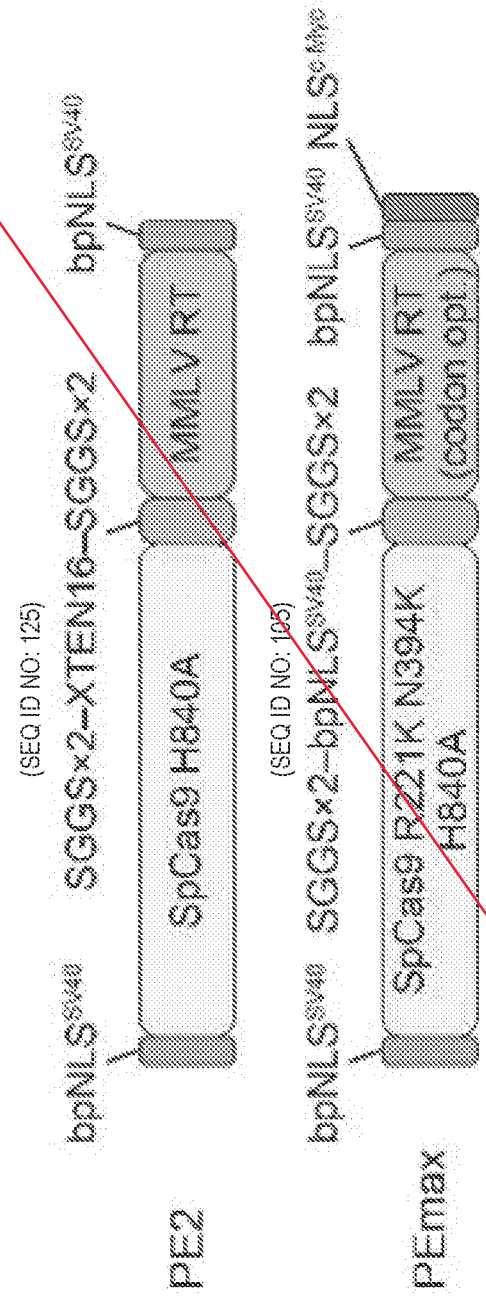

PE2
bpNLS^SV40
SGGSx2–XTEN16–SGGSx2
(SEQ ID NO: 125)
SpCas9 H840A
MMLV RT
bpNLS^SV40
PEmax
bpNLS^SV40
SGGSx2–bpNLS^SV40–SGGSx2
(SEQ ID NO: 126)
SpCas9 R221K N394K H840A
MMLV RT (codon opt.)
bpNLS^SV40
NLS^Myo

FIG. 46A

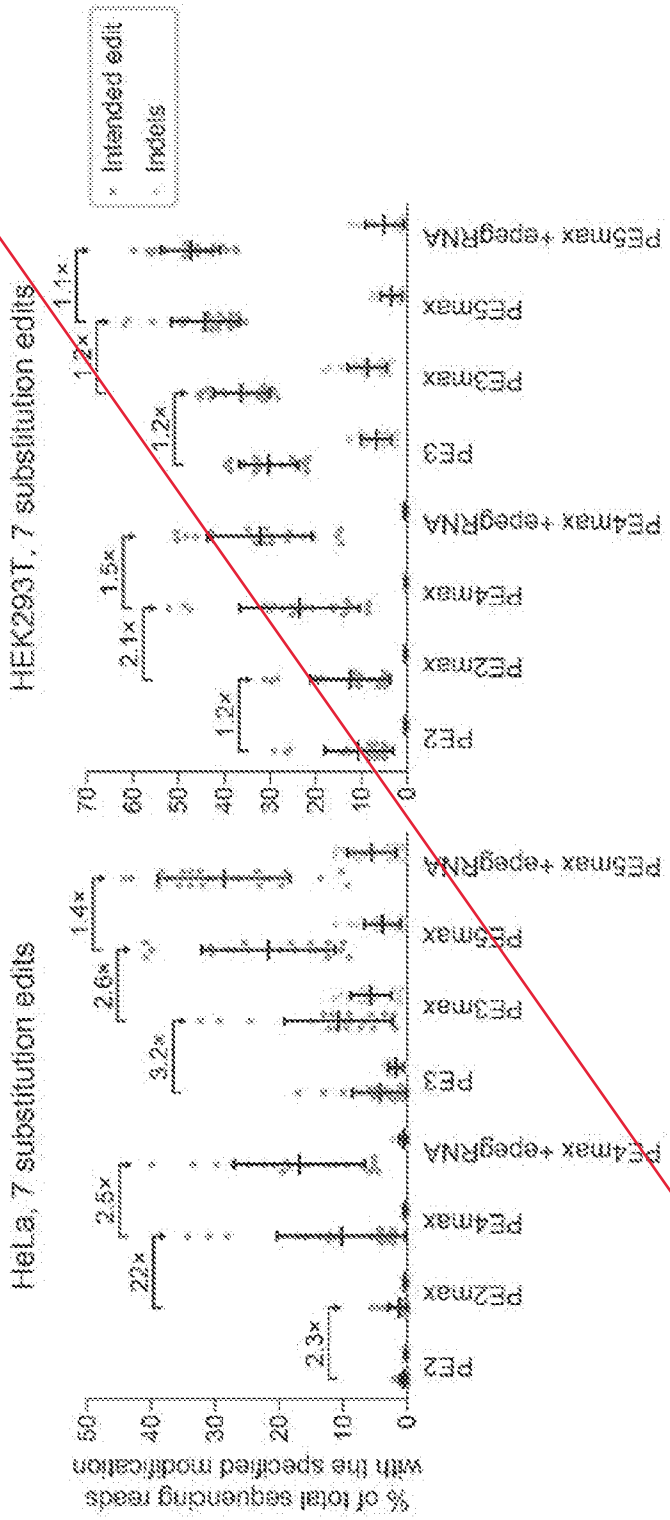

HeLa, 7 substitution edits
HEK293T, 7 substitution edits
% of total sequencing reads with the specified modification
Intended edit
Indels
PE2
PE2max
PE4max
PE4max +epegRNA
PE3
PE3max
PE5max
PE5max +epegRNA
2.3×
22×
2.5×
3.2×
2.6×
1.4×
1.2×
2.1×
1.5×
1.2×
1.2×
1.1×

FIG. 46B

pegRNA
5'
3'
epegRNA
5'
3'
spacer
scaffold
3' extension
structural motif

FIG. 46C

% of total sequencing reads with the specified modification in HEK293T cells
% of total sequencing reads with the specified modification in HeLa cells
HBB +5 G to A (E6V site)
PPMP G127V install (+6 G to T)
CCR5 +1 C to T (c.1412delA site)
CXCR4 P191A install (+5 C to G)
IL2RB H134D Y135F install (+1 T to A, +5 C to G)
HBG1/2 BCL11A to GATA1 motif (+6 C to A, +8 C to A)
PE2
PE2max
PE4max
PE3
PE3max
PE5max

FIG. 46D

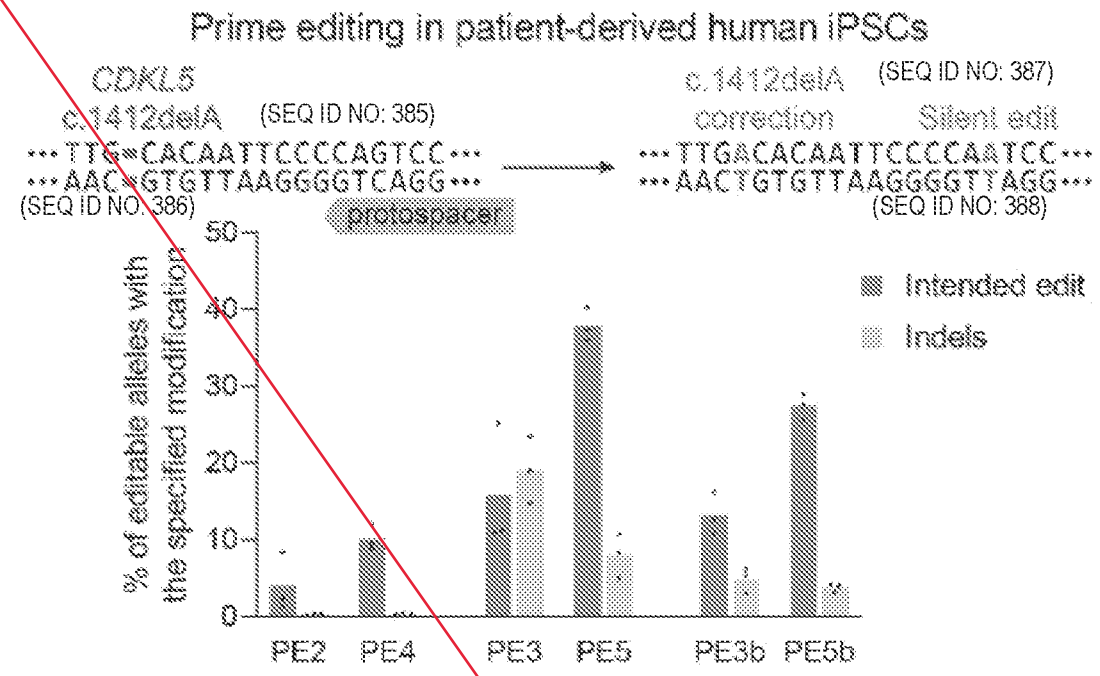


FIG. 46E

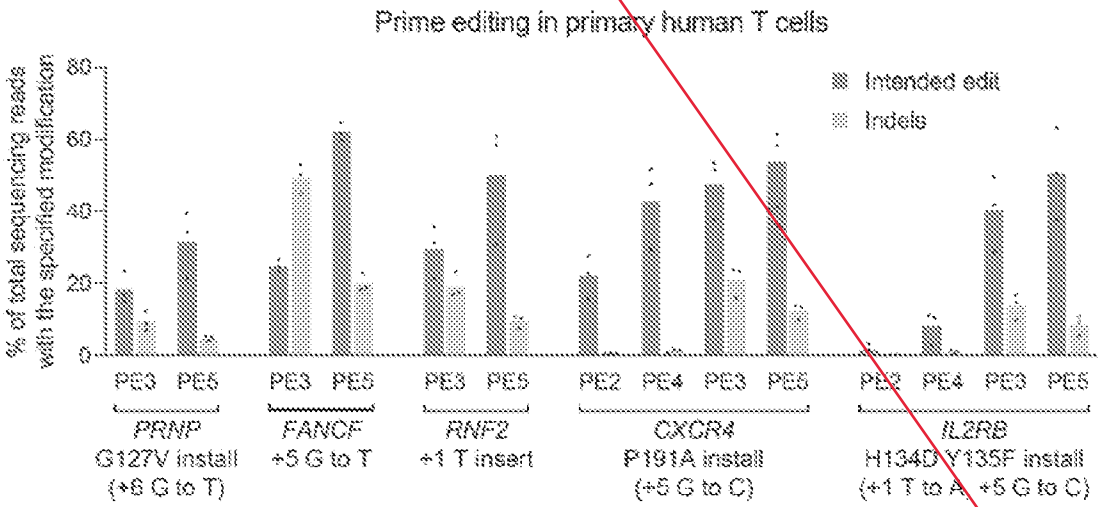


FIG. 46F

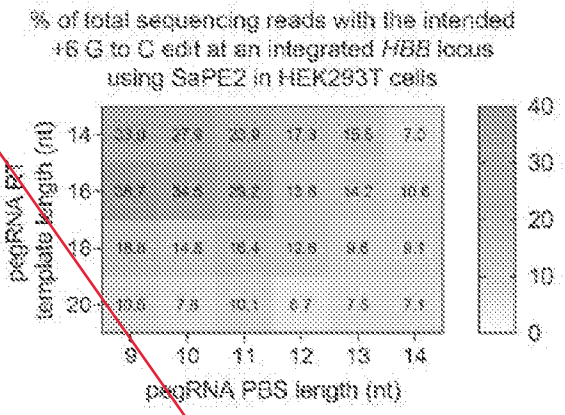


FIG. 47A

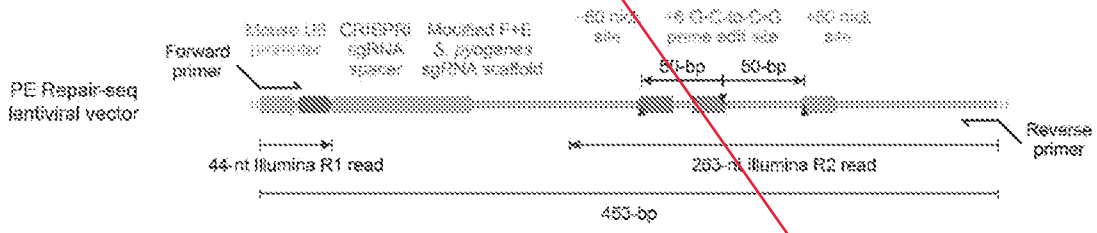


FIG. 47B

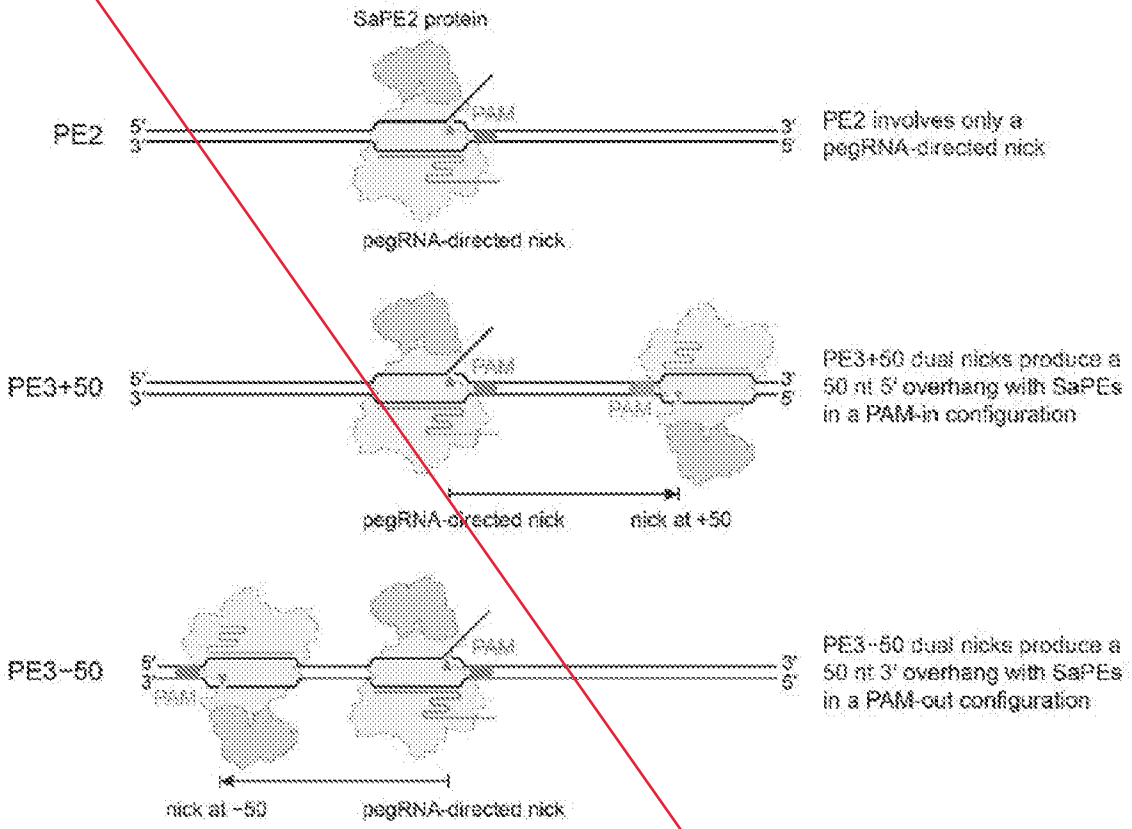

SaPE2 protein
PE2
5'
3'
PAM
3'
5'
PE2 involves only a pegRNA-directed nick
pegRNA-directed nick
PE3+50
PE3+50 dual nicks produce a 50 nt 5' overhang with SaPEs in a PAM-in configuration
nick at +50
PE3~50
PE3~50 dual nicks produce a 50 nt 3' overhang with SaPEs in a PAM-out configuration
nick at ~50

FIG. 47C

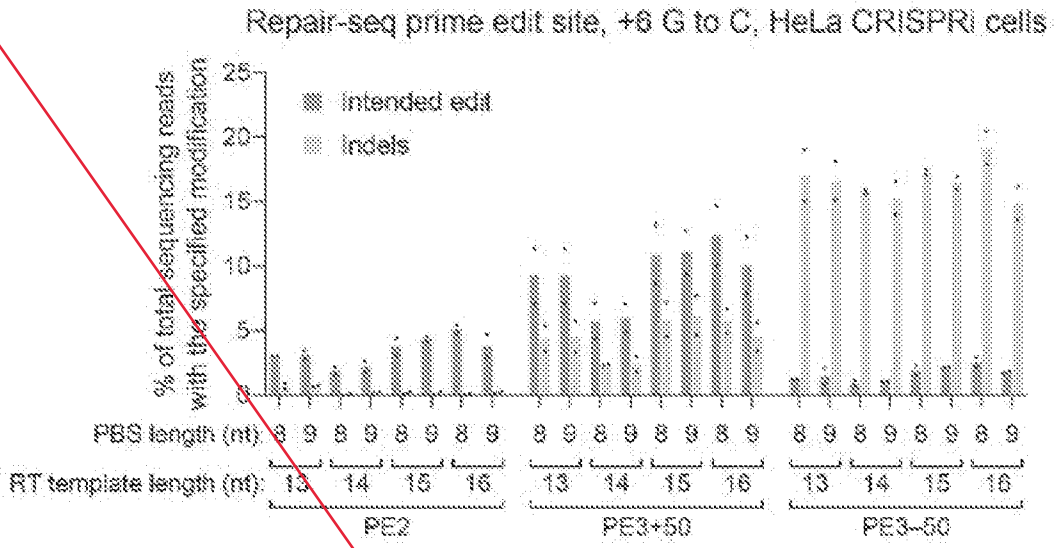


FIG. 47D

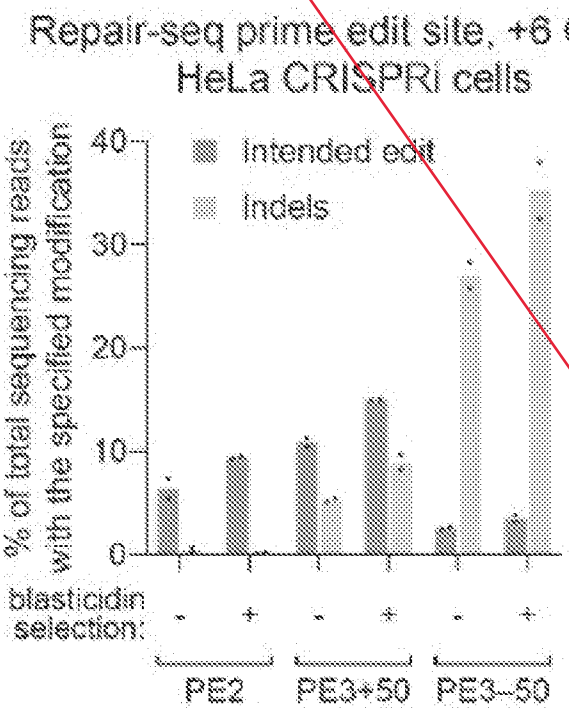


FIG. 47E

Functional annotation class	Number of genes in CRISPRi sgRNA library
Recombinational repair	41
Double-strand break repair	36
Nucleotide-excision repair	30
Base-excision repair	23
Postreplication repair	16
Mismatch repair	16
DNA repair (other)	50
Cellular response to stimulus	4
Nuclear division	12
Cell cycle checkpoint	12
mRNA processing	21
DNA replication	23
Protein modification	24
Other	39
Regulation of transcription	46
Chromatin processes	83
Total	476

FIG. 47F

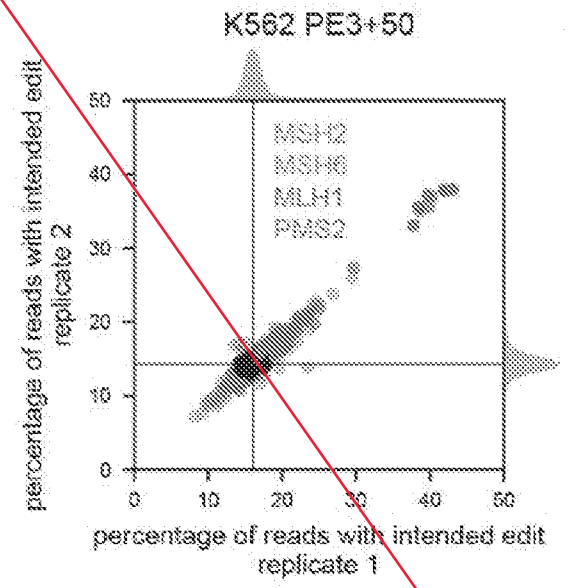


FIG. 47G

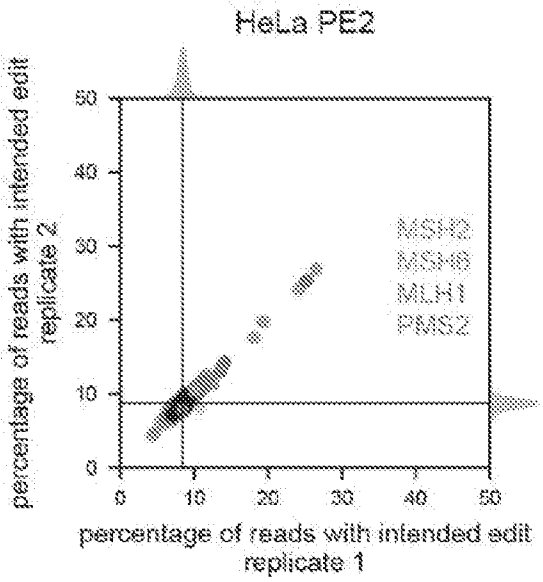


FIG. 47H

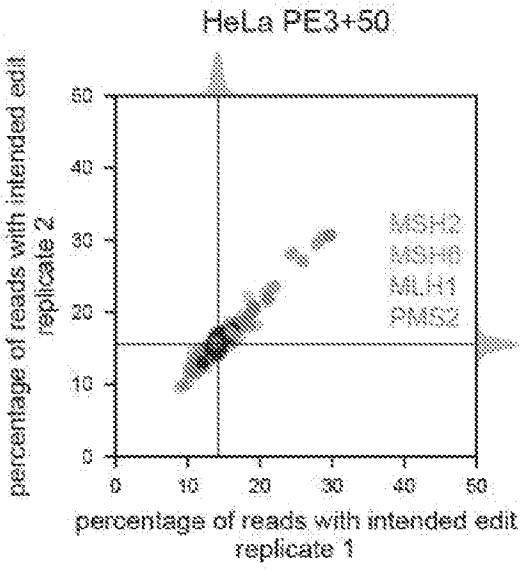


FIG. 47I

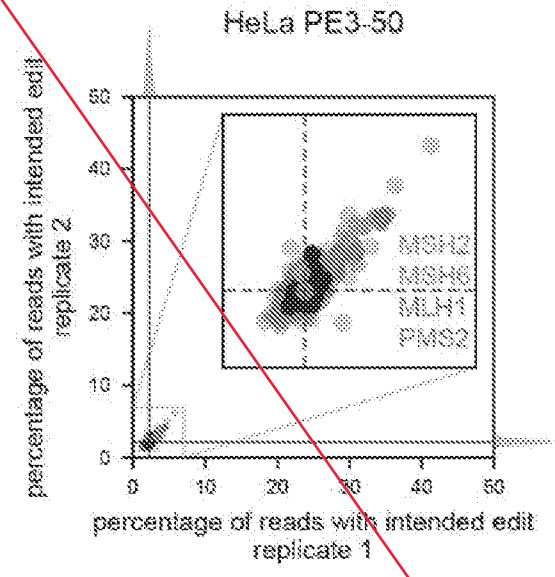


FIG. 47J

FIG. 48A

FIG. 48B

K562 PE3+50
intended edit
joining of reverse-transcribed sequence at unintended location
deletion
installation of additional edits from nearby-matched scaffold
tandem duplication
all others
percentage of reads (linear scale)
percentage of reads (log scale)
log2 fold-change from non-targeting
LIG1 protein
DNA2 protein
POLD3 protein
non-targeting guides
MLH1
MSH2
MSH6
EXO1
DNA2
POLD1
POLD2
POLD3

FIG. 48C

FIG. 48D FIG. 48E FIG. 48F FIG. 48G FIG. 48H FIG. 48I

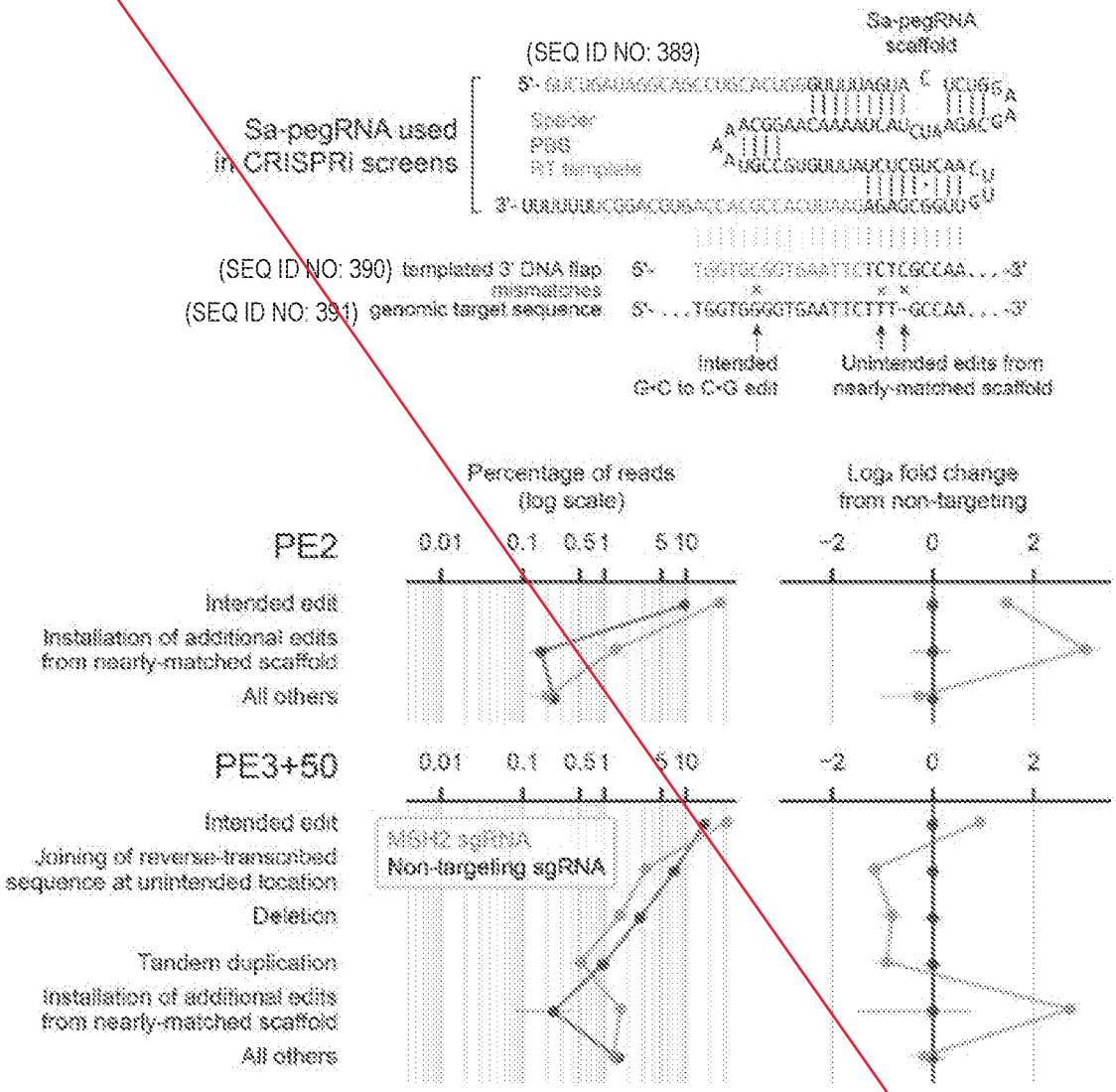

Sa-pegRNA used in CRISPRi screens
(SEQ ID NO: 389)
Sa-pegRNA scaffold
Spacer
PBS
RT template
(SEQ ID NO: 390) templated 3' DNA flap
(SEQ ID NO: 391) genomic target sequence
mismatches
Intended G•C to C•G edit
Unintended edits from nearly-matched scaffold
Percentage of reads (log scale)
Log2 fold change from non-targeting
PE2
PE3+50
Intended edit
Installation of additional edits from nearly-matched scaffold
All others
Joining of reverse-transcribed sequence at unintended location
Deletion
Tandem duplication
MSH2 sgRNA
Non-targeting sgRNA

FIG. 49A

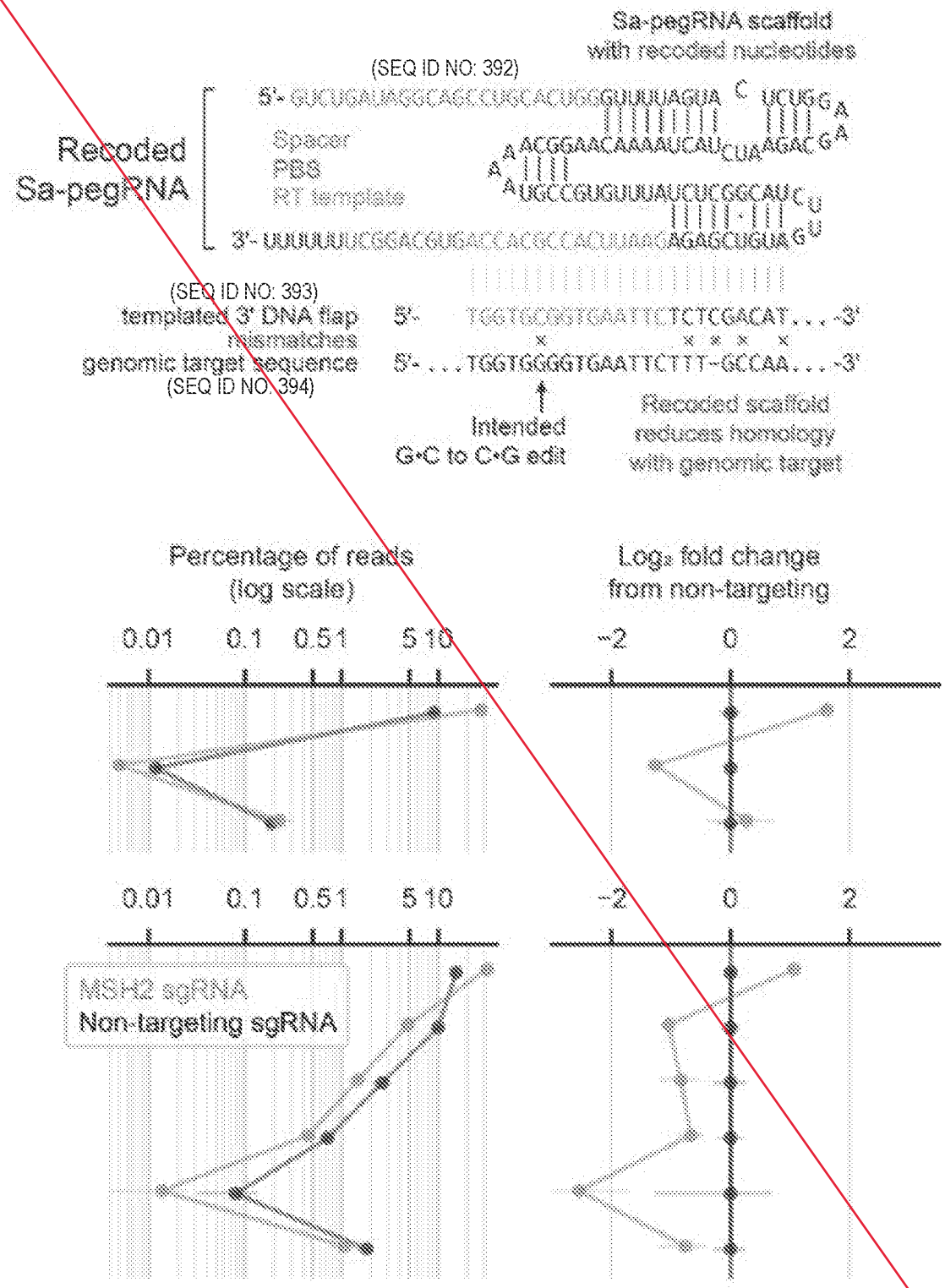

Sa-pegRNA scaffold
with recoded nucleotides
(SEQ ID NO: 392)
Recoded
Sa-pegRNA
Spacer
PBS
RT template
(SEQ ID NO: 393)
templated 3' DNA flap
mismatches
genomic target sequence
(SEQ ID NO: 394)
Intended
G•C to C•G edit
Recoded scaffold
reduces homology
with genomic target
Percentage of reads
(log scale)
Log2 fold change
from non-targeting
0.01
0.1
0.5
1
5
10
−2
0
2
MSH2 sgRNA
Non-targeting sgRNA

FIG. 49B

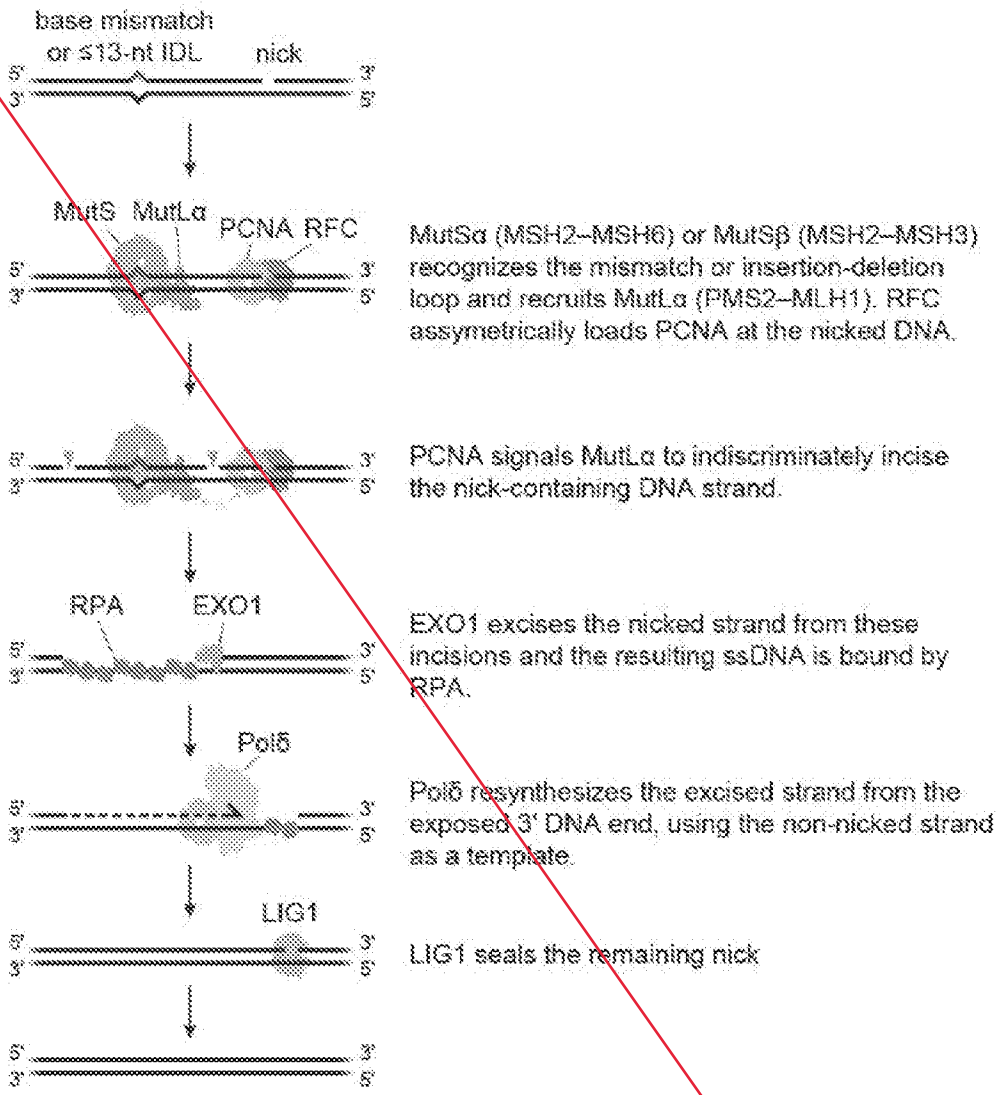

base mismatch
or ≤13-nt IDL
nick
5'
3'
3'
5'
MutS
MutLα
PCNA
RFC
MutSα (MSH2–MSH6) or MutSβ (MSH2–MSH3) recognizes the mismatch or insertion-deletion loop and recruits MutLα (PMS2–MLH1). RFC assymetrically loads PCNA at the nicked DNA.
PCNA signals MutLα to indiscriminately incise the nick-containing DNA strand.
RPA
EXO1
EXO1 excises the nicked strand from these incisions and the resulting ssDNA is bound by RPA.
Polδ
Polδ resynthesizes the excised strand from the exposed 3' DNA end, using the non-nicked strand as a template.
LIG1
LIG1 seals the remaining nick

FIG. 49C

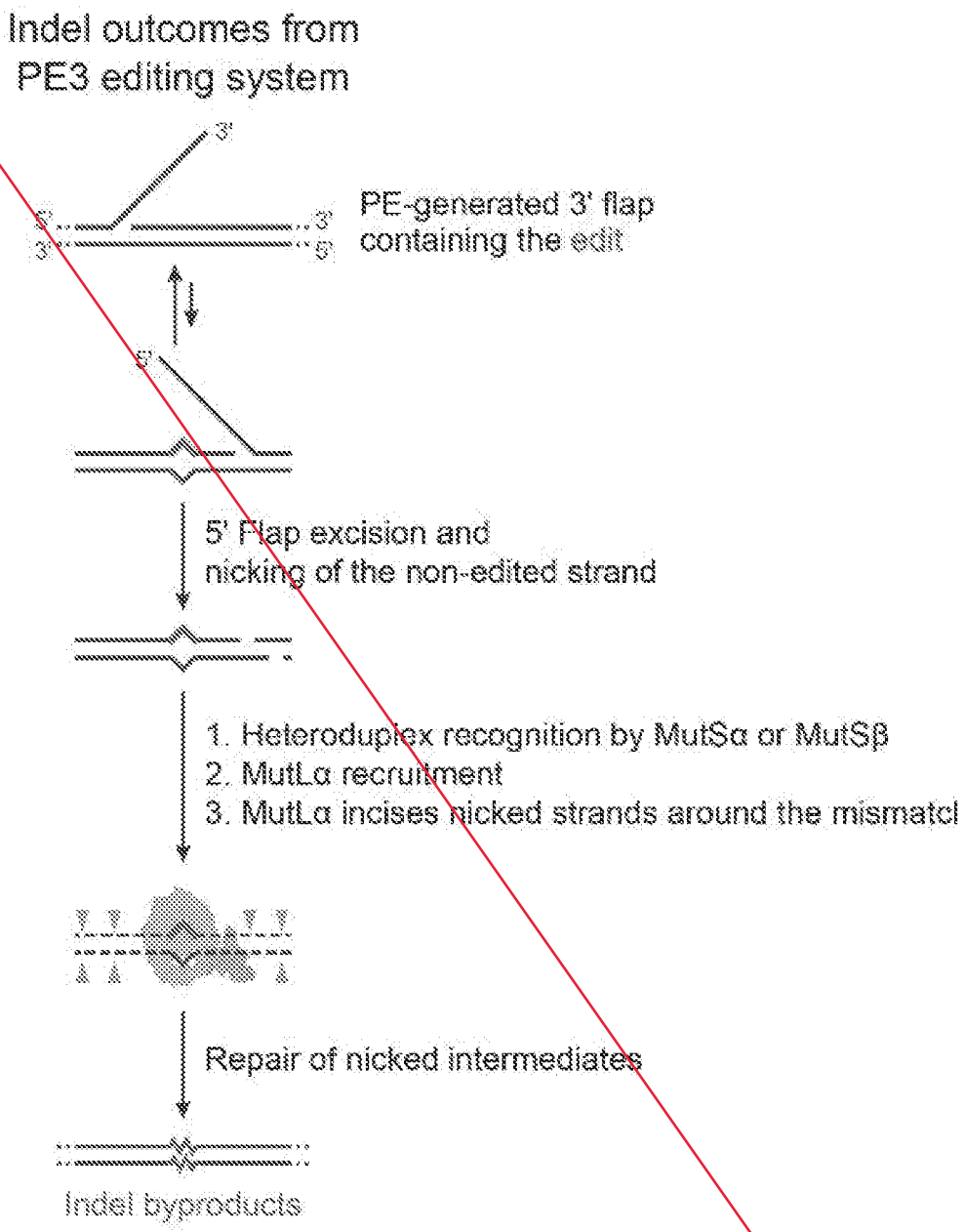

Indel outcomes from
PE3 editing system
3'
5'
3'
3'
5'
PE-generated 3' flap
containing the edit
5'
5' Flap excision and
nicking of the non-edited strand
1. Heteroduplex recognition by MutSα or MutSβ
2. MutLα recruitment
3. MutLα incises nicked strands around the mismatch
Repair of nicked intermediates
Indel byproducts

FIG. 49D

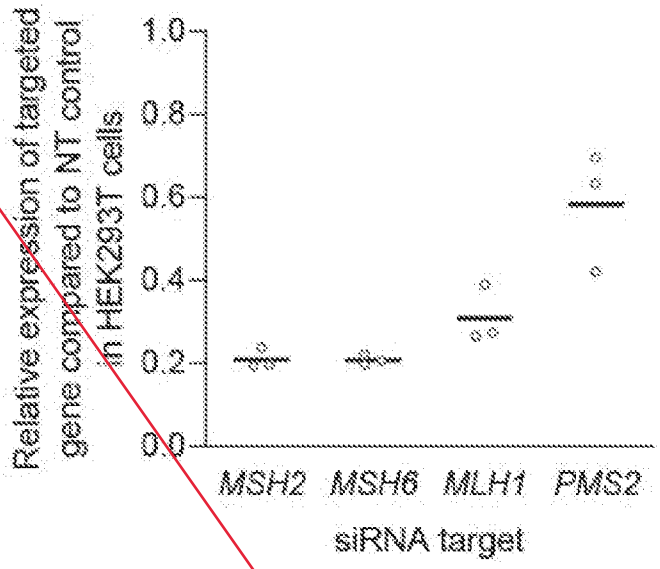


FIG. 49E

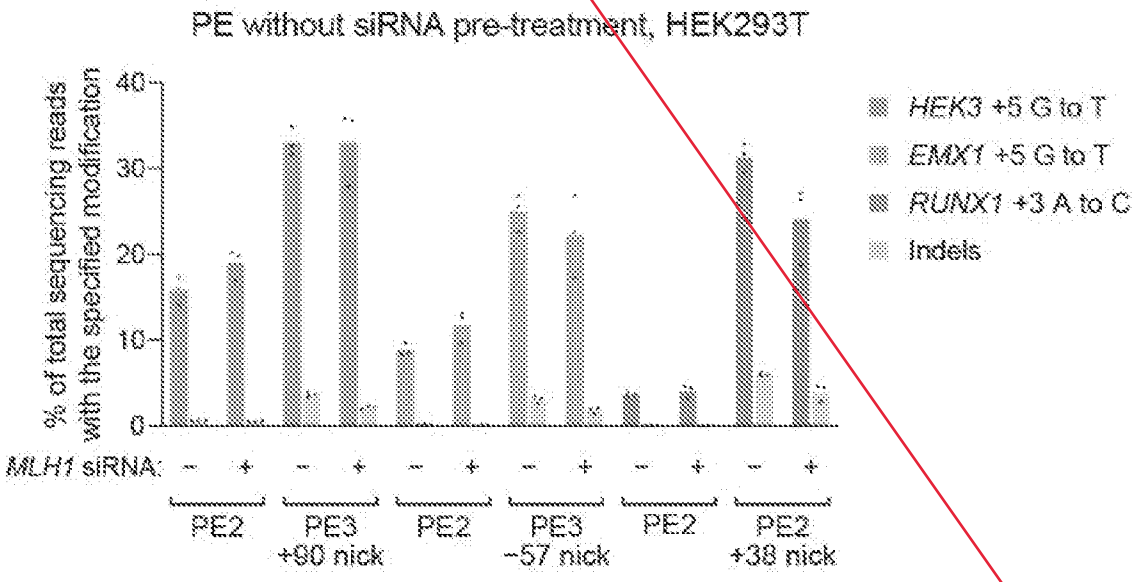


FIG. 49F

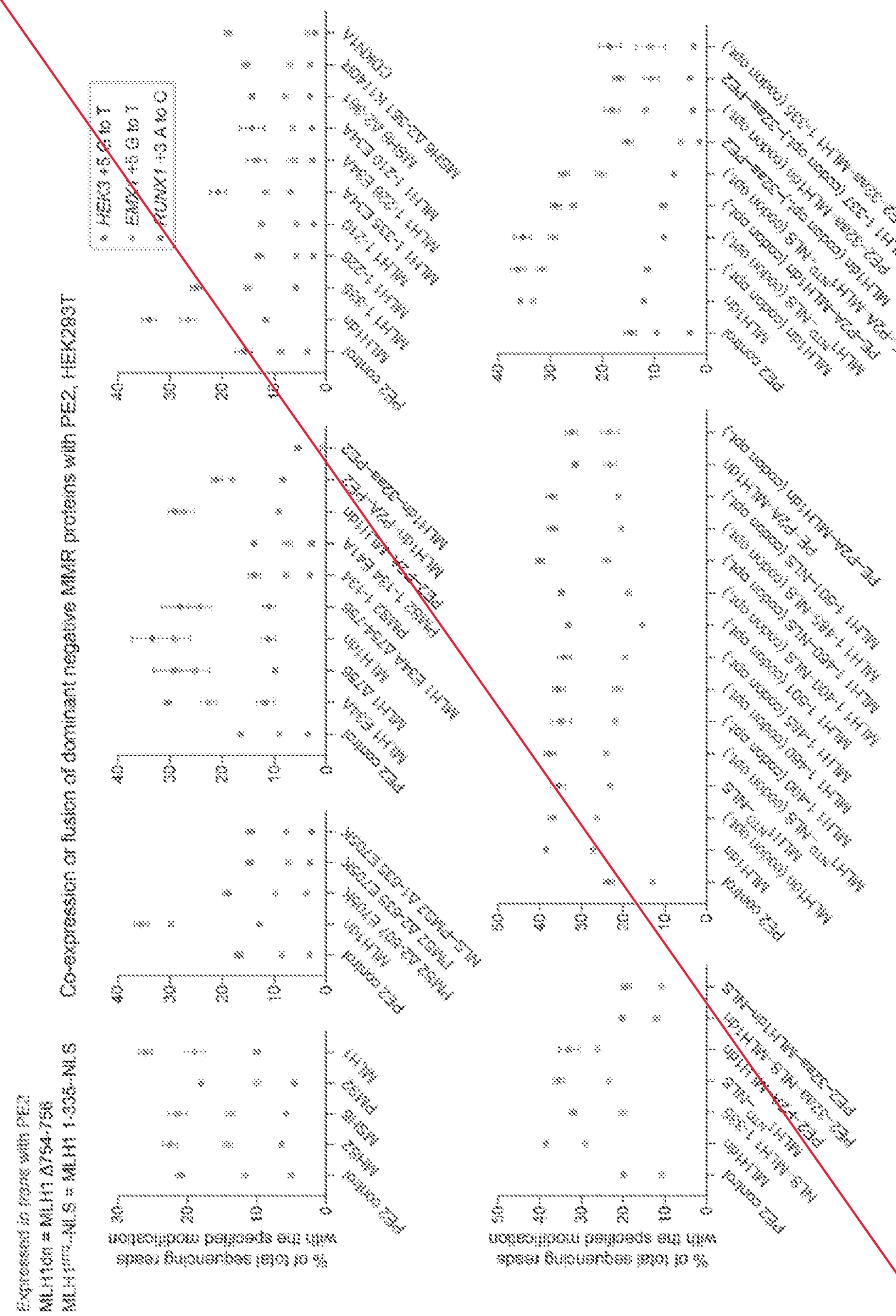

Co-expression or fusion of dominant negative MMR proteins with PE2, HEK293T
HEK3 +5 G to T
EMX1 +5 G to T
RUNX1 +3 A to C
Expressed in trans with PE2
MLH1dn = MLH1 Δ754-756
MLH1 1-335-NLS
% of total sequencing reads with the specified modification
PE2 control
MSH2
MSH6
PMS2
MLH1
MLH1dn
MLH1 Δ756
CRNAVA
FIG. 50A

FIG. 50A

PE2 +MLH1dn, HEK293T
HEK3 +5 G to T, HEK293T
EMX1 +5 G to T, HEK293T
HEK3 +5 G to T
EMX1 +5 G to T
PE2 +RFP
PE2 +MLH1dn
% of total sequencing reads with the specified modification
% of total sequencing reads with the intended edit
MLH1dn transfection dose (ng)
PE2 +expression plasmid total transfection dose (ng)
0
10
20
30
40
0.1
1
10
100

FIG. 50B

FIG. 50C

FIG. 50D

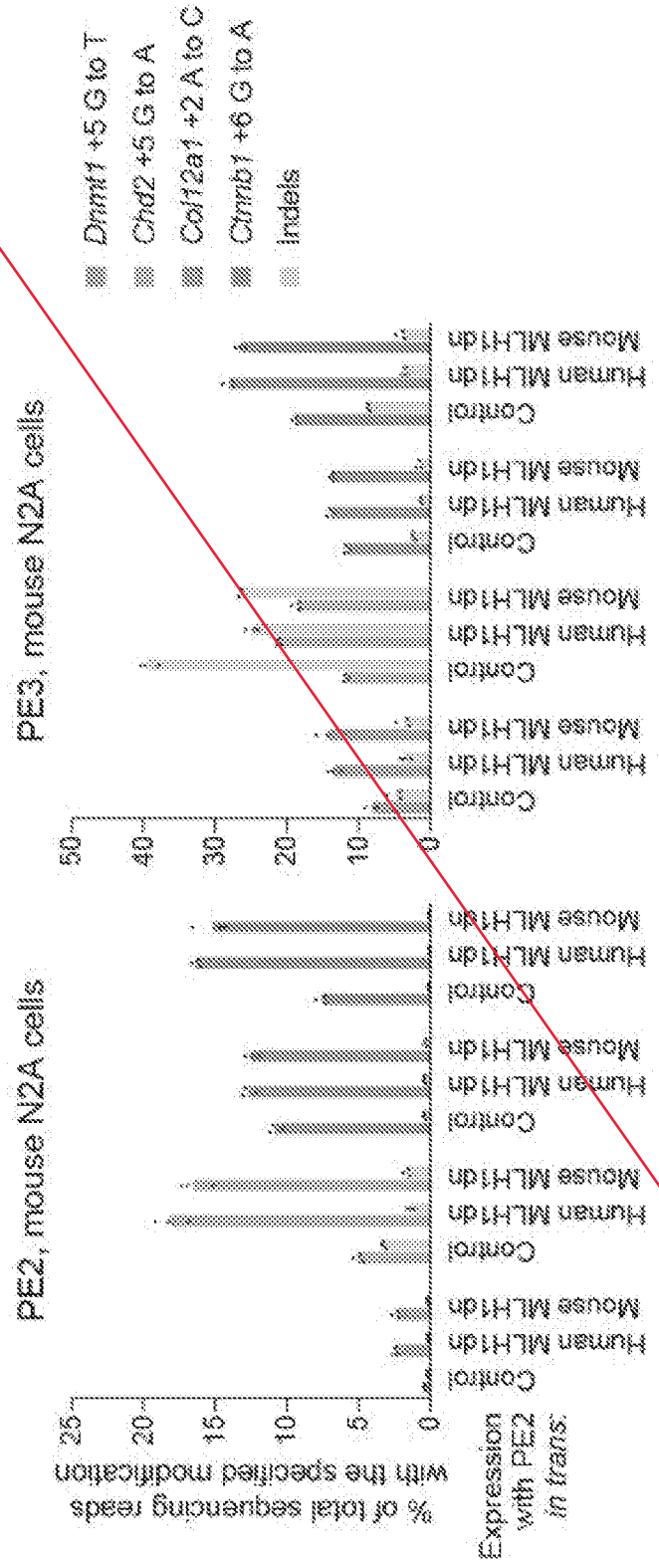

Dnmt1 +5 G to T
Chd2 +5 G to A
Col12a1 +2 A to C
Ctnnb1 +6 G to A
Indels
PE2, mouse N2A cells
PE3, mouse N2A cells
% of total sequencing reads with the specified modification
Expression with PE2 in trans:
Control
Human MLH1dn
Mouse MLH1dn
0
5
10
15
20
25
30
40
50

FIG. 50E

Comparison of MLH1dn and MLH1 knockout
% of total sequencing reads with the intended edit in clonal HeLa cells
HEK3 +1 T to A
EMX1 +5 G to T
FANCF +3 C to T
RUNX1 +5 G to T
Indels
0
5
10
15
20
Control
+MLH1dn
ΔMLH1 (clone 1)
ΔMLH1 (clone 2)
PE without complementary nick (PE2/PE4)
PE with complementary nick (PE3/PE5)

FIG. 50F

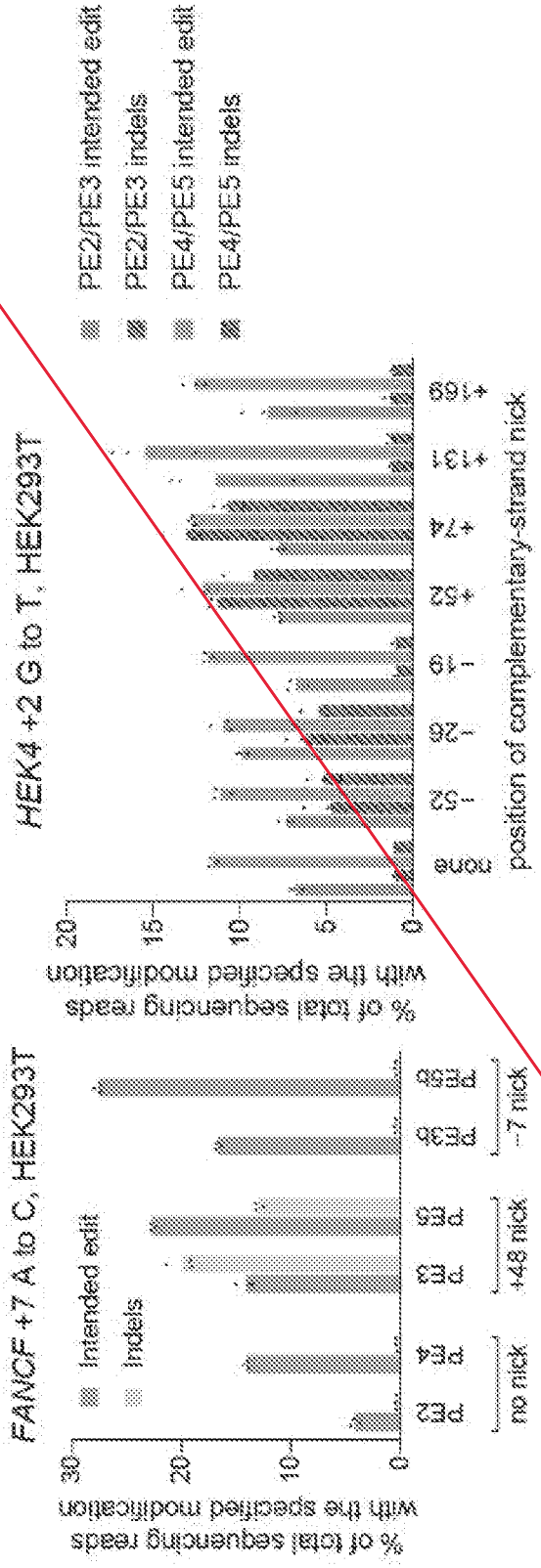


FIG. 50G

FIG. 50H

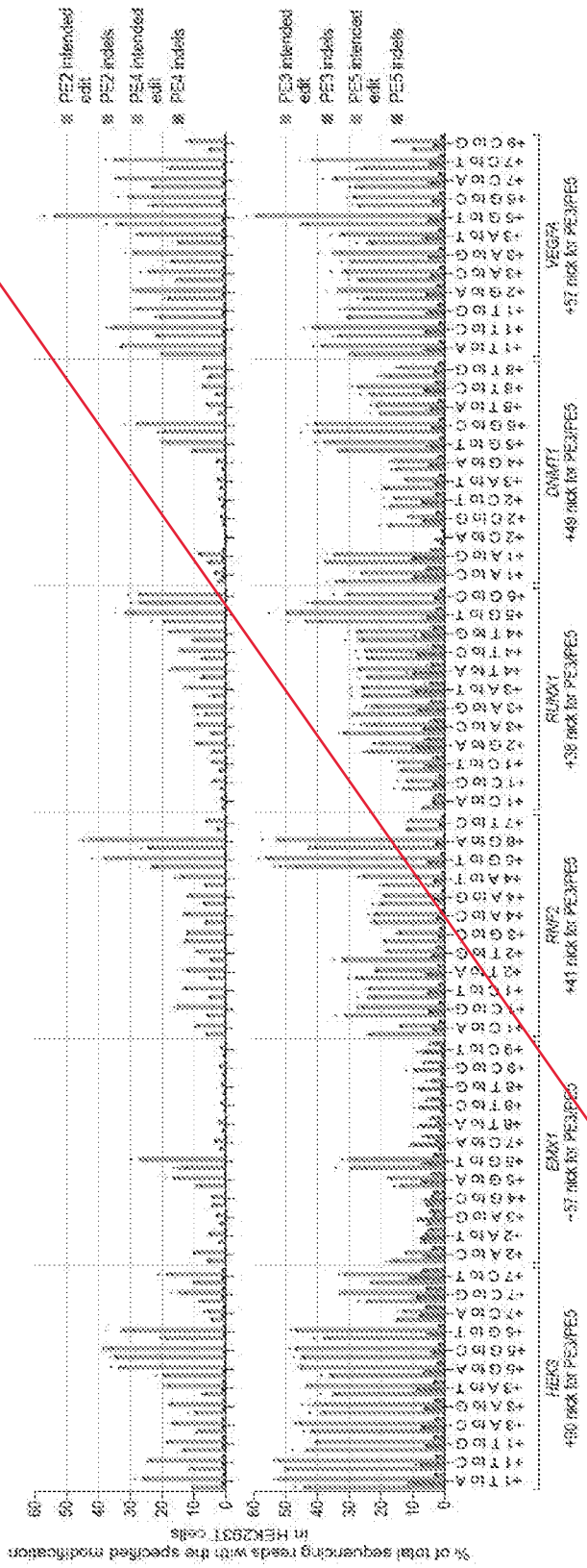

PE2 intended edit
PE2 indels
PE4 intended edit
PE4 indels
PE3 intended edit
PE3 indels
PE5 intended edit
PE5 indels
% of total sequencing reads with the specified modification in HEK293T cells
HEK3
+90 nick for PE3/PE5
EMX1
+57 nick for PE3/PE5
RNF2
+41 nick for PE3/PE5
RUNX1
+38 nick for PE3/PE5
DNMT1
+49 nick for PE3/PE5
VEGFA
+67 nick for PE3/PE5

FIG. 51A

FIG. 51B

FIG. 51C

FIG. 51D

FIG. 51E

FIG. 51F

FIG. 51G

FIG. 51H

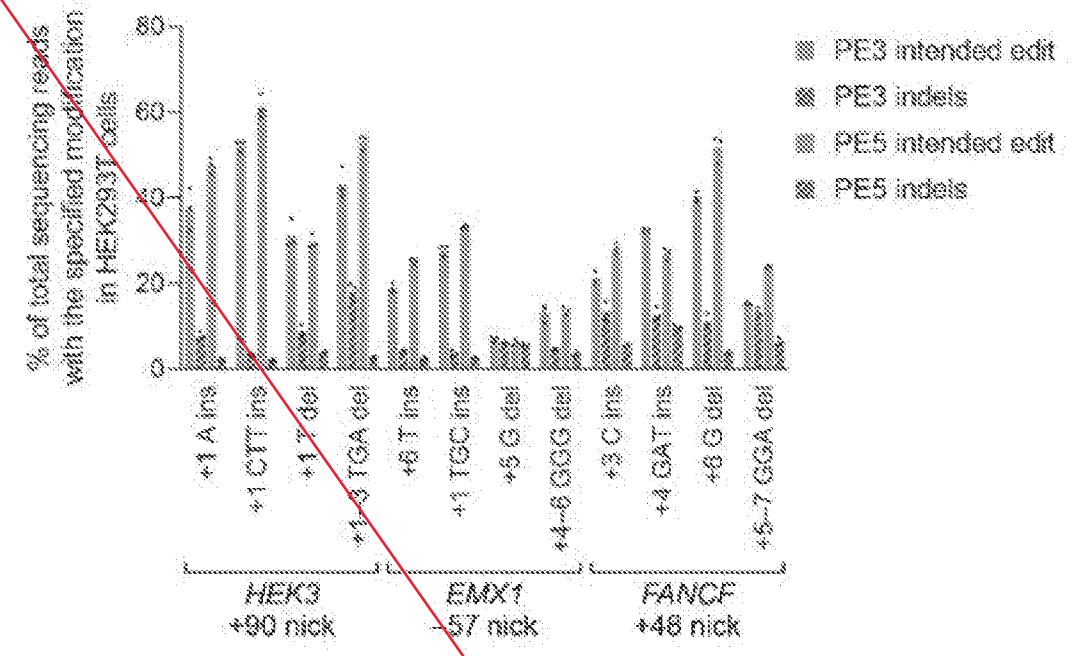


FIG. 51I

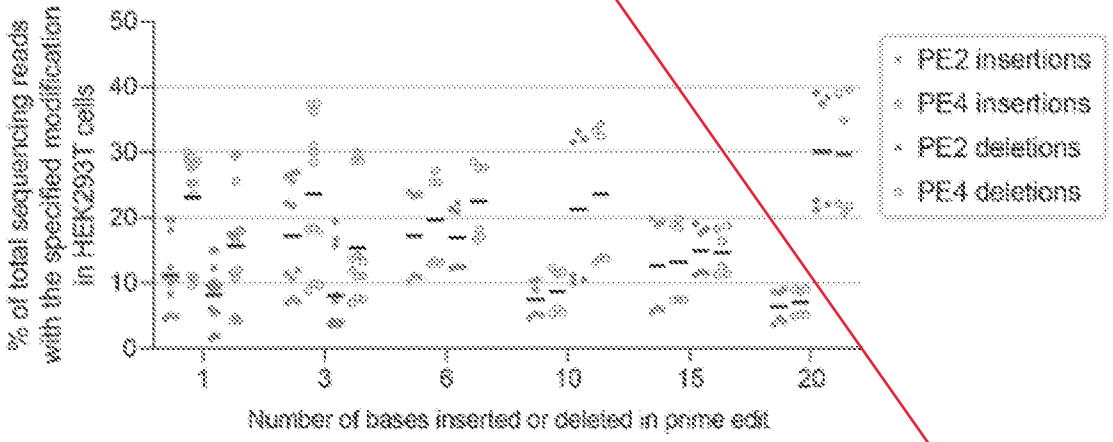


FIG. 51J

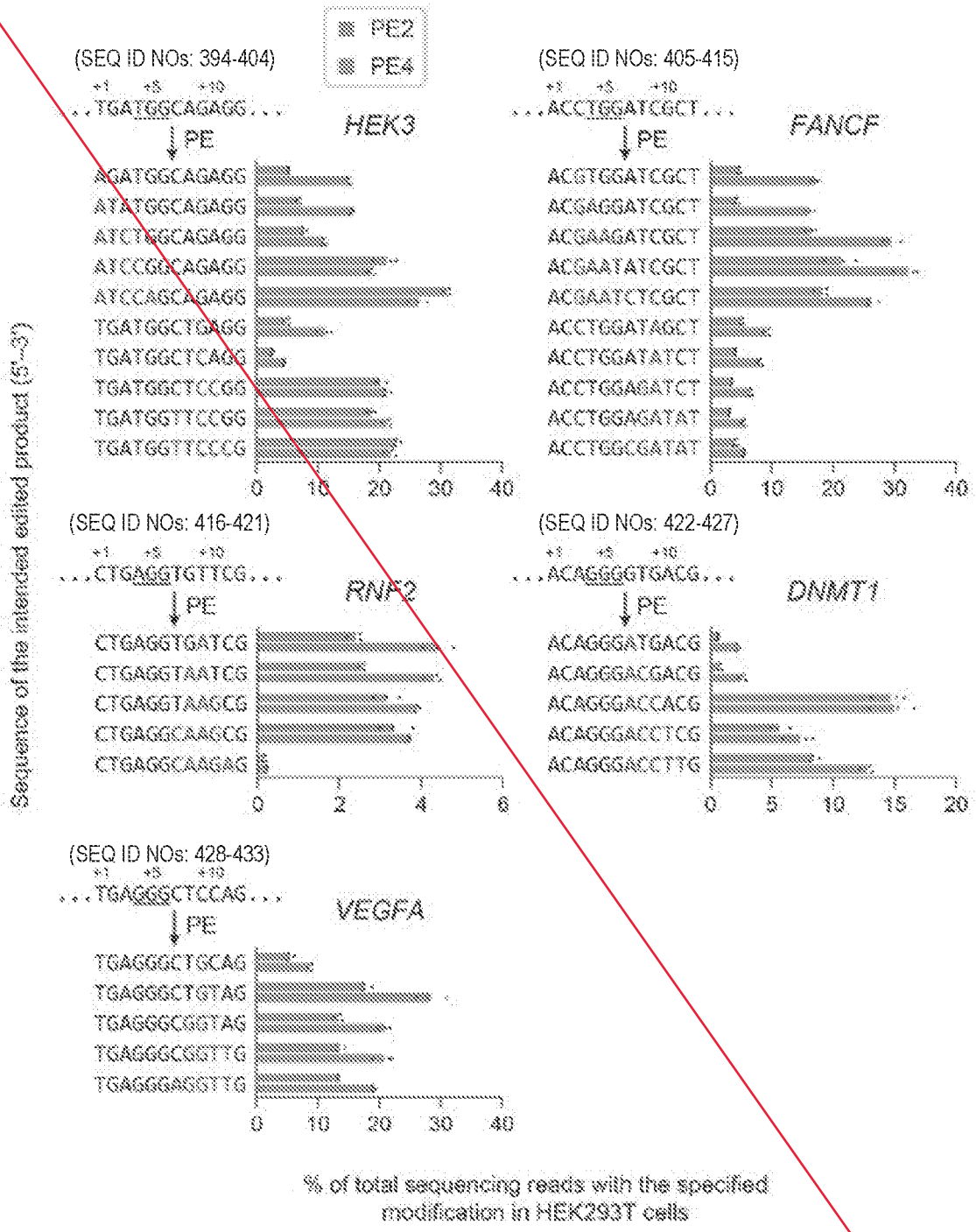

PE2
PE4
(SEQ ID NOs: 394-404)
HEK3
(SEQ ID NOs: 405-415)
FANCF
(SEQ ID NOs: 416-421)
RNF2
(SEQ ID NOs: 422-427)
DNMT1
(SEQ ID NOs: 428-433)
VEGFA
PE
Sequence of the intended edited product (5'-3')
% of total sequencing reads with the specified modification in HEK293T cells

FIG. 52A

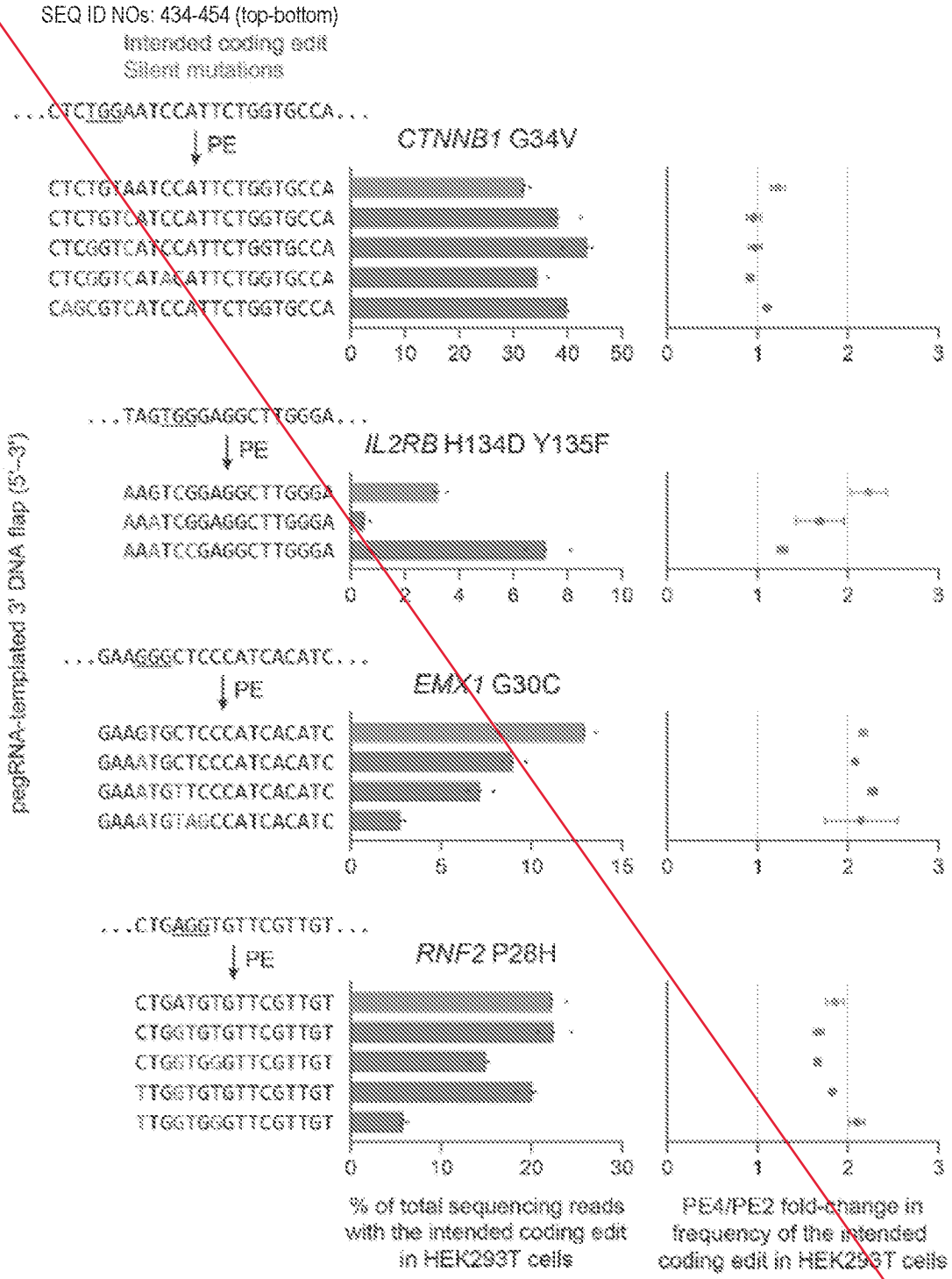

SEQ ID NOs: 434-454 (top-bottom)
Intended coding edit
Silent mutations
...CTCTGGAATCCATTCTGGTGCCA...
PE
CTNNB1 G34V
CTCTGTAATCCATTCTGGTGCCA
CTCTGTCATCCATTCTGGTGCCA
CTCGGTCATCCATTCTGGTGCCA
CTCGGTCATAGATTCTGGTGCCA
CAGCGTCATCCATTCTGGTGCCA
0 10 20 30 40 50
0 1 2 3
...TAGTGGGAGGCTTGGGA...
PE
IL2RB H134D Y135F
AAGTCGGAGGCTTGGGA
AAATCGGAGGCTTGGGA
AAATCCGAGGCTTGGGA
0 2 4 6 8 10
0 1 2 3
...GAAGGGCTCCCATCACATC...
PE
EMX1 G30C
GAAGTGCTCCCATCACATC
GAAATGCTCCCATCACATC
GAAATGTTCCCATCACATC
GAAATGTAGCCATCACATC
0 5 10 15
0 1 2 3
...CTGAGGTGTTCGTTGT...
PE
RNF2 P28H
CTGATGTGTTCGTTGT
CTGGTGTGTTCGTTGT
CTGGTGGGTTCGTTGT
TTGGTGTGTTCGTTGT
TTGGTGGGTTCGTTGT
0 10 20 30
0 1 2 3
pegRNA-templated 3' DNA flap (5'→3')
% of total sequencing reads with the intended coding edit in HEK293T cells
PE4/PE2 fold-change in frequency of the intended coding edit in HEK293T cells

FIG. 52B

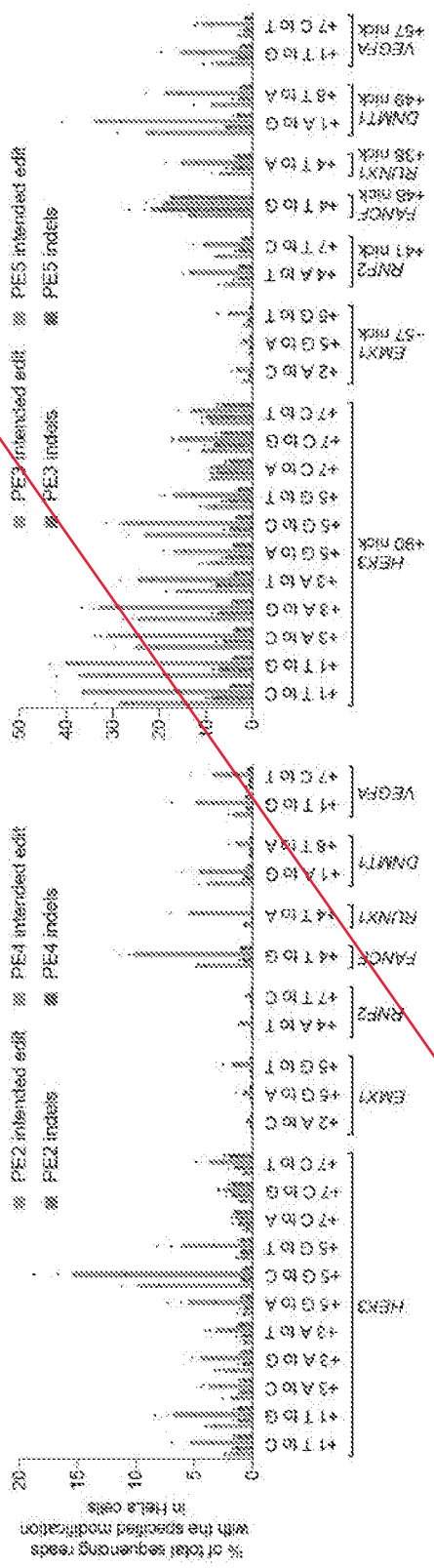

% of total sequencing reads with the specified modification in HeLa cells
PE2 intended edit
PE2 indels
PE3 intended edit
PE4 indels
PE3 intended edit
PE3 indels
+1 T to C
+1 T to G
+3 A to C
+3 A to G
+3 A to T
+5 G to A
+5 G to C
+5 G to T
+7 C to A
+7 C to G
+7 C to T
HEK3
+2 A to C
+5 G to A
+5 G to T
EMX1
+4 A to T
+7 T to C
RNF2
+4 T to G
FANCF
+4 T to A
RUNX1
+1 A to G
+8 T to A
DNMT1
+1 T to G
+7 C to T
VEGFA

FIG. 52C

% of total sequencing reads with indels in HEK293T cells
Untreated
PE2
PE2 +MLH1dn
PE2, non-editing pegRNA
PE2 +MLH1dn, non-editing pegRNA
PE2-dRT
PE2-dRT +MLH1dn
SpCas9 H840A
SpCas9 H840A +MLH1dn
HEK3 +1 T to A +90 nick
EMX1 +5 G to T +53 nick
FANCF +3 C to T +48 nick
DNMT1 +1 A to C +49 nick

FIG. 53A

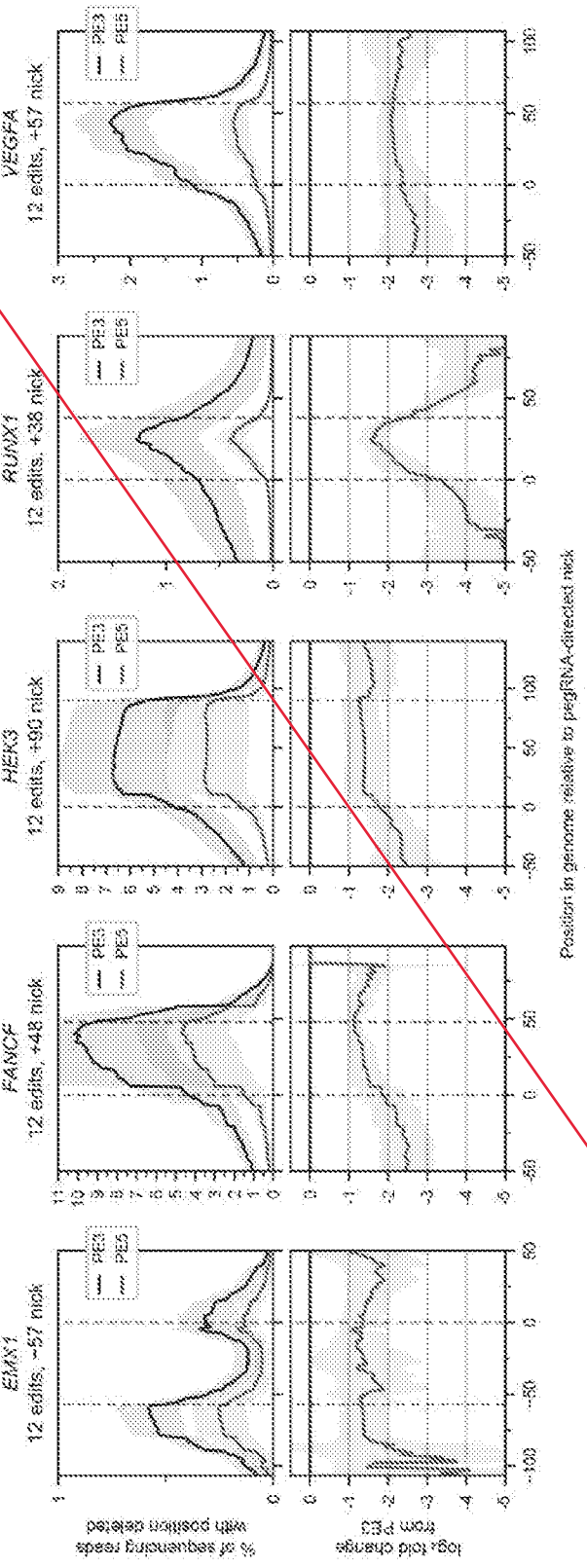

EMX1
12 edits, −57 nick
FANCF
12 edits, +48 nick
HEK3
12 edits, +90 nick
RUNX1
12 edits, +38 nick
VEGFA
12 edits, +57 nick
PE3
PE5
% of sequencing reads with position deleted
log2 fold change from PE3
Position in genome relative to pegRNA-directed nick

FIG. 53B

FIG. 53C

FIG. 53D

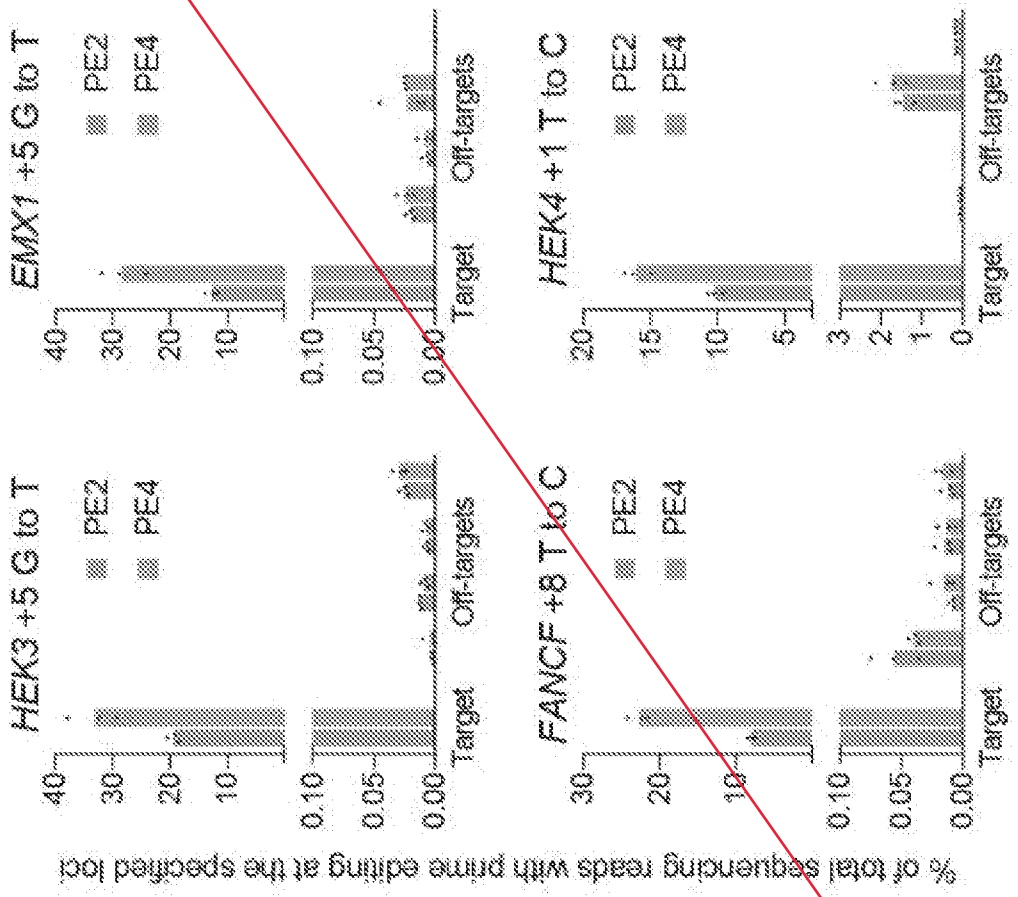

HEK3 +5 G to T
EMX1 +5 G to T
FANCF +8 T to C
HEK4 +1 T to C
PE2
PE4
Target
Off-targets
% of total sequencing reads with prime editing at the specified loci

FIG. 53E

FIG. 53G

FIG. 53F

FIG. 54A

FIG. 54B

HeLa, 7 substitution edits
HEK293T, 7 substitution edits
% of total sequencing reads with the specified modification
Intended edit
Indels
PE2
PE2max
PE4max
PE4max +epegRNA
PE3
PE3max
PE5max
PE5max +epegRNA
2.3×
22×
2.6×
3.2×
2.6×
1.4×
1.2×
2.1×
1.5×
1.2×
1.2×
1.1×

FIG. 54C

PE2
PE2max
PE4max
PE3
PE3max
PE5max
Transfected cell types
% of total sequencing reads with the specified modification in HEK293T cells
% of total sequencing reads with the specified modification in HeLa cells
HBB
-6 G to A
(E6V site)
PRNP
G127V install
(+6 G to T)
CDKL5
+1 C to T
(c.1412delA site)
CXCR4
P191A install
(+5 G to C)
IL2RB
H134D Y135F install
(+1 T to A, +5 G to C)
HBG1/2
BCL11A to GATA1 motif
(+6 G to A, +8 C to A)

FIG. 54D

FIG. 54E

FIG. 54F

PE architectures
PE2
Genscript codon RT
Cas9 R221K N394K
Additional C-term NLS
nCas9–RT linker
% of total sequencing reads with the specified modification in HeLa cells
Fold change in frequency of the intended edit in HeLa cells
HEK3 +1 T to A, +90 nick
FANCF +3 C to T, +48 nick
RUNX1 +3 A to C, +38 nick
VEGFA +7 C to A, +57 nick

FIG. 55A

PE architectures
Cas9
Genscript codon RT
R221K N394K
nCas9–RT linker
N-term NLS
C-term NLS
% of total sequencing reads with the specified modification in HeLa cells
Fold change in frequency of the intended edit in HeLa cells
PE2
PEmax
HEK3 +1 T to A, +40 nick
RNF2 +6 G to A, +41 nick
FANCF +3 C to T, +48 nick
RUNX1 +3 A to C, +38 nick
DNMT1 +6 G to C, +48 nick
VEGFA +7 C to A, +57 nick

FIG. 55B

PE architectures
Codon optimized RT
Cas9 R221K N385K
Additional C-term NLS
Linker between nCas9 and RT
Fold change in frequency of the intended edit in HeLa cells
PE2
PEmax
SGGS×2-XTEN16-SGGS×2 (SEQ ID NO: 125)
SGGS×3-XTEN16-SGGS×2 (SEQ ID NO: 125)
SGGS×2-XTEN16-SGGS×2 (SEQ ID NO: 125)
SGGS×2-XTEN16-SGGS×2 (SEQ ID NO: 125)
SGGS×2-XTEN16-SGGS×2 (SEQ ID NO: 125)
SGGS×2-XTEN16-SGGS×2 (SEQ ID NO: 105)
SGGS×8 (SEQ ID NO: 456)
SGGS×8 (SEQ ID NO: 456)
SGGS×10 (SEQ ID NO: 457)
SGGS×8 (SEQ ID NO: 456)
SGGS×2-SGGSGGS-SGGS×2 (SEQ ID NO: 459)
SGGS×3-XTEN16-SGGS×3 (SEQ ID NO: 111)
SGGS×2-NLS^nuc-SGGS×2 (SEQ ID NO: 105)
SGGS×3-NLS^nuc-SGGS×3 (SEQ ID NO: 112)
SGGS×3-NLS^nuc-SGGS×3 (SEQ ID NO: 112)
NLS^nuc
SGGS×2-XTEN16-SGGS×2 (SEQ ID NO: 125)
SGGS×2-XTEN16-SGGS×2 (SEQ ID NO: 125)
SGGS×2-NLS^nuc-SGGS×2 (SEQ ID NO: 105)
SGGS×2-NLS^nuc-SGGS×2 (SEQ ID NO: 105)
SGGS×2-NLS^nuc-SGGS×2 (SEQ ID NO: 105)
SGGS×2-NLS^nuc-SGGS×2 (SEQ ID NO: 105)
SGGS×2-NLS^nuc-SGGS×2 (SEQ ID NO: 105)
NLS^nuc-NLS^nuc
NLS^nuc-NLS^nuc
NLS^nuc
NLS^nuc
NLS^nuc
NLS^nuc
NLS^nuc
NLS^nuc-NLS^nuc
NLS^nuc-NLS^nuc
NLS^nuc
NLS^nuc-NLS^nuc
NLS^nuc-NLS^nuc
HEK3 +1 T to A, +20 nick
PAPP1 +5 G to A, +41 nick
FANCF +3 C to T, +48 nick
RUNX1 +3 A to C, +38 nick
DNMT1 +5 G to C, +49 nick
VEGFA +7 C to A, +37 nick

FIG. 56A

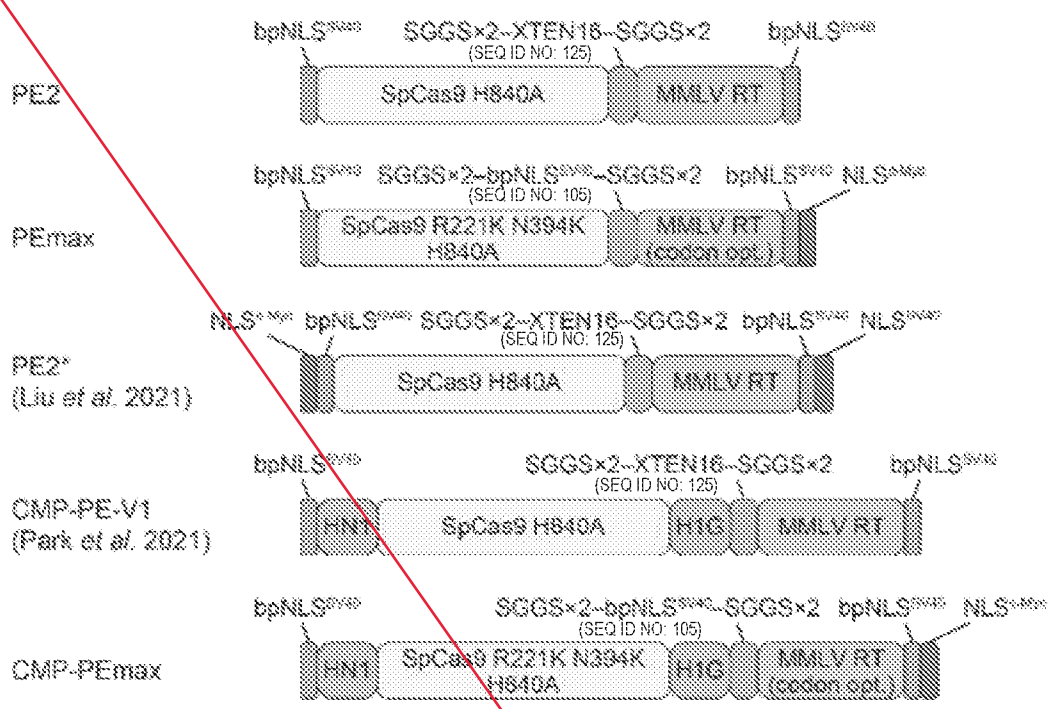

PE2
bpNLS
SGGS×2-XTEN16-SGGS×2
(SEQ ID NO: 125)
bpNLS
SpCas9 H840A
MMLV RT
PEmax
bpNLS
SGGS×2-bpNLS
SGGS×2
(SEQ ID NO: 105)
bpNLS
NLS
SpCas9 R221K N394K H840A
MMLV RT (codon opt.)
PE2*
(Liu et al. 2021)
NLS
bpNLS
SGGS×2-XTEN16-SGGS×2
(SEQ ID NO: 125)
bpNLS
NLS
SpCas9 H840A
MMLV RT
CMP-PE-V1
(Park et al. 2021)
bpNLS
SGGS×2-XTEN16-SGGS×2
(SEQ ID NO: 125)
bpNLS
HN1
SpCas9 H840A
H1G
MMLV RT
CMP-PEmax
bpNLS
SGGS×2-bpNLS
SGGS×2
(SEQ ID NO: 105)
bpNLS
NLS
HN1
SpCas9 R221K N394K H840A
H1G
MMLV RT (codon opt.)

FIG. 56B

% of total sequencing reads with the specified modification in HeLa cells
0
10
20
30
40
Intended edit
Indels
%
PE2
PEmax
PE2*
CMP-PE-V1
CMP-PEmax
HEK3 +1 T to A, +90 nick
EMX1 +5 G to T, +53 nick
RNF2 +6 G to A, +41 nick
FANCF +3 C to T, +48 nick
RUNX1 +3 A to C, +38 nick
DNMT1 +6 G to C, +48 nick
VEGFA +7 C to A, +57 nick

FIG. 56C

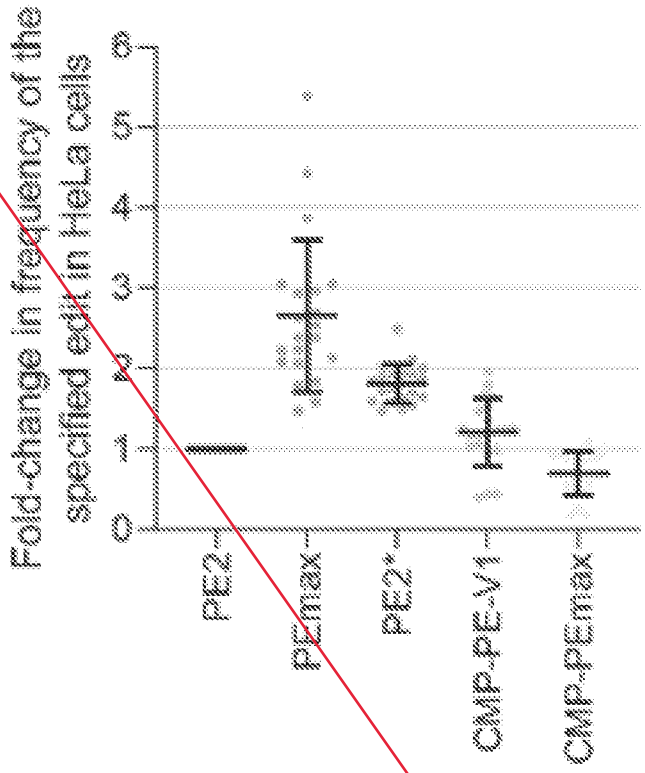

Fold-change in frequency of the specified edit in HeLa cells
0
1
2
3
4
5
6
PE2
PEmax
PE2*
CMP-PE-V1
CMP-PEmax

FIG. 56D

FIG. 56E

FIG. 56F

PRNP G127V install (+6 G to T)
FANCF +5 G to T
RNF2 +1 T insertion
CXCR4 P191A install (+5 G to C)
IL2RB H134D Y135F install (+1 T to A, +5 G to C)
% of total sequencing reads with the specified modifications in primary T cells
Intended edit
Indels
PE2 mRNA (pg):
MLH1dn mRNA (pg):
+72 nick
+48 nick
+41 nick
No nick
+43 nick
No nick
+55 nick

FIG. 56G

3' DNA flap templated by pegRNA used in CRISPRi screens
Genomic target sequence
pegRNA RT template
pegRNA scaffold
3'
5'
(SEQ ID NO: 460)
5' TGGTTGGGGTTGAATTCTTGCCAAAAA 3'
5' TGGTTGGGGTTGAATTCTTGCCAAAGT 3'
5-nt homology between reverse transcribed pegRNA scaffold and genomic sequence
3' DNA flap templated by pegRNA with recoded scaffold
Genomic target sequence
recoded
pegRNA RT template
pegRNA scaffold
5'
3'
(SEQ ID NO: 462)
5' TGGTTGGGGTTGAATTCTTGCCAAAAA 3'
5' TGGTTGGGGTTGAATTCTTGCCAAAGT 3'
Position of intended G•C to C•G edit
Position of unintended edits from reverse transcription into pegRNA scaffold

FIG. 57A

Percentages of reads (log scale)
0.01
0.1
0.51
5
10
Log2 fold change from non-targeting
-2
0
2
MSH2 sgRNA
Non-targeting sgRNA
PE2 screen pegRNA
Intended edit
Installation of additional edits from nearby-matched scaffold
All others
PE2 pegRNA with recoded scaffold
Intended edit
Installation of additional edits from nearby-matched scaffold
All others
PE3+50 screen pegRNA
Intended edit
Joining of reverse-transcribed sequences at unintended location
Deletion
Tandem duplication
Installation of additional edits from nearby-matched scaffold
All others
PE2 pegRNA with recoded scaffold
Intended edit
Joining of reverse-transcribed sequences at unintended location
Deletion
Tandem duplication
Installation of additional edits from nearby-matched scaffold
All others

FIG. 57B

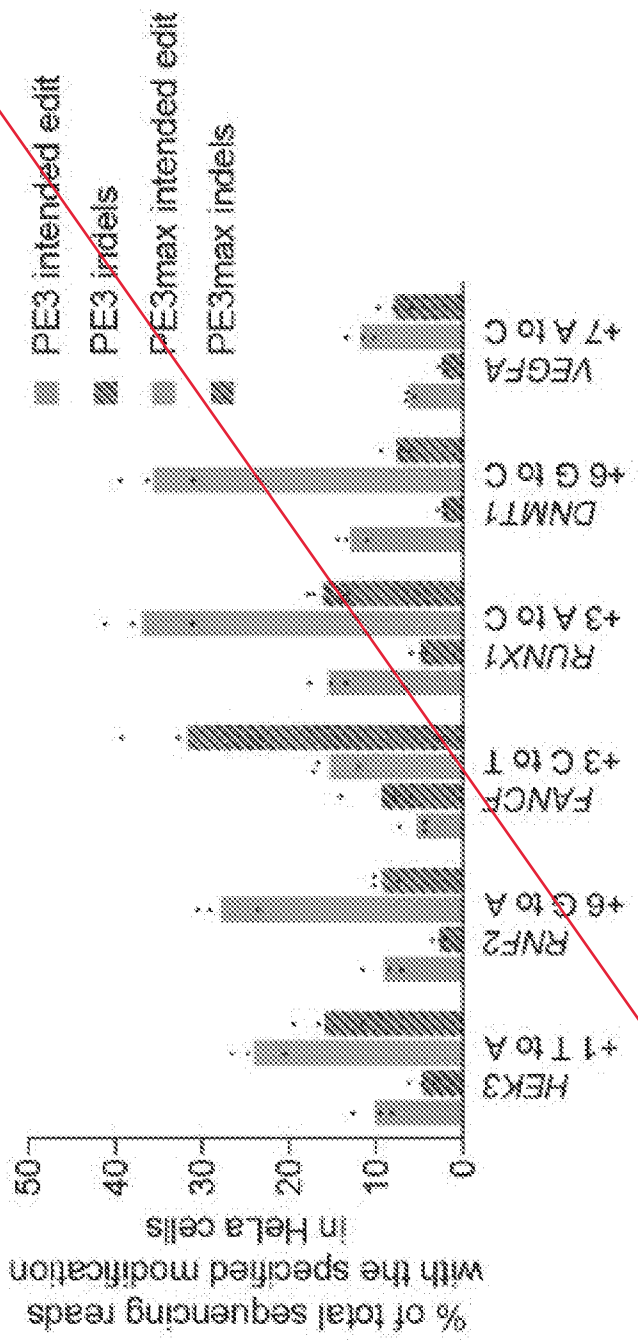

PE3 intended edit
PE3 indels
PE3max intended edit
PE3max indels
% of total sequencing reads with the specified modification in HeLa cells
0
10
20
30
40
50
HEK3 +1 T to A
RNF2 +6 G to A
FANCF +3 C to T
RUNX1 +3 A to C
DNMT1 +6 G to C
VEGFA +7 A to C

FIG. 58

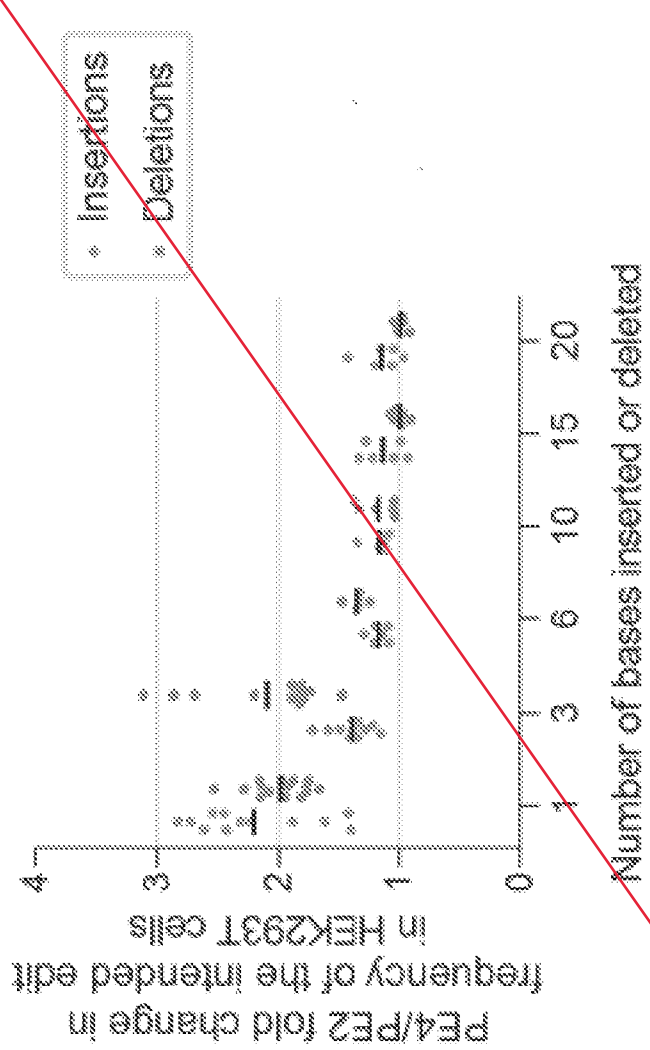

Insertions
Deletions
PE4/PE2 fold change in
frequency of the intended edit
in HEK293T cells
0
1
2
3
4
1
3
6
10
15
20
Number of bases inserted or deleted
in prime edit

FIG. 59

HEK293T cells
pegRNA-encoding lentivirus
transduction at low MOI (1 d)
selection (4 d)
cells with single pegRNA integrant
PE plasmid
transfection (5 d)
edited cells
sequence edited site
PE2
PE2 +epegRNA
PE4 +epegRNA
% of total sequencing reads with the intended edit
0
20
40
60
80
HEK3 +1 T to A
HEK3 +1 CTT ins
DNMT1 +5 G to T
VEGFA +5 G to T

FIG. 60

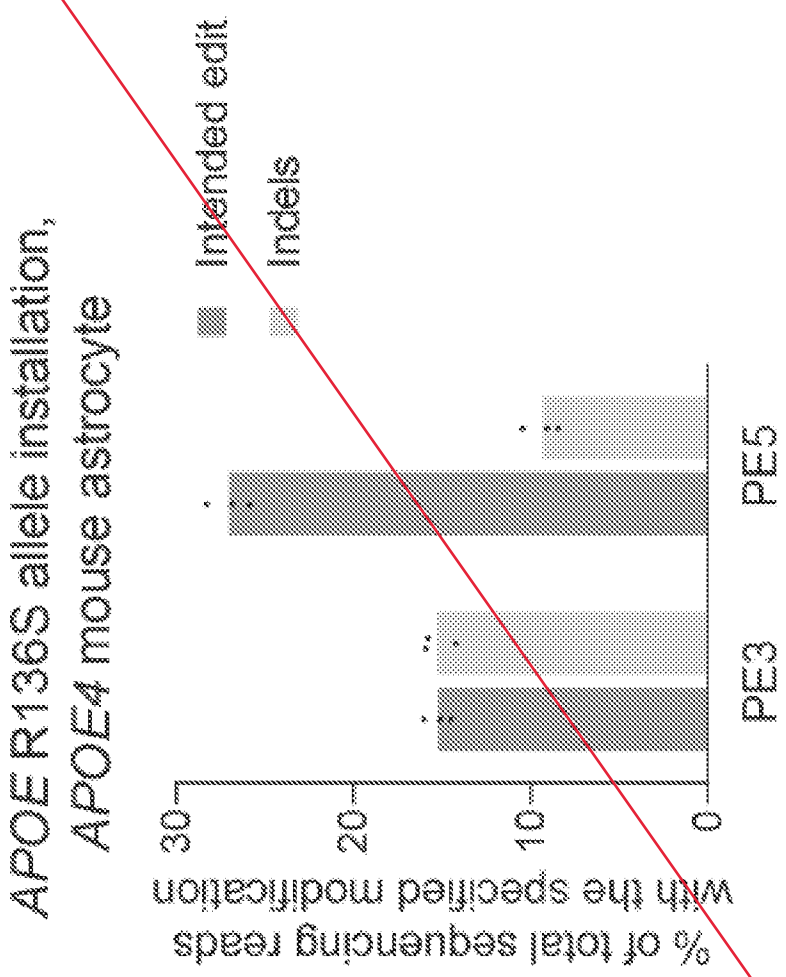

APOE R136S allele installation,
APOE4 mouse astrocyte
Intended edit
Indels
% of total sequencing reads with the specified modification
30
20
10
0
PE3
PE5

FIG. 61

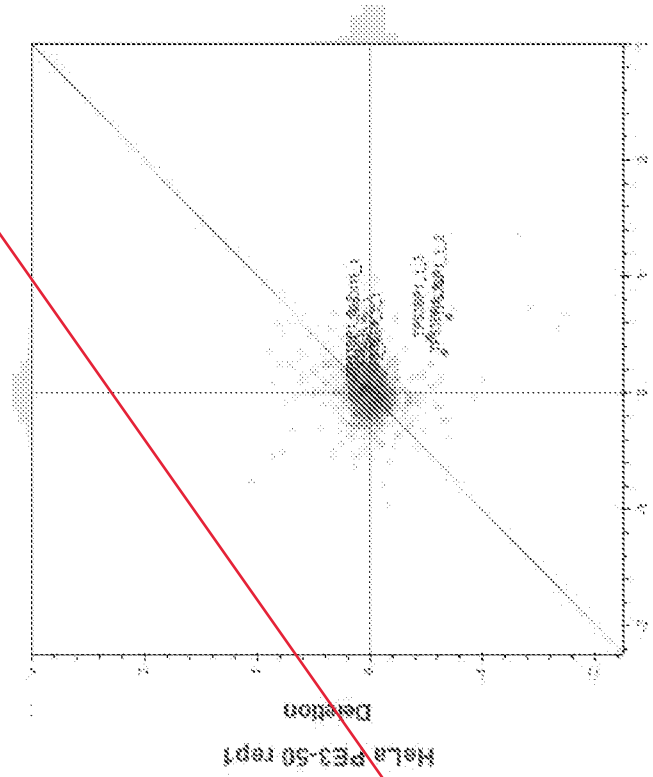


FIG. 62B

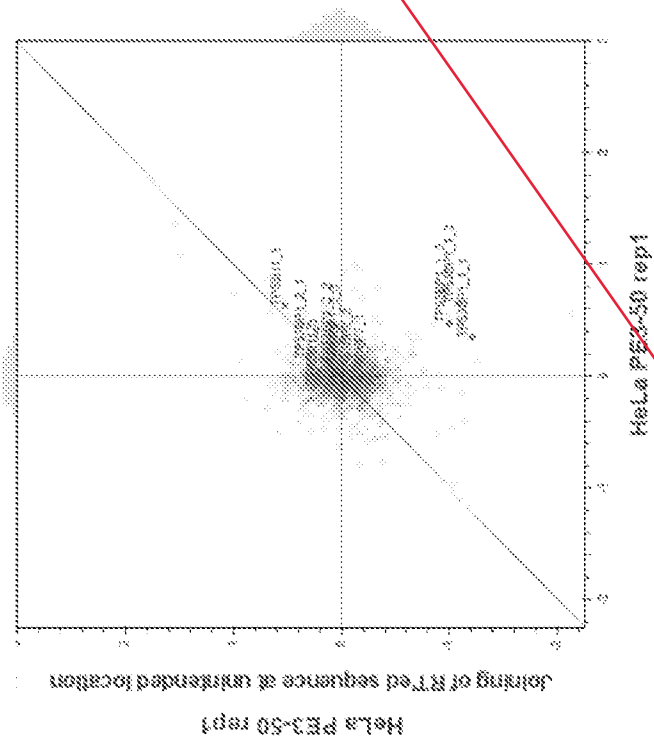


FIG. 62A

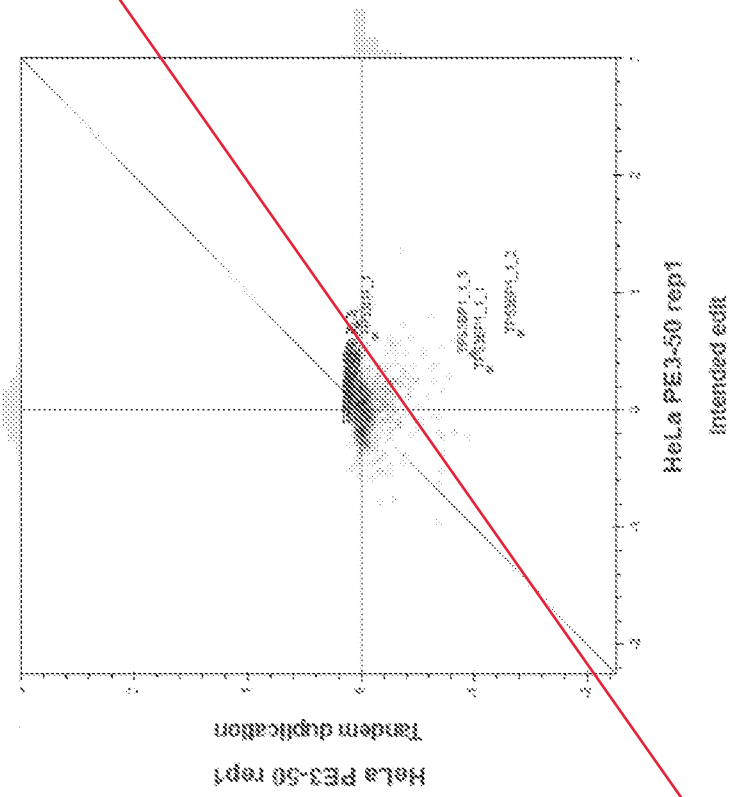

HeLa PE3-30 rep1
Tandem duplication
HeLa PE3-30 rep1
Intended edit

FIG. 62C

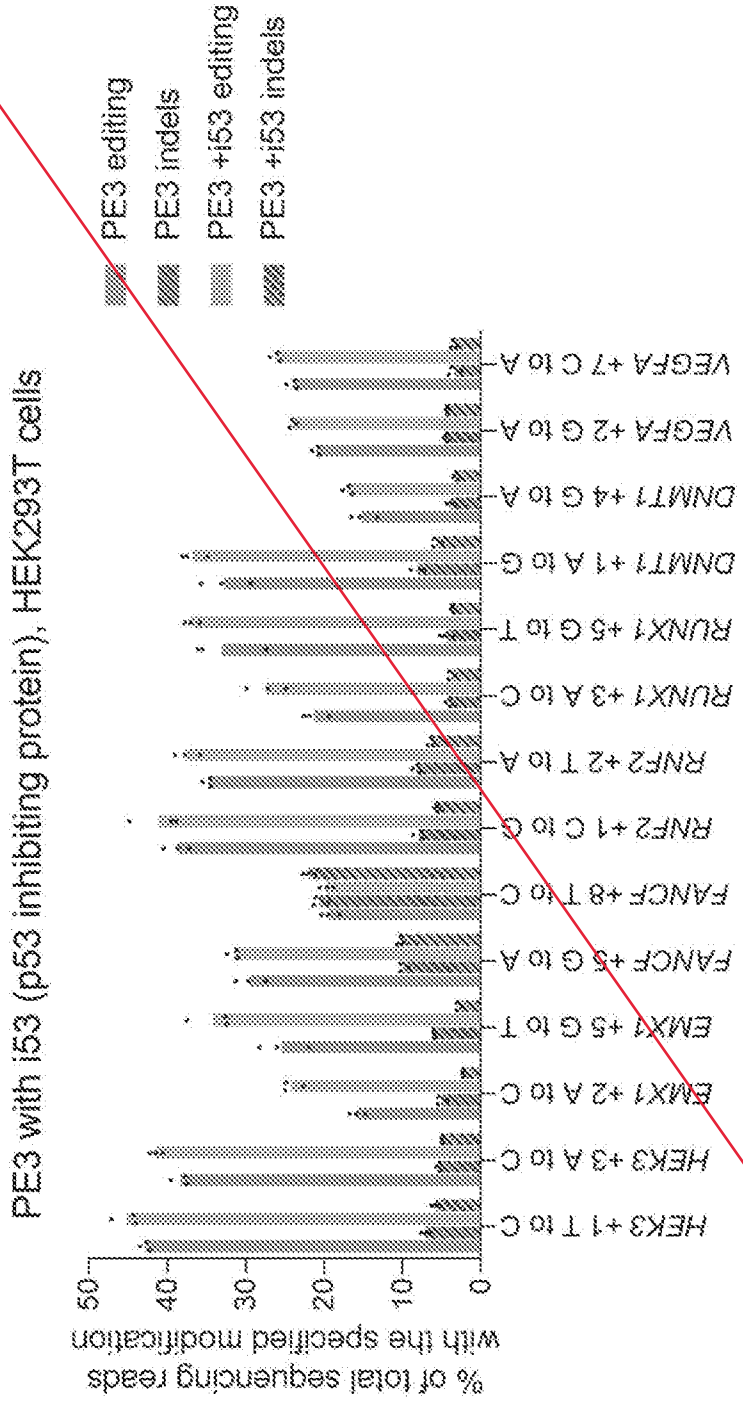

PE3 with i53 (p53 inhibiting protein), HEK293T cells
% of total sequencing reads with the specified modification
0
10
20
30
40
50
PE3 editing
PE3 indels
PE3 +i53 editing
PE3 +i53 indels
HEK3 +1 T to C
HEK3 +3 A to C
EMX1 +2 A to C
EMX1 +5 G to T
FANCF +5 G to A
FANCF +8 T to C
RNF2 +1 C to G
RNF2 +2 T to A
RUNX1 +3 A to C
RUNX1 +5 G to T
DNMT1 +1 A to G
DNMT1 +4 G to A
VEGFA +2 G to A
VEGFA +7 C to A

FIG. 63

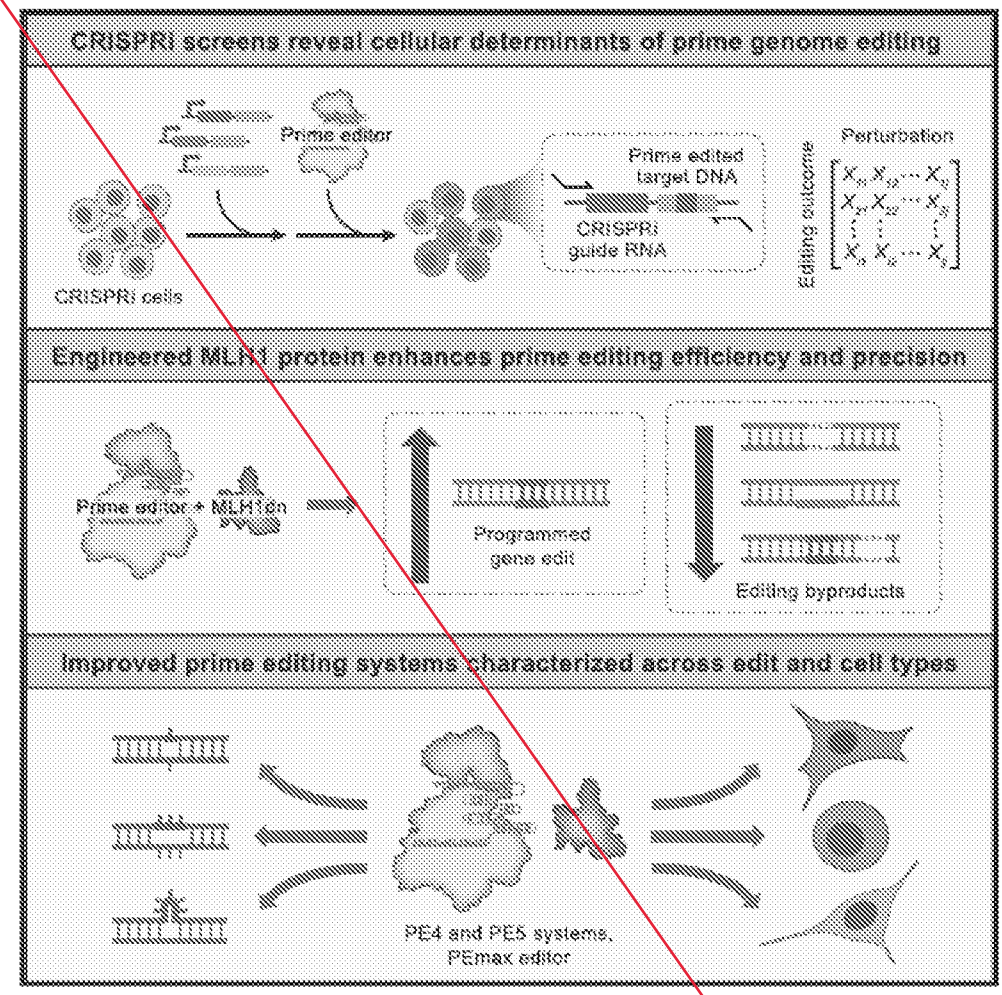

CRISPRi screens reveal cellular determinants of prime genome editing
Prime editor
CRISPRi cells
Prime edited target DNA
CRISPRi guide RNA
Perturbation
Editing outcome
Engineered MLH1d protein enhances prime editing efficiency and precision
Prime editor + MLH1dn
Programmed gene edit
Editing byproducts
Improved prime editing systems characterized across edit and cell types
PE4 and PE5 systems, PEmax editor

FIG. 64

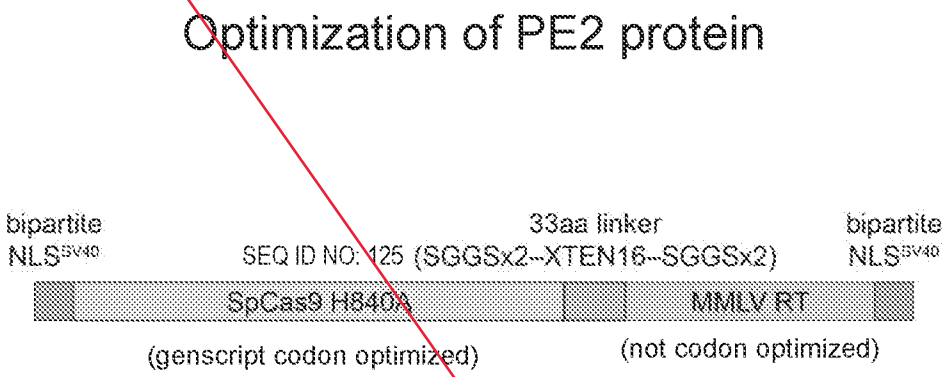

Optimization of PE2 protein
bipartite NLS^{SV40}
33aa linker
SEQ ID NO: 125 (SGGSx2-XTEN16-SGGSx2)
bipartite NLS^{SV40}
SpCas9 H840A
MMLV RT
(genscript codon optimized)
(not codon optimized)
1. Codon optimize MMLV RT
2. Mutations to SpCas9 to improve activity
3. Linker optimization
4. Improve nuclear localization

FIG. 65

PE2 in HEK293T cells (low plasmid)
PE constructs
default PE2 linker
PE2
Genscript codon RT
Cas9 R221K N394K
additional C-term NLS
nCas9–RT linker
SGGSx2–XTEN16–SGGSx2 (SEQ ID NO: 125)
SGGSx2–XTEN16–SGGSx2 (SEQ ID NO: 125)
SGGSx2–XTEN16–SGGSx2 (SEQ ID NO: 125)
SGGSx2–XTEN16–SGGSx2 (SEQ ID NO: 125)
SGGSx2–XTEN16–SGGSx2 (SEQ ID NO: 105)
SGGSx6 (SEQ ID NO: 455)
SGGSx8 (SEQ ID NO: 456)
SGGSx10 (SEQ ID NO: 457)
SGGSGx6 (SEQ ID NO: 458)
SGGSx2–GGSGx4–SGGSx2 (SEQ ID NO: 459)
SGGSx3–XTEN16–SGGSx3 (SEQ ID NO: 111)
SGGSx2–MLS^opt4–SGGSx2 (SEQ ID NO: 105)
SGGSx3–MLS^opt9–SGGSx3 (SEQ ID NO: 112)
GGSGx3–MLS^opt4–GGSGx3 (SEQ ID NO: 112)
fold change in frequency of the intended edit in HEK293T cells (low plasmid dose)
0.0
0.5
1.0
1.5
2.0
HEK3 +1 T to A
EMX1 +5 G to T
RNF2 +6 G to A
FANCF +3 C to T
RUNX1 +4 T to C
DNMT1 +5 G to T
VEGFA +7 A to C
HEK4 +4 G to T
FRNP G127V install
APOE R136S install
CXCR4 P191A install
HEK3 +1 FLAG ins
1.2-fold improvement with best linker
+ Cas9 R221K N394K
+ codon optimization
in HEK293T cells

FIG. 66

PE3 in HeLa cells
default PE2 linker
PE constructs
Genscript codon RT
Cas9 R221K N394K
additional C-term NLS
nCas9–RT linker
fold change in frequency of the intended edit in HeLa cells
0
1
2
3
4
SGGSx2-XTEN16-SGGSx2 (SEQ ID NO: 125)
SGGSx2–XTEN16–SGGSx2 (SEQ ID NO: 125)
SGGSx2–XTEN16–SGGSx2 (SEQ ID NO: 125)
SGGSx2–XTEN16–SGGSx2 (SEQ ID NO: 125)
SGGSx2–XTEN16–SGGSx2 (SEQ ID NO: 105)
SGGSx6 (SEQ ID NO: 455)
SGGSx8 (SEQ ID NO: 456)
SGGSx10 (SEQ ID NO: 457)
GGGGx6 (SEQ ID NO: 458)
SGGSx2–GGGGx4–SGGSx2 (SEQ ID NO: 459)
SGGSx3–XTEN16–SGGSx3 (SEQ ID NO: 111)
SGGSx2–NLS^SV40–SGGSx2 (SEQ ID NO: 105)
SGGSx3–NLS^SV40–SGGSx3 (SEQ ID NO: 112)
GGGGx3–NLS^SV40–GGGGx3 (SEQ ID NO: 112)
NLS^Nucl
NLS^SV40
1.4-fold improvement from codon optimized RT
1.2-fold improvement from additional NLS
1.4-fold improvement from Cas9 R221K N394K mutations
(SEQ ID NO: 105)
SGGSx2–NLS^SV40–SGGSx2 is the best linker
(2.1-fold improvement with linker + codon + Cas9 mutations)
• HEK3 +1 T to A
• FANCF +3 C to T
• RUNX1 +3 A to C
• VEGFA +7 C to A

FIG. 67

HEK293T vs. HeLa editing
default PE2 linker
PE constructs
Genscript codon RT
Cas9 R221K N394K
additional C-term NLS
nCas9–RT linker
fold change in frequency of intended prime edits under the specified conditions
0
1
2
3
4
PE2 in HEK293T (low plasmid)
PE3 in HeLa
improvements matter much more in HeLa cells
SGGSx2–XTEN16–SGGSx2 (SEQ ID NO: 125)
SGGSx2–XTEN16–SGGSx2 (SEQ ID NO: 125)
SGGSx2–XTEN16–SGGSx2 (SEQ ID NO: 125)
SGGSx2–XTEN16–SGGSx2 (SEQ ID NO: 125)
SGGSx2–XTEN16–SGGSx2 (SEQ ID NO: 105)
SGGSx6 (SEQ ID NO: 455)
SGGSx8 (SEQ ID NO: 456)
SGGSx10 (SEQ ID NO: 457)
GGSGx6 (SEQ ID NO: 458)
SGGSx2–GSGSx4–SGGSx2 (SEQ ID NO: 459)
SGGSx3–XTEN16–SGGSx3 (SEQ ID NO: 111)
SGGSx2–NLS^nuc–SGGSx2 (SEQ ID NO: 105)
SGGSx3–NLS^nuc–SGGSx3 (SEQ ID NO: 112)
GGSGx3–NLS^nuc–GGSGx3 (SEQ ID NO: 112)

FIG. 68

PE3 in HeLa cells, NLS architecture optimization
default PE2 linker
*NLSnp48 is missing 1aa
HEK3 +1 T to A
RNF2 +6 G to A
FANCF +3 C to T
RUNX1 +3 A to C
DNMT1 +6 G to C
VEGFA_7 C to A
PE constructs
Gensript codon RT
Cas9 R221K N394K
nCas9–RT linker
N-term NLS
C-term NLS
% of total sequencing reads with the specified modification in HeLa cells
fold change in frequency of the intended edit in HeLa cells
0 10 20 30 40 50
0 1 2 3 4
- - SGGSx2–XTEN16–SGGSx2 (SEQ ID NO: 125) NLSnp48 *NLSnp48
+ - SGGSx2–XTEN16–SGGSx2 (SEQ ID NO: 125) NLSnp48 *NLSnp48
+ + SGGSx2–XTEN16–SGGSx2 (SEQ ID NO: 125) NLSnp48 *NLSnp48
+ + SGGSx2–NLSnp48–SGGSx2 (SEQ ID NO: 105) NLSnp48 *NLSnp48
+ + SGGSx2–NLSnp48–SGGSx2 (SEQ ID NO: 105) NLSnp48 NLSnp48
+ + SGGSx2–NLSnp48–SGGSx2 (SEQ ID NO: 105) NLSnp48 NLSnp48–NLSnp48
+ + SGGSx2–NLSnp48–SGGSx2 (SEQ ID NO: 105) NLSnp48 NLSnp48–NLSnp48
+ + SGGSx2–NLSnp48–SGGSx2 (SEQ ID NO: 105) none NLSnp48
+ + SGGSx2–NLSnp48–SGGSx2 (SEQ ID NO: 105) none NLSnp48–NLSnp48
+ + SGGSx2–NLSnp48–SGGSx2 (SEQ ID NO: 105) none NLSnp48–NLSnp48
"PEmax"

FIG. 69

Final PEmax construct
Streptococcus pyogenes Cas9
(R221K N394K H840A)
bipartite
NLS^SV40
SGGSx2-NLS^nucleoplasmin-SGGSx2
(SEQ ID NO: 105) linker
Genscript codon optimized
MMLV RT pentamutant
(D200N T306K W313F T330P L603W)
bipartite
NLS^SV40
NLS^c-myc
PEmax improves prime editing efficiency
by 2.59-fold across 6 sites in HeLa cells

FIG. 70

PEmax increases indels in addition to the intended edit
% of total sequencing reads with the specified modification in HeLa cells
0
10
20
30
40
50
PE3 intended edit
PE3 indels
PE3max intended edit
PE3max indels
HEK3 +1 T to A
RNF2 +6 G to A
FANCF +3 C to T
RUNX1 +3 A to C
DNMT1 +6 G to C
VEGFA +7 A to C

FIG. 71

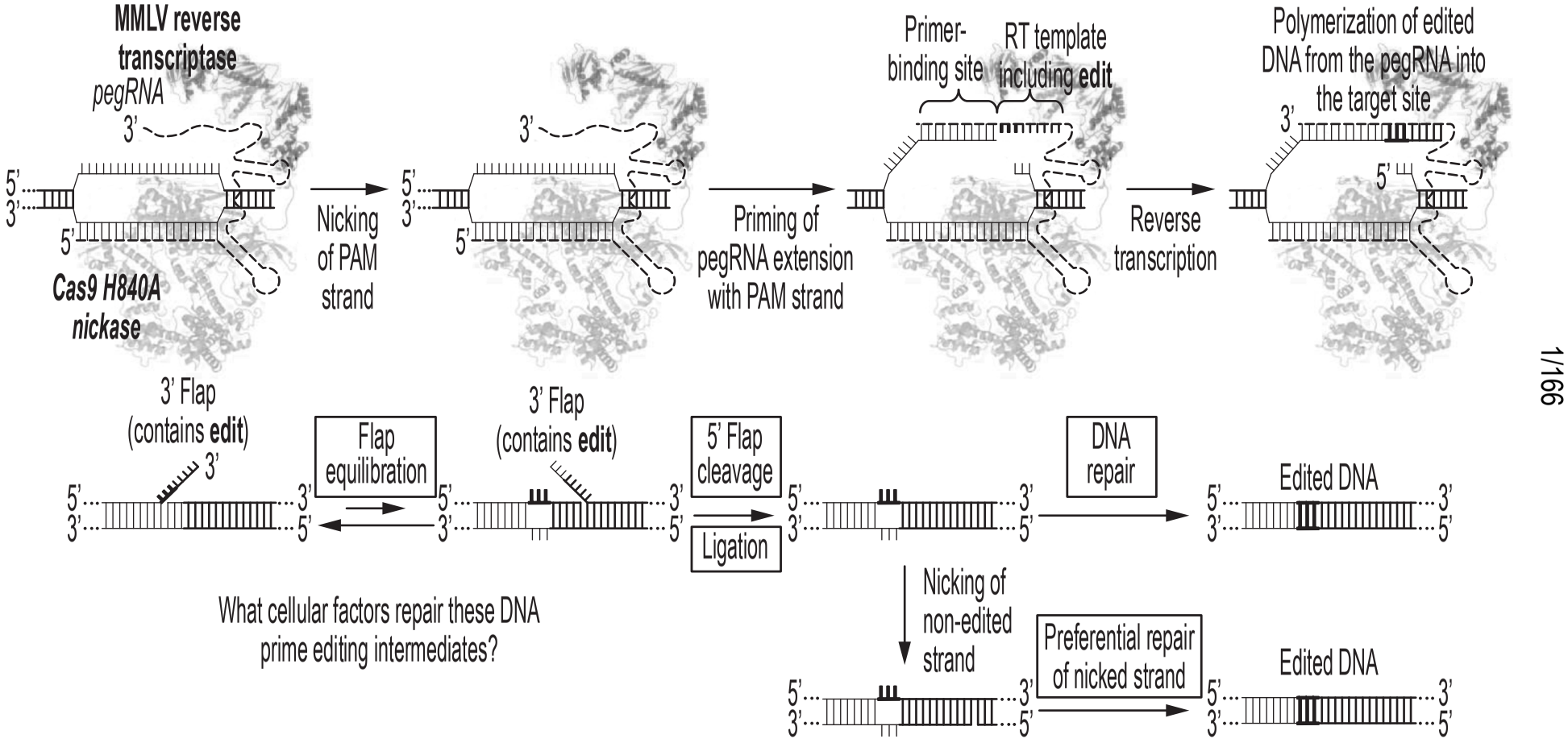

MMLV reverse transcriptase
pegRNA
3'
5'
3'
5'
Cas9 H840A nickase
Nicking of PAM strand
Priming of pegRNA extension with PAM strand
Primer-binding site
RT template including edit
Reverse transcription
Polymerization of edited DNA from the pegRNA into the target site
3' Flap (contains edit)
Flap equilibration
5' Flap cleavage
Ligation
DNA repair
Edited DNA
What cellular factors repair these DNA prime editing intermediates?
Nicking of non-edited strand
Preferential repair of nicked strand
Edited DNA

FIG. 1

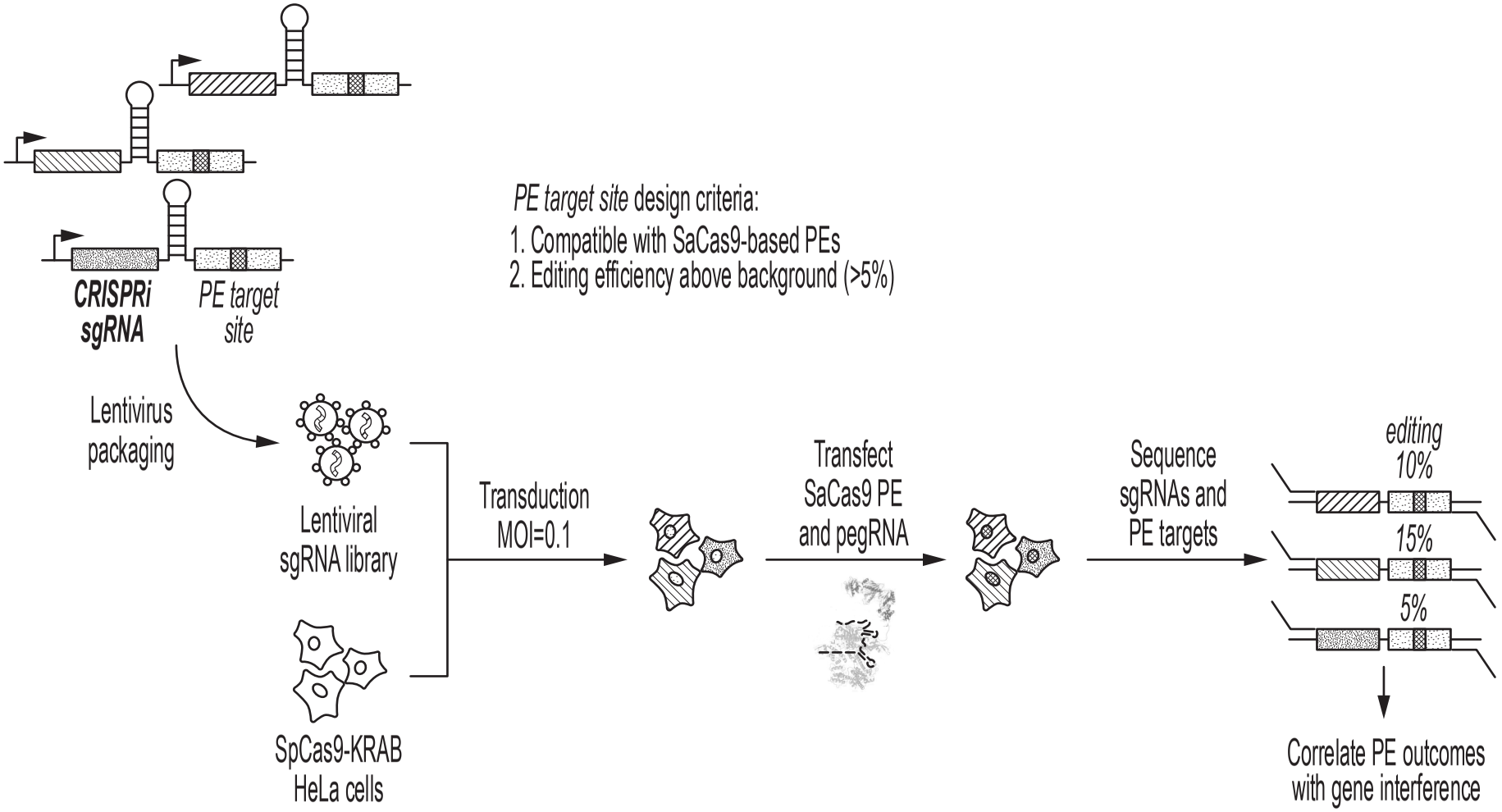

CRISPRi sgRNA
PE target site
PE target site design criteria:
1. Compatible with SaCas9-based PEs
2. Editing efficiency above background (>5%)
Lentivirus packaging
Lentiviral sgRNA library
SpCas9-KRAB HeLa cells
Transduction MOI=0.1
Transfect SaCas9 PE and pegRNA
Sequence sgRNAs and PE targets
editing 10%
15%
5%
Correlate PE outcomes with gene interference

FIG. 2

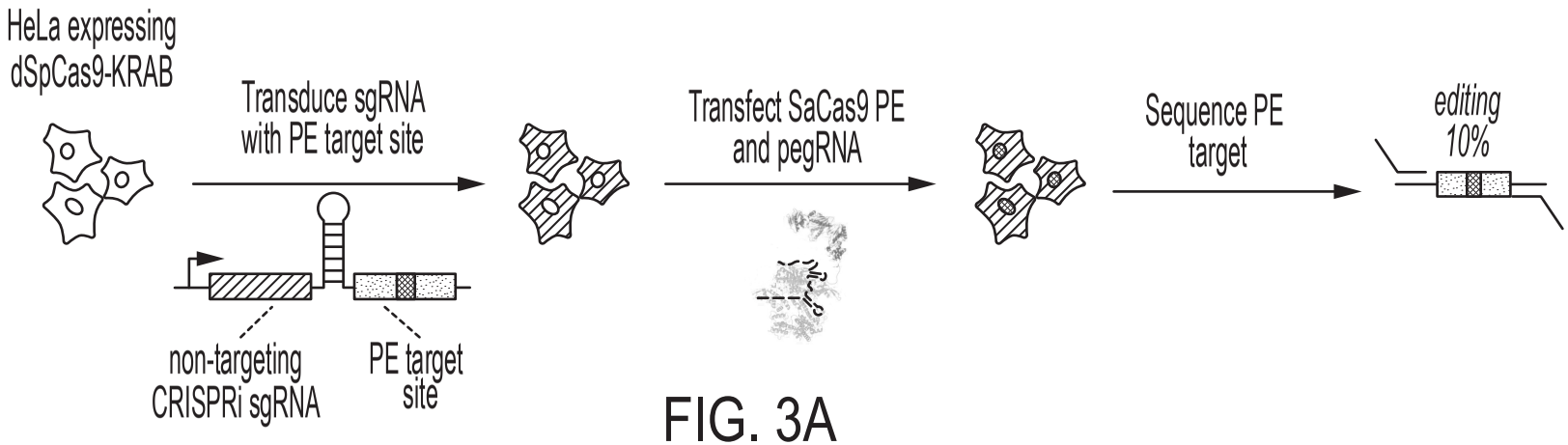

HeLa expressing
dSpCas9-KRAB
Transduce sgRNA
with PE target site
non-targeting
CRISPRi sgRNA
PE target
site
Transfect SaCas9 PE
and pegRNA
Sequence PE
target
editing
10%

FIG. 3A

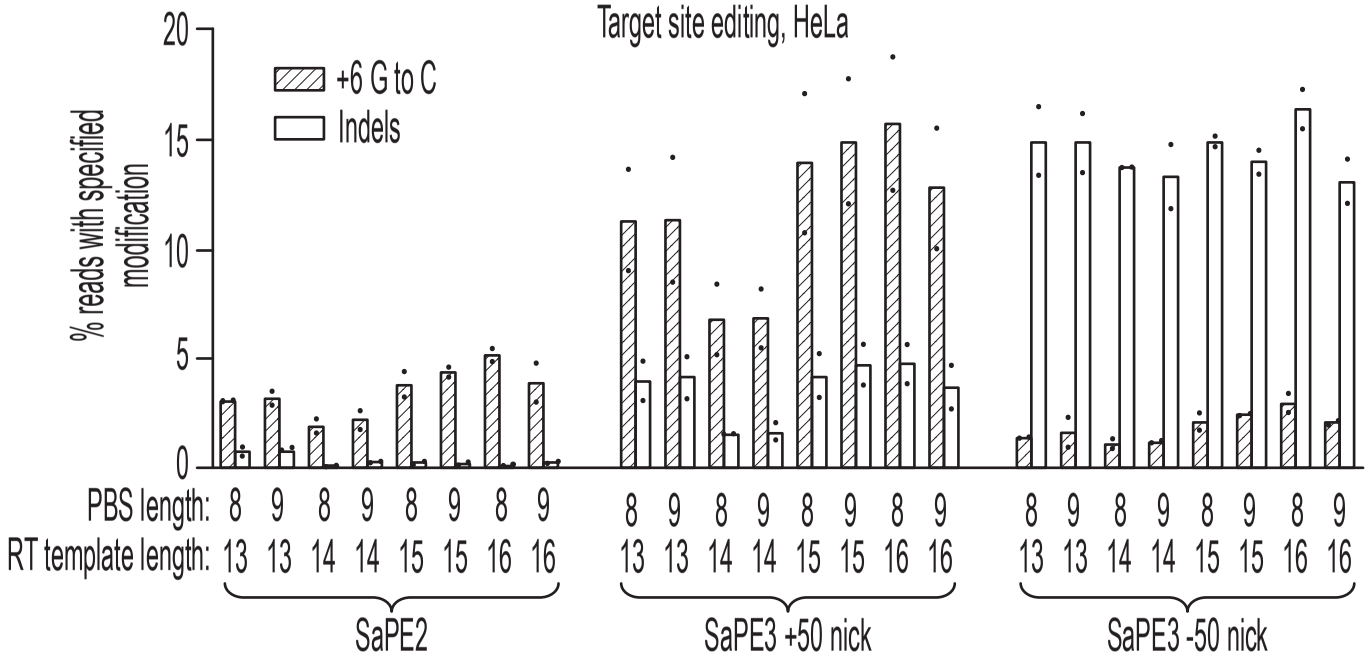

Target site editing, HeLa
+6 G to C
Indels
% reads with specified modification
20
15
10
5
0
PBS length: 8 9 8 9 8 9 8 9 8 9 8 9 8 9 8 9 8 9 8 9 8 9 8 9
RT template length: 13 13 14 14 15 15 16 16 13 13 14 14 15 15 16 16 13 13 14 14 15 15 16 16
SaPE2
SaPE3 +50 nick
SaPE3 -50 nick

FIG. 3B

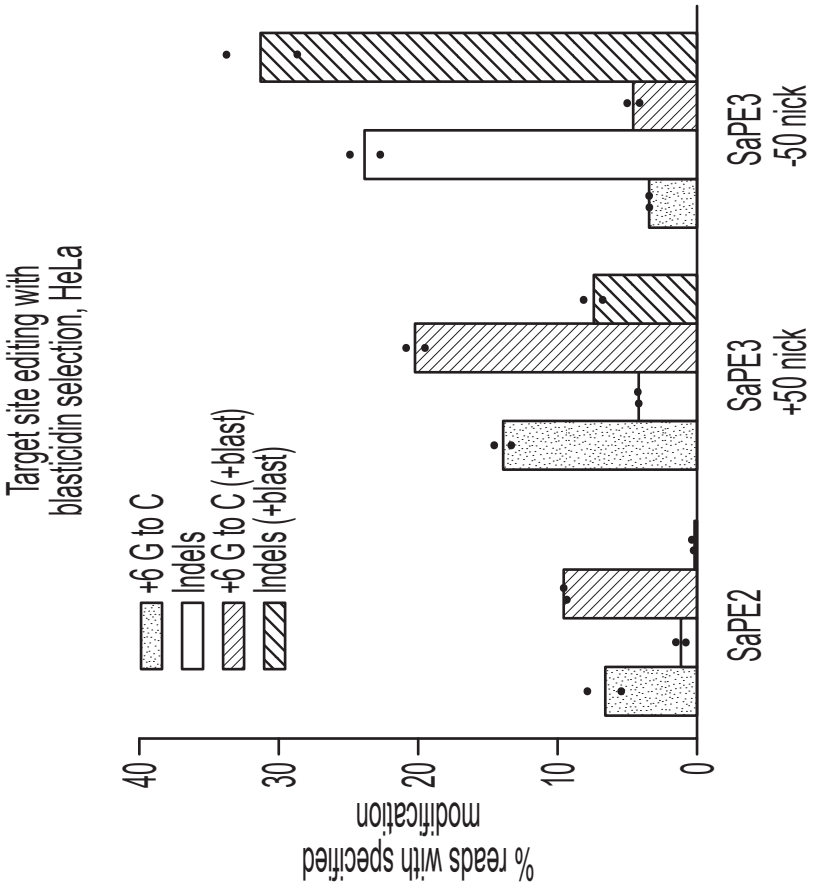

Target site editing with
blasticidin selection, HeLa
+6 G to C
Indels
+6 G to C (+blast)
Indels (+blast)
% reads with specified
modification
40
30
20
10
0
SaPE2
SaPE3
+50 nick
SaPE3
-50 nick

FIG. 3C

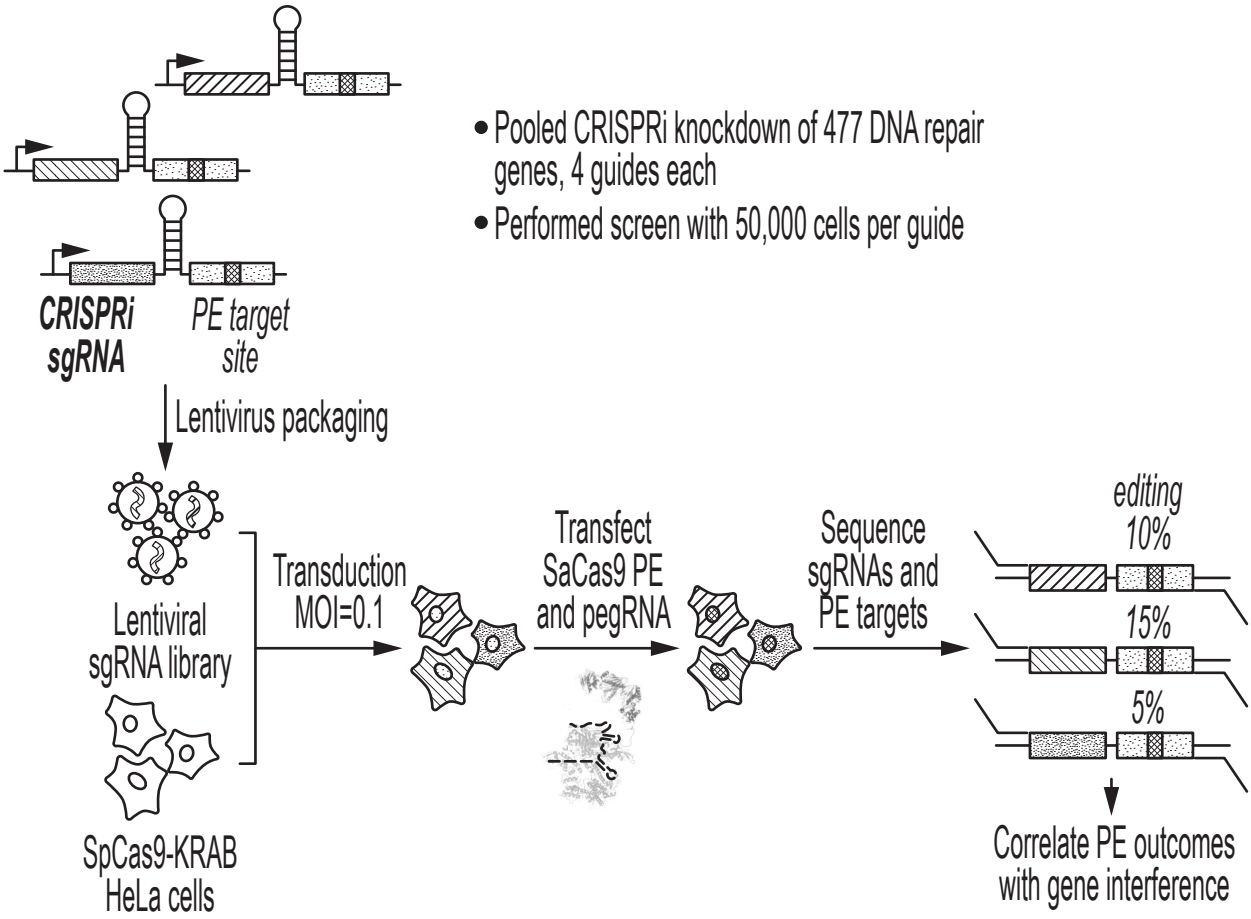

• Pooled CRISPRi knockdown of 477 DNA repair genes, 4 guides each
• Performed screen with 50,000 cells per guide
CRISPRi sgRNA
PE target site
Lentivirus packaging
Lentiviral sgRNA library
SpCas9-KRAB HeLa cells
Transduction MOI=0.1
Transfect SaCas9 PE and pegRNA
Sequence sgRNAs and PE targets
editing 10%
15%
5%
Correlate PE outcomes with gene interference

FIG. 4A

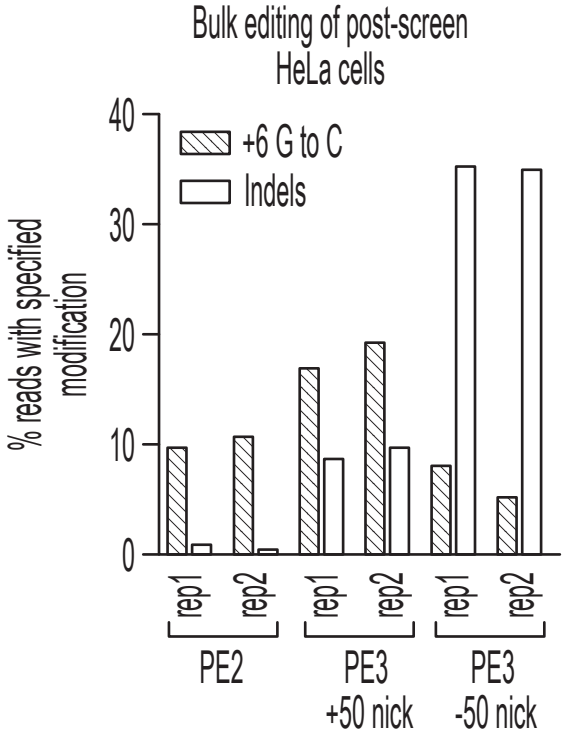

Bulk editing of post-screen HeLa cells
% reads with specified modification
40
30
20
10
0
+6 G to C
Indels
rep1
rep2
PE2
PE3 +50 nick
PE3 -50 nick

FIG. 4B

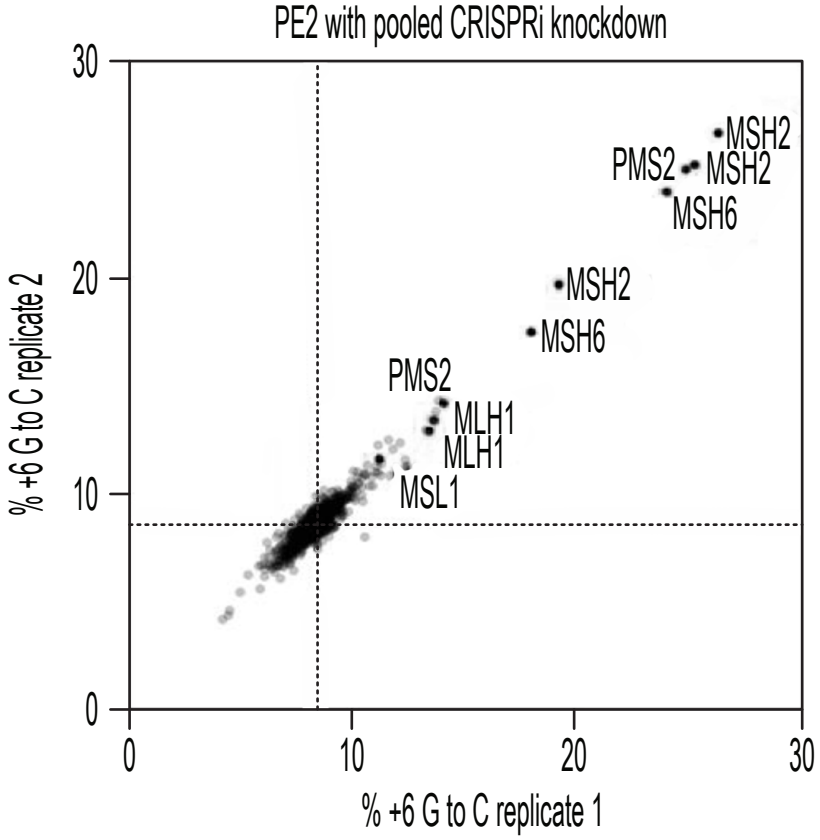

PE2 with pooled CRISPRi knockdown
% +6 G to C replicate 2
% +6 G to C replicate 1
0
10
20
30
MSH2
PMS2
MSH2
MSH6
MSH2
MSH6
PMS2
MLH1
MLH1
MSL1

FIG. 5A

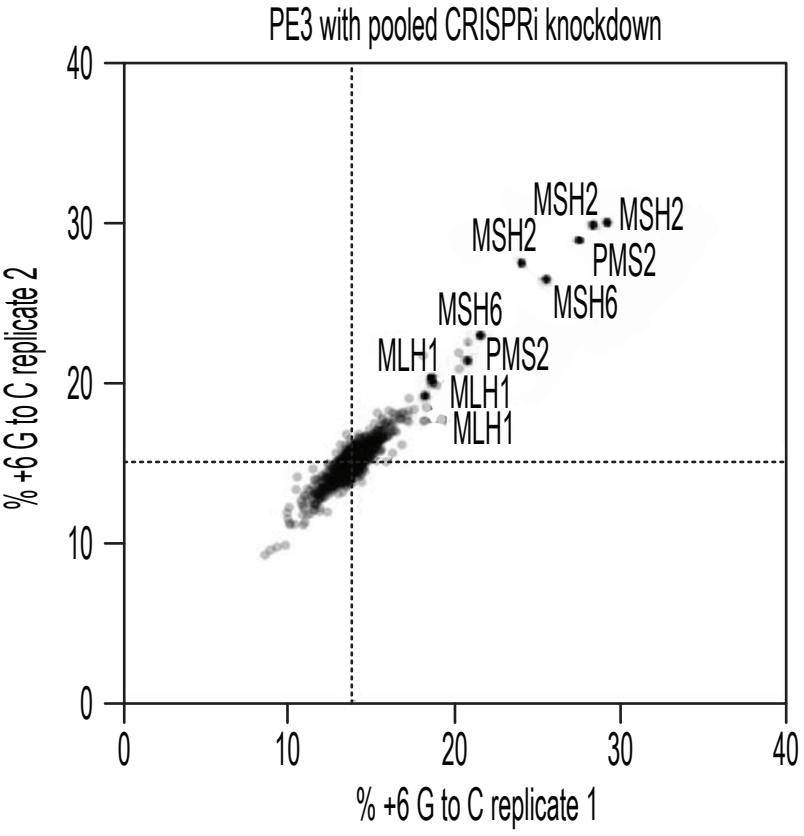

PE3 with pooled CRISPRi knockdown
% +6 G to C replicate 2
% +6 G to C replicate 1
0
10
20
30
40
MSH2
MSH2
MSH2
PMS2
MSH6
MSH6
MLH1
PMS2
MLH1
MLH1

FIG. 5B

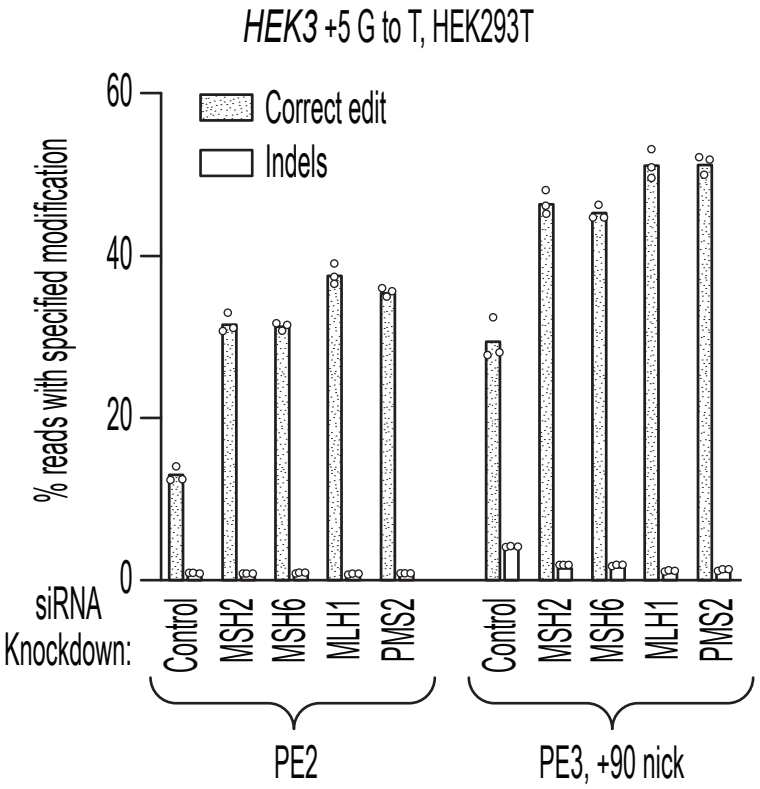

HEK3 +5 G to T, HEK293T
Correct edit
Indels
% reads with specified modification
60
40
20
0
siRNA Knockdown:
Control
MSH2
MSH6
MLH1
PMS2
Control
MSH2
MSH6
MLH1
PMS2
PE2
PE3, +90 nick

FIG. 6A

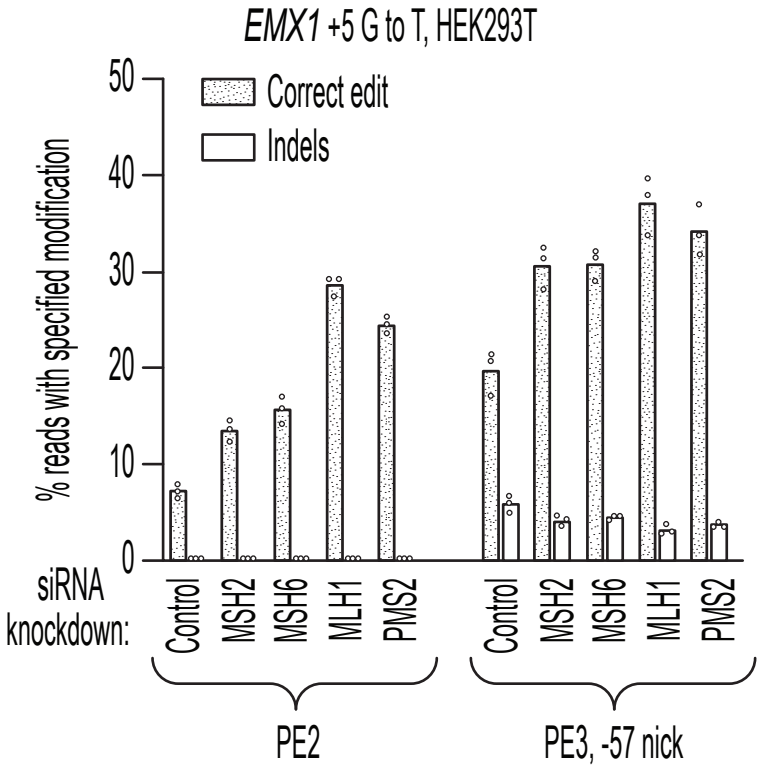

EMX1 +5 G to T, HEK293T
Correct edit
Indels
% reads with specified modification
50
40
30
20
10
0
siRNA knockdown:
Control
MSH2
MSH6
MLH1
PMS2
Control
MSH2
MSH6
MLH1
PMS2
PE2
PE3, -57 nick

FIG. 6B

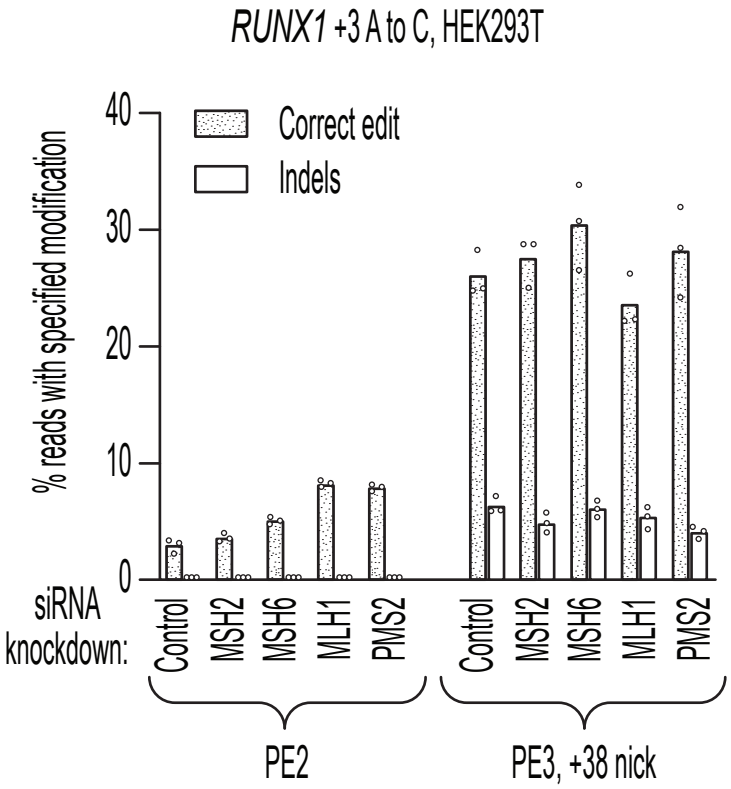

RUNX1 +3 A to C, HEK293T
Correct edit
Indels
% reads with specified modification
40
30
20
10
0
siRNA knockdown:
Control
MSH2
MSH6
MLH1
PMS2
Control
MSH2
MSH6
MLH1
PMS2
PE2
PE3, +38 nick

FIG. 6C

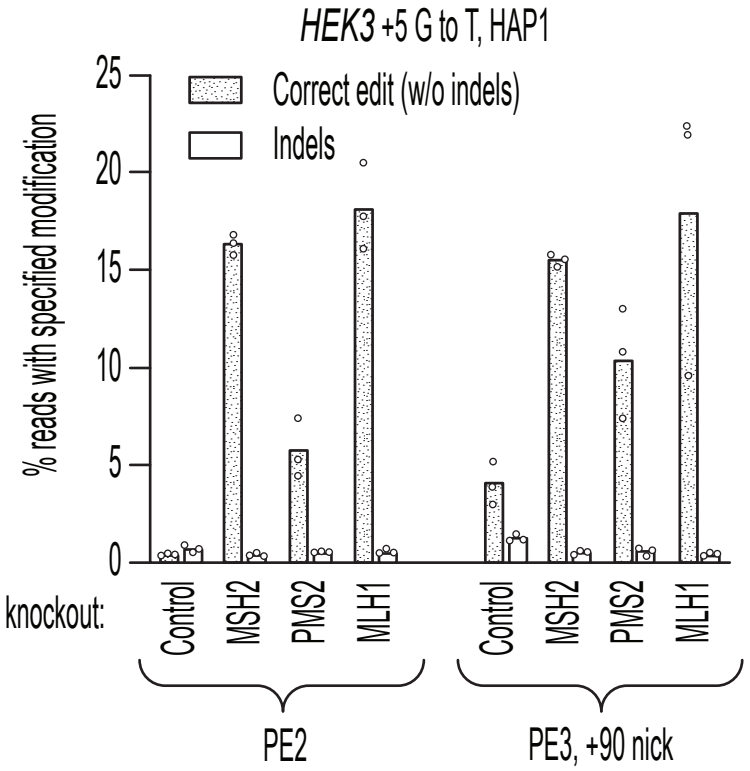

HEK3 +5 G to T, HAP1
Correct edit (w/o indels)
Indels
% reads with specified modification
25
20
15
10
5
0
knockout:
Control
MSH2
PMS2
MLH1
Control
MSH2
PMS2
MLH1
PE2
PE3, +90 nick

FIG. 7A

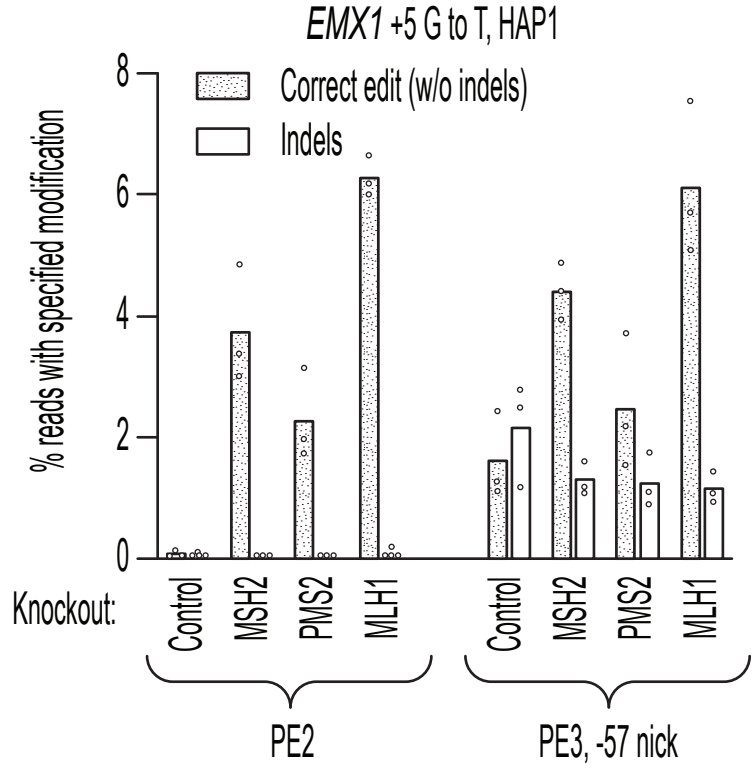

EMX1 +5 G to T, HAP1
Correct edit (w/o indels)
Indels
% reads with specified modification
8
6
4
2
0
Knockout:
Control
MSH2
PMS2
MLH1
Control
MSH2
PMS2
MLH1
PE2
PE3, -57 nick

FIG. 7B

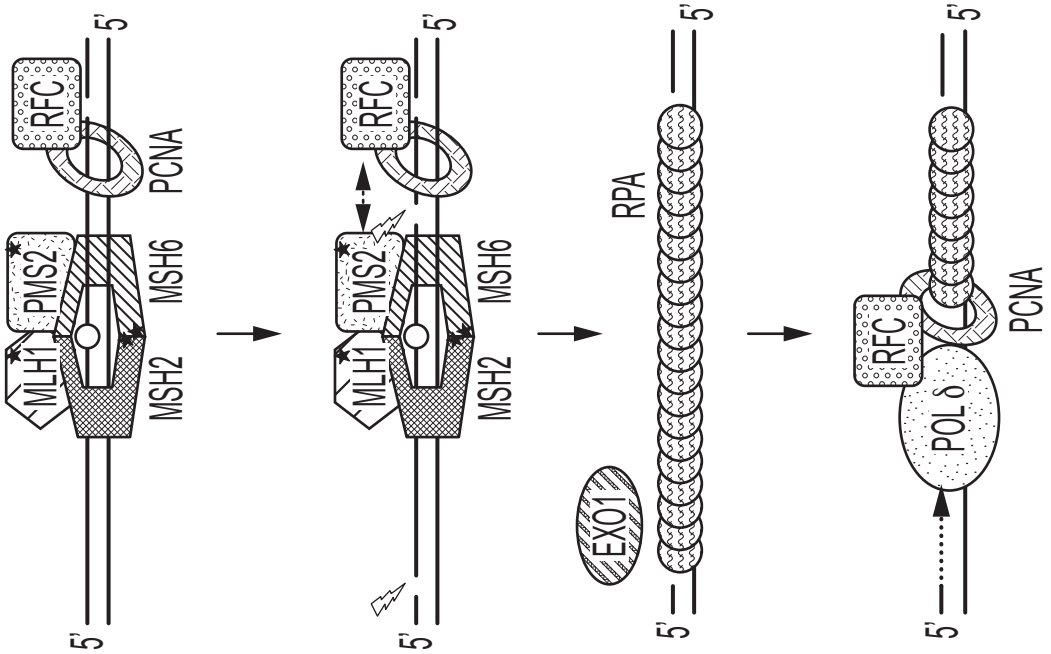

5'
RFC
PCNA
PMS2
MSH6
MLH1
MSH2
5'
RFC
PMS2
MSH6
MLH1
MSH2
5'
RPA
5'
EXO1
5'
RFC
POL δ
PCNA
5'

FIG. 8A

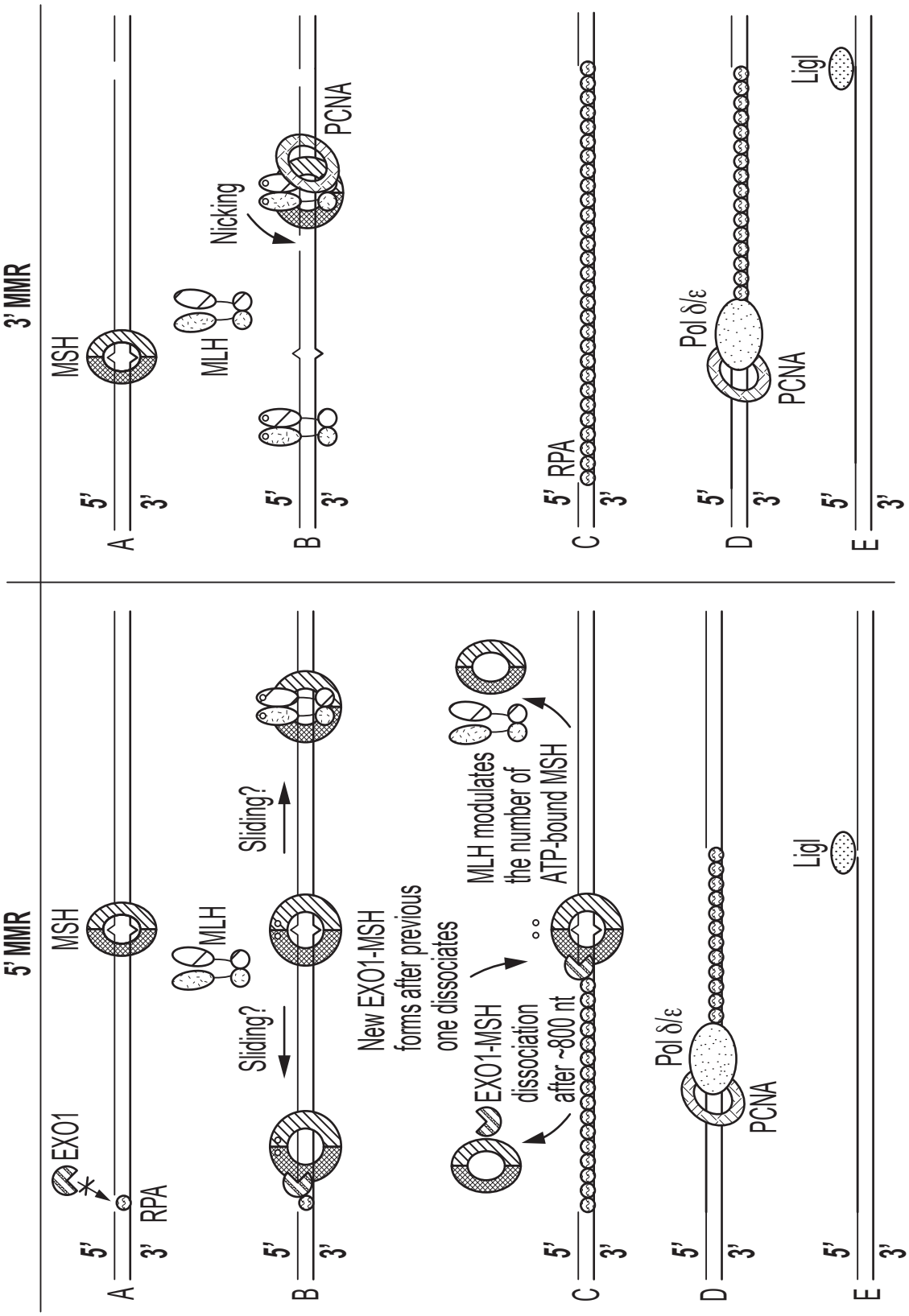

5' MMR
3' MMR
EXO1
MSH
MLH
RPA
Sliding?
Sliding?
New EXO1-MSH
forms after previous
one dissociates
MLH modulates
the number of
ATP-bound MSH
EXO1-MSH
dissociation
after ~800 nt
Nicking
PCNA
Pol δ/ε
PCNA
LigI
RPA
5'
3'
A
B
C
D
E

FIG. 8B

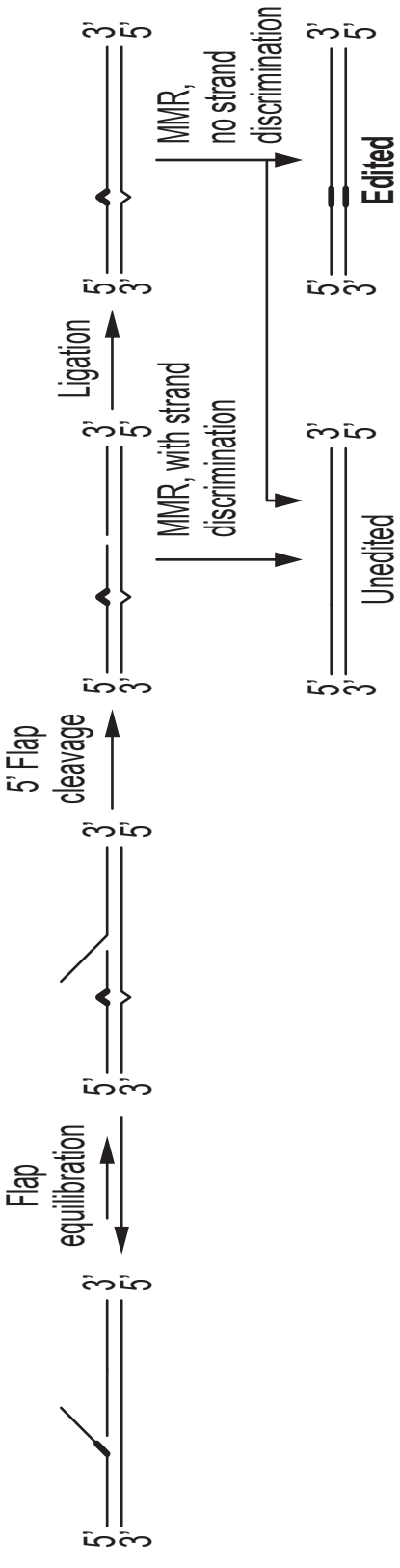

Flap equilibration
5' Flap cleavage
Ligation
MMR, with strand discrimination
MMR, no strand discrimination
Unedited
Edited
5' 3'
3' 5'

FIG. 9A

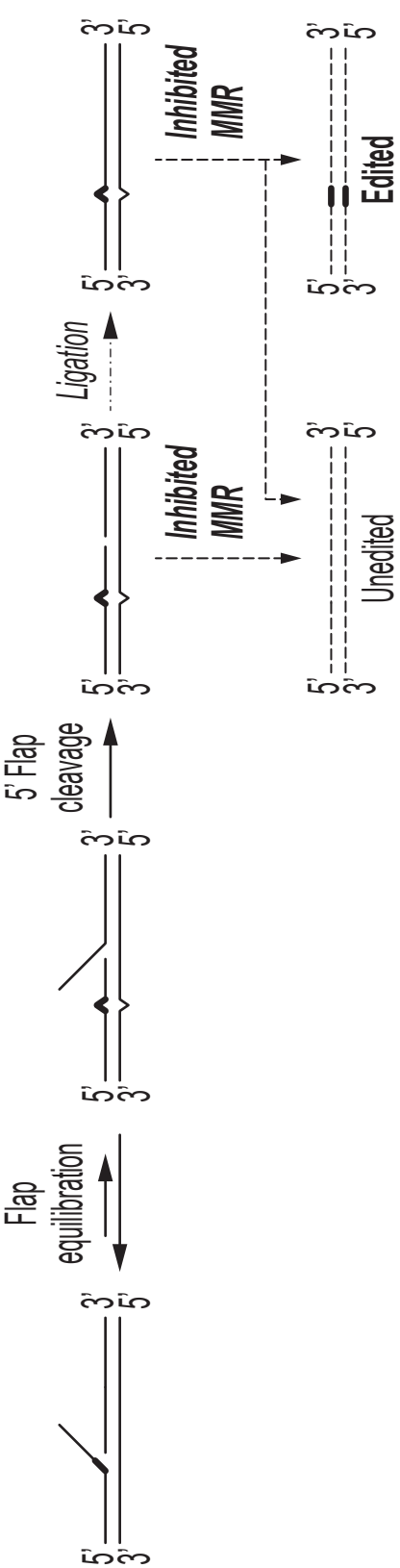

Flap equilibration
5' Flap cleavage
Ligation
Inhibited MMR
Inhibited MMR
Unedited
Edited
5' 3'
3' 5'

FIG. 9B

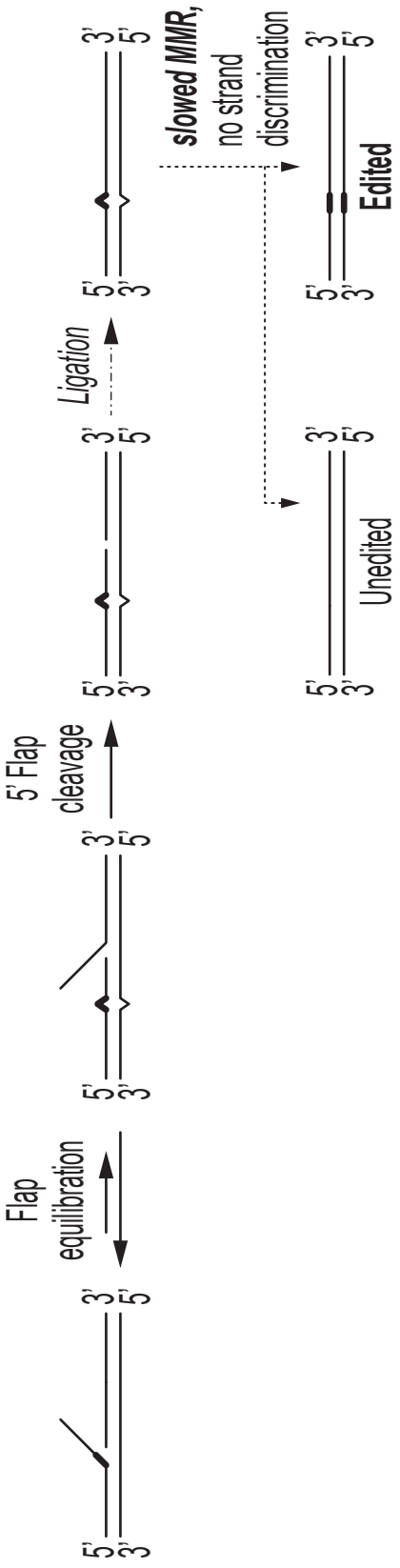

Flap equilibration
5' Flap cleavage
Ligation
slowed MMR, no strand discrimination
Edited
Unedited
5'
3'

FIG. 9C

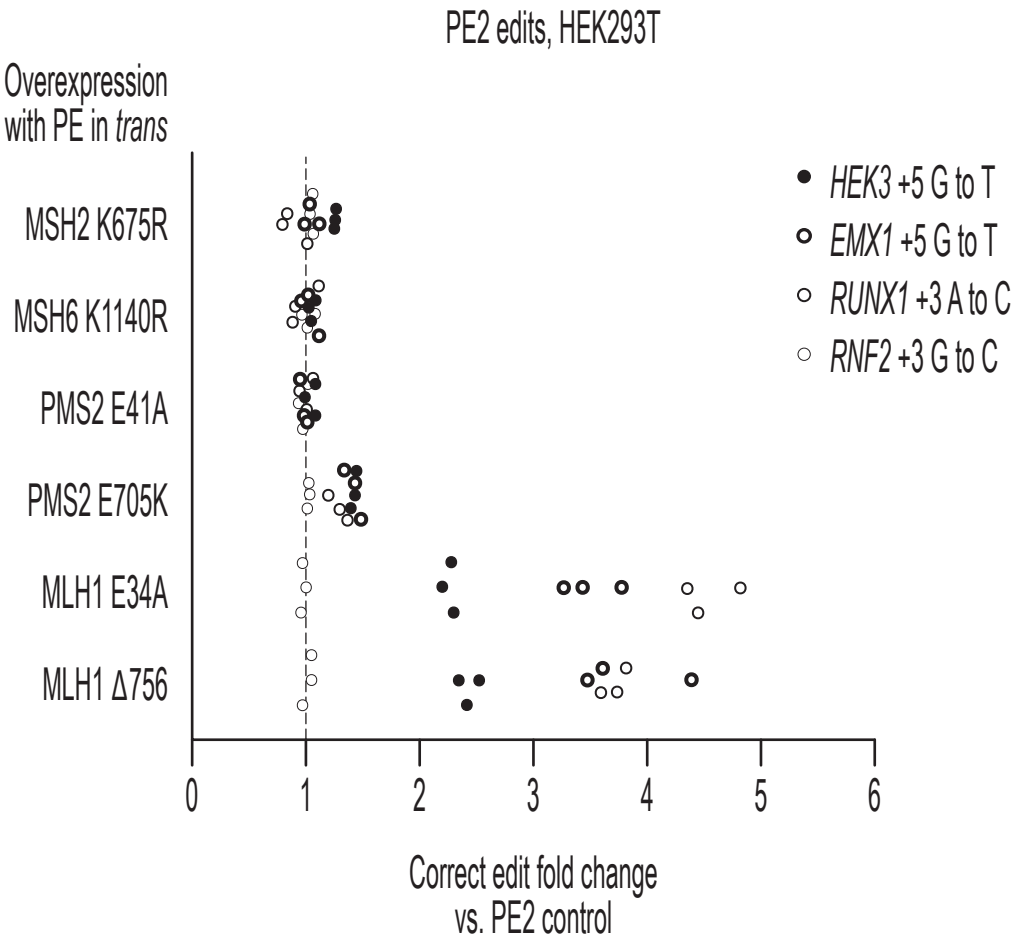

PE2 edits, HEK293T
Overexpression with PE in *trans*
MSH2 K675R
MSH6 K1140R
PMS2 E41A
PMS2 E705K
MLH1 E34A
MLH1 Δ756
HEK3 +5 G to T
EMX1 +5 G to T
RUNX1 +3 A to C
RNF2 +3 G to C
0
1
2
3
4
5
6
Correct edit fold change vs. PE2 control

FIG. 10

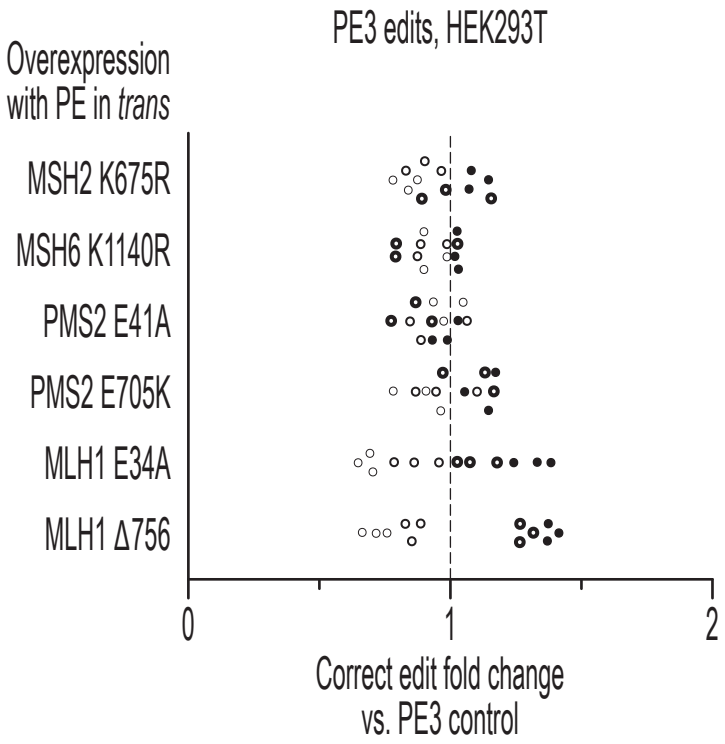


FIG. 11A

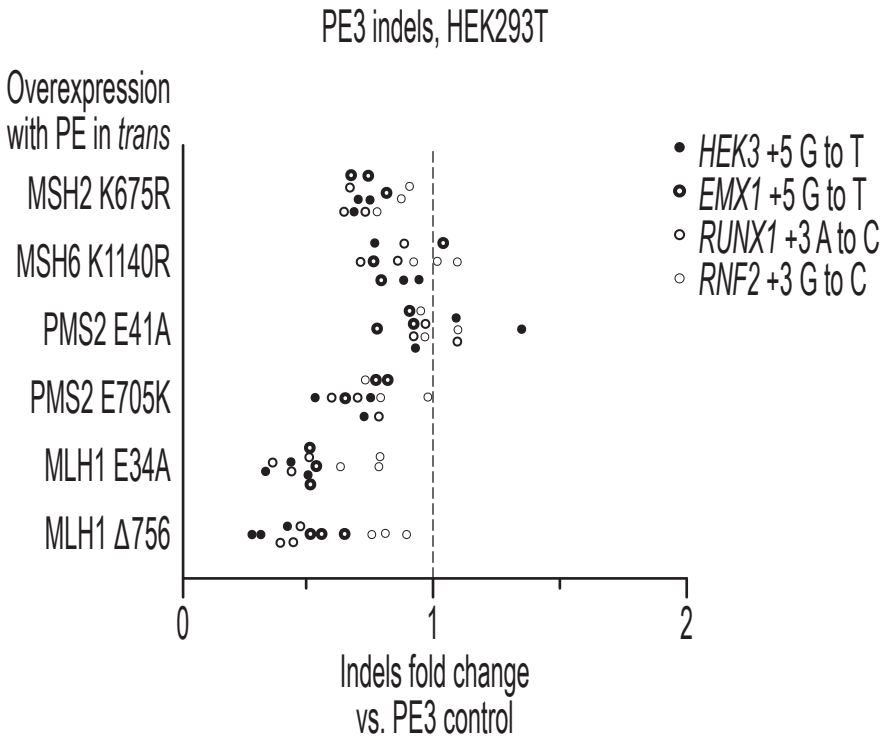


FIG. 11B

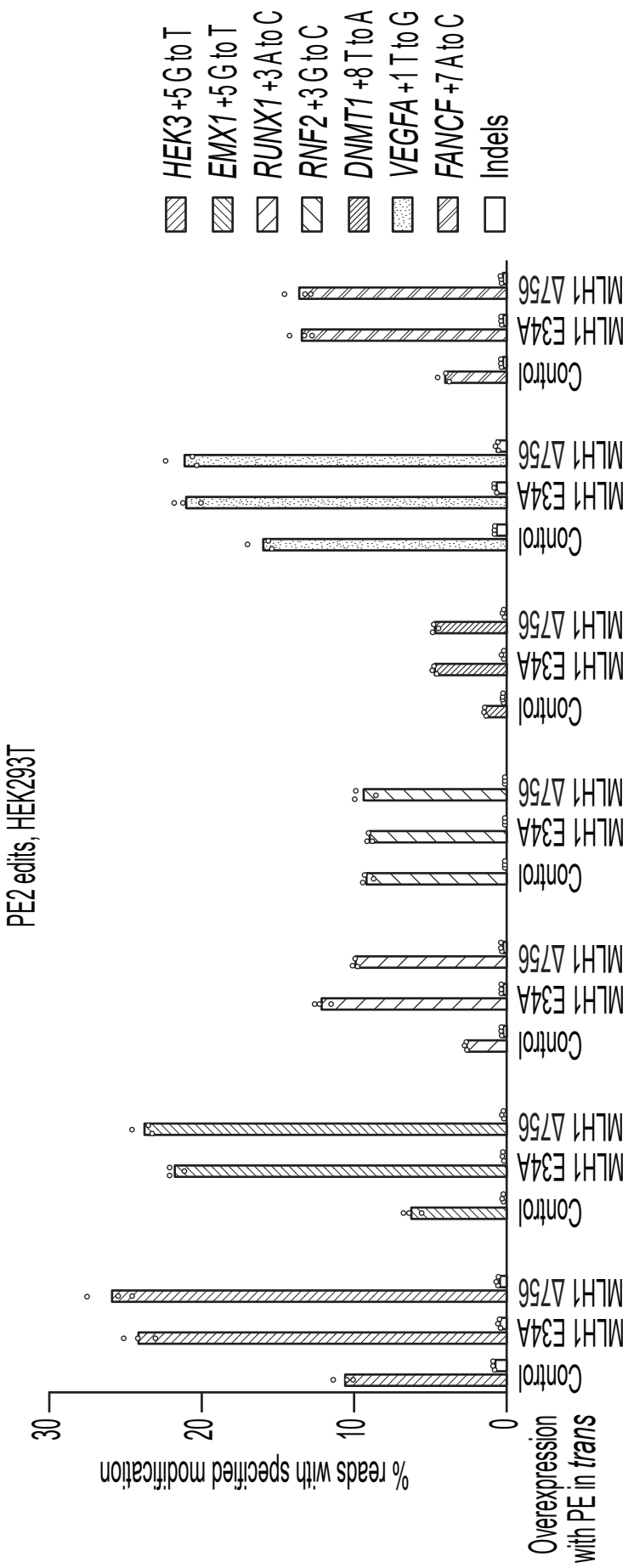

HEK3 +5 G to T
EMX1 +5 G to T
RUNX1 +3 A to C
RNF2 +3 G to C
DNMT1 +8 T to A
VEGFA +1 T to G
FANCF +7 A to C
Indels
PE2 edits, HEK293T
% reads with specified modification
0
10
20
30
Overexpression with PE in trans
Control
MLH1 E34A
MLH1 Δ756

FIG. 12A

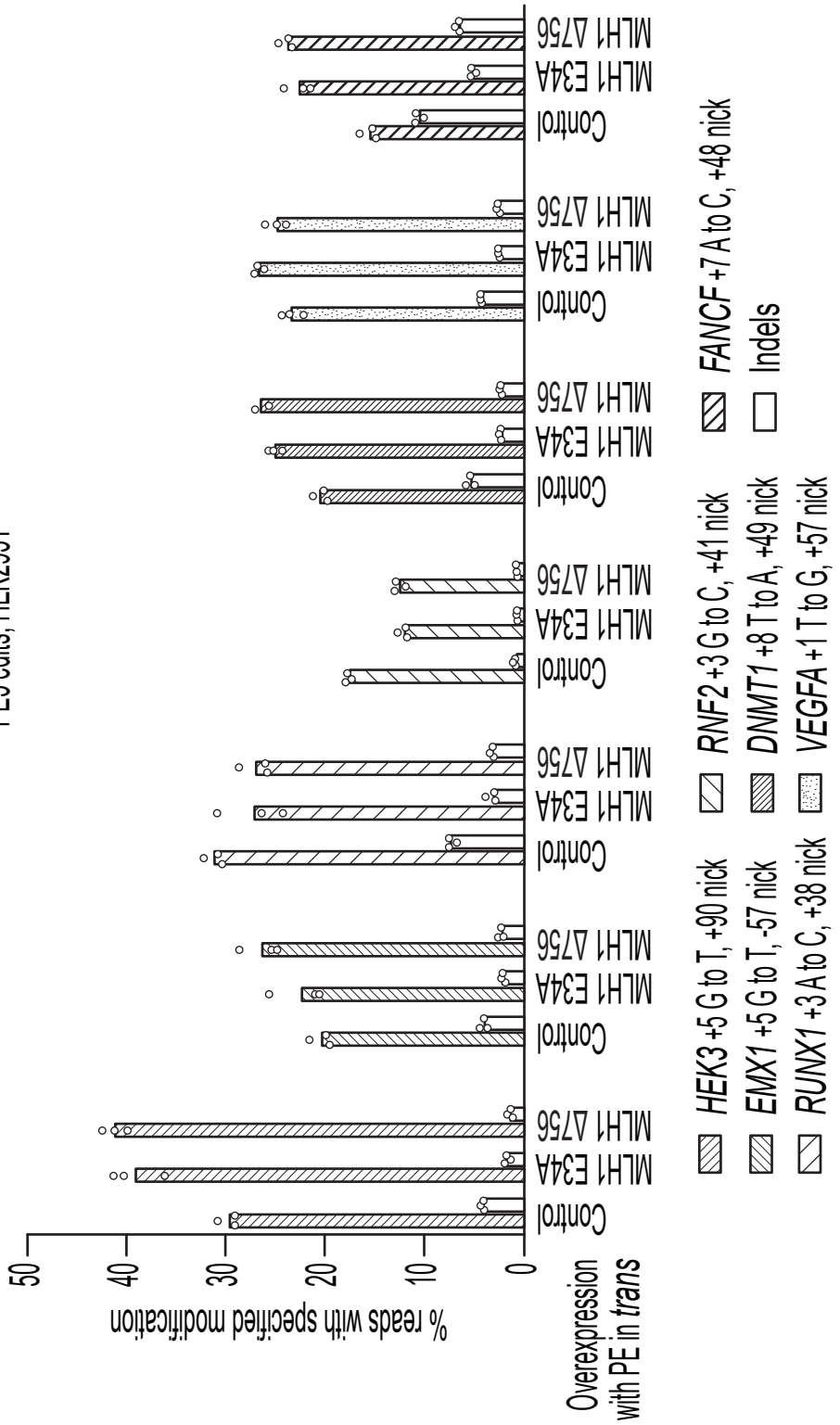

PE3 edits, HEK293T
% reads with specified modification
50
40
30
20
10
0
Overexpression with PE in trans
Control
MLH1 E34A
MLH1 Δ756
HEK3 +5 G to T, +90 nick
EMX1 +5 G to T, -57 nick
RUNX1 +3 A to C, +38 nick
RNF2 +3 G to C, +41 nick
DNMT1 +8 T to A, +49 nick
VEGFA +1 T to G, +57 nick
FANCF +7 A to C, +48 nick
Indels

FIG. 12B

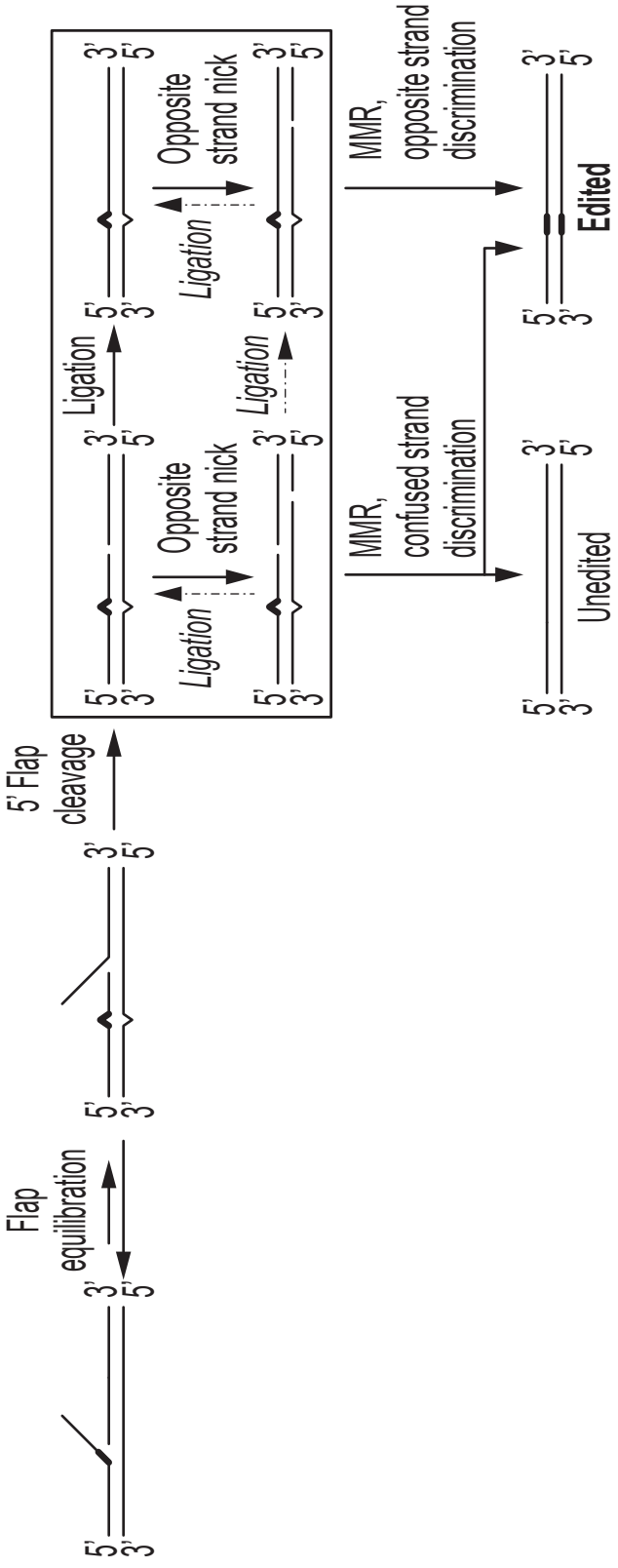

Flap equilibration
5' Flap cleavage
Ligation
Opposite strand nick
Ligation
Ligation
Opposite strand nick
Ligation
MMR, confused strand discrimination
MMR, opposite strand discrimination
Unedited
Edited
5'
3'

FIG. 13

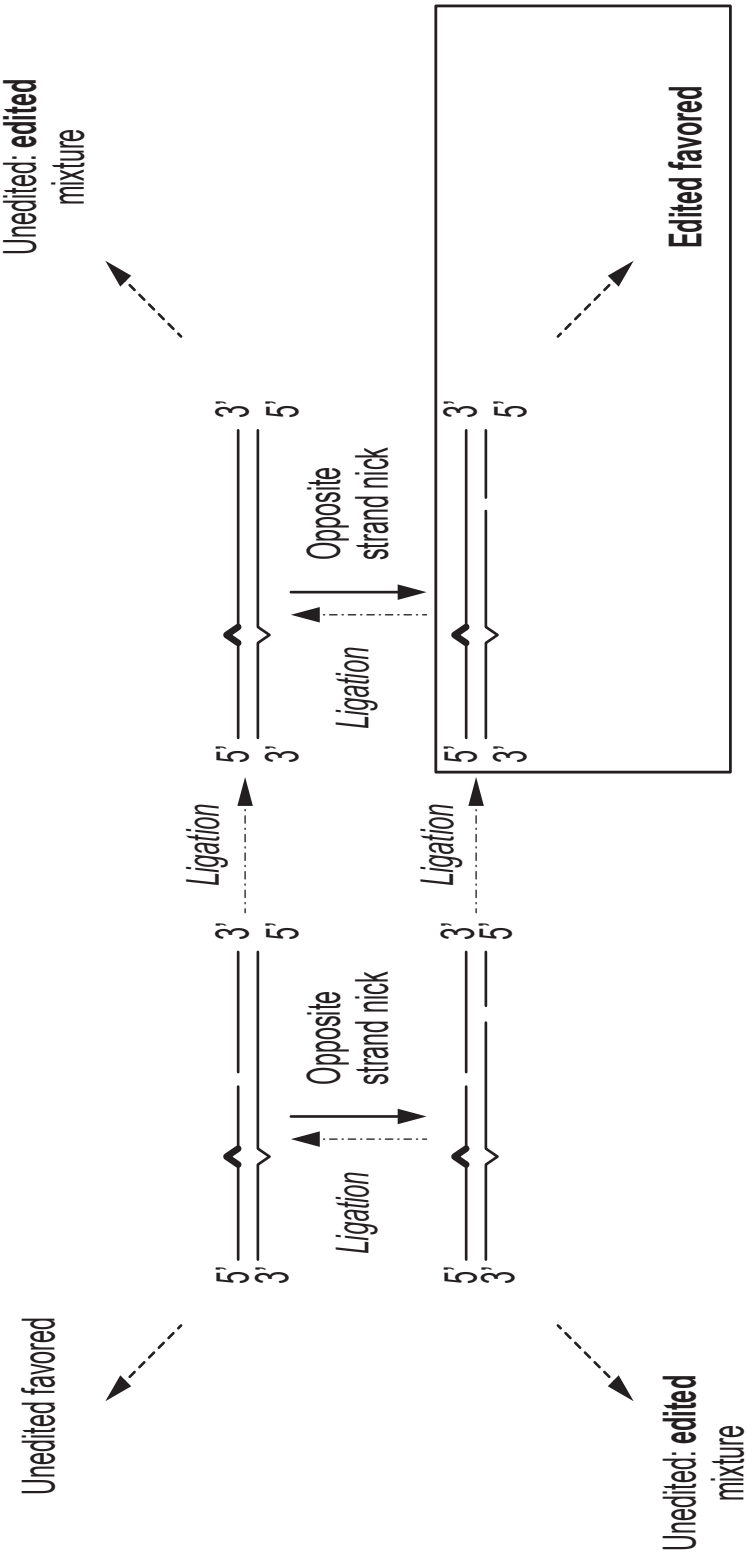

Unedited: edited mixture
Edited favored
Unedited favored
Unedited: edited mixture
Opposite strand nick
Ligation
Opposite strand nick
Ligation
Ligation
Ligation
3'
5'
5'
3'

FIG. 14

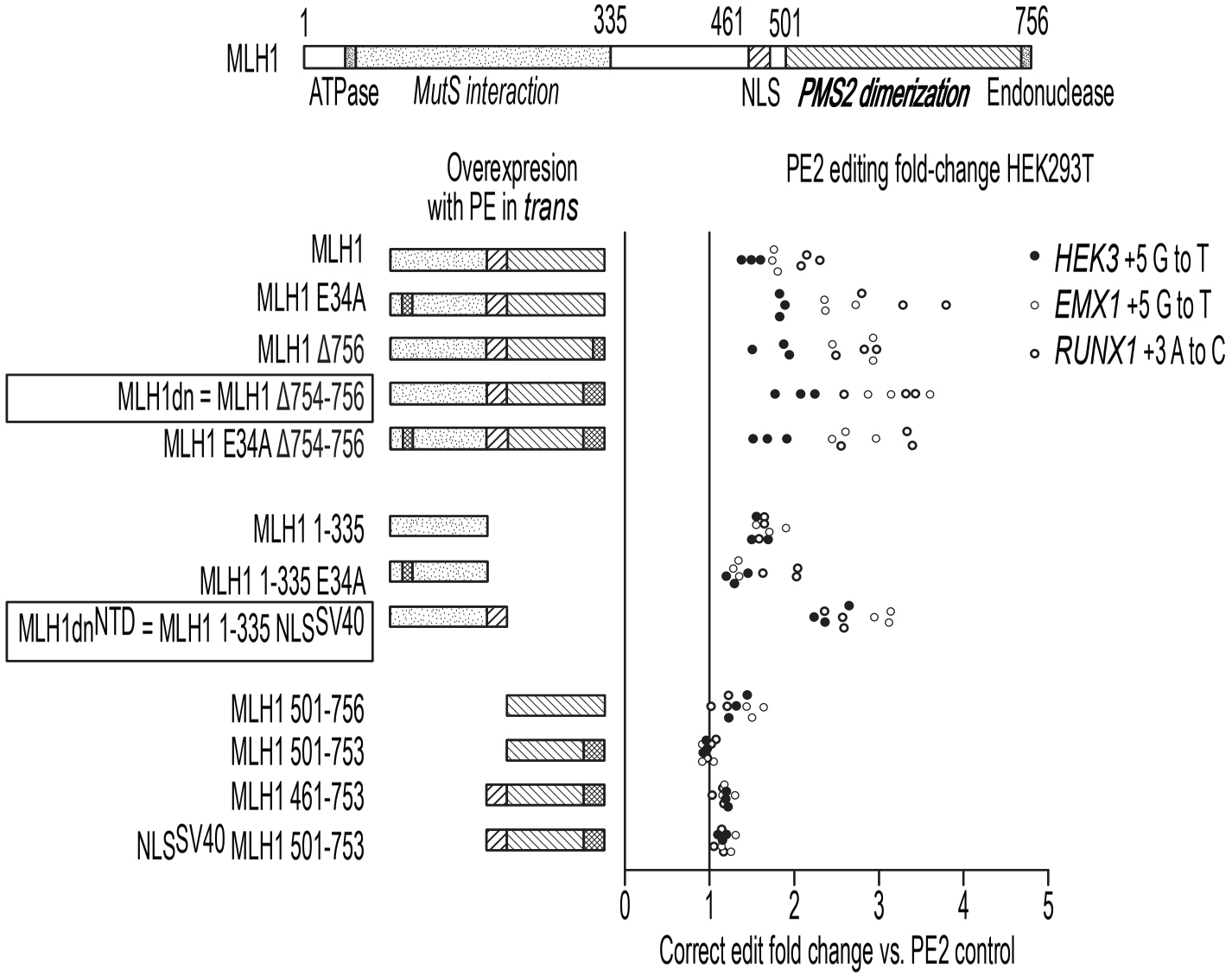

1
335
461
501
756
MLH1
ATPase
MutS interaction
NLS
PMS2 dimerization
Endonuclease
Overexpresion
with PE in trans
PE2 editing fold-change HEK293T
MLH1
MLH1 E34A
MLH1 Δ756
MLH1dn = MLH1 Δ754-756
MLH1 E34A Δ754-756
MLH1 1-335
MLH1 1-335 E34A
MLH1dnNTD = MLH1 1-335 NLSSV40
MLH1 501-756
MLH1 501-753
MLH1 461-753
NLSSV40 MLH1 501-753
HEK3 +5 G to T
EMX1 +5 G to T
RUNX1 +3 A to C
0
1
2
3
4
5
Correct edit fold change vs. PE2 control

FIG. 15A

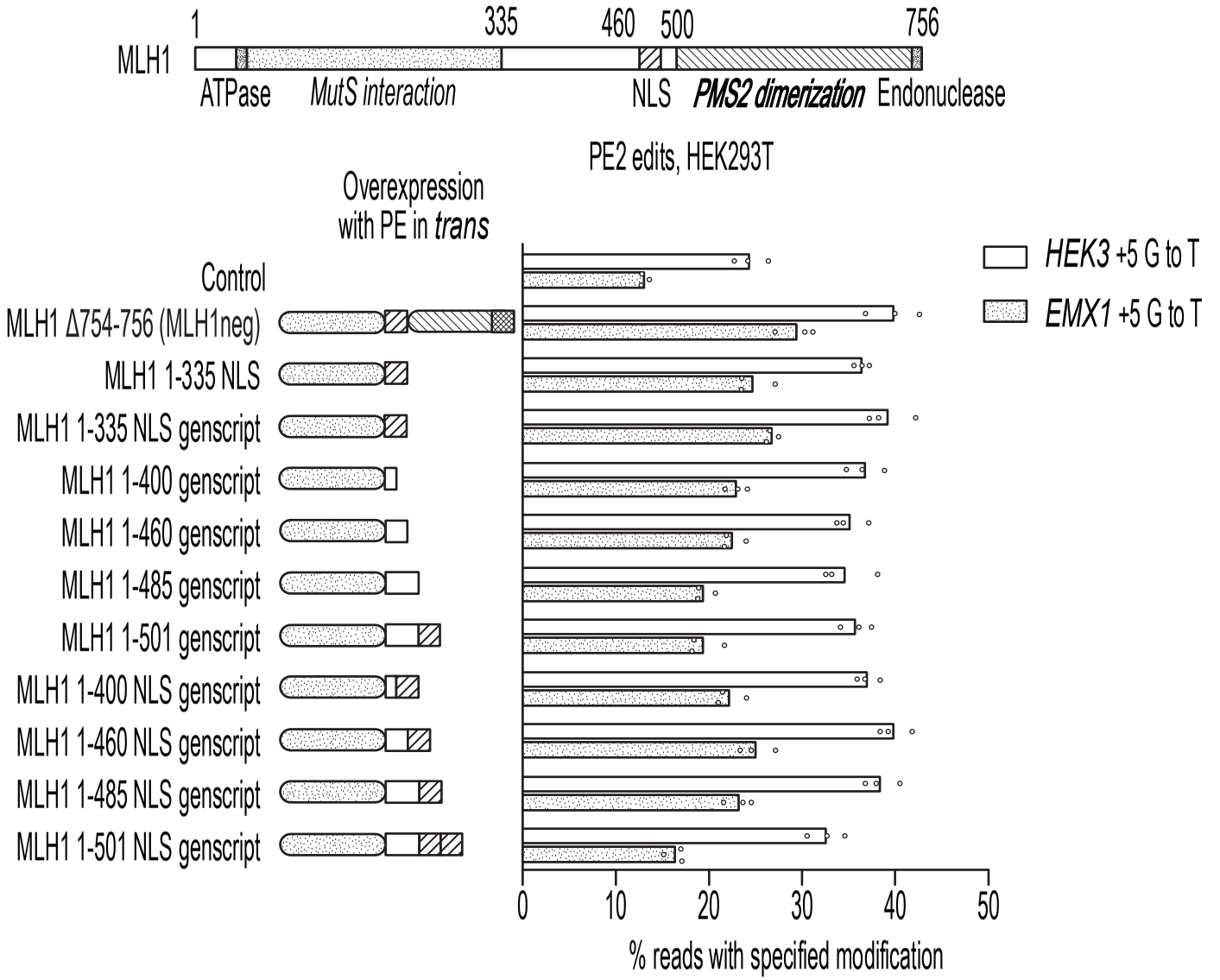

1
335
460
500
756
MLH1
ATPase
MutS interaction
NLS
PMS2 dimerization
Endonuclease
PE2 edits, HEK293T
Overexpression with PE in trans
Control
MLH1 Δ754-756 (MLH1neg)
MLH1 1-335 NLS
MLH1 1-335 NLS genscript
MLH1 1-400 genscript
MLH1 1-460 genscript
MLH1 1-485 genscript
MLH1 1-501 genscript
MLH1 1-400 NLS genscript
MLH1 1-460 NLS genscript
MLH1 1-485 NLS genscript
MLH1 1-501 NLS genscript
HEK3 +5 G to T
EMX1 +5 G to T
0
10
20
30
40
50
% reads with specified modification

FIG. 15B

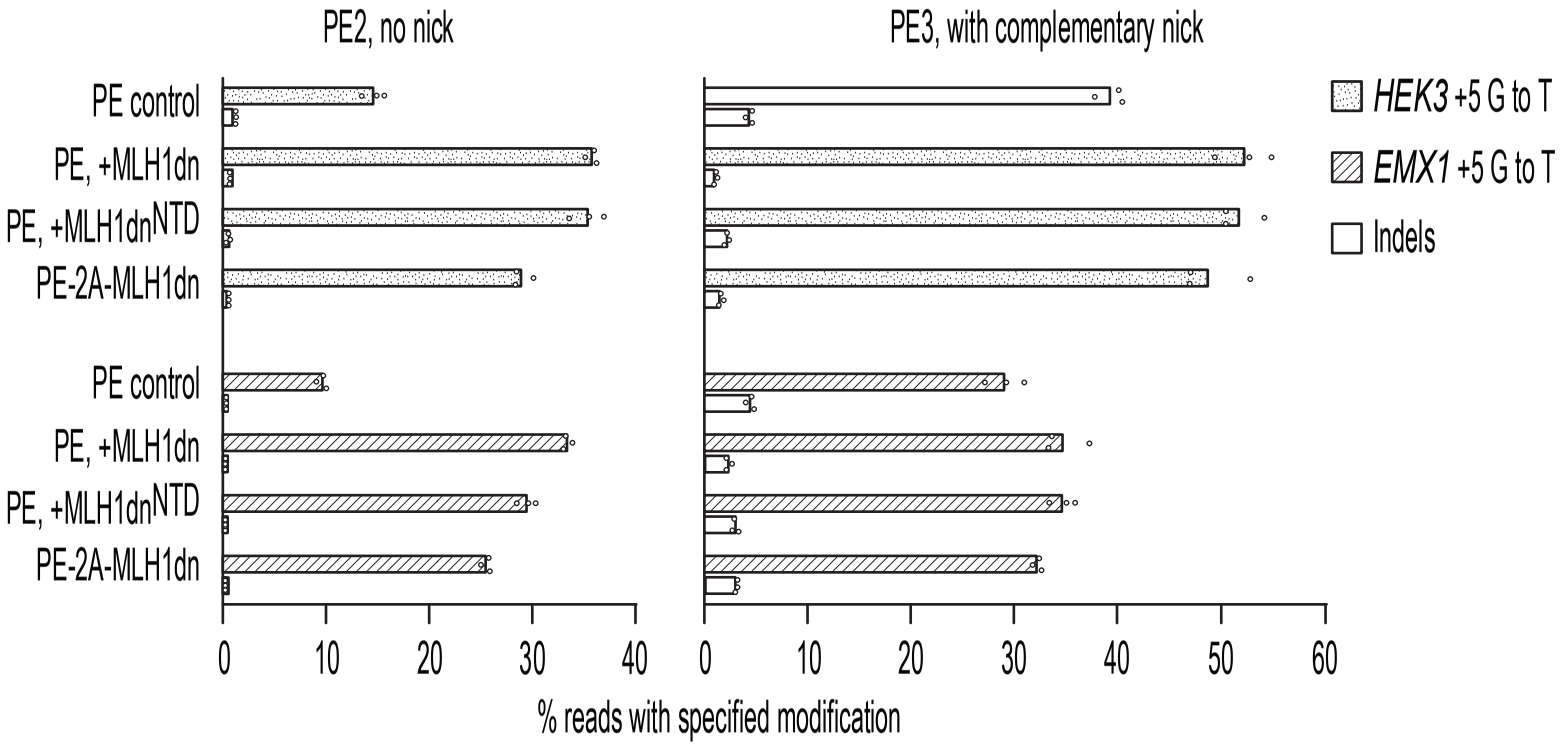

PE2, no nick
PE3, with complementary nick
PE control
PE, +MLH1dn
PE, +MLH1dnNTD
PE-2A-MLH1dn
PE control
PE, +MLH1dn
PE, +MLH1dnNTD
PE-2A-MLH1dn
HEK3 +5 G to T
EMX1 +5 G to T
Indels
0
10
20
30
40
0
10
20
30
40
50
60
% reads with specified modification

FIG. 15C

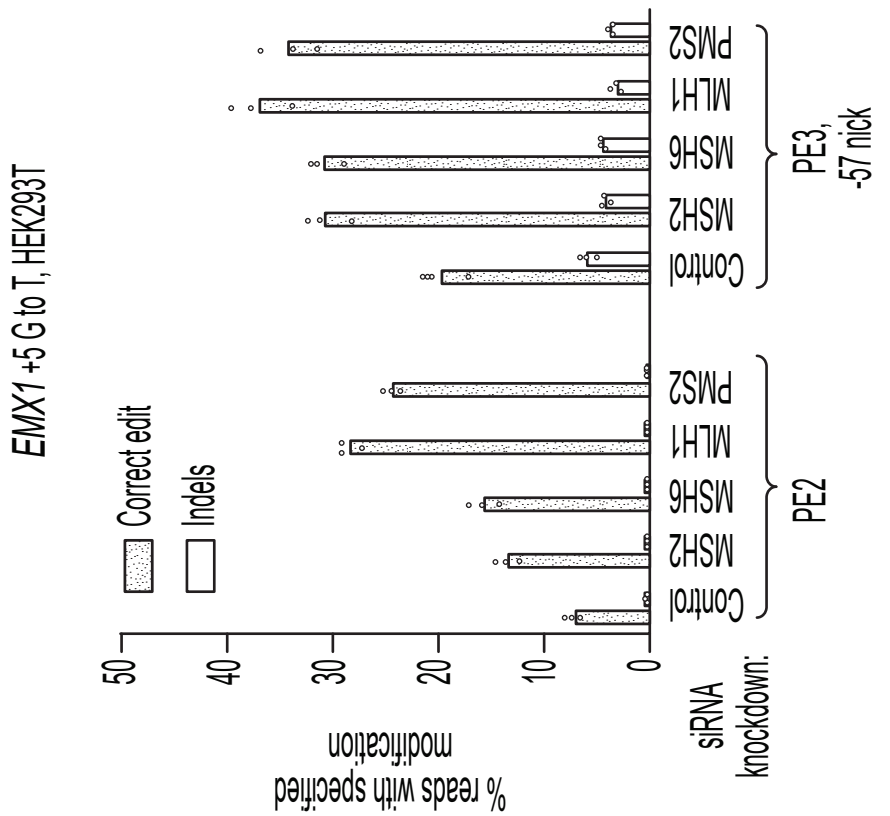


FIG. 15E

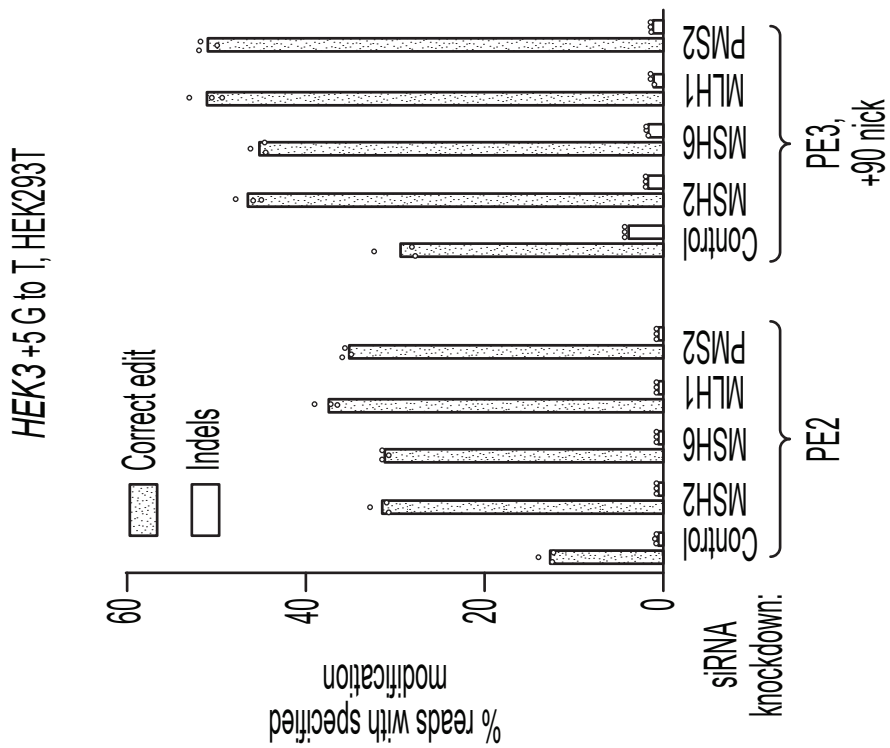


FIG. 15D

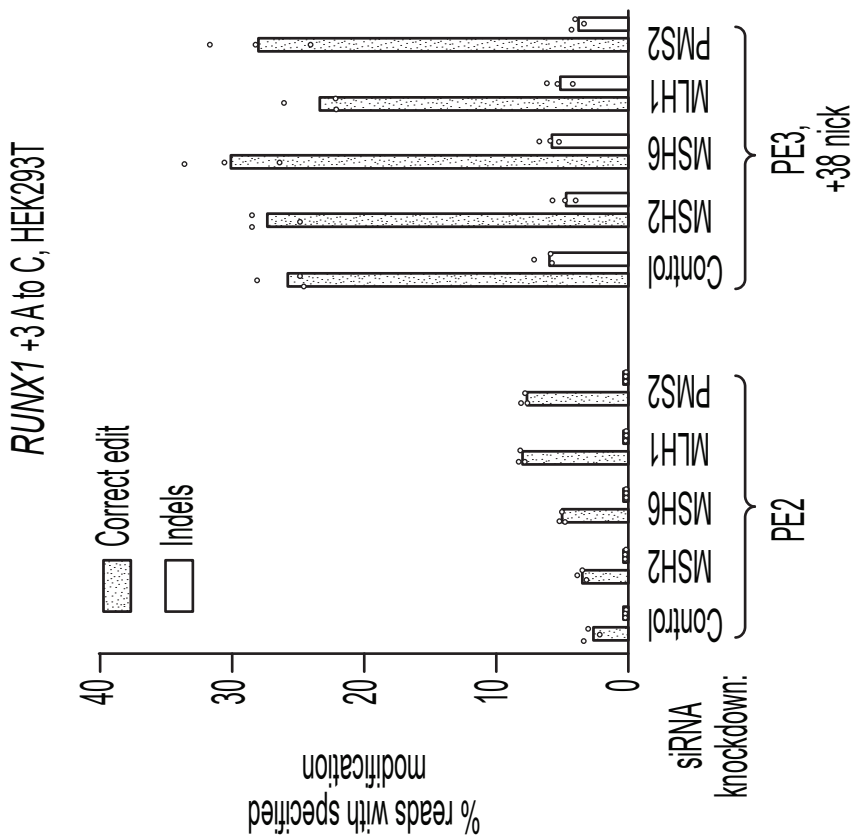

RUNX1 +3 A to C, HEK293T
Correct edit
Indels
% reads with specified modification
40
30
20
10
0
siRNA knockdown:
Control
MSH2
MSH6
MLH1
PMS2
PE2
Control
MSH2
MSH6
MLH1
PMS2
PE3, +38 nick

FIG. 15F

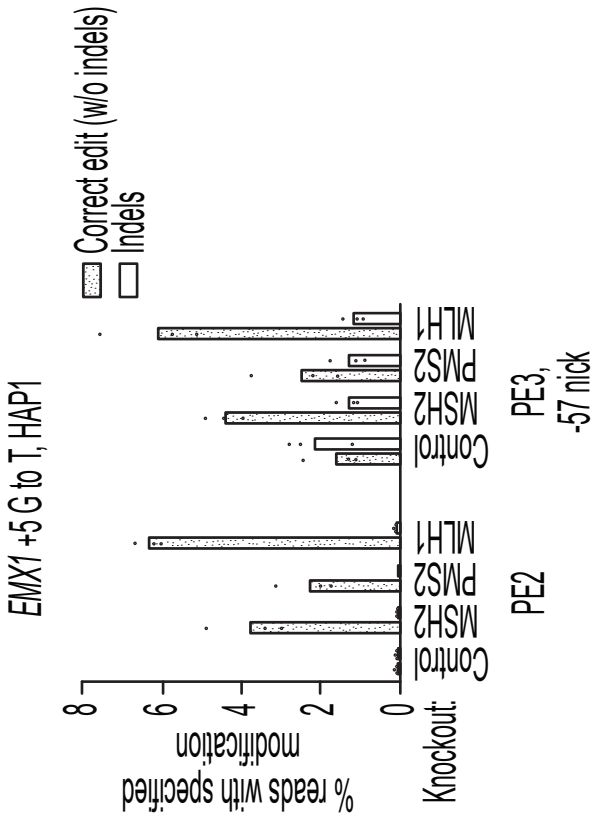


FIG. 15H

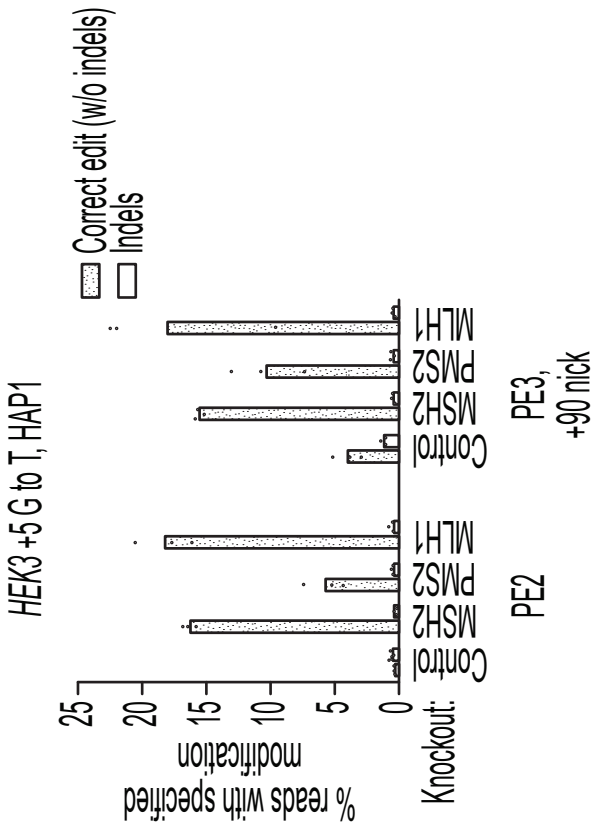


FIG. 15G

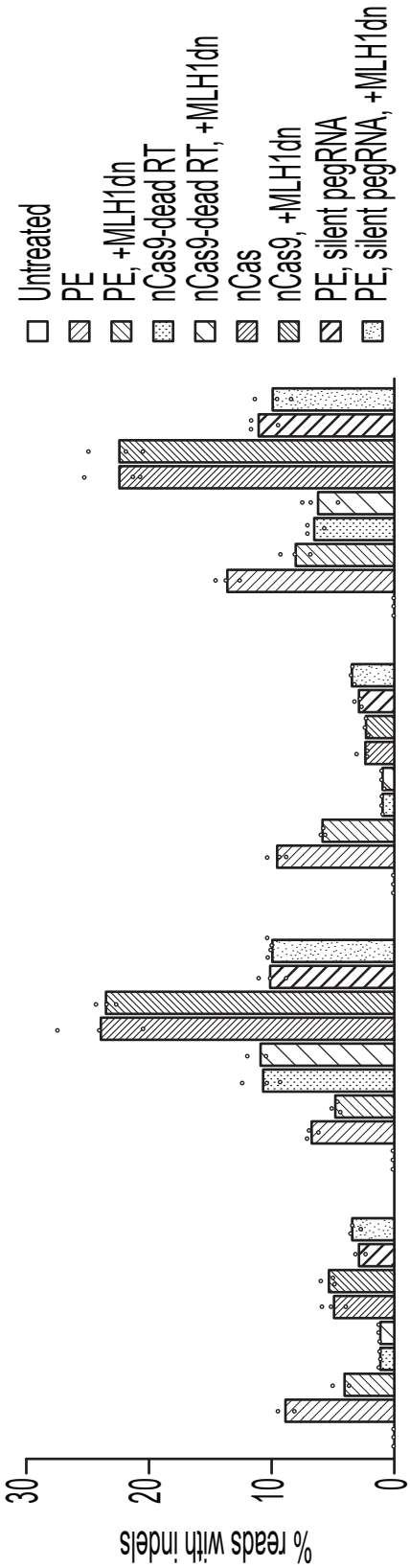

PE3 indels, HEK293T
% reads with indels
30
20
10
0
Untreated
PE
PE, +MLH1dn
nCas9-dead RT
nCas9-dead RT, +MLH1dn
nCas
nCas9, +MLH1dn
PE, silent pegRNA
PE, silent pegRNA, +MLH1dn
HEK3 +1 T to A
+90 nick
EMK1 +5 G to T
-57 nick
DNMT1 +1 A to C
+49 nick
FANCF +3 C to T
+48 nick

FIG. 16

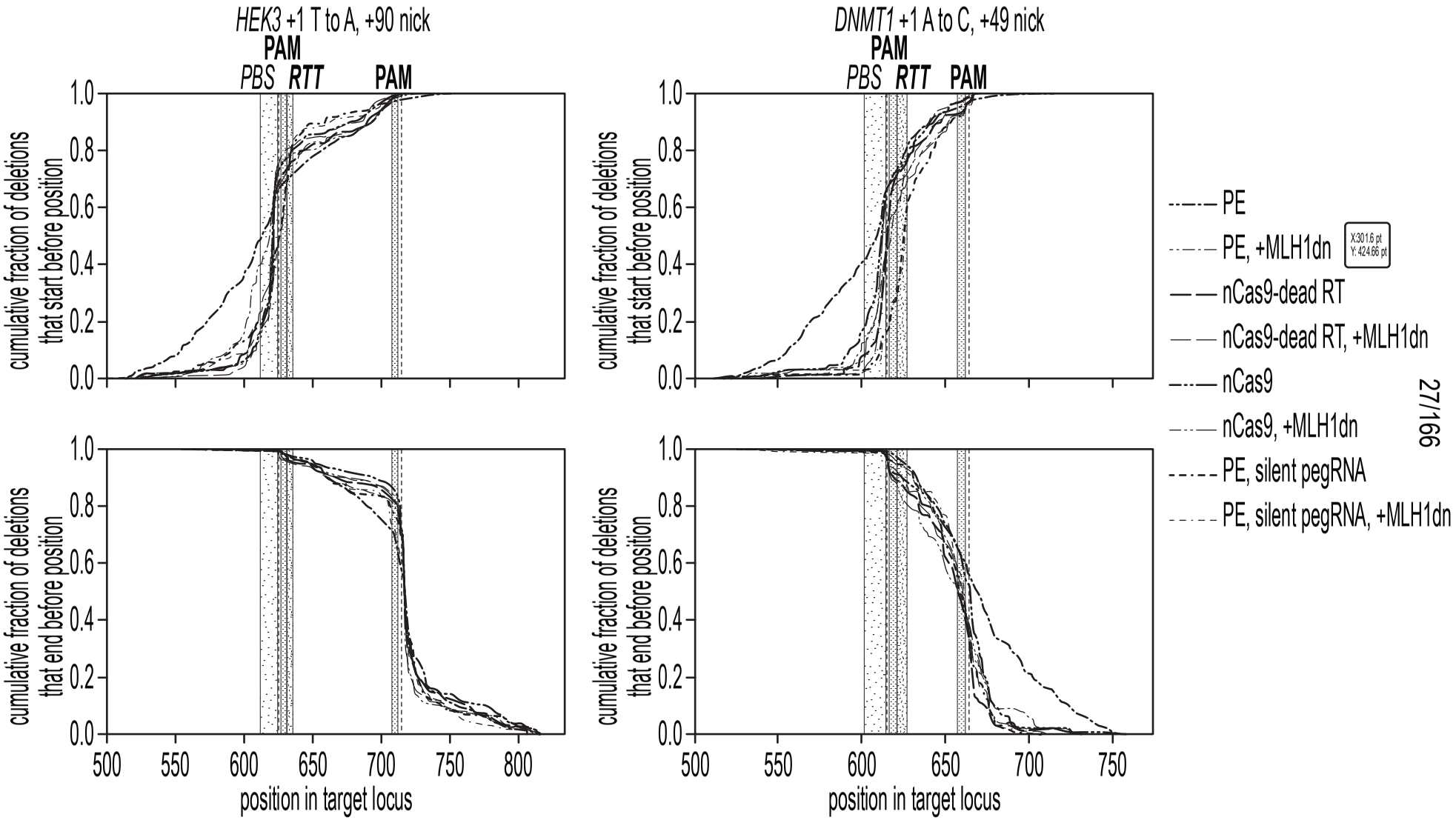

HEK3 +1 T to A, +90 nick
PAM
PBS RTT PAM
DNMT1 +1 A to C, +49 nick
PAM
PBS RTT PAM
cumulative fraction of deletions that start before position
cumulative fraction of deletions that end before position
1.0
0.8
0.6
0.4
0.2
0.0
500
550
600
650
700
750
800
position in target locus
PE
PE, +MLH1dn
nCas9-dead RT
nCas9-dead RT, +MLH1dn
nCas9
nCas9, +MLH1dn
PE, silent pegRNA
PE, silent pegRNA, +MLH1dn
X:301.6 pt
Y:424.66 pt

FIG. 17

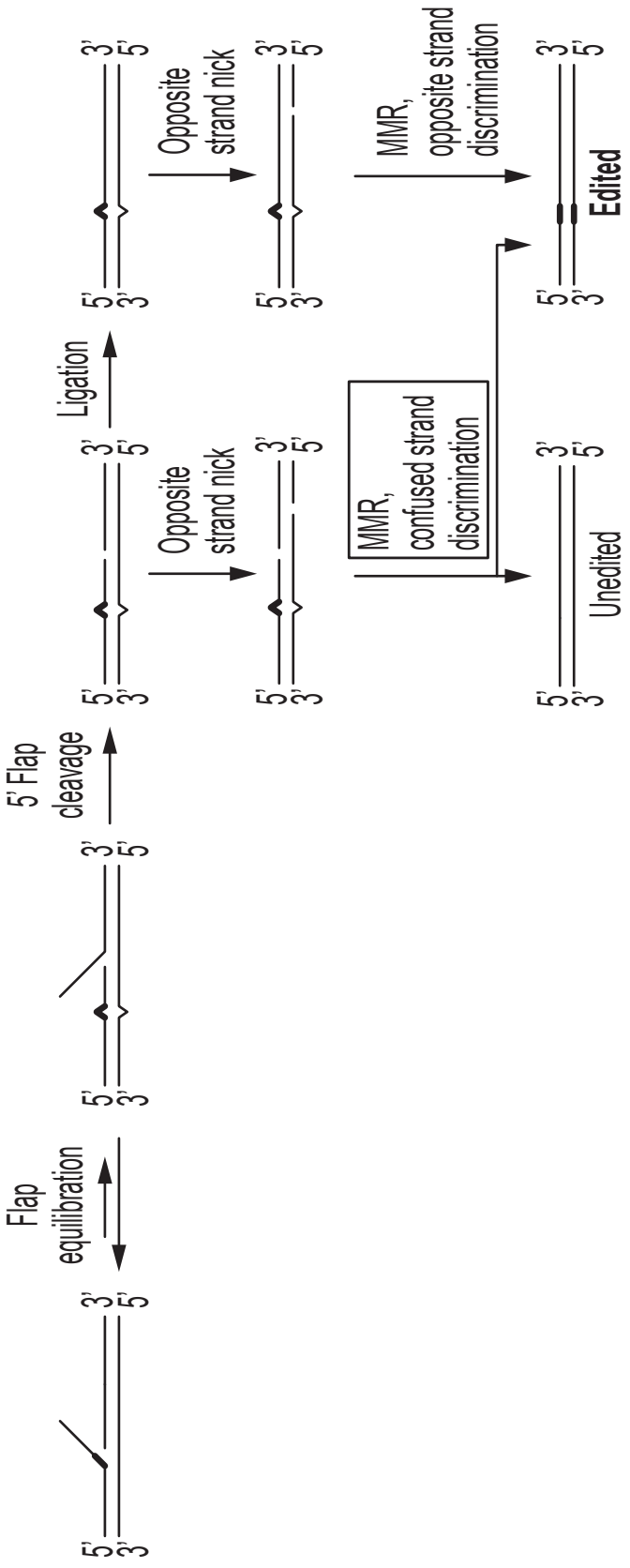

Flap equilibration
5' Flap cleavage
Ligation
Opposite strand nick
Opposite strand nick
MMR, confused strand discrimination
MMR, opposite strand discrimination
Edited
Unedited
5'
3'

FIG. 18

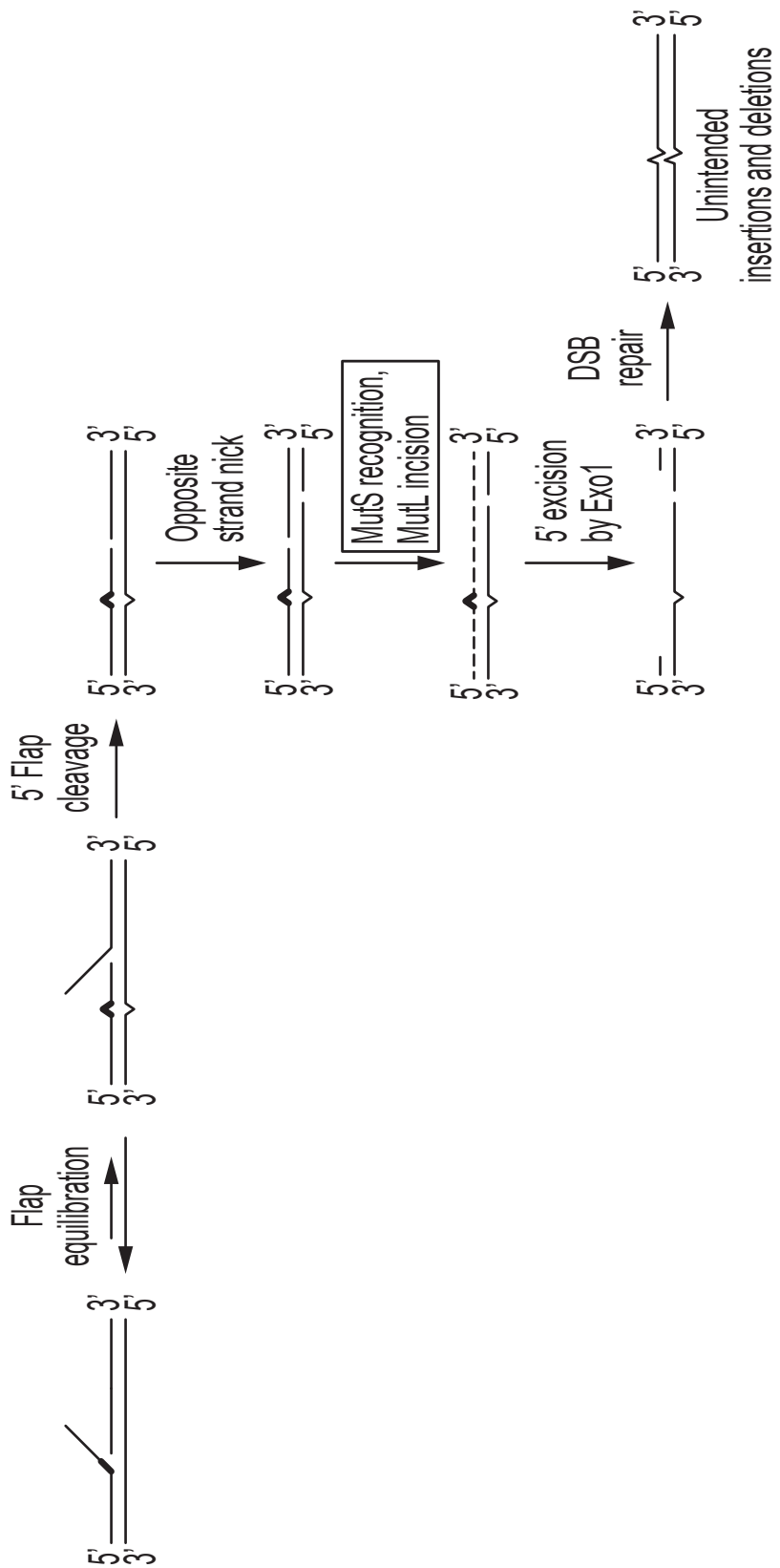

Flap equilibration
5' Flap cleavage
Opposite strand nick
MutS recognition, MutL incision
5' excision by Exo1
DSB repair
Unintended insertions and deletions
5'
3'

FIG. 19A

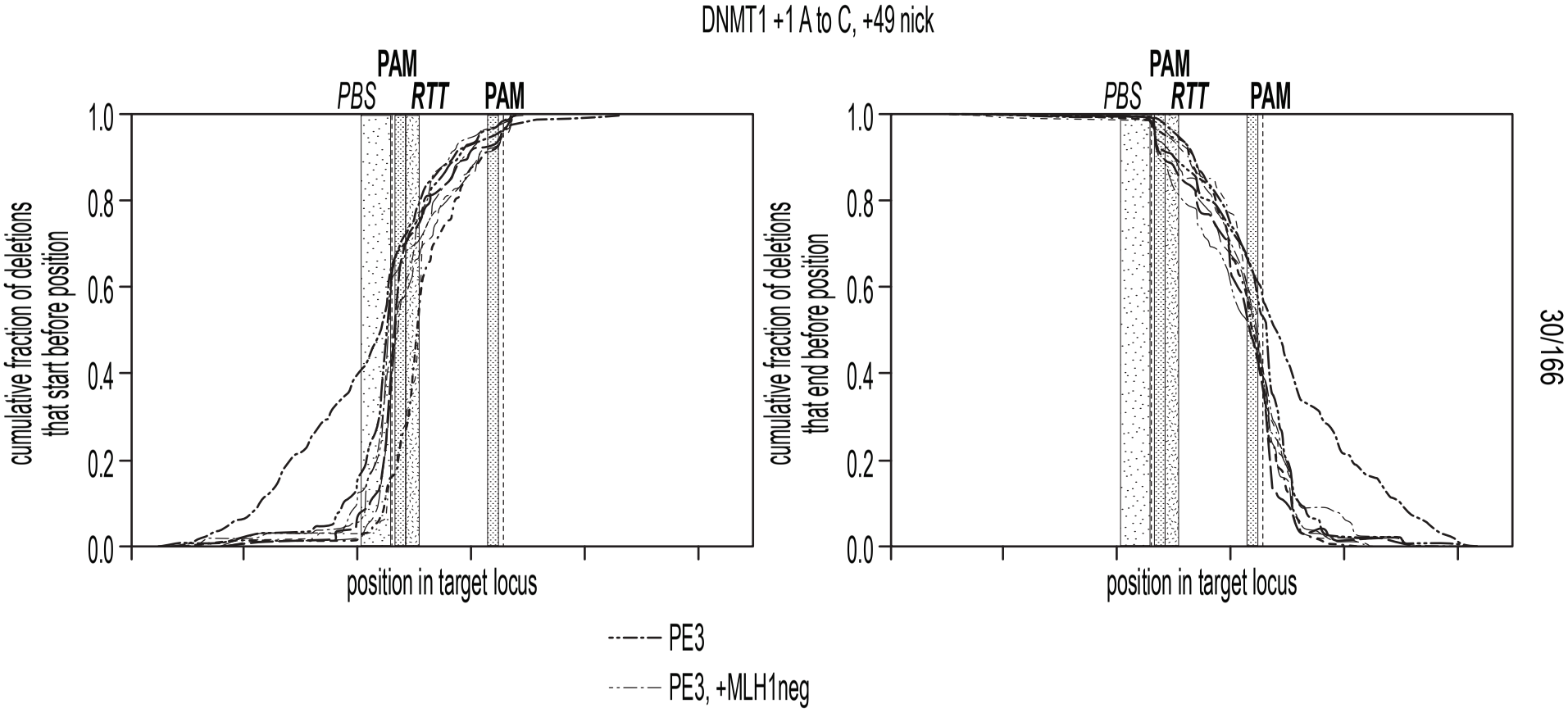

DNMT1 +1 A to C, +49 nick
PAM
PBS
RTT
PAM
1.0
0.8
0.6
0.4
0.2
0.0
cumulative fraction of deletions that start before position
position in target locus
PAM
PBS
RTT
PAM
1.0
0.8
0.6
0.4
0.2
0.0
cumulative fraction of deletions that end before position
position in target locus
PE3
PE3, +MLH1neg

FIG. 19B

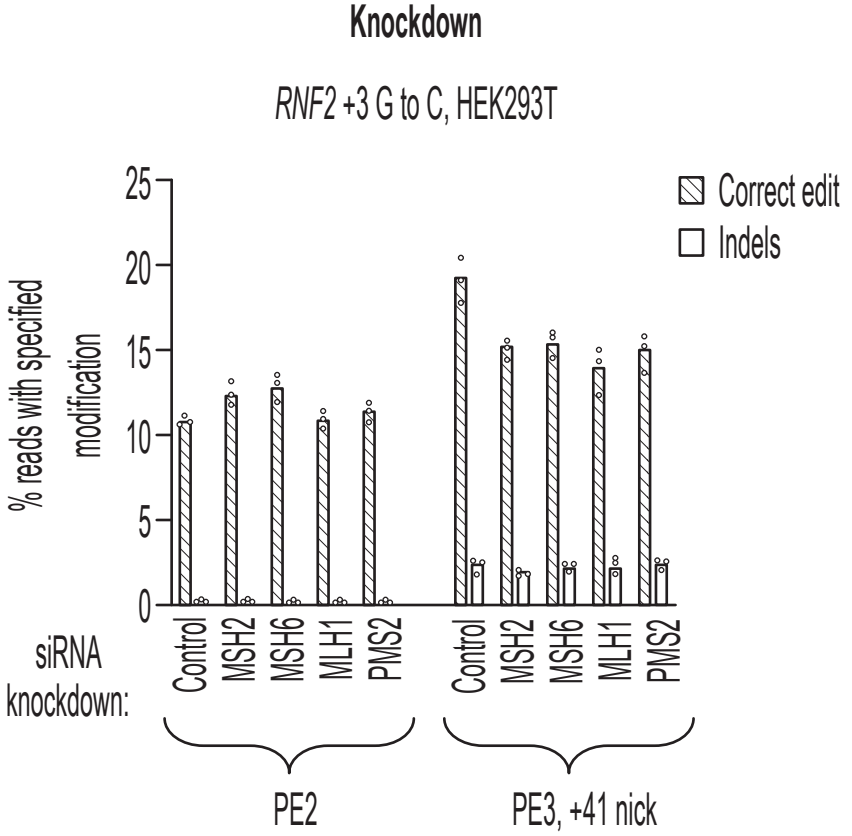


FIG. 20A

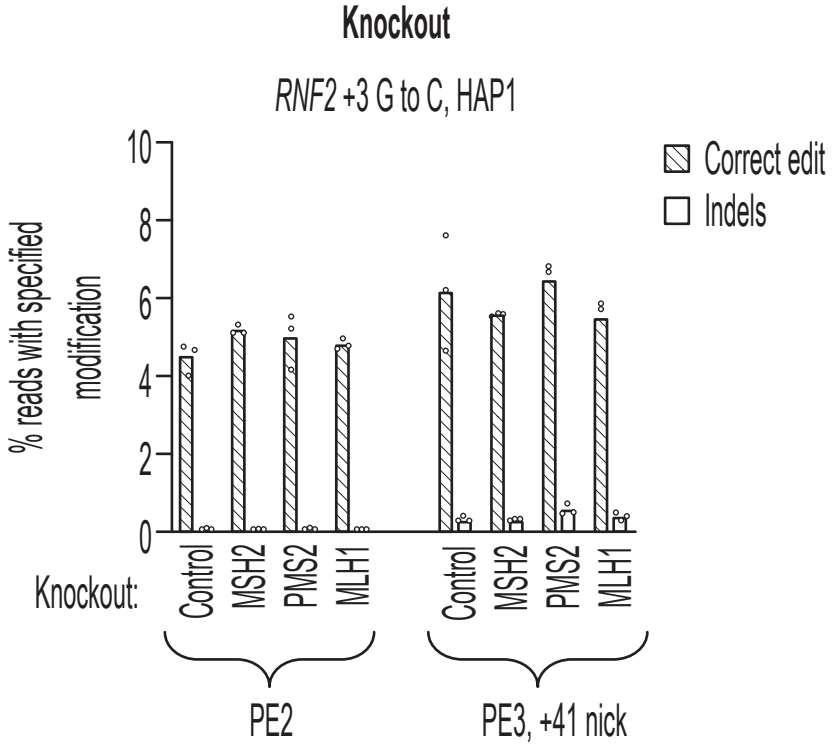


FIG. 20B

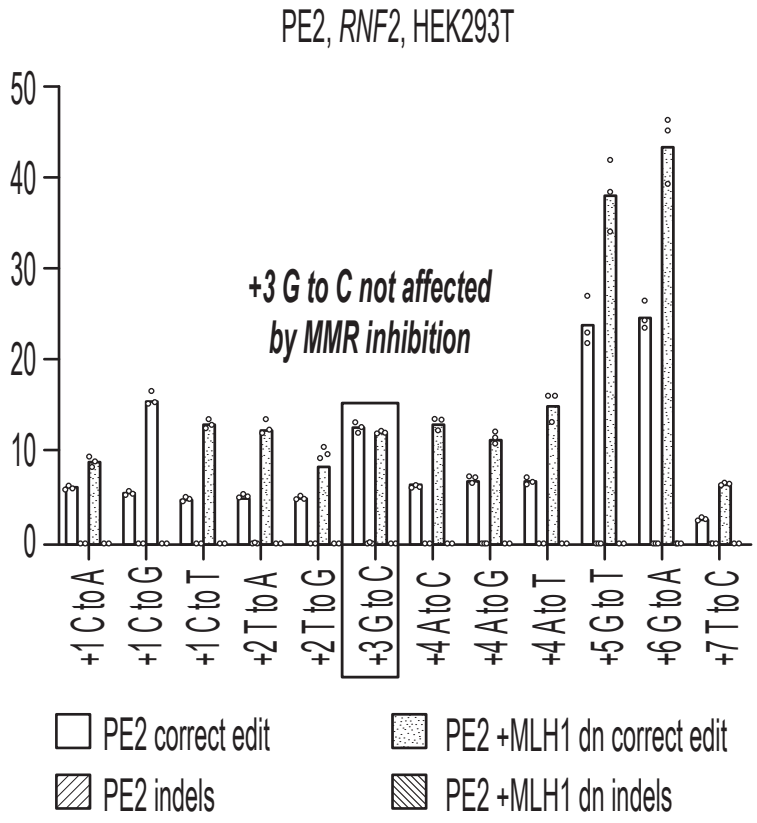


FIG. 21A

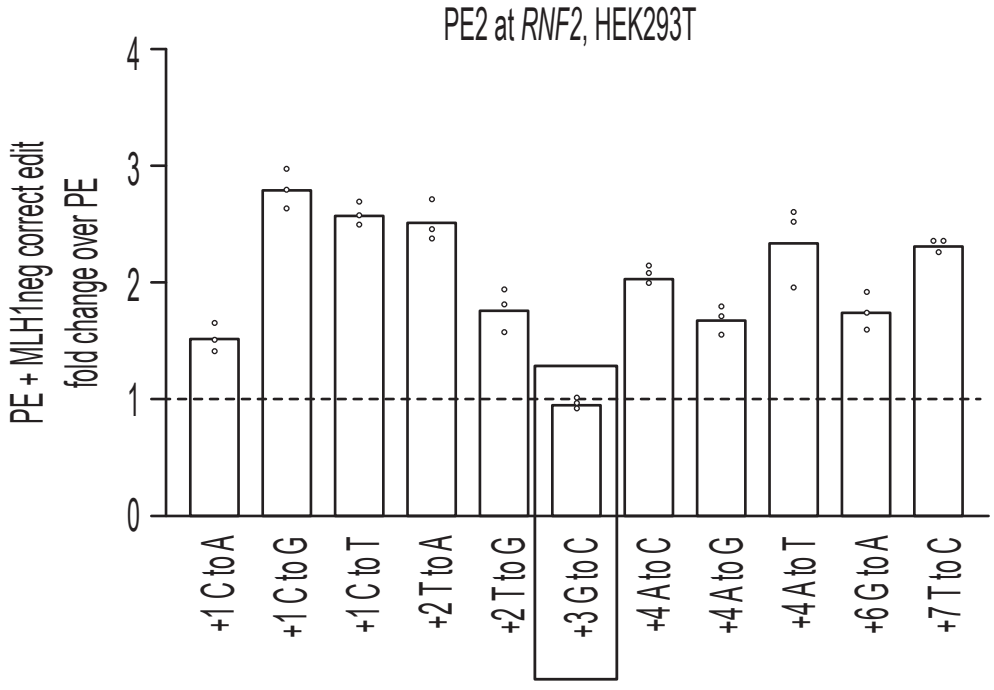


FIG. 21B

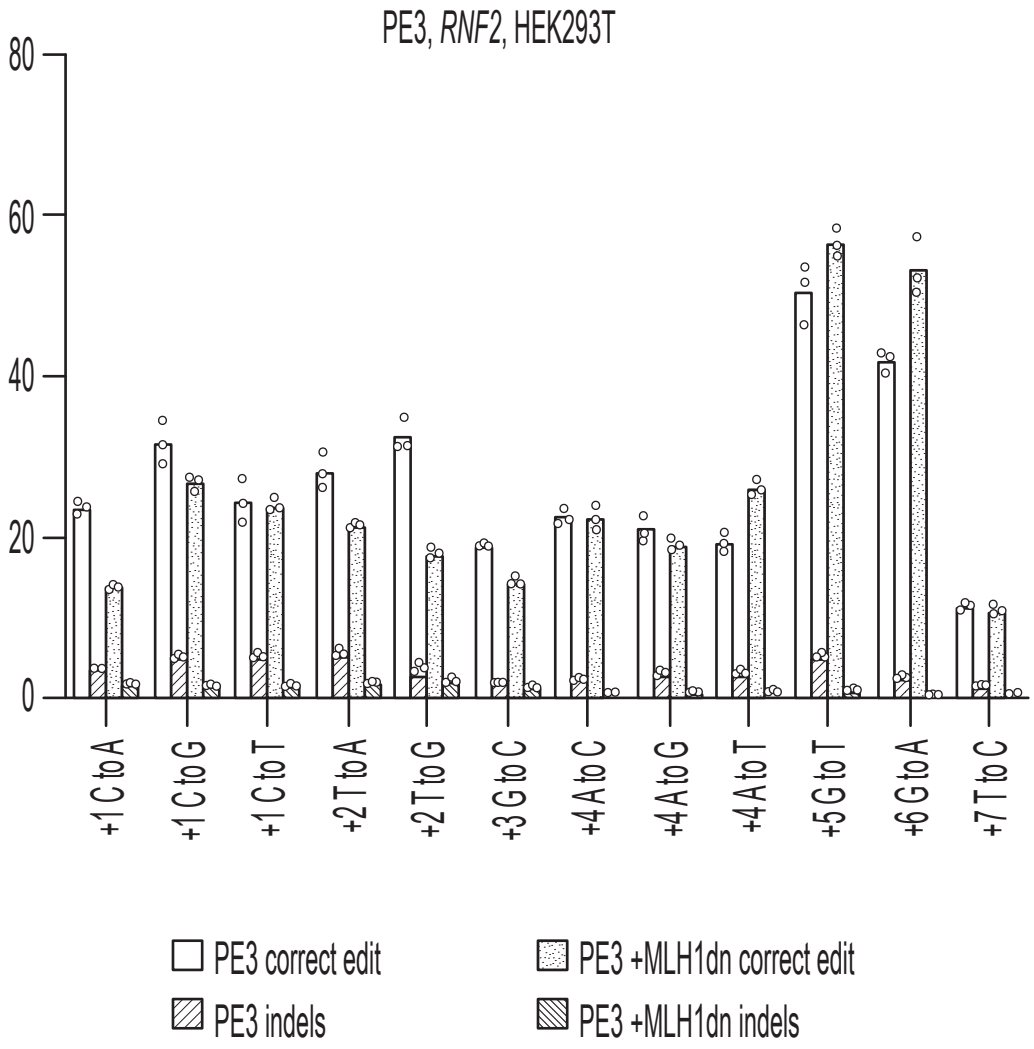

PE3, *RNF2*, HEK293T
80
60
40
20
0
+1 C to A
+1 C to G
+1 C to T
+2 T to A
+2 T to G
+3 G to C
+4 A to C
+4 A to G
+4 A to T
+5 G to T
+6 G to A
+7 T to C
PE3 correct edit
PE3 indels
PE3 +MLH1dn correct edit
PE3 +MLH1dn indels

FIG. 21C

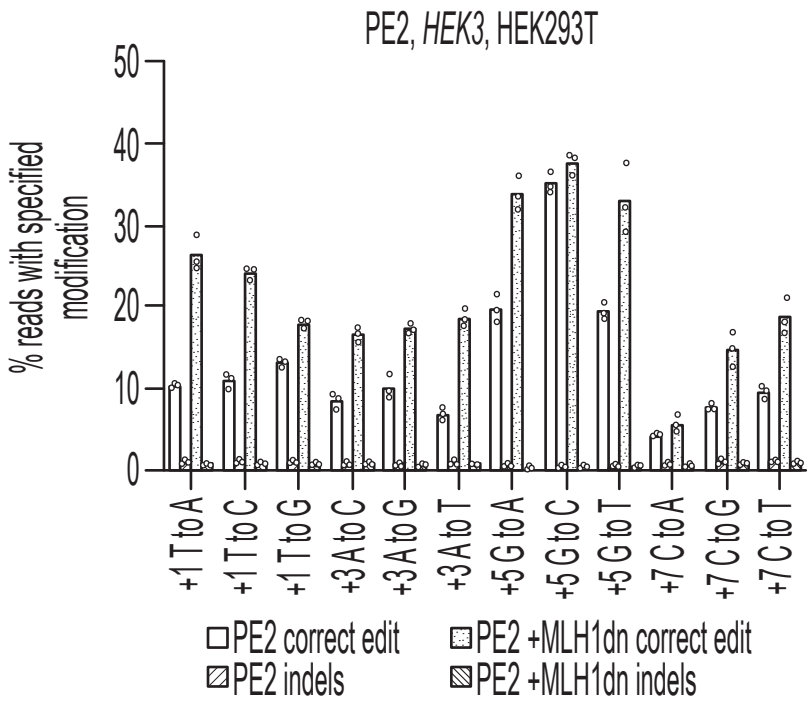


FIG. 22A

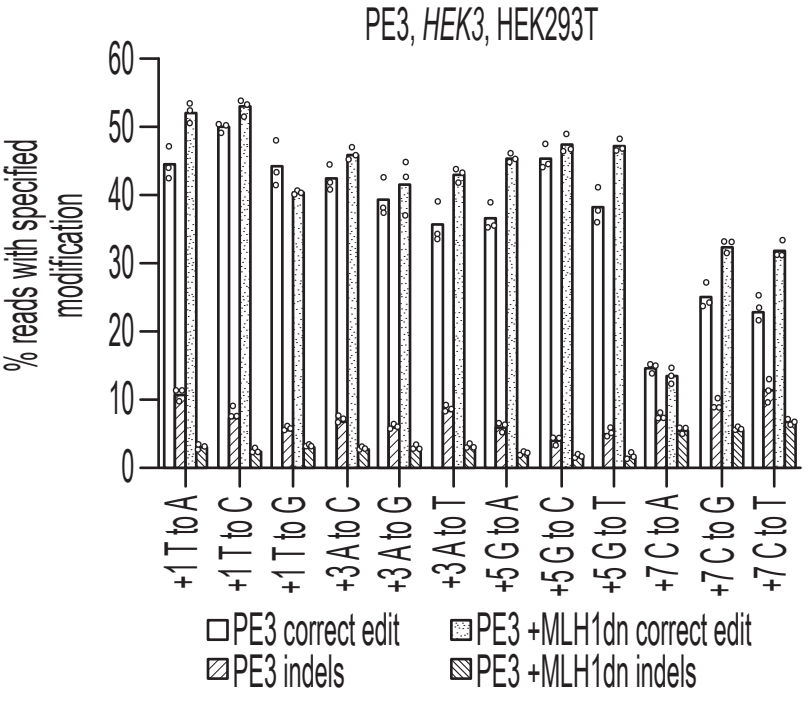


FIG. 22B

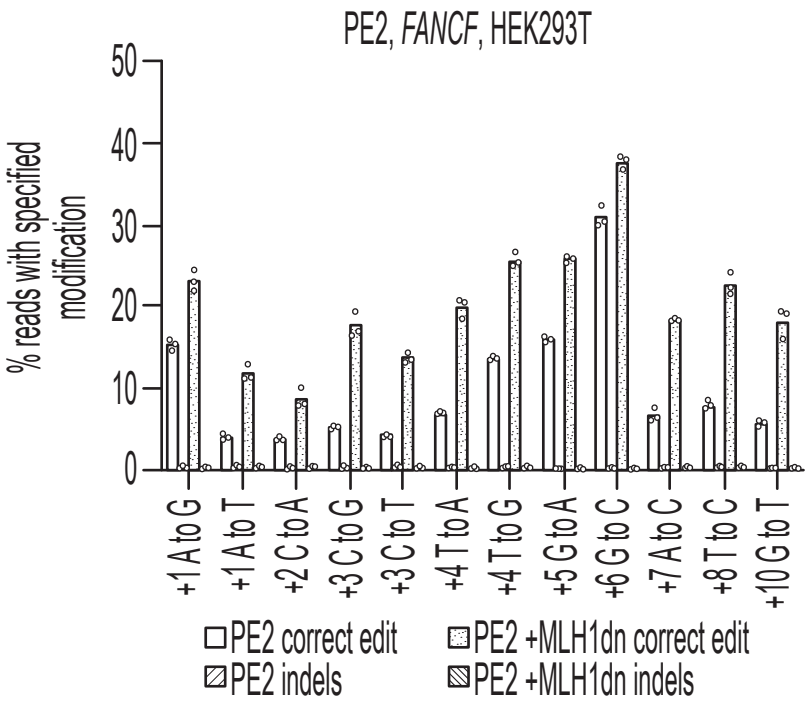


FIG. 23A

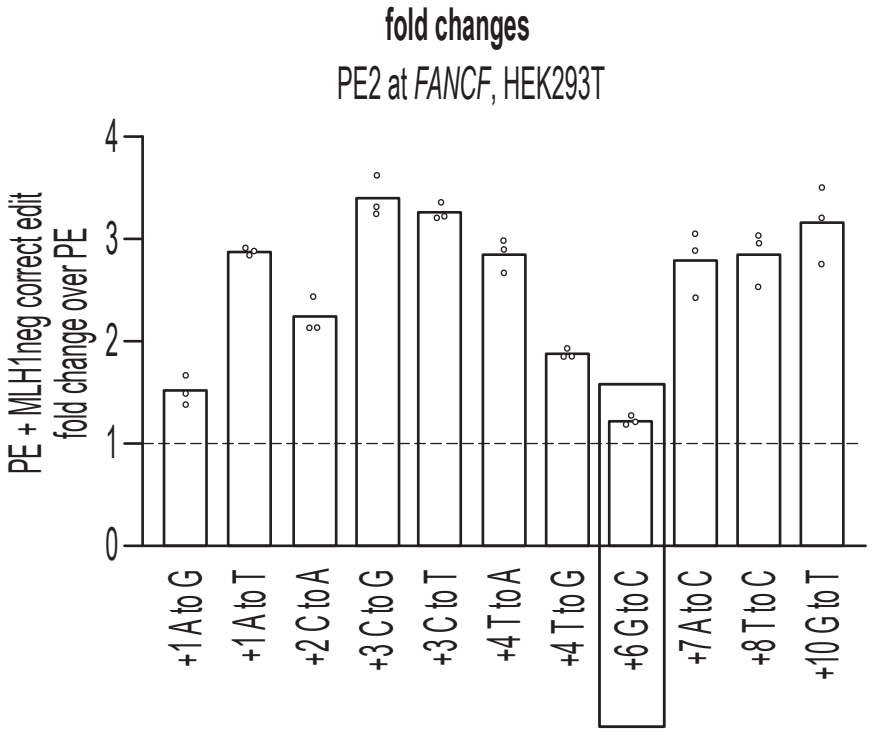


FIG. 23B

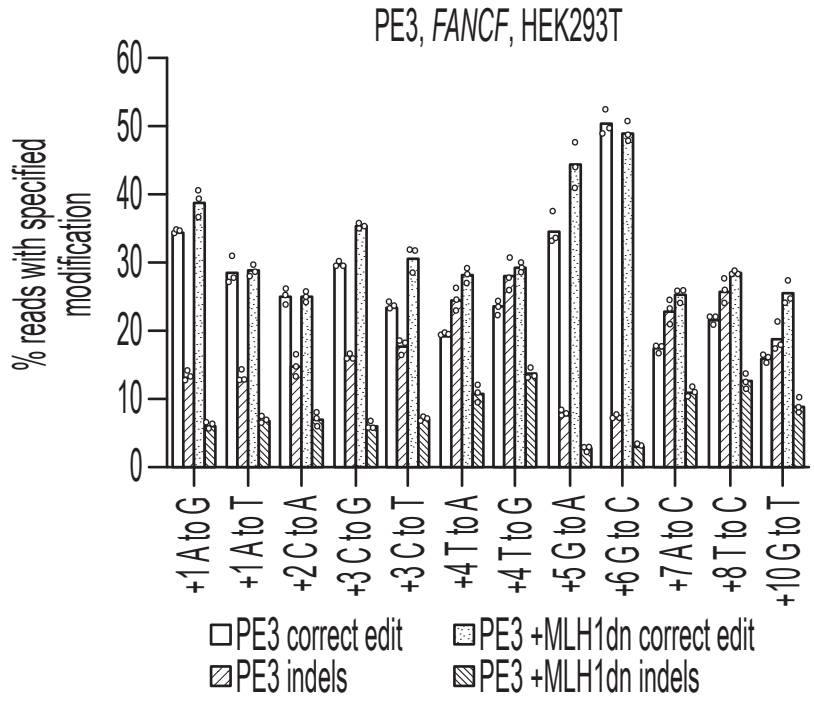


FIG. 23C

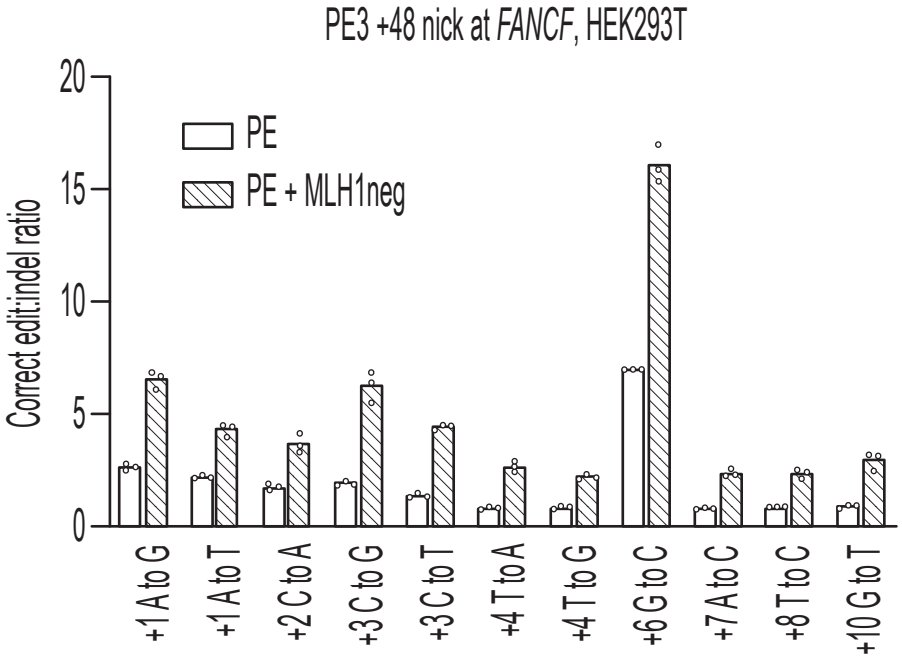


FIG. 23D

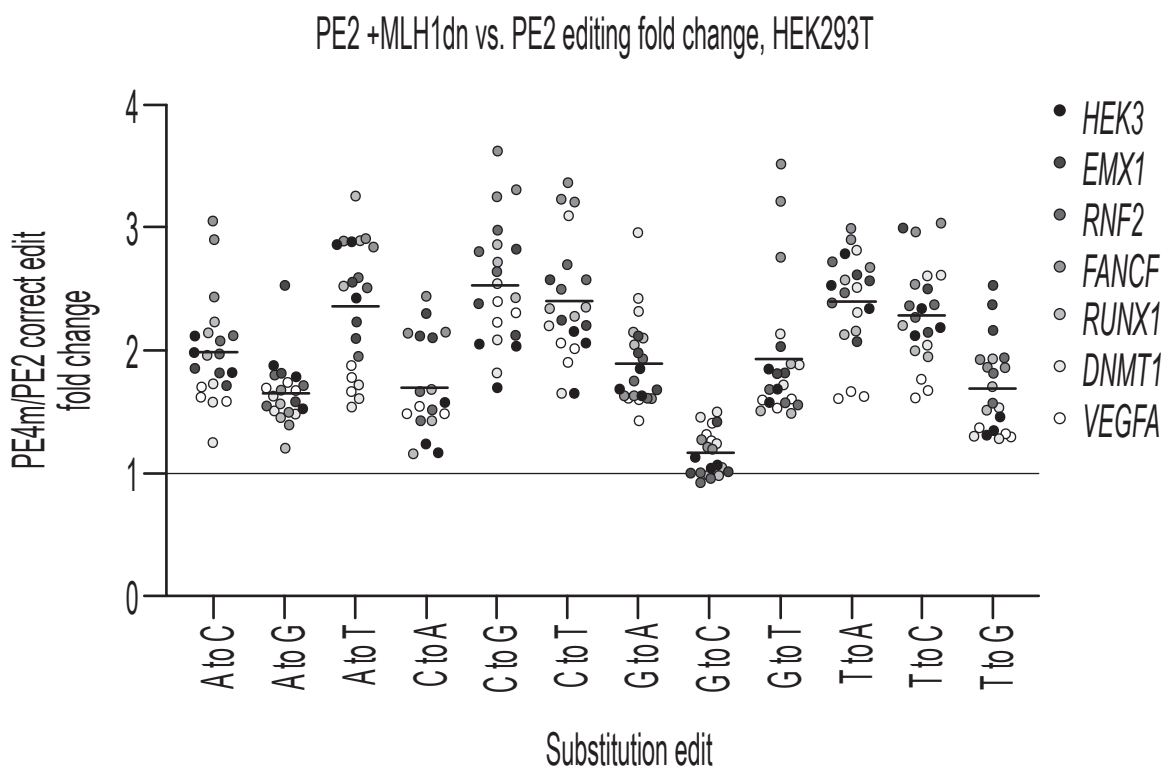


FIG. 24A

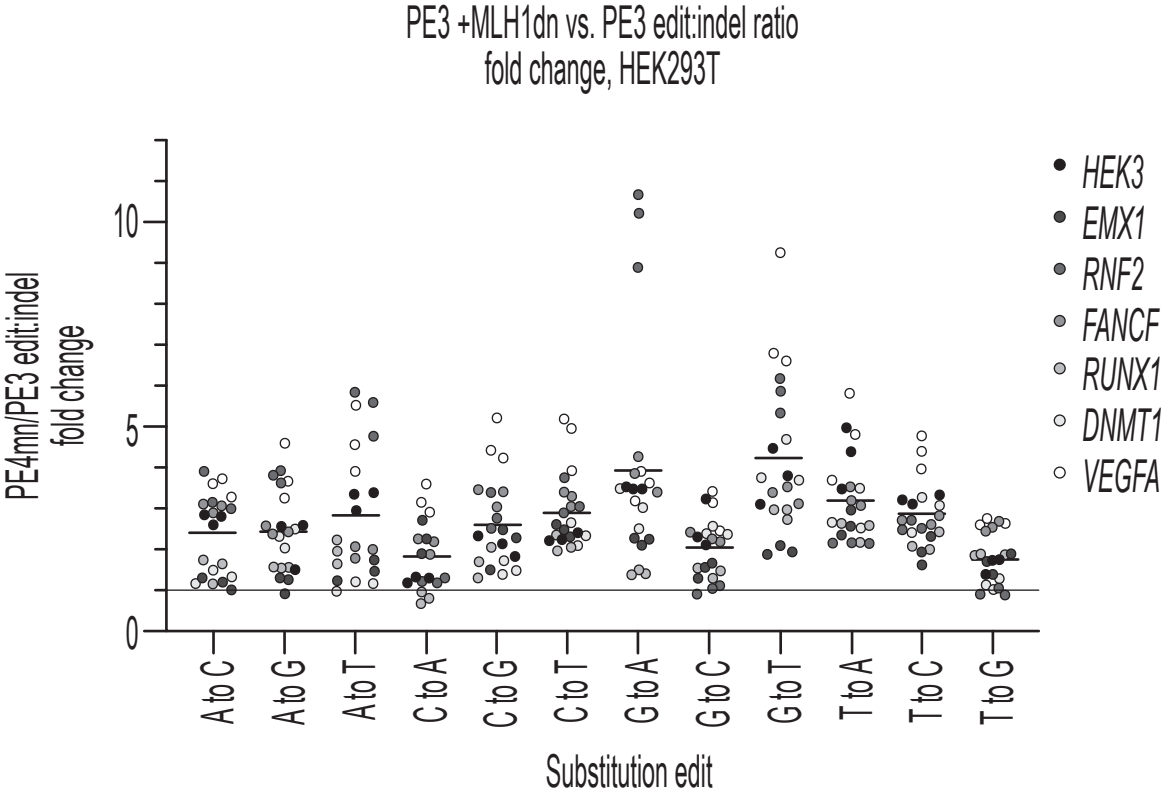


FIG. 24B

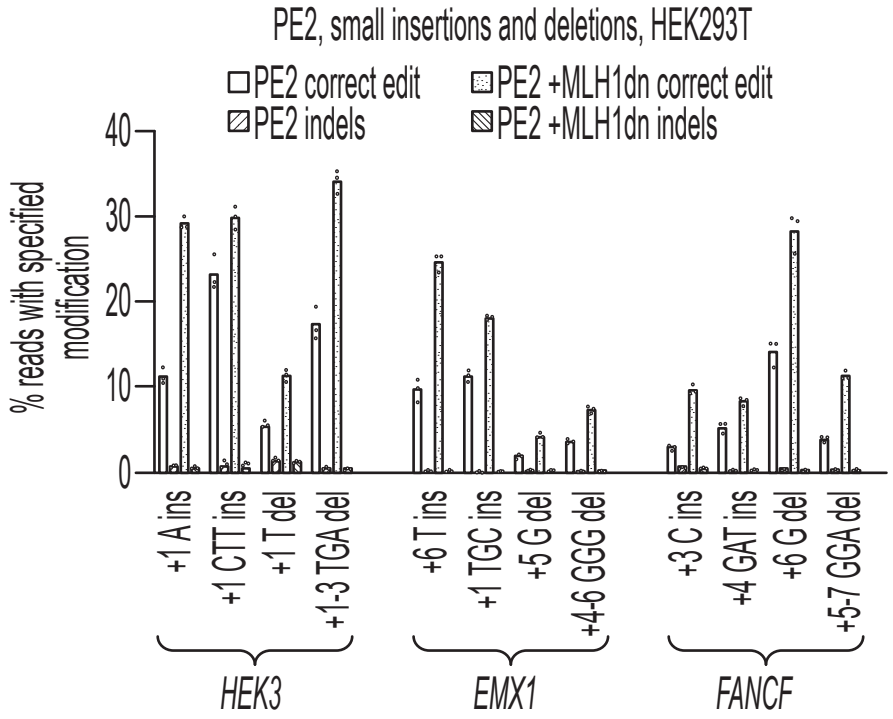


FIG. 25A

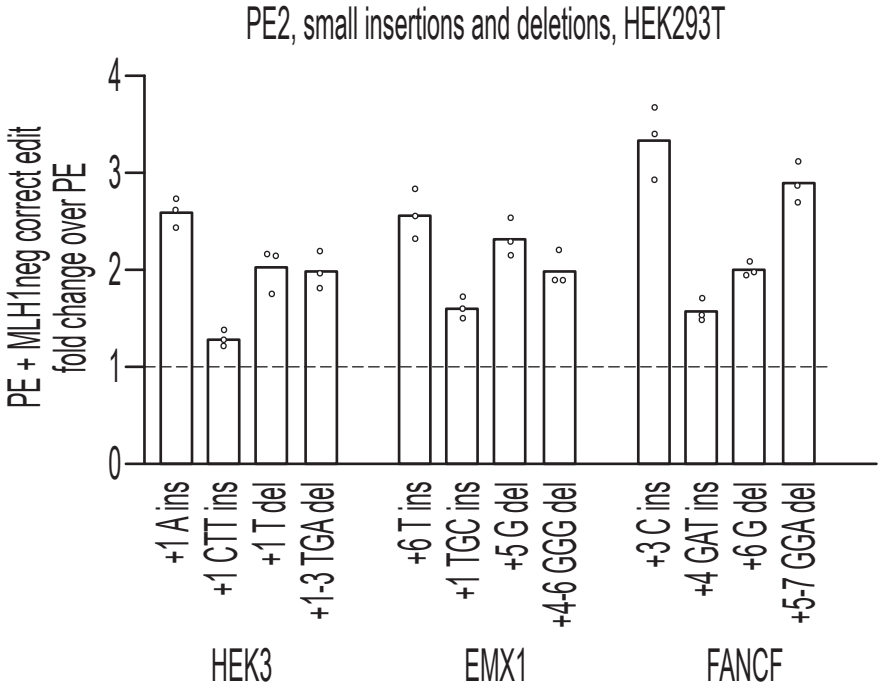


FIG. 25B

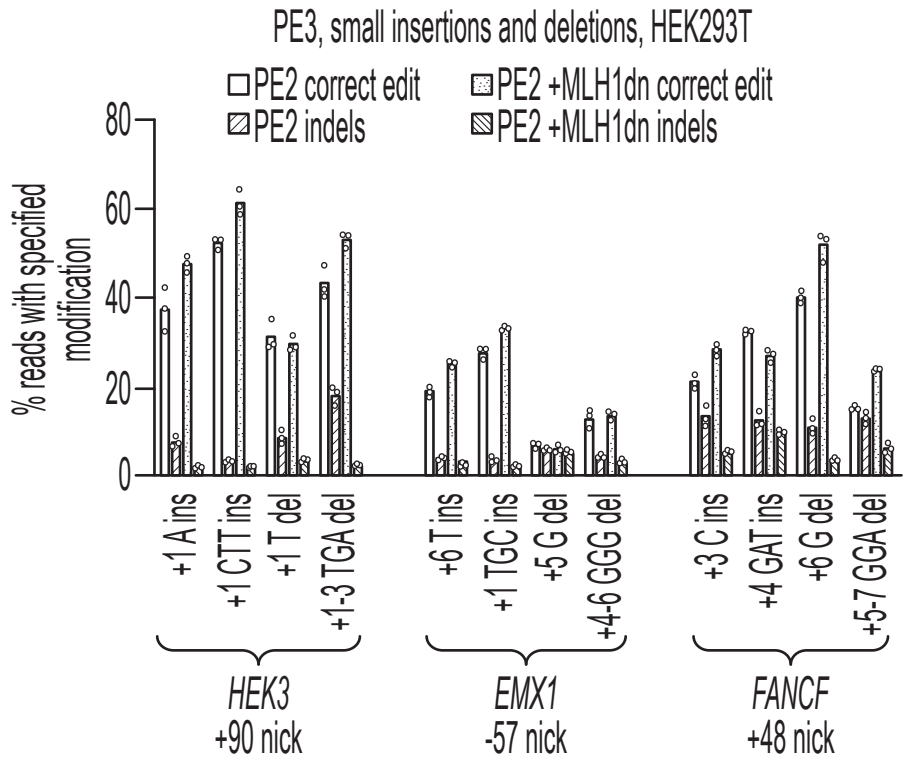


FIG. 25C

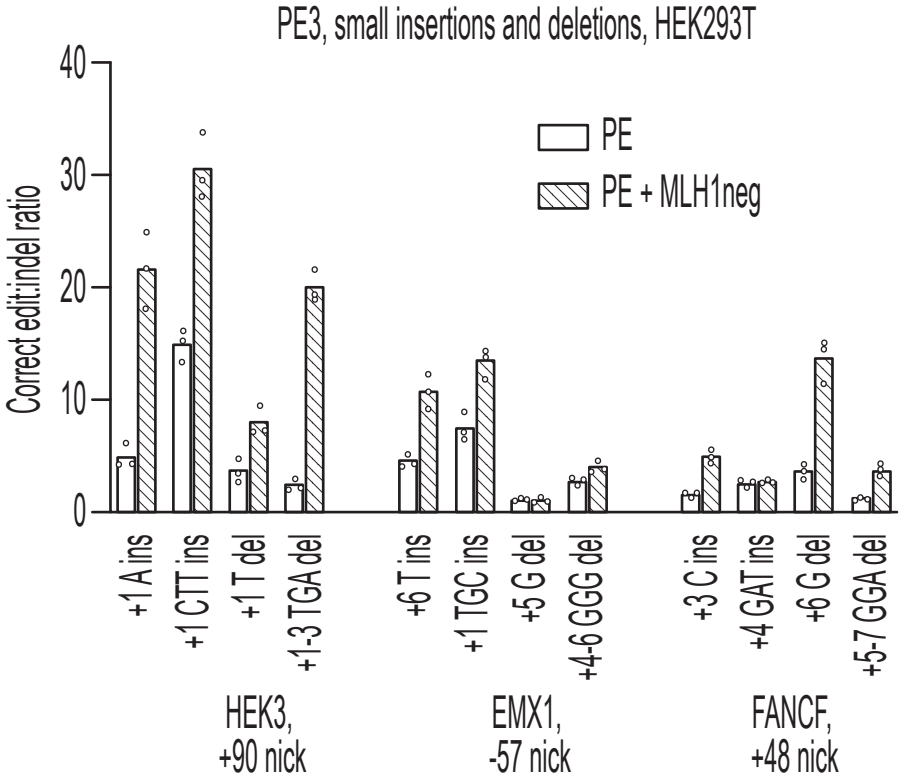


FIG. 25D

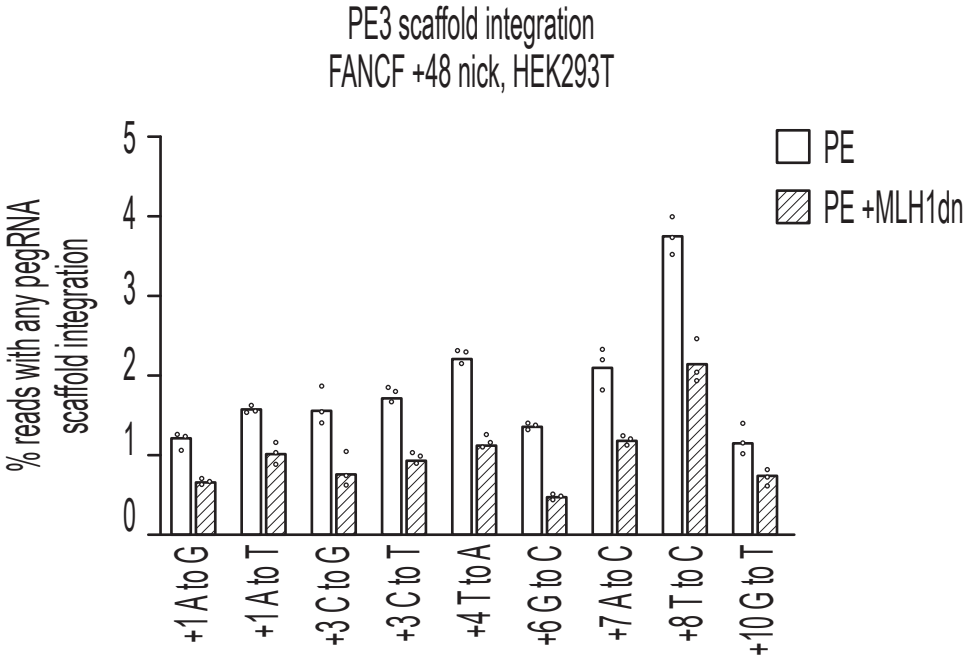


FIG. 26A

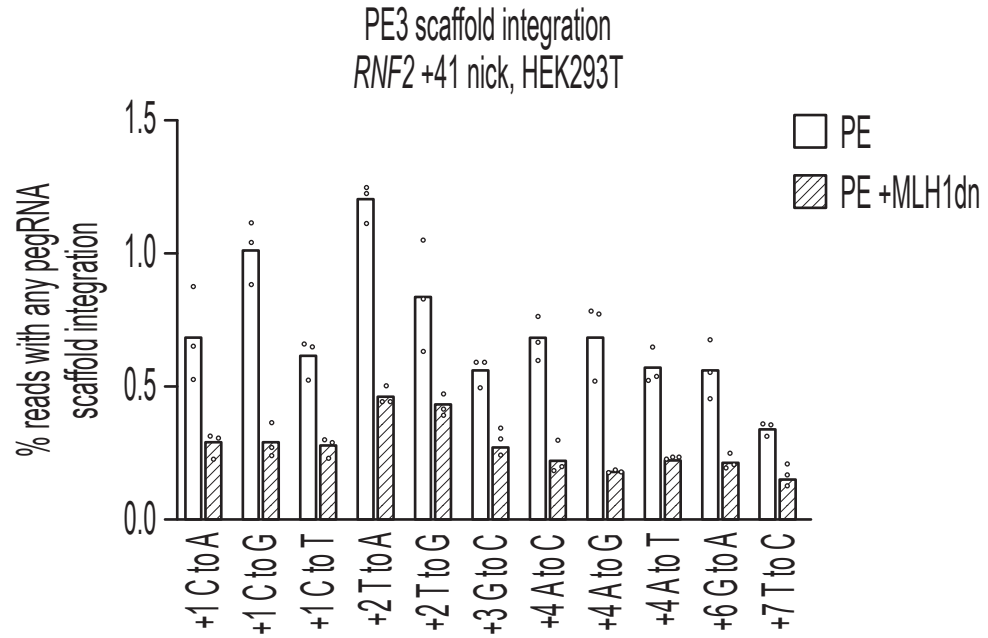


FIG. 26B

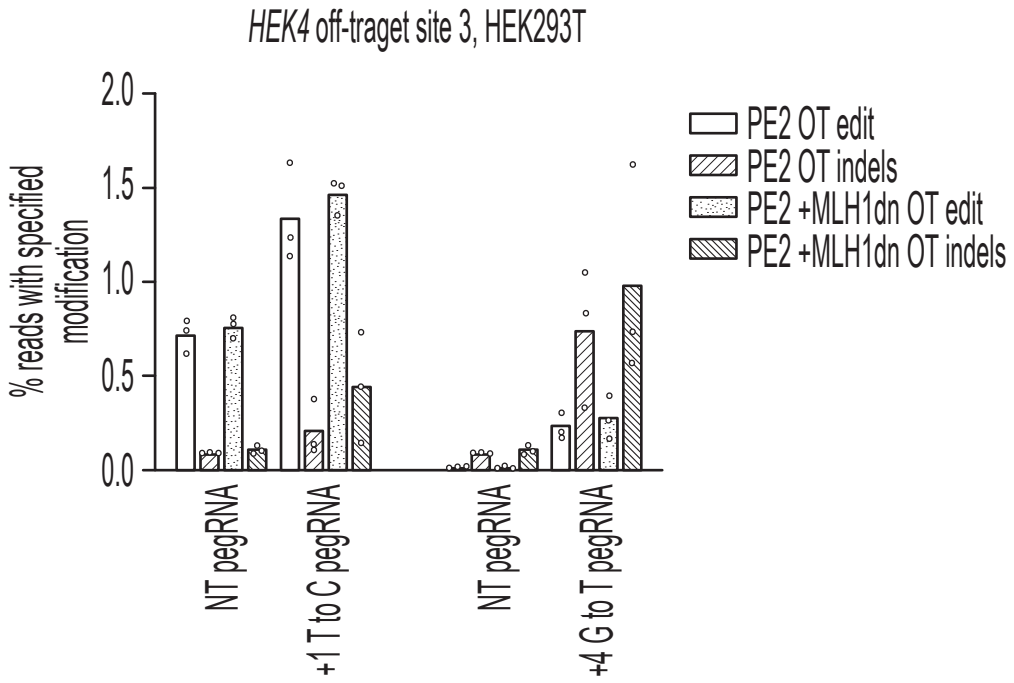

HEK4 off-traget site 3, HEK293T
% reads with specified modification
2.0
1.5
1.0
0.5
0.0
PE2 OT edit
PE2 OT indels
PE2 +MLH1dn OT edit
PE2 +MLH1dn OT indels
NT pegRNA
+1 T to C pegRNA
NT pegRNA
+4 G to T pegRNA

FIG. 27

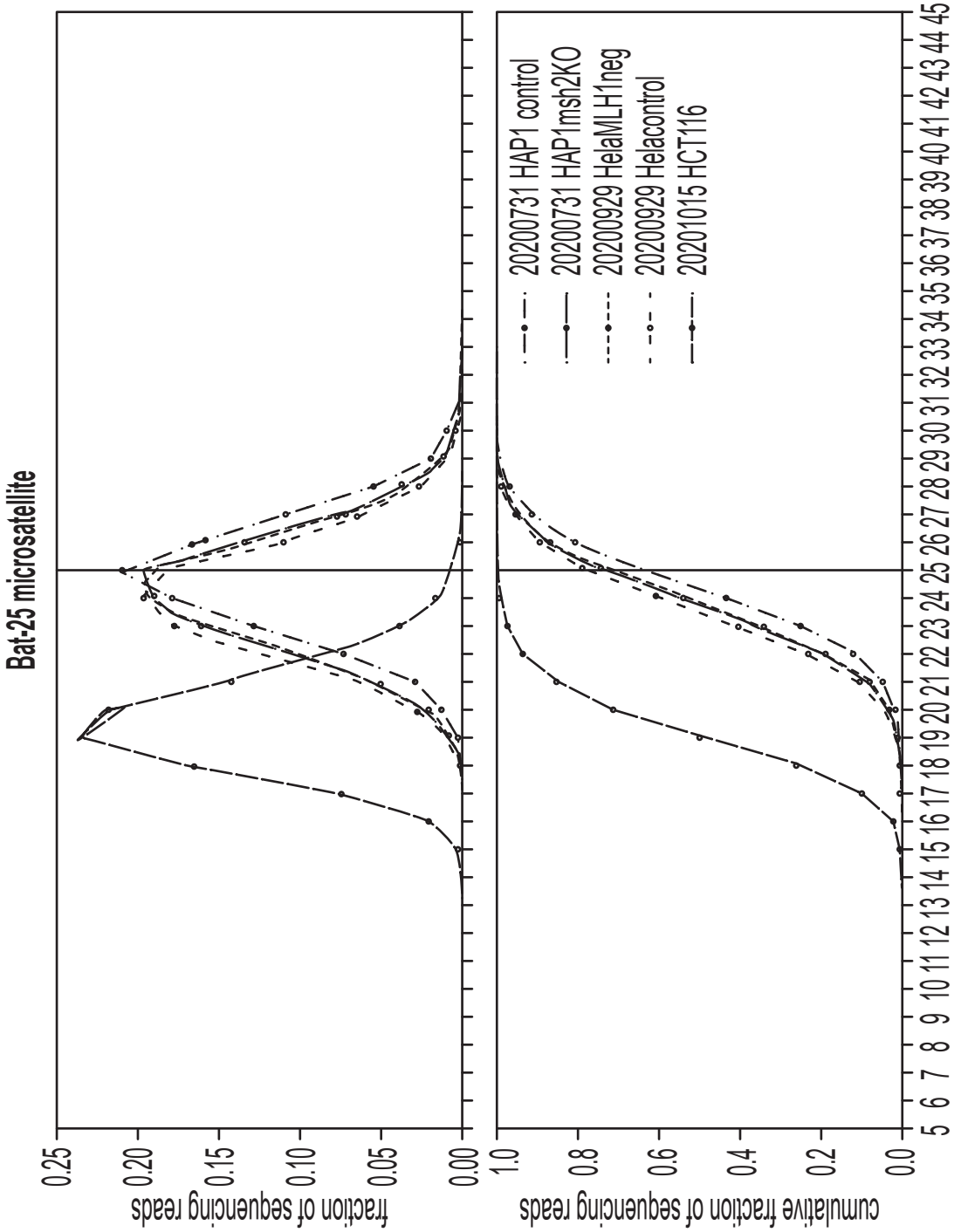

Bat-25 microsatellite
fraction of sequencing reads
0.25
0.20
0.15
0.10
0.05
0.00
cumulative fraction of sequencing reads
1.0
0.8
0.6
0.4
0.2
0.0
5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
distribution of microsatellite lengths (nt)
20200731 HAP1 control
20200731 HAP1msh2KO
20200929 HelaMLH1neg
20200929 Helacontrol
20201015 HCT116

FIG. 28A

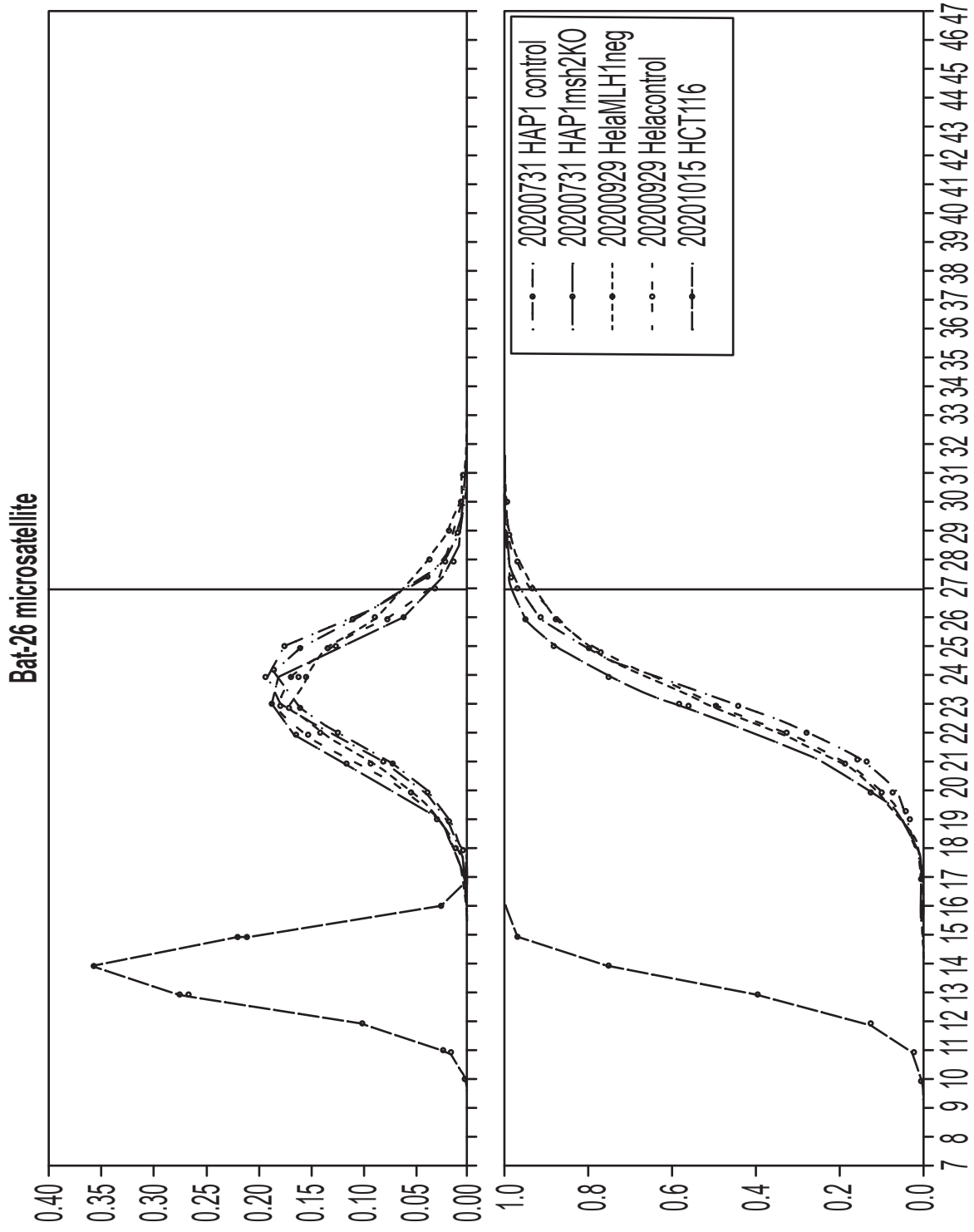

Bat-26 microsatellite
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
1.0
0.8
0.6
0.4
0.2
0.0
7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47
distribution of microsatellite lengths (nt)
20200731 HAP1 control
20200731 HAP1msh2KO
20200929 HelaMLH1neg
20200929 Helacontrol
20201015 HCT116

FIG. 28B

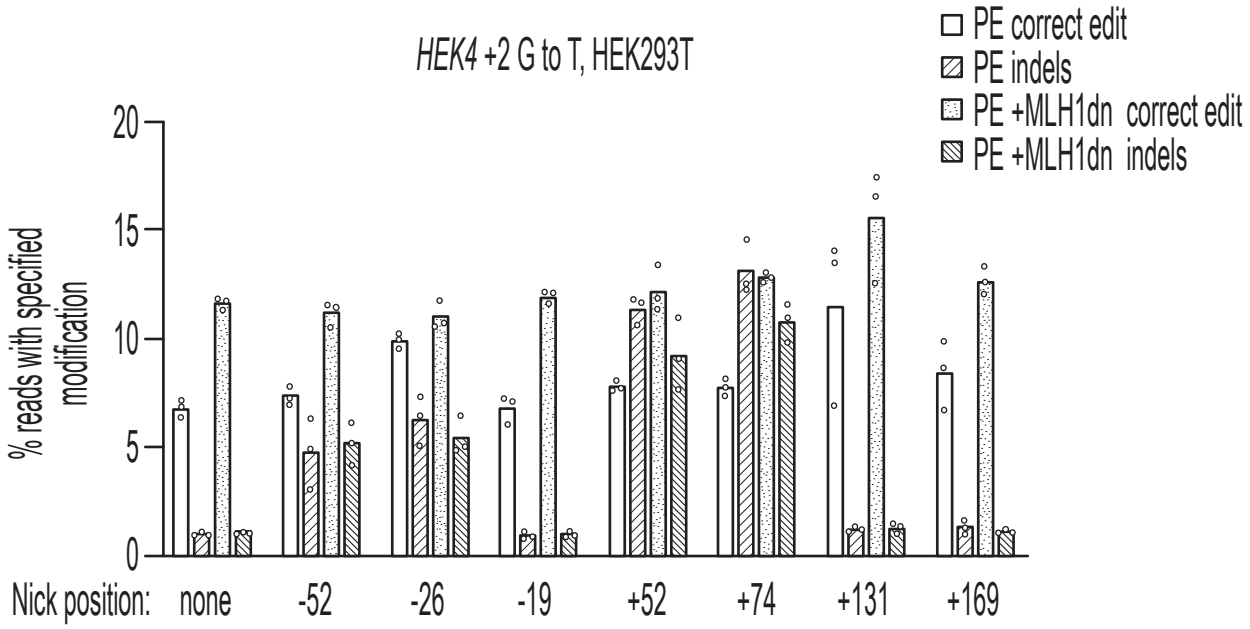

HEK4 +2 G to T, HEK293T
PE correct edit
PE indels
PE +MLH1dn correct edit
PE +MLH1dn indels
% reads with specified modification
20
15
10
5
0
Nick position:
none
-52
-26
-19
+52
+74
+131
+169

FIG. 29

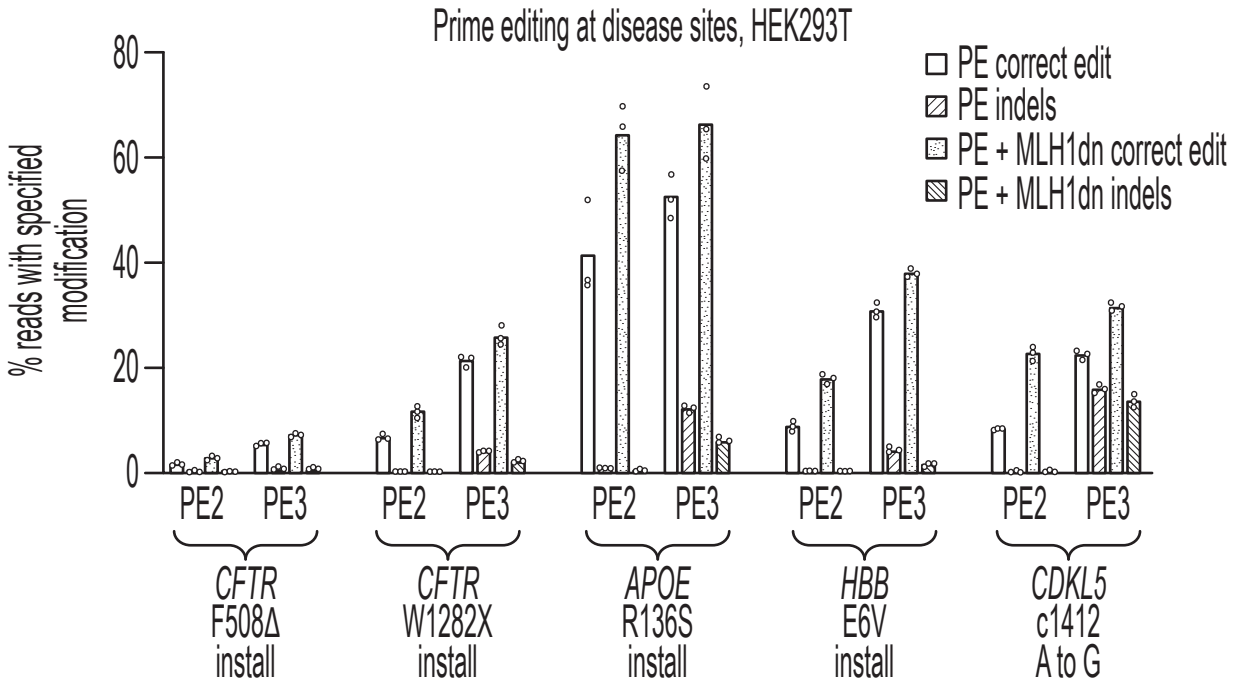

Prime editing at disease sites, HEK293T
% reads with specified modification
80
60
40
20
0
PE correct edit
PE indels
PE + MLH1dn correct edit
PE + MLH1dn indels
PE2
PE3
CFTR F508Δ install
CFTR W1282X install
APOE R136S install
HBB E6V install
CDKL5 c1412 A to G

FIG. 30

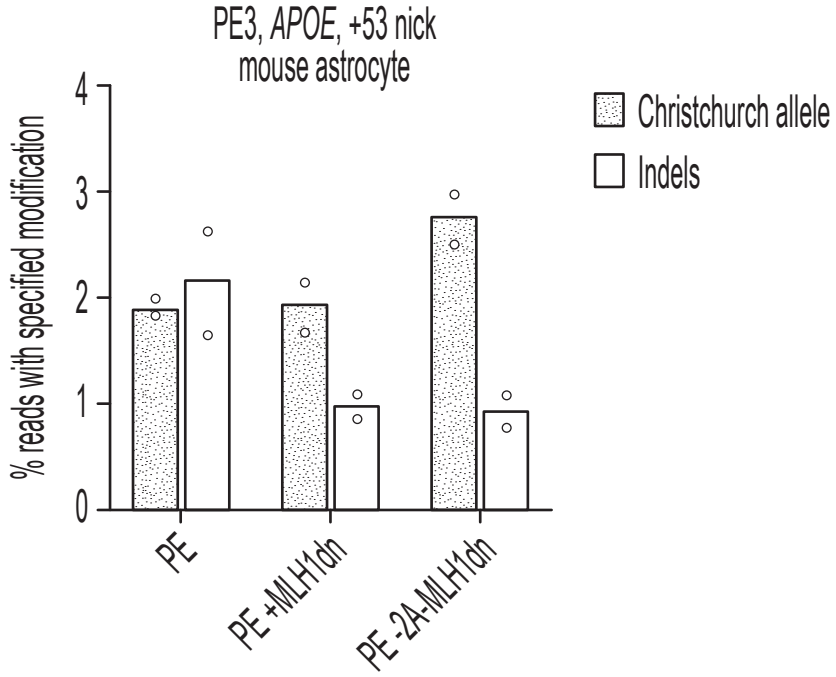

PE3, APOE, +53 nick
mouse astrocyte
% reads with specified modification
0
1
2
3
4
PE
PE+MLH1dn
PE-2A-MLH1dn
Christchurch allele
Indels

FIG. 31

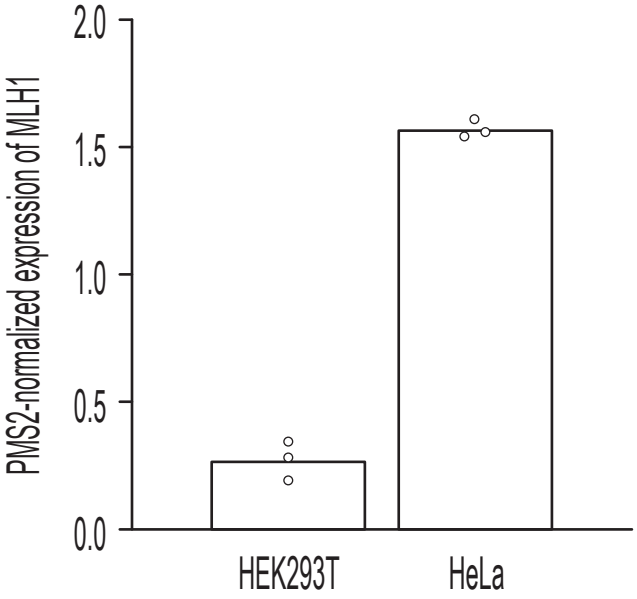

PMS2-normalized expression of MLH1
0.0
0.5
1.0
1.5
2.0
HEK293T
HeLa

FIG. 32

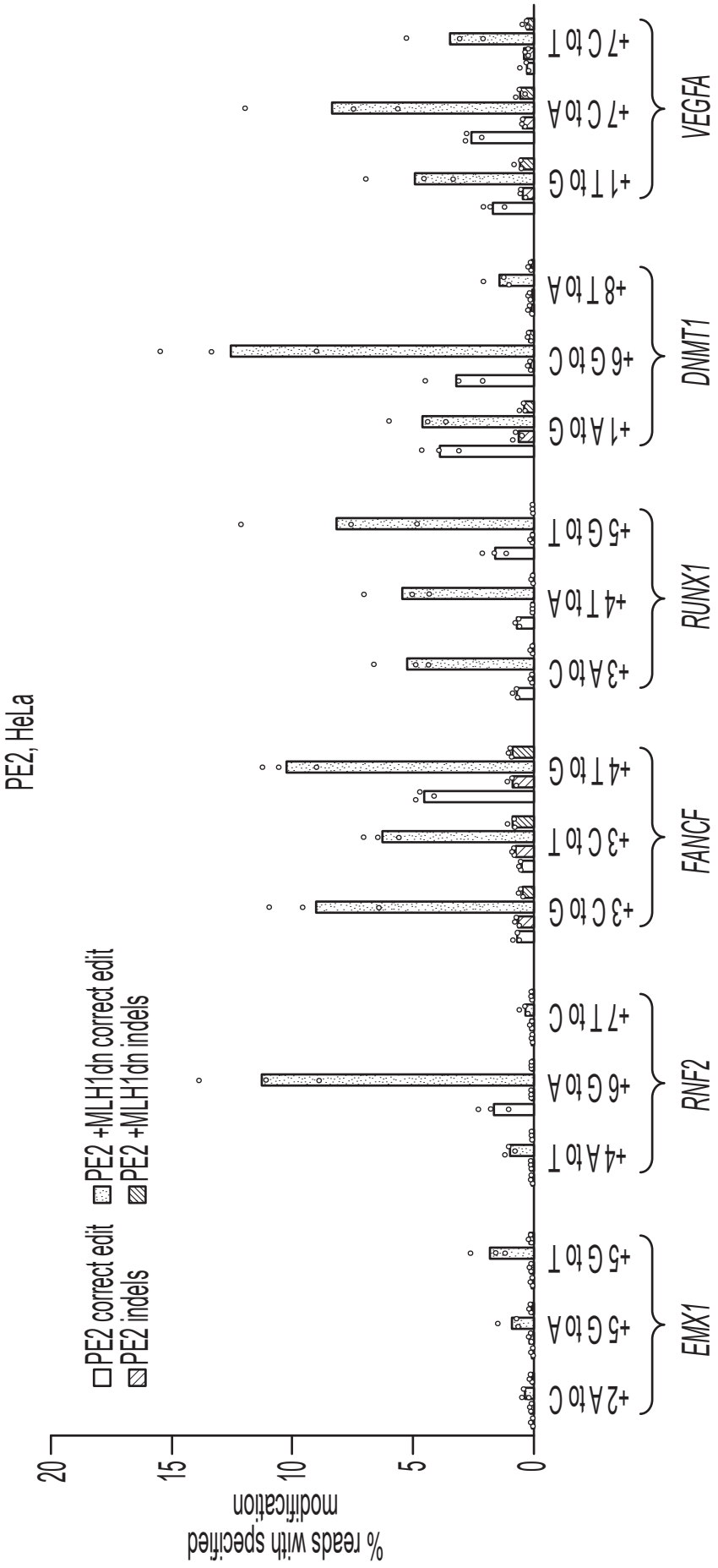

PE2, HeLa
% reads with specified modification
20
15
10
5
0
PE2 correct edit
PE2 indels
PE2 +MLH1dn correct edit
PE2 +MLH1dn indels
+2A to C
+5G to A
+5G to T
EMX1
+4A to T
+6G to A
+7T to C
RNF2
+3C to G
+3C to T
+4T to G
FANCF
+3A to C
+4T to A
+5G to T
RUNX1
+1A to G
+6G to C
+8T to A
DNMT1
+1T to G
+7C to A
+7C to T
VEGFA

FIG. 33A

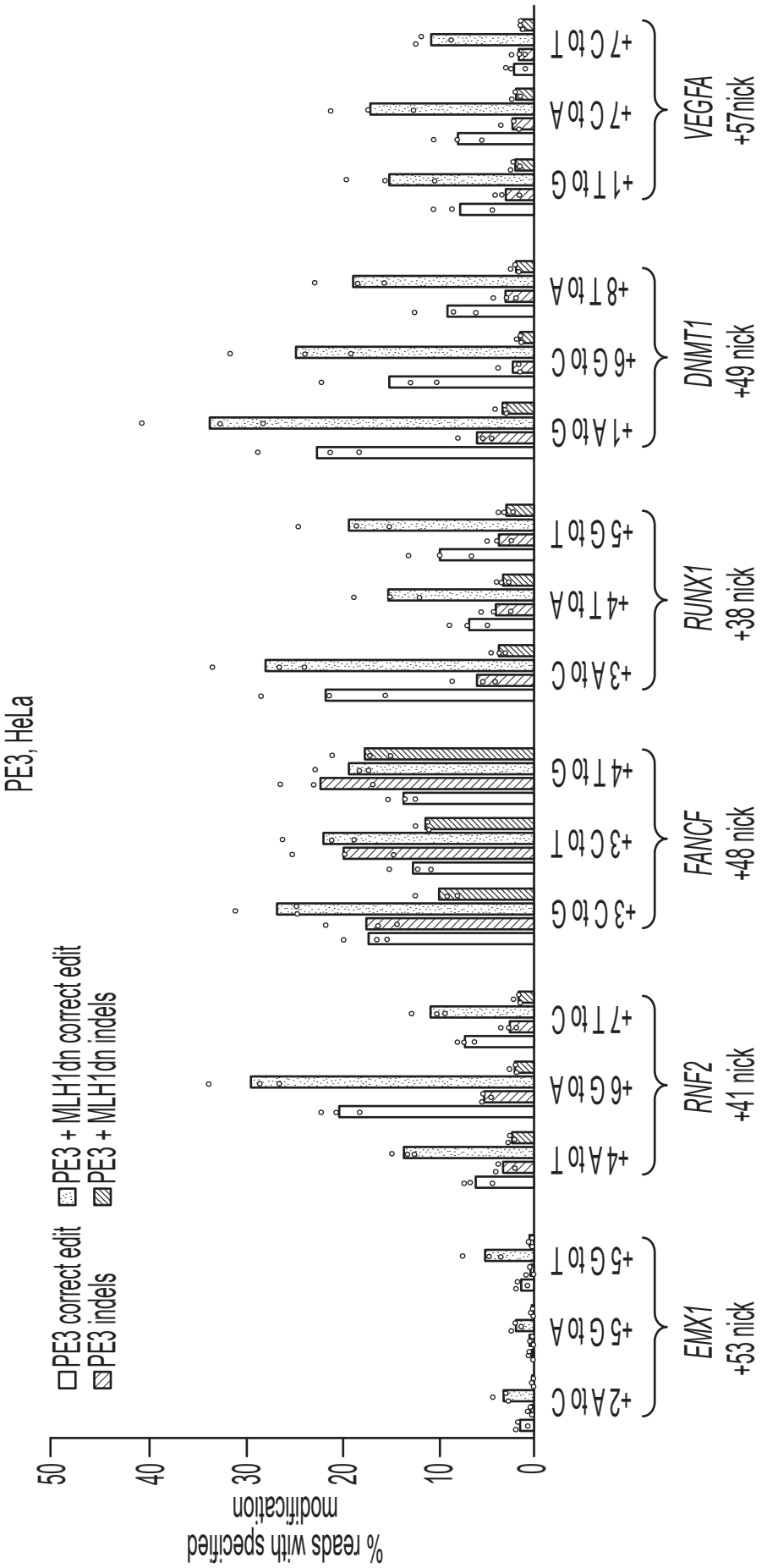

PE3, HeLa
PE3 correct edit
PE3 indels
PE3 + MLH1dn correct edit
PE3 + MLH1dn indels
% reads with specified modification
0
10
20
30
40
50
+2A to C
+5G to A
+5G to T
EMX1
+53 nick
+4A to T
+6G to A
+7T to C
RNF2
+41 nick
+3C to G
+3C to T
+4T to G
FANCF
+48 nick
+3A to C
+4T to A
+5G to T
RUNX1
+38 nick
+1A to G
+6G to C
+8T to A
DNMT1
+49 nick
+1T to G
+7C to A
+7C to T
VEGFA
+57nick

FIG. 33B

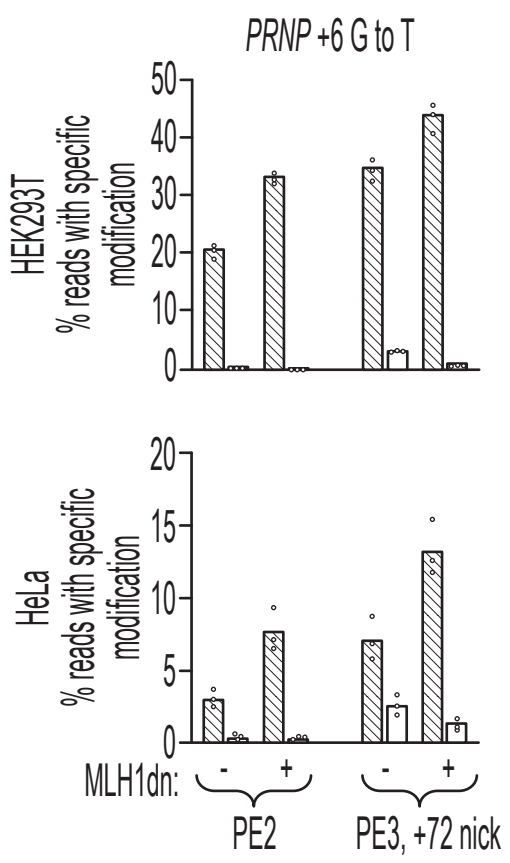


FIG. 34A

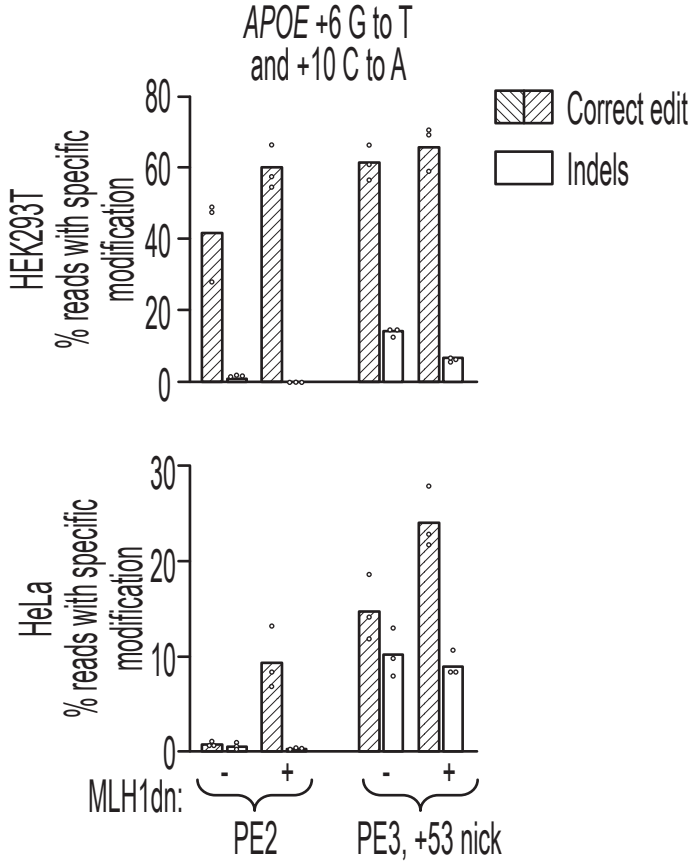


FIG. 34B

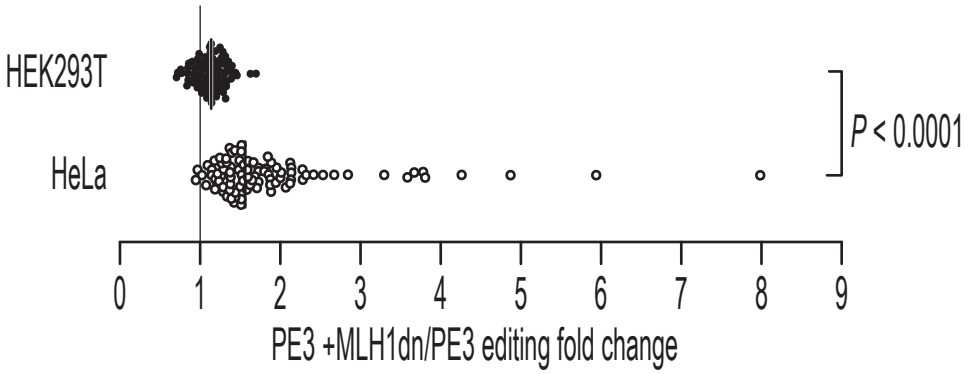


FIG. 35A

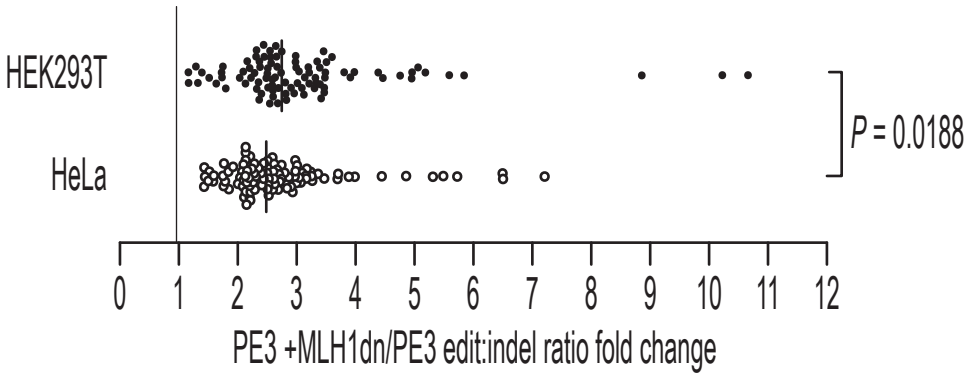


FIG. 35B

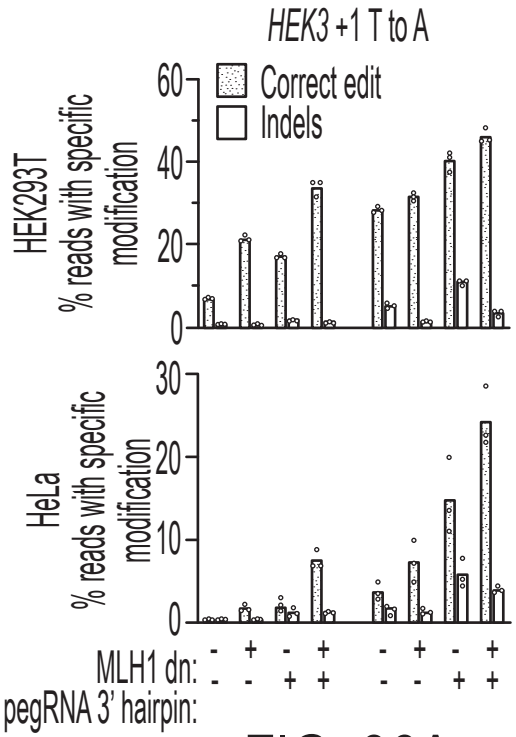


FIG. 36A

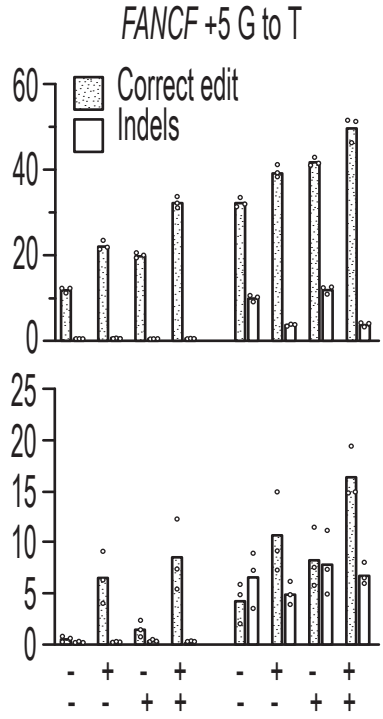


FIG. 36B

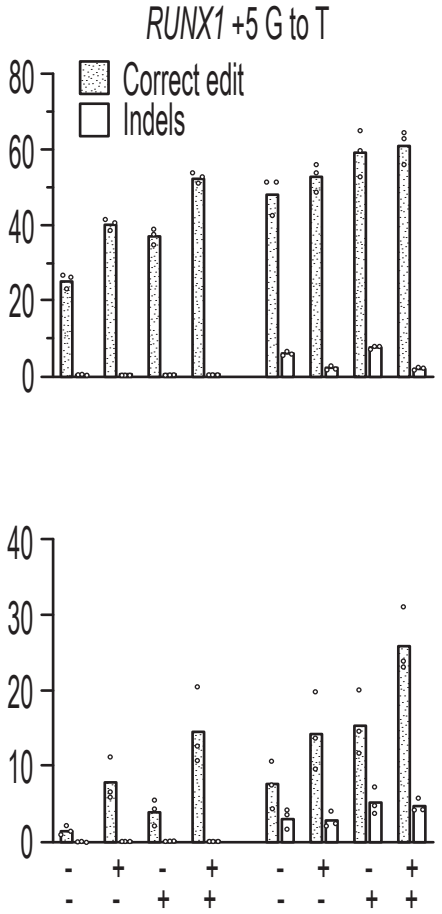


FIG. 36C

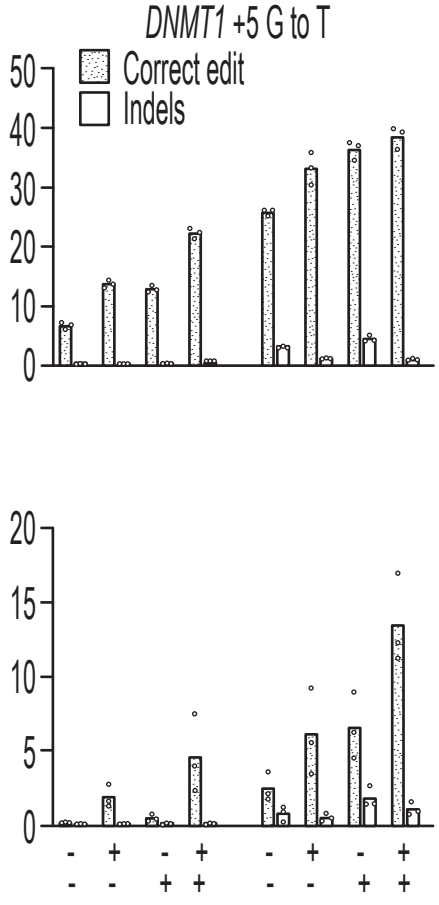


FIG. 36D

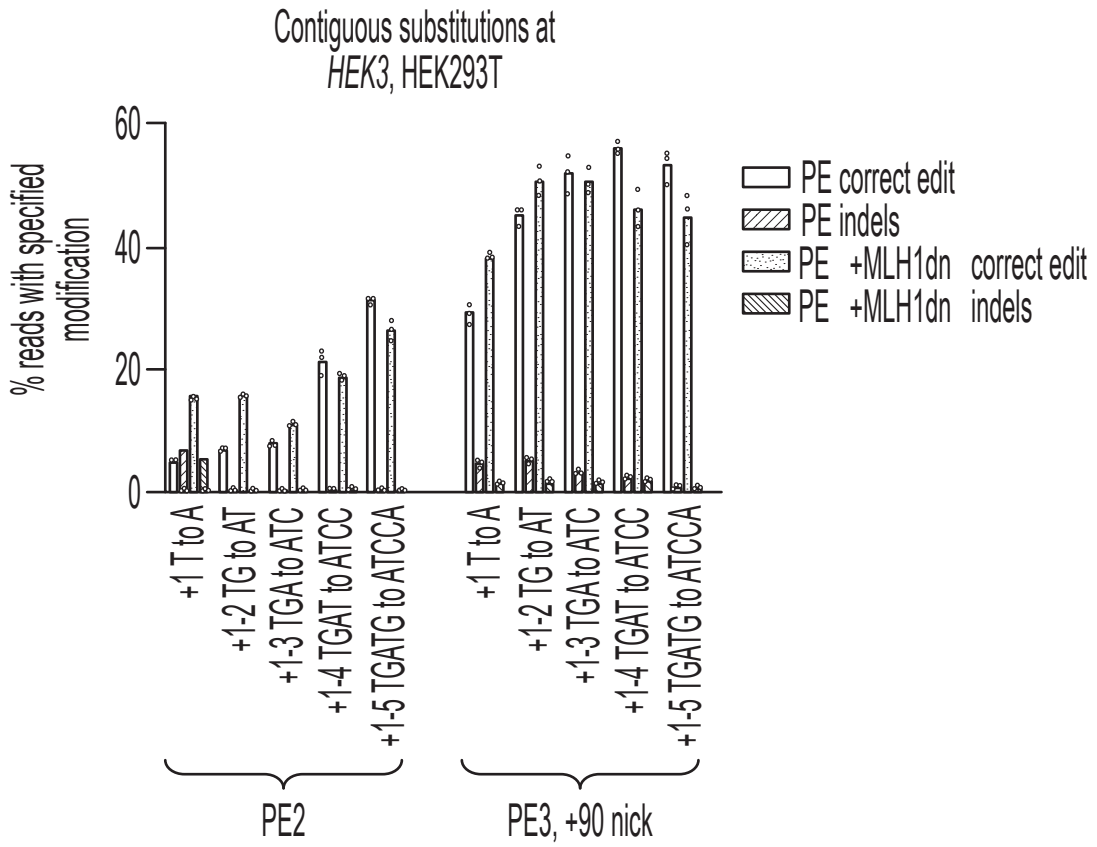


FIG. 37A

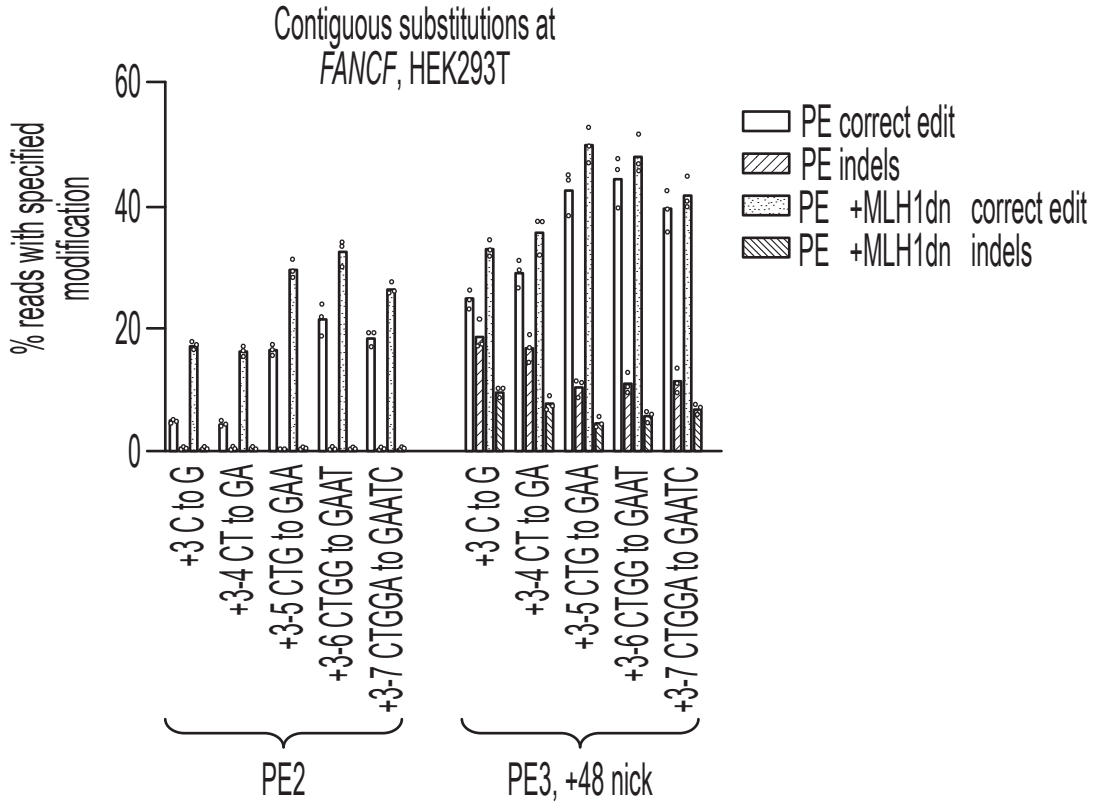


FIG. 37B

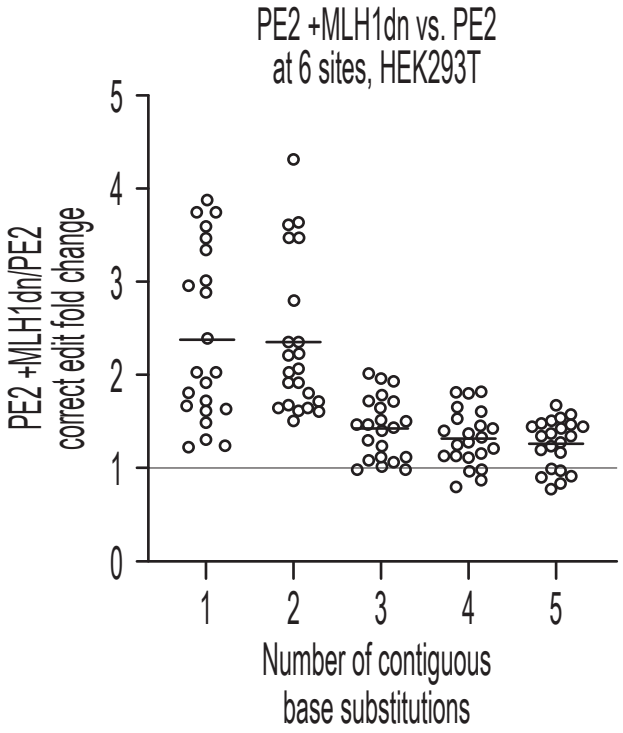


FIG. 38

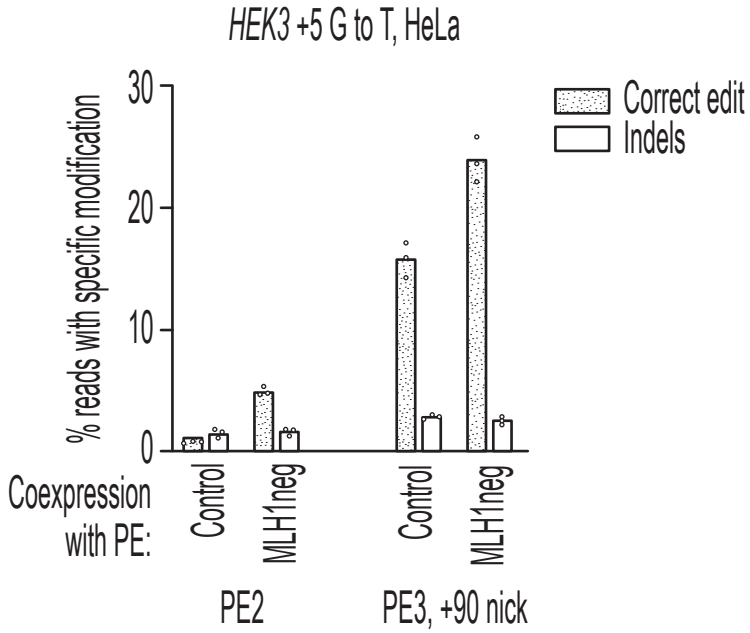


FIG. 39A

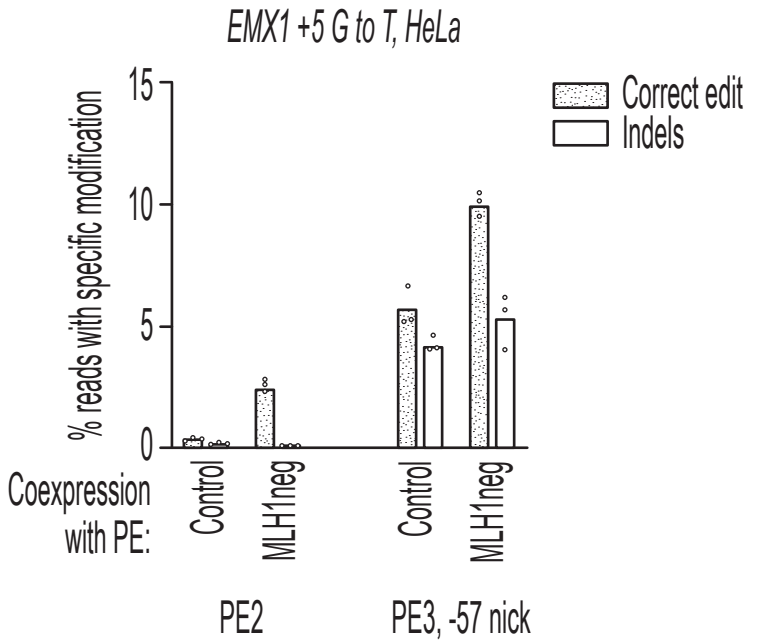


FIG. 39B

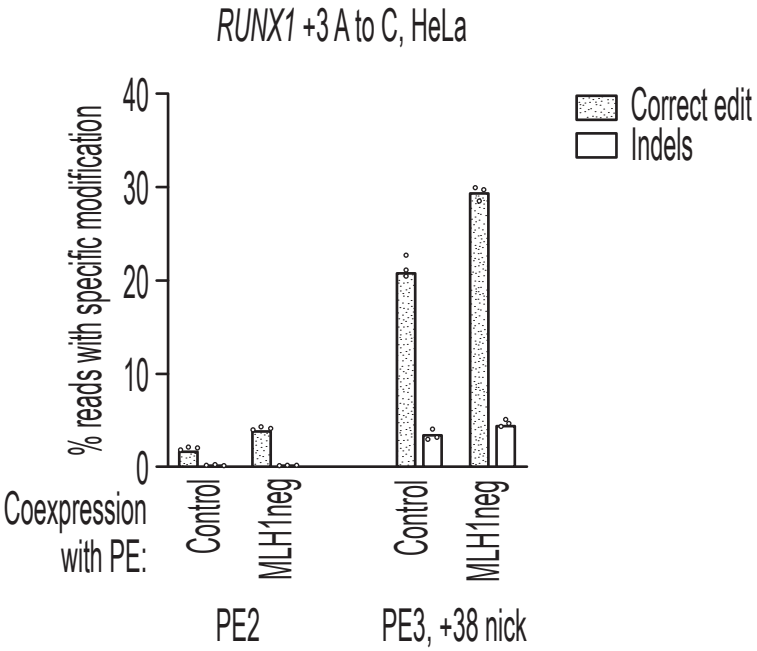


FIG. 39C

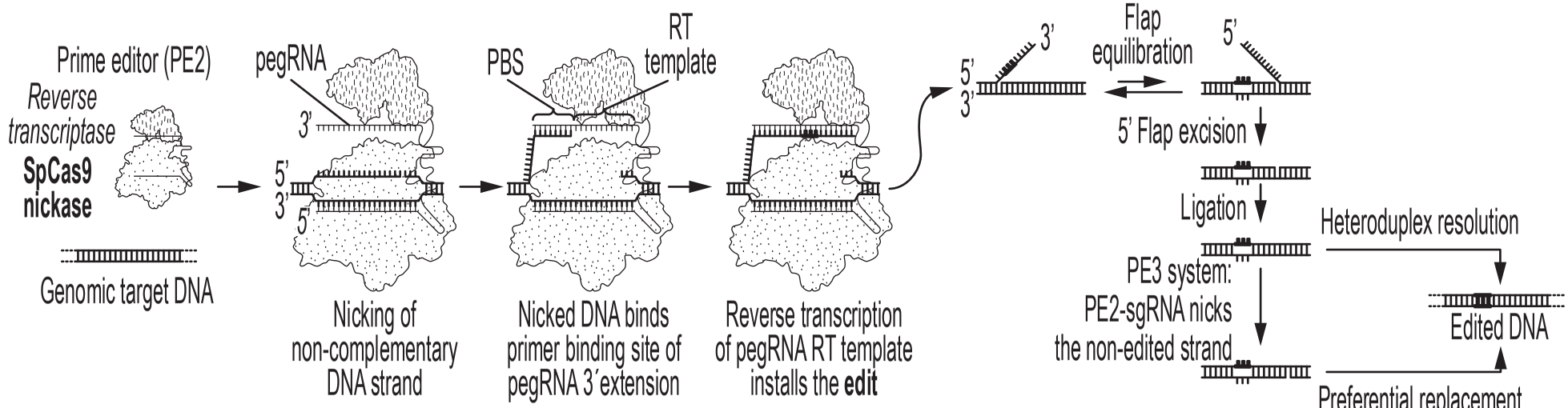


FIG. 40A

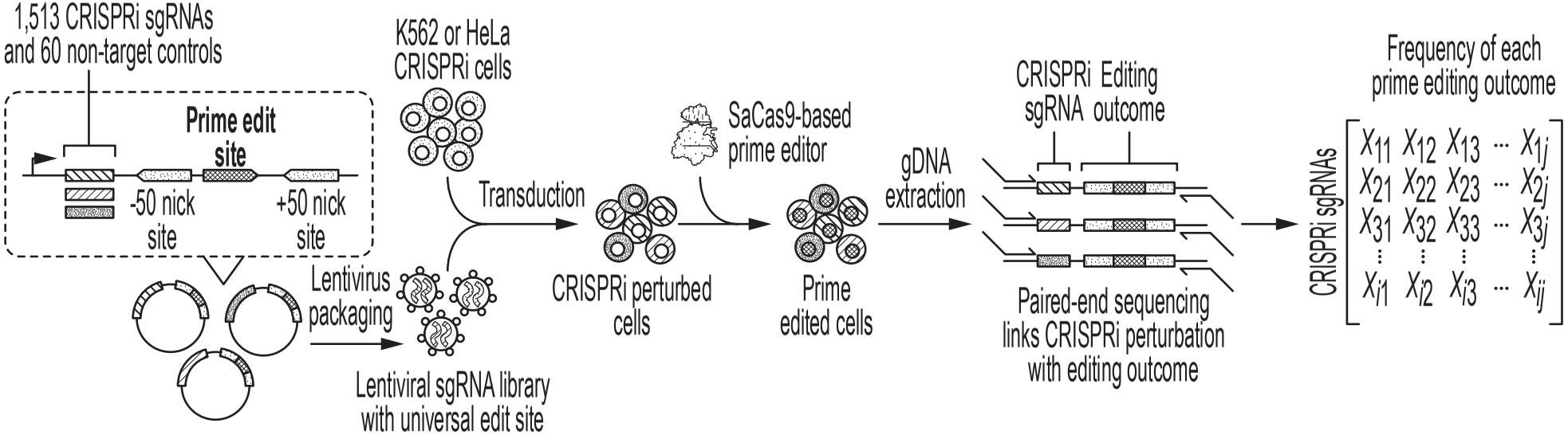


FIG. 40B

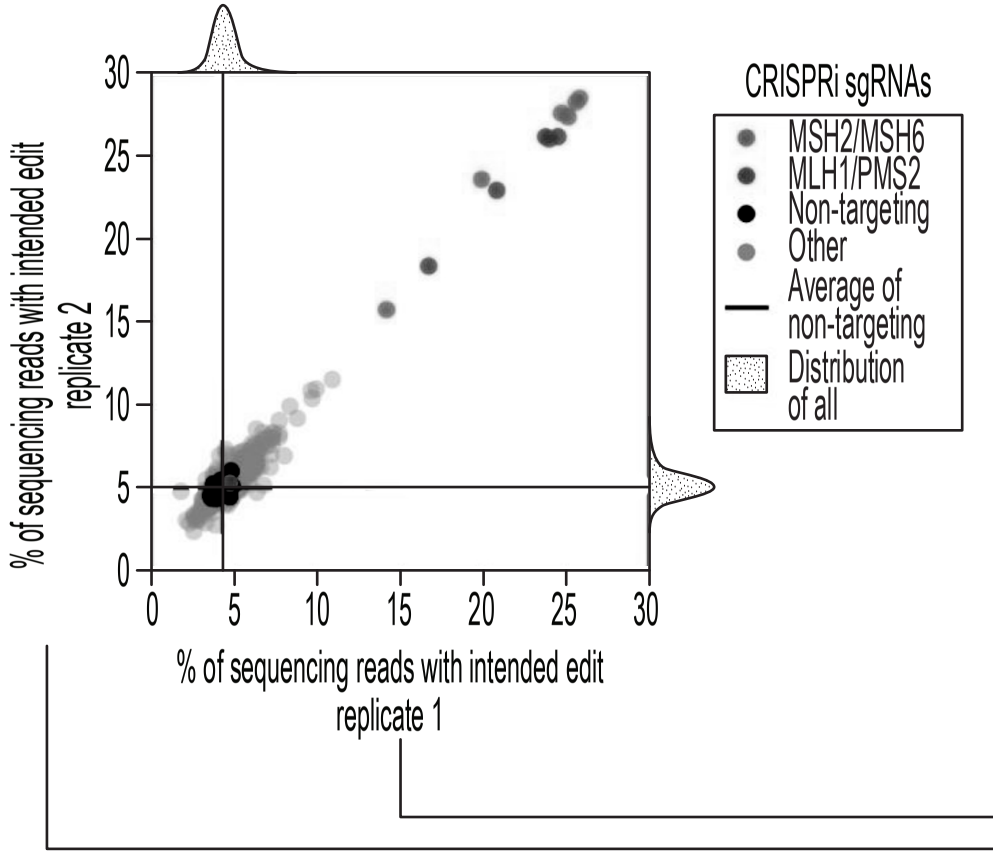


FIG. 40C

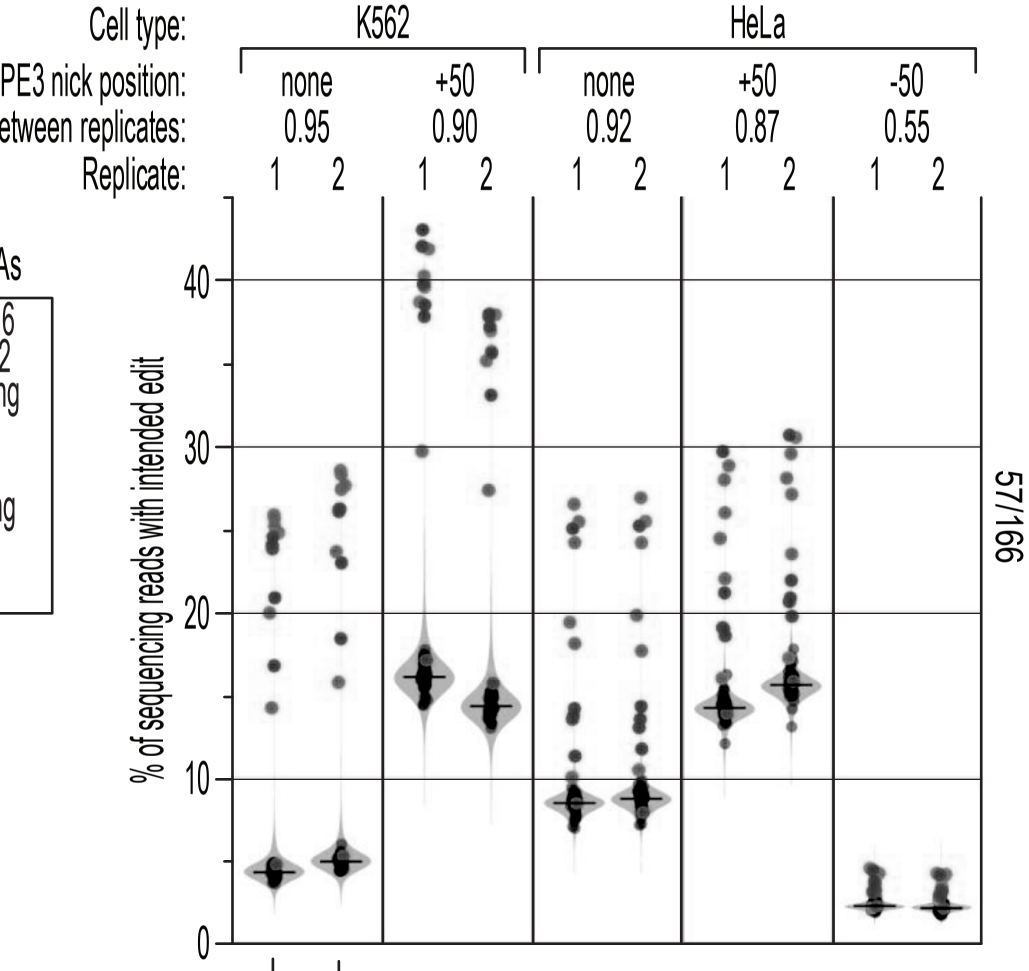


FIG. 40D

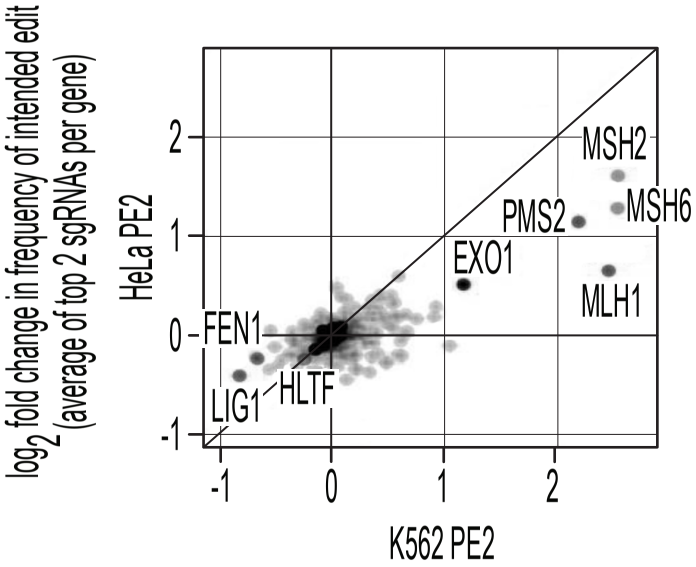


FIG. 40E

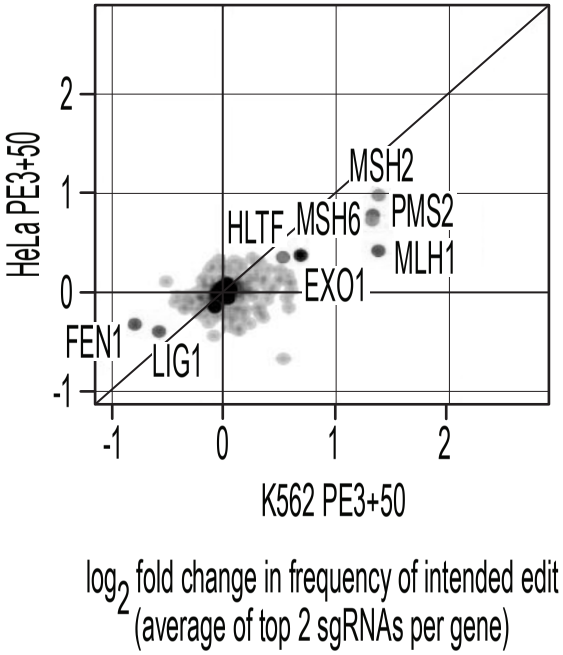


FIG. 40F

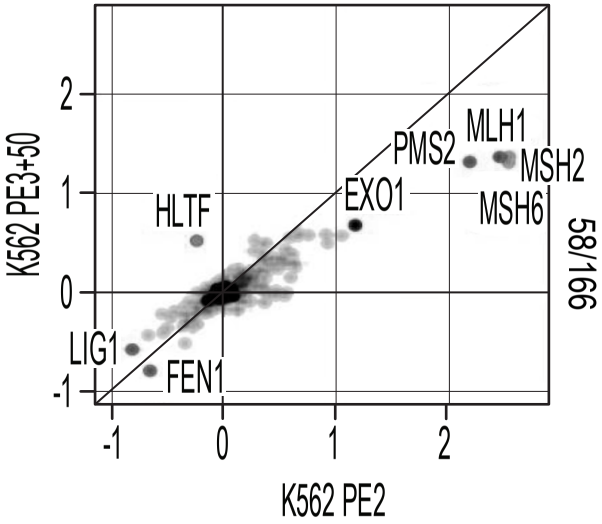


FIG. 40G

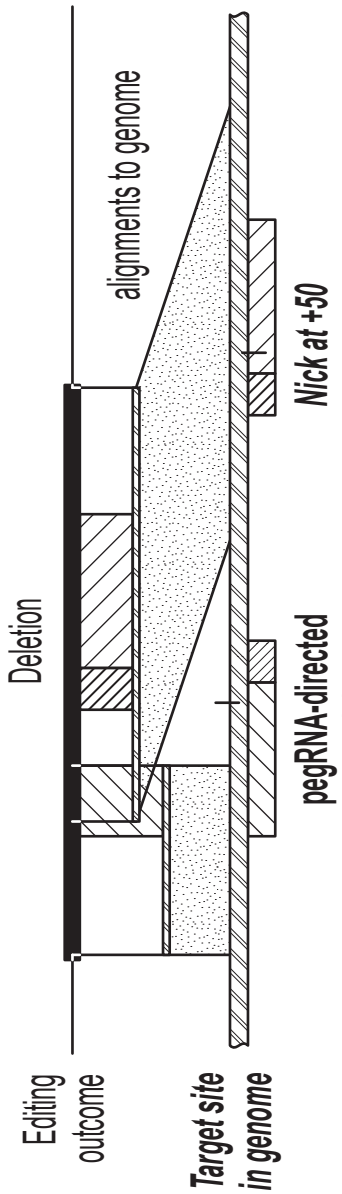


FIG. 41A

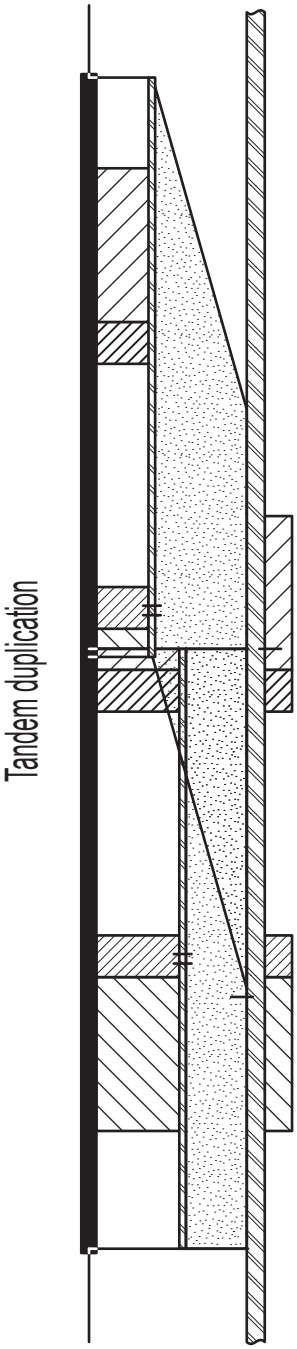


FIG. 41B

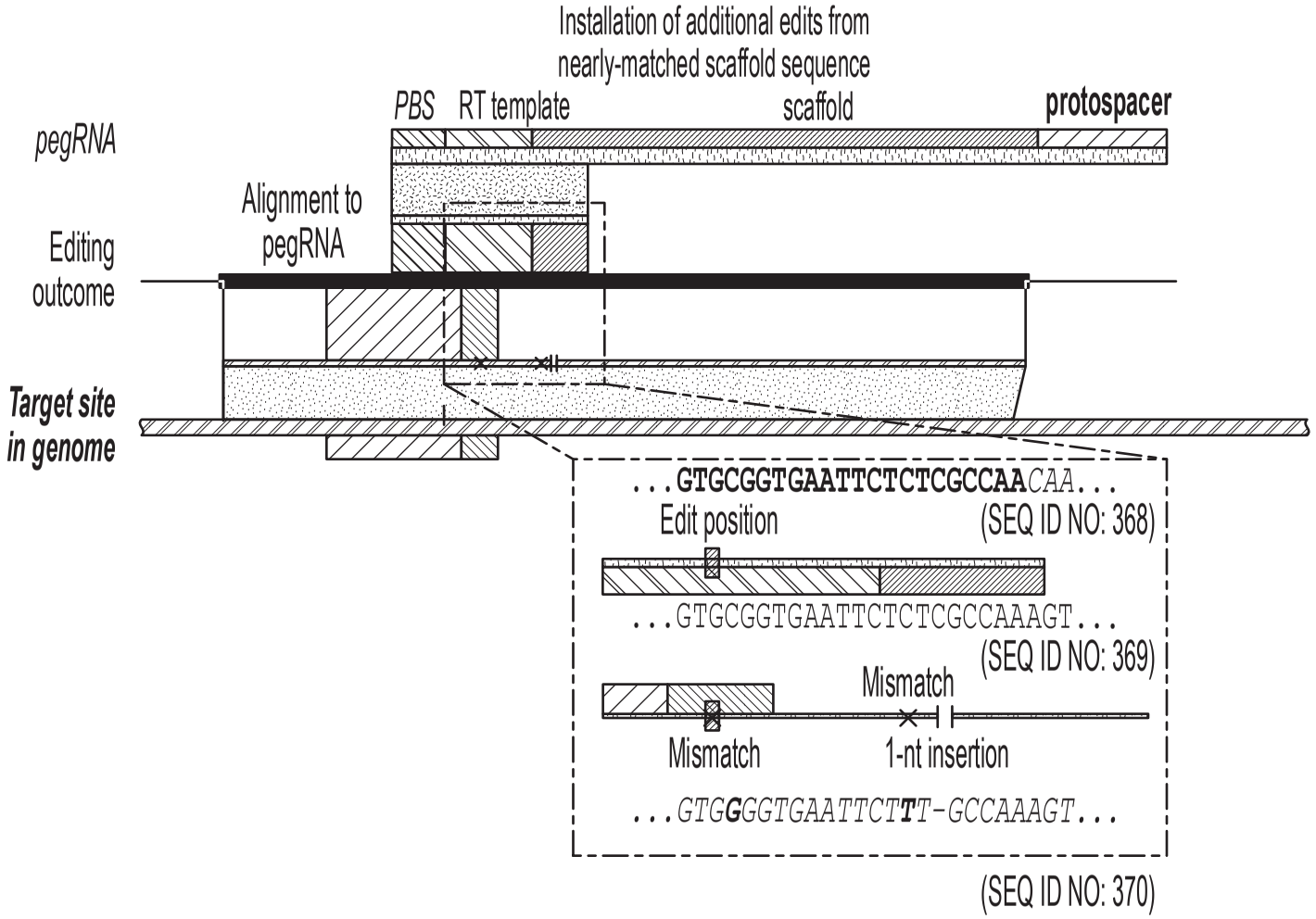

Installation of additional edits from
nearly-matched scaffold sequence
PBS
RT template
scaffold
protospacer
pegRNA
Alignment to
pegRNA
Editing
outcome
Target site
in genome
...GTGCGGTGAATTCTCTCGCCAACAA...
Edit position
(SEQ ID NO: 368)
...GTGCGGTGAATTCTCTCGCCAAAGT...
(SEQ ID NO: 369)
Mismatch
Mismatch
1-nt insertion
...GTGGGGTGAATTCTTT-GCCAAAGT...
(SEQ ID NO: 370)

FIG. 41C

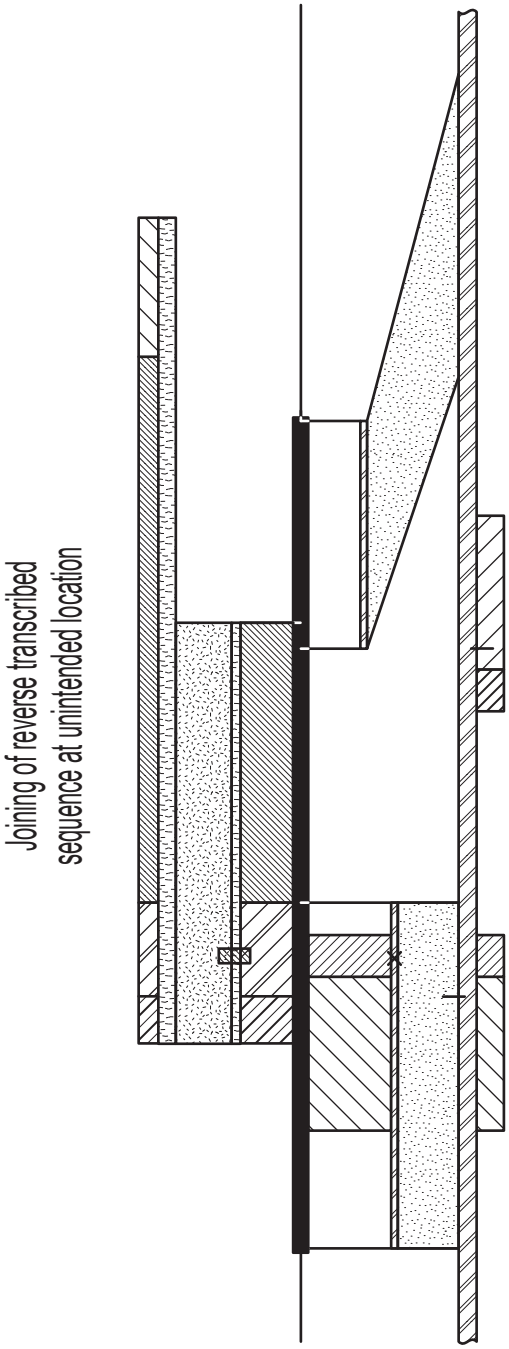

Joining of reverse transcribed
sequence at unintended location

FIG. 41D

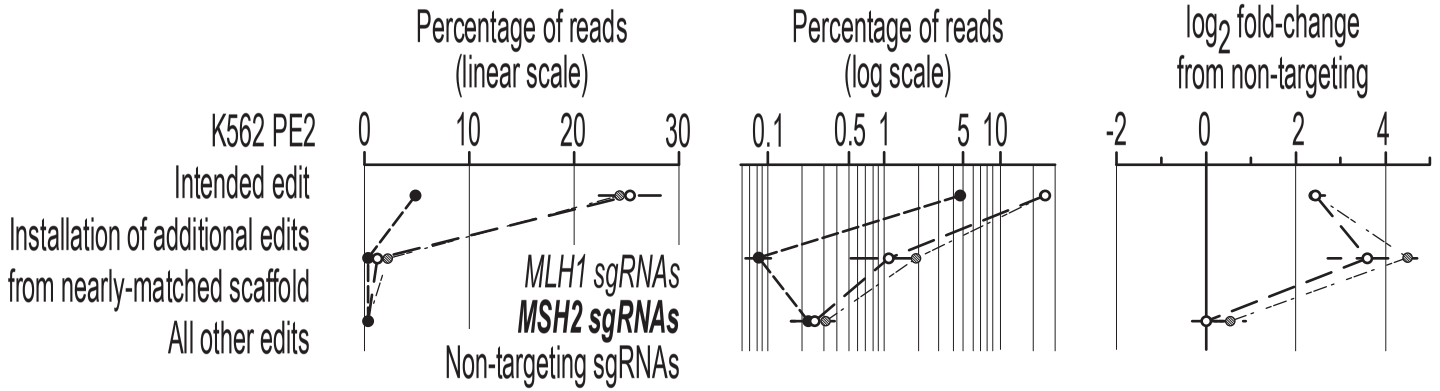


FIG. 41E

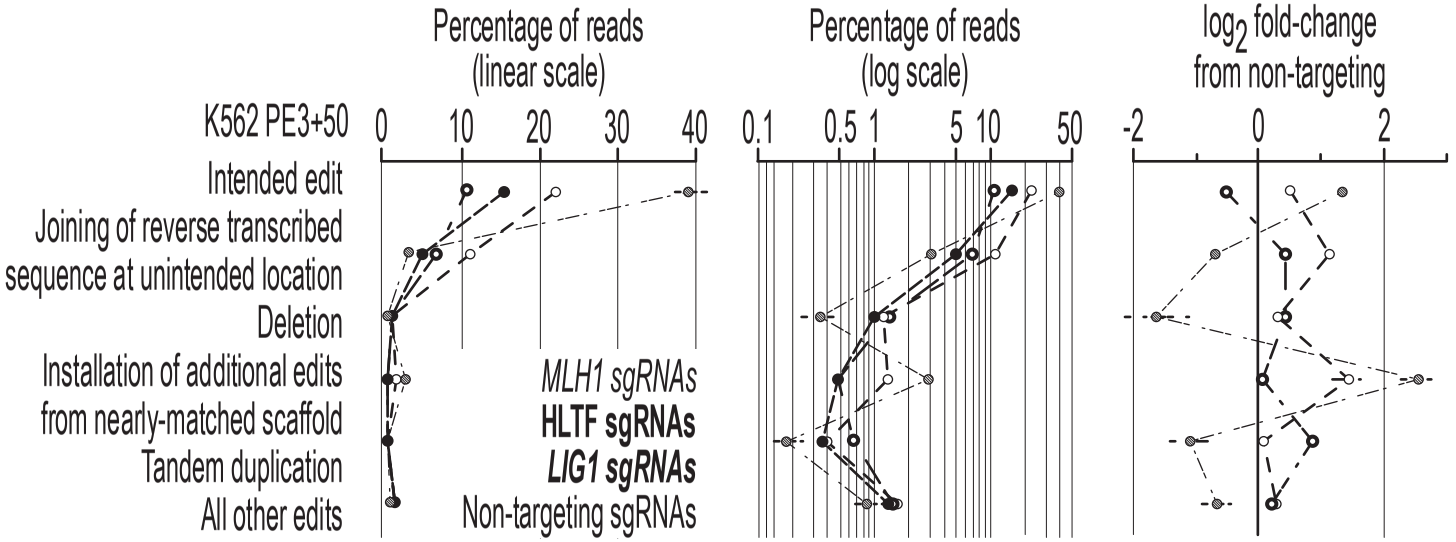


FIG. 41F

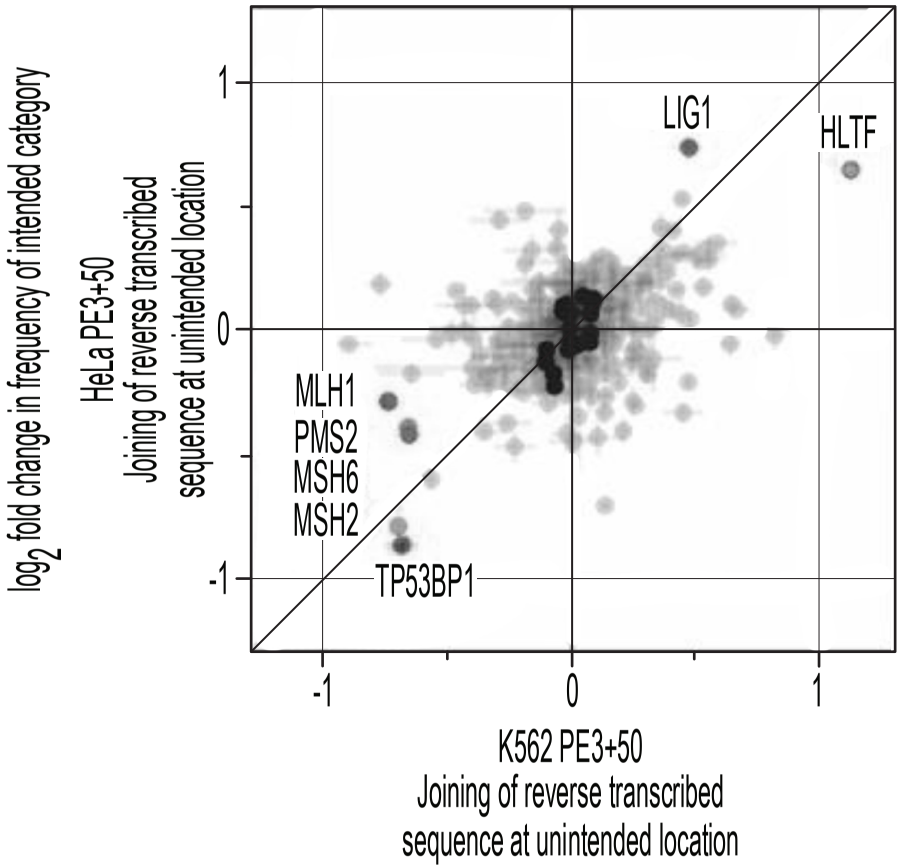


FIG. 41G

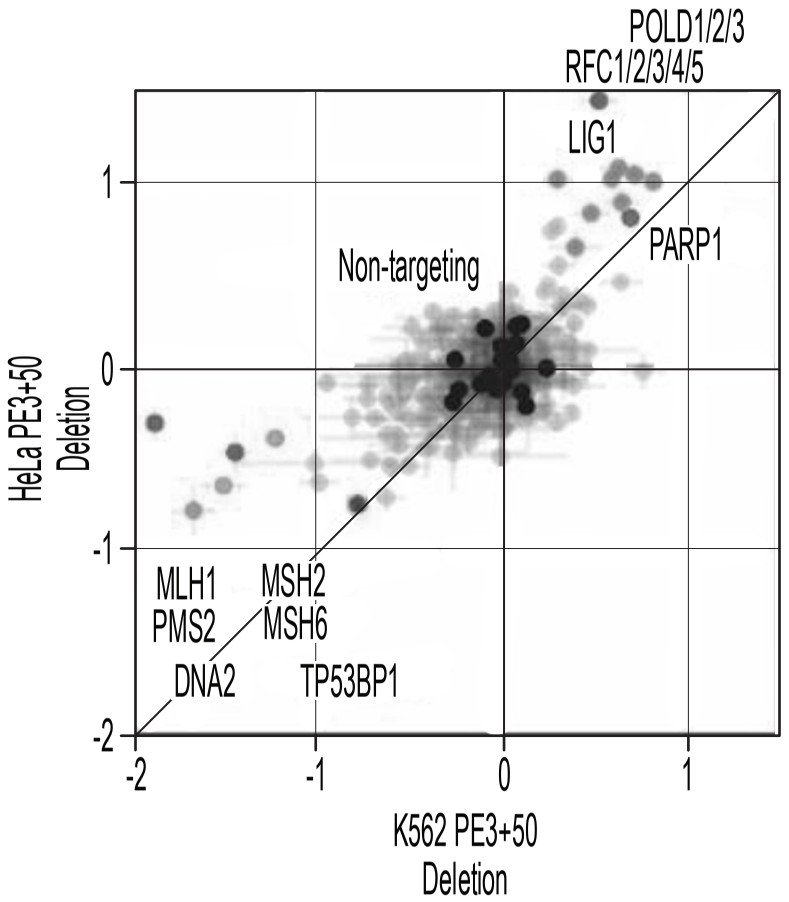


FIG. 41H

Log_2 fold change in frequency of indicated category

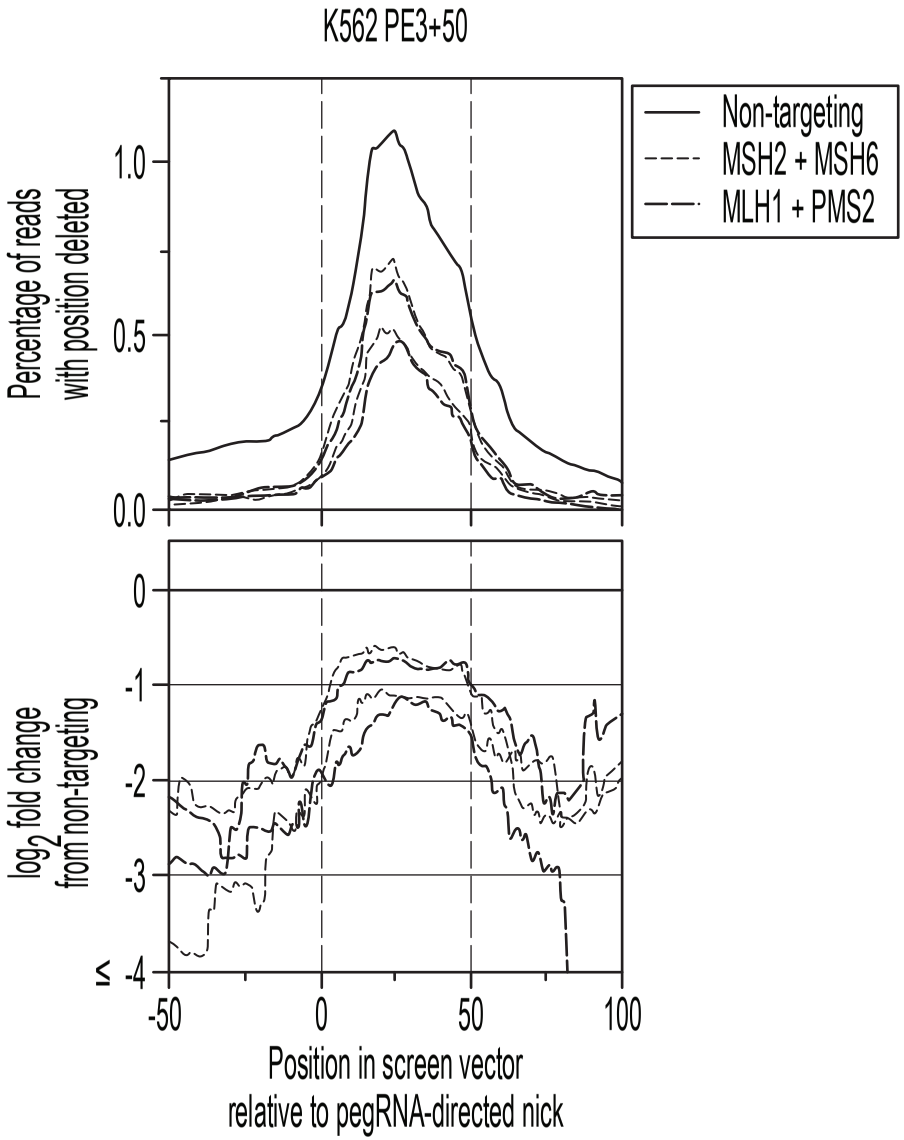


FIG. 41I

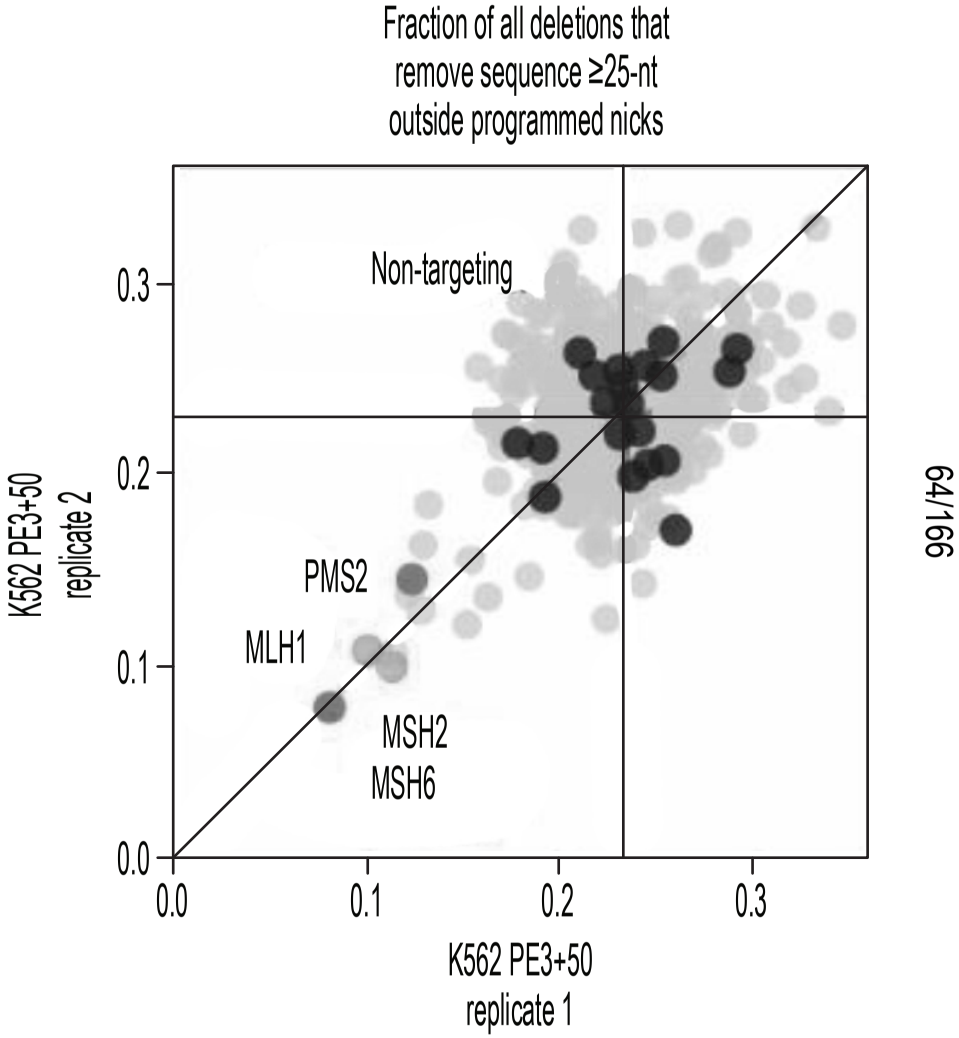


FIG. 41J

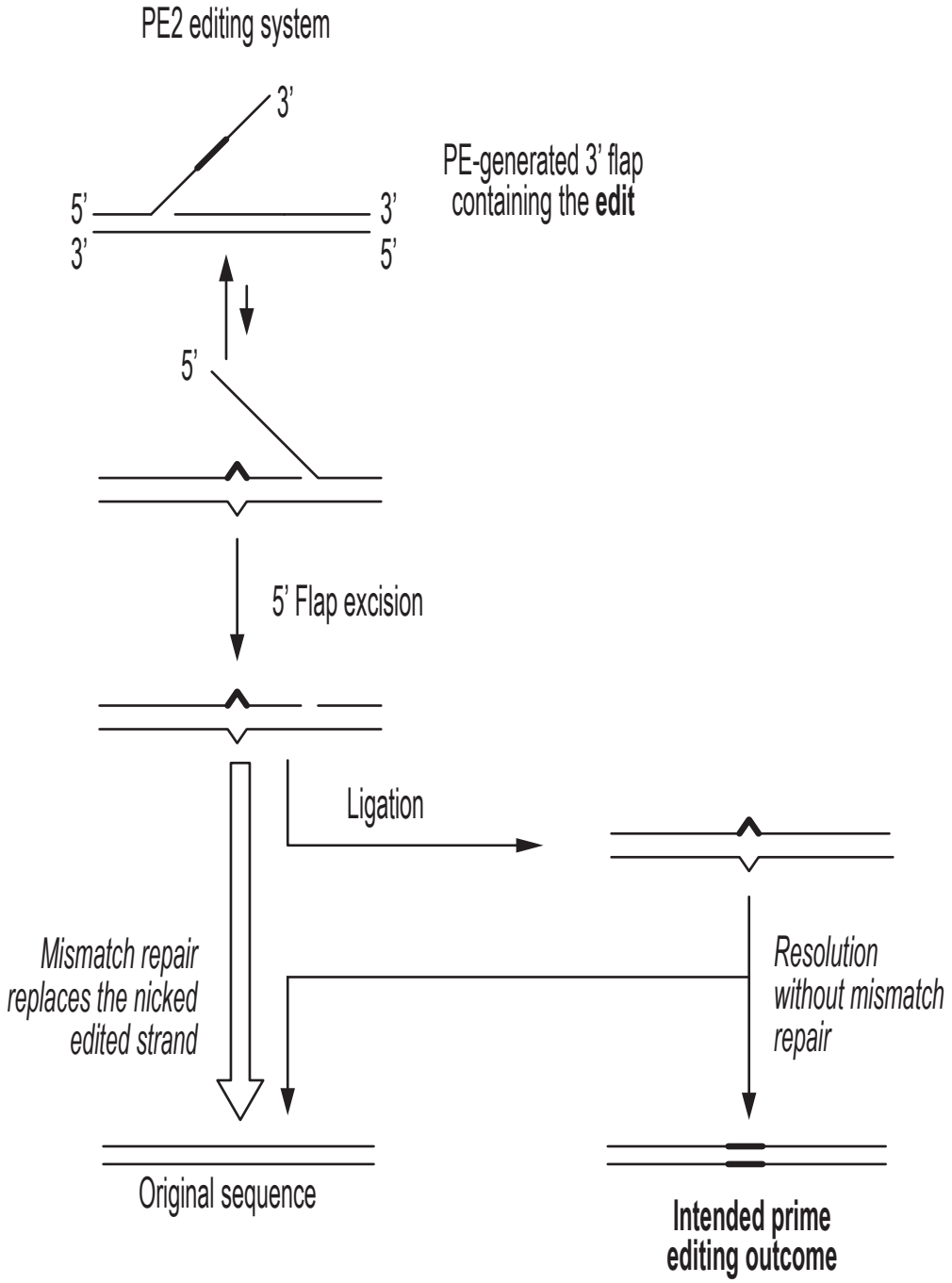

PE2 editing system
3'
5'
3'
3'
5'
PE-generated 3' flap
containing the **edit**
5'
5' Flap excision
Ligation
Mismatch repair replaces the nicked edited strand
Resolution without mismatch repair
Original sequence
Intended prime editing outcome

FIG. 42A

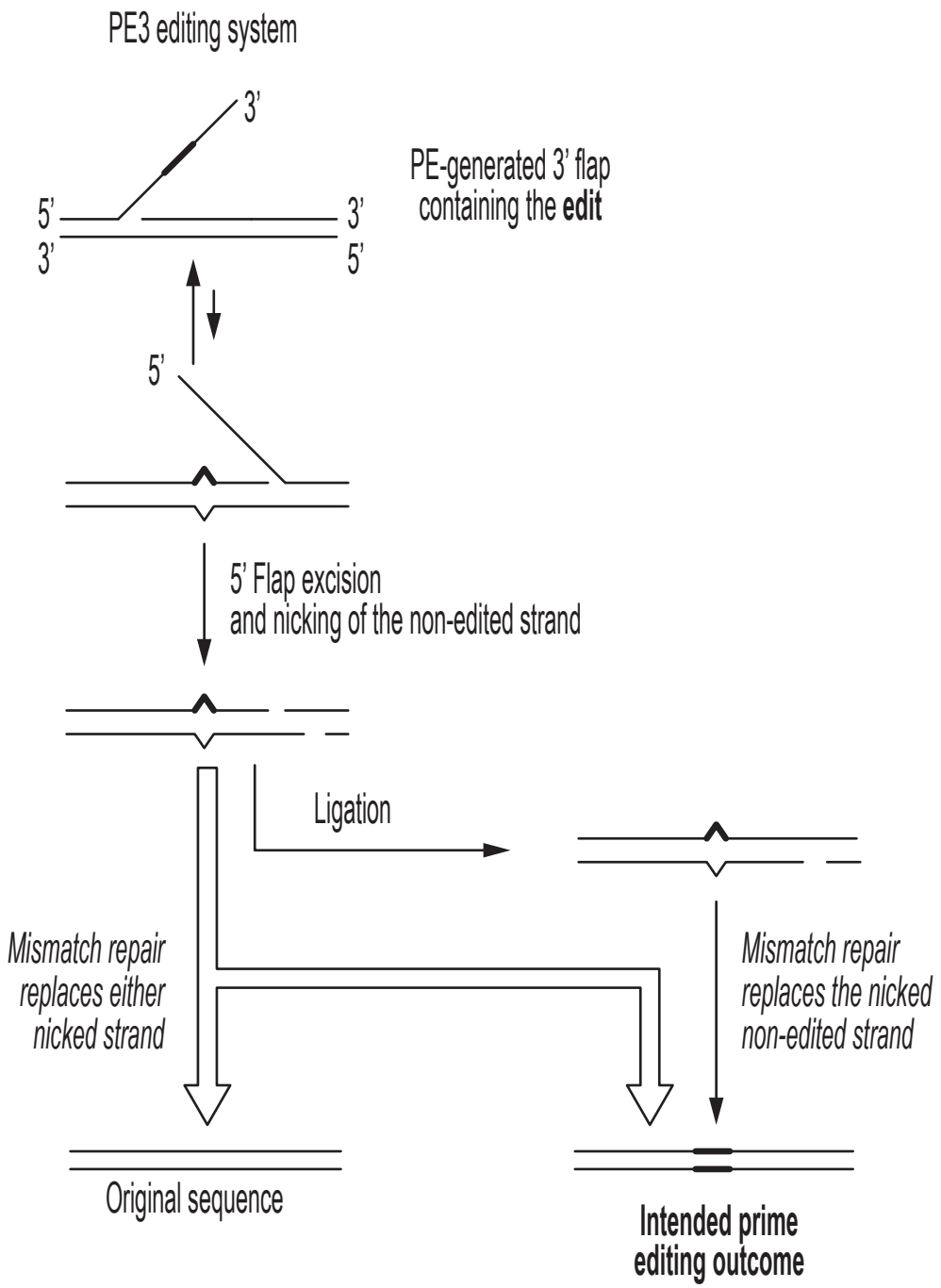

PE3 editing system
3'
5'
3'
3'
5'
PE-generated 3' flap
containing the **edit**
5'
5' Flap excision
and nicking of the non-edited strand
Ligation
Mismatch repair replaces either nicked strand
Mismatch repair replaces the nicked non-edited strand
Original sequence
Intended prime editing outcome

FIG. 42B

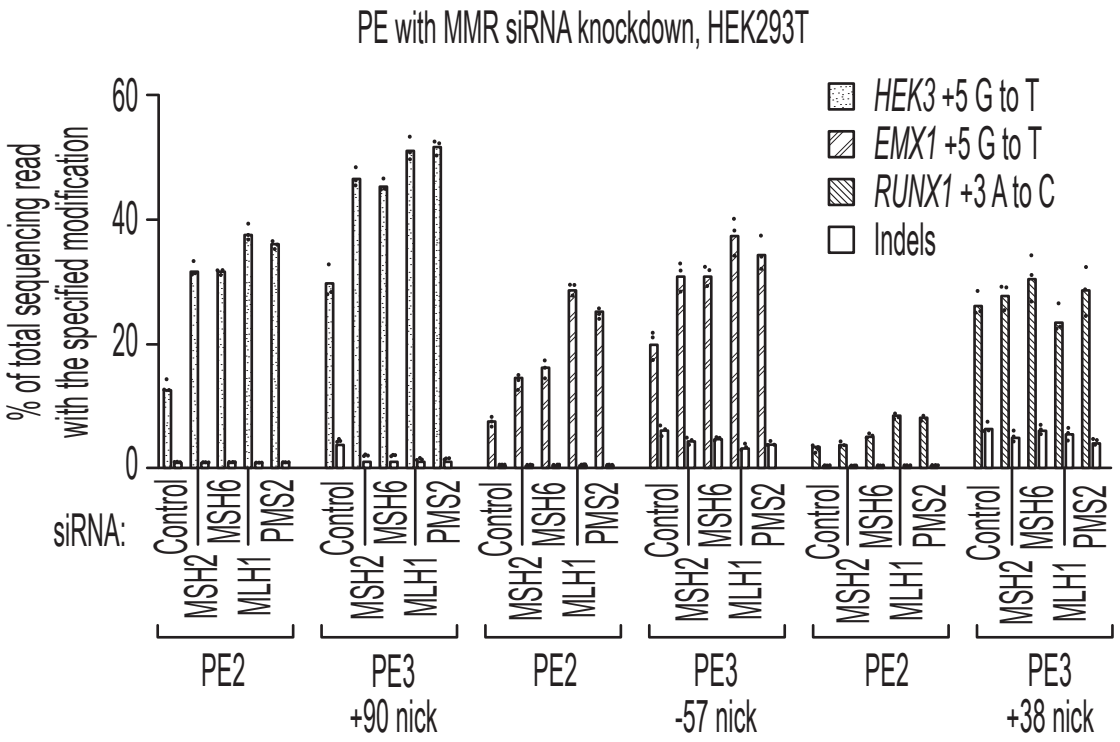


FIG. 42C

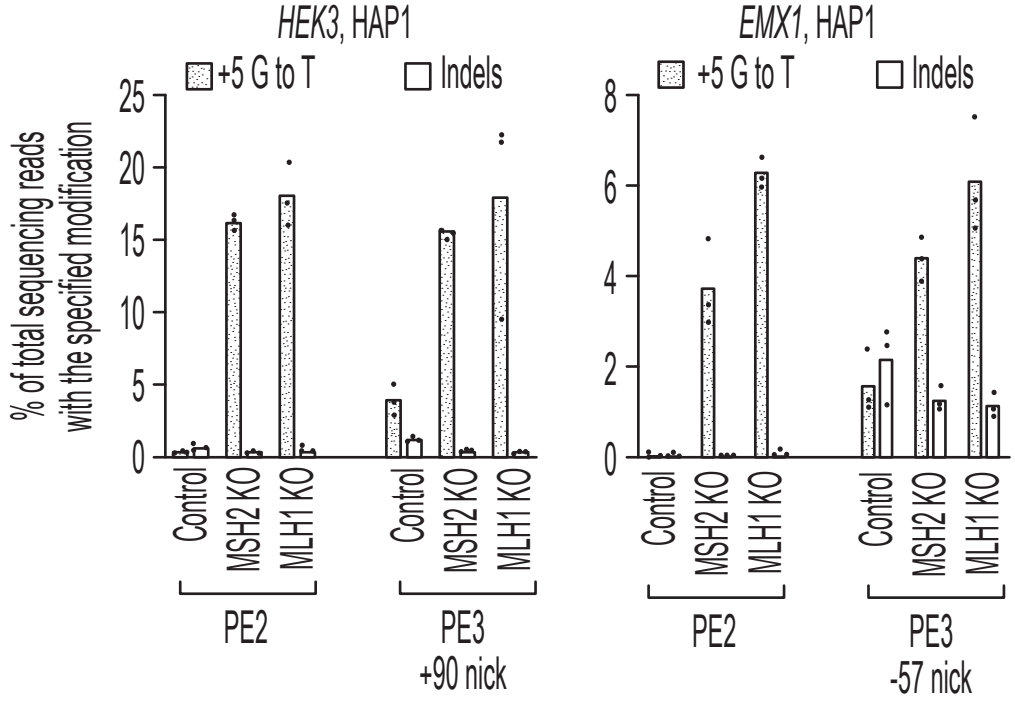


FIG. 42D

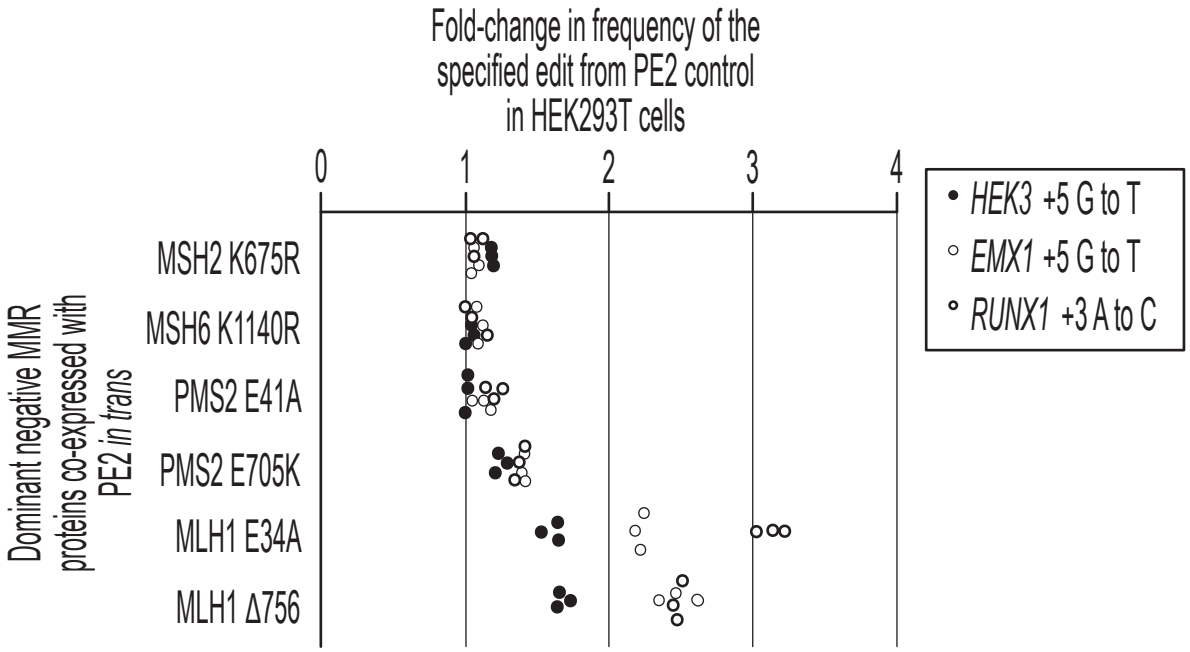


FIG. 43A

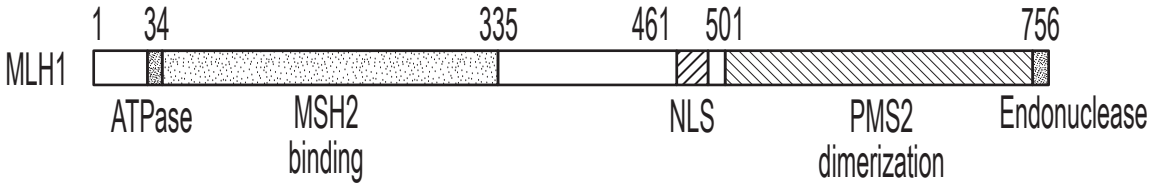


FIG. 43B

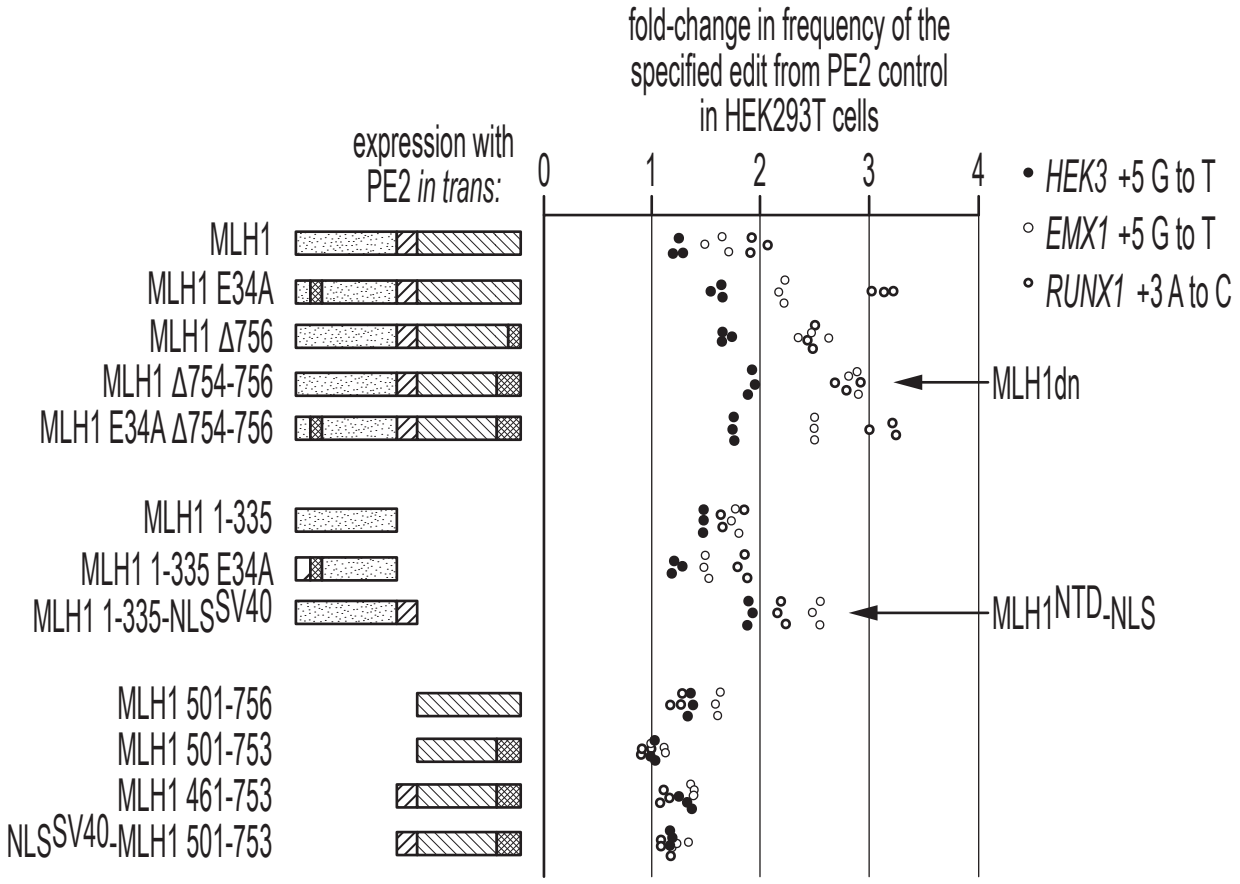


FIG. 43C

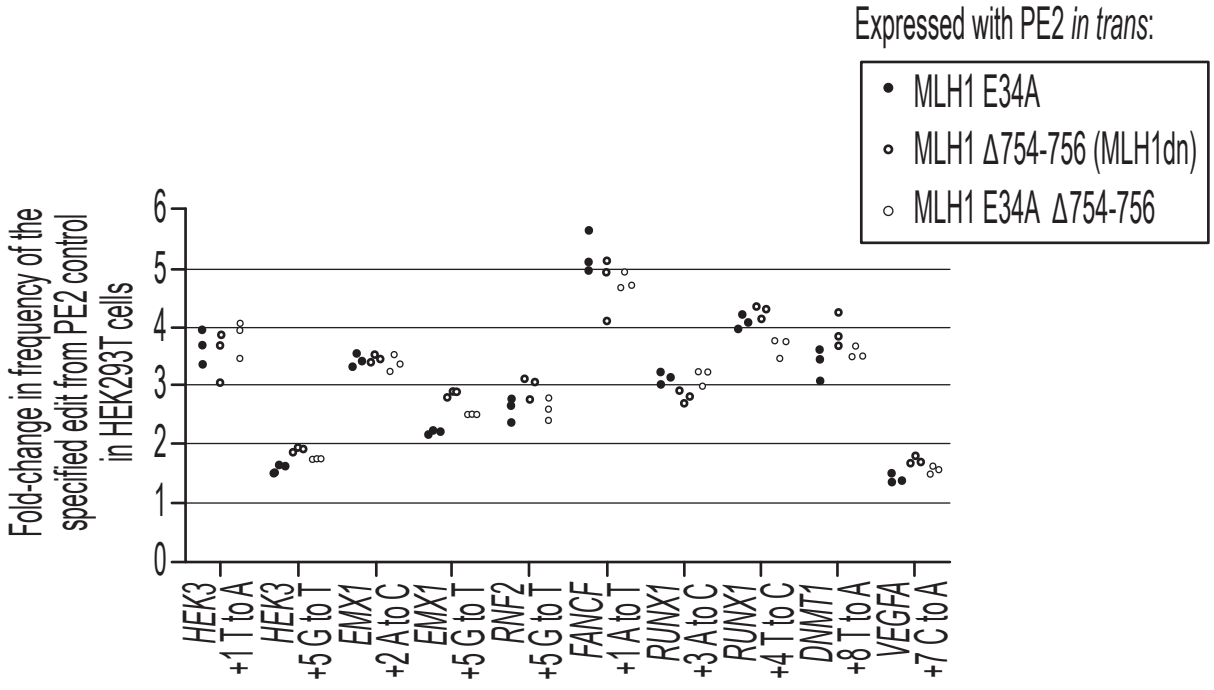


FIG. 43D

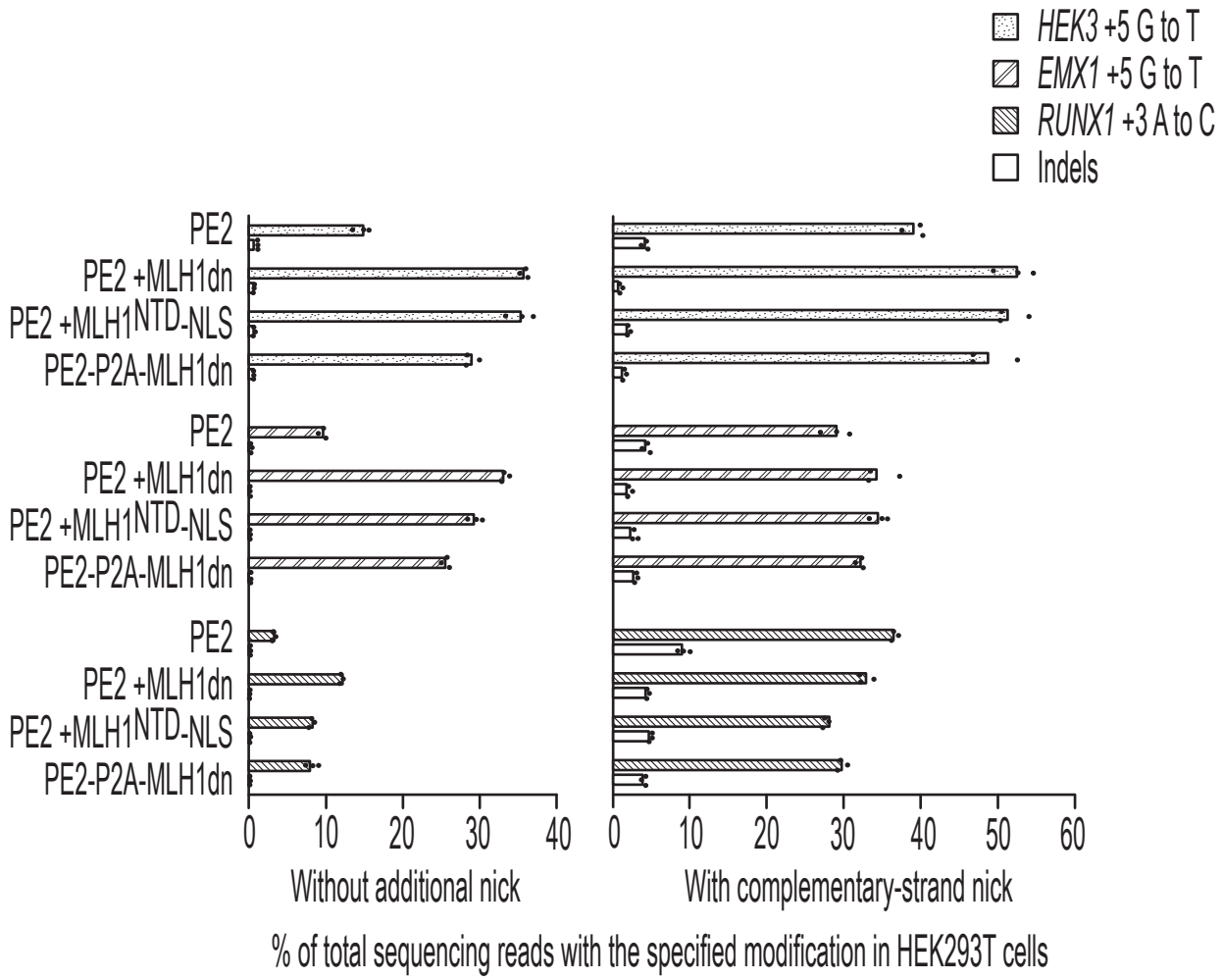

HEK3 +5 G to T
EMX1 +5 G to T
RUNX1 +3 A to C
Indels
PE2
PE2 +MLH1dn
PE2 +MLH1NTD-NLS
PE2-P2A-MLH1dn
PE2
PE2 +MLH1dn
PE2 +MLH1NTD-NLS
PE2-P2A-MLH1dn
PE2
PE2 +MLH1dn
PE2 +MLH1NTD-NLS
PE2-P2A-MLH1dn
0
10
20
30
40
0
10
20
30
40
50
60
Without additional nick
With complementary-strand nick
% of total sequencing reads with the specified modification in HEK293T cells

FIG. 43E

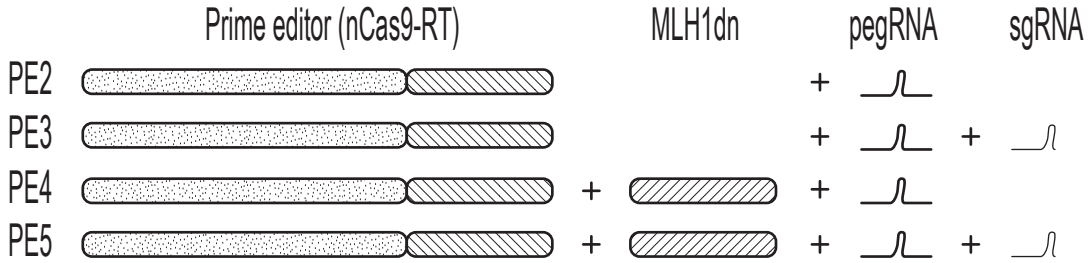


FIG. 43F

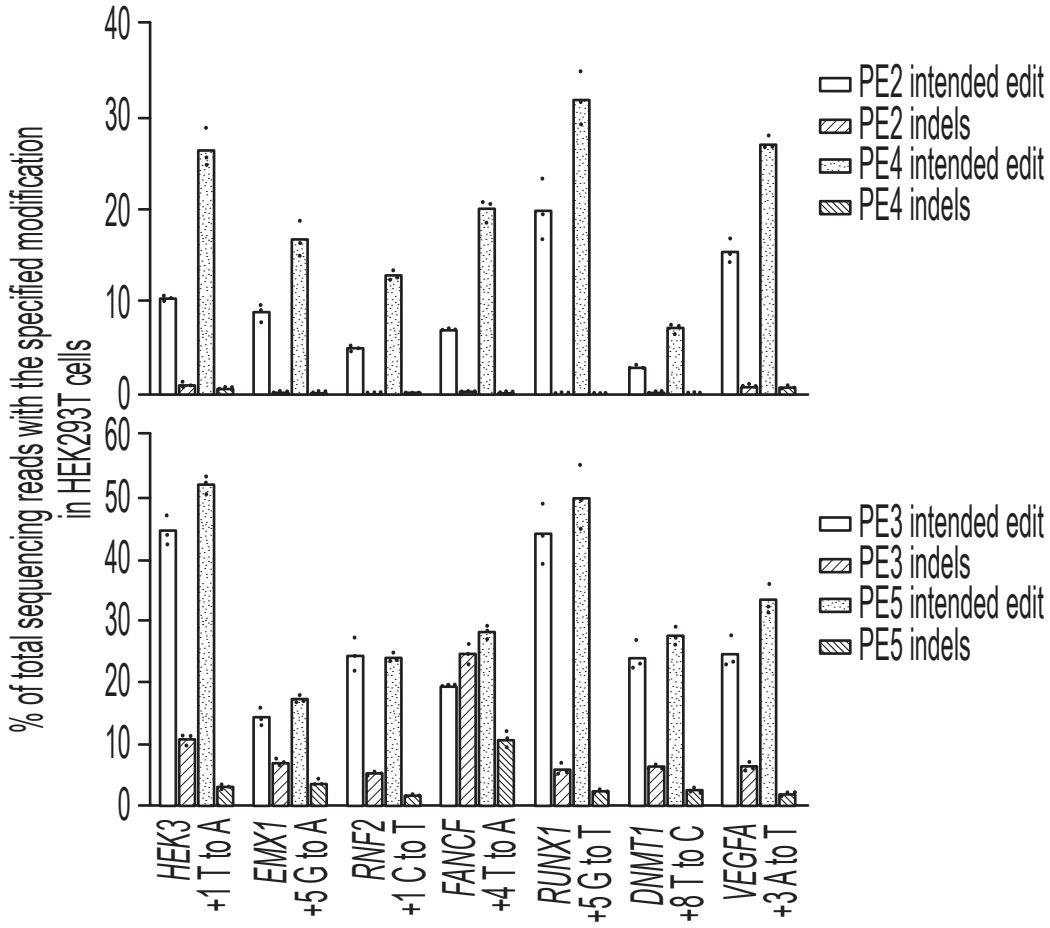


FIG. 43G

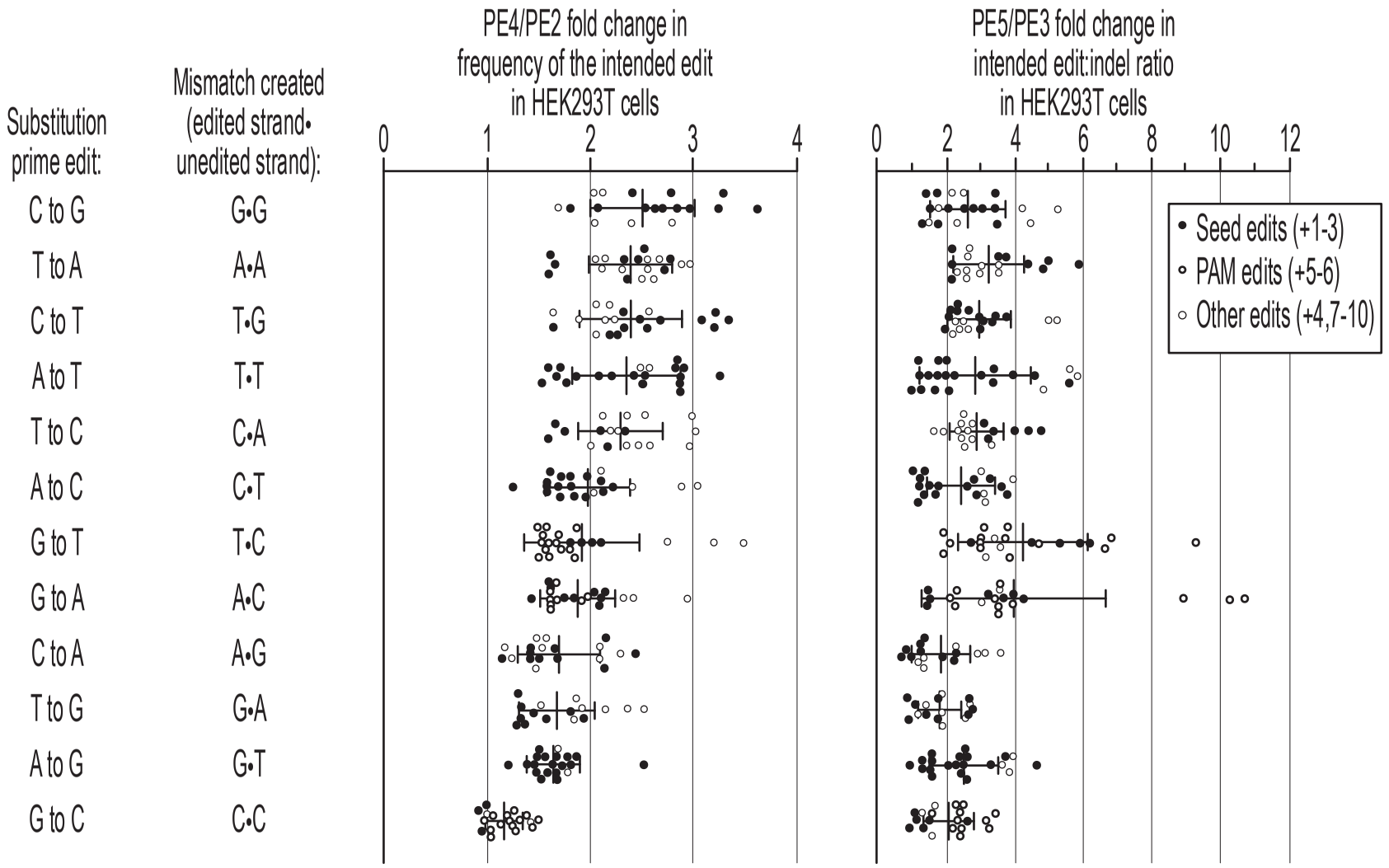

Substitution prime edit:
Mismatch created (edited strand• unedited strand):
C to G
G•G
T to A
A•A
C to T
T•G
A to T
T•T
T to C
C•A
A to C
C•T
G to T
T•C
G to A
A•C
C to A
A•G
T to G
G•A
A to G
G•T
G to C
C•C
PE4/PE2 fold change in frequency of the intended edit in HEK293T cells
0
1
2
3
4
PE5/PE3 fold change in intended edit:indel ratio in HEK293T cells
0
2
4
6
8
10
12
Seed edits (+1-3)
PAM edits (+5-6)
Other edits (+4,7-10)

FIG. 44A

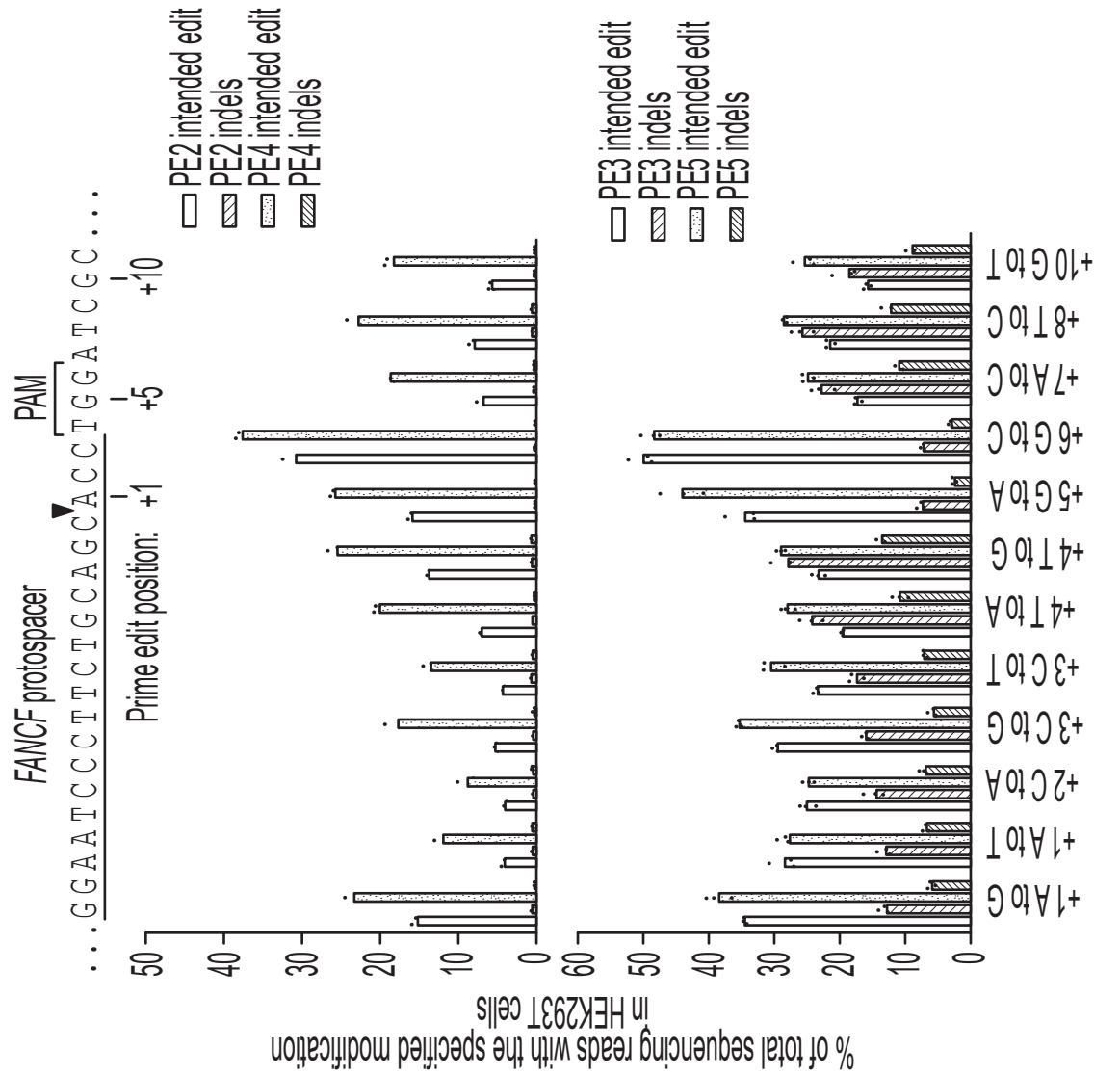

FANCF protospacer
PAM
...GGAATCCCTTCTGCAGCACCTGGATCGGC...
Prime edit position: +1 +5 +10
PE2 intended edit
PE2 indels
PE4 intended edit
PE4 indels
PE3 intended edit
PE3 indels
PE5 intended edit
PE5 indels
% of total sequencing reads with the specified modification in HEK293T cells
+1A to G
+1A to T
+2C to A
+3C to G
+3C to T
+4T to A
+4T to G
+5G to A
+6G to C
+7A to C
+8T to C
+10G to T

FIG. 44B

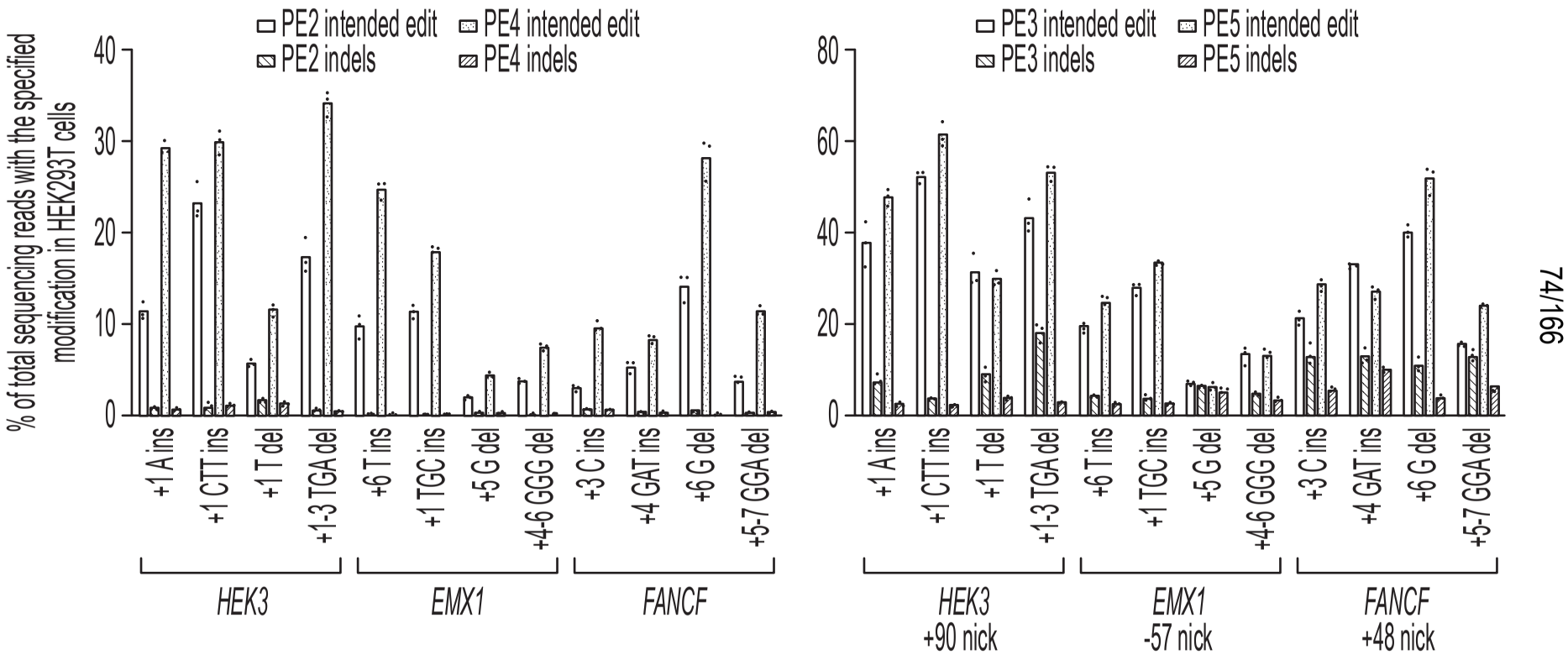

% of total sequencing reads with the specified modification in HEK293T cells
PE2 intended edit
PE2 indels
PE4 intended edit
PE4 indels
PE3 intended edit
PE3 indels
PE5 intended edit
PE5 indels
+1 A ins
+1 CTT ins
+1 T del
+1-3 TGA del
+6 T ins
+1 TGC ins
+5 G del
+4-6 GGG del
+3 C ins
+4 GAT ins
+6 G del
+5-7 GGA del
HEK3
EMX1
FANCF
HEK3 +90 nick
EMX1 -57 nick
FANCF +48 nick

FIG. 44C

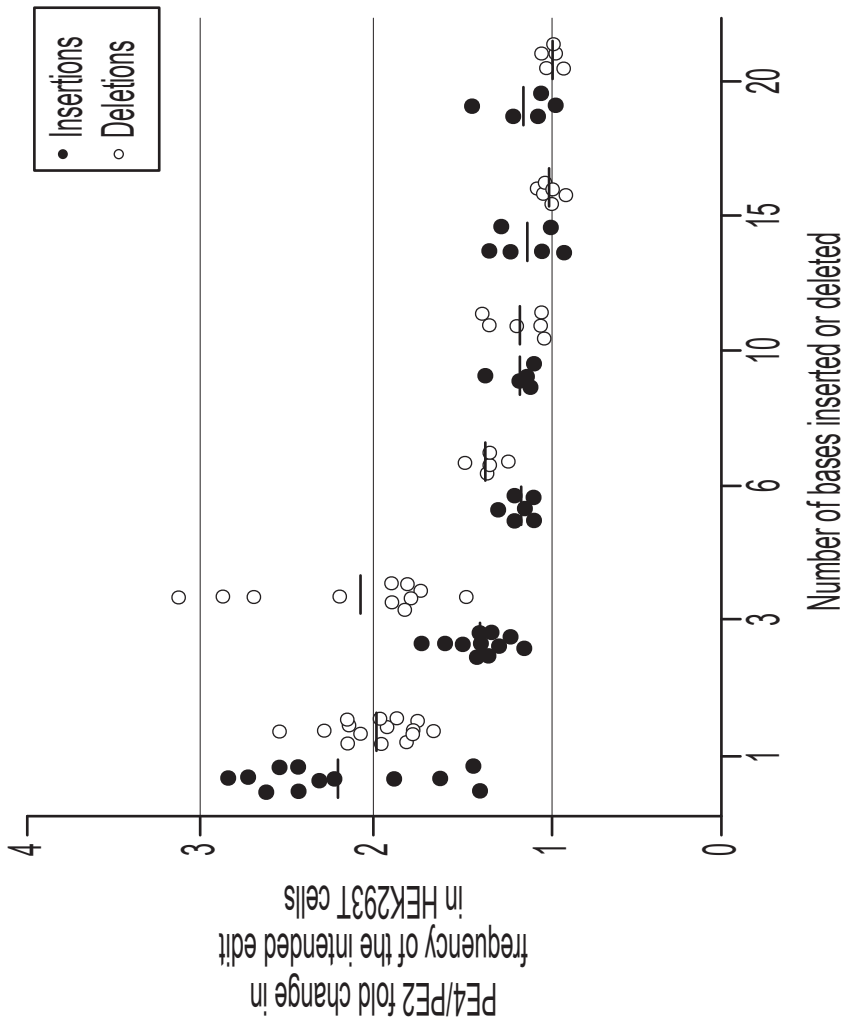

• Insertions
◦ Deletions
PE4/PE2 fold change in
frequency of the intended edit
in HEK293T cells
0
1
2
3
4
1
3
6
10
15
20
Number of bases inserted or deleted
in prime edit

FIG. 44D

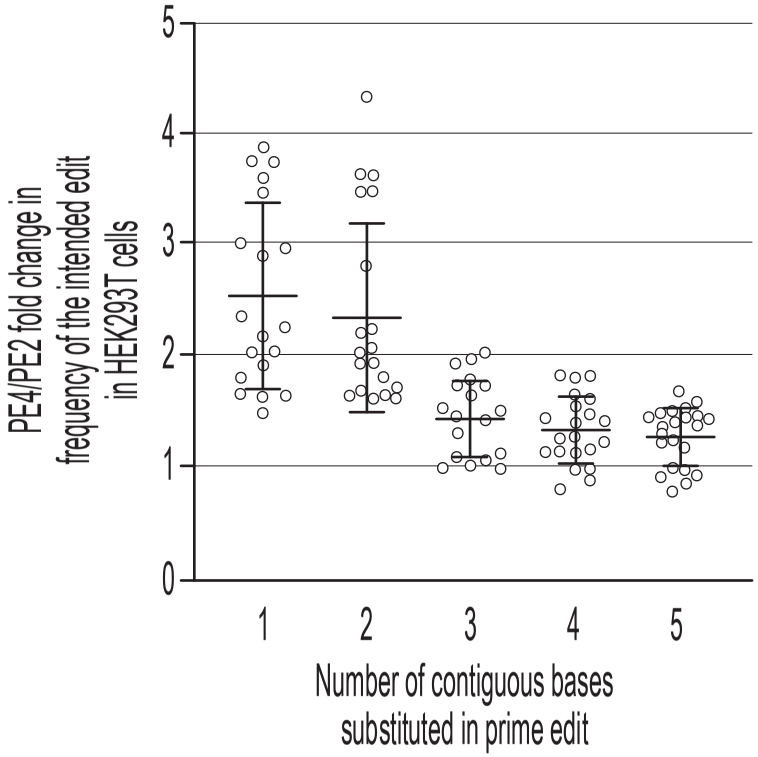


FIG. 44E

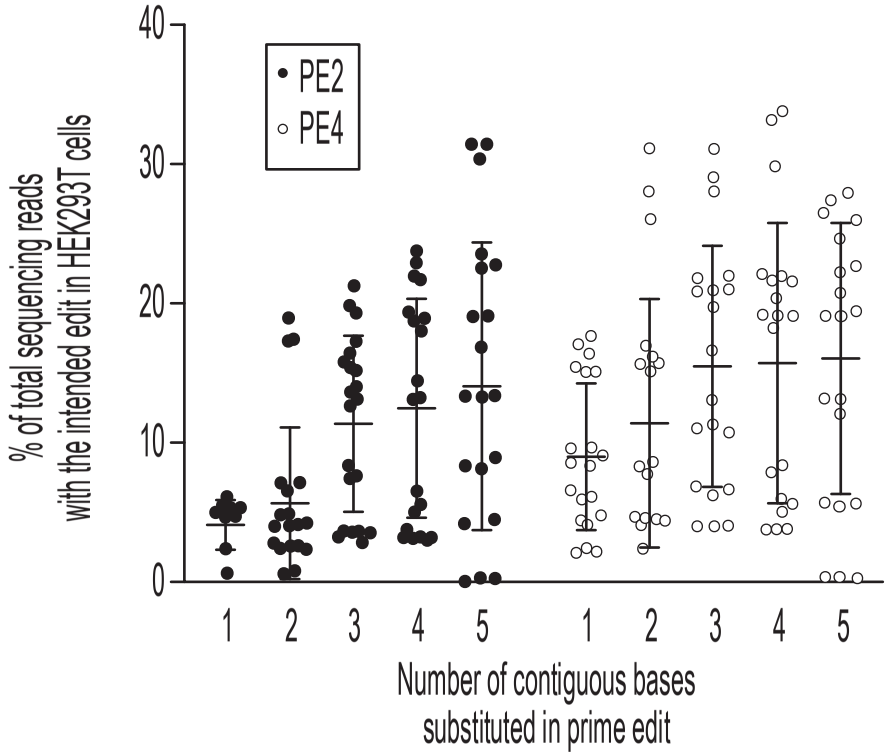


FIG. 44F

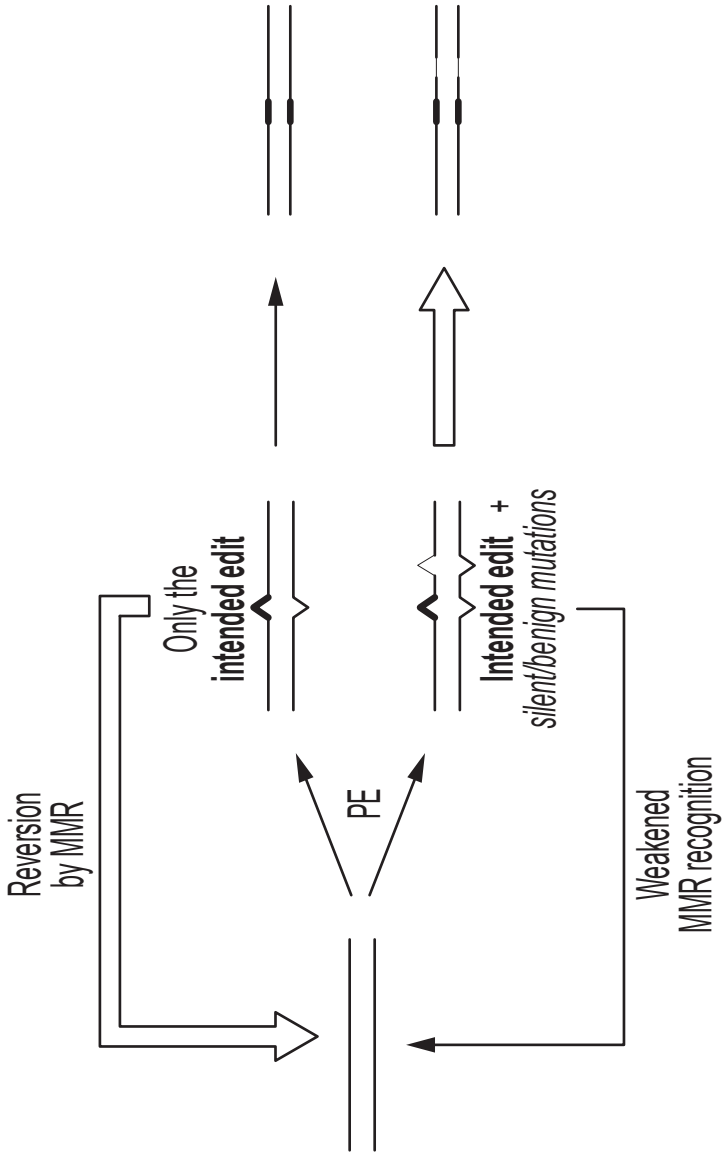

Only the
intended edit
Intended edit +
silent/benign mutations
PE
Reversion
by MMR
Weakened
MMR recognition

FIG. 44G

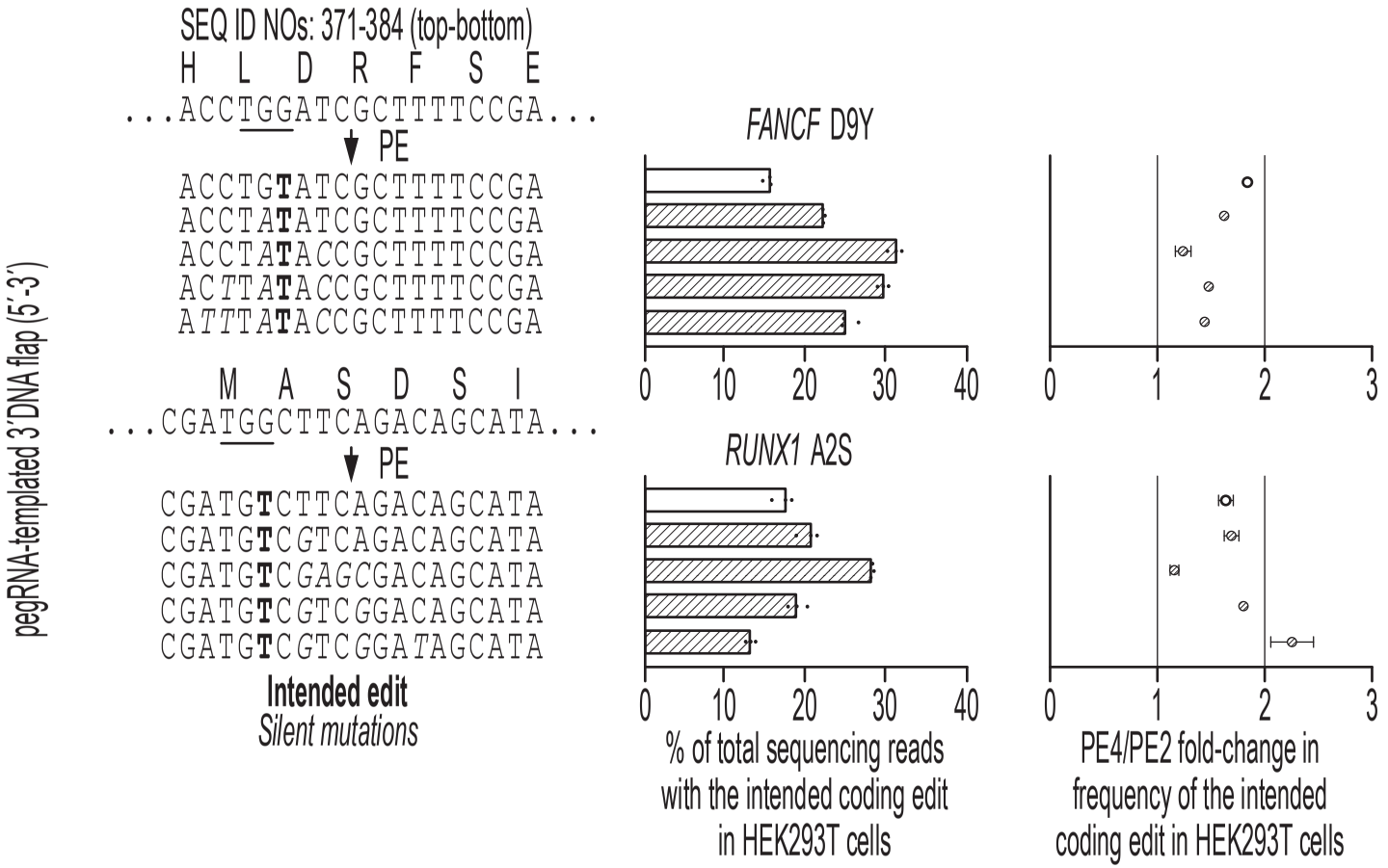

SEQ ID NOs: 371-384 (top-bottom)
H L D R F S E
...ACCTGGATCGCTTTTCCGA...
PE
ACCTGTATCGCTTTTCCGA
ACCTATATCGCTTTTCCGA
ACCTATACCGCTTTTCCGA
ACTTATACCGCTTTTCCGA
ATTTATACCGCTTTTCCGA
M A S D S I
...CGATGGCTTCAGACAGCATA...
PE
CGATGTCTTCAGACAGCATA
CGATGTCGTCAGACAGCATA
CGATGTCGAGCGACAGCATA
CGATGTCGTCGGACAGCATA
CGATGTCGTCGGATAGCATA
Intended edit
Silent mutations
pegRNA-templated 3'DNA flap (5'-3')
FANCF D9Y
RUNX1 A2S
0 10 20 30 40
0 1 2 3
% of total sequencing reads with the intended coding edit in HEK293T cells
PE4/PE2 fold-change in frequency of the intended coding edit in HEK293T cells

FIG. 44H

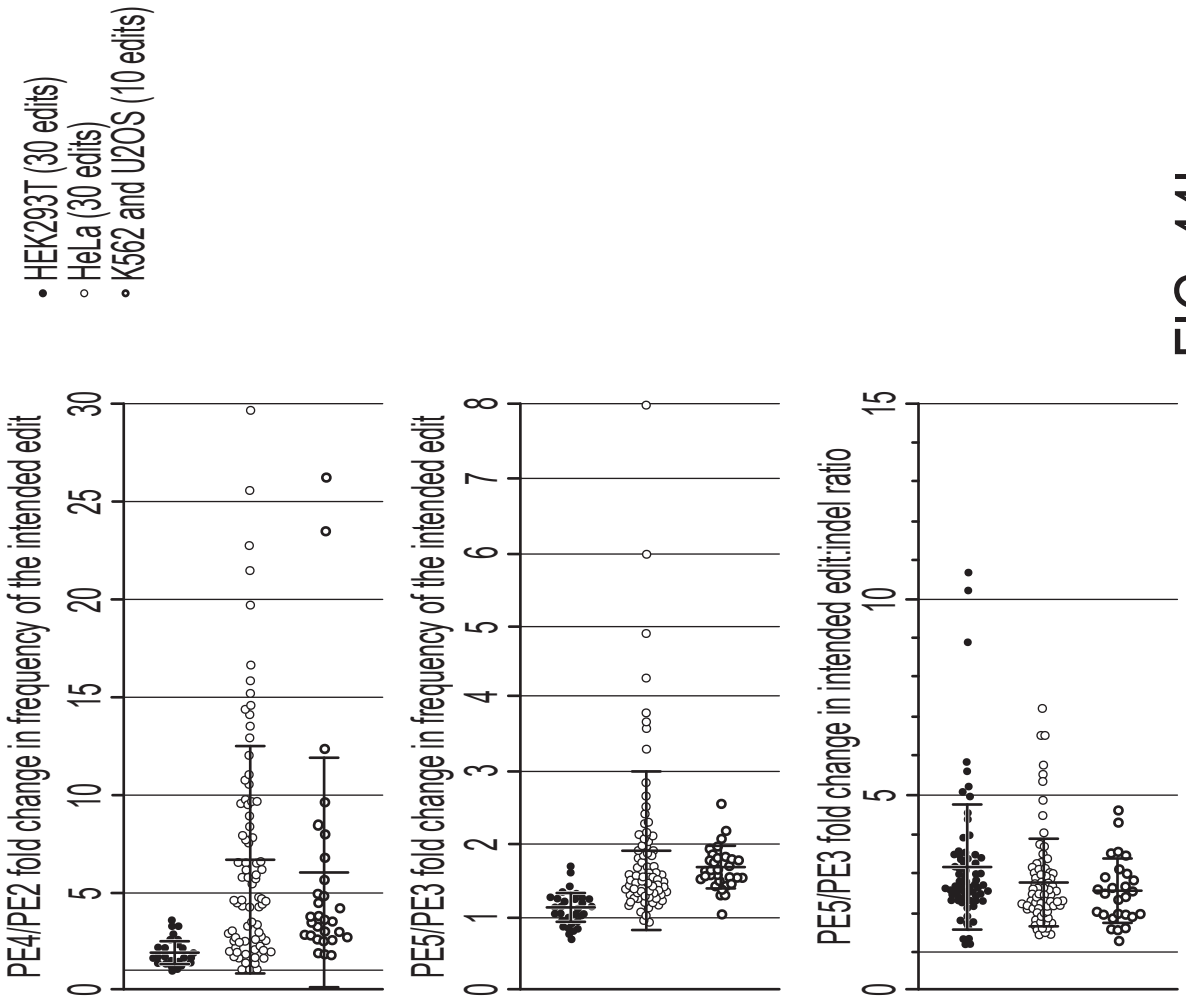

• HEK293T (30 edits)
◦ HeLa (30 edits)
• K562 and U2OS (10 edits)
PE4/PE2 fold change in frequency of the intended edit
0
5
10
15
20
25
30
PE5/PE3 fold change in frequency of the intended edit
0
1
2
3
4
5
6
7
8
PE5/PE3 fold change in intended edit:indel ratio
0
5
10
15

FIG. 44I

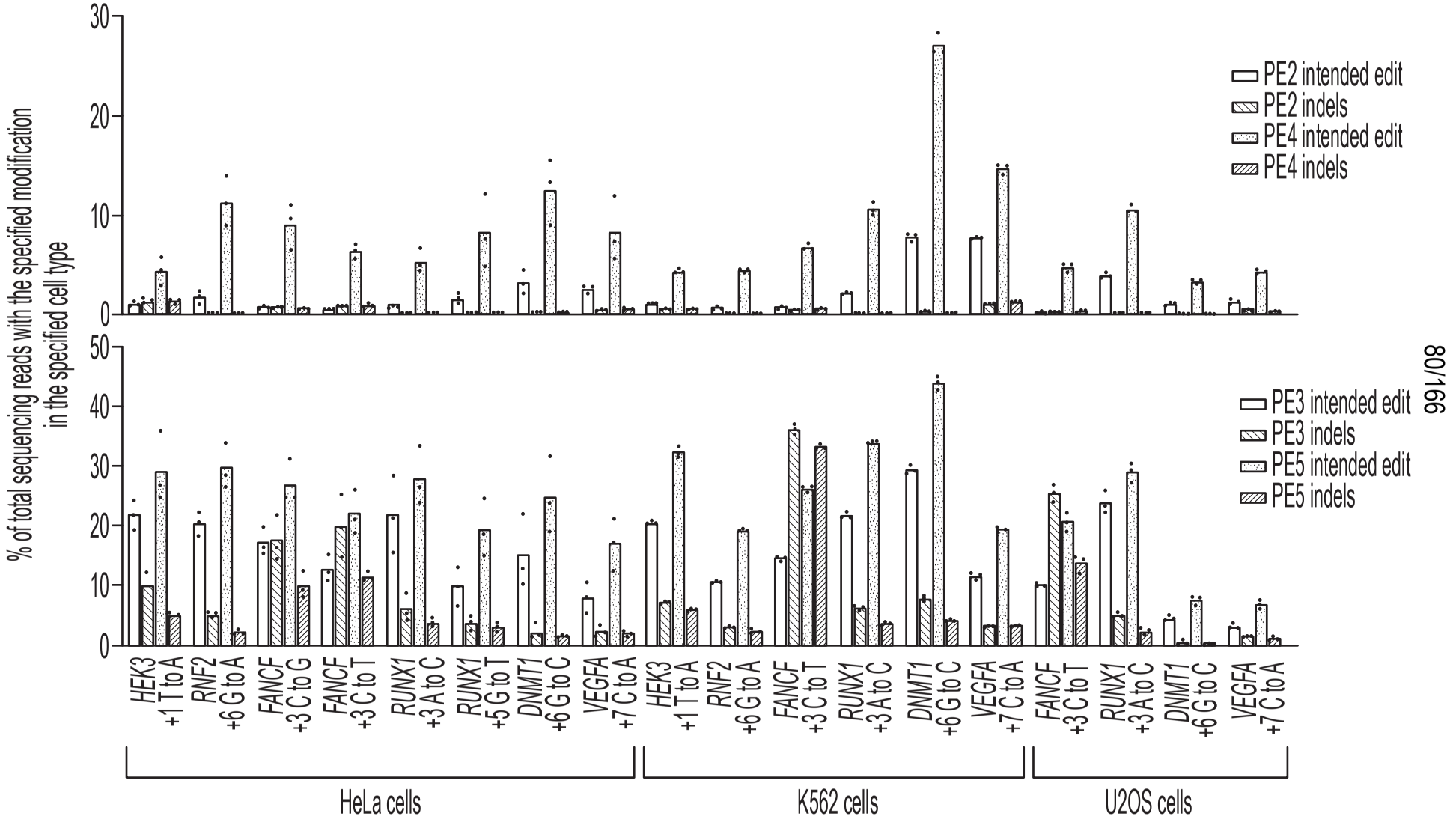

% of total sequencing reads with the specified modification in the specified cell type
PE2 intended edit
PE2 indels
PE4 intended edit
PE4 indels
PE3 intended edit
PE3 indels
PE5 intended edit
PE5 indels
HEK3 +1 T to A
RNF2 +6 G to A
FANCF +3 C to G
FANCF +3 C to T
RUNX1 +3 A to C
RUNX1 +5 G to T
DNMT1 +6 G to C
VEGFA +7 C to A
HEK3 +1 T to A
RNF2 +6 G to A
FANCF +3 C to T
RUNX1 +3 A to C
DNMT1 +6 G to C
VEGFA +7 C to A
FANCF +3 C to T
RUNX1 +3 A to C
DNMT1 +6 G to C
VEGFA +7 C to A
HeLa cells
K562 cells
U2OS cells

FIG. 44J

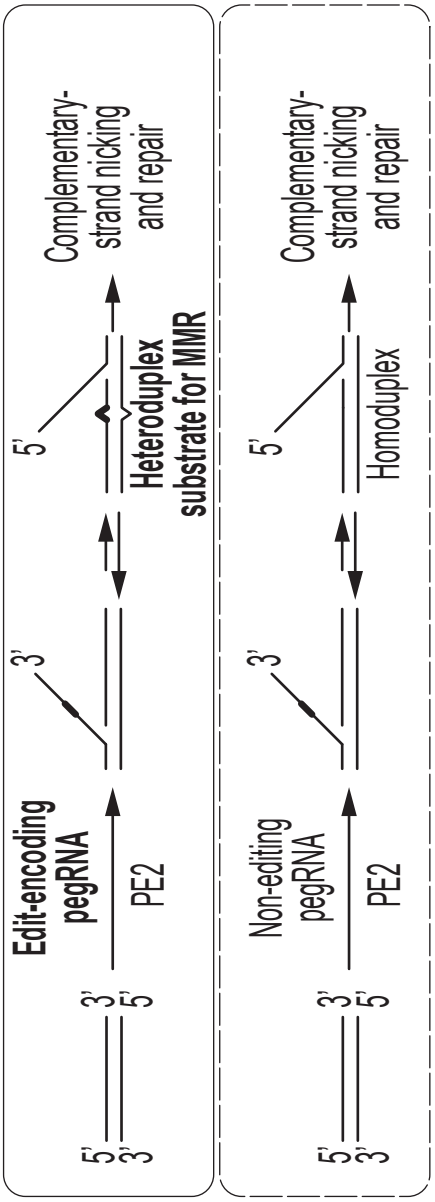


FIG. 45A

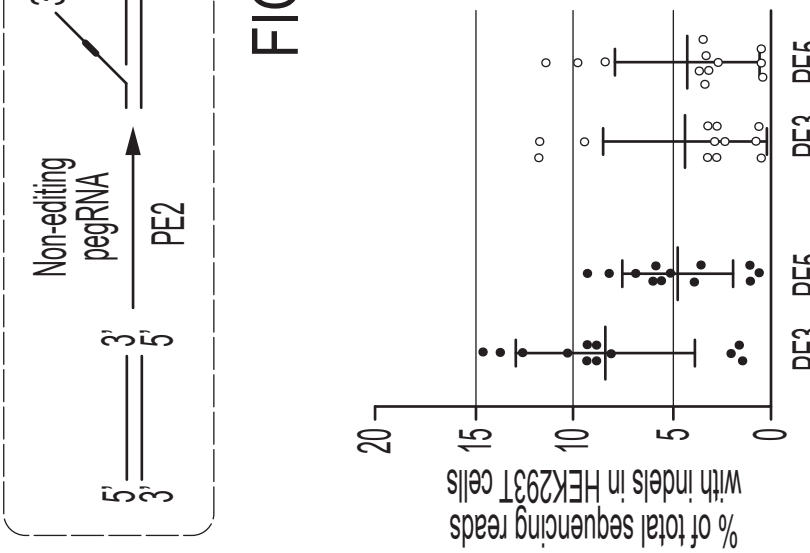


FIG. 45B

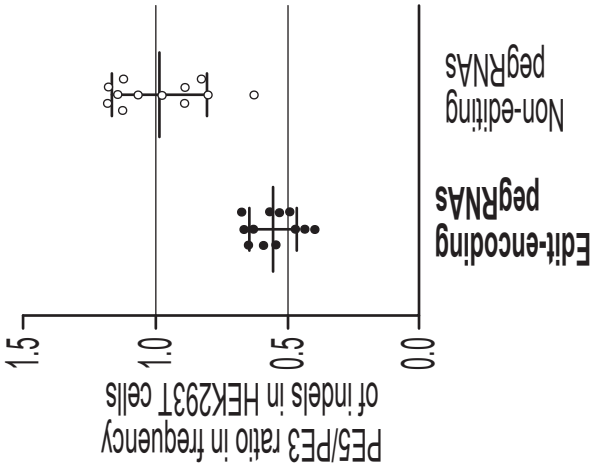


FIG. 45C

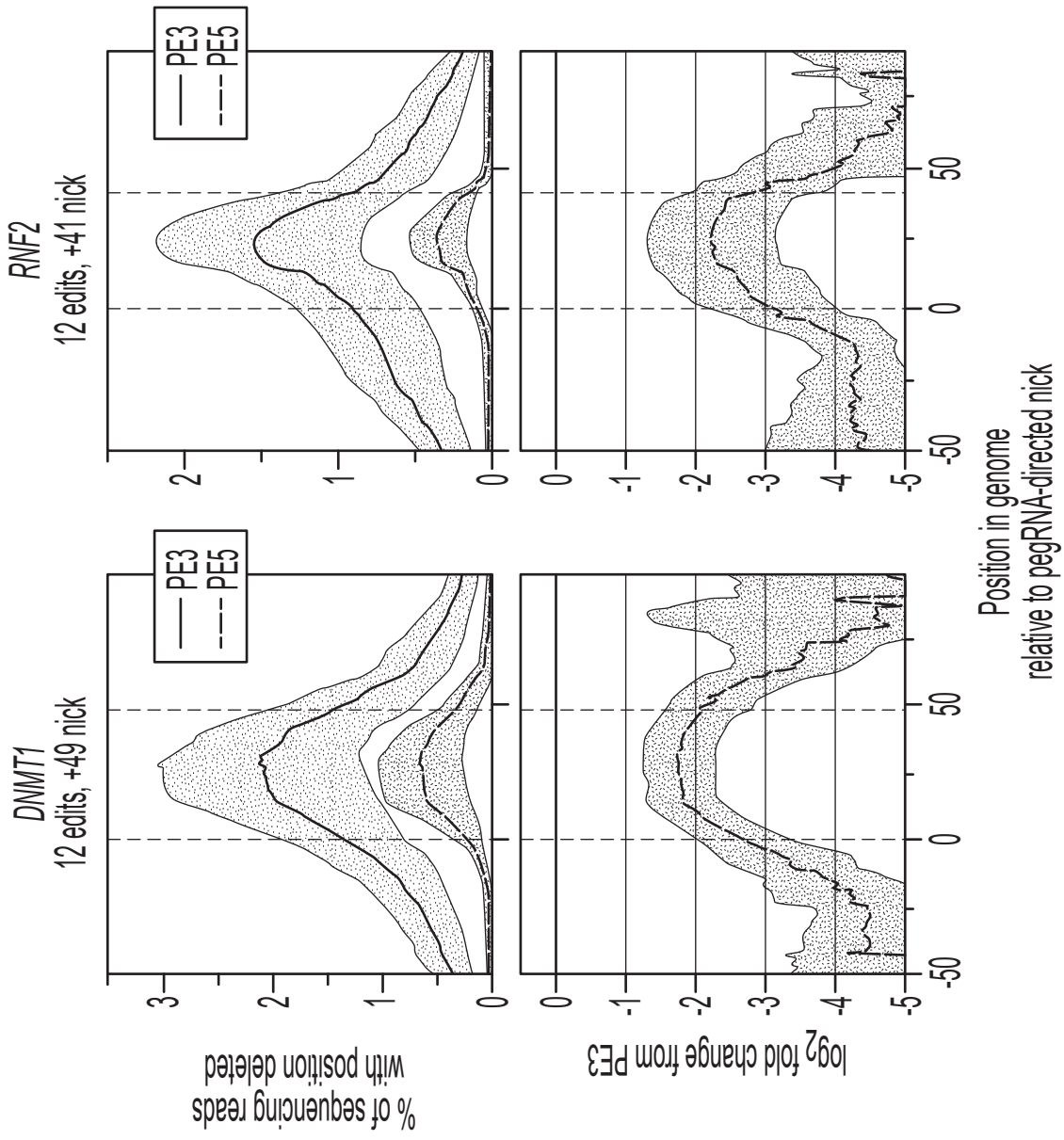

DNMT1
12 edits, +49 nick
RNF2
12 edits, +41 nick
PE3
PE5
% of sequencing reads with position deleted
log2 fold change from PE3
Position in genome relative to pegRNA-directed nick

FIG. 45D

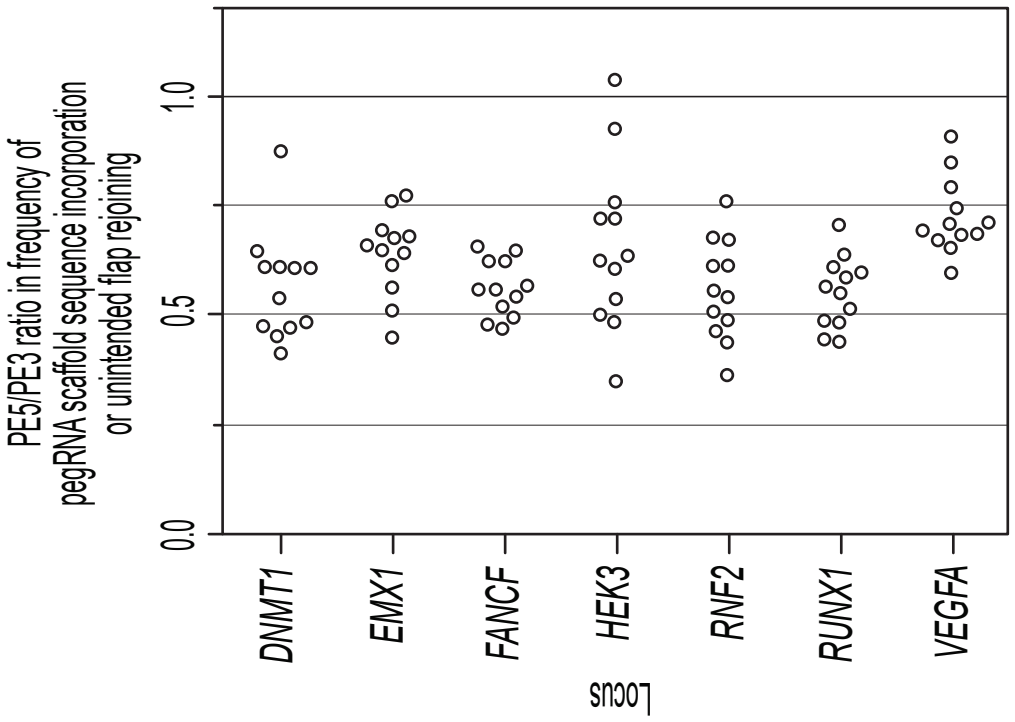


FIG. 45F

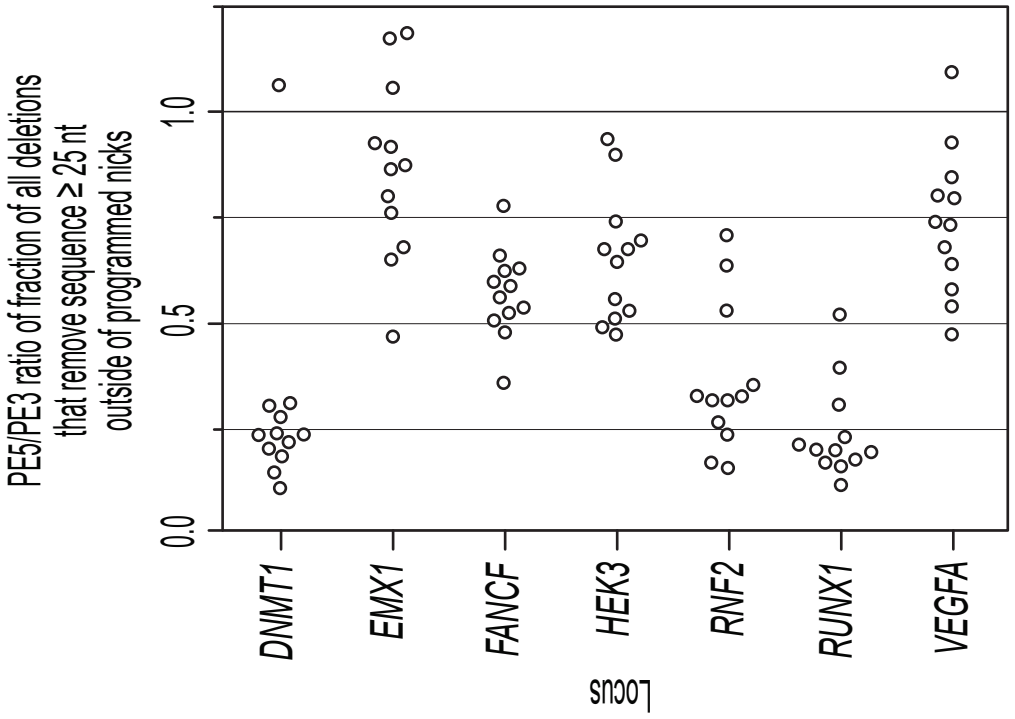


FIG. 45E

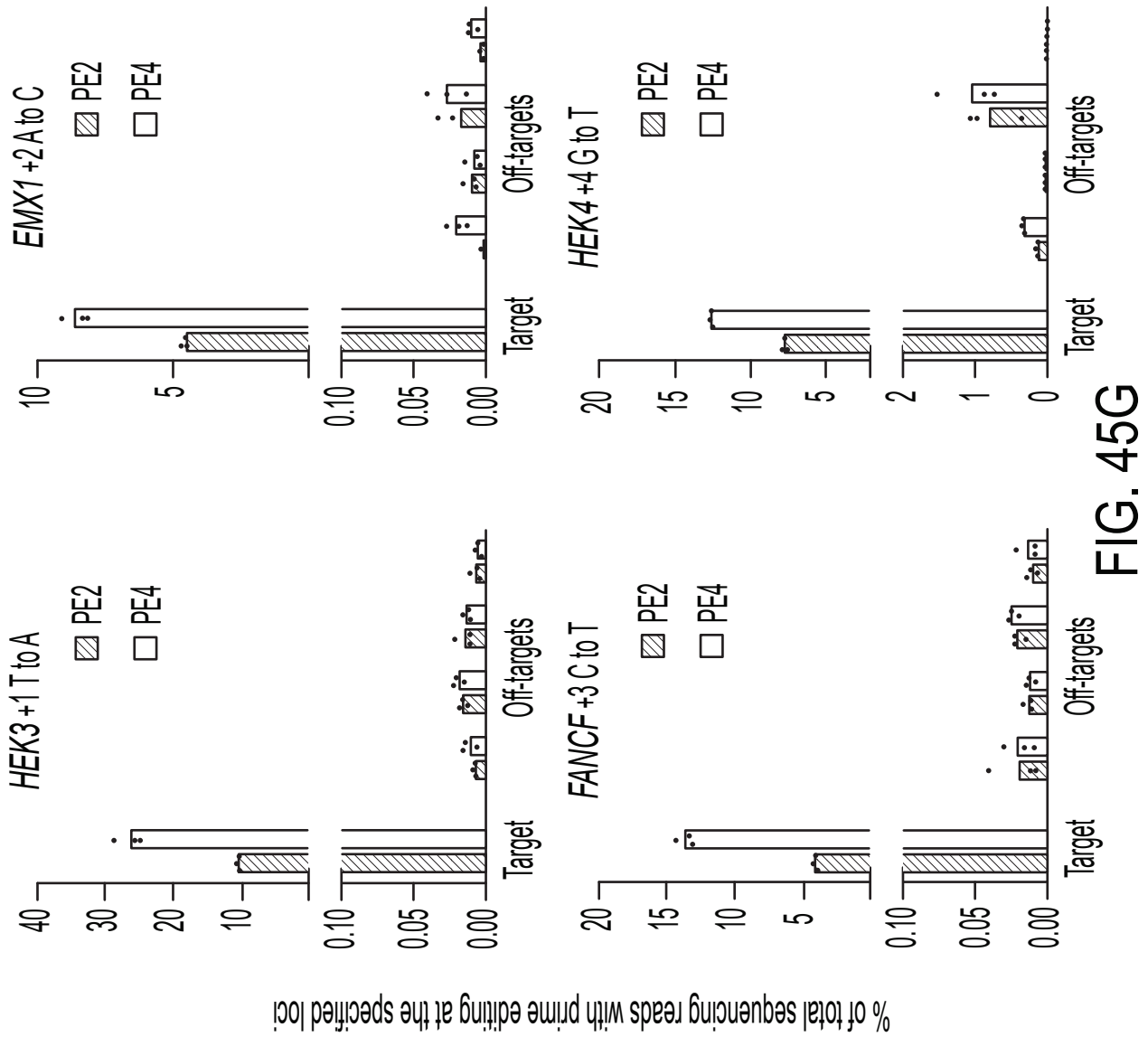

HEK3 +1 T to A
EMX1 +2 A to C
FANCF +3 C to T
HEK4 +4 G to T
PE2
PE4
Target
Off-targets
% of total sequencing reads with prime editing at the specified loci

FIG. 45G

HAP1 wt
HAP1 ΔMSH2
HCT116 (MLH1-/-)
HeLa PE2
HeLa PE4
Average microsatellite repeat length (nt)
30
25
20
15
10
HAP1 ΔMSH2 - HAP1 wt
HeLa PE4 - HeLa PE2
MMR inhibition-induced change in average microsatellite repeat length (nt)
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
Bat-25
Bat-26
MONO-27
NR-21
NR-24
HSPH1-T17
EWSR1
MSI-01
MSI-03
MSI-04
MSI-06
MSI-07
MSI-08
MSI-09
MSI-11
MSI-12
MSI-13
Microsatellite locus

FIG. 45H

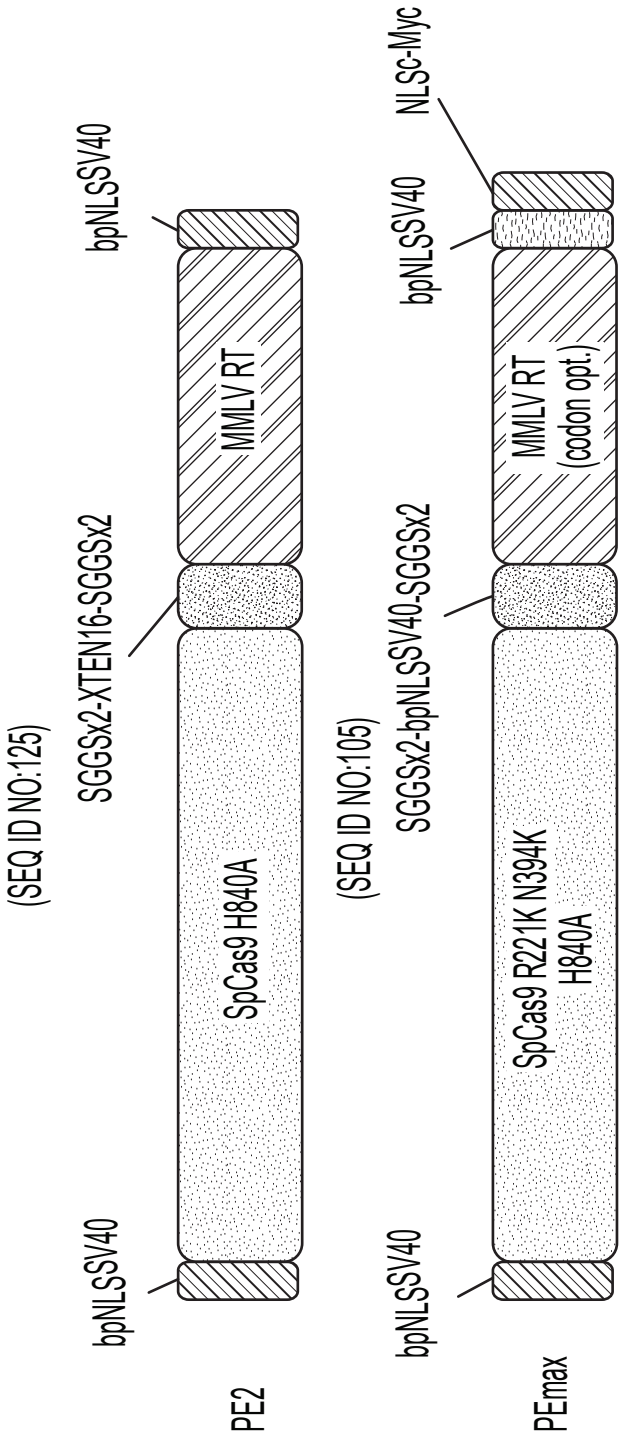

bpNLSSV40
SpCas9 H840A
SGGSx2-XTEN16-SGGSx2
(SEQ ID NO:125)
MMLV RT
bpNLSSV40
PE2
bpNLSSV40
SpCas9 R221K N394K
H840A
SGGSx2-bpNLSSV40-SGGSx2
(SEQ ID NO:105)
MMLV RT
(codon opt.)
bpNLSSV40
NLSc-Myc
PEmax

FIG. 46A

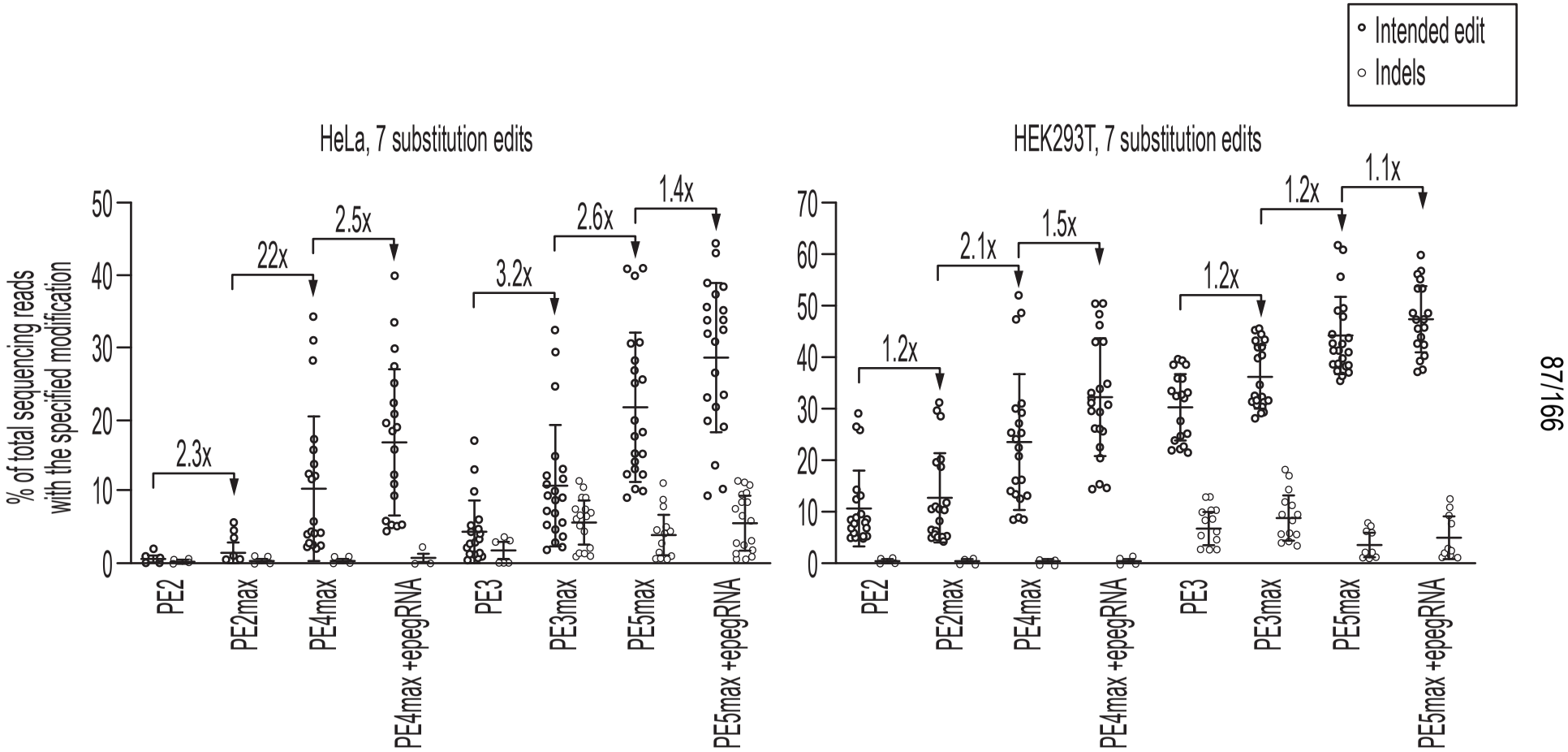

Intended edit
Indels
HeLa, 7 substitution edits
HEK293T, 7 substitution edits
% of total sequencing reads with the specified modification
0
10
20
30
40
50
0
10
20
30
40
50
60
70
2.3x
22x
2.5x
3.2x
2.6x
1.4x
1.2x
2.1x
1.5x
1.2x
1.2x
1.1x
PE2
PE2max
PE4max
PE4max +epegRNA
PE3
PE3max
PE5max
PE5max +epegRNA

FIG. 46B

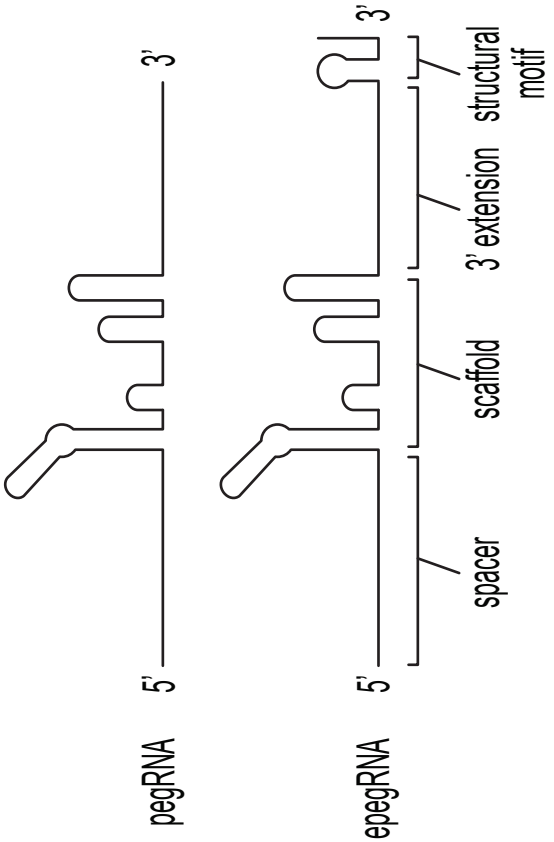

pegRNA
5'
3'
epegRNA
5'
3'
spacer
scaffold
3' extension
structural motif

FIG. 46C

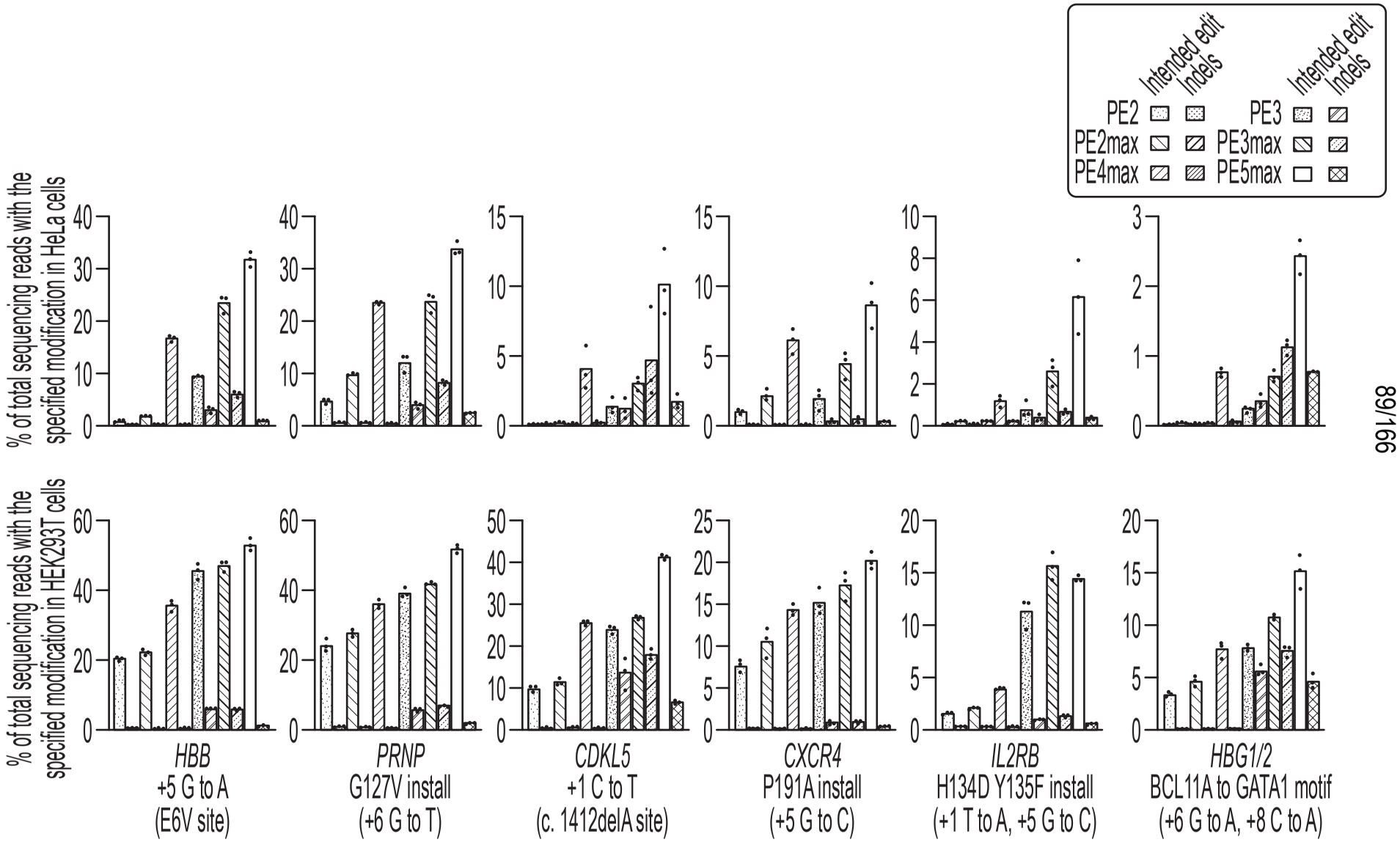

Intended edit
Indels
PE2
PE3
PE2max
PE3max
PE4max
PE5max
% of total sequencing reads with the specified modification in HeLa cells
% of total sequencing reads with the specified modification in HEK293T cells
HBB
+5 G to A
(E6V site)
PRNP
G127V install
(+6 G to T)
CDKL5
+1 C to T
(c. 1412delA site)
CXCR4
P191A install
(+5 G to C)
IL2RB
H134D Y135F install
(+1 T to A, +5 G to C)
HBG1/2
BCL11A to GATA1 motif
(+6 G to A, +8 C to A)

FIG. 46D

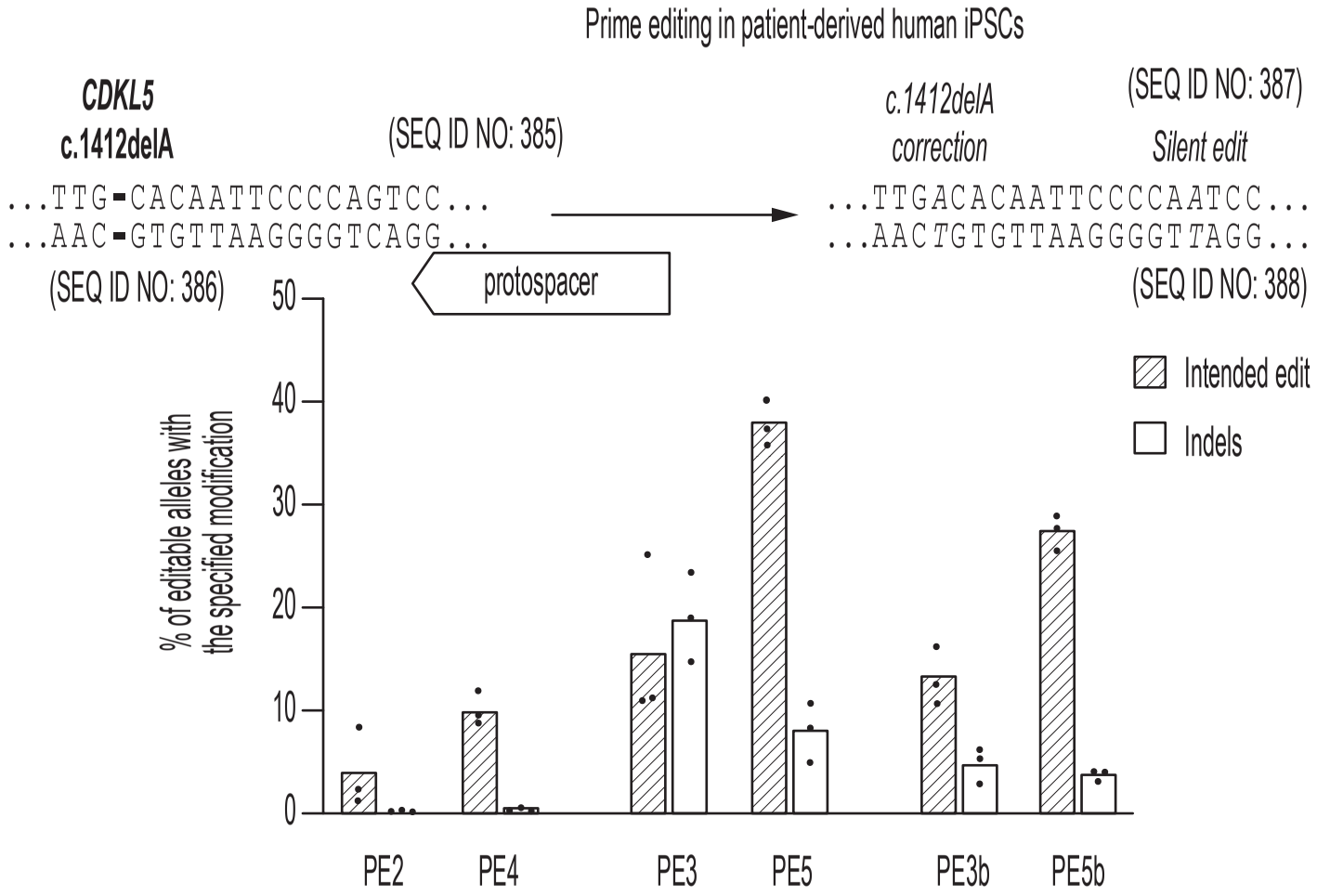

Prime editing in patient-derived human iPSCs
CDKL5
c.1412delA
(SEQ ID NO: 385)
...TTG-CACAATTCCCCAGTCC...
...AAC-GTGTTAAGGGGTCAGG...
(SEQ ID NO: 386)
protospacer
c.1412delA
correction
(SEQ ID NO: 387)
Silent edit
...TTGACACAATTCCCCAATCC...
...AACTGTGTTAAGGGGTTAGG...
(SEQ ID NO: 388)
Intended edit
Indels
% of editable alleles with the specified modification
0
10
20
30
40
50
PE2
PE4
PE3
PE5
PE3b
PE5b

FIG. 46E

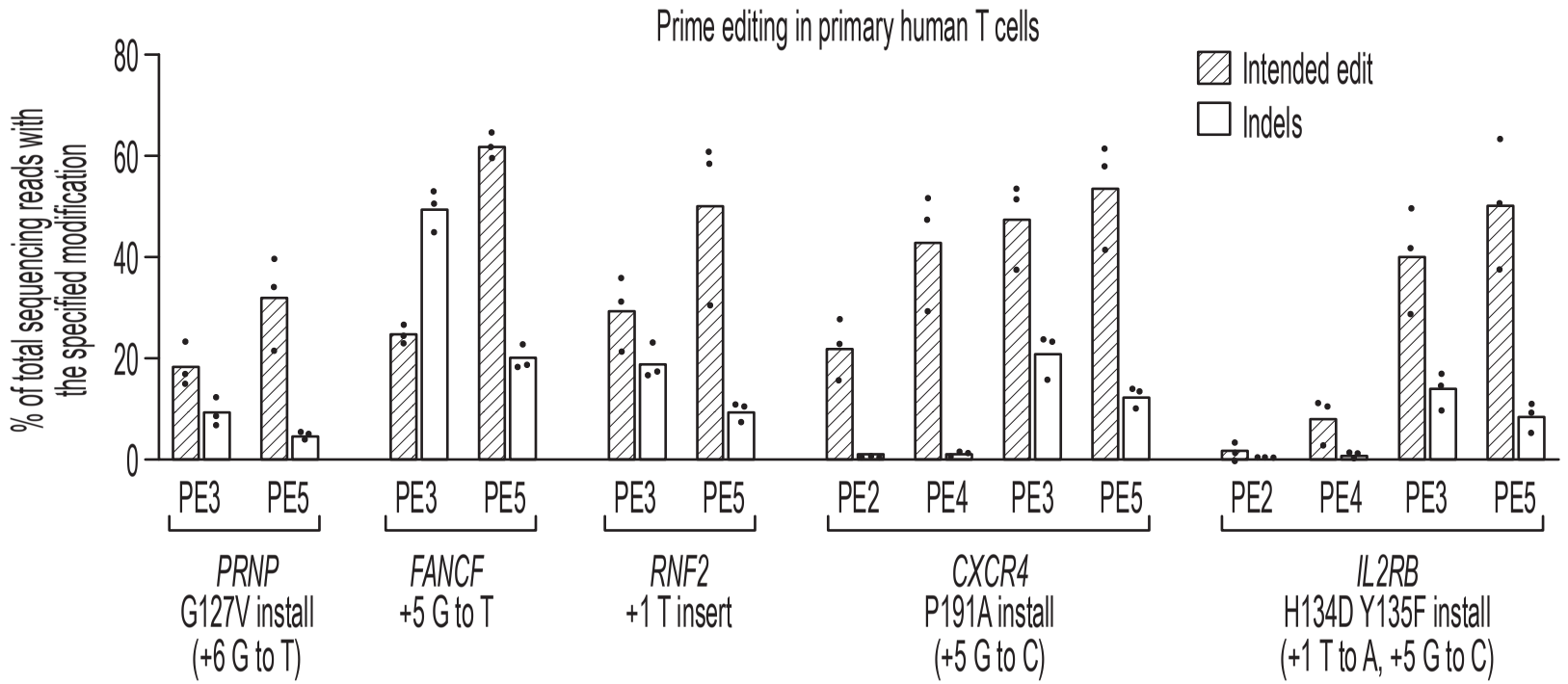

Prime editing in primary human T cells
Intended edit
Indels
% of total sequencing reads with the specified modification
80
60
40
20
0
PE3
PE5
PE3
PE5
PE3
PE5
PE2
PE4
PE3
PE5
PE2
PE4
PE3
PE5
PRNP
G127V install
(+6 G to T)
FANCF
+5 G to T
RNF2
+1 T insert
CXCR4
P191A install
(+5 G to C)
IL2RB
H134D Y135F install
(+1 T to A, +5 G to C)

FIG. 46F

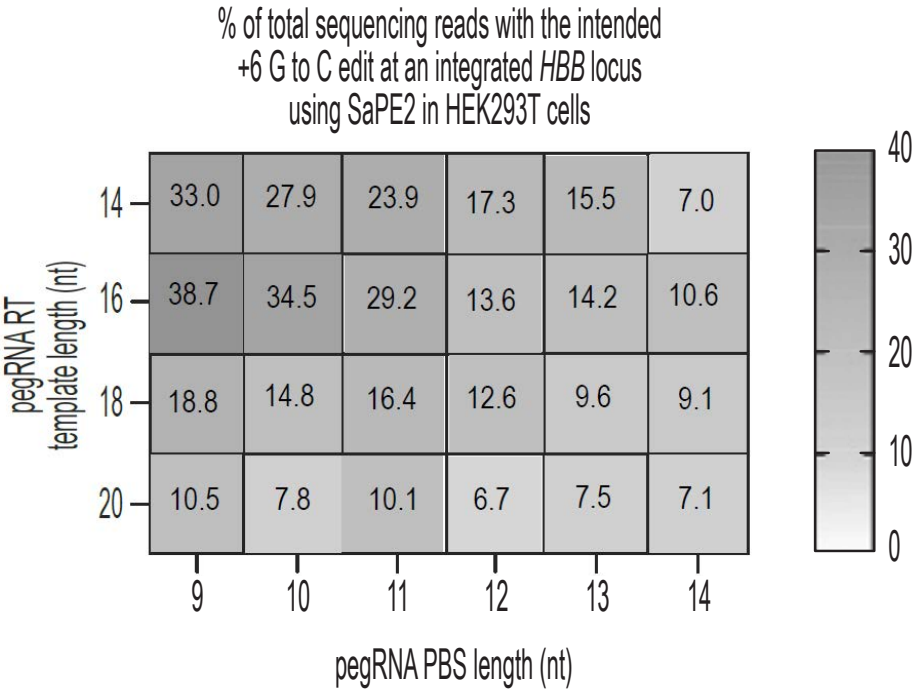


FIG. 47A

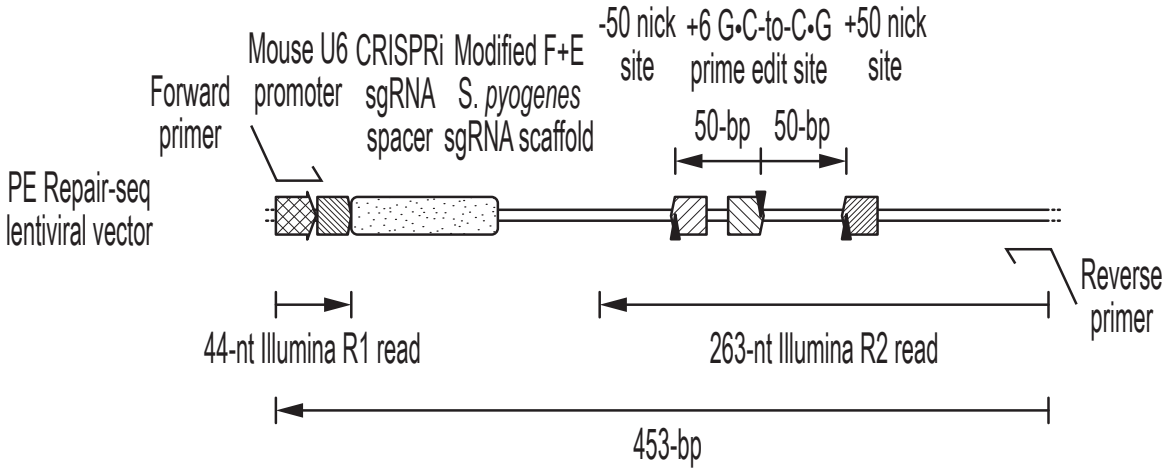


FIG. 47B

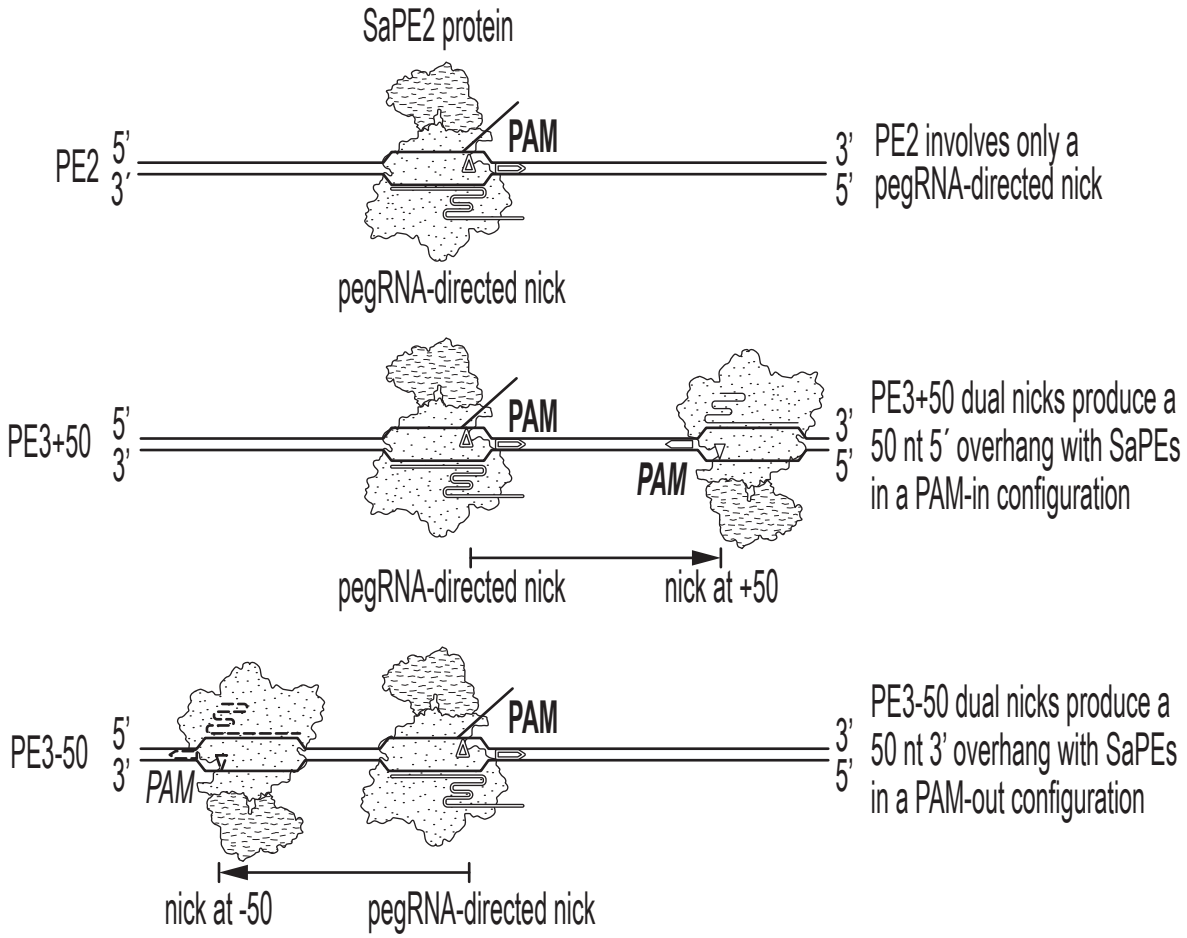

SaPE2 protein
PE2
5'
3'
PAM
3'
5'
PE2 involves only a pegRNA-directed nick
pegRNA-directed nick
PE3+50
5'
3'
PAM
PAM
3'
5'
PE3+50 dual nicks produce a 50 nt 5' overhang with SaPEs in a PAM-in configuration
pegRNA-directed nick
nick at +50
PE3-50
5'
3'
PAM
PAM
3'
5'
PE3-50 dual nicks produce a 50 nt 3' overhang with SaPEs in a PAM-out configuration
nick at -50
pegRNA-directed nick

FIG. 47C

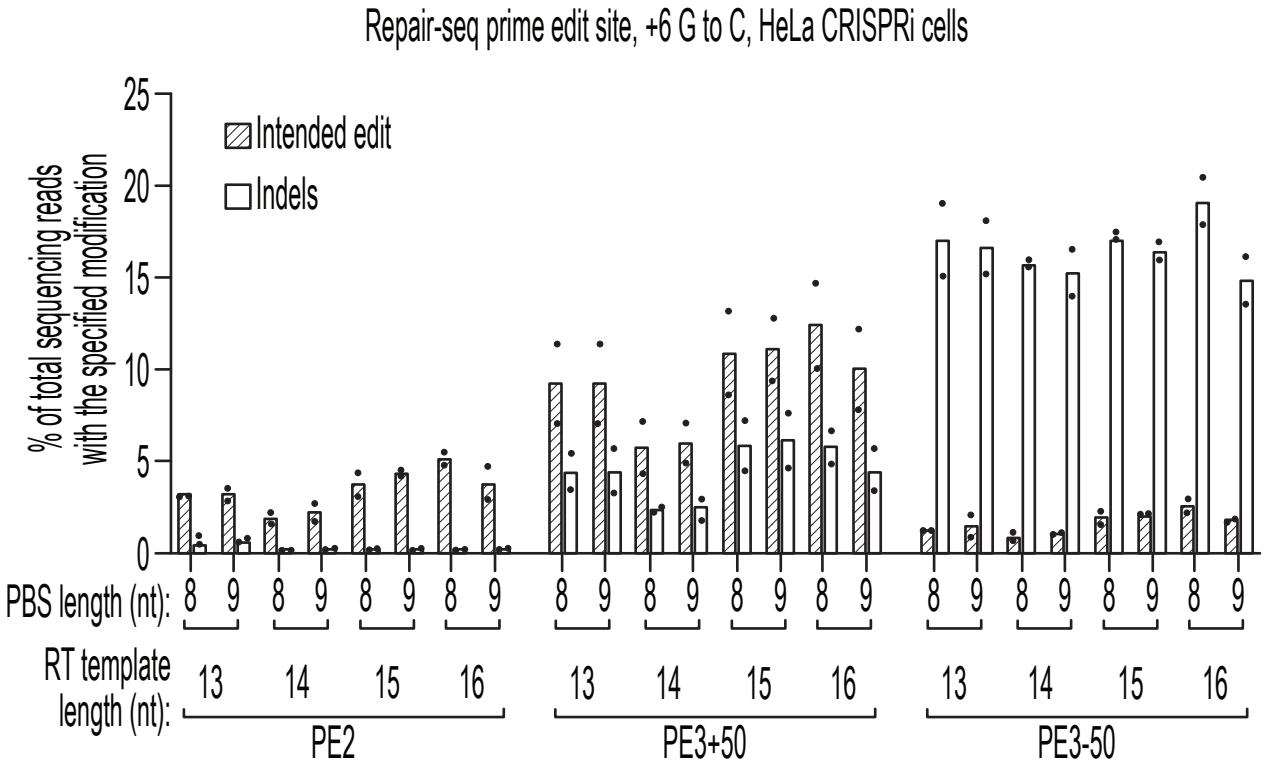


FIG. 47D

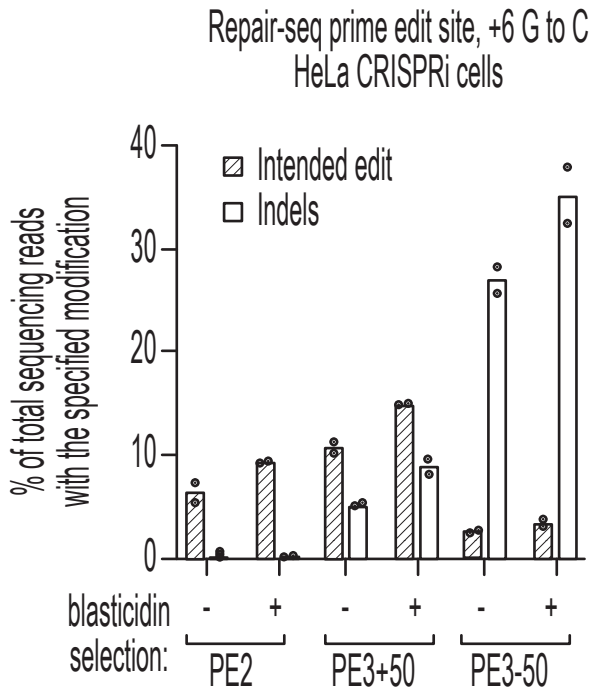


FIG. 47E

Functional annotation class	Number of genes in CRISPRi sgRNA library
Recombinational repair	41
Double-strand break repair	36
Nucleotide-excision repair	30
Base-excision repair	23
Postreplication repair	16
Mismatch repair	16
DNA repair (other)	50
Cellular response to stimulus	4
Nuclear division	12
Cell cycle checkpoint	12
mRNA processing	21
DNA replication	23
Protein modification	24
Other	39
Regulation of transcription	46
Chromatin processes	83
Total	476

FIG. 47F

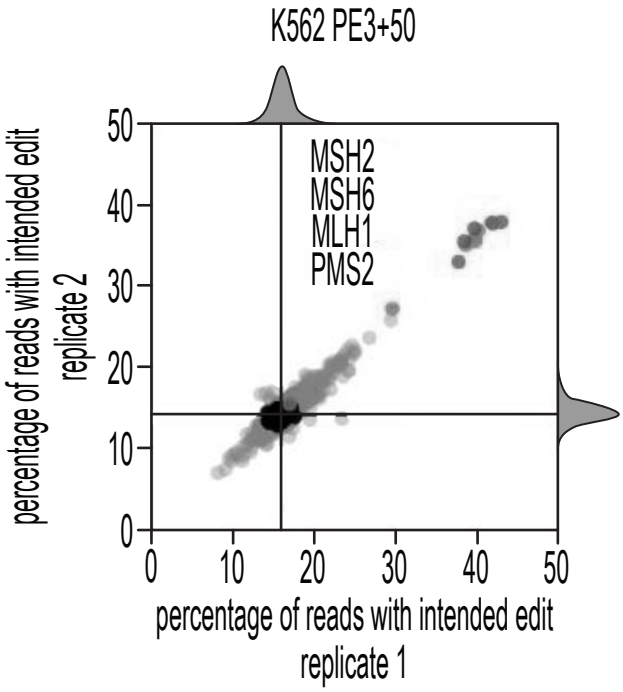


FIG. 47G

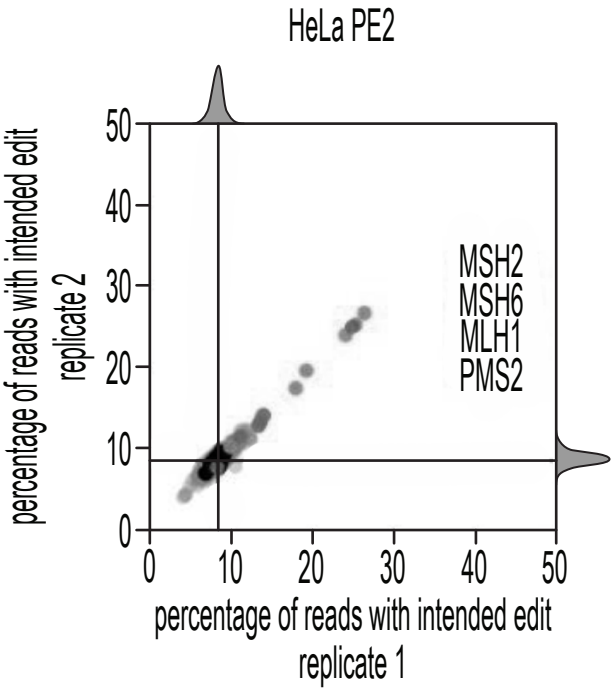


FIG. 47H

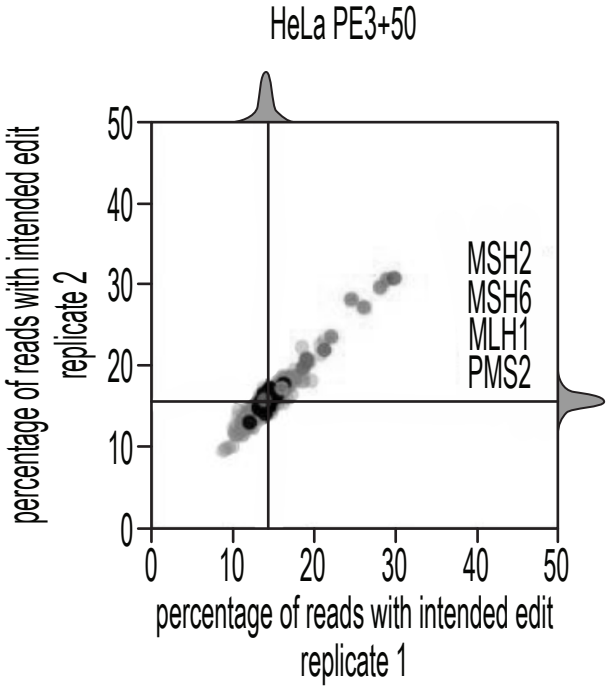


FIG. 47I

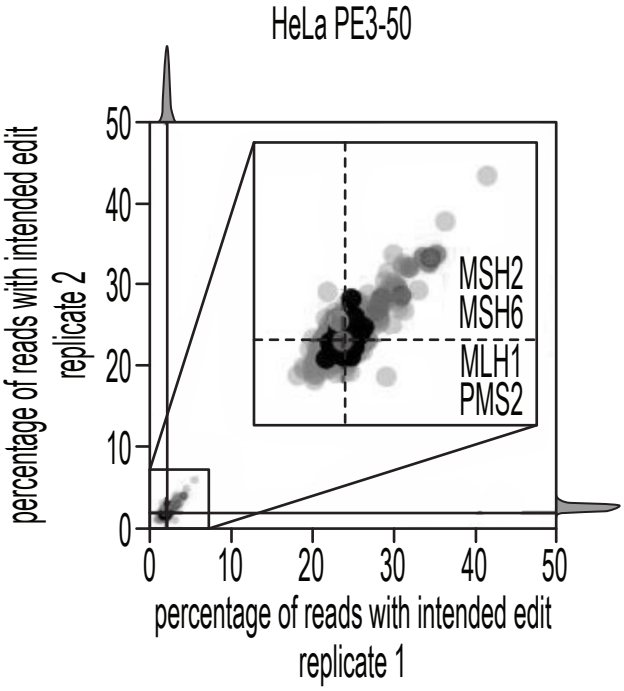


FIG. 47J

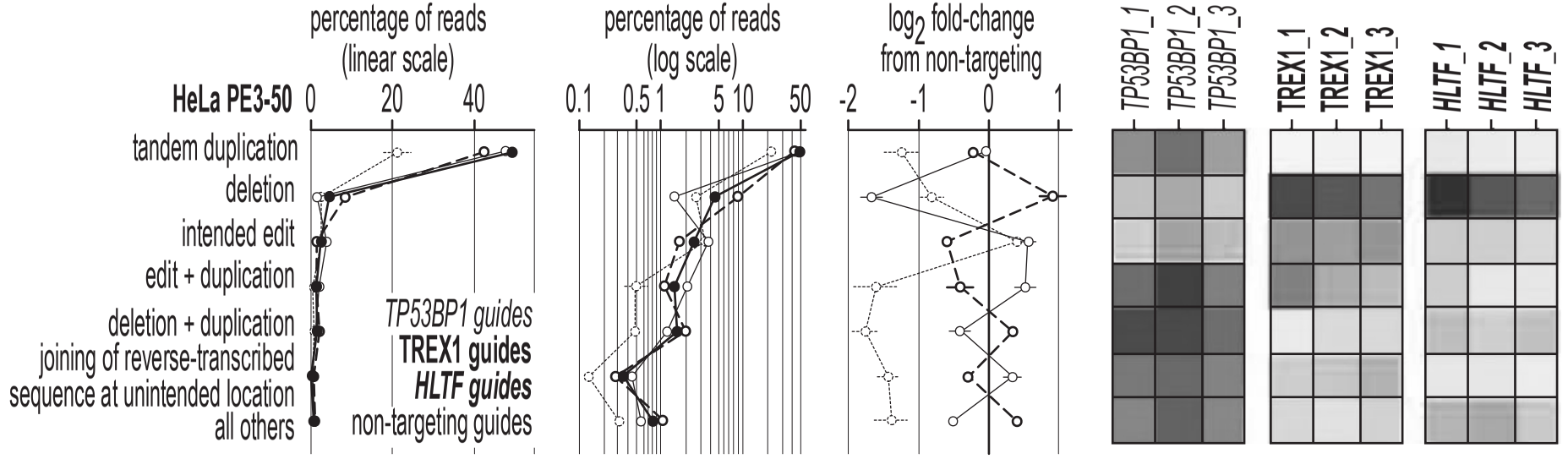


FIG. 48A

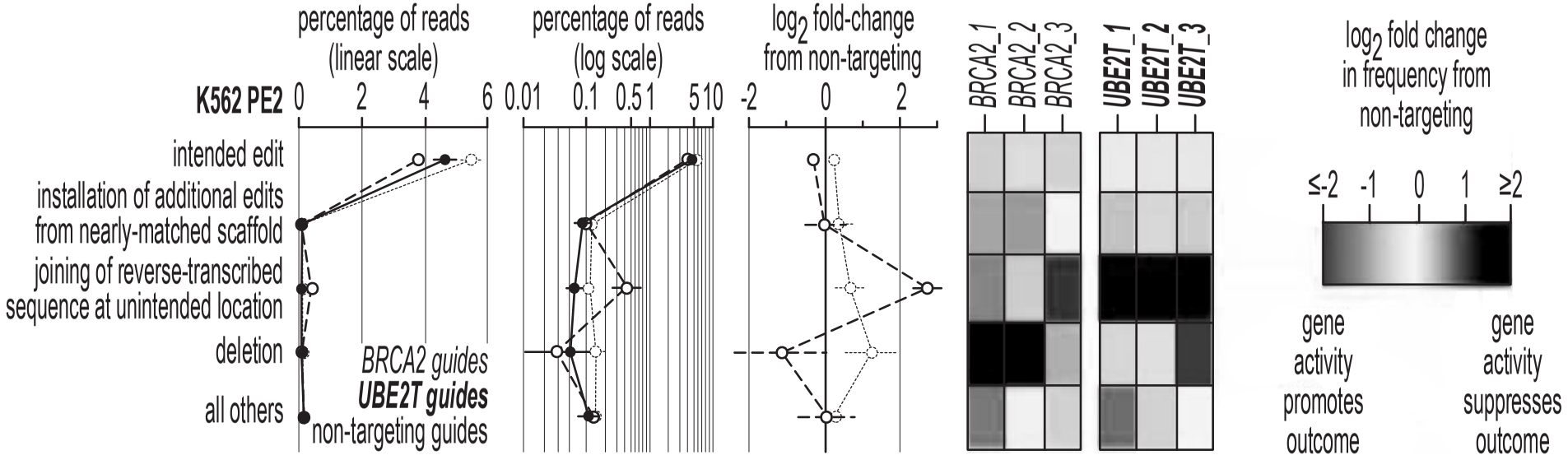
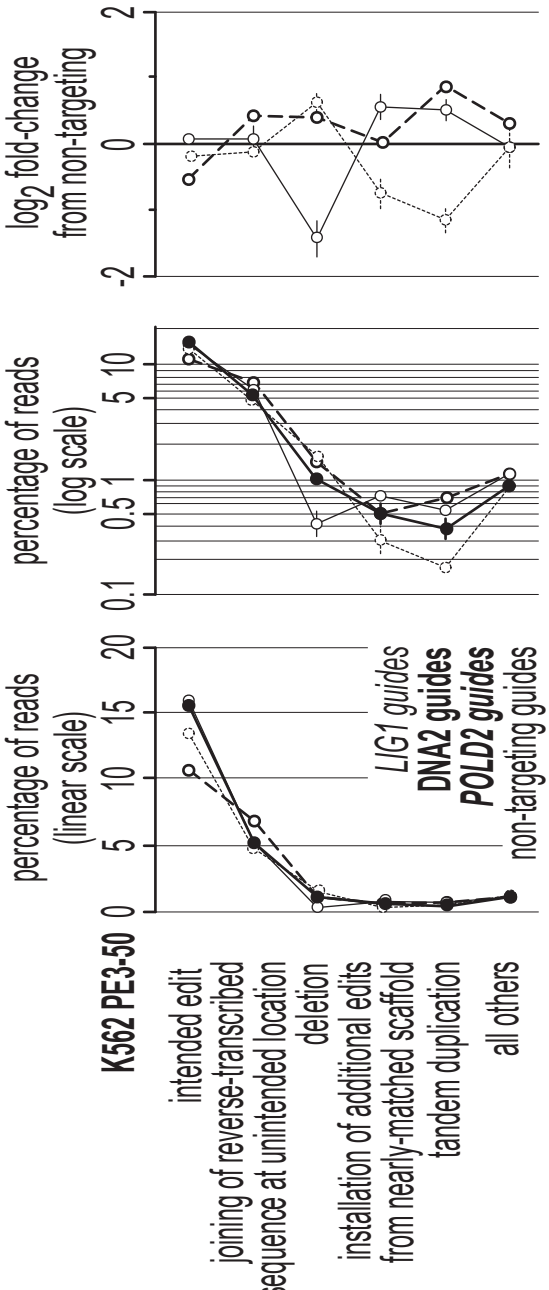


FIG. 48B


K562 PE3-50
percentage of reads (linear scale)
0
5
10
15
20
percentage of reads (log scale)
0.1
0.5
1
5
10
log2 fold-change from non-targeting
-2
0
2
intended edit
joining of reverse-transcribed sequence at unintended location
deletion
installation of additional edits from nearly-matched scaffold
tandem duplication
all others
LIG1 guides
DNA2 guides
POLD2 guides
non-targeting guides

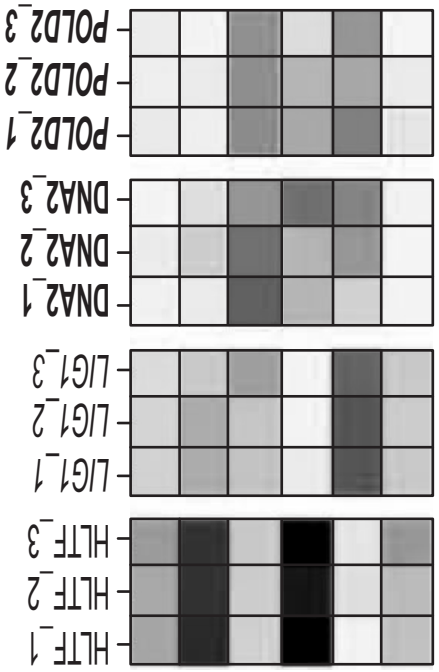

HLTF_1
HLTF_2
HLTF_3
LIG1_1
LIG1_2
LIG1_3
DNA2_1
DNA2_2
DNA2_3
POLD2_1
POLD2_2
POLD2_3

FIG. 48C

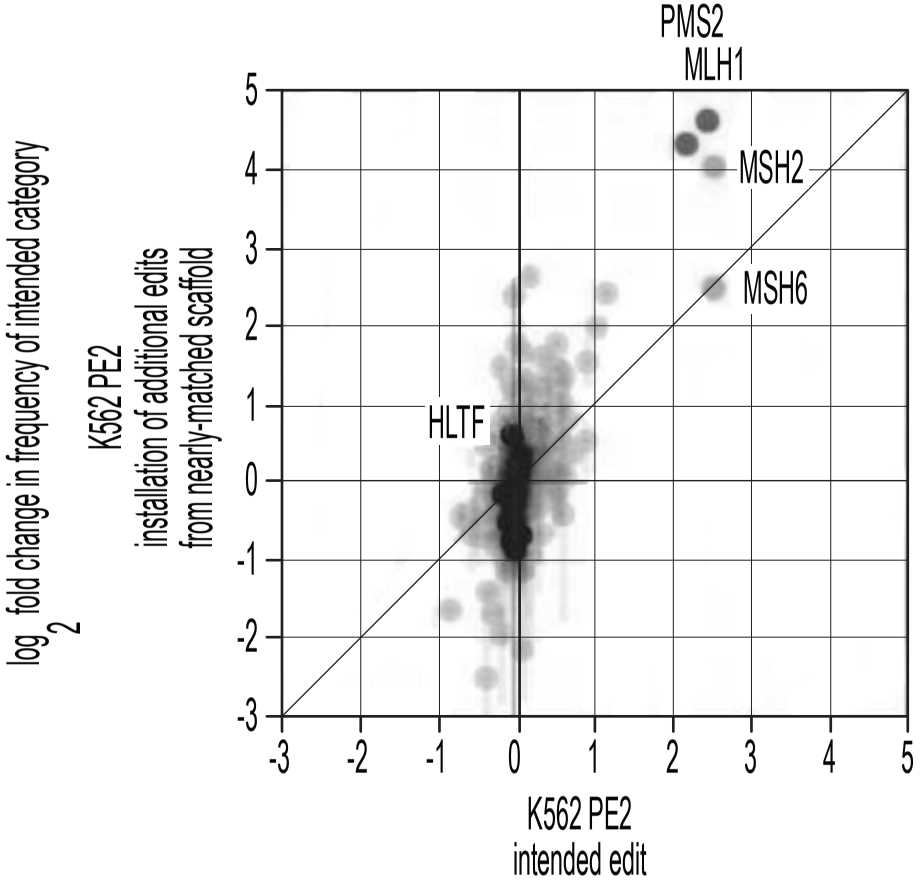


FIG. 48D

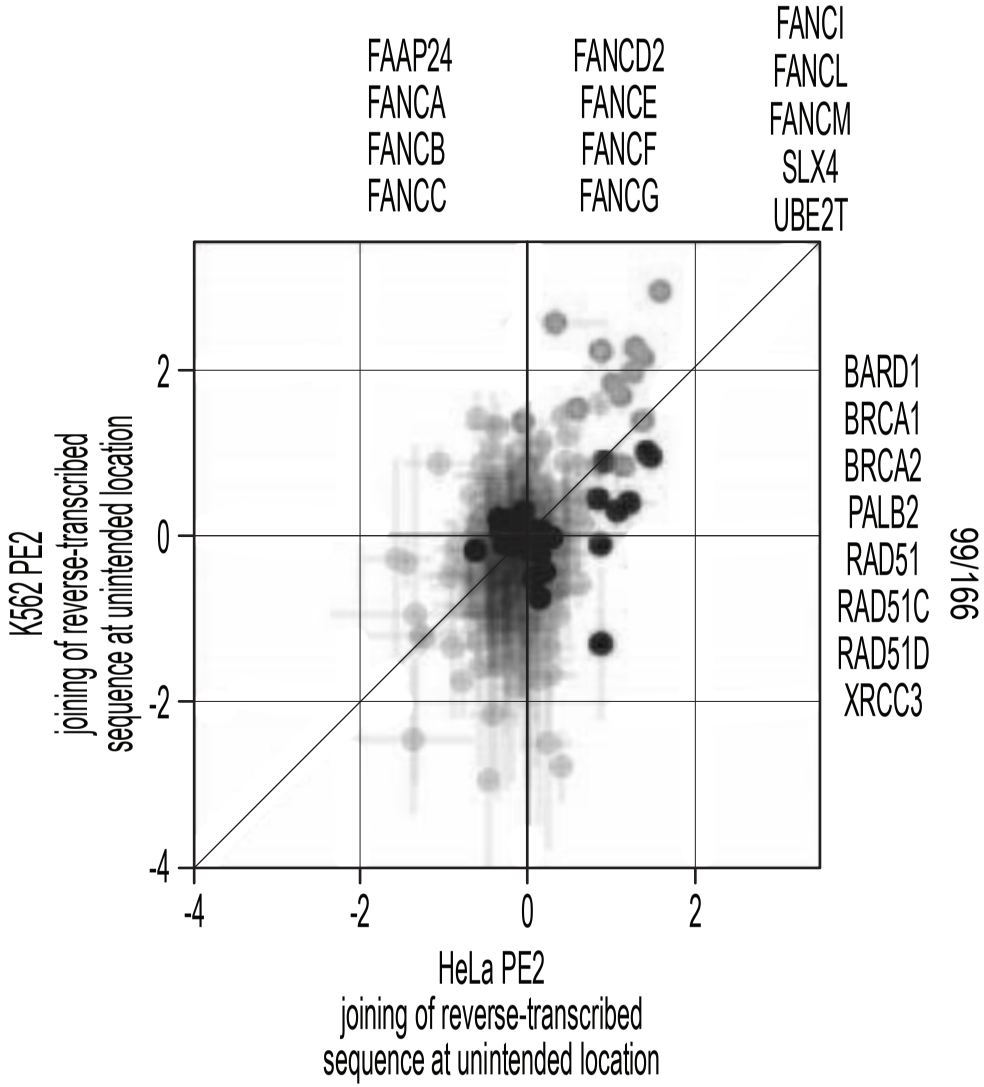


FIG. 48E

$\log_2$ fold change in frequency of indicated category

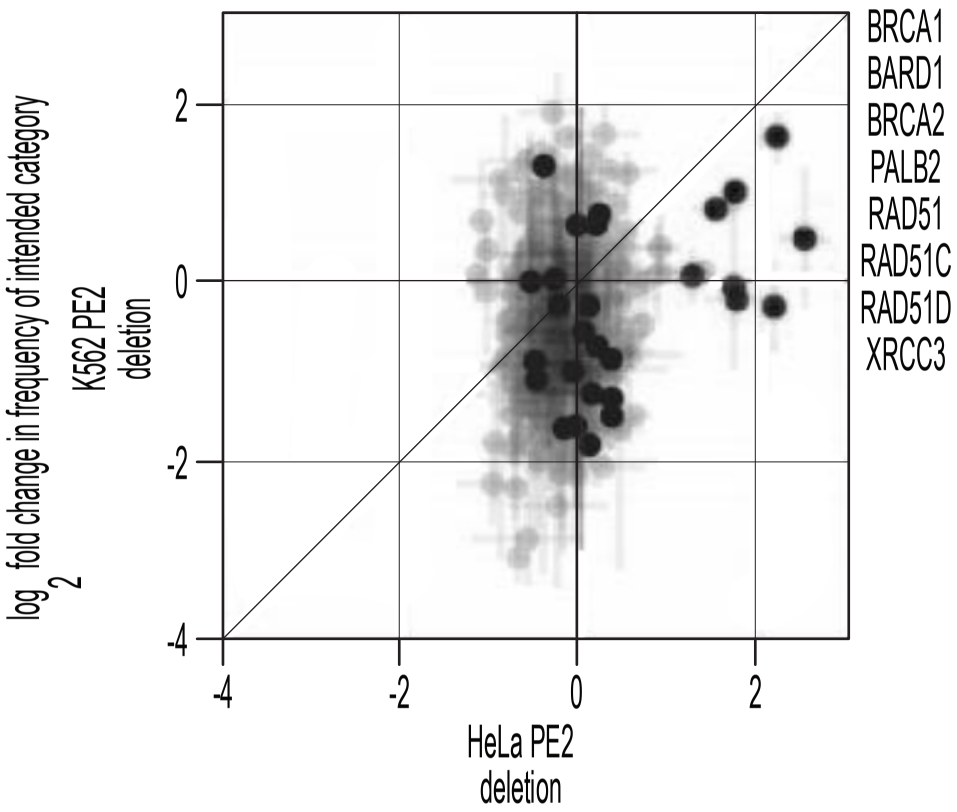


FIG. 48F

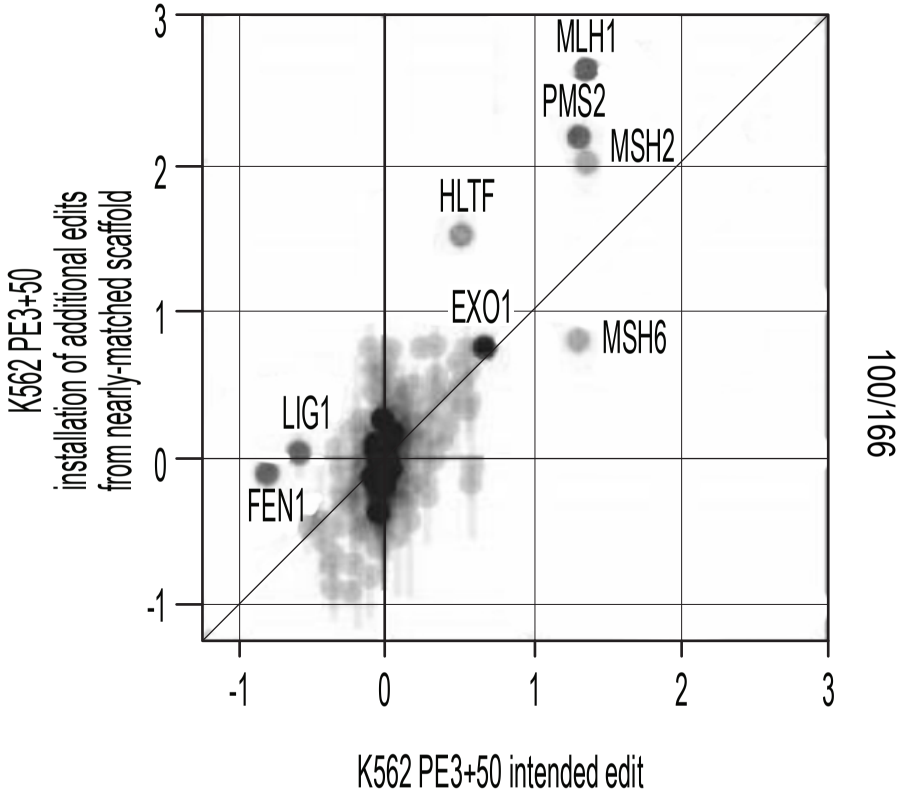


FIG. 48G

log2 fold change in frequency of indicated category

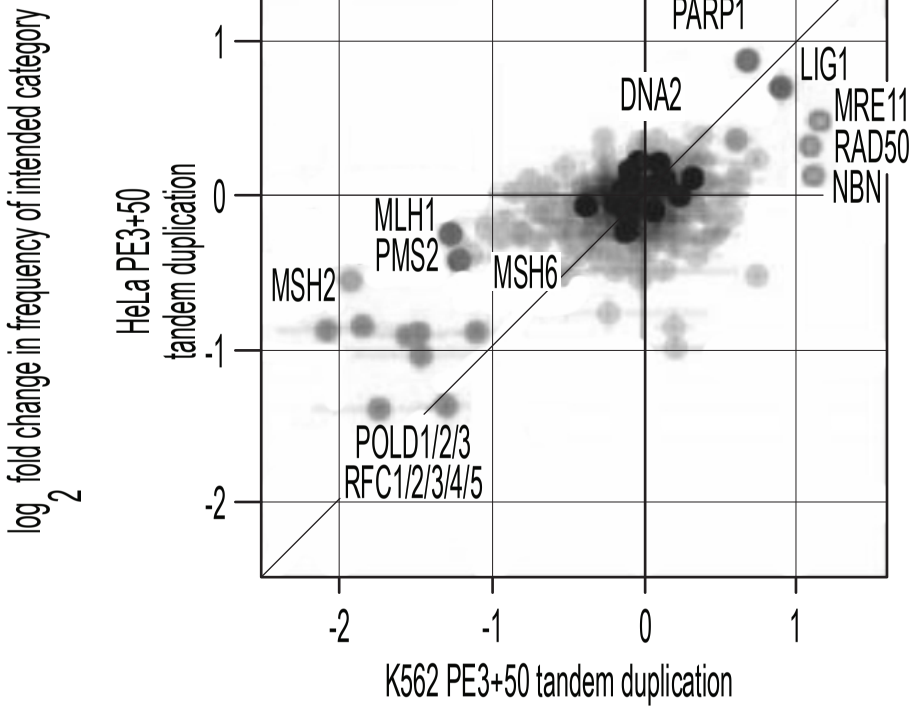


FIG. 48H

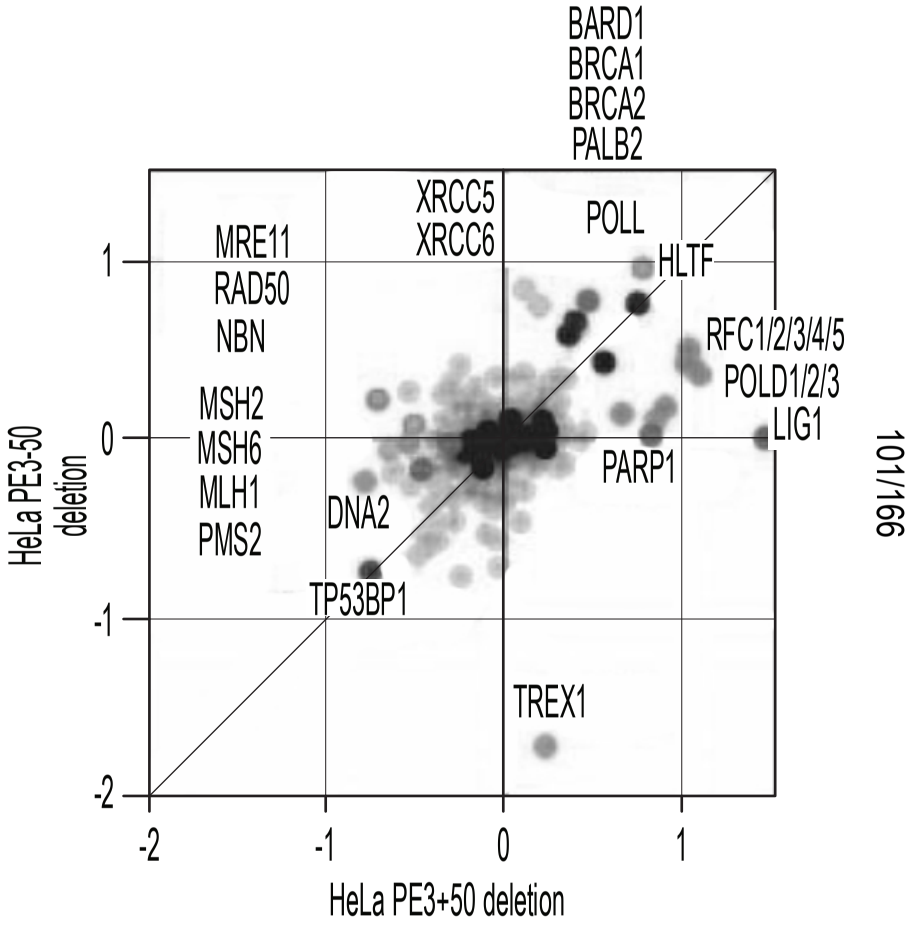
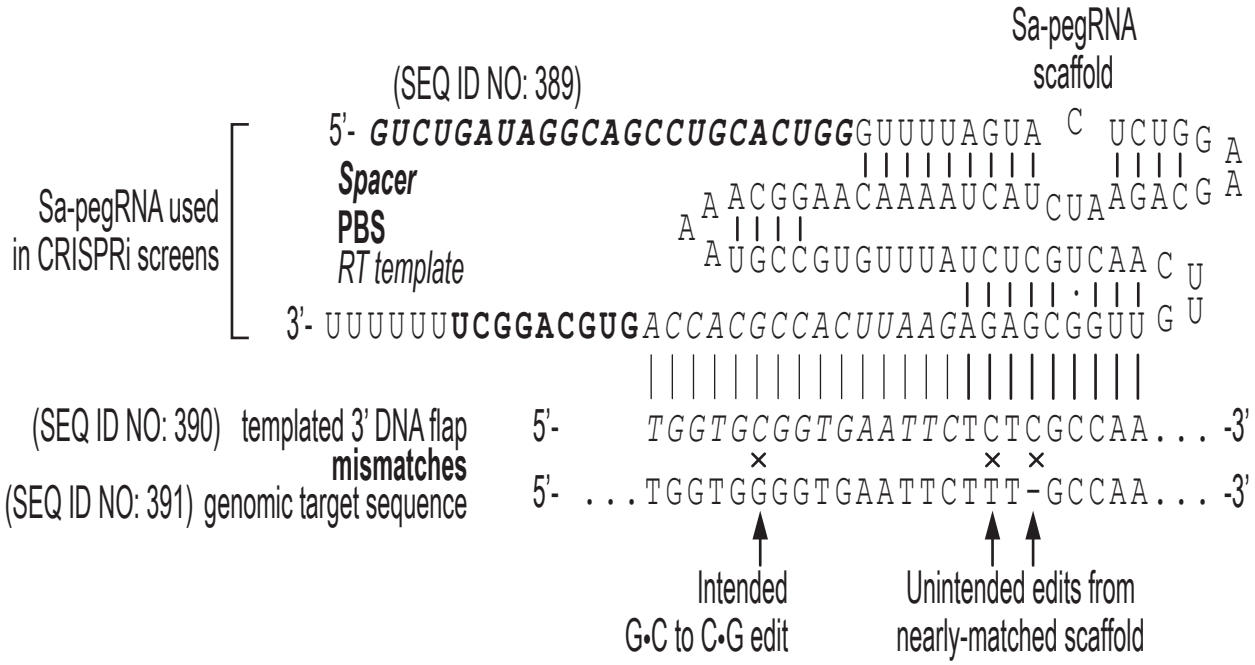


FIG. 48I

log2 fold change in frequency of indicated category


Sa-pegRNA scaffold
(SEQ ID NO: 389)
5'- GUCUGAUAGGCAGCCUGCACUGGGUUUUAGUA C UCUGG A
Spacer
PBS
RT template
A ACGGAACAAAAUCAU CUA AGACG A
A AUGCCGUGUUUAUCUCGUCAA C U
3'- UUUUUUUCGGACGUGACCACGCCACUUAAGAGAGCGGUU G U
Sa-pegRNA used in CRISPRi screens
(SEQ ID NO: 390) templated 3' DNA flap
mismatches
(SEQ ID NO: 391) genomic target sequence
5'- TGGTGCGGTGAATTCTCTCGCCAA... -3'
5'- ...TGGTGGGGTGAATTCTTT-GCCAA... -3'
Intended G•C to C•G edit
Unintended edits from nearly-matched scaffold

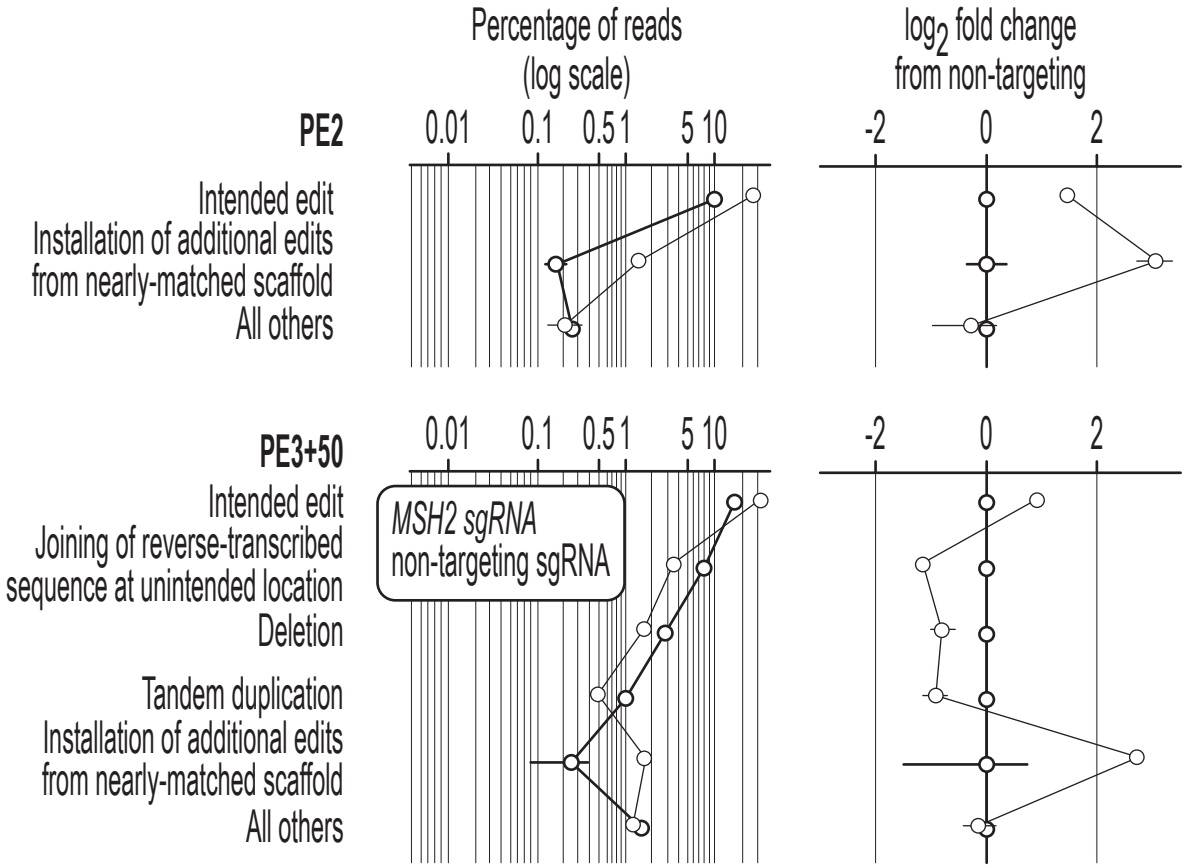
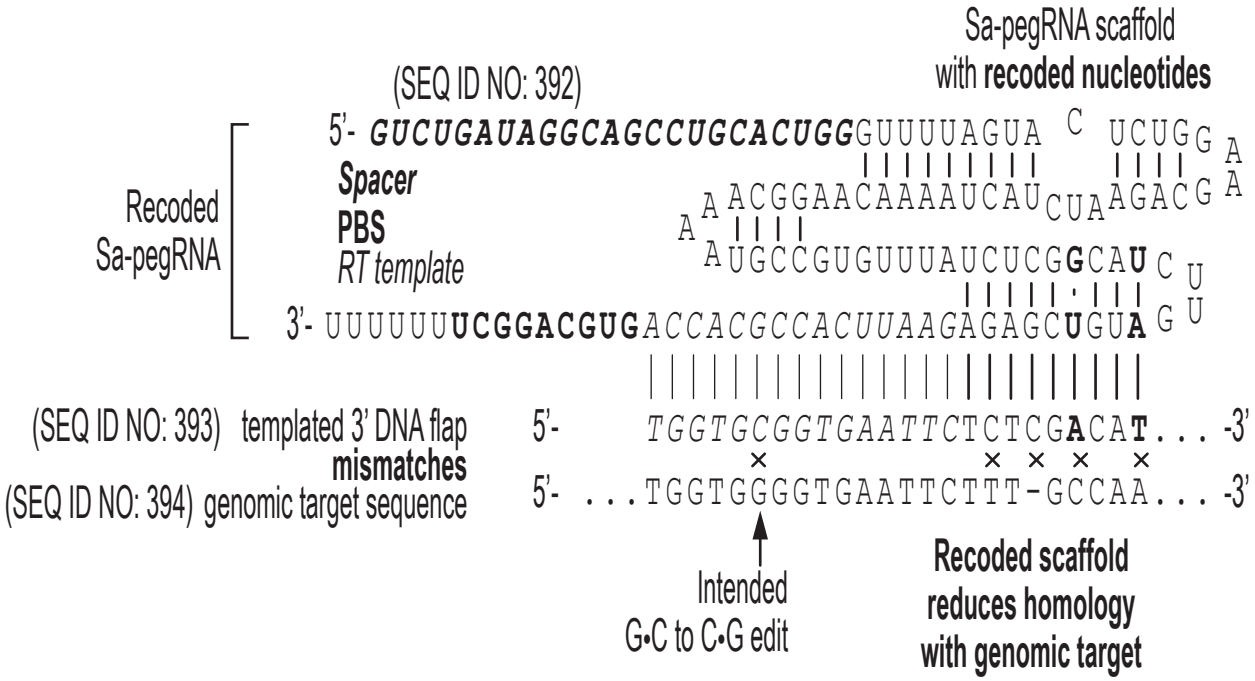

Percentage of reads (log scale)
log2 fold change from non-targeting
PE2
Intended edit
Installation of additional edits from nearly-matched scaffold
All others
0.01 0.1 0.5 1 5 10
-2 0 2
PE3+50
Intended edit
Joining of reverse-transcribed sequence at unintended location
Deletion
Tandem duplication
Installation of additional edits from nearly-matched scaffold
All others
MSH2 sgRNA
non-targeting sgRNA

FIG. 49A


Sa-pegRNA scaffold
with recoded nucleotides
(SEQ ID NO: 392)
5'- GUCUGAUAGGCAGCCUGCACUGGGUUUUAGUA C UCUGG A
Spacer
PBS
RT template
Recoded
Sa-pegRNA
A ACGGAACAAAAUCAU CUA AGACG A
A AUGCCGUGUUUAUCUCGGCAU C U
3'- UUUUUUUCGGACGUGACCACGCCACUUAAGAGAGCUGUA G U
(SEQ ID NO: 393) templated 3' DNA flap
mismatches
(SEQ ID NO: 394) genomic target sequence
5'- TGGTGCGGTGAATTCTCTCGACAT... -3'
5'- ...TGGTGGGGTGAATTCTTT-GCCAA... -3'
Intended
G•C to C•G edit
Recoded scaffold
reduces homology
with genomic target

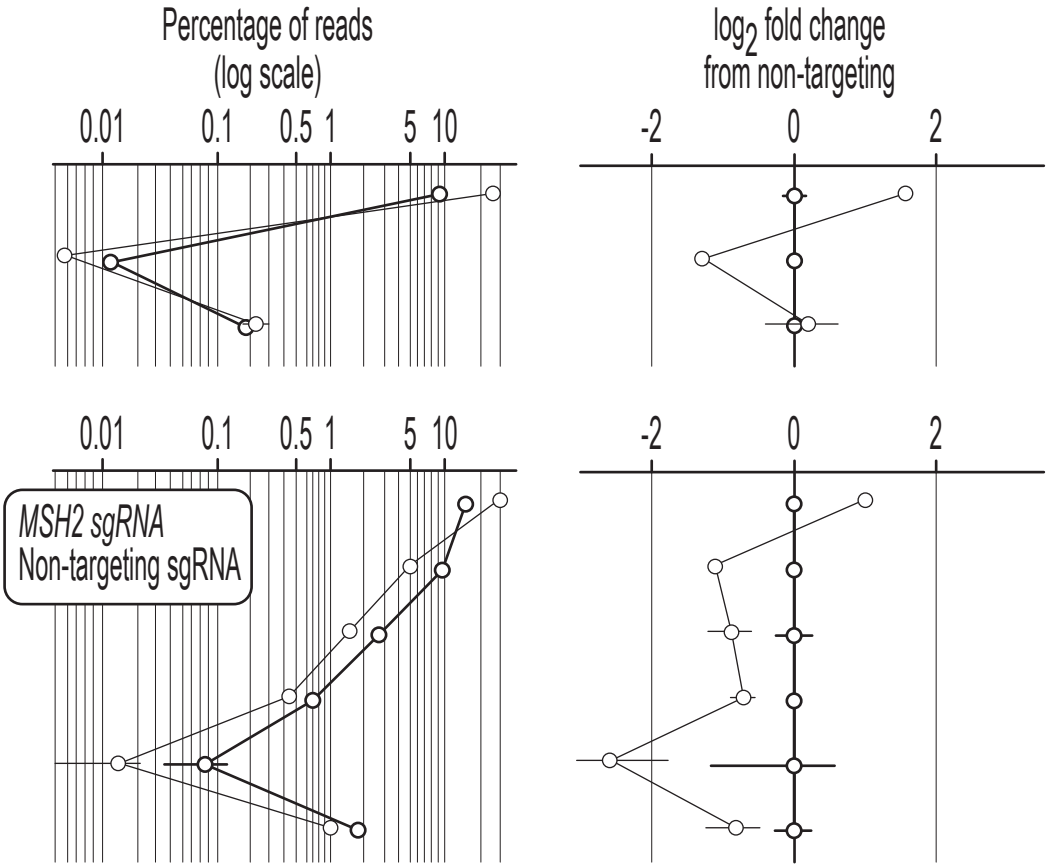
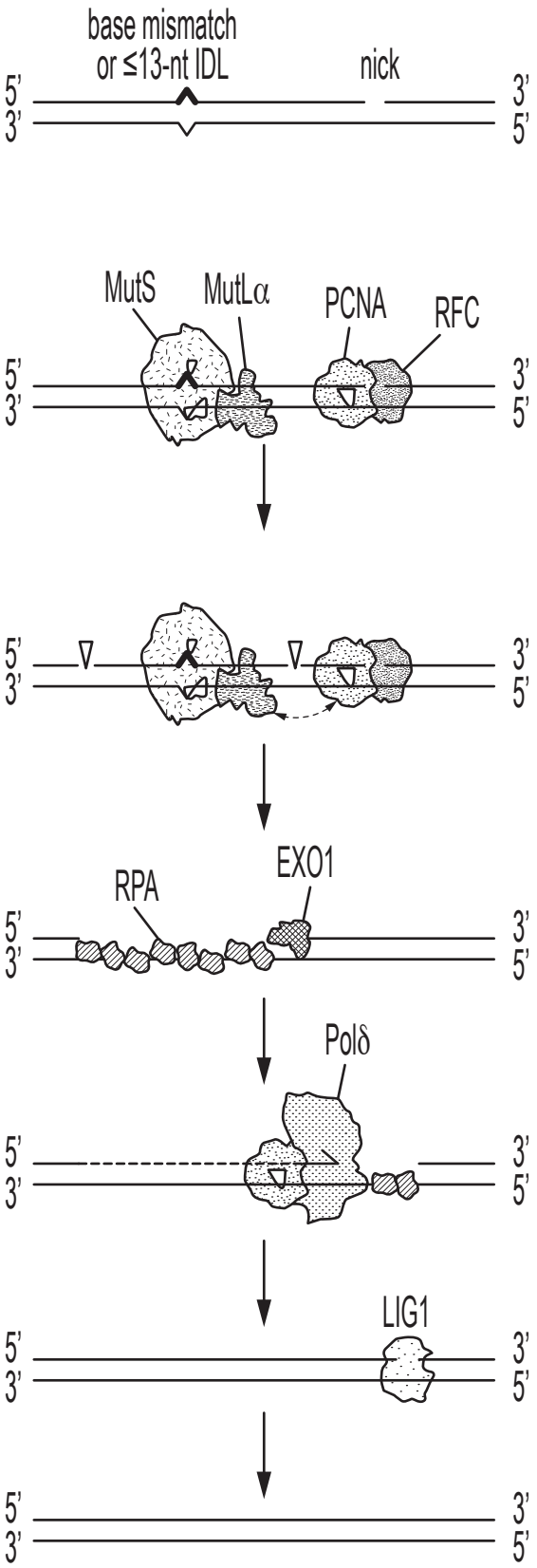

Percentage of reads
(log scale)
0.01
0.1
0.5
1
5
10
log2 fold change
from non-targeting
-2
0
2
MSH2 sgRNA
Non-targeting sgRNA

FIG. 49B



MutLα (MSH2-MSH6) or MutSβ (MSH2-MSH3) recognizes the mismatch or insertion-deletion loop and recruits MutLα (PMS2-MLH1). RFC asymmetrically loads PCNA at the nicked DNA.

PCNA signals MutLα to indiscriminately incise the nick-containing DNA strand.

EXO1 excises the nicked strand from these incisions and the resulting ssDNA is bound by RPA.

Polδ resynthesizes the excised strand from the exposed 3' DNA end, using the non-nicked strand as a template.

LIG1 seals the remaining nick

FIG. 49C

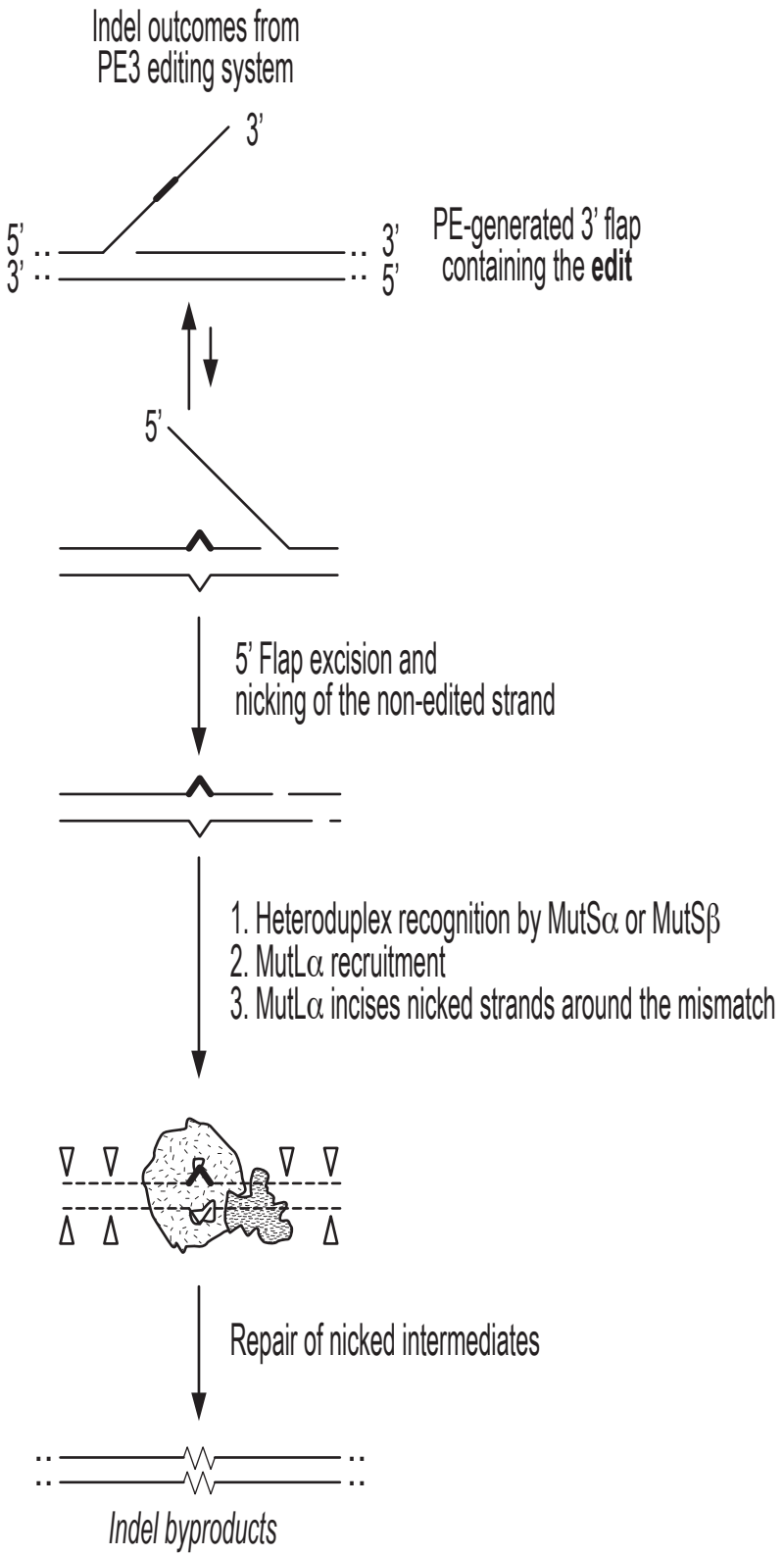

Indel outcomes from
PE3 editing system
3'
5'
3'
3'
5'
PE-generated 3' flap
containing the **edit**
5'
5' Flap excision and
nicking of the non-edited strand
1. Heteroduplex recognition by MutSα or MutSβ
2. MutLα recruitment
3. MutLα incises nicked strands around the mismatch
Repair of nicked intermediates
Indel byproducts

FIG. 49D

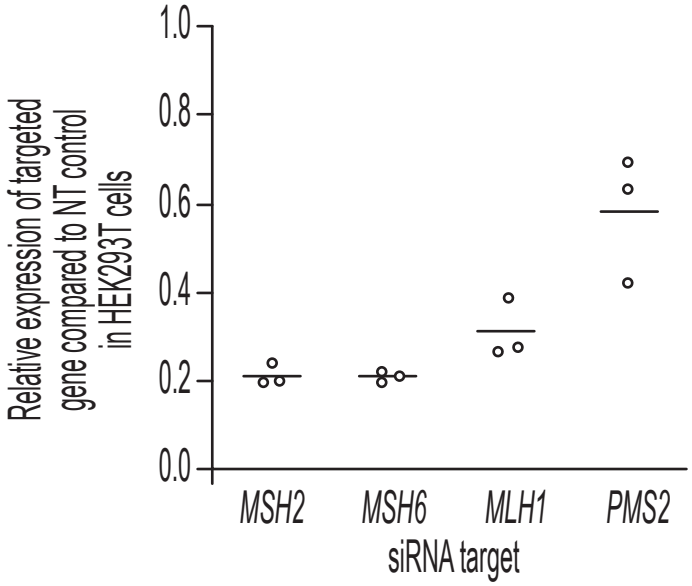


FIG. 49E

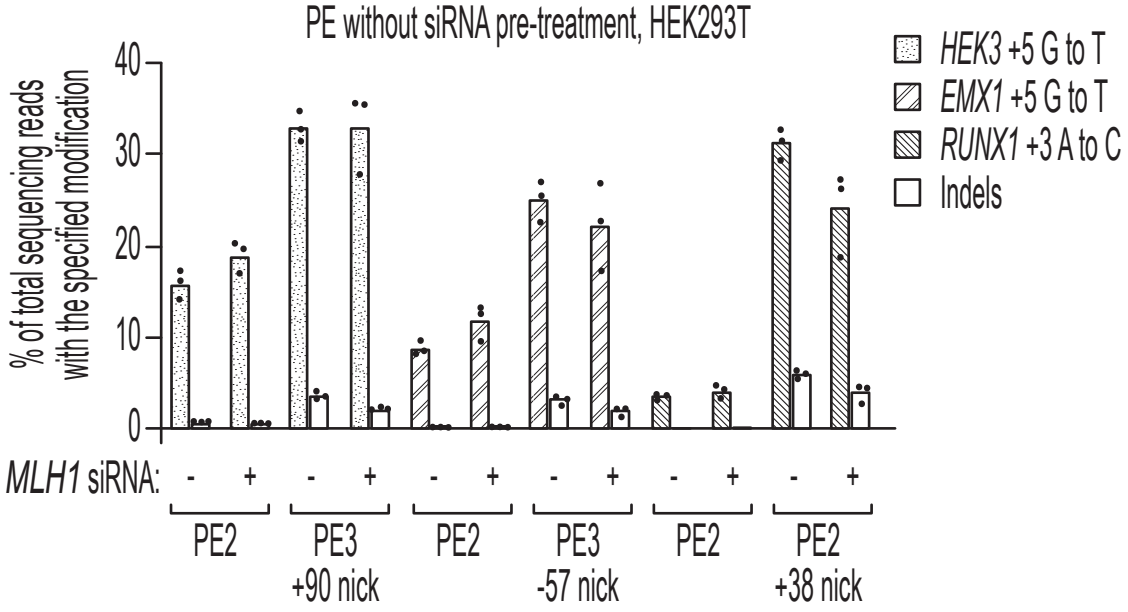


FIG. 49F

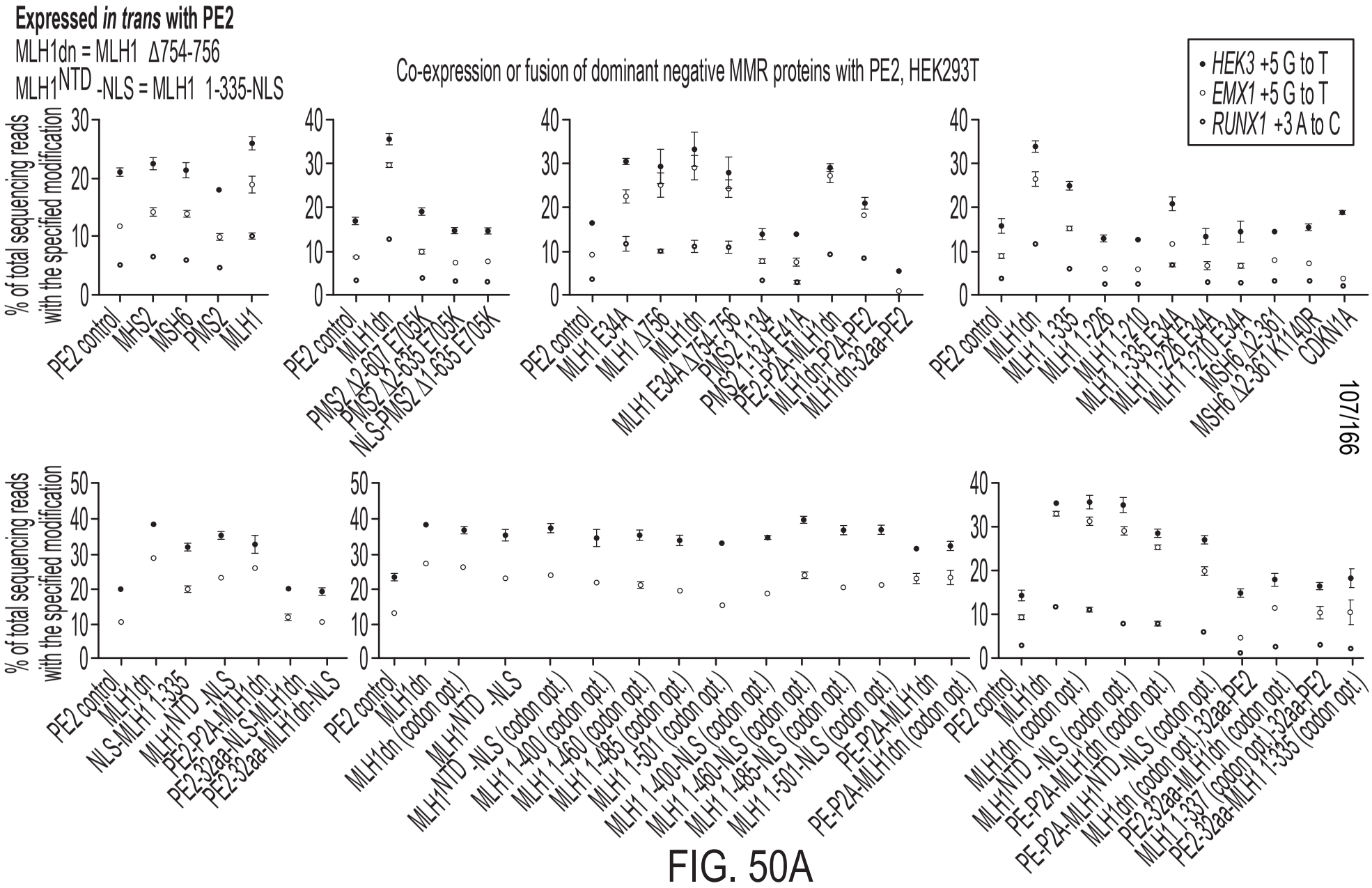

Expressed *in trans* with PE2
MLH1dn = MLH1 Δ754-756
MLH1NTD -NLS = MLH1 1-335-NLS
Co-expression or fusion of dominant negative MMR proteins with PE2, HEK293T
• HEK3 +5 G to T
◦ EMX1 +5 G to T
∘ RUNX1 +3 A to C
% of total sequencing reads with the specified modification
PE2 control
MHS2
MSH6
PMS2
MLH1
MLH1dn
PMS2 Δ2-607 E705K
PMS2 Δ2-635 E705K
NLS-PMS2 Δ1-635 E705K
MLH1 E34A
MLH1 Δ756
MLH1dn
MLH1 E34A Δ754-756
PMS2 1-134
PMS2 1-134 E41A
PE2-P2A-MLH1dn
MLH1dn-P2A-PE2
MLH1dn-32aa-PE2
MLH1 1-335
MLH1 1-226
MLH1 1-210
MLH1 1-335 E34A
MLH1 1-226 E34A
MLH1 1-210 E34A
MSH6 Δ2-361
MSH6 Δ2-361 K1140R
CDKN1A
NLS-MLH1 1-335
MLH1NTD -NLS
PE2-P2A-MLH1dn
PE2-32aa-NLS-MLH1dn
PE2-32aa-MLH1dn-NLS
MLH1dn (codon opt.)
MLH1NTD -NLS (codon opt.)
MLH1 1-400 (codon opt.)
MLH1 1-460 (codon opt.)
MLH1 1-485 (codon opt.)
MLH1 1-501 (codon opt.)
MLH1 1-400-NLS (codon opt.)
MLH1 1-460-NLS (codon opt.)
MLH1 1-485-NLS (codon opt.)
MLH1 1-501-NLS (codon opt.)
PE-P2A-MLH1dn
PE-P2A-MLH1dn (codon opt.)
PE-P2A-MLH1NTD -NLS (codon opt.)
MLH1dn (codon opt.)-32aa-PE2
MLH1 1-337 (codon opt.)-32aa-PE2
MLH1 1-335 (codon opt.)

FIG. 50A

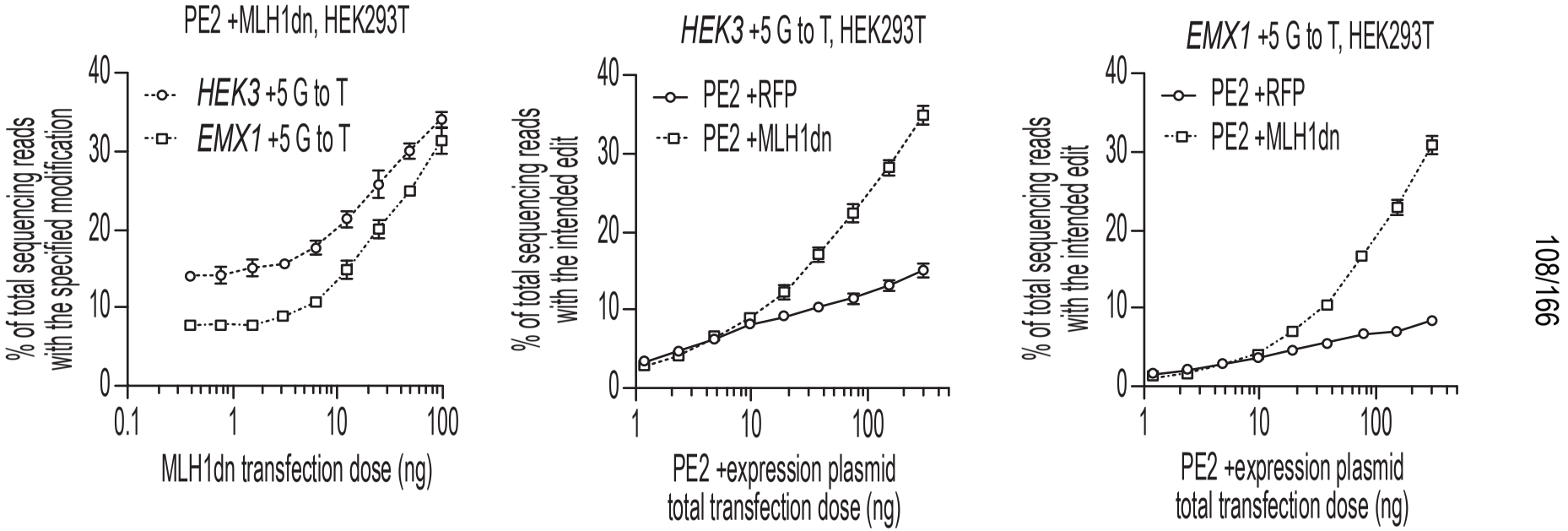

PE2 +MLH1dn, HEK293T
HEK3 +5 G to T
EMX1 +5 G to T
% of total sequencing reads with the specified modification
MLH1dn transfection dose (ng)
HEK3 +5 G to T, HEK293T
PE2 +RFP
PE2 +MLH1dn
% of total sequencing reads with the intended edit
PE2 +expression plasmid total transfection dose (ng)
EMX1 +5 G to T, HEK293T
PE2 +RFP
PE2 +MLH1dn
% of total sequencing reads with the intended edit
PE2 +expression plasmid total transfection dose (ng)

FIG. 50B

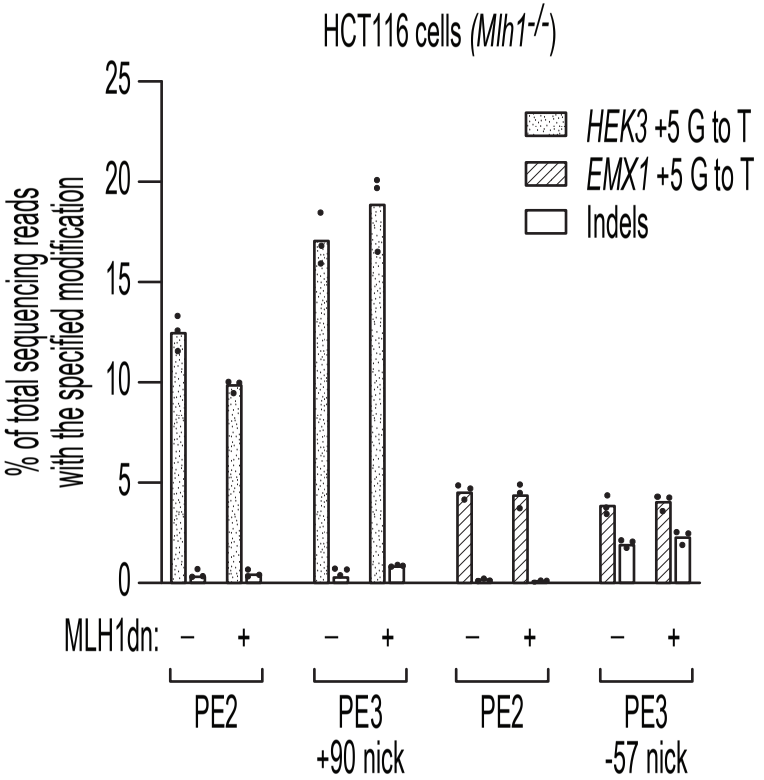


FIG. 50C

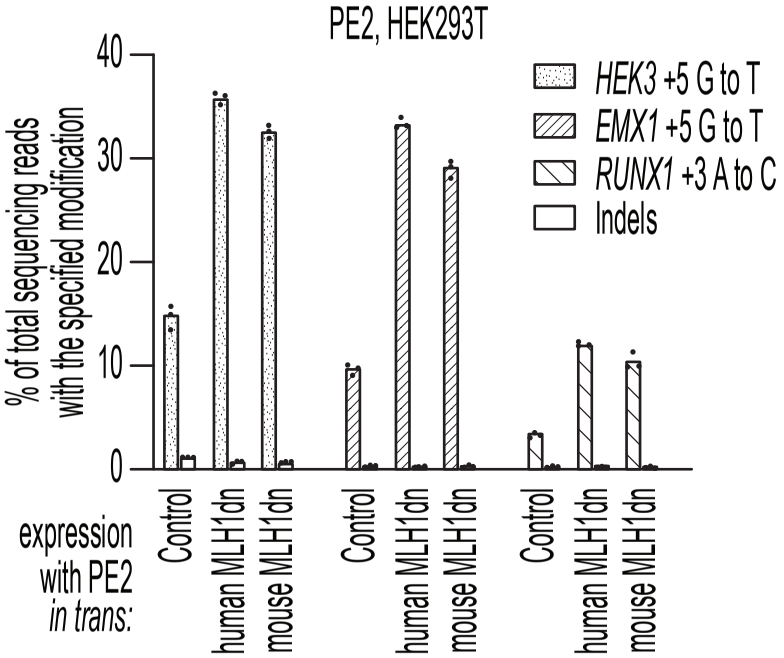


FIG. 50D

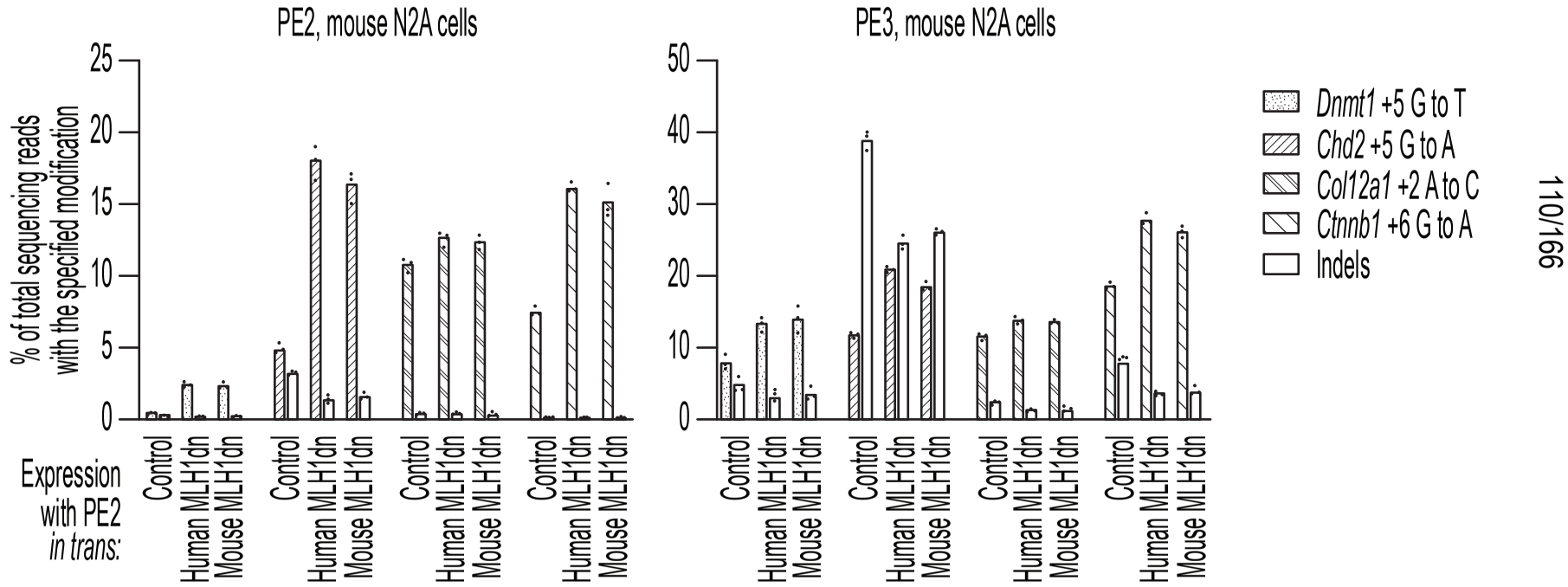

PE2, mouse N2A cells
PE3, mouse N2A cells
% of total sequencing reads with the specified modification
25
20
15
10
5
0
50
40
30
20
10
0
Expression with PE2 in trans:
Control
Human MLH1dn
Mouse MLH1dn
Dnmt1 +5 G to T
Chd2 +5 G to A
Col12a1 +2 A to C
Ctnnb1 +6 G to A
Indels

FIG. 50E

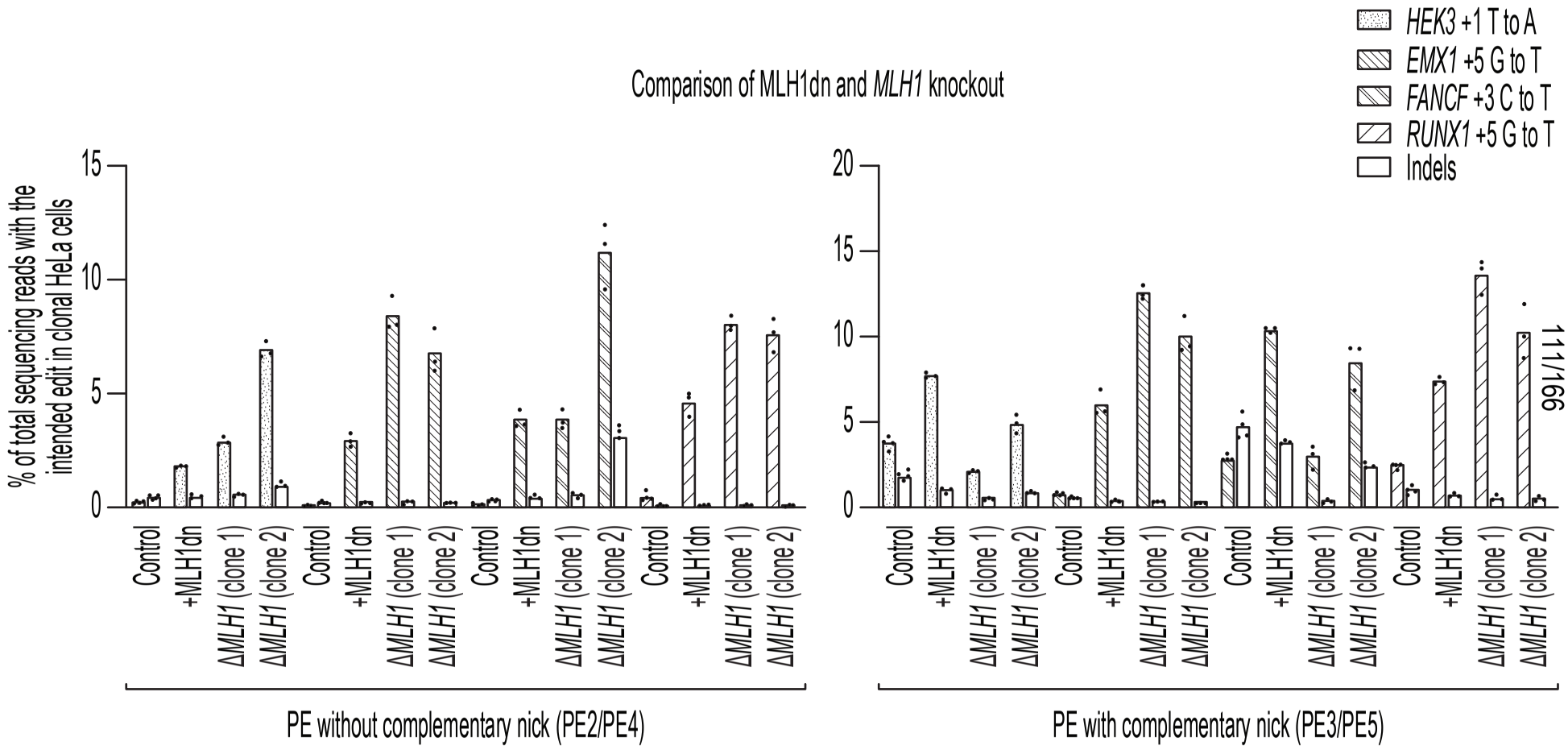

Comparison of MLH1dn and MLH1 knockout
HEK3 +1 T to A
EMX1 +5 G to T
FANCF +3 C to T
RUNX1 +5 G to T
Indels
% of total sequencing reads with the intended edit in clonal HeLa cells
0
5
10
15
20
Control
+MLH1dn
ΔMLH1 (clone 1)
ΔMLH1 (clone 2)
PE without complementary nick (PE2/PE4)
PE with complementary nick (PE3/PE5)
111/166

FIG. 50F

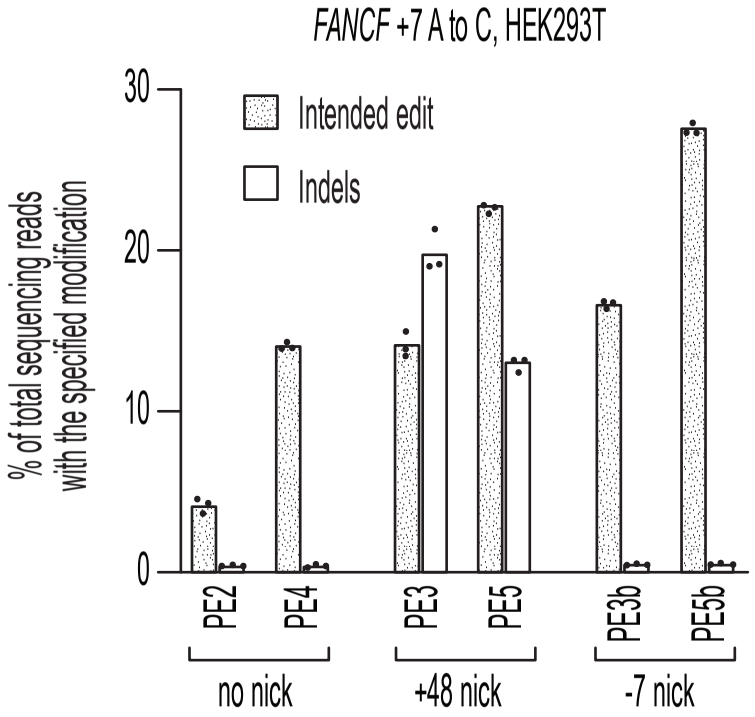


FIG. 50G

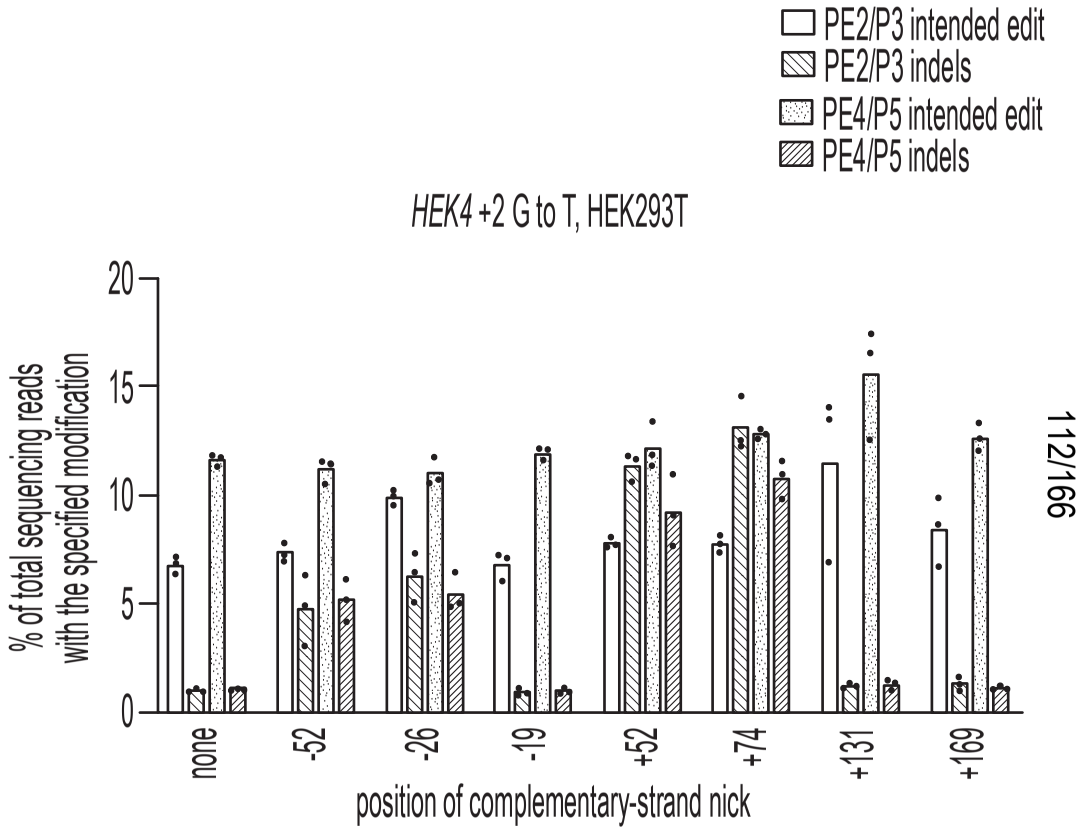


FIG. 50H

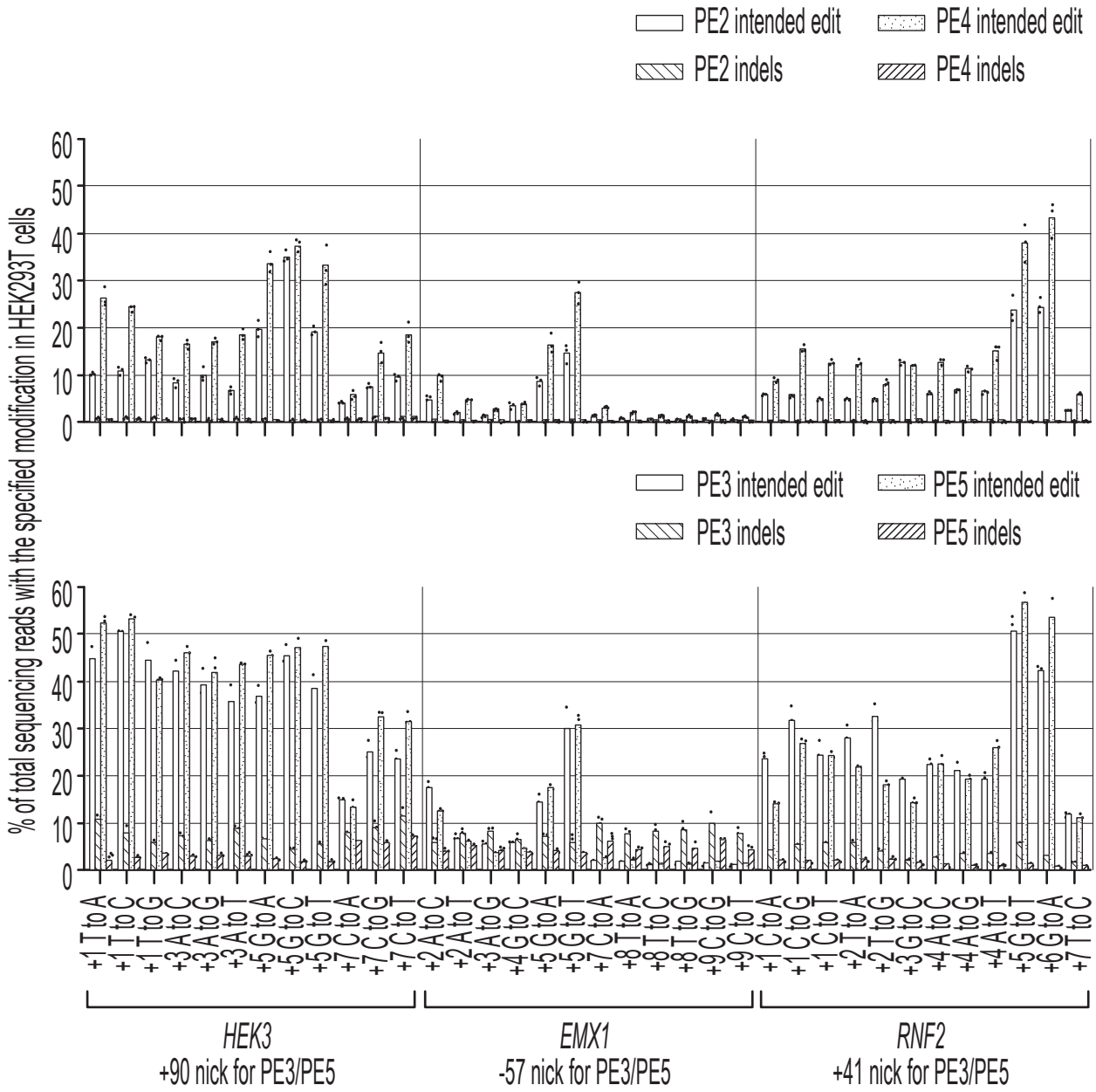

PE2 intended edit
PE4 intended edit
PE2 indels
PE4 indels
PE3 intended edit
PE5 intended edit
PE3 indels
PE5 indels
% of total sequencing reads with the specified modification in HEK293T cells
0
10
20
30
40
50
60
+1T to A
+1T to C
+1T to G
+3A to C
+3A to G
+3A to T
+5G to A
+5G to C
+5G to T
+7C to A
+7C to G
+7C to T
+2A to C
+2A to T
+2A to G
+3A to G
+4G to C
+5G to A
+5G to T
+7C to A
+8T to A
+8T to C
+8T to G
+9C to T
+1C to A
+1C to G
+1C to T
+2T to A
+2T to G
+3G to C
+4A to C
+4A to G
+4A to T
+5G to T
+6G to A
+7 to C
HEK3
+90 nick for PE3/PE5
EMX1
-57 nick for PE3/PE5
RNF2
+41 nick for PE3/PE5

FIG. 51A

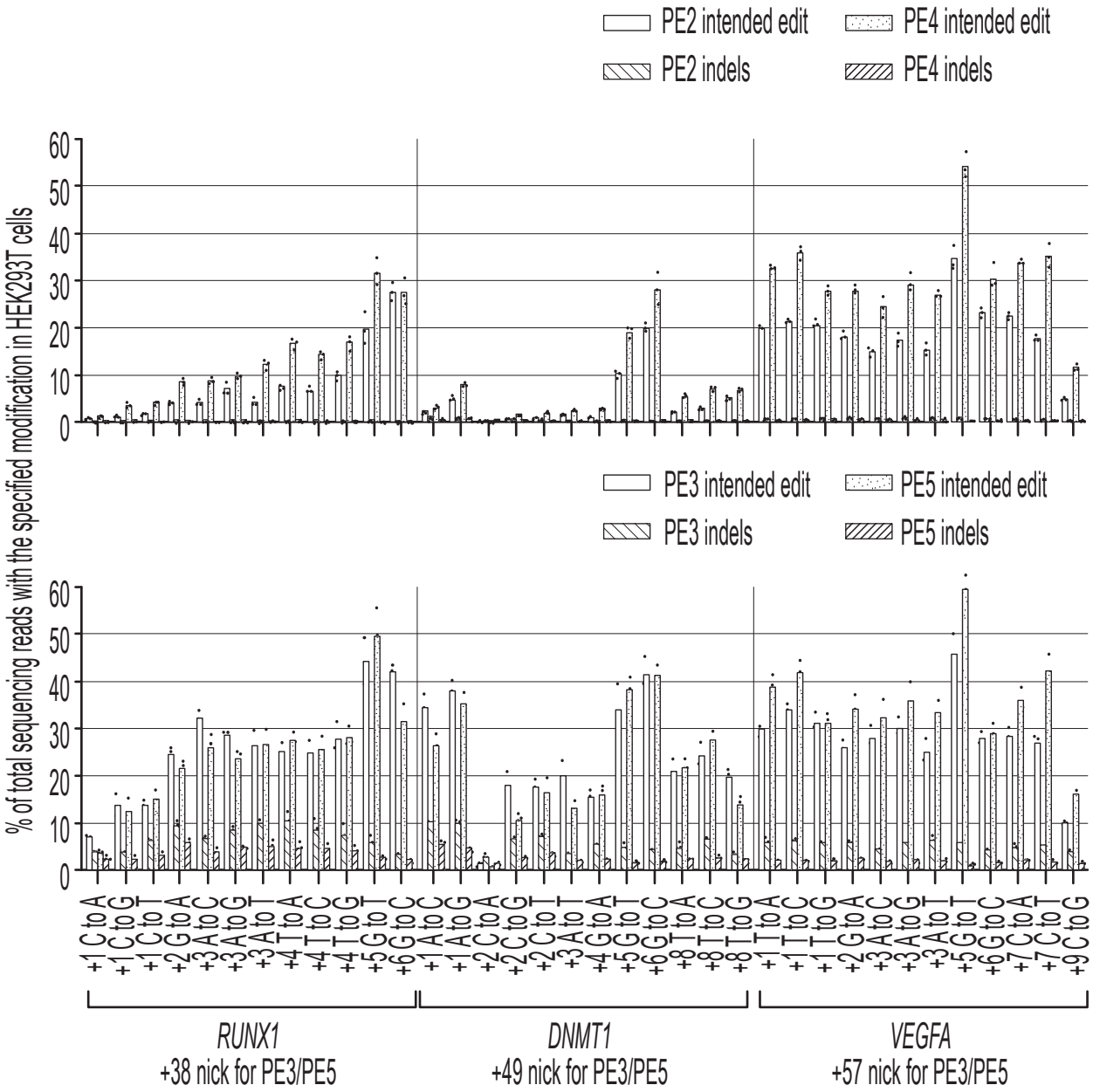


FIG. 51A CONTINUED

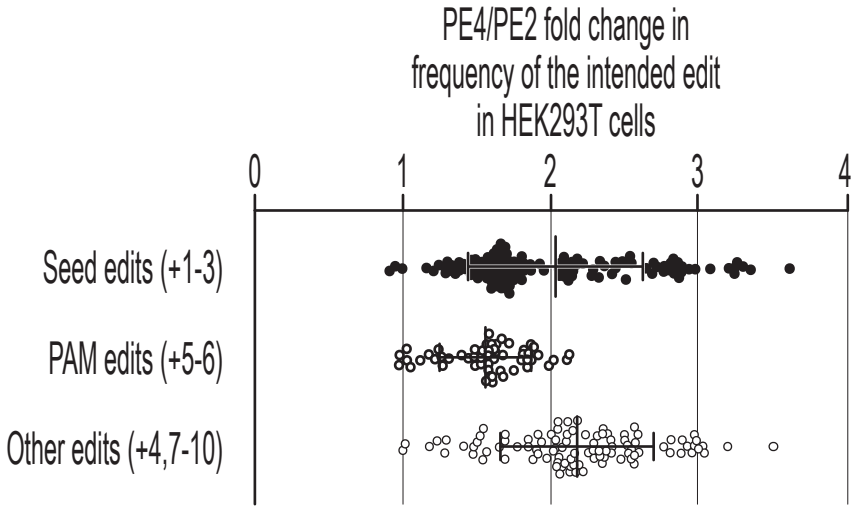


FIG. 51B

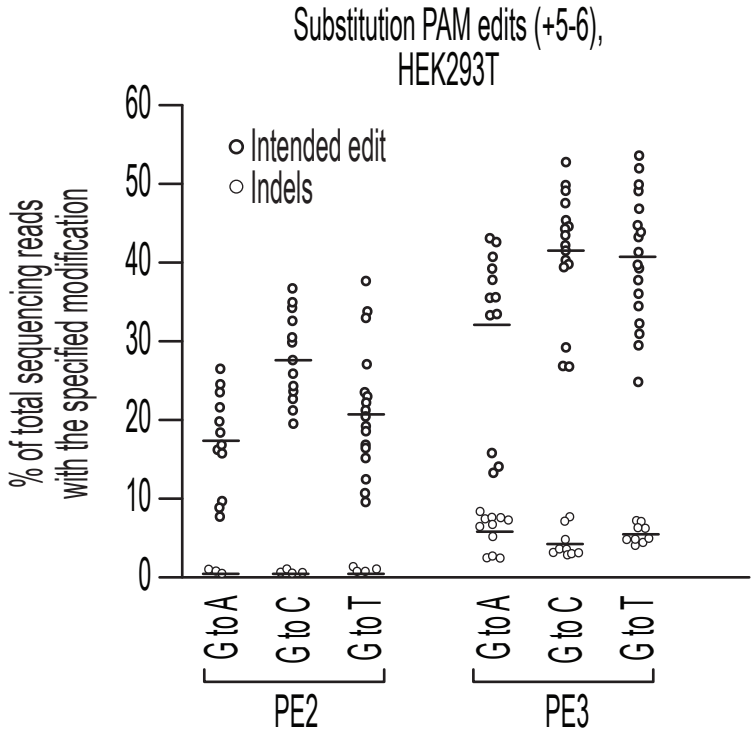


FIG. 51C

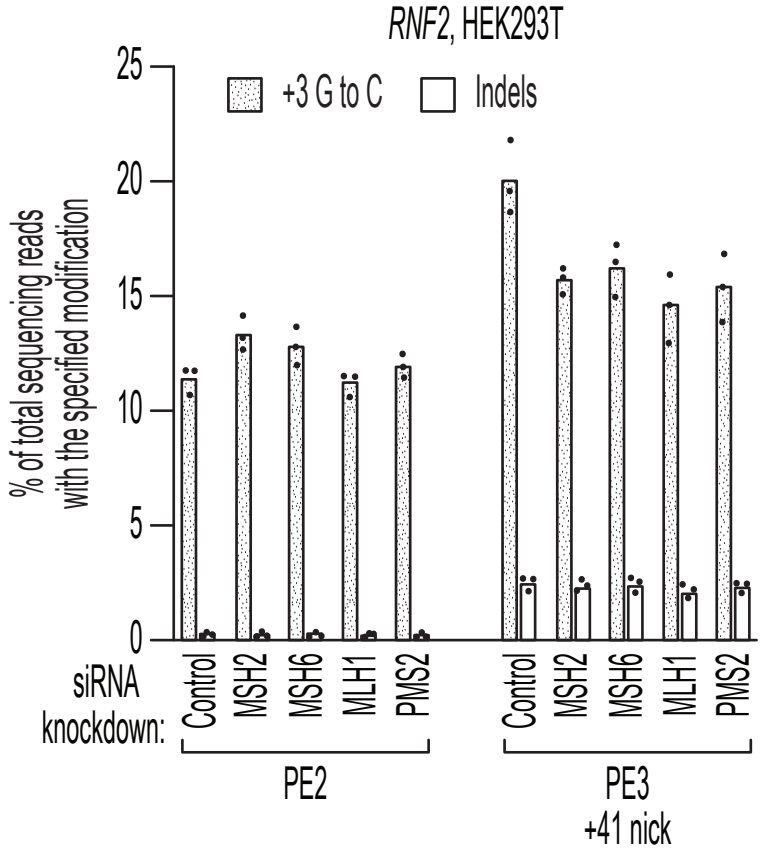


FIG. 51D

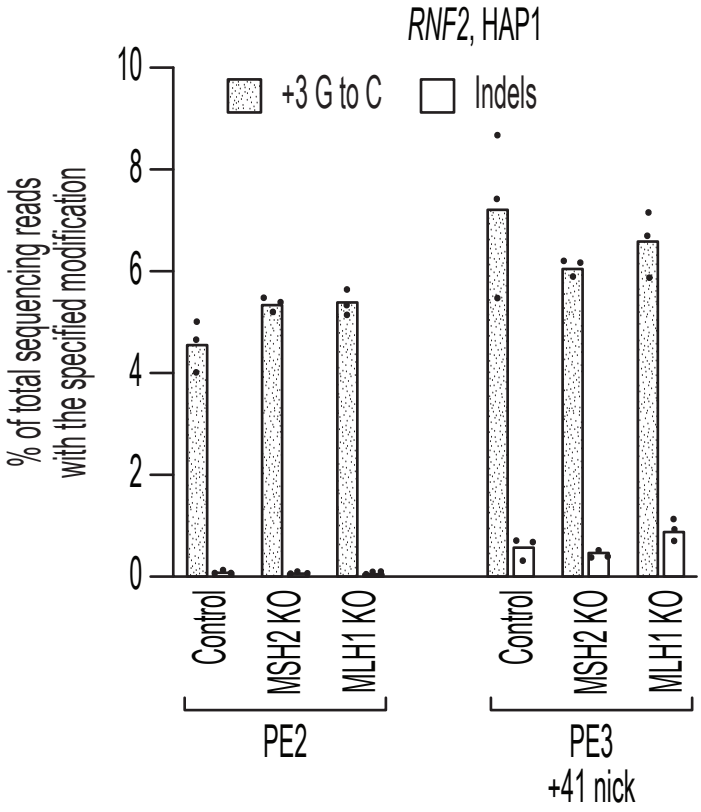


FIG. 51E

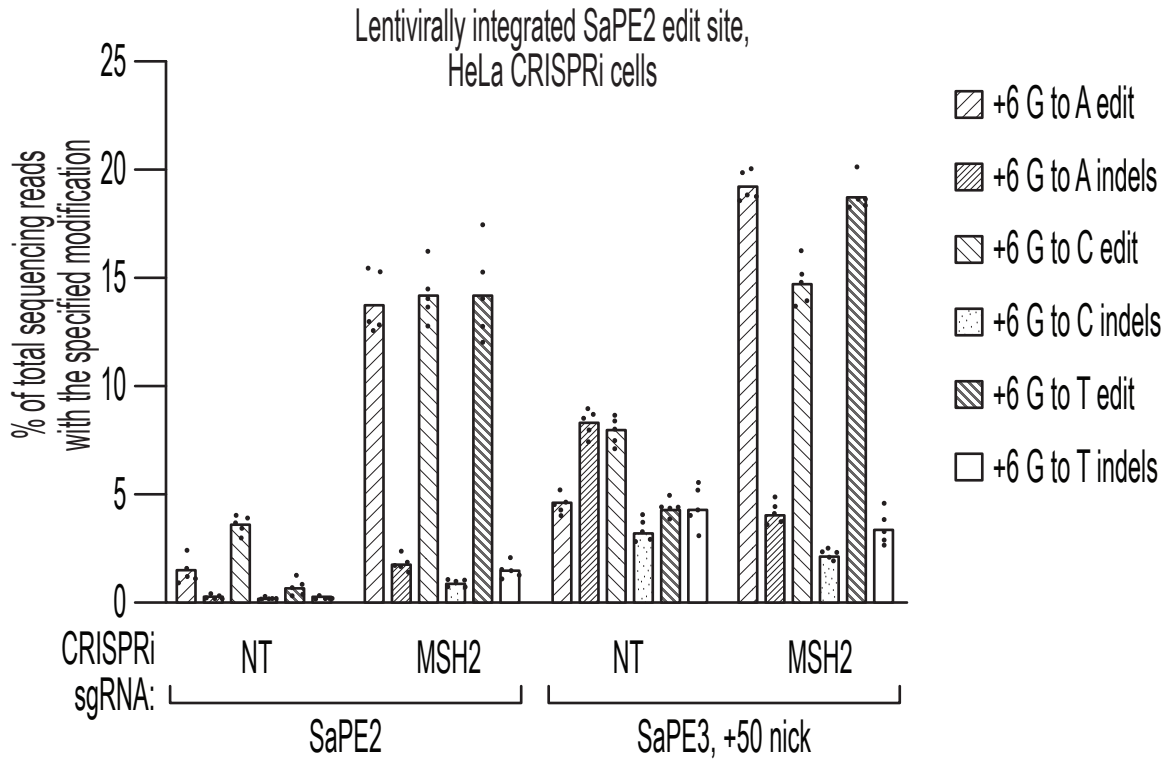


FIG. 51F

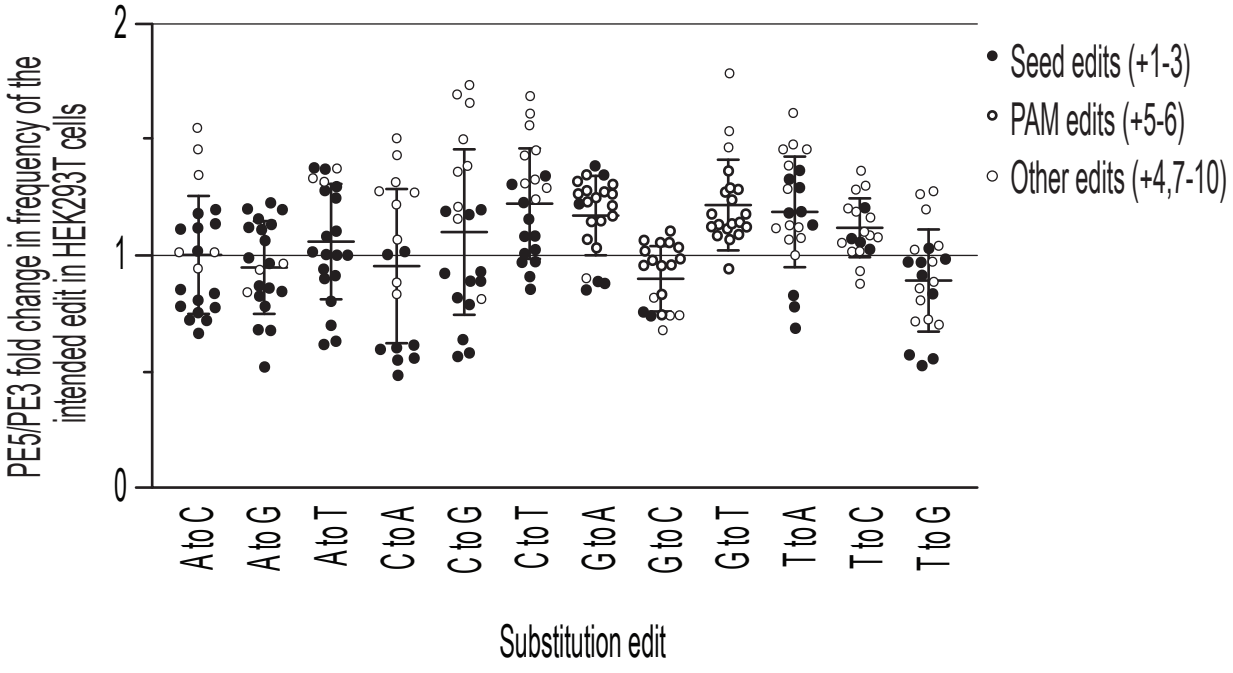


FIG. 51G

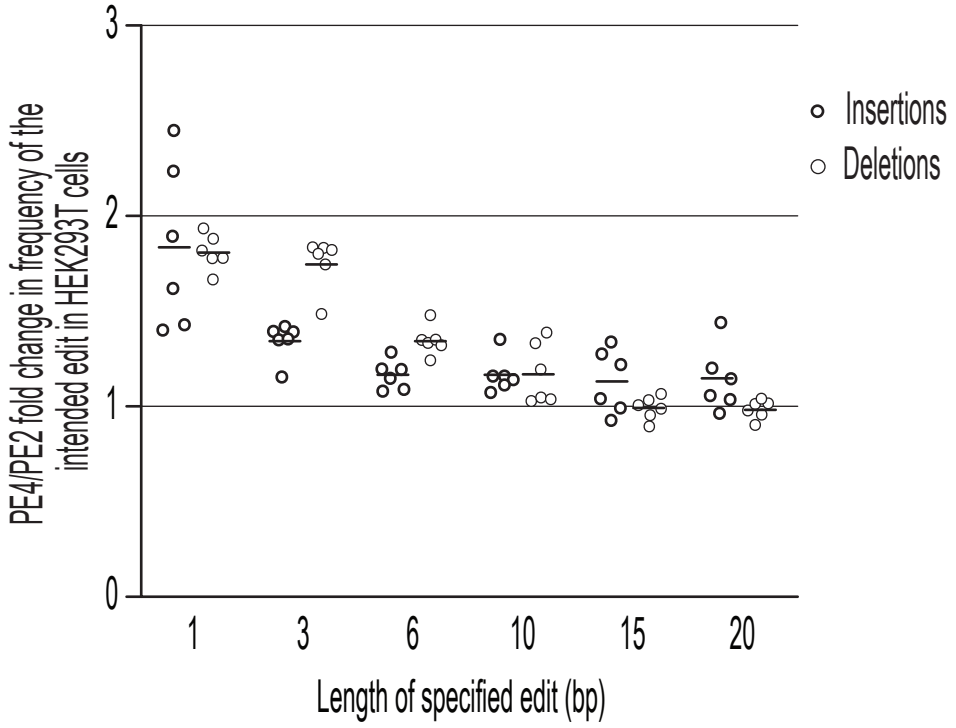

PE4/PE2 fold change in frequency of the intended edit in HEK293T cells
3
2
1
0
1
3
6
10
15
20
Length of specified edit (bp)
Insertions
Deletions

FIG. 51H

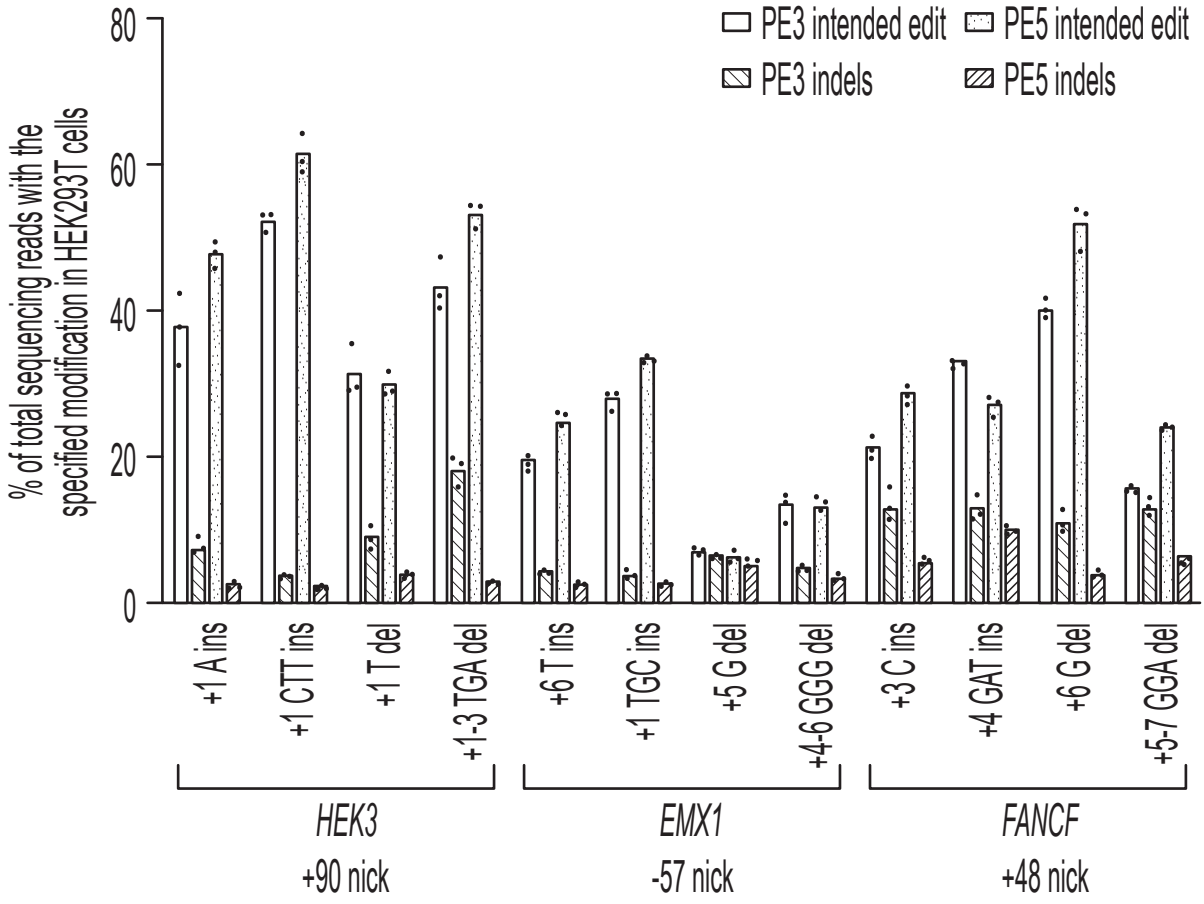


FIG. 51I

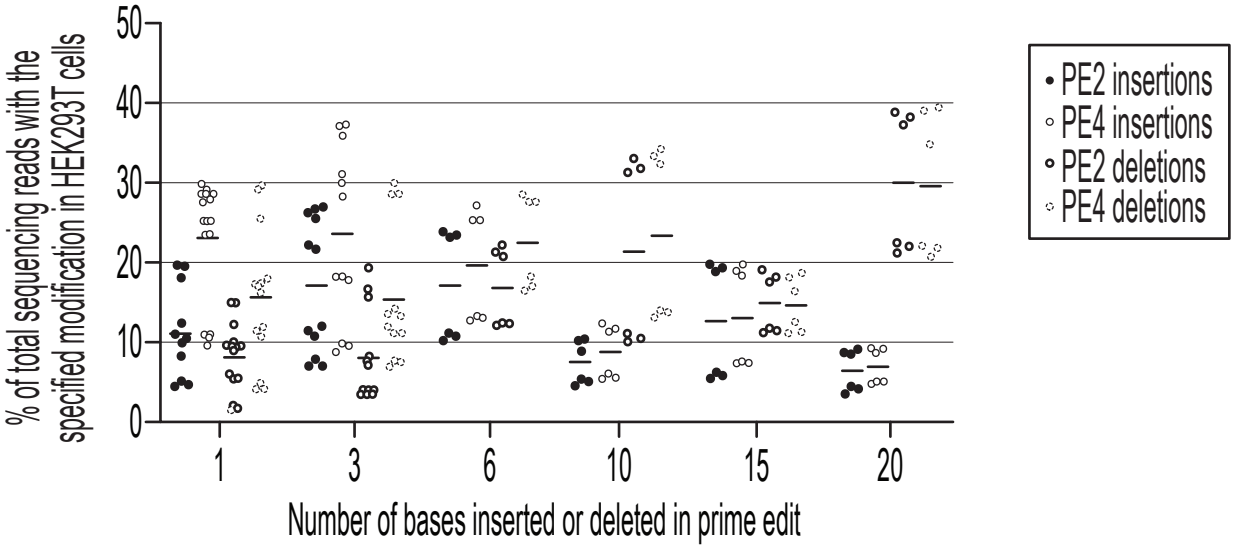


FIG. 51J

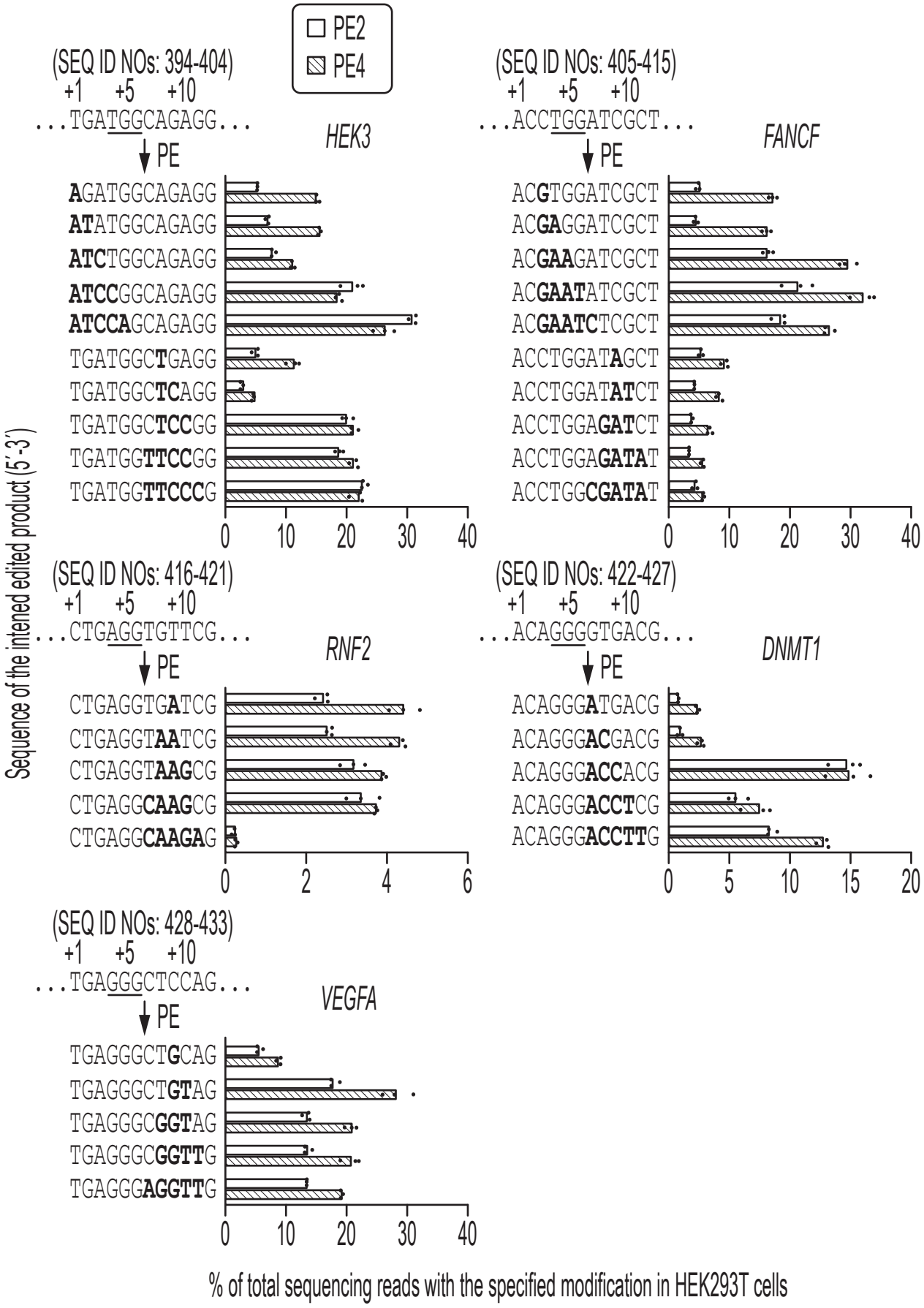

PE2
PE4
(SEQ ID NOs: 394-404)
+1 +5 +10
...TGATGGCAGAGG...
PE
HEK3
AGATGGCAGAGG
ATATGGCAGAGG
ATCTGGCAGAGG
ATCCGGCAGAGG
ATCCAGCAGAGG
TGATGGCTGAGG
TGATGGCTCAGG
TGATGGCTCCGG
TGATGGTTCCGG
TGATGGTTCCCG
0 10 20 30 40
(SEQ ID NOs: 405-415)
+1 +5 +10
...ACCTGGATCGCT...
PE
FANCF
ACGTGGATCGCT
ACGAGGATCGCT
ACGAAGATCGCT
ACGAATATCGCT
ACGAATCTCGCT
ACCTGGATAGCT
ACCTGGATATCT
ACCTGGAGATCT
ACCTGGAGATAT
ACCTGGCGATAT
0 10 20 30 40
(SEQ ID NOs: 416-421)
+1 +5 +10
...CTGAGGTGTTCG...
PE
RNF2
CTGAGGTGATCG
CTGAGGTAATCG
CTGAGGTAAGCG
CTGAGGCAAGCG
CTGAGGCAAGAG
0 2 4 6
(SEQ ID NOs: 422-427)
+1 +5 +10
...ACAGGGGTGACG...
PE
DNMT1
ACAGGGATGACG
ACAGGGACGACG
ACAGGGACCACG
ACAGGGACCTCG
ACAGGGACCTTG
0 5 10 15 20
(SEQ ID NOs: 428-433)
+1 +5 +10
...TGAGGGCTCCAG...
PE
VEGFA
TGAGGGCTGCAG
TGAGGGCTGTAG
TGAGGGCGGTAG
TGAGGGCGGTTG
TGAGGGAGGTTG
0 10 20 30 40
Sequence of the intened edited product (5'-3')
% of total sequencing reads with the specified modification in HEK293T cells

FIG. 52A

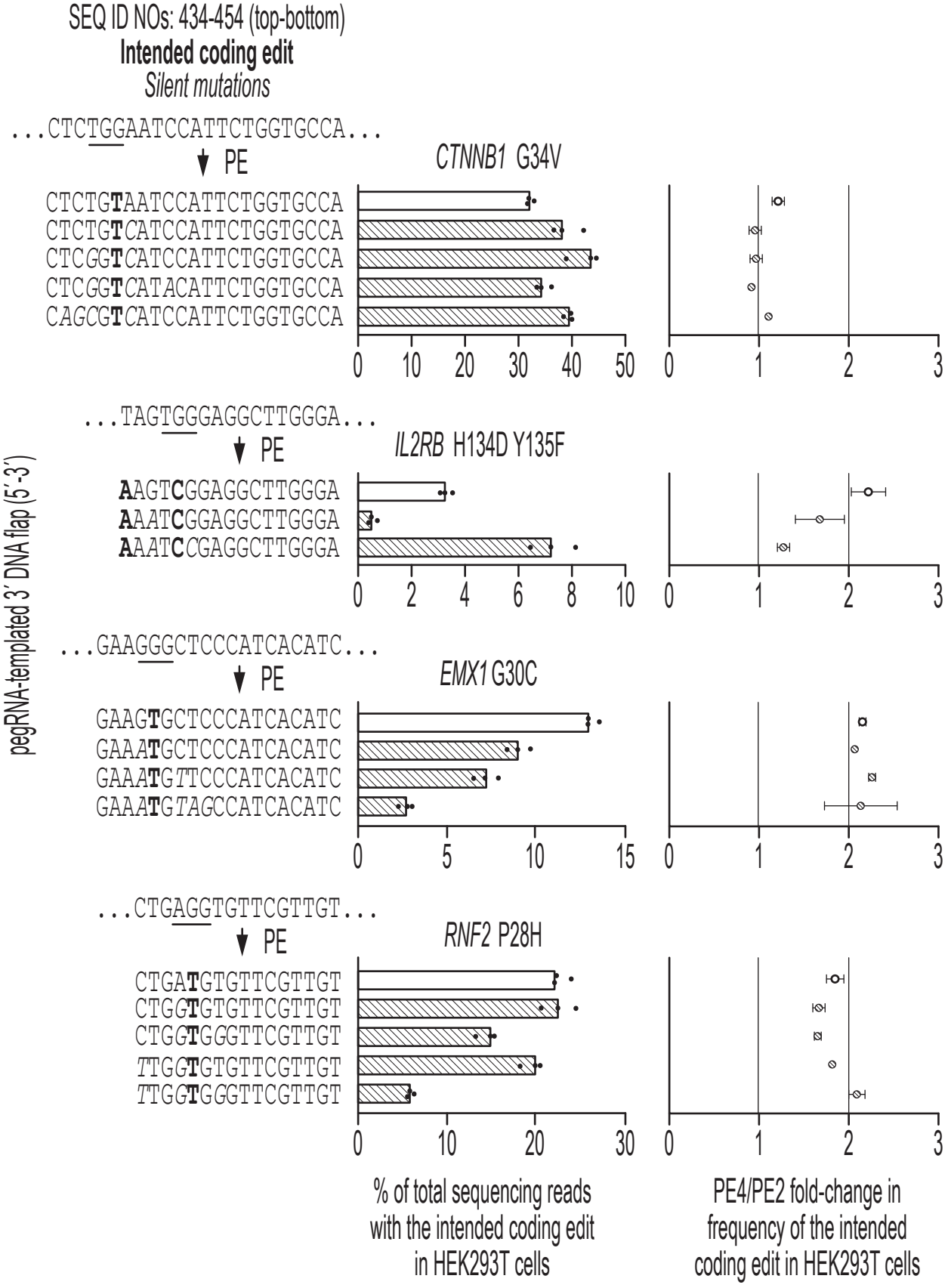

SEQ ID NOs: 434-454 (top-bottom)
Intended coding edit
Silent mutations
...CTCTGGAATCCATTCTGGTGCCA...
PE
CTCTGTAATCCATTCTGGTGCCA
CTCTGTCATCCATTCTGGTGCCA
CTCGGTCATCCATTCTGGTGCCA
CTCGGTCATACATTCTGGTGCCA
CAGCGTCATCCATTCTGGTGCCA
CTNNB1 G34V
...TAGTGGGAGGCTTGGGA...
PE
AAGTCGGAGGCTTGGGA
AAATCGGAGGCTTGGGA
AAATCCGAGGCTTGGGA
IL2RB H134D Y135F
...GAAGGGCTCCCATCACATC...
PE
GAAGTGCTCCCATCACATC
GAAATGCTCCCATCACATC
GAAATGTTCCCATCACATC
GAAATGTAGCCATCACATC
EMX1 G30C
...CTGAGGTGTTCGTTGT...
PE
CTGATGTGTTCGTTGT
CTGGTGTGTTCGTTGT
CTGGTGGGTTCGTTGT
TTGGTGTGTTCGTTGT
TTGGTGGGTTCGTTGT
RNF2 P28H
pegRNA-templated 3' DNA flap (5'-3')
0 10 20 30 40 50
0 1 2 3
0 2 4 6 8 10
0 1 2 3
0 5 10 15
0 1 2 3
0 10 20 30
0 1 2 3
% of total sequencing reads with the intended coding edit in HEK293T cells
PE4/PE2 fold-change in frequency of the intended coding edit in HEK293T cells

FIG. 52B

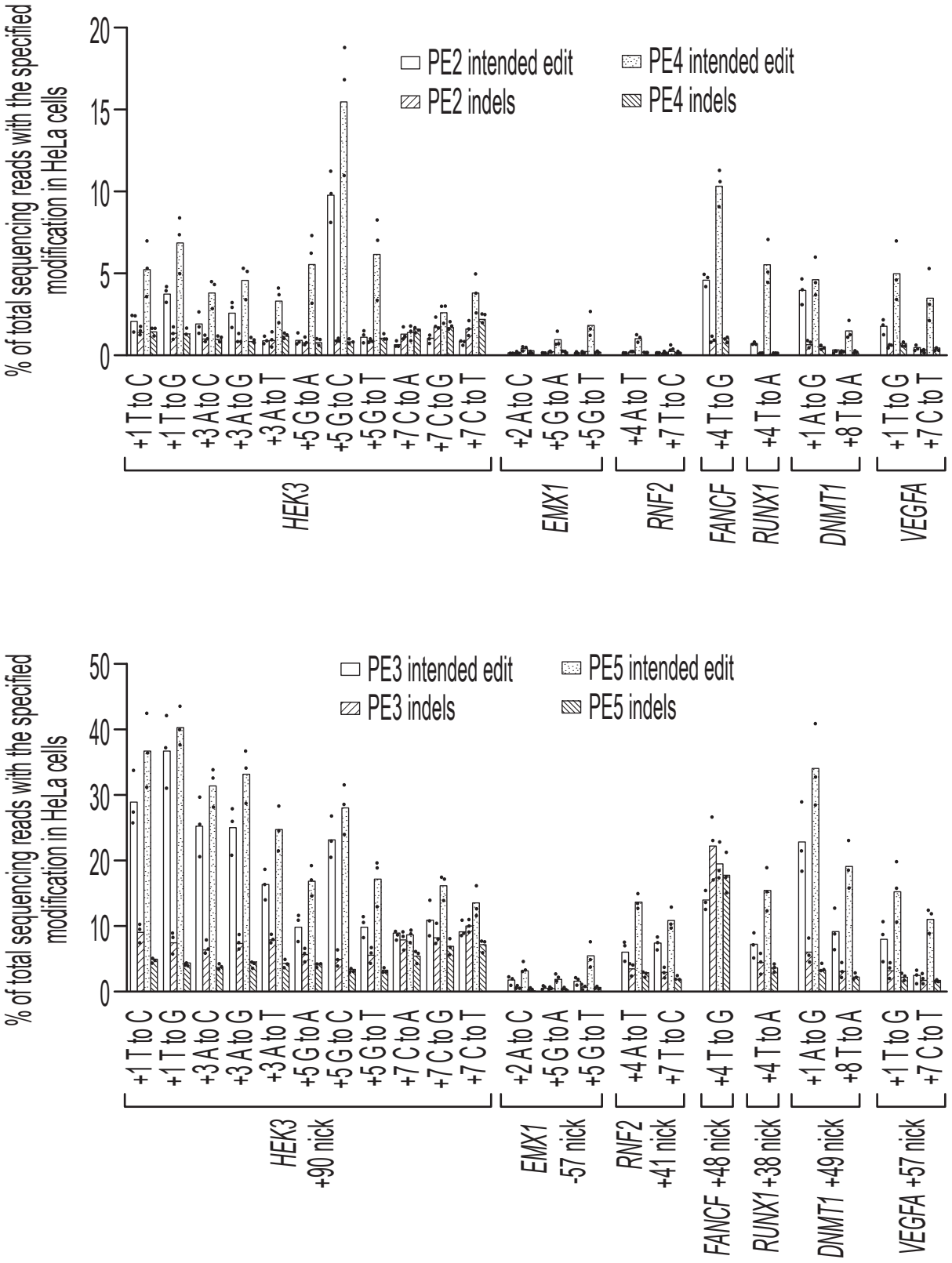

% of total sequencing reads with the specified modification in HeLa cells
PE2 intended edit
PE2 indels
PE4 intended edit
PE4 indels
PE3 intended edit
PE3 indels
PE5 intended edit
PE5 indels
+1 T to C
+1 T to G
+3 A to C
+3 A to G
+3 A to T
+5 G to A
+5 G to C
+5 G to T
+7 C to A
+7 C to G
+7 C to T
+2 A to C
+5 G to A
+5 G to T
+4 A to T
+7 T to C
+4 T to G
+4 T to A
+1 A to G
+8 T to A
+1 T to G
+7 C to T
HEK3
EMX1
RNF2
FANCF
RUNX1
DNMT1
VEGFA
HEK3 +90 nick
EMX1 -57 nick
RNF2 +41 nick
FANCF +48 nick
RUNX1 +38 nick
DNMT1 +49 nick
VEGFA +57 nick

FIG. 52C

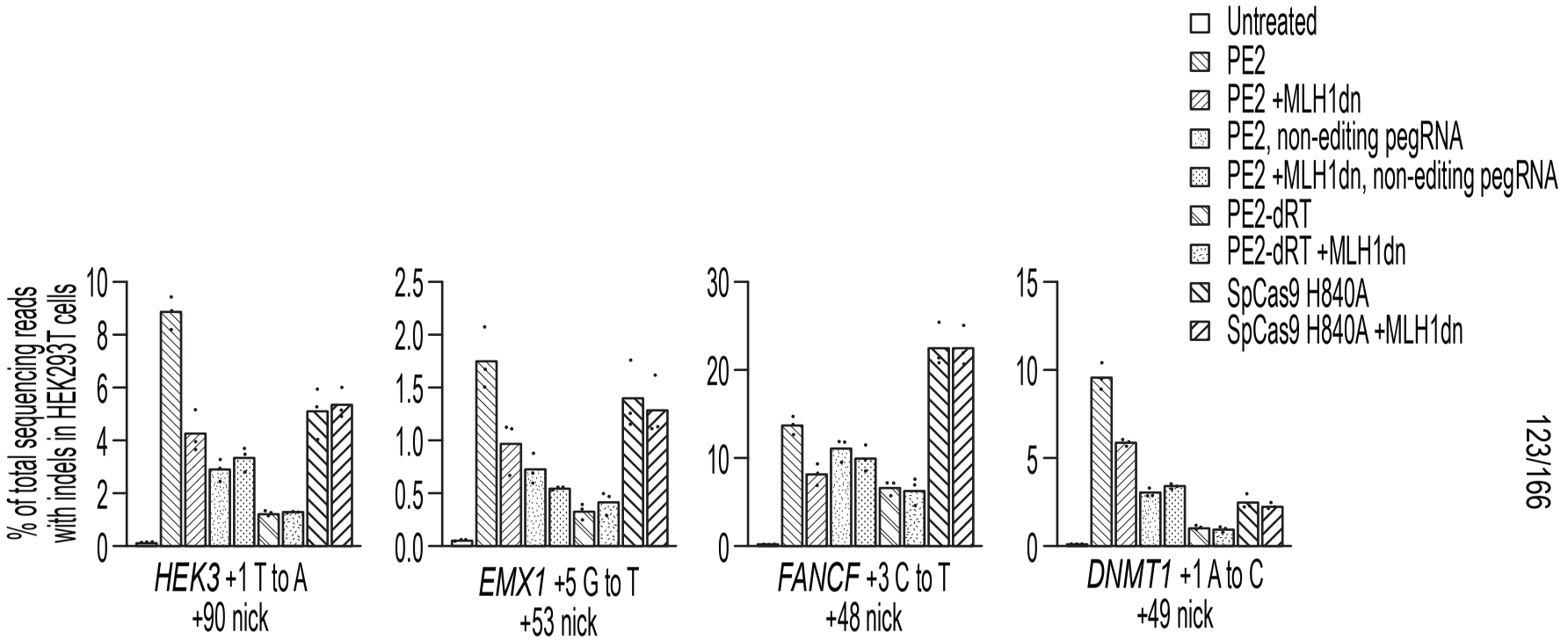

% of total sequencing reads with indels in HEK293T cells
10
8
6
4
2
0
HEK3 +1 T to A
+90 nick
2.5
2.0
1.5
1.0
0.5
0.0
EMX1 +5 G to T
+53 nick
30
20
10
0
FANCF +3 C to T
+48 nick
15
10
5
0
DNMT1 +1 A to C
+49 nick
Untreated
PE2
PE2 +MLH1dn
PE2, non-editing pegRNA
PE2 +MLH1dn, non-editing pegRNA
PE2-dRT
PE2-dRT +MLH1dn
SpCas9 H840A
SpCas9 H840A +MLH1dn

FIG. 53A

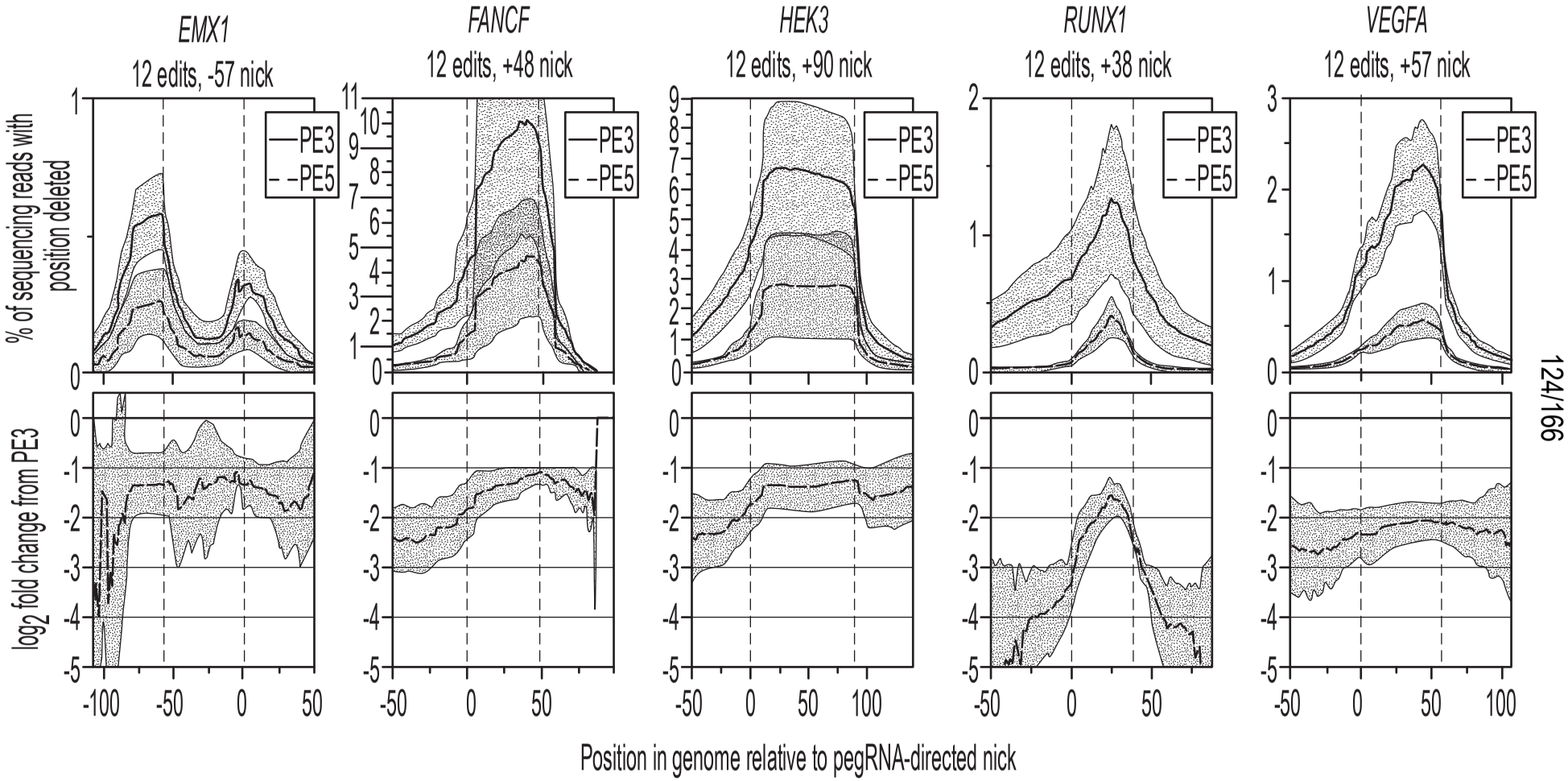

EMX1
12 edits, -57 nick
FANCF
12 edits, +48 nick
HEK3
12 edits, +90 nick
RUNX1
12 edits, +38 nick
VEGFA
12 edits, +57 nick
PE3
PE5
% of sequencing reads with position deleted
log2 fold change from PE3
Position in genome relative to pegRNA-directed nick

FIG. 53B

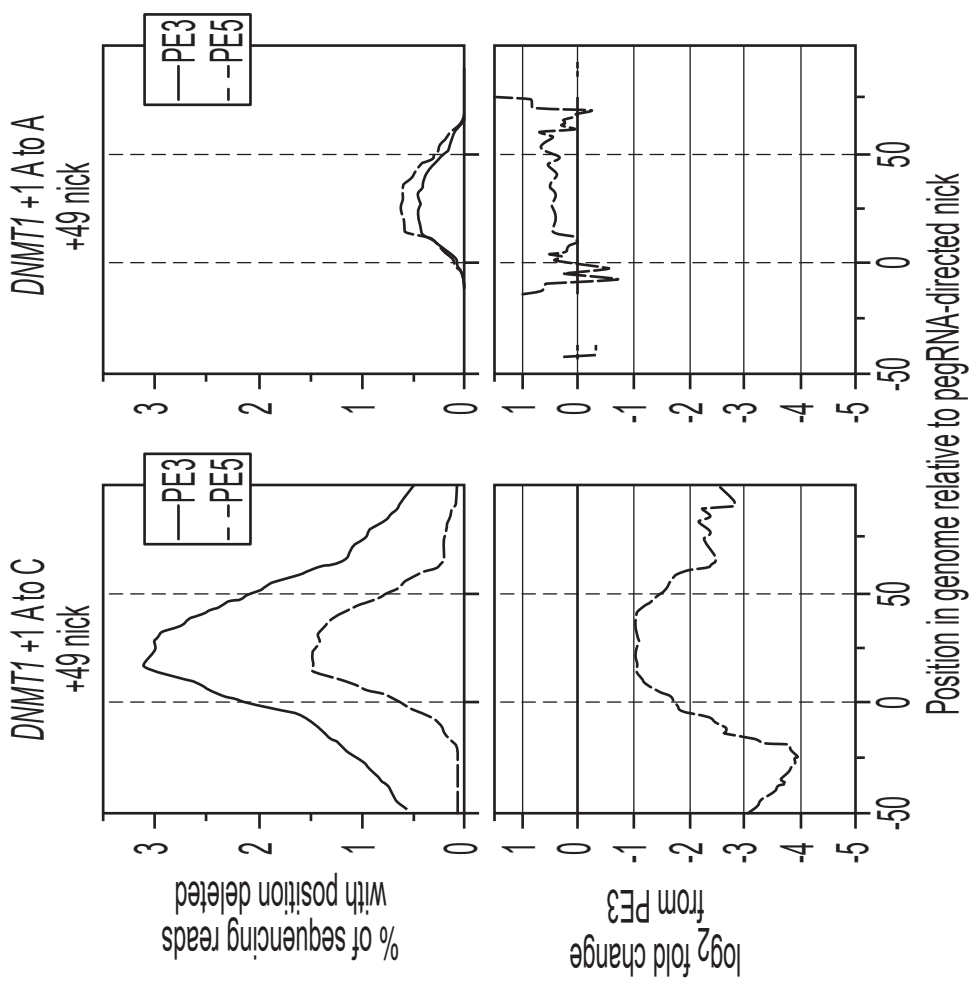


FIG. 53C

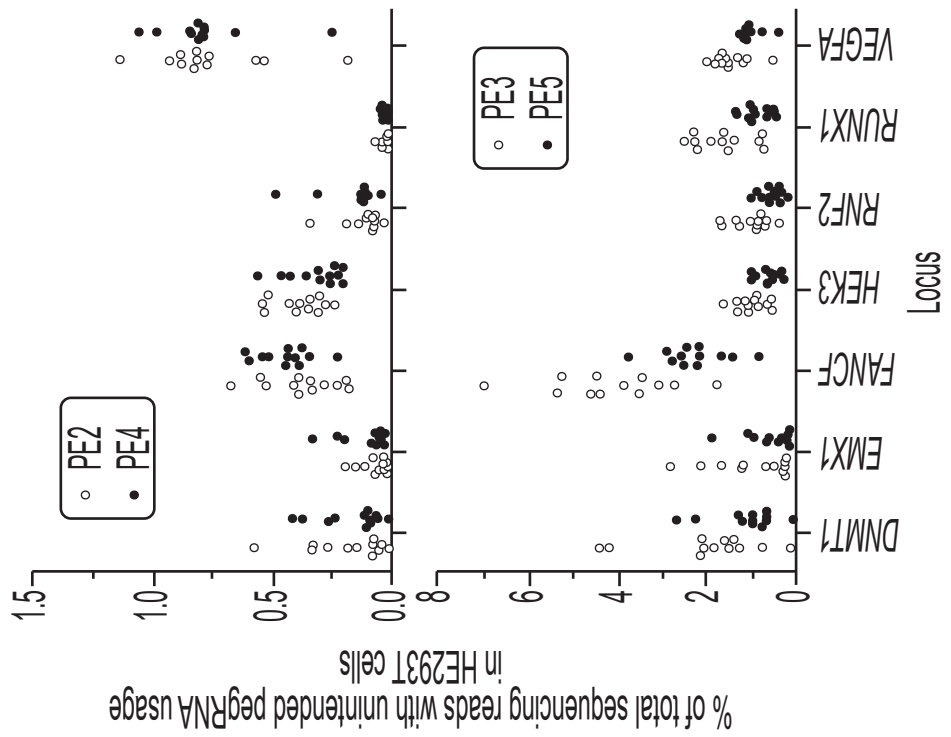


FIG. 53D

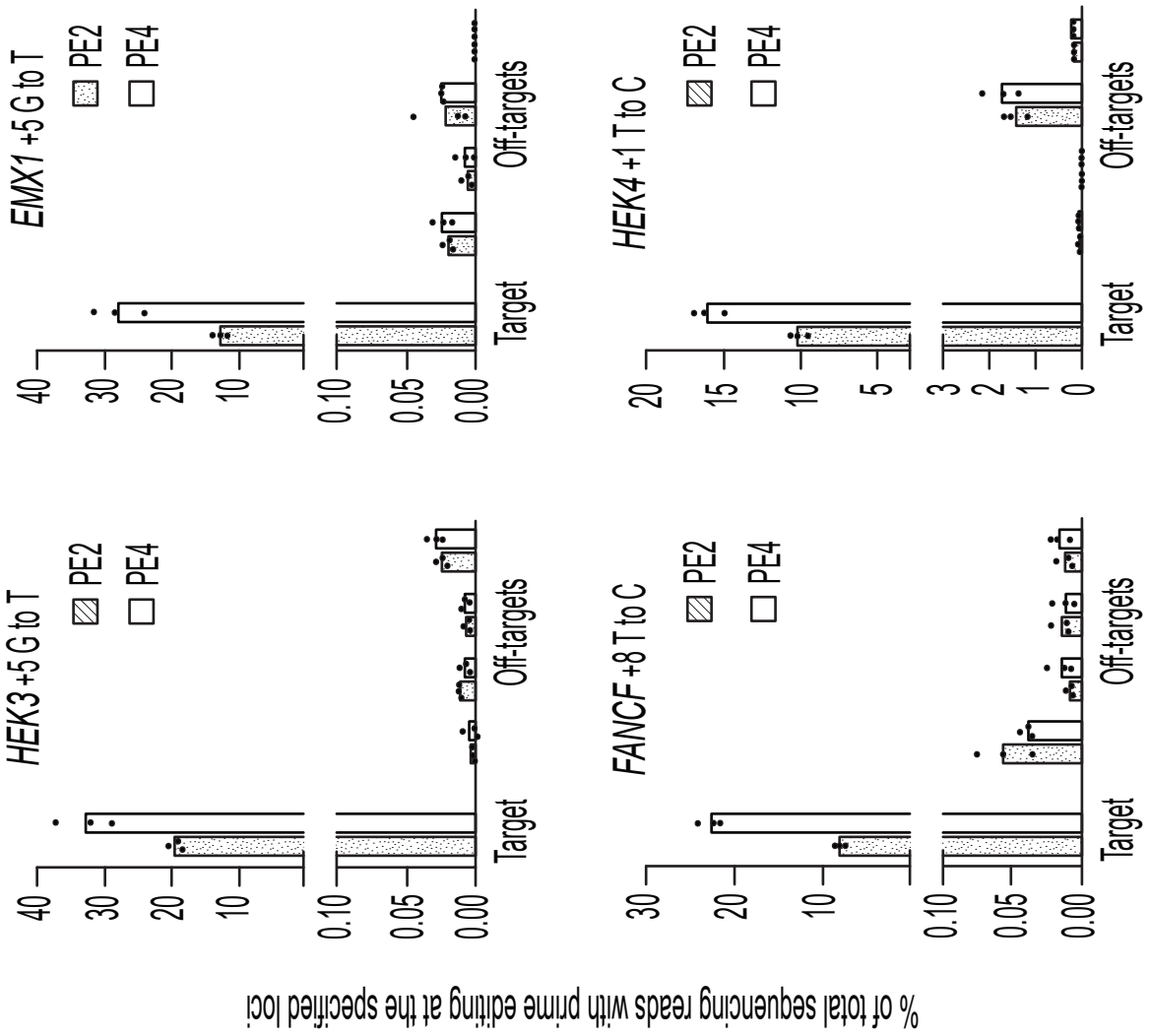

HEK3 +5 G to T
PE2
PE4
EMX1 +5 G to T
PE2
PE4
FANCF +8 T to C
PE2
PE4
HEK4 +1 T to C
PE2
PE4
Target
Off-targets
% of total sequencing reads with prime editing at the specified loci

FIG. 53E

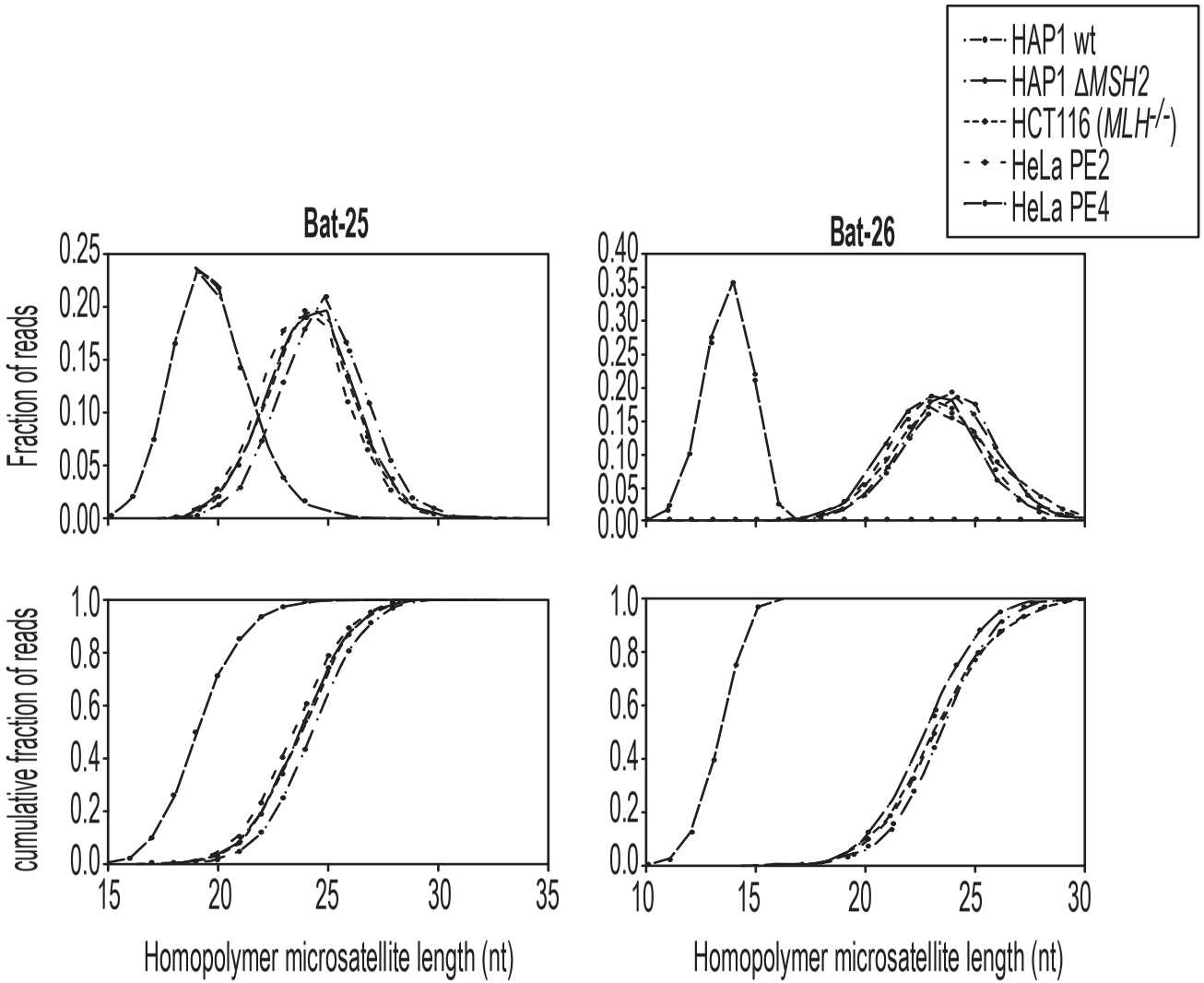


FIG. 53F

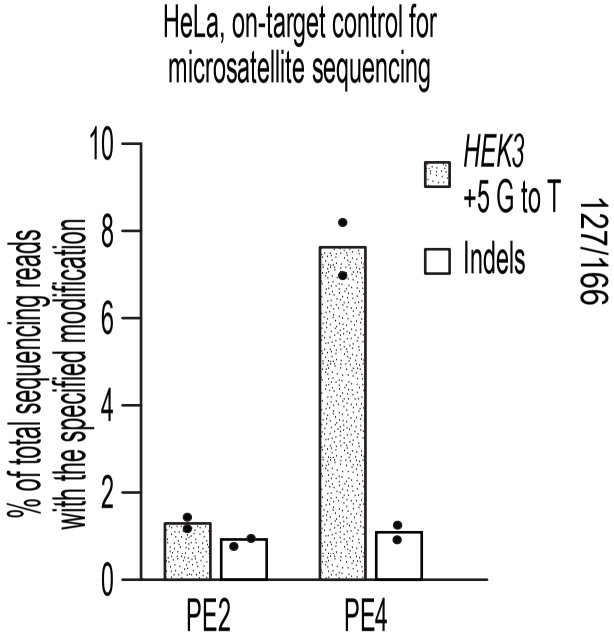


FIG. 53G

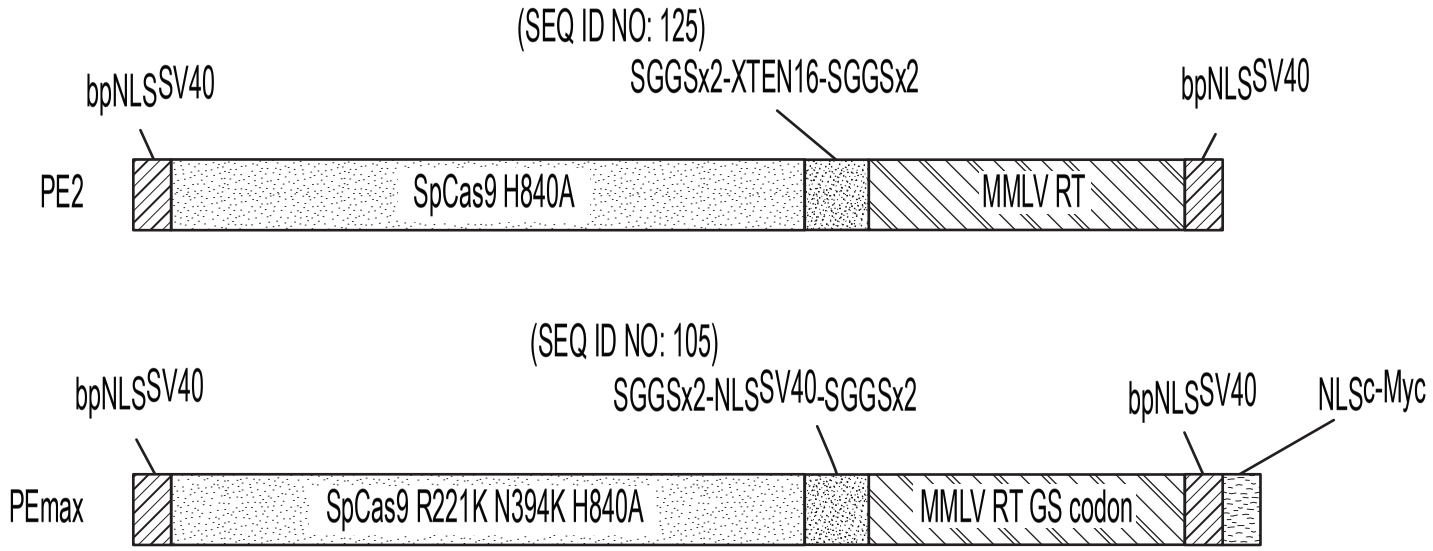


FIG. 54A

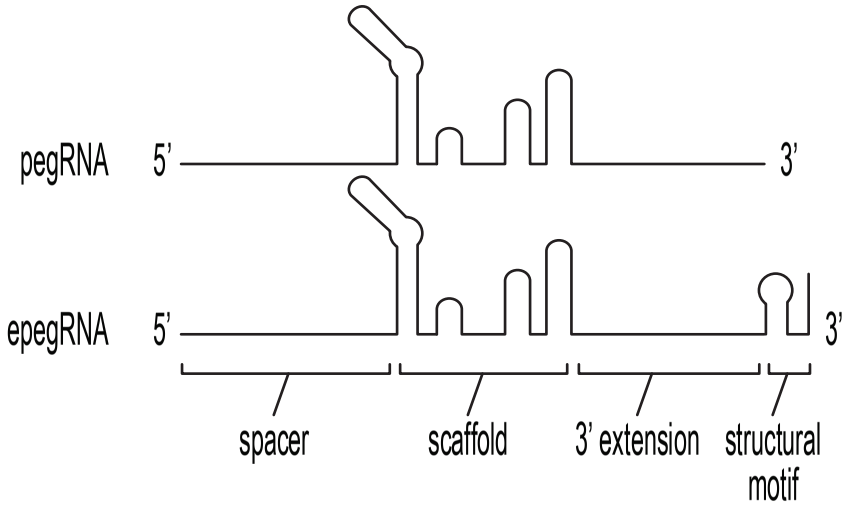


FIG. 54B

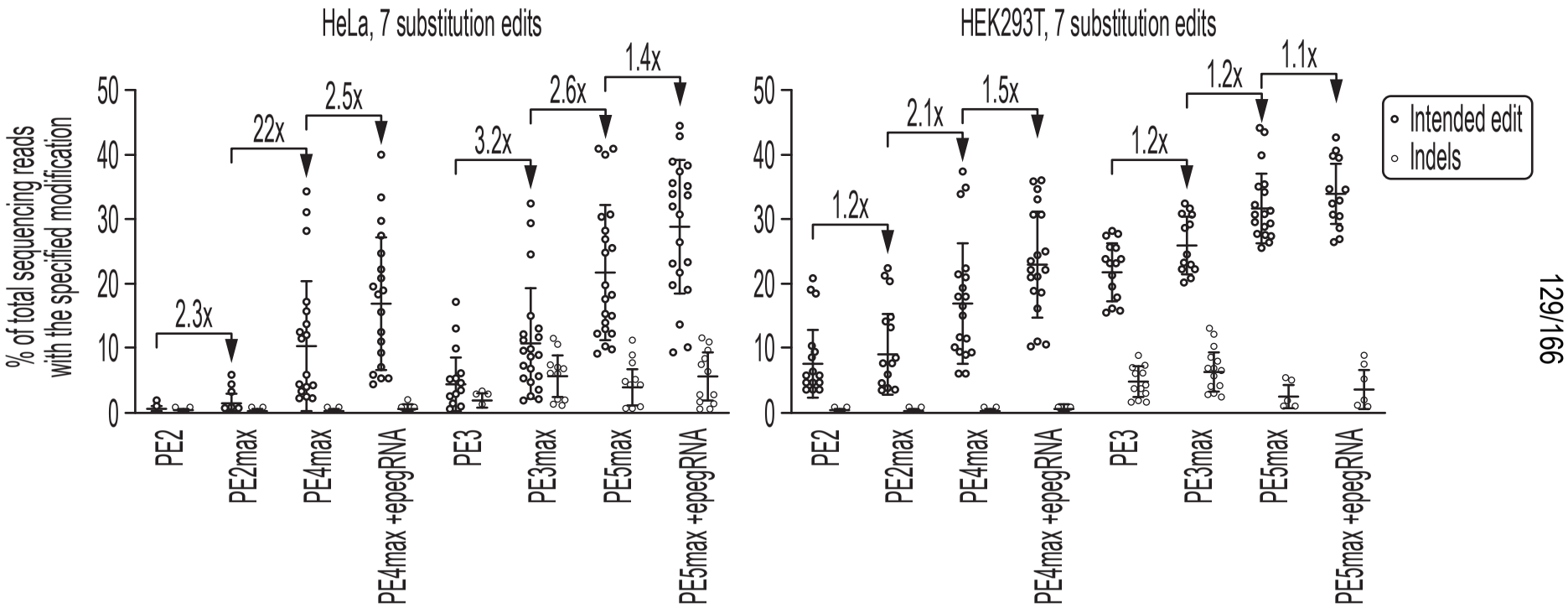

HeLa, 7 substitution edits
HEK293T, 7 substitution edits
% of total sequencing reads with the specified modification
0
10
20
30
40
50
2.3x
22x
2.5x
3.2x
2.6x
1.4x
1.2x
2.1x
1.5x
1.2x
1.2x
1.1x
Intended edit
Indels
PE2
PE2max
PE4max
PE4max +epegRNA
PE3
PE3max
PE5max
PE5max +epegRNA

FIG. 54C

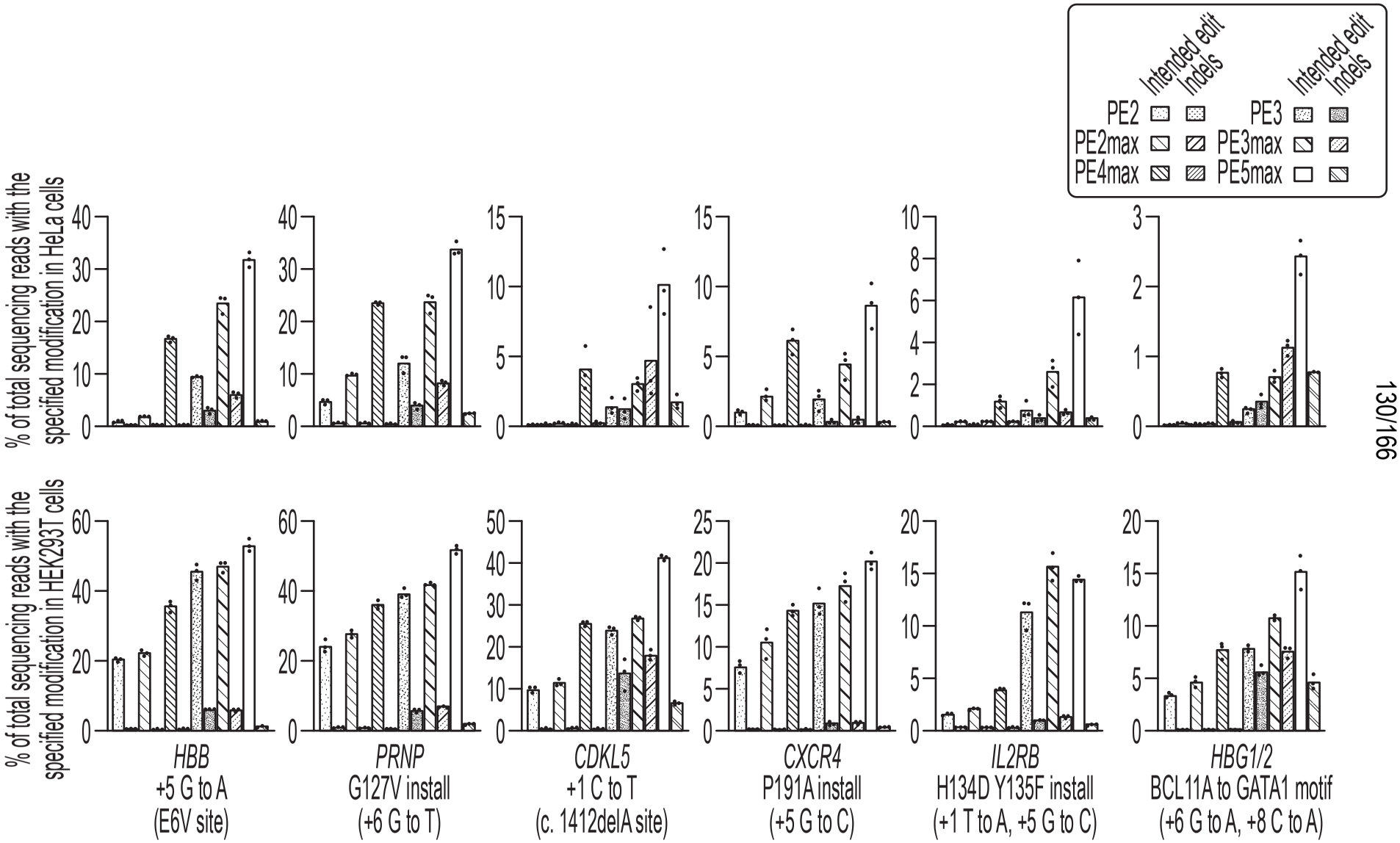

Intended edit
Indels
PE2
PE2max
PE4max
PE3
PE3max
PE5max
% of total sequencing reads with the specified modification in HeLa cells
% of total sequencing reads with the specified modification in HEK293T cells
HBB
+5 G to A
(E6V site)
PRNP
G127V install
(+6 G to T)
CDKL5
+1 C to T
(c. 1412delA site)
CXCR4
P191A install
(+5 G to C)
IL2RB
H134D Y135F install
(+1 T to A, +5 G to C)
HBG1/2
BCL11A to GATA1 motif
(+6 G to A, +8 C to A)

FIG. 54D

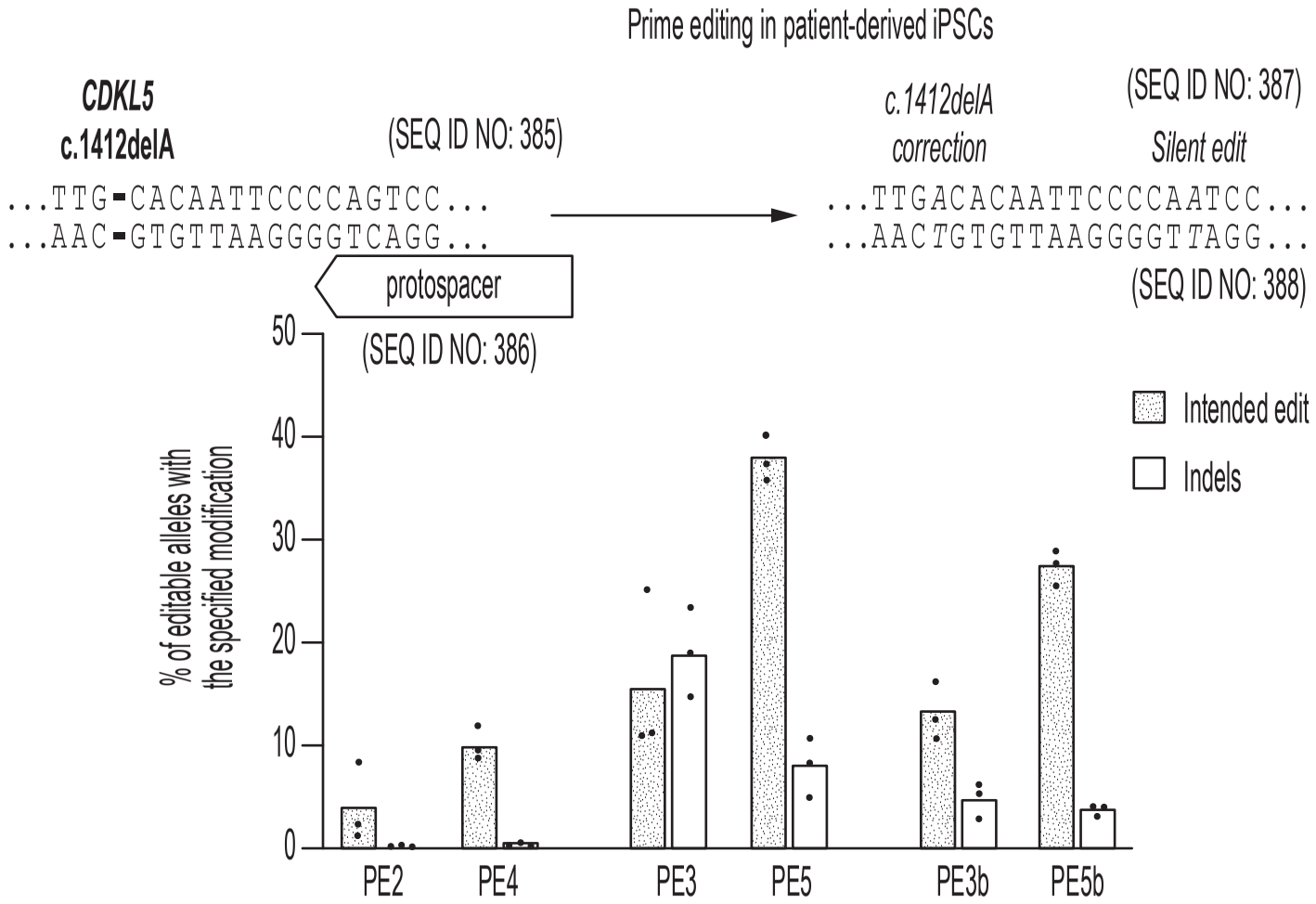

Prime editing in patient-derived iPSCs
CDKL5
c.1412delA
(SEQ ID NO: 385)
...TTG-CACAATTCCCCAGTCC...
...AAC-GTGTTAAGGGGTCAGG...
protospacer
(SEQ ID NO: 386)
c.1412delA
correction
(SEQ ID NO: 387)
Silent edit
...TTGACACAATTCCCCAATCC...
...AACTGTGTTAAGGGGTTAGG...
(SEQ ID NO: 388)
Intended edit
Indels
% of editable alleles with the specified modification
0
10
20
30
40
50
PE2
PE4
PE3
PE5
PE3b
PE5b

FIG. 54E

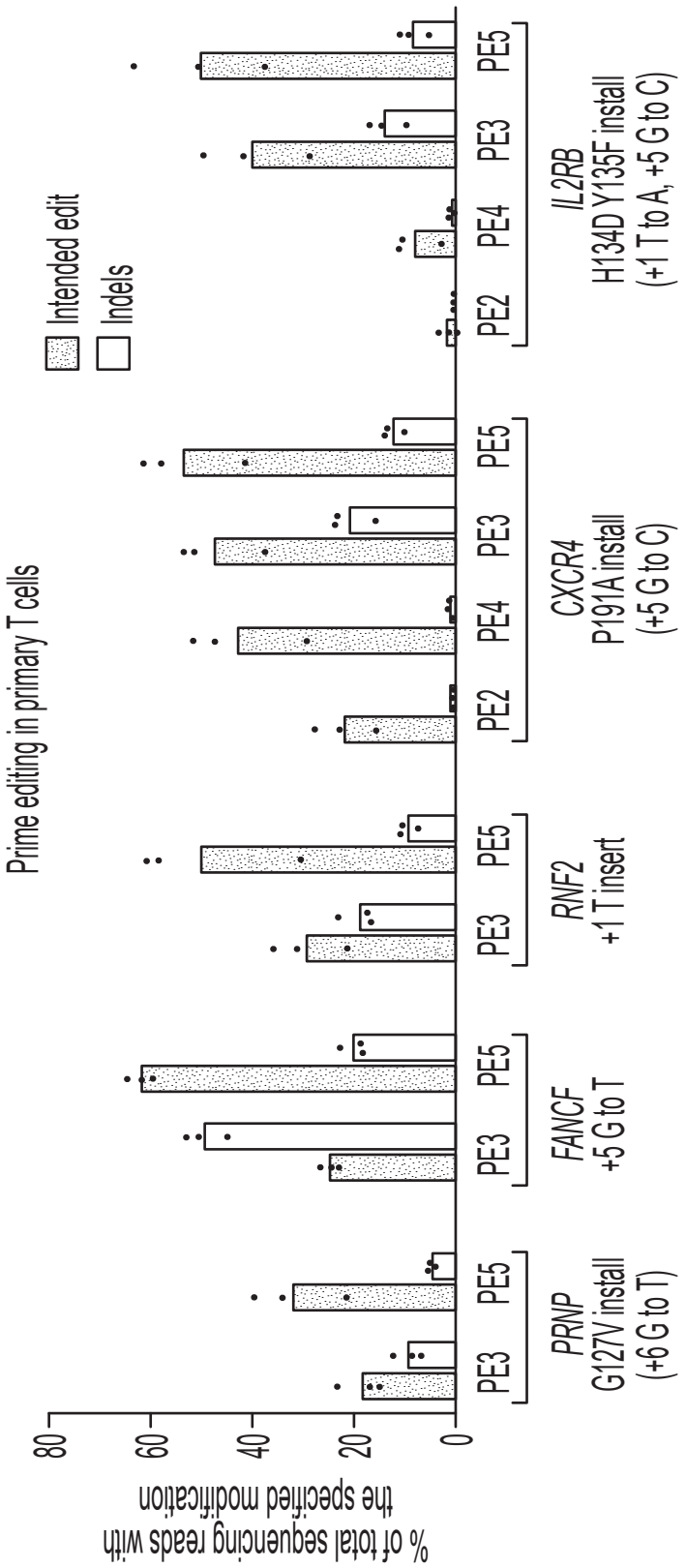

Prime editing in primary T cells
% of total sequencing reads with the specified modification
0
20
40
60
80
Intended edit
Indels
PE3
PE5
PRNP
G127V install
(+6 G to T)
PE3
PE5
FANCF
+5 G to T
PE3
PE5
RNF2
+1 T insert
PE2
PE4
PE3
PE5
CXCR4
P191A install
(+5 G to C)
PE2
PE4
PE3
PE5
IL2RB
H134D Y135F install
(+1 T to A, +5 G to C)

FIG. 54F

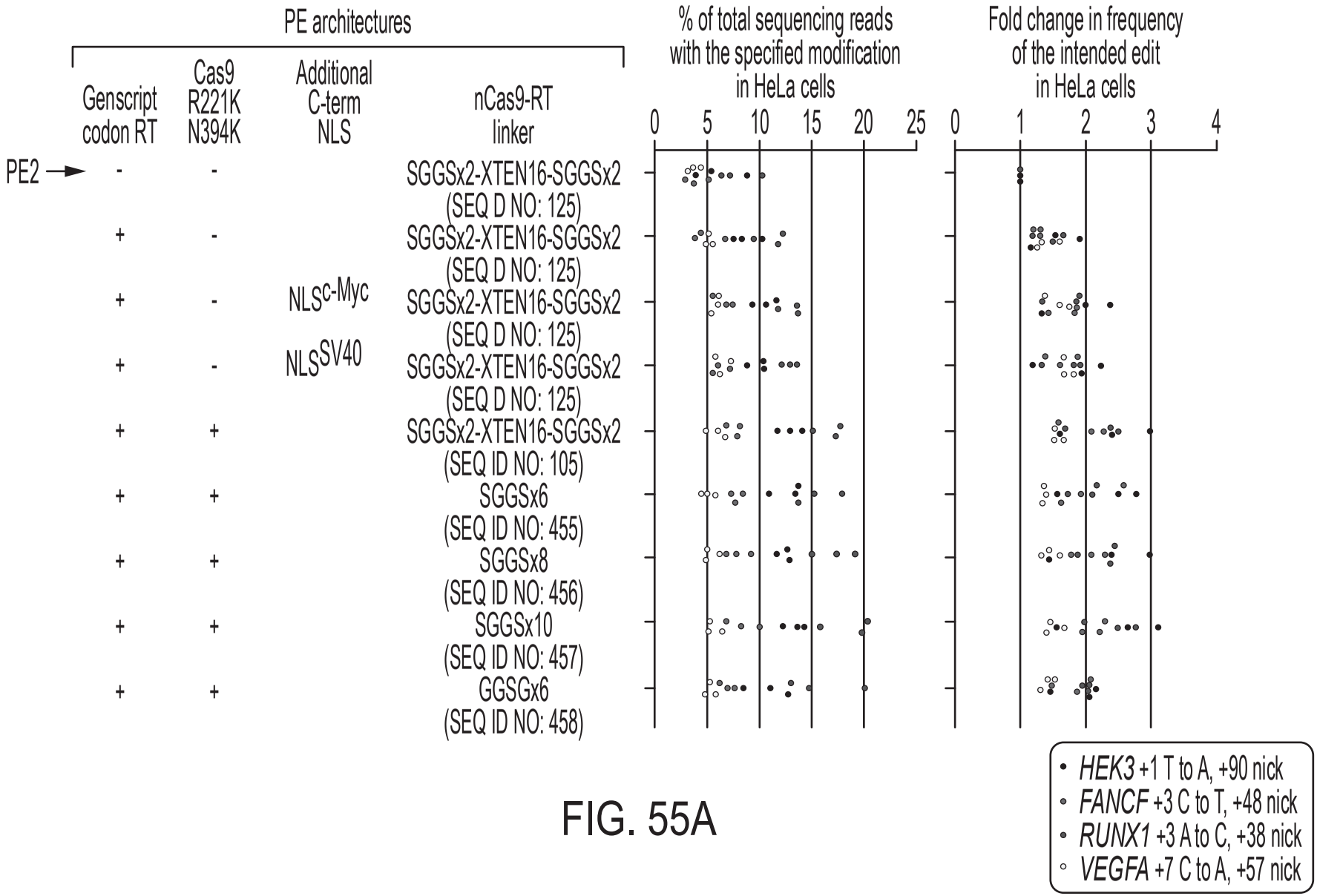

PE architectures
Genscript codon RT
Cas9 R221K N394K
Additional C-term NLS
nCas9-RT linker
PE2
-
-
SGGSx2-XTEN16-SGGSx2 (SEQ D NO: 125)
+
-
SGGSx2-XTEN16-SGGSx2 (SEQ D NO: 125)
+
-
NLSc-Myc
SGGSx2-XTEN16-SGGSx2 (SEQ D NO: 125)
+
-
NLSSV40
SGGSx2-XTEN16-SGGSx2 (SEQ D NO: 125)
+
+
SGGSx2-XTEN16-SGGSx2 (SEQ ID NO: 105)
+
+
SGGSx6 (SEQ ID NO: 455)
+
+
SGGSx8 (SEQ ID NO: 456)
+
+
SGGSx10 (SEQ ID NO: 457)
+
+
GGSGx6 (SEQ ID NO: 458)
% of total sequencing reads with the specified modification in HeLa cells
0
5
10
15
20
25
Fold change in frequency of the intended edit in HeLa cells
0
1
2
3
4
HEK3 +1 T to A, +90 nick
FANCF +3 C to T, +48 nick
RUNX1 +3 A to C, +38 nick
VEGFA +7 C to A, +57 nick

FIG. 55A

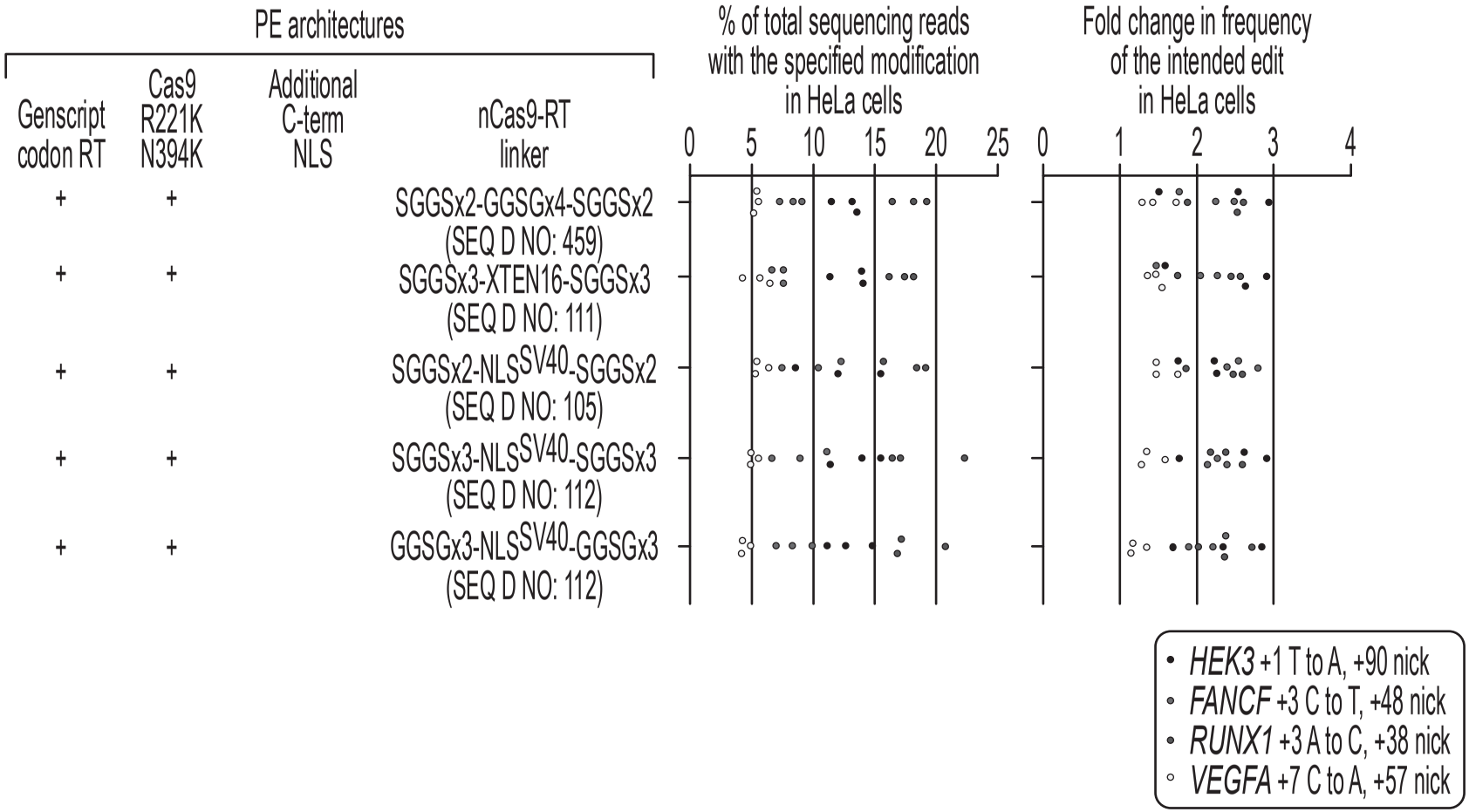


FIG. 55A CONTINUED

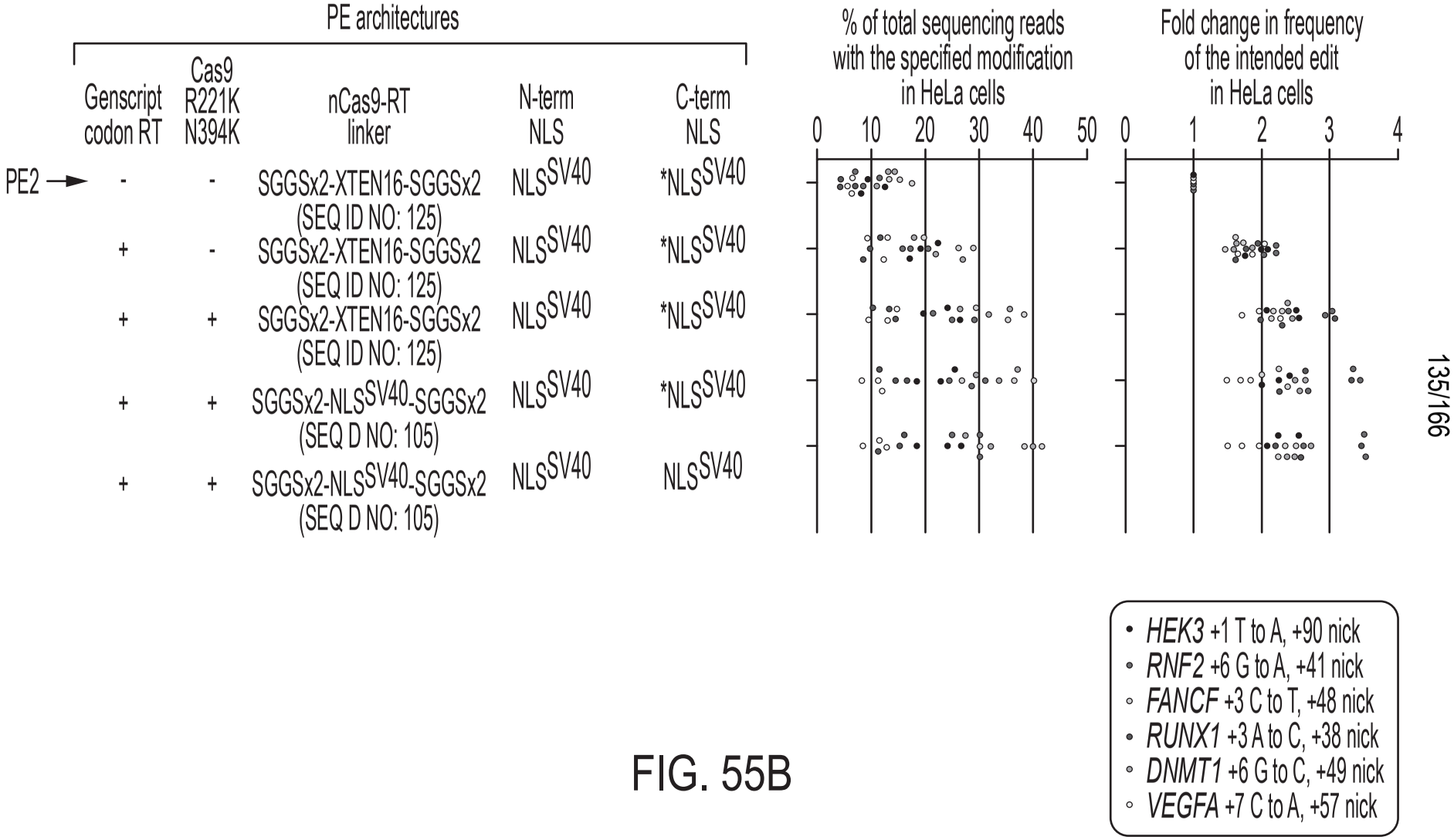

PE architectures
Genscript codon RT
Cas9 R221K N394K
nCas9-RT linker
N-term NLS
C-term NLS
PE2
-
-
SGGSx2-XTEN16-SGGSx2 (SEQ ID NO: 125)
NLSSV40
*NLSSV40
+
-
SGGSx2-XTEN16-SGGSx2 (SEQ ID NO: 125)
NLSSV40
*NLSSV40
+
+
SGGSx2-XTEN16-SGGSx2 (SEQ ID NO: 125)
NLSSV40
*NLSSV40
+
+
SGGSx2-NLSSV40-SGGSx2 (SEQ D NO: 105)
NLSSV40
*NLSSV40
+
+
SGGSx2-NLSSV40-SGGSx2 (SEQ D NO: 105)
NLSSV40
NLSSV40
% of total sequencing reads with the specified modification in HeLa cells
0
10
20
30
40
50
Fold change in frequency of the intended edit in HeLa cells
0
1
2
3
4
HEK3 +1 T to A, +90 nick
RNF2 +6 G to A, +41 nick
FANCF +3 C to T, +48 nick
RUNX1 +3 A to C, +38 nick
DNMT1 +6 G to C, +49 nick
VEGFA +7 C to A, +57 nick

FIG. 55B

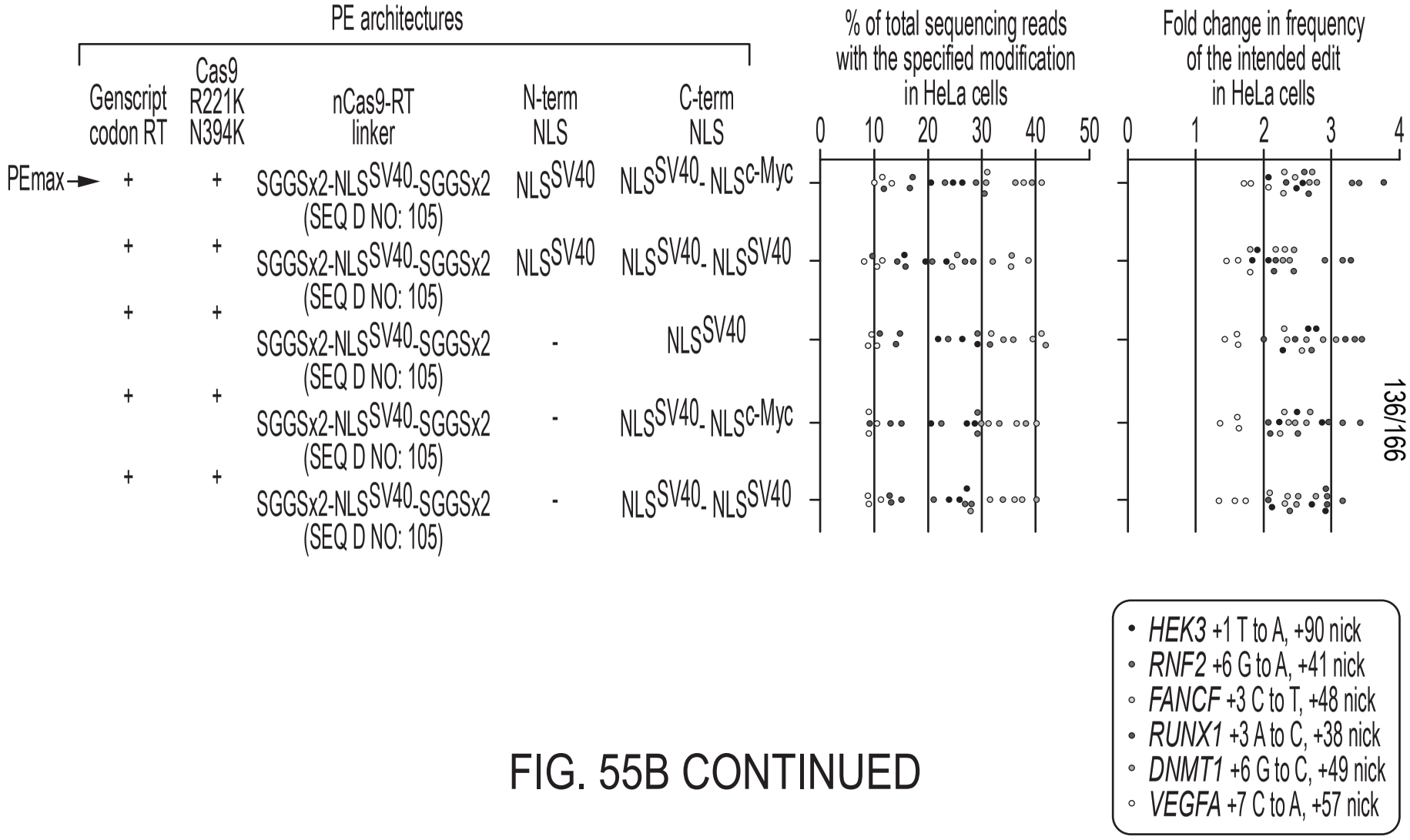


FIG. 55B CONTINUED

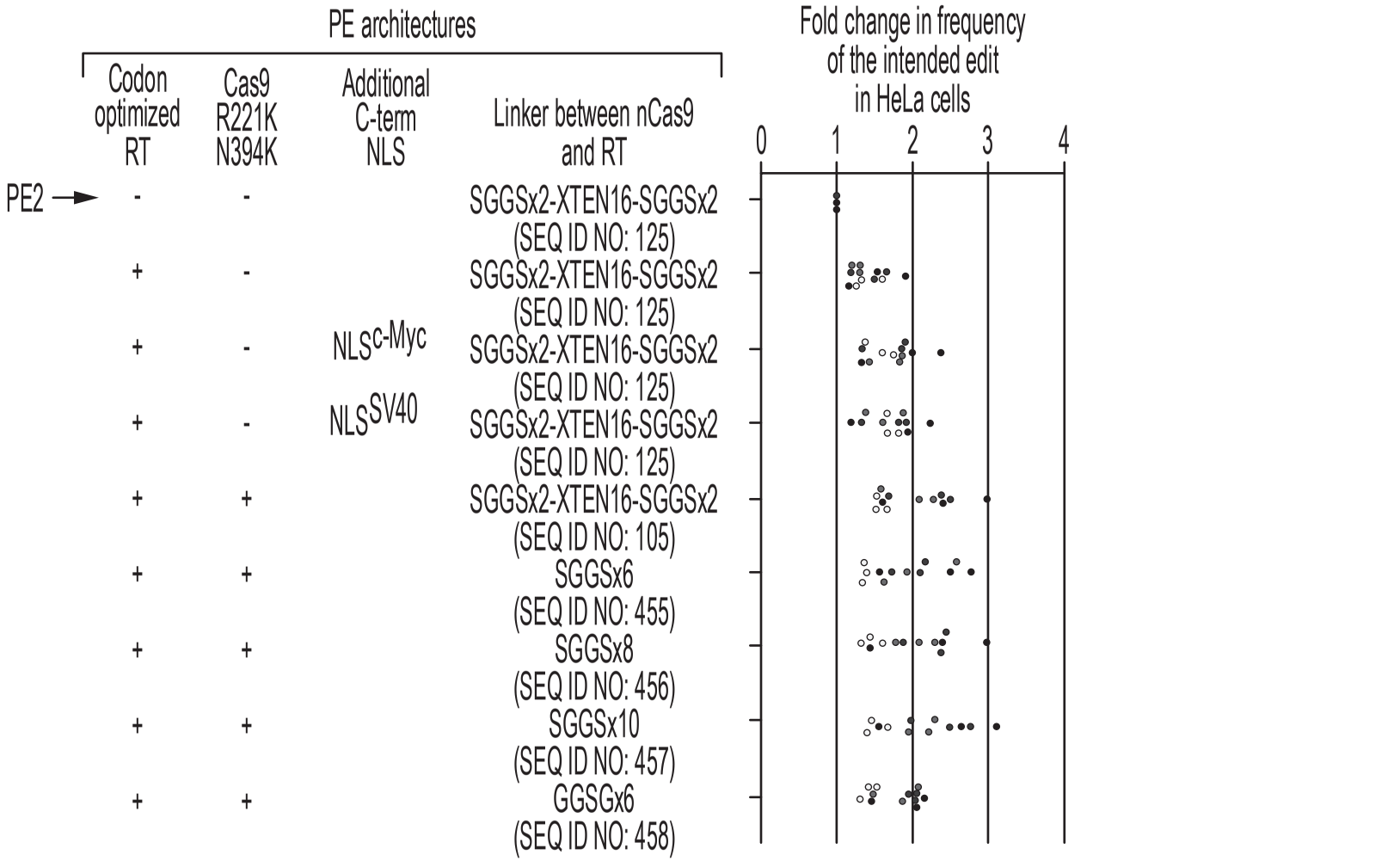

PE architectures
Codon optimized RT
Cas9 R221K N394K
Additional C-term NLS
Linker between nCas9 and RT
Fold change in frequency of the intended edit in HeLa cells
0
1
2
3
4
PE2
-
-
SGGSx2-XTEN16-SGGSx2 (SEQ ID NO: 125)
+
-
SGGSx2-XTEN16-SGGSx2 (SEQ ID NO: 125)
+
-
NLSc-Myc
SGGSx2-XTEN16-SGGSx2 (SEQ ID NO: 125)
+
-
NLSSV40
SGGSx2-XTEN16-SGGSx2 (SEQ ID NO: 125)
+
+
SGGSx2-XTEN16-SGGSx2 (SEQ ID NO: 105)
+
+
SGGSx6 (SEQ ID NO: 455)
+
+
SGGSx8 (SEQ ID NO: 456)
+
+
SGGSx10 (SEQ ID NO: 457)
+
+
GGSGx6 (SEQ ID NO: 458)
• HEK3 +1 T to A, +90 nick
• RNF2 +6 G to A, +41 nick
• FANCF +3 C to T, +48 nick
• RUNX1 +3 A to C, +38 nick
• DNMT1 +6 G to C, +49 nick
• VEGFA +7 C to A, +57 nick

FIG. 56A

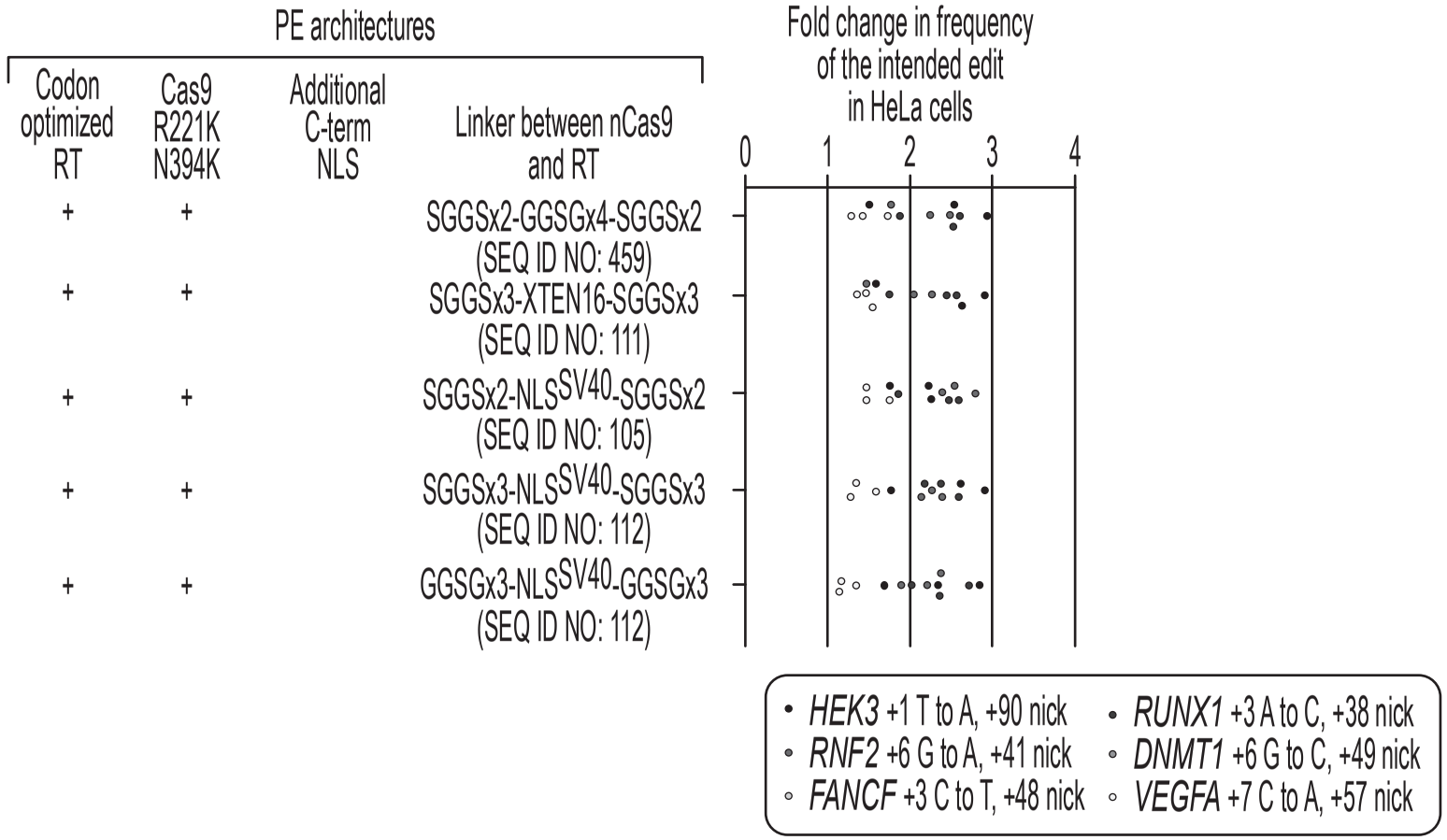


FIG. 56A CONTINUED

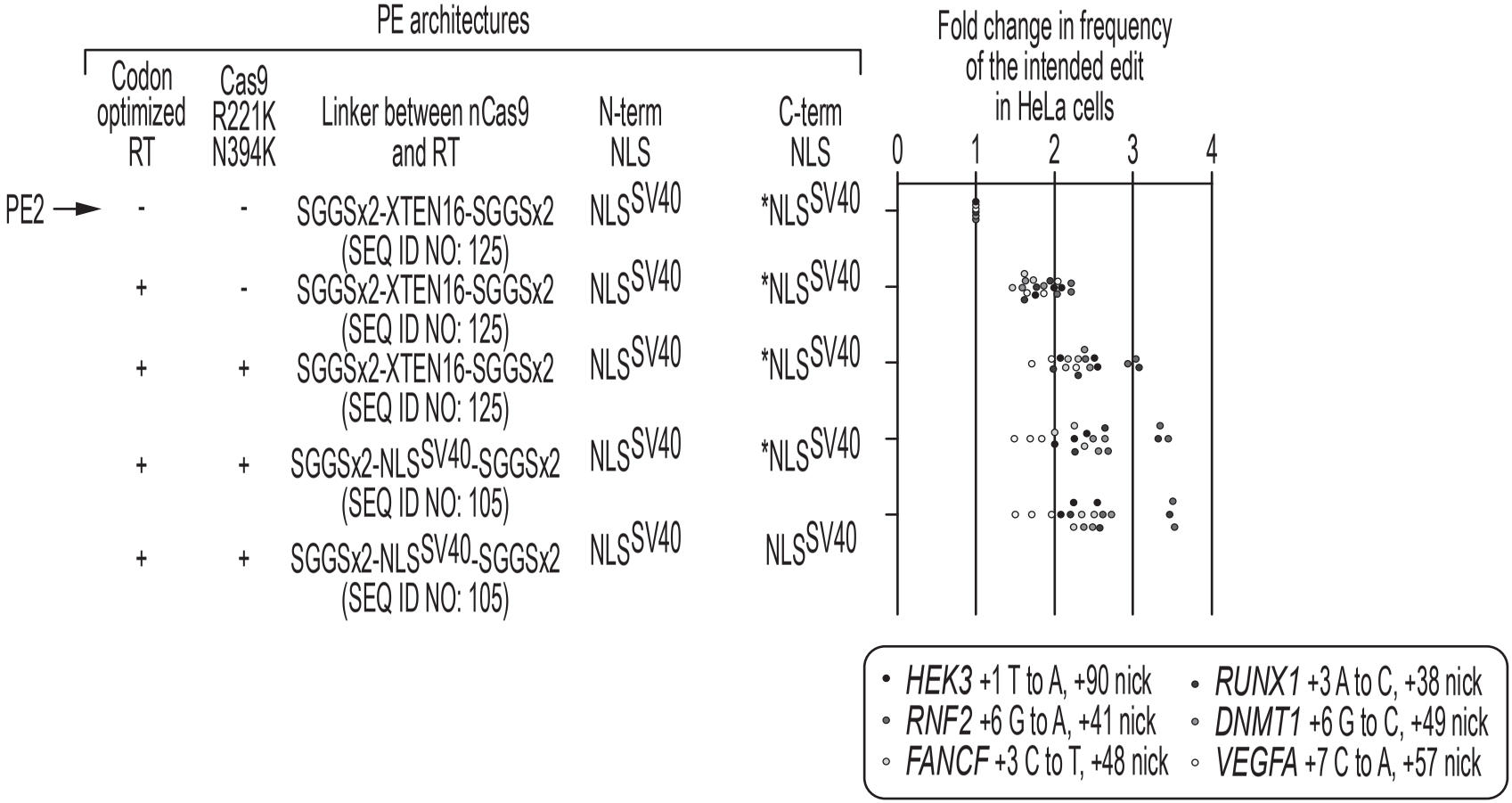


FIG. 56A CONTINUED

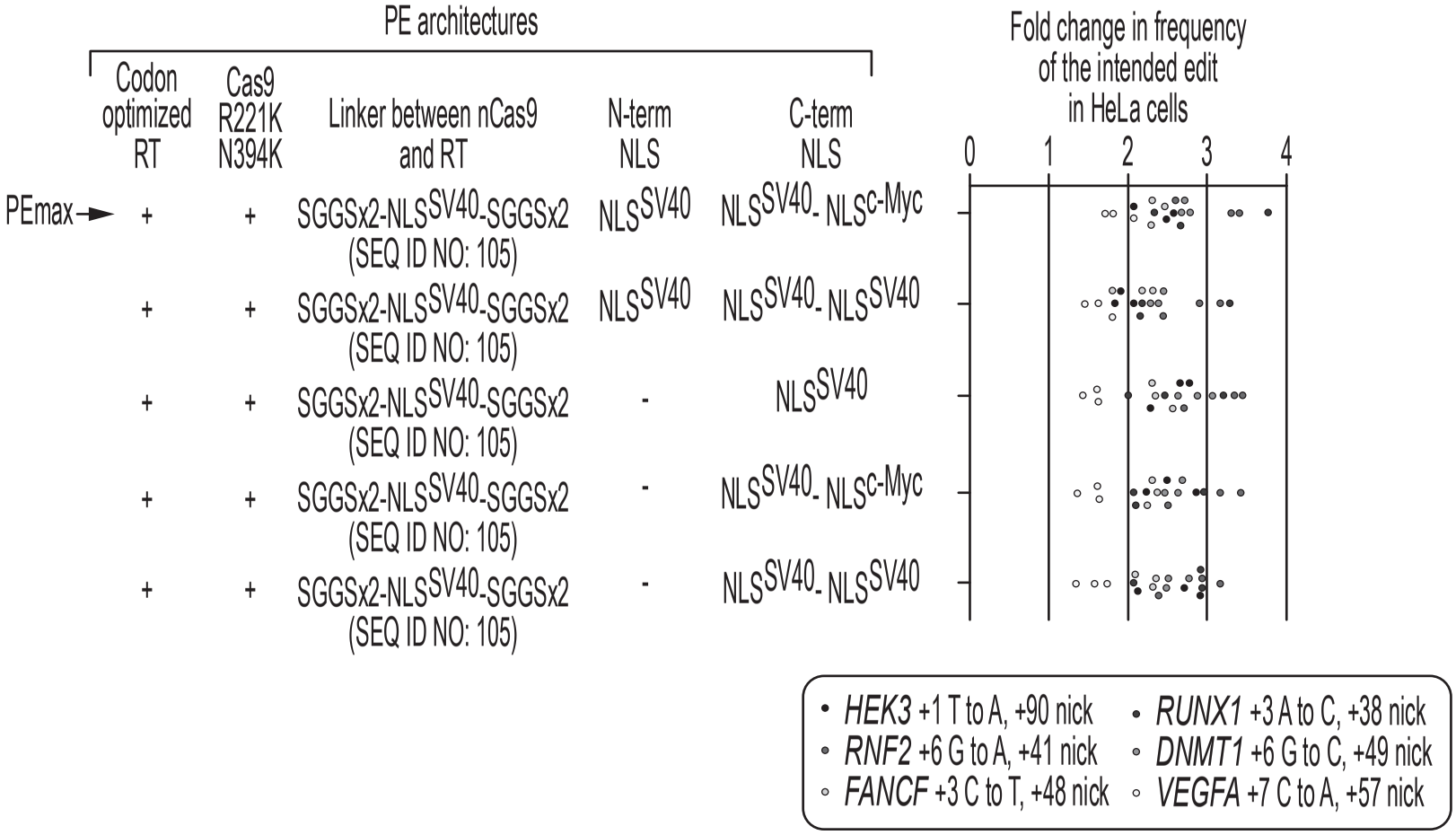


FIG. 56A CONTINUED

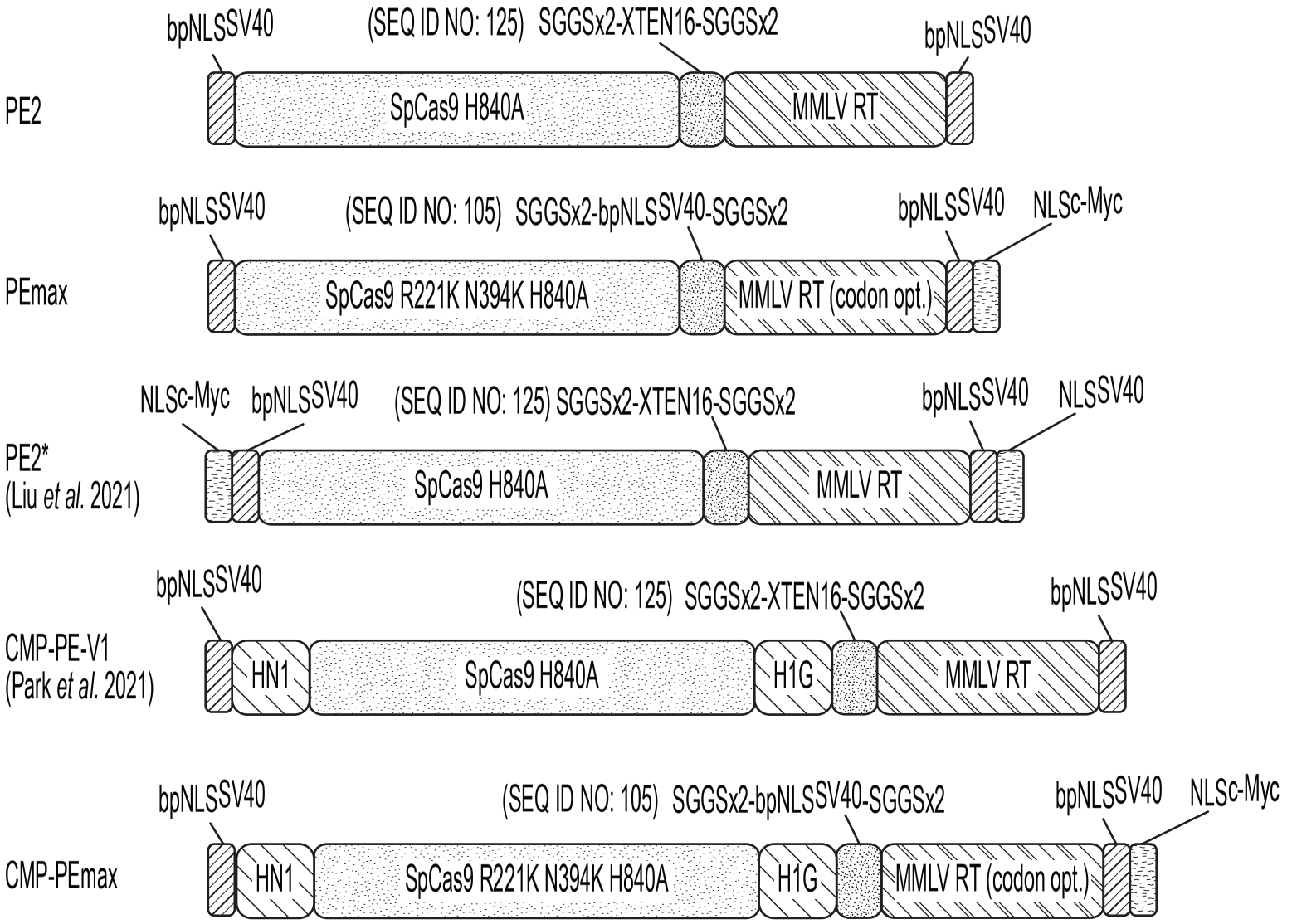

PE2
bpNLSSV40
(SEQ ID NO: 125) SGGSx2-XTEN16-SGGSx2
bpNLSSV40
SpCas9 H840A
MMLV RT
PEmax
bpNLSSV40
(SEQ ID NO: 105) SGGSx2-bpNLSSV40-SGGSx2
bpNLSSV40
NLSc-Myc
SpCas9 R221K N394K H840A
MMLV RT (codon opt.)
PE2*
(Liu et al. 2021)
NLSc-Myc
bpNLSSV40
(SEQ ID NO: 125) SGGSx2-XTEN16-SGGSx2
bpNLSSV40
NLSSV40
SpCas9 H840A
MMLV RT
CMP-PE-V1
(Park et al. 2021)
bpNLSSV40
(SEQ ID NO: 125) SGGSx2-XTEN16-SGGSx2
bpNLSSV40
HN1
SpCas9 H840A
H1G
MMLV RT
CMP-PEmax
bpNLSSV40
(SEQ ID NO: 105) SGGSx2-bpNLSSV40-SGGSx2
bpNLSSV40
NLSc-Myc
HN1
SpCas9 R221K N394K H840A
H1G
MMLV RT (codon opt.)

FIG. 56B

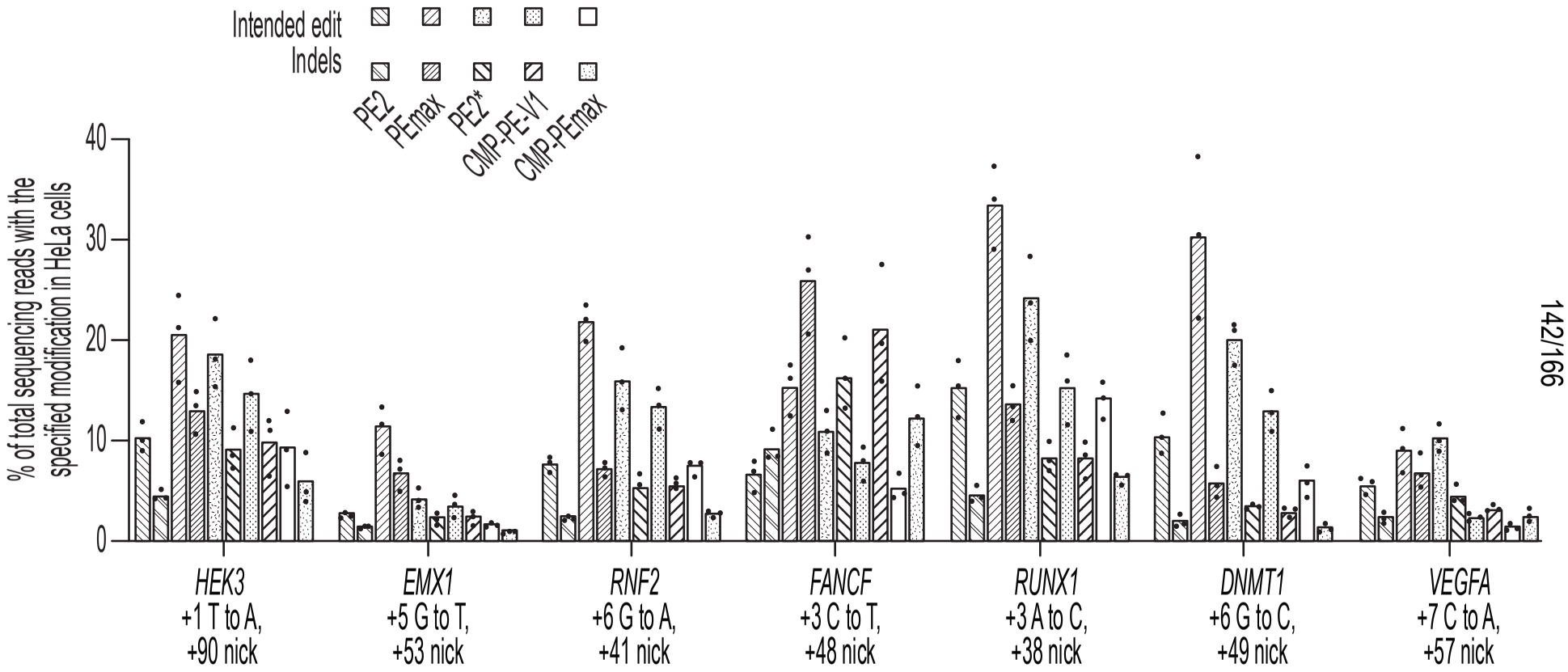

Intended edit
Indels
PE2
PEmax
PE2*
CMP-PE-V1
CMP-PEmax
% of total sequencing reads with the specified modification in HeLa cells
40
30
20
10
0
HEK3
+1 T to A,
+90 nick
EMX1
+5 G to T,
+53 nick
RNF2
+6 G to A,
+41 nick
FANCF
+3 C to T,
+48 nick
RUNX1
+3 A to C,
+38 nick
DNMT1
+6 G to C,
+49 nick
VEGFA
+7 C to A,
+57 nick

FIG. 56C

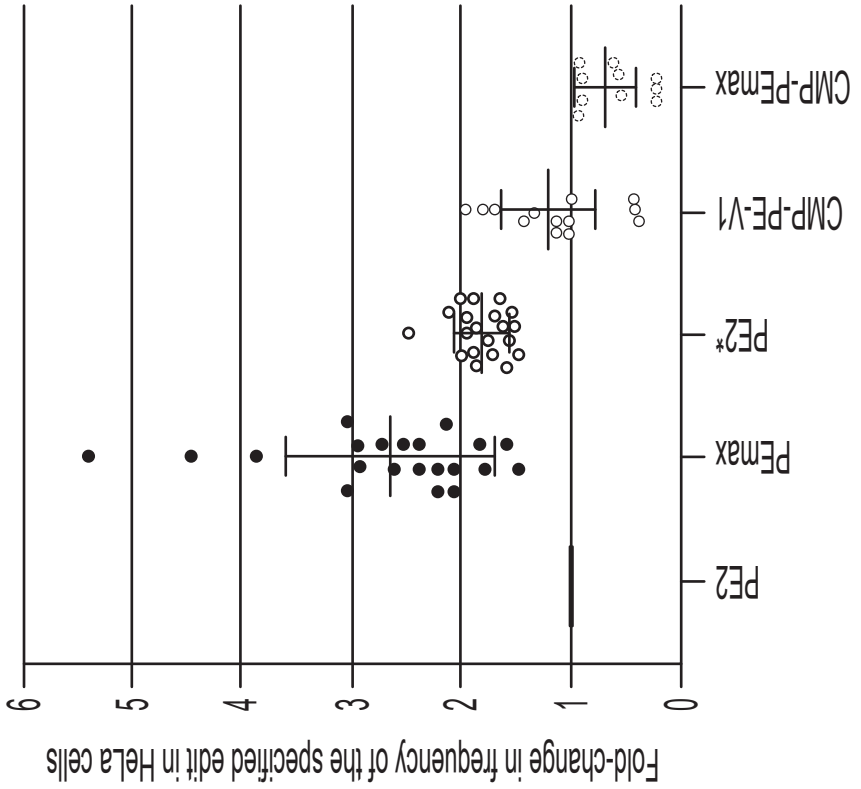

Fold-change in frequency of the specified edit in HeLa cells
0
1
2
3
4
5
6
PE2
PEmax
PE2*
CMP-PE-V1
CMP-PEmax

FIG. 56D

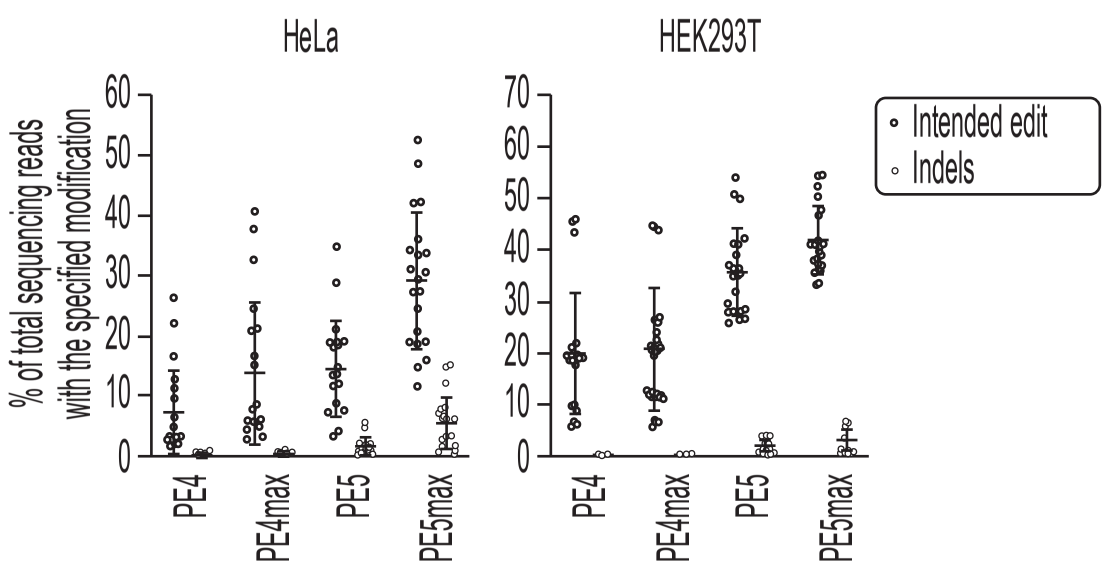


FIG. 56E

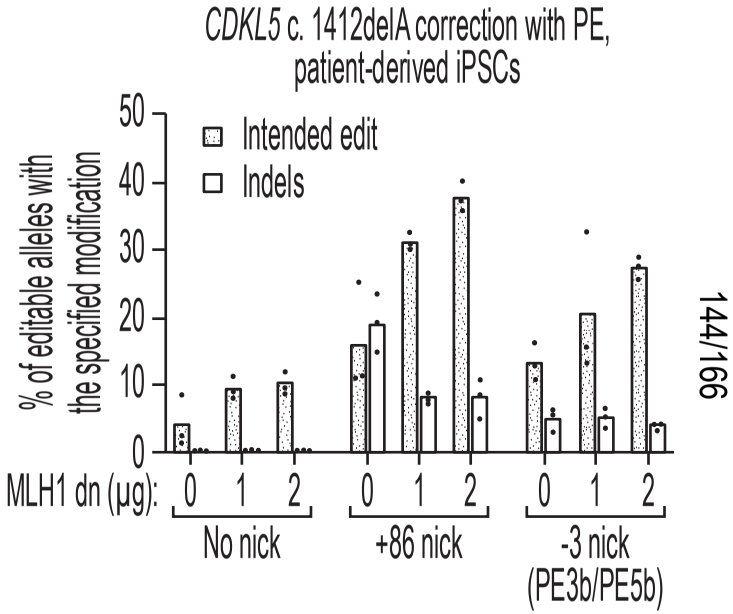


FIG. 56F

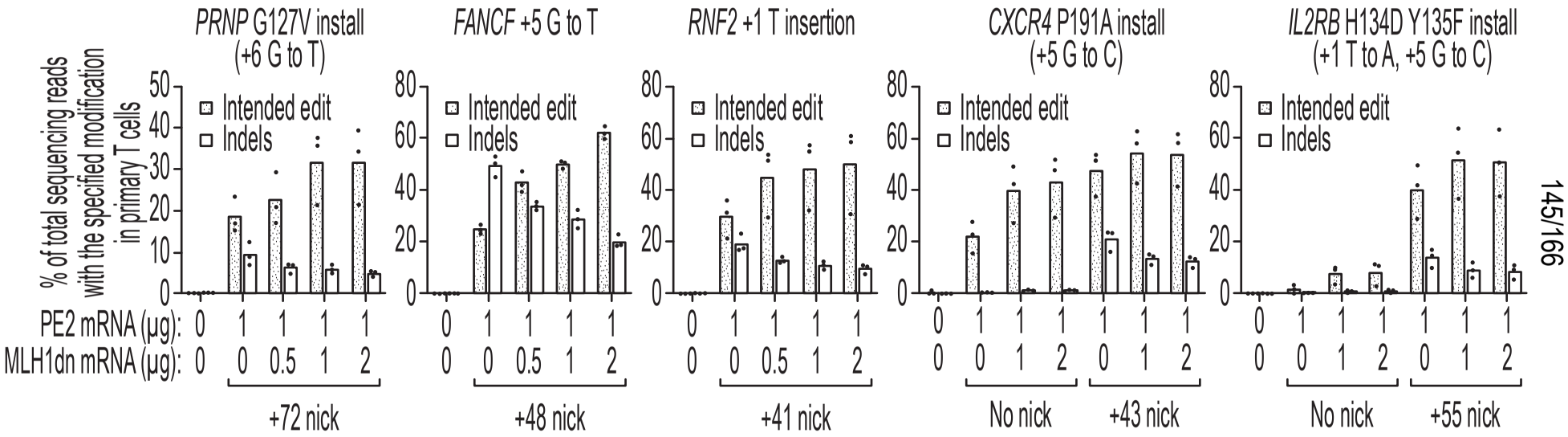

PRNP G127V install
(+6 G to T)
FANCF +5 G to T
RNF2 +1 T insertion
CXCR4 P191A install
(+5 G to C)
IL2RB H134D Y135F install
(+1 T to A, +5 G to C)
% of total sequencing reads
with the specified modification
in primary T cells
Intended edit
Indels
PE2 mRNA (µg):
MLH1dn mRNA (µg):
+72 nick
+48 nick
+41 nick
No nick
+43 nick
No nick
+55 nick

FIG. 56G

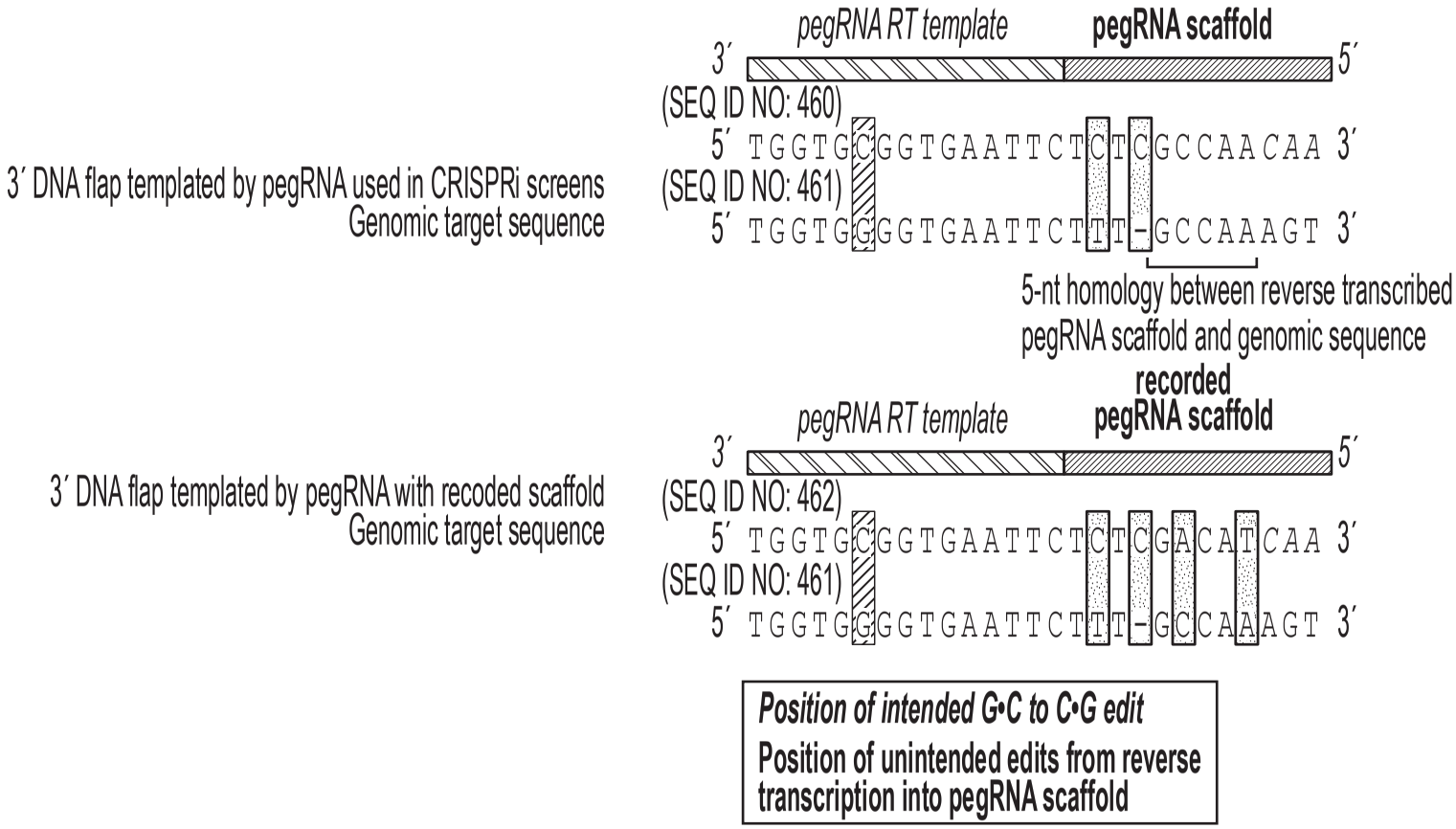

pegRNA RT template
pegRNA scaffold
3′
5′
(SEQ ID NO: 460)
5′ TGGTGCGGTGAATTCTCTCGCCAACAA 3′
3′ DNA flap templated by pegRNA used in CRISPRi screens
Genomic target sequence
(SEQ ID NO: 461)
5′ TGGTGGGGTGAATTCTTT-GCCAAAGT 3′
5-nt homology between reverse transcribed pegRNA scaffold and genomic sequence
recorded
pegRNA RT template
pegRNA scaffold
3′
5′
(SEQ ID NO: 462)
5′ TGGTGCGGTGAATTCTCTCGACATCAA 3′
3′ DNA flap templated by pegRNA with recoded scaffold
Genomic target sequence
(SEQ ID NO: 461)
5′ TGGTGGGGTGAATTCTTT-GCCAAAGT 3′
Position of intended G•C to C•G edit
Position of unintended edits from reverse transcription into pegRNA scaffold

FIG. 57A

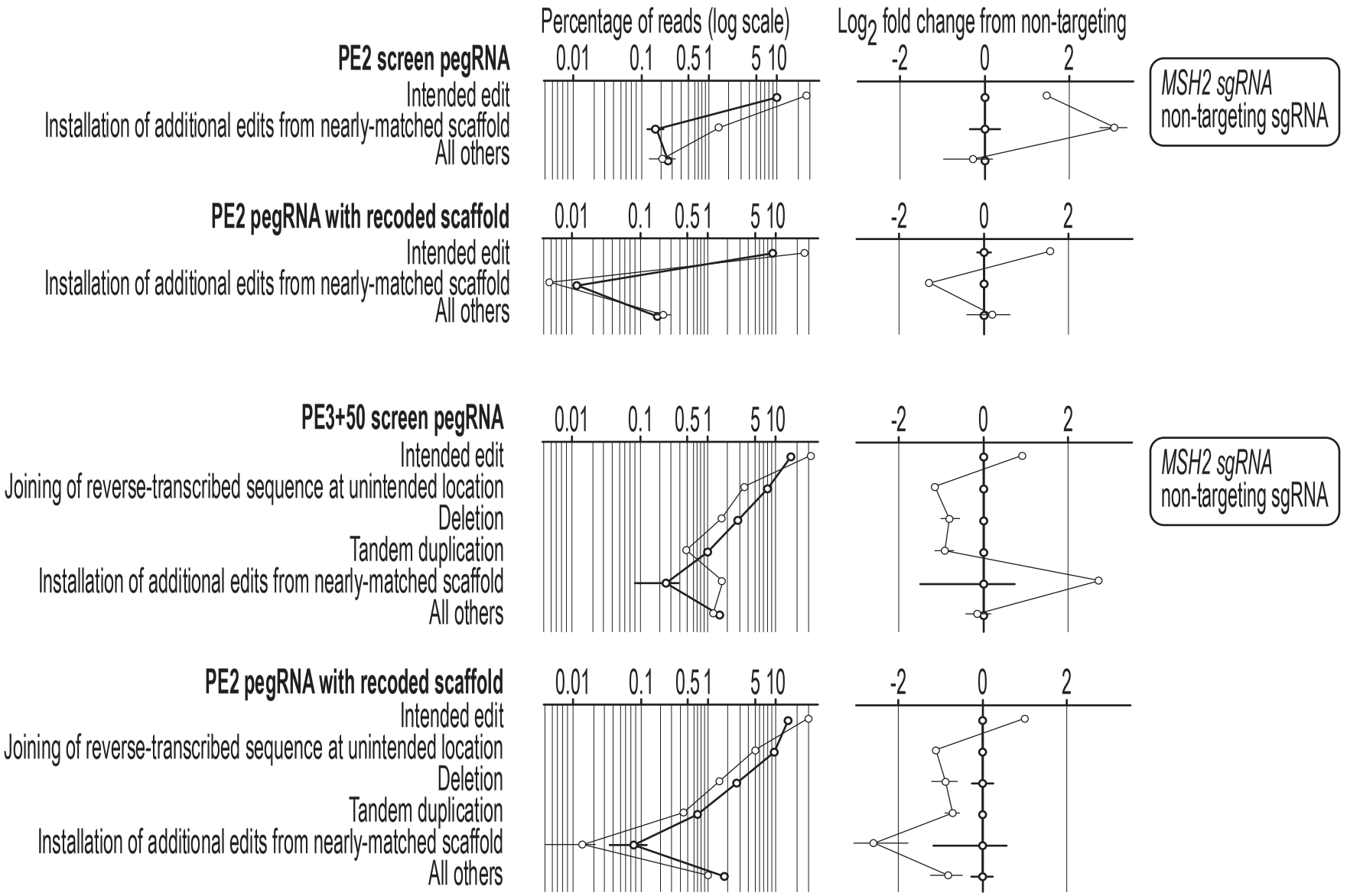

Percentage of reads (log scale)
Log2 fold change from non-targeting
PE2 screen pegRNA
Intended edit
Installation of additional edits from nearly-matched scaffold
All others
PE2 pegRNA with recoded scaffold
Intended edit
Installation of additional edits from nearly-matched scaffold
All others
PE3+50 screen pegRNA
Intended edit
Joining of reverse-transcribed sequence at unintended location
Deletion
Tandem duplication
Installation of additional edits from nearly-matched scaffold
All others
PE2 pegRNA with recoded scaffold
Intended edit
Joining of reverse-transcribed sequence at unintended location
Deletion
Tandem duplication
Installation of additional edits from nearly-matched scaffold
All others
0.01
0.1
0.5
1
5
10
-2
0
2
MSH2 sgRNA
non-targeting sgRNA

FIG. 57B

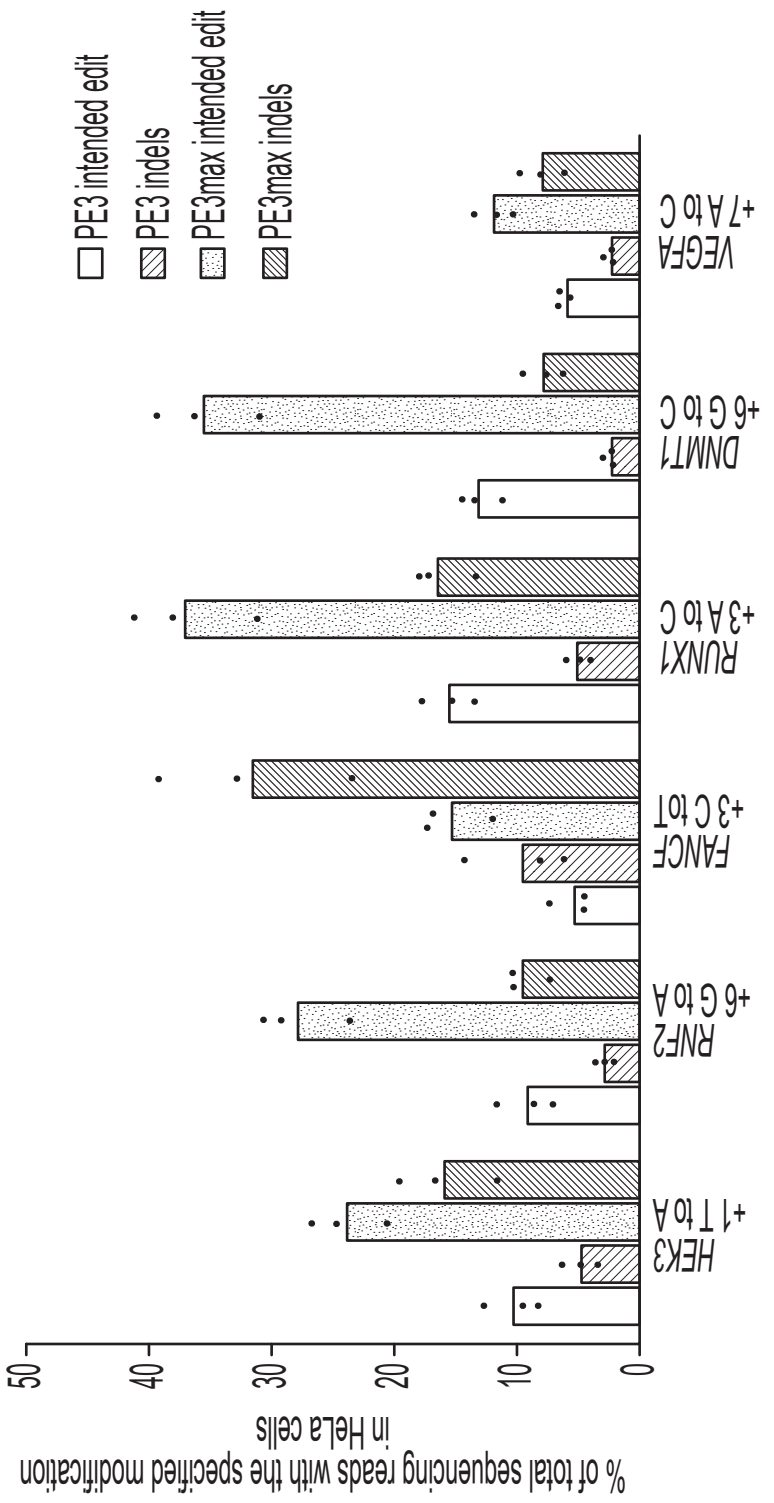

PE3 intended edit
PE3 indels
PE3max intended edit
PE3max indels
% of total sequencing reads with the specified modification in HeLa cells
0
10
20
30
40
50
HEK3 +1 T to A
RNF2 +6 G to A
FANCF +3 C to T
RUNX1 +3 A to C
DNMT1 +6 G to C
VEGFA +7 A to C

FIG. 58

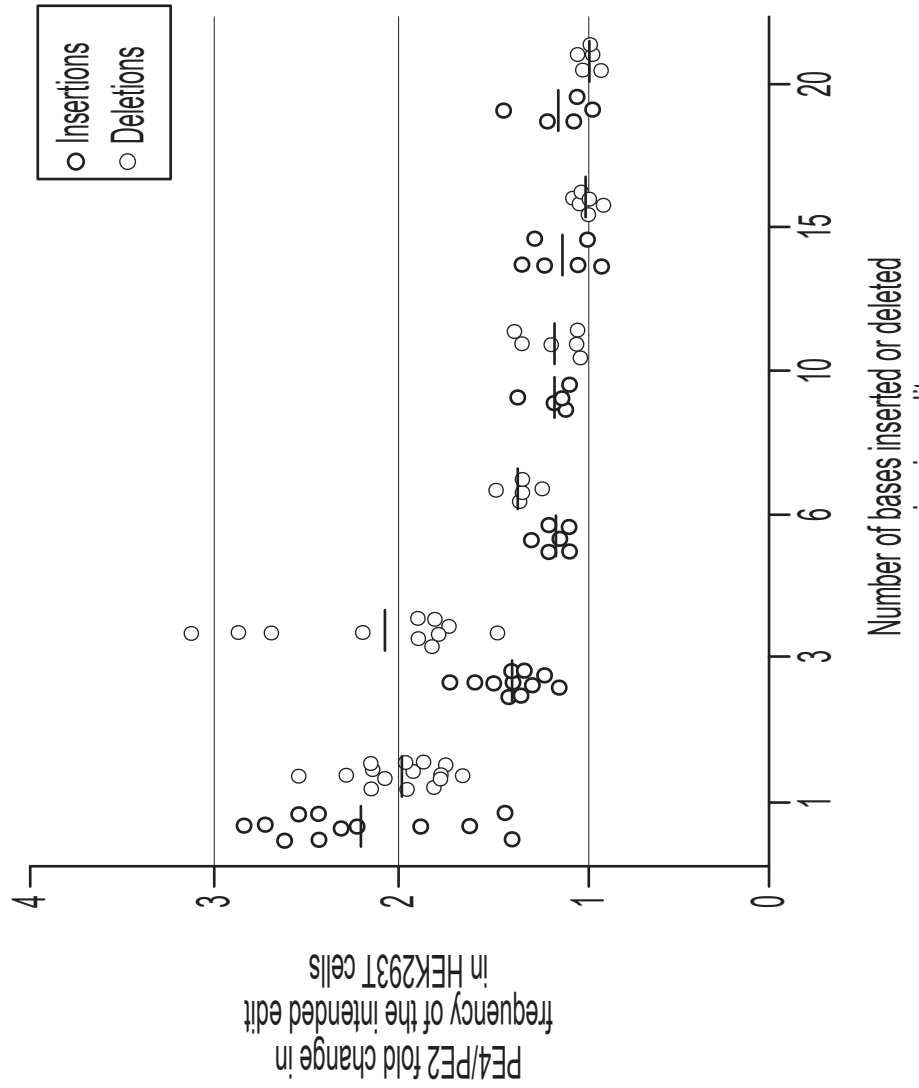

Insertions
Deletions
PE4/PE2 fold change in
frequency of the intended edit
in HEK293T cells
0
1
2
3
4
1
3
6
10
15
20
Number of bases inserted or deleted
in prime edit

FIG. 59

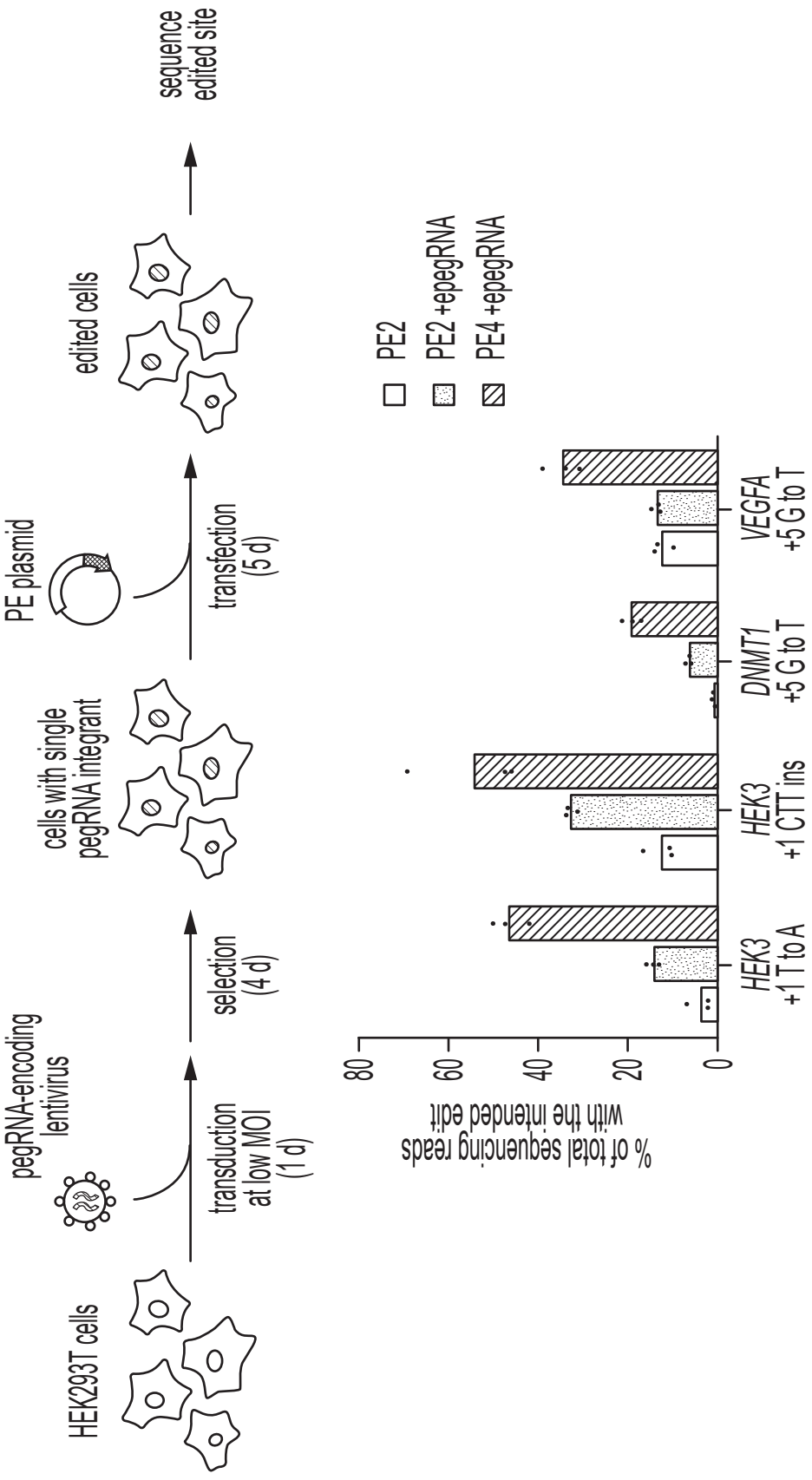

HEK293T cells
pegRNA-encoding lentivirus
transduction at low MOI (1 d)
selection (4 d)
cells with single pegRNA integrant
PE plasmid
transfection (5 d)
edited cells
sequence edited site
PE2
PE2 +epegRNA
PE4 +epegRNA
% of total sequencing reads with the intended edit
0
20
40
60
80
HEK3 +1 T to A
HEK3 +1 CTT ins
DNMT1 +5 G to T
VEGFA +5 G to T

FIG. 60

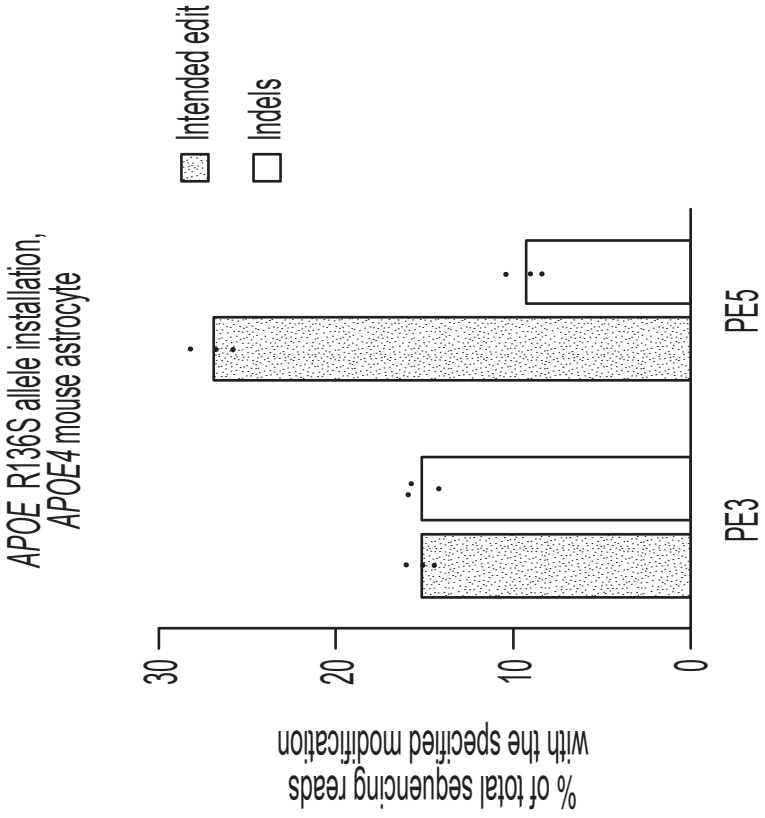

APOE R136S allele installation,
APOE4 mouse astrocyte
Intended edit
Indels
% of total sequencing reads
with the specified modification
30
20
10
0
PE3
PE5

FIG. 61

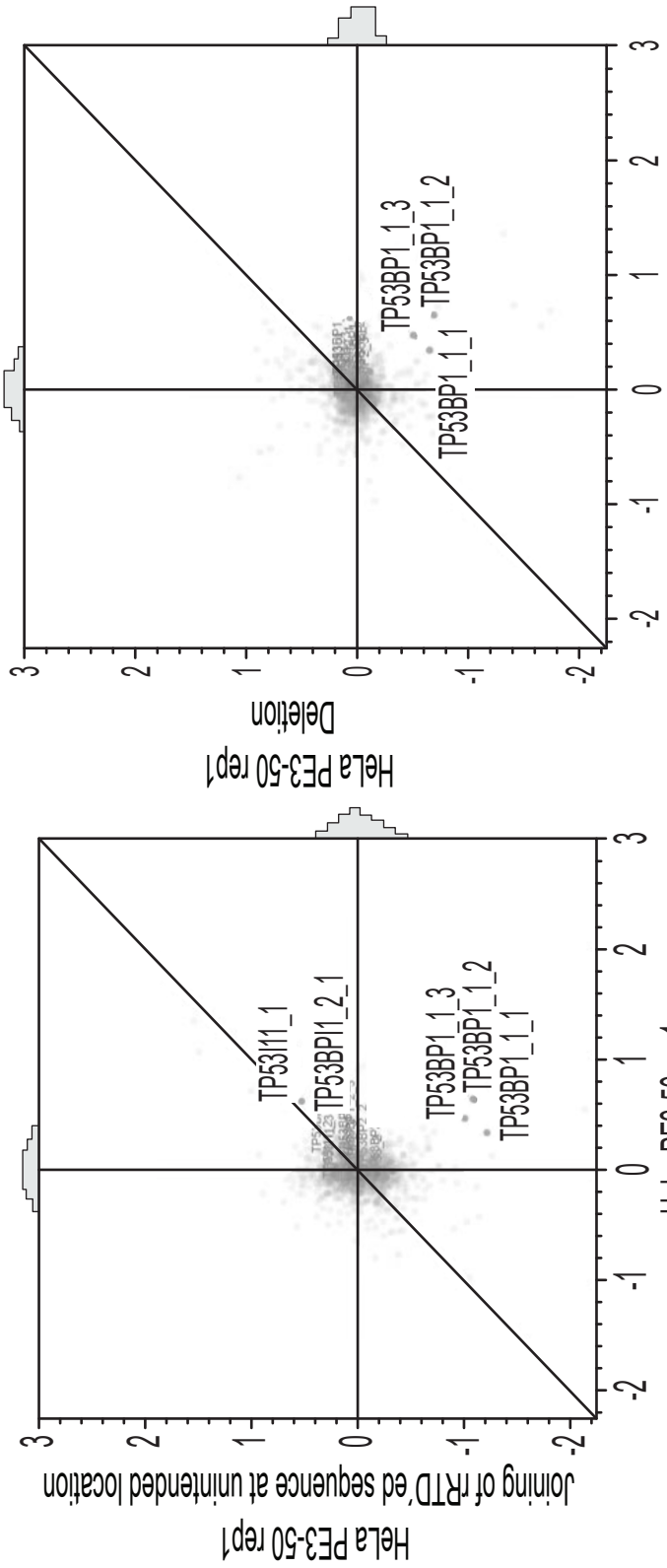


FIG. 62A

FIG. 62B

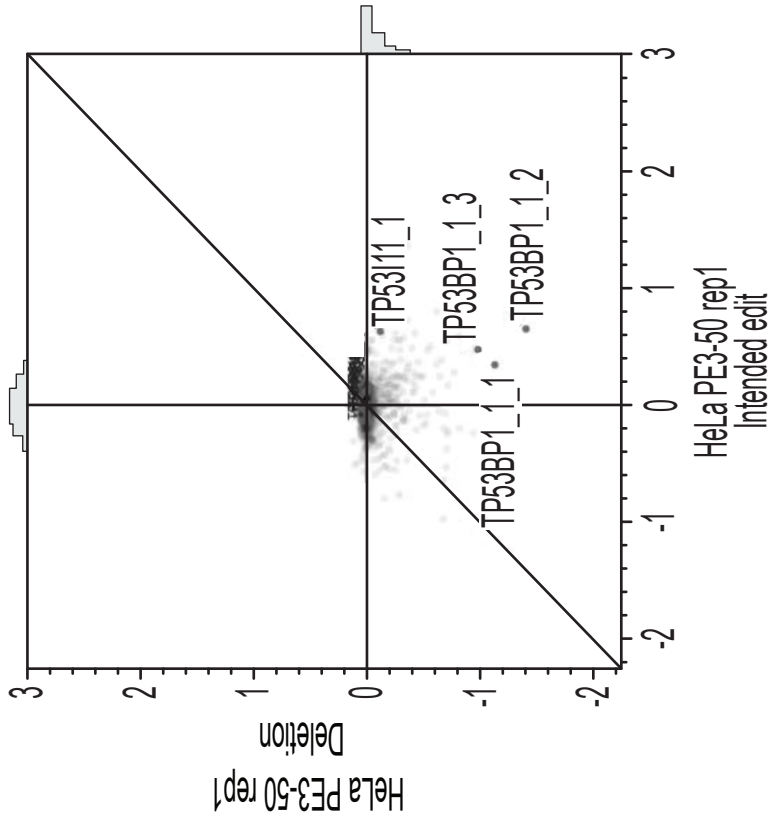

HeLa PE3-50 rep1
Deletion
3
2
1
0
-1
-2
TP53I11_1
TP53BP1_1_3
TP53BP1_1_2
TP53BP1_1_1
HeLa PE3-50 rep1
Intended edit

FIG. 62C

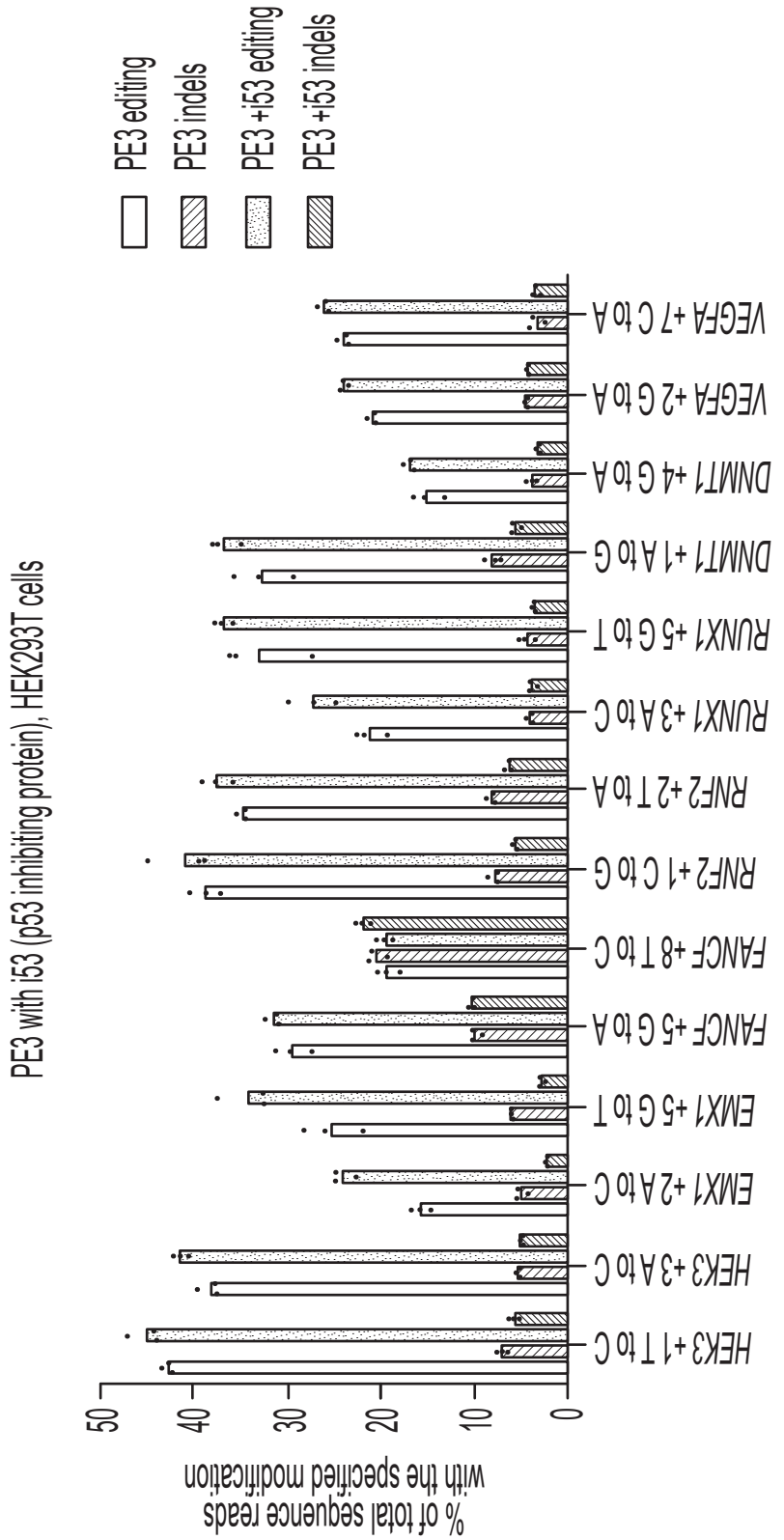

PE3 editing
PE3 indels
PE3 +i53 editing
PE3 +i53 indels
PE3 with i53 (p53 inhibiting protein), HEK293T cells
% of total sequence reads with the specified modification
0
10
20
30
40
50
HEK3 +1 T to C
HEK3 +3 A to C
EMX1 +2 A to C
EMX1 +5 G to T
FANCF +5 G to A
FANCF +8 T to C
RNF2 +1 C to G
RNF2 +2 T to A
RUNX1 +3 A to C
RUNX1 +5 G to T
DNMT1 +1 A to G
DNMT1 +4 G to A
VEGFA +2 G to A
VEGFA +7 C to A

FIG. 63

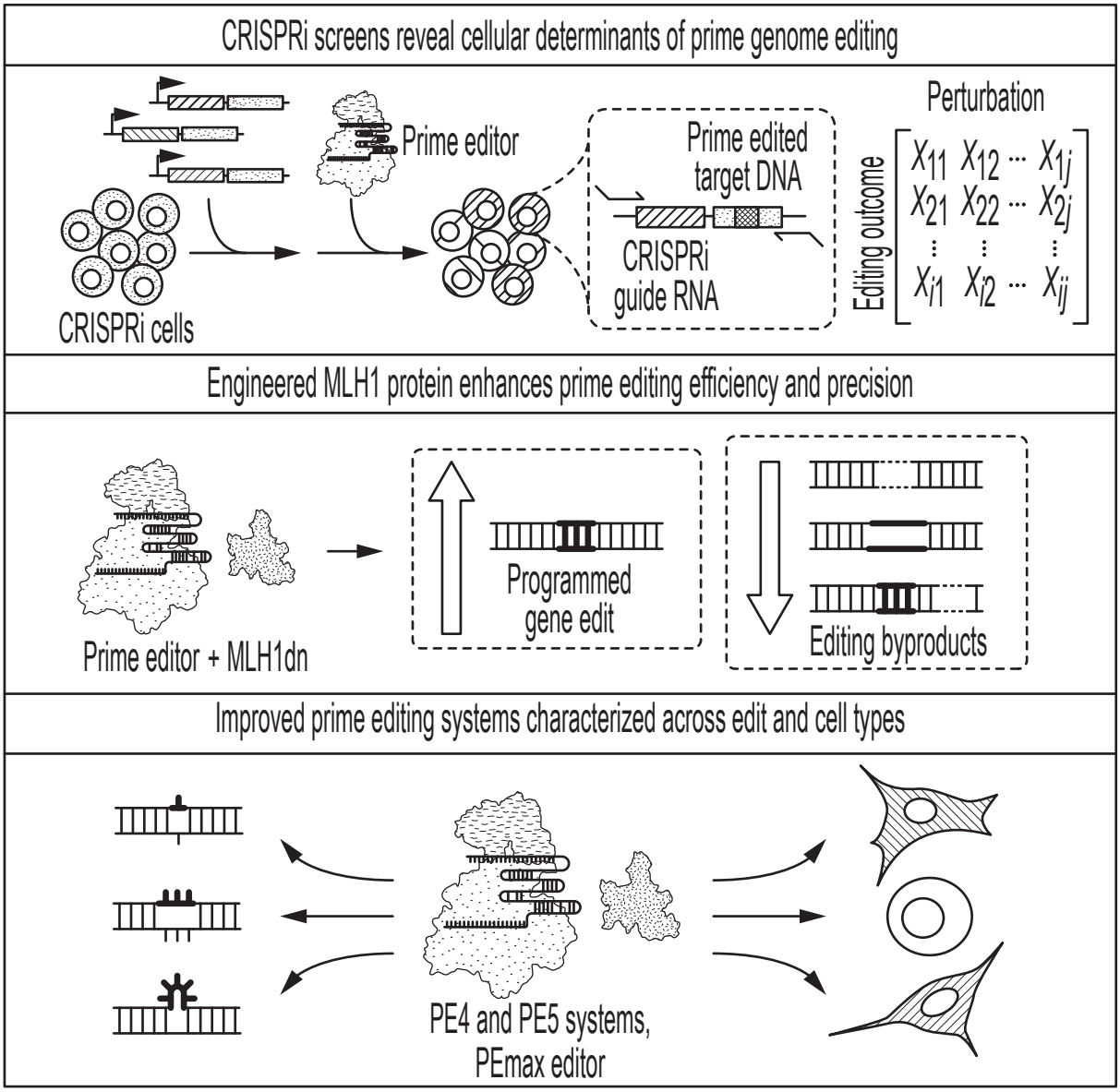
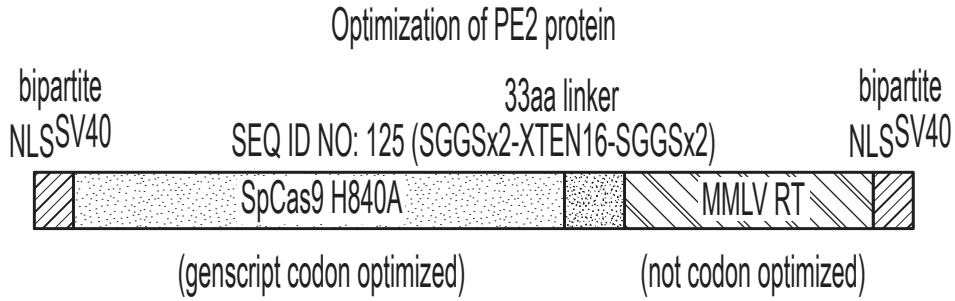

CRISPRi screens reveal cellular determinants of prime genome editing
Prime editor
Prime edited target DNA
CRISPRi guide RNA
CRISPRi cells
Perturbation
Editing outcome
Engineered MLH1 protein enhances prime editing efficiency and precision
Prime editor + MLH1dn
Programmed gene edit
Editing byproducts
Improved prime editing systems characterized across edit and cell types
PE4 and PE5 systems, PEmax editor

FIG. 64



1. Codon optimize MMLV RT
2. Mutations to SpCas9 to improve activity
3. Linker optimization
4. Improve nuclear localization

FIG. 65

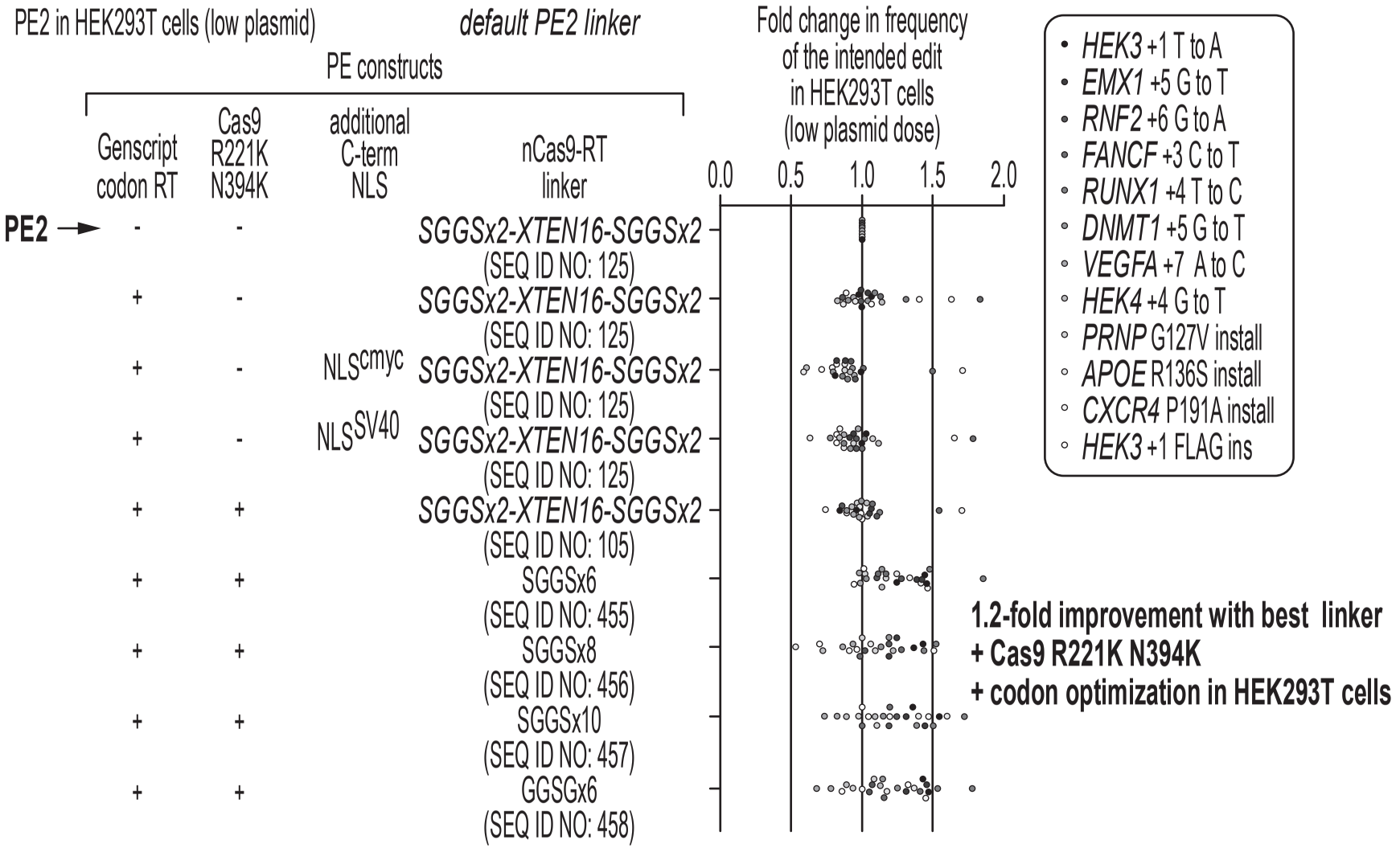

PE2 in HEK293T cells (low plasmid)
default PE2 linker
PE constructs
Genscript codon RT
Cas9 R221K N394K
additional C-term NLS
nCas9-RT linker
PE2
SGGSx2-XTEN16-SGGSx2 (SEQ ID NO: 125)
NLScmyc
NLSSV40
SGGSx2-XTEN16-SGGSx2 (SEQ ID NO: 105)
SGGSx6 (SEQ ID NO: 455)
SGGSx8 (SEQ ID NO: 456)
SGGSx10 (SEQ ID NO: 457)
GGSGx6 (SEQ ID NO: 458)
Fold change in frequency of the intended edit in HEK293T cells (low plasmid dose)
0.0
0.5
1.0
1.5
2.0
HEK3 +1 T to A
EMX1 +5 G to T
RNF2 +6 G to A
FANCF +3 C to T
RUNX1 +4 T to C
DNMT1 +5 G to T
VEGFA +7 A to C
HEK4 +4 G to T
PRNP G127V install
APOE R136S install
CXCR4 P191A install
HEK3 +1 FLAG ins
1.2-fold improvement with best linker
+ Cas9 R221K N394K
+ codon optimization in HEK293T cells

FIG. 66

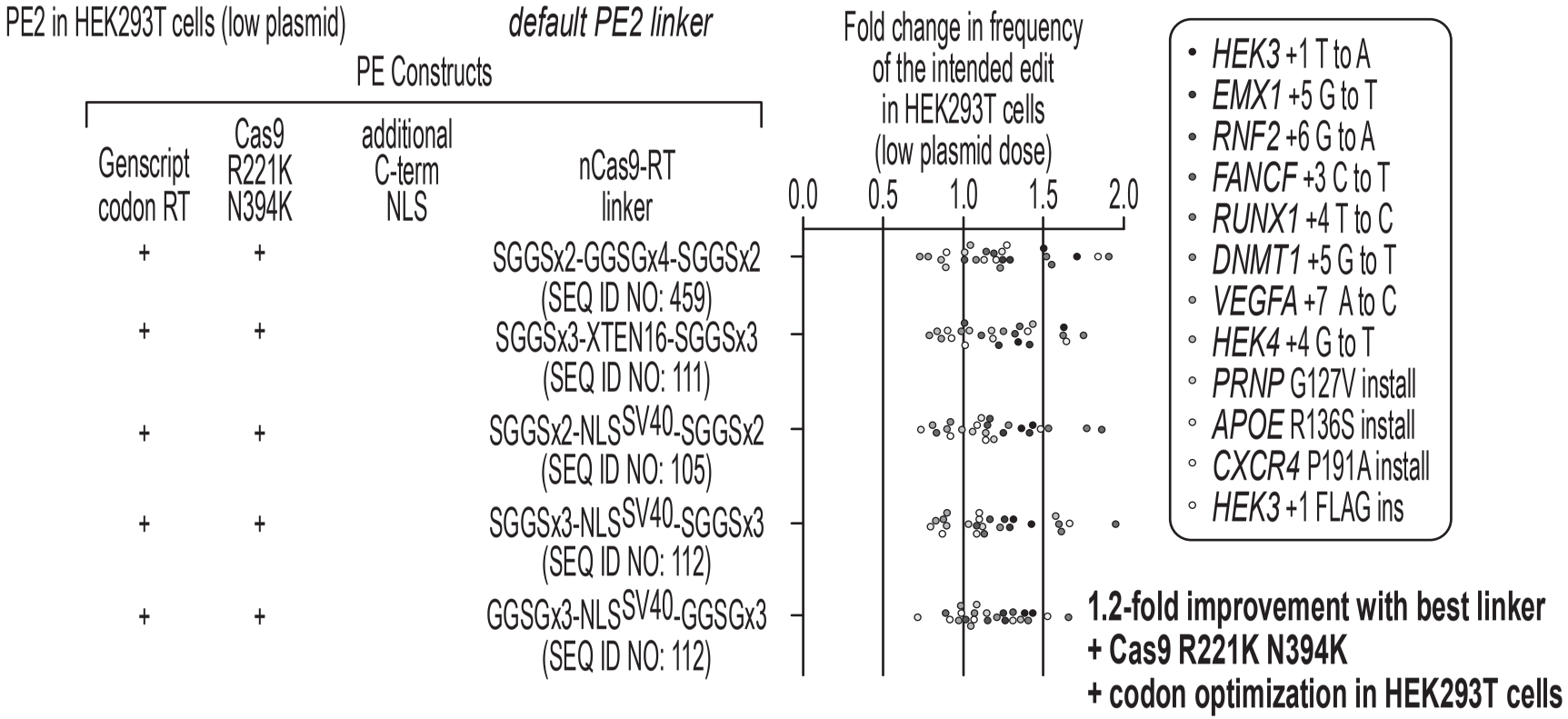


FIG. 66 CONTINUED

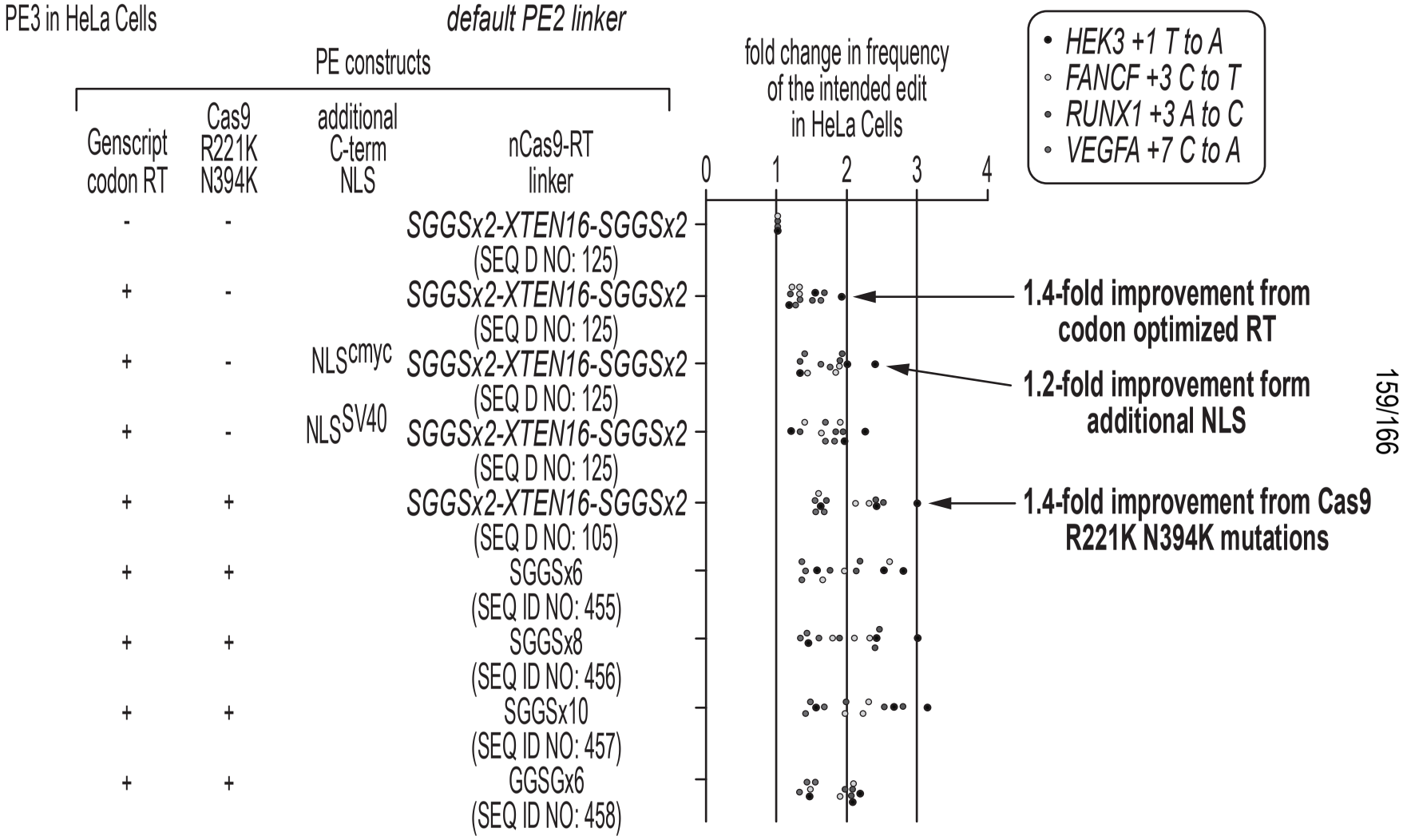

PE3 in HeLa Cells
default PE2 linker
PE constructs
Genscript codon RT
Cas9 R221K N394K
additional C-term NLS
nCas9-RT linker
fold change in frequency of the intended edit in HeLa Cells
0
1
2
3
4
HEK3 +1 T to A
FANCF +3 C to T
RUNX1 +3 A to C
VEGFA +7 C to A
- - SGGSx2-XTEN16-SGGSx2 (SEQ D NO: 125)
+ - SGGSx2-XTEN16-SGGSx2 (SEQ D NO: 125)
+ - NLScmyc SGGSx2-XTEN16-SGGSx2 (SEQ D NO: 125)
+ - NLSSV40 SGGSx2-XTEN16-SGGSx2 (SEQ D NO: 125)
+ + SGGSx2-XTEN16-SGGSx2 (SEQ D NO: 105)
+ + SGGSx6 (SEQ ID NO: 455)
+ + SGGSx8 (SEQ ID NO: 456)
+ + SGGSx10 (SEQ ID NO: 457)
+ + GGSGx6 (SEQ ID NO: 458)
1.4-fold improvement from codon optimized RT
1.2-fold improvement form additional NLS
1.4-fold improvement from Cas9 R221K N394K mutations

FIG. 67

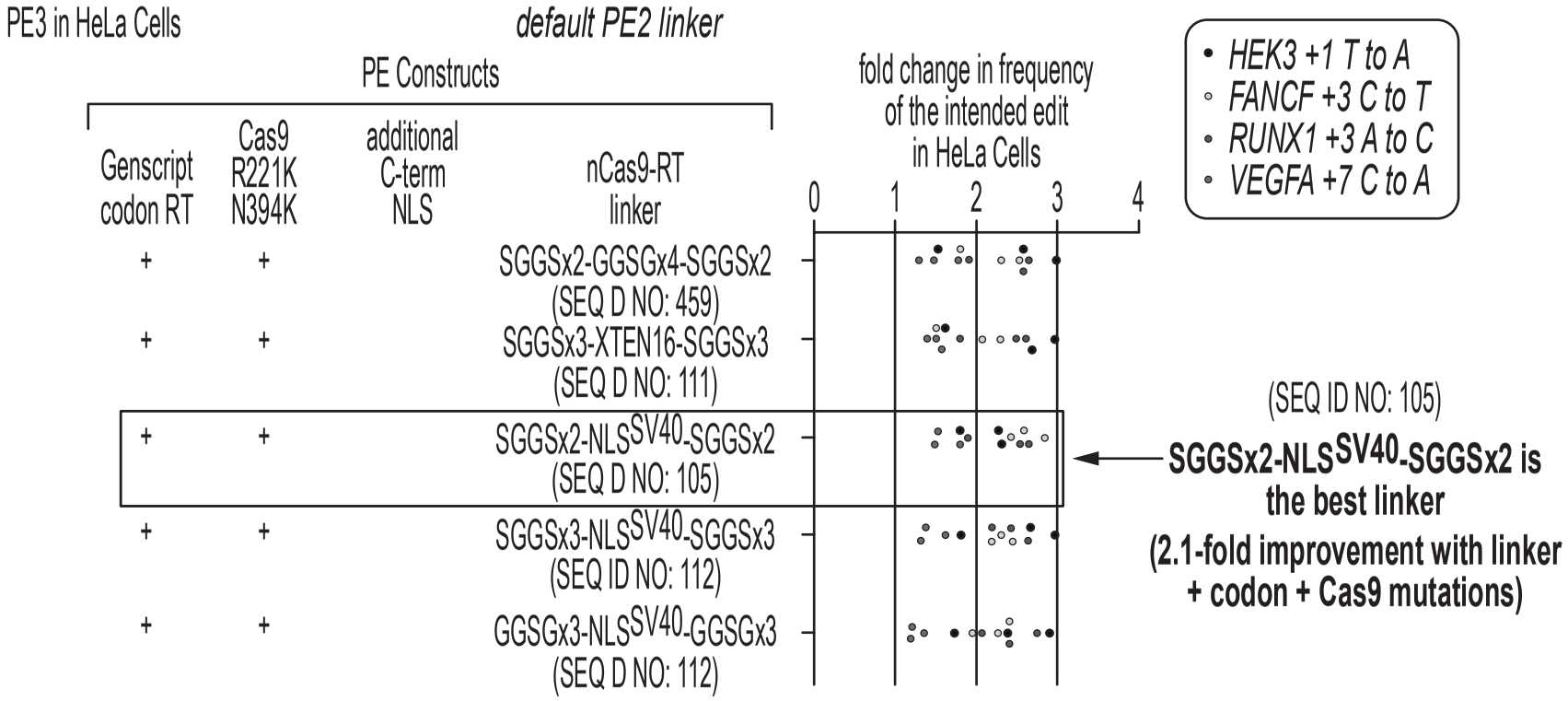


FIG. 67 CONTINUED

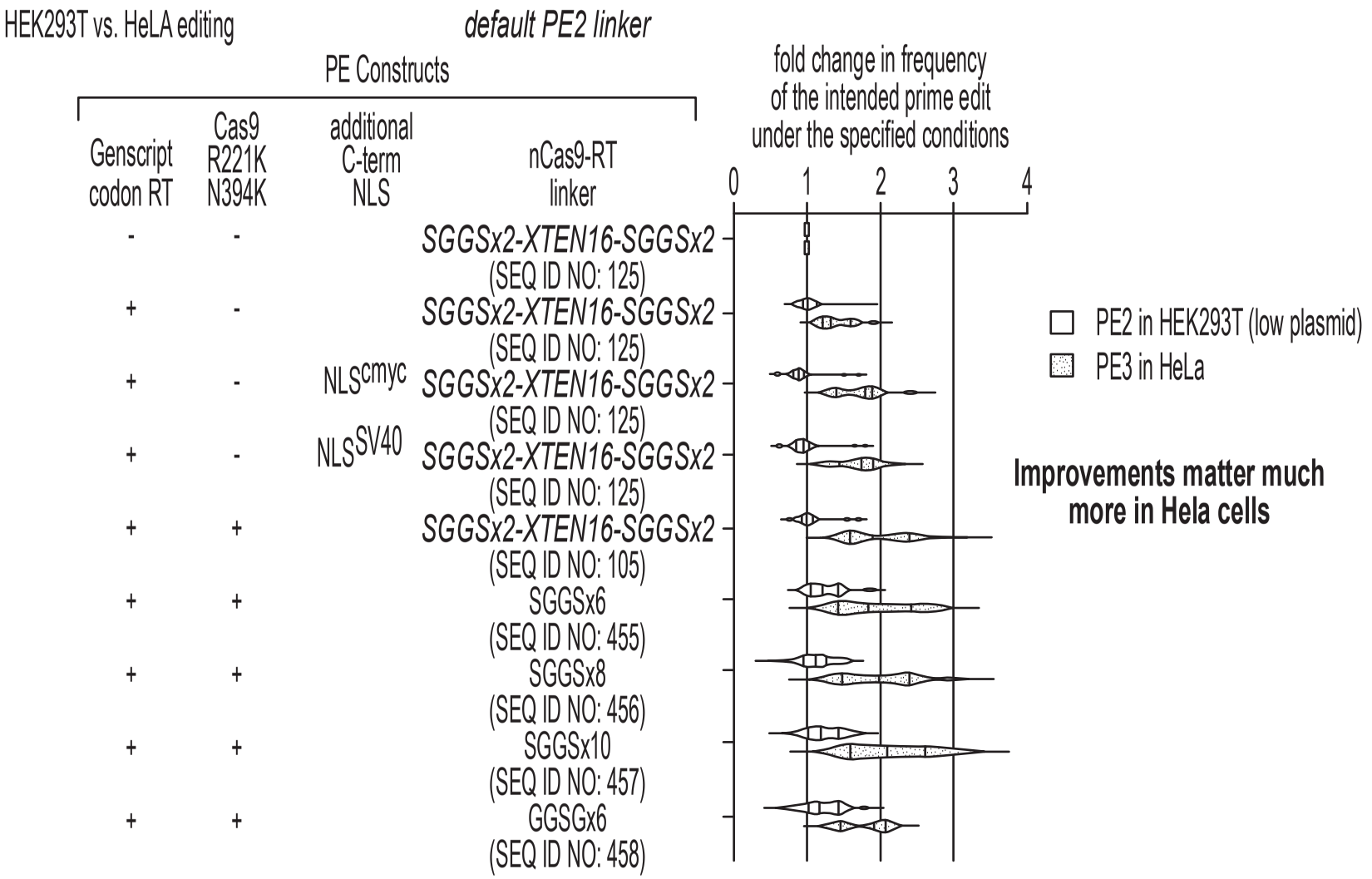

HEK293T vs. HeLA editing
default PE2 linker
PE Constructs
Genscript codon RT
Cas9 R221K N394K
additional C-term NLS
nCas9-RT linker
fold change in frequency of the intended prime edit under the specified conditions
0
1
2
3
4
- - SGGSx2-XTEN16-SGGSx2 (SEQ ID NO: 125)
+ - SGGSx2-XTEN16-SGGSx2 (SEQ ID NO: 125)
+ - NLScmyc SGGSx2-XTEN16-SGGSx2 (SEQ ID NO: 125)
+ - NLSSV40 SGGSx2-XTEN16-SGGSx2 (SEQ ID NO: 125)
+ + SGGSx2-XTEN16-SGGSx2 (SEQ ID NO: 105)
+ + SGGSx6 (SEQ ID NO: 455)
+ + SGGSx8 (SEQ ID NO: 456)
+ + SGGSx10 (SEQ ID NO: 457)
+ + GGSGx6 (SEQ ID NO: 458)
PE2 in HEK293T (low plasmid)
PE3 in HeLa
Improvements matter much more in Hela cells

FIG. 68

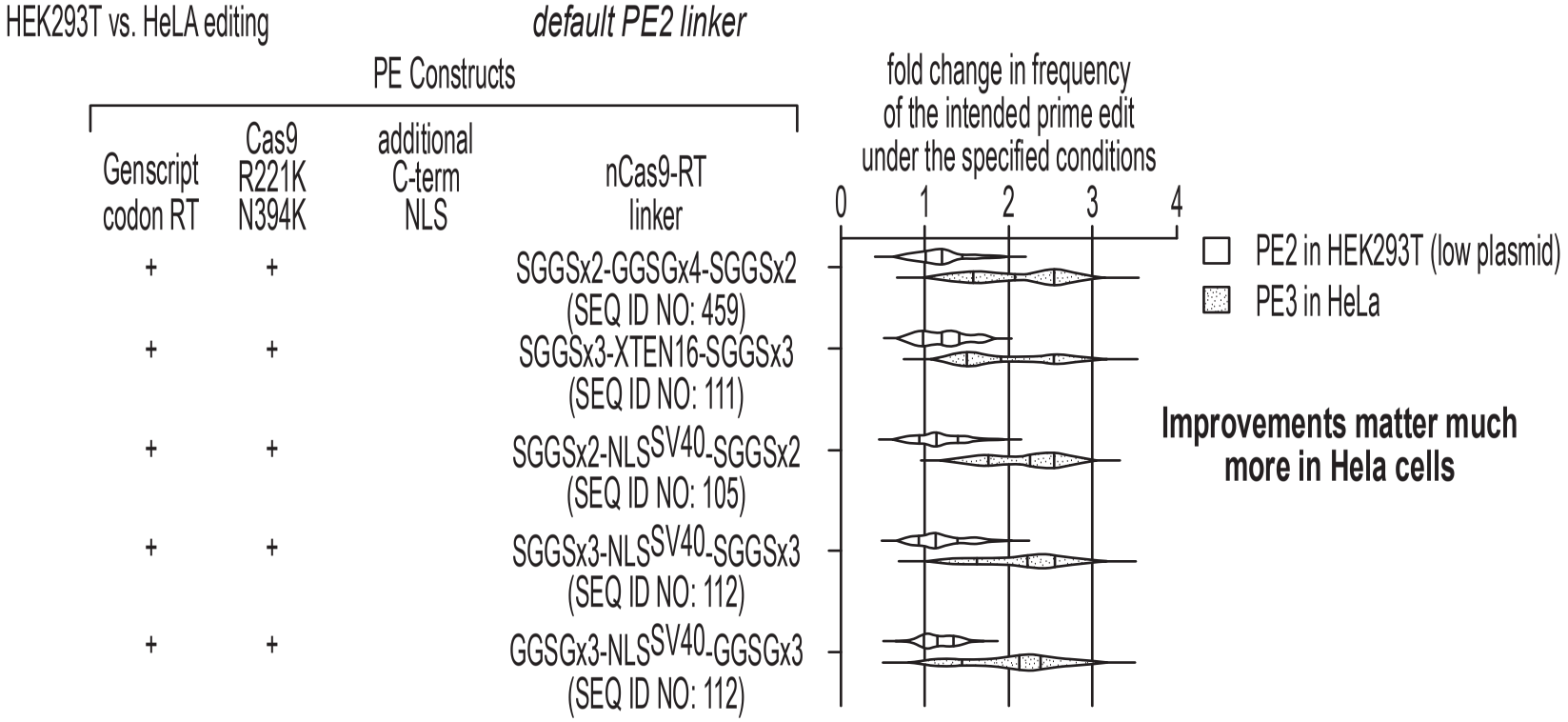


FIG. 68 CONTINUED

PE3 in HeLa cells, NLS architecture optimization

PE constructs				
Genscript codon RT	Cas9 R221K N394K	nCas9-RT linker	N-term NLS	C-term NLS
-	-	SGGSx2-XTEN16-SGGSx2 (SEQ ID NO: 125)	NLS^{SV40}	$*NLS^{SV40}$
+	-	SGGSx2-XTEN16-SGGSx2 (SEQ ID NO: 125)	NLS^{SV40}	$*NLS^{SV40}$
+	+	SGGSx2-XTEN16-SGGSx2 (SEQ ID NO: 125)	NLS^{SV40}	$*NLS^{SV40}$
+	+	SGGSx2-NLS^{SV40}-SGGSx2 (SEQ ID NO: 105)	NLS^{SV40}	$*NLS^{SV40}$
+	+	SGGSx2-NLS^{SV40}-SGGSx2 (SEQ D NO: 105)	NLS^{SV40}	NLS^{SV40}

default PE2 linker
$*NLS^{SV40}$ is missing 1aa

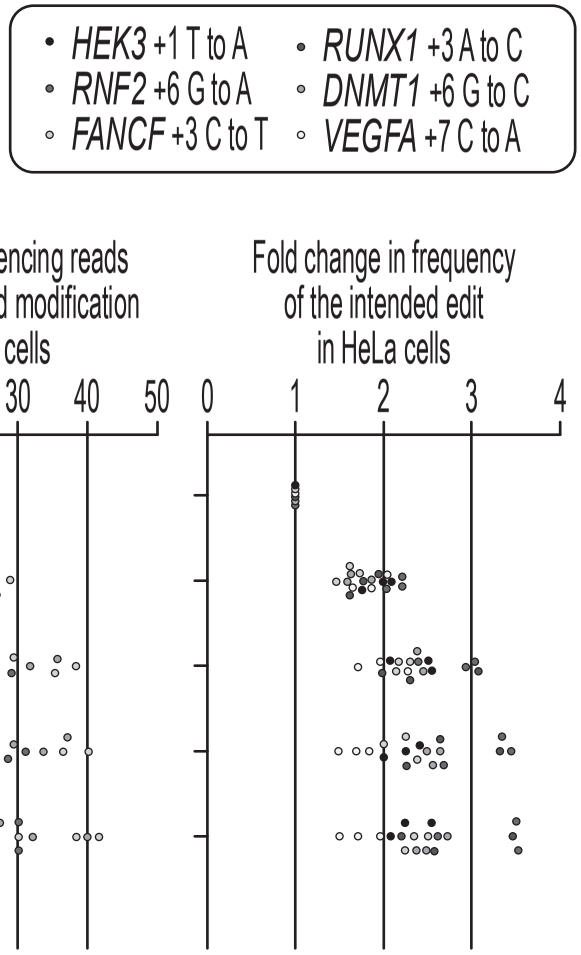


FIG. 69

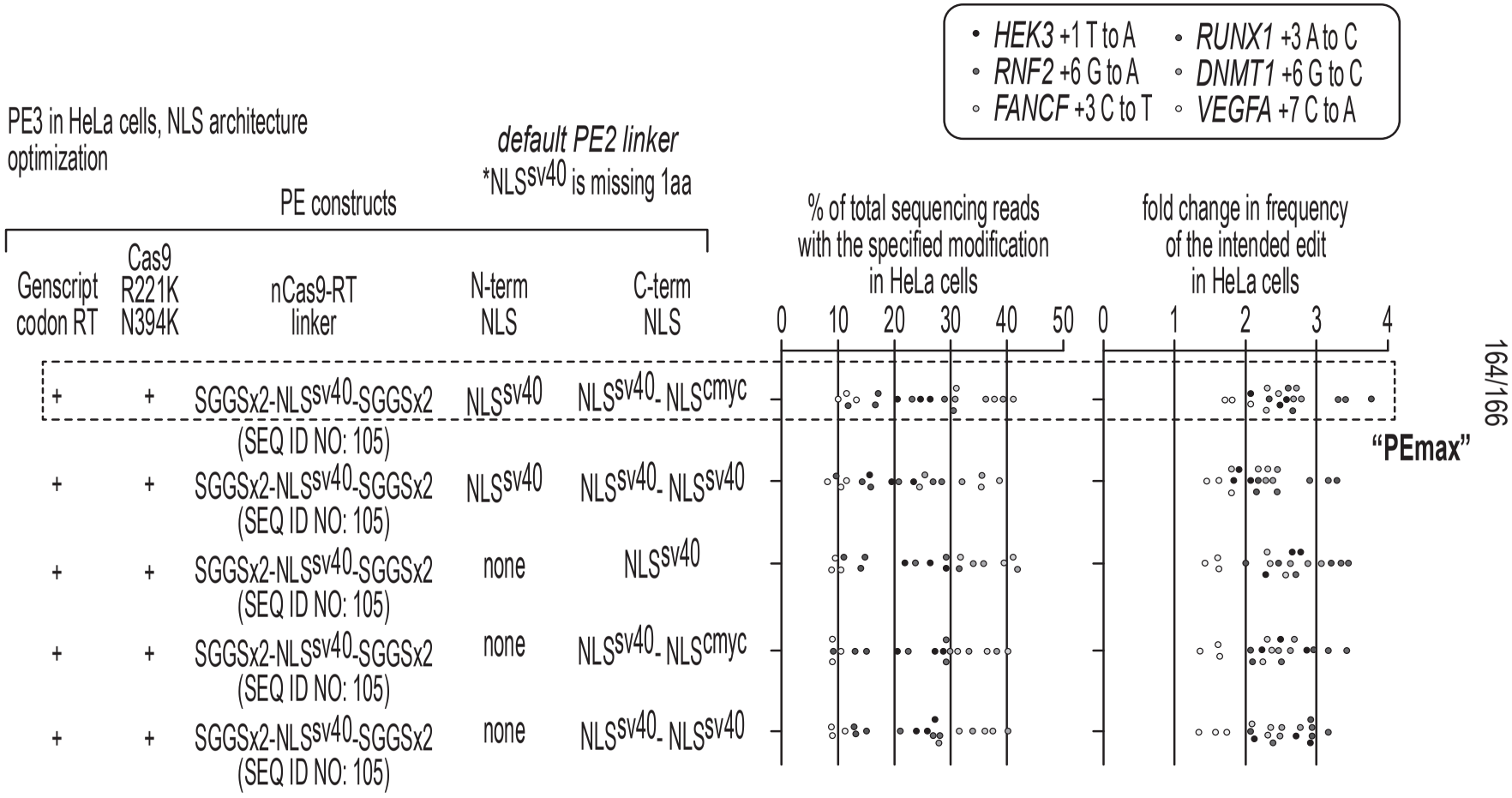


FIG. 69 CONTINUED

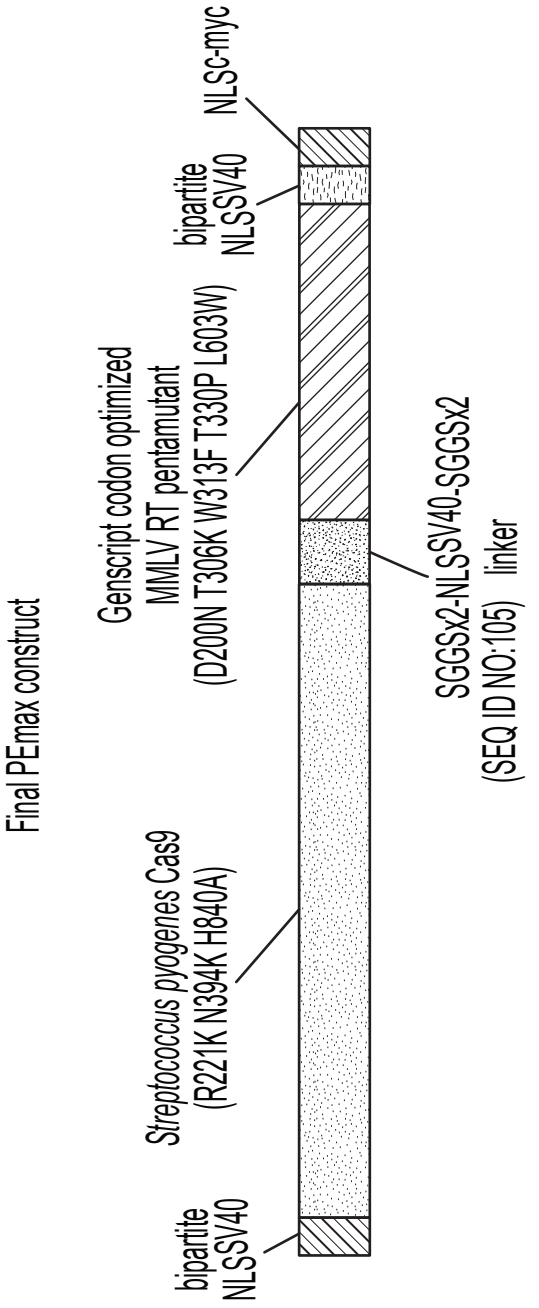

Final PEmax construct
bipartite NLSSV40
Streptococcus pyogenes Cas9 (R221K N394K H840A)
Genscript codon optimized MMLV RT pentamutant (D200N T306K W313F T330P L603W)
bipartite NLSSV40
NLSc-myc
SGGSx2-NLSSV40-SGGSx2 (SEQ ID NO:105) linker
PEmax improves prime editing efficiency by 2.59-fold across 6 sites in HeLa cells

FIG. 70

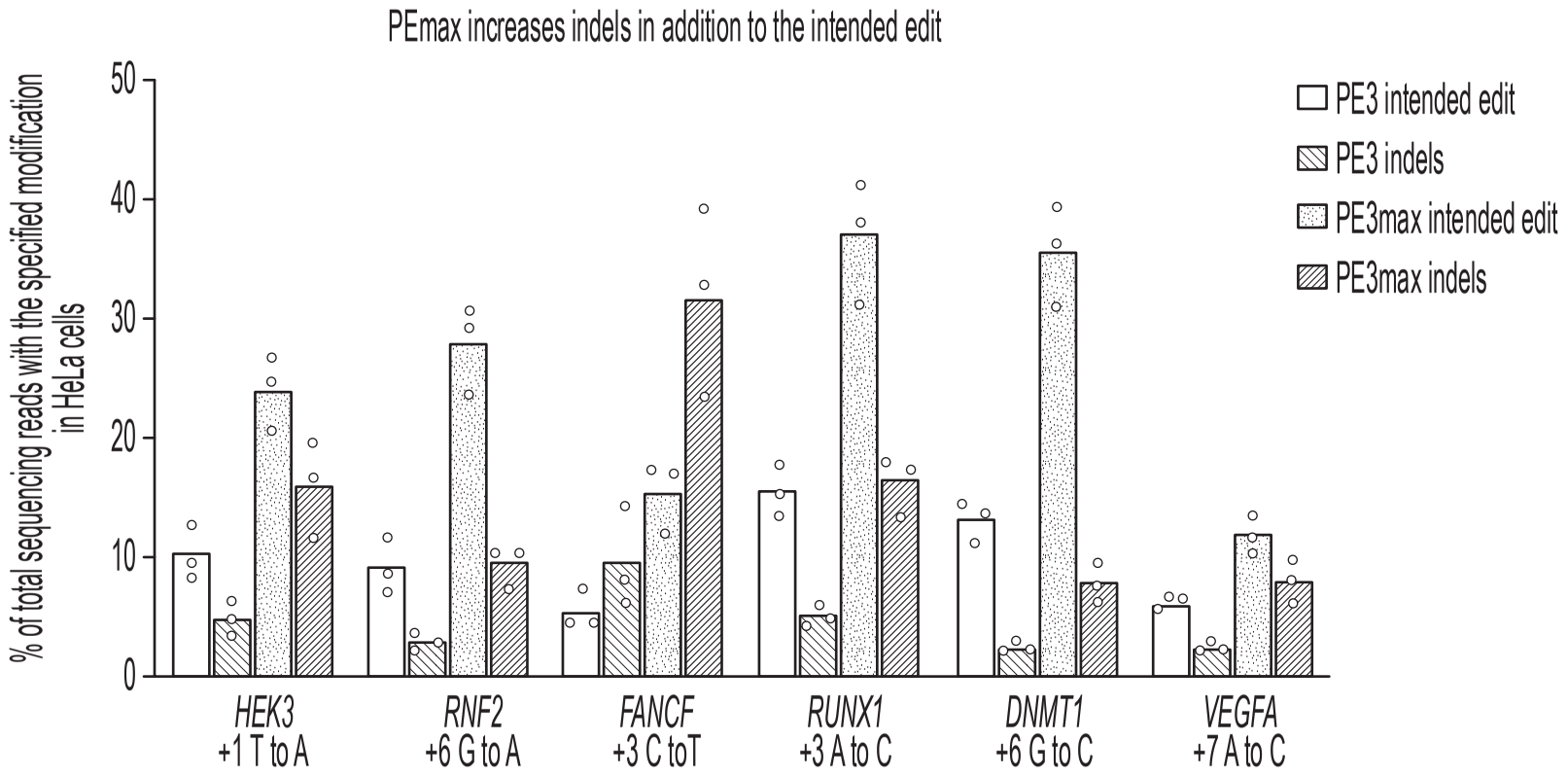

PEmax increases indels in addition to the intended edit
% of total sequencing reads with the specified modification in HeLa cells
50
40
30
20
10
0
PE3 intended edit
PE3 indels
PE3max intended edit
PE3max indels
HEK3 +1 T to A
RNF2 +6 G to A
FANCF +3 C toT
RUNX1 +3 A to C
DNMT1 +6 G to C
VEGFA +7 A to C

FIG. 71