

specific DSB by non-homologous end joining (NHEJ) can also result in gene disruption. Creation of two such DSBs results in deletion of arbitrarily large regions. The modular DNA recognition preferences of zinc-fingers protein allows for the rational design of site-specific multi-finger DNA binding proteins. Fusion of the nuclease domain from the Type II restriction enzyme *Fok I* to site-specific zinc-finger proteins allows for the creation of site-specific nucleases. See, for example, United States Patent Publications 20030232410; 20050208489; 20050026157; 20050064474; 20060188987; 20060063231; 20070134796; 2008015164; 20080131962; 2008015996 and International Publication WOs 07/014275 and 2008/133938, which all describe use of zinc-finger nucleases and which are incorporated by reference in their entireties for all purposes.

SUMMARY

[0006] Disclosed herein are compositions and methods for targeted insertion into a *Rosa* gene locus. The compositions and methods described herein can be used for genome editing, including, but not limited to: cleaving of one or more genes in an animal cell resulting in targeted alteration (insertion, deletion and/or substitution mutations) in one or more genes, including the incorporation of these targeted alterations into the germline; targeted introduction of non-endogenous nucleic acid sequences, the partial or complete inactivation of one or more genes in an animal; methods of inducing homology-directed repair, generation of transgenic animals (e.g., rodents) and/or generation of random mutations encoding novel allelic forms of animal genes.

[0006a] In one aspect, the present invention refers to an isolated fusion protein comprising a nuclease and an engineered zinc finger domain, wherein the engineered zinc finger domain binds to a target site in a *Rosa* gene and the zinc finger domain comprises four, five or six zinc finger recognition regions designated and ordered F1 to F4, or F1 to F5 or F1 to F6 from N-terminus to C-terminus and further wherein the zinc finger DNA-binding domains comprise recognition regions selected from the group consisting of:

- (i) F1: DRSDLSR (SEQ ID NO:38);
F2: RSDDLTR (SEQ ID NO:39);
F3: TSGHLSR (SEQ ID NO:40);
F4: RSDNLSV (SEQ ID NO:41); and
F5: RSANLTR (SEQ ID NO:42);
- (ii) F1: QSDHLTK (SEQ ID NO:43);
F2: NSSNLSR (SEQ ID NO:44);
F3: RSDHLTK (SEQ ID NO:45);
F4: NSDHLSR (SEQ ID NO:46); and
F5: RSDHLSR (SEQ ID NO:47);
- (iii) F1: RSDHLSE (SEQ ID NO:48);
F2: RSAALAR (SEQ ID NO:49);
F3: RSDHLST (SEQ ID NO:50);
F4: QNAHRIT(SEQ ID NO:51); and
F5: RSAVLSE (SEQ ID NO:52);
- (iv) F1: QSGDLTR (SEQ ID NO:17);
F2: TSGSLTR (SEQ ID NO:18);
F3: RSANLTR (SEQ ID NO:42);
F4: RSDHLTK (SEQ ID NO:45); and
F5: NSDHLSR (SEQ ID NO:46);
- (v) F1: RSANLTR (SEQ ID NO:42);
F2: QSGDLTR (SEQ ID NO:17);
F3: QSGDLTR (SEQ ID NO:17);
F4: RSANLAR (SEQ ID NO:53); and

- F5: RSDNLRE (SEQ ID NO:54);
- (vi) F1: RSDHLST (SEQ ID NO:50);
F2: DNRDRIK (SEQ ID NO:55);
F3: RSDTLSE (SEQ ID NO:56);
F4: QSSHLAR (SEQ ID NO:57); and
F5: QNAHRKT (SEQ ID NO:22);
- (vii) F1: QSGDLTR (SEQ ID NO:17);
F2: QSGDLTR (SEQ ID NO:17);
F3: RSDNLTR (SEQ ID NO:58);
F4: RSDNLSE (SEQ ID NO:21); and
F5: QNAHRKT (SEQ ID NO:22);
- (viii) F1: DRSDLSR (SEQ ID NO:38);
F2: RSDHLST (SEQ ID NO:50);
F3: DNRDRIK (SEQ ID NO:55);
F4: RSDTLSE (SEQ ID NO:56); and
F5: QSSHLAR (SEQ ID NO:57);
- (ix) F1: QSGDLTR (SEQ ID NO:17);
F2: RSDNLTR (SEQ ID NO:58);
F3: RSDNLSE (SEQ ID NO:21);
F4: QNAHRKT (SEQ ID NO:22);
F5: RSDHLSE (SEQ ID NO:48); and
F6: TSSTRKT (SEQ ID NO:59);
- (x) F1: TSGNLTR (SEQ ID NO:60);
F2: QSGNLAR (SEQ ID NO:61);
F3: RSDALSV (SEQ ID NO:62);
F4: DSSHRTR (SEQ ID NO:63); and
F5: RSDVLSE (SEQ ID NO:64)
- (xi) F1: RSDNLSE (SEQ ID NO:21);
F2: QNAHRKT (SEQ ID NO:22);
F3: RSDHLSE (SEQ ID NO:48);
F4: TSSTRKT (SEQ ID NO:59); and
F5: TSGHLR (SEQ ID NO:40);
- (xii) F1: TSGNLTR (SEQ ID NO:60);
F2: QSGNLAR (SEQ ID NO:61);

F3: RSDALSV (SEQ ID NO:62); and
F4: DSSHRTR (SEQ ID NO:63);
 (xiii) F1: QRSNLVR (SEQ ID NO:65);
F2: RSDHLTQ (SEQ ID NO:66);
F3: QSGHLQR (SEQ ID NO:67); and
F4: DRSHLAR (SEQ ID NO:68);
 (xiv) F1: RSDVLSE (SEQ ID NO:64);
F2: QRNHRTT (SEQ ID NO:69);
F3: TKRSLIE (SEQ ID NO:70);
F4: TSSNLSR (SEQ ID NO:71);
F5: RSDDLK (SEQ ID NO:25); and
F6: DNRDRIK (SEQ ID NO:55);
 (xv) F1: RSDHLSA (SEQ ID NO:72);
F2: QSGHLR (SEQ ID NO:24);
F3: RSDHLSR (SEQ ID NO:47);
F4: QNDNRIK (SEQ ID NO:73); and
F5: QSGNLAR (SEQ ID NO:61); and
 (xvi) F1: NNRDLIN (SEQ ID NO:74);
F2: TSSNLSR (SEQ ID NO:71);
F3: RSDVLSE (SEQ ID NO:64);
F4: QRNHRTT (SEQ ID NO:69);
F5: TKRSLIE (SEQ ID NO:70); and
F6: TSSNLSR (SEQ ID NO:71).

[0007] In one example aspect, described herein is a zinc-finger protein (ZFP) that binds to target site in a *Rosa* gene in a genome (*e.g.*, a rodent genome), wherein the ZFP comprises one or more engineered zinc-finger binding domains. In one example embodiment, the ZFP is a zinc-finger nuclease (ZFN) that cleaves a target genomic region of interest, wherein the ZFN comprises one or more engineered zinc-finger binding domains and a nuclease cleavage domain or cleavage half-domain. Cleavage domains and cleavage half domains can be obtained, for example, from various restriction endonucleases and/or homing endonucleases. In one example embodiment, the cleavage half-domains are derived from a Type IIS restriction endonuclease (*e.g.*, *Fok I*). In certain example embodiments, the zinc finger domain recognizes a target site in a *Rosa* gene, for example *Rosa26*.

[0008] The ZFN may bind to and/or cleave a *Rosa* gene within the coding region of the gene or in a non-coding sequence within or adjacent to the gene, such as, for example, a leader sequence, trailer sequence or intron, or within a non-transcribed region, either upstream or downstream of the coding region.

[0009] In another exampleaspect, described herein are compositions comprising one or more of the zinc-finger nucleases described herein. In certain exampleembedments, the composition comprises one or more zinc-finger nucleases in combination with a pharmaceutically acceptable excipient.

[0009a] In another aspect, the present invention refers to a polynucleotide encoding an isolated fusion protein as defined herein.

[0010] In another exampleaspect, described herein is a polynucleotide encoding one or more ZFNs described herein. The polynucleotide may be, for example, mRNA.

[0011] In another exampleaspect, described herein is a ZFN expression vector comprising a polynucleotide, encoding one or more ZFNs described herein, operably linked to a promoter.

[0011a] In yet another aspect, the present invention refers to an isolated cell comprising a fusion protein as defined herein or a polynucleotide as defined herein.

[0012] In another exampleaspect, described herein is a host cell comprising one or more ZFN expression vectors as described herein. The host cell may be stably transformed or transiently transfected or a combination thereof with one or more ZFP expression vectors. In one exampleembedment, the host cell is an embryonic stem cell. In other exampleembedments, the one or more ZFP expression vectors express one or more ZFNs in the host cell. In another exampleembedment, the host cell may further comprise an exogenous polynucleotide donor sequence. In any of the exampleembedments, described herein, the host cell can comprise an embryo cell, for example a one or more mouse, rat, rabbit or other mammal cell embryo.

[0012a] In a further aspect, the present invention refers to an *in vitro* method for cleaving one or more *Rosa* genes in an isolated cell, the method comprising: introducing, into the isolated cell, one or more fusion proteins as defined herein or one or more polynucleotides as defined herein into the isolated cell, such that the one or more *Rosa* gene are cleaved.

[0013] In another exampleaspect, described herein is a method for cleaving one or more *Rosa* genes in a cell, the method comprising: (a) introducing, into the cell, one or more polynucleotides encoding one or more ZFNs that bind to a target site in the one or more genes under conditions such that the ZFN(s) is (are) expressed and the one or more genes are cleaved.

[0013a] In another aspect, the present invention refers to an *in vitro* method of introducing the exogenous polynucleotide sequence into the genome of an isolated cell, the method comprising: cleaving one or more *Rosa* genes as defined herein; and contacting the isolated cell with an exogenous polynucleotide sequence; wherein cleavage of the one or more genes stimulates integration of the exogenous polynucleotide sequence into the genome by homologous recombination.

[0014] In yet another ~~example~~ aspect, described herein is a method for introducing an exogenous sequence into the genome of a cell, the method comprising the steps of: (a) introducing, into the cell, one or more polynucleotides encoding one or more ZFNs that bind to a target site in a *Rosa* gene under conditions such that the ZFN(s) is (are) expressed and the one or more genes are cleaved; and (b) contacting the cell with an exogenous polynucleotide; such that cleavage of the gene(s) stimulates integration of the exogenous polynucleotide into the genome by homologous recombination. In

certain exampleembodiments, the exogenous polynucleotide is integrated physically into the genome. In other exampleembodiments, the exogenous polynucleotide is integrated into the genome by copying of the exogenous sequence into the host cell genome via nucleic acid replication processes (e.g., homology-directed repair of the double strand break). In yet other exampleembodiments, integration into the genome occurs through non-homology dependent targeted integration (e.g. "end-capture"). In certain exampleembodiments, the one or more nucleases are fusions between the cleavage domain of a Type IIS restriction endonuclease and an engineered zinc-finger binding domain. In certain exampleembodiments, the exogenous sequence is integrated into a small mammal (e.g. rabbit or rodent such as mouse, rat, etc.) *Rosa* gene.

[0015] In another exampleembodiment, described herein is a method for modifying one or more *Rosa* gene sequence(s) in the genome of cell, the method comprising (a) providing a cell comprising one or more *Rosa* sequences; and (b) expressing first and second zinc-finger nucleases (ZFNs) in the cell, wherein the first ZFN cleaves at a first cleavage site and the second ZFN cleaves at a second cleavage site, wherein the gene sequence is located between the first cleavage site and the second cleavage site, wherein cleavage of the first and second cleavage sites results in modification of the gene sequence by non-homologous end joining and/or homology directed repair. Optionally, the cleavage results in insertion of an exogenous sequence (transgene) also introduced into the cell. In other exampleembodiments, non-homologous end joining results in a deletion between the first and second cleavage sites. The size of the deletion in the gene sequence is determined by the distance between the first and second cleavage sites. Accordingly, deletions of any size, in any genomic region of interest, can be obtained. Deletions of 25, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1,000 nucleotide pairs, or any integral value of nucleotide pairs within this range, can be obtained. In addition deletions of a sequence of any integral value of nucleotide pairs greater than 1,000 nucleotide pairs can be obtained using the methods and compositions disclosed herein.

[0015a] In one aspect, the present invention refers to a method for generating a non-human transgenic animal, wherein the method comprises: modifying a *Rosa* gene sequence in a non-human embryo cell as described herein; and allowing the non-human embryo to develop into an animal.

[0016] Methods of modifying an endogenous *Rosa* gene as described herein can be used to create models of animal (e.g., human) disease, for example by inactivating (partially or fully) a gene or by creating random mutations at defined positions of genes that allow for the identification or selection of transgenic animals (e.g., rats, rabbits or mice) carrying novel allelic forms of those genes, by insertion of humanized genes (to study, by way of a non-limiting example, drug metabolism) or

by insertion of a mutant alleles of interest to examine, for example, the phenotypic affect of such a mutant allele.

[0017] In yet another exampleaspect, described herein is a method for germline disruption of one or more target *Rosa* genes, the method comprising modifying one or more *Rosa* sequences in the genome of one or more cells of an embryo by any of the methods described herein and allowing the embryo to develop, wherein that the modified gene sequences are present in at least a portion of gametes of the sexually mature animal. In certain exampleembodiments, the animal is a small mammal, such as a rodent or rabbit.

[0017a] In one aspect, the present invention refers to a method of creating one or more heritable mutant alleles in at least one *Rosa* locus of interest, the method comprising: generating a non-human transgenic animal as defined herein, wherein the non-human embryo is raised to sexual maturity; and allowing the sexually mature non-human animal to produce offspring; wherein at least some of the offspring comprise the mutant alleles.

[0018] In another exampleaspect, described herein is a method of creating one or more heritable mutant alleles in at least one *Rosa* locus of interest, the method comprising modifying one or more *Rosa* loci in the genome of one or more cells of an animal embryo by any of the methods described herein; raising the embryo to sexual maturity; and allowing the sexually mature animal to produce offspring; wherein at least some of the offspring comprise the mutant alleles. In certain exampleembodiments, the animal is a small mammal, for example a rabbit or a rodent such as rat, a mouse or a guinea pig.

[0019] In any of the methods described herein, the polynucleotide encoding the zinc finger nuclease(s) can comprise DNA, RNA or combinations thereof. In certain exampleembodiments, the polynucleotide comprises a plasmid. In other exampleembodiments, the polynucleotide encoding the nuclease comprises mRNA.

[0020] In a still further exampleaspect, ~~disclosed~~provided herein is a method for site specific integration of a nucleic acid sequence into a *Rosa* locus of a chromosome. In certain exampleembodiments, the method comprises: (a) injecting an embryo with (i) at least one DNA vector, wherein the DNA vector comprises an upstream sequence and a downstream sequence flanking the nucleic acid sequence to be integrated, and (ii) at least one RNA molecule encoding a zinc finger nuclease that recognizes the site of integration in the *Rosa* locus, and (b) culturing the embryo to allow expression of the zinc finger nuclease, wherein a double stranded break introduced into the site of integration by the zinc finger nuclease is repaired, via homologous recombination with the DNA vector, so as to integrate the nucleic acid sequence into the chromosome.

[0021] Suitable embryos may be derived from several different vertebrate species, including mammalian, bird, reptile, amphibian, and fish species. Generally

speaking, a suitable embryo is an embryo that may be collected, injected, and cultured to allow the expression of a zinc finger nuclease. In some ~~example~~embodiments, suitable embryos may include embryos from small mammals (*e.g.*, rodents, rabbits, etc.), companion animals, livestock, and primates. Non-limiting examples of rodents may include mice, rats, hamsters, gerbils, and guinea pigs. Non-limiting examples of companion animals may include cats, dogs, rabbits, hedgehogs, and ferrets. Non-limiting examples of livestock may include horses, goats, sheep, swine, llamas, alpacas, and cattle. Non-limiting examples of primates may include capuchin monkeys, chimpanzees, lemurs, macaques, marmosets, tamarins, spider monkeys, squirrel monkeys, and vervet monkeys. In other embodiments, suitable embryos may include embryos from fish, reptiles, amphibians, or birds. Alternatively, suitable embryos may be insect embryos, for instance, a *Drosophila* embryo or a mosquito embryo.

[0022] Also ~~disclosed~~provided is an embryo comprising at least one DNA vector, wherein the DNA vector comprises an upstream sequence and a downstream sequence flanking the nucleic acid sequence to be integrated, and at least one RNA molecule encoding a zinc finger nuclease that recognizes the chromosomal site of integration. Organisms derived from any of the embryos as described herein are also provided.

[0022a] In one aspect, the present invention refers to a composition comprising a fusion protein as defined herein or a polynucleotide as defined herein and a pharmaceutically acceptable excipient.

[0022b] In another aspect, the present invention refers to a kit comprising an isolated fusion protein as defined herein or a polynucleotide as defined herein.

[0023] A kit, comprising the ZFPs of the invention, is also ~~disclosed~~provided. The kit may comprise nucleic acids encoding the ZFPs, (*e.g.* RNA molecules or ZFP encoding genes contained in a suitable expression vector), donor molecules, suitable host cell lines, instructions for performing the methods of the invention, and the like.

BRIEF DESCRIPTION OF THE DRAWINGS

[0024] Figure 1 depicts a Southern blot demonstrating the results of NHEJ repair following cleavage of the rat *rosa26* locus as assayed by the Surveyor™ (Transgenomic) mismatch assay. "G" indicates reactions where cells were transfected with GFP ZFNs, and numbered lanes indicate specific ZFN pairs. Arrows indicate lanes where NHEJ has occurred.

[0025] Figure 2 depicts insertion of *Rosa*-targeted donor nucleotides into mouse genomic DNA.