



(51) International Patent Classification:

A61K 9/51 (2006.01) C12N 9/22 (2006.01)
A61K 38/46 (2006.01) C12N 15/113 (2010.01)
C07K 14/315 (2006.01) C12N 15/62 (2006.01)
C07K 19/00 (2006.01) C12N 15/85 (2006.01)

(21) International Application Number:

PCT/US2023/080718

(22) International Filing Date:

21 November 2023 (21.11.2023)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/427,370 22 November 2022 (22.11.2022) US

(71) Applicant: **FC IP HOLDINGS, LLC** [US/US]; c/o Co-
agency Global Inc., 850 New Burton Road, Suite 201, Dover,
Delaware 19904 (US).

(72) Inventors: **WU, Hao**; 6 Stimson Avenue, Lexington, Mass-
achusetts 02421 (US). **SHAN, Richard**; 625 Mount Auburn
Street, Suite 105, Cambridge, Massachusetts 02138 (US).

(74) Agent: **BEHROOZ, Alireza** et al.; Saul Ewing LLP, Centre
Square West, 1500 Market Street, 38th Floor, Philadelphia,
Pennsylvania 19102 (US).

(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,

AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ,
CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG,
KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY,
MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA,
NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO,
RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH,
TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS,
ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, CV,
GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST,
SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ,
RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ,
DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT,
LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE,
SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN,
GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- as to applicant's entitlement to apply for and be granted a
patent (Rule 4.17(ii))
- as to the applicant's entitlement to claim the priority of the
earlier application (Rule 4.17(iii))

Published:

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: DNA EDITING CONSTRUCTS AND METHODS USING THE SAME

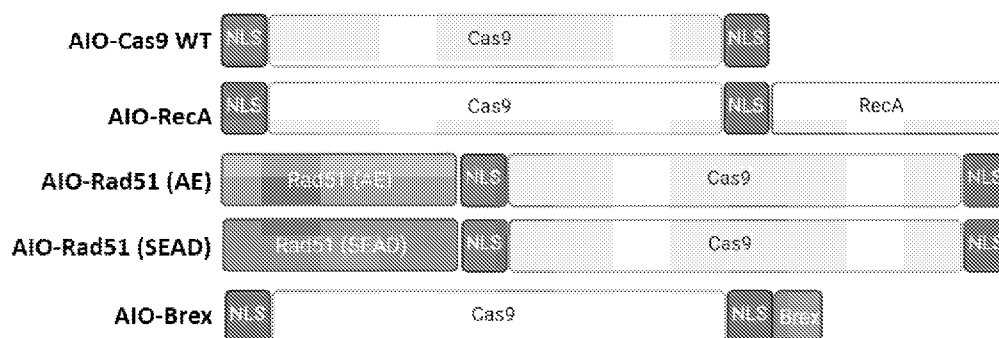


Fig. 1A

(57) **Abstract:** Described herein are DNA editing constructs including a DNA editing motif, a nuclear localization signal, and a single stranded DNA binding motif. The DNA editing constructs, when used together with a single stranded DNA donor template, can achieve improved editing efficiency. Also described herein are nucleic acids for encoding the DNA editing construct, DNA editing compositions and kits including the DNA editing constructs, as well as methods of editing DNA using the DNA editing constructs.

DNA EDITING CONSTRUCTS AND METHODS USING THE SAME

CROSS-REFERENCE TO RELATED APPLICATIONS

[001] This application claims priority under 35 U.S.C. § 119(e) to U.S. Provisional Patent Application No. 63/427,370, filed November 22, 2022, which is incorporated herein by reference in its entirety.

REFERENCE TO AN ELECTRONIC SEQUENCE LISTING

[002] The present application contains a Sequence Listing, which has been submitted in XML (ST.26) format and is incorporated herein by reference in its entirety. Said XML copy, created on November 20, 2023, is named 385722-1001WO1_Seq_Listing.xml and is 55,596 bytes in size.

BACKGROUND

[003] Unsatisfactory editing efficiency has always been an issue in the art of gene editing. Therefore, there is a need for DNA editing constructs and DNA editing methods that can improve editing efficiency. The present invention addresses this need.

SUMMARY

[004] In some aspects, the present invention is directed to the following non-limiting embodiments:

DNA editing construct

[005] In some aspects, the present invention is directed to a DNA editing construct.

[006] In some embodiments, the DNA editing construct comprises a DNA editing motif having a site-specific DNA endonuclease activity; a nuclear localization signal (NLS); and a single stranded DNA (ssDNA) binding motif.

[007] In some embodiments, the DNA editing motif introduces a double strand break, a single strand break, or no double strand break in double stranded DNA molecules.

[008] In some embodiments, the DNA editing motif introduces the double-strand break in double stranded DNA molecules, and the double-strand break is a blunt-ended break or a sticky-ended break.

[0009] In some embodiments, the DNA editing motif comprises a Cas nuclease, a zinc-finger nuclease, a meganuclease, or a transcription activator-like effector nuclease (TALEN).

[0010] In some embodiments, the DNA editing motif comprises a class I Cas protein or a class II Cas protein.

[0011] In some embodiments, the DNA editing motif comprises Cas9, Cas3, Cas10, Cas11, CasX, Cas12a, or MAD7. In some embodiments, the DNA editing motif comprises a catalytic dead form thereof. In some embodiments, the DNA editing motif comprises nickase forms thereof.

[0012] In some embodiments, the nuclear localization signal comprises a 53BP1 NLS, a CXCR4 NLS, an ERK5 NLS, an EWS NLS, an IER5 NLS, an ING4 NLS, an Hrp1 NLS, an MSX1 NLS, an NLS-RAR α NLS, a nucleoplasmin NLS, a Pho4 NLS, a PTHrP NLS, an rpL23a NLS, a simian virus 40 (SV40) large T antigen NLS, a UL79 NLS, a VACM-1/CUL5 NLS, a VP1 NLS, a PABPN1 NLS, a STAT1 NLS, a FGF2 NLS, a RAR α NLS, a tandem phenylalanine-glycine repeat (FG repeat) from a nuclear pore complex, a histone H2a NLS, or a histone H2b NLS.

[0013] In some embodiments, the construct comprises a plurality of the nuclear localization signals.

[0014] In some embodiments, the ssDNA binding motif is or comprises a K homology (KH) domain ssDNA binding motif, an oligonucleotide/oligosaccharide binding (OB) fold ssDNA binding motif, a RecA/Rad51/DMC1 loop L2 homologous ssDNA binding motif, an RNA recognition motif (RRM) that binds ssDNA, or a whirly domain ssDNA binding motif.

[0015] In some embodiments, the ssDNA binding motif is or comprises a bacterial single-stranded DNA-binding (SSB) protein or ssDNA binding portions thereof, a RecA/Rad51/DMC family protein or ssDNA binding portions thereof, a replication protein A (RPA) or ssDNA binding portions thereof, or a single-stranded annealing protein (SSAP) or ssDNA binding portions thereof.

[0016] In some embodiments, the ssDNA binding motif is or comprises an ssDNA binding motif derived from a RecA/Rad51/DMC family protein, and wherein the ssDNA binding motif does not have recombinase activity.

[0017] In some embodiments, the ssDNA binding motif is or comprises a polypeptide having the amino acid sequence set forth in SEQ ID NO:11.

[0018] In some embodiments, the construct comprises a plurality of the ssDNA binding motifs.

[0019] In some embodiments, the construct further comprises one or more linkers connecting the DNA editing motif, the nuclear localization signal, and the ssDNA binding motif to each other.

[0020] In some embodiments, the one or more linkers connects the DNA editing motif, the nuclear localization signal, and the ssDNA binding motif to each other covalently.

[0021] In some embodiments, the one or more linkers independently are polypeptide linkers.

[0022] In some embodiments, each of the polypeptide linkers independently has a length ranging from about 2 amino acid residues to about 100 amino acid residues.

[0023] In some embodiments, the construct is a protein.

Nucleic acid

[0024] In some aspects, the present invention is directed to a nucleic acid.

[0025] In some embodiments, the nucleic acid encodes the construct here, such as the protein construct herein.

[0026] In some embodiments, the nucleic acid further encodes a guide RNA (gRNA) that provides the site-specificity to the DNA editing motif.

[0027] In some embodiments, the nucleic acid is comprised in an expression vector.

[0028] In some embodiments, the expression vector is a plasmid vector, an RNA vector, or a viral vector.

[0029] In some embodiments, the nucleic acid is packaged in a viral particle, a lipid nanoparticle, or a non-lipid nanoparticle.

Composition

[0030] In some aspects, the present invention is directed to a composition.

[0031] In some embodiments, the composition is a composition for editing DNA.

[0032] In some embodiments, the composition comprises a DNA editing construct; and a single stranded DNA (ssDNA) having a DNA insert, a 5'-homology arm, and a 3'-homology arm.

[0033] In some embodiments, the DNA editing construct is the same as or similar to those described elsewhere herein.

[0034] In some embodiments, the 5' homology arm and the 3' homology arm are complementary to the nuclear DNA in a target region.

[0035] In some embodiments, the ssDNA is a circular ssDNA (cssDNA), a linear ssDNA (lssDNA), a double stranded DNA/circular single stranded DNA (dsDNA/cssDNA) hybrid having an ssDNA section, or an RNA/DNA hybrid having an ssDNA section.

[0036] In some embodiments, the ssDNA is chemically modified.

[0037] In some embodiments, the ssDNA is produced from a cell free system or a cell culture system.

[0038] In some embodiments, each of the 5'-homology arm and the 3'-homology arm independently has a length ranging from about 25 nucleotides to about 5000 nucleotides.

[0039] In some embodiments, the DNA insert has a length ranging from 1 nucleotide to about 50,000 nucleotides.

Kit

[0040] In some aspects, the present invention is directed to a kit.

[0041] In some embodiments, the kit is a kit for editing DNA.

[0042] In some embodiments, the kit comprises a construct and/or a nucleic acid of; and a single stranded DNA (ssDNA), a nucleic acid that encodes the ssDNA, or a nucleic acid for constructing the ssDNA.

[0043] In some embodiments, the construct is a DNA editing construct here. In some embodiments, the nucleic acid is a nucleic acid encoding a DNA editing construct herein.

[0044] In some embodiments, the ssDNA comprises a DNA insert, a 5'-homology arm, and a 3'-homology arm, wherein the 5' homology arm and the 3' homology arm are complementary to the nuclear DNA in a target region.

[0045] In some embodiments, the ssDNA is a circular ssDNA (cssDNA), a linear ssDNA (lssDNA), a double stranded DNA/circular single stranded DNA (dsDNA/cssDNA) hybrid having an ssDNA section, or an RNA/DNA hybrid having an ssDNA section.

[0046] In some embodiments, the ssDNA is chemically modified;

[0047] In some embodiments, the ssDNA is produced from a cell free system or a cell culture system.

[0048] In some embodiments, each of the 5'-homology arm and the 3'-homology arm independently has a length ranging from about 25 nucleotides to about 5000 nucleotides.

[0049] In some embodiments, the DNA insert has a length ranging from 1 nucleotide to about 50,000 nucleotides.

[0050] In some embodiments, the nucleic acid for constructing the ssDNA donor template encodes every elements of the ssDNA donor template, except for the DNA insert.

[0051] In some embodiments, the nucleic acid for constructing ssDNA donor template encodes every elements of the ssDNA donor template, except for the 5' homology arm and the 3' homology arm.

[0052] In some embodiments, the nucleic acid for constructing ssDNA donor template does not encode the 5' homology arm, the 3' homology arm, and the DNA insert.

[0053] In some embodiments, the nucleic acid includes a multiple cloning site for inserting the 5' homology arm, the 3' homology arm, and/or the DNA insert.

[0054] In some embodiments, the nucleic acid for constructing ssDNA donor template is a plasmid or a viral vector.

[0055] In some embodiments, the kit further comprises a component for delivering the construct, the nucleic acid for encoding the construct, the ssDNA, and/or the nucleic acid for encoding or constructing the ssDNA into a cell.

Method of editing a nuclear DNA

[0056] In some aspects, the present invention is directed to a method of editing a nuclear DNA, such as a nuclear DNA in a cell.

[0057] In some embodiments, the method comprises delivering into the cell the construct herein and/or the nucleic acid herein; and a single stranded DNA (ssDNA) having a DNA insert, a 5'-homology arm, and a 3'-homology arm, wherein the 5' homology arm and the 3' homology arm are complementary to the nuclear DNA in a target region.

[0058] In some embodiments, the ssDNA is a circular ssDNA (cssDNA), a linear ssDNA (lssDNA), a double stranded DNA/circular single stranded DNA (dsDNA/cssDNA) hybrid having an ssDNA section, or an RNA/DNA hybrid having an ssDNA section.

[0059] In some embodiments, the ssDNA is chemically modified.

[0060] In some embodiments, the ssDNA is produced from a cell free system or a cell culture system.

[0061] In some embodiments, each of the 5'-homology arm and the 3'-homology arm independently has a length ranging from about 25 nucleotides to about 5000 nucleotides.

[0062] In some embodiments, the DNA insert has a length ranging from 1 nucleotide to about 50,000 nucleotides.

[0063] In some embodiments, the construct and/or nucleic acid and the ssDNA are delivered into the cell by a viral vector, a lipid or non-lipid nanoparticle delivery, an electroporation delivery, a gene gun delivery, or an injection.

[0064] In some embodiments, the nuclear DNA is a genomic DNA.

[0065] In some embodiments, the cell is a prokaryotic cell, a cultured plant cell, a eukaryotic cell, a cultured mammalian cell, a primary cell, an embryonic stem cell, an adult stem cell, a hematopoietic stem cell, an induced pluripotent stem cell, or a CAR-T cell.

[0066] In some embodiments, the cell is in the body of a subject.

[0067] In some embodiments, the subject is a mammal, such as a human.

Non-natural polypeptide

[0068] In some aspects, the present invention is directed to a non-natural polypeptide comprising an amino acid sequence having 90% or higher identity to the amino acid sequence set forth in SEQ ID NO:11.

[0069] In some aspects, the non-natural polypeptide is a ssDNA-binding polypeptide.

BRIEF DESCRIPTION OF THE DRAWINGS

[0070] The following detailed description of exemplary embodiments will be better understood when read in conjunction with the appended drawings. For the purpose of illustrating, non-limiting embodiments are shown in the drawings. It should be understood, however, that the instant specification is not limited to the precise arrangements and instrumentalities of the embodiments shown in the drawings.

[0071] Figs. 1A-1F demonstrate that fusion of Cas9 and homologous recombination proteins enhance the ssDNA mediated knock-in in K562 cells, in accordance with some embodiments.

Fig. 1A: Schematic diagram of various Cas9-homologous recombination protein fusion

constructs (enGagers) in all-in-one (AIO) plasmid format modified from Addgene plasmid #42230. Two nuclear localization signals were added to the N' and C'-termini of the Cas9 protein. RecA is the bacteria homologous DNA repair protein and Rad51(AE)/Rad51(SEAD) are two mutant variants from eukaryotes. Brex is a 36 amino acid peptide reported to recruit Rad51 in mammalian cells. Fig. 1B: Schematic diagram of knock-in strategy of a 2Kb (left) and 4Kb (right) cssDNA donor template for *RAB11A* locus. Fig. 1C: representative FACS profiles with gating strategy showing % of GFP transgene cassette knock-in on *RAB11A* locus at day 3 post electroporation for various enGagers listed in A. Fig. 1D: Quantification of 2Kb GFP transgene cassette knock-in fold change of various enGagers as compared to wild type (WT) Cas9 at day 3 (left), 8 (middle) and 14 (right) post electroporation. Fig. 1E: representative FACS profiles with gating strategy showing % of 4Kb GFP transgene cassette knock-in on *RAB11A* locus at day 3 post electroporation for various enGagers listed in A. Fig. 1F: Quantification of 4Kb GFP transgene cassette knock-in fold change of various enGagers as compared to WT Cas9 at day 3 (left), 7 (right) post electroporation. Note that Brex enGager does not enhance knock-in efficiency. Rad51 mutants and RecA enGagers increase both small and large transgene cassette knock-in by 1.55 to 3.00-fold. RecA enGager outperforms among the enGagers tested. Bars represent mean $\pm$ SD from 3 biological replicates.

[0072] Figs. 2A-2D: Identification of small enGagers with Cas9 fused ssDNA binding motifs in K562 cells, in accordance with some embodiments. Fig. 2A: Schematic diagram of various Cas9-ssDNA binding motifs fusion constructs (enGagers) in all-in-one (AIO) plasmid format modified from Addgene plasmid #42230. Two nuclear localization signals were added to the N' and C'-termini of the Cas9 protein. Cas9-RecA fusion construct was used as a positive control. FECO, WECO and YECO are 20 amino acid sequences previously identified as ssDNA binding motifs in various bacteria species of RecA, FECO3X, WECO3X and YECO3X are 3 tandem copies of the 20 AA peptides separated by multi-GS peptide linkers. Fig. 2B: Schematic diagram of knock-in strategy of a 2Kb cssDNA donor template for *RAB11A* locus. Fig. 2C: representative FACS profiles with gating strategy showing % of GFP transgene cassette knock-in on *RAB11A* locus at day 7 post electroporation for various small enGagers listed in A. Fig. 2D: Quantification of 2Kb GFP transgene cassette knock-in fold change of various enGagers as compared to WT Cas9 at day 7 post electroporation. Note that Cas9-FECO fusion performs similarly with Cas9-RecA fusion in cssDNA mediated transgene integration (1.59- vs 1.57-fold).

In most cases, enGagers with 3X tandem ssDNA binding peptides neither further enhanced nor significantly reduced knock-in efficiency. Bars represent mean $\pm$ SD from 3 biological replicates.

[0073] Figs. 3A-3E: Identification of additional novel enGagers from Cas9-ssDNA binding module chimeras in K562 cells, in accordance with some embodiments. Fig. 3A: Schematic diagram of various Cas9-ssDNA binding protein and peptide fusion constructs (enGagers) in all-in-one (AIO) plasmid format modified from Addgene plasmid #42230. Two nuclear localization signals were added to the N' and C'-termini of the Cas9 protein. Cas9-RecA and Cas9-FECO fusion constructs were used as positive controls. The fusion protein or peptides include DrRecA, 20AA motif identified from DrRecA, SSAP, Lambda Beta and RecT, an *E. coli* single stranded DNA annealing protein (SSAP). Fig. 3B: Schematic diagram of knock-in strategy of a 2Kb cssDNA donor template for *RAB11A* locus. Fig. 3C: Quantification of 2Kb GFP transgene cassette knock-in fold change of various enGagers as compared to WT Cas9 at day 7 post electroporation. Note that Cas9-DrRecA and Cas9-DrRecA20AA fusion has the highest performance in knock-in with 2.17- and 2.43-fold as compared to WT Cas9, respectively. Fig. 3D: Quantification of cell viability day 7 post electroporation. Bars represents mean $\pm$ SD from 3 biological replicates. Figs. 3E: Amino acid sequence alignment of 20AA of multiple *E. coli* RecA mutant variants and RecA from archaea and mammalian organism. Dr: *Deinococcus radiodurans*; Ec: *Escherichia coli*; Sc: *Saccharomyces cerevisiae*; Hs: *Homo sapiens*; Pf: *Pyrococcus furiosus*; Mv: *Methanococcus voltae*; Sso: *Saccharolobus solfataricus*. The amino acid sequences of Figs. 3E are listed below in Table 1:

Table 1

<i>S. solfataricus</i> RadA ssDNA binding motif: NQVMARPDPMFYGDPTVAVGG (SEQ ID NO:1)
<i>M. voltae</i> RadA ssDNA binding motif: NQVSAKPDAFFGMAEQAIGG (SEQ ID NO:2)
<i>P. furiosus</i> RadA ssDNA binding motif: NQVQARPDFAFFGDPTRPIGG (SEQ ID NO:3)
<i>H. sapiens</i> Rad51 ssDNA binding motif: NQVVAQVDGAAMFAADPKKPIGG (SEQ ID NO:4)
<i>S. cerevisiae</i> Rad51 ssDNA binding motif: NQVVAQVDGGMAFNPDPPKKPIGG (SEQ ID NO:5)

<i>H. sapiens</i> DMC1 ssDNA binding motif: NQMTADPGATMTFQADPKKPIGG (SEQ ID NO:6)
<i>S. cerevisiae</i> DMC1 ssDNA binding motif: NQVQSDPGASALFASADGRKPIGG (SEQ ID NO:7)
<i>E. coli</i> RecA loop L2 region (“FECO peptide”): NQIRMKIGVMFGNPETTTGG (SEQ ID NO:8)
<i>E. coli</i> RecA loop L2 region with F to W mutation (“WECO” peptide): NQIRMKIGVMWGNPETTTGG (SEQ ID NO:9)
<i>E. coli</i> RecA loop L2 region with F to Y mutation (“YECO” peptide): NQIRMKIGWMYGNPETTTGG (SEQ ID NO:10)
<i>D. radiodurans</i> RecA L2 region: NQVREKIGVMYGNPETTTGG (SEQ ID NO:11)

[0074] Figs. 4A-4I demonstrate that enGager mediated genome integration enhancement is ssDNA dependent, in accordance with some embodiments. Fig. 4A: Schematic diagram of various Cas9-ssDNA binding protein and peptide fusion constructs (enGagers) in mRNA form. Two nuclear localization signals were added to the N' and C'-termini of the Cas9 protein. GS is a shortened multi-GS peptide linker “EFGGGGS” (SEQ ID NO:12). Fig. 4B: Schematic diagram of knock-in strategy of a 2Kb cssDNA donor template for *RAB11A* locus. Fig. 4C: Dose titration of cssDNA at 0.3, 1, 1.5, 2 and 3ug for 2Kb GFP transgene knock-in at day 3 post electroporation. Cas9-RecA mRNA enhance around 25-30% knock-in efficiency than WT Cas9 mRNA at all the cssDNA dose tested in K562 cells. Fig. 4D: representative FACS profiles with gating strategy showing % of 2Kb GFP transgene cassette knock-in on *RAB11A* locus by cssDNA donor at day 5 post electroporation for various enGagers listed in Fig. 4A in K562 cells. Fig. 4E: Quantification of 2Kb GFP transgene cassette knock-in fold change (left) and cell viability (right) of various enGagers as compared to WT Cas9 at day 5 post electroporation from Fig. 4D. Fig. 4F: representative FACS profiles with gating strategy showing % of 2Kb GFP transgene cassette knock-in on *RAB11A* locus by dsDNA donor at day 5 post electroporation for various enGagers listed in Fig. 4A in K562 cells. Fig. 4G: Quantification of 2Kb GFP transgene cassette knock-in fold change (left) and cell viability (right) of various enGagers as compared to WT Cas9 at day 5 post electroporation from Fig. 4F. Fig. 4H: Quantification of 2Kb GFP

transgene cassette knock-in fold change (left) and cell viability (right) of various enGagers mRNA as compared to WT Cas9 mRNA with cssDNA donor at 5 days post-delivery in HEK293 cells by lipofectamine 3000 transfection. Fig. 4I: Quantification of 2Kb GFP transgene cassette knock-in fold change (left) and cell viability (right) of various enGagers mRNA as compared to WT Cas9 mRNA with dsDNA donor at 5 days post-delivery in HEK293 cells by lipofectamine 3000 transfection. Bars represents mean $\pm$ SD from 3 biological replicates.

[0075] Figs. 5A-5F demonstrate that enGagers in mRNA form enhance genome integration on various locus and large payload, in accordance with some embodiments. Fig. 5A: Schematic diagram of knock-in strategy of 2Kb (css76), 4Kb (css116) and 8Kb (css167) cssDNA donor templates for *RAB11A* locus. Fig. 5B: Quantification of % of 2Kb (left), 4Kb (middle) and 8Kb (right) GFP transgene cassette knock-in for various mRNA enGagers *RAB11A* day 13 post electroporation in K562 cells. Fig. 5C: Quantification of cell viability for 2Kb (left), 4Kb (middle) and 8Kb (right) GFP transgene cassette knock-in for various mRNA enGagers on *RAB11A* locus day 13 post electroporation in K562 cells. Fig. 5D: Schematic diagram of knock-in strategy of 2Kb (css27) and 4Kb (css88) cssDNA donor templates for *B2M* locus. Fig. 5E: Quantification of % of 2Kb (left) and 4Kb (right) GFP transgene cassette knock-in for various mRNA enGagers on *B2M* locus day 13 post electroporation in K562 cells. Fig. 5F: Quantification of cell viability for 2Kb (left) and 4Kb (right) GFP transgene cassette knock-in for various mRNA enGagers on *B2M* locus day 13 post electroporation in K562 cells. Bars represents mean $\pm$ SD from 3 biological replicates.

[0076] Figs. 6A-6F demonstrate that CAR-T engineered by enGager has superior efficiency than those engineered by WT Cas9, in accordance with some embodiments. Fig. 6A: Schematic diagram of knock-in strategy of 3Kb CD19/CD22 dual CAR (css62) cssDNA donor templates for *TRAC* locus in primary T cells. Fig. 6B: Quantification of % of CD19/CD22 CAR knock-in using cssDNA donor analyzed by Protein-L for various doses of WT Cas9 and Cas9-GS-FECO enGager mRNA at 1 ug, 2ug and 4ug at day 7 post electroporation of primary T cells. Cas9-GS-FECO enGager achieves ~ 4- to 6-fold higher CAR-T engineering efficiency than WT Cas9. Fig. 6C: Characterization of the engineered CAR-T cells or mock treated T cells for total cell count, cell proliferation fold change and cell viability over time (at day 1, 4, 7 and 11 days post electroporation). Fig. 6D: NALM6 leukemia lymphocyte killing curve of unengineered T cells, CD19-CD22 dual CAR-T cells engineered with 2 ug of WT Cas9 mRNA and 2 ug of GS-FECO

enGager mRNA over the course of 96 hrs. Effect (T cells): Target (NALM6 cells) are at 2.25:1 for left panel, 4.5:1 for middle panel and 9:1 for right panel. Bars represents mean $\pm$ SD from 3 biological replicates. Fig. 6E: NALM6 cell killing function of CAR-T cells at 24 hrs for E:T ratio at 2.25:1, 4.5:1 and 9:1. Fig. 6F: Schematic diagram of engineered enGagers with single stranded DNA binding protein (SSBP) can recruit cssDNA donor template and form a tripartite editing machinery for efficient translocation of the entire editing complex from cytoplasm to nucleus and tethering to the targeted genomic locus. As a result, the donor DNA has higher effective local concentration in the nucleus for more efficient homologous directed genome integration. This process works more prominently with cssDNA.

[0077] Figs. 7A-7B: illustration of a non-limiting DNA editing construct, as well as method of editing nuclear DNA using the construct, in accordance with some embodiments. Fig. 7A: the DNA editing construct 100 includes a DNA editing motif 110, a nuclear localization signal 130, and a single stranded DNA binding motif 150 with linkers 170. Fig. 7B: the DNA editing construct 100, when introduced into a cell 500 together with a ssDNA 200 (shown as a circular single stranded DNA or cssDNA) that serves as a donor template for a homology-directed repair, are able to bind to the ssDNA 200 and forms the DNA editing complex 300. The DNA editing complex 300 moves into the nucleus 510 via active transportation in addition to passive diffusion, and edits the nuclear DNA 530 (at a locus 531) inside the nucleus 510 through the combined effects of the DNA editing motif 110 and the ssDNA 200.

DETAILED DESCRIPTION

[0078] The following disclosure provides many different embodiments, or examples, for implementing different features of the provided subject matter. Specific examples of components and arrangements are described below to simplify the present disclosure. These are, of course, merely examples and are not intended to be limiting. For example, the formation of a first feature over or on a second feature in the description that follows may include embodiments in which the first and second features are formed in direct contact, and may also include embodiments in which additional features may be formed between the first and second features, such that the first and second features may not be in direct contact. In addition, the present disclosure may repeat reference numerals and/or letters in the various examples. This repetition is for the purpose of

simplicity and clarity and does not in itself dictate a relationship between the various embodiments and/or configurations discussed.

Definitions

[0079] As used herein, each of the following terms has the meaning associated with it in this section. Unless defined otherwise, all technical and scientific terms used herein generally have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. Generally, the nomenclature used herein and the laboratory procedures in animal pharmacology, pharmaceutical science, peptide chemistry, and organic chemistry are those well-known and commonly employed in the art. It should be understood that the order of steps or order for performing certain actions is immaterial, so long as the present teachings remain operable. Any use of section headings is intended to aid reading of the document and is not to be interpreted as limiting; information that is relevant to a section heading may occur within or outside of that particular section. All publications, patents, and patent documents referred to in this document are incorporated by reference herein in their entirety, as though individually incorporated by reference.

[0080] In the application, where an element or component is said to be included in and/or selected from a list of recited elements or components, it should be understood that the element or component can be any one of the recited elements or components and can be selected from a group consisting of two or more of the recited elements or components.

[0081] In the methods described herein, the acts can be carried out in any order, except when a temporal or operational sequence is explicitly recited. Furthermore, specified acts can be carried out concurrently unless explicit claim language recites that they be carried out separately. For example, a claimed act of doing X and a claimed act of doing Y can be conducted simultaneously within a single operation, and the resulting process will fall within the literal scope of the claimed process.

[0082] In this document, the terms "a," "an," or "the" are used to include one or more than one unless the context clearly dictates otherwise. The term "or" is used to refer to a nonexclusive "or" unless otherwise indicated. The statement "at least one of A and B" or "at least one of A or B" has the same meaning as "A, B, or A and B."

[0083] "About" as used herein when referring to a measurable value such as an amount, a temporal duration, and the like, is meant to encompass variations of $\pm 20\%$ or $\pm 10\%$, in certain embodiments $\pm 5\%$, in certain embodiments $\pm 1\%$, in certain embodiments $\pm 0.1\%$ from the specified value, as such variations are appropriate to perform the disclosed methods.

[0084] Abbreviations: cssDNA: Circular single stranded DNA. DSB: Double stranded DNA break. gRNA: guide RNA. HDR: Homology directed repair. HR: Homologous recombination. KI: knock-in. sgRNA: Single guide RNA. ssDNA: Single stranded DNA.

Constructs for DNA Editing

[0085] The study described herein ("the present study") developed non-limiting novel DNA editing constructs that edit DNA using single stranded DNA (ssDNA) as a donor template. Without wishing to be bound by theory, it is hypothesized that novel DNA editing constructs described herein are able to form a complex with the ssDNA donor template in the cell, such as the cytoplasm. The complexes are then efficiently transported into the cell nucleus where the DNA editing constructs edit nuclear DNA using the ssDNA complexed with the constructs as the donor template. The present study further discovered that the novel DNA editing constructs have significantly higher DNA editing efficiency than similar DNA editing constructs that are unable to bind to the ssDNA.

[0086] Accordingly, in some embodiments, the present invention is directed to DNA editing constructs. In some embodiments, the present invention is directed to DNA editing constructs that edit nuclear DNA using a ssDNA as a template.

[0087] Referring to Fig. 7A, in some embodiments, the DNA editing construct 100 includes a DNA editing motif 110, a nuclear localization signal 130, and a single stranded DNA binding motif.

[0088] Referring to Fig. 7B, in some embodiments, the DNA editing construct 100, when introduced into a cell 500 together with a ssDNA 200 (shown as a circular single stranded DNA or cssDNA) that serves as a donor template for a homology-directed repair, are able to bind to the ssDNA 200 and forms the DNA editing complex 300. The DNA editing complex 300 moves into the nucleus 510 via active transportation in addition to passive diffusion, and edits the nuclear DNA 530 inside the nucleus 510 through the combined effects of the DNA editing motif 110 and the ssDNA 200.

DNA Editing Motifs

[0089] In some embodiments, the DNA editing motif 110 edits the DNA molecule via homologous recombination (HR) or homology-directed repair (HDR) using the ssDNA as the template.

[0090] In some embodiments, the DNA editing motif 110 has a nuclease activity, such as a DNA nuclease activity, such as a DNA endonuclease activity, such as a site-specific DNA endonuclease activity. In some embodiments, the DNA editing motif 110 is, or includes, a nuclease, such as a DNA nuclease, such as a DNA endonuclease, such as a site-specific DNA endonuclease.

[0091] As used herein, the term “DNA endonuclease activity” refers to the enzymatic activity of introducing a break, such as a double strand break (either a blunt-ended break or a sticky-ended break) or a single strand break (also referred to as a “nick”) into a DNA molecule, such as a double stranded DNA molecule, at locations other than the two ends thereof. In some embodiments, the term “site-specific DNA endonuclease activity” refers to the enzymatic activity of introducing a break into a DNA molecule at or near a specific site, such as a site as defined by a specific nucleotide sequence. Any proteins that possess DNA endonuclease activity is referred to herein as a DNA endonuclease, regardless of their origins (natural, recombinant, engineered, truncated, *etc.*)

[0092] It is worth noting that, the sequence specificities of some site-specific DNA endonucleases are, at least partially, derived from external components that are not considered as part of the nucleases *per se* by the art. For example, the site-specificities of some Cas proteins are largely derived from guide RNAs, which are not normally considered part of the Cas proteins. For the purposes of this application, however, as long as a nuclease is able to gain site-specificity from such external components, the nuclease is considered a “site-specific” regardless of whether the external components are actually present.

[0093] One of ordinary skill in the art, reading in light of the instant specification, would understand that the DNA editing construct 100 herein is able to bring ssDNA donor templates to the proximity of the target site on the DNA molecule. Since DNA strand breaks (as well as homologous recombination/homology directed repair) occur frequently and naturally without the presence of nucleases, enzyme dead form of site-specific DNA endonucleases can be used as

DNA editing motif 110, as such site-specific DNA endonucleases would still perform homologous recombination or homology directed repair-based DNA editing. Accordingly, in some embodiments, the DNA editing motif 110 includes a catalytic dead form of the site-specific nuclease.

[0094] In some embodiments, the DNA editing motif is not a catalytic dead forms of a site-specific DNA endonuclease, as such form of endonucleases generally result in lower editing efficiencies than their enzymatically active counterparts.

[0095] In some embodiments, the DNA editing motif 110 is a site-specific DNA endonuclease, or portions thereof that possess the site-specific DNA endonuclease activity, such as one of more protein domains (such as the enzyme domains) of the site-specific DNA endonuclease.

[0096] Non-limiting examples of site-specific DNA endonucleases include the Cas nucleases, the zinc-finger nucleases, the meganucleases, the transcription activator-like effector nucleases (TALENs), and the like. A large number of site-specific DNA endonucleases are genetically engineered proteins, such as those described in Bogdanove *et al.* (*Nucleic Acids Research*, Volume 46, Issue 10, 1 June 2018, Pages 4845–4871, the entirety of this reference is hereby incorporated herein by reference).

[0097] Non-limiting examples of Cas nucleases include class I and Class II Cas proteins such as Cas9, CasX, Cas12a, MAD7, or derivatives thereof, such as catalytic dead forms or nickase forms thereof. Class I and class II proteins are described in, for example, Makarova *et al.* (*Nature Reviews Microbiology* volume 18, pages67–83 (2020)). The entirety of this reference is hereby incorporated herein by reference.

Nuclear Localization Signals

[0098] In some embodiments, the nuclear localization signal 130 includes or is a nuclear localization signal commonly known in the art, which is a peptide motif that mediates the transportation, such as the importin-mediated transportation, of proteins from the cytoplasm into the nucleus.

[0099] Non-limiting examples of nuclear localization signals include a 53BP1 NLS, a CXCR4 NLS, an ERK5 NLS, an EWS NLS, an IER5 NLS, an ING4 NLS, an Hrp1 NLS, an MSX1 NLS, an NLS-RAR α NLS, a nucleoplasmin NLS, a Pho4 NLS, a PTHrP NLS, an rpl23a NLS, a simian virus 40 (SV40) large T antigen NLS, a UL79 NLS, a VACM-1/CUL5 NLS, a VP1 NLS,

PABPN1 NLS, STAT1 NLS, FGF2 NLS, RAR α NLS or the like, and their homologous sequences from different species. Nuclear localization signals also include one or more tandem phenylalanine-glycine repeats (FG repeats) from nuclear pore complex such as importin- α/β and their homologous sequence from different species. Nuclear localization signals further include those found in histone H2a and H2b, which are described in, for example Mosammaparast *et al.* (*J Cell Biol.* 2001 Apr 16; 153(2): 251–262). The entirety of this reference is hereby incorporated herein by reference.

[00100] Nuclear localization signals are further described in, for example, Lu *et al.* (*Cell Communication and Signaling* volume 19, Article number: 60 (2021)).

[00101] In some embodiments, the construct 100 includes a plurality of the nuclear localization signals 130, such as 2 or more, 3 or more, 4 or more, or 5 or more of the nuclear localization signals 130.

ssDNA Binding Motifs

[00102] In some embodiments, the ssDNA binding motif 150 is a motif that directly binds to ssDNA molecules. In some embodiments, the ssDNA binding motif 150 is not an adaptor for recruiting intrinsic ssDNA binding proteins in the cell.

[00103] The present study has discovered that, when an ssDNA binding motif 150 is included in the construct 100, the DNA editing efficiency can increase significantly, sometimes by about 50% or more, about 100% or more, or about 150% or more, as compared to similar DNA editing constructs without the ssDNA binding motif (see e.g., Figs. 1D, 1E, 3C, 4E, 4H, 5B, and 5D).

[00104] Furthermore, the present study further discovered that, as long as a ssDNA binding motif 150 is present, the DNA editing efficiencies can be increased. The structure or evolutionary paths of the ssDNA binding motif, or manners by which the ssDNA binding motif interacts with ssDNA molecules do not appear to matter. For example, RecA protein interact with ssDNA molecules via a 20 amino acid residue loop (“loop L2”), and bacteria SSB protein interact with ssDNA molecules via an oligonucleotide/oligosaccharide binding (OB) fold structural feature, which is unrelated to the loop L2 of RecA either structurally or evolutionarily. Yet the inclusion of either RecA (as well as the loop L2 portions thereof) or bacteria SSB in DNA editing constructs was able to significantly increase the DNA editing efficiency (see e.g., Figs. 3A and 3C).

[00105] In some embodiments, the ssDNA Binding Motif 150 is a K homology (KH) domain ssDNA binding motif, an oligonucleotide/oligosaccharide binding (OB) fold ssDNA binding motif, a RecA/Rad51/DMC1 loop L2 homologous ssDNA binding motif, an RNA recognition motif (RRM) that binds ssDNA, a whirly domain ssDNA binding motif, or the like. KH domain, OB folds, RRM and whirly domain are described in, for example, Dickey *et al.* (*Structure*, Volume 21, Issue 7, 2 July 2013, Pages 1074-1084, the entirety of the reference is hereby incorporated herein by reference). RecA/Rad51/DMC1 loop L2 homologous ssDNA binding motif is described herein (see e.g., Fig. 3E).

[00106] In some embodiments, the ssDNA binding motif 150 is a bacterial single-stranded DNA-binding (SSB) protein or ssDNA binding portions thereof (such as the OB fold), a DNA meiotic recombinase 1 (DMC1) protein or ssDNA binding portions thereof (such as the L2 homologous motif), a RecA/Rad51 family protein or ssDNA binding portions thereof (such as the L2 homologous motif), a replication protein A (RPA) or ssDNA binding portions thereof (such as the RPA1, RPA2, or RPA3 subunit of RPA, or the OB folds thereof), a single-stranded annealing protein (SSAP), SAAV2152, or a ring-shaped MCM helicase or ssDNA binding portions, and their interspecies homologs thereof.

[00107] In some embodiments, the ssDNA binding motif 150 is the RecA DNA recombinase, a functional derivative, or a protein from the RecA phylogeny (including paralogous and orthologous proteins from the prokaryotes or the eukaryotes, such as Rad51 or DMC1), or the ssDNA binding portions thereof.

[00108] The present study discovered that the ssDNA binding loop L2 of some non-limiting RecA protein (e.g., WECO peptide) increases DNA editing efficiency at levels comparable to the full length RecA protein (see e.g., Fig. 2D), which suggests that the recombinase activity of RecA is not important. Accordingly, in some embodiments, the RecA/Rad51/DMC1 proteins or ssDNA binding portions thereof ssDNA binding motif 150 does not have recombinase activity. In some embodiments, the ssDNA binding motif 150 is the loop L2 of the RecA/Rad51/DMC1 proteins.

[00109] In some embodiments, the binding motif 150 includes, or is a polypeptide set forth in SEQ ID NOs:1-11. In some embodiments, the binding motif 150 includes, or is a polypeptide set forth in SEQ ID:1.

[00110] In some embodiments, the ssDNA binding motif 150 is the bacterial SSB protein, a functional derivative, a protein from the SSB phylogeny, or the ssDNA binding portions thereof.

[00111] In some embodiments, the ssDNA binding motif 150 is the RecT protein, a functional derivative, a protein from the RecT phylogeny, or the ssDNA binding portions thereof.

[00112] In some embodiments, a length of the ssDNA binding motif 150 ranges from 5-25 amino acid residues. In some embodiments, the length of the ssDNA binding motif 150 ranges from 25-100 amino acid residues. In some embodiments, the length of the ssDNA binding motif 150 has more than 100 amino acid residues.

[00113] The present study tested the inclusion of multiple ssDNA binding motifs 150 in one DNA editing construct 100, and found that the multiple ssDNA binding motifs 150 at least did not affect the DNA editing efficiency when comparing to constructs 100 that include only one ssDNA binding motif 150 (see e.g., Fig. 2D). Accordingly, in some embodiments, the construct 100 includes a plurality of the ssDNA binding motifs 150. In some embodiments, the construct 100 includes 2 or more of the ssDNA binding motifs 150, such as 3 or more of the ssDNA binding motifs 150, 4 or more of the ssDNA binding motifs 150, or 5 or more of the ssDNA binding motifs 150. In some embodiments, the DNA editing construct 100 includes two different types or more of the ssDNA binding motifs 150.

Linkers

[00114] Referring to Fig. 7A, in some embodiments, the construct 100 further include one or more linkers 170 connecting the DNA editing motif 110, the nuclear localization signal 130, and the ssDNA binding motif 150 to each other.

[00115] The nature of the linkers 170 is not limited, as long as they are able to survive the intracellular environment (and the extracellular environment if the storage, delivery, etc. of the construct 100 requires substantial exposure to such environments) without breaking substantially, and strong enough that the DNA editing motif 110, the NLS 130 and the ssDNA binding motif 150 do not substantially dissociated with each other before or during DNA editing.

[00116] In some embodiments, the one or more linkers 170 connect the DNA editing motif 110, the nuclear localization signal 130, and the ssDNA binding motif 150 to each other covalently.

[00117] In some embodiments, the one or more linkers 170 are polypeptide linkers. In some embodiments, each of the polypeptide linkers has the length ranging from about 2 amino acid residues to about 100 amino acid residues.

[00118] In some embodiments, the DNA binding motif 110, the NLS 130, the ssDNA binding motif 150, and the linker 170 are each of them independently selected polypeptides.

[00119] In some embodiments, the DNA binding motif 110, the NLS 130, the ssDNA binding motif 150, and the linker 170 are all part of the same polypeptide chain, and the construct 100 is one protein.

[00120] In some embodiments, the DNA binding motif 110, the NLS 130, the ssDNA binding motif 150, and the linker 170 are not part of the same polypeptide chain and the construct 100 includes two or more separate polypeptide chains.

[00121] In some embodiments, the construct 110 includes two or more proteins.

[00122] In some embodiments, one or more of the linkers 170 are multi-piece design such that components of the construct 100 can be delivered separately and form an integral construct 100 when inside cells or before entering the cytoplasm. In some embodiments, one or more of the linkers 170 are formed by two or more components that are able to associate with each other to form the linkers 170. Examples of multi-piece configurations of linkers 170 include linkers composed of protein A/G and antibody Fc domain, linkers composed of a single-chain variable fragment (scFv) and the target thereof, *etc.*

Nucleic Acid Encoding Constructs for DNA Editing

[00123] The study described herein (“the present study”) developed a novel DNA editing construct that edits DNA using a circular single stranded DNA (cssDNA) as a donor template, which has significantly higher DNA editing efficiency than similar conventional DNA editing constructs that also use cssDNA as donor templates. The novel DNA editing construct can be one or more proteins, which can be encoded by a nucleic acid such as a DNA molecule or an RNA molecule.

[00124] Accordingly, in some aspects, the present invention is directed to a nucleic acid encoding a construct for DNA editing. In some embodiments, the construct for DNA editing comprises one or more proteins that are the same as or similar to those described elsewhere, such as in the “Construct for DNA Editing” section.

[00125] As detailed elsewhere herein, the DNA editing motif 110 sometimes do not have site-specificity alone and depends on external components to provide the specificity. The external components are sometimes guide RNAs. Accordingly, in some embodiments, the nucleic acid further encodes a guide RNA (gRNA) that provides the site-specificity to the DNA editing motif.

[00126] As detailed elsewhere herein, the function of the construct 100 sometimes includes more than one protein, and/or requires guide RNA. In some embodiments, the nucleic acid encodes more than one protein, or encodes one or more proteins and at least one guide RNA. In some embodiments, the nucleic acid is polycistronic.

[00127] In some embodiments, the nucleic acid is comprised in an expression vector. Non-limiting examples of expression vector include plasmid vectors, RNA vectors, viral vectors, or the like.

[00128] In some embodiments, the nucleic acid is packaged in a viral particle, a lipid nanoparticle, a non-lipid nanoparticle, or the like.

Composition for Editing DNA

[00129] The study described herein (“the present study”) developed novel compositions for editing DNA molecules, which include the novel DNA editing constructs and a single stranded DNA (ssDNA) as a donor template. When introduced into cells, the DNA editing composition herein achieved significantly higher DNA editing efficiency than those achieved with conventional DNA editing compositions including the same ssDNA and conventional DNA editing constructs.

[00130] Accordingly, in some embodiments, the present invention is directed to a composition for editing DNA. In some embodiments, the composition edits nuclear DNA, such as genomic DNA. In some embodiments, the composition edits DNA via homologous recombination or homology directed repair, using the ssDNA as a donor template.

[00131] In some embodiments, the composition for editing DNA includes a DNA editing construct and a ssDNA donor template.

[00132] In some embodiments, the composition including both the DNA editing construct and the ssDNA donor template is prepared prior to the delivery into a cell to be edited.

[00133] In some embodiments, the DNA editing construct (or a nucleic acid encoding the same) and the ssDNA donor template are prepared separately and the composition is formed when, for example, both of the components are delivered into the cell.

[00134] In some embodiments, when the composition is delivered into a cell and/or formed inside the cell, the DNA editing construct (sometimes in concert with cellular machineries endogenous to the cell) inserts the DNA insert into the target location in the nuclear DNA via homologous recombination or homology directed repair.

[00135] In some embodiments, the DNA editing construct is the same as or similar to those as described elsewhere herein, such as in the “Constructs for DNA Editing” section. In some embodiments, the nucleic acid encoding the DNA editing construct is the same as or similar to those described elsewhere herein, such as in the “Nucleic Acid Encoding Constructs for DNA Editing” section.

[00136] In some embodiments, the ssDNA includes a DNA insert, a 5'-homology arm, and a 3'-homology arm. In some embodiments, the 5' homology arm and the 3' homology arm are complementary to the nuclear DNA in a target region.

[00137] In some embodiments, the ssDNA is a circular ssDNA (cssDNA), a linear ssDNA (lssDNA), a double stranded DNA/circular single stranded DNA (dsDNA/cssDNA) hybrid including a single stranded DNA portion, an RNA/DNA hybrid including a single stranded DNA portion, or the like.

[00138] In some embodiments, the ssDNA is chemically modified. In some embodiments, the ssDNA includes a modified base, a modified sugar group, a modified backbone, or the like.

[00139] In some embodiments, the ssDNA is produced from a cell free system or a cell culture system.

[00140] In some embodiments, each of the 5'-homology arm and the 3'-homology arm has a length ranging from about 25 nucleotides to about 5000 nucleotides, such as from about 50 nucleotides to about 5000 nucleotides, from about 75 nucleotides to about 2500 nucleotides, from about 100 nucleotides to about 2000 nucleotides, from about 150 nucleotides to about 1500 nucleotides, from about 200 nucleotides to about 1000 nucleotides, or from about 300 nucleotides to about 500 nucleotides.

[00141] The present study demonstrated that the DNA editing composition including the novel DNA editing construct herein can increase editing efficiencies for all tested insert sizes (~2kb,

~4kb, ~8kb). Accordingly, it is expected that the composition herein can be used to insert DNA inserts of all size insertable by conventional DNA editing compositions. In some embodiments, the DNA insert in the ssDNA donor template has a length ranging from 1 nucleotide to about 50,000 nucleotides, such as from 1 nucleotide to about 40,000 nucleotides, from 1 nucleotide to about 30,000 nucleotides, from 1 nucleotide to about 20,000 nucleotides, from 1 nucleotide to about 10,000 nucleotides, or from 1 nucleotide to about 8,000 nucleotides.

[00142] The ssDNA, such as cssDNA suitable for being included in the DNA editing composition herein are also described in, for example, WO2020142730A1. The entirety of this reference is hereby incorporated herein by reference.

Kit for Editing DNA

[00143] The study described herein (“the present study”) developed novel kit for editing DNA molecules, which include the novel DNA editing constructs (or nucleic acids encoding the constructs) and a single stranded DNA (ssDNA) as a donor template. When both components were introduced into cells, the combination of the DNA editing construct herein and the ssDNA donor template achieved significantly higher DNA editing efficiency than those achieved with conventional DNA editing combinations including the same ssDNA and conventional DNA editing constructs (or nucleic acids encoding such conventional editing constructs).

[00144] Accordingly, in some embodiments, the present invention is directed to a kit for editing DNA. In some embodiments, the components of the kit, when combined, edit nuclear DNA, such as genomic DNA. In some embodiments, the components of the kit edit DNA via homologous recombination or homology directed repair, using the ssDNA as a donor template.

[00145] In some embodiments, the kit for editing DNA includes a DNA editing construct (or a nucleic acid encoding the same) and a ssDNA donor template.

[00146] In some embodiments, the DNA editing construct is the same as or similar to those as described elsewhere herein, such as in the “Constructs for DNA Editing” section. In some embodiments, the nucleic acid encoding the DNA editing construct is the same as or similar to those described elsewhere herein, such as in the “Nucleic Acid Encoding Constructs for DNA Editing” section.

[00147] In some embodiments, the DNA editing construct (or a nucleic acid encoding the same) and the ssDNA donor templates are mixed together prior to delivery, or are delivered together or

separately into the cell. Either way, a composition for editing DNA (which, in some embodiments, is the same as or similar to those described elsewhere herein, such as in the “Composition for Editing DNA” section) is formed before, during, or after the components are delivered into the cell.

[00148] In some embodiments, when the components of the kits are delivered into a cell and/or formed inside the cell, the DNA editing construct (sometimes in concert with cellular machineries endogenous to the cell) inserts the DNA insert into the target location in the nuclear DNA via homologous recombination or homology directed repair.

[00149] In some embodiments, the DNA editing construct is the same as or similar to those as described elsewhere herein, such as in the “Constructs for DNA Editing” section. In some embodiments, the nucleic acid encoding the DNA editing construct is the same as or similar to those described elsewhere herein, such as in the “Nucleic Acid Encoding Constructs for DNA Editing” section.

[00150] In some embodiments, the kit for editing DNA does not include the ssDNA donor template. Rather, the kit includes a nucleic acid for producing the ssDNA donor template, such as in a cell free system or a cell culture system.

[00151] In some embodiments, the nucleic acid for producing the ssDNA donor template needs to be engineered to construct a nucleic acid that encodes the ssDNA donor template.

[00152] In some embodiments, the nucleic acid for constructing the ssDNA donor template encodes every elements of the ssDNA donor template, except for the DNA insert. This way, different DNA inserts can be incorporated into the nucleic acid, and the resulting nucleic acid can produce ssDNA that inserts the customized DNA into a predetermined location (such as a location that DNA insertions are not expected to cause any significant disruption to cellular functions) in the nuclear DNA using the 5' homology arm and the 3' homology arm are complementary to sequences adjacent to this location.

[00153] In some embodiments, the nucleic acid for constructing ssDNA donor template encodes every elements of the ssDNA donor template, except for the 5' homology arm and the 3' homology arm. This way, different 5' homology arm and 3' homology arm can be inserted into the nucleic acid, and the resulting nucleic acid can produce ssDNA that are able to insert DNA inserts at desired regions in the nuclear DNA.

[00154] In some embodiments, the nucleic acid for constructing ssDNA donor template encodes every elements of the ssDNA donor template, except for the 5' homology arm, the 3' homology arm, and the DNA insert. This way, the nucleic acid for constructing ssDNA donor template can be fully customized for the site of the editing as well as the sequence to be inserted.

[00155] In some embodiments, the nucleic acid for constructing ssDNA donor template includes a multiple cloning site for inserting the 5' homology arm, the 3' homology arm, and/or the DNA insert.

[00156] In some embodiments, the nucleic acid for constructing ssDNA donor template is a plasmid or a viral vector.

[00157] In some embodiments, the kit further includes a component for delivering the DNA editing construct (or the nucleic acid encoding the same) and the ssDNA. Non-limiting examples of components for delivering including a microinjector (for injecting directly into cells), a lipid for forming lipid nanoparticles, a gene gun and metal nanoparticles (such as gold or tungsten nanoparticles), a cuvette for electroporation, and the like.

Methods of Editing DNA

[00158] The study described herein ("the present study") developed novel DNA editing constructs that edit DNA using a single stranded DNA (ssDNA) as a donor template. The DNA editing constructs (or nucleic acids encoding the same) can be introduced into cells together with the ssDNA, and the DNA editing efficiency of this method is significantly higher than the similar conventional DNA editing methods that also use ssDNA as donor templates.

[00159] Accordingly, in some aspects, the present invention is directed to methods of editing DNA, such as nuclear DNA in a cell.

[00160] In some embodiments, the method includes introducing a DNA editing construct or a nucleic acid encoding the DNA editing construct, as well as a ssDNA donor template into the cell. In some embodiments, the DNA editing construct and/or the nucleic acid is the same as or similar to those described elsewhere herein, such as in the "Construct for DNA Editing" and "Nucleic Acid Encoding Construct for DNA Editing" sections.

[00161] In some embodiments, the ssDNA includes a DNA insert, a 5'-homology arm, and a 3'-homology arm. In some embodiments, the 5' homology arm and the 3' homology arm are complementary to the nuclear DNA in a target region. In some embodiments, the method inserts

the DNA insert into the target location in the nuclear DNA via homologous recombination/homology directed repair.

[00162] In some embodiments, the ssDNA is a circular ssDNA (cssDNA), a linear ssDNA (lssDNA), a double stranded DNA/circular single stranded DNA (dsDNA/cssDNA) hybrid including a single stranded DNA portion, an RNA/DNA hybrid including a single stranded DNA portion, or the like.

[00163] In some embodiments, the ssDNA is chemically modified. In some embodiments, the ssDNA includes a modified base, a modified sugar group, a modified backbone, or the like.

[00164] In some embodiments, the ssDNA is produced from a cell free system or a cell culture system.

[00165] In some embodiments, each of the 5'-homology arm and the 3'-homology arm has a length ranging from about 25 nucleotides to about 5000 nucleotides, such as from about 50 nucleotides to about 3000 nucleotides, from about 75 nucleotides to about 2500 nucleotides, from about 100 nucleotides to about 2000 nucleotides, from about 150 nucleotides to about 1500 nucleotides, from about 200 nucleotides to about 1000 nucleotides, or from about 300 nucleotides to about 500 nucleotides.

[00166] The present study demonstrated that the DNA editing method using the novel DNA editing construct herein can increase editing efficiencies for all tested insert sizes (~2kb, ~4kb, ~8kb). Accordingly, it is expected that the method herein can be used to insert DNA inserts of all size insertable by conventional methods. In some embodiments, the DNA insert has a length ranging from 1 nucleotide to about 50,000 nucleotides, such as from 1 nucleotide to about 40,000 nucleotides, from 1 nucleotide to about 30,000 nucleotides, from 1 nucleotide to about 20,000 nucleotides, from 1 nucleotide to about 10,000 nucleotides, or from 1 nucleotide to about 8,000 nucleotides.

[00167] The ssDNA, such as cssDNA suitable for used in the DNA editing method herein are also described in, for example, WO2020142730A1. The entirety of this reference is hereby incorporated herein by reference.

[00168] The present study demonstrated that the DNA editing constructs and the ssDNAs can be delivered into the cell using various methods. Accordingly, in some embodiments, the constructs or nucleic acids and the ssDNAs are delivered into the cell by a viral vector, a lipid or

non-lipid nanoparticle delivery, a gold particle mediated delivery, an electroporation delivery, and/or an injection.

[00169] In some embodiments, the DNA editing construct is delivered into the cell as the construct (such as one or more proteins), or as a nucleic acid such as an RNA molecule or a DNA molecule.

[00170] In some embodiments the nuclear DNA is a genomic DNA.

[00171] In some embodiments, the cell is a prokaryotic cell, a cultured plant cell, a eukaryotic cell, a cultured mammalian cell, a primary cell, an embryonic stem cell, an adult stem cell, an induced pluripotent stem cell, a hematopoietic stem cell or a CAR-T cell.

[00172] In some embodiments, the cell is in the body of a subject. In some embodiments, the subject is a mammal, such as a human. In some embodiments, the method is a method for *in vivo* DNA engineering.

[00173] In some embodiments, instead of editing a nuclear DNA in a cell, the method is an *ex vivo* DNA editing method. In some embodiments, the method edits DNA in a cell free system.

Non-Natural Polypeptides

[00174] The present study discovered that L2 region of the RecA protein from the bacterium *D. radiodurans* is able to significantly increase the DNA editing efficiency when incorporated into the DNA editing construct herein (see e.g., Fig. 3C). Without wishing to be bound by theory, it is hypothesized that the L2 region of *D. radiodurans* RecA (having the amino acid sequence NQVREKIGVMYGNPETTTGG, SEQ ID NO:11) can form stable complex with ssDNA molecules.

[00175] Accordingly, in some aspects, the present invention is directed to a non-natural polypeptide including the *D. radiodurans* RecA L2 region, or amino sequences derived from the *D. radiodurans* RecA L2 region.

[00176] In some embodiments, the non-natural polypeptide includes an amino acid sequence having 90% or higher identity, such as 95% or higher identity or identical to the amino acid sequence set forth in SEQ ID NO:11.

[00177] In some embodiments, the non-natural polypeptide is a ssDNA-binding polypeptide.

Examples

[00178] The instant specification further describes in detail by reference to the following experimental examples. These examples are provided for purposes of illustration only, and are not intended to be limiting unless so specified. Thus, the instant specification should in no way be construed as being limited to the following examples, but rather, should be construed to encompass any and all variations which become evident as a result of the teaching provided herein.

Example 1: Engineering Tripartite Gene Editing Machinery to Optimize Targeted Genome Integration

[00179] Non-viral DNA donor template has been widely used for targeted genomic integration by homologous recombination (HR). This process has become more efficient with RNA guided endonuclease editor system such as CRISPR/Cas9. Circular single stranded DNA (cssDNA) has been harnessed previously as a superior genome engineering catalyst (GATALYST) for efficient and safe targeted gene knock-in. However, the engineering efficiency is bottlenecked by the nucleoplasm delivery of cssDNA donor, especially for large transgene integration. The present study developed enGager, a set of enhanced GATALYST associated genome editors by fusion of nucleus localization single (NLS) peptide tagged Cas9 with various single stranded DNA binding protein modules. These enGagers form tripartite complex with sgRNA and cssDNA donors as integrative genome integration machinery, thereby facilitate the nucleus shuttling of DNA donors and increase their active local concentration. When applied for targeted genome integration with cssDNA donor templates to diverse genomic loci in various cell types, these enGagers outperform unfused editors. The enhancement of integration efficiency ranges from 1.5- to more than 6-fold, with the effect being more prominent for larger than 4Kb transgene knock-in. The present study further demonstrated that enGager mediated enhancement for genome integration is ssDNA, but not dsDNA dependent. Using one of the small enGagers, the present study demonstrated large chimeric antigen receptor (CAR) transgene integration in primary T cells with exceptional efficiency and anti-tumor cell function. These tripartite editors with ssDNA optimized genome engineering system (TESOGENASETM) add a set of novel endonucleases into the gene-editing toolbox for potential cell and gene therapeutic development based on ssDNA mediated non-viral genome engineering.

Example 2

[00180] Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and CRISPR-associated protein 9 (Cas9) is an effective and precise gene-editing tool that is revolutionizing diverse fields of biotechnology, medical and agricultural research. Cas9 protein combined with 20 nucleotides (nt) guide RNA finds and induces precise double-stranded breaks (DSB) to the complementary region of the guide RNA for gene deletion, repair or insertion in prokaryotes and eukaryotes. Stemmed from this innovative tool, researchers around the world have developed many different applications of CRISPR/Cas9 such as gene knock-out using Non-Homogenous End Joining (NHEJ), correction of single point mutation by base editors that convert C to T or A to G nucleotide in a site specific manner, and targeted transcription regulation using catalytically dead Cas9 (dCas9) fused epigenetic regulators such as histone modifiers and DNA methylation modifiers. However, targeted insertion of a new DNA sequence has been more challenging than the other CRISPR/Cas9 mediated applications. Conventional targeted gene knock-in (KI) is achieved by homologous recombination (HR) which is extremely inefficient, especially for longer sequences. Homology directed repair (HDR), a highly regulated spatial and temporal cellular process induced by double stranded DNA break (DSB), significantly increases the targeted genome integration frequency. Currently, HDR is one of the most efficient ways to insert a few bases to 2Kb length of genes but is extremely inefficient for > 4Kb large payload.

[00181] To achieve efficient targeted genome integration, Adeno-Associated Virus Type 6 (AAV6) has been demonstrated as a prevalent donor carrier for HDR mediated genome engineering. For instance, AAV6 have been demonstrated for efficient beta-globin gene integration in human and rodent hematopoietic cells as well as human stem cells. However, recombinant AAV has been reported with safety issues and manufacturing challenges and succumbed for the use of large genome integration due to its packaging limit of ~ 4.7 Kb. Non-viral targeted gene KI relies primarily on double-stranded DNA (dsDNA) and linear single-stranded DNA (lssDNA) donors. These nucleic acid donor templates have been demonstrated relatively efficient gene targeting in various primary cells, including T cells and natural killer cells. Recently, CD19 CAR-T cells engineered on an immune checkpoint PD-1 locus with linear dsDNA donor template have been reported to treat B cell Non-Hodgkin lymphoma (B-NHL) patients with desired safety and efficacy. However, both forms of DNA template are not ideal HDR donor payloads. While dsDNA donors are easy to access, it succumbs from low editing

efficiency and cGAS dependent cytotoxicity. lssDNA has better performance for targeted gene integration but has major manufacture challenges. Most recently, circular single stranded DNA (cssDNA) has been demonstrated as a superior donor template for highly efficient genome targeting in mammalian cells. However, non-viral DNA donor template mediated genome integration is largely bottlenecked by the nucleus delivery of the donor DNAs, especially for large transgene integration. Therefore, developing a more efficient gene integration method, especially for long sequences with non-viral donor payload is an essential milestone to unlock the full potential for gene editing.

[00182] As HDR mediated genome integration is a highly regulated spatial and temporal process in nucleoplasm, the strategy to enhance the KI efficiency using DNA donor is 1) to efficiently deliver donor DNA into the nucleus and 2) to recruit donor DNA onto the targeted genomic locus, both of which are critical steps to increase the local concentration of the donor template to ensure effective homologous searching and pairing to complete the HDR process. This idea has been verified by multiple studies using different strategies. For instance, physically coupling dsDNA donor templates together with Cas9 endonuclease by fusing Cas9 together with transcription factors or homologous recombinase such as Rad51, Rad52, POLD3 and the dominant negative form of 53BP1 has been shown to increase dsDNA mediated transgene KI. Other studies engineered the single or double stranded donor DNA with chemical modification such as biotinylation, 5'-triethylene glycol (TEG) modification or covalent conjugation of donor DNA with Cas9 endonuclease. Similar strategy has been taken to create a sgRNA/donor DNA hybrid or a truncated Cas9 binding site on the donor DNA to achieve more efficient HDR mediated genome integration. Overall, majority the strategies focus on dsDNA as donor templates, with limited optimization on ssDNA donor templates.

[00183] Circular single stranded DNA (cssDNA) was harnessed as a superior genome engineering catalyst (GATALYST) for efficient and safe targeted gene KI across multiple cell types to many targeted loci in the mammalian genome. Here it was hypothesized that coupling GATALYST donor template with endonuclease editor as a protein/sgRNA/ssDNA tripartite complex can effectively increase the spatial and temporal enrichment of ssDNA donor template in the targeted genome microenvironment, thereby enhance the ssDNA mediated genome integration. Using a functional screening method based on a GFP reporter KI, the present study developed a set of enhanced GATALYST associated genome editors (enGager) by fusion of

nucleus localization signal (NLS) peptide tagged wild type Cas9 together with various single stranded DNA binding protein modules. These enGagers form tripartite complex with sgRNA and cssDNA donors as integrative genome integration machinery, thus facilitate the nucleus translocation and increase the active local concentration of donor DNA. When applied for targeted genome integration with cssDNA donor templates to diverse genomic loci in various cell types, these enGagers outperform unfused canonical endonuclease editors, in plasmid, mRNA and protein forms. The enhancement of integration efficiency ranges from 1.5- to more than 6-fold, with the effect being more prominent for larger than 4Kb transgene KI and in primary cells such as human T cells. These rationally designed tripartite editors with ssDNA optimized genome engineering system (TESOGENASE™) add a set of novel endonucleases into the gene-editing toolbox for potential cell and gene therapeutic development based on ssDNA mediated non-viral genome engineering.

Example 3: Fusion of Cas9 and homologous recombination proteins enhance the ssDNA mediated knock-in

[00184] The cssDNA acts as genome engineering catalyst (GATALYST) to enable safe and efficient targeted genome KI for ultra-large transgene integration. To further optimize the cssDNA mediated gene knock-in, the present study sought to develop a set of enhanced GATALYST associated genome editors (enGager) that could form an endonuclease/sgRNA/ssDNA tripartite complex by physically recruiting the cssDNA donors, and facilitate the nucleus shuttling of the DNA donor by virtue of the nucleus localization signal (NLS) peptide tagged on the endonuclease editors. To develop these enGagers, the present study initially utilized an all-in-one (AIO) plasmid construct to engineer dual NLS tagged Cas9 fusion constructs for a GFP reporter screening (Cong et al., *Science* 339, 819–823, 2013). These enGager constructs contain bi-cistronic expression cassettes for Cas9-fusion endonuclease proteins and a single-guide RNA (sgRNA) targeting against the human *RAB11A* locus (Roth et al., *Nature* 559, 405–409, 2018). The present study initially tested the Cas9 fusion with full length *E. coli* RecA and its human homolog protein Rad51. For potential better performance of Rad51-Cas9 fusion, the present study used two mutant forms of Rad51 (Rees et al., *Nat Commun* 10, 2212, 2019). As the full length RecA and Rad51 has both recombinase activity and DNA binding functions, to uncouple these two functions, the present study also engineered an AIO

construct by fusion of Cas9 with a 36 amino acid (AA) peptide “Brex” which is known as a RecA binding motif to recruit endogenous RecA/Rad51 recombinase, but does not have DNA binding capability in the cells (Fig. 1A).

[00185] To directly compare the knock-in efficiency with various AIO enGager constructs, 2 ug of various size of GFP reporter cssDNA donor templates flanked with 306 nt 5' homology arm and 315 nt 3' homology arm targeting *RAB11A* locus were tested by co-electroporation delivery in K562 cells (Fig. 1B). As illustrated in Fig. C, when DNA donor templates were co-electroporated with AIO constructs consisting of spCas9 fusion enGagers and *RAB11A* sgRNA, precise HDR of the locus can be evaluated by GFP reporter expression using flow cytometry. Day 3 post-electroporation, knock-in efficiency as measured by GFP+ cell percentage was much higher for Cas9-RecA, Cas9-Rad51 (AE) and Cas9-Rad51 (SEAD) enGager constructs (11.5%-15.4%) than Cas9 only construct (6.8%) for a 2 µg of cssDNA donor with 2Kb payload (Fig. 1C). In contrast, no knock-in efficiency enhancement was observed with Cas9-Brex fusion construct. Interestingly, the KI efficiency by Cas9-Brex fusion nuclease is even lower than Cas9 alone (4.9% vs 6.8%) indicating that only deploying the recombination function is not sufficient to enhance cssDNA mediated genome targeting. The present study found that the enGager induced knock-in enhancement is durable over time at 8 days and 14 days post-electroporation. In particular, the Cas9-RecA fusion enGager is able to induce >3-fold KI enhancement as compared to unfused Cas9 (Fig. 1D). The present study further tested the knock-in efficiency at *RAB11A* locus in K562 cells for a larger cssDNA donor with a 4Kb GFP payload. EnGagers Rad51(AE), Rad51(SEAD) and RecA induced GFP KI enhancement by 1.78- to 2.35-fold compared to Cas9 alone (Fig. 1E-1F). Consistently, the Cas9-Brex fusion does not produce benefit for 4 Kb payload KI.

Example 4: Engineering small enGagers with Cas9 fused ssDNA binding motifs

[00186] Based on the Brex- and full-length recombinase-Cas9 fusion results, it was hypothesized that the rate-limiting step to enhance the ssDNA mediated HDR is to increase the effective nucleoplasm delivery of the donor template, which in turn increase the local concentration of the tripartite HDR machinery on the targeted genomic locus including the DNA donor template. RecA family DNA recombinases have an evolutionarily conserved ssDNA binding L2 loop peptide that encompasses 20-24 amino acid with an extremely conserved central

aromatic residue. To test the hypothesis, the present study engineered smaller enGager editors that only capture the DNA binding feature. The present study selected the 20 amino acid peptide sequences FECO, WECO and YECO from Voloshin *et al.* which demonstrated several fold higher binding affinity to ssDNA than to dsDNA in sequence independent manner (Science 272, 868–872, 1996). The present study engineered smaller AIO constructs with Cas9 fused to either one copy or three tandem copies of these DNA binding motifs linked with multiple GS linker sequences (Fig.2A). When co-electroporated together with 2 ug of 2 Kb cssDNA *RAB11A* GFP donor template into K562 cells, these newly designed smaller enGagers demonstrated comparable level of enhanced knock-in efficiency to RecA full length enGager. In most of the cases, the fold changes range from 1.40 to 1.60, whereas the RecA enGager enhances 1.57-fold KI compared to Cas9 alone (Figs. 2C-2D). These data validated the hypothesis that just installing the sequence independent DNA binding capability to nuclease editors could effectively stimulate the homology recombination mediated gene insertion. By doing so, the present study also discovered a set of smaller rationally designed enGagers that could be powerful editing tools for genome writing.

Example 5: Identification of additional enGagers from Cas9-ssDNA binding module chimeras

[00187] To engineer a bigger collection of AIO enGager constructs, the present study performed an extensive search of RecA family recombinase and other protein families that has ssDNA binding capability, either small motif peptide or structured protein domain. The present study tested DrRecA full length and its 20 amino acid L2 peptide, *E. coli* ssDNA binding protein, Lambda Red subunit of bacteriophage recombineering protein, *E. coli* ssDNA annealing protein RecT. (Figs .3A-3B). These fusion constructs were screened using the 2 Kb cssDNA *RAB11A* GFP payload in K562 cells. The majority of the fusion constructs outperformed Cas9 WT by 1.79- to 2.43-fold without compromising the cell viability (Figs. 3C-3D). Interestingly, the enGager with DrRecA full length fusion stimulates *RAB11A* GFP KI by 2.17-fold and the homologous 20 amino acid L2 peptide from DrRecA further improve the KI efficiency level to 2.43-fold of that of WT Cas9. As a comparison, FECO and RecA enGagers increase KI efficiency by 1.59- and 1.87-fold respectively. It is worth noting that *Deinococcus Radiodurans* Bacterium is one of the most radiation-resistant organism known to date due to its

exceptional DNA repair capability by its DNA recombinase DrRecA. It appears that the L2 loop peptide from the whole RecA phylogeny can be good candidate fusion peptides for small version of enGagers to stimulate HDR mediated genome integration. Figs. 3E-3F highlighted a few RecA L2 peptide sequences from archaea bacteria, *E. coli* and mammalian organism with a central amino acid with aromatic ring.

Example 6: EnGager mediated genome integration enhancement is specific to ssDNA

[00188] mRNA has been an attractive approach to reduce cytotoxicity and enhance delivery using lipid nanoparticle technology for vaccine development and gene therapy. Hereafter, except otherwise mentioned, the present study used mRNA form of RecA and FECO enGagers for the remaining of this study (Fig. 4A). Initially, the RecA enGager mediated enhancement of a 2Kb *RAB11A* GFP reporter integration in K562 cells by electroporation delivery was validated. When co-electroporated with various dose of ssDNA donor templates at 0.3, 1, 1.5, 2 to 3 ug per reaction, RecA mRNA enGager (at 1 ug per reaction) significantly increases the GFP KI compared to WT Cas9 mRNA (Figs. 4B-4C). As RecA L2 peptides also have low binding affinity to dsDNA, the present study sought to investigate if these enGagers can facilitate HDR mediated KI by dsDNA donor templates. Using co-electroporation of 1 ug mRNA enGagers/sgRNA and 1 ug of 2Kb *RAB11A* GFP donor template either in ssDNA or dsDNA form, the present study found that enGagers stimulate significant KI efficiency only for ssDNA but not dsDNA donor. Specifically, the FECO, RecA, GS-FECO and Brex-FECO mRNA enGagers are able to enhance GFP KI to 1.68- to 1.99-fold of a WT Cas9 mRNA 5 days post-electroporation (Figs. 4D-4G). Consistently, when the present study tested the liposome mediated co-transfection of enGager mRNA/sgRNA together with ssDNA or dsDNA donor templates in HK293 cells, only ssDNA but not dsDNA donors demonstrated 1.35- to 1.99-fold KI enhancement by FECO, Brex-FECO and GS-FECO mRNA enGagers 5 days post-delivery (Figs. 4H-4I). As a reference, WT Cas9 and Cas9-Brex fusion were used as controls in these experiments.

Example 7: Engagers in mRNA form enhance ultra-large transgene integration on various genomic loci

[00189] The present study then assessed the effect of enGagers more broadly on various genomic loci and for donor payload with various size, especially 4 and 8Kb payloads that are around AAV6 and lentivirus packaging limit, respectively. Utilizing highly purified mRNA form of FECO and RecA enGagers, the present study tested the efficiency of GFP transgene KI on two genomic loci with various length of payload. For *RAB11A* target locus, the present study designed 2Kb, 4Kb and 8Kb cssDNA payloads, for another target site, the clinically relevant immune cell therapy *B2M* locus (Ren et al., *Clinical Cancer Research* 23, 2255–2266, 2017), the present study designed 2Kb and 4Kb cssDNA payloads (Figs. 5A and 5D). As illustrated in Fig. 5B, when mRNA enGager/sgRNA with cssDNA donor templates were co-electroporated to K562 cells, FECO and RecA enGagers increase the GFP transgene KI to 44.6–48.5% from 30.10% by WT Cas9 for 2Kb *RAB11A* GFP payload. For 4Kb payload, FECO and RecA enGagers increase the GFP transgene KI to 11.07–13.77% from 6.17% by WT Cas9. For 8Kb payload, FECO and RecA enGagers increase the GFP transgene KI to 3.73–5.17% from 2.97% by WT Cas9. Similarly, for *B2M* locus, FECO and RecA enGagers increase the KI efficiency to 39.6–43.97% from 30.13% by WT Cas9 for 2Kb GFP transgene, and increase the KI efficiency to 10.67–14.07% from 6.27% by WT Cas9 for 4Kb GFP transgene (Fig. 5E). As a negative control, Cas9-Brex fusion did not enhance cssDNA mediated transgene KI. Cell viability was not affected by 1 ug Cas9 or enGager mRNA (Figs. 5B–5C and 5E–5F) as compared to Mock electroporation.

Example 8: CAR-T engineering by enGager with superior efficiency than WT Cas9

[00190] Finally, the present study sought to use the newly developed FECO enGager to engineer primary T cells by inserting a functional chimeric antigen receptor (CAR) transgene cassette on T cell receptor alpha constant chain (*TRAC*) locus (Eyquem et al., *Nature* 543, 113–117, 2017). The present study chose a ~3Kb CD19-CD22 dual CAR construct that has been demonstrated anti-tumor function for potential treatment of patient with Acute Lymphoblastic Leukemia (ALL) (Fig. 6A) (Fry et al., *Nat Med* 24, 20–28, 2018). The present study used PE conjugated Protein-L binder to analyze the KI efficiency. When delivered by co-electroporation with CD19-CD22 dual CAR cssDNA doner template in CD4⁺/CD8⁺ primary pan-T cells, GS-FECO mRNA enGager achieved 30.2% to 33.4% targeted CAR KI at day 7 post-delivery, whereas WT Cas9 mRNA only achieved 5% to 14.1% of CAR KI with various mRNA dosages.

Especially when lower dose of mRNA editors was used, GS-FECO enGager enhance cssDNA mediated CAR-T engineered by >6 fold (Fig. 6B). Importantly, GS-FECO enGager mediated enhancement of dual CAR-T engineering did not compromise T cell counts, proliferation and cell viability compared to WT Cas9 (Fig. 6C). To test the anti-tumor function of the engineered CAR-T cells, the present study used the Incucyte live imaging approach to monitor the cell killing functions over the course of 94 hours. The present study titrated the E:T ratio at 2:25:1, 4.5:1 and 9:1. When co-cultured with a leukemia lymphocyte NALM6 cells, GS-FECO enGager mRNA engineered CD19-CD22 dual CAR-T cells demonstrated more effective and durable cancer cell killing function compared to WT Cas9 mRNA engineered dual CAR-T cells at every E:T ratio tested (Figs. 6D-6E).

Example 9

[00191] In the present study, a set of enGager endonucleases by fusing Cas9 with various single stranded DNA binding protein modules, including some small versions based on as few as 20 amino acid motifs, was presented. When applied for targeted genome integration with cssDNA donor templates to diverse genomic loci in various cell types, these enGagers outperform unfused editors by up to 6-fold for small and large transgene knock-in. The present study further demonstrated that enGager mediated enhancement for genome integration is ssDNA dependent, with minimal effect for dsDNA donors. Using one of the small enGagers, it was demonstrated that large chimeric antigen receptor (CAR) transgene integration in primary T cells resulted in exceptional efficiency and anti-tumor cell function. As shown in Fig. 6F schematic diagram, the engineered enGagers with single stranded DNA binding protein (SSBP) can recruit cssDNA donor template and form a tripartite editing machinery for efficient translocation of the entire tripartite editing complex from cytoplasm to nucleoplasm. As a result, the donor DNA has higher effective local concentration in the nucleus and to the targeted genomic locus for more efficient homologous directed genome integration (bottom). In contrast, Cas9 alone is not able to engage and recruit cssDNA donor template, therefore the donor DNA templates are restricted in the cytoplasm with relatively low level of gene KI efficiency. These newly designed enGager editors expanded the gene-editing toolbox for potential cell and gene therapeutic development based on ssDNA mediated non-viral genome engineering.

Example 9: Materials and Methods*Plasmid constructs*

[00192] For all-in-one (AIO) plasmid construct design for various enGagers (Table 1), ssDNA binding protein and peptide sequences were synthesized and assembled into an AIO plasmid modified from Addgene plasmid #42230 by golden gate cloning. Complex fusion constructs were cloned using multi-piece DNA Gibson assembly approach. For mRNA constructs, various enGager coding sequences were cloned into a modified vector plasmid from pGEM®-4Z vector containing T7 promoter, 5' UTR, Cas9 fusion sequence ORF 3' UTR, bovine growth hormone polyadenylation signal (bGH polyA), and 64 poly adenine sequence (Promega). Table 1 listed the amino acid sequences of the enGager fusion protein and WT Cas9 endonuclease.

Table 1: Amino Acid Sequences of enGager Fusion Proteins Used in Present Study

1. Cas9 WT
MDYKDHDGDYKDHDIDYKDDDDKMAPKKRKVGIHGVPAADKKYSIGLDIGTNSV GWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTR RKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP TIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQT YNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRV NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPKEYKEIFFDQSKNGYAGYIDGG ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR QEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDK GASAQSFIERMTNFDKNLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLS GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLK IHKDKDFLDNEENEDILEDIVLTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTG WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQ GDSLHEHIANLAGSPAIKKGILQTVKVVDLVKVMGRHKPENIVIAMARENQTTQKG QKNSRERMKRIEELGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYD ENDKLIREVKVITLKSCLVSDFRKDFQFYK VREINNYHHAHDAYLNAVVG TALIKKYP

KLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR
 PLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDK
 LIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSE
 KNPIDFLEAKGYKEVKKDLIIKLPKYSLEFENGRKRMLASAGELQKGNELALPSKYV
 NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLS
 AYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQSI
 TGLYETRIDLSQLGGDKRPAATKKAGQAKKKK (SEQ ID NO:13)

Residues 1-23: Flag tag

Residues 24-40: SV40 large T antigen nuclear localization signal (SV40 NLS)

Residues 41-1407: Cas9

Residues 1408-1423: Nucleoplasmin nuclear localization signal (Nucleoplasmin NLS)

2. Cas9-RecA

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAAADKKYSIGLDIGTNSV
 GWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTR
 RKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP
 TIYHLRKKLVSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQT
 YNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTP
 NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRV
 NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
 ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR
 QEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDK
 GASAQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLS
 GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLK
 IIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG
 WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQ
 GDSLHEHIANLAGSPAIKKGILQTVKVVDLVKVMGRHKPENIVIEMARENQTTQKG
 QKNSRERMKRIEELGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
 NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL
 LNAKLITQRKFDNLTKAERGGSELDAKAGFIKRQLVETRQITKHVAQILDSRMNTKYD
 ENDKLIREVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYP
 KLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR

PLIETNGETGEIVWDKGRDFATVRKVLSPQVNVKKTEVQTGGFSKESILPKRNSDK
 LIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSE
 KNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYV
 NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLS
 AYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI
 TGLYETRIDLSQLGGDKRPAATKKAGQAKKKKEFGGGGSMADENKQKALAAALGQ
 IEKQFGKGSIMRLGEDRSMDETISTGSLSLDIALGAGGLPMGRIVEIYGPESSGKTTLT
 LQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPDGTGEQALEICDALARS
 GAVDVIVVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKL AGNLQSNNTLLIFI
 NQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGSETRVKVVKV
 KIAAPFKQAEFQILYGEINFIYGELVDLGVKEKLEKAGAWYSYKGEKIGQGKANAT
 AWLKDNPETAKEIEKKVRELLLSNPNSTPDFSVDDSEGVAETNEDFPKKKRKV (SEQ
 ID NO:14)

Residues 1-23: Flag tag

Residues 24-40: SV40 large T antigen NLS

Residues 41-1407: Cas9

Residues 1408-1423: Nucleoplasmin NLS

Residues 1424-1430: GS linker

Residues 1431-1783: RecA

Residues 1784-1790: SV40 NLS

3. Cas9-FECO

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAAADKKYSIGLDIGTNSV
 GWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTR
 RKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP
 TIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQT
 YNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTP
 NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRV
 NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
 ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR
 QEDFYFPFLKDNREKIEKILTFRIPIYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDK
 GASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLS

GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLK
 IIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG
 WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQ
 GDSLHEHIANLAGSPAIIKKGILQTVKVVDDELVKVMGRHKPENIVIEMARENQTTQKG
 QKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
 NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL
 LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYD
 ENDKLIREVKVITLKSCLVSDFRKDFQFYKVRINNYYHHAHDAYLNAVVGTAIIKKYP
 KLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR
 PLIETNGETGEIVWDKGRDFATVRKVLSPQVNVKKTEVQTGGFSKESILPKRNSDK
 LIARKKDWDPKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSE
 KNPIDFLEAKGYKEVKKDLIIKLPKYSLENGRKRMLASAGELQKGNELALPSKYV
 NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLS
 AYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQSI
 TGLYETRIDLSQLGGDKRPAATKKAGQAKKKKNQIRMKIGVMFGNPETTTGGPKKKR
 KV (SEQ ID NO:15)

Residues 1-23: Flag tag

Residues 24-40: SV40 large T antigen NLS

Residues 41-1407: Cas9

Residues 1408-1423: Nucleoplasmin NLS

Residues 1424-1443: FECO peptide

Residues 1444-1450: SV40 large T antigen NLS

4. Cas9-3XFECO

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIIHGVPAAADKKYSIGLDIGTNSV
 GWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTR
 RKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP
 TIYHLRKKLVSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQT
 YNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLFGNLIASLGLTP
 NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRV
 NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
 ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR

QEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEFVVDK
 GASAQSFIERMTNFDKNLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLS
 GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLK
 IIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG
 WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQ
 GDSLHEHIANLAGSPAIIKKGILQTVKVVDDELVKVMGRHKPENIVIAMARENQTTQKG
 QKNSRERMKRIEEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
 NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL
 LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYD
 ENDKLIREVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYP
 KLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR
 PLIETNGETGEIVWDKGRDFATVRKVLSPQVNVKKTEVQTGGFSKESILPKRNSDK
 LIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKGKSKKLKSVKELLGITIMERSSE
 KNPIDFLEAKGYKEVKKDLIIKLPKYSLENGRKRMLASAGELQKGNELALPSKYV
 NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLS
 AYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQSI
 TGLYETRIDLSQLGGDKRPAATKKAGQAKKKKNQIRMKIGVMFGNPETTTGGSGGSS
 GGSSGSETPGTSESATPESSGGSSGGSNQIRMKIGVMFGNPETTTGGSGGSSGGSSGSE
 TPGTSESATPESSGGSSGGSNQIRMKIGVMFGNPETTTGGPKKKRKV (SEQ ID NO:16)

Residues 1-23: Flag tag

Residues 24-40: SV40 large T antigen NLS

Residues 41-1407: Cas9

Residues 1408-1423: Nucleoplasmin NLS

Residues 1424-1443: FECO peptide

Residues 1444-1475: GS linker

Residues 1476-1495: FECO peptide

Residues 1496-1527: GS linker

Residues 1528-1547: FECO peptide

Residues 1548-1554: SV40 NLS

5. Cas9-WECO

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAAADKKYSIGLDIGTNSV
 GWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTR
 RKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP
 TIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQT
 YNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLFGNLIALSLGLTP
 NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRV
 NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
 ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR
 QEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDK
 GASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLS
 GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLK
 IIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG
 WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQ
 GDSLHEHIANLAGSPAIIKKGILQTVKVVDL VKVMGRHKPENIVIEMARENQTTQKG
 QKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
 NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL
 LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYD
 ENDKLIREVKVITLKS KLVSDFRKDFQFYK VREINNYHHAHDAYLNAVVG TALIKKYP
 KLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR
 PLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDK
 LIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSE
 KNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYV
 NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLS
 AYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI
 TGLYETRIDLSQLGGDKRPAATKKAGQAKKKKNQIRMKIGVMWGNPETTTPKKKRK
 V (SEQ ID NO:17)

Residues 1-23: Flag tag

Residues 24-40: SV40 large T antigen NLS

Residues 41-1407: Cas9

Residues 1408-1423: Nucleoplasmin NLS

Residues 1424-1443: WECO peptide

Residues 1444-1450: SV40 large T antigen NLS

6. Cas9-3XWECO

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAADKKYSIGLDIGTNSV
 GWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTR
 RKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP
 TIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQT
 YNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
 NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLS DAILLSDILRV
 NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
 ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR
 QEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDK
 GASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLS
 GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLK
 IIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG
 WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF MQLIHDDSLTFKEDIQKAQVSGQ
 GDSLHEHIANLAGSPAIIKKGILQTVKVVDDELVKVMGRHKPENIVIEMARENQTTQKG
 QKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
 NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL
 LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYD
 ENDKLIREVKVITLKS KLVSDFRKDFQFYK VREINNYHHAHDAYLNAVVG TALIKKYP
 KLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR
 PLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDK
 LIARKKDWDPKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERS SFE
 KNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYV
 NFLYLASHYEKLGKSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKR VILADANLDKVLS
 AYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI
 TGLYETRIDLSQLGGDKRPAATKKAGQAKKKKNQIRMKIGVMWGNPETTTGGSGGS
 SGGSSGSETPGTSESATPESSGGSSGGSNQIRMKIGVMWGNPETTTGGSGGSSGGSSGS
 ETPGTSESATPESSGGSSGGSNQIRMKIGVMWGNPETTTGGPKKKRKV (SEQ ID
 NO:18)

Residues 1-23: Flag tag

Residues 24-40: SV40 large T antigen NLS

Residues 41-1407: Cas9

Residues 1408-1423: Nucleoplasmin NLS

Residues 1424-1443: WECO peptide

Residues 1444-1475: GS linker

Residues 1476-1495: WECO peptide

Residues 1496-1527: GS linker

Residues 1528-1547: WECO peptide

Residues 1548-1554: SV40 NLS

7. Cas9-YECO

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAAADKKYSIGLDIGTNSV
 GWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTR
 RKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP
 TIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQT
 YNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
 NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSAAILSDILRV
 NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPKEYKEIFFDQSKNGYAGYIDGG
 ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR
 QEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDK
 GASAQSFIERMTNFDKNLPNEKVLPKHSLLEYEFTVYNELTKVKYVTEGMRKPAFLS
 GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLK
 IIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTG
 WGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFQMQLIHDDSLTFKEDIQKAQVSGQ
 GDSLHEHIANLAGSPAIIKKGILQTVKVVDDELVKVMGRHKPENIVIAMARENQTTQKG
 QKNSRERMKRIEIEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
 NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWRQL
 LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYD
 ENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYP
 KLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR
 PLIETNGETGEIVWDKGRDFATVRKVLSPQVNVKKTEVQTGGFSKESILPKRNSDK
 LIARKKDWDPKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSE

KNPIDFLEAKGYKEVKKDLIIKLPKYSLENGRKRMLASAGELQKGNELALPSKYV
 NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLS
 AYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI
 TGLYETRIDLSQLGGDKRPAATKKAGQAKKKKNQIRMKIGVMYGNPETTTTPKKKKR
 V (SEQ ID NO:19)

Residues 1-23: Flag tag

Residues 24-40: SV40 large T antigen NLS

Residues 41-1407: Cas9

Residues 1408-1423: Nucleoplasmin NLS

Residues 1424-1443: YECO peptide

Residues 1444-1450: SV40 large T antigen NLS

8. Cas9-3XYECO

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAAADKKYSIGLDIGTNSV
 GWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTR
 RKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP
 TIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQT
 YNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTP
 NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRV
 NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
 ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR
 QEDFYFPFLKDNREKIEKILTFRIPIYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDK
 GASAQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLS
 GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLK
 IIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG
 WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQ
 GDSLHEHIANLAGSPAIKKGILQTVKVVDLVKVMGRHKPENIVIEMARENQTTQKG
 QKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
 NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL
 LNAKLITQRKFDNLTKAERGGSELDAKAGFIKRQLVETRQITKHVAQILDSRMNTKYD
 ENDKLIREVKVITLKS KLVSDFRKDFQFYK VREINNYHHAHDAYLNAVVG TALIKKYP
 KLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR

PLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDK
 LIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSE
 KNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYV
 NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLS
 AYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI
 TGLYETRIDLSQLGGDKRPAATKKAGQAKKKKNQIRMKIGVMYGNPETTTGGSGGSS
 GGSSGSETPGTSESATPESSGGSSGGSNQIRMKIGVMYGNPETTTGGSGGSSGGSSGSE
 TPGTSESATPESSGGSSGGSNQIRMKIGVMYGNPETTTGGPKKKRKV (SEQ ID NO:20)

Residues 1-23: Flag tag

Residues 24-40: SV40 large T antigen NLS

Residues 41-1407: Cas9

Residues 1408-1423: Nucleoplasmin NLS

Residues 1424-1443: YECO peptide

Residues 1444-1475: GS linker

Residues 1476-1495: YECO peptide

Residues 1496-1527: GS linker

Residues 1528-1547: YECO peptide

Residues 1548-1554: SV40 NLS

9. Rad51(AE)-Cas9

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAAAMQMQLLEANADTSV
 EEESFGPQPISRLEQCGINANDVKKLEEAGFHTVEAVAYAPKKELINIKGISEAKADKIL
 AEA AKLVPMGFTTETEFHQRRSEIIQITTGSKELDKLLQGGIETGSITEMFGEFRTGKTQ
 ICHTLAVTCQLPIDRGGGEGKAMYIDTEGTFRPERLLAVAERYGLSGSDVLDNVAYA
 RAFNTDHQTQLLYQASAMMVESRYALLIVDSATALYRTDYSGRGELSARQMHLARF
 LRMLLR LADEFGVAVVITNQVVAQVDGAAMFAADPKKPIGGNIIAHASTTRLYLKRG
 RGETRICKIYDSPCLPEAEAMFAINADGVGDAKDSGGSSGGSSGSETPGTSESATPESS
 GGSSGGSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALL
 FDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFHRLEESFLVEED
 KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LIYLA LAHMIKFRGHF
 LIEGDLNPDNSDV DKLFIQLVQTYNQLFEE NPINASGVDAKAIL SARLSKSRRL ENLIA
 QLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLDNLLAQIGDQ

YADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
 EKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQ
 RTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRF
 AWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLEYEFT
 VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
 SVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL
 KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF
 MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDELVKVM
 GRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNE
 KLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGK
 SDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE
 TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVBREINNY
 HHAHDAYLNAVVGITALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYF
 FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIK
 KTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKGGK
 SKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIILPKYSLFELENGRKR
 MLASAGELQKGNELALPSKYVNFLYLASHYEKLGSPEDNEQKQLFVEQHKHYLDEI
 IEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDT
 TIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDKRPAATKKAGQAKKKK

(SEQ ID NO:21)

Residues 1-23: Flag tag

Residues 24-40: SV40 large T antigen NLS

Residues 41-378: Rad51(AE)

Residues 379-410: GS linker

Residues 411-1777: Cas 9

Residues 1778-1793: Nucleoplasmin NLS

10. Rad51(SEAD)-Cas9

MDYKDHDGDYKDHDIDYKDDDDKMAPKKRKVGIHGVPAAAMQMQLANADTSV
 EEESFGPQPISRLEQCGINANDVKKLEEAGFHTVEAVAYAPKKELINIKGISEAKADKIL
 AEA AKLVPMGFTTETEFHQRRSAIIQITTGSKELDKLLQGGIETGSITEMFGEFRTGKTQ
 ICHTLAVTCQLPIDRGGGEGKAMYIDTEGTFRPERLLAVAERYGLSGSDVLDNVAYA

RAFNTDHQTQLLYQAEDMMVESRYALLIVDSATALYRTDYSGRGELSARQMHLARF
LRMLLRLADEFQVAVVITNQVVAQVDGAAMFAADPKKPIGGNIIAHASTTRLYLRKG
RGETRICKIYDSPCLPEAEAMFAINADGVGDAKDSGGSSGGSSGSETPGTSESATPESS
GGSSGGSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALL
FDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEED
KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHMIKFRGHF
LIEGDLNPDNSDVDKLFIQLVQTYNQLFEEENPINASGVDAKAILSARLSKSRLENLIA
QLPGEKKNGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQ
YADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
EKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQ
RTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRF
AWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHSLLEYEFT
VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
SVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL
KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF
MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKGILQTVKVDELVKVM
GRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNE
KLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGK
SDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE
TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVRINNY
HHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYF
FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIK
KTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEVGK
SKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIILPKYSLFELENGRKR
MLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEI
IEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDT
TIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQLGGDKRPAATKKAGQAKKKK

(SEQ ID NO:22)

Residues 1-23: Flag tag

Residues 24-40: SV40 large T antigen NLS

Residues 41-378: Rad51(SEAD)

Residues 379-1777: Cas9

Residues 1778- 1793: Nucleoplasmin NLS

11. Cas9-Brex27

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAAADKKYSIGLDIGTNSV
 GWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTR
 RKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP
 TIYHLRKKLVSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQT
 YNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
 NFKSNFDLAEDAQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRV
 NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
 ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR
 QEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDK
 GASAQSFIERMTNFDKNLPNEKVLPHKSLLYEYFTVYNELTKVKYVTEGMRKPAFLS
 GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLK
 IIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG
 WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQ
 GDSLHEHIANLAGSPAIIKKGILQTVKVVDDELVKVMGRHKPENIVIEMARENQTTQKG
 QKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
 NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL
 LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYD
 ENDKLIREVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYP
 KLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR
 PLIETNGETGEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDK
 LIARKKDWDPKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSE
 KNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYV
 NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLS
 AYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQSI
 TGLYETRIDLSQLGGDKRPAATKKAGQAKKKKDYKDDDDKAAGALDFLSRLPLPPP
 SPICTFVSPAAQKAFQPPRSCG

(SEQ ID NO:23)

Residues 1-23: Flag tag

Residues 24-40: SV40 large T antigen NLS

Residues 41-1407: Cas9

Residues 1408-1423: Nucleoplasmin NLS

Residues 1424-1434: linker

Residues 1435-1470: Brex27

12. Cas9-Brex27-FECO

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAAADKKYSIGLDIGTNSV
 GWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTR
 RKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP
 TIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQT
 YNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
 NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRV
 NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
 ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR
 QEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDK
 GASAQSFIERMTNFDKNLPNEKVLPHKSLLYEYFTVYNELTKVKYVTEGMRKPAFLS
 GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLK
 IIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG
 WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQ
 GDSLHEHIANLAGSPAIIKKGILQTVKVVDDELVKVMGRHKPENIVIEMARENQTTQKG
 QKNSRERMKRIIEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
 NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL
 LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYD
 ENDKLIREVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYP
 KLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR
 PLIETNGETGEIVWDKGRDFATVRKVLSPQVNVKKTEVQTGGFSKESILPKRNSDK
 LIARKKDWDPKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSE
 KNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYV
 NFLYLASHYEKLGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLS
 AYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQSI
 TGLYETRIDLSQLGGDKRPAATKKAGQAKKKKEFGGGGSNQIRMKIGVMFGNPETTT

GGDYKDDDDKAAGALDFLSRLPLPPPVSPICTFVSPAAQKAFQPPRSCG (SEQ ID NO:24)

Residues 1-23: Flag tag

Residues 24-40: SV40 large T antigen NLS

Residues 41-1407: Cas9

Residues 1408-1423: Nucleoplasmin NLS

Residues 1424-1430: GS linker

Residues 1431-1450: FECO

Residues 1451-1461: linker

Residues 1462-1497: Brex

13. Cas9-DrRecA

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGHHGVPAADKKYSIGLDIGTNSV
GWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTR
RKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP
TIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQT
YNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRV
NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR
QEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDK
GASAQSFIERMTNFDKNLPNEKVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLS
GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLK
IHKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG
WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQ
GDSLHEHIANLAGSPAIIKKGILQTVKVVDDELVKVMGRHKPENIVIEMARENQTTQKG
QKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL
LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYD
ENDKLIREVKVITLKSCLVSDFRKDFQFYKVRINNYYHHAHDAYLNAVVG TALIKKYP
KLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR
PLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDK

LIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKGKSKKLKSVKELLGITIMERSSE
 KNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYV
 NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLS
 AYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQSI
 TGLYETRIDLSQLGGDKRPAATKKAGQAKKKKEFGGGGSGSHMSKDATKEISAPTD
 KERSKAJETAMSQIEKAFGKG SIMKLGAESKLDVQVVSTGSLSLDLALGVGGIPRGRIT
 EIYGPESGGKTTLALAIVAQAQKAGGTCAFIDAEHALDPVYARALGVNTDELLVSQPD
 NGEQALEIMELLVRSGAIDVVVVDSVAALTPRAEIEGDMGDSLPLGLQARLMSQALRK
 LTAILSKTGTAAIFINQVREKIGVMYGNPETTTGGRALKFYASVRLDVRKIGQPTKVG
 NDAVANTVKIKTVKNKVAAPFKEVELALVYGKGFQDLSDLVGLAADMDIHKAGSF
 YSYGDERIGQGKEKTIAIYAERPHEMEQEIRDRVMAAIRAGNAGEAPALAPAPAAPEAA
 EAPKKKRKV (SEQ ID NO:25)

Residues 1-23: Flag tag

Residues 24-40: SV40 large T antigen NLS

Residues 41-1407: Cas9

Residues 1408-1423: Nucleoplasmin NLS

Residues 1424-1430: GS linker

Residues 1431-1796: DrRecA

Residues 1797- 1803: SV40 large T antigen NLS

14. Cas9-DrRecA 20AA

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAADKKYSIGLDIGTNSV
 GWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTR
 RKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP
 TIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQT
 YNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLFGNLIALSLGLTP
 NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRV
 NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
 ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR
 QEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDK
 GASAQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLS
 GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLK

IIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG
 WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQ
 GDSLHEHIANLAGSPAIIKKGILQTVKVVDDELVKVMGRHKPENIVIEMARENQTTQKG
 QKNSRERMKRIIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
 NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL
 LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYD
 ENDKLIREVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYP
 KLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR
 PLIETNGETGEIVWDKGRDFATVRKVLSPQVNVKKTEVQTGGFSKESILPKRNSDK
 LIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKGKSKKLKSVKELLGITIMERSSE
 KNPIDFLEAKGYKEVKKDLIIKLPKYSLENGRKRMLASAGELQKGNELALPSKYV
 NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLS
 AYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLHQS
 TGLYETRIDLSQLGGDKRPAATKKAGQAKKKKEFGGGGSNQVREKIGVMYGNPETT
 TGGPKKKRKV (SEQ ID NO:26)

Residues 1-23: Flag tag

Residues 24-40: SV40 large T antigen NLS

Residues 41-1407: Cas9

Residues 1408-1423: Nucleoplasmin NLS

Residues 1424-1430: GS linker

Residues 1431-1450: DrRecA 20AA

Residues 1451- 1457: SV40 large T antigen NLS

15. Cas9-RecT

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAAADKKYSIGLDIGTNSV
 GWAVITDEYKVPSSKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTR
 RKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP
 TIYHLRKKLVSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQT
 YNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTP
 NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILSDILRV
 NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
 ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR

QEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEFVVDK
 GASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLS
 GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLK
 IIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG
 WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQ
 GDSLHEHIANLAGSPAIAKKGILQTVKVDELVKVMGRHKPENIVIAMARENQTTQKG
 QKNSRERMKRIEEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
 NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL
 LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYD
 ENDKLIREVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYP
 KLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR
 PLIETNGETGEIVWDKGRDFATVRKVLSPQVNVKKTEVQTGGFSKESILPKRNSDK
 LIARKKDWDPKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSE
 KNPIDFLEAKGYKEVKKDLIIKLPKYSLENGRKRMLASAGELQKGNELALPSKYV
 NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLS
 AYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQSI
 TGLYETRIDLSQLGGDKRPAATKKAGQAKKKKEFGGGGSMKQPPIAKADLQKTQG
 NRAPAAVKNSDVISFINQPSMKEQLAAALPRHMTAERMIRIATTEIRKVPALGNCDTM
 SFVSAIVQCSQLGLEPGSALGHAYLLPFGNKNEKSGKKNVQLIIGYRGMIDLARRSGQI
 ASLSARVVREGDEFSEFGLDEKLIHRPGENEDAPVTHVYAVARLKDGGTQFEVMTR
 KQIELVRSLSKAGNNGPWVTHWEEMAKKTAIRRLFKYLPVSIEIQRAVSMDEKEPLTI
 DPADSSVLTGEYSVIDNSEEPKKKRKV (SEQ ID NO:27)

Residues 1-23: Flag tag

Residues 24-40: SV40 large T antigen NLS

Residues 41-1407: Cas9

Residues 1408-1423: Nucleoplasmin NLS

Residues 1424-1430: GS linker

Residues 1431-1699: RecT

Residues 1670-1676: SV40 large T antigen NLS

16. Cas9-Lambda Beta

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAAADKKYSIGLDIGTNSV
GWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTR
RKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP
TIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQT
YNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRV
NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR
QEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDK
GASAQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLS
GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLK
IHKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG
WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQ
GDSLHEHIANLAGSPAIIKKGILQTVKVVDDELVKVMGRHKPENIVIEMARENQTTQKG
QKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL
LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYD
ENDKLIREVKVITLKSCLVSDFRKDFQFYK VREINNYHHAHDAYLNAVVG TALIKKYP
KLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR
PLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDK
LIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSE
KNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYV
NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLS
AYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI
TGLYETRIDLSQLGGDKRPAATKKAGQAKKKKEFGGGGSMSTALATLAGKLAERVG
MDSVDPQELITT LRQTAFKGDASDAQFIALLIVANQYGLNPWTKEIYAFDPKQNGIVP
VVGVDGWSRIINENQQFDGMDFEQDNESCTCRIYRKDRNHPICVTEWMDECRREPFK
TREGREITGPWQSHPKRMLRHKAMIQCARLAFGFAGIYDKDEAERIVENTAYTAERQ
PERDITPVNDETMQEINTLLIALDKTWDDDLLPLCSQIFRRDIRASSETQAEAVKALG
FLKQKAAEQKVAAPKKKRKV (SEQ ID NO:28)

Residues 1-23: Flag tag

Residues 24-40: SV40 large T antigen NLS

Residues 41-1407: Cas9

Residues 1408-1423: Nucleoplasmin NLS

Residues 1424-1430: GS linker

Residues 1431-1691: Lambda Beta

Residues 1692-1698: SV40 large T antigen NLS

17. Cas9-SSBP

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAAADKKYSIGLDIGTNSV
 GWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTR
 RKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP
 TIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQT
 YNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTP
 NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILSDILRV
 NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG
 ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR
 QEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVDK
 GASAQSFIERMTNFDKNLPNEKVLPKHSLLEYEFTVYNELTKVKYVTEGMRKPAFLS
 GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLK
 IIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG
 WGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFQMQLIHDDSLTFKEDIQKAQVSGQ
 GDSLHEHIANLAGSPAIKKGILQTVKVVDDELVKVMGRHKPENIVIAMARENQTTQKG
 QKNSRERMKRIEEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
 NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL
 LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYD
 ENDKLIREVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYP
 KLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR
 PLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDK
 LIARKKDWDPKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEFE
 KNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYV
 NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLS
 AYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI

TGLYETRIDLSQLGGDKRPAATKKAGQAKKKKEFGGGGSMASRGVNKVILVGNLGQ
DPEVRYMPNGGAVANITLATSESWRDKATGEMKEQTEWHRVVLFGKLAEVASEYLR
KGSQVYIEGQLRTRKWTQSGQDRYTTEVVVNVGGMQMLGGRQGGGAPAGGNIG
GGQPQGGWGQPQQPQGGNQFSGGAQSRPQQSAPAAPSNPPMDFDDEIPFPKKKRK
V (SEQ ID NO:29)

Residues 1-23: Flag tag

Residues 24-40: SV40 large T antigen NLS

Residues 41-1407: Cas9

Residues 1408-1423: Nucleoplasmin NLS

Residues 1424-1430: GS linker

Residues 1431-1608: SSBP

Residues 1609-1615: SV40 large T antigen NLS

18. Cas9-GS-FECO

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAADKKYSIGLDIGTNSV
GWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTR
RKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYP
TIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQT
YNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTP
NFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRV
NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPKEYKEIFFDQSKNGYAGYIDGG
ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR
QEDFYFPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEDEVVDK
GASAQSFIERMTNFDKNLPNEKVLPKHSLLEYEFTVYNELTKVKYVTEGMRKPAFLS
GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLK
IHKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG
WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQ
GDSLHEHIANLAGSPAIKKGILQTVKVVDDELVKVMGRHKPENIVIAMARENQTTQKG
QKNSRERMKRIEELGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI
NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL
LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYD
ENDKLIREVKVITLKSCLVSDFRKDFQFYK VREINNYHHAHDAYLNAVVGTA LIKKYP

KLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKR
 PLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDK
 LIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSE
 KNPIDFLEAKGYKEVKKDLIIKLPKYSLENGRKRMLASAGELQKGNELALPSKYV
 NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIIEQISEFSKRVILADANLDKVLS
 AYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQSI
 TGLYETRIDLSQLGGDKRPAATKKAGQAKKKKEFGGGGSNQIRMKIGVMFGNPETTT
 GGPKKKRKV (SEQ ID NO:30)

Residues 1-23: Flag tag

Residues 24-40: SV40 large T antigen NLS

Residues 41-1407: Cas9

Residues 1408-1423: Nucleoplasmin NLS

Residues 1424-1430: GS linker

Residues 1431-1450: FECO

Residues 1451-1457: SV40 NLS

mRNA purification

[00193] mRNA in this study was produced through in vitro transcription from plasmid templates. First plasmids containing T7 promoter, 5' UTR, Cas9 fusion sequence ORF 3' UTR, bovine growth hormone polyadenylation signal (bGH polyA), and 64 poly adenine sequence were cloned. To linearize the DNA, restriction sites were inserted at the end of the 64 poly adenines. After plasmids were produced, ~50µg of plasmids were digested using a restriction enzyme to cut the plasmids at the added restriction site for linearization. After 24 hours, the resulting linear DNA was cleaned through a DNA clean-up kit from zymo (Cat# D4029). After DNA cleanup, the linear DNA fragments were introduced to HISCRIBETM T7 mRNA Kit with CLEANCAP® Reagent AG (Cat# E2080S) from New England Biolabs. This kit used AG CLEANCAP® Reagent from TriLink Biotechnologies and yielded ~ 100µg per 1µg of linear DNA templates. For each reaction, 1µg of templates were mixed with 2µl 10x reaction buffer, 2µl of ATP (60mM), 2µl of GTP (50mM), 2µl of UTP (50mM), 2µl of CTP (50mM), 2µl of Cap Analog (40mM) and 2µl of RNA polymerase Mix. Then Nuclease-free water was added to the final volume of 20µl to the mixtures. The mixture was then incubated at 37C° for 2 hours. After

2 hours of *in vitro* transcription, the resulting mRNA was purified using RNA Clean & concentrator from Zymo (Cat # R1019). DNase was also used during the purification to remove residual DNA templates from the solutions. After *in vitro* transcriptions, each mRNA was analyzed using agarose gel analysis. 500ng of mRNA were diluted to 20ul of nuclease-free water then 40ul of 2 x RNA loading dye from ThermoFisher Scientific (Cat# R0641) was added. The mixture was then incubated at 80C° for 15 minutes. After incubation, the mRNA mixtures were immediately transferred to Ice and ran on the clean 1% agarose gel made from TAE buffer.

HEK293 lipofectamine transfection

[00194] For lipofectamine transfection of mRNA editors/sgRNA (mRNA cocktail) and DNA donors, 24 well plates are coated with PLO for 2hrs. Plates were washed 2X and dried before they were plated with cells. For each well 2×10^5 cells were plated on Day 0 using 500 ul of complete DMEM media. On Day 1, 250ul of media was slowly replaced with equal volume of serum free and antibiotic free DMEM media. Both mRNA cocktail and DNA were prepared separately as per the conditions. To prepare the mRNA cocktail, individual mRNA construct (1ug/well) and *RAB11A* sgRNA (2.54ul/well from a stock of 80uM) were diluted in 1:1 ratio with lipofectamine 3000 and incubated for 10-15 minutes; for DNA prep, both single stranded and double stranded DNAs were packaged separately. For both, concentration of 1ug/well was used. Respective DNA was mixed with P3000 reagent (1ug/well) and the whole mix was diluted with lipofectamine 3000 in 1:1 dilution. The DNA cocktail was also incubated at RT for 10-15 mins. After 15 mins of incubation, mRNA cocktail and DNA cocktail were added in 1:1 ratio as per the plate map. Cells were then allowed to sit for 48 hours after which they were transferred from 24 well plate format to 6 well plate with 1ml media. FLO analysis was performed on D5, D9 and D14 to check on the knock-in efficiency of the constructs.

Generation of circular single-stranded DNA

[00195] Donor template sequences (transgene sequence flanked with 5' and 3' homology arms at 300-500 nt in length) are constructed as dsDNA and placed into phagemid vector by Golden Gate Assembly. An XL1-Blue *E. Coli* Strain was co-transformed with the M13 helper plasmid and phagemid containing donor template sequences and selected in agar plate with kanamycin (50 µg/mL), carbenicillin (100 µg/mL). Single colony was selected and grown for ~24 hours (37

°C, 225 rpm) in 250 mL 2xYT media (1.6% tryptone, 1% yeast extract, 0.25% NaCl) to reach OD₆₀₀ between 2.5-3.0. The bacteria were pelleted by centrifugation and the phage particles in the supernatant were precipitated by PEG-8000 by centrifugation, washed and lysed in 20 mM MOPS, 1M Guanidine-HCl and 2% Triton X-100. The cssDNA released from phage were then extracted with NucleoBond Xtra Midi EF kit (Macherey-Nagel) following the manufacturer's instructions. The concentration of cssDNA is determined by Nanodrop. Ratios of Absorbance (A_{260 nm}/A_{280 nm} and A_{260 nm}/A_{230 nm}) will reflect consistent purity (1.8 and > 2, respectively) from serial preps. Recombinant cssDNA is verified by DNA sequencing using custom-designed staggered sequencing primers for complete coverage.

Cell culture

[00196] K562 cells (ATCC, CCL-243) were maintained in RPMI-1640 media with 10% FBS and 1% penicillin and streptomycin. HEK293T (ATCC) cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum and 1% penicillin and streptomycin (Gibco). iPSC (Thermo Fisher) was cultured in StemFlex (Thermo Fisher) media in vitronectin-coated flask. iPSC colonies were checked regularly and passaged using ReLeSR (StemCell Technologies) every 3-4 days of culture. iPSCs were ready for electroporation after 2-3 passages. T cells were cultured and expanded in TexMACS Medium (Miltenyi) supplemented with 200 IU/mL Human IL-2 IS (Miltenyi). T cells were activated for 2 days with T Cell TransAct (Miltenyi) before electroporation. All cells were maintained in a humidified incubator at 37 °C and 5 % CO₂, unless otherwise specified. Cell count viability were determined using Via2-Cassette in NUCLEOCOUNTER® NC-202 (ChemoMetec) on specified days after engineering.

Electroporation of Cas9 RNP, AIO plasmid and mRNA/sgRNA complex with DNA donor

[00197] All K562, HEK293T, iPSC and T cell electroporation were performed using the Amaxa™ 96-well Shuttle™ with the 4D Nucleofector (Lonza). Cas9 nucleases and sgRNAs were precomplexed in supplemented NUCLEOFECTOR® Solution for 20 min at room temperature and the RNP solution was made up to a final volume of 2.5 µL (10X). For electroporating K562 cells, SF Cell Line 4D-Nucleofector™ Kit and 250,000-500,000 cells per reaction were used with program FF-120. For electroporating HEK293T cells, SF Cell Line 4D-

NUCLEOFECTOR™ Kit and 200,000-300,000 cells per reaction were used with program FS-100. For electroporating iPSC cells, 100,000 cells per reaction were used with P3 Primary Cell 4D-Nucleofector™ Kit and program CA-137. For electroporating primary T cells, 2x10⁶ cells per reaction were used with P3 Primary Cell 4D-NUCLEOFECTOR™ Kit and program EO-115. Indicated amount of HDR donor template (cssDNA or dsDNA) were co-electroporation with RNP. For mRNA or AIO plasmid enGager electroporation, 1 ug-1.5ug of mRNA or DNA were electroporated together with indicated amount of HDR donor templates, with the same electroporation parameters used for RNP electroporation. After electroporation, cells were incubated in humidified 32°C incubator with 5 % CO₂ for 12-24 hours followed up transferring to 37°C incubator for additional days.

An example of electroporation for gene KI in K562 cells

Reagents:

1. SF 4D-Nucleofector SF cell line solution and Supplement (Lonza, 4D-NUCLEOFECTOR™ X kit)
2. Add the supplement to the Nucleofector solution. The ratio of Nucleofector solution to supplement is 4.5: 1
3. Cas9 (10µg/ul stock, Aldevron)
4. *RAB11A* sgRNA (80µM)
5. Trypan Blue Solution (0.4%) (Sigma-Aldrich #T8154)
6. Different Cas9-All in one plasmid
 - Control Cas9-AIO (plasmid/mRNA conc-1.8µg/ul)
 - Cas9 Rad51 SEAD (2µg/ul)
 - Cas9 Rad51 AE (2µg/ul)
 - Cas9 RecA (2µg/ul)
 - Cas9 Brex (1.8µg/ul)
7. cssDNA stock css76 (RAB11-GFP 2Kb payload) and css116 (RAB11-GFP 4kb payload). The stock concentration is >2 ug/ul.
8. K562 cells, MEM with 10% FBS media

Procedure:

1. Prepare electroporation SF buffer

2. Prepare 10X RNP mix using Cas9, *RAB11A* sgRNA and SF buffer with supplement. Incubate RNPs for 20mins @ RT followed by ice.
3. To prepare the cells for electroporation, resuspend cells in growth media and count viable cells.
4. Electroporation is done in triplicate condition.
5. Calculate the number of cells needed per condition of electroporation (0.5×10^6 cells/well) and pellet required number of cells at 300g for 5 mins.
6. Wash cells with OPTI-MEM and centrifuge again.
7. Resuspend cells in required amount of SF buffer.
8. cssDNA should be diluted into desired concentrations with ddH₂O.
9. Prepare sample in PCR strips.
10. Add 25µl of the sample prep into designated well of the cuvette strips in 96 well Lonza EP plate.
11. Gently tap the Nucleocuvette™ Vessels to make sure the sample covers the bottom of the cuvette. For K562 cells, FF-120 program is used for electroporation.
12. After electroporation, add 160µl of prewarmed K562 culture media into the designated cuvette strips (total 185µl)
13. Mix cells by gently pipetting up and down 2-3 times; incubate at RT for 10 minutes.
14. Transfer 40µl of the same into 96 well round bottom plate prefilled with 160µl of K562 culture media. Outer wells should have water jacket to prevent electroporation.
15. Incubate the cells in humidified 32°C incubator with 5 % CO₂ for 12-24 hours after which transfer to regular 37°C incubator.
16. After 48hrs change media and keep it back again @37°C
17. After 72hrs of electroporation keep a fraction of cell for propagation for Day7 FLO analysis and the rest use for Day3 FLO analysis.
 - a) For FLO: Prepare cells for FLO analysis. KI efficiency is checked by calculating the increase in GFP+ve cells over the control condition.
 - b) To propagate cells for Day 7, take 40µl of cells solution and add to 160µl of respective media in the fresh 96 well round bottom plate. Outer wells should have water jacket to prevent evaporation.

18. On Day6, again change the media of the cells and follow the same protocol for FLO on Day7.

19. Data analysis is performed using FlowJo_v10.8.0_CL software.

Flow Cytometry analysis

[00198] All flow cytometry was performed on an Attune NxT flow cytometer with a 96-well autosampler (ThermoFisher Scientific). Unless otherwise indicated, cells were collected 3-14 days post electroporation, resuspended in fluorescence-activated cell sorting (FACS) buffer (2% BSA in PBS) and stained with 7-AAD (BioLegend) as a dye for cell viability assay, and the indicated cell-surface marker, such as PE labelled Protein L (ACROBiosystems), or FITC-Labeled Human CD19 (20-291) Protein, Fc Tag (ACROBiosystems). To obtain comparable live cell counts between conditions, events were recorded from an equivalent fixed volume for all samples. Data analysis was performed using FlowJo_v10.8.0_CL software with exclusion of subcellular debris, singlet gating and live/dead staining. Data analysis is performed using FlowJo_v10.8.0_CL software. Data were plotted using Prism Graphpad.

CAR-T cell engineering

[00199] Primary T cells were isolated and enriched from Leukopak using Pan T Cell MicroBead Cocktail with MultiMCAS Cell24 Separator Plus (Miltenyi), CD4⁺ and CD8⁺ pan T cells were cryopreserved for later use. T cells were cultured and expanded in TexMACS Medium (Miltenyi) supplemented with 200 IU/mL Human IL-2 IS (Miltenyi). T cells were activated for 2 days with T Cell TransAct (Miltenyi) before electroporation. 48 h after initiating T-cell initiation and activation, T cells were electroporated using Amaxa™ 96-well Shuttle™ in 4D Nucleofector. 2×10⁶ cells were mixed with 25 pmol of Cas9 WT protein or 1 ug of enGager mRNA and 50 pmol of gRNA (RNP) into each well. For cssDNA engineered cells, 2 µg of cssDNA encoding bi-specific CD19×CD22 CAR was electroporated with RNP or mRNA/sgRNA cocktail targeting *TRAC* locus. Following electroporation, cells were diluted into culture medium in the presence of T Cell TransAct with 1 µM DNA-PK inhibitor M-3814, and incubated at 32 °C, 5% CO₂ for 24 hours. Cells were then washed and subsequently transferred into G-Rex 24 Multi-Well Cell Culture Plate (Wilson Wolf) in standard culture conditions at 37 °C, 5% CO₂ in IL-2 supplemented TexMACS medium and replenished every 3-4 days. CAR

expressions were determined using PE labelled Protein L (ACROBiosystems), or FITC-Labeled Human CD19 (20-291) Protein, Fc Tag (ACROBiosystems).

Enumerated Embodiments

[00200] The following enumerated embodiments are provided, the numbering of which is not to be construed as designating levels of importance:

[00201] Embodiment 1. A DNA editing construct, comprising:
a DNA editing motif having a site-specific DNA endonuclease activity;
a nuclear localization signal (NLS); and
a single stranded DNA (ssDNA) binding motif.

[00202] Embodiment 2. The construct of Embodiment 1, wherein the DNA editing motif introduces a double strand break, a single strand break, or no double strand break in double stranded DNA molecules.

[00203] Embodiment 3. The construct of Embodiment 2, wherein the DNA editing motif introduces the double-strand break in double stranded DNA molecules, and wherein the double-strand break is a blunt-ended break or a sticky-ended break.

[00204] Embodiment 4. The construct of any one of Embodiments 1-3, wherein the DNA editing motif is at least one site-specific nuclease selected from the group consisting of a Cas nuclease, a zinc-finger nuclease, a meganuclease, and a transcription activator-like effector nuclease (TALEN).

[00205] Embodiment 5. The construct of any one of Embodiments 1-4, wherein the DNA editing motif is a class I Cas protein or a class II Cas protein.

[00206] Embodiment 6. The construct of any one of Embodiments 1-5, wherein the DNA editing motif is Cas9, Cas3, Cas10, Cas11, CasX, Cas12a, MAD7, or catalytic dead forms or nickase forms thereof.

[00207] Embodiment 7. The construct of any one of Embodiments 1-6, wherein the nuclear localization signal is at least one selected from the group consisting of a 53BP1 NLS, a CXCR4 NLS, an ERK5 NLS, an EWS NLS, an IER5 NLS, an ING4 NLS, an Hrp1 NLS, an MSX1 NLS, an NLS-RAR α NLS, a nucleoplasmin NLS, a Pho4 NLS, a PTHrP NLS, an rpl23a NLS, a simian virus 40 (SV40) large T antigen NLS, a UL79 NLS, a VACM-1/CUL5 NLS, a VP1 NLS,

a PABPN1 NLS, a STAT1 NLS, a FGF2 NLS, a RARa NLS, a tandem phenylalanine-glycine repeat (FG repeat) from a nuclear pore complex, a histone H2a NLS, and a histone H2b NLS.

[00208] Embodiment 8. The construct of any one of Embodiments 1-7, wherein the construct comprises a plurality of the nuclear localization signals.

[00209] Embodiment 9. The construct of any one of Embodiments 1-8, wherein at least one of the following applies:

- (a) the ssDNA binding motif is or comprises a K homology (KH) domain ssDNA binding motif, an oligonucleotide/oligosaccharide binding (OB) fold ssDNA binding motif, a RecA/Rad51/DMC1 loop L2 homologous ssDNA binding motif, an RNA recognition motif (RRM) that binds ssDNA, or a whirly domain ssDNA binding motif,
- (b) the ssDNA binding motif is or comprises a bacterial single-stranded DNA-binding (SSB) protein or ssDNA binding portions thereof, a RecA/Rad51/DMC family protein or ssDNA binding portions thereof, a replication protein A (RPA) or ssDNA binding portions thereof, or a single-stranded annealing protein (SSAP) or ssDNA binding portions thereof,
- (c) the ssDNA binding motif is or comprises an ssDNA binding motif derived from a RecA/Rad51/DMC family protein, and wherein the ssDNA binding motif does not have recombinase activity.

[00210] Embodiment 10. The construct of any one of Embodiments 1-9, wherein the ssDNA binding motif is or comprises a polypeptide having the amino acid sequence set forth in SEQ ID NO:11.

[00211] Embodiment 11. The construct of any one of Embodiments 1-10, wherein the construct comprises a plurality of the ssDNA binding motifs.

[00212] Embodiment 12. The construct of any one of Embodiments 1-11, further comprising one or more linkers connecting the DNA editing motif, the nuclear localization signal, and the ssDNA binding motif to each other.

[00213] Embodiment 13. The construct of Embodiment 12, wherein the one or more linkers connects the DNA editing motif, the nuclear localization signal, and the ssDNA binding motif to each other covalently.

[00214] Embodiment 14. The construct of any one of Embodiments 12-13, wherein the one or more linkers independently are polypeptide linkers.

[00215] Embodiment 15. The construct of Embodiment 14, wherein each of the polypeptide linkers independently has a length ranging from about 2 amino acid residues to about 100 amino acid residues.

[00216] Embodiment 16. The construct of any one of Embodiments 1-15, wherein the construct is a protein.

[00217] Embodiment 17. A nucleic acid encoding the construct of Embodiment 16.

[00218] Embodiment 18. The nucleic acid of Embodiment 17, further encodes a guide RNA (gRNA) that provides the site-specificity to the DNA editing motif.

[00219] Embodiment 19. The nucleic acid of any one of Embodiments 17-18, wherein the nucleic acid is comprised in an expression vector.

[00220] Embodiment 20. The nucleic acid of Embodiment 19, wherein the expression vector is a plasmid vector, an RNA vector, or a viral vector.

[00221] Embodiment 21. The nucleic acid of any one of Embodiments 17-20, wherein the nucleic acid is packaged in a viral particle, a lipid nanoparticle, or a non-lipid nanoparticle.

[00222] 22. A composition for editing DNA, comprising:

the construct of any one of Embodiments 1-16; and

a single stranded DNA (ssDNA) having a DNA insert, a 5'-homology arm, and a 3'-homology arm, wherein the 5' homology arm and the 3' homology arm are complementary to the nuclear DNA in a target region.

[00223] Embodiment 23. The composition of Embodiment 22, wherein least one of the following applies:

(a) the ssDNA is a circular ssDNA (cssDNA), a linear ssDNA (lssDNA), a double stranded DNA/circular single stranded DNA (dsDNA/cssDNA) hybrid having an ssDNA section, or an RNA/DNA hybrid having an ssDNA section;

(b) the ssDNA is chemically modified;

(c) the ssDNA is produced from a cell free system or a cell culture system.

[00224] Embodiment 24. The composition of any one of Embodiments 22-23, wherein each of the 5'-homology arm and the 3'-homology arm independently has a length ranging from about 25 nucleotides to about 5000 nucleotides.

[00225] Embodiment 25. The composition of any one of Embodiments 22-24, wherein the DNA insert has a length ranging from 1 nucleotide to about 50,000 nucleotides.

[00226] Embodiment 26. A kit for editing DNA, comprising:
the construct of any one of Embodiments 1-16 and/or the nucleic acid of any one of Embodiments 17-21; and
a single stranded DNA (ssDNA), a nucleic acid that encodes the ssDNA, or a nucleic acid for constructing the ssDNA,
wherein the ssDNA comprises a DNA insert, a 5'-homology arm, and a 3'-homology arm,
wherein the 5' homology arm and the 3' homology arm are complementary to the nuclear DNA in a target region.

[00227] Embodiment 27. The kit of Embodiment 26, wherein least one of the following applies:

- (a) the ssDNA is a circular ssDNA (cssDNA), a linear ssDNA (lssDNA), a double stranded DNA/circular single stranded DNA (dsDNA/cssDNA) hybrid having an ssDNA section, or an RNA/DNA hybrid having an ssDNA section;
- (b) the ssDNA is chemically modified;
- (c) the ssDNA is produced from a cell free system or a cell culture system.

[00228] Embodiment 28. The kit of any one of Embodiments 26-27, wherein each of the 5'-homology arm and the 3'-homology arm independently has a length ranging from about 25 nucleotides to about 5000 nucleotides.

[00229] Embodiment 29. The kit of any one of Embodiments 26-28, wherein the DNA insert has a length ranging from 1 nucleotide to about 50,000 nucleotides.

[00230] Embodiment 30. The kit of any one of Embodiments 26-29, wherein at least one of the following applies:

- (a) the nucleic acid for constructing the ssDNA donor template encodes every elements of the ssDNA donor template, except for the DNA insert,
- (b) the nucleic acid for constructing ssDNA donor template encodes every elements of the ssDNA donor template, except for the 5' homology arm and the 3' homology arm,
- (c) the nucleic acid for constructing ssDNA donor template does not encode the 5' homology arm, the 3' homology arm, and the DNA insert.

[00231] Embodiment 31. The kit of Embodiment 30, wherein the nucleic acid includes a multiple cloning site for inserting the 5' homology arm, the 3' homology arm, and/or the DNA insert.

[00232] Embodiment 32. The kit of any one of Embodiments 30-31, wherein the nucleic acid for constructing ssDNA donor template is a plasmid or a viral vector.

[00233] Embodiment 33. The kit of any one of Embodiment 27-32, further comprising a component for delivering the construct, the nucleic acid for encoding the construct, the ssDNA, and/or the nucleic acid for encoding or constructing the ssDNA into a cell.

[00234] Embodiment 34. A method of editing a nuclear DNA in a cell, the method comprising delivering into the cell the construct of any one of Embodiments 1-16 and/or the nucleic acid of any one of Embodiments 17-21; and a single stranded DNA (ssDNA) having a DNA insert, a 5'-homology arm, and a 3'-homology arm, wherein the 5' homology arm and the 3' homology arm are complementary to the nuclear DNA in a target region.

[00235] Embodiment 35. The method of Embodiment 34, wherein at least one of the following applies:

- (a) the ssDNA is a circular ssDNA (cssDNA), a linear ssDNA (lssDNA), a double stranded DNA/circular single stranded DNA (dsDNA/cssDNA) hybrid having an ssDNA section, or an RNA/DNA hybrid having an ssDNA section;
- (b) the ssDNA is chemically modified;
- (c) the ssDNA is produced from a cell free system or a cell culture system.

[00236] Embodiment 36. The method of any one of Embodiments 34-35, wherein each of the 5'-homology arm and the 3'-homology arm independently has a length ranging from about 25 nucleotides to about 5000 nucleotides.

[00237] Embodiment 37. The method of any one of Embodiments 34-36, wherein the DNA insert has a length ranging from 1 nucleotide to about 50,000 nucleotides.

[00238] Embodiment 38. The method of any one of Embodiments 34-37, wherein the construct and/or nucleic acid and the ssDNA are delivered into the cell by a viral vector, a lipid or non-lipid nanoparticle delivery, an electroporation delivery, a gene gun delivery, or an injection.

[00239] Embodiment 39. The method of any one of Embodiments 34-38, wherein the nuclear DNA is a genomic DNA.

[00240] Embodiment 40. The method of any one of Embodiments 34-39, wherein the cell is a prokaryotic cell, a cultured plant cell, a eukaryotic cell, a cultured mammalian cell, a primary

cell, an embryonic stem cell, an adult stem cell, a hematopoietic stem cell, an induced pluripotent stem cell, or a CAR-T cell.

[00241] Embodiment 41. The method of any one of Embodiments 34-40, wherein the cell is in the body of a subject.

[00242] Embodiment 42. The method of Embodiment 41, wherein the subject is a mammal, such as a human.

[00243] Embodiment 43. A non-natural polypeptide comprising an amino acid sequence having 90% or higher identity to the amino acid sequence set forth in SEQ ID NO:11.

[00244] Embodiment 44. The non-natural polypeptide of Embodiment 43, wherein the non-natural polypeptide is a ssDNA-binding polypeptide.

[00245] The foregoing outlines features of several embodiments so that those skilled in the art may better understand the aspects of the present disclosure. Those skilled in the art should appreciate that they may readily use the present disclosure as a basis for designing or modifying other processes and structures for carrying out the same purposes and/or achieving the same advantages of the embodiments introduced herein. Those skilled in the art should also realize that such equivalent constructions do not depart from the spirit and scope of the present disclosure, and that they may make various changes, substitutions, and alterations herein without departing from the spirit and scope of the present disclosure.

CLAIMS

What is claimed is:

1. A DNA editing construct, comprising:
a DNA editing motif having a site-specific DNA endonuclease activity;
a nuclear localization signal (NLS); and
a single stranded DNA (ssDNA) binding motif.
2. The construct of claim 1, wherein the DNA editing motif introduces a double strand break, a single strand break, or no double strand break in double stranded DNA molecules.
3. The construct of claim 2, wherein the DNA editing motif introduces the double-strand break in double stranded DNA molecules, and wherein the double-strand break is a blunt-ended break or a sticky-ended break.
4. The construct of any one of claims 1-3, wherein the DNA editing motif is at least one site-specific nuclease selected from the group consisting of a Cas nuclease, a zinc-finger nuclease, a meganuclease, and a transcription activator-like effector nuclease (TALEN).
5. The construct of any one of claims 1-4, wherein the DNA editing motif is a class I Cas protein or a class II Cas protein.
6. The construct of any one of claims 1-5, wherein the DNA editing motif is Cas9, Cas3, Cas10, Cas11, CasX, Cas12a, MAD7, or catalytic dead forms or nickase forms thereof.
7. The construct of any one of claims 1-6, wherein the nuclear localization signal is at least one selected from the group consisting of a 53BP1 NLS, a CXCR4 NLS, an ERK5 NLS, an EWS NLS, an IER5 NLS, an ING4 NLS, an Hrp1 NLS, an MSX1 NLS, an NLS-RAR α NLS, a nucleoplasmin NLS, a Pho4 NLS, a PTHrP NLS, an rpL23a NLS, a simian virus 40 (SV40) large T antigen NLS, a UL79 NLS, a VACM-1/CUL5 NLS, a VP1 NLS, a PABPN1 NLS, a STAT1 NLS, a FGF2 NLS, a RAR α NLS, a tandem phenylalanine-glycine repeat (FG repeat) from a nuclear pore complex, a histone H2a NLS, and a histone H2b NLS.

8. The construct of any one of claims 1-7, wherein the construct comprises a plurality of the nuclear localization signals.
9. The construct of any one of claims 1-8, wherein at least one of the following applies:
- (a) the ssDNA binding motif is or comprises a K homology (KH) domain ssDNA binding motif, an oligonucleotide/oligosaccharide binding (OB) fold ssDNA binding motif, a RecA/Rad51/DMC1 loop L2 homologous ssDNA binding motif, an RNA recognition motif (RRM) that binds ssDNA, or a whirlly domain ssDNA binding motif,
 - (b) the ssDNA binding motif is or comprises a bacterial single-stranded DNA-binding (SSB) protein or ssDNA binding portions thereof, a RecA/Rad51/DMC family protein or ssDNA binding portions thereof, a replication protein A (RPA) or ssDNA binding portions thereof, or a single-stranded annealing protein (SSAP) or ssDNA binding portions thereof,
 - (c) the ssDNA binding motif is or comprises an ssDNA binding motif derived from a RecA/Rad51/DMC family protein, and wherein the ssDNA binding motif does not have recombinase activity.
10. The construct of any one of claims 1-9, wherein the ssDNA binding motif is or comprises a polypeptide having the amino acid sequence set forth in SEQ ID NO:11.
11. The construct of any one of claims 1-10, wherein the construct comprises a plurality of the ssDNA binding motifs.
12. The construct of any one of claims 1-11, further comprising one or more linkers connecting the DNA editing motif, the nuclear localization signal, and the ssDNA binding motif to each other.
13. The construct of claim 12, wherein the one or more linkers connects the DNA editing motif, the nuclear localization signal, and the ssDNA binding motif to each other covalently.

14. The construct of any one of claims 12-13, wherein the one or more linkers independently are polypeptide linkers.
15. The construct of claim 14, wherein each of the polypeptide linkers independently has a length ranging from about 2 amino acid residues to about 100 amino acid residues.
16. The construct of any one of claims 1-15, wherein the construct is a protein.
17. A nucleic acid encoding the construct of claim 16.
18. The nucleic acid of claim 17, further encodes a guide RNA (gRNA) that provides the site-specificity to the DNA editing motif.
19. The nucleic acid of any one of claims 17-18, wherein the nucleic acid is comprised in an expression vector.
20. The nucleic acid of claim 19, wherein the expression vector is a plasmid vector, an RNA vector, or a viral vector.
21. The nucleic acid of any one of claims 17-20, wherein the nucleic acid is packaged in a viral particle, a lipid nanoparticle, or a non-lipid nanoparticle.
22. A composition for editing DNA, comprising:
the construct of any one of claims 1-16; and
a single stranded DNA (ssDNA) having a DNA insert, a 5'-homology arm, and a 3'-homology arm, wherein the 5' homology arm and the 3' homology arm are complementary to the nuclear DNA in a target region.
23. The composition of claim 22, wherein least one of the following applies:

(a) the ssDNA is a circular ssDNA (cssDNA), a linear ssDNA (lssDNA), a double stranded DNA/circular single stranded DNA (dsDNA/cssDNA) hybrid having an ssDNA section, or an RNA/DNA hybrid having an ssDNA section;

(b) the ssDNA is chemically modified;

(c) the ssDNA is produced from a cell free system or a cell culture system.

24. The composition of any one of claims 22-23, wherein each of the 5'-homology arm and the 3'-homology arm independently has a length ranging from about 25 nucleotides to about 5000 nucleotides.

25. The composition of any one of claims 22-24, wherein the DNA insert has a length ranging from 1 nucleotide to about 50,000 nucleotides.

26. A kit for editing DNA, comprising:

the construct of any one of claims 1-16 and/or the nucleic acid of any one of claims 17-21; and

a single stranded DNA (ssDNA), a nucleic acid that encodes the ssDNA, or a nucleic acid for constructing the ssDNA,

wherein the ssDNA comprises a DNA insert, a 5'-homology arm, and a 3'-homology arm, wherein the 5' homology arm and the 3' homology arm are complementary to the nuclear DNA in a target region.

27. The kit of claim 26, wherein least one of the following applies:

(a) the ssDNA is a circular ssDNA (cssDNA), a linear ssDNA (lssDNA), a double stranded DNA/circular single stranded DNA (dsDNA/cssDNA) hybrid having an ssDNA section, or an RNA/DNA hybrid having an ssDNA section;

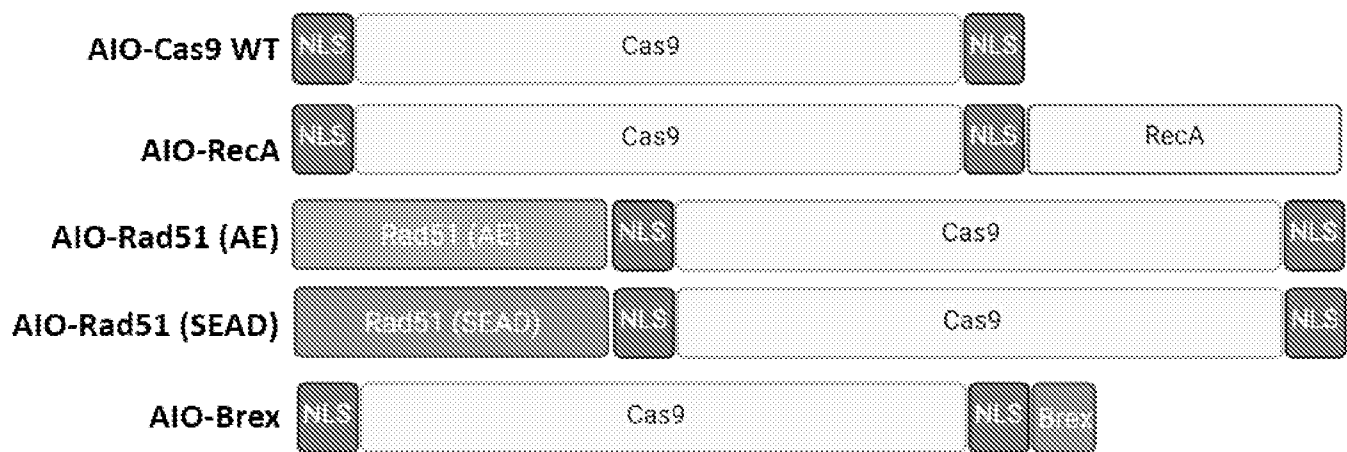
(b) the ssDNA is chemically modified;

(c) the ssDNA is produced from a cell free system or a cell culture system.

28. The kit of any one of claims 26-27, wherein each of the 5'-homology arm and the 3'-homology arm independently has a length ranging from about 25 nucleotides to about 5000 nucleotides.
29. The kit of any one of claims 26-28, wherein the DNA insert has a length ranging from 1 nucleotide to about 50,000 nucleotides.
30. The kit of any one of claims 26-29, wherein at least one of the following applies:
- (a) the nucleic acid for constructing the ssDNA donor template encodes every elements of the ssDNA donor template, except for the DNA insert,
 - (b) the nucleic acid for constructing ssDNA donor template encodes every elements of the ssDNA donor template, except for the 5' homology arm and the 3' homology arm,
 - (c) the nucleic acid for constructing ssDNA donor template does not encode the 5' homology arm, the 3' homology arm, and the DNA insert.
31. The kit of claim 30, wherein the nucleic acid includes a multiple cloning site for inserting the 5' homology arm, the 3' homology arm, and/or the DNA insert.
32. The kit of any one of claims 30-31, wherein the nucleic acid for constructing ssDNA donor template is a plasmid or a viral vector.
33. The kit of any one of claim 27-32, further comprising a component for delivering the construct, the nucleic acid for encoding the construct, the ssDNA, and/or the nucleic acid for encoding or constructing the ssDNA into a cell.
34. A method of editing a nuclear DNA in a cell, the method comprising delivering into the cell the construct of any one of claims 1-16 and/or the nucleic acid of any one of claims 17-21; and a single stranded DNA (ssDNA) having a DNA insert, a 5'-homology arm, and a 3'-homology arm, wherein the 5' homology arm and the 3' homology arm are complementary to the nuclear DNA in a target region.

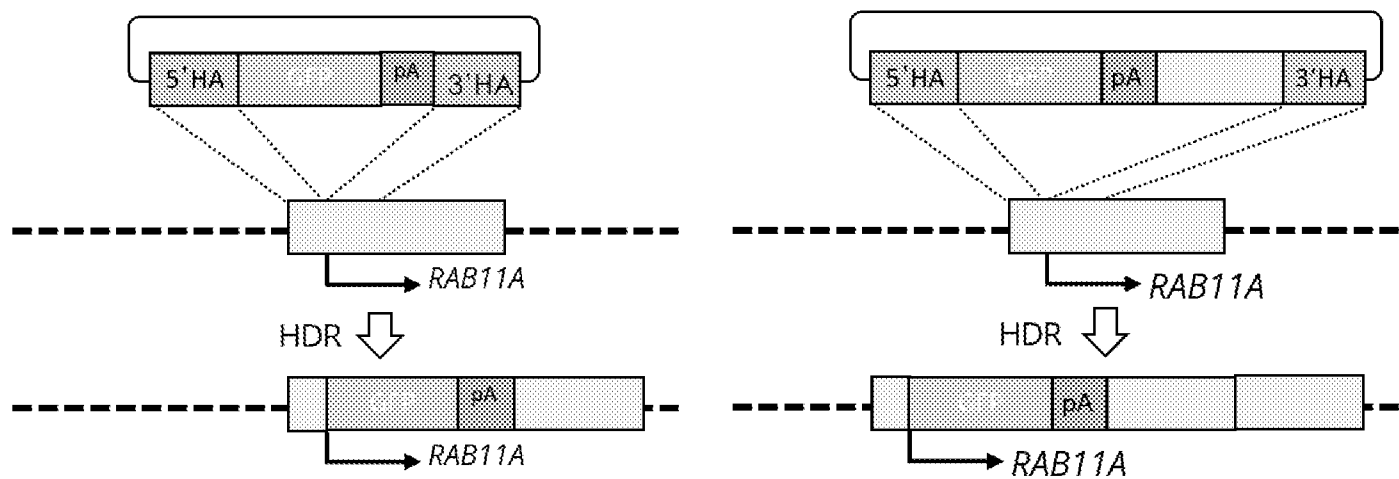
35. The method of claim 34, wherein at least one of the following applies:
- (a) the ssDNA is a circular ssDNA (cssDNA), a linear ssDNA (lssDNA), a double stranded DNA/circular single stranded DNA (dsDNA/cssDNA) hybrid having an ssDNA section, or an RNA/DNA hybrid having an ssDNA section;
 - (b) the ssDNA is chemically modified;
 - (c) the ssDNA is produced from a cell free system or a cell culture system.
36. The method of any one of claims 34-35, wherein each of the 5'-homology arm and the 3'-homology arm independently has a length ranging from about 25 nucleotides to about 5000 nucleotides.
37. The method of any one of claims 34-36, wherein the DNA insert has a length ranging from 1 nucleotide to about 50,000 nucleotides.
38. The method of any one of claims 34-37, wherein the construct and/or nucleic acid and the ssDNA are delivered into the cell by a viral vector, a lipid or non-lipid nanoparticle delivery, an electroporation delivery, a gene gun delivery, or an injection.
39. The method of any one of claims 34-38, wherein the nuclear DNA is a genomic DNA.
40. The method of any one of claims 34-39, wherein the cell is a prokaryotic cell, a cultured plant cell, a eukaryotic cell, a cultured mammalian cell, a primary cell, an embryonic stem cell, an adult stem cell, a hematopoietic stem cell, an induced pluripotent stem cell, or a CAR-T cell.
41. The method of any one of claims 34-40, wherein the cell is in the body of a subject.
42. The method of claim 41, wherein the subject is a mammal, such as a human.
43. A non-natural polypeptide comprising an amino acid sequence having 90% or higher identity to the amino acid sequence set forth in SEQ ID NO:11.

44. The non-natural polypeptide of claim 43, wherein the non-natural polypeptide is a ssDNA-binding polypeptide.

**Fig. 1A**

RAB11A locus
2Kb payload - css76

RAB11A locus
4Kb payload - css116

**Fig. 1B**

Day 3 FACS profile

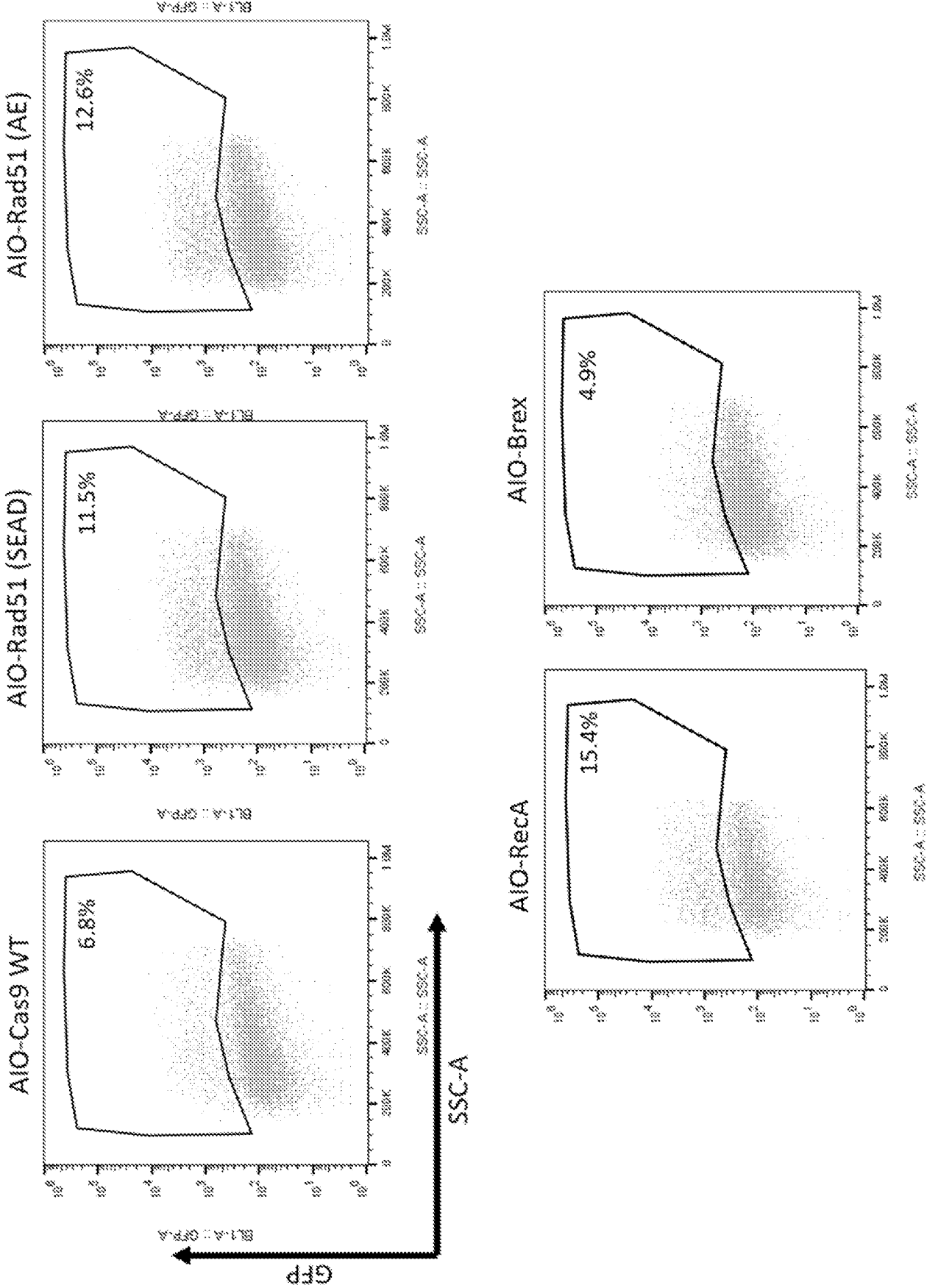


Fig. 1C

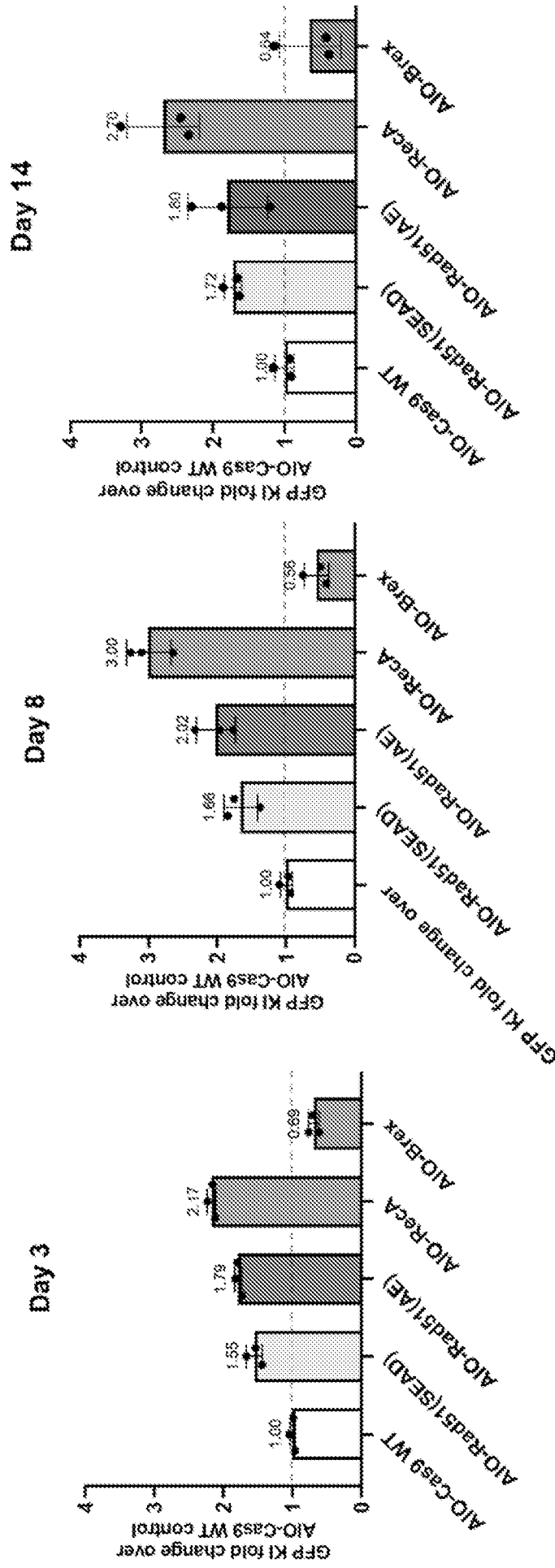


Fig. 1D

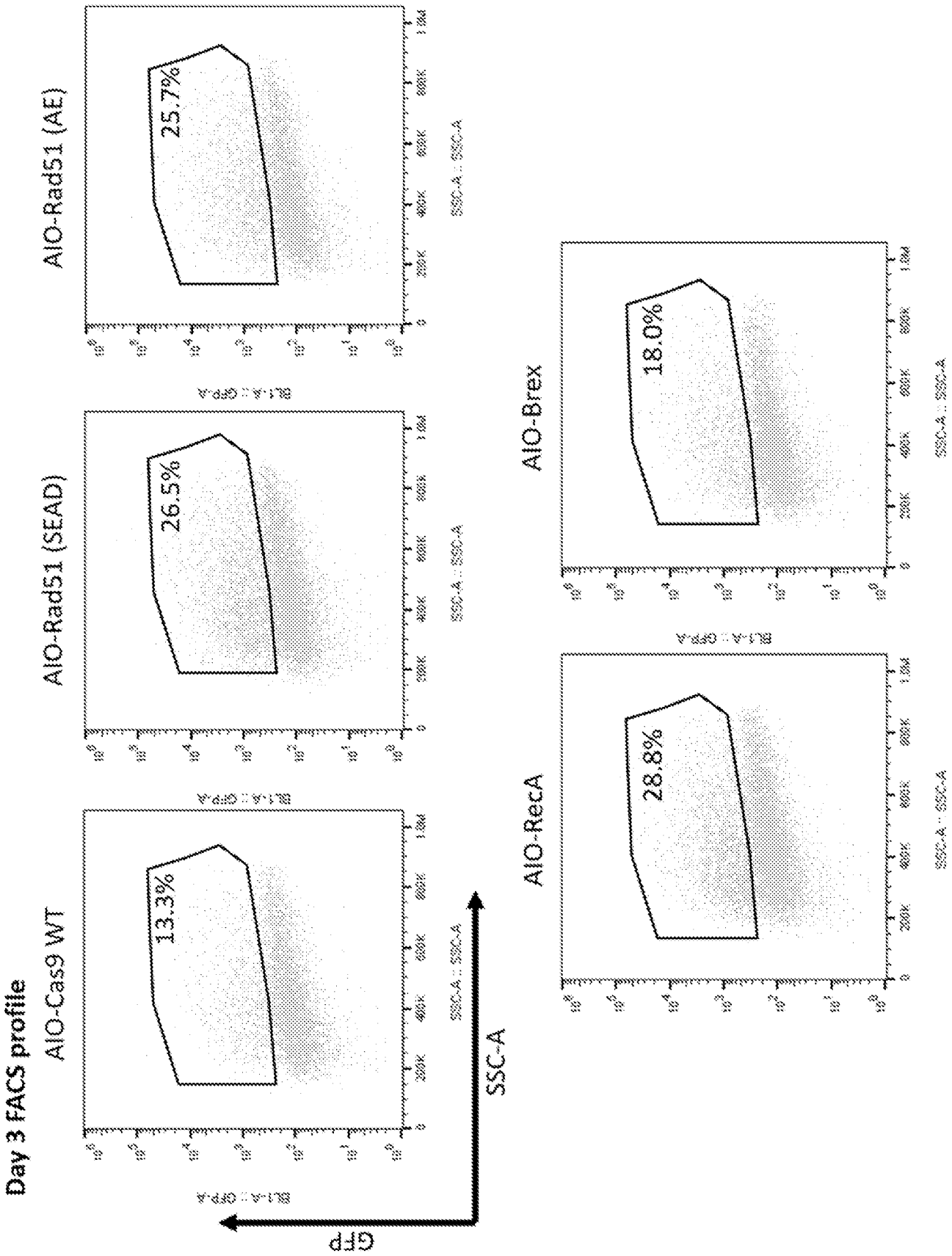


Fig. 1E

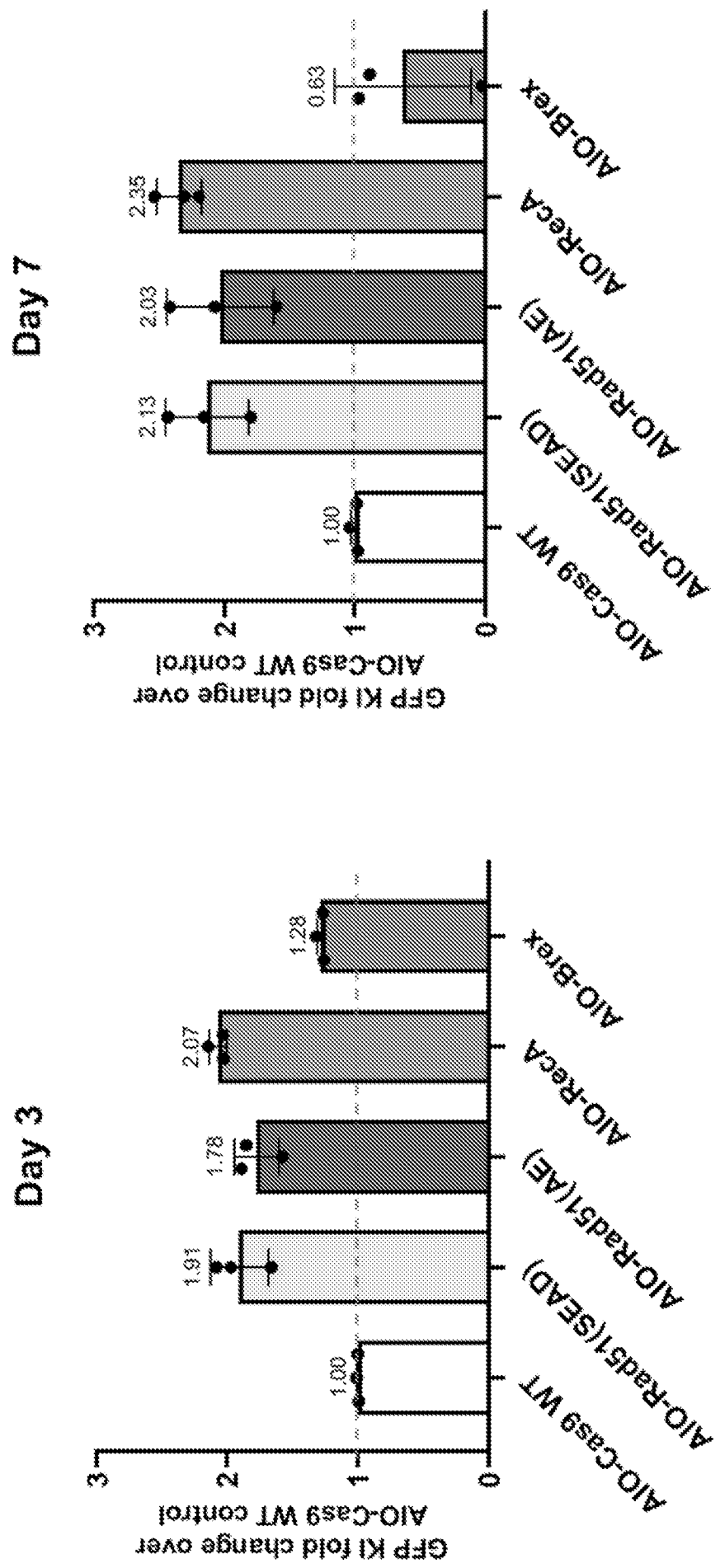
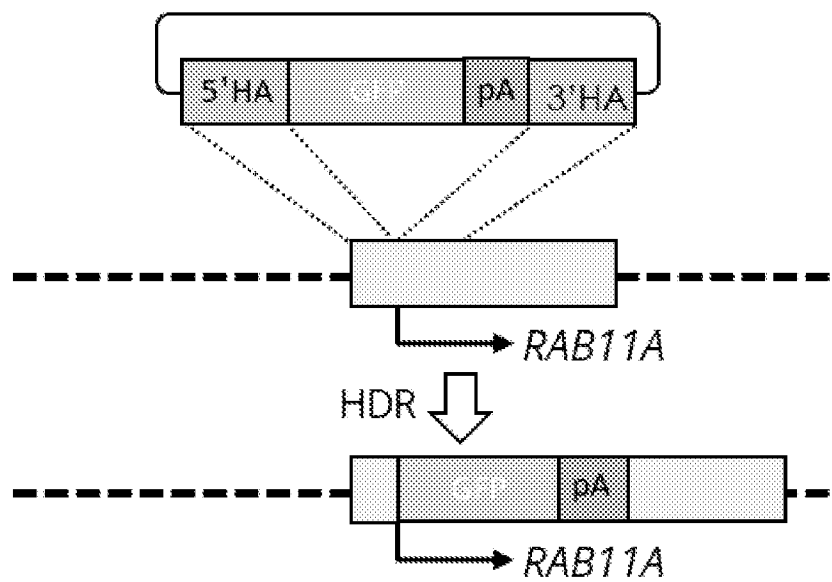


Fig. 1F

**Fig. 2A**

RAB11A locus
2Kb payload – css76

**Fig. 2B**

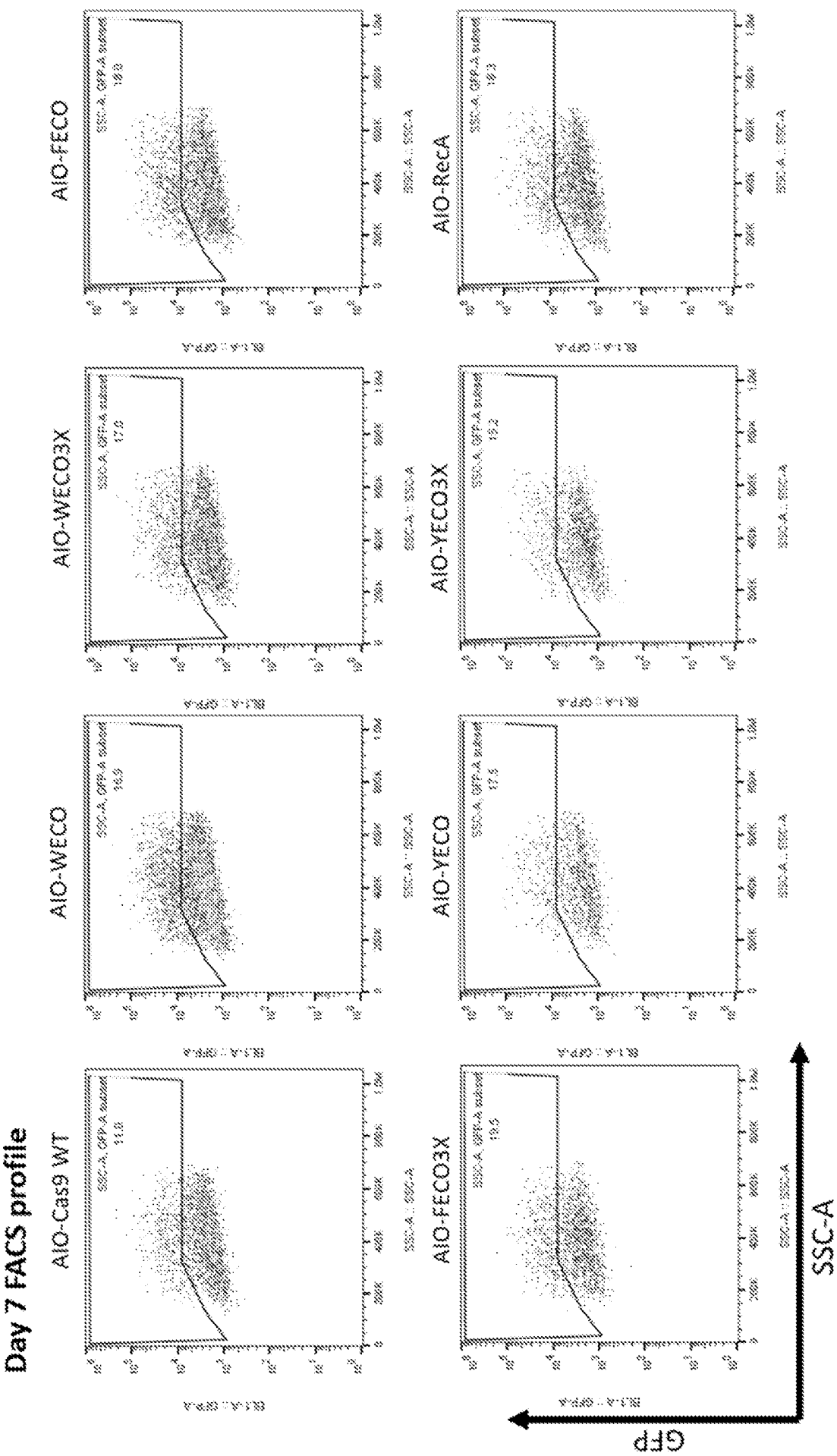


Fig. 2C

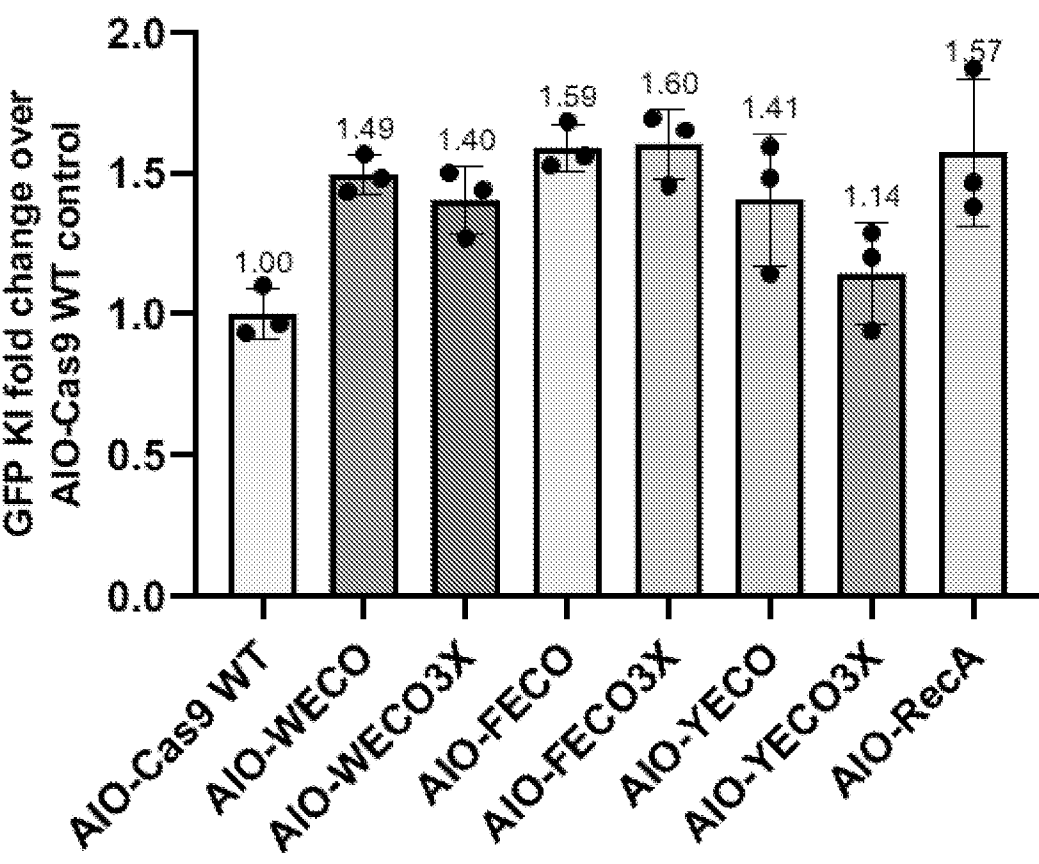
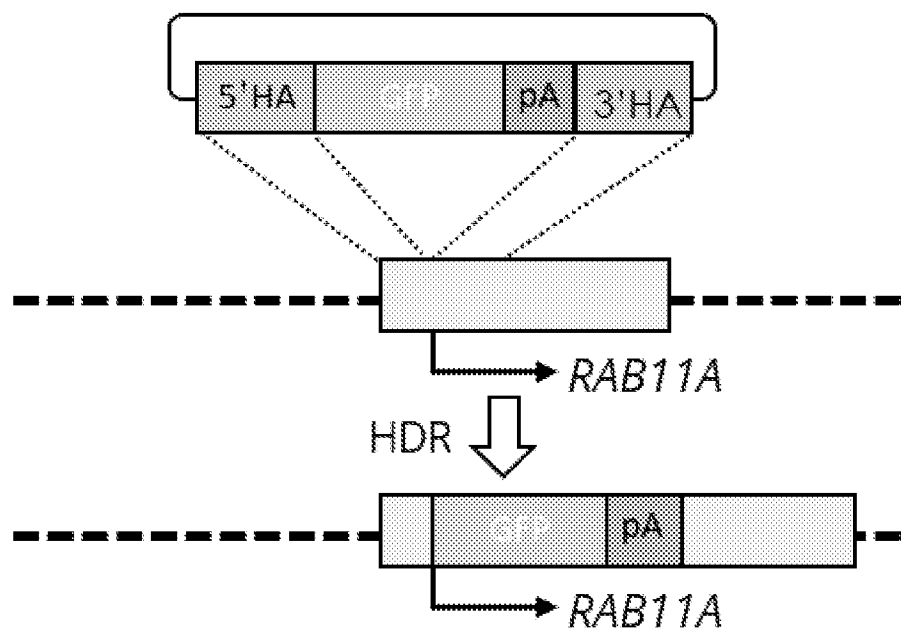


Fig. 2D

**Fig. 3A**

RAB11A locus
2Kb payload – css76

**Fig. 3B**

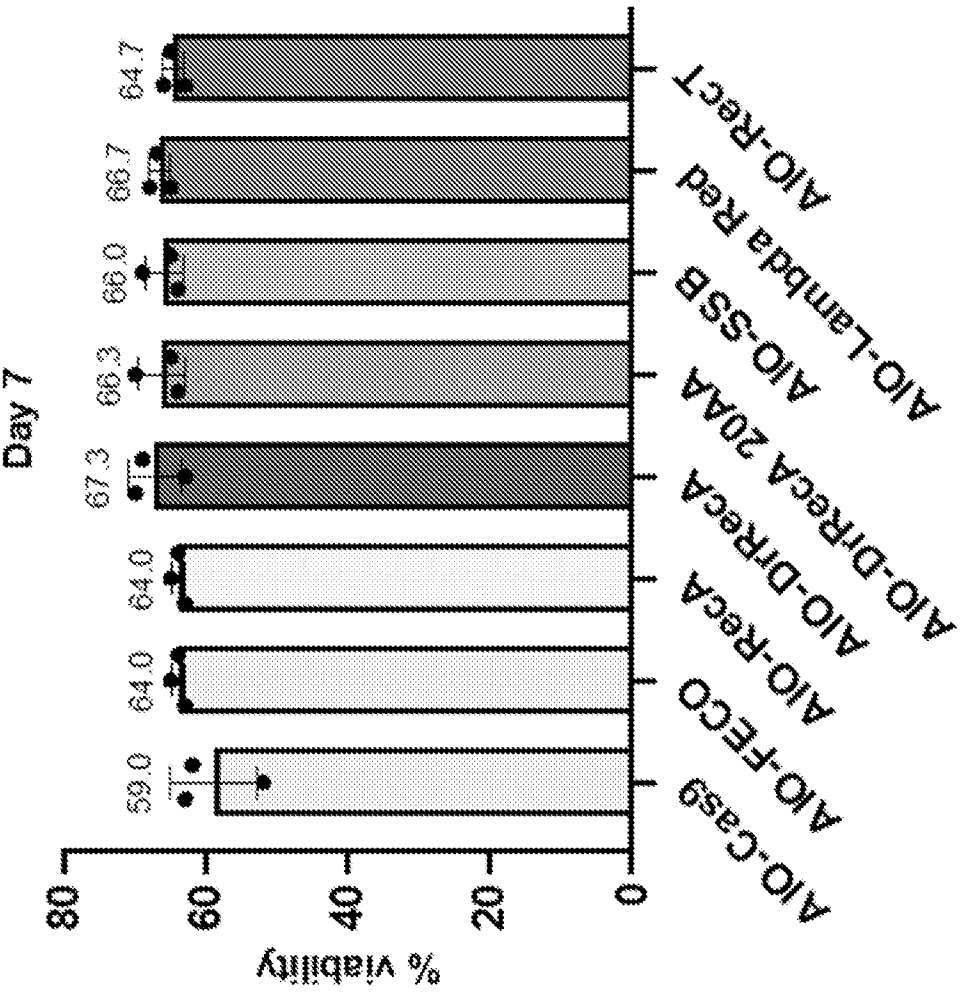


Fig. 3D

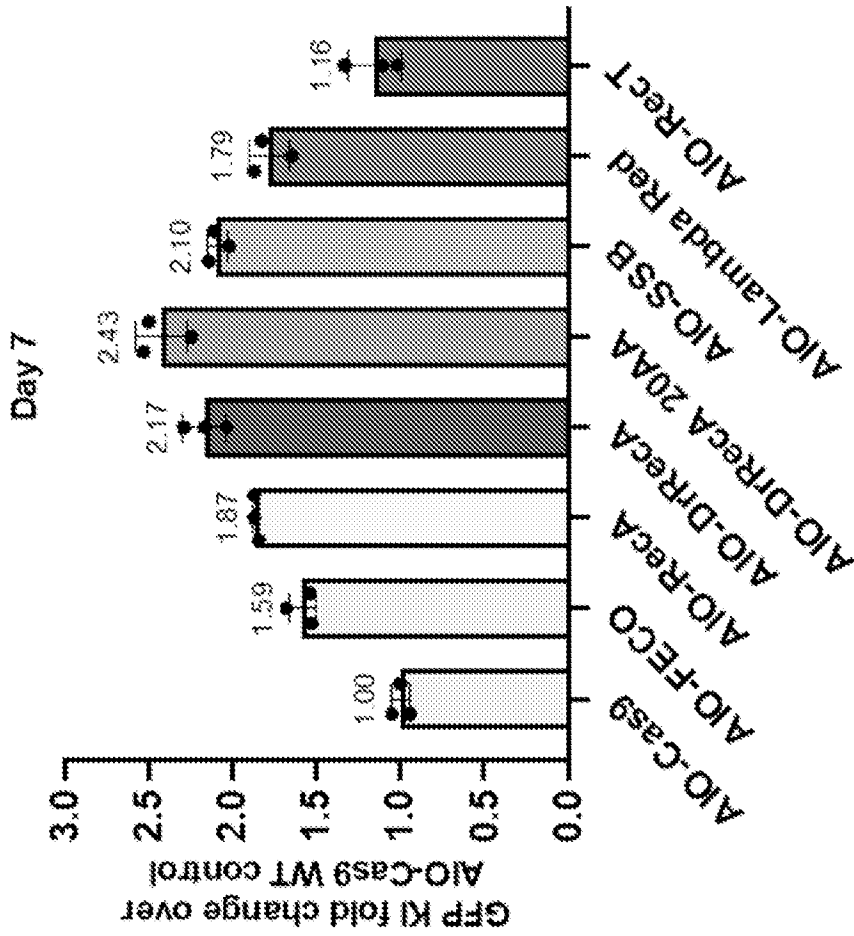


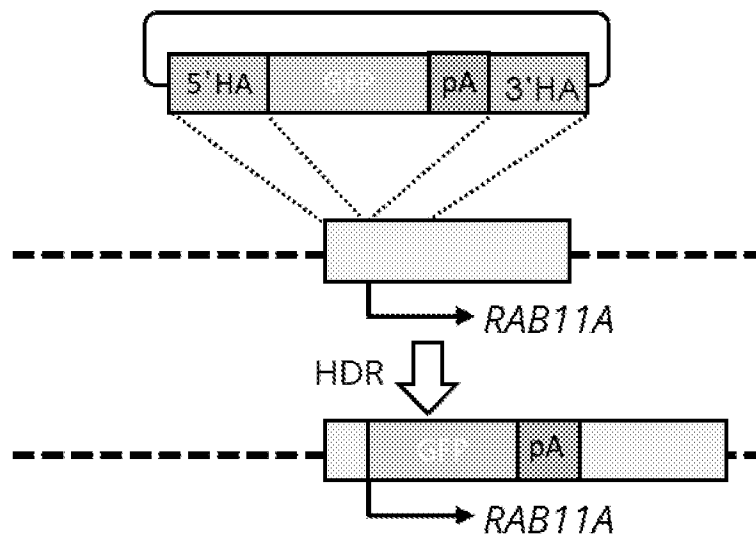
Fig. 3C

SsoRadA	NQVMARPD--MFYG--DPTVAVGG
MvRadA	NQVSAKPD--AFFG--MAEQAIGG
PfRadA	NQVQARPD--AFFG--DPTRPIGG
HsRad51	NQVVAQVDGAAMFA-ADPKKPIGG
ScRad51	NQVVAQVDGGMAFN-PDPKKPIGG
HsDMC1	NQMTADPGATMTFQ-ADPKKPIGG
ScDMC1	NQVQSDPGASALFASADGRKPIGG
EcRecA-FECO	NQIRMKIG--VMFG--NPETTTGG
EcRecA-WECO	NQIRMKIG--VMWG--NPETTTGG
EcRecA-YECO	NQIRMKIG--WMYG--NPETTTGG
DrRecA	NQVREKIG--VMYG--NPETTTGG

Fig. 3E

**Fig. 4A**

RAB11A locus
2Kb payload – css76

**Fig. 4B**

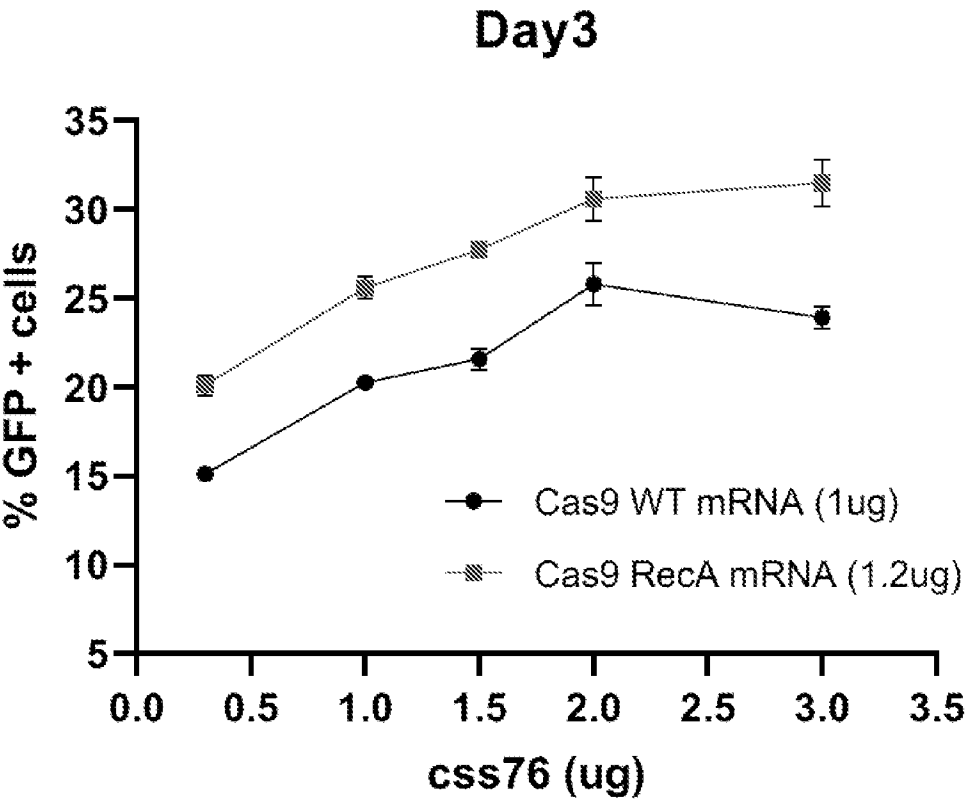


Fig. 4C

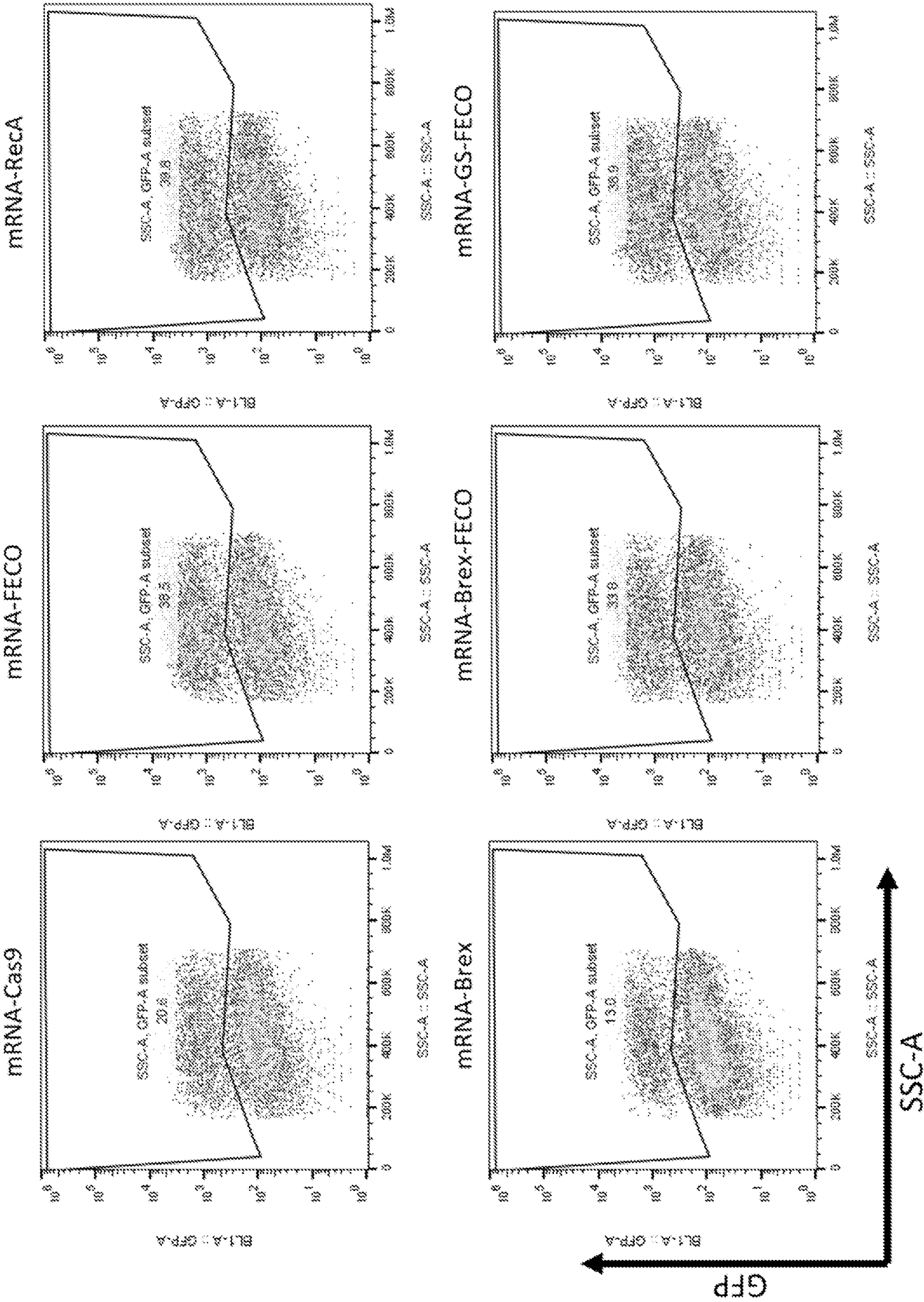


Fig. 4D

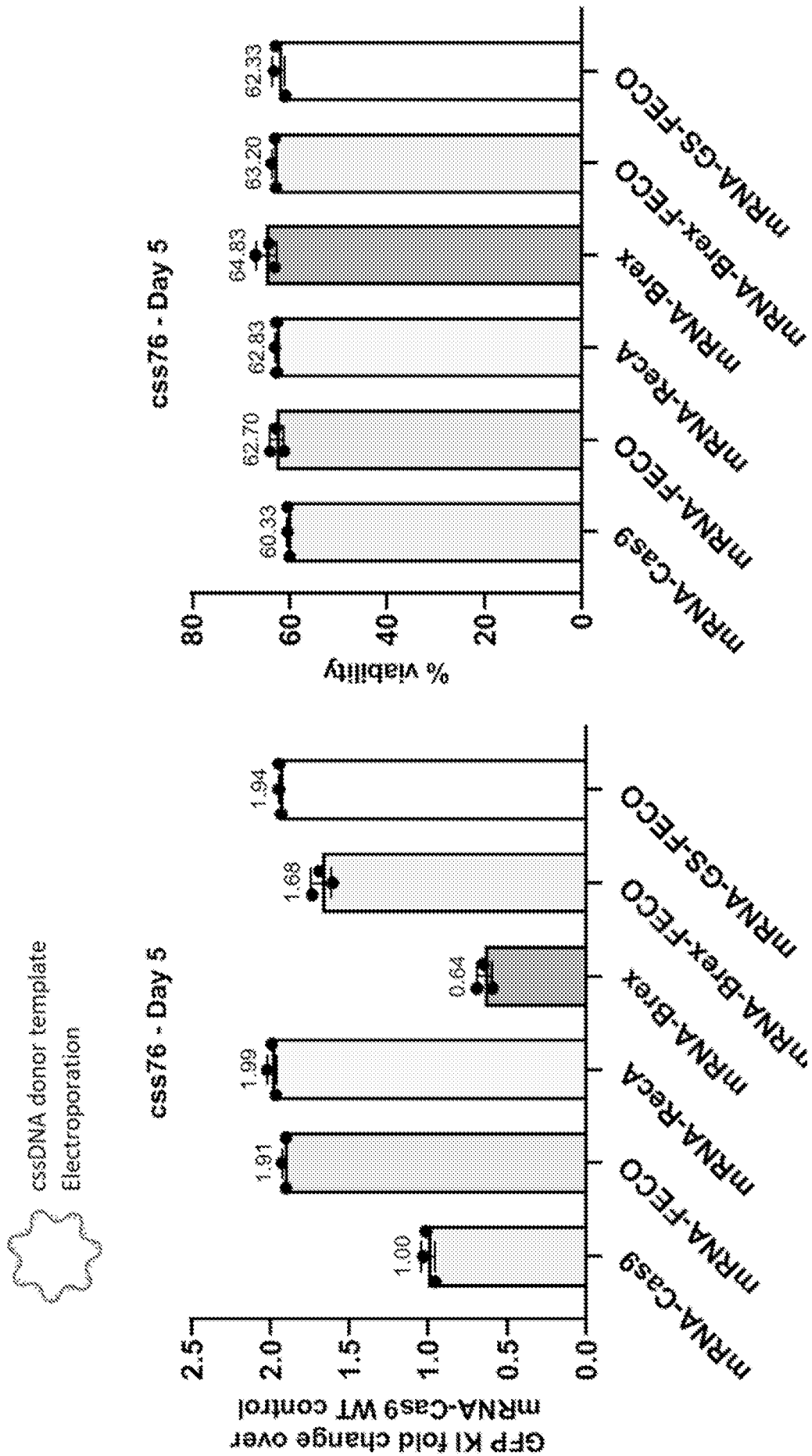


Fig. 4E

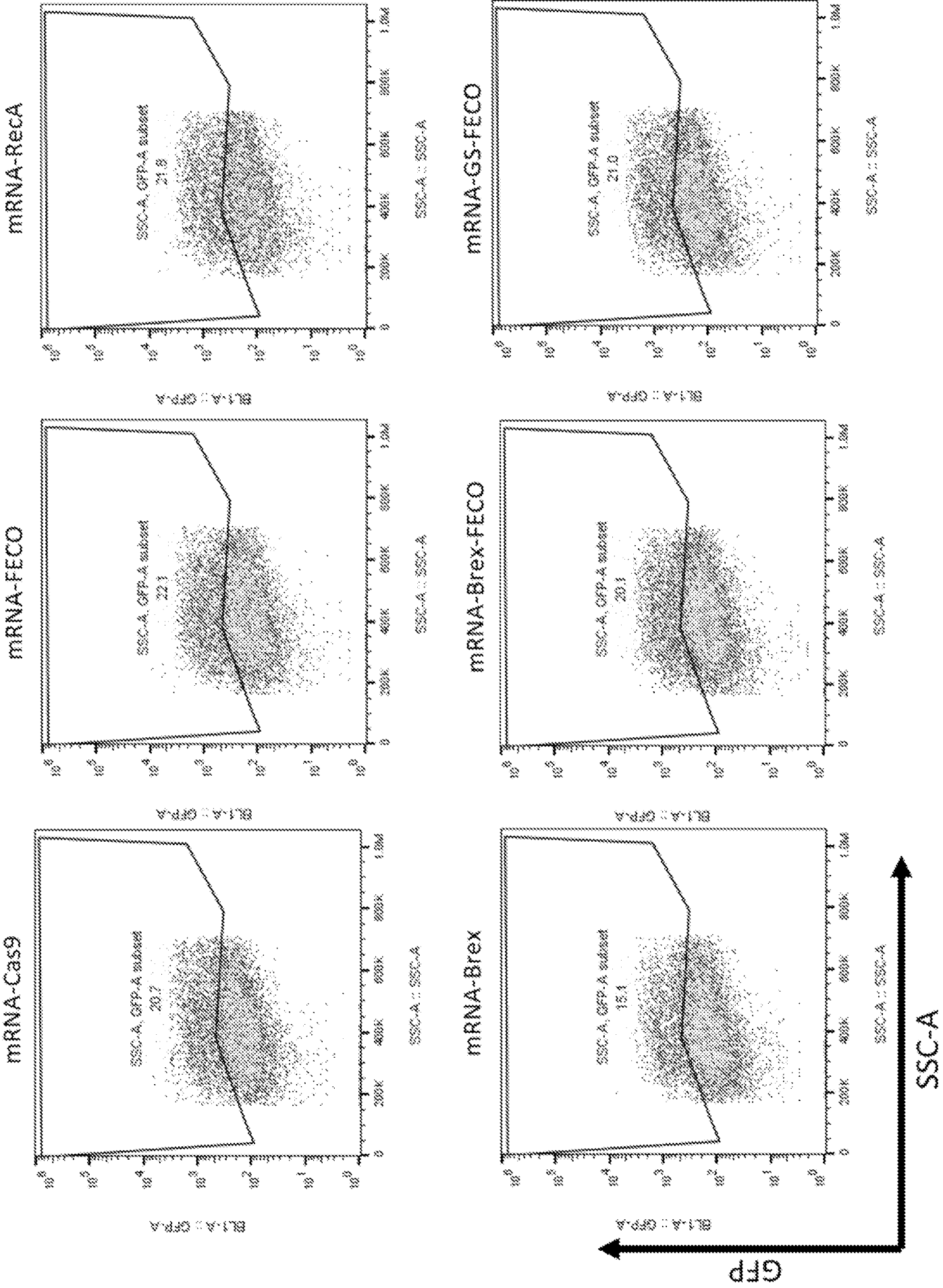


Fig. 4F

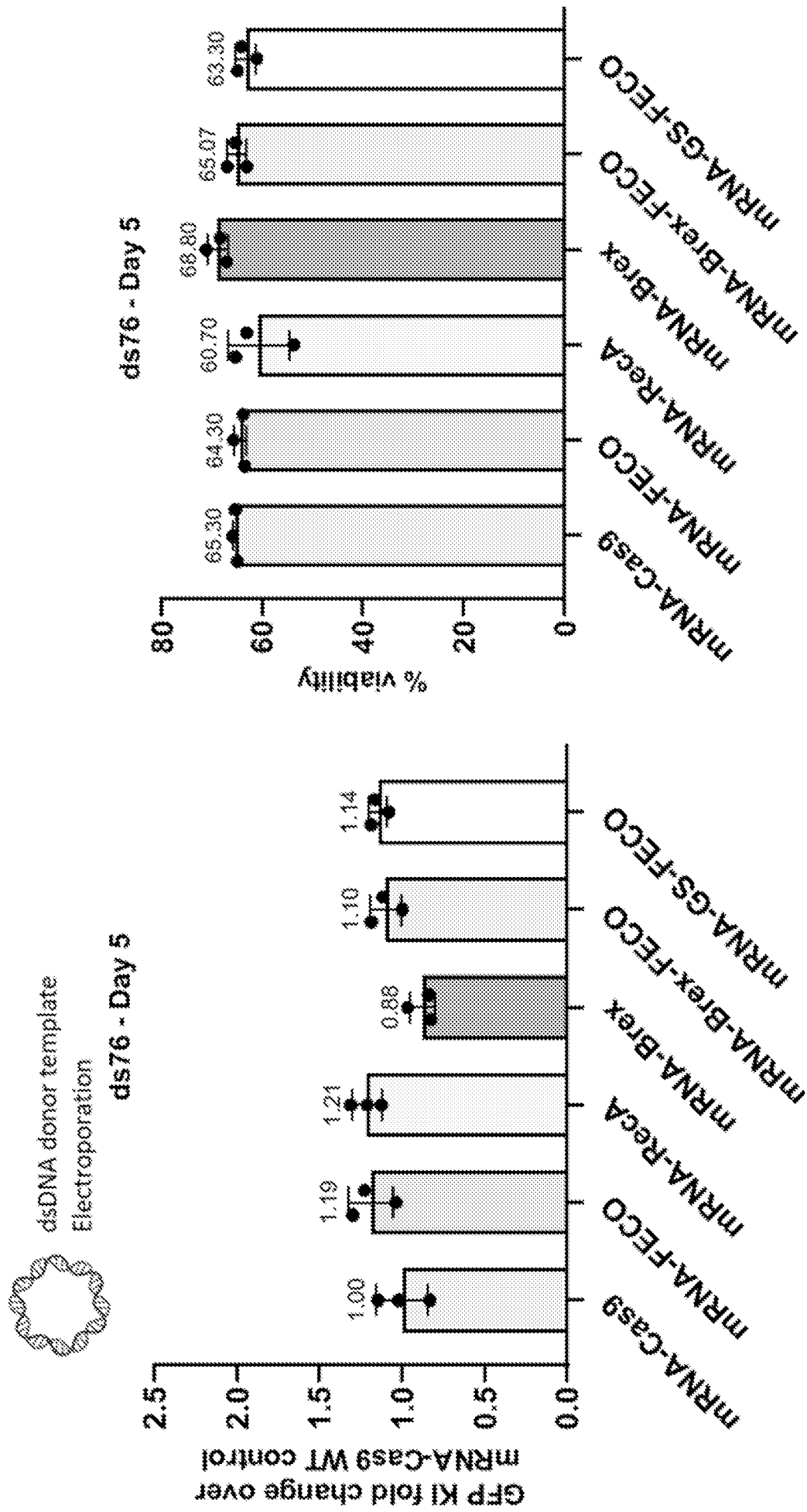


Fig. 4G

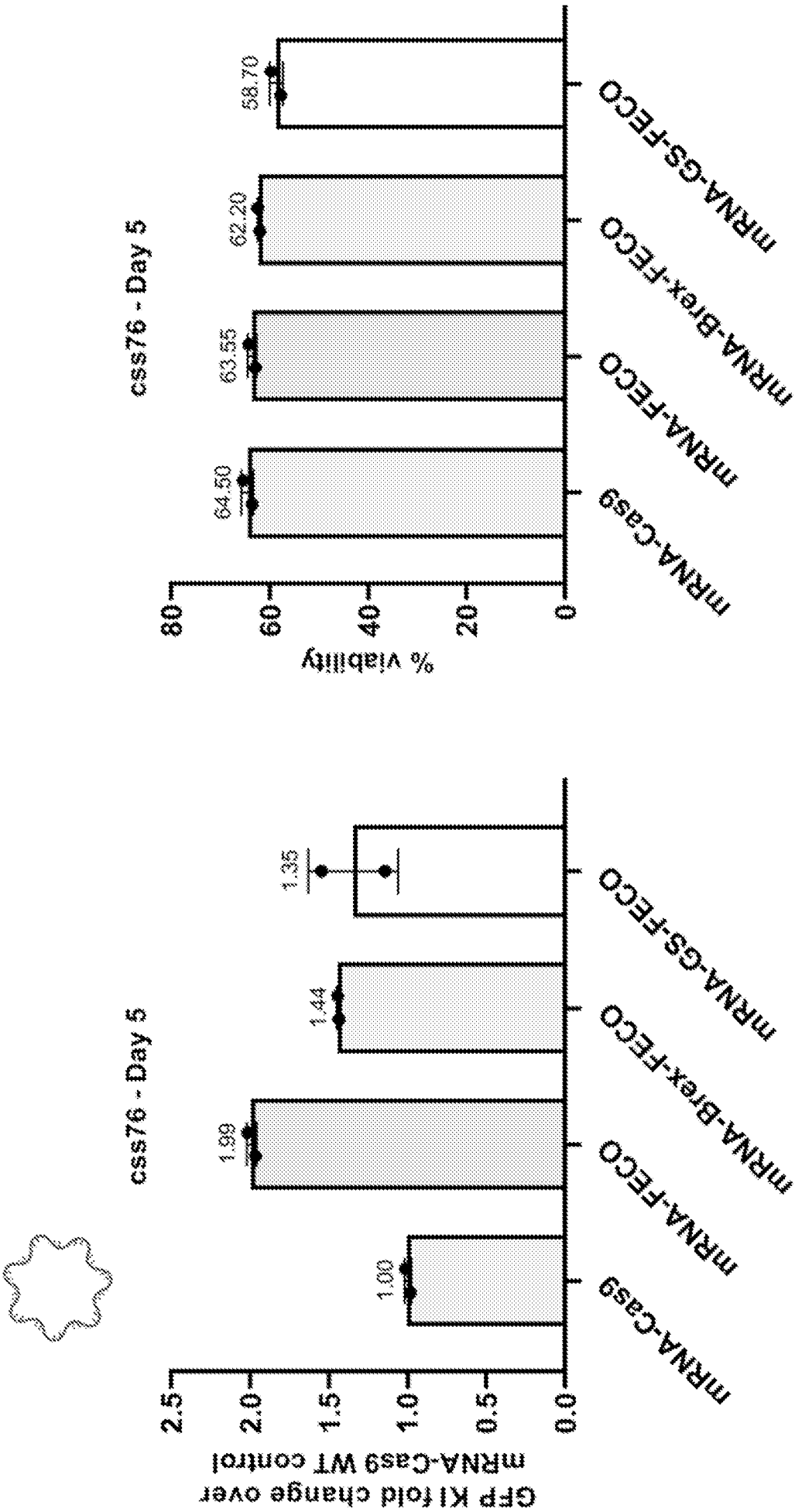


Fig. 4H

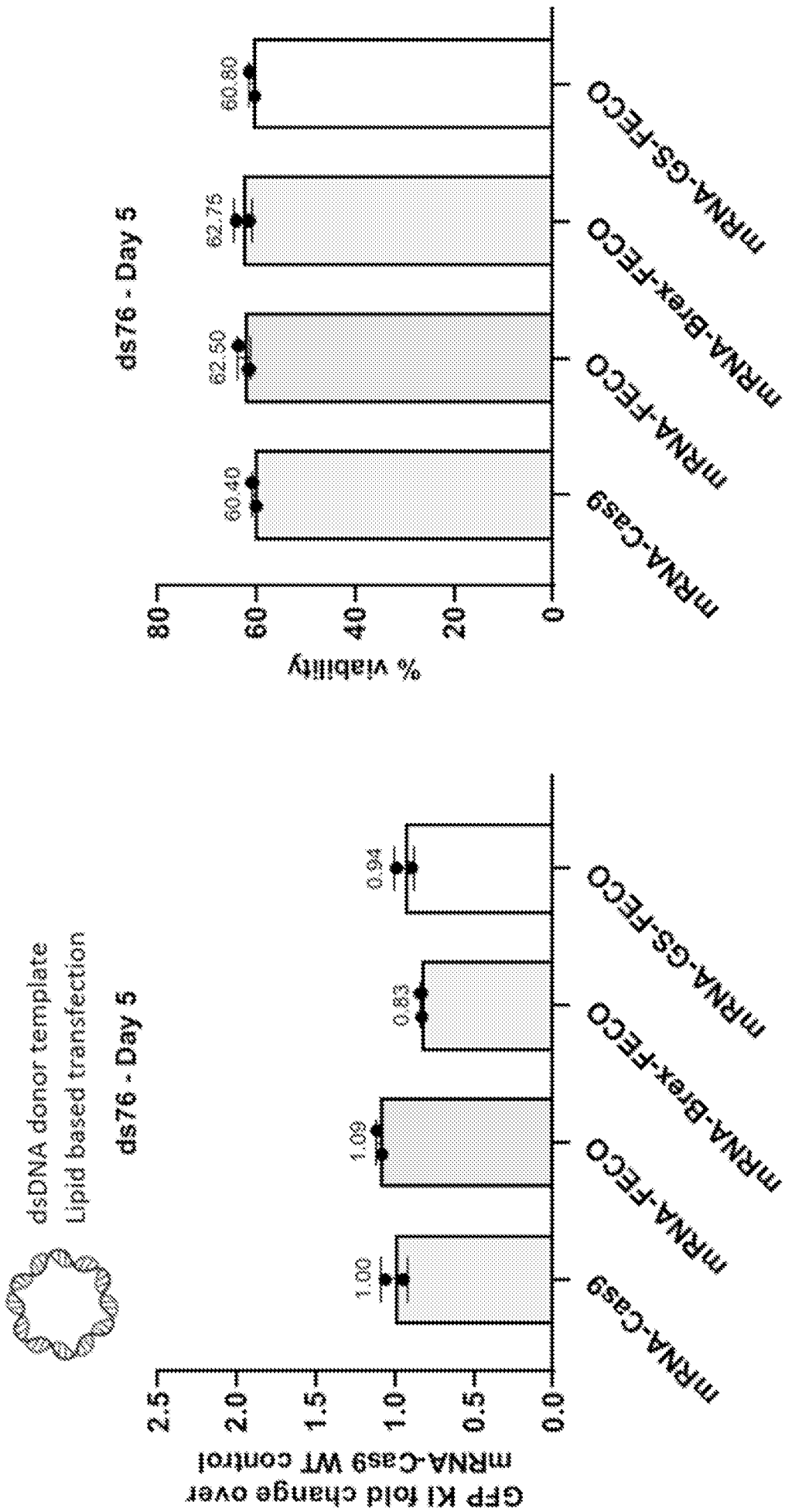


Fig. 4I

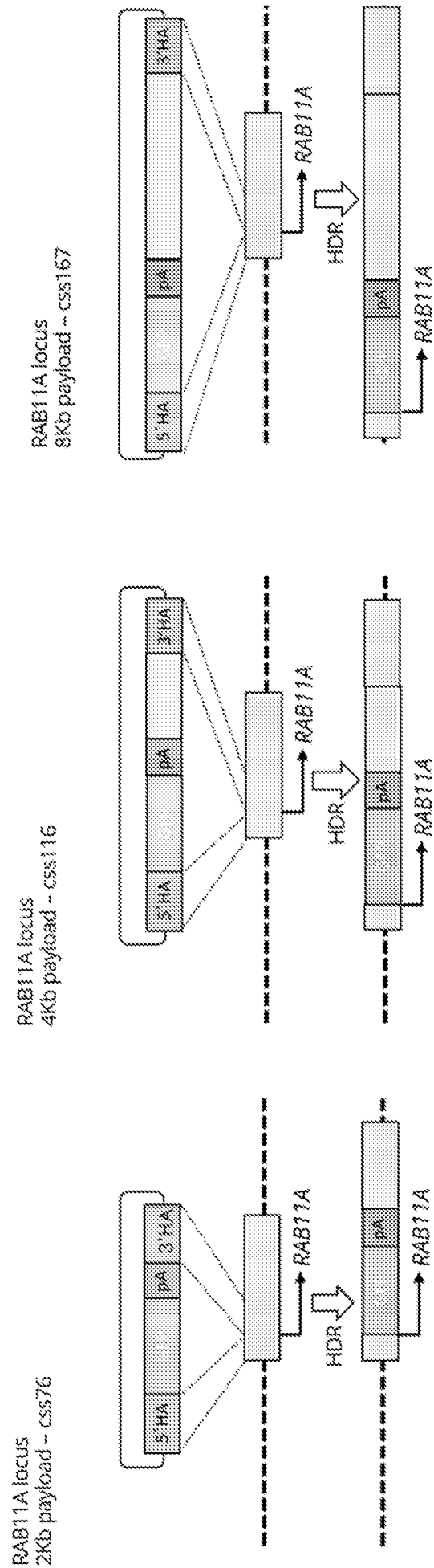


Fig. 5A

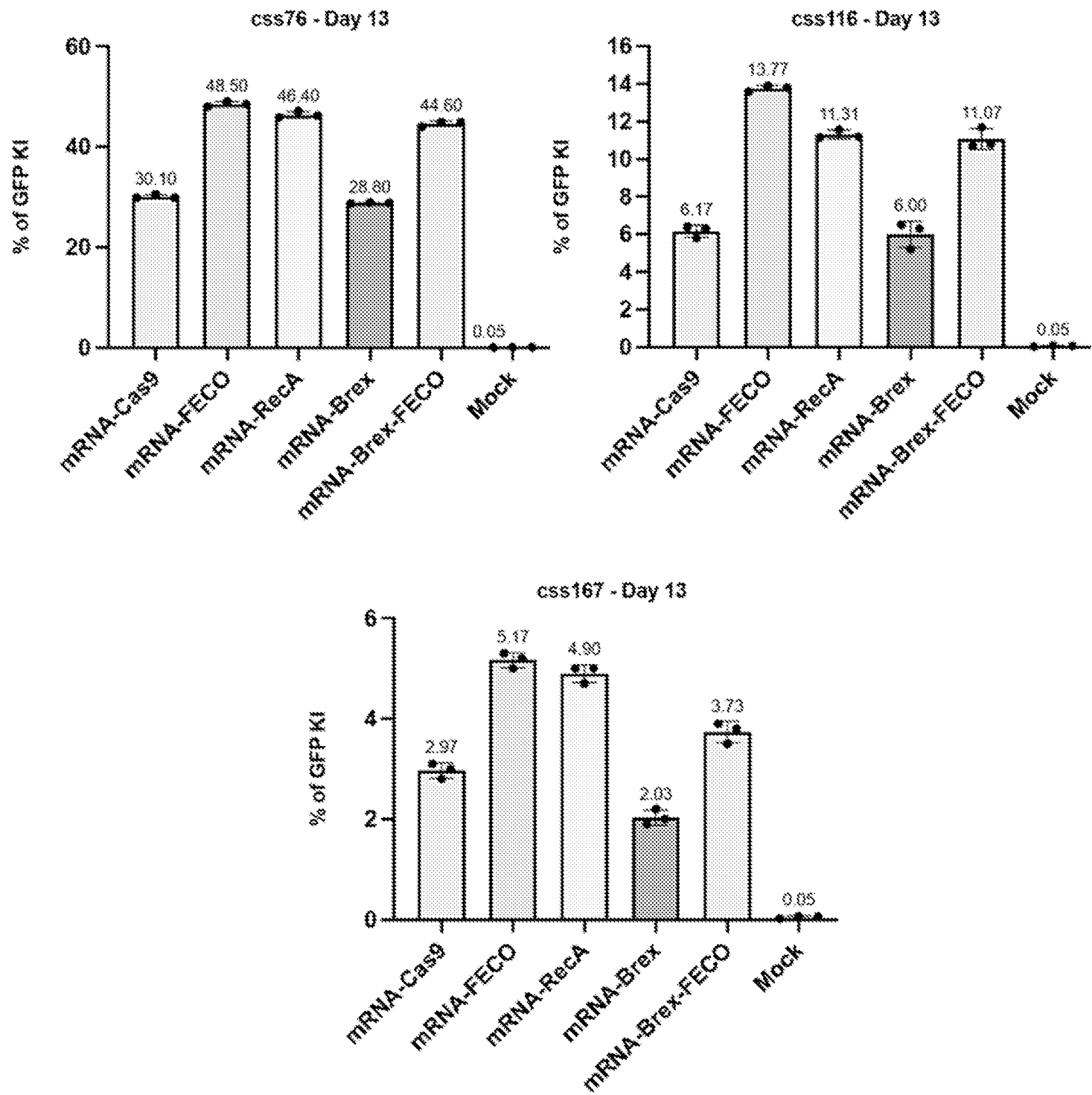


Fig. 5B

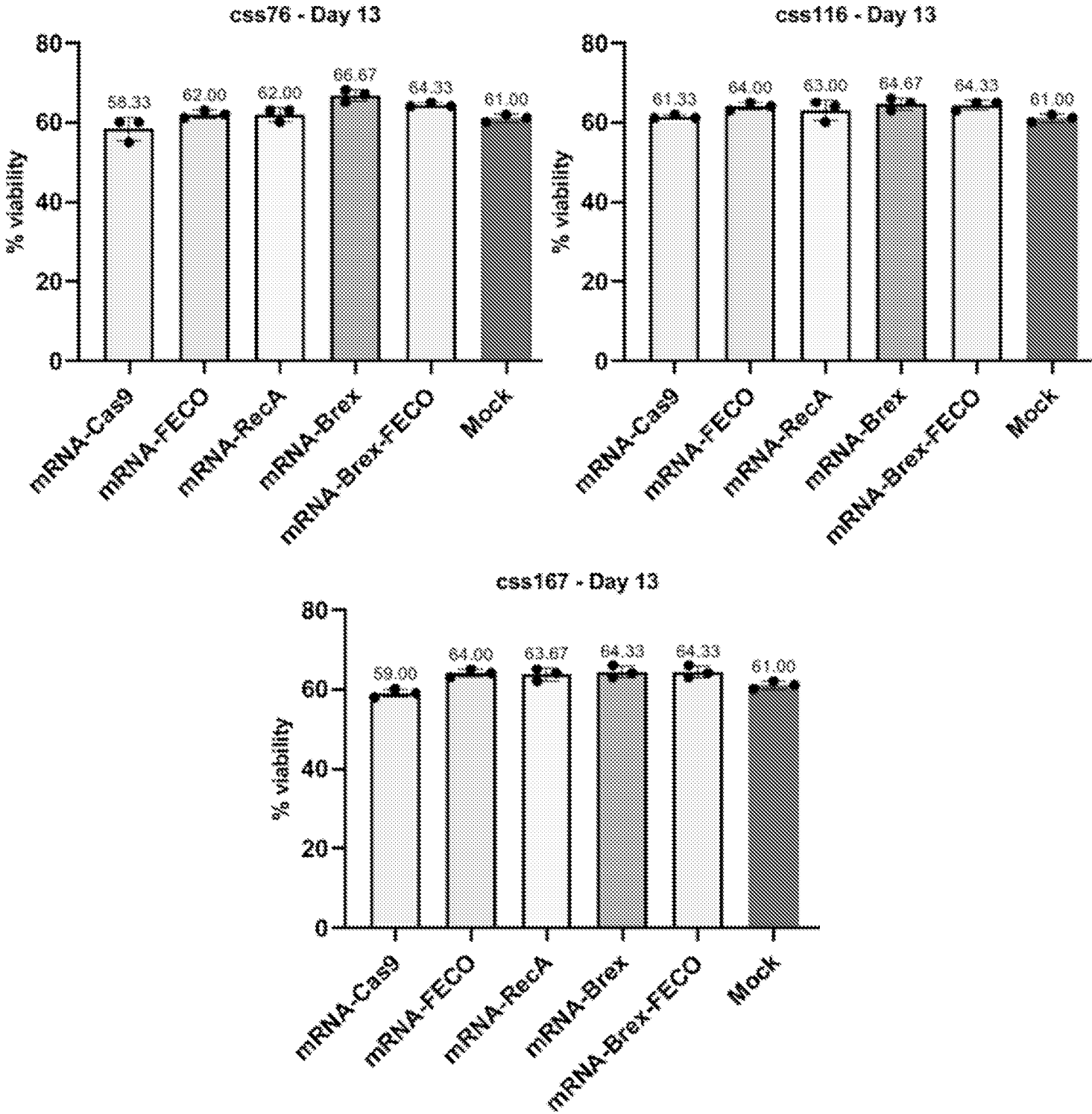


Fig. 5C

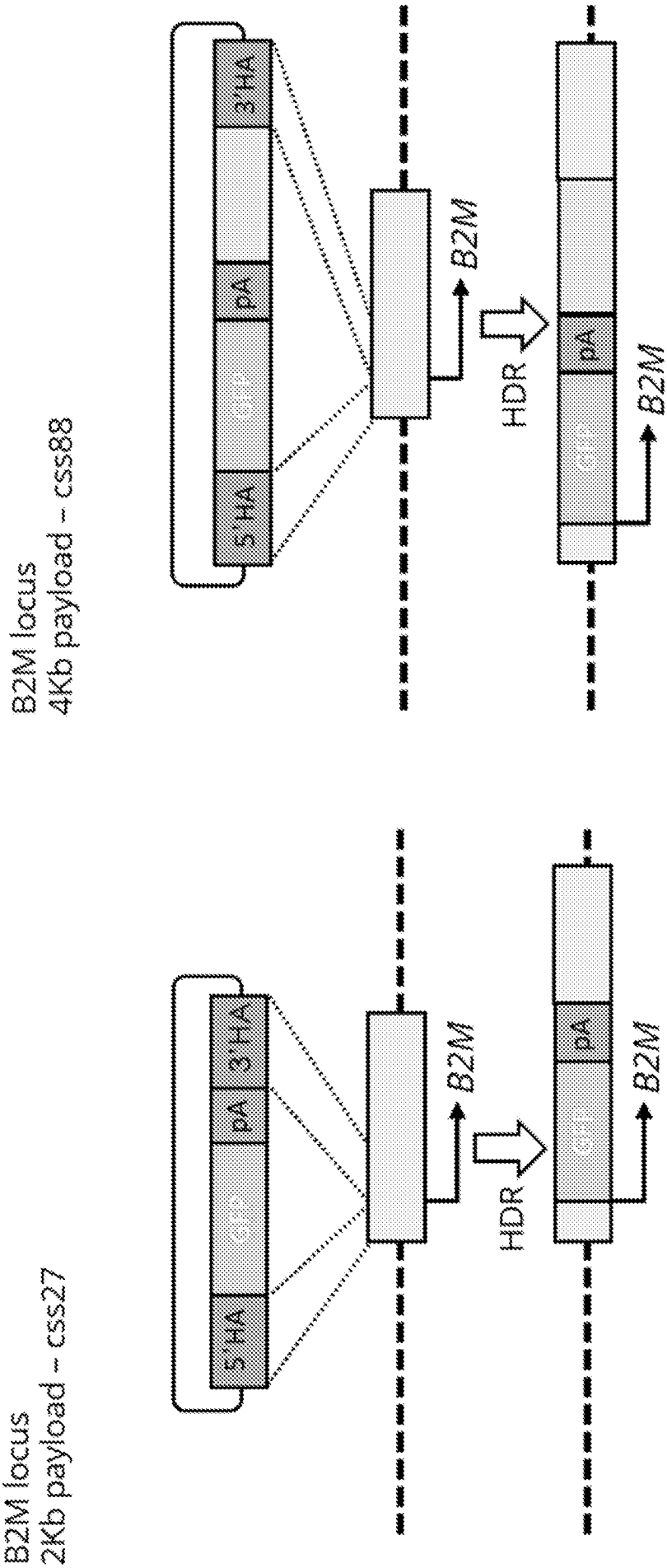


Fig. 5D

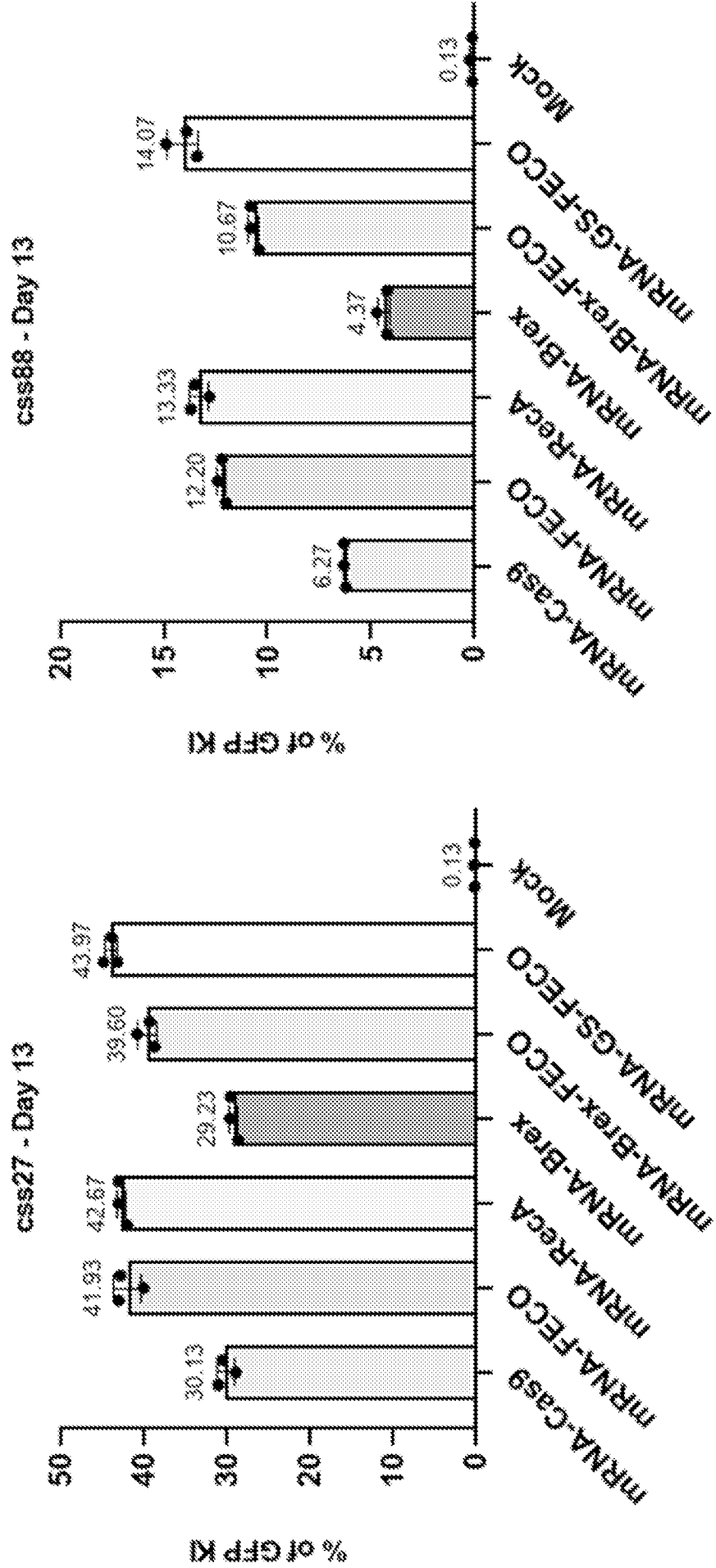


Fig. 5E

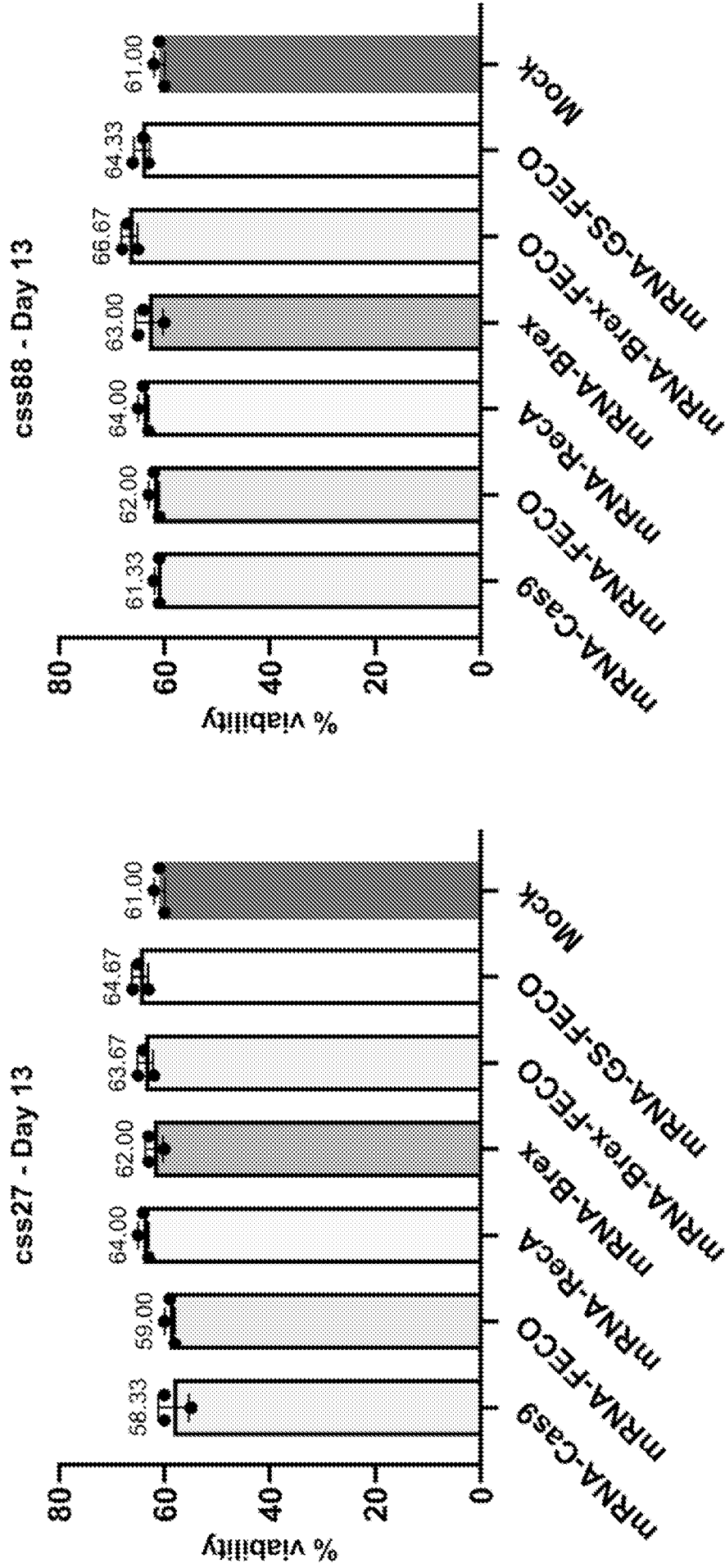


Fig. 5F

TRAC locus
3Kb CD19/CD22 payload
- c5562

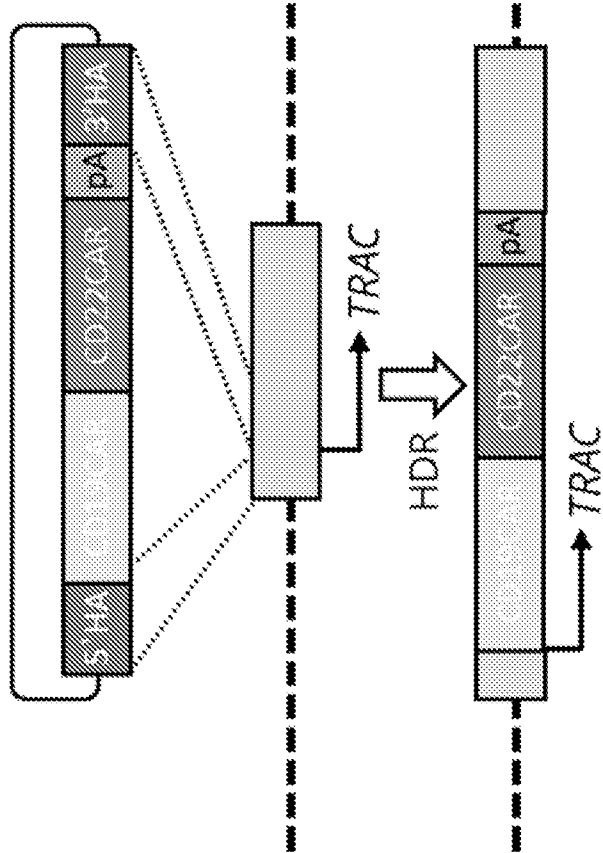


Fig. 6A

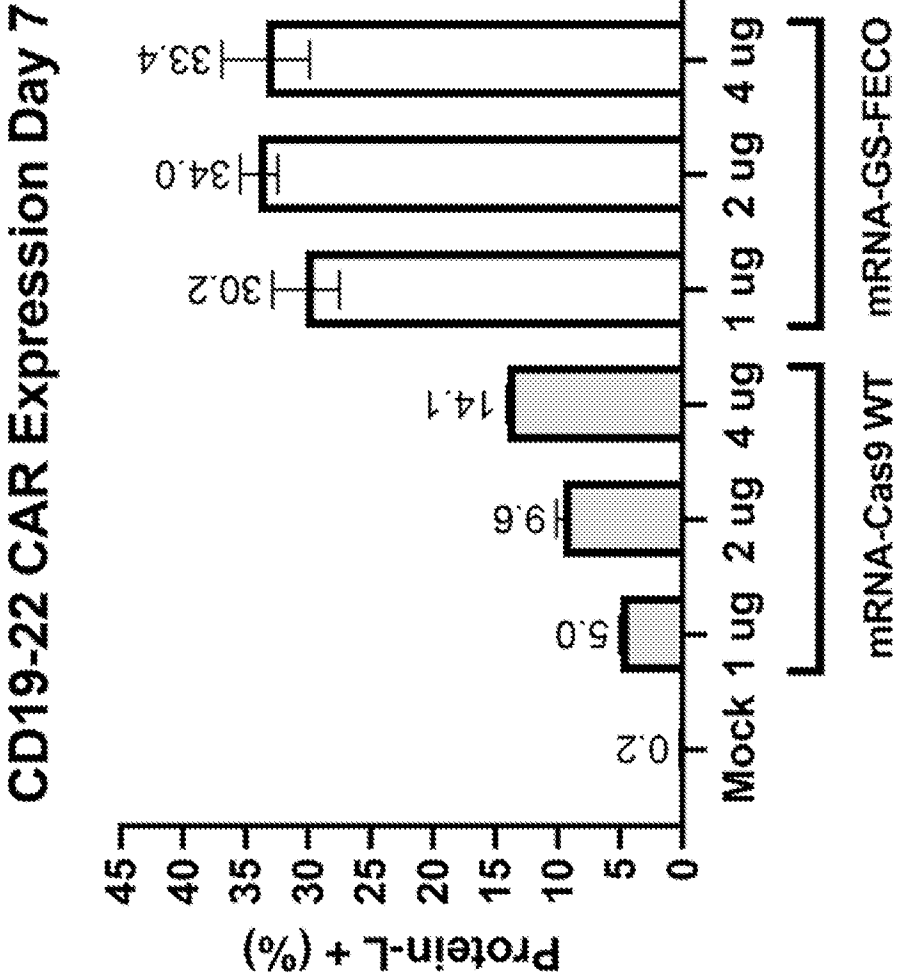


Fig. 6B

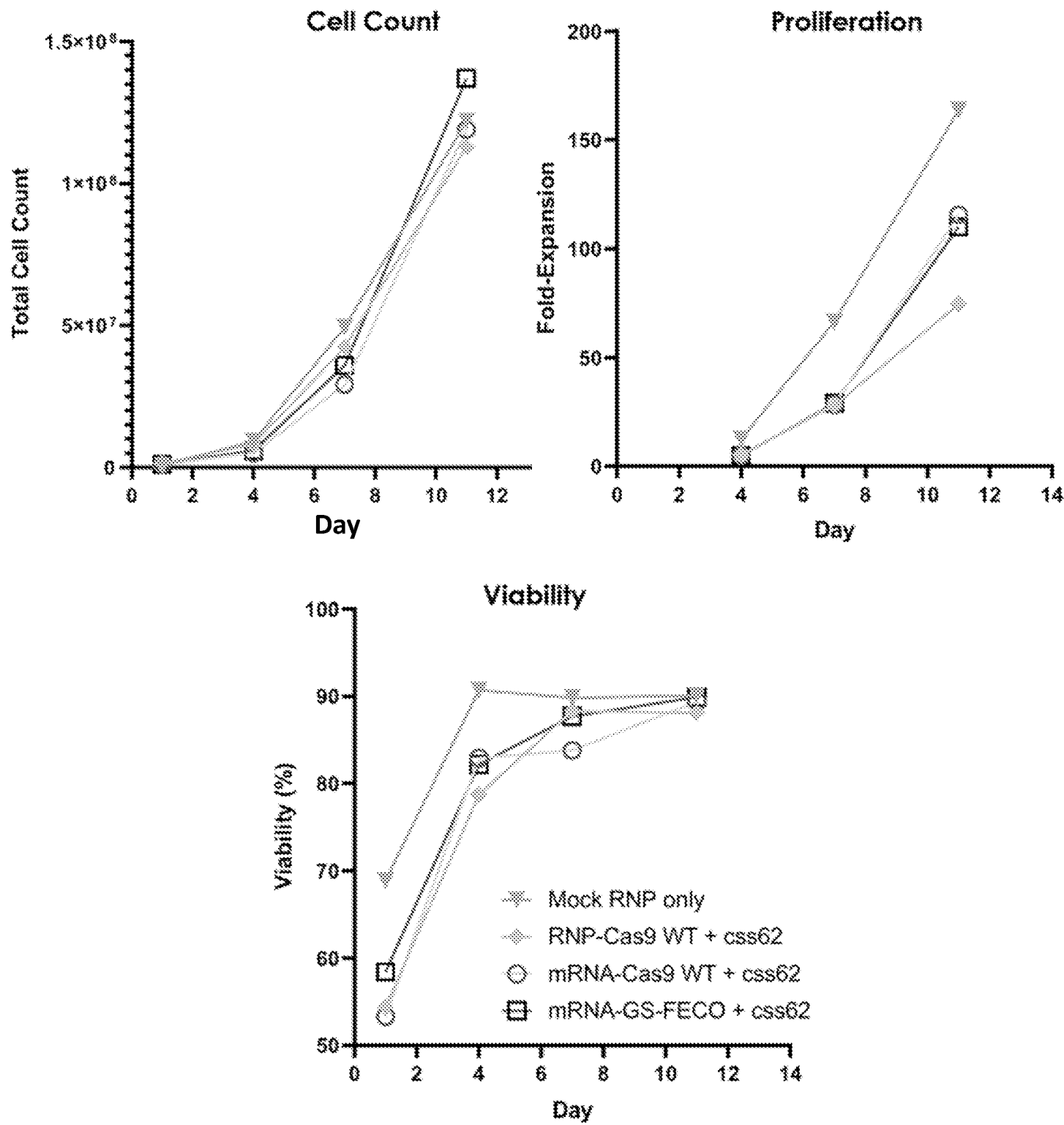


Fig. 6C

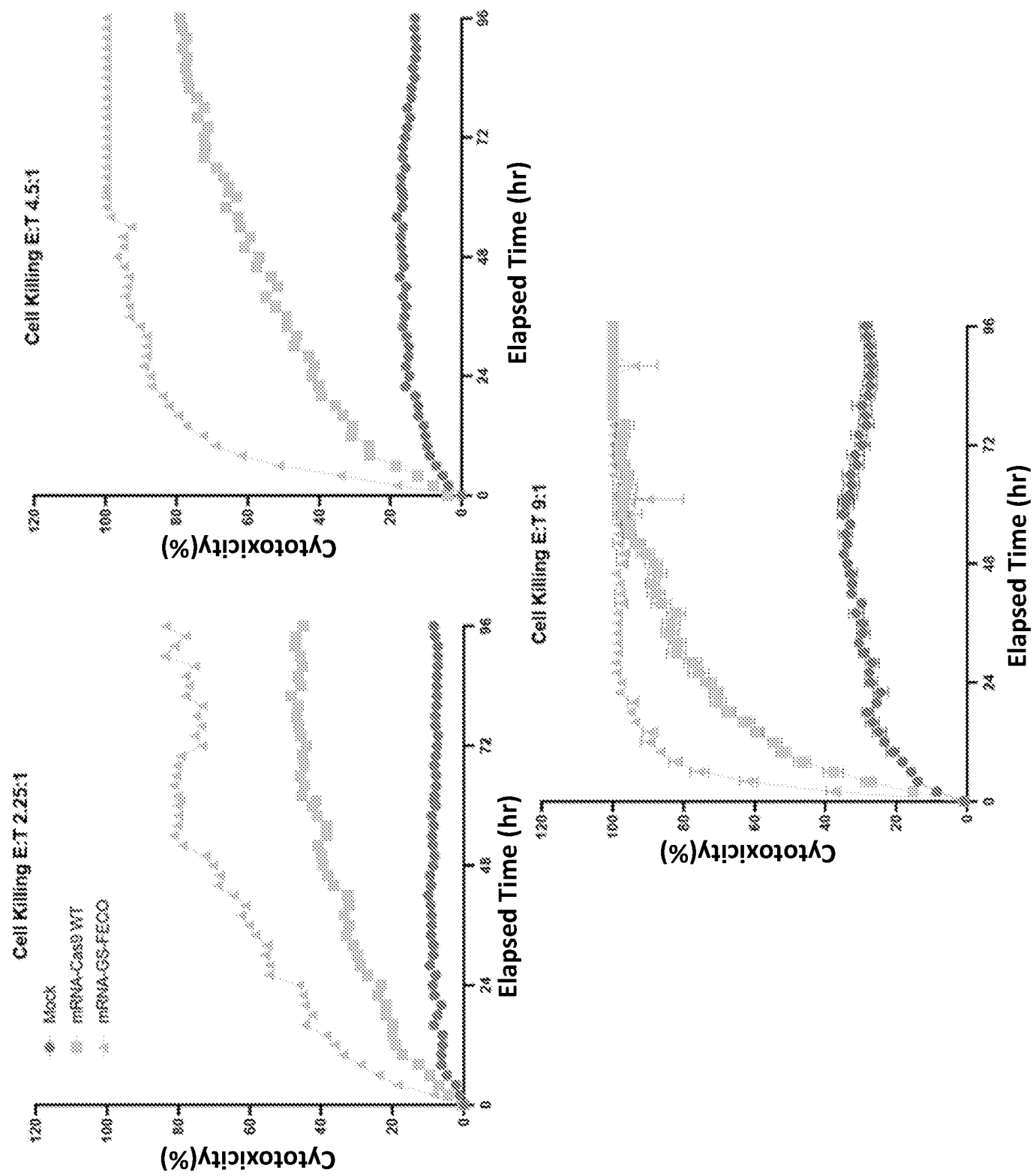


Fig. 6D

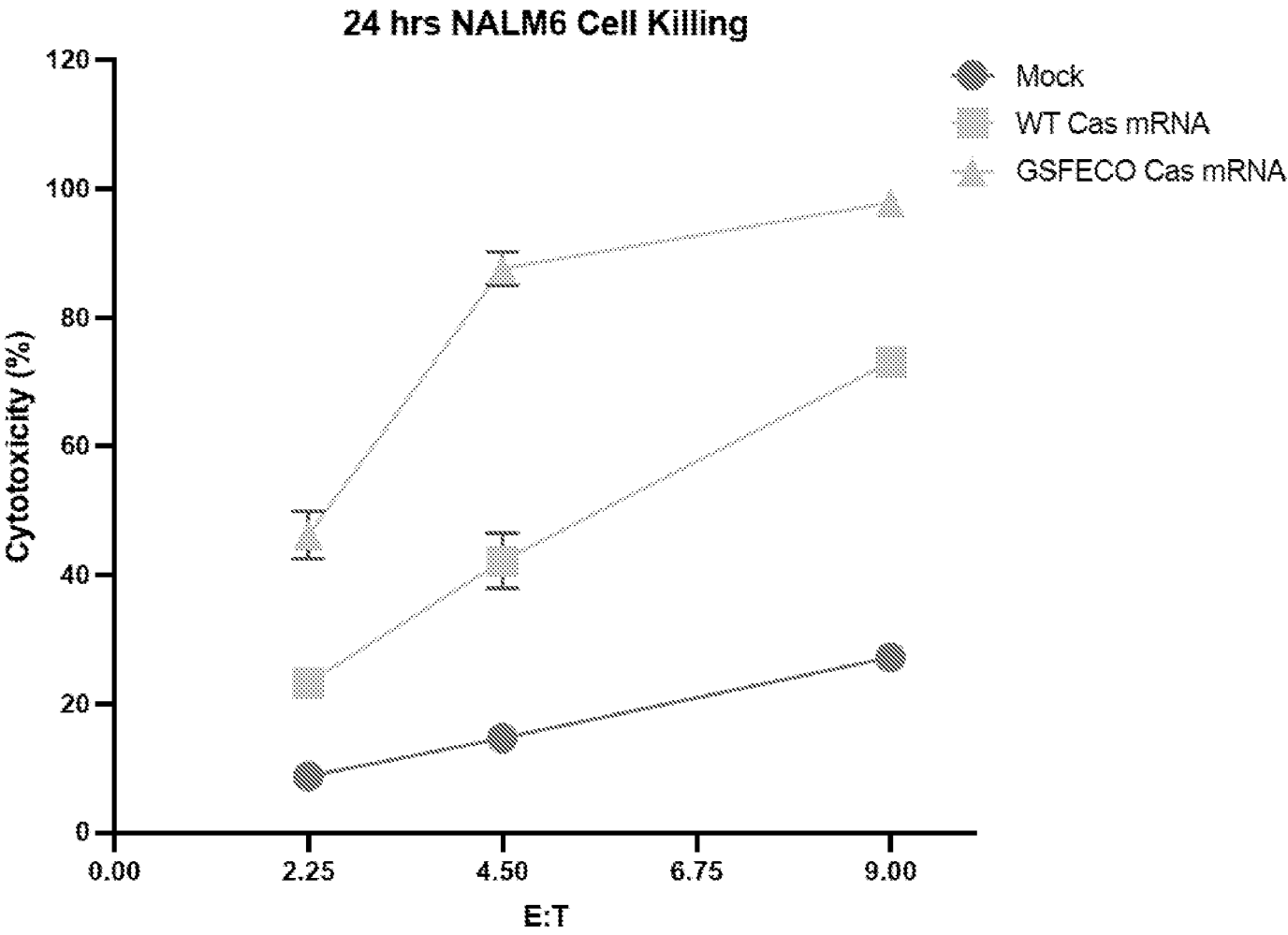


Fig. 6E

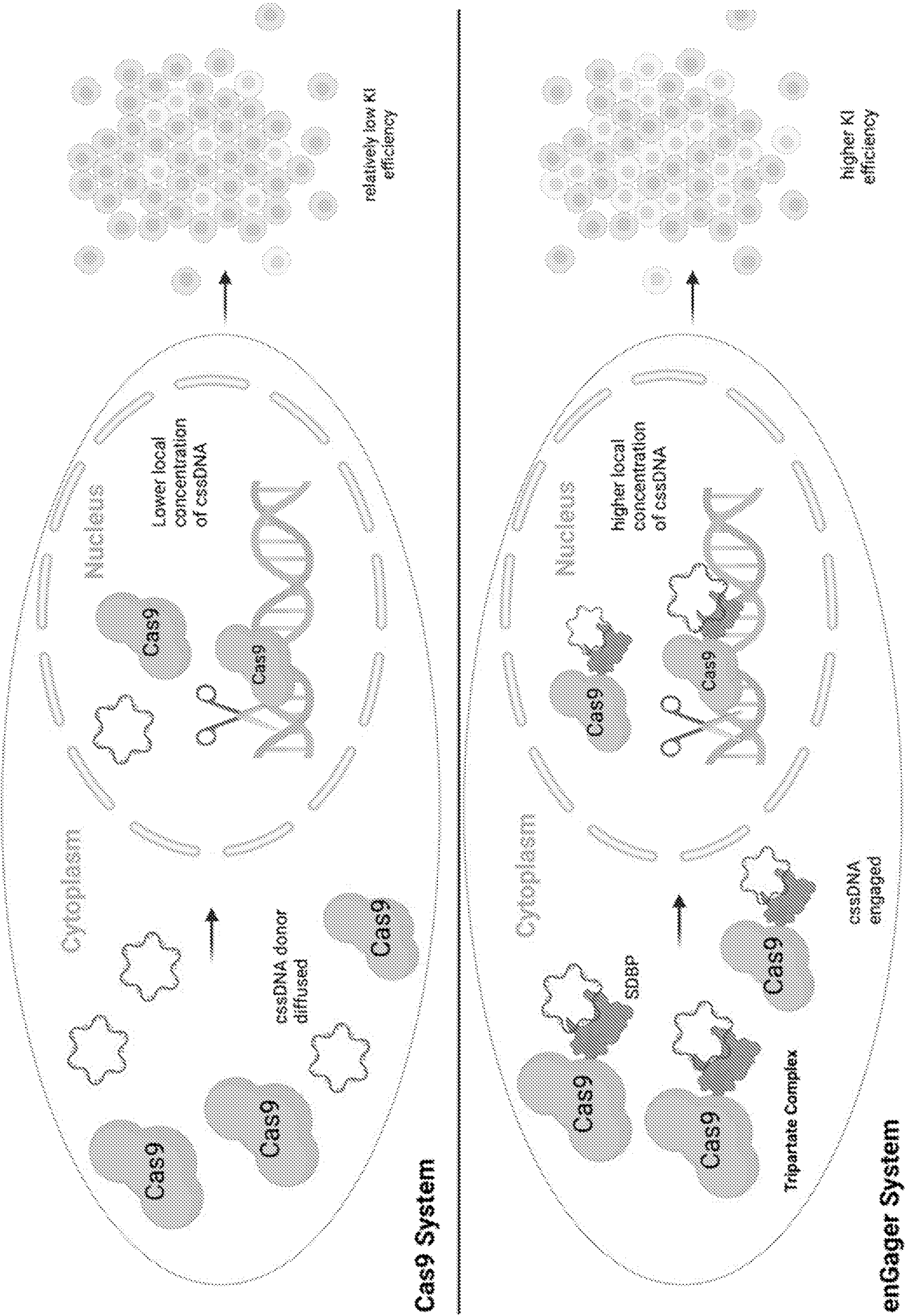


Fig. 6F

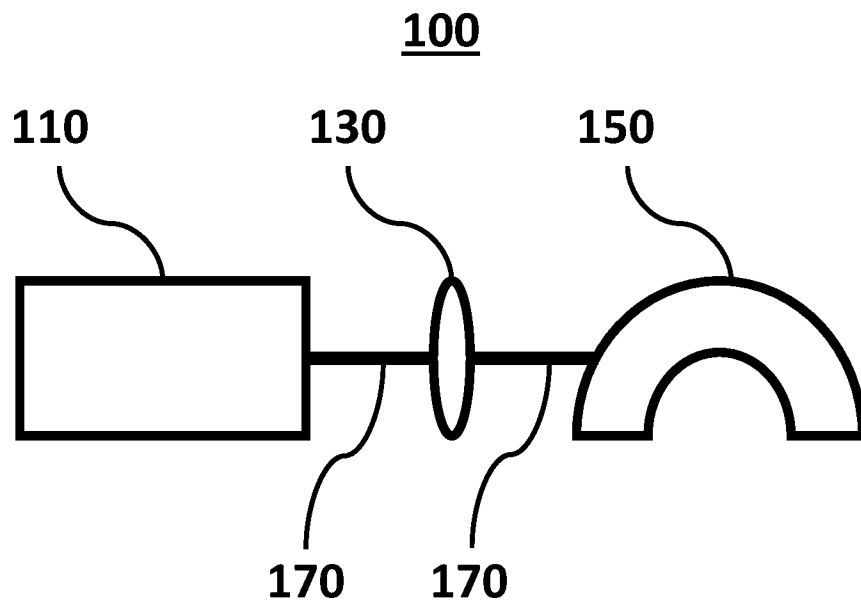


Fig. 7A

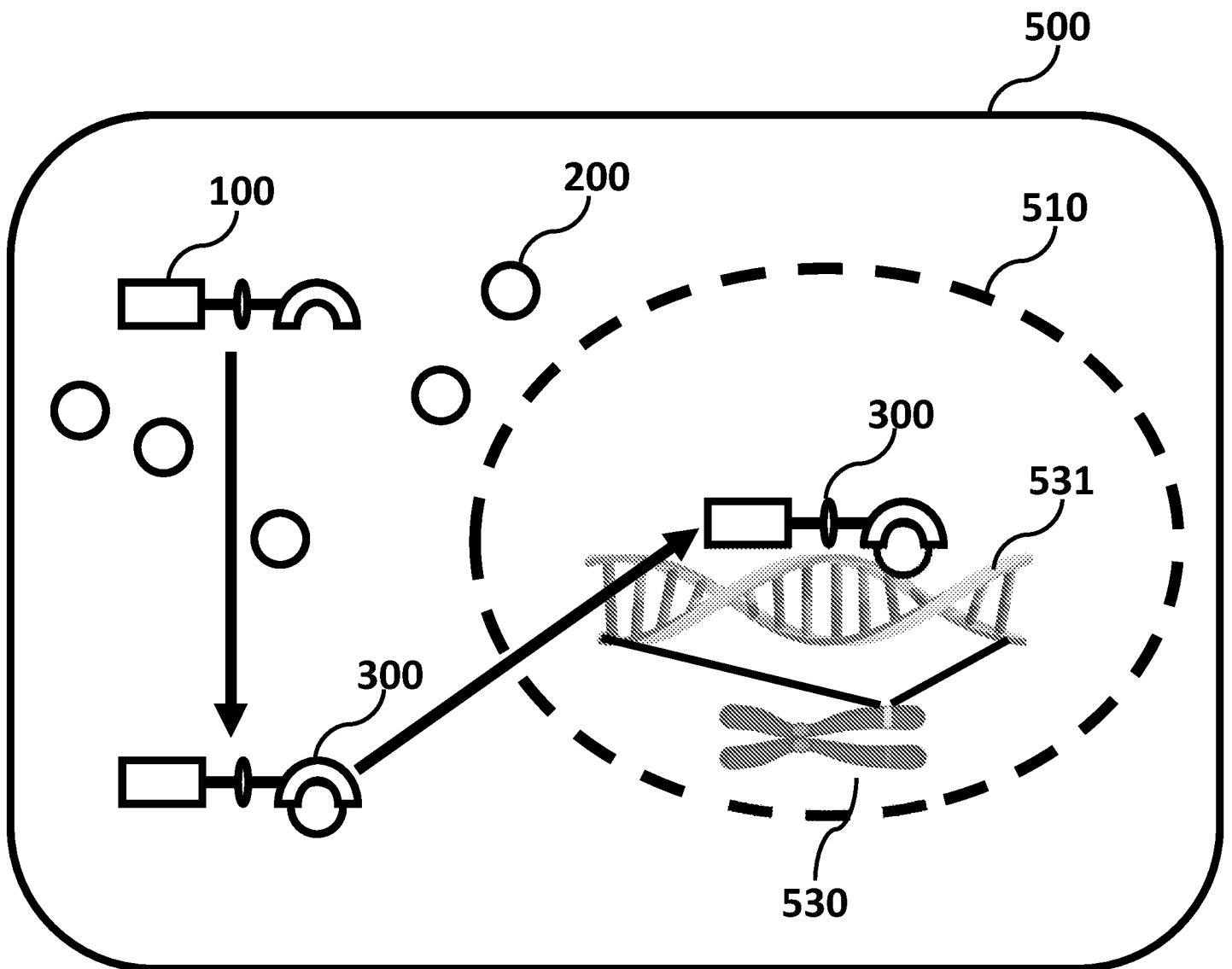


Fig. 7B

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2023/080718

A. CLASSIFICATION OF SUBJECT MATTERIPC: *A61K 9/51* (2024.01); *A61K 38/46* (2024.01); *C07K 14/315* (2024.01); *C07K 19/00* (2024.01); *C12N 9/22* (2024.01); *C12N 15/113* (2024.01); *C12N 15/62* (2024.01); *C12N 15/85* (2024.01)CPC: *C12N 9/22*; *A61K 9/51*; *C07K 14/315*; *A61K 38/465*; *C12N 15/113*; *C12N 15/62*; *C12N 15/85*; *C12N 15/907*; *C12N 2310/20*; *C07K 2319/80*; *C07K 2319/09*

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History Document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

See Search History Document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

See Search History Document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2021/113769 A1 (SCRIBE THERAPEUTICS INC) 10 June 2021 (10.06.2021) entire document	1-4
X	WO 1996/040765 A1 (THE GOVERNMENT OF THE UNITED STATES OF AMERICA REPRESENTED BY THE SECRETARY DEPARTMENT OF HEALTH AND HUMAN SERVICES) 19 December 1996 (19.12.1996) entire document	43, 44
A	US 2018/0171359 A1 (CARIBOU BIOSCIENCES INC.) 21 June 2018 (21.06.2018) entire document	1-4, 43, 44
A	WO 2018/035495 A1 (WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH) 22 February 2018 (22.02.2018) entire document	1-4, 43, 44



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“D” document cited by the applicant in the international application

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

06 February 2024 (06.02.2024)

Date of mailing of the international search report

26 February 2024 (26.02.2024)

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US
Commissioner for Patents
P.O. Box 1450, Alexandria, VA 22313-1450

Facsimile No. 571-273-8300

Authorized officer

MATOS
TAINA

Telephone No. 571-272-4300

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2023/080718**Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)**

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
 - a. ☒ forming part of the international application as filed.
 - b. ☐ furnished subsequent to the international filing date for the purposes of international search (Rule 13*ter*.1(a)),
☐ accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.
2. ☐ With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.
3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2023/080718

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: **5-42**
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).