



- (51) International Patent Classification:
A61K 38/46 (2006.01)
- (21) International Application Number:
PCT/US2012/026653
- (22) International Filing Date:
24 February 2012 (24.02.2012)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
61/447,672 28 February 2011 (28.02.2011) US
- (71) Applicant (for all designated States except US):
SEATTLE CHILDREN'S RESEARCH INSTITUTE
[US/US]; M/s Cw8-4, P.o. Box 5371, Seattle, WA 98145-5005 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **SCHARENBERG, Andrew, M.** [US/US]; 1222 Nw Norcross Way, Seattle, WA 98177 (US). **CERTO, Michael, T.** [US/US]; 528 20th Avenue E, Apt. #4, Seattle, WA 98112 (US). **GWIAZDA, Kamila, Sabina** [CA/US]; 5115 39th Avenue Ne, Seattle, WA 98105 (US).
- (74) Agent: **FURMAN, Eric, S.**; Knobbe Martens Olson & Bear Llp, 2040 Main Street, Fourteenth Floor, Irvine, CA 92614 (US).

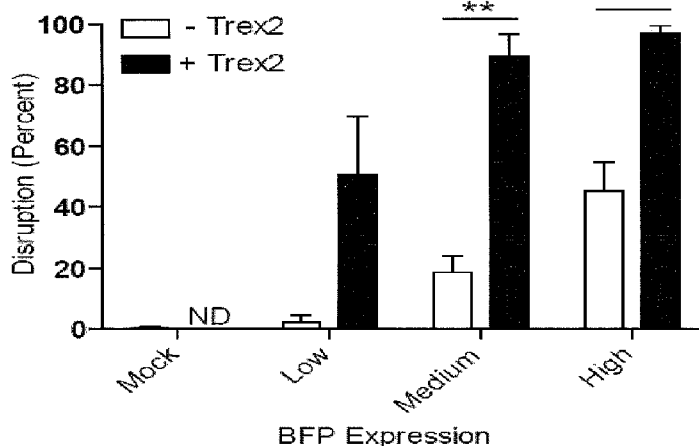
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, 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, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, 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, ML, MR, NE, SN, TD, TG).

Published:

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

(54) Title: COUPLING ENDONUCLEASES WITH END-PROCESSING ENZYMES DRIVE HIGH EFFICIENCY GENE DISRUPTION

Fig. 5B



(57) Abstract: The present disclosure relates to the co-expression of an endonuclease with an end-processing enzyme for the purpose of enhanced processing of the polynucleotide ends generated by endonuclease cleavage.

WO 2012/118717 A2

**COUPLING ENDONUCLEASES WITH END-PROCESSING ENZYMES DRIVES
HIGH EFFICIENCY GENE DISRUPTION**

RELATED APPLICATION

[0001] The present application claims the benefit of priority to U.S. Provisional Patent Application No. 61/447,672, filed February 28, 2011, the entire disclosure of which is incorporated herein by reference in its entirety.

STATEMENT REGARDING FEDERALLY SPONSORED R&D

[0002] This invention was made with government support under Grant No. T32 GM07270 awarded by the U.S. National Institute of General Medical Sciences and Grant Nos. RL1CA133832, UL1DE019582, R01-HL075453, PL1-HL092557, RL1-HL092553, RL1-HL92554, and U19-AI96111 awarded by the National Institutes of Health.

FIELD

[0003] The present disclosure relates to molecular and cellular biology. Some embodiments relate to genome engineering and the introduction of targeted, site-specific DNA breaks mediated by endonucleases to achieve gene disruption or site-specific recombination. Several embodiments relate to compositions and methods for partial or complete inactivation of a target gene. Some embodiments relate to inactivation of a targeted gene for therapeutic purposes and/or to produce cell lines in which a target gene is inactivated.

BACKGROUND

[0004] Targeted gene disruption has wide applicability for research, therapeutic, agricultural, and industrial uses. One strategy for producing targeted gene disruption is through the generation of double-strand DNA breaks caused by site-specific endonucleases. Endonucleases are most often used for targeted gene disruption in organisms that have traditionally been refractive to more conventional gene targeting methods, such as algae, plants, and large animal models, including humans. For example, there are currently human clinical trials underway involving zinc finger nucleases for the treatment and prevention of HIV infection. Additionally, endonuclease engineering is currently being used in attempts to disrupt genes that produce undesirable phenotypes in crops.

[0005] The homing endonucleases, also known as meganucleases, are sequence specific endonucleases that generate double strand breaks in genomic DNA with a high degree of specificity due to their large (e.g., >14 bp) cleavage sites. While the specificity of the homing endonucleases for their target sites allows for precise targeting of the induced DNA breaks,

homing endonuclease cleavage sites are rare and the probability of finding a naturally occurring cleavage site in a targeted gene is low.

[0006] One class of artificial endonucleases is the zinc finger endonucleases. Zinc finger endonucleases combine a non-specific cleavage domain, typically that of FokI endonuclease, with zinc finger protein domains that are engineered to bind to specific DNA sequences. The modular structure of the zinc finger endonucleases makes them a versatile platform for delivering site-specific double-strand breaks to the genome. One limitation of the zinc finger endonucleases is that low specificity for a target site or the presence of multiple target sites in a genome can result in off-target cleavage events. As FokI endonuclease cleaves as a dimer, one strategy to prevent off-target cleavage events has been to design zinc finger domains that bind at adjacent 9 base pair sites.

[0007] Another class of artificial endonucleases is the engineered meganucleases. Engineered homing endonucleases are generated by modifying the specificity of existing homing endonucleases. In one approach, variations are introduced in the amino acid sequence of naturally occurring homing endonucleases and then the resultant engineered homing endonucleases are screened to select functional proteins which cleave a targeted binding site. In another approach, chimeric homing endonucleases are engineered by combining the recognition sites of two different homing endonucleases to create a new recognition site composed of a half-site of each homing endonuclease.

[0008] The mutagenicity of the double strand DNA breaks generated by both the naturally occurring and artificial endonucleases depend upon the precision of DNA repair. The double strand breaks caused by endonucleases are commonly repaired through non-homologous end joining (NHEJ), which is the major DNA double-strand break repair pathway for many organisms. NHEJ is referred to as "non-homologous" because the break ends are ligated directly without the need for a homologous template, in contrast to homologous recombination, which utilizes a homologous sequence to guide repair. Imprecise repair through this pathway can result in mutations at the break site, such as DNA base deletions and insertions as well as translocations and telomere fusion. When the mutations are made within the coding sequence of a gene, they can render the gene and its subsequent protein product non-functional, creating a targeted gene disruption or "knockout" of the gene.

[0009] Double strand DNA break repair through the NHEJ pathway is often not mutagenic. The majority of endonuclease-induced breaks repaired by the NHEJ pathway involve precise re-ligation, resulting in the restoration of the original DNA sequence. This is especially true of the types of DNA breaks created by the current endonuclease platforms available for engineering site-specificity, namely homing endonucleases (meganucleases) and

zinc finger nucleases. Both of these types of enzymes leave compatible base pair overhangs that do not require processing prior to re-ligation by the NHEJ pathway. When the overhangs are compatible, NHEJ repairs the break with a high degree of accuracy. Thus, from a genome engineering standpoint, many of the cleavage events generated by the current site-specific endonuclease platforms are unproductive.

[0010] The need for additional solutions to these problems is manifest.

SUMMARY

[0011] Mutagenesis of cellular DNA can occur when a DNA cleavage event is followed by imprecise end joining during DNA repair. As disclosed herein, one strategy for increasing the frequency of imprecise DNA repair events is by modifying compatible overhangs generated at double-strand DNA breaks with an end-processing enzyme. The methods and compositions described herein are broadly applicable and may involve any agent of interest which generates either blunt ends or compatible overhangs upon cleaving double stranded DNA, for example, nucleases, ionizing radiation, such as x-rays and gamma rays, as well as drugs such as bleomycin, cisplatin, and mitomycin C. Several embodiments disclosed herein relate to methods for coupling the generation of double-strand DNA breaks to modification of compatible overhangs generated at the cleavage site with a DNA end-processing enzyme. Several embodiments disclosed herein relate to methods for coupling the generation of double-strand DNA breaks to modification of blunt ends generated at the cleavage site with an end-processing enzyme. Some embodiments disclosed herein relate to methods for coupling the generation of double-strand DNA breaks to cleavage of the exposed phosphodiester bonds at the DNA break site by an exonuclease. Some embodiments disclosed herein relate to methods for coupling the generation of double-strand DNA breaks to the addition of DNA bases to an exposed DNA end by a non-template polymerase.

[0012] In yet another aspect, the methods and compositions described herein are broadly applicable and may involve any agent of interest which generates breaks in a polynucleotide. Several embodiments disclosed herein relate to methods for coupling the generation of polynucleotide breaks to modification of polynucleotide ends generated at the cleavage site with an end-processing enzyme. In some embodiments, the polynucleotide may be double stranded DNA, single stranded DNA, stranded RNA, single stranded RNA, double stranded DNA/RNA hybrids and synthetic polynucleotides.

[0013] Several embodiments disclosed herein relate to a strategy for increasing the frequency of imprecise DNA repair events by modifying compatible overhangs generated at exonuclease-induced DNA breaks with a DNA end-processing enzyme. Several embodiments disclosed herein relate to methods for coupling site-specific cleavage of a targeted DNA

sequence to modification of compatible overhangs generated at the cleavage site with a DNA end-processing enzyme. Several embodiments disclosed herein relate to methods for coupling site-specific cleavage of a targeted DNA sequence to modification of blunt DNA ends generated at the cleavage site with a DNA end-processing enzyme. Some embodiments disclosed herein relate to methods for coupling site-specific cleavage of a targeted DNA sequence by an endonuclease to cleavage of the exposed phosphodiester bonds at the DNA cleavage site by an exonuclease. Some embodiments disclosed herein relate to methods for coupling site-specific cleavage of a targeted DNA sequence by an endonuclease to the addition of DNA bases to an exposed DNA end by a non-template polymerase. Some embodiments disclosed herein relate to methods for coupling site-specific cleavage of a targeted DNA sequence by an endonuclease to removal of a 5' phosphate at the DNA cleavage site by a 5'-phosphatase. Some embodiments disclosed herein relate to methods for coupling site-specific cleavage of a targeted DNA sequence by an endonuclease to removal of a 3' phosphate at the DNA cleavage site by a 3' phosphatase. Further disclosed herein are fusion proteins, comprising one or more site-specific endonuclease domains tethered to one or more DNA end-processing domains.

[0014] Non-limiting examples of endonucleases include homing endonucleases (meganucleases), zinc finger nucleases and TAL effector nucleases. The endonucleases may comprise heterologous DNA-binding and cleavage domains (e.g., zinc finger nucleases; homing endonuclease DNA-binding domains with heterologous cleavage domains or TAL-effector domain nuclease fusions) or, alternatively, the DNA-binding domain of a naturally-occurring nuclease may be altered to bind to a selected target site (e.g., a homing endonuclease that has been engineered to bind to site different than the cognate binding site or a TAL-effector domain nuclease fusion).

[0015] Non-limiting examples of DNA end-processing enzymes include 5-3' exonucleases, 3-5' exonucleases, 5-3' alkaline exonucleases, 5' flap endonucleases, helicases, phosphatases, hydrolases and template-independent DNA polymerases. The exonucleases may comprise heterologous DNA-binding and end-processing domains (e.g., a zinc finger and an exonuclease domain).

[0016] Several embodiments relate to co-expression of one or more endonucleases (enzymes that incise DNA at a specific internal target site) with one or more end-processing enzymes, in order to achieve enhanced processing of the polynucleotide ends produced by endonuclease-mediated polynucleotide cleavage. Several embodiments relate to co-expression of one or more endonucleases with one or more exonucleases (enzymes that catalyzes the removal of polynucleotide bases from an exposed polynucleotide end) in order to achieve enhanced processing of the polynucleotide ends produced by endonuclease-mediated

polynucleotide cleavage. Several embodiments relate to co-expression of one or more endonucleases with one or more non-templative polymerases (enzymes that catalyze the addition of DNA bases to an exposed DNA end) in order to achieve enhanced processing of the DNA ends produced by endonuclease-mediated DNA cleavage. Several embodiments relate to co-expression of one or more endonucleases with one or more phosphatases that catalyze the removal of a 5' phosphate in order to achieve enhanced processing of the polynucleotide ends produced by endonuclease-mediated polynucleotide cleavage. Several embodiments relate to co-expression of one or more endonucleases with one or more phosphatases that catalyze the removal of a 3' phosphate in order to achieve enhanced processing of the polynucleotide ends produced by endonuclease-mediated polynucleotide cleavage. In some embodiments, an endonuclease is coupled to an end-processing enzyme.

[0017] Several embodiments relate to co-expression of one or more endonucleases (enzymes that incise DNA at a specific internal target site) with one or more DNA end-processing enzymes, in order to achieve enhanced processing of the DNA ends produced by endonuclease-mediated DNA cleavage. Several embodiments relate to co-expression of one or more endonucleases with one or more exonucleases (enzymes that catalyze the removal of DNA bases from an exposed DNA end) in order to achieve enhanced processing of the DNA ends produced by endonuclease-mediated DNA cleavage. Several embodiments relate to co-expression of one or more endonucleases with one or more non-templative polymerases (enzymes that catalyze the addition of DNA bases to an exposed DNA end) in order to achieve enhanced processing of the DNA ends produced by endonuclease-mediated DNA cleavage. Several embodiments relate to co-expression of one or more endonucleases with one or more phosphatases that catalyze the removal of a 5' phosphate in order to achieve enhanced processing of the DNA ends produced by endonuclease-mediated DNA cleavage. Several embodiments relate to co-expression of one or more endonucleases with one or more phosphatases that catalyze the removal of a 3' phosphate in order to achieve enhanced processing of the DNA ends produced by endonuclease-mediated DNA cleavage. In some embodiments, an endonuclease is coupled to a DNA end-processing enzyme.

[0018] In one aspect, a method for improving the mutation frequency associated with endonuclease mediated cleavage of cellular DNA in a region of interest (e.g., a method for targeted disruption of genomic sequences) is provided, the method comprising: (a) selecting a sequence in the region of interest; (b) selecting a site-specific endonuclease which cleaves the sequence within the region of interest; and (c) delivering one or more fusion proteins to the cell, the fusion protein(s) comprising one or more site-specific endonuclease domains and one or more DNA end-processing domains; wherein the endonuclease domain cleaves the DNA in the

region of interest. In some embodiments, a fusion protein can be delivered to a cell by delivering a polynucleotide encoding the fusion protein to a cell. In some embodiments the polynucleotide is DNA. In other embodiments, the polynucleotide is RNA. In some embodiments, a fusion protein can be expressed in a cell by delivering a DNA vector encoding the fusion protein to a cell, wherein the DNA vector is transcribed and the mRNA transcription product is translated to generate the fusion protein. In some embodiments, a fusion protein can be expressed in a cell by delivering an RNA molecule encoding the fusion protein to the cell wherein the RNA molecule is translated to generate the fusion protein. In some embodiments, a fusion protein may be delivered directly to the cell.

[0019] In another aspect, a method for improving the mutation frequency associated with endonuclease mediated cleavage of cellular DNA in a region of interest (e.g., a method for targeted disruption of genomic sequences) is provided, the method comprising: (a) selecting a sequence in the region of interest; (b) selecting one or more site-specific endonucleases which cleaves the sequence within the region of interest; and (c) co-expressing the one or more selected endonuclease and one or more end-processing enzyme in the cell; wherein the endonuclease cleaves the DNA in the region of interest and the end-processing enzyme modifies the DNA ends exposed by the endonuclease. The nucleases and end-processing enzymes can be expressed in a cell, e.g., by delivering the proteins to the cell or by delivering one or more polynucleotides encoding the nucleases to a cell. In some embodiments, a single polynucleotide encodes both the one or more endonucleases and the one or more end-processing enzymes under the control of a single promoter. In some embodiments, one or more endonucleases and one or more end-processing enzymes are coupled by one or more T2A “skip” peptide motifs. In some embodiments, one or more endonucleases and one or more end-processing enzymes are encoded by separate polynucleotides. In some embodiments, expression of the DNA end-processing enzyme precedes that of the endonuclease.

[0020] In yet another aspect, a method for improving the mutation frequency associated with endonuclease mediated cleavage of cellular DNA in multiple regions of interest (e.g., a method for targeted disruption of multiple genomic sequences) is provided, the method comprising: (a) selecting a first sequence in a first region of interest; (b) selecting a first site-specific endonuclease which cleaves the first sequence within the first region of interest; (c) selecting a second sequence in a second region of interest; (d) selecting a second site-specific endonuclease which cleaves the second sequence within the second region of interest and (c) co-expressing the selected endonucleases and one or more end-processing enzymes in the cell; wherein the first endonuclease cleaves the DNA in the first region of interest, the second endonuclease cleaves the DNA in the second region of interest and the one or more end-

processing enzymes modify the exposed DNA ends. The nucleases and end-processing enzyme(s) can be expressed in a cell, e.g., by delivering the proteins to the cell or by delivering one or more polynucleotides encoding the nucleases and end-processing enzyme(s) to a cell. In some embodiments, a single polynucleotide encodes both the first and second endonucleases and the one or more end-processing enzyme under the control of a single promoter. In some embodiments, the endonucleases and the end-processing enzyme(s) are coupled by one or more T2A “skip” peptide motifs. In some embodiments, the first and second regions of interest are in the same gene. In other embodiments, the first and second regions of interest are in different genes. In some embodiments the method further comprises co-expression of a third, fourth, fifth, sixth, seventh, eighth, ninth, and/or tenth endonuclease in the cell.

[0021] In yet another aspect, the disclosure provides a method for treating or preventing, or inhibiting HIV infection or ameliorating a condition associated with HIV in a subject, the method comprising: (a) introducing, into a cell, a first nucleic acid encoding a first polypeptide, wherein the first polypeptide comprises: (i) a zinc finger DNA-binding domain that is engineered to bind to a first target site in the CCR5 gene; and (ii) a cleavage domain; and (iii) an end-processing domain under conditions such that the polypeptide is expressed in the cell, whereby the polypeptide binds to the target site and cleaves the CCR5 gene and end-processing enzyme domain modifies the endonuclease cleavage site; and (b) introducing the cell into the subject. In certain embodiments, the cell is selected from the group consisting of a hematopoietic stem cell, a T-cell, a macrophage, a dendritic cell, and an antigen-presenting cell.

[0022] In yet another aspect, the disclosure provides a method for treating or preventing or inhibiting HIV infection or ameliorating a condition associated with HIV in a subject, the method comprising: (a) introducing, into a cell, a first nucleic acid encoding a first polypeptide and a second polypeptide, wherein the first polypeptide comprises: (i) a zinc finger DNA-binding domain that is engineered to bind to a first target site in the CCR5 gene; and (ii) a cleavage domain; and the second polypeptide comprises a end-processing enzyme under conditions such that the polypeptides are co-expressed in the cell, whereby the first polypeptide binds to the target site and cleaves the CCR5 gene and the end-processing enzyme modifies the exposed DNA ends created at the endonuclease cleavage site; and (b) introducing the cell into the subject. In certain embodiments, the cell is selected from the group consisting of a hematopoietic stem cell, a T-cell, a macrophage, a dendritic cell and an antigen-presenting cell.

[0023] In another aspect, the disclosure provides a method for treating or preventing or inhibiting HIV infection or ameliorating a condition associated with HIV in a subject, the method comprising: (a) introducing, into a cell, a nucleic acid encoding a polypeptide, wherein the polypeptide comprises: (i) a homing endonuclease domain that is engineered to bind to a first

target site in the CCR5 gene; and (ii) a end-processing domain under conditions such that the polypeptide is expressed in the cell, whereby the polypeptide binds to the target site and cleaves the CCR5 gene and modifies the exposed DNA ends created at the cleavage site; and (b) introducing the cell into the subject. In certain embodiments, the DNA end-processing domain comprises an exonuclease.

[0024] In another aspect, the disclosure provides a method for treating or preventing or inhibiting HIV infection or ameliorating a condition associated with HIV in a subject, the method comprising: (a) introducing, into a cell, a nucleic acid encoding a first polypeptide and a second polypeptide, wherein the first polypeptide comprises a homing endonuclease that is engineered to bind to a target site in the CCR5 gene; and the second polypeptide comprises a end-processing enzyme under conditions such that the polypeptides are co-expressed in the cell, whereby the first polypeptide binds to the target site and cleaves the CCR5 gene and the end-processing enzyme modifies the exposed DNA ends created at the endonuclease cleavage site; and (b) introducing the cell into the subject. In certain embodiments, the end-processing enzyme comprises an exonuclease. In some embodiments, the homing endonuclease and the end-processing enzyme are coupled by one or more T2A “skip” peptide motifs.

[0025] In yet another aspect, the disclosure provides a method for treating or preventing or inhibiting HIV infection or ameliorating a condition associated with HIV in a subject, the method comprising: (a) introducing, into a cell, a first nucleic acid encoding a first polypeptide, wherein the first polypeptide comprises: a homing endonuclease that is engineered to bind to a first target site in the CCR5 gene; and (b) introducing, into the cell, a second nucleic acid encoding a second polypeptide, wherein the second polypeptide comprises: a end-processing enzyme; under conditions such that the polypeptides are expressed in the cell, whereby the homing endonuclease binds to the target site and cleaves the CCR5 gene and the end-processing enzyme modifies the exposed DNA ends created at the endonuclease cleavage site; and (b) introducing the cell into the subject. In certain embodiments, the end-processing enzyme comprises an exonuclease. In some embodiments, expression of the end-processing enzyme precedes that of the endonuclease.

[0026] In another aspect, the disclosure provides a method for treating or preventing or inhibiting hyper IGE syndrome or ameliorating a condition associated with hyper IGE syndrome a subject, the method comprising: (a) introducing, into one or more cells, a nucleic acid encoding a polypeptide, wherein the polypeptide comprises: (i) a homing endonuclease domain that is engineered to bind to a first target site in the Stat3 gene; and (ii) a end-processing domain under conditions such that the polypeptide is expressed in the cell, whereby the polypeptide binds to the target site and cleaves the Stat3 gene and modifies the exposed DNA

ends created at the endonuclease cleavage site. In certain embodiments, the end-processing enzyme domain comprises an exonuclease.

[0027] In yet another aspect, the disclosure provides a method for treating or preventing or inhibiting hyper IGE syndrome or ameliorating a condition associated with hyper IGE syndrome a subject, the method comprising: (a) introducing, into a cell, a first nucleic acid encoding a first polypeptide, wherein the first polypeptide comprises: a homing endonuclease that is engineered to bind to a first target site in the STAT3 gene; and (b) introducing, into the cell, a second nucleic acid encoding a second polypeptide, wherein the second polypeptide comprises: a end-processing enzyme; under conditions such that the polypeptides are expressed in the cell, whereby the homing endonuclease binds to the target site and cleaves the STAT3 gene and the end-processing enzyme modifies the exposed DNA ends created at the endonuclease cleavage site. In certain embodiments, the end-processing enzyme comprises an exonuclease. In some embodiments, the expression of the end-processing enzyme precedes that of the endonuclease.

[0028] In yet another aspect, the disclosure provides a method for treating or preventing or inhibiting hyper IGE syndrome or ameliorating a condition associated with hyper IGE syndrome a subject, the method comprising: (a) introducing, into a cell, a nucleic acid encoding a first polypeptide and a second polypeptide, wherein the first polypeptide comprises a homing endonuclease that is engineered to bind to a first target site in the STAT3 gene and the second polypeptide comprises a end-processing enzyme; under conditions such that the polypeptides are co-expressed in the cell, whereby the homing endonuclease binds to the target site and cleaves the STAT3 gene and the end-processing enzyme modifies the exposed DNA ends created at the endonuclease cleavage site. In certain embodiments, the end-processing enzyme comprises an exonuclease. In some embodiments, the homing endonuclease and the end-processing enzyme are coupled by one or more T2A “skip” peptide motifs.

BRIEF DESCRIPTION OF THE DRAWINGS

[0029] Figure 1A shows a schematic of the Traffic Light Reporter system (TLR) for measuring the effectiveness of exonuclease induced gene disruption. mCherry positive cells represent a proportion of the total cells that have undergone gene disruption. Figures 1B-1H show schematic representations of expression vectors for delivery of endonucleases and DNA end-processing enzymes.

[0030] Figure 2A shows representative flow plots of HEK293 cells harboring Traffic light Reporter transfected with expression vectors encoding SceD44A-IRES-BFP, SceD44A-T2A-Trex2-IRES-BFP, I-SceI-IRES-BFP, and I-SceI-T2A-Trex2-IRES-BFP. SceD44A corresponds to an inactive mutant form of I-SceI. Figure 2B shows quantification of

gene disruption in three independent transfections of the vectors indicated in Figure 2A. Error bars represent standard error of the mean (SEM), and *p*-values (with * representing $p < 0.05$, ** $p < 0.005$, and *** $p < 0.0005$) were calculated using the Student's two-tailed unpaired t-test to compare the samples indicated in this and all subsequent figures.

[0031] Figure 3A shows representative flow plots of HEK293 cells harboring Traffic light Reporter transfected with expression vectors encoding I-SceI-IRES-BFP, I-SceI-T2A-Trex2-BFP, or I-SceI-G4S-Trex2-IRES BFP. Figure 3B shows an anti-HA western blot demonstrating equal expression of endonucleases, and stability of the (HA-)I-SceI, (HA-)I-SceI-T2A and (HA-)I-SceI-G4S-Trex2 proteins from Figure 3A. Figure 3C is a licor western blot showing size and stability of the HA-tagged I-SceI in indicated HEK293T lysates.

[0032] Figure 4A shows gating analysis of HEK293 cells harboring Traffic Light Reporter transfected with I-SceI-IRES-BFP. Figure 4B shows a gating analysis of HEK293 cells harboring Traffic Light Reporter transfected with I-SceI-T2A-Trex2-IRES-BFP expression vectors.

[0033] Figure 5A shows an I-SceI restriction digest of amplicons flanking the I-SceI target site from HEK293 cells harboring traffic light reporter sorted by BFP expression levels follow transfection with expression constructs as indicated in Figure 4A and Figure 4B. Figure 5B shows quantification of three independent experiments as described in Figure 5A.

[0034] Figure 6 shows the results of DNA sequencing of amplicons surrounding the I-SceI target site in HEK293 Traffic Light Reporter cells treated with I-SceI-IRES-BFP or I-SceI-T2A-Trex2-IRES-BFP.

[0035] Figure 7 shows a graph scoring observed mutations (deletions are negative, insertions are positive) at the I-SceI target site following transfection of HEK293 Traffic Light Reporter cells with I-SceI-IRES-BFP or I-SceI-T2A-Trex2-IRES-BFP as described in Figure 5.

[0036] Figure 8A shows a kinetic time course analysis demonstrating transient expression of I-SceI-T2A-Trex2-IRES-BFP after transfection into HEK293 cells harboring Traffic Light Reporter. The constructs shown are tagged to BFP by an IRES sequence downstream of either I-SceI or Trex2. Figure 8B shows a graph quantifying 3 experiments of HEK293T cells transfected with the vectors indicated in Figure 8A, analyzed at the indicated time-points. Cherry indicates gene disruption rates observed in transfected cells. Figure 9A shows an I-SceI restriction digest of amplicons from primary murine embryonic fibroblasts spanning an I-SceI target site 72 hours post transduction with I-SceI-IRES BFP or I-SceI-T2A-Trex2-IRES-BFP. Figure 9B shows a graph quantifying cleavage site disruption in 2 independent experiments. Figure 9C shows an I-SceI restriction digest of amplicons from lineage depleted bone marrow spanning an I-SceI target site 72 hours post transduction with I-

SceI-IRES BFP or I-SceI-T2A-Trex2-IRES-BFP. Figure 9D shows quantification of bands from Figure 9C.

[0037] Figure 10 shows a graph quantifying gene disruption rates of several different homing endonucleases with and without Trex2 exonuclease as measured by HEK293 cells harboring Traffic Light Reporters with respective target sites for the indicated homing endonucleases.

[0038] Figure 11A shows representative flow plots and targets sites of HEK293 Traffic Light Reporter cells following transfection with a homing endonuclease with and without Trex2 and a zinc finger nuclease with and without Trex2. Figure 11B shows a graph of an independent experiment examining cleavage site mutation for I-SceI and Zinc Finger Nuclease in the presence and absence of Trex2. Figure 11C shows a graph of HEK293 Traffic Light Reporter cells following co-transfection of an HE with Trex2 or a TALEN with Trex2.

[0039] Figure 12A shows representative flow plots of HEK293 cells harboring Traffic Light Reporters with an I-AniI target site following transfection with either I-AniI-IRES-BFP, I-AniI-T2A-Trex2-IRES-BFP, I-AniIY2-IRES-BFP, I-AniIY2-T2A-Trex2-IRES-BFP. Figure 12B shows a graph quantitating 3 independent experiments as performed in Figure 12A.

[0040] Figure 13 shows graph depicting cell cycle analysis of murine embryonic fibroblasts transduced with Mock, I-SceI-IRES-BFP, or I-SceI-T2A-Trex2-IRES-BFP viruses.

[0041] Figure 14 shows a graph depicting maintenance of BFP expression in cells transduced with an integrating lentivirus containing I-SceID44A-IRES-BFP or I-SceID44A-T2A-Trex2-IRES-BFP.

[0042] Figure 15A shows a graph measuring human CD34+ hematopoietic stem cell survival when transduced with I-SceID44A-IRES-BFP or I-SceID44A-T2A-Trex2-IRES-BFP and challenged with Mitomycin C. Figure 15B shows a graph measuring human CD34+ hematopoietic stem cell survival when transduced with I-SceID44A-IRES-BFP or I-SceID44A-T2A-Trex2-IRES-BFP and challenged with camptothecin. Figure 15C shows a graph measuring human CD34+ hematopoietic stem cell survival when transduced with I-SceID44A-IRES-BFP or I-SceID44A-T2A-Trex2-IRES-BFP and challenged with ionizing radiation.

[0043] Figure 16A shows a graph measuring murine embryonic fibroblast cell survival when transduced with I-SceID44A-IRES-BFP or I-SceID44A-T2A-Trex2-IRES-BFP and challenged with Mitomycin C. Figure 16B shows a graph measuring murine embryonic fibroblast cell survival when transduced with I-SceID44A-IRES-BFP or I-SceID44A-T2A-Trex2-IRES-BFP and challenged with camptothecin.

[0044] Figure 17A shows representative flow plots of HEK293 Traffic Light Reporter cells following co-transfection of I-SceI-IRES-BFP and an expression plasmid coding

for the indicated end-processing enzyme. Figure 17B shows a graph quantifying 3 independent experiments as performed in Figure 17A.

[0045] Figure 18A shows representative flow plots of a gating analysis of I-SceI-IRES-BFP co-transfected with ARTEMIS expression plasmid as indicated in Figure 17A. Figure 18B shows a graph quantifying gating analysis of several end-processing enzymes from 3 independent experiments as indicated in Figure 18A.

[0046] Figure 19A shows a graph of HEK293 Traffic Light Reporter cells following co-transfection with a zinc finger nuclease and the indicated end-processing enzyme expression plasmid. Figure 19B shows a graph of HEK293 Traffic Light Reporter cells following co-transfection with a TALEN and the indicated end-processing enzyme expression plasmid.

[0047] Figure 20 shows a comparison of expression levels and gene disruption rates between integrating lentivirus and integrase deficient lentivirus from I-SceI with and without exonuclease coupling on HEK293 Traffic Light reporter cells.

[0048] Figure 21A shows live cell image of cells 72 hrs post mock transfection or transfection with an expression vectors encoding I-SceI-IRES-BFP or I-SceI-T2A-Trex2-IRES-BFP. Figure 21B shows a graph depicting maintenance of BFP expression in cells transduced with an integrating lentivirus containing BFP alone (no Trex2) or Trex2-BFP.

DETAILED DESCRIPTION

Definitions

[0049] In the description that follows, a number of terms are used extensively. The following definitions are provided to facilitate understanding of the present embodiments.

[0050] As used herein, “a” or “an” may mean one or more than one.

[0051] As used herein, the term “about” indicates that a value includes the inherent variation of error for the method being employed to determine a value, or the variation that exists among experiments.

[0052] As used herein, “nucleic acid” or “nucleic acid molecule” refers to polynucleotides, such as deoxyribonucleic acid (DNA) or ribonucleic acid (RNA), oligonucleotides, fragments generated by the polymerase chain reaction (PCR), and fragments generated by any of ligation, scission, endonuclease action, and exonuclease action. Nucleic acid molecules can be composed of monomers that are naturally-occurring nucleotides (such as DNA and RNA), or analogs of naturally-occurring nucleotides (e.g., enantiomeric forms of naturally-occurring nucleotides), or a combination of both. Modified nucleotides can have alterations in sugar moieties and/or in pyrimidine or purine base moieties. Sugar modifications include, for example, replacement of one or more hydroxyl groups with halogens, alkyl groups, amines, and azido groups, or sugars can be functionalized as ethers or esters. Moreover, the

entire sugar moiety can be replaced with sterically and electronically similar structures, such as aza-sugars and carbocyclic sugar analogs. Examples of modifications in a base moiety include alkylated purines and pyrimidines, acylated purines or pyrimidines, or other well-known heterocyclic substitutes. Nucleic acid monomers can be linked by phosphodiester bonds or analogs of such linkages. Analogs of phosphodiester linkages include phosphorothioate, phosphorodithioate, phosphoroselenoate, phosphorodiselenoate, phosphoroanilothioate, phosphoranilidate, phosphoramidate, and the like. The term “nucleic acid molecule” also includes so-called “peptide nucleic acids,” which comprise naturally-occurring or modified nucleic acid bases attached to a polyamide backbone. Nucleic acids can be either single stranded or double stranded.

[0053] The term “contig” denotes a nucleic acid molecule that has a contiguous stretch of identical or complementary sequence to another nucleic acid molecule. Contiguous sequences are said to “overlap” a given stretch of a nucleic acid molecule either in their entirety or along a partial stretch of the nucleic acid molecule.

[0054] The term “degenerate nucleotide sequence” denotes a sequence of nucleotides that includes one or more degenerate codons as compared to a reference nucleic acid molecule that encodes a polypeptide. Degenerate codons contain different triplets of nucleotides, but encode the same amino acid residue (e.g., GAU and GAC triplets each encode Asp). It will be understood that, as a result of the degeneracy of the genetic code, a multitude of nucleotide sequences encoding a given protein such as an endonuclease, end-processing enzyme, or endonuclease/end-processing enzyme fusion protein of the present embodiments may be produced.

[0055] The term “complementary to” means that the complementary sequence is homologous to all or a portion of a reference polynucleotide sequence. For illustration, the nucleotide sequence “CATTAG” corresponds to a reference sequence “CATTAG” and is complementary to a reference sequence “GTAATC.”

[0056] The term “structural gene” refers to a nucleic acid molecule that is transcribed into messenger RNA (mRNA), which is then translated into a sequence of amino acids characteristic of a specific polypeptide.

[0057] An “isolated nucleic acid molecule” is a nucleic acid molecule that is not integrated in the genomic DNA of an organism. For example, a DNA molecule that encodes a growth factor that has been separated from the genomic DNA of a cell is an isolated DNA molecule. Another non-limiting example of an isolated nucleic acid molecule is a chemically-synthesized nucleic acid molecule that is not integrated in the genome of an organism. A

nucleic acid molecule that has been isolated from a particular species is smaller than the complete DNA molecule of a chromosome from that species.

[0058] “Complementary DNA (cDNA)” is a single-stranded DNA molecule that is formed from an mRNA template by the enzyme reverse transcriptase. Typically, a primer complementary to portions of mRNA is employed for the initiation of reverse transcription. Those skilled in the art may also use the term “cDNA” to refer to a double-stranded DNA molecule consisting of such a single-stranded DNA molecule and its complementary DNA strand. The term “cDNA” may also refer to a clone of a cDNA molecule synthesized from an RNA template.

[0059] A “promoter” is a nucleotide sequence that directs the transcription of a structural gene. In some embodiments, a promoter is located in the 5’ non-coding region of a gene, proximal to the transcriptional start site of a structural gene. Sequence elements within promoters that function in the initiation of transcription are often characterized by consensus nucleotide sequences. These promoter elements include RNA polymerase binding sites, TATA sequences, CAAT sequences, differentiation-specific elements (DSEs; McGehee et al., *Mol. Endocrinol.* 7:551 (1993)), cyclic AMP response elements (CREs), serum response elements (SREs; Treisman, *Seminars in Cancer Biol.* 1:47 (1990)), glucocorticoid response elements (GREs), and binding sites for other transcription factors, such as CRE/ATF (O’Reilly et al., *J. Biol. Chem.* 267:19938 (1992)), AP2 (Ye et al., *J. Biol. Chem.* 269:25728 (1994)), SP1, cAMP response element binding protein (CREB; Loeken, *Gene Expr.* 3:253 (1993)) and octamer factors (see, in general, Watson et al., eds., *Molecular Biology of the Gene*, 4th ed. (The Benjamin/Cummings Publishing Company, Inc. 1987), and Lemaigre and Rousseau, *Biochem. J.* 303:1 (1994)). As used herein, a promoter may be constitutively active, repressible or inducible. If a promoter is an inducible promoter, then the rate of transcription increases in response to an inducing agent. In contrast, the rate of transcription is not regulated by an inducing agent if the promoter is a constitutive promoter. Repressible promoters are also known.

[0060] A “core promoter” contains essential nucleotide sequences for promoter function, including the TATA box and start of transcription. By this definition, a core promoter may or may not have detectable activity in the absence of specific sequences that may enhance the activity or confer tissue specific activity.

[0061] A “regulatory element” is a nucleotide sequence that modulates the activity of a core promoter. For example, a regulatory element may contain a nucleotide sequence that binds with cellular factors enabling transcription exclusively or preferentially in particular cells,

tissues, or organelles. These types of regulatory elements are normally associated with genes that are expressed in a “cell-specific,” “tissue-specific,” or “organelle-specific” manner.

[0062] An “enhancer” is a type of regulatory element that can increase the efficiency of transcription, regardless of the distance or orientation of the enhancer relative to the start site of transcription.

[0063] “Heterologous DNA” refers to a DNA molecule, or a population of DNA molecules, that does not exist naturally within a given host cell. DNA molecules heterologous to a particular host cell may contain DNA derived from the host cell species (e.g., endogenous DNA) so long as that host DNA is combined with non-host DNA (e.g., exogenous DNA). For example, a DNA molecule containing a non-host DNA segment encoding a polypeptide operably linked to a host DNA segment comprising a transcription promoter is considered to be a heterologous DNA molecule. Conversely, a heterologous DNA molecule can comprise an endogenous gene operably linked with an exogenous promoter. As another illustration, a DNA molecule comprising a gene derived from a wild-type cell is considered to be heterologous DNA if that DNA molecule is introduced into a mutant cell that lacks the wild-type gene.

[0064] A “polypeptide” is a polymer of amino acid residues joined by peptide bonds, whether produced naturally or synthetically. Polypeptides of less than about 10 amino acid residues are commonly referred to as “peptides.”

[0065] A “protein” is a macromolecule comprising one or more polypeptide chains. A protein may also comprise non-peptide components, such as carbohydrate groups. Carbohydrates and other non-peptide substituents may be added to a protein by the cell in which the protein is produced, and will vary with the type of cell. Proteins are defined herein in terms of their amino acid backbone structures; substituents such as carbohydrate groups are generally not specified, but may be present nonetheless.

[0066] A peptide or polypeptide encoded by a non-host DNA molecule is a “heterologous” peptide or polypeptide.

[0067] An “integrated genetic element” is a segment of DNA that has been incorporated into a chromosome of a host cell after that element is introduced into the cell through human manipulation. Within the present embodiments, integrated genetic elements are most commonly derived from linearized plasmids that are introduced into the cells by electroporation or other techniques. Integrated genetic elements are passed from the original host cell to its progeny.

[0068] A “cloning vector” is a nucleic acid molecule, such as a plasmid, cosmid, plastome, or bacteriophage that has the capability of replicating autonomously in a host cell. Cloning vectors typically contain one or a small number of restriction endonuclease recognition

sites that allow insertion of a nucleic acid molecule in a determinable fashion without loss of an essential biological function of the vector, as well as nucleotide sequences encoding a marker gene that is suitable for use in the identification and selection of cells transduced with the cloning vector. Marker genes typically include genes that provide tetracycline resistance or ampicillin resistance.

[0069] An “expression vector” is a nucleic acid molecule encoding a gene that is expressed in a host cell. Typically, an expression vector comprises a transcription promoter, a gene, and a transcription terminator. Gene expression is usually placed under the control of a promoter, and such a gene is said to be “operably linked to” the promoter. Similarly, a regulatory element and a core promoter are operably linked if the regulatory element modulates the activity of the core promoter.

[0070] As used herein, “transient transfection” refers to the introduction of exogenous nucleic acid(s) into a host cell by a method that does not generally result in the integration of the exogenous nucleic into the genome of the transiently transfected host cell.

[0071] By the term “host cell” is meant a cell that contains one or more nucleases, for example endonucleases, end-processing enzymes, and/or endonuclease/end-processing enzyme fusion proteins encompassed by the present embodiments or a vector encoding the same that supports the replication, and/or transcription or transcription and translation (expression) of one or more nucleases, for example endonucleases, end-processing enzymes, and/or endonuclease/end-processing enzyme fusion proteins. Host cells for use in the present invention can be prokaryotic cells or eukaryotic cells. Examples of prokaryotic host cells include, but are not limited to *E. coli*, nitrogen fixing bacteria, *Staphylococcus aureus*, *Staphylococcus albus*, *Lactobacillus acidophilus*, *Bacillus anthracis*, *Bacillus subtilis*, *Bacillus thuringiensis*, *Clostridium tetani*, *Clostridium botulinum*, *Streptococcus mutans*, *Streptococcus pneumoniae*, mycoplasmas, and cyanobacteria. Examples of eukaryotic host cells include, but are not limited to, protozoa, fungi, algae, plant, insect, amphibian, avian and mammalian cells.

[0072] “Integrative transformants” are recombinant host cells, in which heterologous DNA has become integrated into the genomic DNA of the cells.

[0073] An “isolated polypeptide” is a polypeptide that is essentially free from contaminating cellular components, such as carbohydrate, lipid, or other proteinaceous impurities associated with the polypeptide in nature. Typically, a preparation of isolated polypeptide contains the polypeptide in a highly purified form, e.g., at least about 80% pure, at least about 90% pure, at least about 95% pure, greater than 95% pure, or greater than 99% pure. One way to show that a particular protein preparation contains an isolated polypeptide is by the appearance of a single band following sodium dodecyl sulfate (SDS)-polyacrylamide gel

electrophoresis of the protein preparation and Coomassie Brilliant Blue staining of the gel. However, the term “isolated” does not exclude the presence of the same polypeptide in alternative physical forms, such as dimers, or alternatively glycosylated or derivative forms.

[0074] The terms “amino-terminal” and “carboxyl-terminal” are used herein to denote positions within polypeptides. Where the context allows, these terms are used with reference to a particular sequence or portion of a polypeptide to denote proximity or relative position. For example, a certain sequence positioned carboxyl-terminal to a reference sequence within a polypeptide is located proximal to the carboxyl terminus of the reference sequence, but is not necessarily at the carboxyl terminus of the complete polypeptide.

[0075] The term “gene expression” refers to the biosynthesis of a gene product. For example, in the case of a structural gene, gene expression involves transcription of the structural gene into mRNA and the translation of mRNA into one or more polypeptides.

[0076] The term “endonuclease” refers to enzymes that cleave the phosphodiester bond within a polynucleotide chain. The polynucleotide may be double-stranded DNA (dsDNA), single-stranded DNA (ssDNA), RNA, double-stranded hybrids of DNA and RNA, and synthetic DNA (for example, containing bases other than A, C, G, and T). An endonuclease may cut a polynucleotide symmetrically, leaving “blunt” ends, or in positions that are not directly opposing, creating overhangs, which may be referred to as “sticky ends.” The methods and compositions described herein may be applied to cleavage sites generated by endonucleases.

[0077] The term “homing endonuclease” refers to double stranded DNases that have large, asymmetric recognition sites (12-40 base pairs). Homing endonuclease recognition sites are extremely rare. For example, an 18 base pair recognition sequence will occur only once in every 7×10^{10} base pairs of random sequence. This is equivalent to only one site in 20 mammalian-sized genomes. Unlike standard restriction endonucleases, however, homing endonucleases tolerate some sequence degeneracy within their recognition sequence. As a result, their observed sequence specificity is typically in the range of 10-12 base pairs. Although the cleavage specificity of most homing endonucleases is not absolute with respect to their recognition sites, the sites are of sufficient length that a single cleavage event per mammalian-sized genome can be obtained by expressing a homing endonuclease in a cell containing a single copy of its recognition site. Examples of homing endonucleases include, but are not limited to, I-Anil, I-SceI, I-CeuI, PI-PspI, PI-Sce, I-SceIV, I-CsmI, I-PanI, I-PanII, I-PanMI, I-SceII, I-PpoI, I-SceIII, I-CreI, I-LtrI, I-GpiI, I-GZeI, I-OnuI, I-HjeMI, I-TevI, I-TevII, and I-TevIII. Their recognition sequences are known. The specificity of homing endonucleases and meganucleases can be engineered to bind non-natural target sites. See, for example, Chevalier et al. (2002) *Molec. Cell* 10:895-905; Epinat et al. (2003) *Nucleic Acids Res.*

31:2952-2962; Ashworth et al. (2006) *Nature* 441:656-659; Paques et al. (2007) *Current Gene Therapy* 7:49-66. The methods and compositions described herein may be applied to cleavage sites generated by homing endonucleases.

[0078] The term “TAL effector nuclease” (TALEN) refers to a nuclease comprising a TAL-effector domain fused to a nuclease domain. TAL-effector DNA binding domains, isolated from the plant pathogen *Xanthomonas* have been described (see Boch et al., (2009) *Science* 29 Oct. 2009 (10.1126/science.117881) and Moscou and Bogdanove, (2009) *Science* 29 Oct. 2009 (10.1126/science.1178817)). These DNA binding domains may be engineered to bind to a desired target and fused to a nuclease domain, such as the FokI nuclease domain, to derive a TAL effector domain-nuclease fusion protein. The methods and compositions described herein may be applied to cleavage sites generated by TAL effector nucleases.

[0079] The term “Zinc-finger nuclease” (ZFN) refers to artificial restriction enzymes generated by fusing a zinc finger DNA-binding domain to a DNA-cleavage domain. Zinc finger domains can be engineered to bind to a desired target site. In some embodiments, the cleavage domain comprises the non-specific cleavage domain of FokI. In other embodiments, the cleavage domain comprises all or an active portion of another nuclease. In some embodiments, the cleavage domain may comprise Trex2 or an active fragment thereof. The methods and compositions described herein may be applied to cleavage sites generated by zinc-finger nucleases

[0080] The term “end-processing enzyme” refers to an enzyme that modifies the exposed ends of a polynucleotide chain. The polynucleotide may be double-stranded DNA (dsDNA), single-stranded DNA (ssDNA), RNA, double-stranded hybrids of DNA and RNA, and synthetic DNA (for example, containing bases other than A, C, G, and T). An end-processing enzyme may modify exposed polynucleotide chain ends by adding one or more nucleotides, removing one or more nucleotides, removing or modifying a phosphate group and/or removing or modifying a hydroxyl group. A end-processing enzyme may modify may modify ends at endonuclease cut sites or at ends generated by other chemical or mechanical means, such as shearing (for example by passing through fine-gauge needle, heating, sonicating, mini bead tumbling, and nebulizing), ionizing radiation, ultraviolet radiation, oxygen radicals, chemical hydrolysis and chemotherapy agents.

[0081] The term “DNA end-processing enzyme” refers to an enzyme that modifies the exposed ends of DNA. A DNA end-processing enzyme may modify blunt ends or staggered ends (ends with 5' or 3' overhangs). A DNA end-processing enzyme may modify single stranded or double stranded DNA. A DNA end-processing enzyme may modify ends at endonuclease cut sites or at ends generated by other chemical or mechanical means, such as

shearing (for example by passing through fine-gauge needle, heating, sonicating, mini bead tumbling, and nebulizing), ionizing radiation, ultraviolet radiation, oxygen radicals, chemical hydrolysis and chemotherapy agents. DNA end-processing enzyme may modify exposed DNA ends by adding one or more nucleotides, removing one or more nucleotides, removing or modifying a phosphate group and/or removing or modifying a hydroxyl group. Non-limiting examples of types of DNA end-processing enzymes include 5-3' exonucleases, 5-3' alkaline exonucleases, 3-5' exonucleases, 5' flap endonucleases, helicases, phosphatases, hydrolases and template-independent DNA polymerases. Examples of DNA end-processing enzymes include, but are not limited to, Trex2, Trex1, Trex1 without transmembrane domain, Apollo, Artemis, DNA2, Exo1, ExoT, ExoIII, Fen1, Fan1, MreII, Rad2, Rad9, TdT (terminal deoxynucleotidyl transferase), PNKP, RecE, RecJ, RecQ, Lambda exonuclease, Sox, Vaccinia DNA polymerase, exonuclease I, exonuclease III, exonuclease VII, NDK1, NDK5, NDK7, NDK8, WRN, T7-exonuclease Gene 6, avian myeloblastosis virus integration protein (IN), Bloom, Antarctic Phosphatase, Alkaline Phosphatase, Poly nucleotide Kinase (PNK), ApeI, Mung Bean nuclease, Hex1, TTRAP (TDP2), Sgs1, Sae2, CtIP, Pol mu, Pol lambda, MUS81, EME1, EME2, SLX1, SLX4 and UL-12. Many DNA end-processing enzymes are highly conserved throughout evolution, and thus likely to function in several different species. Further, homologues of DNA end-processing enzymes may be readily identifiable in organisms of biotechnological interest, including plants, animals, and algae. Contemplated herein are methods of modifying DNA end-processing enzymes to optimize activity or processivity.

[0082] The term "exonuclease" refers to enzymes that cleave phosphodiester bonds at the end of a polynucleotide chain via a hydrolyzing reaction that breaks phosphodiester bonds at either the 3' or 5' end. The polynucleotide may be double-stranded DNA (dsDNA), single-stranded DNA (ssDNA), RNA, double-stranded hybrids of DNA and RNA, and synthetic DNA (for example, containing bases other than A, C, G, and T). The term "5' exonuclease" refers to exonucleases that cleave the phosphodiester bond at the 5' end. The term "3' exonuclease" refers to exonucleases that cleave the phosphodiester bond at the 3' end. Exonucleases may cleave the phosphodiester bonds at the end of a polynucleotide chain at endonuclease cut sites or at ends generated by other chemical or mechanical means, such as shearing (for example by passing through fine-gauge needle, heating, sonicating, mini bead tumbling, and nebulizing), ionizing radiation, ultraviolet radiation, oxygen radicals, chemical hydrolysis and chemotherapy agents. Exonucleases may cleave the phosphodiester bonds at blunt ends or sticky ends. *E. coli* exonuclease I and exonuclease III are two commonly used 3'-exonucleases that have 3'-exonucleolytic single-strand degradation activity. Other examples of 3'-exonucleases include Nucleoside diphosphate kinases (NDKs), NDK1 (NM23-H1), NDK5, NDK7, and NDK8 (Yoon

J-H, et al., Characterization of the 3' to 5' exonuclease activity found in human nucleoside diphosphate kinase 1 (NDK1) and several of its homologues. *Biochemistry* 2005;44(48):15774-15786.), WRN (Ahn, B., et al., Regulation of WRN helicase activity in human base excision repair. *J. Biol. Chem.* 2004, 279:53465–53474) and Three prime repair exonuclease 2 (Trex2) (Mazur, D.J., Perrino, F.W., Excision of 3' termini by the Trex1 and TREX2 3'-->5' exonucleases. Characterization of the recombinant proteins. *J. Biol. Chem.* 2001, 276:17022-17029.). *E. coli* exonuclease VII and T7-exonuclease Gene 6 are two commonly used 5'-3' exonucleases that have 5% exonucleolytic single-strand degradation activity. The exonuclease can be originated from prokaryotes, such as *E. coli* exonucleases, or eukaryotes, such as yeast, worm, murine, or human exonucleases.

[0083] The term “cleavage” refers to the breakage of the covalent backbone of a polynucleotide. Cleavage can be initiated by a variety of methods including, but not limited to, enzymatic or chemical hydrolysis of a phosphodiester bond. Both single-stranded cleavage and double-stranded cleavage are possible, and double-stranded cleavage can occur as a result of two distinct single-stranded cleavage events. Double stranded DNA, RNA, or DNA/RNA hybrid cleavage can result in the production of either blunt ends or staggered ends.

[0084] The terms "target site" or "target sequence" refers to a nucleic acid sequence that defines a portion of a nucleic acid to which a binding molecule will bind, provided sufficient conditions for binding exist. For example, the target sites for several homing endonucleases are shown in Table 1.

Table 1: Examples of Homing Endonucleases and their Target Sites.

Homing Endonucleases	Target
I-SceI	TAGGGATAACAGGGTAAT (SEQ ID No. 1)
I-LtrI	AATGCTCCTATACGACGTTTAG (SEQ ID No. 2)
I-GpiI	TTTTCTGTATATGACTTAAAT (SEQ ID No. 3)
I-GzeI	GCCCCTCATAACCCGTATCAAG (SEQ ID No. 4)
I-xMpeMI	TAGATAACCATAAGTGCTAAT (SEQ ID No.5)
I-PanMI	GCTCCTCATAATCCTTATCAAG (SEQ ID No. 6)
I-CreI	TCAAAACGTCGTGAGACAGTTTGG (SEQ ID No. 7)
I-OnuI	TTTCCACTTATTCAACCTTTTA (SEQ ID No. 8)
I-HjeMI	TTGAGGAGGTTTCTCTGTTAAT (SEQ ID No. 9)
I-AniI	TGAGGAGGTTTCTCTGTAAA (SEQ ID No. 10)

[0085] The term "fusion protein" indicates that the protein includes polypeptide components derived from more than one parental protein or polypeptide. Typically, a fusion protein is expressed from a fusion gene in which a nucleotide sequence encoding a polypeptide sequence from one protein is appended in frame with, and optionally separated by a linker from, a nucleotide sequence encoding a polypeptide sequence from a different protein. The fusion gene can then be expressed by a host cell as a single protein. A fusion protein can comprise at least part of one polypeptide fused with another polypeptide. In some embodiments, a fusion protein can comprise at least a part of one polypeptide fused with at least a part of the same polypeptide. One example of a fusion protein is monomeric Trex2 (at least a part of Trex2 fused to at least a part of Trex2).

[0086] The term "endonuclease/end-processing enzyme fusion protein" or "fusion protein having endonuclease and end-processing activity" refers to an enzyme, which has an endonuclease catalytic domain and an end-processing catalytic domain and exhibits endonuclease and end-processing activity.

[0087] A "domain" of a protein is any portion of the entire protein, up to and including the complete protein, but typically comprising less than the complete protein. A domain can, but need not, fold independently of the rest of the protein chain and/or be correlated with a particular biological, biochemical, or structural function or location (e.g., an endonuclease domain, a polynucleotide binding domain, such as a DNA-binding domain, or an end-processing domain).

[0088] "Prokaryotic" cells lack a true nucleus. Examples of prokaryotic cells are bacteria (e.g., cyanobacteria, *Lactobacillus acidophilus*, Nitrogen-Fixing Bacteria, *Helicobacter pylori*, *Bifidobacterium*, *Staphylococcus aureus*, *Bacillus anthrax*, *Clostridium tetani*, *Streptococcus pyogenes*, *Staphylococcus pneumoniae*, *Klebsiella pneumoniae* and *Escherichia coli*) and archaea (e.g., Crenarchaeota, Euryarchaeota, and Korarchaeota).

[0089] "Eukaryotic" cells include, but are not limited to, algae cells, fungal cells (such as yeast), plant cells, animal cells, mammalian cells, and human cells (e.g., T-cells).

[0090] "Plant" cells include, but are not limited to, cells of monocotyledonous (monocots) or dicotyledonous (dicots) plants. Non-limiting examples of monocots include cereal plants such as maize, rice, barley, oats, wheat, sorghum, rye, sugarcane, pineapple, onion, banana, and coconut. Non-limiting examples of dicots include tobacco, tomato, sunflower, cotton, sugarbeet, potato, lettuce, melon, soybean, canola (rapeseed), and alfalfa. Plant cells may be from any part of the plant and/or from any stage of plant development.

[0091] "Algae" are predominantly aquatic organisms that carry out oxygen-evolving photosynthesis but lack specialized water-conducting and food-conducting tissues. Algae may

be unicellular or multicellular. Algae may be adapted to live in salt water, fresh water and on land. Example of algae include, but are not limited to, diatoms, chlorophyta (for example, volvox, spirogyra), euglenophyta, dinoflagellata, chrysophyta, phaephyta (for example, fucus, kelp, sargassum), and rhodophyta (for example, lemanae).

[0092] The term "subject" as used herein includes all members of the animal kingdom including non-human primates and humans.

Overview

[0093] Several embodiments described herein relate to a method of improving the rate of gene disruptions caused by imprecise repair of DNA double-strand breaks. In some embodiments, DNA end-processing enzymes are provided to enhance the rate of gene disruption. Some aspects of the present embodiments include, without limitation, enhanced rates of DNA end-processing enzyme-mediated processing of DNA ends at the site of a double-strand break.

[0094] Targeted DNA double-strand breaks introduced by rare-cleaving endonucleases can be harnessed for gene disruption applications in diverse cell types by engaging non-homologous end joining DNA repair pathways. However, endonucleases create chemically clean breaks that are often subject to precise repair, limiting the efficiency of targeted gene disruption. Several embodiments described herein relate to a method of improving the rate of targeted gene disruptions caused by imprecise repair of endonuclease-induced site-specific DNA double-strand breaks. In some embodiments, site specific endonucleases are coupled with end-processing enzymes to enhance the rate of targeted gene disruption. Coupling may be, for example, physical, spatial, and/or temporal.

[0095] Some aspects of the present embodiments include, without limitation, enhanced rates of end-processing enzyme-mediated processing of endonuclease-produced DNA ends, leading to enhanced targeted gene disruption at the genomic target site. Using this strategy, embodiments described herein show over 25 fold increased endonuclease-induced disruption rates. Certain embodiments described herein can achieve complete knockout of a target gene within a population. This technology further has the potential to dramatically increase the utility of rare-cleaving endonucleases for genetic knockout applications. Improving the mutation rate associated with endonucleases facilitates endonuclease engineering, as enzymes with different levels of activity can be utilized. In some embodiments, endo-end-processor coupling is used modify DNA ends for endonuclease-induced genome engineering. In some embodiments, expression of exonucleases capable of processive 5' end resection coupled with manipulation of the DNA repair environment can be used to enhance homologous recombination-mediated gene targeting.

[0096] Not to be bound by any particular theory, the resolution of a double-strand DNA breaks by “error-prone” non-homologous end-joining (NHEJ) can be harnessed to create targeted disruptions and genetic knockouts, as the NHEJ process can result in insertions and deletions at the site of the break. NHEJ is mediated by several sub-pathways, each of which has distinct mutational consequences. The classical NHEJ pathway (cNHEJ) requires the KU/DNA-PKcs/Lig4/XRCC4 complex, and ligates ends back together with minimal processing. As the DNA breaks created by designer endonuclease platforms (zinc-finger nucleases (ZFNs), TAL effector nucleases (TALENs), and homing endonucleases (HEs)) all leave chemically clean, compatible overhang breaks that do not require processing prior to ligation, they are excellent substrates for precise repair by the cNHEJ pathway. In the absence or failure of the classical NHEJ pathway to resolve a break, alternative NHEJ pathways (altNHEJ) can substitute; however, these pathways are considerably more mutagenic.

[0097] Not to be bound by any particular theory, modification of DNA double-strand breaks by end-processing enzymes may bias repair towards an altNHEJ pathway. Further, different subsets of end-processing enzymes may enhance disruption by different mechanisms. For example, Trex2, an exonuclease that specifically hydrolyzes the phosphodiester bonds which are exposed at 3' overhangs, biases repair at break sites toward mutagenic deletion. By contrast, terminal deoxynucleotidyl transferase (TdT), a non-templative polymerase, is expected to bias repair at break sites toward mutagenic insertions by promoting the addition of nucleotide bases to alter DNA ends prior to ligation. Accordingly, one of skill in the art may use end-processing enzymes with different activities to provide for a desired engineering outcome. Further one of skill in the art may use synergy between different end-processing enzymes to achieve maximal or unique types of knockout effects.

[0098] Several embodiments described herein couple DNA breaks created by endonucleases with end-processing enzymes is a robust way to improve the rates of targeted disruption in a variety of cell types and species, without associated toxicity to the host. This is an important advance at least because: 1) Double-strand breaks (DSBs) trigger cell cycle checkpoints to arrest division until the break has been resolved; in the case of a “persistent break” (a repetitive cycle of cleaving and precise repair), cells may arrest indefinitely, leading to apoptosis. 2) Engineering applications often utilize transient delivery of an endonuclease, providing only a short window in which enzyme concentration is sufficient to achieve breaks. 3) Persistent breaks can be a source of translocations. Coupling endonucleases to end-processing enzymes prevents the establishment of a persistent break and reduces the incidence of gross chromosomal rearrangements, thereby potentially improving the safety of endonuclease-induced targeted disruption. 4) Multiple changes in a single round of mutagenesis

may be achieved, for use for example, in multi-allelic knockouts and multiplexing, as data described herein suggests that coupling endonucleases to end-processing enzymes improves the mutagenic rate of two given endonucleases 5-fold at their respective targets, a 25-fold improvement may be realized in disrupting both targets simultaneously.

[0099] Any suitable method may be used to provide endonucleases, end-processing enzymes, and/or fusion proteins having endonuclease and end-processing activity to host cells. In some embodiments one or more polypeptides having endonuclease and/or end-processing activity may be provided directly to cells. In some embodiments, expression of endonucleases, end-processing enzymes and/or fusion proteins having endonuclease and end-processing activity in a host cell can result from delivery of one or more polynucleotides encoding one or more endonucleases, end-processing enzymes, and/or fusion proteins having endonuclease and end-processing activity to the host cell. In some embodiments, one or more polynucleotides is a DNA expression vector. In some embodiments, one or more polynucleotides is an RNA expression vector. In some embodiments, trans-splicing, polypeptide cleavage and/or polypeptide ligation can be involved in expression of one or more proteins in a cell. Methods for polynucleotide and polypeptide delivery to cells are well known in the art.

[0100] The compositions and methods described herein are useful for generating targeted disruptions of the coding sequences of genes and in some embodiments, creating gene knockouts. Targeted cleavage by the compositions and methods described herein can also be used to alter non-coding sequences (e.g., regulatory sequences such as promoters, enhancers, initiators, terminators, splice sites) to alter the levels of expression of a gene product. Such methods can be used, for example, for biological research, for biotechnology applications such as crop modification, for therapeutic purposes, functional genomics, and/or target validation studies.

[0101] Targeted mutations resulting from the methods and compositions described herein include, but are not limited to, point mutations (e.g., conversion of a single base pair to a different base pair), substitutions (e.g., conversion of a plurality of base pairs to a different sequence of identical length), insertions of one or more base pairs, deletions of one or more base pairs and any combination of the aforementioned sequence alterations.

[0102] Some embodiments relate to coupling the activity of one or more site-specific endonucleases with one or more end-processing enzymes. In some embodiments, the endonucleases and end-processing enzymes are provided as separate proteins. In some embodiments, the endonucleases and end-processing enzymes are co-expressed in a cell. If expression of the separate endonucleases and end-processing enzymes is by polynucleotide delivery, each of the endonucleases and end-processing enzymes can be encoded by separate

polynucleotides, or by a single polynucleotide. In some embodiments, the endonucleases and end-processing enzymes are encoded by a single polynucleotide and expressed by a single promoter. In some embodiments, an endonuclease and end-processing enzymes are linked by a T2A sequence which allows for two separate proteins to be produced from a single translation. In some embodiments, a different linker sequence can be used. In other embodiments a single polynucleotide encodes the endonucleases and end-processing enzymes separated by an Internal Ribosome Entry Sequence (IRES).

[0103] Several embodiments relate to coupling the activity of one or more site-specific endonucleases selected from the group consisting of: homing endonucleases (meganucleases) (including engineered homing endonucleases), zinc finger nucleases, and TAL effector nucleases with one or more end-processing enzymes. The endonucleases may comprise heterologous DNA-binding and cleavage domains (e.g., zinc finger nucleases; homing endonuclease DNA-binding domains with heterologous cleavage domains or TAL-effector domain nuclease fusions) or, alternatively, the DNA-binding domain of a naturally-occurring nuclease may be altered to bind to a selected target site (e.g., a homing endonuclease that has been engineered to bind to site different than the cognate binding site or a TAL-effector domain nuclease fusion). In some embodiments, the endonucleases and end-processing enzymes are provided as a fusion protein. In some embodiments, the endonucleases and end-processing enzymes are provided as separate proteins. In some embodiments, the endonucleases and end-processing enzymes are co-expressed in a cell.

[0104] Several embodiments relate to coupling the activity of one or more site-specific homing endonucleases selected from the group consisting of: I-Anil, I-SceI, I-CeuI, PI-PspI, PI-Sce, I-SceIV, I-CsmI, I-PanI, I-PanII, I-PanMI, I-SceII, I-PpoI, I-SceIII, I-CreI, I-LtrI, I-GpiI, I-GZeI, I-OnuI, I-HjeMI, I-TevI, I-TevII, and I-TevIII with one or more DNA end-processing enzymes selected from the group consisting of: Trex2, Trex1, Trex1 without transmembrane domain, Apollo, Artemis, DNA2, Exo1, ExoT, ExoIII, Fen1, Fan1, MreII, Rad2, Rad9, TdT (terminal deoxynucleotidyl transferase), PNKP, RecE, RecJ, RecQ, Lambda exonuclease, Sox, Vaccinia DNA polymerase, exonuclease I, exonuclease III, exonuclease VII, NDK1, NDK5, NDK7, NDK8, WRN, T7-exonuclease Gene 6, avian myeloblastosis virus integration protein (IN), Bloom, Antarctic Phosphatase, Alkaline Phosphatase, Poly nucleotide Kinase (PNK), ApeI, Mung Bean nuclease, Hex1, TTRAP (TDP2), Sgs1, Sae2, CtIP, Pol mu, Pol lambda, MUS81, EME1, EME2, SLX1, SLX4 and UL-12. In some embodiments, the homing endonucleases and DNA end-processing enzymes are provided as a fusion protein. In some embodiments, the endonucleases and DNA end-processing enzymes are provided as

separate proteins. In some embodiments, the endonucleases and DNA end-processing enzymes are co-expressed in a host cell.

[0105] Several embodiments relate to coupling the activity of one or more ZFNs with one or more DNA end-processing enzymes selected from the group consisting of: Trex2, Trex1, Trex1 without transmembrane domain, Apollo, Artemis, DNA2, Exo1, ExoT, ExoIII, Fen1, Fan1, MreII, Rad2, Rad9, TdT (terminal deoxynucleotidyl transferase), PNKP, RecE, RecJ, RecQ, Lambda exonuclease, Sox, Vaccinia DNA polymerase, exonuclease I, exonuclease III, exonuclease VII, NDK1, NDK5, NDK7, NDK8, WRN, T7-exonuclease Gene 6, avian myeloblastosis virus integration protein (IN), Bloom, Antartic Phosphatase, Alkaline Phosphatase, Poly nucleotide Kinase (PNK), ApeI, Mung Bean nuclease, Hex1, TTRAP (TDP2), Sgs1, Sae2, CtIP, Pol mu, Pol lambda, MUS81, EME1, EME2, SLX1, SLX4 and UL-12. In some embodiments, the ZFNs and DNA end-processing enzymes are provided as a fusion protein. In some embodiments, the ZFNs and DNA end-processing enzymes are provided as separate proteins. In some embodiments, the ZFNs and DNA end-processing enzymes are co-expressed in a host cell.

[0106] Several embodiments relate to coupling the activity of one or more TALENs with one or more DNA end-processing enzymes selected from the group consisting of: Trex2, Trex1, Trex1 without transmembrane domain, Apollo, Artemis, DNA2, Exo1, ExoT, ExoIII, Fen1, Fan1, MreII, Rad2, Rad9, TdT (terminal deoxynucleotidyl transferase), PNKP, RecE, RecJ, RecQ, Lambda exonuclease, Sox, Vaccinia DNA polymerase, exonuclease I, exonuclease III, exonuclease VII, NDK1, NDK5, NDK7, NDK8, WRN, T7-exonuclease Gene 6, avian myeloblastosis virus integration protein (IN), Bloom, Antartic Phosphatase, Alkaline Phosphatase, Poly nucleotide Kinase (PNK), ApeI, Mung Bean nuclease, Hex1, TTRAP (TDP2), Sgs1, Sae2, CtIP, Pol mu, Pol lambda, MUS81, EME1, EME2, SLX1, SLX4 and UL-12. In some embodiments, the TALENs and DNA end-processing enzymes are provided as a fusion protein. In some embodiments, the TALENs and DNA end-processing enzymes are provided as separate proteins. In some embodiments, the TALENs and DNA end-processing enzymes are co-expressed in a host cell.

[0107] In several embodiments, the activity of one or more site-specific homing endonucleases selected from the group consisting of: I-Anil, I-SceI, I-CeuI, PI-PspI, PI-Sce, I-SceIV, I-CsmI, I-PanI, I-PanII, I-PanMI, I-SceII, I-PpoI, I-SceIII, I-CreI, I-LtrI, I-GpiI, I-GZel, I-OnuI, I-HjeMI, I-TevI, I-TevII, and I-TevIII is coupled with the activity of one or more DNA end-processing enzymes selected from the group consisting of: Apollo, Artemis, Dna2, Exo1, Mre11, Rad2, RecE, Lambda exonuclease, Sox, exonuclease VII, T7-exonuclease Gene 6 and UL-12. In some embodiments, the homing endonucleases and DNA end-processing enzymes are

provided as a fusion protein. In some embodiments, the endonucleases and DNA end-processing enzymes are provided as separate proteins. In some embodiments, the endonucleases and DNA end-processing enzymes are co-expressed in a host cell.

[0108] In several embodiments, the activity of one or more site-specific homing endonucleases selected from the group consisting of: I-Anil, I-SceI, I-CeuI, PI-PspI, PI-Sce, I-SceIV, I-CsmI, I-PanI, I-PanII, I-PanMI, I-SceII, I-PpoI, I-SceIII, I-CreI, I-LtrI, I-GpiI, I-GZeI, I-OnuI, I-HjeMI, I-TevI, I-TevII, and I-TevIII is coupled with the activity of one or more DNA end-processing enzymes selected from the group consisting of: Sox and UL-12. In some embodiments, the homing endonucleases and DNA end-processing enzymes are provided as a fusion protein. In some embodiments, the endonucleases and DNA end-processing enzymes are provided as separate proteins. In some embodiments, the endonucleases and DNA end-processing enzymes are co-expressed in a host cell.

[0109] In several embodiments, the activity of one or more site-specific homing endonucleases selected from the group consisting of: I-Anil, I-SceI, I-CeuI, PI-PspI, PI-Sce, I-SceIV, I-CsmI, I-PanI, I-PanII, I-PanMI, I-SceII, I-PpoI, I-SceIII, I-CreI, I-LtrI, I-GpiI, I-GZeI, I-OnuI, I-HjeMI, I-TevI, I-TevII, and I-TevIII is coupled with the activity of one or more DNA end-processing enzymes selected from the group consisting of: Trex2, Vaccinia DNA polymerase, Mre11, exonuclease I, exonuclease III, NDK1, NDK5, NDK7, NDK8, and WRN. In some embodiments, the homing endonucleases and DNA end-processing enzymes are provided as a fusion protein. In some embodiments, the endonucleases and DNA end-processing enzymes are provided as separate proteins. In some embodiments, the endonucleases and DNA end-processing enzymes are co-expressed in a host cell.

[0110] In several embodiments, the activity of one or more site-specific homing endonucleases selected from the group consisting of: I-Anil, I-SceI, I-CeuI, PI-PspI, PI-Sce, I-SceIV, I-CsmI, I-PanI, I-PanII, I-PanMI, I-SceII, I-PpoI, I-SceIII, I-CreI, I-LtrI, I-GpiI, I-GZeI, I-OnuI, I-HjeMI, I-TevI, I-TevII, and I-TevIII is coupled with the activity of Fen1. In some embodiments, the homing endonucleases and DNA end-processing enzymes are provided as a fusion protein. In some embodiments, the endonucleases and DNA end-processing enzymes are provided as separate proteins. In some embodiments, the endonucleases and DNA end-processing enzymes are co-expressed in a host cell.

[0111] In several embodiments, the activity of one or more site-specific homing endonucleases selected from the group consisting of: I-Anil, I-SceI, I-CeuI, PI-PspI, PI-Sce, I-SceIV, I-CsmI, I-PanI, I-PanII, I-PanMI, I-SceII, I-PpoI, I-SceIII, I-CreI, I-LtrI, I-GpiI, I-GZeI, I-OnuI, I-HjeMI, I-TevI, I-TevII, and I-TevIII is coupled with the activity of TdT. In some embodiments, the homing endonucleases and DNA end-processing enzymes are provided as a

fusion protein. In some embodiments, the endonucleases and DNA end-processing enzymes are provided as separate proteins. In some embodiments, the endonucleases and DNA end-processing enzymes are co-expressed in a host cell.

[0112] Some embodiments relate to coupling the activity of multiple site-specific endonucleases with the activity of one or more end-processing enzymes. The site specific endonucleases may cleave target sites within the same gene or in different genes. In some embodiments, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 site-specific endonucleases may be provided to a cell along with one or more end-processing enzymes. In some embodiments, a combination of homing endonucleases, zinc finger endonucleases, and/or TAL effector endonucleases may be provided to a cell with one or more end-processing enzymes. In some embodiments, the end-processing enzyme is an exonuclease. In some embodiments, a 5' and a 3' exonuclease may be provided. If expression of the multiple endonucleases and one or more exonucleases is by polynucleotide delivery, each of the endonucleases and exonucleases can be encoded by separate polynucleotides, or by a single polynucleotide. In some embodiments, the endonucleases and exonucleases are encoded by a single polynucleotide and expressed by a single promoter. In some embodiments, the endonucleases and exonucleases are linked by a T2A sequence which allows for separate proteins to be produced from a single translation. In some embodiments, different linker sequences can be used. In other embodiments, a single polynucleotide encodes the endonucleases and exonucleases separated by IRESs.

[0113] Several embodiments relate to a heterologous fusion protein, which comprises an endonuclease domain and an end-processing domain or portions thereof. Several embodiments relate to a heterologous fusion construct, which encodes a fusion protein having endonuclease and end-processing activity. The present embodiments also relate to vectors and host cells comprising the heterologous fusion construct as well as methods for producing a fusion protein having endonuclease and end-processing activity and compositions thereof. In one embodiment, the endonuclease domain is coupled to the end-processing domain by recombinant means (e.g., the fusion protein is generated by translation of a nucleic acid in which a polynucleotide encoding all or a portion of a endonuclease is joined in-frame with a polynucleotide encoding all or a portion of a end-processing enzyme). In other embodiments, the endonuclease domain and end-processing domain of a fusion protein may be linked chemically. This chemical linkage can be carried out, for example, by using bifunctional linker molecules, such as, BS3 (Bis[sulfosuccinimidyl] suberate).

[0114] Some embodiments relate to a fusion protein comprising an endonuclease domain and exonuclease domain. In some embodiments the fusion protein comprises at least a fragment or variant of a homing endonuclease and at least a fragment or variant of an

exonuclease, for example a 3' exonuclease, which are associated with one another by genetic or chemical conjugation to one another. In several embodiments, the 3' exonuclease is a Trex2 monomer, dimer, or a variant thereof. In other embodiments, the fusion protein comprises at least a fragment or variant of a zinc finger endonuclease and at least a fragment or variant of a 5' exonuclease, which are associated with one another, by genetic fusion or chemical conjugation to one another. The endonuclease and exonuclease, once part of the fusion protein, may be referred to as a "portion", "region," "domain" or "moiety" of the endo/exo-nuclease fusion protein.

[0115] In some embodiments, an end-processing enzyme (or fragment or variant thereof) is fused directly to an endonuclease (or fragment or variant thereof). The end-processing enzyme (or fragment or variant thereof) may be fused to the amino terminus or the carboxyl terminus of the endonuclease (or fragment or variant thereof).

[0116] An endonuclease/end-processing enzyme fusion protein may optionally include a linker peptide between the endonuclease and end-processing enzyme domains to provide greater physical separation between the moieties and thus maximize the accessibility of the endonuclease portion, for instance, for binding to its target sequence. The linker peptide may consist of amino acids selected to make it more flexible or more rigid depending on the relevant function. The linker sequence may be cleavable by a protease or cleavable chemically to yield separate endonuclease and end-processing enzyme moieties. Examples of enzymatic cleavage sites in the linker include sites for cleavage by a proteolytic enzyme, such as enterokinase, Factor Xa, trypsin, collagenase, and thrombin. In some embodiments, the protease is one which is produced naturally by the host or it is exogenously introduced. Alternatively, the cleavage site in the linker may be a site capable of being cleaved upon exposure to a selected chemical, e.g., cyanogen bromide, hydroxylamine, or low pH. The optional linker sequence may serve a purpose other than the provision of a cleavage site. The linker sequence should allow effective positioning of the endonuclease moiety with respect to the end-processing enzyme moiety so that the endonuclease domain can recognize and cleave its target sequence and the end-processing domain can modify the DNA ends exposed at the cleavage site. The linker may also be a simple amino acid sequence of a sufficient length to prevent any steric hindrance between the endonuclease domain and the end-processing domain. In addition, the linker sequence may provide for post-translational modification including, but not limited to, e.g., phosphorylation sites, biotinylation sites, sulfation sites, γ -carboxylation sites, and the like.

[0117] In some embodiments the linker sequence comprises from about 4 to 30 amino acids, more preferably from about 8 to 22 amino acids. That is, the linker sequence can be any number of amino acids from about 4 to 30, such as at least or equal to 4, 5, 6, 7, 8, 9, 10,

11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 amino acids. In some embodiments, the linker sequence is flexible so as not hold the biologically active peptide in a single undesired conformation. The linker may be predominantly comprised of amino acids with small side chains, such as glycine, alanine, and serine, to provide for flexibility. In some embodiments about 80 or 90 percent or greater of the linker sequence comprises glycine, alanine, or serine residues, particularly glycine and serine residues. In several embodiments, a G4S linker peptide separates the end-processing and endonuclease domains of the fusion protein. In other embodiments, a T2A linker sequence allows for two separate proteins to be produced from a single translation. Suitable linker sequences can be readily identified empirically. Additionally, suitable size and sequences of linker sequences also can be determined by conventional computer modeling techniques.

[0118] Expression of a fusion protein in a cell can result from delivery of the fusion protein to the cell or by delivery of a polynucleotide encoding the fusion protein to a cell, wherein the polynucleotide is transcribed, and the transcript is translated, to generate the fusion protein. Trans-splicing, polypeptide cleavage and polypeptide ligation can also be involved in expression of a protein in a cell. Methods for polynucleotide and polypeptide delivery to cells are well known in the art.

[0119] A variety of DNA molecules encoding the above-described endonucleases, end-processing enzymes and fusion proteins may be constructed for providing the selected proteins or peptides to a cell. The DNA molecules encoding the endonucleases, end-processing enzyme, and fusion proteins may be modified to contain different codons to optimize expression in a selected host cell, as is known in the art.

[0120] A variety of RNA molecules encoding the above-described endonucleases, end-processing enzymes and fusion proteins may be constructed for providing the selected proteins or peptides to a cell. The RNA molecules encoding the endonucleases, end-processing enzyme, and fusion proteins may be modified to contain different codons to optimize expression in a selected host cell, as is known in the art.

[0121] Several embodiments relate to the prevention of precise cNHEJ mediated repair of endonuclease-induced double strand breaks by simultaneous expression of end-processing enzymes capable of recognizing the post-endonuclease break structure, resulting in the modification of DNA ends prior to ligation, promoting a mutagenic outcome. Some embodiments relate to the simultaneous expression exonucleases capable of recognizing the post-endonuclease break structure, resulting in the trimming of DNA ends prior to ligation, promoting a mutagenic outcome. Simultaneous expression of a site-specific endonuclease and an end-processing enzyme improves the efficiency of targeted gene disruption by up to ~70 fold,

essentially fixing a mutagenic outcome in 100% of a population of cells containing the target site in less than 72 hours.

[0122] In some embodiments, effective amounts of endonucleases and end-processing enzymes or an effective amount of a fusion protein are delivered to a cell either directly by contacting the cell with the protein(s) or by transient expression from an expression construct. In such embodiments, cell division reduces the concentration of the nucleases to sub-active levels within a few cell divisions.

[0123] Several embodiments relate to a method of conferring site specificity on a DNA end-processing enzyme by physically tethering an end-processing enzyme domain to a site specific DNA binding domain. In some embodiments, the end-processing enzyme domain is tethered to a DNA binding domain through a linker peptide. The composition and structure of the linker peptide is not especially limited and in some embodiments the linker may be chemically or enzymatically cleavable. The linker peptide may be flexible or rigid and may comprise from about 4 to 30 amino acids. In other embodiments, the end-processing enzyme domain is chemically fused to a DNA binding domain. Not wishing to be bound by a particular theory, imparting site specificity to a end-processing enzyme through tethering the end-processing enzyme to a site specific DNA binding domain decreases toxicity associated with indiscriminate end-processing activity, such as exonuclease activity, and reduces the effective amount of end-processing enzyme required for efficient modification of the exposed double stranded DNA break caused by endonuclease activity compared to untethered end-processing enzyme. In some embodiments, the end-processing enzyme is tethered to a homing endonuclease. In other embodiments, the end-processing enzyme is tethered to zinc finger endonuclease. In some embodiments, an end-processing enzyme domain is tethered to a zinc finger DNA binding domain which binds to a DNA sequence adjacent to the cleavage site of a homing endonuclease or zinc finger endonuclease.

[0124] Several embodiments relate to coupling the activity of one or more site-specific endonucleases with Trex2. Trex2 may be provided as a monomer or dimer. The Trex2 enzyme specifically hydrolyzes the phosphodiester bonds which are exposed at 3' overhangs. While the homing endonucleases generate 3' overhangs which are susceptible to Trex2 exonuclease activity, the zinc finger nucleases, which utilize the Fok1 cleavage domain, generate double strand DNA breaks with 5' overhangs. The homing endonucleases and zinc finger nucleases generate mutations at their cleavage sites at a baseline rate. Co-expression of Trex2 with homing endonucleases increased the mutation rate ~70 fold. Co-expression of Trex2 with zinc finger endonucleases was also observed to effect on the rate of mutation. See Figure 11 A and B. Accordingly, several embodiments described herein relate to improving the

mutation rate associated zinc finger endonuclease targeted cleavage events by coupling zinc finger endonuclease to exonucleases which cleave 5' overhangs. Some embodiments relate to coupling 3' exonucleases to zinc finger endonucleases wherein the nuclease domain of the zinc finger endonuclease generates 3' overhangs.

[0125] Some embodiments relate to the co-expression of a homing endonuclease and the exonuclease, Trex2, via a single promoter linked by a T2A sequence that enables separate polypeptides to be produced from a single translation event. In this way, the endonuclease and exonuclease are provided in a 1 to 1 ratio. Higher rates of modification are achieved using T2A linked expression of the homing endonuclease, I-SceI, and Trex2 than is achieved through co-transduction of separate I-SceI, and Trex2 expression constructs. In some embodiments, a fusion protein comprising one or more endonuclease domains and one or more Trex2 domains may be provided.

[0126] In another aspect, methods of co-expressing an end-processing enzyme with a zinc finger endonuclease capable of mutating the CCR-5 gene and/or inactivating CCR-5 function in a cell or cell line are provided. In some embodiments, a method for improving the inactivation of a CCR-5 gene in a human cell is provided, the method comprising administering to the cell any site specific endonuclease having a target site in a CCR5 coupled to an end-processing enzyme. In some embodiments, a method for improving the inactivation of a CCR-5 gene in a human cell is provided, the method comprising administering to the cell any site specific endonuclease having a target site in a CCR5 coupled to an exonuclease capable of cleaving the phosphodiester bonds created at the site of endonuclease cleavage. In some embodiments, a method for improving the inactivation of a CCR-5 gene in a human cell is provided, the method comprising administering to the cell any site specific endonuclease having a target site in a CCR5 and contemporaneously administering an exonuclease capable of cleaving the phosphodiester bonds created at the site of endonuclease cleavage. Examples of suitable endonucleases include engineered homing endonucleases and meganucleases, which have very long recognition sequences, some of which are likely to be present, on a statistical basis, once in a human-sized genome. Any such nuclease having a unique target site in a CCR5 gene can be used instead of, or in addition to, a zinc finger nuclease, in conjunction with an exonuclease for targeted cleavage in a CCR5 gene. Some embodiments relate to administration of a fusion protein comprising a CCR5-site-specific endonuclease and an exonuclease capable of cleaving the phosphodiester bonds created at the site of endonuclease cleavage.

Expression Vectors

[0127] Expression constructs can be readily designed using methods known in the art. Examples of nucleic acid expression vectors include, but are not limited to: recombinant

viruses, lentiviruses, adenoviruses, plasmids, bacterial artificial chromosomes, yeast artificial chromosomes, human artificial chromosomes, minicircle DNA, episomes, cDNA, RNA, and PCR products. In some embodiments, nucleic acid expression vectors encode a single peptide (e.g., an endonuclease, an end-processing enzyme, or a fusion protein having endonuclease and end-processing activity). In some embodiments, nucleic acid expression vectors encode one or more endonucleases and one or more end-processing enzymes in a single, polycistronic expression cassette. In some embodiments, one or more endonucleases and one or more end-processing enzymes are linked to each other by a 2A peptide sequence or equivalent “autocleavage” sequence. In some embodiments, a polycistronic expression cassette may incorporate one or more internal ribosomal entry site (IRES) sequences between open reading frames. In some embodiments, the nucleic acid expression vectors are DNA expression vectors. In some embodiments, the nucleic acid expression vectors are RNA expression vectors.

[0128] In some embodiments, a nucleic acid expression vector may further comprise one or more selection markers that facilitate identification or selection of host cells that have received and express the endonuclease(s), end-processing enzyme(s), and/or fusion protein(s) having endonuclease and end-processing activity along with the selection marker. Examples of selection markers include, but are not limited to, genes encoding fluorescent proteins, e.g., EGFP, DS-Red, YFP, and CFP; genes encoding proteins conferring resistance to a selection agent, e.g., Puro^R gene, Zeo^R gene, Hygro^R gene, neo^R gene, and the blasticidin resistance gene. In some cases, the selection marker comprises a fluorescent reporter and a selection marker.

[0129] In some embodiments, a DNA expression vector may comprise a promoter capable of driving expression of one or more endonuclease(s), end-processing enzyme(s), and/or fusion protein(s) having endonuclease and end-processing activity. Examples of promoters include, but are not limited to, retroviral LTR elements; constitutive promoters such as CMV, HSV1-TK, SV40, EF-1 α , β -actin; inducible promoters, such as those containing Tet-operator elements; and tissue specific promoters. Suitable bacterial and eukaryotic promoters are well known in the art and described, e.g., in Sambrook et al., *Molecular Cloning, A Laboratory Manual* (2nd ed. 2001); Kriegler, *Gene Transfer and Expression: A Laboratory Manual* (1990); and *Current Protocols in Molecular Biology* (2010). Non-limiting examples of plant promoters include promoter sequences derived from *A. thaliana* ubiquitin-3 (ubi-3).

[0130] In some embodiments, a nucleic acid encoding one or more endonucleases, end-processing enzymes, and/or fusion proteins having endonuclease and end-processing activity can be cloned into a vector for transformation into prokaryotic or eukaryotic cells. In some embodiments, nucleic acids encoding different endonucleases and end-processing enzymes are cloned into the same vector. In such cases, the nucleic acids encoding different

endonucleases and end-processing enzymes may optionally be separated by T2A or IRES sequences. Vectors can be prokaryotic vectors, e.g., plasmids, or shuttle vectors, insect vectors, or eukaryotic vectors, including plant vectors described herein. Expression of the nucleases and fusion proteins may be under the control of a constitutive promoter or an inducible promoter.

[0131] Introduction of polypeptides having endonuclease and/or end-processing activity and/or polynucleotides encoding polypeptides having endonuclease and/or end-processing activity into host cells may use any suitable methods for nucleic acid or protein delivery as described herein or as would be known to one of ordinary skill in the art. The polypeptides and polynucleotides described herein can be delivered into cultured cells in vitro, as well as in situ into tissues and whole organisms. Introduction of the polypeptides and polynucleotides of the present embodiments into a host cell can be accomplished chemically, biologically, or mechanically. This may include, but is not limited to, electroporation, sonoporation, use of a gene gun, lipotransfection, calcium phosphate transfection, use of dendrimers, microinjection, polybrene, protoplast fusion, the use of viral vectors including adenoviral, AAV, and retroviral vectors, and group II ribozymes.

Organisms

[0132] The present invention is applicable to any prokaryotic or eukaryotic organism in which it is desired to create a targeted genetic mutation. Examples of eukaryotic organisms include, but are not limited to, algae, plants, animals (e.g., mammals such as mice, rats, primates, pigs, cows, sheep, rabbits, etc.), fish, and insects. In some embodiments, isolated cells from the organism can be genetically modified as described herein. In some embodiments, the modified cells can develop into reproductively mature organisms. Eukaryotic (e.g., algae, yeast, plant, fungal, piscine, avian, and mammalian cells) cells can be used. Cells from organisms containing one or more additional genetic modifications can also be used.

[0133] Examples of mammalian cells include any cell or cell line of the organism of interest, for example oocytes, somatic cells, K562 cells, CHO (Chinese hamster ovary) cells, HEP-G2 cells, BaF-3 cells, Schneider cells, COS cells (monkey kidney cells expressing SV40 T-antigen), CV-1 cells, HuTu80 cells, NTERA2 cells, NB4 cells, HL-60 cells and HeLa cells, 293 cells and myeloma cells like SP2 or NS0. Peripheral blood mononucleocytes (PBMCs) or T-cells can also be used, as can embryonic and adult stem cells. For example, stem cells that can be used include embryonic stem cells (ES), induced pluripotent stem cells (iPSC), mesenchymal stem cells, hematopoietic stem cells, muscle stem cells, skin stem cells, and neuronal stem cells.

[0134] Examples of target plants and plant cells include, but are not limited to, monocotyledonous and dicotyledonous plants, such as crops including grain crops (e.g., wheat,

maize, rice, millet, barley), fruit crops (e.g., tomato, apple, pear, strawberry, orange), forage crops (e.g., alfalfa), root vegetable crops (e.g., carrot, potato, sugar beets, yam), leafy vegetable crops (e.g., lettuce, spinach); flowering plants (e.g., petunia, rose, chrysanthemum), conifers and pine trees (e.g., pine fir, spruce); plants used in phytoremediation (e.g., heavy metal accumulating plants); oil crops (e.g., sunflower, rape seed) and plants used for experimental purposes (e.g., Arabidopsis). Thus, the disclosed methods and compositions have use over a broad range of plants, including, but not limited to, species from the genera Asparagus, Avena, Brassica, Citrus, Citrullus, Capsicum, Cucurbita, Daucus, Erigeron, Glycine, Gossypium, Hordeum, Lactuca, Lolium, Lycopersicon, Malus, Manihot, Nicotiana, Orychophragmus, Oryza, Persea, Phaseolus, Pisum, Pyrus, Prunus, Raphanus, Secale, Solanum, Sorghum, Triticum, Vitis, Vigna, and Zea. The term plant cells include isolated plant cells as well as whole plants or portions of whole plants such as seeds, callus, leaves, roots, etc. The present disclosure also encompasses seeds of the plants described above. The present disclosure further encompasses the progeny, clones, cell lines, or cells of the plants described.

Generating Homozygously Modified Organisms

[0135] Cells in which one or more endonucleases are co-expressed with one or more end-processing enzyme(s) and/or cells in which one or more fusion proteins having endonuclease and end-processing activity are expressed are then assayed for the presence of mutations at the endonuclease cleavage site(s). Such modified cells can be identified using any suitable method known to the skilled artisan, including sequencing, PCR analysis, southern blotting, and the like. In some embodiments, an amplicon spanning the endonuclease target site is generated by PCR. The amplicon is then exposed to the endonuclease and the ability of the endonuclease to cut the amplicon is assessed. Mutation of the target site is indicated by the absence of endonuclease generated cleavage products.

[0136] Subsequently, cells containing the mutated target site(s) are cultured or otherwise treated such that they generate a whole organism with the mutated target site. For example, traditional methods of pro-nuclear injection or oocyte injection can be used to generate animals with the mutated target site. Likewise, plant cells containing the mutated target site(s) can be cultured to regenerate a whole plant which possesses the mutant genotype and thus the desired phenotype. Regeneration can also be obtained from plant callus, explants, organs, pollens, embryos, or parts thereof. Once the heterozygous organisms containing the mutated target site(s) reach reproductive maturity, they can be crossed to each other, or in some instances, spores may be grown into haploids. Of the resulting progeny from crosses, approximately 25% will be homozygous mutant/mutant at the target locus.

Pharmaceutical Compositions and Administration

[0137] Endonucleases, end-processing enzymes and fusion proteins having endonuclease and end-processing activity and expression vectors encoding endonucleases, end-processing enzymes and fusion proteins having endonuclease and end-processing activity can be administered directly to a patient for targeted cleavage of a DNA sequence and for therapeutic or prophylactic applications, for example, cancer, ischemia, diabetic retinopathy, macular degeneration, rheumatoid arthritis, psoriasis, HIV infection, sickle cell anemia, Alzheimer's disease, muscular dystrophy, neurodegenerative diseases, vascular disease, cystic fibrosis, stroke, hyper IGE syndrome, hemophilia and the like. In some embodiments, the compositions described herein (e.g., endonucleases, end-processing enzymes and fusion proteins having endonuclease and end-processing activity and expression vectors encoding endonucleases, end-processing enzymes and fusion proteins having endonuclease and end-processing activity) can be used in methods of treating, preventing, or inhibiting a disease (e.g., cancer, ischemia, diabetic retinopathy, macular degeneration, rheumatoid arthritis, psoriasis, HIV infection, sickle cell anemia, Alzheimer's disease, muscular dystrophy, neurodegenerative diseases, vascular disease, cystic fibrosis, stroke, hyper IGE syndrome, hemophilia) or ameliorating a disease condition or symptom associated with a disease, such as, cancer, ischemia, diabetic retinopathy, macular degeneration, rheumatoid arthritis, psoriasis, HIV infection, sickle cell anemia, Alzheimer's disease, muscular dystrophy, neurodegenerative diseases, vascular disease, cystic fibrosis, stroke, hyper IGE syndrome, hemophilia. In some embodiments endonucleases, end-processing enzymes and fusion proteins having endonuclease and end-processing activity and expression vectors encoding endonucleases, end-processing enzymes and fusion proteins having endonuclease and end-processing activity are administered to treat, prevent, or inhibit an autosomal dominant disease, such as achondroplasia, pseudoachondroplasia, the multiple epiphyseal dysplasias, chondrodysplasias, osteogenesis imperfecta, Marfan syndrome, polydactyly, hereditary motor sensory neuropathies I and II (Charcot-Marie-Tooth disease), myotonic dystrophy, and neurofibromatosis or ameliorate a disease condition or symptom associated with an autosomal dominant disease, such as achondroplasia, pseudoachondroplasia, the multiple epiphyseal dysplasias, chondrodysplasias, osteogenesis imperfecta, Marfan syndrome, polydactyly, hereditary motor sensory neuropathies I and II (Charcot-Marie-Tooth disease), myotonic dystrophy, and neurofibromatosis. In some embodiments endonucleases, end-processing enzymes and fusion proteins having endonuclease and end-processing activity and expression vectors encoding endonucleases, end-processing enzymes and fusion proteins having endonuclease and end-processing activity are administered to treat, prevent, or inhibit a disease caused by misregulation of genes. In some embodiments endonucleases, end-processing enzymes and fusion proteins having endonuclease

and end-processing activity and expression vectors encoding endonucleases, end-processing enzymes and fusion proteins having endonuclease and end-processing activity are administered to treat, prevent, or inhibit a cancer, such as BCL-2, Bcl-XI, and FLIP, or ameliorate a disease condition or symptom associated with a cancer, such as BCL-2, Bcl-XI, and FLIP, by, for example, increasing the mutation rate of genes with anti-apoptotic activity.

[0138] Examples of microorganisms that can be inhibited by provision of endonucleases, end-processing enzymes and fusion proteins having endonuclease and end-processing activity include pathogenic bacteria, e.g., chlamydia, rickettsial bacteria, mycobacteria, staphylococci, streptococci, pneumococci, meningococci and conococci, klebsiella, proteus, serratia, pseudomonas, legionella, diphtheria, salmonella, bacilli, cholera, tetanus, botulism, anthrax, plague, leptospirosis, and Lyme disease bacteria; infectious fungus, e.g., Aspergillus, Candida species; protozoa such as sporozoa (e.g., Plasmodia), rhizopods (e.g., Entamoeba) and flagellates (Trypanosoma, Leishmania, Trichomonas, Giardia, etc.); viral diseases, e.g., hepatitis (A, B, or C), herpes virus (e.g., VZV, HSV-1, HSV-6, HSV-II, CMV, and EBV), HIV, Ebola, adenovirus, influenza virus, flaviviruses, echovirus, rhinovirus, coxsackie virus, coronavirus, respiratory syncytial virus, mumps virus, rotavirus, measles virus, rubella virus, parvovirus, vaccinia virus, HTLV virus, dengue virus, papillomavirus, poliovirus, rabies virus, and arboviral encephalitis virus, etc.

[0139] Administration of therapeutically effective amounts is by any of the routes normally used for introducing homing endonucleases or zinc finger endonucleases into ultimate contact with the tissue to be treated. The endonucleases, end-processing enzymes and fusion proteins having endonuclease and end-processing activity are administered in any suitable manner, and in some embodiments with pharmaceutically acceptable carriers. Suitable methods of administering such proteins or polynucleotides are available and well known to those of skill in the art, and, although more than one route can be used to administer a particular composition, a particular route can often provide a more immediate and more effective reaction than another route.

[0140] Pharmaceutically acceptable carriers are determined in part by the particular composition being administered, as well as by the particular method used to administer the composition. Accordingly, there is a wide variety of suitable formulations of pharmaceutical compositions that are available (see, e.g., Remington's Pharmaceutical Sciences).

[0141] The endonucleases, end-processing enzymes and fusion proteins having endonuclease and end-processing activity or vectors encoding endonucleases, end-processing enzymes and fusion proteins having endonuclease and end-processing activity, alone or in combination with other suitable components, can be made into aerosol formulations (e.g., they

can be "nebulized") to be administered via inhalation. Aerosol formulations can be placed into pressurized acceptable propellants, such as dichlorodifluoromethane, propane, nitrogen, and the like.

[0142] Formulations suitable for parenteral administration, such as, for example, by intravenous, intramuscular, intradermal, and subcutaneous routes, include aqueous and non-aqueous, isotonic sterile injection solutions, which can contain antioxidants, buffers, bacteriostats, and solutes that render the formulation isotonic with the blood of the intended recipient, and aqueous and non-aqueous sterile suspensions that can include suspending agents, solubilizers, thickening agents, stabilizers, and preservatives. The disclosed compositions can be administered, for example, by intravenous infusion, orally, topically, intraperitoneally, intravesically or intrathecally. The formulations of compounds can be presented in unit-dose or multi-dose sealed containers, such as ampules and vials. Injection solutions and suspensions can be prepared from sterile powders, granules, and tablets of the kind previously described.

Kits

[0143] Also provided are kits for performing any of the above methods. The kits typically contain one or more endonucleases, end-processing enzymes and/or fusion proteins having endonuclease and end-processing activity or expression vectors encoding endonucleases, end-processing enzymes and/or fusion proteins having endonuclease and end-processing activity as described herein. The kits may also contain a reporter construct, such as the mCherry+ reporter construct described herein, containing a cloning site for insertion of the target site for a selected endonuclease of interest. In some embodiments, kits may contain one or more plasmids according to SEQ ID NOs: 110-145. For example, kits for screening mutagenesis produced by coupled endonuclease and end-processing activity and/or fusion proteins with activity to a particular gene are provided with one or more reporter constructs containing the desired target site(s). Similarly, kits for enriching cells for a population of cells having a endonuclease-mediated genomic modification may comprise a reporter construct comprising a target site present in the genome of the cells and one or more endonuclease specific to the target site of interest and one or more selected end-processing enzymes and/or one or more fusion proteins specific to the target site of interest.

[0144] The kits can also contain cells, buffers for transformation of cells, culture media for cells, and/or buffers for performing assays. Typically, the kits also contain a label, which includes any material such as instructions, packaging or advertising leaflet that is attached to or otherwise accompanies the other components of the kit.

[0145] While the foregoing written description enables one of ordinary skill to make and use what is considered presently to be the best mode thereof, those of ordinary skill will

understand and appreciate the existence of variations, combinations, and equivalents of the specific embodiment, method, and examples herein. The present embodiments should therefore not be limited by the above described embodiment, method, and examples, but by all embodiments and methods within the scope and spirit of the present embodiments.

[0146] The following Examples are presented for the purposes of illustration and should not be construed as limitations.

EXAMPLE 1

Co-Expression of the Homing Endonuclease, I-SceI, and Trex2

Exonuclease Increases the Rate at Which I-SceI Induces Mutations

[0147] To determine if coupling an exonuclease with a site-specific endonuclease could enhance targeted gene disruption efficiency, we assessed the effect of Trex2 on the mutagenic repair of DSBs generated by I-SceI. To ensure that Trex2 would be co-expressed with I-SceI, we developed expression vectors that drive coupled expression of both an endonuclease and an end-processing enzyme from a single promoter via a T2A “skip” peptide motif. We also included mTagBFP fluorescent protein co-expression by an internal ribosomal entry site (IRES) for tracking transfection efficiency.

[0148] To measure the rate of nuclease-induced targeted disruption, a mutNHEJ reporter construct (Traffic Light Reporter (TLR)) was constructed by placing the I-SceI target site, SEQ ID NO: 146 5'-AGTTACGCTAGGGATAACAGGGTAATATAG-3', in front of the mCherry fluorescent protein ORF in the +3 reading frame. See Figure 1A. When an endonuclease-induced DNA cleavage event results in a frameshift into the +3 reading frame, the mCherry fluorescent protein is placed in frame and correctly translated, resulting in red fluorescent cells that may be easily detected by flow cytometry. HEK cell lines harboring the TLR were generated by plating 0.1×10^6 HEK293 cells 24hrs prior to transduction in a 24 well plate. mutNHEJ (TLR) reporter cell lines were made by transducing HEK293 cells at limiting titer (~5%) with ~25ngs of an integrating lentivirus containing the reporter construct with 4ug/ml polybrene. Media was changed 24hrs after transduction.

[0149] Expression vectors comprising the homing endonuclease, I-SceI, a fluorescent protein (BFP), and optionally Trex2 with either a T2A or G4S linker peptide were constructed according to the schematics provided in Figure 1 B-H.

[0150] 0.1×10^6 HEK293 cells containing a genomically-integrated mutNHEJ (TLR) reporter cassette were plated 24hrs prior to transfection in a 24 well plate. The HEK 293 cells were transfected with expression constructs comprising the I-SceI mutant D44A alone, the I-SceI mutant D44A coupled to Trex2 via a T2A linker, I-SceI alone or I-SceI coupled to Trex2 via a T2A linker using Fugene transfection reagent according to manufacture's protocol. 72

hours following transduction of the cell line with the expression vectors, the cells were analyzed by flow cytometry on a BD LSRII or BD FACS ARII. The mCherry fluorophore was excited using a 561nm laser and acquired with a 610/20 filter. The mTagBFP fluorophore was excited on a 405nm laser with a 450/50 filter. Data was analyzed using FlowJo software (FlowJo, Ashland OR).

[0151] The plot shown in Figure 2 demonstrates that I-SceI expression induced mutagenic NHEJ events as visualized by mCherry+ expression and that the rate of mutagenic NHEJ events (mCherry+) was significantly increased following co-expression of I-SceI with the exonuclease Trex2. See Figure 2. While neither I-SceI D44A (catalytically inactive) nor I-SceI D44A coupled to Trex2 was able to induce any measurable gene disruption, I-SceI coupled to Trex2 via T2A linkage exhibited a substantial increase in mCherry positive cells compared to I-SceI alone. See Figure 2.

[0152] Following co-expression of I-SceI endonuclease and Trex2 exonuclease, genomic DNA was extracted from the HEK 293 reporter cells using Qiagen's DNA easy kit. Amplicons spanning the I-SceI target site were generated by PCR, cloned into a shuttle vector and subjected to DNA sequencing of the I-SceI target site. The sequencing demonstrated that essentially every cell in the population contains a mutated I-SceI target site, as predicted by the reporter readout. See Figure 6.

[0153] HEK 293 cells were transduced with expression constructs comprising the I-SceI mutant D44A alone, I-SceI alone or I-SceI coupled to Trex2 via a T2A linker. Following transduction of the cell line with the expression vectors, the cells were analyzed by visual inspection daily. Live cell images were taken 72 hours post transduction with the expression vectors. The cells treated in each manner appeared indistinguishable, and there is no overt toxicity associated with Trex2 co-expression. See Figure 21.

[0154] To assess the total gene disruption rate, I-SceI and I-SceI-T2A-Trex2 transfected cells were sorted based on varying BFP expression levels. HEK 293 cells containing a genomically-integrated cassette corresponding to the targeted disruption reporter illustrated in Figure 1A (TLR) were transduced with expression constructs comprising I-SceI-IRES-BFP (blue fluorescent protein) or I-SceI-T2A-Trex2-IRES-BFP. Expression of I-SceI-IRES-BFP and I-SceI-T2A-Trex2-IRES-BFP was measured in the transduced cells by a gating analysis of flow cytometry plots of BFP activity. Cells with low, low-medium, medium and high levels of BFP expression (corresponding to different levels of I-SceI endonuclease or I-SceI endonuclease/Trex2 exonuclease expression) were then assayed for induced mutagenic NHEJ events as visualized by mCherry+ expression. The data demonstrated that low levels of I-SceI alone resulted in lower mutation levels, while expression of I-SceI in combination with Trex2

result in high modification rates even at low levels of expression from the I-SceI-T2A-Trex2-IRES-BFP construct. See Figure 4.

[0155] After the I-SceI and I-SceI-T2A-Trex2 transfected cells were sorted based on varying BFP expression levels, the area flanking the I-SceI target was amplified from each of the populations by PCR. 100ng of each PCR product was digested *in vitro* with recombinant I-SceI (New England Biolabs) for 6 hours at 37°C. DNA was separated using a 1% agarose gel stained with ethidium bromide to look for a resistant band, indicative of a mutagenic event at the locus that destroyed the I-SceI target site. See Figure 5A. Percent disruption was calculated by quantifying band intensity using Image J software, and dividing the intensity of the undigested band by the total. At low endonuclease expression levels, a 25-fold increase in total gene disruption between I-SceI and I-SceI coupled to Trex2 (2.2 to 50.2% respectively) was observed, and nearly 100% of targets were disrupted in the medium and high expression gates of I-SceI T2A Trex2 (90.3, and 97.1% respectively) See Figure 5B.

[0156] These experiments indicate that while I-SceI exhibits a dose dependent increase in gene disruption, I-SceI coupled to Trex2 quickly becomes saturated. Sequence analysis of the I-SceI target site in high expressing cells confirmed that 100% of cells were modified in the I-SceI-T2A-Trex2 treated cells. See Figure 6. Comparison of the mutation spectra between I-SceI alone and I-SceI.T2A.Trex2 showed a trend towards small deletion events in the exonuclease treated cells. See Figures 6 and 7. In a kinetic analysis, while all constructs exhibited similar expression patterns, Trex2 expression coincided with the appearance of disruption events at earlier time-points. See Figure 8. In sum, coupling of endonucleases to Trex2 expression in a single open reading frame resulted in up to 25-fold enhancement in the efficiency of targeted gene disruption in cells from multiple species and in primary cell types, and is able to drive targeted knockout rates to near completion within 72hrs.

EXAMPLE 2

Trex2 Exonuclease Increases the Mutation Rate of a Variety of Homing Endonucleases

[0157] The applicability of Trex2-enhanced disruption to multiple different nuclease scaffolds was evaluated. Targeted disruption reporter cassettes (mutNHEJ reporter cassettes) with target cleavage sites for I-Ltr, I-Gpi, I-Gze, I-MpeMI, I-PanMI, I-Cre, I-OnuI, I-HjeMI, and I-Anil (See Table 1) were generated by placing the endonuclease target site of interest placed in front of the mCherry fluorescent protein ORF in the +3 reading frame. HEK293T Reporter cell lines containing genomically-integrated I-Ltr, I-Gpi, I-Gze, I-MpeMI, I-PanMI, I-Cre, I-OnuI, I-HjeMI, and I-Anil TLR reporter cassettes were then generated. Each cell line was transfected with an expression construct for its respective enzyme with or without co-transfection of an expression construct encoding Trex2, and disruption rates were measured.

[0158] The effect of Trex2 co-expression with each of I-Ltr, I-Gpi, I-Gze, I-MpeMI, I-PanMI, I-Cre, I-OnuI, I-HjeMI, and I-AniI homing endonucleases was analyzed by flow cytometry. For each of the different Homing Endonucleases tested, disruption rates increased when coupled to Trex2, demonstrating that the Trex2 exonuclease can facilitate gene disruption from breaks generated by a variety of different homing endonucleases, which leave different 3' 4bp overhangs and possess varying enzyme kinetics. See Figure 10. This data demonstrates that Trex2 expression increases the mutagenesis rates associated with targeted DNA cleavage by a variety of homing endonucleases. Further, co-expression of Trex2 with I-Gze increased mCherry+ expression significantly over the background levels observed with I-Gze expression alone. See Figure 10.

[0159] Homing Endonucleases in the panel having very low activity were rescued by coupling to Trex2. See Figure 10. This suggests that Homing Endonucleases that appear inactive may be generating breaks at an undetectable rate, and that addition of Trex2 reveals these breaks by catalyzing end processing prior to break ligation. This is consistent with the observation that Trex2 can increase disruption rates of a higher activity enzyme, such as I-SceI, even at very low expression levels.

[0160] To test the ability of Trex2 to reveal breaks caused by Homing Endonucleases having very low activity, the effect of coupling Trex2 on the gene disruption rate of the I-AniI Homing Endonucleases was analyzed by flow cytometry. WT I-AniI exhibits very little activity in cells and expression of WT I-AniI alone does not exhibit targeted disruption activity. See Figure 12. Coupling of Trex2 to WT I-AniI increases its gene disruption capacity to that of the highly active I-AniI variant, I-AniI Y2. See Figure 12. I-AniI Y2 was subjected to several rounds of directed evolution to improve its activity. Coupling of Trex2 to an inactive form of I-AniI, I-AniI E148D, shows no increase in reporter expression. This data demonstrates that Trex2 expression increases the mutagenesis rates associated with targeted DNA cleavage by sub-active homing endonucleases.

[0161] Together, these results show that Trex2 can increase disruption rates for a variety of homing endonucleases and rescue low-activity endonucleases, effectively lowering the engineering bar for enzymes designed to produce gene disruption at novel target sites.

EXAMPLE 3

Co-Expression of Trex2 Exonuclease Affects the Mutation Rate Associated With FokI Zinc Finger Nuclease Mediated Breaks

[0162] A reporter cell line was generated that harbors a 5' ACC ATC TTC ttcaag GAC GAC GGC 3' (SEQ ID NO. 147) target site for a corresponding zinc finger nuclease containing a FokI nuclease domain. Expression vectors encoding the zinc finger nuclease were

transduced into reporter cell lines harboring the TLR-FokI reporter cassette with and without Trex2. Co-expression of Trex2 with the zinc finger nuclease results in an increased mutation rate. See Figure 11B.

EXAMPLE 4

The Chimeric I-SceI-G4s-Trex2 Endo/Exo-Nuclease Fusion Protein Improves the Rate of Targeted Disruption

[0163] Expression vectors comprising HA-I-SceI-BFP, (HA-I-SceI)-T2A-Trex2-BFP or (HA-I-SceI)-G4S-Trex2-BFP were constructed as described in Example 1. The I-SceI gene used to construct the expression vectors further encoded an N-terminal HA epitope tag. The (HA-I-SceI)-T2A-(HA-Trex2-BFP) expression vector expresses HA-I-SceI and Trex2 in a 1 to 1 ratio from a single promoter, but the T2A linker sequence allows for two separate proteins to be produced from a single translation. The (HA-I-SceI)-G4S-(HA-Trex2)-BFP expression vector produces an endo/exo-nuclease fusion protein where HA-I-SceI and Trex2 proteins are coupled together by a G4S linker peptide. The HA-I-SceI-BFP, (HA-I-SceI)-T2A-Trex2-BFP and (HA-I-SceI)-G4S-Trex2-BFP expression vectors were transduced into HEK293 cells containing a genomically-integrated cassette corresponding to the targeted disruption reporter illustrated in Figure 1A.

[0164] Following transduction of the cell line with the expression vectors, the cells were analyzed for mCherry+ expression by flow cytometry. The plot shown in Figure 3A demonstrated that I-SceI-G4S-Trex2 endo/exo fusion proteins are active and increase targeted disruption rates over provision of I-SceI alone. See Figure 3.

[0165] However, Sce-G4S-Trex2, despite stable fusion protein expression, was inferior at inducing gene disruption compared to Sce-T2A-Trex2, possibly due to steric hindrance. See Figure 3A.

[0166] An anti-HA western blot was performed to assess the stability of the HA-I-SceI, HA-I-SceI-T2A and (HA-I-SceI)-G4S-Trex2 proteins in the expressing cells. As shown in Figure 3B and 3C, the chimeric (HA-I-SceI)-G4S-Trex2 endo-exo fusion protein was expressed at the same levels as I-SceI alone, or I-SceI containing a residual T2A tag peptide.

EXAMPLE 5

Co-Expression of I-SceI and Trex2 Exonuclease Increases the Rate of I-SceI-Induced Mutations in Primary Cells

[0167] To determine if Trex2 would increase gene disruption rates in primary cells, primary murine embryonic fibroblasts (MEFs) were isolated from a mouse with an I-SceI site “knocked into” the *Interleukin-2 receptor subunit gamma (IL2RG)* locus (“Sce-SCID” mouse, unpublished data, G.C., D.J.R., A.M.S). MEFs were isolated from Sce-SCID embryos at 12-14

days gestation. Briefly, individual embryos were removed from the uterus and washed with PBS. The head and red tissue were removed from the embryo, and the remaining tissue was minced. The tissue was incubated with trypsin-EDTA for 10 minutes at 37°C, followed by centrifugation at 10,000xG for 5 minutes. The pellet was re-suspended in MEF media and plated at 37°C. MEF cells were cultured in glutamine-free Dulbecco's modified Eagle's medium supplemented with 2mM L-glutamine, 10% Fetal Bovine Serum (FBS) and 1% penicillin/streptomycin.

[0168] 1.0×10^5 Sce-SCID MEF cells were seeded in a 24-well plate 24 hours prior to transduction with I-SceI or I-SceI.T2A.Trex2 expressing recombinant lentiviral vectors (LV). 0.5µg DNA was used for each expression vector, and transfected using Fugene6 or XtremeGene9 (Roche) according to the manufacture's protocol. Cells were passaged 24 hours later and analyzed 72 hours post transduction. Total gene disruption at the I-SceI target site was assayed using the digestion assay described in Example 1. A 6-fold increase in disruption at the common gamma chain locus was observed with I-SceI coupled to Trex2 (I-SceI = 15.8, I-SceI.T2A.Trex2 = 88.7). See Figures 9A and 9B. Additionally, since *IL2RG* is only expressed in a subset of differentiated hematopoietic cells, these experiments demonstrate Trex2 can facilitate high frequency disruption at unexpressed loci.

EXAMPLE 6

Effect of Exonuclease Over-expression on Repair of Endogenous DNA Damage

[0169] To determine if exonuclease over-expression alters the cells ability to repair other types of endogenous DNA damage, Trex2 expressing cells are treated with model DNA damage inducing agents. 1.0×10^6 Sce-SCID MEFs were seeded in a 10cm dish 24 hours before transduction. 500µL of 10x LV (pCVL.SFFV.sceD44A.IRES.BFP or T2A.TREX2.IRES.BFP) was added to the culture with 4µg/mL polybrene. 24 hours post-transduction, cells were passaged to 15cm plates. 72 hours post-transduction, 1.0×10^5 Sce-SCID MEFs were seeded in a 12-well plate with 1mL media and treated as indicated with DNA damage inducing agents: Mitomycin C (Sigma Aldrich, St. Louis), Camptothecin (Sigma Aldrich, St. Louis), or ionizing radiation. 48 hours after exposure, cells were incubated in 0.5µg/mL PI as above and analyzed by flow cytometry. For CD34+ cells, 72 hours post-transduction with Trex2 expressing LV, 2.0×10^5 CD34+ HSCs were seeded in a 96-well plate in 200µL of media, DNA damaging agents were added to the media, and plates analyzed as above. Over-expression of Trex2 had no adverse effect on cell cycle or sensitivity to model DNA damaging agents, suggesting cells maintain high fidelity DNA repair at lesions occurring independently of those created by the endonuclease. See Figures 13-16.

EXAMPLE 7

Co-Expression of I-SceI and End-Processing Enzymes Increases the Rate of I-SceI-Induced Mutations

[0170] To determine if the results of coupling homing endonucleases with Trex2 could be extended to other DNA modifying enzymes, a library of 13 candidate enzymes possessing an array of biochemical end-processing activities derived from mammalian, bacterial or viral species was generated. See Table 2. The library of DNA end-processing enzymes was cloned into the pExodus vector with genes synthesized by Genscript (Piscataway, NJ) as cDNA codon-optimized for human expression. See SEQ ID NOs. 110-145.

[0171] The library of DNA end-processing enzymes was screened by co-expressing each enzyme with either the homing endonuclease, I-SceI, or the Zinc Finger Nuclease, VF2468, in the respective HEK293T TLR cells. See Figures 17-19. Five of DNA end-processing enzymes (Artemis, Tdt, Apollo, Rad2, and Exo1) robustly increased the gene disruption efficiency of I-SceI. See Figure 17. Additionally, the gene disruption activity of these five enzymes was analyzed at three levels of I-SceI expression (quantified by the mean fluorescence intensity, MFI, of the BFP fluorophore). Coexpression of these enzymes with I-SceI increased I-SceI's mutagenic efficiency, even at low levels of endonuclease expression. See Figure 18. In contrast, although several of the DNA end-processing enzymes possess 5' exonuclease activity, a significant effect of any enzyme on increasing the gene disruption efficiency of the VF2468 ZFN was not observed. See Figure 19A.

[0172] In addition, the library of DNA end-processing enzymes was screened by co-expressing each enzyme with TALEN. See Figure 19B.

Table 2: Library of DNA End-Processing Enzymes.

Enzyme	Gene name	Activity	Species of origin	NLS added	Reference
Apollo	SNM1B	5-3' exonuclease	Human	No	Lenain, C. et al., The Apollo 5' exonuclease functions together with TRF2 to protect telomeres from DNA repair. <i>Curr. Biol.</i> 16, 1303-1310 (2006).
Artemis	Artemis	5-3' exonuclease	Human	No	Kurosawa, A., and Adachi, N. Functions and regulation of Artemis: a goddess in the maintenance of genome integrity.

Enzyme	Gene name	Activity	Species of origin	NLS added	Reference
					<i>J Radiat. Res. (Tokyo)</i> 51, 503-509 (2010).
Dna2	DNA2	5-3' exonuclease, helicase	Human	No	Nimonkar, A.V., et al. BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. <i>Genes Dev</i> 25, 350-362 (2011).
Exo1	EXO1	5-3' exonuclease	Human	No	Nimonkar, A.V. et al. BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. <i>Genes Dev</i> 25, 350-362 (2011). Orans, J., et al. Structures of human exonuclease 1 DNA complexes suggest a unified mechanism for nuclease family. <i>Cell</i> 145, 212-223 (2011).
Fen1	FEN1	5' flap endonuclease	Human	No	Jagannathan, I., Pepenella, S. Hayes, J.J. Activity of FEN1 endonuclease on nucleosome substrates is dependent upon DNA sequence but not flap orientation. <i>J. Biol. Chem.</i> 286, 17521-17529 (2011). Tsutakawa, S.E., et al., Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily. <i>Cell</i> 145, 198-211 (2011).
Mre11	MRE11	5-3' and 3-5'	Human	No	Garcia, V., Phelps,

Enzyme	Gene name	Activity	Species of origin	NLS added	Reference
		exonuclease			S.E., Gray, S., and Neale, M.J. Bidirectional resection of DNA double-strand breaks by Mre11 and Exo1. <i>Nature</i> 479, 241-244 (2011).
Rad2	n/a (catalytic domain of Exo1)	5-3' exonuclease (Exo1 catalytic domain)	Human	No	Lee, B.I., and Wilson, D.M., 3rd The RAD2 domain of human exonuclease 1 exhibits 5' to 3' exonuclease and flap structure-specific endonuclease activities. <i>J Bio.l Chem.</i> 274, 37763-37769 (1999).
TdT (terminal deoxynucleotidyl transferase)	TdT	Single-stranded Template independent DNA polymerase	Human	No	Mahajan, K.N., et al., Association of terminal deoxynucleotidyl transferase with Ku. <i>Proc. Natl.Acad. Sci. USA</i> 96, 13926-13931 (1999).
RecE	RecE	5-3' exonuclease	<i>E. coli</i>	Yes	Zhang, J., Xing, X., Herr, A.B., and Bell, C.E. Crystal structure of <i>E. coli</i> RecE protein reveals a toroidal tetramer for processing double-stranded DNA breaks. <i>Structure</i> 17, 690-702 (2009).
Lambda exonuclease	λ exonuclease	5-3' exonuclease	Bacteriophage λ	Yes	Zhang, J., McCabe, K.A., and Bell, C.E. Crystal structures of lambda exonuclease in complex with DNA suggest an electrostatic ratchet mechanism for processivity. <i>Proc. Natl. Acad. Sci. USA</i> 108, 11872-11877 (2011).

Enzyme	Gene name	Activity	Species of origin	NLS added	Reference
Sox (T24I mutation)	SOX	5-3' alkaline exonuclease	Kaposi's sarcoma associated herpes virus	Yes	<p>Glaunsinger, B., Chavez, L., and Ganem, D., The exonuclease and host shutoff functions of the SOX protein of Kaposi's sarcoma-associated herpesvirus are genetically separable. <i>J Virol.</i> 79, 7396-7401 (2005).</p> <p>Dahlroth, S.L., et al., Crystal structure of the shutoff and exonuclease protein from the oncogenic Kaposi's sarcoma-associated herpes virus. <i>FEBS J</i> 276, 6636-6645 (2009).</p>
Vaccinia DNA polymerase	E9L	3-5' exonuclease	Vaccinia poxvirus	Yes	<p>Gammon, D.B., and Evans, D.H., The 3'-to-5' exonuclease activity of vaccinia virus DNA polymerase is essential and plays a role in promoting virus genetic recombination. <i>J. Virol.</i> 83, 4236-4250 (2009).</p>
UL-12	UL12	5-3' alkaline exonuclease	Herpes simplex virus (HSV)-1	Yes	<p>Reuven, N.B., et al. The herpes simplex virus type 1 alkaline nuclease and single-stranded DNA binding protein mediate strand exchange in vitro. <i>J. Virol.</i> 77, 7425-7433 (2003).</p> <p>Balasubramanian, N., et al. Physical interaction between the herpes simplex virus type 1 exonuclease, UL12, and the DNA double-strand break-sensing MRN</p>

Enzyme	Gene name	Activity	Species of origin	NLS added	Reference
					complex. <i>J. Virol.</i> 84, 12504-12514 (2010).

EXAMPLE 8

Exonuclease Screen

[0173] An expression library containing both 3' and 5' specific exonucleases is screened by expressing the exonucleases in cells containing a targeted disruption reporter harboring a homing endonuclease target site, for example an I-SceI target site. The exonucleases are co-expressed in the reporter cells with a homing endonuclease, for example I-Sce-I, which generates 3' overhangs upon cleaving its target site. Exonucleases which increase the rate of disruption, as visualized by mCherry+ expression, of the homing endonuclease target site over expression of the homing endonuclease alone are then identified.

[0174] An expression library containing both 3' and 5' specific exonucleases is additionally screened by expressing the exonucleases in cells containing a targeted disruption reporter harboring a zinc finger endonuclease target site. The exonucleases are co-expressed in the reporter cells with a zinc finger endonuclease, which generates 5' overhangs upon cleaving its target site with Fok1. Exonucleases which increase the rate of disruption, as visualized by mCherry+ expression, of the zinc finger endonuclease target site over expression of the zinc finger endonuclease alone are identified.

EXAMPLE 9

Trex-Multiplex

[0175] Increasing disruption rates for individual nucleases by coupling endonuclease activity with exonuclease activity, enables multiple simultaneous changes to a genome (multiplexing).

[0176] Three homing endonuclease are designed to knock out three different genes (x, y, and z). In the absence of exonuclease co-expression, the efficiency of producing a disruptive mutation, knockout, for each gene individually is 10%, which means that the chance of successfully producing all three disruptive mutations in a single cell with a single round of endonuclease expression is .1%. An exonuclease, for example Trex2, is co-expressed with the three homing endonucleases to increase the rate of mutagenesis induced by the homing endonucleases. A 5-fold increase in the mutagenesis rate, to 50% for each individual gene, improves the chance of disrupting all three in a single cell, in a single round to 12.5%, a 125-fold difference.

EXAMPLE 10

Reduction of Chromosomal Abnormalities During Endonuclease Mediated Targeted Disruption

[0177] Endonucleases, such as homing endonucleases, zinc finger nucleases, and TAL effector nucleases, induce indiscriminate chromosomal abnormalities, such as translocations. To test the ability of co-expression of an exonuclease that facilitates disruption

of an endonuclease target site to decrease the incidence of indiscriminate chromosomal abnormalities, an endonuclease, or a series of endonucleases are expressed in the presence and absence of Trex2. Karyotyping analysis or GCH array analysis is performed to determine if the incidence of genomic abnormalities induced by the endonucleases is reduced.

EXAMPLE 11

Imparting Site-Specificity to Exonucleases

[0178] An exonuclease of interest, for example Trex2, is directly fused or coupled through a linker peptide to an endonuclease or to a DNA binding domain which specifically binds to a target site adjacent to the site where exonuclease activity is desired.

EXAMPLE 12

Method of Treating, Preventing, or Inhibiting HIV infection in a Human Patient

[0179] Hematopoietic stem cells are isolated from bone marrow obtained from a human subject. The isolated stem cells are contacted with an effective amount of a zinc finger nuclease (ZFN) having target sites in the human CCR-5 gene and contemporaneously contacted with a 5' exonuclease. The contacted cells are allowed to recover in media for 72 hrs and then screened for targeted disruption of the CCR-5 gene. Cells containing a targeted disruption in CCR-5 are then propagated under appropriate conditions. The subject is given a daily intravenous (i.v.) injection of about 20 million cells containing the targeted disruption in the CCR-5 gene. This dosage can be adjusted based on the results received and the judgment of the attending physician. The protocol is preferably continued for at least about 1 or 2 weeks, preferably at least about 1 or 2 months, and may be continued on a chronic basis.

EXAMPLE 13

Method of Treating, Preventing, or Inhibiting HIV Infection in a Human Patient

[0180] Hematopoietic stem cells are isolated from bone marrow obtained from a human subject. The isolated stem cells are contacted with an effective amount of a homing endonuclease engineered to cleave a target site in the human CCR-5 gene and contemporaneously contacted with Trex2 exonuclease. The contacted cells are allowed to recover in media for 72 hrs and then screened for targeted disruption of the CCR-5 gene. Cells containing a targeted disruption in CCR-5 are then propagated under appropriate conditions. The subject is given a daily intravenous (i.v.) injection of about 20 million cells containing the targeted disruption in the CCR-5 gene. This dosage can be adjusted based on the results of the treatment and the judgment of the attending physician. The protocol is preferably continued for at least about 1 or 2 weeks, preferably at least about 1 or 2 months, and may be continued on a chronic basis.

EXAMPLE 14

Method of Treating, Preventing, or Inhibiting HIV Infection in a Human Patient

[0181] Hematopoietic stem cells are isolated from bone marrow obtained from a human subject. The isolated stem cells are contacted with an effective amount of a fusion protein comprising an endonuclease domain linked to an exonuclease domain wherein the endonuclease domain comprises a homing endonuclease engineered to cleave a target site in the human CCR-5 gene or fragment thereof and wherein the exonuclease domain comprises Trex2 exonuclease or a fragment thereof. The contacted cells are allowed to recover in media for 72 hrs and then screened for targeted disruption of the CCR-5 gene. Cells containing a targeted disruption in CCR-5 are then propagated under appropriate conditions. The subject is given a daily intravenous (i.v.) injection of about 20 million cells containing the targeted disruption in the CCR-5 gene. This dosage can be adjusted based on the results of the treatment and the judgment of the attending physician. The protocol is preferably continued for at least about 1 or 2 weeks, preferably at least about 1 or 2 months, and may be continued on a chronic basis.

EXAMPLE 15

End-Modifying Enzyme Screen

[0182] An expression library containing end-modifying enzymes is screened by expressing the end-modifying enzymes in cells containing a targeted disruption reporter harboring a homing endonuclease target site, for example an I-SceI target site. The end-modifying enzymes are co-expressed in the reporter cells with a homing endonuclease, for example I-Sce-I, which generates 3' overhangs upon cleaving its target site. End-modifying enzymes which increase the rate of disruption, as visualized by mCherry+ expression, of the homing endonuclease target site over expression of the homing endonuclease alone are then identified.

[0183] An expression library containing end-modifying enzymes is additionally screened by expressing the exonucleases in cells containing a targeted disruption reporter harboring a zinc finger endonuclease target site. The end-modifying enzymes are co-expressed in the reporter cells with a zinc finger endonuclease, which generates 5' overhangs upon cleaving its target site with Fok1. End-modifying enzymes which increase the rate of disruption, as visualized by mCherry+ expression, of the zinc finger endonuclease target site over expression of the zinc finger endonuclease alone are identified.

EXAMPLE 16

Method of Treating, Preventing, or Inhibiting Cancer in a Human Patient

[0181] A patient having cancer is identified. The isolated an effective amount of an endonuclease targeting a site within the regulatory or coding sequence of an anti-apoptotic gene

is administered in combination with an end processing enzyme. The patient is monitored for increased apoptosis and or decreased malignant cell proliferation. In some embodiments, tumor growth is monitored. The protocol may be administered on a periodic or chronic basis.

WHAT IS CLAIMED IS:

1. A method of increasing the mutagenic activity of an endonuclease comprising: providing the endonuclease concurrently with an end-processing enzyme.
2. The method of Claim 2 wherein the endonuclease and the end-processing enzyme are encoded by a single polynucleotide.
3. The method of Claims 1 or 2 wherein the endonuclease and the end-processing enzyme are linked by a T2A linker sequence.
4. The method of Claims 1 or 2 wherein the endonuclease and exonuclease are provided as a fusion protein.
5. The method of any one of Claims 1-4 wherein the end-processing enzyme exhibits 5-3' exonuclease, 5-3' alkaline exonuclease, 3-5' exonuclease, 5' flap endonuclease, helicase or template-independent DNA polymerases activity.
6. A polypeptide comprising an endonuclease domain, and end-processing domain.
7. The polypeptide of Claim 6, further comprising a linker domain.
8. The polypeptide of Claims 6 or 7 wherein the endonuclease domain comprises an engineered homing endonuclease or a biologically active fragment thereof.
9. The polypeptide of Claims 6, 7, or 8 wherein the end-processing domain exhibits 5-3' exonuclease, 5-3' alkaline exonuclease, 3-5' exonuclease, 5' flap endonuclease, helicase, or template-independent DNA polymerases activity.
10. The polypeptide of any one of Claims 6-9 wherein the end-processing domain comprises Trex2 or a biologically active fragment thereof.
11. The polypeptide of any one of Claims 6-10 wherein the polypeptide is capable of cleaving CCR-5.
12. A method of treating a disease in a subject comprising: contacting the cells of the subject with a polypeptide of any one of Claims 6-11,
wherein the polypeptide binds to a site within the cells involved in the proliferation of the disease.
13. A method of increasing site specific mutagenesis of a nucleotide sequence comprising: coupling the activity of one or more endonucleases that cleave the nucleotide sequence and the activity of one or more end-processing enzymes.
14. The method of Claim 13 wherein the endonuclease is engineered to bind the nucleotide sequence.
15. The method of Claim 13 or 14 wherein the end-processing enzyme exhibits 5-3' exonuclease, 5-3' alkaline exonuclease, 3-5' exonuclease, 5' flap endonuclease, helicase, or template-independent DNA polymerases activity.

16. The method of any one of Claims 13-15 wherein the nucleotide sequence encodes CCR-5 or a fragment thereof.

17. The method of any one of Claims 13-16 wherein the end-processing enzyme is Trex2 or a biologically active fragment thereof.

18. The method of any one of Claims 13-16 wherein a fusion protein having endonuclease and end-processing activity is provided.

19. The method of Claim 18 wherein the fusion protein comprises a linker domain.

20. The method of any one of Claims 13-16 wherein one or more proteins having endonuclease activity and one or more proteins having end-processing activity are provided.

21. The method of Claim 20 wherein the one or more proteins having endonuclease activity and the one or more proteins having end-processing activity are encoded by one or more expression vectors.

22. The method of Claim 16 wherein the one or more proteins having endonuclease activity and the one or more proteins having end-processing activity are encoded by a single expression vector comprising a T2A sequence.

23. The method of any one of Claims 18-22 wherein at least one protein comprises a zinc finger domain.

24. A method of treating a disease in a subject comprising: contacting one or more cells of the subject with one or more polypeptides having endonuclease activity and one or more polypeptides having end-processing activity,

wherein the one or more polypeptides having endonuclease activity bind to a portion of a nucleotide sequence within the cells encoding a gene involved in the disease or a regulatory sequence thereof.

25. The method of Claim 24 wherein the disease is HIV.

26. The method of Claim 24 wherein the disease is hyper IGE syndrome.

27. The method of any one of Claims 24-26 wherein one or more polypeptides having endonuclease activity is an engineered homing endonuclease or a biologically active fragment thereof.

28. The method of any one of Claims 20-27 wherein one or more polypeptides having end-processing activity is an exonuclease or a biologically active fragment thereof.

29. The method of any one of Claims 20-28 wherein one or more polypeptides having end-processing activity is Trex2 exonuclease or a biologically active fragment thereof.

30. The method of any one of Claims 20-29 wherein one or more polypeptides comprise a zinc finger domain.

Fig. 1A



Fig. 1B

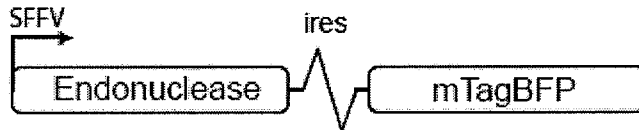


Fig. 1C

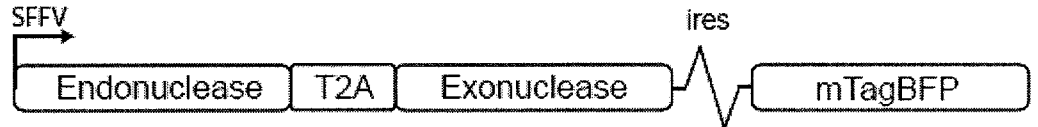


Fig. 1D

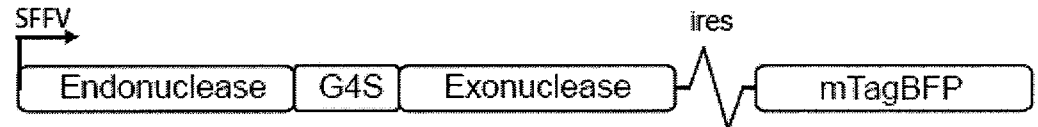


Fig. 1E

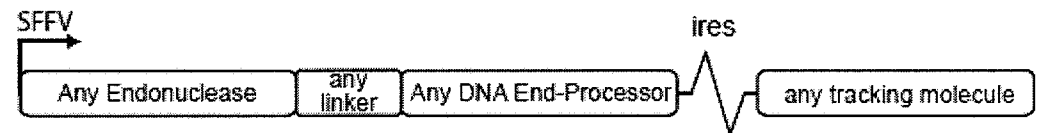


Fig. 1F

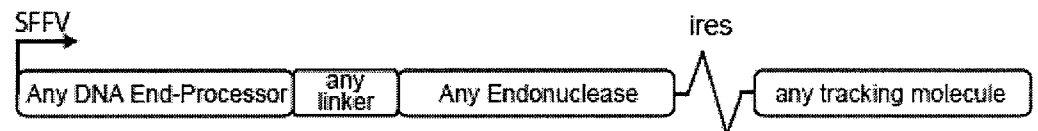


Fig. 1G



Fig. 1H

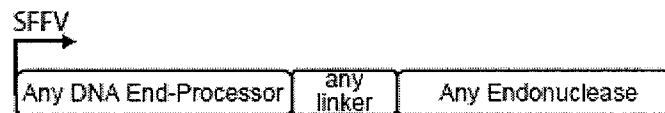


Fig. 2A

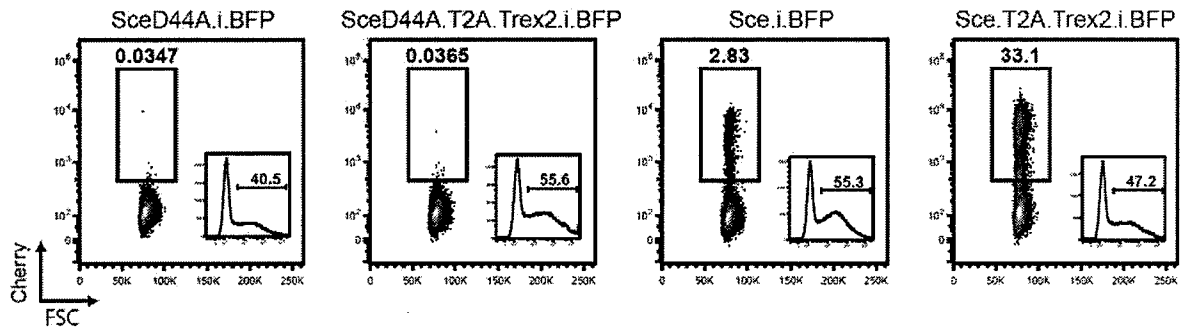


Fig. 2B

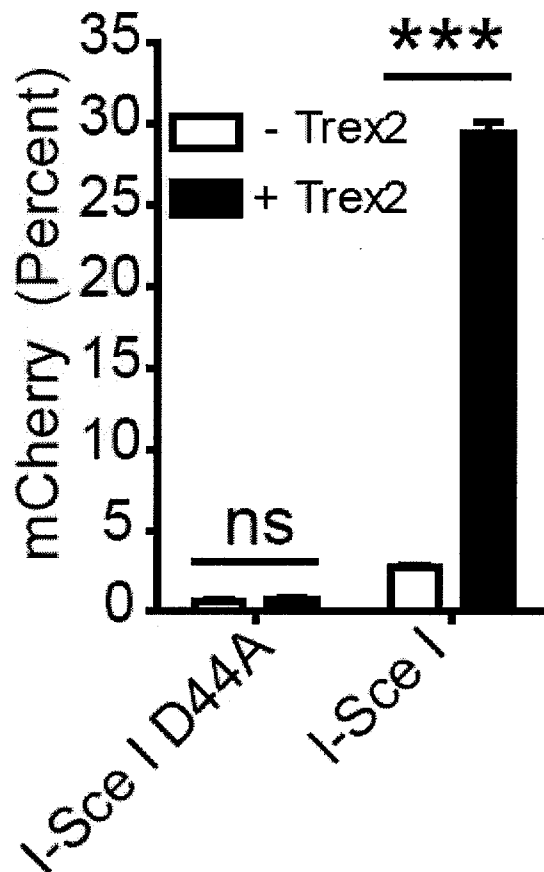


Fig. 3A

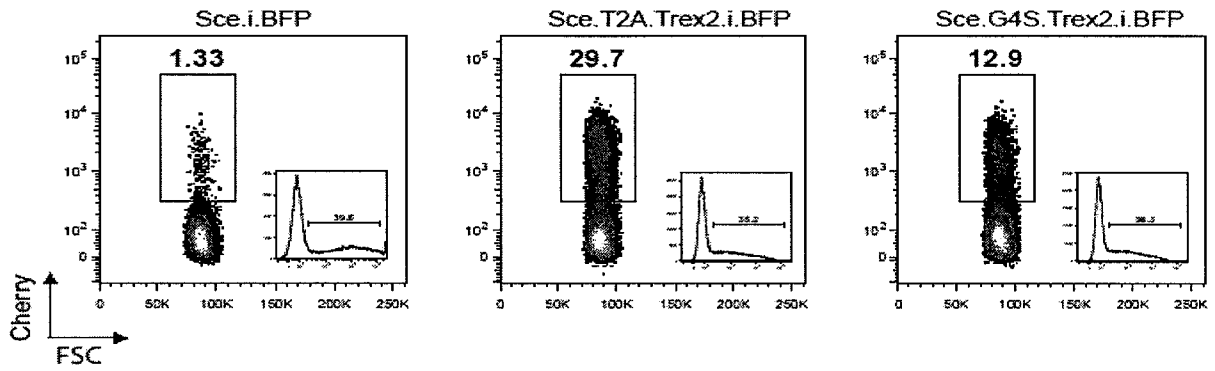


Fig. 3B

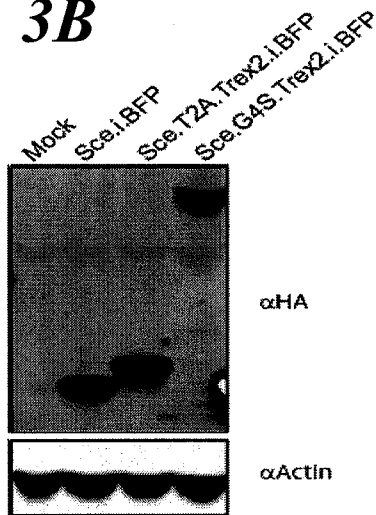


Fig. 3C

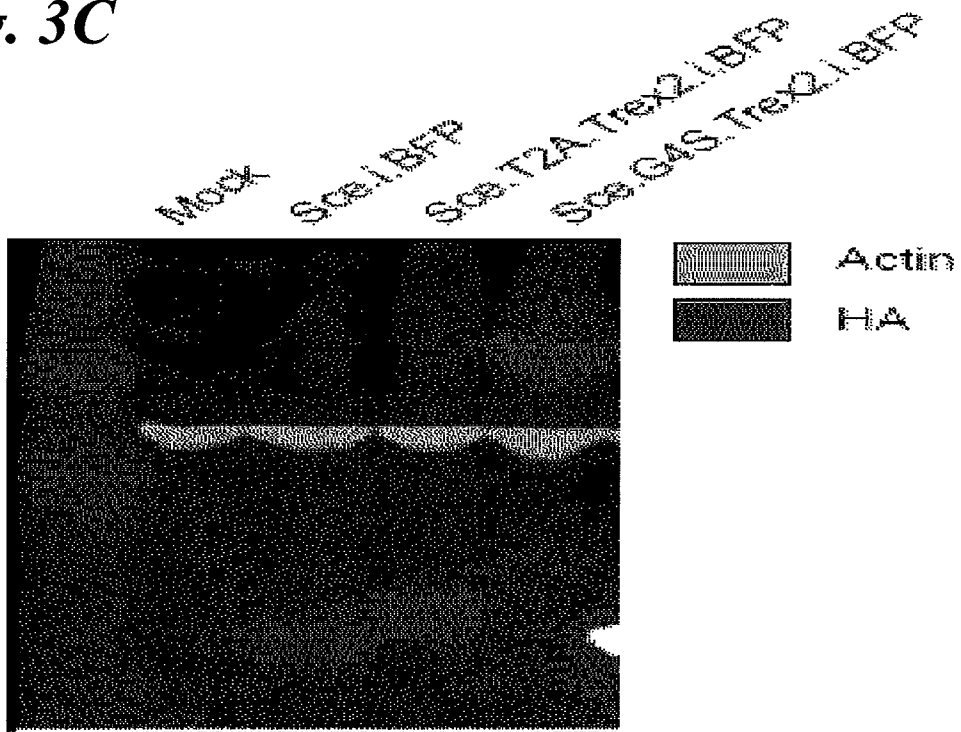


Fig. 4A

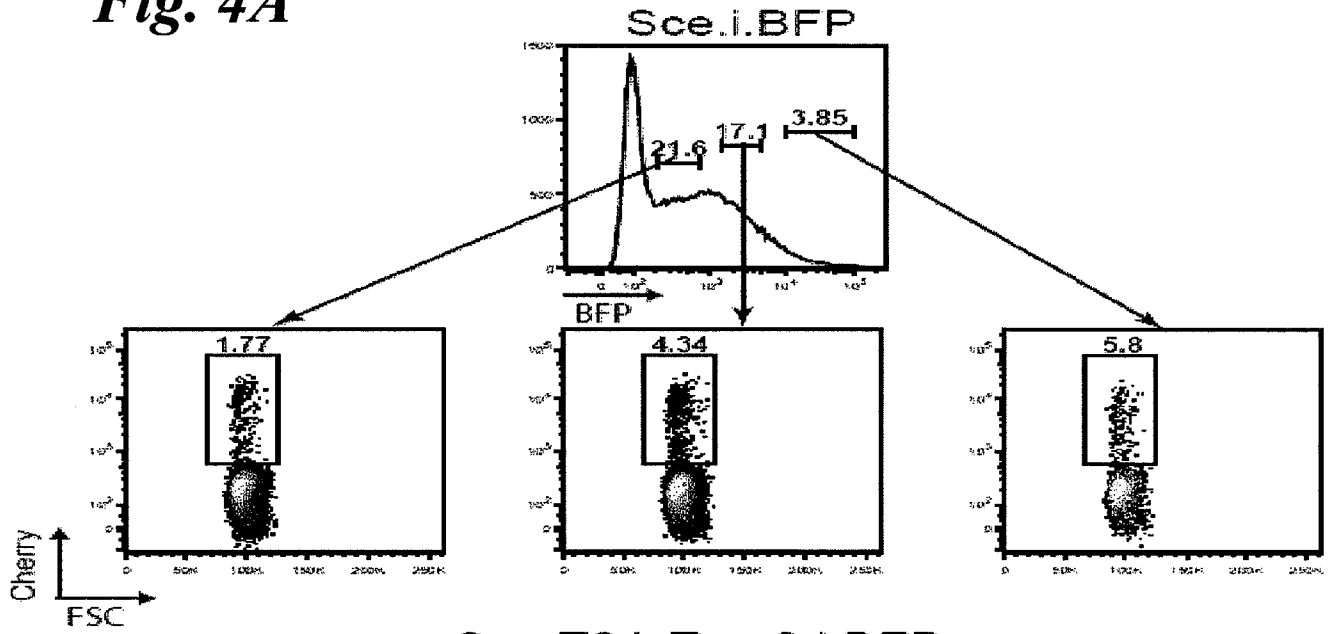


Fig. 4B

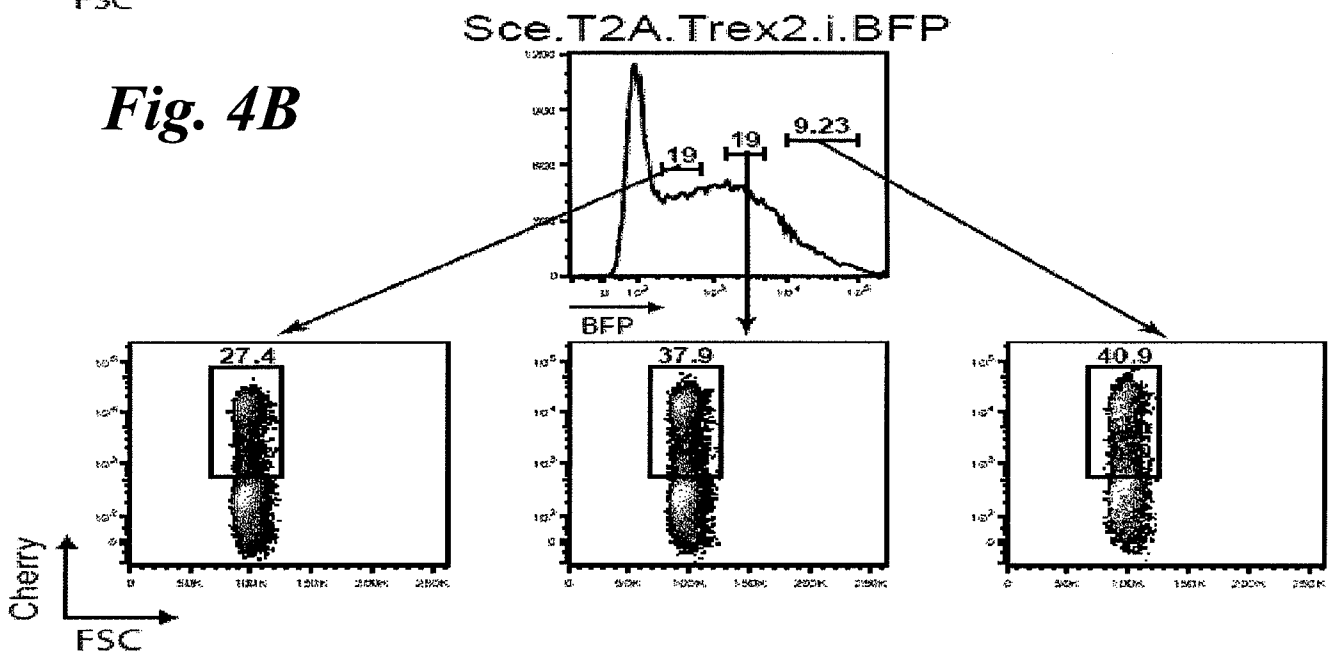


Fig. 5A

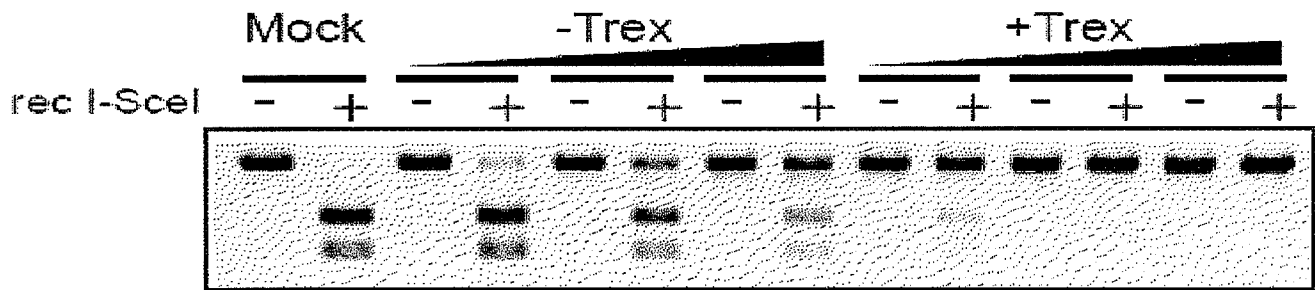


Fig. 5B

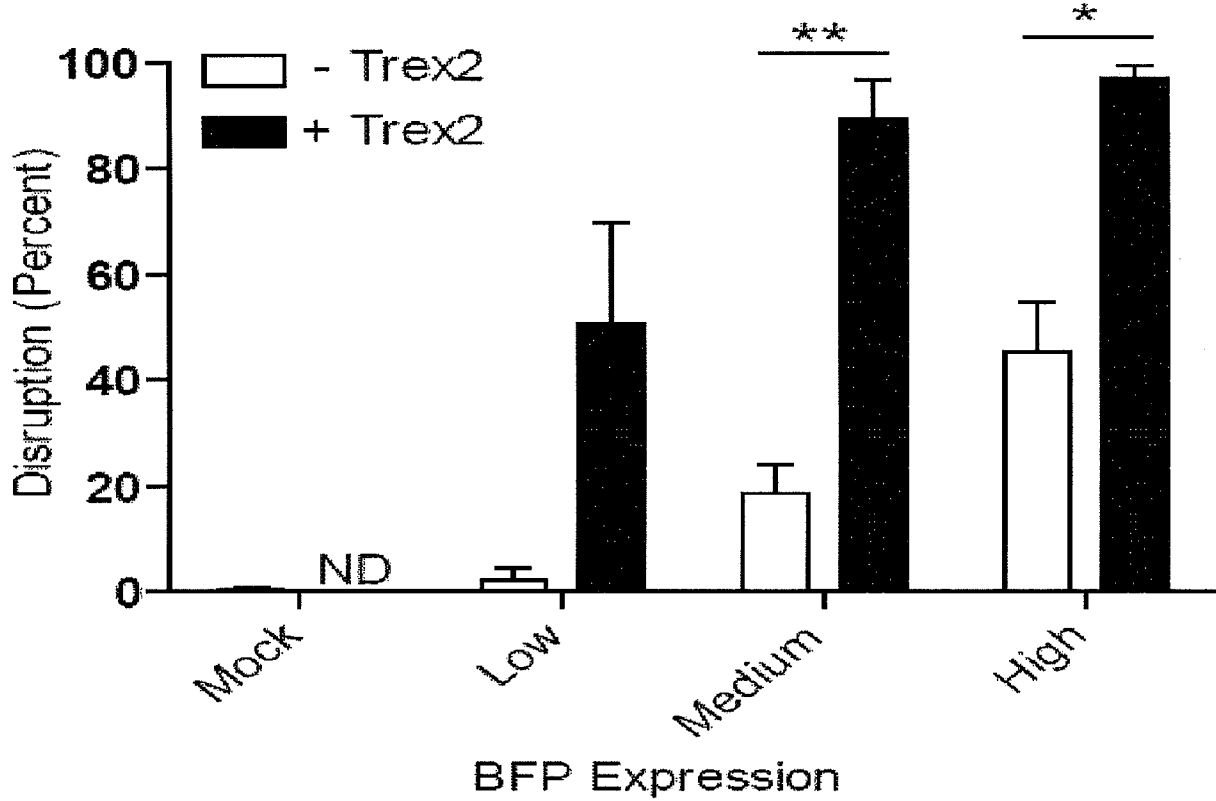


Fig. 7

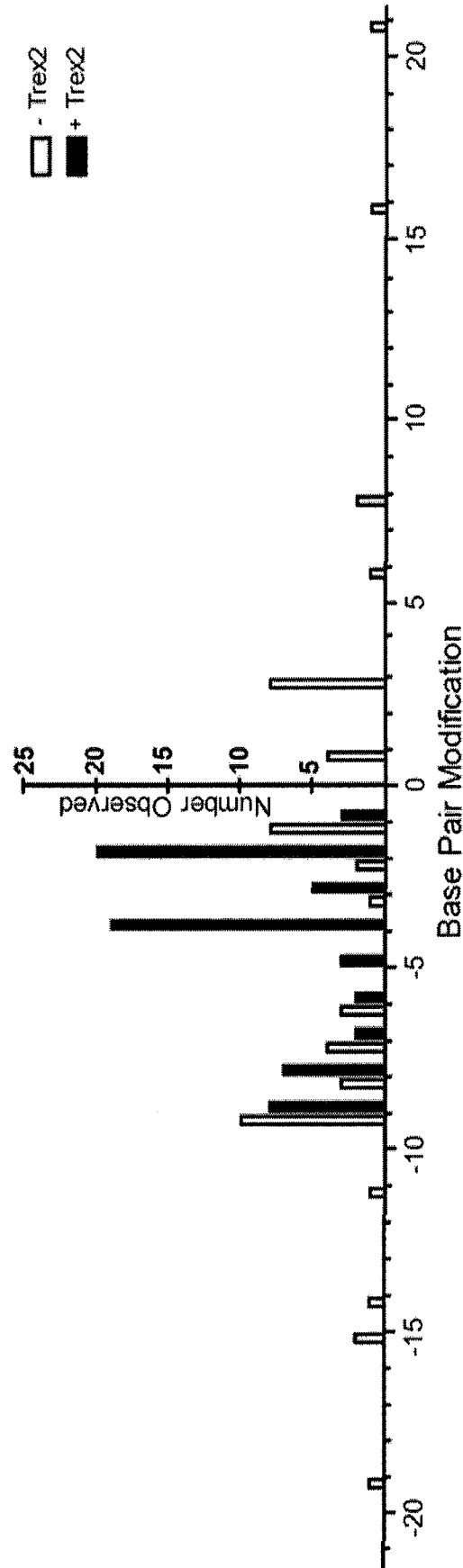


Fig. 8A

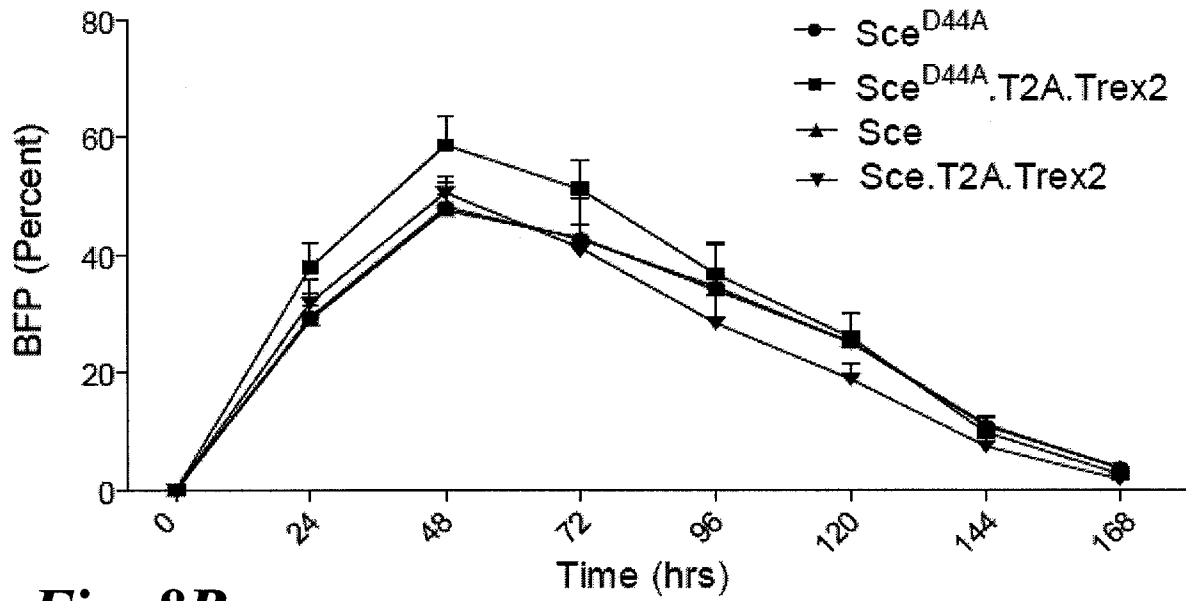


Fig. 8B

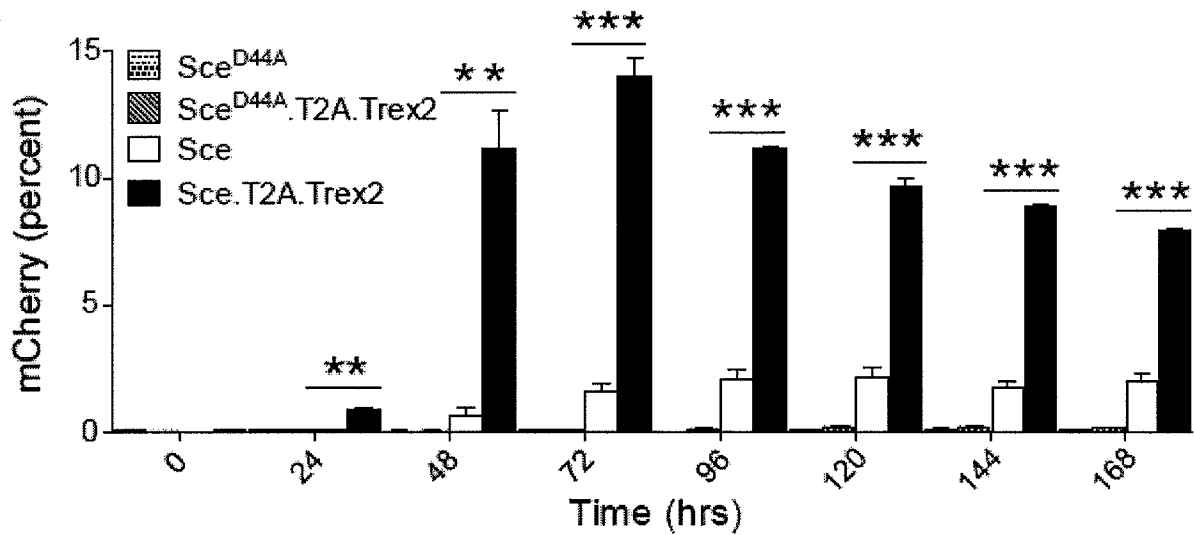


Fig. 9A

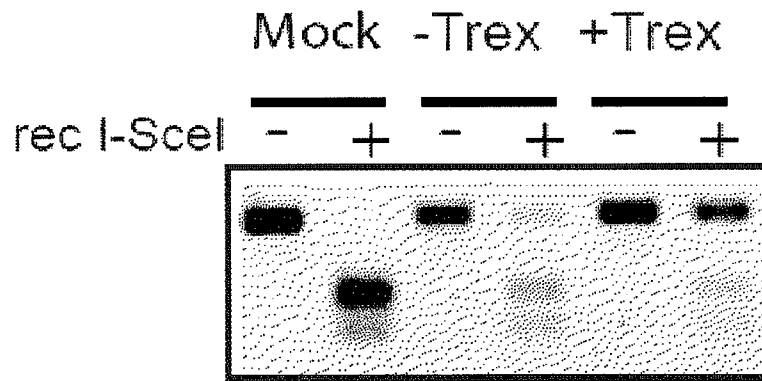


Fig. 9B

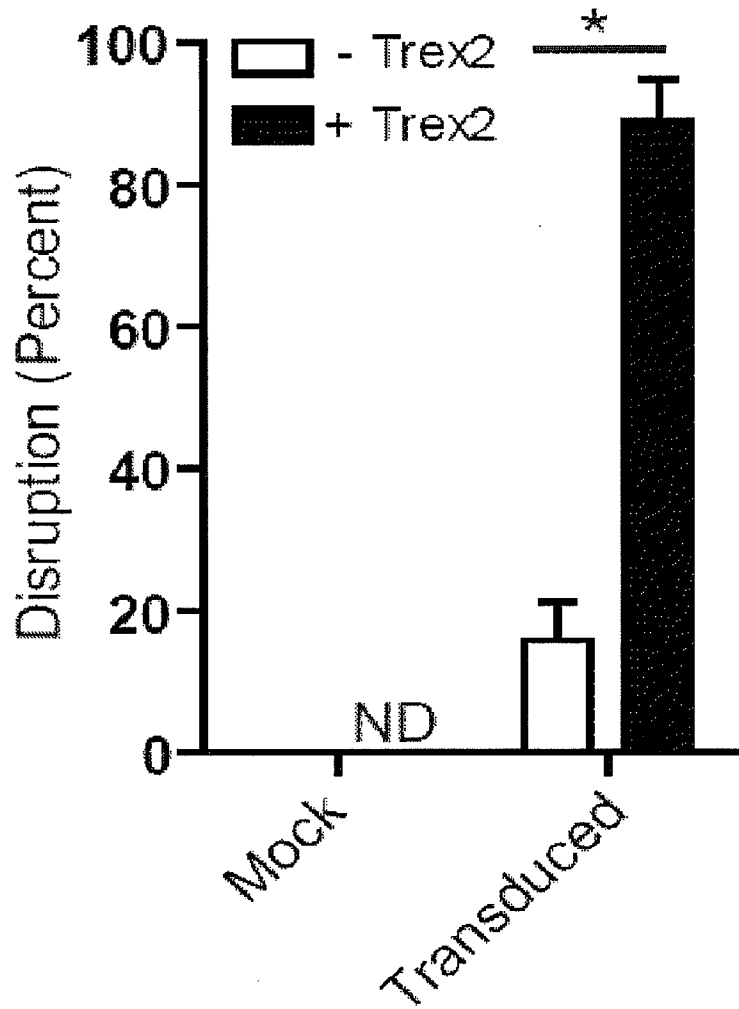


Fig. 9C

murine Lin- Bone Marrow

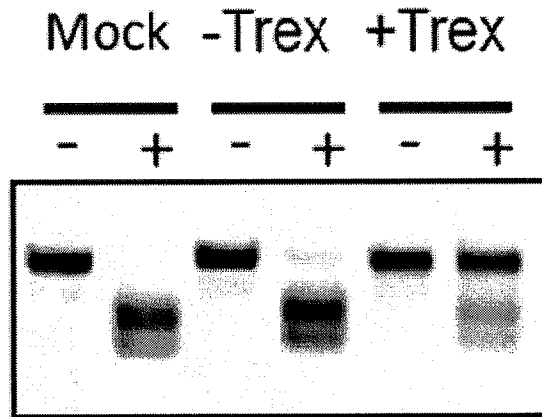


Fig. 9D

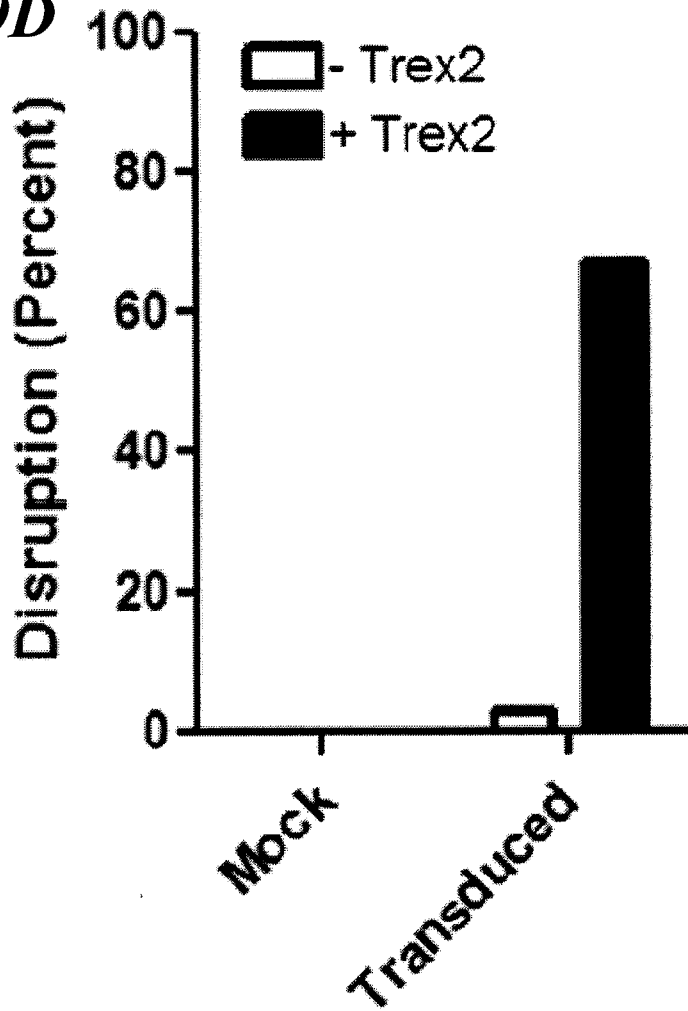


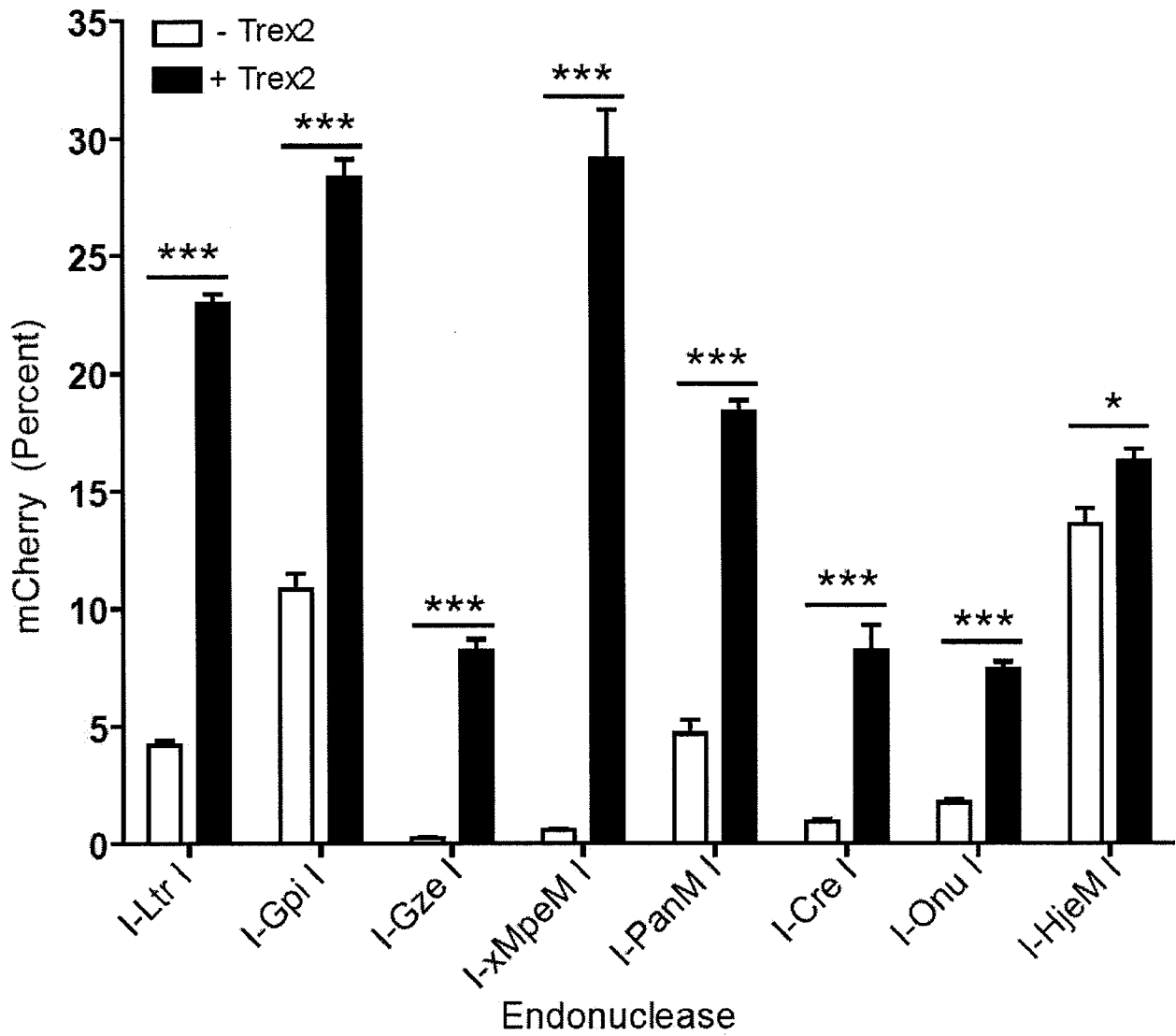
Fig. 10

Fig. 11A

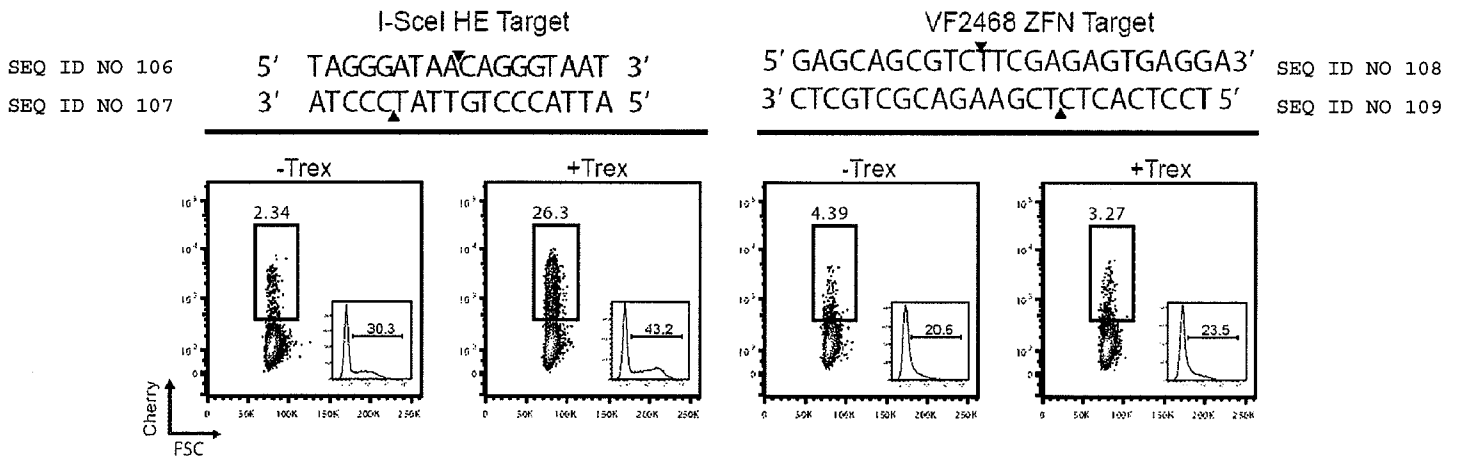


Fig. 11B

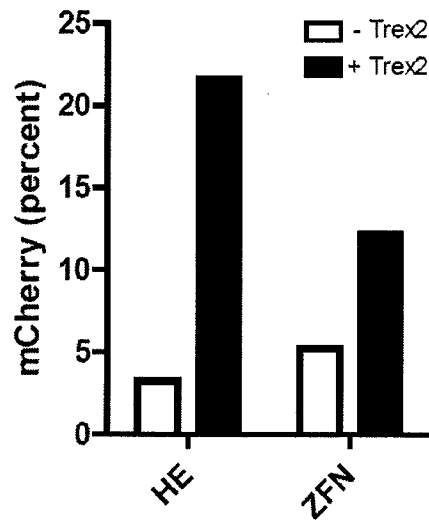


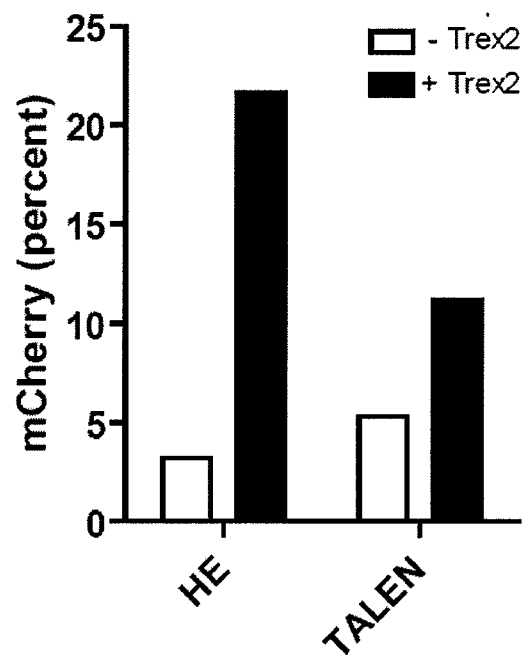
Fig. 11C

Fig. 12A

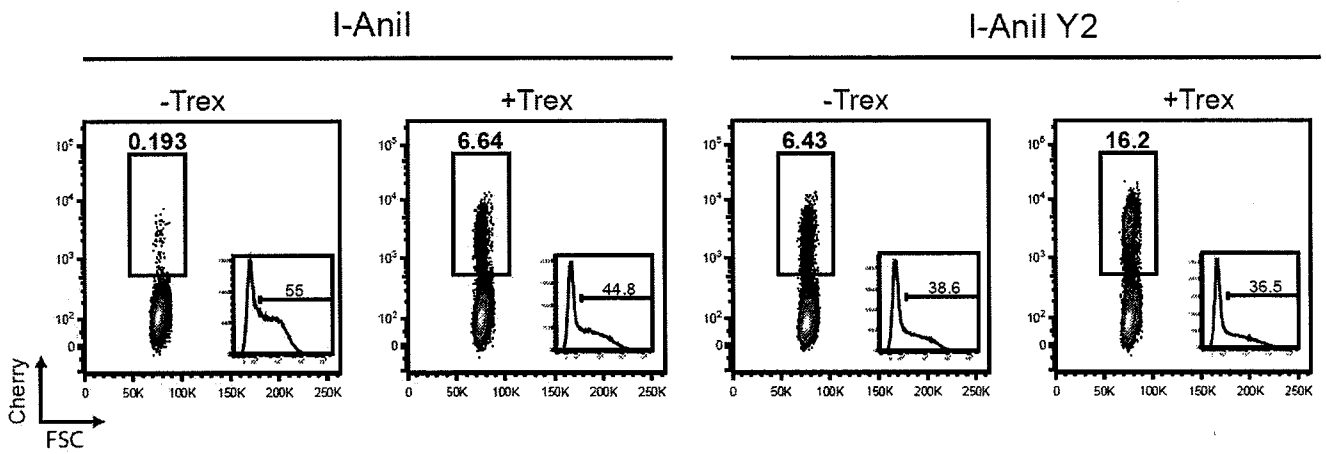


Fig. 12B

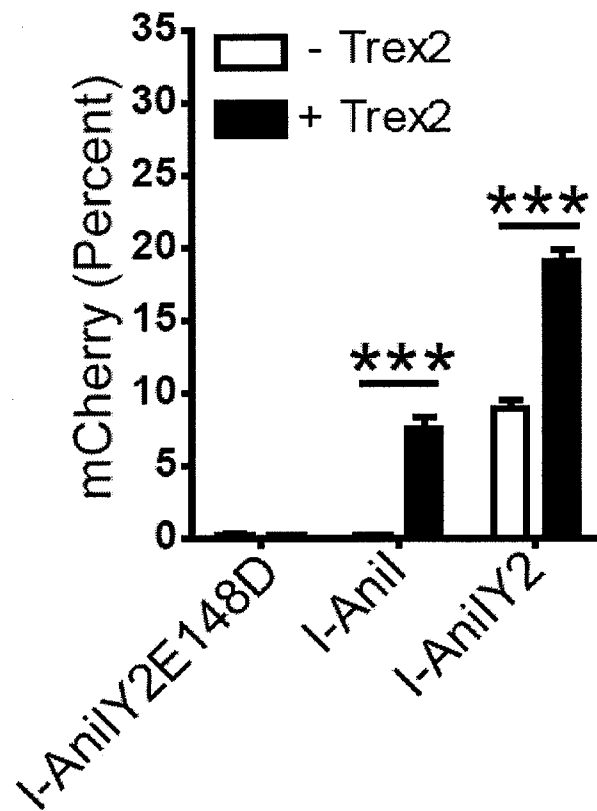


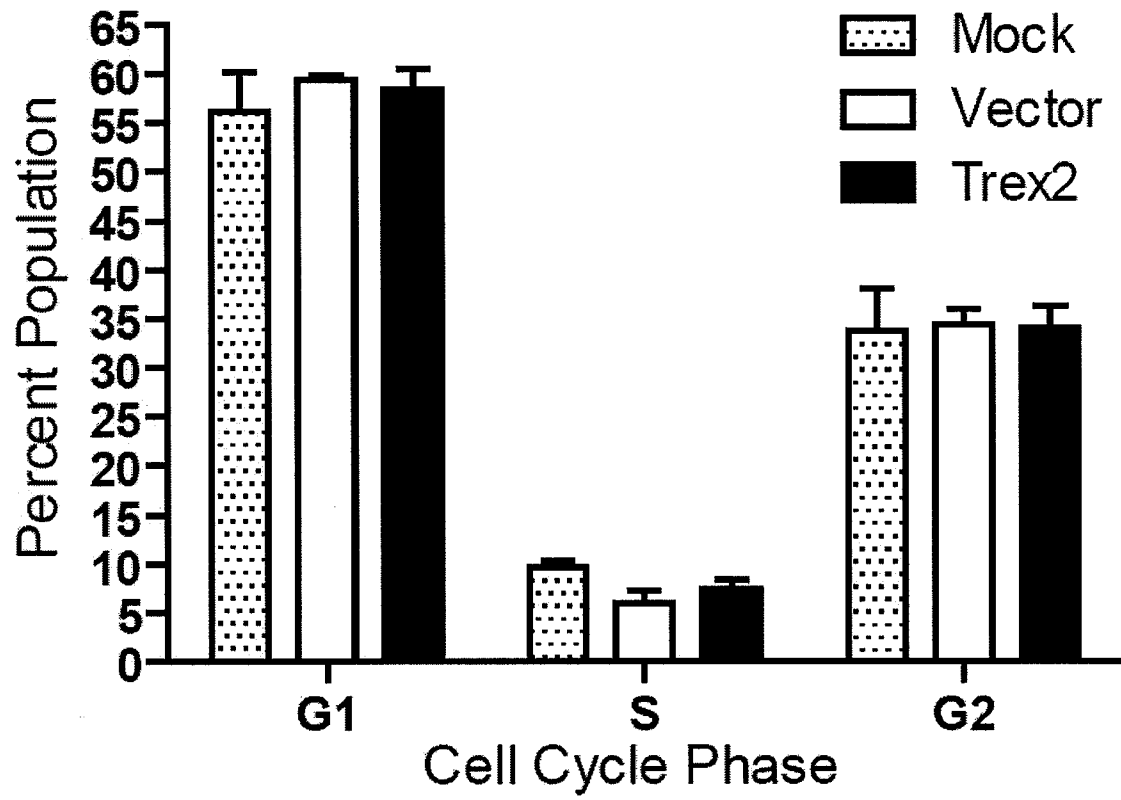
Fig. 13

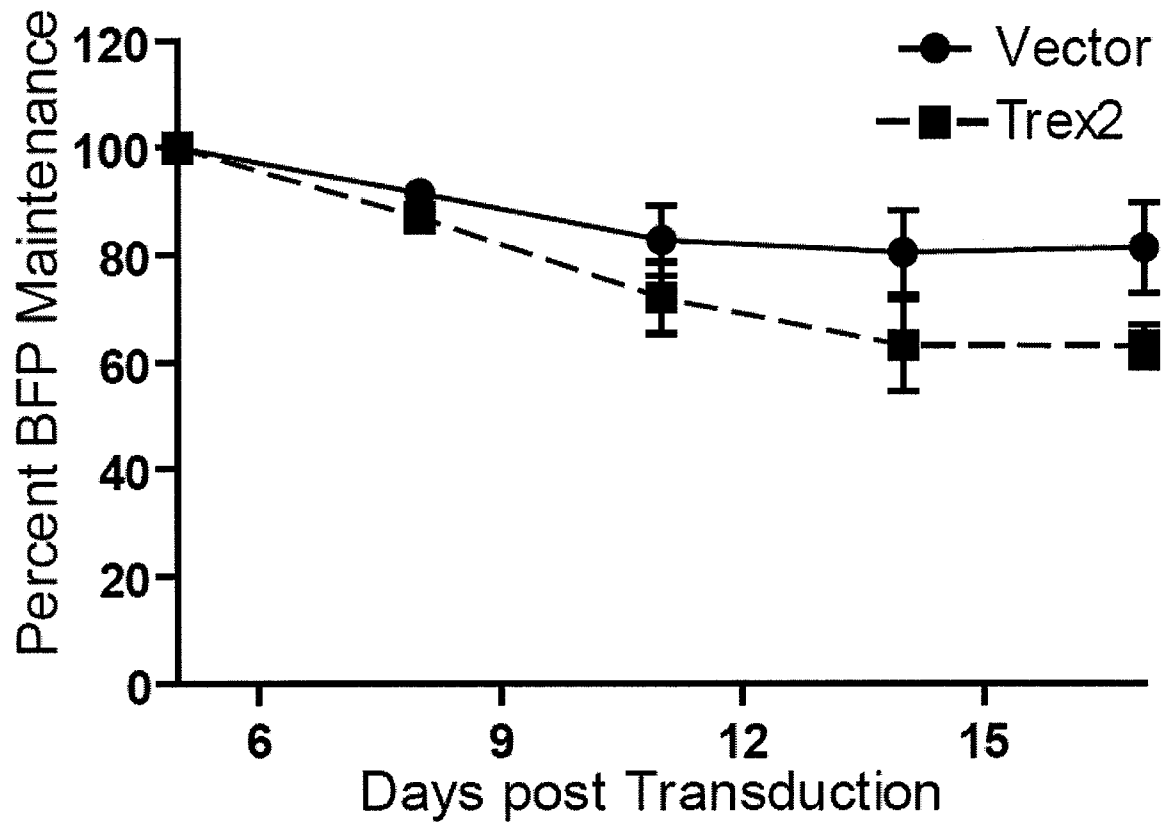
Fig. 14

Fig. 15A

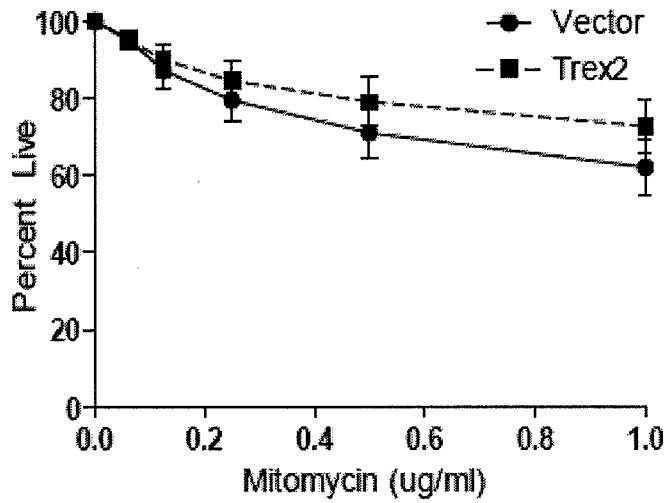


Fig. 15B

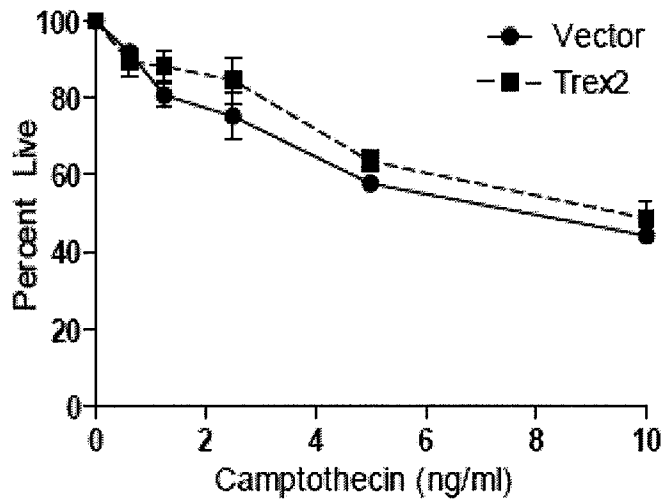


Fig. 15C

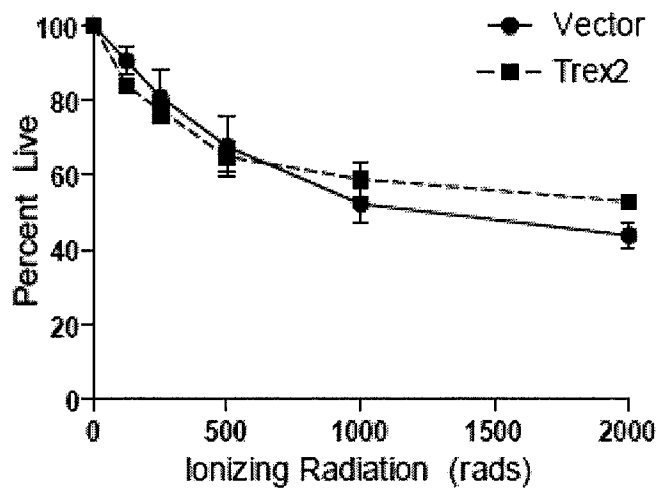


Fig. 16A

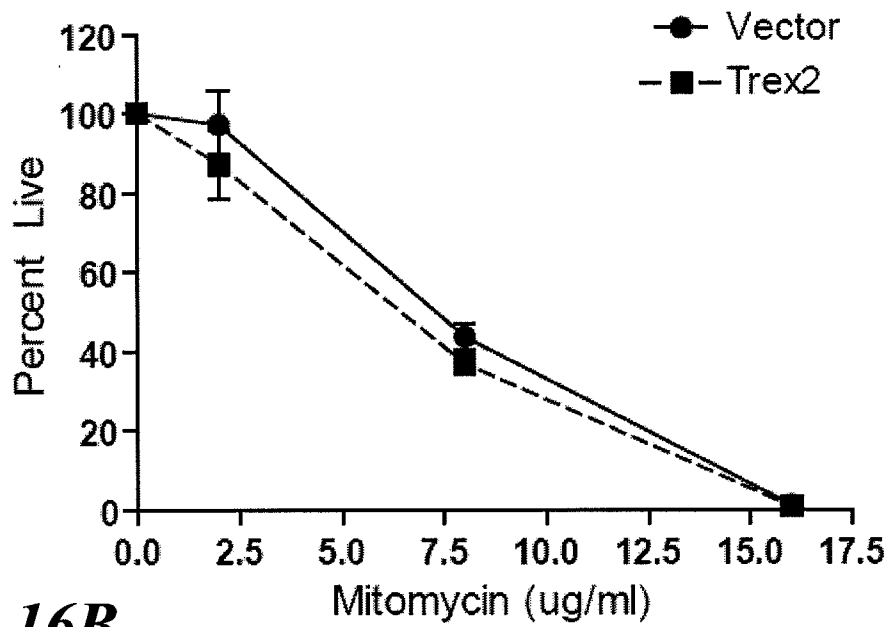


Fig. 16B

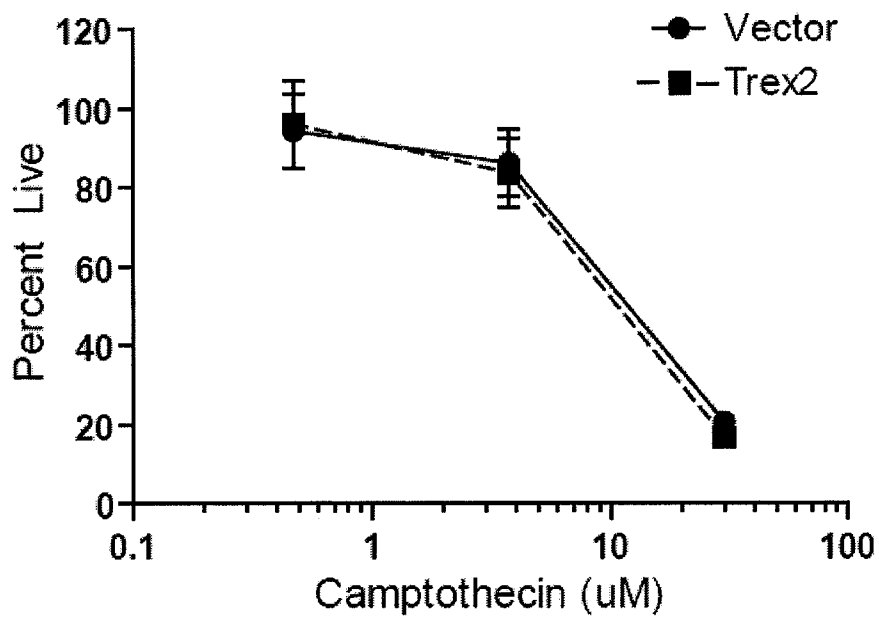


Fig. 17A

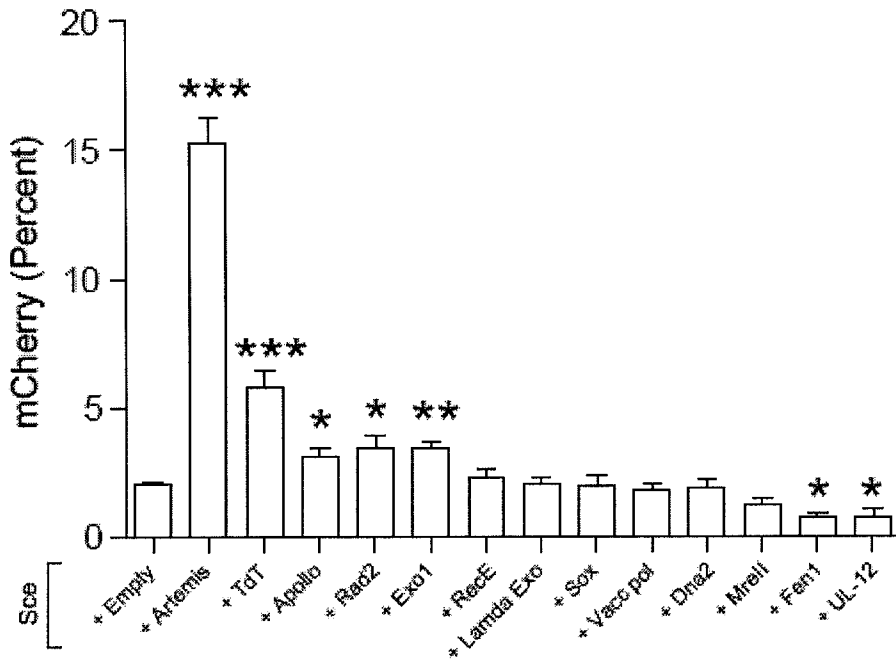
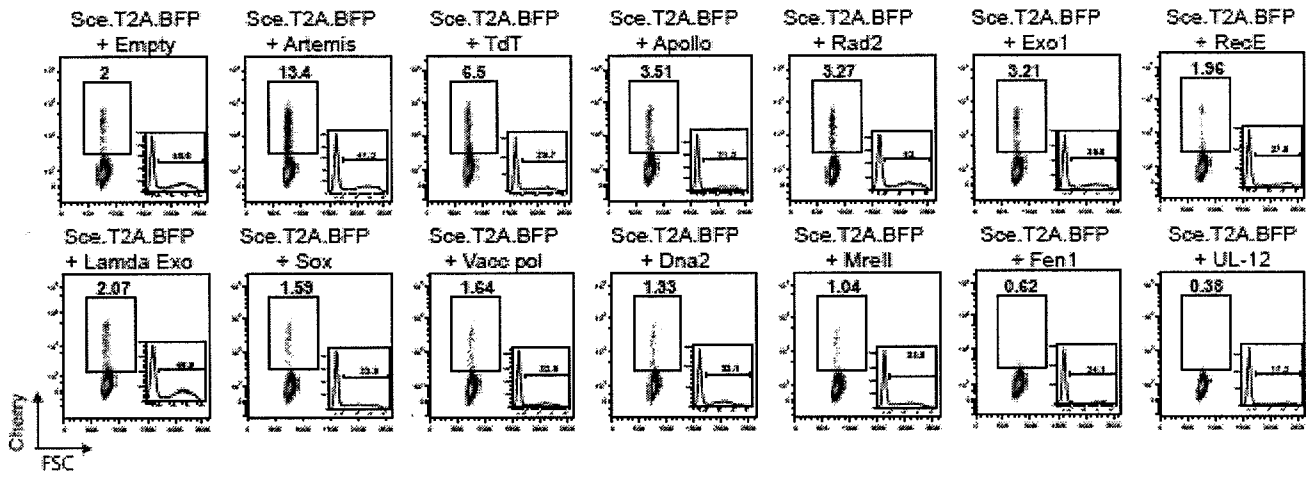


Fig. 17B

Fig. 18A

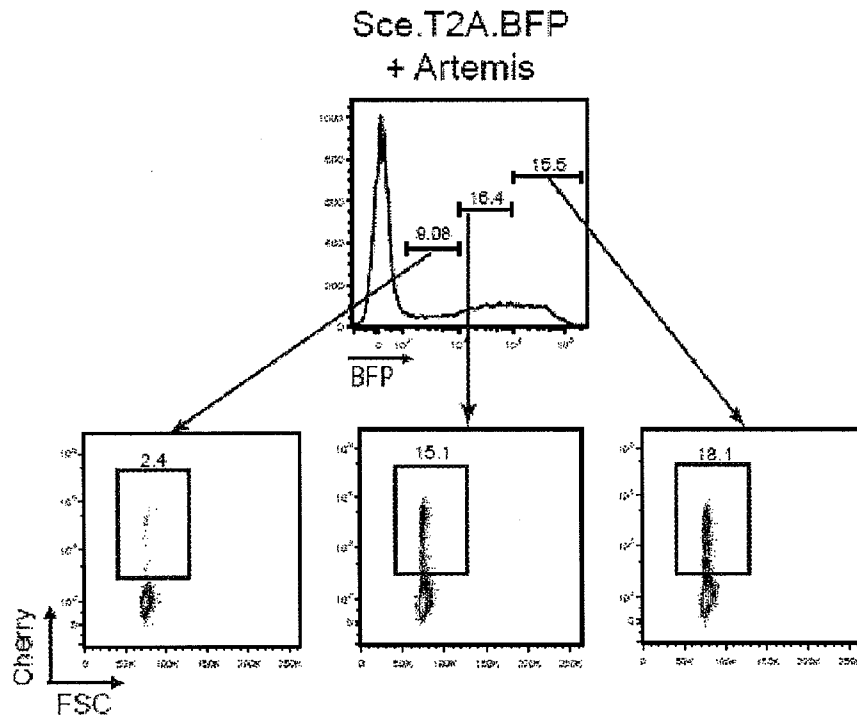


Fig. 18B

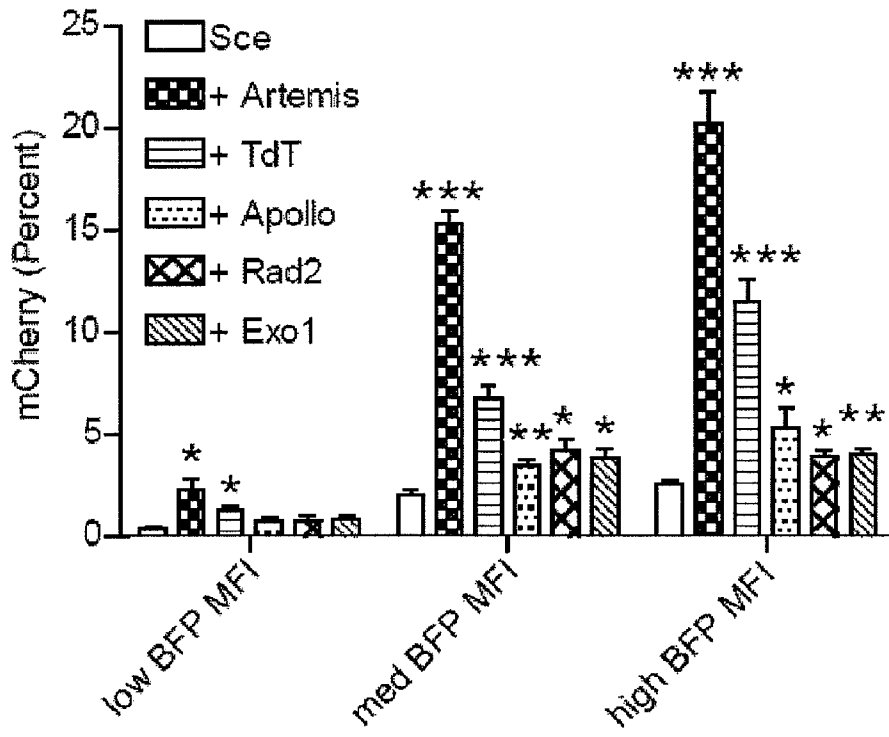


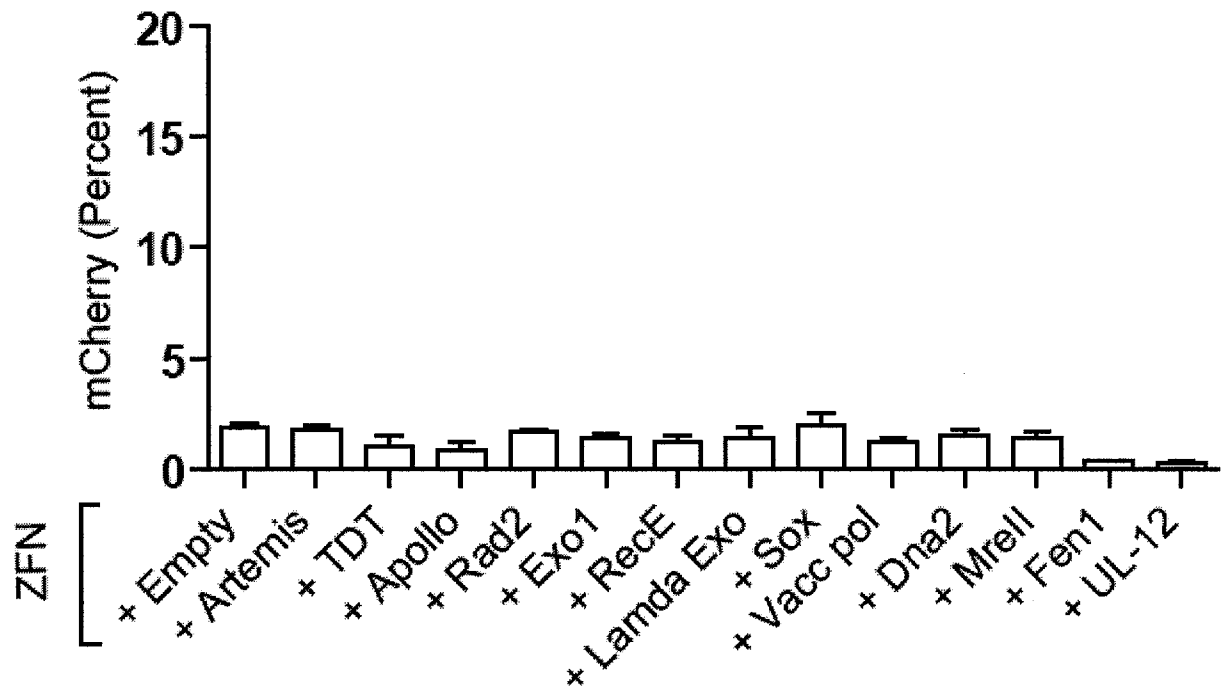
Fig. 19A

Fig. 19B

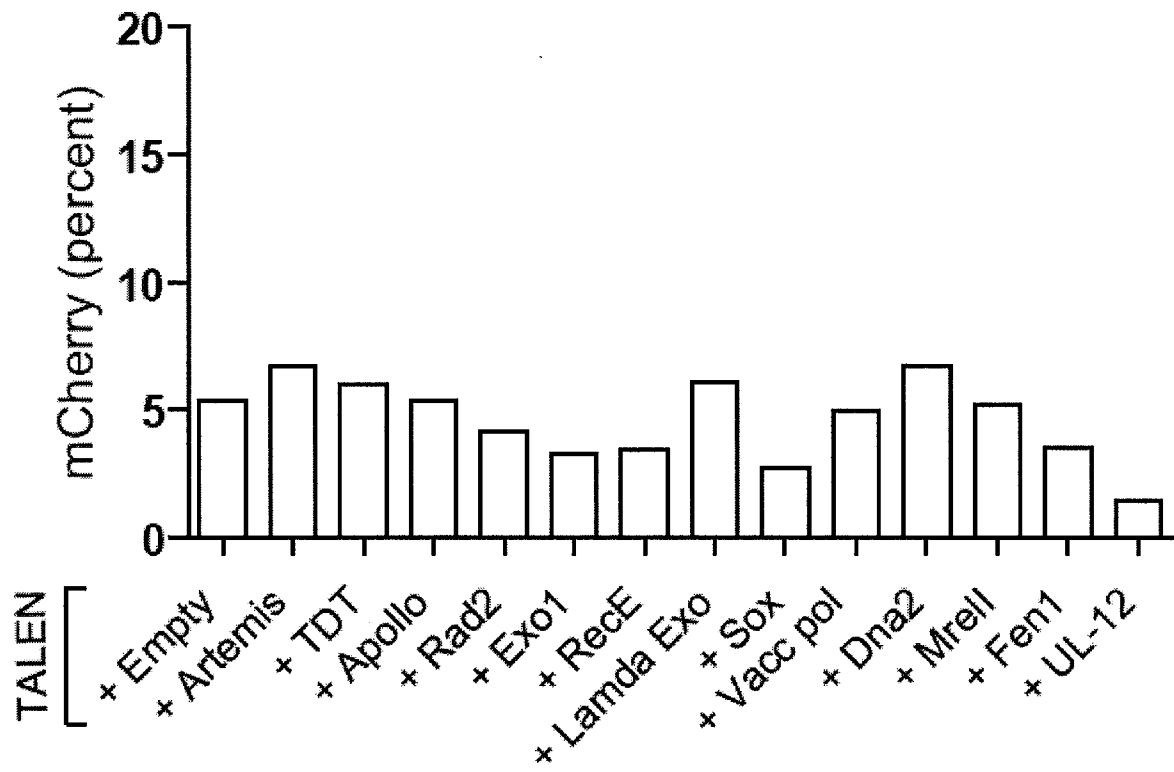


Fig. 20

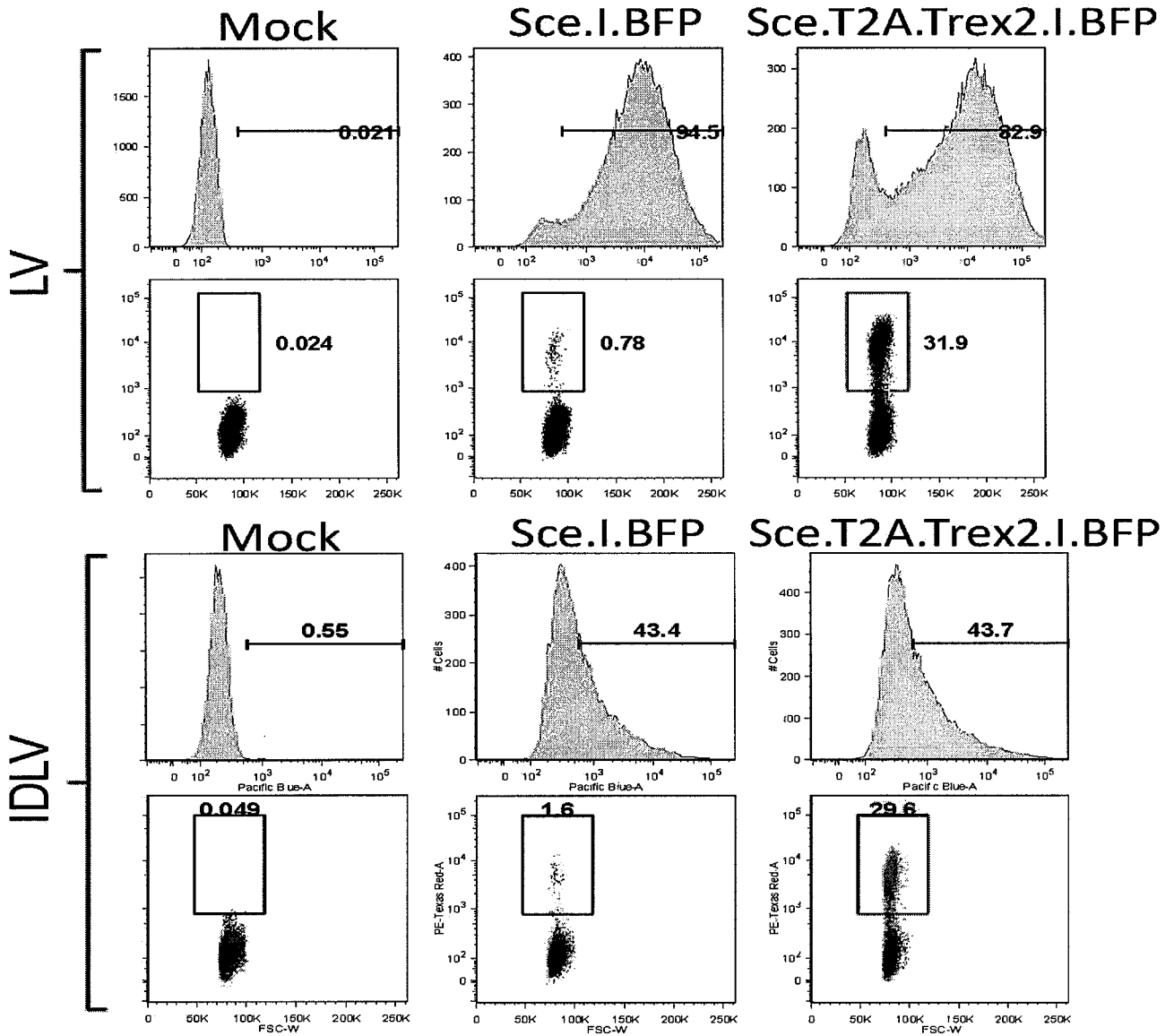


Fig. 21A

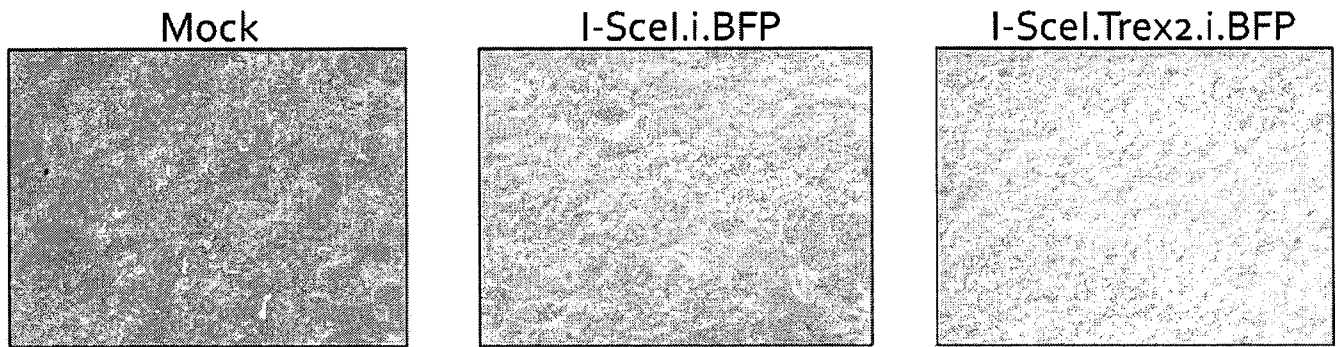


Fig. 21B

Effect of Trex2 Overexpression on Cell Growth

