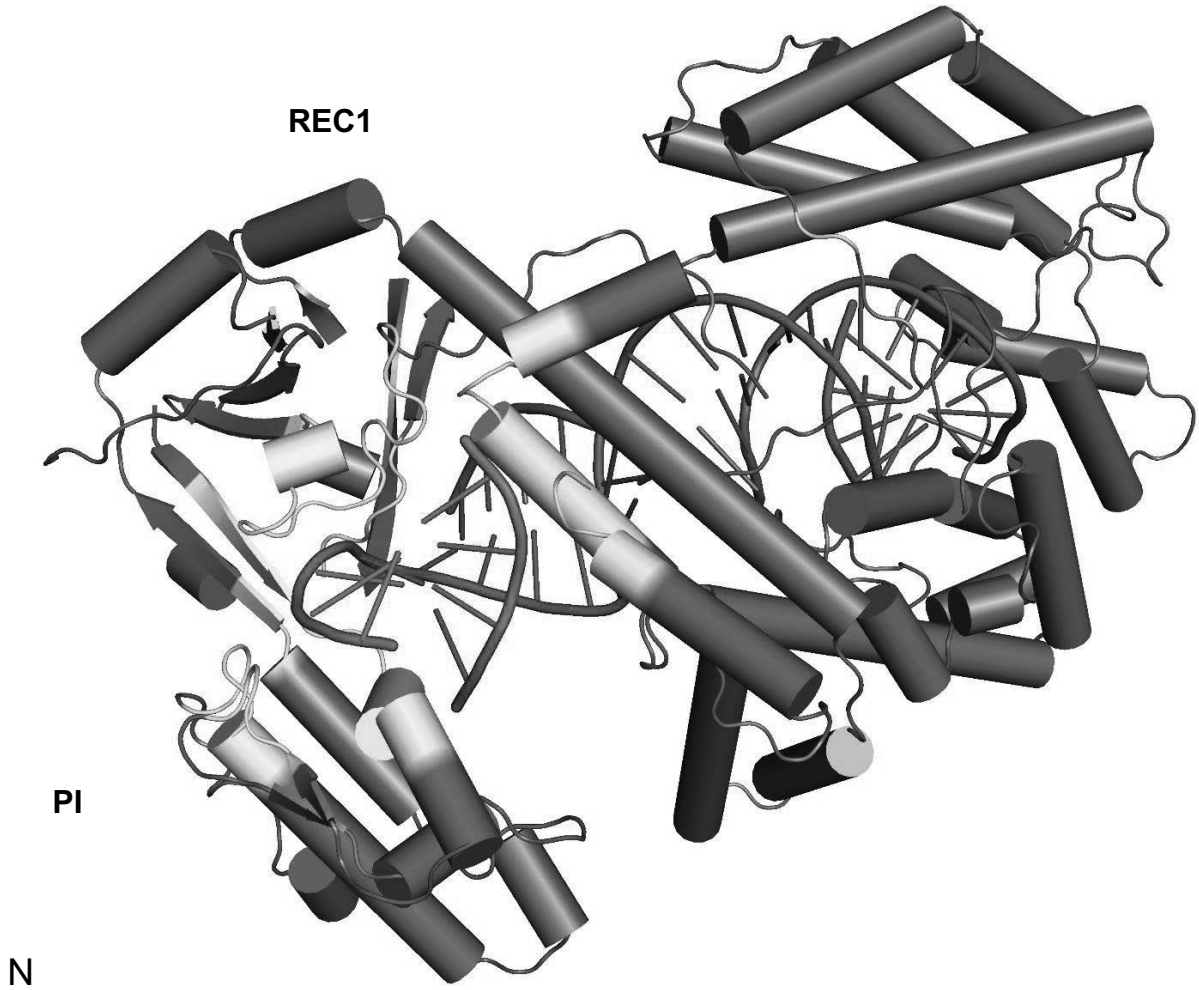


FIG. 1

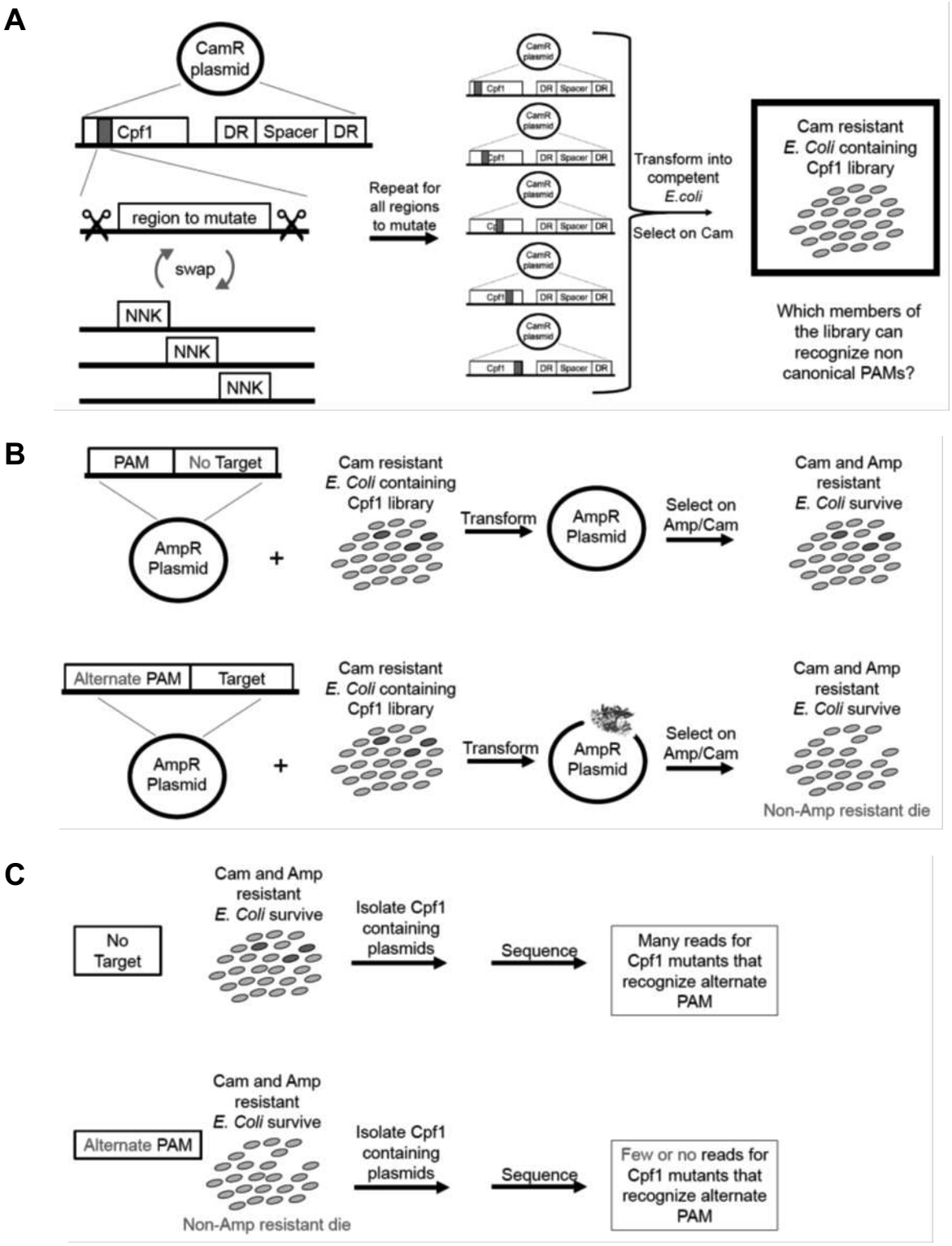

A
CamR plasmid
Cpf1
DR
Spacer
DR
region to mutate
swap
NNK
NNK
NNK
Repeat for all regions to mutate
Transform into competent E.coli
Select on Cam
Cam resistant E. Coli containing Cpf1 library
Which members of the library can recognize non canonical PAMs?
B
PAM
No Target
AmpR Plasmid
+
Cam resistant E. Coli containing Cpf1 library
Transform
AmpR Plasmid
Select on Amp/Cam
Cam and Amp resistant E. Coli survive
Alternate PAM
Target
AmpR Plasmid
+
Cam resistant E. Coli containing Cpf1 library
Transform
AmpR Plasmid
Select on Amp/Cam
Cam and Amp resistant E. Coli survive
Non-Amp resistant die
C
No Target
Cam and Amp resistant E. Coli survive
Isolate Cpf1 containing plasmids
Sequence
Many reads for Cpf1 mutants that recognize alternate PAM
Alternate PAM
Cam and Amp resistant E. Coli survive
Isolate Cpf1 containing plasmids
Sequence
Few or no reads for Cpf1 mutants that recognize alternate PAM
Non-Amp resistant die

FIG. 2A-2C

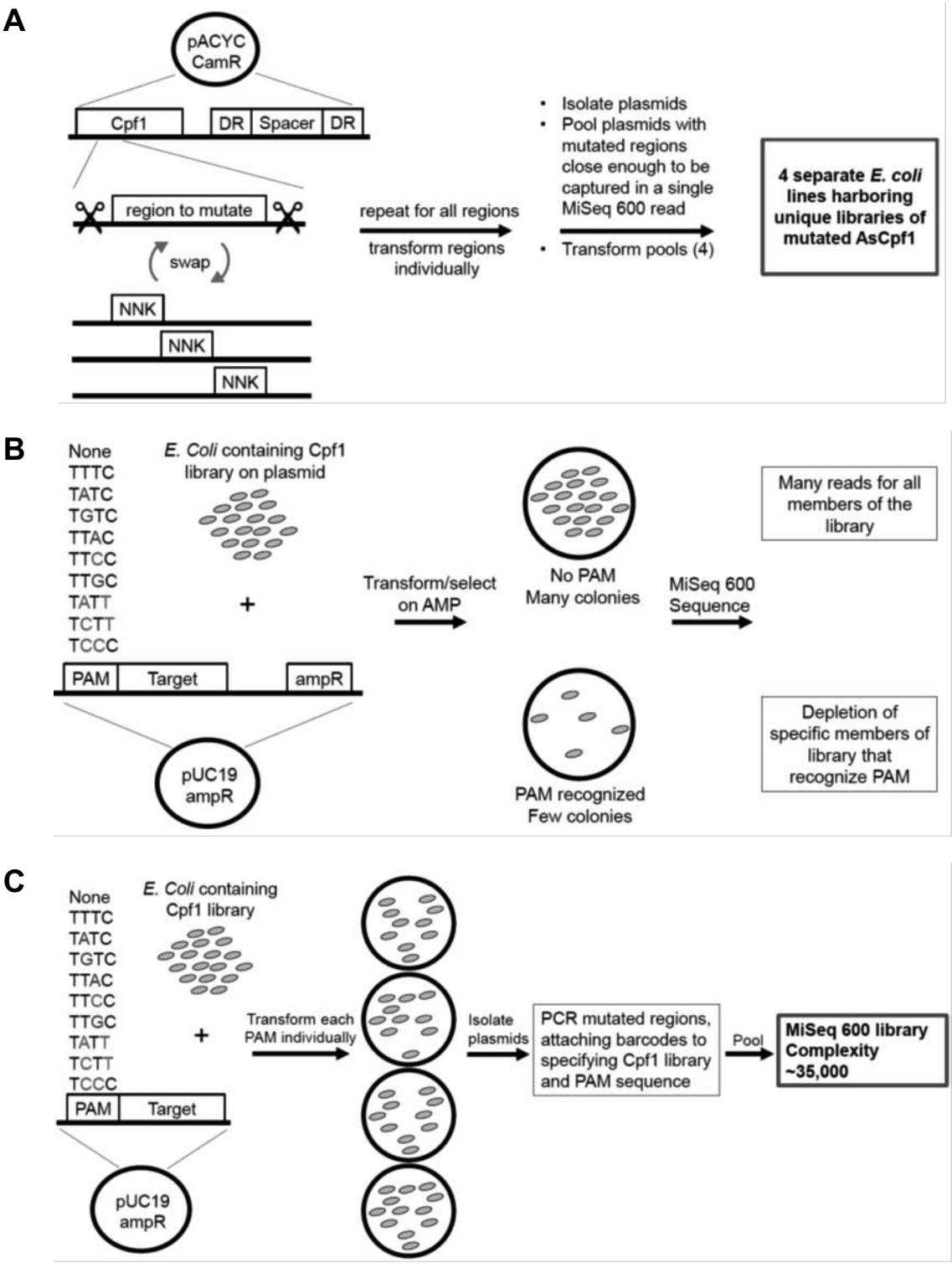
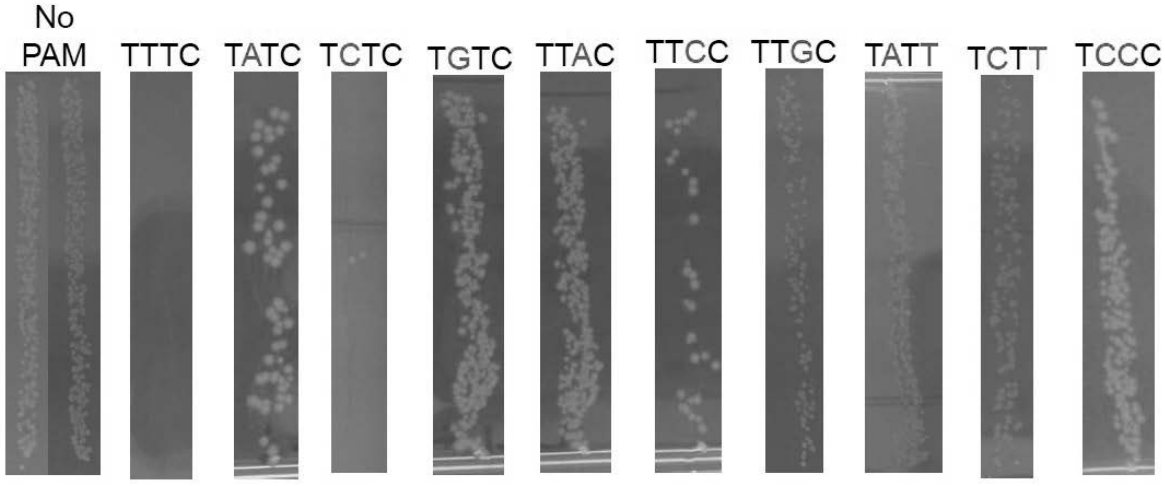

A
pACYC
CamR
Cpf1
DR
Spacer
DR
region to mutate
swap
NNK
NNK
NNK
repeat for all regions
transform regions individually
Isolate plasmids
Pool plasmids with mutated regions close enough to be captured in a single MiSeq 600 read
Transform pools (4)
4 separate E. coli lines harboring unique libraries of mutated AsCpf1
B
None
TTTC
TATC
TGTC
TTAC
TTCC
TTGC
TATT
TCTT
TCCC
E. Coli containing Cpf1 library on plasmid
+
PAM
Target
ampR
pUC19
ampR
Transform/select on AMP
No PAM
Many colonies
MiSeq 600 Sequence
Many reads for all members of the library
PAM recognized
Few colonies
Depletion of specific members of library that recognize PAM
C
None
TTTC
TATC
TGTC
TTAC
TTCC
TTGC
TATT
TCTT
TCCC
E. Coli containing Cpf1 library
+
PAM
Target
pUC19
ampR
Transform each PAM individually
Isolate plasmids
PCR mutated regions, attaching barcodes to specifying Cpf1 library and PAM sequence
Pool
MiSeq 600 library Complexity ~35,000

FIG. 3A-3C

FIG. 4A


Cm/Amp

FIG. 4B


No PAM
TTTC
TATC
TCTC
TGTC
TTAC
TTCC
TTGC
TATT
TCTT
TCCC

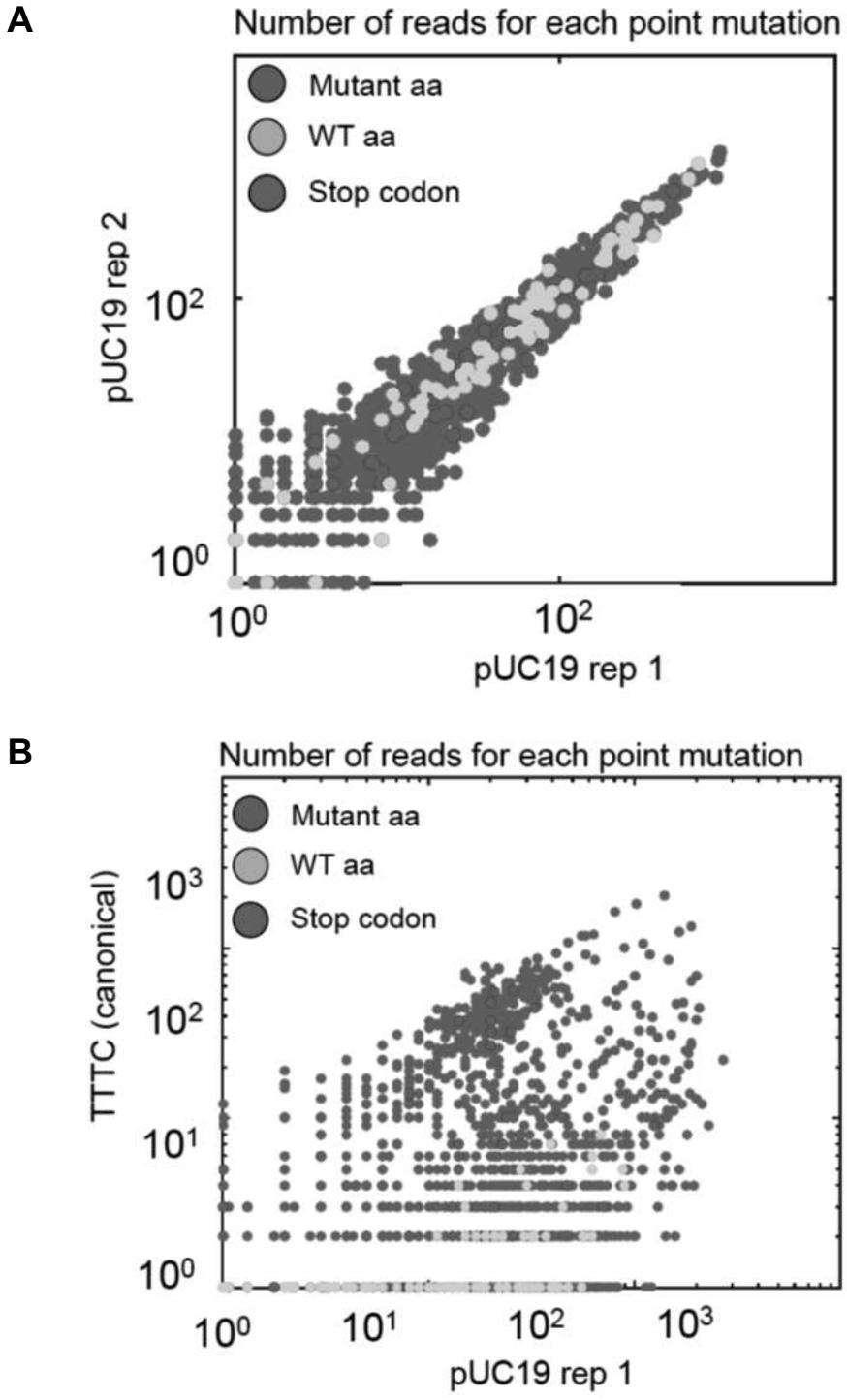

A
Number of reads for each point mutation
Mutant aa
WT aa
Stop codon
pUC19 rep 2
10^2
10^0
10^0
10^2
pUC19 rep 1
B
Number of reads for each point mutation
Mutant aa
WT aa
Stop codon
TTTC (canonical)
10^3
10^2
10^1
10^0
10^0
10^1
10^2
10^3
pUC19 rep 1

FIG. 5A-5B

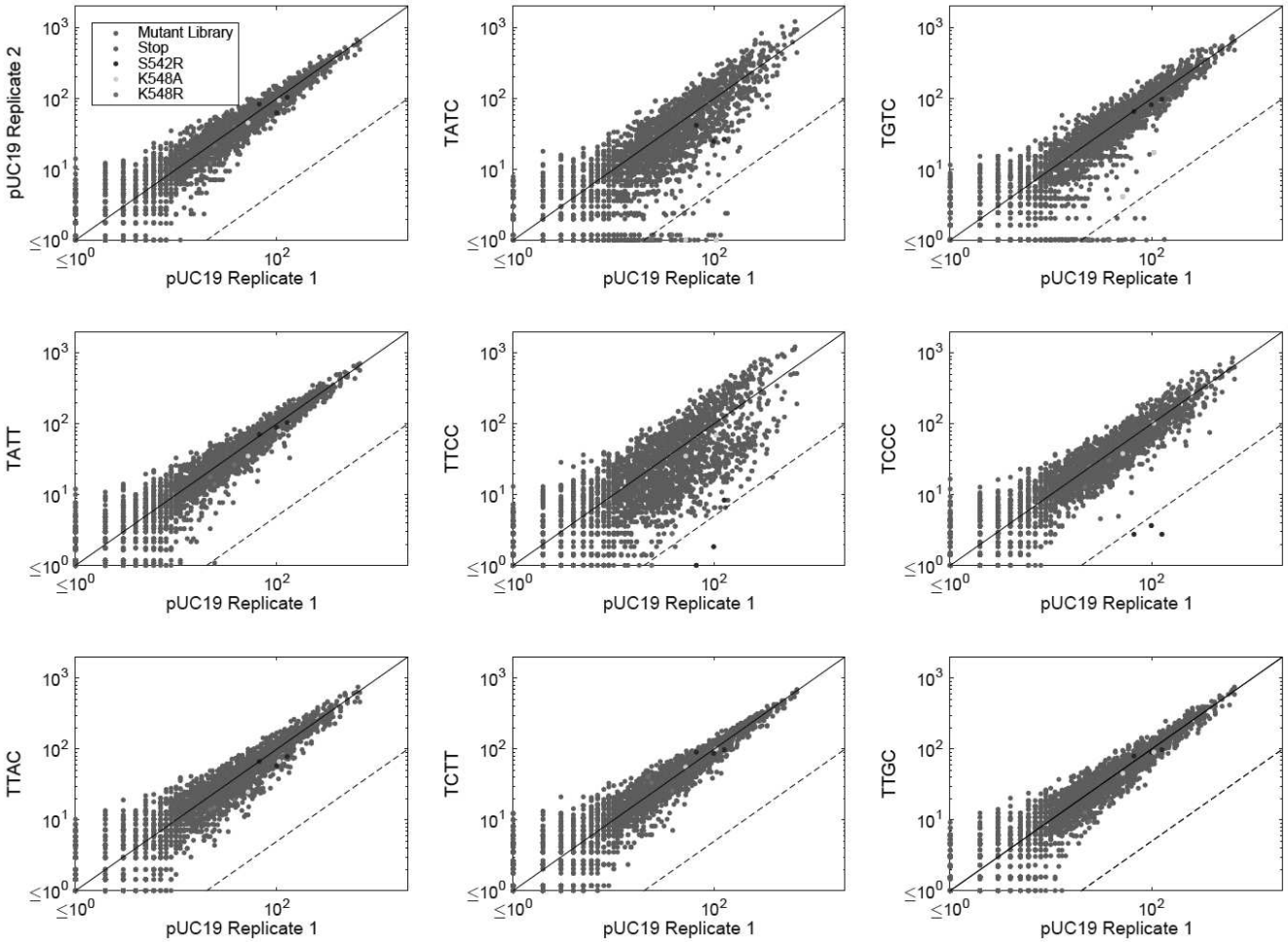

Mutant Library
Stop
S542R
K548A
K548R
pUC19 Replicate 2
pUC19 Replicate 1
TATC
TGTC
TATT
TTCC
TCCC
TTAC
TCTT
TTGC

FIG. 6

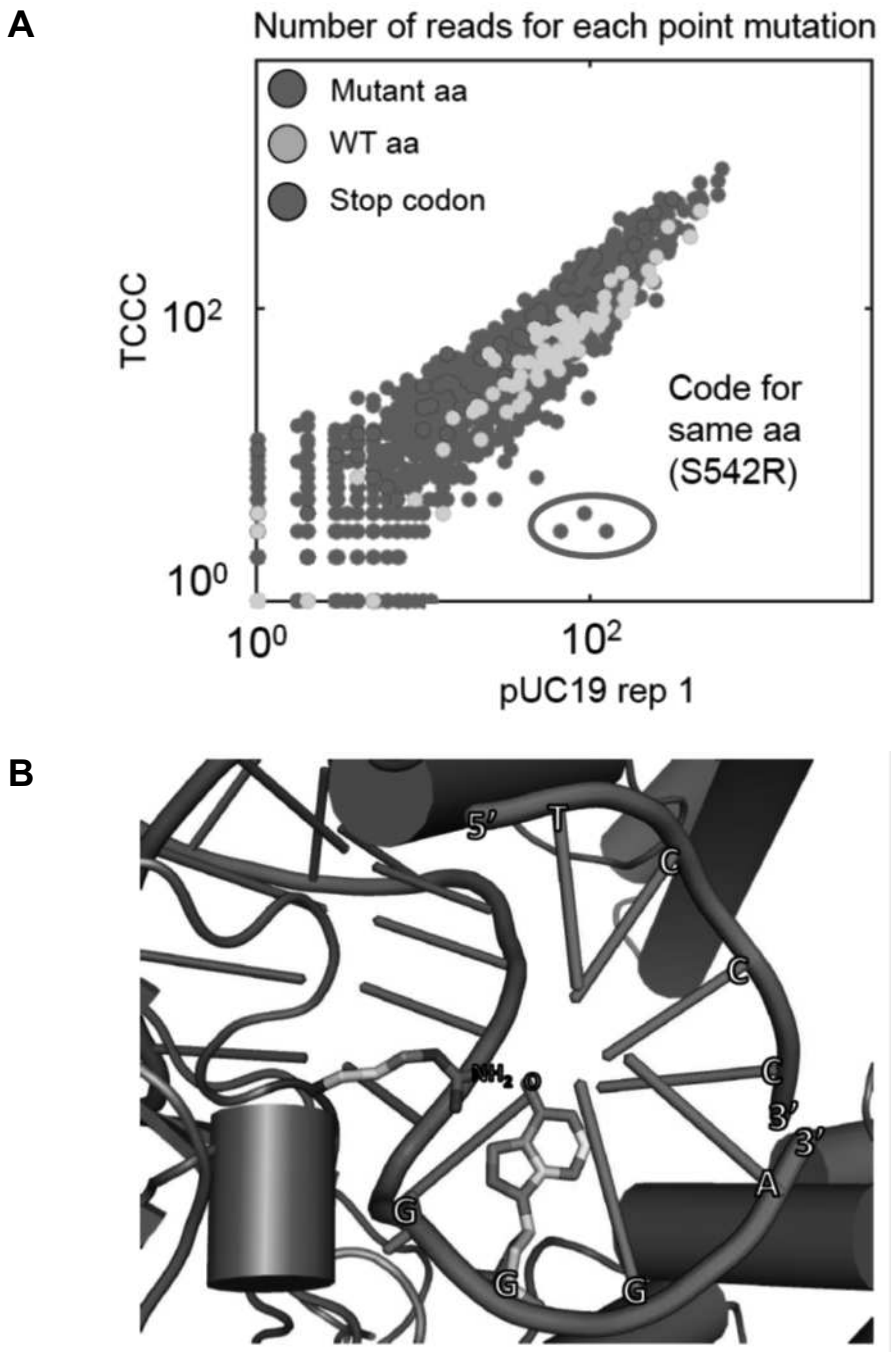

A
Number of reads for each point mutation
Mutant aa
WT aa
Stop codon
TCCC
10^2
10^0
Code for
same aa
(S542R)
10^0
10^2
pUC19 rep 1
B
5′
T
C
C
C
3′
3′
A
G
G
G
NH_2
O

FIG. 7A-7B

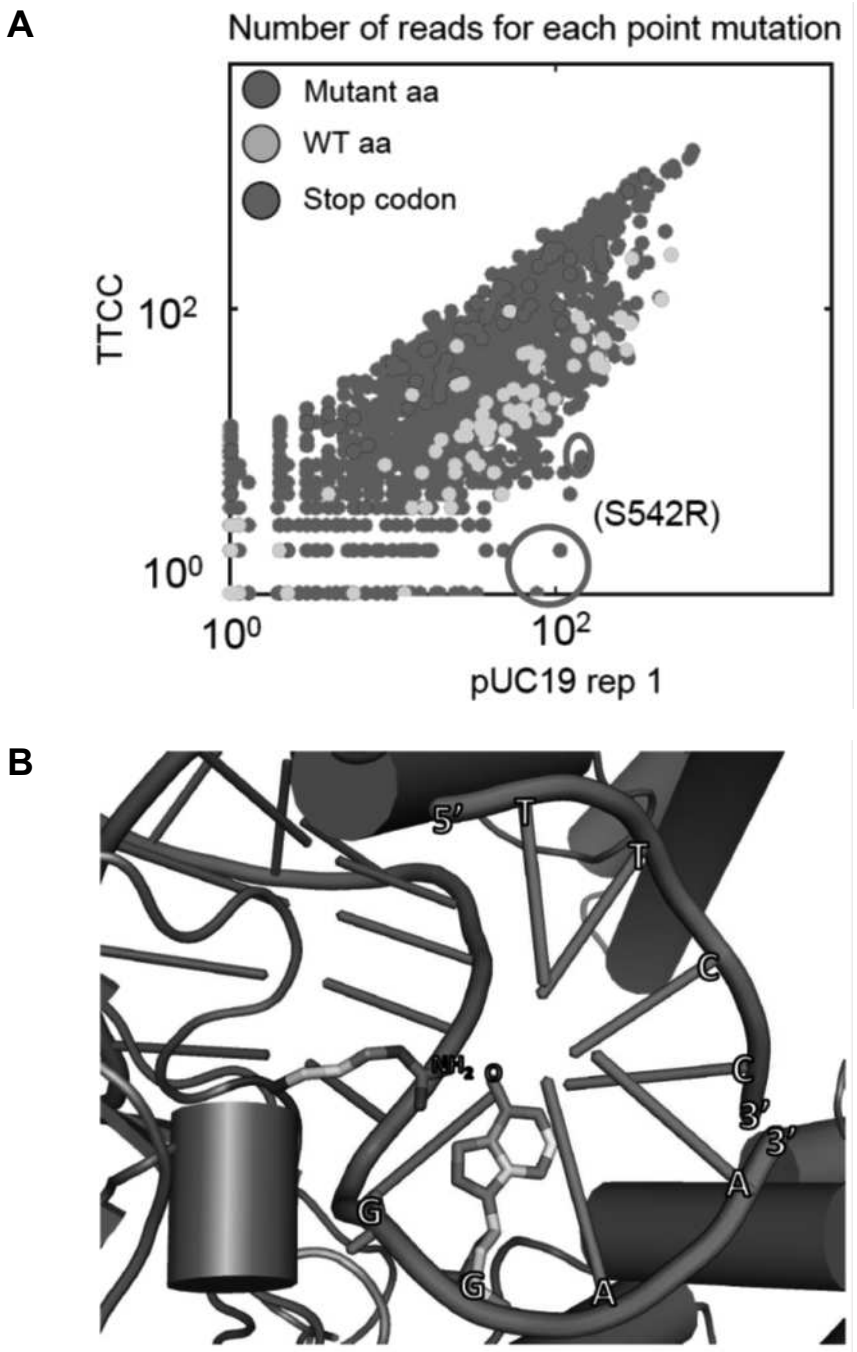
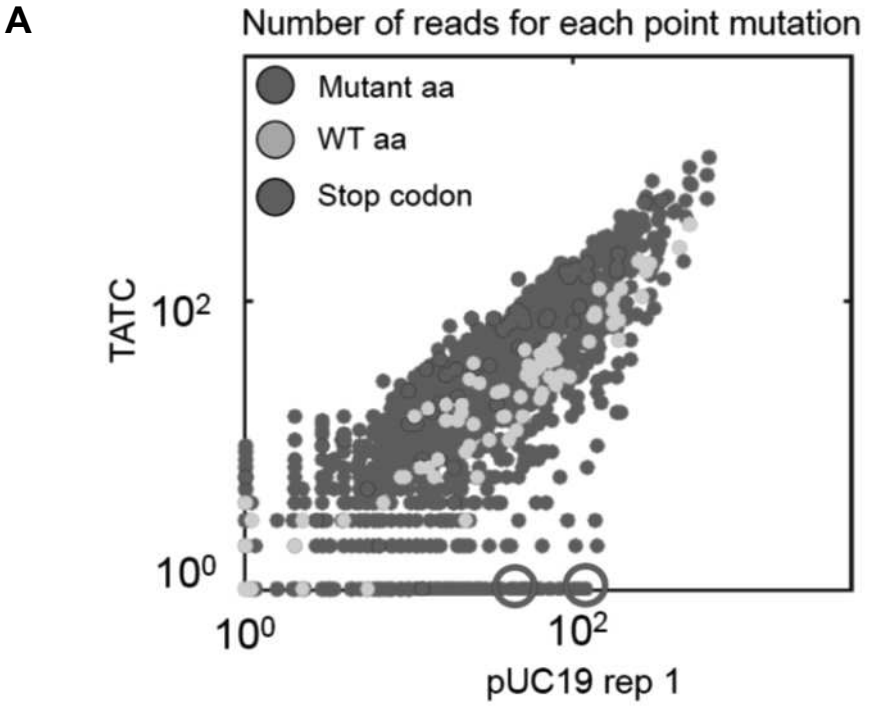

A
Number of reads for each point mutation
Mutant aa
WT aa
Stop codon
TTCC
10^2
10^0
(S542R)
10^0
10^2
pUC19 rep 1
B
5′
T
T
C
C
3′
3′
A
G
G
A

FIG. 8A-8B


A
Number of reads for each point mutation
Mutant aa
WT aa
Stop codon
TATC
10^2
10^0
10^0
10^2
pUC19 rep 1

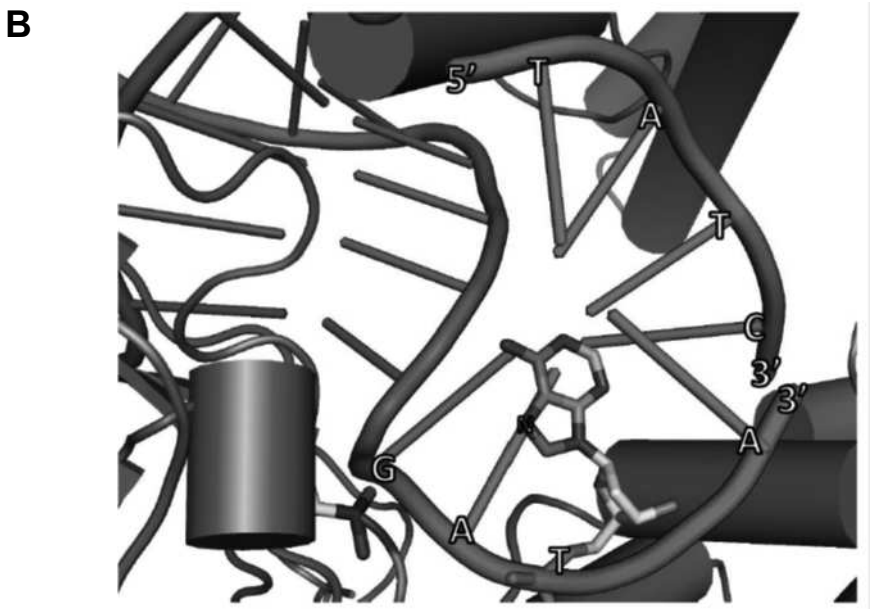
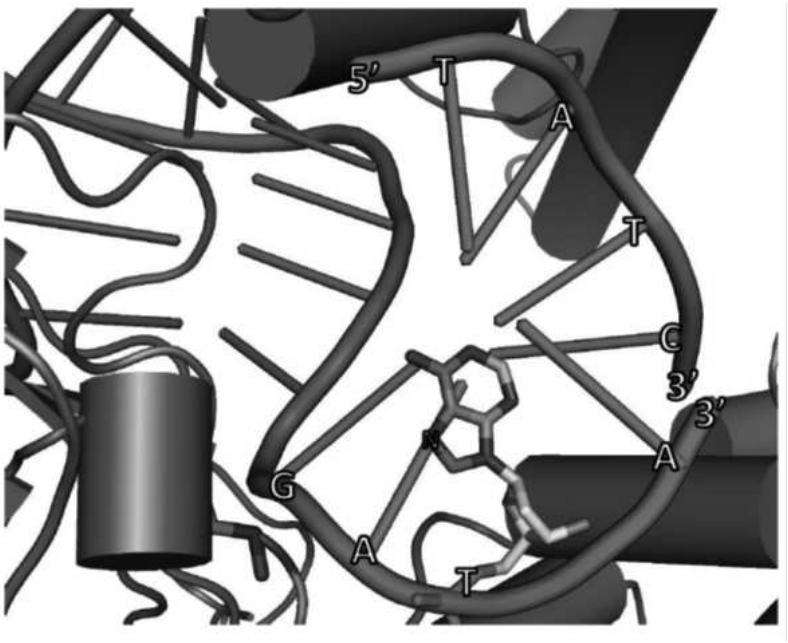

B
5′
T
A
T
C
3′
3′
A
G
A
T

FIG. 9A-9B

C



D

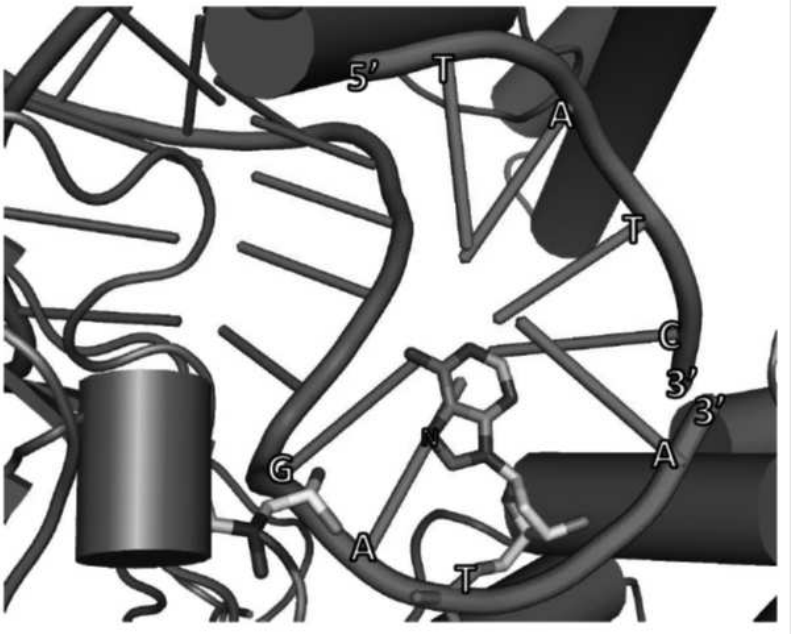
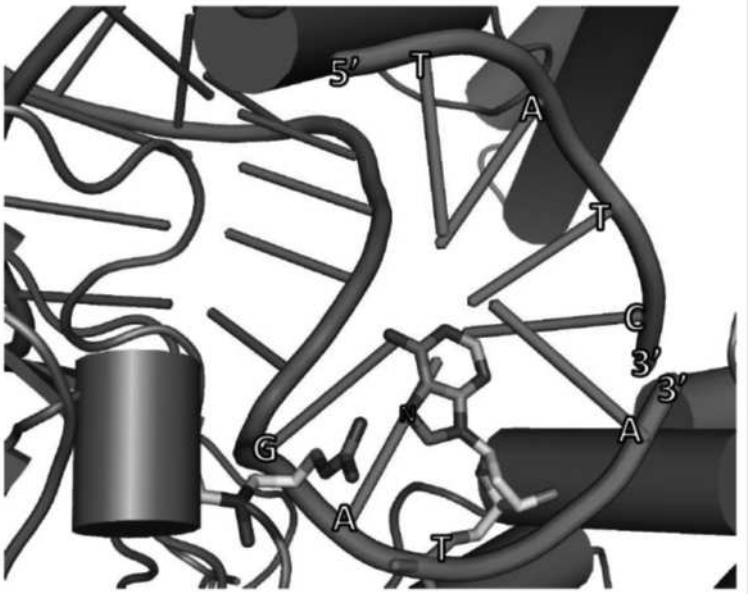


FIG. 9C-9D

E


5′
T
A
T
C
3′
3′
A
G
A
T

F

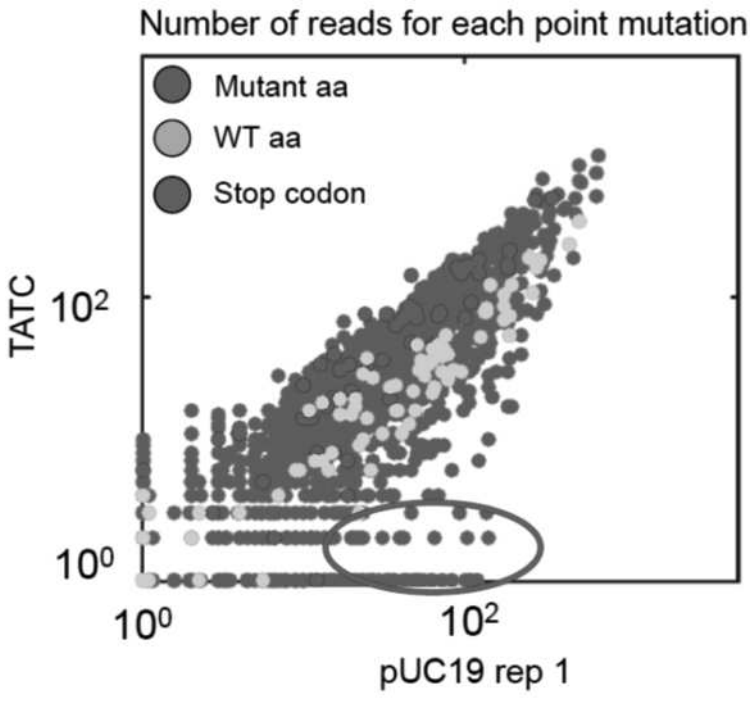

Number of reads for each point mutation
Mutant aa
WT aa
Stop codon
TATC
10²
10⁰
10⁰
10²
pUC19 rep 1

FIG. 9E-9F

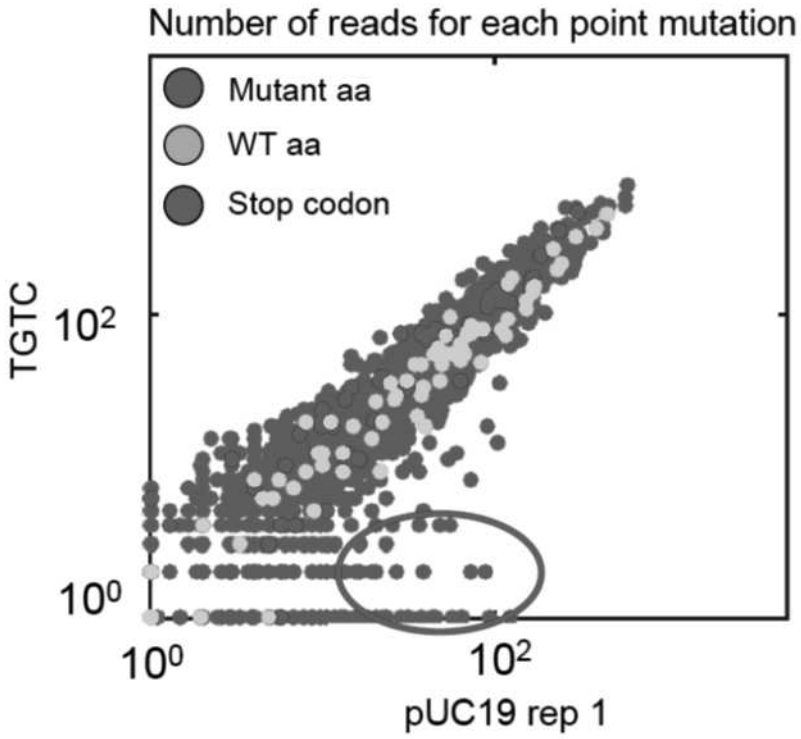

Number of reads for each point mutation
Mutant aa
WT aa
Stop codon
TGTC
10^2
10^0
10^0
10^2
pUC19 rep 1

FIG. 10

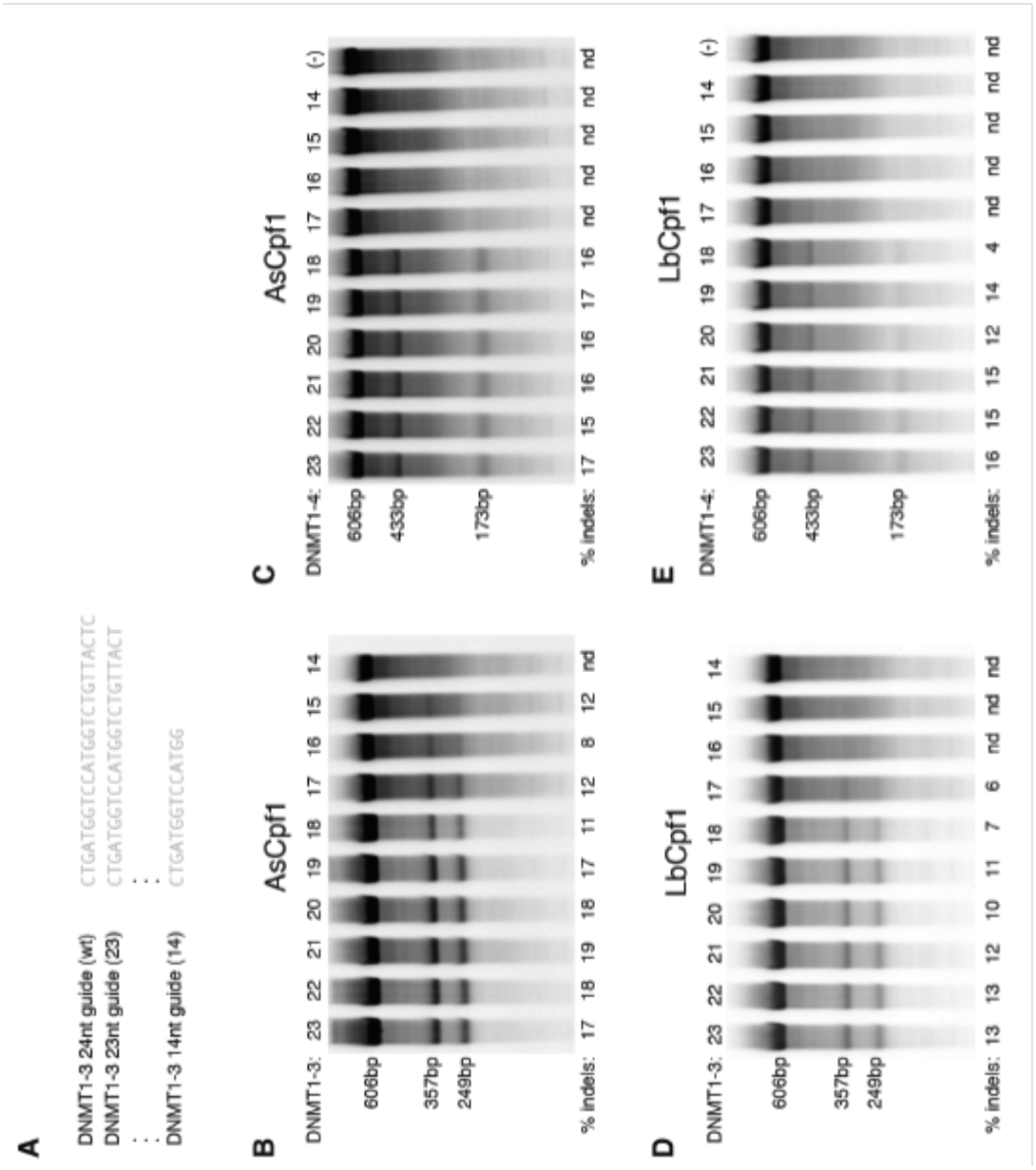

A
DNMT1-3 24nt guide (wt) CTGATGGTCCATGGTCTGTTACTC
DNMT1-3 23nt guide (23) CTGATGGTCCATGGTCTGTTACT
DNMT1-3 14nt guide (14) CTGATGGTCCATGG
B
AsCpf1
DNMT1-3: 23 22 21 20 19 18 17 16 15 14
606bp
357bp
249bp
% indels: 17 18 19 18 17 11 12 8 12 nd
C
AsCpf1
DNMT1-4: 23 22 21 20 19 18 17 16 15 14 (-)
606bp
433bp
173bp
% indels: 17 15 16 16 17 16 nd nd nd nd nd
D
LbCpf1
DNMT1-3: 23 22 21 20 19 18 17 16 15 14
606bp
357bp
249bp
% indels: 13 13 12 10 11 7 6 nd nd nd
E
LbCpf1
DNMT1-4: 23 22 21 20 19 18 17 16 15 14 (-)
606bp
433bp
173bp
% indels: 16 15 15 12 14 4 nd nd nd nd nd

FIG. 11A-11E

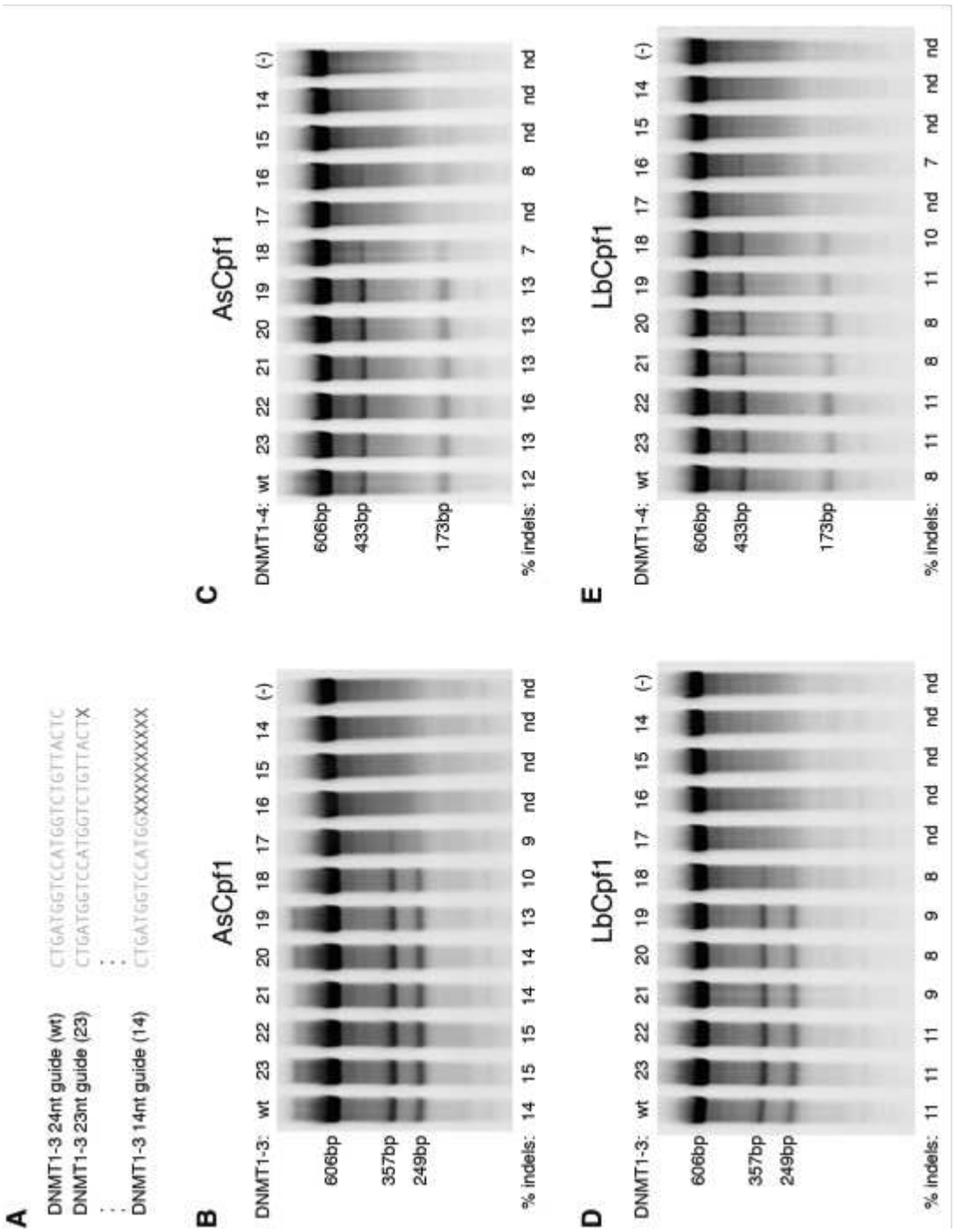
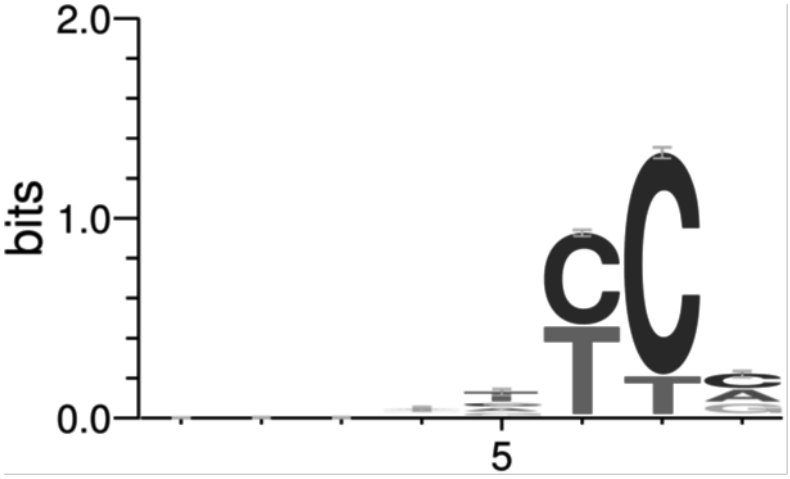

A
DNMT1-3 24nt guide (wt) CTGATGGTCCATGGTCTGTTACTC
DNMT1-3 23nt guide (23) CTGATGGTCCATGGTCTGTTACTX
DNMT1-3 14nt guide (14) CTGATGGTCCATGGXXXXXXXXXX
B
AsCpf1
DNMT1-3: wt 23 22 21 20 19 18 17 16 15 14 (-)
606bp
357bp
249bp
% indels: 14 15 15 14 14 13 10 9 nd nd nd nd
C
AsCpf1
DNMT1-4: wt 23 22 21 20 19 18 17 16 15 14 (-)
606bp
433bp
173bp
% indels: 12 13 16 13 13 13 7 nd 8 nd nd nd
D
LbCpf1
DNMT1-3: wt 23 22 21 20 19 18 17 16 15 14 (-)
606bp
357bp
249bp
% indels: 11 11 11 9 8 9 8 nd nd nd nd nd
E
LbCpf1
DNMT1-4: wt 23 22 21 20 19 18 17 16 15 14 (-)
606bp
433bp
173bp
% indels: 8 11 11 8 8 11 10 nd 7 nd nd nd

FIG. 12A-12^{E}

A



B

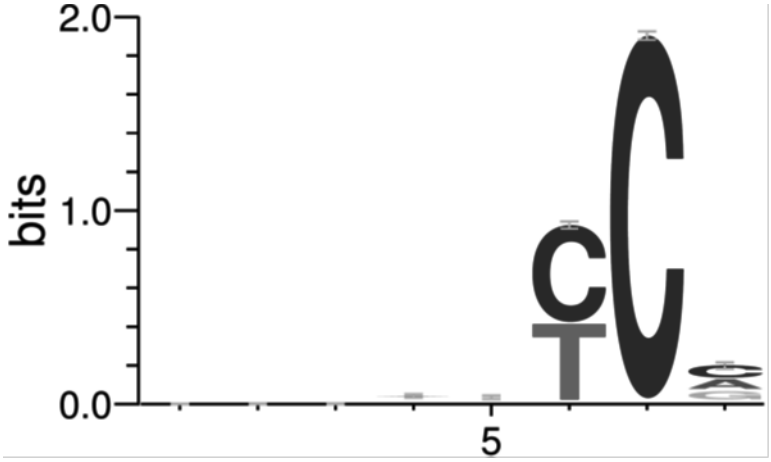
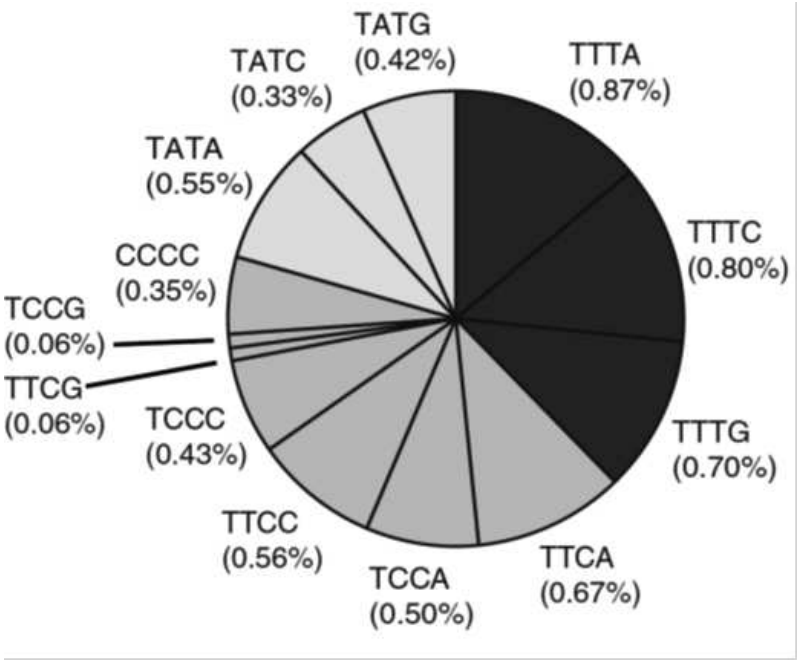


FIG. 13A-13B

C



D

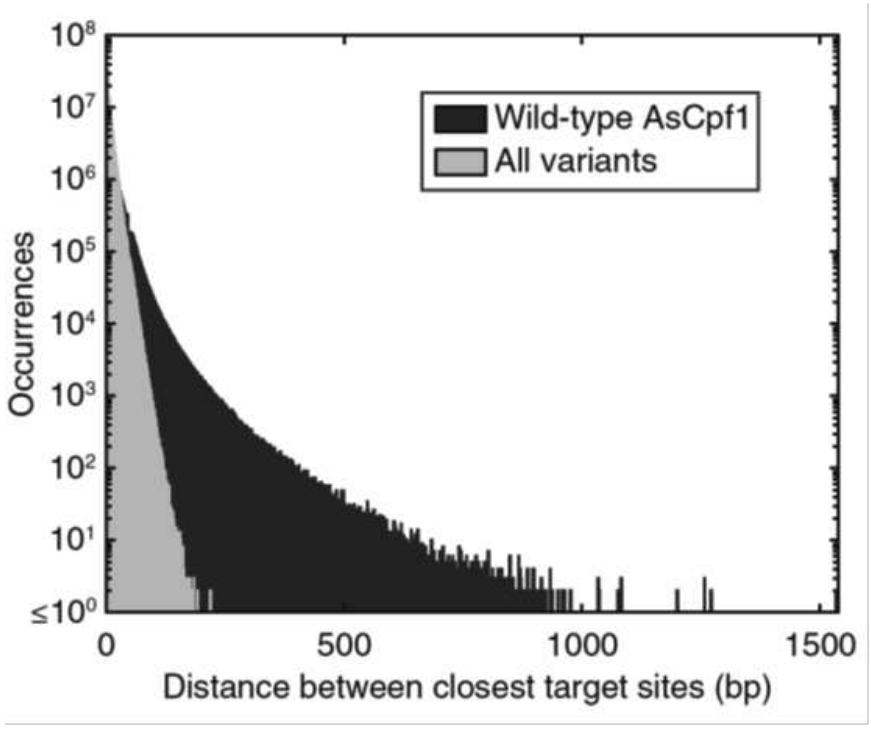


FIG. 13C-13D

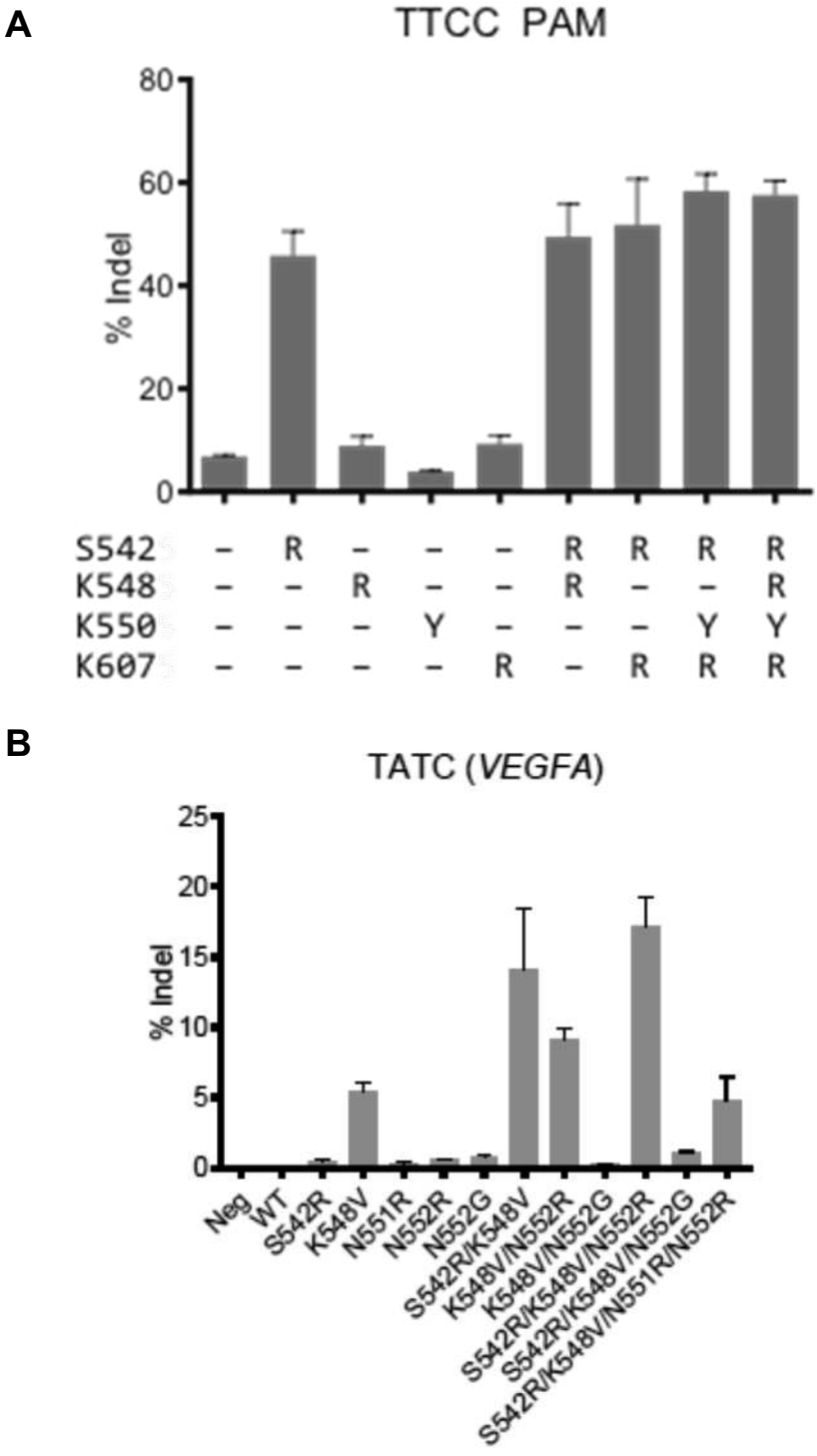
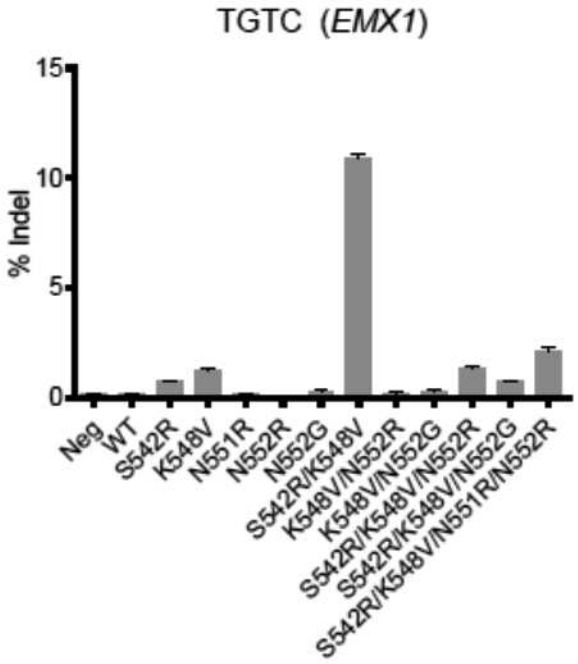

A
TTCC PAM
% Indel
80
60
40
20
0
S542 – R – – – R R R R
K548 – – R – – R – – R
K550 – – – Y – – – Y Y
K607 – – – – R – R R R
B
TATC (VEGFA)
% Indel
25
20
15
10
5
0
Neg
WT
S542R
K548V
N551R
N552R
N552G
S542R/K548V
K548V/N552R
K548V/N552G
S542R/K548V/N552R
S542R/K548V/N552G
S542R/K548V/N551R/N552R

FIG. 14A-14B

C



D

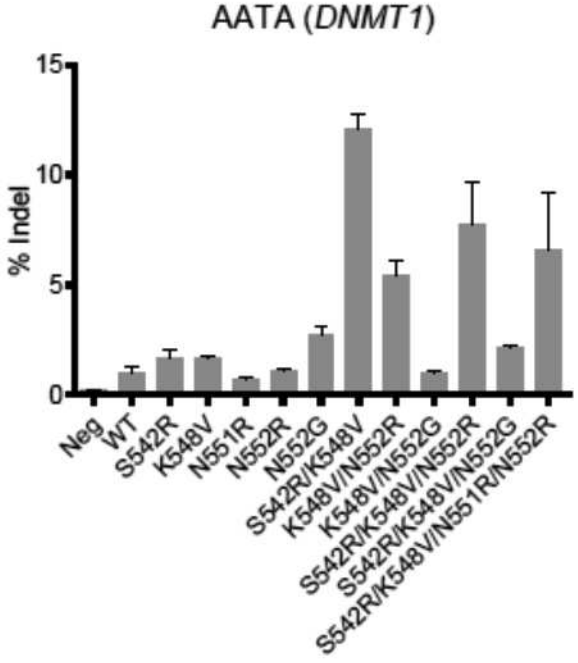


FIG. 14C-14D

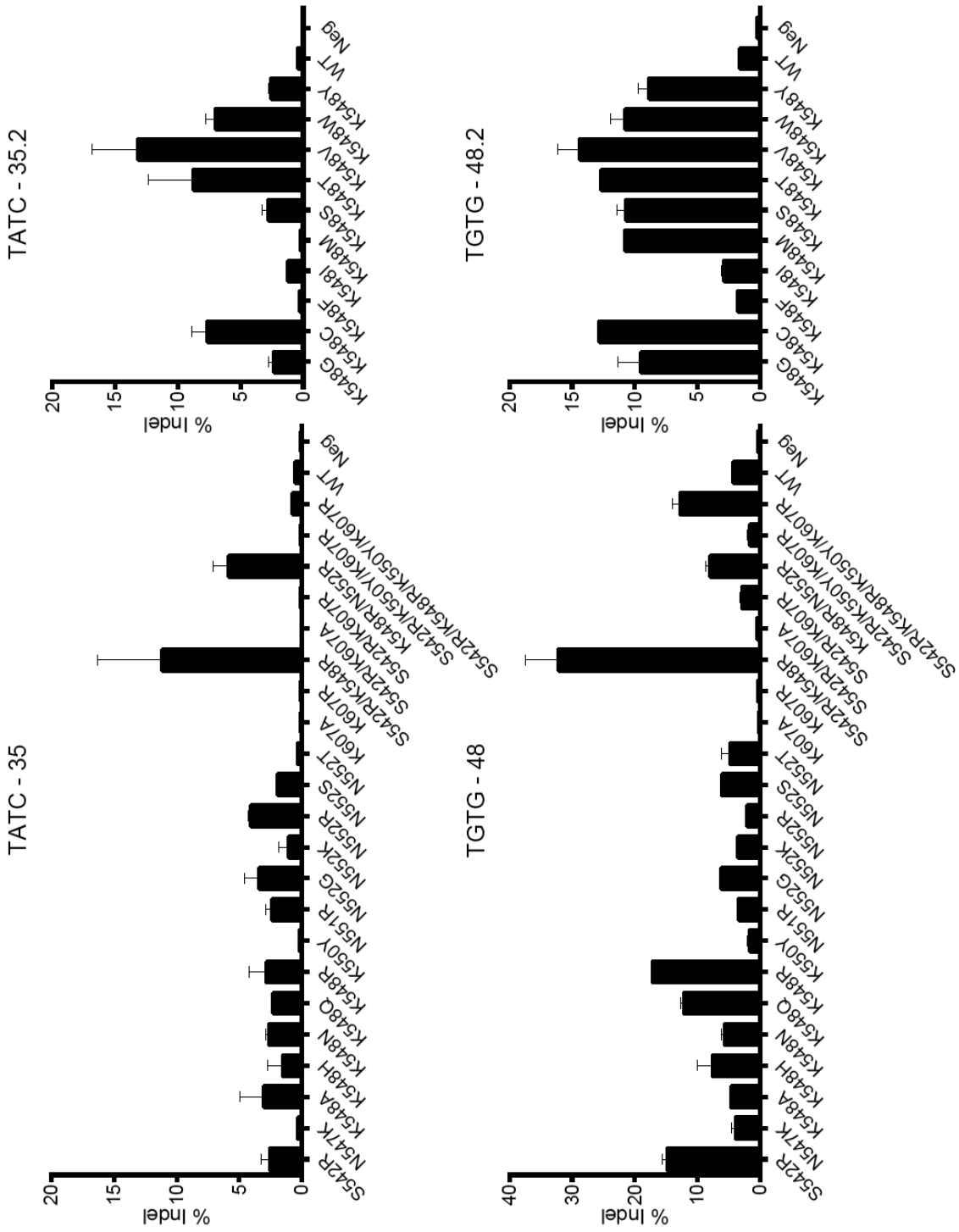


FIG. 14D (continued)

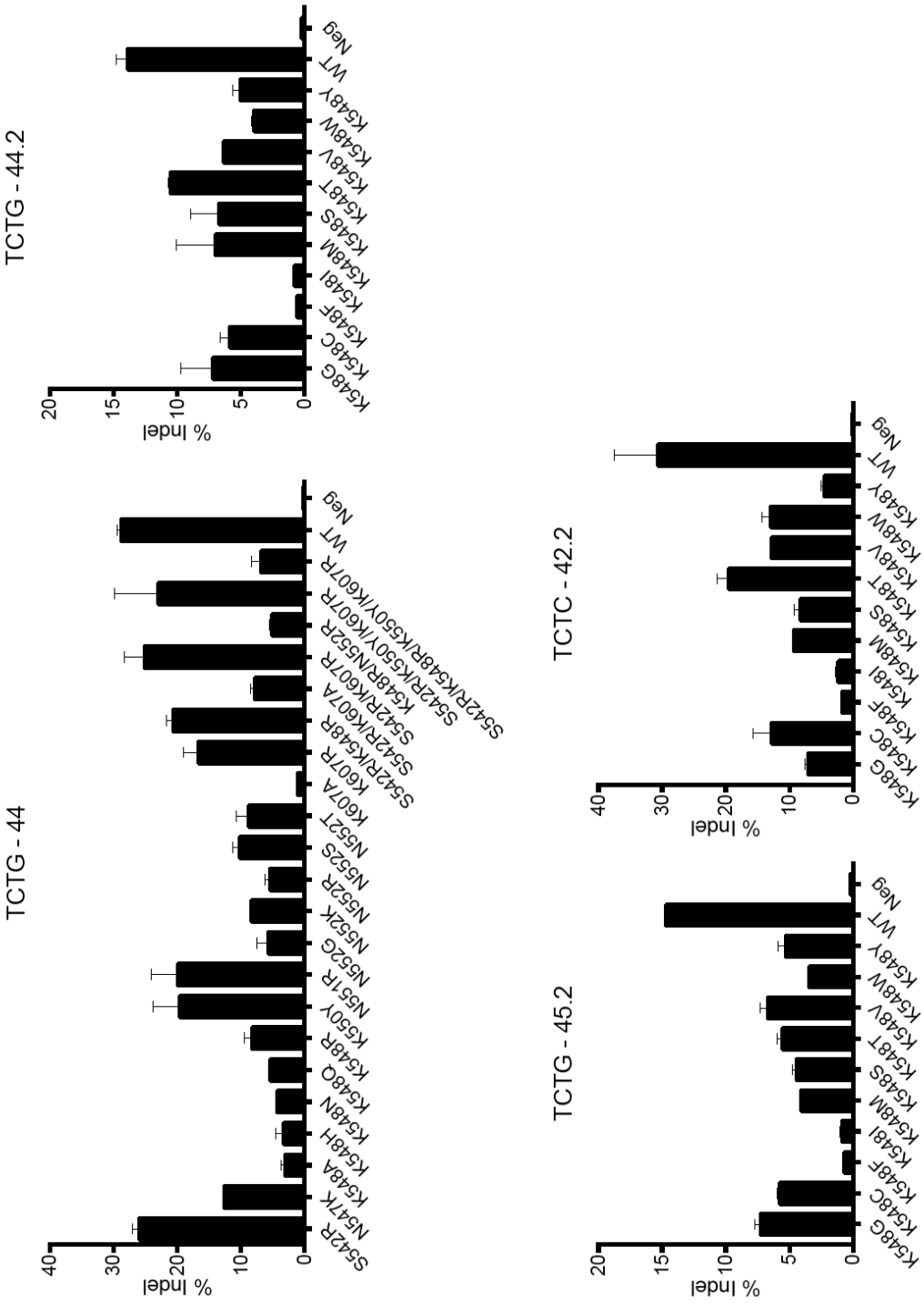


FIG. 14D (continued)

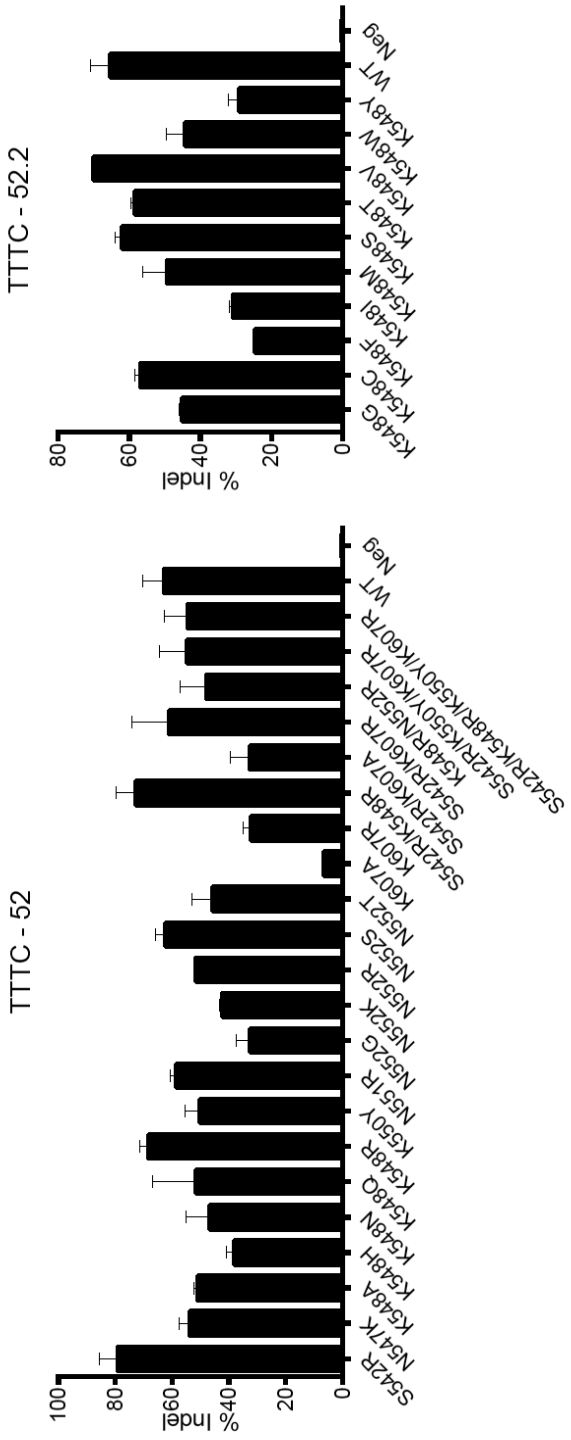


FIG. 14D (continued)

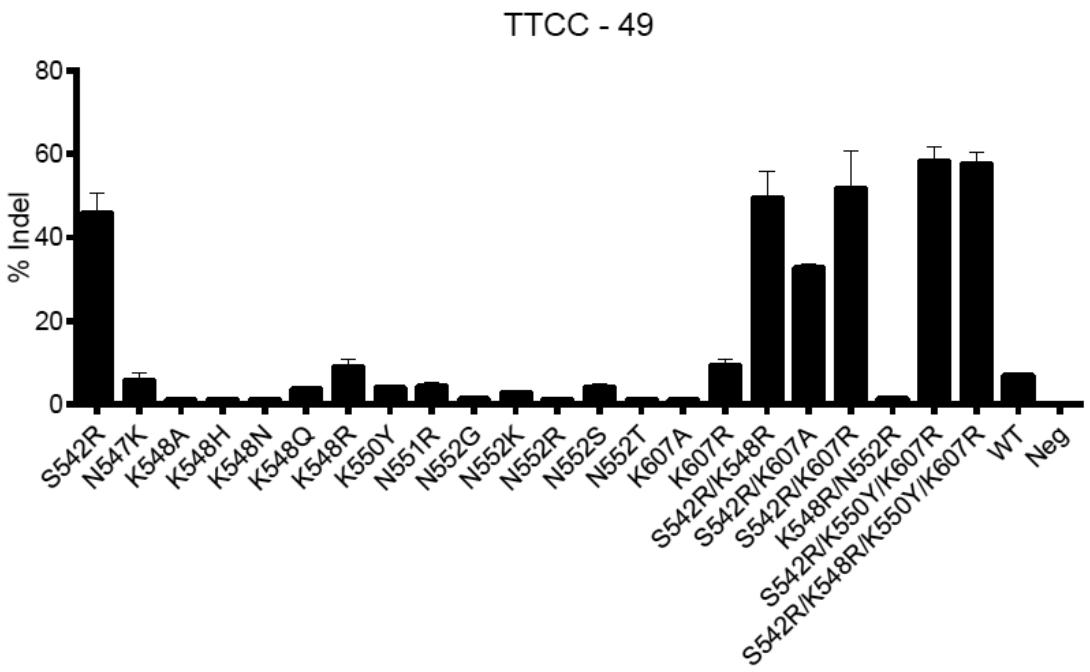
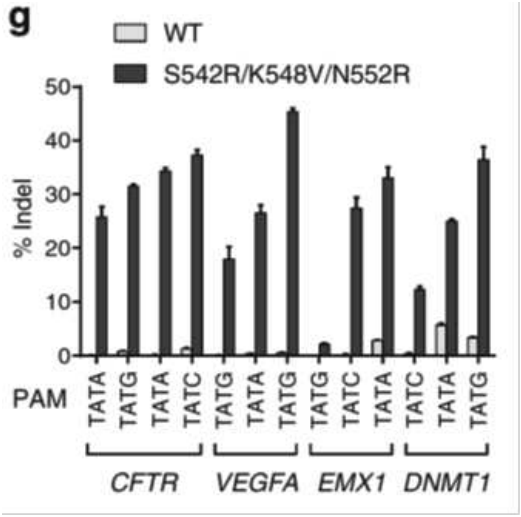
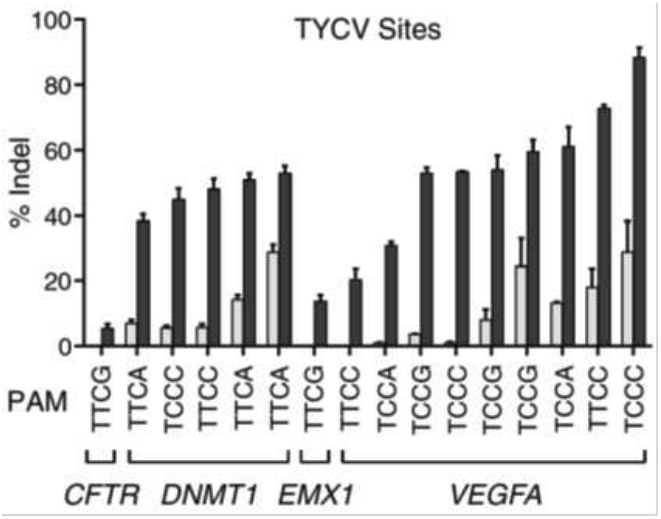


FIG. 14D (continued)

A



B



C

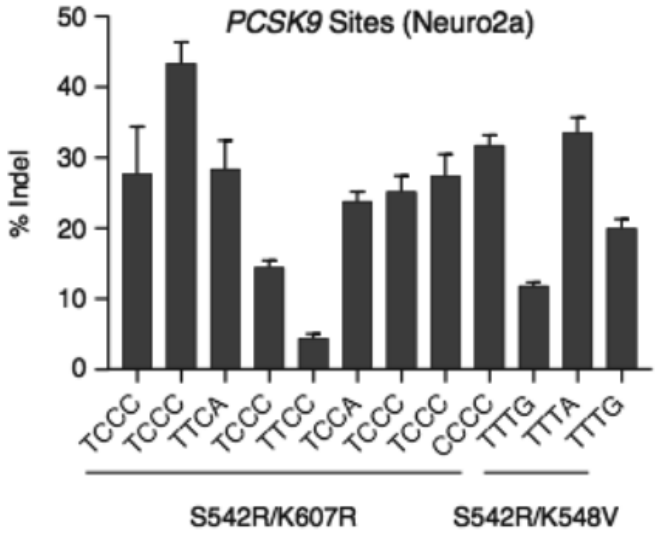


FIG. 15A-15C

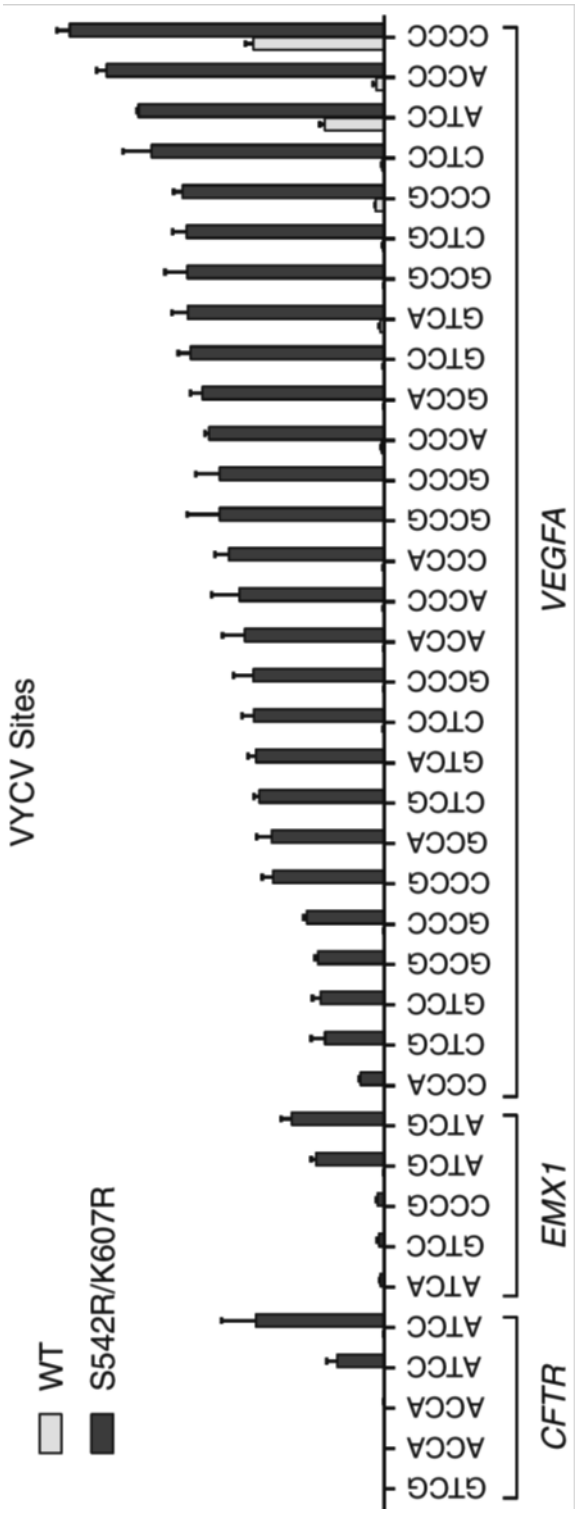

WT
S542R/K607R
VYCV Sites
CCCC
ACCC
ATCC
CTCC
CCCG
CTCG
GCCG
GTCA
GTCC
GCCA
ACCC
GCCC
GCCG
CCCA
ACCC
ACCA
GCCC
CTCC
GTCA
CTCG
GCCA
CCCG
GCCC
GCCG
GTCC
CTCG
CCCA
ATCG
ATCG
CCCG
GTCC
ATCA
ATCC
ATCC
ACCA
ACCA
GTCG
VEGFA
EMX1
CFTR

FIG. 15D

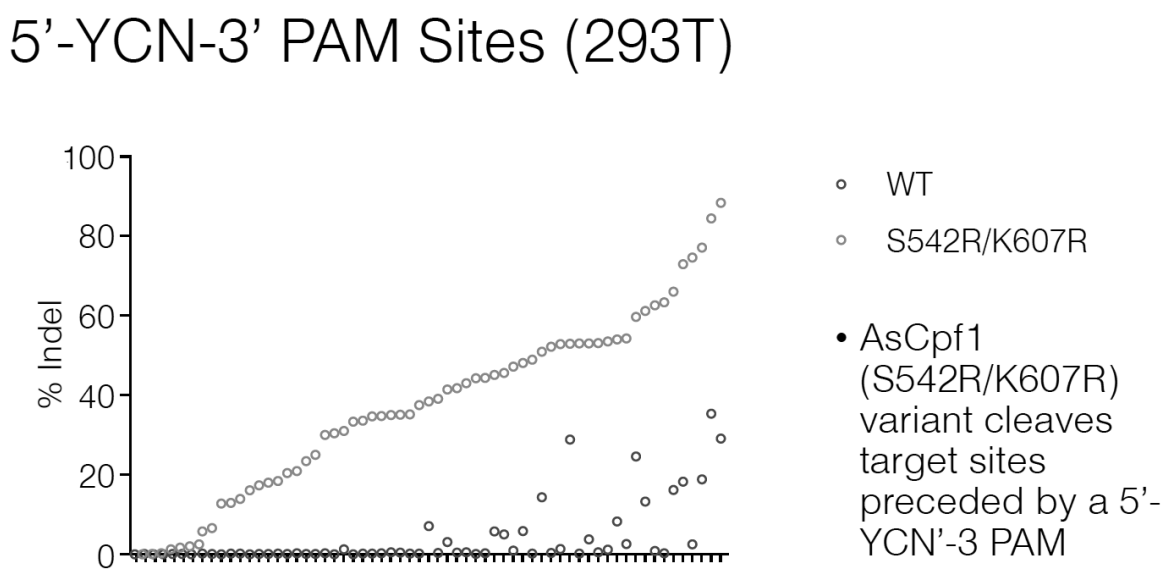

5’-YCN-3’ PAM Sites (293T)
100
80
60
40
20
0
% Indel
WT
S542R/K607R
• AsCpf1 (S542R/K607R) variant cleaves target sites preceded by a 5’-YCN’-3 PAM

FIG. 16

```
AsCpf1     1 MTQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQGFIEEDKARNDHYKELKPIIDRIYKT   60
LbCpf1     1 MSKLEKFTNCYSLSKTLRFKAIPVGKTQENIDNKRLLVEDEKRAEDYKGVKKLLDRYYLS   60
AsCpf1    61 YADQCLQLVQLDWENLSAAIDSYRKEKTEETRNALIEE-QATYRNAIHDYFIGRTDNLTD  119
LbCpf1    61 FINDVLHSIKL--KNLNNYISLFRKKTRTEKENKELENLEINLRKEIAKAFKGN------  112
AsCpf1   120 AINKRHAEIYKGLFKAELFNGKVLKQLGTVTTTEHENALLRSFDKFTTYFSGFYENRKNV  179
LbCpf1   113 -------EGYKSLFKKDI----IETILPEFLDDKDEIALVNSFNGFTTAFTGFFDNRENM  161
AsCpf1   180 FSAEDISTAIPHRIVQDNFPKFKENCHIFTRLITAVPSLREHFENVKKAIGIFVSTSIEE  239
LbCpf1   162 FSEEAKSTSIAFRCINENLTRYISNMDIFEKVDAIFD--KHEVQEIKEKI-LNSDYDVED  218
AsCpf1   240 VFSFPFYNQLLTQTQIDLYNQLLGGISREAGTEKIKGLNEVLNLAIQKNDETAHIIASLP  299
LbCpf1   219 FFEGEFFNFVLTQEGIDVYNAIIGGFVTESG-EKIKGLNEYINLYNQKTKQ------KLP  271
AsCpf1   300 HRFIPLFKQILSDRNTLSFILEEFKSDEEVIQSFCKYKTLLRNENVLETA---EALFNEL  356
LbCpf1   272 -KFKPLYKQVLSDRESLSFYGEGYTSDEEVLEVF--RNTLNKNSEIFSSIKKLEKLFKNF  328
AsCpf1   357 NSIDLTHIFISHK-KLETISSALCDHWDTLR---NALYERRISELTGKITKSAKEKVQRS  412
LbCpf1   329 DEYSSAGIFVKNGPAISTISKDIFGEWNVIRDKWNAEYDDIHLKKKAVVTEKYEDDRRKS  388
AsCpf1   413 LKH-EDINLQEIISAAGKELS------EAFKQKTSEIL-----SHAHAALDQPLPTTLKK  460
LbCpf1   389 FKKIGSFSLEQLQEYADADLSVVEKLKEIIIQKVDEIYKVYGSSEKLFDADFVLEKSLKK  448
AsCpf1   461 QEEK-EILKSQLDSLLGLYHLLD-WFAVDESNEVDPEFSARLT---GIKLEMEPSLSFYN  515
LbCpf1   449 NDAVVAIMKDLLDSVKSFENYIKAFFGEGKETNRDESFYGDFVLAYDILLKVD---HIYD  505
AsCpf1   516 KARNYATKKPYSVEKFKLNFQMPTLASGWDVNKEKNNGAILFVKNGLYYLGIMPKQKGRY  575
LbCpf1   506 AIRNYVTQKPYSKDKFKLYFQNPQFMGGWDKDKETDYRATILRYGSKYYLAIMDKKYA--  563
AsCpf1   576 KALSFEPTEKTSEGFDKMYYDYFPDAAKMIPKCSTQLKAVTAHFQTHTTPILLSNNFIEP  635
LbCpf1   564 KCLQKIDKDDVNGNYEKINYKLLPGPNKMLPKVFFSKKWMAYY----------------  606
AsCpf1   636 LEITKEIYDLNNPEKEPKKFQTAYAKKTGDQKGYREALCKWIDFTRDFLSKYTKTTSIDL  695
LbCpf1   607 -----------NPSEDIQKIYKNGTFKKGDMFNLNDC-HKLIDFFKDSISRYPKWSNAYD  654
```

FIG. 17

```
AsCpf1   696 SSLRPSSQYKDLGEYYAELNPLLYHISFQRIAEKEIMDAVETGKLYLFQIYNKDFAKGHH   755
LbCpf1   655 FNFSETEKYKDIAGFYREVEEQGYKVSFESASKKEVDKLVEEGKLYMFQIYNKDFSDKSH   714
AsCpf1   756 GKPNLHTLYWTGLFSPENLAKTSIKLNGQAELFYRPKSRMKR--MAHRLGEKMLNKKLKD   813
LbCpf1   715 GTPNLHTMYFKLLFDENNHGQ--IRLSGGAELFMRRASLKKEELVVHPANSPIANKN---   769
AsCpf1   814 QKTPIPDTLYQELYDYVNHRLSHDLSDEARALLPNVITKEVSHEIIKDRRFTSDKFFFHV   873
LbCpf1   770 -----PDNPKK-------------------------TTTLSYDVYKDKRFSEDQYELHI   798
AsCpf1   874 PITLNYQAANSPSKFNQRVNAYLKEHPETP-IIGIDRGERNLIYITVIDSTGKILEQRSL   932
LbCpf1   799 PIAIN-KCPKNIFKINTEVRVLLK-HDDNPYVIGIDRGERNLLYIVVVDGKGNIVEQYSL   856
AsCpf1   933 NTI--------QQFDYQKKLDNREKERVAARQAWSVVGTIKDLKQGYLSQVIHEIVDLMI   984
LbCpf1   857 NEIINNFNGIRIKTDYHSLLDKKEKERFEARQNWTSIENIKELKAGYISQVVHKICELVE   916
AsCpf1   985 HYQAVVVLENLNFGFKSKRTGIAEKAVYQQFEKMLIDKLNCLVLKDYPAEKVGGVLNPYQ  1044
LbCpf1   917 KYDAVIALEDLNSGFKNSRVKV-EKQVYQKFEKMLIDKLNYMVDKKSNPCATGGALKGYQ   975
AsCpf1  1045 LTDQFTSFAKMGTQSGFLFYVPAPYTSKIDPLTGFVDPFVWKTIKNHESRKHFLEGFDFL  1104
LbCpf1   976 ITNKFESFKSMSTQNGFIFYIPAWLTSKIDPSTGFVNLLKTKYTSIADSKK-FISSFDRI  1034
AsCpf1  1105 HY--DVKTGDFILHFKMNRNLSFQRGLPGFMPAWDIVFEKNETQFDAKGTPFIAGKRIVP  1162
LbCpf1  1035 MYVPEEDLFEFALDYK-----NFSRTDADYIKKWKLYSYGNRIRIFRN-----PKKNNVF  1084
AsCpf1  1163 VIENHRFTGRYRDLYPANELIALLEEKGIVFRDGSNILPKLLENDDSHAIDTMVALIRSV  1222
LbCpf1  1085 DWEEVCLTSAYKELF---------NKYGINYQQG-DIRALLCEQSDKAFYSSFMALMSLM  1134
AsCpf1  1223 LQMRNS-NAATGEDYINSPVRDLNGVCFDSRF----QNPEWPMDADANGAYHIALKGQLL  1277
LbCpf1  1135 LQMRNSITGRTDVDFLISPVKNSDGIFYDSRNYEAQENAILPKNADANGAYNIARKVLWA  1194
AsCpf1  1278 LNHLKESKDLKLQN---GISNQDWLAYIQ                                 1303
LbCpf1  1195 IGQFKKAEDEKLDKVKIAISNKEWLEYAQ                                 1223
```

FIG. 17 (continued)

A

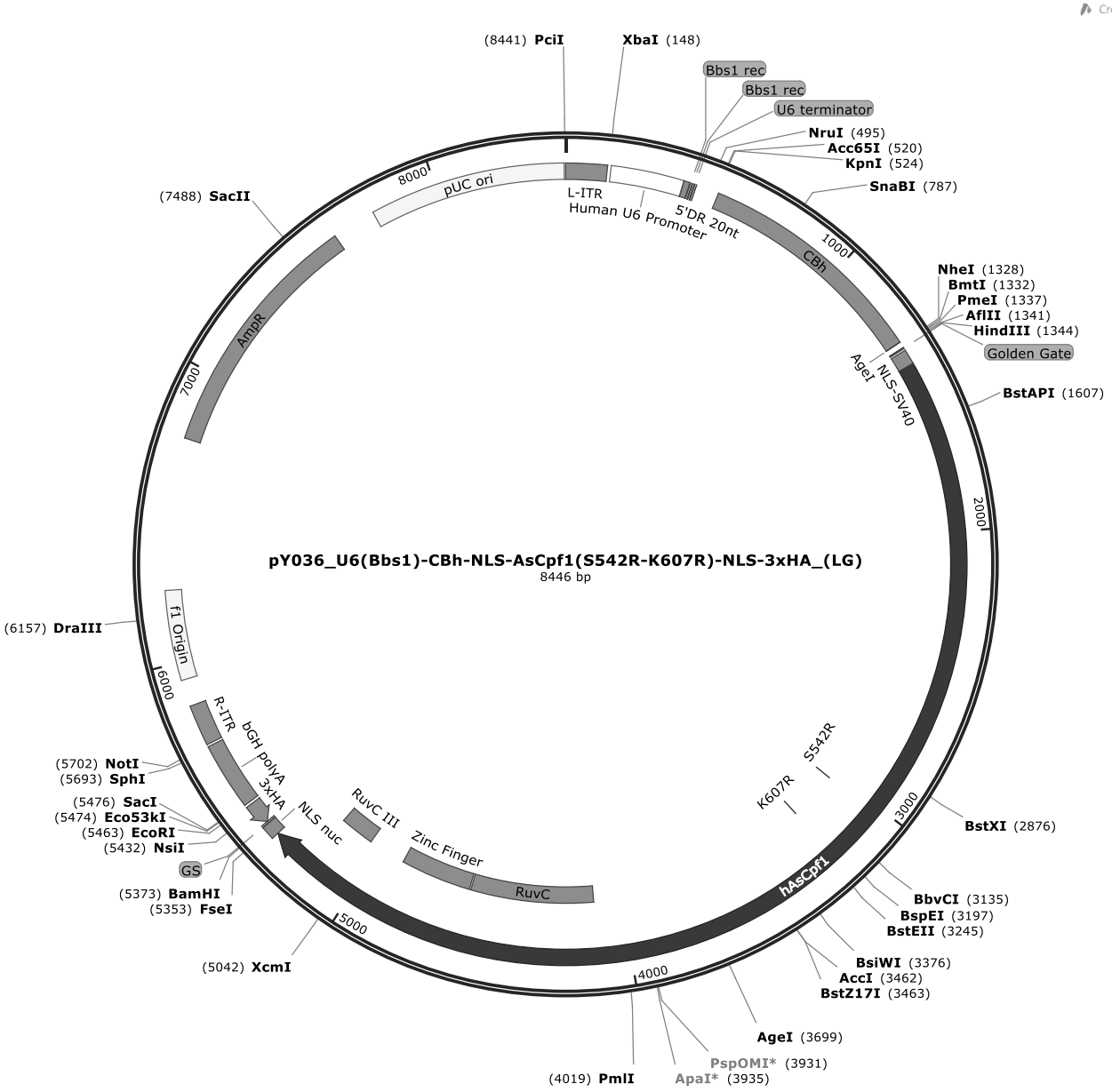

(8441) PciI
XbaI (148)
Bbs1 rec
Bbs1 rec
U6 terminator
NruI (495)
Acc65I (520)
KpnI (524)
SnaBI (787)
(7488) SacII
8000
pUC ori
L-ITR
Human U6 Promoter
5'DR 20nt
1000
CBh
NheI (1328)
BmtI (1332)
PmeI (1337)
AflII (1341)
HindIII (1344)
Golden Gate
AgeI
NLS-SV40
BstAPI (1607)
AmpR
7000
2000
pY036_U6(Bbs1)-CBh-NLS-AsCpf1(S542R-K607R)-NLS-3xHA_(LG)
8446 bp
(6157) DraIII
f1 Origin
6000
R-ITR
bGH polyA
3xHA
NLS nuc
RuvC III
Zinc Finger
RuvC
S542R
K607R
3000
BstXI (2876)
(5702) NotI
(5693) SphI
(5476) SacI
(5474) Eco53kI
(5463) EcoRI
(5432) NsiI
GS
(5373) BamHI
(5353) FseI
hASCpf1
BbvCI (3135)
BspEI (3197)
BstEII (3245)
5000
(5042) XcmI
4000
BsiWI (3376)
AccI (3462)
BstZ17I (3463)
AgeI (3699)
PspOMI* (3931)
(4019) PmlI
ApaI* (3935)

FIG. 18A


5′
3′
L-ITR
XbaI
Human U6 Promoter
NruI
5'DR 20nt
Bbs1 rec
Bbs1 rec
U6 terminator
Acc65I
KpnI
CBh
SnaBI
NheI
BmtI
PmeI
AflII
HindIII
Thr Gln Phe Glu Gly Phe Thr Asn
AgeI
Golden Gate
NLS-SV40
hAsCpf1

FIG. 18B

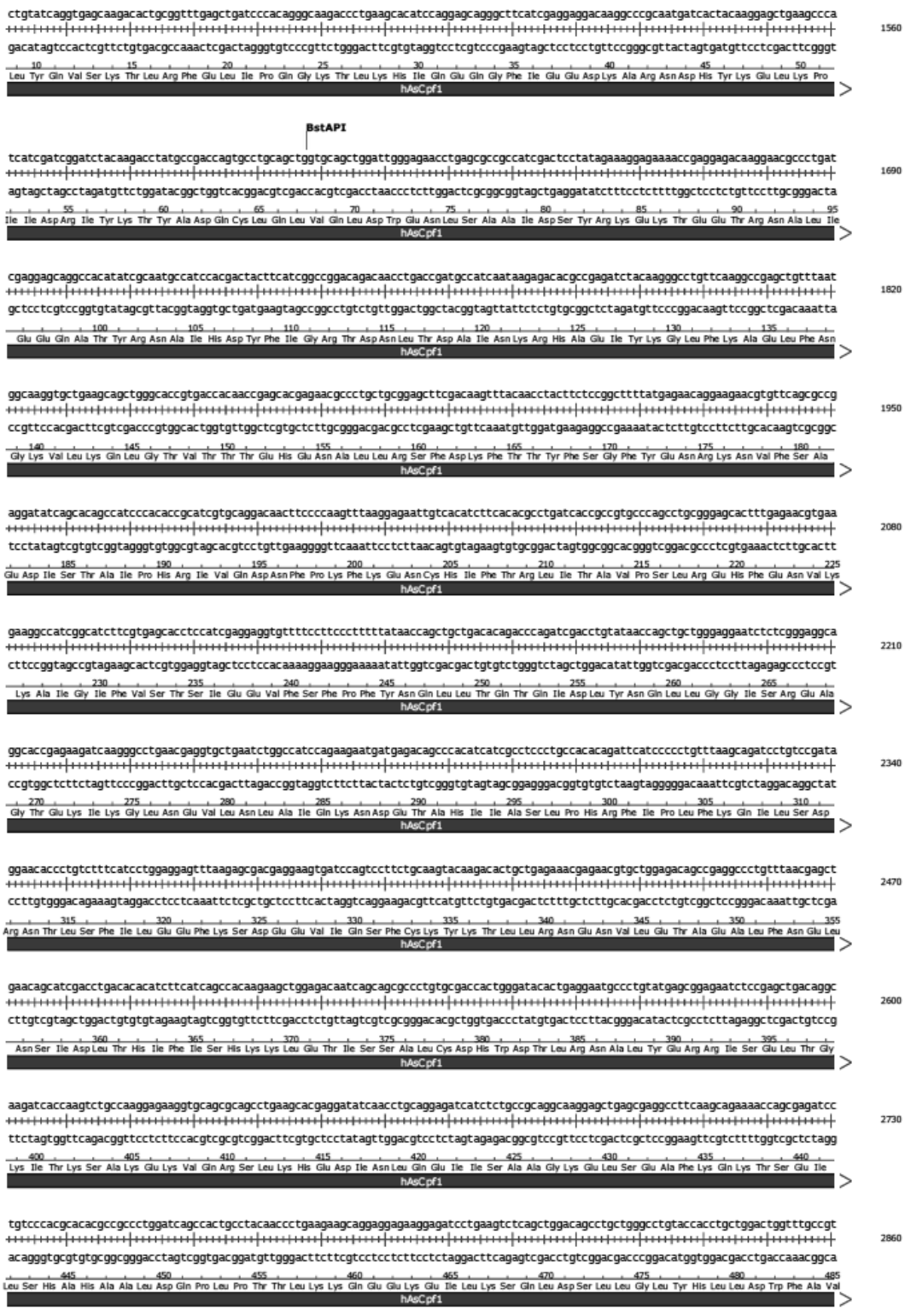


FIG. 18B (continued)



FIG. 18B (continued)

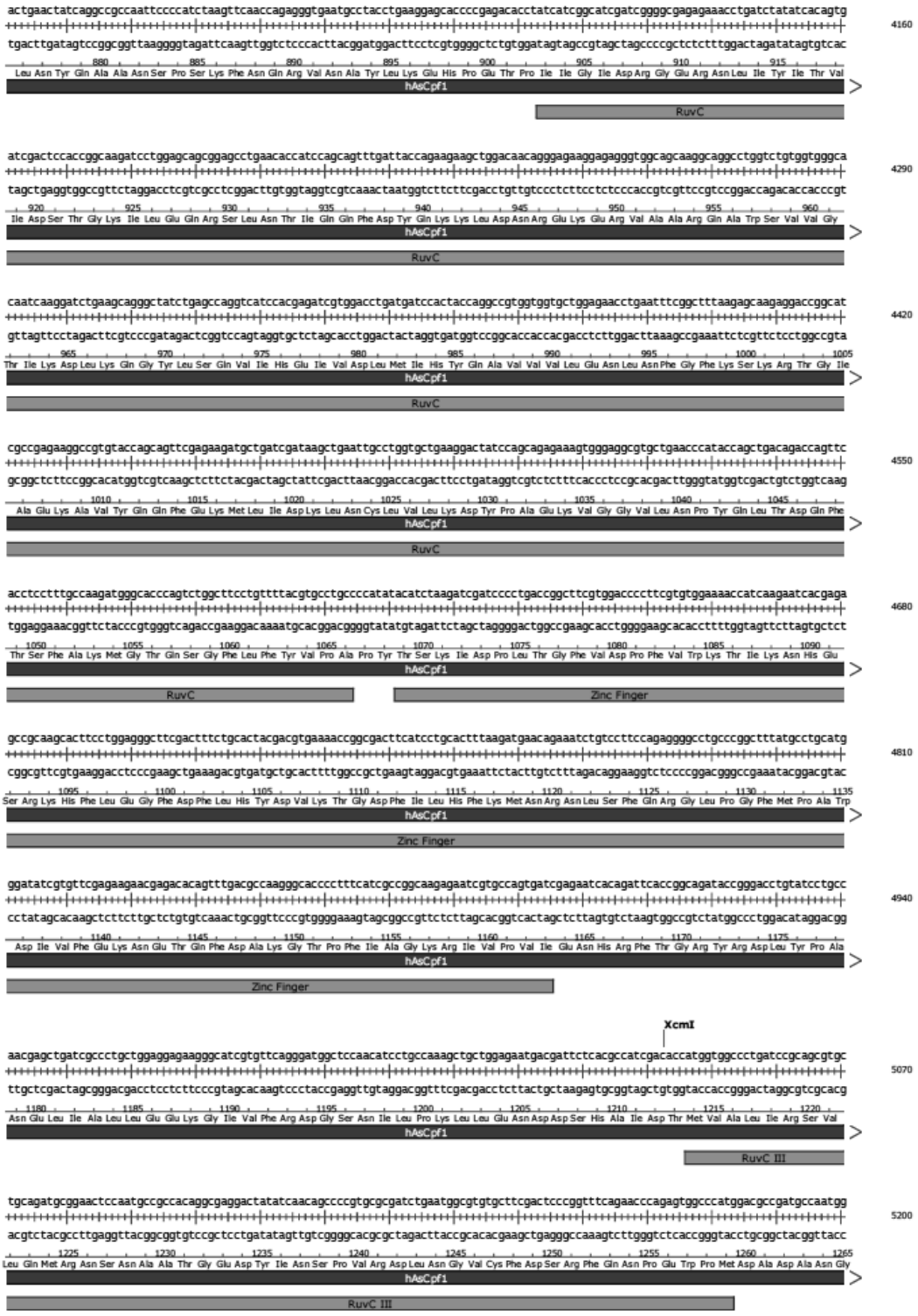


FIG. 18B (continued)



FIG. 18B (continued)



FIG. 18B (continued)

B

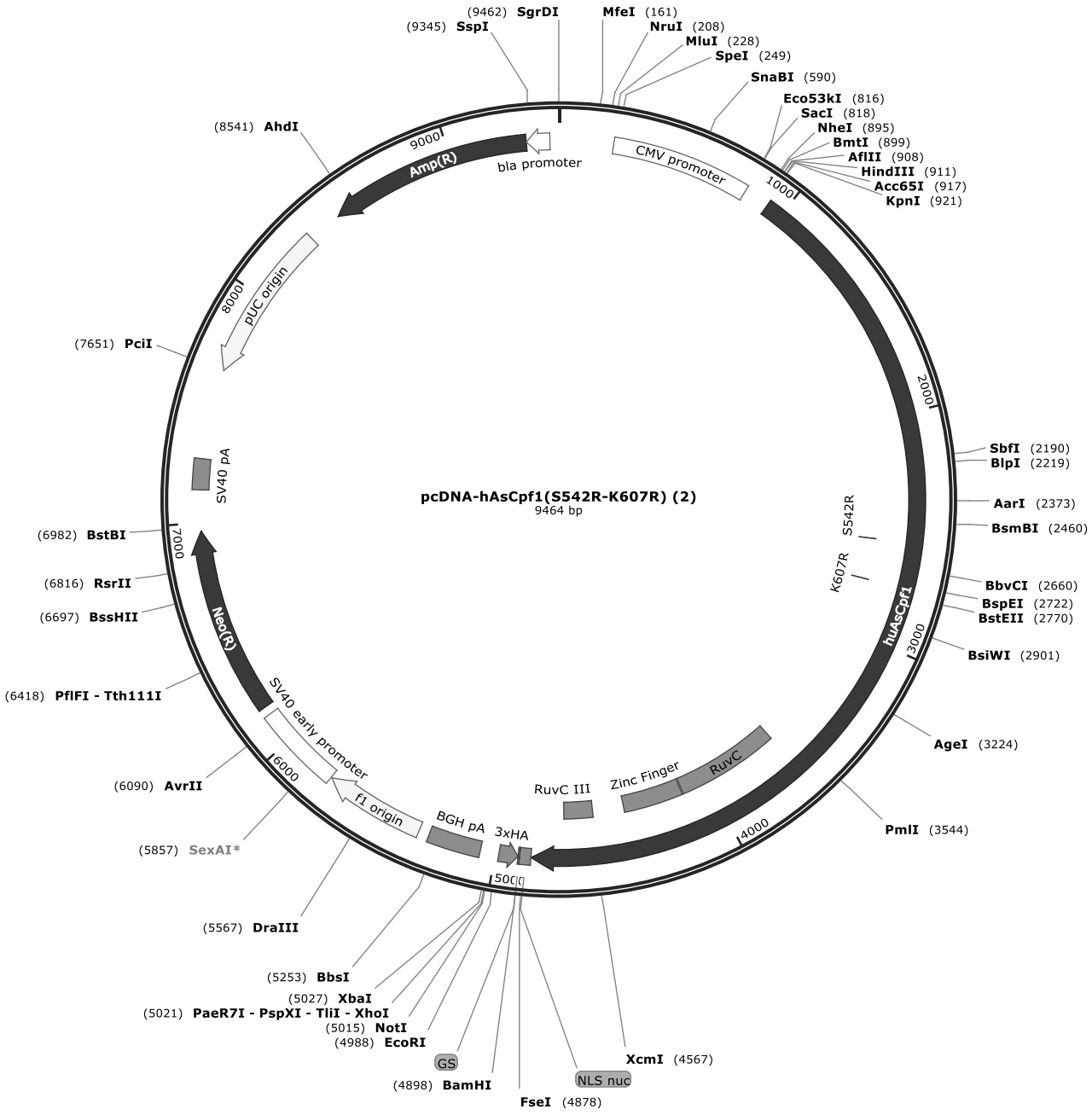

pcDNA-hAsCpf1(S542R-K607R) (2)
9464 bp
(9462) SgrDI
(9345) SspI
MfeI (161)
NruI (208)
MluI (228)
SpeI (249)
SnaBI (590)
Eco53kI (816)
SacI (818)
NheI (895)
BmtI (899)
AflII (908)
HindIII (911)
Acc65I (917)
KpnI (921)
(8541) AhdI
9000
Amp(R)
bla promoter
CMV promoter
1000
pUC origin
8000
(7651) PciI
SV40 pA
2000
SbfI (2190)
BlpI (2219)
AarI (2373)
BsmBI (2460)
S542R
K607R
BbvCI (2660)
BspEI (2722)
BstEII (2770)
huAsCpf1
3000
BsiWI (2901)
(6982) BstBI
7000
(6816) RsrII
(6697) BssHII
Neo(R)
(6418) PflFI - Tth111I
SV40 early promoter
AgeI (3224)
(6090) AvrII
6000
F1 origin
RuvC III
Zinc Finger
RuvC
(5857) SexAI*
BGH pA
3xHA
4000
PmlI (3544)
5000
(5567) DraIII
(5253) BbsI
(5027) XbaI
(5021) PaeR7I - PspXI - TliI - XhoI
(5015) NotI
(4988) EcoRI
GS
(4898) BamHI
FseI (4878)
NLS nuc
XcmI (4567)

FIG. 18C

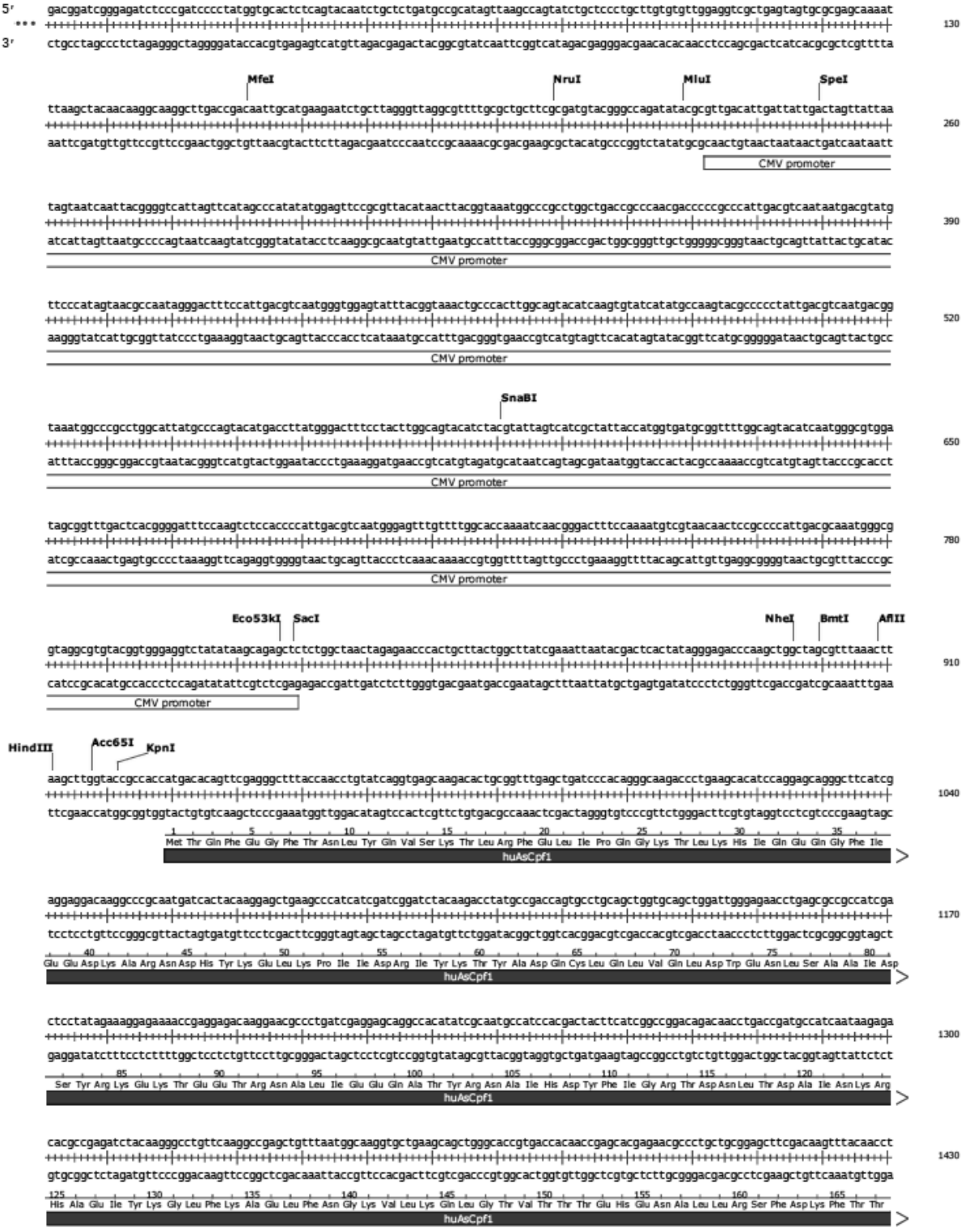

5′
3′
MfeI
NruI
MluI
SpeI
CMV promoter
SnaBI
Eco53kI
SacI
NheI
BmtI
AflII
HindIII
Acc65I
KpnI
huAsCpf1

FIG. 18D

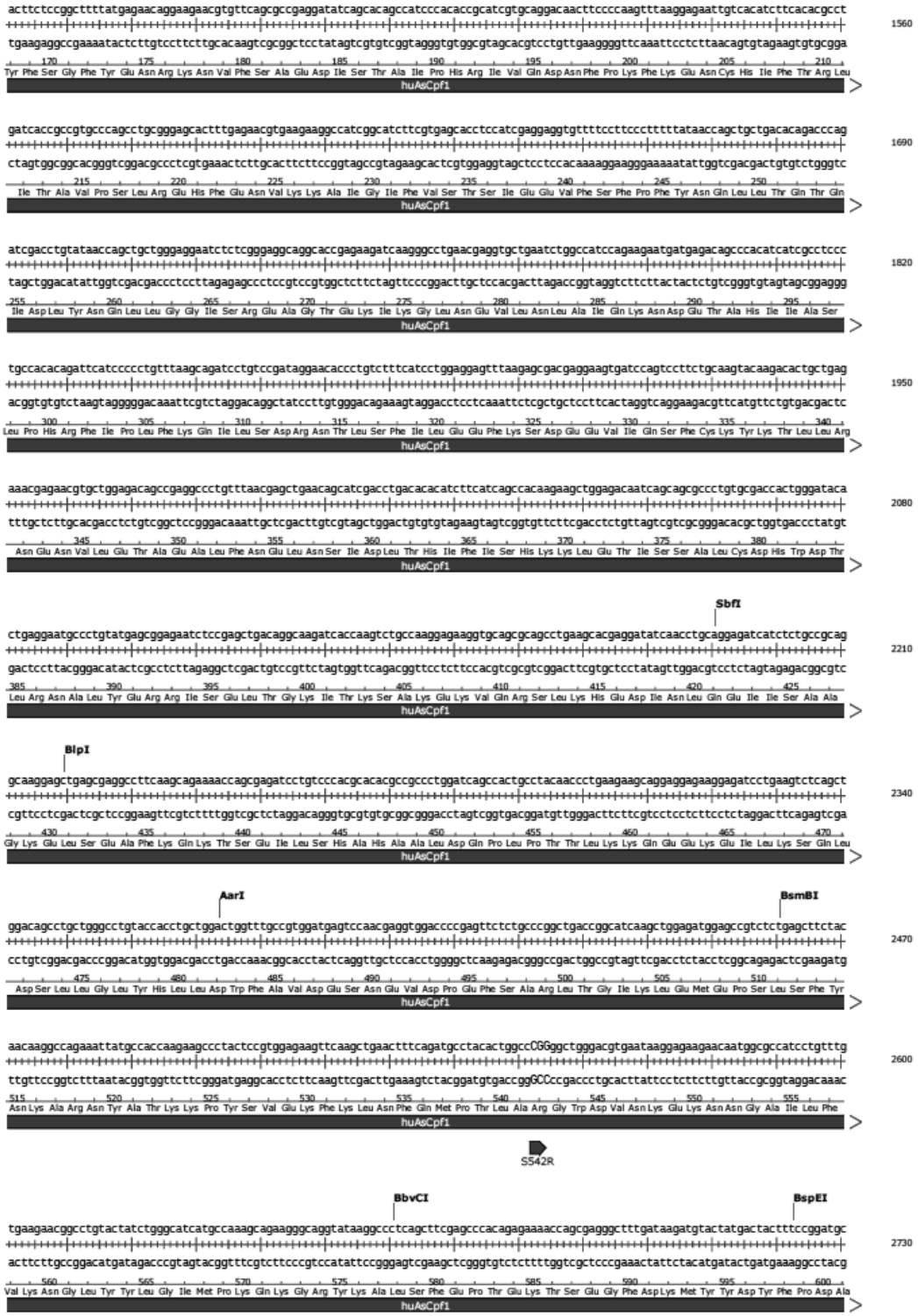


FIG. 18D (continued)

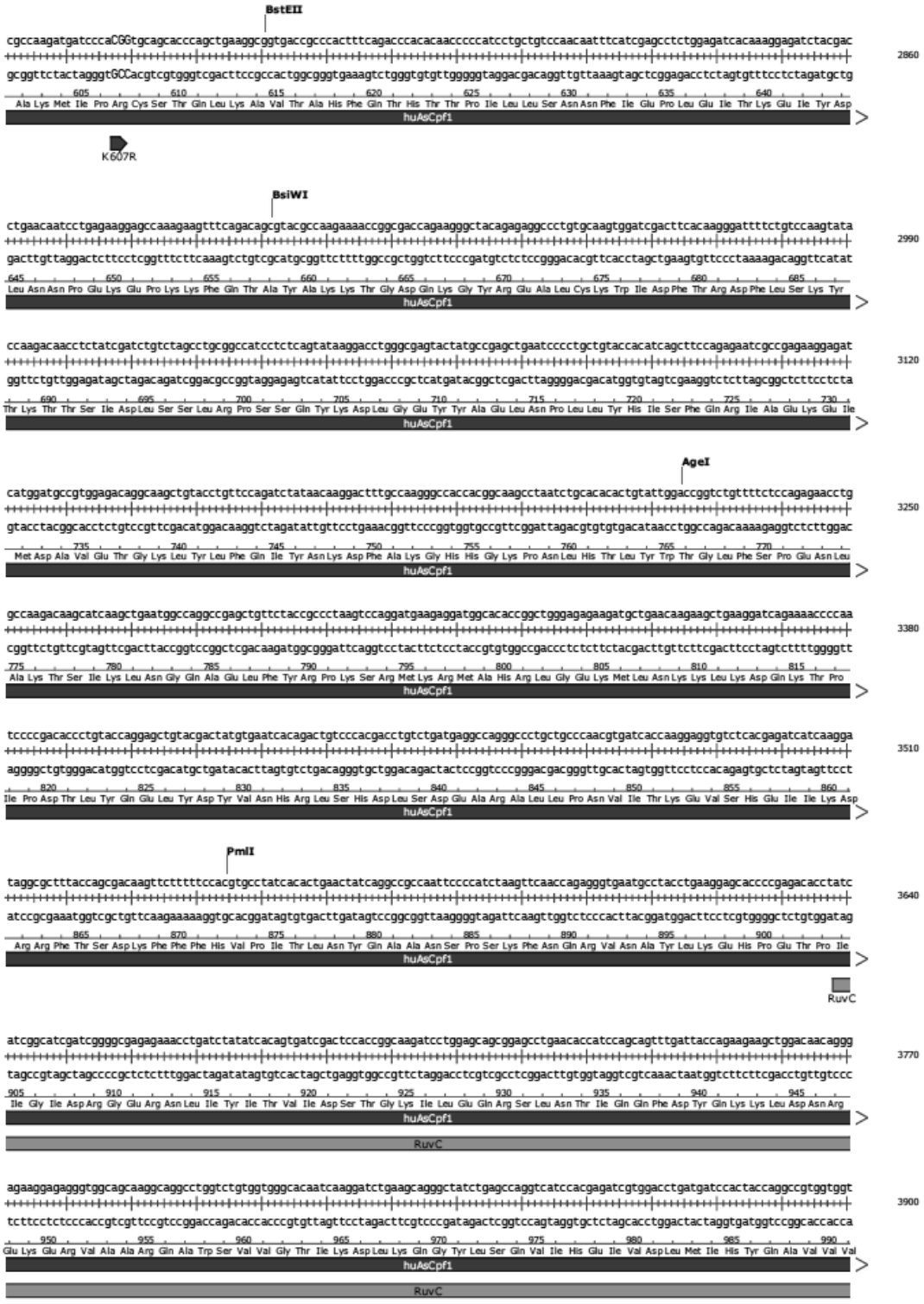


FIG. 18D (continued)

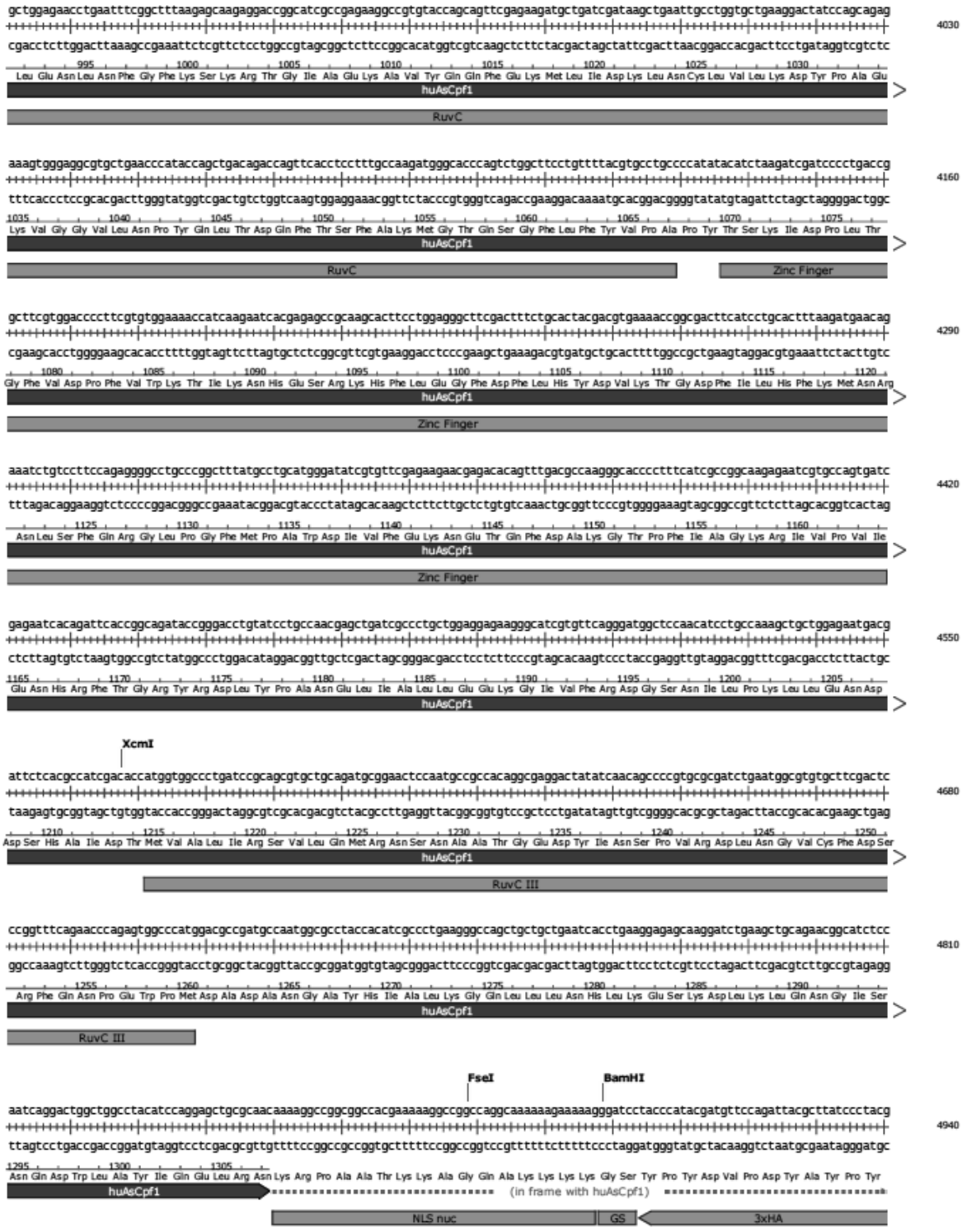


FIG. 18D (continued)

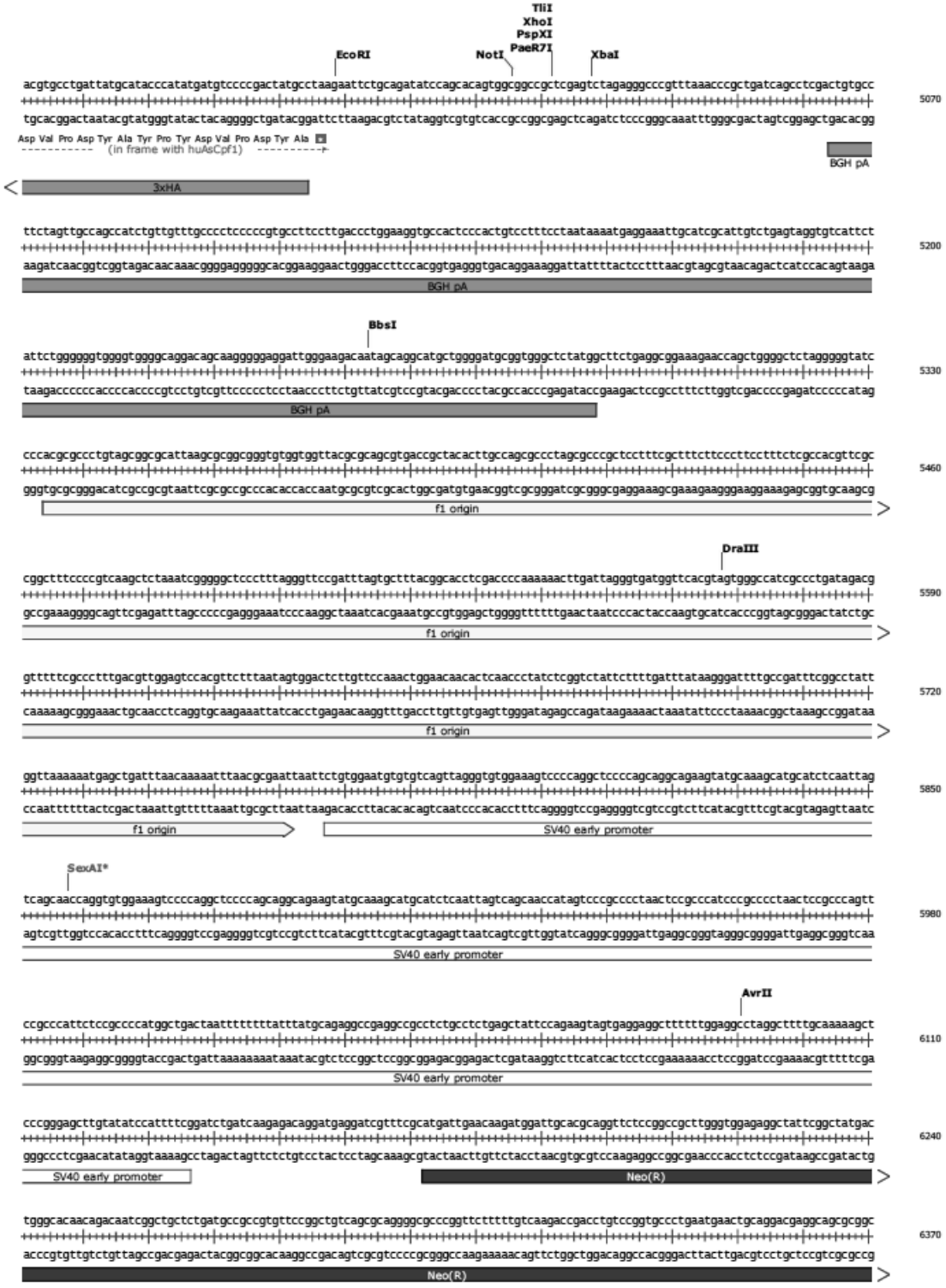


FIG. 18D (continued)



FIG. 18D (continued)



FIG. 18D (continued)

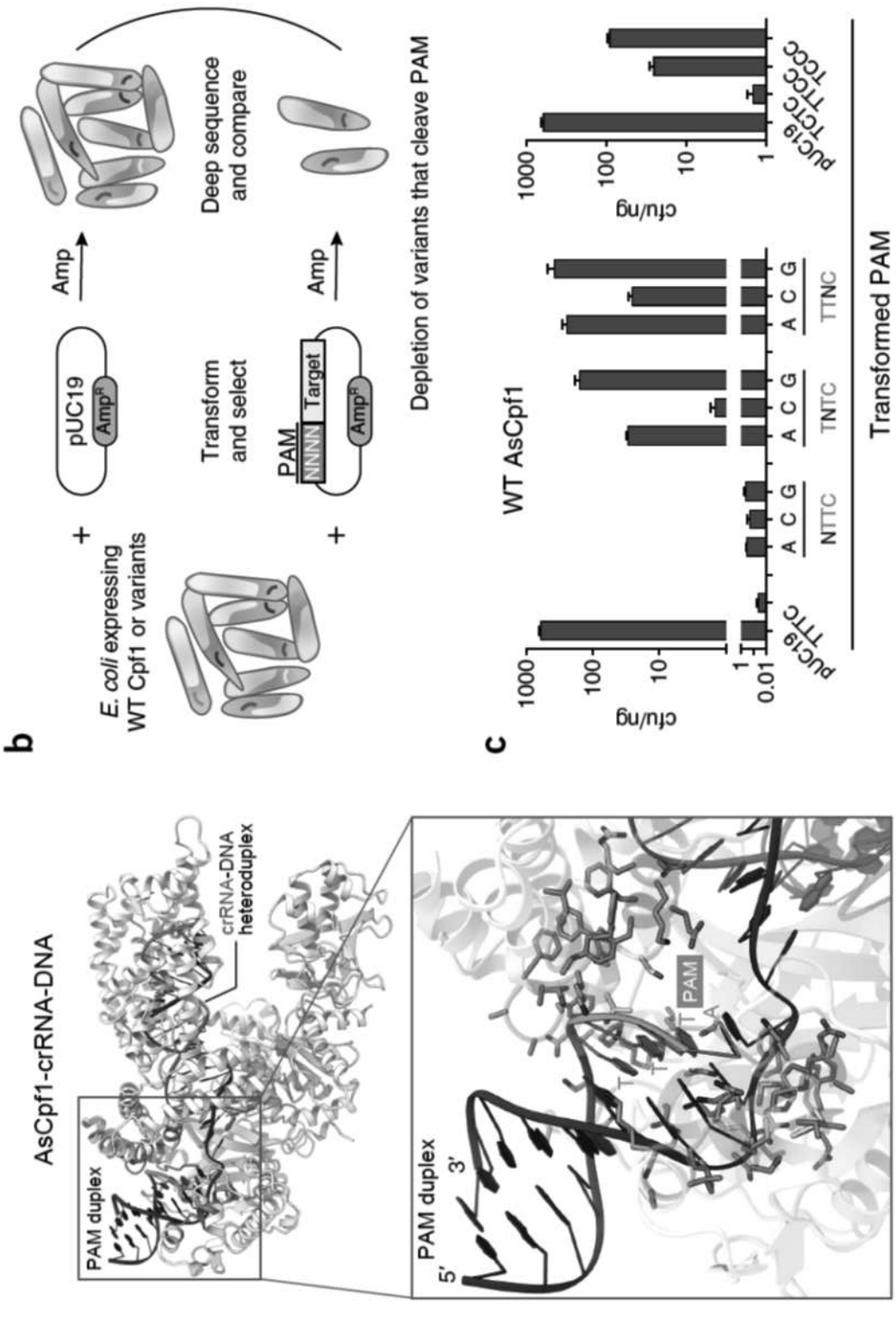

a
AsCpf1-crRNA-DNA
PAM duplex
crRNA-DNA heteroduplex
5′
3′
PAM duplex
PAM
b
pUC19
Amp^R
Amp
Deep sequence and compare
Transform and select
PAM
NNNN
Target
Amp^R
Amp
+
E. coli expressing WT Cpf1 or variants
+
Depletion of variants that cleave PAM
c
WT AsCpf1
cfu/ng
1000
100
10
1
0.01
pUC19 TTTC
A C G
NTTC
A C G
TNTC
A C G
TTNC
pUC19 TCCC
Transformed PAM

FIG. 19A-19C

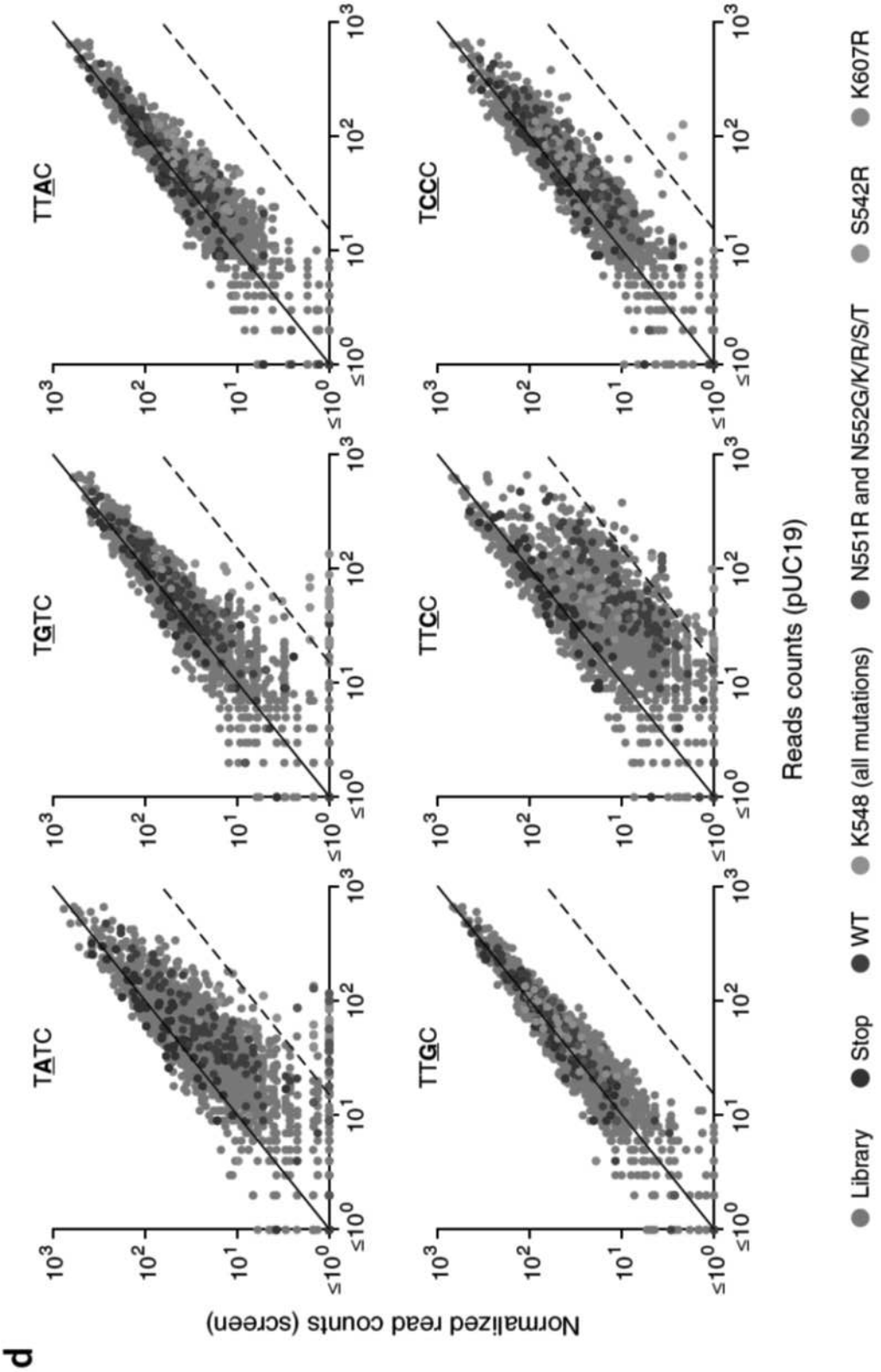

d
Normalized read counts (screen)
TATC
TGTC
TTAC
TTGC
TTCC
TCCC
Reads counts (pUC19)
Library
Stop
WT
K548 (all mutations)
N551R and N552G/K/R/S/T
S542R
K607R

FIG. 19D

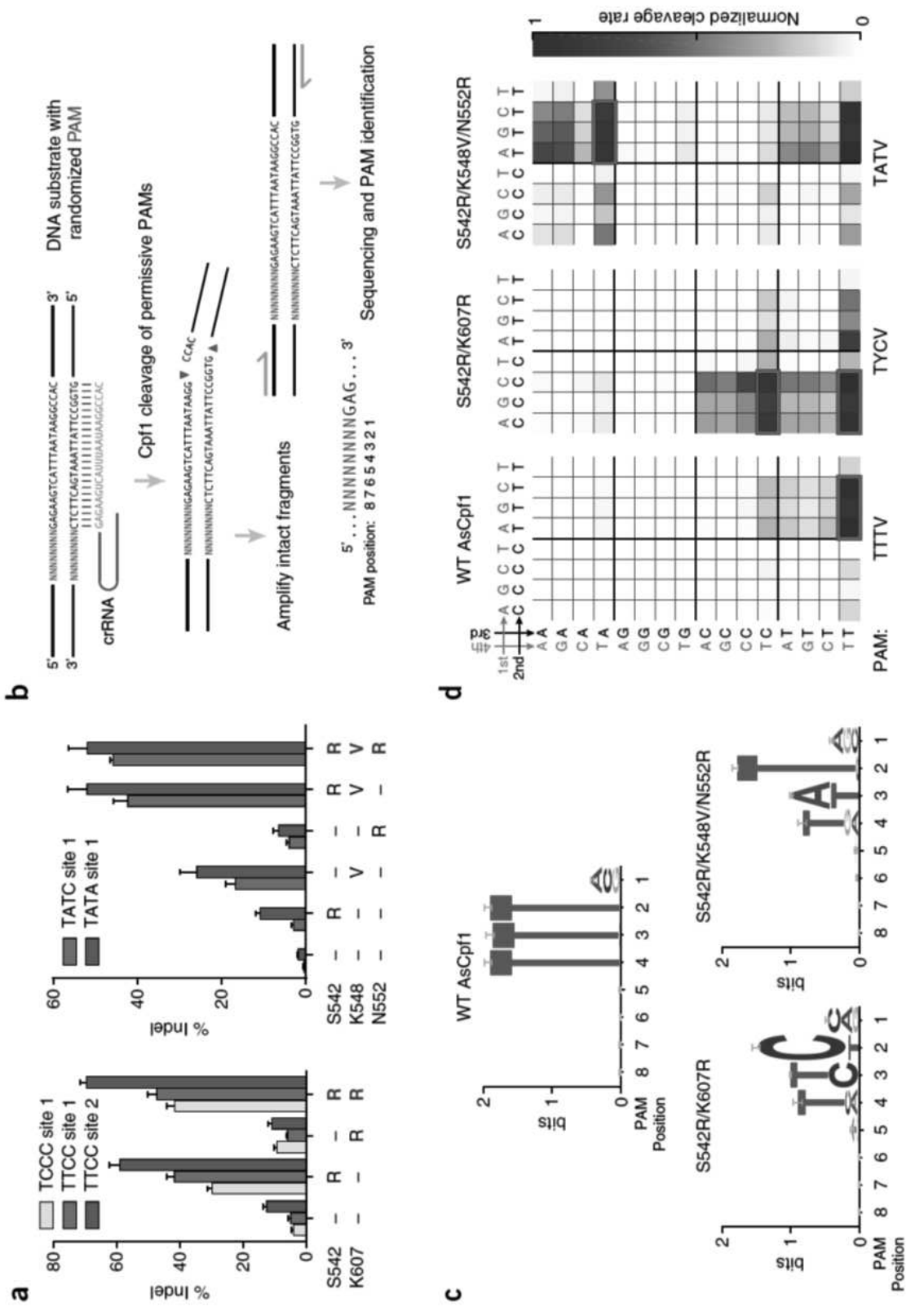

a
TCCC site 1
TTCC site 1
TTCC site 2
% Indel
S542
K607
TATC site 1
TATA site 1
S542
K548
N552
b
DNA substrate with randomized PAM
crRNA
Cpf1 cleavage of permissive PAMs
Amplify intact fragments
Sequencing and PAM identification
PAM position: 8 7 6 5 4 3 2 1
c
WT AsCpf1
S542R/K607R
S542R/K548V/N552R
bits
PAM Position
d
WT AsCpf1
S542R/K607R
S542R/K548V/N552R
PAM:
TTTV
TYCV
TATV
Normalized cleavage rate

FIG. 20A-20D

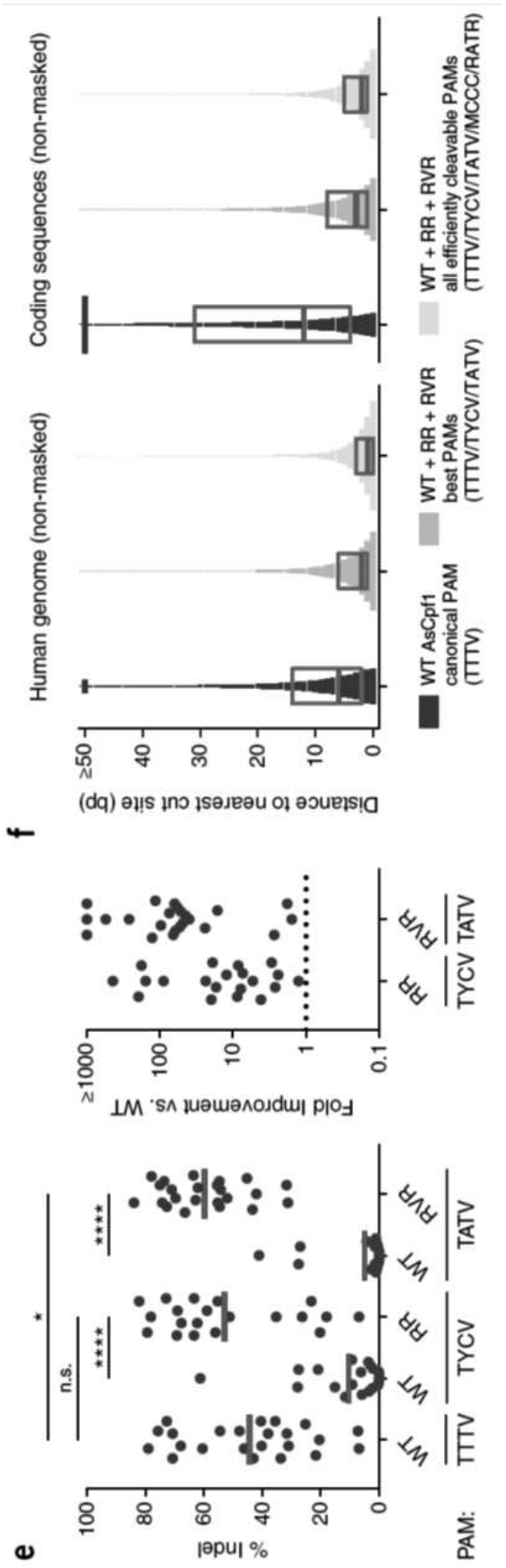

e
% Indel
PAM:
TTTV
TYCV
TATV
WT
RR
RVR
n.s.
*
****
Fold improvement vs. WT
f
Distance to nearest cut site (bp)
Human genome (non-masked)
Coding sequences (non-masked)
WT AsCpf1 canonical PAM (TTTV)
WT + RR + RVR best PAMs (TTTV/TYCV/TATV)
WT + RR + RVR all efficiently cleavable PAMs (TTTV/TYCV/TATV/MCCC/RATR)

FIG. 20E-20F

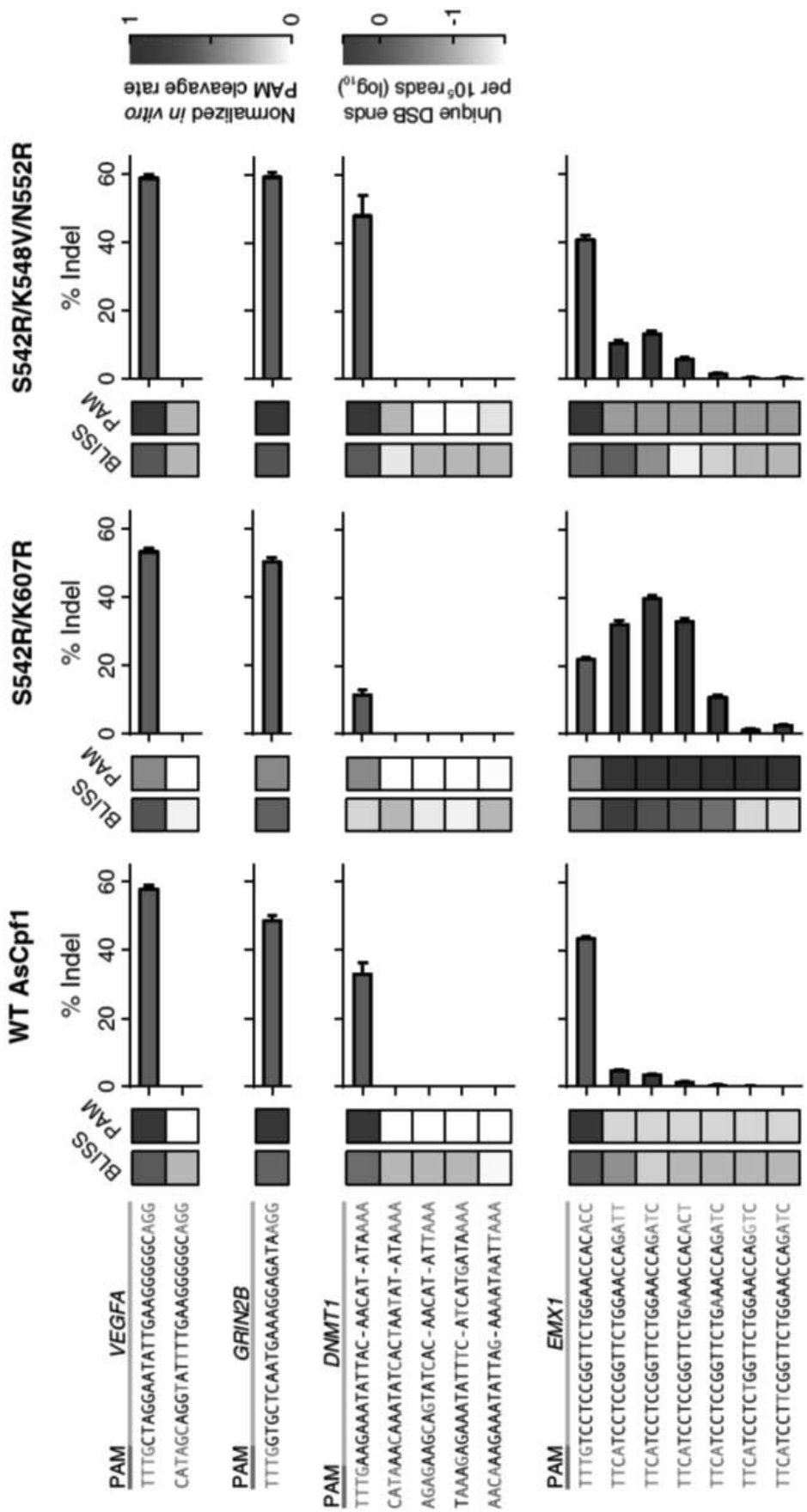

WT AsCpf1
S542R/K607R
S542R/K548V/N552R
% Indel
0
20
40
60
BLISS
PAM
Normalized in vitro
PAM cleavage rate
1
0
Unique DSB ends
per 10^5 reads (log10)
0
-1
PAM
VEGFA
TTTGCTAGGAATATTGAAGGGGGCAGG
CATAGCAGGTATTTTGAAGGGGGCAGG
PAM
GRIN2B
TTTGGTGCTCAATGAAAGGAGATAAGG
PAM
DNMT1
TTTGAAGAAATATTAC-AACAT-ATAAAA
CATAAACAAATATCACTAATAT-ATAAAA
AGAGAAGCAGTATCAC-AACAT-ATTAAA
TAAAGAGAAATATTTTC-ATCATGATAAAA
AACAAAGAAATATTAG-AAAATAATTAAA
PAM
EMX1
TTTGTCCTCCGGTTCTGGAACCACACC
TTCATCCTCCGGTTCTGGAACCAGATT
TTCATCCTCCGGTTCTGGAACCAGATC
TTCATCCTCCGGTTCTGGAACCACACT
TTCATCCTCCGGTTCTGAAACCAGATC
TTCATCCTCCGGTTCTGAAACCAGATC
TTCATCCTCTGGTTCTGGAACCAGGTC
TTCATCCTTCGGTTCTGGAACCAGATC

FIG. 21A

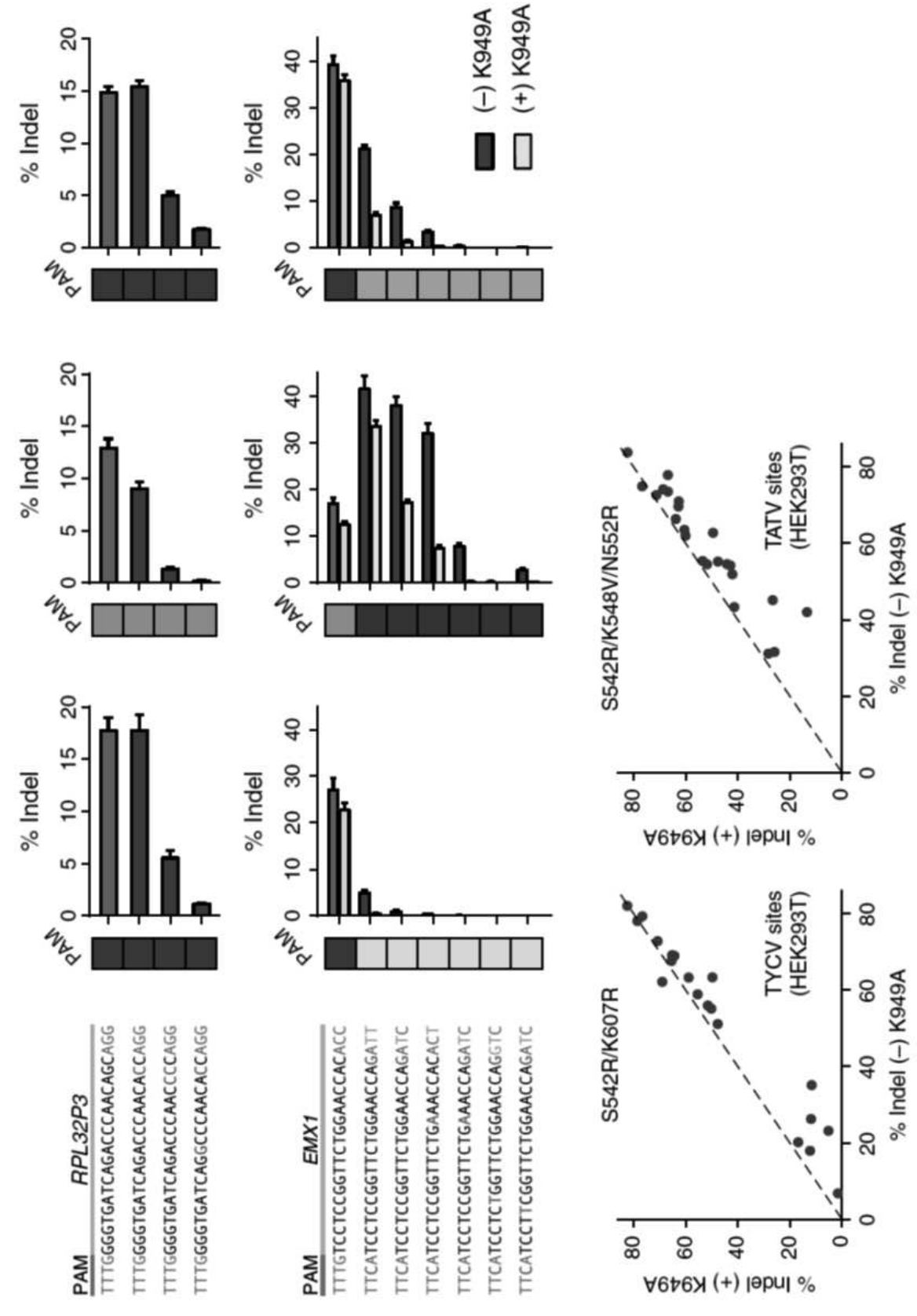

b
c
d
PAM
RPL32P3
EMX1
% Indel
(–) K949A
(+) K949A
S542R/K607R
S542R/K548V/N552R
TYCV sites (HEK293T)
TATV sites (HEK293T)
% Indel (+) K949A
% Indel (–) K949A

FIG. 21B-21D

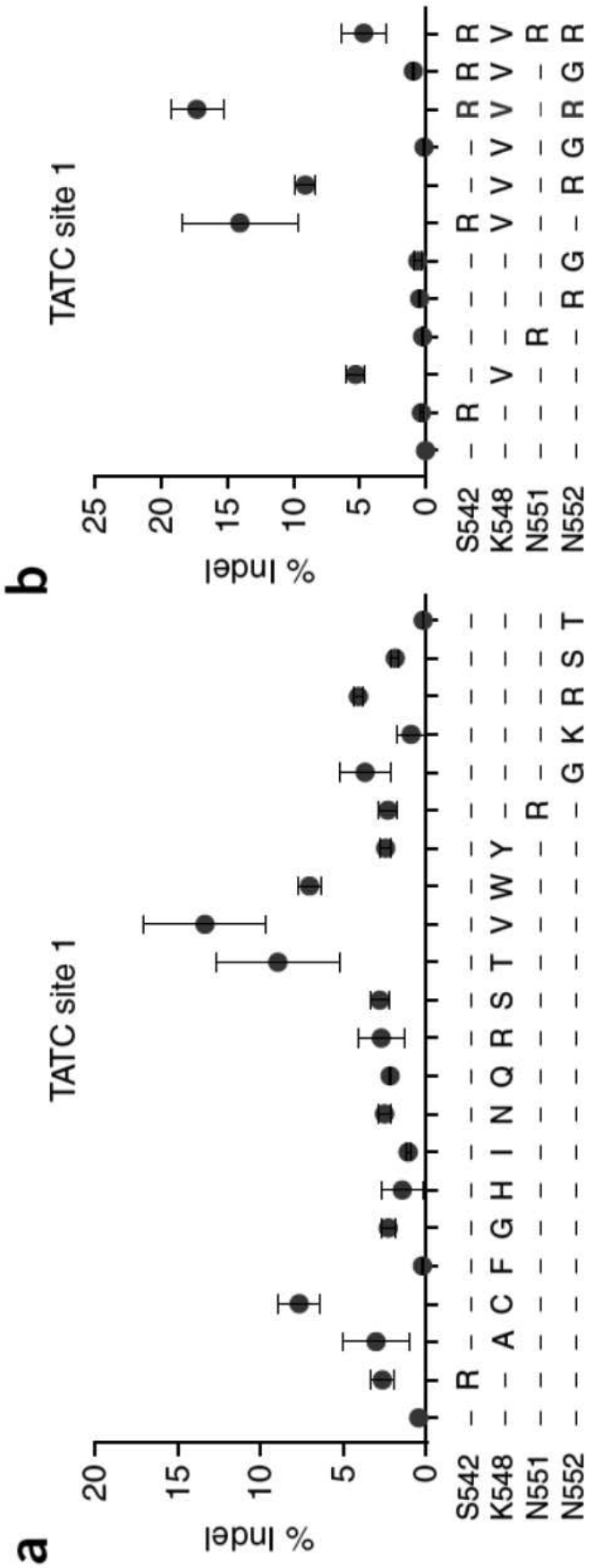

a
TATC site 1
% Indel
20
15
10
5
0
S542 – R –
K548 – – A C F G H I N Q R S T V W Y – – – – – –
N551 – – – – – – – – – – – – R – – – – – – – – –
N552 – – – – – – – – – – – – – – – – G K R S T
b
TATC site 1
% Indel
25
20
15
10
5
0
S542 – R – – – – R – – – R R R
K548 – – V – – – V V V V V V
N551 – – – – – – R – – – – – R
N552 – – – – – – R G – R G R R

FIG. 22A-22B

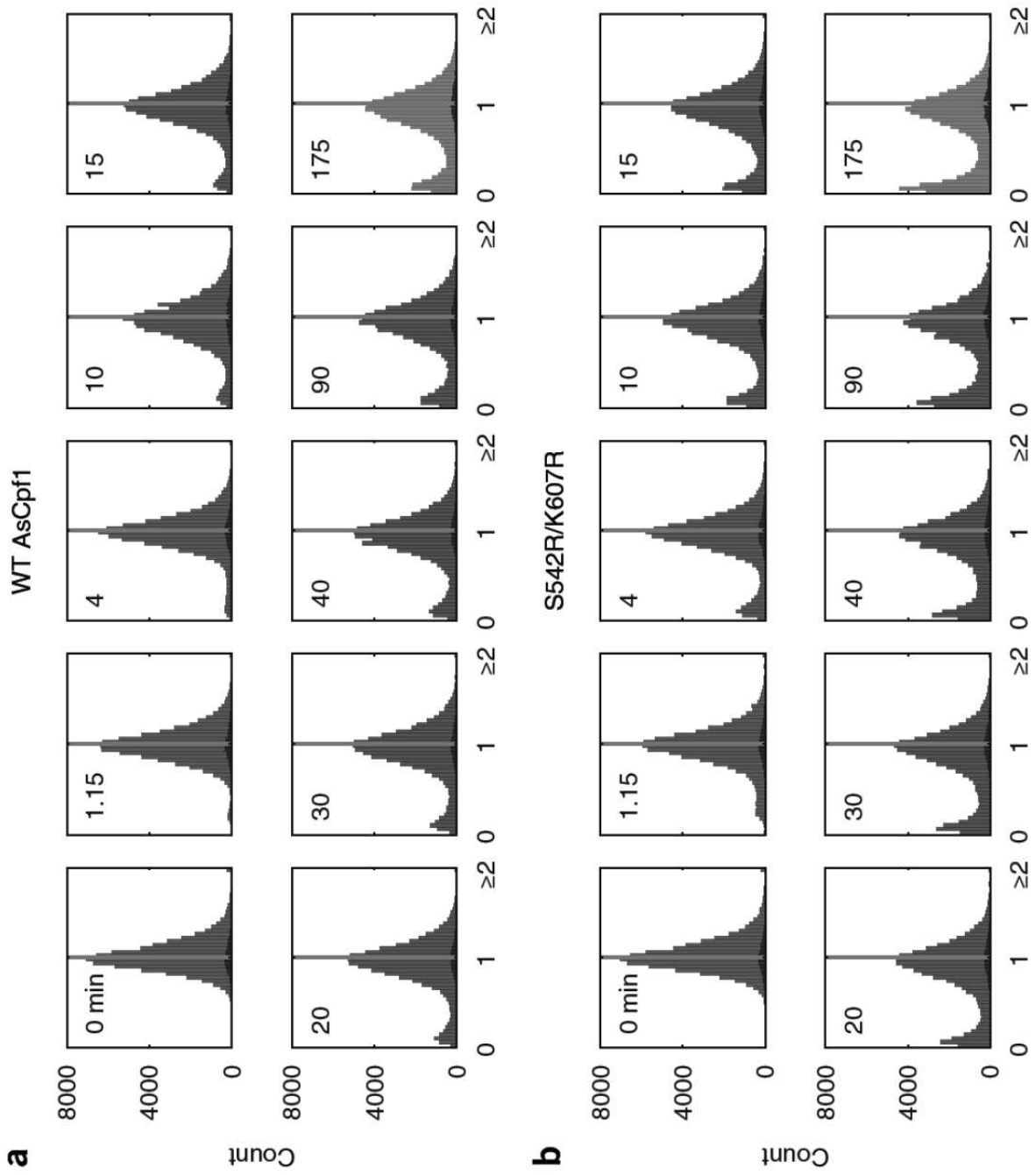

a
WT AsCpf1
b
S542R/K607R
Count
8000
4000
0
0 min
1.15
4
10
15
20
30
40
90
175
0
1
≥2

FIG. 23A-23B

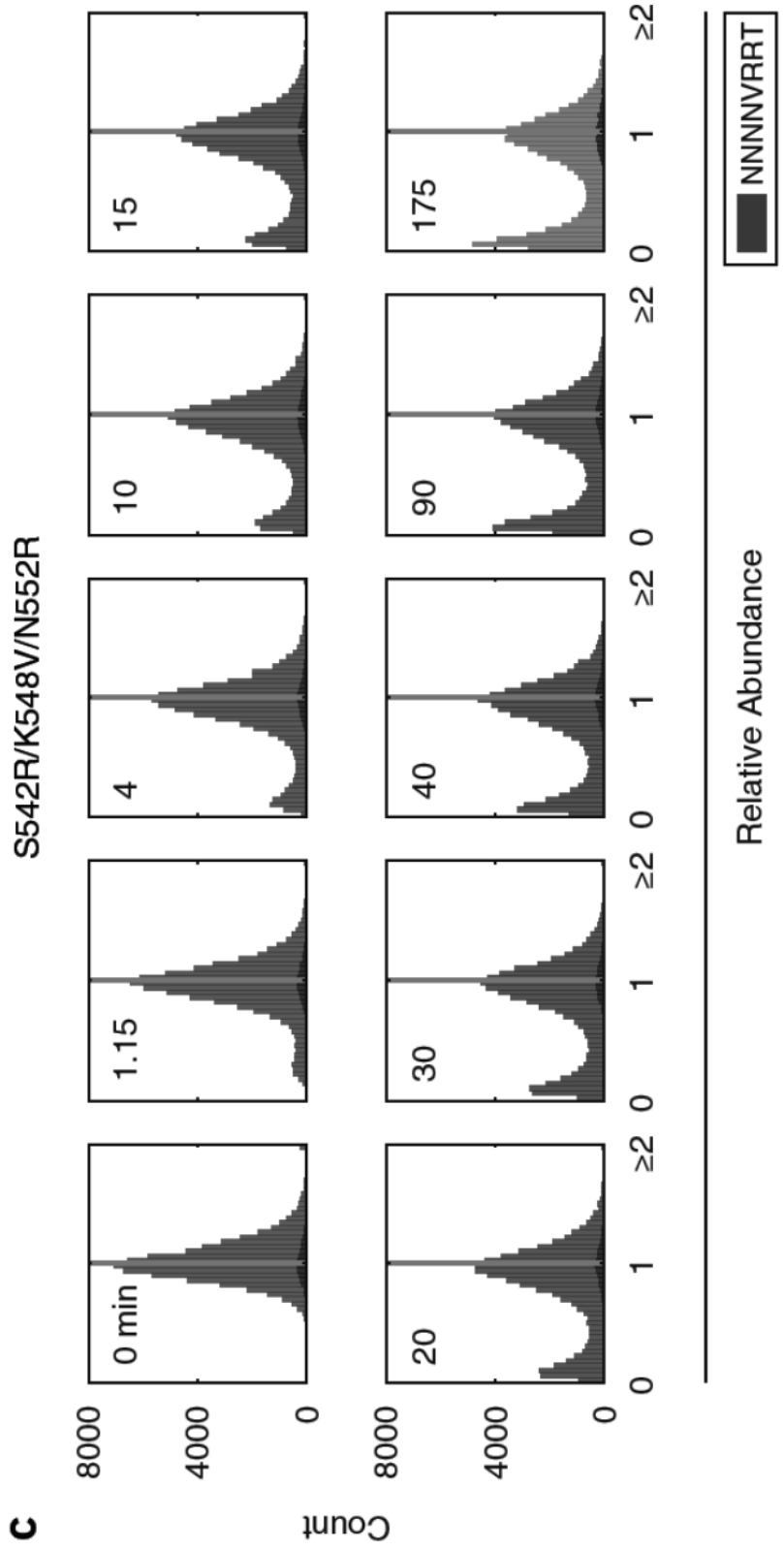

c
S542R/K548V/N552R
Count
8000
4000
0
0 min
1.15
4
10
15
20
30
40
90
175
0
1
≥2
Relative Abundance
NNNNVRRT

FIG. 23C

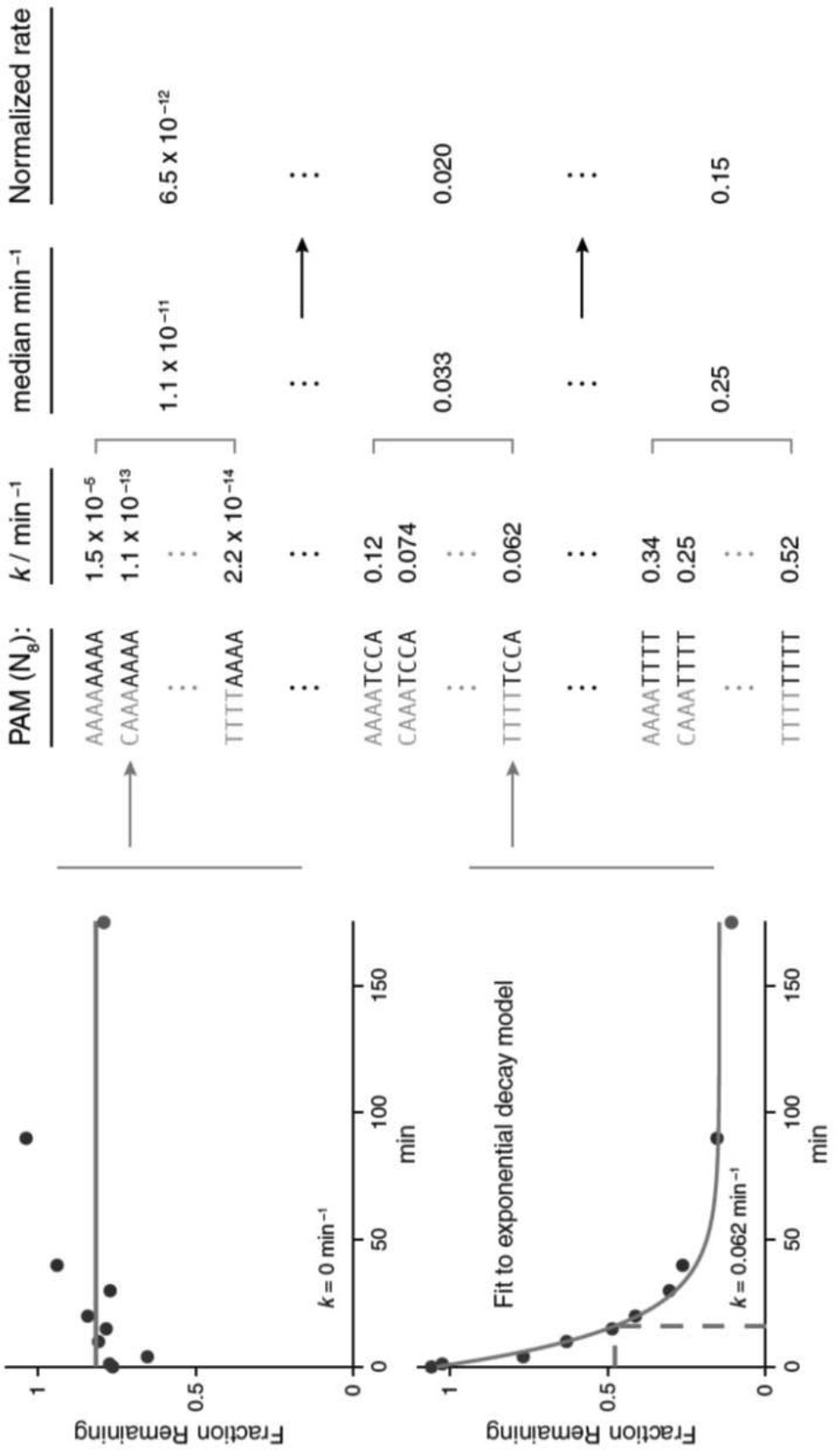

PAM (N8):
k / min−1
median min−1
Normalized rate
AAAAAAAA
CAAAAAAA
TTTTAAAA
AAAATCCA
CAAATCCA
TTTTTCCA
AAAATTTT
CAAATTTT
TTTTTTTT
1.5 x 10−6
1.1 x 10−13
2.2 x 10−14
0.12
0.074
0.062
0.34
0.25
0.52
1.1 x 10−11
0.033
0.25
6.5 x 10−12
0.020
0.15
Fraction Remaining
k = 0 min−1
min
Fit to exponential decay model
k = 0.062 min−1

FIG. 24

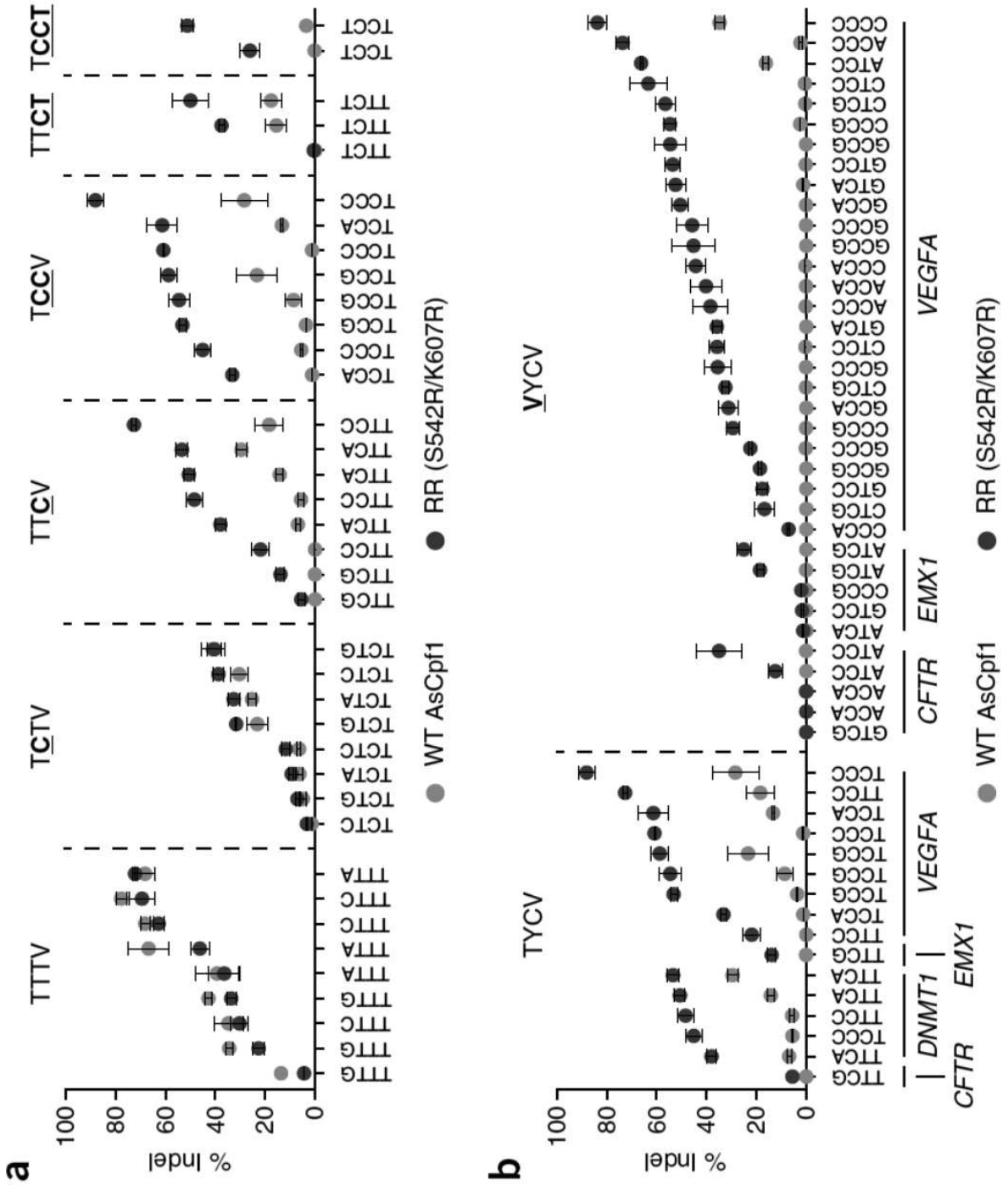

a
% Indel
100
80
60
40
20
0
TTTV
TCTV
TTCV
TCCV
TTCT
TCCT
WT AsCpf1
RR (S542R/K607R)
b
% Indel
TYCV
VYCV
CFTR
DNMT1
EMX1
VEGFA
CFTR
EMX1
CFTR
EMX1
VEGFA
WT AsCpf1
RR (S542R/K607R)

FIG. 25A-25B

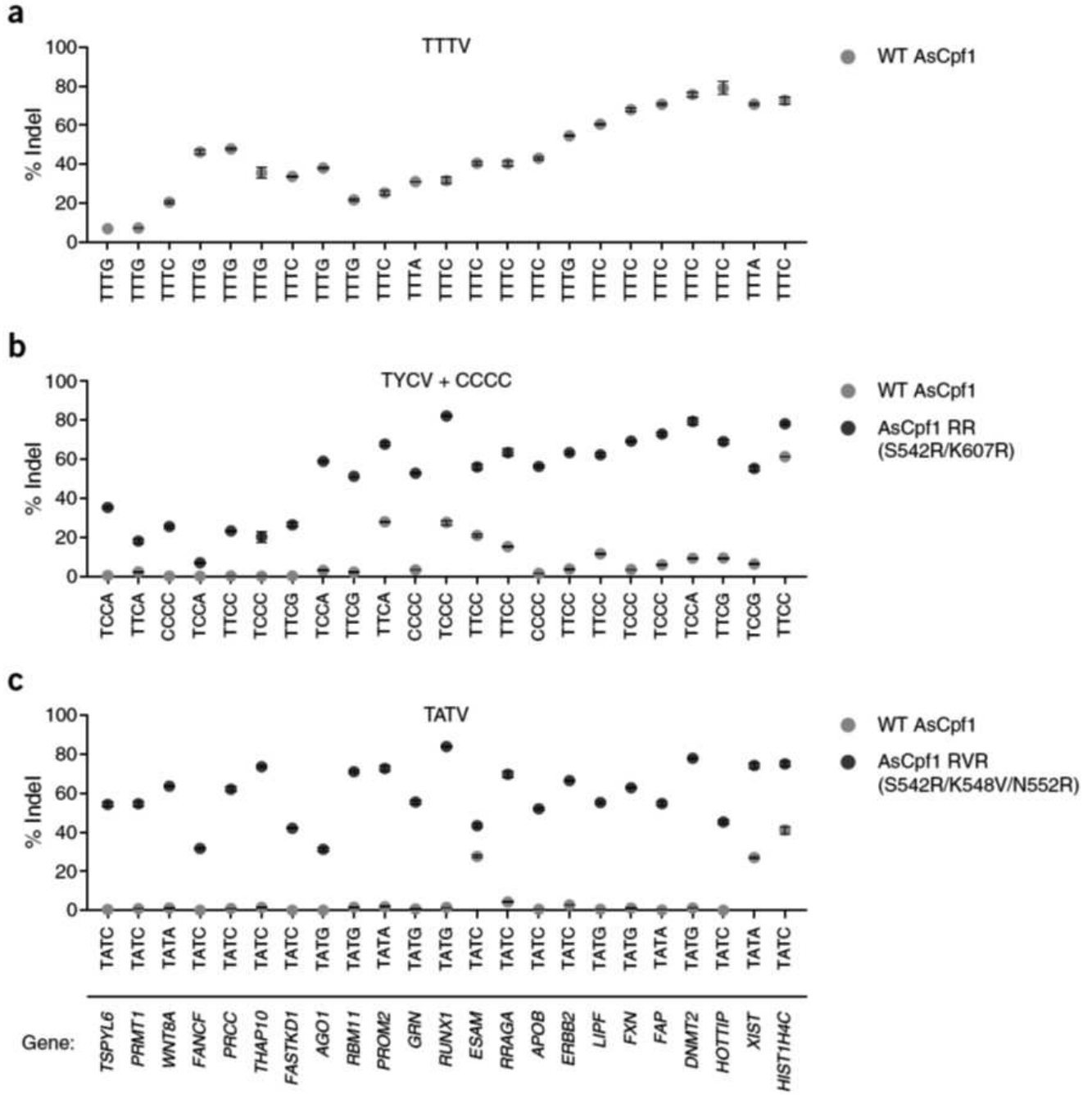

a
TTTV
% Indel
100
80
60
40
20
0
TTTG TTTG TTTC TTTG TTTG TTTG TTTC TTTG TTTG TTTC TTTA TTTC TTTC TTTC TTTC TTTG TTTC TTTC TTTC TTTC TTTC TTTA TTTC
WT AsCpf1
b
TYCV + CCCC
% Indel
100
80
60
40
20
0
TCCA TTCA CCCC TCCA TTCC TCCC TTCG TCCA TTCG TTCA CCCC TCCC TTCC TTCC CCCC TTCC TTCC TCCC TCCC TCCA TTCG TCCG TTCC
WT AsCpf1
AsCpf1 RR
(S542R/K607R)
c
TATV
% Indel
100
80
60
40
20
0
TATC TATC TATA TATC TATC TATC TATC TATG TATG TATA TATG TATG TATC TATC TATC TATC TATG TATG TATA TATG TATC TATA TATC
Gene:
TSPYL6 PRMT1 WNT8A FANCF PRCC THAP10 FASTKD1 AGO1 RBM11 PROM2 GRN RUNX1 ESAM RRAGA APOB ERBB2 LIPF FXN FAP DNMT2 HOTTIP XIST HIST1H4C
WT AsCpf1
AsCpf1 RVR
(S542R/K548V/N552R)

FIG. 26A-26C

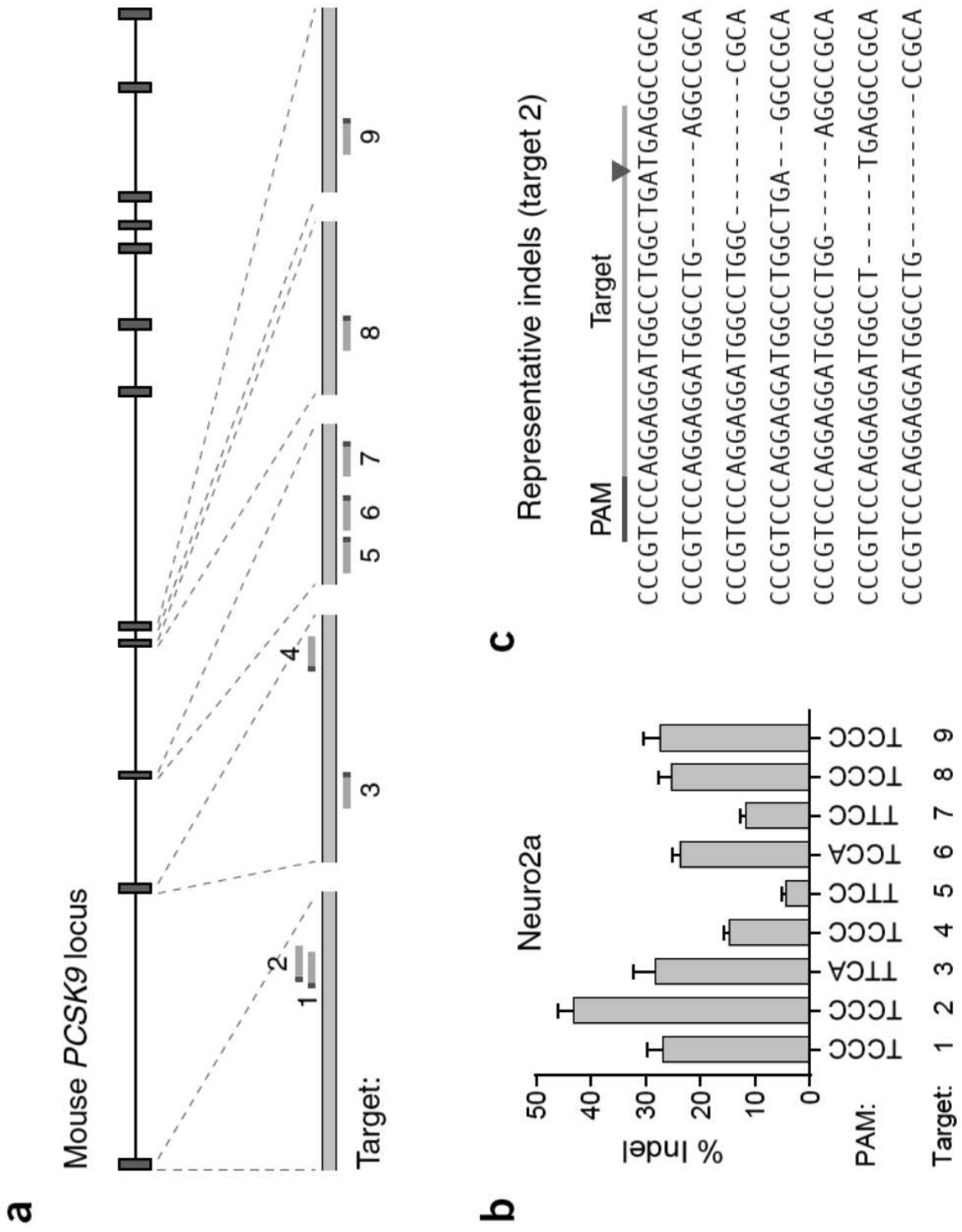

a
Mouse PCSK9 locus
Target:
1
2
3
4
5
6
7
8
9
b
Neuro2a
% Indel
0
10
20
30
40
50
PAM:
Target: 1 2 3 4 5 6 7 8 9
c
Representative indels (target 2)
PAM
Target
CCCGTCCCAGGAGGATGGCCTGGCTGATGAGGCCGCA
CCCGTCCCAGGAGGATGGCCTG-------AGGCCGCA
CCCGTCCCAGGAGGATGGCCTGGC---------CGCA
CCCGTCCCAGGAGGATGGCCTGGCTGA---GGCCGCA
CCCGTCCCAGGAGGATGGCCTGG-----AGGCCGCA
CCCGTCCCAGGAGGATGGCCT------TGAGGCCGCA
CCCGTCCCAGGAGGATGGCCTG----------CCGCA

FIG. 27A-27C

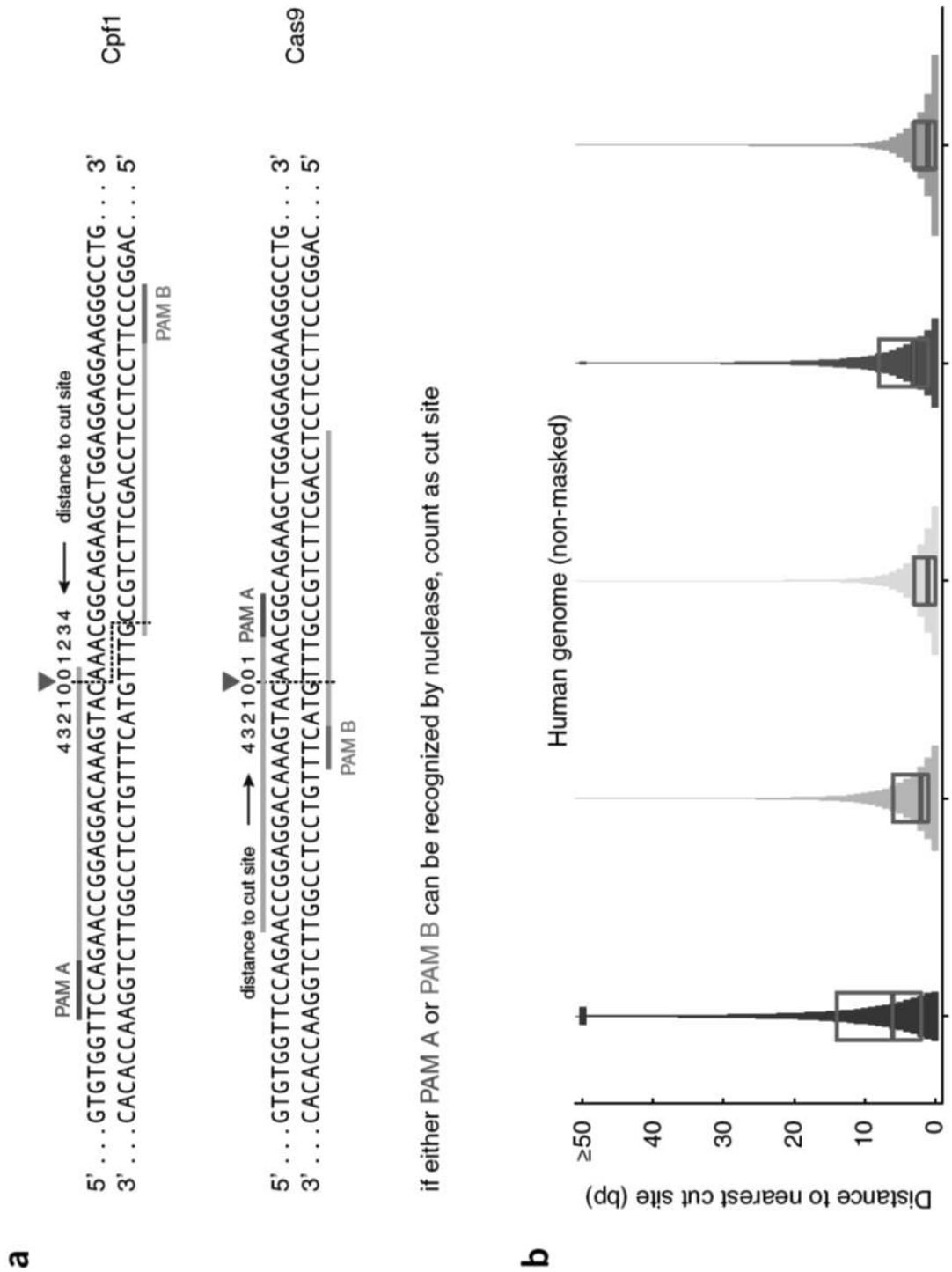

a
Cpf1
PAM A
4 3 2 1 0 0 1 2 3 4
distance to cut site
5' . . . GTGTGGTTCCAGAACCGGAGGACAAAGTAC AAACGGCAGAAGCTGGAGGAGGAAGGGCCTG . . . 3'
3' . . . CACACCAAGGTCTTGGCCTCCTGTTTCATGTTTG CCGTCTTCGACCTCCTCCTTCCCGGAC . . . 5'
PAM B
Cas9
distance to cut site
4 3 2 1 0 0 1
PAM A
5' . . . GTGTGGTTCCAGAACCGGAGGACAAAGTAC AAACGGCAGAAGCTGGAGGAGGAAGGGCCTG . . . 3'
3' . . . CACACCAAGGTCTTGGCCTCCTGTTTCATG TTTGCCGTCTTCGACCTCCTCCTTCCCGGAC . . . 5'
PAM B
if either PAM A or PAM B can be recognized by nuclease, count as cut site
b
Human genome (non-masked)
Distance to nearest cut site (bp)
≥50
40
30
20
10
0

FIG. 28A-28B

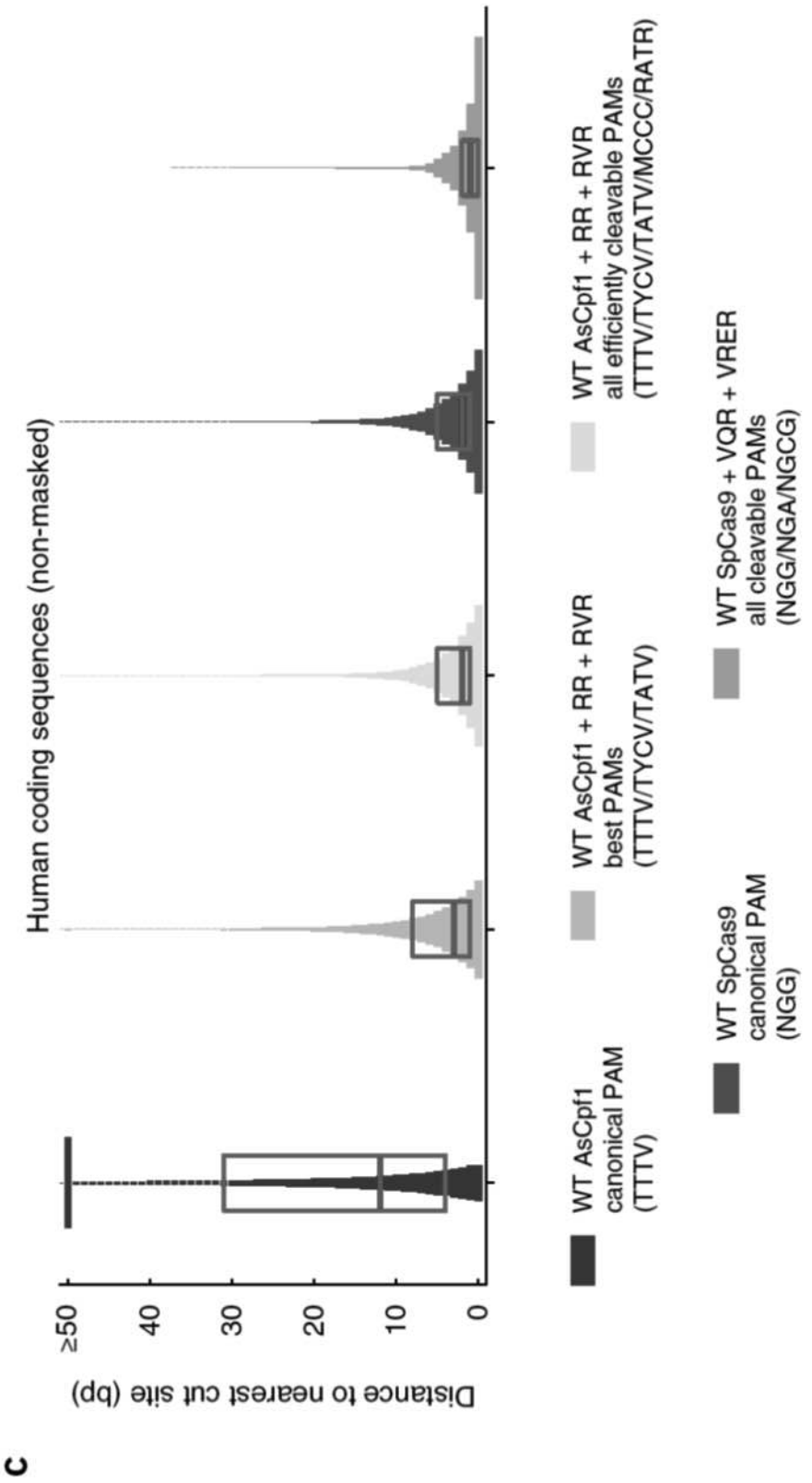

c
Distance to nearest cut site (bp)
≥50
40
30
20
10
0
Human coding sequences (non-masked)
WT AsCpf1 canonical PAM (TTTV)
WT AsCpf1 + RR + RVR best PAMs (TTTV/TYCV/TATV)
WT AsCpf1 + RR + RVR all efficiently cleavable PAMs (TTTV/TYCV/TATV/MCCC/RATR)
WT SpCas9 canonical PAM (NGG)
WT SpCas9 + VQR + VRER all cleavable PAMs (NGG/NGA/NGCG)

FIG. 28C

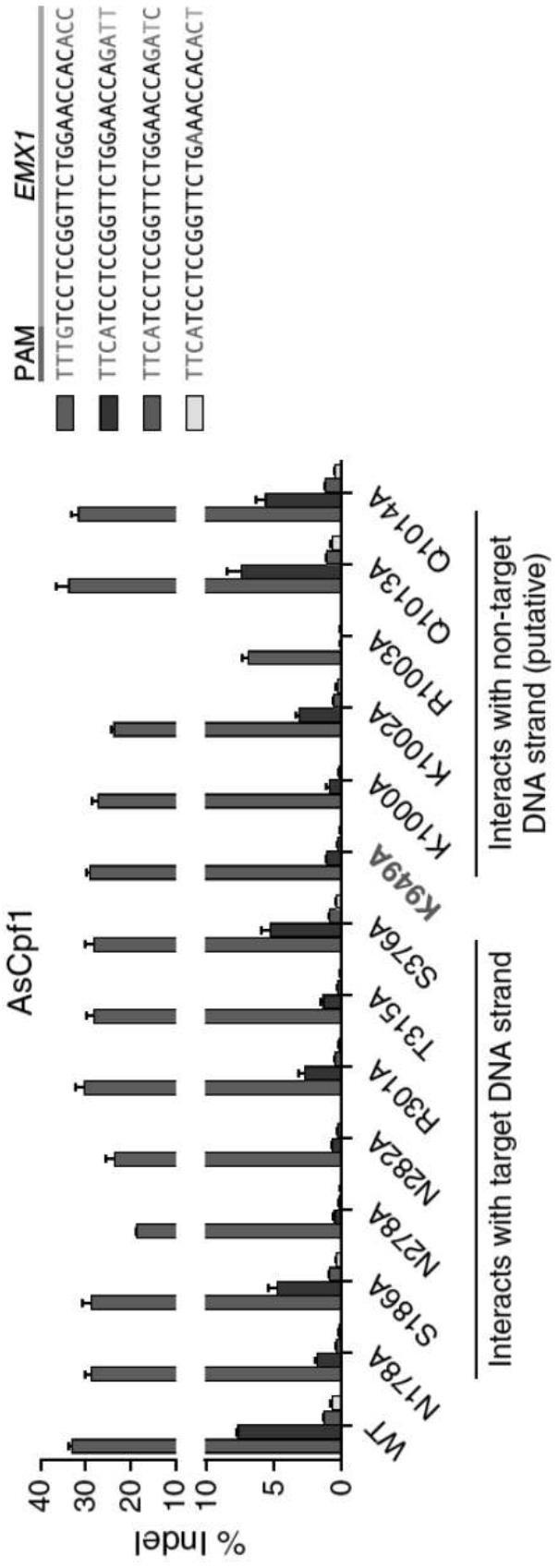

PAM
EMX1
TTTGTCCTCCGGTTCTGGAACCACACC
TTCATCCTCCGGTTCTGGAACCAGATT
TTCATCCTCCGGTTCTGGAACCAGATC
TTCATCCTCCGGTTCTGGAAACCACACT
AsCpf1
% Indel
40
30
20
10
10
5
0
WT
N178A
S186A
N278A
N282A
R301A
T315A
S376A
K949A
K1000A
K1002A
R1003A
Q1013A
Q1014A
Interacts with target DNA strand
Interacts with non-target DNA strand (putative)

FIG. 29

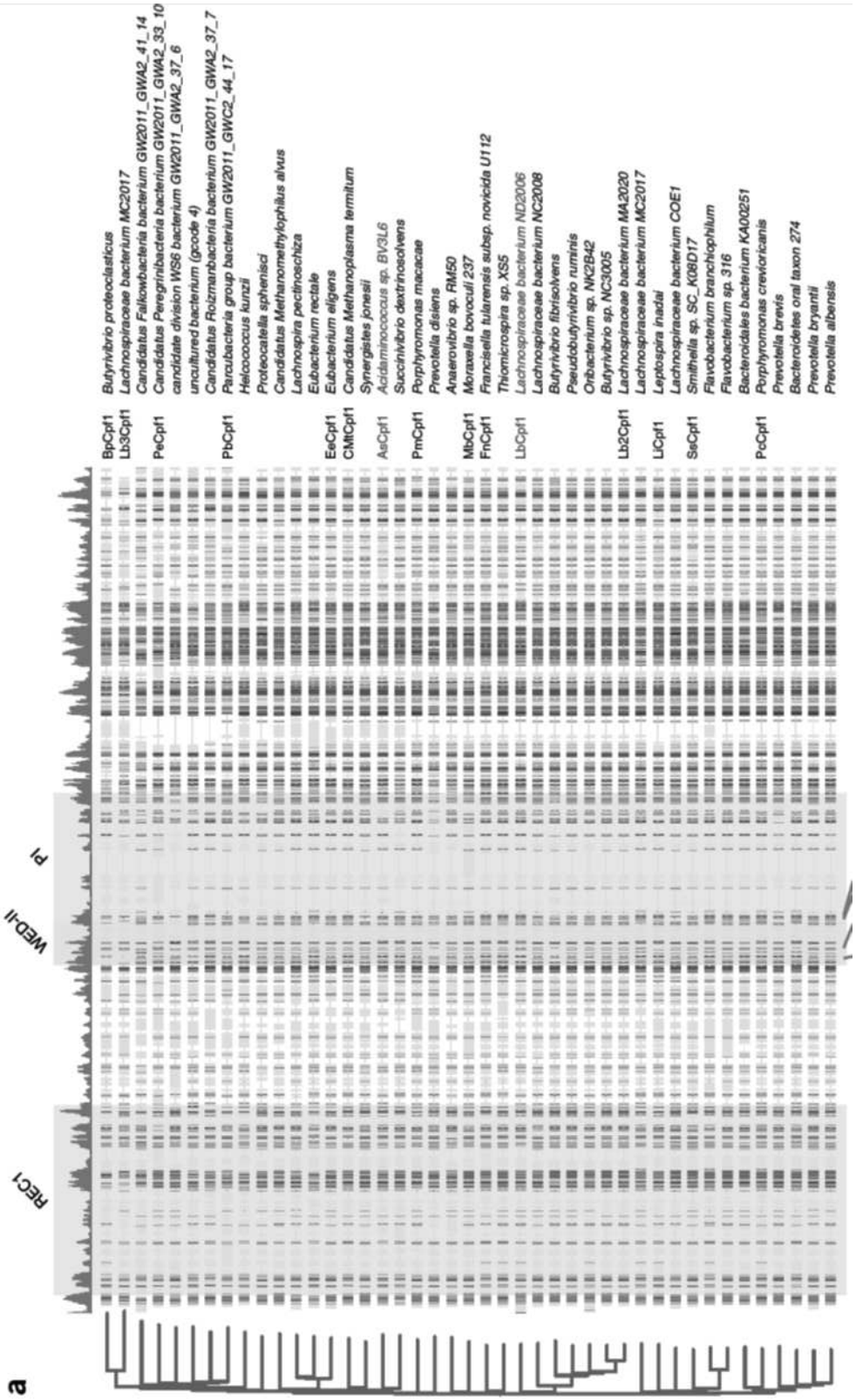

a
REC1
WED-II
PI
BpCpf1
Lb3Cpf1
PeCpf1
PbCpf1
EeCpf1
CMtCpf1
AsCpf1
PmCpf1
MbCpf1
FnCpf1
LbCpf1
Lb2Cpf1
LiCpf1
SsCpf1
PoCpf1
Butyrivibrio proteoclasticus
Lachnospiraceae bacterium MC2017
Candidatus Falkowbacteria bacterium GW2011_GWA2_41_14
Candidatus Peregrinibacteria bacterium GW2011_GWA2_33_10
candidate division WS6 bacterium GW2011_GWA2_37_6
uncultured bacterium (gcode 4)
Candidatus Roizmanbacteria bacterium GW2011_GWA2_37_7
Parcubacteria group bacterium GW2011_GWC2_44_17
Helcococcus kunzii
Proteocatella sphenisci
Candidatus Methanomethylophilus alvus
Lachnospira pectinoschiza
Eubacterium rectale
Eubacterium eligens
Candidatus Methanoplasma termitum
Synergistes jonesii
Acidaminococcus sp. BV3L6
Succinivibrio dextrinosolvens
Porphyromonas macacae
Prevotella disiens
Anaerovibrio sp. RM50
Moraxella bovoculi 237
Francisella tularensis subsp. novicida U112
Thiomicrospira sp. XS5
Lachnospiraceae bacterium ND2006
Lachnospiraceae bacterium NC2008
Butyrivibrio fibrisolvens
Pseudobutyrivibrio ruminis
Oribacterium sp. NK2B42
Butyrivibrio sp. NC3005
Lachnospiraceae bacterium MA2020
Lachnospiraceae bacterium MC2017
Leptospira inadai
Lachnospiraceae bacterium COE1
Smithella sp. SC_K08D17
Flavobacterium branchiophilum
Flavobacterium sp. 316
Bacteroidales bacterium KA00251
Porphyromonas crevioricanis
Prevotella brevis
Bacteroidetes oral taxon 274
Prevotella bryantii
Prevotella albensis

FIG. 30A

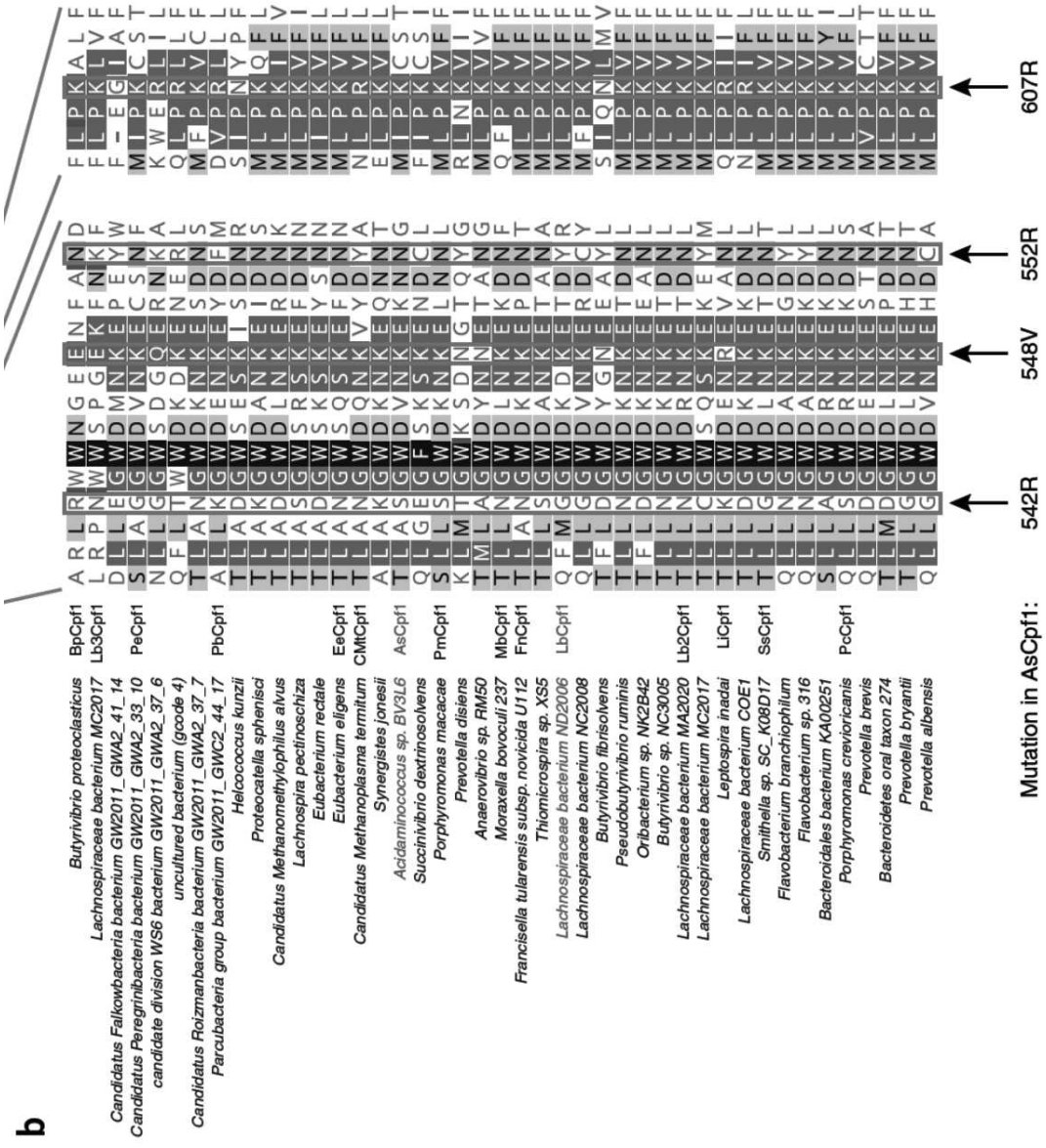

b
BpCpf1
Lb3Cpf1
PeCpf1
PbCpf1
EeCpf1
CMtCpf1
AsCpf1
PmCpf1
MbCpf1
FnCpf1
LbCpf1
Lb2Cpf1
LiCpf1
SsCpf1
PcCpf1
Butyrivibrio proteoclasticus
Lachnospiraceae bacterium MC2017
Candidatus Falkowbacteria bacterium GW2011_GWA2_41_14
Candidatus Peregrinibacteria bacterium GW2011_GWA2_33_10
candidate division WS6 bacterium GW2011_GWA2_37_6
uncultured bacterium (gcode 4)
Candidatus Roizmanbacteria bacterium GW2011_GWA2_37_7
Parcubacteria group bacterium GW2011_GWC2_44_17
Helcococcus kunzii
Proteocatella sphenisci
Candidatus Methanomethylophilus alvus
Lachnospira pectinoschiza
Eubacterium rectale
Eubacterium eligens
Candidatus Methanoplasma termitum
Synergistes jonesii
Acidaminococcus sp. BV3L6
Succinivibrio dextrinosolvens
Porphyromonas macacae
Prevotella disiens
Anaerovibrio sp. RM50
Moraxella bovoculi 237
Francisella tularensis subsp. novicida U112
Thiomicrospira sp. XS5
Lachnospiraceae bacterium ND2006
Lachnospiraceae bacterium NC2008
Butyrivibrio fibrisolvens
Pseudobutyrivibrio ruminis
Oribacterium sp. NK2B42
Butyrivibrio sp. NC3005
Lachnospiraceae bacterium MA2020
Lachnospiraceae bacterium MC2017
Leptospira inadai
Lachnospiraceae bacterium COE1
Smithella sp. SC_K08D17
Flavobacterium branchiophilum
Flavobacterium sp. 316
Bacteroidales bacterium KA00251
Porphyromonas crevioricanis
Prevotella brevis
Bacteroidetes oral taxon 274
Prevotella bryantii
Prevotella albensis
Mutation in AsCpf1:
542R
548V
552R
607R

FIG. 30B

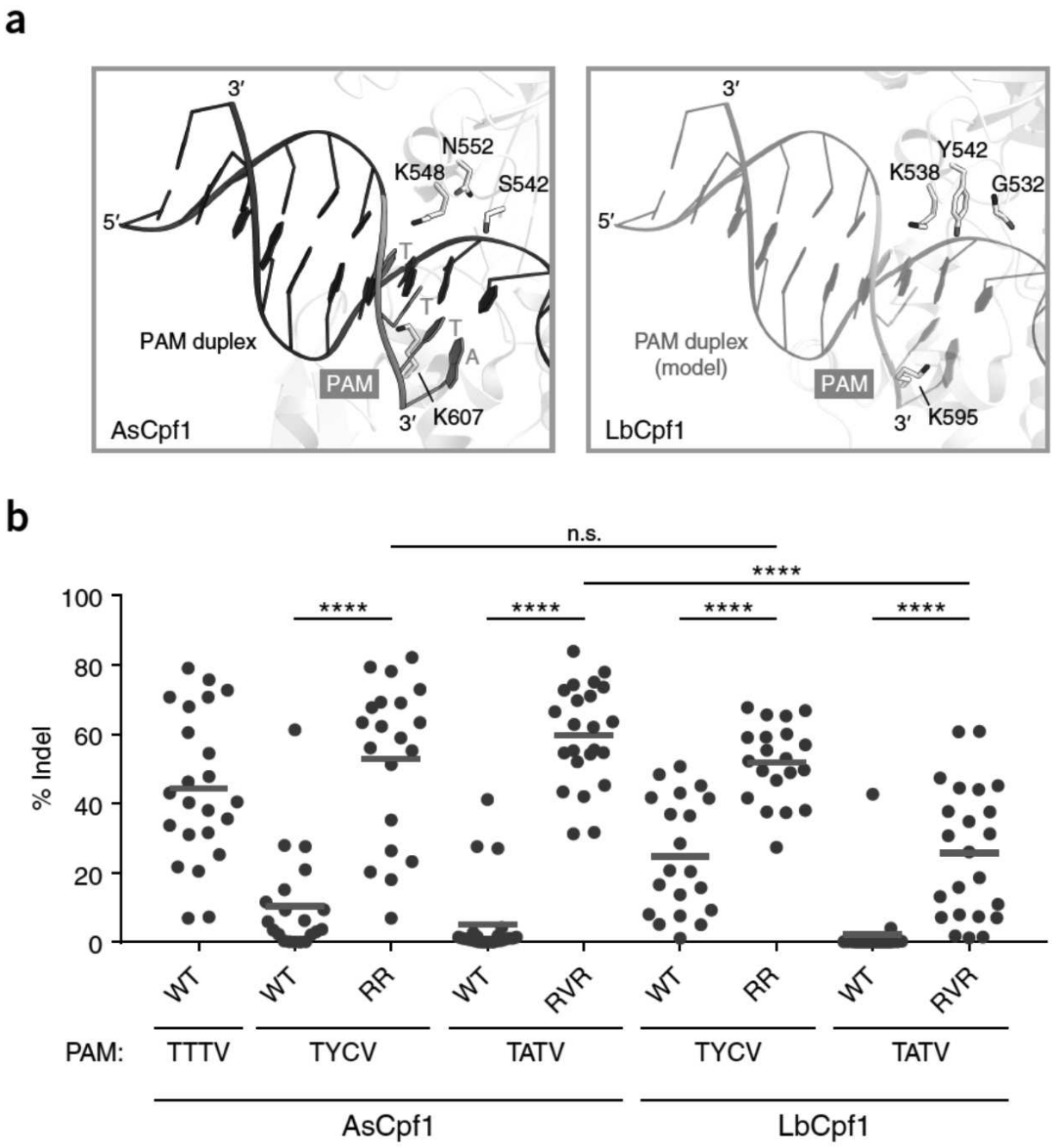

a
3′
5′
N552
K548
S542
T
T
T
A
PAM duplex
PAM
3′
K607
AsCpf1
3′
5′
Y542
K538
G532
PAM duplex
(model)
PAM
3′
K595
LbCpf1
b
n.s.
****
****
****
****
****
100
80
60
40
20
0
% Indel
WT
WT
RR
WT
RVR
WT
RR
WT
RVR
PAM:
TTTV
TYCV
TATV
TYCV
TATV
AsCpf1
LbCpf1

FIG. 31A-31B

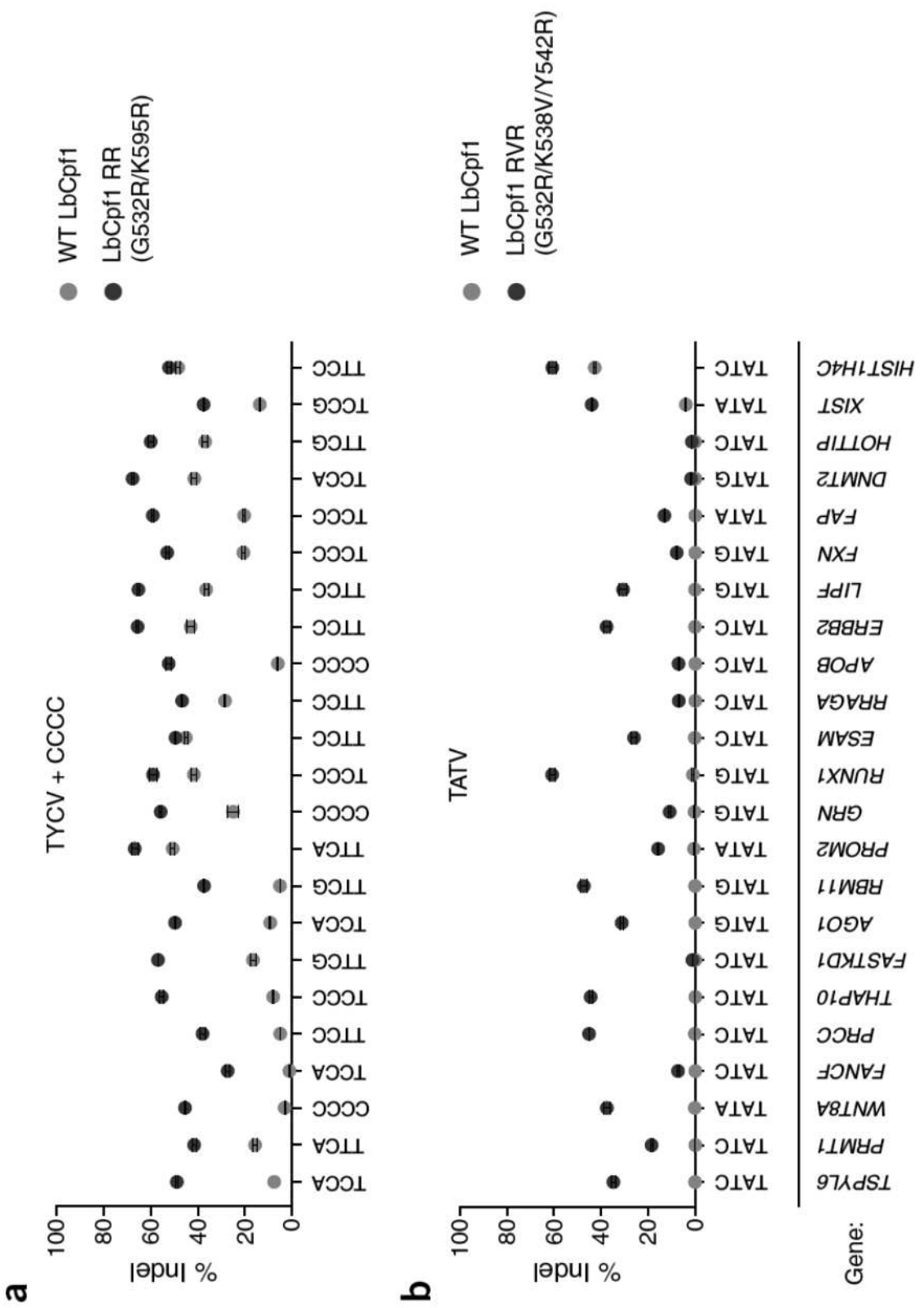

a
WT LbCpf1
LbCpf1 RR (G532R/K595R)
TYCV + CCCC
% Indel
100
80
60
40
20
0
TCCA
TTCA
CCCC
TTCA
TCCC
TTCC
TCCA
TTCG
TCCG
TTCC
CCCC
TTCC
TTCC
TCCC
TCCC
CCCC
TTCA
TTCG
TCCA
TTCG
TCCC
TTCC
TCCA
TTCC
TCCC
TTCG
TCCG
TTCC
b
WT LbCpf1
LbCpf1 RVR (G532R/K538V/Y542R)
TATV
% Indel
100
80
60
40
20
0
TATC
TATA
TATG
TATA
TATG
TATG
TATC
TATC
TATC
TATG
TATG
TATA
TATC
TATC
TATC
TATG
TATA
TATC
TATC
TATA
TATC
Gene:
TSPYL6
PRMT1
WNT8A
FANCF
PRCC
THAP10
FASTKD1
AGO1
RBM11
PROM2
GRN
RUNX1
ESAM
RRAGA
APOB
ERBB2
LIPF
FXN
FAP
DNMT2
HOTTIP
XIST
HIST1H4C

FIG. 32A-32B