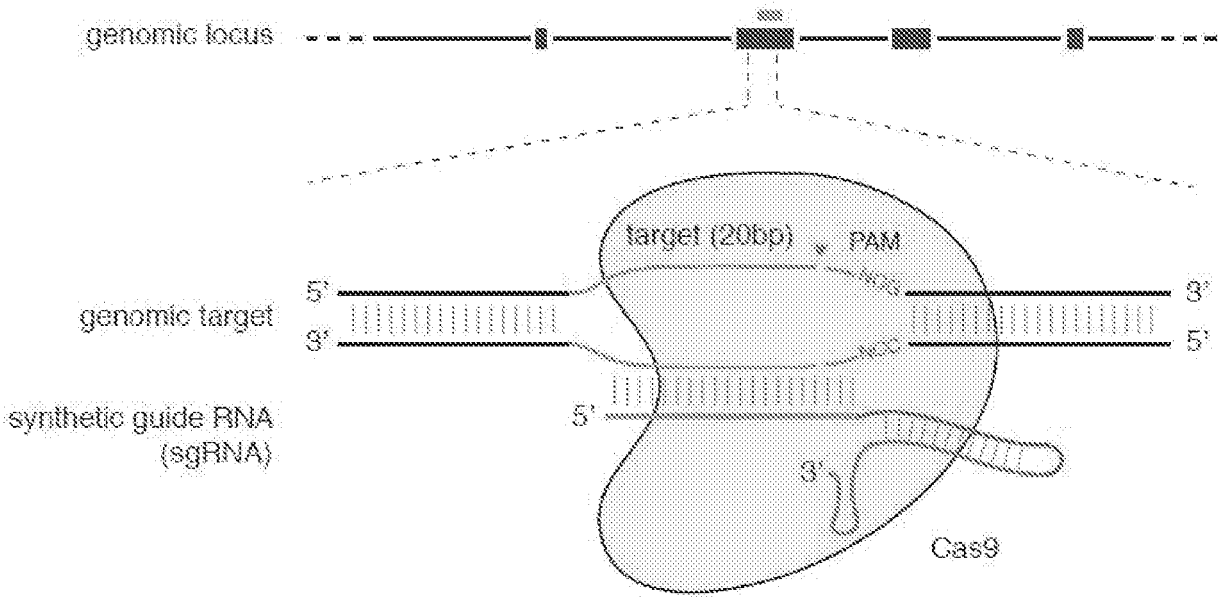

genomic locus
target (20bp)
PAM
NNG
genomic target
5'
3'
3'
5'
NCC
synthetic guide RNA
(sgRNA)
5'
3'
Cas9

FIG. 1

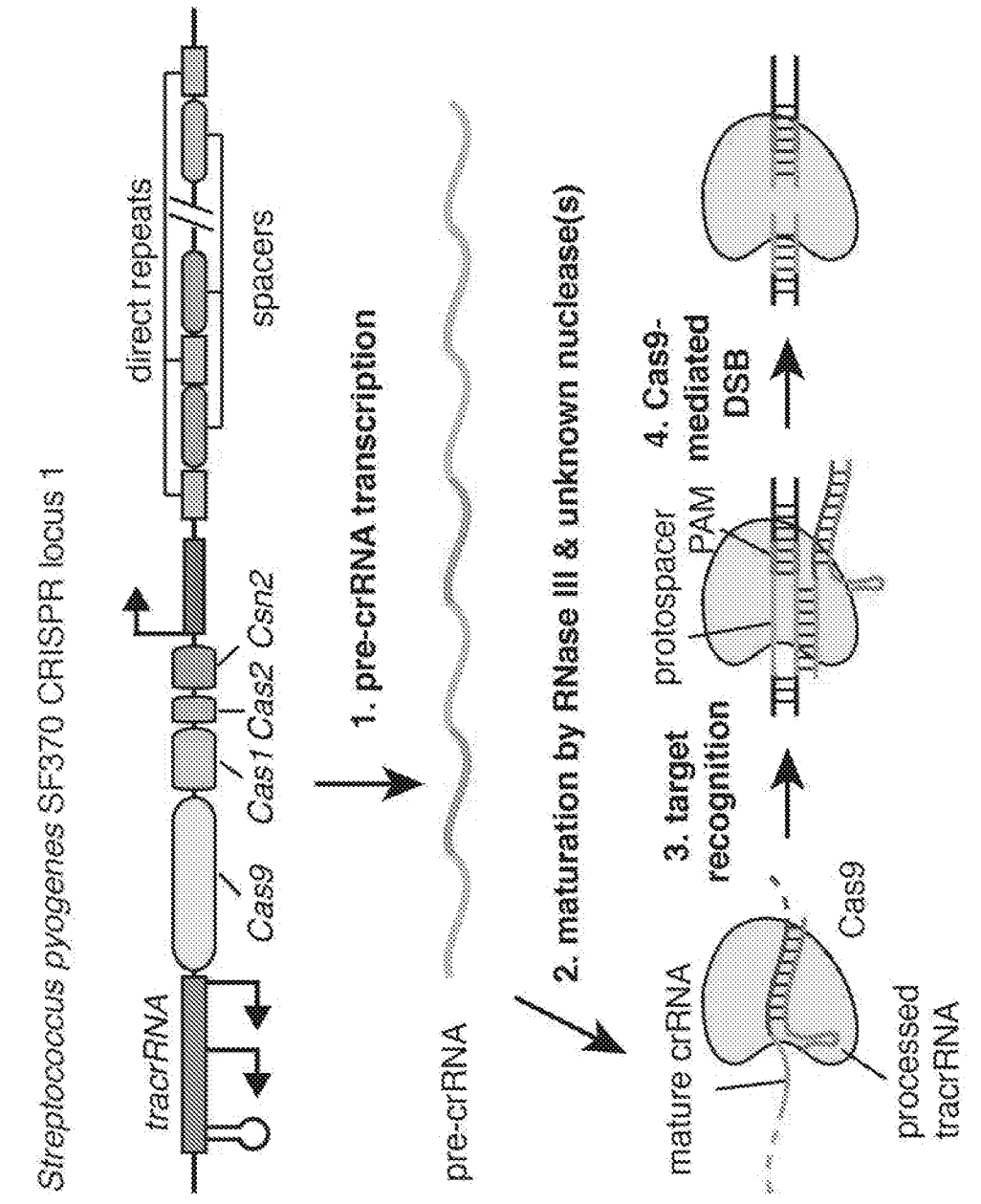

A
Streptococcus pyogenes SF370 CRISPR locus 1
direct repeats
spacers
tracrRNA
Cas9
Cas1
Cas2
Csn2
1. pre-crRNA transcription
pre-crRNA
2. maturation by RNase III & unknown nuclease(s)
mature crRNA
3. target recognition
protospacer
PAM
4. Cas9-mediated DSB
Cas9
processed tracrRNA

FIG. 2A

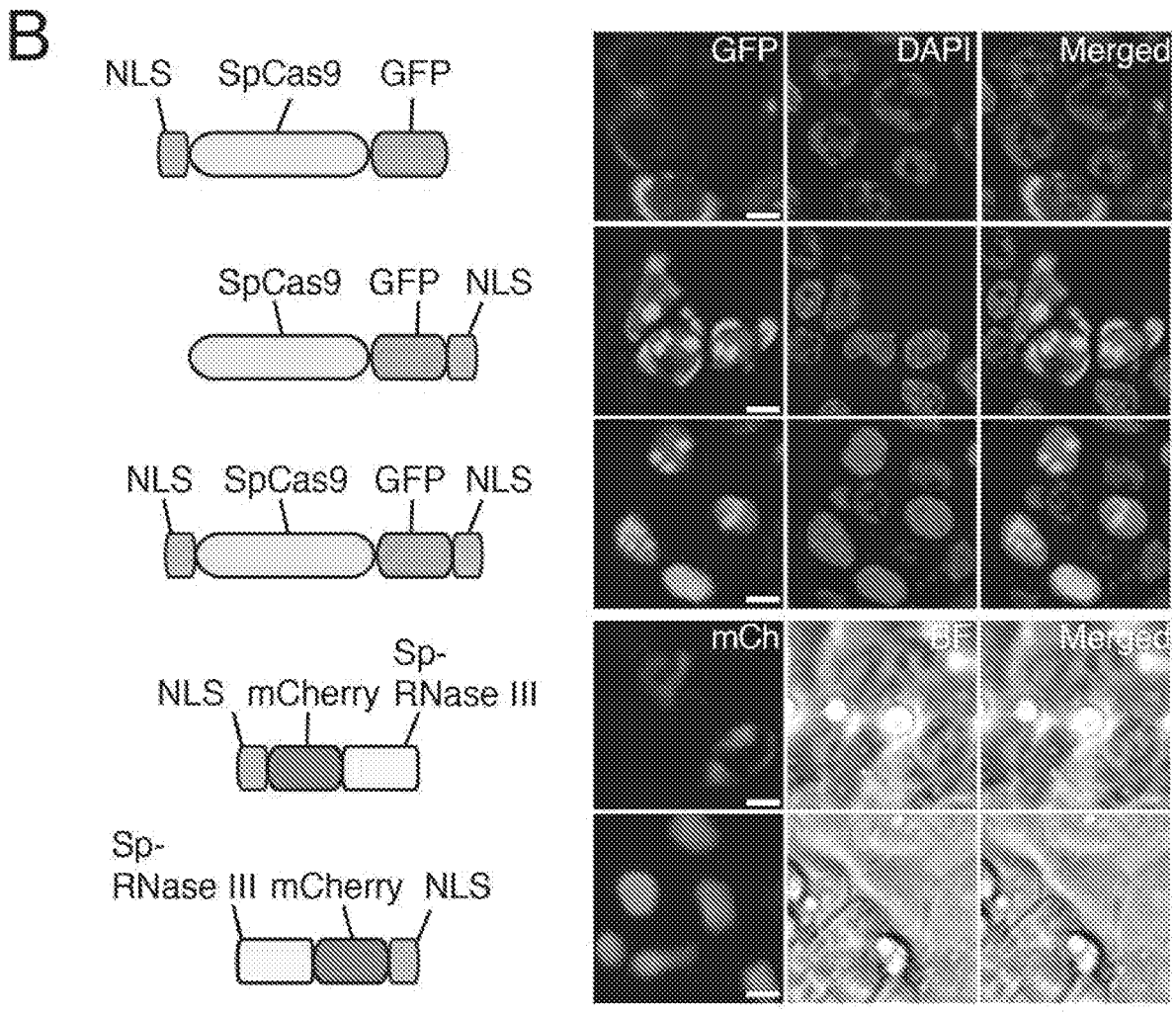

B
NLS SpCas9 GFP
SpCas9 GFP NLS
NLS SpCas9 GFP NLS
Sp-
NLS mCherry RNase III
Sp-
RNase III mCherry NLS
GFP
DAPI
Merged
mCh

FIG. 2B

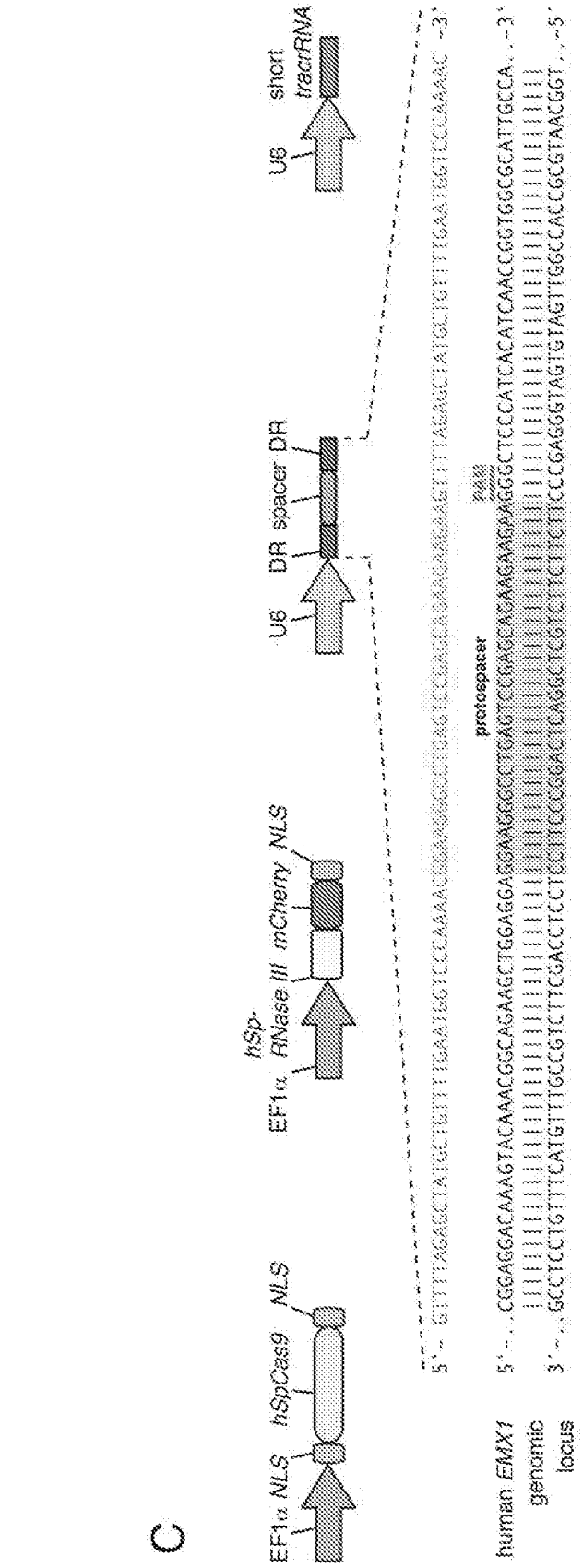

C
EF1α NLS hSpCas9 NLS
EF1α hSp. RNase III mCherry NLS
U6 DR spacer DR
U6 short tracrRNA
5'- GTTTTAGAGCTATGCTGTTTTGAATGGTCCCAAAACGGAAGGCCTGAGTCCGAGCAGAAGAAGAAGGGCTCCCATCACATCAACCGGTGGCGCATTGCCA -3'
protospacer
PAM
human EMX1 genomic locus
5'-...CGGAGGACAAAGTACAAACGGCAGAAGCTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGAAGAAGGGCTCCCATCACATCAACCGGTGGCGCATTGCCA...-3'
3'-...GCCTCCTGTTTCATGTTTGCCGTCTTCGACCTCCTCCCGGACTCAGGCTCGTCTTCTTCTTCCCGAGGGTAGTGTAGTTGGCCACCGCGTAACGGT...-5'

FIG. 2C

D

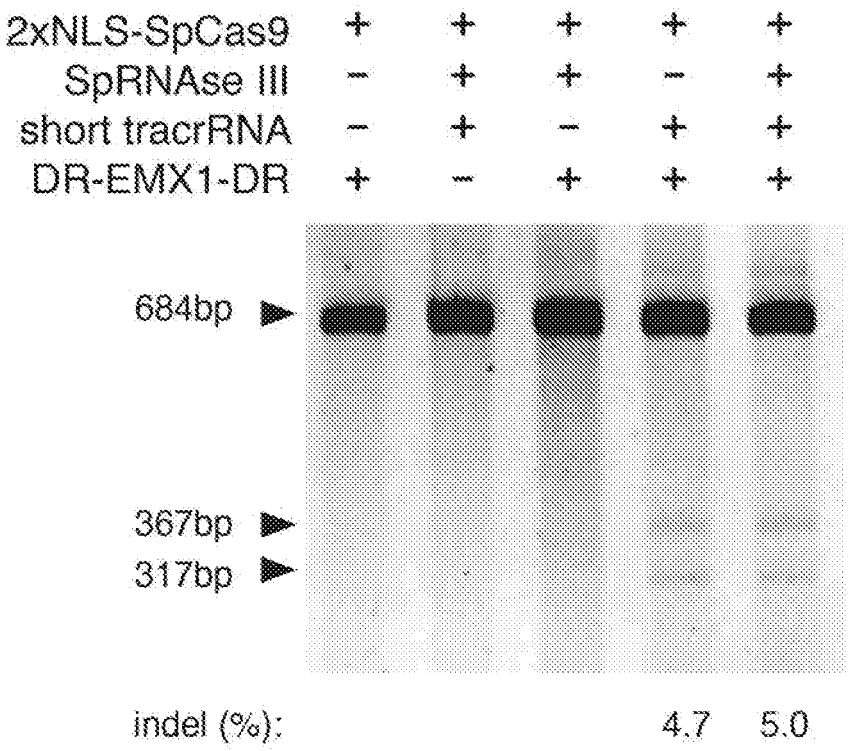

2xNLS-SpCas9 + + + + +
SpRNAse III − + + − +
short tracrRNA − + − + +
DR-EMX1-DR + − + + +
684bp
367bp
317bp
indel (%): 4.7 5.0

FIG. 2D

E

F

human *EMX1* protospacer target (mutation in 5 of 43 sequenced clones = 11.6%)

WT	5' - . . CTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAG AAGAAGGGCTCCCATCACAT . . -3'
Δ1	CTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAG---AGAAGGGCTCCCATCACAT
+1	CTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGAAAGAAGGGCTCCCATCACAT
Δ3	CTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAG-----AAGGGCTCCCATCACAT
m1, Δ6	CTGGAGGAGGAAGGGCCTGAGCCCGAGCAGAAG----------GGCTCCCATCACAT

FIG. 2E-F

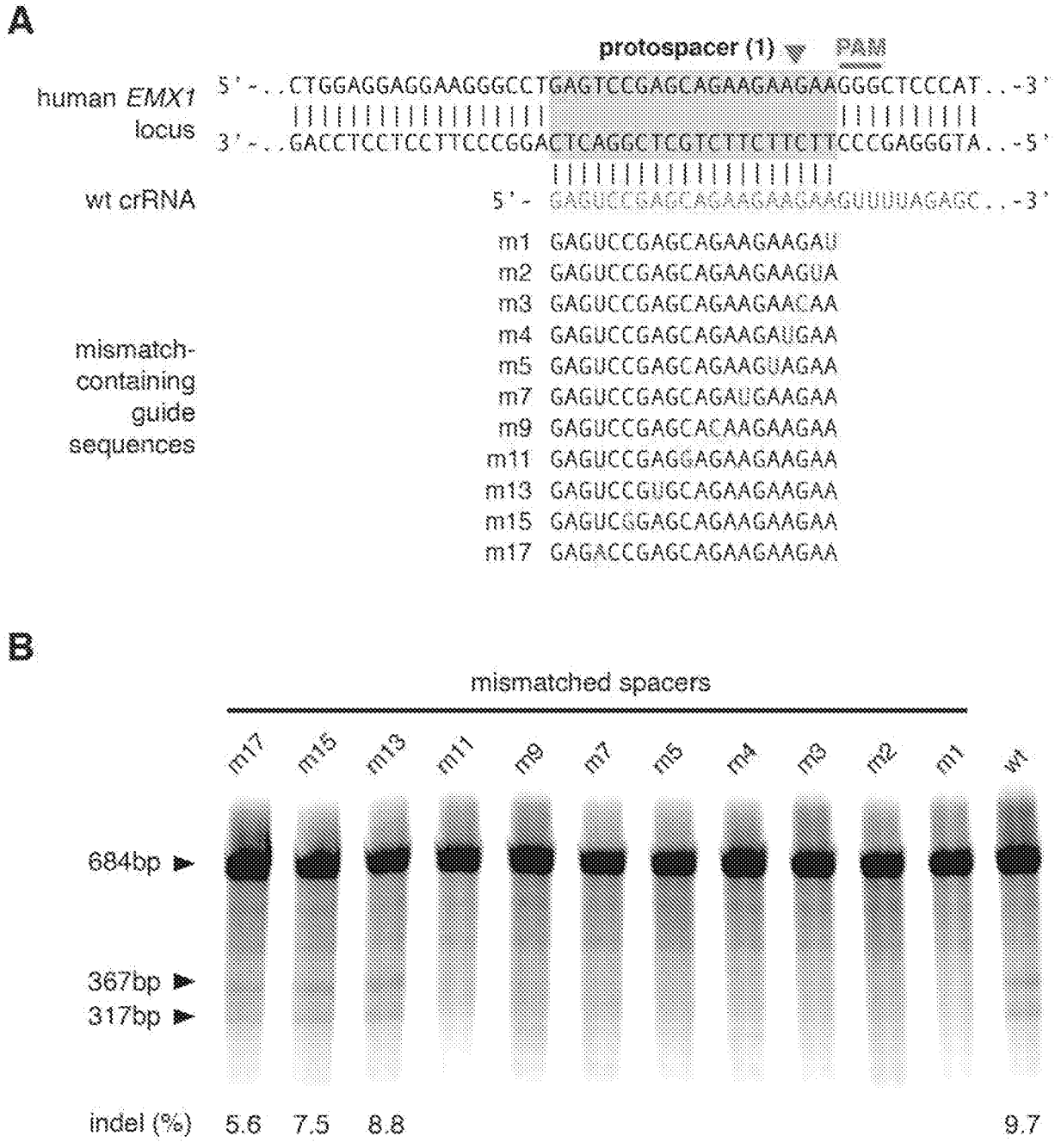

A
protospacer (1)
PAM
human EMX1 locus
5'-..CTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGAAGAAGGGCTCCCAT..-3'
3'-..GACCTCCTCCTTCCCGGACTCAGGCTCGTCTTCTTCTTCCCGAGGGTA..-5'
wt crRNA
5'- GAGUCCGAGCAGAAGAAGAAGUUUUAGAGC..-3'
mismatch-containing guide sequences
m1 GAGUCCGAGCAGAAGAAGAU
m2 GAGUCCGAGCAGAAGAAGUA
m3 GAGUCCGAGCAGAAGAACAA
m4 GAGUCCGAGCAGAAGAUGAA
m5 GAGUCCGAGCAGAAGUAGAA
m7 GAGUCCGAGCAGAUGAAGAA
m9 GAGUCCGAGCACAAGAAGAA
m11 GAGUCCGAGGAGAAGAAGAA
m13 GAGUCCGUGCAGAAGAAGAA
m15 GAGUCGGAGCAGAAGAAGAA
m17 GAGACCGAGCAGAAGAAGAA
B
mismatched spacers
m17
m15
m13
m11
m9
m7
m5
m4
m3
m2
m1
wt
684bp
367bp
317bp
indel (%)
5.6
7.5
8.8
9.7
C
left TALEN binding site
protospacer (1)
PAM
human EMX1 locus
5'-..CTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGAAGAAGGGCTCCCAT..-3'
3'-..GACCTCCTCCTTCCCGGACTCAGGCTCGTCTTCTTCTTCCCGAGGGTA..-5'
right TALEN binding site

FIG. 3A-C

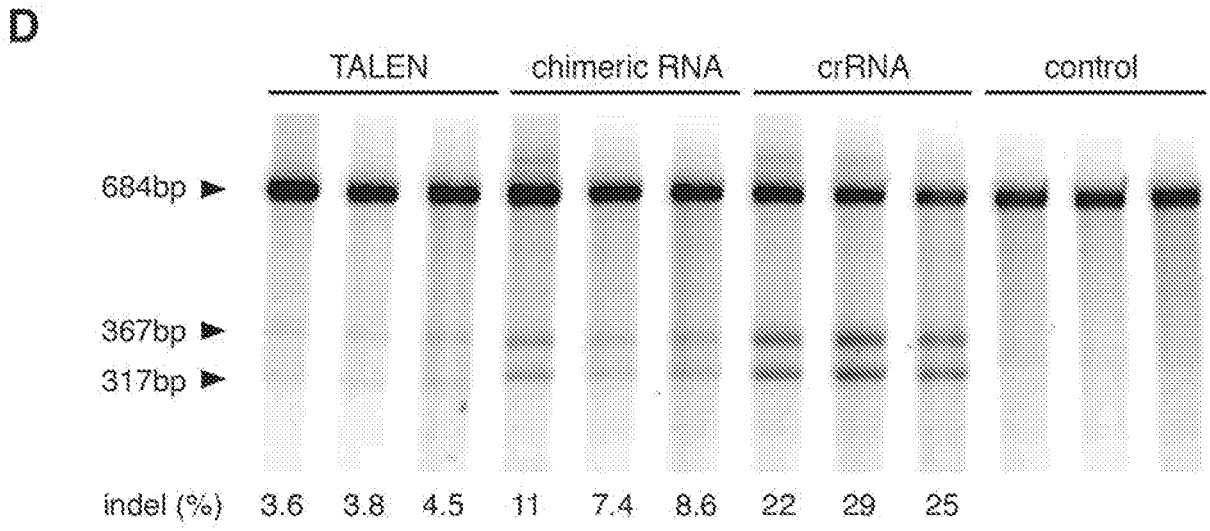

D
TALEN
chimeric RNA
crRNA
control
684bp
367bp
317bp
indel (%)
3.6
3.8
4.5
11
7.4
8.6
22
29
25

FIG. 3D

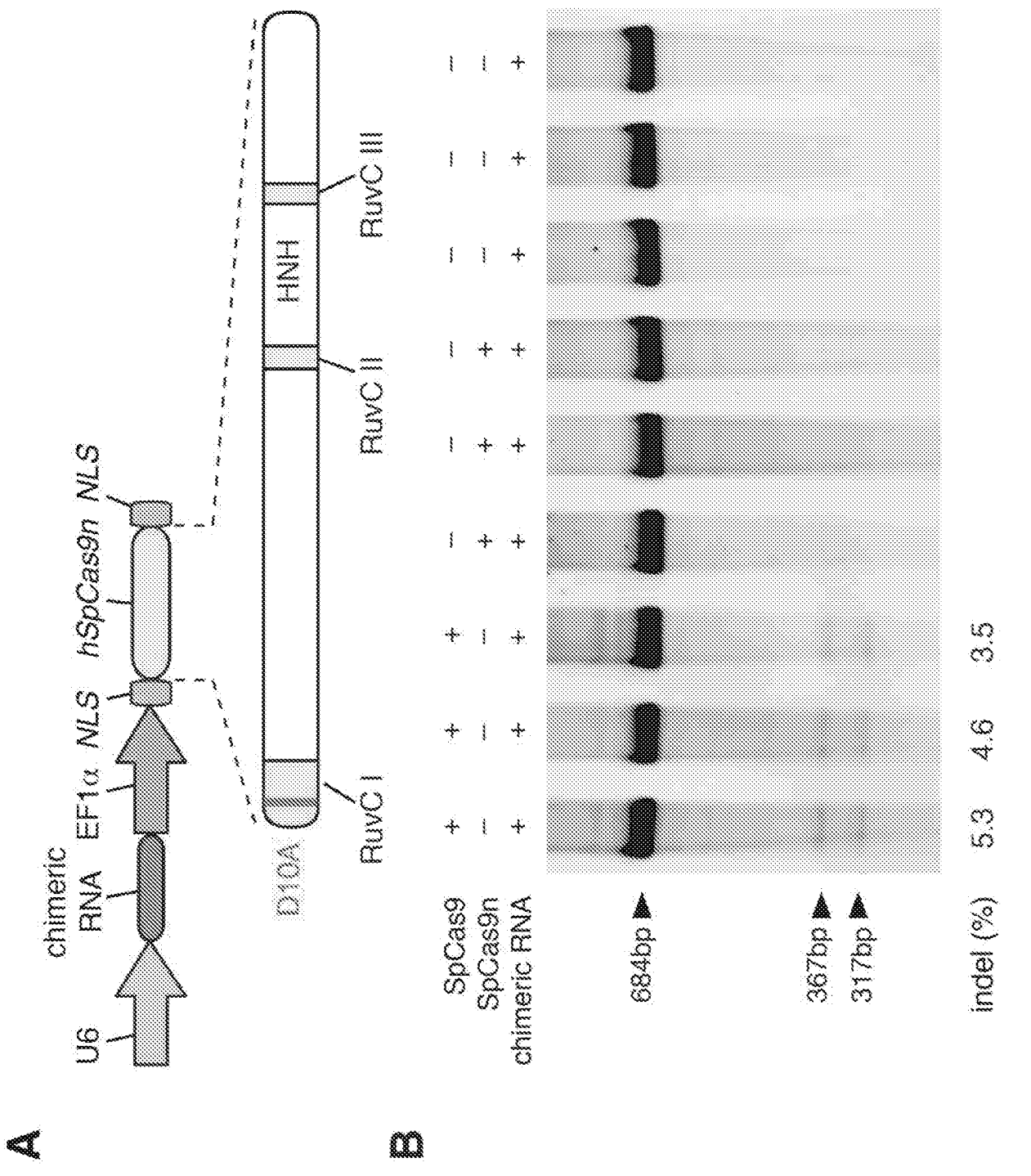

A
U6
chimeric RNA
EF1α
NLS
hSpCas9n
NLS
D10A
RuvC I
RuvC II
HNH
RuvC III
B
SpCas9
SpCas9n
chimeric RNA
684bp
367bp
317bp
indel (%)
5.3
4.6
3.5

FIG. 4A-B

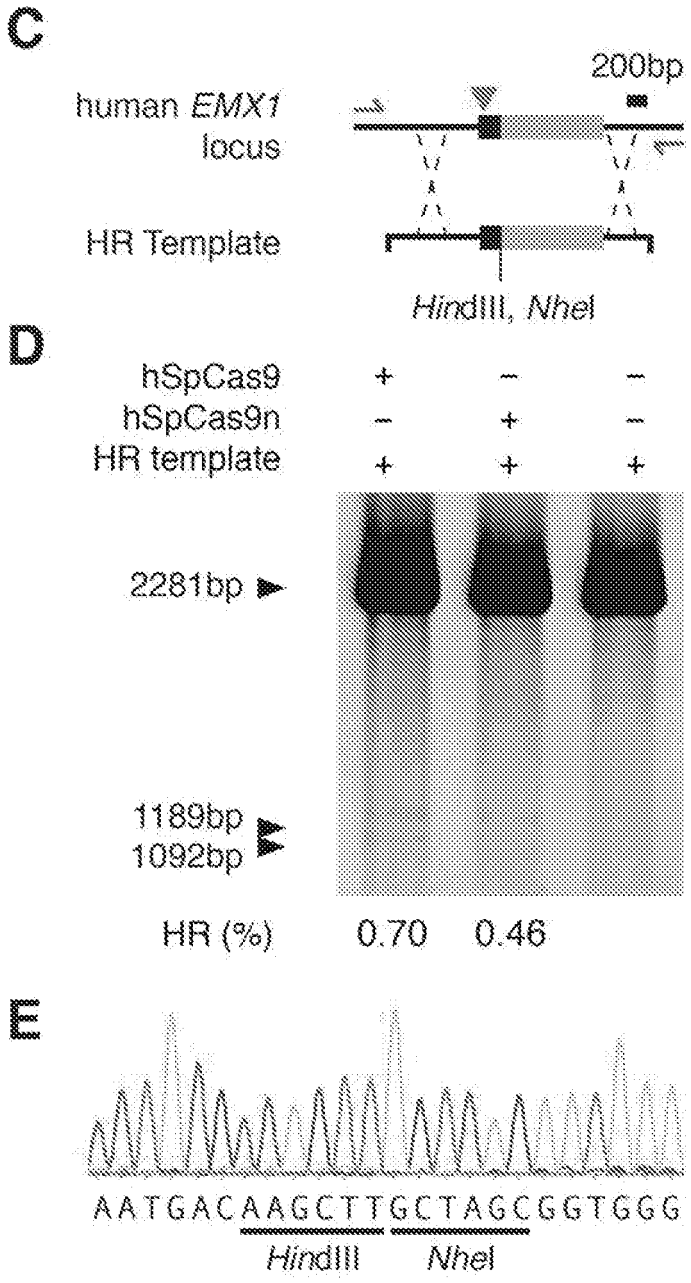

C
200bp
human EMX1 locus
HR Template
HindIII, NheI
D
hSpCas9
hSpCas9n
HR template
+ – –
– + –
+ + +
2281bp
1189bp
1092bp
HR (%)
0.70 0.46
E
AATGACAAGCTTGCTAGCGGTGGG
HindIII
NheI

FIG. 4C-E

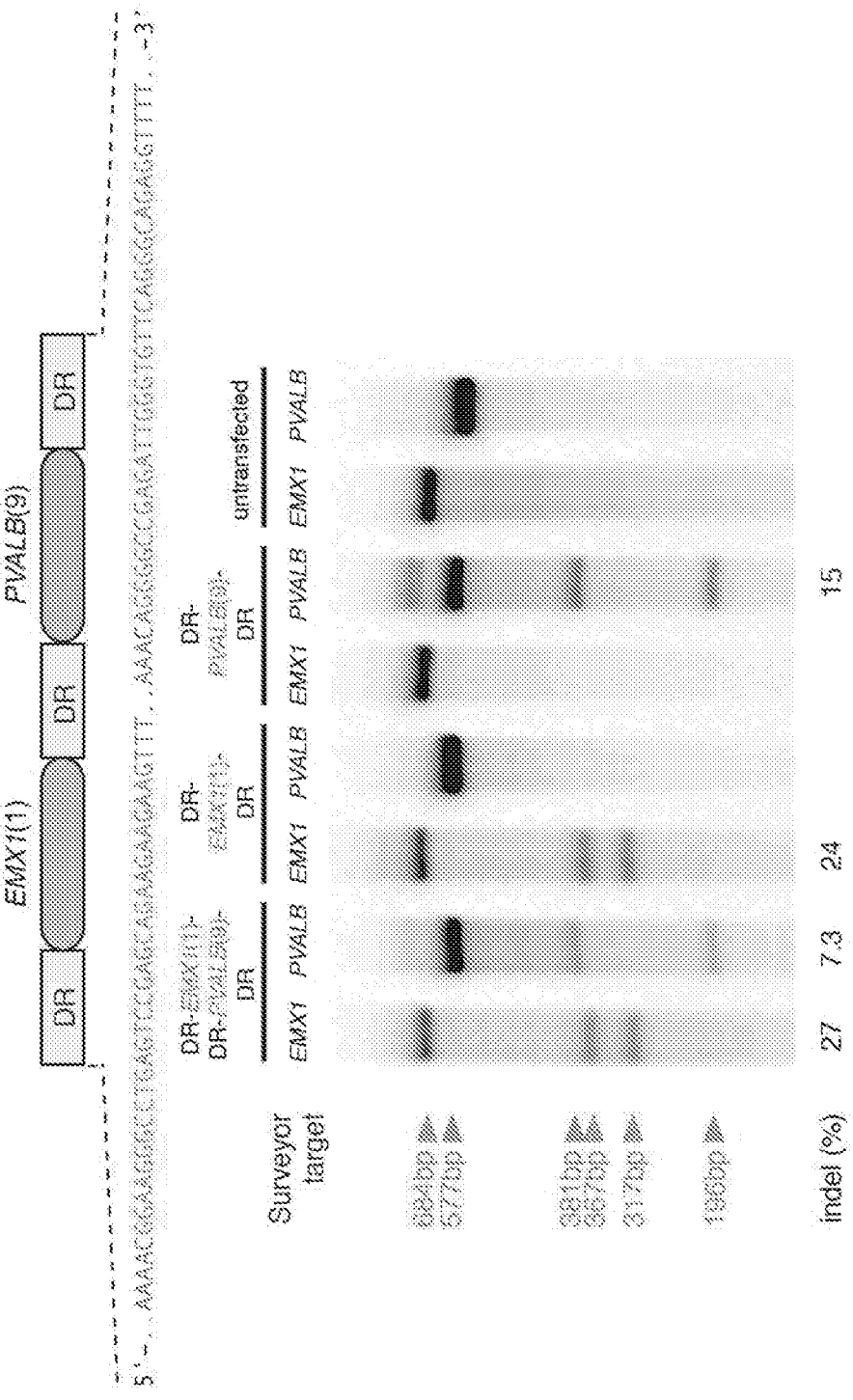

F
crRNA array design
EMX1(1)
PVALB(9)
DR
DR
DR
DR
5'-...AAAACAGAAGGGCCTGAGTCCGAGCAGAAGAAGAAGTTT...AAACAGGGGCCCGAGATTGGGTGTTCAGGGCAGAAGGTTTT...-3'
Surveyor target
DR-EMX1(1)-DR-PVALB(9)-DR
DR-EMX1(1)-DR
DR-PVALB(9)-DR
unitransfected
EMX1
PVALB
EMX1
PVALB
EMX1
PVALB
EMX1
PVALB
694bp
577bp
631bp
557bp
317bp
196bp
indel (%)
27
7.3
24
15

FIG. 4F

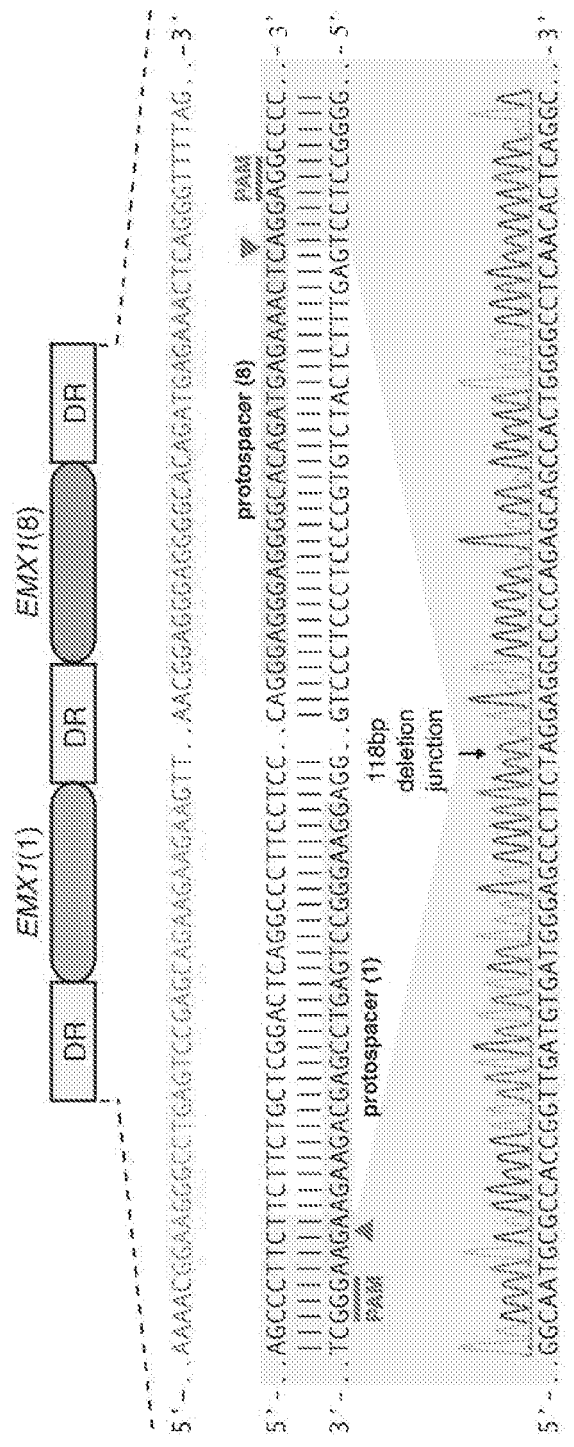

G
crRNA array design
EMX1(1)
EMX1(8)
DR
DR
DR
5' ~ . . AAAACGGAAGGGGCCTTCAGTCCGAGCAGAAGAAGAAGTT . . AACGGAGGGAGGGGCACAGATGAGAAACTCAGGGTTTTAG . . -3'
human EMX1 locus
protospacer (8)
PAM
5' - . . AGCCCTTCTTCTTCTGCTCGGACTCAGGCCCTTCCTCC . . CAGGGAGGGAGGGGCACAGATGAGAAACTCAGGAGGCCCC . . - 3'
3' - . . TCGGGAAGAAGACGAGTCCTGAGTCCGGGAAGGAGG . . GTCCCTCCCTCCCCGTGTCTACTCTTTGAGTCCTCCGGGG . . - 5'
protospacer (1)
PAM
118bp deletion junction
deletion result
5' - . . GGCAATGCGCCACCGGTTGATGTGATGGGAGCCCTTCTAGGAGGCCCCCAGAGCAGCCACTGGGGCCTCAACACTCAGGC . . - 3'

FIG. 4G

Cas9	target species	gene	protospacer ID	protospacer sequence (5' to 3')	PAM	strand	cell line tested	% indel (pre-crRNA + tracrRNA)	% indel (chimeric RNA)
S. pyogenes SF370 type II CRISPR	*Homo sapiens*	*EMX1*	1	GGAGGGCCTGAGTCCGAGCAGAAGAA	GGG	+	293FT	20 ± 1.8	6.7 ± 0.82
		EMX1	2	CATTGGAGGTGACATCGATGTCCTCCCCAT	TGG	–	293FT	2.1 ± 0.31	N.D.
		EMX1	3	GGACATCGATGTCACCTCCAATGACTAGGG	TGG	+	293FT	14 ± 1.1	N.D.
		EMX1	4	CATCGATGTCCTCCCCATTGGCCTGCTTCG	TGG	–	293FT	11 ± 1.7	N.D.
		EMX1	5	TTCGTGGCAATGCGCCACCGGTGATGTGA	TGG	–	293FT	4.3 ± 0.46	2.1 ± 0.51
		EMX1	6	TCGTGGCAATGCGCCACCGGTGATGTGAT	GGG	–	293FT	4.0 ± 0.66	0.41 ± 0.25
		EMX1	7	TCCAGCTTCTGCCGTTTGTACTTGTCCTC	CGG	–	293FT	1.5 ± 0.12	N.D.
		EMX1	8	GGAGGAGGGCACAGATGAGAAACTCAGG	AGG	–	293FT	7.8 ± 0.83	2.3 ± 1.2
	Homo sapiens	*PVALB*	9	AGGGCCCAGATTGGGTGTTCAGGCAGAG	AGG	+	293FT	21 ± 2.6	6.5 ± 0.32
		PVALB	10	ATGCAGCAGGGTGCCGAGAGGGCCGAGAT	TGG	+	293FT	N.D.	N.D.
		PVALB	11	GGTGCCGAGAGGGCCGAGATTGGGTGTTC	AGG	+	293FT	N.D.	N.D.
	Mus musculus	*Th*	12	CAAGCACTGAGTGCCATTAGCTAAATGCAT	AGG	–	Neuro2A	27 ± 4.3	4.1 ± 2.2
		Th	13	AATGCATAGGGTACCACCCACAGGTGCCAG	GGG	–	Neuro2A	4.8 ± 1.2	N.D.
		Th	14	ACACAATGGGAAAGCCTCTGGGCCAGGAA	AGG	+	Neuro2A	11.3 ± 1.3	N.D.
S. thermophilus LMD-9 CRISPR1	*Homo sapiens*	*EMX1*	15	GGAGGAGGTAGTATACAGAACACAGAGAA	GTAGAAT	–	293FT	14 ± 0.88	N.T.
		EMX1	16	AGAATGTAGAGCAGTCACAGAAACTCAGCA	CTAGAAA	–	293FT	7.8 ± 0.77	N.T.

FIG. 5

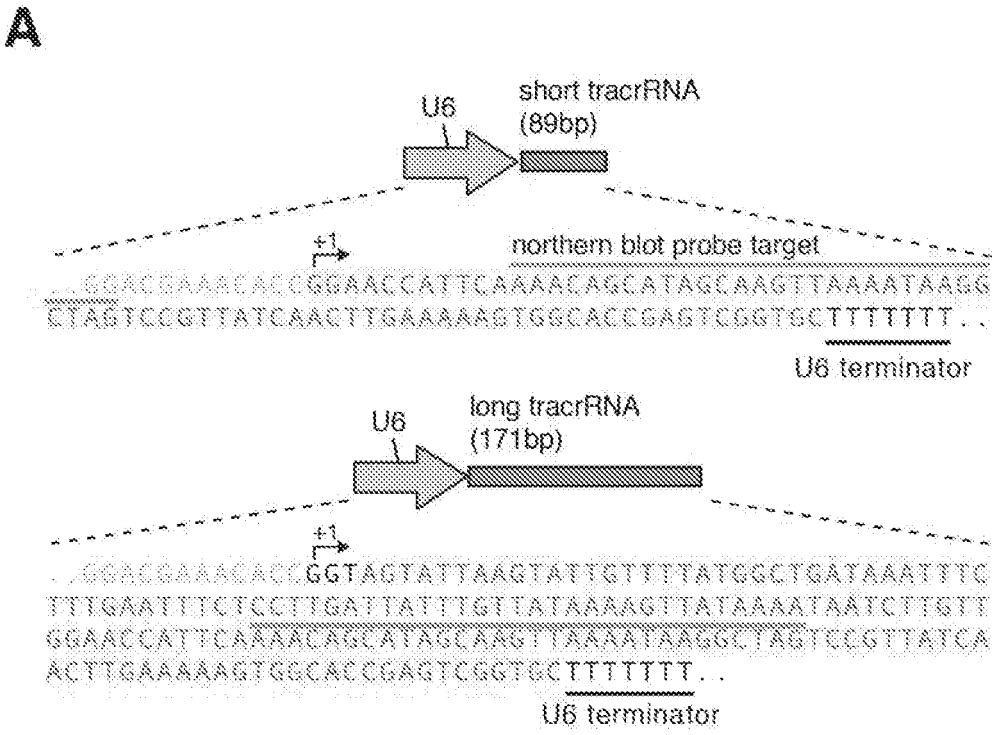

A
U6
short tracrRNA
(89bp)
+1
northern blot probe target
GGACGAAACACCGGAACCATTCAAAACAGCATAGCAAGTTAAAATAAGG
CTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTT..
U6 terminator
U6
long tracrRNA
(171bp)
+1
GGACGAAACACCGGTAGTATTAAGTATTGTTTTATGGCTGATAAATTTC
TTTGAATTTCTCCTTGATTATTTGTTATAAAAGTTATAAAATAATCTTGTT
GGAACCATTCAAAACAGCATAGCAAGTTAAAATAAGGCTAGTCCGTTATCA
ACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTTT..
U6 terminator
B
SpCas9 + + + +
long tracrRNA + + – –
short tracrRNA – – + +
SpRNase III + + + +
DR-EMX1(1)-DR + + + +
684bp
367bp
317bp
indel (%) 0.60 14 11

FIG. 6A-B

C

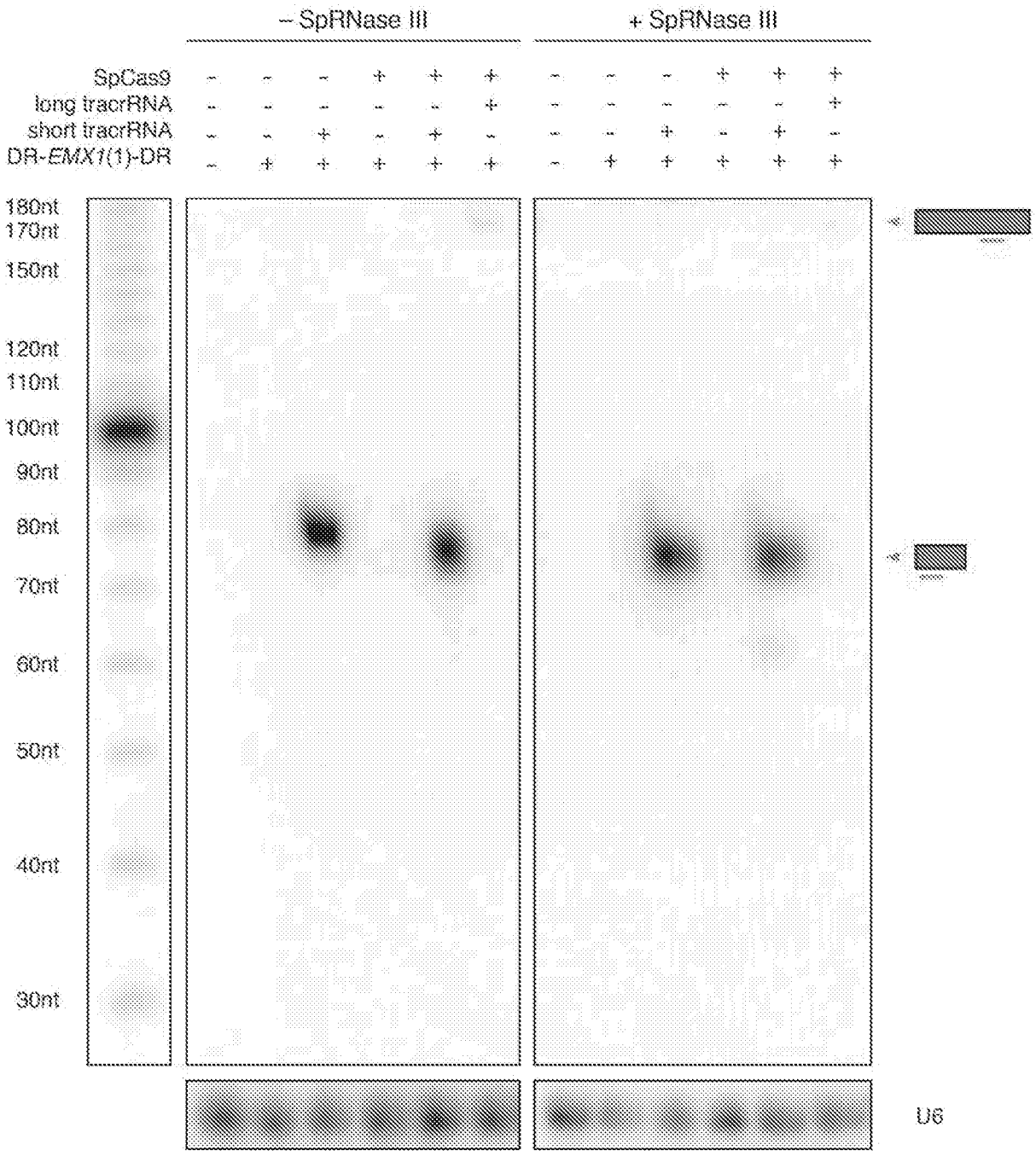

– SpRNase III
+ SpRNase III
SpCas9
long tracrRNA
short tracrRNA
DR-EMX1(1)-DR
180nt
170nt
150nt
120nt
110nt
100nt
90nt
80nt
70nt
60nt
50nt
40nt
30nt
U6

FIG. 6C

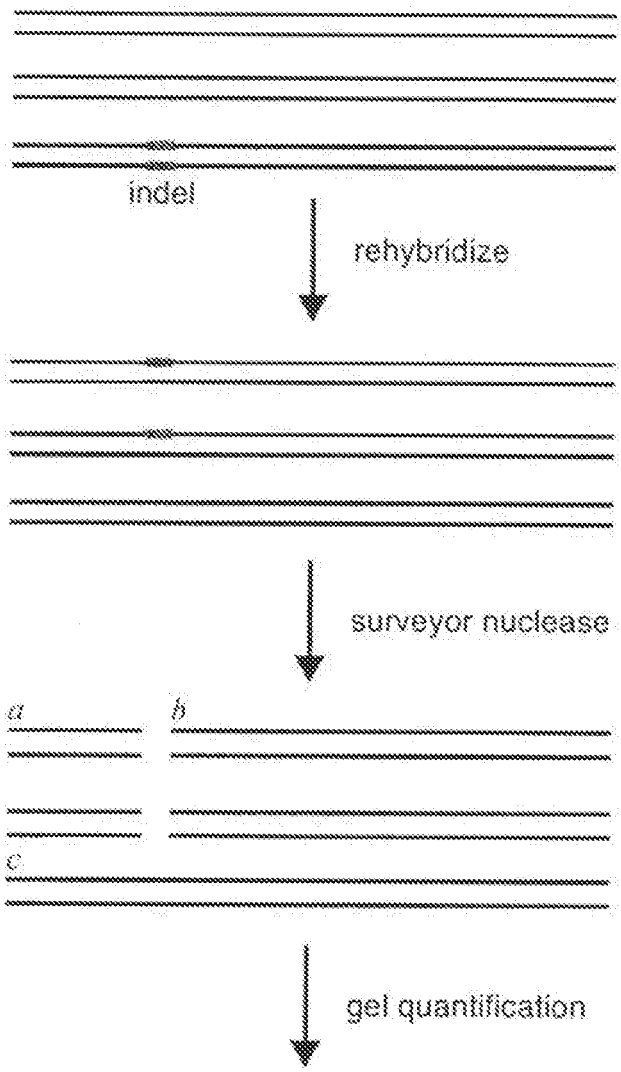

indel
rehybridize
surveyor nuclease
a
b
c
gel quantification
% indel = (1- √1 - (a + b)/(a + b + c)) *100

FIG. 7

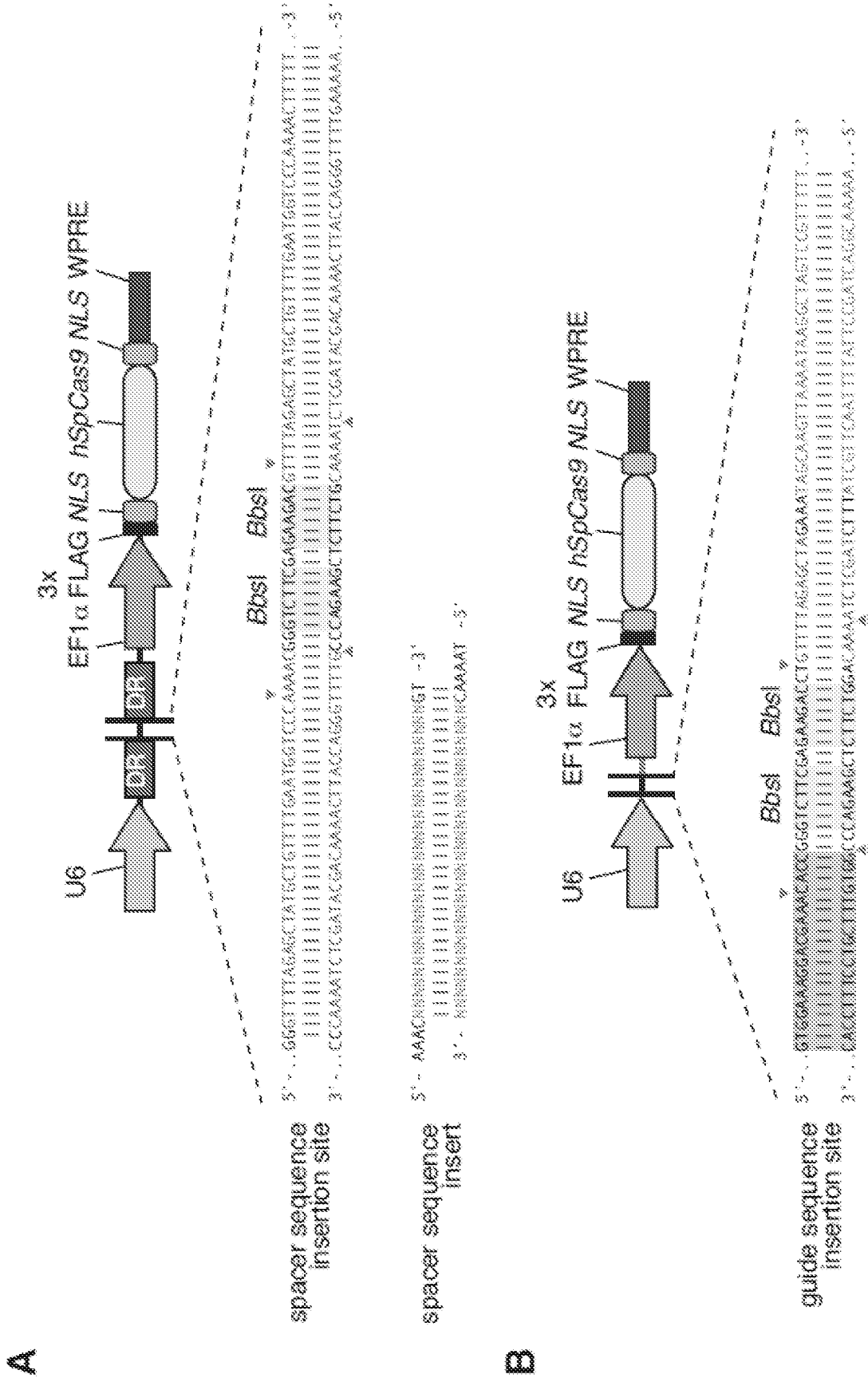
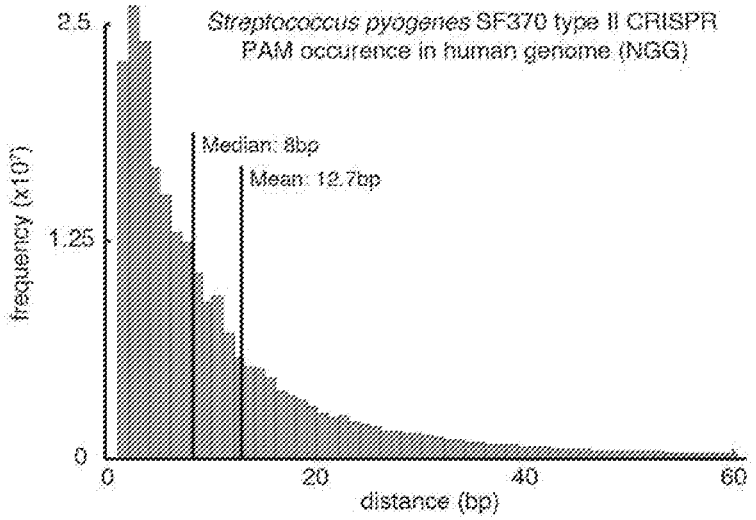
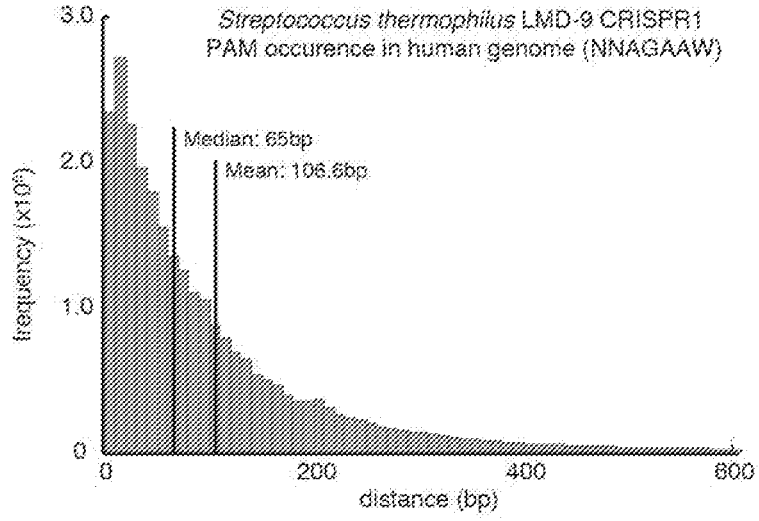

A
U6
3x
EF1α FLAG NLS hSpCas9 NLS WPRE
BbsI
BbsI
spacer sequence insertion site
spacer sequence insert
B
U6
3x
EF1α FLAG NLS hSpCas9 NLS WPRE
BbsI
BbsI
guide sequence insertion site
guide sequence insert

FIG. 8A-B

A



B



C

	NGG		NNAGAAW	
Chr	median	mean	median	mean
1	7	12.8	67	115.8
2	8	12.7	64	100.8
3	8	13.0	63	98.5
4	9	14.0	61	94.5
5	8	13.1	63	97.9
6	8	13.1	63	98.5
7	8	12.4	64	102.9
8	8	12.8	64	100.9
9	7	13.9	65	120.5
10	7	12.1	66	107.0
11	7	12.0	65	105.8
12	8	12.4	65	103.5
13	8	13.6	62	94.6
14	8	12.0	65	101.5
15	7	11.5	68	107.7
16	7	11.7	74	136.8
17	6	10.3	76	127.9
18	8	13.4	63	101.8
19	6	9.4	82	145.4
20	7	11.1	72	121.8
21	7	13.4	64	111.4
22	6	9.2	85	140.3
X	8	13.2	63	99.0
Y	8	29.2	62	223.7

FIG. 9A-C

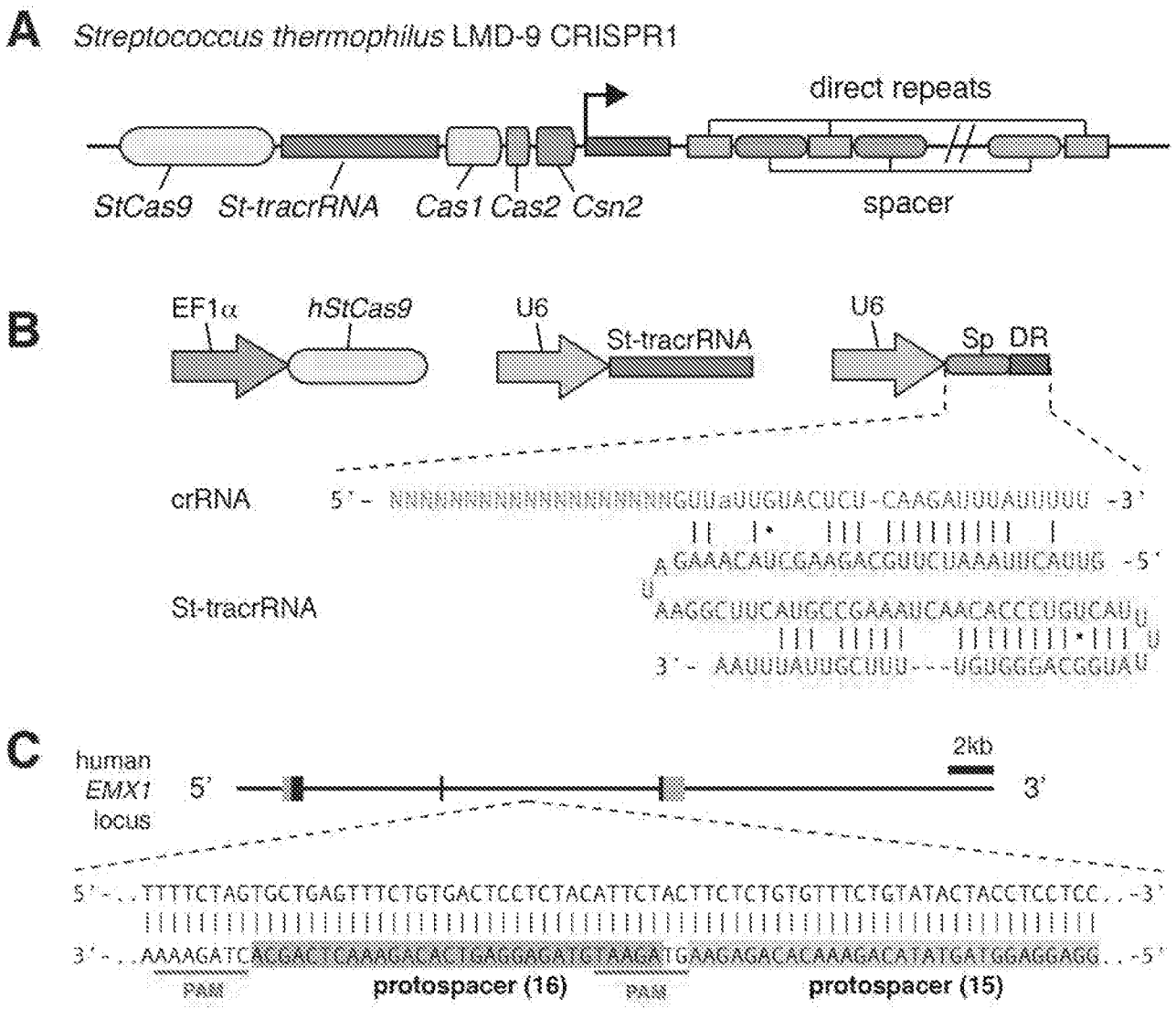

A
Streptococcus thermophilus LMD-9 CRISPR1
direct repeats
StCas9
St-tracrRNA
Cas1
Cas2
Csn2
spacer
B
EF1α
hStCas9
U6
St-tracrRNA
U6
Sp
DR
crRNA
St-tracrRNA
C
human
EMX1
locus
5'
2kb
3'
PAM
protospacer (16)
PAM
protospacer (15)

FIG. 10A-C

D

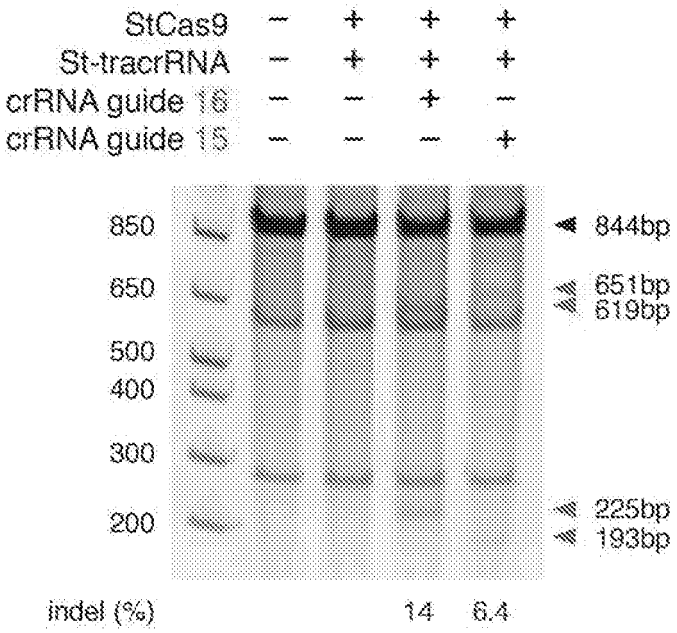

StCas9 – + + +
St-tracrRNA – + + +
crRNA guide 16 – – + –
crRNA guide 15 – – – +
850
650
500
400
300
200
844bp
651bp
619bp
225bp
193bp
indel (%)
14
6.4

FIG. 10D

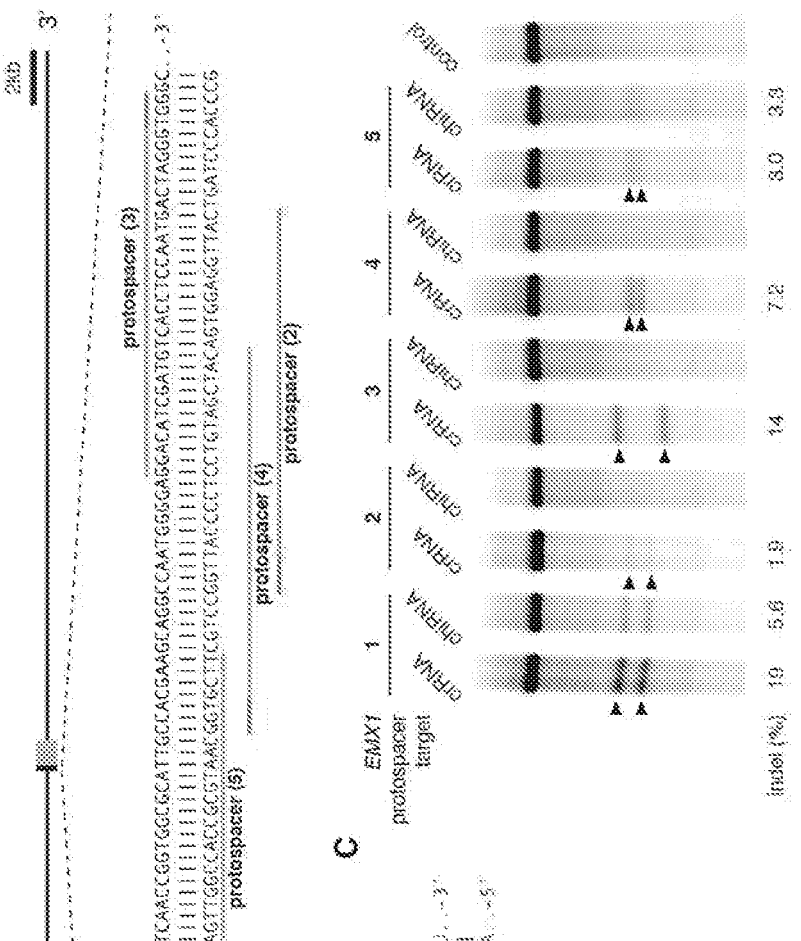

A
human EMX1 locus
5'
2kb
3'
protospacer (1)
PAM
protospacer (3)
protospacer (5)
protospacer (4)
protospacer (2)
B
pre-crRNA + tracrRNA processing
spacer (20 bp)
pre-crRNA
tracrRNA
chimeric RNA
guide sequence (20 bp)
C
EMX1 protospacer target
crRNA
chiRNA
1
2
3
4
5
control
indel (%)
19
5.6
1.9
14
7.2
3.0
3.8

FIG. 11A-C

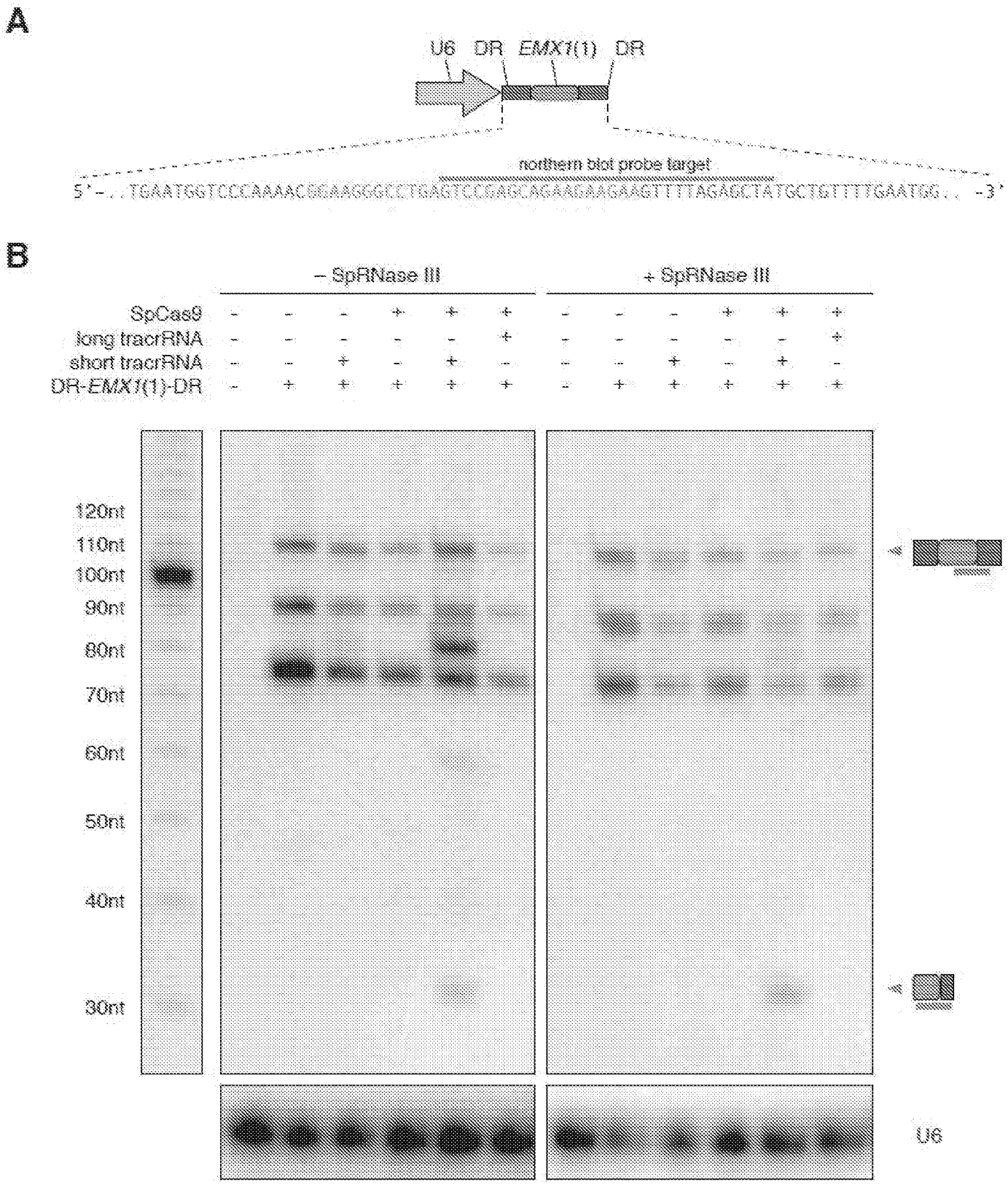

A
U6 DR EMX1(1) DR
northern blot probe target
5'-...TGAATGGTCCCAAAACGGAAGGGCCTGAGTCCGAGCAGAAGAAGAAGTTTTAGAGCTATGCTGTTTTGAATGG... -3'
B
– SpRNase III
+ SpRNase III
SpCas9
long tracrRNA
short tracrRNA
DR-EMX1(1)-DR
120nt
110nt
100nt
90nt
80nt
70nt
60nt
50nt
40nt
30nt
U6

FIG. 12A-B

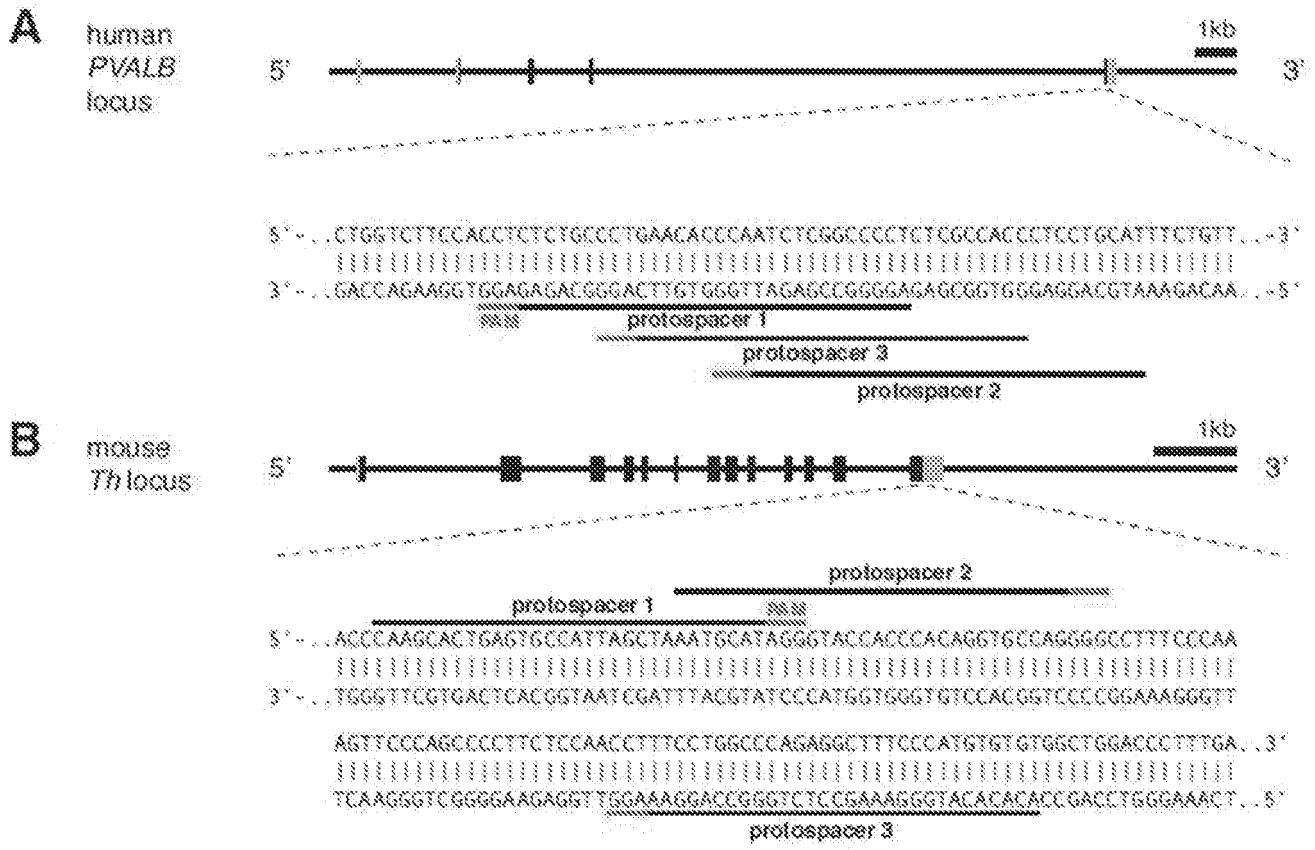

A
human
PVALB
locus
5'
1kb
3'
5'-..CTGGTCTTCCACCTCTCTGCCCTGAACACCCAATCTCGGCCCCTCTCGCCACCCTCCTGCATTTCTGTT..-3'
3'-..GACCAGAAGGTGGAGAGACGGGACTTGTGGGTTAGAGCCGGGGAGAGCGGTGGGAGGACGTAAAGACAA..-5'
protospacer 1
protospacer 3
protospacer 2
B
mouse
Th locus
5'
1kb
3'
protospacer 2
protospacer 1
5'-..ACCCAAGCACTGAGTGCCATTAGCTAAATGCATAGGGTACCACCCACAGGTGCCAGGGGCCTTTCCCAA
3'-..TGGGTTCGTGACTCACGGTAATCGATTTACGTATCCCATGGTGGGTGTCCACGGTCCCCGGAAAGGGTT
AGTTCCCAGCCCCTTCTCCAACCTTTCCTGGCCCAGAGGCTTTCCCATGTGTGTGGCTGGACCCTTTGA..3'
TCAAGGGTCGGGGGAAGAGGTTGGGAAAGGACCGGGTCTCCGAAAGGGTACACACACCGACCTGGGAAACT..5'
protospacer 3

FIG. 13A-B

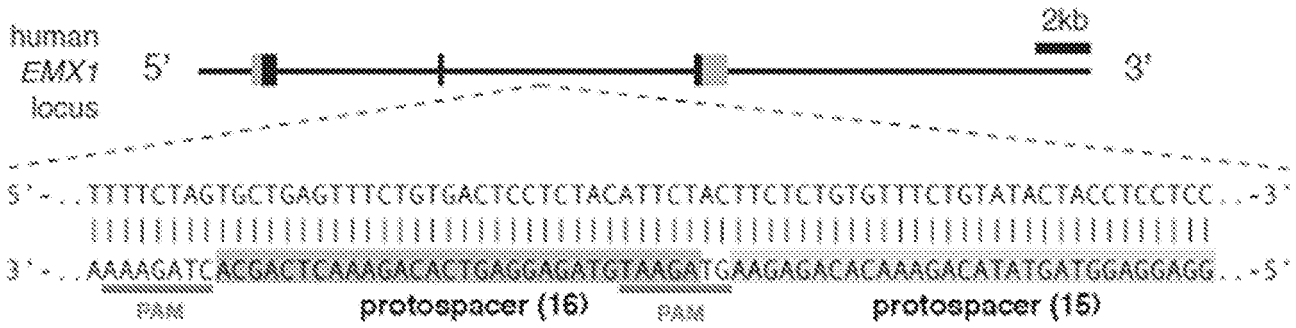

human
EMX1
locus
5'
2kb
3'
5'-..TTTTCTAGTGCTGAGTTTCTGTGACTCCTCTACATTCTACTTCTCTGTGTTTCTGTATACTACCTCCTCC..-3'
3'-..AAAAGATCACGACTCAAAGACACTGAGGAGATGTAAGATGAAGAGACACAAAGACATATGATGGAGGAGG..-5'
PAM
protospacer (16)
PAM
protospacer (15)

FIG. 14

Primer name	Assay	Genomic Target	Primer sequence
Sp-EMX1-F	SURVEYOR assay, sequencing	*EMX1*	AAAACCACCCTTCTCTCTGGC
Sp-EMX1-R	SURVEYOR assay, sequencing	*EMX1*	GGAGATTGGAGACACGGAGAG
Sp-PVALB-F	SURVEYOR assay, sequencing	*PVALB*	CTGGAAAGCCAATGCCTGAC
Sp-PVALB-R	SURVEYOR assay, sequencing	*PVALB*	GGCAGCAAACTCCTTGTCCT
Sp-Th-F	SURVEYOR assay, sequencing	*Th*	GTGCTTTGCAGAGGCCTACC
Sp-Th-R	SURVEYOR assay, sequencing	*Th*	CCTGGAGCGCATGCAGTAGT
St-EMX1-F	SURVEYOR assay, sequencing	*EMX1*	ACCTTCTGTGTTTCCACCATTC
St-EMX1-R	SURVEYOR assay, sequencing	*EMX1*	TTGGGGAGTGCACAGACTTC
Sp-EMX1-RFLP-F	RFLP, sequencing	*EMX1*	GGCTCCCTGGGTTCAAAGTA
Sp-EMX1-RFLP-R	RFLP, sequencing	*EMX1*	AGAGGGGTCTGGATGTCGTAA
Pb_EMX1_sp1	Northern Blot Probe	Not applicable	TAGCTCTAAAACTTCTTCTTCTGCTCGGAC
Pb_tracrRNA	Northern Blot Probe	Not applicable	CTAGCCTTATTTTAACTTGCTATGCTGTTT

FIG. 15


HNH endonuclease [Acidothermus cellulolyticus 11B]
hypothetical protein Amuc_2010 [Akkermansia muciniphila ATCC BAA-835]

FIG.16A



FIG. 16B

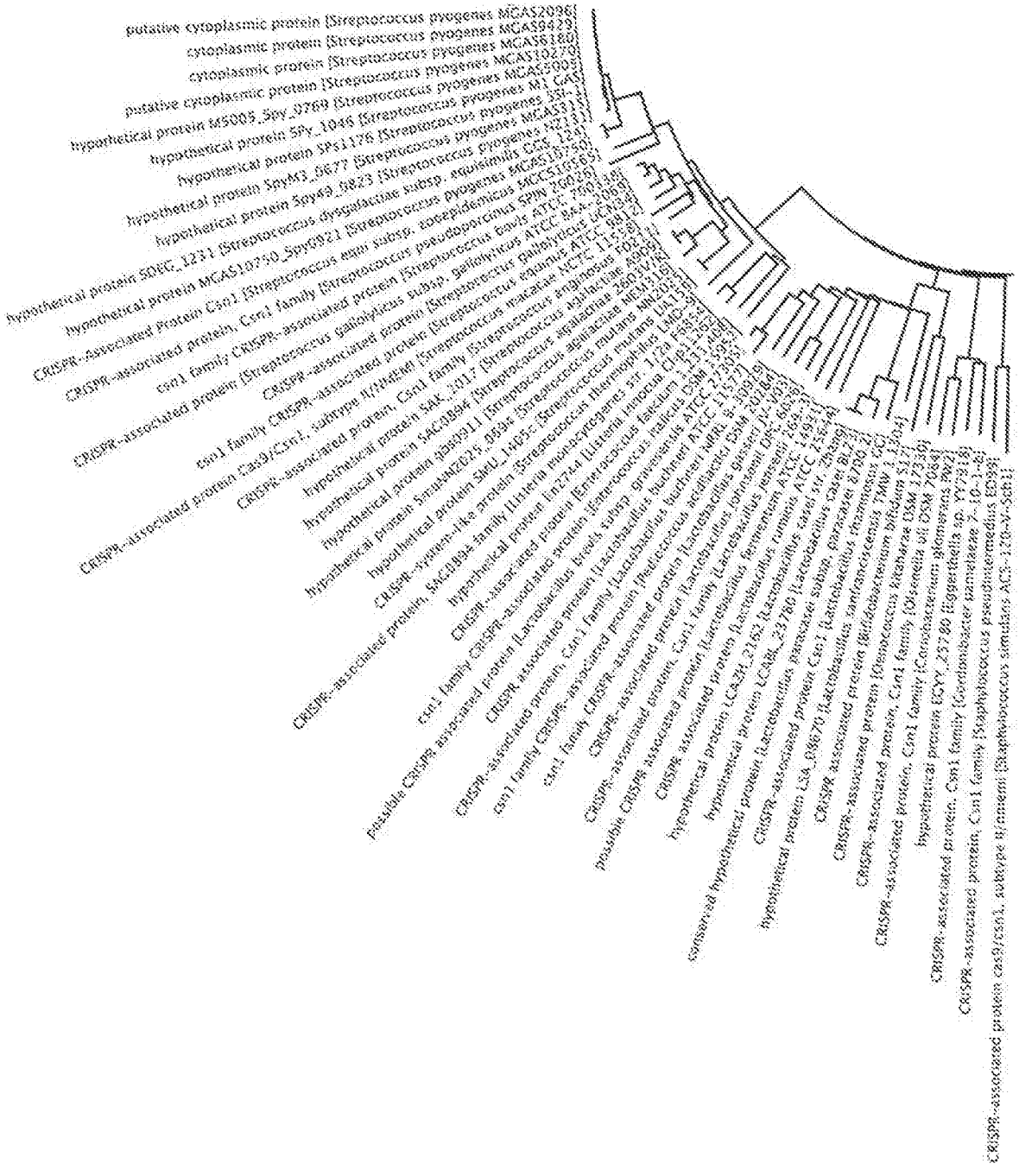

putative cytoplasmic protein [Streptococcus pyogenes MGAS2096]
cytoplasmic protein [Streptococcus pyogenes MGAS9429]
cytoplasmic protein [Streptococcus pyogenes MGAS6180]
putative cytoplasmic protein [Streptococcus pyogenes MGAS10270]
hypothetical protein MS005_Spy_0769 [Streptococcus pyogenes MGAS5005]
hypothetical protein Spy_1046 [Streptococcus pyogenes M1 GAS]
hypothetical protein SPs1176 [Streptococcus pyogenes MGAS315]
hypothetical protein SpyM3_0677 [Streptococcus pyogenes MGAS10750]
hypothetical protein Spy49_0823 [Streptococcus pyogenes MGAS10394]
hypothetical protein SDEG_1231 [Streptococcus dysgalactiae subsp. equisimilis GGS_124]
hypothetical protein MGAS10750_Spy0921 [Streptococcus pyogenes MGAS10750]
CRISPR-associated protein, Csn1 family [Streptococcus equi subsp. zooepidemicus MGCS10565]
CRISPR-associated protein, Csn1 family [Streptococcus pseudoporcinus SPIN 20026]
csn1 family CRISPR-associated protein [Streptococcus gallolyticus subsp. gallolyticus ATCC BAA-2069]
CRISPR-associated protein Cas9/Csn1, subtype II/NMEN [Streptococcus equinus ATCC 9812]
CRISPR-associated protein [Streptococcus macacae NCTC 11558]
hypothetical protein SAK_2017 [Streptococcus agalactiae A909]
hypothetical protein gbs0911 [Streptococcus agalactiae 2603V/R]
CRISPR-system-like protein Smu1462c [Streptococcus agalactiae NEM316]
hypothetical protein SAG0896 [Streptococcus agalactiae 18RS21]
CRISPR-associated protein [Streptococcus mutans NN2025]
csn1 family CRISPR-associated protein [Streptococcus thermophilus LMD-9]
CRISPR-associated protein [Streptococcus sp. 1_1_36]
possible CRISPR associated protein [Streptococcus salivarius JIM8777]
CRISPR-associated protein [Lactobacillus brevis subsp. gravesensis ATCC 27305]
csn1 family CRISPR-associated protein [Lactobacillus buchneri NRRL B-30929]
CRISPR-associated protein, Csn1 family [Lactobacillus buchneri ATCC 11577]
CRISPR-associated protein [Pediococcus acidilactici DSM 20284]
possible CRISPR associated protein [Lactobacillus gasseri JV-V03]
CRISPR associated protein, Csn1 family [Lactobacillus hamsteri DSM 5661]
hypothetical protein LCABL_23760 [Lactobacillus rhamnosus LRHMDP2]
hypothetical protein LCA_23760 [Lactobacillus rhamnosus LRHMDP3]
conserved hypothetical protein [Lactobacillus fermentum ATCC 14931]
hypothetical protein LSA_08670 [Lactobacillus salivarius CECT 5713]
CRISPR-associated protein [Lactobacillus ruminis ATCC 25644]
CRISPR-associated protein, Csn1 family [Lactobacillus sakei subsp. sakei 23K]
CRISPR-associated protein [Lactobacillus casei BL23]
CRISPR-associated protein [Lactobacillus rhamnosus GG]
CRISPR-associated protein [Lactobacillus zeae KCTC 3804]
CRISPR-associated protein [Bifidobacterium bifidum S17]
CRISPR-associated protein, Csn1 family [Olsenella uli DSM 7084]
hypothetical protein [Coriobacterium glomerans PW2]
CRISPR-associated protein, Csn1 family [Gordonibacter pamelaeae 7-10-1-b]
CRISPR-associated protein, Csn1 family [Staphylococcus pseudintermedius ED99]
CRISPR-associated protein cas9/csn1, subtype II/nmeni [Staphylococcus simulans ACS-120-V-Sch1]

FIG. 16C

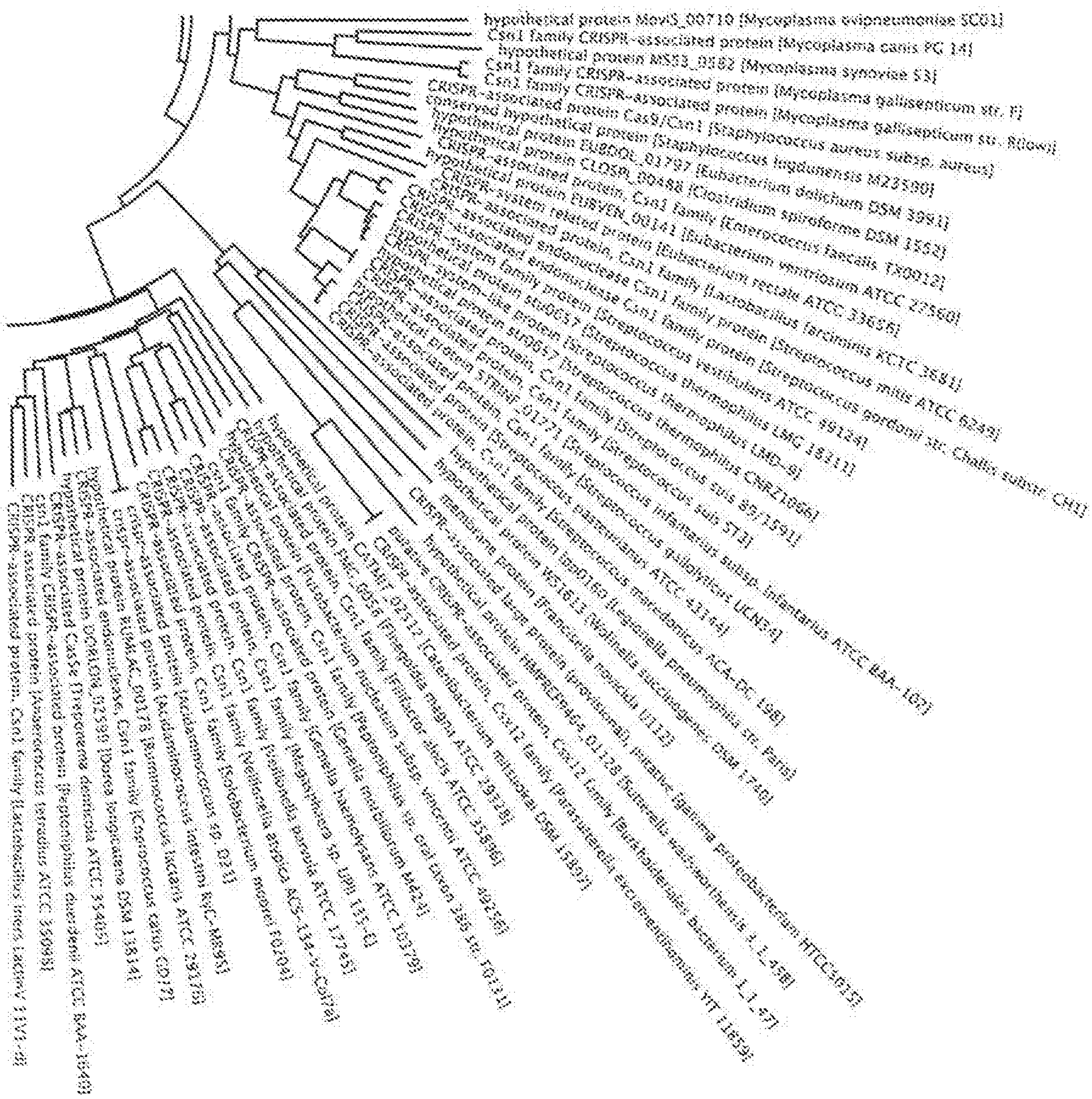

hypothetical protein MoviS_00710 [Mycoplasma ovipneumoniae SC01]
Csn1 family CRISPR-associated protein [Mycoplasma canis PG 14]
hypothetical protein MS53_0582 [Mycoplasma synoviae 53]
Csn1 family CRISPR-associated protein [Mycoplasma gallisepticum str. F]
CRISPR-associated protein [Mycoplasma gallisepticum str. R(low)]
conserved hypothetical protein Cas9/Csn1 [Staphylococcus aureus subsp. aureus]

FIG. 16D



FIG.17A



FIG. 17B

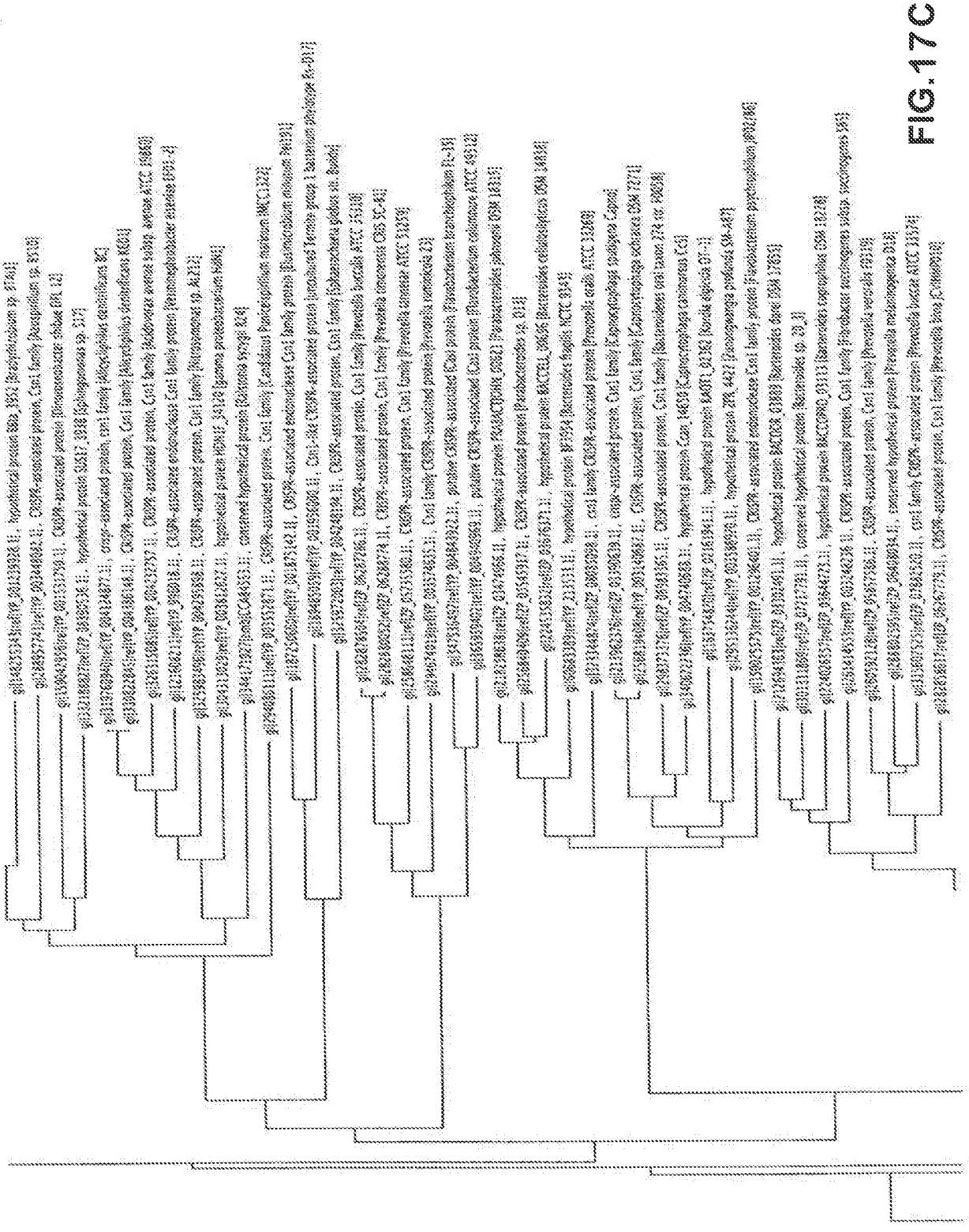


FIG.17C

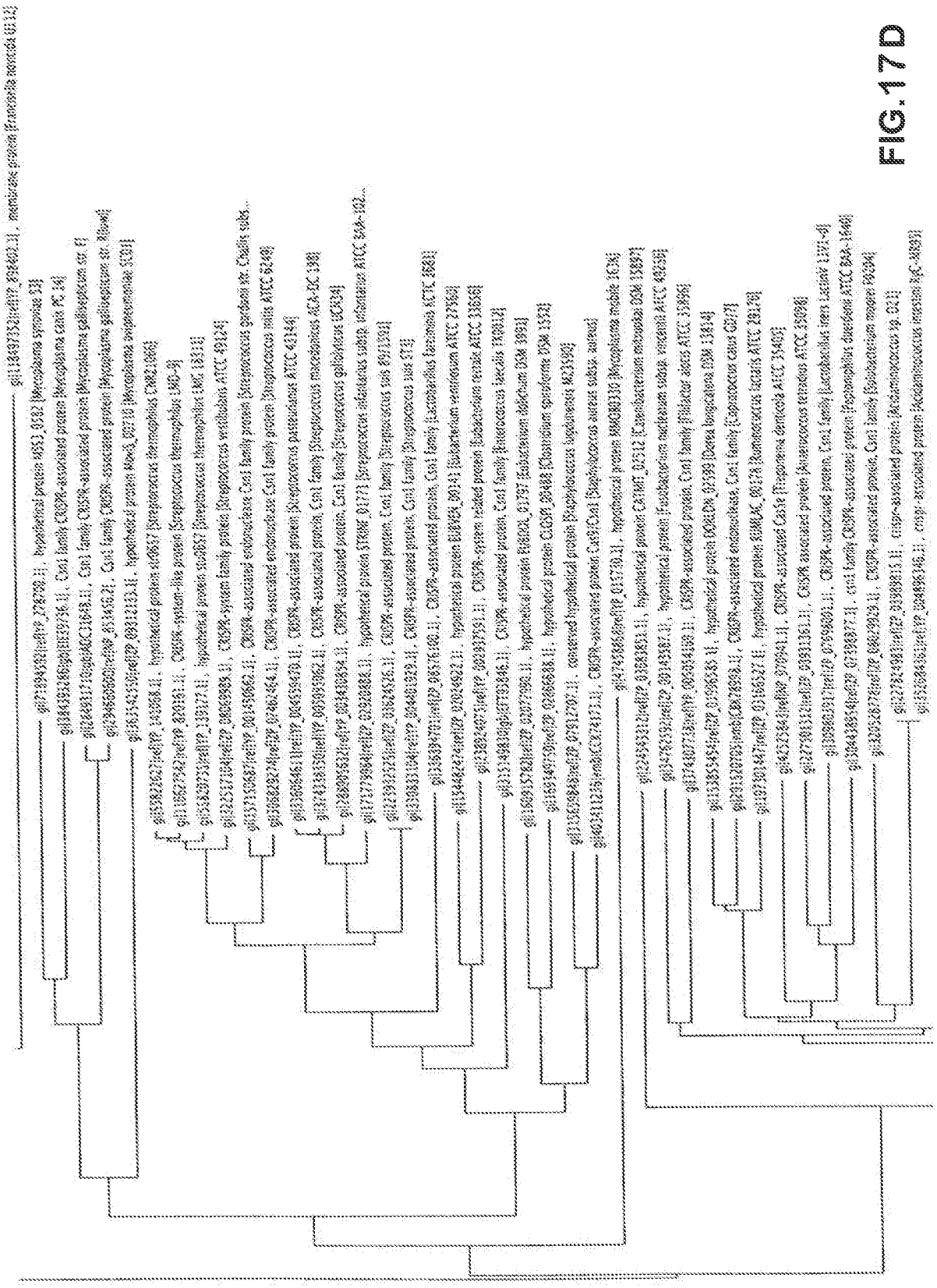


FIG.17D

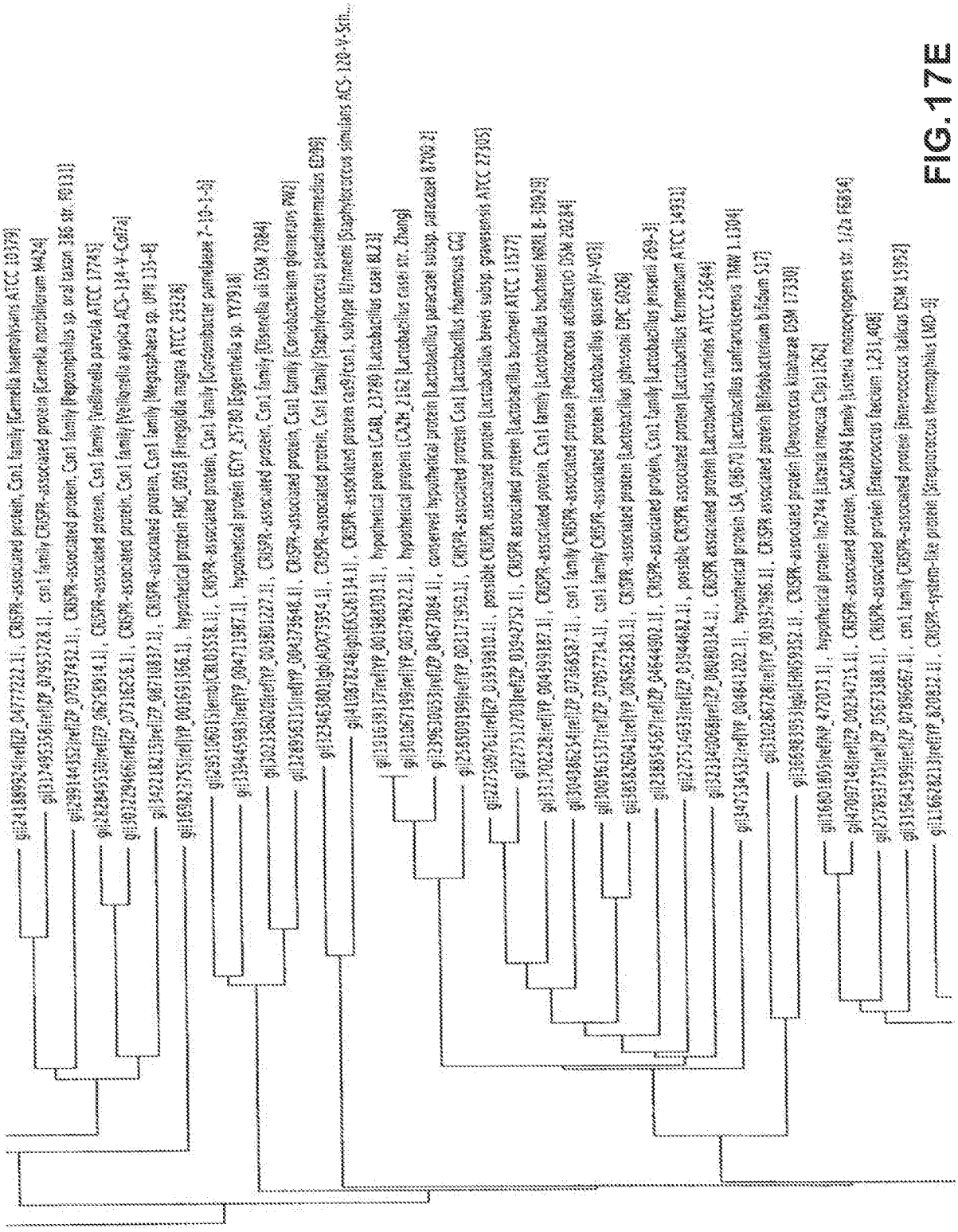

gi|24188892|ref|NP_047722.1|, CRISPR-associated protein, Csn1 family [Escherichia haemolytica ATCC 10379]
gi|31749538|ref|NP_855728.1|, csn1 family CRISPR-associated protein [Escherichia marbiliorum M424]
gi|29914432|ref|NP_070372.1|, CRISPR-associated protein, Csn1 family [Pepromphilus sp. oral taxon 386 str. F0131]
gi|28284933|ref|NP_782849.1|, CRISPR-associated protein, Csn1 family [Halanaella parvula ATCC 17745]
gi|30122046|ref|NP_078256.1|, CRISPR-associated protein, Csn1 family [Halanaella angusta ACS-134-V-Col2a]
gi|34221818|ref|NP_087188.1|, CRISPR-associated protein, Csn1 family [Mesopshaera sp. UPII 135-E]
gi|16082275|ref|NP_001611366.1|, hypothetical protein PAE_0038 [Pyrogalda magna ATCC 29328]
gi|28211803|emb|CAD63558.1|, CRISPR-associated protein, Csn1 family [Corynebacter parvulum 2-10-1-4]
gi|13944598|ref|NP_004731987.1|, hypothetical protein ECTY_28780 [Cyperthela sp. YYT918]
gi|30213602|ref|NP_803051227.1|, CRISPR-associated protein, Csn1 family [Shewella oli OSK 7084]
gi|11295611|ref|NP_004275646.1|, CRISPR-associated protein, Csn1 family [Corynebacterium glutamicum PW1]
gi|12345680|gb|AAK71954.1|, CRISPR-associated protein, Csn1 family [Staphylococcus pseudintermedius ED99]
gi|41037824|dbj|BAD32134.1|, CRISPR-associated protein csn1/Csn1, subtype Ilhaensis [Staphylococcus simulans ACS-120-V-Sch...
gi|15913913|ref|NP_360868203.1|, hypothetical protein CA0_21760 [Lactobacillus casei BL23]
gi|30136719|ref|NP_883769222.1|, hypothetical protein CA2h_2162 [Lactobacillus casei str. Zhang]
gi|23963033|ref|NP_84671084.1|, conserved hypothetical protein [Lactobacillus paracasei subsp. paracasei 8700:2]
gi|25450919|ref|NP_83171959.1|, CRISPR-associated protein Csn1 [Lactobacillus rhamnosus GG]
gi|27350976|ref|NP_231910410.1|, possible CRISPR associated protein [Lactobacillus brevis subsp. gravesensis ATCC 27305]
gi|27351703|ref|NP_81942752.1|, CRISPR associated protein [Lactobacillus buchneri ATCC 11577]
gi|31702228|ref|NP_804393187.1|, CRISPR-associated protein, Csn1 family [Lactobacillus buchneri NRRL B-30929]
gi|30406624|ref|NP_97366532.1|, csn1 family CRISPR-associated protein [Pediococcus acidilactici DSM 20284]
gi|30361533|ref|NP_87877714.1|, csn1 family CRISPR-associated protein [Lactobacillus gasseri JV-V03]
gi|30362660|ref|NP_005662383.1|, CRISPR-associated protein [Lactobacillus johnsonii DPC 6026]
gi|23861457|ref|NP_84644802.1|, CRISPR-associated protein, Csn1 family [Lactobacillus jensenii 269-3]
gi|22751463|ref|NP_81944882.1|, possible CRISPR associated protein [Lactobacillus fermentum ATCC 14931]
gi|31114206|ref|NP_80880134.1|, CRISPR associated protein [Lactobacillus ruminis ATCC 25644]
gi|34221345|ref|NP_84641282.1|, hypothetical protein LSA_08670 [Lactobacillus sakei subsp. sakei 23K]
gi|31062678|ref|NP_80192386.1|, CRISPR associated protein [Bifidobacterium bifidum S17]
gi|30660855|gb|AAP20932.1|, CRISPR-associated protein [Pediococcus kimchicus DSM 17338]
gi|16801695|ref|NP_472071.1|, hypothetical protein lin2744 [Listeria innocua Clip11262]
gi|47097218|ref|ZP_00234715.1|, CRISPR-associated protein, SAG0894 family [Listeria monocytogenes str. 1/2a F6854]
gi|29376317|ref|NP_815571388.1|, CRISPR-associated protein [Enterococcus faecium 1,231,408]
gi|31316419|ref|NP_87866667.1|, csn1 family CRISPR-associated protein [Enterococcus italicus DSM 15952]
gi|31166782|ref|NP_820022.1|, CRISPR-system-like protein [Streptococcus thermophilus LMD-9]

FIG.17E

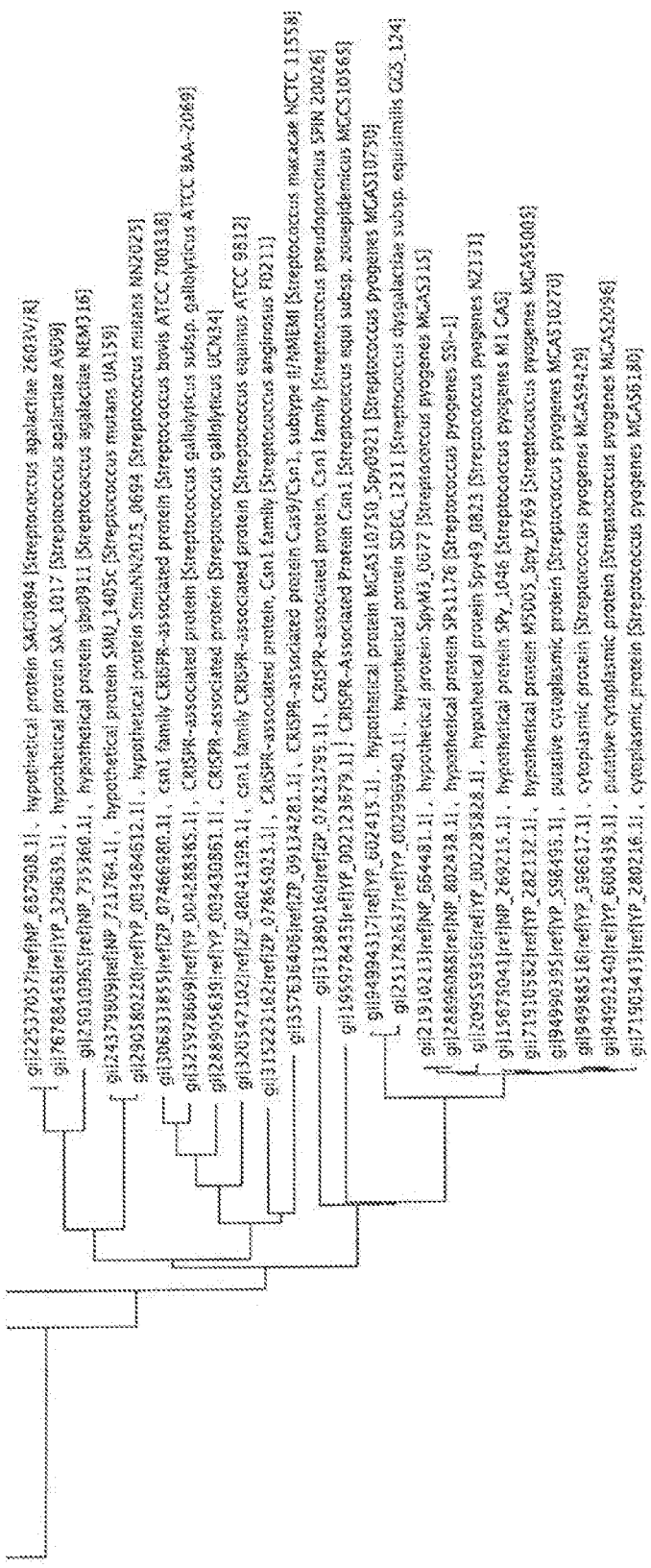

gi|22537057|ref|NP_687908.1| hypothetical protein SAG0894 [Streptococcus agalactiae 2603V/R]
gi|76788438|ref|YP_328639.1| hypothetical protein SAK_1017 [Streptococcus agalactiae A909]
gi|23101955|ref|NP_735960.1| hypothetical protein gbs0911 [Streptococcus agalactiae NEM316]
gi|24379809|ref|NP_721764.1| hypothetical protein SMU_1405c [Streptococcus mutans UA159]
gi|29380228|ref|YP_003484932.1| hypothetical protein SmuNN2025_0694 [Streptococcus mutans NN2025]
gi|30683318|ref|ZP_07466880.1| Cas1 family CRISPR-associated protein [Streptococcus bovis ATCC 700338]
gi|32379866|ref|YP_004288365.1| CRISPR-associated protein [Streptococcus gallolyticus subsp. gallolyticus ATCC BAA-2069]
gi|28895639|ref|YP_003430861.1| CRISPR-associated protein [Streptococcus gallolyticus UCN34]
gi|32634102|ref|ZP_08041998.1| Cas1 family CRISPR-associated protein [Streptococcus equinus ATCC 9812]
gi|31322162|ref|ZP_07865023.1| CRISPR-associated protein, Cas1 family [Streptococcus anginosus F0211]
gi|33763640|ref|ZP_09134281.1| CRISPR-associated protein Cas9/Csn1, subtype II/NMEN [Streptococcus mascae NCTC 11558]
gi|31389016|ref|ZP_07823795.1| CRISPR-associated protein, Cas1 family [Streptococcus pseudoporcinus SPIN 20026]
gi|19559784|ref|YP_002123679.1| CRISPR-Associated Protein Cas1 [Streptococcus equi subsp. zooepidemicus MGCS10565]
gi|49904317|ref|YP_602419.1| hypothetical protein MGAS10750_Spy0921 [Streptococcus pyogenes MGAS10750]
gi|21178637|ref|YP_002996940.1| hypothetical protein SDEG_1231 [Streptococcus dysgalactiae subsp. equisimilis GGS_124]
gi|21910213|ref|NP_664481.1| hypothetical protein SpyM3_0677 [Streptococcus pyogenes MGAS315]
gi|28896088|ref|NP_802828.1| hypothetical protein SPs1176 [Streptococcus pyogenes SSI-1]
gi|13095593|ref|YP_002285828.1| hypothetical protein Spy49_0833 [Streptococcus pyogenes NZ131]
gi|15675041|ref|NP_269216.1| hypothetical protein SPy_1046 [Streptococcus pyogenes M1 GAS]
gi|71910582|ref|YP_282132.1| hypothetical protein MS005_Spy_0769 [Streptococcus pyogenes MGAS5005]
gi|94990395|ref|YP_598493.1| putative cytoplasmic protein [Streptococcus pyogenes MGAS10270]
gi|94988516|ref|YP_596617.1| cytoplasmic protein [Streptococcus pyogenes MGAS9429]
gi|94992340|ref|YP_600439.1| putative cytoplasmic protein [Streptococcus pyogenes MGAS2096]
gi|71906341|ref|YP_280216.1| cytoplasmic protein [Streptococcus pyogenes MGAS315]

FIG.17F

a

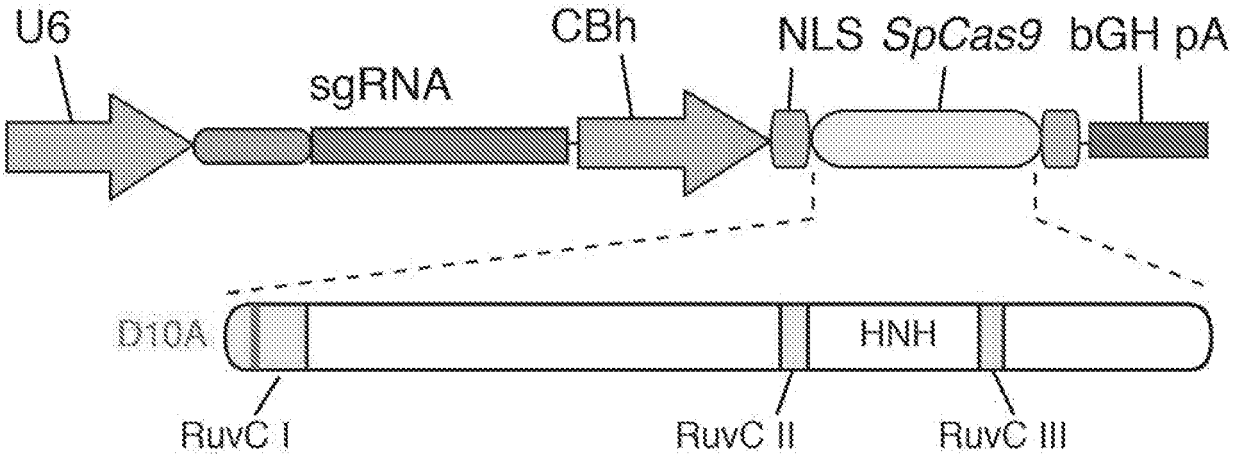


FIG. 18A

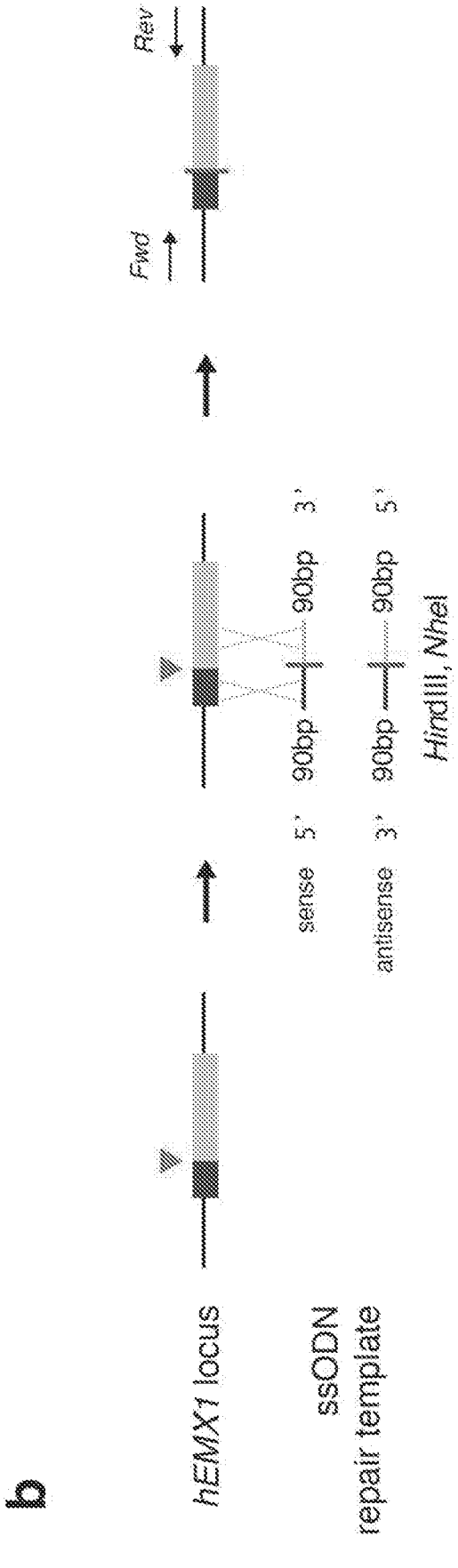

b
hEMX1 locus
ssODN
repair template
Fwd
Rev
sense
5'
90bp
90bp
3'
antisense
3'
90bp
90bp
5'
HindIII, NheI

FIG. 18B

C
hEMX1 locus
genomic target
PAM
5'-..CAGAAGAAGAAGGGC...CCAATGGGAGGACATCGATGTCACCTCCAATGACTAGGGTGGTGGGCAAC...CTCTGGCCACTCCCT..-3'
3'-..GTCTTCTTCTTCCCG...GGTTACCCCCTCCTGTAGCTACAGTGGAGGTTACTGATCCCACCACCCGTTG...GAGACCGGGTGAGGGA..-5'
sense 5'- CAGAAGAAGAAGGGC...ACATCGATGTCACCTCCAATGACAAGCTTGCTAGCGGTGGGCAACCACAAAC...CTCTGGCCACTCCCT -3'
HindIII NheI
antisense 3'- GTCTTCTTCTTCCCG...TGTAGCTACAGTGGAGGTTACTGTTCGAACGATCGCCACCCGTTGGTGTTTG...GAGACCGGGTGAGGGA -5'
HindIII NheI

FIG. 18C

d

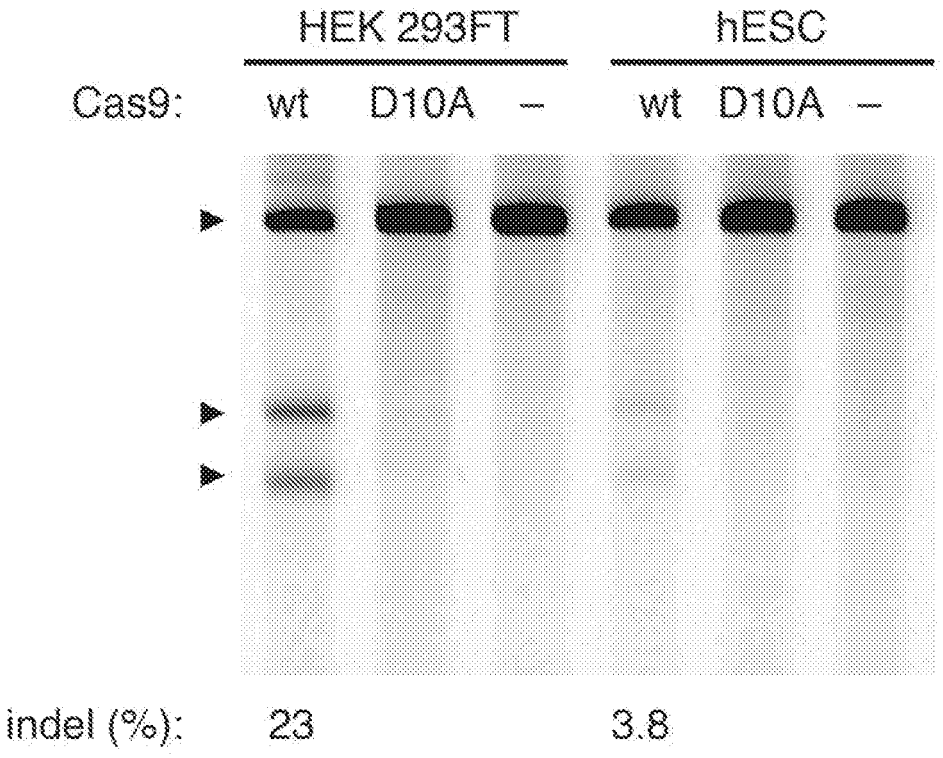


FIG. 18D

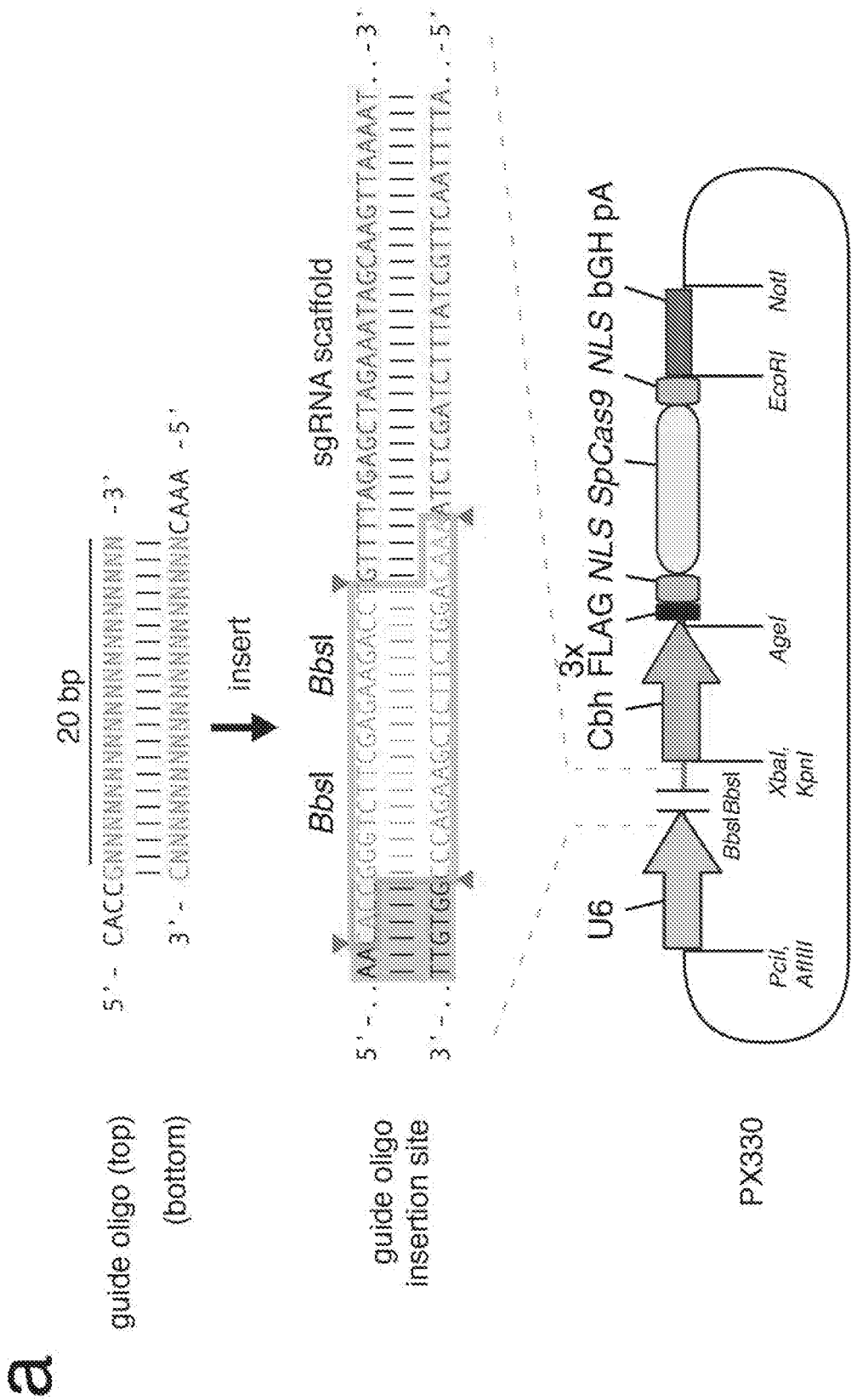

a
guide oligo (top)
(bottom)
20 bp
5' - CACCGNNNNNNNNNNNNNNNNNNN - 3'
3' - CNNNNNNNNNNNNNNNNNNNCAAA - 5'
insert
BbsI
BbsI
sgRNA scaffold
guide oligo insertion site
5' - . . AACNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAAGTTAAAAT . . - 3'
3' - . . TTGTGCNNNNNNNNNNNNNNNNNNNCAAAATCTCGATCTTTATCGTTCAATTTTA . . - 5'
PX330
U6
Cbh
3x FLAG NLS SpCas9 NLS bGH pA
PciI, AflIII
BbsI BbsI
XbaI, KpnI
AgeI
EcoRI
NotI

FIG. 19A

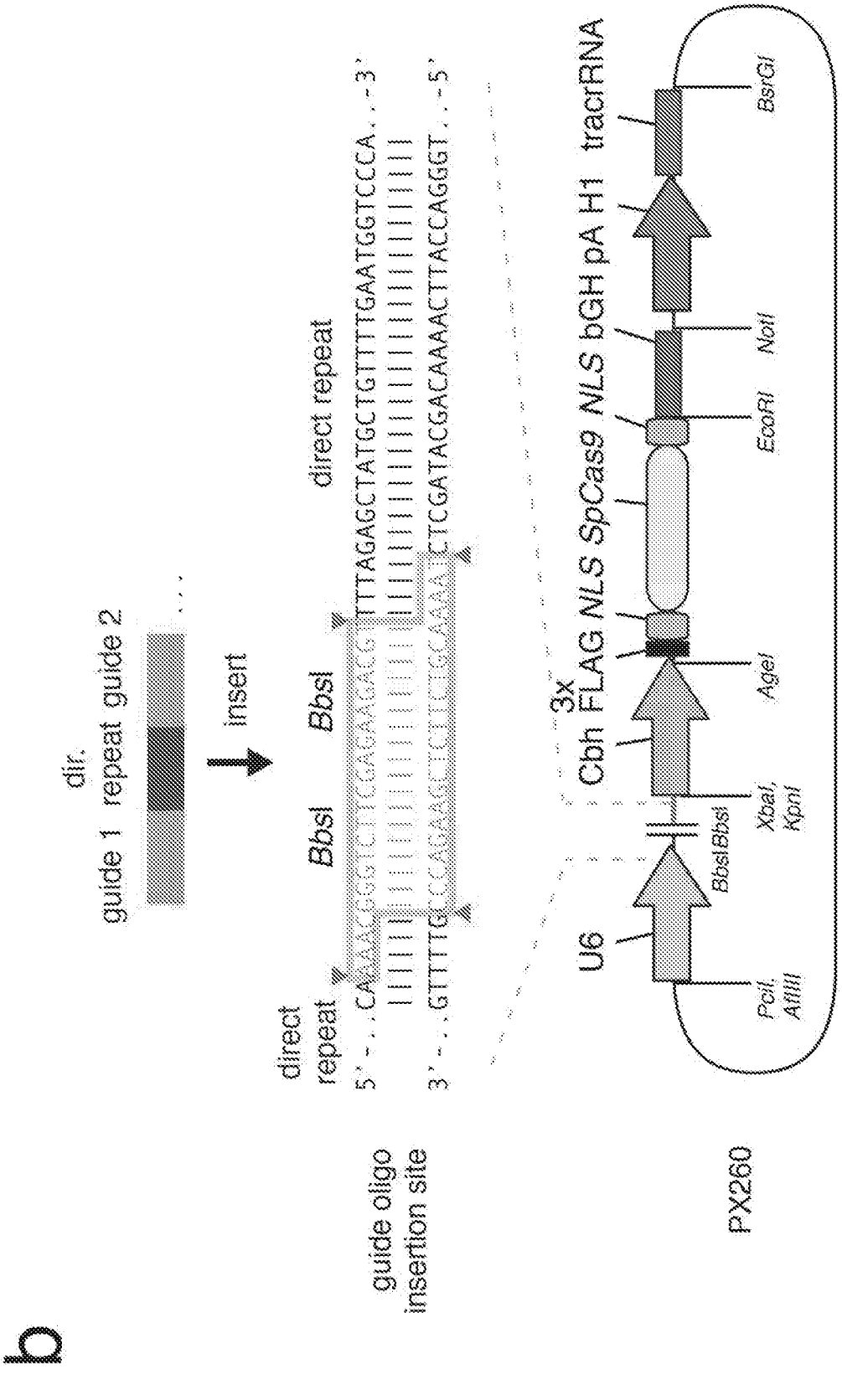

b
dir.
guide 1 repeat guide 2
insert
BbsI
BbsI
direct repeat
direct repeat
5' - . . CAAACGGTCTTCGAGAAGACG TTTAGAGCTATGCTGTTTTGAATGGTCCCA. . -3'
3' - . . GTTTTGCCAGAAGCTCTTCTGCAAAATCTCGATACGACAAAACTTACCAGGGT. . -5'
guide oligo insertion site
3x
U6
Cbh FLAG NLS SpCas9 NLS bGH pA H1 tracrRNA
PX260
PciI, AflIII
BbsI BbsI
XbaI, KpnI
AgeI
EcoRI
NotI
BsrGI

FIG.19B

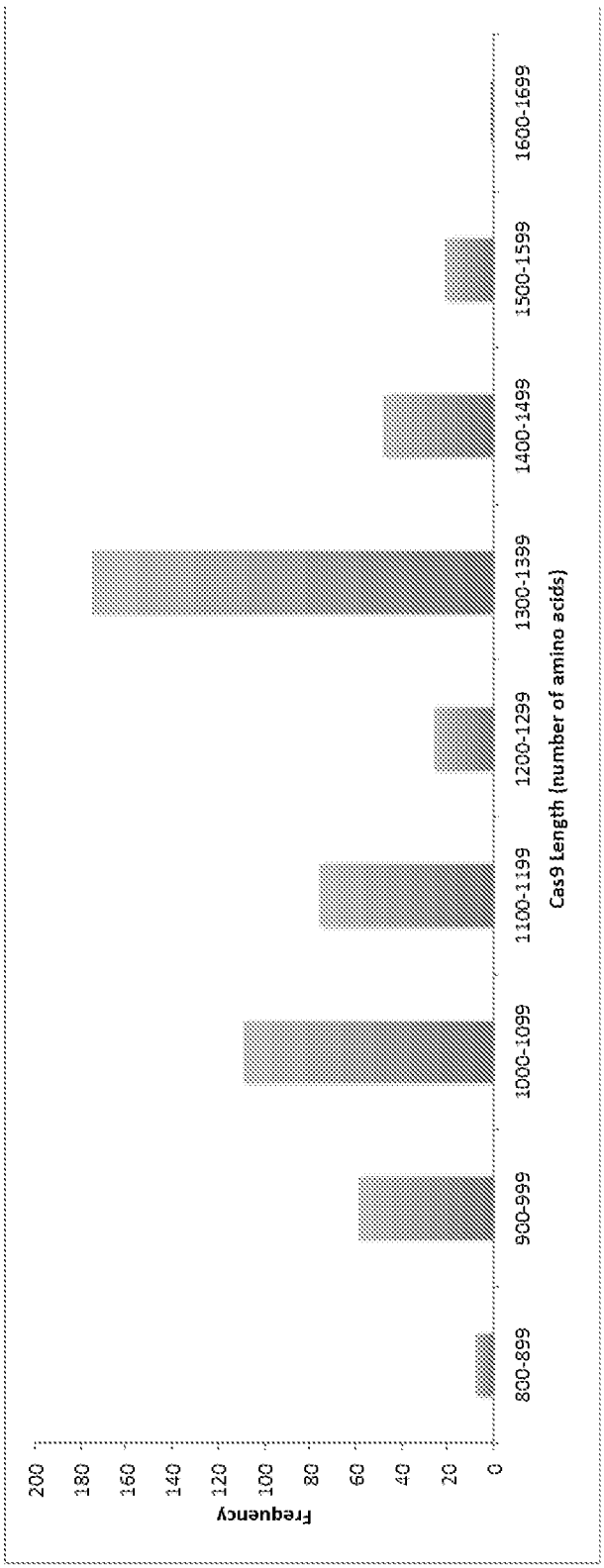

Frequency
200
180
160
140
120
100
80
60
40
20
0
800-899
900-999
1000-1099
1100-1199
1200-1299
1300-1399
1400-1499
1500-1599
1600-1699
Cas9 Length (number of amino acids)

FIG. 20

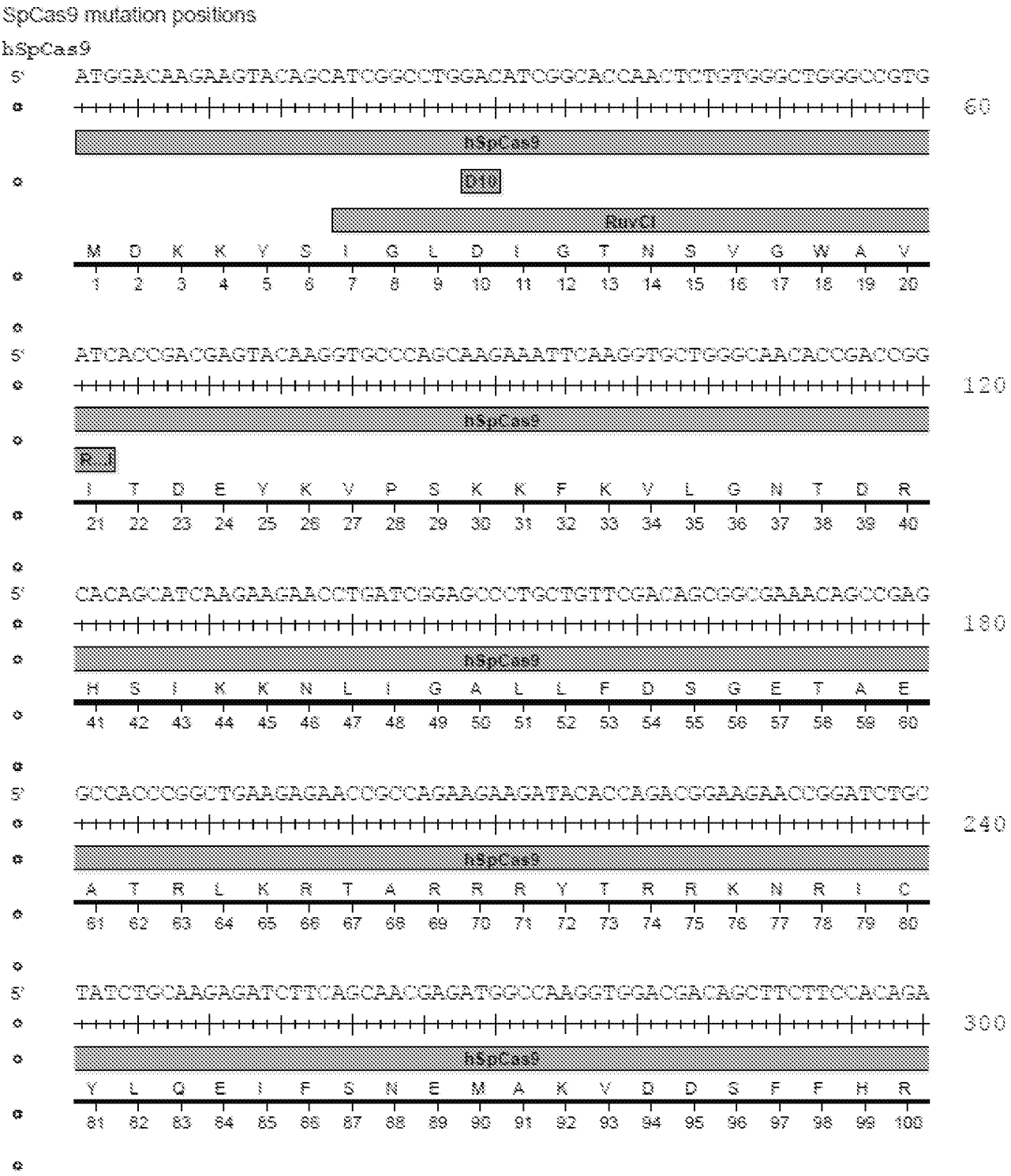

SpCas9 mutation positions
hSpCas9
5'
ATGGACAAGAAGTACAGCATCGGCCTGGACATCGGCACCAACTCTGTGGGCTGGGCCGTG
60
hSpCas9
RuvC1
M D K K Y S I G L D I G T N S V G W A V
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
5'
ATCACCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGG
120
hSpCas9
I T D E Y K V P S K K F K V L G N T D R
21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
5'
CACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAG
180
hSpCas9
H S I K K N L I G A L L F D S G E T A E
41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
5'
GCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTGC
240
hSpCas9
A T R L K R T A R R R Y T R R K N R I C
61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80
5'
TATCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCCACAGA
300
hSpCas9
Y L Q E I F S N E M A K V D D S F F H R
81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

FIG. 21A

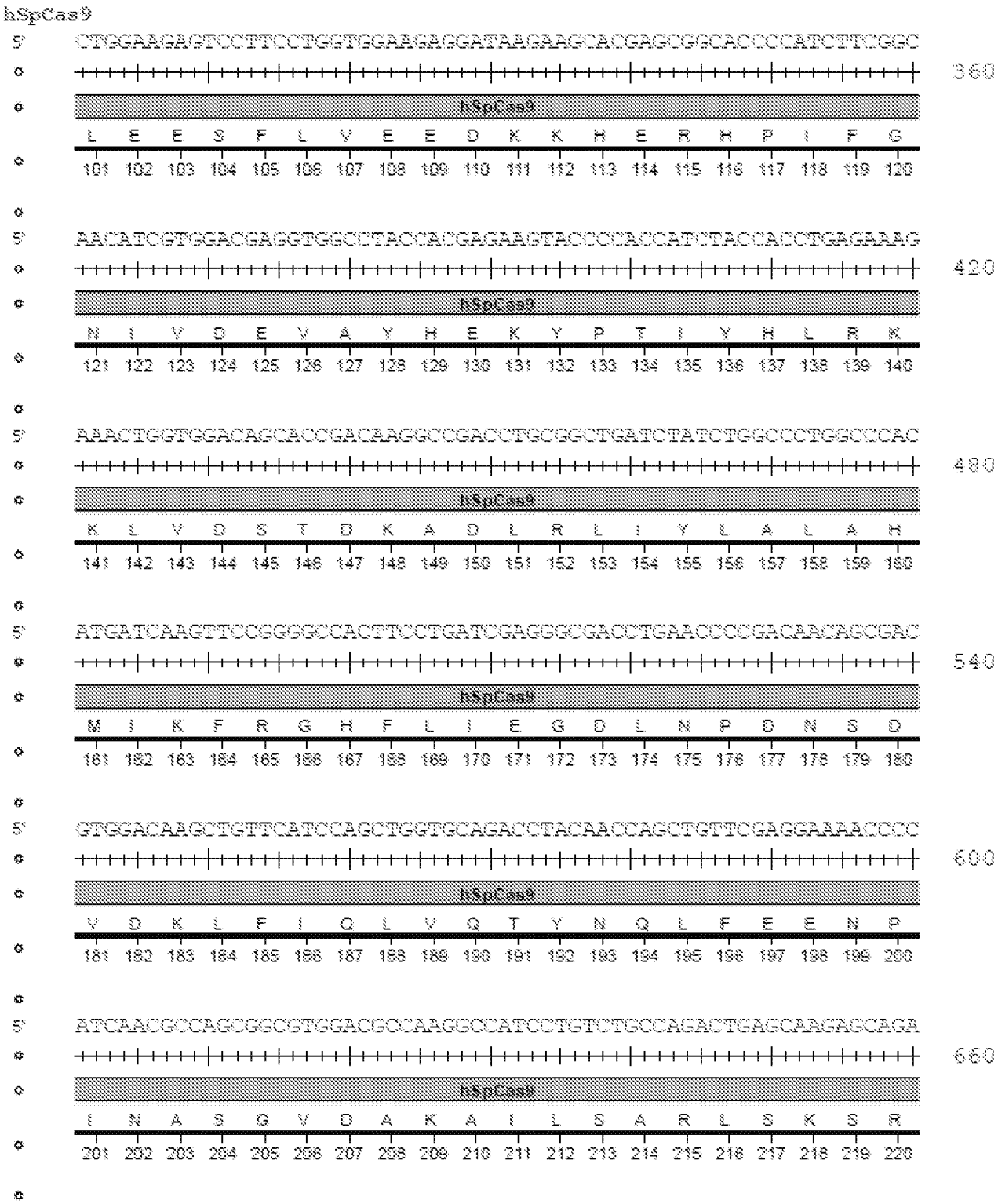

hSpCas9
5' CTGGAAGAGTCCTTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGC
360
hSpCas9
L E E S F L V E E D K K H E R H P I F G
101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
5' AACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAAG
420
hSpCas9
N I V D E V A Y H E K Y P T I Y H L R K
121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140
5' AAACTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCAC
480
hSpCas9
K L V D S T D K A D L R L I Y L A L A H
141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160
5' ATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGAC
540
hSpCas9
M I K F R G H F L I E G D L N P D N S D
161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180
5' GTGGACAAGCTGTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCC
600
hSpCas9
V D K L F I Q L V Q T Y N Q L F E E N P
181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200
5' ATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGA
660
hSpCas9
I N A S G V D A K A I L S A R L S K S R
201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220

FIG. 21B

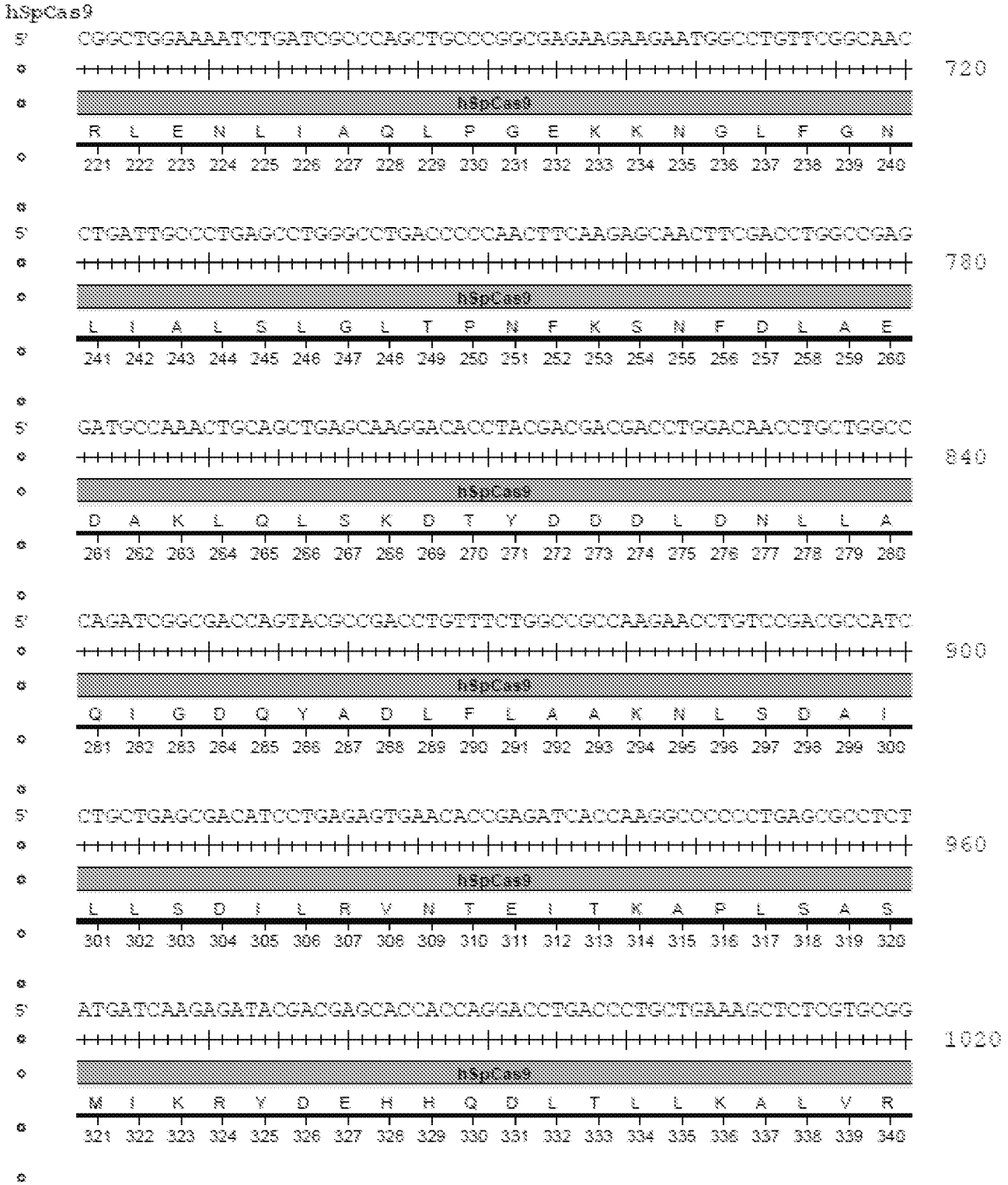

hSpCas9
5' CGGCTGGAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGCAAC
720
hSpCas9
R L E N L I A Q L P G E K K N G L F G N
221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240
5' CTGATTGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAG
780
hSpCas9
L I A L S L G L T P N F K S N F D L A E
241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260
5' GATGCCAAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCC
840
hSpCas9
D A K L Q L S K D T Y D D D L D N L L A
261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280
5' CAGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCCGACGCCATC
900
hSpCas9
Q I G D Q Y A D L F L A A K N L S D A I
281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
5' CTGCTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCCCCCCTGAGCGCCTCT
960
hSpCas9
L L S D I L R V N T E I T K A P L S A S
301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320
5' ATGATCAAGAGATACGACGAGCACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCGG
1020
hSpCas9
M I K R Y D E H H Q D L T L L K A L V R
321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340

FIG. 21C

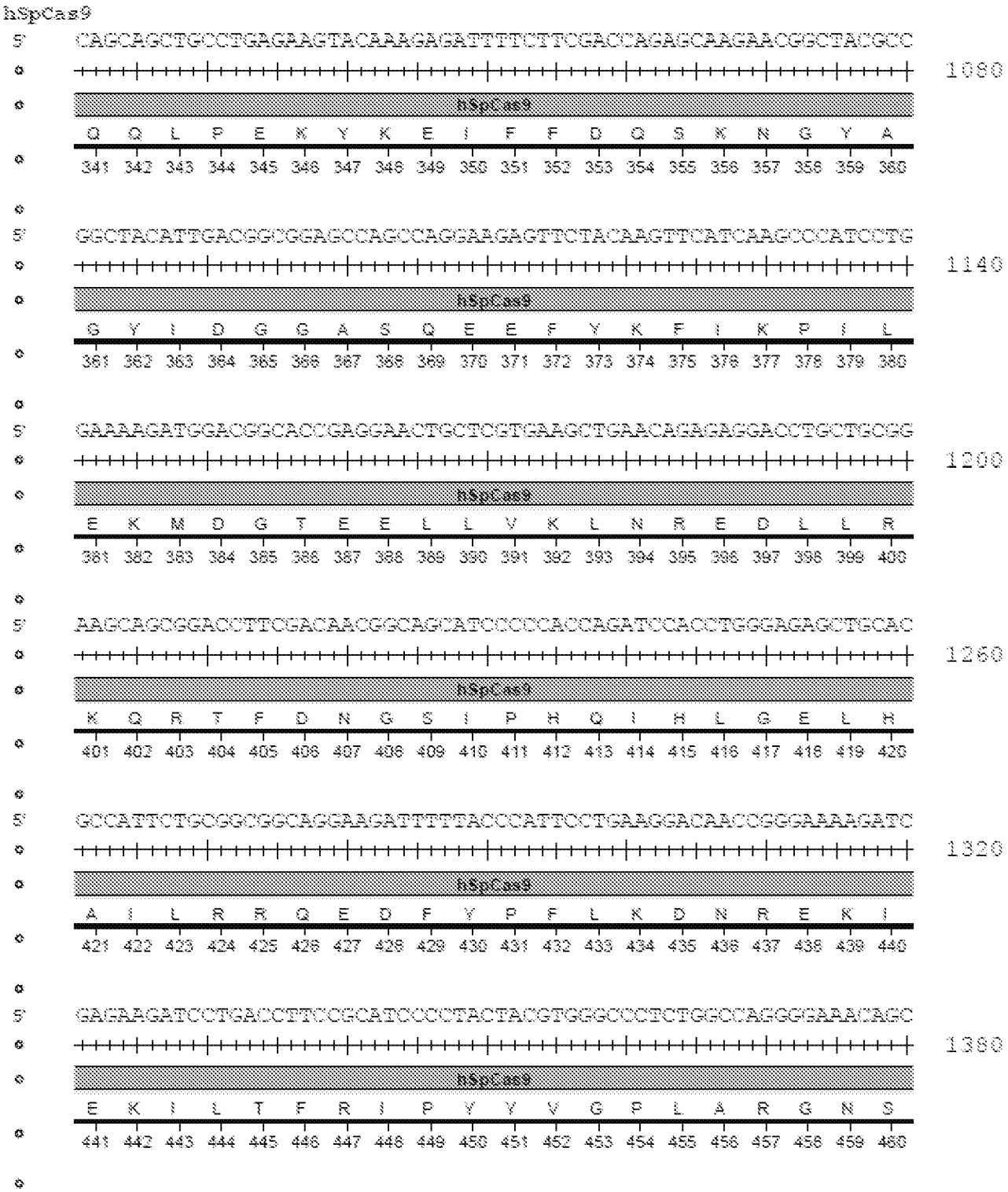

hSpCas9
5' CAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCC
1080
hSpCas9
Q Q L P E K Y K E I F F D Q S K N G Y A
341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360
5' GGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTG
1140
hSpCas9
G Y I D G G A S Q E E F Y K F I K P I L
361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380
5' GAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACCTGCTGCGG
1200
hSpCas9
E K M D G T E E L L V K L N R E D L L R
381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400
5' AAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTGGGAGAGCTGCAC
1260
hSpCas9
K Q R T F D N G S I P H Q I H L G E L H
401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420
5' GCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACAACCGGGAAAAGATC
1320
hSpCas9
A I L R R Q E D F Y P F L K D N R E K I
421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440
5' GAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACAGC
1380
hSpCas9
E K I L T F R I P Y Y V G P L A R G N S
441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460

FIG. 21D

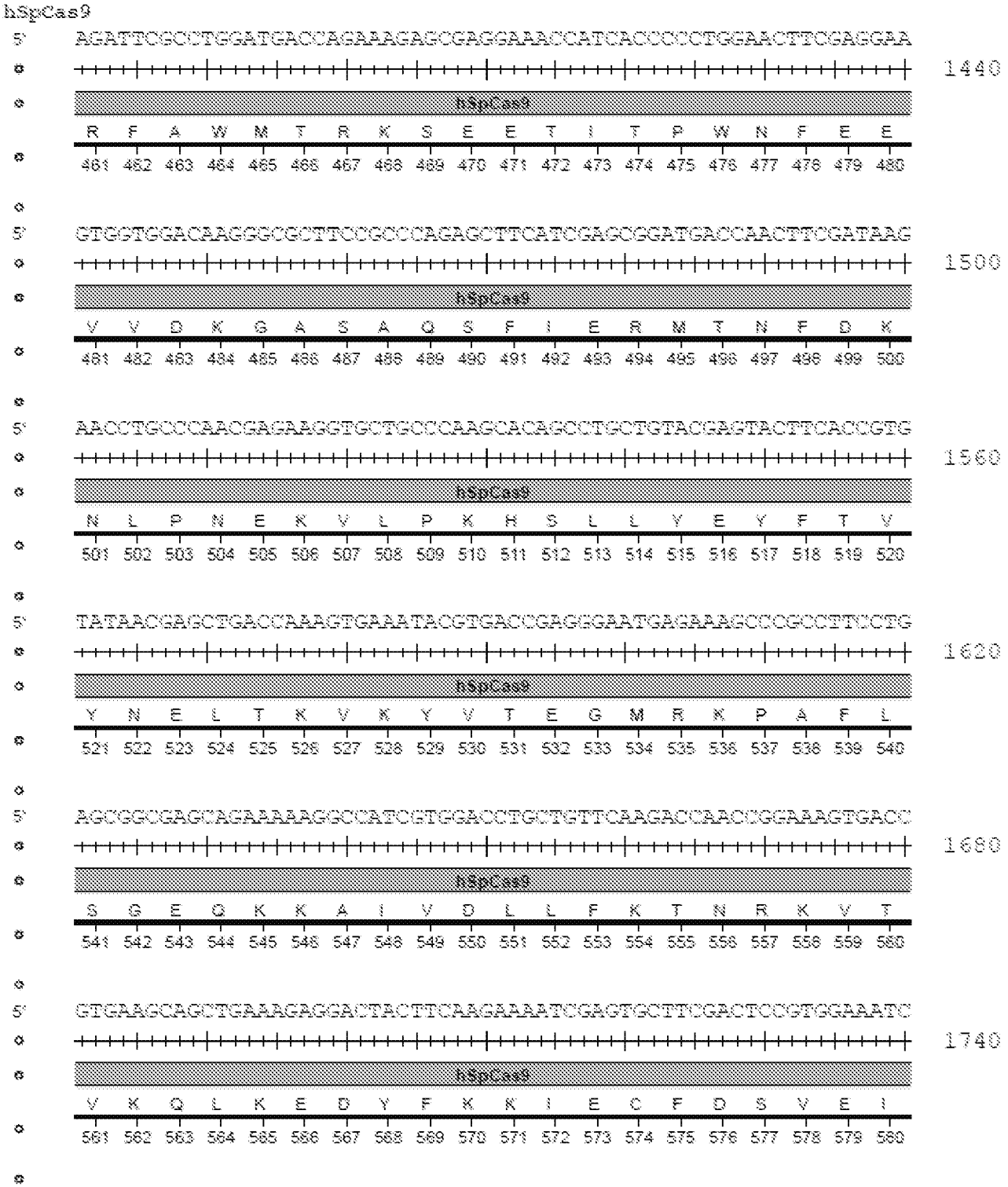

hSpCas9
5' AGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAA
1440
hSpCas9
R F A W M T R K S E E T I T P W N F E E
461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480
5' GTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAAG
1500
hSpCas9
V V D K G A S A Q S F I E R M T N F D K
481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500
5' AACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTG
1560
hSpCas9
N L P N E K V L P K H S L L Y E Y F T V
501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520
5' TATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTG
1620
hSpCas9
Y N E L T K V K Y V T E G M R K P A F L
521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540
5' AGCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACC
1680
hSpCas9
S G E Q K K A I V D L L F K T N R K V T
541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560
5' GTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATC
1740
hSpCas9
V K Q L K E D Y F K K I E C F D S V E I
561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580

FIG. 21E

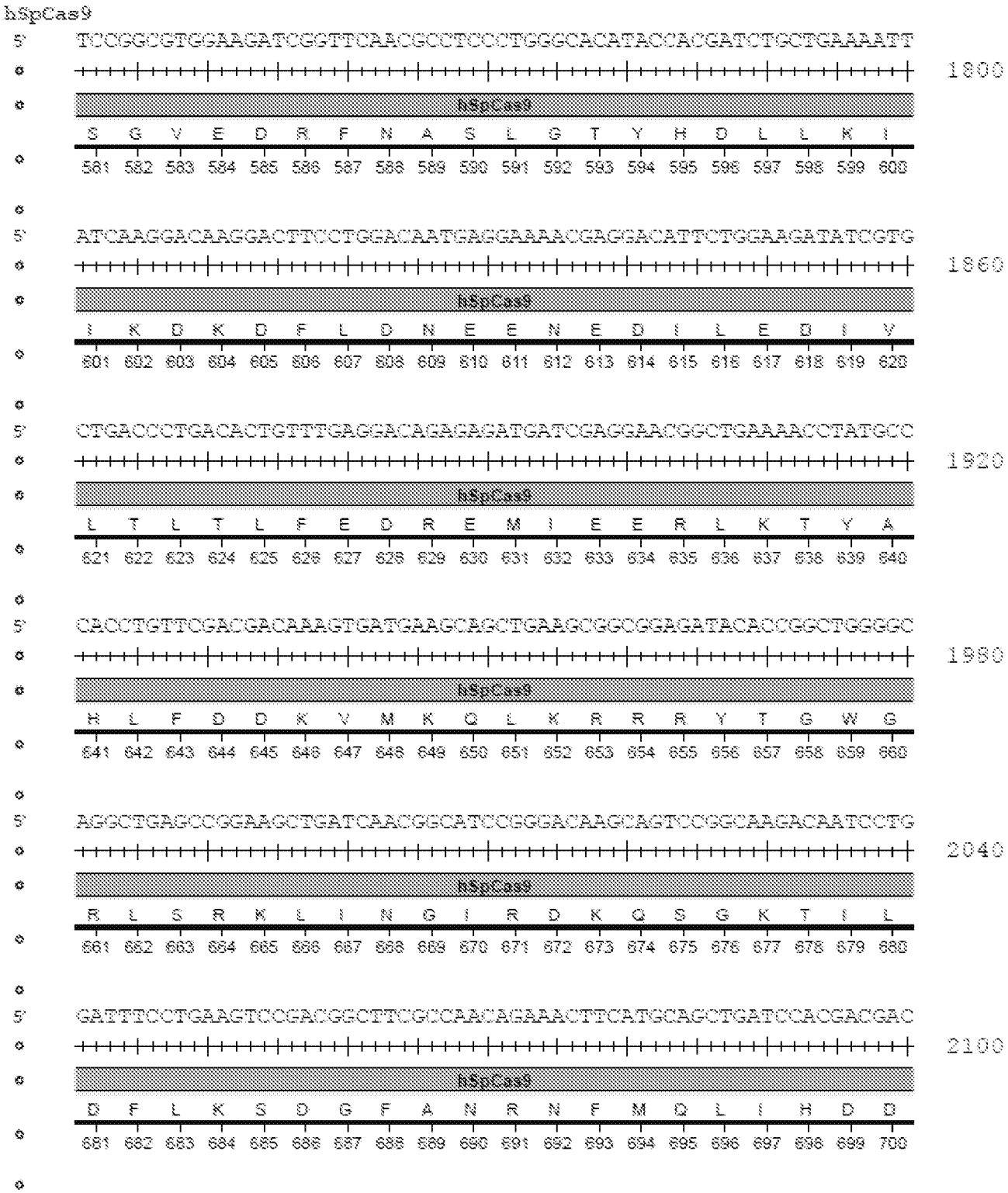

hSpCas9
5' TCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGAAAATT
1800
hSpCas9
S G V E D R F N A S L G T Y H D L L K I
561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580
5' ATCAAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTG
1860
hSpCas9
I K D K D F L D N E N E D I L E D I V
581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600
5' CTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCC
1920
hSpCas9
L T L T L F E D R E M I E E R L K T Y A
601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620
5' CACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGC
1980
hSpCas9
H L F D D K V M K Q L K R R R Y T G W G
621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640
5' AGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTG
2040
hSpCas9
R L S R K L I N G I R D K Q S G K T I L
641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660
5' GATTTCCTGAAGTCCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGAC
2100
hSpCas9
D F L K S D G F A N R N F M Q L I H D D
661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680

FIG. 21F

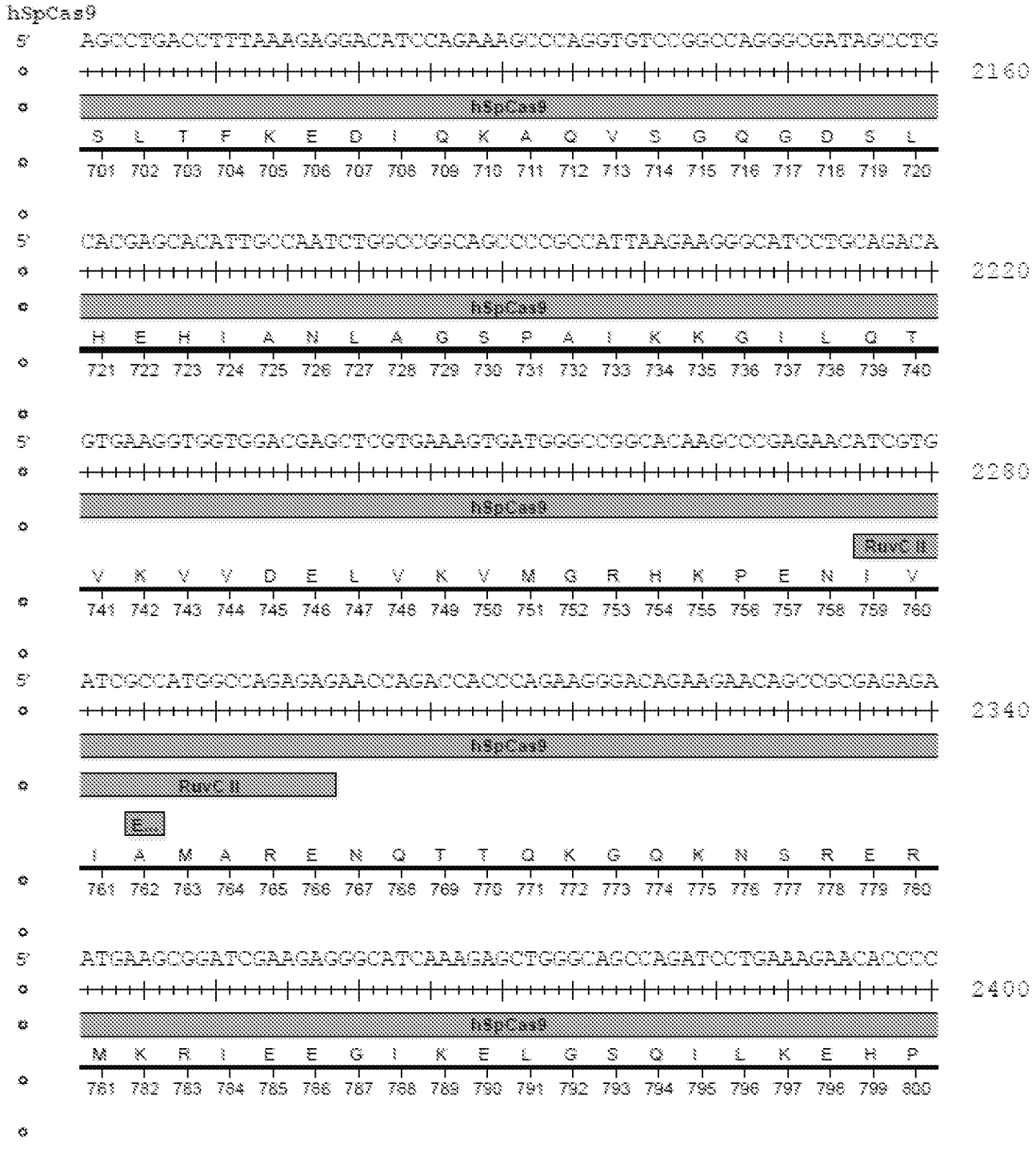

hSpCas9
5' AGCCTGACCTTTAAAGAGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTG
2160
hSpCas9
S L T F K E D I Q K A Q V S G Q G D S L
701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720
5' CACGAGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACA
2220
hSpCas9
H E H I A N L A G S P A I K K G I L Q T
721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740
5' GTGAAGGTGGTGGACGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTG
2280
hSpCas9
RuvC II
V K V V D E L V K V M G R H K P E N I V
741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760
5' ATCGCCATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGA
2340
hSpCas9
RuvC II
E
I A M A R E N Q T T Q K G Q K N S R E R
761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780
5' ATGAAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCC
2400
hSpCas9
M K R I E E G I K E L G S Q I L K E H P
781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800

FIG. 21G

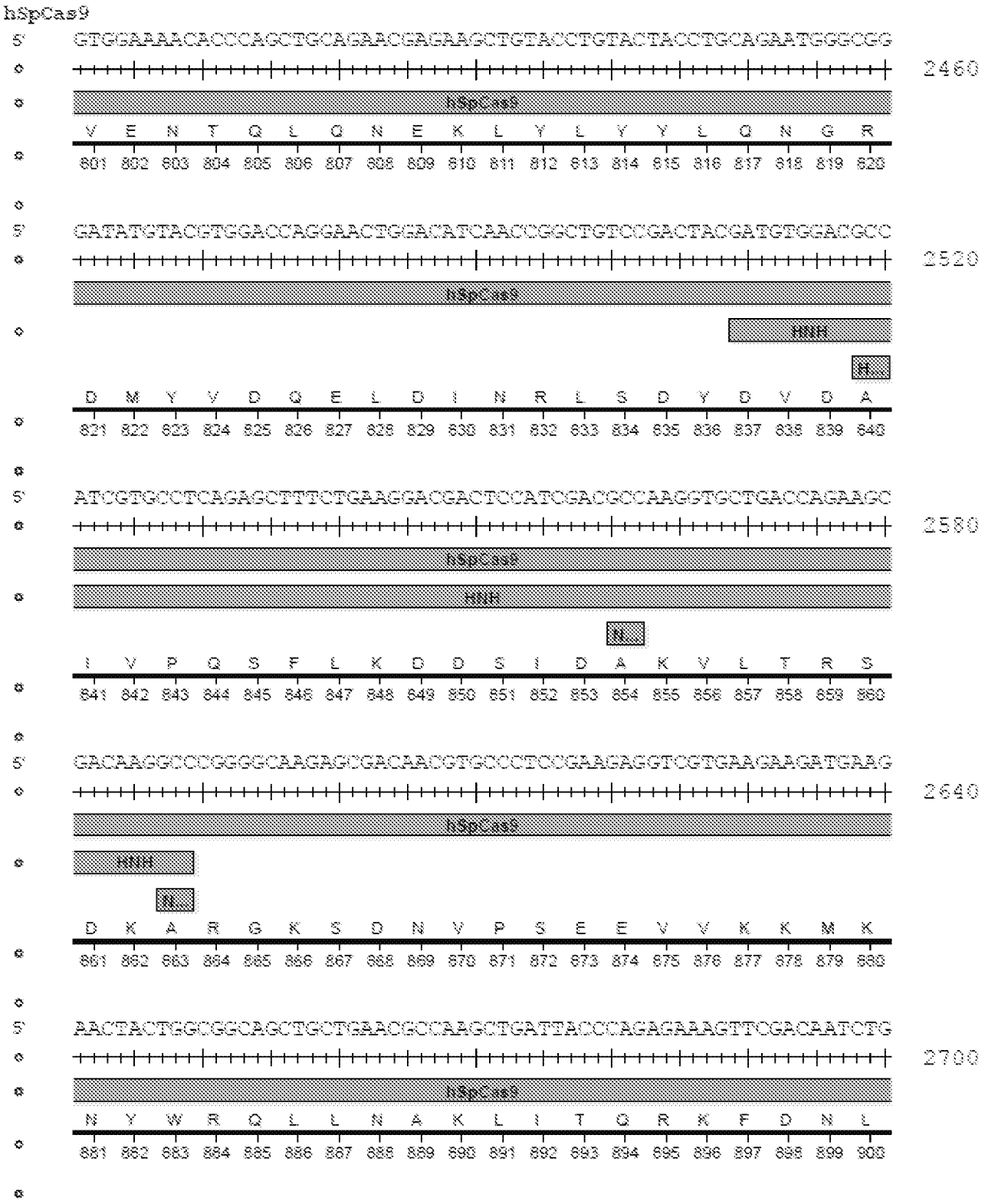

hSpCas9
5' GTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGG
2460
hSpCas9
V E N T Q L Q N E K L Y L Y Y L Q N G R
801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820
5' GATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATGTGGACGCC
2520
hSpCas9
HNH
H...
D M Y V D Q E L D I N R L S D Y D V D A
821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840
5' ATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACGCCAAGGTGCTGACCAGAAGC
2580
hSpCas9
HNH
N...
I V P Q S F L K D D S I D A K V L T R S
841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860
5' GACAAGGCCGGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAAG
2640
hSpCas9
HNH
N...
D K A R G K S D N V P S E E V V K K M K
861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880
5' AACTACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACCCAGAGAAAGTTCGACAATCTG
2700
hSpCas9
N Y W R Q L L N A K L I T Q R K F D N L
881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900

FIG. 21H

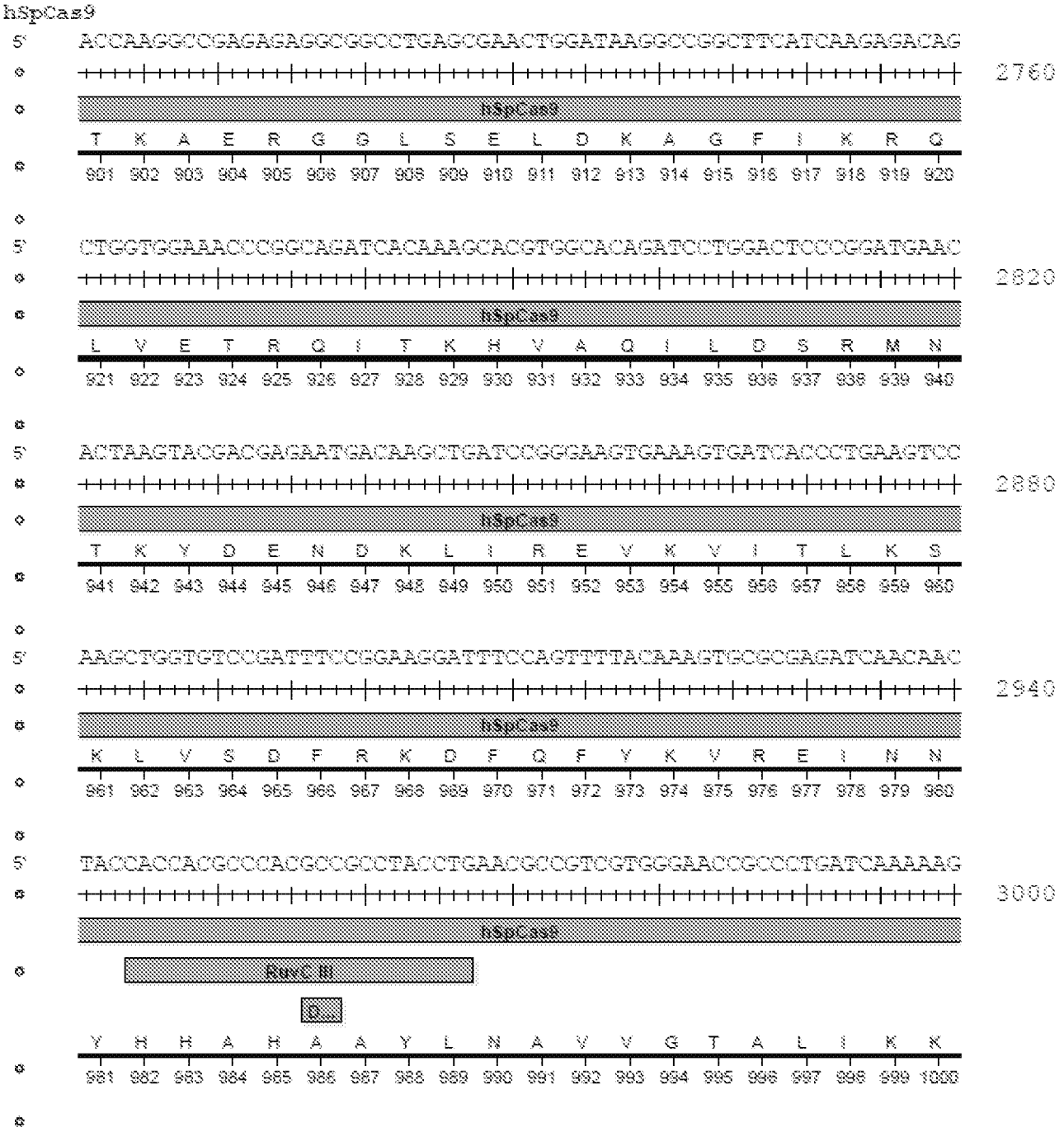

hSpCas9
5' ACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAG
2760
hSpCas9
T K A E R G G L S E L D K A G F I K R Q
901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920
5' CTGGTGGAAACCCGGCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAAC
2820
hSpCas9
L V E T R Q I T K H V A Q I L D S R M N
921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940
5' ACTAAGTACGACGAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCC
2880
hSpCas9
T K Y D E N D K L I R E V K V I T L K S
941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960
5' AAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACAAC
2940
hSpCas9
K L V S D F R K D F Q F Y K V R E I N N
961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980
5' TACCACCACGCCCACGCCGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAAAAAG
3000
hSpCas9
RuvC III
D
Y H H A H A A Y L N A V V G T A L I K K
981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

FIG. 21I

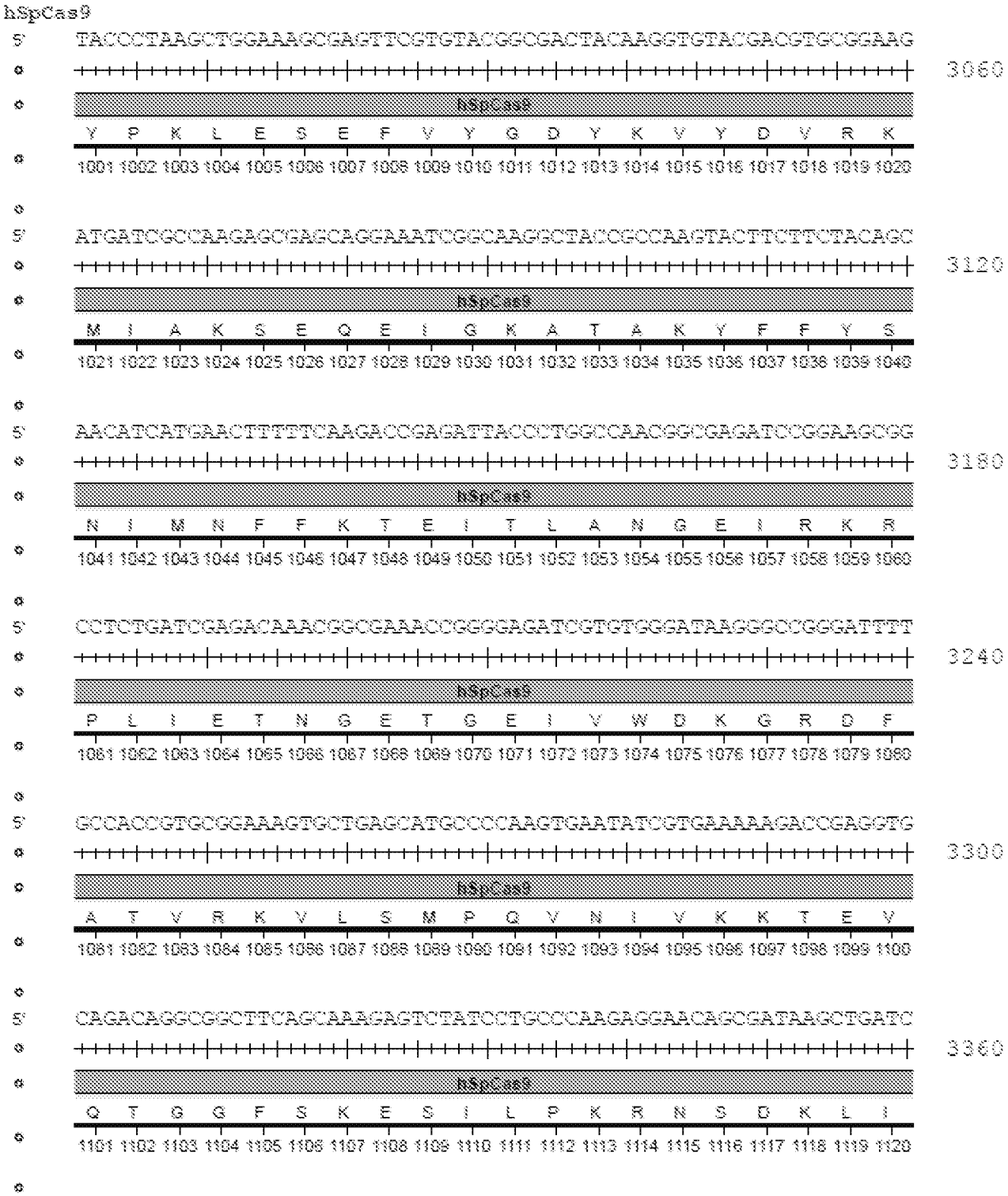

hSpCas9
5' TACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGAAG
3060
hSpCas9
Y P K L E S E F V Y G D Y K V Y D V R K
1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020
5' ATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGC
3120
hSpCas9
M I A K S E Q E I G K A T A K Y F F Y S
1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040
5' AACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGG
3180
hSpCas9
N I M N F F K T E I T L A N G E I R K R
1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060
5' CCTCTGATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTT
3240
hSpCas9
P L I E T N G E T G E I V W D K G R D F
1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080
5' GCCACCGTGCGGAAAGTGCTGAGCATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTG
3300
hSpCas9
A T V R K V L S M P Q V N I V K K T E V
1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100
5' CAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATC
3360
hSpCas9
Q T G G F S K E S I L P K R N S D K L I
1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120

FIG. 21J

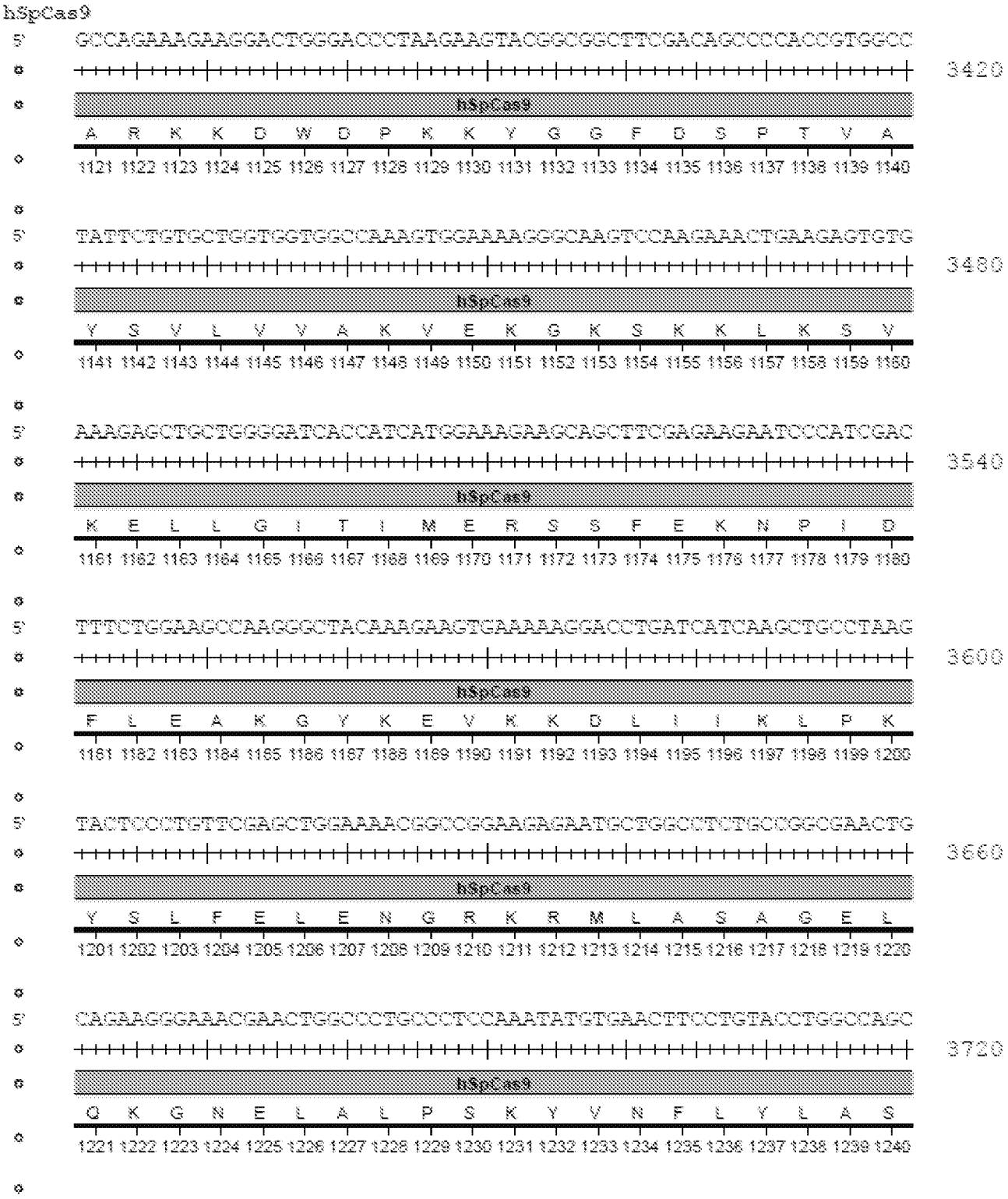

hSpCas9
5' GCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCC
3420
hSpCas9
A R K K D W D P K K Y G G F D S P T V A
1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140
5' TATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTG
3480
hSpCas9
Y S V L V V A K V E K G K S K K L K S V
1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160
5' AAAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGAC
3540
hSpCas9
K E L L G I T I M E R S S F E K N P I D
1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180
5' TTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTAAG
3600
hSpCas9
F L E A K G Y K E V K K D L I I K L P K
1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200
5' TACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTG
3660
hSpCas9
Y S L F E L E N G R K R M L A S A G E L
1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220
5' CAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGC
3720
hSpCas9
Q K G N E L A L P S K Y V N F L Y L A S
1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240

FIG. 21K

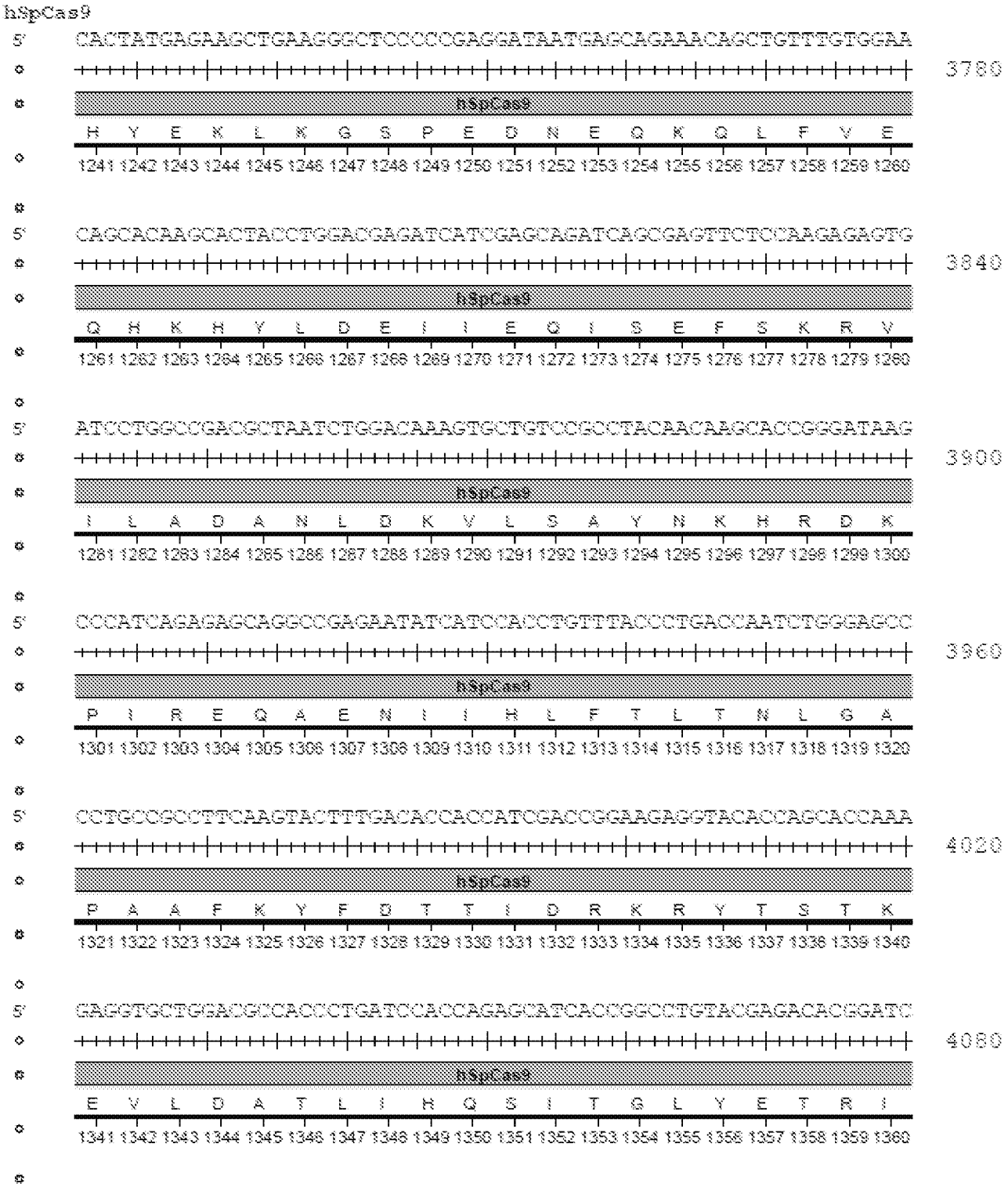

hSpCas9
5' CACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAA
3780
hSpCas9
H Y E K L K G S P E D N E Q K Q L F V E
1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260
5' CAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTG
3840
hSpCas9
Q H K H Y L D E I I E Q I S E F S K R V
1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280
5' ATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAG
3900
hSpCas9
I L A D A N L D K V L S A Y N K H R D K
1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298 1299 1300
5' CCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCC
3960
hSpCas9
P I R E Q A E N I I H L F T L T N L G A
1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320
5' CCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAA
4020
hSpCas9
P A A F K Y F D T T I D R K R Y T S T K
1321 1322 1323 1324 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340
5' GAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGATC
4080
hSpCas9
E V L D A T L I H Q S I T G L Y E T R I
1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354 1355 1356 1357 1358 1359 1360

FIG. 21L

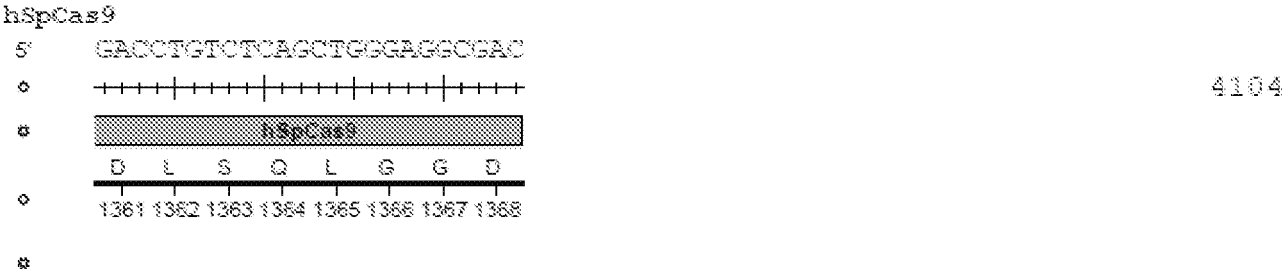

hSpCas9
5'
GACCTGTCTTCAGCTGGGAGGCGAC
4104
hSpCas9
D L S Q L G G D
1361 1362 1363 1364 1365 1366 1367 1368

FIG. 21M

Conditional Cas9, Rosa26 targeting vector map

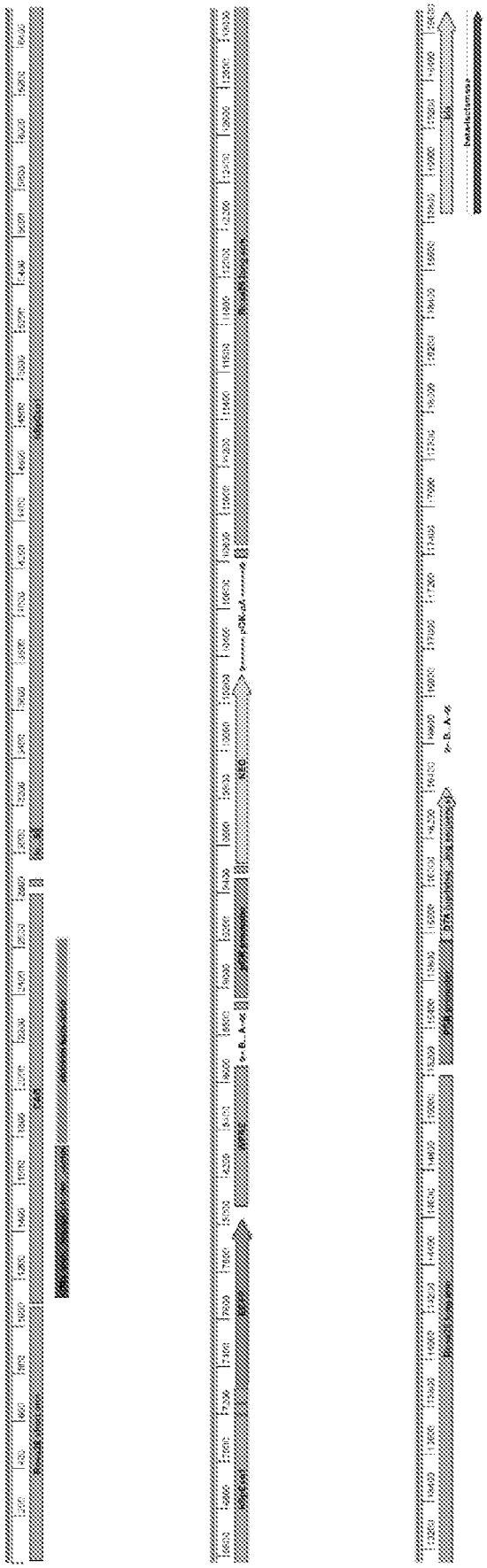
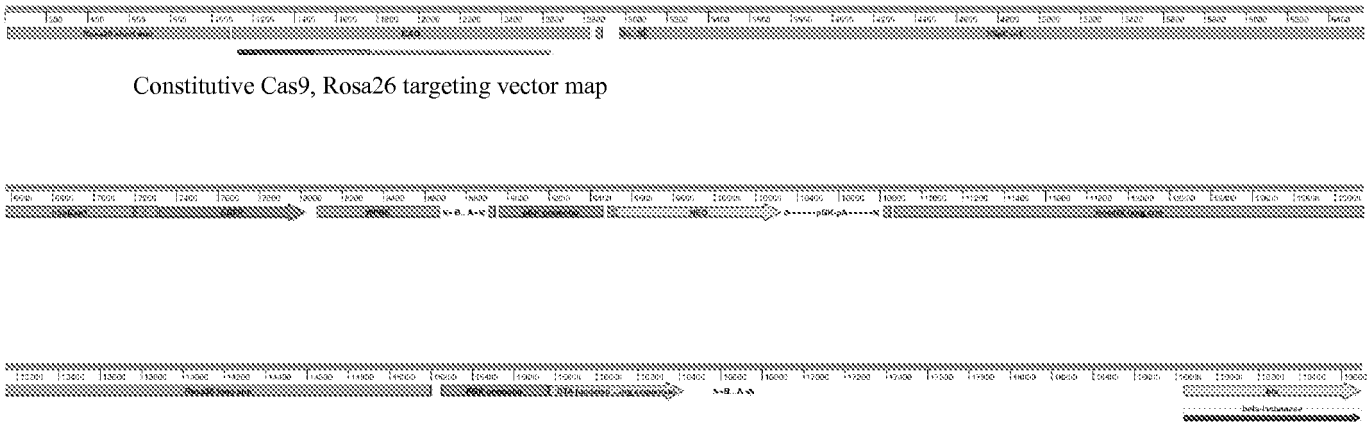


FIG. 22A



Constitutive Cas9, Rosa26 targeting vector map

FIG. 22B

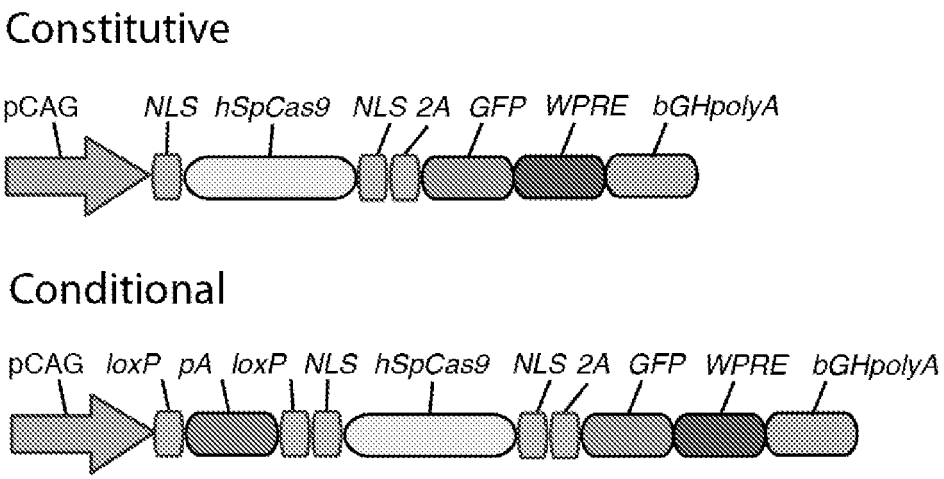

Constitutive
pCAG
NLS
hSpCas9
NLS
2A
GFP
WPRE
bGHpolyA
Conditional
pCAG
loxP
pA
loxP
NLS
hSpCas9
NLS
2A
GFP
WPRE
bGHpolyA

FIG. 23

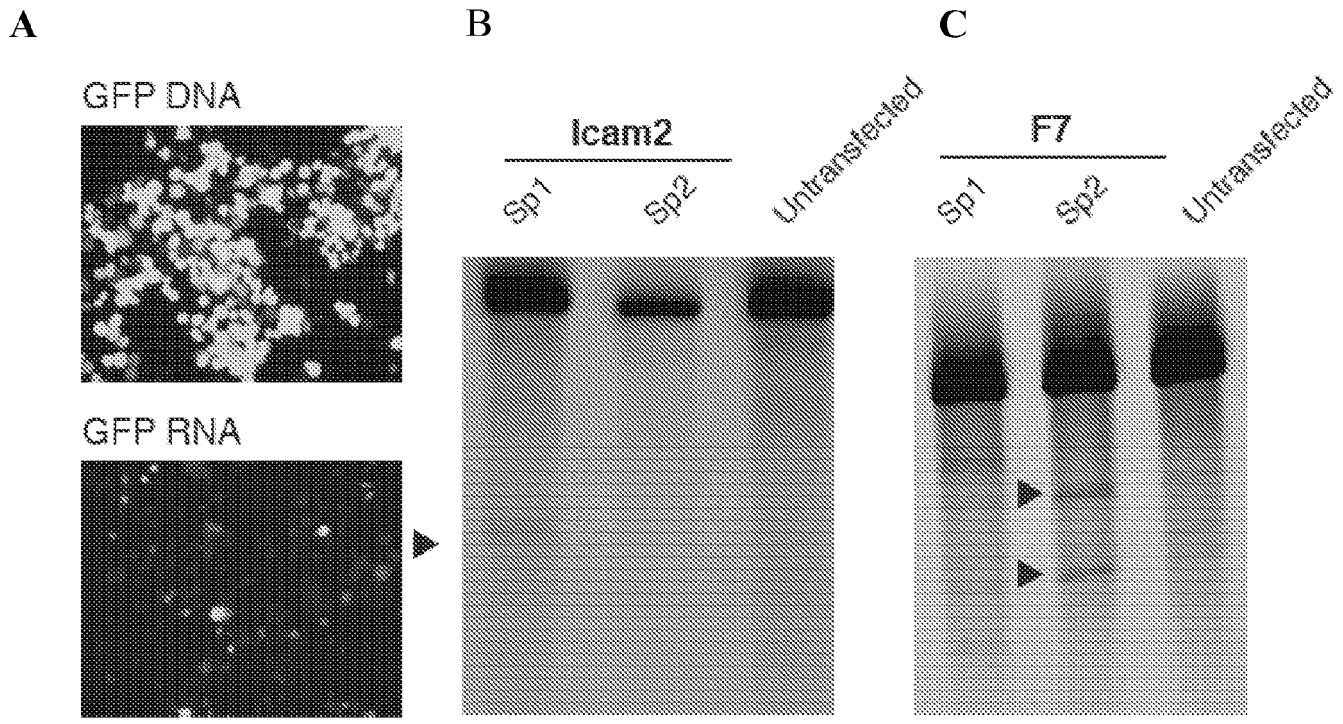

A
GFP DNA
GFP RNA
B
Icam2
Sp1
Sp2
Untransfected
C
F7
Sp1
Sp2
Untransfected

FIG. 24A-C

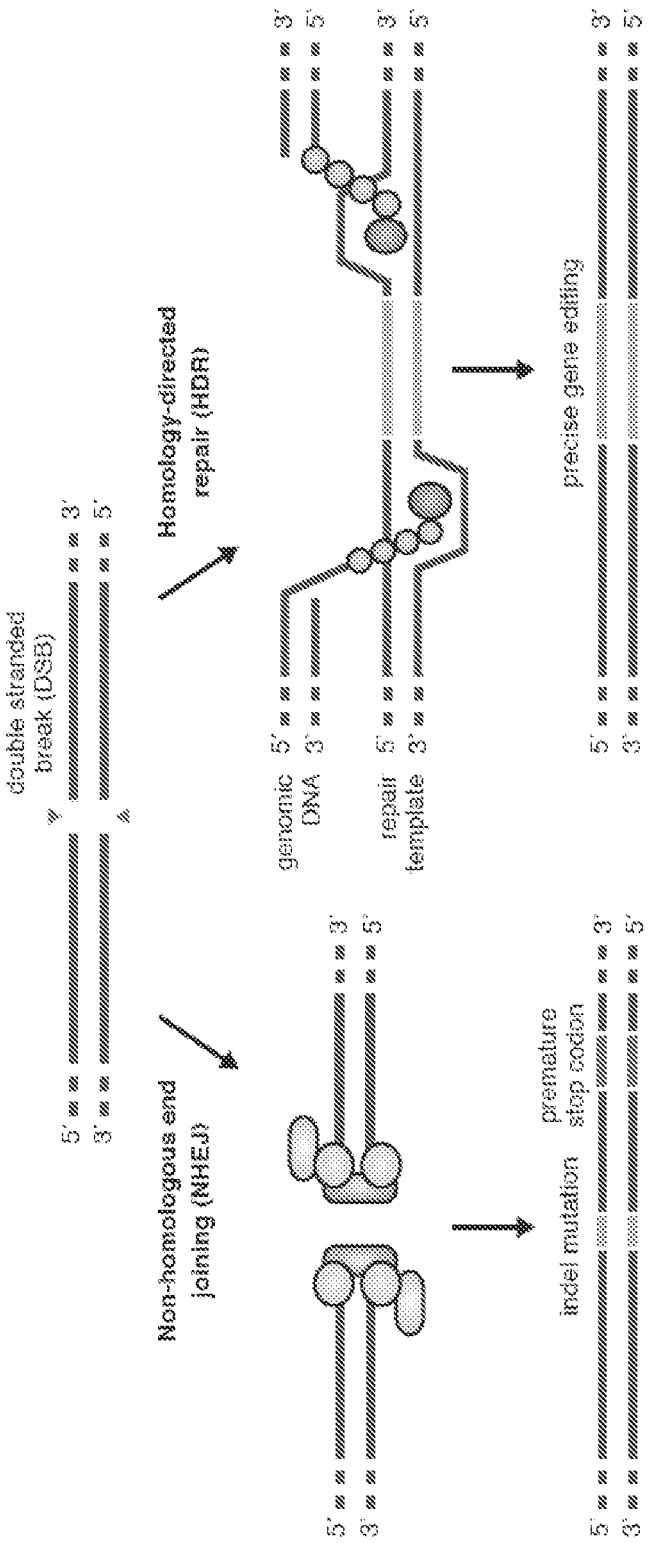

double stranded break (DSB)
5'
3'
Homology-directed repair (HDR)
genomic DNA
repair template
precise gene editing
Non-homologous end joining (NHEJ)
indel mutation
premature stop codon

FIG. 25

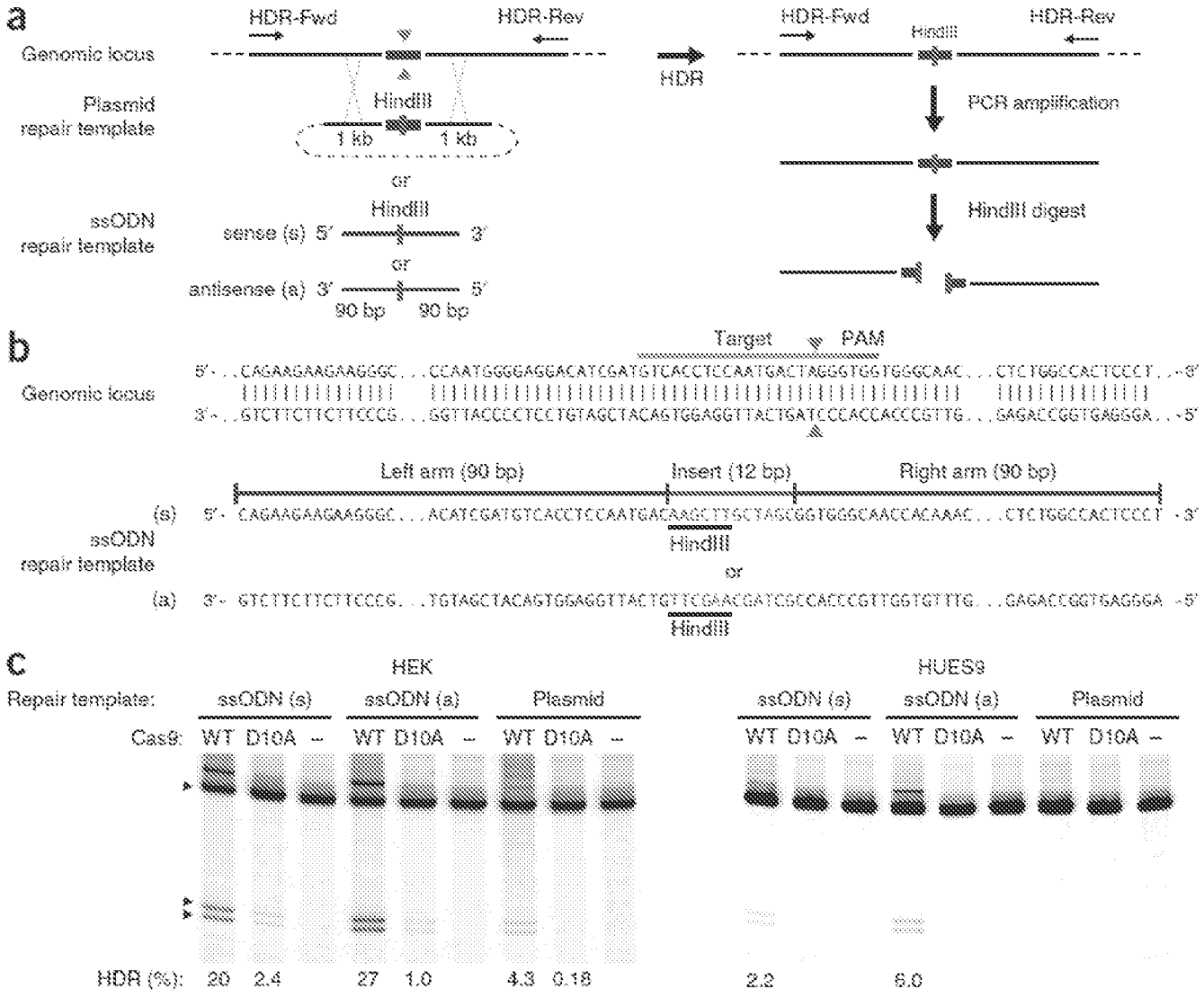

a
HDR-Fwd
HDR-Rev
Genomic locus
HDR
HindIII
PCR amplification
Plasmid repair template
1 kb
1 kb
or
ssODN repair template
sense (s) 5′
3′
antisense (a) 3′
5′
90 bp
90 bp
HindIII digest
b
Target
PAM
Genomic locus
Left arm (90 bp)
Insert (12 bp)
Right arm (90 bp)
(s)
(a)
ssODN repair template
HindIII
or
c
HEK
HUES9
Repair template:
ssODN (s)
ssODN (a)
Plasmid
Cas9:
WT
D10A
–
HDR (%):
20
2.4
27
1.0
4.3
0.18
2.2
6.0

FIG. 26A-C

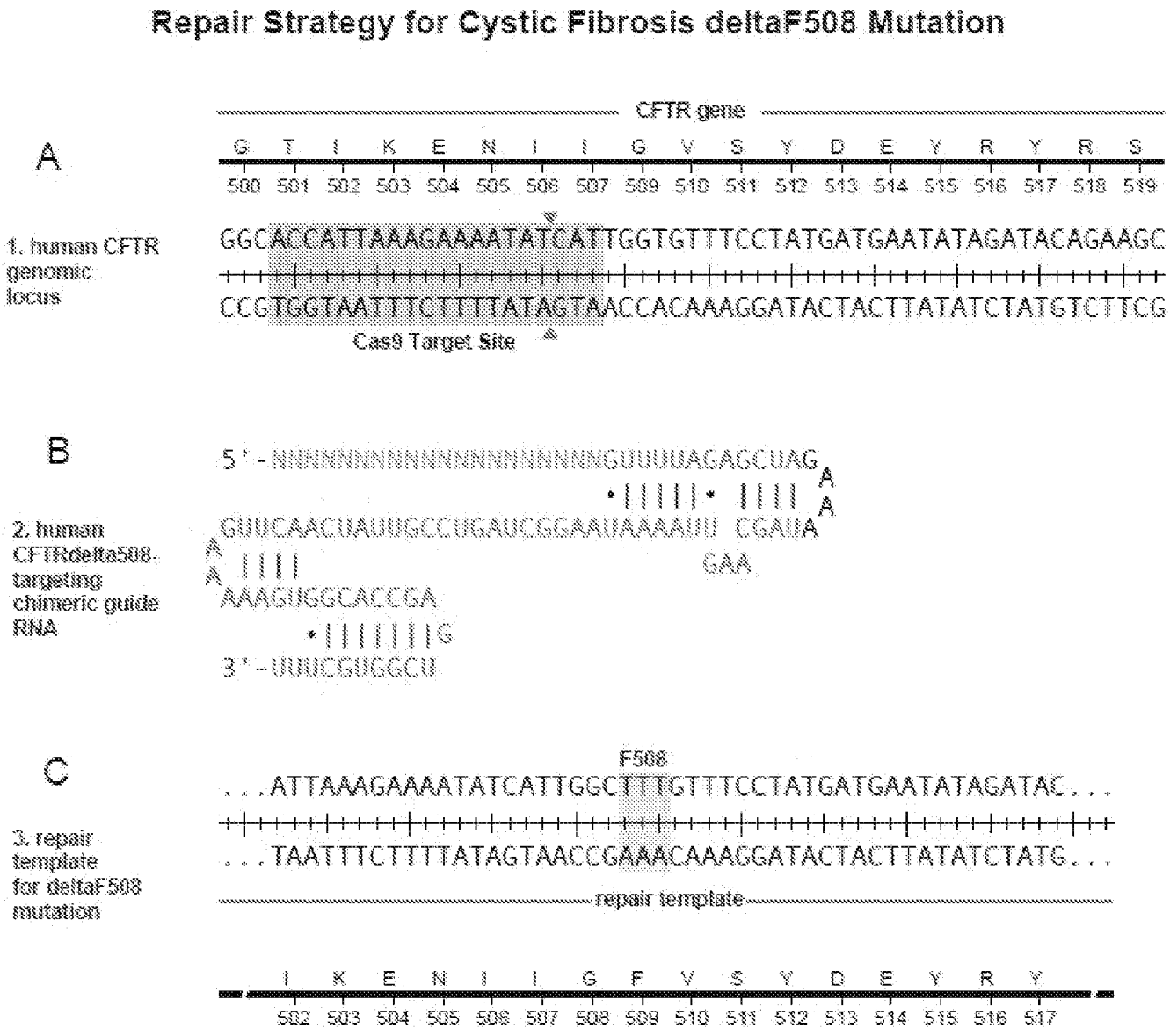

Repair Strategy for Cystic Fibrosis deltaF508 Mutation
A
CFTR gene
G T I K E N I I G V S Y D E Y R Y R S
500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519
1. human CFTR genomic locus
GGCACCATTAAAGAAAATATCATTGGTGTTTCCTATGATGAATATAGATACAGAAGC
CCGTGGTAATTTCTTTTATAGTAACCACAAAGGATACTACTTATATCTATGTCTTCG
Cas9 Target Site
B
2. human CFTRdelta508-targeting chimeric guide RNA
5'-NNNNNNNNNNNNNNNNNNNGUUUUAGAGCUAG
GUUCAACUAUUGCCUGAUCGGAAUAAAAUU CGAUA
AAAGUGGCACCGA
3'-UUUCGUGGCU
GAA
C
3. repair template for deltaF508 mutation
F508
...ATTAAAGAAAATATCATTGGCTTTGTTTCCTATGATGAATATAGATAC...
...TAATTTCTTTTATAGTAACCGAAACAAAGGATACTACTTATATCTATG...
repair template
I K E N I I G F V S Y D E Y R Y
502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517

FIG. 27A-C

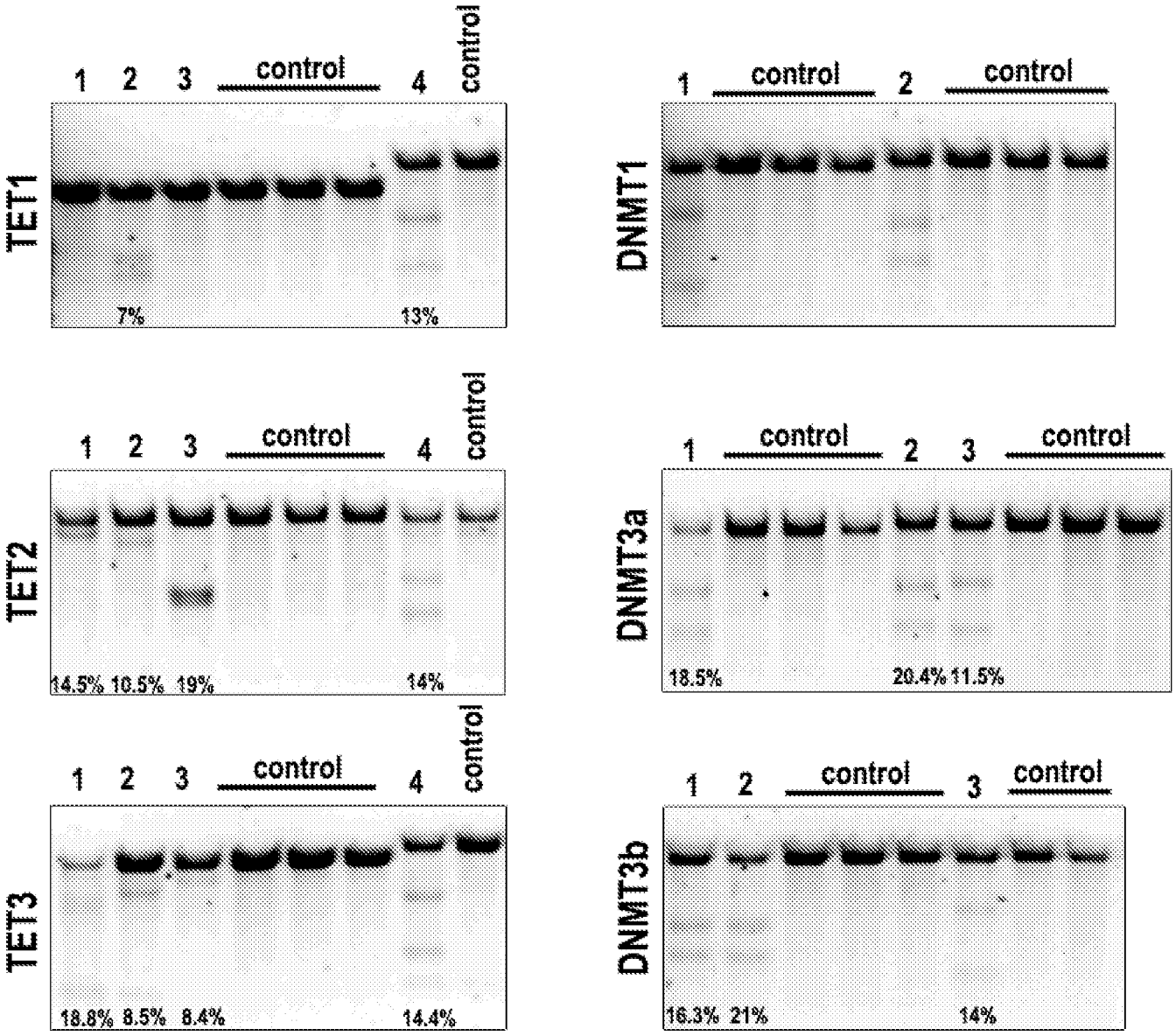

TET1
1 2 3 control 4 control
7%
13%
DNMT1
1 control 2 control
TET2
1 2 3 control 4 control
14.5% 10.5% 19%
14%
DNMT3a
1 control 2 3 control
18.5%
20.4% 11.5%
TET3
1 2 3 control 4 control
18.8% 8.5% 8.4%
14.4%
DNMT3b
1 2 control 3 control
16.3% 21%
14%

FIG. 28

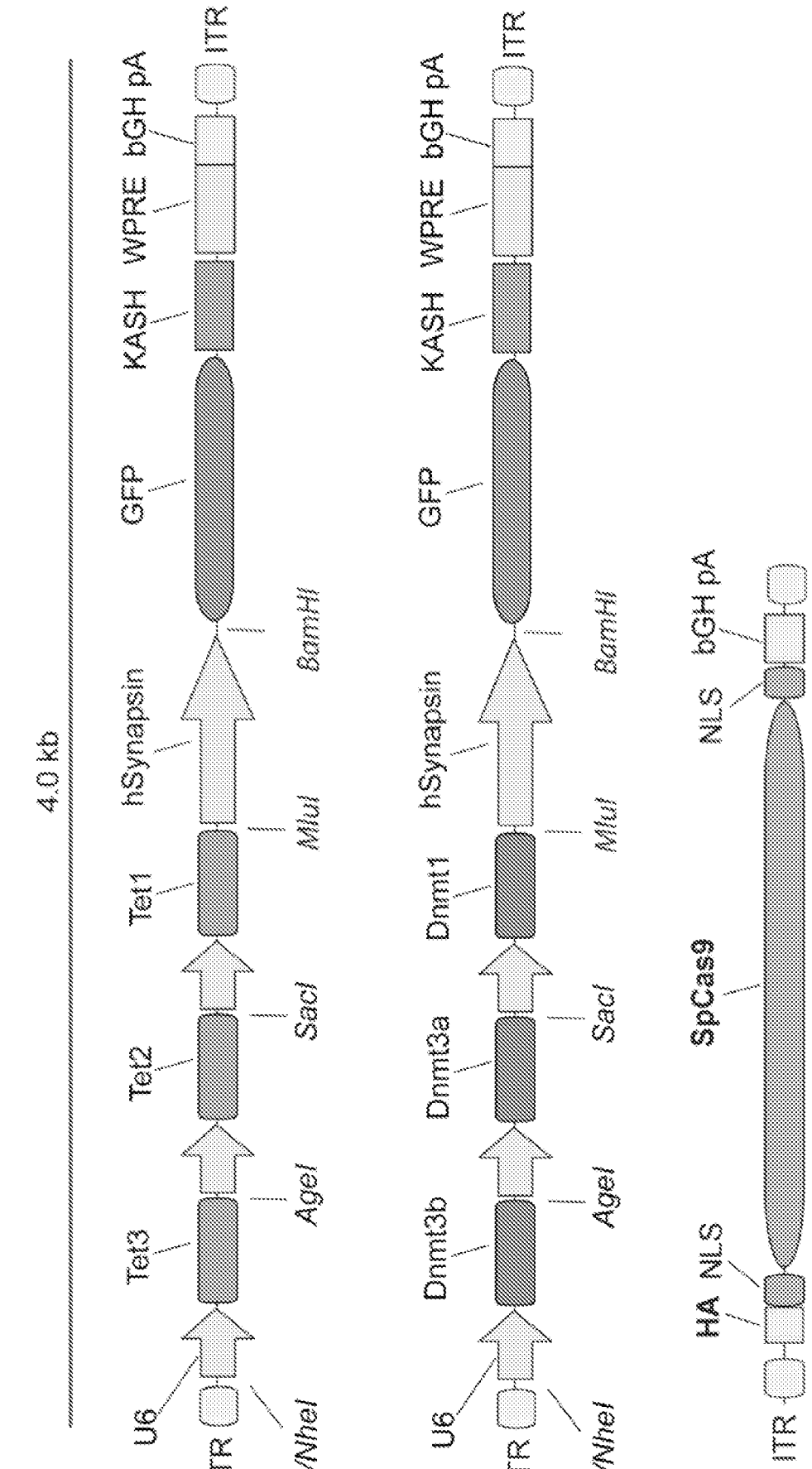

a
4.0 kb
U6
ITR
NotI/NheI
Tet3
AgeI
Tet2
SacI
Tet1
MluI
hSynapsin
BamHI
GFP
KASH
WPRE
bGH pA
ITR
U6
ITR
NotI/NheI
Dnmt3b
AgeI
Dnmt3a
SacI
Dnmt1
MluI
hSynapsin
BamHI
GFP
KASH
WPRE
bGH pA
ITR
HA
NLS
SpCas9
NLS
bGH pA
ITR

FIG. 29

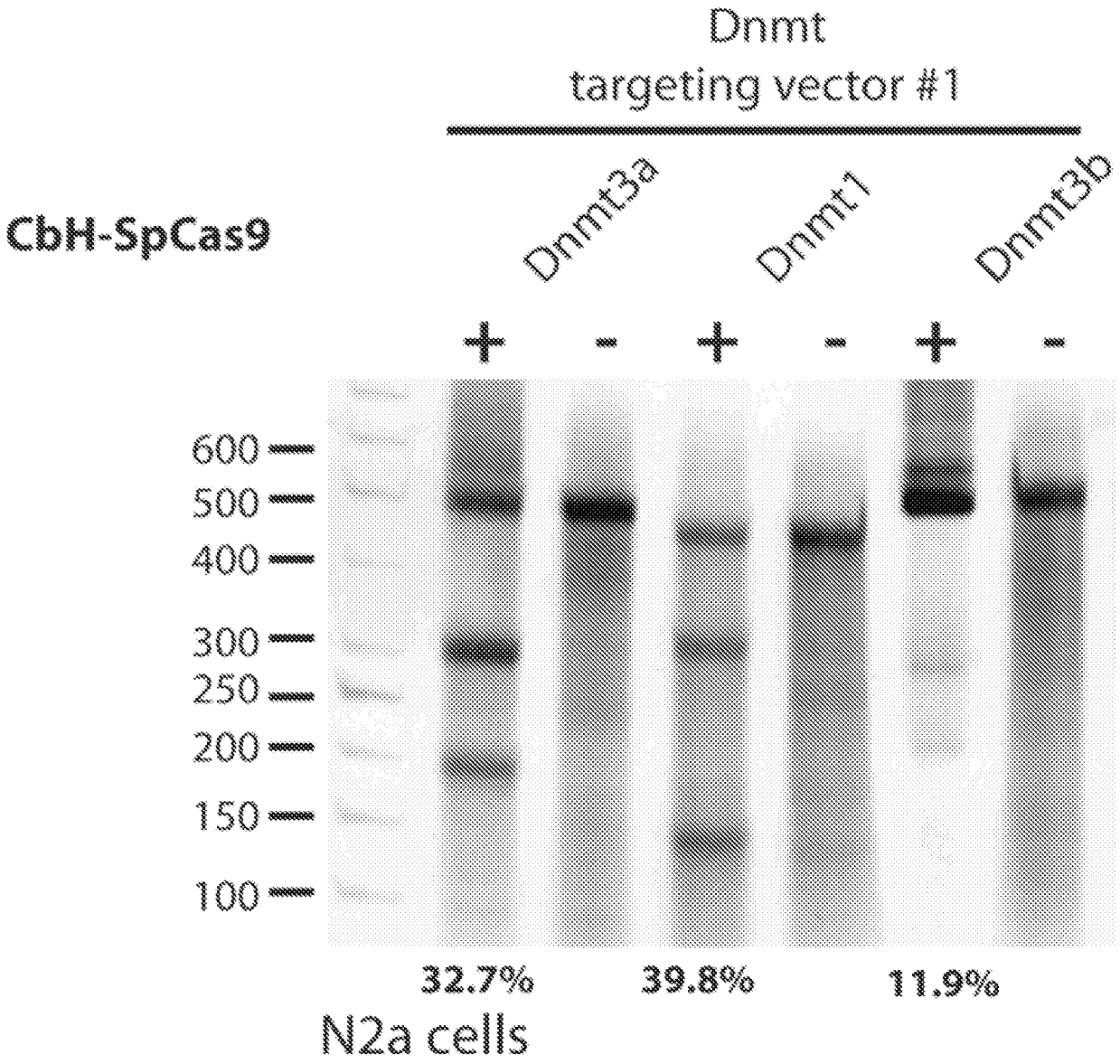

Dnmt
targeting vector #1
CbH-SpCas9
Dnmt3a
Dnmt1
Dnmt3b
+
-
+
-
+
-
600
500
400
300
250
200
150
100
32.7%
39.8%
11.9%
N2a cells

FIG. 30

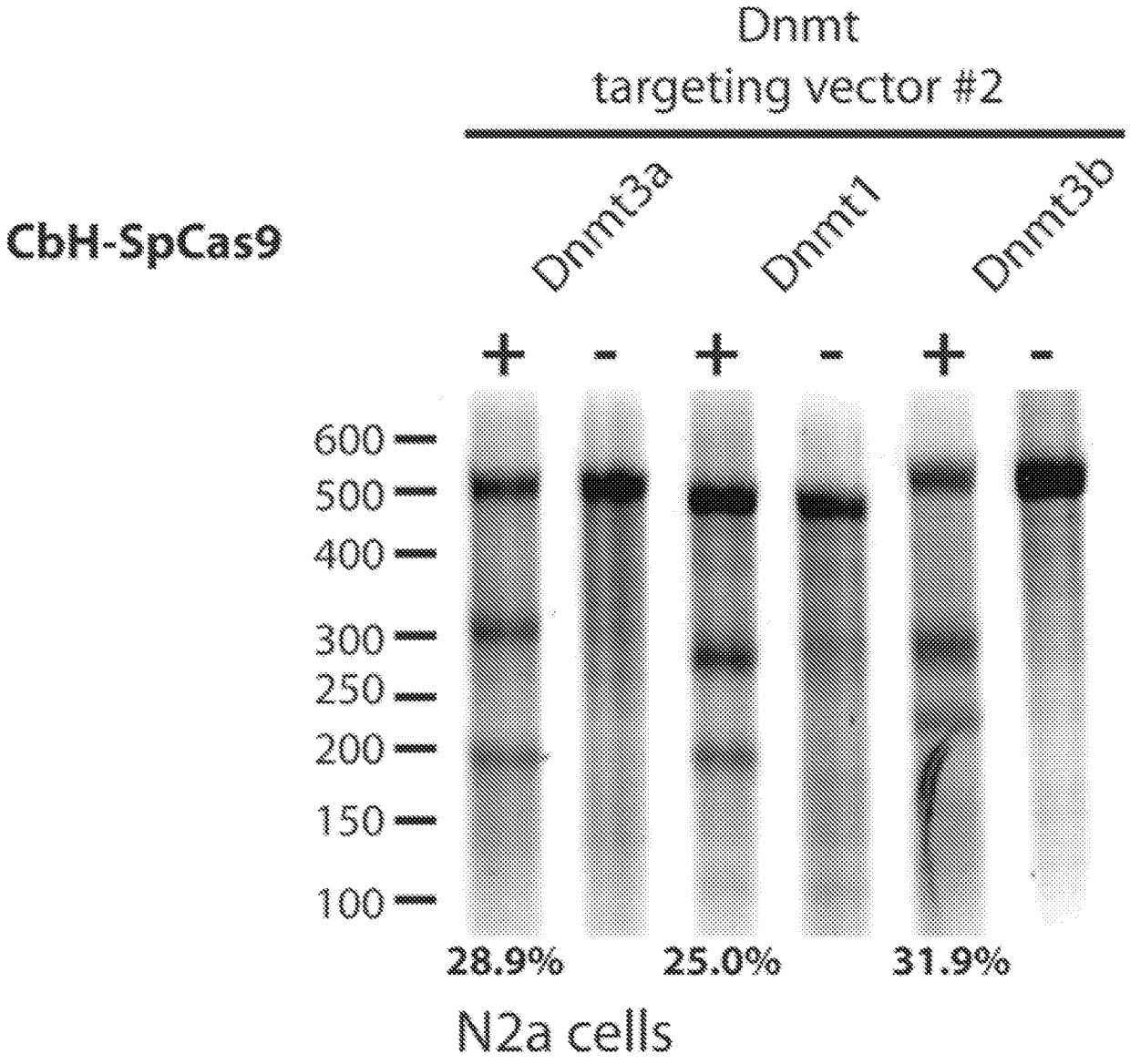

Dnmt
targeting vector #2
CbH-SpCas9
Dnmt3a
Dnmt1
Dnmt3b
+
-
+
-
+
-
600
500
400
300
250
200
150
100
28.9%
25.0%
31.9%
N2a cells

FIG. 31

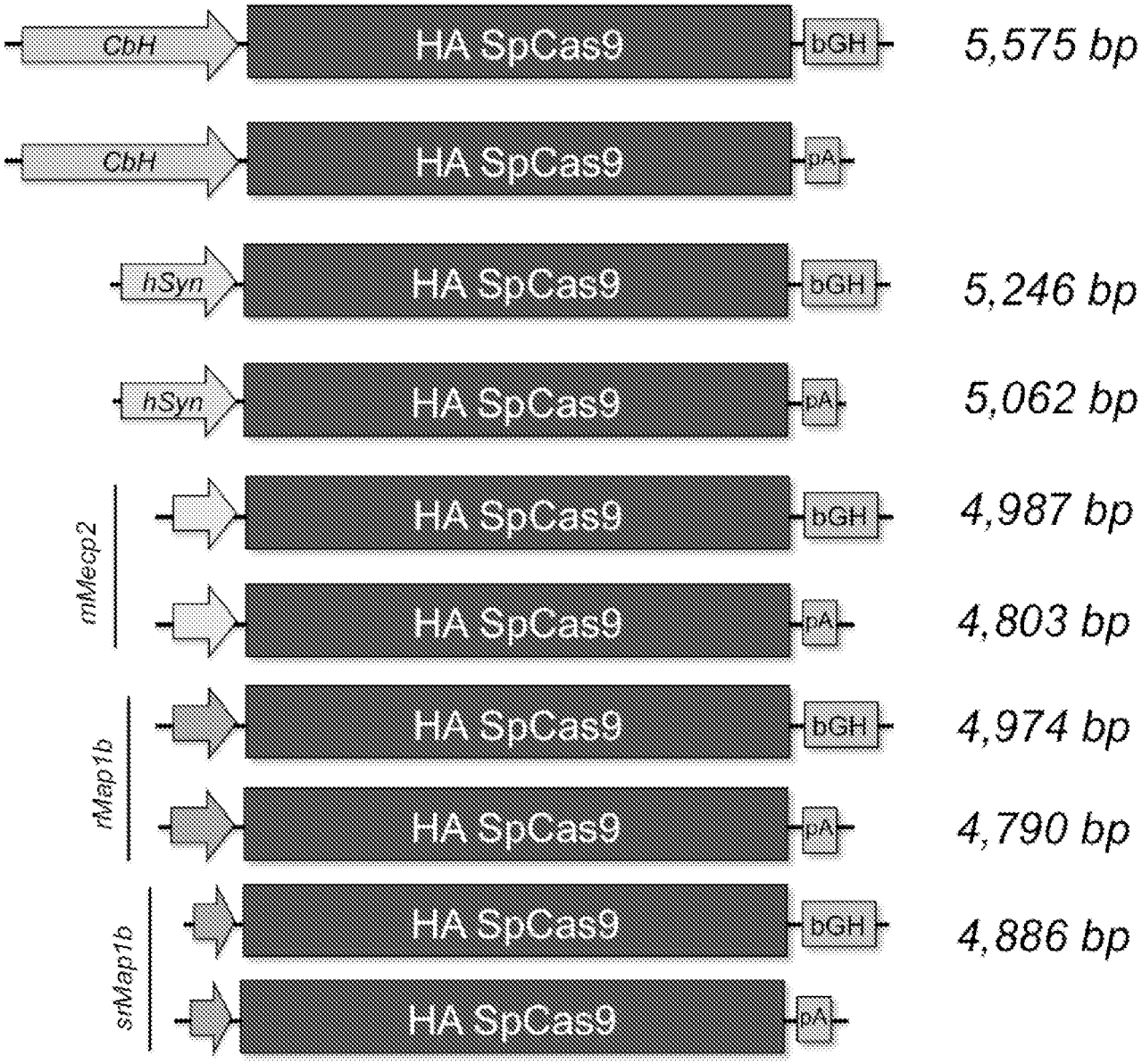

CbH
HA SpCas9
bGH
5,575 bp
CbH
HA SpCas9
pA
hSyn
HA SpCas9
bGH
5,246 bp
hSyn
HA SpCas9
pA
5,062 bp
mMecp2
HA SpCas9
bGH
4,987 bp
HA SpCas9
pA
4,803 bp
rMap1b
HA SpCas9
bGH
4,974 bp
HA SpCas9
pA
4,790 bp
smMap1b
HA SpCas9
bGH
4,886 bp
HA SpCas9
pA

FIG. 32

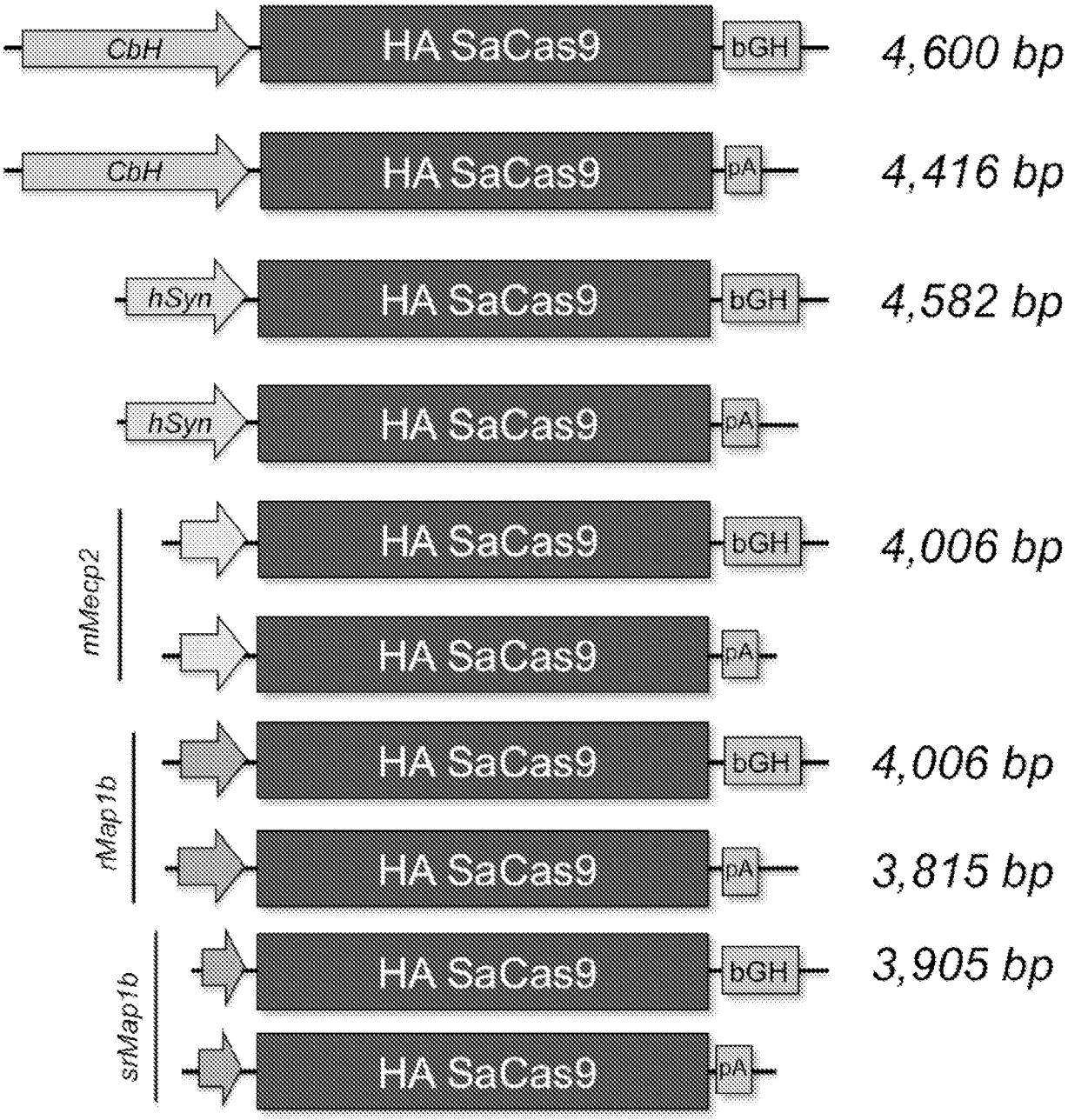

CbH
HA SaCas9
bGH
4,600 bp
CbH
HA SaCas9
pA
4,416 bp
hSyn
HA SaCas9
bGH
4,582 bp
hSyn
HA SaCas9
pA
mMecp2
HA SaCas9
bGH
4,006 bp
HA SaCas9
pA
rMap1b
HA SaCas9
bGH
4,006 bp
HA SaCas9
pA
3,815 bp
srMap1b
HA SaCas9
bGH
3,905 bp
HA SaCas9
pA

FIG. 33

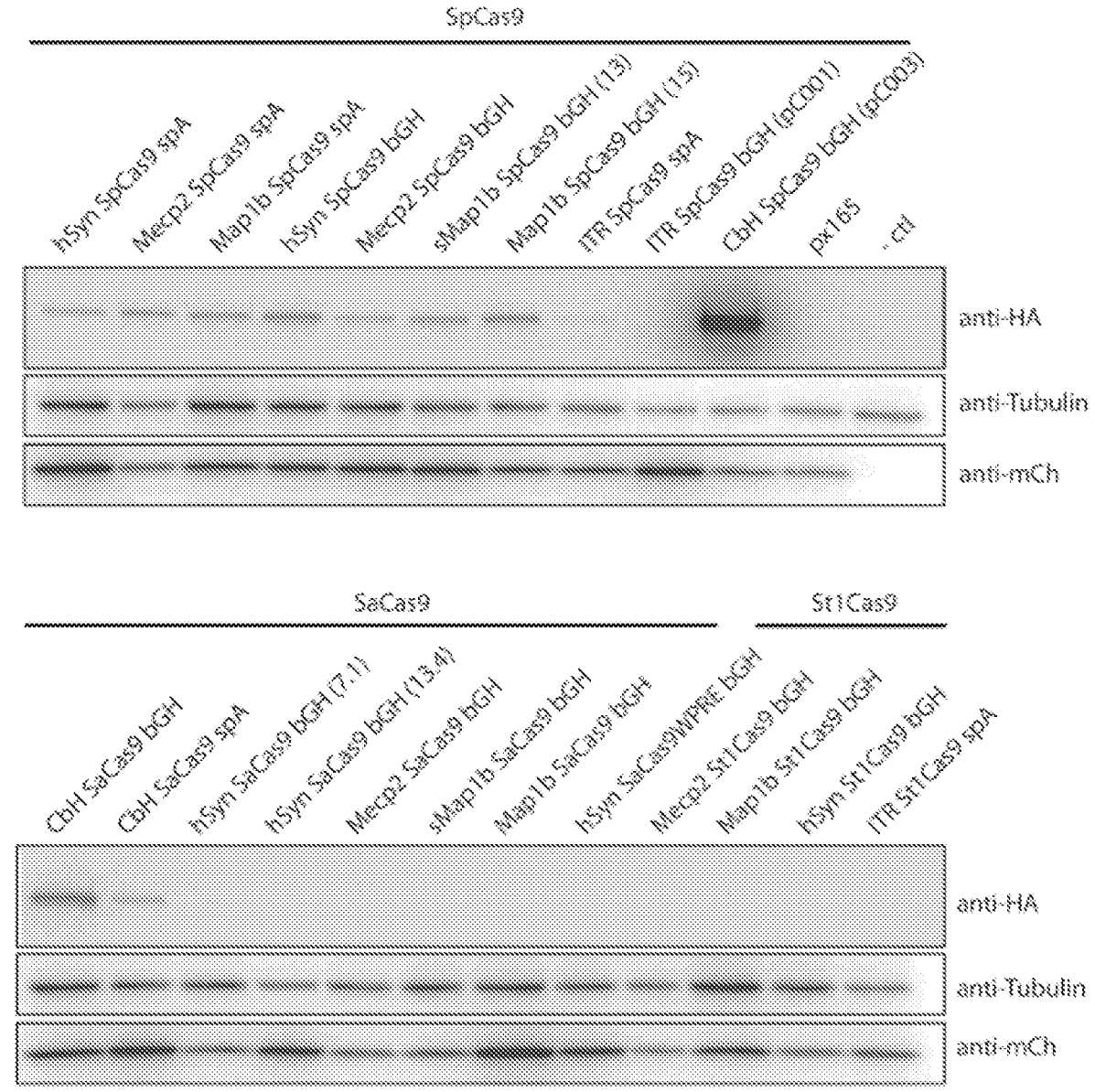

Expression of SpCas9 & SaCas9 in N2a cells
SpCas9
hSyn SpCas9 spA
Mecp2 SpCas9 spA
Map1b SpCas9 spA
hSyn SpCas9 bGH
Mecp2 SpCas9 bGH
sMap1b SpCas9 bGH (13)
Map1b SpCas9 bGH (15)
ITR SpCas9 bGH
ITR SpCas9 bGH (pC001)
CbH SpCas9 bGH (pC003)
px165
- ctl
anti-HA
anti-Tubulin
anti-mCh
SaCas9
St1Cas9
CbH SaCas9 bGH
CbH SaCas9 spA
hSyn SaCas9 bGH (7.1)
hSyn SaCas9 bGH (13.4)
Mecp2 SaCas9 bGH
sMap1b SaCas9 bGH
Map1b SaCas9 bGH
hSyn SaCas9WPRE bGH
Mecp2 St1Cas9 bGH
Map1b St1Cas9 bGH
hSyn St1Cas9 bGH
ITR St1Cas9 spA
anti-HA
anti-Tubulin
anti-mCh

FIG. 34

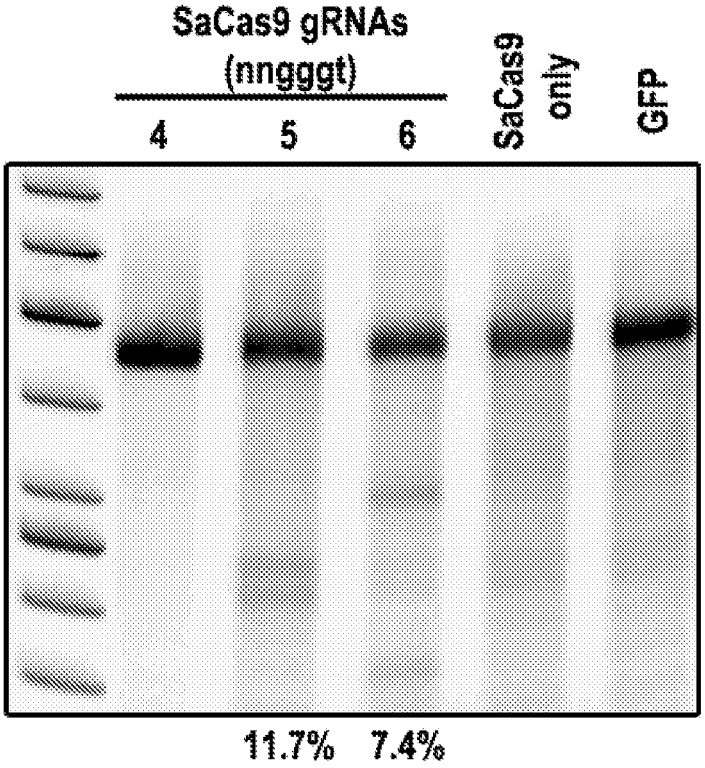


FIG. 35

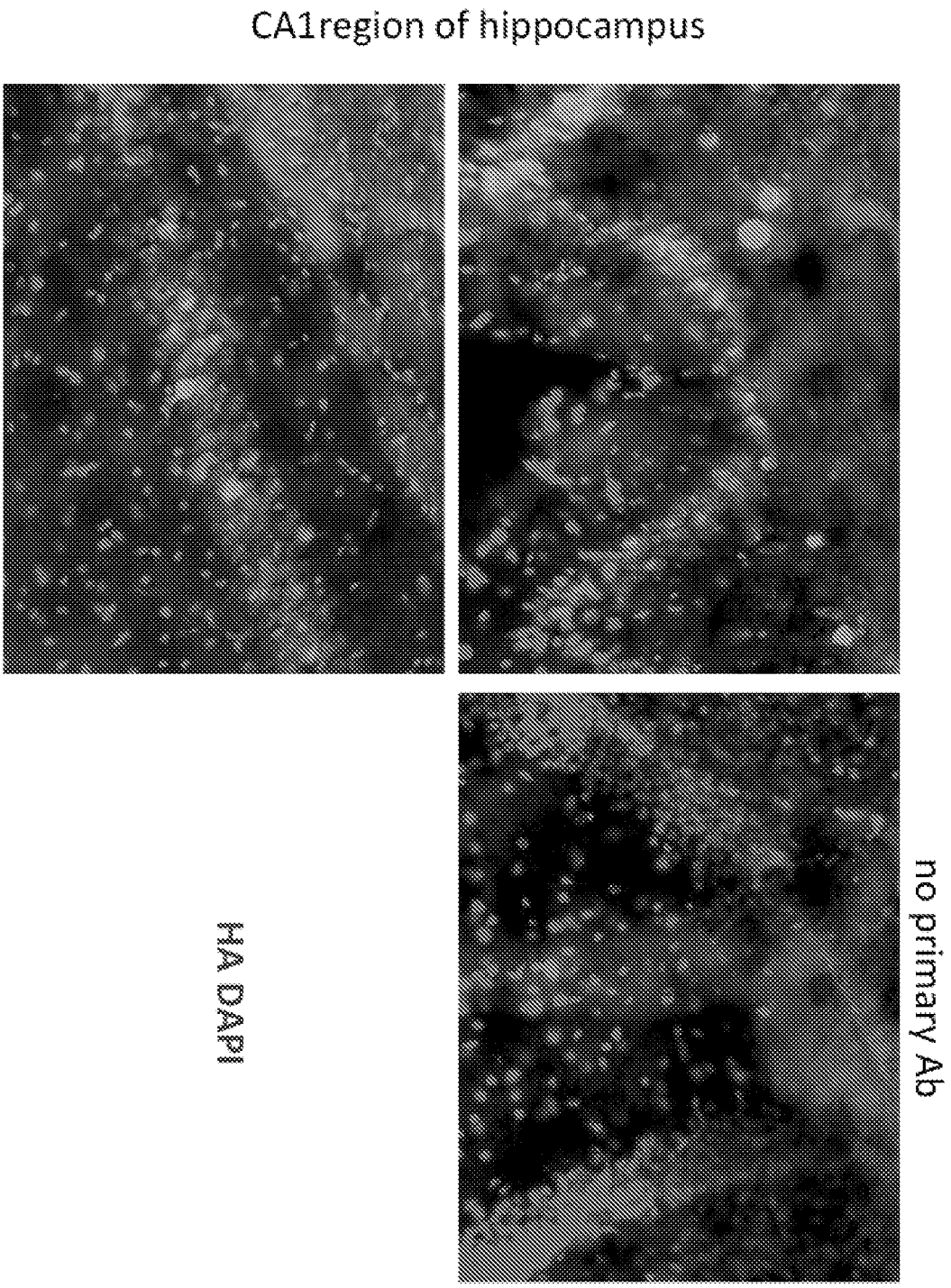


FIG. 36

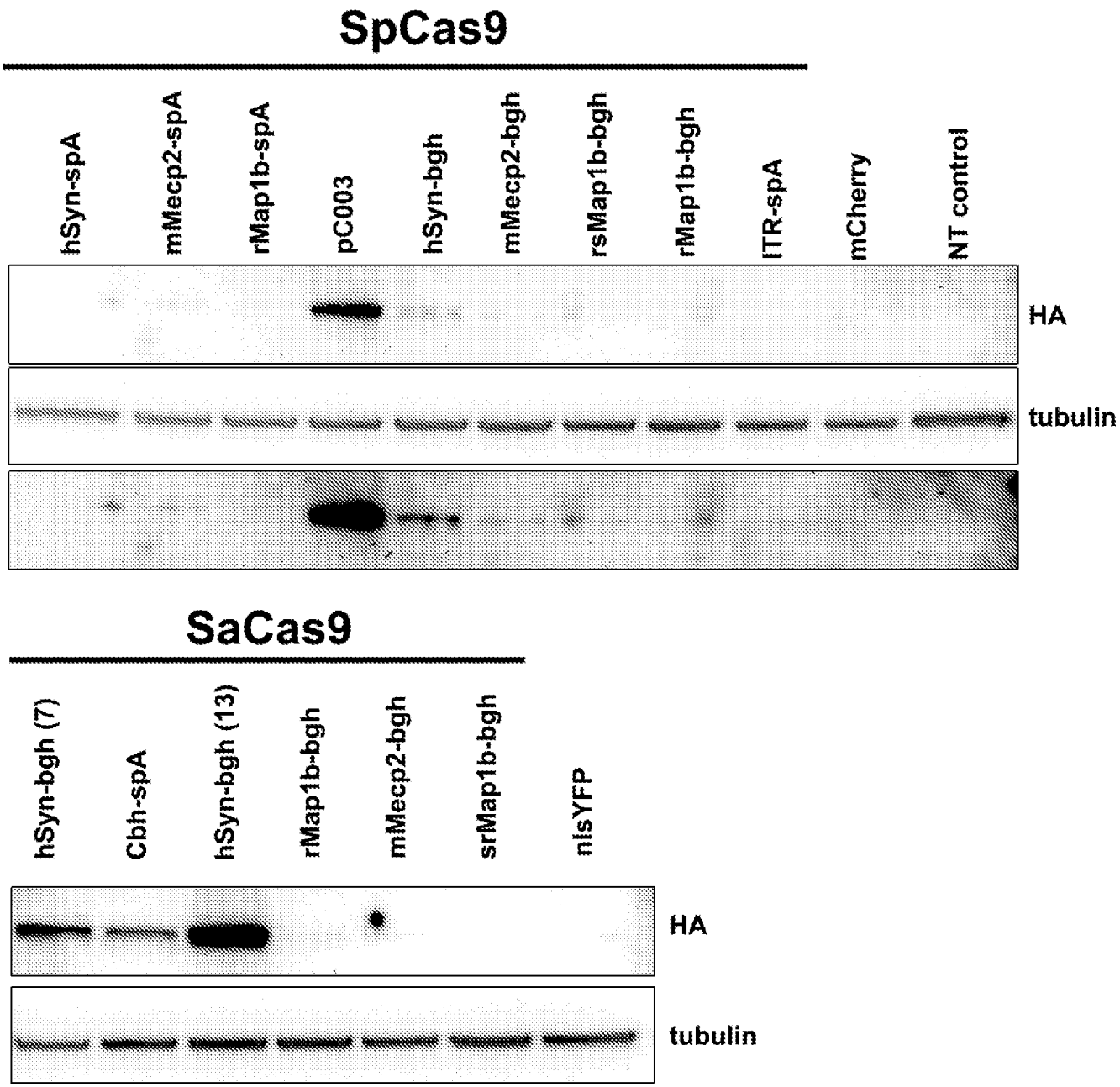

SpCas9
hSyn-spA
mMecp2-spA
rMap1b-spA
pC003
hSyn-bgh
mMecp2-bgh
rsMap1b-bgh
rMap1b-bgh
ITR-spA
mCherry
NT control
HA
tubulin
SaCas9
hSyn-bgh (7)
Cbh-spA
hSyn-bgh (13)
rMap1b-bgh
mMecp2-bgh
srMap1b-bgh
nlsYFP
HA
tubulin

FIG. 37

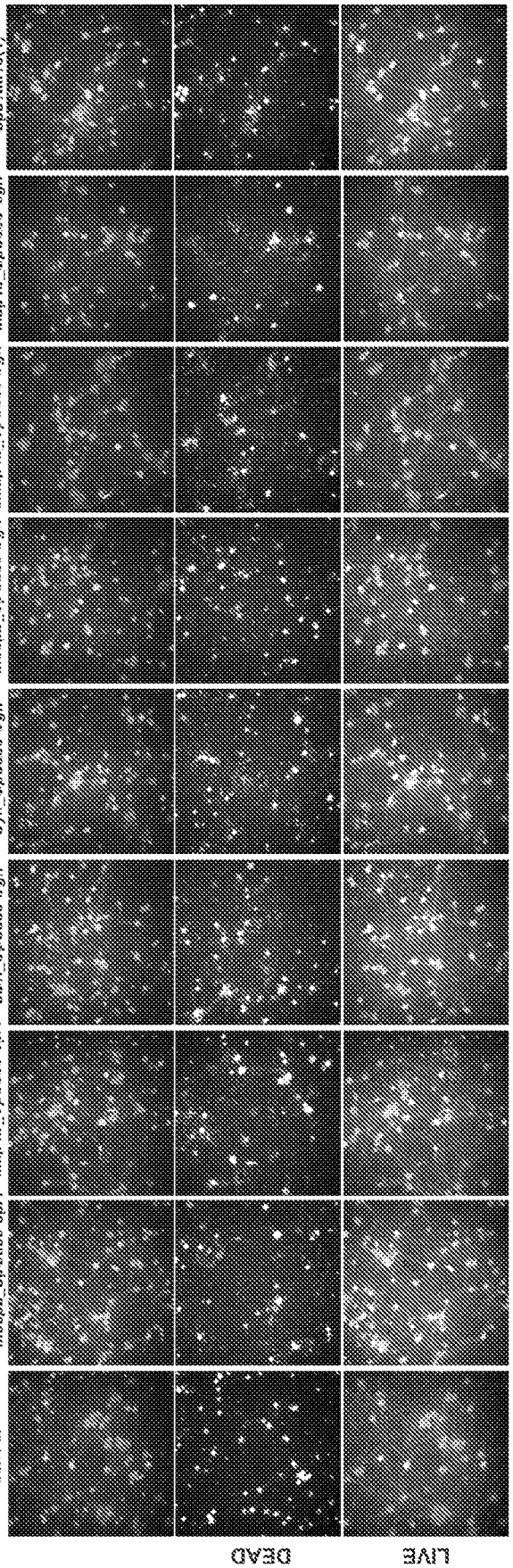

control
Mecp2_SpCas9-spA
Map1b_SpCas9-spA
CBH_SpCas9-bgh
Syn_SpCas9-bgh
Mecp2_SpCas9-bgh
sMap1b_SpCas9-bgh
Map1b_SpCas9-bgh
SpDNMTs(1)
DEAD
LIVE

FIG. 38

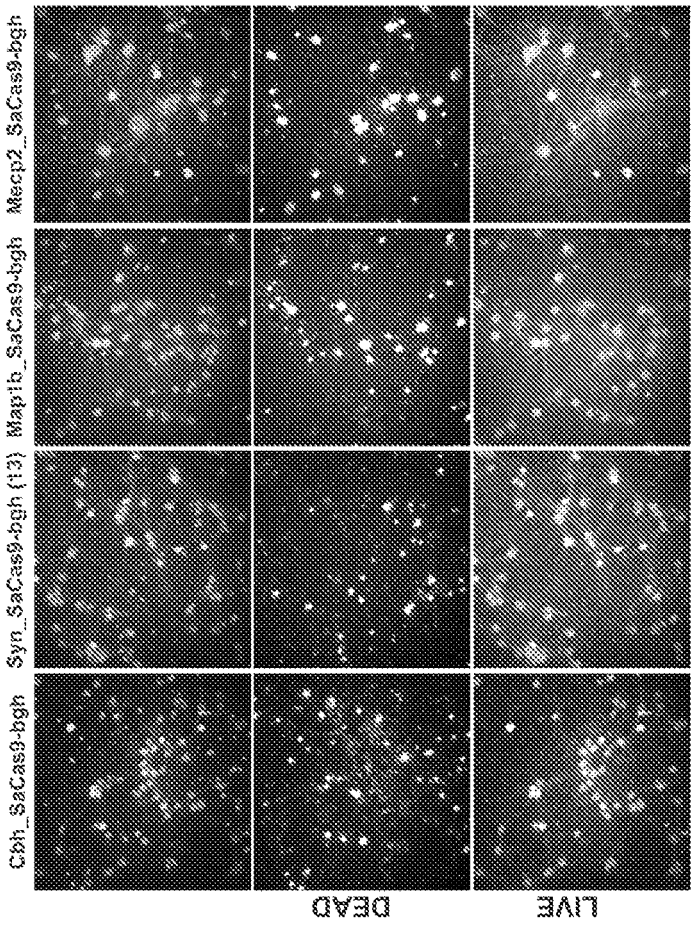

Cbh_SaCas9-bgh
Syn_SaCas9-bgh (13)
Map1b_SaCas9-bgh
Mecp2_SaCas9-bgh
DEAD
LIVE

FIG. 39

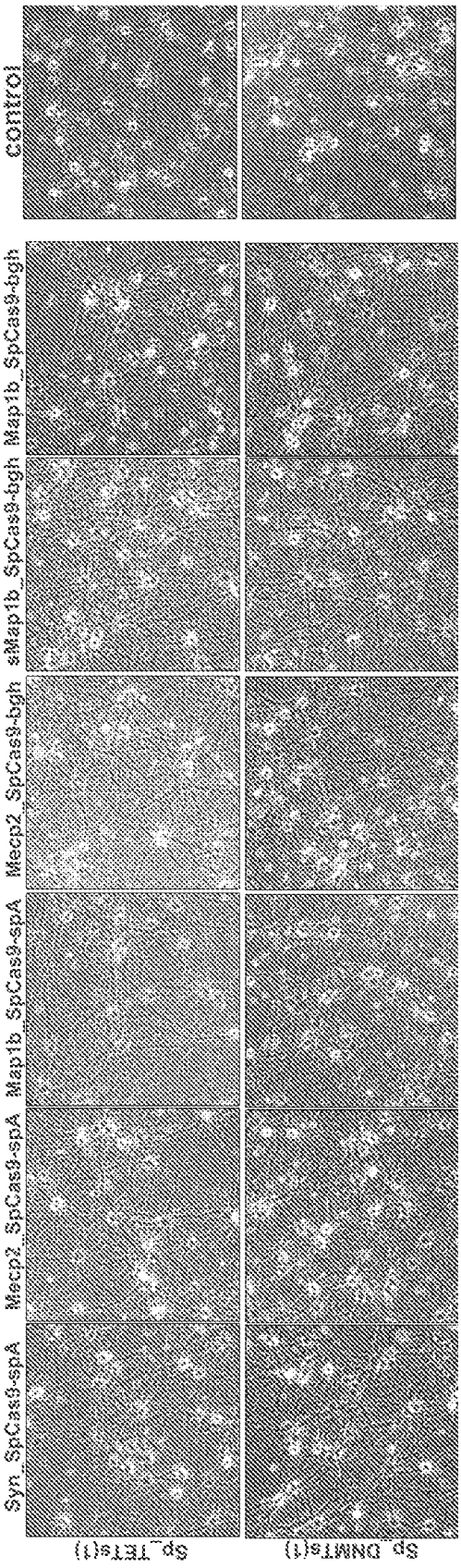

control
Map1b_SpCas9-bgh
sMap1b_SpCas9-bgh
Mecp2_SpCas9-bgh
Map1b_SpCas9-spA
Mecp2_SpCas9-spA
Syn_SpCas9-spA
Sp_DNMTs(1)
Sp_TETs(1)

FIG. 40

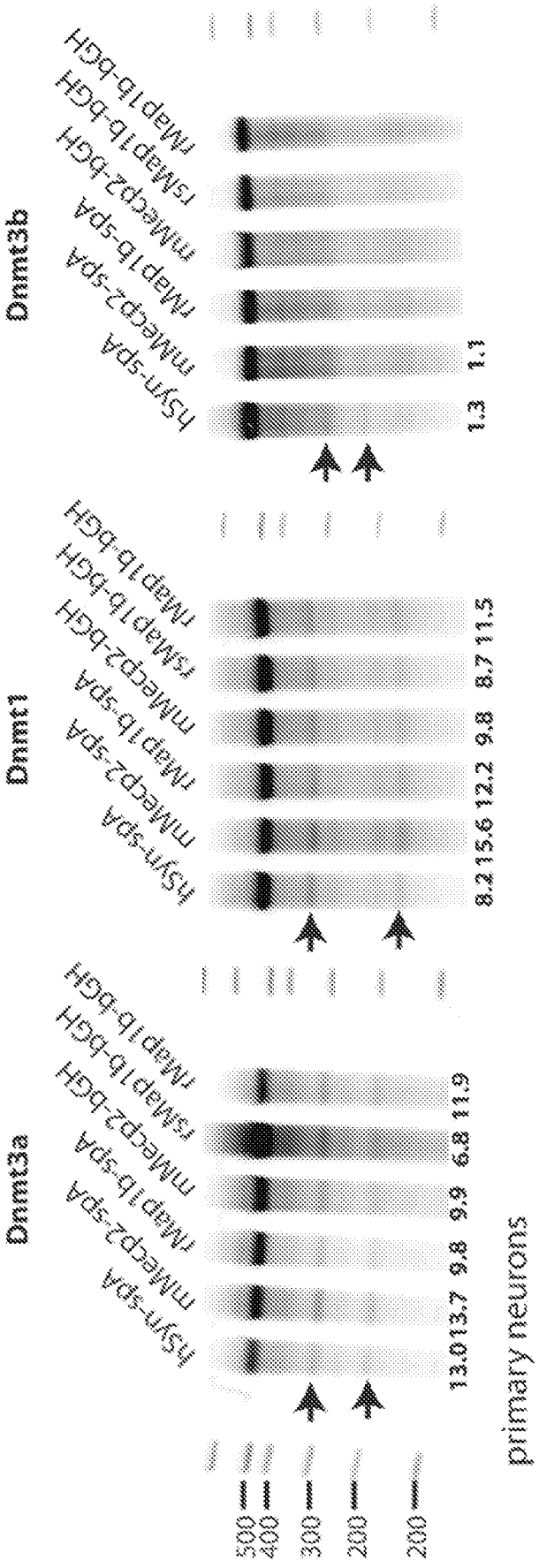

Dnmt3a
Dnmt1
Dnmt3b
hSyn-sPA
mMeCP2-sPA
mMap1b-sPA
rSMap1b-bGH
rMeCP2-bGH
rMap1b-bGH
13.0 13.7 9.8 9.9 6.8 11.9
8.2 15.6 12.2 9.8 8.7 11.5
1.3 1.1
500
400
300
200
200
primary neurons

FIG. 41

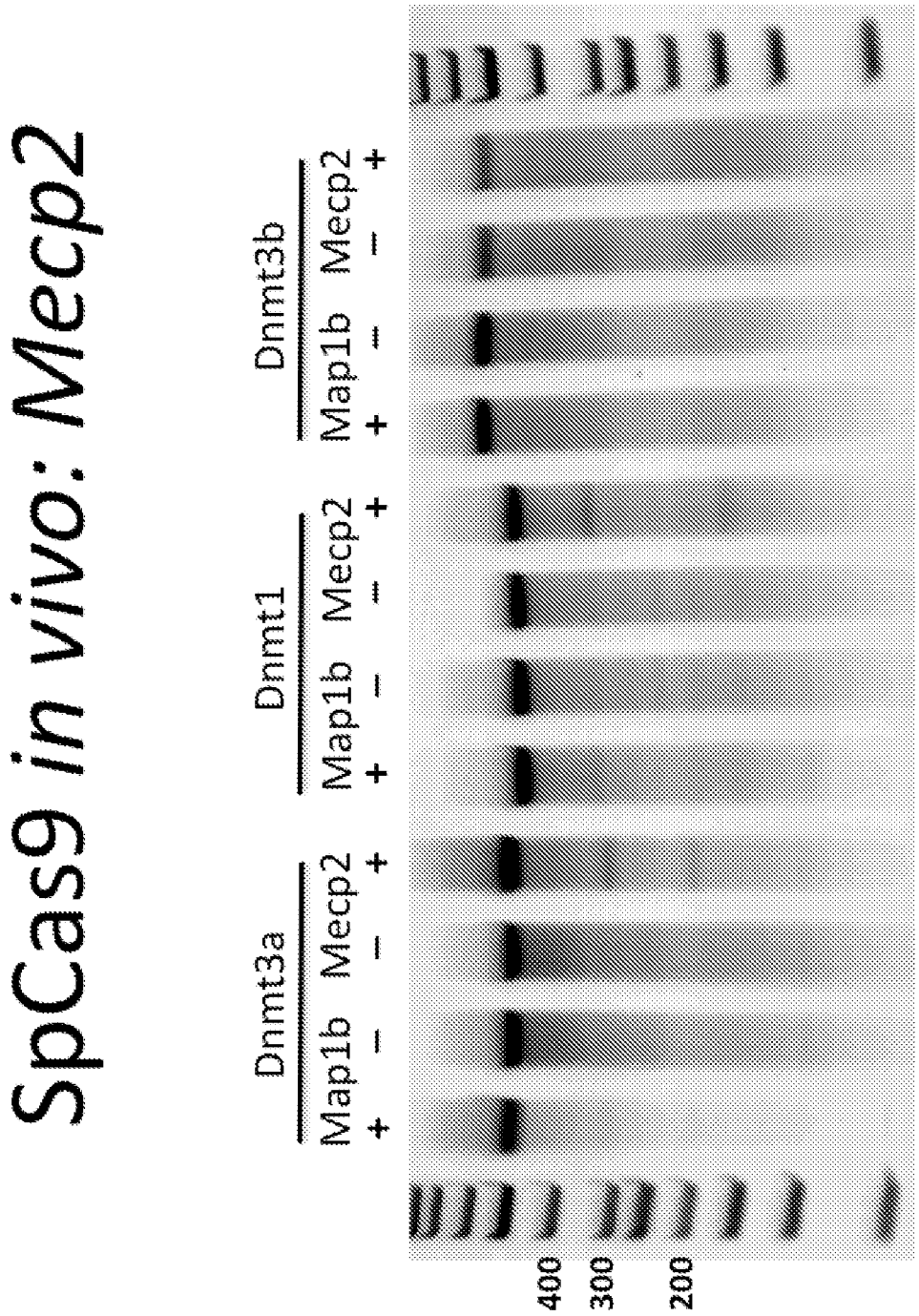

SpCas9 in vivo: Mecp2
Dnmt3a
Map1b + + Mecp2 – +
Dnmt1
Map1b + + Mecp2 – +
Dnmt3b
Map1b + + Mecp2 – +
400
300
200
12.0%
17.0%

FIG. 42

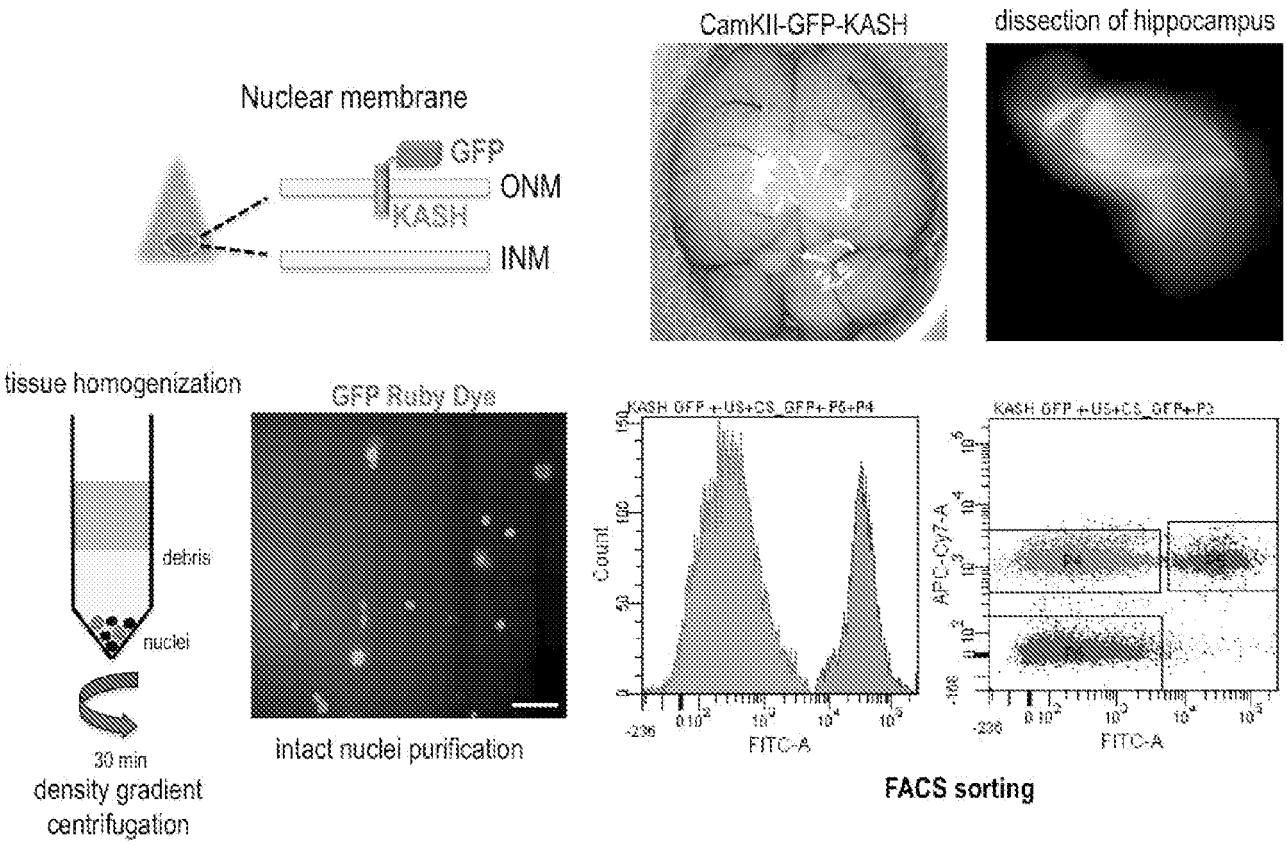

Purification of cell nuclei from brain
Nuclear membrane
GFP
ONM
KASH
INM
CamKII-GFP-KASH
dissection of hippocampus
tissue homogenization
debris
nuclei
30 min
density gradient centrifugation
GFP Ruby Dye
intact nuclei purification
KASH GFP+US+CS_GFP+P6+P4
Count
FITC-A
KASH GFP+US+CS_GFP+P3
APC-Cy7-A
FITC-A
FACS sorting

FIG. 43

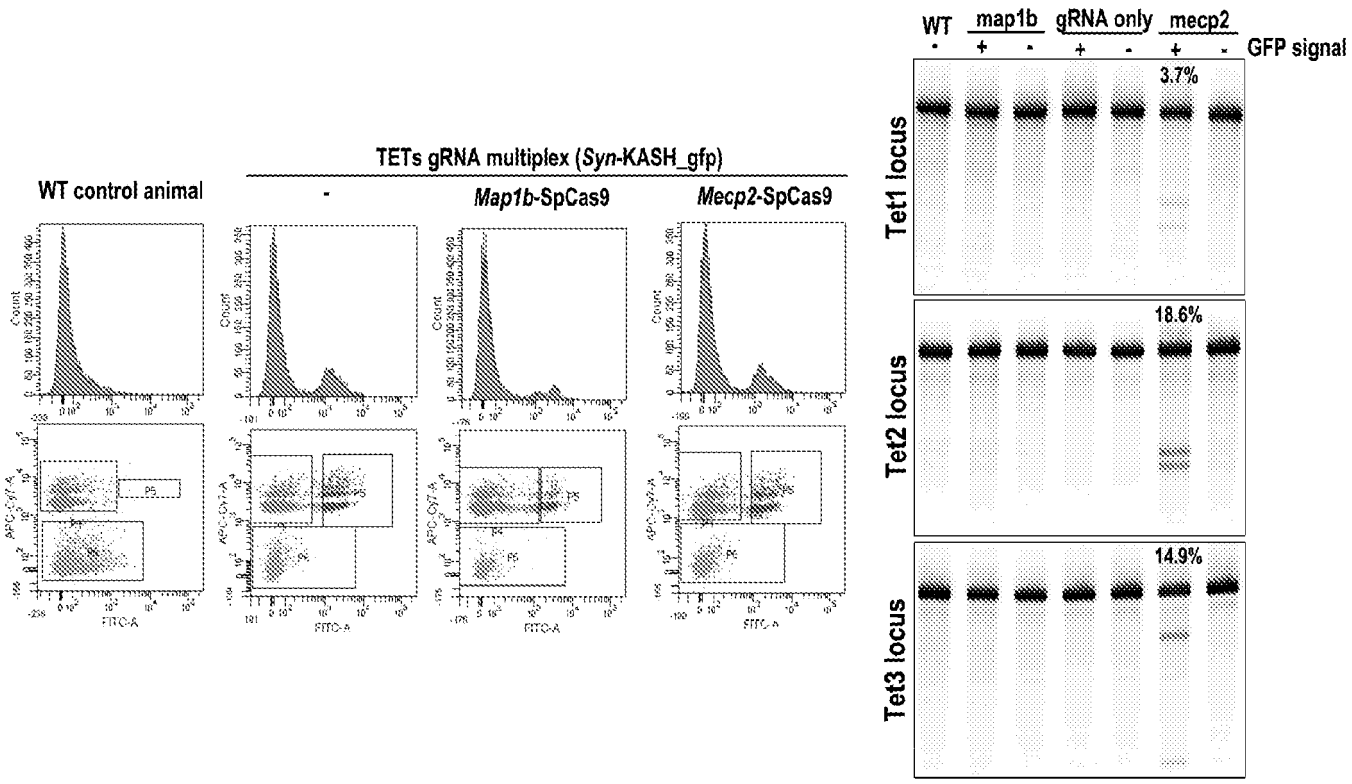

TETs gRNA multiplex (Syn-KASH_gfp)
WT control animal
-
Map1b-SpCas9
Mecp2-SpCas9
WT
map1b
gRNA only
mecp2
GFP signal
Tet1 locus
Tet2 locus
Tet3 locus
3.7%
18.6%
14.9%

FIG. 44

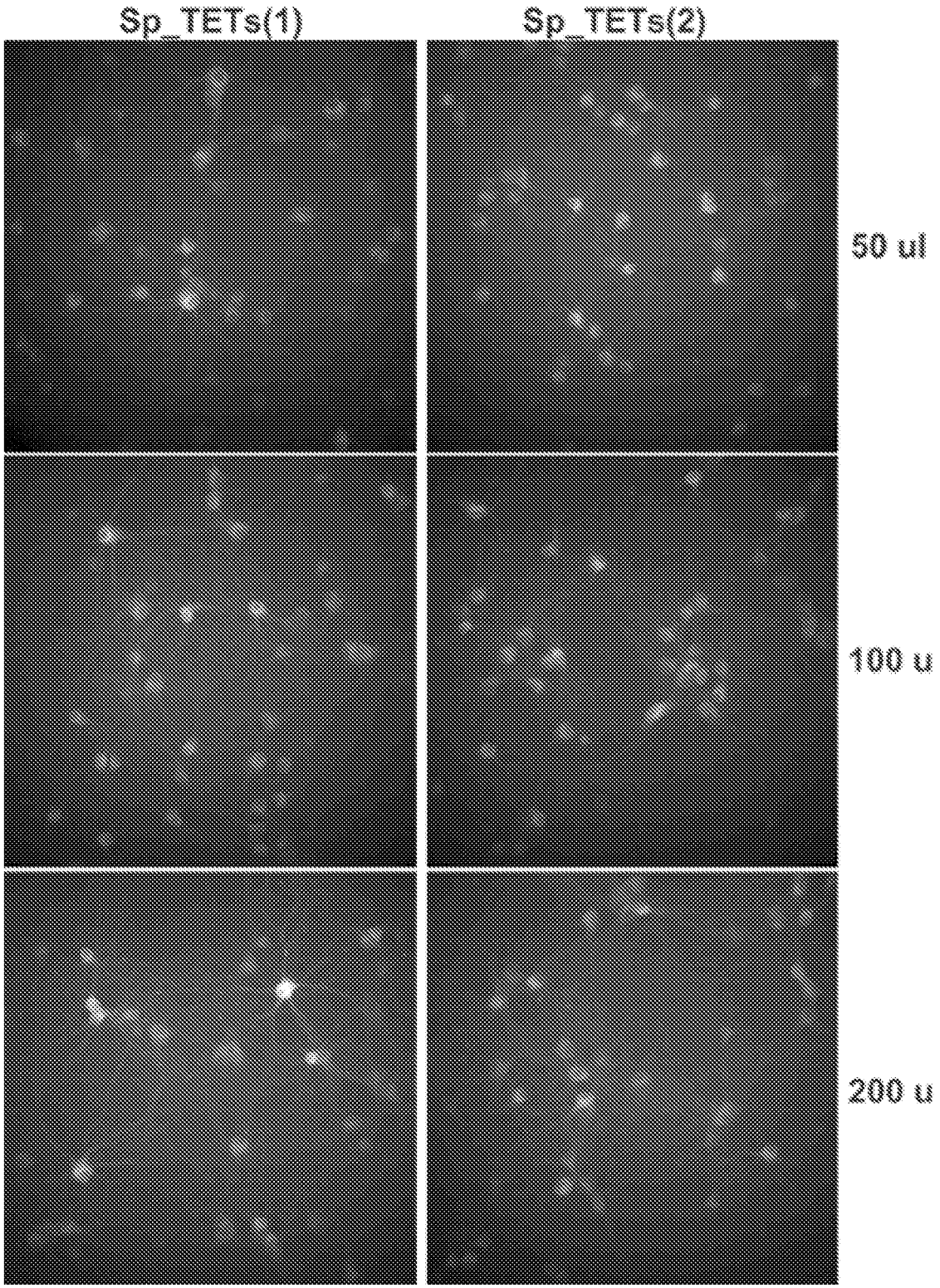


FIG. 45

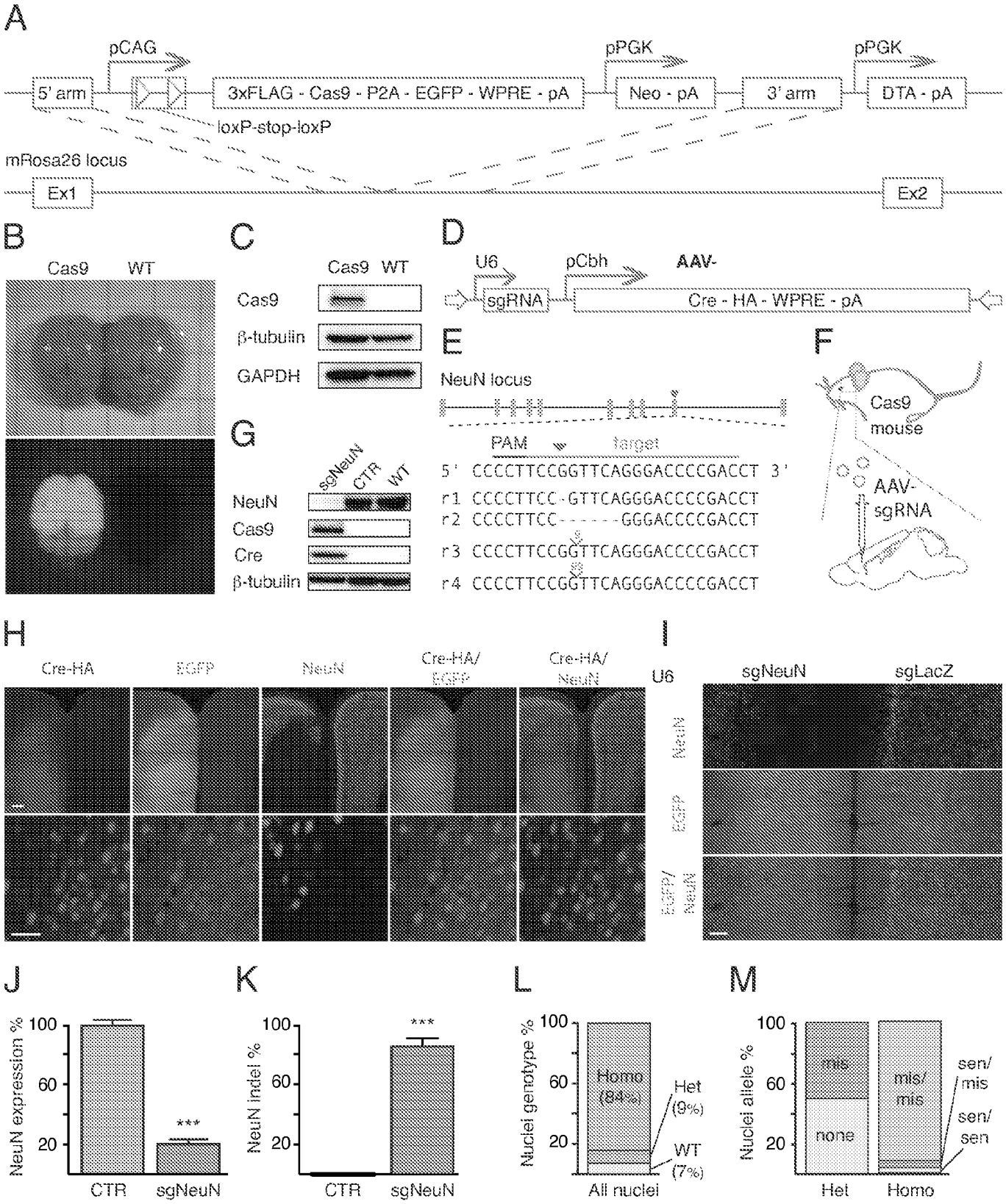

A
pCAG
pPGK
pPGK
5' arm
3xFLAG - Cas9 - P2A - EGFP - WPRE - pA
Neo - pA
3' arm
DTA - pA
loxP-stop-loxP
mRosa26 locus
Ex1
Ex2
B
Cas9
WT
C
Cas9 WT
Cas9
β-tubulin
GAPDH
D
U6
pCbh
AAV-
sgRNA
Cre - HA - WPRE - pA
E
NeuN locus
PAM
target
5' CCCCTTCCGGTTCAGGGACCCCGACCT 3'
r1 CCCCTTCC-GTTCAGGGACCCCGACCT
r2 CCCCTTCC------GGGACCCCGACCT
r3 CCCCTTCCGGTTCAGGGACCCCGACCT
r4 CCCCTTCCGGTTCAGGGACCCCGACCT
F
Cas9 mouse
AAV-sgRNA
G
sgNeuN
CTR
WT
NeuN
Cas9
Cre
β-tubulin
H
Cre-HA
EGFP
NeuN
Cre-HA/EGFP
Cre-HA/NeuN
I
U6
sgNeuN
sgLacZ
NeuN
EGFP
EGFP/NeuN
J
NeuN expression %
100
60
20
***
CTR
sgNeuN
K
NeuN indel %
100
60
20
***
CTR
sgNeuN
L
Nuclei genotype %
100
60
20
Homo (84%)
Het (9%)
WT (7%)
All nuclei
M
Nuclei allele %
100
60
20
mis
none
mis/mis
sen/mis
sen/sen
Het
Homo

FIG. 46A-M

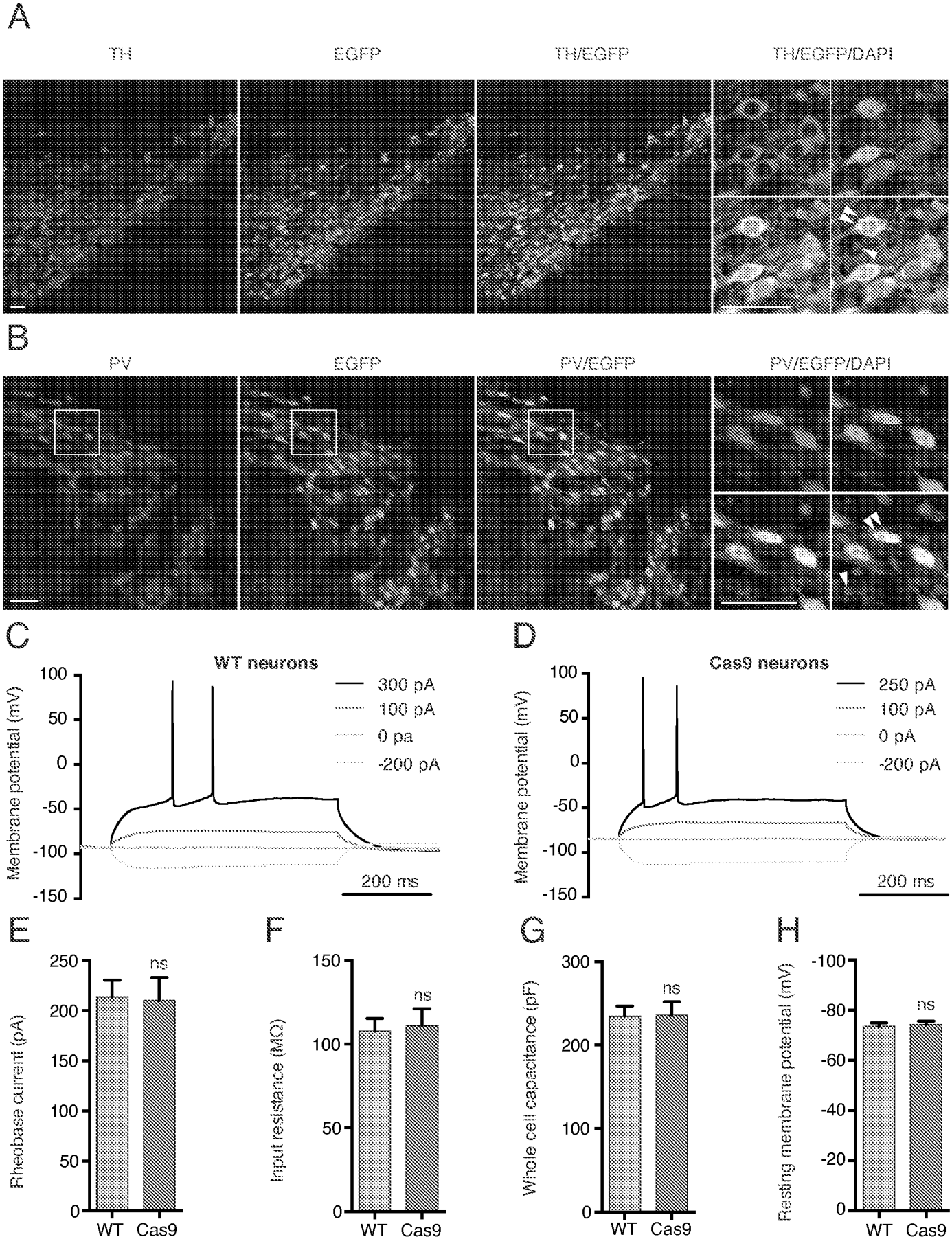

A
TH
EGFP
TH/EGFP
TH/EGFP/DAPI
B
PV
EGFP
PV/EGFP
PV/EGFP/DAPI
C
WT neurons
Membrane potential (mV)
300 pA
100 pA
0 pa
-200 pA
200 ms
D
Cas9 neurons
Membrane potential (mV)
250 pA
100 pA
0 pA
-200 pA
200 ms
E
Rheobase current (pA)
ns
WT
Cas9
F
Input resistance (MΩ)
ns
WT
Cas9
G
Whole cell capacitance (pF)
ns
WT
Cas9
H
Resting membrane potential (mV)
ns
WT
Cas9

FIG. 47A-H

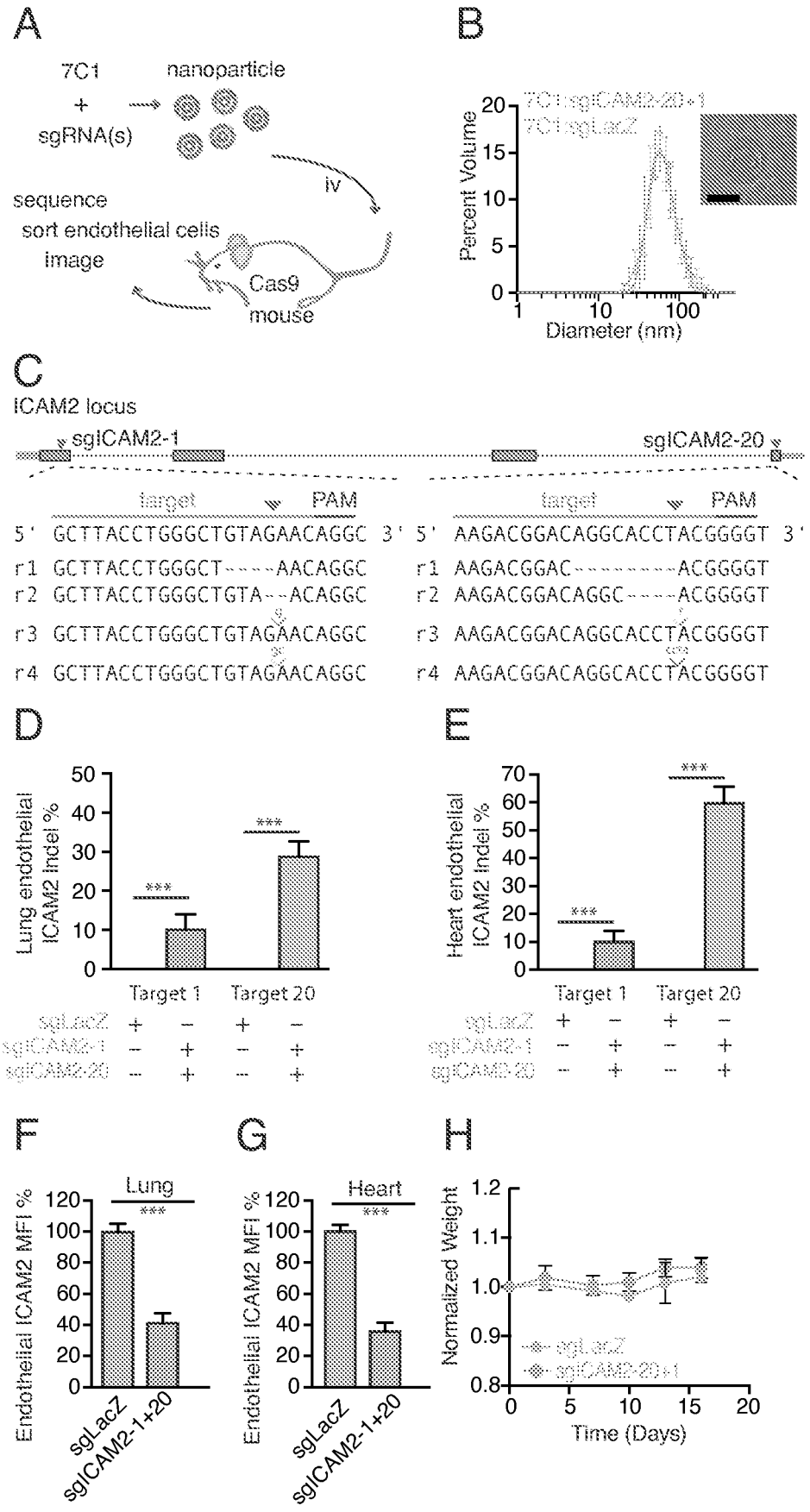

A
7C1
+
sgRNA(s)
nanoparticle
iv
sequence
sort endothelial cells
image
Cas9
mouse
B
7C1:sgICAM2-20+1
7C1:sgLacZ
Percent Volume
Diameter (nm)
C
ICAM2 locus
sgICAM2-1
sgICAM2-20
target
PAM
5' GCTTACCTGGGCTGTAGAACAGGC 3'
r1 GCTTACCTGGGCT----AACAGGC
r2 GCTTACCTGGGCTGTA--ACAGGC
r3 GCTTACCTGGGCTGTAGAACAGGC
r4 GCTTACCTGGGCTGTAGAACAGGC
5' AAGACGGACAGGCACCTACGGGGT 3'
r1 AAGACGGAC---------ACGGGGT
r2 AAGACGGACAGGC-----ACGGGGT
r3 AAGACGGACAGGCACCTACGGGGT
r4 AAGACGGACAGGCACCTACGGGGT
D
Lung endothelial ICAM2 Indel %
Target 1
Target 20
sgLacZ
sgICAM2-1
sgICAM2-20
E
Heart endothelial ICAM2 Indel %
Target 1
Target 20
sgLacZ
sgICAM2-1
sgICAM2-20
F
Lung
Endothelial ICAM2 MFI %
sgLacZ
sgICAM2-1+20
G
Heart
Endothelial ICAM2 MFI %
sgLacZ
sgICAM2-1+20
H
Normalized Weight
sgLacZ
sgICAM2-20+1
Time (Days)

FIG. 48A-H

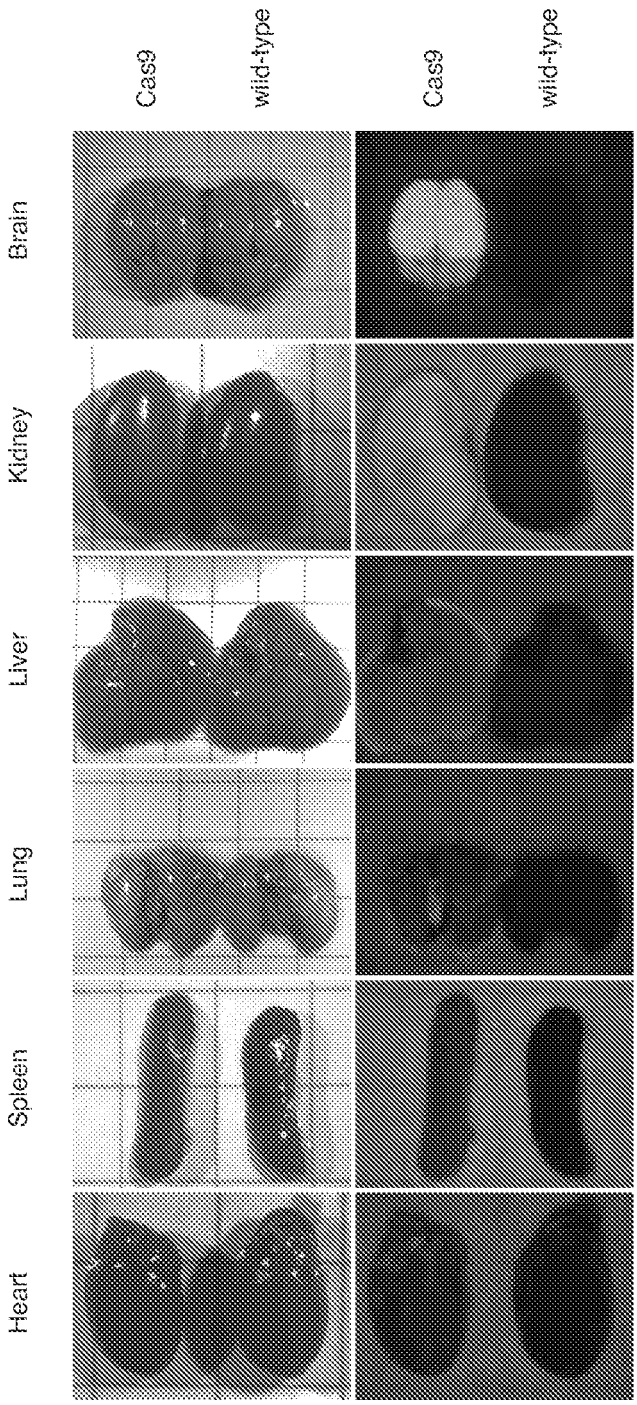

Cas9
wild-type
Cas9
wild-type
Brain
Kidney
Liver
Lung
Spleen
Heart

FIG. 49

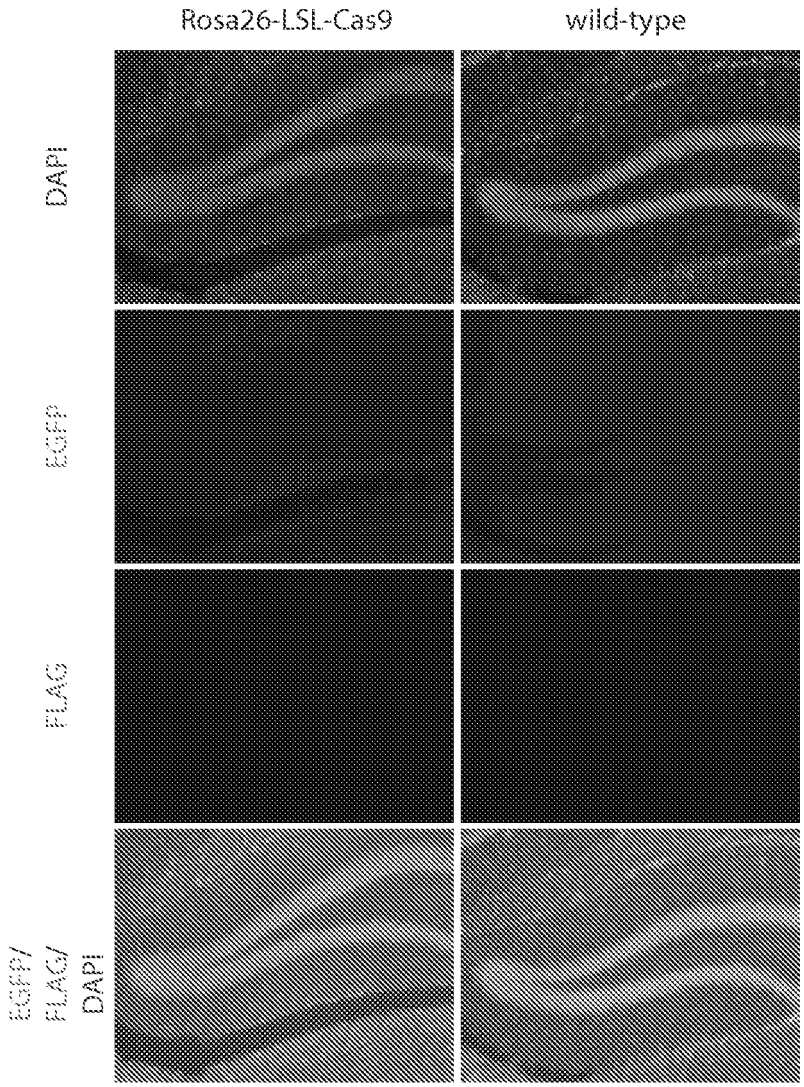


FIG. 50

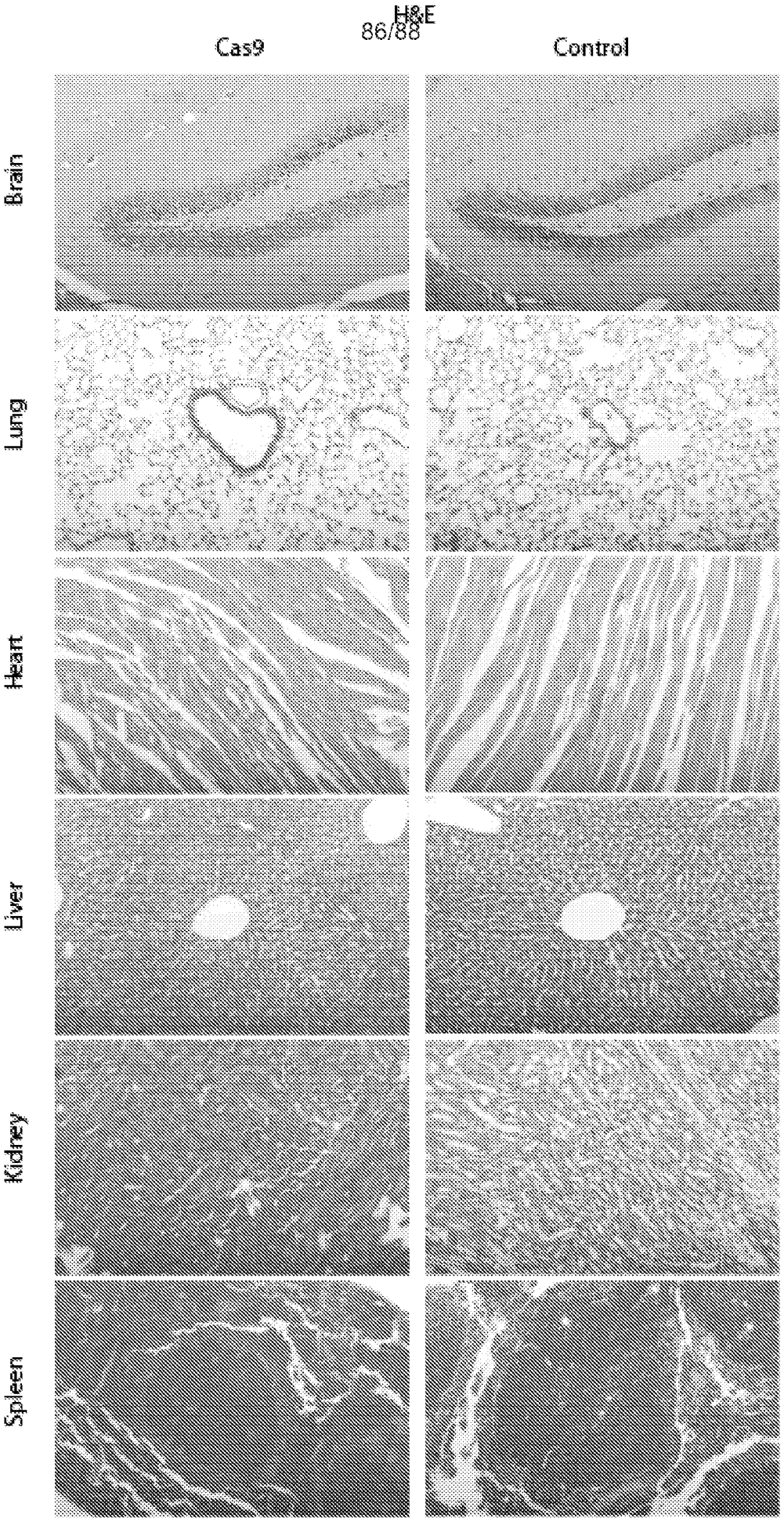


FIG. 51

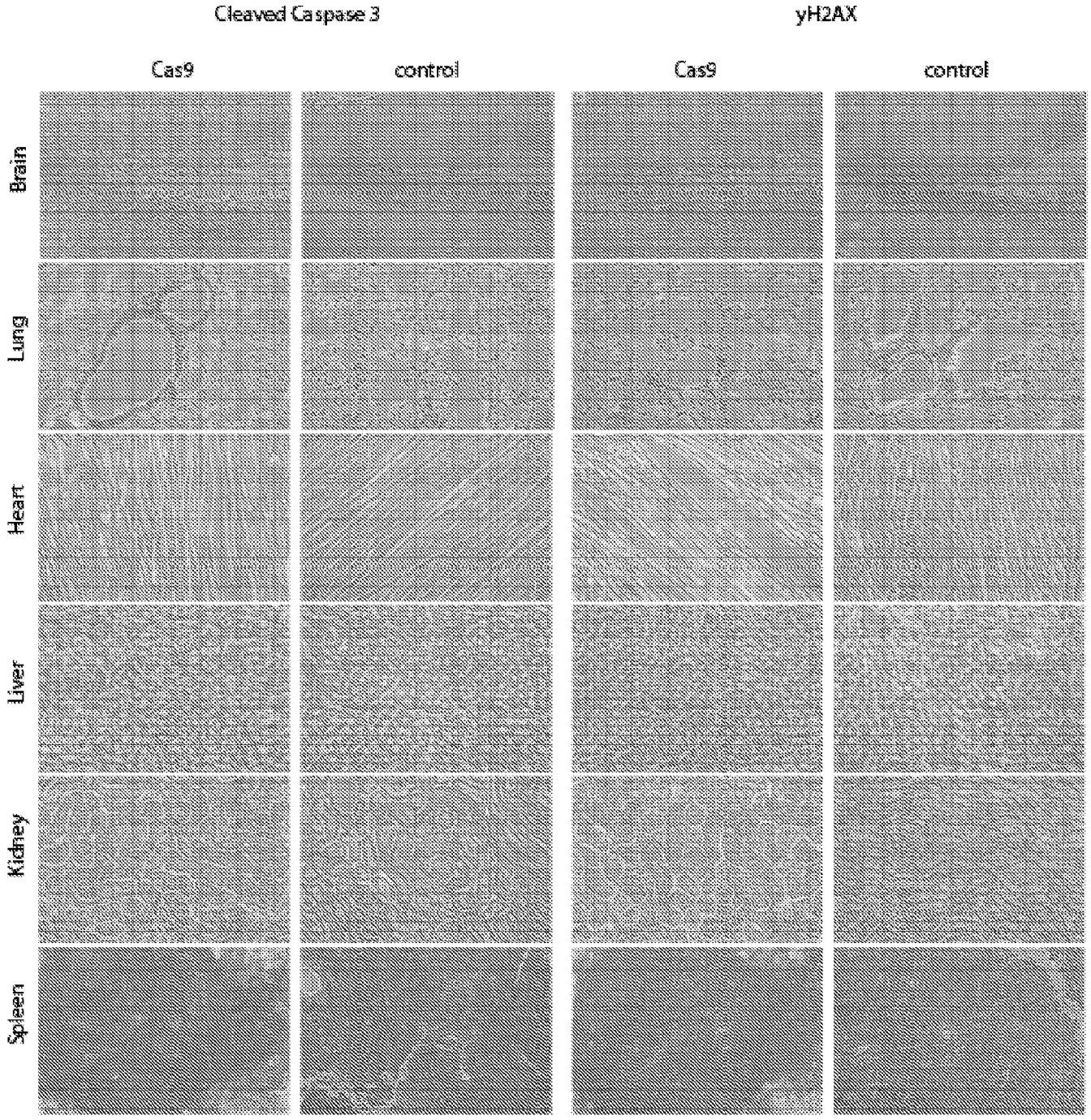

Cleaved Caspase 3
yH2AX
Cas9
control
Cas9
control
Brain
Lung
Heart
Liver
Kidney
Spleen

FIG. 52

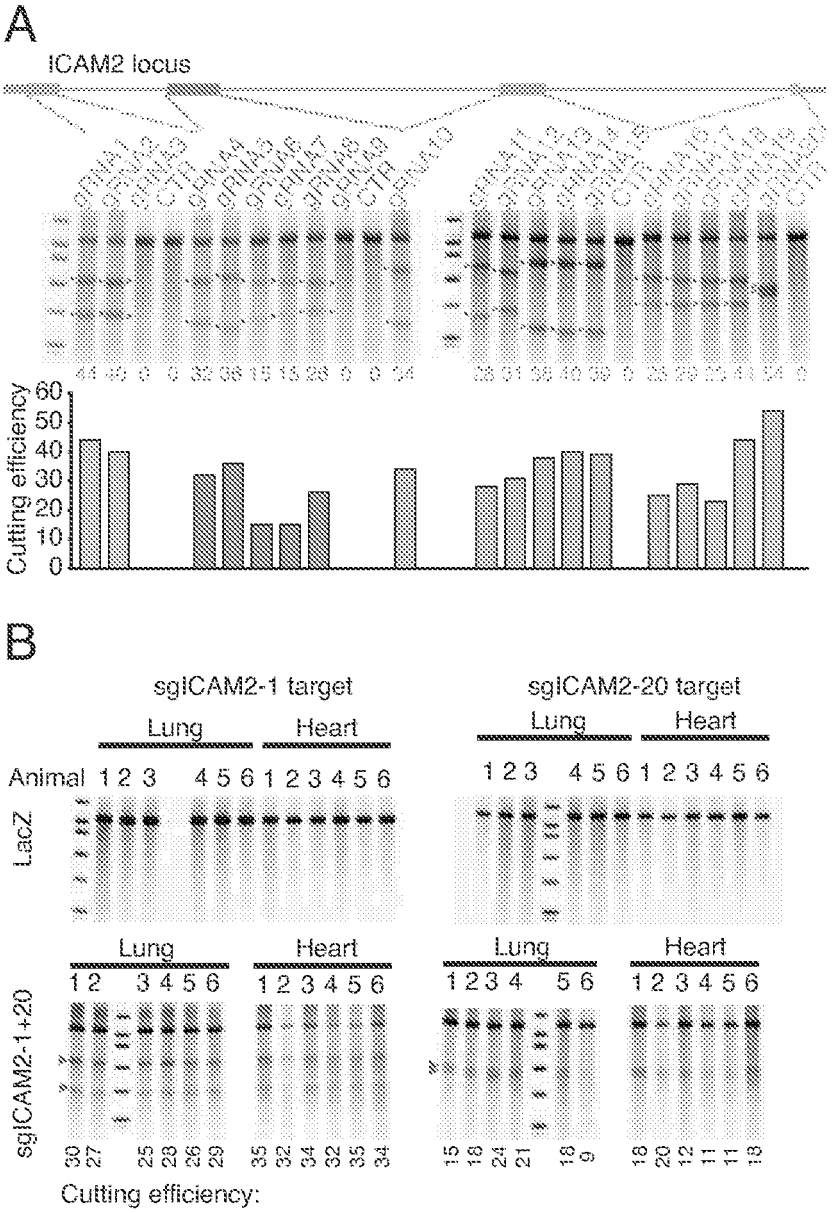

A
ICAM2 locus
gRNA1
gRNA2
gRNA3
CTR
gRNA4
gRNA5
gRNA6
gRNA7
gRNA8
gRNA9
CTR
gRNA10
gRNA11
gRNA12
gRNA13
gRNA14
gRNA15
CTR
gRNA16
gRNA17
gRNA18
gRNA19
CTR
Cutting efficiency
B
sgICAM2-1 target
sgICAM2-20 target
Lung
Heart
Animal 1 2 3 4 5 6 1 2 3 4 5 6
LacZ
sgICAM2-1+20
Cutting efficiency:

FIG. 53A-B