(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 May 2011 (19.05.2011)

PCT

(10) International Publication Number
WO 2011/058555 A1

(51) **International Patent Classification**:
C12N 15/10 (2006.01) *C12N 15/113* (2010.01)

(21) **International Application Number**:
PCT/IL2010/000934

(22) **International Filing Date**:
11 November 2010 (11.11.2010)

(25) **Filing Language**: English

(26) **Publication Language**: English

(30) **Priority Data**:
61/272,854 12 November 2009 (12.11.2009) US

(71) **Applicant** *(for all designated States except US)*: **YEDA RESEARCH AND DEVELOPMENT CO. LTD.** [IL/IL]; at the Weizmann Institute of Science, P.O. Box 95, 76100 Rehovot (IL).

(72) **Inventors; and**
(75) **Inventors/Applicants** *(for US only)*: **SOEN, Yoav** [IL/IL]; 2 Neve Matz, The Weizmann Institute of Science, 76100 Rehovot (IL). **ELGART, Michael** [IL/IL]; c/o Yeda Research and Development Co. Ltd., at the Weizmann Institute of Science, P.O. Box 95, 76100 Rehovot (IL).

(74) **Agents**: **G.E EHRLICH (1995) LTD.** et al.; 11 Menachem Begin Street, 52521 Ramat Gan (IL).

(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, RO, RS, RU, 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, 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).

Declarations under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

Published:

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

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

— *with sequence listing part of description (Rule 5.2(a))*

WO 2011/058555 A1

(54) **Title**: A METHOD OF EDITING DNA IN A CELL AND CONSTRUCTS CAPABLE OF SAME

(57) **Abstract**: ABSTRACT A method of inserting a nucleic acid sequence of interest into genomic DNA in a cell is disclosed. The method comprises introducing at least one single-stranded DNA polynucleotide into the cell, wherein the at least one single-stranded DNA polynucleotide comprises a recombination site flanked by two targeting sequences, wherein the two targeting sequences hybridize under physiological conditions with a continuous sequence of the genomic DNA. This recombination site can further be recombined with a second polynucleotide, thus inserting the latter into the genomic DNA. The second polynucleotide can comprise various useful features such as genes, selection markers etc. Isolated polynucleotides used for same and further uses thereof are also disclosed.

A METHOD OF EDITING DNA IN A CELL AND CONSTRUCTS CAPABLE OF SAME

FIELD AND BACKGROUND OF THE INVENTION

The present invention, in some embodiments thereof, relates to a method of editing DNA in a cell and constructs capable of same.

Transgenic technologies have emerged as invaluable tools to manipulate the genome in biomedical, veterinary, and agricultural research. Since the early work by Brackett et al. [Proc Natl Acad Sci USA 68: 353–357, 1971] in which SV40 DNA was transfected into rabbit spermatozoa and the subsequent work by Jaenisch and colleagues which took advantage of wild-type retroviruses [Proc Natl Acad Sci USA 73: 1260–1264, 1976] to manipulate embryonic stem cells, there has been considerable interest in developing tools and approaches to modify the genome so as to generate transgenic animals and/or cell-lines. Presently, the most commonly accepted method to generate transgenic animals or cell lines remains the pronuclear injection of plasmid DNA. Unfortunately, this method is still hampered by a relatively low efficiency, particularly in species other than mice. Although these inefficiencies can be overcome using high-throughput screening for DNA integration, this type of approach becomes technically and economically very challenging.

The low efficiency of current methods motivates the development of other strategies to enhance the ability to modify and edit DNA in intact mammalian cells. One such technique involves the use of lentiviral vectors, which were initially developed for gene therapy to correct genetic disorders in somatic cells. There are a plethora of viral vector systems that have been developed for gene therapy applications, but the vast majority of these vectors are non-integrating, which precludes their usefulness for transgenesis. The best characterized integrating viral vectors originate from the *retroviridae* family, the members of which are ideal for the genetic manipulation of mammalian cells because of their intrinsic ability to integrate into genomic DNA. The lentivirus genome derived from immunodeficiency viruses, such as human immunodeficiency virus-1 (HIV-1), has been split into multiple fragments to minimize the potential formation of replication-competent viruses. Such vectors target heterogeneous cell populations and lead to semi-random insertion of transgenes into

cellular chromosomes. Somatic gene transfer thus results in a complex mosaic with coexistence of initially millions of cell clones that are distinguished not only by the site and number of transgene insertions, but potentially also by their intrinsic developmental potential. Moreover, depending on the vector type, the transgene sequence including its epigenetic modulation and the level of transgene expression may differ between individual gene-modified cells. Owing to the multifaceted nature of this selection process, presently available approaches do not allow reliable predictions of potential consequences of insertional mutagenesis.

Recently a new method for improving the recombination frequency in mammalian cells has been developed. It involves design of DNA-nicking enzymes fused to Zinc-Finger DNA binding proteins that recognize specific genomic sequences [Science, 2003, 300 (5620), 764]. Zinc finger artificial transcription factors (ZF-ATFs), consisting of a polydactyl zinc finger (PZF) DNA-binding domain linked to a protein domain that either activates or represses gene expression near the PZF binding site, have received considerable attention. ZF-ATFs are generally constructed using Cys2His2 ZF domains, the most thoroughly characterized ZF moieties to date. Apart from being small, just 30 amino acids, Cys2His2 ZF domains typically bind three contiguous DNA bases per ZF and are highly modular. These features allow construction of more complex PZF fusion proteins, which in principle can recognize unique sites within a complex genome. For GNN, ANN and recently also CNN target sites, where N represents any of the four bases, the most optimal zinc finger designs have been characterized, making ZF-ATF construction possible for almost any target site. More recently Zinc-Finger-Recombinase chimeras were also demonstrated. While being efficient in cases where the targeting is specific, the design of the above proteins remains very challenging and is currently expensive and lengthy process of trial-and-error with no guarantee of a successful outcome.

An additional experimental strategy was reported for repair of specific mutations as an alternative to the gene augmentation strategy currently employed in most somatic gene therapy approaches. The technique relies on the use of chimeric RNA/DNA oligonucleotides (RDOs) or modified single-stranded oligodeoxyribonucleotides (ssODNs) designed to correct single-base mutations by generation of a mismatch between the RDO and the target point mutation. The mismatched base pair was

anticipated to cause helical distortion recognized by the genetic repair system, thus converting the targeted base in the genome using the information provided by the RDO template. By this approach mutations in a multitude of genes have been corrected, thereby providing a proof of principle for the method.

5 The RDOs applied in the first gene correction experiments were all comprised of DNA and RNA in a complementary hairpin configuration. These 68-mers contain a "mutator" region of 5 DNA nucleotides complementary to the target sequence, except for the targeted point mutation flanked by two blocks of 2 prime-O-methyl RNA residues of ten nucleotides each, also complementary to the target locus.

10 The 2'-O-methylation serves as protection from nucleases present in the cell nuclei. A nick is embedded in the duplex by extending the 3' end of the oligonucleotide to form a 5-bp-long GC clamp bridged by four T residues (a T-loop) at the end. The nick is provided as a swivel to permit topological intertwining of the RDO into the target double stranded DNA. It is recessed within the duplex to enhance thermal 15 stability and nuclease resistance.

To further examine the function of individual sections of the RDO, single-stranded 25-mer all-DNA, all-RNA, or RNA/DNA oligonucleotides were designed to correct a mutant kanamycin gene using mammalian and plant cell-free extracts [Igoucheva O et al., 2001, Gene Ther 8:391-399]. Whereas all-RNA and RNA/DNA 20 oligonucleotides exhibited low repair potential, the all-DNA strand was found to be three times more efficient than the traditional RDO design confirming previous findings that the all-DNA strand directs the actual repair. Using a similar approach, gene conversion of a mutant episomal and chromosomal bgr-galactosidase gene with ssODNs of varying length and polarity was accomplished with the same efficiency as 25 RDOs [Igoucheva O et al., *supra*]. An increase in correction frequency was observed with the homology region extended up to 45 nucleotides. It was also shown that the conversion was 1000-fold higher when using an antisense oligonucleotide than a sense oligonucleotide [Igoucheva O et al., *supra*]. Additional observations indicate that DNA with an open configuration due to transcriptional activity is more amenable to complex 30 formation with an ssODN and that the presence of RNA polymerase and associated protein factors causes steric hindrance on the transcribed strand in the gene repair process. These results therefore encourage targeting of transcriptional active genes with

the sense (non-transcribed) strand being the target strand. Finally, correction of the episomal mutant lacZ gene has been shown to occur at a higher frequency (1%) than chromosomal correction (0.1%) suggesting also a limited accessibility of chromosomal DNA as compared to episomally expressed genes [Igoucheva O et al., *supra*].

5 International Patent Application WO 2004/015117 teaches insertion of up to 12 base pair (bp) fragments using ssDNA oligonucleotides by disruption of one or more MMR genes e.g. MSH2, MLH1, PMS1 and PMS2.

U.S. Patent Application No. 20040014057 teaches that modified ssDNA oligonucleotides can be used to modify genomic DNA. The modified ssDNA 10 oligonucleotides comprise nuclease resistant residues, preferably at or near the termini of the oligonucleotides.

U.S. Patent Application No. 20070122822 teaches compositions and methods for enhancing oligonucleotide-directed nucleic acid sequence alteration. These methods and compositions involve cells with altered levels or activities of a protein from the 15 RAD52 epistasis group.

International Patent Application WO 1990/014092 teaches single-stranded oligonucleotides for site-directed modification in mammalian cells. International Patent Application WO 1990/014092 teaches that the oligonucleotides comprise sequences which have substantial homology with the site of insertion, but may differ by one or 20 more bases, usually not more than about 10.

International Patent Application WO 2003/076574 teaches single-stranded oligonucleotides for site-directed modification in plant cells, whereby the oligonucleotides are limited in that they comprise at least one cationic phosphoramidate internucleoside linkage.

25

SUMMARY OF THE INVENTION

According to an aspect of some embodiments of the present invention there is provided an isolated polynucleotide comprising a single-stranded DNA polynucleotide encoding a regulatory site, the regulatory site being longer than 15 bases and being 30 flanked by two targeting sequences, wherein the two targeting sequences hybridize with a continuous sequence of mammalian genomic DNA and wherein each of the targeting sequence is at least 10 bases.

According to an aspect of some embodiments of the present invention there is provided a method of inserting a nucleic acid sequence of interest into genomic DNA in a cell, the method comprising introducing at least one single-stranded DNA polynucleotide into the cell, wherein the at least one single-stranded DNA polynucleotide comprises the nucleic acid sequence of interest flanked by two targeting sequences, wherein the two targeting sequences hybridize under physiological conditions with a continuous sequence of the genomic DNA and wherein the nucleic acid sequence of interest is longer than 15 bases, thereby inserting the nucleic acid sequence into genomic DNA of a cell.

According to an aspect of some embodiments of the present invention there is provided a method of inserting a nucleic acid of interest into genomic DNA of a population of cells, the method comprising:

(a) introducing the isolated polynucleotide of the present invention into the cell population; and

(b) introducing:

(i) a polynucleotide which encodes the recombination site and comprises the nucleic acid of interest; and

(ii) a polynucleotide encoding a recombinase enzyme into the cell population, thereby inserting the nucleic acid of interest into genomic DNA of the population of cells.

According to an aspect of some embodiments of the present invention there is provided a method of deleting a target sequence from genomic DNA in a cell, the method comprising:

(a) introducing the isolated polynucleotide of the present invention into a plurality of cells;

(b) introducing a polynucleotide encoding a recombinase enzyme into the plurality of cells, thereby deleting a target sequence from genomic DNA of a cell.

According to an aspect of some embodiments of the present invention there is provided a method of treating a trinucleotide repeat-associated disease in a subject in need thereof, the method comprising:

(a) introducing a single-stranded DNA polynucleotide into cells of the subject, wherein the single-stranded DNA polynucleotide encodes a recombinase site,

the site being flanked by two targeting sequences, wherein the two targeting sequences hybridize under physiological conditions with a continuous sequence of genomic DNA of the cells and wherein the two targeting sequences each hybridize with a trinucleotide repeating motif of the trinucleotide repeat associated disease; and

5 (b) introducing a polynucleotide encoding a recombinase enzyme into the cells, thereby treating Huntington's chorea in a subject.

According to some embodiments of the invention, the regulatory site comprises a recombination site.

According to some embodiments of the invention, the isolated polynucleotide is 10 attached to a targeting moiety.

According to some embodiments of the invention, the recombination site is selected from the group consisting of a *hix* site, an *att* site, a *lox* site and a *frt* site.

According to some embodiments of the invention, the lox site comprises a sequence as set forth in SEQ ID NOs: 9-15.

15 According to some embodiments of the invention, each of the two targeting sequences is 20 bases.

According to some embodiments of the invention, at least one of the two targeting sequences comprises a modified base.

According to some embodiments of the invention, the continuous sequence of 20 genomic DNA repeats at least twenty times in a single chromosome of the cell.

According to some embodiments of the invention, the at least one single-stranded DNA polynucleotide comprises at least two identical single-stranded DNA polynucleotides.

According to some embodiments of the invention, the at least one single-25 stranded DNA polynucleotide comprises two non-identical single-stranded DNA polynucleotides.

According to some embodiments of the invention, the two non-identical single-stranded DNA polynucleotides comprise non-identical nucleic acid sequences of interest.

30 According to some embodiments of the invention, the single-stranded DNA polynucleotide comprises non-identical targeting sequences.

According to some embodiments of the invention, the genomic DNA is selected from the group consisting of nuclear DNA, mitochondrial DNA and chloroplast DNA.

According to some embodiments of the invention, the nucleic acid sequence of interest comprises at least one stem-loop structure.

5 According to some embodiments of the invention, the nucleic acid sequence of interest comprises less than 50 bases.

According to some embodiments of the invention, the nucleic acid sequence of interest encodes a recombination site.

According to some embodiments of the invention, at least one of the two 10 targeting sequences is greater than 10 bases.

According to some embodiments of the invention, at least one of the two targeting sequences is 20 bases.

According to some embodiments of the invention, at least one of the two targeting sequences comprises a modified base.

15 According to some embodiments of the invention, the cell is a eukaryotic cell.

According to some embodiments of the invention, the cell is a mammalian cell.

According to some embodiments of the invention, the cell is a plant cell.

According to some embodiments of the invention, the cell is an intact cell.

According to some embodiments of the invention, the mammalian cell is a 20 human cell.

According to some embodiments of the invention, the cell is a stem cell.

According to some embodiments of the invention, the stem cell is an embryonic stem cell.

According to some embodiments of the invention, the cell is a diseased cell.

25 According to some embodiments of the invention, the method is effected in vivo.

According to some embodiments of the invention, the method is effected ex vivo.

According to some embodiments of the invention, the single stranded DNA 30 polynucleotide is associated with a cell-targeting moiety.

According to some embodiments of the invention, step (a) is affected prior to step (b).

According to some embodiments of the invention, step (a) is effected concomitant to step (b).

According to some embodiments of the invention, the genomic DNA is selected from the group consisting of nuclear DNA, mitochondrial DNA and chloroplast DNA.

5 According to some embodiments of the invention, the recombinase enzyme is selected from the group consisting of Hin invertase, λ integrase, Cre and Flp.

According to some embodiments of the invention, the nucleic acid of interest encodes at least one detectable moiety.

According to some embodiments of the invention, the method further comprises 10 selecting cells which comprise the nucleic acid of interest.

According to some embodiments of the invention, the method further comprises selecting at least one of the plurality of cells, prior to step (b) wherein the at least one of the plurality of cells comprises at least two recombinase sites integrated into a single chromosome thereof.

15 According to some embodiments of the invention, the trinucleotide repeat-associated disease is Huntington's chorea.

According to some embodiments of the invention, the method further comprises selecting at least one of the cells, prior to step (b) wherein the at least one of the plurality of cells comprises at least two recombinase sites integrated into a single chromosome 20 thereof.

According to some embodiments of the invention, the cells comprise neuronal cells.

According to some embodiments of the invention, the cells comprise stem cells.

25 Unless otherwise defined, all technical and/or scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which the invention pertains. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of embodiments of the invention, exemplary methods and/or materials are described below. In case of conflict, the patent 30 specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and are not intended to be necessarily limiting.

BRIEF DESCRIPTION OF THE DRAWINGS

Some embodiments of the invention are herein described, by way of example only, with reference to the accompanying drawings. With specific reference now to the drawings in detail, it is stressed that the particulars shown are by way of example and for purposes of illustrative discussion of embodiments of the invention. In this regard, the description taken with the drawings makes apparent to those skilled in the art how embodiments of the invention may be practiced.

In the drawings:

FIGs. 1A-C are schemes and photographs illustrating fragment incorporation into a 40bp sequence (which was repeated 22 times) in HeLa cells according to an embodiment of the present invention. (A) Schematic representation of the genomic regions containing multiple repeats, the primers used in this study and the fragment containing LoxP (SEQ ID NO: 1) which was incorporated. (B) Schematic representation of the expected PCR products (C) Experimental PCR products following transfection with the fragment that contains loxP (SEQ ID NO: 1), or GFP plasmid into genomic DNA of HeLa cells.

FIG. 2A is a photograph of the experimental PCR products following transfection with the fragment that contains loxP (SEQ ID NO: 1), or GFP plasmid into genomic DNA of HeLa cells (identical to Figure 1C).

FIG. 2B is a sequence (SEQ ID NO: 2) of the PCR product following transfection with the fragment (SEQ ID NO: 1). The band was excised from gel, ligated to pGEM-t-easy, grown in bacteria and sequenced. LoxP incorporation sites are highlighted in green.

FIGs. 3A-D are photographs illustrating fragment (SEQ ID NO: 1) in different cell lines and under different conditions. "N" stands for "No Primers", "O" for OFI. (A) Standard procedure as in Figure 2A. (B) Same as (Figure 2A) after 6 cell divisions. (C) Same as (Figure 2A) in H1 line of human ESCs. (D) Same as (Figure 2A) after irradiation (60Rad) of the HeLa cells prior to the transfection of the fragment.

FIG. 4 are photographs and schematic representations illustrating the dependence of the efficiency of fragment incorporation on the insert. Fragments with different inserts but same flanking region (targeted to the 22 repeat loci) were transfected to HeLa cells. Shown here are PCR results for the incorporation detection

(top), the secondary structure and sequence, respectively SEQ ID NOs: 3, 4, 1, and 5) of each of the constructs (middle) and its free energy (bottom).

FIGs. 5A-B are photographs illustrating the contribution of genomic location on the efficiency of fragment incorporation. (A) fragment with a single loxP insert was incorporated into two different 22 repeats of a 40bp sequence: rpt22-1(SEQ ID NO: 1) and rpt22-2 (SEQ ID NO: 6). (B) fragment with a single lox66 insert was incorporated into 3 different genomic repeats: rpt22 (22 repeats, SEQ ID NO: 4), rpt 5188 (5188 repeats, SEQ ID NO: 7) and rpt5192 (5192 repeats, SEQ ID NO: 8).

FIG. 6A is a schematic representation of a vector plasmid according to one embodiment of the present invention.

FIG. 6B is a photograph of FACS positive cells expressing the plasmid 38 days following transfection.

FIG. 7 is a schematic representation of the incorporation of the fragments of embodiments of the present invention. (1) Dividing cells are transfected with the fragment. Cells are allowed to grow for over two passages to verify stable incorporation into the genome. (2) The same cells are then transfected with a plasmid carrying the Cre recombinase and a second plasmid carrying the gene of interest, loxP site and selection marker. The subpopulation of cells that by now contains the intragenic loxP site will integrate the second Plasmid (3) and start expressing the selection marker;

FIGs. 8A-C are schematic representations of constructs designed to identify cells that underwent genomic incorporation of the plasmid. (A) The tet-off system. tetR protein binds to tetO sequences only upon introduction of Dox, thus leading to the repression of the downstream gene. (B) Schematics of a modified system expressing the selection marker only upon its integration into the genome. A mutant loxRE site is inserted between an upstream tetO sequences and a downstream selection marker (eGFP). In the presence of tetR (introduced on a different plasmid) and Dox, the expression of eGFP from the plasmid is prevented by the repressor. (C) Integration of the plasmid into the genome displaces the repressor-bound tetO sequence (and two mutated loxP sites, LE and RE) downstream to the selection marker, which is now driven by fully active, CMV promoter.

FIGs. 9A-B are schematic representations illustrating the mechanism of Cre/loxP-mediated plasmid integration into the genome. (A) The standard

recombination model, exhibiting equilibrium between integration and excision reactions with a strong preference for excision. (B) Integration with mutant lox sites. Integration creates two different lox sites (WT and full-mutant) which can rarely be excised, thus tilting the equilibrium almost solely towards integration.

5 FIGs. 10A-B are schematic representations for identifying and analyzing fragment incorporation into hES cells. (A) The vector plasmid carries two different fluorophore genes – GFP and TFP – downstream of a constitutive PGK promoter. The GFP gene is separated from its promoter by a lox site. Successful recombination into genomic lox site would therefore eliminate the GFP promoter and will turn off the 10 expression of GFP. (B) The vector plasmid carries two different fluorophore genes – GFP and TFP – downstream of a constitutive PGK promoter. The TFP gene is separated from its promoter by a lox site. Successful recombination into genomic lox site would therefore eliminate the TFP promoter and will turn off the expression of TFP. Genetically modified cells will be GFP+/TFP- while unmodified cells will express both 15 GFP and TFP.

FIGs. 11A-G are schematic representations and images illustrating the transformation of cells with exemplary fragments according to Figure 10B (A) Schematic representation of the processes leading to transformation of cells. First a cell incorporates the fragment, and after the recombination of the Vector Plasmid, only GFP 20 is expressed (B-D) hES cells on matrigel 24h after the p7.2LTG Vector Plasmid and CRE recombinase co-transfection. Both GFP and TFP are present in all transfected cells (E-G) Same cells after 3 more days. Most of the cells lost fluorescence. Some are still both GFP and TFP positive (blue circle) which indicate that in these cells the vector plasmid is episomal. Others are only GFP positive (red circles) which indicates 25 successful genomic integration of the Vector Plasmid.

FIGs. 12A-B are schematic representations illustrating Random intergenic recombinations between wild-type loxP sites. Upon introduction of multiple genomic lox sites, In the presence of Cre recombinase, any number of recombinations can occur leading to unwanted effects such as deletions of part of chromosomes and combining 30 two or more chromosomes into one. Lines denote chromosomes and triangles denote lox sites. (A) The lox sites after fragment introductions but before the introduction of

Cre. Blue arrows indicate possible lox recombination pairs (B) Some of the possible ill effects following Cre introduction.

FIGs. 13A-E are schematic representations of mutant lox sites that only allow plasmid integration. (A) Wild type lox sites that allow any recombination. (B) Insertion only lox mutants that allow only plasmid integration but not excision. The mutation is located at the flanks of the lox sequence. (C) lox Mutants that can not recombine with themselves, but are able to undergo recombination with another lox site. The mutation is located in the center (spacer) part of the lox sequence. (D) Combined mutant that can not recombine with itself, and is not excitable (E) More detailed representation and sequence of the double-mutant lox site from (D).

FIGs. 14A-F are schematic representations and photographs illustrating the effectiveness of the mammalian genome targeting approach of embodiments of the present invention. (Figures 14A-B) Schematics of targeting repeated (14A) or unique genomic sequences (14B), either within genes or outside genes. (Figure 14C) Schematics of detection of exemplary fragment incorporation in (14B) using PCR. (14D) Successful incorporation of exemplary fragments into the Pdx1 intron of Hela cells, as determined by PCR bands. (14E) Sequence-based verification of successful targeting of the Pdx1 locus – SEQ ID NO: 27. (14F) Schematics of a 5-day protocol resulting in fragments incorporation (day 1), its exchanges with an arbitrary plasmid (day 3), and selection of modified cells by antibiotic or by FACS.

FIGs. 15A-C are photographs illustrating fragment-dependent survival of Hela cells under G418 selection. Hela cells were targeted as shown in Figure 14F and were exposed for 1 week to 1mg/ml of G418. Resulting colonies were stained with methylene blue. Both the number and size of surviving colonies were clearly dependent on the dosage of fragments in the first step of the incorporation, indicating that the subsequent incorporation of the neo plasmid (2nd step) was indeed mediated by successful incorporation of the fragment in the preceding step. Note that all three cases were transfected with the same amount of plasmid ($3\mu g/10^6$ cells) carrying *neo*.

FIGs. 16A-B are images of a colony of human iPS cells that were derived from a Huntington's patient and were targeted with OFI directed against the 180 tri-nucleotide repeats in the Huntington gene. The incorporated fragment was then exchanged with a plasmid carrying GFP. Green colonies were FACS-sorted and expanded using standard

techniques of growing human embryonic stem cells. Shown are a phase contrast (Figure 16A), and a fluorescent image (Figure 16B) of a single colony.

FIGs.17A-B illustrate OFI-mediated disruption of CXCR4 and protection from HIV-1 infection in human cells. (A) The HIV virus surface receptor gp120 binds CD4 and CXCR4 in order to infect cells. Once Infected the HIV inserts its genome into the target cell genome at a random position. The cell then continues to produce virions. A mixture of OFIs can be used to disrupt both the integrated viral sequence and the endogenous co-receptor CXCR4. This allows establishing a virus-free population of T-cells or T-cell precursors which are resistant to HIV infection due to their lacking of the CXCR4 co-receptor. (B) Example of targeting of the CXCR4 gene using a single OFI. The depicted sequence (SEQ ID NO: 36) is part of the second exon of CXCR4 gene with OFI-mediated lox66 incorporation. The incorporation causes frameshift rendering the gene inactive. Additionally the lox66 sequence can be further used to recombine a sequence of interest, such as selection marker, into this genomic location. Lane 1 of the gel is a ladder. Lane 2 is a PCR of the OFI-treated cells using genomic primers to verify genomic DNA integrity. Lane 3 is a PCR of the OFI-treated cells with genomic forward primer and lox66 reverse primer which works only if the lox66 sequence was successfully incorporated into the genomic locus. Lane 4 is a PCR performed on non-treated, wild-type genomic DNA with genomic forward primer and lox66 reverse primer. Lane 3 produces a band of the correct size while the negative control Lane4 is empty as expected.

DESCRIPTION OF SPECIFIC EMBODIMENTS OF THE INVENTION

The present invention, in some embodiments thereof, relates to a method of editing DNA in a cell and constructs capable of same.

Before explaining at least one embodiment of the invention in detail, it is to be understood that the invention is not necessarily limited in its application to the details set forth in the following description or exemplified by the Examples. The invention is capable of other embodiments or of being practiced or carried out in various ways.

Chimeric RNA/DNA oligonucleotides (RDOs) or modified single-stranded oligodeoxyribonucleotides (ssODNs) have been designed to correct single-base mutations by generation of a mismatch between the RDO and the target point mutation.

The present inventors have devised ssDNA inserts designed to achieve a dramatic improvement in the ability to modify the mammalian genome. Specifically, the present inventors designed oligonucleotides such that a 34 base pair LoxP fragment was stably inserted into the cell genome at a targeted site (Figures 2A-B, 3A-D and 14B-E). Subsequent transfection with a plasmid encoding a CRE recombinase enabled the exchange of the initial insert with a (co-transfected) DNA sequence of choice (Figures 6B, 15A-C).

Thus, according to one aspect of the present invention, a method is provided for targeted insertion of a nucleic acid sequence of interest into DNA in a cell. The method comprises introduction of a single-stranded DNA polynucleotide into the cell, wherein the single-stranded DNA polynucleotide comprises an exogenous nucleic acid sequence of interest flanked by two targeting sequences, wherein the two targeting sequences hybridize under physiological conditions with a continuous sequence of the genomic DNA and wherein the exogenous nucleic acid sequence of interest is longer than 15 bases, thereby inserting the nucleic acid sequence into genomic DNA of a cell.

Cells into which the nucleic acid sequence of interest may be inserted are typically eukaryotic cells, including for example mammalian cells (e.g. human) and plant cells.

According to one embodiment, the mammalian cells are stem cells, such as embryonic stem cells or mesenchymal stem cells. In addition, the present invention contemplates insertion of the nucleic acid sequence of interest into induced pluripotent cells (iPS) cells. Such cells are further described herein below.

It will be appreciated that the cells may be isolated e.g. part of a cell line or primary cells. Alternatively, the cells may be comprised in a tissue, organ or organism.

According to one embodiment, the cells are diseased cells.

According to another embodiment, the cells are proliferating.

According to another embodiment, the cells are intact cells (i.e. not part of a cell extract).

The genomic DNA into which the nucleic acid of interest is inserted includes nuclear DNA and organellar DNA, such as mitochondrial DNA and chloroplast DNA.

It will be appreciated that the nucleic acid of interest may be inserted into any part of the DNA – e.g. in the regulatory portion of a gene, in the promoter region, in the

3' or 5' non- translated region of a gene, in the coding region of a gene, in the junction between an intron and an exon, or in a transformed region, i.e. non-native DNA. The insertion site may be a unique sequence of the genome or may be repeated multiple times. In addition, the nucleic acid of interest may be inserted into the genome in a 5 region which is devoid of genes, regardless of the chromatin state. Specifically, the method applies to both open and closed chromatin regions.

The present invention also contemplates inserting more than one type of single stranded polynucleotide into the cell at the same time. Thus, for example, the present invention contemplates inserting an identical sequence of interest into the genome of a 10 cell using non-identical targeting sequences (see Example 6 herein below). Additionally, or alternatively, the present invention contemplates inserting non-identical sequences of interest into the genome using identical targeting sequences.

Whilst reducing the present invention to practice, the present inventors showed a strong, positive dependency of the efficiency of incorporation on the number of target 15 sites (Figure 5B).

Thus, according to one embodiment, the nucleic acid of interest is inserted into a region of genomic DNA which repeats at least 5 times in a single chromosome. According to another embodiment, the nucleic acid of interest is inserted into a region of genomic DNA which repeats at least 10 times in a single chromosome. According to 20 yet another embodiment, the nucleic acid of interest is inserted into a region of genomic DNA which repeats at least 20 times in a single chromosome.

As mentioned, the method is affected by introducing a single-stranded DNA polynucleotide comprising the nucleic acid sequence of interest flanked by two targeting sequences into the cell, wherein the two targeting sequences hybridize under 25 physiological conditions with a continuous sequence of the genomic DNA and wherein the nucleic acid sequence of interest is longer than 15 bases.

Polynucleotides designed according to the teachings of the present invention can be generated according to any polynucleotide/oligonucleotide synthesis method known in the art such as enzymatic synthesis or solid phase synthesis. Equipment and reagents 30 for executing solid-phase synthesis are commercially available from, for example, Applied Biosystems. Any other means for such synthesis may also be employed; the actual synthesis of the oligonucleotides is well within the capabilities of one skilled in

the art and can be accomplished via established methodologies as detailed in, for example, “Molecular Cloning: A laboratory Manual" Sambrook et al., (1989); "Current Protocols in Molecular Biology" Volumes I-III Ausubel, R. M., ed. (1994); Ausubel et al., "Current Protocols in Molecular Biology", John Wiley and Sons, Baltimore, Maryland (1989); Perbal, "A Practical Guide to Molecular Cloning", John Wiley & Sons, New York (1988) and "Oligonucleotide Synthesis" Gait, M. J., ed. (1984) utilizing solid phase chemistry, e.g. cyanoethyl phosphoramidite followed by deprotection, desalting and purification by for example, an automated trityl-on method or HPLC.

The single-stranded targeting DNA polynucleotides are typically greater than 60 nucleotides in length (combined length). According to one embodiment, the single-stranded targeting DNA polynucleotide is about 80 nucleotides in length. According to one embodiment, the single-stranded targeting DNA polynucleotide is about 90 nucleotides in length. According to one embodiment, the single-stranded targeting DNA polynucleotide is about 100 nucleotides in length. According to yet another embodiment, the single-stranded targeting DNA polynucleotide is between about 60 nucleotides in length and 400 nucleotides in length.

According to an embodiment of this aspect of the present invention, the nucleic acid sequence of interest is at least 10 nucleotides in length, at least 15 nucleotides in length, at least 20 nucleotides in length, at least 25 nucleotides in length or at least 30 nucleotides in length.

According to another embodiment, the nucleic acid sequence of interest is no longer than 50 nucleotides in length.

According to one embodiment, the nucleic acid sequence of interest comprises at least one stem loop structure. Analysis of the sequence of a particular nucleic acid enables the determination of its structure. Determination of structure of a particular nucleic acid can be aided by the use of specific computer programs, such as the Mfold web server for nucleic acid folding and hybridization prediction (Nucleic Acids Res. 31 (13), 3406-15, 2003) and Vienna RNA webservers.

According to still another embodiment, the nucleic acid sequence of interest is selected such that it avoids identification by the DNA repair machinery, which might

otherwise operate to excise and/or degrade the inserted sequence. This protection may increase the efficiency of insertion.

According to one embodiment, the nucleic acid sequence of interest encodes a regulatory site. Examples of such regulatory sites include, but are not limited to recombination sites, lac Operator sites (e.g. SEQ ID NO: 16), tet Operator sites (e.g. SEQ ID NO: 17), a UAS site (e.g. SEQ ID NO: 18), and any other type of binding site for any type of gene or another molecule (e.g. transcription factor, chromatin modifying enzyme, co-repressor, co-activator, RNA molecule, DNA molecule, etc). Such sequences could then allow targeting of these incorporated exogenic sites by highly specific (exogenous and/or endogenous) DNA-binding proteins (and/or other types of molecules) from either the same or different organisms with added function such as DNA methylation, gene expression repression, direct DNA labeling etc.

Examples of nucleic acid sequences of interest that comprise at least one stem loop structure are recombination sites and mutations thereof, including, but not limited to hix sites (e.g. SEQ ID NO: 19,20), att sites (e.g. SEQ ID NO: 21, 22), lox sites (e.g. SEQ ID NOs: 9-15) and frt sites (e.g. SEQ ID NO: 23).

The two targeting sequences on either side of the nucleic acid sequence of interest are designed such that they are capable of hybridizing under physiological conditions with a continuous sequence of the genomic DNA. Preferably, the targeting sequences are designed such that they hybridize with the lagging strand of the genomic DNA (i.e. the targeting sequences comprise antisense oligonucleotides)

As used herein, the phrase “capable of hybridizing” refers to the ability to base pair with one strand of the genomic DNA. “Physiological conditions” refer to the conditions present in cells, tissue or a whole organism or body. Preferably, the physiological conditions used by the present invention include a temperature between 34-40 °C, more preferably, a temperature between 35-38 °C, more preferably, a temperature between 36 and 37.5 °C, most preferably, a temperature between 37 to 37.5 °C; salt concentrations (e.g., sodium chloride NaCl) between 0.8-1 %, more preferably, about 0.9 %; and/or pH values in the range of 6.5-8, more preferably, 6.5-7.5, most preferably, pH of 7-7.5.

According to one embodiment, at least one of the two targeting sequences are greater than 10 nucleotides in length - for example about 20 nucleotides in length.

According to another embodiment, each of the two targeting sequences is greater than 10 nucleotides in length - for example about 20 nucleotides in length.

According to yet another embodiment, at least one of the two targeting sequences comprises a modified base or linkage. The modification may be at the ends of the single-stranded targeting DNA polynucleotide, however the present invention also contemplates modification along a portion or even the entire length of the single-stranded targeting DNA polynucleotide.

For example, U.S. Patent Application No. 20040014057, incorporated herein by reference teaches that modified ssDNA oligonucleotides can be used to modify genomic DNA. The modified ssDNA oligonucleotides comprise nuclease resistant residues (e.g. 2'-O-methyl analogs or locked nucleotide analogs), preferably at or near the termini of the oligonucleotides.

In addition, WO 2003/076574, incorporated herein by reference, teaches single-stranded oligonucleotides for site-directed modification in plant cells, whereby the oligonucleotides are modified to comprise at least one cationic phosphoramidate internucleoside linkage. Methods for preparing cationic oligonucleotides having such linkages are found in Dagle & Weeks, Nucl. Acids Res. 24: 2143-2149,1996 ; Dagle et al., Nucl. Acids Res 28: 2153-2157,2000 ; US 6,274, 313 and US 5,734,040 the disclosures of which are incorporated herein by reference. The cationic phosphoramidate may be incorporated during oligonucleotide synthesis via oxidative amidation. In one embodiment the cationic phosphoramidate have both a primary and a tertiary amine. Positively charged internucleoside linkages can be generated by using a diethylethylenediamine such as N, N-diethylethylenediamine (DEED) and methoxyehtylamine phosphoramidates.

In addition, the polynucleotides of the present invention may comprise heterocylic nucleosides consisting of purines and the pyrimidines bases, bonded in a 3' to 5' phosphodiester linkage.

Specific examples of polynucleotides which may be useful according to this aspect of the present invention include oligonucleotides containing modified backbones or non-natural internucleoside linkages. Oligonucleotides having modified backbones include those that retain a phosphorus atom in the backbone, as disclosed in U.S. Pat. NOs: 4,469,863; 4,476,301; 5,023,243; 5,177,196; 5,188,897; 5,264,423; 5,276,019;

5,278,302; 5,286,717; 5,321,131; 5,399,676; 5,405,939; 5,453,496; 5,455,233; 5,466, 677; 5,476,925; 5,519,126; 5,536,821; 5,541,306; 5,550,111; 5,563,253; 5,571,799; 5,587,361; and 5,625,050.

Additional modified oligonucleotide backbones include, for example, phosphorothioates, chiral phosphorothioates, phosphorodithioates, phosphotriesters, aminoalkyl phosphotriesters, methyl and other alkyl phosphonates including 3'-alkylene phosphonates and chiral phosphonates, phosphinates, phosphoramidates including 3'-amino phosphoramidate and aminoalkylphosphoramidates, thionophosphoramidates, thionoalkylphosphonates, thionoalkylphosphotriesters, and boranophosphates having normal 3'-5' linkages, 2'-5' linked analogs of these, and those having inverted polarity wherein the adjacent pairs of nucleoside units are linked 3'-5' to 5'-3' or 2'-5' to 5'-2'. Various salts, mixed salts and free acid forms can also be used.

Alternatively, modified oligonucleotide backbones that do not include a phosphorus atom therein have backbones that are formed by short chain alkyl or cycloalkyl internucleoside linkages, mixed heteroatom and alkyl or cycloalkyl internucleoside linkages, or one or more short chain heteroatomic or heterocyclic internucleoside linkages. These include those having morpholino linkages (formed in part from the sugar portion of a nucleoside); siloxane backbones; sulfide, sulfoxide and sulfone backbones; formacetyl and thioformacetyl backbones; methylene formacetyl and thioformacetyl backbones; alkene containing backbones; sulfamate backbones; methyleneimino and methylenehydrazino backbones; sulfonate and sulfonamide backbones; amide backbones; and others having mixed N, O, S and CH2 component parts, as disclosed in U.S. Pat. Nos. 5,034,506; 5,166,315; 5,185,444; 5,214,134; 5,216,141; 5,235,033; 5,264,562; 5,264,564; 5,405,938; 5,434,257; 5,466,677; 5,470,967; 5,489,677; 5,541,307; 5,561,225; 5,596,086; 5,602,240; 5,610,289; 5,602,240; 5,608,046; 5,610,289; 5,618,704; 5,623, 070; 5,663,312; 5,633,360; 5,677,437; and 5,677,439.

Other oligonucleotides which can be used according to the present invention, are those modified in both sugar and the internucleoside linkage, i.e., the backbone, of the nucleotide units are replaced with novel groups. The base units are maintained for complementation with the appropriate polynucleotide target. An example for such an oligonucleotide mimetic, includes peptide nucleic acid (PNA). A PNA oligonucleotide

refers to an oligonucleotide where the sugar-backbone is replaced with an amide containing backbone, in particular an aminoethylglycine backbone. The bases are retained and are bound directly or indirectly to aza nitrogen atoms of the amide portion of the backbone. United States patents that teach the preparation of PNA compounds include, but are not limited to, U.S. Pat. Nos. 5,539,082; 5,714,331; and 5,719,262, each of which is herein incorporated by reference. Other backbone modifications, which can be used in the present invention are disclosed in U.S. Pat. No: 6,303,374.

Oligonucleotides of the present invention may also include base modifications or substitutions. As used herein, "unmodified" or "natural" bases include the purine bases adenine (A) and guanine (G), and the pyrimidine bases thymine (T), cytosine (C) and uracil (U). Modified bases include but are not limited to other synthetic and natural bases such as 5-methylcytosine (5-me-C), 5-hydroxymethyl cytosine, xanthine, hypoxanthine, 2-aminoadenine, 6-methyl and other alkyl derivatives of adenine and guanine, 2-propyl and other alkyl derivatives of adenine and guanine, 2-thiouracil, 2-thiothymine and 2-thiocytosine, 5-halouracil and cytosine, 5-propynyl uracil and cytosine, 6-azo uracil, cytosine and thymine, 5-uracil (pseudouracil), 4-thiouracil, 8-halo, 8-amino, 8-thiol, 8-thioalkyl, 8-hydroxyl and other 8-substituted adenines and guanines, 5-halo particularly 5-bromo, 5-trifluoromethyl and other 5-substituted uracils and cytosines, 7-methylguanine and 7-methyladenine, 8-azaguanine and 8-azaadenine, 7-deazaguanine and 7-deazaadenine and 3-deazaguanine and 3-deazaadenine. Further bases include those disclosed in U.S. Pat. No: 3,687,808, those disclosed in The Concise Encyclopedia Of Polymer Science And Engineering, pages 858-859, Kroschwitz, J. I., ed. John Wiley & Sons, 1990, those disclosed by Englisch et al., Angewandte Chemie, International Edition, 1991, 30, 613, and those disclosed by Sanghvi, Y. S., Chapter 15, Antisense Research and Applications, pages 289-302, Crooke, S. T. and Lebleu, B. , ed., CRC Press, 1993. Such bases are particularly useful for increasing the binding affinity of the oligomeric compounds of the invention. These include 5-substituted pyrimidines, 6-azapyrimidines and N-2, N-6 and O-6 substituted purines, including 2-aminopropyladenine, 5-propynyluracil and 5-propynylcytosine. 5-methylcytosine substitutions have been shown to increase nucleic acid duplex stability by 0.6-1.2°C. [Sanghvi YS et al. (1993) Antisense Research and Applications, CRC

Press, Boca Raton 276-278] and are presently preferred base substitutions, even more particularly when combined with 2'-O-methoxyethyl sugar modifications.

The polynucleotides of the present invention may be attached (either directly, or non-directly) to a targeting moiety. The targeting moiety may be a cell targeting moiety or an intracellular targeting moiety (e.g. nuclear targeting moiety). Thus, for example the present invention contemplates attaching the polynucleotide of the present invention to an antibody or fragment thereof, a receptor ligand or a glycosaminoglycan. Further, the polynucleotides of the present invention may be attached (either directly, or non-directly) to a detectable moiety – e.g. a fluorophore.

According to the teachings of the present invention, non-covalent association of the polynucleotide of the present invention with a targeting moiety may be effected via a nucleic acid carrier which is associated with both the polynucleotide of the present invention and the targeting moiety. Suitable nucleic acid carriers for practicing the method of the present invention are described further herein below.

Covalent association of molecules such as the nucleic acid carrier, the targeting moiety, and the polynucleotide of the present invention may be achieved using any of various chemical and biological methods well known to the ordinarily skilled artisan. For general guidance regarding the practice of such chemical methods, refer, for example, to the extensive guidelines provided by The American Chemical Society (http://worldwidewebdotchemistrydotorg/portal/Chemistry). One of ordinary skill in the art, such as, for example, a chemist, will possess the required expertise for practicing chemical techniques suitable for covalently associating molecules for practicing the present invention.

Introducing the polynucleotide of the present invention into genomic DNA in a cell can be effected using any method known in the art. Depending on the location of the cell, the polynucleotides may be introduced in vivo, ex vivo or in vitro. Selection of a particular method of introduction is typically dependent on the cell type being targeted.

Thus, for example if a mammalian cell is being targeted, the polynucleotide of the present invention may be incorporated into its genome using calcium phosphate method, electroporation or via a nucleic acid carrier.

It will be appreciated that a nucleic acid carrier which prevents chemical interaction of the polynucleotide of the present invention with a non cytoplasmic environment, such as serum in particular, can serve to prevent, *in-vivo* and *in-vitro*, degradation of the DNA molecule by the non cytoplasmic environment.

5 Exemplary nucleic acid carriers that may be used to transfect the polynucleotides of the present invention into mammalian cells include, but are not limited to liposomes, cationic lipids and synthetic polymers such as polyethylenimine (PEI) or poly-L-lysine.

Exemplary cationic lipids that may be used as carriers include, but are not limited 10 to 1,2-Dilauroyl-sn-Glicero-3-Phosphoethanolamine (DLPE) and 1,2-Dilauroyl-sn-Glicero-3-Glycerol (DLPG), dioleoyl-1,2-diacyl-3-trimethylammonium-propane (DOTAP, at 18:1; 14:0; 16:0, 18:0) and *N*-[1-(2,3-dioleyloxy)propyl]-*N*,*N*,*N*-trimethlylammonium chloride (DOTMA); dimethyldioctadecylammonium (DDAB); 1,2-dilauroyl-sn-glycero-3-ethylphosphocholine (Ethyl PC , at 12:0; 14:0; 16:0; 18:0; 15 18:1; 16:0-18:1); 1,2-di-(9Z-octadecenoyl)-3-dimethylammonium-propane and 3ß-[N-(N',N'-dimethylaminoethane)-carbamoyl]cholesterol hydrochloride (DC-Cholesterol).

The polynucleotides of the present invention may be introduced into plants by one or more techniques typically used for direct DNA delivery into cells. Such protocols may vary depending on the type of plant or plant cell, i. e. monocot or dicot, targeted for 20 gene modification. Suitable methods of transforming plant cells include microinjection (Crossway et al. (1986) Biotechniques 4: 320-334), electroporation (Riggs et al. (1986) Proc. Natl. Acad. Sci. USA 83: 5602-5606, direct gene transfer (Paszkowski et al. (1984) EMBO J. 3: 2717-2722), and ballistic particle acceleration (see, for example, Sanford et al., U. S. Patent No. 4,945,050; WO 91/10725 and McCabe et al. (1988) 25 Biotechnology 6: 923-926).

As mentioned, one embodiment of this aspect of the present invention comprises inserting a recombinase site into the genome of a cell, thus allowing the subsequent or concomitant insertion of an additional nucleic acid of interest.

This may be effected by a recombinase enzyme-mediated site-specific 30 integration of transfected DNA into the chromosomally positioned recombinase site.

Site-specific recombination is a process involving reciprocal exchange between specific nucleic acid sites (referred to as target sites) catalyzed by specialized proteins

known as site-specific recombinases. Site-specific recombinases recognize specific DNA sequences, and in the presence of specific recombination sites they catalyze the recombination of DNA strands. In these site-specific recombination systems, recombinases can catalyze excision or inversion of a DNA fragment according to the orientation of their specific target sites. Recombination between directly oriented sites leads to excision of the DNA between them, whereas recombination between inverted target sites causes inversion of the DNA between them.

The lambda integrase family of site-specific recombination systems consists of more than 100 different members. Among the most prominent of these are lambda Int, Cre/lox, FLP/FRT, R/RS, and Gin/gix. Recombinases such as Cre, FLP, R and Gin catalyze DNA recombination between their respective DNA substrates or target sites, loxP, FRT, RS and gix. These recombination systems use a common reaction pathway to carry out very different biological functions. They utilize a single polypeptide recombinase capable of recognizing a small DNA sequence without requiring any accessory factors. In heterologous systems, Cre/lox, FLP/FRT, R/RS, Gin/gix carry out a freely reversible reaction, whereas lamda-Int requires additional factors to carry out the reverse reaction.

According to this aspect of the present invention, following, or concomitant to insertion of the recombinase site (e.g. LoxP site) as described herein above, the cells are transfected with a mixture of two polynucleotides (plasmids): *(i)* a polynucleotide encoding the recombinase enzyme (e.g. CRE recombinase) and *(ii)* a polynucleotide comprising the nucleic acid of interest and an additional recombinase site (e.g. loxP site). As a result of the recombinase activity, the nucleic acid of interest will be incorporated into the loci where loxP resides (see Figure 7 and Figure 9A).

Exemplary recombinase enzymes include, but are not limited to Hin invertase, λ integrase, Cre and Flp, mutated forms thereof, temperature sensitive forms thereof as well as self destructing forms thereof (i.e. plasmids with delayed expression of RNAi against the recombinase.

It will be appreciated that the nucleic of interest may encode a selectable marker or detectable moiety, such that selection can be applied to acquire a homogenous cell population, or so that the success rate of integration may be followed, or even to visualize the number of sites that were incorporated in specific cells (e.g. by swapping

the recombination site with a cascade of DNA sequences (denoted "tagging cascade") that can be bound by specific proteins carrying a visible tag, such as the lacO/lacR-GFP system or the UAS/GAL4-GFP system). Alternatively, the polynucleotide (ii) may encode a selectable marker in addition to the nucleic acid of interest.

5 Typically, a selectable marker encodes a product that the host cell cannot make, such that the cell acquires resistance to a specific compound or is able to survive under specific conditions. For example, the marker may code for an antibiotic resistance factor. Suitable examples of antibiotic resistance markers include, but are not limited to, those coding for proteins that impart resistance to kanamycin, spectomycin, puromycin, 10 neomycin, geneticin (G418), ampicillin, tetracycline, and chlorampenicol. The selectable marker may code for proteins that confer resistance to herbicides, such as chlorsulfuron or phosphinotricin acetyltransferase. Other appropriate selectable markers include the thymidine kinase (tk) and the adenine phosphoribosyltransferase (apr) genes, which enable selection in $tk^-$and apr^- cells, respectively, and the dihydrofloate reductase (dhfr) 15 genes that confer resistance to methotrexate or trimethoprim.

Several approaches may be taken in order to select the cells that incorporated the nucleic acid of interest into their genome.

One approach is described in Example 3, herein below. Another approach is illustrated in Figures 8A-C.

20 This latter approach uses the Tet-Off system in combination with the Cre/loxP system. The Tet-Off system carries a tetR repressor which shuts the transcription from the plasmid by binding to a tetO sequences in the presence of Doxycycline (Dox). If the tetO sequences are located upstream of a promoter of a gene, then in the presence of tetR and Dox, the gene is silenced (Figure 8A). Upon plasmid integration, it undergoes 25 linearization at the lox site, placing the repressor site away from the region to be expressed (Figure 8C). A plasmid may be engineered with at least one tetO sequence upstream of a gene encoding a detectable moiety (e.g. eGFP gene) driven by a constitutive promoter (e.g. CMV promoter). In between the tetO repeats and the CMV promoter a lox site may be placed (Figure 15B). The expression of eGFP from this 30 plasmid is repressed by tetR in the presence of Dox. Upon plasmid integration, the tetO sequences will move far downstream of the CMV promoter, thus removing the inhibition on the expression of eGFP (Figure 8C).

Yet another approach is illustrated in Figures 10A. In this example, the plasmid has two different fluorophores (e.g. a GFP and Tomato – TFP –) under a constitutive promoter (e.g. CMV promoter or a PGK promoter which was shown to work in ES cells (among others)). The GFP is separated from its promoter by the lox site. Successful recombination of the plasmid into the genomic lox site separates the GFP from its promoter thereby turning of the green fluorescence while retaining the red fluorescence. The cells that have incorporated the plasmid are then sorted for TFP positive and GFP negative signals.

It will be appreciated that insertion of more than one recombinase site into a single chromosome of a cell allows for a deletion of a portion of that chromosome between the two recombinase sites.

Thus, according to another aspect of the present invention, there is provided a method of deleting a target sequence from genomic DNA in a cell, the method comprising:

(a) introducing the isolated polynucleotide which comprises a recombinase site into a plurality of cells; and

(b) introducing a polynucleotide encoding a recombinase enzyme into the cell, thereby deleting a target sequence from genomic DNA of an intact cell.

It will be appreciated that step (b) can be effected following step (a) or alternatively concomitant with step (a).

In the case of the first scenario, following introduction of the first polynucleotide, and prior to introduction of the second polynucleotide, cells may be selected that comprise more than one inserted recombinase site in a particular chromosome – e.g. by PCR or other such technique, or by insertion of a tagging cascade (as described above).

It will be appreciated that in order for the recombinase site to integrate into more than one site on a particular chromosome, the targeting sequences flanking the recombinase site have to target sites which repeat at least once along the length of the chromosome. Alternatively, a cocktail of targeting sites, each being homologous to a different locus can be used in order to integrate a single nucleic acid of interest into multiple sites that are different in sequence.

It will be appreciated however, that upon introduction of multiple genomic recombination sites, in the presence of a recombinase enzyme, any number of

recombinations can occur leading to unwanted (or alternatively desired) effects such as deletions of part of chromosomes and combining two or more chromosomes into one.

If the above scenarios are meant to be avoided, it is preferable that the recombinase site which is selected possesses two main qualities: (a) Uni-directionality; 5 which means that the incorporation reaction would be preferred to the excision reaction (Figures 9A-B and Figure 13B) (b) Inability to recombine with itself; this way the genomic mutant recombination sites would be unable to recombine among themselves while still being able to recombine with plasmid based recombination sites (Figure 13C).

Examples of recombination sites that possess these qualities are loxP mutants 10 (SEQ ID NO: 14, 15)). Such mutants should allow efficient recombination of all and any plasmids carrying complementary lox into the genome whilst at the same time avoid all non-specific by-products, such as those illustrated in Figures 12A-B.

According to one embodiment, the targeting sequences are engineered such that they are capable of hybridizing to a repeating motif.

15 Thus, according to another aspect of the present invention, there is provided a method of treating a trinucleotide repeat-associated disease in a subject in need thereof, the method comprising:

(a) introducing a single-stranded DNA polynucleotide into cells of the subject, wherein the single-stranded DNA polynucleotide encodes a recombinase site, 20 the site being flanked by two targeting sequences, wherein the two targeting sequences hybridize under physiological conditions with a continuous sequence of genomic DNA of the cells and wherein the two targeting sequences each hybridize with a trinucleotide repeating motif of the trinucleotide repeat associated disease.

(b) introducing a polynucleotide encoding a recombinase enzyme into the 25 cells, thereby treating the trinucleotide repeat-associated disease in the subject.

It will be appreciated that step (b) may be effected following or concomitant to step (a), as described earlier for other aspects of the present invention.

The present invention also contemplates using other single-stranded DNA polynucleotides of the present invention (which do not encode a recombination site) in 30 order to disrupt the trinucleotide repeating motif of the genomic DNA of the patient.

Diseases associated with repeating DNA sequences are described for example in Seminars in Pediatric Neurology, Volume 14, Issue 1, March 2007, Pages 26-33, incorporated herein by reference.

According to one embodiment, the trinucleotide repeat-associated disease is 5 Huntington's chorea.

According to one embodiment, the cells are neuronal cells derived from the subject.

According to another embodiment the cells are pluripotent cells (e.g. induced pluripotent cells) derived from the subject. Methods of inducing pluripotent stem cells 10 are known in the art – see for example U.S. Patent Application 20090047263. Example 7 of the Examples section herein below demonstrates successful incorporation of a fragment using the methods described herein into in iPS cell derived from a Huntington's patient.

It will be appreciated that following manipulation of iPS cells using the 15 constructs of the present invention, the cells are typically re-differentiated towards a particular phenotype – e.g. neuronal phenotype. Methods of differentiating cells towards a neuronal phenotype are known in the art, including those methods used for differentiating embryonic or mesenchymal stem cells towards a neuronal phenotype. It should be noted that the polynucleotide of interest incorporated into the iPS cells can 20 also be exchanged (by recombination) with another sequence carrying a fluorescent reporter and/or a selection gene downstream of a promoter specific to the desired cell type. This could be used to isolate the manipulated cells after they have re-differentiated to the desired type of tissue.

Following re-differentiation, the cells are typically transplanted back into the 25 patient at the appropriate site.

As used herein, the term "treating" includes abrogating, substantially inhibiting, slowing or reversing the progression of a condition, substantially ameliorating clinical or aesthetical symptoms of a condition or substantially preventing the appearance of clinical or aesthetical symptoms of a condition.

30 As used herein, the term "subject" includes mammals, preferably human beings at any age which suffer from the pathology. Preferably, this term encompasses individuals who are at risk to develop the pathology.

According to one embodiment of the invention the described method can be used to target specific genomic locations in order to allow controlled incorporation of the gene of interest for example in order to treat enzyme deficiencies.

According to another embodiment of the invention the described method can be used to overexpress a gene in human cells thus causing differentiation or transdifferentiation of the cells into a different cell type. For example primary liver cells can be extracted from a patient with Diabetes, grown in culture, and a PDX1 gene (under strong inducible promoter) may be introduced thereto using methods described herein. Thus as previously shown (Shternhall-Ron Keren; Journal of autoimmunity 2007;28(2-3):134-42.) the cells would transdifferentiate into β-cells which can produce insulin in response to blood glucose levels. These cells can then be reintroduced to the patient thus aiding in the treatment of Diabetes. These cells should not be predisposed to cancer as in cases where the gene is introduced virally since the introduction of PDX1 is location specific and does not interfere with other genes. Alternatively the promoter region of the endogenous gene (e.g. a master regulator gene such as PDX1 gene) can be modified using the current method to regulate expression (either up-regulate or down-regulate).

According to another embodiment, the methods described herein can be used in order to clean HIV-infected T-cell progenitors from the HIV virus by cutting out the incorporated HIV genome (Figure 16). This is currently impossible to do in any other way since the HIV incorporates randomly into the genome; its size doesn't allow the usage of homologous recombination to target it, and Zinc-Fingers are inapplicable because of the high HIV mutation rate which renders Zinc-Finger technology inapplicable. Additionally it is possible to knock-out the HIV co-receptors (CXCR4,CCR5) in the T-cell progenitors thus rendering them immune to further HIV infections (Figure 16).

The polynucleotides of the present invention may be introduced to the subject per se, or as part of a pharmaceutical composition.

As used herein, a "pharmaceutical composition" refers to a preparation of one or more of the active ingredients described herein with other chemical components such as physiologically suitable carriers and excipients. The purpose of a pharmaceutical composition is to facilitate administration of a compound to an organism.

Herein the term "active ingredient" refers to the polynucleotides accountable for the biological effect.

Hereinafter, the phrases "physiologically acceptable carrier" and "pharmaceutically acceptable carrier" which may be interchangeably used refer to a carrier or a diluent that does not abrogate the biological activity and properties of the administered compound. An adjuvant is included under these phrases.

Herein the term "excipient" refers to an inert substance added to a pharmaceutical composition to further facilitate administration of an active ingredient. Examples, without limitation, of excipients include calcium carbonate, calcium phosphate, various sugars and types of starch, cellulose derivatives, gelatin, vegetable oils and polyethylene glycols.

Techniques for formulation and administration of drugs may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition, which is incorporated herein by reference.

Suitable routes of administration may, for example, include oral, rectal, transmucosal, especially transnasal, intestinal or parenteral delivery, including intramuscular, subcutaneous and intramedullary injections as well as intrathecal, direct intraventricular, intracardiac, e.g., into the right or left ventricular cavity, into the common coronary artery, intravenous, inrtaperitoneal, intranasal, or intraocular injections.

Conventional approaches for drug delivery to the central nervous system (CNS) include: neurosurgical strategies (e.g., intracerebral injection or intracerebroventricular infusion); molecular manipulation of the agent (e.g., production of a chimeric fusion protein that comprises a transport peptide that has an affinity for an endothelial cell surface molecule in combination with an agent that is itself incapable of crossing the BBB) in an attempt to exploit one of the endogenous transport pathways of the BBB; pharmacological strategies designed to increase the lipid solubility of an agent (e.g., conjugation of water-soluble agents to lipid or cholesterol carriers); and the transitory disruption of the integrity of the BBB by hyperosmotic disruption (resulting from the infusion of a mannitol solution into the carotid artery or the use of a biologically active agent such as an angiotensin peptide). However, each of these strategies has limitations, such as the inherent risks associated with an invasive surgical procedure, a size

limitation imposed by a limitation inherent in the endogenous transport systems, potentially undesirable biological side effects associated with the systemic administration of a chimeric molecule comprised of a carrier motif that could be active outside of the CNS, and the possible risk of brain damage within regions of the brain where the BBB is disrupted, which renders it a suboptimal delivery method.

Alternately, one may administer the pharmaceutical composition in a local rather than systemic manner, for example, via injection of the pharmaceutical composition directly into a tissue region of a patient.

The term "tissue" refers to part of an organism consisting of an aggregate of cells having a similar structure and/or a common function. Examples include, but are not limited to, brain tissue, retina, skin tissue, hepatic tissue, pancreatic tissue, bone, cartilage, connective tissue, blood tissue, muscle tissue, cardiac tissue brain tissue, vascular tissue, renal tissue, pulmonary tissue, gonadal tissue, hematopoietic tissue.

Pharmaceutical compositions of the present invention may be manufactured by processes well known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes.

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in conventional manner using one or more physiologically acceptable carriers comprising excipients and auxiliaries, which facilitate processing of the active ingredients into preparations which, can be used pharmaceutically. Proper formulation is dependent upon the route of administration chosen.

For injection, the active ingredients of the pharmaceutical composition may be formulated in aqueous solutions and/or in biocompatible gels, preferably in physiologically compatible buffers such as Hank's solution, Ringer's solution, or physiological salt buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

For oral administration, the pharmaceutical composition can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the pharmaceutical composition to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions,

and the like, for oral ingestion by a patient. Pharmacological preparations for oral use can be made using a solid excipient, optionally grinding the resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, 5 including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carbomethylcellulose; and/or physiologically acceptable polymers such as polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as cross-linked polyvinyl pyrrolidone, agar, or 10 alginic acid or a salt thereof such as sodium alginate.

Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, titanium dioxide, lacquer solutions and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to 15 the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical compositions which can be used orally, include push-fit capsules made of gelatin as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules may contain the active ingredients in 20 admixture with filler such as lactose, binders such as starches, lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active ingredients may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for the chosen route of administration.

25 For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by nasal inhalation, the active ingredients for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from a pressurized pack or a nebulizer with the use of a suitable propellant, 30 e.g., dichlorodifluoromethane, trichlorofluoromethane, dichloro-tetrafluoroethane or carbon dioxide. In the case of a pressurized aerosol, the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, e.g.,

gelatin for use in a dispenser may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

The pharmaceutical composition described herein may be formulated for parenteral administration, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampoules or in multidose containers with optionally, an added preservative. The compositions may be suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical compositions for parenteral administration include aqueous solutions of the active preparation in water-soluble form. Additionally, suspensions of the active ingredients may be prepared as appropriate oily or water based injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acids esters such as ethyl oleate, triglycerides or liposomes. Aqueous injection suspensions may contain substances, which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the active ingredients to allow for the preparation of highly concentrated solutions.

Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile, pyrogen-free water based solution, before use.

The pharmaceutical composition of the present invention may also be formulated in rectal compositions such as suppositories or retention enemas, using, e.g., conventional suppository bases such as cocoa butter or other glycerides.

Pharmaceutical compositions suitable for use in context of the present invention include compositions wherein the active ingredients are contained in an amount effective to achieve the intended purpose. More specifically, a therapeutically effective amount means an amount of active ingredients effective to prevent, alleviate or ameliorate symptoms of a disorder (e.g., Huntington's Chorea) or prolong the survival of the subject being treated.

Determination of a therapeutically effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein.

For any preparation used in the methods of the invention, the therapeutically effective amount or dose can be estimated initially from in vitro and cell culture assays. For example, a dose can be formulated in animal models to achieve a desired concentration or titer. Such information can be used to more accurately determine useful doses in humans.

Toxicity and therapeutic efficacy of the active ingredients described herein can be determined by standard pharmaceutical procedures in vitro, in cell cultures or experimental animals. The data obtained from these in vitro and cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage may vary depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. (See e.g., Fingl, et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1).

Dosage amount and interval may be adjusted individually to provide levels of the active ingredient that are sufficient to induce or suppress the biological effect (minimal effective concentration, MEC). The MEC will vary for each preparation, but can be estimated from in vitro data. Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. Detection assays can be used to determine plasma concentrations.

Depending on the severity and responsiveness of the condition to be treated, dosing can be of a single or a plurality of administrations, with course of treatment lasting from several days to several weeks or until cure is effected or diminution of the disease state is achieved.

The amount of a composition to be administered will, of course, be dependent on the subject being treated, the severity of the affliction, the manner of administration, the judgment of the prescribing physician, etc.

Compositions of the present invention may, if desired, be presented in a pack or dispenser device, such as an FDA approved kit, which may contain one or more unit dosage forms containing the active ingredient. The pack may, for example, comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. The pack or dispenser may also be accommodated by a notice associated with the container in a form prescribed by a

governmental agency regulating the manufacture, use or sale of pharmaceuticals, which notice is reflective of approval by the agency of the form of the compositions or human or veterinary administration. Such notice, for example, may be of labeling approved by the U.S. Food and Drug Administration for prescription drugs or of an approved product insert. Compositions comprising a preparation of the invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition, as is further detailed above.

As used herein the term “about” refers to ± 10 %.

The terms "comprises", "comprising", "includes", "including", “having” and their conjugates mean "including but not limited to".

The term “consisting of means “including and limited to”.

The term "consisting essentially of" means that the composition, method or structure may include additional ingredients, steps and/or parts, but only if the additional ingredients, steps and/or parts do not materially alter the basic and novel characteristics of the claimed composition, method or structure.

As used herein, the singular form "a", "an" and "the" include plural references unless the context clearly dictates otherwise. For example, the term "a compound" or "at least one compound" may include a plurality of compounds, including mixtures thereof.

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

Whenever a numerical range is indicated herein, it is meant to include any cited numeral (fractional or integral) within the indicated range. The phrases “ranging/ranges

between" a first indicate number and a second indicate number and "ranging/ranges from" a first indicate number "to" a second indicate number are used herein interchangeably and are meant to include the first and second indicated numbers and all the fractional and integral numerals therebetween.

5 As used herein the term "method" refers to manners, means, techniques and procedures for accomplishing a given task including, but not limited to, those manners, means, techniques and procedures either known to, or readily developed from known manners, means, techniques and procedures by practitioners of the chemical, pharmacological, biological, biochemical and medical arts.

10

It is appreciated that certain features of the invention, which are, for clarity, described in the context of separate embodiments, may also be provided in combination in a single embodiment. Conversely, various features of the invention, which are, for brevity, described in the context of a single embodiment, may also be provided
15 separately or in any suitable subcombination or as suitable in any other described embodiment of the invention. Certain features described in the context of various embodiments are not to be considered essential features of those embodiments, unless the embodiment is inoperative without those elements.

Various embodiments and aspects of the present invention as delineated
20 hereinabove and as claimed in the claims section below find experimental support in the following examples.

EXAMPLES

Reference is now made to the following examples, which together with the above
25 descriptions illustrate some embodiments of the invention in a non limiting fashion.

Generally, the nomenclature used herein and the laboratory procedures utilized in the present invention include molecular, biochemical, microbiological and recombinant DNA techniques. Such techniques are thoroughly explained in the literature. See, for example, "Molecular Cloning: A laboratory Manual" Sambrook et
30 al., (1989); "Current Protocols in Molecular Biology" Volumes I-III Ausubel, R. M., ed. (1994); Ausubel et al., "Current Protocols in Molecular Biology", John Wiley and Sons, Baltimore, Maryland (1989); Perbal, "A Practical Guide to Molecular Cloning", John

Wiley & Sons, New York (1988); Watson et al., "Recombinant DNA", Scientific American Books, New York; Birren et al. (eds) "Genome Analysis: A Laboratory Manual Series", Vols. 1-4, Cold Spring Harbor Laboratory Press, New York (1998); methodologies as set forth in U.S. Pat. Nos. 4,666,828; 4,683,202; 4,801,531; 5,192,659
5 and 5,272,057; "Cell Biology: A Laboratory Handbook", Volumes I-III Cellis, J. E., ed. (1994); "Culture of Animal Cells - A Manual of Basic Technique" by Freshney, Wiley-Liss, N. Y. (1994), Third Edition; "Current Protocols in Immunology" Volumes I-III Coligan J. E., ed. (1994); Stites et al. (eds), "Basic and Clinical Immunology" (8th Edition), Appleton & Lange, Norwalk, CT (1994); Mishell and Shiigi (eds), "Selected
10 Methods in Cellular Immunology", W. H. Freeman and Co., New York (1980); available immunoassays are extensively described in the patent and scientific literature, see, for example, U.S. Pat. Nos. 3,791,932; 3,839,153; 3,850,752; 3,850,578; 3,853,987; 3,867,517; 3,879,262; 3,901,654; 3,935,074; 3,984,533; 3,996,345; 4,034,074; 4,098,876; 4,879,219; 5,011,771 and 5,281,521; "Oligonucleotide
15 Synthesis" Gait, M. J., ed. (1984); "Nucleic Acid Hybridization" Hames, B. D., and Higgins S. J., eds. (1985); "Transcription and Translation" Hames, B. D., and Higgins S. J., eds. (1984); "Animal Cell Culture" Freshney, R. I., ed. (1986); "Immobilized Cells and Enzymes" IRL Press, (1986); "A Practical Guide to Molecular Cloning" Perbal, B., (1984) and "Methods in Enzymology" Vol. 1-317, Academic Press; "PCR
20 Protocols: A Guide To Methods And Applications", Academic Press, San Diego, CA (1990); Marshak et al., "Strategies for Protein Purification and Characterization - A Laboratory Course Manual" CSHL Press (1996); all of which are incorporated by reference as if fully set forth herein. Other general references are provided throughout this document. The procedures therein are believed to be well known in the art and are
25 provided for the convenience of the reader. All the information contained therein is incorporated herein by reference.

EXAMPLE 1

Incorporation of LoxP using ssDNA oligonucleotides (incorporating sequence = SEQ ID NO: 1)

30 **MATERIALS AND METHODS**

Selection of target site: A 40 b.p. target site was selected which repeated 22 times in a 27kb region of chromosome 1 (chr1:2576961-2603453). This region is

non-conserved and not recognized as a simple repeat. It is devoid of genes but is flanked with active genes both immediately upstream (MMEL1) and downstream (ACTRT2), both of which were reported to be expressed in HeLa cells. The incorporating sequence used to target this repeat was SEQ ID NO: 1.

5 ***Transfection of cells with the fragments of the present invention:*** Several approaches were used to transfect the DNA in order to rule out influence of the transfection step. These include electroporation and using chemical agents such as PEI, JetPEI, MIRUS Transit, FuGENE HD.

Extraction of genomic DNA: DNA was extracted in several ways to exclude the 10 influence of specific choice of DNA extraction. For example, the DNA was extracted by lysis followed by phenol-chloroform extraction; In addition, the DNA was extracted using Sigma Genelute mammalian genomic DNA miniprep kit.

Testing incorporation of the loxP site: A pair of primers was selected such that both the forward and the reverse PCR primers are 20bp sequences appearing in multiple 15 repeats in the same region as the targeted 40bp repeat (Figure 1A). Upon incorporation of loxP, the PCR primer pair: "loxP F" (ATAACTTCGTATAGCATACATTATACGAAGTTAT – SEQ ID NO: 24) & "Genomic R" (ATCTGGGCGTATGGGTGCTG – SEQ ID NO: 25) should yield multiple bands, while "Genomic R" and "Genomic F" 20 (GAGCATCTGACAGCCTGGAA – SEQ ID NO: 26) serve as internal PCR control (Figure 1B).

RESULTS

A clear difference between the bands was observed following transfection of the fragment (SEQ ID NO: 1) or of GFP, suggesting successful loxP incorporation events 25 into the genome (Figure 1C). To verify this possibility, a band was cut from the gel, ligated to pGEM-t-easy vector and transformed into E-coli. Sequencing plasmids from several colonies revealed the presence of incorporated loxP sites (Figures 2A-B). All the incorporations were detected at the expected locations. Importantly, one of the colonies had two insertions indicating that some cells incorporated the fragment at least twice per 30 cell.

To test the stability of the incorporation, the cells were analyzed 6 generations following transfection of the fragment. Genomic DNA was extracted, the incorporation

was analyzed as described above and similar results were found indicating stable incorporation (Figure 3B).

To examine the efficiency of the procedure in other cell lines, the procedure was repeated for the H1 line of human Embryonic Stem cells (hESCs), resulting in lower, yet still detectable yields (Figure 3C). Other cell lines (Lung Cancer, Human Foreskin Fibroblast) also showed loxP incorporation (not shown).

To evaluate the likelihood that the present incorporation method is similar to current ssDNA-based techniques, the dependency of the efficiency of incorporation on replication or DNA damage dependent machineries was tested. Incorporation by a similar mechanism is expected to increase upon delay in proliferation and induction of DNA damage. The present inventors therefore irradiated HeLa cells prior to transfection with SEQ ID NO: 1, so as to stop their proliferating and induce DNA damage. It was found that the irradiation abolished the incorporation (Figure 3D). This result is in contrast to reported cases in which the incorporation was improved by irradiation. This result is in line with the sequence being incorporated as an Okazaki-like fragment during replication. Once, replication is blocked, there could be no more incorporation. This experiment provides indirect evidence to suggest that the present method of incorporation is different from previously published methods.

EXAMPLE 2

Incorporation of lacO, lox66, loxP and lowE using ssDNA oligonucleotides

To evaluate the dependency of incorporation on the insert to be incorporated four different inserts were tested (lacO - SEQ ID NO: 3, lox66 - SEQ ID NO: 4, loxP - SEQ ID NO: 1 and lowE - SEQ ID NO: 5 each with different secondary structures, but otherwise flanked by the same 20bp regions

MATERIALS AND METHODS

Transfection: The fragments were transfected as described in Example 1.

RESULTS

It appears that the free energy of the fragment and length of the fragment might play a role in the efficiency of incorporation (Figure 4).

EXAMPLE 3

Relationship between incorporation efficiency and the targeting bases of the ssDNA oligonucleotides

5 **MATERIALS AND METHODS**

Fragments: A fragment with a single loxP insert was incorporated into two different 22 repeats of a 40bp sequence: rpt22-1(SEQ ID NO: 1) and rpt22-2 (SEQ ID NO: 6). Additionally, a fragment with a single lox 66 insert was incorporated into 3 different genomic repeats: rpt22 (22 repeats, SEQ ID NO: 4), rpt 5188 (5188 repeats, 10 SEQ ID NO: 7) and rpt5192 (5192 repeats, SEQ ID NO: 8).

Transfection: The transfection procedure was effected as described in Example 1 herein above.

RESULTS

To analyze the effect of the genomic target site on the efficiency of fragment 15 incorporation, another 40bp genomic repeat was selected with the exact same number of repeats as the first repeat described above: rpt22-2 (22 repeats). The transfection and detection procedure was repeated for both fragments (Figure 5A). The incorporation frequency was clearly higher in the first repeat tested (rpt 22-1). When the homologous flanks of the fragment were shortened to 10bp (instead of 20bp), no detectable 20 incorporation occured (data not shown).

To further test the dependency of incorporation on the number of repeats the present inventors targeted fragments with lox66 insert into three different repeats (Figure 5B): *(i)* rpt22-1 – the 22 repeat sequence described herein above *(ii)* rpt 5188 – a repeat that appears 5188 times in the genome (evenly distributed across all the 25 chromosomes) and *(iii)*, rpt 5192 – which appears 5192 times in the genome (also evenly distributed). A strong, positive dependency of the efficiency of incorporation on the number of target sites (Figure 5B) was found. Moreover, the strong signal obtained for very high repeat numbers suggests that incorporation takes place in most of the cells.

EXAMPLE 4

Validating Plasmid Integration

MATERIALS AND METHODS

Fragments and plasmids: (i) A vector plasmid carrying the eGFP gene under a CMV promoter (Figure 6A). (ii) a fragment comprising a sequence as set forth in SEQ ID NO: 4.

Transfection: Briefly, cells were transfected on day 1 with the fragment (SEQ ID NO: 4) via electroporation. On day 3 (48 hours later), the same cells were transfected again with mixture of plasmids carrying, respectively, eGFP and the Cre recombinase gene.

FACS analysis: FACS was performed on FACS ARIA. Only cells with GFP levels above background were collected and re-plated.

RESULTS

FACS was performed on day 7 (4 days following eGFP and Cre recominase trasfection) to collect only the GFP positive cells, which at this point indicate the cells that successfully underwent transfection. At this stage there were 70 % GFP positive cells overall. These cells were passaged (1:10 dilution each ~3 days) for two more weeks without selection, and re-sorted by FACS on day 22. By that time, ~0.5 % of the cells were GFP positive. The cells were then passaged for 16 more days (without selection) and imaged on day 38. Figure 6B shows that almost all the cells were still expressing the eGFP gene after 38 days, a period much longer than that of plasmid retention in mammalian cells indicating that the plasmid was retained in the genome of the cell.

EXAMPLE 5

Validation of an exemplary selection scheme according to one embodiment of the present invention

Human Embryonic Stem Cells (hES) were passaged on Mouse Embryonic Fibroblasts (MEFs) and subsequently were transferred to matrigel and transfected with the fragment comprising a sequence as set forth in SEQ ID NO: 4. Following 2 passages the cells were subsequently transfected with the improved Vector Plasmid p7.2LTG (LTG=PGK Lox TFP STOP PGK GFP) as illustrated in Figure 10B.

RESULTS

One day following the second transfection, all of the transfected cells exhibited both GFP and TFP fluorescence (Figures 11B-D). However 4 days post the second transfection, three distinct populations of cells can be seen: (a) cells that lost all fluorescence (b) cells that kept both GFP (Figure 11F blue circle) and TFP (Figure 11G blue circle) fluorescence and therefore hasn't integrated into the genome (c) and cells that are still GFP positive, but TFP negative (Figure 11F,G red circles) which is indicative of the plasmids' genomic integration through the lox site.

EXAMPLE 6

Targeting of Pdx1 in Hela cells

A "cocktail" of 5 fragments (SEQ ID NOs: 28-32) were used each with complementary sequences to different regions located close-by within 1.5 kbase of the Pdx1 intron (Figure 14B). Inserting the loxP site into the intron allowed for its subsequent exchange with a fluorescent protein (XFP) flanked by sequences that promote its splicing into the Pdx1 gene as an additional exon, see for example [Frenkel-Morgenstern M., et al., Nucleic Acids Research, October 9, 2009. Vol 38(Database issue 2010):D508-D512]).

Successful targeting of Pdx1 was demonstrated by several means. First, LoxP incorporation was tested by PCR using probes against the LoxP on one end, and Pdx1 on the other end (Figure 14C). The sequence of the primers are provided in SEQ ID NOs: 33-34. Successful incorporations are expected to yield bands with lengths which depend on the distance from the Pdx1 probe sequence. Multiple bands indicate multiple sites of incorporation. Higher efficiencies of incorporation are expected to yield stronger and potentially higher number of bands. Consistent with this expectation a dose dependent occurrence and strength of specific bands was observed, the number of bands depending on the number of different fragments in the cocktail with larger diversity giving rise to more bands (Figure 14D). To further verify that the incorporation is specific to the Pdx1 locus, the bands were extracted from the gel, cloned using pgem-t-easy vector, and the colonies sequenced. As illustrated in Figure 14E, correct insertion of the loxP site into the intronic region of Pdx1 was found. Finally, the exchange of the

loxP site with a plasmid carrying eGFP driven by a PGK promoter and G418 resistance gene (*neo*) driven by SV40 promoter was tested. Both the selection and fluorescent markers were stably exchanged into the Hela cells following transient transfection of the plasmid together with the Cre recombinase (Figure 14F). The entire procedure of OFI incorporation and exchange takes no more than 5 days. To further verify that the stable incorporation resulted from a specific exchange with the loxP site, the survival rate in G418 as a function of loxP dosage used in the preceding step of OFI incorporation was examined. It was found that the G418 tolerance of the cells was clearly dependent on the concentration of OFI, indicating that the selection plasmid was indeed exchanged with the loxP sites (Figures 15A-C). In parallel the present inventors showed that the targeting of Pdx1 works with as few as 3 different OFIs in the cocktail.

EXAMPLE 7

Targeting of pluripotent stem cells containing long tri-nucleotide repeats

The present inventors targeted human induced pluripotent cells (hiPSCs) containing abnormally long tri-nucleotide repeats (~180 repeats) of the Huntington gene.

MATERIALS AND METHODS

Generation of iPCSs: Skin cells from a Huntington patient were induced to pluripotency using the method described by Takahashi et al., j.cell.2007.11.019.

Incorporation of the fragment: Using a single fragment against the repeats (SEQ ID NO: 35) the cells were targeted with a fragment comprising LoxP which was subsequently exchanged with GFP.

RESULTS

As illustrated in Figures 16A-B, GFP was successfully incorporated into the induced pluripotent stem cells.

EXAMPLE 8

Targeting CXCR4 (the co-receptor mediating HIV infection) using a single OFI

A single OFI was used to target CXCR4 gene – SEQ ID NO: 36 (Figure 17B). As in the case of PDX1 targeting described in Example 7, a single probe (SEQ ID NO:

exon of human CXCR4 gene. This gene is a co-receptor for HIV X4 strain entry into T-Cells, and without which the infection is impossible (Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases Nature Biotechnology 26, 808 - 816 (2008) Published online: 29 June 2008 doi:10.1038/nbt1410).

RESULTS

The cells were transfected with the OFI and allowed to proliferate for 2 days. Genomic DNA was extracted. PCR analysis of the genomic DNA (using the primers set forth in SEQ ID NOs: 38 and 39), showed correct single incorporation of lox66 site into desired locus (Figure 17B).

Although the invention has been described in conjunction with specific embodiments thereof, it is evident that many alternatives, modifications and variations will be apparent to those skilled in the art. Accordingly, it is intended to embrace all such alternatives, modifications and variations that fall within the spirit and broad scope of the appended claims.

All publications, patents and patent applications mentioned in this specification are herein incorporated in their entirety by reference into the specification, to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated herein by reference. In addition, citation or identification of any reference in this application shall not be construed as an admission that such reference is available as prior art to the present invention. To the extent that section headings are used, they should not be construed as necessarily limiting.

WHAT IS CLAIMED IS:

1. An isolated polynucleotide comprising a single-stranded DNA polynucleotide encoding a regulatory site, said regulatory site being longer than 15 bases and being flanked by two targeting sequences, wherein said two targeting sequences hybridize with a continuous sequence of mammalian genomic DNA and wherein each of said targeting sequence is at least 10 bases.

2. The isolated polynucleotide of claim 1, wherein said regulatory site comprises a recombination site.

3. The isolated polynucleotide of claim 1, being attached to a targeting moiety.

4. The isolated polynucleotide of claim 2, wherein said recombination site is selected from the group consisting of a *hix* site, an *att* site, a *lox* site and a *frt* site.

5. The isolated polynucleotide of claim 4, wherein said lox site comprises a sequence as set forth in SEQ ID NOs: 9-15.

6. The isolated polynucleotide of claim 1, wherein each of said two targeting sequences is 20 bases.

7. The isolated polynucleotide of claim 1, wherein at least one of said two targeting sequences comprises a modified base.

8. The isolated polynucleotide of claim 1, wherein said continuous sequence of genomic DNA repeats at least twenty times in a single chromosome of the cell.

9. A method of inserting a nucleic acid sequence of interest into genomic DNA in a cell, the method comprising introducing at least one single-stranded DNA

polynucleotide into the cell, wherein said at least one single-stranded DNA polynucleotide comprises the nucleic acid sequence of interest flanked by two targeting sequences, wherein said two targeting sequences hybridize under physiological conditions with a continuous sequence of the genomic DNA and wherein the nucleic acid sequence of interest is longer than 15 bases, thereby inserting the nucleic acid sequence into genomic DNA of a cell.

10. The method of claim 9, wherein said at least one single-stranded DNA polynucleotide comprises at least two identical single-stranded DNA polynucleotides.

11. The method of claim 9, wherein said at least one single-stranded DNA polynucleotide comprises two non-identical single-stranded DNA polynucleotides.

12. The method of claim 11, wherein said two non-identical single-stranded DNA polynucleotides comprise non-identical nucleic acid sequences of interest.

13. The method of claim 11, wherein said two non-identical single-stranded DNA polynucleotides comprise non-identical targeting sequences.

14. The method of claim 9, wherein the genomic DNA is selected from the group consisting of nuclear DNA, mitochondrial DNA and chloroplast DNA.

15. The method of claim 9, wherein the nucleic acid sequence of interest comprises at least one stem-loop structure.

16. The method of claim 9, wherein the nucleic acid sequence of interest comprises less than 50 bases.

17. The method of claim 15, wherein the nucleic acid sequence of interest encodes a recombination site.

18. The method of claim 17, wherein said recombination site is selected from the group consisting of a *hix* site, an *att* site, a *lox* site and a *frt* site.

19. The method of claim 18, wherein said lox site comprises a sequence as set forth in SEQ ID NOs: 9-15.

20. The method of claim 9, wherein at least one of said two targeting sequences is greater than 10 bases.

21. The method of claim 9, wherein at least one of said two targeting sequences is 20 bases.

22. The method of claim 9, wherein at least one of said two targeting sequences comprises a modified base.

23. The method of claim 9, wherein said continuous sequence of genomic DNA repeats at least twenty times in a single chromosome of the cell.

24. The method of claim 9, wherein the cell is a eukaryotic cell.

25. The method of claim 9, wherein the cell is a mammalian cell.

26. The method of claim 9, wherein the cell is a plant cell.

27. The method of claim 9, wherein the cell is an intact cell.

28. The method of claim 25, wherein said mammalian cell is a human cell.

29. The method of claim 9, wherein the cell is a stem cell.

30. The method of claim 29, wherein said stem cell is an embryonic stem cell.

31. The method of claim 9, wherein the cell is a diseased cell.

32. The method of claim 9, being effected in vivo.

33. The method of claim 9, being effected ex vivo.

34. The method of claim 9, wherein said single stranded DNA polynucleotide is associated with a cell-targeting moiety.

35. A method of inserting a nucleic acid of interest into genomic DNA of a population of cells, the method comprising:

(a) introducing the isolated polynucleotide of claim 2 into the cell population; and

(b) introducing:

(i) a polynucleotide which encodes said recombination site and comprises the nucleic acid of interest; and

(ii) a polynucleotide encoding a recombinase enzyme into the cell population, thereby inserting the nucleic acid of interest into genomic DNA of the population of cells.

36. The method of claim 35, wherein step (a) is affected prior to step (b).

37. The method of claim 35, wherein step (a) is effected concomitant to step (b).

38. The method of claim 35, wherein the genomic DNA is selected from the group consisting of nuclear DNA, mitochondrial DNA and chloroplast DNA.

39. The method of claim 35, wherein said recombinase enzyme is selected from the group consisting of Hin invertase, λ integrase, Cre and Flp.

40. The method of claim 35, wherein the cell is a eukaryotic cell.

41. The method of claim 35, wherein the cell is an intact cell.

42. The method of claim 40, wherein the eukaryotic cell is a mammalian cell.

43. The method of claim 42, wherein said mammalian cell is a human cell.

44. The method of claim 35, wherein the cell is a stem cell.

45. The method of claim 44, wherein the stem cell is an embryonic stem cell.

46. The method of claim 35, wherein the cell is a diseased cell.

47. The method of claim 35, being effected in vivo.

48. The method of claim 35, being effected ex vivo.

49. The method of claim 35, wherein said nucleic acid of interest encodes at least one detectable moiety.

50. The method of claim 49, further comprising selecting cells which comprise the nucleic acid of interest.

51. A method of deleting a target sequence from genomic DNA in a cell, the method comprising:

(a) introducing the isolated polynucleotide of claim 2 into a plurality of cells;

(b) introducing a polynucleotide encoding a recombinase enzyme into said plurality of cells, thereby deleting a target sequence from genomic DNA of a cell.

52. The method of claim 51, wherein step (a) is effected prior to step (b).

53. The method of claim 51, wherein step (a) is effected concomitant to step (b).

54. The method of claim 51 further comprising selecting at least one of said plurality of cells, prior to step (b) wherein said at least one of said plurality of cells comprises at least two recombinase sites integrated into a single chromosome thereof.

55. The method of claim 51, wherein said recombinase enzyme is selected from the group consisting of Hin invertase, λ integrase, Cre and Flp.

56. The method of claim 51, wherein the cell is a mammalian cell.

57. The method of claim 56, wherein said mammalian cell is a human cell.

58. The method of claim 51, wherein the cell is a stem cell.

59. The method of claim 58, wherein said stem cell is an embryonic stem cell.

60. The method of claim 51, wherein the cell is a diseased cell.

61. A method of treating a trinucleotide repeat-associated disease in a subject in need thereof, the method comprising:

(a) introducing a single-stranded DNA polynucleotide into cells of the subject, wherein the single-stranded DNA polynucleotide encodes a recombinase site, said site being flanked by two targeting sequences, wherein said two targeting sequences hybridize under physiological conditions with a continuous sequence of genomic DNA of said cells and wherein said two targeting sequences each hybridize with a trinucleotide repeating motif of the trinucleotide repeat associated disease; and

(b) introducing a polynucleotide encoding a recombinase enzyme into the cells, thereby treating Huntington's chorea in a subject.

62. The method of claim 61, wherein step (a) is effected prior to step (b).

63. The method of claim 61, wherein step (a) is effected concomitant to step (b).

64. The method of claim 61, wherein the trinucleotide repeat-associated disease is Huntington's chorea.

65. The method of claim 62, further comprising selecting at least one of said cells, prior to step (b) wherein said at least one of said cells comprises at least two recombinase sites integrated into a single chromosome thereof.

66. The method of claim 61, wherein said cells comprise neuronal cells.

67. The method of claim 61, wherein said cells comprise stem cells.

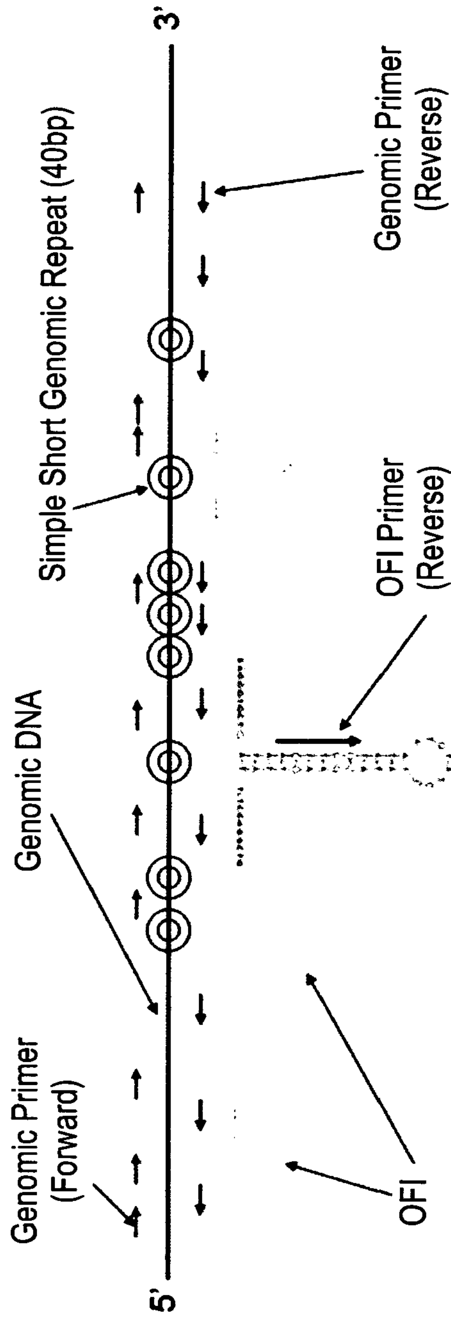


FIG. 1A

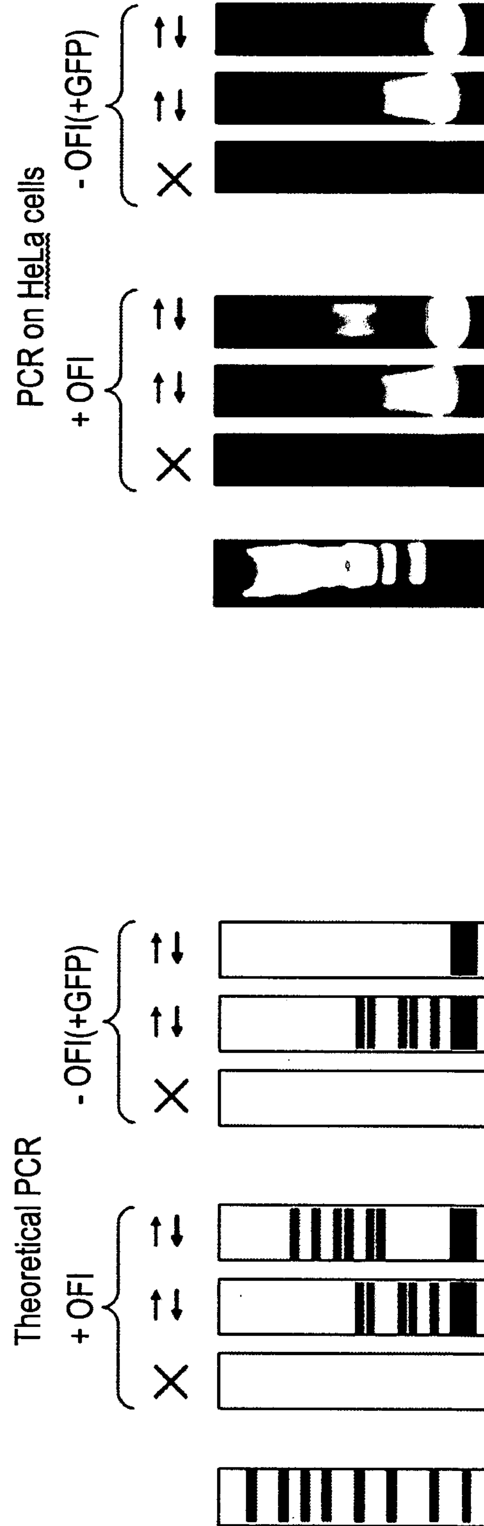
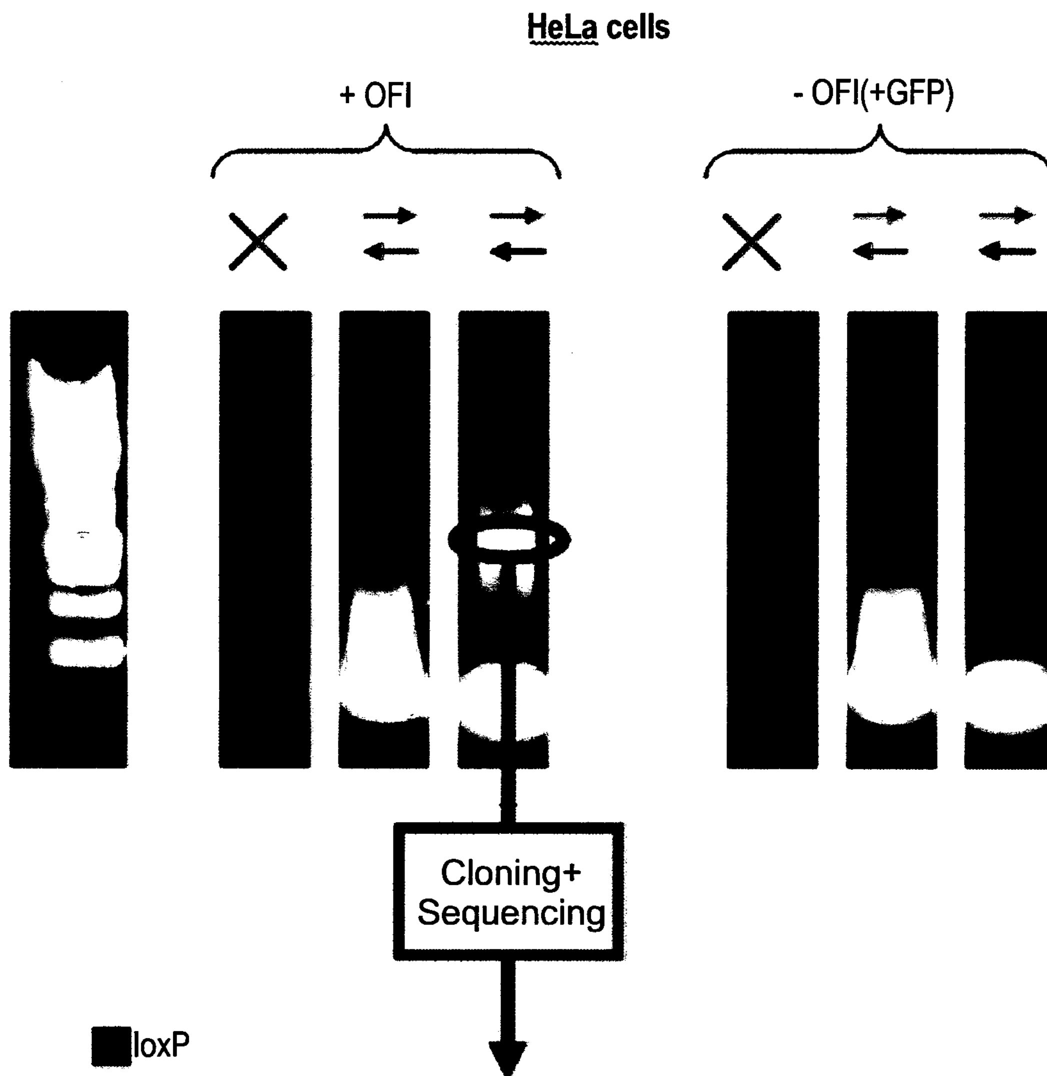


FIG. 1B

FIG. 1C

FIG. 2A




```
  1  NNNNNNNNNC CNNNTCGNAT GCTCCCGGCC GNNNGGCGGC CGCGGGAATT
 51  CGATTCCCCC AGGTGAGTAT CTGACGGCCT GGAATAGCAC CCACACCCGC
101  AGGTGAGCAT CTGACATCGT GGAGCAGCAC CCACACCCAC AAGTGCGCAT
151  CTGACAGCCT GGAA                                    CA
201  GCACCCTGCA CCCCCAGGCG AGCATCTGAA CTCACAGAGC AGCAACCACA
251  CCCCCAGGTG AGCATCTGAC AGCCTGGAA
301                          CAGCACC CTGCACCCCC AGGTGAGGAT CTGACAGCCT
351  GAAACAGCAC CCTCCACCTT CAGGTGAGAA TATGACAGCC TGAAACAGCA
401  CCCCTACTGG CAATCACTAG TGAATTCGCG GGCGCCTGCA GGTCGACCAT
```

FIG. 2B

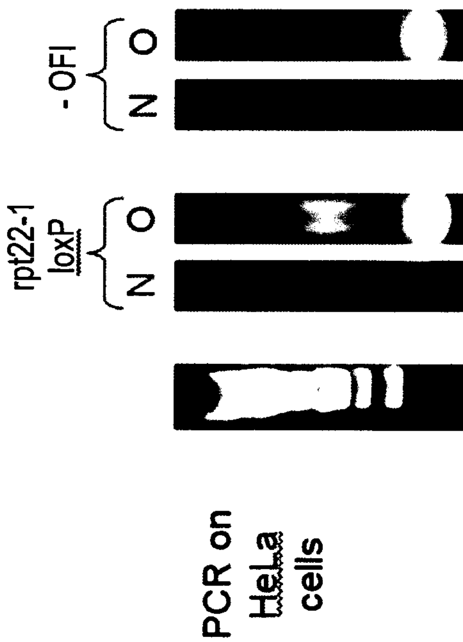


FIG. 3A

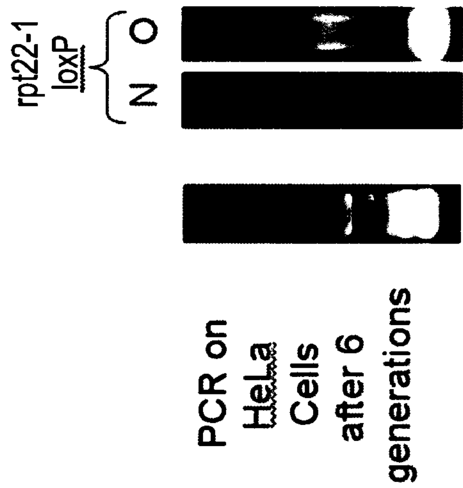


FIG. 3B

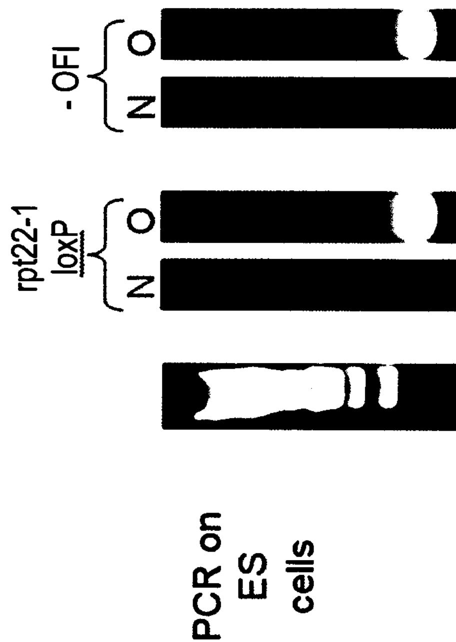


FIG. 3C

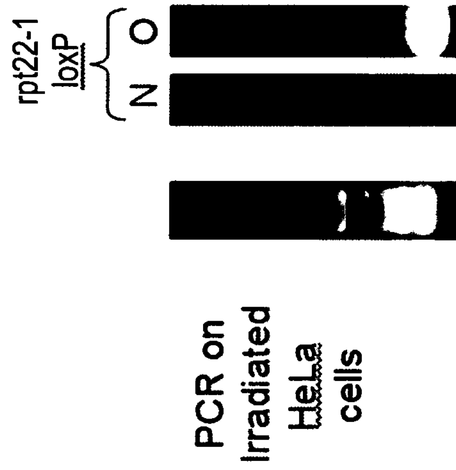


FIG. 3D

FIG. 4

HeLa cells

lacO N O > lox66 N O ≥ loxP N O > lowE N O

dG=-16.25 08Oct06-09-59-10
-16 kcal/mole

dG=-8.47 09Jan04-10-41-49
-8.47 kcal/mole

dG=-10.86 08Oct07-09-17-57
-11 kcal/mole

dG=-2.59 08Oct30-11-11-55
-2.59 kcal/mole

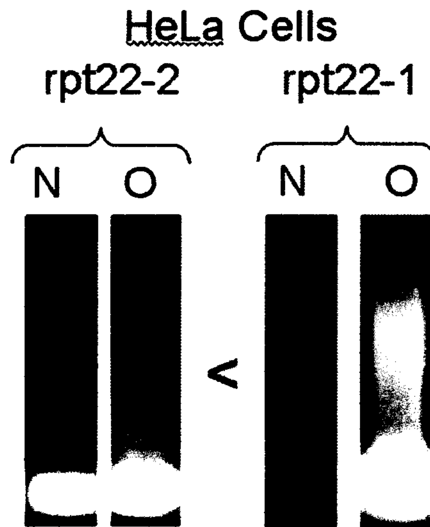


FIG. 5A

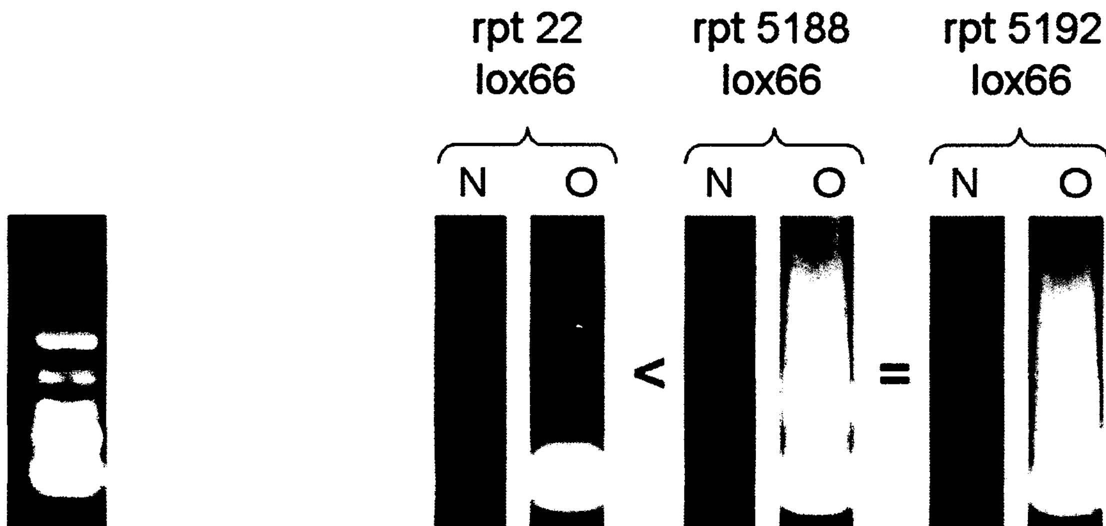
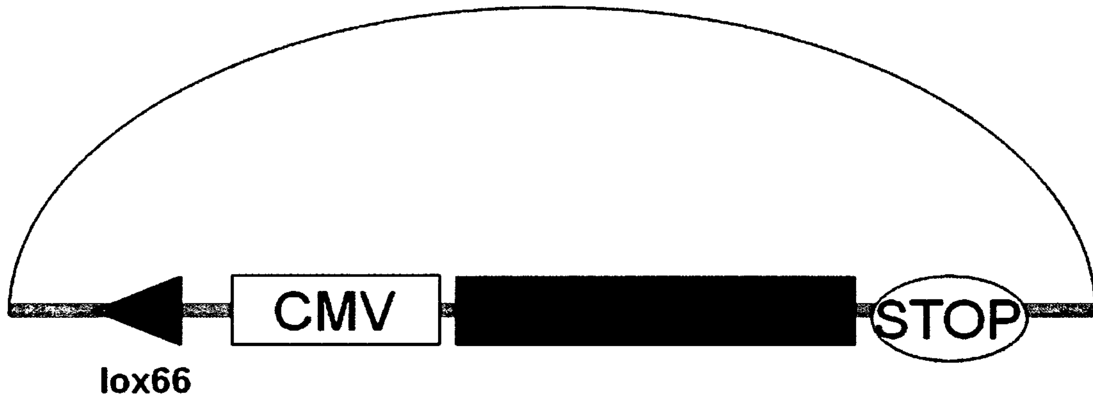


FIG. 5B

FIG. 6A



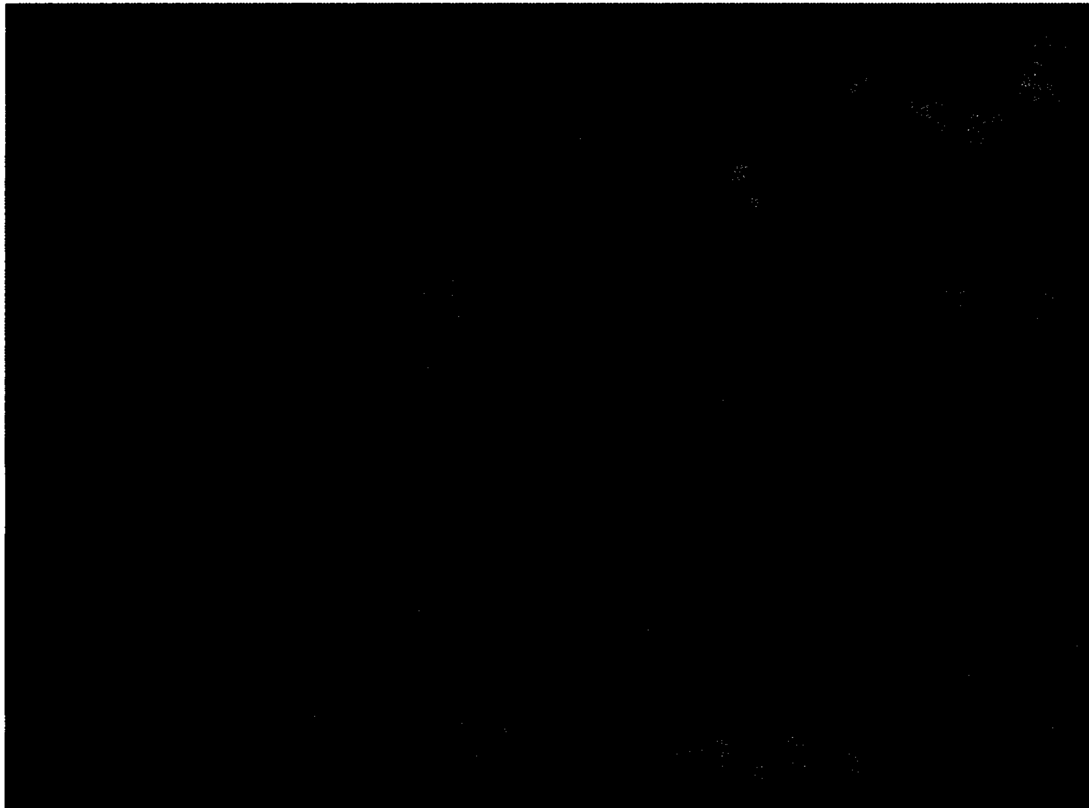


FIG. 6B

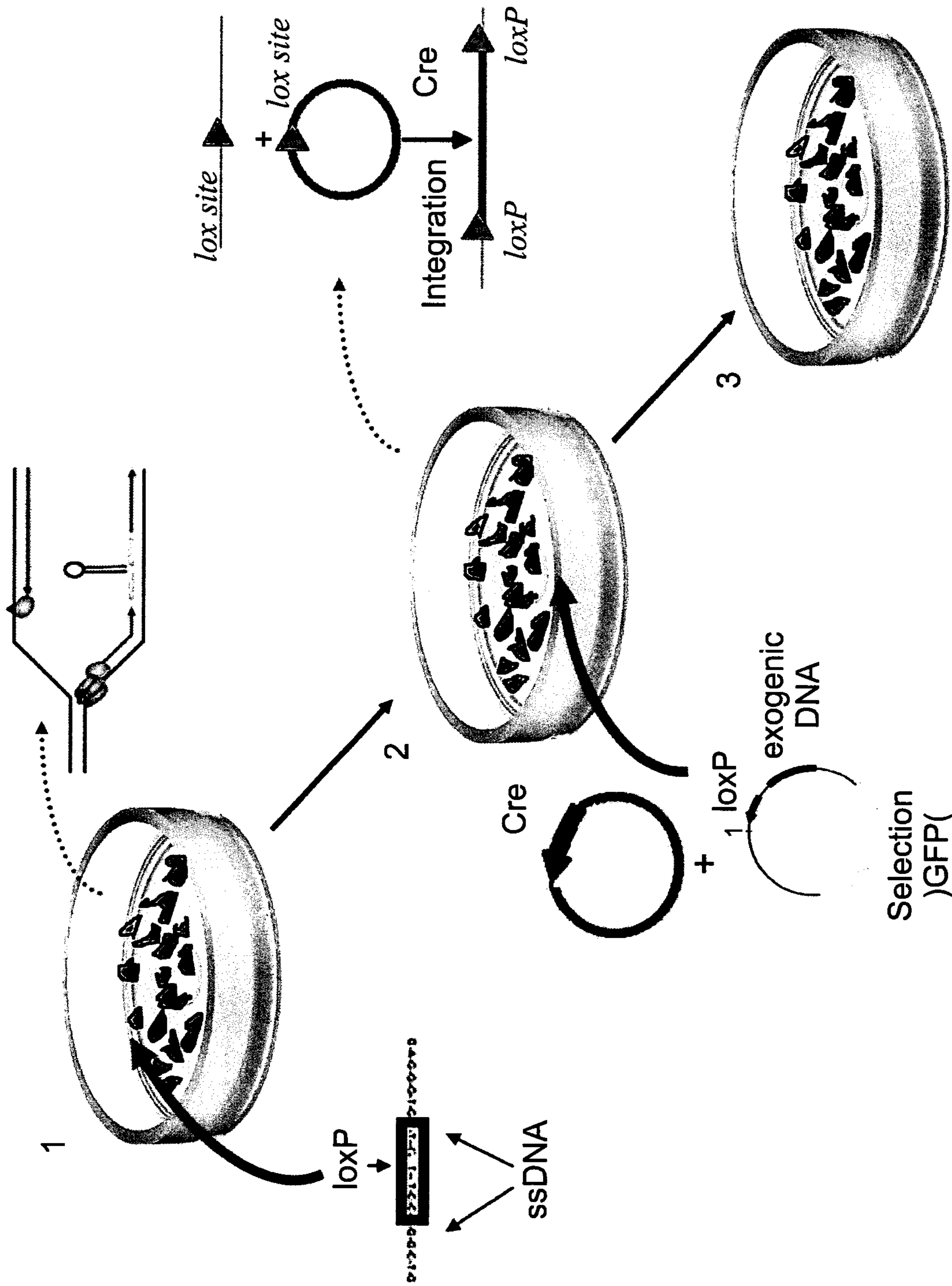


FIG. 7

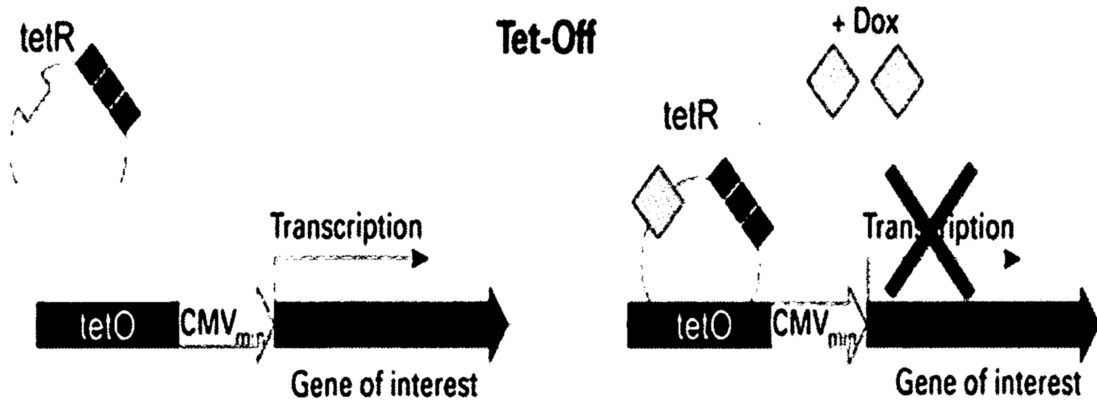


FIG. 8A

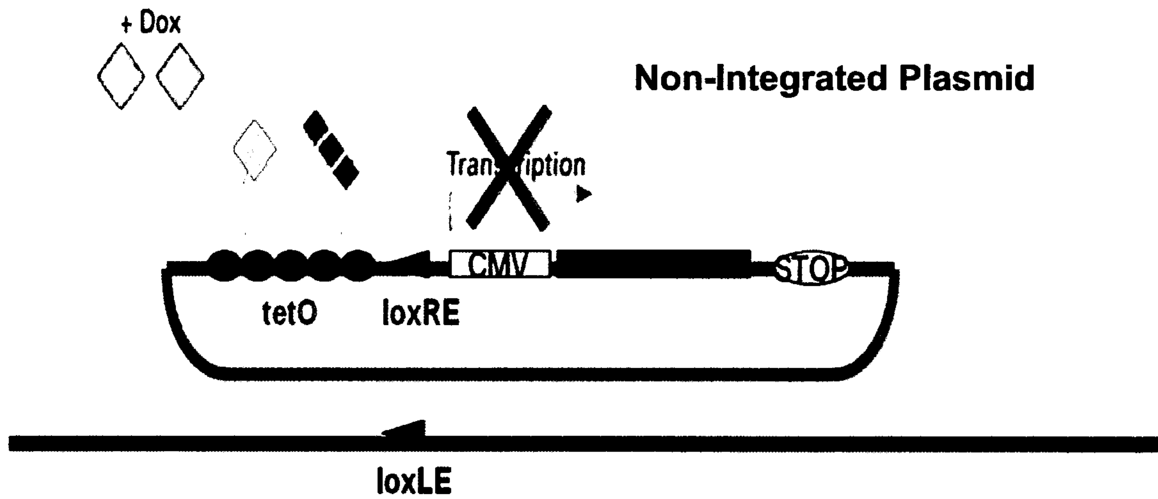


FIG. 8B

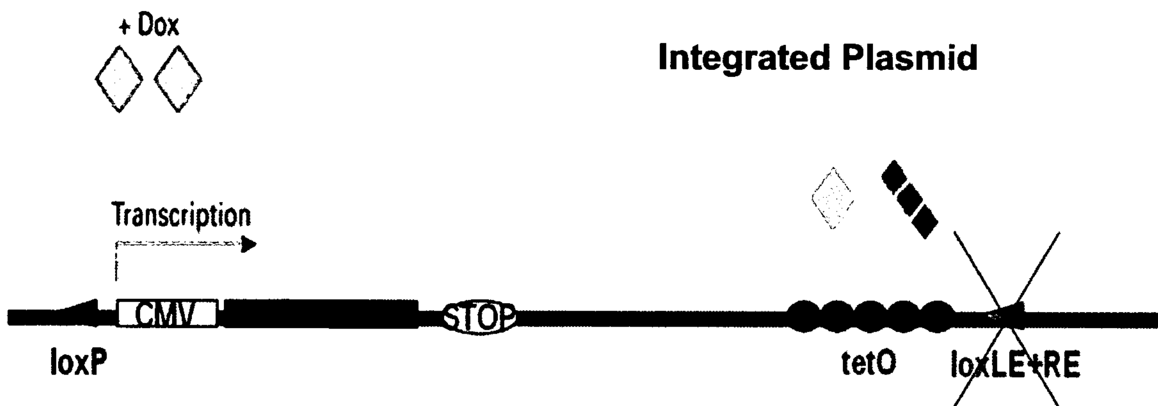


FIG. 8C

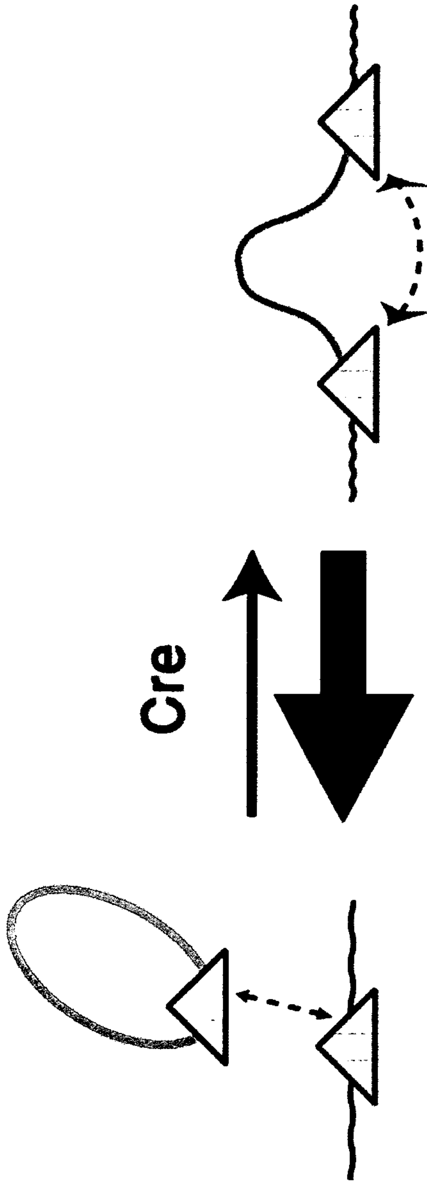


FIG. 9A

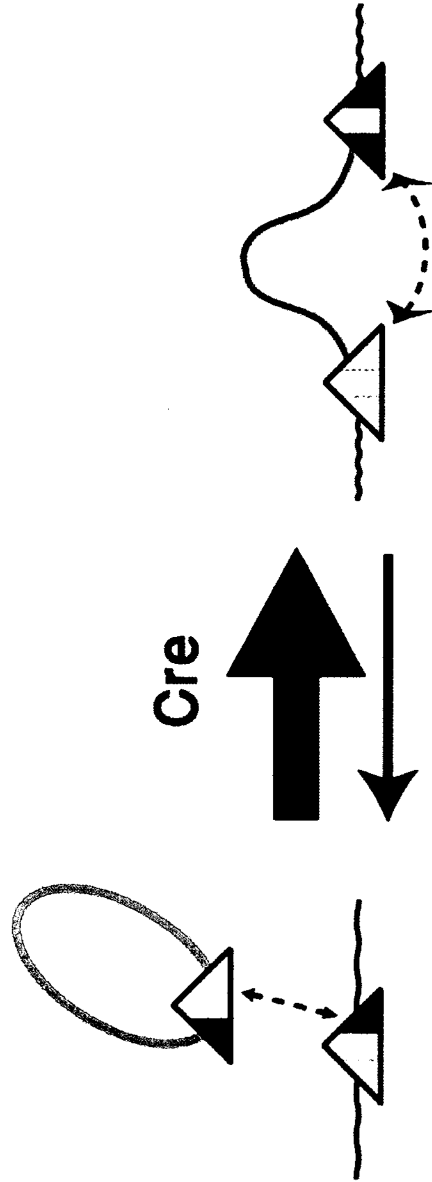


FIG. 9B

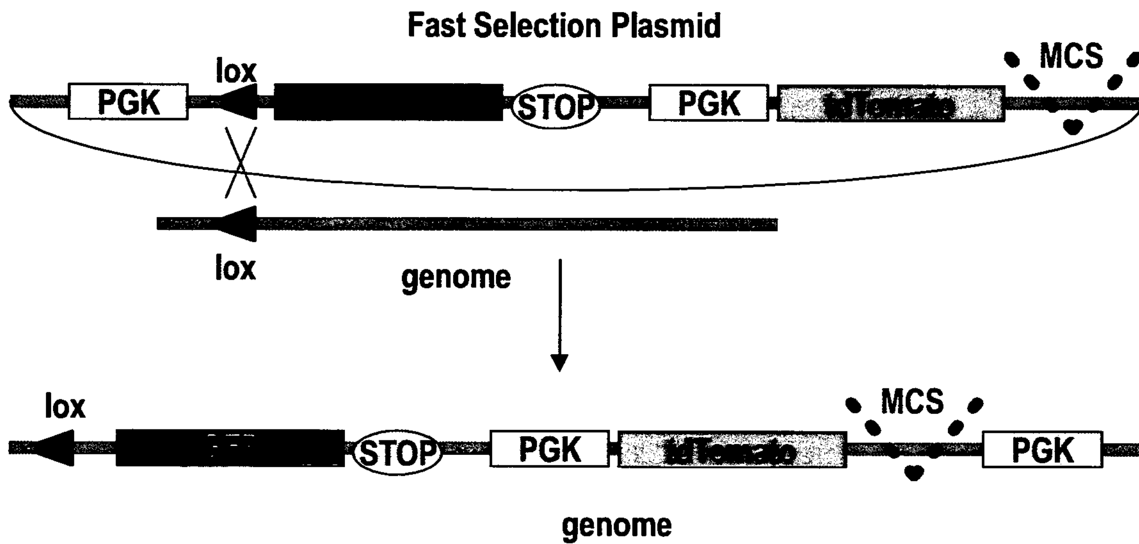


FIG. 10A

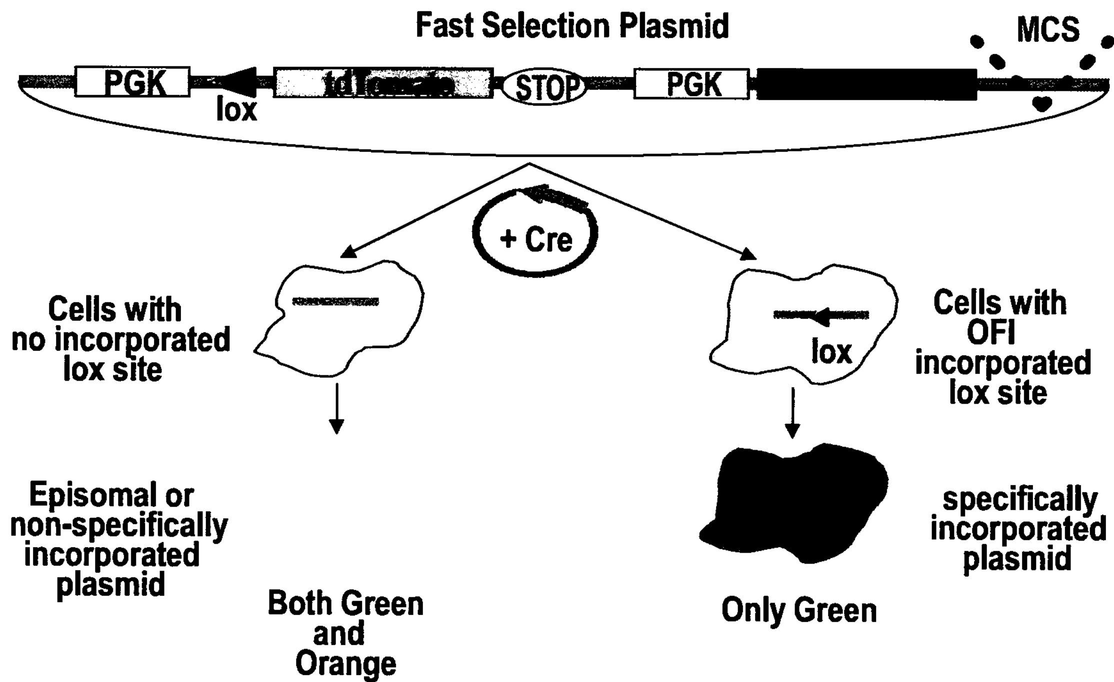


FIG. 10B

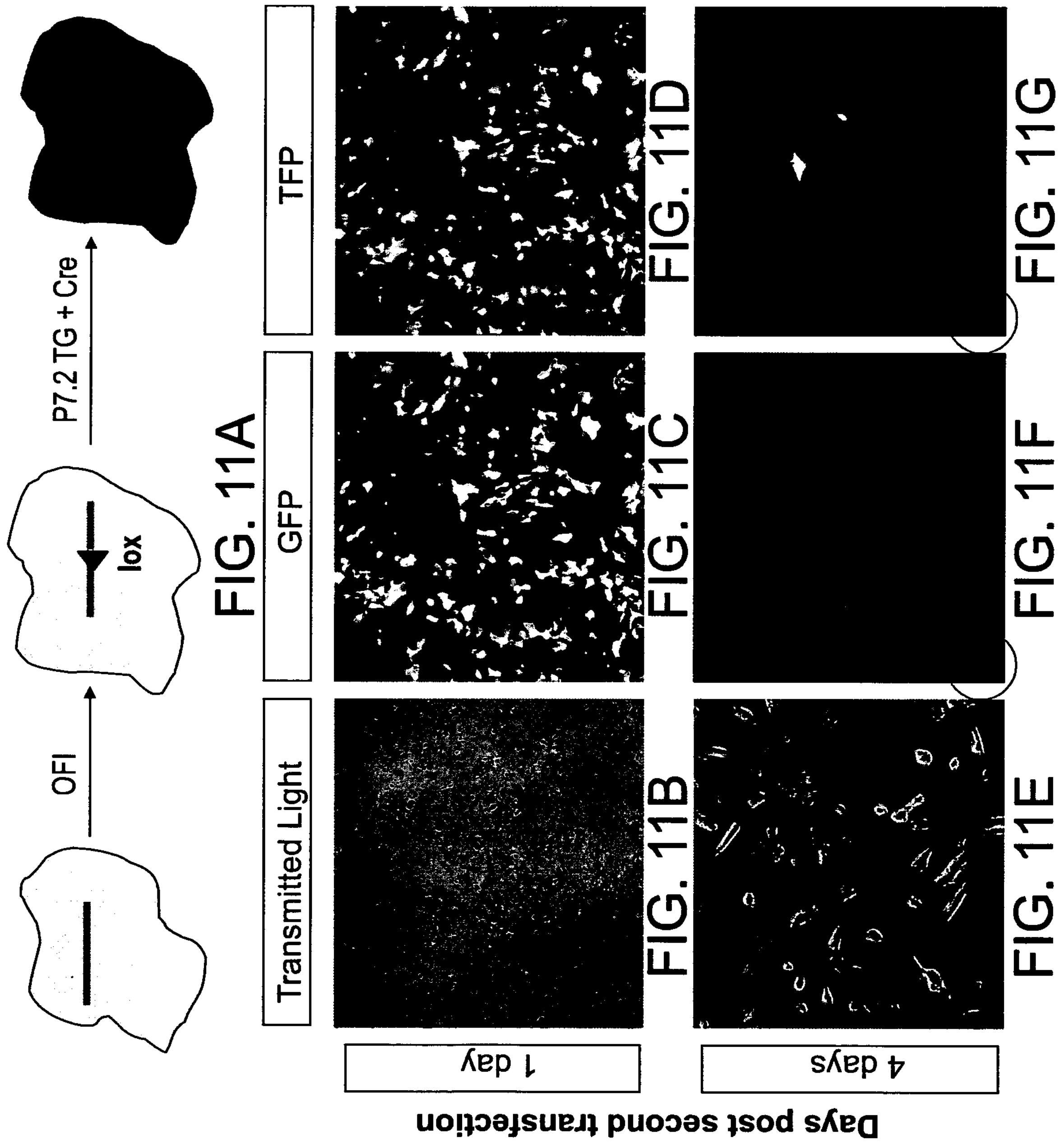


FIG. 11A

FIG. 11B

FIG. 11C

FIG. 11D

FIG. 11E

FIG. 11F

FIG. 11G

FIG. 12A

FIG. 12B

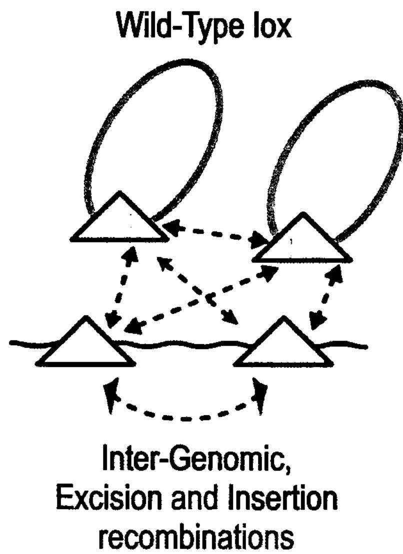


FIG. 13A

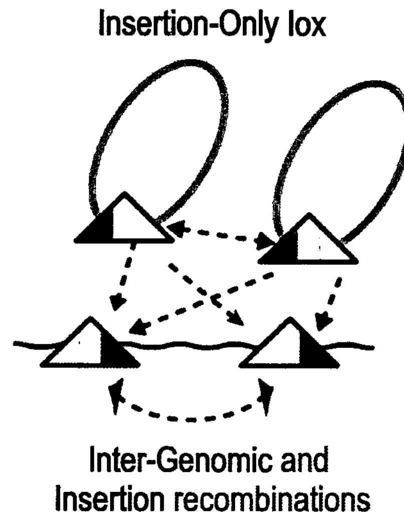


FIG. 13B

Self Non-Recombinating lox

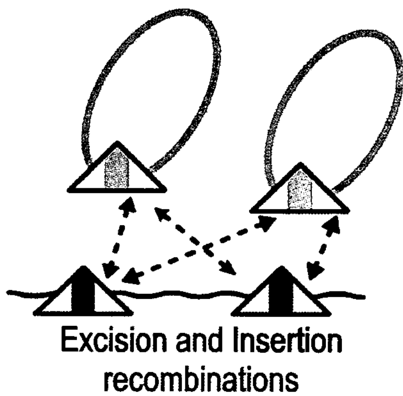


FIG. 13C

Self Non-Recombinating,
Insertion Only lox

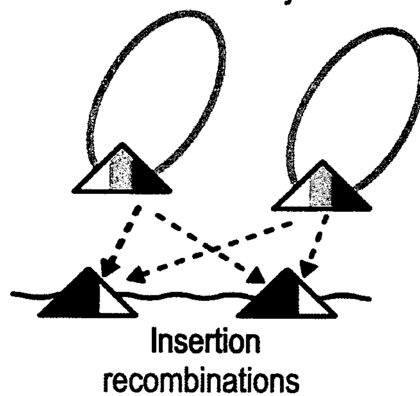


FIG. 13D

Self Non-Recombinating, Insertion-Only lox

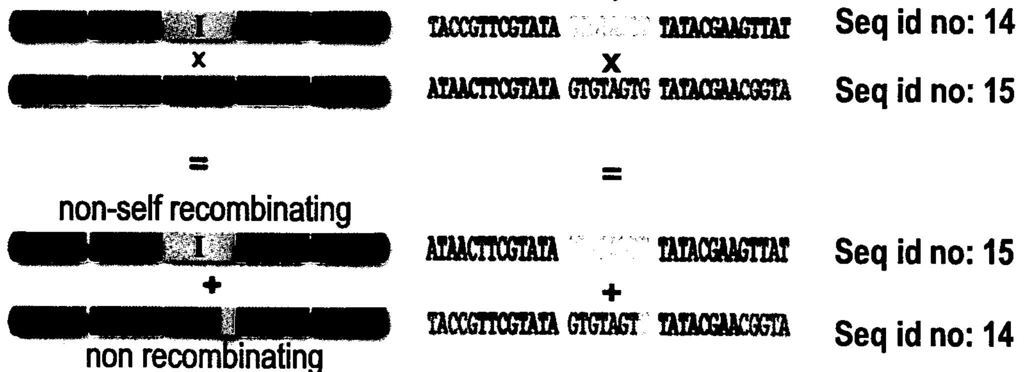


FIG. 13E

FIG. 14A

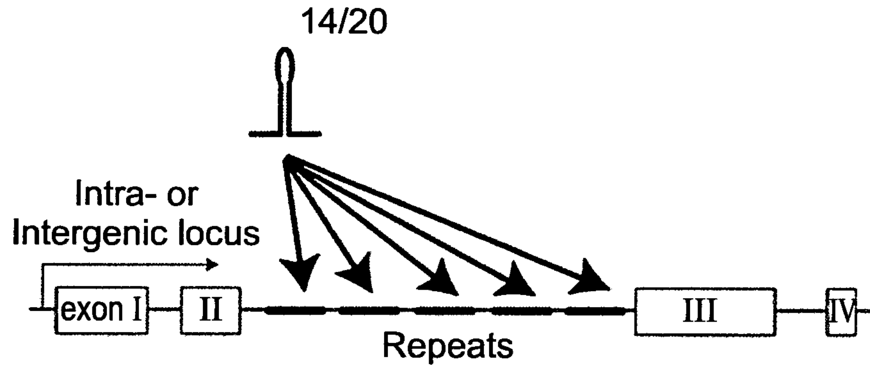


FIG. 14B

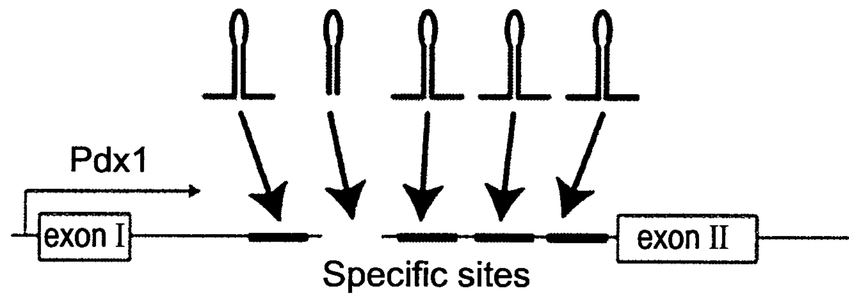


FIG. 14C

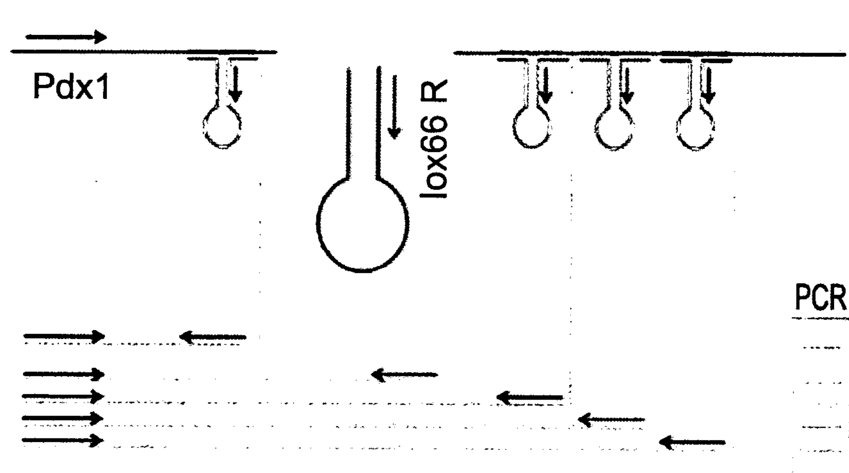
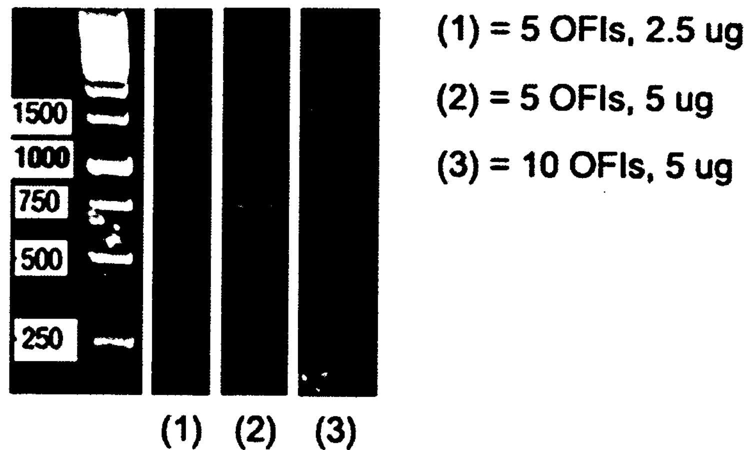


FIG. 14D



15/20

```
1    NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNN
51              GCTGGCTC TCAGGTTGGG ACTAATTACG ATATAACTAT
101  AGAGAGAGGA AACACATGGT CAGATATAAC AAAATGTGTC ACAGTCTCCA
151  TTAGCACAAA GATTTTCAAA CTGCAGGTTG CACCCATTCG CAGGTCATAA
201  AATCAATTTA CTACGTTGAG AT
251             GTAGATAG CATCATACGT GGTAAACGTT TGTTTTATGT
301  CCTTAAGATT TGTCAGTATA AGTGACCTGC AGTGTCCGTG TGTGAACTAC
351  ACAACGATCC GAAATGTATT TCTCACATTT GTGGGTCACC ATCAGGAGGT
401  TTTTTTTAGC CCTGGATTAA AGGCGTTTTA TTGCCTTTGT AGGATCCAGC
451  CGGTTTAACT ATTATATACA TTTAAATCAA CATCAATCAG TTGATTAACA
```

Y58
Y59
Y57_second_half
lox66_f

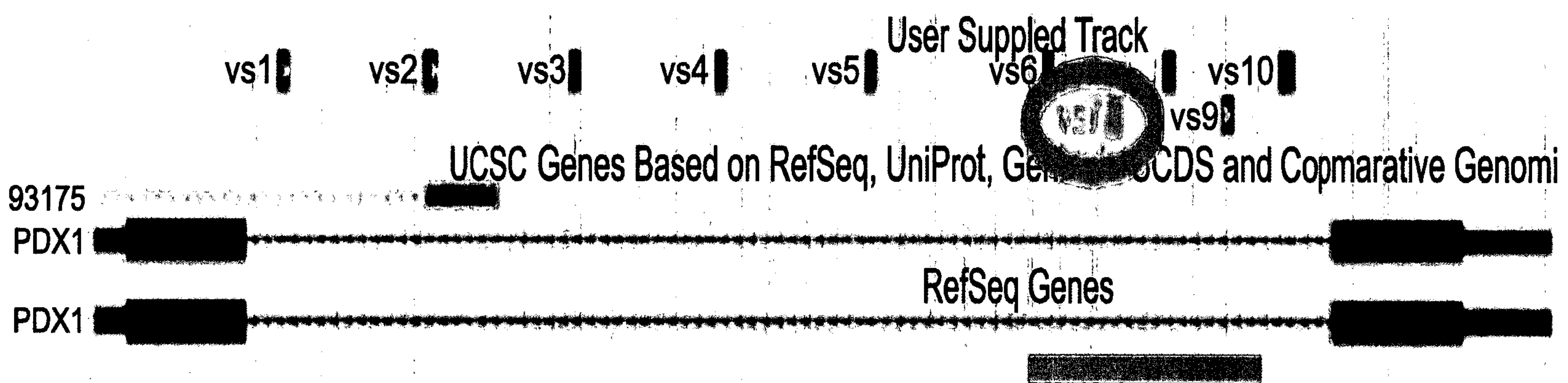


FIG. 14E

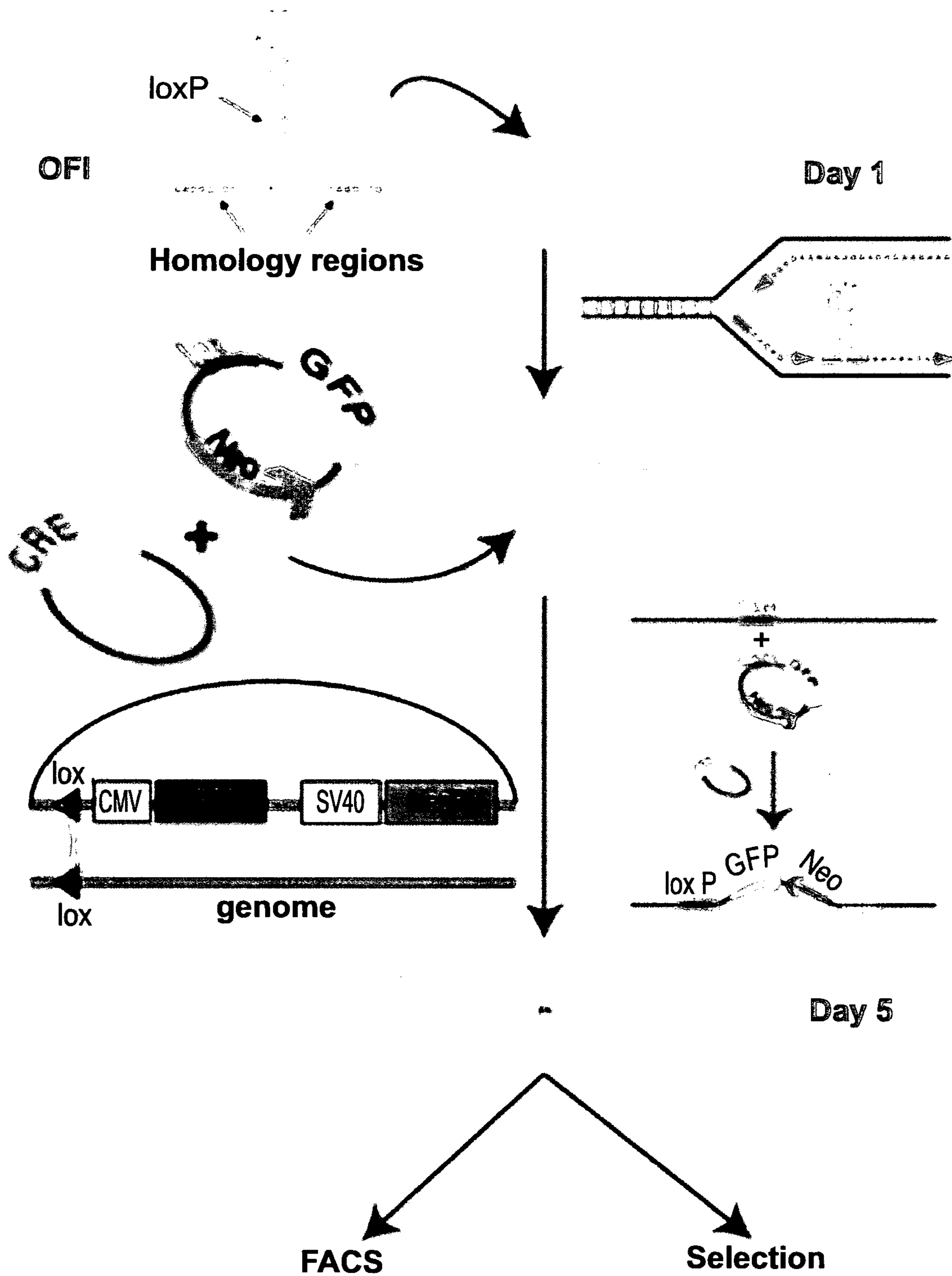


FIG. 14F

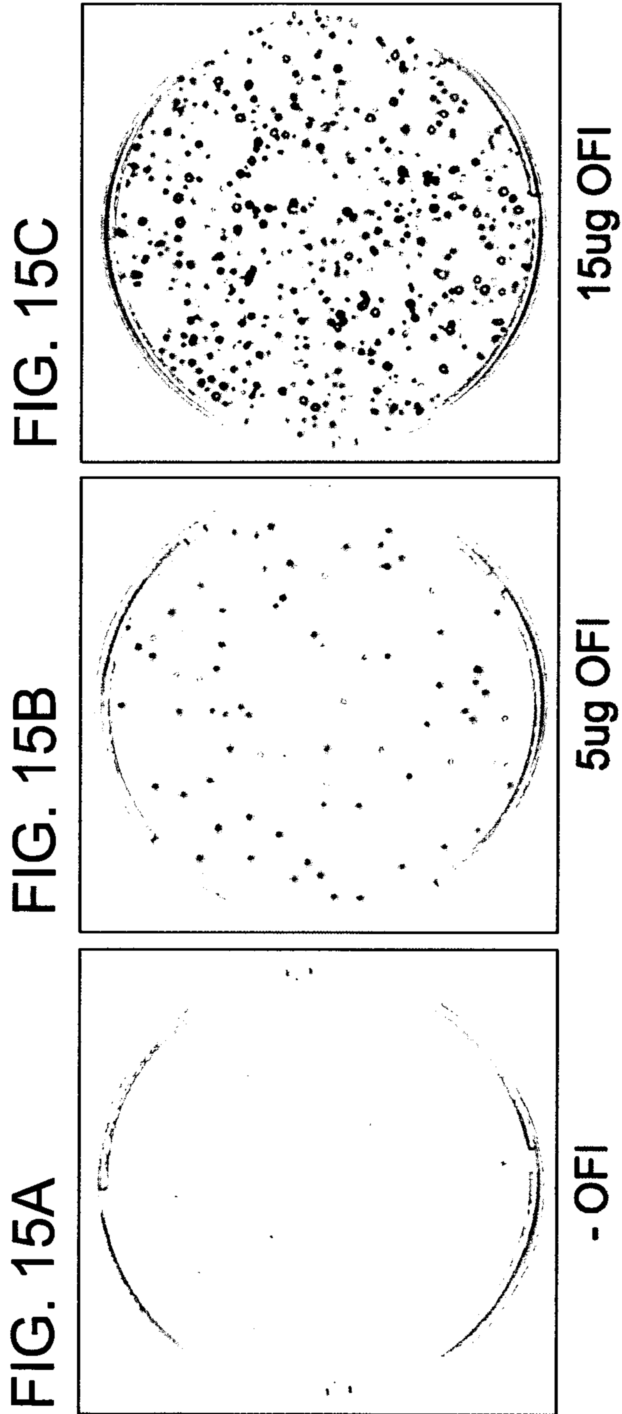


FIG. 15A FIG. 15B FIG. 15C

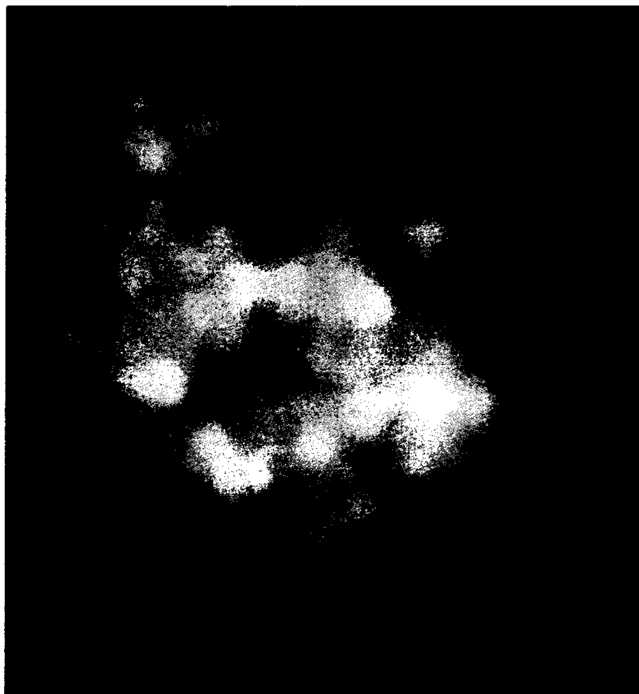


FIG. 16B

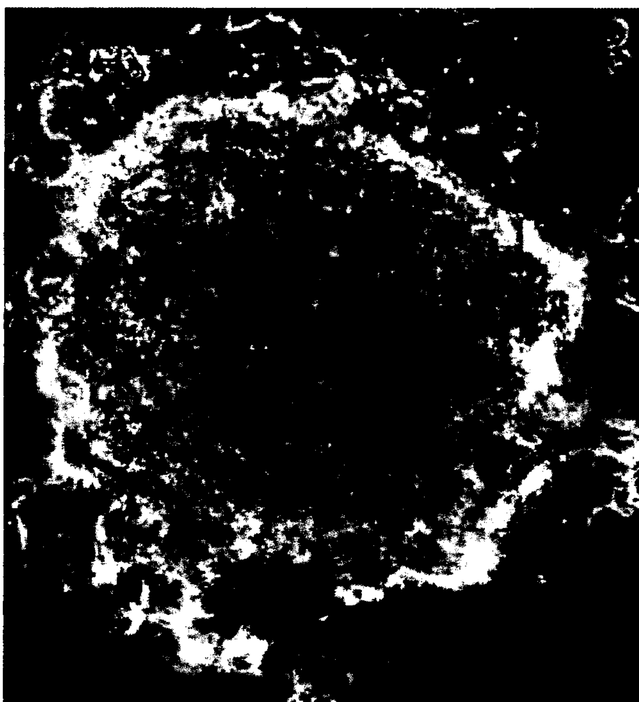


FIG. 16A

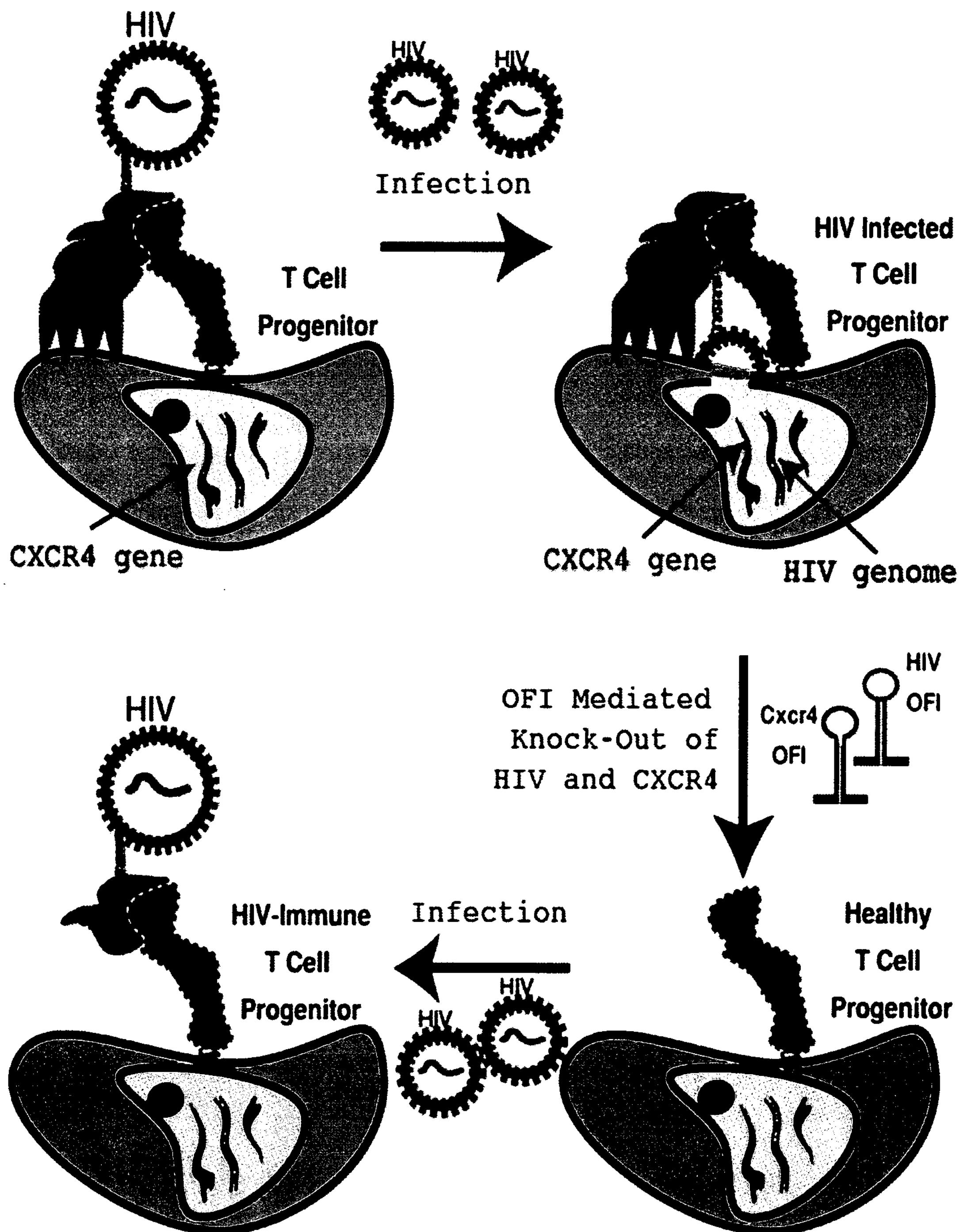
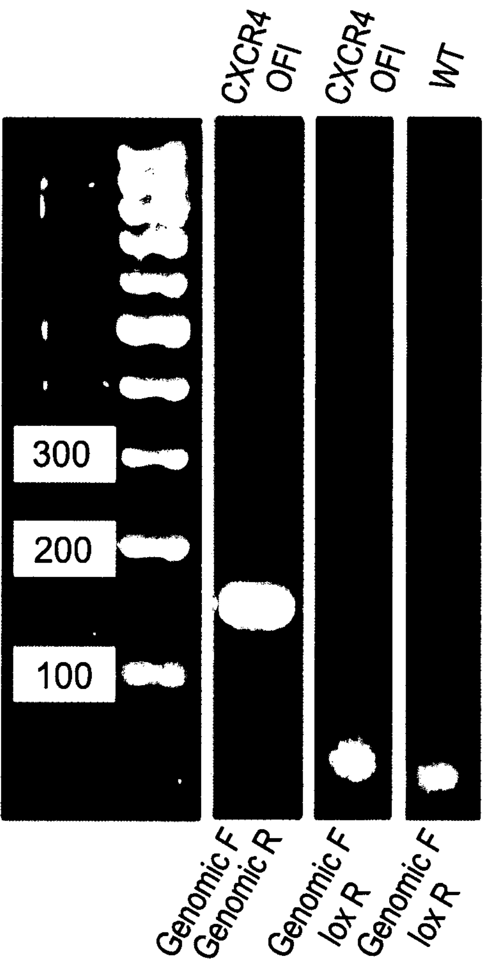


FIG. 17A




```
1    GG TCGGCCACTG ACAGGTGCAG CCTGTACTTG
51   TCCGTCATGC TTCTCAGTTT CTTCTGGTAA CCCATGACAT AACTTCGTAT
101  AGCATACATT ATACGAACGC TACAGGATGA CCAATCCATT GCCCTCAATG
151  CCAGTT AGA AGATGATGGA GTAGATGGTG GGC
```

lox66_f
cxcr4_OFB_F
cxcr4_OFB_R

FIG. 17B

INTERNATIONAL SEARCH REPORT

International application No
PCT/IL2010/000934

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N15/10 C12N15/113
ADD.

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12N

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

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

EPO-Internal, Sequence Search

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2009/104094 A2 (GENE BRIDGES GMBH [DE]; MARESCA MARCELLO [DE]; ERLER AXEL STEFFEN [DE]) 27 August 2009 (2009-08-27)	1-9, 35-38
Y	in particular see claims and pages 6-12, 36-37 and 42-44; the whole document	1-67

X	WO 2004/015117 A2 (NL KANKER I [NL]; TE RIELE HENRICUS PETRUS JOSEP [NL]) 19 February 2004 (2004-02-19)	1-9
Y	in particular see claims and pages 1-3; the whole document	1-67

X	US 2004/014057 A1 (KMIEC ERIC B [US] ET AL) 22 January 2004 (2004-01-22)	1
Y	in particular see pages 3-6; the whole document	1-67

	-/--	

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
7 March 2011	25/03/2011

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Vix, Olivier

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 03/062435 A1 (MORPHOTEK INC [US]; GRASSO LUIGI [US]; KLINE J BRADFORD [US]; NICOLAID) 31 July 2003 (2003-07-31) in particular see claims 1, 16, 24 and pages 6-10; the whole document	1-67
Y	EP 1 229 113 A2 (INVITROGEN CORP [US]) 7 August 2002 (2002-08-07) in particular see claims and pages 14-15 and 9-10	1-67
Y	IGOUCHEVA O ET AL: "Targeted gene correction by small single-stranded oligonucleotides in mammalian cells", GENE THERAPY, MACMILLAN PRESS LTD., BASINGSTOKE, GB, vol. 8, no. 5, 1 March 2001 (2001-03-01), pages 391-399, XP001206900, ISSN: 0969-7128, DOI: DOI:10.1038/SJ.GT.3301414 the whole document	1-67
X	ZORIN B ET AL: "NUCLEAR-GENE TARGETING BY USING SINGLE-STRANDED DNA AVOIDS ILLEGITIMATE DNA INTEGRATION IN CHLAMYDOMONAS REINHARDTII", EUKARYOTIC CELL, AMERICAN SOCIETY FOR MICROBIOLOGY, US, vol. 4, no. 7, 1 July 2005 (2005-07-01), pages 1264-1272, XP009053638, ISSN: 1535-9778, DOI: DOI:10.1128/EC.4.7.1264-1272.2005	1
Y	the whole document	1-67
Y	OLSEN PETTER ANGELL ET AL: "Genomic sequence correction by single-stranded DNA oligonucleotides: role of DNA synthesis and chemical modifications of the oligonucleotide ends.", THE JOURNAL OF GENE MEDICINE DEC 2005 LNKD- PUBMED:16025558, vol. 7, no. 12, December 2005 (2005-12), pages 1534-1544, XP002626342, ISSN: 1099-498X the whole document	1-67

-/--

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	ARAKI K ET AL: "TARGETED INTEGRATION OF DNA USING MUTANT LOX SITES IN EMBRYONIC STEM CELLS", NUCLEIC ACIDS RESEARCH, OXFORD UNIVERSITY PRESS, SURREY, GB, vol. 25, no. 4, 1 January 1997 (1997-01-01), pages 868-872, XP002934226, ISSN: 0305-1048, DOI: DOI:10.1093/NAR/25.4.868 the whole document -----	1-67
X	SHANG XI-YING ET AL: "Improvement of SSO-mediated gene repair efficiency by nonspecific oligonucleotides.", BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS 7 NOV 2008 LNKD- PUBMED:18771655, vol. 376, no. 1, 7 November 2008 (2008-11-07), pages 74-79, XP002626343, ISSN: 1090-2104	1
Y	the whole document -----	1-67
X	OLSEN P A ET AL: "Implications of cell cycle progression on functional sequence correction by short single-stranded DNA oligonucleotides.", GENE THERAPY MAR 2005 LNKD- PUBMED:15674399, vol. 12, no. 6, March 2005 (2005-03), pages 546-551, XP002626344, ISSN: 0969-7128 the whole document -----	1
Y	MISSIRLIS PERSEUS I ET AL: "A high-throughput screen identifying sequence and promiscuity characteristics of the loxP spacer region in Cre-mediated recombination", BMC GENOMICS, BIOMED CENTRAL, LONDON, GB, vol. 7, no. 1, 4 April 2006 (2006-04-04), page 73, XP021014674, ISSN: 1471-2164, DOI: DOI:10.1186/1471-2164-7-73 the whole document -----	1-67

-/--

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	SAUER B: "Inducible Gene Targeting in Mice Using the Cre/loxSystem", METHODS : A COMPANION TO METHODS IN ENZYMOLOGY, ACADEMIC PRESS INC., NEW YORK, NY, US, vol. 14, no. 4, 1 April 1998 (1998-04-01), pages 381-392, XP004466618, ISSN: 1046-2023, DOI: DOI:10.1006/METH.1998.0593 the whole document -----	1-67
X	MOROZOV V ET AL: "Single-strand DNA-mediated targeted mutagenesis of genomic DNA in early mouse embryos is stimulated by Rad51/54 and by Ku70/86 inhibition.", GENE THERAPY MAR 2008 LNKD- PUBMED:18079752, vol. 15, no. 6, March 2008 (2008-03), pages 468-472, XP002626721, ISSN: 1476-5462 see abstract and page 470-471; the whole document -----	35-38
Y	KANG JEE HYUN ET AL: "Gene targeting in mouse embryos mediated by RecA and modified single-stranded oligonucleotides.", IN VITRO CELLULAR & DEVELOPMENTAL BIOLOGY. ANIMAL 2008 MAR-APR LNKD- PUBMED:18266050, vol. 44, no. 3-4, March 2008 (2008-03), pages 57-62, XP002626722, ISSN: 1071-2690 the whole document -----	1-67
X,P	KAMEYAMA YUJIRO ET AL: "An accumulative site-specific gene integration system using Cre recombinase-mediated cassette exchange.", BIOTECHNOLOGY AND BIOENGINEERING 15 APR 2010 LNKD- PUBMED:19998287, vol. 105, no. 6, 15 April 2010 (2010-04-15), pages 1106-1114, XP002626345, ISSN: 1097-0290 the whole document -----	1-67

-/--

INTERNATIONAL SEARCH REPORT

International application No
PCT/IL2010/000934

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MITAS M: "Trinucleotide repeats associated with human disease", NUCLEIC ACIDS RESEARCH, OXFORD UNIVERSITY PRESS, SURREY, GB, vol. 25, no. 12, 1 January 1997 (1997-01-01), pages 2245-2253, XP002484842, ISSN: 0305-1048, DOI: DOI:10.1093/NAR/25.12.2245 the whole document -----	1-67

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/IL2010/000934

Patent document cited in search report		Publication date	Patent family member(s)			Publication date
WO 2009104094	A2	27-08-2009	AU	2009215361	A1	27-08-2009
			CA	2715871	A1	27-08-2009
			EP	2255001	A2	01-12-2010
WO 2004015117	A2	19-02-2004	AU	2003253197	A1	25-02-2004
US 2004014057	A1	22-01-2004	NONE			
WO 03062435	A1	31-07-2003	AU	2003205173	B2	22-05-2008
			CA	2473741	A1	31-07-2003
			EP	1474522	A1	10-11-2004
			US	2003176386	A1	18-09-2003
EP 1229113	A2	07-08-2002	EP	1227147	A2	31-07-2002
			SI	937098	T1	28-02-2003