



(51) International Patent Classification:

C07K 14/005 (2006.01) C12N 15/10 (2006.01)

C07K 14/47 (2006.01) C12N 15/11 (2006.01)

C12N 9/12 (2006.01) C12N 7/00 (2006.01)

C12N 9/22 (2006.01) C12N 9/78 (2006.01)

(21) International Application Number:

PCT/US2022/080834

(22) International Filing Date:

02 December 2022 (02.12.2022)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/285,995 03 December 2021 (03.12.2021) US

63/298,621 11 January 2022 (11.01.2022) US

(71) Applicants: **THE BROAD INSTITUTE, INC.** [US/US]; 415 Main Street, Cambridge, MA 02142 (US). **PRESIDENT AND FELLOWS OF HARVARD COLLEGE** [US/US]; 17 Quincy Street, Cambridge, MA 02138 (US).

(72) Inventors: **LIU, David, R.**; c/o The Broad Institute, Inc., 415 Main Street, Cambridge, MA 02142 (US). **RAGURAM, Aditya**; c/o The Broad Institute, Inc., 415 Main Street, Cambridge, MA 02142 (US). **BANSKOTA, Samagya**; c/o The Broad Institute, Inc., 415 Main Street, Cambridge, MA 02142 (US).

(74) Agent: **GODDARD, Tyler, N.** et al.; Wolf, Greenfield & Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210-2206 (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, 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, 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).

Published:

- without international search report and to be republished upon receipt of that report (Rule 48.2(g))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: SELF-ASSEMBLING VIRUS-LIKE PARTICLES FOR DELIVERY OF NUCLEIC ACID PROGRAMMABLE FUSION PROTEINS AND METHODS OF MAKING AND USING SAME

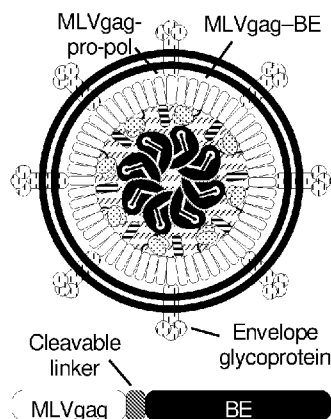


FIG. 1A

(57) Abstract: The present disclosure provides virus-like particles for delivering gene editing agents such as nucleic acid-programmable DNA-binding proteins (napDNAbps) and base editor fusion proteins ("BE-VLPs" or "eVLPs"), and systems comprising such eVLPs. The present disclosure also provides polynucleotides encoding the eVLPs described herein, which may be useful for producing said eVLPs. Also provided herein are methods for editing the genome of a target cell by introducing the presently described eVLPs into the target cell. The present disclosure also provides fusion proteins that make up a component of the eVLPs described herein, as well as polynucleotides, vectors, cells, and kits.

SELF-ASSEMBLING VIRUS-LIKE PARTICLES FOR DELIVERY OF NUCLEIC ACID PROGRAMMABLE FUSION PROTEINS AND METHODS OF MAKING AND USING SAME

RELATED APPLICATIONS

[0001] This application claims priority under 35 U.S.C. § 119(e) to U.S. Provisional Application, U.S.S.N. 63/285,995, filed December 3, 2021, and U.S. Provisional Application, U.S.S.N. 63/298,621, filed January 11, 2022, each of which is incorporated herein by reference.

FEDERALLY SPONSORED RESEARCH

[0002] This invention was made with government support under Grant Nos. UG3AI150551, U01AI142756, R35GM118062, RM1HG009490, R01EY009339, and T32GM095450 awarded by the National Institutes of Health. The government has certain rights in the invention.

BACKGROUND OF THE INVENTION

[0003] Recently developed gene editing agents enable the precise manipulation of genomic DNA in living organisms and raise the possibility of treating the root cause of many genetic diseases (Anzalone *et al.*, 2020; Doudna, 2020). Base editors (BEs) mediate targeted single-nucleotide conversions without requiring double-stranded DNA breaks (DSBs), and thereby minimize undesired consequences of editing such as indels, large deletions (Kosicki *et al.*, 2018; Song *et al.*, 2020), translocations (Giannoukos *et al.*, 2018; Stadtmauer *et al.*, 2020; Webber *et al.*, 2019), chromothripsis (Leibowitz *et al.*, 2021), or other chromosomal abnormalities. Cytosine base editors (CBEs) (Komor *et al.*, 2016; Nishida *et al.*, 2016) and adenine base editors (ABEs) (Gaudelli *et al.*, 2017) in principle can together correct the majority of known disease-causing single-nucleotide variants (Anzalone *et al.*, 2020; Rees and Liu, 2018). Previously, BEs have been applied to correct pathogenic point mutations and rescue disease phenotypes in mice and non-human primates (Levy *et al.*, 2020; Yeh *et al.*, 2020), highlighting the potential of *in vivo* base editing as a therapeutic strategy.

[0004] The broad therapeutic application of *in vivo* base editing requires safe and efficient methods for delivering BEs to multiple tissues and organs. The most robust approaches for delivering BEs *in vivo* reported to date involve the use of viruses, such as adeno-associated viruses (AAVs) or lentivirus (LV), to deliver BE-encoding DNA to target tissues (Levy *et al.*,

2020; Newby and Liu, 2021). However, viral delivery of DNA encoding editing agents leads to prolonged expression in transduced cells, which increases the frequency of off-target editing (Akcakaya *et al.*, 2018; Davis *et al.*, 2015; Wang *et al.*, 2020; Yeh *et al.*, 2018). In addition, viral delivery of DNA raises the possibility of viral vector integration into the genome of transduced cells, both of which can promote oncogenesis or other adverse effects (Anzalone *et al.*, 2020; Chandler *et al.*, 2017). Further, in spite of the constant evolution of transfection methods and performances of viral delivery vectors (*e.g.*, AAV or LV), the efficiency of these approaches can vary dramatically, especially in primary cells that are highly sensitive to modifications of their environment and may be altered in response to transfection agents and/or vectors.

[0005] One alternate method for delivering gene editing agents (*e.g.*, BEs) *in vivo* would be to directly deliver proteins (*e.g.*, a BE) or ribonucleoproteins (RNPs) (*e.g.*, a BE complexed with a guide RNA) instead of DNA. The short lifespan of RNPs in cells limits opportunities for off-target editing, as demonstrated by previous reports that delivering BE RNPs instead of BE-encoding DNA or mRNA leads to substantially reduced off-target editing, typically without sacrificing on-target editing efficiency (Doman *et al.*, 2020; Rees *et al.*, 2017). While successful base editing has previously been reported in the mouse inner ear and retina following local administration of lipid-encapsulated BE RNPs (Yeh *et al.*, 2018), no generalizable strategy for delivering BE RNPs to multiple tissues and organs *in vivo* has been reported previously. Accordingly, there is a need for a system/method that effectively delivers BE ribonucleoproteins (RNPs) into cells, tissues, or organs of subjects in need thereof, and in a manner which improves the overall safety by limiting and/or avoiding off-target editing without sacrificing target edits.

SUMMARY OF THE INVENTION

[0006] Virus-like particles (VLPs), assemblies of viral proteins that can infect cells but lack viral genetic material, have emerged as potentially promising vehicles for delivering gene editing agents as ribonucleoproteins (RNPs) (Campbell *et al.*, 2019; Choi *et al.*, 2016; Gee *et al.*, 2020; Hamilton *et al.*, 2021; Indikova and Indik, 2020; Lyu *et al.*, 2019; Lyu *et al.*, 2021; Mangeot *et al.*, 2019; Yao *et al.*, 2021). VLPs that deliver RNP cargos exploit the efficiency and tissue targeting advantages of viral delivery but avoid the risks associated with viral genome integration and prolonged expression of the editing agent. However, existing VLP-mediated strategies for delivering gene editing agent RNPs thus far support low to moderate

editing efficiencies or limited validation of their therapeutic efficacy *in vivo* (Campbell *et al.*, 2019; Choi *et al.*, 2016; Gee *et al.*, 2020; Hamilton *et al.*, 2021; Indikova and Indik, 2020; Lyu *et al.*, 2019; Lyu *et al.*, 2021; Mangeot *et al.*, 2019; Yao *et al.*, 2021). Indeed, therapeutic levels of post-natal *in vivo* gene editing using RNP-packaging VLPs have not been previously reported.

[0007] The present disclosure is based on the development and application of engineered virus-like particles (referred to herein as either “VLPs” or “eVLPs” interchangeably) for packaging and delivering therapeutic RNPs, including Cas9 and base editors (or “BEs” as disclosed herein), *in vitro* and *in vivo* that offer key advantages of both viral and non-viral delivery strategies. In various embodiments, extensive VLP architecture engineering of initial designs that were based on previously reported VLPs (Mangeot *et al.*, "Genome editing in primary cells and in vivo using viral-derived Nanoblades loaded with Cas9-sgRNA ribonucleoproteins," *Nature Communications*, 2019) yielded first, second, third, and fourth generation eVLPs capable of delivering ribonucleoproteins, such as Cas9 and BEs complexed with sgRNAs, to cells, tissue, or subjects. By iteratively engineering VLP architectures to overcome cargo packaging, release, and localization bottlenecks, optimized eVLPs were generated that mediate efficient on-target base editing *in vitro* across a variety of cell types and endogenous genomic loci with minimal detected off-target editing, as well as higher editing efficiencies of eVLP-delivered BE cargoes.

[0008] As described in various embodiments in the Examples, such eVLPs enable highly efficient base editing with minimal off-target editing in a variety of cell types, including multiple immortalized cell lines, primary human and mouse fibroblasts, and primary human T cells, as well as 4.7-fold improved Cas9 nuclease-mediated indel formation compared with a previously reported Cas9-VLP. Exemplary applications of use of the presently described BE-VLPs show in the Examples that single *in vivo* injections of eVLPs into mice mediated efficient base editing of various target genes in multiple organs, strongly knocked down serum Pcsk9 levels, and partially restored visual function in a mouse model of genetic blindness. The present disclosure, including the Examples, establish eVLPs as a useful platform for transiently delivering gene editing agents (*e.g.*, Cas9 or BE ribonucleoproteins) *in vitro* and *in vivo* with therapeutically relevant efficiencies and with minimized risk of off-target editing or DNA integration and similarly improves the *in vivo* delivery of other proteins and RNPs.

[0009] In various embodiments, the eVLPs (*e.g.*, BE-VLPs) comprise a supra-molecular assembly comprising (a) an envelope comprising (i) a lipid membrane (*e.g.*, single-layer or bi-layer membrane) and (ii) a viral envelope glycoprotein, and (b) a multi-protein core region enclosed by the envelope and comprising (i) a Gag protein, (ii) a Gag-Pro-Pol protein, and (iii) a Gag-cargo fusion protein comprising a Gag protein fused to a cargo protein (*e.g.*, a napDNAbp, such as Cas9, or BE) via a cleavable linker (*e.g.*, a protease-cleavable linker, *e.g.*, an MMLV protease-cleavable linker). In various embodiments, the cargo protein is a napDNAbp (*e.g.*, Cas9). In other embodiments, the cargo protein is a base editor. In various other embodiments, the multi-protein core region of the VLPs further comprises one or more guide RNA molecules which are complexed with the napDNAbp or the base editor to form a ribonucleoprotein (RNP). In various embodiments, the VLPs are prepared in a producer cell that is transiently transformed with plasmid DNA that encodes the various protein and nucleic acid (sgRNA) components of the VLPs. Without being bound by theory, the components self-assemble at the cell membrane and bud out in accordance with the naturally occurring mechanism of budding (*e.g.*, retroviral budding or the budding mechanism of other envelope viruses) in order to release from the cell fully-matured VLPs. Once formed, the Gag-Pol-Pro cleaves the protease-sensitive linker of the Gag-cargo (*i.e.*, [Gag]-[cleavable linker]-[cargo], wherein the cargo can be BE-RNP or a napDNAbp RNP), thereby releasing the BE RNP and/or napDNAbp RNA, as the case may be, within the VLP. Thus, in various embodiments, the present disclosure also provides VLPs in which the protease-sensitive linker has been cleaved (*e.g.*, producing two cleavage products comprising (i) a fusion protein comprising a gag nucleocapsid protein and a nuclear export sequence, and (ii) a napDNAbp, which may be fused to additional domains such as one or more NLS and/or a deaminase (*i.e.*, to form a base editor)). For example, the present disclosure provides VLPs comprising a group-specific antigen (gag) protease (pro) polyprotein, a nucleic acid programmable DNA binding protein (napDNAbp), and a fusion protein comprising a gag nucleocapsid protein and a nuclear export sequence (NES), encapsulated by a lipid membrane and a viral envelope glycoprotein. In some embodiments, the present disclosure provides VLPs comprising a mixture of cleaved and uncleaved products (*i.e.*, some of the napDNAbps or BEs have been cleaved from the gag proteins and are free, while some have not yet been cleaved from the gag proteins). In some embodiments, more than 50%, more than 60%, more than 70%, more than 80%, or more than 90% of the napDNAbp or BE has been cleaved from the gag protein inside the VLP. Once the VLP is administered to a recipient cell and taken up by said

recipient cell, the contents of the VLP are released, *e.g.*, released BE RNP and/or napDNAbp RNP. Once in the cell, the RNPs may translocate to the nucleus of the cell (in particular, where NLSs are included as part the RNPs), where DNA editing, cleavage, or other modification may occur at target site(s) specified by the guide RNA. The present disclosure also provides polynucleotides and vectors encoding various components of the VLPs described herein.

[0010] In another aspect, the present disclosure provides compositions (*e.g.*, pharmaceutical compositions) comprising a virus-like particle (VLP) comprising a group-specific antigen (gag) protease (pro) polyprotein and a fusion protein encapsulated by a viral envelope glycoprotein, wherein the fusion protein comprises: (i) a gag nucleocapsid protein; (ii) a nucleic acid programmable DNA binding protein (napDNAbp); (iii) a cleavable linker; and (iv) a nuclear export sequence (NES). In some embodiments, the napDNAbp is fused to one or more additional domains such as one or more NLS and/or one or more deaminase (*i.e.*, to form a base editor). In some embodiments, the pharmaceutical composition comprises a VLP comprising a group-specific antigen (gag) protease (pro) polyprotein, a nucleic acid programmable DNA binding protein (napDNAbp), and a fusion protein comprising a gag nucleocapsid protein and a nuclear export sequence (NES), encapsulated by a lipid membrane and a viral envelope glycoprotein (*i.e.*, a VLP in which the cleavable linker has been cleaved by a protease). In some embodiments, the napDNAbp is fused to one or more additional domains such as one or more NLS and/or one or more deaminase (*i.e.*, to form a base editor). Each component of the pharmaceutical compositions provided herein may comprise any of the options described above in reference to the VLPs, or any of the other options provided by the present disclosure. In some embodiments, a pharmaceutical composition further comprises a pharmaceutically acceptable excipient.

[0011] In another aspect, the present disclosure provides methods for editing a nucleic acid molecule in a target cell by base editing comprising contacting the target cell with any of the compositions provided herein, thereby installing one or more modifications to the nucleic acid molecule at a target site. In some embodiments, the cell is a mammalian cell (*e.g.*, a human cell). In some embodiments, the cell is a cell from an animal relevant for veterinary or agricultural use. In some embodiments, the cell is in a subject. In certain embodiments, the subject is a human. In some embodiments, the one or more modifications to the nucleic acid molecule are associated with reducing, relieving, or preventing the symptoms of a disease or disorder.

[0012] In another aspect, the present disclosure provides fusion proteins comprising: (i) a group-specific antigen (gag) nucleocapsid protein; (ii) a nucleic acid programmable DNA binding protein (napDNAbp); (iii) a cleavable linker; and (iv) a nuclear export sequence (NES). Each component of the fusion proteins provided herein may comprise any of the options described herein in reference to the BE-VLPs, or any of the other options provided by the present disclosure. In other aspects, the present disclosure also provides polynucleotides encoding any of the eVLP components, including the fusion proteins provided herein, vectors comprising such polynucleotides, cells comprising any of the eVLP proteins, including fusion proteins, polynucleotides, or vectors provided herein, and kits comprising any of the pluralities of polynucleotides or eVLP proteins, including fusion proteins, provided herein.

[0013] In another aspect, the present disclosure provides VLPs produced by transfecting, transducing, electroporating, or otherwise inserting any of the polynucleotides or vectors disclosed herein into a cell and expressing the components of the VLPs from the polynucleotides or vectors, thereby allowing the virus-like particle to spontaneously assemble in the cell. In some embodiments, any of the compositions, methods, or cells provided herein may be used to produce the VLPs described herein.

[0014] In another aspect, the present disclosure provides compositions comprising any of the VLPs, polynucleotides, vectors, and fusion proteins provided herein.

[0015] In another aspect, the present disclosure provides methods of editing a nucleic acid molecule in a target cell using any of the VLPs, polynucleotides, compositions, and fusion proteins provided herein.

[0016] In another aspect, the present disclosure provides cells comprising any of the VLPs, polynucleotides, vectors, compositions, and fusion proteins described herein.

[0017] In another aspect, the present disclosure provides kits comprising any of the VLPs, polynucleotides, vectors, compositions, and fusion proteins described herein.

[0018] It should be appreciated that the foregoing concepts, and additional concepts discussed below, may be arranged in any suitable combination, as the present disclosure is not limited in this respect. Further, other advantages and novel features of the present disclosure will become apparent from the following detailed description of various non-limiting embodiments when considered in conjunction with the accompanying figures.

BRIEF DESCRIPTION OF THE DRAWINGS

[0019] The following drawings form part of the present specification and are included to further demonstrate certain aspects of the present disclosure, which can be better understood by reference to one or more of these drawings in combination with the detailed description of specific embodiments presented herein.

[0020] **FIGS. 1A-1D:** BE-VLP architecture and initial (v1) editing efficiencies. **FIG. 1A:**

Schematic of BE-VLPs. Base editor protein is fused to the C-terminus of murine leukemia virus (MLV) gag polyprotein via a linker that is cleaved by the MLV protease upon particle maturation. BE = base editor. **FIG. 1B:** Adenine base editing efficiencies of v1 BE-VLPs at two genomic loci in HEK293T cells. The protospacer positions of the target adenines are denoted by subscripts (*i.e.*, A₅ = adenine at position 5), where the PAM is positions 21-23. Data are shown as individual data points and mean $\pm$ s.e.m for $n = 3$ independent biological replicates. **FIG. 1C** provides a generalized structure for the virus-like particles contemplated herein, which includes (a) a lipid membrane which is derived from the cell membrane of the producer cell as a result of the retroviral budding process, (b) a viral envelope glycoprotein (which facilitates binding to a recipient cell and effects of tropism), and (c) a protein core or shell comprising an assembly of proteins comprising retroviral Gag proteins, wherein a portion of the Gag proteins are fused to a cleavable protein cargo (*e.g.*, a napDNAbp or BE) or Pro-Pol (comprising a protease activity). The cleavable protein cargo is joined to the Gag protein by a protease-cleavable linker and becomes cleaved by Pro-Pol at some point following the assembly of the VLP. As background, **FIG. 1D** provides a schematic depicting the budding out process of a typical retrovirus and the involvement of the Gag polyprotein, which includes the “MA” domain (matrix domain), the “CA” domain (capsid domain), and the “NC” domain (nucleocapsid domain). Without being bound by theory, it is believed that the Gag, Gag-Pro-Pol, and Gag-cargo fusions of the eVLPs described herein drive a similar budding out process to form the mature eVLPs which are released from the producer cells.

[0021] **FIGS. 2A-2G:** Optimization of BE-VLPs (identifying and engineering solutions to bottlenecks that limit VLP potency results in v2, v3, and v4 eVLPs). **FIG. 2A:** More efficient linker cleavage leads to improved cargo release after VLP maturation. **FIG. 2B:** Adenine base editing efficiencies of v1 and v2 BE-eVLPs at position A7 of the BCL11A enhancer site in HEK293T cells. Optimization of protease-cleavable linker sequence is shown (see also **FIG. 8**). **FIG. 2C:** Improved localization of cargo in producer cells leads to more efficient incorporation into eVLPs. **FIG. 2D:** Installing a 3xNES motif upstream of the cleavable

linker encourages cytoplasmic localization of gag–3xNES–cargo in producer cells but nuclear localization of free ABE cargo in transduced cells. **FIG. 2E:** Optimization of gag–ABE localization (see also **FIGS. 9A-9B**). Adenine base editing efficiencies of v2.4 and v3 BE-eVLPs at position A₇ of the BCL11A enhancer site in HEK293T cells. **FIG. 2F:** The optimal gag–cargo:gag–pro–pol stoichiometry balances the amount of cargo protein per particle with the amount of MMLV protease required for efficient particle maturation. **FIG. 2G:** Optimization of gag–ABE:gag–pro–pol ratio. Adenine base editing efficiencies of v3.4 eVLPs with different gag–ABE:gag–pro–pol stoichiometries at position A₇ of the BCL11A enhancer site in HEK293T cells. Legend denotes % gag–ABE plasmid of the total amount of gag–ABE and gag–pro–pol plasmids. **FIGS. 2B, 2E, and 2G:** Values and error bars reflect mean $\pm$ s.e.m. of $n = 3$ independent biological replicates. Data were fit to 4-parameter logistic curves using nonlinear regression.

[0022] FIGS. 3A-3J: Characterization of BE-eVLPs. **FIG. 3A:** Quantification of BE molecules per eVLP by anti-Cas9 and anti-MLV (p30) ELISA (see also **FIGS. 10A-10C**). Values and error bars reflect mean $\pm$ s.e.m. of $n=3$ independent replicates. **FIG. 3B:** Quantification of relative sgRNA abundance by RT-qPCR using sgRNA-specific primers, normalized relative to v1 sgRNA abundance. Values and error bars reflect mean $\pm$ s.e.m. of $n=3$ technical replicates. **FIGS. 3C-3D:** Comparison of editing efficiencies with v1, v2.4, v3.4, and v4 BE-eVLPs at the *BCL11A* enhancer site in HEK293T cells (**FIG. 3C**) and at the *Dnmt1* site in NIH 3T3 cells (**FIG. 3D**). Values and error bars reflect mean $\pm$ s.e.m. of $n=3$ independent biological replicates. Data were fit to four-parameter logistic curves using nonlinear regression. **FIG. 3E:** Adenine base editing efficiencies in HEK293T cells of single BE-eVLPs targeting either the *HEK2* or *BCL11A* enhancer loci separately or multiplex v4 BE-eVLPs targeting both loci simultaneously. Data are shown as individual data points and mean $\pm$ s.e.m for $n=3$ independent biological replicates. **FIG. 3F:** Adenine base editing efficiencies of FuG-B2-pseudotyped v4 BE-eVLPs in Neuro-2a cells or 3T3 fibroblasts. Data are shown as individual data points and mean $\pm$ s.e.m for $n=3$ independent biological replicates. **FIG. 3G:** Adenine base editing efficiencies at three on-target genomic loci and their corresponding Cas-dependent off-target sites in HEK293T cells treated with v4 BE-eVLPs or ABE8e plasmid. OT1 = off-target site 1, OT2 = off-target site 2, OT3 = off-target site 3. **FIG. 3H:** Cas-independent off-target editing frequencies at six off-target R-loops in HEK293T cells treated with v4 BE-eVLPs or ABE8e plasmid. OTRL = off-target R-loop. (see also FIG. 11A for the experimental timeline, and FIG. 11B for on-target editing

controls). **FIG. 3I:** Molecules of BE-encoding DNA per v4 BE-eVLP detected by qPCR of lysed VLPs or lysis buffer only. **FIG. 3J:** Amount of BE-encoding DNA detected by qPCR of lysate from cells that were either treated with BE-VLPs or transfected with BE-encoding plasmids. **FIGS. 3E-3J:** Data are shown as individual data points and mean $\pm$ s.e.m. for $n = 3$ independent biological replicates.

[0023] FIGS. 4A-4C: Base editing in primary human and mouse cells using v4 BE-eVLPs.

FIG. 4A: Correction efficiencies of the *COL7A1*(R185X) mutation in patient-derived primary human fibroblasts. Genomic DNA was harvested from cells 48 h post transduction with v4 BE-VLPs. Values and error bars reflect mean $\pm$ s.e.m. of $n=3$ independent biological replicates. Data were fit to four-parameter logistic curves using nonlinear regression. **FIG. 4B:** Correction efficiencies of the *Idua*(W392X) mutation in primary mouse fibroblasts. Genomic DNA was harvested from cells 48 h post transduction with v4 BE-VLPs. Values and error bars reflect mean $\pm$ s.e.m. of $n=3$ independent biological replicates. Data were fit to four-parameter logistic curves using nonlinear regression. **FIG. 4C:** Adenine base editing efficiencies at the *B2M* and *CIITA* loci in primary human T cells. T cells were transduced twice with v4 BE-VLPs, and genomic DNA was harvested from cells 48 h after the second transduction (see Examples). Data are shown as individual data points and mean $\pm$ s.e.m for $n = 3$ independent biological replicates.

[0024] FIGS. 5A-5B: *In vivo* base editing in the central nervous system using v4 BE-eVLPs.

FIG. 5A: Schematic of P0 ICV injections of v4 BE-eVLPs. *Dnmt1*-targeting v4 BE-eVLPs were co-injected with a lentivirus encoding EGFP-KASH. Tissue was harvested 3 weeks post-injection, and cortex and mid-brain were separated. Nuclei were dissociated for each tissue and analyzed by high-throughput sequencing as bulk unsorted (all nuclei) or GFP+ nuclei. **FIG. 5B:** Adenine base editing efficiencies at the *Dnmt1* locus in bulk unsorted (all nuclei) and GFP+ populations. Data are shown as individual data points and mean $\pm$ s.e.m for $n = 4$ mice.

[0025] FIGS. 6A-6E: *In vivo* knockdown of *Pcsk9* from a single systemic injection of v4 BE-eVLPs.

FIG. 6A: Schematic of systemic injections of BE-eVLPs. *Pcsk9*-targeting BE-eVLPs were injected retro-orbitally into 6- to 7-week-old C57BL/6J mice. Organs were harvested one week after injection, and the genomic DNA of unsorted cells was sequenced. **FIG. 6B:** Adenine base editing efficiencies at the *Pcsk9* exon 1 splice donor in the mouse liver after systemic injection of v1 BE-VLPs or v4 BE-eVLPs. Data are shown as individual data points and mean $\pm$ s.e.m for $n=3$ mice (v1 BE-VLP and v4 BE-eVLP at 4×10^{11} VLPs) or $n=4$ mice

(v4 BE-eVLP at 7×10^{11} eVLPs). **FIG. 6C:** Adenine base editing efficiencies at the *Pcsk9* exon 1 splice donor in the mouse heart, kidney, liver, lungs, muscle, and spleen after systemic injection of 7×10^{11} v4 BE-eVLPs. Data are shown as individual data points and mean $\pm$ s.e.m for $n=4$ mice (treated) or $n=3$ mice (untreated). **FIG. 6D:** DNA sequencing reads containing A•T-to-G•C mutations within protospacer positions 4–10 for the fourteen CIRCLE-seq-nominated off-target loci from the livers of v4 BE-eVLP-treated, AAV-treated, and untreated mice. Data are shown as individual data points and mean $\pm$ s.e.m for $n=4$ mice (BE-eVLP), $n=5$ mice (AAV), or $n=3$ mice (untreated). vg = viral genomes. **FIG. 6E:** Serum *Pcsk9* levels as measured by ELISA. Data are shown as individual data points and mean $\pm$ s.e.m for $n=4$ mice (treated) or $n=3$ mice (untreated).

[0026] FIGs. 7A-7J: *In vivo* base editing by v4 BE-eVLPs in a mouse model of genetic blindness. **FIG. 7A:** Schematic of *Rpe65* exon 3 surrounding the R44X mutation (in gray and italicized under the label “R44X”), which can be corrected by an A•T-to-G•C conversion at position A6 in the protospacer (shaded grey, PAM underlined). Sequences shown are SEQ ID NO: 497 (top) and SEQ ID NO: 498 (bottom). **FIG. 7B:** Schematic of subretinal injections. Five weeks post-injection, phenotypic rescue was assessed via electroretinogram (ERG), and tissues were subsequently harvested for sequencing. **FIG. 7C:** Adenine base editing efficiencies at positions A₃, A₆, and A₈ of the protospacer in genomic DNA harvested from rd12 mice. Data are shown as individual data points and mean $\pm$ s.e.m for $n = 6$ mice (both treated groups) or $n = 4$ mice (untreated). **FIG. 7D:** Allele frequency distributions of genomic DNA harvested from treated rd12 mice. Data are shown as mean $\pm$ s.e.m for $n = 6$ mice. 8e-LV = ABE8e-NG-LV, 8e-eVLP = v4 ABE8e-NG-eVLP. **FIG. 7E:** Scotopic a-wave and b-wave amplitudes measured by ERG following overnight dark adaptation. Data are shown as individual data points and mean $\pm$ s.e.m for $n = 8$ mice (wild-type), $n = 6$ mice (ABE8e-NG-LV and v4 ABE8e-NG-eVLP) or $n = 4$ mice (untreated). **FIG. 7F:** Adenine base editing efficiencies at positions A₃, A₆, and A₈ of the protospacer in genomic DNA harvested from rd12 mice. Data are shown as individual data points and mean $\pm$ s.e.m for $n = 6$ mice (v4 ABE7.10-NG-eVLP) or $n = 4$ mice (ABE7.10-NG-LV and untreated). P values were calculated using a two-sided t-test. **FIG. 7G:** Allele frequency distributions of genomic DNA harvested from treated rd12 mice. Data are shown as mean $\pm$ s.e.m for $n = 6$ mice (v4 ABE7.10-NG-eVLP) or $n = 4$ mice (ABE7.10-NG-LV and untreated). 7.10-LV = ABE7.10-NG-LV, 7.10-eVLP = v4 ABE7.10-NG-eVLP. **FIG. 7H:** Scotopic a-wave and b-wave amplitudes measured by ERG following overnight dark adaptation. Data are shown as

individual data points and mean $\pm$ s.e.m for n=8 mice (wild-type), n=7 mice (v4 ABE7.10-NG-eVLP), n=5 mice (ABE7.10-NG-LV), or n=4 mice (untreated). P values were calculated using a two-sided t-test. **FIG. 7I:** Western blot of protein extracts from RPE tissues of wild-type, untreated, v4 ABE7.10-NG-eVLP-treated, and ABE7.10-NG-LV-treated mice. **FIG. 7J:** Representative ERG waveforms from wild-type, untreated, ABE7.10-NG-LV-treated, and v4 ABE7.10-NG-eVLP-treated mice.

[0027] **FIGS. 8A-8E:** Engineering and characterization of v1 BE-VLPs and v2 BE-eVLPs.

FIG. 8A: Validation of VLP production. Immunoblot analysis of proteins from purified BE-VLPs using anti-Cas9, anti-p30, and anti-VSV-G antibodies. **FIG. 8B:** Adenine base editing efficiencies of v1 BE-VLPs at position A₇ of the *BCL11A* enhancer site in HEK293T cells.

Values and error bars reflect mean $\pm$ s.e.m. of n=3 independent biological replicates. Data were fit to four-parameter logistic curves using nonlinear regression. **FIG. 8C:** Schematic of an immature BE-VLP with ABE8e fused to the gag structural protein. Various MMLV protease cleavage sites were inserted between the gag and ABE8e to determine the optimal cleavable sequence that promotes liberation of ABE8e from the gag during proteolytic virion maturation. Arrows indicate the cleavage site. Sequences shown are PRSSLY (SEQ ID NO: 499), PALTP (SEQ ID NO: 500), VQAL (SEQ ID NO: 501), VLTQ (SEQ ID NO: 502), PLQVL (SEQ ID NO: 503), TLNIERR (SEQ ID NO: 504), TSTLL (SEQ ID NO: 505), and MENSS (SEQ ID NO: 506). **FIG. 8D:** Representative western blot evaluating cleaved ABE8e versus full-length gag-ABE8e in purified v2 BE-VLPs variants. **FIG. 8E:** Densitometry-based quantification of the cleaved ABE8e fraction from western blots. Data are shown as mean values $\pm$ s.e.m. for n=3 technical replicates.

[0028] **FIGS. 9A-9D:** Improving gag-ABE localization in producer cells. **FIG. 9A:**

Schematic showing the localization of BE-RNP cargo in the producer cells with (right) and without (left) nuclear exclusion signal (NES). **FIG. 9B:** v2.4 and v3 BE-eVLP constructs.

Three HIV NESs were fused to either the C-terminus or N-terminus of the gag-ABE fusion.

A protease cleavable linker was incorporated between ABE and the NES sequences such that the final BE cargo will be devoid of NESs following proteolytic virion maturation. Protease cleavage sequences shown are TSTLL (SEQ ID NO: 505), MENSS (SEQ ID NO: 506), MSKLL (SEQ ID NO: 507), ATVVS (SEQ ID NO: 508), PLQVL (SEQ ID NO: 503), TLNIERR (SEQ ID NO: 504), IRKIL (SEQ ID NO: 509), and FLDG (SEQ ID NO: 510).

FIG. 9C: Representative immunofluorescence image of producer cells transfected with the v2.4 gag-ABE construct or the v3.4 gag-3xNES-ABE construct. After 48 h post-

transfection, cells were fixed in paraformaldehyde and stained with anti-tubulin antibody to stain the cytoskeleton, DAPI for nuclei staining, and anti-Cas9 antibody to visualize gag-ABE fusion, as shown in the legend provided. Scale bars denote 50 μ m. **FIG. 9D:** Automated image analysis-based quantification of cytoplasmic localization of the v2.4 gag-ABE construct or the v3.4 gag-3xNES-ABE construct. Data are shown as mean values $\pm$ s.e.m. for n=3 technical replicates. P values were calculated using a two-sided t-test.

[0029] FIGs. 10A-10G: Characterization of BE-eVLPs. **FIG. 10A:** Representative negative-stain transmission electron micrograph (TEM) of v4 BE-eVLPs. Scale bar denotes 200 nm. **FIGS. 10B-10C:** Protein content for v1, v2.4, v3.4, and v4 BE-eVLPs was measured by anti-Cas9 or anti-MLV(p30) ELISA. Data are shown as individual data points and mean values $\pm$ s.e.m. for n=3 technical replicates. **FIG. 10D:** Comparison of editing efficiencies with particle number-normalized v1, v2.4, v3.4, and v4 BE-VLPs at the *BCL11A* enhancer site in HEK293T cells. Data are shown as mean values $\pm$ s.e.m. for n=3 biological replicates. **FIG. 10E:** Cell viability after v4 BE-eVLP treatment of HEK293T cells and NIH 3T3 fibroblasts. Data are shown as values $\pm$ s.e.m. for n=3 biological replicates. **FIG. 10F:** Indels frequencies generated by v1 Cas9-VLP and v4 Cas9-eVLPs at the *EMX1* locus in HEK293T cells. Data are shown as values $\pm$ s.e.m. for n=3 biological replicates. **FIG. 10G:** Adenine base editing efficiencies of VSV-G-pseudotyped v4 BE-eVLPs in Neuro-2a cells or 3T3 fibroblasts. Data are shown as individual data points and mean values $\pm$ s.e.m. for n=3 biological replicates.

[0030] FIGS. 11A-11D: Evaluation of off-target editing by v4 BE-eVLPs. **FIG. 11A:** Experimental timeline for the orthogonal R-loop assay. **FIG. 11B:** On-target editing controls for the orthogonal R-loop experiment. Data are shown as individual data points and mean values $\pm$ s.e.m. for n=3 biological replicates. **FIG. 11C:** Cell viability following v4 BE-VLP treatment of RDEB fibroblasts. Data are shown as mean values $\pm$ s.e.m. for n=3 biological replicates. **FIG. 11D:** DNA sequencing reads containing A•T-to-G•C mutations within protospacer positions 4–10 for ten previously identified off-target loci from the genomic DNA of v4-BE-eVLP treated RDEB patient-derived fibroblasts. The dotted grey line represents the highest observed background mutation rate of 0.1%. Data are shown as individual data points and mean values $\pm$ s.e.m. for n=3 biological replicates.

[0031] FIG. 12: Editing efficiencies of BE-VLPs in Neuro2a cells at Dnmt1.

[0032] FIGs. 13A-13B: Flow cytometry analysis for nuclei sorting from the mouse brain after P0 ICV injection. **FIG. 13A:** Singlet nuclei were gated based on FSC/BSC ratio and DyeCycle Ruby signal. The first row demonstrates the gating strategy on a GFP-negative

sample. Bulk nuclei correspond to events that passed gate D for singlet nuclei. **FIG. 13B:** Percentage of GFP-positive nuclei measured by flow cytometry following P0 ICV injection. Data are shown as mean values $\pm$ s.e.m. for $n=3$ biological replicates.

[0033] FIGs. 14A-14C: Assessment of liver toxicity following systemic v4 BE-eVLP injection. **FIG. 14A:** Plasma aspartate transaminase (AST) and alanine transaminase (ALT) levels one week after v4 BE-eVLP injection. **FIGS. 14B-14C:** Histopathological assessment by haematoxylin and eosin staining of livers at 1-week post-injection of (**FIG. 14B**) untreated mice and (**FIG. 14C**) v4 BE-eVLP treated mice. A representative example of each is shown. Scale bars denote 50 μ m.

[0034] FIG. 15A-15C: Sequencing analysis of RPE cDNA after v4 BE-eVLP or lentivirus treatment. **FIG. 15A:** v4 BE-eVLP and lentivirus treatment led to 50-60% of A•T-to-G•C conversion at the target adenine (A_6). Data are shown as individual data points and mean values $\pm$ s.e.m. for $n = 6$ (ABE8e-NG-LV, ABE8e-NG-eVLP, and ABE7.10-NG-eVLP), or $n = 4$ (ABE7.10-NG-LV and untreated) replicates. **FIGs. 15B-15C:** Off-target A-to-G RNA editing by v4 BE-eVLPs and lentiviruses as measured by high-throughput sequencing of the (**FIG. 15B**) *Mcm3ap* and (**FIG. 15C**) *Perp* transcripts. Data are shown as mean values $\pm$ s.e.m. for $n = 6$ (ABE8e-NG-LV, ABE8e-NG-eVLP, and ABE7.10-NG-eVLP), or $n = 4$ (ABE7.10-NG-LV and untreated) replicates.

[0035] FIG. 16. Overview of an embodiment of the manufacture of eVLPs comprising BE RNPs (*e.g.*, BE-VLPs) in a producer cell using a set of expression plasmids which encode the various self-assembling components of the eVLPs: (a) plasmid encoding a Gag-BE fusion protein (*e.g.*, a retroviral Gag, MMLV-Gag-BE fusion protein); (b) plasmid encoding a Gag-Pro-Pol protein (*e.g.*, a retroviral protein, such as a MMLV protease precursor); (c) a plasmid encoding a BE sgRNA; and (d) a plasmid encoding an envelope glycoprotein (*e.g.*, the spike glycoprotein of the vesicular stomatitis virus (VSV-G)). The plasmids are transiently co-transfected into the producer cell, and the encoded protein and sgRNA products are encoded. In some embodiments, such as the fourth-generation eVLPs described herein, the inventors found an optimized stoichiometry ratio of Gag-cargo fusion to Gag-Pro-Pol fusion protein which balances the amount of Gag-cargo available to be packaged into VLPs with the amount of retrovirus protease (the “Pro” in the Gag-Pro-Pol fusion) required for VLP maturation. In one embodiment, the optimized ratio of Gag-cargo fusion to Gag-Pro-Pol fusion protein is achieved by the appropriate ratio of plasmids encoding each component which are transiently delivered to the producer cells. In one embodiment, to modulate the stoichiometry of the

Gag-cargo fusion to Gag-Pro-Pol fusion, the ratio of the plasmid encoding Gag-cargo (*e.g.*, Gag–3xNES–ABE8e) to wild-type MMLV gag-pro-pol plasmids transfected for VLP production was varied. It was found that increasing the amount of gag–cargo plasmid beyond the original proportion used for producing v3.4 BE-eVLPs (38% Gag–cargo plasmid and 62% gag-pro-pol plasmid) did not improve editing efficiencies (**FIG. 2G**). Decreasing the proportion of gag–cargo plasmid from 38% to 25% modestly improved editing efficiencies (**FIG. 2G**). However, further decreasing the proportion of gag–cargo plasmid below 25% reduced editing efficiencies (**FIG. 2G**). These results are consistent with a model in which an optimal gag–cargo:gag-pro-pol stoichiometry balances the amount of gag–cargo available to be packaged into VLPs with the amount of MMLV protease (the “pro” in gag-pro-pol) required for VLP maturation. In one embodiment, the results of this final round of optimization revealed a fourth-generation (v4) BE-eVLP formulation (**FIG. 2G**), which combines the optimal gag–BE:gag-pro-pol stoichiometry (25% gag–BE) with the v3.4 BE-eVLP architecture.

[0036] As depicted in **FIG. 16**, the present disclosure provides pluralities of polynucleotides encoding the eVLP (*e.g.*, BE-VLP) self-assembling component as described herein. In some embodiments, the present disclosure provides pluralities of polynucleotides comprising: (i) a first polynucleotide (*e.g.*, a plasmid) comprising a nucleic acid sequence encoding a viral envelope glycoprotein; (ii) a second polynucleotide (*e.g.*, a plasmid) comprising a nucleic acid sequence encoding a group-specific antigen (gag) protease (pro) polyprotein; (iii) a third polynucleotide (*e.g.*, a plasmid) comprising a nucleic acid sequence encoding a fusion protein, wherein the fusion protein comprises: (a) a group-specific antigen (gag) nucleocapsid protein; (b) a nucleic acid programmable DNA binding protein (napDNAbp); (c) a cleavable linker; and (d) a nuclear export sequence (NES); and (iv) a fourth polynucleotide (*e.g.*, a plasmid) comprising a nucleic acid sequence encoding a guide RNA (gRNA). In some embodiments, the gRNA binds to the napDNAbp of the fusion protein encoded by the third polynucleotide. In some embodiments, the ratio of the second polynucleotide to the third polynucleotide is approximately 10:1, approximately 9:1, approximately 8:1, approximately 7:1, approximately 6:1, approximately 5:1, approximately 4:1, approximately 3:1, approximately 2:1, approximately 1.5:1, approximately 1:1, or approximately 0.5:1. In certain embodiments, the ratio of the second polynucleotide to the third polynucleotide is approximately 3:1.

[0037] FIGs. 17A-17B: v4 BE-eVLPs can efficiently edit primary human hematopoietic stem cells (HSCs). **FIG. 17A:** Four-marker sort for HSCs. Hematopoietic progenitor cells (HPC): CD34⁺/CD38⁺. HSC: CD34⁺/CD38⁻/CD90⁺/CD45RA⁻. **FIG. 17B:** Adenine base editing at the *BCL11A* enhancer locus.

[0038] FIG. 18: v4 BE-eVLPs minimally perturb HSC cellular viability.

[0039] FIGs. 19A-19B: v4 BE-eVLPs enable efficient on-target editing with minimal off-target editing. Lower Cas-dependent off-target editing was observed compared to previous base editing approaches targeting the same site (*e.g.*, Zeng *et al.*, *Nat. Med.* (2020)).

DEFINITIONS

[0040] Unless defined otherwise, all technical and scientific terms used herein have the meaning commonly understood by a person skilled in the art to which this invention belongs. The following references provide one of skill with a general definition of many of the terms used in this invention: Singleton *et al.*, *Dictionary of Microbiology and Molecular Biology* (2nd ed. 1994); *The Cambridge Dictionary of Science and Technology* (Walker ed., 1988); *The Glossary of Genetics*, 5th Ed., R. Rieger *et al.* (eds.), Springer Verlag (1991); and Hale & Marham, *The Harper Collins Dictionary of Biology* (1991). As used herein, the following terms have the meanings ascribed to them unless specified otherwise.

Adenosine deaminase

[0041] As used herein, the term “adenosine deaminase” or “adenosine deaminase domain” refers to a protein or enzyme that catalyzes a deamination reaction of an adenosine (or adenine). The terms are used interchangeably. In certain embodiments, the disclosure provides nucleobase editor fusion proteins comprising one or more adenosine deaminase domains. For instance, an adenosine deaminase domain may comprise a heterodimer of a first adenosine deaminase and a second deaminase domain, connected by a linker. Adenosine deaminases (*e.g.*, engineered adenosine deaminases or evolved adenosine deaminases) provided herein may be enzymes that convert adenine (A) to inosine (I) in DNA or RNA. Such adenosine deaminases can lead to an A:T to G:C base pair conversion. In some embodiments, the deaminase is a variant of a naturally-occurring deaminase from an organism. In some embodiments, the deaminase does not occur in nature. For example, in some embodiments, the deaminase is at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at

least 97%, at least 98%, at least 99%, or at least 99.5% identical to a naturally-occurring deaminase.

[0042] In some embodiments, the adenosine deaminase is derived from a bacterium, such as, *E. coli*, *S. aureus*, *S. typhi*, *S. putrefaciens*, *H. influenzae*, or *C. crescentus*. In some embodiments, the adenosine deaminase is a TadA deaminase. In some embodiments, the TadA deaminase is an *E. coli* TadA deaminase (ecTadA). In some embodiments, the TadA deaminase is a truncated *E. coli* TadA deaminase. For example, the truncated ecTadA may be missing one or more N-terminal amino acids relative to a full-length ecTadA. In some embodiments, the truncated ecTadA may be missing 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 6, 17, 18, 19, or 20 N-terminal amino acid residues relative to the full length ecTadA. In some embodiments, the truncated ecTadA may be missing 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 6, 17, 18, 19, or 20 C-terminal amino acid residues relative to the full length ecTadA. In some embodiments, the ecTadA deaminase does not comprise an N-terminal methionine. In some embodiments, the adenosine deaminase comprises ecTadA(8e) (*i.e.*, as used in the base editor ABE8e) as described further herein. Reference is made to U.S. Patent Publication No. 2018/0073012, published March 15, 2018, which is incorporated herein by reference.

Base editing

[0043] “Base editing” refers to genome editing technology that involves the conversion of a specific nucleic acid base into another at a targeted genomic locus. In certain embodiments, this can be achieved without requiring double-stranded DNA breaks (DSB), or single stranded breaks (*i.e.*, nicking). To date, other genome editing techniques, including CRISPR-based systems, begin with the introduction of a DSB at a locus of interest. Subsequently, cellular DNA repair enzymes mend the break, commonly resulting in random insertions or deletions (indels) of bases at the site of the DSB. However, when the introduction or correction of a point mutation at a target locus is desired rather than stochastic disruption of the entire gene, these genome editing techniques are unsuitable, as correction rates are low (*e.g.*, typically 0.1% to 5%), with the major genome editing products being indels. In order to increase the efficiency of gene correction without simultaneously introducing random indels, the CRISPR/Cas9 system is modified to directly convert one DNA base into another without DSB formation. See, Komor, A.C., *et al.*, Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420-424 (2016), the entire contents of which is incorporated by reference herein.

Base editors

[0044] The terms “base editor (BE)” and “nucleobase editor,” which are used interchangeably herein, refer to an agent comprising a polypeptide that is capable of making a modification to a base (*e.g.*, A, T, C, G, or U) within a nucleic acid sequence (*e.g.*, DNA or RNA) that converts one base to another (*e.g.*, A to G, A to C, A to T, C to T, C to G, C to A, G to A, G to C, G to T, T to A, T to C, or T to G). In some embodiments, the nucleobase editor is capable of deaminating a base within a nucleic acid such as a base within a DNA molecule. In the case of an adenosine nucleobase editor, the nucleobase editor is capable of deaminating an adenine (A) in DNA. Such nucleobase editors may include a nucleic acid programmable DNA binding protein (napDNAbp) fused to an adenosine deaminase. Some nucleobase editors include CRISPR-mediated fusion proteins that are utilized in the base editing methods described herein. In some embodiments, the nucleobase editor comprises a nuclease-inactive Cas9 (dCas9) fused to a deaminase which binds a nucleic acid in a guide RNA-programmed manner via the formation of an R-loop, but does not cleave the nucleic acid. For example, the dCas9 domain of the fusion protein may include a D10A and a H840A mutation (which renders Cas9 capable of cleaving only one strand of a nucleic acid duplex), as described in PCT/US2016/058344, which published as WO 2017/070632 on April 27, 2017, and is incorporated herein by reference. The DNA cleavage domain of *S. pyogenes* Cas9 includes two subdomains, the HNH nuclease subdomain and the RuvC1 subdomain. The HNH subdomain cleaves the strand complementary to the gRNA (the “targeted strand,” or the strand in which editing or deamination occurs), whereas the RuvC1 subdomain cleaves the non-complementary strand containing the PAM sequence (the “non-edited strand”). The RuvC1 mutant D10A generates a nick in the targeted strand, while the HNH mutant H840A generates a nick on the non-edited strand (see Jinek *et al.*, *Science*, 337:816-821(2012); Qi *et al.*, *Cell*. 28;152(5):1173-83 (2013)).

[0045] In some embodiments, a nucleobase editor is a macromolecule or macromolecular complex that results primarily (*e.g.*, more than 80%, more than 85%, more than 90%, more than 95%, more than 99%, more than 99.9%, or 100%) in the conversion of a nucleobase in a polynucleotide sequence into another nucleobase (*i.e.*, a transition or transversion) using a combination of 1) a nucleotide-, nucleoside-, or nucleobase-modifying enzyme and 2) a nucleic acid binding protein that can be programmed to bind to a specific nucleic acid sequence.

[0046] In some embodiments, the nucleobase editor comprises a DNA binding domain (*e.g.*, a programmable DNA binding domain such as a dCas9 or nCas9) that directs it to a target sequence. In some embodiments, the nucleobase editor comprises a nucleobase modification domain fused to a programmable DNA binding domain (*e.g.*, dCas9 or nCas9). The terms “nucleobase modifying enzyme” and “nucleobase modification domain,” which are used interchangeably herein, refer to an enzyme that can modify a nucleobase and convert one nucleobase to another (*e.g.*, a deaminase such as a cytidine deaminase or an adenosine deaminase). The nucleobase modifying enzyme of the nucleobase editor may target cytosine (C) bases in a nucleic acid sequence and convert the C to a thymine (T) base. In some embodiments, C to T editing is carried out by a deaminase, *e.g.*, a cytidine deaminase. In some embodiments, A to G editing is carried out by a deaminase, *e.g.*, an adenosine deaminase. Nucleobase editors that can carry out other types of base conversions (*e.g.*, C to G) are also contemplated.

[0047] A “split nucleobase editor” refers to a nucleobase editor that is provided as an N-terminal portion (also referred to as a N-terminal half) and a C-terminal portion (also referred to as a C-terminal half) encoded by two separate nucleic acids. The polypeptides corresponding to the N-terminal portion and the C-terminal portion of the nucleobase editor may be combined to form a complete nucleobase editor. In some embodiments, for a nucleobase editor that comprises a dCas9 or nCas9, the “split” is located in the dCas9 or nCas9 domain, at positions as described herein in the split Cas9. Accordingly, in some embodiments, the N-terminal portion of the nucleobase editor contains the N-terminal portion of the split Cas9, and the C-terminal portion of the nucleobase editor contains the C-terminal portion of the split Cas9. Similarly, intein-N or intein-C may be fused to the N-terminal portion or the C-terminal portion of the nucleobase editor, respectively, for the joining of the N- and C-terminal portions of the nucleobase editor to form a complete nucleobase editor.

[0048] In some embodiments, a nucleobase editor converts a C to a T. In some embodiments, the nucleobase editor comprises a cytosine deaminase. A “cytosine deaminase”, or “cytidine deaminase,” refers to an enzyme that catalyzes the chemical reaction “cytosine + H₂O → uracil + NH₃” or “5-methyl-cytosine + H₂O → thymine + NH₃.” As may be apparent from the reaction formula, such chemical reactions result in a C to U/T nucleobase change. In the context of a gene, such a nucleotide change, or mutation, may in turn lead to an amino acid change in the protein, which may affect the protein’s function, *e.g.*, loss-of-function or gain-of-function. In some embodiments, the C to T nucleobase editor comprises a dCas9 or nCas9

fused to a cytidine deaminase. In some embodiments, the cytidine deaminase domain is fused to the N-terminus of the dCas9 or nCas9. In some embodiments, the nucleobase editor further comprises a domain that inhibits uracil glycosylase, and/or a nuclear localization signal. Such nucleobase editors have been described in the art, e.g., in Rees & Liu, *Nat Rev Genet.* 2018;19(12):770-788 and Koblan *et al.*, *Nat Biotechnol.* 2018;36(9):843-846; as well as U.S. Patent Publication No. 2018/0073012, published March 15, 2018, which issued as U.S. Patent No. 10,113,163 on October 30, 2018; U.S. Patent Publication No. 2017/0121693, published May 4, 2017, which issued as U.S. Patent No. 10,167,457 on January 1, 2019; PCT Publication No. WO 2017/070633, published April 27, 2017; U.S. Patent Publication No. 2015/0166980, published June 18, 2015; U.S. Patent No. 9,840,699, issued December 12, 2017; U.S. Patent No. 10,077,453, issued September 18, 2018; PCT Publication No. WO 2019/023680, published January 31, 2019; PCT Publication No. WO 2018/0176009, published September 27, 2018, PCT Application No PCT/US2019/033848, filed May 23, 2019, PCT Application No. PCT/US2019/47996, filed August 23, 2019; PCT Application No. PCT/US2019/049793, filed September 5, 2019; International Patent Application No. PCT/US2020/028568, filed April 17, 2020; PCT Application No. PCT/US2019/61685, filed November 15, 2019; PCT Application No. PCT/US2019/57956, filed October 24, 2019; PCT Application No. PCT/US2019/58678, filed October 29, 2019, the contents of each of which are incorporated herein by reference.

[0049] In some embodiments, a nucleobase editor converts an A to a G. In some embodiments, the nucleobase editor comprises an adenosine deaminase. An “adenosine deaminase” is an enzyme involved in purine metabolism. It is needed for the breakdown of adenosine from food and for the turnover of nucleic acids in tissues. Its primary function in humans is the development and maintenance of the immune system. An adenosine deaminase catalyzes hydrolytic deamination of adenosine (forming inosine, which base pairs as G) in the context of DNA. There are no known natural adenosine deaminases that act on DNA. Instead, known adenosine deaminase enzymes only act on RNA (tRNA or mRNA). Evolved deoxyadenosine deaminase enzymes that accept DNA substrates and deaminate dA to deoxyinosine have been described, e.g., in PCT Application PCT/US2017/045381, filed August 3, 2017, which published as WO 2018/027078, PCT Application No. PCT/US2019/033848, which published as WO 2019/226953, PCT Application No. PCT/US2019/033848, filed May 23, 2019, and PCT Patent Application No.

PCT/US2020/028568, filed April 17, 2020; each of which is herein incorporated by reference.

[0050] Exemplary adenosine and cytidine nucleobase editors are also described in Rees & Liu, Base editing: precision chemistry on the genome and transcriptome of living cells, *Nat. Rev. Genet.* 2018;19(12):770-788; as well as U.S. Patent Publication No. 2018/0073012, published March 15, 2018, which issued as U.S. Patent No. 10,113,163 on October 30, 2018; U.S. Patent Publication No. 2017/0121693, published May 4, 2017, which issued as U.S. Patent No. 10,167,457 on January 1, 2019; PCT Publication No. WO 2017/070633, published April 27, 2017; U.S. Patent Publication No. 2015/0166980, published June 18, 2015; U.S. Patent No. 9,840,699, issued December 12, 2017; and U.S. Patent No. 10,077,453, issued September 18, 2018, the contents of each of which are incorporated herein by reference in their entireties.

Cytosine deaminase

[0051] As used herein, a “cytosine deaminase” encoded by the CDA gene is an enzyme that catalyzes the removal of an amine group from cytidine (*i.e.*, the base cytosine when attached to a ribose ring) to uridine (C to U) and deoxycytidine to deoxyuridine (C to U). A non-limiting example of a cytosine deaminase is APOBEC1 (“apolipoprotein B mRNA editing enzyme, catalytic polypeptide 1”). Another example is AID (“activation-induced cytosine deaminase”). Under standard Watson-Crick hydrogen bond pairing, a cytosine base hydrogen bonds to a guanine base. When cytidine is converted to uridine (or deoxycytidine is converted to deoxyuridine), the uridine (or the uracil base of uridine) undergoes hydrogen bond pairing with the base adenine. Thus, a conversion of “C” to uridine (“U”) by cytosine deaminase will cause the insertion of “A” instead of a “G” during cellular repair and/or replication processes. Since the adenine “A” pairs with thymine “T”, the cytosine deaminase in coordination with DNA replication causes the conversion of a C·G pairing to a T·A pairing in the double-stranded DNA molecule.

Cas9

[0052] The term “Cas9” or “Cas9 nuclease” refers to an RNA-guided nuclease comprising a Cas9 domain, or a fragment thereof (*e.g.*, a protein comprising an active or inactive DNA cleavage domain of Cas9, and/or the gRNA binding domain of Cas9). A “Cas9 domain” as used herein, is a protein fragment comprising an active or inactive cleavage domain of Cas9 and/or the gRNA binding domain of Cas9. A “Cas9 protein” is a full length Cas9 protein. A Cas9 nuclease is also referred to sometimes as a casn1 nuclease or a CRISPR (Clustered

Regularly Interspaced Short Palindromic Repeat)-associated nuclease. CRISPR is an adaptive immune system that provides protection against mobile genetic elements (viruses, transposable elements, and conjugative plasmids). CRISPR clusters contain spacers, sequences complementary to antecedent mobile elements, and target invading nucleic acids. CRISPR clusters are transcribed and processed into CRISPR RNA (crRNA). In type II CRISPR systems, correct processing of pre-crRNA requires a trans-encoded small RNA (tracrRNA), endogenous ribonuclease 3 (rnc), and a Cas9 domain. The tracrRNA serves as a guide for ribonuclease 3-aided processing of pre-crRNA. Subsequently, Cas9/crRNA/tracrRNA endonucleolytically cleaves a linear or circular dsDNA target complementary to the spacer. The target strand not complementary to crRNA is first cut endonucleolytically, then trimmed 3'-5' exonucleolytically. In nature, DNA-binding and cleavage typically requires protein and both RNAs. However, single guide RNAs ("sgRNA", or simply "gRNA") can be engineered so as to incorporate aspects of both the crRNA and tracrRNA into a single RNA species. See, *e.g.*, Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. *Science* 337:816-821(2012), the contents of which are incorporated herein by reference. Cas9 recognizes a short motif in the CRISPR repeat sequences (the PAM or protospacer adjacent motif) to help distinguish self versus non-self. Cas9 nuclease sequences and structures are well known to those of skill in the art (see, *e.g.*, "Complete genome sequence of an M1 strain of *Streptococcus pyogenes*." Ferretti *et al.*, J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K., Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P., Qian Y., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L., White J., Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E., *Proc. Natl. Acad. Sci. U.S.A.* 98:4658-4663(2001); "CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III." Deltcheva E., Chylinski K., Sharma C.M., Gonzales K., Chao Y., Pirzada Z.A., Eckert M.R., Vogel J., Charpentier E., *Nature* 471:602-607(2011); and "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. *Science* 337:816-821(2012), the entire contents of each of which are incorporated herein by reference). Cas9 orthologs have been described in various species, including, but not limited to, *S. pyogenes* and *S. thermophilus*. Additional suitable Cas9 nucleases and sequences will be apparent to those of skill in the art based on this disclosure, and such Cas9 nucleases and sequences include Cas9 sequences from the organisms and loci disclosed in Chylinski, Rhun, and Charpentier, "The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems"

(2013) *RNA Biology* 10:5, 726-737; the entire contents of which are incorporated herein by reference. In some embodiments, a Cas9 nuclease comprises one or more mutations that partially impair or inactivate the DNA cleavage domain.

[0053] A nuclease-inactivated Cas9 domain may interchangeably be referred to as a “dCas9” protein (for nuclease-“dead” Cas9). Methods for generating a Cas9 domain (or a fragment thereof) having an inactive DNA cleavage domain are known (see, *e.g.*, Jinek *et al.*, *Science*. 337:816-821(2012); Qi *et al.*, “Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression” (2013) *Cell*. 28;152(5):1173-83, the entire contents of each of which are incorporated herein by reference). For example, the DNA cleavage domain of Cas9 is known to include two subdomains, the HNH nuclease subdomain and the RuvC1 subdomain. The HNH subdomain cleaves the strand complementary to the gRNA, whereas the RuvC1 subdomain cleaves the non-complementary strand. Mutations within these subdomains can silence the nuclease activity of Cas9. For example, the mutations D10A and H840A completely inactivate the nuclease activity of *S. pyogenes* Cas9 (Jinek *et al.*, *Science*. 337:816-821(2012); Qi *et al.*, *Cell*. 28;152(5):1173-83 (2013)). In some embodiments, proteins comprising fragments of Cas9 are provided. For example, in some embodiments, a protein comprises one of two Cas9 domains: (1) the gRNA binding domain of Cas9; or (2) the DNA cleavage domain of Cas9. In some embodiments, proteins comprising Cas9 or fragments thereof are referred to as “Cas9 variants.” A Cas9 variant shares homology to Cas9, or a fragment thereof. For example, a Cas9 variant is at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 96% identical, at least about 97% identical, at least about 98% identical, at least about 99% identical, at least about 99.5% identical, at least about 99.8% identical, or at least about 99.9% identical to wild type Cas9 (*e.g.*, SpCas9 of SEQ ID NO: 13). In some embodiments, the Cas9 variant may have 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, or more amino acid changes compared to wild type Cas9 (*e.g.*, SpCas9 of SEQ ID NO: 13). In some embodiments, the Cas9 variant comprises a fragment of SEQ ID NO: 13 Cas9 (*e.g.*, a gRNA binding domain or a DNA-cleavage domain), such that the fragment is at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 96% identical, at least about 97% identical, at least about 98% identical, at least about 99% identical, at least about 99.5% identical, or at least about 99.9% identical to the corresponding

fragment of wild type Cas9 (*e.g.*, SpCas9 of SEQ ID NO: 13). In some embodiments, the fragment is at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95% identical, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% of the amino acid length of a corresponding wild type Cas9 (*e.g.*, SpCas9 of SEQ ID NO: 13).

CRISPR

[0054] CRISPR is a family of DNA sequences (*i.e.*, CRISPR clusters) in bacteria and archaea that represent snippets of prior infections by a virus that have invaded the prokaryote. The snippets of DNA are used by the prokaryotic cell to detect and destroy DNA from subsequent attacks by similar viruses and effectively compose, along with an array of CRISPR-associated proteins (including Cas9 and homologs thereof) and CRISPR-associated RNA, a prokaryotic immune defense system. In nature, CRISPR clusters are transcribed and processed into CRISPR RNA (crRNA). In certain types of CRISPR systems (*e.g.*, type II CRISPR systems), correct processing of pre-crRNA requires a trans-encoded small RNA (tracrRNA), endogenous ribonuclease 3 (*rnc*) and a Cas9 protein. The tracrRNA serves as a guide for ribonuclease 3-aided processing of pre-crRNA. Subsequently, Cas9/crRNA/tracrRNA endonucleolytically cleaves a linear or circular dsDNA target complementary to the RNA. Specifically, the target strand not complementary to crRNA is first cut endonucleolytically, then trimmed 3'-5' exonucleolytically. In nature, DNA-binding and cleavage typically requires protein and both RNAs. However, single guide RNAs ("sgRNA", or simply "gRNA") can be engineered so as to incorporate aspects of both the crRNA and tracrRNA into a single RNA species – the guide RNA. See, *e.g.*, Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. *Science* 337:816-821(2012), the entire contents of which is hereby incorporated by reference. Cas9 recognizes a short motif in the CRISPR repeat sequences (the PAM or protospacer adjacent motif) to help distinguish self versus non-self. CRISPR biology, as well as Cas9 nuclease sequences and structures are well known to those of skill in the art (see, *e.g.*, "Complete genome sequence of an M1 strain of *Streptococcus pyogenes*." Ferretti *et al.*, J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K., Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P., Qian Y., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L., White J., Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E., *Proc. Natl. Acad. Sci. U.S.A.* 98:4658-4663(2001); "CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III."

Deltcheva E., Chylinski K., Sharma C.M., Gonzales K., Chao Y., Pirzada Z.A., Eckert M.R., Vogel J., Charpentier E., *Nature* 471:602-607(2011); and “A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.” Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. *Science* 337:816-821(2012), the entire contents of each of which are incorporated herein by reference). Cas9 orthologs have been described in various species, including, but not limited to, *S. pyogenes* and *S. thermophilus*. Additional suitable Cas9 nucleases and sequences will be apparent to those of skill in the art based on this disclosure, and such Cas9 nucleases and sequences include Cas9 sequences from the organisms and loci disclosed in Chylinski, Rhun, and Charpentier, “The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems” (2013) *RNA Biology* 10:5, 726-737; the entire contents of which are incorporated herein by reference.

[0055] In certain types of CRISPR systems (*e.g.*, type II CRISPR systems), correct processing of pre-crRNA requires a trans-encoded small RNA (tracrRNA), endogenous ribonuclease 3 (*rnc*), and a Cas9 protein. The tracrRNA serves as a guide for ribonuclease 3-aided processing of pre-crRNA. Subsequently, Cas9/crRNA/tracrRNA endonucleolytically cleaves a linear or circular nucleic acid target complementary to the RNA. Specifically, the target strand not complementary to crRNA is first cut endonucleolytically, then trimmed 3'-5' exonucleolytically. In nature, DNA-binding and cleavage typically requires protein and both RNAs. However, single guide RNAs (“sgRNA”, or simply “gRNA”) can be engineered so as to incorporate embodiments of both the crRNA and tracrRNA into a single RNA species — the guide RNA.

[0056] In general, a “CRISPR system” refers collectively to transcripts and other elements involved in the expression of or directing the activity of CRISPR-associated (“Cas”) genes, including sequences encoding a Cas gene, a tracr (trans-activating CRISPR) sequence (*e.g.*, tracrRNA or an active partial tracrRNA), a tracr mate sequence (encompassing a “direct repeat” and a tracrRNA-processed partial direct repeat in the context of an endogenous CRISPR system), a guide sequence (also referred to as a “spacer” in the context of an endogenous CRISPR system), or other sequences and transcripts from a CRISPR locus. The tracrRNA of the system is complementary (fully or partially) to the tracr mate sequence present on the guide RNA.

Deaminase

[0057] The term “deaminase” or “deaminase domain” refers to a protein or enzyme that catalyzes a deamination reaction. In some embodiments, the deaminase is an adenosine (or

adenine) deaminase, which catalyzes the hydrolytic deamination of adenine or adenosine. In some embodiments, the adenosine deaminase catalyzes the hydrolytic deamination of adenine or adenosine in deoxyribonucleic acid (DNA) to inosine. In other embodiments, the deaminase is a cytidine (or cytosine) deaminase, which catalyzes the hydrolytic deamination of cytidine or cytosine.

[0058] The deaminases provided herein may be from any organism, such as a bacterium. In some embodiments, the deaminase or deaminase domain is a variant of a naturally-occurring deaminase from an organism. In some embodiments, the deaminase or deaminase domain does not occur in nature. For example, in some embodiments, the deaminase or deaminase domain is at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to a naturally-occurring deaminase.

Fusion protein

[0059] The term “fusion protein” as used herein refers to a hybrid polypeptide which comprises protein domains from at least two different proteins. One protein may be located at the amino-terminal (N-terminal) portion of the fusion protein or at the carboxy-terminal (C-terminal) protein thus forming an “amino-terminal fusion protein” or a “carboxy-terminal fusion protein,” respectively. A protein may comprise different domains, for example, a nucleic acid binding domain (*e.g.*, the gRNA binding domain of Cas9 that directs the binding of the protein to a target site) and a nucleic acid cleavage domain or a catalytic domain of a nucleic-acid editing protein. Another example includes fusion of a Cas9 or equivalent thereof to a deaminase. Any of the proteins provided herein may be produced by any method known in the art. For example, the proteins provided herein may be produced via recombinant protein expression and purification, which is especially suited for fusion proteins comprising a peptide linker. Methods for recombinant protein expression and purification are well known, and include those described by Green and Sambrook, *Molecular Cloning: A Laboratory Manual* (4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012)), the entire contents of which is incorporated herein by reference.

Group-specific antigen (gag)

[0060] Without being limited by theory, and in the context of typical envelope virus lifecycle, Gag is the primary structural protein responsible for orchestrating the majority of steps in viral assembly, including budding out of fully-formed enveloped virions having an (i) envelope (comprising a lipid membrane formed from cell membrane during budding out, and

one or more glycoproteins inserted therein), and (ii) a capsid, which is the internal protein shell. Most of these assembly steps occur via interactions with three Gag subdomains – matrix (MA), capsid (CA), and nucleocapsid (NC; Figure 1). These three regions have a low level of sequence conservation among the different retroviral genera, which belies the observed high level of structural conservation. Outside of these three domains, Gag proteins can vary widely. For example, HIV-1 Gag additionally codes for a C-terminal p6 protein as well as two spacer proteins, SP1 and SP2, which demarcate the CA–NC and NC–p6 junctions, but HTLV-1 contains no additional sequences outside of MA, CA, and NC (Oroszlan and Copeland, 1985; Henderson *et al.*, 1992).

[0061] Gag is also referred to as a “viral structural protein.” As used herein, the term “viral structural protein” refers to viral proteins that contribute to the overall structure of the capsid protein or of the protein core of a virus. The term “viral structural protein” further includes functional fragments or derivatives of such viral protein contributing to the structure of a capsid protein or of protein core of a virus. An example of viral structural protein is MMLV Gag. The viral membrane fusion proteins are not considered as viral structural proteins. Typically, said viral structural proteins are localized inside the core of the virus.

Group-specific antigen (gag) nucleocapsid protein

[0062] The term “group-specific antigen nucleocapsid protein” or “gag nucleocapsid protein” refers to a protein that makes up the core structural component of the inner shell of many viruses, including retroviruses. The gag nucleocapsid proteins used in the BE-VLPs of the present disclosure may be an MMLV gag nucleocapsid protein, an FMLV gag nucleocapsid protein, or a nucleocapsid protein from any other virus that produces such proteins.

Group-specific antigen (gag) protease (pro) polyprotein

[0063] A “group-specific antigen (gag) protease (pro) polyprotein” or “gag-pro polyprotein” refers to a gag nucleocapsid protein further comprising a viral protease linked thereto. Gag-pro polyproteins mediate proteolytic cleavage of gag and gag-pol polyproteins or nucleocapsid proteins during or shortly after the release of a virion from the plasma membrane. In the BE-VLPs described herein, the protease of a gag-pro polyprotein is responsible for cleaving a cleavable linker in the fusion protein to release a base editor following delivery of the BE-VLP to a target cell. In some embodiments, a gag-pro polyprotein is an MMLV gag-pro polyprotein or an FMLV gag-pro polyprotein.

Guide RNA (“gRNA”)

[0064] As used herein, the term “guide RNA” is a particular type of guide nucleic acid which is mostly commonly associated with a Cas protein of a CRISPR-Cas9 and which associates with Cas9, directing the Cas9 protein to a specific sequence in a DNA molecule that includes complementarity to the protospacer sequence of the guide RNA. However, this term also embraces the equivalent guide nucleic acid molecules that associate with Cas9 equivalents, homologs, orthologs, or paralogs, whether naturally occurring or non-naturally occurring (*e.g.*, engineered or recombinant), and which otherwise program the Cas9 equivalent to localize to a specific target nucleotide sequence. The Cas9 equivalents may include other Cas proteins from any type of CRISPR system (*e.g.*, type II, V, VI), including Cpf1 (a type-V CRISPR-Cas system), C2c1 (a type V CRISPR-Cas system), C2c2 (a type VI CRISPR-Cas system), and C2c3 (a type V CRISPR-Cas system). Further Cas-equivalents are described in Makarova *et al.*, “C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector,” *Science* 2016; 353(6299), the contents of which are incorporated herein by reference. Exemplary sequences and structures of guide RNAs are provided herein.

[0065] A guide RNA is a particular type of guide nucleic acid which is most commonly associated with a Cas protein of a CRISPR-Cas9 and which associates with Cas9, directing the Cas9 protein to a specific sequence in a DNA molecule that includes complementarity to the protospacer sequence for the guide RNA. Functionally, guide RNAs associate with Cas9, directing (or programming) the Cas9 protein to a specific sequence in a DNA molecule that includes a sequence complementary to the protospacer sequence for the guide RNA. A gRNA is a component of the CRISPR/Cas system. Typically, a guide RNA comprises a fusion of a CRISPR-targeting RNA (crRNA) and a trans-activation crRNA (tracrRNA), providing both targeting specificity and scaffolding/binding ability for Cas9 nuclease. A “crRNA” is a bacterial RNA that confers target specificity and requires tracrRNA to bind to Cas9. A “tracrRNA” is a bacterial RNA that links the crRNA to the Cas9 nuclease and typically can bind any crRNA. The sequence specificity of a Cas DNA-binding protein is determined by gRNAs, which have nucleotide base-pairing complementarity to target DNA sequences. The native gRNA comprises a 20 nucleotide (nt) Specificity Determining Sequence (SDS), or spacer, which specifies the DNA sequence to be targeted, and is immediately followed by an 80 nt scaffold sequence, which associates the gRNA with Cas9. In some embodiments, an SDS of the present disclosure has a length of 15 to 100 nucleotides, or more. For example, an SDS may have a length of 15 to 90, 15 to 85, 15 to 80, 15 to 75, 15 to 70, 15 to 65, 15 to 60,

15 to 55, 15 to 50, 15 to 45, 15 to 40, 15 to 35, 15 to 30, or 15 to 20 nucleotides. In some embodiments, the SDS is 20 nucleotides long. For example, the SDS may be 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 nucleotides long. At least a portion of the target DNA sequence is complementary to the SDS of the gRNA. For Cas9 to successfully bind to the DNA target sequence, a region of the target sequence is complementary to the SDS of the gRNA sequence and is immediately followed by the correct protospacer adjacent motif (PAM) sequence (*e.g.*, NGG for Cas9 and TTN, TTTN, or YTN for Cpf1). In some embodiments, an SDS is 100% complementary to its target sequence. In some embodiments, the SDS sequence is less than 100% complementary to its target sequence and is, thus, considered to be partially complementary to its target sequence. For example, a targeting sequence may be 99%, 98%, 97%, 96%, 95%, 94%, 93%, 92%, 91%, or 90% complementary to its target sequence. In some embodiments, the SDS of template DNA or target DNA may differ from a complementary region of a gRNA by 1, 2, 3, 4, or 5 nucleotides.

[0066] In some embodiments, the guide RNA is about 15-120 nucleotides long and comprises a sequence of at least 10 contiguous nucleotides that is complementary to a target sequence. In some embodiments, the guide RNA is 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, or 120 nucleotides long. In some embodiments, the guide RNA comprises a sequence of 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more contiguous nucleotides that is complementary to a target sequence. Sequence complementarity refers to distinct interactions between adenine and thymine (DNA) or uracil (RNA), and between guanine and cytosine.

Linker

[0067] The term “linker,” as used herein, refers to a molecule linking two other molecules or moieties. The linker can be an amino acid sequence in the case of a linker joining two fusion proteins. For example, a Cas9 can be fused to a deaminase (*e.g.*, an adenosine deaminase or a cytosine deaminase) by an amino acid linker sequence. The linker can also be a nucleotide sequence in the case of joining two nucleotide sequences together (*e.g.*, in a gRNA). In other embodiments, the linker is an organic molecule, group, polymer, or chemical moiety. In some embodiments, the linker is 5-200 amino acids in length, for example, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 30-35, 35-40, 40-45, 45-

50, 50-60, 60-70, 70-80, 80-90, 90-100, 100-150, or 150-200 amino acids in length. Longer or shorter linkers are also contemplated.

[0068] A “cleavable linker” refers to a linker that can be split or cut by any means. The linker can be an amino acid sequence. In some embodiments, the linker between the NES and the napDNAbp of the BE-VLPs provided herein comprises a cleavable linker. A cleavable linker may comprise a self-cleaving peptide (*e.g.*, a 2A peptide such as EGRGSLTTCGDVEENPGP (SEQ ID NO: 9), ATNFSLLKQAGDVEENPGP (SEQ ID NO: 10), QCTNYALLKLAGDVESNPGP (SEQ ID NO: 11), or VKQTLNFDLLKLAGDVESNPGP (SEQ ID NO: 12)). In some embodiments, a cleavable linker comprises a protease cleavage site that is cut after being contacted by a protease. For example, the present disclosure contemplates that use of cleavable linkers comprising a protease cleavage site of amino acid sequences TSTLLMENSS (SEQ ID NO: 1), PRSSLYPALTP (SEQ ID NO: 2), VQALVLTQ (SEQ ID NO: 3), PLQVLTNIERR (SEQ ID NO: 4), or an amino acid sequence at least 90% identical to any one of SEQ ID NOs: 1-4. In certain embodiments, a cleavable linker comprises an MMLV protease cleavage site of an FMLV protease cleavage site.

napDNAbp

[0069] As used herein, the term “nucleic acid programmable DNA binding protein” or “napDNAbp,” of which Cas9 is an example, refers to a protein that uses RNA:DNA hybridization to target and bind to specific sequences in a DNA molecule. Each napDNAbp is associated with at least one guide nucleic acid (*e.g.*, guide RNA), which localizes the napDNAbp to a DNA sequence that comprises a DNA strand (*i.e.*, a target strand) that is complementary to the guide nucleic acid, or a portion thereof (*e.g.*, the protospacer of a guide RNA). In other words, the guide nucleic-acid “programs” the napDNAbp (*e.g.*, Cas9 or equivalent) to localize and bind to a complementary sequence.

[0070] Without being bound by theory, the binding mechanism of a napDNAbp – guide RNA complex, in general, includes the step of forming an R-loop whereby the napDNAbp induces the unwinding of a double-strand DNA target, thereby separating the strands in the region bound by the napDNAbp. The guide RNA protospacer then hybridizes to the “target strand.” This displaces a “non-target strand” that is complementary to the target strand, which forms the single strand region of the R-loop. In some embodiments, the napDNAbp includes one or more nuclease activities, which then cut the DNA, leaving various types of lesions. For example, the napDNAbp may comprise a nuclease activity that cuts the non-target strand at a

first location, and/or cuts the target strand at a second location. Depending on the nuclease activity, the target DNA can be cut to form a “double-stranded break” whereby both strands are cut. In other embodiments, the target DNA can be cut at only a single site, *i.e.*, the DNA is “nicked” on one strand. Exemplary napDNAbp with different nuclease activities include “Cas9 nickase” (“nCas9”) and a deactivated Cas9 having no nuclease activities (“dead Cas9” or “dCas9”). Exemplary sequences for these and other napDNAbp are provided herein.

Nickase

[0071] As used herein, a “nickase” refers to a napDNAbp (*e.g.*, a Cas protein) which is capable of cleaving only one of the two complementary strands of a double-stranded target DNA sequence, thereby generating a nick in that strand. In some embodiments, the nickase cleaves a non-target strand of a double stranded target DNA sequence. In some embodiments, the nickase comprises an amino acid sequence with one or more mutations in a catalytic domain of a canonical napDNAbp (*e.g.*, a Cas protein), wherein the one or more mutations reduces or abolishes nuclease activity of the catalytic domain. In some embodiments, the nickase is a Cas9 that comprises one or more mutations in a RuvC-like domain relative to a wild type Cas9 sequence or to an equivalent amino acid position in other Cas9 variants or Cas9 equivalents. In some embodiments, the nickase is a Cas9 that comprises one or more mutations in an HNH-like domain relative to a wild type Cas9 sequence or to an equivalent amino acid position in other Cas9 variants or Cas9 equivalents. In some embodiments, the nickase is a Cas9 that comprises an aspartate-to-alanine substitution (D10A) in the RuvC I catalytic domain of Cas9 relative to a canonical Cas9 sequence or to an equivalent amino acid position in other Cas9 variants or Cas9 equivalents. In some embodiments, the nickase is a Cas9 that comprises a H840A, N854A, and/or N863A mutation relative to a canonical Cas9 sequence, or to an equivalent amino acid position in other Cas9 variants or Cas9 equivalents. In some embodiments, the term “Cas9 nickase” refers to a Cas9 with one of the two nuclease domains inactivated. This enzyme is capable of cleaving only one strand of a target DNA. In some embodiments, the nickase is a Cas protein that is not a Cas9 nickase.

Nuclear export sequence (NES)

[0072] The term “nuclear export sequence” or “NES” refers to an amino acid sequence that promotes transport of a protein out of the cell nucleus to the cytoplasm, for example, through the nuclear pore complex by nuclear transport. Nuclear export sequences are known in the art and would be apparent to the skilled artisan. For example, NES sequences are described in Xu, D. *et al.* Sequence and structural analyses of nuclear export signals in the NESdb

database. *Mol Biol. Cell.* 2012, 23(18) 3677-3693, the contents of which are incorporated herein by reference.

Nuclear localization sequence (NLS)

[0073] The term “nuclear localization sequence” or “NLS” refers to an amino acid sequence that promotes import of a protein into the cell nucleus, for example, by nuclear transport. Nuclear localization sequences are known in the art and would be apparent to the skilled artisan. For example, NLS sequences are described in Plank *et al.*, International PCT Application, PCT/EP2000/011690, filed November 23, 2000, published as WO/2001/038547 on May 31, 2001, the contents of which are incorporated herein by reference for its disclosure of exemplary nuclear localization sequences. In some embodiments, an NLS comprises the amino acid sequence PKKKRKV (SEQ ID NO: 204).

Nucleic acid molecule

[0074] The term “nucleic acid,” as used herein, refers to a polymer of nucleotides. The polymer may include natural nucleosides (*i.e.*, adenosine, thymidine, guanosine, cytidine, uridine, deoxyadenosine, deoxythymidine, deoxyguanosine, and deoxycytidine), nucleoside analogs (*e.g.*, 2-aminoadenosine, 2-thiothymidine, inosine, pyrrolo-pyrimidine, 3-methyl adenosine, 5-methylcytidine, C5 bromouridine, C5 fluorouridine, C5 iodouridine, C5 propynyl uridine, C5 propynyl cytidine, C5 methylcytidine, 7 deazaadenosine, 7 deazaguanosine, 8 oxoadenosine, 8 oxoguanosine, O(6) methylguanine, 4-acetylcytidine, 5-(carboxyhydroxymethyl)uridine, dihydrouridine, methylpseudouridine, 1-methyl adenosine, 1-methyl guanosine, N6-methyl adenosine, and 2-thiocytidine), chemically modified bases, biologically modified bases (*e.g.*, methylated bases), intercalated bases, modified sugars (*e.g.*, 2'-fluororibose, ribose, 2'-deoxyribose, 2'-O-methylcytidine, arabinose, and hexose), or modified phosphate groups (*e.g.*, phosphorothioates and 5' N phosphoramidite linkages).

Protease cleavage site

[0075] The term “protease cleavage site,” as used herein, refers to an amino acid sequence that is recognized and cleaved by a protease, *i.e.*, an enzyme that catalyzes proteolysis and breaks down proteins into smaller polypeptides, or single amino acids. In some embodiments, a protease cleavage site is included in a cleavable linker in a fusion protein, as described herein. In certain embodiments, a protease cleavage site is cleaved by the protease of a gag-pro polyprotein. In some embodiments, a protease cleavage site comprises an MMLV protease cleavage site or an FMLV protease cleavage site. In certain embodiments, a protease cleavage site comprises one of the amino acid sequences TSTLLMENSS (SEQ ID NO: 1),

PRSSLYPALTP (SEQ ID NO: 2), VQALVLTQ (SEQ ID NO: 3), PLQVLTNIERR (SEQ ID NO: 4), or an amino acid sequence at least 90% identical to any one of SEQ ID NOs: 1-4. In some embodiments, a protease cleavage site comprises an amino acid sequence of any one of SEQ ID NOs: 1-8 or 499-510, or an amino acid sequence at least 90% identical to any one of SEQ ID NOs: 1-8 or 499-510.

Protein, peptide, and polypeptide

[0076] The terms “protein,” “peptide,” and “polypeptide” are used interchangeably herein and refer to a polymer of amino acid residues linked together by peptide (amide) bonds. The terms refer to a protein, peptide, or polypeptide of any size, structure, or function. Typically, a protein, peptide, or polypeptide will be at least three amino acids long. A protein, peptide, or polypeptide may refer to an individual protein or a collection of proteins. One or more of the amino acids in a protein, peptide, or polypeptide may be modified, for example, by the addition of a chemical entity such as a carbohydrate group, a hydroxyl group, a phosphate group, a farnesyl group, an isofarnesyl group, a fatty acid group, a linker for conjugation, functionalization, or other modification, etc. A protein, peptide, or polypeptide may also be a single molecule or may be a multi-molecular complex. A protein, peptide, or polypeptide may be just a fragment of a naturally occurring protein or peptide. A protein, peptide, or polypeptide may be naturally occurring, recombinant, or synthetic, or any combination thereof. Any of the proteins provided herein may be produced by any method known in the art. For example, the proteins provided herein may be produced via recombinant protein expression and purification, which is especially suited for fusion proteins comprising a peptide linker. Methods for recombinant protein expression and purification are well known, and include those described by Green and Sambrook, *Molecular Cloning: A Laboratory Manual* (4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012)), the contents of which are incorporated herein by reference.

Subject

[0077] The term “subject,” as used herein, refers to an individual organism, for example, an individual mammal. In some embodiments, the subject is a human. In some embodiments, the subject is a non-human mammal. In some embodiments, the subject is a non-human primate. In some embodiments, the subject is a rodent. In some embodiments, the subject is a sheep, a goat, a cattle, a cat, or a dog. In some embodiments, the subject is a vertebrate, an amphibian, a reptile, a fish, an insect, a fly, or a nematode. In some embodiments, the subject is a research animal. In some embodiments, the subject is genetically engineered, *e.g.*, a

genetically engineered non-human subject. The subject may be of either sex and at any stage of development.

Treatment

[0078] The terms “treatment,” “treat,” and “treating,” refer to a clinical intervention aimed to reverse, alleviate, delay the onset of, or inhibit the progress of a disease or disorder, or one or more symptoms thereof, as described herein. As used herein, the terms “treatment,” “treat,” and “treating” refer to a clinical intervention aimed to reverse, alleviate, delay the onset of, or inhibit the progress of a disease or disorder, or one or more symptoms thereof, as described herein. In some embodiments, treatment may be administered after one or more symptoms have developed and/or after a disease has been diagnosed. In other embodiments, treatment may be administered in the absence of symptoms, *e.g.*, to prevent or delay onset of a symptom or inhibit onset or progression of a disease. For example, treatment may be administered to a susceptible individual prior to the onset of symptoms (*e.g.*, in light of a history of symptoms and/or in light of genetic or other susceptibility factors). Treatment may also be continued after symptoms have resolved, for example, to prevent or delay their recurrence.

Variant

[0079] As used herein, the term “variant” should be taken to mean the exhibition of qualities that have a pattern that deviates from what occurs in nature, *e.g.*, a variant Cas9 is a Cas9 comprising one or more changes in amino acid residues as compared to a wild type Cas9 amino acid sequence. The term “variant” encompasses homologous proteins having at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% identity with a reference sequence and having the same or substantially the same functional activity or activities as the reference sequence. The term also encompasses mutants, truncations, or domains of a reference sequence that display the same or substantially the same functional activity or activities as the reference sequence.

Vector

[0080] The term “vector,” as used herein, refers to a nucleic acid that can be modified to encode a gene of interest and that is able to enter a host cell, mutate, and replicate within the host cell, and then transfer a replicated form of the vector into another host cell. Exemplary suitable vectors include viral vectors, such as retroviral vectors or bacteriophages and filamentous phage, and conjugative plasmids. Additional suitable vectors will be apparent to those of skill in the art based on the instant disclosure.

Viral envelope glycoprotein

The term “viral envelope glycoprotein” refers to oligosaccharide-containing proteins that form a part of the viral envelope, *i.e.*, the outermost layer of many types of viruses that protects the viral genetic materials when traveling between host cells. Glycoproteins may assist with identification and binding to receptors on a target cell membrane so that the viral envelope fuses with the membrane, allowing the contents of the viral particle (which may comprise, *e.g.*, a BE-VLP as described herein) to enter the host cell. This property may also be referred to as “tropism.” The viral envelope glycoproteins used in the BE-VLPs (or aka the eVLPs) of the present disclosure may comprise any glycoprotein from an enveloped virus. In some embodiments, a viral envelope glycoprotein is an adenoviral envelope glycoprotein, an adeno-associated viral envelope glycoprotein, a retroviral envelope glycoprotein, or a lentiviral envelope glycoprotein. In certain embodiments, a viral envelope glycoprotein is a vesicular stomatitis virus G protein (VSV-G), a baboon retroviral envelope glycoprotein (BaEVRless), a FuG-B2 envelope glycoprotein, an HIV-1 envelope glycoprotein, or an ecotropic murine leukemia virus (MLV) envelope glycoprotein.

Virus-like particles (VLPs)

[0081] As used herein, a virus-like particle consists of a supra-molecular assembly comprising (a) an envelope comprising (i) a lipid membrane (*e.g.*, single-layer or bi-layer membrane) and a (ii) viral envelope glycoprotein, and (b) a multi-protein core region comprising (ii) a Gag protein, (ii) a first fusion protein comprising a Gag protein and Pro-Pol, and (iii) a second fusion protein comprising a Gag protein fused to a cargo protein via a protease-cleavable linker. In various embodiments, the cargo protein is a napDNAbp (*e.g.*, Cas9). In other embodiments, the cargo protein is a base editor. In various other embodiments, the multi-protein core region of the VLPs further comprises one or more guide RNA molecules which are complexed with the napDNAbp or the base editor to form a ribonucleoprotein (RNP). In various embodiments, the VLPs are prepared in a producer cell that is transiently transformed with plasmid DNA that encodes that various protein and nucleic acid (sgRNA) components of the VLPs. The components self-assemble at the cell membrane and bud out in accordance with the naturally occurring mechanism of retroviral budding in order to release from the cell fully-matured VLPs. Once formed, the Pol-Pro cleaves the protease-sensitive linker joining the Gag-cargo linker (*e.g.*, the linker joining a Gag to a BE RNP or a napDNAbp RNP) to release the BE RNP and/or napDNAbp RNA as the case may be within the VLP. Once the VLP is administered to a recipient cell and take up

by said cell, the contents of the VLP are released, including free BE RNP and/or napDNAbp RNA. Once in the cell, the RNPs may translocate to the nucleus of the cell (in particular, where NLSs are included on the RNPs), where DNA editing may occur at target sites specified by the guide RNA. Various embodiments comprise one or more improvements.

[0082] In one embodiment, the protease-cleavable linker is optimized to improve cleavage efficiency after VLP maturation, as demonstrated herein for v.2 VLPs (or “second generation” VLPs).

[0083] In another embodiment, the Gag-cargo fusion (e.g., Gag::BE) further comprises one or more nuclear export signals at one or more locations along the length of the fusion polypeptide protein which may be joined by a cleavable linker such that during VLP assembly in the producer cell, the Gag-cargo fusions (due to presence of competing NLS signals) do not accumulate in the nucleus of the producer cells but instead are available in the cytoplasm to undergo the VLP assembly process at the cell membrane. Once inside the matured VLPs following release from the producer cell, the NES may be cleaved by Pro-Pol thereby separating the cargo (e.g., napDNAbp or a BE) from the NES. Upon delivery to a recipient cell, therefore, the cargo (e.g., napDNAbp or BE, typically flanked with one or more NLS elements) will not comprise an NES element, which may otherwise prohibit the transport of the cargo into the nucleus and hinder gene editing activity. This is exemplified as v.3 VLPs described herein (or “third generation” VLPs).

[0084] In another embodiment, as demonstrated by v.4 VLPs (or “fourth generation” VLPs) described herein, the inventors found an optimized stoichiometry ratio of Gag-cargo fusion to Gag-Pro-Pol fusion protein which balances the amount of Gag-cargo available to be packaged into VLPs with the amount of retrovirus protease (the “Pro” in the Gag-Pro-Pol fusion) required for VLP maturation. In one embodiment, the optimized ratio of Gag-cargo fusion to Gag-Pro-Pol fusion protein is achieved by the appropriate ratio of plasmids encoding each component which are transiently delivered to the producer cells. In one embodiment, to modulate the stoichiometry of the Gag-cargo fusion to Gag-Pro-Pol fusion, the ratio of the plasmid encoding Gag-cargo (e.g., Gag-3xNES-ABE8e) to wild-type MMLV gag-pro-pol plasmids transfected for VLP production was varied. It was found that increasing the amount of gag-cargo plasmid beyond the original proportion used for producing v3.4 BE-eVLPs (38% Gag-cargo plasmid and 62% gag-pro-pol plasmid) did not improve editing efficiencies (**FIG. 2G**). Decreasing the proportion of gag-cargo plasmid from 38% to 25% modestly improved editing efficiencies (**FIG. 2G**). However, further decreasing the

proportion of gag–cargo plasmid below 25% reduced editing efficiencies (**FIG. 2G**). These results are consistent with a model in which an optimal gag–cargo:gag-pro-pol stoichiometry balances the amount of gag–cargo available to be packaged into VLPs with the amount of MMLV protease (the “pro” in gag-pro-pol) required for VLP maturation. In one embodiment, the results of this final round of optimization revealed a fourth-generation (v4) BE-eVLP formulation (**FIG. 2G**), which combines the optimal gag–BE:gag-pro-pol stoichiometry (25% gag–BE) with the v3.4 BE-eVLP architecture.

[0085] In some embodiments, a VLP comprises additional agents for targeting the VLP for delivery to particular cell types. For example, such additional targeting agents may be incorporated into the outer lipid membrane encapsulation layer of the VLP. In some embodiments, the additional targeting agent is a protein. In certain embodiments, the additional targeting agent is an antibody.

[0086] Thus, as used herein, a virus-derived particle comprises a virus-like particle formed by one or more virus-derived protein(s), which virus-derived particle is substantially devoid of a viral genome such that the VLP is replication-incompetent when delivered to a recipient cell.

Wild type

[0087] As used herein the term “wild type” is a term of the art understood by skilled persons and means the typical form of an organism, strain, gene or characteristic as it occurs in nature as distinguished from mutant or variant forms.

DETAILED DESCRIPTION

[0088] The present disclosure is based on the development and application of an engineered VLP (eVLPs) platform for packaging and delivering a ribonucleoprotein cargo, such as a napDNAbp-guide RNA cargo or a base editor-guide RNA cargo, *in vitro* and/or *in vivo*. In embodiments which deliver base editor-guide RNA ribonucleoprotein cargo, the eVLPs may be referred to as base editor virus-like proteins (BE-VLPs). In various embodiments, the optimized BE-VLPs enable highly efficient base editing with minimal off-target editing in a variety of cell types. In particular, the BE-VLPs described herein are based on the surprising discovery that both nuclear-export sequences (NES) and nuclear localization sequences (NLS) may be included on the same fusion protein to promote trafficking of the fusion protein to different parts of a cell during production and during delivery. The presently described BE-VLPs are produced in viral producer cells and exported from the nucleus due to

the presence of one or more NES sequences in the fusion proteins inside the BE-VLPs. Following delivery to a target cell, the NES is cleaved from the fusion protein when the BE is released from the VLP, allowing the BE (which comprises one or more NLS sequences) to enter the nucleus of a target cell and edit the genome. The present disclosure also describes the optimization of a protease cleavage site which separates the NES and VLP proteins from the rest of the base editor to promote highly efficient cleavage and delivery of the BE. Finally, the present disclosure also describes the optimization of the ratios of various components of the BE-VLPs, ensuring high efficiency of BE-VLP production.

[0089] Accordingly, the present disclosure provides virus-like particles for delivering base editor fusion proteins (BE-VLPs) and systems comprising such BE-VLPs. The present disclosure also provides polynucleotides encoding the BE-VLPs described herein, which may be useful for producing said VLPs. Also provided herein are methods for editing the genome of a target cell by introducing the presently described BE-VLPs into the target cell. The present disclosure also provides fusion proteins that make up a component of the BE-VLPs described herein, as well as polynucleotides, vectors, cells, and kits.

eVLPs

[0090] In various embodiments, the eVLPs (*e.g.*, BE-VLPs) comprise a supra-molecular assembly comprising (a) an envelope comprising (i) a lipid membrane (*e.g.*, single-layer or bi-layer membrane) and a (ii) viral envelope glycoprotein and (b) a multi-protein core region enclosed by the envelope and comprising (i) a Gag protein, (ii) a Gag-Pro-Pol protein, and (iii) a Gag-cargo fusion protein comprising a Gag protein fused to a cargo protein (*e.g.*, a napDNAbp or BE) via a cleavable linker (*e.g.*, a protease-cleavable linker). In various embodiments, the cargo protein is a napDNAbp (*e.g.*, Cas9). In other embodiments, the cargo protein is a base editor. In various other embodiments, the multi-protein core region of the VLPs further comprises one or more guide RNA molecules which are complexed with the napDNAbp or the base editor to form a ribonucleoprotein (RNP). In various embodiments, the VLPs are prepared in a producer cell that is transiently transformed with plasmid DNA that encodes the various protein and nucleic acid (sgRNA) components of the VLPs. Without being bound by theory, the components self-assemble at the cell membrane and bud out in accordance with the naturally occurring mechanism of budding (*e.g.*, retroviral budding or the budding mechanism of other envelope viruses) in order to release from the cell fully-matured VLPs. Once formed, the Gag-Pol-Pro cleaves the protease-sensitive linker of the Gag-cargo (*i.e.*, [Gag]-[cleavable linker]-[cargo], wherein the cargo can be BE-RNP or a

napDNAbp RNP) thereby releasing the BE RNP and/or napDNAbp RNA, as the case may be, within the VLP. Thus, in various embodiments, the present disclosure also provides VLPs in which the napDNAbp or base editor has been cleaved off of the gag protein and released within the VLP. For example, the present disclosure provides VLPs comprising a group-specific antigen (gag) protease (pro) polyprotein, a nucleic acid programmable DNA binding protein (napDNAbp), and a fusion protein comprising a gag nucleocapsid protein and a nuclear export sequence (NES), encapsulated by a lipid membrane and a viral envelope glycoprotein. In some embodiments, the present disclosure provides VLPs comprising a mixture of cleaved and uncleaved products (*i.e.*, a mixture of napDNAbps that have been cleaved from the gag protein and that have not yet been cleaved from the gag protein). In some embodiments, the napDNAbp is fused to one or more additional domains such as one or more NLS and/or a deaminase (*e.g.*, to form a base editor).

[0091] Once the VLP is administered to a recipient cell and taken up by said recipient cell, the contents of the VLP are released, *e.g.*, released BE RNP and/or napDNAbp RNP. Once in the cell, the RNPs may translocate to the nucleus of the cell (in particular, where NLSs are included on the RNPs), where DNA editing may occur at target sites specified by the guide RNA. Various embodiments comprise one or more improvements.

[0092] In one embodiment, the protease-cleavable linker is optimized to improve cleavage efficiency after VLP maturation, as demonstrated herein for v.2 VLPs (or “second generation” VLPs).

[0093] In another embodiment, the Gag-cargo fusion (*e.g.*, Gag-BE) further comprises one or more nuclear export signals at one or more locations along the length of the fusion polypeptide protein which may be joined by a cleavable linker such that during VLP assembly in the producer cell, the Gag-cargo fusions (due to presence of competing NLS signals) do not accumulate in the nucleus of the producer cells but instead are available in the cytoplasm to undergo the VLP assembly process at the cell membrane. Once inside the matured VLPs following release from the producer cell, the NES may be cleaved by Gag-Pro-Pol thereby separating the cargo (*e.g.*, napDNAbp or a BE) from the NES. Upon delivery to a recipient cell, therefore, the cargo (*e.g.*, napDNAbp or BE, typically flanked with one or more NLS elements) will not comprise an NES element, which may otherwise prohibit the transport of the cargo into the nucleus and hinder gene editing activity. This is exemplified as v.3 VLPs described herein (or “third generation” VLPs).

[0094] In another embodiment, as demonstrated by v.4 VLPs (or “fourth generation” VLPs) described herein, the inventors found an optimized stoichiometry ratio of Gag-cargo fusion to Gag-Pro-Pol fusion protein which balances the amount of Gag-cargo available to be packaged into VLPs with the amount of retrovirus protease (the “Pro” in the Gag-Pro-Pol fusion) required for VLP maturation. In one embodiment, the optimized ratio of Gag-cargo fusion to Gag-Pro-Pol fusion protein is achieved by the appropriate ratio of plasmids encoding each component which are transiently delivered to the producer cells. In one embodiment, to modulate the stoichiometry of the Gag-cargo fusion to Gag-Pro-Pol fusion, the ratio of the plasmid encoding Gag-cargo (*e.g.*, Gag-3xNES-ABE8e) to wild-type MMLV gag-pro-pol plasmids transfected for VLP production was varied. It was found that increasing the amount of gag-cargo plasmid beyond the original proportion used for producing v3.4 BE-eVLPs (38% Gag-cargo plasmid, and 62% gag-pro-pol plasmid) did not improve editing efficiencies (**FIG. 2G**). Decreasing the proportion of gag-cargo plasmid from 38% to 25% modestly improved editing efficiencies (**FIG. 2G**). However, further decreasing the proportion of gag-cargo plasmid below 25% reduced editing efficiencies (**FIG. 2G**). These results are consistent with a model in which an optimal gag-cargo:gag-pro-pol stoichiometry balances the amount of gag-cargo available to be packaged into VLPs with the amount of MMLV protease (the “pro” in gag-pro-pol) required for VLP maturation. In one embodiment, the results of this final round of optimization revealed a fourth-generation (v4) BE-eVLP formulation (**FIG. 2G**), which combines the optimal gag-BE:gag-pro-pol stoichiometry (25% gag-BE) with the v3.4 BE-eVLP architecture. In some embodiments, the ratio of gag-pro-polyprotein to gag-cargo is approximately 10:1, approximately 9:1, approximately 8:1, approximately 7:1, approximately 6:1, approximately 5:1, approximately 4:1, approximately 3:1, approximately 2:1, approximately 1.5:1, approximately 1:1, or approximately 0.5:1

[0095] Accordingly, in one aspect, the present disclosure provides an eVLP comprising an (a) envelope, and (b) a multi-protein core, wherein the envelope comprises a lipid membrane (*e.g.*, a lipid mono- or bi-layer membrane) and a viral envelope glycoprotein, and wherein the multi-protein core comprises a Gag (*e.g.*, a retroviral Gag), a group-specific antigen (gag) protease (pro) polyprotein (*i.e.*, “Gag-Pro-Pol”), and a fusion protein comprising a Gag-cargo (*e.g.*, Gag-napDNAbp or Gag-BE). In various embodiments, the Gag-cargo may comprise a ribonucleoprotein cargo, *e.g.*, a napDNAbp or a BE complexed with a guide RNA. In still further embodiments, the Gag-cargo (*e.g.*, Gag fused to a napDNAbp or a BE) may comprise one or more NLS sequences and/or one or more NES sequences to regulate the cellular

location of the cargo in a cell. An NLS sequence will facilitate the transport of the cargo into the cell's nucleus to facilitate editing. A NES will do the opposite, *i.e.*, transport the cargo out from the nucleus, and/or prevent the transport of the cargo into the nucleus. In certain embodiments, the NES may be coupled to the fusion protein by a cleavable linker (*e.g.*, a protease linker) such that during assembly in a producer cell, the NES signals operates to keep the cargo in the cytoplasm and available for the packaging process. However, once matured VLPs are budded out or released from a producer cell in a mature form, the cleavable linker joining the NES may be cleaved, thereby removing the association of NES with the cargo. Thus, without an NES, the cargo will translocate to the nucleus with its NLS sequences, thereby facilitating editing. Various napDNAbps may be used in the systems of the present disclosure. In some embodiments, the napDNAbp is a Cas9 protein (*e.g.*, a Cas9 nickase, dead Cas9 (dCas9), or another Cas9 variant as described herein). In some embodiments, the Cas9 protein is bound to a guide RNA (gRNA). The fusion protein may further comprise other protein domains, such as effector domains. In some embodiments, the fusion protein further comprises a deaminase domain (*e.g.*, an adenosine deaminase domain or a cytosine deaminase domain). In certain embodiments, the fusion protein comprises a base editor, such as ABE8e, or any of the other base editors described herein or known in the art.

[0096] In some embodiments, the fusion protein comprises more than one NES (*e.g.*, two NES, three NES, four NES, five NES, six NES, seven NES, eight NES, nine NES, or ten or more NES). In certain embodiments, the fusion protein further comprises a nuclear localization sequence (NLS), or more than one NLS (*e.g.*, two NLS, three NLS, four NLS, five NLS, six NLS, seven NLS, eight NLS, nine NLS, or ten or more NLS). In certain embodiments, the fusion protein may comprise at least one NES and one NLS.

[0097] The Gag-cargo fusion proteins described herein comprise one or more cleavable linkers. In one embodiment, the Gag-cargo fusion proteins comprise a cleavable linker joining the Gag to the cargo, such that once the Gag-cargo fusion has been packaged in mature VLPs (which will also contain the Gag-Pro-Pol, the protease activity can cleave the Gag-cargo cleavable linker, thereby releasing the cargo. In some embodiments, a cleavable linker may also be provided in such a location such that when the cleavable linker is cleaved (*e.g.*, by the Gag-Pro-Pol protein), the NES is separated away from the cargo protein. Such an arrangement of the fusion protein allows the fusion protein to be exported from the nucleus of a producing cell during BE-VLP production, and the NES can later be cleaved from the

fusion protein after delivery to a target cell, or prior to delivery to the target cell but after packaging into the VLP, releasing the BE and allowing it to enter the nucleus of the target cell. In some embodiments, the cleavable linker comprises a protease cleavage site (*e.g.*, a Moloney murine leukemia virus (MMLV) protease cleavage site or a Friend murine leukemia virus (FMLV) protease cleavage site). Various protease cleavage sites can be used in the fusion proteins of the present disclosure. In certain embodiments, the protease cleavage site comprises the amino acid sequence TSTLLMENSS (SEQ ID NO: 1), PRSSLYPALTP (SEQ ID NO: 2), VQALVLTQ (SEQ ID NO: 3), PLQVLTNIERR (SEQ ID NO: 4), or an amino acid sequence at least 90% identical to any one of SEQ ID NOs: 1-4. In some embodiments, the protease cleavage site comprises the amino acid sequence of any one of SEQ ID NOs: 1-4 comprising one mutation, two mutations, three mutations, four mutations, five mutations, or more than five mutations relative to one of SEQ ID NOs: 1-4. In some embodiments, the cleavable linker of the fusion protein is cleaved by the protease of the gag-pro polyprotein. In certain embodiments, the cleavable linker of the fusion protein is not cleaved by the protease of the gag-pro polyprotein until the BE-VLP has been assembled and delivered into a target cell. In some embodiments, the gag-pro polyprotein of the BE-VLPs described herein comprises an MMLV gag-pro polyprotein or an FMLV gag-pro polyprotein. In some embodiments, the gag nucleocapsid protein of the fusion protein in the BE-VLPs described herein comprises an MMLV gag nucleocapsid protein or an FMLV gag nucleocapsid protein.

[0098] In certain embodiments, the fusion protein comprises the following non-limiting structures:

[gag nucleocapsid protein]-[1X-3X NES]-[cleavable linker]-[NLS]-[deaminase domain]-[napDNAbp]-[NLS], wherein]-[comprises an optional linker (*e.g.*, an amino acid linker, or any of the linkers provided herein);

[1X-3X NES]-[gag nucleocapsid protein]-[cleavable linker]-[NLS]-[deaminase domain]-[napDNAbp]-[NLS], wherein]-[comprises an optional linker (*e.g.*, an amino acid linker, or any of the linkers provided herein); or

[gag nucleocapsid protein]-[1X-3X NES]-[cleavable linker]-[NLS]-[deaminase domain]-[napDNAbp]-[NLS]-[cleavable linker]-[1X-3X NES], wherein]-[comprises an optional linker (*e.g.*, an amino acid linker, or any of the linkers provided herein).

[0099] In embodiments in which the cleavable linker has been cleaved by the protease within the VLP, the VLP may comprise a fusion protein comprising the structure [gag nucleocapsid protein]-[1X-3X NES], and a free napDNAbp or base editor. In certain embodiments, the base editor comprises the structure [NLS]-[deaminase domain]-[napDNAbp]-[NLS], wherein each instance of]-[comprises an optional linker (*e.g.*, an amino acid linker, or any of the linkers provided herein).

[0100] In some embodiments, any of the constructs above comprise 3X NES.

[0101] The eVLPs (*e.g.*, the BE-VLPs) provided by the present disclosure comprise an outer encapsulation layer (or envelope layer) comprising a viral envelope glycoprotein. Any viral envelope glycoprotein described herein, or known in the art, may be used in the BE-VLPs of the present disclosure. In some embodiments, the viral envelope glycoprotein is an adenoviral envelope glycoprotein, an adeno-associated viral envelope glycoprotein, a retroviral envelope glycoprotein, or a lentiviral envelope glycoprotein. In certain embodiments, the viral envelope glycoprotein is a retroviral envelope glycoprotein. In some embodiments, the viral envelope glycoprotein is a vesicular stomatitis virus G protein (VSV-G), a baboon retroviral envelope glycoprotein (BaEVRless), a FuG-B2 envelope glycoprotein, an HIV-1 envelope glycoprotein, or an ecotropic murine leukemia virus (MLV) envelope glycoprotein. In some embodiments, the viral envelope glycoprotein targets the system to a particular cell type (*e.g.*, immune cells, neural cells, retinal pigment epithelium cells, *etc.*). For example, using different envelope glycoproteins in the eVLPs described herein may alter their cellular tropism, allowing the BE-VLPs to be targeted to specific cell types. In some embodiments, the viral envelope glycoprotein is a VSV-G protein, and the VSV-G protein targets the system to retinal pigment epithelium (RPE) cells. In some embodiments, the viral envelope glycoprotein is an HIV-1 envelope glycoprotein, and the HIV-1 envelope glycoprotein targets the system to CD4+ cells. In some embodiments, the viral envelope glycoprotein is a FuG-B2 envelope glycoprotein, and the FuG-B2 envelope glycoprotein targets the system to neurons.

[0102] It will be appreciated that general methods are known in the art for producing viral vector particles, which generally contain coding nucleic acids of interest, and may also be used for producing the virus-derived particles according to the present invention, which do not contain coding nucleic acids of interest but instead are designed to deliver a protein cargo (*e.g.*, a BE RNP).

[0103] Conventional viral vector particles encompass retroviral, lentiviral, adenoviral, and adeno-associated viral vector particles that are well known in the art. For a review of various

viral vector particles that may be used, the one skilled in the art may notably refer to Kushnir *et al.* (2012, *Vaccine*, Vol. 31: 58-83), Zeltons (2013, *Mol Biotechnol*, Vol. 53: 92-107), Ludwig *et al.* (2007, *Curr Opin Biotechnol*, Vol. 18(no 6): 537-55) and Naskalaska *et al.* (2015, Vol. 64 (no 1): 3-13). Further, references to various methods using virus-derived particles for delivering proteins to cells are found by the one skilled in the art in the article of Maetzig *et al.* (2012, *Current Gene Therapy*, Vol. 12: 389-409), as well as the article of Kaczmarczyk *et al.* (2011, *Proc Natl Acad Sci USA*, Vol. 108 (no 41): 16998-17003).

[0104] Generally, a virus-like particle that is used according to the present disclosure, which virus-like particle may also be termed “virus-derived particle,” is formed by one or more virus-derived structural protein(s) and/or one more virus-derived envelope protein.

[0105] A virus-like particle that is used according to the present invention is replication incompetent in a host cell wherein it has entered.

[0106] In preferred embodiments, a virus-like particle is formed by one or more retrovirus-derived structural protein(s) and optionally one or more virus-derived envelope protein(s).

[0107] In preferred embodiments, the virus-derived structural protein is a retroviral Gag protein or a peptide fragment thereof. As it is known in the art, Gag and Gag/pol precursors are expressed from full length genomic RNA as polyproteins, which require proteolytic cleavage, mediated by the retroviral protease (PR), to acquire a functional conformation. Further, Gag, which is structurally conserved among the retroviruses, is composed of at least three protein units: matrix protein (MA), capsid protein (CA) and nucleocapsid protein (NC), whereas Pol consists of the retroviral protease, (PR), the retrotranscriptase (RT), and the integrase (IN).

[0108] In some embodiments, a virus-derived particle comprises a retroviral Gag protein but does not comprise a Pol protein.

[0109] As it is known in the art, the host range of retroviral vector, including lentiviral vectors, may be expanded or altered by a process known as pseudotyping. Pseudotyped lentiviral vectors consist of viral vector particles bearing glycoproteins derived from other enveloped viruses. Such pseudotyped viral vector particles possess the tropism of the virus from which the glycoprotein is derived.

[0110] In some embodiments, a virus-like particle is a pseudotyped virus-like particle comprising one or more viral structural protein(s) or viral envelope protein(s) imparting a tropism to the said virus-like particle for certain eukaryotic cells. A pseudotyped virus-like particle as described herein may comprise, as the viral protein used for pseudotyping, a viral

envelope protein selected in a group comprising VSV-G protein, Measles virus HA protein, Measles virus F protein, Influenza virus HA protein, Moloney virus MLV-A protein, Moloney virus MLV-E protein, Baboon Endogenous retrovirus (BAEV) envelope protein, Ebola virus glycoprotein, and foamy virus envelope protein, or a combination of two or more of these viral envelope proteins.

[0111] A well-known illustration of pseudotyping viral vector particles consists of the pseudotyping of viral vector particles with the vesicular stomatitis virus glycoprotein (VSV-G). For the pseudotyping of viral vector particles, one skilled in the art may notably refer to Yee *et al.* (1994, *Proc Natl Acad Sci, USA*, Vol. 91: 9564-9568) Cronin *et al.* (2005, *Curr Gene Ther*, Vol. 5(no 4): 387-398), which are incorporated herein by reference.

[0112] For producing virus-like particles, and more precisely VSV-G pseudotyped virus-like particles, for delivering protein(s) of interest into target cells, one skilled in the art may refer to Mangeot *et al.* (2011, *Molecular Therapy*, Vol. 19 (no 9): 1656-1666).

[0113] In some embodiments, a virus-like particle further comprises a viral envelope protein, wherein either (i) the said viral envelope protein originates from the same virus as the viral structural protein, *e.g.*, originates from the same virus as the viral Gag protein, or (ii) the said viral envelope protein originates from a virus distinct from the virus from which originates the viral structural protein, *e.g.*, originates from a virus distinct from the virus from which originates the viral Gag protein.

[0114] As is readily understood by one skilled in the art, a virus-like particle that is used according to the disclosure may be selected in a group comprising Moloney murine leukemia virus-derived vector particles, Bovine immunodeficiency virus-derived particles, Simian immunodeficiency virus-derived vector particles, Feline immunodeficiency virus-derived vector particles, Human immunodeficiency virus-derived vector particles, Equine infection anemia virus-derived vector particles, Caprine arthritis encephalitis virus-derived vector particle, Baboon endogenous virus-derived vector particles, Rabies virus-derived vector particles, Influenza virus-derived vector particles, Norovirus-derived vector particles, Respiratory syncytial virus-derived vector particles, Hepatitis A virus-derived vector particles, Hepatitis B virus-derived vector particles, Hepatitis E virus-derived vector particles, Newcastle disease virus-derived vector particles, Norwalk virus-derived vector particles, Parvovirus-derived vector particles, Papillomavirus-derived vector particles, Yeast retrotransposon-derived vector particles, Measles virus-derived vector particles, and bacteriophage-derived vector particles.

[0115] In particular, a virus-like particle that is used according to the invention is a retrovirus-derived particle. Such retrovirus may be selected among Moloney murine leukemia virus, Bovine immunodeficiency virus, Simian immunodeficiency virus, Feline immunodeficiency virus, Human immunodeficiency virus, Equine infection anemia virus, and Caprine arthritis encephalitis virus.

[0116] In another embodiment, a virus-like particle that is used according to the disclosure is a lentivirus-derived particle. Lentiviruses belong to the retroviruses family, and have the unique ability of being able to infect non-dividing cells.

[0117] Such lentivirus may be selected among Bovine immunodeficiency virus, Simian immunodeficiency virus, Feline immunodeficiency virus, Human immunodeficiency virus, Equine infection anemia virus, and Caprine arthritis encephalitis virus.

[0118] For preparing Moloney murine leukemia virus-derived vector particles, one skilled in the art may refer to the methods disclosed by Sharma *et al.* (1997, *Proc Natl Acad Sci USA*, Vol. 94: 10803+-10808), Guibingua *et al.* (2002, *Molecular Therapy*, Vol. 5(no 5): 538-546), which are incorporated herein by reference. Moloney murine leukemia virus-derived (MLV-derived) vector particles may be selected in a group comprising MLV-A-derived vector particles and MLV-E-derived vector particles.

[0119] For preparing Bovine Immunodeficiency virus-derived vector particles, one skilled in the art may refer to the methods disclosed by Rasmussen *et al.* (1990, *Virology*, Vol. 178(no 2): 435-451), which is incorporated herein by reference.

[0120] For preparing Simian immunodeficiency virus-derived vector particles, including VSV-G pseudotyped SIV virus-derived particles, one skilled in the art may notably refer to the methods disclosed by Mangeot *et al.* (2000, *Journal of Virology*, Vol. 71(no 18): 8307-8315), Negre *et al.* (2000, *Gene Therapy*, Vol. 7: 1613-1623) Mangeot *et al.* (2004, *Nucleic Acids Research*, Vol. 32 (no 12), e102), which are incorporated herein by reference.

[0121] For preparing Feline Immunodeficiency virus-derived vector particles, one skilled in the art may notably refer to the methods disclosed by Saenz *et al.* (2012, *Cold Spring Harb Protoc*, (1): 71-76; 2012, *Cold Spring Harb Protoc*, (1): 124-125; 2012, *Cold Spring Harb Protoc*, (1): 118-123), which are incorporated herein by reference.

[0122] For preparing Human immunodeficiency virus-derived vector particles, one skilled in the art may notably refer to the methods disclosed by Jalaguier *et al.* (2011, *PlosOne*, Vol. 6(no 11), e28314), Cervera *et al.* (*J Biotechnol*, Vol. 166(no 4): 152-165), Tang *et al.* (2012, *Journal of Virology*, Vol. 86(no 14): 7662-7676), which are incorporated herein by reference.

[0123] For preparing Equine infection anemia virus-derived vector particles, one skilled in the art may notably refer to the methods disclosed by Olsen (1998, *Gene Ther*, Vol. 5(no 11): 1481-1487), which are incorporated herein by reference.

[0124] For preparing Caprine arthritis encephalitis virus-derived vector particles, one skilled in the art may notably refer to the methods disclosed by Mselli-Lakhal *et al.* (2006, *J Virol Methods*, Vol. 136(no 1-2): 177-184), which are incorporated herein by reference.

[0125] For preparing Baboon endogenous virus-derived vector particles, one skilled in the art may notably refer to the methods disclosed by Girard-Gagnepain *et al.* (2014, *Blood*, Vol. 124(no 8): 1221-1231), which is incorporated herein by reference.

[0126] For preparing Rabies virus-derived vector particles, one skilled in the art may notably refer to the methods disclosed by Kang *et al.* (2015, *Viruses*, Vol. 7: 1134-1152, doi:10.3390/v7031134), Fontana *et al.* (2014, *Vaccine*, Vol. 32(no 24): 2799-27804) or to the PCT application published under no WO 2012/0618, which is incorporated herein by reference.

[0127] For preparing Influenza virus-derived vector particles, one skilled in the art may notably refer to the methods disclosed by Quan *et al.* (2012, *Virology*, Vol. 430: 127-135) and to Latham *et al.* (2001, *Journal of Virology*, Vol. 75(no 13): 6154-6155), which is incorporated herein by reference.

[0128] For preparing Norovirus-derived vector particles, one skilled in the art may notably refer to the methods disclosed by Tomé-Amat *et al.*, (2014, *Microbial Cell Factories*, Vol. 13: 134-142), which is incorporated herein by reference.

[0129] For preparing Respiratory syncytial virus-derived vector particles, one skilled in the art may notably refer to the methods disclosed by Walpita *et al.* (2015, *PlosOne*, DOI: 10.1371/journal.pone.0130755), which is incorporated herein by reference.

[0130] For preparing Hepatitis B virus-derived vector particles, one skilled in the art may notably refer to the methods disclosed by Hong *et al.* (2013, Vol. 87(no 12): 6615-6624), which is incorporated herein by reference.

[0131] For preparing Hepatitis E virus-derived vector particles, one skilled in the art may notably refer to the methods disclosed by Li *et al.* (1997, *Journal of Virology*, Vol. 71(no 10): 7207-7213), which is incorporated herein by reference.

[0132] For preparing Newcastle disease virus-derived vector particles, one skilled in the art may notably refer to the methods disclosed by Murawski *et al.* (2010, *Journal of Virology*, Vol. 84(no 2): 1110-1123), which is incorporated herein by reference.

[0133] For preparing Norwalk virus-derived vector particles, one skilled in the art may notably refer to the methods disclosed by Herbst-Kralovetz *et al.* (2010, *Expert Rev Vaccines*, Vol. 9(no 3): 299-307), which is incorporated herein by reference.

[0134] For preparing Parvovirus-derived vector particles, one skilled in the art may notably refer to the methods disclosed by Ogasawara *et al.* (2006, *In Vivo*, Vol. 20: 319-324), which is incorporated herein by reference.

[0135] For preparing Papillomavirus-derived vector particles, one skilled in the art may notably refer to the methods disclosed by Wang *et al.* (2013, *Expert Rev Vaccines*, Vol. 12(no 2): doi:10.1586/erv.12.151), which is incorporated herein by reference.

[0136] A virus-like particle that is used herein comprises a Gag protein, and most preferably a Gag protein originating from a virus selected from a group consisting of Rous Sarcoma Virus (RSV), Feline Immunodeficiency Virus (FIV), Simian Immunodeficiency Virus (SIV), Moloney Leukemia Virus (MLV), and Human Immunodeficiency Viruses (HIV-1 and HIV-2), especially Human Immunodeficiency Virus of type 1 (HIV-1).

[0137] In some embodiments, a virus-like particle may also comprise one or more viral envelope protein(s). The presence of one or more viral envelope protein(s) may impart to the said virus-derived particle a more specific tropism for the cells which are targeted, as it is known in the art. The one or more viral envelope protein(s) may be selected from a group consisting of envelope proteins from retroviruses, envelope proteins from non-retroviral viruses, and chimeras of these viral envelope proteins with other peptides or proteins. An example of a non-lentiviral envelope glycoprotein of interest is the lymphocytic choriomeningitis virus (LCMV) strain WE54 envelope glycoprotein. These envelope glycoproteins increase the range of cells that can be transduced with retroviral derived vectors.

napDNAbp

[0138] In various embodiments, the BE-VLPs disclosed herein, as well as the fusion proteins that make up the core component of the presently described BE-VLPs, comprise a nucleic acid programmable DNA binding protein (napDNAbp).

[0139] In various embodiments, the BE-VLPs and fusion proteins may include a napDNAbp domain having a wild type Cas9 sequence, including, for example the canonical *Streptococcus pyogenes* Cas9 sequence of SEQ ID NO: 13, shown as follows:

Description	Sequence	SEQ ID NO:
SpCas9	MDKKYSIGLDIGTNSVGWAVITDEYKVPSSKFKVLGNTDRHSIKKNLIGALLF DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLLEESFLV	13

<i>Streptococcus pyogenes</i> M1 SwissProt Accession No. Q99ZW2 Wild type	EEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHA MIKFRGHFLIEGDLNPDNSDVDFKLFQVLVQTYNQLFEENPINASGVDAKILSA RLSKSRRLLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSK DTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMI KRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKF IKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDF YPFLKDNREKIEKILTFRIPIYVGPLARGNSRFAWMTRKSEETITPWNFEEVVD KGASAQSFIERMTNFDKNLPNEKVLPHKSLLYEYFTVYNELTKVKYVTEGMR KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN ASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTFEDREMIEERLKYAHLF DDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQM LIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVK VMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIIEGKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDN KVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAER GGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITL KSKLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVY GDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIE TNGETGEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSD KLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIM ERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLELENGRKRMLASAGELQK GNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEF SKRVILADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTT IDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD	
---	--	--

[0140] In other embodiments, the BE-VLPs and fusion proteins may include a napDNAbp domain having a modified Cas9 sequence, including, for example the nickase variant of *Streptococcus pyogenes* Cas9 of SEQ ID NO: 14 having an H840A substitution relative to the wild type SpCas9 (of SEQ ID NO: 13), shown as follows:

Cas9 nickase	MDKKYSIGLDIGTNSVGWAVITDEYKVPSSKFKVLGNTDRHSIKKNLIGAL LFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLLEE SFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIY LALAHMIKFRGHFLIEGDLNPDNSDVDFKLFQVLVQTYNQLFEENPINASGVDA AKILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAE DAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEI TKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYI DGGASQEEFYKFIPKILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIH LGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYVGPLARGNSRFAWMTR KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSLLYEYFT VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYF KKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTL TLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIAN LAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNS RERMKRIIEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL DINRLSDYDVAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKH VAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVRINNY HHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVR KVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSEKNPIDFLEAKGY KEVKKDLIIKLPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFLYL ASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKV LSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD	SEQ ID NO: 14
--------------	--	------------------

[0141] The BE-VLPs and fusion proteins described herein may include any of the modified Cas9 sequences described above, or any variant thereof having at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence identity thereto. In some embodiments, the base editor fusion proteins described herein include any of the following other wild type SpCas9 sequences, which may be modified with one or more of the mutations described herein at corresponding amino acid positions:

Description	Sequence
SpCas9 <i>Streptococcus pyogenes</i> MGAS1882 wild type NC_017053.1	ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAAATAGCGTCGGATGGGCG GTGATCACTGATGATTATAAGGTTCCGTCTAAAAAGTTCAAGGTTCTGGGAAATACA GACCGCCACAGTATCAAAAAAATCTTATAGGGGCTCTTTTATTTGGCAGTGGAGAG ACAGCGGAAGCGACTCGTCTCAAACGGACAGCTCGTAGAAGGTATACACGTCGGAA GAATCGTATTTGTTATCTACAGGAGATTTTTTCAAATGAGATGGCGAAAGTAGATGAT AGTTTCTTTCATCGACTTGAAGAGTCTTTTTTGGTGGGAAGAAGACAAGAAGCATGA ACGTCATCCTATTTTTGGGAAATATAGTAGATGAAGTTGCTTATCATGAGAAATATCCAA CTATCTATCATCTGCGAAAAAATTTGGCAGATTCTACTGATAAAGCGGATTTGCGCTT AATCTATTTGGCCTTAGCGCATATGATTAAGTTTCGTGGTCATTTTTTGATTGAGGGAG ATTTAAATCCTGATAATAGTGATGTGGACAACTATTTATCCAGTTGGTACAAATCTAC AATCAATTATTTGAAGAAAACCTATTAAACGCAAGTAGAGTAGAGTAAAGCGATT CTTTCTGCACGATTGAGTAAATCAAGACGATTAGAAAATCTCATTGCTCAGCTCCCC GGTGAGAAGAGAAATGGCTTGTGTTGGGAATCTCATTGCTTTGTCATTGGGATTGACC CCTAATTTTAAATCAAATTTTGATTGTTGGCAGAAGATGCTAAATTACAGCTTTCAAAAG ATACTTACGATGATGATTTAGATAATTTATTTGGCGCAAATTGGAGATCAATATGCTGAT TTGTTTTTGGCAGCTAAGAATTTATCAGATGCTATTTTACTTTTACAGATATCCTAAGAGT AAATAGTGAAATAACTAAGGCTCCCCTATCAGCTTCAATGATTAAGCGCTACGATGAA CATCATCAAGACTTGACTCTTTTAAAAGCTTTAGTTTCGACAACAACCTCCAGAAAAG TATAAAGAAATCTTTTTTGATCAATCAAAAAACGGATATGCAGGTTATATTGATGGGG GAGCTAGCCAAGAAGAATTTTATAAATTTATCAAACCAATTTTAGAAAAAATGGATG GTACTGAGGAATTATTTGGTGAACTAAATCGTGAAGATTTGCTGCGCAAGCAACGGA CCTTTGACAACGGCTCTATTTCCCATCAAATTCACCTGGGTGAGCTGCATGCTATTTT GAGAAGACAAGAAGACTTTTATCCATTTTAAAAGACAATCGTGAGAAGATTGAAA AAATCTTGACTTTTCGAATTCCTTATTATGTTGGTCCATTGGCGCGTGGCAATAGTCG TTTTGCATGGATGACTCGGAAGTCTGAAGAAACAATTACCCCATGGAATTTTGAAGA AGTTGTCGATAAAGGTGCTTCAGCTCAATCATTTATTGAACGCATGACAACTTTGAT AAAAATCTTCCAAATGAAAAAGTACTACCAAAACATAGTTTGTCTTATGAGTATTTTA CGGTTTATAACGAATTGACAAAGGTCAAATATGTTACTGAGGGAATGCGAAAACCAG CATTTCTTTCAGGTGAACAGAAAGAAAGCCATTGTTGATTTACTCTTCAAAACAAATC GAAAAGTAACCGTTAAGCAATTAAAAGAAGATTATTTCAAAAAAATAGAATGTTTTG ATAGTGTTGAAATTTACAGGAGTTGAAGATAGATTTAATGCTTCATTAGGCGCCTACCA TGATTTGCTAAAAATTATTAAGATAAAGATTTTTTGGATAATGAAGAAAATGAAGAT ATCTTAGAGGATATTGTTTTAACATTGACCTTATTTGAAGATAGGGGGATGATTGAGG AAAGACTTAAAACATATGCTCACCTCTTTGATGATAAGGTGATGAAACAGCTTAAAC GTCGCCGTTTACTGTTTGGGGACGTTTGTCTCGAAAATTGATTAATGTTATTAGGGA TAAGCAATCTGGCAAAACAATATTAGATTTTTTGAATCAGATGGTTTTGCCAATCGC AATTTTATGCAGCTGATCCATGATGATAGTTTGAATTTAAAGAAGATATTCAAAAAG CACAGGTGTCTGGACAAGGCCATAGTTTACATGAACAGATTGCTAACTTAGCTGGCA GTCTTGCTATTAAAAAAGGTATTTTACAGACTGTAAAAATGTTGATGAACTGGTCA AAGTAATGGGGCATAAGCCAGAAAATATCGTTATTGAAATGGCACGTGAAAATCAGA CAACTCAAAGGGCCAGAAAATTCGCGAGAGCGTATGAAACGAATCGAAGAAGG TATCAAAGAATTAGGAAGTCAGATTCTTAAAGAGCATCCTGTTGAAAATACTCAATT GCAAAATGAAAAGCTCTATCTCTATTATCTACAAAATGGAAGAGACATGTATGTGGA CCAAGAATTAGATATTAATCGTTTAAAGTGATTATGATGTCGATCACATTGTTCCACAAA GTTTCATTAAAGACGATTCAATAGACAATAAGGTACTAACGCGTTCTGATAAAAATCG TGGTAAATCGGATAACGTTCCAAGTGAAGAAGTAGTCAAAAAGATGAAAAACTATT

	GGAGACAACCTTCTAAACGCCAAGTTAATCACTCAACGTAAGTTTGATAATTTAACGA AAGCTGAACGTGGAGGTTTGTAGTGAACCTTGATAAAGCTGGTTTTATCAAACGCCAAT TGGTTGAAACTCGCCAAATCACTAAGCATGTGGCACAATTTTGGATAGTCGCATGA ATACTAAATACGATGAAAATGATAAACTTATTCGAGAGGTTAAAGTGATTACCTTAAA ATCTAAATTAGTTTCTGACTTCCGAAAAGATTTCCAATTCTATAAAGTACGTGAGATT AACAATTACCATCATGCCCATGATGCGTATCTAAATGCCGTCGTTGGAACCTGCTTTGA TTAAGAAATATCCAAAACCTTGAATCGGAGTTTGTCTATGGTGATTATAAAGTTTATGAT GTTTCGTAAAATGATTGCTAAGTCTGAGCAAGAAATAGGCAAAGCAACCGCAAAAATAT TTCTTTTACTCTAATATCATGAACCTTCTTCAAAAACAGAAATTACACTTGCAAATGGAG AGATTTCGCAAACGCCCTCTAATCGAAACTAATGGGGAAACTGGAGAAATTGTCTGG GATAAAGGGCGAGATTTTGGCACAGTGCGCAAAGTATTGTCCATGCCCAAGTCAAT ATTGTCAAGAAAACAGAAAGTACAGACAGGCGGATTCTCCAAGGAGTCAATTTTACC AAAAAGAAATTTCGGACAAGCTTATTGCTCGTAAAAAAGACTGGGATCCAAAAAAT ATGGTGGTTTTGATAGTCCAACGGTAGCTTATTCAGTCCTAGTGTTGCTAAGGTGGA AAAAGGGAAATCGAAGAAGTTAAATCCGTTAAAGAGTTACTAGGGATCACAAATTAT GGAAAGAAGTTCCTTTGAAAAAATCCGATTGACTTTTTAGAAAGCTAAAGGATATAA GGAAGTTAAAAAAGACTTAATCATTAAACTACCTAAATATAGTCTTTTTGAGTTAGAA AACGGTCGTAAACGGATGCTGGCTAGTGCCGGAGAATTACAAAAAGGAAATGAGCT GGCTCTGCCAAGCAAATATGTGAATTTTTTATATTTAGCTAGTCATTATGAAAAGTTGA AGGGTAGTCCAGAAGATAACGAACAAAAACAATTGTTTGTGGAGCAGCATAAGCAT TATTTAGATGAGATTATTGAGCAAATCAGTGAATTTTCTAAGCGTGTTATTTTAGCAGA TGCCAATTTAGATAAAGTTCTTAGTGCAATAACAAACATAGAGACAAACCAATACGT GAACAAGCAGAAAATATTATTCATTTATTTACGTTGACGAATCTTGGAGCTCCCGCTG CTTTTAAATATTTGATACAACAATTGATCGTAAACGATATACGTCTACAAAAGAAGT TTTAGATGCCACTCTTATCCATCAATCCATCACTGGTCTTTATGAAACACGCATTGATT TGAGTCAGCTAGGAGGTGACTGA (SEQ ID NO: 15)
SpCas9 <i>Streptococcus</i> <i>pyogenes</i> MGAS1882 wild type NC_017053.1	MDKKYSIGLDIGTNSVGVAVITDDYKVPSKKFKVLGNTDRHSIKKNLIGALLFGSGETA EATRLKRTARRRYTRRKNRICYLQEISNEMAKVDDSFHRLSEESFLVEEDKKHERHPIF GNIVDEVAYHEKYPTIYHLRKKLADSTDKADLRILIYLAHAMIKFRGHFLIEGDLNPDN SDVDKLFQILVQIYNQLFEENPINASRVDAKILSARLSKSRRLLENLIAQLPGEKRNGLF GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLS DAILLSILRVNSEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLG ELHAILRRQEDFYPLKDNREKIEKILTFRIPIYVGPLARGNSRFAWMTRKSEETITPWN FEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSLLYEYFTVYNELTKVKYVTEGMR KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGAY HDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDRGMIEERLKYAHLFFDKVMKQLKRR RYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFMLIHDDSLTFKEDIQKAQV SGQGHSLHEQIANLAGSPAIIKKGILQTVKIVDELVKVMGHKPENIVIMARENQTTQKG QKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDN RLSDYDVDHIVPQSFIKDDSIDNKVLTNRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNA KLITQRKFDNLTKAERGGLSELDKAGFIKRLVETRQITKHVAQVILDSRMNTKYDENDK LIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESE FVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNG ETGEIVWDKGRDFATVRKVLSPQVNVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEA KGYKEVKKDLIKLPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFYLLASHY EKLKGSPEDEQKQLFVEQHKHYLDEIEQISEFSKRVILADANLDKVL SAYNKH RD KPI REQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQ LGGD (SEQ ID NO: 16)
SpCas9 <i>Streptococcus</i> <i>pyogenes</i> wild type SWBC2D7W01 4	ATGGATAAAAAGTATTCTATTGGTTTAGACATCGGCACTAATTCCGTTGGATGGGCTG TCATAACCGATGAATACAAAGTACCTTCAAAGAAATTTAAGGTGTTGGGGAACACAG ACCGTCATTCGATTAAAAAGAATCTTATCGGTGCCCTCCTATTGATAGTGGCGAAAC GGCAGAGGCGACTCGCCTGAAACGAACCGCTCGGAGAAGGTATACACGTGCGCAAG AACC GAATATGTTACTTACAAGAAATTTTTAGCAATGAGATGGCCAAAGTTGACGAT TCTTTCTTTCACCGTTTGGAAAGAGTCCTTCTTGTGCAAGAGGACAAGAAACATGAA CGGCACCCCATCTTTGGAAACATAGTAGATGAGGTGGCATATCATGAAAAGTACCCA ACGATTTATCACCTCAGAAAAAAGCTAGTTGACTCAACTGATAAAGCGGACCTGAG GTTAATCTACTTGGCTCTTGCCCATATGATAAAGTTCCGTGGGCACTTTCTCATTGAG GGTGATCTAAATCCGGACAACCTCGGATGTCGACAAACTGTTTCATCCAGTTAGTACAA

ACCTATAATCAGTTGTTTGAAGAGAACCCTATAAATGCAAGTGGCGTGGATGCGAAG GCTATTCTTAGCGCCCGCCTCTCTAAATCCCGACGGCTAGAAAACCTGATCGCACAA TTACCCGGAGAGAAGAAAAATGGGTGTTTCGGTAACCTTATAGCGCTCTCACTAGGC CTGACACCAAATTTTAAGTCGAACCTTCGACTTAGCTGAAGATGCCAAATTGCAGCTT AGTAAGGACACGTACGATGACGATCTCGACAATCTACTGGCACAATTGGAGATCAG TATGCGGACTTATTTTTGGCTGCCAAAACCTTAGCGATGCAATCCTCCTATCTGACA TACTGAGAGTTAATACTGAGATTACCAAGGCGCCGTTATCCGCTTCAATGATCAAAA GGTACGATGAACATCACCAAGACTTGACACTTCTCAAGGCCCTAGTCCGTCAGCAA CTGCCTGAGAAATATAAGGAAATATTCTTTGATCAGTCGAAAAACGGGTACGCAGGT TATATTGACGGCGGAGCGAGTCAAGAGGAATTCTACAAGTTTATCAAACCCATATTAG AGAAGATGGATGGGACGGAAGAGTTGCTTGTA AAACTCAATCGCGAAGATCTACTG CGAAAGCAGCGGACTTTCGACAACGGTAGCATTCACATCAAATCCACTTAGGCGA ATTGCATGCTATACTTAGAAGGCAGGAGGATTTTTATCCGTTCTCAAAGACAATCGT GAAAAGATTGAGAAAATCCTAACCTTTCGCATACCTTACTATGTGGGACCCCTGGCC CGAGGGAACCTCTCGGTTTCGCATGGATGACAAGAAAGTCCGAAGAAACGATTACTCC ATGGAATTTTGAGGAAGTTGTCGATAAAGGTGCGTCAGCTCAATCGTTTCATCGAGAG GATGACCAACTTTGACAAGAATTTACCGAACGAAAAAGTATTGCCTAAGCACAGTTT ACTTTACGAGTATTTACAGTGTACAATGAACCTCACGAAAGTTAAGTATGTCACTGA GGGCATGCGTAAACCCGCCTTTCTAAGCGGAGAACAGAAGAAAGCAATAGTAGATC TGTTATTCAAGACCAACCGCAAAGTGACAGTTAAGCAATTGAAAGAGGACTACTTTA AGAAAATTGAATGCTTCGATTCTGTGCGAGATCTCCGGGGTAGAAGATCGATTTAATG CGTCACTTGGTACGTATCATGACCTCCTAAAGATAATTAAGATAAGGACTTCCTGGA TAACGAAGAGAATGAAGATATCTTAGAAGATATAGTGTGACTCTTACCCTCTTTGAA GATCGGGAAATGATTGAGGAAAGACTAAAAACATACGCTCACCTGTTTCGACGATAA GGTTATGAAACAGTTAAAGAGGCGTCGCTATACGGGCTGGGGACGATTGTGCGCGA AACTTATCAACGGGATAAGAGACAAGCAAAGTGGTAAAACCTATTCTCGATTTTCTAA AGAGCGACGGCTTCGCCAATAGGAACTTTATGCAGCTGATCCATGATGACTCTTTAA CCTTCAAAGAGGATATACAAAAGGCACAGGTTTCCGGACAAGGGGACTCATTGCAC GAACATATTGCGAATCTTGCTGGTTTCGCCAGCCATCAAAAAGGGCATACTCCAGACA GTCAAAGTAGTGGATGAGCTAGTTAAGGTCATGGGACGTCACAAACCGGAAAACAT TGTAATCGAGATGGCACGCGAAAATCAAACGACTCAGAAGGGGCAAAAAAACAGT CGAGAGCGGATGAAGAGAATAGAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTT AAAGGAGCATCCTGTGGAATAACCCAATTGCAGAACGAGAACTTTACCTCTATTA CCTACAAAATGGAAGGGACATGTATGTTGATCAGGAACCTGGACATAAACCGTTTATC TGATTACGACGTCGATCACATTGTACCCCAATCCTTTTTGAAGGACGATTCAATCGAC AATAAAGTGCTTACACGCTCGGATAAGAACCGAGGGAAAAGTGACAATGTTCCAAG CGAGGAAGTCGTAAAGAAAATGAAGAACTATTGGCGGCAGCTCCTAAATGCGAAAC TGATAACGCAAAGAAAGTTCGATAACTTAACTAAAGCTGAGAGGGGTGGCTTGCT GAACCTTGACAAGGCCGGATTATTAAACGTCAGCTCGTGAAACCCGCCAAATCAC AAAGCTGTTGCACAGATACTAGATTCCCGAATGAATACGAATAACGACAGAACGA TAAGCTGATTGCGGAAGTCAAAGTAATCACTTTAAAGTCAAAATTGGTGTGCGACTT CAGAAAGGATTTTCAATTCTATAAAGTTAGGGAGATAAATACTACCACCATGCGCA CGACGCTTATCTTAATGCCGTGCTAGGGACCGCACTCATTAAGAAATACCCGAAGCT AGAAAGTGAGTTTGTGTATGGTGATTACAAAGTTTATGACGTCCGTAAGATGATCGC GAAAAGCGAACAGGAGATAGGCAAGGCTACAGCCAAATACTTCTTTTATTCTAACAT TATGAATTTCTTTAAGACGGAAATCACTCTGGCAAACGGAGAGATACGCAAACGACC TTTAATTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAAGGGCCGGGACT TCGCGACGGTGAGAAAAGTTTTGTCCATGCCCAAGTCAACATAGTAAAGAAAACCT GAGGTGCAGACCGGAGGGTTTTCAAAGGAATCGATTCTTCCAAAAAGGAATAGTGA TAAGCTCATCGCTCGTAAAAAGGACTGGGACCCGAAAAAGTACGGTGGCTTCGATA GCCCTACAGTTGCCTATTCTGTCTAGTAGTGGCAAAGTTGAGAAGGGAAAATCCA AGAACTGAAGTCAGTCAAAGAATTATTGGGGATAACGATTATGGAGCGCTCGTCTT TTGAAAAGAACCCCATCGACTTCCTTGAGGCGAAAGGTTACAAGGAAGTAAAAAAG GATCTCATAATTAAACTACCAAAGTATAGTCTGTTTGAGTTAGAAAATGGCCGAAAA CGGATGTTGGCTAGCGCCGGAGAGCTTCAAAGGGGAACGAACTCGCACTACCGTC TAAATACGTGAATTTCTGTATTTAGCGTCCCATACGAGAAGTTGAAAGGTTACCT GAAGATAACGAACAGAAGCAACTTTTTGTTGAGCAGCACAAACATTATCTCGACGA AATCATAGAGCAAATTTTCGGAATTCAGTAAGAGAGTCATCCTAGCTGATGCCAATCT GGACAAAGTATTAAGCGCATACAACAAGCACAGGGATAAACCCATACGTGAGCAGG CGGAAAATATTATCCATTTGTTTACTCTTACCAACCTCGGCGCTCCAGCCGCATTCAA
--

	GTATTTTGACACAACGATAGATCGCAAACGATACACTTCTACCAAGGAGGTGCTAGACGCGACACTGATTCACCAATCCATCACGGGATTATATGAAACTCGGATAGATTTGTACAGCTTGGGGGTGACGGATCCCCAAGAAGAAGAGGAAAGTCTCGAGCGACTACAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGGCTGCAGGA (SEQ ID NO: 17)
SpCas9 <i>Streptococcus pyogenes</i> wild type Encoded product of SWBC2D7W014	MDKKYSIGLDIGTNSVGVAVITDEYKVPSSKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEISNEMAKVDDSFHRLVESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADRLIYLALAHMIKFRGHFLIEGDLNPDN SDVDKLFQILVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLS DAILLSLILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKNLREDLLRKQRTFDNGSIPHQIHLG ELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYVGPLARGNSRFAWMTRKSEETITPWN FEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSLLYEYFTVYNELTKVKYVTEGMR KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNANSLGTY HDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRR RYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFMLIHDDSLTFKEDIQKAQV SGQGDLSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVEMARENQTTQ KGQKNSRERMKRIIEGKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELD INRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWRQLL NAKLITQRKFDNLTKAERGGLSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDEN DKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAVVG TALIKKYPKLE SEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET NGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK DWDPKKYGGFDSPTVAYSVLVAKVEKGSKKLKSVKELLGITIMERSSEKPNIDFLE AKGYKEVKKDLIIKLPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH YEKLKGSPEDEQKQLFVEQHKHYLDEHIEQISEFSKRVLADANLDKVL SAYNKHHRDK PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSTGLYETRIDL SQLGGDGSPKKKRKVSSDYKDHDGDYKDHDIDYKDDDDKAAG (SEQ ID NO: 18)
SpCas9 <i>Streptococcus pyogenes</i> M1GAS wild type NC_002737.2	ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAAATAGCGTCGGATGGGCG GTGATCACTGATGAATATAAGGTTCCGTCTAAAAAGTTCAAGGTTCTGGGAAATACA GACCGCCACAGTATCAAAAAAATCCTTATAGGGGCTCTTTTATTTGACAGTGGAGAG ACAGCGGAAGCGACTCGTCTCAAACGGACAGCTCGTAGAAGGTATACACGTCGGAA GAATCGTATTTGTTATCTACAGGAGATTTTTTCAAATGAGATGGCGAAAGTAGATGAT AGTTTCTTTCATCGACTTGAAGAGTCTTTTTTGGTGGGAAGAAGACAAGAAGCATGA ACGTCATCCTATTTTTGGAAATATAGTAGATGAAGTTGCTTATCATGAGAAATATCCAA CTATCTATCATCTGCGAAAAAATTTGGTAGATTCTACTGATAAAGCGGATTTGCGCTT AATCTATTTGGCCTTAGCGCATATGATTAAGTTTCGTGGTCATTTTTTGATTGAGGGAG ATTTAAATCCTGATAATAGTGATGTGGACAACTATTTATCCAGTTGGTACAAACCTA CAATCAATTATTTGAAGAAAACCTATTAACGCAAGTGGAGTAGATGCTAAAGCGAT TCTTTTGACACGATTGAGTAAATCAAGACGATTAGAAAATCTCATTGCTCAGTCCCC GGTGAGAAGAAAAATGGCTTATTTGGGAATCTCATTGCTTTGCTATGGGTTTGACC CCTAATTTTAAATCAAATTTTGATTTGGCAGAAGATGCTAAATTACAGCTTTCAAAG AATACTTACGATGATGATTTAGATAATTTATTTGGCGCAAATTGGAGATCAATATGCTGAT TTGTTTTTGGCAGCTAAGAATTTATCAGATGCTATTTTACTTTTACAGATATCCTAAGAGT AAATACTGAAATAACTAAGGCTCCCCTATCAGCTTCAATGATTAAACGCTACGATGAA CATCATCAAGACTTGACTCTTTTAAAAGCTTTAGTTTCGACAACAACCTCCAGAAAAG TATAAAGAAATCTTTTTTGATCAATCAAAAAACGGATATGCAGGTTATATTGATGGGG GAGCTAGCCAAGAAGAATTTTATAAATTTATCAAACCAATTTAGAAAAAATGGATG GTACTGAGGAATTATTTGGTGAACTAAATCGTGAAGATTGCTGCGCAAGCAACGGA CCTTTGACAACGGCTCTATTCCCCATCAAATTCACCTGGGTGAGCTGCATGCTATTTT GAGAAGACAAGAAGACTTTTATCCATTTTAAAAGACAATCGTGAGAAGATTGAAA AAATCTTGACTTTTCGAATTCCTTATTATGTTGGTCCATTGGCGCGTGGCAATAGTCG TTTTGCATGGATGACTCGGAAGTCTGAAGAAACAATTACCCCATGGAATTTTGAAGA AGTTGTCGATAAAGGTGCTTCAGCTCAATCATTATTGAACGCATGACAACTTTGAT AAAAATCTTCAAATGAAAAAGTACTACCAAAAACATAGTTTGCTTTATGAGTATTTTA CGGTTTATAACGAATTGACAAAGGTCAAATATGTTACTGAAGGAATGCGAAAACAG CATTTCTTTCAGGTGAACAGAAAGAAAGCCATTGTTGATTTACTCTTCAAACAAATC GAAAAGTAACCGTTAAGCAATTAAAAGAAGATTATTTCAA AAAAATAGAAATGTTTG ATAGTGTTGAAATTCAGGAGTTGAAGATAGATTTAATGCTTCATTAGGTACCTACCA

	TGATTTGCTAAAAATTATTAAAGATAAAGATTTTTGGATAATGAAGAAAATGAAGAT ATCTTAGAGGATATTGTTTTAACATTGACCTTATTTGAAGATAGGGAGATGATTGAGG AAAGACTTAAAACATATGCTCACCTCTTTGATGATAAGGTGATGAAACAGCTTAAAC GTCGCCGTTATACTGGTTGGGGACGTTTGTCTCGAAAATTGATTAATGGTATTAGGGA TAAGCAATCTGGCAAAACAATATTAGATTTTTTGAATCAGATGGTTTTGCCAATCGC AATTTTATGCAGCTGATCCATGATGATAGTTTGACATTTAAAGAAGACATTCAAAAAG CACAAGTGTCTGGACAAGGCGATAGTTTACATGAACATATTGCAAATTTAGCTGGTA GCCCTGCTATTAAAAAAGGTATTTTACAGACTGTAAAAGTTGTTGATGAATTGGTCA AAGTAATGGGGCGGCATAAGCCAGAAAATATCGTTATTGAAATGGCACGTGAAAATC AGACAACTCAAAAGGGCCAGAAAAATTCGCGAGAGCGTATGAAACGAATCGAAGA AGGTATCAAAGAATTAGGAAGTCAGATTCTTAAAGAGCATCCTGTTGAAAATACTCA ATTGCAAAATGAAAAGCTCTATCTCTATTATCTCCAAAATGGAAGAGACATGTATGTG GACCAAGAATTAGATATTAATCGTTTAAAGTGATTATGATGTCGATCACATTGTTCCACA AAGTTTCCTTAAAGACGATTCAATAGACAATAAGGTCTTAACGCGTTCTGATAAAAA TCGTGGTAAATCGGATAACGTTCCAAGTGAAGAAGTAGTCAAAAAGATGAAAAACT ATTGGAGACAACCTCTAAACGCCAAGTTAATCACTCAACGTAAGTTTGATAATTTAA CGAAAGCTGAACGTGGAGGTTTGAGTGAACCTTGATAAAGCTGGTTTTATCAAACGC CAATTGGTTGAAACTCGCCAAATCACTAAGCATGTGGCACAAATTTTGGATAGTCGC ATGAATACTAAATACGATGAAAATGATAAACTTATTCGAGAGGTTAAAGTGATTACCT TAAAATCTAAATTAGTTTCTGACTTCCGAAAAGATTTCCAATTCTATAAAGTACGTGA GATTAACAATTACCATCATGCCCATGATGCGTATCTAAATGCCGTCGTTGGAAGTCTG TTGATTAAGAAATATCCAAAACCTGAATCGGAGTTTGTCTATGGTGATTATAAAGTTT ATGATGTTTCGTAAAATGATTGCTAAGTCTGAGCAAGAAATAGGCAAAGCAACCGCA AAATATTTCTTTTACTCTAATATCATGAACCTCTTCAAAACAGAAATTACACTTGCAA ATGGAGAGATTCGCAAACGCCCTCTAATCGAAACTAATGGGGAAACTGGAGAAATT GTCTGGGATAAAGGGCGAGATTTTGCCACAGTGCAGCAAGTATTGTCCATGCCCCAA GTCAATATTGTCAAGAAAACAGAAGTACAGACAGGCGGATTCTCCAAGGAGTCAAT TTTACCAAAAAGAAATTCGACAAAGCTTATTGCTCGTAAAAAAGACTGGGATCCAA AAAAATATGGTGGTTTTGATAGTCCAACGGTAGCTTATTCAAGTCCTAGTGGTTGCTAA GGTGGAAAAAGGGAAATCGAAGAAGTTAAATCCGTTAAAGAGTTACTAGGGATCA CAATTATGGAAAGAAGTTCCTTTGAaaaaaaATCCGATTGACTTTTTTGAAGCTAAAG GATATAAGGAAGTTAAAAAAGACTTAATCATTAAACTACCTAAATATAGTCTTTTTGA GTTAGAAAACGGTTCGTAAACGGATGCTGGCTAGTGCCGGAGAATTACAAAAAGGAA ATGAGCTGGCTCTGCCAAGCAAATATGTGAATTTTTTATATTTAGCTAGTCATTATGAA AAGTTGAAGGGTAGTCCAGAAGATAACGAACAAAAACAATTGTTTGTGGAGCAGCA TAAGCATTATTTAGATGAGATTATTGAGCAAATCAGTGAATTTTCTAAGCGTGTATTT TAGCAGATGCCAATTTAGATAAAGTTCTTAGTGCATATAACAAACATAGAGACAAAC CAATACGTGAACAAGCAGAAAATATTATTCATTTATTTACGTTGACGAATCTTGGAGC TCCCGCTGCTTTTAAATATTTGATACAACAATTGATCGTAAACGATATACGCTACAA AAGAAGTTTATAGATGCCACTTATCCATCAATCCATCACTGCTTTTATGAAACACG CATTGATTTGAGTCAGCTAGGAGGTGACTGA (SEQ ID NO: 19)
SpCas9 <i>Streptococcus</i> <i>pyogenes</i> M1GAS wild type Encoded product of NC_002737.2 (100% identical to the canonical Q99ZW2 wild type)	MDKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETA EATRLKRTARRRYTRRNKRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIF GNIVDEVAYHEKYPTIYHLRKKLVSDTKADLRILIYLAHMIKFRGHFLIEGDLNPDN SDVDKLFQILVQTYNQLFEENPINASGVDAKILSARLSKSRRLLENLIAQLPGEKKNGLF GNLIASLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLS DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVVLNREDLLRKQRTFDNGSIPHQIHLG ELHAILRRQEDFYFPLKDNREKIEKILTRIPYYVGPLARGNSRFAWMTRKSEETITPWN FEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHSLLYEYFTVYNELTKVKYVTEGMR KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGT HDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQLKRR RYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFMLIHDDSLTFKEDIQKAQV SGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIMARENQTTQ KGQKNSRERMKRIIEGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELD INRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNPSEEVVKMKMKNYWRQLL NAKLITQRKFDNLTKAERGGLSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDEN DKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAVVGTAIIKKYPKLE SEFVYGDYKVYDVRKMIKSEQEI GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET NGETGEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK

	DWDPKKYGGFDSPTVAYSVLVAKVEKGSKKLKSVKELLGITIMERSSSFENPIDFLE AKGYKEVKKDLIIKLPKYSLEFENGRKRMLASAGELQKGNELALPSKYVNFYLLASH YEKLKGSPEDNEQKQLFVEQHKHYLDEHIEQISEFSKRVLADANLDKVL SAYNKHHRDK PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSTGLYETRIDL SQLGGD (SEQ ID NO: 13)
--	---

[0142] The BE-VLPs and fusion proteins described herein may include any of the above SpCas9 sequences, or any variant thereof having at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence identity thereto. In other embodiments, the Cas9 protein can be a wild type Cas9 ortholog from another bacterial species different from the canonical Cas9 from *S. pyogenes*. For example, modified versions of the following Cas9 orthologs can be used in connection with the BE-VLPs and fusion proteins described in this specification by making mutations at positions corresponding to H840A or any other amino acids of interest in wild type SpCas9. In addition, any variant Cas9 orthologs having at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence identity to any of the below orthologs may also be used with the base editors.

Description	Sequence
LfCas9 <i>Lactobacillus fermentum</i> wild type GenBank: SNX31424.1 1	MKEYHIGLDIGTSSIGWAVTDSQFKLMRIKGKTAIGVRLFEEGKTAERRTFRTTRRRLLKR RKWRLHYLDEIFAPHLQEVDENFLRRLKQSNHPEDPTKNQAFIGKLLFPDLLKKNRGY PTLIKMRDELPEQRAHYPMNIYKLREAMINEDRQFDLREYVLAVHHIVKYRGHFLNN ASVDKFKVGRIDFDKSFNVLNEAYEELQNGEGSFTIEPSKVEKIGQLLLDTKMRKLDQR KAVAKLLEVKVADKEETKRNKQIATAMSKLVLGKADFATVAMANGNEWKIDLSSETSE DEIEKFREELESDAQNDILTEITSLFSQIMLNEIVPNGMSISESMMDRYWTHERRQLAEVKEY LATQPASARKEFDQVYNKYIGQAPKERGFDEKGLKKILSKKENWKEIDELLKAGDFLP KQRTSANGVIPHQMHQQELDRIEKQAKYYPWLATENPATGERDRHQAKYELDQLVSFR IPYYVGPLVTPEVQKATSGAKFAWAKRKEDGEITPWNLWDKIDRAESAEAFIKRMTVKD TYLLNEDVLPANSLLYQKYNVLNELNNVRVNGRRLSVGIKQDIYTELFKKKKTVKASDV ASLVMAKTRGVNKPVSVEGLSDPKKFNSNLATYLDLKSIVGDKVDDNRYQTDLENHIEWR SVFEDGEIFADKLTEVEWLTDEQRSALVKKRYKGGWGRSLSKLLTGIVDENGQRIDLMW NTDQNFKEIVDQPVFKEQIDQLNQKAITNDGMTLRERVESVLD DAYTSPQNKKAIWQVV RVVEDIVKAVGNAPKSISIEFARNEGNGKEITRSRRTQLQKLFEDQAHLEVKDTSLTEELE KAPDLSDRYYFYFTQGGKDMYTGDPIFDEISTKYDIDHILPQSFKVNDNSLDNRVLTSRK ENNKSDQVPKLYAAKMKPYWNQLLKQGLITQRKFENLT KDVDQNIKYRSLGFVKRQ LVETRQVIKLTANILGSMYQEAGTEIETRAGLTKQLREEFDLPKRVENDYHHAVDAYL TTFAGQYLNRRYPKLRSFFVYGEYMKFKHGS DLKLRNFFHHELMEGDKSQGKVVDQ TGELITRDEVAKSFDRLNMYMLVSKEVHDRSDQLYGATIVTAKESGKLTSPIEIKNR LVDLYGAYTNGTSAFMTIIFKFTGNKPKYKVIPTTSAASLKRAGKPGSESYNQLHRIK SNPKVKKGFIVVPHVSYGQLIVDGDCKFTLASPTVQHPATQLVLSKKSLETISSGYKILK DKPAIANERLIRVFDEVVGQMNRYFTIFDQRSNRQKVADARDKFLSLPTESKYEGAKKV QVGKTEVITNLLMGLHANATQGD LKVLGLATFGFFQSTTGLSLSEDTMIVYQSPTGLFER RICLKDI (SEQ ID NO: 20)
SaCas9 <i>Staphylococcus aureus</i> wild type GenBank: AYD60528.1	MDKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAE ATRLKRTARRRYTRRNKRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGN IVDEVAYHEKYPTIYHLRKKLVDSTDKADRLRIYLALAHMIKFRGHFLIEGDLNPDNSDV DKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLFGNLI ALSGLTPNFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADFLAAKNLSDAIL LSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAG YIDGGASQEEFYKFIKPILEKMDGTEELLVKNREDLLRKQRTFDNGSIPHQIHLGELHAI LRRQEDFYPLKDNREKIEKILTFRIPIYVGPLARGNSRFAWMTRKSEETITPWNFEVV KGASAQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSG EQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIK

Description	Sequence
	DKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDLSH EHIANLAGSPAIIKKGILQTVKVVDDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER MKRIIEGKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVD HIVPQSFLKDDSIDNKVLTRSDKNRGKSDNPSEEVVKMKMKNYWRQLLNAKLITQRKF DNLTKAERGGSELDAKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT LKSKLVSDFRKDFQFYKVRINNYYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVY DVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKG RDFATVRKVLSPQVNVVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDS PTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIK LPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQ LFVEQHKHYLDEIIIEQISEFSKRVILADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNL GAPAAFKYFDTTIDRKRYTSTKEVLDA TLIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 13)
SaCas9 <i>Staphylococcus aureus</i>	MGKRNILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLR RRHRIQRVKKLLFDYNLLTDHSELGINPYEARVKGLSQKLSSEEFSAALLHLAKRRGVH NVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSYDV EAKQLLKVKQAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGH CTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYEKFQIENVFKQKKKPTL KQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIENAEELLDQIAKILTIYQ SSEDIQEELTNLSELTEQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRL KLVPKVKVDSQKKEIPTTLVDDFILSPVVKRSFIQSIKVINAIKKYGLPNDIIELAREKNSK DAQKMINEMQKRNRQTNERIEEIIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLED LLNNPFNYEVDHIIPRSVSFDNSFNKVLVKQEENSCKGNRTPFQYLSSSDSKISYETFKK HILNLA KGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLRSYFRV NNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAEALIANADFIKEWKKLDKA KKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKPKNRKLI NDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLLKLINKSPEKLLMYHHDPPQTYQKLKL IMEQYGDENPLYKYEEETGNYLTYSKKDNGPVIKKIKYYGNKLN AHLDTDDYPNSR NKVVKLSLKPYPFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKNISNQ AEFIASFYKNDLIKINGELYRVIGVNNDDLNRIEVNMIDITYREYLENMNDKRPPHIKTIA SKTQSIKKYSTDILGNLYEVKSKKHPQIIKK (SEQ ID NO: 21)
StCas9 <i>Streptococcus thermophilus</i> UniProtKB/Swiss-Prot: G3ECR1.2 Wild type	MLFNKCIHISNLDNFSNKEKCMTPYSIGLDIGTNSVGWAVITDNYKVPSSKMKVLGNTS KKYIKKNLLGVLLFDSGITAEGRRRLKRTARRRYTRRRNRILYLQEIFSTEMATLDDAFFQR LDDSLFVPDDKRDSKYPIFGNLVEEKVYHDEFPTIYHLRKYLA DSTKKADRLVYLALA HMIKYRGHFLIEGEFNSKNNDIQKNFQDFLDTYNAIFESDLSLENSKQLEEIVKDKISKLE KKDRILKLPGEKNSGIFSEFLKLIVGNQADFRKCFNLDEKASLHFSKESYDEDLETLLGY IGDDYSDVFLKAKKLYDAILLSGFLTVDNETEAPLSSAMIKRYNEHKEDLALLKEYIRNI SLKTYNEVFKDDTKNGYAGYIDGKTNQEDFYVYLKNLLAEFEGADYFLEKIDREDFLRK QRTFDNGSIPYQIHLQEMRAILDKQAKFYPFLAKNKERIEKILTRIPYYVGPLARGNSDF AWSIRKRNEKITPWNFEDVIDKESSAEAFINRMTSFDLYLPEEKVLPKHSLLYETFNVYNE LTKVRFIAESMRDYQFLDSKQKKDIVRLYFKDKRKVTDKDIIEYLHAIYGYDGIELKGIE KQFNSSLTYHDLLNIINDKEFLDDSSNEAIIIEIHTLTIFEDREMILKQRLSKFENIFDKSVL KKLSRRHYTGWGKLSAKLINGIRDEKSGNTILDYLIIDGINSRNFQMQLIHDDALSFKKKI QKAQIIGDEDKGNIEVVKSLPGSPAIIKKGILQSIKIVDELVKVMGGRRKPESIVVEMAREN QYTNQGKSNSQQLRKRLEKSLKELGSKILKENIPAKLSKIDNNALQNDRLYLYLQNGK DMYTGDLDIDRLSNYDIDHIIPQAFKLDNSIDNKVLVSSASNRGKSDDFPSLEVVKRK TFWYQLLKS KLISQRKFDNLTKAERGGLLPEDKAGFIQRLVETRQITKHVARLLDEKFN NKKDENNRARTVKIITLKTSLVSQFRKDFELYKVRINDFHHAHDAYLNAVIA SALLKK YPKLEPEFVYGDYPKYNSFRERKSATEKVYFYSNIMNIFKKSISLADGRVIERPLIEVNEET GESVWNKESDLATVRRVLSYPQVNVVKKVEEQNHGLDRGPKGLFNANLSSKPKPSN ENLVGAKEYLDPKKYGGYAGISNSFAVLVKG TIEKGAKKKITNVLEFQGISILDRINYRKD KLNFLLEKGYKDIELIHELKYSLEFELSDGSRMLASILSTNNKRGEIHKGNQIFLSQKFVK LLYHAKRISNTINENHRKYVENHKKFEELFYIILEFNENYVGAKKNGKLLNSAFQSWQ NHSIDELCSSFIGPTGSEKGLFELTSRGSAA DFEFLGVKIPRYRDYTPSSLLKDATLIHQS VTGLYETRIDLAKLGEG (SEQ ID NO: 22)
LcCas9 <i>Lactobacillus crispatus</i>	MKIKNYNLALTPTS AVGHVEVDDDLNILEPVHHQKAIGVAKFGEGETAEARRLARSAR RTTKRRANRINH YFNEIMKPEIDKVDPLMFDRIKQAGLSPLDERKEFRTVIFDRPNIASYY HNQFPTIWLHLQKYL MITDEKADIRLIYWALHSLKLRHGHFFNTTPMSQFKPGKLNKDD

Description	Sequence
NCBI Reference Sequence: WP_13347804 4.1 Wild type	MLALDDYNDLEGLSFAVANSPEIEKVIKDRSMHKKKEKIAELKKLIVNDVPDKDLAKRNN KIITQIVNAIMGNSFHLNFIFDMDLTKLAWSKFLDDPELDTKFDASGSMTDNQIGIFE TLQKIYSAISLLDILNGSSNVDAKNALYDKHKRDLNLYFKFLNLTLPDEIAKTLKAGYTL YIGNRKDLLAARKLLKVNVAKNFSQDDFYKLINKELKSIDKQGLQTRFSEKVGELVAQ NNFLPVQRSSDNVFIYPQLNAITFNKILENQGGKYDFLVKPNPAKKDRKNAPYELSOLM QFTIPYYVGPLVTPPEQVKSGIPKTSRFAWMVRKDNGAITPWNFYDKVDIEATADKFIKR SIKDSYLLSELVLPKHSLLEYEYVFNELSNVSLDGGKLSGGVKQILFNEVFKKTNKVN TSRILKALAKHNIPGSKITGLSNPEEFTSSLQTYNAWKKYFPNQIDNFAYQQDLEKMIIEWS TVFEDHKILAKKLDEIEWLDDDDQKKFVANTRLRGWGRLSKRLLTGLKDNYGKSIMQRL ETTKANFQQIVYKPEFREIDKISQAAAKNQSLLEDILANSYTPSPNRKAIRKTMSSVVDEYI KLNHGKEPKIFLMFQRSEQEKGKQTEARSKQLNRILSQLKADKSANKLFSKQLADEFS NAIKKSKYKLNDKQYFYFQQLGRDALTEVIDYDELYKYTVLHIIPRSKLTDDSQNNKV LTKYKIVDGSVALKFGNSYSDALGMPKAFWTELNRLLKIPKGKLLNLTDFSTLNKYQR DGYIARQLVETQQIVKLLATIMQSRFKHTKIEVRNSQVANIRYQFDYFRIKNLNEYRGRF DAYLAADVGTLYKVVYPKARRLFVYGGYLPKPKTNQENQDMHLDSEKKSQGFNFWLN LLYGKQDQIFVNGTDVIAFNRKDLITKMNTVYNYKSQKISLAIDYHNGAMFKATLFPNR DRDTAKTRKLIPKKKDYDTDIYGGYTSNVDGYMLLAELIKRDGNKQYGFYGVPSRLVSE LDTLKKTRYTEYEEKLKEIKPELGVDLKKIKILKNKVPFNQVIIDKGSKFFITSTSYR WNYRQLLSAESQQTLMDLVVDPDFSNHKAARKDARKNADERLIKVYEEILYQVKNYMP MFVELHRCYEKLVDAAQTKFSKLSKISDKAMVLNQILILLHSNATSPVLEKLGYHTRFTLGK KHNLISENAVLTQSITGLKENHVSQML (SEQ ID NO: 23)
PdCas9 <i>Pedicoccus damnosus</i> NCBI Reference Sequence: WP_06291327 3.1 Wild type	MTNEKYSIGLDIGTSSIGFAVNDNNRVRVKGKNAIGVRLFDEGKAAADRRSFRTTRRS FRTRRRRLSRRRWRLKLLREIFDAYITPVDEAFFIRLKESNLSPKDSKKQYSGDILFNDRS DKDFYEKYPTIYHLRNALMTEHRKFDVREIYLAIHHIMKFRGHFLNATPANNFKVGRNL LEEKFEELNDIYQRVFPDESIEFRTDNLEQIKEVLLDNKRSRADRQRTLVSIDIYQSSDKDI EKRNKAVATEILKASLGNKAKLVITNVEVDKEAAKEWSITFDESIDDDLAKIEGQMTD DGHEIIEVLRSLYSGITLSAIVPENHTLSQSMVAKYDLHKDHLKLFKKLINGMTDTKKAK NLRAAYDGYIDGVKGVLPQEDFYKQVQVNLDDSAEANEIQTIDQDIFMPKQRTKAN GSIHQQLQQELDQIENQKAYYPWLAELNPNPDKKRQQLAKYKLDELVTFRVPYVYVGP MITAKDQKNQSGAEFAWMIRKEPKNITPWNFDQKVDRMATANQFIKRMTTTTDTYLLGE DVLPAQSLLYQKFEVLNELNKIRIDHKPISIEQKQIFNDLFKQFKNVTKHLQDYLVSSQ QYSKRPLIEGLADEKRFNSSLSYSTDLGIFGAKLVEENDRQEDLEKIEWSTIFEDKKIYR AKLNDLTWLTDDQKEKLATKRYQGWGRLSRKLVLGLKNSEHRNIMDILWITNENFMQI QAEPDFAKLVTDANKGMLEKTDSDQDVINDLYTSPQNKKAIRQILLVVDIQQNAMHGQAP AKIHVEFARGEERNPRRSVQRQRQVEAAYEKVSNELVSAKVRQEFKEAINNKRDFKDRL FLYFMQGGIDIYTGKQLNIDQLSSYQIDHILPQAFVKDDSLTNRVLTNENQVKADSVPIDI FGKKMLSVWGRMKDQGLISKGYRNLTMPNENISAHTENGFINRQLVETRQVIKLAVNI LADEYGDSTQIISVKADLSHQMRDEFELLKNRDVNDYHHAFDAYLAFIGNYLLKRYPK LESYFVYGDFFKFTQKETKMRRFNFIYDLKHCDQVVNKETGEILWTKDEDIKYIRHLFA YKILVSHEVREKRGALYNQTIYKAKDDKSGSGQESKKLIRIKDDKETKIYGGYSGKSLAY MTIVQITKKNKVSIRVIGIPTLALARLNKLASDSTENNGELYKIIKPKFTHYKVDDKNGEI IETDDFKIVVSKVRFQQLIDDAGQFFMLASDITYKNNAAQQLVISNNALKAINNTNITDCP RDDLERLDNLRDLSAFDEIVKKMDKYFSAYDANNFREKIRNSNLIFYQLPVEDQWENNK ITELGKRTVLTRILQGLHANATTTDMSIFKIKTPFGQLRQRSGISLSENAQLIYQSPTGLFER RVQLNKIK (SEQ ID NO: 24)
FnCas9 <i>Fusobacterium nucleatum</i> NCBI Reference Sequence: WP_06079898 4.1	MKKQKFSDDYLLGFDIGTNSVGWCVTDLDYNNVLRFNKKDMWGSRLFEEAKTAAERRVQ RNSRRRLKRRKWRNLNLEEIFSNEILKIDSNFFRRLKESSLWLEDKSSKEKFTLFNDDNYK DYDFYKQYPTIFHLRNELIKNPEKKDIRLVYLAHSIFKSRGHFLFEGQNLKEIKNFETLYN NLIAFLEDNGINKIIDKNIEKLEKIVCDSSKGLKDKEKEFEIFNSDKQLVAIFKLSVGSS VSLNDLFDTDYKKGVEVEKEKISFREQIYEDDKPIYYISILGEKIELLDIAKTFYDFMVLNN ILADSQYISEAKVKLYEEHKKDLKNLKYIRKYNKGNKYDKLFDKDNENNYSAIYGLNKE KSKKEVIEKSRLKIDDLIKNIKGYLPKVEEIEEKDAIFNKILNKIELKTILPKQRISDNGTL PYQIHEAELEKILENQSKYYDFLNYEENGHITKDKLLMTFKFRIPYVGPLNSYHKDKGG NSWIVRKEEGKILPWNFEQKVDIEKSAEEFIKRMNTNCTYLNGEDVIPKDTFLYSEYVIL NELNKVQVNDEFLENEENKRKIIDELEFKENKKVSEKKFKEYLLVKQIVDGTIELKGVKDSF NSNYISYIRFKDIFGEKLNLDIYKEISEKSILWKCLYGDDKKIFEKKIKNEYGDILTKEIKK INTFKFNWGRLSEKLLTGIEFINLETGECYSSVMDALRRTNYNLMELLSSKFTLQESINN ENKEMNEASYRDLIEESYVSPSLKRAIFQTLKIYEEIRKITGRVPKKVFIEMARGGDESMK NKKIPARQEQLKKLYDSCGNDIANFSIDIKEMKNSLISYDNNSLRQKKLYLYLQFGKCM

Description	Sequence
	YTGREIDLDRLQNNNDTYDIDHIYPRSKVIKDDSFNVLVLKKNENAEKSNEYPVKKEIQ EKMKSFWRFLKEKNFISDEKYKRLTGKDDFELRGFMARQLVNVVRQTTKEVGKILQQIEP EIKIVYSKAEIASSFREMDFIKVRELNDTHHAKDAYLNIVAGNVYNTKFTKPYRYLQEI KENYDVKKIYNIDIKNAWDKENSLEIVKKNMEKNTVNITRFIKEKKGQFLDNLPIKKGE TSNEHSIKPKVYNGKDDKLNEKYGYYSKSLNPAYFLYVEHKEKNKRIKSFERNLVLDVNN IKDEKSLVKYLIENKKLVEPRVIKVKYKRQVILINDYPYSIVTLDNKLMDFENLKPLFLE NKYEKILKNVIKFLDNQKGSEENYKFYILKKKDRYEKNETLESVKDRYNLEFNEMYDK FLEKLDISKDYKNYMNNKKYQELLVDVEKFIKLNLFDKAFTLKSFLDLFNRKTMADFSK VGLTKYLGKIQKISSNVLSKNELYLLEESVTGLFVKIKL (SEQ ID NO: 25)
EcCas9 <i>Enterococcus cecorum</i> NCBI Reference Sequence: WP_04733850 1.1 Wild type	RRKQRIQILQELLGEEVLKTDPGFFHRMKESRYVVEDKRTLDGKQVELPYALFVDKDYT DKEYYKQFPTINHLIVYLMTTSDTPDIRLVYALHYYMKNRGNFLHSGDINNVDKINDIL EQLDNVLETFLDGWNLKLSYVEDIKNIYNRDLGRGERKKAFTNTLGAKTKAEKAFCS LISGGSTNLAELFDDSSLKEIETPKIEFASSLEDKIDGIEALEDRFAVIEAAKRLYDWKTL TDILGDSSSLAEARVNSYQMHHEQLLELKSLSVKEYLDRKVFQEVFVSLNVANNYPAYIG HTKINGKKKELEVKRTKRNDYFYSYVKKQVIEPIKKKVSDEAVTLKSEIESLIEVDKYLPL QVNSDNGVIPYQVKLNELTRIFDNLNENRIPVLRENDRDKIITKFRIPYVYVGSNGVVKNG KCTNWMVRKEEGKIYPWNFEDKVDLEASAEQFIRMTNKTCTYLVNEDVLPKYSLYSK YLVLSLNNLRIDGRPLDVKIKQDIYENVFKNRKVTLLKKIKYLLKEGHTDDDELSGLA DDVKSSLTAYRDFEKLGHLDLSEAQMENIILNITLFGDDKLLKKRLAALYPFIDDKSL NRIATLNYRDWGRLSERFLSGITSVDQETGELRTIIQCMYETQANLMQLLAEPYHFVEAI EKENPKVDLESISYRIVNDLYVSPAVKRQIWQTLVLIKDIKQVMKHDPERIFIEMAREKQE SKKTKSRKQVLSEVYKKAKEYEHLFEKLNLSLTEEQLRSKKIYLYFTQLGKCMYSGEPIDF ENLVSANSNYDIDHIYPQSKTIDDSFNNIVLVKKSLNAYKSNHYPIDKNIRDNEKVKTLW NTLVSKGLITKEKYERLIRSTPFSDEELAGFIARQLVETRQSTKAVAEILSNWFPSEIVYSK AKNVSNFRQDFEILKVRELNDCHHAHDAYLNIVVGNAYHTKFTNSPYRFIKNKANQEYN LRKLLQKVNKIESNGVVAWVGQSENNPGTIATVKKVIRRNTVLISRMVKEVDGQFLDLT LMKKGKGQVPIKSSDERLTDISKYGGYNKATGAYFTFVSKSKRGKVRSFEYVPLHLSK QFENNELLKEYIEKDRGLTDVEILIPKVLINSLFRYNGSLVRITGRGDTRLLLVHEQPLYV SNSFVQQLKSVSSYKLKSENDNAKLTKTATEKLSNIDELYDGLLRKLDLPIYSYWFSSIK EYLVESRTKYIKLSIEEALVIFEILHLFQSDAQVPLKILGLSTKPSRIRIQKNLKDTDKMS IHKQSPSGIFEHEIELTSL (SEQ ID NO: 26)
AhCas9 <i>Anaerostipes hadrus</i> NCBI Reference Sequence: WP_04492427 8.1 Wild type	MQNGFLGITVSSEQVGWAVTNPKYELERASRKDLWGVRLFDKAETAEDRRMFRTNRRRL NQRKKNRIHYLRDIFHEEVNQKDPNFFQQLDESNFCEDDRTVEFNFDNLYKNQFPTVY HLRKYLMETKDKPDIRLVYLAFSKFMKNRGHFLYKGNLGEVMDFENSMKGFCESELEKF NIDFPTLSDEQVKEVRDILCDHKIAKTIVKKKNITITKVKSKTAKAWIGLFCGCSVPVKVL FQDIDEEIVTDPEKISFEDASYDDYIANIEKGVGIYYEAIVSAMKLFWDWSILNEILGDHQLL SDAMIAEYNKHDDDLKRLQKIKGTGSRRELYQDIFINDVSGNYVCYVGHAKTMSSADQK QFYTFILKNRLKNVNGISSEDAEWIDTEIKNGTLLPKQTKRDNVIPHQLQLREFELILDN MQEMYPLKENREKLLKIFNFVIPYVGPLKGVVRKGESTNWMVPKKDGVHPWNFDE MVDKEASAECFISRMGTGNCYSYLFNEKVLPKNSLLYETFEVLNENLPLKINGEPISVELKQ RIYEQLFLTGKKVTKKSLTKYLIKNGYDKDIELSGIDNEFHSNLKSHIDFEDYDNLSDDEEV EQILRITVFEDKQLLDYLNREFVKLSSEDERKQICLSYKGWGNLSEMLLNGITVTDNS GVEVSVMMDMLWNTNLNLMQILSKKYGYKAEIEHYNKEHEKTIYNREDLMDYLNIPPAQ RRKVNQLITIVKSLKKTGYVGNPKIFFKISREHQDDPKRTSSRKEQLKYLYKSLKSEDEKHL MKELDELNDHELSDNKVYLYFLQKGRCIYSGKKLNSRLRKSQNDIDYIYPLSAVND RSMNNKVLGTIQENRADKYTYFPVDSEIQKKMKGFWMELVLQGFMTKEKYFRLSREND FSKSELVSFIEREISDNQQSGRMIAVLQYYPESKIVFVKEKLISSEFKRDFHLISSYGHNL QAAKDAYITIVVGNVYHTKFTMDPAIYFKNHKKRDYDLNRLFLENISRDGQIAWESGPY GSIQTVRKEYAQNHIAVTKRVEVKGGFLKQMPKKGHGEYPLKTNDRFGNIAQYGG YTNVTGSYFVLVESMEKGGKRISLEYVPVYLHERLEDDPGHKLLKEYLVDHRKLNHPKI LLAKVRKNSLLKIDGFYRLNGRSGNALILTNAVELIMDDWQTKTANKISGYMKRRAD KKARVYQNEFHIQELEQLYDFYLDKLNKGVYKNRKNNQAEIHNEKEQFMELKTEDQC VLLTEIKKLFVCSMQADLTIGGSKHTGMIAMSSNVTKADFAVIAEDPLGLRNKVIYSH KGEK (SEQ ID NO: 27)
KvCas9 <i>Kandleria vitulina</i>	MSQNNNKIYNIGLDIGDASVGWAVVDEHYNLLKRHGKHMWGSRLFTQANTAVERRSSR STRRRYNKRERRIRLLREIMEDMVLVDVPTFFIRLANVSFLDQEDKKDYLKENYHSNYN LFIDKDFNDKTYDYKPTIYHLRKHLCESEKEDPRLIYALHHIVKYRGNFLYEGQKFS MDVSNIEDKMIDVLRQFNEINLFYVEDRKKKIDEVLNVLKEPLSKKHKAFAFALFDTT

Description	Sequence
NCBI Reference Sequence: WP_03158996 9.1 Wild type	KDNKAAYKELCAALAGNKFNVTKMLKEAELHDEDEKDISFKFSDATFDDAFVEKQPLL GDCVEFIDLLHDIYSWVELQNILGSAHTSEPSISAAMIQRYEDHKNDLKLKDVIRKYL KKYFEVFRDEKSKNNYCNYINHPSKTPVDEFYKYIKKLEKIDDPDVKTILNKIELESFM LKQNSRTNGAVPYQMQLDELNILENQSVYYSDDLKDNEDKIRSILTFRIPY YFGPLNITKD RQFDWIIKKEGKENERILPWNANEIVDVDKTADDEFIKRMNFCTYFDPDEPVMAKNSLTVS KYEVLNEINKLRINDHLIKRDMKDKMLHTLFMDHKSISANAMKKWLVKNQYFSNTDDI KIEGFQKENACSTSLTPWIDFTKIFGKINESNYDFIEKIIYDVTVFEDKKILRRRLKKEYDL DEEKIKKILKLKYSWWSRLSKKLLSGIKTKYKDSTRTPETVLEVMEERTNMNLMQVINDE KLGFKKTIDDANSTSVSGKFSYAEVQELAGSPAIKRGIWQALLIVDEIKKIMKHEPAHVYI EFARNEDEKERKDSFVNQMLKLYKDYDFEDETEKEANKHLKGEDAKSKIRSERLKLKYT QMGKCMYTGKSLDIDRLDTYQVDHIVPQSLLKDDSIDNKVLVLSSSENQRKLDLVIPSSI RNKMYGFWEKLFNNKIISPKKFYSLIKTEFNEKDQERFINRQIVETRQITKHVAQIIDNH ENTKVVTVRADLSHQFRERYHIYKNRDINDFHHAHDAYIATILGTIYIGHRFESLDAKYI GEYKRIFRNQKNKGKEMKKNNNDGFIENSMRNADKDTGEIVWDPNYIDRIKKCFYK DCFVTKKLEENNGTFFNVTVLPNDTNSDKDNTLATVPVKNYRSNVNKYGGFSGVNSFIV AIKGGKKKGGKVVIEVNKLTIPLMYKNADDEIKINYLKQAEDELEVQIGKEILKNQLIEK DGLLYIVAPTEIINAKQLILNESQTKLVCEIYKAMKYKNYDNLDSEKIIDLYRLLINKME LYYPEYRKQLVKKFEDRYEQLKVISIEEKCNIQKILATLHCNSSIGKIMYSDFKISTTIGRL NGRTISLDDISFIAESPTGMYSKKYKL (SEQ ID NO: 28)
EfCas9 <i>Enterococcus faecalis</i> NCBI Reference Sequence: WP_01663104 4.1 Wild type	MRLFEEGHTAEDRRLKRTARRRISRRRNRLRYLQAFFEEAMTDLDENFFARLQESFLVPE DKKWHHRHPIFAKLEDEVAYHETYPTIYHLRKKLADSSEQADRLRIYLALAHIVKYRGHFL IEGKLSTENTSVDQFQQFMVIYNQTFVNGESRLVSAPLPESVLIEEELTEKASRTKKSEK VLQQFPQEKANGLFGQFLKLMVGNKADFKKVFGLLEEAKITYASESYEEDLEGILAKVG DEYSDVFLAAKNVYDAVELSTILADSDKKSHAKLSSSMIVRFTEHQEDLKKFKRFIRENC PDEYDNLFKNEQKDGAGYIAHAGKVSQKLFYQYVKKIQQDIAGAEYFLEKIAQENFLR KQRTFDNGVIPHQIHLAELQAIHRQAAYYPFLKENQEIEQLVTFRIPY YVGLPSKGDAS TFAWLKRQSEPIRPWNLQETVDLDQSATAFIERMNTNFDTYLPSEKVLPHKSLLYEKFMV FNELTKISYTDGRIKANFSGKEKEKIFDYLFKTRRKVKKKDIIQFYRNEYNTEIVTSLGL EEDQFNASFSTYQDLLKCGLTRAELEDHPDNAEKLEDIKILTIFEDRQRIQTQLSTFKGQFS AEVLKKLERKHYTGWGRLSKKLINGIYDKESGKTILDYLVKDDGVSKHYNRNFMQLIN DSQLSFKNAIQKAQSSSEHEETLSETVNELAGSPAIKKGIYQSLKIVDELVAIMGYAPKRIV VEMARENQTTSTGKRRSIQRLKIVEKAMAEIGSNLLKEQPTTNEQLRDLTRLFLYYMQNG KDMYTGDELSLHRLSHYDIDHIIPQSFMDKDDSLDNLVLVGSTENRGKSDDVPSKEVVKD MKAYWEKLYAAGLISQRKFQRLTKGEQGGLTLEDKAHFIQRQLVETRQITKNVAGILDQR YNAKSKEKKVQIITLKASLTSQFRSIFGLYKRVENDYHHGQDAYLNCVAVTLLKVYPN LAPEFVYGEYPKFQTFKENKATAKAIITNLRRFFTEDEPRFTKDGEILWSNSYLKTIKKE LNYHQMNIVKKVEVQKGGFSKESIKPKGPSNKLIPVKNGLDPQKYGGFDSPPVAYTVLF THEKGKKPLIKQEILGITIMEKTRFEQNPIFLFEKGLRPRVLMKLPKYTLYEFPEGRRL LASAKEAQKGNQMVLPEHLLTLLYHAKQCLLPNQSESLAYVEQHQPFEQILERVVDFA EVHTLAKSKVQQIVKLFEANQADVKEIAASFIQLMQFNAMGAPSTFKFFQKDIERARYT SIKEIFDATHIYQSPTGLYETRRKVVVD (SEQ ID NO: 29)
<i>Staphylococcus aureus</i> Cas9	KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLRRRR HRIQRVKLLFDYNLLTDHSELGINPYEARVGLSQKLSSEEFSAALLHLAKRRGVHNV NEVEEDTGNELSTKEQISRSKALEEKYVAELQLERLKKDGEVRGSINRFKTSYDVKEA KQLLKVKQAYHQLDQSFIDTYIDLLETRRTYEEGPGEGSPFGWKDIKEWYEMLMGHCT YFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYEYEFQIENVFQKQKKPTLKQ IAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIENAEELLDQIAKILTIYQSS EDIQEELTNLSELTEQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKL VPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIKKYGLPNDIIELAREKNSKD AQKMINEMQKRNRTNERIEEIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLL NNPFNYEVDHIIPRSVSFDNSFNKVLVKQEENSCKGNRTPFQYLSSSDSKISYETFKKHI LNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLRSYFRVNN LDVKVKSINGGFTSFLRRKWKFKKERNKGYKHAEDALIINANADFIFKEWKLDKAKK VMENQMFEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKNRELIND TLYSTRKDDKGNLIVNNLNGLYDKDNDKLKLINKSPEKLLMYHHDPPQTYQKLKLIM EQYGDEKNPLYKYEEETGNYLTYSKKDNGPVIKKIKYYGNKLNALHDLITDDYPNSRN KVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKENYYEVNSKCYEEAKKLKKISNQA EFIASFYNNDLIKINGELYRVIGVNNLLNRIEVMIDITYREYLENMNDKRPRIIKTIAS KTQSIKKYSTDILGNLYEVKSKKHPQIIKKG (SEQ ID NO: 30)

Description	Sequence
<i>Geobacillus thermodenitrificans</i> Cas9	MKYKIGLDIGITSIGWAVINLDIPRIEDLGVRIFDRAENPKTGESLALPRRLARSARRRLRR RKHLERIRRLFVREGILTKEELNKLFEKKHEIDVWQLRVEALDRKLNDELARILLHLA KRRGFRSNRKSERTNKENSTMLKHIEENQSSILSSYRTVAEMVVKDPKPSLHKNKEDNY TNTVARDDLEREIKLIFAKQREYGNIVCTEAFEHEYISIWASQRPFASKDDIEKKVGFCTFE PKEKRAPKATYTFQSFTVWEHINKLRLVSPGGIRALTDDERRLIYKQAFHKNKITFHDVR TLLNLPDDTRFKGLLYDRNTTLKENEKVRFLELGAYHKIRKAIDS VYKGGA AKSFRPIDF DTFGYALTMFKDDTDIRS YLRNEYEQNGKR MENLADKVYDEELIEELLNLSFSKFGHLS LKALRNILPYMEQGEVYSTACERAGYTFTGPKKKQKT VLLPNIPPIANPVVMRALTQAR KVVNAIIKKYGSPVSIHIELARELSQSFDERRKMQKEQEGNRKKNETAIRQLVEYGLTLNP TGLDIVKFKLWSEQNGKCAYS LQPIEIERLLEPGYTEVDHVIPYSRSLDDSYTNKVLVLT ENREKGNRTPAEYLG LGSERWQQFETFVLTNKQFSKKRDRLLRLHYDENEENEFKNRN LNDTRYISRFLANFIREHLKFADSDDKQKVYTVNGRITAHLSRWNFNKNREESNLHHA VDAIVACTTPSDIARVTA FYQRREQNKELSKKTD PQFPQWPWFAD ELQARLSKNPKES IKALNLGNYDNEKLESLQPVFVSRMPKRSITGA AHQETLRRYIGIDERSGKIQT VVKKKL SEIQLDKGTGHFPMY GKESDPRTYE AIRQRLL EHNNDPKKAFQEPLYKPKNGELGPIIRTI KIIDTTNQVIPLNDGKT VAYNSNIVRVDFEKGDKY YCVPIYTIDMMKGILPNKAIEPNKP YSEWKEMTEDYTFRFSLYPNDLIRIEFPREKTIKTAVGEEIKIKDLFAYYQTIDSSNGGLSL VSHDNNFSLRSIGSRTLKRFEKYQVDVLGNIYKVRGEKRVGVASSSHSKAGETIRPL (SEQ ID NO: 31)
ScCas9 <i>S. canis</i> 1375 AA 159.2 kDa	MEKKYSIGLDIGTNSVGWAVITDDYKVP SKKFKVLGNTNRKSIKKNLMGALLFDSGETA EATRLKRTARRRYTRRKNRIRYLQEIFANEMAKLDDSFQRLEESFLVEEDKKNERHPIFG NLADDEVAYHRNYPTIYHLRKKLADSPEKADRLIYLALAHIIKFRGHFLIEGKLN AENS VAKLFYQLIQTYNQLFEESPLDEIEVDAKGILSARLSKSKRLEKLI AVFPNEKKNGLFGNII ALALGLTPNFKSNFDLTEDAKLQLSKD TYDDDLDELLGQIGDQYADLFSA AKNLS DAILL SDILRSNSEVTKAPLSASMVKRYDEHHQDLALLKTLVRQQFPEKYAEIFKDDTKNGYAG YVGIGIKHRKRTTKLATQEEFYKFIKPILEKMDGAEELLAKLNRD DLRKQRTFDNGSIPH QIHLKELHAILRRQEEFY PFLKENREKIEKILTFRIPY YVGPLARGNSRFAWLTRKSEEAITP WNFEEVVDKGAS AQSFIERMTNFDEQLPNKKVLPKHSLLYEYFTVYNELTKVKYVTER MRKPEFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEIIGVEDRFNASLGT YHDLKIIKDKDFLDNEENEDILEDIVLT LTLFEDREMIEERLKYAHLFDDKVMKQLKR RHYTGWGRLSRKMINGIRDKQSGKTILDFL KSDGFSNRNFMQLIHDDSLTFKEEIEKAQV SGQGD SLHEQIADLAGSPA IKKGILQTVKIVDELVKVMGHK PENIVIAMARENQTTTKGL QQSRERKKRIE EGIKELESQILKENPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLS DYDVDHIVPQSF IKDDSIDNKVLTRSVENRGKSDNVPSEEVVKMKKNYWRQLLNAKLIT QRKFDNLTKAERGG LSEADKAGFIK RQLVETRQITKHVARILDSRMNTKRDKN DKPIRE VKVITLKS KLVSDFRKDFQLYKVRDINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYG DYKVYDVRKMI AKSEQEIGKATAKRFFYSNIMNFFKTEVK LANGEIRKRPLIETNGETGE VVWNKEKDFATVRKVLAMPQVNIVKKTEVQTGGFSKESILSKRES AKLIPRKKGWDR KYGGFGSPTVAYSILV VAKVEKGKAKKLKSVKVLVGITIMEKGSYEKDPIGFLEAKGYKD IKKELIFKLPKYS LFELENGRRRMLASATELQKANELVLPQHLVRLLYYTQNISATTGSNN LGYIEQHREEFKEIFEKIIDFSEKYILKNKVNSNLKSSFDEQFAVSDSILLSNSFVSLLKYTS FGASGFTFLDL DVKQGR LRYQTVTEVLDATLIYQSITGLYETRTDLSQLGGD (SEQ ID NO: 32)

[0143] The napDNAbp used in the BE-VLPs and fusion proteins described herein may include any suitable homologs and/or orthologs or naturally occurring enzymes, such as, Cas9. Cas9 homologs and/or orthologs have been described in various species, including, but not limited to, *S. pyogenes* and *S. thermophilus*. The Cas moiety may be configured (*e.g.*, mutagenized, recombinantly engineered, or otherwise obtained from nature) as a nickase, *i.e.*, capable of cleaving only a single strand of the target double-stranded DNA. Additional suitable Cas9 nucleases and sequences will be apparent to those of skill in the art based on this disclosure, and such Cas9 nucleases and sequences include Cas9 sequences from the

organisms and loci disclosed in Chylinski, Rhun, and Charpentier, “The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems” (2013) *RNA Biology* 10:5, 726-737; the entire contents of which are incorporated herein by reference. In some embodiments, a Cas9 nuclease has an inactive (*e.g.*, an inactivated) DNA cleavage domain; that is, the Cas9 is a nickase. In some embodiments, the Cas9 protein comprises an amino acid sequence that is at least 85%, at least 90%, at least 92%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to the amino acid sequence of a Cas9 protein as provided by any one of the Cas9 orthologs in the above tables.

[0144] In some embodiments, the VLPs described herein can be used for delivery of any Cas9 equivalent to a target cell. As used herein, the term “Cas9 equivalent” is a broad term that encompasses any nDnAbp protein that serves the same function as Cas9 despite that its amino acid primary sequence and/or its three-dimensional structure may be different and/or unrelated from an evolutionary standpoint. Thus, while Cas9 equivalents include any Cas9 orthologs, homologs, mutants, or variants described or embraced herein that are evolutionarily related, the Cas9 equivalents also embrace proteins that may have evolved through convergent evolution processes to have the same or similar function as Cas9, but which do not necessarily have any similarity with regard to amino acid sequence and/or three-dimensional structure. The VLPs described here may be used to deliver any Cas9 equivalent that would provide the same or similar function as Cas9 despite that the Cas9 equivalent may be based on a protein that arose through convergent evolution. For instance, if Cas9 refers to a type II enzyme of the CRISPR-Cas system, a Cas9 equivalent can refer to a type V or type VI enzyme of the CRISPR-Cas system.

[0145] For example, Cas12e (CasX) is a Cas9 equivalent that reportedly has the same function as Cas9, but which evolved through convergent evolution. Thus, the Cas12e (CasX) protein described in Liu *et al.*, “CasX enzymes comprise a distinct family of RNA-guided genome editors,” *Nature*, 2019, Vol.566: 218-223, is contemplated to be delivered using the VLPs described herein. In addition, any variant or modification of Cas12e (CasX) is conceivable and within the scope of the present disclosure.

[0146] Cas9 is a bacterial enzyme that evolved in a wide variety of species. However, the Cas9 equivalents contemplated herein may also be obtained from archaea, which constitute a domain and kingdom of single-celled prokaryotic microbes different from bacteria.

[0147] In some embodiments, Cas9 equivalents may refer to Cas12e (CasX) or Cas12d (CasY), which have been described in, for example, Burstein *et al.*, “New CRISPR–Cas

systems from uncultivated microbes.” *Cell Res.* 2017 Feb 21. Doi: 10.1038/cr.2017.21, the entire contents of which is hereby incorporated by reference. Using genome-resolved metagenomics, a number of CRISPR–Cas systems were identified, including the first reported Cas9 in the archaeal domain of life. This divergent Cas9 protein was found in little-studied nanoarchaea as part of an active CRISPR–Cas system. In bacteria, two previously unknown systems were discovered: CRISPR–Cas12e and CRISPR–Cas12d, which are among the most compact systems yet discovered. In some embodiments, Cas9 refers to Cas12e, or a variant of Cas12e. In some embodiments, Cas9 refers to a Cas12d, or a variant of Cas12d. It should be appreciated that other RNA-guided DNA binding proteins may be used as a nucleic acid programmable DNA binding protein (napDNAbp) and are within the scope of this disclosure. Also see Liu *et al.*, “CasX enzymes comprises a distinct family of RNA-guided genome editors,” *Nature*, 2019, Vol.566: 218-223. Any of these Cas9 equivalents are contemplated by the present disclosure.

[0148] In some embodiments, the Cas9 equivalent comprises an amino acid sequence that is at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to a naturally-occurring Cas12e (CasX) or Cas12d (CasY) protein. In some embodiments, the napDNAbp is a naturally-occurring Cas12e (CasX) or Cas12d (CasY) protein. In some embodiments, the napDNAbp comprises an amino acid sequence that is at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to a wild-type Cas moiety or any Cas moiety provided herein.

[0149] In various embodiments, the nucleic acid programmable DNA binding proteins include, without limitation, Cas9 (*e.g.*, dCas9 and nCas9), Cas12e (CasX), Cas12d (CasY), Cas12a (Cpf1), Cas12b1 (C2c1), Cas13a (C2c2), Cas12c (C2c3), Argonaute, and Cas12b1. One example of a nucleic acid programmable DNA-binding protein that has different PAM specificity than Cas9 is Clustered Regularly Interspaced Short Palindromic Repeats from *Prevotella* and *Francisella* 1 (*i.e.*, Cas12a (Cpf1)). Similar to Cas9, Cas12a (Cpf1) is also a Class 2 CRISPR effector, but it is a member of type V subgroup of enzymes, rather than the type II subgroup. It has been shown that Cas12a (Cpf1) mediates robust DNA interference with features distinct from Cas9. Cas12a (Cpf1) is a single RNA-guided endonuclease lacking tracrRNA, and it utilizes a T-rich protospacer-adjacent motif (TTN, TTTN, or YTN). Moreover, Cpf1 cleaves DNA via a staggered DNA double-stranded break. Out of 16 Cpf1-

family proteins, two enzymes from *Acidaminococcus* and *Lachnospiraceae* are shown to have efficient genome-editing activity in human cells. Cpf1 proteins are known in the art and have been described previously, for example, in Yamano *et al.*, “Crystal structure of Cpf1 in complex with guide RNA and target DNA.” *Cell* (165) 2016, p. 949-962; the entire contents of which is hereby incorporated by reference.

[0150] In still other embodiments, the Cas protein may include any CRISPR associated protein, including but not limited to, Cas12a, Cas12b1, Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (also known as Csn1 and Csx12), Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, homologs thereof, or modified versions thereof, and preferably comprising a nickase mutation (*e.g.*, a mutation corresponding to the D10A mutation of the wild type Cas9 polypeptide of SEQ ID NO: 13).

[0151] In various other embodiments, the napDNAbp can be any of the following proteins: a Cas9, a Cas12a (Cpf1), a Cas12e (CasX), a Cas12d (CasY), a Cas12b1 (C2c1), a Cas13a (C2c2), a Cas12c (C2c3), a GeoCas9, a CjCas9, a Cas12g, a Cas12h, a Cas12i, a Cas13b, a Cas13c, a Cas13d, a Cas14, a Csn2, an xCas9, an SpCas9-NG, a circularly permuted Cas9, or an Argonaute (Ago) domain, or a variant thereof.

[0152] The VLPs described herein may also be used for delivery of Cas12a (Cpf1) (dCpf1) variants that may be used as a guide nucleotide sequence-programmable DNA-binding protein domain. The Cas12a (Cpf1) protein has a RuvC-like endonuclease domain that is similar to the RuvC domain of Cas9 but does not have an HNH endonuclease domain, and the N-terminus of Cas12a (Cpf1) does not have the alpha-helical recognition lobe of Cas9. It was shown in Zetsche *et al.*, *Cell*, 163, 759–771, 2015 (which is incorporated herein by reference) that the RuvC-like domain of Cas12a (Cpf1) is responsible for cleaving both DNA strands, and inactivation of the RuvC-like domain inactivates Cas12a (Cpf1) nuclease activity.

[0153] In some embodiments, the napDNAbp is a single effector of a microbial CRISPR-Cas system. Single effectors of microbial CRISPR-Cas systems include, without limitation, Cas9, Cas12a (Cpf1), Cas12b1 (C2c1), Cas13a (C2c2), and Cas12c (C2c3). Typically, microbial CRISPR-Cas systems are divided into Class 1 and Class 2 systems. Class 1 systems have multi-subunit effector complexes, while Class 2 systems have a single protein effector. For example, Cas9 and Cas12a (Cpf1) are Class 2 effectors. In addition to Cas9 and Cas12a

(Cpf1), three distinct Class 2 CRISPR-Cas systems (Cas12b1, Cas13a, and Cas12c) have been described by Shmakov *et al.*, “Discovery and Functional Characterization of Diverse Class 2 CRISPR Cas Systems”, *Mol. Cell*, 2015 Nov 5; 60(3): 385–397, the entire contents of which are hereby incorporated by reference.

[0154] Effectors of two of the systems, Cas12b1 and Cas12c, contain RuvC-like endonuclease domains related to Cas12a. A third system, Cas13a, contains an effector with two predicated HEPN Rnase domains. Production of mature CRISPR RNA is tracrRNA-independent, unlike production of CRISPR RNA by Cas12b1. Cas12b1 depends on both CRISPR RNA and tracrRNA for DNA cleavage. Bacterial Cas13a has been shown to possess a unique Rnase activity for CRISPR RNA maturation distinct from its RNA-activated single-stranded RNA degradation activity. These Rnase functions are different from each other and from the CRISPR RNA-processing behavior of Cas12a. See, *e.g.*, East-Seletsky, *et al.*, “Two distinct Rnase activities of CRISPR-Cas13a enable guide-RNA processing and RNA detection”, *Nature*, 2016 Oct 13;538(7624):270-273, the entire contents of which are hereby incorporated by reference. *In vitro* biochemical analysis of Cas13a in *Leptotrichia shahii* has shown that Cas13a is guided by a single CRISPR RNA and can be programmed to cleave ssRNA targets carrying complementary protospacers. Catalytic residues in the two conserved HEPN domains mediate cleavage. Mutations in the catalytic residues generate catalytically inactive RNA-binding proteins. See *e.g.*, Abudayyeh *et al.*, “C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector”, *Science*, 2016 Aug 5; 353(6299), the entire contents of which are hereby incorporated by reference.

[0155] The crystal structure of *Alicyclobacillus acidoterrastris* Cas12b1 (AacC2c1) has been reported in complex with a chimeric single-molecule guide RNA (sgRNA). See *e.g.*, Liu *et al.*, “C2c1-sgRNA Complex Structure Reveals RNA-Guided DNA Cleavage Mechanism”, *Mol. Cell*, 2017 Jan 19;65(2):310-322, the entire contents of which are hereby incorporated by reference. The crystal structure has also been reported in *Alicyclobacillus acidoterrestris* C2c1 bound to target DNAs as ternary complexes. See *e.g.*, Yang *et al.*, “PAM-dependent Target DNA Recognition and Cleavage by C2c1 CRISPR-Cas endonuclease”, *Cell*, 2016 Dec 15;167(7):1814-1828, the entire contents of which are hereby incorporated by reference. Catalytically competent conformations of AacC2c1, both with target and non-target DNA strands, have been captured independently positioned within a single RuvC catalytic pocket, with C2c1-mediated cleavage resulting in a staggered seven-nucleotide break of target DNA.

Structural comparisons between C2c1 ternary complexes and previously identified Cas9 and Cpf1 counterparts demonstrate the diversity of mechanisms used by CRISPR-Cas9 systems. [0156] In some embodiments, the napDNAbp may be a C2c1, a C2c2, or a C2c3 protein. In some embodiments, the napDNAbp is a C2c1 protein. In some embodiments, the napDNAbp is a Cas13a protein. In some embodiments, the napDNAbp is a Cas12c protein. In some embodiments, the napDNAbp comprises an amino acid sequence that is at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to a naturally-occurring Cas12b1 (C2c1), Cas13a (C2c2), or Cas12c (C2c3) protein. In some embodiments, the napDNAbp is a naturally-occurring Cas12b1 (C2c1), Cas13a (C2c2), or Cas12c (C2c3) protein.

Other programmable nucleases

[0157] In various embodiments described herein, the presently disclosed VLPs are used to deliver a napDNAbp, such as a Cas9 protein, alone or as a part of a fusion protein (*e.g.*, a base editor). These proteins are “programmable” by way of their becoming complexed with a guide RNA, which guides the Cas9 protein to a target site on the DNA that possesses a sequence that is complementary to the spacer portion of the gRNA, and that also possesses the required PAM sequence. However, in certain embodiments envisioned here, the napDNAbp may be substituted with a different type of programmable protein, such as a zinc finger nuclease (ZFN) or a transcription activator-like effector nuclease (TALEN), which may be delivered to a target cell using the presently described VLPs.

[0158] As such, it is contemplated that suitable nucleases for delivery using the presently described VLPs do not necessarily need to be “programmed” by a nucleic acid targeting molecule (such as a guide RNA), but rather, may be programmed by defining the specificity of a DNA-binding domain, such as and in particular, a nuclease. Just as with napDNAbp moieties, it may be preferable that such alternative programmable nucleases be modified such that only one strand of a target DNA is cut. In other words, the programmable nucleases may function as nickases.

[0159] Suitable alternative programmable nucleases are well known in the art. TALENS are artificial restriction enzymes generated by fusing the TAL effector DNA binding domain to a DNA cleavage domain. These reagents enable efficient, programmable, and specific DNA cleavage and represent powerful tools for genome editing *in situ*. Transcription activator-like effectors (TALEs) can be quickly engineered to bind practically any DNA sequence. The

term TALEN, as used herein, is broad and includes a monomeric TALEN that can cleave double stranded DNA without assistance from another TALEN. The term TALEN is also used to refer to one or both members of a pair of TALENs that are engineered to work together to cleave DNA at the same site. TALENs that work together may be referred to as a left-TALEN and a right-TALEN, which references the handedness of DNA. See U.S. Ser. No. 12/965,590; U.S. Ser. No. 13/426,991 (U.S. Pat. No. 8,450,471); U.S. Ser. No. 13/427,040 (U.S. Pat. No. 8,440,431); U.S. Ser. No. 13/427,137 (U.S. Pat. No. 8,440,432); and U.S. Ser. No. 13/738,381, all of which are incorporated by reference herein in their entirety. In addition, TALENS are described in WO 2015/027134, U.S. 9,181,535, Boch *et al.*, “Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors”, *Science*, vol. 326, pp. 1509-1512 (2009), Bogdanove *et al.*, TAL Effectors: Customizable Proteins for DNA Targeting, *Science*, vol. 333, pp. 1843-1846 (2011), Cade *et al.*, “Highly efficient generation of heritable zebrafish gene mutations using homo- and heterodimeric TALENs”, *Nucleic Acids Research*, vol. 40, pp. 8001-8010 (2012), and Cermak *et al.*, “Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting”, *Nucleic Acids Research*, vol. 39, No. 17, e82 (2011), each of which are incorporated herein by reference.

[0160] Zinc finger nucleases may also be used as alternative programmable nucleases and delivered using the VLPs described herein. Like with TALENS, the ZFN proteins may be modified such that they function as nickases, *i.e.*, engineering the ZFN such that it cleaves only one strand of the target DNA. ZFN proteins have been extensively described in the art, for example, in Carroll *et al.*, “Genome Engineering with Zinc-Finger Nucleases,” *Genetics*, Aug 2011, Vol. 188: 773-782; Durai *et al.*, “Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells,” *Nucleic Acids Res*, 2005, Vol. 33: 5978-90; and Gaj *et al.*, “ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering,” *Trends Biotechnol.* 2013, Vol.31: 397-405, each of which are incorporated herein by reference in their entireties.

Deaminase domains

[0161] In some embodiments, the BE-VLPs and fusion proteins described herein further comprise a deaminase domain (*e.g.*, when a base editor is being encapsulated and delivered in the VLP). A deaminase domain may be a cytosine deaminase domain or an adenosine deaminase domain.

[0162] Base editors that convert a C to T, in some embodiments, comprise a cytosine deaminase. A “cytosine deaminase” refers to an enzyme that catalyzes the chemical reaction “cytosine + H₂O → uracil + NH₃” or “5-methyl-cytosine + H₂O → thymine + NH₃.” As it may be apparent from the reaction formula, such chemical reactions result in a C to U/T nucleobase change. In the context of a gene, such a nucleotide change, or mutation, may in turn lead to an amino acid change in the protein, which may affect the protein’s function, *e.g.*, loss-of-function or gain-of-function. In some embodiments, the C to T base editor comprises a dCas9 or nCas9 fused to a cytosine deaminase. In some embodiments, the cytosine deaminase domain is fused to the N-terminus of the dCas9 or nCas9.

[0163] Non-limiting examples of suitable cytosine deaminase domains are provided below, as SEQ ID NOs: 33-56.

[0164] Human AID

MDSLLMNRRKFLYQFKNVRWAKGRRETYLCYVVKRRDSATSFSLDFGYLRNKNGC
HVELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGPNLRLRIFTAR
LYFCEDRKAEPGLRRLHRAGVQIAIMTFKDYFYCWNTFVENHERTFKAWEGLHEN
SVRLSRQLRRILLPLYEVDDLRLDAFRTLGL (SEQ ID NO: 33)

[0165] Mouse AID

MDSLLMKQKKFLYHFKNVRWAKGRHETLYCYVVKRRDSATSCSLDFGHLRNKSGC
HVELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVAEFLRWPNLRLRIFTAR
LYFCEDRKAEPGLRRLHRAGVQIGIMTFKDYFYCWNTFVENRERTFKAWEGLHEN
SVRLTRQLRRILLPLYEVDDLRLDAFRMLGF (SEQ ID NO: 34)

[0166] Dog AID

MDSLLMKQRKFLYHFKNVRWAKGRHETLYCYVVKRRDSATSFSLDFGHLRNKSGC
HVELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGYPNLRLRIFAAR
LYFCEDRKAEPGLRRLHRAGVQIAIMTFKDYFYCWNTFVENREKTFKAWEGLHEN
SVRLSRQLRRILLPLYEVDDLRLDAFRTLGL (SEQ ID NO: 35)

[0167] Bovine AID

MDSLLKKQRQFLYQFKNVRWAKGRHETLYCYVVKRRDSPTSFSLSDFGHLRNKAGC
HVELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGYPNLRLRIFTAR
LYFCDKERKAEPGLRRLHRAGVQIAIMTFKDYFYCWNTFVENHERTFKAWEGLHE
NSVRLSRQLRRILLPLYEVDDLRLDAFRTLGL (SEQ ID NO: 36)

[0168] Mouse APOBEC-3

MGPFCCLGCSHRKCYSPIRNLISQETFKFHFKNLGYAKGRKDTFLCYEVTRKDCDSPV
SLHHGVFKNKDNIHAEICFLYWFHDKVLKVLSPREEFKITWYMSWSPCFECAEQIVR
FLATHHNLSLDIFSSRLYNVQDPETQQNLRLVQEGAQVAAMDLYEFKKCWKKFVD
NGGRRFRPWKRLLTNFRYQDSKLQEILRPCYIPVSSSSSTLSNICLTGKLPETRFCVE
GRRMDPLSEEFYSQFYNQVRVKHLCYYHRMKPYLCYQLEQFNGQAPLKGCLLSEKG
KQHAEIFLDKIRSMELSQVTITCYLTWSPCPNCAWQLAAFKRDRPDLILHIYTSRLY

FHWKRPFQKGLCSLWQSGILVDVMDLPQFTDCWTN FVNPKRPFWPWKGLEIISRRT
QRRLRRIKESWGLQDLVND FGNLQLGPPMS (SEQ ID NO: 37)

[0169] Rat APOBEC-3

MGPFC LGCSHRKCYSPIRNLISQETFKFHFKNLRYAIDRKDTFLCYEVTRKDCDSPVS
LHHGV FKNKDNIHAEICFLYWFHDKVLKVLSPREEFKITWYMSWSPCFECAEQVLR
LATHHNLSLDIFSSRLYNIRDPENQQNL CRLVQEGAQVAAMDLYEFKKCWKKFVDN
GGRRFRPWKKLLTNFRYQDSKLQEILRPCYIPVPSSSSSTLSNICLTKGLPETRFCVER
RRVHLLSEEEFY SQFYNQVRVKHLCYYHGVKPYLCYQLEQFNGQAPLKGCLLSEK GK
QHAEILFLDKIRSMELSQVIITCYLTWSPCPNCAWQLAAFKRDRPD LILHIYTSRLYFH
WKRPFQKGLCSLWQSGILVDVMDLPQFTDCWTN FVNPKRPFWPWKGLEIISRRTQR
RLHRIKESWGLQDLVND FGNLQLGPPMS (SEQ ID NO: 38)

[0170] Rhesus macaque APOBEC-3G

MVEPMDPRTFVS NFNNRPILSGLNTVWLCCEVKT KDPSGPPLDAKIFQGKVYSKAKY
HPEMRFLRWFHKWRQLHHDQEYKVTWYVSWSPCTRCANSVATFLAKDPKVTLTIF
VARLYYFWKPDYQQALRILCQKRG GPHATMKIMNYNEFQDCWNKFVDGRGKPFKP
RNNLPKH YTL LQATLGELLRHLM DPGTFTSNFNNKPWVSGQHET YLCYKVERLHND
TWVPLNQHRGFLRNQAPNIHGFPKGRHAELCFLDLIPFWKLDGQQYRVTCFTSWSPC
FSCAQEMAKFISNNEHVSLCIFAARIYDDQGRYQEGLRALHRDGAKIAMMNYSEFEY
CWDTFVDRQGRPFQPWDGLDEHSQALSGRLRAI (SEQ ID NO: 39)

[0171] Chimpanzee APOBEC-3G

MKPHFRNPVERMYQDTFSDNFYNRPILSHRNTVWLCYEVKT KGPSRPPLDAKIFRGQ
VYSKLKYHPEMRFFHWFSKWRKLHRDQEYEV TWYISWSPCTKCTRDVATFLAEDP
KVTLTIFVARLYYFWDPDYQEALRSLCQKRDGPRATMKIMNYDEFQHCWSKFVYS
QRELFEPWNNLPKYYILLHIMLGEILRHSM DPTFTSNFNNELWVRGRHET YLCYEV
ERLHN D TWVLLNQRRGFLCNQAPHKHGFLEGRHAELCFLDVIPFWKLDLHQDYRV
T CFTSWSPCFSCAQEMAKFISNNKHVSLCIFAARIYDDQGR CQEGLRTLAKAGAKISIM
TYSEFKHCWDTFVDHQGCPFPQWDGLEEHSQALSGRLRAILQNQGN (SEQ ID NO:
40)

[0172] Green monkey APOBEC-3G

MNPQIRNMVEQMEPDIFVYYFNRPILSGRNTVWLCYEVKT KDPSGPPLDANIFQGK
LYPEAKDHPEMKFLHWFRKWRQLHRDQEYEV TWYVSWSPCTRCANSVATFLAEDP
KVTLTIFVARLYYFWKPDYQQALRILCQERGGPHATMKIMNYNEFQHCWNEFVDG
QGKPKPRKNLPKH YTL L HATLGELLRHVMDPGTFTSNFNNKPWVSGQRETYLCYK
VERSHND TWVLLNQHRGFLRNQAPDRHGFPKGRHAELCFLDLIPFWKLDDQYRV
T CFTSWSPCFSCAQKMAKFISNNKHVSLCIFAARIYDDQGR CQEGLRTLHRDGAKIAV
MNYSEFEYCWDTFVDRQGRPFQPWDGLDEHSQALSGRLRAI (SEQ ID NO: 41)

[0173] Human APOBEC-3G

MKPHFRNTVERMYRDTFSYNFYNRPILSRNTVWLCYEVKT KGPSRPPLDAKIFRGQ
VYSELKYHPEMRFFHWFSKWRKLHRDQEYEV TWYISWSPCTKCTRDMATFLAEDP
KVTLTIFVARLYYFWDPDYQEALRSLCQKRDGPRATMKIMNYDEFQHCWSKFVYS
QRELFEPWNNLPKYYILLHIMLGEILRHSM DPTFTFNFNNEPWVRGRHET YLCYEV
ERMHN D TWVLLNQRRGFLCNQAPHKHGFLEGRHAELCFLDVIPFWKLDLDQDYRV

TCFTSWSPCFSCAQEMAKFISK NKHVSLCIFTARIYDDQGRQCQEGRLTLAEAGAKISI
MTYSEFKHCWDTFVDHQGCPFQPWDGLDEHSQDLSGRLRAILQNQEN (SEQ ID NO:
42)

[0174] Human APOBEC-3F

MKPHFRNTVERMYRDTFSYNFYNRPILSRRNTVWLCYEYVKTGKPSRPRLDAKIFRGQ
VYSQPEHHAEMCFLSWFCGNQLPAYKCFQITWVSWTPCPDCVAKLAEFLAHPNV
TLTISAARLYYYWERDYRRALCRLSQAGARVKIMDDEEFAYCWENFVYSEGQPFMP
WYKFDDNYAFLHRTLKEILRNPMEAMYPHIFYFHFKNLRKAYGRNESWLCFTMEV
VKHHSPVSWKRGVFRNQVDPETHCHAERCFLSWFCDDILSPNTNYEVTWYTSWSPC
PECAGEVAEFLARHSNVNLTIFTARLYYFWDTDYQEGRLSLSQEGASVEIMGYKDFK
YCWENFVYNDDEPFKPWKGLKYNFLFLDSKLQEILE (SEQ ID NO: 43)

[0175] Human APOBEC-3B

MNPQIRNPMERMYRDTFYDNFENEPILYGRSYTWLCYEYVKIKRGRSNLLWDTGVFR
GQVYFKPQYHAEMCFLSWFCGNQLPAYKCFQITWVSWTPCPDCVAKLAEFLSEHP
NVTLTISAARLYYYWERDYRRALCRLSQAGARVTIMDYEEFAYCWENFVYNEGQQ
FMPWYKF DENYAFLHRTLKEILRYLMDPDTFTFNFNNDPLVLRRTQTYLCYEYERL
DNGTWVLMDQHMGFLCNEAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYRVTWFI
SWSPCFSWGCAGEVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSI
MTYDEFEYCWDTFVYRQGCPFQPWDGLEEHSQALSGLRLRAILQNQGN (SEQ ID NO:
44)

[0176] Human APOBEC-3C

MNPQIRNPMKAMYPGTFYFQFKNLWEANDRNETWLCFTVEGIKRRSVVSWKTGVF
RNQVDSETHCHAERCFLSWFCDDILSPNTKYQVTWYTSWSPCPDCAGEVAEFLARH
SNVNLTIFTARLYYFQYPCYQEGRLSLSQEGVAVEIMDYEDFKYCWENFVYNDNEPF
KPWKGLKTNFRLKRRRLRESLQ (SEQ ID NO: 45)

[0177] Human APOBEC-3A

MEASPASGPRHLMDPHIFTSNFFNNGIGRHKTYLCYEYERLDNGTSVKMDQHRGFLH
NQAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYRVTWFI SWSPCFSWGCAGEVRAFL
QENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQ
GCPFQPWDGLDEHSQALSGLRLRAILQNQGN (SEQ ID NO: 46)

[0178] Human APOBEC-3H

MALLTAETFRLQFNKRRLRRPYYPKALLCYQLTPQNGSTPTRGYFENKKKCHAEI
CFINEIKSMGLDETQCYQVTCYLTWSPCSSCAWELVDFIKAHDHLNLGIFASRLYYH
WCKPQQKGLRLLCGSQVPVEVMGF PKFADCWENFVDHEKPLSFNPYKMLEELDKN
SRAIKRRLERIKIPGVRAQGRYMDILCDAEV (SEQ ID NO: 47)

[0179] Human APOBEC-3D

MNPQIRNPMERMYRDTFYDNFENEPILYGRSYTWLCYEYVKIKRGRSNLLWDTGVFR
GPVLPKRQSNHRQEYVYFRFENHAEMCFLSWFCGNRLPANRRFQITWVSWNPCLPC
VVKVTKFLAHPNVTLTISAARLYYYRDRDWRVLLRLHKAGARVKIMDYEDFAY
CWENFVCNEGQPFMPWYKFDDNYASLHRTLKEILRNPMEAMYPHIFYFHFKNLLKA
CGRNESWLCFTMEVTKHSAVFRKRGVFRNQVDPETHCHAERCFLSWFCDDILSPN

TNYEVTWYTSWSPCECAGEVAEFLARHSNVNLTIFTARLCYFWDTDYQEGLCSL
QEGASVKIMGYKDFVSCWKNFVYSDDEPFKPWKGLQTNFRLKRRRLREILQ (SEQ
ID NO: 48)

[0180] Human APOBEC-1

MTSEKGPSTGDPTLRRRIEPWEFDVFYDPRELRLKEACLLYEIKWGMSRKIWRSSGKN
TTNHVEVNFIIKFTSERDFHPSMSCSITWFLSWSPCWECQAIREFLSRHPGVTLVIYV
ARLFWHMDQQNRQGLRDLVNSGVTIQIMRASEYYHCWRNFVNYPGDEAHWPQY
PPLWMMLYALELHCILSLPPCLKISRRWQNHLTFFRLHLQNCHYQTIPPHILLATGLI
HPSVAWR (SEQ ID NO: 49)

[0181] Mouse APOBEC-1

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRLKETCLLYEINWGGRHSVWRHTSQN
TSNHVEVNFLEKFTTERYFRPNTRCSITWFLSWSPCGECSRAITEFLSRHPYVTLFIYIA
RLYHHTDQRNRQGLRDLISSGVTIQIMTEQEYCYCWRNFVNYPSPNEAYWPRYPHL
WVKLYVLELYCIILGLPPCLKILRRKQPQLTFFTITLQTCHYQRIPPHLLWATGLK
(SEQ ID NO: 50)

[0182] Rat APOBEC-1

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRLKETCLLYEINWGGRHSIWRHTSQNT
NKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIAR
LYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYPSPNEAHWPYPHLW
VRLYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLK (SEQ
ID NO: 51)

[0183] Petromyzon marinus CDA1 (pmCDA1)

MTDAEYVRIHEKLDIYTFKKQFFNNKKS VSHRCYVLFELKRRGERRACFWGYAVNK
PQSGTERGIHAEIFSIRKVEEYLRDNPGQFTINWYSSWSPCADCAEKILEWYNQELRG
NGHTLKIWACKLYYEKNARNQIGLWNLRDNGVGLNVMVSEHYQCCRKIFIQSSHN
QLNENRWLEKTLKRAEKRRSELSIMI QVKILHTTKSPAV (SEQ ID NO: 52)

[0184] Evolved pmCDA1 (evoCDA1)

MTDAEYVRIHEKLDIYTFKKQFSNNKKS VSHRCYVLFELKRRGERRACFWGYAVNK
PQSGTERGIHAEIFSIRKVEEYLRDNPGQFTINWYSSWSPCADCAEKILEWYNQELRG
NGHTLKIWVCKLYYEKNARNQIGLWNLRDNGVGLNVMVSEHYQCCRKIFIQSSHN
QLNENRWLEKTLKRAEKRRSELSIMFQVKILHTTKSPAV (SEQ ID NO: 53)

[0185] Human APOBEC3G D316R_D317R

MKPHFRNTVERMYRDTFSYNFYNRPILSRRNTVWLCYEVKTKGPSRPPLDAKIFRGQ
VYSELKYHPEMRFFHWFSKWRKLHRDQEYEV TWYISWSPCTKCTRD MATFLAEDP
KVTLTIFVARLYYFWD PDYQEALRSLCQKRDGPRATMKIMNYDEFQHCWSKFVYS
QRELFE PWNNLPKYYILLHIMLGEILRHSM DPPTFTFNFNNEPWVRGRHET YLCYEV
ERMHNDTWVLLNQRRGFLCNQAPHKHGFLEGRHAELCFLDVIPFWKLDLDQDYRV
TCFTSWSPCFSCAQEMAKFISK NKHVS LCIFTARIYRRQGR CQEGLRTLAEAGAKISI
MTYSEFKHCWDTFVDHQGCPFPWDGLDEHSQDLSGRLRAILQNQEN (SEQ ID NO:
54)

[0186] Human APOBEC3G chain A

MDPPTFTFNFNNEPWVRGRHETLYLCYEVERMHNDTWVLLNQRRGFLCNQAPHKHG
 FLEGRHAELCFLDVIPFWKLDLDQDYRVTCFTSWSPCFSCAQEMAKFISKKNKHVSLCI
 FTARIYDDQGRCQEGLRTLAEAGAKISIMTYSEFKHCWDTFVDHQGCPFQPWDGLD
 EHSQDLSGRLRAILQ (SEQ ID NO: 55)

[0187] Human APOBEC3G chain A D120R_D121R

MDPPTFTFNFNNEPWVRGRHETLYLCYEVERMHNDTWVLLNQRRGFLCNQAPHKHG
 FLEGRHAELCFLDVIPFWKLDLDQDYRVTCFTSWSPCFSCAQEMAKFISKKNKHVSLCI
 FTARIYRRQGRCQEGLRTLAEAGAKISIMTYSEFKHCWDTFVDHQGCPFQPWDGLD
 EHSQDLSGRLRAILQ (SEQ ID NO: 56)

[0188] In some embodiments, a base editor converts an A to G. In some embodiments, the base editor comprises an adenosine deaminase. An “adenosine deaminase” is an enzyme involved in purine metabolism. It is needed for the breakdown of adenosine from food and for the turnover of nucleic acids in tissues. Its primary function in humans is the development and maintenance of the immune system. An adenosine deaminase catalyzes hydrolytic deamination of adenosine (forming inosine, which base pairs as G) in the context of DNA. There are no known adenosine deaminases that act on DNA. Instead, known adenosine deaminase enzymes only act on RNA (tRNA or mRNA). Evolved deoxyadenosine deaminase enzymes that accept DNA substrates and deaminate dA to deoxyinosine and here use in adenosine nucleobase editors have been described, *e.g.*, in PCT Application PCT/US2017/045381, filed August 3, 2017, which published as WO 2018/027078, PCT Application No. PCT/US2019/033848, which published as WO 2019/226953, PCT Application No. PCT/US2019/033848, filed May 23, 2019, and PCT Application No. PCT/US2020/028568, filed April 17, 2020; each of which is herein incorporated by reference. Non-limiting examples of evolved adenosine deaminases that accept DNA as substrates are provided below. In some embodiments, an adenosine deaminase comprises any of the following amino acid sequences, or an amino acid sequence at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or at least 99.9% identical to any of the following amino acid sequences:

[0189] ecTadA

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
 HAEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTG
 AAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD
 (SEQ ID NO: 57)

[0190] ecTadA (D108N)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
 HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKTG
 AAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD
 (SEQ ID NO: 58)

[0191] ecTadA (D108G)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
 HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARGAKTG
 AAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD
 (SEQ ID NO: 59)

[0192] ecTadA (D108V)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
 HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARVAKTG
 AAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD
 (SEQ ID NO: 60)

[0193] ecTadA (H8Y, D108N, N127S)

SEVEFSYEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
 HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKTG
 AAGSLMDVLHHPGMSHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD
 (SEQ ID NO: 61)

[0194] ecTadA (H8Y, D108N, N127S, E155D)

SEVEFSYEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
 HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKTG
 AAGSLMDVLHHPGMSHRVEITEGILADECAALLSDFFRMRRQDIKAQKKAQSSTD
 (SEQ ID NO: 62)

[0195] ecTadA (H8Y, D108N, N127S, E155G)

SEVEFSYEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
 HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKTG
 AAGSLMDVLHHPGMSHRVEITEGILADECAALLSDFFRMRRQGIKAQKKAQSSTD
 (SEQ ID NO: 63)

[0196] ecTadA (H8Y, D108N, N127S, E155V)

SEVEFSYEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
 HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKTG
 AAGSLMDVLHHPGMSHRVEITEGILADECAALLSDFFRMRRQVIKAQKKAQSSTD
 (SEQ ID NO: 64)

[0197] ecTadA (A106V, D108N, D147Y, and E155V)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
 HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGVRNAKTG
 AAGSLMDVLHHPGMNHRVEITEGILADECAALLSYFFRMRRQVIKAQKKAQSSTD
 (SEQ ID NO: 65)

[0198] ecTadA (S2A, I49F, A106V, D108N, D147Y, E155V)

AEVEFSSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPFGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGVRNAKTG
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSYFFRMRRQVIKAQKKAQSSTD
 (SEQ ID NO: 66)

[0199] ecTadA (H8Y, A106T, D108N, N127S, K160S)

SEVEFSYEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
 HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGTRNAKTG
 AAGSLMDVLHHPGMSHRVEITEGILADECAALLSDFFRMRRQEIKAQSKAQSSTD
 (SEQ ID NO: 67)

[0200] ecTadA (R26G, L84F, A106V, R107H, D108N, H123Y, A142N, A143D, D147Y, E155V, I156F)

SEVEFSHEYWMRHALTLAKRAWDEGEVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
 HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVHNAKTG
 AAGSLMDVLHYPGMNHRVEITEGILADECNDLLSYFFRMRRQVFKAQKKAQSSTD
 (SEQ ID NO: 68)

[0201] ecTadA (E25G, R26G, L84F, A106V, R107H, D108N, H123Y, A142N, A143D, D147Y, E155V, I156F)

SEVEFSHEYWMRHALTLAKRAWDGGEVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
 HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVHNAKTG
 AAGSLMDVLHYPGMNHRVEITEGILADECNDLLSYFFRMRRQVFKAQKKAQSSTD
 (SEQ ID NO: 69)

[0202] ecTadA (E25D, R26G, L84F, A106V, R107K, D108N, H123Y, A142N, A143G, D147Y, E155V, I156F)

SEVEFSHEYWMRHALTLAKRAWDDGEVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
 HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVKNAKTG
 AAGSLMDVLHYPGMNHRVEITEGILADECNGLLSYFFRMRRQVFKAQKKAQSSTD
 (SEQ ID NO: 70)

[0203] ecTadA (R26Q, L84F, A106V, D108N, H123Y, A142N, D147Y, E155V, I156F)

SEVEFSHEYWMRHALTLAKRAWDEQEVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
 HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
 AAGSLMDVLHYPGMNHRVEITEGILADECNALLSYFFRMRRQVFKAQKKAQSSTD
 (SEQ ID NO: 71)

[0204] ecTadA (E25M, R26G, L84F, A106V, R107P, D108N, H123Y, A142N, A143D, D147Y, E155V, I156F)

SEVEFSHEYWMRHALTLAKRAWDMGEVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
 HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVPNAKTG

AAGSLMDVLHYPGMNHRVEITEGILADECNDLLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 72)

[0205] ecTadA (R26C, L84F, A106V, R107H, D108N, H123Y, A142N, D147Y, E155V, I156F)

SEVEFSHEYWMRHALTLAKRAWDECEVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVHNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECNALLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 73)

[0206] ecTadA (L84F, A106V, D108N, H123Y, A142N, A143L, D147Y, E155V, I156F)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECNLLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 74)

[0207] ecTadA (R26G, L84F, A106V, D108N, H123Y, A142N, D147Y, E155V, I156F)

SEVEFSHEYWMRHALTLAKRAWDEGEVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECNALLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 75)

[0208] ecTadA (R51H, L84F, A106V, D108N, H123Y, D147Y, E155V, I156F, K157N)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGHHHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFNAQKKAQSSTD
(SEQ ID NO: 76)

[0209] ecTadA (E25A, R26G, L84F, A106V, R107N, D108N, H123Y, A142N, A143E, D147Y, E155V, I156F)

SEVEFSHEYWMRHALTLAKRAWDAGEVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVNNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECNELLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 77)

[0210] ecTadA (N37T, P48T, L84F, A106V, D108N, H123Y, D147Y, E155V, I156F)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHTNRVIGEGWNRTIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 78)

[0211] ecTadA (N37S, L84F, A106V, D108N, H123Y, D147Y, E155V, I156F)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHSNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG

AAGSLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 79)

[0212] ecTadA (H36L, L84F, A106V, D108N, H123Y, D147Y, E155V, I156F)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVLNRRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 80)

[0213] ecTadA (H36L, P48L, L84F, A106V, D108N, H123Y, D147Y, E155V, I156F)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVLNRRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 81)

[0214] ecTadA (H36L, L84F, A106V, D108N, H123Y, D147Y, E155V, K57N, I156F)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVLNRRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFNAQKKAQSSTD
(SEQ ID NO: 82)

[0215] ecTadA (H36L, L84F, A106V, D108N, H123Y, S146C, D147Y, E155V, I156F)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVLNRRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 83)

[0216] ecTadA (L84F, A106V, D108N, H123Y, S146R, D147Y, E155V, I156F)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLRYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 84)

[0217] ecTadA (N37S, R51H, L84F, A106V, D108N, H123Y, D147Y, E155V, I156F)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGHHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 85)

[0218] ecTadA (R51L, L84F, A106V, D108N, H123Y, D147Y, E155V, I156F, K157N)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGLHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFNAQKKAQSSTD
(SEQ ID NO: 86)

[0219] saTadA (D108N)

GSHMTNDIYFMTLAIEEAKKAAQLGEVPIGAIITKDDEVIARAHNLRETLQQPTAHAE
HIAIERAAKVLGSRLEGCTLYVTLEPCVMCAGTIVMSRIPRVVYGADNPKGGCSGS
LMNLLQQSNFNHRAIVDKGVLKEACSTLLTFFKNLRANKKSTN (SEQ ID NO: 87)

[0220] saTadA (D107A_D108N)

GSHMTNDIYFMTLAIEEAKKAAQLGEVPIGAIITKDDEVIARAHNLRETLQQPTAHAE
HIAIERAAKVLGSRLEGCTLYVTLEPCVMCAGTIVMSRIPRVVYGAANPKGCSGS
LMNLLQQSNFNHRAIVDKGVLKEACSTLLTFFKNLRANKKSTN (SEQ ID NO: 88)

[0221] saTadA (G26P_D107A_D108N)

GSHMTNDIYFMTLAIEEAKKAAQLPEVPIGAIITKDDEVIARAHNLRETLQQPTAHAE
HIAIERAAKVLGSRLEGCTLYVTLEPCVMCAGTIVMSRIPRVVYGAANPKGCSGS
LMNLLQQSNFNHRAIVDKGVLKEACSTLLTFFKNLRANKKSTN (SEQ ID NO: 89)

[0222] saTadA (G26P_D107A_D108N_S142A)

GSHMTNDIYFMTLAIEEAKKAAQLPEVPIGAIITKDDEVIARAHNLRETLQQPTAHAE
HIAIERAAKVLGSRLEGCTLYVTLEPCVMCAGTIVMSRIPRVVYGAANPKGCSGS
LMNLLQQSNFNHRAIVDKGVLKEACATLLTFFKNLRANKKSTN (SEQ ID NO: 90)

[0223] saTadA (D107A_D108N_S142A)

GSHMTNDIYFMTLAIEEAKKAAQLGEVPIGAIITKDDEVIARAHNLRETLQQPTAHAE
HIAIERAAKVLGSRLEGCTLYVTLEPCVMCAGTIVMSRIPRVVYGAANPKGCSGS
LMNLLQQSNFNHRAIVDKGVLKEACATLLTFFKNLRANKKSTN (SEQ ID NO: 91)

[0224] ecTadA (P48S)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRSIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTG
AAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD
(SEQ ID NO: 92)

[0225] ecTadA (P48T)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRTIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTG
AAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD
(SEQ ID NO: 93)

[0226] ecTadA (P48A)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRAIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTG
AAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD
(SEQ ID NO: 94)

[0227] ecTadA (A142N)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTG

AAGSLMDVLHHPGMNHRVEITEGILADECANALLSDFFRMRRQEIKAQKKAQSSTD
(SEQ ID NO: 95)

[0228] ecTadA (W23R)

SEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTG
AAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD
(SEQ ID NO: 96)

[0229] ecTadA (W23L)

SEVEFSHEYWMRHALTLAKRALDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTG
AAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD
(SEQ ID NO: 97)

[0230] ecTadA (R152P)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTG
AAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMPRQEIKAQKKAQSSTD
(SEQ ID NO: 98)

[0231] ecTadA (R152H)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTG
AAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMHRQEIKAQKKAQSSTD
(SEQ ID NO: 99)

[0232] ecTadA (L84F, A106V, D108N, H123Y, D147Y, E155V, I156F)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPVCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 100)

[0233] ecTadA (H36L, R51L, L84F, A106V, D108N, H123Y, S146C, D147Y, E155V,
I156F, K157N)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVLNNRVIGEGWNRPIGLHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPVCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMRRQVFNAQKKAQSSTD
(SEQ ID NO: 101)

[0234] ecTadA (H36L, P48S, R51L, L84F, A106V, D108N, H123Y, S146C, D147Y,
E155V, I156F, K157N)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVLNNRVIGEGWNRSIGLHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPVCVMCAGAMIHSRIGRVVFGVRNAKTG

AAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMRRQVFNAQKKAQSSTD
(SEQ ID NO: 102)

[0235] ecTadA (H36L, P48A, R51L, L84F, A106V, D108N, H123Y, S146C, D147Y,
E155V, I156F, K157N)

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMRRQVFNAQKKAQSSTD
(SEQ ID NO: 103)

[0236] ecTadA (W23L, H36L, P48A, R51L, L84F, A106V, D108N, H123Y, S146C,
D147Y, R152P, E155V, I156F, K157N)

SEVEFSHEYWMRHALTLAKRALDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD
(SEQ ID NO: 104)

[0237] ecTadA (W23R, H36L, P48A, R51L, L84F, A106V, D108N, H123Y, S146C,
D147Y, R152P, E155V, I156F, K157N)

SEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD
(SEQ ID NO: 113)

[0238] *Staphylococcus aureus* TadA:

MGSHMTNDIYFMTLAIEEAKKAAQLGEVPIGAITKDDEVIARAHNLRETLQQPTAH
AEHIAIERAAKVLGSRLEGCTLYVTLEPCVMCAGTIVMSRIPRVVYGADDPKGGCS
GSLMNLQSNFNRHRAIVDKGVLKEACSTLLTFFKNLRANKKSTN (SEQ ID NO:
105)

[0239] *Bacillus subtilis* TadA:

MTQDELYMKEAIKEAKKAEKGEVPIGAVLVINGEIIARAHNLRETEQRSIAHAEML
VIDEACKALGTWRLEGATLYVTLEPCPMCAGAVVLSRVEKVVFAGFDPKGGCSGTL
MNLLQEERFNHQAEEVSGVLEEECGGMLS AFFREL RKKKKAARKNLSE (SEQ ID
NO: 106)

[0240] *Salmonella typhimurium* (*S. typhimurium*) TadA:

MPPAFITGVTSLSDVELDHEYWMRHALTLAKRAWDEREVPVGAVLVHNHRVIGEG
WNRPIGRHDPTAHAEIMALRQGGLVLQNYRLDITLYVTLEPCVMCAGAMVHSRIG
RVVFGARDAKTGAAGSLIDVLHHPGMNHRVEIIEGVLRDECATLLSDFFRMRRQEIK
ALKKADRAEGAGPAV (SEQ ID NO: 107)

[0241] *Shewanella putrefaciens* (*S. putrefaciens*) TadA:

MDEYWMQVAMQMAEKAEAAAGEVPVGAVLVKDGQQIATGYNLSISQHDPTAHAEI
LCLRSAGKKLENYRLLDATLYITLEPCAMCAGAMVHSRIARVVYGARDEKTGAAGT

VVNLLQHAPFNHQVEVTSGVLAEACSAQLSRFFKRRRDEKKALKLAQRAQQGIE
(SEQ ID NO: 108)

[0242] *Haemophilus influenzae* F3031 (*H. influenzae*) TadA:

MDAAKVRSEFDEKMMRYALELADKAEALGEIPVGAVLVDDARNIIGEGWNLSIVQS
DPTAHAEIIALRNGAKNIQNYRLNSTLYVTLEPCTMCAGAILHSRIKRLVFGASDYK
TGAIGSRFHFFDDYKMNHTLEITSGVLAEECSQLSTFFQKRREEKKIEKALLKSLSD
K (SEQ ID NO: 109)

[0243] *Caulobacter crescentus* (*C. crescentus*) TadA:

MRTDESEDQDHRMMRLALDAARAAAEAGETPVGAVILDPSTGEVIATAGNGPIAAH
DPTAHAEIAAMRAAAAKLGNRYRLDTLTVTLEPCAMCAGAISHARIGRVVFGADD
PKGGAVVHGPKFFAQPTCHWRPEVTGGVLADESADLLRGFFRARRKAKI (SEQ ID
NO: 110)

[0244] *Geobacter sulfurreducens* (*G. sulfurreducens*) TadA:

MSSLKKTPIRDDAYWMGKAIREAAKAAARDEVPIGAVIVRDGAVIGRGHNLREGSN
DPSAHAEMIAIRQAARRSANWRLTGATLYVTLEPCLMCMGAILARLERVVFGCYDP
KGGAAGSLYDLSADPRLNHQVRLSPGVCQEECGTMLSDFFRDLRRRKAKATPALF
IDERKVPPEP (SEQ ID NO: 111)

[0245] *Streptococcus pyogenes* (*S. pyogenes*) TadA

MPYSLEEQTYFMQEALKEAEKSLQKAEIPIGCVIVKDGEIIGRGHNAREESNQAIMHA
EIMAINANAHEGNWRLLDTTLFVTIEPCVMCSGAIGLARIPHVIYGASNQKFGGADS
LYQILTDERLNHRVQVERGLLAADCANIMQTFFRQGRERKKIAKHLIKEQSDPFD
(SEQ ID NO: 112)

[0246] TadA 7.10:

SEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD
(SEQ ID NO: 113)

[0247] TadA 7.10 (V106W) (*E. coli*)

SEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGWRNAKT
GAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD
(SEQ ID NO: 114)

[0248] TadA-8e (*E. coli*)

SEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRG
AAGSLMNVNLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSIN
(SEQ ID NO: 115)

[0249] TadA-8e(V106W) (*E. coli*)

SEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGWRNSKR
GAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSIN
(SEQ ID NO: 116)

Base editors

[0250] In some aspects, the present disclosure provides eVLPs and fusion proteins for delivering base editors. Base editors are known in the art, and the presently described BE-VLPs may be used to deliver any base editor that is already known, or that is developed in the future. The base editors contemplated for delivery may comprise an amino acid sequence that is at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to any one of the base editor sequences provided herein.

[0251] In some aspects, the BE-VLPs of the present disclosure comprise cytidine base editors (CBEs) comprising a napDNAbp domain and a cytosine deaminase domain that enzymatically deaminates a cytosine nucleobase of a C:G nucleobase pair to a uracil. The uracil may be subsequently converted to a thymine (T) by the cell's DNA repair and replication machinery. The mismatched guanine (G) on the opposite strand may subsequently be converted to an adenine (A) by the cell's DNA repair and replication machinery. In this manner, a target C:G nucleobase pair is ultimately converted to a T:A nucleobase pair.

[0252] In some aspects, the BE-VLPs of the disclosure comprise the use of a cytidine base editor. Exemplary cytidine base editors include, but are not limited to, BE3, BE3.9max, BE4max, BE4-SaKKH, BE3.9-NG, BE3.9-NRRH, or BE4max-VRQR. Other cytidine base editors are known in the art, and a person of ordinary skill in the art would recognize which cytidine base editors could be delivered using the BE-VLPs of the present disclosure.

[0253] The CBEs in the BE-VLPs described herein may further comprise one or more nuclear localization signals (NLSs) and/or one or more uracil glycosylase inhibitor (UGI) domains. Thus, the base editors may comprise the structure: NH₂-[first nuclear localization sequence]-[cytosine deaminase domain]-[napDNAbp domain]-[first UGI domain]-[second UGI domain]-[second nuclear localization sequence]-COOH, wherein each instance of “]-[” indicates the presence of an optional linker sequence. Exemplary CBEs may have a structure that comprises the “BE4max” architecture, with an NH₂-[NLS]-[cytosine deaminase]-[Cas9 nickase]-[UGI domain]-[UGI domain]-[NLS]-COOH structure, having optimized nuclear localization signals and wherein the napDNAbp domain comprises a Cas9 nickase. This BE4max structure was reported to have optimized codon usage for expression in human cells,

as reported in Koblan *et al.*, *Nat Biotechnol.* 2018;36(9):843-846, incorporated herein by reference.

[0254] In other embodiments, CBEs may have a structure that comprises a modified BE4max architecture that contains a napDNAbp domain comprising a Cas9 variant other than Cas9 nickase, such as SpCas9-NG, xCas9, or circular permutant CP1028. Accordingly, exemplary CBEs may comprise the structure: NH₂-[NLS]-[cytosine deaminase]-[xCas9]-[UGI domain]-[UGI domain]-[NLS]-COOH; or NH₂-[NLS]-[cytosine deaminase]-[SpCas9-NG]-[UGI domain]-[UGI domain]-[NLS]-COOH, , wherein each instance of “[]-[” indicates the presence of an optional linker sequence.

[0255] The CBEs in the presently disclosed BE-VLPs may comprise modified (or evolved) cytosine deaminase domains, such as deaminase domains that recognize an expanded PAM sequence, have improved efficiency of deaminating 5'-GC targets, and/or make edits in a narrower target window. In some embodiments, the disclosed cytidine base editors comprise evolved nucleic acid programmable DNA binding proteins (napDNAbp), such as an evolved Cas9.

[0256] Exemplary cytidine base editors are disclosed herein and may also comprise amino acid sequences that are at least 85%, at least 90%, at least 95%, at least 98%, at least 99%, or at least 99.5% identical to any one of the amino acid sequences disclosed herein. In particular embodiments, the cytidine base editors comprise an amino acid sequence that is at least 90% identical to any one of the CBE sequences disclosed herein. In particular embodiments, the disclosed cytidine nucleobase editors comprise the amino acid sequence of any one of the CBE sequences disclosed herein. Non-limiting examples of C to T nucleobase editors are provided below:

[0257] His₆-rAPOBEC1-XTEN-dCas9 for Escherichia coli expression

MGSSHHHHHHMSSETGPVAVDPTLRRRIEPHEFEVFFDPREL RKETCLLYEINWGGR
HSIWRHTSQNTNKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSR
YPHVTLFIYIARLYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNE
AHWPRYPHLWVRLYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHIL
WATGLKSGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVP SKKFKVLG
NTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNE MAKVD
DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLR
LIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKA
ILSARLSKSRLENLIAQLPGEKKNLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD
TYDDDLNLLAQIGDQYADLFLAAKNLSDAILSDILRVNTEITKAPLSASMIKRYDE
HHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKI
LTFRIPIYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVLPHKSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN

RKVTVKQLKEDYFKKIECFDSVEISGVEDRNFNASLGTYHDLLKIIKDKDFLDNEENED
 ILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRD
 KQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAG
 SPAIKKGILQTVKVVDDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEE
 GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVP
 QSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDN
 LTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
 LKSKLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDY
 KVDYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEI
 VWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDP
 KKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAK
 GYKEVKKDLIIKLPKYSLEFENGGRKMLASAGELQKGNELALPSKYVNFLYLASHY
 EKLKGSPEDNEQKQLFVEQHKHYLDEIIIEQISEFSKRVLADANLDKVL SAYNKHARDK
 PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSIITGLYETRI
 DLSQLGGDSGGSPKKKRKV (SEQ ID NO: 117)

[0258] rAPOBEC1-XTEN-dCas9-NLS for mammalian expression

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELKRETCLLYEINWGGRHSIWRHTSQNT
 NKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIAR
 LYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNEAHWPYRPHLW
 VRLYVLELYCIIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSET
 PGTSSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLI
 GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFHRLSEESFL
 VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHAMIK
 FRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSR
 LENLIAQLPGEKKNGFLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLNL
 LAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLK
 ALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN
 REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYYVGP
 LARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPK
 HSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE
 DYFKKIECFDSVEISGVEDRNFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLF
 EDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFL
 KSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTV
 KVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKE
 HPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDN
 KVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLS
 ELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFR
 KDFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVDYDVRKMI
 AKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
 ATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPT
 TVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLI
 IKLPKYSLEFENGGRKMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDN
 EQKQLFVEQHKHYLDEIIIEQISEFSKRVLADANLDKVL SAYNKHARDKPIREQAENIIH
 LFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSIITGLYETRIDLSQLGGDSG
 GSPKKKRKV (SEQ ID NO: 118)

[0259] hAPOBEC1-XTEN-dCas9-NLS for Mammalian expression

MTSEKGPSTGDPTLRRRIEPWEFDVFYDPRELKKEACLLYEIKWGMSRKIWRSSGKN
 TTNHVEVNFIIKFTSERDFHPSMSCSITWFLSWSPCWECSQAIREFLSRHPGVTLVIYV
 ARLFWHMDQQNRQGLRDLVNSGVTIQIMRASEYYHCWRNFVNYPGDEAHWPQY
 PPLWMMLYALELHCILSLPPCLKISRRWQNHLTFFRLHLQNCHYQTIPPHILLATGLI
 HPSVAWRSGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLG
 NTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD
 DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR
 LIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKA
 ILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD
 TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDE
 HHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
 GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKI
 LTRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
 NLPNEKVLPHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN
 RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED
 ILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRD
 KQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAG
 SPAIKKGILQTVKVVDDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEE
 GIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVP
 QSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDN
 LTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
 LKSKLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDY
 KVDYVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEI
 VWDKGRDFATVRKVLSPQVNVKTEVQTGGFSKESILPKRNSDKLIARKKDWDP
 KKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAK
 GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFYLYLASHY
 EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHDR
 KPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQSIITGLYETR
 IDLSQLGGDSGGSPKKKRKV (SEQ ID NO: 119)

[0260] rAPOBEC1-XTEN-dCas9-UGI-NLS

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELKKECTCLLYEINWGGRHSIWRHTSQNT
 NKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIAR
 LYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNPSNEAHWPYRPHLW
 VRLYVLELYCIIHLGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSET
 PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLI
 GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDSDSFFHRLEESFL
 VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHMIK
 FRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILLSARLSKSRR
 LENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLNLL
 LAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLK
 ALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN
 REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTRIPYYVG
 PLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPH
 HSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE
 DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTF
 EDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFL
 KSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTV
 KVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKE

HPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDN
 KVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLS
 ELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFR
 KDFQFYKVRINNYHHAHDAYLNAVVGTAIHKYPKLESEFVYGDYKVYDVRKMI
 AKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
 ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSP
 TVAYSVLVVAKEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLI
 IKLPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDN
 EQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLSAYNKHARDKPIREQAENIIH
 LFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSTGLYETRIDLSQLGGDSG
 GSTNLSDIIEKETGKQLVIQESILMLPEEVVEEVIGNKPESDILVHTAYDESTDENVMMLT
 SDAPEYKPWALVIQDSNGENKIKMLSGGSPKKRKV (SEQ ID NO: 120)

[0261] rAPOBEC1-XTEN-SpCas9 nickase-UGI-NLS (BE3)

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELKRETCCLYEINWGGRHSIWRHTSQNT
 NKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIAR
 LYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNPSNEAHWPYPHLLW
 VRLYVLELYCIIILGLPCLNLRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSET
 PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNI
 GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEAMAKVDDSFHRLSEESFL
 VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHAMIK
 FRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRR
 LENLIAQLPGEKKNGFLGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLN
 LAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLK
 ALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN
 REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYVGP
 LARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPK
 HSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE
 DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTITLF
 EDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFL
 KSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDLSLHEHIANLAGSPAIIKKGILQTV
 KVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKE
 HPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVHDHIVPQSFLKDDSIDN
 KVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLS
 ELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFR
 KDFQFYKVRINNYHHAHDAYLNAVVGTAIHKYPKLESEFVYGDYKVYDVRKMI
 AKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
 ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSP
 TVAYSVLVVAKEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLI
 IKLPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDN
 EQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLSAYNKHARDKPIREQAENIIH
 LFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSTGLYETRIDLSQLGGDSG
 GSTNLSDIIEKETGKQLVIQESILMLPEEVVEEVIGNKPESDILVHTAYDESTDENVMMLT
 SDAPEYKPWALVIQDSNGENKIKMLSGGSPKKRKV (SEQ ID NO: 121)

[0262] pmCDA1-XTEN-dCas9-UGI (bacteria)

MTDAEYVRIHEKLDIYTFKKQFFNNKKSVSHRCYVLFELKRRGERRACFWGYAVNK
 PQSGTERGIHAEIFSIRKVEEYLRDNPQGFTINWYSSWSPCADCAEKILEWYNQELRG
 NGHTLKIWACKLYYEKNARNQIGLWNLRDNGVGLNVMVSEHYQCCRKIFIQSSH

QLNENRWLEKTLKRAEKRRSELSIMIQVKILHTTKSPAVSGSETPGTSESATPESDKK
 YSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEAT
 RLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHERHPIFG
 NIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHMIKFRGHFLIEGDLNPD
 NSDVDKLFQILVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKN
 GLFGNLIALLSLGLTPNFKSNDLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLA
 AKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
 FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
 GSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT
 RKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPHSLLYEYFTVYNE
 LTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEI
 SGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTY
 AHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFQM
 LIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKGILQTVKVVDELVKVMG
 RHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEK
 LYLYYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLTRSDKNRGK
 SDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRLV
 ETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVRINN
 YHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSESEQEIGKATAK
 YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVN
 IVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVE
 KGKSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKKDLIIKLPKYSLFELENG
 RKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKH
 YLDEIIEQISEFSKRVLADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAF
 KYFDTTIDRKRYTSTKEVLDTLHQITGLYETRIDLSQLGGDSGGSMTNLSDIIEKE
 TGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMMLTSDAPEYKPWA
 LVIQDSNGENKIKML (SEQ ID NO: 122)

[0263] pmCDA1-XTEN-nCas9-UGI-NLS (mammalian construct)

MTDAEYVRIHEKLDIYTFKKQFFNNKKSVSHRICYVLFELKRRGERRACFWGYAVNK
 PQSGTERGIHAEIFSIRKVEEYLRDNPQGFTINWYSSWSPCADCAEKILEWYNQELRG
 NGHTLKIWACKLYYEKNARNQIGLWNLRDNGVGLNVMVSEHYQCCRKIFIQSSHN
 QLNENRWLEKTLKRAEKRRSELSIMIQVKILHTTKSPAVSGSETPGTSESATPESDKK
 YSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEAT
 RLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHERHPIFG
 NIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHMIKFRGHFLIEGDLNPD
 NSDVDKLFQILVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKN
 GLFGNLIALLSLGLTPNFKSNDLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLA
 AKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI
 FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
 GSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT
 RKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPHSLLYEYFTVYNE
 LTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEI
 SGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTY
 AHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFQM
 LIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKGILQTVKVVDELVKVMG
 RHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEK
 LYLYYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLTRSDKNRGK
 SDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRLV

ETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVVREINN
 YHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVVYDVRKMIKSEQEIGKATAK
 YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVN
 IVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVE
 KGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLEFENG
 RKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKH
 YLDEIIEQISEFSKRVLADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAF
 KYFDTTIDRKRYTSTKEVLDATLIHQSTGLYETRIDLSQLGGDSGGSTNLSDIIEKETG
 KQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMMLTSDAPEYKPWALV
 IQDSNGENKIKMLSGGSPKKKRKV (SEQ ID NO: 123)

[0264] huAPOBEC3G-XTEN-dCas9-UGI (bacteria)

MDPPTFTFNFNNEPWVRGRHETLYCYEVERMHNDTWVLLNQRRGFLCNQAPHKHG
 FLEGRHAELCFLDVIPFWKLDLDQDYRVTCFTSWSPCFSCAQEMAKFISKKNHVSICI
 FTARIYDDQGRQCQGLRTLAEAGAKISIMTYSEFKHCWDTFVDHQQGCPFPWDGLD
 EHSQDLGRLRAILQSGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPS
 KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS
 NEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVD
 STDKADLRILIYLAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPIN
 ASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAE
 DAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS
 ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKF
 IKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLK
 DNREKIEKILTRIPYYVGPLARGNSRFAMTRKSEETITPWNFEEVVDKGASAQSF
 ERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA
 VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDF
 LDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWRSL
 RKLINGIRDKQSGKTILDFLKSDFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLH
 EHLANLAGSPAIIKGLITVKVDELVKVMGRHKPENIVIEARENQTTQKGQKNSR
 ERMKRIEIEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSD
 YDVAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNPSEEVVKMKKNYWRQLLNAK
 LITQRKFDNLTKAERGGLSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDEND
 KLIREVKVITLKSCLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKL
 ESEFVYGDYKVVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPL
 IETNGETGEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLI
 ARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSEFEK
 NPIDFLEAKGYKEVKKDLIIKLPKYSLEFENG RKRMLASAGELQKGNELALPSKYV
 NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL
 SAYNKHDKPIREQAENIIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVLDATLIH
 QSITGLYETRIDLSQLGGDSGGSMTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKP
 ESDILVHTAYDESTDENVMMLTSDAPEYKPWALVIQDSNGENKIKML (SEQ ID NO:
 124)

[0265] huAPOBEC3G-XTEN-nCas9-UGI-NLS (mammalian construct)

MDPPTFTFNFNNEPWVRGRHETLYCYEVERMHNDTWVLLNQRRGFLCNQAPHKHG
 FLEGRHAELCFLDVIPFWKLDLDQDYRVTCFTSWSPCFSCAQEMAKFISKKNHVSICI
 FTARIYDDQGRQCQGLRTLAEAGAKISIMTYSEFKHCWDTFVDHQQGCPFPWDGLD
 EHSQDLGRLRAILQSGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPS
 KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS

NEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVD
 STDKADLRILIYLAHAMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPIN
 ASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAE
 DAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS
 ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKF
 IKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLK
 DNREKIEKILTRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSF
 ERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI
 VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDF
 LDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLS
 RKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLH
 EHIANLAGSPAIIKGILQTVKVVDDELVKVMGRHKPENIVIEEMARENQTTQKGQKNSR
 ERMKRIEIEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSD
 YDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAK
 LITQRKFDNLTKAERGGLSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDEND
 KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKL
 ESEFVYGDYKVYDVRKMIKSEGEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPL
 IETNGETGEIVWDKGRDFATVRKVLSPQVNVKKTEVQTGGFSKESILPKRNSDKLI
 ARKKDWDPKKYGGFDSPTVAYSVLVVAKEKKGSKKLKSVKELLGITIMERSSSF
 NPIDFLEAKGYKEVKKDLIIKLPKYSLEFENGRKRMLASAGELQKGNELALPSKYV
 NFLYLASHYEKLKGPEDNEQKQLFVEQHKHYLDEIEQISEFSKRVLADANLDKVL
 SAYNKHARDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLH
 QSITGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPE
 SDILVHTAYDESTDENVMMLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRK
 V (SEQ ID NO: 125)

[0266] huAPOBEC3G (D316R_D317R)-XTEN-nCas9-UGI-NLS (mammalian construct)

MDPPTFTFNFNNEPWVRGRHETLYCYEVERMHNDTWVLLNQRRGFLCNQAPHKHG
 FLEGRHAELCFLDVIPFWKLDLDQDYRVTCFTSWSPCFSCAQEMAKFISKNNKHVSLCI
 FTARIYRRQGRCEGLRTLAEAGAKISIMTYSEFKHCWDTFVDHQQCPFPWDGLD
 EHSQDLSGRLRAILQSGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPS
 KKFVKLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS
 NEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVD
 STDKADLRILIYLAHAMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPIN
 ASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAE
 DAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS
 ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKF
 IKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLK
 DNREKIEKILTRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSF
 ERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI
 VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDF
 LDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLS
 RKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLH
 EHIANLAGSPAIIKGILQTVKVVDDELVKVMGRHKPENIVIEEMARENQTTQKGQKNSR
 ERMKRIEIEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSD
 YDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAK
 LITQRKFDNLTKAERGGLSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDEND
 KLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKL
 ESEFVYGDYKVYDVRKMIKSEGEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPL

IETNGETGEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLI
 ARKKDWDPKKYGGFDSPTVAYSVLVVAKEVEKGKSKKLKSVKELLGITIMERSSEFEK
 NPIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGELQKGNELALPSKYV
 NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL
 SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH
 QSITGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPE
 SDILVHTAYDESTDENVMMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRK
 V (SEQ ID NO: 126)

[0267] High fidelity nucleobase editor

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELKRETCCLLYEINWGGRHSIWRHTSQNT
 NKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIAR
 LYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNPSNEAHWPYRPHLW
 VRLYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSET
 PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLI
 GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFL
 VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHAMIK
 FRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRR
 LENLIAQLPGEKKNGFLGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLN
 LAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLK
 ALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN
 REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYYVGP
 LARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTAFDKNLPNEKVLPK
 HSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE
 DYFKKIECFDSVEISGVEDRFNASLGTYHDLKIKDKDFLDNEENEDILEDIVLTTLTF
 EDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGALSRLKINGIRDKQSGKTILDFL
 KSDGFANRNFMAIHDDSLTFKEDIQKAQVSGQGDLSHEHIANLAGSPAIIKKGILQTV
 KVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIIEGIKELGSQILKE
 HPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDN
 KVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLS
 ELDKAGFIKQQLVETRAITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFR
 KDFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMI
 AKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
 ATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPT
 TVAYSVLVVAKEVEKGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLI
 IKLPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDN
 EQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKHRDKPIREQAENIIH
 LFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGD
 (SEQ ID NO: 127)

[0268] rAPOBEC1-XTEN-SaCas9n-UGI-NLS) (SaBE3 and SaBE3.9max)

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELKRETCCLLYEINWGGRHSIWRHTSQNT
 NKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIAR
 LYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNPSNEAHWPYRPHLW
 VRLYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSET
 PGTSESATPESKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRS
KRGARRLKRRRRHRIQRVKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFS
AALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDG
EVRGSINRFKTSYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGECS

PFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEK
LEYEYEFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIK
DITARKEIENAEELLDQIAKILTIYQSSEDIQEELTNLNSLTQEEIEQISNLKGYTGTHN
LSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQOKEIPTTLVDDFILSPVVKRS
FIQSIKVINAIIKKYGLPNDIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIIRTGK
ENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNK
VLVKQEEASKKGNRTPFQYLSSSDSKISYETFKKHILNLAAGKGRISKTKEYLLEER
DINRFSVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKVKSINGGFTSFLRRK
WKFKKERNKGYKHHAEDALIINANADFIFKEWKKLDKAKKVMENQMFECKQAESMP
EIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYSTRKDDKGNTLI
VNNLNGLYDKDNDKLKLINKSPEKLLMYHHDPTQYQKLKLIMEQYGDENPLYK
YYEETGNYLTKYSKKDNGPVIKKIKYGGNKLNAHLDTDDYPNSRNKVVKLSLKP
RFDVYLDNGVYKFVTVKNLDVIKKENYEVNSKCYEEAKKLKISNQAEIFASFYN
NDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIASKTQSIK
KYSTDILGNLYEVKSKKHPQIIKKGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEV
IGNKPESDILVHTAYDESTDENVMMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSP
 KKKRKV (SEQ ID NO: 128)

[0269] rAPOBEC1-XTEN-SaCas9n-UGI-NLS

MSSETGPVAVDPTLRRRIEPHEFEVFFDPREL RKETCLLYEINWGGRHSIWRHTSQNT
 NKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIAR
 LYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNEAHWPYRPHLW
 VRLYVLELYCIILGLPPCLNLRKQPKLTFTIALQSCHYQRLPPHILWATGLKSGSET
 PGTSESATPESKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRS
 KRGARRLKRRRRHRIQRVKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFS
AALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDG
EVRGSINRFKTSYVKEAKQLLKVKAYHQLDQSFIDTYIDLLETRRTYYEGPGEFS
PFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEK
LEYEYEFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIK
DITARKEIENAEELLDQIAKILTIYQSSEDIQEELTNLNSLTQEEIEQISNLKGYTGTHN
LSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQOKEIPTTLVDDFILSPVVKRS
FIQSIKVINAIIKKYGLPNDIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIIRTGK
ENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNK
VLVKQEEASKKGNRTPFQYLSSSDSKISYETFKKHILNLAAGKGRISKTKEYLLEER
DINRFSVQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKVKSINGGFTSFLRRK
WKFKKERNKGYKHHAEDALIINANADFIFKEWKKLDKAKKVMENQMFECKQAESMP
EIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLI
VNNLNGLYDKDNDKLKLINKSPEKLLMYHHDPTQYQKLKLIMEQYGDENPLYK
YYEETGNYLTKYSKKDNGPVIKKIKYGGNKLNAHLDTDDYPNSRNKVVKLSLKP
RFDVYLDNGVYKFVTVKNLDVIKKENYEVNSKCYEEAKKLKISNQAEIFASFYK
NDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPHIIKTIAASKTQSIK
KYSTDILGNLYEVKSKKHPQIIKKGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEV
IGNKPESDILVHTAYDESTDENVMMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSP
 KKKRKV (SEQ ID NO: 129)

[0270] Nucleobase Editor 4-SSB

MSSETGPVAVDPTLRRRIEPHEFEVFFDPREL RKETCLLYEINWGGRHSIWRHTSQNT
 NKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIAR
 LYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNEAHWPYRPHLW

VRLYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSET
PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI
GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFL
VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHAMIK
FRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRR
LENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDAEDAKLQLSKDITYDDDLN
LAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLK
ALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN
REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGP
LARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPK
HSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE
DYFKKIECFDSVEISGVEDRNFASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTTLF
EDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFL
KSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTV
KVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKE
HPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDN
KVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLS
ELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFR
KDFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMI
AKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
ATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSP
TVAYSVLVVAKEKGKSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKKDLI
IKLPKYSLEFENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGPEDN
EQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHARDKPIREQAENIIH
LFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQSTGLYETRIDLSQLGGDSG
GSGGSGGSASRGVNKVILVGNLQDPEVRYMPNGGAVANITLATSESWRDKATGE
MKEQTEWHRVVLFGKLAEVASEYLRKGSQVYIEGQLRTRKWDQSGQDRYTTEVV
VNVGGTMQMLGGRQGGGAPAGGNIGGGQPQGGWGQPQQPQGGNQFSGGAQSRPQ
QSAPAAPSNEPPMDFDDIPFSGGSPKKRKV (SEQ ID NO: 130)

[0271] Nucleobase Editor 4-(GGS)₃

MSSETGPVAVDPTLRRIEPHEFEVFFDPRELKRETCLLYEINWGGRHSIWRHTSQNT
NKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIAR
LYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNPSNEAHWPYRPHLW
VRLYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSET
PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI
GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFL
VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHAMIK
FRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRR
LENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDAEDAKLQLSKDITYDDDLN
LAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLK
ALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN
REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGP
LARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPK
HSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE
DYFKKIECFDSVEISGVEDRNFASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTTLF
EDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFL
KSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTV
KVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKE

HPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDN
 KVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLS
 ELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFR
 KDFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMI
 AKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
 ATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSP
 TVAYSVLVVAKEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLI
 IKLPKYSLFELNGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDN
 EQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHARDKPIREQAENIIH
 LFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSTGLYETRIDLSQLGGDSG
 GSGGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDE
 NVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKRKV (SEQ ID NO: 131)

[0272] Nucleobase Editor 4-XTEN

MSSETGPVAVDPTLRRRIEPHEFEVFFDPREL RKETCLLYEINWGGRHSIWRHTSQNT
 NKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFYIAR
 LYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNPSNEAHWPYPHLLW
 VRLYVLELYCIIILGLPPCLNLRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSET
 PGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI
 GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFHRLSEESFL
 VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHAMIK
 FRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRR
 LENLIAQLPGEKKNGFLGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLNLL
 LAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLK
 ALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN
 REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYVGP
 LARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPK
 HSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE
 DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTTLTF
 EDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFL
 KSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDLSLHEHIANLAGSPAIIKKGILQTV
 KVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKE
 HPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDN
 KVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLS
 ELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFR
 KDFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMI
 AKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
 ATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSP
 TVAYSVLVVAKEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLI
 IKLPKYSLFELNGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDN
 EQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHARDKPIREQAENIIH
 LFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSTGLYETRIDLSQLGGDSG
 SETPGTSESATPESTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAY
 DESTDENVMMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKRKV (SEQ ID
 NO: 132)

[0273] Nucleobase Editor 4-32aa linker

MSSETGPVAVDPTLRRRIEPHEFEVFFDPREL RKETCLLYEINWGGRHSIWRHTSQNT
 NKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFYIAR

LYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNEAHWPYPHLLW
VRLYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGGS
SGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKF
KVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNE
MAKVDDSSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTD
KADLRILIYALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASG
VDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAK
LQLSKDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLK
ALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN
REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTRIPYYVGP
LARGNSRFAWMTRKSEETITPWNFEEVVDKGASQSFIERMTNFDKNLPNEKVLPKHS
LLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFK
KIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTTLTFEDRE
MIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN
RNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDEL
VKVMGRHHPENIVIAMARENQTTQKGQKNSRERMKRIIEGIKELGSQILKEHPVENTQL
QNEKLYLYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLTRSDKNRGK
SDNVPSEEVVKMKKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETR
QITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDA
YLNNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTE
ITLANGEIRKRIPIETNGETGEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESIL
PKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERS
SSFENPIDFLEAKGYKEVKKDLIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYV
NFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL
SAYNKHHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL
DATLIHQSGITGLYETRIDLSQLGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEE
VIGNKPESDILVHTAYDESTDENVMMLTSDAPEYKPWALVIQDSNGENKIKMLSGGSP
KKKRKV (SEQ ID NO: 133)

[0274] Nucleobase Editor 4-2X UGI

MSSETGPVAVDPTLRRIEPHEFEVFFDPRELKRETCCLLYEINWGGRHSIWRHTSQNT
NKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIAR
LYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNEAHWPYPHLLW
VRLYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLKSGSET
PGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDR
HSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFF
HRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYAL
AHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKS
RRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDQYADLFLAA
KNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFD
QSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH
QIHLGELHAILRRQEDFYFPLKDNREKIEKILTRIPYYVGPLARGNSRFAWMTRKSEET
ITPWNFEEVVDKGASQSFIERMTNFDKNLPNEKVLPKHSLLEYFTVYNELTKVKYVTEG
MRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGT
YHDLLKIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQLKRR
RYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQV
SGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHHPENIVIAMARENQTTQK
GQKNSRERMKRIIEGIKELGSQILKE

HPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDN
 KVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLS
 ELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFR
 KDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMI
 AKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
 ATVRKVLSPQVNIKKTEVQTTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPT
 TVAYSVLVVAKEKGKSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKKDLI
 IKLPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDN
 EQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHARDKPIREQAENIIH
 LFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSTGLYETRIDLSQLGGDSG
 GSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMML
 TSDAPEYKPWALVIQDSNGENKIKMLSGGSTNLSDIIEKETGKQLVIQESILMLPEEVE
 EVIGNKPESDILVHTAYDESTDENVMMLTSDAPEYKPWALVIQDSNGENKIKMLSGG
 SPKKKRKV (SEQ ID NO: 134)

[0275] Nucleobase Editor 4 (BE4)

MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELKRETCLLYEINWGGRHSIWRHTSQNT
 NKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIAR
 LYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSNEAHWPYPHLLW
 VRLYVLELYCIIILGLPPCLNLRKQQLTFFTIALQSCHYQRLPPHILWATGLKSGGS
 SGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPKSKF
 KVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNE
 AKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVSTD
 KADRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFQILVQTYNQLFEENPINASG
 VDAKILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAK
 LQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILSDILRVNTEITKAPLSASM
 IKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI
 LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDN
 EKIEKILTFRIPYYVGPLARGNSRFWMTRKSEETITPWNFEEVVDKKGASQAQSFIERM
 TNFDKNLPNEKVLPHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDL
 LFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDN
 EENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKL
 INGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHI
 ANLAGSPAIIKKGILQTVKVVDELVKVMGRHKKPENIVIAMARENQTTQKGQKNSRER
 MKRIEELGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYD
 VDHIWVQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLIT
 QRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
 REVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESE
 FVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET
 NGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTTGGFSKESILPKRNSDKLIAR
 KKDWDPPKYGGFDSPTVAYSVLVVAKEKGKSKKLKSVKELLGITIMERSSEKPNID
 FLEAKGYKEVKKDLIIKLPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFL
 YLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAY
 NKHARDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQST
 GLYETRIDLSQLGGDSGGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGN
 KPESDILVHTAYDESTDENVMMLTSDAPEYKPWALVIQDSNGENKIKMLSGGSGGSGG
 GSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMML
 TSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV (SEQ ID NO: 135)

[0276] BE4max (also AncBE4max)

MKRTADGSEFESPKKKRKVSSETGPVAVDPTLRRRIEPHEFEVFFDPREL RKETCLLY
 EINWGGGRHSIWRHTSQNTNKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSR
 AITEFLSRYPHVTLFIYIARLYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFV
 NYSPSNEAHWPRYPHLWVRLYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHY
 QRLPPHILWATGLKSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTN
 SVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRR
 YTRRKNRICYLQEIFSNEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVAYH
 EKYPTIYHLRKKLVDSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQ
 LVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALS
 LGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILL
 SDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA
 GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGE
 LHAILRRQEDFY PFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPW
 NFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTE
 GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNA
 SLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKV
 MKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS DGFANRNFQMQLIHDDSLTF
 KEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDDELVKVMGRHKPENIVI
 EMARENQTTQKGQKNSRERMKRIE EGikelGSQILKEHPVENTQLQNEKLYLYYLQN
 GRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEV
 VKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHV
 AQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHHAHDAY
 LNAVVG TALIKKYPKLESEFVYGDYKVYDV RKMIKSEQEIGKATAKYFFYSNIMNF
 FKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTG
 GFSKESILPKRNSDKLIARKKDWD PKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKS
 VKELLGITIMERSSSF EKNPIDFLEAKGYKEVKKDLI IKLPKYSLFELENGRKRMLASAG
 ELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEF
 SKRVILADANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK
 RYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDSGGSGSGGSTNLSDIIEKETGKQL
 VIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQD
 SNGENKIKMLSGGSGSGSGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDI
 LVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSKRTADGSEF
 EPKKKRKV (SEQ ID NO: 136)

[0277] AID-BE4max

MDSLLMNR RKFLYQFKNVRWAKGRRETYLCYVVKRRDSATSFSLDFGYLRNKNKC
 HVELLFLRYISDWLDLPGRCYRV TWFTSWSPCYDCARHVADFLRGNP NLSLRIFTAR
 LYFCEDRKA EPEGLRRLHRAGVQIAIMTFKDYFYCWNTFVENHERTFKAWEGLHEN
 SVRLSRQLRRILLPLYEVDDL RDAFRTLGLSGGSSGGSSGSETPGTSESATPESSGGSS
 GGSDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDS
 GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRL EESFLVEEDKK
 HERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHMIKFRGHFLI
 EGDLPDNDSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQ
 LPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQ
 YADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQL
 PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRK
 QRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPYYVGPLARGNS
 RFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY

FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIE
 CFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIE
 ERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFA
 NRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDE
 LVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGKELGSQILKEHPVEN
 TQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTR
 SDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTAKAERGGLSELDKA
 GFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQF
 YKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
 EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRK
 VLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS
 VLVVAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPK
 YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ
 LFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHARDKPIREQAENIIHLFTL
 TNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSTGLYETRIDLSQLGGDSGGSG
 GSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENV
 MLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSGGGSGGSTNLSDIIEKETGKQLVIQ
 SILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGE
 NKIKMLSGGSPKKRKV (SEQ ID NO: 137)

[0278] AID-VRQR-BE4max

MDSLLMNRKFLYQFKNVRWAKGRRETYLCYVVKRRDSATSFSLDFGYLRNKNKC
 HVELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGPNPNSLRIFTAR
 LYFCEDRKAEPGLRRLHRAGVQIAIMTFKDYFYCWNTFVENHERTFKAWEGLHEN
 SVRLSRQLRRILLPLYEVDDL RDAFRTLGLSGGSSGGSSGSETPGTSESATPESSGGSS
 GGSDDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS
 GETAEATRLKRTARRRYTRRKNRICYLQEFSNEMAKVDDSFHRLEESFLVEEDKK
 HERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHAMIKFRGHFLI
 EGDLPDNDSDVDKLFQQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQ
 LPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLDNLLAQIGDQ
 YADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQL
 PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRK
 QRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYVGPLARGNS
 RFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHSLLEYEY
 FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIE
 CFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIE
 ERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFA
 NRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDE
 LVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGKELGSQILKEHPVEN
 TQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTR
 SDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTAKAERGGLSELDKA
 GFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQF
 YKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
 EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRK
 VLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFVSPTVAYS
 VLVVAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPK
 YSLFELENGRKRMLASARELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ
 LFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHARDKPIREQAENIIHLFTL
 TNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHQSTGLYETRIDLSQLGGDSGGSG

GSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENV
MLLTSDAPEYKPWALVIQDSNGENKIKMLSGSGSGSGGSTNLSDIIEKETGKQLVIQ
SILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGE
NLIKMLSGGSKRTADGSEFEPKKRKV (SEQ ID NO: 138)

[0279] AncBE4max 689

MKRTADGSEFESPKKKRKVSSETGPVAVDPTLRRRIEPHEFEVFFDPRELKRETCCLY
EIKWGTSHKIWRHSSKNNTTKHVEVNFIEKFTSERHFCPSTSCSITWFLSWSPCGECSK
AITEFLSQHPNVTLYIYVARLYHHMDQQNRQGLRDLVNSGVTIQIMTAPEYDYCWR
NFVNYPPGKEAHWPYPPLWMKLYALELHAGILGLPPCLNILRRKQPQLTFFTIALQS
CHYQRLPPHILWATGLKSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAI
GTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTA
RRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEV
AYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVK
LFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLI
ALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSD
AILSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN
GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIH
LGELHAILRRQEDFYPPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETIT
PWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPHSLLYEYFTVYNELTKVKYV
TEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRF
NASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDK
VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLT
FKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIV
IEMARENQTTQKGQKNSRERMKRIEEGKELGSQILKEHPVENTQLQNEKLYLYYLQ
NGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE
VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRLVETRQITKH
VAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDA
YLNNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIM
NFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNVKKTEVQ
TGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKL
KSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLEFENGRKRMLAS
AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIIEQI
SEFSKRVLADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID
RKRYTSTKEVLDTLIHQSTGLYETRIDLSQLGGDSGGSGSGGSTNLSDIIEKETGK
QLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVI
QDSNGENKIKMLSGSGSGSGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPE
SDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSKRTADGS
EFEPKKRKV (SEQ ID NO: 139)

[0280] YE1-BE4

MKRTADGSEFESPKKKRKVSSETGPVAVDPTLRRRIEPHEFEVFFDPRELKRETCCLY
EINWGGRHSIWRHTSQNTNKHVEVNFIEKFTTERYFCPNTRCSITWFLSYSPCGECSR
AITEFLSRYPHVTLFIYIARLYHHADPENRQGLRDLISSGVTIQIMTEQESGYCWRNFV
NYSPTSNEAHWPYPHLLWVRLYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHY
QRLPPHILWATGLKSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNS
VGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK
NRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHL
RKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEN

PINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTPNFKSNFDLAED
AKLQLSKDTYDDDLNLLAQIGDQYADFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK
RYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIPILEKM
DGTEELLVKNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKIL
TFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNE
KVLPHKSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE
DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFED
REMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGF
ANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAICKGILQTVKVVDELV
KVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGKELGSQILKEHPVENTQLQN
EKLYLYYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLTRSDKNRGKSD
NVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQIT
KHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYL
NAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEIT
LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILP
KRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKGKSKKLKSVKELLGITIMERSS
FEKNPIDFLEAKGYKEVKKDLIKLPKYSLENGRKRMLASAGELQKGNELALPSKYVN
FLYLASHYEKLKGGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNK
HRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRI
DLSQLGGDSGGSGGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILV
HTAYDESTDENVMMLTSDAPEYKPWALVIQDSNGENKIKMLSGGSGGSGGSTNLSDI
IEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMMLTSDAPEYK
PWALVIQDSNGENKIKMLSGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 140)

[0281] YE2-BE4

MKRTADGSEFESPKKKRKVSSETGPVAVDPTLRRRIEPHEFEVFFDPREL RKETCLLY
EINWGGRHSIWRHTSQNTNKHVEVNFIEKFTTERYFCPNTRCSITWFLSYSPCGEC SR
AITEFLSRYPHVTLFIYIARLYHHADPRNRQGLEDLISSGV TIQIMTEQESGYCWRNFV
NYSPTSNEAHWPYPH LWRLYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHY
QRLPPHILWATGLKSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNS
VGWAVITDEYKVP SKKFV LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK
NRICYLQEIFS NEMAKVDDSF FHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHL
RKKLV DSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEN
PINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTPNFKSNFDLAED
AKLQLSKDTYDDDLNLLAQIGDQYADFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK
RYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIPILEKM
DGTEELLVKNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKIL
TFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNE
KVLPHKSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE
DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFED
REMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGF
ANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAICKGILQTVKVVDELV
KVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGKELGSQILKEHPVENTQLQN
EKLYLYYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLTRSDKNRGKSD
NVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQIT
KHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYL
NAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEIT
LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILP
KRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKGKSKKLKSVKELLGITIMERSS

FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVN
 FLYLASHYEKLKGGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNK
 HRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRI
 DLSQLGGDSGGSGGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILV
 HTAYDESTDENVMMLTSDAPEYKPWALVIQDSNGENKIKMLSGGSGGSGGSTNLSDI
 IEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMMLTSDAPEYK
 PWALVIQDSNGENKIKMLSGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 141)

[0282] YEE-BE4

MKRTADGSEFESPKKKRKVSSETGPVAVDPTLRRRIEPHEFEVFFDPREL RKETCLLY
 EINWGGRHSIWRHTSQNTNKHVEVNFIEKFTTERYFCPNTRCSITWFLSYSPCGECSR
 AITEFLSRYPHVTLFIYIARLYHHADPENRQGLEDLISSGVTIQIMTEQESGYCWRNFV
 NYSPSNEAHWP RYPHLWVRLYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHY
 QRLPPHILWATGLKSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNS
 VGWAVITDEYKVP SKKFVVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK
 NRICYLQEIFS NEMAKVDDSF FHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHL
 RKKLVDSTDKADRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFQVLVQTYNQ LFEEN
 PINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTPNFKSNFDLAED
 AKLQLSKDTYDDDLNLLAQIGDQYADFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK
 RYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKM
 DGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKIL
 TFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNE
 KVLPHKSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE
 DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTTLTFED
 REMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGF
 ANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDDELV
 KVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQN
 EKLYLYYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLT RSDKNRGKSD
 NVPSEEVVKMKKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQIT
 KHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHHAHDAYL
 NAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEIT
 LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILP
 KRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERS
 FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVN
 FLYLASHYEKLKGGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNK
 HRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRI
 DLSQLGGDSGGSGGSGGSGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILV
 HTAYDESTDENVMMLTSDAPEYKPWALVIQDSNGENKIKMLSGGSGGSGGSTNLSDI
 IEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMMLTSDAPEYK
 PWALVIQDSNGENKIKMLSGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 142)

[0283] EE-BE4

MKRTADGSEFESPKKKRKVSSETGPVAVDPTLRRRIEPHEFEVFFDPREL RKETCLLY
 EINWGGRHSIWRHTSQNTNKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSR
 AITEFLSRYPHVTLFIYIARLYHHADPENRQGLEDLISSGVTIQIMTEQESGYCWRNFV
 NYSPSNEAHWP RYPHLWVRLYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHY
 QRLPPHILWATGLKSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNS
 VGWAVITDEYKVP SKKFVVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK
 NRICYLQEIFS NEMAKVDDSF FHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHL

RKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEEN
 PINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTPNFKSNDLAED
 AKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK
 RYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKM
 DGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKIL
 TFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNE
 KVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE
 DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFED
 REMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGF
 ANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELV
 KVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQN
 EKLYLYYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLTTRSDKNRGKSD
 NVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQIT
 KHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYL
 NAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEIT
 LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILP
 KRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERS
 FEKNPIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGELQKGNELALPSKYVN
 FLYLASHYEKLKGGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNK
 HRDKPIREQAENIIHLFTLTNLGAPAAFYFDTTIDRKRYTSTKEVLDTLIHQSTGLYETRI
 DLSQLGGDSGGSGGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILV
 HTAYDESTDENVMMLTSDAPEYKPWALVIQDSNGENKIKMLSGGSGGSGGSTNLSDI
 IEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMMLTSDAPEYK
 PWALVIQDSNGENKIKMLSGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 143)

[0284] R33A-BE4

MKRTADGSEFESPKKKRKVSSETGPVAVDPTLRRRIEPHEFEVFFDPRELAKETCLLY
 EINWGGRHSIWRHTSQNTNKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSR
 AITEFLSRYPHVTLFIYIARLYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFV
 NYSPSNEAHWPRYPHLWVRLYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHY
 QRLPPHILWATGLKSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNS
 VGWAVITDEYKVPSSKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK
 NRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHL
 RKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEEN
 PINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTPNFKSNDLAED
 AKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK
 RYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKM
 DGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKIL
 TFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNE
 KVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE
 DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFED
 REMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGF
 ANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELV
 KVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQN
 EKLYLYYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLTTRSDKNRGKSD
 NVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQIT
 KHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYL
 NAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEIT
 LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILP

KRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERS
FEKNPIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGELQKGNELALPSKYVN
FLYLASHYEKLKGGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNK
HRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRI
DLSQLGGDSGGSGGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILV
HTAYDESTDENVMMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSGGSGGSTNLSDI
IEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMMLLTSDAPEYK
PWALVIQDSNGENKIKMLSGGSGKRTADGSEFEPKKKRKV (SEQ ID NO: 144)

[0285] R33A+K34A-BE4

MKRTADGSEFESPKKKRKVSSETGPVAVDPTLRRRIEPHEFEVFFDPRELA AETCLLY
EINWGGRHSIWRHTSQNTNKHVEVNFIEKFTTTERYFCPNTRCSITWFLSWSPCGECSR
AITEFLSRYPHVTLFIYIARLYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFV
NYSPTSNEAHWPYPHLYVRLYVLELYCIIILGLPPCLNILRRKQPQLTFFTIALQSCHY
QRLPPHILWATGLKSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNS
VGWAVITDEYKVP SKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK
NRICYLQEIFS NEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHL
RKKLVDSTDKADLRILIYLAHAMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEEN
PINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAED
AKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK
RYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKM
DGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKIL
TFRIPYVVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNE
KVLPHKSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE
DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTTLTFED
REMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGF
ANRNFMQLIHDDSLTFKEDIQKAQVSGQGD SLHEHIANLAGSPA IKKGILQTVKVVD ELV
KVMGRHKPENIVIE MARENQTTQKGQKNSRERMKRIEEG IKELGSQILKEHPVENTQLQN
EKLYLYYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLT RSDKNRGKSD
NVPSEEVVKMKKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQIT
KHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHHAH DAYL
NAVVG TALIKKYPKLESEFVYG DYKVYDVRKMI AKSEQEIGKATAKYFFYSNIMNFFKTEIT
LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILP
KRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERS
FEKNPIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGELQKGNELALPSKYVN
FLYLASHYEKLKGGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNK
HRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRI
DLSQLGGDSGGSGGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILV
HTAYDESTDENVMMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSGGSGGSTNLSDI
IEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMMLLTSDAPEYK
PWALVIQDSNGENKIKMLSGGSGKRTADGSEFEPKKKRKV (SEQ ID NO: 145)

[0286] FERNY-BE4

MKRTADGSEFESPKKKRKVFERNYDPREL RKETYLLYEIKWGKSGKLWRHWCQNN
RTQHA EYV FLENIFNARRFNPSTHCSITWYLSWSPCAECSQKIVDFLKEHPNVNLEIY
VARLYYHEDERNRQGLRDLVNSGV TIRIMDL PDYNYCWKTFVSDQGGDEDYWP GH
FAPWIKQYSLKLSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVG
WAVITDEYKVP SKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK NR
CYLQEIFS NEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRK

KLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEENPI
 NASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTPNFKSNDLAEDAK
 LQLSKDQYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY
 DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDG
 TEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTR
 IPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVL
 PKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYF
 KKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREM
 IEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANR
 NFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDELVKVM
 GRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGKELGSQILKEHPVENTQLQNEKL
 YLYYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLTSDKNRGKSDNVP
 SEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHV
 AQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYLNAV
 VGTALIKKYPKLESEFVYGDYKVYDVVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
 GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRN
 SDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSEFEK
 NPIDFLEAKGYKEVKKDLIKLPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFYLY
 LASHYEKLKSGPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH
 DKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQITGLYETRIDL
 SQLGGDSGGSGGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHT
 AYDESTDENVMMLTSDAPEYKPWALVIQDSNGENKIKMLSGGSGGSGGSTNLSDIIE
 KETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMMLTSDAPEYKP
 WALVIQDSNGENKIKMLSGGSKRTADGSEFEPKKRKY (SEQ ID NO: 146)

[0287] AALN-BE4

MKRTADGSEFESPKKKRKVSSETGPVAVDPTLRRRIEPHEFEVFFDPRELA AETCLLY
 EINWGGRHSIWRHTSQNTNKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECRSR
 AITEFLSRYPHVTLFIYIARLYHLANPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFV
 NYSPSNEAHWPRYPHLWVRLYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHY
 QRLPPHILWATGLKSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNS
 VGWAVITDEYKVPSSKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK
 NRICYLQEIFS NEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHL
 RKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEEN
 PINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTPNFKSNDLAED
 AKLQLSKDQYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK
 RYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKM
 DGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKIL
 TRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNE
 KVLPKHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE
 DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFED
 REMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGF
 ANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDELV
 KVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGKELGSQILKEHPVENTQLQN
 EKLYLYYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLTSDKNRGKSD
 NVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQIT
 KHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYL
 NAVVG TALIKKYPKLESEFVYGDYKVYDVVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEIT
 LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILP

KRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSS
FEKNPIDFLEAKGYKEVKKDLIHKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVN
FLYLASHYEKLKGGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNK
HRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRI
DLSQLGGDSGGSGGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILV
HTAYDESTDENVMMLTSDAPEYKPWALVIQDSNGENKIKMLSGGSGGSGGSTNLSDI
IEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMMLTSDAPEYK
PWALVIQDSNGENKIKMLSGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 147)

[0288] BE4max, modified with SpCas9-NG ("BE4-NG")

MKRTADGSEFESPKKKRKVSSETGPVAVDPTLRRRIEPHEFEVFFDPREL RKETCLLY
EINWGGRHSIWRHTSQNTNKHVEVNFIEKFTTTERYFCPNTRCSITWFLSWSPCGECSR
AITEFLSRYPHVTLFIYIARLYHHADPRNRQGLRDLISSGV TIQIMTEQESGYCWRNFV
NYSPTSNEAHWP RYPHLWVRLYVLELYCIILGLPPCLNLRKQ PQLTFFTIALQSCHY
QRLPPHILWATGLKSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNS
VGWAVITDEYKVP SKKFV LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK
NRICYLQEIFS NEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHL
RKKLVDSTDKADLR LIYLALAHMIKFRGHFLIEGDLNPDNSD VDKLFIQLVQTYNQLFEEN
PINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTPNFKSNFDLAED
AKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK
RYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKM
DGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKIL
TFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVV DKGASAQSFIERMTNFDKNLPNE
KVLPHKSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE
DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTTLTFED
REMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGF
ANRNFMQLIHDDSLTFKEDIQKAQVSGQGD SLHEHIANLAGSPA IKKGILQTVKVVD ELV
KVMGRHKPENIVIE MARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQN
EKLYLYYLQNGRDMYVDQELDINRLSDYDV DHIVPQSFLKDDSIDNKVLT RSDKNRGKSD
NVPSEEVVKMKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQIT
KHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHHAHDAYL
NAVVG TALIKKYPKLESEFVYG DYKVYDVRKMI AKSEQEIGKATAKYFFYSNIMNFFKTEIT
LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESIRP
KRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSS
FEKNPIDFLEAKGYKEVKKDLIHKLPKYSLFELENGRKRMLASARFLQKGNELALPSKYVN
FLYLASHYEKLKGGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNK
HRDKPIREQAENIIHLFTLTNLGAPRAFKYFDTTIDRKVYRSTKEVL DATLIHQ SITGLYETRI
DLSQLGGDSGGSGGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILV
HTAYDESTDENVMMLTSDAPEYKPWALVIQDSNGENKIKMLSGGSGGSGGSTNLSDI
IEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMMLTSDAPEYK
PWALVIQDSNGENKIKMLSGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 148)

[0289] BE4max-SaKKH

MKRTADGSEFESPKKKRKVSSETGPVAVDPTLRRRIEPHEFEVFFDPREL RKETCLLY
EINWGGRHSIWRHTSQNTNKHVEVNFIEKFTTTERYFCPNTRCSITWFLSWSPCGECSR
AITEFLSRYPHVTLFIYIARLYHHADPRNRQGLRDLISSGV TIQIMTEQESGYCWRNFV
NYSPTSNEAHWP RYPHLWVRLYVLELYCIILGLPPCLNLRKQ PQLTFFTIALQSCHY
QRLPPHILWATGLKSGGSSGGSSGSETPGTSESATPESSGGSSGGSGKRN YILGLAIGITS
VGYGIIDYETRDVIDAGVRLFK EANVENNEGRRSKRGARRLKRRRRHRIQRVKKLLFDYNL

LTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQ
 ISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVQKAYHQLDQSFI
 DTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNA
 LNDLNNLVITRDENEKLEYEKFQIENVFVKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGK
 PEFTNLKVYHDIKDITARKEIENAEELLDQIAKILTIYQSSEDIQEELTNLNSLTQEEIEQISN
 LKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILS
 PVVKRSFIQSIKVINAIKKYGLPNDIIELAREKNSKDAQKMINEMQKRNROQTNERIEEIIRT
 TGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIIPRSVSFDNSFNKVK
 LVKQEENSCKGNRTPFQYLSSSDSKISYETFKKHILNLA KGKGRISKTKKEYLLEERDINRFS
 VQKDFINRNLVDTRYATRGLMNLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNK
 GYKHAEDALIINANADFIFKEWKKLDKAKKVMENQMFEKQAESMPEIETE QEYKEIFIT
 PHQIKHIKDFKDYKYSHRVDKKNRKLINDTLYSTRKDDKGNLIVNNLNGLYDKDNDKL
 KKLINKSPEKLLMYHHDPTQYQKLKLIMEQYQYDEKNPLYKYEEETGNLYTKYSKKDNGPV
 IKKIKYYGNKLNALHDITDDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKE
 NYEVENSKCYEEAKLKKISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEVMIDIT
 YREYLENMNDKRPPHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKSGSGSGSGG
 STNLSDIIEKETGKQLVQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAP
 EYKPWALVIQDSNGENKIKMLSGSGSGSGSTNLSDIIEKETGKQLVQESILMLPEEVEEV
 IGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGSGSKRTAD
 GSEFEPKKKRKV (SEQ ID NO: 149)

[0290] BE4max-NRRH

MKRTADGSEFESPKKKRKVSSETGPVAVDPTLRRRIEPHEFEVFFDPREL RKETCLLY
 EINWGGRHSHWRHTSQNTNKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSR
 AITEFLSRYPHVTLFIYIARLYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFV
 NYSPSNEAHWPRYPHLWVRLYVLELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHY
 QRLPPHILWATGLKSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLTIGTNS
VGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRK
NRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHL
RKKLVDSTDKADRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFQOLVOTYNOLFEEN
PINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKS NFDLAED
AKLQLSKDTYDDDLDNLLAQIGDOYADFLAAKNLSDAILLSDILRVNTEITKAPLSASMVK
RYDEHHQDLTLLKALVRQOLPEKYKEIFFDOSKNYAGYIDGGASQEEFYKFIKPILEKM
DGTEELLVKLNREDLLRKORTFDNGIIPHQIHLGELHAILRROGDFYPFLKDNREKIEKILT
FRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASQSFIERMTNFDKNLPNE
KVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKOLKE
DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFED
REMIEERLKTYAHLFDDKVMKOLKRLRYTGWGRLSRKLINGIRDKOSGKTILDFLKSDGF
ANRNFQMOLIHDDSLTFKEDIQKAQVSGOGDSLHEHIANLAGSPAIKKGILOTVKVVDLV
KVMGGHKPENIVIEMARENOTTOKGOKNSRERMKRIEEGKELGSOILKEHPVENTOLON
EKLYLYYLONGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLTRSDKNRGKSD
NVPSEEVVKKMKNYWRQLLNAKLITORKFDNLTKAERGGLSELDKAGFIKROLVETROIT
KHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYL
NAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEOEIGKATAKYFFYSNIMNFFKTEIT
LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVOTGGFSKESILP
KGNSDKLIARKKDWDPKKYGGFNSPTAAYSVLVAKVEKGKSKKLKSVKELLGITIMERSS
FEKNPIGFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGVLHKGNELALPSKYVN
FLYLASHYEKLKGSPEDEOKOLFVEOHKHYLDEIIEQISEFSKRVLADANLDKVLSAYNK
HRDKPIREQAENIIHLFTLTNLGVPAAFKYFDTTIDKKRYTSTKEVLDTLIHQITGLYETRI

DLSOLGGDSGGSGGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSGGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSGKRTADGSEFEPKKRKV (SEQ ID NO: 150)

[0291] BE4max-VQR

MKRTADGSEFESP~~KKRK~~VSSSETGPVAVDPTLRRRIEPHEFEVFFDPREL~~RKETCLLY~~
EINWGGRHSIWRHTSQNTNKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSR
AITEFLSRYPHVTLFIYIARLYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFV
NYSPTSNEAHWPYPHLLWRLYVLELYCIIILGLPPCLNILRRKQPQLTFFTIALQSCHY
QRLPPHILWATGLKSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNS
VGWAVITDEYKVP~~SKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK~~
NRICYLQEIFSNEMAKVDDSF~~FHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHL~~
RKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEN
PINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAED
AKLQLSKDTYDDDLNLLAQIGDOYADFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK
RYDEHHQDLTLLKALVROQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKM
DGTEELLVKLNREDLLRKORTFDNGSIPHQIHLGELHAILRROEDFYFPLKDNREKIEKIL
TFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNE
KVLPKHSLLYEFYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKOLKE
DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTTLTFED
REMIERLKYAHLFDDKVMKOLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGF
ANRNFMOIHDDSLTFKEDIQKAQVSGOGDSLHEHIANLAGSPAIKKGILOTVKVVDDELV
KVMGRHKPENIVIMARENOTTOKGOKNSRERMKRIEEGIKELGSOILKEHPVENTOLON
EKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSD
NVPSEEVVKKMKNYWRQLLNAKLITORKFDNLTKAERGGLSELDKAGFIKROLVETROIT
KHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHAHDAYL
NAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEOEIGKATAKYFFYSNIMNFFKTEIT
LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILP
KRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAKEGKSKKLKSVKELLGITIMERS
FEKNPIDFLEAKGYKEVKKDLIKLPKYSLELENGRKRMLASAGELQGNELALPSKYVN
FLYLASHYEKLKGSPEDNEOKOLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL SAYNK
HRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDTLIHQSIITGLYETR
IDLSOLGGDSGGSGGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSGKRTADGSEFEPKKRKV (SEQ ID NO: 151)

[0292] BE4max-VRQR

MKRTADGSEFESP~~KKRK~~VSSSETGPVAVDPTLRRRIEPHEFEVFFDPREL~~RKETCLLY~~
EINWGGRHSIWRHTSQNTNKHVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSR
AITEFLSRYPHVTLFIYIARLYHHADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFV
NYSPTSNEAHWPYPHLLWRLYVLELYCIIILGLPPCLNILRRKQPQLTFFTIALQSCHY
QRLPPHILWATGLKSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNS
VGWAVITDEYKVP~~SKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK~~
NRICYLQEIFSNEMAKVDDSF~~FHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHL~~
RKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEN
PINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAED
AKLQLSKDTYDDDLNLLAQIGDOYADFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK

RYDEHHQDLTLLKALVROQLPEKYKEIFFDOSKNGYAGYIDGGASQEEFYKFIKPILEKM
DGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKIL
TRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKGASQSFIERMTNFDKNLPNE
KVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKOLKE
DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFED
REMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGF
ANRNFMOIHHDDSLTFKEDIQKAQVSGOGDSLHEHIANLAGSPAIIKKGILOTVKVVDDELV
KVMGRHKPENIVIMARENOTTOKGOKNSRERMKRIEEGKELGSQILKEHPVENTOLON
EKLYLYYLQNGRDMYVDOELDINRLSDYDVIDHIVPOSFLKDDSIDNKVLTTRSDKNRGKSD
NVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKROLVETROIT
KHVAOILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFOFYKVREINNYHHAHDAYL
NAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEI GKATAKYFFYSNIMNFFKTEIT
LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQOVNIVKKTEVOTGGFSKESILP
KRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAKEKGKSKKLKSVKELLGITIMERSS
FEKNPIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASARELOKGNELALPSKYVN
FLYLASHYEKLKGGSPEDNEOKOLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNK
HRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVL DATLIHOSITGLYETR
IDLSQLGGDSGGSGGGSGGSTNLSDIIEKETGKQLVIQESILMLPEEVVEEVIGNKPESDIL
VHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSGGGSGGSTNLS
DIIEKETGKQLVIQESILMLPEEVVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPE
YKPWALVIQDSNGENKIKMLSGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 152)

[0293] In some aspects, the BE-VLPs of the disclosure comprise an adenine base editor.

Exemplary adenine nucleobase editors include, but are not limited to, ABE7.10 (or ABE_{max}), ABE8e, ABE8e-SaKKH, ABE8e-NG, ABE-xCas9, ABE7.10-SaKKH, ABE7.10-NG, ABE7.10-VRQR, ABE7.10-VQR, ABE8e-NRTH, ABE8e-NRRH, ABE8e-VQR, or ABE8e-VRQR. In certain embodiments, the adenine base editor delivered by the BE-VLPs is an ABE8e or an ABE7.10. ABE8e is sometimes referred to herein as “ABE8” or “ABE8.0”. The ABE8e base editor and variants thereof may comprise an adenosine deaminase domain containing a TadA-8e adenosine deaminase monomer (monomer form) or a TadA-8e adenosine deaminase homodimer or heterodimer (dimer form). Other ABEs may be used to deaminate an A nucleobase.

[0294] Some aspects of the disclosure provide fusion proteins that comprise a nucleic acid programmable DNA binding protein (napDNA_{bp}) and at least two adenosine deaminase domains. Without wishing to be bound by any particular theory, dimerization of adenosine deaminases (*e.g.*, in *cis* or in *trans*) may improve the ability (*e.g.*, efficiency) of the fusion protein to modify a nucleic acid base, for example to deaminate adenine. In some embodiments, any of the fusion proteins may comprise 2, 3, 4 or 5 adenosine deaminase domains. In some embodiments, any of the fusion proteins provided herein comprises two adenosine deaminases. In some embodiments, any of the fusion proteins provided herein contains only two adenosine deaminases. In some embodiments, the adenosine deaminases

are the same. In some embodiments, the adenosine deaminases are any of the adenosine deaminases provided herein. In some embodiments, the adenosine deaminases are different.

[0295] In some embodiments, the general architecture of exemplary fusion proteins with a first adenosine deaminase, a second adenosine deaminase, and a napDNAbp comprises any one of the following structures, where NLS is a nuclear localization sequence (*e.g.*, any NLS provided herein), NH₂ is the N-terminus of the fusion protein, and COOH is the C-terminus of the fusion protein: NH₂-[first adenosine deaminase]-[second adenosine deaminase]-[napDNAbp]-COOH;

[0296] NH₂-[first adenosine deaminase]-[napDNAbp]-[second adenosine deaminase]-COOH; NH₂-[napDNAbp]-[first adenosine deaminase]-[second adenosine deaminase]-COOH; NH₂-[second adenosine deaminase]-[first adenosine deaminase]-[napDNAbp]-COOH; NH₂-[second adenosine deaminase]-[napDNAbp]-[first adenosine deaminase]-COOH; NH₂-[napDNAbp]-[second adenosine deaminase]-[first adenosine deaminase]-COOH.

[0297] In some embodiments, the fusion proteins provided herein do not comprise a linker. In some embodiments, a linker is present between one or more of the domains or proteins (*e.g.*, first adenosine deaminase, second adenosine deaminase, and/or napDNAbp). In some embodiments, the “]-[” used in the general architecture above indicates the presence of an optional linker. Exemplary fusion proteins comprising a first adenosine deaminase, a second adenosine deaminase, a napDNAbp, and an NLS are provided: NH₂-[NLS]-[first adenosine deaminase]-[second adenosine deaminase]-[napDNAbp]-COOH; NH₂-[first adenosine deaminase]-[NLS]-[second adenosine deaminase]-[napDNAbp]-COOH; NH₂-[first adenosine deaminase]-[second adenosine deaminase]-[NLS]-[napDNAbp]-COOH; NH₂-[first adenosine deaminase]-[second adenosine deaminase]-[napDNAbp]-[NLS]-COOH; NH₂-[NLS]-[first adenosine deaminase]-[napDNAbp]-[second adenosine deaminase]-COOH; NH₂-[first adenosine deaminase]-[NLS]-[napDNAbp]-[second adenosine deaminase]-COOH; NH₂-[first adenosine deaminase]-[napDNAbp]-[NLS]-[second adenosine deaminase]-COOH; NH₂-[first adenosine deaminase]-[napDNAbp]-[second adenosine deaminase]-[NLS]-COOH; NH₂-[NLS]-[napDNAbp]-[first adenosine deaminase]-[second adenosine deaminase]-COOH; NH₂-[napDNAbp]-[NLS]-[first adenosine deaminase]-[second adenosine deaminase]-COOH; NH₂-[napDNAbp]-[first adenosine deaminase]-[NLS]-[second adenosine deaminase]-COOH; NH₂-[napDNAbp]-[first adenosine deaminase]-[second adenosine deaminase]-[NLS]-COOH; NH₂-[NLS]-[second adenosine deaminase]-[first adenosine deaminase]-[napDNAbp]-COOH;

NH₂-[second adenosine deaminase]-[NLS]-[first adenosine deaminase]-[napDNAbp]-COOH;
 NH₂-[second adenosine deaminase]-[first adenosine deaminase]-[NLS]-[napDNAbp]-COOH;
 NH₂-[second adenosine deaminase]-[first adenosine deaminase]-[napDNAbp]-[NLS]-COOH;
 NH₂-[NLS]-[second adenosine deaminase]-[napDNAbp]-[first adenosine deaminase]-COOH;
 NH₂-[second adenosine deaminase]-[NLS]-[napDNAbp]-[first adenosine deaminase]-COOH;
 NH₂-[second adenosine deaminase]-[napDNAbp]-[NLS]-[first adenosine deaminase]-COOH;
 NH₂-[second adenosine deaminase]-[napDNAbp]-[first adenosine deaminase]-[NLS]-COOH;
 NH₂-[NLS]-[napDNAbp]-[second adenosine deaminase]-[first adenosine deaminase]-COOH;
 NH₂-[napDNAbp]-[NLS]-[second adenosine deaminase]-[first adenosine deaminase]-COOH;
 NH₂-[napDNAbp]-[second adenosine deaminase]-[NLS]-[first adenosine deaminase]-COOH;
 NH₂-[napDNAbp]-[second adenosine deaminase]-[first adenosine deaminase]-[NLS]-COOH.

[0298] Exemplary ABEs include, without limitation, the following fusion proteins.

[0299] In some embodiments, an A to G base editor comprises the structure of NH₂-[second adenosine deaminase]-[first adenosine deaminase]-[dCas9]-COOH. In some embodiments, the second adenosine deaminase is a wild-type ecTadA (SEQ ID NO: 153). In some embodiments, a linker is used between each domain. In some embodiments, the linker is 32 amino acids long and comprises the amino acid sequence of SGGSSGGSSGSETPGTSESATPESSGGSSGGS (SEQ ID NO: 306).

Exemplary adenine base editors comprise amino acid sequences that are at least 85%, at least 90%, at least 95%, at least 98%, at least 99%, or at least 99.5% identical to any one of the amino acid sequences SEQ ID NOs: 153-203. In particular embodiments, the disclosed adenine base editors comprise an amino acid sequence that is at least 90% identical to any of SEQ ID NOs: 153-203. In particular embodiments, the disclosed adenine base editors comprise an amino acid sequence of any of SEQ ID NOs: 153-203.

[0300] Non-limiting examples of A to G base editors are provided below, as SEQ ID NOs: 153-203.

[0301] ecTadA(wt)-XTEN-nCas9-NLS

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTD
 SGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSI
 KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLAL
 AHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDKAILSARL
 SKSRRLLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDD

LDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL
 TLLKALVRQQLEPKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL
 VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIP
 YYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNE
 KVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV
 KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV
 LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
 KTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAK
 KGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKEL
 GSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFL
 KDDSIDNKVLTRSDKNRGKSDNPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK
 AERGGSELKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 KLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVY
 DVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
 KGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY
 GGFDSPVAVSVLVVAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKE
 VKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLK
 GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLSAYNKHRRDKPIRE
 QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSIITGLYETRIDLQS
 LGGDSGGSPKKRKV (SEQ ID NO: 153)

[0302] ecTadA(D108N)-XTEN-nCas9-NLS: (mammalian construct, active on DNA)

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAAQKKAQSSTD
 SGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI
 KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADRLIYLAL
 AHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARL
 SKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDD
 LDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL
 TLLKALVRQQLEPKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL
 VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIP
 YYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNE
 KVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV
 KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV
 LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
 KTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAK
 KGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKEL
 GSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFL
 KDDSIDNKVLTRSDKNRGKSDNPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK
 AERGGSELKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 KLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVY
 DVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
 KGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY
 GGFDSPVAVSVLVVAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKE
 VKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLK
 GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLSAYNKHRRDKPIRE

QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDL SQ
LGGDSGGSPKKKRKV (SEQ ID NO: 154)

[0303] ecTadA(D108G)-XTEN-nCas9-NLS: (mammalian construct, active on DNA, A to G editing)

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARGAKT
GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTD
SGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI
KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSSFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLAL
AHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARL
SKSRRLLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDD
LDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL
TLLKALVRQQLP EKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL
VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIP
YYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNE
KVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV
KQLKEDYFKKIECFDSVEISGVEDRFNASLGT YHDL LKIIKDKDFLDNEENEDILEDIV
LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
KTILDFLKSDGFANRNF MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA IK
KGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKEL
GSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYD VDHIVPQSFL
KDDSIDNKVLTRSDKNRGKSDNPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK
AERGGSEL DKA GFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVY
DVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
KGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWD PKKY
GGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERS SFEKNPIDFLEAKGYKE
VKKDLI IKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEK LK
GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKR VILADANLDKVL SAYNKH RDKPIRE
QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDL SQ
LGGDSGGSPKKKRKV (SEQ ID NO: 155)

[0304] ecTadA(D108V)-XTEN-nCas9-NLS: (mammalian construct, active on DNA, A to G editing)

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARVAKT
GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTD
SGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI
KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSSFFHR
LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLAL
AHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARL
SKSRRLLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDD
LDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL
TLLKALVRQQLP EKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL
VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIP
YYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNE

KVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV
 KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV
 LTLTLFEDREMIEERLKTYYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
 KTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIK
 KGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGKEL
 GSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFL
 KDDSIDNKVLTRSDKNRGKSDNPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK
 AERGGELSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 KLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVY
 DVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
 KGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY
 GGFDSP TVAYSVLVVAKEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKE
 VKKDLIIKLPKYSLEFENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLG
 GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLSAYNKHDKPIRE
 QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI TGLYETRIDL SQ
 LGGDSGGSPKKKRKV (SEQ ID NO: 156)

[0305] ecTadA(D108N)-XTEN-nCas9-UGI-NLS (BE3 analog of A to G editor)

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAAQKKAQSSD
 SGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSI
 KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADRLIYLAL
 AHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARL
 SKSRRLLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDD
 LDNLLAQIGDQYADLFLAAKNLSDAILSDILRVNTEITKAPLSASMIKRYDEHHQDL
 TLLKALVRQQLPKEYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL
 VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIP
 YYVGPLARGNSRFWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNE
 KVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV
 KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV
 LTLTLFEDREMIEERLKTYYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
 KTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIK
 KGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGKEL
 GSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFL
 KDDSIDNKVLTRSDKNRGKSDNPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK
 AERGGELSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 KLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVY
 DVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
 KGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY
 GGFDSP TVAYSVLVVAKEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKE
 VKKDLIIKLPKYSLEFENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLG
 GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLSAYNKHDKPIRE
 QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI TGLYETRIDL SQ
 LGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDE
 NVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV (SEQ ID NO: 157)

[0306] ecTadA(D108G)-XTEN-nCas9-UGI-NLS (BE3 analog of A to G editor)

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARGAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD
 SGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSI
 KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLAL
 AHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARL
 SKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKNFDLAEDAKLQLSKD TYDDD
 LDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL
 TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL
 VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIP
 YYVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDKNLPNE
 KVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV
 KQLKEDYFKKIECFDSVEISGVEDRFNASLGT YHDLLKIIKDKDFLDNEENEDILEDIV
 LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
 KTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA IK
 KGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEG IKEL
 GSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYD VDHIVPQSFL
 KDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK
 AERGGSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVY
 DVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
 KGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWD PKKY
 GGFDSP TVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSSF EKNPIDFLEAKGYKE
 VKKDLIIKL PKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEK LK
 GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIRE
 QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDL SQ
 LGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDE
 NVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKRKV (SEQ ID NO: 158)

[0307] ecTadA(D108V)-XTEN-nCas9-UGI-NLS (BE3 analog of A to G editor)

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARVAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD
 SGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSI
 KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLAL
 AHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARL
 SKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKNFDLAEDAKLQLSKD TYDDD
 LDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL
 TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL
 VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIP
 YYVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDKNLPNE
 KVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV
 KQLKEDYFKKIECFDSVEISGVEDRFNASLGT YHDLLKIIKDKDFLDNEENEDILEDIV
 LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
 KTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA IK
 KGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEG IKEL
 GSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYD VDHIVPQSFL

KDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK
 AERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVY
 DVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
 KGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY
 GGFDSP TVAYSVLVVAKEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKE
 VKKDLIIKLPKYSLEFENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLG
 GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLSAYNKHRDKPIRE
 QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI TGLYETRIDL SQ
 LGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDE
 NVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV (SEQ ID NO: 159)

[0308] ecTadA(D108N)-XTEN-dCas9-UGI-NLS (mammalian cells, BE2 analog of A to G editor)

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAAQKKAQSSD
 SGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSI
 KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADRLRIYLAL
 AHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARL
 SKSRRLLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDD
 LDNLLAQIGDQYADLFLAAKNLS DAILSDILRVNTEITKAPLSASMIKRYDEHHQDL
 TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL
 VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIP
 YYVGPLARGNSRFWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNE
 KVLPHKSLLEYEFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV
 KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV
 LTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
 KTILDFLKSDGFANRNF MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA
 IKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE EGKEL
 GSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFL
 KDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK
 AERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVY
 DVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
 KGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY
 GGFDSP TVAYSVLVVAKEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKE
 VKKDLIIKLPKYSLEFENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLG
 GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLSAYNKHRDKPIRE
 QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI TGLYETRIDL SQ
 LGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDE
 NVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV (SEQ ID NO: 160)

[0309] ecTadA(D108G)-XTEN-dCas9-UGI-NLS (mammalian cells, BE2 analog of A to G editor)

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARGAKT

GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD
 SGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI
 KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADRLIYLAL
 AHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARL
 SKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKS NFDAEDAKLQLSKD TYDDD
 LDNLLAQIGDQYADLFLAAKNLSDAILSDILRVNTEITKAPLSASMIKRYDEHHQDL
 TLLKALVRQQLPKEYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL
 VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIP
 YYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNE
 KVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV
 KQLKEDYFKKIECFDSVEISGVEDRFNASLGT YHDLLKIIKDKDFLDNEENEDILEDIV
 LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
 KTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIK
 KGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE EGIKEL
 GSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFL
 KDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK
 AERGGSELDDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 KLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVY
 DVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
 KGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY
 GGFDSP TVAYSVLVVAKEKGKSKKLKSVKELLGITIMERS SFEKNPIDFLEAKGYKE
 VKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEK LK
 GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLSAYNKH RDKPIRE
 QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQ
 LGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDE
 NVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV (SEQ ID NO: 161)

[0310] ecTadA(D108V)-XTEN-dCas9-UGI-NLS (mammalian cells, BE2 analog of A to G editor)

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARVAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD
 SGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI
 KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADRLIYLAL
 AHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARL
 SKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKS NFDAEDAKLQLSKD TYDDD
 LDNLLAQIGDQYADLFLAAKNLSDAILSDILRVNTEITKAPLSASMIKRYDEHHQDL
 TLLKALVRQQLPKEYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL
 VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTFRIP
 YYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNE
 KVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV
 KQLKEDYFKKIECFDSVEISGVEDRFNASLGT YHDLLKIIKDKDFLDNEENEDILEDIV
 LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
 KTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIK
 KGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE EGIKEL
 GSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFL
 KDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK

AERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVY
 DVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
 KGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY
 GGFDSP TVAYSVLVVAKEVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKE
 VKKDLIIKLPKYSLEFENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLK
 GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLSAYNKHRRDKPIRE
 QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDL SQ
 LGGDSGGSTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDE
 NVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGSPKKKRKV (SEQ ID NO: 162)

[0311] ecTadA(D108N)-XTEN-nCas9-AAG(E125Q)-NLS – cat. alkyladenosine glycosylase

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTD
 SGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI
 KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADRLIYLAL
 AHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARL
 SKSRRLLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDD
 LDNLLAQIGDQYADLFLAAKNLS DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL
 TLLKALVRQQLPKEYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL
 VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIP
 YYVGPLARGNSRFAMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNE
 KVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV
 KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV
 LTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
 KTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIK
 KGILQTVKVVDELVKVMGRHKPENIVIMARENQTTQKGQKNSRERMKRIE EGikel
 GSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFL
 KDDSIDNKVLTRSDKNRGKSDNPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK
 AERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVY
 DVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
 KGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY
 GGFDSP TVAYSVLVVAKEVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKE
 VKKDLIIKLPKYSLEFENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLK
 GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLSAYNKHRRDKPIRE
 QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDL SQ
 LGGDSGGSKGHLTRLGLEFFDQPAVPLARAFLGQVLVRRLPNGTELGRIVETQAYL
 GPEDEAAHSRGGRQTPRNRGMFMKPGTLYVYIYGYFCMNISSQGDGACVLLRAL
 EPLEGLETMRQLRSTLRKGTASRVLKDRELCSGPSKLCQALAINKSFDQRDLAQDEA
 VWLERGPLEPSEPAVVAARVGVGHAGEWARKPLRFYVRGSPWVSVVDRVAEQDT
 QASGGSPKKKRKV (SEQ ID NO: 163)

[0312] ecTadA(D108G)-XTEN-nCas9-AAG(E125Q)-NLS – cat. alkyladenosine glycosylase

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARGAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTD

SGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI
 KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL
 AHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARL
 SKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDD
 LDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL
 TLLKALVRQQLPKEYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL
 VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIP
 YYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNE
 KVLPHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV
 KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV
 LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
 KTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIK
 KGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEIGIKEL
 GSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFL
 KDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK
 AERGGSELDDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVY
 DVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
 KGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY
 GGFDSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKE
 VKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLG
 GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHDKPIRE
 QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQSTGLYETRIDLQ
 LGGDSGGSKGHLTRLGLEFFDQPAVPLARAFLGQVLVRRLPNGTELGRIVETQAYL
 GPEDEAAHSRGGRTPRNRGMFMKPGTLYVYIYGYFCMNISSQGDGACVLLRAL
 EPLEGLETMRQLRSTLRKGTASRVLDRELCSGPSKLCQALAINKSFDQRDLAQDEA
 VWLERGPLEPSEPAVVAARVGVGHAGEWARKPLRFYVRGSPWVSVDRAEQDT
 QASGGSPKKRKV (SEQ ID NO: 164)

[0313] ecTadA(D108V)-XTEN-nCas9-AAG(E125Q)-NLS – cat. alkyladenosine glycosylase

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARVAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAAQKKAQSSD
 SGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI
 KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL
 AHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARL
 SKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDD
 LDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL
 TLLKALVRQQLPKEYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL
 VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIP
 YYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNE
 KVLPHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV
 KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV
 LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
 KTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIK
 KGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEIGIKEL
 GSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFL

KDDSIDNKVLTRSDKNRGKSDNPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK
 AERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVY
 DVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
 KGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY
 GGFDSPVAVSVLVVAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKE
 VKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLG
 GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLSAYNKHDKPIRE
 QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQSIITGLYETRIDLSQ
 LGGDSGGSGKHLTRLGLEFFDQPAVPLARAFLGQVLVRRLPNGTELGRIVETQAYL
 GPEDEAAHSRGGRTPRNRGMFMKPGTLYVYIYGYFCMNISSQGDGACVLLRAL
 EPLEGLETMRQLRSTLRKGTASRVLKDRELCSGPSKLCQALAINKSFDQRDLAQDEA
 VWLERGPLEPSEPAVVAAARVGVGHAGEWARKPLRFYVRGSPWVSVVDRVAEQDT
 QASGGSPKKRKV (SEQ ID NO: 165)

[0314] ecTadA(D108N)-XTEN-nCas9-EndoV(D35A)-NLS: contains cat. endonuclease V

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAAQKKAQSSD
 SGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSI
 KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEAKVDDSSFFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDDSTDKADRLIYLAL
 AHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARL
 SKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDD
 LDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL
 TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL
 VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIP
 YYVGPLARGNSRFAMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNE
 KVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV
 KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV
 LTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
 KTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAK
 KGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEIEGKEL
 GSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFL
 KDDSIDNKVLTRSDKNRGKSDNPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK
 AERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVY
 DVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
 KGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY
 GGFDSPVAVSVLVVAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKE
 VKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLG
 GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLSAYNKHDKPIRE
 QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQSIITGLYETRIDLSQ
 LGGDSGGSDLASLRAQQIELASSVIREDRDKDPPDLIAGAAVGFEQGGEVTRAAMV
 LLKYPSELVEYKVARIATTMPYIPGFLSFREYPALLAAWEMLSQKPDLVFVDGHGIS
 HPRRLGVASHFGLLVDPVTIGVAKKRLCGKFEPLSSEPGALAPLMDKGEQLAWVWR
 SKARCNPLFIATGHRVSVDSALAWVQRCMKGYRLPEPTRWADAVASERPAFVRYT
 ANQPSGGSPKKRKV (SEQ ID NO: 166)

[0315] ecTadA(D108G)-XTEN-nCas9-EndoV (D35A)-NLS: contains cat. endonuclease V

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARGAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAAQKKAQSSTD
 SGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI
 KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYAL
 AHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARL
 SKSRRLLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDD
 LDNLLAQIGDQYADLFLAAKNLSDAILSDILRVNTEITKAPLSASMIKRYDEHHQDL
 TLLKALVRQQLPKEYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL
 VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIP
 YYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNE
 KVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV
 KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV
 LTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
 KTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIK
 KGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEIEGKEL
 GSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFL
 KDDSIDNKVLTRSDKNRGKSDNPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK
 AERGGSELDDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVY
 DVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
 KGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY
 GGFDSPTVAYSVLVVAKEKKGSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKE
 VKKDLIIKLPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFYLYLASHYEKLEK
 GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLSAYNKHRRDKPIRE
 QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQSIITGLYETRIDLQ
 LGGDSGGSDLASLRAQQIELASSVIREDRDKDPPDLIAGAAVGFEQGGGEVTRAAMV
 LLKYPSELVEYKVARIATTMPYIPGFLSFREYPALLAAWEMLSQKPDLVFVDGHGIS
 HPRRLGVASHFGLLDVPTIGVAKKRLCGKFEPLSSEPGALAPLMDKGEQLAWVWR
 SKARCNPLFIATGHRVSVDSALAWVQRCMKGYRLPEPTRWADAVASERPAFVRYT
 ANQPSGGSPKKKRKV (SEQ ID NO: 167)

[0316] ecTadA(D108V)-XTEN-nCas9-EndoV(D35A)-NLS: contains cat. endonuclease V

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARVAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAAQKKAQSSTD
 SGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI
 KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYAL
 AHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARL
 SKSRRLLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDD
 LDNLLAQIGDQYADLFLAAKNLSDAILSDILRVNTEITKAPLSASMIKRYDEHHQDL
 TLLKALVRQQLPKEYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL
 VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIP
 YYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNE
 KVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV
 KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV

LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
 KTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAICK
 KGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKEL
 GSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFL
 KDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK
 AERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVY
 DVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
 KGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY
 GGFDSPVAVSVLVVAKVEKGKSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKE
 VKKDLIIKLPKYSLEFENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLK
 GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIRE
 QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQ
 LGGDSGGSDLASLRAQQIELASSVIREDRDKDPDLIAGAAVGFEQGGEVTRAAMV
 LLKYPSELEVEYKVARIAATMPYIPGFLSFREYPALLAAWEMLSQKPD LVFDGHGIS
 HPRRLGVASHFGLLVDPVTIGVAKKRLCGKFEPLSSEPGALAPLMDKGEQLAWVWR
 SKARCNPLFIATGHRVSVDSALAWVQRCMKGYRLPEPTRWADAVASERPAFVRYT
 ANQPSGGSPKKRKV (SEQ ID NO: 168)

[0317] Variant resulting from first round of evolution (in bacteria)

ecTadA(H8Y_D108N_N127S)-XTEN-dCas9

MSEVEFSYEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKT
 GAAGSLMDVLHHPGMSHRVEITEGILADECAALLSDFFRMRRQEIKAKKKAQSSTDS
 GSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIK
 KNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFHRL E
 ESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLA LAH
 MIKFRGHFLIEGDLNPDNSDV DKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK
 SRLENLIAQLPGEKKNGLFGNLI ALSGLTPNFKS NFDLAEDAKLQLSKD TYDDDL
 DNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLT
 LLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV
 KLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPY
 YVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEK
 VLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVK
 QLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVL
 TLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGK
 TILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAICK
 GILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELG
 SQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYD VDAIVPQSFLK
 DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKA
 ERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 LVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYD
 VRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDK
 GRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYG
 GFDSPTVAVSVLVVAKVEKGKSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEV
 KKDIIKLPKYSLEFENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLK
 SPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQ
 AENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQL
 GGD (SEQ ID NO: 169)

[0318] Enriched variants from second round of evolution (in bacteria) ecTadA

(H8Y_D108N_N127S_E155X)-XTEN-dCas9; X=D, G or V

MSEVEFSYEWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKT
 GAAGSLMDVLHHPGMSHRVEITEGILADECAALLSDFFRMRRQXIKAKKKAQSSTD
 SGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI
 KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL
 AHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARL
 SKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDD
 LDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL
 TLLKALVRQQLPKEYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL
 VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIP
 YYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNE
 KVLPHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV
 KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIV
 LTLTLFEDREMIEERLKTIAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
 KTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIK
 KGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEIGIKEL
 GSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFL
 KDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK
 AERGGSELDDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 KLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVY
 DVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
 KGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY
 GGFDSPTVAYSVLVVAKEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKE
 VKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFYLYLASHYEKLLK
 GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHDKPIRE
 QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSTGLYETRIDLSQL
 LGGD (SEQ ID NO: 170)

[0319] pNMG-160: ecTadA(D108N)-XTEN-nCas9-GGS-AAG*(E125Q)-GGS-NLS

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKKAQSSTD
 SGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI
 KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLAL
 AHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARL
 SKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDD
 LDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL
 TLLKALVRQQLPKEYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL
 VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIP
 YYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNE
 KVLPHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV
 KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIV
 LTLTLFEDREMIEERLKTIAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG

KTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAK
 KGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKEL
 GSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFL
 KDDSIDNKVLTRSDKNRGKSDNPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK
 AERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 KLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVY
 DVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
 KGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY
 GGFDSPVAVSVLVVAKVEKGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKE
 VKKDLIIKLPKYSLEFENGKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLK
 GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLSAYNKHDKPIRE
 QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQSIITGLYETRIDLQS
 LGGDGGSGKHLTRLGLEFFDQPAVPLARAFGLQVLRRLPNGTELGRIVETQAYLG
 PEDEAAHSRGGRTPRNRGMFMKPGTLYVYIIYGYFCMNISSQGDGACVLLRALE
 PLEGLTMRQLRSTLRKGTASRVLKDRELCSGPSKLCQALAINKSFDQRDLAQDEAV
 WLERGPLEPSEPAVVAAARVGVGHAGEWARKPLRFYVRGSPWVSVVDRVAEQDTQ
 AGGSPKKKRKV (SEQ ID NO: 171)

[0320] pNMG-161: ecTadA(D108N)-XTEN-nCas9-GGS-EndoV*(D35A)-GGS-NLS

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAAQKKAQSSD
 SGSETPGTSESATPESDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI
 KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHR
 LEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDDSTDKADRLIYLAL
 AHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARL
 SKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDD
 LDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL
 TLLKALVRQQLPKEYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL
 VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIP
 YYVGPLARGNSRFAMTRKSEETITPWNFEVVDKGASQSFIERMTNFDKNLPNE
 KVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV
 KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV
 LTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSG
 KTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAK
 KGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKEL
 GSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFL
 KDDSIDNKVLTRSDKNRGKSDNPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK
 AERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 KLVSDFRKDFQFYKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVY
 DVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
 KGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY
 GGFDSPVAVSVLVVAKVEKGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKE
 VKKDLIIKLPKYSLEFENGKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLK
 GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLSAYNKHDKPIRE
 QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQSIITGLYETRIDLQS
 LGGDGGSDLASLRAQQIELASSVIREDRDKDPPDLIAGAAVGFEQGGEVTRAAMVL
 LKYPSELEVEYKVARIATTMPYIPGFLSFREYPALLAAWEMLSQKPDLVFVDGHGIS
 HPRRLGVASHFGLLDVPTIGVAKKRLCGKFEPLSSEPGALAPLMDKGEQLAWVWR

SKARCNPLFIATGHRVSVDSALAWVQRCMKGYRLPEPTRWADAVASERPAFVRYT
ANQPGGSPKKRKV (SEQ ID NO: 172)

[0321] pNMG-371: ecTadA(L84F_A106V_D108N_H123Y_D147Y_E155V_I156F)-SGGS-
SGGS-XTEN-SGGS-SGGS-

ecTadA(L84F_A106V_D108N_H123Y_D147Y_E155V_I156F)-SGGS-SGGS-XTEN-
SGGS-SGGS-nCas9-SGGS- NLS

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTDS
GGSSGGSSGSETPGTSESATPESSGGSSGGSSSEVEFSHEYWMRHALTLAKRAWDERE
VPVGAVLVHNNRVIGEGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVT
FEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILAD
ECAALLSYFFRMRRQVFKAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSG
GSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHE
RHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRILIYLAHMIKFRGHFLIEG
DLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLDNLLAQIGDQYA
DLFLAAKNLSDAILSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQ
RTFDNGSIPHQIHLGELHAILRRQEDFYPLKDNREKIEKILTFRIPYYVGPLARGNSRF
AWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFT
VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECF
DSVEISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEE
RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN
RNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAICKGILQTVKVVDEL
VKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE EGIKELGSQILKEHPVENT
QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYVDVHVPQSFLKDDSIDNKVLTRS
DKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAG
FIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFY
KVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEI
GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL
SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL
VVAKEVGKSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKKDLIILPKYS
LFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGS PEDNEQKQLF
VEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTN
LGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDSGGSPKK
KRKV (SEQ ID NO: 173)

[0322] pNMG-616 amino acid sequence: ecTadA(wild type)-(SGGS)2-XTEN-(SGGS)2-
ecTadA(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P
_E155V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD

SGGSSGGSSGSETPGTSESATPESSGGSSGSSGSEVEFSHEYWMRHALTLAKRALDERE
 VPVGAVLVNLRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVT
 FEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILAD
 ECAALLCYFFRMPRQVFNAQKKAQSSTDSSGGSSGGSSGSETPGTSESATPESSGGSSG
 GSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
 TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLVESFLVEEDKKHE
 RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDDSTDKADLRILIYLAHMIKFRGHFLIEG
 DLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLP
 GEKKNGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQYA
 DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
 KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQ
 RTFDNGSIPHQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTRIPYYVGPLARGNSRF
 AWMTRKSEETITPWNFEVVDKGAQAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFT
 VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECF
 DSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEE
 RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN
 RNFMQLIHDDSLTFKEDIQKAQVSGQGDLSHEHIANLAGSPAIIKKGILQTVKVVDEL
 VKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENT
 QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRS
 DKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAG
 FIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFY
 KVINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEI
 GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL
 SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL
 VVAKVEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIILPKYS
 LFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLF
 VEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHARDKPIREQAENIIHLFTLTN
 LGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQSIITGLYETRIDLSQLGGDSGGSPKK
 KRKV (SEQ ID NO: 174)

[0323] pNMG-624 amino acid sequence: ecTadA(wild type)-32 a.a. linker-

ecTadA(W23R_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P
 _E155V_I156F_K157N)-24 a.a. linker_nCas9_SGGS_NLS

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTD
 SGGSSGGSSGSETPGTSESATPESSGGSSGSSGSEVEFSHEYWMRHALTLAKRARDERE
 VPVGAVLVNLRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVT
 FEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILAD
 ECAALLCYFFRMPRQVFNAQKKAQSSTDSSGGSSGGSSGSETPGTSESATPESDKKYSI
 GLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRL
 KRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLVESFLVEEDKKHERHPIFGNI
 VDEVAYHEKYPTIYHLRKKLVDDSTDKADLRILIYLAHMIKFRGHFLIEGDLNPDNS
 DVDKLFILQVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGL
 FGNLIASLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAK
 NLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPKEYKEIFFD
 QSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIP
 HQIHLGELHAILRRQEDFYFPFLKDNREKIEKILTRIPYYVGPLARGNSRFAWMTRKS

EETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTK
VKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGV
EDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHL
FDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIH
DDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDDELVKVMGRH
KPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLY
LYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSD
NVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVET
RQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINNY
HHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYF
FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIV
KKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEK
GKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGS PEDNEQKQLFVEQHKHYL
DEIIEQISEFSKRVILADANLDKVL SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFK
YFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGDSGGSPKKKRKV (SEQ
ID NO: 175)

[0324] pNMG-476 amino acid sequence (evolution #3 hetero dimer, wt TadA + TadA evo #3
mutations): ecTadA(wild-type)-(SGGS)2-XTEN-(SGGS)2-
ecTadA(L84F_A106V_D108N_H123Y_D147Y_E155V_I156F)-(SGGS)2-XTEN-
(SGGS)2_nCas9_SGGS_NLS

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAAQKKAQSSD
SGGSSGGSSGSETPGTSESATPESGGSSGSSGSEVEFSHEYWMRHALTLAKRAWDER
EVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYV
TFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHHPGMNHRVEITEGILA
DECAALLSYFFRMRRQVFKAQKKAQSSDSSGGSSGGSSGSETPGTSESATPESGGSS
GGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS
GETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFHRL EESFLVEEDKK
HERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLALAHMIKFRGHFLI
EGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQ
LPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLDNLLAQIGDQ
YADLFLAAKNLSDAILSDILRVNTEITKAPLSASMIKRYDEHHQD LTLKALVRQQL
PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRK
QRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPYYVGPLARGNS
RFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY
FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIE
CFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIE
ERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFA
NRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVD
LVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVEN
TQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTR
SDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKA
GFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQF
YKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRK

VLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS
VLVVAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPK
YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ
LFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLFTL
TNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDSGGSP
KKKRKV (SEQ ID NO: 176)

[0325] pNMG-477 amino acid sequence: ecTadA(wild-type)-(SGGS)2-XTEN-(SGGS)2-
ecTadA(H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F
_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTD
SGGSSGGSSGSETPGTSESATPESSGGSSGGSSSEVEFSHEYWMRHALTLAKRAWDER
EVPVGAVLVHNNRVIGEGWNRPIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYV
TFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILA
DECAALLCYFFRMRRQVFNAQKKAQSSTD SGGSSGGSSGSETPGTSESATPESSGGSS
GGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS
GETAEATRLKRTARRRYTRRKNRICYLQEFSNEMAKVDDSFHRLEESFLVEEDKK
HERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHAMIKFRGHFLI
EGDLNPDNSDV DKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQ
LPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQ
YADLFLAAKNLS DAILSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQL
PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIPILEKMDGTEELLVKLNREDLLRK
QRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYVVGPLARGNS
RFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY
FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIE
CFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIE
ERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFA
NRNFMQLIHDDSLTFKEDIQKAQVSGQGD SLHEHIANLAGSPA I KKGILQTVKV VDE
LVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE EG I KELGSQILKEHPVEN
TQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTR
SDKNRGKSDNPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL DKA
GFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQF
YKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRK
VLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS
VLVVAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLPK
YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ
LFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLFTL
TNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDSGGSP
KKKRKV (SEQ ID NO: 177)

[0326] pNMG-558 amino acid sequence: ecTadA(wild-type)- 32 a.a. linker-
ecTadA(H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F
_K157N)- 24 a.a. linker_nCas9_SGGS_NLS

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTD
 SGGSSGGSSGSETPGTSESATPESSGGSSGSSGSEVEFSHEYWMRHALTLAKRAWDER
 EVPVGAVLVLNNRVIGEGWNRPIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYV
 TFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILA
 DECAALLCYFFRMRRQVFNAQKKAQSSTDSSGGSSGGSSGSETPGTSESATPESSDKKY
 SIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATR
 LKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGN
 IVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNS
 DVDKLFQILVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGL
 FGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAK
 NLSDAILSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFD
 QSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIP
 HQIHLGELHAILRRQEDFYFPLKDNREKIEKILTRIPYYVGPLARGNSRFAWMTRKS
 EETITPWNFEVVDKGSASQSFIERMTNFDKNLPNEKVLPHSLLYEYFTVYNELTK
 VKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGV
 EDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHL
 FDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLQIH
 DDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDLVKVMGRH
 KPENIVIEMARENQTTQKGQKNSRERMKRIEIEGKELGSQILKEHPVENTQLQNEKLY
 LYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSD
 NVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDAKAGFIKRQLVET
 RQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNY
 HHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYF
 FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNV
 KKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAVKVEK
 GKSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKKDLIIKLPKYSLELENGR
 KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYL
 DEIIIEQISEFSKRVILADANLDKVLSAYNKHDKPIREQAENIIHLFTLTNLGAPAAFK
 YFDTTIDRKRYTSTKEVLDATLIHQSTGLYETRIDLSQLGGDSGGSPKKKRKV (SEQ
 ID NO: 178)

[0327] pNMG-576 amino acid sequence: ecTadA(wild-type)-(SGGS)2-XTEN-(SGGS)2-
 ecTadA(H36L_P48S_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F
 _K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_GGS_NLS

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTD
 SGGSSGGSSGSETPGTSESATPESSGGSSGSSGSEVEFSHEYWMRHALTLAKRAWDER
 EVPVGAVLVLNNRVIGEGWNRPIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYV
 TFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILA
 DECAALLCYFFRMRRQVFNAQKKAQSSTDSSGGSSGGSSGSETPGTSESATPESSGGSS
 GGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS
 GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKK
 HERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLI
 EGDLPDNDSDVDKLFQILVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQ
 LPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQ

YADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQL
 PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRK
 QRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTRIPYYVGPLARGNS
 RFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY
 FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIE
 CFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIE
 ERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFA
 NRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDE
 LVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEELGKELGSQILKEHPVEN
 TQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTR
 SDKNRGKSDNPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTAKAERGGLSELDKA
 GFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQF
 YKVVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
 EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRK
 VLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS
 VLVVAKVEKGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPK
 YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ
 LFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHDKPIREQAENIIHLFTL
 TNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLHQSIITGLYETRIDLSQLGGDSGGSP
 KKKRKV (SEQ ID NO: 179)

[0328] pNMG-577 amino acid sequence: ecTadA(wild-type)-(SGGS)2-XTEN-(SGGS)2-
 ecTadA(H36L_P48S_R51L_L84F_A106V_D108N_H123Y_S146C_A142N_D147Y_E155
 V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_GGS_NLS

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAAQKKAQSSTD
 SGGSSGGSSGSETPGTSESATPESSGGSSGGSSSEVEFSHEYWMRHALTLAKRAWDER
 EVPVGAVLVHNNRVIGEGWNRSIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYV
 TFEPVCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILA
 DECNALLCYFFRMRRQVFNAQKKAQSSTDSSGGSSGGSSGSETPGTSESATPESSGGSS
 GGSDDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS
 GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKK
 HERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLI
 EGDLPDNDSDVDKLFQQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQ
 LPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQ
 YADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQL
 PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRK
 QRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTRIPYYVGPLARGNS
 RFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY
 FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIE
 CFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIE
 ERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFA
 NRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDE
 LVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEELGKELGSQILKEHPVEN
 TQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTR
 SDKNRGKSDNPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTAKAERGGLSELDKA
 GFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQF

YKVVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRK
VLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS
VLVVAKEVGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPK
YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ
LFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLFTL
TNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDSGGSP
KKKRKV (SEQ ID NO: 180)

[0329] pNMG-586 amino acid sequence: ecTadA(wild-type)-(SGGS)2-XTEN-(SGGS)2-
ecTadA(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F
_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_GGS_NLS

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAAQKKAQSSTD
SGGSSGGSSGSETPGTSESATPESSGGSSGGSSSEVEFSHEYWMRHALTLAKRAWDER
EVPVGAVLVHNNRVIGEGWNRRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYV
TFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHHPGMNHRVEITEGIL
DECAALLCYFFRMRRQVFNAQKKAQSSTDSSGGSSGGSSGSETPGTSESATPESSGGSS
GGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS
GETAEATRLKRTARRRYTRRKNRICYLQEISNEMAKVDDSFHRLSEESFLVEEDKK
HERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHAMIKFRGHFLI
EGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQ
LPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLDNLLAQIGDQ
YADLFLAAKNLS DAILSDILRVNTEITKAPLSASMIKRYDEHHQD TLTKALVRQQ
PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRK
QRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNS
RFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY
FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIE
CFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIE
ERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFGA
NRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDE
LVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE EGIKELGSQILKEHPVEN
TQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVHDHVPQSFLKDDSIDNKVLTR
SDKNRGKSDNVPSEEVVKMKKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL DKA
GFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQF
YKVVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRK
VLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS
VLVVAKEVGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPK
YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ
LFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLFTL
TNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDSGGSP
KKKRKV (SEQ ID NO: 181)

[0330] pNMG-588 amino acid sequence: ecTadA(wild-type)-(SGGS)2-XTEN-(SGGS)2-ecTadA(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_A142N_D147Y_E155V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_GGS_NLS

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTD
 SGGSSGGSSGSETPGTSESATPESSGGSSGGSSSEVEFSHEYWMRHALTLAKRAWDER
 EVPVGAVLVHNNRVIGEGWNRRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYV
 TFPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILA
 DECNALLCYFFRMRRQVFNAQKKAQSSTDSSGGSSGGSSGSETPGTSESATPESSGGSS
 GSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS
 GETAEATRLKRTARRRYTRRKNRICYLQEFSNEMAKVDDSFHRLEESFLVEEDKK
 HERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHAMIKFRGHFLI
 EGDLPDNDSDVDKLFQILVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQ
 LPGEKKNGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQ
 YADFLAAKNLSAAILSDILRVNTEITKAPLSASMIKRYDEHHQDILTLLKALVRQQ
 PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRK
 QRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYYVGPLARGNS
 RFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY
 FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIE
 CFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIE
 ERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFA
 NRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDE
 LVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEELGELGSQILKEHPVEN
 TQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTR
 SDKNRGKSDNPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTAKAERGGLSELDA
 GFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQF
 YKVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
 EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRK
 VLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS
 VLVVAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDLIIKLKPK
 YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ
 LFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHARDKPIREQAENIIHLFTL
 TNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQSTGLYETRIDLSQLGGDSGGSP
 KKKRKV (SEQ ID NO: 182)

[0331] pNMG-620 amino acid sequence: ecTadA(wild-type)-(SGGS)2-XTEN-(SGGS)2-ecTadA(W23R_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_GGS_NLS

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTD
 SGGSSGGSSGSETPGTSESATPESSGGSSGGSSSEVEFSHEYWMRHALTLAKRARDERE
 VPVGAVLVHNNRVIGEGWNRRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVT
 FEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILAD
 ECAALLCYFFRMPRQVFNAQKKAQSSTDSSGGSSGGSSGSETPGTSESATPESSGGSSG
 GSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE

TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHE
 RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHAHMIKFRGHFLIEG
 DLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLP
 GEKKNGLFGNLIASLGLTPNFKSNDLAEDAQLQLSKDQYDDDLNLLAQIGDQYA
 DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQPE
 KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQ
 RTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRF
 AWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSLLYEYFT
 VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECF
 DSVEISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEE
 RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN
 RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDEL
 VKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGKELGSQILKEHPVENT
 QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYVDHIVPQSFLKDDSIDNKVLTRS
 DKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAG
 FIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFY
 KVINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKAKSEQEI
 GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL
 SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL
 VVAKVEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKDLIKLPKYS
 LFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLF
 VEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHARDKPIREQAENIIHLFTLTN
 LGAPAAFKYFDTTIDRKRYTSTKEVLDTLHQSITGLYETRIDLSQLGGDSGGSPKK
 KRKV (SEQ ID NO: 183)

[0332] pNMG-617 amino acid sequence: ecTadA(wild-type)-(SGGS)2-XTEN-(SGGS)2-
 ecTadA(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142A_S146C_D147Y
 _E155V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_GGS_NLS

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAAQKKAQSSTD
 SGGSSGGSSGSETPGTSESATPESSGGSSGGSSSEVEFSHEYWMRHALTLAKRALDERE
 VPVGAVLVHNNRVIGEGWNRVIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVT
 FEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILAD
 ECNALLCYFFRMRRQVFNAQKKAQSSTDSSGGSSGGSSGSETPGTSESATPESSGGSSG
 GSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
 TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHE
 RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHAHMIKFRGHFLIEG
 DLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLP
 GEKKNGLFGNLIASLGLTPNFKSNDLAEDAQLQLSKDQYDDDLNLLAQIGDQYA
 DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQPE
 KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQ
 RTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRF
 AWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSLLYEYFT
 VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECF
 DSVEISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEE
 RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN
 RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDEL

VKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENT
 QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRS
 DKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAG
 FIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFY
 KVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEI
 GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL
 SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL
 VVAKVEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYS
 LFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLF
 VEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHARDKPIREQAENIIHLFTLTN
 LGAPAAFKYFDTTIDRKRYTSTKEVLDTLHQSIITGLYETRIDLSQLGGDSGGSPKK
 KRKV (SEQ ID NO: 184)

[0333] pNMG-618 amino acid sequence: ecTadA(wild-type)-(SGGS)2-XTEN-(SGGS)2-
 ecTadA(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142A_S146C_D147Y
 _R152P_E155V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_GGS_NLS

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFRMRREQEIKAKKAQSSTD
 SGGSSGGSSGSETPGTSESATPESGGSSGGSSSEVEFSHEYWMRHALTLAKRALDERE
 VPVGAVLVHNNRVIGEGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVT
 FEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILAD
 ECNALLCYFFRMPRQVFNAQKKAQSSTDSSGGSSGGSSGSETPGTSESATPESGGSSG
 GSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
 TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHE
 RHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRILIYLAHMIKFRGHFLIEG
 DLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLP
 GEKKNGLFNGNLIASLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQYA
 DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
 KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQ
 RTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRF
 AWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFT
 VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECF
 DSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEE
 RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN
 RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDEL
 VKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENT
 QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRS
 DKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAG
 FIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFY
 KVREINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEI
 GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL
 SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL
 VVAKVEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYS
 LFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLF
 VEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHARDKPIREQAENIIHLFTLTN
 LGAPAAFKYFDTTIDRKRYTSTKEVLDTLHQSIITGLYETRIDLSQLGGDSGGSPKK
 KRKV (SEQ ID NO: 185)

[0334] pNMG-620 amino acid sequence: ecTadA(wild-type)-(SGGS)2-XTEN-(SGGS)2-ecTadA(W23R_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_GGS_NLS

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAAQKKAQSSTD
 SGGSSGGSSGSETPGTSESATPESSGGSSGGSSSEVEFSHEYWMRHALTLAKRAWDERE
 VPGAVLVHNNRVIGEGWNRRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVT
 FEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILAD
 ECAALLCYFFRMMPRQVFNAQKKAQSSTDSSGGSSGGSSGSETPGTSESATPESSGGSSG
 GSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
 TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHE
 RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYALAHMIKFRGHFLIEG
 DLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLP
 GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYA
 DLFLAAKNLSDAILSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQPE
 KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQ
 RTFDNGSIPHQIHLGELHAILRRQEDFYPLKDNREKIEKILTFRIPYYVGPLARGNSRF
 AWMTRKSEETITPWNFEVVDKGAQAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFT
 VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECF
 DSVEISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEE
 RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN
 RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDEL
 VKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGKELGSQILKEHPVENT
 QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYVDHVPQSFLKDDSIDNKVLTRS
 DKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDDKAG
 FIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFY
 KVINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEI
 GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL
 SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL
 VVAKVEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIILPKYS
 LFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLF
 VEQHKHYLDEIIEQISEFSKRVLADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN
 LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSIITGLYETRIDLSQLGGDSGGSPKK
 KRKV (SEQ ID NO: 183)

[0335] pNMG-621 amino acid sequence: ecTadA(wild-type)- 32 a.a. linker-ecTadA(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F_K157N)- 24 a.a. linker_nCas9_GGS_NLS

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAAQKKAQSSTD
 SGGSSGGSSGSETPGTSESATPESSGGSSGGSSSEVEFSHEYWMRHALTLAKRAWDER
 EVPVGAVLVHNNRVIGEGWNRRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYV
 TFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILA
 DECAALLCYFFRMMPRQVFNAQKKAQSSTDSSGGSSGGSSGSETPGTSESATPESDKKY

SIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATR
 LKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGN
 IVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNS
 DVDKLFQILVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGL
 FGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAK
 NLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFD
 QSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIP
 HQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKS
 EETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHSLLYEYFTVYNELTK
 VKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGV
 EDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHL
 FDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLQIH
 DDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRH
 KPENIVIEMARENQTTQKGQKNSRERMKRIEIGIKELGSQILKEHPVENTQLQNEKLY
 LYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSD
 NVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTAKAERGGLSELDKAGFIKRQLVET
 RQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNY
 HHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYF
 FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNV
 KKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEK
 GKSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKKDLIILPKYSLFELENGR
 KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYL
 DEIIQISEFSKRVLADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFK
 YFDTTIDRKRYTSTKEVLDTLHQITGLYETRIDLSQLGGDSGGSPKKKRKV (SEQ
 ID NO: 186)

[0336] pNMG-622 amino acid sequence: ecTadA(wild-type)- 32 a.a. linker-

ecTadA(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_R152
 P_E155V_I156F_K157N)- 24 a.a. linker_nCas9_GGS_NLS

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSD
 SGGSSGGSSGSETPGTSESATPESSGGSSGSSGSEVEFSHEYWMRHALTLAKRAWDER
 EVPVGAVLVHNNRVIGEGWNRRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYV
 TFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHHPGMNHRVEITEGIL
 DECNALLCYFFRMPRQVFNAQKKAQSSDSSGGSSGGSSGSETPGTSESATPESDKKY
 SIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATR
 LKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGN
 IVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNS
 DVDKLFQILVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGL
 FGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAK
 NLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFD
 QSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIP
 HQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKS
 EETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHSLLYEYFTVYNELTK
 VKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGV
 EDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHL
 FDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLQIH

DDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRH
 KPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLY
 LYYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLTRSDKNRGKSD
 NVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVET
 RQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVINNY
 HHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYF
 FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNI
 KKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEK
 GKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLEENGR
 KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYL
 DEIIIEQISEFSKRVILADANLDKVL SAYNKHRRDKPIREQAENIIHLFTLTNLGAPAAFK
 YFDTTIDRKRYTSTKEVLDTLHQISITGLYETRIDLSQLGGDSGGSPKKKRKV (SEQ
 ID NO: 187)

[0337] pNMG-623 amino acid sequence: ecTadA(wild-type)- 32 a.a. linker-

ecTadA(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P
 _E155V_I156F_K157N)- 24 a.a. linker_nCas9_GGS_NLS

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTD
 SGGSSGGSSGSETPGTSESATPESSGGSSGGSSSEVEFSHEYWMRHALTLAKRALDERE
 VPVGAVLVHNNRVIGEGWNRVIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVT
 FEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILAD
 ECAALLCYFFRMPRQVFNAQKKAQSSTDSSGGSSGGSSGSETPGTSESATPESDKKYSI
 GLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRL
 KRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLLESFLVEEDKKHERHPIFGNI
 VDEVAYHEKYPTIYHLRKKLV DSTDKADLRILIYALAHMIKFRGHFLIEGDLNPDNS
 DVDKLFQQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGL
 FGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLDNLLAQIGDQYADLFLAAK
 NLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPKEYKEIFFD
 QSKNGYAGYIDGGASQEEFYKFIPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIP
 HQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYVGPLARGNSRFAWMTRKS
 EETITPWNFEVVDKGASQSFIERMTNFDKNLPNEKVLPHSLLYEYFTVYNELTK
 VKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGV
 EDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLEFEDREMIEERLKYAHL
 FDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFMQLIH
 DDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRH
 KPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLY
 LYYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLTRSDKNRGKSD
 NVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVET
 RQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVINNY
 HHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYF
 FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNI
 KKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEK
 GKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLEENGR
 KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYL
 DEIIIEQISEFSKRVILADANLDKVL SAYNKHRRDKPIREQAENIIHLFTLTNLGAPAAFK

YFDTTIDRKRYTSTKEVLDTLIHQSI TGLYETRIDLSQLGGDSGGSPKKKRKV (SEQ ID NO: 188)

[0338] ABE6.3 ecTadA(wild-type)-(SGGS)2-XTEN-(SGGS)2-

ecTadA(H36L_P48S_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTD
SGGSSGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRAWDER
EVPVGAVLVHNNRVIGEGWNRSIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYV
TFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILA
DECAALLCYFFRMRRQVFNAQKKAQSSTDSSGGSSGGSSGSETPGTSESATPESSGGSS
GGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS
GETAEATRLKRTARRRYTRRKNRICYLQEISNEMAKVDDSFHRLEESFLVEEDKK
HERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLA LAHMIKFRGHFLI
EGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQ
LPGEKKNGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQ
YADFLAAKNLSDAILSDILRVNTEITKAPLSASMIKRYDEHHQD TLTKALVRQQ
PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRK
QRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNS
RFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHSLLEYEY
FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIE
CFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIE
ERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFA
NRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDE
LVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE EGikelGSQILKEHPVEN
TQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTR
SDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL DKA
GFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQF
YKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRK
VLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWD PKKYGGFDSPTVAYS
VLVVAKEVGKSKKLKSVKELLGITIMERSSSF EKNPIDFLEAKGYKEVKKDLIIKLPK
YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ
LFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAENIIHLFTL
TNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQSI TGLYETRIDLSQLGGDSGGSP
KKKRKV* (SEQ ID NO: 189)

[0339] ABE7.8 ecTadA(wild-type)-(SGGS)2-XTEN-(SGGS)2-

ecTadA(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_E155V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTD
SGGSSGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRALDERE
VPVGAVLVHNNRVIGEGWNR AIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVT

FEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILAD
 ECNALLCYFFRMRRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSG
 GSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
 TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHE
 RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEG
 DLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLP
 GEKKNGLFGNLIALSLGLTPNFKS NFDAEDAKLQLSKDTYDDDLDNLLAQIGDQYA
 DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
 KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQ
 RTFDNGSIPHQIHLGELHAILRRQEDFYPLKDNREKIEKILTFRIPYYVGPLARGNSRF
 AWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFT
 VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECF
 DSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEE
 RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN
 RNFMQLIHDDSLTFKEDIKKAQVSGQGDLSHEHIANLAGSPAICKGILQTVKVVDEL
 VKVMGRHHPENIVIEMARENQTTQKGQKNSRERMKRIEEGKELGSQILKEHPVENT
 QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVHDHVPQSFLKDDSIDNKVLT
 RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAG
 FIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFY
 KVINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEI
 GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL
 SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL
 VVAKVEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIKLPKYS
 LFELENGRKRMLASAGELQKGNELALPSKYVNFYLASHYEKLKGSPEDENEQKQLF
 VEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHDKPIREQAENIIHLFTLTN
 LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSTGLYETRIDLSQLGGDSGGSPKK
 KRKV* (SEQ ID NO: 190)

[0340] ABE7.9 ecTadA(wild-type)-(SGGS)2-XTEN-(SGGS)2-

ecTadA(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y
 _R152P-_E155V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKKAQKKAQSST
 SGGSSGGSSGSETPGTSESATPESSGGSSGGSSSEVEFSHEYWMRHALTLAKRALDERE
 VPVGAVLVHNNRVIGEGWNRRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVT
 FEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILAD
 ECNALLCYFFRMRRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSG
 GSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
 TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHE
 RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEG
 DLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLP
 GEKKNGLFGNLIALSLGLTPNFKS NFDAEDAKLQLSKDTYDDDLDNLLAQIGDQYA
 DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
 KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQ
 RTFDNGSIPHQIHLGELHAILRRQEDFYPLKDNREKIEKILTFRIPYYVGPLARGNSRF
 AWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFT
 VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECF

DSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEE
 RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN
 RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKGILQTVKVVDL
 VKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIIEGKELGSQILKEHPVENT
 QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYVDHIVPQSFLKDDSIDNKVLTRS
 DKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAG
 FIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFY
 KVINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEI
 GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL
 SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL
 VVAKVEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKDLIIKLPKYS
 LFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLF
 VEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKHRRDKPIREQAENIIHLFTLTN
 LGAPAAFKYFDTTIDRKRYTSTKEVLDTLHQSITGLYETRIDLSQLGGDSGGSPKK
 KRKV* (SEQ ID NO: 191)

[0341] ABE7.10 ecTadA(wild-type)-(SGGS)2-XTEN-(SGGS)2-

ecTadA(W23R_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P
 -E155V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTD
 SGGSSGGSSGSETPGTSESATPESSGGSSGSSSEVEFSHEYWMRHALTLAKRARDERE
 VPVGAVLVHNNRVIGEGWNRALGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVT
 FEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILAD
 ECAALLCYFFRMPRQVFNAQKKAQSSTDSSGGSSGGSSGSETPGTSESATPESSGGSSG
 GSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
 TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHE
 RHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRILIYLAHMIKFRGHFLIEG
 DLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLP
 GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQYA
 DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQPE
 KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQ
 RTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRF
 AWMTRKSEETITPWNFEVVDKGAQAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFT
 VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECF
 DSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEE
 RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN
 RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKGILQTVKVVDL
 VKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIIEGKELGSQILKEHPVENT
 QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYVDHIVPQSFLKDDSIDNKVLTRS
 DKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAG
 FIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFY
 KVINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEI
 GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL
 SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL
 VVAKVEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKDLIIKLPKYS
 LFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLF

VEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLFTLTN
LGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDSGGSPKK
KRKV* (SEQ ID NO: 192)

[0342] ABE6.4: ecTadA(wild-type)-(SGGS)2-XTEN-(SGGS)2-

ecTadA(H36L_P48S_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_E155
V_I156F_K157N)-(SGGS)2-XTEN-(SGGS)2_nCas9_SGGS_NLS

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAKKAQSSTD
SGGSSGGSSGSETPGTSESATPESSGGSSGGSSSEVEFSHEYWMRHALTLAKRAWDER
EVPVGAVLVHNNRVIGEGWNRSIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYV
TFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILA
DECNALLCYFFRMRRQVFNAQKKAQSSTDSSGGSSGGSSGSETPGTSESATPESSGGSS
GGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS
GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKK
HERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLA LAHMIKFRGHFLI
EGDLNPDNSDV DKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQ
LPGEKKNGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQ
YADFLAAKNLS DAILSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ
PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRK
QRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTRIPYYVGPLARGNS
RFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHSLLYEY
FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIE
CFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIE
ERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFA
NRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDE
LVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE EG I KELGSQILKEHPVEN
TQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTR
SDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL DKA
GFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQF
YK VREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRK
VLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS
VLVVAKEVGKSKKLKSVKELLGITIMERSSSF EKNPIDFLEAKGYKEVKKDLIILPK
YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ
LFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLFTL
TNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDSGGSP
KKKRKV (SEQ ID NO: 180)

[0343] ABE_{max}

MKRTADGSEFESPKKKRKVMSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLV
HNNRVIGEGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCA
GAMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSD
FFRMRRQEIKAKKAQSSTDSSGGSSGGSSGSETPGTSESATPESSGGSSGGSSSEVEFSH
EYWMRHALTLAKRARDEREVPVGAVLVHNNRVIGEGWNR AIGLHDPTAHAEIMAL
RQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLM
DVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDSSGGSSGG

SSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKV
 LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK
 VDDSFHRLLESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKA
 DLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVD
 AKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTPNFKS NFDAEDAKLQ
 LSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK
 RYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILE
 KMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNRE
 KIEKILTRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMT
 NFDKNLPNEKVLPHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLL
 FKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNE
 ENEDILEDIVLTLTLFEDREMIEERLKTIAHLFDDKVMKQLKRRRYTGWGRLSRKL
 NGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIA
 NLAGSPAIIKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERM
 KRIEELGELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDV
 DHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ
 RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIR
 EVKVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVGTAIIKKYPKLESEF
 VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETN
 GETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARK
 KDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSSFENPID
 FLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFY
 LASHYEKLKGSPEDEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL SAYN
 KHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITG
 LYETRIDLSQLGGDKRTADGSEFEPKKKRKV (SEQ ID NO: 193)

[0344] ABE8c (monomer)

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVN
 NRVIGEGWNRAIGLHDPTAHAEIMALRQGGGLVMQNYRLIDATLYVTFEPCVMCAGA
 MIHSRIGRVVFGVRNSKRGAAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFY
 RMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGL
 AIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLK
 TARRRYTRRKNRICYLQEIFSNEMAK VDDSFHRLLESFLVEEDKKHERHPIFGNIV
 EVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDV
 DKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFG
 NLIASLGLTPNFKS NFDAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNL
 SDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQ
 KNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH
 QIHLGELHAILRRQEDFYFPLKDNREKIEKILTRIPYYVGPLARGNSRFAWMTRKSE
 ETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHSLLEYFTVYNELTKV
 KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVE
 DRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTIAHLF
 DDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIH
 DDLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKGILQTVKVVDELVKVMGRHKP
 ENIVIAMARENQTTQKGQKNSRERMKRIEELGELGSQILKEHPVENTQLQNEKLYLY
 YLQNGRDMYVDQELDINRLSDYDV DHIVPQSFLKDDSIDNKVLTRSDKNRGKSDN
 VPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQ
 ITHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVRINNYHH
 AHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFY

SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKK
TEVQTGGFSKESILPKRNSDKLIARKKDWD PKKYGGFDSPTVAYSVLVVAKEVGK
SKKLKSVKELLGITIMERSSSF EKNPIDFLEAKGYKEVKKDLIIKLPKYSLFEL ENGRKR
MLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDE
IIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLGAPAAFKYF
DTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQLGGDSGGSKRTADGSEFEPKK
KRKV (SEQ ID NO: 194)

[0345] ABE8e (dimer)

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHN
NRVIGEGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGA
MIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFR
MRRQEIKAKKAQSSTDSSGGSSGGSSGSETPGTSESATPESSGGSSGGSSSEVEFSHEY
WMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQ
GGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAGSLMN
LNYPGMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSINSGGSSGGSSG
SETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVP SKKFKVLGN
TDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDD
SFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRL
IYLA LAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAI
LSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD
TYDDDLNLLAQIGDQYADLFLAAKNLS DAILSDILRVNTEITKAPLSASMIKRYDE
HHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKI
LTFRIPIYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVLPHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN
RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENED
ILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRD
KQSGKTILDFLKS DGFANRNF MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAG
SPAIKKGILQTVKVVDLVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE
GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVP
QSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDN
LTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
LKSKLVSDFRKDFQFYK VREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDY
KVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEI
VWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWD
PKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSSF EKNPIDFLEAK
GYKEVKKDLIIKLPKYSLFEL ENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY
EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD
KPIREQAENIIHLFTLTNLGAPAAFKYFD TTTIDRKRYTSTKEVL DATLIHQ SITGLYETR
IDLSQLGGDSGGSKRTADGSEFEPKKRKV (SEQ ID NO: 195)

[0346] SaABE8e

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVN
NRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGA
MIHSRIGRVVFGVRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFY
RMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGTSESATPESSGGSSGGSGKRN YILG
LAIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKR RRRRHRIQR
VKKLLFDYNLLTDHSELGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNE

VEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTS DYVKEA
 KQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGH
 CTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYEYEFQIEN VFKQKKK
 PTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIENAE LLDQIAK
 ILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDN
 QIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDII
 IELAREKNSKDAQKMINEMQKRNRQTNERIEEII RTTGKENAKYLIEKIKLHDMQEGK
 CLYSLEAIPLEDLLNPNFNYEVDHIIPRSVSFDNSFNKVLVKQEENSKKGNRTPFQY
 LSSSDSKISYETFKKHILNLA KGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTR
 YATRGLMNNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAE DA
 LIANADFIFKEWKKLDKAKKVMENQMFE EKQAESMPEIETE QEYKEIFITPHQIKHIK
 DFKDYKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDK LKLI
 NKSPEKLLMYHHDPPQTYQKLKLIMEQYGDEKNPLYKYEETGNYLT KYSKKDN GP
 VIKKIKYYYGNKLN AHLDDITDDYPNSRNKVVKL SLKPYRFDVYLDNGVYKFVTVKNL
 DVIKKENYYYEVNSKCYEEAKK LKKISNQAEFIASFYNNDLIKINGEL YRVIGVNNDDL
 NRIEVNMIDITYREYLENMNDKRPPRII KTIASKTQSIKKYSTDILGNLYEVKSKKHPQI
 IKKGSGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 196)

[0347] SpCas9NG-ABE8e (“ABE8e-NG”)

MKRTADGSEFESP KKKRKVSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLV LN
 NRVIGEGWNRAIGLHDPTAHAEIMALRQGG LVMQNYRLIDATLYVTFEPCVMCAGAMI
 HSRIGRVVFGVRNSKRGAAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFYRMP
 RQVFNAQKKAQSSINSGGSSGGSSGSETPGTSESATPES SGGSSGSDKKYSIGLAIGT
 NSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARR
 RYTRRKNRICYLQEIFS NEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVAY
 HEKYPTIYHLRKKLV DSTDKADLR LIYLA LAHMIKFRGHFLIEGDLNPDNSDVKLFI
 QLVQTYNQLFEENPINASGVDAKAILSARLSKSRREN LIAQLPGEKKNGLFGNLIAL
 SLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLS DAIL
 LSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGY
 AGYIDGGASQEEFYKFIKPILEKMDGTEELLV KLNREDLLRKQRTFDNGSIPHQIHLG
 ELHAILRRQEDFY PFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPW
 NFEEVVDKGASQAQSFIERMTNFDKNLPNEKVL PKHSLLEYEYFTVYNELTKVKYVTE
 GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNA
 SLGTYHDL LKIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKV
 MKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFL KSDGFANRNFQMQLIHDDSLTF
 KEDIQKAQVSGQGDSLHEHIANLAGSPA IKKGILQTVKVVD ELVKVMGRHHPENIVIE
 MARENQTTQKGQKNSRERMKRIEEG IKELGSQILKEHPVENTQLQNEKLYLYYLQN
 GRDMYVDQELDINRLSDYD VDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEV
 VKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL DKAGFIKRQLVETRQITKHV
 AQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHHAHDAY
 LNAVVG TALIKKYPKLESEFVYGDYKVYDV RKMIKSEQEIGKATAKYFFYSNIMNF
 FKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFS
 KESIRPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVA KVEKGKSKKLKSVKE
 LLGITIMERSSFEKNPIDFLEAKGYKEVKKDLI IKLPKYSLFELNGRKRMLASARFLQ
 KGNELALPSKYVNF LYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKR
 VILADANLDKVL SAYNKH RDKPIREQAENIIHLFTL TNLGAPRAFKYFDTTIDRKVYR
 STKEVLDATLIHQ SITGLYETRIDLSQLGGDSGGSKRTADGSEFEPKKKRKV (SEQ ID
 NO: 197)

[0348] SaKKH-ABE8e (“ABE8e-KKH”)

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLN
 NRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGA
 MIHSRIGRVVFGVRNSKRGAAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFY
 RMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGTSESATPESSGGSSGGSGKRNILG
 LAIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRRHRIQR
 VKKLLFDYNLLTDHSELGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNE
 VEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYVKEA
 KQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGH
 CTYFPEELRSVKYAYNADLYNALNDLNLVITRDENEKLEYEYEFQIENVFQKQKK
 PTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIENAEELLDQIAK
 ILTIYQSSEDIQEELTNLSELTEQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDN
 QIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIKKYGLPNDII
 IELAREKNSKDAQKMINEMQKRNRQTNERIEEIRTTGKENAKYLIEKIKLHDMQEGK
 CLYSLEAIPLEDLLNPNFNYEVDHIIPRSVSFDNSFNKVLVKQEENSKKGNRTPFQY
 LSSSDSKISYETFKKHILNLAAGKGRISKTKKEYLLEERDINRFSVQKDFINRLVDTR
 YATRGLMNLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAEDA
 LIIANADFIKFEWKLDKAKKVMENQMFEKQAESMPEIETEYQYKEIFITPHQIKHIK
 DFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLLKKLI
 NKSPEKLLMYHHDPPQTYQKLKLIMEQYGDENPLYKYEEETGNYLTKYSKKDNBP
 VIKKIKYYYGNKLNALHDITDDYPNSRNKVVKLKSLKPYRFDVYLDNGVYKFVTVKNL
 DVIKKENYYYEVNSKCYEEAKKLLKISNQAEFIASFYKNDLIKINGELRYRIGVNNDDL
 NRIEVMIDITYREYLENMNDKRPPHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQ
 IIKKSGSGSKRTADGSEFESPKKKRKV (SEQ ID NO: 198)

[0349] ABE8-NRTH: NLS, TadA, linker, TadA, NRTH

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLN
 NRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGA
 MIHSRIGRVVFGVRNSKRGAAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFY
 RMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGTSESATPESSGGSSGGSSSEVEFSHEY
 WMRHALTLAKRARDEREVPVGAVLVLN NRVIGEGWNRAIGLHDPTAHAEIMALRQGGLV
 MQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAAGSLMNVLNYPGMNH
 RVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGTSESATP
 ESSGGSSGGSDKKYSIGLTIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF
 DSGETAETRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFHRLSEESFLVEEDKKHE
 RHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRILIYLAHAMIKFRGHFLIEGDLNP
 DNSDVKDLFIQLVQTYNQLFEEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF
 GNLIASLSGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADFLAAKNLSD
 AILLSDILRVNTEITKAPLSASMVKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA
 GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGIIPHQIHLGELHAI
 LRRQGDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVD
 KGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEFYFTVYNELTKVKYVTEGMRKPAFLSGE
 QKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKKIKDK
 DFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRLRYTGWRSLR
 KLINGIRDKQSGKTILDFLKSDGFANRFMQLIHDDSLTFKEDIQKAQVSCQGDSLHEHIA
 NLAGSPAIIKKGILQTVKVVDELVKVMGGHKPENIVIAMARENQTTQKGQKNSRERMKRIE
 EGikelGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVHDIVPQSF
 LKDDSIDNKVLTRSDKNRGKSDNPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERG
 GLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRK

DFQFYKVVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEI
 GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMP
 QVNIVKKTEVQTGGFSKESILPKGNSDKLIARKKDWDPKKYGGFNSPTVAYSVLVVAKEK
 GKSKKLKSVKELLGITIMERSSEFKNPIGFLEAKGYKEVKKDLIIKLPKYSLELENGRKR
 LASASVLHKGNELALPSKYVNFLYLASHYEKLKGSSSEDNKQKQLFVEQHKHYLDEIIEQISE
 FSKRVILADANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLGASAAFKYFDTTIGRKLYTS
 TKEVL DATLIHQ SITGLYETRIDLSQLGGDSGGSKRTADGSEFEPKKKRKV (SEQ ID NO:
 199)

[0350] ABE8-NRRH: NLS, TadA, linker, TadA, NRRH

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLV LN
 NRVIGEGWNRAIGLHDPTAHAEIMALRQGGGLVMQNYRLIDATLYVTFEPCVMCAGA
 MIHSRIGRVVFGVRNSKRGAAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFY
 RMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGTSESATPESSGGSSGGSSSEVEFSHEY
 WMRHALTLAKRARDEREVPVGAVLV LN NRVIGEGWNRAIGLHDPTAHAEIMALRQGGGLV
 MQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAAGSLMNVLNYPGMNH
 RVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGTSESATP
 ESSGGSSGGSDKKYSIGLTIGTNSVGWAVITDEYKVP SKKFKVLGNTDRHSIKKNLIGALLF
 DSGETAETRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFHRLEESFLVEEDKKHE
 RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LIYLALAHMIKFRGHFLIEGDLNP
 DNSDVKDLFIQLVQTYNQ LFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF
 GNLIASLSGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADFLAAKNLSDA
 ILLSDILRVNTEITKAPLSASMVKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAG
 YIDGGASQEEFYKFIKPILEKMDGTEELLVKNLREDLLRKQRTFDNGIIPHQIHLGELHAIL
 RRQGD FYPFLKDNREKIEKILTFRIPYVVGPLARGNSRFAWMTRKSEETITPWNFEVVVDK
 GASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQK
 KAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDFL
 DNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRLRYTGWGRLSRKLIN
 GIRDKQSGKTILDFLKS DGFANRNF MQLIHDDSLTFKEDIQKAQVSCQGD SLHEHIANLA
 GSPAIIKKGILQTVKVVDELVKVMGGGHKPENIVIE MARENQTTQKGQKNSRERMKRIEEGI
 KELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK
 DDSIDNKVLT RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL
 SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDF
 QFYKVVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGK
 ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVN
 IVKKTEVQTGGFSKESILPKGNSDKLIARKKDWDPKKYGGFNSPTAAYSVLVVAKEK GK
 KKLKSVKELLGITIMERSSEFKNPIGFLEAKGYKEVKKDLIIKLPKYSLELENGRKRMLAS
 AGVLHKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFS
 KRVILADANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLGVPAAFKYFDTTIDKKRYTSTK
 EVLDATLIHQ SITGLYETRIDLSQLGGDSGGSKRTADGSEFEPKKKRKV (SEQ ID NO:
 200)

[0351] xCas9(3.7)-ABE(7.10): (ecTadA(wt)–linker(32 aa)–ecTadA*(7.10)–linker(32 aa)–
 nxCas9(3.7)–NLS):

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
 AHAEIMALRQGGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
 GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAAQKKAQSSD
 SGGSSGGSSGSETPGTSESATPESSGGSSGGSSSEVEFSHEYWMRHALTLAKRARDERE

VPVGAVLVNLRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVT
 FEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILAD
 ECAALLCYFFRMPRQVFNAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSG
 GSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
 TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHE
 RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEG
 DLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLP
 GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDTKLQLSKDQYDDDLNLLAQIGDQYA
 DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKLYDEHHQDLTLLKALVRQQLP
 KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQ
 RTFDNGIIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRF
 AWMTRKSEETITPWNFEKVVDKGAQAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFT
 VYNELTKVKYVTEGMRKPAFLSGDQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECF
 DSVEISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEE
 RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN
 RNFIQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELV
 KVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEIEGKELGSQILKEHPVENTQ
 LQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLT
 RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL
 DKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
 KLVSDFRKDFQFYK VREINNYHHAHDAYLNAVVG
 TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIG
 KATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS
 MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS
 VLVAKVEKGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPKYS
 LFELENGRKRMLASAGVLQKGNELALPSKYVNFYLASHYEKLKGSPE
 DNEQKQLFVE QHKHYLDEIIEQISEFSKR
 VILADANLDKVLSAYNKH
 RDKPIREQAENIIHLFTLTNLG
 APAAFKYFDTTIDRKRYTSTKEVLDATLIHQ
 SITGLYETRIDLSQLGGDEGADKRTAD
 GSEFESPKKKRKV (SEQ ID NO: 201)

[0352] ABE8-VRQR: NLS, TadA, linker, TadA, SpCas9-VRQR

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVN
 NRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGA
 MIHSRIGRVVFGVRNSKRGAAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFY
 RMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGTSESATPESSGGSSGSSGSSEVEFSHEY
 WMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHAEIMALRQGGLV
 MQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAAGSLMNVLNYPGMNHR
 VEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGTSESATPE
 SSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFD
 SGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFFHRLEESFLVEEDKKHER
 HPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPD
 NSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFG
 NLIALSLGLTPNFKSNFDLAEDAKLQLSKDQYDDDLNLLAQIGDQYADLFLAAKNLSDAI
 LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLP
 EKYKEIFFDQSKNGYAGY IDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQ
 RTFDNGSIPHQIHLGELHAILR RQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRF
 AWMTRKSEETITPWNFEKVVDKGAQAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFT
 VYNELTKVKYVTEGMRKPAFLSGEQKK AIVDLLFKTNRKVTVKQLKEDYFKKIECF
 DSVEISGVEDRFNASLGTYHDLLKIKDKDFLD NEENEDILEDIVLTLTLFEDREMIEE
 RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLING IRDKQSGKTILDFLKSDGFAN
 RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGS

PAIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGKEL
 GSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSI
 DNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL
 DKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFY
 KVINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATA
 KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIVK
 KTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVAKVEKGKSKKL
 KSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLELENGRKRMLASAREL
 QKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIIEQISEFSKRVIL
 ADANLDKVL SAYNKHRRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVLD
 ATLIHQSTGLYETRIDLSQLGGDSGGSKRTADGSEFEPKKRKV (SEQ ID NO: 202)

[0353] ABE8e(TadA-8e V106W)

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVN
 NRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGA
 MIHSRIGRVVFGWRNSKRGAAGSLMNVLNYPGMNHRVEITEGILADECAALLCDFY
 RMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGL
 AIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKR
 TARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVD
 EVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDV
 DKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFG
 NLIALSLGLTPNFKSNFDLAEDAKLQLSKDQYDDDLNLLAQIGDQYADLFLAAKNL
 SDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQS
 KNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH
 QIHLGELHAILRRQEDFYPLKDNREKIEKILTRIPYYVGPLARGNSRFAWMTRKSE
 ETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
 KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVE
 DRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLF
 DDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHD
 DSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKP
 ENIVIAMARENQTTQKGQKNSRERMKRIEEGKELGSQILKEHPVENTQLQNEKLYLY
 YLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNV
 PSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQ
 ITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHH
 AHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFY
 SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIVKK
 TEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGK
 SKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLELENGRKR
 MLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDE
 IIEQISEFSKRVILADANLDKVL SAYNKHRRDKPIREQAENIIHLFTLTNLGAPAAFKYF
 DTTIDRKRYTSTKEVLDATLIHQSTGLYETRIDLSQLGGDSGGSKRTADGSEFEPKK
 KRKV (SEQ ID NO: 203)

Nuclear localization sequences (NLS)

[0354] In various embodiments, the fusion proteins delivered by the BE-VLPs described herein may comprise one or more nuclear localization sequences (NLS), which help promote

translocation of a protein into the cell nucleus. Such sequences are well-known in the art and can include the following examples:

Description	Sequence	SEQ ID NO:
NLS of SV40 large T-Ag	PKKKRKV	204
NLS	MKRTADGSEFESPKKKRKV	205
NLS	MDSLLMNRKFLYQFKNVRWAKGRRET YLC	206
NLS of nucleoplasmin	AVKRPAATKKAGQAKKKKLD	207
NLS of EGL-13	MSRRRKANPTKLSENAKKLAKEVEN	208
NLS of c-MYC	PAAKRVKLD	209
NLS of TUS-protein	KLKIKRPVK	210
NLS of polyoma large T-Ag	VSRKRPRP	211
NLS of Hepatitis D virus antigen	EGAPPAKRAR	212
NLS of murine p53	PPQPKKKPLDGE	213
NLS of PE1 and PE2	SGGSKRTADGSEFEPKKKKRKV	214
Bipartite sv40 nls	KRTADGSEFESPKKKRKV	215

[0355] The NLS examples above are non-limiting. The fusion proteins delivered by the presently described BE-VLPs may comprise any known NLS sequence, including any of those described in Cokol *et al.*, “Finding nuclear localization signals,” *EMBO Rep.*, 2000, 1(5): 411-415 and Freitas *et al.*, “Mechanisms and Signals for the Nuclear Import of Proteins,” *Current Genomics*, 2009, 10(8): 550-7, each of which are incorporated herein by reference.

[0356] In various embodiments, the fusion proteins, constructs encoding the fusion proteins, and BE-VLPs disclosed herein further comprise one or more, preferably, at least two nuclear localization sequences. In certain embodiments, the fusion proteins comprise at least two NLSs. In embodiments with at least two NLSs, the NLSs can be the same NLSs or they can be different NLSs. In some embodiments, one or more of the NLSs are bipartite NLSs (“bpNLS”). In certain embodiments, the disclosed fusion proteins comprise two bipartite NLSs. In some embodiments, the disclosed fusion proteins comprise more than two bipartite NLSs.

[0357] The location of the NLS fusion can be at the N-terminus, the C-terminus, or within a sequence of a fusion protein (*e.g.*, inserted between the encoded napDNAbp component (*e.g.*, Cas9) and a deaminase domain (*e.g.*, an adenosine or cytosine deaminase)).

[0358] The NLSs may be any known NLS sequence in the art. The NLSs may also be any future-discovered NLSs for nuclear localization. The NLSs also may be any naturally-occurring NLS, or any non-naturally occurring NLS (*e.g.*, an NLS with one or more desired mutations).

[0359] The term “nuclear localization sequence” or “NLS” refers to an amino acid sequence that promotes import of a protein into the cell nucleus, for example, by nuclear transport. Nuclear localization sequences are known in the art and would be apparent to the skilled artisan. For example, NLS sequences are described in Plank *et al.*, International PCT application PCT/EP2000/011690, filed November 23, 2000, published as WO/2001/038547 on May 31, 2001, the contents of which are incorporated herein by reference. In some embodiments, an NLS comprises the amino acid sequence PKKKRKV (SEQ ID NO: 204), MDSLLMNRKFLYQFKNVRWAKGRRETYLC (SEQ ID NO: 206), KRTADGSEFESPKKKRKV (SEQ ID NO: 215), or KRTADGSEFEPKKKKRKV (SEQ ID NO: 216). In other embodiments, NLS comprises the amino acid sequences NLSKRPAAIKKAGQAKKKK (SEQ ID NO: 217), PAAKRVKLD (SEQ ID NO: 209), RQRNELKRSF (SEQ ID NO: 218), or NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY (SEQ ID NO: 219).

[0360] In one aspect of the disclosure, a base editor or other fusion protein may be modified with one or more nuclear localization sequences (NLS), preferably at least two NLSs. In certain embodiments, the fusion proteins are modified with two or more NLSs. The disclosure contemplates the use of any nuclear localization sequence known in the art at the time of the disclosure, or any nuclear localization sequence that is identified or otherwise made available in the state of the art after the time of the instant filing. A representative nuclear localization sequence is a peptide sequence that directs the protein to the nucleus of the cell in which the sequence is expressed. A nuclear localization signal is predominantly basic, can be positioned almost anywhere in a protein’s amino acid sequence, generally comprises a short sequence of four amino acids (Autieri & Agrawal, (1998) *J. Biol. Chem.* 273: 14731-37, incorporated herein by reference) to eight amino acids, and is typically rich in lysine and arginine residues (Magin *et al.*, (2000) *Virology* 274: 11-16, incorporated herein by reference). Nuclear localization sequences often comprise proline residues. A variety of nuclear localization sequences have been identified and have been used to effect transport of biological molecules from the cytoplasm to the nucleus of a cell. See, *e.g.*, Tinland *et al.*, (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89:7442-46; Moede *et al.*, (1999) *FEBS Lett.* 461:229-34, which is incorporated herein by reference. Translocation is currently thought to involve nuclear pore proteins.

[0361] Most NLSs can be classified in three general groups: (i) a monopartite NLS exemplified by the SV40 large T antigen NLS (PKKKRKV (SEQ ID NO: 204)); (ii) a

bipartite motif consisting of two basic domains separated by a variable number of spacer amino acids and exemplified by the *Xenopus nucleoplasmin* NLS (KRXXXXXXXXXXXXKKKL (SEQ ID NO: 220)); and (iii) noncanonical sequences such as M9 of the hnRNP A1 protein, the influenza virus nucleoprotein NLS, and the yeast Gal4 protein NLS (Dingwall and Laskey 1991).

[0362] Nuclear localization sequences appear at various points in the amino acid sequences of proteins. NLS have been identified at the N-terminus, the C-terminus, and in the central region of proteins. Thus, the disclosure provides fusion proteins that may be modified with one or more NLSs at the C-terminus and/or the N-terminus, as well as at internal regions of the fusion protein. The residues of a longer sequence that do not function as component NLS residues should be selected so as not to interfere, for example, tonically or sterically, with the nuclear localization signal itself. Therefore, although there are no strict limits on the composition of an NLS-comprising sequence, in practice, such a sequence can be functionally limited in length and composition.

[0363] The present disclosure contemplates any suitable means by which to modify a fusion protein to include one or more NLSs. In one aspect, the fusion proteins may be engineered to express a fusion protein that is translationally fused at its N-terminus or its C-terminus (or both) to one or more NLSs, *i.e.*, to form a base editor-NLS fusion construct. In other embodiments, a fusion protein-encoding nucleotide sequence may be genetically modified to incorporate a reading frame that encodes one or more NLSs in an internal region of the encoded base editor. In addition, the NLSs may include various amino acid linkers or spacer regions encoded between the base editor and the N-terminally, C-terminally, or internally-attached NLS amino acid sequence, *e.g.*, and in the central region of proteins. Thus, the present disclosure also provides for nucleotide constructs, vectors, and host cells for expressing fusion proteins that comprise a base editor and one or more NLSs, among other components.

[0364] The fusion proteins delivered by the BE-VLPs described herein may also comprise nuclear localization sequences that are linked to a base editor through one or more linkers, *e.g.*, a polymeric, amino acid, nucleic acid, polysaccharide, chemical, or nucleic acid linker element. The linkers within the contemplated scope of the disclosure are not intended to have any limitations and can be any suitable type of molecule (*e.g.*, polymer, amino acid, polysaccharide, nucleic acid, lipid, or any synthetic chemical linker domain) and can be

joined to the base editor by any suitable strategy that effectuates forming a bond (*e.g.*, covalent linkage, hydrogen bonding) between the base editor and the one or more NLSs.

Nuclear export sequences (NES)

[0365] In various embodiments, the fusion proteins delivered by the BE-VLPs described herein may comprise one or more nuclear export sequences (NES), which help promote translocation of a protein out of the cell nucleus. Nuclear export sequences (or nuclear export signals) have the opposite function of nuclear localization signals (NLSs). Such sequences are well-known in the art (*e.g.*, Xu et al., “Sequence and structural analyses of nuclear export signals in the NESdb database,” *Mol. Biol. Cell*, 2012, 23(18): 3677-3693, the contents of which are incorporated herein by reference) and can include the following examples:

SEQUENCE:	SEQ ID NO:
MEELSQLASSFSV	221
PLQLPPLERLTL	222
NELALKLAGLDI	223
ERFEMFRELNEALEL	224
DHAEKVAEKLALS SV	225
QLVEELLKIIICAFQL	226
TNLEALQKKLEEL	227
DVKEEMTSALATMRV	228
STNGSLAAEFRHLQL	229
PSVQELTEQIHRLLM	230
MNFKELKDFLKELNI	231
ENFEILMKLKESLEL	232
FETVYELTKMCTIR	233
SGKASSSLGLQDFDL	234
PKYSDIDVDGLCSEL	235
VDLACTPTDVRDVI	236
YGEKTTQRDLTELEI	237
RRIYDITNVLEGIGL	238
AKIIPYSGLLLIVITV	239
LRSEEVHVLHVDMGV	240
LQSEEVHVLHLDMGV	241
LQVRKYSLDLASLIL	242
AGVEAIIRILQQLLF	243
TGVEALIRILQQLLF	244
IVLNQLCVRFFGLDL	245
SLGGFEITPPVVLRL	246
EAIQDLCLAVEEVSL	247
DELLQVLRMMVGVNI	248
SVMLAVQEGIDLLTF	249
LSSHFAQELSI	250
QSTHVDIRTLEDLLM	251

ESSAEDLRTLQQLFL	252
EFSLPTHHTVRLIRV	253
MSSGYLGEILRLAL	254
DTVLDILRDFELRL	255
NSVNEILSEFYVRL	256
CAFLSVKKQFEELTL	257
ISPEHVIQALESLGF	258
AHWMRQLVSFQKLKL	259
ATRELDELMASLSDF	260
YQNIELITFINALKL	261
FNATAVVRHMRKLQL	262
SGIFGLVTNLEELEV	263
EESYTLNSDLARLGV	264
EESYDLTSHLARLGV	265
GIQQAHAEQLANMRI	266
DVKEEMTSALATMRV	228
AAEPVILDLRDLFQL	267
MEGCVSNLMV	268
EGCVSNLMV	269
DMDFLRNLFSQTLSL	270
EQLLEIVHDLENLSL	271
NVMKYFTDLFDYLPL	272
KVYPILRLGSNLSL	273
YAGFSLPHAILRIDL	274
EIVRDIKEKLCYVAL	275
EAINKLESNLRELQI	276
EAINKLENNLRELQI	277
SDQKQEQLLLKKMYL	278
KQVLWDRTFSLFQQL	279
AQLQNLTKRIDSPL	280
NDENEHQLSLRTVSL	281
ISFTEFVKVLEKVDV	282
MESAITLWQFLQL	283
VPKELMQQIENFEKI	284
QARFILEKIDGKIII	285
QVKFIKMIIEKELTV	286
NHRMKNLREISQLGI	287
NHRVKKLNEISKLGI	288
TEKHLQKYLRQDLRL	289
RQERKRPLLDLHIEL	290
ANMRIQDLKVSLKPL	291
ATMRVDYEQIKIKKI	292
LQGEEFVCLKSIILL	293
THYGQKAILFLPLPV	294
PSAHEITGLADSLQL	295
URLHDVLHSDKKLTL	296
LINRNGELKLANFGL	297
LEPLKKLECLKSLDL	298

[0366] The NES examples above are non-limiting. The fusion proteins delivered by the presently described BE-VLPs may comprise any known NES sequence, including any of those described in Xu, D. *et al.* Sequence and structural analyses of nuclear export signals in the NESdb database. *Mol. Biol. Cell.* 2012, 23(18), 3677-3693; Fung, H. Y. J. *et al.* Structural determinants of nuclear export signal orientation in binding to exportin CRM1. *eLife.* 2015, 4:e10034; and Kosugi, S. *et al.* Nuclear Export Signal Consensus Sequences Defined Using a Localization-based Yeast Selection System. *Traffic.* 2008, 9(12), 2053-2062, each of which are incorporated herein by reference.

[0367] In various embodiments, the fusion proteins, constructs encoding the fusion proteins, and BE-VLPs disclosed herein further comprise one or more, preferably, at least three nuclear export sequences. In certain embodiments, the fusion proteins comprise at least three NESs. In embodiments with at least three NESs, the NESs can be the same NESs or they can be different NESs. In certain other embodiments, the fusion proteins, constructs encoding the fusion proteins, and BE-VLPs may comprise at least one, at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, at least ten or more NESs. In general, the one or more NESs are of sufficient strength to drive accumulation of the BE-VLPs proteins (e.g., the Gag-cargo) in a detectable amount respectively in the cytoplasm of a producer cell.

[0368] The location of the NES fusion can be at the N-terminus, the C-terminus, or within a sequence of a fusion protein (e.g., inserted between the encoded napDNAbp component (e.g., Cas9) and the gag nucleocapsid protein). In certain preferred embodiments, the NES (or multiple NESs, e.g., three NESs) are positioned between the napDNAbp and the gag nucleocapsid protein such that they can be cleaved from the napDNAbp upon delivery of the fusion protein to a target cell. NES sequences may preferably be joined to a fusion protein via a cleavable linker, such as protease-cleavable linker (e.g., the Gag-Pro-Pol). In this way, as shown in the fourth generation eVLPs described herein, the NES may be removed from the cargo protein (e.g., a BE or napDNAbp) after VLP maturation so that the BE and/or napDNAbp cargo may be free to translocate to the nucleus once delivered to a recipient cell.

[0369] The NESs may be any known NES sequence in the art. The NESs may also be any future-discovered NESs for nuclear export. The NESs also may be any naturally-occurring NES, or any non-naturally occurring NES (e.g., an NES with one or more desired mutations).

[0370] The term “nuclear export sequence” or “NES” refers to an amino acid sequence that promotes export of a protein from the cell nucleus, for example, by nuclear transport. Nuclear export sequences are known in the art and would be apparent to the skilled artisan.

[0371] In one aspect of the disclosure, a base editor or other fusion protein may be modified with one or more nuclear export sequences (NES), preferably at least three NESs. In certain embodiments, the fusion proteins are modified with two or more, three or more, five or more, six or more, seven or more, eight or more, nine or more, or ten or more NESs. The disclosure contemplates the use of any nuclear export sequence known in the art at the time of the disclosure, or any nuclear export sequence that is identified or otherwise made available in the state of the art after the time of the instant filing. A representative nuclear export sequence is a peptide sequence that directs the protein out of the nucleus of the cell in which the sequence is expressed. NESs commonly contain hydrophobic amino acid residues in the sequence LXXXLXXLXL, where L is a hydrophobic residue (frequently leucine), and X represents any amino acid. Nuclear export sequences often comprise leucine residues.

[0372] The fusion proteins delivered by the BE-VLPs described herein may also comprise nuclear export sequences that are linked to a base editor through one or more linkers, *e.g.*, a polymeric, amino acid, nucleic acid, polysaccharide, chemical, or nucleic acid linker element. The linkers within the contemplated scope of the disclosure are not intended to have any limitations and can be any suitable type of molecule (*e.g.*, polymer, amino acid, polysaccharide, nucleic acid, lipid, or any synthetic chemical linker domain) and can be joined to the base editor by any suitable strategy that effectuates forming a bond (*e.g.*, covalent linkage, hydrogen bonding) between the base editor and the one or more NESs. In some embodiments, the linker joining one or more NES and a base editor is a cleavable linker, as described further herein, such the one or more NES can be cleaved from the base editor, *e.g.*, upon delivery of the base editor to a target cell.

[0373] In various embodiments it may be useful to monitor the accumulation of a BE-VLP protein in the cytoplasm and/or nucleus, for example, to confirm that a protein cargo (*e.g.*, a Gag-BE is accumulating in the cytoplasm (not the nucleus) during the process of VLP production in a producer cell. In other embodiments it may be useful to monitor the accumulation of a BE-VLP protein in the nucleus and/or nucleus, for example, to confirm in a recipient cell that receives an eVLP for BE delivery that the delivered BE actually ends up being transported to the nucleus where it may edit DNA. Detection of accumulation in the nucleus or cytoplasm, as the case may be, can be performed by any suitable technique. For

example, a detectable marker may be fused to a BE such that location within a cell may be visualized, such as in combination with a means for detecting the location of the nucleus (e.g., a stain specific for the nucleus such as DAPI). Examples of detectable markers include fluorescent proteins (such as Green fluorescent proteins, or GFP; RFP; CFP), and epitope tags (HA tag, flag tag, SNAP tag). Cell nuclei may also be isolated from cells, the contents of which may then be analyzed by any suitable process for detecting protein, such as immunohistochemistry, Western blot, or enzyme activity assay.

Linkers

[0374] The fusion proteins and BE-VLPs described herein may include one or more linkers. As defined above, the term “linker,” as used herein, refers to a chemical group or a molecule linking two molecules or moieties, *e.g.*, a binding domain and a cleavage domain of a nuclease. In some embodiments, a linker joins a gRNA binding domain of an RNA-programmable nuclease and the catalytic domain of a deaminase (*e.g.*, a cytosine deaminase or an adenosine deaminase). In some embodiments, a linker joins a dCas9 and a deaminase. Typically, the linker is positioned between, or flanked by, two groups, molecules, or other moieties and connected to each one via a covalent bond, thus connecting the two. In some embodiments, the linker is an amino acid or a plurality of amino acids (*e.g.*, a peptide or protein). In some embodiments, the linker is an organic molecule, group, polymer, or chemical moiety. In some embodiments, the linker is 5-100 amino acids in length, for example, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 30-35, 35-40, 40-45, 45-50, 50-60, 60-70, 70-80, 80-90, 90-100, 100-150, or 150-200 amino acids in length. Longer or shorter linkers are also contemplated.

[0375] The linker may be as simple as a covalent bond, or it may be a polymeric linker many atoms in length. In certain embodiments, the linker is a polypeptide, or amino acid-based. In other embodiments, the linker is not peptide-like. In certain embodiments, the linker is a covalent bond (*e.g.*, a carbon-carbon bond, disulfide bond, carbon-heteroatom bond, etc.). In certain embodiments, the linker is a carbon-nitrogen bond of an amide linkage. In certain embodiments, the linker is a cyclic or acyclic, substituted or unsubstituted, branched or unbranched aliphatic or heteroaliphatic linker. In certain embodiments, the linker is polymeric (*e.g.*, polyethylene, polyethylene glycol, polyamide, polyester, etc.). In certain embodiments, the linker comprises a monomer, dimer, or polymer of aminoalkanoic acid. In certain embodiments, the linker comprises an aminoalkanoic acid (*e.g.*, glycine, ethanoic acid, alanine, beta-alanine, 3-aminopropanoic acid, 4-aminobutanoic acid, 5-pentanoic acid,

etc.). In certain embodiments, the linker comprises a monomer, dimer, or polymer of aminohexanoic acid (Ahx). In certain embodiments, the linker is based on a carbocyclic moiety (*e.g.*, cyclopentane, cyclohexane). In other embodiments, the linker comprises a polyethylene glycol moiety (PEG). In other embodiments, the linker comprises amino acids. In certain embodiments, the linker comprises a peptide. In certain embodiments, the linker comprises an aryl or heteroaryl moiety. In certain embodiments, the linker is based on a phenyl ring. The linker may include functionalized moieties to facilitate attachment of a nucleophile (*e.g.*, thiol, amino) from the peptide to the linker. Any electrophile may be used as part of the linker. Exemplary electrophiles include, but are not limited to, activated esters, activated amides, Michael acceptors, alkyl halides, aryl halides, acyl halides, and isothiocyanates.

[0376] In some other embodiments, the linker comprises the amino acid sequence (GGGS)_n (SEQ ID NO: 299), (G)_n (SEQ ID NO: 300), (EAAAK)_n (SEQ ID NO: 301), (GGS)_n (SEQ ID NO: 302), (SGGS)_n (SEQ ID NO: 303), (XP)_n (SEQ ID NO: 304), or any combination thereof, wherein n is independently an integer between 1 and 30, and wherein X is any amino acid. In some embodiments, the linker comprises the amino acid sequence (GGS)_n (SEQ ID NO: 302), wherein n is 1, 3, or 7. In some embodiments, the linker comprises the amino acid sequence SGSETPGTSESATPES (SEQ ID NO: 305). In some embodiments, the linker comprises the amino acid sequence SGGSSGGSSGSETPGTSESATPESGGSSGGSS (SEQ ID NO: 306). In some embodiments, the linker comprises the amino acid sequence SGGSSGGSSGGSS (SEQ ID NO: 307). In some embodiments, the linker comprises the amino acid sequence SGGS (SEQ ID NO: 303). In other embodiments, the linker comprises the amino acid sequence

SGGSSGGSSGSETPGTSESATPESAGSYPYDVPDYAGSAAPAAKKKKLDGSGSGSSGGSS (SEQ ID NO: 308, 60AA). In some embodiments, the linker comprises the amino acid sequence GGS (SEQ ID NO: 302), GGSGGS (SEQ ID NO: 309), GGSGSGGS (SEQ ID NO: 310), SGGSSGGSSGSETPGTSESATPESGGSSGGSS (SEQ ID NO: 311), SGSETPGTSESATPES (SEQ ID NO: 305), or SGGSSGGSSGSETPGTSESATPESAGSYPYDVPDYAGSAAPAAKKKKLDGSGSGSSGGSS (SEQ ID NO: 312).

[0377] In certain embodiments, linkers may be used to link any of the peptides or peptide domains or moieties of the invention (*e.g.*, a napDNAbp linked or fused to a deaminase domain, and/or a napDNAbp linked to one or more NESs). Any of the domains of the fusion

proteins described herein may also be connected to one another through any of the presently described linkers.

[0378] In some embodiments, a linker is a cleavable linker (*e.g.*, a linker that can be split or cut by any means). A cleavable linker may be an amino acid sequence. In some embodiments, the linker between one or more NES and the napDNAbp of the fusion proteins and BE-VLPs provided herein comprises a cleavable linker. A cleavable linker may comprise a self-cleaving peptide (*e.g.*, a 2A peptide such as EGRGSLLTCTGDVEENPGP (SEQ ID NO: 9), ATNFSLLKQAGDVEENPGP (SEQ ID NO: 10), QCTNYALLKLAGDVESNPGP (SEQ ID NO: 11), or VKQTLNFDLLKLAGDVESNPGP (SEQ ID NO: 12)). In some embodiments, a cleavable linker comprises a protease cleavage site that is cut after being contacted by a protease. For example, the present disclosure contemplates the use of cleavable linkers comprising a protease cleavage site of amino acid sequences TSTLLMENSS (SEQ ID NO: 1), PRSSLYPALTP (SEQ ID NO: 2), VQALVLTQ (SEQ ID NO: 3), PLQVLTLNIERR (SEQ ID NO: 4), or an amino acid sequence at least 90% identical to any one of SEQ ID NOs: 1-4. In certain embodiments, a cleavable linker comprises an MMLV protease cleavage site of an FMLV protease cleavage site. In certain embodiments, the fusion proteins and BE-VLPs described herein comprise the cleavable linker TSTLLMENSS (SEQ ID NO: 1) joining one or more NES and a napDNAbp. In some embodiments, the linker is cleaved upon delivery of the BE-VLP/fusion protein to a target cell, releasing a free base editor that is capable of translocating into the nucleus of the target cell.

[0379] The protease cleavage site may be any known in the art, or any sequence yet to be discovered, so long as the corresponding protease may be co-packaged in the eVLPs to allow for post-maturation cleavage within the mature eVLP particles. Such cleavage sites and their corresponding proteases include but are not limited to: (a) granzyme A, which recognizes and cleaves a sequence comprising ASPRAGGK (SEQ ID NO: 5), (b) granzyme B, which recognizes and cleaves a sequence comprising YEADSLEE (SEQ ID NO: 6), (c) granzyme K, which recognizes and cleaves a sequence comprising YQYRAL (SEQ ID NO: 7), (d) Cathepsin D, which recognizes and cleaves a sequence comprising LGVLIV (SEQ ID NO: 8). Many other combinations of specific proteases and protease cleavage sites may be used in connection with the present disclosure by co-packing a specific protease during the eVLP manufacture process. Such proteases can include, without limitation, Arg-C proteinase, Asp-N Endopeptidase, Caspase 1, Caspase 2, Caspase 3, Caspase 4, Caspase 5, Caspase 7,

Caspase 8, Caspase 9, Caspase 10, Chymotrypsin, Clostripain, Enterokinase, Factor Xa, Glutamyl endopeptidase, Granzyme B, Neutrophil elastase, Pepsin, Prolyl-endopeptidase, Proteinase K, Staphylococcal peptidase I, Thermolysin, Thrombin, and Trypsin. Any protease paired with its cognate recognition sequence may be used in the present disclosure protease-sensitive linkers, including any serine protease, cysteine protease, aspartic protease, threonine protease, glutamic protease, metalloprotease, or asparagine peptide lyase (which constitute major classifications of known proteases). The specific protease cleavage sites for said enzymes are well-known in the art and may be utilized in the linkers herein to provide protease-susceptible linkers.

Group-specific antigen (gag) proteins and viral envelope glycoproteins

[0380] The BE-VLPs described herein include various viral envelope and capsid components, which are used to encapsulate and deliver the base editor fusion proteins described herein. The use of viral envelope and capsid components for nucleic acid and protein delivery is known in the art, and a person of ordinary skill in the art would readily appreciate the various options known in the art that could be used or substituted for these components in the presently described BE-VLPs. The use of such viral components for nucleic acid and/or protein delivery (*e.g.*, delivery of Cas9) is described, for example, in Mangeot *et al.*, *Nat. Commun.* 10, 45 (2019); Gutkin, *et al. Nat. Biotechnol.* (2021); and Hamilton, J. R. *et al. Cell Reports* 35(9), 109207 (2021), each of which is incorporated herein by reference.

[0381] In some embodiments, the BE-VLPs described herein comprise a viral envelope glycoprotein layer as the outermost layer of the BE-VLP. Viral envelope glycoproteins are oligosaccharide-containing proteins that form a part of the viral envelope, *i.e.*, the outermost layer of many types of viruses that protects the viral genetic materials when traveling between host cells. Glycoproteins may assist with identification and binding to receptors on a target cell membrane so that the viral envelope fuses with the membrane, allowing the contents of the viral particle (which may comprise, *e.g.*, a fusion protein in a BE-VLP as described herein) to enter the host cell.

[0382] The viral envelope glycoproteins used in the BE-VLPs of the present disclosure may comprise any glycoprotein from an enveloped virus. In some embodiments, a viral envelope glycoprotein is an adenoviral envelope glycoprotein, an adeno-associated viral envelope glycoprotein, a retroviral envelope glycoprotein, or a lentiviral envelope glycoprotein. In certain embodiments, a viral envelope glycoprotein is a vesicular stomatitis virus G protein

(VSV-G), a baboon retroviral envelope glycoprotein (BaEVRless), a FuG-B2 envelope glycoprotein, or an ecotropic murine leukemia virus (MLV) envelope glycoprotein.

[0383] Any known viral envelope glycoprotein can be used in the eVLPs of the present disclosure. Any viral envelope glycoprotein discovered or characterized in the future can also be used in the eVLPs of the present disclosure. A person of ordinary skill in the art would readily be able to find additional viral envelope glycoproteins that could be used in the eVLPs described herein. For example, viral envelope glycoproteins are described in Banerjee, V. and Mukhopadhyay, S. *VirusDisease* (2016), 27(1), 1-11 and Li, Y. *et al. Front. Immunol.* (2021), 12, 1-12, each of which is incorporated herein by reference.

[0384] The viral envelope glycoproteins used in the VLPs described herein may also be capable of targeting the VLPs to a particular cell type (*e.g.*, immune cells, neural cells, retinal pigment epithelium cells, *etc.*). For example, using different envelope glycoproteins in the eVLPs described herein may alter their cellular tropism, allowing the eVLPs to be targeted to specific cell types. The process of producing a viral vector in combination with foreign viral envelope proteins is known as pseudotyping. Using pseudotyping, foreign viral envelope glycoproteins can be used to alter the cellular tropism of a VLP. Envelope glycoproteins incorporated into the VLP allow it to readily enter different cell types with the corresponding host receptor. Pseudotyping of viral vector systems is known in the art and is described further, for example, in Hamilton, J. R. *et al. Targeted delivery of CRISPR-Cas9 and transgenes enables complex immune cell engineering. Cell Reports.* 2021, 35, 109207; Kato, S. *et al. Selective Neural Pathway Targeting Reveals Key Roles of Thalamostriatal Projection in the Control of Visual Discrimination. J. Neurosci.* 2011, 31(47), 17169-17179; and Kato, S. *et al. A lentiviral strategy for highly efficient retrograde gene transfer by pseudotyping with fusion envelope glycoprotein. Human Gene Ther.* 2011, 22(2), 197-206, each of which is incorporated herein by reference.

[0385] Thus, the use of different glycoproteins in the VLPs described herein may be employed to alter their cellular tropism. Retrovirus tropisms may be readily modulated by pseudotyping virions with different envelope glycoproteins, enabling targeting of VLPs to specific cell types. In some embodiments, the viral envelope glycoprotein is a VSV-G protein, and the VSV-G protein targets the VLP to retinal pigment epithelium (RPE) cells. In some embodiments, the viral envelope glycoprotein is an HIV-1 envelope glycoprotein, and the HIV-1 envelope glycoprotein targets the VLP to CD4+ cells. In some embodiments, the

viral envelope glycoprotein is a FuG-B2 envelope glycoprotein, and the FuG-B2 envelope glycoprotein targets the VLP to neurons.

[0386] In some embodiments, exemplary viral envelope glycoproteins that may be used to target the presently described VLPs to particular cell types include, but are not limited to, glycoproteins of the following amino acid sequences:

HIV-1 envelope glycoprotein	MRVKEKYQHLWRWGWKWGIMLLGILMICSATENLWVTVYYGVPVWKEATTTLFCASDAK AYDTEVHNVCATHACVPTDPNPQEVILNVNTENFDMWKNDMVEQMHEDIISLWDQSLKPCV KLTPLCVNLKCTDLKNDTNTNSSNGRMIMEKGEIKNCSFNISTSIRNKVQKEYAFFYKLDIRPI DNTTYRLISCNTSVITQACPKVSFEPIPIHYCAPAGFAILKCNKDTFNGTGPCTNVSTVQCTHGI RPVVSTQLLNLGSLAEEEGVIRSANFTDNAKTIIVQLNTSVEINCTRPNNNTRKSIRIQRGPR AFVTIGKIGNMRQAHCNISRAKWMSTLQKQASKLREQFGNNKTVIFKQSSGGDPEIVTHSFNC GGEFFYCNSTQLFNSTWFNSTWSTEGSNNTGSDTITLPCRIKQFINMWQEVGKAMYAPPISG QIRCSSNITGLLLTRDGGKNTNESEVFRPGGDMRDNRSELYKYKVVKIETLGVAPTAKR RVVQREKRAVGIGALFLGFLGAAGSTMTGAASMTLTVQARQLLSGIVQQNNLLRAIEAQQH LLQLTVWGKQLQARILAVERYLKDQQLLGIWGCSGKLICTTAVPWNASWSNKSLEQFWNN MTWMEWDREINNYTSLIHSIDESQNNQKEKNEQELLELDKWASLWNWFNITNWLWYIKIFIM IVGGLVGLRIVFAVLSIVNRVRQGYSPFSQTHLPNRGGPDRPEGIEEEGGERDRDRSVRLVNG SLALIWDLLRSLCLFSYHRLRDLILLIVTRIVELLGRRGWEALKYWWNLLQYWSQELKNSAVS LLNATAIAVAEGTDRVIEVVQGAYRAIRHIPRRIRQGLERIL (SEQ ID NO: 313)
FuG-B2 envelope glycoprotein	MVPQALLFVPLLVPPLCFGKFPYITPDKLGWSPIDIHHLSCPNLNVVEDEGCTNLSGFSYME LKVGYSIAIKMNGFTCTGVVTEAETYTNFVGYYTTTTFKRKHFRPTPDACRAAYNWKMAGDP RYEESLHNPYPDYHWLRTVKTTKESLVIISPSVADLDPYDRSLHSPVFPGGNCSGVAVSSTYCS TNHDYTIWMPENPRLGMSCDIFTNSRGKRASKGSETCGFVDERGLYKSLKGACKLKLCGVL GLRLMDGTWVAMQTSNETKWCPGQLVNLHDFRSDEIEHLVVEELVKKREECLDALESIMTT KSVSFRRSLHLRKLVPFGFKAYTIFNKTLMEDIAHYKSVRTWNEIIPSKGCLRVGGRCRPHV NGVFFNGIILGPDGNVLIPEMQSSLLQQHMELLVSSVIPLMHPLADPSTVFKNGDEAEDFVEV HLPDVHERISGVDLGLPNWKGKYLVSAGALTALMLIIFLMTWCVRVGIHLCLIKLHTKKRQIYT DIEMNRLGK (SEQ ID NO: 314)
VSV-G protein	MKCLLYLAFLFIGVNCKFTIVFPHNQKGNWKNVPSNYHYCPSSSDLNWHNDLIGTAIQVKMP KSHKAIQADGWMCHASKWVTTCDFRWYGPKYITQSIRSFTPSVEQCKESIEQTKQGTWLN GFPPQSCGYATVTDAAEAVIVQVTPHHVLVDEYTGGEWVDSQFINGKCSNYICPTVHNSTTWHS DYKVKGLCDNLISMDITFFSEDGELSSLGKEGTGFRSNYFAYETGGKACKMQYCKHWGVR LPSGVWFEMADKDLFAAARFPECPEGSSISAPSQTSVDVSLIQDVERILDYSLCQETWSKIRAG LPISPVDLSYLAPKNPGTGPAFTIINGTLKYFETRYIRVDIAAPILSRMVGMISSGTTTERELWDD WAPYEDVEIGPNGVLRRTSSGYKFLYMIGHGMLDSDLHLSSKAQVFEHPHIQDAASQLPDDE SLFFGDTGLSKNPIELVEGWFWSSWKSSIASFFFIIGLIIGLFLVLRVGIHLCLIKLHTKKRQIYTDI EMNRLGK (SEQ ID NO: 315)

[0387] In some embodiments, the eVLPs described herein further comprise an inner encapsulation layer comprising components from viral capsids. These components include gag-pro polyproteins (*e.g.*, gag nucleocapsid proteins further comprising a viral protease linked thereto) and gag nucleocapsid proteins (*e.g.*, proteins that make up the core structural component of the inner shell of many viruses, lacking the protease of the gag-pro polyproteins) as described herein.

[0388] Gag-Pro polyproteins mediate proteolytic cleavage of Gag and Gag-Pol polyproteins or nucleocapsid proteins during or shortly after the release of a virion from the plasma membrane. In the eVLPs described herein, the protease of a gag-pro polyprotein is

responsible for cleaving a cleavable linker in the fusion protein to release a base editor following delivery of the BE-VLP to a target cell. In some embodiments, a gag-pro polyprotein is an MMLV gag-pro polyprotein or an FMLV gag-pro polyprotein.

[0389] The gag nucleocapsid proteins used in the eVLPs of the present disclosure may be an MMLV gag nucleocapsid protein, an FMLV gag nucleocapsid protein, or a nucleocapsid protein from any other virus that produces such proteins. In some embodiments, gag nucleocapsid proteins are fused to napDNABps (*e.g.*, as part of a base editor). In some embodiments, the fusion further comprises an NES as described herein. In certain embodiments, the gag nucleocapsid protein and the NES are located on one side of a cleavable linker as described herein, and the napDNABp or base editor is located on the other side of the cleavable linker, such that the base editor can be released from the gag nucleocapsid protein upon cleavage of the cleavable linker by the protease of the gag-pro polyprotein following delivery of the BE-VLP to a target cell.

[0390] Both the gag-pro polyprotein and the gag nucleocapsid protein form the inner encapsulation layer of the presently described eVLPs, as shown in FIG. 1. Any ratio of the gag-pro polyprotein to the gag nucleocapsid protein (*i.e.*, as part of the fusion proteins described herein) is contemplated in the eVLPs of the present disclosure. In some embodiments, the ratio of the gag-pro polyprotein to the fusion protein comprising a gag nucleocapsid protein is approximately 10:1, approximately 9:1, approximately 8:1, approximately 7:1, approximately 6:1, approximately 5:1, approximately 4:1, approximately 3:1, approximately 2:1, approximately 1.5:1, approximately 1:1, or approximately 0.5:1. In certain embodiments, the ratio is approximately 3:1.

Methods for Producing eVLPs

[0391] In one aspect, as exemplified in **FIG. 16**, the present disclosure relates to methods for producing the eVLPs described herein. In some embodiments, a method for producing the presently described eVLPs comprises transfecting, transducing, electroporating, or otherwise inserting into a producer cell one or more polynucleotides that together encode all the components of the eVLPs (*e.g.*, any of the pluralities of polynucleotides described herein, or any of the vectors described herein). In some embodiments, the polynucleotides which are transfected, transduced, electroporated, or otherwise inserted into a producer cell comprise: (i) a first polynucleotide comprising a nucleic acid sequence encoding a viral envelope glycoprotein; (ii) a second polynucleotide comprising a nucleic acid sequence encoding a

group-specific antigen (gag) protease (pro) polyprotein; (iii) a third polynucleotide comprising a nucleic acid sequence encoding a fusion protein, wherein the fusion protein comprises: (a) a group-specific antigen (gag) nucleocapsid protein; (b) a nucleic acid programmable DNA binding protein (napDNAbp); (c) a cleavable linker; and (d) a nuclear export sequence (NES); and (iv) a fourth polynucleotide comprising a nucleic acid sequence encoding a guide RNA (gRNA), wherein the gRNA binds to the napDNAbp of the fusion protein encoded by the third polynucleotide. In some embodiments, the present disclosure provides one or more vectors comprising one, two, three, or all four of the plurality of polynucleotides provided herein. In certain embodiments, each of the first, second, third, and fourth polynucleotides are on separate vectors. In certain embodiments, one or more of the first, second, third, and fourth polynucleotides are on the same vector.

[0392] In some embodiments, once the producer cell expresses the polynucleotides, the various components of the eVLPs self-assemble spontaneously within the producer cells. Assembly of the eVLPs relies on multimerization of the gag polyproteins encoded on the polynucleotides as described above. The gag polyproteins (some of which are fused to a gene editing agent, such as a Cas9 protein or a base editor) multimerize at the cell membrane of a producer cell and are subsequently released into the producer cell supernatant spontaneously. Thus, BE-eVLPs may be produced by transient transfection of producer cells (for example, Gesicle Producer 293T cells) as described in the Examples herein. All of the polynucleotides required for production of the eVLPs may be transfected into the producer cells simultaneously, or each polynucleotide needed may be transfected one at a time. In some embodiments, a single polynucleotide encodes all the components needed to produce the eVLPs described herein. Following transfection and incubation of the producer cells (*e.g.*, for about 2 hours, about 3 hours, about 4 hours, about 5 hours, about 6 hours, about 7 hours, about 8 hours, about 9 hours, about 10 hours, about 15 hours, about 24 hours, about 36 hours, about 48 hours, or more than 48 hours), producer cell supernatant may be harvested, and eVLPs may be purified therefrom.

[0393] Any cell capable of expressing a foreign polynucleotide may be used to produce the eVLPs described herein. For example, the present disclosure contemplates the use of any of the cells listed in the Kits and Cells section herein for production of the eVLPs, or any other cell known in the art capable of expressing a foreign polynucleotide.

[0394] Overview of an embodiment of the manufacture of eVLPs comprising BE RNPs (*e.g.*, BE-VLPs) in a producer cell using a set of expression plasmids which encode the various

self-assembling components of the eVLPs: (a) plasmid encoding a Gag-BE fusion protein (*e.g.*, a retroviral Gag, MMLV-Gag-BE fusion protein); (b) plasmid encoding a Gag-Pro-Pol protein (*e.g.*, a retroviral proteins, such as a MMLV protease precursor); (c) a plasmid encoding a BE sgRNA; and (d) a plasmid encoding an envelope glycoprotein (*e.g.*, the spike glycoprotein of the vesicular stomatitis virus (VSV-G)). The plasmids are transiently co-transfected into the producer cell and the encoded protein and sgRNA products are encoded. In some embodiments, such as the fourth-generation eVLPs described herein, the inventors found an optimized stoichiometry ratio of Gag-cargo fusion to Gag-Pro-Pol fusion protein which balances the amount of Gag-cargo available to be packaged into VLPs with the amount of retrovirus protease (the “Pro” in the Gag-Pro-Pol fusion) required for VLP maturation. In one embodiment, the optimized ratio of Gag-cargo fusion to Gag-Pro-Pol fusion protein is achieved by the appropriate ratio of plasmids encoding each component which are transiently delivered to the producer cells. In one embodiment, to modulate the stoichiometry of the Gag-cargo fusion to Gag-Pro-Pol fusion, the ratio of the plasmid encoding Gag-cargo (*e.g.*, Gag-3xNES-ABE8e) to wild-type MMLV gag-pro-pol plasmids transfected for VLP production was varied. It was found that increasing the amount of gag-cargo plasmid beyond the original proportion used for producing v3.4 BE-eVLPs (38% Gag-cargo plasmid and 62% gag-pro-pol plasmid) did not improve editing efficiencies (**FIG. 2G**). Decreasing the proportion of gag-cargo plasmid from 38% to 25% modestly improved editing efficiencies (**FIG. 2G**). However, further decreasing the proportion of gag-cargo plasmid below 25% reduced editing efficiencies (**FIG. 2G**). These results are consistent with a model in which an optimal gag-cargo:gag-pro-pol stoichiometry balances the amount of gag-cargo available to be packaged into VLPs with the amount of MMLV protease (the “pro” in gag-pro-pol) required for VLP maturation. In one embodiment, the results of this final round of optimization revealed a fourth-generation (v4) BE-eVLP formulation (**FIG. 2G**), which combines the optimal gag-BE:gag-pro-pol stoichiometry (25% gag-BE) with the v3.4 BE-eVLP architecture.

[0395] As depicted in **FIG. 16**, the present disclosure provides pluralities of polynucleotides encoding the eVLP (*e.g.*, BE-VLP) self-assembling component as described herein. In some embodiments, the present disclosure provides pluralities of polynucleotides comprising: (i) a first polynucleotide (*e.g.*, a plasmid) comprising a nucleic acid sequence encoding a viral envelope glycoprotein; (ii) a second polynucleotide (*e.g.*, a plasmid) comprising a nucleic acid sequence encoding a group-specific antigen (gag) protease (pro) polyprotein; (iii) a third

polynucleotide (*e.g.*, a plasmid) comprising a nucleic acid sequence encoding a fusion protein, wherein the fusion protein comprises: (a) a group-specific antigen (gag) nucleocapsid protein; (b) a nucleic acid programmable DNA binding protein (napDNAbp); (c) a cleavable linker; and (d) a nuclear export sequence (NES); and (iv) a fourth polynucleotide (*e.g.*, a plasmid) comprising a nucleic acid sequence encoding a guide RNA (gRNA). In some embodiments, the gRNA binds to the napDNAbp of the fusion protein encoded by the third polynucleotide. In some embodiments, the ratio of the second polynucleotide to the third polynucleotide is approximately 10:1, approximately 9:1, approximately 8:1, approximately 7:1, approximately 6:1, approximately 5:1, approximately 4:1, approximately 3:1, approximately 2:1, approximately 1.5:1, approximately 1:1, or approximately 0.5:1. In certain embodiments, the ratio of the second polynucleotide to the third polynucleotide is approximately 3:1.

Pharmaceutical compositions

[0396] Other aspects of the present disclosure relate to pharmaceutical compositions comprising any of the eVLPs, fusion proteins, and polynucleotides/pluralities of polynucleotides or vectors described herein. The term “pharmaceutical composition”, as used herein, refers to a composition formulated for pharmaceutical use. In some embodiments, the pharmaceutical composition further comprises a pharmaceutically acceptable carrier. In some embodiments, the pharmaceutical composition comprises additional agents (*e.g.*, for specific delivery, increasing half-life, or other therapeutic compounds).

[0397] As used here, the term “pharmaceutically-acceptable carrier” means a pharmaceutically-acceptable material, composition or vehicle, such as a liquid or solid filler, diluent, excipient, manufacturing aid (*e.g.*, lubricant, talc magnesium, calcium or zinc stearate, or steric acid), or solvent encapsulating material, involved in carrying or transporting the compound from one site (*e.g.*, the delivery site) of the body, to another site (*e.g.*, organ, tissue or portion of the body). A pharmaceutically acceptable carrier is “acceptable” in the sense of being compatible with the other ingredients of the formulation and not injurious to the tissue of the subject (*e.g.*, physiologically compatible, sterile, physiologic pH, etc.). Some examples of materials which can serve as pharmaceutically-acceptable carriers include: (1) sugars, such as lactose, glucose and sucrose; (2) starches, such as corn starch and potato starch; (3) cellulose, and its derivatives, such as sodium carboxymethyl cellulose, methylcellulose, ethyl cellulose, microcrystalline cellulose and cellulose acetate; (4) powdered tragacanth; (5) malt; (6) gelatin; (7) lubricating agents, such as magnesium

stearate, sodium lauryl sulfate and talc; (8) excipients, such as cocoa butter and suppository waxes; (9) oils, such as peanut oil, cottonseed oil, safflower oil, sesame oil, olive oil, corn oil and soybean oil; (10) glycols, such as propylene glycol; (11) polyols, such as glycerin, sorbitol, mannitol and polyethylene glycol (PEG); (12) esters, such as ethyl oleate and ethyl laurate; (13) agar; (14) buffering agents, such as magnesium hydroxide and aluminum hydroxide; (15) alginic acid; (16) pyrogen-free water; (17) isotonic saline; (18) Ringer's solution; (19) ethyl alcohol; (20) pH buffered solutions; (21) polyesters, polycarbonates and/or polyanhydrides; (22) bulking agents, such as polypeptides and amino acids; (23) serum component, such as serum albumin, HDL and LDL; (22) C2-C12 alcohols, such as ethanol; and (23) other non-toxic compatible substances employed in pharmaceutical formulations. Wetting agents, coloring agents, release agents, coating agents, sweetening agents, flavoring agents, perfuming agents, preservatives, and antioxidants can also be present in the formulation. The terms such as “excipient”, “carrier”, “pharmaceutically acceptable carrier” or the like are used interchangeably herein.

[0398] In some embodiments, the pharmaceutical composition is formulated for delivery to a subject, *e.g.*, for gene editing. Suitable routes of administering the pharmaceutical composition described herein include, without limitation: topical, subcutaneous, transdermal, intradermal, intralesional, intraarticular, intraperitoneal, intravesical, transmucosal, gingival, intradental, intracochlear, transtympanic, intraorgan, epidural, intrathecal, intramuscular, intravenous, intravascular, intraosseous, periocular, intratumoral, intracerebral, and intracerebroventricular administration.

[0399] In some embodiments, the pharmaceutical composition described herein is administered locally to a diseased site (*e.g.*, tumor site). In some embodiments, the pharmaceutical composition described herein is administered to a subject by injection, by means of a catheter, by means of a suppository, or by means of an implant, the implant being of a porous, non-porous, or gelatinous material, including a membrane, such as a sialastic membrane, or a fiber.

[0400] In other embodiments, the pharmaceutical composition described herein is delivered in a controlled release system. In one embodiment, a pump may be used (see, *e.g.*, Langer, 1990, *Science* 249:1527-1533; Sefton, 1989, *CRC Crit. Ref. Biomed. Eng.* 14:201; Buchwald *et al.*, 1980, *Surgery* 88:507; Saudek *et al.*, 1989, *N. Engl. J. Med.* 321:574). In another embodiment, polymeric materials can be used. (See, *e.g.*, *Medical Applications of Controlled Release* (Langer and Wise eds., CRC Press, Boca Raton, Fla., 1974); *Controlled Drug*

Bioavailability, Drug Product Design and Performance (Smolen and Ball eds., Wiley, New York, 1984); Ranger and Peppas, 1983, *Macromol. Sci. Rev. Macromol. Chem.* 23:61. See also Levy *et al.*, 1985, *Science* 228:190; During *et al.*, 1989, *Ann. Neurol.* 25:351; Howard *et al.*, 1989, *J. Neurosurg.* 71:105). Other controlled release systems are discussed, for example, in Langer, *supra*.

[0401] In some embodiments, the pharmaceutical composition is formulated in accordance with routine procedures as a composition adapted for intravenous or subcutaneous administration to a subject, *e.g.*, a human. In some embodiments, pharmaceutical compositions for administration by injection are solutions in sterile isotonic aqueous buffer. Where necessary, the pharmaceutical composition can also include a solubilizing agent and a local anesthetic such as lignocaine to ease pain at the site of the injection. Generally, the ingredients are supplied either separately or mixed together in unit dosage form, for example, as a dry lyophilized powder or water free concentrate in a hermetically sealed container such as an ampoule or sachette indicating the quantity of active agent. Where the pharmaceutical composition is to be administered by infusion, it can be dispensed with an infusion bottle containing sterile pharmaceutical grade water or saline. Where the pharmaceutical composition is administered by injection, an ampoule of sterile water for injection or saline can be provided so that the ingredients can be mixed prior to administration.

[0402] A pharmaceutical composition for systemic administration may be a liquid, *e.g.*, sterile saline, lactated Ringer's or Hank's solution. In addition, the pharmaceutical composition can be in solid forms and re-dissolved or suspended immediately prior to use. Lyophilized forms are also contemplated.

[0403] The pharmaceutical composition can be contained within a lipid particle or vesicle, such as a liposome or microcrystal, which is also suitable for parenteral administration. The particles can be of any suitable structure, such as unilamellar or plurilamellar, so long as compositions are contained therein. Compounds can be entrapped in "stabilized plasmid-lipid particles" (SPLP) containing the fusogenic lipid dioleoylphosphatidylethanolamine (DOPE), low levels (5-10 mol%) of cationic lipid and stabilized by a polyethyleneglycol (PEG) coating (Zhang Y. P. *et al.*, *Gene Ther.* 1999, 6:1438-47). Positively charged lipids such as N-[1-(2,3-dioleoyloxy)propyl]-N,N,N-trimethyl-ammoniummethylsulfate, or "DOTAP," are particularly preferred for such particles and vesicles. The preparation of such lipid particles is well known. See, *e.g.*, U.S. Patent Nos. 4,880,635; 4,906,477; 4,911,928; 4,917,951; 4,920,016; and 4,921,757; each of which is incorporated herein by reference.

[0404] The pharmaceutical compositions described herein may be administered or packaged as a unit dose, for example. The term “unit dose” when used in reference to a pharmaceutical composition of the present disclosure refers to physically discrete units suitable as unitary dosage for the subject, each unit containing a predetermined quantity of active material calculated to produce the desired therapeutic effect in association with the required diluent; *i.e.*, carrier, or vehicle.

[0405] Further, the pharmaceutical composition can be provided as a pharmaceutical kit comprising (a) a container containing a compound of the invention in lyophilized form and (b) a second container containing a pharmaceutically acceptable diluent (*e.g.*, sterile water) for injection. The pharmaceutically acceptable diluent can be used for reconstitution or dilution of the lyophilized compound of the invention. Optionally associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use, or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use, or sale for human administration.

[0406] In another aspect, an article of manufacture containing materials useful for the treatment of the diseases described above is included. In some embodiments, the article of manufacture comprises a container and a label. Suitable containers include, for example, bottles, vials, syringes, and test tubes. The containers may be formed from a variety of materials such as glass or plastic. In some embodiments, the container holds a composition that is effective for treating a disease and may have a sterile access port. For example, the container may be an intravenous solution bag or a vial having a stopper pierce-able by a hypodermic injection needle. The active agent in the composition is a compound of the invention. In some embodiments, the label on or associated with the container indicates that the composition is used for treating the disease of choice. The article of manufacture may further comprise a second container comprising a pharmaceutically-acceptable buffer, such as phosphate-buffered saline, Ringer's solution, or dextrose solution. It may further include other materials desirable from a commercial and user standpoint, including other buffers, diluents, filters, needles, syringes, and package inserts with instructions for use.

Kits and cells

[0407] The fusion proteins, eVLPs, and compositions of the present disclosure may be assembled into kits. In some embodiments, the kit comprises polynucleotides for expression and assembly of the eVLPs described herein. In other embodiments, the kit further comprises appropriate guide nucleotide sequences or nucleic acid vectors for the expression of such

guide nucleotide sequences, to target the Cas9 protein of the base editors being delivered by the eVLPs to the desired target sequence.

[0408] The kit described herein may include one or more containers housing components for performing the methods described herein, and optionally instructions for use. Any of the kits described herein may further comprise components needed for performing the base editing methods described herein. Each component of the kits, where applicable, may be provided in liquid form (*e.g.*, in solution) or in solid form, (*e.g.*, a dry powder). In certain cases, some of the components may be reconstitutable or otherwise processible (*e.g.*, to an active form), for example, by the addition of a suitable solvent or other species (for example, water), which may or may not be provided with the kit.

[0409] In some embodiments, the kits may optionally include instructions and/or promotion for use of the components provided. As used herein, “instructions” can define a component of instruction and/or promotion, and typically involve written instructions on or associated with packaging of the disclosure. Instructions also can include any oral or electronic instructions provided in any manner such that a user will clearly recognize that the instructions are to be associated with the kit, for example, audiovisual (*e.g.*, videotape, DVD, *etc.*), Internet, and/or web-based communications, *etc.* The written instructions may be in a form prescribed by a governmental agency regulating the manufacture, use, or sale of pharmaceuticals or biological products, which can also reflect approval by the agency of manufacture, use or sale for animal administration. As used herein, “promoted” includes all methods of doing business including methods of education, hospital and other clinical instruction, scientific inquiry, drug discovery or development, academic research, pharmaceutical industry activity including pharmaceutical sales, and any advertising or other promotional activity including written, oral and electronic communication of any form, associated with the disclosure. Additionally, the kits may include other components depending on the specific application, as described herein.

[0410] The kits may contain any one or more of the components described herein in one or more containers. The components may be prepared sterilely, packaged in a syringe, and shipped refrigerated. Alternatively, they may be housed in a vial or other container for storage. A second container may have other components prepared sterilely. Alternatively, the kits may include the active agents premixed and shipped in a vial, tube, or other container.

[0411] The kits may have a variety of forms, such as a blister pouch, a shrink-wrapped pouch, a vacuum sealable pouch, a sealable thermoformed tray, or a similar pouch or tray

form, with the accessories loosely packed within the pouch, one or more tubes, containers, a box, or a bag. The kits may be sterilized after the accessories are added, thereby allowing the individual accessories in the container to be otherwise unwrapped. The kits can be sterilized using any appropriate sterilization techniques, such as radiation sterilization, heat sterilization, or other sterilization methods known in the art. The kits may also include other components, depending on the specific application, for example, containers, cell media, salts, buffers, reagents, syringes, needles, a fabric, such as gauze, for applying or removing a disinfecting agent, disposable gloves, a support for the agents prior to administration, *etc.* Some aspects of this disclosure provide kits comprising a nucleic acid construct comprising a nucleotide sequence encoding the various components of the eVLPs described herein (*e.g.*, including, but not limited to, the napDNABps, deaminase domains, gag proteins, gRNAs, and viral envelope glycoproteins. In some embodiments, the nucleotide sequence(s) comprises a heterologous promoter (or more than a single promoter) that drives expression of the BE-VLP system components.

[0412] Other aspects of this disclosure provide kits comprising one or more nucleic acid constructs encoding the various components of the BE-VLP system described herein, *e.g.*, a nucleotide sequence encoding the components of the BE-VLP system capable of delivering a base editor to a target cell. In some embodiments, the nucleotide sequence comprises a heterologous promoter that drives expression of the BE-VLP system components.

[0413] Cells that may contain any of the eVLPs, fusion proteins, and compositions described herein include prokaryotic cells and eukaryotic cells. In various aspects relating to the production of eVLPs, the disclosure provides for any suitable cells for use as a VLP-producer cell line, *i.e.*, the cell line that in various embodiments becomes transiently transformed with the plasmids encoding the protein and nucleic acid components of the eVLPs. In various other aspects relating to applications of eVLPs, the disclosure provides for any suitable target or recipient cells, *e.g.*, a diseased cell or tissue in a subject in need of treatment by way of base editing as delivered by a BE-VLP. The methods described herein may be used to deliver a base into a eukaryotic cell (*e.g.*, a mammalian cell, such as a human cell). In some embodiments, the cell is *in vitro* (*e.g.*, cultured cell). In some embodiments, the cell is *in vivo* (*e.g.*, in a subject such as a human subject). In some embodiments, the cell is *ex vivo* (*e.g.*, isolated from a subject and may be administered back to the same or a different subject).

[0414] Typically, the eukaryotic cell is a mammalian cell, such as a human cell, a chicken cell or an insect cell. Examples of suitable mammalian cells are, but are not limited to HEK-

293T cells, COS7 cells, Hela cells and HEK-293 cells. Examples of suitable insect cells include, but are not limited to, High5 cells and Sf9 cells. In some embodiment, insect cells as they are devoid of undesirable human protein, and their culture does not require animal serum.

[0415] Mammalian cells of the present disclosure include human cells, primate cells (*e.g.*, vero cells), rat cells (*e.g.*, GH3 cells, OC23 cells) or mouse cells (*e.g.*, MC3T3 cells). There are a variety of human cell lines, including, without limitation, human embryonic kidney (HEK) cells, HeLa cells, cancer cells from the National Cancer Institute's 60 cancer cell lines (NCI60), DU145 (prostate cancer) cells, Lncap (prostate cancer) cells, MCF-7 (breast cancer) cells, MDA-MB-438 (breast cancer) cells, PC3 (prostate cancer) cells, T47D (breast cancer) cells, THP-1 (acute myeloid leukemia) cells, U87 (glioblastoma) cells, SHSY5Y human neuroblastoma cells (cloned from a myeloma) and Saos-2 (bone cancer) cells. In some embodiments, eVLPs are delivered into human embryonic kidney (HEK) cells (*e.g.*, HEK 293 or HEK 293T cells). In some embodiments, eVLPs are delivered into stem cells (*e.g.*, human stem cells) such as, for example, pluripotent stem cells (*e.g.*, human pluripotent stem cells including human induced pluripotent stem cells (hiPSCs)). A stem cell refers to a cell with the ability to divide for indefinite periods in culture and to give rise to specialized cells. A pluripotent stem cell refers to a type of stem cell that is capable of differentiating into all tissues of an organism, but not alone capable of sustaining full organismal development. A human induced pluripotent stem cell refers to a somatic (*e.g.*, mature or adult) cell that has been reprogrammed to an embryonic stem cell-like state by being forced to express genes and factors important for maintaining the defining properties of embryonic stem cells (see, *e.g.*, Takahashi and Yamanaka, *Cell* 126 (4): 663–76, 2006, incorporated by reference herein). Human induced pluripotent stem cell cells express stem cell markers and are capable of generating cells characteristic of all three germ layers (ectoderm, endoderm, mesoderm).

[0416] Additional non-limiting examples of cell lines that may be used in accordance with the present disclosure include 293-T, 293-T, 3T3, 4T1, 721, 9L, A-549, A172, A20, A253, A2780, A2780ADR, A2780cis, A431, ALC, B16, B35, BCP-1, BEAS-2B, bEnd.3, BHK-21, BR 293, BxPC3, C2C12, C3H-10T1/2, C6, C6/36, Cal-27, CGR8, CHO, CML T1, CMT, COR-L23, COR-L23/5010, COR-L23/CPR, COR-L23/R23, COS-7, COV-434, CT26, D17, DH82, DU145, DuCaP, E14Tg2a, EL4, EM2, EM3, EMT6/AR1, EMT6/AR10.0, FM3, H1299, H69, HB54, HB55, HCA2, Hepa1c1c7, High Five cells, HL-60, HMEC, HT-29, HUVEC, J558L cells, Jurkat, JY cells, K562 cells, KCL22, KG1, Ku812, KYO1, LNCap,

Ma-Mel 1, 2, 3....48, MC-38, MCF-10A, MCF-7, MDA-MB-231, MDA-MB-435, MDA-MB-468, MDCK II, MG63, MONO-MAC 6, MOR/0.2R, MRC5, MTD-1A, MyEnd, NALM-1, NCI-H69/CPR, NCI-H69/LX10, NCI-H69/LX20, NCI-H69/LX4, NIH-3T3, NW-145, OPCN/OPCT Peer, PNT-1A/PNT 2, PTK2, Raji, RBL cells, RenCa, RIN-5F, RMA/RMAS, S2, Saos-2 cells, Sf21, Sf9, SiHa, SKBR3, SKOV-3, T-47D, T2, T84, THP1, U373, U87, U937, VCaP, WM39, WT-49, X63, YAC-1, and YAR cells.

[0417] Some aspects of this disclosure provide cells comprising any of the constructs disclosed herein. In some embodiments, a host cell is transiently or non-transiently transfected with one or more vectors described herein. In some embodiments, a cell is transfected as it naturally occurs in a subject. In some embodiments, a cell that is transfected is taken from a subject. In some embodiments, the cell is derived from cells taken from a subject, such as a cell line. A wide variety of cell lines for tissue culture are known in the art. Examples of cell lines include, but are not limited to, C8161, CCRF-CEM, MOLT, mIMCD-3, NHDF, HeLa-S3, Huh1, Huh4, Huh7, HUVEC, HASMC, HEK_n, HEK_a, MiaPaCell, Panc1, PC-3, TF1, CTLL-2, C1R, Rat6, CV1, RPTE, A10, T24, J82, A375, ARH-77, Calu1, SW480, SW620, SKOV3, SK-UT, CaCo2, P388D1, SEM-K2, WEHI-231, HB56, TIB55, Jurkat, J45.01, LRMB, Bcl-1, BC-3, IC21, DLD2, Raw264.7, NRK, NRK-52E, MRC5, MEF, Hep G2, HeLa B, HeLa T4, COS, COS-1, COS-6, COS-M6A, BS-C-1 monkey kidney epithelial, BALB/3T3 mouse embryo fibroblast, 3T3 Swiss, 3T3-L1, 132-d5 human fetal fibroblasts; 10.1 mouse fibroblasts, 293-T, 3T3, 721, 9L, A2780, A2780ADR, A2780cis, A172, A20, A253, A431, A-549, ALC, B16, B35, BCP-1 cells, BEAS-2B, bEnd.3, BHK-21, BR 293, BxPC3, C3H-10T1/2, C6/36, Cal-27, CHO, CHO-7, CHO-IR, CHO-K1, CHO-K2, CHO-T, CHO Dhfr ^{-/-}, COR-L23, COR-L23/CPR, COR-L23/5010, COR-L23/R23, COS-7, COV-434, CML T1, CMT, CT26, D17, DH82, DU145, DuCaP, EL4, EM2, EM3, EMT6/AR1, EMT6/AR10.0, FM3, H1299, H69, HB54, HB55, HCA2, HEK-293, HeLa, Hepa1c1c7, HL-60, HMEC, HT-29, Jurkat, JY cells, K562 cells, Ku812, KCL22, KG1, KYO1, LNCap, Ma-Mel 1-48, MC-38, MCF-7, MCF-10A, MDA-MB-231, MDA-MB-468, MDA-MB-435, MDCK II, MDCK 11, MOR/0.2R, MONO-MAC 6, MTD-1A, MyEnd, NCI-H69/CPR, NCI-H69/LX10, NCI-H69/LX20, NCI-H69/LX4, NIH-3T3, NALM-1, NW-145, OPCN/OPCT cell lines, Peer, PNT-1A/PNT 2, RenCa, RIN-5F, RMA/RMAS, Saos-2 cells, Sf-9, SkBr3, T2, T-47D, T84, THP1 cell line, U373, U87, U937, VCaP, Vero cells, WM39, WT-49, X63, YAC-1, YAR, and transgenic varieties thereof.

[0418] Cell lines are available from a variety of sources known to those with skill in the art (see, *e.g.*, the American Type Culture Collection (ATCC) (Manassus, Va.)). In some embodiments, a cell transfected with one or more vectors described herein is used to establish a new cell line comprising one or more vector-derived sequences. In some embodiments, a cell transiently transfected with the components of a CRISPR system as described herein (such as by transient transfection of one or more vectors, or transfection with RNA), and modified through the activity of a CRISPR complex, is used to establish a new cell line comprising cells containing the modification but lacking any other exogenous sequence. In some embodiments, cells transiently or non-transiently transfected with one or more vectors described herein, or cell lines derived from such cells, are used in assessing one or more test compounds.

EXAMPLES

Example 1. Therapeutic in vivo base editing with minimal off-target activity using engineered DNA-free virus-like particles (eVLPs)

[0419] Base editors (BEs) enable the therapeutic correction of pathogenic point mutations in the genomic DNA of living organisms. While various strategies have been used to deliver BEs *in vivo*, a method that delivers BE ribonucleoproteins (RNPs) into tissues in animals would offer important safety advantages over existing approaches that deliver DNA or mRNA. The extensive engineering and application of engineered VLPs (eVLPs, also referred to herein as BE-VLPs), virus-like particles that efficiently package and deliver BE or Cas9 RNPs without DNA delivery or the possibility of unwanted DNA integration, is reported herein. By iteratively engineering VLP architectures to overcome cargo packaging, release, and localization bottlenecks, optimized fourth-generation eVLPs were generated that mediate efficient on-target base editing *in vitro* across a variety of cell types and endogenous genomic loci with minimal detected off-target editing, as well as 4.7-fold higher editing following Cas9 nuclease delivery compared with first-generation VLPs. Using different glycoproteins in eVLPs alters their cellular tropism. Optimized eVLPs also supported *in vivo* base editing in multiple organs following single injections into mice, resulting in 26-fold higher editing efficiency in the liver than a previously described VLP architecture and 78% knockdown of serum Pcsk9 levels, as well as partial restoration of visual function in a mouse model of genetic blindness. Frequencies of off-target editing following treatment with eVLPs were substantially lower both in cultured cells and *in vivo* than base editor delivery with plasmid DNA or AAV. eVLPs do not affect cell viability or induce detected liver pathology. Cell-type

tropism of eVLPs can be controlled by pseudotyping with different envelope glycoproteins. These results establish eVLPs as a promising method for therapeutic base editing *in vivo* that minimizes risks of off-target editing or DNA integration.

[0420] Virus-like particles (VLPs), assemblies of viral proteins that can infect cells but lack viral genetic material, have emerged as potentially promising vehicles for delivering gene editing agents as RNPs (Campbell *et al.*, 2019; Choi *et al.*, 2016; Gee *et al.*, 2020; Hamilton *et al.*, 2021; Indikova and Indik, 2020; Lyu *et al.*, 2019; Lyu *et al.*, 2021; Mangeot *et al.*, 2019; Yao *et al.*, 2021). VLPs that deliver RNP cargos exploit the efficiency and tissue targeting advantages of viral delivery but avoid the risks associated with viral genome integration and prolonged expression of the editing agent. However, existing VLP-mediated strategies for delivering gene editing agent RNPs thus far support low to moderate editing efficiencies or limited validation of their therapeutic efficacy *in vivo* (Campbell *et al.*, 2019; Choi *et al.*, 2016; Gee *et al.*, 2020; Hamilton *et al.*, 2021; Indikova and Indik, 2020; Lyu *et al.*, 2019; Lyu *et al.*, 2021; Mangeot *et al.*, 2019; Yao *et al.*, 2021). Indeed, therapeutic levels of post-natal *in vivo* gene editing using RNP-packaging VLPs have not been previously reported.

[0421] Described herein is the development and application of eVLPs, an engineered VLP platform for packaging and delivering therapeutic RNPs, including Cas9 nuclease and base editors, *in vitro* and *in vivo* that offers key advantages of both viral and non-viral delivery strategies. Extensive VLP architecture engineering yielded fourth-generation eVLPs that package an average of 16-fold more BE RNP compared to initial designs that were based on previously reported VLPs (Mangeot *et al.*, 2019). These eVLPs enable highly efficient base editing with minimal off-target editing in a variety of cell types, including multiple immortalized cell lines, primary human and mouse fibroblasts, and primary human T cells, as well as 4.7-fold improved Cas9 nuclease-mediated indel formation compared with a previously reported Cas9-VLP. Single *in vivo* injections of eVLPs into mice mediated efficient base editing of various target genes in multiple organs, strongly knocked down serum Pcsk9 levels, and partially restored visual function in a mouse model of genetic blindness. These results establish eVLPs as a useful platform for transiently delivering gene editing agents (e.g., BEs) *in vivo* with therapeutically relevant efficiencies and minimized risk of off-target editing or DNA integration, and the eVLPs described herein may similarly improve the *in vivo* delivery of other proteins and RNPs.

Results

A retroviral scaffold supports efficient base editor VLPs

[0422] It was hypothesized that retroviruses would be an attractive scaffold for engineering base editor VLPs (BE-VLPs, aka “eVLPs”). Retroviral capsids generally lack the rigid symmetry requirements of many non-enveloped icosahedral viruses (Zhang *et al.*, 2015), suggesting increased structural flexibility to incorporate non-native protein cargos. Additionally, retrovirus tropisms can be readily modulated by pseudotyping virions with different envelope glycoproteins, which could enable targeting of eVLPs to specific cell types (Cronin *et al.*, 2005). Previous work has demonstrated that fusing a desired protein cargo to the C-terminus of retroviral gag polyproteins is sufficient to direct packaging of that cargo protein within retroviral particles (Kaczmarczyk *et al.*, 2011; Voelkel *et al.*, 2010). More recently, similar strategies have been applied to package Cas9 RNPs within retroviral particles (Hamilton *et al.*, 2021; Mangeot *et al.*, 2019). Therefore, whether retroviral scaffolds could support efficient BE-VLP formation in a manner that preserves BE activity was investigated.

[0423] As an initial (v1) BE-VLP design, ABE8e, a highly active adenine base editor (Richter *et al.*, 2020), was fused to the C-terminus of the Friend murine leukemia virus (FMLV) gag polyprotein via a linker peptide that would be cleaved by the FMLV protease upon particle maturation (**FIG. 1A**). FMLV-based VLPs were previously used successfully to package and deliver Cas9 RNPs (Mangeot *et al.*, 2019). eVLPs were produced by transfecting Gesicle 293T producer cells with plasmids expressing this FMLV gag–ABE8e fusion construct, wild-type FMLV gag-pro-pol polyprotein, the VSV-G envelope glycoprotein, and an sgRNA targeting HEK293T cell genomic site 2 or site 3, hereafter referred to as *HEK2* or *HEK3*.

[0424] After harvesting eVLPs from producer cell supernatant, HEK293T cells were transduced *in vitro* with concentrated eVLPs. Encouragingly, v1 eVLPs robustly edited the *HEK2* and *HEK3* genomic loci with efficiencies >97% at the highest doses in unsorted cells (**FIG. 1B**). It was confirmed via immunoblotting that these eVLPs contained Cas9, the MLV capsid, and VSV-G proteins (**FIG. 8A**). These observations indicated that the FMLV retroviral scaffold supports BE-VLP formation and that v1 eVLPs can efficiently transduce and edit HEK293T cells *in vitro*.

Improving cargo release after VLP maturation

[0425] While v1 eVLPs robustly edited the *HEK2* and *HEK3* loci in HEK293T cells, these commonly used test loci are especially amenable to gene editing and lack therapeutic relevance (Anzalone *et al.*, 2020). To begin to evaluate the therapeutic potential of eVLPs, their ability to install mutations in the *BCL11A* erythroid-specific enhancer that upregulate the expression of fetal hemoglobin in erythrocytes, an established base editing strategy for the treatment of β -hemoglobinopathies (Richter *et al.*, 2020; Zeng *et al.*, 2020), was assessed. It was observed that v1 eVLPs achieved 73% editing efficiency at the *BCL11A* enhancer locus in HEK293T cells at high doses, but editing levels dropped steeply with decreasing doses (FIG. 8B). These results indicated that v1 BE-VLP activity could be improved.

[0426] Cleavage of the gag–ABE8e linker by the MLV protease after particle maturation is required to liberate free ABE8e RNP. It was reasoned that linker cleavage efficiency might bottleneck BE-VLP editing (FIG. 2A). To test this hypothesis, a series of second-generation (v2) engineered BE-eVLPs were constructed that contain a variety of protease-cleavable linker sequences between the MLV gag and ABE8e (FIG. 8C). First, the retroviral scaffold was switched from Friend MLV to Moloney MLV (MMLV), a similar MLV strain whose protease substrate specificity has been extensively characterized (Feher *et al.*, 2006). Four different linker sequences were then screened that were known to be cleaved with varying efficiencies by the MMLV protease, and several new gag–ABE8e linkers that improved editing efficiencies compared to v1 eVLPs were identified (FIG. 2B). Specifically, v2.4 BE-eVLPs exhibited 1.2–1.5-fold higher editing efficiencies at all doses tested relative to v1 eVLPs (FIG. 2B). To investigate the cleavage efficiencies of the linker sequences in v2.1–v2.4 BE-eVLPs, western blots were performed to determine the fraction of cleaved ABE8e versus full-length gag–ABE8e present in purified eVLPs. This analysis revealed that the v2.4 linker is cleaved more efficiently than the v2.1 and v2.2 linkers, but less efficiently than the v2.3 linker (FIGS. 8D–8E).

[0427] These findings support a model in which the linker sequence in v2.4 BE-eVLPs is cleaved at an optimal rate that supports efficient release of ABE8e RNP after VLP maturation but precludes premature release of ABE8e RNP prior to its incorporation into VLPs. These findings demonstrate that the gag–cargo protein linker sequence is an important parameter of VLP architectures and that optimizing this sequence to balance the linker cleavage kinetics between these two constraints can improve eVLP activity.

Improving cargo localization and loading into eVLPs

[0428] Previously optimized BEs are fused at their N- and C-termini to bipartite nuclear localization signals (NLSs), which promotes nuclear import of BEs and enhances their access to genomic DNA (Koblan *et al.*, 2018). However, gag–BE fusions must be localized to the cytoplasm and outer membrane of producer cells in order to be incorporated into VLPs as they form (**FIGS. 2C**). The presence of two NLSs within the gag–BE fusion may hamper gag–BE localization to the outer membrane and impede BE incorporation into VLPs.

[0429] To encourage cytosolic gag–cargo localization in producer cells, third-generation (v3) eVLP architectures that contain nuclear export signals (NESs) in addition to NLSs were designed. Previous work demonstrated that MLV-based VLPs can tolerate the addition of NESs at multiple locations within the gag protein (Wu and Roth, 2014). In the v3 designs, MMLV protease-cleavable linker sequences were placed at locations next to NESs to ensure that the NESs would be cleaved from the cargo following VLP maturation (**FIGS. 2D** and **9B**), thereby liberating NLS-flanked cargo proteins that could be efficiently imported into the nucleus of the transduced cells.

[0430] All v3 BE-eVLP architectures contained the optimal gag–ABE8e linker sequence from v2.4 BE-eVLPs. BE-eVLPs v3.1, v3.2, and v3.3 harbor a 3xNES motif fused at the C-terminus of ABE8e via an additional MMLV protease-cleavable linker and exhibited comparable or lower efficiencies relative to v2.4 BE-eVLPs (**FIG. 2E**). However, v3.4 BE-eVLPs, which contain a 3xNES motif at the C-terminus of MMLV gag immediately before the v2.4 optimized cleavable linker sequence, exhibited 1.1–2.1-fold improvements in editing efficiencies at the *BCL11A* enhancer locus at all doses tested relative to v2.4 BE-eVLPs (**FIG. 2E**). Notably, v3.4 BE-eVLPs require only a single viral protease cleavage event to liberate NLS-flanked, NES-free BEs (**FIGS. 2D** and **9B**), compared to the two distinct cleavage events required in v3.1, v3.2, and v3.3 BE-eVLPs, which might explain their superior efficiency. To further investigate the effect of NES addition on gag–ABE localization, immunofluorescence microscopy of producer cells transfected with the v3.4 gag–3xNES–ABE construct or the v2.4 gag–ABE construct was performed. This analysis revealed a 1.3-fold increase in cytoplasmic localization of ABE protein detected in v3.4-transfected producer cells relative to v2.4-transfected producer cells (**FIGS. 10C** and **10D**). These results demonstrate that BE-eVLP activity can be improved by promoting the extranuclear localization of the gag–BE fusion in producer cells while maintaining the nuclear localization of the BEs released into transduced cells.

Improving component stoichiometry of eVLPs

[0431] Finally, the gag–cargo:gag–pro–pol stoichiometry of v3.4 eVLPs was optimized. It was hypothesized that an optimal gag–cargo:gag–pro–pol stoichiometry would balance the amount of gag–cargo available to be packaged into VLPs with the amount of MMLV protease (“pro” in gag–pro–pol) required for VLP maturation (**FIG. 2F**). To modulate this stoichiometry, the ratio of gag–3xNES–ABE8e to wild-type MMLV gag–pro–pol plasmids transfected for VLP production was varied. It was found that increasing the amount of gag–BE plasmid beyond the original proportion used for producing v3.4 BE-eVLPs (38% gag–BE plasmid and 62% gag–pro–pol plasmid) did not improve editing efficiencies (**FIG. 2G**). Decreasing the proportion of gag–BE plasmid from 38% to 25% modestly improved editing efficiencies (**FIG. 2G**). However, further decreasing the proportion of gag–BE plasmid below 25% reduced editing efficiencies (**FIG. 2G**). These results are consistent with a model in which an optimal gag–BE:gag–pro–pol stoichiometry balances the amount of gag–BE available to be packaged into VLPs with the amount of MMLV protease (the “pro” in gag–pro–pol) required for VLP maturation.

[0432] The results of this final round of optimization revealed a fourth-generation (v4) BE-eVLP formulation (**FIG. 2G**), which combines the optimal gag–BE:gag–pro–pol stoichiometry (25% gag–BE) with the v3.4 BE-eVLP architecture. The v4 BE-eVLPs were visualized by transmission electron microscopy, and their spherical morphology and approximate particle diameter of 100-150 nm was confirmed (**FIG. 10A**).

[0433] Next, the effects of this architecture engineering on the protein content of BE-eVLPs was determined. Anti-Cas9 and anti-MLV(p30) ELISAs were performed to quantify the number of BE molecules and p30 (MLV capsid) molecules present in v1 through v4 BE-eVLPs (**FIG. 10B-10C**). These experiments revealed that v2.4, v3.4, and v4 BE-eVLPs contain 1.8-, 19.2-, and 11-fold more BE cargo protein molecules per particle respectively compared to v1 eVLPs (**FIG. 3A**). This increase in BE protein content per particle correlates with an increase in the relative amount of sgRNAs per particle as measured by targeted RT-qPCR of lysed VLPs (**FIG. 3B**). Interestingly, v4 BE-eVLPs contain fewer BE protein molecules per particle than v3.4 BE-eVLPs but the same amount of sgRNA molecules, which suggests that v3.4 and v4 BE-eVLPs may contain similar amounts of active BE RNPs per particle. Additionally, v4 BE-eVLPs are produced at higher titer than v3.4 BE-eVLPs (**FIG. 10C**).

[0434] These results support a model in which increasing the number of active BE RNP molecules per particle can improve BE-eVLP editing efficiencies. However, increasing the number of BE molecules per particle beyond a certain threshold can be harmful, since these additional BE molecules do not appear to be complexed with sgRNAs, and there is an apparent trade-off between the number of cargo molecules incorporated per VLP and overall VLP titers. Together, these results reveal additional important parameters that influence eVLP efficiencies and demonstrate how these parameters can be improved by modulating gag–cargo localization and gag–BE:gag-pro-pol stoichiometry.

v4 eVLPs support potent, high-efficiency gene editing

[0435] The successive VLP engineering efforts described above substantially improved editing efficiencies of v4 BE-eVLPs at the *BCL11A* enhancer locus in HEK293T cells to 95% at the maximal dose (**FIG. 3C**). v4 BE-eVLPs exhibit a 5.6-fold improvement in editing efficiency per unit volume compared to v1 eVLPs and a 2.2-fold improvement compared to v2.4 BE-eVLPs (**FIG. 3C**). It was also observed that v4 BE-eVLPs exhibit 8.5-fold improvements in base editing activity per viral particle in HEK293T cells (**FIG. 10D**). To confirm that v4 VLP engineering supported general base editing improvements that were not restricted to one particular genomic locus or target cell line, v1, v2.4, v3.4, and v4 BE-eVLPs targeting the *Dnmt1* locus in 3T3 mouse fibroblasts were tested. A very similar trend in the editing efficiencies of the four eVLP architectures was observed with an 8.6-fold improvement in editing efficiency per unit volume of v4 eVLPs compared to v1 eVLPs in 3T3 cells (**FIG. 3D**). Additionally, treatment with v4 eVLPs had no negative impact on the viability of HEK293T or 3T3 cells (**FIG. 10E**). v4 BE-eVLPs also supported robust multiplex editing of the *BCL11A* enhancer and *HEK2* genomic loci in HEK293T cells (**FIG. 3E**). These results show that v4 eVLPs mediate high-efficiency base editing while being minimally perturbative to the treated cells.

[0436] It was hypothesized that the engineered v4 eVLP architecture might similarly improve VLP-mediated delivery of other proteins in addition to base editors. To test this possibility, v1 and v4 VLPs were constructed that packaged Cas9 nuclease (Cas9-VLPs) and an sgRNA targeting the *EMX1* genomic locus. A 4.7-fold improvement in indel frequencies per unit volume generated by v4 Cas9-eVLPs compared to v1 Cas9-VLPs in HEK293T cells (**FIG. 10F**) was observed. This observation suggests that the optimized v4 eVLP architecture offers

generalizable improvements to VLP-mediated delivery of proteins that are not limited to base editors.

[0437] An attractive feature of eVLPs is that their cellular tropism in principle can be modulated by producing them with different envelope glycoproteins. A similar strategy was used previously to modulate the tropism of Cas9-VLPs (Hamilton *et al.*, 2021). To investigate whether eVLPs can be programmed to target certain cell types, we produced v4 eVLPs pseudotyped with the FuG-B2 envelope glycoprotein (Kato *et al.*, 2011). FuG-B2 is an engineered envelope glycoprotein that contains the extracellular and transmembrane domains of the rabies virus envelope glycoprotein and the cytoplasmic domain of VSV-G, and can be used to pseudotype lentiviral vectors for neuron-specific transduction (Kato *et al.*, 2011). Indeed, it was observed that FuG-B2-pseudotyped v4 BE-eVLPs efficiently transduce and edit Neuro-2a cells (a mouse neuroblastoma cell line) but not mouse 3T3 fibroblasts (**FIGS. 3F** and **10G**). These results validate that the tissue specificity of eVLPs can be targeted by swapping in other glycoproteins such as those used to pseudotype lentiviruses to transduce specific cell populations.

[0438] Collectively, these findings identify factors that influence VLP activity, and demonstrate that extensively engineering the protease-cleavable linker sequence, gag–cargo localization, and gag–cargo:gag-pro-pol stoichiometry can overcome bottlenecks that limit VLP potency. These results also reveal novel insights into the factors that influence VLP activity and establish v4 BE-eVLPs as a robust method for delivering BE RNPs in cultured cells.

v4 BE-eVLPs show minimal off-target editing or DNA integration

[0439] Given that v4 BE-eVLPs exhibit robust on-target base editing at several endogenous genomic loci in multiple cell types, their off-target editing profiles were next assessed. BEs can mediate Cas-dependent off-target editing at a subset of Cas9 off-target binding sites, as well as Cas-independent off-target editing at a low level throughout the genome (Anzalone *et al.*, 2020). To evaluate Cas-dependent off-target editing by v4 BE-eVLPs relative to ABE8e plasmid transfection in HEK293T cells, targeted amplicon sequencing of known Cas9 off-target sites associated with three different sgRNAs targeting the *HEK2*, *HEK3*, and *BCL11A* enhancer loci was performed. It was observed that v4 eVLPs exhibited comparable or higher on-target editing efficiency from v4 BE-eVLPs compared to plasmid transfection at these three genomic loci, but 12- to 900-fold lower Cas-dependent off-target editing compared to v4 BE-eVLPs (**FIG. 3G**).

[0440] To evaluate Cas-independent off-target DNA editing, an orthogonal R-loop assay was performed, which was previously validated as a strategy for assessing the ability of a base editor to deaminate DNA in an unguided manner without requiring whole-genome sequencing (Doman *et al.*, 2020; Yu *et al.*, 2020). Compared with transfection of DNA plasmid encoding the same BE, v4 BE-eVLPs exhibited a >100-fold reduction in Cas-independent off-target editing, down to virtually undetected levels (**FIG. 3H**, **FIG. 11B**). These results confirm and extend previous findings that off-target editing by highly active BEs can be substantially minimized with RNP delivery (Doman *et al.*, 2020; Jang *et al.*, 2021; Lyu *et al.*, 2021; Newby *et al.*, 2021; Rees and Liu, 2018; Richter *et al.*, 2020; Yeh *et al.*, 2018) and highlight the ability of eVLPs to support highly efficient on-target base editing with minimal off-target editing.

[0441] The DNA-free nature of eVLPs in principle avoids the possibility of DNA integration into the genomes of transduced cells, an important safety advantage over existing viral delivery modalities (David and Doherty, 2017; Milone and O'Doherty, 2018). qPCR was used to verify that purified v4 BE-eVLPs contain < 0.03 molecules of BE-encoding DNA per VLP (**FIG. 3I**). Additionally, while substantial amounts (8.7 ng/μL) of BE-encoding DNA was detected in cellular lysate from HEK293T cells that were transfected with BE-encoding plasmids, BE-encoding DNA was not detected in cellular lysate from v4 BE-eVLP-treated HEK293T cells above background levels in samples from untreated cells (< 0.02 ng/μL) (**FIG. 3J**). These results demonstrate that BE-eVLPs do not expose transduced cells to detected levels of DNA encoding base editors, thereby minimizing the possibility of genomic integration of cargo DNA.

v4 BE-eVLPs efficiently edit primary human and mouse cells

[0442] To further explore the utility of v4 BE-eVLPs, their ability to target and edit a variety of primary human or mouse cells *ex vivo* was assessed. ABE-mediated correction of nonsense mutations in *COL7A1* that cause recessive dystrophic epidermolysis bullosa (RDEB) in primary human patient-derived fibroblasts has previously been demonstrated (Osborn *et al.*, 2020). After transducing primary fibroblasts harboring a homozygous *COL7A1*(R185X) mutation with v4 BE-eVLPs, >95% editing was observed at the target adenine base with no difference in the cellular viability between VLP-treated and untreated cells (**FIG. 4A** and **FIG. 11C**). Additionally, minimal Cas-dependent off-target editing was observed at ten previously identified off-target sites (Osborn *et al.*, 2020) (**FIG. 11D**). The ability of v4 BE-eVLPs to correct a nonsense mutation in primary fibroblasts derived from a mouse model of

Mucopolysaccharidosis type IH (Wang *et al.*, 2010) was also assessed. Again, >95% correction of the *Idua*(W392X) mutation was observed following v4 BE-eVLP transduction (**FIG. 4B**). These results validate that BE-VLP activity is not restricted to immortalized cell lines and demonstrate that v4 BE-eVLPs can achieve levels of base editing in primary human and mouse fibroblasts approaching 100%.

[0443] Next, BE-eVLP-mediated editing in primary human T cells was investigated. Gene editing strategies that reduce the expression of immunomodulatory proteins on the surface of T cells, including MHC class I and MHC class II, could advance T-cell therapies by enabling “off-the-shelf” allogeneic chimeric antigen receptor (CAR) T cells. Previous reports have shown that disrupting splice sites in the *B2M* and *CIITA* genes reduces expression of MHC class I and MHC class II in primary human T cells (Gaudelli *et al.*, 2020; LeibundGut-Landmann *et al.*, 2004; Serreze *et al.*, 1994). Treating primary human T cells with v4 BE-eVLPs led to 45–60% disruption of *B2M* and *CIITA* splice sites (**FIG. 4C**). Collectively, these results confirm that eVLPs can efficiently edit clinically relevant primary human cell types *ex vivo* and lay a foundation for the further optimization of BE-VLP editing efficiencies in primary human T cells.

In vivo base editing in the CNS with eVLPs

[0444] The robust activity of eVLPs *ex vivo* suggested that they might be promising vehicles for delivering BE RNPs *in vivo*. To begin to assess their *in vivo* efficacy, the ability of eVLPs to enable base editing within the mouse central nervous system (CNS) was first investigated. v4 BE-eVLPs were produced that install a silent mutation in mouse *Dnmt1* at a genomic locus known to be amenable to nuclease-mediated indel formation and adenine base editing *in vivo* (Levy *et al.*, 2020; Swiech *et al.*, 2015). To deliver BE-eVLPs to the CNS, neonatal cerebroventricular (P0 ICV) injections were performed, which are direct injections into cerebrospinal fluid that bypass the blood–brain barrier, similar to the intrathecal injections currently used to deliver nusinersen in patients with spinal muscular atrophy (Mercuri *et al.*, 2018).

[0445] v4 BE-eVLPs were co-injected into each hemisphere together with a VSV-G-pseudotyped lentivirus encoding EGFP fused to a nuclear membrane-localized Klarsicht/ANC-1/Syne-1 homology (KASH) domain (**FIG. 5A**). It was reasoned that this strategy would enable the isolation of GFP-positive nuclei as a way to enrich cells that were exposed to eVLPs. This approach is particularly useful to determine editing efficiencies following injection in the brain, where many cells may not be accessible. Three weeks post-

injection, bulk unsorted (all nuclei) and GFP-positive nuclei from cortical and mid-brain tissues were analyzed, and base editing was assessed by high-throughput sequencing (**FIG. 5A**).

[0446] The frequencies of GFP-positive nuclei in both cortical and mid-brain tissues were low (**FIG. 13B**), consistent with previous reports that the cells transduced by VSV-G-pseudotyped lentiviruses injected into the mouse brain are localized near the injection site (Humbel *et al.*, 2021; Parr-Brownlie *et al.*, 2015), possibly because the size of the viral particles, which have an average diameter ~3-fold larger than the width of the brain extracellular space (Thorne and Nicholson, 2006), may hinder diffusion through bulk brain tissue. Encouragingly, 53% and 55% editing in GFP-positive cortex and mid-brain cells was observed, respectively, corresponding to 6.1% and 4.4% editing of bulk cortex and mid-brain (**FIG. 5B**). These data establish BE-eVLPs as a new non-viral delivery system for CNS base editing applications that deliver robust levels of active BE RNP per transduction event, although improvements in transduction efficiency are needed to achieve high levels of editing in bulk brain tissue.

In vivo liver base editing with eVLPs leads to efficient knockdown of Pcsk9

[0447] To further explore the utility of BE-eVLPs *in vivo*, their ability to mediate therapeutic base editing in adult animals was investigated. First, proprotein convertase subtilisin/kexin type 9 (*Pcsk9*), a therapeutically relevant gene involved in cholesterol homeostasis (Abifadel *et al.*, 2003; Fitzgerald *et al.*, 2014), was targeted. Loss-of-function *PCSK9* mutations occur naturally without apparent adverse health consequences (Abifadel *et al.*, 2003; Cohen *et al.*, 2005; Cohen *et al.*, 2006; Hooper *et al.*, 2007; Rao *et al.*, 2018). These individuals have lower levels of low-density lipoprotein (LDL) cholesterol in the blood and a reduced risk of atherosclerotic cardiovascular disease, suggesting that disrupting the *PCSK9* gene could be a promising strategy for the treatment of familial hypercholesterolemia (Musunuru *et al.*, 2021; Rothgangl *et al.*, 2021). The optimized v4 BE-VLP architecture supported much more robust editing in the liver than a previously described VLP architecture (v1 BE-VLP), which mediated only 1.5% editing, 26-fold less than v4 eVLPs at the same dose (**FIG. 6B**).

[0448] BE-eVLPs that target and disrupt the splice donor at the boundary of *Pcsk9* exon 1 and intron 1, a previously established base editing strategy for *Pcsk9* knockdown in the mouse liver (Musunuru *et al.*, 2021; Rothgangl *et al.*, 2021), were designed and produced. Systemic (retro-orbital) injections of the eVLPs into 6- to 7-week-old adult C57BL/6 mice were performed, and base editing in the bulk liver was measured one week after injection

(**FIG. 6A**). 63% editing efficiency in the bulk liver was observed following treatment with the highest dose (7×10^{11} eVLPs) of v4 BE-eVLPs (**FIG. 6B**), which is comparable to editing efficiencies typically achieved at this site with optimized, state-of-the-art AAV-based delivery modalities and lipid nanoparticle (LNP)-based mRNA delivery systems (Musunuru *et al.*, 2021; Rothgangl *et al.*, 2021). The engineered v4 BE-eVLP architecture supported 26-fold higher editing levels in the liver than the VLP architecture based on a previously reported design (v1 BE-VLP) at the same dose (**FIG. 6B**). These results establish efficient base editing by RNPs at a therapeutically relevant locus in the mouse liver.

[0449] In mice treated with the highest dose of v4 BE-eVLPs, base editing efficiencies were also assessed in non-liver tissues, including the heart, skeletal muscle, lungs, kidney, and spleen. 4.3% base editing in the spleen was observed, and no editing above background levels was observed in the lungs, kidneys, heart, and muscle. This pattern of editing across tissues is consistent with the previously characterized tissue tropism of intravenously administered VSV-G-pseudotyped particles (Pan *et al.*, 2002).

[0450] To assess whether treatment with BE-eVLPs resulted in Cas-dependent off-target editing in liver tissue, we performed CIRCLE-seq (Tsai *et al.*, 2017) to nominate potential off-target loci. From the nominated loci, 14 candidate off-target sites were selected and examined by targeted high-throughput sequencing based on homology near the PAM-proximal region of the protospacer. No detectable off-target editing above background levels was observed at any of these loci in genomic DNA isolated from livers of mice treated with 7×10^{11} v4 BE-eVLPs (**FIG. 6D**). In contrast, low but detectable (0.1–0.3%) levels of off-target editing were observed at three of these loci in genomic DNA isolated from livers of mice treated with dual AAV8 vectors (1×10^{11} viral genomes) encoding ABE8e and the same *Pcsk9*-targeting sgRNA (**FIG. 6D**). These results demonstrate that v4 BE-eVLPs can offer comparable on-target editing but minimal off-target editing *in vivo*, an improvement compared to existing viral based delivery approaches.

[0451] Phenotypic analyses performed one-week post-injection revealed a 78% reduction in serum *Pcsk9* protein level in mice treated with 7×10^{11} v4 BE-eVLPs compared to untreated mice (**FIG. 6E**). To assess the potential toxicity of systemically administered eVLPs, one-week after injection of 7×10^{11} v4 BE-eVLPs, serum alanine aminotransferase (ALT) and aspartate transaminase (AST) levels, important biomarkers of hepatocellular injury (Meunier and Larrey, 2019), were evaluated. All mice exhibited AST and ALT levels within the normal range, and there were no discernible differences between the untreated mice and the

BE-eVLP-treated mice (**FIG. 14A**). Additionally, liver histology was performed on samples from eVLP-treated and untreated mice, and no evident morphological differences due to BE-eVLP treatment were found (**FIGS. 14B-14C**). Together, these results demonstrate that v4 BE-eVLPs can mediate efficient, therapeutically relevant base editing in the mouse liver with no apparent adverse consequences and no detected off-target editing.

v4 BE-eVLPs restore visual function in a mouse model of genetic blindness

[0452] Finally, BE-eVLPs were applied to correct a disease-causing point mutation in an adult mouse model of a genetic retinal disorder. Loss-of-function mutations in multiple genes are associated with various forms of Leber congenital amaurosis (LCA), a family of monogenic retinal disorders that involve retinal degeneration, early-onset visual impairment, and eventual blindness (Cideciyan, 2010; den Hollander *et al.*, 2008). Gene editing approaches hold promise to treat and cure congenital blindness; an ongoing clinical trial (NCT03872479) uses AAV-delivered Cas9 nucleases to disrupt an aberrant splice site in *CEP290* that is associated with rare Leber congenital amaurosis 10 (LCA10). Loss-of-function mutations in other genes, including the retinoid isomerohydrolase *RPE65*, are also candidates for *in vivo* correction using precision gene editing agents (Sodi *et al.*, 2021; Suh *et al.*, 2021).

[0453] It was investigated whether v4 BE-eVLPs can restore visual function in a mouse model of LCA. *rd12* mice harbor a nonsense mutation in exon 3 of *Rpe65* (c.130C > T; p.R44X) that causes a near-complete loss of visual function (Pang *et al.*, 2005; Suh *et al.*, 2021). A homologous mutation responsible for LCA has recently been identified in people (Zhong *et al.*, 2019), highlighting the clinical relevance of the *rd12* model.

[0454] v4 BE-eVLPs encapsulating ABE8e-NG RNPs and an sgRNA (**FIG. 7A**) that targets the *Rpe65*(R44X) mutation (hereafter referred to as ABE8e-NG-eVLPs) were designed and produced. ABE8e-NG-eVLPs were pseudotyped with VSV-G to enable them to efficiently transduce retinal pigment epithelium (RPE) cells (Puppo *et al.*, 2014; Suh *et al.*, 2021). ABE8e-NG-eVLPs were injected subretinally into 4-week-old *rd12* mice. In a separate cohort, replication-incompetent lentivirus encoding the identical ABE8e-NG and sgRNA constructs (ABE8e-NG-LV) were also subretinally injected. It was previously reported that lentiviral delivery of ABEs can successfully restore visual function in *rd12* mice (Suh *et al.*, 2021).

[0455] Five weeks post-injection, RPE tissue was harvested, and high-throughput sequencing of RPE genomic DNA was performed (**FIG. 7B**). Encouragingly, sequencing analysis

revealed that ABE8e-NG-VLPs and ABE8e-NG-LV successfully mediated 21% and 11.5% correction respectively of the R44X mutation at position A₆ of the protospacer (**FIG. 7C**). Notably, ABE8e-NG-VLPs achieved 1.8-fold higher editing at the target base compared to ABE8e-NG-LV, even though BE-VLP delivery is transient. These results demonstrate that eVLPs enable highly efficient correction of a pathogenic mutation in the mouse RPE.

[0456] While highly efficient correction of the target mutation was observed, it was also observed that both ABE8e-NG-eVLP and ABE8e-NG-LV induced substantial levels of bystander editing (**FIG. 7C**) due to the wide editing window of ABE8e-NG (Richter *et al.*, 2020), such that the majority of edited alleles contained conversions at A₃, A₆, and/or A₈ as opposed to A₆ alone (**FIG. 7D**). The bystander edits at positions A₃ and A₈ lead to *Rpe65* missense mutations C45R and L43P respectively. It was previously shown that the L43P mutation renders the *Rpe65* enzyme inactive (Suh *et al.*, 2021). Indeed, after performing scotopic electroretinography (ERG) to assess retinal cell response, minimal rescue of visual function in both ABE8e-NG-eVLP-injected and ABE8e-NG-LV-injected eyes was observed (**FIG. 7E**). These results suggested that the wide base editing window of ABE8e-NG is not well-suited to precisely correct the *Rpe65*(R44X) mutation.

[0457] To address this limitation, v4 BE-eVLPs that encapsulate ABE7.10-NG, which exhibits a narrower editing window compared to ABE8e-NG (Huang *et al.*, 2019; Richter *et al.*, 2020), were designed and produced. Subretinal injection of ABE7.10-NG-eVLPs into adult *rd12* mice led to 12% correction of the R44X mutation in RPE genomic DNA with virtually no bystander editing (**FIG. 7F**). Specifically, it was observed that ABE7.10-NG-eVLP treatment resulted in 11% perfect R44X correction without bystander edits, a 9-fold improvement in perfect correction relative to ABE8e-NG-eVLP treatment (**FIG. 7G**). Furthermore, treatment with ABE7.10-NG-eVLPs resulted in a 1.4-fold improvement in bystander-free correction relative to treatment with ABE7.10-NG-LV, a lentivirus encoding the identical ABE7.10-NG and sgRNA constructs, an additional demonstration that v4 BE-eVLP transient delivery can achieve comparable or higher editing efficiencies compared to lentiviral BE delivery (**FIG. 7G**).

[0458] It was confirmed via western blot that ABE7.10-NG-eVLP treatment restored the expression of *Rpe65* protein. Notably, ABE7.10-NG-LV-treated eyes still expressed BE protein 5-weeks post-injection, while ABE7.10-NG-eVLP-treated eyes did not (**FIG. 7I**), demonstrating the transient exposure of cells *in vivo* to base editors delivered using eVLPs. Importantly, ABE7.10-NG-eVLPs successfully rescued visual function to similar levels

relative to ABE7.10-NG-LV as measured by ERG of the treated eyes (**FIGS. 7H** and **7J**). It was previously shown that this level of ERG rescue corresponds to other improvements in visual function, including restoration of the visual chromophore and recovery of visual cortical responses (Suh *et al.*, 2021). These results demonstrate that eVLPs can mediate efficient correction of a pathogenic mutation in the mouse RPE with amelioration of the disease phenotype.

[0459] To further analyze editing outcomes, RNA was extracted from treated eyes, and targeted high-throughput sequencing of specific cDNAs was performed. As expected, in the eVLP treated eyes, up to 64% of A•T-to-G•C conversion of the target adenine (A₆) in the on-target *Rpe65* transcript was observed (**FIG. 15A**). The higher proportion of corrected *Rpe65* transcripts compared to *Rpe65* genomic loci potentially reflects nonsense-mediated decay of uncorrected mRNAs.

[0460] BEs are known to exhibit low-level transcriptome-wide Cas-independent off-target RNA editing (Anzalone *et al.*, 2020). To investigate this possibility, off-target RNA editing by ABE-eVLPs and ABE-LVs was assessed by sequencing the *Mcm3ap* and *Perp* transcripts from treated eyes, two transcripts that were previously identified as potential candidates for off-target RNA editing based on their sequence similarity to the native TadA deaminase substrate (Jo *et al.*, 2021). RNA off-target editing by ABE8e-NG-LV in both transcripts and low but detectable RNA off-target editing by ABE7.10-NG-LV at one adenine in *Perp* was observed (**FIGS. 15B-15C**). In contrast, there was no detection of any RNA off-target editing above background in these two transcripts by ABE8e-NG-eVLPs or ABE7.10-NG-eVLPs (**FIGS. 15B-15C**). Collectively, these findings highlight the therapeutic utility of eVLPs as a DNA-free method for transiently delivering BE RNPs *in vivo* with high on-target editing and minimal off-target editing.

Discussion

[0461] Presented herein is an efficient engineered VLP platform that can safely deliver RNPs for therapeutically relevant *ex vivo* and *in vivo* applications. Through identifying and engineering solutions to three distinct bottlenecks to VLP delivery efficiency, protein loading was improved within v4 eVLPs by an average of 16-fold and base editing efficiencies by an average of 8-fold compared to initial designs based on previously reported VLP scaffolds. These findings suggest that v4 eVLPs are highly versatile and suitable for a wide range of both *ex vivo* and *in vivo* base editing applications. It is also anticipated that the eVLP

architecture will serve as a modular platform for delivering other proteins or RNPs of interest in addition to BEs and nucleases.

[0462] The results presented herein highlight the potential therapeutic benefit of using rational engineering to further advance delivery platforms for gene editing agents. While VLPs have been used previously to deliver Cas9 nuclease RNPs (Campbell *et al.*, 2019; Choi *et al.*, 2016; Gee *et al.*, 2020; Hamilton *et al.*, 2021; Indikova and Indik, 2020; Lyu *et al.*, 2019; Mangeot *et al.*, 2019), and a recent study used VLPs to deliver BE RNPs to HEK293T cells with lower efficiencies than the eVLPs described here (Lyu *et al.*, 2021), no previous study has reported therapeutic levels of post-natal *in vivo* gene editing of any type using RNP-delivering VLPs. The eVLP platform developed in this work uses a rationally engineered architecture that was customized to package increased amounts of cargo and improve particle titers. These eVLPs can mediate therapeutic levels of *in vivo* base editing across multiple organs and routes of administration in mice, achieving the highest levels of post-natal *in vivo* gene editing using RNPs reported to date.

[0463] A single intravenous injection of eVLPs mediated base editing of *Pcsk9* in the mouse liver at efficiencies >60%, comparable to those achieved at the same target by current state-of-the-art BE delivery methods, including AAV-mediated delivery of BE-encoding DNA (Rothgangl *et al.*, 2021) and LNP-mediated delivery of BE-encoding mRNA (Musunuru *et al.*, 2021; Rothgangl *et al.*, 2021). However, eVLPs offer key advantages over both AAV-mediated DNA delivery and LNP-mediated mRNA delivery strategies. AAV-mediated delivery can lead to detectable levels of viral genome integration into the genomes of transduced cells, which can lead to oncogenesis (Chandler *et al.*, 2017; Koblan *et al.*, 2021), while eVLPs lack DNA and therefore should avoid the possibility of insertional mutagenesis. Additionally, AAV-mediated delivery leads to prolonged cargo expression, increasing the frequency of off-target editing, but transient eVLP-mediated delivery of BE RNPs greatly reduces the opportunity for off-target editing, as was shown both *in vitro* and *in vivo* (**FIGs. 3G, 3H, and 6D**). While LNP-mediated delivery of BE-encoding mRNA is also transient, delivering BE RNPs offers even shorter exposures to editing agents and lower off-target editing opportunities due to the shorter lifetime of RNPs in cells compared with mRNA, each copy of which generates cellular RNPs throughout the lifetime of the mRNA (Newby *et al.*, 2021).

[0464] While LNPs can efficiently package mRNAs, packaging gene editing agent RNPs within LNPs is substantially more challenging (Wei *et al.*, 2020). Because eVLPs can achieve

comparable levels of editing in the liver as these other strategies but possess the important advantages mentioned above, they are a particularly attractive option for further development as a therapeutic modality for *in vivo* editing approaches to treat genetic liver diseases. The v4 eVLP architecture was critical for achieving robust editing in the mouse liver and improved *in vivo* editing efficiency by 26-fold compared to a previously reported (v1) VLP design (**FIG. 6B**), underscoring the importance of engineering VLP architectures for *in vivo* editing. The observed degree of base editing at this *Pcsk9* splice donor with v4 BE-eVLPs (>60%) is thought to be sufficient for the reduction of serum LDL and treatment of hypercholesterolemia (Musunuru *et al.*, 2021).

[0465] A single subretinal injection of v4 BE-eVLPs in a mouse model of LCA efficiently corrected the disease-causing point mutation and restored visual function. In this model, once again, eVLPs achieved editing efficiencies and levels of rescue that are comparable or higher than those previously achieved using viral delivery methods, including lentiviral BE delivery (Suh *et al.*, 2021) and AAV-mediated BE delivery (Jo *et al.*, 2021). The accessibility of the eyes and their immune-privileged status (Taylor, 2009) may more readily enable the translation of new delivery modalities into pre-clinical and clinical studies. These data provide evidence of the therapeutic potential of BE-eVLPs as a means to correct pathogenic point mutations that cause ocular disorders.

[0466] The developments reported herein combine the one-time treatment potential of gene editing agents and the transient nature of RNPs to minimize the opportunity for unwanted off-target editing or DNA integration with the efficient, tissue-targeted nature of viral transduction. These findings thus suggest that eVLPs are an attractive alternative to other delivery strategies for the *in vivo* or *ex vivo* delivery of base editors, nucleases, and other proteins of therapeutic interest.

Methods

Materials Availability

[0467] Plasmids generated in this Example are available from Addgene (additional details provided in the Table 1).

Table 1. Key Resources

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse anti-Cas9 monoclonal antibody	Thermo Fisher Scientific	Cat#MA5-23519
Mouse anti-MLV p30 monoclonal antibody	Abcam	Cat#ab130757
Mouse anti-VSVG monoclonal antibody	Sigma-Aldrich	Cat#V5507

IRDye 680RD goat anti-mouse antibody	LI-COR	Cat#926-68070
Mouse anti-Rpe65 monoclonal antibody	(Golczak et al., 2010)	
Goat anti-mouse IgG-HRP antibody	Cell Signaling Technology	Cat#7076S
Mouse anti-Cas9 monoclonal antibody	Invitrogen	Cat#MA523519
Rabbit anti- β -actin polyclonal antibody	Cell Signaling Technology	Cat#7076S
Goat anti-rabbit IgG-HRP antibody	Cell Signaling Technology	Cat#7074S
Bacterial and Virus Strains		
One Shot Mach1 T1 Phage-Resistant Chemically Competent <i>E. coli</i>	Thermo Fisher Scientific	Cat#C862003
NEB Stable Competent <i>E. coli</i>	New England BioLabs	Cat#C3040H
Chemicals, Peptides, and Recombinant Proteins		
USER enzyme	New England BioLabs	Cat#M5505S
DpnI	New England BioLabs	Cat#R0176S
KLD Enzyme Mix	New England BioLabs	Cat#M0554S
Lipofectamine 2000	Thermo Fisher Scientific	Cat#11668019
jetPRIME Transfection Reagent	Polyplus	Cat#114-75
FuGENE HD Transfection Reagent	Promega	Cat#E2312
PEG-it Virus Precipitation Solution	System Biosciences	Cat#LV825A-1
Recombinant Cas9 (<i>S. pyogenes</i>) nuclease	New England BioLabs	Cat#M0386
SYBR green dye	Lonza	Cat#50512
Proteinase K	Thermo Fisher Scientific	Cat#EO0492
Proteinase K	New England BioLabs	Cat#P8107S
Human AB Serum	Valley Biomedical	Cat#HP1022HI
N-Acetyl-L-cysteine	Sigma-Aldrich	Cat#A7250-100G
Recombinant Human IL-2	Peprotech	Cat#200-02
Recombinant Human IL-7	Peprotech	Cat#200-07
Recombinant Human IL-15	Peprotech	Cat#200-15
RetroNectin [®]	Clontech/Takara	Cat#T100A/B
Dynabeads [™] Human T-Expander CD3/CD28 beads	Thermo Fisher Scientific	Cat#1161D
QuickExtract [™] DNA Extraction Solution	Lucigen	Cat#QE09050
Salt Active Nuclease	ArcticZymes	Cat#70910-202
BSA	New England BioLabs	Cat#B9000S
0.9% NaCl	Fresenius Kabi	Cat#918610
Critical Commercial Assays		
Phusion U Multiplex PCR Master Mix	Thermo Fisher Scientific	Cat#F562L
Phusion High-Fidelity DNA Polymerase	New England BioLabs	Cat#M0530S
QIAquick PCR Purification Kit	QIAGEN	Cat#28104
QIAquick Gel Extraction Kit	QIAGEN	Cat#28704
QIAGEN Plasmid Plus Midi Kit	QIAGEN	Cat#12943
QIAGEN Plasmid Plus Maxi Kit	QIAGEN	Cat#12963
FastScan [™] Cas9 (<i>S. pyogenes</i>) ELISA Kit	Cell Signaling Technology	Cat#29666C
MuLV Core Antigen ELISA Kit	Cell Biolabs	Cat#VPK-156
QIAmp Viral RNA Mini Kit	QIAGEN	Cat#52904
SuperScript [™] III First-Strand Synthesis SuperMix	Thermo Fisher Scientific	Cat#18080400
EasySep Human T Cell Isolation Kit	STEMCELL Technologies	Cat#17951
AAVpro Titration Kit version 2	Clontech/Takara	Cat#6233
Agencourt DNAdvance Kit	Beckman	Cat#V10309
Total Cholesterol Reagents	Thermo Fisher Scientific	Cat#TR13421
Mouse Proprotein Convertase 9/PCSK9 Quantikine ELISA Kit	R&D Systems	Cat#MPC900
QuickTiter [™] Lentivirus Titer Kit	Cell Biolabs	Cat#VPK-107-5
AllPrep DNA/RNA Mini Kit	QIAGEN	Cat#80284
MiSeq Reagent Kit v2 (300-cycles)	Illumina	Cat#MS-102-2002
MiSeq Reagent Micro Kit v2 (300-cycles)	Illumina	Cat#MS-103-1002
Deposited Data		

Targeted amplicon sequencing data	This study	PRJNA768458
Experimental Models: Cell Lines		
Human: HEK293T	ATCC	Cat#CRL-3216
Human: Gesicle Producer 293T	Takara	Cat#632617
Mouse: NIH/3T3	ATCC	Cat#CRL-1658
Mouse: Neuro-2a	ATCC	Cat#CCL-131
Experimental Models: Organisms		
Timed pregnant C57BL/6J mice	Charles River Laboratories	Cat#027
C57BL/6J mice	Jackson Laboratory	Cat#000664
<i>rd12</i> mice	Jackson Laboratory	Cat#005379
Recombinant DNA		
pCMV-VSV-G	Addgene	8454
psPAX2	Addgene	12260
pBS-CMV-gagpol	Addgene	35614
BIC-Gag-Cas9	Addgene	119942
lentiCRISPRv2	Addgene	135955
v4 BE-VLP	Addgene (this study)	TBA
Software and Algorithms		
CRISPResso2	(Clement et al., 2019)	github.com/pinellolab/CRISPResso2
Prism	GraphPad	graphpad.com

Data and Code Availability

[0468] The sequencing data generated in this Example is deposited at the NCBI Sequence Read Archive database under PRJNA768458. The code used for data processing and analysis are available at github.com/pinellolab/CRISPResso2.

Experimental Model and Subject Details

Cell culture conditions

[0469] HEK293T cells (ATCC; CRL-3216), Gesicle Producer 293T cells (Takara; 632617), 3T3 cells (ATCC; CRL-1658), and Neuro-2a cells (ATCC; CCL-131) were maintained in DMEM + GlutaMAX (Life Technologies) supplemented with 10% (v/v) fetal bovine serum. Primary human and mouse fibroblasts were maintained in MEM alpha media (Thermo Fisher; 12571063) containing 20% (v/v) FBS, 2 mM GlutaMAX (Thermo Fisher; 35050061), 1 % penicillin and streptomycin (Thermo Fisher; 15070063), 1X Nonessential amino acids (Thermo Fisher; 11140050), 1X Antioxidant Supplement (Sigma Aldrich; A1345), 10 ng/mL Epidermal Growth Factor from murine submaxillary gland (Sigma Aldrich; E4127) and 0.5 ng/mL Fibroblast Growth Factor (Sigma Aldrich; F3133). Cells were cultured at 37 °C with 5% carbon dioxide and were confirmed to be negative for mycoplasma by testing with MycoAlert (Lonza Biologics).

Isolation of primary human T cells

[0470] Primary human T cells were isolated as described previously (Chen et al., 2021). Buffy coats were obtained from Memorial Blood Centers (St. Paul, MN) and peripheral blood mononuclear cells were isolated using SepMate tubes (STEMCELL Technologies; 85450). The EasySep Human T-cell Isolation Kit was used to enrich for T-cells that were then frozen for long-term storage.

Method Details*Cloning*

[0471] All plasmids used in this Example were cloned using either USER cloning or KLD cloning as described previously (Doman et al., 2020). DNA was PCR-amplified using PhusionU Green Multiplex PCR Master Mix (Thermo Fisher Scientific). Mach1 (Thermo Fisher Scientific) chemically competent *E. coli* were used for plasmid propagation.

BE-eVLP production and purification

[0472] As depicted in the embodiment of FIG. 16, BE-eVLPs were produced by transient transfection of Gesicle Producer 293T cells. Gesicle cells were seeded in T-75 flasks (Corning) at a density of 5×10^6 cells per flask. After 20–24 h, cells were transfected using the jetPRIME transfection reagent (Polyplus) according to the manufacturer's protocols. For producing v1–v3 BE-eVLPs, a mixture of plasmids expressing VSV-G (400 ng), MLVgag-pro-pol (2,800 ng), MLVgag–ABE8e (1,700 ng), and an sgRNA (4,400 ng) were co-transfected per T-75 flask. For MLVgag–ABE8e:MLVgag-pro-pol stoichiometry optimization, the total amount of plasmid DNA for these two components was fixed at 4,500 ng, and the relative amounts of each were varied. For producing v4 BE-eVLPs, a mixture of plasmids expressing VSV-G (400 ng), MMLVgag-pro-pol (3,375 ng), MMLVgag–3xNES–ABE8e (1,125 ng), and an sgRNA (4,400 ng) were co-transfected per T-75 flask. Exemplary BE-eVLP construct protein sequences are provided in Table 4.

[0473] 40–48 h post-transfection, producer cell supernatant was harvested and centrifuged for 5 min at 500 g to remove cell debris. The clarified eVLP-containing supernatant was filtered through a 0.45 μ m PVDF filter. For BE-eVLPs that were used in cell culture, unless otherwise stated, the filtered supernatant was concentrated 100-fold using PEG-it Virus Precipitation Solution (System Biosciences; LV825A-1) according to the manufacturer's protocols. For BE-eVLPs that were injected into mice, the filtered supernatant was concentrated 1000–3000-fold by ultracentrifugation using a cushion of 20% (w/v) sucrose in

PBS. Ultracentrifugation was performed at 26,000 rpm for 2 h (4°C) using either an SW28 rotor in an Optima XPN Ultracentrifuge (Beckman Coulter) or an AH-629 rotor in a Sorvall WX+ Ultracentrifuge (Thermo Fisher Scientific). Following ultracentrifugation, BE-eVLP pellets were resuspended in cold PBS (pH 7.4) and centrifuged at 1,000 g for 5 min to remove debris. BE-eVLPs were frozen at a rate of 1°C/min and stored at -80°C. eVLPs were thawed on ice immediately prior to use.

BE-eVLP transduction in cell culture and genomic DNA isolation

[0474] Cells were plated for transduction in 48-well plates (Corning) at a density of 30,000–40,000 cells per well. After 20–24 h, BE-eVLPs were added directly to the culture media in each well. 48–72 h post-transduction, cellular genomic DNA was isolated as previously reported (Doman et al., 2020). Briefly, cells were washed once with PBS and lysed in 150 µL of lysis buffer (10 mM Tris-HCl pH 8.0, 0.05% SDS, 25 µg mL⁻¹ Proteinase K (Thermo Fisher Scientific)) at 37 °C for 1 h followed by heat inactivation at 80°C for 30 min.

High-throughput sequencing of genomic DNA

[0475] Genomic DNA was isolated as described above. Following genomic DNA isolation, 1 µL of the isolated DNA (1–10 ng) was used as input for the first of two PCR reactions. Genomic loci were amplified in PCR1 using PhusionU polymerase (Thermo Fisher Scientific). PCR1 primers for genomic loci are listed in Table 3 under the HTS_fwd and HTS_rev columns. PCR1 was performed as follows: 95 °C for 3 min; 30–35 cycles of 95 °C for 15 s, 61 °C for 20 s, and 72 °C for 30s; 72°C for 1 min. PCR1 products were confirmed on a 1% agarose gel. Then, 1 µL of PCR1 was used as an input for PCR2 to install Illumina barcodes. PCR2 was conducted for 9 cycles of amplification using a Phusion HS II kit (Life Technologies). Following PCR2, samples were pooled and gel purified in a 1% agarose gel using a Qiaquick Gel Extraction Kit (Qiagen). Library concentration was quantified using the Qubit High-Sensitivity Assay Kit (Thermo Fisher Scientific). Samples were sequenced on an Illumina MiSeq instrument (paired-end read, read 1: 200–280 cycles, read 2: 0 cycles) using an Illumina MiSeq 300 v2 Kit (Illumina).

High-throughput sequencing data analysis

[0476] Sequencing reads were demultiplexed using the MiSeq Reporter software (Illumina) and were analyzed using CRISPResso2 (Clement et al., 2019) as previously described (Doman et al., 2020). Batch analysis mode (one batch for each unique amplicon and sgRNA combination analyzed) was used in all cases. Reads were filtered by minimum average quality score (Q > 30) prior to analysis. The following quantification window parameters

were used: -w 20 -wc -10. Base editing efficiencies are reported as the percentage of sequencing reads containing a given base conversion at a specific position. Prism 9 (GraphPad) was used to generate dot plots and bar plots.

Immunoblot analysis of BE-eVLP protein content

[0477] BE-eVLPs were lysed in Laemmli sample buffer (50 mM Tris-HCl pH 7.0, 2% sodium dodecyl sulfate (SDS), 10% (v/v) glycerol, 2 mM dithiothreitol (DTT)) by heating at 95°C for 15 min. Lysed BE-eVLPs were spotted onto a dry nitrocellulose membrane (Thermo Fisher Scientific) and dried for 30 min. The membrane was blocked for 1 h at room temperature with rocking in blocking buffer: 1% bovine serum albumin (BSA) in TBST (150 mM NaCl, 0.5% Tween-20, and 50 mM Tris-HCl). After blocking, the membrane was incubated overnight at 4°C with rocking with one of the following primary antibodies diluted in blocking buffer: mouse anti-Cas9 (Thermo Fisher; MA5-23519, 1:1000 dilution), mouse anti-MLV p30 (Abcam; ab130757, 1:1500 dilution), or mouse anti-VSV-G (Sigma Aldrich; V5507, 1:50000 dilution). The membrane was washed three times with 1xTBST (Tris-buffered saline + 0.5% Tween-20) for 10 min each time at room temperature, then incubated with goat anti-mouse antibody (LI-COR IRDye 680RD; 926-68070, 1:10000 dilution in blocking buffer) for 1 h at room temperature with rocking. The membrane was washed as before and imaged using an Odyssey Imaging System (LI-COR).

Western blot analysis of BE-eVLP protein content

[0478] BE-eVLPs were lysed as described above. Protein extracts were separated by electrophoresis at 150 V for 45 min on a NuPAGE 3–8% Tris-Acetate gel (Thermo Fisher Scientific) in NuPAGE Tris-Acetate SDS running buffer (Thermo Fisher Scientific). Transfer to a PVDF membrane was performed using an iBlot 2 Gel Transfer Device (Thermo Fisher Scientific) at 20 V for 7 min. The membrane was blocked for 1 h at room temperature with rocking in blocking buffer: 1% bovine serum albumin (BSA) in TBST (150 mM NaCl, 0.5% Tween-20, and 50 mM Tris-HCl). After blocking, the membrane was incubated overnight at 4°C with rocking with mouse anti-Cas9 (Cell Signaling Technology; 14697, 1:1000 dilution). The membrane was washed three times with 1xTBST for 10 min each time at room temperature, then incubated with goat anti-mouse antibody (LI-COR IRDye 680RD; 926-68070, 1:10000 dilution in blocking buffer) for 1 h at room temperature with rocking. The membrane was washed as before and imaged using an Odyssey Imaging System (LI-COR). The relative amounts of cleaved ABE and full-length gag–ABE were quantified by

densitometry using ImageJ, and the fraction of cleaved ABE relative to total (cleaved + full-length) ABE was calculated.

Immunofluorescence microscopy of producer cells

[0479] Gesicle Producer 293T cells were seeded at a density of 15,000 cells per well in PhenoPlate™ 96-well microplates coated with poly-D-lysine (PerkinElmer). After 24 h, cells were co-transfected with 1 ng of v2.4 or v3.4 BE-VLP plasmids, 40 ng of mouse *Dnmt1*-targeting sgRNA plasmid, and 40 ng of pUC19 plasmid using the jetPRIME transfection reagent (Polyplus) according to the manufacturer's protocols. After 40 h, 32% aqueous paraformaldehyde (Electron Microscopy Sciences) was added dropwise directly into the cellular media to a final concentration of 4% paraformaldehyde. Cells were subsequently fixed for 20 min at room temperature. After fixation, cells were washed three times with PBS and then permeabilized with 1xPBST (PBS + 0.1% Triton X-100) for 30 min at room temperature. Cells were then blocked in blocking buffer (3% w/v BSA in 1xPBST) for 30 min at room temperature. After blocking, cells were incubated overnight at 4°C with mouse anti-Cas9 (Cell Signaling Technology; 14697, 1:250 dilution) and rabbit anti-tubulin (abcam; 52866, 1:400 dilution) diluted in blocking buffer. Cells were washed four times with 1xPBST, then incubated for 1 h at room temperature with goat anti-mouse AlexaFluor® 647-conjugated antibody (abcam; 150115, 1:500 dilution), goat anti-rabbit AlexaFluor® 488-conjugated antibody (abcam; 150077, 1:500 dilution), and 1 µM DAPI diluted in blocking buffer. Cells were washed three times with 1xPBST and two times with PBS before imaging using an Opera Phenix High-Content Screening System (PerkinElmer). Images were acquired using a 20x water immersion objective in a confocal mode. Automated image analysis was performed using the Harmony software (PerkinElmer). The normalized cytoplasmic intensity was determined by calculating the ratio of the mean cytoplasmic intensity of Cas9 signal per cell to the mean cytoplasmic intensity of tubulin signal per cell.

Negative-stain transmission electron microscopy

[0480] Negative-stain TEM was performed at the Koch Nanotechnology Materials Core Facility of MIT. BE-eVLPs were centrifuged for 5 min at 15,000 g to remove debris. From the clarified supernatant, 10 µL of sample and buffer containing solution was added to 200 mesh copper grid coated with a continuous carbon film. The sample was allowed to adsorb for 60 seconds after which excess solution was removed with kimwipes. 10 µL of negative staining solution containing 1% aqueous phosphotungstic acid was added to the TEM grid and the stain was immediately blotted off with kimwipes. The grid was then air-dried at room

temperature in the chemical hood. The grid was then mounted on a JEOL single tilt holder equipped within the TEM column. The specimen was cooled down by liquid-nitrogen and then observed using JEOL 2100 FEG microscope at 200kV with a magnification of 10,000-60,000. Images were taken using Gatan 2kx2k UltraScan CCD camera.

BE-eVLP protein content quantification

[0481] For protein quantification, BE-eVLPs were lysed in Laemmli sample buffer as described above. The concentration of BE protein in purified BE-eVLPs was quantified using the FastScanTM Cas9 (*S. pyogenes*) ELISA kit (Cell Signaling Technology; 29666C) according to the manufacturer's protocols. Recombinant Cas9 (*S. pyogenes*) nuclease protein (New England Biolabs; M0386) was used to generate the standard curve for quantification. The concentration of MLV p30 protein in purified BE-eVLPs was quantified using the MuLV Core Antigen ELISA kit (Cell Biolabs; VPK-156) according to the manufacturer's protocols. The concentration of VLP-associated p30 protein was calculated with the assumption that 20% of the observed p30 in solution was associated with eVLPs, as was previously reported for MLV particles (Renner *et al.*, 2020). The number of BE protein molecules per VLP was calculated by assuming a copy number of 1800 molecules of p30 per eVLP, as was previously reported for MLV particles (Renner *et al.*, 2020). The same analysis was used to determine VLP titers for all therapeutic application experiments. The same analysis was used to determine eVLP titers for all therapeutic application experiments.

BE-eVLP sgRNA extraction and quantification

[0482] RNA was extracted from BE-eVLPs using the QIAmp Viral RNA Mini Kit (Qiagen; 52904) according to the manufacturer's protocols. Extracted RNA was reverse transcribed using SuperScriptTM III First-Strand Synthesis SuperMix (Thermo Fisher Scientific; 18080400) and an sgRNA-specific DNA primer (Table 2) according to the manufacturer's protocols. qPCR was performed using a CFX96 Touch Real-Time PCR Detection System (Bio-Rad) with SYBR green dye (Lonza; 50512). The amount of cDNA input was normalized to MLV p30 content, and the sgRNA abundance per eVLP was calculated as $\log_2[\text{fold change}] (\Delta C_q)$ relative to v1 eVLPs.

Cell viability assays

[0483] Cell viability was quantified using a Promega CellTiter-Glo luminescent cell viability kit (Promega; G7570). 4×10^4 cells (for HEK293T and NIH 3T3) and 2.5×10^4 cells (for RDEB patient fibroblasts) were seeded in 250 μ L of media per well. The cells were allowed to adhere for 16-18 h before treatment with BE-eVLPs. After 48 h of transduction, 100 μ L of

CellTiter-Glo reagent was added to each well in the dark. Cells were incubated for 10 min at room temperature and the 80 μ L of solution was transferred into black 96-well flat bottom plates (Greiner Bio-one; 655096), and the luminescence was measured on a M1000 Pro microplate reader (Tecan) with a 1-second integration time. Cells treated with Opti-MEM were defined as 100% viable. The percentage of viable cells in BE-eVLP treated wells was calculated by normalizing the luminescence reading from each treatment well to the luminescence of PBS treated cells.

Table 2

Description	Forward primer sequence	Reverse primer sequence
qPCR detection of sgRNA	ACACTCTTTCCCTACACGACGCTC TTCCGATCTNNNGTTTATCACAG GCTCCAGGAAG (SEQ ID NO: 316)	TGGAGTTCAGACGTGTGCTCTTCCGATCT GGTGCCACTTTTCAAGTTGATAAC (SEQ ID NO: 318)
qPCR detection of BE- encoding DNA	ACGAGCACATTGCCAATCTG (SEQ ID NO: 317)	GCCATTTCGATCACGATGTTC (SEQ ID NO: 319)

BE-VLP transduction in cell culture and genomic DNA isolation

[0484] Cells were plated for transduction in 48-well plates (Corning) at a density of 30,000–40,000 cells per well. After 20–24 h, eVLPs were added directly to the culture media in each well. 48–72 h post-transduction, cellular genomic DNA was isolated as previously reported (Doman *et al.*, 2020). Briefly, cells were washed once with PBS and lysed in 150 μ L of lysis buffer (10 mM Tris-HCl pH 8.0, 0.05% SDS, 25 μ g mL⁻¹ Proteinase K (Thermo Fisher Scientific)) at 37 °C for 1 h followed by heat inactivation at 80°C for 30 min.

Plasmid transfections

[0485] Plasmid transfections were performed as described previously (Doman *et al.*, 2020). Plasmids were prepared for transfection using a PlasmidPlus Midi Kit (Qiagen) with endotoxin removal. HEK293T cells were plated for transfection in 48-well plates (Corning) at a density of 40,000 cells per well. After 20–24 h, cells were transfected with 1 μ g total DNA using 1.5 μ L of Lipofectamine 2000 (Thermo Fisher Scientific) per well according to the manufacturer's protocols. Unless otherwise specified, 750 ng of base editor plasmid and 250 ng of guide RNA plasmid were co-transfected per well. Genomic DNA was isolated from transfected cells at 72 h post-transfection as described above.

High-throughput sequencing of genomic DNA

[0486] Genomic DNA was sequenced as described above. Following genomic DNA isolation, 1 μ L of the isolated DNA (1–10 ng) was used as input for the first of two PCR reactions. Genomic loci were amplified in PCR1 using PhusionU polymerase (Thermo Fisher

Scientific). PCR1 primers for genomic loci are listed in Table 3 under the HTS_fwd and HTS_rev columns. PCR1 was performed as follows: 95 °C for 3 min; 30–35 cycles of 95 °C for 15 s, 61 °C for 20 s, and 72 °C for 30s; 72°C for 1 min. PCR1 products were confirmed on a 1% agarose gel. Then, 1 µL of PCR1 was used as an input for PCR2 to install Illumina barcodes. PCR2 was conducted for 9 cycles of amplification using a Phusion HS II kit (Life Technologies). Following PCR2, samples were pooled and gel purified in a 1% agarose gel using a Qiaquick Gel Extraction Kit (Qiagen). Library concentration was quantified using the Qubit High-Sensitivity Assay Kit (Thermo Fisher Scientific). Samples were sequenced on an Illumina MiSeq instrument (paired-end read, read 1: 200–280 cycles, read 2: 0 cycles) using an Illumina MiSeq 300 v2 Kit (Illumina).

Table 3

Name	Protospacer sequence	HTS_fwd	HTS_rev	Amplicon
HEK2	GAACACA AAGCATA GACTGC (SEQ ID NO: 320)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNNCCAG CCCCATCT GTCAAAC (SEQ ID NO: 356)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTTG AATGGATTCCT TGGAAACAAT GA (SEQ ID NO: 394)	TGAATGGATTCCTTGGAAACAATGATAACA AGACCTGGCTGAGCTAACTGTGACAGCAT GTGGTAATTTTCCAGCCCGCTGGCCCTGTA AAGGAAACTGGAACACAAAGCATAGACTG CGGGGCGGGCCAGCCTGAATAGCTGCAAA CAAGTGCAGAATATCTGATGATGTCATACG CACAGTTTGACAGATGGGGCTGG (SEQ ID NO: 432)
HEK3	GGCCCAG ACTGAGC ACGTGA (SEQ ID NO: 321)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNNATGT GGGCTGCC TAGAAAGG (SEQ ID NO: 357)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTCC CAGCCAACT TGTCAACC (SEQ ID NO: 395)	ATGTGGGCTGCCTAGAAAGGCATGGATGA GAGAAGCCTGGAGACAGGGATCCCAGGG AAACGCCCATGCAATTAGTCTATTTCTGCT GCAAGTAAGCATGCATTTGTAGGCTTGATG CTTTTTTCTGCTTCTCCAGCCCTGGCCTG GGTCAATCCTTGGGGCCAGACTGAGCAC GTGATGGCAGAGGAAAGGAAGCCCTGCTT CCTCCAGAGGGCGTTCGAGGACAGCTTTT CCTAGACAGGGGCTAGTATGTGCAGCTCCT GCACCGGGATACTGGTTGACAAGTTTGGCT GGG (SEQ ID NO: 433)
BCL11A enhancer	TTTATCA CAGGCTC CAGGAA (SEQ ID NO: 322)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNGCCA GAAAAGAG ATATGGCAT C (SEQ ID NO: 358)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTAG AGAGCCTTCC GAAAGAGG (SEQ ID NO: 396)	GCCAGAAAAGAGATATGGCATCTACTCTTA GACATAACACACCAGGGTCAATACAACCTTT GAAGCTAGTCTAGTGCAAGCTAACAGTTG CTTTTATCACAGGCTCCAGGAAGGGTTTGG CCTCTGATTAGGGTGGGGCGTGGGTGGG GTAGAAGAGGACTGGCAGACCTCTCCATC GGTGCCGTTTGCCCAGGGGGGCCTCTTT CGGAAGGCTCTCT (SEQ ID NO: 434)
mDnmt1	AACAGCT CTGAACG AGACCC (SEQ ID NO: 323)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNNCCTT CGGGCATA GCATGGTC (SEQ ID NO: 359)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTTA TATGCCTCGGC ATCGGTCC (SEQ ID NO: 397)	TATATGCCTCGGCATCGGTCCCCGCCCTCA CCCCACCCTGCGTGGCACCTACCGCCTGC GGACATGGTCCGGGAGCGAGCCTGCCGGG GAGGCAAGCGCAGGCACTCGGGCTGGAG CTGTTTCGCGCTGGCATCTTGACAGTTGACG ACGACAGAACAGCTCTGAACGAGACCCCG GCTTTTTTCGCGCGCGCGGAAACCAATTGG GAGGGGGCGGCGCAAGCGGAAGCAGCAT GTACCACACAGGGCAAGAGAGTGGGGGA

				AGACCATGCTATGCCCCGAAGG (SEQ ID NO: 435)
BCL11A off-target 1	CCTATCACTGGCTCAGGAA (SEQ ID NO: 324)	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNACCTGTGGGCATCCTGAGTGC (SEQ ID NO: 360)	TGGAGTTCAGACGTGTGCTCTTCCGATCTTACGGCCCCAC TCCTCTCA (SEQ ID NO: 398)	ACCTGTGGGCATCCTGAGTTGCTTCTGATGTCCCACCCATCACCTTGACCTGCTCAGAGCAGAGCATTGTTCTGAAATCTGAGGCATTGTCTGCCCCTGGCCTATCACTGGCTCCAGGAAGGGCCTAGTGTCTCTGACCAGCTCTAGATCACCTCCTCCTCCTGAGCCCTGTACGTTGCCAGGCTGATGAGAGGAGTGGGGCCGTGA (SEQ ID NO: 436)
BCL11A off-target 2	CTTATCATAGGCCCCAGGAA (SEQ ID NO: 325)	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCTTGCGCAGTTCTGTGTATG (SEQ ID NO: 361)	TGGAGTTCAGACGTGTGCTCTTCCGATCTTACATGCTGTGA GAAAATGAAGTGT (SEQ ID NO: 399)	CTTGGCGCAGTTCTGTGTATGGATAATCTTACAGAATCGCTACTCTCCCTCTCCTTTGAGCTGGCCTAGCTTTGGCTTATCATAGGCCCCAGGAAAGGCCAGGGGACTGGGGTACCGGT TAGAGGGATATAAAAGTTTATTCTGCCTTG TACGTATGTTTAAATTGATTAGAACACTTCAT TTTCTCACAGCATGTA (SEQ ID NO: 437)
HEK2 off-target 1	GAACACAATGCATAGATTGC (SEQ ID NO: 326)	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNGTGTGGAGAGTGAAGTAAGCC A (SEQ ID NO: 362)	TGGAGTTCAGACGTGTGCTCTTCCGATCTACGGTAGGATGATTTTTCAGGCA (SEQ ID NO: 400)	GTGTGGAGAGTGAGTAAGCCAGAACACAA TGCATAGATTGCCGGTAAATAGGTTTAGATT CATCCATTTTAAAAAATGGTGTGGGAGCA TTAATATGTATATAGTAGATATGGAAAAAT GATTCTCATAATAACTGACATTTCTGTTTCA CAAGAAAATTATTTTACATTATATGTATATTT TACATAAATTATACATAGTCATTTAAAAAGC TCAAATAGTGCAAAAACAATATGGAGAATT GCCTGAAATCATCTACCGT (SEQ ID NO: 438)
HEK3 off-target 1	CACCCAGACTGAGCACGTGC (SEQ ID NO: 327)	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNTCCCCTGTTGACCTGGAGAA (SEQ ID NO: 363)	TGGAGTTCAGACGTGTGCTCTTCCGATCTCACTGTACTTGCCCTGACCA (SEQ ID NO: 401)	TCCCCTGTTGACCTGGAGAAGCATGAACCAGTCAAAAAGTTTAAAGACAAGAGCATTAACTGCACCAAGTGGGCAGCTCAGCTCAGAC ACCAGTAGCGTGGGCACCCAGACTGAGCA CGTGCTGGAGCCCAAGAAATGCAGAGACC TGTGCACCTCTGGTCAGGGCAAGTACAGT G (SEQ ID NO: 439)
HEK3 off-target 2	GACACAGACTGGGCACGTGA (SEQ ID NO: 328)	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNTTGGTGTGACAGGGAGCAA (SEQ ID NO: 364)	TGGAGTTCAGACGTGTGCTCTTCCGATCTCTGAGATGTGGG CAGAAGGG (SEQ ID NO: 402)	TTGGTGTGACAGGGAGCAACTTCACAGTCCCAGGCATCAGGACACAGACTGGGCACGTGAGGGAAAGCCCAAGGGAGAGGACTGGT GTAATCGAGGCTGACTCCACTTTTAAATGTTGACTGATGATAGGTTTCAAGTCTCACTAA GTCTCCTTCCCCTTCTGCCACATCTCAG (SEQ ID NO: 440)
HEK3 off-target 3	AGCTCAGACTGAGCAAGTGA (SEQ ID NO: 329)	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNTGAGAGGGAACA GAAGGGCT (SEQ ID NO: 365)	TGGAGTTCAGACGTGTGCTCTTCCGATCTGTCCAAAGGCCAAGAACCT (SEQ ID NO: 403)	TGAGAGGGAAACAGAAGGGCTAAGACTAA AAGGAACAGAGGAGTTTCATAGTGAGCGGT AAAGAGCTCAGACTGAGCAAGTGAGGGG CTCAGCCTCCCATGGAGGACAGGGGGCTG GGGCCCCTGGCTGATGTCTGGACTGAAGC CCCCACGCCCAGAGGTTCTTGGGCCTTTG GAC (SEQ ID NO: 441)
dSaCas9 R-loop 1	GTGGTAGACAGCAT	ACACTCTTTCCCTACAC	TGGAGTTCAGACGTGTGCTC	TGGTGGAGTGCTCTGTGTTGTCTTTATAA ACCCAGATGAGAGGATGAAGGCAACAAGC

	GTGTCCT A (SEQ ID NO: 330)	CGACGCTC TTCCGATCT NNNNTGCA GTCTCCTG CTTCTCTG (SEQ ID NO: 366)	TTCCGATTGGT GGAGTGCTCT GTGTTTG (SEQ ID NO: 404)	TTCTGTACCAACATACATGCCCCCTTTGCCTC AAGTCTGGTTATTTTAGGGGGATGCTAGGT TGCTTTGGGTCTACCTTACTGAGAAAATGG CCCCAGGTCATTGTTCATGTCCAGTTGTGGT AGACAGCATGTGTCCTAAAGGGTATATTCA CATGCATGTGCAAAAATACAGGGGTCCTTC TAACCTATCACAGAGAAGCAGGAGACTG C (SEQ ID NO: 442)
dSaCas9 R- loop 2	ATTTACA GCCTGGC CTTTGGG G (SEQ ID NO: 331)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNNGGAC ATTTCACC GCAAAATG (SEQ ID NO: 367)	TGGAGTTCAG ACGTGTGCTC TTCCGATGCTA CAGAAAGGTC AGCAGC (SEQ ID NO: 405)	GCTACAGAAAGGTCAGCAGCTATATTTAAC CTCAGACCAGGGTGCGGTGGGAGATCTGG TTTCCGGAAGACGGAATGGGGAGAAGGGC AGGTTCCCCGAGGCGCCAGACACCCAAT CCTCCCGGTGACATTTACAGCCTGGCCTTT GGGGTCGGGTCAACGCTAGGCTGGCAGGG GAAGGGCGGGGCCGTGAGGTGAGCCGGC GCTGCAGGAAGGGGCCACCACCAGAGGG GCCATTTGCGGTGGAATGTCC (SEQ ID NO: 443)
dSaCas9 R- loop 3	GTGTCAG GTAATGT GCTAAAC A (SEQ ID NO: 332)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNNCTGC ACCTAGCC TCCATGTC (SEQ ID NO: 368)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTGC TGTGGCATCC AGAGACAT (SEQ ID NO: 406)	AAGTGTTTCAGCTGCTTTTCTTTTCATTTATTC CACATATAATTACTATAATTGCTAAACATTT ATTTAGTGTTCAGGTAATGTGCTAAACAGAG AGTTACTGCTCAGACATGTAATAATAATAA ATAACACATCAAATAACCATAACCATTTTAAG CTGTAGTATTATGAAGGGAAATCTGGAGCA AAGAGAATAGACTGTAGGGAAACCAGTTA AGAAATAGGACATGGAGGCTAGGTGCAG (SEQ ID NO: 444)
dSaCas9 R- loop 4	GGTGGAG GAGGGTG CATGGGG T (SEQ ID NO: 333)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNNGGAG GTGGAGAG AGGATGT (SEQ ID NO: 369)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTTC CTGAGGTCTA GGAACCCG (SEQ ID NO: 407)	TTTGCTTATCCAGAAAAGGGAGTGATTGCT TCCAGGGGCTCAGGGGAATAAATCATAG AATCCTGGACAAGGTTTGAAGGACAGGTA GGATTTGGGTGGGTGGAGGAGGGTGCATG GGGTCAGAATTGTAACCGAAAACCTCATTC AGGTGGATAGAGAAAATTTCTAGTGTTGTT GTTTTTAAACTATTTGGGGGACTGGCACAG ACCTTTTTTGAATACCTGATGGGCTCACAT TTCTGTCTGAATCCCAG (SEQ ID NO: 445)
dSaCas9 R- loop 5	TCTGCTT CTCCAGC CCTGGC (SEQ ID NO: 334)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNNATGT GGGCTGCC TAGAAAGG (SEQ ID NO: 357)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTCC CAGCCAACT TGTCAACC (SEQ ID NO: 395)	ATGTGGGCTGCCTAGAAAAGGCATGGATGA GAGAAGCCTGGAGACAGGGATCCCAGGG AAACGCCCATGCAATTAGTCTATTTCTGCT GCAAGTAAGCATGCATTTGTAGGCTTGATG CTTTTTTCTGCTTCTCCAGCCCTGGCCTG GGTCAATCCTTGGGGCCAGACTGAGCAC GTGATGGCAGAGGAAAGGAAGCCCTGCTT CCTCCAGAGGGCGTCGCAGGACAGCTTTT CCTAGACAGGGGCTAGTATGTGCAGCTCCT GCACCGGGATACTGGTTGACAAGTTTGGCT GGG (SEQ ID NO: 433)
dSaCas9 R- loop 6	GATGTTT CAATCAG TACGCA (SEQ ID NO: 335)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNNCAT GCAGAGAG GCGTATC (SEQ ID NO: 370)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTG GGGTCCCAGG TGCTGAC (SEQ ID NO: 408)	CATTGCAGAGAGGCGTATCATTTTCGCGGAT GTTCCAATCAGTACGCAGAGAGTCGCCGT CTCCAAGGTGAAAGCGGAAGTAGGGCCTT CGCGCACCTCATGGAATCCCTTCTGCAGCA CCTGGATCGCTTTTCCGAGCTTCTGGCGGT CTCAAGCACTACCTACGTCAGCACCTGGG ACCCC (SEQ ID NO: 446)
COL7A1 (R185X)	CAACTCA CTTCAGC TCCTCA	ACACTCTT TCCCTACA CGACGCTC	TGGAGTTCAG ACGTGTGCTC TTCCGATCTAG	CTCTCCGGGAAACGAGGGGCGAGTAGTGTC CTCAAGATGCTGAAGTCATTGACGAAGAA GAAGAAGTCACTGGTGGGCTGTGAGGCAA

	(SEQ ID NO: 336)	TTCCGATCT NNNNCTGG TGGACACA GCTG (SEQ ID NO: 371)	CAGTCGTGCA CAC (SEQ ID NO: 409)	CTCACTTCAGCTCCTCAGGGTCAGCATTCT TGATCCCTGAAGTGACGACCCATCAGGAC TCAGTCACCCACATGCTCTCTGACTGCCCC CACCCCCAGCTGACCTGTCACTCCTGTCTC GGTCCTTACCCACAGCAAATAGCTTGACCC CCTGCCCTTCAGCCTTTGGGCAGCTGTGT CCACCAG (SEQ ID NO: 447)
Idua (W392X)	ACTCTAG GCAGAG GTCTCAA (SEQ ID NO: 337)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNNTTAG GGTAGGAA GCCAGATG CTA (SEQ ID NO: 372)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTGT GTGCGTGGGT GTCATC (SEQ ID NO: 410)	TTAGGGTAGGAAGCCAGATGCTAGGTATGA GAGAGCCAACAGCCTCAGCCCTCTGCTTG GCTTATAGATGGAGAACAACCTCTAGGCAGA GGTCTCAAAGGCTGGGGCTGTGTTGGACA GCAATCATACAGTGGGTGTCTGGCCAGCA CCCATCACCCCTGAAGGCTCCGCAGCGGCC TGGAGTACCACAGTCTCTATCTACACTAGT GATGACACCCACGCACAC (SEQ ID NO: 448)
B2M	CTTACCC CACTTAA CTATCT (SEQ ID NO: 338)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNNTGTC TTTCAGCA AGGACTGG TCT (SEQ ID NO: 373)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTG GGACTCATTC AGGGTAGTAT GGC (SEQ ID NO: 411)	GGGACTCATTACAGGGTAGTATGGCCATAGA CCTTTTTTATATCAAAGCAGCTTTATGATAT GACTACTCATACACAACCTTTCAGCAGCTTA CAAAAGAATGTAAGACTTACCCCACTTAAC TATCTTGGGCTGTGACAAAGTCACATGGTT CACACGGCAGGCATACTCATCTTTTTTCAGT GGGGGTGAATTCAGTGTAGTACAAGAGAT AGAAAGACCAGTCCTTGCTGAAAGACA (SEQ ID NO: 449)
CIITA	CACTCAC CTTAGCC TGAGCA (SEQ ID NO: 339)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNNAGGC ATGCAAGT TTGGTCCT GA (SEQ ID NO: 374)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTCC CTGCAGCCAG CACGATGT (SEQ ID NO: 412)	CCCTGCAGCCAGCACGATGTGGGTTCCT GCGCTCTGCAGCCCCCAGCTCAGCACCT GACCGGTATCCGGGGCCCCACTCACCTTAG CCTGAGCAGGGATGCAGCGAGCGAAGGCA GGGCCTCGGCGAGTTTGTAGGCACCCAGG TCAGTGATGTTGTTCTGGGACAGACTGCG GGGACACAGTGAGGGGGAGGGCTCAGGA CCAACTTGCATGCCT (SEQ ID NO: 450)
mPcsk9	CCCATAC CTTGGAG CAACGG (SEQ ID NO: 340)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNNGGCT GCACTTAG AGACCACC (SEQ ID NO: 375)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTAT GAAGAGCTGA TGCTCGCC (SEQ ID NO: 413)	GGCTGCACTTAGAGACCACCAGACGGCTA GATGAGCAGAGAAGACCCCCGAAGAGCAT CACCCCAACCCCAAAGCAACGCCGTTGCC TGGCACCCATACCTTGGAGCAACGGCGGA AGGTGGCGGTGGCCACATGTGCGGCCTCA TCAGCCAGGCCATCCTCCTGGGACGGGAG GGCGAGCATCAGCTCTTCAT (SEQ ID NO: 451)
mPcsk9 OT1	CCCCTAC CTTGGGG CAACAG (SEQ ID NO: 341)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNNAAGT ATGTTGGG ACCCTTGG CTGG (SEQ ID NO: 376)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTTG GCCTGTTCTAC TGACTATGGG G (SEQ ID NO: 414)	AAGTATGTTGGGACCCTTGGCTGGGCTTCT TGCCCTCTCTAGAACCAAGATGTCATTCT GCACACCAAGAGCTACCCCTACCTTGGGG CAACAGTGGAAGCCATGGCTGGAGAAAGC AAACAATTCCTGAAGGTGACAGATTCTCCT GGGAAGGGACTTAGCCCCATAGTCAGTAG AACAGGCCA (SEQ ID NO: 452)
mPcsk9 OT2	ACCATAC CTAAGAG CAAACCT (SEQ ID NO: 342)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNNGACA GACACAGG GAAGCCTT	TGGAGTTCAG ACGTGTGCTC TTCCGATCTAA CCTTCCAGGA GAGAGAAACC TGT (SEQ ID NO: 415)	GACAGACACAGGGAAGCCTTGGGGAGCC GGAGGCTTGGCCAGGAGCTCAGGGGTCCC TGGGCAGATGCTCACACTGGGCAGAAGGT CACACCATACCTAAGAGCAAACCTGGGGCC CAAACGACTGAGTGTTGCTGAGAGCCATC CTTGGCTCATTCTCAAAAAACAGGTTTCTC TCTCCTGGAAGGTT (SEQ ID NO: 453)

		GGG (SEQ ID NO: 377)		
mPcsk9 OT3	CCCACCC TTTGGAG AACGG (SEQ ID NO: 343)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNNTGGC AAGGGACA GGGTCAGC (SEQ ID NO: 378)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTAG CTGGTGGCAG AGGTGTGG (SEQ ID NO: 416)	TGGCAAGGGACAGGGTCAGCTCTTCACTC CCATTCCATCTGGGGCAGCTCACCTGCATC CAAGCCAATAGAGACAGCCCTACTGTGTT GCTCAGTTGAGGTACGGGGCCCACCCTTT GGAGAACGGTGGGGGTGGGAGCTATGCCA ACACTTCTGCTCTAACACCCTCACAGCTAG CTACCCACACCTCTGCCACCAGCT (SEQ ID NO: 454)
mPcsk9 OT4	CCCAGCC TTGGGGC AACGG (SEQ ID NO: 344)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNNTTCA AGCAATCA CGAGACAC TCAG (SEQ ID NO: 379)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTCC CACCACCCAG CAGCTTTATTG (SEQ ID NO: 417)	TTCAAGCAATCACGAGACACTCAGTTTGG ATCCCCAGAGCCACATAAAAGATCAGAC ACAGAGTGCATGCCTGTAAACCCAGCCTTG GGGCAACGGAGGCTCTGAAGCTCGTCGGT TAGCCAGCTGAAGCATATCCATGAGGTTTA GTGTTGGAGCCTGTCTCAATAAAGCTGCTG GGTGGTGGG (SEQ ID NO: 455)
mPcsk9 OT5	CACATAT CTAGGAG CAAGG (SEQ ID NO: 345)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNNTCTC AGGCGACC TGGTTTCT GC (SEQ ID NO: 380)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTTC TGCCAGATGC GTCCGATCA (SEQ ID NO: 418)	TCTCAGGCGACCTGGTTTCTGCAAAGGGC AGGGTTGGCTTTATGCTGAGTCCCTACAGAT CTTAGACCCCCCCCCCAAACCTAAACACA TATCTAGGAGCAAGGAGGGGTCATGAAAA GATAGAGCCTGCTTTGGCAGACTATAGAAC AGAACACTAAGGATTTAACTTACTAGTGAA ATGATCGGACGCATCTGGCAGA (SEQ ID NO: 456)
mPcsk9 OT6	CCCACAC CCGGAGC AACGG (SEQ ID NO: 346)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNNGCCA GCCCTGCC TGGAAGTT AG (SEQ ID NO: 381)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTTG ACCTCCGGGA TTCTCAGCCC (SEQ ID NO: 419)	GCCAGCCCTGCCTGGAAGTTAGCCATGGA GGATGGAGCTGAACTTGACCTTTGCGGTTT ACAGCCACACCCGGAGCAACGGGGAGG TCGTCTGAGCCAGTCAGTCGTTTGGTTG CAAAGAAGCTTTTAATAAGGGAAGTTTCA GTCATGGAATGAGAGGTGAGGTGAAGTGG GCTGAGAATCCCGGAGGTCA (SEQ ID NO: 457)
mPcsk9 OT7	TCCATAC CCGGAGC AACGA (SEQ ID NO: 347)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNNGCTT CCTGTCTG CAATTGGG GTCT (SEQ ID NO: 382)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTAG TAGGTTGCGG GGCTCAGGA (SEQ ID NO: 420)	GCTTCCTGTCTGCAATTGGGGTCTTTGTTG TCCTTCTGGCTGTCTTCTCCTCTTCATCAA CAAGAAGCTATGCTCTGAAAACCTAAGAG GGCATCCATACCCGGAGCAACGAGGGAAG AGAAAGCACTCGAGAGACAAGACTGGAG GCCACACAGGAAGTGGTAAGCACCATGCT TTATGTTTTCTGAGCCCCGCAACCTACT (SEQ ID NO: 458)
mPcsk9 OT8	TTCATCC TTGGAGC AACGG (SEQ ID NO: 348)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNNCCAG CAGGTCCC CAGTGACG (SEQ ID NO: 383)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTTA CCCACCTGGG TGTGTCCA (SEQ ID NO: 421)	CCAGCAGGTCCCCAGTGACGCAAGCCAGC AGGGGGTGGGAAGCTTCAGGAGAAAAGG ACATGGAGCAGTAGGGTATGACATTCAAA GCCTGACAGCGTCTCTACCAGCCCTTCATC CTTGAGCAACGGTGAGATGAACATTATG TTCATACTGCAGAGTTGAACAGAATCCAGA ACAGCCAGCCTTTTGGAGCTACATAACAAA AGTATCATGTGCACATGTGGACACACCCAG GTGGGTA (SEQ ID NO: 459)
mPcsk9 OT9	TCTGTAC CATGGAG CAAAGG	ACACTCTT TCCCTACA CGACGCTC	TGGAGTTCAG ACGTGTGCTC TTCCGATCTAC	AACCTCCACGGGGGTATCTGAGGTCTTCTG CTGTAGTGTGTCTTTTCAGTCATCAATAAC ATGGGCAGGTACCATCCCCCTCCGATGTGGG

	(SEQ ID NO: 349)	TTCCGATCT NNNNAACC TCCACGGG GGTATCTG AGG (SEQ ID NO: 384)	CTGGCAAGTG GGGTACTGG (SEQ ID NO: 422)	CGAGTACCACAAGTTTGAAGGTCACAGG GCTGCTCTGTACCATGGAGCAAAGGCGGA AAGGAAACCTTGGGTGTCTGATGCATTGG AACCCAGTACCCCACTTGCCAGGT (SEQ ID NO: 460)
mPcsk9 OT10	ACCATAA CCAAGAG CAACAG (SEQ ID NO: 350)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNGTCT AAATGGGC AAGCAATC CCCT (SEQ ID NO: 385)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTCC AGGATCCAC AGGGTCCTTC T (SEQ ID NO: 423)	GTCTAAATGGGCAAGCAATCCCCTGTCCAG GGTCGATTCAGGGCTGTCTGTGAGAAGTC TCGGTGTCTTATGGAGGATTTCTACTGATG AGTAAACACCATAACCAAGAGCAACAGG GGAGGGAAGGGTCTCCTGCAGCTTACATC TGACAGTCATCCAGGGTAGTCAGTGAAGG GACTCTCTCAGAAGGACCCTGTGGGATCC TGG (SEQ ID NO: 461)
mPcsk9 OT11	TCCATAA CTCAGAG CAACAG (SEQ ID NO: 351)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNNTCCC CAGAGCCC AGGGAATA TCA (SEQ ID NO: 386)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTTG TTGCTCCGAT GGAAGGATGG G (SEQ ID NO: 424)	TCCCCAGAGCCCAGGGAATATCATGGGGG AATATAAGAGCTATAGGATGAGAATTGGTG CTACAGAATGCTGTTTGTGGATAAGACATG GCTGATGCATCCATAACTCAGAGCAACAGT GGTGACTTGCTCAAGACCTTCACAAGACT GAGCTGTCAACCTTCTACCCTGGATGGAAG ACGGGATGGTAAGATCCCATCCTTCCATCG GAGCAACA (SEQ ID NO: 462)
mPcsk9 OT12	GCCATAC CCTGGGG CAGCAG (SEQ ID NO: 352)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNNTGTG GAACCCAC CCCCGATA CA (SEQ ID NO: 387)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTAG TGCTGATGGG CAAGGCATTT G (SEQ ID NO: 425)	TGTGGAACCCACCCCGATACACACACAC CTTAAGTCGTACCTCTCTCAACATGTCTGC TGAAGCCACCTGCCCGCAGAGTAAGCA GGCGCCATACCCTGGGGCAGCAGTGGAGG CTATGATTTAGAATAACTGTGGTCCGGTCTC TCTAACATTTGCCGCTGTATTCAATTCTAAGT TTAATGAGGGACAAATGCCTTGCCCATCAG CACT (SEQ ID NO: 463)
mPcsk9 OT13	GCAACAC CTTGGAG CAACTG (SEQ ID NO: 353)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNNCCAC CAGAAGCG CCCCAGAA (SEQ ID NO: 388)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTG GGGAATCGCC TCCACTGCC (SEQ ID NO: 426)	CCACCAGAAGCGCCCCAGAACTCCTTGCT GGCTAGTTGGCTCTCATCAGCTCAGCCTG CCCAACTCAGCGTGGGGCTGTAGGTGCAA CACCTTGGAGCAACTGAGGTATCAACAGC AGAGATAGAGATGGAGGAAGCTGCAGCAA CAGAGGCAGTGGAGGCGATTCCCC (SEQ ID NO: 464)
mPcsk9 OT14	GACATCC TTGGAGC AACTG (SEQ ID NO: 354)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNGTTC TTATTGGCC AGGGAGCC TT (SEQ ID NO: 389)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTCT CCCCAAGTGA CAGGAACCAC G (SEQ ID NO: 427)	GTTCTTATTGGCCAGGGAGCCTTTCTGCAG TTCTTTGTAAATCCAGCTAAAATGCAAACA CTGACATCAATCATTTGAAATGAGGTGGCT GTCAGGTCCTCAGACATCCTTGAGGCAAC TGTGGGTGAGTATTCCTGATGGGAATTTTC TCTCTTCATCCAGGAGTGAGGGCTCACTTG GTGCCCAACCTACAGGCTGGGTGGAGGGC TGGGCACCACGTGGTTTCTGTCACTTGGG GAG (SEQ ID NO: 465)
Rpe65 (R44X) gDNA	ACATCAG AGGAGA CTGCCAG (SEQ ID NO: 355)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNAGCT GACAAATA ACAAATAG GCACA	TGGAGTTCAG ACGTGTGCTC TTCCGATCTG GCTCTACTCTG GTGAGGTCAG (SEQ ID NO: 428)	GGCTCTACTCTGGTGAGGTCAGTCATGGAC TTACCTTCTGTGGTATGTGACATGGCCCTC CTTGAAGTCAAACCTTGTGAAAAGGGCTT GTCCATCAAACAGGTGATAGAAAGGCTCA GATCCAACCTCAAAGAGCCCTGGCCCACA TCAGAGGAGACTGCCAGTGAGCCAGAGG GGAATCCTGCCTGCAGCAAAGTGAGATATC AGGTGGTACTACTACTAGATTTTCTATGTG

		(SEQ ID NO: 390)		CCTATTTGTTATTTGTCAGCT (SEQ ID NO: 466)
Rpe65 (R44X) cDNA	ACATCAG AGGAGA CTGCCAG (SEQ ID NO: 355)	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNTGTC CTCACCAC TAACAGCT (SEQ ID NO: 391)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTCT TCTCAGTCATT GCTCGAACA (SEQ ID NO: 429)	CTTCTCAGTCATTGCTCGAACATAAGCATC AGTGC GGATGAATCTTCTGTGGTATGTGAC ATGGCCCTCCTTGAAGTCAAACCTTGTGCAA AAGGGCTTGTCCATCAAACAGGTGATAGA AAGGCTCAGATCCAACCTTCAAAGAGCCCT GGCCACATCAGAGGAGACTGCCAGTGAG CCAGAGGGGAATCCTGCCTGTGACATGAG CTGTTAGTGGTGAGGACA (SEQ ID NO: 467)
Mcm3ap cDNA	n/a	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNGCTT CCAAAGCC TGCGCCTG (SEQ ID NO: 392)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTCC ATGGAACTT CCTCAGCGGC (SEQ ID NO: 430)	GCTTCCAAAGCCTGCGCCTGTGTACTCTGA CTCGGACCTGGTACAGGTGGTGGACGAGC TCATCCAGGAGGCTCTGCAAGTGGACTGT GAGGAAGTCAGCTCCGCTGGGGCAGCCTA CGTAGCCGCAGCTCTGGGCGTTTCCAATGC TGCTGTGGAGGATCTGATTACTGCTGCGAC CACGGGCATTCTGAGGCACGTTGCCGCTG AGGAAGTTTCCATGG (SEQ ID NO: 468)
Perp cDNA	n/a	ACACTCTT TCCCTACA CGACGCTC TTCCGATCT NNNGCCA TCGCCTTC GACATCAT (SEQ ID NO: 393)	TGGAGTTCAG ACGTGTGCTC TTCCGATCTAA CAAGCATCTG GGGTCCAC (SEQ ID NO: 431)	GCCATCGCCTTCGACATCATCGCGCTGGCC GGCCGCGGCTGGCTGCAGTCTAGCAACCA CATCCAGACATCGTCGCTTTGGTGGAGGTG TTTCGACGAGGGCGGCGGCAGCGGCTCCT ACGACGATGGCTGCCAGAGCCTCATGGAG TACGCATGGGGACGAGCAGCTGCAGCCAC GCTTTTCTGTGGCTTTATCATCCTGTGCATC TGCTTCATTCTCTCGTTCTTCGCCCTGTGTG GACCCAGATGCTTGTT (SEQ ID NO: 469)

High-throughput sequencing data analysis

[0487] Sequencing reads were demultiplexed using the MiSeq Reporter software (Illumina) and were analyzed using CRISPResso2 (Clement *et al.*, 2019) as previously described (Doman *et al.*, 2020). Batch analysis mode (one batch for each unique amplicon and sgRNA combination analyzed) was used in all cases. Reads were filtered by minimum average quality score (Q > 30) prior to analysis. The following quantification window parameters were used: -w 20 -wc -10. Base editing efficiencies are reported as the percentage of sequencing reads containing a given base conversion at a specific position. Prism 9 (GraphPad) was used to generate dot plots and bar plots.

Assessment of off-target DNA base editing in HEK293T cells

[0488] HEK293T cells were transduced with v4 BE-eVLPs or transfected with BE-encoding plasmid as described above. To assess Cas-dependent off-target editing, cells were transfected or transduced with 1 μ L of v4 BE-eVLPs on the same day and genomic DNA was isolated 72 h post treatment in both cases. On-target and off-target loci were amplified and sequenced as described above. Orthogonal R-loop assays were performed as described previously (Doman *et al.*, 2020) to assess Cas-independent off-target editing. To allow time

for expression of SaCas9 and formation of the off-target R-loops following plasmid transfection, cells were transduced with 1 μ L of PEG-concentrated v4 BE-eVLPs at 24 h post-transfection with dSaCas9- and orthogonal sgRNA-encoding plasmids. Genomic DNA was isolated 72 h post-transfection (48 h post-transduction) and sequenced as described above. See also FIG. 12A for an experimental schematic. See also FIG. 11A for an experimental schematic.

Quantification of BE-encoding DNA

[0489] For quantifying the amount of BE-encoding DNA in BE-eVLP preparations, v4 BE-eVLPs were lysed as described above, and the lysate was used as input into a qPCR reaction with BE-specific primers (Table 2). For quantifying the amount of BE-encoding DNA in eVLP-transduced vs. plasmid-transfected HEK293T cells, DNA was isolated from cell lysate as described above and used as input into a qPCR reaction with BE-specific primers (Table 2). In both cases, a standard curve was generated with BE-encoding plasmid standards of known concentration and was used to infer the amount of BE-encoding DNA present in the original samples.

Transduction of T cells and genomic DNA preparation

[0490] Thawed cells (day 0) were rested for 24 h in basal T-cell media comprised of X-VIVO™ 15 Serum-free Hematopoietic Cell Medium (Lonza; BE02-0606F) with 10% AB human serum (Valley Biomedical; HP1022), 2 mg/mL N-acetyl-cysteine (Sigma Aldrich; A7250), 300 IU/mL recombinant human IL-2 (Peprotech ; 200-02) and 5 ng/mL recombinant human IL-7 (Peprotech ; 200-07) and 5 ng/mL IL-15 (Peprotech; 500-P15). On day 1, 50,000 cells in 50 μ L of T-cell media were plated in 96-well-plates coated with 10 μ g/cm² RetroNectin® (Clontech/Takara; catalog number T100A/B). 5 μ L (3.0×10^{10} eVLPs) of ultracentrifuge-purified v4 BE-eVLPs were used to transduce the cells on day 1 and on day 2 the cells were stimulated with Dynabeads™ Human T-Expander CD3/CD28 beads (Thermo Fisher; 11161D). Beads were added at a bead to cell ratio of 3:1 in a volume of 50 μ L. On day 3, the cells were transduced for a second time with 5 μ L (3.0×10^{10} eVLPs) of v4 BE-eVLPs in a total media volume of 200 μ L. Twenty-four hours later (day 4) the cells were resuspended in 1 mL of fresh T-cell media and re-plated in wells of a 48 well plate. On day 6 the cells were harvested, and genomic DNA was isolated using the QuickExtract™ DNA Extraction Solution (Lucigen; QE09050).

Lentiviral vector cloning and production

[0491] Lentiviral vectors were constructed via USER cloning into the lentiCRISPRv2 backbone (Addgene #135955). Lentiviral transfer vectors were propagated in NEB Stable Competent *E. coli* (New England Biolabs). HEK293T/17 (ATCC CRL-11268) cells were maintained in antibiotic-free DMEM supplemented with 10% fetal bovine serum (v/v). On day 1, 5×10^6 cells were plated in 10 mL of media in T75 flasks. The following day, cells were transfected with 6 μ g of VSV-G envelope plasmid, 9 μ g of psPAX2 (plasmid encoding viral packaging proteins) and 9 μ g of transfer vector plasmid (plasmid encoding the gene of interest) diluted in 1,500 μ L Opti-MEM with 70 μ L of FuGENE. Two days after transfection, media was centrifuged at 500 g for 5 min to remove cell debris following filtration using 0.45- μ m PVDF vacuum filter. The lentiviruses were further concentrated by ultracentrifugation with a 20% (w/v) sucrose cushion as described above for eVLP production.

AAV production

[0492] AAV production was performed as previously described (Deverman *et al.*, 2016; Levy *et al.*, 2020) with some alterations. HEK293T/17 cells were maintained in DMEM with 10% fetal bovine serum without antibiotics in 150-mm dishes (Thermo Fisher Scientific; 157150) and passaged every 2–3 days. Cells for production were split 1:3 one day before polyethylenimine transfection. Then, 5.7 μ g AAV genome, 11.4 μ g pHelper (Clontech), and 22.8 μ g AAV8 rep-cap plasmid were transfected per plate. The day after transfection, media was exchanged for DMEM with 5% fetal bovine serum. Three days after transfection, cells were scraped with a rubber cell scraper (Corning), pelleted by centrifugation for 10 min at 2,000 g, resuspended in 500 μ L hypertonic lysis buffer per plate (40 mM Tris base, 500 mM NaCl, 2 mM MgCl_2 and 100 U mL^{-1} salt active nuclease (ArcticZymes; 70910-202)) and incubated at 37 °C for 1 h to lyse the cells. The media was decanted, combined with a 5X solution of 40% poly(ethylene glycol) (PEG) in 2.5 M NaCl (final concentration: 8% PEG/500 mM NaCl), incubated on ice for 2 h to facilitate PEG precipitation, and centrifuged at 3,200 g for 30 min. The supernatant was discarded, and the pellet was resuspended in 500 μ L lysis buffer per plate and added to the cell lysate. Crude lysates were either incubated at 4 °C overnight or directly used for ultracentrifugation.

[0493] Cell lysates were clarified by centrifugation at 2,000 g for 10 min and added to Beckman Quick-Seal tubes via 16-gauge 5" disposable needles (Air-Tite N165). A discontinuous iodixanol gradient was formed by sequentially floating layers: 9 mL 15%

iodixanol in 500 mM NaCl and 1× PBS-MK (1× PBS plus 1 mM MgCl₂ and 2.5 mM KCl), 6 mL 25% iodixanol in 1× PBS-MK, and 5 mL each of 40 and 60% iodixanol in 1× PBS-MK. Phenol red at a final concentration of 1 µg mL⁻¹ was added to the 15, 25 and 60% layers to facilitate identification. Ultracentrifugation was performed using a Ti 70 rotor in an Optima XPN-100 Ultracentrifuge (Beckman Coulter) at 58,600 rpm for 2 h 15 min at 18 °C. Following ultracentrifugation, 3 mL of solution was withdrawn from the 40–60% iodixanol interface via an 18-gauge needle, dialyzed with PBS containing 0.001% F-68 using 100-kD MWCO columns (EMD Millipore). The concentrated viral solution was sterile filtered using a 0.22-µm filter. The final AAV preparation was quantified via qPCR (AAVpro Titration Kit version 2; Clontech) and stored at 4 °C until use.

Animals

[0494] Timed pregnant C57BL/6J mice for P0 studies were purchased from Charles River Laboratories (027). Wild-type adult C57BL/6J mice (000664) and pigmented *rd12* mice (005379) were purchased from the Jackson Laboratory. All mice were housed in a room maintained on a 12 h light and dark cycle with *ad libitum* access to standard rodent diet and water. Animals were randomly assigned to various experimental groups.

P0 ventricle injections

[0495] P0 ventricle injections were performed as described previously (Levy *et al.*, 2020). Drummond PCR pipettes (5-000-1001-X10) were pulled at the ramp test value on a Sutter P1000 micropipette puller and passed through a Kimwipe three times, resulting in a tip size of ~100 µm. A small amount of Fast Green was added to the BE-eVLP injection solution to assess ventricle targeting. The injection solution was loaded via front filling using the included Drummond plungers. P0 pups were anaesthetized by placement on ice for 2–3 min until they were immobile and unresponsive to a toe pinch. Then, 2 µL of injection mix (containing 2.6×10¹⁰ eVLPs encapsulating a total of 3.2 pmol of BE protein) was injected freehand into each ventricle. Ventricle targeting was assessed by the spread of Fast Green throughout the ventricles via transillumination of the head.

Nuclear isolation and sorting

[0496] Nuclei were isolated from the cortex and the mid-brain as previously described (Levy *et al.*, 2020). Briefly, dissected cortex and mid-brain were homogenized using a glass Dounce homogenizer (Sigma-Aldrich; D8938) with 20 strokes using pestle A followed by 20 strokes from pestle B in 2 mL of ice-cold EZ-PREP buffer (Sigma-Aldrich; NUC-101). Samples

were then decanted into a new tube containing an additional 2 mL of EZ-PREP buffer on ice. After 5 min, homogenized tissues were centrifuged for 5 min at 500 g at 4°C. The nuclei pellet was resuspended in 4 mL of ice-cold Nuclei Suspension Buffer (NSB) consisting of 100 µg/mL BSA (NEB; B9000S) and 3.33 µM Vybrant DyeCycle Ruby (Thermo Fisher; V10309) in PBS followed by centrifugation at 500 g for 5 min at 4°C. After centrifugation, the supernatant was removed, and nuclei were resuspended in 1-2 mL of NSB, passed through 35-µm cell strainer, followed by flow sorting using the Sony MA900 Cell Sorter (Sony Biotechnology) at the Broad Institute flow cytometry core. See FIG. 13A for example FACS gating. Nuclei were sorted into DNAdvance lysis buffer, and the genomic DNA was purified according to the manufacturer's protocol (Beckman Coulter; A48705).

Retro-orbital injections

[0497] 50 µL of VLPs (containing 4×10^{11} or 7×10^{11} VLPs) were centrifuged for 10 min at 15,000 g to remove debris. The clarified supernatant was diluted to 120 µL in 0.9% NaCl (Fresenius Kabi; 918610) right before injection. 1×10^{11} viral genomes (vg) of total AAV was diluted to 120 µL in 0.9% NaCl (Fresenius Kabi; 918610) right before injection. Anesthesia was induced with 4% isoflurane. Following induction, as measured by unresponsiveness to bilateral toe pinch, the right eye was protruded by gentle pressure on the skin, and an insulin syringe was advanced, with the bevel facing away from the eye, into the retrobulbar sinus where VLP or AAV mix was slowly injected. One drop of Proparacaine Hydrochloride Ophthalmic Solution (Patterson Veterinary; 07-885-9765) was then applied to the eye as an analgesic. Genomic DNA was purified from various tissue using Agencourt DNAdvance kits (Beckman Coulter; A48705) following the manufacturer's instructions.

Histology and staining

[0498] Liver tissue was fixed in 4% PFA overnight at 4°C. The next day, fixed liver was transferred into 1x PBS with 10 mM glycine to quench free aldehyde for at least 24 h followed by paraffinization at the Rodent Histopathology Core of Harvard Medical School. Liver paraffin block was then cut into 5 µm sections followed by hematoxylin and eosin staining for histopathological examination.

Alanine Aminotransferase (ALT) and Aspartate Aminotransferase (AST) assay

[0499] Blood was collected 7 days after injection via submandibular bleeding and allowed to clot at room temperature for 1 h. The serum was then separated by centrifugation at 2000 g for 15 min and sent to IDEXX Bioanalytics, MA, for analysis.

Serum Pcsk9 measurements

[0500] To track serum levels of Pcsk9, blood was collected using a submandibular bleed in a serum separator tube. Serum was separated by centrifugation at 2000 g for 15 min and stored at -80°C. Pcsk9 levels were determined by ELISA using the Mouse Proprotein Convertase 9/PCSK9 Quantikine ELISA Kit (R&D Systems; MPC900) following the manufacturer's instructions.

CIRCLE-seq

[0501] Circularization for *In vitro* Reporting of Cleavage Effects by sequencing (CIRCLE-seq) was performed and analyzed as described previously (Tsai *et al.*, 2017), save for the following modifications: For the Cas9 cleavage step, guide denaturation, incubation, and proteinase K treatment was conducted using the more efficient method described in the CHANGE-seq protocol (Lazzarotto *et al.*, 2020). Specifically, the sgRNA with the guide sequence "GCCCATACCTTGGAGCAACGG" (SEQ ID NO: 496) was ordered from Synthego with their standard chemical modifications, 2'-O-Methyl for the first three and last three bases, and phosphorothioate bonds between the first three and last two bases. A 5' "G" nucleotide was included with the 20-nucleotide specific guide sequence to recapitulate the sequence expressed and packaged into VLPs. The sgRNA was diluted to 9 µM in nuclease-free water and re-folded by incubation at 90°C for 5 min followed by a slow annealing down to 25°C at a ramp rate of 0.1 °C/second. The sgRNA was complexed with Cas9 nuclease (NEB; M0386T) via a 10 min room temperature incubation after mixing 5 µL of 10x Cas9 Nuclease Reaction Buffer provided with the nuclease, 4.5 µL of 1 µM Cas9 nuclease (diluted from the 20µM stock in 1x Cas9 Nuclease Reaction Buffer), and 1.5 µL of 9 µM annealed sgRNA. Circular DNA from mouse N2A cells was added to a total mass of 125 ng and diluted to a final volume of 50 µL. Following 1 h of incubation at 37°C, Proteinase K (NEB; P8107S) was diluted 4-fold in water, and 5 µL of the diluted mixture was added to the cleavage reaction. Following a 15 min Proteinase K treatment at 37°C, DNA was A-tailed, adapter ligated, USER-treated, and PCR-amplified as described in the CIRCLE-seq protocol (Tsai *et al.*, 2017). Following PCR, samples were loaded on a preparative 1% agarose gel and DNA was extracted between the 300bp and 1kb range to eliminate primer dimers before sequencing on an Illumina MiSeq. Data was processed using the CIRCLE-seq analysis pipeline and aligned to the human genome Hg19 (GRCh37) with parameters: "read_threshold: 4; window_size: 3; mapq_threshold: 50; start_threshold: 1; gap_threshold: 3; mismatch_threshold: 6; merged_analysis: True".

Amplicon sequencing of off-target sites nominated by CIRCLE-seq

[0502] It has previously been observed with exhaustively assessed ABE8e off-target sites nominated by CIRCLE-seq that off-target editing efficiency did not track well with the CIRCLE-seq read count (Newby *et al.*, 2021). However, nominated off-target sites where editing was observed shared some striking similarities. Namely, over 90.7% of the 54 off-target sites with validated off-target editing had zero mismatches or one mismatch to the guide in the 9 nucleotides proximal to the PAM. The few sites with more than 1 mismatch in this region were all edited with low efficiency (the bottom half of sites, when ranked by editing efficiency). Based on this knowledge, 14 off-target sites were chosen to be assessed in the CIRCLE-seq list that showed one or fewer mismatches in the 9 nucleotides of the protospacer proximal to the PAM to increase the chance that a true off-target site is sequenced (Table 5).

Mouse subretinal injection

[0503] Mice were anesthetized by intraperitoneal injection of a cocktail consisting of 20 mg/mL ketamine and 1.75 mg/mL xylazine in phosphate-buffered saline at a dose of 0.1 mL per 20 g body weight, and their pupils were dilated with topical administration of 1% tropicamide ophthalmic solution (Akorn; 17478-102-12). Subretinal injections were performed under an ophthalmic surgical microscope (Zeiss). An incision was made through the cornea adjacent to the limbus at the nasal side using a 25-gauge needle. A 34-gauge blunt-end needle (World Precision Instruments; NF34BL-2) connected to an RPE-KIT (World Precision Instruments, no. RPE-KIT) by SilFlex tubing (World Precision Instruments; SILFLEX-2) was inserted through the corneal incision while avoiding the lens and advanced through the retina. Each mouse was injected with 1 μ L of experimental reagent (lentivirus or eVLPs) per eye. Lentivirus titer was $>1 \times 10^9$ TU/mL as measured by the QuickTiterTM Lentivirus Titer Kit (Cell Biolabs; VPK-107-5). BE-eVLPs were normalized to a titer of 4×10^{10} eVLPs/ μ L, corresponding to an encapsulated BE protein content of 3 pmol/ μ L. After injections, pupils were hydrated with the application of GenTeal Severe Lubricant Eye Gel (0.3% Hypromellose, Alcon) and kept for recovery.

RPE dissociation and genomic DNA and RNA preparation

[0504] Under a light microscope, mouse eyes were dissected to separate the posterior eyecup (containing RPE, choroid, and sclera) from the retina and anterior segments. Each posterior eyecup was immediately immersed in 350 μ L of RLT Plus tissue lysis buffer provided with AllPrep DNA/RNA Mini Kit (Qiagen; 80284). After 1 min incubation, RPE cells were

detached in the lysis buffer from the posterior eyecup by gentle pipetting, followed by a removal of the remaining posterior eyecup. The lysis buffer containing RPE cells was further processed for DNA and RNA extraction using the AllPrep DNA/RNA Mini Kit protocol. The final DNA and RNA were eluted in 30 μ L and 15 μ L water, respectively. cDNA synthesis was performed using the SuperScript[™] III First-Strand Synthesis SuperMix (Thermo Fisher; 18080400).

Western blot analysis of mouse RPE tissue extracts

[0505] To prepare the protein lysate from the mouse RPE tissue, the dissected mouse eyecup, consisting of RPE, choroid, and sclera, was transferred to a microcentrifuge tube containing 30 μ L of RIPA buffer with protease inhibitors and homogenized with a motor tissue grinder (Fisher Scientific; K749540-0000) and centrifuged for 30 min at 20,000 g at 4°C. The resulting supernatant was pre-cleared with Dynabeads Protein G (Thermo Fisher; 10003D) to remove contaminants from blood prior to gel loading. Twenty μ L of RPE lysates pre-mixed with NuPAGE LDS Sample Buffer (Thermo Fisher; NP0007) and NuPAGE Sample Reducing Agent (Thermo Fisher; NP0004) was loaded into each well of a NuPAGE 4-12% Bis-Tris gel (Thermo Fisher; NP0321BOX), separated for 1 h at 130 V and transferred onto a PVDF membrane (Millipore; IPVH00010). After 1 h blocking in 5% (w/v) non-fat milk in PBS containing 0.1% (v/v) Tween-20 (PBS-T), the membrane was incubated with primary antibody, mouse anti-RPE65 monoclonal antibody (1:1,000; in-house production) (Golczak *et al.*, 2010), diluted in 1% (w/v) non-fat milk in PBS-T overnight at 4°C. After overnight incubation, membranes were washed three times with PBS-T for 5 min each and then incubated with goat anti-mouse IgG-HRP antibody (1:5,000; Cell Signaling Technology; 7076S) for 1 h at room temperature. After washing the membrane three times with PBS-T for 5 min each, protein bands were visualized after exposure to SuperSignal West Pico Chemiluminescent substrate (Thermo Fisher; 34580). Membranes were stripped and reprobed for ABE and β -actin expression using mouse anti-Cas9 monoclonal antibody (1:1,000; Invitrogen; MA523519) and rabbit anti- β -actin polyclonal antibody (1:1,000; Cell Signaling Technology; 4970S), following the same protocol. Corresponding secondary antibodies were goat anti-mouse IgG-HRP antibody (1:5,000; Cell Signaling Technology; 7076S) and goat anti-rabbit IgG-HRP antibody (1:5,000; Cell Signaling Technology; 7074S).

Electroretinography

[0506] Prior to recording, mice were dark adapted for 24 h overnight. Under a safety light, mice were anesthetized by intraperitoneal injection of a cocktail consisting of 20 mg/mL

ketamine and 1.75 mg/mL xylazine in phosphate-buffered saline at a dose of 0.1 mL per 20 g body weight, and their pupils were dilated with topical administration of 1% tropicamide ophthalmic solution (Akorn; 17478-102-12) followed by 2.5% hypromellose (Akorn; 9050-1) for hydration. The mouse was placed on a heated Diagnosys Celeris rodent ERG device (Diagnosys LLC). Ocular electrodes were placed on the corneas, and the reference electrode was positioned subdermally between the ears. The eyes were stimulated with a green light (peak emission 544 nm, bandwidth ~160 nm) stimulus of $-0.3 \log (\text{cd} \cdot \text{s}/\text{m}^2)$. The responses for 10 stimuli with an inter-stimulus interval of 10 s were averaged together, and the a- and b-wave amplitudes were acquired from the averaged ERG waveform. The ERGs were recorded with the Celeris rodent electrophysiology system (Diagnosys LLC) and analyzed with Espion V6 software (Diagnosys LLC).

Quantification And Statistical Analysis

[0507] Data are presented as mean and standard error of the mean (s.e.m.). No statistical methods were used to predetermine sample size. Statistical analysis was performed using GraphPad Prism software. Sample size and the statistical tests used are described in the figure legends.

Additional Sequences

Table 4.

Description	Protein sequence
v1 BE-VLP	MGQAVTTPLSLTLDHWKDVERTAHNLSVEVRKRRWVTFCSAEWPTFNVGWPRDGT NPDIITQVKIKVFSPGPHGHPDQVPYIVTWEAIAVDPPPWVRPFVHPKPPLSLPPSAPSLP PEPLSTPPQSSLYPALTSPLNTKPRPQVLPDSGGPLIDLLTEDPPPYRDPGPPSPDGNDS GEVAPTEGAPDPSPMVSRRLGRKEPPVADSTTSQAFPLRLGGNGQYQYWPFSDDL WKNNNPSFSEDPAKLTALIESVLLTHQPTWDDCQQLGTLTGEEKQRVLLLEARKAVR GEDGRPTQLPNDINDAFPLERPDWDYNTQGRNHLVHYRQLLLAGLQNAGRSPTNLA KVKGITQGPNEPSAFLERLKEAYRRYTPYDPEDPGQETNVAMSFQWQAPDIGRKL LEDLKSCTLGDLVREAEKIFNKRETPEEREERIRRETEEKEERRRAEDVQREKERDRR HREMSKLLATVVSQQRQDRQGGERRRPQLDHDQCA YCKEKGHWARDCPKKPRGPRG PRPQASLLTRSSLYPALTPTGDKDDDDKKRTADGSEFESPKKKRKVSEVEFSHEYWM RHALTLAKRARDEREVPVGAVLVLNRRVIGEGWNRAIGLHDPTAHAEIMALRQGGLV MQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAAGSLMNVNLNYPGM NHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGTSES ATPESSGGSSGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLI GALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEFSNEMAKVDDSFHRLSEFLV EEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSADKADLRILIYLAHMIKFRG HFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLNLI AQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGD QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQL PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQ RTFDNGSIPHQIHLGELHAILRRQEDFYFLKDNREKIEKILTFRIPIYYVGPLARGNSRFA WMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLEYEFTVY NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEI SGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYA HLFDDKVMKQLKRRRYTGWGRSLRKLINGIRDKQSGKTILDFLKSDFANRNFMLQIH

	DDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDDELVKVMGRHKP ENIVIAMARENQTTQKGQKNSRERMKRIIEGKELGSQILKEHPVENTQLQNEKLYLY LQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNPSE EVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETROITKHV AQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAYL NAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFK TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFS KESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKGKSKKLKSVKELL GITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGN ELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIEQISEFSKRVLAD ANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDSGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 470)
v2.1 BE-VLP	MGQTVTTPLSLTLGHWKDVERIAHNQSV DVKKRRWVTFCSAEWPTFNVGWPRDGT NRDLITQVKIKVFSPGPHGHPDQVPYIVTWEALAFDPPPWVKPFVHPKPPPPLPPSAPSLP LEPPRSTPPRSSLYPALTPSLGAKPKPQVLSDSGGPLIDLTTEDPPPYRDRPPPSDRDGN GGEATPAGEAPDPSPMASRLRGRREPPVADSTTSQAFPLRAGGNGQLQYWPFSDDL NWKNNPFSFSEDPGKLTALIESVLITHQPTWDDCQQLGLTLLTGEEKQRVLLLEARKAV RGDDGRPTQLPNEVDAAFLERPDWDYTTQAGRNHLVHYRQLLLAGLQNAGRSPTNL AKVKGITQGPNESPSAFLERLKEAYRRYTPYDPEDPGQETNVSMFSIWQSAPDIGRKL RLEDLKNKTLGDLVREAEKIFNKRETPEEREERIRRETEEKEERRRTEDEQKEKERDRRR HREMSKLLATVVSQKQDRQGGERRRSQLD RDQCA YCKEKGHWAKDCPKKPRGPRG PRPQTSLLTLDDPRSSLYPALTPGDYKDDDDKKRTADGSEFESPKKKRVSEVEFSHEY WMRHALTLAKRARDEREVPVGAVLVLN RVIGEGWNRAIGLHDPTAHAEIMALRQGG LVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAAGSLMNVLNYP GMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGT SESATPESGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKK NLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFHRLSEF LVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRILIYLAHAMIKF RGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLEN LIAQLPGEKKNGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIG DQYADLFLAAKNLS DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRK QRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYVGPLARGNSRF AWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSV EISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKY AHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLI HDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDDELVKVMGRHK PENIVIAMARENQTTQKGQKNSRERMKRIIEGKELGSQILKEHPVENTQLQNEKLYLY YLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNPSE EVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETROITKH VAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFF KTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGF SKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKGKSKKLKSVKEL LGITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQK NELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIEQISEFSKRVL DANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEV LDATLIHQ SITGLYETRIDLSQLGGDSGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 471)
v2.2 BE-VLP	MGQTVTTPLSLTLGHWKDVERIAHNQSV DVKKRRWVTFCSAEWPTFNVGWPRDGT NRDLITQVKIKVFSPGPHGHPDQVPYIVTWEALAFDPPPWVKPFVHPKPPPPLPPSAPSLP LEPPRSTPPRSSLYPALTPSLGAKPKPQVLSDSGGPLIDLTTEDPPPYRDRPPPSDRDGN GGEATPAGEAPDPSPMASRLRGRREPPVADSTTSQAFPLRAGGNGQLQYWPFSDDL NWKNNPFSFSEDPGKLTALIESVLITHQPTWDDCQQLGLTLLTGEEKQRVLLLEARKAV RGDDGRPTQLPNEVDAAFLERPDWDYTTQAGRNHLVHYRQLLLAGLQNAGRSPTNL AKVKGITQGPNESPSAFLERLKEAYRRYTPYDPEDPGQETNVSMFSIWQSAPDIGRKL RLEDLKNKTLGDLVREAEKIFNKRETPEEREERIRRETEEKEERRRTEDEQKEKERDRRR HREMSKLLATVVSQKQDRQGGERRRSQLD RDQCA YCKEKGHWAKDCPKKPRGPRG PRPQTSLLTLDDVQALVLTQGDYKDDDDKKRTADGSEFESPKKKRVSEVEFSHEYW MRHALTLAKRARDEREVPVGAVLVLN RVIGEGWNRAIGLHDPTAHAEIMALRQGG

	VMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAAGSLMNVLNYPG MNRHVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSINSGSSGGSSGSETPGTS ESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSSKFKVLGNTRHSIKK NLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFHRLSEESF LVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADRLIYLALAHMIKF RGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLEN LIAQLPGEKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKDTYDDDLNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRK QRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYYVGPLARGNSRF AWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSLLEYEFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSV EISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKY AHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLQI HDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHK PENIVIMARENQTTQKGQKNSRERMKRIEIEGKELGSQILKEHPVENTQLQNEKLYL YLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPS EEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDAKAGFIKRQLVETRQITKH VAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKFDFQFYK VREINNYHHAHDAY LNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFF KTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGF SKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKEL LGITIMERSSEKPNIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIIEQISEFSKR VILA DANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEV LDATLIHQ SITGLYETRIDLSQLGGDSGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 472)
v2.3 BE-VLP	MGQTVTTPLSLTLGHWKDVERIAHNQSV DVKKRRWVTFCSAEWPTFNVGWPRDGT NRDLITQVKIKVFSPGPHGHPDQVPYIVTWEALAFDPPPWVKPFVHPKPPPLPPSAPSLP LEPPRSTPPRSSLYPALTPSLGAKPKPQVLSDSGGPLIDLLEDPPPYRDRPPPSDRDGN GGEATPAGEAPDPSPMASRLRGRREPPVADSTTSQAFPLRAGGNGQLQYWPSSSDLY NWKNNNPSFSEDPGKLTALIESVLITHQPTWDDCQQLGTLTGEEKQRVLLLEARKAV RGDDGRPTQLPNEVDAAFLERPDWDYTTQAGRNHLVHYRQLLLAGLQNAGRSPTNL AKVKGITQGPNESPSAFLERLKEAYRRYTPYDPEDPGQETNVSMSFIWQSAPDIGRKL RLEDLKNKTLGDLVREAEKIFNKRETPEEREERIRRETEEKEERRRTEDEQKEKERDRR HREMSKLLATVVSQGKQDRQGGERRRSQLD RDQCA YCKEKGHWAKDCPKKPRGPRG PRPQTSLLTLDDPLQVLTNIERRGDYKDDDDKKRTADGSEFESPKKKRKVSEVEFSHE YWMRHALTLAKRARDEREVPVGAVLV LNNRVIGEGWNRAIGLHDPTAHAEIMALRQG GLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAAGSLMNVLNYP PGMNRHVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSINSGSSGGSSGSETPG TSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSSKFKVLGNTRHSIK KNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSSFHRLSEES FLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADRLIYLALAHMIK FRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL NLIAQLPGEKKNGLFGNLIALSLGLTPNFKS NFDLAEDAKLQLSKDTYDDDLNLLAQI GDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQ QLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR KQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYYVGPLARGNSR FAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSLLEYEFTV VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKT Y AHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLQ LIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRH K PENIVIMARENQTTQKGQKNSRERMKRIEIEGKELGSQILKEHPVENTQLQNEKLYL YYLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPS SEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDAKAGFIKRQLVETRQITK HVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKFDFQFYK VREINNYHHAHDA YLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFF FKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGG FSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKE LLGITIMERSSEKPNIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGELQK

	GNELALPSKYVNFLYLASHYEKLKGSPEDENEQKQLFVEQHKHYLDEIEQISEFSKRVIL ADANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKE VLDATLIHQ SITGLYETRIDLSQLGGDSGGSKRTADGSEFEPK KKRKV (SEQ ID NO: 473)
v2.4 BE-VLP	MGQTVTTPLSLTLGHWKDVERIAHNQSV DVKKRRWVTFCSAEWPTFNVGWPRDGT NRDLITQVKIKVFSPGPHGHPDQVPYIVTWEALAFDPPPWVKPFVHPKPPPPLPPSAPSLP LEPPRSTPPRSSLYPALTPSLGAKPKPQVLSDSGGPLIDLTTEDPPPYRDRPPPSDRDGN GGEATPAGEAPDPSPMASRLRGRREPPVADSTTSQAFPLRAGGNGQLQYWPFS SSDLY NWKNNNPSFSEDPGKLTALIESVLITHQPTWDDCQQLGTLTGEEKQRVLLEARKAV RGDDGRPTQLPNEVDAAFLERPDWDYTTQAGRNHLVHYRQLLLAGLQNAGRSPTNL AKVKGITQGPNESPSAFLERLKEAYRRYTPYDPEDPGQETNVSMSFIWQSAPDIGR KLE RLEDLKNKTLGDLVREAEKIFNKRETPEEREERIRRETEEKEERRRTEDEQKEKERDRRR HREMSKLLATVVSGQKQDRQGGERRRSQ LDRDQCA YCKEKGHWAKDCPKKPRGPRG PRPQTSLLTLDDTSTLLMENSSGDYKDDDDKKRTADGSEFESPKKKR KVSEVEFSHEY WMRHALTLAKRARDEREVPVGAVLV LNNRVIGEGWNRAIGLHDPTAHAEIMALRQGG LVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAAGSLMNVLNYP GMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGT SESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKK NLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFHRL EESF LVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYLALAHMIKF RGHFLIEGDLNPDNSDVKLFQILVQTYNQLFEENPINASGVDAKAILSARLSKSRRL EN LIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV KLNREDLLRK QRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPIYVGPLARGNSRF AWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSV EISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLT LTLFEDREMIEERLKY AHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS DGFANRNFMQLI HDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA IKKGILQTVKVVDELVKVMGRHK PENIVEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLY YLQNGRDMYVDQELDINRLSDYDV DHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVP S EEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGSEL DKA GFIKRQLVETRQITKH VAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYK VREINNYHHAHDAY LNAVVGTA LIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFF KTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGF SKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKEL LGITIMERS SFEKNPIDFLEAKGYKEVKKDLI IKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPEDENEQKQLFVEQHKHYLDEIEQISEFSKRVILA DANLDKVL SAYNKH RD KPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEV L DATLIHQ SITGLYETRIDLSQLGGDSGGSKRTADGSEFEPK KKRKV (SEQ ID NO: 474)
v3.1 BE-VLP	MGQTVTTPLSLTLGHWKDVERIAHNQSV DVKKRRWVTFCSAEWPTFNVGWPRDGT NRDLITQVKIKVFSPGPHGHPDQVPYIVTWEALAFDPPPWVKPFVHPKPPPPLPPSAPSLP LEPPRSTPPRSSLYPALTPSLGAKPKPQVLSDSGGPLIDLTTEDPPPYRDRPPPSDRDGN GGEATPAGEAPDPSPMASRLRGRREPPVADSTTSQAFPLRAGGNGQLQYWPFS SSDLY NWKNNNPSFSEDPGKLTALIESVLITHQPTWDDCQQLGTLTGEEKQRVLLEARKAV RGDDGRPTQLPNEVDAAFLERPDWDYTTQAGRNHLVHYRQLLLAGLQNAGRSPTNL AKVKGITQGPNESPSAFLERLKEAYRRYTPYDPEDPGQETNVSMSFIWQSAPDIGR KLE RLEDLKNKTLGDLVREAEKIFNKRETPEEREERIRRETEEKEERRRTEDEQKEKERDRRR HREMSKLLATVVSGQKQDRQGGERRRSQ LDRDQCA YCKEKGHWAKDCPKKPRGPRG PRPQTSLLTLDDTSTLLMENSSGDYKDDDDKKRTADGSEFESPKKKR KVSEVEFSHEY WMRHALTLAKRARDEREVPVGAVLV LNNRVIGEGWNRAIGLHDPTAHAEIMALRQGG LVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAAGSLMNVLNYP GMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGT SESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKK NLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFHRL EESF LVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYLALAHMIKF RGHFLIEGDLNPDNSDVKLFQILVQTYNQLFEENPINASGVDAKAILSARLSKSRRL EN LIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV KLNREDLLRK

	QRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRF AWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSHLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSV EISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKY AHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFQMQLI HDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDELVKVMGRHK PENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLY YLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPS EEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKH VAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTAIIKKYPKLESEFVYGDYKVDVRKMIKSEQEIGKATAKYFFYSNIMNFF KTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNVKKTEVQTGGF SKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKEL LGITIMERSSEKPNIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIIEQISEFSKRVL DANLDKVLSAYNKHRDKPIREQAENIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEV LDATLIHQSTGLYETRIDLSQLGGDSGSKRTADGSEFEPKKKRKVSOGGMSKLLATV VSSGSLQLPPLERLTGSLQLPPLERLTGSLQLPPLERLT (SEQ ID NO: 475)
v3.2 BE-VLP	MGQTVTTPLSLTLGHWKDVERIAHNQSV DVKKRRWVTFCSAEWPTFNVGWPRDGT NRDLITQVKIKVFSPGPHGHPDQVPYIVTWEALAFDPPPWVKPFVHPKPPPLPPSAPSLP LEPPRSTPPRSSLYPALTPSLGAKPKQVLSDSGGPLIDLTTEDPPPYRDRPPPSDRDGN GGEATPAGEAPDPSPMASRLRGRREPPVADSTTSQAFPLRAGGNGQLQYWPSSSDLY NWKNNNPSFSEDPGKLTAIESVLITHQPTWDDCQQLGTLTGEEKQVRVLEARKAV RGDDGRPTQLPNEVDAAFLERPDDYTTQAGRNHLVHYRQLLLAGLQNAGRSPTNL AKVKGITQGPNESPSAFLERLKEAYRRYTPYDPEDPGQETNVSMSFIWQSAPDIGRKE RLEDLKNKTLGDLVREAEKIFNKRETPEEREERIRRETEEKEERRRTEDEQKEKERDRR HREMSKLLATVVSGQKQDRQGGERRRSQLD RDQCA YCKEKGHWAKDCPKKPRGPRG PRPQTSLLTDDTSTLLMENS SSGDYKDDDDKKRTADGSEFESPKKKRKVSEVEFSHEY WMRHALTLAKRARDEREVPVGA VLVNLRVIGEGWNRAIGLHDPTAHAEIMALRQGG LVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAGSLMNVNLNYP GMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGT SESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKK NLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESF LVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRILIYLAHAMIKF RGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLEN LIAQLPGEKKNGFLGNLIALSLGLTPNFKS NFDLAEDAKLQLSKDTYDDDLNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRK QRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLARGNSRF AWMTRKSEETITPWNFEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSHLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSV EISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKY AHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFQMQLI HDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKKGILQTVKVVDELVKVMGRHK PENIVIAMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLY YLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPS EEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKH VAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTAIIKKYPKLESEFVYGDYKVDVRKMIKSEQEIGKATAKYFFYSNIMNFF KTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNVKKTEVQTGGF SKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGKSKKLKSVKEL LGITIMERSSEKPNIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIIEQISEFSKRVL DANLDKVLSAYNKHRDKPIREQAENIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEV LDATLIHQSTGLYETRIDLSQLGGDSGSKRTADGSEFEPKKKKRVSSGGSPLQVLTNIE RRSGGSLQLPPLERLTGSLQLPPLERLTGSLQLPPLERLT (SEQ ID NO: 476)
v3.3 BE-VLP	MGQTVTTPLSLTLGHWKDVERIAHNQSV DVKKRRWVTFCSAEWPTFNVGWPRDGT NRDLITQVKIKVFSPGPHGHPDQVPYIVTWEALAFDPPPWVKPFVHPKPPPLPPSAPSLP LEPPRSTPPRSSLYPALTPSLGAKPKQVLSDSGGPLIDLTTEDPPPYRDRPPPSDRDGN GGEATPAGEAPDPSPMASRLRGRREPPVADSTTSQAFPLRAGGNGQLQYWPSSSDLY

	NWKNNNPSFSEDPGKLTALIESVLITHQPTWDDCQQLLGTLTGEEKQRVLLLEARKAV RGDDGRPTQLPNEVDAAFLERPDWDYTTQAGRNHLVHYRQLLLAGLQNAGRSPTNL AKVKGITQGPNESPSAFLERLKEAYRRYTPYDPEDPGQETNVSMSFIWQSAPDIGRKL RLEDLKNKTLGDLVREAEIFNKRETPEEREERIRRETEEKEERRRTEDEQKEKERDRRR HREMSKLLATVVSGQKQDRQGGERRRSQLDQCAQYCKEKGHWAKDCPKKPRGPRG PRPQTSLLTLDDTSTLLMENS SGDYKDDDDKKRTADGSEFESPCKKRKVSEVEFSHEY WMRHALTLAKRARDEREVPVGAVLVNLRVIGEGWNRAIGLHDPTAHAEIMALRQGG LVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNSKRGAAAGSLMNVLNYP GMNHRVEITEGILADECAALLCDFYRMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGT SESATPESGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKK NLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDDSFFHRLEESF LVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYLALAHMIKF RGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLEN LIAQLPGEKKNGFLGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRK QRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYVGPLARGNSRF AWMTRKSEETITPWNFEVVDKGASQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSV EISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKY AHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFMQLI HDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHK PENIVIEMARENQTTQKGQKNSRERMKRIEELGELGSQILKEHPVENTQLQNEKLYLY YLQNGRDMYVDQELDINRLSDYDVHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPS EEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKH VAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHDAY LNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFF KTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGF SKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVAVKEVGKSKKLKSVKEL LGITIMERSSEFKNPIDFLEAKGYKEVKKDLIKLPKYSLFELENGRKRMLASAGELQKG NELALPSKYVNFLYLASHYEKLKGSPEDEQKQLFVEQHKHYLDEIEQISEFSKRVILA DANLDKVLSAYNKHRDKPIREQAENIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEV LDATLIHQSIITGLYETRIDLSQLGGDSGGSKRTADGSEFEPKCKRKVS GGSIRKIFLDGSG GSLQLPPLERLTGSLQLPPLERLTGSLQLPPLERLT (SEQ ID NO: 477)
v3.4/v4 BE- VLP	MGQTVTTPLSLTLGHWKDVERIAHNQSV DVKKRRWVTFCSAEWPTFNVGWPRDGT NRDLITQVKIKVFSPPHGHDPQVPYIVTWEALAFDPPPWKPFVHPKPPPLPPSAPSLP LEPPRSTPPRSSLYPALTPSLGAKPKQVLSDSGGPLIDLLEDPPPYRDRPPPSDRDGN GGEATPAGEAPDPSPMASRLRGRREPPVADSTTSQAFPLRAGGNGQLQYWPFSDDL NWKNNNPSFSEDPGKLTALIESVLITHQPTWDDCQQLLGTLTGEEKQRVLLLEARKAV RGDDGRPTQLPNEVDAAFLERPDWDYTTQAGRNHLVHYRQLLLAGLQNAGRSPTNL AKVKGITQGPNESPSAFLERLKEAYRRYTPYDPEDPGQETNVSMSFIWQSAPDIGRKL RLEDLKNKTLGDLVREAEIFNKRETPEEREERIRRETEEKEERRRTEDEQKEKERDRRR HREMSKLLATVVSGQKQDRQGGERRRSQLDQCAQYCKEKGHWAKDCPKKPRGPRG PRPQTSLLTLDDSGGSLQLPPLERLTGSLQLPPLERLTGSLQLPPLERLTSTLLMEN SSGDYKDDDDKKRTADGSEFESPCKKRKVSEVEFSHEYWMRHALTLAKRARDEREVP VGAVLVNLRVIGEGWNRAIGLHDPTAHAEIMALRQGG LVMQNYRLIDATLYVTFEPC VMCAGAMIHSRIGRVVFGVRNSKRGAAAGSLMNVLNYPGMNHRVEITEGILADECAAL LCDFYRMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGTSESATPESGGSSGGSDKKYS IGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKR TARRRYTRRKNRICYLQEIFSNEMAKVDDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEV AYHEKYPTIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLF IQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGFLGNLIALSL GLTPNFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDI LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ LPEKYKEIFFDQSKNGYAGYID GGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILR RQEDFYFPLKDNREKIEKILTFRIPIYVGPLARGNSRF AWMTRKSEETITPWNFEVVDK GASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSG EQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIK DKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGW RLSRKLINGIRDKQSGKTILDFLKSDFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL

	HEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSR ERMKRIIEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDY DVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV KVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYG DYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGE IVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYK EVKKDLIIKLPKYSLEFENGKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLG SPEDNEQKQLFVEQHKHYLDEIIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAE NIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGD SGGSKRTADGSEFEPKKRKV (SEQ ID NO: 478)
v4 BE-VLP (ABE8e-NG)	MGQTVTTPLSLTLGHWKDVERIAHNQSV DVKKRRWVTFCSAEWPTFNVGWPRDGT NRDLITQVKIKVFSPGPHGHPDQVPYIVTWEALAFDPPPWVKPFVHPKPPPPLPPSAPSLP LEPPRSTPPRSSLYPALTPSLGAKPKQVLSDSGGPLIDLLEDPPPYRDRPPPSDRDGN GGEATPAGEAPDPSPMASRLRGRREPPVADSTTSQAFPLRAGGNGQLQYWPFSDDL NWKNNPSPFSEDPGKLTALIESVLITHQPTWDDCQQLGLTLLTGEEKQRVLLERKAV RGDDGRPTQLPNEVDAAFLERPDWDYTTQAGRNHLVHYRQLLLAGLQNAGRSPTNL AKVKGITQGPNESPSAFLERLKEAYRRYTPYDPEDPGQETNVSMFSIWQSAPDIGRKL RLEDLKNKTLGDLVREAEKIFNKRETPEEREERIRRETEEKEERRRTEDEQKEKERDRR HREMSKLLATVVSQKQDRQGGERRRSQLD RDQCA YCKEKGHWAKDCPKKPRGPRG PRPQTSLLTLDDSGGSLQLPPLERLTGSLQLPPLERLTGSLQLPPLERLTSTLLMEN SSGDYKDDDDKKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRARDEREVP VGAVLVNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPC VMCAGAMIHSRIGRVVFGVRNSKRGAAAGSLMNVNYPGMNHRVEITEGILADECAAL LCDFYRMPRQVFNAQKKAQSSINSGGSSGGSSGSETPGTSESATPSSGGSSGGSDKKYS IGLAIGTNSVGWAVITDEYKVP SKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKR TARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEV AYHEKYPTIYHLRKKLV DSTDKADLRLIY LALAHMIKFRGHFLIEGDLNPDNSDVKLF IQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASL GLTPNFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDI LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYID GGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILR RQEDFYFPFLKDNREKIEKILTFRIPIYVGPLARGNSRF AWMTRKSEETITPWNFEVVDK GASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSG EQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIK DKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWG RLSRKLINGIRDKQSGKTILDFLKS DGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSL HEHIANLAGSPAIIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSR ERMKRIIEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDY DVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV KVITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYG DYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGE IVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESIRPKRNSDKLIARKKDWDPK KYGGFVSPTVAYSVLVVAKEVGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYK EVKKDLIIKLPKYSLEFENGKRMLASARFLQKGNELALPSKYVNFLYLASHYEKLG SPEDNEQKQLFVEQHKHYLDEIIIEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQAE NIIHLFTLTNLGAPRAFKYFDTTIDRKVYRSTKEVLDATLIHQ SITGLYETRIDLSQLGGD SGGSKRTADGSEFEPKKRKV (SEQ ID NO: 479)
v4 BE-VLP (ABE7.10-NG)	MGQTVTTPLSLTLGHWKDVERIAHNQSV DVKKRRWVTFCSAEWPTFNVGWPRDGT NRDLITQVKIKVFSPGPHGHPDQVPYIVTWEALAFDPPPWVKPFVHPKPPPPLPPSAPSLP LEPPRSTPPRSSLYPALTPSLGAKPKQVLSDSGGPLIDLLEDPPPYRDRPPPSDRDGN GGEATPAGEAPDPSPMASRLRGRREPPVADSTTSQAFPLRAGGNGQLQYWPFSDDL NWKNNPSPFSEDPGKLTALIESVLITHQPTWDDCQQLGLTLLTGEEKQRVLLERKAV RGDDGRPTQLPNEVDAAFLERPDWDYTTQAGRNHLVHYRQLLLAGLQNAGRSPTNL AKVKGITQGPNESPSAFLERLKEAYRRYTPYDPEDPGQETNVSMFSIWQSAPDIGRKL RLEDLKNKTLGDLVREAEKIFNKRETPEEREERIRRETEEKEERRRTEDEQKEKERDRR HREMSKLLATVVSQKQDRQGGERRRSQLD RDQCA YCKEKGHWAKDCPKKPRGPRG PRPQTSLLTLDDSGGSLQLPPLERLTGSLQLPPLERLTGSLQLPPLERLTSTLLMEN

SSGDYKDDDDKKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVP VGAVLVHNNRVIGEGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPC VMCAGAMIHSRIGRVVFGARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAAL LSDFFRMRRQEIKAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGSSSEVEFS HEYWMRHALTLAKRARDEREVPVGAVLVNNRVIGEGWNRRAIGLHDPTAHAEIMALR QGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVL HYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDSGGSSGGSSGSET PGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRH SIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDDTKADLRILIYLAHAH MIKFRGHFLIEGDLNPDNSDVKLFQVLVQTYNQLFEENPINASGVDAKAILSARLSKSR RLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLNLL AQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALV RQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDL LRKQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTRIPYYVGPLARGN SRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHKSLLEYEF TVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDILLFKTNRKVTVKQLKEDYFKKIECFD SVEISGVEDRFNASLGTYHDLKIKDKDFLDNEENEDILEDIVLTTLFEDREMIEERLK TYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLSDGFANRNF QLIHDDSLTFKEDIQKAQVSGQGDLSLHEHIANLAGSPAIIKKGILQTVKVVDELVKVMGR HKPENIVIAMARENQTTQKGQKNSRERMKRIEELGSGILKEHPVENTQLQNEKLY LYYLQNGRDMYVDQELDINRLSDYDVDHVPQSFLKDDSIDNKVLTRSDKNRGKSDNV PSEEVVKMKMNYWRQLLNAKLITQRKFDNLTAKERGGLSELDKAGFIKRQLVETRQIT KHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVINNYHHAHD AYLNAVVGTAIIKKYPKLESEFVYGDYKVDVRKMIKSEQEIGKATAKYFFYSNIMN FFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNVKKTEVQTG GFSKESIRPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAKEVGKSKKLKSVK ELLGITIMERSSEKPNIDFLEAKGYKEVKKDLIIKLPKYSLEFENGRKRMLASARFLQK GNELALPSKYVNFLYLASHYEKLKGPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVIL ADANLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPRAFKYFDTTIDRKVYRSTKE VLDATLIHQSIITGLYETRIDLSQLGGDSGGSKRTADGSEFEPKKRKV (SEQ ID NO: 480)

Table 5.

Description	Spacer	SEQ ID NO:	Gene
On-target	CCCATACCTTGGAGCAACGG <i>CGG</i>	481	Pcsk9
OT1	<i>G</i> ACATACCTTAAAGCAAAGG <i>AGG</i>	482	Intron; <i>ELP3</i>
OT2	CCCCTACCTTGGGGCAACAG <i>TGG</i>	483	Intergenic
OT3	CCCACCCTTTGGAG-AACGG <i>TGG</i>	484	LncRNA; LINC02006
OT4	CCCAG-CCTTGGGGCAACGG <i>AGG</i>	485	Intergenic
OT5	CACATATCTAGGAGCAA-GG <i>AGG</i>	486	Intergenic
OT6	CCCACACCC- G GAGCAACGG <i>GGA</i>	487	Intron; <i>DDX6</i>
OT7	T CCATACCC-GGAGCAACGA <i>GGG</i>	488	LncRNA; RP11-314D7.4
OT8	T TCAT-CCTTGGAGCAACGG <i>TGA</i>	489	LncRNA; FAM66D
OT9	T CTGTACCATGGAGCAAAGG <i>CGG</i>	490	LncRNA; RIKEN cDNA 4933424G05 gene
OT10	ACCATAACCAAGAGCAACAG <i>GGG</i>	491	Intron; <i>Klhl3</i>
OT11	T CCATAACTCAGAGCAACAG <i>TGG</i>	492	Intergenic
OT12	G CCATACCTTGGGGCAGCAG <i>TGG</i>	493	Intron; <i>NCAM1</i>
OT13	G CAACACCTTGGAGCAACTG <i>AGG</i>	494	Intron; <i>SNRNP40</i>
OT14	G ACAT-CCTTGGAGCAACTG <i>TGG</i>	495	Intron; <i>Fry</i>

*Mismatches are denoted in bold italic.

REFERENCES

- [0508] Abifadel, M., Varret, M., Rabes, J.P., Allard, D., Ouguerram, K., Devillers, M., Cruaud, C., Benjannet, S., Wickham, L., Erlich, D., *et al.* (2003). Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nat Genet* 34, 154-156.
- [0509] Akcakaya, P., Bobbin, M.L., Guo, J.A., Malagon-Lopez, J., Clement, K., Garcia, S.P., Fellows, M.D., Porritt, M.J., Firth, M.A., Carreras, A., *et al.* (2018). *In vivo* CRISPR editing with no detectable genome-wide off-target mutations. *Nature* 561, 416-419.
- [0510] Alanis-Lobato, G., Zohren, J., McCarthy, A., Fogarty, N.M.E., Kubikova, N., Hardman, E., Greco, M., Wells, D., Turner, J.M.A., and Niakan, K.K. (2021). Frequent loss of heterozygosity in CRISPR-Cas9-edited early human embryos. *Proc Natl Acad Sci U S A* 118.
- [0511] Anzalone, A.V., Koblan, L.W., and Liu, D.R. (2020). Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors. *Nat Biotechnol* 38, 824-844.
- [0512] Campbell, L.A., Coke, L.M., Richie, C.T., Fortuno, L.V., Park, A.Y., and Harvey, B.K. (2019). Vesicle-Mediated Delivery of CRISPR/Cas9 Ribonucleoprotein Complex for Inactivating the HIV Provirus. *Mol Ther* 27, 151-163.
- [0513] Chandler, R.J., Sands, M.S., and Venditti, C.P. (2017). Recombinant Adeno-Associated Viral Integration and Genotoxicity: Insights from Animal Models. *Hum Gene Ther* 28, 314-322.
- [0514] Chen, P.J., Hussmann, J.A., Yan, J., Knipping, F., Ravisankar, P., Chen, P.F., Chen, C., Nelson, J.W., Newby, G.A., Sahin, M., *et al.* (2021). Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. *Cell* 184, 5635-5652 e5629.
- [0515] Choi, J.G., Dang, Y., Abraham, S., Ma, H., Zhang, J., Guo, H., Cai, Y., Mikkelsen, J.G., Wu, H., Shankar, P., *et al.* (2016). Lentivirus pre-packed with Cas9 protein for safer gene editing. *Gene Ther* 23, 627-633.
- [0516] Cideciyan, A.V. (2010). Leber congenital amaurosis due to RPE65 mutations and its treatment with gene therapy. *Prog Retin Eye Res* 29, 398-427.
- [0517] Clement, K., Rees, H., Canver, M.C., Gehrke, J.M., Farouni, R., Hsu, J.Y., Cole, M.A., Liu, D.R., Joung, J.K., Bauer, D.E., *et al.* (2019). CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat Biotechnol* 37, 224-226.
- [0518] Cohen, J., Pertsemlidis, A., Kotowski, I.K., Graham, R., Garcia, C.K., and Hobbs, H.H. (2005). Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. *Nat Genet* 37, 161-165.

- [0519] Cohen, J.C., Boerwinkle, E., Mosley, T.H., Jr., and Hobbs, H.H. (2006). Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. *N Engl J Med* 354, 1264-1272.
- [0520] Cronin, J., Zhang, X.Y., and Reiser, J. (2005). Altering the tropism of lentiviral vectors through pseudotyping. *Curr Gene Ther* 5, 387-398.
- [0521] David, R.M., and Doherty, A.T. (2017). Viral Vectors: The Road to Reducing Genotoxicity. *Toxicol Sci* 155, 315-325.
- [0522] Davis, K.M., Pattanayak, V., Thompson, D.B., Zuris, J.A., and Liu, D.R. (2015). Small molecule-triggered Cas9 protein with improved genome-editing specificity. *Nat Chem Biol* 11, 316-318.
- [0523] den Hollander, A.I., Roepman, R., Koenekoop, R.K., and Cremers, F.P. (2008). Leber congenital amaurosis: genes, proteins and disease mechanisms. *Prog Retin Eye Res* 27, 391-419.
- [0524] Deverman, B.E., Pravdo, P.L., Simpson, B.P., Kumar, S.R., Chan, K.Y., Banerjee, A., Wu, W.L., Yang, B., Huber, N., Pasca, S.P., *et al.* (2016). Cre-dependent selection yields AAV variants for widespread gene transfer to the adult brain. *Nat Biotechnol* 34, 204-209.
- [0525] Doman, J.L., Raguram, A., Newby, G.A., and Liu, D.R. (2020). Evaluation and minimization of Cas9-independent off-target DNA editing by cytosine base editors. *Nat Biotechnol* 38, 620-628.
- [0526] Doudna, J.A. (2020). The promise and challenge of therapeutic genome editing. *Nature* 578, 229-236.
- [0527] Feher, A., Boross, P., Sperka, T., Miklossy, G., Kadas, J., Bagossi, P., Oroszlan, S., Weber, I.T., and Tozser, J. (2006). Characterization of the murine leukemia virus protease and its comparison with the human immunodeficiency virus type 1 protease. *J Gen Virol* 87, 1321-1330.
- [0528] Fitzgerald, K., Frank-Kamenetsky, M., Shulga-Morskaya, S., Liebow, A., Bettencourt, B.R., Sutherland, J.E., Hutabarat, R.M., Clausen, V.A., Karsten, V., Cehelsky, J., *et al.* (2014). Effect of an RNA interference drug on the synthesis of proprotein convertase subtilisin/kexin type 9 (PCSK9) and the concentration of serum LDL cholesterol in healthy volunteers: a randomised, single-blind, placebo-controlled, phase 1 trial. *Lancet* 383, 60-68.
- [0529] Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I., and Liu, D.R. (2017). Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* 551, 464-471.

- [0530] Gaudelli, N.M., Lam, D.K., Rees, H.A., Sola-Esteves, N.M., Barrera, L.A., Born, D.A., Edwards, A., Gehrke, J.M., Lee, S.J., Liguori, A.J., *et al.* (2020). Directed evolution of adenine base editors with increased activity and therapeutic application. *Nat Biotechnol* 38, 892-900.
- [0531] Gee, P., Lung, M.S.Y., Okuzaki, Y., Sasakawa, N., Iguchi, T., Makita, Y., Hozumi, H., Miura, Y., Yang, L.F., Iwasaki, M., *et al.* (2020). Extracellular nanovesicles for packaging of CRISPR-Cas9 protein and sgRNA to induce therapeutic exon skipping. *Nat Commun* 11, 1334.
- [0532] Giannoukos, G., Ciulla, D.M., Marco, E., Abdulkerim, H.S., Barrera, L.A., Bothmer, A., Dhanapal, V., Gloskowski, S.W., Jayaram, H., Maeder, M.L., *et al.* (2018). UDiTaS, a genome editing detection method for indels and genome rearrangements. *BMC Genomics* 19, 212.
- [0533] Golczak, M., Kiser, P.D., Lodowski, D.T., Maeda, A., and Palczewski, K. (2010). Importance of Membrane Structural Integrity for RPE65 Retinoid Isomerization Activity. *Journal of Biological Chemistry* 285, 9667-9682.
- [0534] Hamilton, J.R., Tsuchida, C.A., Nguyen, D.N., Shy, B.R., McGarrigle, E.R., Sandoval Espinoza, C.R., Carr, D., Blaeschke, F., Marson, A., and Doudna, J.A. (2021). Targeted delivery of CRISPR-Cas9 and transgenes enables complex immune cell engineering. *Cell Rep* 35, 109207.
- [0535] Hooper, A.J., Marais, A.D., Tanyanyiwa, D.M., and Burnett, J.R. (2007). The C679X mutation in PCSK9 is present and lowers blood cholesterol in a Southern African population. *Atherosclerosis* 193, 445-448.
- [0536] Huang, T.P., Zhao, K.T., Miller, S.M., Gaudelli, N.M., Oakes, B.L., Fellmann, C., Savage, D.F., and Liu, D.R. (2019). Circularly permuted and PAM-modified Cas9 variants broaden the targeting scope of base editors. *Nat Biotechnol* 37, 626-631.
- [0537] Humbel, M., Ramosaj, M., Zimmer, V., Regio, S., Aeby, L., Moser, S., Boizot, A., Sipion, M., Rey, M., and Deglon, N. (2021). Maximizing lentiviral vector gene transfer in the CNS. *Gene Ther* 28, 75-88.
- [0538] Indikova, I., and Indik, S. (2020). Highly efficient 'hit-and-run' genome editing with unconcentrated lentivectors carrying Vpr.Prot.Cas9 protein produced from RRE-containing transcripts. *Nucleic Acids Res* 48, 8178-8187.

- [0539] Jang, H.K., Jo, D.H., Lee, S.N., Cho, C.S., Jeong, Y.K., Jung, Y., Yu, J., Kim, J.H., Woo, J.S., and Bae, S. (2021). High-purity production and precise editing of DNA base editing ribonucleoproteins. *Sci Adv* 7.
- [0540] Jo, D.H., Jang, H.-K., Cho, C.S., Han, J.H., Ryu, G., Jung, Y., Bae, S., and Kim, J.H. (2021). Therapeutic adenine base editing corrects nonsense mutation and improves visual function in a mouse model of Leber congenital amaurosis. *bioRxiv*.
- [0541] Johnson, S., Wheeler, J.X., Thorpe, R., Collins, M., Takeuchi, Y., and Zhao, Y. (2018). Mass spectrometry analysis reveals differences in the host cell protein species found in pseudotyped lentiviral vectors. *Biologicals* 52, 59-66.
- [0542] June, C.H., O'Connor, R.S., Kawalekar, O.U., Ghassemi, S., and Milone, M.C. (2018). CAR T cell immunotherapy for human cancer. *Science* 359, 1361-1365.
- [0543] Kaczmarczyk, S.J., Sitaraman, K., Young, H.A., Hughes, S.H., and Chatterjee, D.K. (2011). Protein delivery using engineered virus-like particles. *Proc Natl Acad Sci U S A* 108, 16998-17003.
- [0544] Kato, S., Kuramochi, M., Kobayashi, K., Fukabori, R., Okada, K., Uchigashima, M., Watanabe, M., Tsutsui, Y., and Kobayashi, K. (2011). Selective neural pathway targeting reveals key roles of thalamostriatal projection in the control of visual discrimination. *J Neurosci* 31, 17169-17179.
- [0545] Koblan, L.W., Doman, J.L., Wilson, C., Levy, J.M., Tay, T., Newby, G.A., Maianti, J.P., Raguram, A., and Liu, D.R. (2018). Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nat Biotechnol* 36, 843-846.
- [0546] Koblan, L.W., Erdos, M.R., Wilson, C., Cabral, W.A., Levy, J.M., Xiong, Z.M., Tavaréz, U.L., Davison, L.M., Gete, Y.G., Mao, X., *et al.* (2021). *In vivo* base editing rescues Hutchinson-Gilford progeria syndrome in mice. *Nature* 589, 608-614.
- [0547] Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., and Liu, D.R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420-424.
- [0548] Kosicki, M., Tomberg, K., and Bradley, A. (2018). Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nat Biotechnol* 36, 765-771.
- [0549] Lazzarotto, C.R., Malinin, N.L., Li, Y., Zhang, R., Yang, Y., Lee, G., Cowley, E., He, Y., Lan, X., Jividen, K., *et al.* (2020). CHANGE-seq reveals genetic and epigenetic effects on CRISPR-Cas9 genome-wide activity. *Nature Biotechnology* 38, 1317-1327.

- [0550] Leibowitz, M.L., Papathanasiou, S., Doerfler, P.A., Blaine, L.J., Sun, L., Yao, Y., Zhang, C.Z., Weiss, M.J., and Pellman, D. (2021). Chromothripsis as an on-target consequence of CRISPR-Cas9 genome editing. *Nat Genet* 53, 895-905.
- [0551] LeibundGut-Landmann, S., Waldburger, J.M., Krawczyk, M., Otten, L.A., Suter, T., Fontana, A., Acha-Orbea, H., and Reith, W. (2004). Mini-review: Specificity and expression of CIITA, the master regulator of MHC class II genes. *Eur J Immunol* 34, 1513-1525.
- [0552] Levy, J.M., Yeh, W.H., Pendse, N., Davis, J.R., Hennessey, E., Butcher, R., Koblan, L.W., Comander, J., Liu, Q., and Liu, D.R. (2020). Cytosine and adenine base editing of the brain, liver, retina, heart and skeletal muscle of mice via adeno-associated viruses. *Nat Biomed Eng* 4, 97-110.
- [0553] Lyu, P., Javidi-Parsijani, P., Atala, A., and Lu, B. (2019). Delivering Cas9/sgRNA ribonucleoprotein (RNP) by lentiviral capsid-based bionanoparticles for efficient 'hit-and-run' genome editing. *Nucleic Acids Res* 47, e99.
- [0554] Lyu, P., Lu, Z., Cho, S.I., Yadav, M., Yoo, K.W., Atala, A., Kim, J.S., and Lu, B. (2021). Adenine Base Editor Ribonucleoproteins Delivered by Lentivirus-Like Particles Show High On-Target Base Editing and Undetectable RNA Off-Target Activities. *CRISPR J* 4, 69-81.
- [0555] Mangeot, P.E., Dollet, S., Girard, M., Cancia, C., Joly, S., Peschanski, M., and Lotteau, V. (2011). Protein transfer into human cells by VSV-G-induced nanovesicles. *Mol Ther* 19, 1656-1666.
- [0556] Mangeot, P.E., Risson, V., Fusil, F., Marnef, A., Laurent, E., Blin, J., Mournetas, V., Massourides, E., Sohier, T.J.M., Corbin, A., *et al.* (2019). Genome editing in primary cells and in vivo using viral-derived Nanoblades loaded with Cas9-sgRNA ribonucleoproteins. *Nat Commun* 10, 45.
- [0557] Mercuri, E., Darras, B.T., Chiriboga, C.A., Day, J.W., Campbell, C., Connolly, A.M., Iannaccone, S.T., Kirschner, J., Kuntz, N.L., Saito, K., *et al.* (2018). Nusinersen versus Sham Control in Later-Onset Spinal Muscular Atrophy. *N Engl J Med* 378, 625-635.
- [0558] Meunier, L., and Larrey, D. (2019). Drug-Induced Liver Injury: Biomarkers, Requirements, Candidates, and Validation. *Front Pharmacol* 10, 1482.
- [0559] Milone, M.C., and O'Doherty, U. (2018). Clinical use of lentiviral vectors. *Leukemia* 32, 1529-1541.

- [0560] Musunuru, K., Chadwick, A.C., Mizoguchi, T., Garcia, S.P., DeNizio, J.E., Reiss, C.W., Wang, K., Iyer, S., Dutta, C., Clendaniel, V., *et al.* (2021). *In vivo* CRISPR base editing of PCSK9 durably lowers cholesterol in primates. *Nature* 593, 429-434.
- [0561] Newby, G.A., and Liu, D.R. (2021). *In vivo* somatic cell base editing and prime editing. *Mol Ther*.
- [0562] Newby, G.A., Yen, J.S., Woodard, K.J., Mayuranathan, T., Lazzarotto, C.R., Li, Y., Sheppard-Tillman, H., Porter, S.N., Yao, Y., Mayberry, K., *et al.* (2021). Base editing of haematopoietic stem cells rescues sickle cell disease in mice. *Nature* 595, 295-302.
- [0563] Nishida, K., Arazoe, T., Yachie, N., Banno, S., Kakimoto, M., Tabata, M., Mochizuki, M., Miyabe, A., Araki, M., Hara, K.Y., *et al.* (2016). Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* 353.
- [0564] Osborn, M.J., Newby, G.A., McElroy, A.N., Knipping, F., Nielsen, S.C., Riddle, M.J., Xia, L., Chen, W., Eide, C.R., Webber, B.R., *et al.* (2020). Base Editor Correction of COL7A1 in Recessive Dystrophic Epidermolysis Bullosa Patient-Derived Fibroblasts and iPSCs. *J Invest Dermatol* 140, 338-347 e335.
- [0565] Pan, D., Gunther, R., Duan, W., Wendell, S., Kaemmerer, W., Kafri, T., Verma, I.M., and Whitley, C.B. (2002). Biodistribution and toxicity studies of VSVG-pseudotyped lentiviral vector after intravenous administration in mice with the observation of *in vivo* transduction of bone marrow. *Mol Ther* 6, 19-29.
- [0566] Pang, J.J., Chang, B., Hawes, N.L., Hurd, R.E., Davisson, M.T., Li, J., Noorwez, S.M., Malhotra, R., McDowell, J.H., Kaushal, S., *et al.* (2005). Retinal degeneration 12 (rd12): a new, spontaneously arising mouse model for human Leber congenital amaurosis (LCA). *Mol Vis* 11, 152-162.
- [0567] Parr-Brownlie, L.C., Bosch-Bouju, C., Schoderboeck, L., Sizemore, R.J., Abraham, W.C., and Hughes, S.M. (2015). Lentiviral vectors as tools to understand central nervous system biology in mammalian model organisms. *Front Mol Neurosci* 8, 14.
- [0568] Puppo, A., Cesi, G., Marrocco, E., Piccolo, P., Jacca, S., Shayakhmetov, D.M., Parks, R.J., Davidson, B.L., Colloca, S., Brunetti-Pierri, N., *et al.* (2014). Retinal transduction profiles by high-capacity viral vectors. *Gene Ther* 21, 855-865.
- [0569] Rao, A.S., Lindholm, D., Rivas, M.A., Knowles, J.W., Montgomery, S.B., and Ingelsson, E. (2018). Large-Scale Phenome-Wide Association Study of PCSK9 Variants Demonstrates Protection Against Ischemic Stroke. *Circ Genom Precis Med* 11, e002162.

- [0570] Rees, H.A., Komor, A.C., Yeh, W.H., Caetano-Lopes, J., Warman, M., Edge, A.S.B., and Liu, D.R. (2017). Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery. *Nat Commun* 8, 15790.
- [0571] Rees, H.A., and Liu, D.R. (2018). Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat Rev Genet* 19, 770-788.
- [0572] Renner, T.M., Tang, V.A., Burger, D., and Langlois, M.A. (2020). Intact Viral Particle Counts Measured by Flow Virometry Provide Insight into the Infectivity and Genome Packaging Efficiency of Moloney Murine Leukemia Virus. *J Virol* 94.
- [0573] Richter, M.F., Zhao, K.T., Eton, E., Lapinaite, A., Newby, G.A., Thuronyi, B.W., Wilson, C., Koblan, L.W., Zeng, J., Bauer, D.E., *et al.* (2020). Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nat Biotechnol*.
- [0574] Rothgangl, T., Dennis, M.K., Lin, P.J.C., Oka, R., Witzigmann, D., Villiger, L., Qi, W., Hruzova, M., Kissling, L., Lenggenhager, D., *et al.* (2021). *In vivo* adenine base editing of PCSK9 in macaques reduces LDL cholesterol levels. *Nat Biotechnol* 39, 949-957.
- [0575] Serreze, D.V., Leiter, E.H., Christianson, G.J., Greiner, D., and Roopenian, D.C. (1994). Major histocompatibility complex class I-deficient NOD-B2mnull mice are diabetes and insulinitis resistant. *Diabetes* 43, 505-509.
- [0576] Sodi, A., Banfi, S., Testa, F., Della Corte, M., Passerini, I., Pelo, E., Rossi, S., Simonelli, F., and Italian, I.R.D.W.G. (2021). RPE65-associated inherited retinal diseases: consensus recommendations for eligibility to gene therapy. *Orphanet J Rare Dis* 16, 257.
- [0577] Song, Y., Liu, Z., Zhang, Y., Chen, M., Sui, T., Lai, L., and Li, Z. (2020). Large-Fragment Deletions Induced by Cas9 Cleavage while Not in the BEs System. *Mol Ther Nucleic Acids* 21, 523-526.
- [0578] Stadtmauer, E.A., Fraietta, J.A., Davis, M.M., Cohen, A.D., Weber, K.L., Lancaster, E., Mangan, P.A., Kulikovskaya, I., Gupta, M., Chen, F., *et al.* (2020). CRISPR-engineered T cells in patients with refractory cancer. *Science* 367.
- [0579] Suh, S., Choi, E.H., Leinonen, H., Foik, A.T., Newby, G.A., Yeh, W.H., Dong, Z., Kiser, P.D., Lyon, D.C., Liu, D.R., *et al.* (2021). Restoration of visual function in adult mice with an inherited retinal disease via adenine base editing. *Nat Biomed Eng* 5, 169-178.
- [0580] Swiech, L., Heidenreich, M., Banerjee, A., Habib, N., Li, Y., Trombetta, J., Sur, M., and Zhang, F. (2015). *In vivo* interrogation of gene function in the mammalian brain using CRISPR-Cas9. *Nat Biotechnol* 33, 102-106.
- [0581] Taylor, A.W. (2009). Ocular immune privilege. *Eye (Lond)* 23, 1885-1889.

- [0582] Thorne, R.G., and Nicholson, C. (2006). *In vivo* diffusion analysis with quantum dots and dextrans predicts the width of brain extracellular space. *Proc Natl Acad Sci U S A* 103, 5567-5572.
- [0583] Tsai, S.Q., Nguyen, N.T., Malagon-Lopez, J., Topkar, V.V., Aryee, M.J., and Joung, J.K. (2017). CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR-Cas9 nuclease off-targets. *Nat Methods* 14, 607-614.
- [0584] Turchiano, G., Andrieux, G., Klermund, J., Blattner, G., Pennucci, V., El Gaz, M., Monaco, G., Poddar, S., Mussolino, C., Cornu, T.I., et al. (2021). Quantitative evaluation of chromosomal rearrangements in gene-edited human stem cells by CAST-Seq. *Cell Stem Cell* 28, 1136-1147 e1135.
- [0585] Voelkel, C., Galla, M., Maetzig, T., Warlich, E., Kuehle, J., Zychlinski, D., Bode, J., Cantz, T., Schambach, A., and Baum, C. (2010). Protein transduction from retroviral Gag precursors. *Proc Natl Acad Sci U S A* 107, 7805-7810.
- [0586] Wang, D., Shukla, C., Liu, X., Schoeb, T.R., Clarke, L.A., Bedwell, D.M., and Keeling, K.M. (2010). Characterization of an MPS I-H knock-in mouse that carries a nonsense mutation analogous to the human IDUA-W402X mutation. *Mol Genet Metab* 99, 62-71.
- [0587] Wang, D., Zhang, F., and Gao, G. (2020). CRISPR-Based Therapeutic Genome Editing: Strategies and *In Vivo* Delivery by AAV Vectors. *Cell* 181, 136-150.
- [0588] Webber, B.R., Lonetree, C.L., Kluesner, M.G., Johnson, M.J., Pomeroy, E.J., Diers, M.D., Lahr, W.S., Draper, G.M., Slipek, N.J., Smeester, B.A., et al. (2019). Highly efficient multiplex human T cell engineering without double-strand breaks using Cas9 base editors. *Nat Commun* 10, 5222.
- [0589] Wei, T., Cheng, Q., Min, Y.L., Olson, E.N., and Siegwart, D.J. (2020). Systemic nanoparticle delivery of CRISPR-Cas9 ribonucleoproteins for effective tissue specific genome editing. *Nat Commun* 11, 3232.
- [0590] Wheeler, J.X., Jones, C., Thorpe, R., and Zhao, Y. (2007). Proteomics analysis of cellular components in lentiviral vector production using Gel-LC-MS/MS. *Proteomics Clin Appl* 1, 224-230.
- [0591] Wu, D.T., and Roth, M.J. (2014). MLV based viral-like-particles for delivery of toxic proteins and nuclear transcription factors. *Biomaterials* 35, 8416-8426.

- [0592] Yao, X., Lyu, P., Yoo, K., Yadav, M.K., Singh, R., Atala, A., and Lu, B. (2021). Engineered extracellular vesicles as versatile ribonucleoprotein delivery vehicles for efficient and safe CRISPR genome editing. *J Extracell Vesicles* 10, e12076.
- [0593] Yeh, W.H., Chiang, H., Rees, H.A., Edge, A.S.B., and Liu, D.R. (2018). *In vivo* base editing of post-mitotic sensory cells. *Nat Commun* 9, 2184.
- [0594] Yeh, W.H., Shubina-Oleinik, O., Levy, J.M., Pan, B., Newby, G.A., Wornow, M., Burt, R., Chen, J.C., Holt, J.R., and Liu, D.R. (2020). *In vivo* base editing restores sensory transduction and transiently improves auditory function in a mouse model of recessive deafness. *Sci Transl Med* 12.
- [0595] Yu, Y., Leete, T.C., Born, D.A., Young, L., Barrera, L.A., Lee, S.J., Rees, H.A., Ciaramella, G., and Gaudelli, N.M. (2020). Cytosine base editors with minimized unguided DNA and RNA off-target events and high on-target activity. *Nat Commun* 11, 2052.
- [0596] Zeng, J., Wu, Y., Ren, C., Bonanno, J., Shen, A.H., Shea, D., Gehrke, J.M., Clement, K., Luk, K., Yao, Q., *et al.* (2020). Therapeutic base editing of human hematopoietic stem cells. *Nat Med* 26, 535-541.
- [0597] Zhang, W., Cao, S., Martin, J.L., Mueller, J.D., and Mansky, L.M. (2015). Morphology and ultrastructure of retrovirus particles. *AIMS Biophys* 2, 343-369.
- [0598] Zhong, Z., Rong, F., Dai, Y., Yibulayin, A., Zeng, L., Liao, J., Wang, L., Huang, Z., Zhou, Z., and Chen, J. (2019). Seven novel variants expand the spectrum of RPE65-related Leber congenital amaurosis in the Chinese population. *Mol Vis* 25, 204-214.

EQUIVALENTS AND SCOPE

- [0599] In the claims articles such as “a,” “an,” and “the” may mean one or more than one unless indicated to the contrary or otherwise evident from the context. Claims or descriptions that include “or” between one or more members of a group are considered satisfied if one, more than one, or all of the group members are present in, employed in, or otherwise relevant to a given product or process unless indicated to the contrary or otherwise evident from the context. The invention includes embodiments in which exactly one member of the group is present in, employed in, or otherwise relevant to a given product or process. The invention includes embodiments in which more than one, or all of the group members are present in, employed in, or otherwise relevant to a given product or process.
- [0600] Furthermore, the invention encompasses all variations, combinations, and permutations in which one or more limitations, elements, clauses, and descriptive terms from

one or more of the listed claims is introduced into another claim. For example, any claim that is dependent on another claim can be modified to include one or more limitations found in any other claim that is dependent on the same base claim. Where elements are presented as lists, *e.g.*, in Markush group format, each subgroup of the elements is also disclosed, and any element(s) can be removed from the group. It should be understood that, in general, where the invention, or aspects of the invention, is/are referred to as comprising particular elements and/or features, certain embodiments of the invention or aspects of the invention consist, or consist essentially of, such elements and/or features. For purposes of simplicity, those embodiments have not been specifically set forth *in haec verba* herein. It is also noted that the terms “comprising” and “containing” are intended to be open and permits the inclusion of additional elements or steps. Where ranges are given, endpoints are included. Furthermore, unless otherwise indicated or otherwise evident from the context and understanding of one of ordinary skill in the art, values that are expressed as ranges can assume any specific value or sub-range within the stated ranges in different embodiments of the invention, to the tenth of the unit of the lower limit of the range, unless the context clearly dictates otherwise.

[0601] This application refers to various issued patents, published patent applications, journal articles, and other publications, all of which are incorporated herein by reference. If there is a conflict between any of the incorporated references and the instant specification, the specification shall control. In addition, any particular embodiment of the present invention that falls within the prior art may be explicitly excluded from any one or more of the claims. Because such embodiments are deemed to be known to one of ordinary skill in the art, they may be excluded even if the exclusion is not set forth explicitly herein. Any particular embodiment of the invention can be excluded from any claim, for any reason, whether or not related to the existence of prior art.

[0602] Those skilled in the art will recognize or be able to ascertain using no more than routine experimentation many equivalents to the specific embodiments described herein. The scope of the present embodiments described herein is not intended to be limited to the above Description, but rather is as set forth in the appended claims. Those of ordinary skill in the art will appreciate that various changes and modifications to this description may be made without departing from the spirit or scope of the present invention, as defined in the following claims.

CLAIMS

What is claimed is:

1. A virus-like particle comprising a group-specific antigen (gag) protease (pro) polyprotein and a fusion protein encapsulated by a lipid membrane and a viral envelope glycoprotein, wherein the fusion protein comprises:
 - (i) a gag nucleocapsid protein;
 - (ii) a nucleic acid programmable DNA binding protein (napDNAbp);
 - (iii) a cleavable linker; and
 - (iv) a nuclear export sequence (NES).
2. The virus-like particle of claim 1, wherein the napDNAbp is a Cas9 protein.
3. The virus-like particle of claim 2, wherein the Cas9 protein is a Cas9 nickase.
4. The virus-like particle of claim 2, wherein the Cas9 protein is a nuclease-inactive Cas9 (dCas9).
5. The virus-like particle of any one of claims 2-4, wherein the Cas9 protein is bound to a guide RNA (gRNA).
6. The virus-like particle of any one of claims 1-5, wherein the fusion protein further comprises a deaminase domain.
7. The virus-like particle of claim 6, wherein the deaminase domain is an adenosine deaminase domain.
8. The virus-like particle of claim 6, wherein the deaminase domain is a cytosine deaminase domain.
9. The virus-like particle of any one of claims 6-8, wherein the fusion protein comprises a base editor.

10. The virus-like particle of claim 9, wherein the base editor is ABE8e.
11. The virus-like particle of any one of claims 1-10, wherein the fusion protein comprises two NES, three NES, four NES, five NES, six NES, seven NES, eight NES, nine NES, or ten NES.
12. The virus-like particle of any one of claims 1-11, wherein the fusion protein further comprises a nuclear localization sequence (NLS).
13. The virus-like particle of claim 12, wherein the fusion protein further comprises two NLS.
14. The virus-like particle of any one of claims 1-13, wherein the cleavable linker is located between the napDNAbp and the NES.
15. The virus-like particle of any one of claims 1-14, wherein the cleavable linker comprises a protease cleavage site.
16. The virus-like particle of claim 15, wherein the protease cleavage site is a Moloney murine leukemia virus (MMLV) protease cleavage site or a Friend murine leukemia virus (FMLV) protease cleavage site.
17. The virus-like particle of claim 15 or 16, wherein the protease cleavage site comprises the amino acid sequence TSTLLMENSS (SEQ ID NO: 1), PRSSLYPALTP (SEQ ID NO: 2), VQALVLTQ (SEQ ID NO: 3), PLQVLTLNIERR (SEQ ID NO: 4), or an amino acid sequence at least 90% identical to any one of SEQ ID NOs: 1-4.
18. The virus-like particle of any one of claims 1-17, wherein the gag-pro polyprotein comprises an MMLV gag-pro polyprotein or an FMLV gag-pro polyprotein.

19. The virus-like particle of any one of claims 1-18, wherein the gag nucleocapsid protein comprises an MMLV gag nucleocapsid protein or an FMLV gag nucleocapsid protein.
20. The virus-like particle of any one of claims 1-19, wherein the fusion protein comprises the structure:
[gag nucleocapsid protein]-[1X-3X NES]-[cleavable linker]-[NLS]-[deaminase domain]-[napDNAbp]-[NLS] wherein]-[comprises an optional linker.
21. A virus-like particle comprising a group-specific antigen (gag) protease (pro) polyprotein, a nucleic acid programmable DNA binding protein (napDNAbp), and a fusion protein comprising a gag nucleocapsid protein and a nuclear export sequence (NES), encapsulated by a lipid membrane and a viral envelope glycoprotein.
22. The virus-like particle of claim 21, wherein the napDNAbp is a Cas9 protein.
23. The virus-like particle of claim 22, wherein the Cas9 protein is a Cas9 nickase.
24. The virus-like particle of claim 22, wherein the Cas9 protein is a nuclease-inactive Cas9 (dCas9).
25. The virus-like particle of any one of claims 22-24, wherein the Cas9 protein is bound to a guide RNA (gRNA).
26. The virus-like particle of any one of claims 21-25, wherein the napDNAbp is fused to a deaminase domain.
27. The virus-like particle of claim 26, wherein the deaminase domain is an adenosine deaminase domain.
28. The virus-like particle of claim 26, wherein the deaminase domain is a cytosine deaminase domain.

29. The virus-like particle of any one of claims 26-28, wherein the napDNAbp fused to a deaminase comprises a base editor.
30. The virus-like particle of claim 29, wherein the base editor is ABE8e.
31. The virus-like particle of any one of claims 21-30, wherein the fusion protein comprises two NES, three NES, four NES, five NES, six NES, seven NES, eight NES, nine NES, or ten NES.
32. The virus-like particle of any one of claims 21-31, wherein the napDNAbp and/or the deaminase is fused to a nuclear localization sequence (NLS).
33. The virus-like particle of claim 32, wherein the napDNAbp and/or the deaminase is fused to two NLS, or the napDNAbp is fused to a first NLS and the deaminase is fused to a second NLS.
34. The virus-like particle of any one of claims 21-33, wherein the napDNAbp and the fusion protein were previously fused via a cleavable linker located between the napDNAbp and the NES, and the cleavable linker has subsequently been cleaved by the protease of the gag-pro-polyprotein.
35. The virus-like particle of claim 34, wherein the cleavable linker comprises a protease cleavage site.
36. The virus-like particle of claim 35, wherein the protease cleavage site is a Moloney murine leukemia virus (MMLV) protease cleavage site or a Friend murine leukemia virus (FMLV) protease cleavage site.
37. The virus-like particle of claim 35 or 36, wherein the protease cleavage site comprises the amino acid sequence TSTLLMENSS (SEQ ID NO: 1), PRSSLYPALTP (SEQ ID NO: 2), VQALVLTQ (SEQ ID NO: 3), PLQVLTLNIERR (SEQ ID NO: 4), or an amino acid sequence at least 90% identical to any one of SEQ ID NOs: 1-4.

38. The virus-like particle of any one of claims 21-37, wherein the gag-pro polyprotein comprises an MMLV gag-pro polyprotein or an FMLV gag-pro polyprotein.
39. The virus-like particle of any one of claims 21-38, wherein the gag nucleocapsid protein comprises an MMLV gag nucleocapsid protein or an FMLV gag nucleocapsid protein.
40. The virus-like particle of any one of claims 21-39, wherein the fusion protein comprises the structure:
[gag nucleocapsid protein]-[1X-3X NES] wherein]-[comprises an optional linker.
41. The virus-like particle of any one of claims 29-40, wherein the base editor comprises the structure:
[NLS]-[deaminase domain]-[napDNAbp]-[NLS] wherein]-[comprises an optional linker.
42. The virus-like particle of any one of claims 1-41, wherein the viral envelope glycoprotein is an adenoviral envelope glycoprotein, an adeno-associated viral envelope glycoprotein, a retroviral envelope glycoprotein, or a lentiviral envelope glycoprotein.
43. The virus-like particle of claim 42, wherein the viral envelope glycoprotein is a retroviral envelope glycoprotein.
44. The virus-like particle of claim 42, wherein the viral envelope glycoprotein is a vesicular stomatitis virus G protein (VSV-G), a baboon retroviral envelope glycoprotein (BaEVRless), a FuG-B2 envelope glycoprotein, an HIV-1 envelope glycoprotein, or an ecotropic murine leukemia virus (MLV) envelope glycoprotein.
45. The virus-like particle of any one of claims 1-44, wherein the viral envelope glycoprotein targets the virus-like particle to a particular cell type.
46. The virus-like particle of claim 45, wherein the cell type is immune cells, neural cells, or retinal pigment epithelium cells.

47. The virus-like particle of claim 45, wherein the viral envelope glycoprotein is a VSV-G protein, and wherein the VSV-G protein targets the virus-like particle to retinal pigment epithelium (RPE) cells.
48. The virus-like particle of claim 45, wherein the viral envelope glycoprotein is an HIV-1 envelope glycoprotein, and wherein the HIV-1 envelope glycoprotein targets the virus-like particle to CD4+ cells.
49. The virus-like particle of claim 45, wherein the viral envelope glycoprotein is a FuG-B2 envelope glycoprotein, and wherein the FuG-B2 envelope glycoprotein targets the virus-like particle to neurons.
50. A cell comprising the virus-like particle of any one of claims 1-49.
51. A plurality of polynucleotides comprising:
- (i) a first polynucleotide comprising a nucleic acid sequence encoding a viral envelope glycoprotein;
 - (ii) a second polynucleotide comprising a nucleic acid sequence encoding a group-specific antigen (gag) protease (pro) polyprotein;
 - (iii) a third polynucleotide comprising a nucleic acid sequence encoding a fusion protein, wherein the fusion protein comprises:
 - (a) a group-specific antigen (gag) nucleocapsid protein;
 - (b) a nucleic acid programmable DNA binding protein (napDNAbp);
 - (c) a cleavable linker; and
 - (d) a nuclear export sequence (NES); and
 - (iv) a fourth polynucleotide comprising a nucleic acid sequence encoding a guide RNA (gRNA), wherein the gRNA binds to the napDNAbp of the fusion protein encoded by the third polynucleotide.
52. The plurality of polynucleotides of claim 51, wherein the ratio of the second polynucleotide to the third polynucleotide is approximately 10:1, approximately 9:1, approximately 8:1, approximately 7:1, approximately 6:1, approximately 5:1, approximately

4:1, approximately 3:1, approximately 2:1, approximately 1.5:1, approximately 1:1, or approximately 0.5:1.

53. The plurality of polynucleotides of claim 52, wherein the ratio of the second polynucleotide to the third polynucleotide is approximately 3:1.

54. The plurality of polynucleotides of any one of claims 51-53, wherein the napDNAbp is a Cas9 protein.

55. The plurality of polynucleotides of claim 54, wherein the Cas9 protein is a Cas9 nickase.

56. The plurality of polynucleotides of claim 54, wherein the Cas9 protein is a nuclease-inactive Cas9 (dCas9).

57. The plurality of polynucleotides of any one of claims 51-56, wherein the fusion protein further comprises a deaminase domain.

58. The plurality of polynucleotides of claim 57, wherein the deaminase domain is an adenosine deaminase domain.

59. The plurality of polynucleotides of claim 57, wherein the deaminase domain is a cytosine deaminase domain.

60. The plurality of polynucleotides of any one of claims 57-59, wherein the fusion protein comprises a base editor.

61. The plurality of polynucleotides of claim 60, wherein the base editor is ABE8e.

62. The plurality of polynucleotides of any one of claims 51-61, wherein the fusion protein comprises two NES, three NES, four NES, five NES, six NES, seven NES, eight NES, nine NES, or ten NES.

63. The plurality of polynucleotides of any one of claims 51-62, wherein the fusion protein further comprises a nuclear localization sequence (NLS).
64. The plurality of polynucleotides of claim 63, wherein the fusion protein further comprises two NLS.
65. The plurality of polynucleotides of any one of claims 51-64, wherein the cleavable linker is located between the napDNAbp and the NES.
66. The plurality of polynucleotides of any one of claims 51-65, wherein the cleavable linker comprises a protease cleavage site.
67. The plurality of polynucleotides of claim 66, wherein the protease cleavage site is a Moloney murine leukemia virus (MMLV) protease cleavage site or a Friend murine leukemia virus (FMLV) protease cleavage site.
68. The plurality of polynucleotides of claim 66 or 67, wherein the protease cleavage site comprises the amino acid sequence TSTLLMENSS (SEQ ID NO: 1), PRSSLYPALTP (SEQ ID NO: 2), VQALVLTQ (SEQ ID NO: 3), PLQVLTNIERR (SEQ ID NO: 4), or an amino acid sequence at least 90% identical to any one of SEQ ID NOs: 1-4.
69. The plurality of polynucleotides of any one of claims 51-68, wherein the gag-pro polyprotein comprises an MMLV gag-pro polyprotein or an FMLV gag-pro polyprotein.
70. The plurality of polynucleotides of any one of claims 51-69, wherein the gag nucleocapsid protein comprises an MMLV gag nucleocapsid protein or an FMLV gag nucleocapsid protein.
71. The plurality of polynucleotides of any one of claims 51-70, wherein the fusion protein comprises the structure:
[gag nucleocapsid protein]-[1X-3X NES]-[cleavable linker]-[NLS]-[deaminase domain]-[napDNAbp]-[NLS] wherein]-[comprises an optional linker.

72. The plurality of polynucleotides of any one of claims 51-71, wherein the viral envelope glycoprotein is an adenoviral envelope glycoprotein, an adeno-associated viral envelope glycoprotein, a retroviral envelope glycoprotein, or a lentiviral envelope glycoprotein.
73. The plurality of polynucleotides of claim 72, wherein the viral envelope glycoprotein is a retroviral envelope glycoprotein.
74. The plurality of polynucleotides of claim 72, wherein the viral envelope glycoprotein is a vesicular stomatitis virus G protein (VSV-G), a baboon retroviral envelope glycoprotein (BaEVRless), a FuG-B2 envelope glycoprotein, an HIV-1 viral envelope glycoprotein, or an ecotropic murine leukemia virus (MLV) envelope glycoprotein.
75. One or more vectors comprising the plurality of polynucleotides of any one of claims 51-74.
76. The one or more vectors of claim 75, wherein each of the first, second, third, and fourth polynucleotides are on separate vectors.
77. The one or more vectors of claim 75, wherein one or more of the first, second, third, and fourth polynucleotides are on the same vector.
78. A cell comprising the plurality of polynucleotides of any one of claims 51-74 or the one or more vectors of any one of claims 75-77.
79. A method of making a virus-like particle (VLP) comprising transfecting the plurality of polynucleotides of any one of claims 51-74 or the one or more vectors of any one of claims 75-77 into a cell.
80. The method of claim 79, wherein the viral envelope glycoprotein targets the VLP to a particular cell type.

81. The method of claim 80, wherein the cell type is immune cells, neural cells, or retinal pigment epithelium cells.
82. The method of claim 80, wherein the viral envelope glycoprotein is a VSV-G protein, and wherein the VSV-G protein targets the virus-like particle to retinal pigment epithelium (RPE) cells.
83. The method of claim 80, wherein the viral envelope glycoprotein is an HIV-1 envelope glycoprotein, and wherein the HIV-1 envelope glycoprotein targets the virus-like particle to CD4+ cells.
84. The method of claim 80, wherein the viral envelope glycoprotein is a FuG-B2 envelope glycoprotein, and wherein the FuG-B2 envelope glycoprotein targets the virus-like particle to neurons.
85. A pharmaceutical composition comprising a virus-like particle (VLP) comprising a group-specific antigen (gag) protease (pro) polyprotein and a fusion protein encapsulated by a viral envelope glycoprotein, wherein the fusion protein comprises:
- (i) a gag nucleocapsid protein;
 - (ii) a nucleic acid programmable DNA binding protein (napDNAbp);
 - (iii) a cleavable linker; and
 - (iv) a nuclear export sequence (NES).
86. The pharmaceutical composition of claim 85, wherein the napDNAbp is a Cas9 protein.
87. The pharmaceutical composition of claim 86, wherein the Cas9 protein is a Cas9 nickase.
88. The pharmaceutical composition of claim 86, wherein the Cas9 protein is a nuclease-inactive Cas9 (dCas9).

89. The pharmaceutical composition of any one of claims 85-88, wherein the Cas9 protein is bound to a guide RNA (gRNA).
90. The pharmaceutical composition of any one of claims 85-89, wherein the fusion protein further comprises a deaminase domain.
91. The pharmaceutical composition of claim 90, wherein the deaminase domain is an adenosine deaminase domain.
92. The pharmaceutical composition of claim 90, wherein the deaminase domain is a cytosine deaminase domain.
93. The pharmaceutical composition of any one of claims 90-92, wherein the fusion protein comprises a base editor.
94. The pharmaceutical composition of claim 93, wherein the base editor is ABE8e.
95. The pharmaceutical composition of any one of claims 85-94, wherein the fusion protein comprises two NES, three NES, four NES, five NES, six NES, seven NES, eight NES, nine NES, or ten NES.
96. The pharmaceutical composition of any one of claims 85-95, wherein the fusion protein further comprises a nuclear localization sequence (NLS).
97. The pharmaceutical composition of claim 96, wherein the fusion protein further comprises two NLS.
98. The pharmaceutical composition of any one of claims 85-97, wherein the cleavable linker is located between the napDNAbp and the NES.
99. The pharmaceutical composition of any one of claims 85-98, wherein the cleavable linker comprises a protease cleavage site.

100. The pharmaceutical composition of claim 99, wherein the protease cleavage site is a Moloney murine leukemia virus (MMLV) protease cleavage site or a Friend murine leukemia virus (FMLV) protease cleavage site.

101. The pharmaceutical composition of claim 99 or 100, wherein the protease cleavage site comprises the amino acid sequence TSTLLMENS (SEQ ID NO: 1), PRSSLYPALTP (SEQ ID NO: 2), VQALVLTQ (SEQ ID NO: 3), PLQVLTNIERR (SEQ ID NO: 4), or an amino acid sequence at least 90% identical to any one of SEQ ID NOs: 1-4.

102. The pharmaceutical composition of any one of claims 85-101, wherein the gag-pro polyprotein comprises an MMLV gag-pro polyprotein or an FMLV gag-pro polyprotein.

103. The pharmaceutical composition of any one of claims 85-102, wherein the gag nucleocapsid protein comprises an MMLV gag nucleocapsid protein or an FMLV gag nucleocapsid protein.

104. The pharmaceutical composition of any one of claims 85-103, wherein the fusion protein comprises the structure:

[gag nucleocapsid protein]-[1X-3X NES]-[cleavable linker]-[NLS]-[deaminase domain]-[napDNAbp]-[NLS] wherein]-[comprises an optional linker.

105. A pharmaceutical composition comprising a virus-like particle comprising a group-specific antigen (gag) protease (pro) polyprotein, a nucleic acid programmable DNA binding protein (napDNAbp), and a fusion protein encapsulated by a lipid membrane and a viral envelope glycoprotein, wherein the fusion protein comprises a gag nucleocapsid protein and a nuclear export sequence (NES).

106. The pharmaceutical composition of claim 105, wherein the napDNAbp is a Cas9 protein.

107. The pharmaceutical composition of claim 106, wherein the Cas9 protein is a Cas9 nickase.

108. The pharmaceutical composition of claim 106, wherein the Cas9 protein is a nuclease-inactive Cas9 (dCas9).
109. The pharmaceutical composition of any one of claims 106-108, wherein the Cas9 protein is bound to a guide RNA (gRNA).
110. The pharmaceutical composition of any one of claims 105-109, wherein the napDNAbp is fused to a deaminase domain.
111. The pharmaceutical composition of claim 110, wherein the deaminase domain is an adenosine deaminase domain.
112. The pharmaceutical composition of claim 110, wherein the deaminase domain is a cytosine deaminase domain.
113. The pharmaceutical composition of any one of claims 110-112, wherein the napDNAbp fused to a deaminase comprises a base editor.
114. The pharmaceutical composition of claim 113, wherein the base editor is ABE8e.
115. The pharmaceutical composition of any one of claims 105-114, wherein the fusion protein comprises two NES, three NES, four NES, five NES, six NES, seven NES, eight NES, nine NES, or ten NES.
116. The pharmaceutical composition of any one of claims 105-115, wherein the napDNAbp and/or the deaminase is fused to a nuclear localization sequence (NLS).
117. The pharmaceutical composition of claim 116, wherein the napDNAbp and/or the deaminase is fused to two NLS, or the napDNAbp is fused to a first NLS and the deaminase is fused to a second NLS.
118. The pharmaceutical composition of any one of claims 105-117, wherein the napDNAbp and the fusion protein were previously fused via a cleavable linker located

between the napDNAbp and the NES, and the cleavable linker has subsequently been cleaved by the protease of the gag-pro-polyprotein.

119. The pharmaceutical composition of claim 118, wherein the cleavable linker comprises a protease cleavage site.

120. The pharmaceutical composition of claim 119, wherein the protease cleavage site is a Moloney murine leukemia virus (MMLV) protease cleavage site or a Friend murine leukemia virus (FMLV) protease cleavage site.

121. The pharmaceutical composition of claim 119 or 120, wherein the protease cleavage site comprises the amino acid sequence TSTLLMENSS (SEQ ID NO: 1), PRSSLYPALTP (SEQ ID NO: 2), VQALVLTQ (SEQ ID NO: 3), PLQVLTNIERR (SEQ ID NO: 4), or an amino acid sequence at least 90% identical to any one of SEQ ID NOs: 1-4.

122. The pharmaceutical composition of any one of claims 105-121, wherein the gag-pro polyprotein comprises an MMLV gag-pro polyprotein or an FMLV gag-pro polyprotein.

123. The pharmaceutical composition of any one of claims 105-122, wherein the gag nucleocapsid protein comprises an MMLV gag nucleocapsid protein or an FMLV gag nucleocapsid protein.

124. The pharmaceutical composition of any one of claims 105-123, wherein the fusion protein comprises the structure:

[gag nucleocapsid protein]-[1X-3X NES] wherein]-[comprises an optional linker.

125. The pharmaceutical composition of any one of claims 113-124, wherein the base editor comprises the structure:

[NLS]-[deaminase domain]-[napDNAbp]-[NLS] wherein]-[comprises an optional linker.

126. The pharmaceutical composition of any one of claims 85-125, wherein the viral envelope glycoprotein is an adenoviral envelope glycoprotein, an adeno-associated viral

envelope glycoprotein, a retroviral envelope glycoprotein, or a lentiviral envelope glycoprotein.

127. The pharmaceutical composition of claim 126, wherein the viral envelope glycoprotein is a retroviral envelope glycoprotein.

128. The pharmaceutical composition of claim 126, wherein the viral envelope glycoprotein is a vesicular stomatitis virus G protein (VSV-G), a baboon retroviral envelope glycoprotein (BaEVRless), a FuG-B2 envelope glycoprotein, an HIV-1 viral envelope glycoprotein, or an ecotropic murine leukemia virus (MLV) envelope glycoprotein.

129. The pharmaceutical composition of any one of claims 85-128, wherein the viral envelope glycoprotein targets the virus-like particle to a particular cell type.

130. The pharmaceutical composition of claim 129, wherein the cells are immune cells, neural cells, or retinal pigment epithelium cells.

131. The pharmaceutical composition of claim 129, wherein the viral envelope glycoprotein is a VSV-G protein, and wherein the VSV-G protein targets the virus-like particle to retinal pigment epithelium (RPE) cells.

132. The pharmaceutical composition of claim 129, wherein the viral envelope glycoprotein is an HIV-1 envelope glycoprotein, and wherein the HIV-1 envelope glycoprotein targets the virus-like particle to CD4+ cells.

133. The pharmaceutical composition of claim 129, wherein the viral envelope glycoprotein is a FuG-B2 envelope glycoprotein, and wherein the FuG-B2 envelope glycoprotein targets the virus-like particle to neurons.

134. A method of editing a nucleic acid molecule in a target cell by base editing comprising contacting the target cell with the virus-like particle of any one of claims 1-49 or the pharmaceutical composition of any one of claims 85-133, thereby installing one or more modifications to the nucleic acid molecule at a target site.

135. The method of claim 134, wherein the cell is a mammalian cell.
136. The method of claim 134 or 135, wherein the cell is a human cell.
137. The method of any one of claims 134-136, wherein the cell is in a subject.
138. The method of claim 137, wherein the subject is a human.
139. The method of any one of claims 134-138, wherein the one or more modifications to the nucleic acid molecule are associated with reducing, relieving, or preventing the symptoms of a disease or disorder, optionally wherein the disease or disorder is a CNS disorder, liver disorder, or ocular disorder.
140. A fusion protein comprising:
- (i) a group-specific antigen (gag) nucleocapsid protein;
 - (ii) a nucleic acid programmable DNA binding protein (napDNAbp);
 - (iii) a cleavable linker; and
 - (iv) a nuclear export sequence (NES).
141. The fusion protein of claim 140, wherein the napDNAbp is a Cas9 protein.
142. The fusion protein of claim 141, wherein the Cas9 protein is a Cas9 nickase.
143. The fusion protein of claim 141, wherein the Cas9 protein is a nuclease-inactive Cas9 protein.
144. The fusion protein of any one of claims 141-143, wherein the Cas9 protein is bound to a guide RNA (gRNA).
145. The fusion protein of any one of claims 140-144 further comprising a deaminase domain.

146. The fusion protein of claim 145, wherein the deaminase domain is an adenosine deaminase domain.
147. The fusion protein of claim 145, wherein the deaminase domain is a cytosine deaminase domain.
148. The fusion protein of any one of claims 145-147, wherein the fusion protein comprises a base editor.
149. The fusion protein of claim 148, wherein the base editor is ABE8e.
150. The fusion protein of any one of claims 140-149, wherein the fusion protein comprises two NES, three NES, four NES, five NES, six NES, seven NES, eight NES, nine NES, or ten NES.
151. The fusion protein of any one of claims 140-150 further comprising a nuclear localization sequence (NLS).
152. The fusion protein of claim 151 further comprising two NLS.
153. The fusion protein of any one of claims 140-152, wherein the cleavable linker is located between the napDNAbp and the NES.
154. The fusion protein of any one of claims 140-153, wherein the cleavable linker comprises a protease cleavage site.
155. The fusion protein of claim 154, wherein the protease cleavage site is a Moloney murine leukemia virus (MMLV) protease cleavage site or a Friend murine leukemia virus (FMLV) protease cleavage site.
156. The fusion protein of claim 154 or 155, wherein the protease cleavage site comprises the amino acid sequence TSTLLMENSS (SEQ ID NO: 1), PRSSLYPALTP (SEQ ID NO: 2),

VQALVLTQ (SEQ ID NO: 3), PLQVLTLNIERR (SEQ ID NO: 4), or an amino acid sequence at least 90% identical to any one of SEQ ID NOs: 1-4.

157. The fusion protein of any one of claims 140-156, wherein the gag nucleocapsid protein comprises an MMLV gag nucleocapsid protein or an FMLV gag nucleocapsid protein.

158. The fusion protein of any one of claims 140-157, wherein the fusion protein comprises the structure:

[gag nucleocapsid protein]-[1X-3X NES]-[cleavable linker]-[NLS]-[deaminase domain]-[napDNAbp]-[NLS] wherein]-[comprises an optional linker.

159. A polynucleotide encoding the fusion protein of any one of claims 140-158.

160. A vector comprising the polynucleotide of claim 159.

161. A cell comprising the fusion protein of any one of claims 140-158, the polynucleotide of claim 159, or vector of claim 160.

162. A kit comprising the virus-like particle of any one of claims 1-49, the plurality of polynucleotides of any one of claims 85-133, or the fusion protein of any one of claims 140-158.

163. A virus-like particle of any one of claims 1-49 produced by transfecting, transducing, electroporating, or otherwise inserting the plurality of polynucleotides of any one of claims 51-74 or the one or more vectors of any one of claims 75-77 into a cell and expressing the components of the virus-like particle from the plurality of polynucleotides or one or more vectors in the cell, thereby allowing the virus-like particle to spontaneously assemble in the cell.

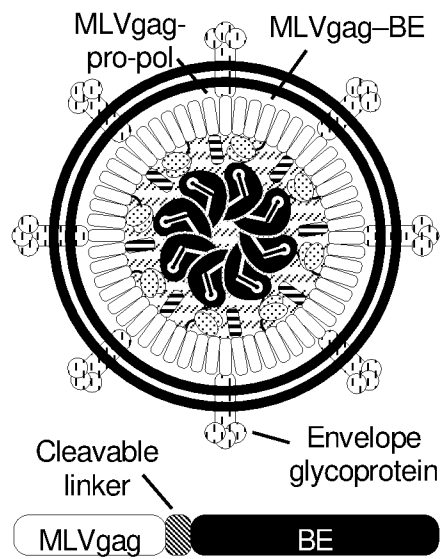


FIG. 1A

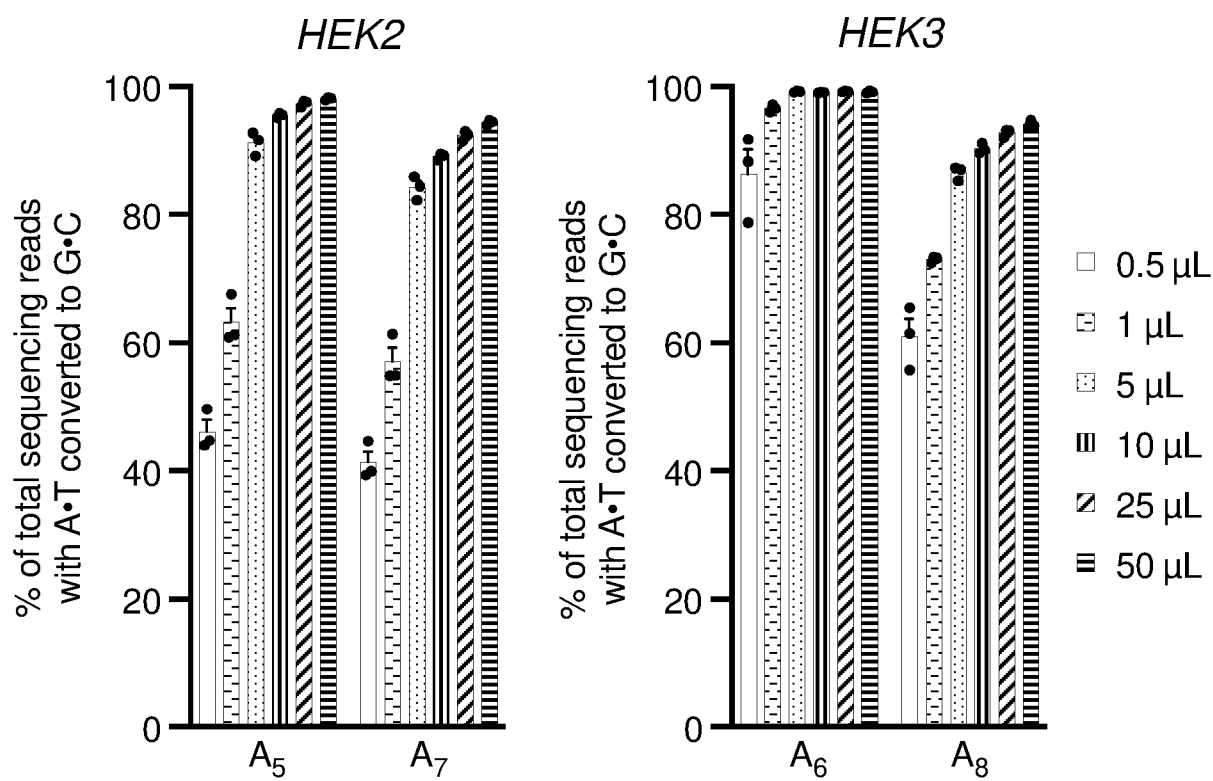
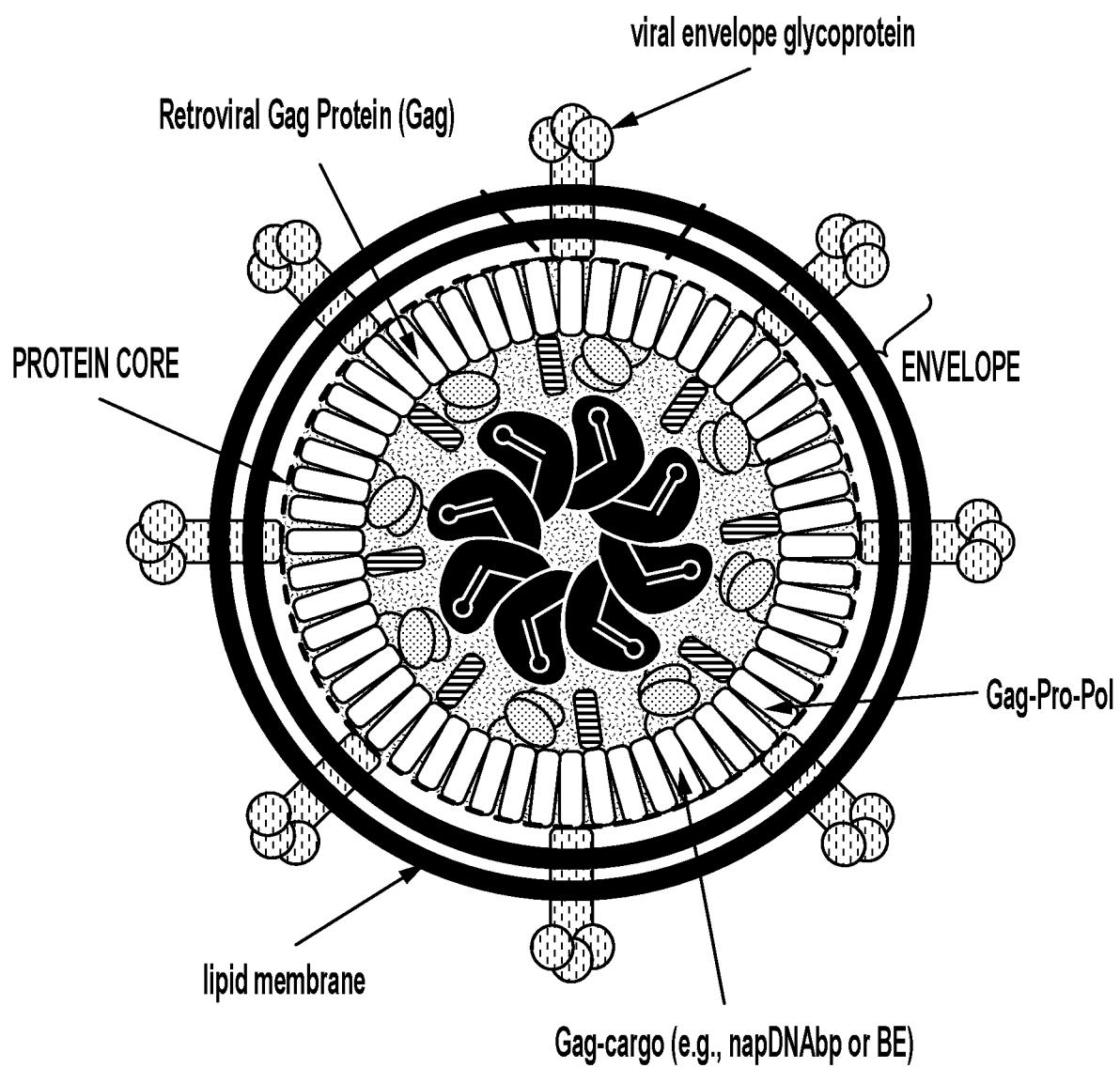
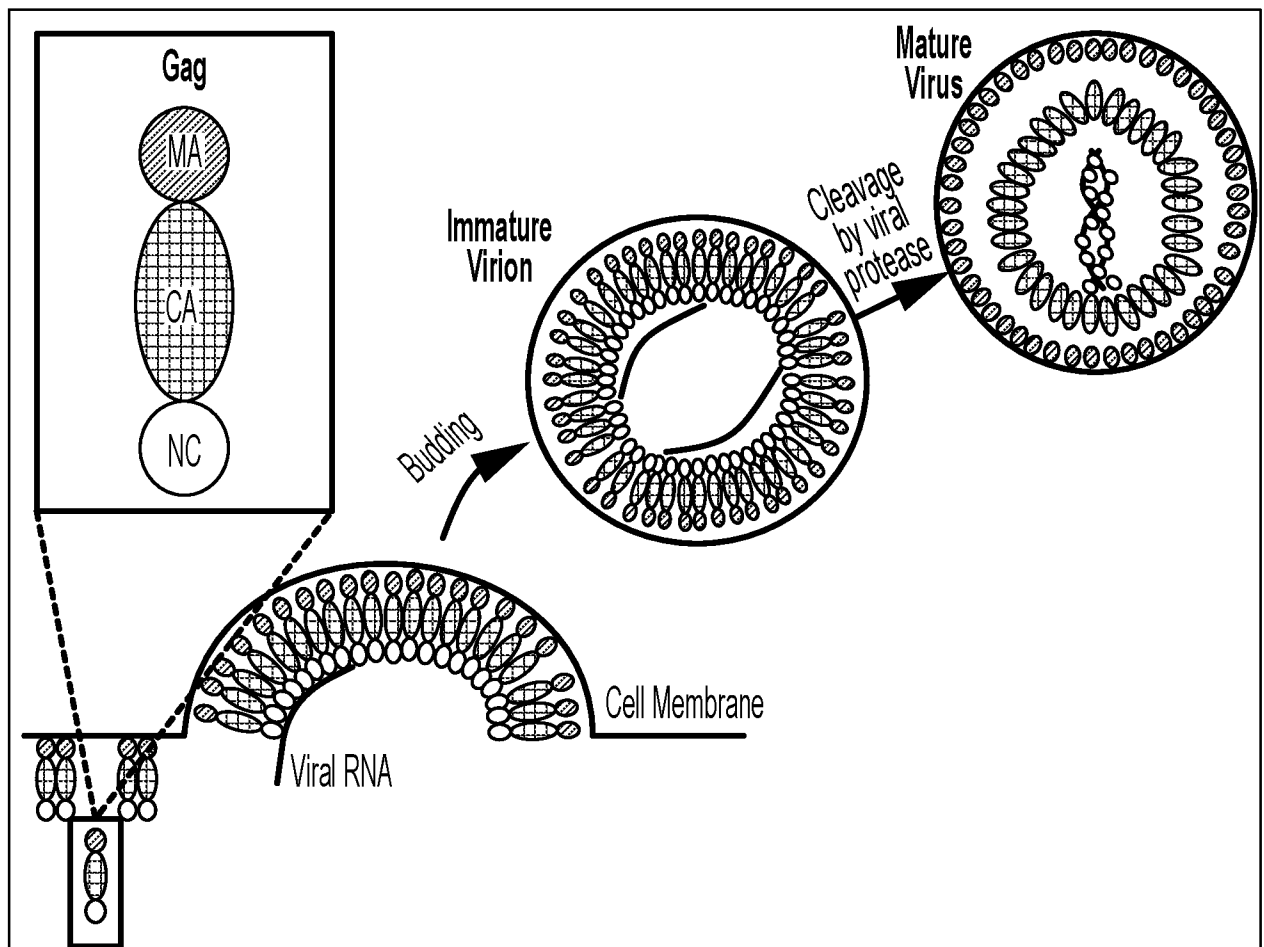


FIG. 1B

**FIG. 1C**

**FIG. 1D**

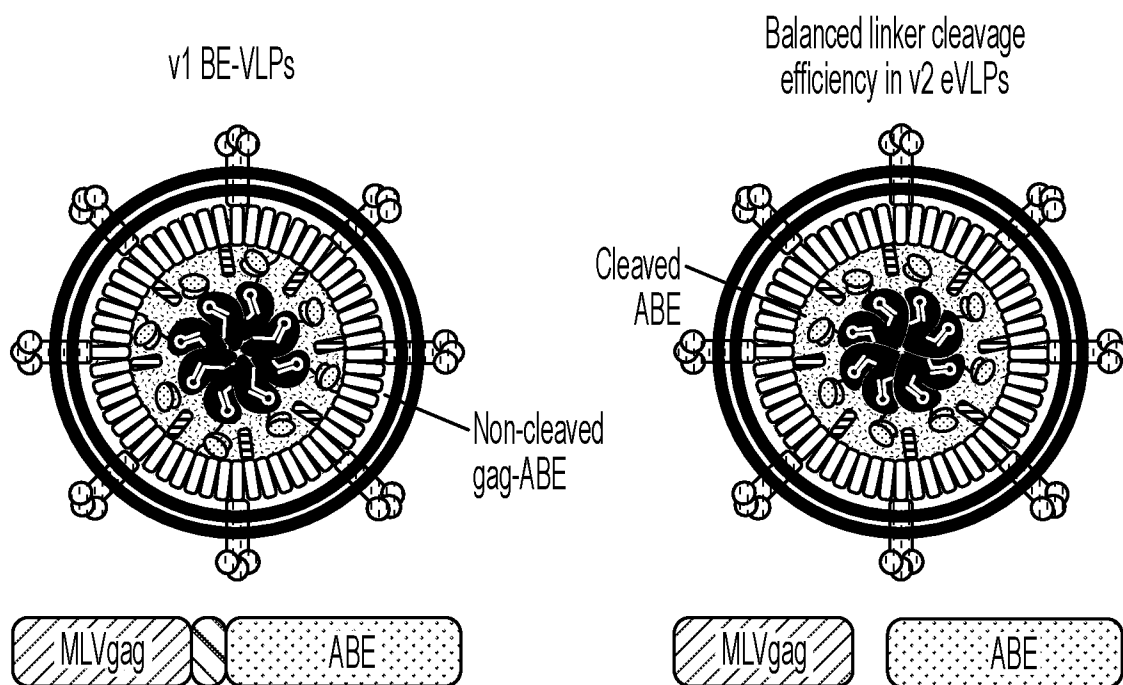


FIG. 2A

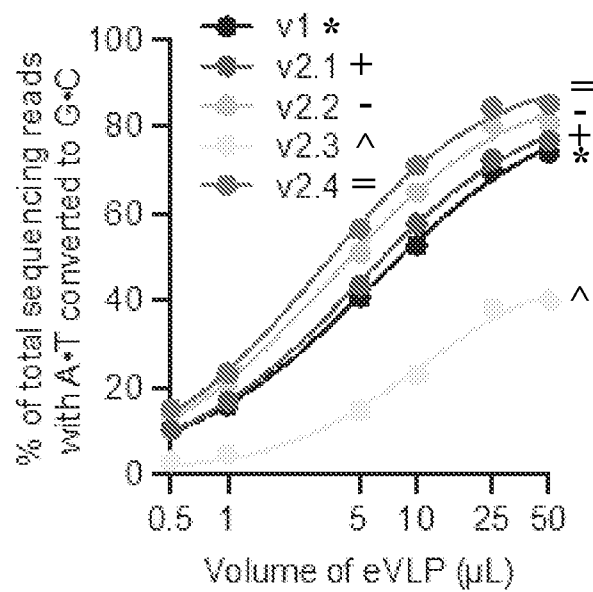
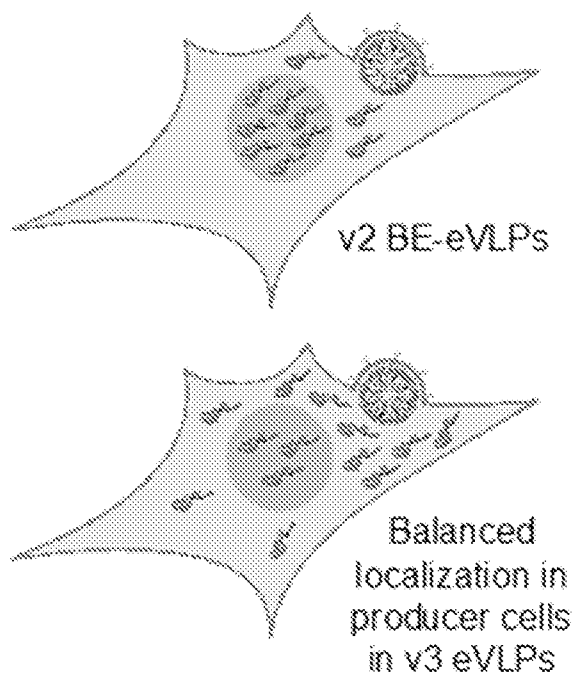
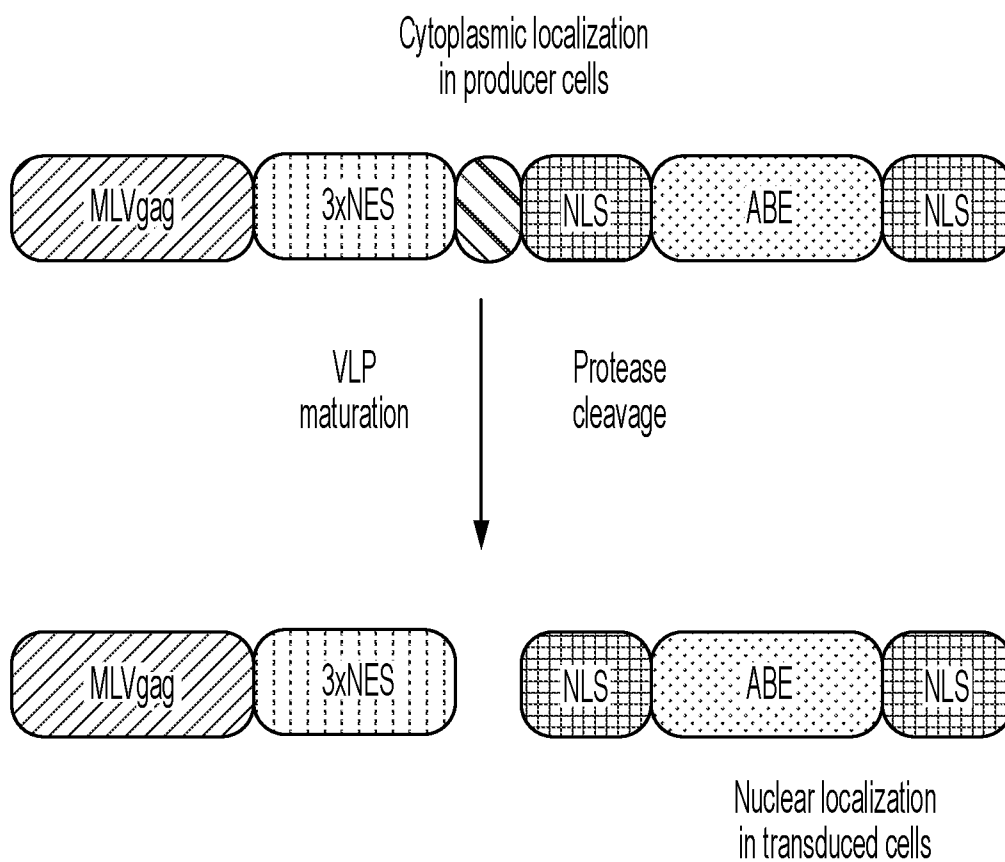


FIG. 2B

**FIG. 2C****FIG. 2D**

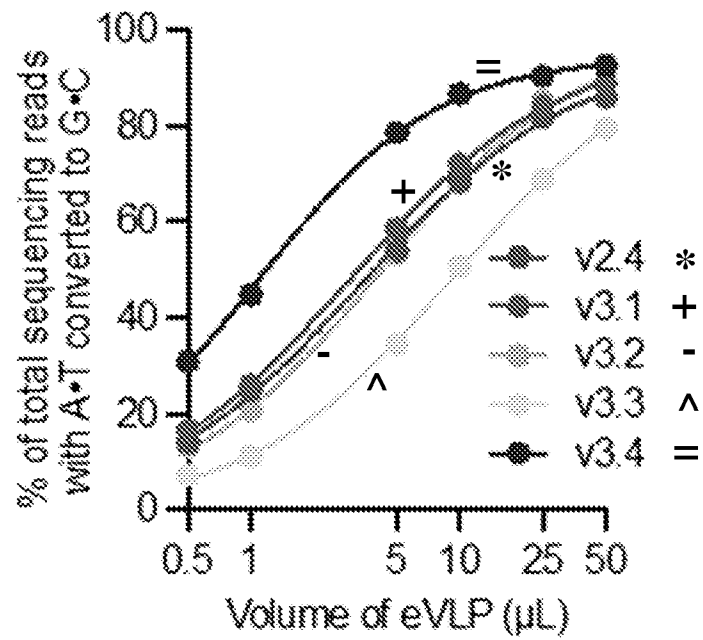


FIG. 2E

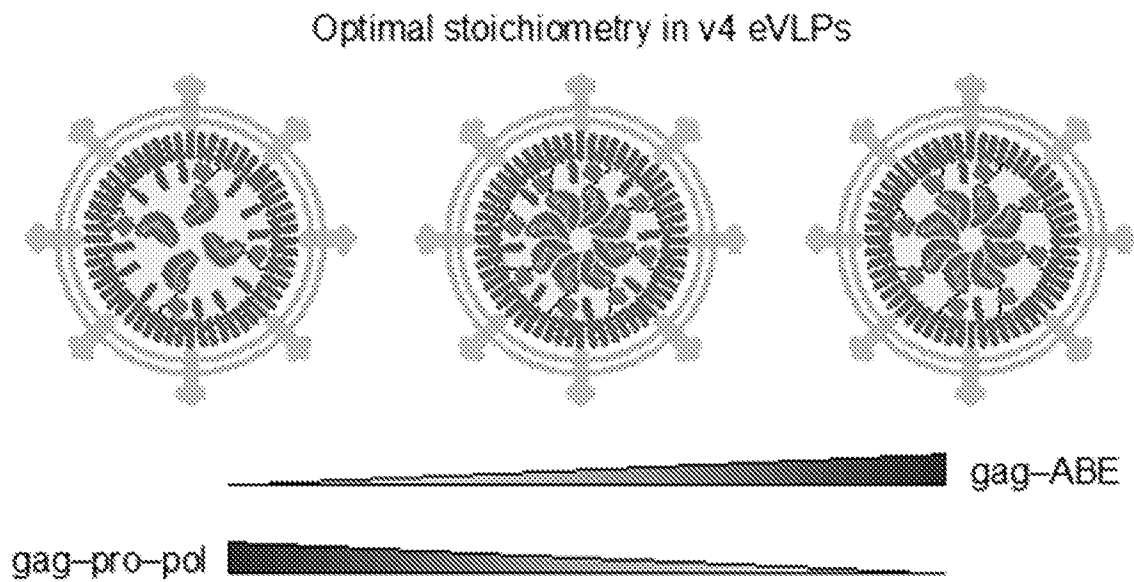


FIG. 2F

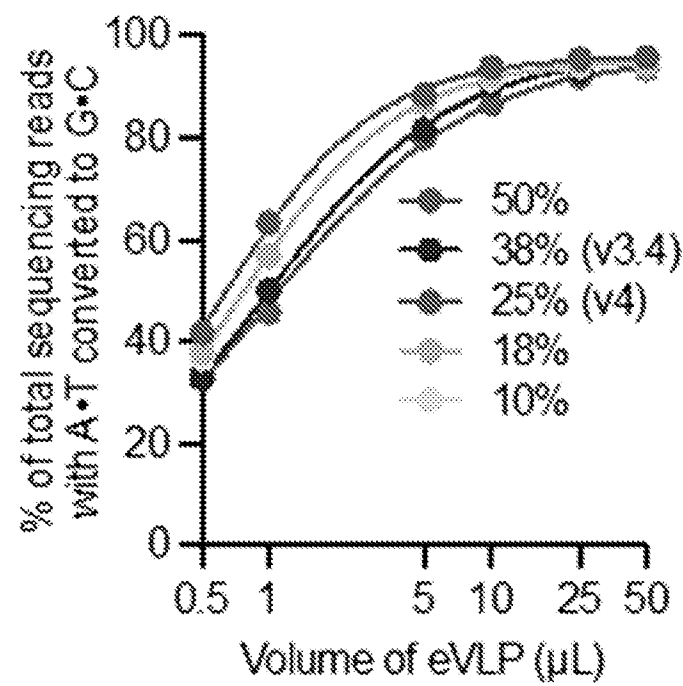


FIG. 2G

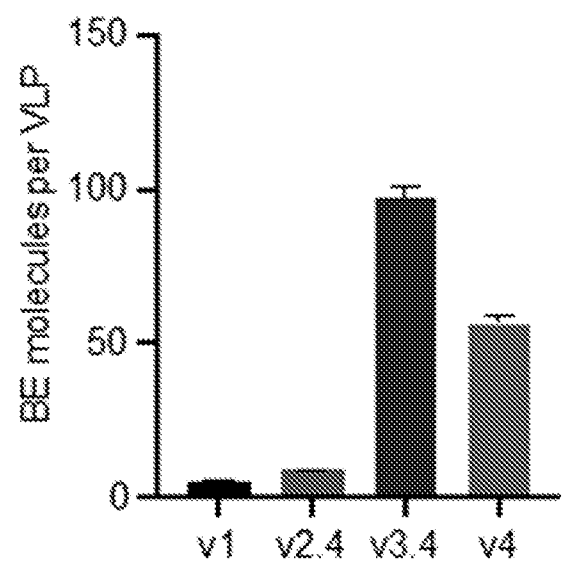


FIG. 3A

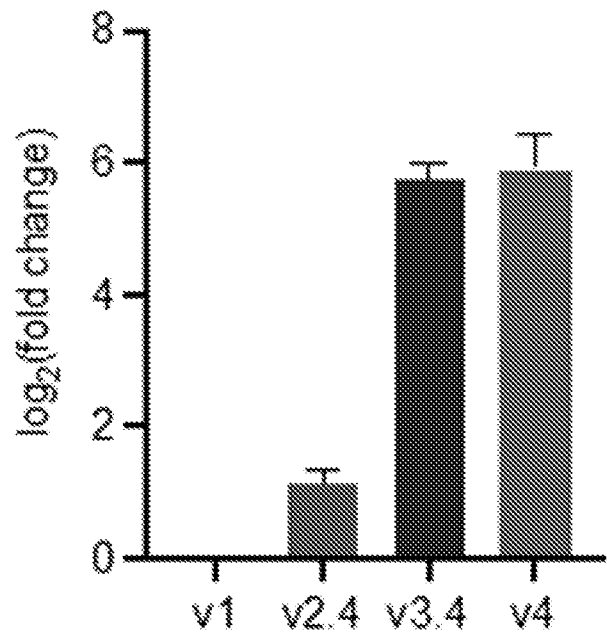


FIG. 3B

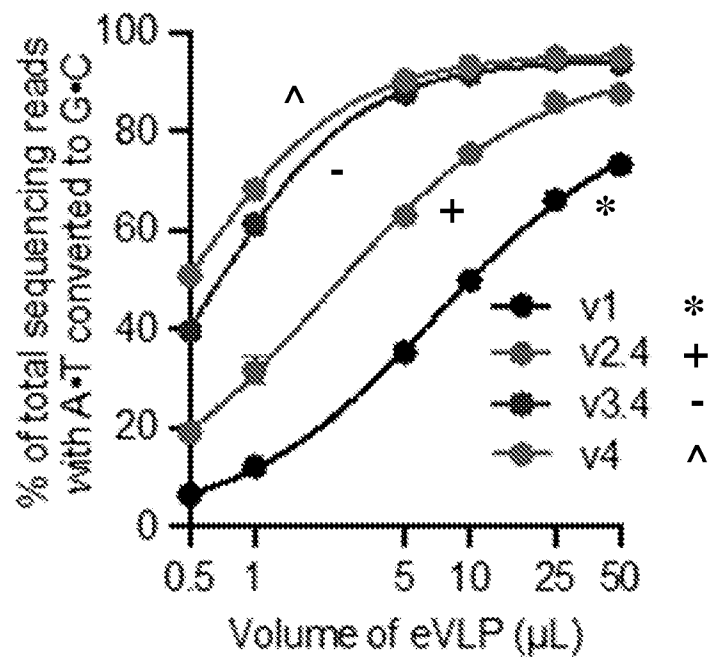


FIG. 3C

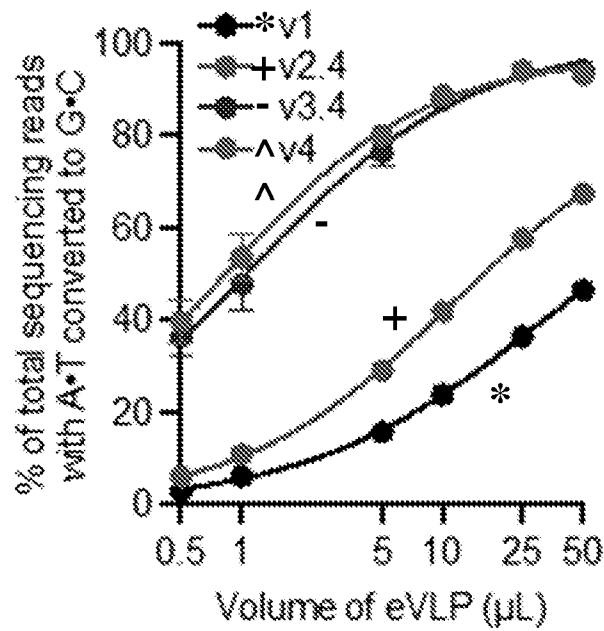


FIG. 3D

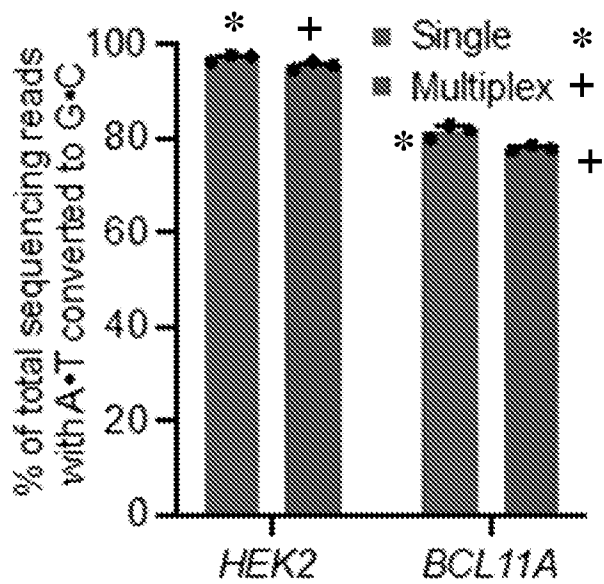


FIG. 3E

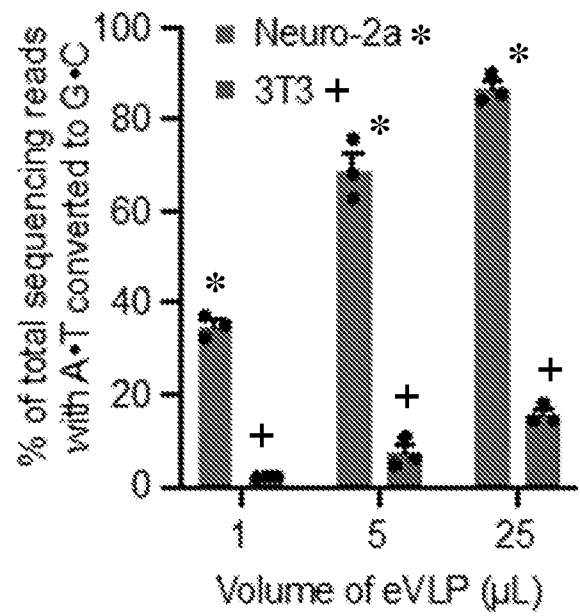


FIG. 3F

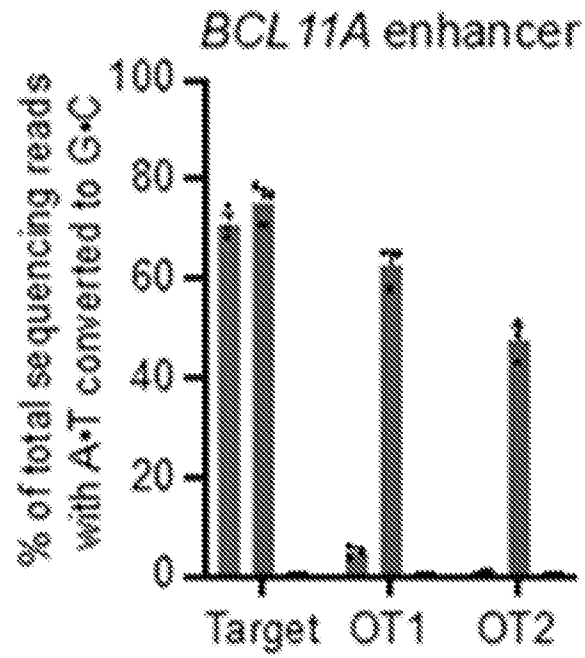


FIG. 3G

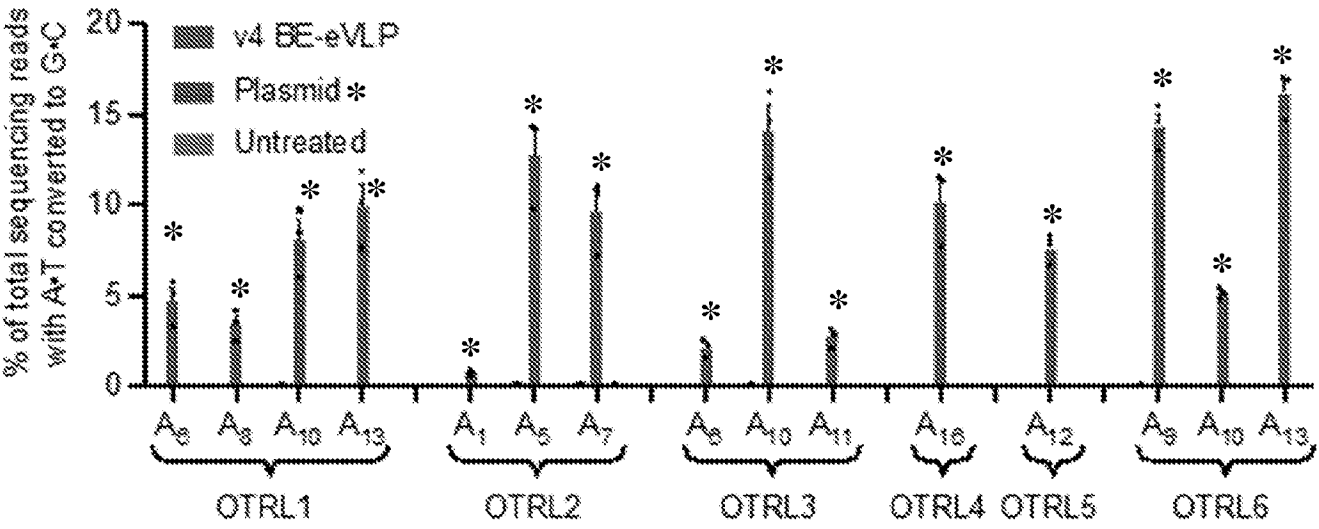


FIG. 3H

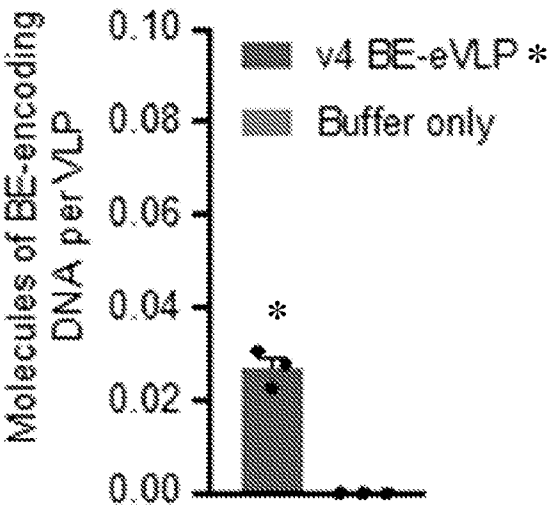


FIG. 3I

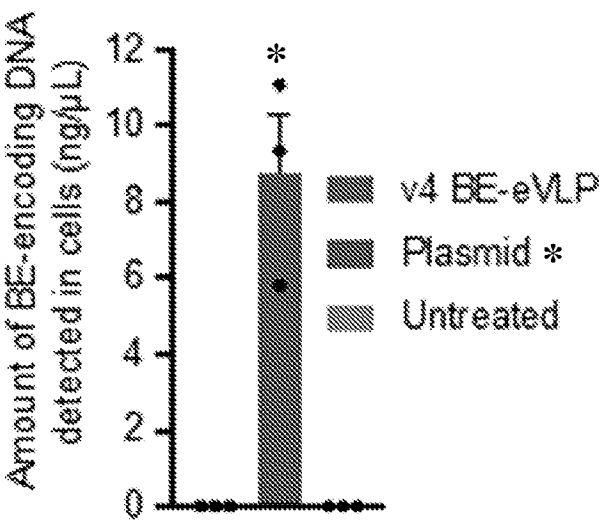


FIG. 3J

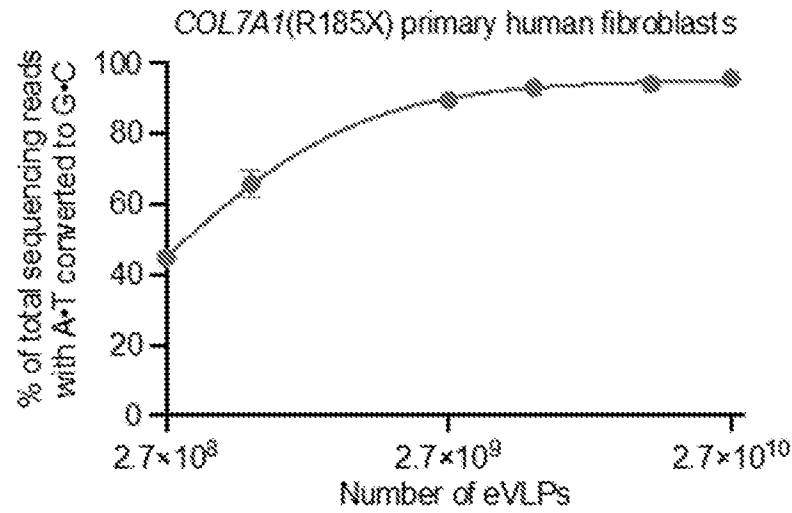


FIG. 4A

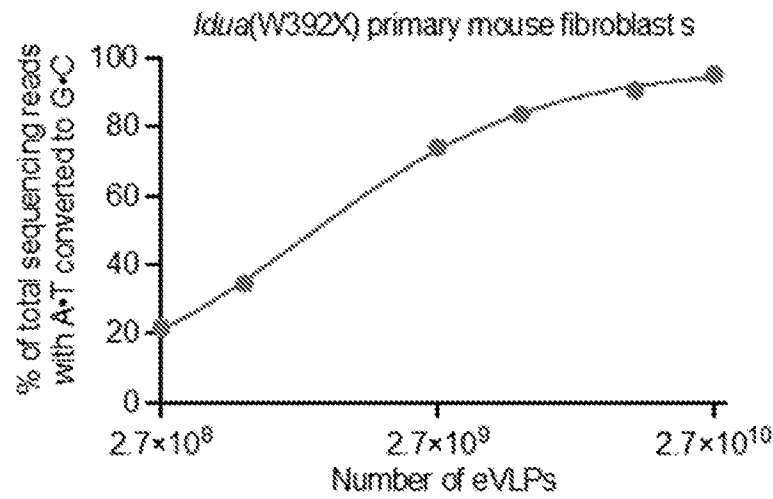


FIG. 4B

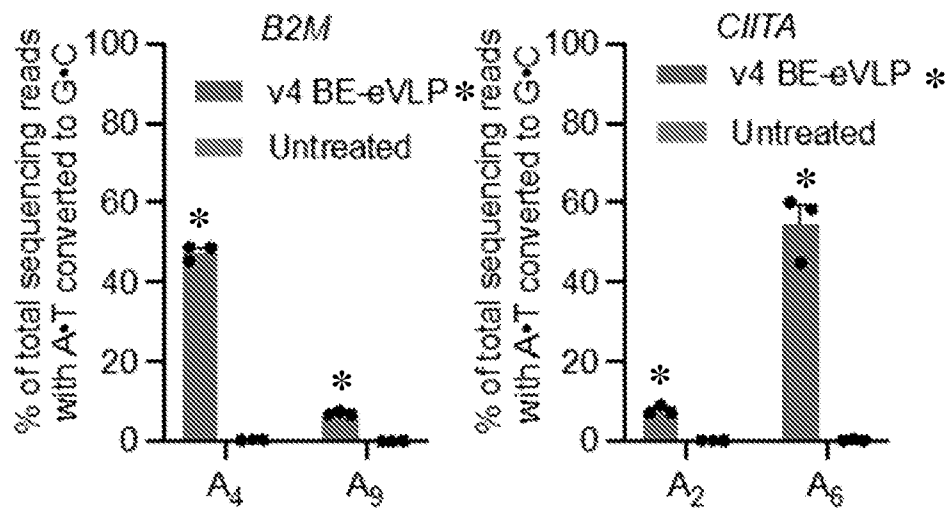


FIG. 4C

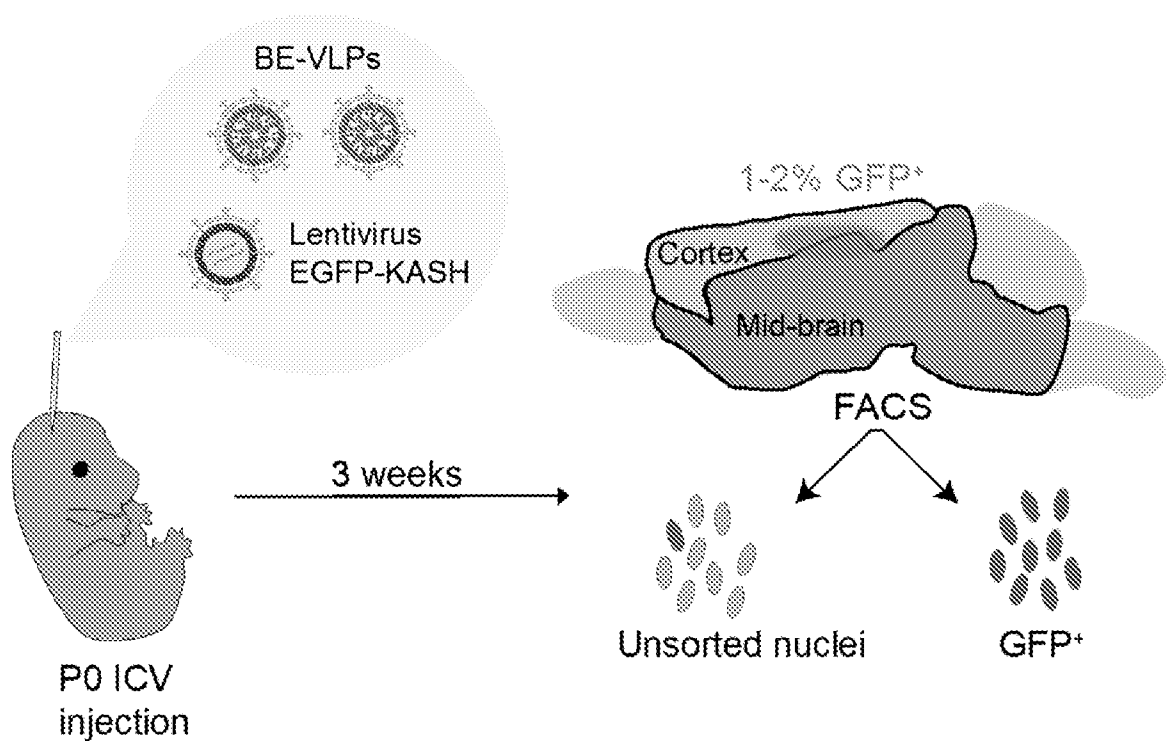


FIG. 5A

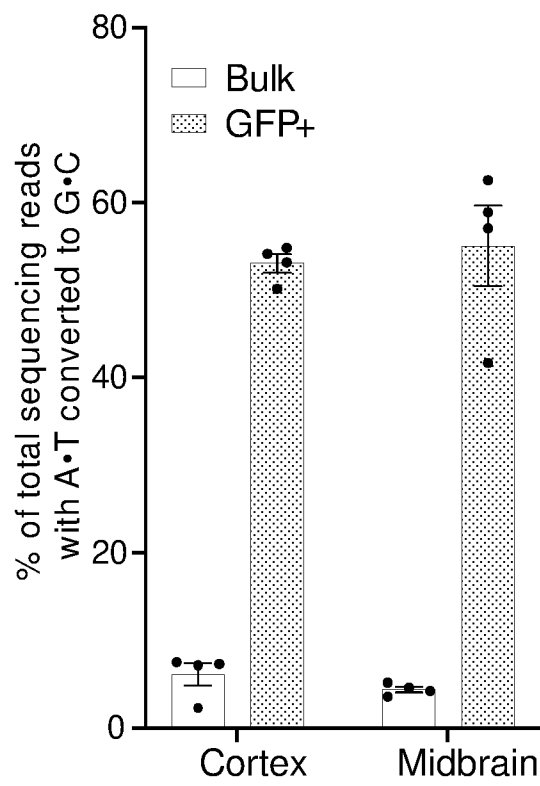
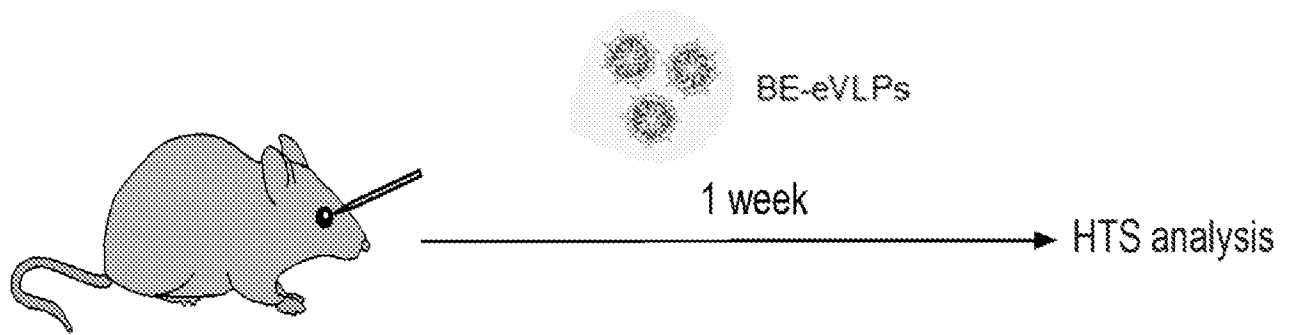
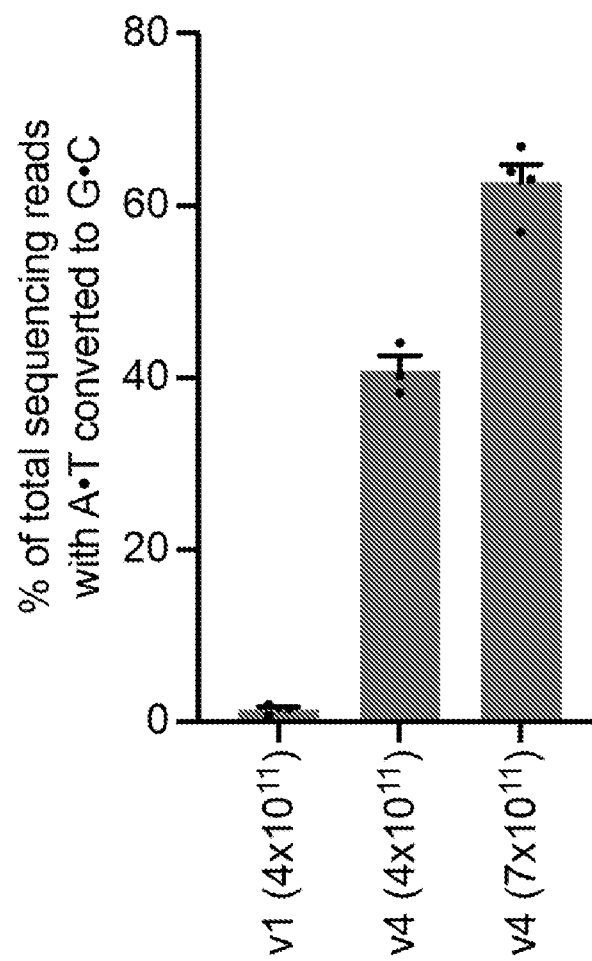


FIG. 5B

**FIG. 6A****FIG. 6B**

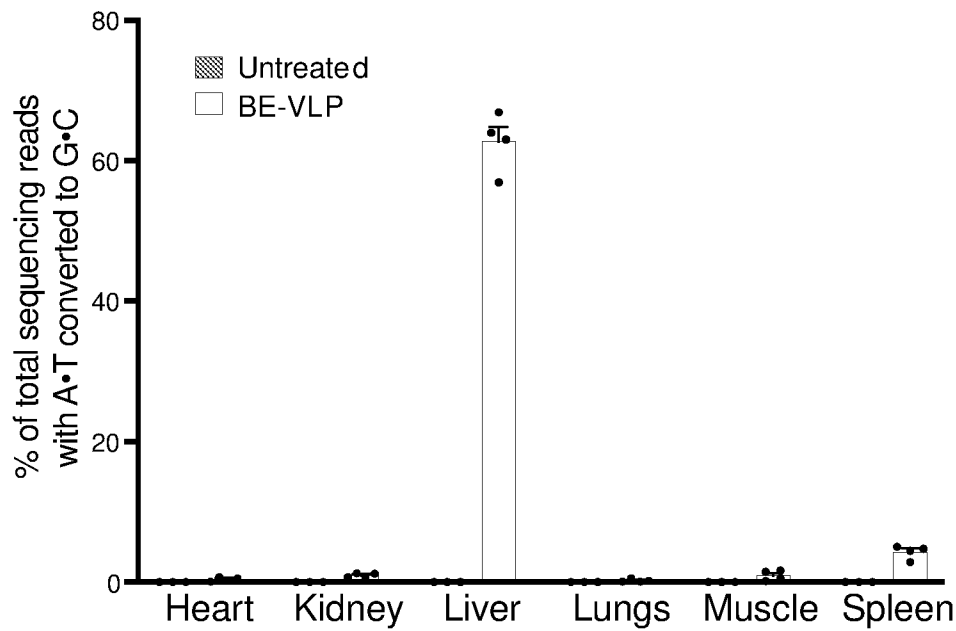


FIG. 6C

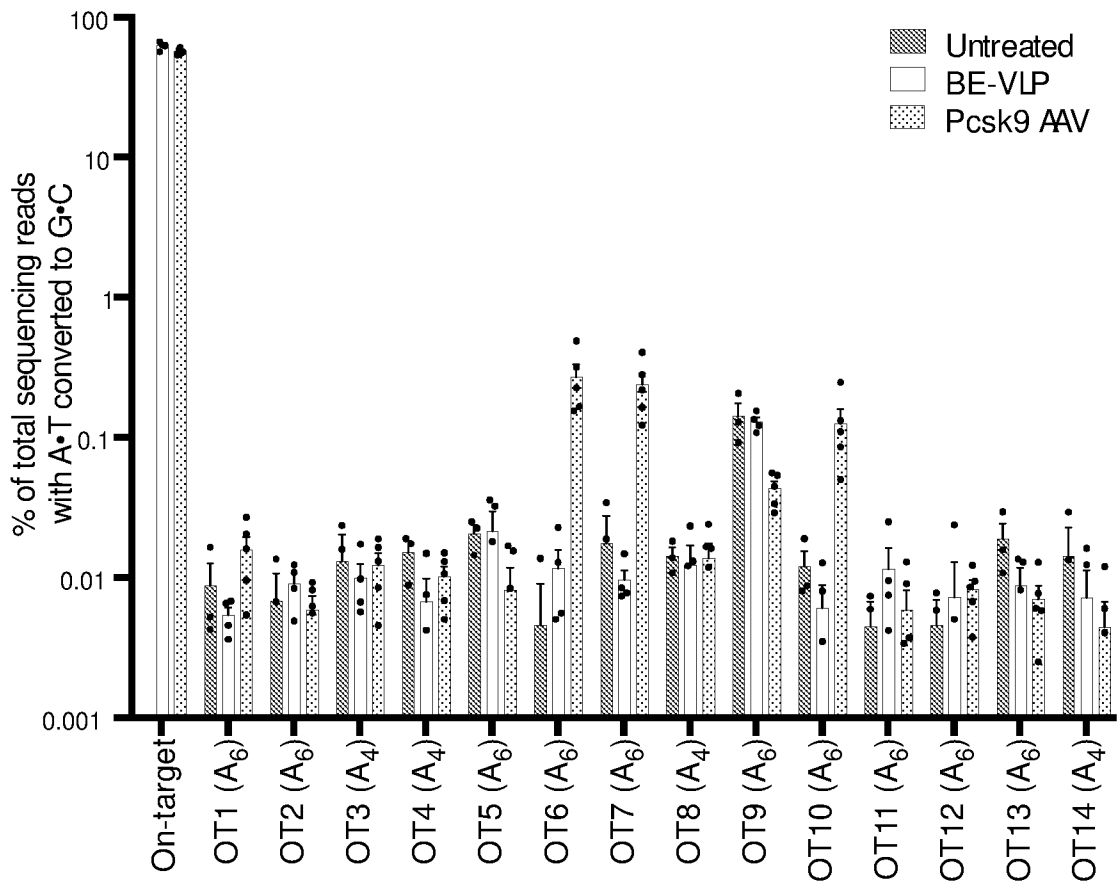


FIG. 6D

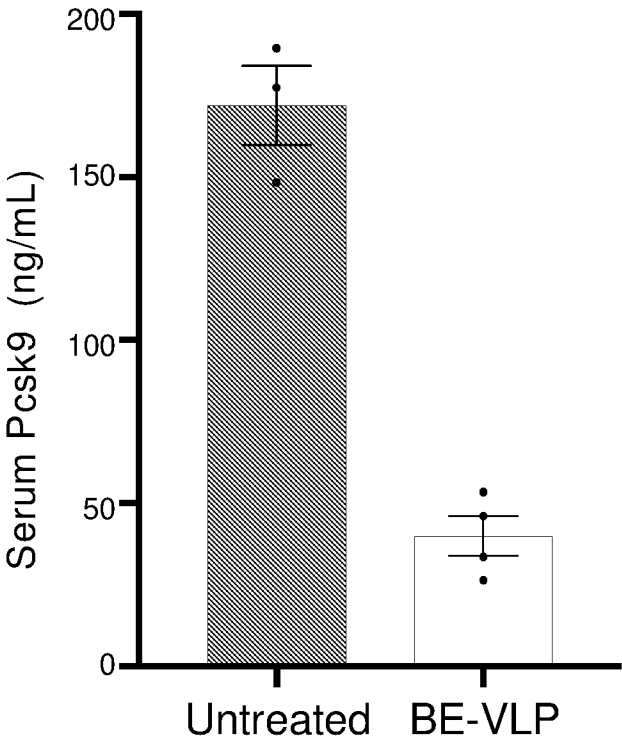


FIG. 6E

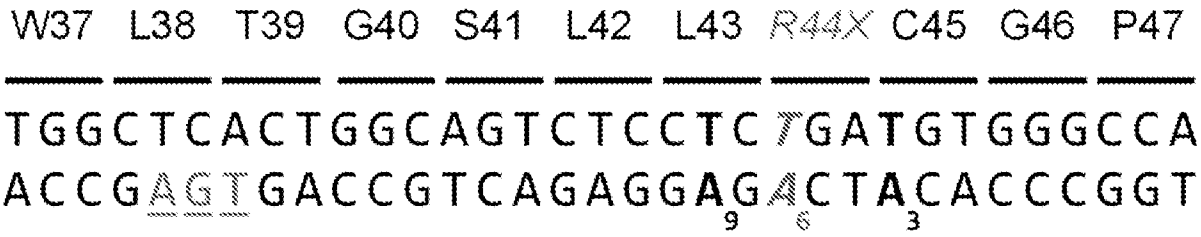


FIG. 7A

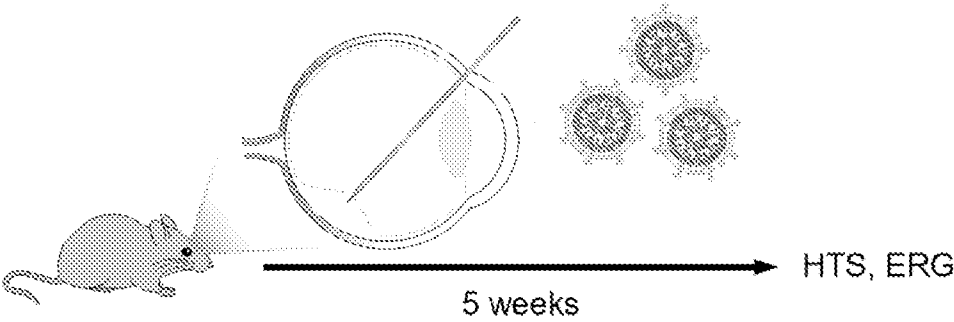


FIG. 7B

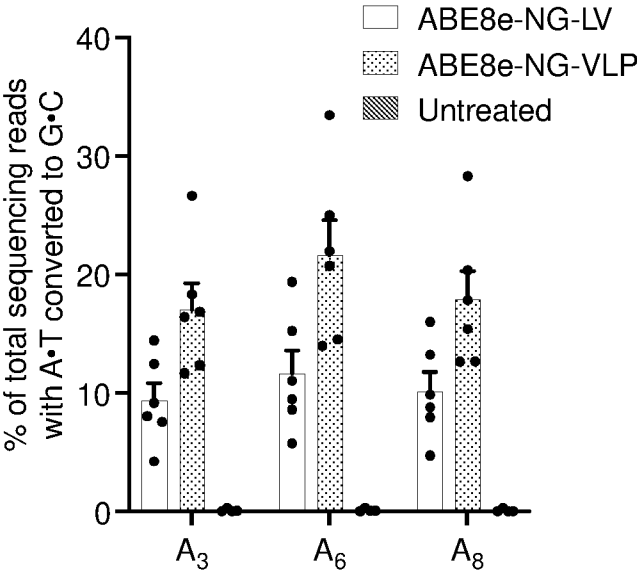


FIG. 7C

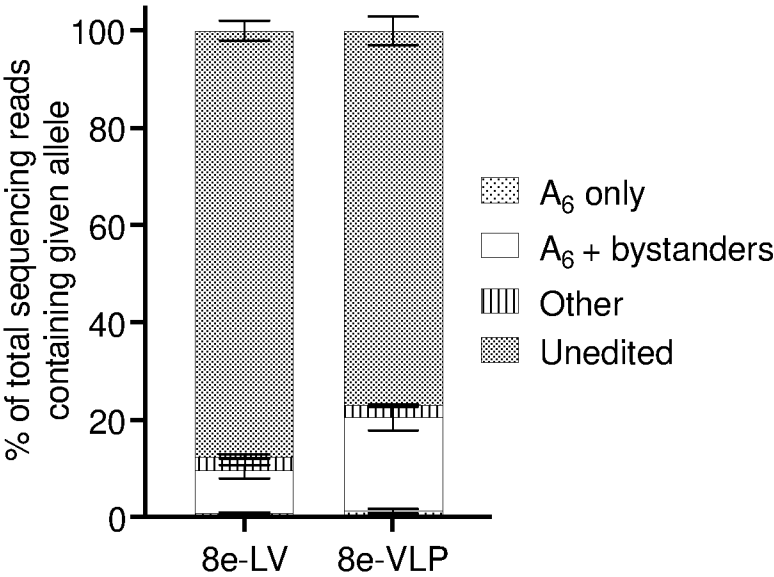


FIG. 7D

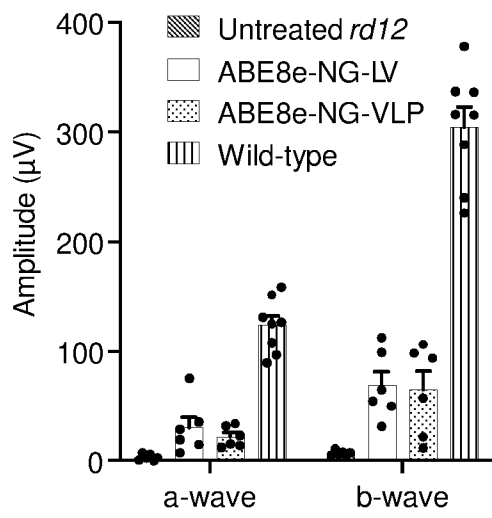


FIG. 7E

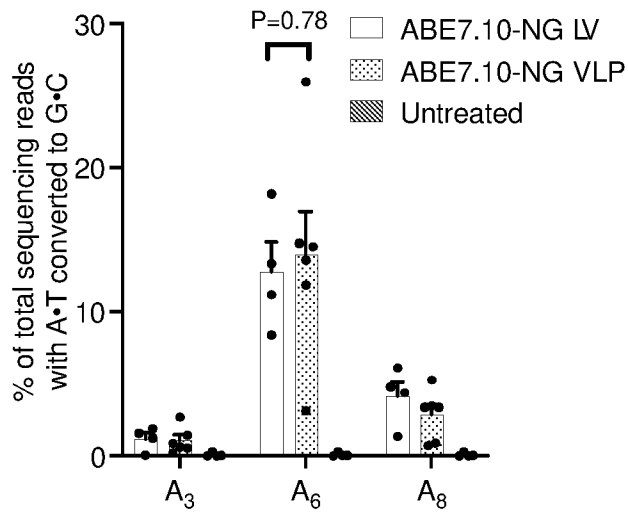


FIG. 7F

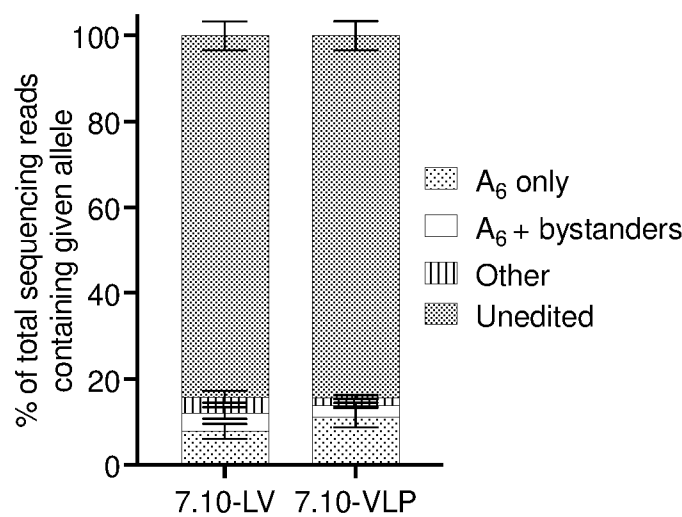


FIG. 7G

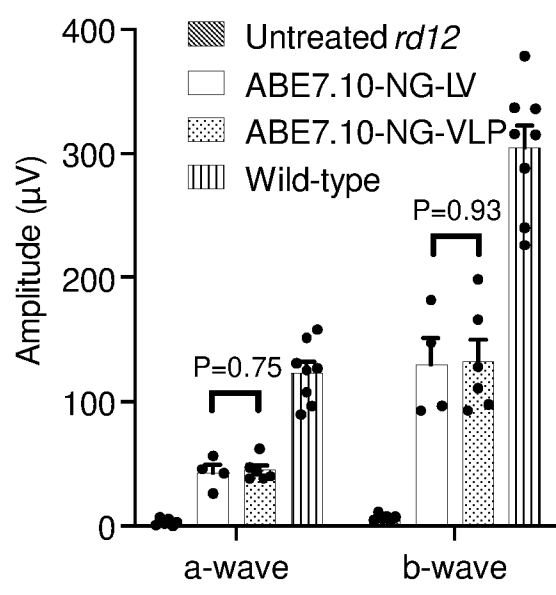


FIG. 7H

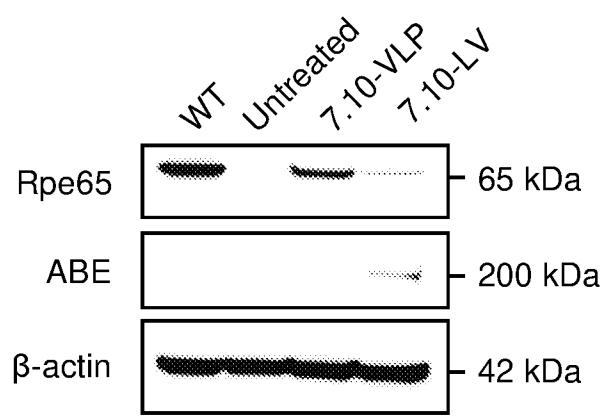


FIG. 7I

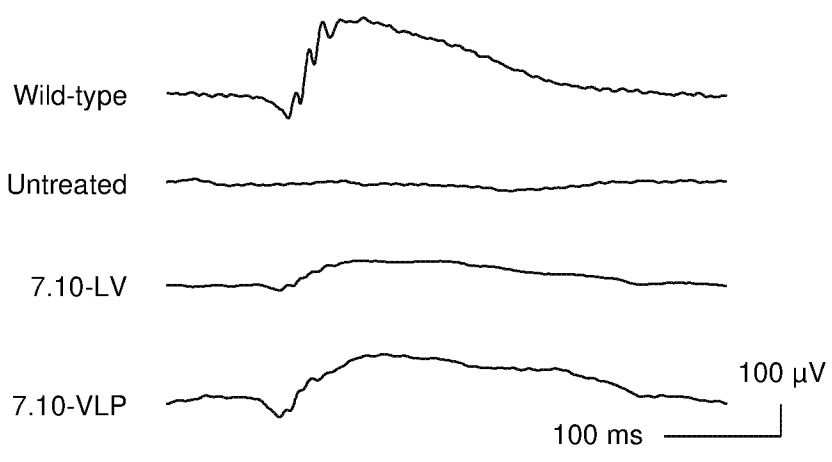
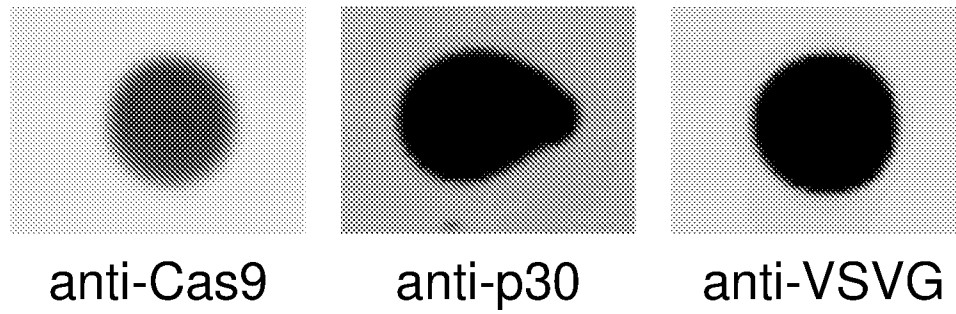
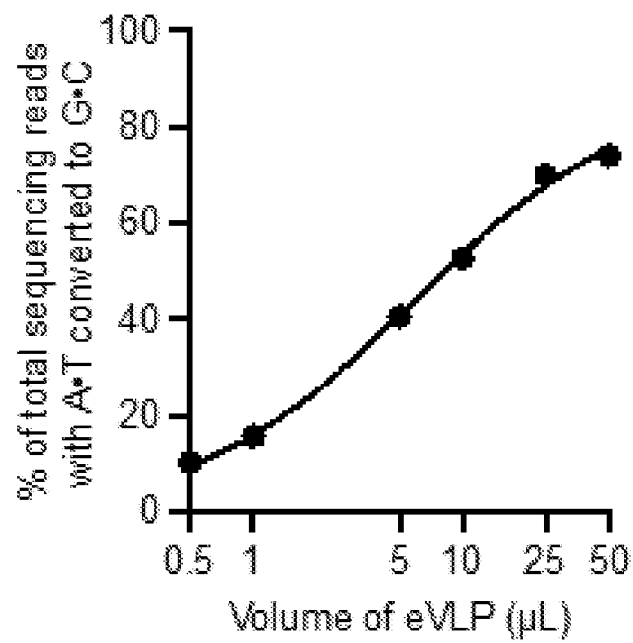


FIG. 7J

**FIG. 8A****FIG. 8B**

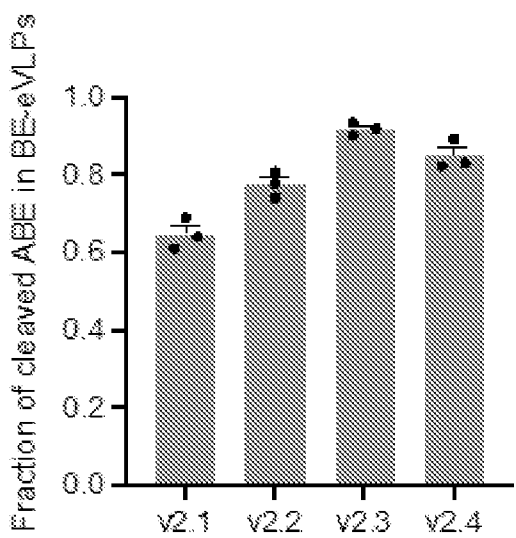
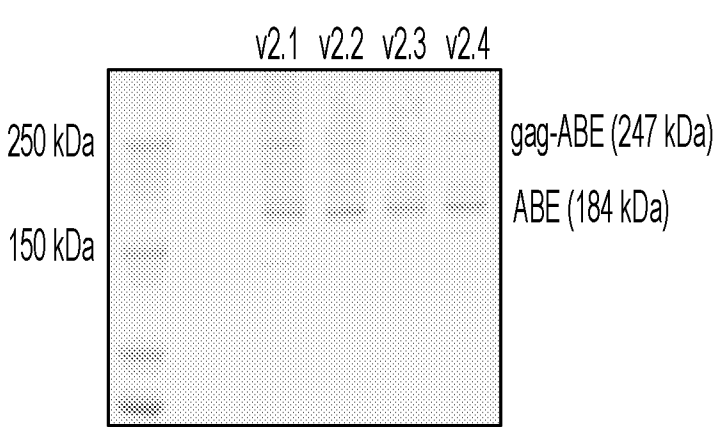
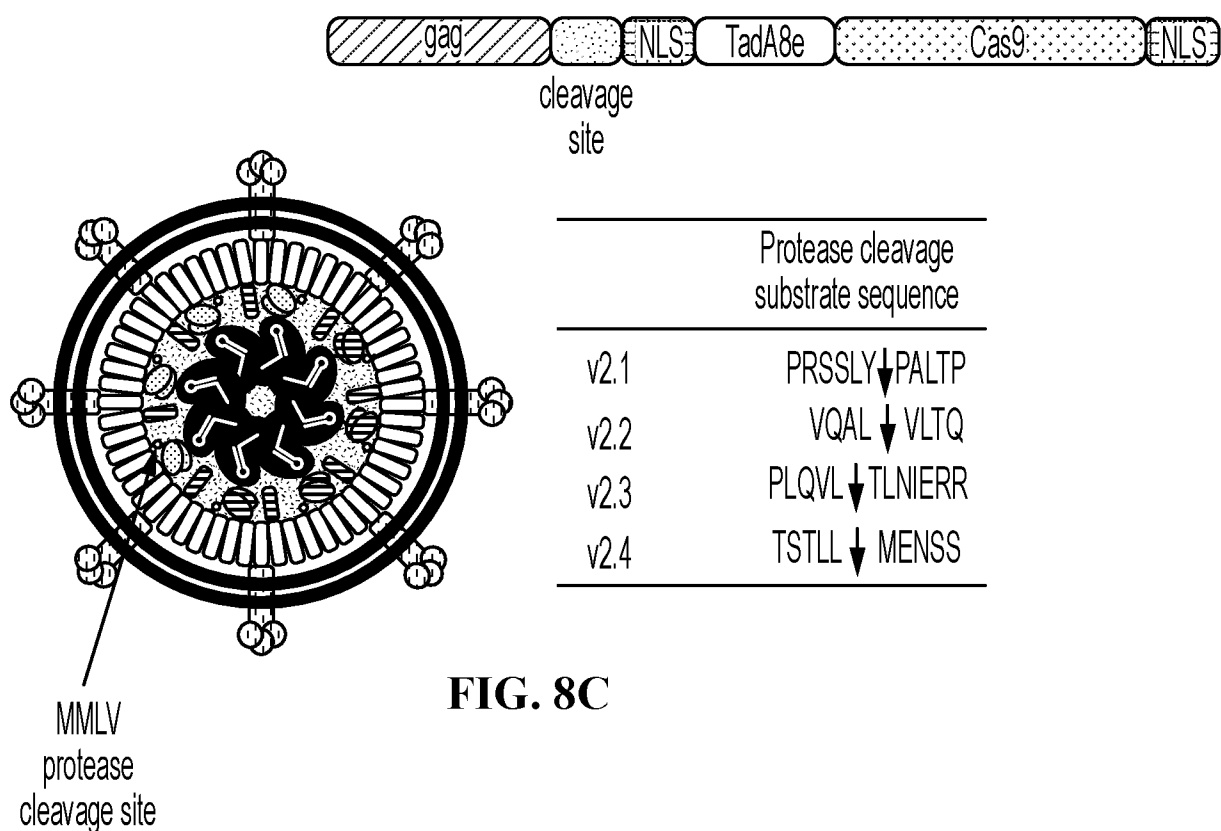


FIG. 8D

FIG. 8E

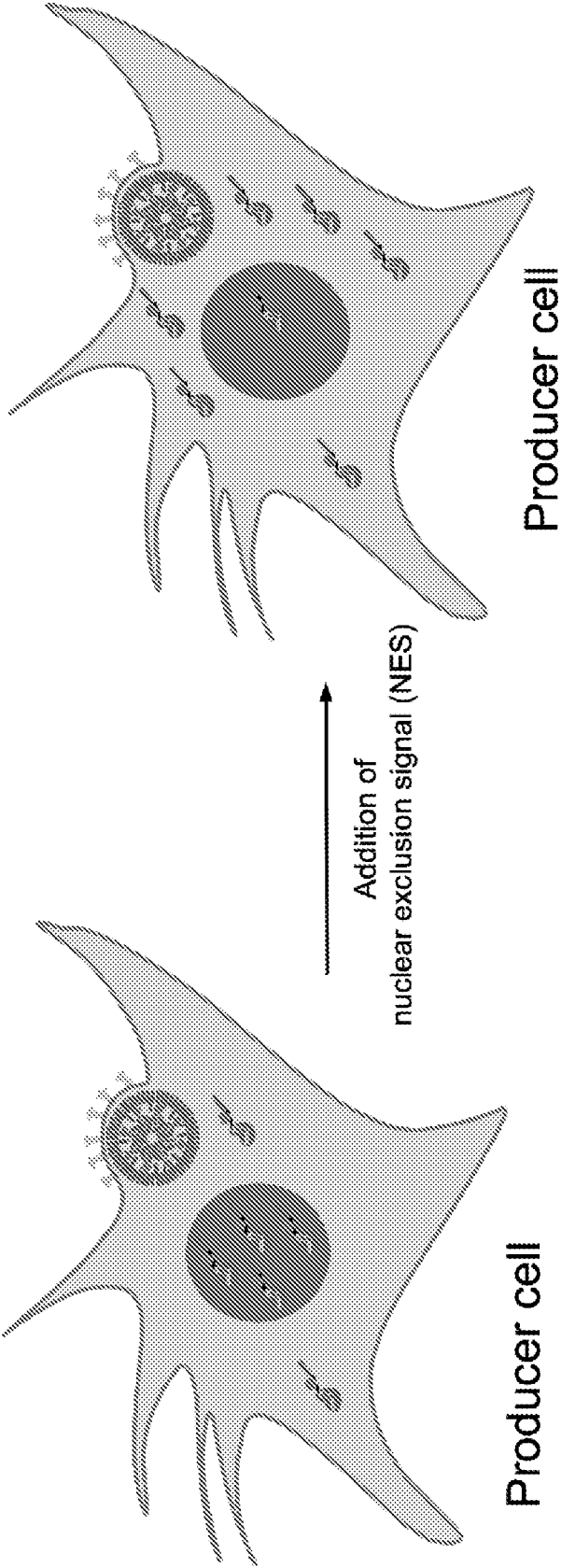


FIG. 9A

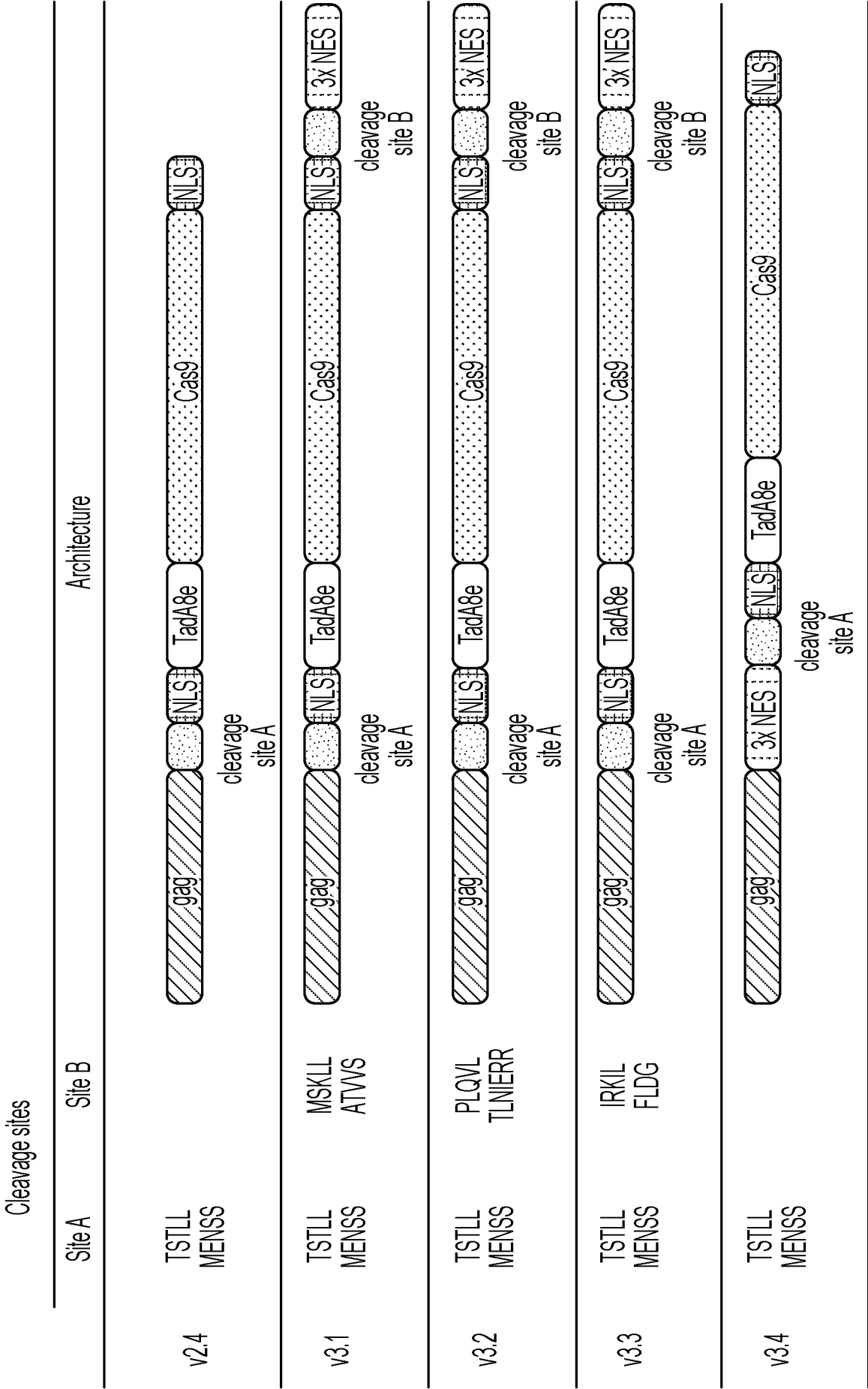


FIG. 9B

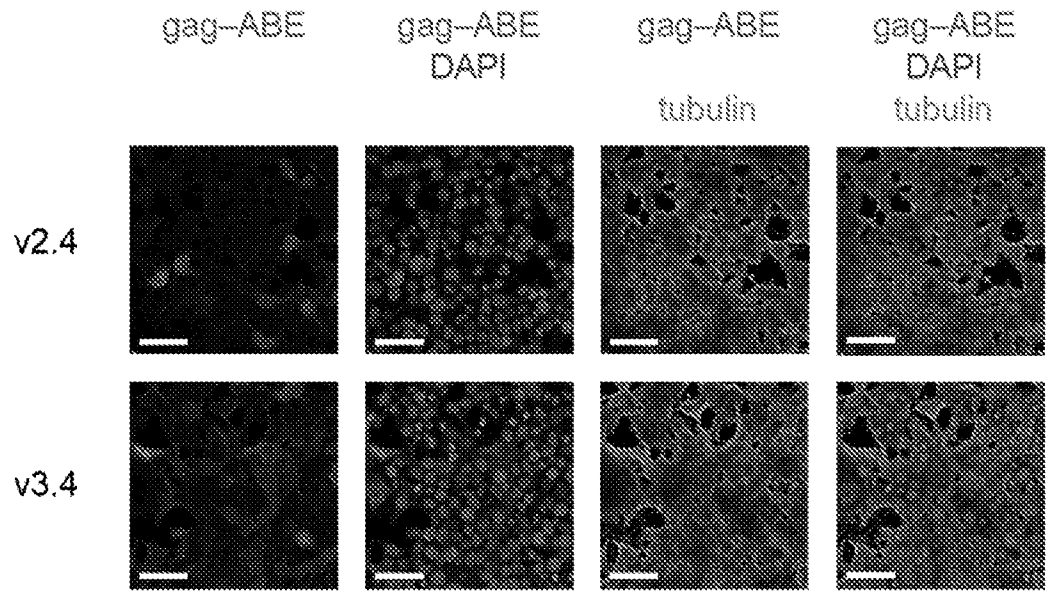


FIG. 9C

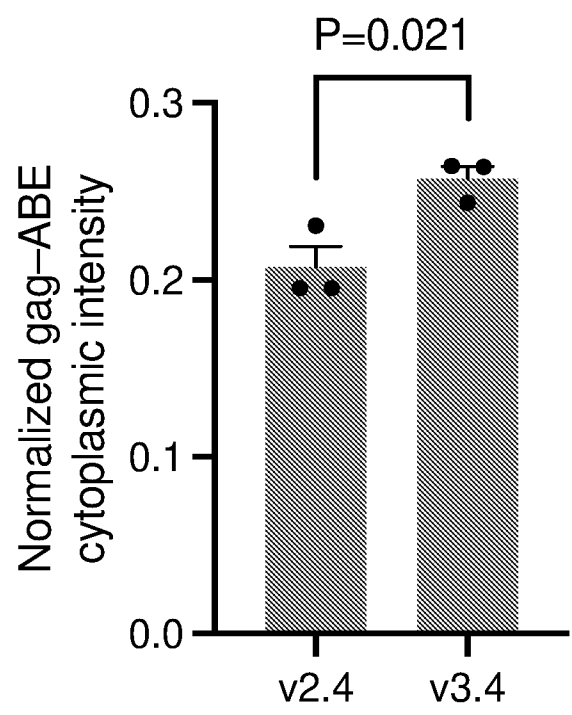
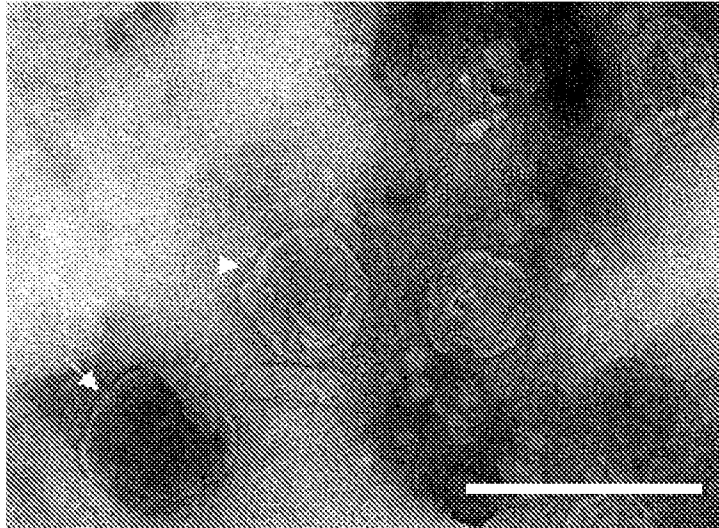
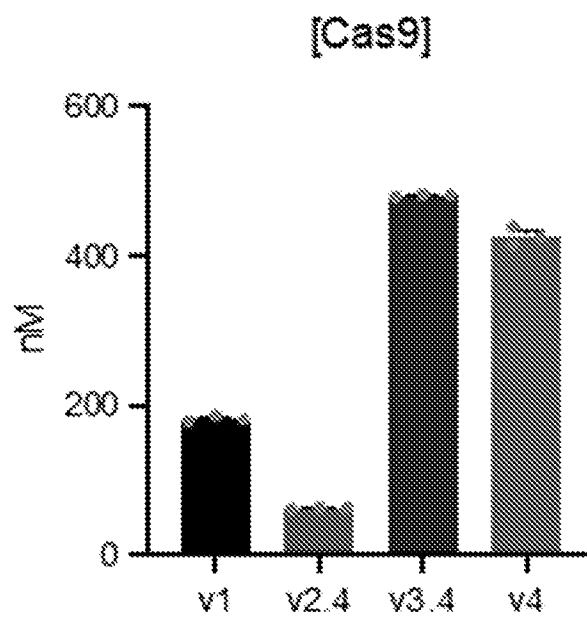


FIG. 9D

**FIG. 10A****FIG. 10B**

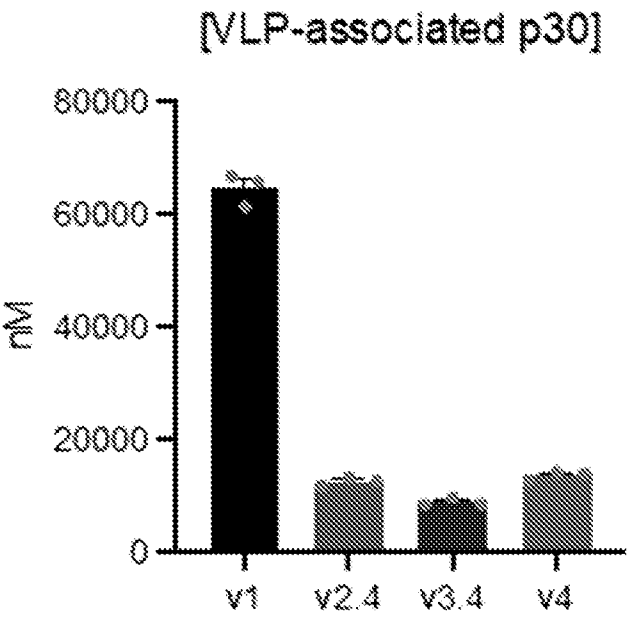


FIG. 10C

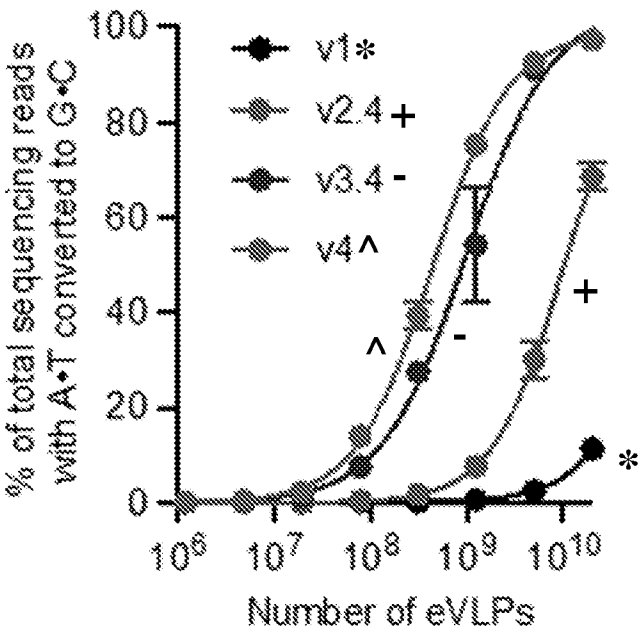


FIG. 10D

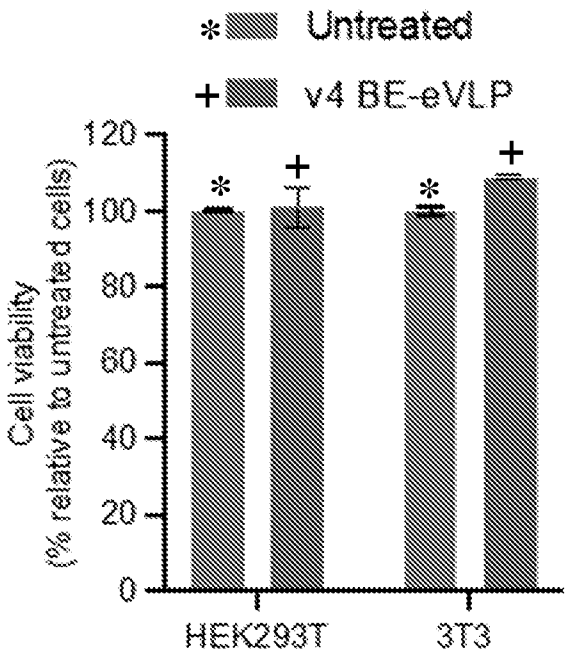


FIG. 10E

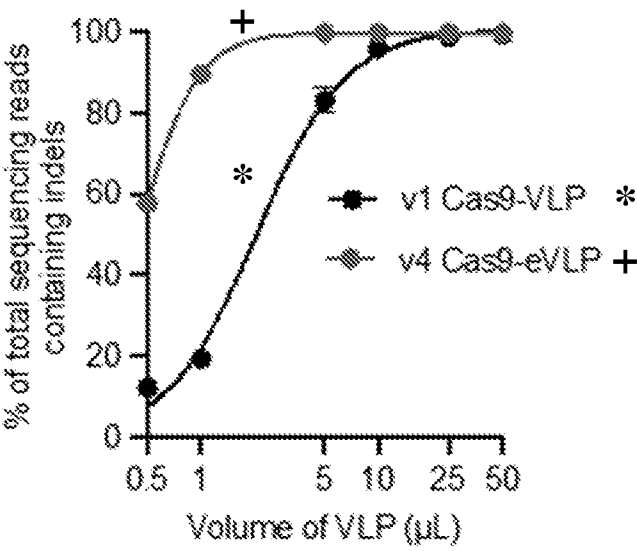


FIG. 10F

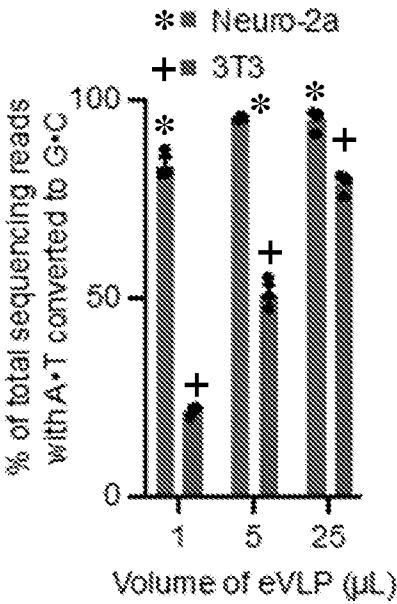
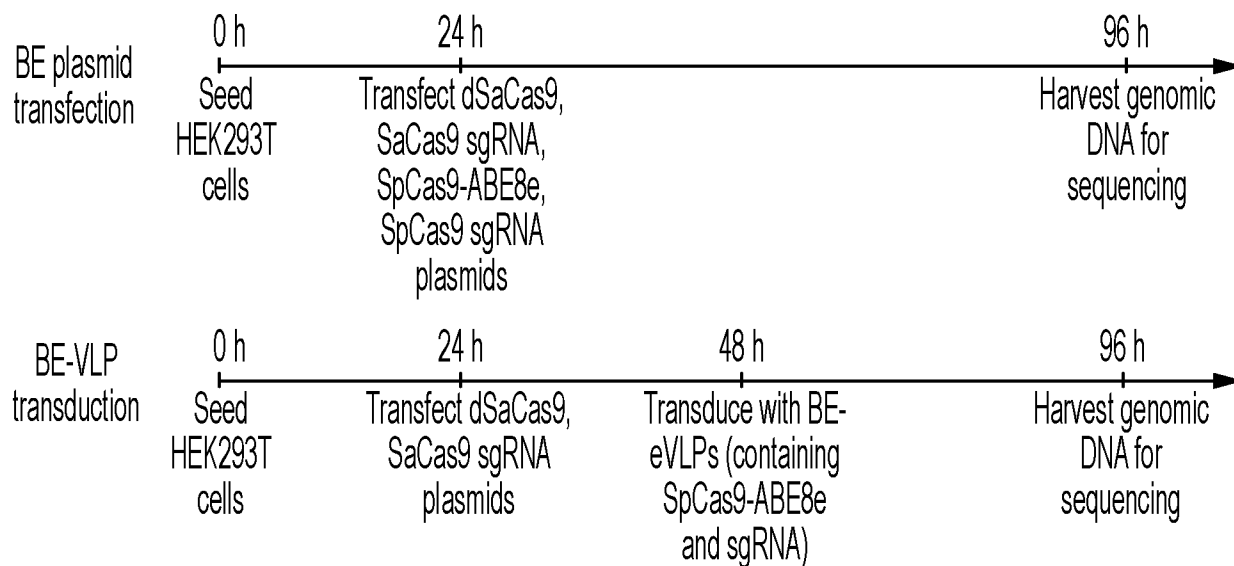
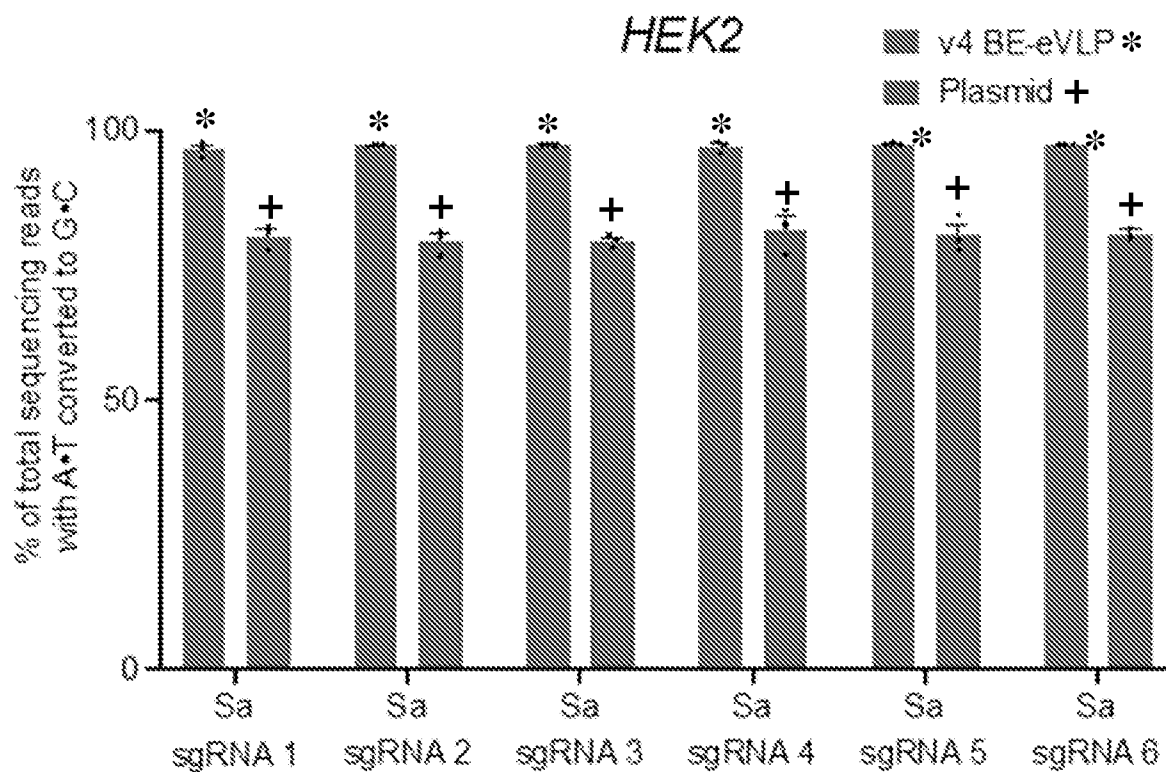


FIG. 10G

**FIG. 11A****FIG. 11B**

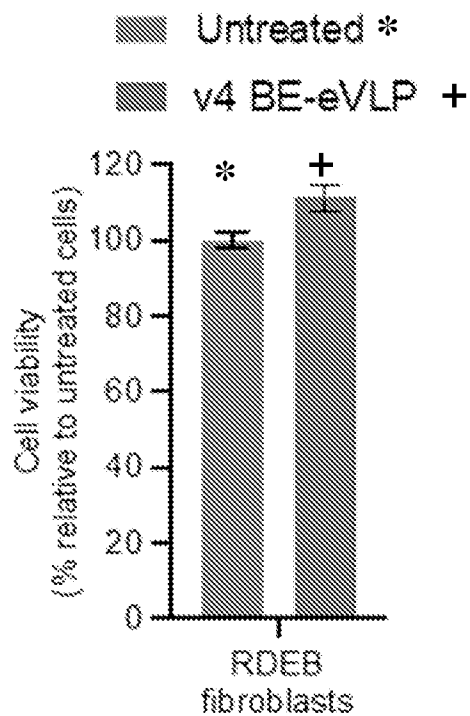


FIG. 11C

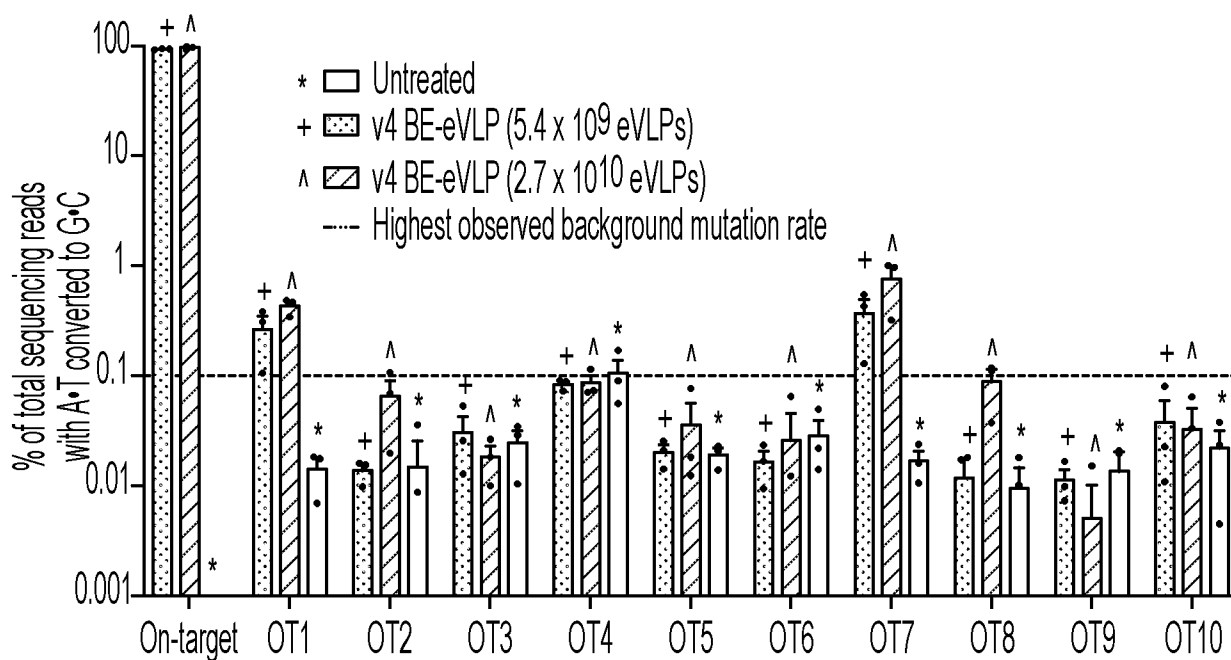


FIG. 11D

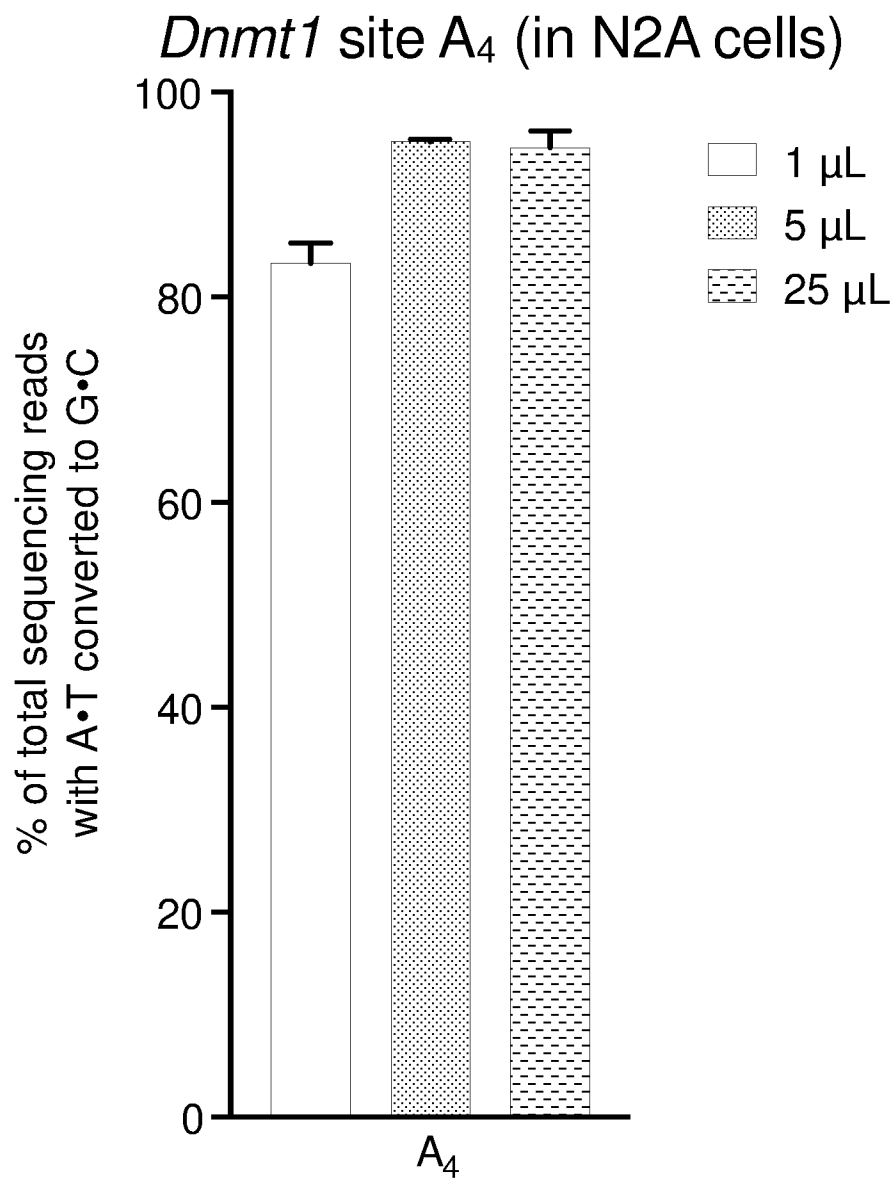
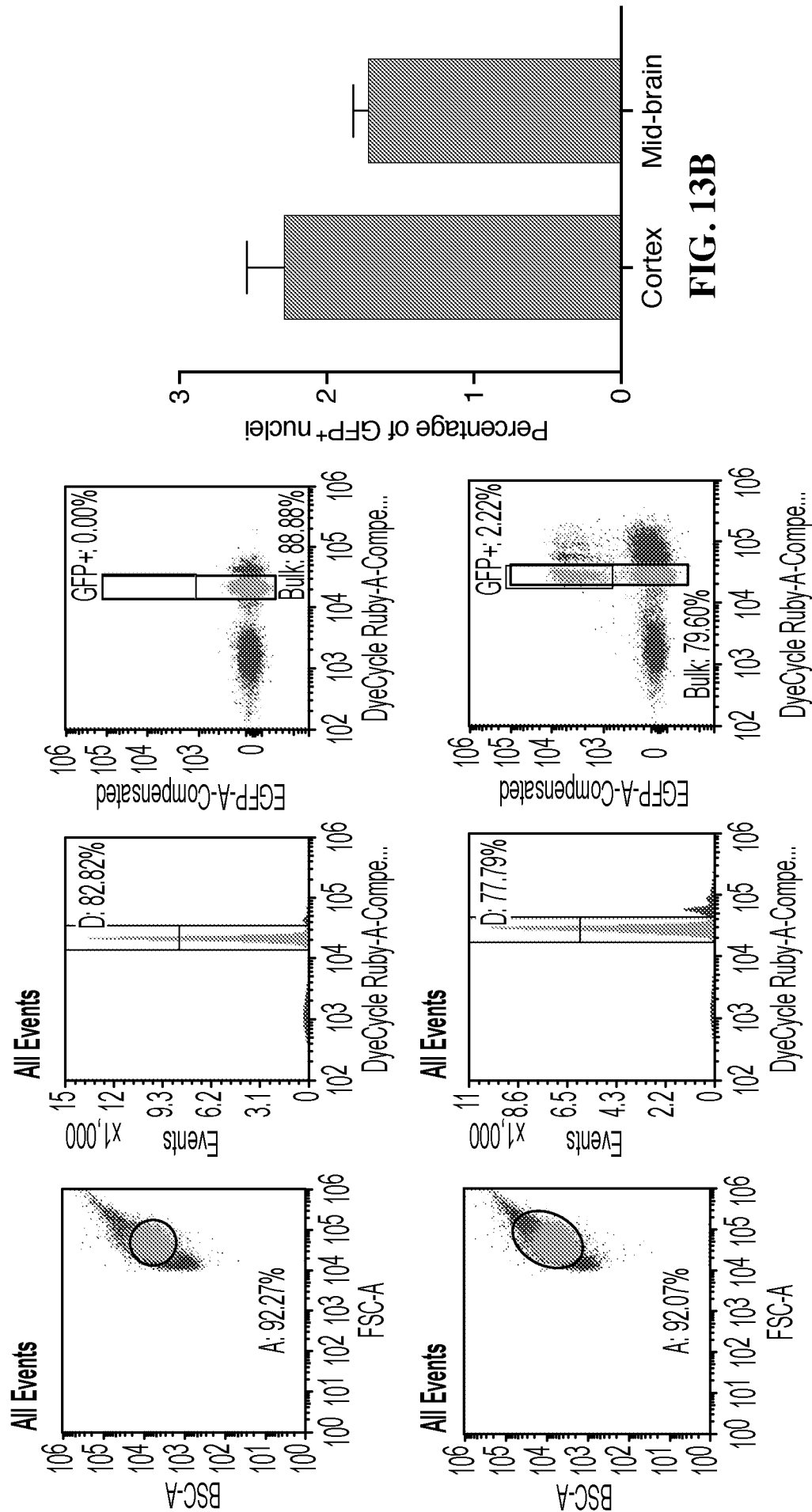
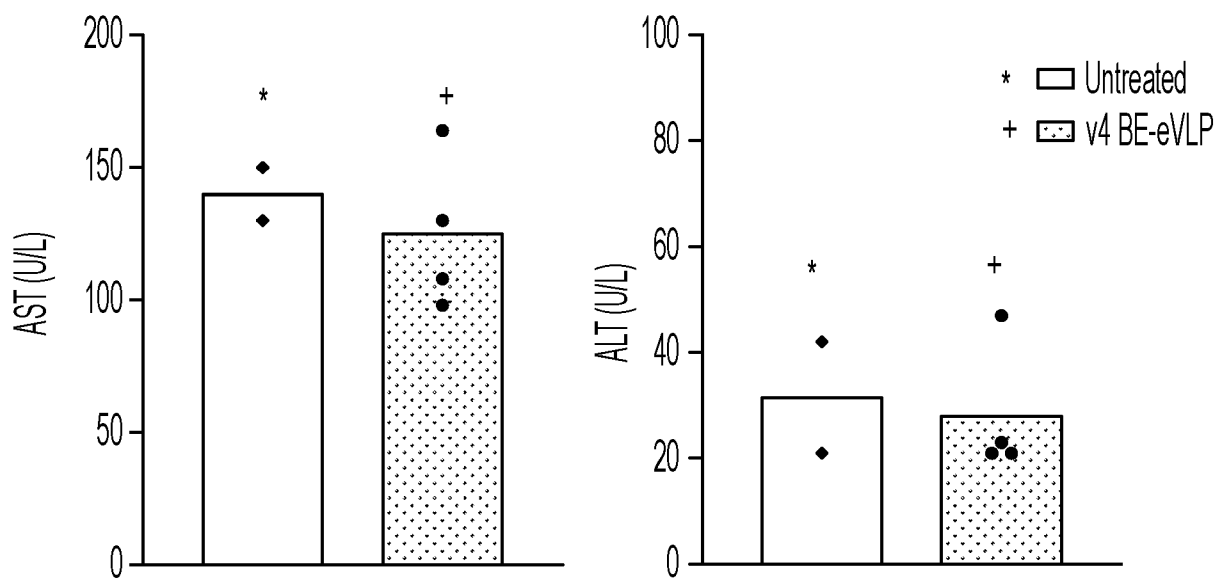
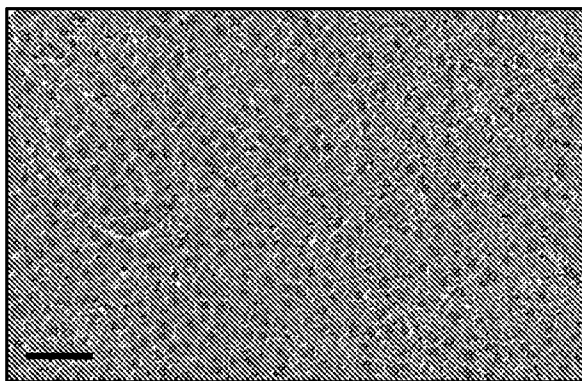
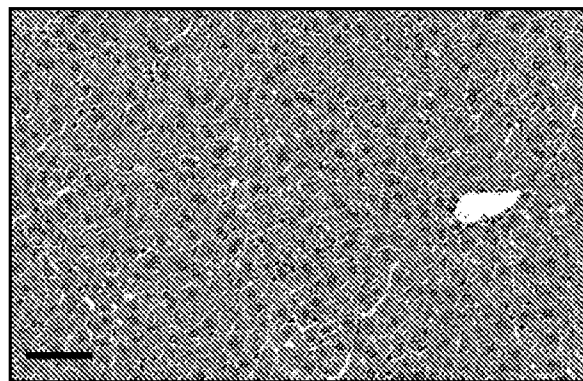


FIG. 12



**FIG. 14A****FIG. 14B****FIG. 14C**

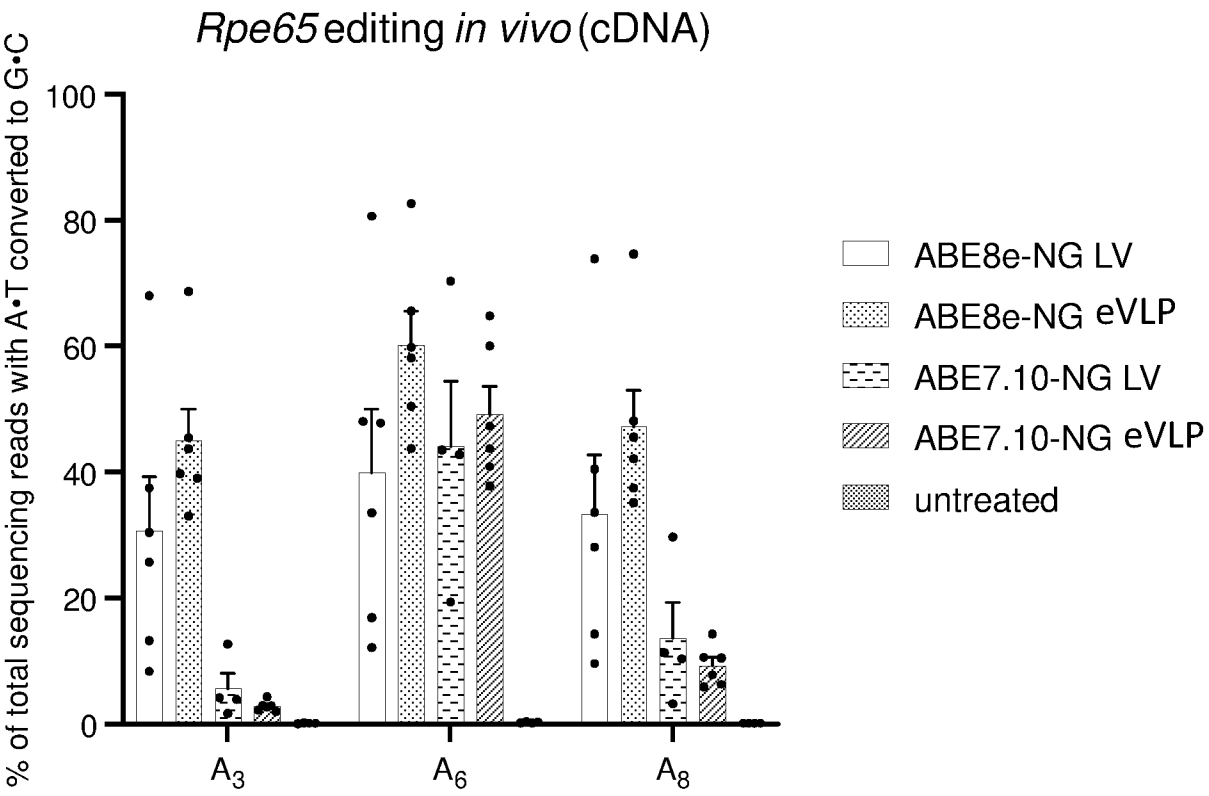


FIG. 15A

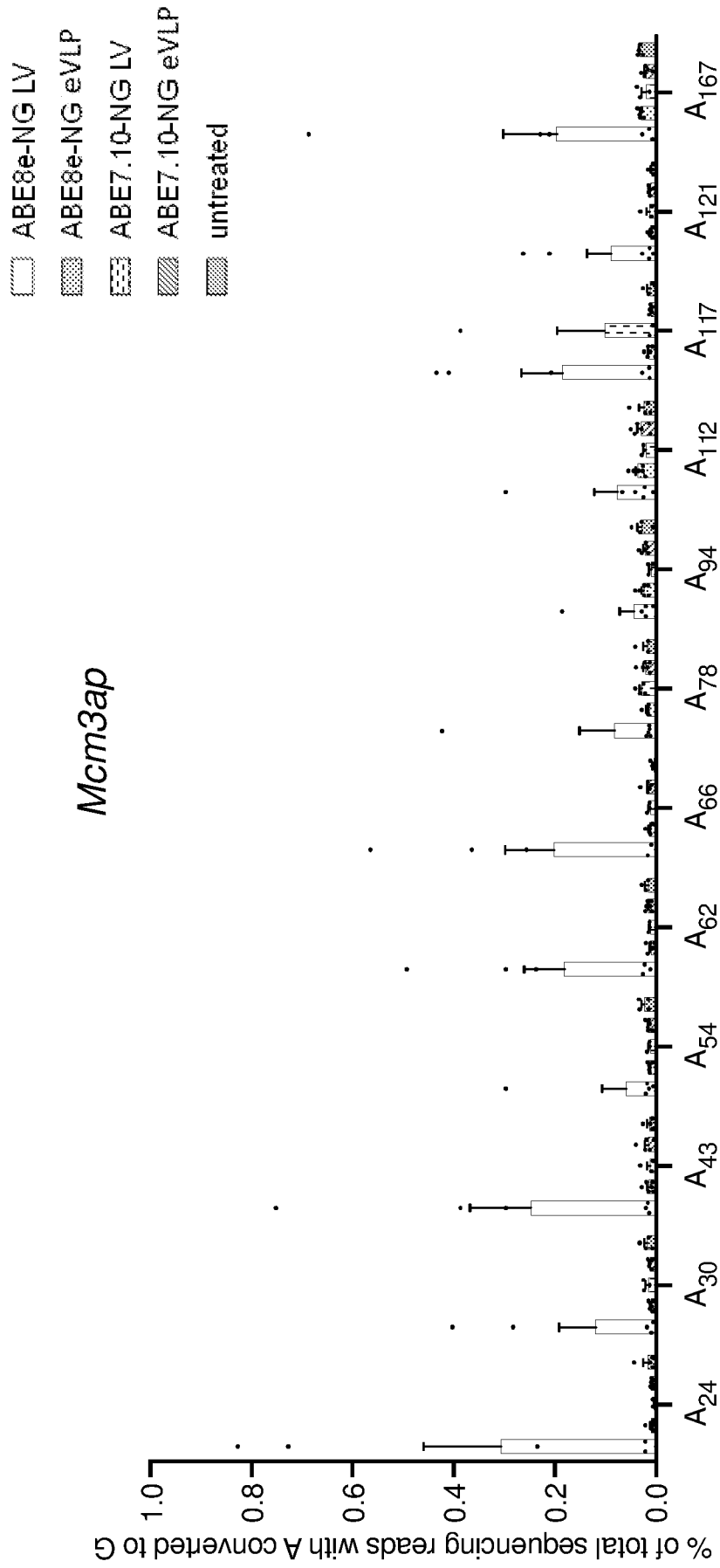


FIG. 15B

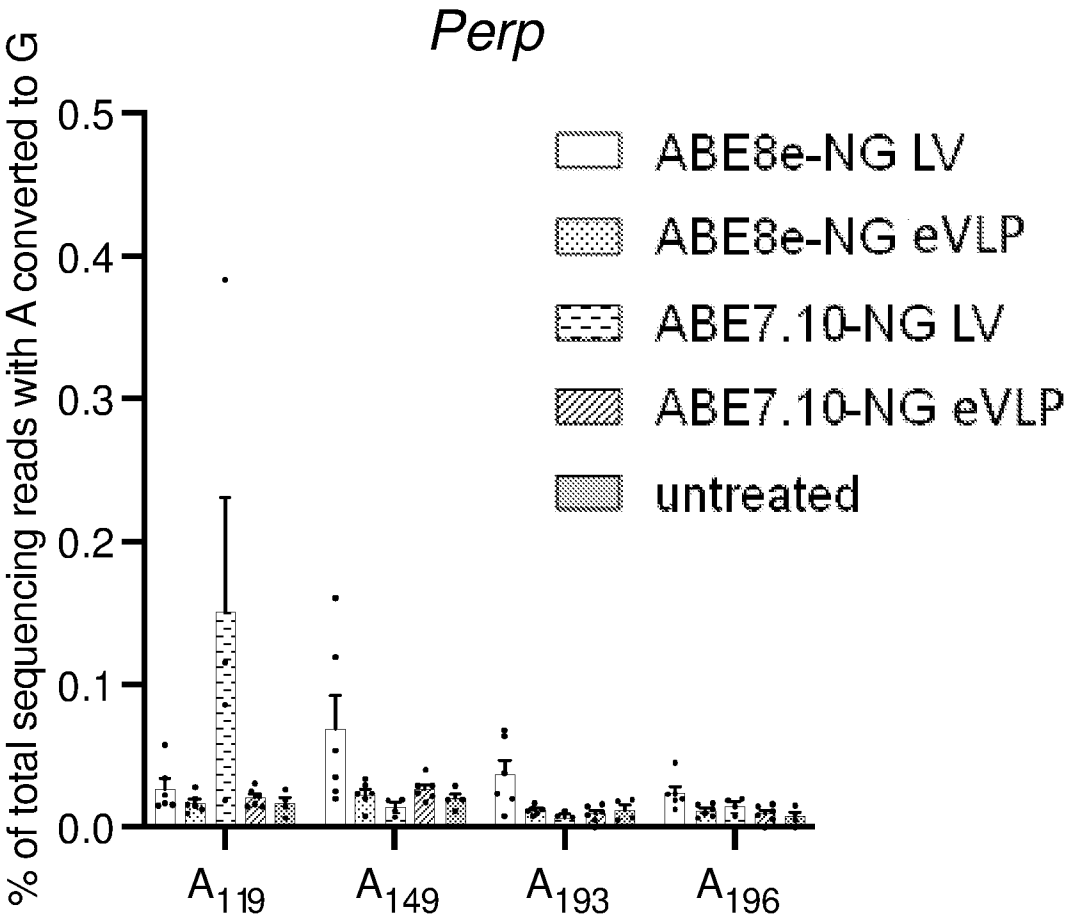


FIG. 15C

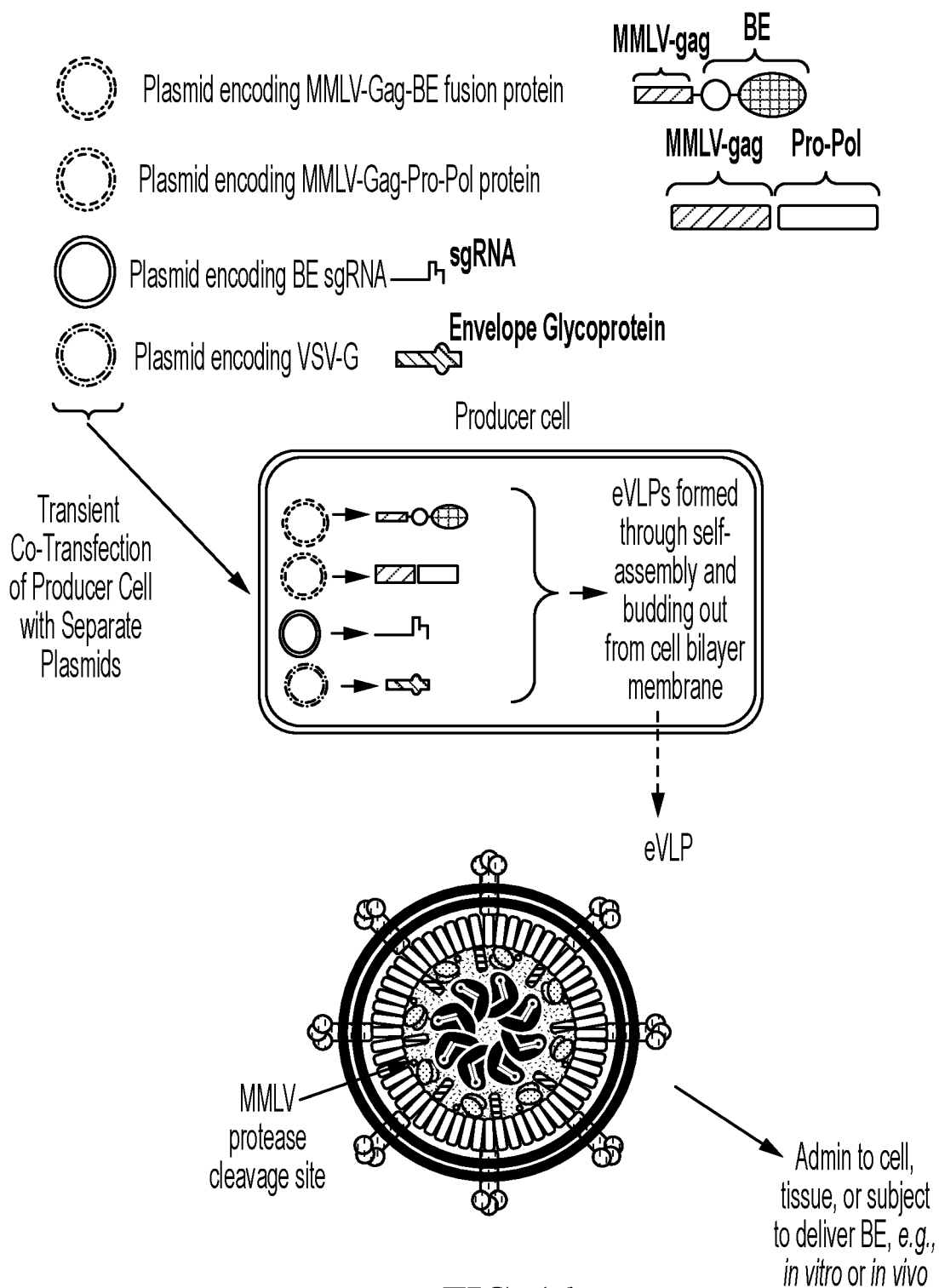


FIG. 16

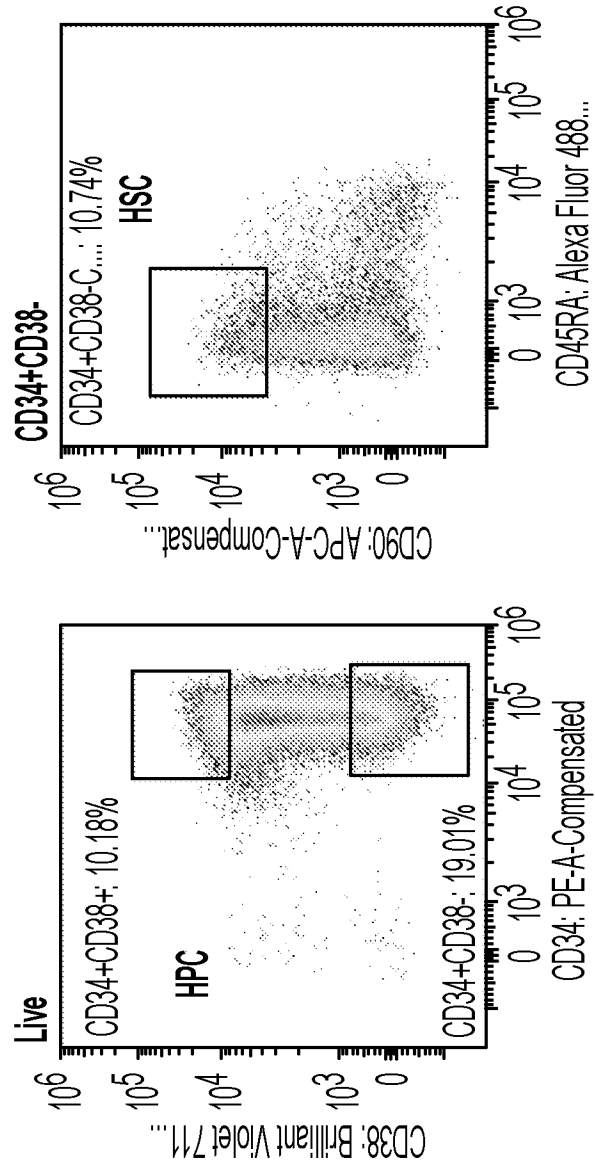


FIG. 17A

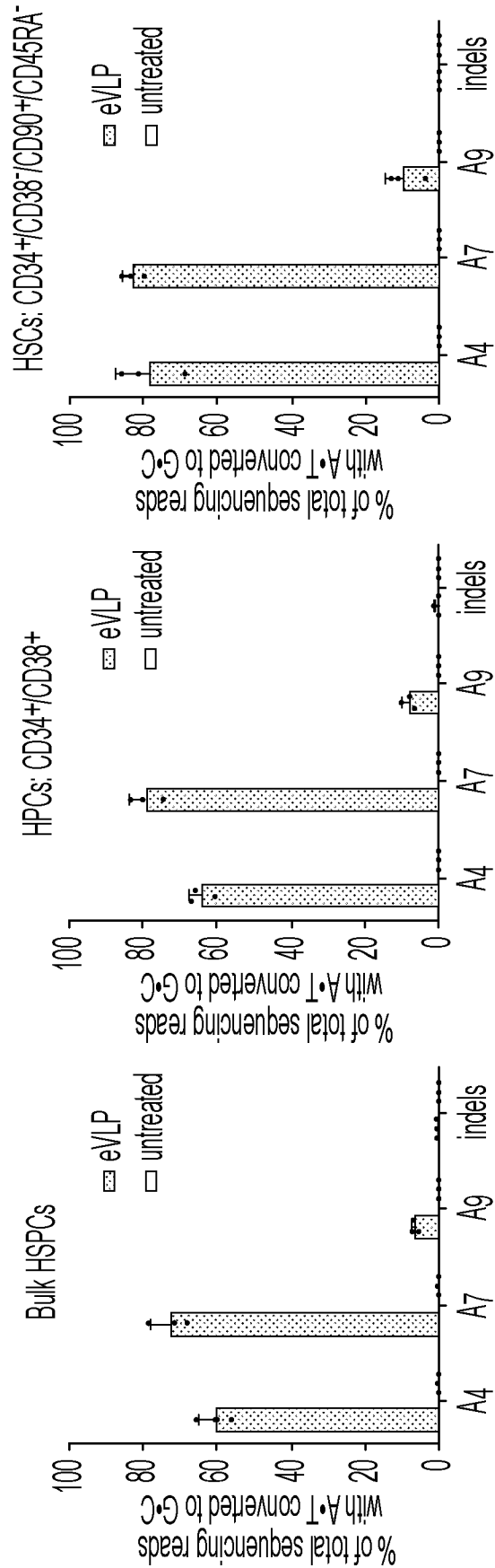
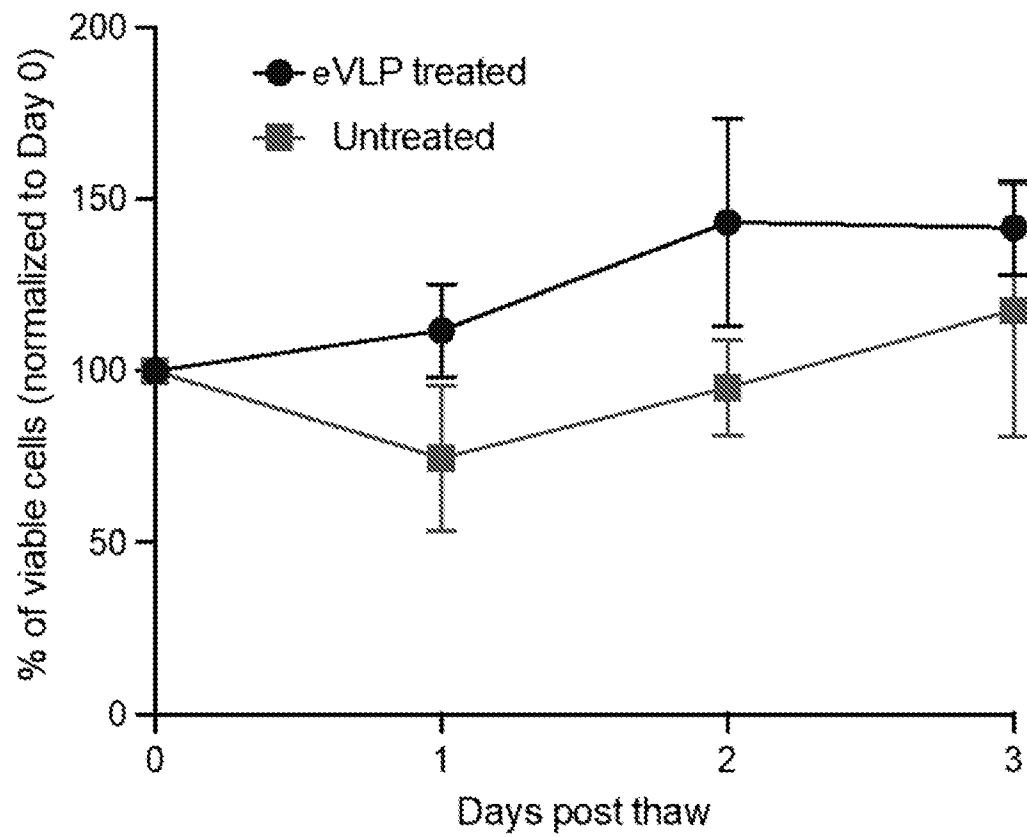
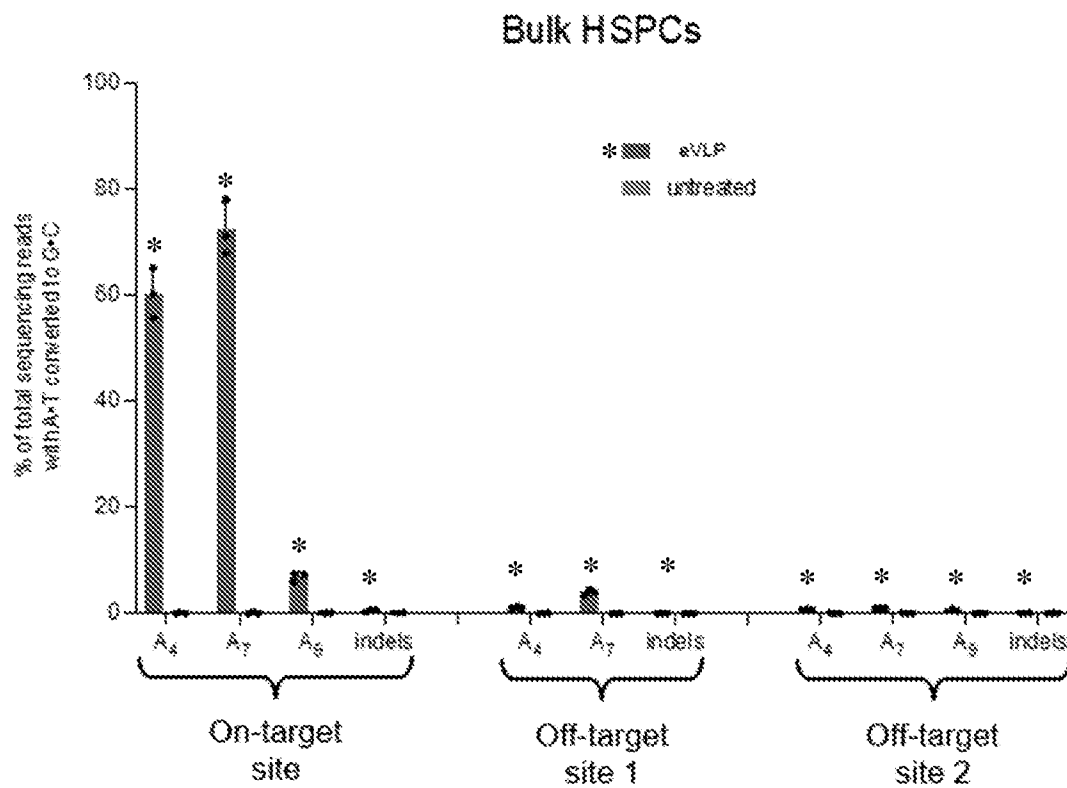
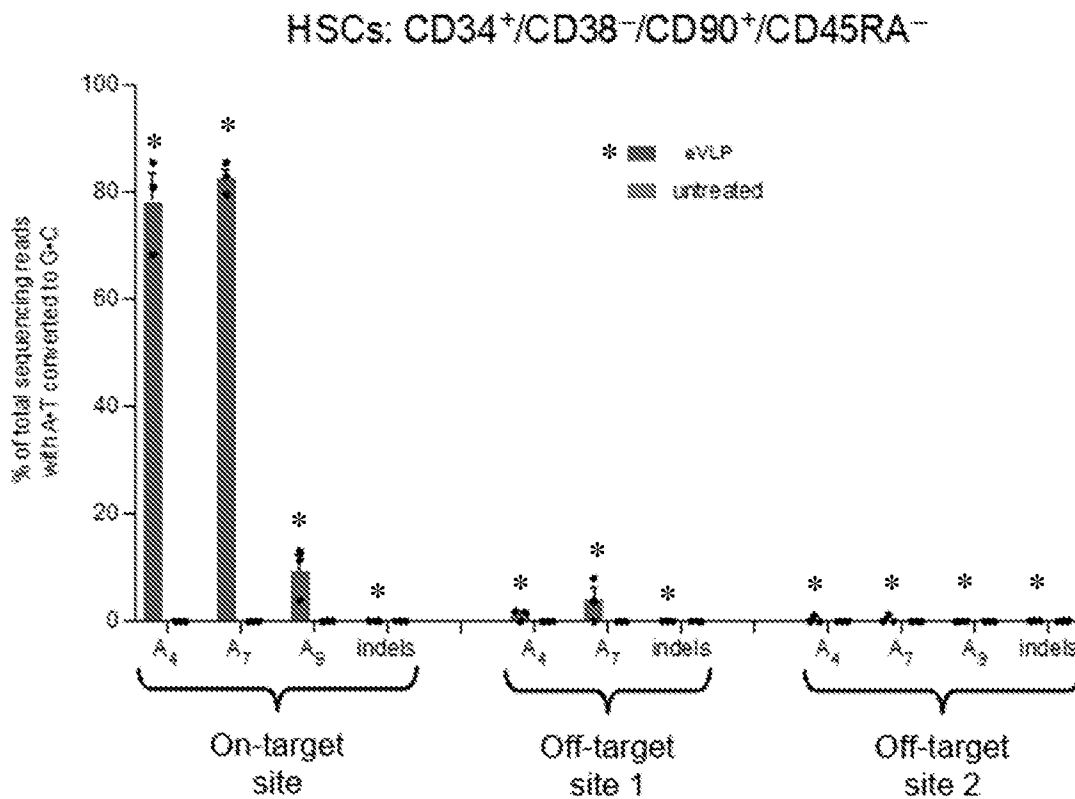


FIG. 17B

**FIG. 18**

**FIG. 19A****FIG. 19B**