



(51) International Patent Classification:

A61K 45/06 (2006.01) C12N 15/00 (2006.01)

A61P 31/12 (2006.01) C12N 15/86 (2006.01)

C07K 14/16 (2006.01) C12Q 1/68 (2018.01)

(21) International Application Number:

PCT/US2018/026716

(22) International Filing Date:

09 April 2018 (09.04.2018)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/486,237 17 April 2017 (17.04.2017) US

62/524,218 23 June 2017 (23.06.2017) US

(71) Applicants: TEMPLE UNIVERSITY- OF THE COMMONWEALTH SYSTEM OF HIGHER EDUCATION

[US/US]; Broad Street And Montgomery Avenue, Philadelphia, PA 19122 (US). BOARD OF REGENTS OF THE UNIVERSITY OF NEBRASKA [US/US]; Verner Hall, 3835 Holdrege Street, Lincoln, Nebraska 68583-0745 (US).

(72) Inventors: KHALILI, Kamel; 190 Presidential Boulevard, #718, Bala Cynwyd, PA 19004 (US). GENDELMAN, Howard, E.; 125 South 127th Plaza, Omaha, NE 68154 (US). EDGAWA, Benson; 310 N. 40th Street, Apt. 9, Omaha, Nebraska 68131 (US).

(74) Agent: ZACHARIADES, Nicholas, A. et al.; Mintz Levin Cohn Ferris Glovsky Popeo, P.C., One Financial Center, Boston, MA 02111 (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, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH,

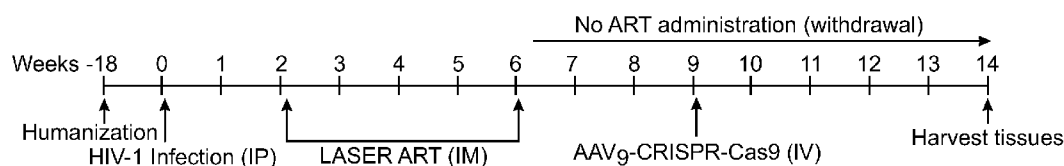
GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA,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

Published:

— with international search report (Art. 21(3))

(54) Title: AN HIV-1 ERADICATION STRATEGY EMPLOYING NANOFORMULATED ANTI-RETROVIRAL DRUGS AND GENE EDITING AGENTS

FIG. 1A



(57) Abstract: Methods of eliminating a retrovirus from a subject utilize nanoformulated anti-retroviral compounds and gene editing agents. Compositions comprise at least one anti-retroviral compounds, at least one gene-editing agent, or combinations thereof.



AN HIV-1 ERADICATION STRATEGY EMPLOYING NANOFORMULATED ANTI-RETROVIRAL DRUGS AND GENE EDITING AGENTS

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This Application claims the benefit of U.S. Provisional Application 62/486,237 filed on April 17, 2017, the entire contents of which are incorporated herein by reference in its entirety.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH

[0002] This invention was made with government support under Grant No. P30MH092177 (Khalili) awarded by the National Institutes of Health. The government has certain rights in this invention.

FIELD OF THE INVENTION

[0003] A combination therapy for the elimination and eradication of a retrovirus, for example, HIV, from an infected subject. In particular, the therapeutic approach utilizes long-acting slow effective release antiretroviral therapy (called LASER ART) and a gene editing agent.

BACKGROUND

[0004] The elimination of the human immunodeficiency virus (HIV) from its viral reservoirs is a requirement for disease cure. Cure is defined as undetectable viremia measured in time periods of years in the absence of antiretroviral therapy (ART).

SUMMARY

[0005] Embodiments of the invention are directed to a combination therapy comprising antiretroviral therapy (ART) along with gene editing.

[0006] In certain embodiments, a method of eradicating a retrovirus in a subject, comprises administering to a patient a composition comprising a therapeutically effective amount of at least one antiretroviral agent and/or a composition comprising a therapeutically effective amount of at least one gene editing agent. In certain embodiments, the antiretroviral

or anti-viral agent is formulated as a long-acting slow effective release (LASER) antiretroviral agent. In certain embodiments, the at least one antiretroviral or anti-viral agent is nanoformulated. In certain embodiments, the at least one antiretroviral or anti-viral agent comprises: myristoylated dolutegravir, lamivudine, abacavir, rilpivirine or combinations thereof.

[0007] In certain embodiments, at least one antiretroviral agent is administered to the subject prior to administering the at least one gene editing agent. In certain embodiments, the at least one antiretroviral agent and at least one gene-editing agent are co-administered. In certain embodiments, the at least one antiretroviral agent and at least one gene-editing agent are administered sequentially.

[0008] In certain embodiments, the at least one gene editing agent comprises: an isolated nucleic acid sequence encoding a Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR)-associated endonuclease/Cas (CRISPR/Cas) and at least one guide RNA (gRNA), the gRNA being complementary to a target nucleic acid sequence in a retroviral genome.

[0009] In certain embodiments, the CRISPR/Cas fusion protein comprises catalytically deficient Cas protein (dCas), orthologs, homologs, mutants variants or fragments thereof.

[0010] In certain embodiments, the at least one gRNA includes at least a first gRNA that is complementary to a target sequence in the integrated retroviral DNA; and a second gRNA that is complementary to another target sequence in the integrated retroviral DNA, whereby the intervening sequences between the two gRNAs are removed.

[0011] In certain embodiments, the isolated nucleic acid is included in at least one expression vector. In certain embodiments, the expression vector comprises a lentiviral vector, an adenoviral vector, or an adeno-associated virus vector. In certain embodiments the vector is an adeno-associated vector, e.g. AAV₉.

[0012] In certain embodiments, the retrovirus is a human immunodeficiency virus (HIV).

[0013] In certain embodiments, the target sequences comprise one or more nucleic acid sequences in HIV comprising: long terminal repeat (LTR) nucleic acid sequences, nucleic acid sequences encoding structural proteins, non-structural proteins or combinations thereof.

[0014] In certain embodiments, the sequences encoding structural proteins comprise nucleic acid sequences encoding: Gag, Gag-Pol precursor, Pro (protease), Reverse Transcriptase (RT), integrase (In), Env or combinations thereof. In certain embodiments, the sequences encoding non-structural proteins comprise nucleic acid sequences encoding: regulatory proteins, accessory proteins or combinations thereof. In certain embodiments, the regulatory proteins comprise: Tat, Rev or combinations thereof. In certain embodiments, the accessory proteins comprise Nef, Vpr, Vpu, Vif or combinations thereof.

[0015] In certain embodiments, a gRNA comprises at least one nucleic acid sequence set forth in Tables 1- 5 or combinations of gRNAs.

[0016] In certain embodiments, a composition further comprises a therapeutically effective amount of a non-nucleoside reverse transcriptase inhibitor (NNRTI), and/or a nucleoside reverse transcriptase inhibitor (NRTI) and/or a protease inhibitor. In certain embodiments, the NNRTI comprises: etravirine, efavirenz, nevirapine, rilpivirine, delavirdine, or nevirapine. In certain embodiments, the NRTI comprises: lamivudine, zidovudine, emtricitabine, abacavir, zalcitabine, dideoxycytidine, azidothymidine, tenofovir disoproxil fumarate, didanosine (ddI EC), dideoxyinosine, stavudine, abacavir sulfate or combinations thereof. In certain embodiments, a protease inhibitor comprises: amprenavir, tipranavir, indinavir, saquinavir mesylate, lopinavir and ritonavir (LPV/RTV), Fosamprenavir Calcium (FOS-APV), ritonavir, darunavir, atazanavir sulfate, nelfinavir mesylate or combinations thereof.

[0017] In certain embodiments, the pharmaceutical composition comprising a therapeutically effective amount of a nanoformulated long-acting slow effective release antiretroviral agent. In certain embodiments, the nanoformulated antiretroviral agent comprises: myristoylated dolutegravir, lamivudine, abacavir, rilpivirine or combinations thereof. In certain embodiments, the pharmaceutical composition comprises at least one an isolated nucleic acid sequence encoding a Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR)-associated endonuclease; at least one isolated nucleic acid sequence encoding at least one guide RNA (gRNA) that is complementary to a target sequence in retroviral DNA; said isolated nucleic acid sequences being included in at least one expression vector. In certain embodiments the pharmaceutical composition comprise the gene-editing agent.

[0018] In certain embodiments, the integrated retroviral DNA is human immunodeficiency virus (HIV) DNA, and said at least one gRNA includes a first gRNA that is complementary to a first target sequence in the HIV DNA, and a second gRNA that is complementary to a second target sequence in the HIV DNA.

[0019] Other aspects are described *infra*.

[0020] *Definitions*

[0021] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which the invention pertains. Although any methods and materials similar or equivalent to those described herein can be used in the practice for testing of the present invention, the preferred materials and methods are described herein. In describing and claiming the present invention, the following terminology will be used. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting.

[0022] The articles “a” and “an” are used herein to refer to one or to more than one (i.e., to at least one) of the grammatical object of the article. By way of example, “an element” means one element or more than one element. Thus, recitation of “a cell”, for example, includes a plurality of the cells of the same type. Furthermore, to the extent that the terms “including”, “includes”, “having”, “has”, “with”, or variants thereof are used in either the detailed description and/or the claims, such terms are intended to be inclusive in a manner similar to the term “comprising.”

[0023] “About” as used herein when referring to a measurable value such as an amount, a temporal duration, and the like, is meant to encompass variations of +/- 20%, +/- 10%, +/- 5%, +/- 1%, or +/- 0.1% from the specified value, as such variations are appropriate to perform the disclosed methods. Alternatively, particularly with respect to biological systems or processes, the term can mean within an order of magnitude within 5-fold, and also within 2-fold, of a value. Where particular values are described in the application and claims, unless otherwise stated the term “about” meaning within an acceptable error range for the particular value should be assumed.

[0024] The term “anti-viral agent” or “anti-retroviral agent” as used herein, refers to any molecule that is used for the treatment of a virus and include agents which alleviate any symptoms associated with the virus, for example, anti-pyretic agents, anti-inflammatory

agents, chemotherapeutic agents, and the like. An antiviral agent includes, without limitation: antibodies, aptamers, adjuvants, anti-sense oligonucleotides, chemokines, cytokines, immune stimulating agents, immune modulating agents, B-cell modulators, T-cell modulators, NK cell modulators, antigen presenting cell modulators, enzymes, siRNA's, ribavirin, protease inhibitors, helicase inhibitors, polymerase inhibitors, neuraminidase inhibitors, nucleoside reverse transcriptase inhibitors, non-nucleoside reverse transcriptase inhibitors, purine nucleosides, chemokine receptor antagonists, interleukins, or combinations thereof. The term also refers to non-nucleoside reverse transcriptase inhibitors (NNRTIs), nucleoside reverse transcriptase inhibitors (NRTIs), analogs, variants etc.

[0025] As used herein, the terms “comprising,” “comprise” or “comprised,” and variations thereof, in reference to defined or described elements of an item, composition, apparatus, method, process, system, etc. are meant to be inclusive or open ended, permitting additional elements, thereby indicating that the defined or described item, composition, apparatus, method, process, system, etc. includes those specified elements--or, as appropriate, equivalents thereof--and that other elements can be included and still fall within the scope/definition of the defined item, composition, apparatus, method, process, system, etc.

[0026] The term “eradication” of a retrovirus, *e.g.* human immunodeficiency virus (HIV), as used herein, means that that virus is unable to replicate, the genome is deleted, fragmented, degraded, genetically inactivated, or any other physical, biological, chemical or structural manifestation, that prevents the virus from being transmissible or infecting any other cell or subject resulting in the clearance of the virus *in vivo*. In some cases, fragments of the viral genome may be detectable, however, the virus is incapable of replication, or infection etc. The presence or absence of the HIV virus can be determined via any means, such as for example, p24 detection or lack thereof, etc.

[0027] An “effective amount” as used herein, means an amount which provides a therapeutic or prophylactic benefit.

[0028] “Encoding” refers to the inherent property of specific sequences of nucleotides in a polynucleotide, such as a gene, a cDNA, or an mRNA, to serve as templates for synthesis of other polymers and macromolecules in biological processes having either a defined sequence of nucleotides (*i.e.*, rRNA, tRNA and mRNA) or a defined sequence of amino acids and the biological properties resulting therefrom. Thus, a gene encodes a protein if transcription and translation of mRNA corresponding to that gene produces the protein in a

cell or other biological system. Both the coding strand, the nucleotide sequence of which is identical to the mRNA sequence and is usually provided in sequence listings, and the non-coding strand, used as the template for transcription of a gene or cDNA, can be referred to as encoding the protein or other product of that gene or cDNA.

[0029] The term “expression” as used herein is defined as the transcription and/or translation of a particular nucleotide sequence driven by its promoter.

[0030] “Expression vector” refers to a vector comprising a recombinant polynucleotide comprising expression control sequences operatively linked to a nucleotide sequence to be expressed. An expression vector comprises sufficient cis-acting elements for expression; other elements for expression can be supplied by the host cell or in an *in vitro* expression system. Expression vectors include all those known in the art, such as cosmids, plasmids (*e.g.*, naked or contained in liposomes) and viruses (*e.g.*, lentiviruses, retroviruses, adenoviruses, and adeno-associated viruses) that incorporate the recombinant polynucleotide.

[0031] “Isolated” means altered or removed from the natural state. For example, a nucleic acid or a peptide naturally present in a living animal is not “isolated,” but the same nucleic acid or peptide partially or completely separated from the coexisting materials of its natural state is “isolated.” An isolated nucleic acid or protein can exist in substantially purified form, or can exist in a non-native environment such as, for example, a host cell.

[0032] An “isolated nucleic acid” refers to a nucleic acid segment or fragment which has been separated from sequences which flank it in a naturally occurring state, *i.e.*, a DNA fragment which has been removed from the sequences which are normally adjacent to the fragment, *i.e.*, the sequences adjacent to the fragment in a genome in which it naturally occurs. The term also applies to nucleic acids which have been substantially purified from other components which naturally accompany the nucleic acid, *i.e.*, RNA or DNA or proteins, which naturally accompany it in the cell. The term therefore includes, for example, a recombinant DNA which is incorporated into a vector, into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote, or which exists as a separate molecule (*i.e.*, as a cDNA or a genomic or cDNA fragment produced by PCR or restriction enzyme digestion) independent of other sequences. It also includes: a recombinant DNA which is part of a hybrid gene encoding additional polypeptide sequence, complementary DNA (cDNA), linear or circular oligomers or polymers of natural and/or modified monomers or linkages, including deoxyribonucleosides, ribonucleosides, substituted

and alpha-anomeric forms thereof, peptide nucleic acids (PNA), locked nucleic acids (LNA), phosphorothioate, methylphosphonate, and the like.

[0033] The nucleic acid sequences may be “chimeric,” that is, composed of different regions. In the context of this invention “chimeric” compounds are oligonucleotides, which contain two or more chemical regions, for example, DNA region(s), RNA region(s), PNA region(s) etc. Each chemical region is made up of at least one monomer unit, *i.e.*, a nucleotide. These sequences typically comprise at least one region wherein the sequence is modified in order to exhibit one or more desired properties.

[0034] Unless otherwise specified, a “nucleotide sequence encoding” an amino acid sequence includes all nucleotide sequences that are degenerate versions of each other and that encode the same amino acid sequence. The phrase nucleotide sequence that encodes a protein or an RNA may also include introns to the extent that the nucleotide sequence encoding the protein may in some version contain an intron(s).

[0035] “Optional” or “optionally” means that the subsequently described event or circumstance can or cannot occur, and that the description includes instances where the event or circumstance occurs and instances where it does not.

[0036] As used in this specification and the appended claims, the term “or” is generally employed in its sense including “and/or” unless the content clearly dictates otherwise.

[0037] “Parenteral” administration of an immunogenic composition includes, *e.g.*, subcutaneous (s.c.), intravenous (i.v.), intramuscular (i.m.), or intrasternal injection, or infusion techniques.

[0038] The terms “patient” or “individual” or “subject” are used interchangeably herein, and refers to a mammalian subject to be treated, with human patients being preferred. In some cases, the methods of the invention find use in experimental animals, in veterinary application, and in the development of animal models for disease, including, but not limited to, rodents including mice, rats, and hamsters, and primates.

[0039] The term “percent sequence identity” or having “a sequence identity” refers to the degree of identity between any given query sequence and a subject sequence.

[0040] As used herein, a “pharmaceutically acceptable” component/carrier etc. is one that is suitable for use with humans and/or animals without undue adverse side effects (such

as toxicity, irritation, and allergic response) commensurate with a reasonable benefit/risk ratio.

[0041] The term “target nucleic acid” sequence refers to a nucleic acid (often derived from a biological sample), to which the oligonucleotide is designed to specifically hybridize. The target nucleic acid has a sequence that is complementary to the nucleic acid sequence of the corresponding oligonucleotide directed to the target. The term target nucleic acid may refer to the specific subsequence of a larger nucleic acid to which the oligonucleotide is directed or to the overall sequence (*e.g.*, gene or mRNA). The difference in usage will be apparent from context.

[0042] To “treat” a disease as the term is used herein, means to reduce the frequency or severity of at least one sign or symptom of a disease or disorder experienced by a subject. Treatment of a disease or disorders includes the eradication of a virus.

[0043] “Treatment” is an intervention performed with the intention of preventing the development or altering the pathology or symptoms of a disorder. Accordingly, “treatment” refers to both therapeutic treatment and prophylactic or preventative measures. “Treatment” may also be specified as palliative care. Those in need of treatment include those already with the disorder as well as those in which the disorder is to be prevented. Accordingly, “treating” or “treatment” of a state, disorder or condition includes: (1) eradicating the virus; (2) preventing or delaying the appearance of clinical symptoms of the state, disorder or condition developing in a human or other mammal that may be afflicted with or predisposed to the state, disorder or condition but does not yet experience or display clinical or subclinical symptoms of the state, disorder or condition; (3) inhibiting the state, disorder or condition, *i.e.*, arresting, reducing or delaying the development of the disease or a relapse thereof (in case of maintenance treatment) or at least one clinical or subclinical symptom thereof; or (4) relieving the disease, *i.e.*, causing regression of the state, disorder or condition or at least one of its clinical or subclinical symptoms. The benefit to an individual to be treated is either statistically significant or at least perceptible to the patient or to the physician.

[0044] As defined herein, a “therapeutically effective” amount of a compound or agent (*i.e.*, an effective dosage) means an amount sufficient to produce a therapeutically (*e.g.*, clinically) desirable result. The compositions can be administered from one or more times per day to one or more times per week; including once every other day. The skilled artisan will appreciate that certain factors can influence the dosage and timing required to effectively

treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and other diseases present. Moreover, treatment of a subject with a therapeutically effective amount of the compounds of the invention can include a single treatment or a series of treatments.

[0045] Where any amino acid sequence is specifically referred to by a Swiss Prot. or GENBANK Accession number, the sequence is incorporated herein by reference. Information associated with the accession number, such as identification of signal peptide, extracellular domain, transmembrane domain, promoter sequence and translation start, is also incorporated herein in its entirety by reference.

[0046] Genes: All genes, gene names, and gene products disclosed herein are intended to correspond to homologs from any species for which the compositions and methods disclosed herein are applicable. It is understood that when a gene or gene product from a particular species is disclosed, this disclosure is intended to be exemplary only, and is not to be interpreted as a limitation unless the context in which it appears clearly indicates. Thus, for example, for the genes or gene products disclosed herein, are intended to encompass homologous and/or orthologous genes and gene products from other species.

[0047] Ranges: throughout this disclosure, various aspects of the invention can be presented in a range format. It should be understood that the description in range format is merely for convenience and brevity and should not be construed as an inflexible limitation on the scope of the invention. Accordingly, the description of a range should be considered to have specifically disclosed all the possible subranges as well as individual numerical values within that range. For example, description of a range such as from 1 to 6 should be considered to have specifically disclosed subranges such as from 1 to 3, from 1 to 4, from 1 to 5, from 2 to 4, from 2 to 6, from 3 to 6 etc., as well as individual numbers within that range, for example, 1, 2, 2.7, 3, 4, 5, 5.3, and 6. This applies regardless of the breadth of the range.

BRIEF DESCRIPTION OF THE DRAWINGS

[0048] FIGS. 1A-1G show the viral and immune profiles from sequential LASER ART and AAV₉-CRISPR-Cas9 treatments of HIV-1 infected humanized mice. FIG. 1A: After infection at week 0 and confirmation of VL, mice were administered 45 mg/kg nanoformulated myristoylated DTG (NMDTG), nanoformulated RPV (NRPV) and 40 mg/kg NM3TC, NMABC. Three weeks after the last LASER ART treatment, a single IV dose of

AAV₉-CRISPR-Cas9 (10^{12} GC units) was administered and left without antiretroviral drugs for an additional five weeks. FIG. 1B: Evaluation of human CD4⁺ T cell numbers in humanized mice by flow cytometry tests on days 0, 3, 5, 7, and 14 of infection. FIG. 1C: Viral load assessment by determining viral RNA copies in plasma at day-14 after HIV-1 NL4-3 infection and prior to LASER ART treatment. FIG. 1D: Detection of human cells and viral infection in various tissues at day-14 after infection. Stains of human HLA-DR in lymph nodes, spleen, and lung show significant human immune cell reconstitution in infected animals. Replicate slides demonstrate HIV-1 p24⁺ stained cells in tissue sections. FIG. 1E: Detection of HIV-1 DNA by semi-nested real-time q-PCR assay in different tissues of HIV-1 infected animals at day-14 of infection. FIG. 1F: Evaluation of viral load shows that after administration of AAV₉-CRISPR-Cas9, two out of seven mice showed no evidence for viral rebound at week-14. Viral load in untreated animals remained high during the course of study. FIG. 1G: FACS analyses of human CD4⁺ T cells are shown with increased numbers in the LASER ART and AAV₉-CRISPR-Cas9 groups. A one-way ANOVA and Bonferroni's post-hoc tests for multiple comparisons and a two-tailed Student's *t-test* was used for statistical analyses in FIG. 1B. * $P < 0.05$, ** $P < 0.01$.

[0049] FIGS. 2A-2C show the excision of the viral DNA fragments by CRISPR-Cas9 in tissues from HIV-1 infected humanized mice treated with LASER ART. FIG. 2A: Schematic illustration of HIV-1_{NL4-3} DNA highlighting the positions of gRNA LTR1 and gRNA Gag D target sites, their nucleotide compositions, and the three possible CRISPR-Cas9 induced break points leading to the excisions of various length of viral DNA fragments. FIG. 2B: Total DNA from spleen, GALT, and kidney from three groups of animals were used for PCR genotyping using a set of primers derived from the 5'LTR, 3'LTR, and gag gene in reaction conditions that are calibrated for efficient amplification of short (less than 600bp) or large DNA fragments. Predicted amplicons of 193bp and 523bp, which result from the excisions of DNA fragments between 5' LTR to gag and gag to 3'LTR, respectively, were selected for DNA sequencing. The fragment of 396bp represents both populations of full length LTRs, as well as the chimeric of both 5' and 3' LTR after excision of entire genome by gRNA LTR1/Cas9 and re-joining of the residual segments of cleaved 5' LTR and 3' LTRs. Several other fragments with closely similar size, caused by InDel mutations, were detected and further analyzed by sequencing. Single asterisks on top of the bands point to the specificity of the fragmental HIV DNA excision by CRISPR-Cas9 as verified by Sanger sequencing (also illustrated in FIGS. 12A-12C and 13A-13M). The double asterisk depicts

non-specific amplification of unrelated DNA or randomly amplified segment of truncated HIV-1 sequence (also see FIGS. 14A-14F). FIG. 2C: Representative DNA sequences from each group were aligned to the reference LTR-Gag region of the HIV-1_{NL4-3} sequence. The positions and nucleotide compositions of targets for gRNAs LTR1 and GagD are shown in green, PAM in red and insertion sequences in yellow. Arrows highlighted positions of small and large deletions.

[0050] FIGS. 3A- 3E show the detection of viral DNA and RNA in various tissues after sequential LASER ART and AAV₉-CRISPR-Cas9 treatments of infected humanized mice. (FIG 3A) HIV-1 DNA and (FIG. 3D) HIV-1 RNA analyses using ultrasensitive semi-nested real-time PCR assays from spleen, bone marrow (BM), GALT, brain, liver, kidney, and lung from treatment groups. The data are expressed as total HIV-1 DNA (FIG. 3A) or HIV-1 RNA (FIG. 3D) copies/10⁶ human CD45⁺ cells. Two animals, #4346 and #4349 [shown by the red squares below the dashed lines (detection limit)], with dual treatments, showed sterilization of virus from all tissues analyzed. FIGS. 3B and 3C: Quantitative PCR showed complete elimination of signals corresponding to *pol* (FIG. 3B) and *env* (FIG. 3C) DNA sequences of HIV-1 in mice #4346 and #4349 (shown by red arrows). FIG. 3E: Representative results from RNAscope assay revealed the detection of single or clusters of brown dots corresponding to HIV-1 RNA in 5 pm-thick spleen sections of infected animals receiving either LASER ART or CRISPR-Cas9, but not both (#4346). *E1* are representative spleen sections obtained from humanized mice infected with HIV-1 (controls); *E2* are HIV-1 infected animals treated only with CRISPR-Cas9; *E3* are HIV-1 infected LASER ART treated animals demonstrating viral rebound after cessation of therapy, and *E4* are infected animals treated first with LASER ART followed by CRISPR-Cas9. *E1-E4* are representative tissue sections taken from each of the animal groups. In these assays, we used the antisense V-HIV1-Clade-B targeting 854-8291 base pairs of HIV-1 as the probe. Human peptidylprolyl isomerase B (PPIB) was used as a positive control for all tissues analyzed. Images are 40x magnification. A one-way ANOVA and Bonferroni's post-hoc tests for multiple comparisons and a two-tailed Student's *t-test* was used for comparisons between two groups as in FIGS. 3A and 3D for statistical analyses. **P* < 0.05, ***P* < 0.01, ****P* < 0.001.

[0051] FIGS. 4A-4G show the viral sterilization in HIV-1_{ADA} infected humanized mice in LASER ART and CRISPR-Cas9 (dual treated) by measures of viral, immune profile and excision profiles. FIG. 4A: The timeline of the experiment showing the temporal administration of LASER ART and CRISPR-Cas9 treatments, and animal sacrifice. FIG 4B:

The percentage of human CD4⁺ T-cells and (FIG. 4C) viral loads measured in the HIV-1 infected and HIV-1 infected and dual treated animal groups. Dual treated animals that showed no or viral rebound are illustrated. FIG. 4D: HIV-1 DNA analysis was performed using ultrasensitive semi-nested real-time q-PCR assays from spleen, GALT, liver, lung, brain and bone marrow from infected and infected and dual treated mice. The data are expressed as total HIV-1 DNA copies/10⁶ human CD45⁺ cells. Two animals, #3319 and #3336 (illustrated by the red squares) were below the dashed lines for virus detection as measured by plasma VL. These animals had no detectable viral DNA after dual treatments demonstrating viral sterilization from all analyzed tissues. A single animal (#3324) is illustrated by a half-red-black designation that had an undetectable VL but viral DNA was observed. FIG. 4E: Ultrasensitive ddPCR, with sensitivity of detecting 1-2 viral copies, was used in cross validation tests for viral DNA detections and performed in all tissues of infected and infected/dual treated animals. As a positive control, one animal each from the HIV-1 infected and HIV-1 and LASER ART groups are illustrated as open structures together. These were placed together with the six infected animals from the dual treatment group illustrated as closed structures. Dashed line represents the limit of detection. FIG. 4F: Agarose gel analyses of the PCR assay of DNA from various tissues of two animals with no rebound shows the presence of segments of HIV-1 LTR DNA and detection of a 121 bp amplicon, indicative of excision of a DNA fragment between the LTR and the *gag* gene (top). The histogram illustrates representative results from sequencing of the 121 bp fragment highlighting the position of the 5' LTR breakpoint, and Gag and PAM trinucleotide on the GagD RNA. FIG. 4G: Splenocytes and bone marrow cells were isolated from HIV-1 infected mice with or without prior LASER ART and/or CRISPR-Cas9 treatments. These cells were then used in adoptive transfers performed in uninfected and drug naïve mice. These transfer experiments performed in CD34⁺ HSC humanized mice were used to examine potential rebound from latent reservoirs not detected by standard ddPCR and nested PCR tests. In addition, as positive controls, two animals from an HIV-1 infected group and one from the LASER ART “alone” treatment group are shown as open circles and boxes. Five animals from the dual treatment group are illustrated as closed circles and boxes. Mice were sacrificed after 30 days and analyzed for plasma viral RNA using the Roche Amplicrep/Taqman-48 V2.0 detection assay. Virus was not detected in 2 “dual-treated” animals (#3319 and #3336, red circles and boxes below the dotted line for the cutoffs for viral detection) in all tests. This was used as the definition of “viral eradication” in these experiments. In contrast, virus was readily identified in all other infected and treated groups.

A one-way ANOVA and Bonferroni's post-hoc tests for multiple comparisons and a two-tailed Student's *t*-test was used for comparisons between two groups as in FIGS. 4B and 4D for statistical analyses. **P* < 0.05.

[0052] FIGS. 5A-5F show the results from a study of viral and CD4⁺ T cell profiles from simultaneously treated HIV-1 infected Hu-PBL mice with LASER ART and AAV₉-CRISPR-Cas9. FIG. 5A: Study scheme illustrates time of human cells reconstitution, HIV-1 infection, LASER ART administration, AAV₉-CRISPR-Cas9 injection (50 µl of 2 × 10¹³ GC/ml), and time of sacrifice and flow cytometric evaluation of pan-human CD4⁺ T cells. FIG. 5B: Peripheral blood cells were assayed prior to and after (day 14) HIV-1 infection. FIG. 5C: Plasma viral load was detected using Roche ampliprep V2.0/Taqman 48 system from different mice groups. FIG. 5D: The DNA analysis from gag region from spleen tissues showed reduced HIV-1 in LASER ART alone which were further decreased in LASER ART plus AAV₉-CRISPR-Cas9 treated groups as compared to HIV-infected but untreated controls and the AAV₉-CRISPR-Cas9 group. FIG. 5E: HIV-1 RNA was analyzed using highly sensitive semi-nested real-time PCR assays from spleen samples of all four groups of mice at the end of the study (day-14 after infection). Significant decreases in HIV-1 RNA in LASER ART alone and LASER ART plus AAV₉-CRISPR-Cas9 treated groups compared to HIV-1 infected but untreated controls were observed. The data are expressed as the ratio of total HIV-1 RNA copies/10⁶ human CD45⁺ cells. FIG. 5F: Quantitative PCR of viral RNA from 14 days HIV-1 infected humanized mice. HIV-1 RNA was analyzed using ultrasensitive semi-nested real-time PCR assays from spleen, lymph node, bone marrow, lung, liver, and GALT obtained from HIV-1 infected humanized mice at day 14.

[0053] FIGS. 6A-6C show the combined effect of ART and CRISPR/Cas9 on HIV-1 infection of Jurkat T cell line. FIG. 6A: Experimental design and procedure. Jurkat cells were infected with HIV-1NL₄₋₃-GFP-P2A-Nef at multiplicity of infection (MOI) 0.001 and 0.01. Next, cells were divided into four groups: one control, DMSO treated and three treated with the cocktail composed of four antiretroviral drugs (ART) at the concentrations of 5x EC₉₀ values (dolutegravir (DTG) 11.1 ng/ml, rilpivirine (RPV) 3.3 ng/ml, lamivudine (3TC) 17.2 µg/ml and abacavir (ABC) 8.3 µg/ml. Second set of experiments was performed using myristoylated, precursor antiretroviral drugs (LASER ART) similarly, at the doses 5x EC₉₀ values (myristoylated dolutegravir (MDTG) 16.7 ng/ml, rilpivirine (RPV) 3.3 ng/ml, myristoylated lamivudine (M3TC) 32.9 µg/ml and myristoylated abacavir (MABC) 14.4 µg/ml). ART/LASER ART treatment was started at day 1, 3 or 5 after infection and fresh

drugs were added daily. At day 6 of infection the drugs were removed to allow efficient lentiviral transduction of Cas9 and gRNAs LTR A and LTR B which was conducted at day 7. At day 8 antiretrovirals were added back and continued for another 4 days. Twelve days after HIV-1 infection cells were collected, genomic DNA was extracted and analyzed by PCR for CRISPR-Cas9 mediated cleavage of viral LTR sequences. FIG. 6B: Quantification of the level of infection at day 7. Cells were fixed with 2% PFA and FACS analysis was performed to measure GFP expressing population for HIV infection/replication *in vitro*. FIG. 6C: Similar to FIG. 6B with exception that cells were treated with modified ART.

[0054] FIGS. 7A-7E show the excision of HIV DNA fragment by CRISPR-Cas9 in ART treated T cells and Patient driven PBMCs. Results from standard PCRs of genomic DNA obtained from infected and treated T cells. The presence of full length LTR (357bp) and truncated, CRISPR-Cas9 induced products (167bp) was examined (FIGS. 7A and 7B) and aligned to HIV genome after Sanger sequencing. FIG. 7C: Results of the truncated PCR product obtained after purification from the agarose gel and TA cloning. gRNAs target sequences are shown in green, PAM sequences in red and PCR primers in blue. Below, a representative example of Sanger sequence tracing of truncated product. The HIV-1 LTR sequence was cleaved by Cas9 at target sites LTR A and LTR B and then joined together, resulting in deletion of 190bp proviral DNA segment. The double cleaved/end-joined site is shown as a breaking point in red. FIG. 7D: PCR results of genomic DNA from PBMC's obtained from HIV positive individual. The presence of full length LTR (357bp) and truncated, CRISPR-Cas9 induced products (167bp) was examined. The cells were pretreated with ART for 5 (line 4.), 3 (line 3.) or 1 day (line 2.) or control, DMSO treated (line 1.). At day 6 drugs were removed and next day Cas9 and gRNAs were delivered by lentiviral transduction. At day-8 ART was resumed for another 4 days when cells were collected and processed same way like Jurkat cells above. FIG. 7E: Alignment of a representative Sanger sequencing results of the truncated PCR products obtained after purification from the agarose gel and TA cloning. gRNAs target sequences are shown in green, PAM sequences in red and PCR primers in blue. Below a representative examples of Sanger sequence tracing of truncated products. The HIV-1 LTR sequence was cleaved by Cas9 at target sites LTR A and LTR B and then joined together, resulting in deletion of 190bp proviral DNA segment. The double cleaved/end-joined site is shown as a breaking point in red. In the case of second clone a short: 5 bp deletion was detected at the cut site (in grey).

[0055] FIGS. 8A, 8B are flow cytometric evaluations of human leukocyte reconstitution in humanized mice. Peripheral blood of human stem cell reconstituted mice was assayed before and after (weeks 2, 6, 9, and 14) HIV-1 infection for the presence of human CD45⁺ (FIG. 8A) and CD3⁺ (FIG. 8B) cells. These experiments were performed to assess levels of humanization throughout the study. Numbers of human CD45⁺ and CD3⁺ cells were consistent within all the treated groups. These included animals not treated, treated with LASER ART or CRISPR-Cas9 alone or in combinations of LASER ART and CRISPR-Cas9. Notably, in the HIV-1 infected mice group, the numbers of CD45⁺ and CD3⁺ human cells in blood of mice were comparable to each of the treatment groups.

[0056] FIG. 9 shows the immunohistology of spleens from HIV-1 infected humanized mice. These mice were administered LASER ART or were left untreated. Animals were sacrificed at the time of CRISPR-Cas9 treatment to determine the presence of human CD4⁺ viral target T cells. Representative images are shown from mice infected with HIV-1NL₄₋₃ with or without LASER ART. Significant reductions in CD4⁺ T cells numbers (brown stained cells) are readily seen in the HIV-1-infected group compared to HIV-1 infected animals treated with LASER ART. Duplicate treatments groups demonstrate adequacy or randomization for CRISPR-Cas9 therapy.

[0057] FIG. 10 shows the verification of the presence of human cells in the spleens of humanized mice. PCR analysis of genomic DNA isolated from the spleens of humanized animals using primer sets specific to human and mouse (for a control) beta-globin genes.

[0058] FIG. 11 shows the excision of the viral DNA fragments by CRISPR-Cas9 in tissues from HIV-1 infected humanized mice with and without treatments with LASER ART. Results from standard PCRs of genomic DNA obtained from lungs, livers and brains of treated animals. The presence of full length LTR (396bp) and truncated, CRISPR-Cas9 induced products (193bp for 5'LTR-gag and 523bp for gag-3'LTR) were tested. * CRISPR-Cas9 mediated excision products. ** Non-related.

[0059] FIGS. 12A-12C show the Sanger sequencing results of the truncated, CRISPR-Cas9 excised HIV-1 genomes. FIG. 12A: Representative examples of canonical, InDel free, CRISPR-Cas9 induced, double cleaved/end-joined HIV-1 genome truncations observed in majority of the tissues of AAV₉-CRISPR-Cas9/gRNA treated animals. On the left, result obtained from the spleen of mouse #4356 using 5'LTR-gag specific primers and on the right sequence from the spleen of mouse #4375 using gag-3'LTR specific

amplification. FIG. 12B: Verification of the presence of 41bp insertion at the CRISPR-Cas9 mediated cleavage site in the viral sequence observed in GALT sample from mouse #4349. FIG. 12C: Sequence of the longer, 160bp insertion found at the Cas9 cleavage site in the kidney sample from the same mouse #4349.

[0060] FIGS. 13A-13M show the Sanger sequencing tracing results of the truncated, CRISPR-Cas9 excised HIV-1 genomes. Representative examples of canonical, InDel free, CRISPR-Cas9 induced, double cleaved/end-joined HIV-1 genome truncations observed in majority of the tissues of AAV₉-Cas9/gRNA treated animals (FIGS. 13A, 13B: GALT; FIGS. 13C, 13D: Kidney; FIGS. 13E, 13F: Lung; FIGS. 13G, 13H: Liver and FIGS. 13I, 13J: Brain). InDel mutation detected at the cleavage/end-joining sites in several tissues are shown in FIGS. 13K, 13L for spleen, FIG. 13M for kidney.

[0061] FIGS. 14A-14F show Sanger sequencing results of a few, non-related to CRISPR-Cas9, truncated HIV-1 amplicons detected in some of the samples. Sequences were aligned to HIV-1NL₄₋₃ sequence as a reference. The positions and nucleotide compositions of targets for gRNAs LTR1 and GagD are shown in green, PAMs in red. The sequencing data revealed lack of CRISPR-Cas9 specific cleavage (3 nucleotides from PAM) at the target sites LTR 1 (5'LTR in FIG. 14A for spleen lane 3 and 7, 3'LTR for kidney in FIG. 14C, lane 11, in FIG. 14D for lung lane 4 and FIG. 14F for brain, lane 9) or GagD (in FIG. 14C for kidney lane 7 and in FIG. 14E for liver lanes 2 and 16). Partial 3'LTR sequence was obtained for spleen lane 16 (FIG. 14B).

[0062] FIG. 15 shows the Cas9/gRNAs expression in the spleens of treated animals. Reverse transcription-PCR analysis of RNA extracted from spleens of treated animals to represent SaCas9 mRNA (top panels), single guide RNAs: LTR 1 (second row panels) and Gag D (third row panels) and a control beta-actin mRNA (bottom panels) were detected using primer sets specific to each target.

[0063] FIGS. 16A-16C show the hierarchical clustering analysis of the truncation efficiencies across different animals, treatments, tissues and HIV-1 gene segments. Probabilities are shown with the numbers as well as the heat-map intensities. Most similar groups are clustered together. Dendrograms indicate the hierarchy of clusters for each axis. FIG. 16A: Clustering of truncation efficiencies of different HIV-1 segments in different tissues under ART, CRISPR-Cas9 and ART plus CRISPR-Cas9 treatments. The clustering reveals the most similarity between ART plus CRISPR-Cas9-mediated editing in GALT,

spleen and lung. FIG. 16B: Clustering of truncation efficiencies of different HIV-1 segments and qPCR data in different animals under ART, CRISPR-Cas9 and ART plus CRISPR-Cas9 treatments. The clustering scheme has recognized the similarity patterns and grouped the animals with the similar treatments under the same clusters. FIG. 16C: Clustering of truncation efficiencies in different tissues of the animals under the aforementioned treatments. Note that the animals with no rebound (treated with both LASER ART and AAV₉-CRISPR-Cas9) exhibit similar patterns in excision probabilities both across different HIV-1 segments and across different tissues. These analyses are later used in drawing the significance levels of combined treatment in viral genome eradication compared to the control groups. S1 refers to 5' LTR-Gag and S2 refers to Gag-3' LTR of the HIV-1 gene, respectively.

[0064] FIGS. 17A-17C show the Off target effect in cell model (FIG. 17A) of genomic DNA obtained from TZM-bl single cell clones: two controls (C1-2) and six Cas9/gRNA LTR 1 + Gag D treated (E1-6). The presence of full length LTR -454/+43 (497bp) was examined. Amplicons containing CRISPR-Cas9 specific InDel mutations at the LTR 1 target site in integrated HIV-1 LTR sequence are pointed by asterisks. Single asterisks indicate deletions, double asterisks insertions. FIG. 17B: Alignment of a representative Sanger sequencing results of HIV-1 LTR specific amplicons. The positions and nucleotide compositions of target for gRNA LTR1 is shown in green, PAM in red, sequence deletions in grey and sequence insertions in yellow, PCR primers in blue. FIG. 17C: Representative Sanger sequencing tracing of LTR 1 region of HIV-1 LTRs obtained for each single cell clone. The positions and nucleotide compositions of target for gRNAs LTR1 is shown in green, PAM in red, sequence deletions in grey.

[0065] FIGS. 18A-18F show representative Sanger sequencing tracing of predicted three Off target regions for gRNAs LTR 1 and Gag D obtained for each single cell clone. The positions and nucleotide compositions of Off target sites are shown in green, PAMs in red. Red squares point mismatched nucleotides comparing to target sequences. LTR 1 off target sites: TSC2 (FIG. 17A), TUB (FIG. 17B) and ch8 (FIG. 17C). Gag D off target sites: TACC2 (FIG. 17D), ADNP (FIG. 17E) and ch3 (FIG. 17F). No any InDel mutations at the predicted off target sites was detected. See also Tables 4 and 5.

[0066] FIGS. 19A-19C show the appearance of Somatic mutations in humanized mice. FIG. 19A: Sequence of NGS data analysis steps used for off target detection. FIG. 19B: Number of different types of somatic structural variations (SV) in each sample.

Abbreviations: TRA: (Translocation) the number of translocations, INV: (Inversions) the number of inversions, DEL: (Deletion) the number of deletions, DUP: (Tandem duplication) the number of tandem duplications, INS: (Insertion) the number of insertions. FIG. 19C: The size of genomic regions affected by somatic CNVs in each sample.

[0067] FIGS. 20A-20D are Circos diagrams of the animals (FIG. 20A), #3539 ((LASER ART), (FIG. 20B) #4346 and, (FIG. 20C) #4349 (CRISPR-Cas9+LASER ART), and (FIG. 20D) #4356 (CRISPR-Cas9). The diagrams consist of seven rings. From outer to inner rings: (1) the outer circle (the first circle) is chromosome information. (2) The second ring represents the read coverage in histogram style. A histogram is the average coverage of a 0.5Mbp region. (3) The third ring represents InDel density in scatter style. A black dot is calculated as InDel number in a range of 1Mbp. (4) the fourth ring represents SNP density in scatter style. A green dot is calculated as SNP number in a range of 1Mbp. (5) the fifth ring represents the proportion of homozygous SNP (orange) and heterozygous SNP (grey) in histogram style. A histogram is calculated from a 1Mbp region. (6) The sixth ring represents the CNV inference. Red means gain, and green means loss. (7) The most central ring represents the SV inference in exonic and splicing regions. TRA (orange), INS (green), DEL (grey), DUP (pink) and INV (blue).

[0068] FIGS. 21A, 21B are an analysis of humanized mice tissues using highly sensitive ddPCR assay to detect HIV-1. Ultrasensitive droplet digital PCR (ddPCR) with sensitivity of detecting 1-2 copies was used to detect viral DNA in spleen of the infected animals belonging to 4 groups, control infected, LASER ART or AAV9-CRISPR-Cas9 alone treated and dual treatment (LASER ART+Cas9) (FIG. 21A) and the various organs of the two mice with no viral rebound (FIG. 21B). Note that the two animals with double treatment group (group-4, #4346 and #4349) showed complete elimination of virus in spleen and the other tissues (Lung, liver, GALT, brain and Kidney) tested.

[0069] FIGS. 22A and 22B show a viral recovery assay using co-culture method. FIG. 22A: Splenocytes and bone marrow cells were isolated from HIV-1 infected mice with or without prior LASER ART and/or CRISPR-Cas9 treatments then co-cultivated with PHA/IL-2 stimulated human PBMCs. Cells were harvested 12 days post-cocultivation for HIV-1 DNA (FIG. 22A) and (FIG. 22B) RNA and looked to examine rebound virus using highly sensitive semi-nested real-time q-PCR assay. Data are expressed as total viral copies/ 10^6 human CD45+ cells. Dual LASER ART and CRISPR-Cas9 treatments mice

resulted in no detection of viral nucleic acids, which were also confirmed by reverse transcriptase assay of culture supernatants. Virus was detected in all other groups of animals.

[0070] FIG. 23 shows tissue analyses of HIV-1_{ADA} infected and treated humanized mice by RNAscope. RNAscope was used to detect viral RNA in spleens and demonstrating single brown dots or cluster of dots in 5- μ m thick sections. The assays used antisense probeV-HIV-1-Calde-B targeting 854-8291 base pairs of the HIV-1 genome. Mouse #3319 which received LASER ART and AAV₉-CRISPR-Cas9, showed no viral detection signals. Viral RNA was detected in other 2 groups of humanized mice spleen (HIV-1_{ADA} infected and infected + LASER ART treated) as shown. The photomicrographs are representative images from each group. Human peptidyl Isomerase B (PIIB) was used as a positive control for every tissue analyzed. Images are 40 x magnifications.

[0071] FIGS. 24A-24E show the detection of HIV-1_{ADA} DNA and RNA in spleen tissues in adoptively transferred humanized mice. Splenocytes and bone marrow cells were isolated from HIV-1 infected mice with or without prior LASER ART and or CRISPR-Cas9 treatments. These were for adoptive transfers into “new” CD34⁺ NSG-humanized mice. The intent was to perform cross disciplinary viral amplification from known infectious cell reservoirs. FIGS. 24A, 24B and 24C: HIV-1 DNA and (FIG. 24D) RNA analyses using ultrasensitive semi-nested real-time qPCR assays from spleen, bone marrow and lung tissues of adoptively transferred humanized mice. The data are expressed as total HIV-1 DNA or RNA copies/10⁶ human CD45⁺ cells. Four animals (splenocyte and bone marrow cells isolated and adoptively transferred from #3319 and #3336) mice (shown by red circles and squares below dotted line), showed no viral recovery. The above data was further confirmed using ultrasensitive ddPCR assay (with sensitivity of 1-2 copies), where the same four target adoptively transferred recipient animals showed no HIV-1 and (FIG. 24E) indicating complete elimination of virus. In mice from HIV-1_{ADA} infected with or without LASER ART treatment showed easily recovered virus in the spleen tissues. These results provide definitive testing of viral eradication in the two tested and the assayed mice (#3319 and #3336).

[0072] FIGS. 25A-25C show the excision of HIV proviral DNA by CRISPR-Cas9 in HIV_{ADA}-infected humanized mice. A much shorter fragment (193 bp) of excised HIV proviral DNA from the 5’LTR to gag region was amplified by nested-PCR in total genomic DNA extracted from various tissues of each humanized mice (#3324 and #3349) (FIG. 25A) along with the presence of SaCas9 DNA in each tissue (FIG. 25B). HIV excision was not detected in the humanized mouse treated with LASER ART only (#3357) even though a full

length of HIV-1 LTR could be amplified abundantly to reveal the existence of HIV proviral DNA (FIG. 25C).

[0073] FIG. 26 shows liver tissue histology following therapy in humanized mice. Hematoxylin and eosin staining of representative sections from liver tissues in uninfected, HIV-1_{ADA}-infected, infected and LASER ART treated and dual treated (LASER ART+AAV₉-CRISPR-Cas9) humanized mice at the endpoint of the study. Tissue pathology was not observed in LASER ART alone nor the dual treatment mice group. All images were captured at 10-X magnification.

[0074] FIG. 27 shows the gating strategy. Blood cells were first gated for mononuclear cells and lymphocytes using forward and side scattered panels (FSC and SSC). From the gated lymphocyte population, human CD45⁺ cells were re-gated in side-scatter panel. Gated human CD45⁺ mononuclear cells were assessed for expression of human CD3 (T cells) and CD19 (B cells). CD3⁺ T cells were further gated to assess the expression of CD4 and CD8 cells.

DETAILED DESCRIPTION

[0075] Embodiments of the invention are directed in general to nanoparticle delivery of long-acting, slow effective release (LASER) antiretroviral therapy (ART) and gene editing technologies.

[0076] Briefly, the invention is based, in part, on the finding that treatment of HIV-1 infected humanized mice with CRISPR-Cas9 designed to edit the HIV-1 genome following two months treatment with the newly developed long-acting, slow effective release ART (LASER ART) eradicated HIV-1 infection in twenty-nine percent of infected animals with restored CD4⁺ T cells. Ultrasensitive nested and digital droplet PCR and RNA scope assays failed to detect HIV-1 in blood, spleen, lung, kidney, liver, gut-associated lymphoid tissue and brain. Excision of proviral DNA fragments spanning the LTRs and the Gag gene by CRISPR/Cas9, in the absence of any off target effects, along with the lack of viral rebound following cessation of ART with no progeny virus recovery verified HIV-1 eradication. Thus, the sequential application of antiretroviral agents and CRISPR-Cas9 therapies administered to HIV-1 infected humanized mice provided the first proof of concept that viral sterilization is possible.

[0077] *LASER ART*: Long-acting slow effective release ART (LASER ART) enable improved pharmacokinetic profiles and reservoir targeting. These antiretrovirals (ARVs)

overcome limitations of current drugs associated with *in vivo* delivery and tissue penetrance. The gene editing agent also had improved delivery and improved the therapeutic index of the drugs.

[0078] Dolutegravir, lamivudine, abacavir and rilpivirine (DTG, 3TC, ABC and RPV respectively) were transformed into long-acting drugs. Drug solubility, dissolution, metabolism, protein-binding, and excretion rates for each of the antiretroviral drugs were optimized and each were shown to influence the drug's half-life and biodistribution profiles. These studies provided the means to transform standard daily or twice-daily antiretroviral drugs into hydrophobic drug crystals to extend the drug's half-life and alter its solubility and metabolic patterns. The drugs were found to possess significant antiretroviral efficacy and high tolerability for conversion into a long-acting compound. Reversible chemical modification and polymer coating techniques were developed to convert each into a long-acting nanoformulation. Change of the antiretroviral drug (ARV) structure was made through reversible myristoylation of the native compound creating a water insoluble prodrug with commensurate crystal formation. When the drug crystals were packaged into a nanoparticle, they were rapidly taken up by human monocyte-derived macrophages (MDM), slowly released from the cells, and retained for a prolonged period inside the macrophage. These chemical and biological outcomes improved drug bioavailability and increased *in vitro* antiretroviral activity up to 100-fold. Pharmacokinetic and pharmacodynamic profiles were improved up to 10-fold over a native drug formulation, exhibiting broad tissue distribution and increased potency. The studies herein provide evidence that ARV conversion into a long-acting slow release formulation is readily achieved. As such, the drug-encased nanoparticles were employed as a "first-step" measure to facilitate drug penetrance into viral reservoirs to facilitate the actions of the excision Cas9 system.

[0079] Accordingly, in certain embodiments, the anti-retroviral agents are formulated into long-acting nanoformulated agents or compounds.

[0080] *Gene Editing Agents:* The application of Cas9 technology in eradicating HIV-1 reservoir, particularly targeting LTR, has been shown to be a promising strategy for treating and possibly curing AIDS. Hu, *et al.*, *PNAS* 2014, 111:114616, disclosed that stable transfection of human cell cultures with plasmids expressing Cas9/gRNAs targeted to sites in the HIV-1 LTR successfully eradicated part and/or the entire HIV-1 genome without compromising host cell function. The targeted sites were termed LTR-A, LTR-B, LTR-C, and LTR-D. The targeting of two different sites in the LTR was particularly effective at

producing the deletions sufficiently extensive to constitute the excision of all or substantially all of the proviral DNA sequence. The pre-existence of Cas9/gRNAs in cells also prevented new HIV-1 infection.

[0081] HIV and other retroviruses are highly mutable, so there is a need for a broader spectrum of Cas9/gRNA reagents and methods for targeting the integrated HIV genome. Of particular use would be Cas9/gRNA reagents that effectively target various genes in the viral genome, such as for example, structural genes of HIV, such as gag and pol; genes that encode ligands that allow for viral entry into cells, etc..

[0082] Accordingly, embodiments of the invention are directed to compositions and methods for the treatment and eradication of highly mutable and/or latent viruses from a host cell *in vitro* or *in vivo*. Methods of the invention may be used to remove viral or other foreign genetic material from a host organism, without interfering with the integrity of the host's genetic material. A nuclease may be used to target viral nucleic acid, thereby interfering with viral replication or transcription or even excising the viral genetic material from the host genome. The nuclease may be specifically targeted to remove only the viral nucleic acid without acting on host material either when the viral nucleic acid exists as a particle within the cell or when it is integrated into the host genome. Targeting the viral nucleic acid can be done using a sequence-specific moiety such as a guide RNA that targets viral genomic material for destruction by the nuclease and does not target the host cell genome. In some embodiments, a CRISPR/Cas nuclease and guide RNA (gRNA) that together target and selectively edit or destroy viral genomic material is used. The CRISPR (clustered regularly interspaced short palindromic repeats) is a naturally-occurring element of the bacterial immune system that protects bacteria from phage infection. The guide RNA localizes the CRISPR/Cas complex to a viral target sequence. Binding of the complex localizes the Cas endonuclease to the viral genomic target sequence causing breaks in the viral genome. Other nuclease systems can be used including, for example, zinc finger nucleases, transcription activator-like effector nucleases (TALENs), meganucleases, or any other system that can be used to degrade or interfere with viral nucleic acid without interfering with the regular function of the host's genetic material.

[0083] The compositions embodied herein, can be used to target viral nucleic acid in any form or at any stage in the viral life cycle. Together, with the combination of LASER-ART therapeutics, renders these compositions formidable in the treatment and/or prevention of infection by a retrovirus, e.g. HIV. The targeted viral nucleic acid may be present in the

host cell as independent particles. In a preferred embodiment, the viral infection is latent and the viral nucleic acid is integrated into the host genome. Any suitable viral nucleic acid may be targeted for cleavage and digestion.

[0084] *CRISPR/Cas Systems:* The CRISPR-Cas system includes a gene editing complex comprising a CRISPR-associated nuclease, *e.g.*, Cas9, and a guide RNA complementary to a target sequence situated on a DNA strand, such as a target sequence in proviral DNA integrated into a mammalian genome. The gene editing complex can cleave the DNA within the target sequence. This cleavage can in turn cause the introduction of various mutations into the proviral DNA, resulting in inactivation of HIV provirus. The mechanism by which such mutations inactivate the provirus can vary. For example, the mutation can affect proviral replication, and viral gene expression. The mutations may be located in regulatory sequences or structural gene sequences and result in defective production of HIV. The mutation can comprise a deletion. The size of the deletion can vary from a single nucleotide base pair to about 10,000 base pairs. In some embodiments, the deletion can include all or substantially all of the integrated retroviral nucleic acid sequence. In some embodiments the deletion can include the entire integrated retroviral nucleic acid sequence. The mutation can comprise an insertion, that is, the addition of one or more nucleotide base pairs to the pro-viral sequence. The size of the inserted sequence also may vary, for example from about one base pair to about 300 nucleotide base pairs. The mutation can comprise a point mutation, that is, the replacement of a single nucleotide with another nucleotide. Useful point mutations are those that have functional consequences, for example, mutations that result in the conversion of an amino acid codon into a termination codon or that result in the production of a nonfunctional protein.

[0085] In general, CRISPR/Cas proteins comprise at least one RNA recognition and/or RNA binding domain. RNA recognition and/or RNA binding domains interact with guide RNAs. CRISPR/Cas proteins can also comprise nuclease domains (*i.e.*, DNase or RNase domains), DNA binding domains, helicase domains, RNase domains, protein-protein interaction domains, dimerization domains, as well as other domains. Active DNA-targeting CRISPR-Cas systems use 2 to 4 nucleotide protospacer-adjacent motifs (PAMs) located next to target sequences for self versus non-self discrimination. ARMAN-1 has a strong 'NGG' PAM preference. Cas9 also employs two separate transcripts, CRISPR RNA (crRNA) and trans-activating CRISPR RNA (tracrRNA), for RNA-guided DNA cleavage. Putative tracrRNA was identified in the vicinity of both ARMAN-1 and ARMAN-4 CRISPR-Cas9

systems (Burstein, D. *et al.* New CRISPR-Cas systems from uncultivated microbes. *Nature*. 2017 Feb 9;542(7640):237-241. doi: 10.1038/nature21059. Epub 2016 Dec 22).

[0086] In embodiments, the CRISPR/Cas-like protein can be a wild type CRISPR/Cas protein, a modified CRISPR/Cas protein, or a fragment of a wild type or modified CRISPR/Cas protein. The CRISPR/Cas-like protein can be modified to increase nucleic acid binding affinity and/or specificity, alter an enzymatic activity, and/or change another property of the protein. For example, nuclease (i.e., DNase, RNase) domains of the CRISPR/Cas-like protein can be modified, deleted, or inactivated. Alternatively, the CRISPR/Cas-like protein can be truncated to remove domains that are not essential for the function of the fusion protein. The CRISPR/Cas-like protein can also be truncated or modified to optimize the activity of the effector domain of the fusion protein.

[0087] In embodiments, the CRISPR/Cas system can be a type I, a type II, or a type III system. Non-limiting examples of suitable CRISPR/Cas proteins include Cas9, CasX, CasY.1, CasY.2, CasY.3, CasY.4, CasY.5, CasY.6, spCas, eSpCas, SpCas9-HF1, SpCas9-HF2, SpCas9-HF3, SpCas9-HF4, ARMAN 1, ARMAN 4, Cas3, Cas4, Cas5, Cas5e (or CasD), Cas6, Cas6e, Cas6f, Cas7, Cas8a1, Cas8a2, Cas8b, Cas8c, Cas9, Cas10, Cas10d, CasF, CasG, CasH, Csy1, Csy2, Csy3, Cse1 (or CasA), Cse2 (or CasB), Cse3 (or CasE), Cse4 (or CasC), Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csz1, Csx15, Csf1, Csf2, Csf3, Csf4, and Cu1966.

[0088] The Cas9 can be an orthologous. Six smaller Cas9 orthologues have been used and reports have shown that Cas9 from *Staphylococcus aureus* (SaCas9) can edit the genome with efficiencies similar to those of SpCas9, while being more than 1 kilobase shorter.

[0089] In addition to the wild type and variant Cas9 endonucleases described, embodiments of the invention also encompass CRISPR systems including newly developed “enhanced-specificity” *S. pyogenes* Cas9 variants (eSpCas9), which dramatically reduce off target cleavage. These variants are engineered with alanine substitutions to neutralize positively charged sites in a groove that interacts with the non-target strand of DNA. This aim of this modification is to reduce interaction of Cas9 with the non-target strand, thereby encouraging re-hybridization between target and non-target strands. The effect of this modification is a requirement for more stringent Watson-Crick pairing between the gRNA

and the target DNA strand, which limits off-target cleavage (Slaymaker, I.M. *et al.* (2015) DOI:10.1126/science.aad5227).

[0090] In certain embodiments, three variants found to have the best cleavage efficiency and fewest off-target effects: SpCas9 (K855A), SpCas9 (K810A/K1003A/R1060A) (a.k.a. eSpCas9 1.0), and SpCas9(K848A/K1003A/R1060A) (a.k.a. eSPCas9 1.1) are employed in the compositions. The invention is by no means limited to these variants, and also encompasses all Cas9 variants (Slaymaker, I.M. *et al. Science*. 2016 Jan 1;351(6268):84-8. doi: 10.1126/science.aad5227. Epub 2015 Dec 1). The present invention also includes another type of enhanced specificity Cas9 variant, “high fidelity” spCas9 variants (HF-Cas9). Examples of high fidelity variants include SpCas9-HF1 (N497A/R661A/Q695A/Q926A), SpCas9-HF2 (N497A/R661A/Q695A/Q926A/D1135E), SpCas9-HF3 (N497A/R661A /Q695A/ Q926A/ L169A), SpCas9-HF4 (N497A/R661A/Q695A/Q926A/Y450A). Also included are all SpCas9 variants bearing all possible single, double, triple and quadruple combinations of N497A, R661A, Q695A, Q926A or any other substitutions (Kleinstiver, B. P. *et al.*, 2016, *Nature*. DOI: 10.1038/nature16526).

[0091] As used herein, the term “Cas” is meant to include all Cas molecules comprising variants, mutants, orthologues, high-fidelity variants and the like.

[0092] In one embodiment, the endonuclease is derived from a type II CRISPR/Cas system. In other embodiments, the endonuclease is derived from a Cas9 protein and includes Cas9, CasX, CasY.1, CasY.2, CasY.3, CasY.4, CasY.5, CasY.6, spCas, eSpCas, SpCas9-HF1, SpCas9-HF2, SpCas9-HF3, SpCas9-HF4, ARMAN 1, ARMAN 4, mutants, variants, high-fidelity variants, orthologs, analogs, fragments, or combinations thereof. The Cas9 protein can be from *Streptococcus pyogenes*, *Streptococcus thermophilus*, *Streptococcus* sp., *Nocardiopsis dassonvillei*, *Streptomyces pristinaespiralis*, *Streptomyces viridochromogenes*, *Streptomyces viridochromogenes*, *Streptosporangium roseum*, *Alicyclobacillus acidocaldarius*, *Bacillus pseudomycooides*, *Bacillus selenitireducens*, *Exiguobacterium sibiricum*, *Lactobacillus delbrueckii*, *Lactobacillus salivarius*, *Microscilla marina*, *Burkholderiales bacterium*, *Polaromonas naphthalenivorans*, *Polaromonas* sp., *Crocospaera watsonii*, *Cyanothece* sp., *Microcystis aeruginosa*, *Synechococcus* sp., *Acetohalobium arabaticum*, *Ammonifex degensii*, *Caldicelulosiruptor beccsii*, *Candidatus Desulfurudis*, *Clostridium botulinum*, *Clostridium difficile*, *Finegoldia magna*, *Natranaerobius thermophilus*, *Pelotomaculum thermopropionicum*, *Acidithiobacillus caldus*,

Acidithiobacillus ferrooxidans, *Allochromatium vinosum*, *Marinobacter* sp., *Nitrosococcus halophilus*, *Nitrosococcus watsoni*, *Pseudoalteromonas haloplanktis*, *Ktedonobacter racemifer*, *Methanohalobium evestigatum*, *Anabaena variabilis*, *Nodularia spumigena*, *Nostoc* sp., *Arthrospira maxima*, *Arthrospira platensis*, *Arthrospira* sp., *Lyngbya* sp., *Microcoleus chthonoplastes*, *Oscillatoria* sp., *Petrotoga mobilis*, *Thermosipho africanus*, or *Acaryochloris marina*. Included are Cas9 proteins encoded in genomes of the nanoarchaea ARMAN-1 (*Candidatus* Micrarchaeum acidiphilum ARMAN-1) and ARMAN-4 (*Candidatus* Parvarchaeum acidiphilum ARMAN-4), CasY (*Kerfeldbacteria*, *Vogelbacteria*, *Komeilibacteria*, *Katanobacteria*), CasX (*Planctomycetes*, *Deltaproteobacteria*).

[0093] Embodiments of the invention also include a new type of class 2 CRISPR-Cas system found in the genomes of two bacteria recovered from groundwater and sediment samples. This system includes Cas1, Cas2, Cas4 and an approximately ~ 980 amino acid protein that is referred to as CasX. The high conservation (68% protein sequence identity) of this protein in two organisms belonging to different phyla, Deltaproteobacteria and Planctomycetes, suggests a recent cross-phyla transfer. The CRISPR arrays associated with each CasX has highly similar repeats (86% identity) of 37 nucleotides (nt), spacers of 33–34 nt, and a putative tracrRNA between the Cas operon and the CRISPR array. Distant homology detection and protein modeling identified a RuvC domain near the CasX C-terminal end, with organization reminiscent of that found in type V CRISPR-Cas systems. The rest of the CasX protein (630 N-terminal amino acids) showed no detectable similarity to any known protein, suggesting this is a novel class 2 effector. The combination of tracrRNA and separate Cas1, Cas2 and Cas4 proteins is unique among type V systems, and phylogenetic analyses indicate that the Cas1 from the CRISPR-CasX system is distant from those of any other known type V. Further, CasX is considerably smaller than any known type V proteins: 980 aa compared to a typical size of about 1,200 amino acids for Cpf1, C2c1 and C2c3 (Burstein, D. *et al.*, 2017 *supra*).

[0094] Another new class 2 Cas protein is encoded in the genomes of certain candidate phyla radiation (CPR) bacteria. This approximately 1,200 amino acid Cas protein, termed CasY, appears to be part of a minimal CRISPR-Cas system that includes Cas1 and a CRISPR array. Most of the CRISPR arrays have unusually short spacers of 17–19 nt, but one system, which lacks Cas1 (CasY.5), has longer spacers (27–29 nt). Accordingly, in some embodiments of the invention, the CasY molecules comprise CasY.1, CasY.2, CasY.3, CasY.4, CasY.5, CasY.6, mutants, variants, analogs or fragments thereof.

[0095] In some embodiments, the CRISPR/Cas-like protein can be derived from a wild type Cas protein or fragment thereof. In other embodiments, the CRISPR/Cas-like protein can be derived from modified Cas proteins. For example, the amino acid sequence of the Cas9 protein can be modified to alter one or more properties (e.g., nuclease activity, affinity, stability, etc.) of the protein. Alternatively, domains of the Cas9 protein not involved in RNA-guided cleavage can be eliminated from the protein such that the modified Cas9 protein is smaller than the wild type Cas9 protein.

[0096] In some embodiments, the CRISPR-associated endonuclease can be a sequence from another species, for example, other bacterial species, bacteria genomes and archaea, or other prokaryotic microorganisms. Alternatively, the wild type Cas9, CasX, CasY.1, CasY.2, CasY.3, CasY.4, CasY.5, CasY.6, ARMAN 1, ARMAN 4, sequences can be modified. The nucleic acid sequence can be codon optimized for efficient expression in mammalian cells, *i.e.*, “humanized.” A humanized Cas9 nuclease sequence can be for example, the Cas9 nuclease sequence encoded by any of the expression vectors listed in GENBANK accession numbers KM099231.1 GI:669193757; KM099232.1 GI:669193761; or KM099233.1 GI:669193765. Alternatively, the Cas9, CasX, CasY.1, CasY.2, CasY.3, CasY.4, CasY.5, CasY.6, ARMAN 1, ARMAN 4, sequences can be for example, the sequence contained within a commercially available vector such as PX330 or PX260 from Addgene (Cambridge, MA). In some embodiments, the Cas9 endonuclease can have an amino acid sequence that is a variant or a fragment of any of the Cas9 endonuclease sequences of GENBANK accession numbers KM099231.1 GI:669193757; KM099232.1 GI:669193761; or KM099233.1 GI:669193765, or Cas9 amino acid sequence of PX330 or PX260 (Addgene, Cambridge, MA).

[0097] The wild type Cas9, CasX, CasY.1, CasY.2, CasY.3, CasY.4, CasY.5, CasY.6, ARMAN 1, ARMAN 4, sequences can be a mutated sequence. For example, the Cas9 nuclease can be mutated in the conserved HNH and RuvC domains, which are involved in strand specific cleavage. In another example, an aspartate-to-alanine (D10A) mutation in the RuvC catalytic domain allows the Cas9 nickase mutant (Cas9n) to nick rather than cleave DNA to yield single-stranded breaks, and the subsequent preferential repair through HDR can potentially decrease the frequency of unwanted indel mutations from off-target double-stranded breaks. The sequences of Cas9, CasX, CasY.1, CasY.2, CasY.3, CasY.4, CasY.5, CasY.6, spCas, eSpCas, SpCas9-HF1, SpCas9-HF2, SpCas9-HF3, SpCas9-HF4, ARMAN 1, ARMAN 4, mutants, variants, high-fidelity variants, orthologs, analogs, fragments, or

combinations thereof, can be modified to encode biologically active variants, and these variants can have or can include, for example, an amino acid sequence that differs from a wild type by virtue of containing one or more mutations (*e.g.*, an addition, deletion, or substitution mutation or a combination of such mutations). One or more of the substitution mutations can be a substitution (*e.g.*, a conservative amino acid substitution). For example, a biologically active variant of a Cas9, CasX, CasY.1, CasY.2, CasY.3, CasY.4, CasY.5, CasY.6, spCas, eSpCas, SpCas9-HF1, SpCas9-HF2, SpCas9-HF3, SpCas9-HF4, ARMAN 1, ARMAN 4, polypeptides can have an amino acid sequence with at least or about 50% sequence identity (*e.g.*, at least or about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity) to a wild type Cas9, CasX, CasY.1, CasY.2, CasY.3, CasY.4, CasY.5, CasY.6, spCas, eSpCas, SpCas9, ARMAN 1, ARMAN 4 polypeptides. Conservative amino acid substitutions typically include substitutions within the following groups: glycine and alanine; valine, isoleucine, and leucine; aspartic acid and glutamic acid; asparagine, glutamine, serine and threonine; lysine, histidine and arginine; and phenylalanine and tyrosine. The amino acid residues in the Cas9, CasX, CasY.1, CasY.2, CasY.3, CasY.4, CasY.5, CasY.6, spCas, eSpCas, SpCas9-HF1, SpCas9-HF2, SpCas9-HF3, SpCas9-HF4, ARMAN 1, ARMAN 4, amino acid sequence can be non-naturally occurring amino acid residues. Naturally occurring amino acid residues include those naturally encoded by the genetic code as well as non-standard amino acids (*e.g.*, amino acids having the D-configuration instead of the L-configuration). The present peptides can also include amino acid residues that are modified versions of standard residues (*e.g.* pyrrolysine can be used in place of lysine and selenocysteine can be used in place of cysteine). Non-naturally occurring amino acid residues are those that have not been found in nature, but that conform to the basic formula of an amino acid and can be incorporated into a peptide. These include D-alloisoleucine(2R,3S)-2-amino-3-methylpentanoic acid and L-cyclopentyl glycine (S)-2-amino-2-cyclopentyl acetic acid. For other examples, one can consult textbooks or the worldwide web (a site currently maintained by the California Institute of Technology displays structures of non-natural amino acids that have been successfully incorporated into functional proteins).

[0098] Two nucleic acids or the polypeptides they encode may be described as having a certain degree of identity to one another. For example, a Cas9 protein and a biologically active variant thereof may be described as exhibiting a certain degree of identity. Alignments may be assembled by locating short Cas9 sequences in the Protein Information Research

(PIR) site (pir.georgetown.edu), followed by analysis with the “short nearly identical sequences” Basic Local Alignment Search Tool (BLAST) algorithm on the NCBI website (ncbi.nlm.nih.gov/blast).

[0099] A percent sequence identity to Cas9 can be determined and the identified variants may be utilized as a CRISPR-associated endonuclease and/or assayed for their efficacy as a pharmaceutical composition. A naturally occurring Cas9 can be the query sequence and a fragment of a Cas9 protein can be the subject sequence. Similarly, a fragment of a Cas9 protein can be the query sequence and a biologically active variant thereof can be the subject sequence. To determine sequence identity, a query nucleic acid or amino acid sequence can be aligned to one or more subject nucleic acid or amino acid sequences, respectively, using the computer program ClustalW (version 1.83, default parameters), which allows alignments of nucleic acid or protein sequences to be carried out across their entire length (global alignment). See Chenna *et al.*, *Nucleic Acids Res.* 31:3497-3500, 2003.

[0100] The Cas9 nuclease sequence can be a mutated sequence. For example, the Cas9 nuclease can be mutated in the conserved HNH and RuvC domains, which are involved in strand specific cleavage. For example, an aspartate-to-alanine (D10A) mutation in the RuvC catalytic domain allows the Cas9 nickase mutant (Cas9n) to nick rather than cleave DNA to yield single-stranded breaks, and the subsequent preferential repair through HDR can potentially decrease the frequency of unwanted indel mutations from off-target double-stranded breaks.

[0101] *Guide RNA:* A gRNA includes a mature crRNA that contains about 20 base pairs (bp) of unique target sequence (called spacer) and a trans-activated small RNA (tracrRNA) that serves as a guide for ribonuclease III-aided processing of pre-crRNA. The crRNA:tracrRNA duplex directs Cas9 to target DNA via complementary base pairing between the spacer on the crRNA and the complementary sequence (called protospacer) on the target DNA. Cas9 recognizes a trinucleotide (NGG) protospacer adjacent motif (PAM) to specify the cut site (the 3rd nucleotide from PAM). In the present invention, the crRNA and tracrRNA can be expressed separately or engineered into an artificial fusion gRNA via a synthetic stem loop (AGAAAU) to mimic the natural crRNA/tracrRNA duplex. Such gRNA can be synthesized or in vitro transcribed for direct RNA transfection or expressed from U6 or H1-promoted RNA expression vector.

[0102] In the compositions of the present invention, each gRNA includes a sequence that is complementary to a target sequence in a retrovirus. The exemplary target retrovirus is HIV, but the compositions of the present invention are also useful for targeting other retroviruses, such as HIV-2 and simian immunodeficiency virus (SIV)-1.

[0103] Some of the exemplary gRNAs of the present invention are complimentary to target sequences in the long terminal repeat (LTR) regions of HIV. The LTRs are subdivided into U3, R and U5 regions. LTRs contain all of the required signals for gene expression, and are involved in the integration of a provirus into the genome of a host cell. For example, the basal or core promoter, a core enhancer and a modulatory region is found within U3 while the transactivation response element is found within R. In HIV-1, the U5 region includes several sub-regions, for example, TAR or trans-acting responsive element, which is involved in transcriptional activation; Poly A, which is involved in dimerization and genome packaging; PBS or primer binding site; Psi or the packaging signal; DIS or dimer initiation site.

[0104] Accordingly, in some embodiments a gRNA target sequence comprises one or more target sequences in an LTR region of an HIV proviral DNA and one or more targets in a structural gene and/or non-structural gene of the HIV proviral DNA. In other embodiments, a gRNA target sequence comprises one or more target sequences in an LTR region of an HIV proviral DNA and one or more targets in a structural gene. In another embodiment, a gRNA target sequence comprises one or more target sequences in an LTR region of an HIV proviral DNA and one or more targets in a non-structural gene of the HIV proviral DNA. In yet another embodiment, a gRNA target sequence comprises one or more target sequences in an HIV proviral a structural gene and one or more targets in a non-structural gene of the HIV proviral DNA. In yet another embodiment, a gRNA target sequence comprises one or more target sequences in an HIV proviral a non-coding gene and one or more targets in a coding gene of the HIV proviral DNA. In yet another embodiment a gRNA target nucleic acid sequence comprises one or more target nucleic acid sequences in a first gene and one or more target nucleic acid sequences in a second gene; or, one or more target nucleic acid sequences in a first gene and one or more target nucleic acid sequences in a third gene; or, one or more target nucleic acid sequences in a first gene and one or more target nucleic acid sequences in a second gene and one or more target nucleic acid sequences in a third gene; or, one or more target nucleic acid sequences in a second gene and one or more target nucleic acid sequences in a third gene or fourth gene; or, any combinations thereof. As can be seen, any combination

of target nucleic acid sequences can be used and are only limited by the imagination of one of ordinary skill in the art.

[0105] In certain embodiments, target sequences comprise sequences within the U3, R, and U5 regions of the LTR. In certain embodiments the target sequences comprise one or more sequences from: LTR 1, LTR 2, LTR 3, LTR A, LTR B, LTR B', LTR C, LTR D, LTR E, LTR F, LTR G, LTR H, LTR I, LTR J, LTR K, LTR L, LTR M, LTR N, LTR O, LTR P, LTR Q, LTR R, LTR S, AND LTR T. The compositions of the present invention include these exemplary gRNAs, but are not limited to them, and can include gRNAs complimentary to any suitable target site in the HIV LTRs.

[0106] Some of the exemplary gRNAs of the present invention target sequences in the protein coding genome of HIV. Sequences within the gene encoding the structural protein gag were found to be useful target sequences. gRNAs complementary to these target sequences include Gag A, Gag B, Gag C, and Gag D. Useful target sequences were also found within the gene encoding the structural protein pol. gRNAs complementary to these target sequences include Pol A and Pol B.

[0107] Examples of guide RNAs are shown in Tables 1-5. Accordingly, the compositions of the present invention include these exemplary gRNAs, but are not limited to them, and can include gRNAs complimentary to any suitable target site in the protein coding genes of HIV, including but not limited to those encoding the structural protein tat, and the accessory proteins vif, nef (negative factor) vpu (Virus protein U), vpr, and tev.

[0108] Guide RNA sequences according to the present invention can be sense or anti-sense sequences. The guide RNA sequence generally includes a proto-spacer adjacent motif (PAM). The sequence of the PAM can vary depending upon the specificity requirements of the CRISPR endonuclease used. In the CRISPR-Cas system derived from *S. pyogenes*, the target DNA typically immediately precedes a 5'-NGG proto-spacer adjacent motif (PAM). Thus, for the *S. pyogenes* Cas9, the PAM sequence can be AGG, TGG, CGG or GGG. Other Cas9 orthologs may have different PAM specificities. For example, Cas9 from *S. thermophilus* requires 5'-NNAGAA for CRISPR 1 and 5'-NGGNG for CRISPR3) and *Neisseria meningitidis* requires 5'-NNNNGATT). The specific sequence of the guide RNA may vary, but, regardless of the sequence, useful guide RNA sequences will be those that minimize off-target effects while achieving high efficiency and complete ablation of the genomically integrated retrovirus, e.g. HIV. The length of the guide RNA sequence can vary

from about 20 to about 60 or more nucleotides, for example about 20, about 21, about 22, about 23, about 24, about 25, about 26, about 27, about 28, about 29, about 30, about 31, about 32, about 33, about 34, about 35, about 36, about 37, about 38, about 39, about 40, about 45, about 50, about 55, about 60 or more nucleotides. Useful selection methods identify regions having extremely low homology between the foreign viral genome and host cellular genome including endogenous retroviral DNA, include bioinformatic screening using 12-bp+NGG target-selection criteria to exclude off-target human transcriptome or (even rarely) untranslated-genomic sites; avoiding transcription factor binding sites within the HIV LTR promoter (potentially conserved in the host genome); and WGS, Sanger sequencing and SURVEYOR assay, to identify and exclude potential off-target effects.

[0109] The guide RNA sequence can be configured as a single sequence or as a combination of one or more different sequences, *e.g.*, a multiplex configuration. Multiplex configurations can include combinations of two, three, four, five, six, seven, eight, nine, ten, or more different guide RNAs. Combinations of gRNAs are especially effective when expressed in multiplex fashion, that is, simultaneously in the same cell. In many cases, the combinations produced excision of the HIV provirus extending between the target sites. The excisions are attributable to deletions of sequences between the cleavages induced by Cas9 at each of the multiple target sites. These combinations pairs of gRNAs, with one member being complementary to a target site in an LTR of the retrovirus, and the other member being complementary to a gRNA complementary to a target site in a structural gene of the retrovirus. Exemplary effective combinations include Gag D combined with one of LTR 1, LTR 2, LTR 3, LTR A, LTR B, LTR C, LTR D, LTR E, LTR F, LTR G; LTR H, LTR I, LTR J, LTR K, LTR L, LTR M; LTR N, LTR O, LTR P, LTR Q, LTR R, LTR S, or LTR T. Exemplary effective combinations also include LTR 3 combined with one of LTR-1, Gag A; Gag B; Gag C, Gag D, Pol A, or Pol B. In certain embodiments, a gRNA sequence has at least a 75% sequence identity to complementary target nucleic acid sequences encoding T LTR 1, LTR 2, LTR 3, LTR A, LTR B, LTR C, LTR D, LTR E, LTR F, LTR G; LTR H, LTR I, LTR J, LTR K, LTR L, LTR M; LTR N, LTR O, LTR P, LTR Q, LTR R, LTR S, or LTR T. The compositions of present invention are not limited to these combinations, but include any suitable combination of gRNAs complimentary to two or more different target sites in the HIV provirus.

[0110] In certain embodiments, a target nucleic acid sequence comprises one or more nucleic acid sequences in coding and non-coding nucleic acid sequences of the retroviral

genome. The target nucleic acid sequence can be located within a sequence encoding structural proteins, non-structural proteins or combinations thereof. The sequences encoding structural proteins comprise nucleic acid sequences encoding: Gag, Gag-Pol precursor, Pro (protease), Reverse Transcriptase (RT), integrase (In), Env or combinations thereof. The sequences encoding non-structural proteins comprise nucleic acid sequences encoding: regulatory proteins e.g. Tat, Rev, accessory proteins, e.g. Nef, Vpr, Vpu, Vif or combinations thereof.

[0111] In certain embodiments, a gRNA sequence has at least a 75% sequence identity to complementary target nucleic acid sequences encoding Gag, Gag-Pol precursor, Pro, Reverse Transcriptase (RT), integrase (In), Env, Tat, Rev, Nef, Vpr, Vpu, Vif or combinations thereof.

[0112] In certain embodiments, a gRNA sequence is complementary to target nucleic acid sequences encoding Gag, Gag-Pol precursor, Pro, Reverse Transcriptase (RT), integrase (In), Env, Tat, Rev, Nef, Vpr, Vpu, Vif or combinations thereof.

[0113] In certain embodiments, gRNAs in single and multiplex configurations target the retroviral genome as well as the genes encoding receptors used by the virus to infect a cell, e.g. in the case of HIV, the receptor can be CCR5.

[0114] In some embodiments, the one or more isolated nucleic acid sequences are encoded by two or more constructs with one member directed toward a first retroviral target sequence, and the other member toward a second retroviral target sequence excises or eradicates the retroviral genome from an infected cell. Accordingly, the invention features compositions for use in inactivating a proviral DNA integrated into a host cell, including an isolated nucleic acid sequence encoding a CRISPR-associated endonuclease and one or more isolated nucleic acid sequences encoding one or more gRNAs complementary to a target sequence in HIV or another retrovirus. A second isolated nucleic acid sequence encoding a CRISPR-associated endonuclease and one or more isolated nucleic acid sequences encoding one or more gRNAs complementary to a target sequence encoding a receptor used by a virus to infect a cell. The isolated nucleic acid can include one gRNA, two gRNAs, three gRNAs etc. Furthermore, the isolated nucleic acid can include one or more gRNAs complementary to target sequences in the retrovirus and a second isolated nucleic acid can include one or more gRNAs complementary to target sequences encoding receptors used by the virus to infect a cell. Alternatively each isolated nucleic acid can include at least one gRNA

complementary to a target virus sequence and at least one a gRNA complementary to target sequences encoding receptors used by the virus to infect a cell. One of ordinary skill in the art would only be limited by their imagination with respect to the various combinations of gRNAs.

[0115] *Modified or Mutated Nucleic Acid Sequences:* In some embodiments, any of the nucleic acid sequences may be modified or derived from a native nucleic acid sequence, for example, by introduction of mutations, deletions, substitutions, modification of nucleobases, backbones and the like. The nucleic acid sequences include the vectors, gene-editing agents, gRNAs, etc. Examples of some modified nucleic acid sequences envisioned for this invention include those comprising modified backbones, for example, phosphorothioates, phosphotriesters, methyl phosphonates, short chain alkyl or cycloalkyl intersugar linkages or short chain heteroatomic or heterocyclic intersugar linkages. In some embodiments, modified oligonucleotides comprise those with phosphorothioate backbones and those with heteroatom backbones, $\text{CH}_2\text{--NH--O--CH}_2$, $\text{CH}_2\text{--N(CH}_3\text{)--O--CH}_2$ [known as a methylene(methylimino) or MMI backbone], $\text{CH}_2\text{--O--N(CH}_3\text{)--CH}_2$, $\text{CH}_2\text{--N(CH}_3\text{)--N(CH}_3\text{)--CH}_2$ and $\text{O--N(CH}_3\text{)--CH}_2\text{--CH}_2$ backbones, wherein the native phosphodiester backbone is represented as O--P--O--CH_2). The amide backbones disclosed by De Mesmaeker *et al. Acc. Chem. Res.* 1995, 28:366-374) are also embodied herein. In some embodiments, the nucleic acid sequences having morpholino backbone structures (Summerton and Weller, U.S. Pat. No. 5,034,506), peptide nucleic acid (PNA) backbone wherein the phosphodiester backbone of the oligonucleotide is replaced with a polyamide backbone, the nucleobases being bound directly or indirectly to the aza nitrogen atoms of the polyamide backbone (Nielsen *et al. Science* 1991, 254, 1497). The nucleic acid sequences may also comprise one or more substituted sugar moieties. The nucleic acid sequences may also have sugar mimetics such as cyclobutyls in place of the pentofuranosyl group.

[0116] The nucleic acid sequences may also include, additionally or alternatively, nucleobase (often referred to in the art simply as “base”) modifications or substitutions. As used herein, “unmodified” or “natural” nucleobases include adenine (A), guanine (G), thymine (T), cytosine (C) and uracil (U). Modified nucleobases include nucleobases found only infrequently or transiently in natural nucleic acids, *e.g.*, hypoxanthine, 6-methyladenine, 5-Me pyrimidines, particularly 5-methylcytosine (also referred to as 5-methyl-2’ deoxycytosine and often referred to in the art as 5-Me-C), 5-hydroxymethylcytosine (HMC), glycosyl HMC and gentobiosyl HMC, as well as synthetic nucleobases, *e.g.*, 2-aminoadenine,

2-(methylamino)adenine, 2-(imidazolylalkyl)adenine, 2-(aminoalkylamino)adenine or other heterosubstituted alkyladenines, 2-thiouracil, 2-thiothymine, 5-bromouracil, 5-hydroxymethyluracil, 8-azaguanine, 7-deazaguanine, N⁶ (6-aminohexyl)adenine and 2,6-diaminopurine. Kornberg, A., *DNA Replication*, W. H. Freeman & Co., San Francisco, 1980, pp75-77; Gebeyehu, G., *et al. Nucl. Acids Res.* 1987, 15:4513). A “universal” base known in the art, *e.g.*, inosine may be included. 5-Me-C substitutions have been shown to increase nucleic acid duplex stability by 0.6-1.2°C. (Sanghvi, Y. S., in Crooke, S. T. and Lebleu, B., eds., *Antisense Research and Applications*, CRC Press, Boca Raton, 1993, pp. 276-278).

[0117] Another modification of the nucleic acid sequences of the invention involves chemically linking to the nucleic acid sequences one or more moieties or conjugates which enhance the activity or cellular uptake of the oligonucleotide. Such moieties include but are not limited to lipid moieties such as a cholesterol moiety, a cholesteryl moiety (Letsinger *et al.*, *Proc. Natl. Acad. Sci. USA* 1989, 86, 6553), cholic acid (Manoharan *et al. Bioorg. Med. Chem. Lett.* 1994, 4, 1053), a thioether, *e.g.*, hexyl-S-tritylthiol (Manoharan *et al. Ann. N.Y. Acad. Sci.* 1992, 660, 306; Manoharan *et al. Bioorg. Med. Chem. Lett.* 1993, 3, 2765), a thiocholesterol (Oberhauser *et al.*, *Nucl. Acids Res.* 1992, 20, 533), an aliphatic chain, *e.g.*, dodecandiol or undecyl residues (Saison-Behmoaras *et al. EMBO J.* 1991, 10, 111; Kabanov *et al. FEBS Lett.* 1990, 259, 327; Svinarchuk *et al. Biochimie* 1993, 75, 49), a phospholipid, *e.g.*, di-hexadecyl-rac-glycerol or triethylammonium 1,2-di-O-hexadecyl-rac-glycero-3-H-phosphonate (Manoharan *et al. Tetrahedron Lett.* 1995, 36, 3651; Shea *et al. Nucl. Acids Res.* 1990, 18, 3777), a polyamine or a polyethylene glycol chain (Manoharan *et al. Nucleosides & Nucleotides* 1995, 14, 969), or adamantane acetic acid (Manoharan *et al. Tetrahedron Lett.* 1995, 36, 3651). It is not necessary for all positions in a given nucleic acid sequence to be uniformly modified, and in fact more than one of the aforementioned modifications may be incorporated in a single nucleic acid sequence or even at within a single nucleoside within a nucleic acid sequence.

[0118] In some embodiments, the RNA molecules *e.g.* crRNA, tracrRNA, gRNA are engineered to comprise one or more modified nucleobases. For example, known modifications of RNA molecules can be found, for example, in Genes VI, Chapter 9 (“Interpreting the Genetic Code”), Lewis, ed. (1997, Oxford University Press, New York), and Modification and Editing of RNA, Grosjean and Benne, eds. (1998, ASM Press, Washington DC). Modified RNA components include the following: 2'-O-methylcytidine;

N⁴-methylcytidine; N⁴-2'-O-dimethylcytidine; N⁴- acetylcytidine; 5-methylcytidine; 5,2'-O-dimethylcytidine; 5-hydroxymethylcytidine; 5- formylcytidine; 2'-O-methyl-5-formaylcytidine; 3-methylcytidine; 2-thiocytidine; lysidine; 2'-O- methyluridine; 2-thiouridine; 2-thio-2'-O-methyluridine; 3,2'-O-dimethyluridine; 3-(3-amino-3-carboxypropyl)uridine; 4-thiouridine; ribosylthymine; 5,2'-O-dimethyluridine; 5-methyl-2-thiouridine; 5-hydroxyuridine; 5-methoxyuridine; uridine 5-oxyacetic acid; uridine 5-oxyacetic acid methyl ester; 5-carboxymethyluridine; 5-methoxycarbonylmethyluridine; 5-methoxycarbonylmethyl-2'-O-methyluridine; 5-methoxycarbonylmethyl-2'-thiouridine; 5-carbamoylmethyluridine; 5-carbamoylmethyl-2'-O-methyluridine; 5-(carboxyhydroxymethyl)uridine; 5-(carboxyhydroxymethyl) uridinemethyl ester; 5-aminomethyl-2-thiouridine; 5-methylaminomethyluridine; 5-methylaminomethyl-2-thiouridine; 5-methylaminomethyl-2-selenouridine; 5-carboxymethylaminomethyluridine; 5-carboxymethylaminomethyl-2'-O-methyl- uridine; 5-carboxymethylaminomethyl-2-thiouridine; dihydrouridine; dihydroribosylthymine; 2'-methyladenosine; 2-methyladenosine; N⁶Nmethyladenosine; N⁶, N⁶-dimethyladenosine; N⁶,2'-O-trimethyladenosine; 2 methylthio-N⁶Nisopentenyladenosine; N⁶-(cis-hydroxyisopentenyl)-adenosine; 2-methylthio-N⁶-(cis--hydroxyisopentenyl)-adenosine; N⁶-glycincylcarbamoyl)adenosine; N⁶ threonylcarbamoyl adenosine; N⁶-methyl-N⁶-threonylcarbamoyl adenosine; 2-methylthio-N⁶-methyl-N⁶-threonylcarbamoyl adenosine; N⁶-hydroxynorvalylcarbamoyl adenosine; 2-methylthio-N⁶-hydroxynorvalylcarbamoyl adenosine; 2'-O-ribosyladenosine (phosphate); inosine; 2'-O-methyl inosine; 1-methyl inosine; 1,2'-O-dimethyl inosine; 2'-O-methyl guanosine; 1-methyl guanosine; N²-methyl guanosine; N², N²-dimethyl guanosine; N², 2'-O-dimethyl guanosine; N², N², 2'-O-trimethyl guanosine; 2'-O-ribosyl guanosine (phosphate); 7-methyl guanosine; N², 7-dimethyl guanosine; N², N²;7-trimethyl guanosine; wyosine; methylwyosine; under-modified hydroxywybutosine; wybutosine; hydroxywybutosine; peroxywybutosine; queuosine; epoxyqueuosine; galactosyl-queuosine; mannosyl-queuosine; 7-cyano-7-deazaguanosine; arachaeosine [also called 7-formamido-7-deazaguanosine]; and 7-aminomethyl-7-deazaguanosine.

[0119] The isolated nucleic acid molecules of the present invention can be produced by standard techniques. For example, polymerase chain reaction (PCR) techniques can be used to obtain an isolated nucleic acid containing a nucleotide sequence described herein. Various PCR methods are described in, for example, *PCR Primer: A Laboratory Manual*, Dieffenbach and Dveksler, eds., Cold Spring Harbor Laboratory Press, 1995. Generally,

sequence information from the ends of the region of interest or beyond is employed to design oligonucleotide primers that are identical or similar in sequence to opposite strands of the template to be amplified. Various PCR strategies also are available by which site-specific nucleotide sequence modifications can be introduced into a template nucleic acid. Isolated nucleic acids also can be chemically synthesized, either as a single nucleic acid molecule (*e.g.*, using automated DNA synthesis in the 3' to 5' direction using phosphoramidite technology) or as a series of oligonucleotides. For example, one or more pairs of long oligonucleotides (*e.g.*, >50-100 nucleotides) can be synthesized that contain the desired sequence, with each pair containing a short segment of complementarity (*e.g.*, about 15 nucleotides) such that a duplex is formed when the oligonucleotide pair is annealed. DNA polymerase is used to extend the oligonucleotides, resulting in a single, double-stranded nucleic acid molecule per oligonucleotide pair, which then can be ligated into a vector.

[0120] *Recombinant Constructs and Delivery Vehicles*

[0121] Recombinant constructs are also provided herein and can be used to transform cells in order to express the isolated nucleic acid sequences embodied herein. A recombinant nucleic acid construct comprises promoter operably linked to a regulatory region suitable for expressing at least one tRNA, ribozyme, single guide RNA (sgRNA), gene editing agent or combinations thereof.

[0122] It will be appreciated that a number of nucleic acids can encode a polypeptide having a particular amino acid sequence. The degeneracy of the genetic code is well known in the art. For many amino acids, there is more than one nucleotide triplet that serves as the codon for the amino acid. For example, codons in the coding sequence for Cas9 can be modified such that optimal expression in a particular organism is obtained, using appropriate codon bias tables for that organism.

[0123] Nucleic acids as described herein may be contained in vectors. Vectors can include, for example, origins of replication, scaffold attachment regions (SARs), and/or markers. A marker gene can confer a selectable phenotype on a host cell. For example, a marker can confer biocide resistance, such as resistance to an antibiotic (*e.g.*, kanamycin, G418, bleomycin, or hygromycin). An expression vector can include a tag sequence designed to facilitate manipulation or detection (*e.g.*, purification or localization) of the expressed polypeptide. Tag sequences, such as green fluorescent protein (GFP), glutathione S-transferase (GST), polyhistidine, c-myc, hemagglutinin, or FLAGTM tag (Kodak, New

Haven, CT) sequences typically are expressed as a fusion with the encoded polypeptide. Such tags can be inserted anywhere within the polypeptide, including at either the carboxyl or amino terminus.

[0124] Additional expression vectors also can include, for example, segments of chromosomal, non-chromosomal and synthetic DNA sequences. Suitable vectors include derivatives of SV40 and known bacterial plasmids, e.g., *E. coli* plasmids col El, pCR1, pBR322, pMal-C2, pET, pGEX, pMB9 and their derivatives, plasmids such as RP4; phage DNAs, e.g., the numerous derivatives of phage 1, e.g., NM989, and other phage DNA, e.g., M13 and filamentous single stranded phage DNA; yeast plasmids such as the 2 μ plasmid or derivatives thereof, vectors useful in eukaryotic cells, such as vectors useful in insect or mammalian cells; vectors derived from combinations of plasmids and phage DNAs, such as plasmids that have been modified to employ phage DNA or other expression control sequences.

[0125] Several delivery methods may be utilized for *in vitro* (cell cultures) and *in vivo* (animals and patients) systems. In one embodiment, a lentiviral gene delivery system may be utilized. Such a system offers stable, long term presence of the gene in dividing and non-dividing cells with broad tropism and the capacity for large DNA inserts. (Dull *et al*, *J Virol*, 72:8463-8471 1998). In an embodiment, adeno-associated virus (AAV) may be utilized as a delivery method. AAV is a non-pathogenic, single-stranded DNA virus that has been actively employed in recent years for delivering therapeutic gene in *in vitro* and *in vivo* systems (Choi *et al*, *Curr Gene Ther*, 5:299-310, 2005). An example non-viral delivery method may utilize nanoparticle technology. This platform has demonstrated utility as a pharmaceutical *in vivo*. Nanotechnology has improved transcytosis of drugs across tight epithelial and endothelial barriers. It offers targeted delivery of its payload to cells and tissues in a specific manner (Allen and Cullis, *Science*, 303:1818-1822, 1998).

[0126] The vector can also include a regulatory region. The term “regulatory region” refers to nucleotide sequences that influence transcription or translation initiation and rate, and stability and/or mobility of a transcription or translation product. Regulatory regions include, without limitation, promoter sequences, enhancer sequences, response elements, protein recognition sites, inducible elements, protein binding sequences, 5' and 3' untranslated regions (UTRs), transcriptional start sites, termination sequences, polyadenylation sequences, nuclear localization signals, and introns.

[0127] The term “operably linked” refers to positioning of a regulatory region and a sequence to be transcribed in a nucleic acid so as to influence transcription or translation of such a sequence. For example, to bring a coding sequence under the control of a promoter, the translation initiation site of the translational reading frame of the polypeptide is typically positioned between one and about fifty nucleotides downstream of the promoter. A promoter can, however, be positioned as much as about 5,000 nucleotides upstream of the translation initiation site or about 2,000 nucleotides upstream of the transcription start site. A promoter typically comprises at least a core (basal) promoter. A promoter also may include at least one control element, such as an enhancer sequence, an upstream element or an upstream activation region (UAR). The choice of promoters to be included depends upon several factors, including, but not limited to, efficiency, selectability, inducibility, desired expression level, and cell- or tissue-preferential expression. It is a routine matter for one of skill in the art to modulate the expression of a coding sequence by appropriately selecting and positioning promoters and other regulatory regions relative to the coding sequence.

[0128] Vectors include, for example, viral vectors (such as adenoviruses Ad, AAV, lentivirus, and vesicular stomatitis virus (VSV) and retroviruses), liposomes and other lipid-containing complexes, and other macromolecular complexes capable of mediating delivery of a polynucleotide to a host cell. Vectors can also comprise other components or functionalities that further modulate gene delivery and/or gene expression, or that otherwise provide beneficial properties to the targeted cells. As described and illustrated in more detail below, such other components include, for example, components that influence binding or targeting to cells (including components that mediate cell-type or tissue-specific binding); components that influence uptake of the vector nucleic acid by the cell; components that influence localization of the polynucleotide within the cell after uptake (such as agents mediating nuclear localization); and components that influence expression of the polynucleotide. Such components also might include markers, such as detectable and/or selectable markers that can be used to detect or select for cells that have taken up and are expressing the nucleic acid delivered by the vector. Such components can be provided as a natural feature of the vector (such as the use of certain viral vectors which have components or functionalities mediating binding and uptake), or vectors can be modified to provide such functionalities. Other vectors include those described by Chen *et al*; *BioTechniques*, 34: 167-171 (2003). A large variety of such vectors is known in the art and are generally available. A “recombinant viral vector” refers to a viral vector comprising one or more heterologous

gene products or sequences. Since many viral vectors exhibit size-constraints associated with packaging, the heterologous gene products or sequences are typically introduced by replacing one or more portions of the viral genome. Such viruses may become replication-defective, requiring the deleted function(s) to be provided in trans during viral replication and encapsidation (by using, e.g., a helper virus or a packaging cell line carrying gene products necessary for replication and/or encapsidation). Modified viral vectors in which a polynucleotide to be delivered is carried on the outside of the viral particle have also been described (see, e.g., Curiel, D T, *et al. PNAS* 88: 8850-8854, 1991).

[0129] Additional vectors include viral vectors, fusion proteins and chemical conjugates. Retroviral vectors include Moloney murine leukemia viruses and HIV-based viruses. One HIV based viral vector comprises at least two vectors wherein the gag and pol genes are from an HIV genome and the env gene is from another virus. DNA viral vectors include pox vectors such as orthopox or avipox vectors, herpesvirus vectors such as a herpes simplex I virus (HSV) vector [Geller, A.I. *et al., J. Neurochem*, 64: 487 (1995); Lim, F., *et al.*, in *DNA Cloning: Mammalian Systems*, D. Glover, Ed. (Oxford Univ. Press, Oxford England) (1995); Geller, A.I. *et al., Proc Natl. Acad. Sci.: U.S.A.*:90 7603 (1993); Geller, A.I., *et al., Proc Natl. Acad. Sci USA*: 87:1149 (1990)], Adenovirus Vectors [LeGal LaSalle *et al., Science*, 259:988 (1993); Davidson, *et al., Nat. Genet.* 3: 219 (1993); Yang, *et al., J. Virol.* 69: 2004 (1995)] and Adeno-associated Virus Vectors [Kaplitt, M.G., *et al., Nat. Genet.* 8:148 (1994)].

[0130] The polynucleotides disclosed herein may be used with a microdelivery vehicle such as cationic liposomes and adenoviral vectors. For a review of the procedures for liposome preparation, targeting and delivery of contents, see Mannino and Gould-Fogerite, *BioTechniques*, 6:682 (1988). See also, Felgner and Holm, *Bethesda Res. Lab. Focus*, 11(2):21 (1989) and Maurer, R.A., *Bethesda Res. Lab. Focus*, 11(2):25 (1989).

[0131] Replication-defective recombinant adenoviral vectors, can be produced in accordance with known techniques. See, Quantin, *et al., Proc. Natl. Acad. Sci. USA*, 89:2581-2584 (1992); Stratford-Perricadet, *et al., J. Clin. Invest.*, 90:626-630 (1992); and Rosenfeld, *et al., Cell*, 68:143-155 (1992).

[0132] Another delivery method is to use single stranded DNA producing vectors which can produce the expressed products intracellularly. See for example, Chen *et al.*,

BioTechniques, 34: 167-171 (2003), which is incorporated herein, by reference, in its entirety.

[0133] The polynucleotides disclosed herein may be used with a microdelivery vehicle such as cationic liposomes and adenoviral vectors. For a review of the procedures for liposome preparation, targeting and delivery of contents, see Mannino and Gould-Fogerite, *BioTechniques*, 6:682 (1988). See also, Felgner and Holm, *Bethesda Res. Lab. Focus*, 11(2):21 (1989) and Maurer, R. A., *Bethesda Res. Lab. Focus*, 11(2):25 (1989).

[0134] In certain embodiments of the invention, non-viral vectors may be used to effectuate transfection. Methods of non-viral delivery of nucleic acids include lipofection, nucleofection, microinjection, biolistics, virosomes, liposomes, immunoliposomes, polycation or lipid:nucleic acid conjugates, naked DNA, artificial virions, and agent-enhanced uptake of DNA. Lipofection is described in e.g., U.S. Pat. Nos. 5,049,386, 4,946,787; and 4,897,355) and lipofection reagents are sold commercially (e.g., Transfectam and Lipofectin). Cationic and neutral lipids that are suitable for efficient receptor-recognition lipofection of polynucleotides include those described in U.S. Pat. No. 7,166,298 to Jessee or U.S. Pat. No. 6,890,554 to Jesse, the contents of each of which are incorporated by reference. Delivery can be to cells (e.g. *in vitro* or *ex vivo* administration) or target tissues (e.g. *in vivo* administration).

[0135] Synthetic vectors are typically based on cationic lipids or polymers which can complex with negatively charged nucleic acids to form particles with a diameter in the order of 100 nm. The complex protects nucleic acid from degradation by nuclease. Moreover, cellular and local delivery strategies have to deal with the need for internalization, release, and distribution in the proper subcellular compartment. Systemic delivery strategies encounter additional hurdles, for example, strong interaction of cationic delivery vehicles with blood components, uptake by the reticuloendothelial system, kidney filtration, toxicity and targeting ability of the carriers to the cells of interest. Modifying the surfaces of the cationic non-virals can minimize their interaction with blood components, reduce reticuloendothelial system uptake, decrease their toxicity and increase their binding affinity with the target cells. Binding of plasma proteins (also termed opsonization) is the primary mechanism for RES to recognize the circulating nanoparticles. For example, macrophages, such as the Kupffer cells in the liver, recognize the opsonized nanoparticles via the scavenger receptor.

[0136] The anti-retroviral agents and/or the isolated nucleic acid sequences of the invention can be delivered to an appropriate cell of a subject. This can be achieved by, for example, the use of a polymeric, biodegradable microparticle or microcapsule delivery vehicle, sized to optimize phagocytosis by phagocytic cells such as macrophages. For example, PLGA (poly-lacto-co-glycolide) microparticles approximately 1-10 μm in diameter can be used. The polynucleotide is encapsulated in these microparticles, which are taken up by macrophages and gradually biodegraded within the cell, thereby releasing the polynucleotide. Once released, the DNA is expressed within the cell. A second type of microparticle is intended not to be taken up directly by cells, but rather to serve primarily as a slow-release reservoir of nucleic acid that is taken up by cells only upon release from the micro-particle through biodegradation. These polymeric particles should therefore be large enough to preclude phagocytosis (i.e., larger than 5 μm and preferably larger than 20 μm). Another way to achieve uptake of the nucleic acid is using liposomes, prepared by standard methods. The nucleic acids can be incorporated alone into these delivery vehicles or co-incorporated with tissue-specific antibodies, for example antibodies that target cell types that are commonly latently infected reservoirs of HIV infections. Alternatively, one can prepare a molecular complex composed of a plasmid or other vector attached to poly-L-lysine by electrostatic or covalent forces. Poly-L-lysine binds to a ligand that can bind to a receptor on target cells. Delivery of “naked DNA” (i.e., without a delivery vehicle) to an intramuscular, intradermal, or subcutaneous site, is another means to achieve *in vivo* expression. In the relevant polynucleotides (e.g., expression vectors) the nucleic acid sequence encoding an isolated nucleic acid sequence comprising a sequence encoding CRISPR/Cas and/or a guide RNA complementary to a target sequence of HIV, as described above.

[0137] In some embodiments, the compositions of the invention can be formulated as a nanoparticle, for example, nanoparticles comprised of a core of high molecular weight linear polyethylenimine (LPEI) complexed with DNA and surrounded by a shell of polyethyleneglycol modified (PEGylated) low molecular weight LPEI. In some embodiments, the compositions can be formulated as a nanoparticle encapsulating the compositions embodied herein. L-PEI has been used to efficiently deliver genes *in vivo* into a wide range of organs such as lung, brain, pancreas, retina, bladder as well as tumor. L-PEI is able to efficiently condense, stabilize and deliver nucleic acids *in vitro* and *in vivo*.

[0138] In some embodiments, delivery of vectors can also be mediated by exosomes. Exosomes are lipid nanovesicles released by many cell types. They mediate intercellular

communication by transporting nucleic acids and proteins between cells. Exosomes contain RNAs, miRNAs, and proteins derived from the endocytic pathway. They may be taken up by target cells by endocytosis, fusion, or both. Exosomes can be harnessed to deliver nucleic acids to specific target cells.

[0139] The expression constructs of the present invention can also be delivered by means of nanoclews. Nanoclews are a cocoon-like DNA nanocomposites (Sun, *et al.*, *J. Am. Chem. Soc.* 2014, 136:14722–14725). They can be loaded with nucleic acids for uptake by target cells and release in target cell cytoplasm. Methods for constructing nanoclews, loading them, and designing release molecules can be found in Sun, *et al.* (Sun W, *et al.*, *J. Am. Chem. Soc.* 2014, 136:14722–14725; Sun W, *et al.*, *Angew. Chem. Int. Ed.* 2015: 12029–12033.)

[0140] The nucleic acids and vectors may also be applied to a surface of a device (e.g., a catheter) or contained within a pump, patch, or any other drug delivery device. The nucleic acids and vectors disclosed herein can be administered alone, or in a mixture, in the presence of a pharmaceutically acceptable excipient or carrier (e.g., physiological saline). The excipient or carrier is selected on the basis of the mode and route of administration. Suitable pharmaceutical carriers, as well as pharmaceutical necessities for use in pharmaceutical formulations, are described in Remington's Pharmaceutical Sciences (E. W. Martin), a well-known reference text in this field, and in the USP/NF (United States Pharmacopeia and the National Formulary).

[0141] In some embodiments of the invention, liposomes are used to effectuate transfection into a cell or tissue. The pharmacology of a liposomal formulation of nucleic acid is largely determined by the extent to which the nucleic acid is encapsulated inside the liposome bilayer. Encapsulated nucleic acid is protected from nuclease degradation, while those merely associated with the surface of the liposome is not protected. Encapsulated nucleic acid shares the extended circulation lifetime and biodistribution of the intact liposome, while those that are surface associated adopt the pharmacology of naked nucleic acid once they disassociate from the liposome. Nucleic acids may be entrapped within liposomes with conventional passive loading technologies, such as ethanol drop method (as in SALP), reverse-phase evaporation method, and ethanol dilution method (as in SNALP).

[0142] Liposomal delivery systems provide stable formulation, provide improved pharmacokinetics, and a degree of 'passive' or 'physiological' targeting to tissues.

Encapsulation of hydrophilic and hydrophobic materials, such as potential chemotherapy agents, are known. See for example U.S. Pat. No. 5,466,468 to Schneider, which discloses parenterally administrable liposome formulation comprising synthetic lipids; U.S. Pat. No. 5,580,571, to Hostetler *et al.* which discloses nucleoside analogues conjugated to phospholipids; U.S. Pat. No. 5,626,869 to Nyqvist, which discloses pharmaceutical compositions wherein the pharmaceutically active compound is heparin or a fragment thereof contained in a defined lipid system comprising at least one amphiphatic and polar lipid component and at least one nonpolar lipid component.

[0143] Liposomes and polymersomes can contain a plurality of solutions and compounds. In certain embodiments, the complexes of the invention are coupled to or encapsulated in polymersomes. As a class of artificial vesicles, polymersomes are tiny hollow spheres that enclose a solution, made using amphiphilic synthetic block copolymers to form the vesicle membrane. Common polymersomes contain an aqueous solution in their core and are useful for encapsulating and protecting sensitive molecules, such as drugs, enzymes, other proteins and peptides, and DNA and RNA fragments. The polymersome membrane provides a physical barrier that isolates the encapsulated material from external materials, such as those found in biological systems. Polymersomes can be generated from double emulsions by known techniques, see Lorenceau *et al.*, 2005, Generation of Polymersomes from Double-Emulsions, *Langmuir* 21(20):9183-6.

[0144] In some embodiments of the invention, non-viral vectors are modified to effectuate targeted delivery and transfection. PEGylation (i.e. modifying the surface with polyethyleneglycol) is the predominant method used to reduce the opsonization and aggregation of non-viral vectors and minimize the clearance by reticuloendothelial system, leading to a prolonged circulation lifetime after intravenous (i.v.) administration. PEGylated nanoparticles are therefore often referred as “stealth” nanoparticles. The nanoparticles that are not rapidly cleared from the circulation will have a chance to encounter infected cells.

[0145] In some embodiments of the invention, targeted controlled-release systems responding to the unique environments of tissues and external stimuli are utilized. Gold nanorods have strong absorption bands in the near-infrared region, and the absorbed light energy is then converted into heat by gold nanorods, the so-called “photothermal effect”. Because the near-infrared light can penetrate deeply into tissues, the surface of gold nanorod could be modified with nucleic acids for controlled release. When the modified gold nanorods are irradiated by near-infrared light, nucleic acids are released due to thermo-

denaturation induced by the photothermal effect. The amount of nucleic acids released is dependent upon the power and exposure time of light irradiation.

[0146] Regardless of whether compositions are administered as nucleic acids or polypeptides, they are formulated in such a way as to promote uptake by the mammalian cell. Useful vector systems and formulations are described above. In some embodiments the vector can deliver the compositions to a specific cell type. The invention is not so limited however, and other methods of DNA delivery such as chemical transfection, using, for example calcium phosphate, DEAE dextran, liposomes, lipoplexes, surfactants, and perfluoro chemical liquids are also contemplated, as are physical delivery methods, such as electroporation, micro injection, ballistic particles, and “gene gun” systems.

[0147] In other embodiments, the compositions comprise a cell which has been transformed or transfected with one or more vectors encoding the isolated nucleic acids embodied herein. In some embodiments, the methods of the invention can be applied *ex vivo*. That is, a subject's cells can be removed from the body and treated with the compositions in culture to excise, and the treated cells returned to the subject's body. The cell can be the subject's cells or they can be haplotype matched or a cell line. The cells can be irradiated to prevent replication. In some embodiments, the cells are human leukocyte antigen (HLA)-matched, autologous, cell lines, or combinations thereof. In other embodiments the cells can be a stem cell. For example, an embryonic stem cell or an artificial pluripotent stem cell (induced pluripotent stem cell (iPS cell)). Embryonic stem cells (ES cells) and artificial pluripotent stem cells (induced pluripotent stem cell, iPS cells) have been established from many animal species, including humans. These types of pluripotent stem cells would be the most useful source of cells for regenerative medicine because these cells are capable of differentiation into almost all of the organs by appropriate induction of their differentiation, with retaining their ability of actively dividing while maintaining their pluripotency. iPS cells, in particular, can be established from self-derived somatic cells, and therefore are not likely to cause ethical and social issues, in comparison with ES cells which are produced by destruction of embryos. Further, iPS cells, which are self-derived cell, make it possible to avoid rejection reactions, which are the biggest obstacle to regenerative medicine or transplantation therapy.

[0148] Transduced cells are prepared for reinfusion according to established methods. After a period of about 2-4 weeks in culture, the cells may number between 1×10^6 and 1×10^{10} . In this regard, the growth characteristics of cells vary from patient to patient and from

cell type to cell type. About 72 hours prior to reinfusion of the transduced cells, an aliquot is taken for analysis of phenotype, and percentage of cells expressing the therapeutic agent. For administration, cells of the present invention can be administered at a rate determined by the LD₅₀ of the cell type, and the side effects of the cell type at various concentrations, as applied to the mass and overall health of the patient. Administration can be accomplished via single or divided doses. Adult stem cells may also be mobilized using exogenously administered factors that stimulate their production and egress from tissues or spaces that may include, but are not restricted to, bone marrow or adipose tissues.

[0149] *Combination or Alternation Therapy*

[0150] Accordingly, the invention features compositions which include therapeutically effective amounts of at least one antiretroviral agent administered sequentially or alternately or in conjunction with a composition for inactivating a proviral DNA integrated into a host cell. This composition comprises an isolated nucleic acid sequence encoding a CRISPR-associated endonuclease and one or more isolated nucleic acid sequences encoding one or more gRNAs complementary to a target sequence in HIV or another retrovirus.

[0151] In one embodiment, the antiretroviral agent comprises viral entry inhibitors, reverse transcriptase inhibitors, protease inhibitors, and immune-based therapeutic agents.

[0152] For example, when used to treat or prevent HIV infection, the antiretroviral agent or its prodrug or pharmaceutically acceptable salt can be administered in combination or alternation with another anti-HIV agent and/or a gene-editing agent embodied herein. In general, in combination therapy, effective dosages of two or more agents are administered together, whereas during alternation therapy, an effective dosage of each agent is administered serially. The dosage will depend on absorption, inactivation and excretion rates of the drug, as well as other factors known to those of skill in the art. It is to be noted that dosage values will also vary with the severity of the condition to be alleviated. It is to be further understood that for any particular subject, specific dosage regimens and schedules should be adjusted over time according to the individual need and the professional judgment of the person administering or supervising the administration of the compositions.

[0153] Combination therapy may be administered as (a) a single pharmaceutical composition which comprises an antiretroviral agent as described herein, at least one gene editing agent as described herein, and a pharmaceutically acceptable excipient, diluent, or carrier; or (b) two separate pharmaceutical compositions comprising (i) a first composition

comprising an anti-retroviral agent as embodied herein and a pharmaceutically acceptable excipient, diluent, or carrier, and (ii) a second composition comprising at least one gene editing agents as embodied herein. The pharmaceutical compositions can be administered simultaneously or sequentially and in any order.

[0154] In use in treating or preventing viral disease, the antiretroviral(s) can be administered together with at least one gene editing agent as part of a unitary pharmaceutical composition. Alternatively, each can be administered apart from the other antiviral agents. In this embodiment, the antiretroviral(s) and the at least one at least one gene editing agent are administered substantially simultaneously, i.e. the compounds are administered at the same time or one after the other, so long as the compounds reach therapeutic levels for a period of time in the blood. In other embodiments, the antiretroviral agents are administered in one or more doses over a period of time followed by administration of the gene editing agents embodied herein.

[0155] The antiretroviral agents may be a nucleoside reverse transcriptase inhibitor, a nucleotide reverse transcriptase inhibitor, a non-nucleoside reverse transcriptase inhibitor, a protease inhibitor, an integrase inhibitor, a fusion inhibitor, a maturation inhibitor, or a combination thereof.

[0156] In certain embodiments, the at least one antiretroviral agent comprises: myristoylated dolutegravir, lamivudine, abacavir, rilpivirine or combinations thereof.

[0157] In certain embodiments, a composition comprises a therapeutically effective amount of a non-nucleoside reverse transcriptase inhibitor (NNRTI) and/or a nucleoside reverse transcriptase inhibitor (NRTI) ,and/or myristoylated dolutegravir, lamivudine, abacavir, rilpivirine analogs, variants or combinations thereof. In certain embodiments, an NNRTI comprises: etravirine, efavirenz, nevirapine, rilpivirine, delavirdine, or nevirapine. In embodiments, an NRTI comprises: lamivudine, zidovudine, emtricitabine, abacavir, zalcitabine, dideoxycytidine, azidothymidine, tenofovir disoproxil fumarate, didanosine (ddI EC), dideoxyinosine, stavudine, abacavir sulfate or combinations thereof.

[0158] Examples of nucleoside reverse transcriptase inhibitors include zidovudine, didanosine, stavudine, zalcitabine, abacavir, emtricitabine, and lamivudine. Examples of non-nucleoside reverse transcriptase inhibitors include efavirenz, nevirapine, and delaviradine. Examples of protease inhibitors include indinavir, ritonavir, saquinavir, lopinavir, and nelfinavir. Examples of a reverse transcriptase inhibitor, an integrase inhibitor, a fusion

inhibitor, and a maturation inhibitor are tenofovir, raltegravir, mariviroc, and bevirimat, respectively. In some aspects, the antiretroviral agents present in a nanoparticle include, ritonavir, lopinavir, and efavirenz, or efavirenz, abacavir, and lamivudine, or emtricitabine, tenofovir, and raltegravir.

[0159] In certain embodiments, the composition further comprises at least one or more protease inhibitors. In certain embodiments, a protease inhibitor comprises: amprenavir, tipranavir, indinavir, saquinavir mesylate, lopinavir and ritonavir (LPV/RTV), Fosamprenavir Calcium (FOS-APV), ritonavir, darunavir, atazanavir sulfate, nelfinavir mesylate or combinations thereof.

[0160] In certain embodiments, the compositions comprise an anti-retroviral agent, used in HAART, chemotherapeutic agents, activators of HIV transcription, e.g. PMA, TSA, and the like. Antiretroviral agents may include reverse transcriptase inhibitors (e.g., nucleoside/nucleotide reverse transcriptase inhibitors, zidovudine, emtricitabine, lamivudine and tenofovir; and non-nucleoside reverse transcriptase inhibitors such as efavirenz, nevirapine, rilpivirine); protease inhibitors, e.g., tipiravir, darunavir, indinavir; entry inhibitors, e.g., mariviroc; fusion inhibitors, e.g., enfuvirtide; or integrase inhibitors e.g., raltegravir, dolutegravir. Antiretroviral agents may also include multi-class combination agents for example, combinations of emtricitabine, efavirenz, and tenofovir; combinations of emtricitabine; rilpivirine, and tenofovir; or combinations of elvitegravir, cobicistat, emtricitabine and tenofovir.

[0161] In addition, one or more agents which alleviate any other symptoms that may be associated with the virus infection, e.g. fever, chills, headaches, secondary infections, can be administered in concert with, or as part of the pharmaceutical composition or at separate times. These agents comprise, without limitation, an anti-pyretic agent, anti-inflammatory agent, chemotherapeutic agent, or combinations thereof.

[0162] Some antiviral agents which can be used for combination therapy include agents that interfere with the ability of a virus to infiltrate a target cell. The virus must go through a sequence of steps to do this, beginning with binding to a specific “receptor” molecule on the surface of the host cell and ending with the virus “uncoating” inside the cell and releasing its contents. Viruses that have a lipid envelope must also fuse their envelope with the target cell, or with a vesicle that transports them into the cell, before they can uncoat.

[0163] There are two types of active agents which inhibit this stage of viral replication. One type includes agents which mimic the virus-associated protein (VAP) and bind to the cellular receptors, including VAP anti-idiotypic antibodies, natural ligands of the receptor and anti-receptor antibodies, receptor anti-idiotypic antibodies, extraneous receptor and synthetic receptor mimics. The other type includes agents which inhibit viral entry, for example, when the virus attaches to and enters the host cell. For example, a number of “entry-inhibiting” or “entry-blocking” drugs are being developed to fight HIV, which targets the immune system white blood cells known as “helper T cells”, and identifies these target cells through T-cell surface receptors designated “CRX4” and “CCR5”. Thus, CRX4 and CCR5 receptor inhibitors such as amantadine and rimantadine, can be used to inhibit viral infection, such as HIV.

[0164] Further antiviral agents that can be used in combination with the gene-editing agents embodied herein include agents which interfere with viral processes that synthesize virus components after a virus invades a cell. Representative agents include nucleotide and nucleoside analogues that look like the building blocks of RNA or DNA, but deactivate the enzymes that synthesize the RNA or DNA once the analogue is incorporated. Acyclovir is a nucleoside analogue, and is effective against herpes virus infections. Zidovudine (AZT), 3TC, FTC, and other nucleoside reverse transcriptase inhibitors (NRTI), as well as non-nucleoside reverse transcriptase inhibitors (NNRTI), can also be used. Integrase inhibitors can also be used.

[0165] Once a virus genome becomes operational in a host cell, it then generates messenger RNA (mRNA) molecules that direct the synthesis of viral proteins. Production of mRNA is initiated by proteins known as transcription factors, and certain active agents block attachment of transcription factors to viral DNA.

[0166] Other active agents include antisense oligonucleotides and ribozymes (enzymes which cut apart viral RNA or DNA at selected sites). HIV include protease enzymes, which cut viral protein chains apart so they can be assembled into their final configuration. Protease inhibitors are another type of antiviral agent that can be used in combination with the inhibitory compounds described herein. The final stage in the life cycle of a virus is the release of completed viruses from the host cell.

[0167] Still other active agents function by stimulating the patient's immune system. Interferons, including pegylated interferons, are representative compounds of this class.

[0168] In certain embodiments, the anti-viral or antiretroviral agent comprises therapeutically effective amounts of: antibodies, aptamers, adjuvants, anti-sense oligonucleotides, chemokines, cytokines, immune stimulating agents, immune modulating molecules, B-cell modulators, T-cell modulators, NK cell modulators, antigen presenting cell modulators, enzymes, siRNA's, interferon, ribavirin, protease inhibitors, anti-sense oligonucleotides, helicase inhibitors, polymerase inhibitors, helicase inhibitors, neuraminidase inhibitors, nucleoside reverse transcriptase inhibitors, non-nucleoside reverse transcriptase inhibitors, purine nucleosides, chemokine receptor antagonists, interleukins, vaccines or combinations thereof.

[0169] The immune-modulating molecules comprise, but are not limited to cytokines, lymphokines, T cell co-stimulatory ligands, etc. An immune-modulating molecule positively and/or negatively influences the humoral and/or cellular immune system, particularly its cellular and/or non-cellular components, its functions, and/or its interactions with other physiological systems. The immune-modulating molecule may be selected from the group comprising cytokines, chemokines, macrophage migration inhibitory factor (MIF; as described, *inter alia*, in Bernhagen (1998), *Mol Med* 76(3-4); 151-61 or Metz (1997), *Adv Immunol* 66, 197-223), T-cell receptors or soluble MHC molecules. Such immune-modulating effector molecules are well known in the art and are described, *inter alia*, in Paul, "Fundamental immunology", Raven Press, New York (1989). In particular, known cytokines and chemokines are described in Meager, "The Molecular Biology of Cytokines" (1998), John Wiley & Sons, Ltd., Chichester, West Sussex, England; (Bacon (1998). Cytokine Growth Factor Rev 9(2):167-73; Oppenheim (1997). *Clin Cancer Res* 12, 2682-6; Taub, (1994) *Ther. Immunol.* 1(4), 229-46 or Michiel, (1992). *Semin Cancer Biol* 3(1), 3-15).

[0170] Immune cell activity that may be measured include, but is not limited to, (1) cell proliferation by measuring the DNA replication; (2) enhanced cytokine production, including specific measurements for cytokines, such as IFN- γ , GM-CSF, or TNF- α ; (3) cell mediated target killing or lysis; (4) cell differentiation; (5) immunoglobulin production; (6) phenotypic changes; (7) production of chemotactic factors or chemotaxis, meaning the ability to respond to a chemotactin with chemotaxis; (8) immunosuppression, by inhibition of the activity of some other immune cell type; and, (9) apoptosis, which refers to fragmentation of activated immune cells under certain circumstances, as an indication of abnormal activation.

[0171] Also of interest are enzymes present in the lytic package that cytotoxic T lymphocytes or LAK cells deliver to their targets. Perforin, a pore-forming protein, and Fas

ligand are major cytolytic molecules in these cells (Brandau *et al.*, *Clin. Cancer Res.* 6:3729, 2000; Cruz *et al.*, *Br. J. Cancer* 81:881, 1999). CTLs also express a family of at least 11 serine proteases termed granzymes, which have four primary substrate specificities (Kam *et al.*, *Biochim. Biophys. Acta* 1477:307, 2000). Low concentrations of streptolysin O and pneumolysin facilitate granzyme B-dependent apoptosis (Browne *et al.*, *Mol. Cell Biol.* 19:8604, 1999).

[0172] Other suitable effectors encode polypeptides having activity that is not itself toxic to a cell, but renders the cell sensitive to an otherwise nontoxic compound--either by metabolically altering the cell, or by changing a non-toxic prodrug into a lethal drug. Exemplary is thymidine kinase (tk), such as may be derived from a herpes simplex virus, and catalytically equivalent variants. The HSV tk converts the anti-herpetic agent ganciclovir (GCV) to a toxic product that interferes with DNA replication in proliferating cells.

[0173] Any of the above-mentioned compounds can be used in combination therapy with the gene editing agents embodied herein. Concurrent administration of two or more therapeutic agents does not require that the agents be administered at the same time or by the same route, as long as there is an overlap in the time period during which the agents are exerting their therapeutic effect. Simultaneous or sequential administration is contemplated, as is administration on different days or weeks. The therapeutic agents may be administered under a metronomic regimen, e.g., continuous low-doses of a therapeutic agent.

[0174] The compositions described herein are suitable for use in a variety of drug delivery systems described above. Additionally, in order to enhance the *in vivo* serum half-life of the administered compound, the compositions may be encapsulated, introduced into the lumen of liposomes, prepared as a colloid, or other conventional techniques may be employed which provide an extended serum half-life of the compositions. A variety of methods are available for preparing liposomes, as described in, e.g., Szoka, *et al.*, U.S. Pat. Nos. 4,235,871, 4,501,728 and 4,837,028 each of which is incorporated herein by reference. Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with a tissue specific antibody. The liposomes will be targeted to and taken up selectively by the organ.

[0175] The appropriate dose of the compound is that amount effective to prevent occurrence of the symptoms of the disorder or to treat some symptoms of the disorder from which the patient suffers. By “effective amount”, “therapeutic amount” or “effective dose” is

meant that amount sufficient to elicit the desired pharmacological or therapeutic effects, thus resulting in effective prevention or treatment of the disorder.

[0176] When treating viral infections, an effective amount of the inhibitory compound is an amount sufficient to suppress the growth and proliferation of the virus. Viral infections can be prevented, either initially, or from re-occurring, by administering the compounds described herein in a prophylactic manner. Preferably, the effective amount is sufficient to obtain the desired result, but insufficient to cause appreciable side effects.

[0177] Dosage, toxicity and therapeutic efficacy of such compositions can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD₅₀ (the dose lethal to 50% of the population) and the ED₅₀ (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD₅₀/ED₅₀.

[0178] The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compositions lies preferably within a range of circulating concentrations that include the ED₅₀ with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any composition used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC₅₀ (i.e., the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

[0179] As described, a therapeutically effective amount of a composition (i.e., an effective dosage) means an amount sufficient to produce a therapeutically (e.g., clinically) desirable result. The compositions can be administered one from one or more times per day to one or more times per week; including once every other day. The skilled artisan will appreciate that certain factors can influence the dosage and timing required to effectively treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and other diseases present. Moreover, treatment of a subject with a therapeutically effective amount of the compositions of the invention can include a single treatment or a series of treatments.

[0180] The effective dose can vary, depending upon factors such as the condition of the patient, the severity of the viral infection, and the manner in which the pharmaceutical composition is administered. The effective dose of compounds will of course differ from patient to patient, but in general includes amounts starting where desired therapeutic effects occur but below the amount where significant side effects are observed. For human patients, the effective dose of typical compounds generally requires administering the compound in an amount of at least about 1, often at least about 10, and frequently at least about 25 $\mu\text{g}/24$ hr/patient. The effective dose generally does not exceed about 500, often does not exceed about 400, and frequently does not exceed about 300 $\mu\text{g}/24$ hr/patient. In addition, administration of the effective dose is such that the concentration of the compound within the plasma of the patient normally does not exceed 500 ng/mL and frequently does not exceed 100 ng/mL.

[0181] The compounds, when employed in effective amounts in accordance with the method described herein, are effective at eliminating the retrovirus from the subject.

[0182] In some embodiments, the compositions may be formulated as a topical gel, for example, to treat a melanoma after excision, or an autoimmune condition expressed as a skin condition *e.g.* pemphigus. In some embodiments, the compositions can be formulated as a nanoparticle encapsulating a nucleic acid.

[0183] A subject is effectively treated whenever a clinically beneficial result ensues. This may mean, for example, a complete resolution of the symptoms of a disease, a decrease in the severity of the symptoms of the disease, or a slowing of the disease's progression. These methods can further include the steps of a) identifying a subject (*e.g.*, a patient and, more specifically, a human patient) who has a certain disease to be treated; and b) providing to the subject the compositions comprising at least one anti-viral or anti-retroviral agent and/or a composition comprising the gene editing agents embodied herein.

[0184] In methods of treatment of HIV-1 infection, a subject can be identified using standard clinical tests, for example, immunoassays to detect the presence of HIV antibodies or the HIV polypeptide p24 in the subject's serum, or through HIV nucleic acid amplification assays. An amount of such a composition provided to the subject that results in a complete resolution of the symptoms of the infection, a decrease in the severity of the symptoms of the infection, or a slowing of the infection's progression is considered a therapeutically effective amount. The present methods may also include a monitoring step to help optimize dosing

and scheduling as well as predict outcome. In some methods of the present invention, one can first determine whether a patient has a latent HIV infection, and then make a determination as to whether or not to treat the patient with one or more of the compositions described herein. In some embodiments, the methods can further include the step of determining the nucleic acid sequence of the particular HIV harbored by the patient and then designing the guide RNA to be complementary to those particular sequences. For example, one can determine the nucleic acid sequence of a subject's LTR U3, R or U5 region, or pol, gag, or env genes etc., and then design or select one or more gRNAs to be precisely complementary to the patient's sequences. The novel gRNAs provided by the present invention greatly enhance the chances of formulating an effective treatment. The gRNAs targeted to nucleic acid sequences encoding a receptor used by a virus to infect a cell would prevent further infection.

[0185] In methods of reducing the risk of HIV infection, a subject at risk for having an HIV infection can be, for example, any sexually active individual engaging in unprotected sex, i.e., engaging in sexual activity without the use of a condom; a sexually active individual having another sexually transmitted infection; an intravenous drug user; or an uncircumcised man. A subject at risk for having an HIV infection can be, for example, an individual whose occupation may bring him or her into contact with HIV-infected populations, e.g., healthcare workers or first responders. A subject at risk for having an HIV infection can be, for example, an inmate in a correctional setting or a sex worker, that is, an individual who uses sexual activity for income employment or nonmonetary items such as food, drugs, or shelter.

[0186] The methods disclosed herein can be applied to a wide range of species, e.g., humans, non-human primates (e.g., monkeys), horses or other livestock, dogs, cats, ferrets or other mammals kept as pets, rats, mice, or other laboratory animals.

[0187] The methods of the invention can be expressed in terms of the preparation of a medicament. Accordingly, the invention encompasses the use of the agents and compositions described herein in the preparation of a medicament. The compounds described herein are useful in therapeutic compositions and regimens or for the manufacture of a medicament for use in treatment of diseases or conditions as described herein.

[0188] Any composition described herein can be administered to any part of the host's body for subsequent delivery to a target cell. A composition can be delivered to, without limitation, the brain, the cerebrospinal fluid, joints, nasal mucosa, blood, lungs, intestines,

muscle tissues, skin, or the peritoneal cavity of a mammal. In terms of routes of delivery, a composition can be administered by intravenous, intracranial, intraperitoneal, intramuscular, subcutaneous, intramuscular, intrarectal, intravaginal, intrathecal, intratracheal, intradermal, or transdermal injection, by oral or nasal administration, or by gradual perfusion over time. In a further example, an aerosol preparation of a composition can be given to a host by inhalation.

[0189] The dosage required will depend on the route of administration, the nature of the formulation, the nature of the patient's illness, the patient's size, weight, surface area, age, and sex, other drugs being administered, and the judgment of the attending clinicians. Wide variations in the needed dosage are to be expected in view of the variety of cellular targets and the differing efficiencies of various routes of administration. Variations in these dosage levels can be adjusted using standard empirical routines for optimization, as is well understood in the art. Administrations can be single or multiple (e.g., 2- or 3-, 4-, 6-, 8-, 10-, 20-, 50-, 100-, 150-, or more fold). Encapsulation of the compounds in a suitable delivery vehicle (e.g., polymeric microparticles or implantable devices) may increase the efficiency of delivery.

[0190] The duration of treatment with any composition provided herein can be any length of time from as short as one day to as long as the life span of the host (e.g., many years). For example, a compound can be administered once a week (for, for example, 4 weeks to many months or years); once a month (for, for example, three to twelve months or for many years); or once a year for a period of 5 years, ten years, or longer. It is also noted that the frequency of treatment can be variable. For example, the present compounds can be administered once (or twice, three times, *etc.*) daily, weekly, monthly, or yearly.

[0191] An effective amount of any composition provided herein can be administered to an individual in need of treatment. An effective amount can be determined by assessing a patient's response after administration of a known amount of a particular composition. In addition, the level of toxicity, if any, can be determined by assessing a patient's clinical symptoms before and after administering a known amount of a particular composition. It is noted that the effective amount of a particular composition administered to a patient can be adjusted according to a desired outcome as well as the patient's response and level of toxicity. Significant toxicity can vary for each particular patient and depends on multiple factors including, without limitation, the patient's disease state, age, and tolerance to side effects.

[0192] Any method known to those in the art can be used to determine if a particular response is induced. Clinical methods that can assess the degree of a particular disease state can be used to determine if a response is induced. The particular methods used to evaluate a response will depend upon the nature of the patient's disorder, the patient's age, and sex, other drugs being administered, and the judgment of the attending clinician.

[0193] *Kits*

[0194] The compositions described herein can be packaged in suitable containers labeled, for example, for use as a therapy to treat a subject having a viral infection, for example, an HIV infection or a subject at risk of contracting for example, an HIV infection. The containers can include a composition comprising at least one anti-viral or anti-retroviral agent, a gene-editing agent and one or more of a suitable stabilizer, carrier molecule, flavoring, and/or the like, as appropriate for the intended use. In other embodiments, the kit further comprises one or more therapeutic reagents that alleviate some of the symptoms or secondary bacterial infections that may be associated with an HIV infection. Accordingly, packaged products (e.g., sterile containers containing one or more of the compositions described herein and packaged for storage, shipment, or sale at concentrated or ready-to-use concentrations) and kits, including at least one composition of the invention, and instructions for use, are also within the scope of the invention. A product can include a container (e.g., a vial, jar, bottle, bag, or the like) containing one or more compositions of the invention. In addition, an article of manufacture further may include, for example, packaging materials, instructions for use, syringes, delivery devices, buffers or other control reagents for treating or monitoring the condition for which prophylaxis or treatment is required.

[0195] The product may also include a legend (e.g., a printed label or insert or other medium describing the product's use (e.g., an audio- or videotape)). The legend can be associated with the container (e.g., affixed to the container) and can describe the manner in which the compositions therein should be administered (e.g., the frequency and route of administration), indications therefor, and other uses. The compositions can be ready for administration (e.g., present in dose-appropriate units), and may include one or more additional pharmaceutically acceptable adjuvants, carriers or other diluents and/or an additional therapeutic agent. Alternatively, the compositions can be provided in a concentrated form with a diluent and instructions for dilution.

[0196] While various embodiments of the present invention have been described above, it should be understood that they have been presented by way of example only, and not limitation. Numerous changes to the disclosed embodiments can be made in accordance with the disclosure herein without departing from the spirit or scope of the invention. Thus, the breadth and scope of the present invention should not be limited by any of the above described embodiments.

[0197] All documents mentioned herein are incorporated herein by reference. All publications and patent documents cited in this application are incorporated by reference for all purposes to the same extent as if each individual publication or patent document were so individually denoted. By their citation of various references in this document, applicants do not admit any particular reference is “prior art” to their invention.

EXAMPLES

[0198] *Example 1: Combination of CRISPR-Cas9 and Long-acting Antiretroviral Therapy Eliminates HIV-1 Infection in Humanized Mice*

[0199] A cure of HIV-1 infection has been stalled by the absence of a strategy for effective eradication of HIV-1 from the infected tissues and cells serving as viral reservoirs. As such, rebound uniformly occurs after cessation of currently used antiretroviral therapy, ART, that potently controls viral replication but does not eliminate proviral DNA.

[0200] *Materials and Methods*

[0201] *Cell culture:* TZM-bl reporter cell line (AIDS Reagent Program, Division of AIDS, NIAID, NIH, Bethesda, MD) and HEK-293T cells were cultured in DMEM high glucose complemented with 10% FBS and gentamicin (10µg/ml). Jurkat, Clone E6 cells were purchased from ATCC (TIB-152™) and were cultured in RPMI medium containing 10% FBS and gentamicin (10ug/ml). Patient blood samples were obtained through the Comprehensive NeuroAIDS Center (CNAC) Clinical Core (Temple University, Philadelphia, PA, USA). PBMCs were isolated from human peripheral blood by density gradient centrifugation using Ficoll-Paque reagent. Blood sample volume was adjusted to 30ml with HBSS buffer, gently layered on 15 ml of Ficoll-Paque cushion and centrifuged for 30 minutes at 1500 RPM. PBMCs containing layer was collected, washed 3 times in HBSS buffer and counted. Cells were incubated with PHA (5µg/ml) for 24 h and then cultured in RPMI with 10% FBS and gentamicin (10ug/ml) supplemented with human rIL-2 at a concentration of 30ng/ml (STEMCELL Tech.). Fresh media was added every 2-3 days.

[0202] *In vitro infection:* HEK-293T cells were transfected using CaPO₄ precipitation method in the presence of chloroquine (50uM) with 30μg of pNL4-3-EGFP-P2A-Nef plasmid (13) /2.5x10⁶ cells/100mm dish. Next day, medium was replaced; and 24h and 48h later supernatants were collected, clarified at 3000RPM for 10 minutes, filtered through 0.45um filter, and concentrated by ultracentrifugation for 2h with 20% sucrose cushion (25). Viral pellets were resuspended in HBSS by gentle agitation overnight, aliquoted, and tittered in Jurkat cells by FACS for GFP expression. Jurkat cells were infected by spinoculation for 1.5h (26), 32 °C in 500μl inoculum containing 8μg/ml polybrene then resuspended and left for 4h then 500μl of growth medium was added. Next day, cells were washed 3 times with PBS and re-suspended in growth medium.

[0203] *Generation of humanized NSG mice:* NOD/scid-IL-2Rγ^{null} (NSG) mice were obtained from the Jackson Laboratories, Bar Harbor, ME and bred under specific pathogen-free conditions in accordance with the ethical guidelines for care of laboratory animals at the University of Nebraska Medical Center (UNMC) set forth by the National Institutes of Health. CD34⁺ cells were obtained from human cord blood and enriched using immune-magnetic beads (CD34⁺ selection kit; Miltenyi Biotec Inc., Auburn, CA, USA). CD34⁺ cell purity was >90% by flow cytometry. Cells were transplanted into newborn mice irradiated at 1 Gy using a C9 cobalt 60 source (Picker Corporation, Cleveland, OH, USA). Cell suspension was delivered by intrahepatic (i.h.) injection of 10⁴ cells/mouse in 20 μl phosphate-buffered saline (PBS) with a 30-gauge needle. Humanization of the animals was affirmed by flow cytometry (21, 27) for CD45 and CD3 staining of blood immune cells shown in FIGS. 8A, 8B.

[0204] *Drugs and antibodies:* Dolutegravir (DTG), lamivudine (3TC), and abacavir (ABC) were generous gifts from ViiV Healthcare, Research Triangle Park, NC. Rilpivirine (RPV) was purchased from Hangzhou Bingo Chemical Co., Ltd, Hangzhou, China. Poloxamer 407 (P407), HEPES buffer, ciprofloxacin, paraformaldehyde (PFA), and 3,3'-diaminobenzidine (DAB) were purchased from Sigma-Aldrich, St. Louis, MO. Diethyl ether, endotoxin-free water, gentamicin, acetonitrile (ACN), methanol, KH₂PO₄, bovine serum albumin (BSA), Triton X-100, LC-MS-grade water, and TRIzol reagent were purchased from Fisher Scientific, Hampton, NH, USA. FITC-conjugated mouse anti-human CD45, Alexa Fluor-conjugated 700 mouse anti-human CD3, APC-conjugated mouse anti-human CD4, and BV421-conjugated mouse anti-human CD8 antibodies were purchased from BD Biosciences, San Jose, CA. Monoclonal mouse anti-human HIV-1p24 (clone Kal-1), monoclonal mouse

anti-human leukocyte antigen (HLA-DR; clone CR3/43), and the polymer-based HRP-conjugated anti-mouse EnVision+ secondary antibodies were purchased from Dako, Carpinteria, CA.

[0205] *Flow cytometry:* GFP expression in infected cells was quantified using Guava EasyCyte Mini flow cytometer (Guava Technologies, Hayward, CA, USA). Cells were first fixed for 10 minutes in 2% paraformaldehyde then washed 3 times in PBS and analyzed. Peripheral blood was collected from the submandibular vein into ethylenediaminetetraacetic acid (EDTA)-coated tubes or by cardiac puncture at the study end. Blood leukocytes were tested for human pan-CD45, CD3, CD4, CD8, CD14, and CD19 markers as six-color combinations using LSR-II FACS analyzer (BD Biosciences). Antibodies and isotype controls were obtained from BD Pharmingen, San Diego, CA, USA, and staining was analyzed with a FlowJo (BD Immunocytometry Systems, Mountain View, CA, USA). Results were expressed as percentages of total number of gated lymphocytes. The percentages of CD4 and CD8 positive cells were obtained from human CD3+ gate set (10).

[0206] *In vivo HIV-1 infection:* At 18 weeks of age, 25 humanized NSG (NSG-hu) mice were infected intraperitoneally (i.p.) with HIV-1_{NL4-3} (27, 28) at 10⁵ tissue culture infective dose₅₀ (TCID₅₀)/ml and sacrificed at days 1, 3, 7 and 14; n=5 at each time point. Five control-uninfected animals were included in all test evaluations. Levels of viral RNA copies/ml were analyzed with the automated COBAS Ampliprep System V2.0/Taqman-48 system (Roche Molecular Diagnostics, Basel, Switzerland) (17, 29). For this assay, 100 µl of mouse serum was diluted to 1 ml with sterile filtered normal human serum. The detection limit of the assay after dilution is 200 viral RNA copies/ml.

[0207] *Immunohistochemistry examinations:* Spleen, lung, liver, and lymph nodes were perfused with PBS followed by 4% paraformaldehyde and then post fixed overnight and embedded in paraffin. Five-micron thick sections were cut from the paraffin blocks, mounted on glass slides and labeled with mouse monoclonal antibodies (DakoCytomation, Carpinteria, CA, USA) for HLA-DQ/DP/DR (clone CR3/43, 1:100) and HIV-1p24 (1:10). The polymer-based HRP-conjugated anti-mouse Dako EnVision system was used as a secondary detection reagent and developed with 3,3'-diaminobenzidine (DAB). All paraffin-embedded sections were counterstained with Mayer's hematoxylin. Deletion of primary antibodies or mouse IgG served as controls. Images were obtained with a Nikon DS-Fi1 camera fixed to a Nikon Eclipse E800 (Nikon Instruments, Melville, NY) using NIS-Elements F 3.0 software.

[0208] *Preparation of antiretroviral nanoformulations:* Antiretroviral prodrugs and their polymer encasements were performed as previously described (7, 8). Myristoylated modifications for DTG, 3TC and ABC were made (referred to as MDTG, M3TC and MABC) to enhance the incorporation into the nanoparticles and RPV was encased solely by poloxamer 407 (P407) in unmodified form using high pressure homogenization to form crystalline nanoformulated drugs. Particle size, polydispersity index, and zeta potential were determined by dynamic light scattering using a Malvern Nano-ZS (Malvern, Worcestershire, UK) (30). Final drug concentrations in the nanoformulation suspensions and injection solutions were determined by HPLC-UV/Vis and UPLC-MS/MS. A 40-50 μ l volume for each nanoformulation combination (NMDTG/NRPV and NM3TC/NMABC) was administered by intramuscular (IM) injection in opposing thigh muscles of the mice.

[0209] *Nucleic acid extractions and q-PCR assays:* In studies presented in FIGS. 1A-1G, 3A-3E, and 4A-4G total viral nucleic acids (RNA and DNA) extracted from tissue or cells were acquired from the spleen, bone marrow, lung, GALT, liver, kidney, and brain using a Qiagen Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Total cellular DNA and RNA obtained from the HIV-1 infected cell line ACH2 served as a positive control, while human genomic DNA was obtained from uninfected humanized mice as negative controls. Cell-associated HIV-1 RNA and DNA were quantified by q-PCR and droplet digital PCR (ddPCR) assays. Because of extremely low numbers of latently-infected human cells in HIV-infected humanized mice after long-term ART, detection of total HIV-1 DNA, requires two rounds of PCR amplification. The first round of PCR was performed on a conventional PCR machine (T100 Thermal Cycler, Biorad, CA, USA) in 25 μ l of PCR master mix containing 5 μ l of template and 50 ng each of both primers annealing to HIV gag region as follows: 94 °C for 3 min, followed by 15 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 min. The product of the first PCR was subsequently used as a template in the second semi-nested real-time PCR amplification performed on the ABI Prism 7000 real-time PCR machine (Applied Biosystems, MA, USA) using TaqMan detection probe. A total of 2 μ l of the first PCR product was diluted to 50 μ l with PCR master mix containing 0.2 μ M concentrations each of both primers and 0.2 μ M TaqMan dual-labeled fluorescent probe. Real-time PCR settings were as follows: 50 °C for 2 min, then 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min. The amplicon sizes are 221 bp for the first round of PCR and 83 bp for the second round (real-time) PCR. ACH2 cells (8×10^5) containing one integrated copy of HIV-1 per cell were used in triplicate as standards and HIV

copy numbers ranging in serial 10-fold dilutions from 10^5 to 10^1 DNA copies/reaction (17, 18). Integrated DNA provirus was quantified using an adapted alu-PCR assay as described by Agosto *et al.* (31) with modifications for the second round of PCR, following prior published methods (32). Briefly, samples underwent a first-round PCR amplification (95 °C for 2 min; 20 cycles of 95 °C for 15 s, 50 °C for 15 s, and 72 °C for 150 s) using 100 nM alu and 600 nM gag reverse primers. Five µl of the first-round product were amplified in a nested protocol using the assay for HIV-1 gag gene (second PCR primers and probe), as described above. A first-round PCR with 3 replicates using only the gag reverse primer (gag only) acted as background un-integrated control. Serially diluted integration site standards were used to construct a standard curve for each plate. Integration levels per cell were calculated by subtracting gag-only signals from the alu-gag quantification. Semi-nested real-time PCR on HIV-1 RNA was performed as described (17, 18). The eluted cellular RNA was first subjected to DNase treatment to remove HIV-1 DNA to avoid the interference with the quantitation. For reverse transcription assay, random hexamers were used as primers and SuperScript III (Invitrogen, MA, USA) to synthesize first-strand cDNA at 42 °C for 60 min. cDNA was used for the unspliced (usRNA) assay. Two rounds of PCR were performed under the same PCR conditions as described for total viral DNA. For the usRNA assay, real-time PCR was run for 45 cycles. For the usRNA assay, the same primers and fluorescent probe were used as for the total viral DNA assay. The amplicon size was 115 bps.

[0210] In studies presented in FIGS. 2A-2C, 11, 12A-12C, 13A-13M, 14A-14F, 20A-20D, 21A, 21B and 22A, 22B and, frozen tissues were homogenized using Bullet Blender homogenizer (Next Advance, Averill Park, NY, USA) using bead combinations and settings specific for every tissue according to manufacturer protocols. T1 buffer from NucleoSpin Tissue kit (Macherey-Nagel, Duren, Germany) was used for homogenization/initial lysis followed by over-night proteinase K digestion. Extraction of genomic DNA was completed according to the protocol of the manufacturer. For standard PCRs (Table 1.1), 500 ng of extracted DNA were subjected to PCR using Fail Safe PCR kit and buffer D (Epicentre, Madison, WI, USA) under the following PCR conditions: 94°C 5 minutes, 30 cycles (94 °C 30s, 55 °C 30s, 72 °C 30s), 72 °C 7 minutes using 1st round primers followed by nested PCR using diluted 1st round PCR reaction. Nested PCR products were subjected to Sanger sequencing directly if only one amplicon population was detected by agarose gel electrophoresis. For multiple amplicons detected, or to investigate the composition of HIV excision, each amplicon population was separated and purified from an agarose gel

electrophoresis and then cloned into TA vector (Invitrogen, Carlsbad, CA, USA). Plasmid DNA containing excised HIV amplicon was purified from each bacterial colony for Sanger sequencing (Genewiz, South Plainfield, NJ, USA). HIV-1 DNA was quantified using TaqMan qPCR specific for HIV-1 pol and env genes and cellular beta-globin gene as a reference (Table 1.2). Prior to qPCR, genomic DNA was diluted to 10ng/ul and then 5µl (=50 ng) were taken per reaction/well. Reaction mixtures were prepared using Platinum Taq DNA Polymerase (Invitrogen) according to a simplified procedure (33) Standard was prepared from serial dilutions of U1 cells genomic DNA since it contains two single copies of HIV-1 provirus per diploid genome equal to beta-globin gene copy number. qPCR conditions: 98 °C 5 minutes, 45 cycles (98 °C 5 minutes, 45 cycles (98 °C 15s, 60 °C 30s with acquisition, 72 °C 1 minute). Reactions were carried out and data analyzed in a LightCycler96 (Roche, Basel, Switzerland). For RT PCR, TRIzol reagent (Ambion, Austin, TX, USA) was used for initial RNA extraction followed by clean up using RNeasy kit (Qiagen, Hilden, Germany) with DNase I digestion in the extraction column. Total 0.5µg of RNA was used for M-MLV reverse transcription (Invitrogen). For gRNA expression screening specific reverse primer (pX601gRNA scaffold/R, Table 1.3.) was used in RT reaction followed by standard PCR using target LTR 1 or Gag D sense oligos as forward primers (Table 1.3) and agarose gel electrophoresis. For checking saCas9 mRNA expression oligo-dT primer mix was used in RT and cDNA was subjected to PCR using saCas9 specific primer pairs and b-actin as a reference (Table 1.3).

[0211] *Sequence analysis:* Sanger sequencing results were analyzed using Clustal Omega (EMBL-EBI) multiple sequence alignment tool and Sequence Scanner Software 2 (Applied Biosystems).

[0212] *Off-target analysis - Cell culture model:* TZM-bl cells were plated in 6 well plates at 1×10^5 cells/well and co-transfected using Lipofectamine 2000 reagent (Invitrogen) with 1µg of control pX601-AAV-CMV:NLS-SaCas9-NLS-3xHA-bGHpA;U6::BsaI-sgRNA (Addgene #61591) or 1µg of pX601-LTR1-GagD (16) plasmid together with 0.2µg of pKLV-U6gRNA(Bbs1)-PGKpuro2ABFP (Addgene 50946) to provide puromycin selection marker. Next day, cells were transferred into 100-mm dishes and cultured in the presence of puromycin (Sigma) at concentration 1 µg/ml. After two weeks, surviving clones were isolated using cloning cylinders (Corning, Corning, NY, USA). Genomic DNA was prepared from each single cell clone and LTR specific PCRs followed by gel purification; TA cloning and Sanger sequencing were performed. The clones showing the presence of on target CRISPR-

Cas9 induced InDel mutations at target LTR 1 site in integrated HIV-1 LTR sequence (n = 6) together with two control clones were selected for further in vitro off target analysis. The list of potential OFF target sites in human genome for HIV-1 target LTR 1 and Gag D was created using Benchling CRISPR design tool (Benchling, San Francisco, CA 94103 (benchling.com) Tables 2 and 3). Total three potential OFF target sites were chosen (the top scorer plus two top gene specific potential off target sites, see Tables 2 and 3, highlighted in yellow) for PCR based screening in selected single cell clones. The potential OFF target regions (Table 1.4) were PCR amplified, cloned into TA vector, and sent for Sanger sequencing (3-6 sequences/single cell clone/single OFF target).

[0213] *Humanized mice model:* The genetic variation analyses among the three treatments were performed through the next generation sequencing (by the Novogene NGS facility) and bioinformatics tools for four sample animals, one animal from the LASER ART, one animal from CRISPR-Cas9- and two no-rebound animals from the LASER ART+CRISPR-Cas9 groups. The main objective was detecting the possible CRISPR/Cas9 off-target sites. Besides this, some genetic variations such as single nucleotide polymorphisms (SNP), -nsertion-deletions (InDels), structural variants (SVs) and copy number variants (CNVs) were analyzed for those four animals and the results are given in FIGS. 19A-19C, 20A-20D and Tables 4-6. After a thorough quality control step, the resulting paired-end short-reads were mapped to the human reference genome (Human_G1K_V37) utilizing Burrows-Wheeler Aligner (BWA) algorithm. For the animals #4356 (CRISPR-Cas9), #4348 and #4349 (LASER ART+CRISPR-Cas9) and #3539 (LASER ART), the coverages were reported to be 92.01%, 91.97%, 92.01% and 91.92%, while the sequencing depths were 36.08, 63.11, 45.22, and 15.41, respectively.

[0214] *ddPCR for detection of HIV-1 nucleic acids:* ddPCR was performed based on the water-oil emulsion droplet technology and used for viral detection using the outlined primers (Forward - 5'-TCAGCCCAGAAGTAATACCCATGT-3' and Reverse- 5' – CACTGTGTTTAGC ATGGTGTTT- 3') and a TaqMan probe. The ddPCR assay was run with the ddPCR™ SUPERMIX for Probes reagents in the QX200™ DROPLET DIGITAL™ PCR system (Bio-Rad Laboratories, Hercules, CA, USA). For quantification of integrated HIV-1 DNA, the eluted cellular DNA was PCR amplified (17, 18, 31, 32) for integrated viral DNA (iDNA) targeting the HIV-1 gag gene. Total 100 ng of each tissue DNA template were used for ddPCR amplifications and performed on the QX200™ DROPLET DIGITAL™ PCR system (Bio-Rad Laboratories, Hercules, CA, USA) using the ddPCR™ Supermix for Probes

reagents following the thermal cycling conditions for TaqMan detection. Data acquisition and analysis were done using QX200 droplet reader and QUANTASOFT™ software (Bio-Rad Laboratories, Hercules, CA, USA).

[0215] *RNAscope assay:* Viral RNA was detected as single brown dots or cluster of dots in 10 µm thick spleen tissue sections using antisense probe V-HIV1- Clade-B (ACD cat no 416111) targeting 854-8291 bp of HIV-1_{NL4-3} (34) Human peptidylprolyl Isomerase B (PPIB) was used as positive control for the spleen tissue analyzed (images were captured at 40x magnification).

[0216] *Viral recovery:* Phytohemagglutinin (PHA) and interleukin-2 (IL-2) stimulated peripheral blood mononuclear cells (PBMCs) obtained from leukopaks from HIV-1,2 seronegative donors were co-cultured with human bone marrow (BM) or spleen cells recovered from infected and or LASER ART with and without AAV₉-CRISPR-Cas9 treated humanized mice. PBMCs were used in assays after 3-day treatment maintained in 10% RPMI with 30 U/ml of IL-2 then co-cultured with human BM or spleen cells at concentrations of (1:5) (35-37). Cells were harvested eight days later for HIV-1 DNA (A) and RNA (B) using semi-nested real-time PCR assay and supernatant fluids assayed for reverse transcriptase activity. Data are expressed as total HIV-1 DNA (A) or RNA (A) copies/10⁶ human CD45 cells. One of two dual treated animals was tested and confirmed viral sterilization. Viral rescue was observed in other animals tested.

[0217] *Excision efficiencies and hierarchal clustering:* The excision efficiencies for each animal, tissue, and HIV-1 gene segment were calculated as the ratio of the number of the sequencing-verified PCR product to all members in each group with denoted experimental conditions (i.e. treatments, tissues, etc. shown in FIGS. 2A-2C, 11, 12A-12C, 13A-13M). Defined in such a way that the excision efficiencies can be viewed as frequentist probabilities, i.e. the ratio of the frequency of occurrence of the event of interest to the total number of experimental repeats. This interpretation of excision efficiencies provides the user with a predictive value, as they can be used to set a prior expectation on the success rate of each treatment (LASER ART, CRISPR-Cas9, and LASER ART plus CRISPR-Cas9) in excising the desired segments of HIV-1 gene in the studied tissues and further to relate that to the likelihood of cure.

[0218] *Hierarchical clustering* was performed on the efficiency values of truncation events under different treatments and across different animals, tissues, and HIV-1 gene

segments. Once the excision efficiencies were calculated under different combinations of experimental conditions, the hierarchical clustering scheme was employed to group the efficiency values into a multilevel cluster tree represented by a dendrogram. The corresponding efficiency values were listed in heat-map table, to make the clusters visually detectable. To this end, three combinations were considered: i) excision probabilities of different HIV-1 segments in 6 different tissues of animals undergoing antiretroviral treatment, CRISPR-Cas9 mediated editing, and the combined treatments (FIG. 17A); ii) excision probabilities of different segments in different animals under the three treatments (FIG. 17B); and iii) probabilities of observing at least one positive band for each specified tissue in all animals (FIG. 17A). Clusters of FIGS. 17B and 17C also include additional conditions of “Cure” and qPCR data to identify which animals experienced complete cure and highest viral genome eradication. Note that, in all figures, S1 refers to 5’ LTR to Gag excision and S2 refers to Gag to 3’ LTR excision of the HIV-1 genome, respectively.

[0219] *Statistical analyses:* Data were analyzed by GraphPad Prism 7.0 software (La Jolla, CA, USA). Data are represented as the mean +/- the standard error of the mean (SEM). Experiments were performed using a minimum of three biologically distinct replicates. For comparisons of two groups, Student’s t test was used. T cell populations, viral RNA and DNA, and viral load were analyzed by one-way ANOVA with Bonferroni correction for multiple-comparisons. For studies with multiple time points, two-way factorial ANOVA and Bonferroni’s post-hoc tests for multiple comparisons were performed. Multiple comparisons were corrected for the false discovery rate (FDR) using the Benjamini-Hochberg procedure. Animal studies included a minimum of 5-7 animals per group unless otherwise noted. Extreme outliers beyond the 99% confidence interval of the mean and 3-fold greater than the SEM were excluded. Significant differences were determined at a $p < 0.05$.

[0220] *Results*

[0221] A functional cure of HIV infection was documented in a single person (1). However, efforts were stalled by a combination of limited therapeutic access to viral reservoirs, rapid spread of infection, high numbers of virus susceptible cells, and complete inability to eliminate latent integrated proviral DNA. These therapeutic treatments have precluded viral eradication as rebound was seen after cessation of antiretroviral therapy (ART) (2-6).

[0222] To address each of these limitations, highly hydrophobic and lipophilic

antiretroviral prodrugs termed herein “long-acting slow effective release antiretroviral therapy” (LASER ART), were produced to improve drug penetrance across cell and tissue barriers and improve control over ongoing viral infection (7-10). Further, CRISPR-Cas9 technology was employed that specifically and efficiently excised fragments of integrated HIV-1 proviral DNA from the host genome in cell cultures as well as in several tissues from small animal models (11-16). To provide proof of concept that LASER ART and CRISPR-Cas9 treatments could produce synergy towards viral elimination, gut-associated lymphoid tissue (GALT), spleen, lymph nodes, brain, lung, liver, and kidney tissues of NOD.Cg-Prkdc^{scid} Il2rg^{tm1Wjl}/SzJ (NSG) mice were populated with human peripheral blood lymphocytes (PBLs) then infected with 10⁴ tissue culture infective dose₅₀ (TCID₅₀) of HIV-1_{NL4-3} (17, 18). Three days later, animals were divided into four groups (n=7 for each group). Groups were control (uninfected) and infected animals left untreated or treated with LASER ART as defined by combinations of myristoylated dolutegravir (DTG), lamivudine (3TC) (8) and abacavir (ABC) (7) prodrugs and rilpivirine (RPV) nanoformulations with or without AAV₉-CRISPR-Cas9. All treatments were simultaneously administered. After two weeks, CD4⁺ T cells and viral DNA and RNA levels were assessed in blood and tissues. No significant differences in the levels of CD4⁺ T cells and viral DNA and RNA levels were observed between the treatment groups. However, animals treated with both LASER ART and AAV₉-CRISPR-Cas9 viral RNA and DNA levels were decreased more than those receiving LASER ART which by itself restored CD4⁺ T cells and reduced plasma viral RNA to or below baseline (FIGS. 5A-5F). These results provide evidence that CRISPR-Cas9 would be most effective in HIV-1 LTR and the Gag gene excision in an ART setting. Support for this notion was realized from follow up *in vitro* and *ex vivo* studies. These investigations showed significant increases in the proficiency of CRISPR-Cas9 excision of HIV-1 proviral DNA in infected T-cells following ART-induced viral restriction (FIGS. 6A-6C, 7A-7E).

[0223] Based on these observations, an amended treatment strategy was adopted (FIG. 1A) based on the assumption that ART-induced viral suppression would improve CRISPR-Cas9 editing efficiency and facilitate viral elimination without bystander tissue toxicities (16, 19, 20). humanization of the animals was confirmed by flow cytometry for CD45 and CD3 staining of blood immune cells in animals demonstrating human cell survival in all mouse groups (FIGS. 17A and 8B). All infected animals showed marked depletion of CD4⁺ T cells (FIGS. 1B and 1G) and plasma viral RNA levels at a median of 2.2x10⁵ copies/ml (FIGS. 1C and 1F). Immunohistochemical examination for HIV-1p24 antigens in

spleen, lymph node, and lung showed broad distribution of infected HLA-DR reactive human cells (FIG. 1D). Semi-nested real-time q-PCR HIV-1 nucleic acid detections confirmed viral infection in tissue (FIG. 1E). These infected animals were then divided into four groups. The first received no treatment; group 2 received a single intravenous (IV) injection of AAV₉-CRISPR-Cas9, and the third and fourth groups received 40-45 mg/kg LASER ART. The fourth group, in addition, received AAV₉-CRISPR-Cas9 after three weeks of LASER ART. After the last administration of LASER ART, animals were observed for 8 weeks for any evidence of viral rebound, which corresponded to 5 weeks of AAV₉-CRISPR-Cas9 treatment. Such rebound was observed in all groups with the exception of two animals in group 4 animals #4346 and #4349, which received a combination LASER ART and CRISPR-Cas9 (FIG. 1F). Restoration of CD4⁺ T cell counts ($90\% \pm 7\%$) was observed in dual treated group 4 animals, which was higher than those seen in group 3 that received only LASER ART ($82\% \pm 12\%$). In the absence of LASER ART, restoration of CD4⁺ T cells with CRISPR-Cas9 treatment (group 2) remained low ($15\% \pm 6\%$), yet slightly higher than those seen in Group 1 (no treatments, less than 6%) (FIG. 1G). Immunohistochemical evaluation of spleens from HIV-1 infected animals for the presence of CD4⁺ T cells showed comparable increases (FIG. 9). Further, detection of human DNA sequence in spleen confirmed the presence of the human cells in the spleen of the humanized animals (FIG 10).

[0224] Gel electrophoresis analysis of the PCR amplified DNA fragments using specific pairs of primers designed for detection of the various cleavage events (FIG. 2A) revealed robust cleavage and excision of viral DNA fragments obtained from spleen, GALT, and kidney of the group of animals treated with LASER ART and AAV₉-CRISPR-Cas9 (FIG. 2B). Also, it was noted that the type of excision differed in various tissues among the animals as well as in the same individual animals. Efficient excision of the predicted fragment in other tissues including lung, liver, and brain was also observed in some of the animals with dual treatments (FIG. 11). The integrity and precision of the HIV-1 DNA excision by CRISPR-Cas9 were sequence verified (FIG. 2C, and FIGS. 12A-12C, 13A-13M). In mice that received AAV₉-CRISPR-Cas9 without LASER ART fragmental deletion was detected. Several other DNA fragments in tissues from the experimental animals, including those that received only LASER ART, were amplified, which after sequencing were found unrelated to HIV or editing by CRISPR-Cas9 that may represent replication defective HIV-1 (highlighted by double asterisks) (FIG. 2B and FIGS. 14 A-14F). Amplification of the DNA fragments corresponding to control housekeeping actin gene in the various tissues and

expression of gRNAs and Cas9 are shown (FIG. 15).

[0225] Clustering analysis revealed similar excision patterns with high efficiency across the different tissues in the cohort of animals that received combination treatments in comparison to those detected in the groups that received only CRISPR-Cas9 (FIGS. 16A-16C). This hierarchical clustering heat map may offer a predictive capability for viral elimination after the interruption of LASER ART in this model. Bioinformatics analysis of human genome sequence data identified several human genome sites that may serve as targets for gRNAs that are designed for editing of HIV-1 DNA. However, results from sequencing of several selected sites with high scores of specificities and/or their locations in the exons ruled out any off-target effect on genome of human cell line (FIGS. 17A-17C and 18A-18F). Further, deep sequencing of genomic DNA from spleen of the four treated animals, including two that showed no rebound after combination treatment and one from each group with single treatment followed by bioinformatics analysis in search for somatic InDels mutation in the human genome by multiple alignments involving nucleic acids blast revealed no off target effects such as single nucleotide variations, translocation, inversion, deletion, tandem duplication, and insertion in the human genome, that can be attributed to CRISPR-Cas9 (FIGS. 19A-19C, 20A-20D and Tables 6, 7).

[0226] Next, tissue viral DNA and RNA levels were determined in tissues using ultrasensitive semi-nested real time qPCR with primers and probes designed for detection of HIV-1 gag (21, 22) DNA analysis results revealed that combination treatment was more effective than either LASER ART or CRISPR-Cas9 alone in reducing viral DNA copies. The spleen, GALT, and bone marrow of mice #4346 and #4349 showed no rebound (FIG. 3A). Similarly, results from the RNA detection assay corroborated with the data from the DNA study showed the combination of LASER ART and CRISPR-Cas9 reduced HIV-1 RNA production in select animals with complete absence of viral RNA in #4346 and #4349 (FIG. 3B). The presence of HIV-1 RNA was also examined by RNA scope using 10 µm thick spleen sections from infected animals and antisense probe V-HIV-1 Clade-B designed for targeting 854-8291 base pairs of the HIV-1_{NL4-3}. Mouse #4346 with no viral nucleic acid and rebound showed no evidence of viral gene expression (FIG. 4A). Results from the targeted qPCR for DNA sequence detection corresponded to the middle of HIV-1 genome and ruled out the presence of DNA corresponding to the *pol* and *env* genes (FIGS. 4B-4C). Additional evidence for the absence of HIV-1 genomes in the animals #4346 and #4349 was provided by digital droplet PCR (ddPCR) tests. This assay had a sensitivity of detection of < 2 viral

copies. Verifying prior results, no viral DNA was detected in spleens of mice #4346 and #4349 and examination of other tissues showed complete HIV-1 eradication (FIGS. 4D-4E). Finally, a viral rescue assay was performed by co-culturing bone marrow cells and splenocytes of representative samples with PHA/IL-2 PBMCs for an additional two weeks. Representative data (FIG. 4F) showed that while HIV-1 was rescued from 100% of samples with detectable viral DNA and RNA, despite the presence of high number of human cells, no evidence for virus recovery was observed in the samples from the two animals with eradicated HIV-1 DNA and RNA.

[0227] Accordingly, these results provide evidence that the combination of lipophilic LASER ART and AAV₉ delivered CRISPR-Cas9 can potentially lead to the cure of HIV-1 infection by elimination of the replication component of the virus in HIV-1 reservoirs of infected animals as evidenced by the absence of viremia for more than 8 weeks after the last ART treatment. Although the re-appearance of viremia in humans can occasionally be delayed longer (5), the rebound of HIV occurs an average of 2-4 weeks after ART interruption (4, 23) and 5-9 days in animal models (24). These results offer a realistic pathways toward an HIV-1 cure.

[0228] Table 1. PCR primers and probes

	primer	sequence
1. Standard PCRs		
5'LTR-gag	1 st round LTR F	5'- AATTGCGGCCGCTGGAAGGGCTAATTGGTCCC- 3'
	1 st round gag R	5'-TGTCACCTTCCCCTTGGTTCTCTC-3'
	nested 5'LTR F	5'- AAAAGAATTCGTGGATCTACCACACACAAGGC- 3'
	nested gag R	5'-AAAAGGATCCACCATTGCCCCTGGAGGTT-3
Gag-3'LTR	1 st round gag F	5'-GAAAGCGAAAGTAAAGCCAGAGGAGAT-3
	1 st round LTR R	5'-ACACAACAGACGGGCACACACTACTT-3'
	nested gag F	5'AAAAGAATTCGACAGCTACAACCATCCCTTCA GACAG-3'

	nested 3'LTR R	5'- AAAAGGATCCAGCAGTGGGTTCCCTAGTTAGCC AG-3'
LTRs	1 st round LTR -413/S	5'-TTGGCAGAACTACACACCAGGG -3'
	1 st round LTR +43/AS	5'-CCGAGAGCTCCCAGGCTCAGATCT-3'
	nested LTR -374/S	5'-TTAGCAGAACTACACACCAGGGCC-3'
	nested LTR -19/AS	5'-GCTGCTTATATGTAGCATCTGAG-3'
Hs beta-globin	Hs b-globin F	5'-CCCTTGGACCCAGAGGTTCT-3'
	Hs b-globin R	5'-CGAGCACTTTCTTGCCATGA-3'
Mm beta-globin	Mm b-globin F	5'-CCCTTGGACCCAGCGGTACT-3'
	Mm b-globin R	5'-GTTATCACCTTCTTGCCATG-3'
2. Taqman qPCRs		
pol	HIV-1 pol/int F	5'-TCCAGCAGAGACAGGGCAAG-3'
	HIV-1 pol/int R	5'-TGCCAAATTCCTGCTTGATCCC-3'
	HIV-1 pol/int probe	5'-HEX-CGCCCACCAACAGGCGGCCTTAAGT- ZEN-IowaBlackFQ-3'
env	HIV-1 Env F	5'- TCCTTGGGATGTTGATGATCT-3'
	HIV-1 Env R	5'- TGGCCCAAACATTATGTACC-3'
	HIV-1 Env Probe	5'-FAM-TGGTGGTTGCTTCTTTCCACACA-ZEN- IowaBlackFQ-3'
reference	Hs b-globin F	5'-CCCTTGGACCCAGAGGTTCT-3'
	Hs b-globin R	5'-CGAGCACTTTCTTGCCATGA-3'
	Hs b-globin probe:	5'-FAM- GCGAGCATCTGTCCACTCCTGATGCTGTTATGGG CGCTCGC-ZEN-IowaBlackFQ-3'
3. RT-PCRs		
	LTR1/F	5'-GCAGAACTACACACCAGGGCC-3'
	GagD/F	5'-GGATAGATGTAAAAGACACCA-3'
	pX601gRNAscaffold/ R	5'-CGCCAACAAGTTGACGAGAT-3'
	SaCas9/263/F	5'-TCGACTACAACCTGCTGACC-3'
	SaCas9/SEQ1	5'-GGTGGGCTTCTTCTGCTT-3'
	b- actin S	5'-CTACAATGAGCTGCGTGTGGC-3'
	b-actin AS	5'-CAGGTCCAGACGCAGGATGGC-3'
4. In vitro OFF target analysis		
LTR 1 OFF targets	LTR1OFFch8/F	5'-GAGTGACCTTCCCAAATTGC-3'

	LTR1OFFch8/R	5'-ATGGTGAGGTGAGGGATGAG-3'
	TSC2/35001F	5'-CAGACTCTGATGGGTGGCAG-3'
	TSC2/35398R	5'-GCTAAGGAGAGAGGGTGGGA-3'
	TUB/66607F	5'-CCAAGTGGCCCTCAGATTACA-3'
	TUB/67015R	5'-TCATTCACCCCAAATCCTACGG-3'
Gag D OFF targets	GagDOFFch3/F	5'-CATTAACCACCTGGGGAACA-3'
	GagDOFFch3/R	5'-TCTCAGACCCAGGAATGTCA-3'
	TACC2/392F	5'-GAGGACTCTCCAGCCAAAGG-3'
	TACC2/782R	5'-GAGCTGGGGGTCTTAGAGGA-3'
	ADNP/41574F	5'-TGCACCAGCCAAAACCTTAGGA-3'
	ADNP/41996R	5'-TCTAATTAGGTGGCAGCACGTT-3'

[0229] Table 2

HIV-1 LTR 1 target (+ strand)							
GCAGAACTACACACCAGGGGCCAGGGAT							
Sequence	PAM	Score	Gene	Chromosome	Strand	Position	Mismatch On-target
CTCAAGCTGACAGGCGGGGG	ATGAA	3.9		chr8:12315137			
TCAGATCTCAGAGGAGAGG	AGGAG	1.3		chr9:38350394	1	22915337	4 FALSE
ACAGAGAGAGTACACAGGCGG	CAGAG	0.8		chr22:20136985	-1	38360364	4 FALSE
GTAGAGACTATGTCACAGGCGA	AGGAG	3.9		chr8:82102886	-1	20136985	5 FALSE
ACAGAACTACACACAGGCGG	TGGGG	3.8		chr11:75925035	1	52102893	4 FALSE
TGTGAGCTACGCGCCAGGCGG	TGGAG	3.8		chr10:27344725	-1	75925035	4 FALSE
ACAGAGCTGAGAGAGGCGG	CAGGG	3.3		chr10:134817137	1	27344725	5 FALSE
TCAGATCTGAGTCCAGGCGG	ATGGA	3.8		chr15:28948038	1	134817137	5 FALSE
TCAGAGCTTCTTACAGGCGG	ATGGG	0.7		chr1:47888886	-1	28948038	5 FALSE
ACAGAGCTTCTTACAGGCGG	TGGGG	3.7	TSC2 (R800000103187)	chr15:2082581	-1	47888886	5 FALSE
ACAGAGCTTACAGGCGGCGG	AGGAG	3.7		chr7:25575832	1	2082581	5 FALSE
ACAGAGCTAGAGAGGAGGCG	AGGAG	0.7		chr15:54347353	-1	25575832	4 FALSE
TCAGAGCTTACAGGAGGCGG	GGGAA	3.7		chr2:43112424	1	54347353	4 FALSE
GCAGAGCTTACAGGAGGCGG	AGGAA	3.7		chr18:88583864	-1	43112424	5 FALSE
TCAGAGCTTACAGGAGGCGG	AGGAA	0.7		chr15:80881252	1	88583864	5 FALSE
GTAGAGAGTACAGGAGGCGG	TGGGA	3.9		chr10:103794275	-1	80881252	5 FALSE
ACAGAGCTTACAGGAGGCGG	TGGGG	3.5		chr10:57542558	-1	103794275	5 FALSE
TCAGAGCTTACAGGAGGCGG	AGGAG	0.8		chr11:121718784	1	57542558	5 FALSE
ACAGAGCTTACAGGAGGCGG	CAGGA	3.8		chr13:47252260	-1	121718784	5 FALSE
ACAGAGCTTACAGGAGGCGG	TGGGG	3.5		chr2:10844304	-1	47252260	5 FALSE
GCAGAGCTTACAGGAGGCGG	TGGGG	0.8		chr11:78255155	-1	10844304	4 FALSE
TCAGAGCTTACAGGAGGCGG	CAGGA	3.5		chr2:23844805	-1	78255155	3 FALSE
ACAGAGCTTACAGGAGGCGG	AGGGA	3.5		chr2:107589867	-1	23844805	5 FALSE
ACAGAGCTTACAGGAGGCGG	CAGGA	0.8		chr6:154249635	-1	107589867	4 FALSE
GCAGAGCTTACAGGAGGCGG	TGGGG	3.5		chr10:11040802	-1	154249635	5 FALSE
GCAGAGCTTACAGGAGGCGG	CTGGG	3.5		chr1:41442886	1	11040802	4 FALSE
GCAGAGCTTACAGGAGGCGG	AGGAG	0.8		chr11:46723714	-1	41442886	4 FALSE
GCAGAGCTTACAGGAGGCGG	TAGGA	3.4		chr6:13158080	-1	46723714	5 FALSE
GCAGAGCTTACAGGAGGCGG	TGGGG	3.4		chr22:4898738	1	13158080	4 FALSE
GCAGAGCTTACAGGAGGCGG	AGGAA	0.8		chr20:43864342	-1	4898738	4 FALSE
ACAGAGCTTACAGGAGGCGG	CTGAA	3.4		chr3:181013315	-1	43864342	4 FALSE
ACAGAGCTTACAGGAGGCGG	CAGGG	3.4	TUB (R800000204158482)	chr11:8101357	-1	181013315	5 FALSE
GCAGAGCTTACAGGAGGCGG	AGGAG	0.8		chr7:47512275	1	8101357	5 FALSE
GCAGAGCTTACAGGAGGCGG	AGGCG	3.4		chr8:88146884	-1	47512275	5 FALSE
TCAGAGCTTACAGGAGGCGG	CTGGG	3.4		chr9:131113315	-1	88146884	4 FALSE
GCAGAGCTTACAGGAGGCGG	AGGCG	0.8		chr2:5814078	1	131113315	5 FALSE
TCAGAGCTTACAGGAGGCGG	CAGGG	3.3		chr2:120441825	-1	5814078	4 FALSE
GCAGAGCTTACAGGAGGCGG	AGGAA	3.3		chr10:9587680	-1	120441825	5 FALSE
GCAGAGCTTACAGGAGGCGG	ATGGG	0.8		chr2:88175215	-1	9587680	4 FALSE
GCAGAGCTTACAGGAGGCGG	CTGAA	3.3		chr10:43903971	1	88175215	4 FALSE
GCAGAGCTTACAGGAGGCGG	TGGGG	3.3		chr9:5134667	1	43903971	4 FALSE
GCAGAGCTTACAGGAGGCGG	AGGAG	0.8	TBC1D19	chr4:28575715	1	5134667	5 FALSE
GCAGAGCTTACAGGAGGCGG	CTGGA	3.2	ACT1 (R8000001002100)	chr18:15285258	1	28575715	4 FALSE
GCAGAGCTTACAGGAGGCGG	CTGGA	3.2		chr9:13057800	-1	15285258	5 FALSE
GCAGAGCTTACAGGAGGCGG	CTGGA	0.8	GLC4A2	chr12:184858185	1	13057800	5 FALSE
GCAGAGCTTACAGGAGGCGG	CTGGA	3.2	ACT1 (R8000001002100)	chr1:203423454	-1	184858185	5 FALSE
GCAGAGCTTACAGGAGGCGG	TGGGG	3.2		chr10:48193411	-1	203423454	5 FALSE
GCAGAGCTTACAGGAGGCGG	CTGGG	0.8		chr20:41063877	-1	48193411	5 FALSE
GCAGAGCTTACAGGAGGCGG	TGGGA	3.2		chr11:5419275	1	41063877	4 FALSE
GCAGAGCTTACAGGAGGCGG	TGGGA	3.2		chr11:519505	-1	5419275	5 FALSE

[0230] Table 3

HIV-1 Seq ID target (+strand)									
GGATAGATGTAAGACACACCAAGGAAG									
Sequence	PAM	Score	Gene	Chromosome	Strand	Position	Mismatch	On-target	
AGAAAGATGTAAAGACACCT	TGGAA	1.7		chr2:144745442					
TTATACATTATTAAGACACCA	AAGAA	1.5		chr1:194738518	-1	1.45E+08	4	FALSE	
GAATAGATGTAAAGACACCA	GGGGA	1.5		chr16:48642775	-1	1.35E+08	5	FALSE	
TTTACACTTAAGACACACCA	TTGAA	1		chr15:31959865	-1	48934775	3	FALSE	
AGATTGATTTAAAGACACCA	TAGAG	1		chr5:32802184	-1	32059865	5	FALSE	
GGATAGATGTAAAGACACCA	AAGAA	1		chr17:51350241	-1	32002168	5	FALSE	
AAATAGATTTAAAGACACCA	GGGAA	0.8		chr2:175764845	1	51350241	3	FALSE	
AGATTATGTAAAGACACCA	TGGAA	0.5		chr12:44452185	1	1.78E+08	5	FALSE	
TTTACAGTTAAAGACACCA	GGGAG	0.8		chr12:197754848	1	1.44E+08	4	FALSE	
TTTAAATCAACACACACCA	TAAGA	0.3		chr7:141715550	-1	1.89E+08	4	FALSE	
GAAGAGATTAAAGACACCA	AAGAG	0.3		chr2:113158138	-1	1.42E+08	4	FALSE	
AGATAGATTTAAAGACACCA	TAGAG	0.8		chr18:28642157	-1	2.13E+08	4	FALSE	
ATKAGATGTAAAGACACCA	GGGAA	0.8		chr11:28725508	-1	29342353	3	FALSE	
GGTGGATTTAAAGACACCA	TTGAG	0.7		chr12:14366627	-1	2.27E+08	5	FALSE	
GGTACATTTAAAGACACCA	TAGAG	0.7		chr12:38645105	-1	1.43E+08	4	FALSE	
GGATAGATTTAAAGACACCA	GTGAA	0.3		chr15:43602813	-1	38643185	5	FALSE	
GGATATGTGTAAGACACCA	TAGAA	0.5		chr18:51785710	1	43665283	5	FALSE	
AGAAAGAGTAAGACACCA	GGGAG	0.3		chr2:278083925	1	51783716	3	FALSE	
TTATAGTTAAAGACACCA	TTGAA	0.3		chr12:129151012	-1	1.18E+08	5	FALSE	
CCAGATTTAAAGACACCA	GGGAA	0.3	TACD3	chr10:125211347	1	1.26E+08	5	FALSE	
TTATAGTGAAGACACCC	ATGAA	0.3		chr14:48437726	-1	1.22E+08	5	FALSE	
GGTTCAGTAAGACACCA	TTGAA	0.3		chr16:50625737	1	48407726	3	FALSE	
GAAGATGTAAAGACACCA	AAGCA	0.5		chr2:123665433	1	50683737	4	FALSE	
TTATGCAATACACACACCA	TGCGA	0.5		chr12:10750458	1	2.25E+08	3	FALSE	
GAATAGTTGTAAAGACACCA	GTGAA	0.3		chr12:15945004	-1	1.08E+08	5	FALSE	
AGAAAGATTAAGACACCA	GGGAG	0.3		chr11:75323558	-1	1.36E+08	5	FALSE	
TGAAATTTTAAAGACACCA	GGGAG	0.3	ADMP (ENSG00000101126)	chr20:50685247	-1	75273565	5	FALSE	
AGATTGTGTAAAGACACCA	CAGAG	0.3		chr7:114527485	-1	1.15E+08	5	FALSE	
GGAAATTTAAAGACACCA	GGGAA	0.3		chr17:38703771	1	1.15E+08	5	FALSE	
GAATGATTTAAAGACACCA	GTGAA	0.3		chr2:65758275	-1	59103371	4	FALSE	
AAATAGTTTAAAGACACCA	GTGAG	0.3		chr12:65758275	-1	65735276	3	FALSE	
TGTAAGATTAAAGACACCA	TAGAG	0.3		chr16:5117833	-1	65735945	5	FALSE	
GAATGATTTAAAGACACCA	GTGAA	0.3		chr12:79875880	-1	3137938	3	FALSE	
GAATAGTTTAAAGACACCA	GTGAA	0.3		chr12:79875880	-1	79875882	4	FALSE	
AGATAGTTTAAAGACACCA	GTGAA	0.3		chr12:151673060	-1	1.32E+08	5	FALSE	
AGATAGTTTAAAGACACCA	GTGAA	0.4		chr12:152680545	1	57369345	5	FALSE	
TGAAAGATTAAAGACACCA	AAGGA	0.4		chr20:24136741	-1	24138791	5	FALSE	
GAATAGTTTAAAGACACCA	AGGAA	0.4		chr12:62718055	-1	62713605	4	FALSE	
TGAATTTTAAAGACACCA	AGGAG	0.4		chr2:3257648	1	3551548	5	FALSE	
GGTTAGCGTAAGACACCA	CAGGG	0.4		chr14:141535240	-1	1.42E+08	4	FALSE	
GGTTAGCGTAAGACACCA	CAGGG	0.4		chr14:141591639	1	1.42E+08	4	FALSE	
GGTTAGCGTAAGACACCA	CAGGG	0.4		chr14:141582855	-1	1.42E+08	4	FALSE	
GGTTAGCGTAAGACACCA	CAGGG	0.4		chr14:141580587	-1	1.41E+08	4	FALSE	
GGTTAGCGTAAGACACCA	CAGGG	0.4		chr14:141580587	-1	1.41E+08	4	FALSE	
GGATTATGTAAGACACCA	TAGGG	0.4		chr12:85730118	-1	85730118	4	FALSE	
AGAAATTTAAAGACACCA	AGGAG	0.4		chr17:122068645	1	1.73E+08	5	FALSE	
GGAAAGTAAGACACCA	GAAGG	0.4		chr17:81470353	-1	81470363	4	FALSE	
AGATAGTTTAAAGACACCA	AAGAA	0.4		chr12:103514867	-1	1.1E+08	5	FALSE	
AGAAATTTAAAGACACCA	TAGGA	0.3		chr2:123665433	1	2.12E+08	5	FALSE	
GAATAGTAAGACACCA	GTGAA	0.3		chr2:141095048	-1	1.41E+08	5	FALSE	
TTATAGTAAGACACCA	GTGAG	0.3		chr4:146453572	1	1.46E+08	5	FALSE	

[0231] Table 4: HIV-1 LTR1 target single cell clone off target analysis.

OFF TARGET	Target sequence: <i>LTR1</i> <i>PAM</i> GCAGAACTACACACCAAGGCG	Chromosome location/gene	Strand	Position	Score	Mismatches	Single cell clone	Number of sequences analyzed	Indels detected
	Predicted off target sequence:								
1	ACAGACTGCCCACCAAGGCT	Ch 16/TSC2	-	2082981	0.7	5	TOTAL	26	0
							CTRL1	5	0
							CTRL2	3	0
							ERAD1	3	0
							ERAD2	3	0
							ERAD3	3	0
							ERAD4	3	0
							ERAD5	3	0
							ERAD6	3	0
2	ACAGACTGACAGACCAAGGCT	Ch 11/TUB	-	8105357	0.4	5	TOTAL	27	0
							CTRL1	3	0
							CTRL2	3	0
							ERAD1	3	0
							ERAD2	3	0
							ERAD3	3	0
							ERAD4	5	0
							ERAD5	3	0
							ERAD6	3	0
3	CCAGACCGAGACACCAAGGCG	Ch 12/SLC41A2	+	104958165	0.2	5	TOTAL	6	0
							CTRL1		
							CTRL2		
							ERAD1		
							ERAD2	1	0
							ERAD3	2	0
							ERAD4		
							ERAD5	2	0
							ERAD6	1	0

[0232] Table 5 HIV-1 GagD target single cell clone off target analysis

OFF TARGET	Target sequence: <i>gagD</i> FAM GGTACATGTATAAGACACCT	Chromosome location/gene	Strand	Position	Score	Mismatches	Single cell clone	Number of sequences analyzed	Indels detected
	Predicted off target sequence:								
1	CCAAAGATGAAAAAGACACCT	Ch 10/TACC2	+	122211347	0.6	5	TOTAL	34	0
							CTRL1	3	0
							CTRL2	3	0
							ERAD1	6	0
							ERAD2	6	0
							ERAD3	3	0
							ERAD4	4	0
							ERAD5	4	0
							ERAD6	5	0
2	TGAGACTTGTACAAGACACCT	Ch 20/ADNP	-	50889247	0.5	5	TOTAL	23	0
							CTRL1	2	0
							CTRL2	3	0
							ERAD1	3	0
							ERAD2	3	0
							ERAD3	3	0
							ERAD4	3	0
							ERAD5	3	0
							ERAD6	3	0
3	AGAAATGTATAAGACACCT	Ch 3/non gene	-	144746442	1.7	4	TOTAL	24	0
							CTRL1	3	0
							CTRL2	3	0
							ERAD1	3	0
							ERAD2	3	0
							ERAD3	3	0
							ERAD4	3	0
							ERAD5	3	0
							ERAD6	3	0

[0233] Table 6: Number of somatic SNPs in different genomic regions

Sample	#4349	#4345	#4355
CDS	6590	6760	5895
Synonymous_SNP	3003	3053	2588
Missense_SNP	3372	3465	3050
Stopgain	70	86	128
Stoploss	6	6	7
Unknown	140	151	122
Intronic	414960	415812	382465
UTR3	8195	8197	7713
UTR5	1905	1947	1649
Splicing	23	27	30
ncRNA_exonic	3982	4073	3509
ncRNA_intronic	65272	65664	59525
ncRNA_splicing	14	14	13
Upstream	6962	7052	6063
Downstream	7822	7850	7005
Intergenic	664364	669294	602974
Total	1180363	1186971	1077074

Sample: sample name

CDS: the number of somatic SNPs in coding region

Synonymous_SNP: a single nucleotide change that does not cause an amino acid change

Missense_SNP: a single nucleotide change that causes an amino acid change

Stopgain: a nonsynonymous SNP that leads to the immediate creation of stop codon at the variant site

Stoploss: a nonsynonymous SNP that leads to the immediate elimination of stop codon at the variant site

Unknown: unknown function (due to various errors in the gene structure definition in the database file)

Intronic: the number of somatic SNPs in intronic region

UTR3: the number of somatic SNPs in 3'UTR region

UTR5: the number of somatic SNPs in 5'UTR region

Splicing: the number of somatic SNPs within 2-bp of a splicing junction

ncRNA_exonic: the number of somatic SNPs in exonic region of non-coding RNAs

ncRNA_intronic: the number of somatic SNPs in intronic region of non-coding RNAs

ncRNA_splicing: the number of somatic SNPs within 2-bp of a splicing junction of non-coding RNAs

Upstream: the number of somatic SNPs within 1kb away from the transcription start site

Downstream: the number of somatic SNPs within the 1kb away from the transcription termination site

Intergenic: the number of somatic SNPs in intergenic region

Total: the total number of somatic SNPs

[0234] Table 7: Number of somatic InDels in different genomic regions

Sample	JAK2-10	JAK2-11	JAK2-12
CDS	103	124	59
Frameshift_deletion	31	34	21
Frameshift_insertion	14	16	10
Nonframeshift_deletion	36	48	16
Nonframeshift_insertion	18	22	9
Stopgain	2	2	1
Stoploss	0	0	0
Unknown	2	2	2
Intronic	36969	39727	25080
UTR3	946	1003	640
UTR5	134	149	89
Splicing	5	5	3
ncRNA_exonic	285	314	190
ncRNA_intronic	5794	6247	3971
ncRNA_splicing	2	3	1
Upstream	694	771	417
Downstream	879	958	602
Intergenic	56749	61117	38406
Total	102588	110452	69477

Sample: sample name

CDS: the number of somatic InDels in coding region

Frameshift_deletion: a deletion of one or more nucleotides that cause frameshift changes in protein coding sequence

Frameshift_insertion: an insertion of one or more nucleotides that cause frameshift changes in protein coding sequence

Nonframeshift_deletion: a deletion that does not cause frameshift changes

Nonframeshift_insertion: an insertion that does not cause frameshift changes

Stopgain: an insertion or a deletion that leads to the immediate creation of stop codon at the variant site

Stoploss: an insertion or a deletion that leads to the immediate elimination of stop codon at the variant site

Unknown: unknown function (due to various errors in the gene structure definition in the database file)

Intronic: the number of somatic InDels in intronic region

UTR3: the number of somatic InDels in 3'UTR region

UTR5: the number of somatic InDels in 5'UTR region

Splicing: the number of somatic InDels within 2-bp of a splicing junction

ncRNA_exonic: the number of somatic InDels in exonic region of non-coding RNAs

ncRNA_intronic: the number of somatic InDels in intronic region of non-coding RNAs

ncRNA_splicing: the number of somatic InDels within 2-bp of a splicing junction of non-coding RNAs

Upstream: the number of somatic InDels within 1kb away from transcription start site

Downstream: the number of somatic InDels iwithin 1kb away from transcription termination site

Intergenic: the number of somatic InDels in intergenic region

Total: the total number of somatic InDels

[0235] References:

1. G. Hutter *et al.*, Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *The New England journal of medicine* **360**, 692-698 (2009).
2. W. Xu *et al.*, Advancements in Developing Strategies for Sterilizing and Functional HIV Cures. *BioMed research international* **2017**, 6096134 (2017).
3. A. Saez-Cirion *et al.*, Post-treatment HIV-1 controllers with a long-term virological remission after the interruption of early initiated antiretroviral therapy ANRS VISCONTI Study. *PLoS pathogens* **9**, e1003211 (2013).
4. J. Z. Li *et al.*, The size of the expressed HIV reservoir predicts timing of viral rebound after treatment interruption. *AIDS (London, England)* **30**, 343-353 (2016).
5. J. D. Siliciano, R. F. Siliciano, Recent developments in the effort to cure HIV infection: going beyond N = 1. *The Journal of clinical investigation* **126**, 409-414 (2016).
6. A. R. Martin, R. F. Siliciano, Progress Toward HIV Eradication: Case Reports, Current Efforts, and the Challenges Associated with Cure. *Annual review of medicine* **67**, 215-228 (2016).
7. D. Singh *et al.*, Development and characterization of a long-acting nanoformulated abacavir prodrug. *Nanomedicine (London, England)* **11**, 1913-1927 (2016).
8. D. Guo *et al.*, Creation of a Long-Acting Nanoformulated 2',3'-Dideoxy-3'-Thiacytidine. *Journal of acquired immune deficiency syndromes (1999)* **74**, e75-e83 (2017).
9. B. Edagwa, *et al.*, Long-acting slow effective release antiretroviral therapy. *Expert opinion on drug delivery*, 1-11 (2017).
10. P. K. Dash *et al.*, Long-acting nanoformulated antiretroviral therapy elicits potent antiretroviral and neuroprotective responses in HIV-1-infected humanized mice. *AIDS (London, England)* **26**, 2135-2144 (2012).
11. Hu *et.al*, RNA-directed gene editing specifically eradicates latent and prevents new HIV-1 infection. *Proc.Natl.Acad.Sci. USA* **111**,11461-11466 (2014)
12. R. Kaminski *et al.*, Elimination of HIV-1 Genomes from Human T-lymphoid Cells by CRISPR/Cas9 Gene Editing. *Scientific reports* **6**, 22555 (2016).
13. R. Kaminski *et al.*, Negative Feedback Regulation of HIV-1 by Gene Editing Strategy. *Scientific reports* **6**, 31527 (2016).
14. M. K. White, W. Hu, K. Khalili, Gene Editing Approaches against Viral Infections and Strategy to Prevent Occurrence of Viral Escape. *PLoS pathogens* **12**, e1005953 (2016).
15. R. Kaminski *et al.*, Excision of HIV-1 DNA by gene editing: a proof-of-concept in vivo study. *Gene therapy* **23**, 690-695 (2016).
16. C. Yin *et al.*, In Vivo Excision of HIV-1 Provirus by saCas9 and Multiplex Single-Guide RNAs in Animal Models. *Molecular therapy : the journal of the American Society of Gene Therapy* **25**, 1168-1186 (2017).

17. M. Arainga, *et al.*, HIV-1 cellular and tissue replication patterns in infected humanized mice. *Scientific reports* **6**, 23513 (2016).
18. M. Arainga *et al.*, A mature macrophage is a principal HIV-1 cellular reservoir in humanized mice after treatment with long acting antiretroviral therapy. *Retrovirology* **14**, 17 (2017).
19. R. C. Gallo, Shock and kill with caution. *Science (New York, N.Y.)* **354**, 177-178 (2016).
20. S. N. Byraredy *et al.*, Sustained virologic control in SIV+ macaques after antiretroviral and alpha4beta7 antibody therapy. *Science (New York, N.Y.)* **354**, 197-202 (2016).
21. S. Gorantla *et al.*, Human immunodeficiency virus type 1 pathobiology studied in humanized BALB/c-Rag2^{-/-}-gammac^{-/-} mice. *Journal of virology* **81**, 2700-2712 (2007).
22. S. Gorantla *et al.*, Links between progressive HIV-1 infection of humanized mice and viral neuropathogenesis. *Am J Pathol* **177**, 2938-2949 (2010).
23. J. M. Jacobson *et al.*, Evidence that intermittent structured interruption, but not immunization with ALVAC-HIV vCP 1452, promotes control of HIV replication: the results of AIDS Clinical Trials Group. *J. Infect Dis.* **194**, 623-632 (2006).
24. J. B. Honeycutt *et al.*, HIH persistence in tissue macrophages of humanized myeloid-only mice during antiretroviral therapy. *Nat. Med.* **23**, 638-643 (2017).
25. R. H. Kutner *et al.*, Production, concentration and titration of pseudotyped HIV-1-based lentiviral vectors. *Nature Protocols* **4**, 495-505 (2009).
26. O'Doherty U *et al.*, Human Immunodeficiency Virus Type 1 spinoculation enhances infection through virus binding. *Journal of Virology* **74**(21), 10074-10080 (2000).
27. P. K. Dash *et al.*, Loss of neuronal integrity during progressive HIV-1 infection of humanized mice. *J Neurosci* **31**, 3148-3157 (2011).
28. Westervelt, *et al.*, Identification of a determinant within the human immunodeficiency virus 1 surface envelope glycoprotein critical for productive infection of primary monocytes. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 3097-3101 (1991).
29. G. Zhang *et al.*, The mixed lineage kinase-3 inhibitor URM-099 improves therapeutic outcomes for long-acting antiretroviral therapy. *Nanomedicine : nanotechnology, biology, and medicine* **12**, 109-122 (2016).
30. A. S. Nowacek *et al.*, NanoART synthesis, characterization, uptake, release and toxicology for human monocyte-macrophage drug delivery. *Nanomedicine (London, England)* **4**, 903-917 (2009).
31. L. M. Agosto *et al.*, HIV-1 integrates into resting CD4⁺ T cells even at low inoculums as demonstrated with an improved assay for HIV-1 integration. *Virology* **368**, 60-72 (2007).
32. A. O. Pasternak *et al.*, Highly sensitive methods based on seminested real-time reverse transcription-PCR for quantitation of human immunodeficiency virus type 1 unspliced and multiply spliced RNA and proviral DNA. *Journal of clinical microbiology* **46**, 2206-2211 (2008).

33. M. K. Liszewski, *et al*, Detecting HIV-1 integration by repetitive-sampling Alu-gag PCR. *Methods (San Diego, Calif.)* **47**, 254-260 (2009).
34. C. Deleage *et al.*, Defining HIV and SIV Reservoirs in Lymphoid Tissues. *Pathogens &*
35. M. J. Buzon *et al.*, Long-term antiretroviral treatment initiated at primary HIV-1 infection affects the size, composition, and decay kinetics of the reservoir of HIV-1-infected CD4 T cells. *Journal of virology* **88**, 10056-10065 (2014).
36. M. J. Buzon *et al.*, HIV-1 persistence in CD4+ T cells with stem cell-like properties. *Nature medicine* **20**, 139-142 (2014).
37. G. M. Laird *et al.*, Rapid quantification of the latent reservoir for HIV-1 using a viral outgrowth assay. *PLoS pathogens* **9**, e1003398 (2013).

[0236] The invention has been described in an illustrative manner, and it is to be understood that the terminology that has been used is intended to be in the nature of words of description rather than of limitation. Obviously, many modifications and variations of the present invention are possible in light of the above teachings. It is, therefore, to be understood that within the scope of the appended claims, the invention can be practiced otherwise than as specifically described.

CLAIMS

What is claimed:

1. A method of eradicating or eliminating a retrovirus in a subject, comprising administering to a patient a composition comprising a therapeutically effective amount of at least one antiretroviral agent and/or a composition comprising a therapeutically effective amount of at least one gene editing agent, thereby eradicating or eliminating the retrovirus in a subject.
2. The method of claim 1, wherein the antiretroviral agent is formulated as a long-acting slow effective release (LASER) antiretroviral agent.
3. The method of claim 2, wherein the at least one antiretroviral agent is nanoformulated.
4. The method of claim 3, wherein the at least one antiretroviral agent comprises: myristolyated dolutegravir, lamivudine, abacavir, rilpivirine or combinations thereof.
5. The method of claim 1, wherein the at least one antiretroviral agent is administered to the subject prior to administering the at least one gene editing agent.
6. The method of claim 1, wherein the at least one antiretroviral agent and at least one gene-editing agent are co-administered.
7. The method of claim 1, wherein the at least one antiretroviral agent and at least one gene-editing agent are administered sequentially.
8. The method of claim 1, wherein the at least one gene editing agent comprises: an isolated nucleic acid sequence encoding a Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR)-associated endonuclease/Cas (CRISPR/Cas) and at least one guide RNA (gRNA), the gRNA being complementary to a target nucleic acid sequence in a retroviral genome.
9. The method of claim 8, wherein the CRISPR/Cas fusion protein comprises catalytically deficient Cas protein (dCas), orthologs, homologs, mutants variants or fragments thereof.
10. The method of claim 8, wherein the at least one gRNA includes at least a first gRNA that is complementary to a target sequence in the integrated retroviral DNA; and a second

gRNA that is complementary to another target sequence in the integrated retroviral DNA, whereby the intervening sequences between the two gRNAs are removed.

11. The method of claim 8, wherein the isolated nucleic acid is included in at least one expression vector.
12. The method of claim of claim 11, wherein the expression vector comprises a lentiviral vector, an adenoviral vector, or an adeno-associated virus vector.
13. The method of claim 8, wherein the retrovirus is a human immunodeficiency virus (HIV).
14. The method of claim 13, wherein the target sequences comprise one or more nucleic acid sequences in HIV comprising: long terminal repeat (LTR) nucleic acid sequences, nucleic acid sequences encoding structural proteins, non-structural proteins or combinations thereof.
15. The method of claim 14, wherein the sequences encoding structural proteins comprise nucleic acid sequences encoding: Gag, Gag-Pol precursor, Pro (protease), Reverse Transcriptase (RT), integrase (In), Env or combinations thereof.
16. The method of claim 14, wherein the sequences encoding non-structural proteins comprise nucleic acid sequences encoding: regulatory proteins, accessory proteins or combinations thereof.
17. The method of claim 16, wherein regulatory proteins comprise: Tat, Rev or combinations thereof.
18. The method of claim 16, wherein accessory proteins comprise Nef, Vpr, Vpu, Vif or combinations thereof.
19. The method of claim 1, wherein a gRNA comprises at least one nucleic acid sequence set forth in Tables 1- 5 or combinations of gRNAs.
20. The method of claim 1, optionally comprising a therapeutically effective amount of a non-nucleoside reverse transcriptase inhibitor (NNRTI), and/or a nucleoside reverse transcriptase inhibitor (NRTI) and/or a protease inhibitor.
21. The method of claim 20, wherein the NNRTI comprises: etravirine, efavirenz, nevirapine, rilpivirine, delavirdine, or nevirapine.

22. The method of claim 20, wherein the NRTI comprises: lamivudine, zidovudine, emtricitabine, abacavir, zalcitabine, dideoxycytidine, azidothymidine, tenofovir disoproxil fumarate, didanosine (ddI EC), dideoxyinosine, stavudine, abacavir sulfate or combinations thereof.
23. The method of claim 20, wherein a protease inhibitor comprises: amprenavir, tipranavir, indinavir, saquinavir mesylate, lopinavir and ritonavir (LPV/RTV), Fosamprenavir Calcium (FOS-APV), ritonavir, darunavir, atazanavir sulfate, nelfinavir mesylate or combinations thereof.
24. A pharmaceutical composition comprising a therapeutically effective amount of a nanoformulated long-acting slow effective release antiretroviral agent.
25. The pharmaceutical composition of claim 24, wherein the nanoformulated antiretroviral agent comprises: myristoylated dolutegravir, lamivudine, abacavir, rilpivirine or combinations thereof.
26. The pharmaceutical composition of claim 24, further comprising at least one an isolated nucleic acid sequence encoding a Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR)-associated endonuclease; at least one isolated nucleic acid sequence encoding at least one guide RNA (gRNA) that is complementary to a target sequence in retroviral DNA; said isolated nucleic acid sequences being included in at least one expression vector.
27. The pharmaceutical composition of claim 26, wherein the integrated retroviral DNA is human immunodeficiency virus (HIV) DNA, and said at least one gRNA includes a first gRNA that is complementary to a first target sequence in the HIV DNA, and a second gRNA that is complementary to a second target sequence in the HIV DNA.

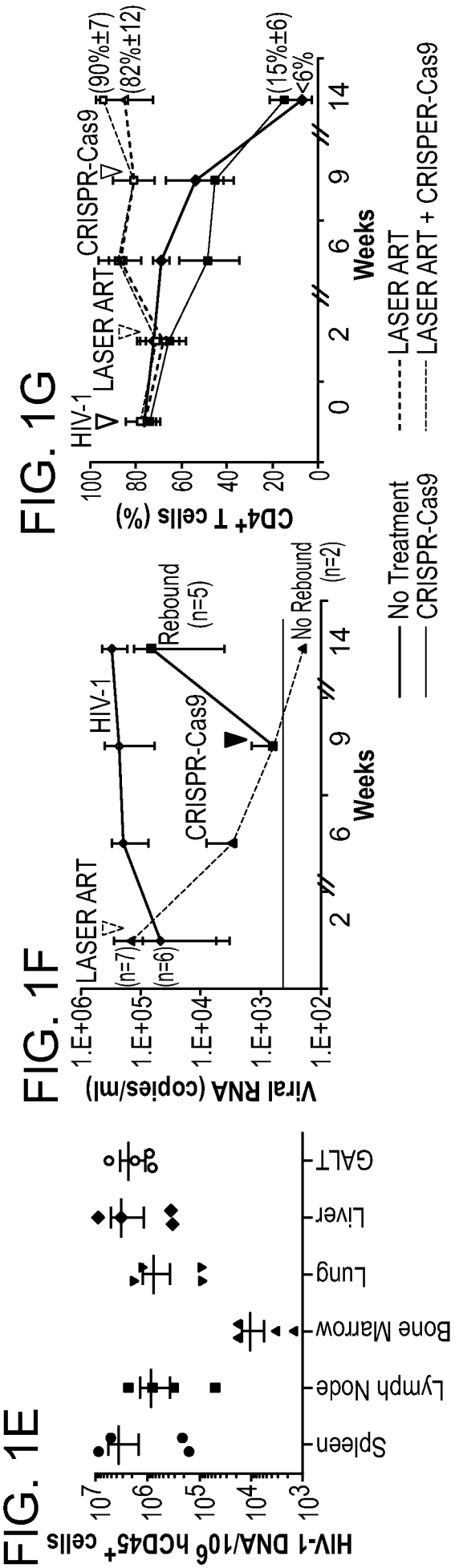
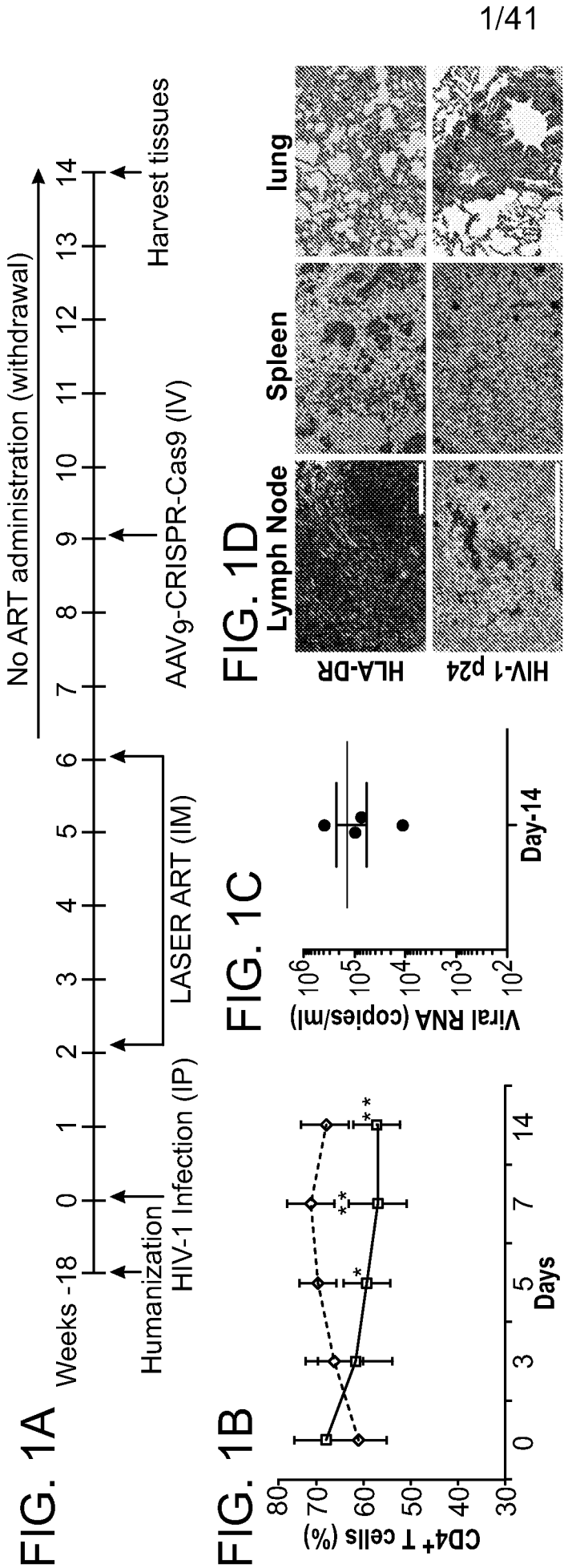
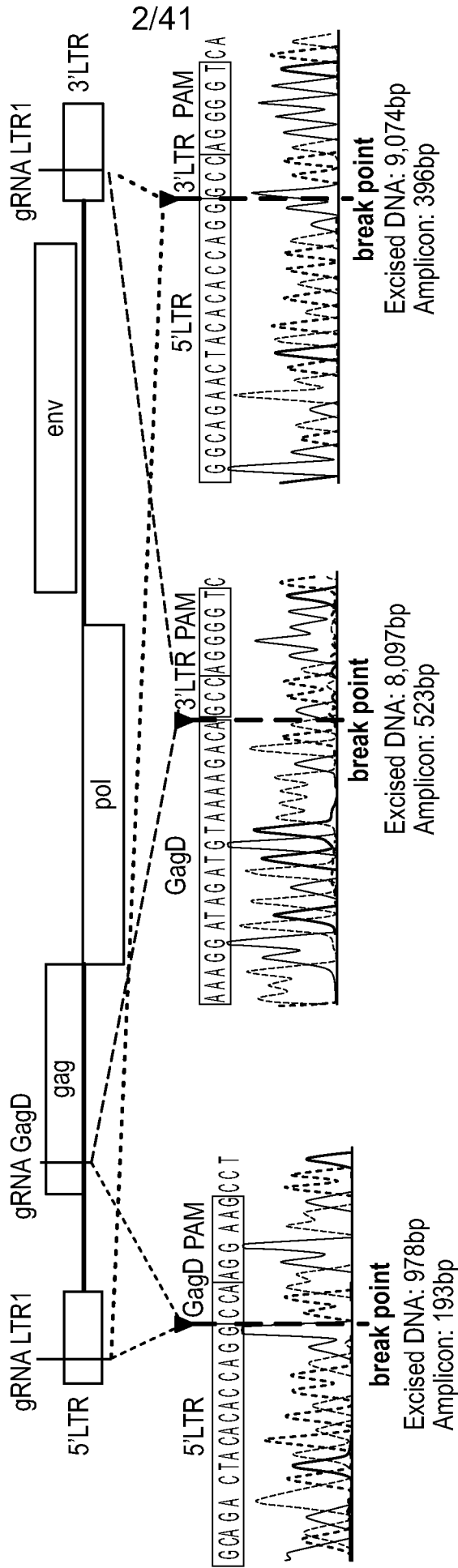


FIG. 2A



3/41

FIG. 2B

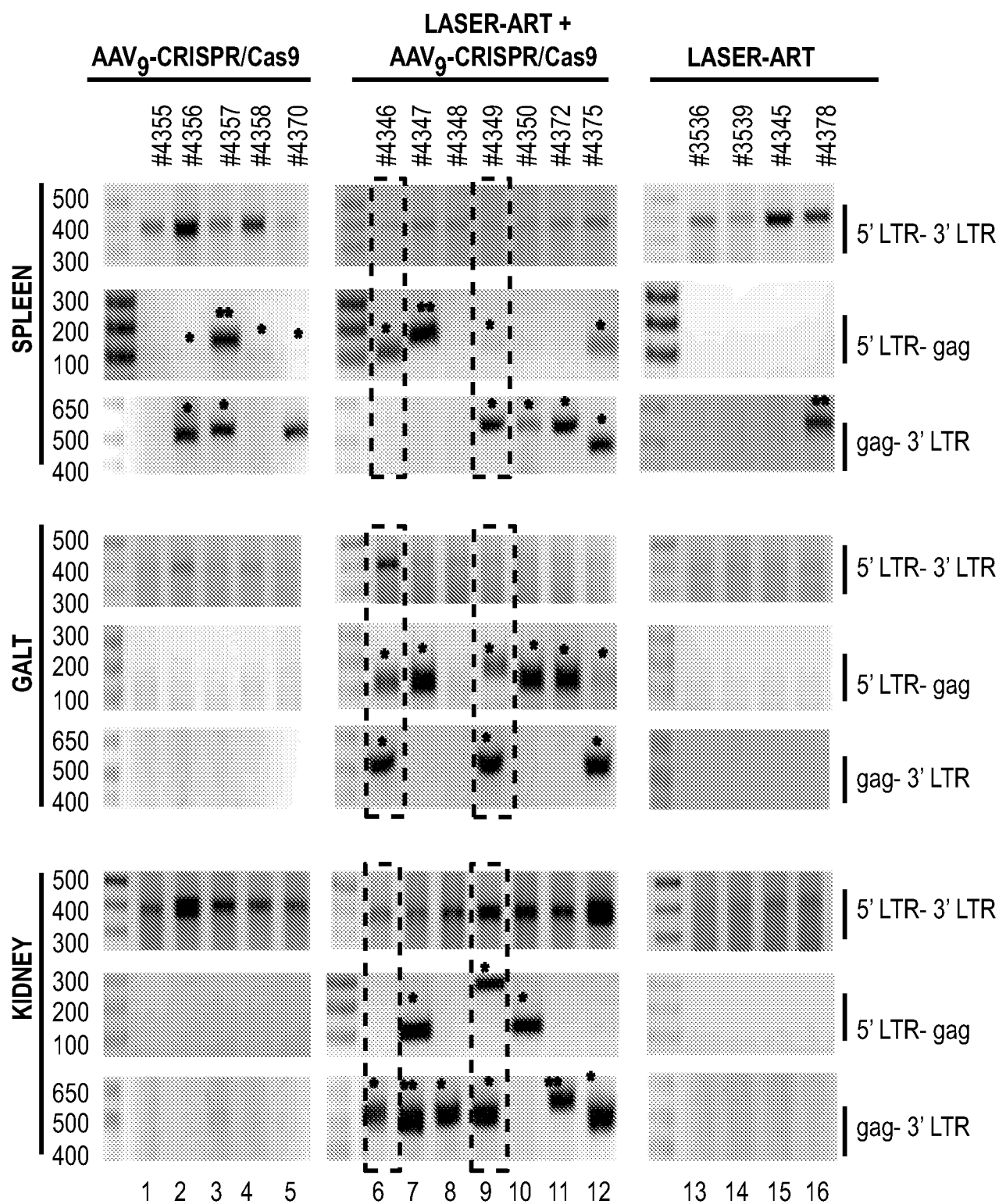
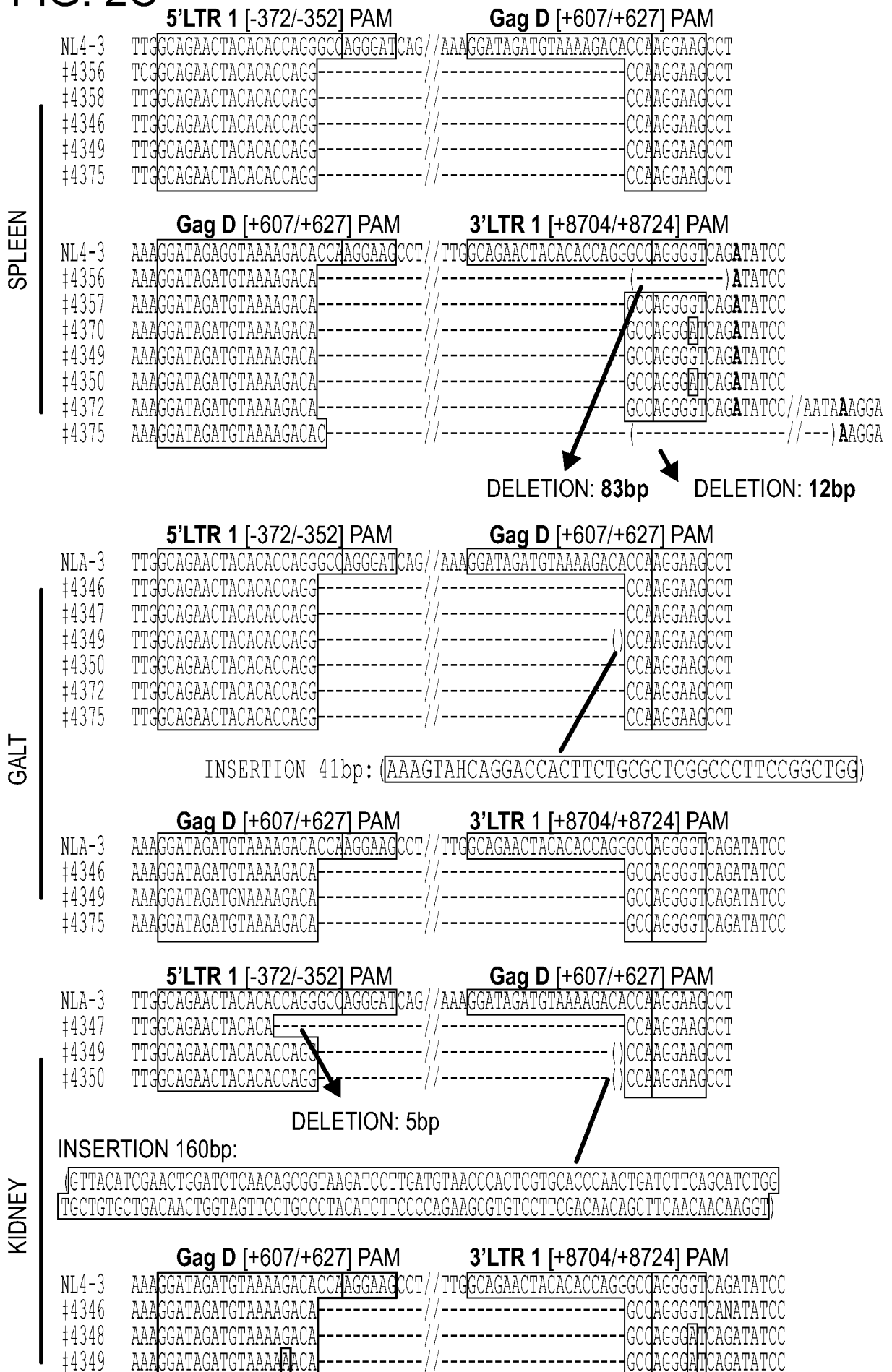
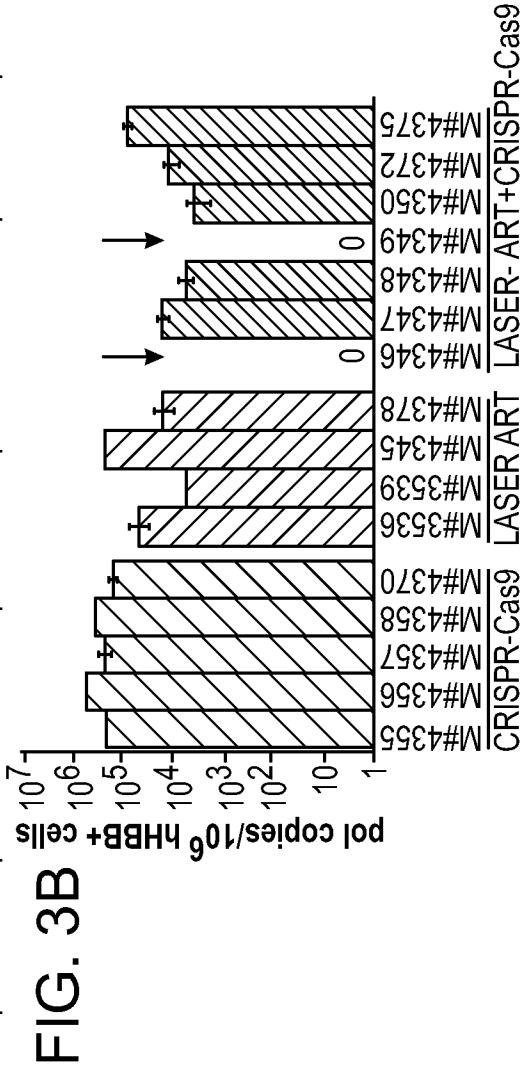
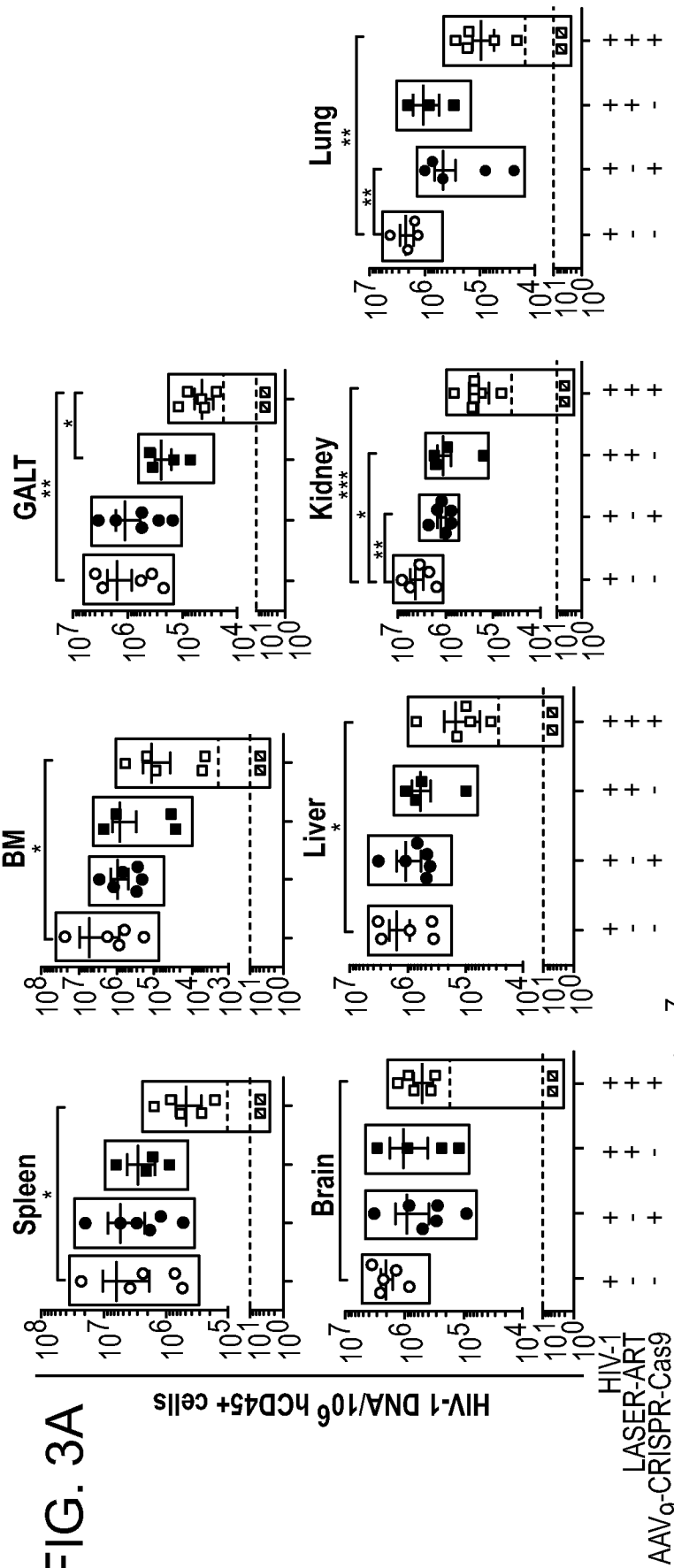
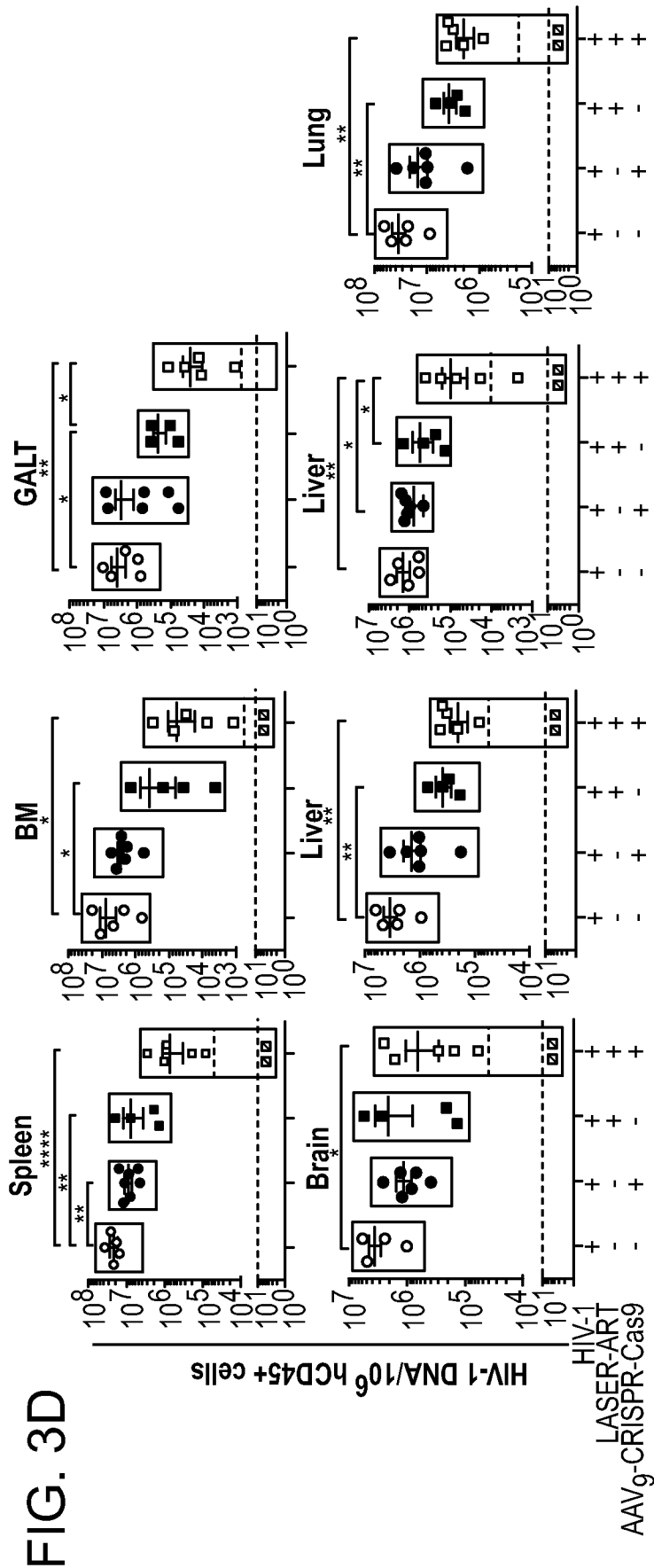
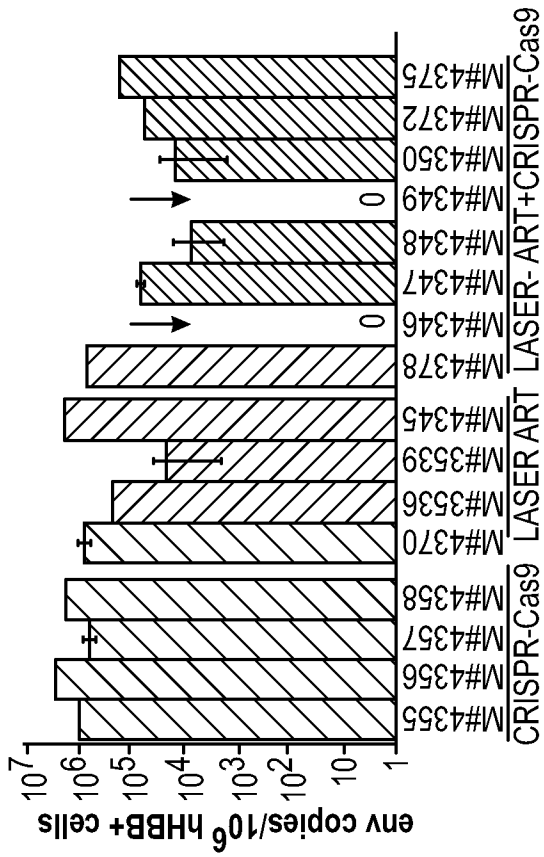


FIG. 2C

4/41







7/41

FIG 3E

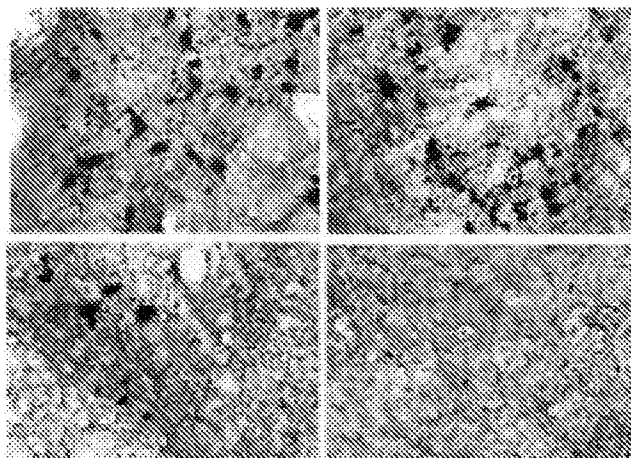


FIG. 4A

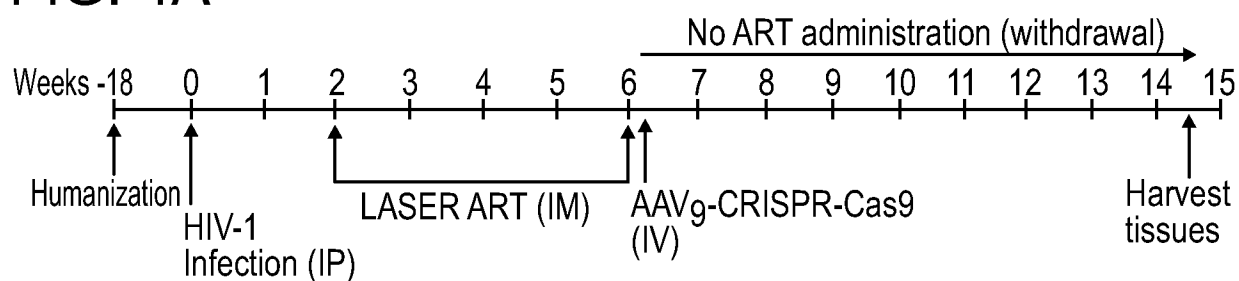


FIG. 4B

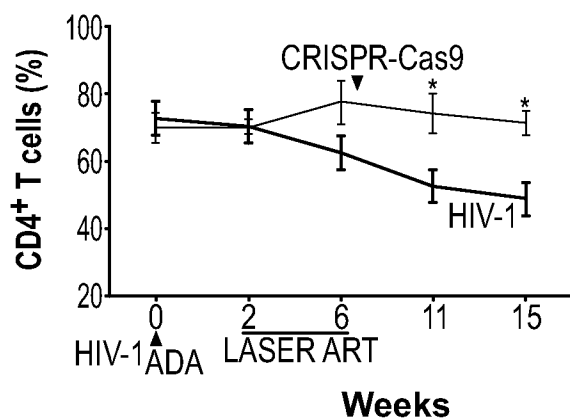


FIG. 4C

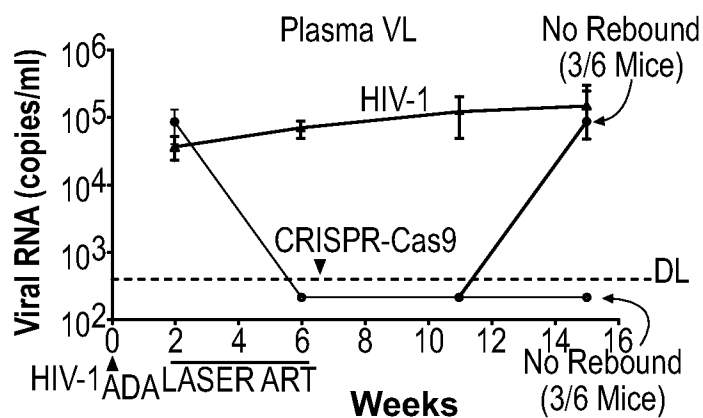


FIG. 4D

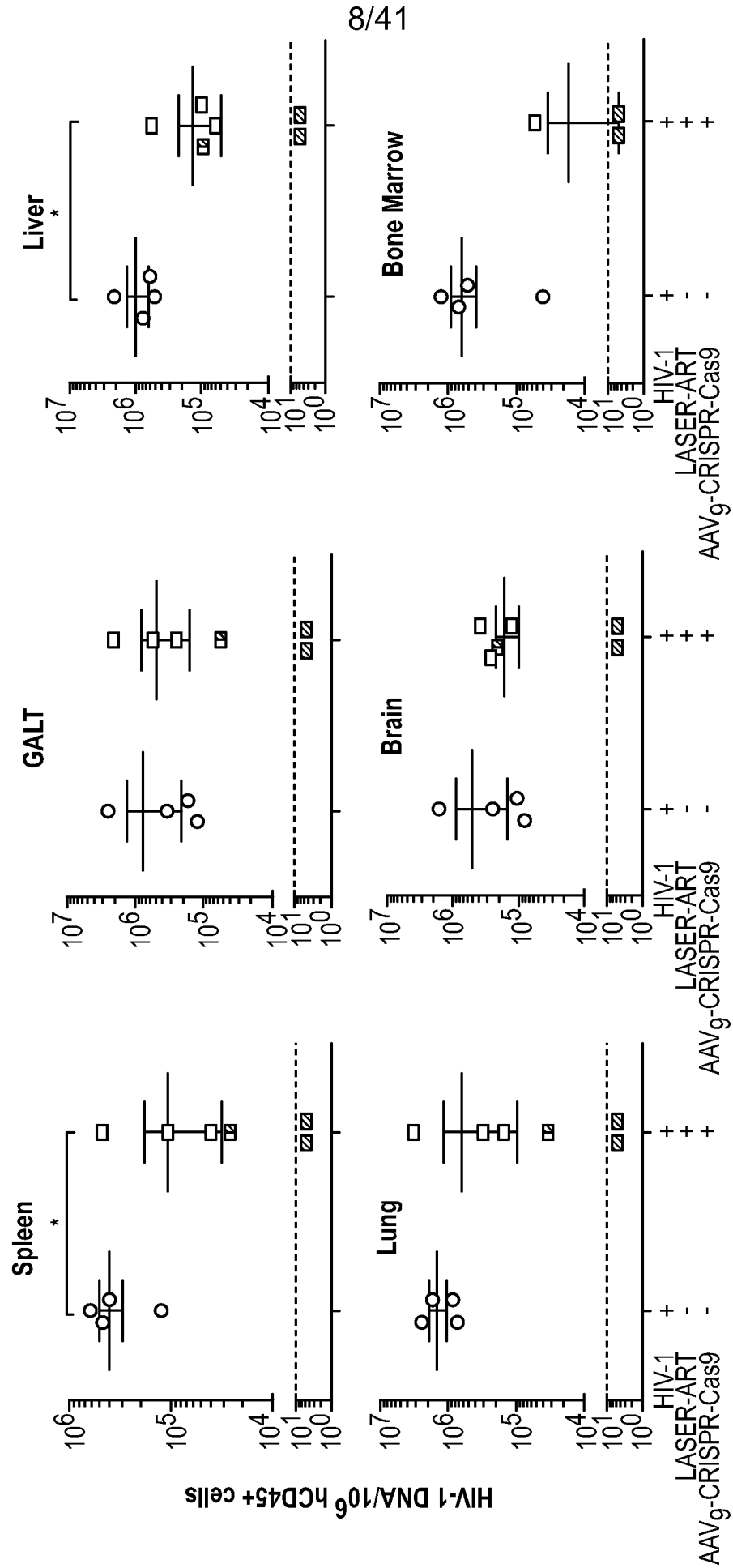


FIG. 4F

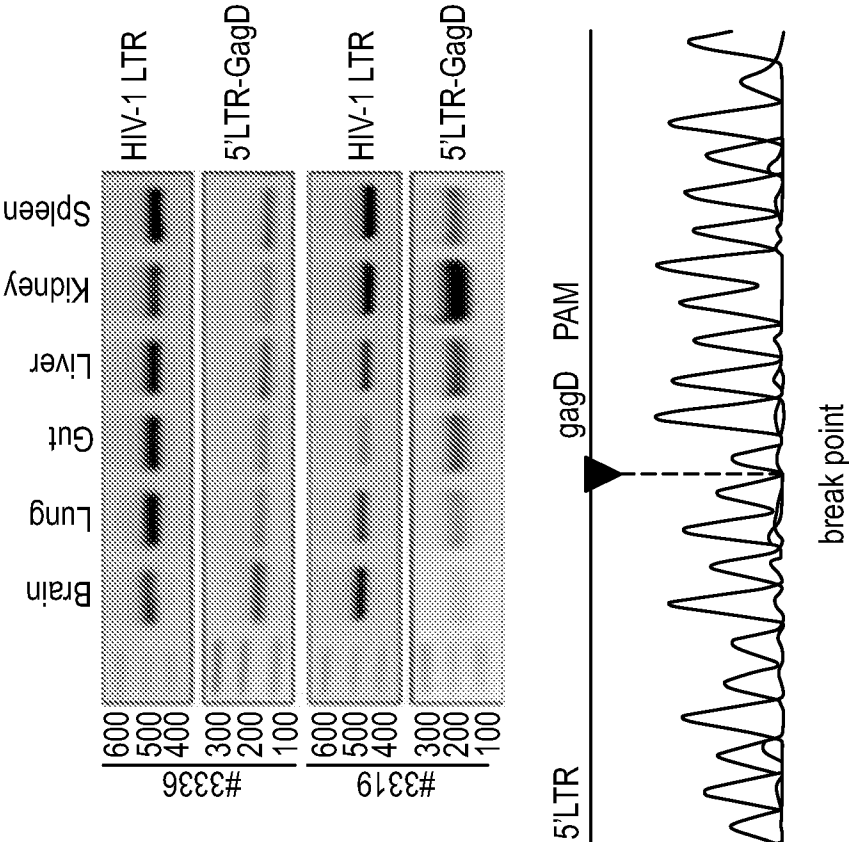


FIG. 4E

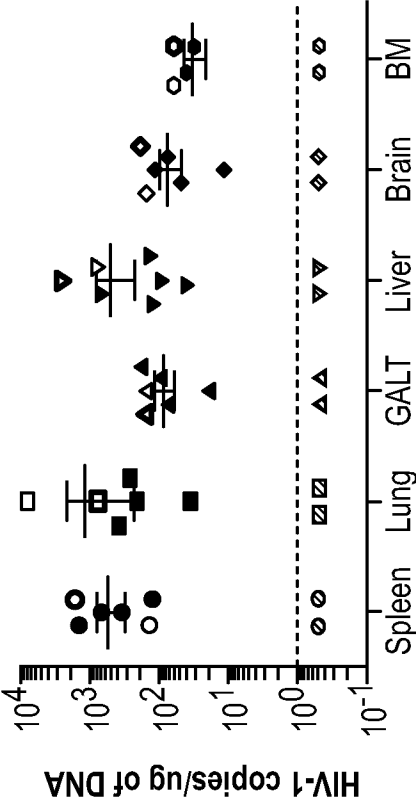


FIG. 4G

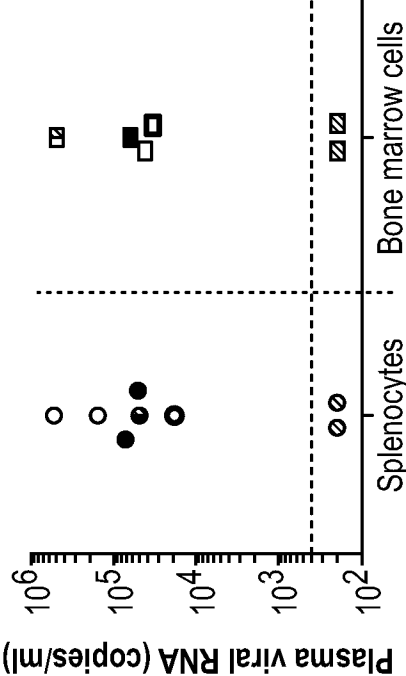


FIG. 5A

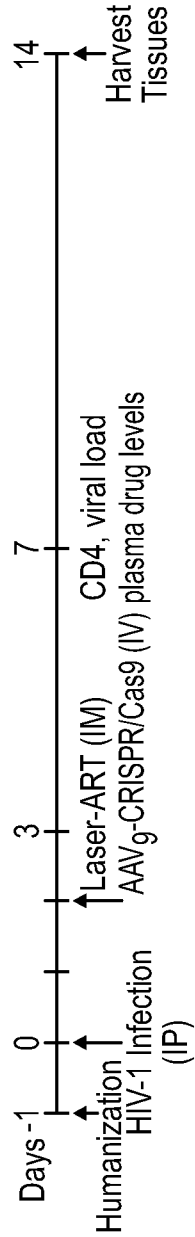


FIG. 5B

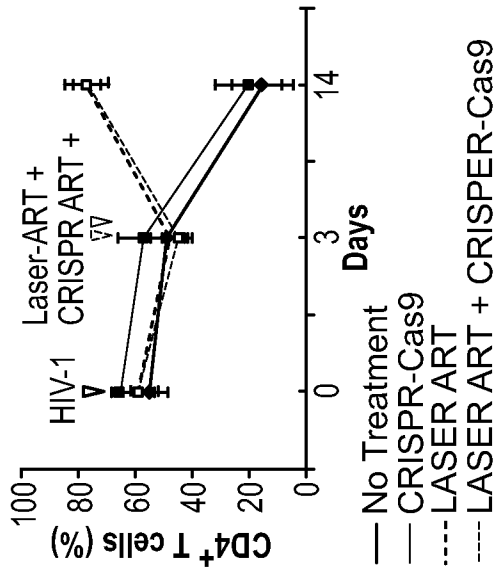


FIG. 5C

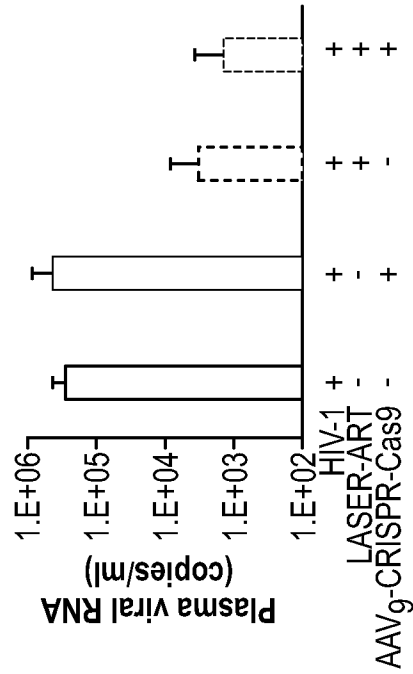


FIG. 5D

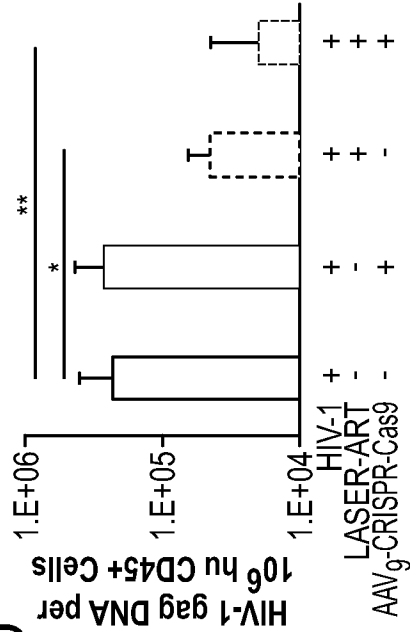


FIG. 5E

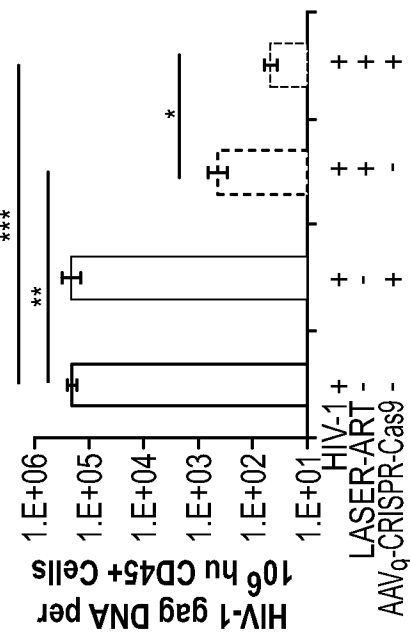


FIG. 5F

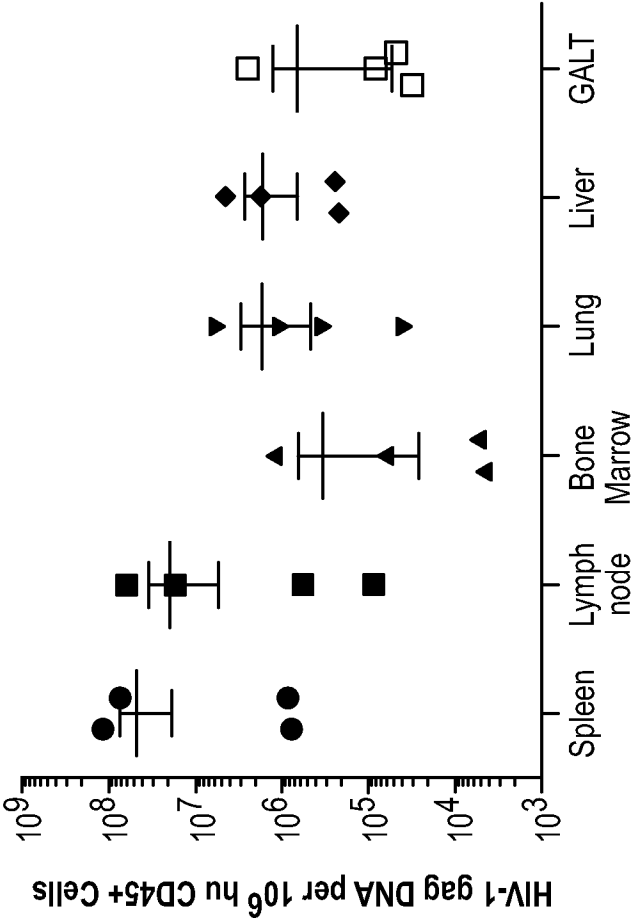
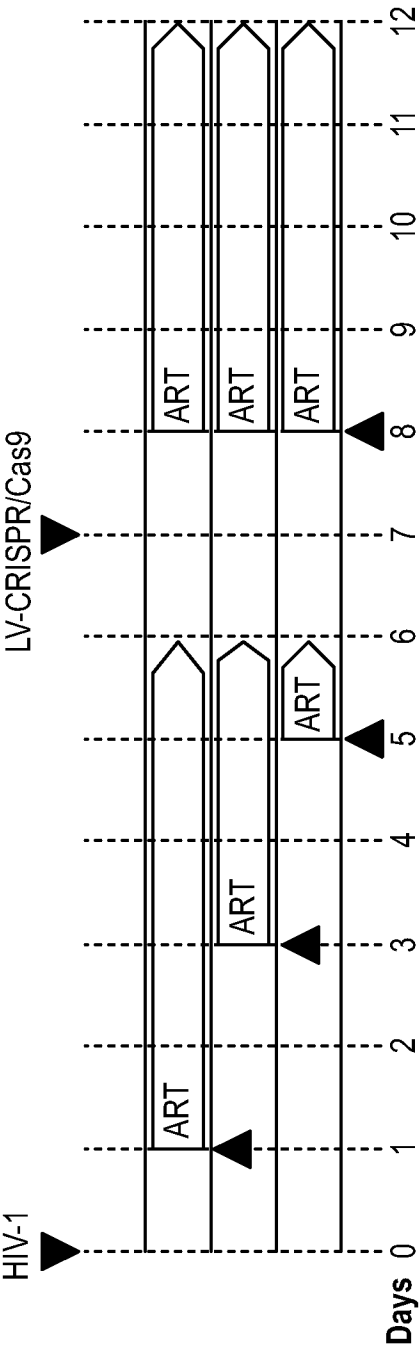


FIG. 6A



12/41

FIG. 6B

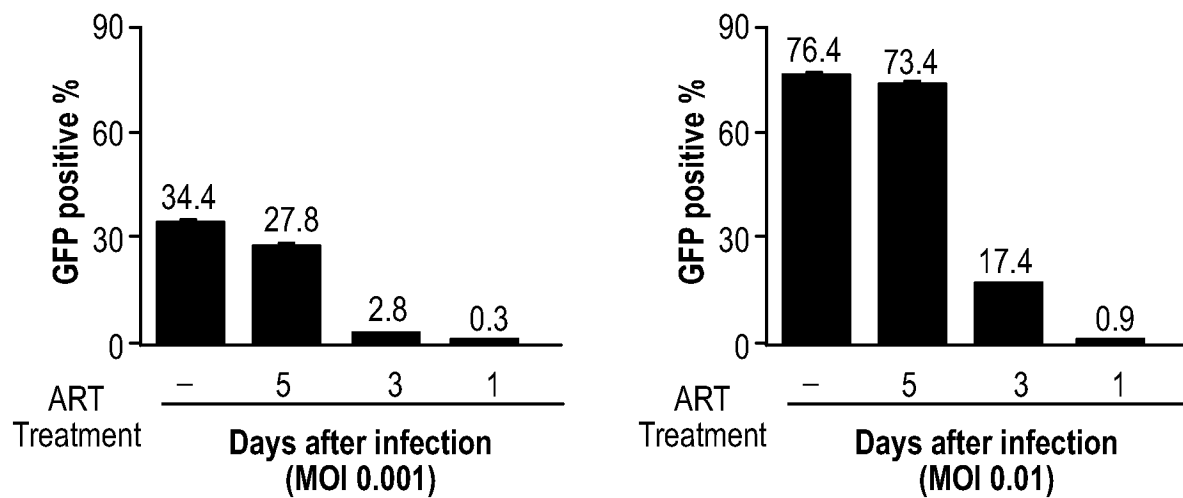


FIG. 6C

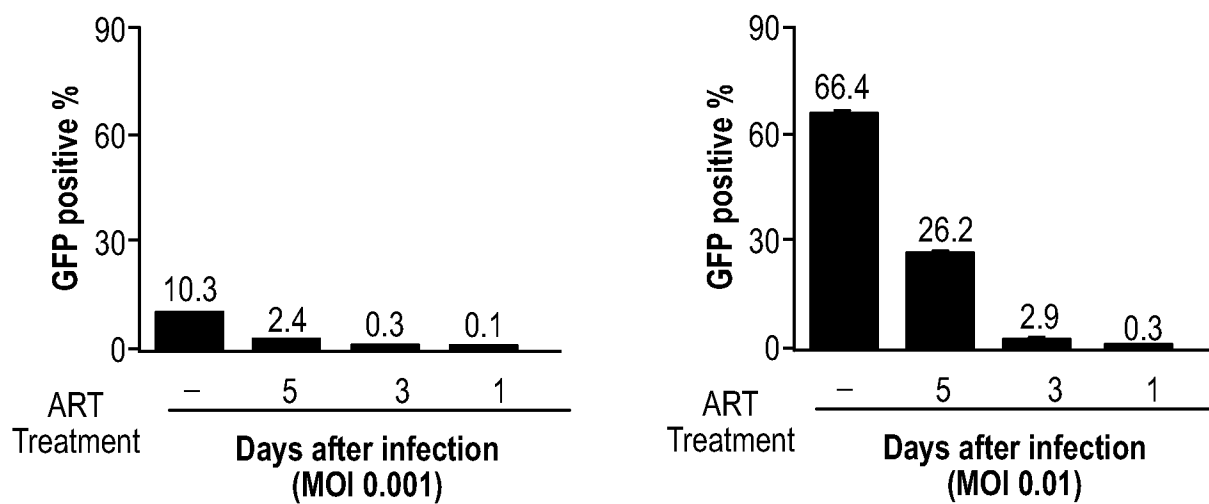


FIG. 7A

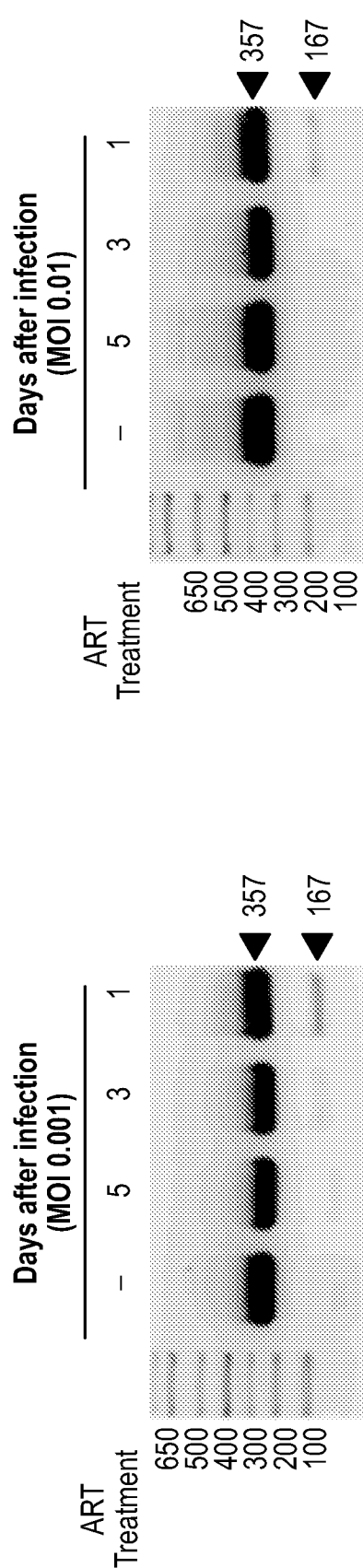


FIG. 7B

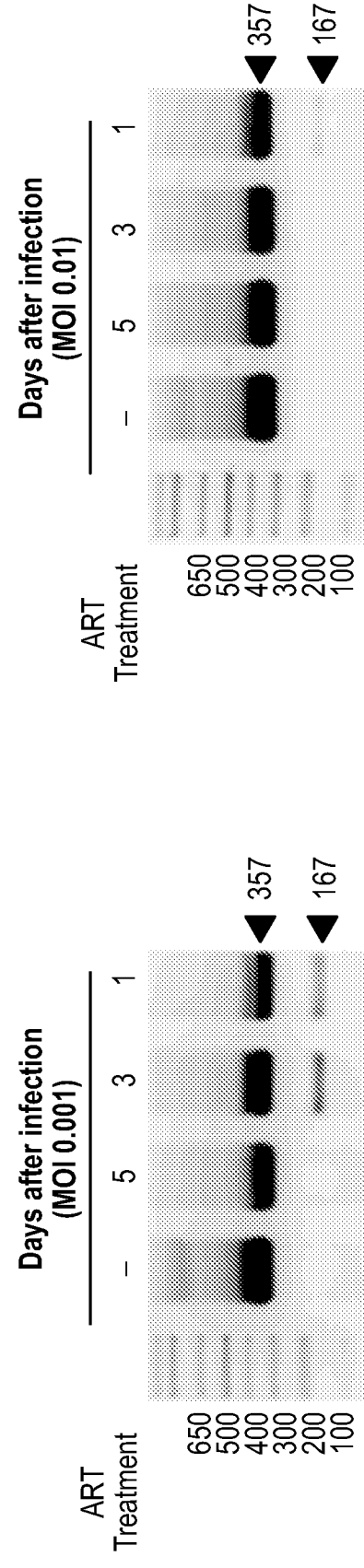


FIG. 7C

Primer F [-375/-354]Target LTR A [-347/-328] PAMPAM Target LTR B [-143/-124]Primer R [-41/-19]

NL4-3Jurc1

TTGGCAGAACTACACACAGGGCCAGGGATCAGATATCCACTGACCTTTGGATG//CATCCGGAGTACTACAAAGACTGCTGACA//GCCCTCAGATGCTACATATAAGCAGC
TTAGCAGAACTACACACAGGGCCAGGGGTCCAGGTCCAGTATCCACTGAC//-----TACTTCAAGAACTGCTGACA//GCCCTCAGATGCTACATATAAGCAGC

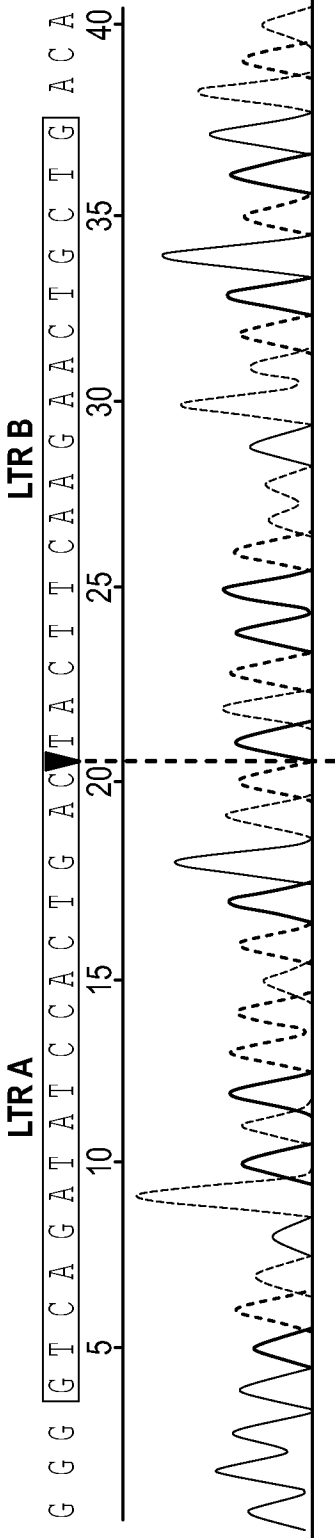


FIG. 7D

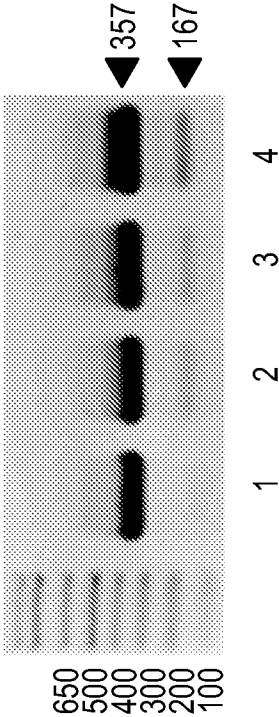


FIG 7E

LTR (HXB2)
TUR92C1
TUR92C2

Primer F [-375/-354]
ttagcagaactacacaccagggtccagggaatcagatatccactgaccttctggtg
TTAGCAGAACTACACACCAGGCTCAGGGATCAGATATCCACTGAC
TTAGCAGAACTACACACCAGGCTCAGGGATCAGATATCCCA

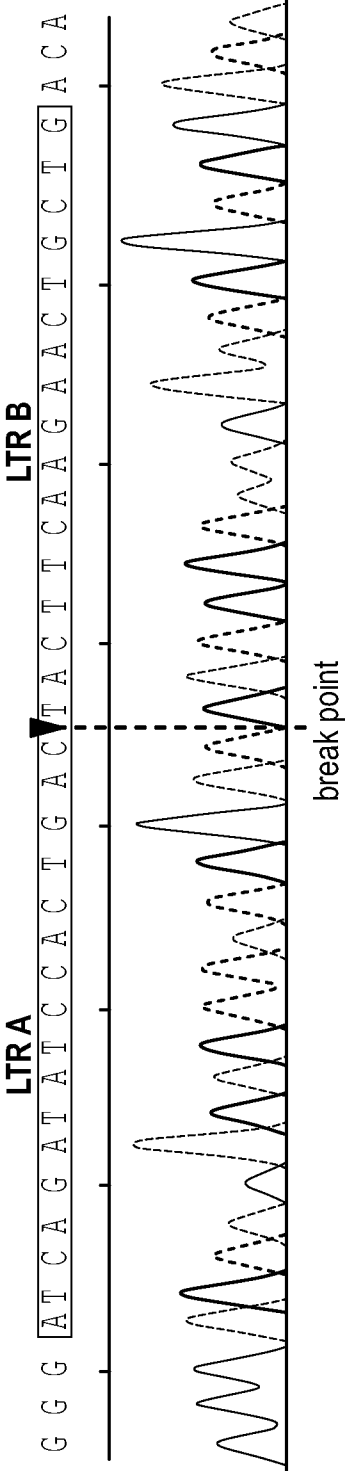
Target LTR A [-347/-328] PAM

PAM Target LTR B [-143/-124]
catgggpcgtacttcaagaactgctgaca//gccctcagatccctgcataataagcagc
TACTTCAAGAACTGCTGACA//GCCCTCAGATGCTACATATAAGCAGC
TACTTCAAGAACTGCTGACA//GCCCTCAGATGCTACATATAAGCAGC

Primer R [-41/-19]

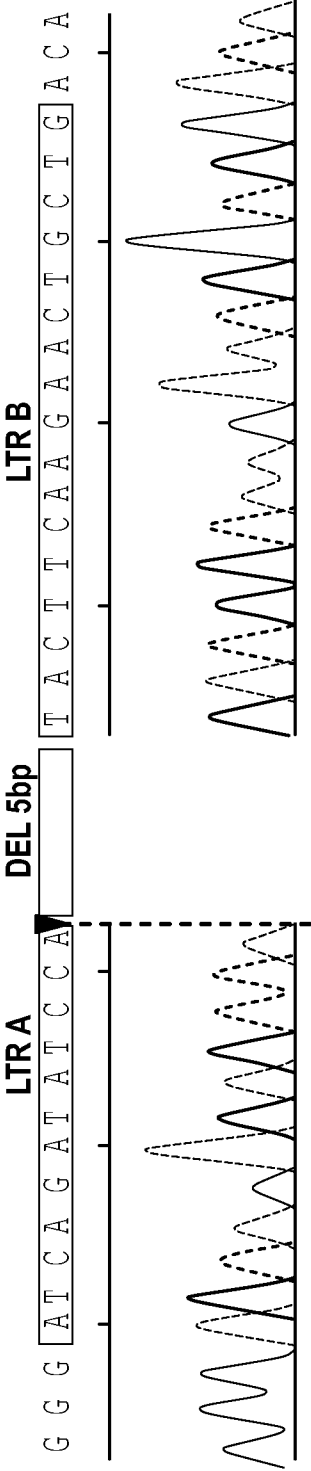
***** ** *****

TUR92c1



15/41

TUR92c2



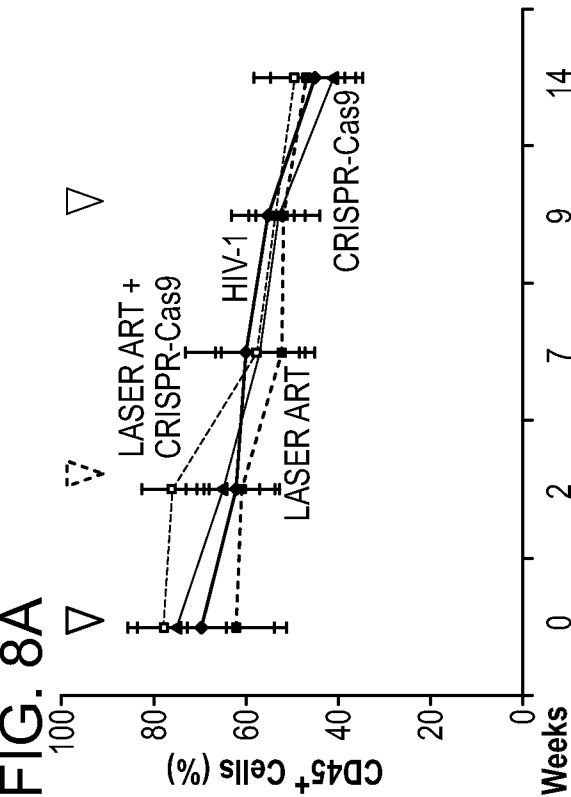
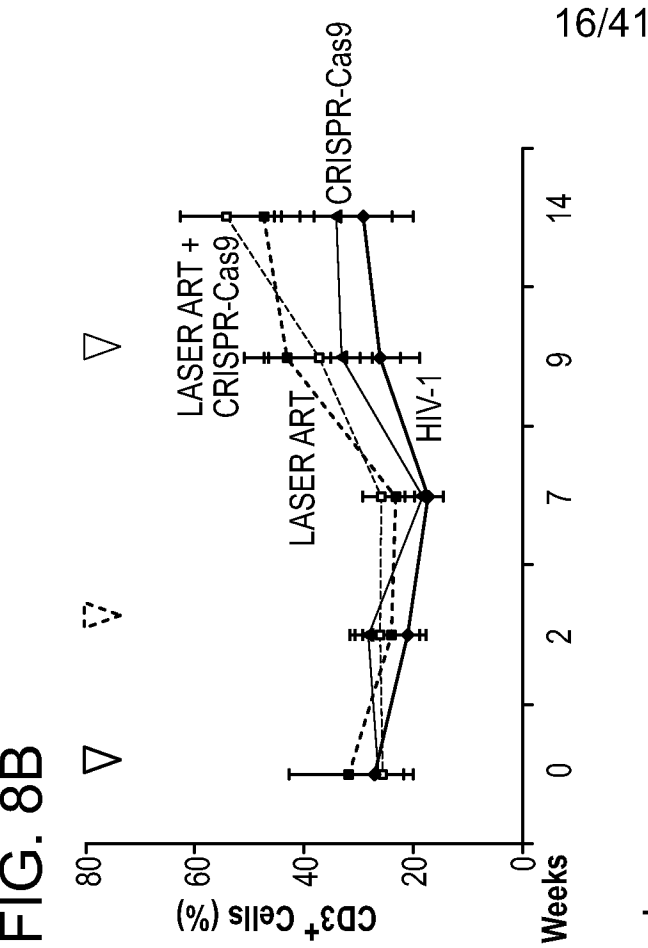


FIG. 9

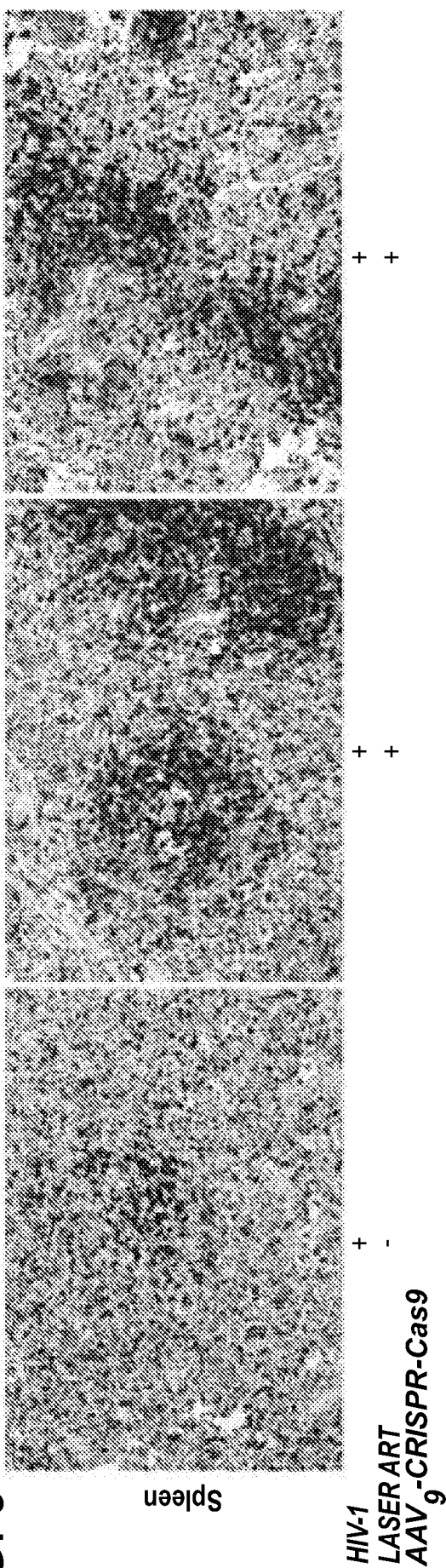


FIG. 10

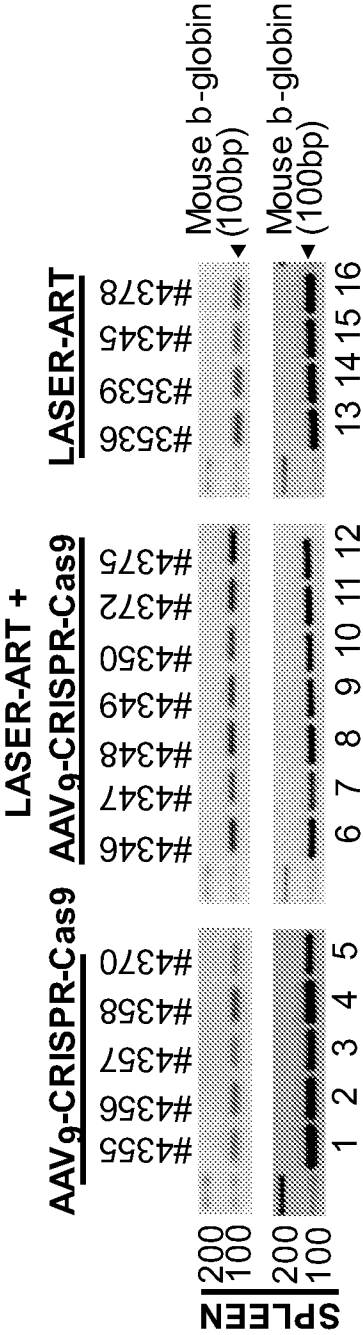


FIG. 11

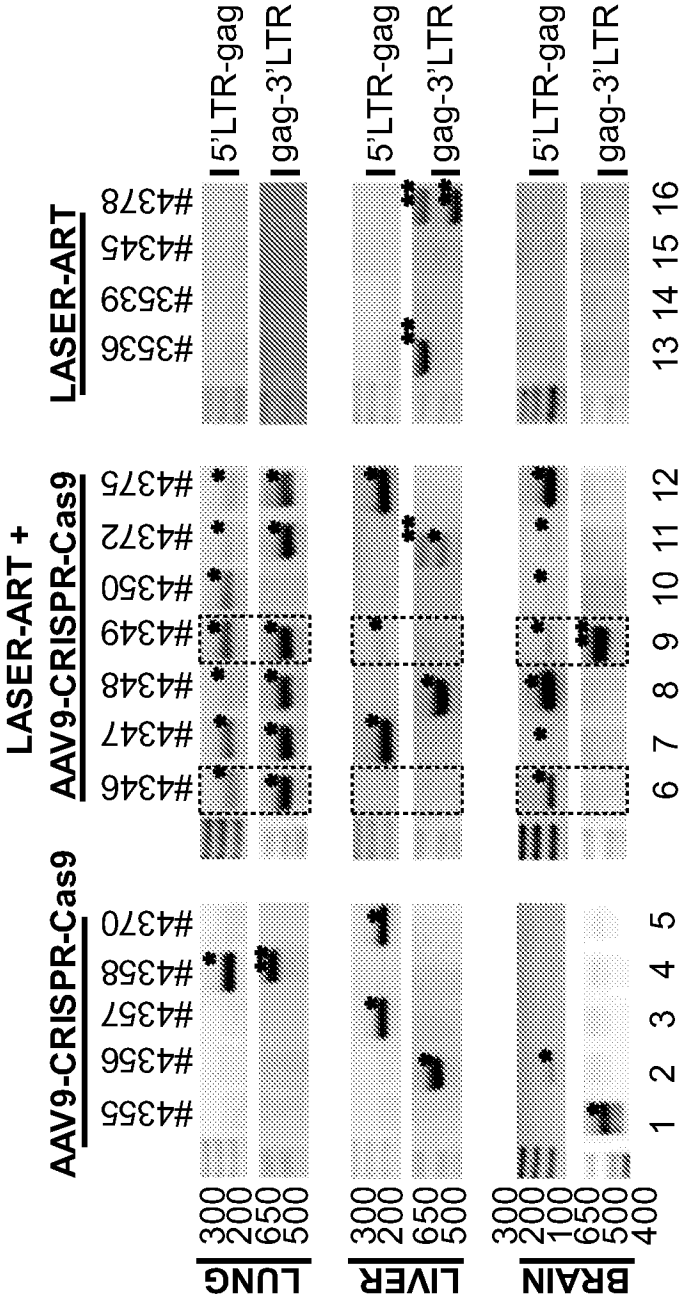


FIG. 12A SPLEEN 5'LTR-gag #4336 (lane 6) and gag-3'LTR #4357 (lane 9)

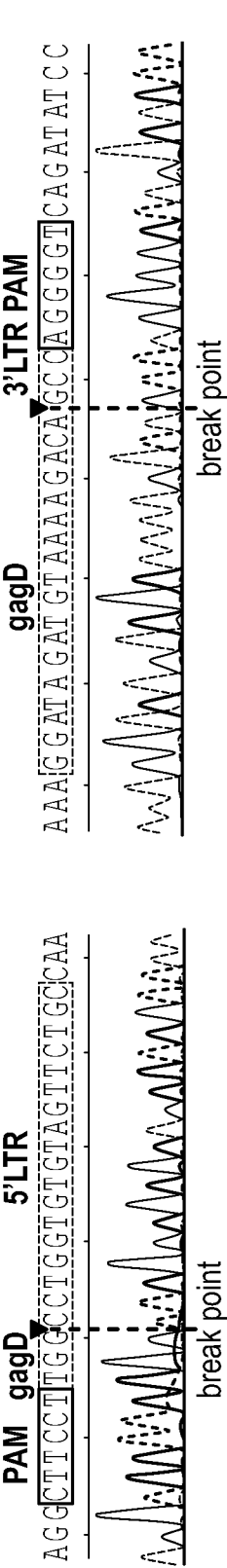


FIG. 12B GALT 5'LTR-gag #4349 (lane 9)

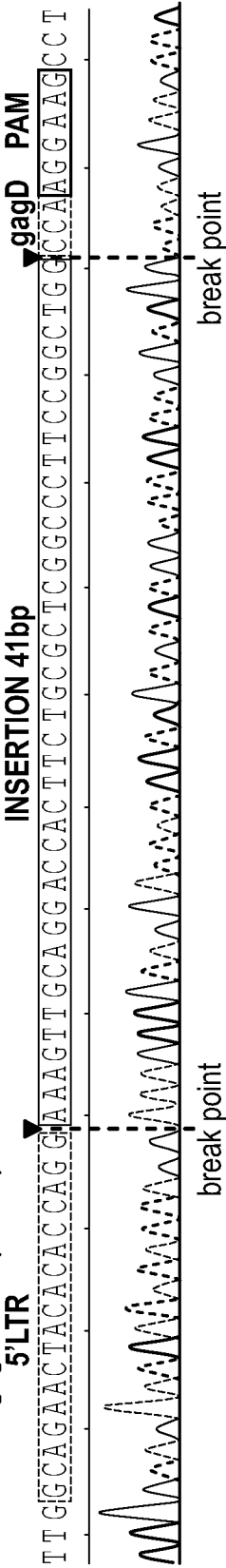
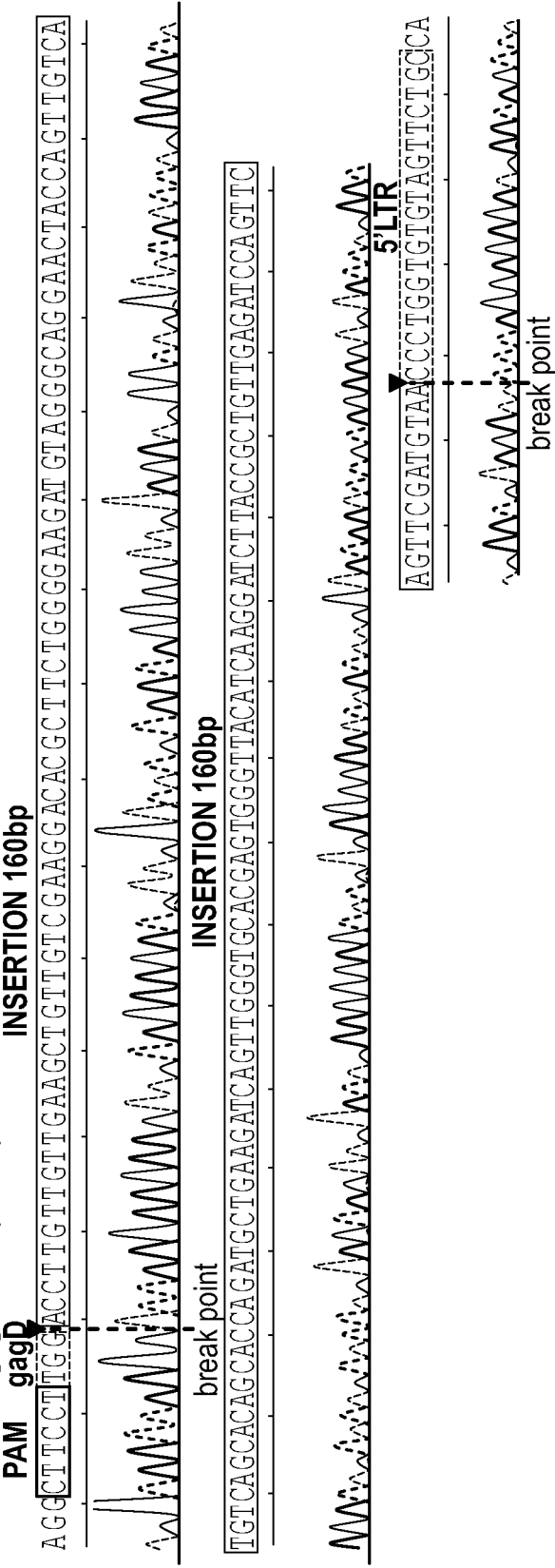
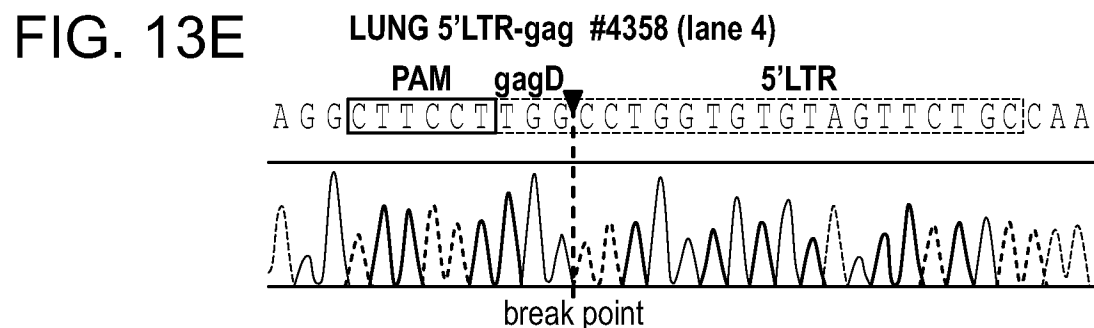
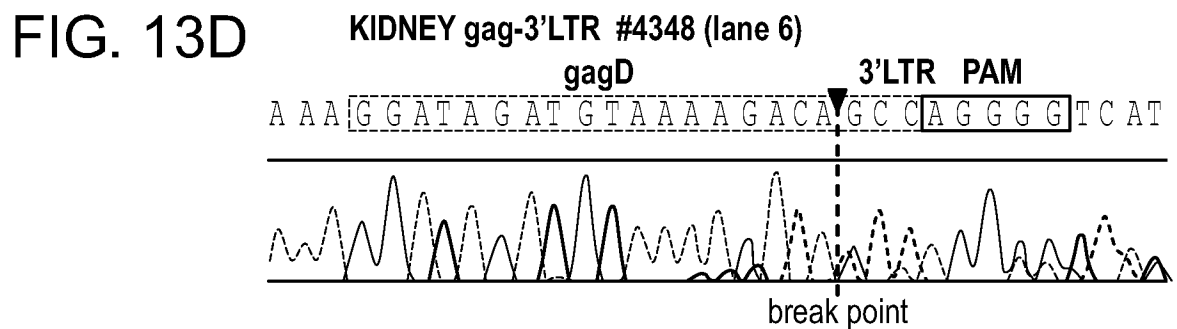
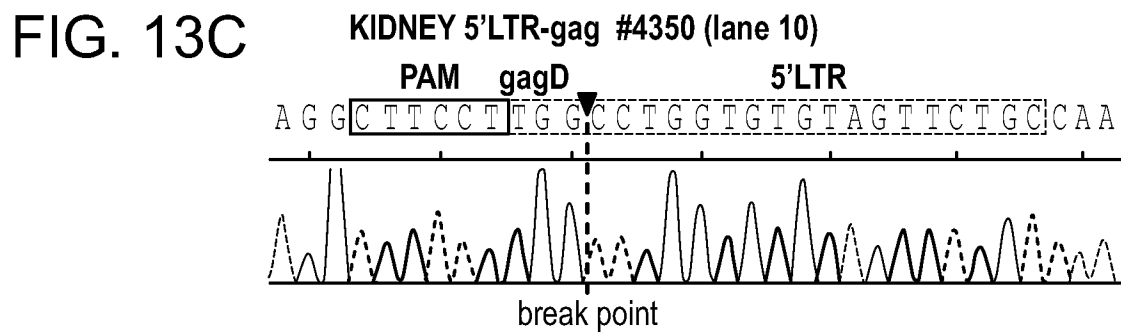
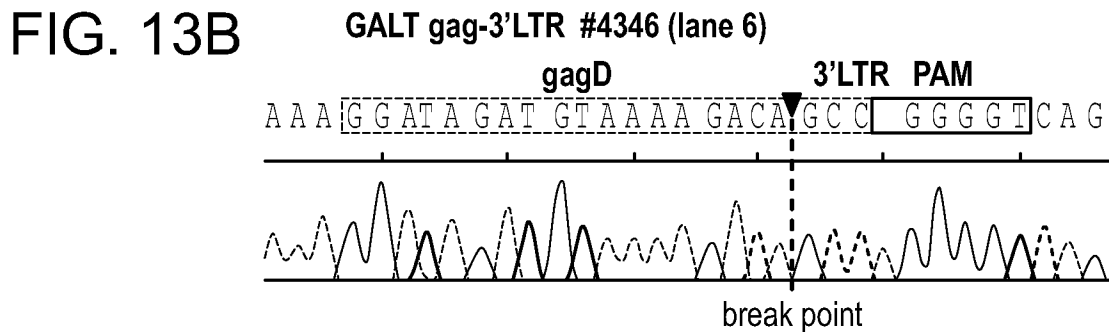
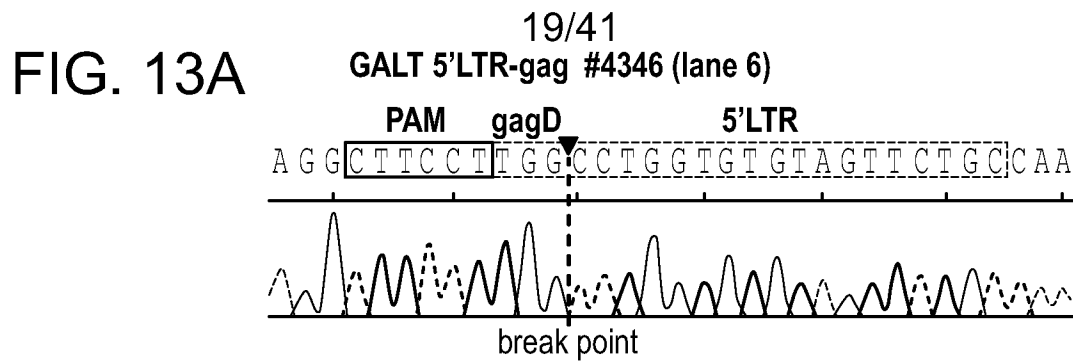


FIG. 12C KIDNEY 5'LTR-gag #4349 (lane 9)





20/41

FIG. 13F

LUNG gag-3'LTR #4346 (lane 6)

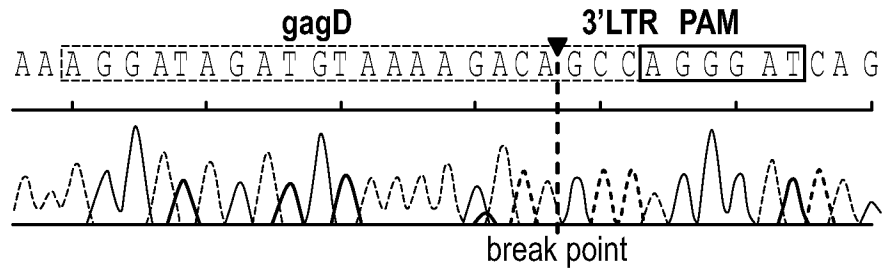


FIG. 13G

LIVER 5'LTR-gag #4357 (lane 3)

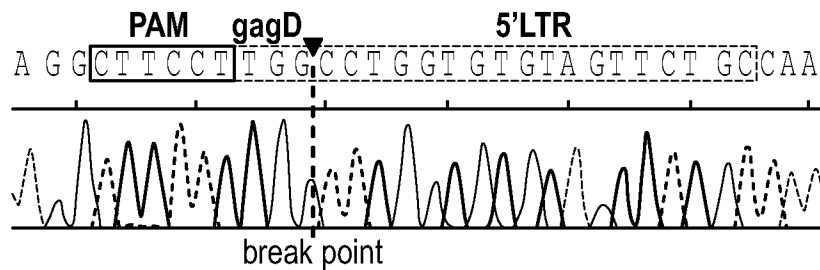


FIG. 13H

LIVER gag-3'LTR- #4348 (lane 8)

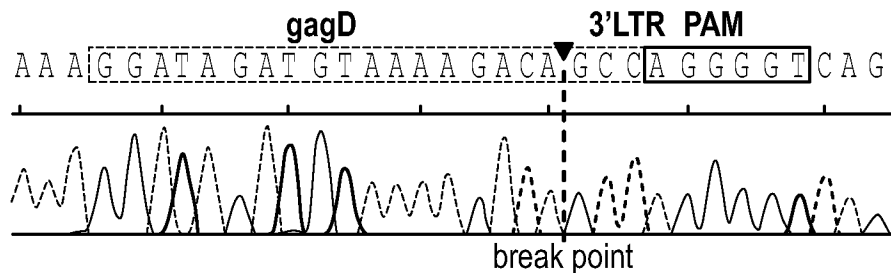


FIG. 13I

BRAIN 5'LTR-gag #4348 (lane 8)

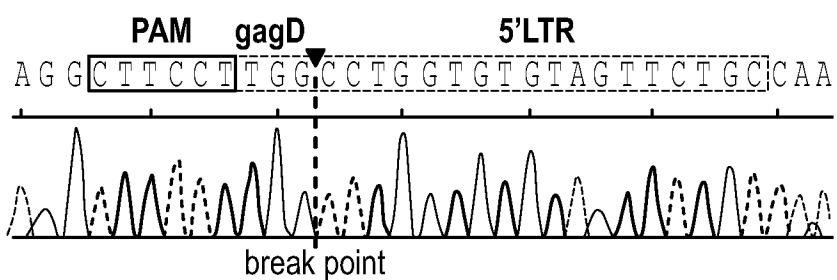
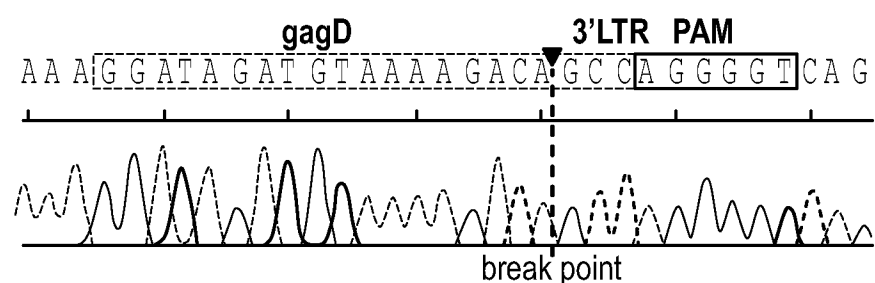


FIG. 13J

BRAIN gag-3'LTR- #4355 (lane 1)



21/41

FIG. 13K

SPLEEN gag-3'LTR #4356 (lane 2)

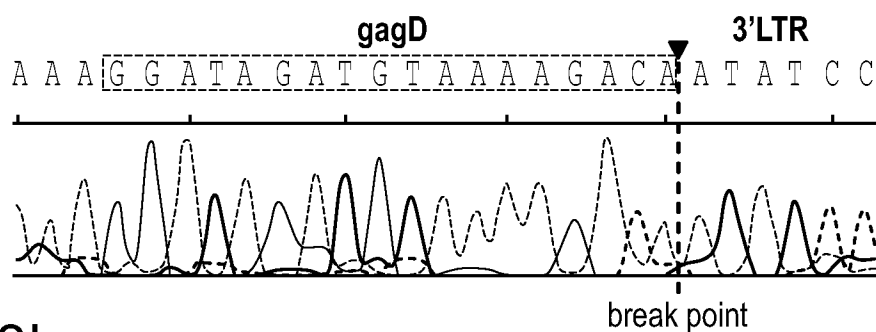


FIG. 13L

SPLEEN gag-3'LTR #4375 (lane 12)

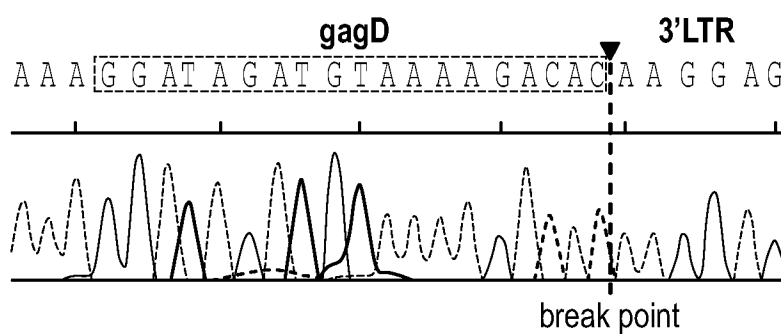


FIG. 13M

KIDNEY 5'LTR-gag #4347 (lane 7)

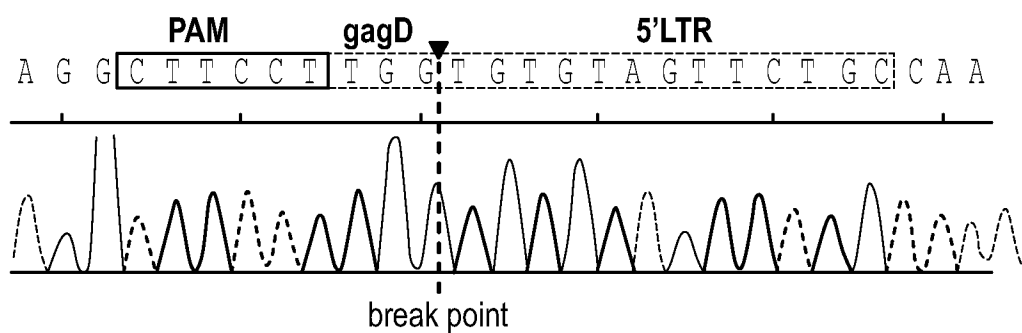


FIG. 14A

SPLEEN 5'LTR-gag #4357 (lane 3) and # 4347 (lane 7)

```

primer F LTR-408/-387] Target LTR 1 [-372/-352] PAM
NL4-3 TGGATCTACACACAAAGGCTACTTCCCTGATTGGCAGAACTACACACAGGGCOAGGGATCAGATATCCACTGACCTTTGGATGGTCTTCAAGTTAGTACCAGTTGAACCCAGA//
3 -----CCACACAAAGGCTACTTCCCTGATTGGCAGAACTACACACAGGGCOAGGGGT CAGATATCCACTGACCTTTGGATGGTCTTCAAGCTAGTACCAGTTGAGCC---//
7 ---AGGNCCACACAAAGGCTACTTCCCTGATTGGCAGAACTACACACAGGGCOAGGGGT CAGATATCCACTGACCTTTGGATGGTCTTCAAGCTAGTACCAGTTGAGCC---//
*****
Target Gag D [+607/+627] PAM
NL4-3 //AAAGGATAGATGTAAAAGACACCAAGGAAGCCTTAGATAAGATAGAGAAAGACAAACAAAAGTAAGAAAAGGCACAGCAAGCAGCAGCTGACACAGGAAACAAACAGCCAGG
3 //-----AGATAAGGTAGAGGAAGAGCAAAACAAAAGTAAGAAAAGGCACAGCAAGCAGCAGCTGACACAGGAAACAAACAGCCAGG
7 //-----AGATAAGGTAGAGGAAGAGCAAAACAAAAGTAAGAAAAGGCACAGCAAGCAGCAGCTGACACAGGAAACAAACAGCCAGN
*****
primer R [gag +744/+763]
NL4-3 TCAGCCAAAATTACCCCTATAGTGCAGAACCTCCAGGGGCAAAATGGT:
3 TCAGCCAAAATTACCCCTATAGTGCAGAACCTCCAGGGGCAAAATGGT:
7 TCAGNCAA-----TATAGTGCAGAACATCCAGGGGCAAAATGGT:
*****

```

22/41

FIG. 14B

SPLEEN gag-3'LTR #4378 (lane 16)

```

Target 3'LTR 1 [+8704/+8724] PAM
NL4-3 //TTGGCAGAACTACACACAGGGCOAGGGGTGAG//GATGACCCCTGAGACAGAGTGTAGAGTGGAGTTTGACAGCCGCTAGCAATTCATCAGTGGCCCCGAGAGCTGCATCCGGA
16 //-----NNNNNNCANANACTTAGATCATTTATATAATACAAATAGCAGTCCCTCTATTTGTGTGACCTGGNCCNNNTGNTGNTCCGGA
*****
NL4-3 GTACTTCAAGAACTGCTGACATCGAGCTTGCTACAAGGACCTTCCAGGGAGGGCTGGCCCTGGGCGGACTGGGAGTGGCGAGCCCTCAGATGCTGCATATAAGC
16 TTACTTCAAGAACTGCTGACATCTAGCTTGCTACAAGGACCTTCCAGGGAGGGCTGGCCCTGGGCGGTACTGGCGGTAGTGGCGATCCCTCCAAAGCTGGAAAAAAC
*****
primer R [LTR-9116/-9140]
NL4-3 AGCTGCTTTTGGCTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAAGTGGGAAACCCACTGCTTAAGCCTCAATAAAGCTTGCCT-TGAGTGTCTCAA
16 GCCTGCTTTTACATGACCTGGGNCCTTTACACCCCTGCCAGATTGANGAANGNCTCTNNGGCTACTAAGNAACCCACTGCTTNAATTTCTTNNNNCTACNTTTTCATCCGTGGCCCCA
*****
NL4-3 GTAGTGTGTGCCCGTCTGTGTGTGACTCTGTGTAAGTCCCTCAGACCCCTTTTAGTCAGTGTGGAAAAATCTCTAGCA-----
16 NCGCTGCATCCCTGAGTNGGTGAAAAAACTGCTGACATCNANCTTGNTACAGGATTTTCCACNGNGGGGTTCCTCCNGGGGNGGTGTGCGCGGGGGCGGANAGNGGGGAGTGGGGACCCCTCA
*****
NL4-3 -----
16 TATGCNACAANAANAGCCGCTTTTNGCGNGVAAGGCTCTCTCTTTTANACCANA TCTGNTNNNGNCTCTCTGCTGTTGGGNGCCCTGCTNATTTTCNNNNNNN
-----

```


FIG. 14E

LIVER gag-3'LTR #3536 (lane 13) and # 4378 (lane 16)

```

NL4-3      primer F [gag +520/+546]      Target Gag D [+607/+627]  PAM
2      :GACAGCTACACCAATCCCTTCAGACAGGAT // AAAGGATAGATGTAAAGACACCAAGGAAGCCCT // TGCAAAATGTTAAA // GAAATCTATAAAA //
      :-----NNNNN // AAAGGATANNNTTAAANACACCAAGGAGTCCT // -----//
16top    :-----NNNNN // AAAGGATAATAGTAAAGACACCAAGGAAGCCCT // TGCAAAATGTTAAA // GAAATCTATA
16bottom :-----NNNNN // AAAGGATAAATGTAAAGACACCAAGGAAGCCCT // TGCAAAATGTT ---//

      ***** * ***** * * *
      Target 3'LTR 1 [+8704/+8724]  PAM      primer R [LTR - 9116/-9140]
NL4-3      // TTGGCAGAACTACACACCAAGGGGTCAG // GGAGCTCTCTGGCTAACTAGGGAACCCACTGCT:
2      // -----// GNACCTCTCTGGCTAANNANGNAACCCACTGTG:
16top     // -----// -----GCTCTCTGGCTAACTAGGGAACCCACTNG:
16bottom  // -----GTGGCGGAGCC // GGAGCTCTCTGGCTAACTAGGGAACCCACTGCT:
      ***** * * ***** * *

```

FIG. 14F

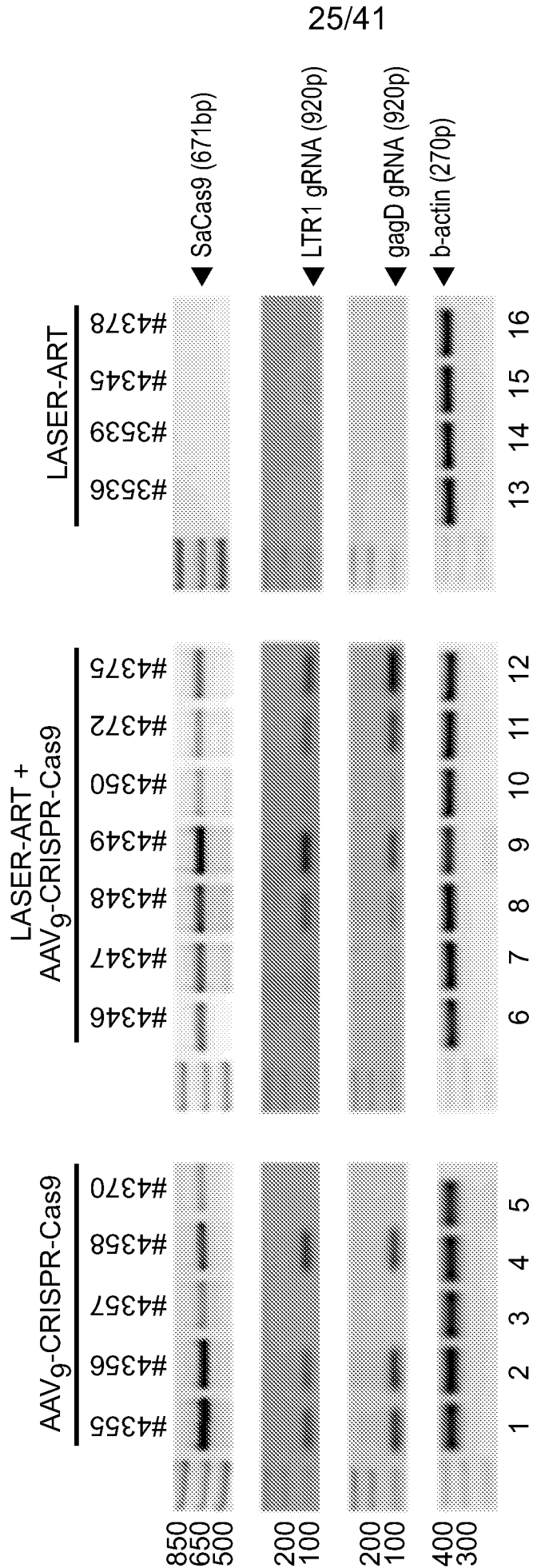
BRAIN gag-3'LTR #4349 (lane 9)

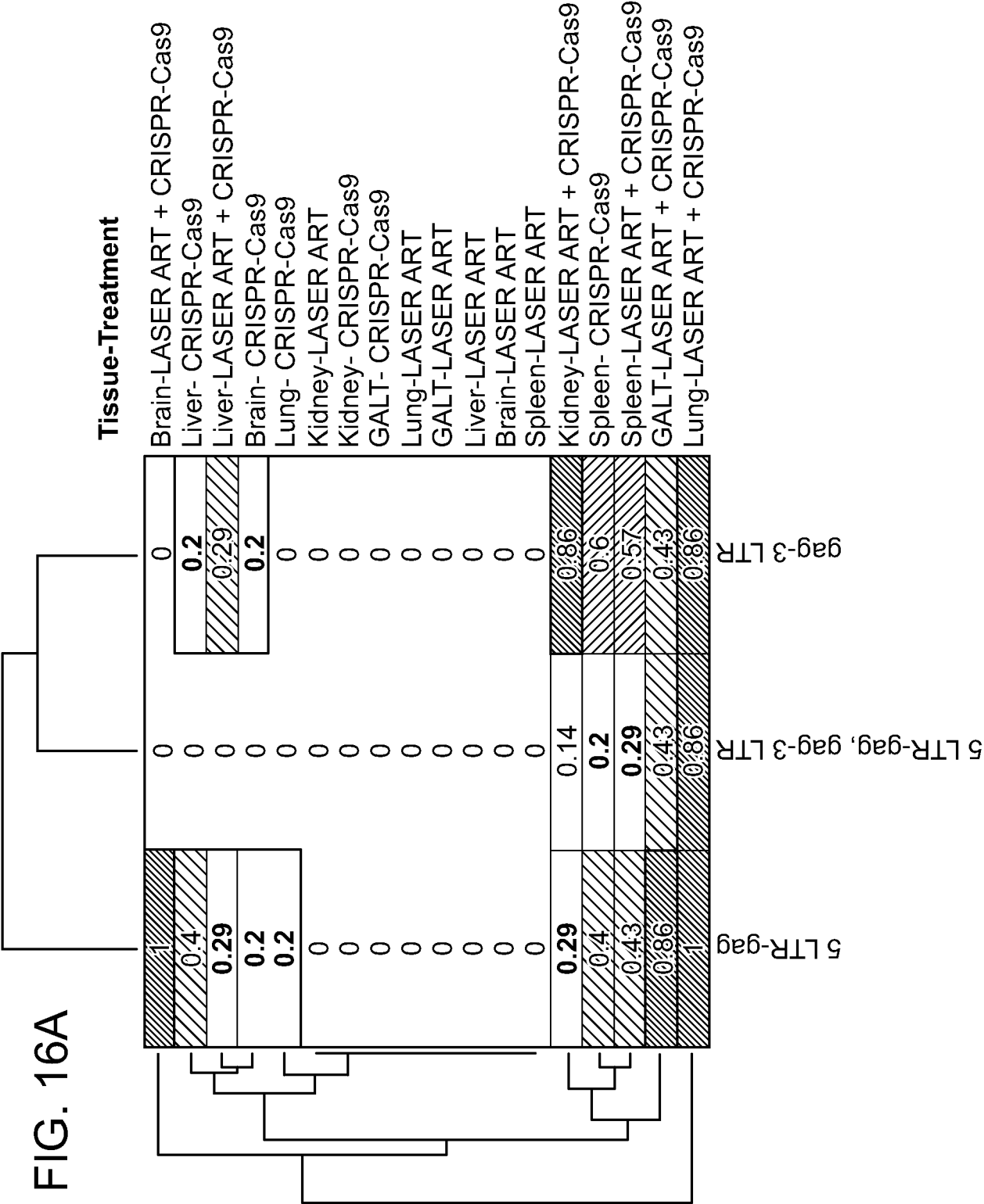
```

NL4-3      primer F [gag +520/+548]      Target Gag D [+607/+627]  PAM
9      :ACAGCTACACCAATCCCTTCAGACAGGAT // AAAGGATAGATGTAAAGACACCAAGGAAGCCCT // AGACCATCAA // TTGGCAGAACTACACACCAAGGGGTCAG //
      :-----NNNNNNN // AAAGGATANNNTTAAAGACACCAAGGAAGCCCT // AGACCAT ---//
      ***** * ***** * *
      Target 3'LTR 1 [+8704/+8724]  PAM      primer R [LTR-9116/-9140]
NL4-3      // GGTCCTCTGTTAGACACAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCT:
9      // ----CTCTCTGTTAGACACAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCT:
      ***** * * ***** * *

```

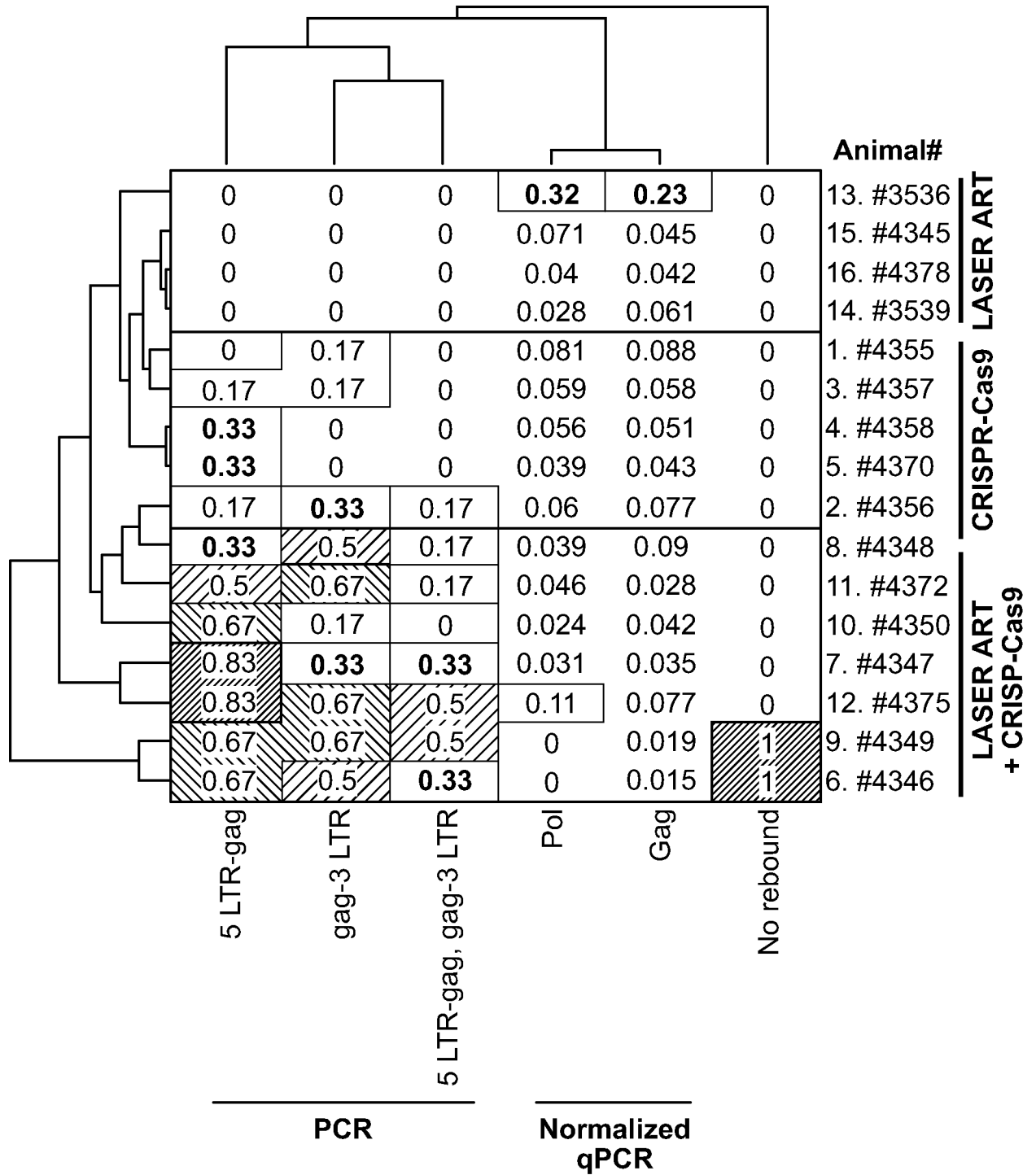
FIG. 15





27/41

FIG. 16B



28/41

FIG. 16C

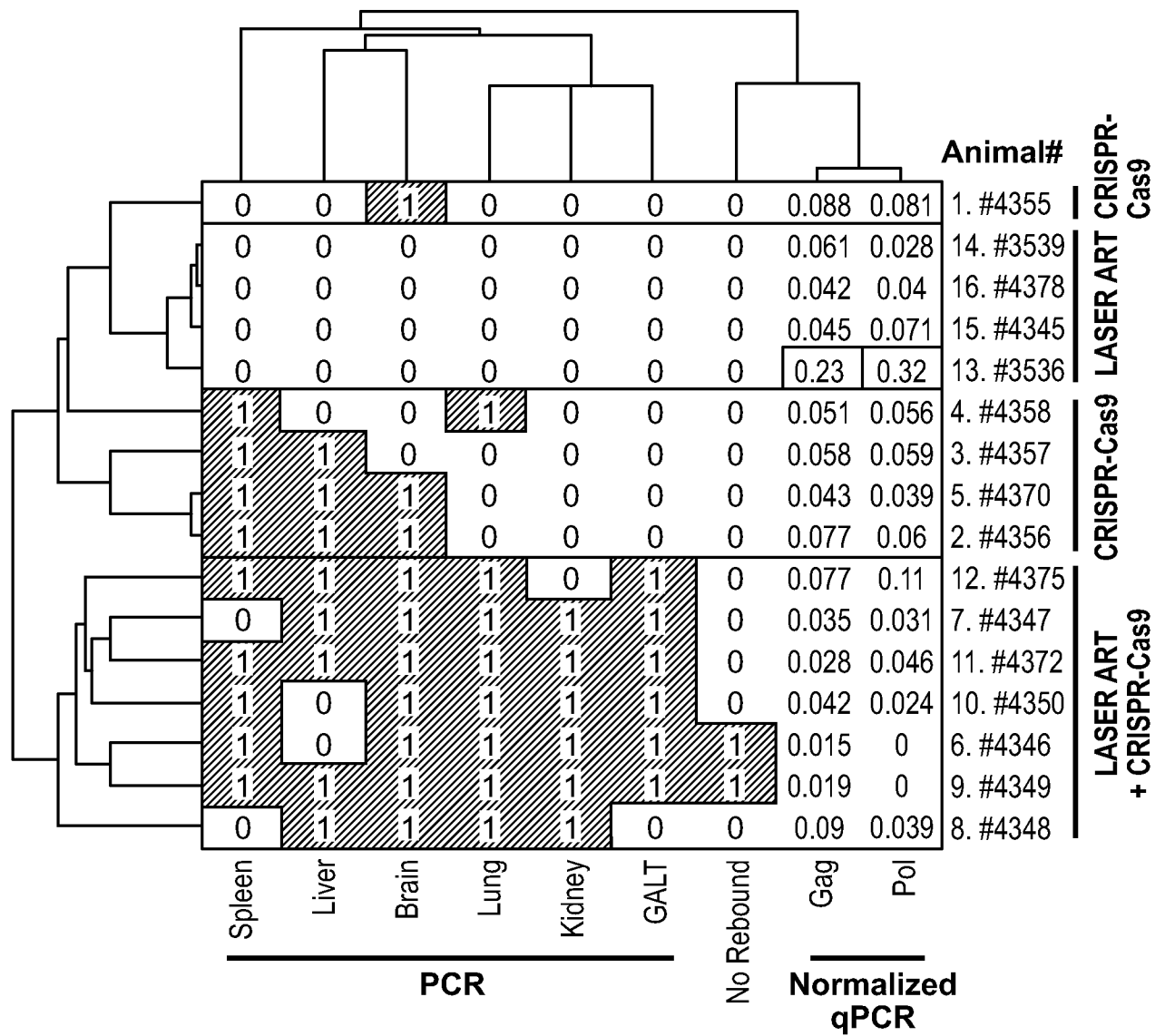


FIG. 17A

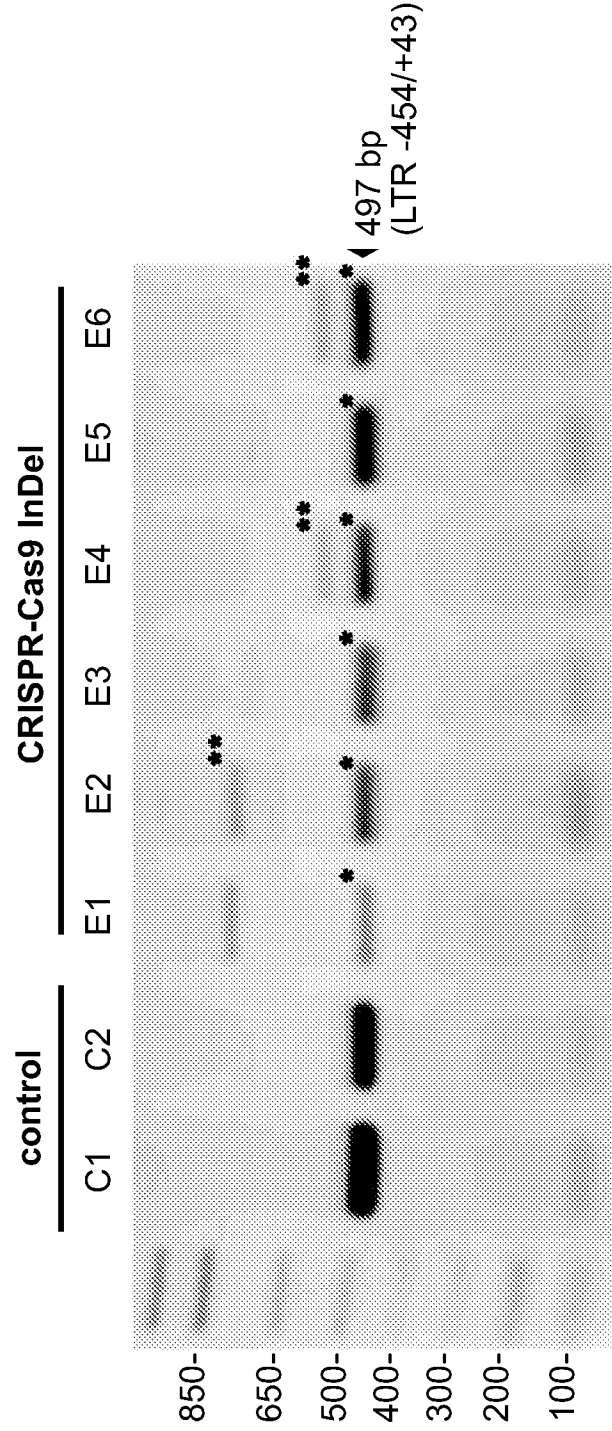


FIG. 17B

	Primer F [-454/-431]	Target LTR 1 [-372/-352] PAM	Primer R [+19/+43]	
NL4-3	TGGAAGGGCTAATTGGTCCCAAAAGAA//TTGGCAGAACTACACACCAGGCGCAGGGATCAGATAT//ACCGATCTGAGCCTGGGAGCTCTCTG			
C1c1	TGGAAGGGCTAATTCACTCCCAACGAA//TTGGCAGAACTACACACCAGGCGCAGGGATCAGATAT//ACCGATCTGAGCCTGGGAGCTCTCGG			
C1c2	TGGAAGGGCTAATTCACTCCCAACGAA//TTGGCAGAACTACACACCAGGCGCAGGGATCAGATAT//ACCGATCTGAGCCTGGGAGCTCTCGG			
C2c1	TGGAAGGGCTAATTCACTCCCAACGAA//TTGGCAGAACTACACACCAGGCGCAGGGATCAGATAT//ACCGATCTGAGCCTGGGAGCTCTCGG			
C2c2	TGGAAGGGCTAATTCACTCCCAACGAA//TTGGCAGAACTACACACCAGGCGCAGGGATCAGATAT//ACCGATCTGAGCCTGGGAGCTCTCGG			
E1*c1	TGGAAGGGCTAATTCACTCCCAACGAA//TTGGCAGAACTACACACCAGGCGCAGGGATCAGATAT//ACCGATCTGAGCCTGGGAGCTCTCGG			
E1*c1	TGGAAGGGCTAATTCACTCCCAACGAA//TTGGCAGAACTACACACCAGGCGCAGGGATCAGATAT//ACCGATCTGAGCCTGGGAGCTCTCGG			
E2*c1	TGGAAGGGCTAATTCACTCCCAACGAA//TTGGCAGAACTACACACCAGGCGCAGGGATCAGATAT//ACCGATCTGAGCCTGGGAGCTCTCGG			
E2*c2	TGGAAGGGCTAATTCACTCCCAACGAA//TTGGCAGAACTACACACCAGGCGCAGGGATCAGATAT//ACCGATCTGAGCCTGGGAGCTCTCGG			
E3*c1	TGGAAGGGCTAATTCACTCCCAACGAA//TTGGCAGAACTACACACCAGGCGCAGGGATCAGATAT//ACCGATCTGAGCCTGGGAGCTCTCGG			
E3*c2	TGGAAGGGCTAATTCACTCCCAACGAA//TTGGCAGAACTACACACCAGGCGCAGGGATCAGATAT//ACCGATCTGAGCCTGGGAGCTCTCGG			
E4*c1	TGGAAGGGCTAATTCACTCCCAACGAA//TTGGCAGAACTACACACCAGGCGCAGGGATCAGATAT//ACCGATCTGAGCCTGGGAGCTCTCGG			
E4*c2	TGGAAGGGCTAATTCACTCCCAACGAA//TTGGCAGAACTACACACCAGGCGCAGGGATCAGATAT//ACCGATCTGAGCCTGGGAGCTCTCGG			
E5*c1	TGGAAGGGCTAATTCACTCCCAACGAA//TTGGCAGAACTACACACCAGGCGCAGGGATCAGATAT//ACCGATCTGAGCCTGGGAGCTCTCGG			
E5*c2	TGGAAGGGCTAATTCACTCCCAACGAA//TTGGCAGAACTACACACCAGGCGCAGGGATCAGATAT//ACCGATCTGAGCCTGGGAGCTCTCGG			
E6*c1	TGGAAGGGCTAATTCACTCCCAACGAA//TTGGCAGAACTACACACCAGGCGCAGGGATCAGATAT//ACCGATCTGAGCCTGGGAGCTCTCGG			
E6*c2	TGGAAGGGCTAATTCACTCCCAACGAA//TTGGCAGAACTACACACCAGGCGCAGGGATCAGATAT//ACCGATCTGAGCCTGGGAGCTCTCGG			
E2**c1	TGGAAGGGCTAATTCACTCCCAACGAA//TTGGCAGAACTACACACCAGGCGCAGGGATCAGATAT//ACCGATCTGAGCCTGGGAGCTCTCGG			
		INSERTION 297bp		
	(CAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCAATTATGACGTCAATGGCGGGGGTGTGGGCGGTGAGCCAGCGGGCCATTACCGTAA			
	GTTATGTACGCGGAACCTCCATATATGGCTATGAACCTAATGACCCCGTAATTGATTACTATTAAATACTAGTCAATAATCAATGTCAACGCCTCGAGTCTAGA			
	GGCCGAGGAACCCCTAGTGTGAGTTGGCCACTCCCTCTCTGCGGCTCGCTCACTGAGCGGGCGACCAAGGTGCCCCGGG)			
E4**c1	TGGAAGGGCTAATTCACTCCCAACGAA//TTGGCAGAACTACACACCAGGCGCAGGGATCAGATAT//ACCGATCTGAGCCTGGGAGCTCTCGG			
E4**c2	TGGAAGGGCTAATTCACTCCCAACGAA//TTGGCAGAACTACACACCAGGCGCAGGGATCAGATAT//ACCGATCTGAGCCTGGGAGCTCTCGG			
		INSERTION 81bp		
	(AAAGTCCGCCGACGCCCGGGCGGCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGCAGGACATGTGAGCAAAAGGCCAGC)			
	(AAAGTCCGCCGACGCCCGGGCGGCTCAGTGAGCGGCGGCGCAGCTGCCTGCAGGACATGTGAGCAAAAGGCCAGC)			
E6**c1	TGGAAGGGCTAATTCACTCCCAACGAA//TTGGCAGAACTACACACCAGGCGCAGGGATCAGATAT//ACCGATCTGAGCCTGGGAGCTCTCGG			
		INSERTION 81bp		
	(AAAGTCCGCCGACGCCCGGGCGGCTCAGTGAGCGGCGGCGCAGCTGCCTGCAGGACATGTGAGCAAAAGGCCAGC)			

30/41

31/41

FIG. 17C

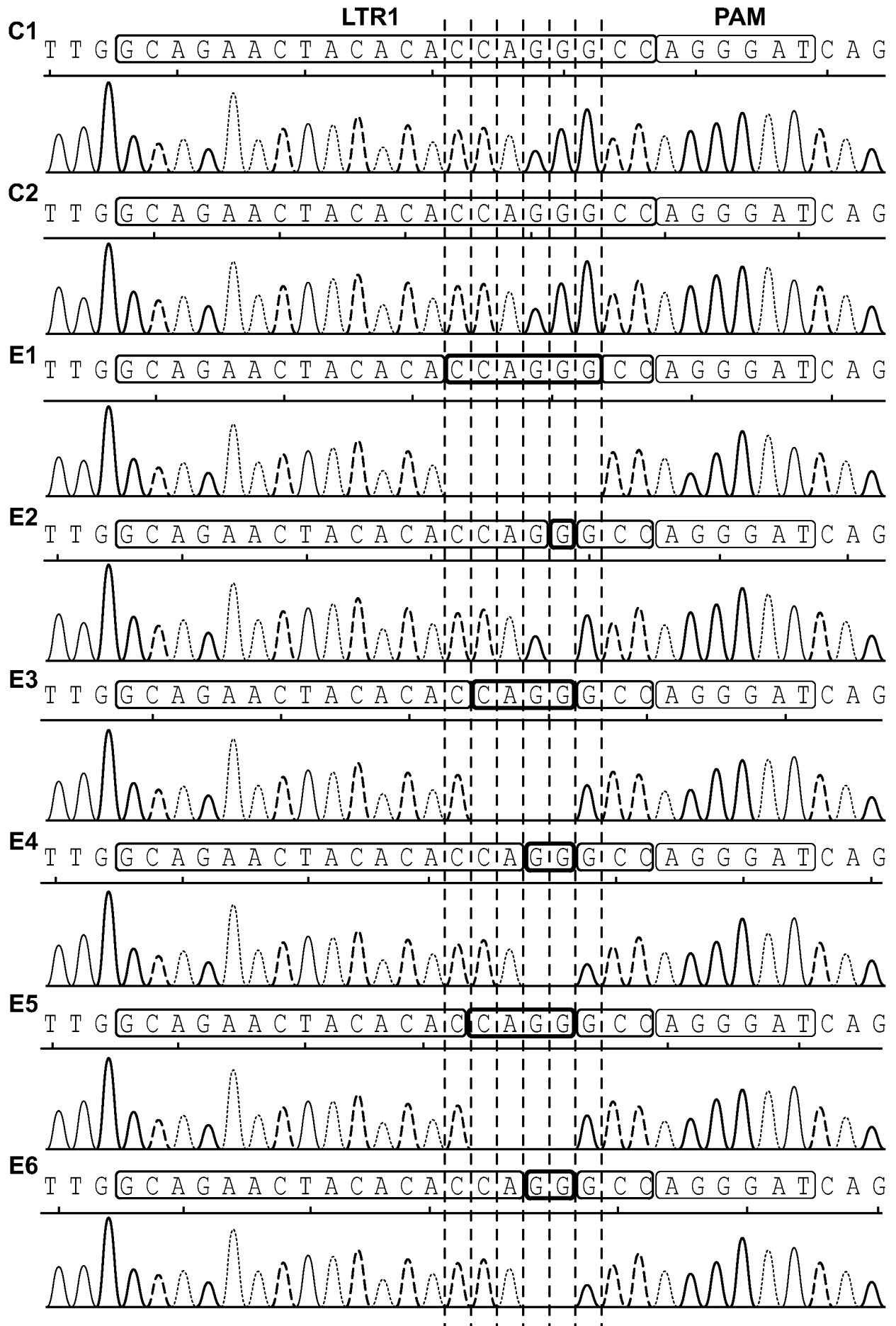


FIG. 18A

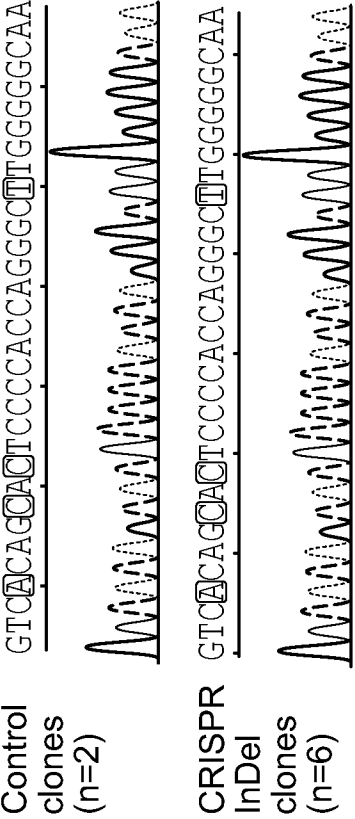


FIG. 18B

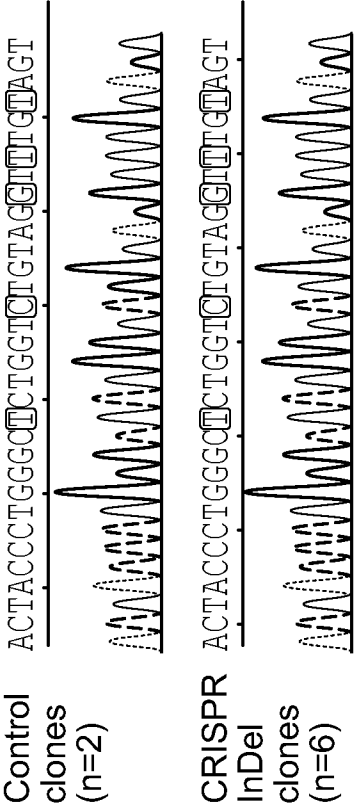


FIG. 18C

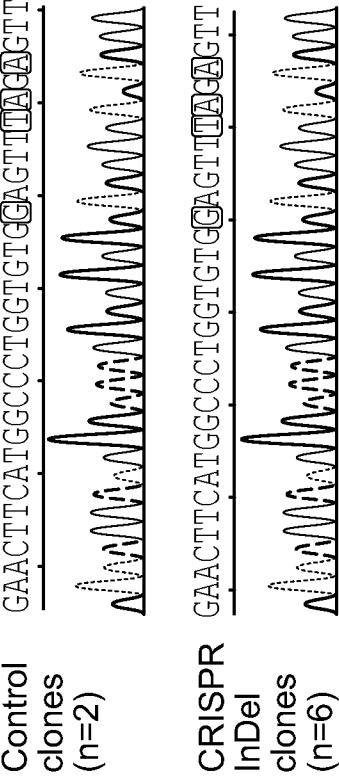


FIG. 18D

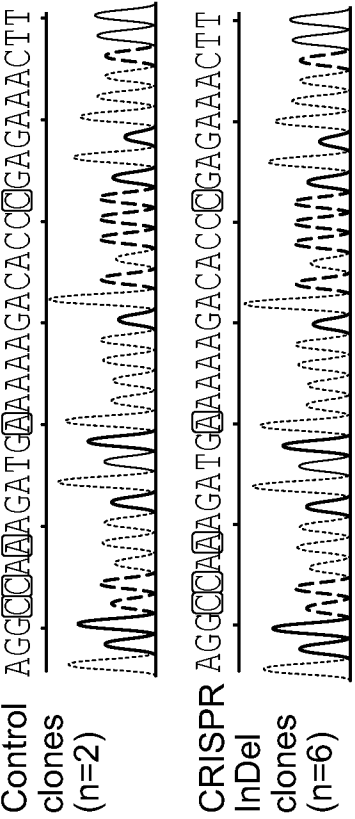


FIG. 18E

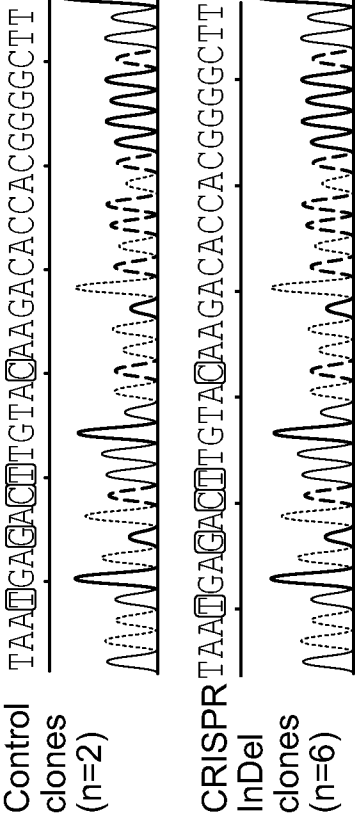
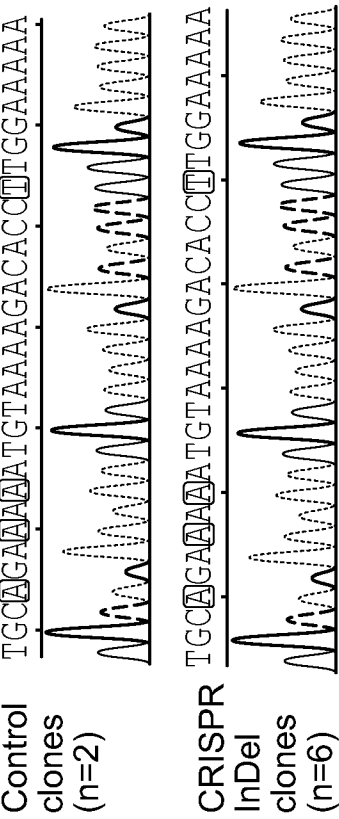
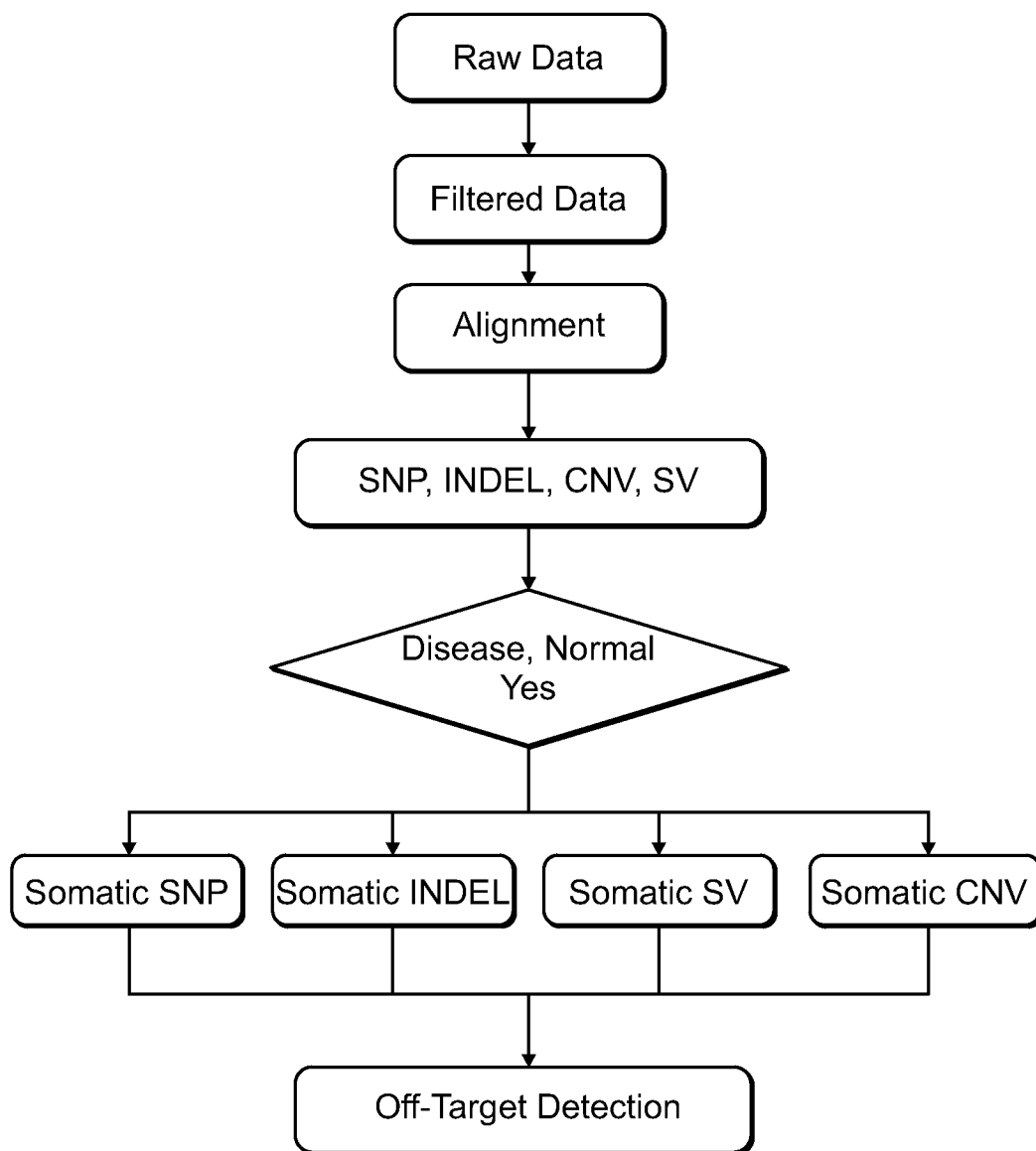


FIG. 18F



33/41

FIG. 19A



34/41

FIG. 19B

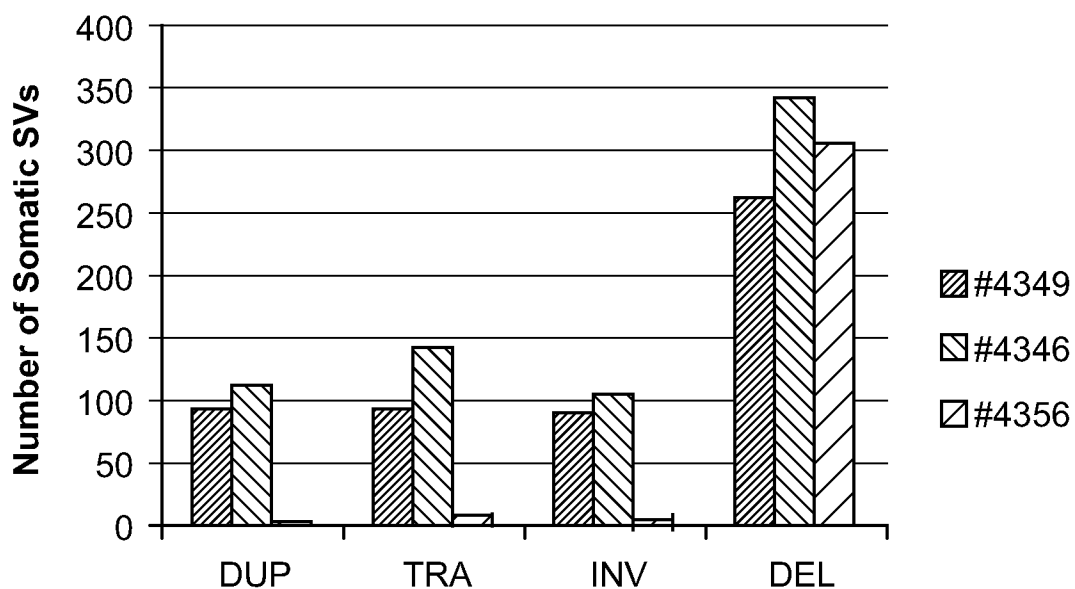
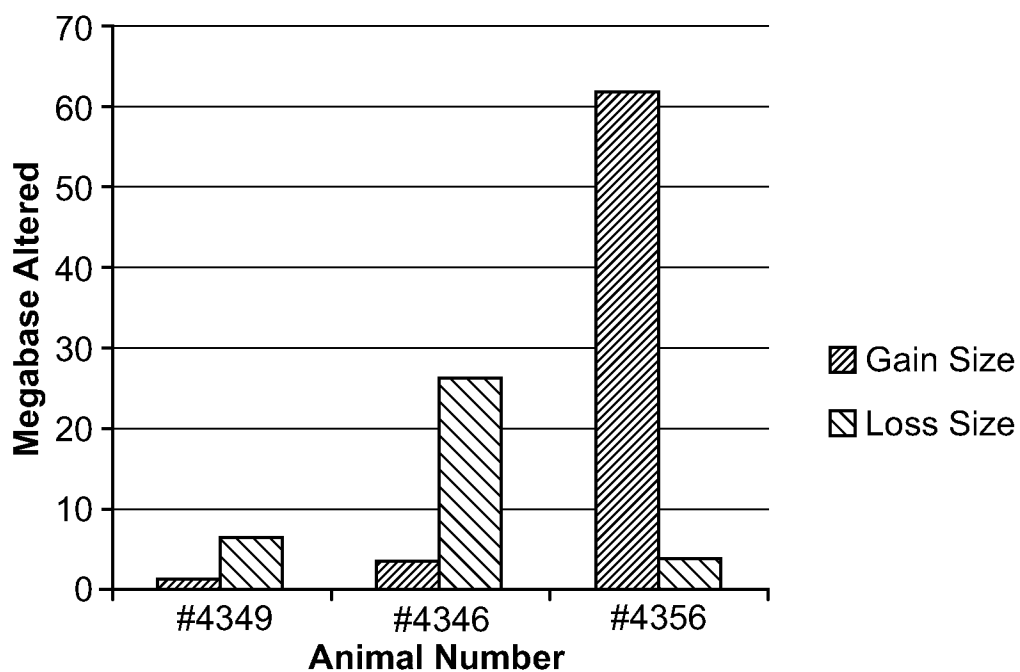


FIG. 19C



35/41

FIG. 20A

#3539

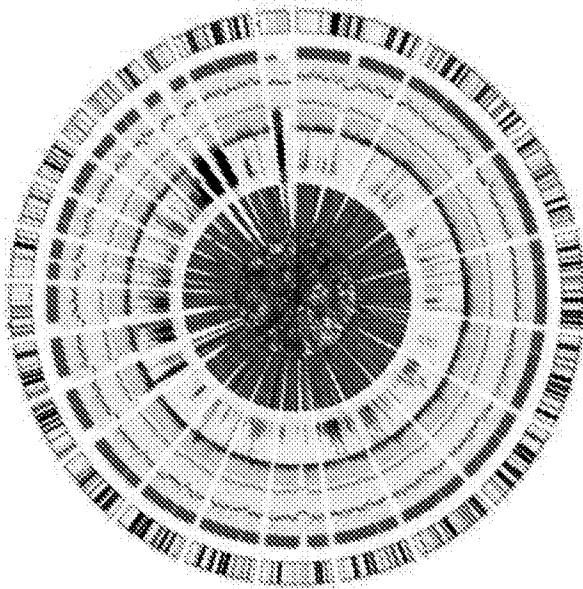


FIG. 20B

#4346

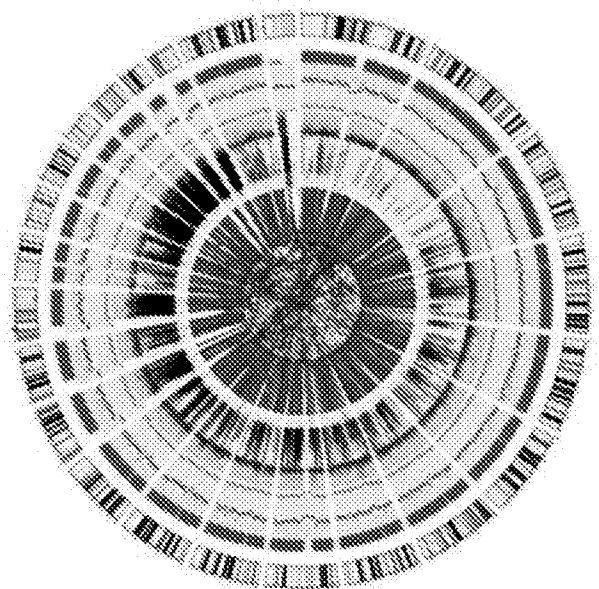


FIG. 20C

#4349

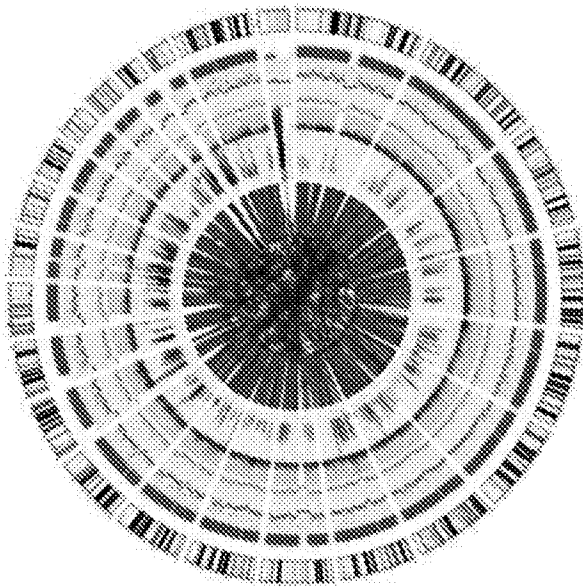
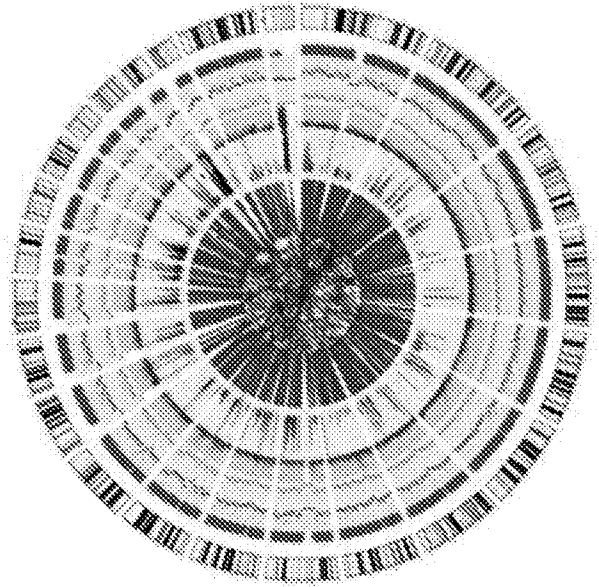


FIG. 20D

#4356



- Chromosome Information
- Read Coverage
- ~ InDel Density
- ~ SNP Density
- Proportion of Homozygous SNP
- CNV Inference
- SV Inference

FIG. 21A

FIG. 21B

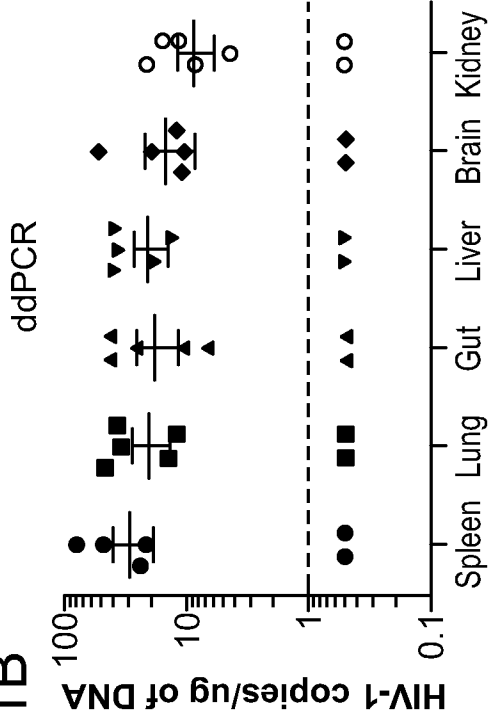
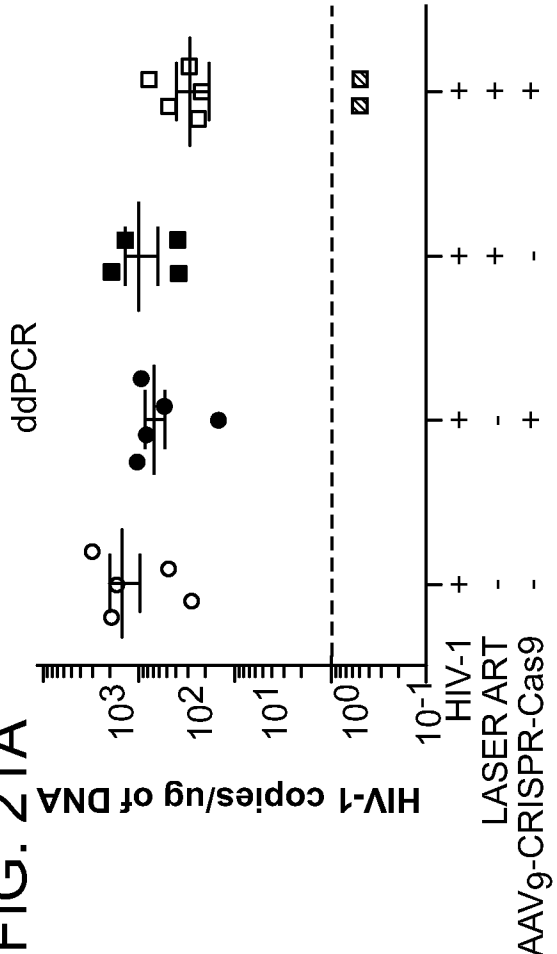


FIG. 22A

FIG. 22B

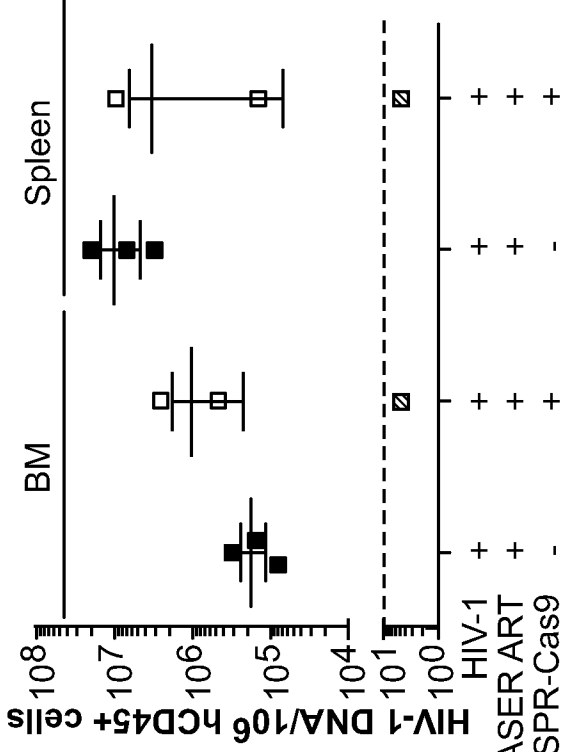
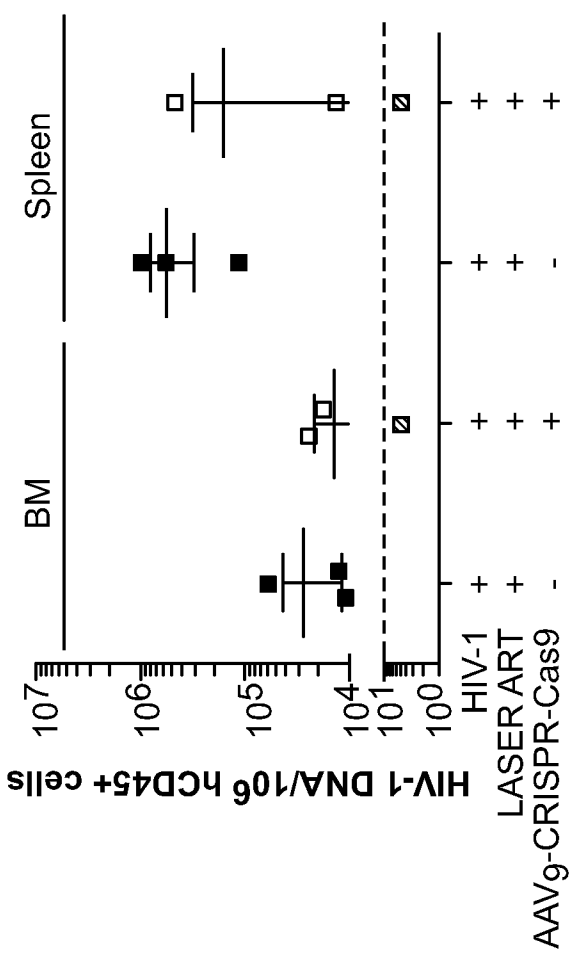
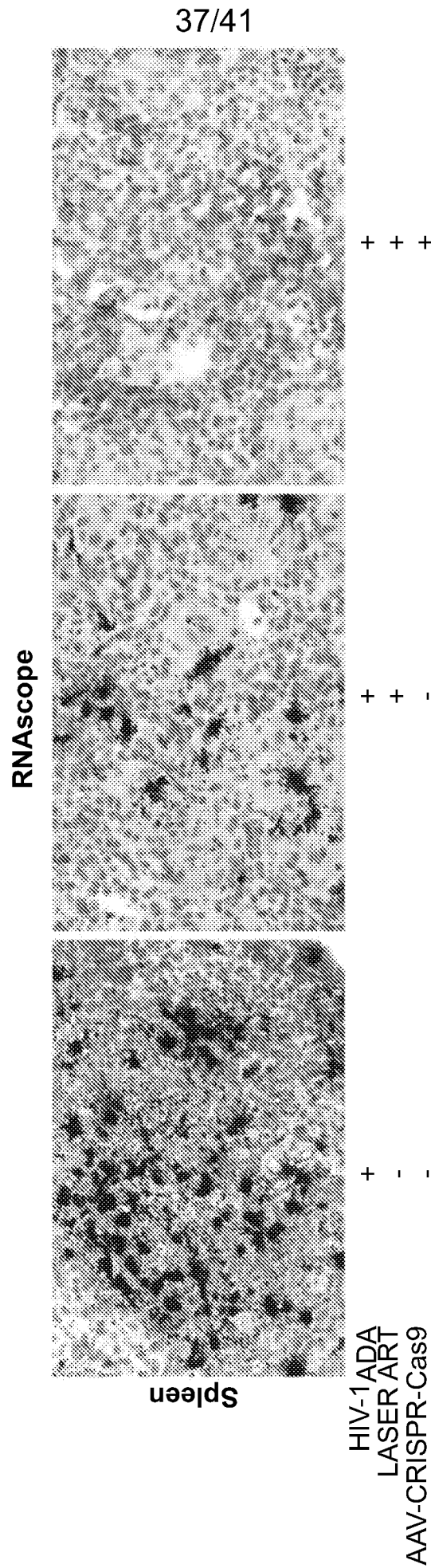
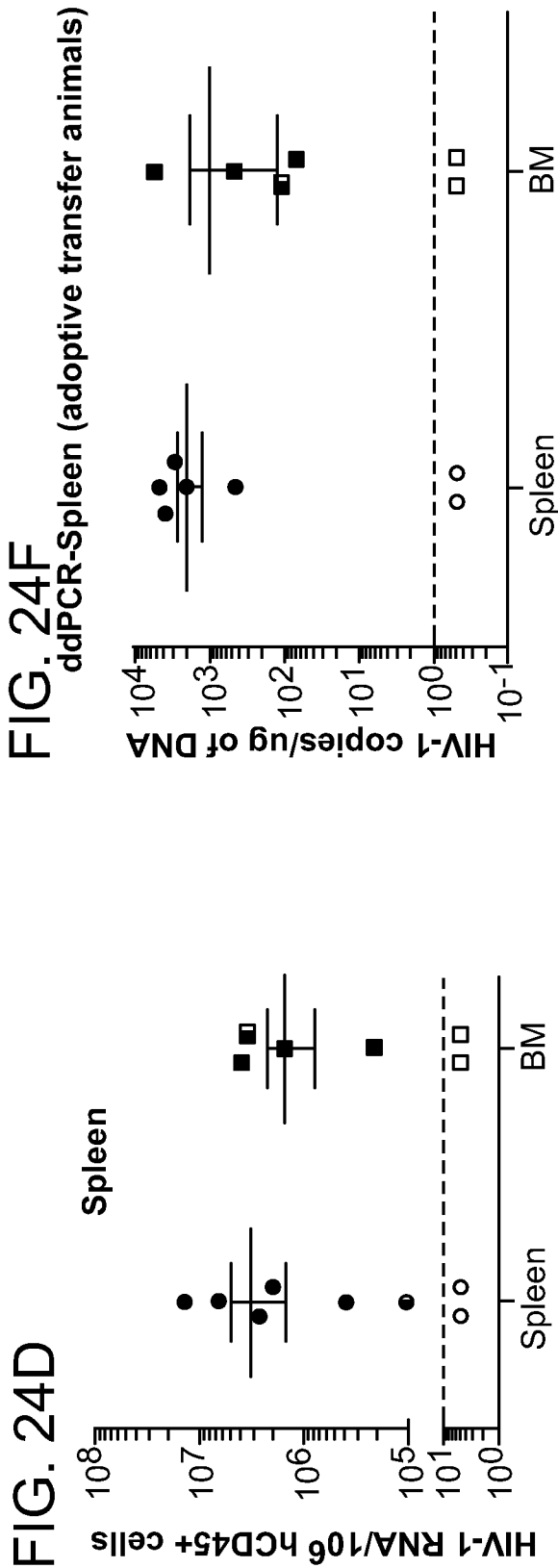
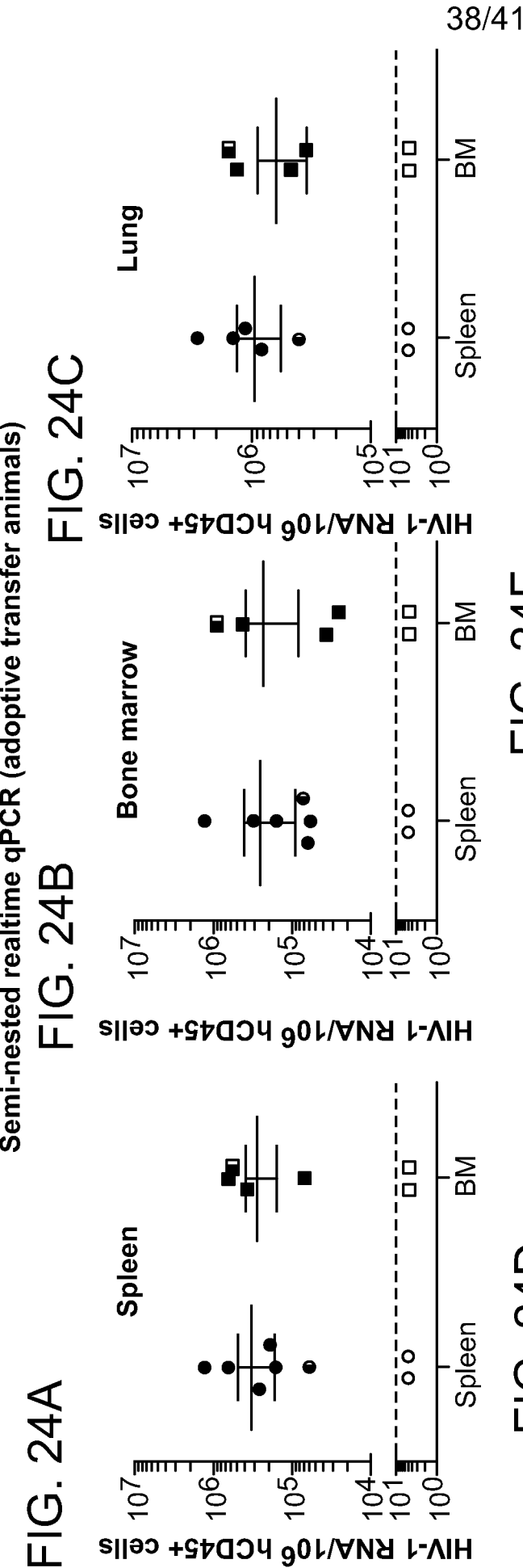
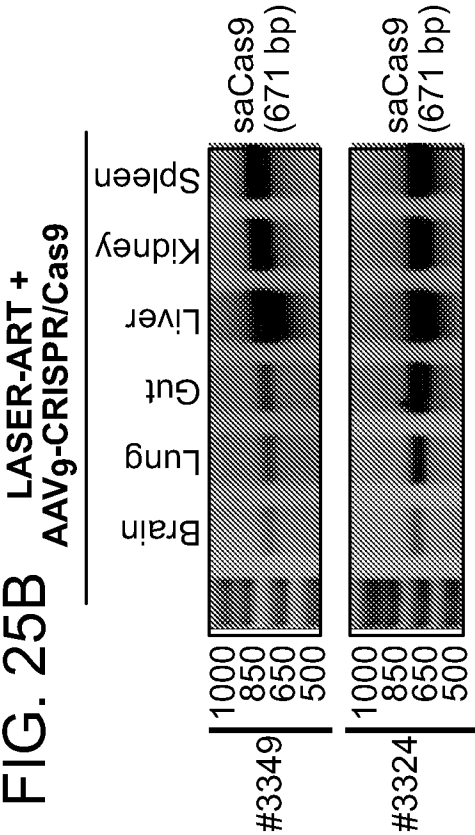
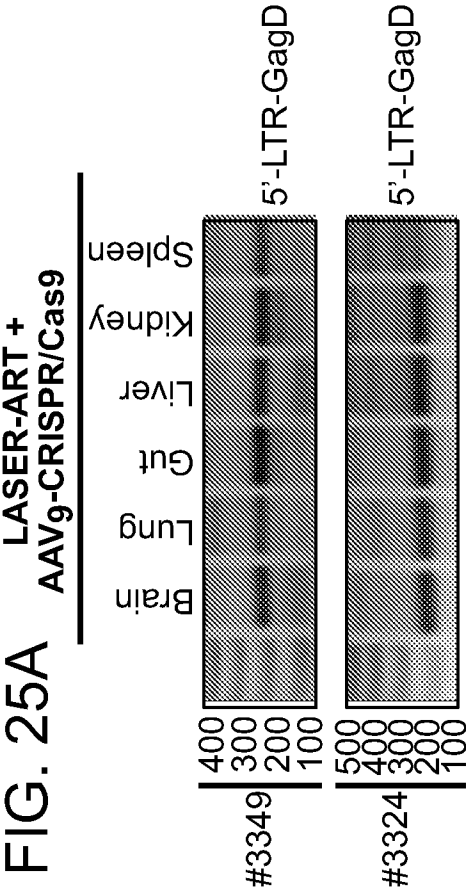
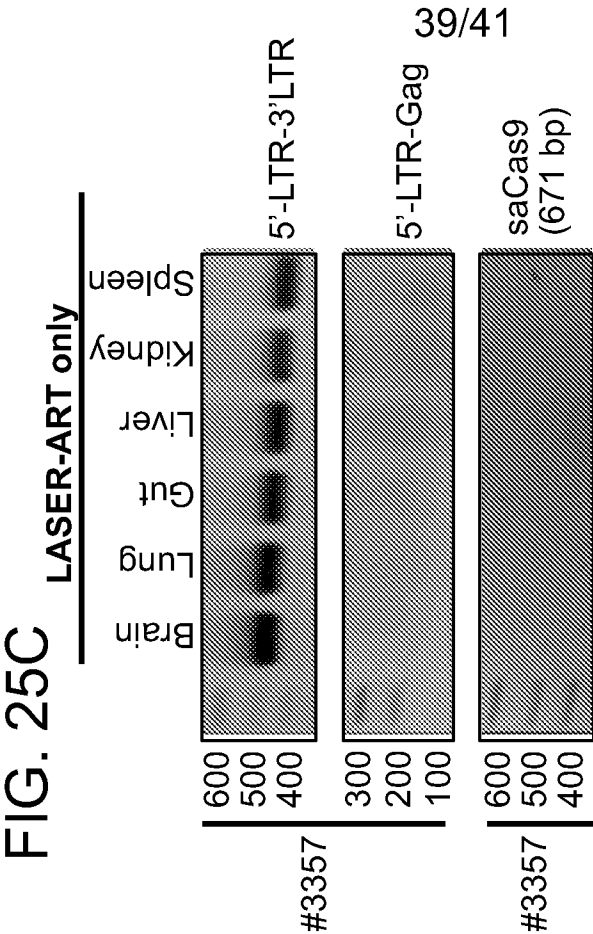


FIG. 23







40/41

FIG. 26

Liver

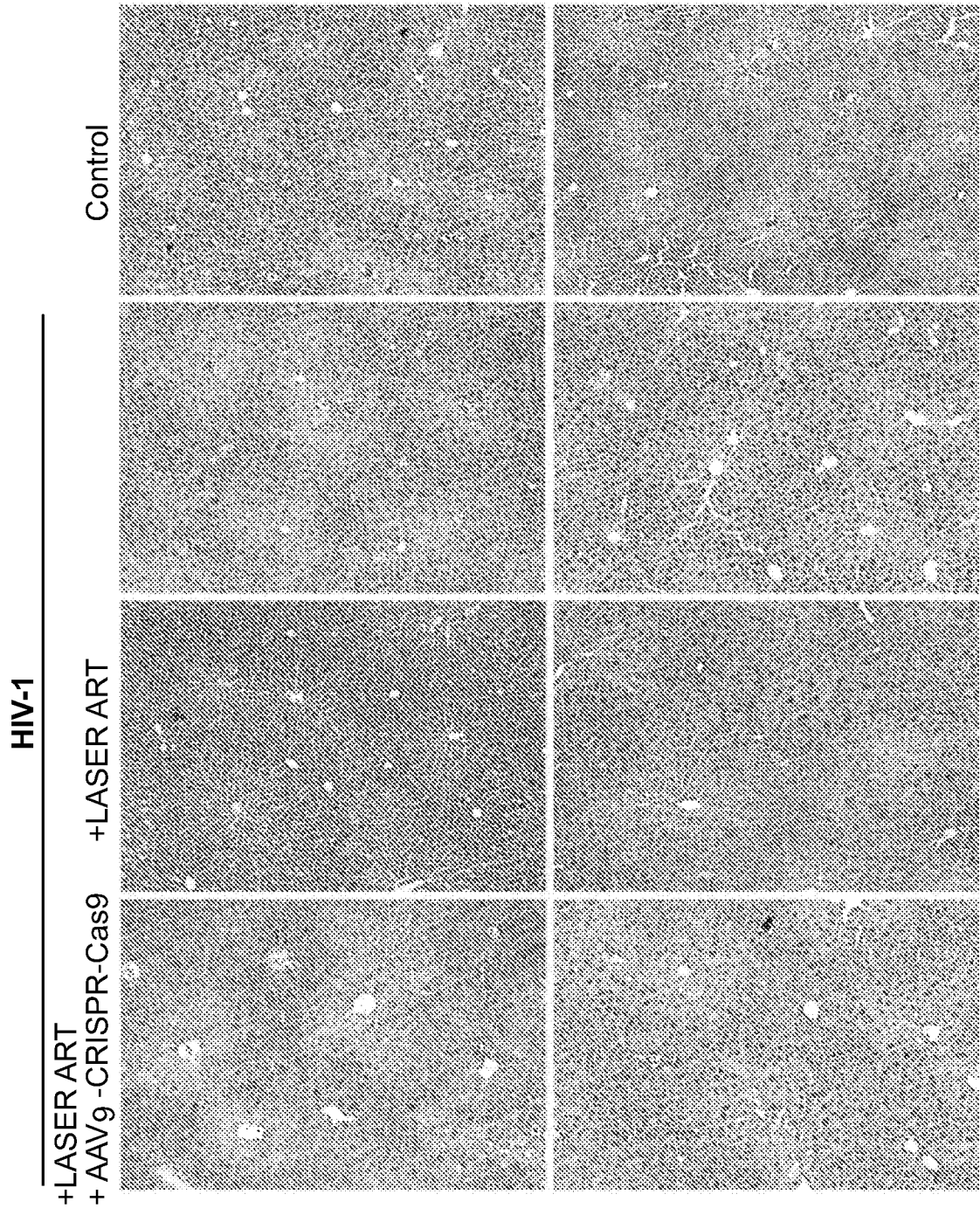
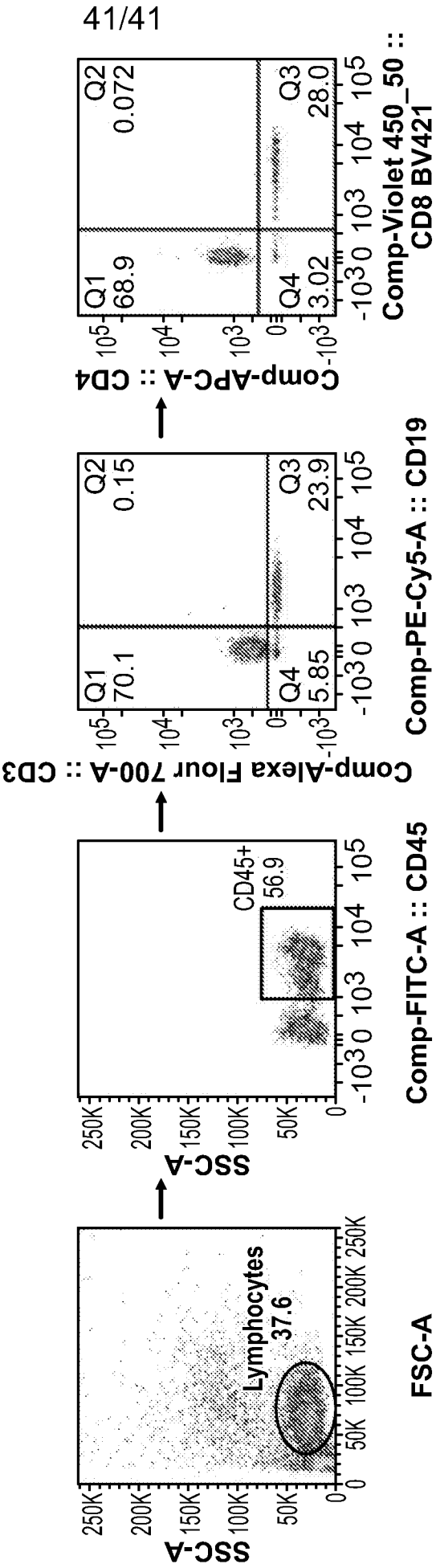


FIG. 27

Gating Strategy



INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2018/026716

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

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

a. ☐ forming part of the international application as filed:

☐ in the form of an Annex C/ST.25 text file.

☐ on paper or in the form of an image file.

b. ☐ furnished together with the international application under PCT Rule 13ter.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

c. ☒ furnished subsequent to the international filing date for the purposes of international search only:

☒ in the form of an Annex C/ST.25 text file (Rule 13ter.1(a)).

☐ on paper or in the form of an image file (Rule 13ter.1(b) and Administrative Instructions, Section 713).

2. ☒ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2018/026716

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - A61K 45/06; A61P 31/12; C07K 14/16; C12N 15/00; C12N 15/86; C12Q 1/68 (2018.01)

CPC - A61K 45/06; C12N 15/102; C12N 15/1131; C12N 15/1132; C12N 2310/20 (2018.05)

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History document

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

USPC - 424/160.1; 424/188.1; 424/192.1 (keyword delimited)

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

See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X — Y	US 2016/0017301 A1 (TEMPLE UNIVERSITY OF THE COMMONWEALTH SYSTEM OF HIGHER EDUCATION) 21 January 2016 (21.01.2016) entire document	1, 5-27 ----- 2-4
Y	EDAGWA et al. "Long-acting slow effective release antiretroviral therapy," Expert Opinion on Drug Delivery, 06 February 2017 (06.02.2017), Vol. 14, No. 11, Pgs. 1281-1291. entire document	2-4
A	US 2015/0232883 A1 (THE BROAD INSTITUTE INC. et al) 20 August 2015 (20.08.2015) entire document	1-27
A	DASH et al. "Long-acting NanoART Elicits Potent Antiretroviral and Neuroprotective Responses in HIV-1 Infected Humanized Mice", AIDS, 13 November 2012 (13.11.2012), Vol. 26, No. 17, Pgs. 2135-2144. entire document	1-27
A	NOWACEK et al. "Nanoformulated Antiretroviral Combinations Extend Drug Release and Antiretroviral Responses in HIV-1 Infected Macrophages: Implications for NeuroAIDS Therapeutics" Journal of Neuroimmune Pharmacol, 17 March 2010 (17.03.2010), Vol. 5, No. 4, Pgs. 592-601. entire document	1-27
A	WO 2017/058658 A2 (TEMPLE UNIVERSITY - OF THE COMMONWEALTH SYSTEM OF HIGHER EDUCATION) 06 April 2017 (06.04.2017) entire document	1-27



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search

05 June 2018

Date of mailing of the international search report

05 JUL 2018

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents

P.O. Box 1450, Alexandria, VA 22313-1450

Facsimile No. 571-273-8300

Authorized officer

Blaine R. Copenheaver

PCT Helpdesk: 571-272-4300

PCT OSP: 571-272-7774