

Meganucleases and Other Tools for Targeted Genome Engineering: Perspectives and Challenges for Gene Therapy

George Silva², Laurent Poirot¹, Roman Galetto¹, Julianne Smith¹, Guillermo Montoya³, Philippe Duchateau² and Frédéric Pâques^{1,2,*}

¹Collectis Genome Surgery, 102 Avenue Gaston Roussel, 93 235 Romainville, Cedex, France; ²Collectis, 102 Avenue Gaston Roussel, 93 235 Romainville, Cedex, France; ³Macromolecular Crystallography Group, Structural Biology and Biocomputing Programme, Spanish National Cancer Centre (CNIO), Melchor Fdez. Almagro 3, 28029 Madrid, Spain

Abstract: The importance of safer approaches for gene therapy has been underscored by a series of severe adverse events (SAEs) observed in patients involved in clinical trials for Severe Combined Immune Deficiency Disease (SCID) and Chronic Granulomatous Disease (CGD). While a new generation of viral vectors is in the process of replacing the classical gamma-retrovirus-based approach, a number of strategies have emerged based on non-viral vectorization and/or targeted insertion aimed at achieving safer gene transfer. Currently, these methods display lower efficacies than viral transduction although many of them can yield more than 1% engineered cells *in vitro*. Nuclease-based approaches, wherein an endonuclease is used to trigger site-specific genome editing, can significantly increase the percentage of targeted cells. These methods therefore provide a real alternative to classical gene transfer as well as gene editing. However, the first endonuclease to be in clinic today is not used for gene transfer, but to inactivate a gene (CCR5) required for HIV infection. Here, we review these alternative approaches, with a special emphasis on meganucleases, a family of naturally occurring rare-cutting endonucleases, and speculate on their current and future potential.

Keywords: Homing endonuclease, Zinc-finger nuclease, Recombinase, Transposons, gene transfer, protein engineering, viral vector.

INTRODUCTION TO TARGETED APPROACHES

Different Strategies for Safer Gene Therapy

At the turn of the last millennium, the successful treatment of several X-SCID patients was a major milestone in the field of gene therapy [1, 2]. However, the importance of safer approaches has been emphasized by a series of severe adverse events (SAEs); e.g. the appearance of leukemia in X-SCID treated patients [3-5] and of myelodysplasia in a clinical trial for Chronic Granulomatous Disease (CGD) [6]. These SAEs were all associated with the integration of the transgene next to a proto-oncogene, and continued efforts have been devoted to the development of new, safer vectors. One such approach involves self-inactivating (SIN) gamma-retrovirus and lentivirus-based vectors, which when integrated lack enhancer and promoter sequences from the LTR [6-13]. These new vectors should decrease or alleviate the risks of proto-oncogene activation by insertional mutagenesis [14]. Meanwhile, these same SAEs fostered a growing interest for alternative strategies, which we will refer to as "targeted approaches" that consist of replacing random integration of therapeutic transgenes by targeted insertion or even correction of the deleterious mutation. Results are beginning to accumulate for clinical trials using new viral vectors [8, 15], and the first clinical trial using a targeted approach has been initiated [16]. Although additional time will

be necessary to evaluate more precisely the potentials and limits of each method, targeted approaches remain today, at least conceptually, more attractive than classical gene transfer.

Targeted Approaches: A Prosperous Methodology

Homologous gene targeting, first described in yeast by the laboratories of Gerald Fink [17], Jack Szostak [18-20] and, in its ultimate design, by Rodney Rothstein [21], was one of the first methods for rational genome engineering. This technique, successfully transferred to mammalian cells by the laboratories of Mario Capecchi [22-24] and Oliver Smithies [25-27], remains to this day a standard for the generation of engineered cells or knock-out mice [28]. An inherently low efficiency has nevertheless prevented it from being used as a routine protocol in most cell types and organisms. As homologous gene targeting can typically be observed in only 10^{-6} to 10^{-9} of treated mammalian cells [29], such frequencies did not obviate its use for gene and cell therapy. To address these issues, an extensive assortment of rational approaches has been proposed with the intent of achieving greater than 1% targeted modifications. The proven benefit of these emerging approaches lies in their ability to engineer the genome of a variety of cell types (including human stem cells). Nevertheless, researchers have invariably tried to assess their potential for therapeutic applications. In the therapeutic field, the replacement of classical gene transfer (resulting in random integration into the genome) with targeted gene insertion, or even targeted correction, has become a

*Address correspondence to this author at the Collectis, 102 Avenue Gaston Roussel, 93 235 Romainville, Cedex, France; Tel: +33 1 41 83 99 00; Fax: +33 1 41 83 99 03; E-mail: paques@collectis.com

commonly accepted potential solution and thus the goal of many laboratories.

Targeted Approaches Based on Recombinases

Several targeting strategies rely on site-specific recombination processes mediated by recombinases or transposases. Such methods have an inherent advantage since these proteins catalyze all the integration steps in a practically autonomous manner (e.g. transposases may require host factors). Their activity thus does not depend on the homologous recombination machinery. Homologous recombination is a complex maintenance system, and many studies argue that it is essentially active in the late S and G2 phases of the cell cycle [30-37]. As a consequence, the efficacy of recombinase or transposase-mediated integration should be relatively cell-type independent, whereas methods based on homologous gene targeting would be very inefficient in post-mitotic or quiescent cells.

The use of the *Streptomyces* phage Φ C31 integrase to incorporate transgenes into the genome of mammalian cells provides an early example of a site-specific recombination approach. The Φ C31 integrase is a serine recombinase that catalyzes the insertion of the phage genome into the bacterial genome via recombination between the phage attachment site (attP) and the bacterial attachment site (attB) [38, 39]. However, as pseudo integration sites can be found in mammalian cells, it was also possible to induce significant integrase-mediated insertion in specific loci in mammalian chromosomes [40]. In subsequent experiments, this technology was used to insert transgenes not only in a series of immortalized cell lines but also primary cells [41-44], including human embryonic stem (hES) cells [41, 45]. The system proved robust enough for *in vivo* use, wherein co-injection of a Φ C31 integrase expression cassette with a plasmid carrying an attB site resulted in high frequencies of integration coupled with stable transgene expression in various tissues, such as mouse muscles [46, 47], rabbit joints [48], rat retina [49] and mouse liver [50, 51]. Using hydrodynamic injection, Olivares and colleagues observed integration in liver tissue at two major sites, one of which had detectable incorporation of a Factor IX transgene in 0.6 to 3.6% of total hepatocytes [51]. This strategy is limited, however, by an inability to choose the integration site. Whereas the most favored unique landing sequence occurred at chromosome 19q13.31, circa one hundred other integration sites have been identified [52]. Recently, mutated derivatives of Φ C31 with enhanced efficiency and specificity have been reported [53], offering great promise for success in the near term.

The Cre and FLP tyrosine recombinases have also been envisioned as reagents to mediate targeted insertions. These proteins have been widely used in genome engineering experiments, albeit always to recombine transgenes carrying their respective target sequences. To this end, several attempts have been made to redesign recombinase site-specificity [54-57], the most remarkable involving a modified recombinase able to recognize sequences from the HIV1 virus and used to excise a provirus from the human genome [54]. Whereas these studies were based on the redesign of the recombinase substrate specificity, an alternative strategy was used by other groups who fused the catalytic domain of the Tn3 resolvase or the Gin recombinase to a Zinc-finger-

based DNA-binding domain [58-60]. However, in these initial proof-of-concept designs the recombinase catalytic domain was still contributing to target preference. Using additional engineering steps, Gersbach and colleagues could suppress the target preference of the Gin moiety, thus producing a chimeric recombinase having essentially the selectivity of the chosen Zinc-finger DNA-binding domain [61].

Targeted Approaches Based on Transposons

The stable introduction of therapeutic transgenes into human cells can be accomplished by transposon-mediated gene transfer and has been proposed as an alternative to viral-mediated gene delivery [62-64]. As transposases can efficiently catalyze gene transfer, transposition has long been used in *Drosophila* [65] to create transgenics. They have also widely been used in other organisms such as plants, fish, bacteria and others [66-68] in mutagenesis experiments. The inability of transposons to introduce a cargo of genetic material into a cell in an autonomous manner is both an advantage and a drawback: unlike viruses they do not present risks of lateral transfer yet one needs to use them in association with a vector to introduce them into the cell. Several transposons have nonetheless been used in gene transfer experiments in vertebrate cells. The most advanced system is perhaps that derived from the Sleeping Beauty (SB) transposon, an ancient mobile genetic element in fish that could be modified to sustain elevated levels of activity in mammalian cells, with up to 5-10% integration [69, 70]. In similar experiments, transposons based on the Piggy Back system were used to make induced pluripotent stem (iPS) cells [71, 72]. Importantly, the efficacy of transposase-mediated gene insertion can be extremely high, and as a consequence high rates of gene transfer can be observed even with a suboptimal vectorization method. For example, the association of electroporation methods and of the SB100X enhanced transposase results in an efficacy of gene insertion of up to 50% in CD34+ blood cells [70]. Transposons remain non-targeted gene transfer vehicles, and as such several attempts have been made to endow them with sequence specificity by fusion with a DNA binding domain. Initial attempts with the SB transposon resulted in an enrichment for targeted integration in human cells [73]. More recently, the ISY100 transposase from a cyanobacteria transposon was fused with a Zinc-finger DNA binding domain (Zif268) and used to efficiently target a specific site in *E. coli*. [74]. While promising as a proof-of-concept, the activity of this transposon in eukaryotic cells is unknown. In terms of production, one of the major drawbacks of viral vectors is the need for a tedious manufacturing process, a difficulty compounded in the case of therapeutic applications where Good Manufacturing Practice (GMP) is compulsory. The ability to achieve similar efficacies while circumventing these issues is a significant plus for transposon-based approaches. A clinical trial using SB-mediated transfer of a chimeric antibody receptor (CAR) for the treatment of leukemia has recently been approved [75, 76].

Targeted Approaches based on Homologous Recombination

In parallel to the development of methods based on site-specific recombination, many groups have focused on en-

hancing the efficacy of homologous gene targeting. Two major disciplines have become apparent: (i) methods that we will call “matrix optimization”, essentially consisting of modifying the targeting vector structure to achieve maximal efficacy, and; (ii) methods involving additional effectors to stimulate homologous recombination (HR), generally sequence-specific endonucleases. The field of matrix optimization has covered a wide range of techniques, with varying degrees of success. For example, attempts to replace the DNA repair matrices used in classical gene targeting experiments by chimeric DNA-RNA modified oligonucleotides, including various modifications and/or specific secondary structures, were initially seen as promising yet eventually proved inefficient in many experiments [77].

Another type of optimization, Small Fragment Homologous Replacement (SFHR), relies on information exchange between short DNA fragments homologous to the targeted sequence and the endogenous locus [78]. The small size of the DNA fragment limits this approach to gene correction and does not allow for gene insertion or gene replacement experiments. All the same, SFHR was successfully applied to correct various mutations associated with monogenetic disorders. Very different frequencies have been reported, and a number of publications report correction of the human HPRT, CFTR, HBB and DMD genes in cell lines and primary cells, with frequencies in the range of 1% or even 10% of transfected cells [78-80]. A recent report describes correction of the CFTR gene in up to 60% of mouse ES cells [81].

The use of Adeno-Associated Virus (AAV) as vectors for the repair matrix provides another way to increase HGT rates. Homologous gene targeting between an AAV vector and a chromosomal sequence can occur at frequencies in the range of 0.1-1.4 % in a variety of cell lines [82-84]. AAV-mediated gene inactivation was also used to disrupt the COL1A1 and COL1A2 genes in human mesenchymal stem cells (MSCs) [85, 86] with an efficacy of 0.06-0.23% per treated cell for COL1A1. In addition, AAV-mediated gene targeting has been used to correct a LacZ marker at a frequency of 0.1% *in vivo* in a mouse model [87]. Research continues towards optimized matrices, although as noted in the next section the use of effectors to stimulate HR is becoming more widespread.

Since the first reports of using the I-SceI meganuclease in mammalian cells, the practice of stimulating HR via nucleases has become standard in many fields [88, 89]. I-SceI itself has been applied to numerous experimental designs in a variety of cell types and organisms (see below, “Use of meganucleases for genome engineering”). In addition to natural endonucleases, other sequence specific nucleases have been investigated. Endonucleases for genome engineering have been classified into three groups by Pingoud and Silva [90]: chemical nucleases, ZFNs and meganucleases. This classification should be updated given the recent emergence of a fourth category, the TALE nucleases (see below).

Chemical endonuclease are synthetic endonucleases resulting from the fusion of DNA-reactive agents to a DNA-binding polymer, such as triplex-forming oligonucleotides (TFO) and can in theory provide a straightforward solution (for review see [91]). Because the sequence specificity of such polymers can be designed *a priori*, it eliminates the

extensive screening process usually necessary for the identification of a new peptidic DNA-binder. A variety of active domains such as a topoisomerase inhibitor (camptothecin) [92], psoralen [93], bypyridine [94], or a restriction enzyme [95] in fusion with TFO have been described as potential sequence-specific tools for genome engineering. While in principle promising, intracellular delivery, efficiency of endogenous gene modifications as well as safety issues still need to be addressed for consideration in therapeutic applications.

To alleviate the delivery and safety issues associated with chemically-derived nucleases, attempts have been made to engineer protein endonucleases with programmable specificities. Among the most prevalent contenders are the Zinc-finger nucleases (ZFN), generated by fusing Zinc-finger-based DNA binding domains to an independent catalytic domain via a flexible linker [96-98]. The archetypal ZFNs are based on the catalytic domain of the Type IIS restriction enzyme FokI and have been successfully used to induce gene correction, gene insertion, and gene deletion. These studies have been reviewed by others [88, 99-102] and will not be described in detail. In summary, several groups have reported targeted HR events occurring at frequencies greater than 1% (and up to 50%) in immortalized human cell lines as well as primary cells including ES cells [103-109]. However, low efficacies have been reported in human ES, iPS [108] and hematopoietic stem cells (HSCs) [103]. Zinc Finger-based DNA binding domains are made of strings of 3 or 4 individual Zinc Fingers, each recognizing a DNA triplet [110]. In theory, one of the major advantages of ZFNs is that they are easy to design, using combinatorial assembly of preexisting Zinc Fingers with known recognition patterns [111-113]. However, close examination of high resolution structures shows that there are actually cross-talks between units [114], and several methods have been used to assemble ZF proteins by choosing individual Zinc Fingers in a context dependant manner [106, 115-117] to achieve better success rates and reagents of better quality.

Recently, a new class of chimeric nucleases using a FokI catalytic domain have been described [118, 119]. The DNA binding domain of these nucleases is derived from Transcription Activator Like Effectors (TALE), a family of proteins used in the infection process by plant pathogens of the *Xanthomonas* genus. In these DNA binding domains, sequence specificity is driven by a series of 33-35 amino acids repeats, differing essentially by two positions [120, 121]. In the DNA target, each base pair is contacted by a single repeat, the specificity resulting from the two variant amino acid of the repeat. The apparent modularity of these DNA binding domains has been confirmed to a certain extent by modular assembly of designed TALE-derived protein with new specificities [120, 121]. However, one cannot yet rule out a certain level of context dependence of individual repeat/base recognition patterns, as was observed for Zinc Finger proteins (see above). TALE-nucleases have been shown to be active in a cell-based assay in yeast [118, 119], and their exact level of engineerability, activity and specificity remains to be explored. Nevertheless, the emergence of this promising new family of endonucleases highlights the fact that the list of reagents continues to grow. The FokI effector appears capable of being fused to any convenient DNA bind-

ing domain, and the creation of functional I-SceI::FokI fusions [122] illustrates this potential diversity.

While ZFNs have shown considerable promise, there has been a continued interest in the natural meganucleases given the extensive use of I-SceI by a large community of scientists. Indeed, I-SceI has been used since the early nineties for genome engineering experiments and remains today the reference (“gold standard”) in the field with regard to activity and specificity. As the meganuclease family of which I-SceI is a member includes hundreds of other natural proteins, a special section has been included below (see “The Meganuclease Family”).

Targeted Approaches based on Non-Homologous End-Joining

In addition to the HR pathway for repairing double-strand breaks (DSBs) in DNA, an error-prone DNA repair mechanism called non-homologous end joining (NHEJ) has been identified [123]. NHEJ seems to comprise at least two different components: (i) a pathway that consists mostly in the direct re-joining of DSB ends, and which depends on the XRCC4, Lig4 and Ku proteins, and; (ii) an alternative NHEJ pathway, which does not depend on XRCC4, Lig4 and Ku, and is especially error-prone, resulting mostly in deletions, with the junctions occurring between micro-homologies [124-129].

For DSBs induced by biological reagents, *e.g.* meganucleases and ZFNs, which cleave DNA by hydrolysis of two phosphodiester bonds, the DNA can be rejoined in a seamless manner by simple re-ligation of the cohesive ends. Alternatively, deleterious insertions or deletions (indels) of various sizes can occur at the breaks, eventually resulting in gene inactivation [16, 106, 111, 130-135]. The nature of this process, which does not rely on site-specific or homologous recombination, gives rise to a third targeted approach based on endonuclease-induced mutagenesis. This approach, as well as the related applications, may be simpler than those based on homologous recombination in that (a) one does not need to introduce a repair matrix, and; (b) efficacy will be less cell-type dependant (in contrast to HR, NHEJ is probably active throughout the cell cycle [35]).

Targeted mutagenesis based on NHEJ has been used to trigger inactivation of single or even multiple genes in immortalized cell lines [136, 137]. In addition, this method opens new perspectives for organisms in which the classical HR-based gene knock-out methods have proven inefficient, or at least difficult to establish, such as rat [138-140], fish [131] and plants [106, 134, 141, 142]. Interestingly, the first clinical trial launched with ZFNs relies on ZFN-mediated mutagenesis of the human CCR5 gene [16] in order to block HIV1 entry into T-cells in AIDS patients (see below, “Targeted mutagenesis of human genes”).

THE MEGANUCLEASE FAMILY OF ENDONUCLEASES: PROPERTIES AND USE

As outlined previously, endonucleases for genome engineering can be classified into four groups: chemical nucleases, ZFNs, meganucleases [90], and now, TALE-nucleases [118, 119]. This section focuses on meganucleases, which

have been used for more than 15 years to induce gene targeting. Recent advances in re-engineering meganuclease specificity have further enhanced their scope of application. The scores of publications and contributions to scientific meetings related to meganucleases reveal a significant evolution in adoption and applicability, with a growing community finding an interest in the exceptional properties of these proteins.

The Meganuclease Family

Meganucleases, also called homing endonucleases, can be divided into five families based on sequence and structure motifs: LAGLIDADG, GIY-YIG, HNH, His-Cys box and PD-(D/E)XK [143, 144]. The most well studied family is that of the LAGLIDADG proteins, which have been found in all kingdoms of life, generally encoded within introns or inteins although freestanding members also exist. To date, a purposeful role within the host has not been identified for these proteins and they tend to be classified as “selfish genetic elements”. With few exceptions, LAGLIDADG proteins exhibit one of two primary activities: (a) they function as RNA maturases involved in facilitating the splicing of their own intron, or; (b) they function as highly specific endonucleases capable of recognizing and cleaving the exon-exon junction sequence wherein their intron resides, thus giving rise to the moniker “homing endonuclease”. It has been hypothesized that homing endonucleases have a so-called “life-cycle” [145] as illustrated in Fig. (1) (i) first, they start out as invasive endonucleases capable of mobilizing their coding sequence; (ii) upon “invasion”, they acquire a concomitant RNA maturase activity to help ensure proper splicing of their intron; (iii) over time, the “invasive” nuclease activity is lost, leaving only the RNA maturase function, and (iv) finally, upon losing the maturase activity, propagation of the intron becomes unviable and the intron is lost. It is thus inferred that the functionality of a given LAGLIDADG protein (endonuclease, maturase, or both) represents a snapshot into the current state of its lifecycle.

Structure of LAGLIDADG Endonucleases

A considerable body of both biochemical and genetic work had established that LAGLIDADG homing endonucleases could be used as molecular tools. It had long been known that the defining sequence motif, LAGLIDADG, represented an essential element for enzymatic activity. Some proteins contained only one such motif, while others contained two; in both cases the motifs were followed by ~75-200 amino acid residues having little to no sequence similarity with other family members. In 1997, two groups shed light on the organization of LAGLIDADG proteins with the structure of both a single-motif protein (I-CreI) [146] and an intein-encoded double-motif protein (PI-SceI) [147]. In both cases the endonuclease domain adopted a similar $\alpha\beta\alpha\beta\alpha$ fold, with the LAGLIDADG motif comprising the terminal region of the first helix and not only contributing to a bipartite catalytic center but also forming the core subunit/subunit interaction. Two such α/β domains assemble to form the functional protein, with the β -strands in each creating a saddle-shaped DNA binding region. That two such subunits were needed for function could be definitively established, as exemplified by I-CreI, a homodimer, having an overall archi-

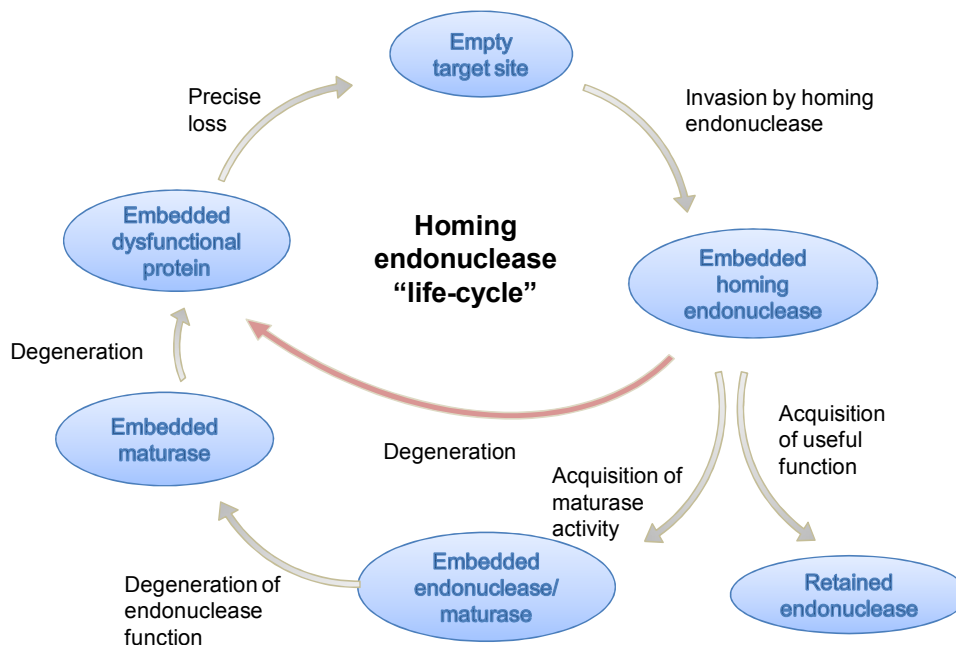


Fig. (1). Proposed life-cycle of a homing endonuclease (see text for details). Adapted from [145].

ture similar to that of the double-motif endonuclease domain of PI-SceI. The α/β fold is a hallmark of LAGLIDADG proteins and has since allowed for better homology modeling and structure/function prediction for proteins lacking structural data.

It has been estimated that LAGLIDADG proteins specifically recognize a DNA sequence ranging from 14 to 40 base pairs in length. Although the initial structures were solved in the absence of target DNA, several labs were able to successfully exploit the structural data to better understand protein-DNA interactions. First, it was clear from the PI-SceI structure that the additional splicing domain characteristic of intein-encoded proteins was flanked by a region capable of specifically contacting DNA [148, 149], a concept borne out when the DNA co-crystal structure was finally solved [150]. These studies addressed the co-evolution of the endonuclease domain (domain II) with the surrounding splicing domain (domain I). Second, having structural data allowed for better predictions regarding studies aimed at investigating the determinants of specificity. Whereas it had previously been known that most homing endonucleases could tolerate single-base changes throughout their recognition sequence, site-sequence degeneracy studies could be rationalized (I-CreI [151], PI-SceI: [152]). Nevertheless, it was only a short time later that DNA co-crystal structures were available, giving not only the first glimpse of LAGLIDADG protein-DNA contacts (I-CreI [153]) but also enabling theories about the nature of catalysis [154, 155]. Despite each subsequent co-crystal structure adding individual idiosyncrasies to the overall picture of LAGLIDADG protein-DNA interactions [150, 155-162], several generalities persist: (i) specificity contacts arise from the burial of the extended β -strands into the major groove of the DNA, with the DNA binding saddle having a pitch and contour mimicking the helical twist of the DNA; (ii) although extensive, the complete complement of hydrogen bonding potential between the protein and DNA is never

fully realized, with many contacts being water-mediated; (iii) cleavage to generate the characteristic 4-nt 3'-OH overhangs occurs across the minor groove, wherein the scissile phosphate bonds are brought closer to the protein catalytic core by a distortion of the DNA in the central "4-base" region; (iv) cleavage occurs via a proposed two-metal mechanism, sometimes involving a unique "metal sharing" paradigm; (v) and finally, additional affinity and/or specificity contacts can arise from "adapted" scaffolds, in regions outside the core α/β fold.

On first inspection it appears that these generalities could define a fixed rule-set, which could help manipulating LAGLIDADG protein-DNA interactions. However, several studies have shown that related proteins can use different subsets of residues to recognize similar DNA [158, 163]. As discussed below, this "relaxed" property of meganucleases has proved to be a double-edge sword, making it at once both easy and problematical to re-engineer specificity.

Engineering Meganucleases

The meganuclease field has recently exploded with novel and interesting engineering applications. In a "bulk approach", several teams tested the concept of subunit interchangeability by fusing both similar and disparate α/β domains, giving rise to hybrid proteins having hybrid specificity derived from each half of the parental protein targets [157, 164, 165]. These proof-of-concept studies addressed not only the feasibility of a "mix and match" approach but also the evolutionary concept that double-motif LAGLIDADG proteins arose from a gene duplication event. Gene fusion alleviates the restriction imposed by homodimers to symmetric or pseudo-symmetric targets by allowing for asymmetric mutations in each subunit and thus recognition of asymmetric DNA targets. This model has undeniably proved valuable in paving the way for selection and screen-

ing studies based on the independence of DNA recognition by individual α/β subunits.

Several strategies for engineering meganucleases involve a semi-rational approach in which specific residues are mutated on the basis of prior structural or functional knowledge to create libraries with limited diversity. Using I-CreI as a model scaffold, several groups have been able to alter its DNA recognition properties both on a small and large scale in terms of total proteins re-engineered. Initial successes with the I-CreI [166-168] and I-SceI scaffolds [169, 170] resulted in the identification of only a few mutants due to the limited screening technologies. To analyze a much larger number of mutant/target combinations, an automated high-throughput screening method was used for the I-CreI scaffold, resulting in the identification of hundreds of mutants with locally altered specificities [171, 172]. These variants were shown to maintain the essential properties of the initial scaffold, i.e. proper folding and stability, cleavage efficiency and a narrow specificity. This has since led to a combinatorial approach in which different sets of I-CreI mutants are combined in the same protein, resulting in entirely redesigned meganucleases cleaving chosen sequences [172, 173].

Currently, engineered meganucleases cleaving two different human genes, XPC and RAG1 [172, 173], as well as a meganuclease cleaving a maize genomic sequence [142], have been described in the literature. Additionally, continued engineering efforts have allowed the production of custom meganucleases cleaving several dozens of chosen natural sequences (our unpublished data). While on the surface these studies solidify the practicability of altering meganucleases, bona fide engineering of specificity has proved to be more complex. Progress remains to be made if one wants to use the full potential of other LAGLIDADG proteins, including

natural single-chain variants. Detailed studies of the LAGLIDADG scaffold have demonstrated noteworthy pitfalls, including differences in how each subunit interacts with DNA, either in a homodimer context to efficiently bind asymmetric DNA [162] or within the same monomer, displaying a disparity in affinity vs. specificity [174]. Moreover, regions outside the core fold have been implicated in being important for target site DNA interactions [175]. Finally, the homodimer nature of the proteins themselves creates an inherent impasse when designing proteins for asymmetric targets, as three species of functional enzyme will arise in the mix.

Fortunately, the future of meganuclease engineering is quite promising as all of these challenges are being overcome and in the process innovative studies have shed light on new and exciting avenues. Comprehensive computational studies have given rise to not only specificity re-engineering [176], but have also addressed the dimerization issue via targeting protein-protein interactions of the subunits [177] as well as through mimicking nature with single-chain designs [178, 179]. Exhaustive analysis of the I-AniI scaffold [180, 181] has led to better computational methods that can exploit binding energy for designs [180]. Through both computational [182] and structural [161] analysis of novel structures of engineered proteins in complex with DNA, insight has been gained into the intricacies of the protein-DNA contacts that appear to involve substantial conformational changes in both the protein and/or DNA; findings that suggest specificity engineering is more than a point-by-point process Fig. (2). Moreover, new opportunities are being explored through the generation of targeted “mega-nickases” (I-SceI [183], I-AniI [184]) that can in principle provide similar levels of induced HR with a minimization in the frequency of NHEJ.

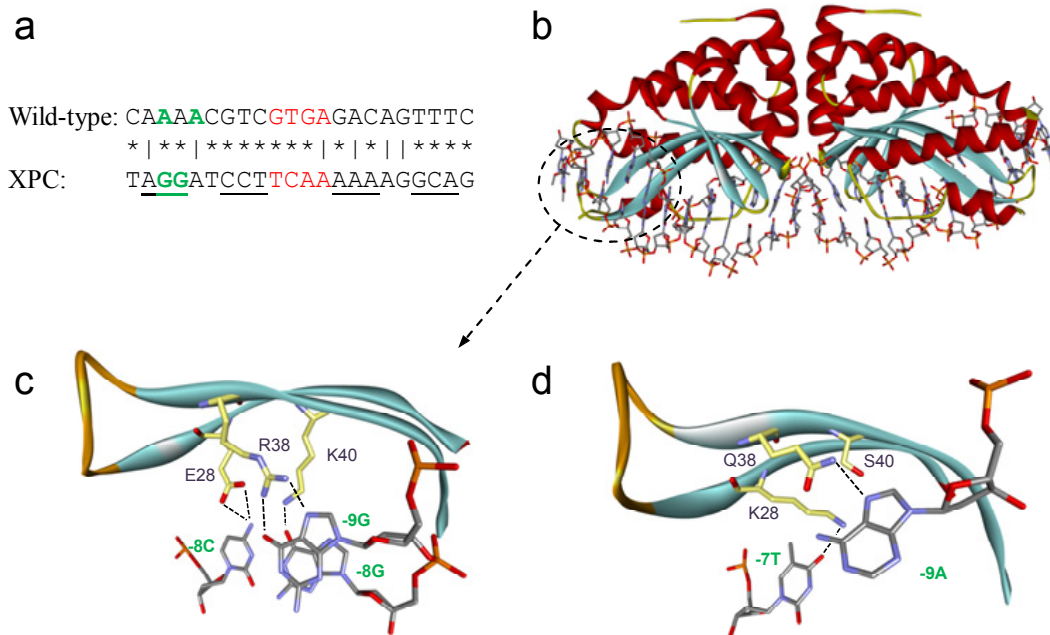


Fig. (2). Re-engineering meganuclease specificity: example of a custom meganuclease targeting a sequence from the human XPC gene. (a) Comparison of the starting and final targets. (b) Crystal structure of the engineered I-CreI variant in complex with target DNA. Details of protein-DNA contacts from the encircled region are highlighted (c) to demonstrate how a clustered approach leads to significant changes in specificity contacts. The native protein-DNA contacts for I-CreI (d) are shown for reference. Adapted from [161].

Use of Meganucleases for Genome Engineering

DNA damage is a naturally occurring event that can lead to chromosomal aberrations or cell death. DSBs are particularly hazardous to the cell as they can potentially lead to genome rearrangements. Not surprisingly, a variety of repair strategies have evolved to ensure genomic integrity. As outlined above, these processes can be either error-prone (NHEJ) or conservative (HR), depending on the mechanisms involved. Understanding the nature of each mechanism can be crucial in the context of gene targeting. Namely, from the genome engineering point of view, the challenge of using the cellular DNA repair machinery lies in creating a precise DSB in a genome.

I-SceI is the prototypical meganuclease used for genome engineering. In nature, the protein stimulates HR by creating a site-specific DSB in the genome, in a process called homing (see above). The discovery of this function in the yeast *Sacharomyces cerevisiae* launched a new era in gene targeting [185]. Researchers studying DNA repair in mammalian cells soon realized the potential of such a tool. In the nineties, pioneering work used a chromosomal neomycin resistant gene interrupted by an I-SceI recognition site as a reporter to monitor gene correction events in mouse cell lines upon introduction of an I-SceI expression vector and a DNA repair matrix. Gene correction could be achieved at frequencies of 3×10^{-5} in ES [186] and 4×10^{-4} in NIH3T3 [187] cells. Further experiments in NIH3T3 and PCVC7 cell lines were able to show that targeted gene insertion could also be obtained using this approach [188, 189]. In contrast to cell populations that received the DNA repair matrix alone, where no targeted events could be detected, frequencies of 1.8×10^{-4} to 4×10^{-4} targeted events per transfected cell could be achieved. Later studies were able to realize an even higher frequency of gene targeting. For example, Donoho and colleagues obtained nearly 1% recombination in ES cells [190], while Szczepiek and coworkers could attain up to 10% with I-SceI (used as positive control) in 293 cells [191]. These analyses paved the way for the DSB-induced gene targeting technology, as they demonstrated a >1000-fold stimulation of events attributable to the activity of the meganuclease. Since then, this simple system has been widely adopted by the scientific community and has allowed for a better understanding of the DNA repair mechanism in a variety of cell types and experimental conditions [89].

Despite the notion that HR precisely repairs the template, in certain cases imperfect targeting events were reported, the consequences of which would be notable for therapeutic applications [187] [186]. These results highlighted the necessity to understand the intricacies of HR. For example, it was found that HR can occur at one end of the DNA junction while non-homologous recombination is carried out at the other. In a similar vein, the possibility to correct (or introduce) a point mutation at a distance from the DSB was studied. It was determined that the double-strand break repair (DSBR) pathway can result in the conversion of sequences adjacent to the DNA break, wherein the efficiency of the conversion decreases as the distance from the DSB increases. Elliott and colleagues observed a very sharp decrease of the conversion efficiency: a polymorphism located at a distance of circa 100 bp from the DSB was corrected in only 13-16%

of the repair events, and by 400 bp correction dropped to only 3%. [192]. Meanwhile, Donoho and coworkers reported longer conversion tracts, with up to 13% correction at a distance of about 4000 bp from the DSB [190]. Finally, experiments in the human cell line 293H have shown that a mutations located 265 bp from the DSB could be efficiently (66%) corrected [193]. Despite the apparent discrepancies in these results, most likely arising in part from experimental design differences, the trend in conversion efficiency observed as a function of distance remains an important point to consider for any gene targeting approach.

Ultimately, the studies made with the I-SceI system to decipher the cellular mechanisms of DSBR have laid the foundation for all the main strategies envisioned today for genome engineering purposes (Fig. (3)). Gene correction [194, 195], gene insertion [189, 190, 196], and gene inactivation through a NHEJ mechanism [130] as well as the elimination of a DNA sequence by single-strand annealing (SSA) assays between direct repeats [130, 190, 197] were all addressed during these early studies. This last approach appears highly efficient as I-SceI is able to induce recombination between direct or inverted repeats at frequencies of up to 10^{-1} events among transfected mammalian cells. A similar frequency was obtained *in vivo* after tail-vein injection of an adenovirus expressing I-SceI into transgenic mice carrying a LacZ-based reporter gene: up to 1.3% of the hepatocytes showed recombination events [198]. Although cell-line development for protein production does not represent the focus of this review, it is worth mentioning that genes inserted by meganuclease-induced recombination show reproducible expression levels among targeted cellular clones [199, 200].

The I-SceI-based HR system has also been used to address fundamental questions that in many ways also gauge the limits of the methodology. For instance, the possibility of utilizing the homologous chromosome or an ectopic chromosomal locus as a repair template for I-SceI induced DSBR was investigated using cultured cells [201, 202]. In terms of gene therapy, this strategy would provide a tremendous advantage in the correction of mutations responsible for dominant monogenetic disease as the introduction into cells of the meganuclease could alone lead to gene correction. Unfortunately, when sequences are located on separate homologous or heterologous chromosomes, the recombination efficiency is extremely low (10^{-6} - 10^{-5} events among transfected cells), making this approach impractical for therapeutic applications. Experiments using the template-based method have nevertheless shown that, in addition to mammalian cells, I-SceI-mediated recombination is robust and efficient in diverse organisms. It has been successfully used to induce various modifications such as mutagenesis, recombination between repeats or gene targeting in bacteria [203-207], mosquito [208, 209], fly [210] and plant [134, 211-216]. I-SceI has also been used to improve transgenesis efficiencies in various organisms such as frog [217], fly [218], fish [219], and sea anemone [220].

An obvious drawback to the use of natural meganucleases lies in the need to first introduce a known cleavage site into the region of interest. Although so-called pseudo-sites within a genome can be exploited by certain methods (e.g. the Φ C31 integrase), such approaches are not always possi-

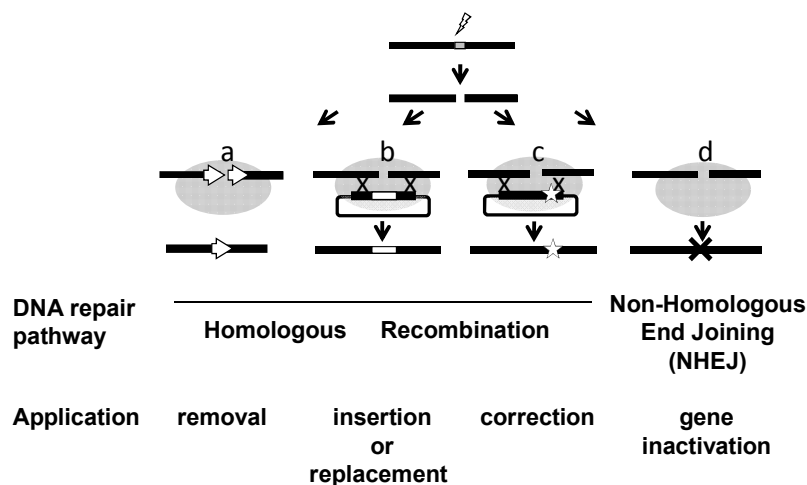


Fig. (3). Endonuclease-induced gene targeting approaches. Upon cleavage, DNA repair mechanisms may result in one of several outcomes. When a double-strand break is targeted between two direct repeats (a), homologous recombination can result in the deletion of one repeat together with the intervening sequence. Gene insertion (b) or correction (c) can be achieved by the introduction of a DNA repair matrix containing sequences homologous to the endogenous sequence surrounding the DNA break. Mutations can be corrected either at or distal to the break, with the frequency of correction decreasing with increasing distance. The misrepair of DNA ends by error-prone non-homologous end joining (d) can result in insertions or deletions of various sizes, leading to gene inactivation.

ble in therapeutic cases where targeting can demand a more stringent level of precision. To this end, engineered meganucleases are poised to address the targeting issue and should allow for a broader use of HR-based technology. Extensive characterization of an engineered I-CreI derivative cleaving the human RAG1 gene [178] has indeed demonstrated that a fully redesigned meganuclease can compare to I-SceI in terms of both efficacy and specificity. This RAG1 endonuclease could be used to induce 6% recombination in 293 cells and its relative specificity was comparable to that of I-SceI [178]. Similar results (1 to 20% targeted recombination) have been obtained in transfection experiments using other immortalized cell lines, not only with the RAG1 meganuclease but with many other I-CreI-derived engineered meganucleases tailored to uniquely cleave diverse targets in the human genome (our unpublished data).

DIFFERENT STRATEGIES FOR GENE THERAPY BASED ON ENDONUCLEASES

In this section, the major strategies envisioned for the therapeutic use of rare-cutting endonucleases are outlined.

Gene Correction Approaches

Several strategies exist for choosing the nature of the meganuclease-induced modification as well as the design of the targeting construct. Meganucleases can in principle be used to accomplish what may be considered as “ideal” gene therapy: true reversion of the deleterious mutation. This approach has practical limitations, since the efficacy of correction decreases rapidly as the distance from the initial DNA double-strand break increases. As a consequence, this approach can only be considered in diseases caused by a prevalent (and, in theory, targetable) mutation, such as sickle cell anemia [221]. An alternative approach along the same lines could involve inserting the whole coding sequence upstream of a deleterious mutation. In this case, although regions of

the damaged gene would remain, they would be effectively “silenced” by the newly inserted corrected gene.

Gene repair strategies can more easily be envisioned wherein the particular disease itself presents a selective advantage upon treatment by gene therapy, such as SCID [222] or Wiskott-Aldrich syndrome (WAS) [223]. Endonucleases targeting two different genes involved in SCID diseases have been described to date: (i) the IL2RG gene that has been targeted by a ZFN [104, 105], and; (ii) the RAG1 gene, for which a meganuclease has been re-engineered to target [172, 178]. *In vivo*, there is a selective advantage for cells with a functional gene because signaling through the γ c cytokine receptor or possessing a functional RAG1 recombinase confers a survival and/or proliferation signal during differentiation of lymphocytes [222, 224]. The IL2RG ZFN has been used in a variety of immortalized and primary cells. An initial design was used to achieve gene targeting in 18% of treated K562 cells, and 5% of treated T cells [104]. Using the latest ZFN variants engineered for minimal toxicity [105], 5-15% gene insertion has been observed in transfection experiments in K562 cells [109]. The use of non-integrative lentiviral vectors resulted in up to 6% of targeted insertion [103]. Likewise, with the meganuclease targeting the RAG1 gene, aiming at the same correction approach, up to 6% of targeted modifications could be observed in transfection experiments using 293 cells [178]. For comparison, the RAG1 meganuclease and the IL2RG ZFN appear to yield the same frequencies of homologous gene targeting in 293 cells (our unpublished data). In principle, these frequencies should be enough for an *ex vivo* treatment of SCID diseases, provided one can achieve them in hematopoietic stem cells (HSCs). However, achieving high efficiencies in HSCs might be problematic (see below, THE EFFICACY AND SAFETY ISSUES), and future studies will tell whether the quality of the nucleases reported above will be high enough to compound the intrinsic difficulties linked with the handling of human HSCs.

Other nucleases have been engineered to target genes whose correction by homologous gene targeting could be of therapeutic relevance. For example, a recent report describes a ZFN targeting the human PIG-A gene, involved in paroxysmal nocturnal hemoglobinuria (PNH) [108]. This ZFN allowed for gene targeting of 3% of transfected 293 cells, but these figures decreased by at least 100 fold in stem cells ([108]. Two meganucleases targeting the human XPC gene have also been described, but these proteins were tested only with reporter systems [161, 173], with targeting frequencies between 0.1% and 1% (our unpublished data). With these frequencies, the use of such nucleases will depend on selection procedures to sort targeted events. For therapeutic purposes, the selection scheme should be both (a) as non-invasive as possible, and; (b) at the expense of the multipotency or progenitor capacity of the treated cells.

Insertion of Therapeutic Genes into a Safe Harbor

An increasingly popular approach in gene therapy today seems to be the targeted insertion into a so-called “safe harbor”. This strategy consists of introducing genetic material into a predefined locus unrelated to the disease-causing gene. In contrast with true gene correction, targeted insertion is not limited to the use of methods based on homologous gene targeting, and targeted recombinases or transposases could perform as well.

For this method to be valid, the locus of interest must possess two essential properties. First, an adequate safe-harbor locus needs to support sufficient and stable gene expression in the modified cells. For the gene therapy to be effective, the inserted element of interest must be available in the host at functional, if not “wild-type”, levels. Second, and most importantly, the safe-harbor locus needs to be “safe” in terms of not interfering with normal cell function and gene regulation. Adverse events could be caused by the deregulation of other genes, either by repression or induction of gene expression, or by inducing chromosomal instability. The difficulty in the safe-harbor approach stems from the ability to predict that a particular locus is safe, and many different strategies are currently being explored. For example, a ideal site for transgene expression was identified in mouse genetics studies of the ROSA26 locus, where insertion of a coding sequence leads to its constitutive expression without altering viability [225]. However, there is no evidence that the homologous human locus is similarly “safe”. A promising safe harbor is the AAVS1 locus, the most common integration site for AAV, a parvovirus lacking a known associated pathology [226]. AAVS1 is located on chromosome 19 within intron 1 of the PPP1R12C gene, and has been shown to allow inserted exogenous DNA to be transcribed [226, 227]. ZFNs targeting AAVS1 have recently been developed and used to induce targeted insertions in immortalized cell lines and primary cells, with stable expression of the transgene [107, 228]. In transfection experiments, frequencies of targeted events could reach circa 10% in K562 cells, 3% in HEK293 and Hep3B cells, and similar values in other cell lines. Gene targeting could also be observed in human stem cells, e.g. ES and iPS cells. However, targeted events were identified after selection for transformants having integrated a reporter gene placed in the repair matrix. Whereas targeted events represented a large fraction

of the integrations (up to 61%), it is difficult to infer the absolute rate of gene targeting (e.g. the ratio of targeted events per transfected cell) in these studies [107, 228]. Nevertheless, one can suspect that it is low, which again raises efficacy issue for real-world therapeutic applications. If the use of a selection scheme has to be envisioned, it is with caution, for it would have to be non invasive and not affect the multipotency of the targeted stem cells. Future studies should tell whether the current targeting protocols can avoid selection procedures that might prove problematic.

While the above examples illustrate active research on potential safe-harbors, the choices of loci in these cases were fortuitous. However, given the constant progress in genome annotation, the mapping and description of multiple features (e.g. GC-clusters, non-coding RNA genes, sequence variations, chromatin-immunoprecipitation, mRNA expression pattern, etc) are available from publicly accessible databases, allowing for semi-rational selection of safe harbor loci. Perhaps the simplest way to keep cellular expression profiles unaltered is to target loci distal from functional genes, with an emphasis on avoiding tumor suppressor genes. While basic in concept, this strategy does not fully address the issue of effects on chromatin remodeling or chromosomal stability. The cell type that needs to be targeted must also be taken into account when investigating the usefulness of a particular locus. For example, a locus that supports stable gene expression in hematopoietic stem cells might not provide the same expression in neurons or hepatocytes. This discrepancy may arise from many factors, one of which being the variability in whether a safe locus is indeed “safe” in different cell types. In all cases, it must be noted that the choice of a viable safe harbor is coupled to its ability to be efficiently targeted, either by preexisting or engineered nucleases. Thus, beyond the issues of safety and stable gene expression, safe harbor accessibility must be considered in light of nuclease targeting potential, which could vary along the genome and between cell types, due to epigenetic modifications (there is at least one example in plants of a nuclease targeting two exactly identical sequences with different efficacies, these two different sequences being found in two different loci, in the SurA and SurB genes [106]). Despite these technical obstacles, the identification of a good safe-harbor/nuclease pair appears to be a readily attainable goal. Notably, an additional advantage of using a safe harbor strategy is that it can also be utilized in multiple indications as different genes can be expressed from the same locus.

Targeted Mutagenesis of Human Genes

The use of meganucleases to target a mutagenic event at a specific locus is actively being investigated for numerous purposes. As evidenced by its widespread use in model systems, targeted mutagenesis can also have therapeutic perspectives. The selective disruption to restore the wild-type phenotype of mutated alleles in dominant diseases remains a challenge: an endonuclease able to discriminate the mutated allele from that of the wild-type is required. Whereas the treatment of monoallelic diseases by targeted mutagenesis represents an interesting concept, more practically this approach can be applied to the inactivation of unwanted genetic elements.

ZFNs have recently been used to inactivate the CCR5 gene, a co-receptor for HIV entry, leading to resistance against HIV infection [16]. Inactivation of 50% of CCR5 alleles could be observed in primary human CD4⁺ cells, providing selective advantage to CCR5^{-/-} double-mutant cells upon HIV1 infection *in vitro* and in an *in vivo* mouse model. Genotoxicity and specificity were assessed by the absence of DNA-repair protein foci formation, and by testing the frequency of induced mutagenesis at other potential cleavage sites. This work led to the opening in 2009 of a new clinical trial for AIDS patients, the first one involving a rare-cutting endonuclease. Recently, ZFN could also be used to inactivate CCR5 in human CD34⁺ cells [229]. These results provide a first step toward a therapeutic approach with the potential to create HIV resistant cells in all the blood cell lineages, instead of only within the T cell population.

Virus Clipping

Nucleases targeted against viral sequences represent a novel class of antiviral agents that could cleave and either partially excise or completely eliminate viral DNA from infected cells, rendering them essentially “virus free”. This strategy mimics the cleavage of foreign DNA by restriction endonucleases in bacteria. The advantage of disrupting the viral genome, and not the various steps of viral replication, lies in the possibility of targeting latent forms of the virus. Latent viruses persist in the cell and are normally not affected by conventional treatments, which are effective only when the virus is actively replicating. This approach can be envisioned for different types of viruses, given that a double-strand DNA (dsDNA) stage is part of its lifecycle. Retroviruses, for example, constitute a large group of RNA viruses that upon entry into the target cell will generate by reverse transcription a dsDNA molecule that integrates into the host-cell genome. Other viruses that can be targeted by engineered meganucleases include the Polyomavirus, Papillomaviruses, Herpesvirus and Hepadnavirus, all of which belong to the group of DNA viruses and for which the genomic DNA remains as an episomal molecule in infected cells. Potential mechanisms are outlined in Fig. (4).

The feasibility of virus clipping is in the process of being validated by recent examples in the literature. For instance, it has been shown that ZFNs targeting Hepatitis B virus (HBV) DNA can efficiently cleave an HBV sequence present within a plasmid transfected in cultured cells [230]. Analysis of the DNA target reveals misrepair of affected HBV sequences within cultured cells coupled with a drop in pre-genomic viral RNA levels, an indicator of replication competence in cellular models that mimic HBV infection. Another promising approach lies in the prevention of viral infection using targeted meganucleases. Using a recombinant Herpes simplex virus (HSV-1) harboring an I-SceI restriction site, it was demonstrated that viral replication is impeded in COS-7 cells that have been transfected with a plasmid coding for the I-SceI meganuclease prior to infection (Smith *et al.* 2010, submitted). Furthermore, redesigned meganucleases derived from I-CreI were generated that are able to recognize and cleave sequences from the HSV genome. These HSV-specific meganucleases, when transfected prior to infection with a wild-type strain of HSV-1, reduce the levels of viral genomic DNA more than 50% compared to cells transfected

with an empty vector. Still, while the above examples of virus clipping appear successful, efficacy remains to be assessed for integrated viral forms. Early experiments suggest that cleavage in both LTRs could result in viral genome loss, by tandem repeat recombination or rejoining of the two DSBs [130, 188, 197]. How these issues will be addressed remains to be seen.

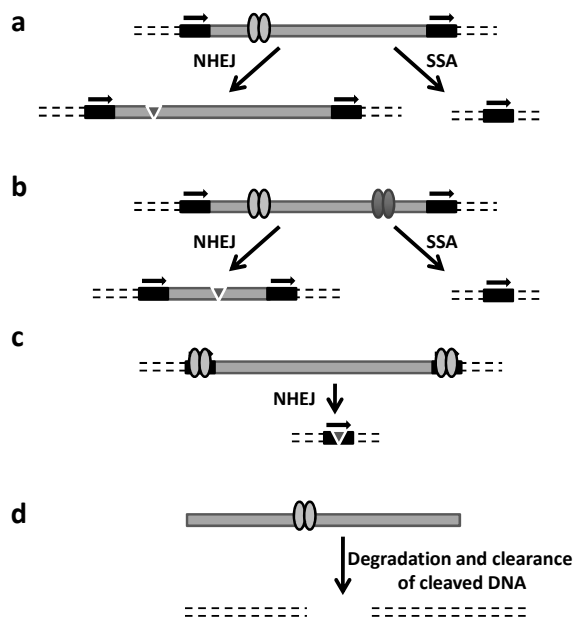


Fig. (4). Meganucleases as antiviral agents. Pathways by which viral sequences belonging either to essential genes or regulatory regions can be inactivated are shown. Gene inactivation can result from small insertions/deletions that introduce lethal mutations in the viral genome by error-prone non-homologous recombination (panel a). Large deletions can be introduced by DNA cleavage and repair when using two different meganucleases targeting the same viral genome at different positions (panel b), or by rejoining DNA ends when cleavage occurs in a repeated region of the viral genome (panel c). Alternatively, when cleavage occurs between two direct repeats (e.g.: the LTR retroviral sequences), deletion of the intervening sequences can be generated by tandem repeat recombination (SSA, panels a and b). While the pathways depicted in panels a-c are valid for integrated as well as episomal viruses, the latter can also be targeted via the degradation and clearance of the viral genome upon DNA cleavage (panel d).

THE EFFICACY AND SAFETY ISSUES

In 2005, Urnov and colleagues created an engineered ZFN targeting the human IL2RG gene. By early 2010, a handful of diverse engineered nucleases able to edit human genes with high efficiency were described in the literature [16, 103, 105-108, 111, 178, 228]. One of these, a ZFN that can induce gene knock-out by NHEJ, is even in clinical trial [16]. These data reflect both the potential and limits of the technology as applied today.

Frequencies of Targeted Genome Modifications

In reviewing the different alternatives to viral gene transfer for gene therapy, it becomes clear that multiple factors

can influence the success of each given approach. All targeted methods have aimed at reaching a symbolic "1% ratio" of targeted integration. Nevertheless, subtle variations such as reagents used, cell-types targeted, vector construction and even data analysis methods can have a significant impact on reported frequencies. Nucleases-based methods as discussed herein can achieve frequencies of >5% in immortalized cells, clearly indicative of the potential these approaches hold. However, a few recent studies suggest that these figures can strongly decrease in stem cells, which might represent one of the major hurdles for such technologies.

Lombardo and co-workers used a ZFN cleaving the human CCR5 gene to induce targeted insertion in different cell types [103]. Using non-integrative lentiviral vectors instead of transfection, up to 50% of targeted events were observed in K562 and Jurkat cells, about 5% in ES cells, and close to 0.1% in CD34+ cord blood progenitor cells [103]. However, Zou and colleagues observed lower figures in human ES cells [108]. Using ZFNs targeting a GFP reporter gene, they obtained frequencies in the range of 3% in 293T cells, but only of 0.1-0.2% in hES and iPS. With a tailored ZFN targeting the PIG-A human gene, the same authors observed even lower frequencies, in the range of $2-4 \times 10^{-4}$ for human ES cells, and 10^{-5} for human iPS. In another study, Hockemeyer *et al.* used a selection process to identify targeted recombination events induced by tailored ZFNs in the human OCT4, AAVS1 and PITX3 endogenous genes in human ES and iPS cells. High frequencies of targeted integrations were observed among transformants (up to 61%), but the absolute rate of gene targeting is not indicated [107].

Discrepancies between the studies of Lombardo and colleagues and those of Zou and coworkers in the observed targeted recombination frequencies in hESCs could be due to the methods used, the quality of the nuclease employed, or even the cell type. In any case, efficacies in stem cells seem to be far below what can be achieved with viral vectors, or even non-viral methods such as those utilizing enhanced versions of the Sleeping Beauty transposon [70].

One hypothesis to explain the low frequencies of HGT events observed by Lombardo *et al.* in CD34+ cells, and by Zou *et al.* in human ES and iPS cells, is that these cells would be less proficient for HR. In principle, endonuclease-mediated mutagenesis should be less limiting than HR-based approaches, for NHEJ is not limited to the S/G2 phases of the cell cycle. Nevertheless, lower efficacies can result from factors other than a limiting repair pathway. The induction of indels by endonucleases has been reported for several cell types and organisms, often with impressive efficacies: the CCR5 gene could be mutated with frequencies in the range of 30-54% in primary T-cells using an adenoviral vector expressing a ZFN [16]. Zou *et al.* also monitored the frequency of targeted mutagenesis induced by their PIG-A ZFN, and observed only 4.6×10^{-6} of indels in ES cells. This frequency, about 100 times lower than the frequency of DSB-induced HGT in the same cells, has to be considered cautiously, for it was monitored by a selection process. However, it suggests that the limiting factor in hES cells could actually be suboptimal vectorization and/or lower expression levels of the endonuclease, which would affect both HGT and targeted mutagenesis.

These data emphasize the need for a more complete body of data regarding targeted recombination in stem cells, and likely reflect the inherent difficulties in handling these cells, regardless of the mechanism of action (targeted mutagenesis or HGT) of the nuclease.

Specificity

Whereas the efficacy of targeted approaches can be measured in terms of directed events (e.g. percent mutagenesis at the targeted locus), a true benchmark for specificity remains elusive as the very nature of monitored events can vary by assay. For instance, the detection of repair protein foci (phosphorylated H2AX or 53BP1) is routinely used to measure nuclease toxicity [16, 105, 178, 191, 231, 232]. The high levels of repair foci that occur naturally in immortalized cells, however, make this assay relatively insensitive despite the assayed proteins being highly specific. A more straightforward assay using immortalized cells involves monitoring cell survival in the presence of increasing doses of nuclease to compare the relative toxicity of different endonucleases [106, 178, 232, 233]. This assay in turn suffers from being hardly predictive of what will happen in true target cells, wherein expression will likely be more limited. Endonuclease-induced indels have also been a major concern and several studies have described the characterization of mutagenesis in promiscuous sequences, typically found to be null or infrequent [16, 107]. However, targeting the CCR5 gene with a dedicated ZFN induced significant mutagenesis in the related CCR2 gene (4-5% for CCR2 versus 36% for CCR5) [16]. Furthermore, studies showing that the occurrence of several DSBs can result in translocation have raised additional concerns [234, 235]. The toolbox for characterizing potential genotoxicity certainly needs to be upgraded to address these issues, and several efforts are currently underway.

A number of standard quality control tests can be used to test the status of stem cells, e.g. monitoring (i) growth rate; (ii) karyotype; (iii) cell morphology; (iv) the expression of genes marking the differentiation status, and; (v) multipotency (differentiation tests *in vitro* or *in vivo*). Such tests have been used *after* stem cell engineering with nucleases [107, 108], but also with other reagents [70], without detection of any gross alteration. Finally, using a humanized mouse model to study the fate of engineered cells [16] appears effective for assessing both activity and safety.

CONCLUSION

The difficulties in harnessing the potential of endonuclease-based targeted approaches can best be placed in perspective when one considers that the first clinical trial using these methods involves a gene knock-out technique. The purpose of replacing current methods that use random transgene insertion by those using targeted recombination is to have better control over the resulting engineered cells. More specifically, the goal is not only to promote and dictate transgene expression, but also to avoid adverse events such as insertional mutagenesis. In addition, these benefits need to be significant. For example, whereas retroviral gene transfer in 20 patients treated for X-SCID [3, 5] and 2 patients treated for CGD [6] resulted in several cases of leukemia (X-SCID)

or myoplasia (CGD), hundreds of patients have received T-cells engineered with similar vectors [236] without a single occurrence of induced malignant transformation. In these cases the integration sites were different from those commonly found in HSCs, perhaps explaining the absence of SAEs [237]. Thus, safety concerns can vary depending on the indication.

The benefits of targeted methods must also not be at the expense of new or unexpected significant drawbacks. In the case of endonuclease-mediated recombination, the obvious potential drawbacks are (i) a potential genotoxicity that remains to be assessed, and; (ii) lower efficacy. Genotoxicity is an inherent problem of enzymes that act on nucleic acids, though one can expect that highly specific endonucleases would reduce or abolish this issue. The lower efficacy, on the other hand, stems from limitations in the levels of homologous recombination within the cells, an obstacle underscored by experiments in stem cells and especially in HSCs. Several studies aspire to alleviate this tissue-specific variability by using iPS cells, although it not clear if (a) they display better recombination rates or; (b) they will be acceptable for cell therapy in the short term.

Currently, random insertion using a quality vector (preferentially a viral one) is more efficient than the highest reported targeted-insertion rate, regardless of the cell type. For any emerging method it is therefore important to strive for both an excellent recombination rate as well as a low threshold in terms of required activity. Whereas the activity level depends on the cell type (for HR proficiency) and vectorization (to bring enough nuclease and repair matrix into the target cell), the threshold of required activity rather depends on the indication. The real question is whether the efficacies described above represent a limit. Differences in efficacy could be less stringent for indications such as haemophilia, wherein expression of 1% of wild-type levels is sufficient to result in a therapeutic effect [238]. If gene correction is associated with a selective advantage (e.g. SCID diseases), a low frequency of repair should be enough to provide a therapeutic effect [89]. Then again, SCID has to be addressed in early progenitor cells, which could prove especially difficult to target. Whether these limitations also apply to the same extent for gene knock-out based strategies is unclear. Although the CCR5 clinical trial cannot provide a benchmark for strategies based on targeted insertion, the absence of toxicity reported so far is an important clue for nuclease-based approaches. Pioneering studies using a good safe harbor/nuclease coupled with one of the more “accessible” diseases (such as the ones quoted above) could represent an essential milestone, answering many important questions. In this respect, the choice of a good nuclease (high activity/toxicity ratio) targeting a good safe-harbor locus (stable expression and no impact on neighboring genes) will be extremely important. However, the choice of the target cell might be at least as important as the mechanism of action of the nuclease: recent publications have shown that the use of stem cells is associated with potential additional difficulties, and it should be noted that the current CCR5 clinical trial is based on the treatment of T cells, and not of HSCs (the possibility to treat HSCs is being currently investigated [229]). Furthermore, one can expect that a humanized mouse model

will be key in deciding whether currently achievable recombination rates in stem cells will be sufficient.

For endonuclease-based approaches, proper expression of the nuclease within the target cell is certainly one of the most important aspects for both efficacy and specificity. A good efficacy relies on the introduction of a sufficient amount of nuclease into the target cell. Although the off-site cleavage of these molecules has been significantly reduced by a variety of strategies, a low level of activity over an extended period of time could have significant consequences. Several methods of vectorization have been used for the generation of nuclease mediated modification events in human cells, including non viral methods (electroporation) and the use of viral vectors (AAV vectors, integration-defective lentiviral vectors (IDLV) and adenovirus vectors, the latter solution being used in the CCR5 clinical trial). Nevertheless, one of the biggest advantages of nuclease-based approaches could be that they could alleviate the need for viral vectors, which represent a real hurdle in terms of GMP manufacturing. In this regard, nuclease-based strategies are more in competition with transposon-based (targeted or non-targeted) rather than with viral vector-based approaches.

ABBREVIATIONS

AAV	=	Adeno-Associated Virus
DSB	=	Double-Strand Break
ES	=	Embryonic Stem (cells)
HGT	=	Homologous Gene Targeting
HR	=	Homologous Recombination
HSC	=	Haematopoietic Stem Cell
iPS	=	Induced Pluripotent Stem (cells)
NHEJ	=	Non-homologous End Joining
SCID	=	Severe Combined Immune Deficiency
SFHR	=	Short Fragment Homologous Replacement
TALE	=	Transcription Activator-Like Effector
TFO	=	Triplex-Forming Oligonucleotides
ZFP	=	Zinc-Finger Protein
ZFN	=	Zinc-Finger Nuclease

REFERENCES

- [1] Cavazzana-Calvo M, Hacein-Bey S, de Saint Basile G, *et al.* Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science* 2000; 288: 669-72.
- [2] Gaspar HB, Parsley KL, Howe S, *et al.* Gene therapy of X-linked severe combined immunodeficiency by use of a pseudotyped gammaretroviral vector. *Lancet* 2004; 364: 2181-7.
- [3] Hacein-Bey-Abina S, Garrigue A, Wang GP, *et al.* Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *J Clin Invest* 2008; 118: 3132-42.
- [4] Hacein-Bey-Abina S, Von Kalle C, Schmidt M, *et al.* LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* 2003; 302: 415-9.
- [5] Howe SJ, Mansour MR, Schwarzwaelder K, *et al.* Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *J Clin Invest* 2008; 118: 3143-50.
- [6] Stein S, Ott MG, Schultze-Strasser S, *et al.* Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation

- after gene therapy for chronic granulomatous disease. *Nat Med* 2010; 16: 198-204.
- [7] Bank A, Dorazio R, Leboulch P. A phase I/II clinical trial of beta-globin gene therapy for beta-thalassemia. *Ann N Y Acad Sci* 2005; 1054: 308-16.
- [8] Cartier N, Hacein-Bey-Abina S, Bartholomae CC, *et al.* Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. *Science* 2009; 326: 818-23.
- [9] Humeau LM, Binder GK, Lu X, *et al.* Efficient lentiviral vector-mediated control of HIV-1 replication in CD4 lymphocytes from diverse HIV+ infected patients grouped according to CD4 count and viral load. *Mol Ther* 2004; 9: 902-13.
- [10] Modlich U, Navarro S, Zychlinski D, *et al.* Insertional transformation of hematopoietic cells by self-inactivating lentiviral and gammaretroviral vectors. *Mol Ther* 2009; 17: 1919-28.
- [11] Titeux M, Pendaries V, Zanta-Boussif MA, *et al.* SIN retroviral vectors expressing COL7A1 under human promoters for *ex vivo* gene therapy of recessive dystrophic epidermolysis bullosa. *Mol Ther* 2010; 18(8): 1509-18.
- [12] Zhou S, Mody D, Deravin SS, *et al.* A self-inactivating lentiviral vector for SCID-X1 gene therapy that does not activate LMO2 expression in human T cells. *Blood* 2010; 116(6): 900-8.
- [13] Cavazzana-Calvo M, Payen E, Negre O, *et al.* Transfusion independence and HMGA2 activation after gene therapy of human beta-thalassaemia. *Nature* 2010; 467: 318-22.
- [14] Montini E, Cesana D, Schmidt M, *et al.* The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy. *J Clin Invest* 2009; 119: 964-75.
- [15] Wang GP, Levine BL, Binder GK, *et al.* Analysis of lentiviral vector integration in HIV+ study subjects receiving autologous infusions of gene modified CD4+ T cells. *Mol Ther* 2009; 17: 844-50.
- [16] Perez EE, Wang J, Miller JC, *et al.* Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat Biotechnol* 2008; 26: 808-16.
- [17] Hinnen A, Hicks JB, Fink GR. Transformation of yeast. *Proc Natl Acad Sci U S A* 1978; 75: 1929-33.
- [18] Orr-Weaver TL, Szostak JW, Rothstein RJ. Yeast transformation: a model system for the study of recombination. *Proc Natl Acad Sci U S A* 1981; 78: 6354-8.
- [19] Orr-Weaver TL, Szostak JW, Rothstein RJ. Genetic applications of yeast transformation with linear and gapped plasmids. *Methods Enzymol* 1983; 101: 228-45.
- [20] Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW. The double-strand-break repair model for recombination. *Cell* 1983; 33: 25-35.
- [21] Rothstein RJ. One-step gene disruption in yeast. *Methods Enzymol* 1983; 101: 202-11.
- [22] Thomas KR, Folger KR, Capecchi MR. High frequency targeting of genes to specific sites in the mammalian genome. *Cell* 1986; 44: 419-28.
- [23] Thomas KR, Capecchi MR. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* 1987; 51: 503-12.
- [24] Mansour SL, Thomas KR, Capecchi MR. Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. *Nature* 1988; 336: 348-52.
- [25] Smithies O, Gregg RG, Boggs SS, Koralewski MA, Kucherlapati RS. Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. *Nature* 1985; 317: 230-4.
- [26] Doetschman T, Gregg RG, Maeda N, *et al.* Targetted correction of a mutant HPRT gene in mouse embryonic stem cells. *Nature* 1987; 330: 576-8.
- [27] Koller BH, Smithies O. Inactivating the beta 2-microglobulin locus in mouse embryonic stem cells by homologous recombination. *Proc Natl Acad Sci U S A* 1989; 86: 8932-5.
- [28] Capecchi MR. Generating mice with targeted mutations. *Nat Med* 2001; 7: 1086-90.
- [29] Capecchi M. Gene targeting. How efficient can you get? *Nature* 1990; 348: 109.
- [30] Gasior SL, Olivares H, Ear U, Hari DM, Weichselbaum R, Bishop DK. Assembly of RecA-like recombinases: distinct roles for mediator proteins in mitosis and meiosis. *Proc Natl Acad Sci U S A* 2001; 98: 8411-8.
- [31] Kim JS, Krasieva TB, Kurumizaka H, Chen DJ, Taylor AM, Yokomori K. Independent and sequential recruitment of NHEJ and HR factors to DNA damage sites in mammalian cells. *J Cell Biol* 2005; 170: 341-7.
- [32] Saintigny Y, Delacote F, Boucher D, Averbek D, Lopez BS. XRCC4 in G1 suppresses homologous recombination in S/G2, in G1 checkpoint-defective cells. *Oncogene* 2007; 26: 2769-80.
- [33] Hinz JM, Yamada NA, Salazar EP, Tebbs RS, Thompson LH. Influence of double-strand-break repair pathways on radiosensitivity throughout the cell cycle in CHO cells. *DNA Repair (Amst)* 2005; 4: 782-92.
- [34] Rothkamm K, Kruger I, Thompson LH, Lohrich M. Pathways of DNA double-strand break repair during the mammalian cell cycle. *Mol Cell Biol* 2003; 23: 5706-15.
- [35] Delacote F, Lopez BS. Importance of the cell cycle phase for the choice of the appropriate DSB repair pathway, for genome stability maintenance: the trans-S double-strand break repair model. *Cell Cycle* 2008; 7: 33-8.
- [36] Aten JA, Stap J, Krawczyk PM, van Oven CH, Hoebe RA, Essers J, *et al.* Dynamics of DNA double-strand breaks revealed by clustering of damaged chromosome domains. *Science* 2004; 303: 92-5.
- [37] Takata M, Sasaki MS, Sonoda E, Morrison C, Hashimoto M, Utsumi H, *et al.* Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *EMBO J* 1998; 17: 5497-508.
- [38] Kuhstoss S, Rao RN. Analysis of the integration function of the streptomyces bacteriophage phi C31. *J Mol Biol* 1991; 222: 897-908.
- [39] Rausch H, Lehmann M. Structural analysis of the actinophage phi C31 attachment site. *Nucleic Acids Res* 1991; 19: 5187-9.
- [40] Thyagarajan B, Olivares EC, Hollis RP, Ginsburg DS, Calos MP. Site-specific genomic integration in mammalian cells mediated by phage phiC31 integrase. *Mol Cell Biol* 2001; 21: 3926-34.
- [41] Lieu PT, Machleidt T, Thyagarajan B, *et al.* Generation of site-specific retargeting platform cell lines for drug discovery using phiC31 and R4 integrases. *J Biomol Screen* 2009; 14: 1207-15.
- [42] Zhang MX, Li ZH, Fang YX, *et al.* TAT-phiC31 integrase mediates DNA recombination in mammalian cells. *J Biotechnol* 2009; 142: 107-13.
- [43] Quenneville SP, Chapdelaine P, Rousseau J, *et al.* Nucleofection of muscle-derived stem cells and myoblasts with phiC31 integrase: stable expression of a full-length-dystrophin fusion gene by human myoblasts. *Mol Ther* 2004; 10: 679-87.
- [44] Calos MP. The phiC31 integrase system for gene therapy. *Curr Gene Ther* 2006; 6: 633-45.
- [45] Liu Y LU, Ozgenç A, Thyagarajan B, *et al.* hESC engineering by integrase-mediated chromosomal targeting. *Methods Mol Biol* 2010; 229-68.
- [46] Quenneville SP, Chapdelaine P, Rousseau J, Tremblay JP. Dystrophin expression in host muscle following transplantation of muscle precursor cells modified with the phiC31 integrase. *Gene Ther* 2007; 14: 514-22.
- [47] Bertoni C, Jarrahan S, Wheeler TM, *et al.* Enhancement of plasmid-mediated gene therapy for muscular dystrophy by directed plasmid integration. *Proc Natl Acad Sci U S A* 2006; 103: 419-24.
- [48] Keravala A, Portlock JL, Nash JA, Vitrant DG, Robbins PD, Calos MP. PhiC31 integrase mediates integration in cultured synovial cells and enhances gene expression in rabbit joints. *J Gene Med* 2006; 8: 1008-17.
- [49] Chalberg TW, Genise HL, Vollrath D, Calos MP. phiC31 integrase confers genomic integration and long-term transgene expression in rat retina. *Invest Ophthalmol Vis Sci* 2005; 46: 2140-6.
- [50] Held PK, Olivares EC, Aguilar CP, Finegold M, Calos MP, Grompe M. *In vivo* correction of murine hereditary tyrosinemia type I by phiC31 integrase-mediated gene delivery. *Mol Ther* 2005; 11: 399-408.
- [51] Olivares EC, Hollis RP, Chalberg TW, Meuse L, Kay MA, Calos MP. Site-specific genomic integration produces therapeutic Factor IX levels in mice. *Nat Biotechnol* 2002; 20: 1124-8.
- [52] Chalberg TW, Portlock JL, Olivares EC, *et al.* Integration specificity of phage phiC31 integrase in the human genome. *J Mol Biol* 2006; 357: 28-48.
- [53] Keravala A, Lee S, Thyagarajan B, *et al.* Mutational derivatives of PhiC31 integrase with increased efficiency and specificity. *Mol Ther* 2009; 17: 112-20.

- [54] Sarkar I, Hauber I, Hauber J, Buchholz F. HIV-1 proviral DNA excision using an evolved recombinase. *Science* 2007; 316: 1912-5.
- [55] Buchholz F, Angrand PO, Stewart AF. Improved properties of FLP recombinase evolved by cycling mutagenesis. *Nat Biotechnol* 1998; 16: 657-62.
- [56] Voziyanov Y, Konieczka JH, Stewart AF, Jayaram M. Stepwise manipulation of DNA specificity in FLP recombinase: progressively adapting FLP to individual and combinatorial mutations in its target site. *J Mol Biol* 2003; 326: 65-76.
- [57] Santoro SW, Schultz PG. Directed evolution of the site specificity of Cre recombinase. *Proc Natl Acad Sci U S A* 2002; 99: 4185-90.
- [58] Akopian A, He J, Boocock MR, Stark WM. Chimeric recombinases with designed DNA sequence recognition. *Proc Natl Acad Sci U S A* 2003; 100: 8688-91.
- [59] Gordley RM, Smith JD, Graslund T, Barbas CF, 3rd. Evolution of programmable zinc finger-recombinases with activity in human cells. *J Mol Biol* 2007; 367: 802-13.
- [60] Gordley RM, Gersbach CA, Barbas CF, 3rd. Synthesis of programmable integrases. *Proc Natl Acad Sci U S A* 2009; 106: 5053-8.
- [61] Gersbach CA, Gaj T, Gordley RM, Barbas CF, 3rd. Directed evolution of recombinase specificity by split gene reassembly. *Nucleic Acids Res* 2010; 40(1): 96-107.
- [62] Ivics Z, Izsvak Z. Transposons for gene therapy! *Curr Gene Ther* 2006; 6: 593-607.
- [63] Izsvak Z, Ivics Z. Sleeping beauty transposition: biology and applications for molecular therapy. *Mol Ther* 2004; 9: 147-56.
- [64] Vandendriessche T, Ivics Z, Izsvak Z, Chuah MK. Emerging potential of transposons for gene therapy and generation of induced pluripotent stem cells. *Blood* 2009; 114: 1461-8.
- [65] Rudin N, Haber JE. Efficient repair of HO-induced chromosomal breaks in *Saccharomyces cerevisiae* by recombination between flanking homologous sequences. *Mol Cell Biol* 1988; 8: 3918-28.
- [66] Candela H, Hake S. The art and design of genetic screens: maize. *Nat Rev Genet* 2008; 9: 192-203.
- [67] Reznikoff WS, Winterberg KM. Transposon-based strategies for the identification of essential bacterial genes. *Methods Mol Biol* 2008; 416: 13-26.
- [68] Sivasubbu S, Balciunas D, Amsterdam A, Ekker SC. Insertional mutagenesis strategies in zebrafish. *Genome Biol* 2007; 8 Suppl 1: S9.
- [69] Ivics Z, Hackett PB, Plasterk RH, Izsvak Z. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell* 1997; 91: 501-10.
- [70] Mates L, Chuah MK, Belay E, *et al.* Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. *Nat Genet* 2009; 41(6): 753-61.
- [71] Kaji K, Norrby K, Paca A, Mileikovsky M, Mohseni P, Woltjen K. Virus-free induction of pluripotency and subsequent excision of reprogramming factors. *Nature* 2009; 458: 771-5.
- [72] Woltjen K, Michael IP, Mohseni P, *et al.* piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature* 2009; 458: 766-70.
- [73] Ivics Z, Katzer A, Stuwe EE, Fiedler D, Knespel S, Izsvak Z. Targeted Sleeping Beauty transposition in human cells. *Mol Ther* 2007; 15: 1137-44.
- [74] Feng X, Bednarz AL, Colloms SD. Precise targeted integration by a chimaeric transposase zinc-finger fusion protein. *Nucleic Acids Res* 2010; 38: 1204-16.
- [75] Hackett PB, Largaespada DA, Cooper LJ. A transposon and transposase system for human application. *Mol Ther* 2010; 18: 674-83.
- [76] Singh H, Manuri PR, Olivares S, *et al.* Redirecting specificity of T-cell populations for CD19 using the Sleeping Beauty system. *Cancer Res* 2008; 68: 2961-71.
- [77] Taubes G. Gene therapy. The strange case of chimeraplasty. *Science* 2002; 298: 2116-20.
- [78] Gruenert DC, Bruscia E, Novelli G, *et al.* Sequence-specific modification of genomic DNA by small DNA fragments. *J Clin Invest* 2003; 112: 637-41.
- [79] Bedayat B, Abdolmohamadi A, Ye L, *et al.* Sequence-specific correction of genomic hypoxanthine-guanine phosphoribosyl transferase mutations in lymphoblasts by small fragment homologous replacement. *Oligonucleotides* 2010; 20: 7-16.
- [80] Bruscia E, Sanguolo F, Sinibaldi P, Goncz KK, Novelli G, Gruenert DC. Isolation of CF cell lines corrected at DeltaF508-CFTR locus by SFHR-mediated targeting. *Gene Ther* 2002; 9: 683-5.
- [81] Sanguolo F, Scaldaferrri ML, Filaretto A, *et al.* Cfr gene targeting in mouse embryonic stem cells mediated by Small Fragment Homologous Replacement (SFHR). *Front Biosci* 2008; 13: 2989-99.
- [82] Russell DW, Hirata RK. Human gene targeting by viral vectors. *Nat Genet* 1998; 18: 325-30.
- [83] Inoue N, Dong R, Hirata RK, Russell DW. Introduction of single base substitutions at homologous chromosomal sequences by adeno-associated virus vectors. *Mol Ther* 2001; 3: 526-30.
- [84] Hirata R, Chamberlain J, Dong R, Russell DW. Targeted transgene insertion into human chromosomes by adeno-associated virus vectors. *Nat Biotechnol* 2002; 20: 735-8.
- [85] Chamberlain JR, Deyle DR, Schwarze U, *et al.* Gene targeting of mutant COL1A2 alleles in mesenchymal stem cells from individuals with osteogenesis imperfecta. *Mol Ther* 2008; 16: 187-93.
- [86] Chamberlain JR, Schwarze U, Wang PR, *et al.* Gene targeting in stem cells from individuals with osteogenesis imperfecta. *Science* 2004; 303: 1198-201.
- [87] Miller DG, Wang PR, Petek LM, Hirata RK, Sands MS, Russell DW. Gene targeting *in vivo* by adeno-associated virus vectors. *Nat Biotechnol* 2006; 24: 1022-6.
- [88] Carroll D. Progress and prospects: zinc-finger nucleases as gene therapy agents. *Gene Ther* 2008; 15: 1463-8.
- [89] Paques F, Duchateau P. Meganucleases and DNA double-strand break-induced recombination: perspectives for gene therapy. *Curr Gene Ther* 2007; 7: 49-66.
- [90] Pingoud A, Silva GH. Precision genome surgery. *Nat Biotechnol* 2007; 25: 743-4.
- [91] Schleifman EB, Chin JY, Glazer PM. Triplex-mediated gene modification. *Methods Mol Biol* 2008; 435: 175-90.
- [92] Arimondo PB, Thomas CJ, Oussedik K, *et al.* Exploring the cellular activity of camptothecin-triple-helix-forming oligonucleotide conjugates. *Mol Cell Biol* 2006; 26: 324-33.
- [93] Majumdar A, Muniandy PA, Liu J, *et al.* Targeted gene knock in and sequence modulation mediated by a psoralen-linked triplex-forming oligonucleotide. *J Biol Chem* 2008; 283: 11244-52.
- [94] Simon P, Cannata F, Perrouault L, *et al.* Sequence-specific DNA cleavage mediated by bipyridine polyamide conjugates. *Nucleic Acids Res* 2008; 36: 3531-8.
- [95] Eisenschmidt K, Lanio T, Simoncsits A, *et al.* Developing a programmed restriction endonuclease for highly specific DNA cleavage. *Nucleic Acids Res* 2005; 33: 7039-47.
- [96] Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc Natl Acad Sci U S A* 1996; 93: 1156-60.
- [97] Smith J, Berg JM, Chandrasegaran S. A detailed study of the substrate specificity of a chimeric restriction enzyme. *Nucleic Acids Res* 1999; 27: 674-81.
- [98] Smith J, Bibikova M, Whitby FG, Reddy AR, Chandrasegaran S, Carroll D. Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. *Nucleic Acids Res* 2000; 28: 3361-9.
- [99] Durai S, Mani M, Kandavelou K, Wu J, Porteus MH, Chandrasegaran S. Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. *Nucleic Acids Res* 2005; 33: 5978-90.
- [100] Carroll D. Using nucleases to stimulate homologous recombination. *Methods Mol Biol* 2004; 262: 195-207.
- [101] Cathomen T, Joung JK. Zinc-finger nucleases: the next generation emerges. *Mol Ther* 2008; 16: 1200-7.
- [102] Porteus MH, Carroll D. Gene targeting using zinc finger nucleases. *Nat Biotechnol* 2005; 23: 967-73.
- [103] Lombardo A, Genovese P, Beausejour CM, *et al.* Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat Biotechnol* 2007; 25: 1298-306.
- [104] Urnov FD, Miller JC, Lee YL, *et al.* Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* 2005; 435(7042): 646-51.
- [105] Miller JC, Holmes MC, Wang J, *et al.* An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol* 2007; 25: 778-85.
- [106] Maeder ML, Thibodeau-Beganny S, Osiak A, *et al.* Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol Cell* 2008; 31: 294-301.

- [107] Hockemeyer D, Soldner F, Beard C, *et al.* Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat Biotechnol* 2009; 27: 851-7.
- [108] Zou J, Maeder ML, Mali P, *et al.* Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. *Cell Stem Cell* 2009; 5: 97-110.
- [109] Moehle EA, Rock JM, Lee YL, *et al.* Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. *Proc Natl Acad Sci U S A* 2007; 104: 3055-60.
- [110] Pabo CO, Peisach E, Grant RA. Design and selection of novel Cys2His2 zinc finger proteins. *Annu Rev Biochem* 2001; 70: 313-40.
- [111] Kim HJ, Lee HJ, Kim H, Cho SW, Kim JS. Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome Res* 2009; 19: 1279-88.
- [112] Choo Y, Klug A. Toward a code for the interactions of zinc fingers with DNA: selection of randomized fingers displayed on phage. *Proc Natl Acad Sci U S A* 1994; 91: 11163-7.
- [113] Choo Y, Klug A. Selection of DNA binding sites for zinc fingers using rationally randomized DNA reveals coded interactions. *Proc Natl Acad Sci U S A* 1994; 91: 11168-72.
- [114] Elrod-Erickson M, Rould MA, Nekludova L, Pabo CO. Zif268 protein-DNA complex refined at 1.6 Å: a model system for understanding zinc finger-DNA interactions. *Structure* 1996; 4: 1171-80.
- [115] Greisman HA, Pabo CO. A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites. *Science* 1997; 275: 657-61.
- [116] Isalan M, Klug A, Choo Y. A rapid, generally applicable method to engineer zinc fingers illustrated by targeting the HIV-1 promoter. *Nat Biotechnol* 2001; 19: 656-60.
- [117] Ramirez CL, Foley JE, Wright DA, *et al.* Unexpected failure rates for modular assembly of engineered zinc fingers. *Nat Methods* 2008; 5: 374-5.
- [118] Li T, Huang S, Jiang WZ, *et al.* TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Res* 2010.
- [119] Christian M, Cermak T, Doyle EL, *et al.* TAL Effector Nucleases Create Targeted DNA Double-strand Breaks. *Genetics* 2010.
- [120] Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. *Science* 2009; 326: 1501.
- [121] Boch J, Scholze H, Schornack S, *et al.* Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 2009; 326: 1509-12.
- [122] Lippow SM, Aha PM, Parker MH, Blake WJ, Baynes BM, Lipovsek D. Creation of a type IIS restriction endonuclease with a long recognition sequence. *Nucleic Acids Res* 2009; 37: 3061-73.
- [123] Lieber M. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem* 2010; 181-211.
- [124] Guirouilh-Barbat J, Huck S, Bertrand P, *et al.* Impact of the KU80 pathway on NHEJ-induced genome rearrangements in mammalian cells. *Mol Cell* 2004; 14: 611-23.
- [125] Guirouilh-Barbat J, Rass E, Plo I, Bertrand P, Lopez BS. Defects in XRCC4 and KU80 differentially affect the joining of distal non-homologous ends. *Proc Natl Acad Sci U S A* 2007; 104: 20902-7.
- [126] McVey M, Lee SE. MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet* 2008; 24: 529-38.
- [127] Frank KM, Sekiguchi JM, Seidl KJ, *et al.* Late embryonic lethality and impaired V(D)J recombination in mice lacking DNA ligase IV. *Nature* 1998; 396: 173-7.
- [128] Gao Y, Sun Y, Frank KM, *et al.* A critical role for DNA end-joining proteins in both lymphogenesis and neurogenesis. *Cell* 1998; 95: 891-902.
- [129] Haber JE. Alternative endings. *Proc Natl Acad Sci U S A* 2008; 105: 405-6.
- [130] Liang F, Han M, Romanienko PJ, Jasin M. Homology-directed repair is a major double-strand break repair pathway in mammalian cells. *Proc Natl Acad Sci U S A* 1998; 95: 5172-7.
- [131] Doyon Y, McCammon JM, Miller JC, *et al.* Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nat Biotechnol* 2008; 26: 702-8.
- [132] Santiago Y, Chan E, Liu PQ, *et al.* Targeted gene knockout in mammalian cells by using engineered zinc-finger nucleases. *Proc Natl Acad Sci U S A* 2008; 105: 5809-14.
- [133] Lloyd A, Plaisier CL, Carroll D, Drews GN. Targeted mutagenesis using zinc-finger nucleases in Arabidopsis. *Proc Natl Acad Sci U S A* 2005; 102: 2232-7.
- [134] Yang M, Djukanovic V, Stagg J, *et al.* Targeted mutagenesis in the progeny of maize transgenic plants. *Plant Mol Biol* 2009; 70: 669-79.
- [135] Lee HJ, Kim E, Kim JS. Targeted chromosomal deletions in human cells using zinc finger nucleases. *Genome Res* 2009; 20: 81-9.
- [136] Liu PQ, Chan EM, Cost GJ, *et al.* Generation of a triple-gene knockout mammalian cell line using engineered zinc-finger nucleases. *Biotechnol Bioeng* 2010; 106: 97-105.
- [137] Cost GJ, Freyvert Y, Vafiadis A, *et al.* BAK and BAX deletion using zinc-finger nucleases yields apoptosis-resistant CHO cells. *Biotechnol Bioeng* 2010; 105: 330-40.
- [138] Geurts AM, Cost GJ, Freyvert Y, *et al.* Knockout rats via embryo microinjection of zinc-finger nucleases. *Science* 2009; 325: 433.
- [139] Ménotet S, Iscache A-L, Tesson M, Rémy S, Usal C, Osborn MJ, *et al.* Characterisation of immunoglobulin heavy chain knockout rats. *Eur J Immunol* 2010; 40(10): 2932-41.
- [140] Mashimo T, Takizawa A, Voigt B, *et al.* Generation of knockout rats with X-linked severe combined immunodeficiency (X-SCID) using zinc-finger nucleases. *PLoS One* 2010; 5: e8870.
- [141] Shukla VK, Doyon Y, Miller JC, *et al.* Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. *Nature* 2009; 459: 437-41.
- [142] Gao H, Smith J, Yang M, *et al.* Heritable targeted mutagenesis in maize using a designed endonuclease. *Plant J* 2010; 61: 176-87.
- [143] Orłowski J, Boniecki M, Bujnicki JM. I-Ssp6803I: the first homing endonuclease from the PD-(D/E)XK superfamily exhibits an unusual mode of DNA recognition. *Bioinformatics* 2007; 23: 527-30.
- [144] Zhao L, Bonocora RP, Shub DA, Stoddard BL. The restriction fold turns to the dark side: a bacterial homing endonuclease with a PD-(D/E)-XK motif. *EMBO J* 2007; 26: 2432-42.
- [145] Gogarten JP, Hilario E. Inteins, introns, and homing endonucleases: recent revelations about the life cycle of parasitic genetic elements. *BMC Evol Biol* 2006; 6: 94.
- [146] Heath PJ, Stephens KM, Monnat RJ Jr., Stoddard BL. The structure of I-CreI, a group I intron-encoded homing endonuclease. *Nat Struct Biol* 1997; 4: 468-76.
- [147] Duan X, Gimble FS, Quiocho FA. Crystal structure of PI-SceI, a homing endonuclease with protein splicing activity. *Cell* 1997; 89: 555-64.
- [148] Grindl W, Wende W, Pingoud V, Pingoud A. The protein splicing domain of the homing endonuclease PI-sceI is responsible for specific DNA binding. *Nucleic Acids Res* 1998; 26: 1857-62.
- [149] He Z, Crist M, Yen H, Duan X, Quiocho FA, Gimble FS. Amino acid residues in both the protein splicing and endonuclease domains of the PI-SceI intein mediate DNA binding. *J Biol Chem* 1998; 273: 4607-15.
- [150] Moure CM, Gimble FS, Quiocho FA. Crystal structure of the intein homing endonuclease PI-SceI bound to its recognition sequence. *Nat Struct Biol* 2002; 9: 764-70.
- [151] Argast GM, Stephens KM, Emond MJ, Monnat RJ, Jr. I-PpoI and I-CreI homing site sequence degeneracy determined by random mutagenesis and sequential *in vitro* enrichment. *J Mol Biol* 1998; 280: 345-53.
- [152] Gimble FS, Moure CM, Posey KL. Assessing the plasticity of DNA target site recognition of the PI-SceI homing endonuclease using a bacterial two-hybrid selection system. *J Mol Biol* 2003; 334: 993-1008.
- [153] Jurica MS, Monnat RJ, Jr., Stoddard BL. DNA recognition and cleavage by the LAGLIDADG homing endonuclease I-CreI. *Mol Cell* 1998; 2: 469-76.
- [154] Chevalier BS, Stoddard BL. Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Res* 2001; 29: 3757-74.
- [155] Chevalier BS, Monnat RJ, Jr., Stoddard BL. The homing endonuclease I-CreI uses three metals, one of which is shared between the two active sites. *Nat Struct Biol* 2001; 8: 312-6.
- [156] Bolduc JM, Spiegel PC, Chatterjee P, *et al.* Structural and biochemical analyses of DNA and RNA binding by a bifunctional homing endonuclease and group I intron splicing factor. *Genes Dev* 2003; 17: 2875-88.
- [157] Chevalier BS, Kortemme T, Chadsey MS, Baker D, Monnat RJ, Stoddard BL. Design, activity, and structure of a highly specific artificial endonuclease. *Mol Cell* 2002; 10: 895-905.

- [158] Chevalier B, Turmel M, Lemieux C, Monnat RJ, Jr., Stoddard BL. Flexible DNA target site recognition by divergent homing endonuclease isoschizomers I-CreI and I-MsoI. *J Mol Biol* 2003; 329: 253-69.
- [159] Moure CM, Gimble FS, Quiocho FA. The crystal structure of the gene targeting homing endonuclease I-SceI reveals the origins of its target site specificity. *J Mol Biol* 2003; 334: 685-95.
- [160] Marcaida MJ, Prieto J, Redondo P, et al. Crystal structure of I-DmoI in complex with its target DNA provides new insights into meganuclease engineering. *Proc Natl Acad Sci U S A* 2008; 105: 16888-93.
- [161] Redondo P, Prieto J, Munoz IG, et al. Molecular basis of xeroderma pigmentosum group C DNA recognition by engineered meganucleases. *Nature* 2008; 456: 107-11.
- [162] Spiegel PC, Chevalier B, Sussman D, Turmel M, Lemieux C, Stoddard BL. The structure of I-CeuI homing endonuclease: Evolving asymmetric DNA recognition from a symmetric protein scaffold. *Structure* 2006; 14: 869-80.
- [163] Lucas P, Otis C, Mercier JP, Turmel M, Lemieux C. Rapid evolution of the DNA-binding site in LAGLIDADG homing endonucleases. *Nucleic Acids Res* 2001; 29: 960-9.
- [164] Epinat JC, Arnould S, Chames P, et al. A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells. *Nucleic Acids Res* 2003; 31: 2952-62.
- [165] Silva GH, Belfort M, Wende W, Pingoud A. From monomeric to homodimeric endonucleases and back: engineering novel specificity of LAGLIDADG enzymes. *J Mol Biol* 2006; 361: 744-54.
- [166] Seligman LM, Chisholm KM, Chevalier BS, et al. Mutations altering the cleavage specificity of a homing endonuclease. *Nucleic Acids Res* 2002; 30: 3870-9.
- [167] Sussman D, Chadsey M, Fauce S, et al. Isolation and characterization of new homing endonuclease specificities at individual target site positions. *J Mol Biol* 2004; 342: 31-41.
- [168] Rosen LE, Morrison HA, Masri S, et al. Homing endonuclease I-CreI derivatives with novel DNA target specificities. *Nucleic Acids Res* 2006.
- [169] Doyon JB, Pattanayak V, Meyer CB, Liu DR. Directed evolution and substrate specificity profile of homing endonuclease I-SceI. *J Am Chem Soc* 2006; 128: 2477-84.
- [170] Chen Z, Wen F, Sun N, Zhao H. Directed evolution of homing endonuclease I-SceI with altered sequence specificity. *Protein Eng Des Sel* 2009; 22: 249-56.
- [171] Arnould S, Chames P, Perez C, et al. Engineering of large numbers of highly specific homing endonucleases that induce recombination on novel DNA targets. *J Mol Biol* 2006; 355: 443-58.
- [172] Smith J, Grizot S, Arnould S, et al. A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences. *Nucleic Acids Res* 2006; 363(2): 283-94.
- [173] Arnould S, Perez C, Cabaniols JP, et al. Engineered I-CreI derivatives cleaving sequences from the human XPC gene can induce highly efficient gene correction in mammalian cells. *J Mol Biol* 2007; 371: 49-65.
- [174] Eastberg JH, McConnell Smith A, Zhao L, Ashworth J, Shen BW, Stoddard BL. Thermodynamics of DNA target site recognition by homing endonucleases. *Nucleic Acids Res* 2007; 35: 7209-21.
- [175] Prieto J, Redondo P, Padro D, et al. The C-terminal loop of the homing endonuclease I-CreI is essential for site recognition, DNA binding and cleavage. *Nucleic Acids Res* 2007; 35: 3262-71.
- [176] Ashworth J, Havranek JJ, Duarte CM, et al. Computational redesign of endonuclease DNA binding and cleavage specificity. *Nature* 2006; 441: 656-9.
- [177] Fajardo-Sanchez E, Stricher F, Paques F, Isalan M, Serrano L. Computer design of obligate heterodimer meganucleases allows efficient cutting of custom DNA sequences. *Nucleic Acids Res* 2008; 36(7): 2163-73.
- [178] Grizot S, Smith J, Daboussi F, et al. Efficient targeting of a SCID gene by an engineered single-chain homing endonuclease. *Nucleic Acids Res* 2009; 37: 5405-19.
- [179] Li H, Pellenz S, Ulje U, Stoddard BL, Monnat RJ, Jr. Generation of single-chain LAGLIDADG homing endonucleases from native homodimeric precursor proteins. *Nucleic Acids Res* 2009; 37(5): 1650-62.
- [180] Thyme SB, Jarjour J, Takeuchi R, et al. Exploitation of binding energy for catalysis and design. *Nature* 2009; 461: 1300-4.
- [181] Jarjour J, West-Foyle H, Certo MT, et al. High-resolution profiling of homing endonuclease binding and catalytic specificity using yeast surface display. *Nucleic Acids Res* 2009; 37: 6871-80.
- [182] Ashworth J, Taylor GK, Havranek JJ, Quadri SA, Stoddard BL, Baker D. Computational reprogramming of homing endonuclease specificity at multiple adjacent base pairs. *Nucleic Acids Res* 2010; 38(16): 5601-8.
- [183] Niu Y, Tenney K, Li H, Gimble FS. Engineering variants of the I-SceI homing endonuclease with strand-specific and site-specific DNA-nicking activity. *J Mol Biol* 2008; 382: 188-202.
- [184] McConnell Smith A, Takeuchi R, Pellenz S, et al. Generation of a nicking enzyme that stimulates site-specific gene conversion from the I-AniI LAGLIDADG homing endonuclease. *Proc Natl Acad Sci U S A* 2009; 106: 5099-104.
- [185] Jacquier A, Dujon B. An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. *Cell* 1985; 41: 383-94.
- [186] Smith F, Rouet P, Romanienko PJ, Jasin M. Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells. *Nucleic Acids Res* 1995; 23: 5012-9.
- [187] Rouet P, Smith F, Jasin M. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proc Natl Acad Sci U S A* 1994; 91: 6064-8.
- [188] Choulouka A, Perrin A, Dujon B, Nicolas JF. The yeast I-Sce I meganuclease induces site-directed chromosomal recombination in mammalian cells. *C R Acad Sci III* 1994; 317: 1013-9.
- [189] Choulouka A, Perrin A, Dujon B, Nicolas JF. Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of *Saccharomyces cerevisiae*. *Mol Cell Biol* 1995; 15: 1968-73.
- [190] Donoho G, Jasin M, Berg P. Analysis of gene targeting and intrachromosomal homologous recombination stimulated by genomic double-strand breaks in mouse embryonic stem cells. *Mol Cell Biol* 1998; 18: 4070-8.
- [191] Szczepek M, Brondani V, Buchel J, Serrano L, Segal DJ, Cathomen T. Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat Biotechnol* 2007; 25: 786-93.
- [192] Elliott B, Richardson C, Winderbaum J, Nickoloff JA, Jasin M. Gene conversion tracts from double-strand break repair in mammalian cells. *Mol Cell Biol* 1998; 18: 93-101.
- [193] Munoz IG, Prieto J, Subramanian S, et al. Molecular basis of engineered meganuclease targeting of the endogenous human RAG1 locus. *Nucleic Acids Res* 2010; [Epub ahead of Print].
- [194] Taghian DG, Nickoloff JA. Chromosomal double-strand breaks induce gene conversion at high frequency in mammalian cells. *Mol Cell Biol* 1997; 17: 6386-93.
- [195] Rouet P, Smith F, Jasin M. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol Cell Biol* 1994; 14: 8096-106.
- [196] Pierce AJ, Hu P, Han M, Ellis N, Jasin M. Ku DNA end-binding protein modulates homologous repair of double-strand breaks in mammalian cells. *Genes Dev* 2001; 15: 3237-42.
- [197] Perez C, Guyot V, Cabaniols JP, et al. Factors affecting double-strand break-induced homologous recombination in mammalian cells. *Biotechniques* 2005; 39: 109-15.
- [198] Gouble A, Smith J, Bruneau S, et al. Efficient in toto targeted recombination in mouse liver by meganuclease-induced double-strand break. *J Gene Med* 2006; 8: 616-22.
- [199] Cabaniols JP, Mathis L, Delenda C. Targeted gene modifications in drug discovery and development. *Curr Opin Pharmacol* 2009; 9: 657-63.
- [200] Cabaniols JP, Paques F. Robust cell line development using meganucleases. *Methods Mol Biol* 2008; 435: 31-45.
- [201] Moynahan ME, Jasin M. Loss of heterozygosity induced by a chromosomal double-strand break. *Proc Natl Acad Sci U S A* 1997; 94: 8988-93.
- [202] Richardson C, Moynahan ME, Jasin M. Double-strand break repair by interchromosomal recombination: suppression of chromosomal translocations. *Genes Dev* 1998; 12: 3831-42.
- [203] Horzempa J, Shanks RM, Brown MJ, Russo BC, O'Dee DM, Nau GJ. Utilization of an unstable plasmid and the I-SceI endonuclease to generate routine markerless deletion mutants in *Francisella tularensis*. *J Microbiol Methods* 2010; 80: 106-8.

- [204] Yu BJ, Kang KH, Lee JH, Sung BH, Kim MS, Kim SC. Rapid and efficient construction of markerless deletions in the *Escherichia coli* genome. *Nucleic Acids Res* 2008; 36: e84.
- [205] Flannagan RS, Linn T, Valvano MA. A system for the construction of targeted unmarked gene deletions in the genus *Burkholderia*. *Environ Microbiol* 2008; 10: 1652-60.
- [206] Cox MM, Layton SL, Jiang T, *et al.* Scarless and site-directed mutagenesis in *Salmonella enteritidis* chromosome. *BMC Biotechnol* 2007; 7: 59.
- [207] Posfai G, Kolisnychenko V, Bereczki Z, Blattner FR. Markerless gene replacement in *Escherichia coli* stimulated by a double-strand break in the chromosome. *Nucleic Acids Res* 1999; 27: 4409-15.
- [208] Windbichler N, Papathanos PA, Catteruccia F, Ranson H, Burt A, Crisanti A. Homing endonuclease mediated gene targeting in *Anopheles gambiae* cells and embryos. *Nucleic Acids Res* 2007; 35: 5922-33.
- [209] Suzuki N, Nonaka H, Tsuge Y, Okayama S, Inui M, Yukawa H. Multiple large segment deletion method for *Corynebacterium glutamicum*. *Appl Microbiol Biotechnol* 2005; 69: 151-61.
- [210] Maggert KA, Gong WJ, Golic KG. Methods for homologous recombination in *Drosophila*. *Methods Mol Biol* 2008; 420: 155-74.
- [211] Siebert R, Puchta H. Efficient repair of genomic double-strand breaks by homologous recombination between directly repeated sequences in the plant genome. *Plant Cell* 2002; 14: 1121-31.
- [212] Puchta H. Gene replacement by homologous recombination in plants. *Plant Mol Biol* 2002; 48: 173-82.
- [213] Puchta H. Use of I-Sce I to induce DNA double-strand breaks in *Nicotiana*. *Methods Mol Biol* 1999; 113: 447-51.
- [214] Puchta H. Double-strand break-induced recombination between ectopic homologous sequences in somatic plant cells. *Genetics* 1999; 152: 1173-81.
- [215] Puchta H, Dujon B, Hohn B. Two different but related mechanisms are used in plants for the repair of genomic double-strand breaks by homologous recombination. *Proc Natl Acad Sci U S A* 1996; 93: 5055-60.
- [216] Puchta H, Dujon B, Hohn B. Homologous recombination in plant cells is enhanced by *in vivo* induction of double strand breaks into DNA by a site-specific endonuclease. *Nucleic Acids Res* 1993; 21: 5034-40.
- [217] Loeber J, Pan FC, Pieler T. Generation of transgenic frogs. *Methods Mol Biol* 2009; 561: 65-72.
- [218] Takeuchi H, Georgiev O, Fetchko M, Kappeler M, Schaffner W, Egli D. *In vivo* construction of transgenes in *Drosophila*. *Genetics* 2007; 175: 2019-28.
- [219] Grabher C, Wittbrodt J. Recent advances in meganuclease-and transposon-mediated transgenesis of medaka and zebrafish. *Methods Mol Biol* 2008; 461: 521-39.
- [220] Renfer E, Amon-Hassenzahl A, Steinmetz PR, Technau U. A muscle-specific transgenic reporter line of the sea anemone, *Nematostella vectensis*. *Proc Natl Acad Sci U S A* 2010; 107: 104-8.
- [221] Sadelain M. Recent advances in globin gene transfer for the treatment of beta-thalassemia and sickle cell anemia. *Curr Opin Hematol* 2006; 13: 142-8.
- [222] Fischer A, Le Deist F, Hacein-Bey-Abina S, *et al.* Severe combined immunodeficiency. A model disease for molecular immunology and therapy. *Immunol Rev* 2005; 203: 98-109.
- [223] Thrasher AJ. New insights into the biology of Wiskott-Aldrich syndrome (WAS). *Hematology Am Soc Hematol Educ Program* 2009: 132-8.
- [224] Bouso P, Wahn V, Douagi I, *et al.* Diversity, functionality, and stability of the T cell repertoire derived *in vivo* from a single human T cell precursor. *Proc Natl Acad Sci U S A* 2000; 97: 274-8.
- [225] Zambrowicz BP, Imamoto A, Fiering S, Herzenberg LA, Kerr WG, Soriano P. Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. *Proc Natl Acad Sci U S A* 1997; 94: 3789-94.
- [226] Smith JR, Maguire S, Davis LA, *et al.* Robust, persistent transgene expression in human embryonic stem cells is achieved with AAVS1-targeted integration. *Stem Cells* 2008; 26: 496-504.
- [227] Lamartina S, Sporeno E, Fattori E, Toniatti C. Characteristics of the adeno-associated virus preintegration site in human chromosome 19: open chromatin conformation and transcription-competent environment. *J Virol* 2000; 74: 7671-7.
- [228] Dekelver RC, Choi VM, Moehle EA, Paschon DE, Hockemeyer D, Meijnsing SH, *et al.* Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome. *Genome Res* 2010; 20(8): 1132-42.
- [229] Holt N, Wang J, Kim K, Friedman G, Wang X, Taupin V, *et al.* Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 *in vivo*. *Nat Biotechnol* 2010; 28: 839-47.
- [230] Cradick TJ, Keck K, Bradshaw S, Jamieson AC, McCaffrey AP. Zinc-finger nucleases as a novel therapeutic strategy for targeting hepatitis B virus DNAs. *Mol Ther* 2010; 18: 947-54.
- [231] Ma DP, King YT, Kim Y, Luckett WS, Jr. The group I intron of apocytocrome b gene from *Chlamydomonas smithii* encodes a site-specific endonuclease. *Plant Mol Biol* 1992; 18: 1001-4.
- [232] Pruett-Miller SM, Reading DW, Porter SN, Porteus MH. Attenuation of zinc finger nuclease toxicity by small-molecule regulation of protein levels. *PLoS Genet* 2009; 5: e1000376.
- [233] Cornu TI, Thibodeau-Beganny S, Guhl E, *et al.* DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. *Mol Ther* 2008; 16: 352-8.
- [234] Richardson C, Jasin M. Frequent chromosomal translocations induced by DNA double-strand breaks. *Nature* 2000; 405: 697-700.
- [235] Brunet E, Simsek D, Tomishima M, *et al.* Chromosomal translocations induced at specified loci in human stem cells. *Proc Natl Acad Sci U S A* 2009; 106: 10620-5.
- [236] Sadelain M, Brentjens R, Riviere I. The promise and potential pitfalls of chimeric antigen receptors. *Curr Opin Immunol* 2009; 21: 215-23.
- [237] Recchia A, Bonini C, Magnani Z, *et al.* Retroviral vector integration deregulates gene expression but has no consequence on the biology and function of transplanted T cells. *Proc Natl Acad Sci U S A* 2006; 103: 1457-62.
- [238] Pierce GF, Lillicrap D, Pipe SW, Vandendriessche T. Gene therapy, bioengineered clotting factors and novel technologies for hemophilia treatment. *J Thromb Haemost* 2007; 5: 901-6.