nicked site, without simultaneously occurrence of error-prone NHEJ repair in mammalian/human cells.

SUMMARY

[0010] Disclosed herein are methods and compositions for inducing a targeted single-stranded break in any double-stranded target sequence of interest. Also described are methods of facilitating homologous recombination and targeted integration following single-stranded cleavage of a target. Thus, targeted modulation of a genome is described.

[0010a] In one aspect, the present invention refers to a protein complex comprising:

(a) a first fusion protein comprising a DNA-binding domain and at least one catalytically inactive *FokI* cleavage domain or cleavage half-domain, wherein the cleavage domain or cleavage half-domain comprises a mutation at one or more of amino acid residues 450, 467 and 469; and

(b) a second fusion protein comprising a zinc finger domain and a *FokI* catalytically active cleavage half domain,

wherein the catalytically inactive *FokI* cleavage half domain forms a heterodimer with the catalytically active *FokI* cleavage half-domain of the second fusion protein and further wherein the protein complex induces a single-stranded break in a target double-stranded sequence.

[0010b] In another aspect, the present invention refers to a composition comprising a protein complex according to the present invention or one or more polynucleotides encoding the first and second fusion proteins of the protein complex of the present invention.

[0010c] In another aspect, the present invention refers to an isolated cell or cell line comprising a protein complex according to the present invention.

[0010d] In another aspect, the present invention refers to an *in vitro* method of generating a single-stranded break in a target double-stranded sequence in a cell, the method comprising:

providing a protein complex according to the present invention, wherein the heterodimer generates a single-stranded break in the target double-stranded sequence.

[0010e] In another aspect, the present invention refers to an *in vitro* method of inserting an exogenous sequence into a region of interest in a cell, the method comprising

generating a single-stranded break in the region of interest according to the method of the present invention; and

providing an exogenous sequence that is inserted into the region of interest following generation of a single-stranded break.

[0010f] In another aspect, the present invention refers to an *in vitro* method of inactivating a target sequence in a region of interest of a cell, the method comprising

generating a single-stranded break in the region of interest according to the method of the present invention, and

providing an exogenous sequence that replaces the wild-type sequence in the region of interest and wherein the exogenous sequence partially or fully inactivates the target sequence.

[0011] In one ~~aspect~~example, artificial (non-naturally occurring) nucleases that generate single-stranded cuts in a desired double-stranded target sequence are provided. The nucleases described herein comprise a DNA-binding domain (*e.g.*, engineered zinc finger protein) and at least one cleavage domain or at least one cleavage half-domain. In certain ~~embodiments~~examples, the nucleases are zinc finger nucleases comprising a zinc finger domain that is engineered to bind any selected sequence (*e.g.*, gene). Any of the zinc finger proteins described herein may include 1, 2, 3, 4, 5, 6 or more zinc fingers, each zinc finger having a recognition helix that binds to a target subsite in the selected sequence(s) (*e.g.*, gene(s)). The cleavage domain can be derived from any nuclease, for example, a cleavage half-domain from a Type IIS nuclease, such as *FokI*. In certain ~~embodiments~~examples, the cleavage half-domain comprises an engineered *FokI* cleavage-half domain that forms a heterodimer with another cleavage half-domain (*e.g.*, engineered or wild-type in the dimerization domain). In still further ~~embodiments~~examples, the cleavage domain (*e.g.*, engineered *FokI* cleavage half-domain) comprises a mutation that inactivates the catalytic domain (enzymatic activity) of the nuclease domain.

[0012] In another ~~aspect~~example, ~~provided~~disclosed herein are complexes and/or compositions comprising a pair of zinc finger nucleases, each nuclease comprising an engineered zinc finger domain and a *FokI* cleavage half-domain, wherein the cleavage half-domain forms a dimer (homodimer or heterodimer). In certain ~~embodiments~~examples, one zinc finger nuclease of the pair comprises a first catalytically active engineered cleavage-half domain and the other zinc finger nuclease comprises a second catalytically inactive engineered cleavage half-domain, in which the first and second engineered cleavage half-domains form an obligate heterodimer.

[0013] In yet another ~~aspect~~example, a polynucleotide encoding any of the nucleases described herein is provided.

[0014] In yet another ~~aspect~~example also provided is an isolated cell comprising any of the proteins and/or polynucleotides described herein. In certain ~~embodiments~~examples, cell lines in which exogenous sequences have been introduced via targeted alteration are provided. In certain ~~embodiments~~examples, one or more selected gene(s) are inactivated (partially or completely) in these cell lines. In other ~~embodiments~~examples, the cells or cell lines include one or more transcribed and/or translated exogenous sequences that have been stably or transiently introduced into the cells. Such cell lines are generated by culturing of cells comprising any of the proteins and/or polynucleotides described herein in a cell line resulting in the inactivation of selected gene(s).

[0015] Also provided are transgenic organisms comprising one or more transcribed and/or translated exogenous sequences that have been stably or transiently introduced into the cells. Further provided are transgenic organisms that contain gene(s) selectively inactivated by the methods provided herein, or organisms containing exogenous sequences capable of altering the expression of endogenous sequences. Transgenic organisms as described herein may be plants (*e.g.* crop plants or tobacco strains) or animals (*e.g.* mice, rats, rabbits, fish, etc.). In certain ~~embodiments~~examples, the transgenic organisms are used to generate lines of organisms carrying one or more sequences encoding nucleases as described herein and/or one or more exogenous sequences (*e.g.*, sequences inserted into the genome via targeted integration using nucleases as described herein). For instance, disclosed herein are transgenic plants and plant lines comprising nucleases as described herein under the control of an inducible promoter. Accordingly, transgenic plants and plant lines comprising episomal or integrated sequences encoding nucleases can be expressed in the plant at the desired time and/or in the desired tissue of the plant.

[0016] In addition, methods of generating a specific single-stranded break in a target double-stranded sequence in a cell are provided. In certain ~~embodiments~~examples, the methods involve introducing one or more pairs of nucleases (proteins or polynucleotide(s) encoding the nucleases) into the cell. Each nuclease pair comprises a first nuclease with a first DNA-binding domain and a first catalytically active cleavage-half domain and a second nuclease with a second DNA-binding domain and second catalytically inactive cleavage-half domain. The one or more pairs of nucleases are introduced into the cells under conditions such that the first and second

cleavage half-domains of each pair form a dimer and generate a single-stranded break in the target sequence.

[0017] In certain ~~embodiments~~examples, the first and/or second DNA-binding domains comprise a zinc finger protein (*e.g.*, an engineered zinc finger protein). Furthermore, in any of the methods described herein, the first and second cleavage half-domains can comprise *FokI* cleavage half-domains, for example *FokI* cleavage half-domains that are engineered to form obligate heterodimers. The target sequence can be any double-stranded sequence, for example, a sequence in cellular chromatin such as a genomic sequence or portion thereof. In addition, the target sequence can be an extrachromosomal double stranded DNA sequence, for example in a plasmid or virus. Similarly, the target sequence can be in any cell type including, prokaryotic and eukaryotic cells, such as fungal cells, plant cells, animal cells, mammalian cells, primate cells and human cells.

[0018] The site of single-stranded break can be coincident with the sequence to which the catalytically active nuclease binds, or it can be adjacent (*e.g.*, separated from the near edge of the binding site by 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or more nucleotides). The fusion proteins can be expressed in a cell, *e.g.*, by delivering the fusion proteins to the cell and/or by delivering a polynucleotide encoding one or more nucleases to a cell, wherein the polynucleotide, if DNA, is transcribed into mRNA which is then translated into the fusion protein. Alternatively, an RNA molecule can be delivered to the cell which then is translated to generate the fusion protein. Methods for polynucleotide and polypeptide delivery to cells are presented elsewhere in this disclosure.

[0019] Methods for targeted recombination (for, *e.g.*, alteration or replacement of a sequence in a chromosome or a region of interest in cellular chromatin) are also provided. For example, a mutant genomic sequence can be replaced by a wild-type sequence, *e.g.*, for treatment of genetic disease or inherited disorders. In addition, a wild-type genomic sequence can be replaced by a mutant sequence, *e.g.*, to prevent function of an oncogene product or a product of a gene involved in an inappropriate inflammatory response. Furthermore, one or more alleles of a gene can be replaced by one or more different alleles (*e.g.*, biallelic targeted integration).

[0020] In the methods of the disclosure, one or more targeted nucleases as described herein create a single-stranded break in the target sequence (*e.g.*, cellular chromatin) at a predetermined site, and a "donor" polynucleotide, having homology to

the nucleotide sequence in the region of the break, can be introduced into the cell. The presence of the single-stranded break has been shown herein to facilitate integration of the donor sequence. The donor sequence may be physically integrated or, alternatively, the donor polynucleotide is used as a template for repair of the break via homologous recombination, resulting in the introduction of all or part of the nucleotide sequence as in the donor into the cellular chromatin. Thus, a first sequence in cellular chromatin can be altered and, in certain ~~embodiments~~examples, can be converted into a sequence present in a donor polynucleotide. Thus, the use of the terms "replace" or "replacement" can be understood to represent replacement of one nucleotide sequence by another, (*i.e.*, replacement of a sequence in the informational sense), and does not necessarily require physical or chemical replacement of one polynucleotide by another.

[0021] In addition, a method for replacement of a region of interest in cellular chromatin (*e.g.*, a genomic sequence) with a first nucleotide sequence is provided, the method comprising: (a) engineering a first zinc finger binding domain to bind to a second sequence in the region of interest; (b) providing a second zinc finger binding domain to bind to a third sequence in the region of interest; (c) expressing a first fusion protein in a cell, the first fusion protein comprising the first zinc finger binding domain and a first catalytically active cleavage half-domain; (d) expressing a second fusion protein in the cell, the second fusion protein comprising the second zinc finger binding domain and a second catalytically inactive cleavage half-domain; and (e) contacting the cell with a polynucleotide comprising the first nucleotide sequence; wherein the first fusion protein binds to the second sequence and the second fusion protein binds to the third sequence, thereby positioning the cleavage half-domains such that a single-stranded break is made in cellular chromatin in the region of interest and a nucleotide sequence in the region of interest is replaced with the first nucleotide sequence. In certain ~~embodiments~~examples, the single-stranded break in cellular chromatin is made in the region of interest at a site between the second and third sequences. The zinc finger nucleases may be provided to the cells as proteins and/or as one or more polynucleotides encoding said zinc finger nuclease(s). For example, two polynucleotides, each comprising sequences encoding one of the two fusion proteins, can be introduced into a cell. Alternatively, a single polynucleotide comprising sequences encoding both fusion proteins can be introduced into the cell.

[0022] In any of the methods described herein, additional pairs of zinc finger proteins can be used for additional double-stranded and/or single-stranded cleavage of additional target sites within the cell.

[0023] Thus, in one ~~embodiment~~example, a method for replacement of a region of interest in cellular chromatin (*e.g.*, a genomic sequence) with a first nucleotide sequence comprises: (a) engineering a first zinc finger binding domain to bind to a second sequence in the region of interest; (b) providing a second zinc finger binding domain to bind to a third sequence; and (c) contacting a cell with: (i) a first polynucleotide comprising the first nucleotide sequence; (ii) a second polynucleotide encoding a first fusion protein, the first fusion protein comprising the first zinc finger binding domain and a first catalytically active cleavage half-domain; and (iii) a third polynucleotide encoding a second fusion protein, the second fusion protein comprising the second zinc finger binding domain and a second catalytically inactive cleavage half-domain; wherein the first and second fusion proteins are expressed, the first fusion protein binds to the second sequence and the second fusion protein binds to the third sequence, thereby positioning the cleavage half-domains such that a single-stranded break is generated in cellular chromatin in the region of interest; and the region of interest is replaced with the first nucleotide sequence.

[0024] In the preferred ~~embodiments~~examples of methods for targeted recombination and/or replacement and/or alteration of a sequence in a region of interest in cellular chromatin, a chromosomal sequence is altered by homologous recombination with an exogenous "donor" nucleotide sequence. Such homologous recombination is stimulated by the presence of a single-stranded break in cellular chromatin, if sequences homologous to the region of the break are present. Notably, single-strand breaks in cellular chromatin do not stimulate cellular mechanisms of non-homologous end joining.

[0025] In any of the methods described herein, the first nucleotide sequence (the "donor sequence") can contain sequences that are homologous, but not identical, to genomic sequences in the region of interest, thereby stimulating homologous recombination to insert a non-identical sequence in the region of interest. Thus, in certain ~~embodiments~~examples, portions of the donor sequence that are homologous to sequences in the region of interest exhibit between about 80 to 99% (or any integer therebetween) sequence identity to the genomic sequence that is replaced. In other ~~embodiments~~examples, the homology between the donor and genomic sequence is higher than

99%, for example if only 1 nucleotide differs as between donor and genomic sequences of over 100 contiguous base pairs. In certain cases, a non-homologous portion of the donor sequence can contain sequences not present in the region of interest, such that new sequences are introduced into the region of interest. In these instances, the non-homologous sequence is generally flanked by sequences of 50-1,000 base pairs (or any integral value therebetween) or any number of base pairs greater than 1,000, for example sequences such as those found in artificial chromosomes, that are homologous or identical to sequences in the region of interest. In other ~~embodiments~~examples, the donor sequence is non-homologous to the first sequence, and is inserted into the genome by non-homologous recombination mechanisms.

[0026] Any of the methods described herein can be used for partial or complete inactivation of one or more target sequences in a cell by targeted integration of donor sequence that disrupts expression of the gene(s) of interest. Cell lines with partially or completely inactivated genes are also provided. Further still, the plant cell lines with partially or completely inactivated genes can be used to generate transgenic plants. Plant cell lines comprising nucleases as described herein, wherein expression of the nuclease is driven by an inducible promoter are also provided. Similarly, mammalian germ cells, such as oocytes, with partially or completely inactivated genes can be used to generate transgenic animals.

[0027] Furthermore, the methods of targeted integration as described herein can also be used to integrate one or more exogenous sequences. The exogenous nucleic acid sequence can comprise, for example, one or more genes or cDNA molecules, or any type of coding or noncoding sequence, as well as one or more control elements (*e.g.*, promoters). In addition, the exogenous nucleic acid sequence may produce one or more RNA molecules (*e.g.*, small hairpin RNAs (shRNAs), inhibitory RNAs (RNAis), microRNAs (miRNAs), etc.).

[0028] In any of the cells, cell lines and methods described herein, the cell or cell line can be a COS, CHO (*e.g.*, CHO-S, CHO-K1, CHO-DG44, CHO-DUXB11, CHO-DUKX, CHOK1SV), VERO, MDCK, WI38, V79, B14AF28-G3, BHK, HaK, NS0, SP2/0-Ag14, HeLa, HEK293 (*e.g.*, HEK293-F, HEK293-H, HEK293-T), PerC.6® (Crucell); EBx™ (Sigma-Aldrich Group), insect cells such as *Spodoptera fugiperda* (Sf), or fungal cells such as *Saccharomyces*, *Pichia* and *Schizosaccharomyces*.

[0029] In another ~~aspect~~example, the ~~invention~~disclosure provides kits comprising one or more nucleases (or polynucleotides encoding these nucleases) as described herein for carrying out the methods as described herein. The kits may optionally comprise reagents, buffers, cells, suitable containers and written instructions.

[0030] These and other aspects will be readily apparent to the skilled artisan in light of disclosure as a whole.

BRIEF DESCRIPTION OF THE DRAWINGS

[0031] **Figure 1** is a schematic depicting an exemplary zinc finger nuclease (ZFN) architecture that includes an inactivated cleavage domain. The ZFN pair cleaves only one DNA strand to form single-stranded breaks (SSBs), also referred to as nicks.

[0032] **Figure 2, panels A to C,** depict single-stranded cleavage of a double-stranded target using a pair of zinc finger nucleases where one nuclease includes a cleavage domain that has been mutated to inactivate cleavage activity of the nuclease domain.

[0033] Fig. 2A depicts analysis of double-stranded cleavage products obtained by cleavage of a target substrate with the indicated ZFN pairs. From left to right, lanes show: molecular weight ladder, control, a CCR5-targeted ZFN pair designated 8196zKK:8267EL, where KK and EL refer to engineered cleavage domains that form obligate heterodimers to generate double-stranded breaks (*see*, U.S. Patent Publication 2008/0131962), referred to as wildtype 8196zKK8267EL (WT) in this application; a CCR5-targeted ZFN pair designated 8196zKK:8267EL, where the 8267EL protein also contains a mutation at position 450 of the catalytic domain (D450N) of one cleavage half-domain that inactivates one cleavage half-domain of the ZFN pair; a CCR5-targeted ZFN pair designated 8196zKK:8267EL, where the 8267EL protein also contains a mutation at position 467 in the catalytic domain (D467A) that inactivates one cleavage half-domain of the ZFN pair; a second molecular weight ladder, a second control; a CCR5-targeted ZFN pair designated 8267RD:8196zDR, where RD and DR refer to engineered cleavage domains that form obligate heterodimers (*see*, U.S. Patent Publication 2008/0131962), referred to as wildtype 8267RD:8196zDR (WT) in this application; a CCR5-targeted ZFN pair designated 8267RD:8196zDR, where the 8196zDR protein also contains a mutation at position 450 of the catalytic domain (D450N) of one cleavage half-domain that inactivates one