Review

CellPress

Genome editing at the crossroads of delivery, specificity, and fidelity

Ignazio Maggio and Manuel A.F.V. Gonçalves

Leiden University Medical Center, Department of Molecular Cell Biology, Einthovenweg 20, 2333 ZC, Leiden, The Netherlands

Genome editing (GE) entails the modification of specific genomic sequences in living cells for the purpose of determining, changing, or expanding their function(s). Typically, GE occurs after delivering sequence-specific designer nucleases (e.g., ZFNs, TALENs, and CRISPR/Cas9) and donor DNA constructs into target cells. These designer nucleases can generate gene knockouts or gene knock-ins when applied alone or in combination with donor DNA templates, respectively. We review progress in this field, with an emphasis on designer nuclease and donor template delivery into mammalian target cell populations. We also discuss the impact that incremental improvements to these tools are having on the specificity and fidelity attainable with state-of-the-art DNA-editing procedures. Finally, we identify areas that warrant further investigation.

Background and scope of GE

Genetic manipulation of higher eukaryote cells plays a crucial role in basic and applied biology (Box 1). The advent and recent diversification of designer nuclease (DN) technologies (see Glossary) and their combination with nucleic acid and protein delivery systems have led to the emergence of a new field interchangeably dubbed genome engineering or GE. This biotechnology is becoming invaluable to not only interrogate but also efficiently rewrite DNA sequences in germ and somatic cells from an increasing number of organisms, including those of mammals [1,2]. Indeed, the universal role played by the genome in biological systems opens up the possibility for adapting the basic principles of GE to many disciplines and applications, including gene therapy, functional genomics, regenerative medicine, synthetic biology, and transgenesis.

Principles of DN-assisted GE

Various genetic engineering methodologies currently fall under the operative definition of GE, such as those based on site-specific recombinases (Box 2), single-stranded oligodeoxyribonucleotides (ssODNs), and recombinant adeno-associated viral vectors (rAAVs). However, we will focus on reviewing the strategies, parameters, and outcomes of GE procedures based on modifying target cell populations through the delivery of DNs, for which there is a growing and versatile portfolio (Box 3). DNs are built to generate double-stranded DNA breaks (DSBs) at predefined chromosomal positions and, in doing so, activate endogenous cellular DNA repair pathways. Indeed, the two main DNA repair pathways responsible for maintaining chromosomal integrity, non-homologous end-joining (NHEJ) and homologous recombination (HR), are activated by DSBs regardless of whether these lesions occur in a random or a site-specific fashion [3,4]. The repair of site-specific DSBs by NHEJ can create knockouts of either coding or *cis*-acting, non-coding sequences. These DSBs can also lead to knock-ins when repaired by HR events involving surrogate DSB repair substrates in the form of foreign donor DNA (Figure 1). Importantly, DNs can increase HR rates from 10^{-8} to 10^{-6} events per treated cell to frequencies as high as 1–30%. These high frequencies avoid the need for complex cell selection schemes in many experimental settings broadening, as a result, the applicability of HR-mediated GE.

Glossary

Adverse genome-modifying event (AGE): undesirable alteration(s) to the cellular DNA or to the integrant structure resulting from a genome editing (GE) intervention.
Ad.iting: designer nuclease-induced GE based on adenoviral vector donor DNA templates.
Designer nuclease (DN): engineered sequence-specific biomolecules (also known as programmable nucleases) consisting of nucleic acid binding and cleaving domains. The most commonly used are ZFNs, TALENs, and RGNs.
Genome editing (GE): the purposeful manipulation of the DNA content of living cells by adding to or removing from specific genomic sequences one or more nucleotides.
GE fidelity: the level of integrant accurateness following a targeted genomic DNA insertion event.
GE specificity: the relative frequencies of on-target versus off-target DN chromosomal cleavage or exogenous DNA chromosomal insertion.
Homologous recombination (HR): cellular DNA break-repairing mechanism involving the copying of genetic information from a donor DNA template (e.g., sister chromatid or homologous chromosome) whose sequence is identical (homologous) to the acceptor, lesion-containing, chromosomal region. HR occurs during the G2 and late S phases of the cell cycle and, in the presence of exogenous donor DNA, can be exploited for introducing genomic changes with nucleotide-level precision.
Integrant: exogenous DNA sequences once integrated in the genome of a transfected or transduced cell.
Non-homologous end-joining (NHEJ): cellular DNA break-repairing mechanism involving end-to-end ligation of DNA termini. NHEJ takes place throughout the cell cycle and can be exploited for disrupting and restoring reading-frames following DN-induced DSBs.
Protein transduction domain (PTD): a peptide (also known as a cell penetrating peptide) whose net positive charge favors plasmalemma interactions and ensuing cellular uptake of a linked cargo (normally a protein).
Transduction: the introduction of foreign nucleic acids into cells by means of viral vector particles.
Transfection: the introduction of foreign nucleic acids into cells by means of chemical or physical methods.

Corresponding author: Gonçalves, M.A.F.V. (m.goncalves@lumc.nl).
Keywords: genome editing; donor DNA; designer nucleases; delivery systems; gene targeting; high fidelity.

0167-7799/
© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
http://dx.doi.org/10.1016/j.tibtech.2015.02.011

CrossMark

Box 1. Classical genome modification technologies

The genetic manipulation of mammalian cells can generically be achieved by non-targeted and targeted chromosomal integration of exogenously added recombinant DNA. The latter genetic engineering procedures are preferable over the former because they result in uniform transgene expression, owing to reduced chromosomal positional effects and predictable phenotypes, owing to decreased risk of endogenous gene disruption (i.e., insertional mutagenesis). However, until the late 1990s the deployment of such precise genome manipulations was restricted to particular HR-based experimental systems, most notably to those involving the generation of knock-in and knockout transgenic mice [99]. In these systems, the very low HR rates and the high frequencies of random non-homologous chromosomal DNA insertions are circumvented by positive/negative selection regimens based on a combination of genetic tools and cytotoxic drugs. These strategies are, however, often difficult to apply in other biotechnological settings. Hence, early approaches aiming at genetic modification of mammalian somatic cells exploited instead the efficient, albeit non-targeted, chromosomal DNA integration capacity of γ-retroviral vectors (γ-RVs) [100]. The γ-RV-mediated genetic modification of hematopoietic stem cells from boys afflicted by X-linked severe combined immunodeficiency provided the first proof-of-concept for gene therapy and, at the same time, materialized genotoxicity risks in the form of leukemogenesis in some of the treated patients [101]. These severe adverse events (SAEs) were linked to the insertion of γ-RV genomes carrying strong promoter/enhancer elements in the vicinity of proto-oncogenes [100,101]. These insertional mutagenesis findings initiated a trend towards HIV-1-based lentiviral vectors (LVs) [102] and the use of self-inactivating retroviral backbones in which viral regulatory sequences are replaced by more physiological cellular promoters [100–102]. Of note, although LVs also display a semi-random integration profile, their proviral insertions are less biased towards the transcription start-sites of host cell genes [103,104]. Furthermore, in contrast to γ-RVs, LVs possess active nuclear import mechanisms leading to efficient transduction of non-cycling cells [105]. Although genotoxic risks associated with LV-induced insertional mutagenesis remain [106], the aforementioned tangible improvements led to therapeutic outcomes in Wiskott–Aldrich syndrome and metachromatic leukodystrophy patients [107,108]. An alternative DNA modification approach consists of adapting transposon/transposase elements from vertebrate genomes. In contrast to retroviral vectors, some of these genetic mobile elements display a truly random chromosomal insertion profile – in other words, do not show a preference for gene bodies and associated regulatory sequences [109].

Although GE has a broad sphere of action in science and technology, it will be most likely in the context of improved gene therapies that GE interventions will be put to the test in the most stringent manner, both in terms of their ultimate efficiency and safety. An initial example of such 'genome surgery', currently being tested in clinical trials, is based on zinc-finger nuclease (ZFN)-mediated *CCR5* (chemokine C-C motif receptor 5) knockout to render $CD4^+$ T cells resistant to HIV-1 infection [5]. In addition to acquired pathologies, 'genome surgery' is also being pursued and developed for tackling inherited disorders. These include the genetic complementation or correction of faulty genes underlying recessive disorders and the knockout of dominant illness-associated alleles. Crucially, several aspects linked to GE technologies require further investigation; these include devising improved methods for delivering the often large and complex GE tools as well as for increasing the specificity and accuracy of the knock-in procedures.

Box 2. Recombinase- and homing endonuclease-assisted genome engineering

The high demand for controlled chromosomal DNA insertion in both scientific and technological settings has been spurring the development of different genome manipulation technologies. In addition to artificial DNs, preeminent examples include the deployment of site-specific recombinases and integrases (e.g., CRE, FLP, and ϕC31) [110], adeno-associated virus (AAV) replicase/integrase complexes (i.e., Rep78/68) [111,112], and intron-encoded homing endonucleases (HEs), also known as meganucleases (e.g., I-*Sce*I) [113]. Although these native proteins are limited to targeting fixed chromosomal positions (e.g., ϕC31 and Rep78/68), and/or require the engineering of their cognate recognition sequences into the target cell DNA in the first place (e.g., CRE, FLP, and I-*Sce*I), they have proved to be very useful tools in particular biotechnology platforms and experimental models. For instance, site-specific recombinases have been thoroughly used for setting up conditional gene activation/deletion systems [110,114], whereas the I-*Sce*I endonuclease has been instrumental in DSB repair studies [3,4]. In fact, the latter studies based on the generation of DSBs at specific model alleles and ensuing activation of DNA repair pathways provided a strong rationale for the development of sequence-tailored designer nucleases. This research, initiated in the 1990s with the introduction of ZFNs [115], heralded the beginning of the DN-assisted GE field. More recently, the tailoring of site-specific recombinases and rare-cutting HEs to new predefined target sequences is also underway. These technologies consist of designing chimeric proteins formed by recombinase or HE domains fused to DNA-binding motifs based on zinc-finger arrays or TALE repeats [116]. In addition, strategies based on complex protein engineering endeavors aiming at altering HE target-site preference have equally been pursued [117].

Delivering the goods: introducing GE tools into target cells

Introducing plasmids encoding DNs into target cell nuclei by electroporation or by transfection based on liposome or cationic polymer formulations are common and rapid procedures, applicable to complex somatic cell populations [6,7]. Frequently, however, these methods are either inefficient in primary cells, especially those that are quiescent or slowly dividing, or lead to substantial cytotoxicity. Therefore, approaches based on delivering DNs directly as proteins or as *in vitro*-transcribed mRNA are being investigated [8,9]. Advantages of protein and mRNA delivery include avoiding insertional mutagenesis risks and lowering off-target effects owing to their shorter half-lives relative to DNA. Related to this, DNs should ideally act in a hit-and-run fashion, in other words generate site-specific DSBs and decay (or cease being expressed) to minimize off-target activities.

Transfection reagent-free strategies based on direct DN delivery can capitalize on protein transduction domains (PTDs) [8]. Genetic fusion of recombinant proteins to these positively supercharged moieties favors their uptake by cellular internalization mechanisms (e.g., lipid raft-dependent macropinocytosis) [10]. There are, however, indications that it may be difficult to generate high yields of soluble and active PTD-containing ZFNs [11,12], transcription activator-like effector nucleases (TALENs) [13], or Cas9 (CRISPR-associated 9) [14] proteins in *Escherichia coli*. Alternative approaches are chemical conjugation of DNs to PTDs, for direct plasmalemma penetration, or to specific ligands, for receptor-mediated endocytosis. Indeed,

Box 3. Characteristics and *modus operandi* of the main classes of DNs

ZFNs (Figure IA) are modular artificial proteins consisting of an array of typically 4–6 synthetic Cys_2-His_2 zinc-finger motifs fused through a linker to the catalytic domain of a restriction enzyme, traditionally that of the type IIS endonuclease *FokI* (the recognition and cleavage DNA sequences of type IIS restriction enzymes are non-overlapping). Functional ZFNs consist of two monomers assembled in a tail-to-tail orientation at the target site on opposite DNA strands. The local dimerization of the *FokI* nuclease domains catalyzes upper and lower strand DNA cleavage at a spacer sequence located between the ZFN half-target sites. As a result, a site-specific DSB is formed.

TALENs (Figure IB) display an architecture generically similar to that of ZFNs because the DNA-binding domain (DBD) is also fused via a linker to the non-specific *FokI* nuclease domain that becomes catalytically active upon *in situ* dimerization. For TALENs, however, the DBD is derived from TALE proteins found in specific phytopathogenic bacteria (e.g., genus *Xanthomomas*) and comprises an array of typically 15.5–19.5 repeat units of approximately 34 residues each. The repeat residues at positions 12 and 13, called repeat-variable di-residues (RVDs), dictate nucleotide recognition (e.g., RVDs NI, NG, and HD recognize preferentially A, T, and C, respectively). Commonly used TALEN scaffolds comprise 17.5 repeats per monomer and are encoded by ORFs of approximately 3 kb.

RGNs (Figure IC) are RNA-dependent nucleases built on components from clustered, regularly interspaced, short palindromic repeat (CRISPR)-associated Cas systems evolved in bacteria and archaea as immune mechanisms against foreign nucleic acids. The most commonly used RGNs are based on the type II CRISPR-Cas9 nuclease system from *Streptococcus pyogenes*. Cas9 is a large protein (~160 kDa) encoded by a 4.1 kb ORF and contains two nuclease domains (RuvC and HNH). This nuclease is addressed to the target site via its association with a single guide RNA (gRNA) molecule. The gRNA component is a bipartite molecule engineered by fusing a sequence-tailored CRISPR RNA (crRNA) to a scaffolding transactivating crRNA (tracrRNA). The target site consists of a nucleotide stretch matching the 5′ terminal gRNA sequence (usually 20 bp in length) followed by a short nucleotide sequence called protospacer adjacent motif (PAM; NGG, in the case of *S. pyogenes* Cas9). The fact that target sequence specificity of RGNs is governed by RNA–DNA hybridization, as opposed to protein–DNA interactions, confers versatility and multiplexing capabilities to RGNs.

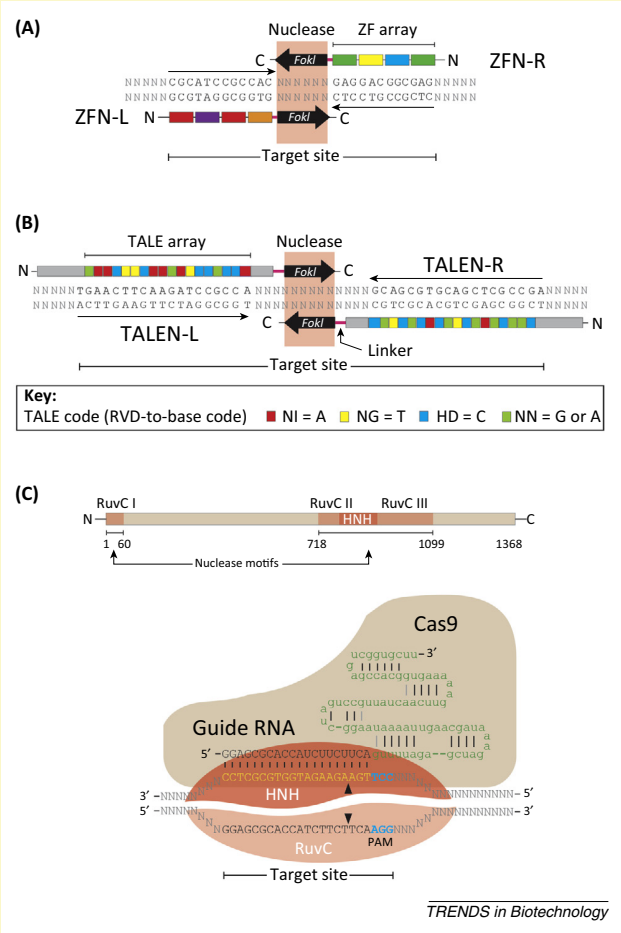


Figure I. The main classes of DNs. **(A)** Zinc-finger nucleases (ZFNs). ZF, zinc-finger; ZFN-L and ZFN-R, 'left' and 'right' ZFN monomers, respectively, bound to their cognate half-target sites. **(B)** Transcription activator-like effector nucleases (TALENs). TALEN-L and TALEN-R, 'left' and 'right' TALEN monomers, respectively, bound to their cognate half target sites. **(C)** RNA-guided nucleases (RGNs). The distribution of the nuclease motifs along the primary *S. pyogenes* Cas9 protein sequence (upper panel) and schematics of a RGN ribonucleoprotein complex bound to its target site (lower panel). Vertical arrowheads mark the position at which the blunt-ended DSB is formed. PAM, protospacer adjacent motif. Guide RNA, fusion product between crRNA and tracrRNA moieties.

gene knockouts in cell lines were detected when using TALENs whose exposed cysteine repeat residues were conjugated to a poly-arginine moiety [13]. In another study, ZFNs conjugated to transferrin were internalized via receptor-mediated endocytosis, and induced about 0.2% homology-directed gene repair of a defective reporter allele in 293 cells [15]. In other work, a cysteine-modified Cas9 nuclease conjugated to an artificial PTD was functional when used in combination with a guide RNA (gRNA) bound to another PTD via ionic interactions [14]. This methodology led to *CCR5* disruption in 8.4% of human fibroblasts and 2.3% of ESCs. Interestingly, owing to the net positive charge of their Cys_2-His_2 zinc-finger motifs, ZFNs display an intrinsic cell penetrating capacity, which can lead to targeted mutagenesis in a variety of cell types [11]. These results are significant, considering the difficulties in producing particular PTD-containing DNs.

Other DN delivery options under investigation include protein transfection procedures. For instance, *in vitro*-assembled Cas9:gRNA ribonucleoprotein complexes induced about 20% targeted mutagenesis in human fibroblasts and ESCs following electroporation [16]. Of note, compared to DNA, protein electroporation yielded approximately twofold more alkaline phosphatase-positive colonies (a marker of pluripotency), presumably reflecting a lower cytotoxicity profile. In another study, enhanced GE was achieved by combining cell cycle synchronization of human cells with electroporation of ssODN donors and Cas9:gRNA ribonucleoprotein complexes [17]. Chemical transfection agents are also being investigated for direct DN transfer. These agents might protect protein cargos from serum inhibition or protease-mediated degradation and, possibly, aid in endosomal escape. In this regard, cationic lipids were successfully used to transfect Cas9.

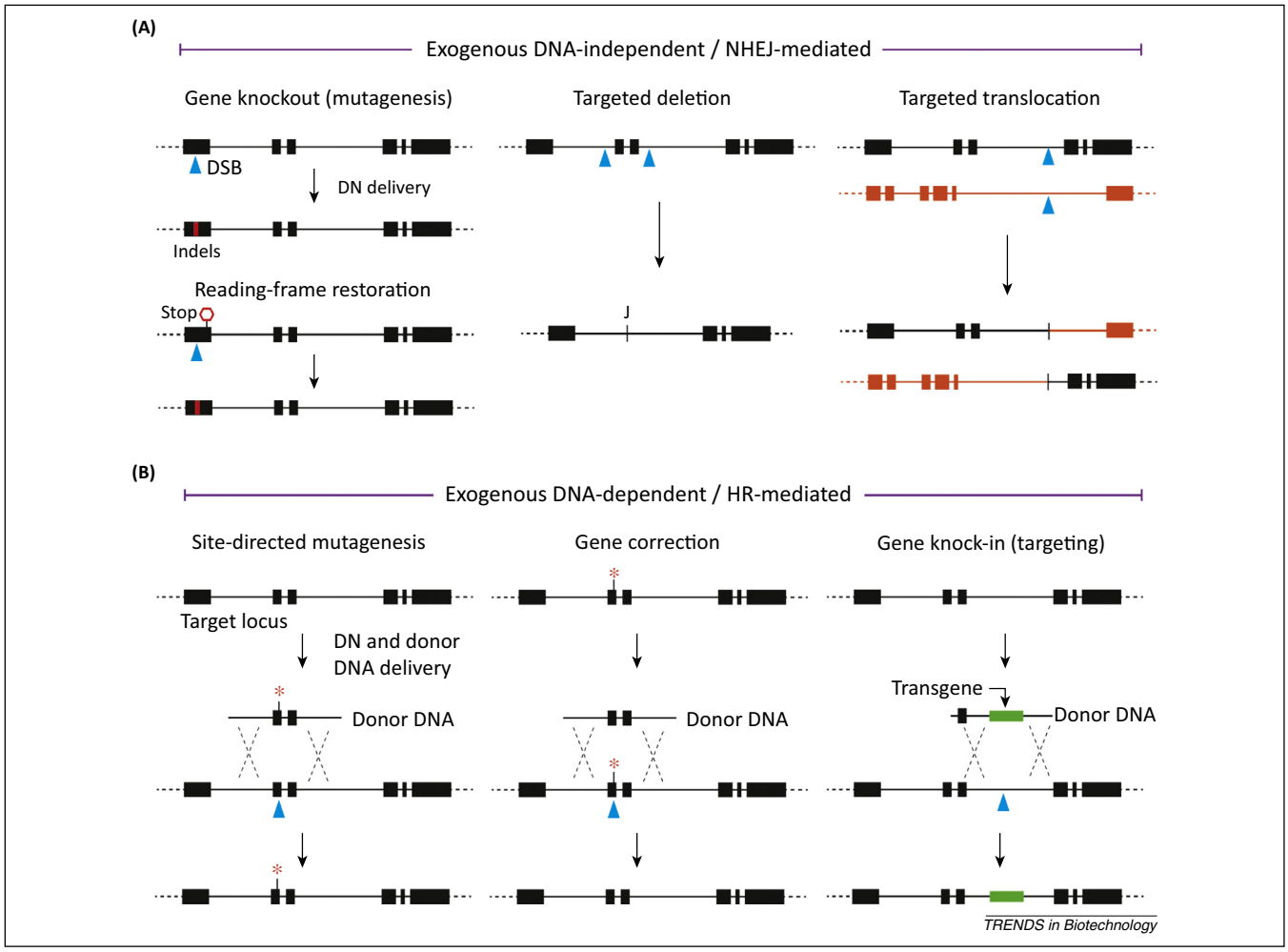


Figure 1. Illustration of the main GE strategies based on DNs. **(A)** NHEJ-mediated GE. Site-specific DSBs (cyan arrowheads) activate NHEJ repair pathways. These processes can be exploited for obtaining different GE outcomes. Left panel, site-specific DSBs can yield reading-frame shifts resulting from indels (red boxes) that disrupt and restore, respectively, in-frame and out-of-frame sequences. Middle panel, simultaneous induction of tandem DSBs by DN pairs (multiplexing) can result in the deletion of the intervening sequence following end-to-end ligation of the distal chromosomal breakpoints. Alternatively, re-insertion of the intervening sequence in an 'antisense' orientation can also occur (not drawn). Of note, indel formation is, in this case, less likely because the newly formed junction (j) should yield a sequence that is not susceptible to DN activity. Right panel, DN multiplexing applied to sequences in different chromosomes can be exploited for studying well-defined translocations. **(B)** HR-mediated GE. Site-specific DSBs can also activate the HR pathway. In the presence of foreign DNA flanked by nucleotide sequences identical to those framing the target site (donor DNA) the HR process yields well-defined junctions between the endogenous and the chromosomally integrated exogenous DNA. By using judiciously constructed donor DNA templates, this DN-induced homology-directed gene targeting can be exploited to introduce or remove specific mutations or single-nucleotide polymorphisms (asterisks) or, alternatively, insert a gene tag (not drawn), a complementary DNA (not drawn), or an entire transgene at a predefined chromosomal position (e.g., a 'safe harbor' whose prototypic example is that of the *AAVS1* locus on the human chromosome 19 at 19q13.42). Solid boxes and continuous lines represent exons and introns, respectively. Abbreviations: *AAVS1*, adeno-associated virus integration site 1; DSB, double-stranded DNA break; DN, designer nuclease; GE, genome editing; HR, homologous recombination; NHEJ, non-homologous end-joining.

To mimic the highly anionic nature of nucleic acids, Cas9 was either fused to a poly-anionic GFP or simply coupled to its naturally negatively charged gRNA partner [18]. Albeit with some cytotoxicity, this approach achieved up to 80% targeted mutagenesis in an osteosarcoma reporter cell line. In further experiments, ssODN-mediated gene repair and *in vivo* gene disruption occurred at a frequency of approximately 8% in the reporter cell line and 20% in neonatal mice, respectively [18].

As with DNA, mRNA transfer can be accomplished by electroporation or by chemical transfection [19]. Electroporation is gaining momentum, especially owing to its efficiency in introducing synthetic mRNAs into specific hematopoietic cell types [19]. In fact, a protocol combining ZFN-encoding mRNA electroporation, integrase-defective lentiviral vector (IDLV) donor DNA delivery, and stem cell viability-maintaining drugs provided a proof-of-principle for gene targeting in $CD34^+$ hematopoietic stem cells [20]. Building upon these promising results, further studies are warranted to establish whether protein- and mRNA-based GE methodologies can become streamlined and broadly applicable, including in post-mitotic cells and in conjunction with different gene knock-in systems.

Although viral vectors are generally less straightforward to produce than most non-viral delivery systems (Box 4), they constitute instrumental gene, and in some cases, protein and mRNA transfer tools [21,22]. This is a consequence of their unsurpassed efficiency in entering into many cell types, both *in vitro* and *in vivo*. Therefore, the ongoing adaptation of viral vectors to GE paradigms is becoming natural and impactful. Indeed, they are suited for many GE applications both dependent and independent

Box 4. The structure of the main gene-editing viral vectors

LVs are made via transient transfection of producer cells (e.g., 293T) with a mixture of recombinant lentiviral, packaging, and pseudotyping plasmids. The recombinant vector contains, in addition to the transgenic sequences, HIV *cis*-acting elements (i.e., 5′ and 3′ long terminal repeats, packaging signal, *Rev*-responsive element, and a central polypurine tract). The packaging and pseudotyping plasmids encode a minimal set of primary HIV products (i.e., Gag, Pol, and Rev) and a heterologous envelope protein (e.g., VSV-G), respectively. The latter moiety directs vector particle–target cell interactions. The full-length genomic RNAs are packaged in the nucleocapsids, which in turn become enveloped upon budding from the producer cell membrane. The resulting vector particles are subsequently collected from the producer cell supernatants and purified. Crucially, for generating episomal IDLVs, the packaging construct harbors specific point mutations in the *pol* region that yield non-functional integrase moieties [118].

AdVs can be divided into helper-independent and helper-dependent systems based on whether their production occurs in the absence or presence, respectively, of a *trans*-complementing helper AdV vector [119]. The former are deleted in one (first-generation) or more (second-generation) early genes essential for *in vitro* replication (i.e., *E1*, *E2A*, *E2B*, and *E4*); the latter wholly lack viral genes, and thus contain from the parental virus genome exclusively the non-coding *cis*-acting sequences involved in DNA replication and packaging (also known as 'gutless' AdVs). The foreign DNA packaging capacities of helper-independent and helper-dependent platforms range from 5–8 kb to 37 kb, respectively. The generation of AdVs starts by transfecting complementing packaging cell lines (e.g., 293, 911, or PER.C6) with recombinant AdV DNA. The rescued AdV particles are subsequently amplified through serial propagation in producer cells. The resulting AdV preparations routinely reach high titers and display high infectious/non-infectious particle ratios.

Normally, rAAV particles are assembled by transient transfection of producer cells (e.g., 293T) with a rAAV plasmid mixed with constructs expressing *rep* and *cap* (replication and packaging functions, respectively) and specific adenoviral genes (helper functions) [120]. Once in the cotransfected producer cells, the rAAV genomes, whose structure consists of inverted terminal repeat (ITR)-flanked transgene sequences, are excised from the plasmid backbone, replicated, and packaged into pre-formed empty viral capsids. The resulting rAAV particles are subsequently harvested and purified. High-titer rAAV preparations can be made despite the fact that scale-up of transient transfection procedures remains challenging with the resulting stocks often containing low infectious/non-infectious particle ratios.

of targeting donor DNA. Because the ideal mode of operation for GE tools is via 'hit-and-run', episomal viral vectors are often preferable over their chromosomally integrating counterparts. Thus, IDLVs [20,23–29], baculoviral vectors (BVs) [30–32], adenoviral vectors (AdVs) [5,28,33–43], and rAAVs [44–51] are all being adapted for transducing DNs and/or donor DNA into mammalian cells (Box 5).

So far, IDLVs have been mostly deployed for ZFN and/or donor template transfer. Albeit involving different DN delivery methods and target loci, the latter studies have established IDLV donor DNA as a prolific HR substrate in many cell types. For instance, gene knock-in frequencies of approximately 3.5% in ESCs [23], 5% in lymphocytes [23], and 9–40% in myoblasts [27,37] have been reported. Unfortunately, the optimal performance of IDLVs as a DN-expressing platform is curtailed [52] due to the susceptibility of IDLV genomes to epigenetic silencing phenomena [53,54]. In addition, *TALEN* open reading frames (ORFs) transduced by lentiviral systems suffer extensive deletions due to reverse transcriptase template switching within *TALE* repeats [34,55]. Importantly, this issue is not insurmountable because repetitive tract minimization via sequence recoding has permitted the transduction of not only intact TALEN monomers but also ZFN dimers in single lentiviral particles [29,55].

BVs have been tested in GE experiments involving the delivery of donor DNA, ZFNs, and TALENs [30,31]. The former experiments, deploying a *GFP* donor cassette, achieved up to 4.4% *CCR5*-targeted cells in ESC cultures [30]; the latter, using a 13.5 kb bicistronic donor cassette, required drug selection to enrich for glioma cell populations containing 95% *AAVS1*-targeted cells [31]. The very large packaging capacity of BV envelopes is being exploited for expressing DN dimers from single viral constructs [30,31]. As an instructive note, recent results have demonstrated that, under particular BV propagation conditions,

Box 5. The viral vector concept and its main scaffolds

Viral vectors are built to exploit the refined cellular infection mechanisms evolved by viruses for transferring their genomes into host cell nuclei. In general, these mechanisms involve the binding of virions to specific cell surface receptors, internalization, cytoplasmic trafficking, and nuclear entry of the genetic material. Viral vectors are in fact engineered viral particles whose genomes contain no or only a subset of viral genes. This renders them replication-incompetent and, simultaneously, creates space in their capsids and envelopes for the packaging of exogenous nucleic acids. The retention of *cis*-acting replication and encapsidation signals in viral vector genomes allows their assembly in so-called packaging cells. These complementing cell lines express in *trans*, either constitutively or transiently, the viral functions for which the recombinant viral genomes are deleted. Often, viral vectors are re-targeted to new cell types through the modification (pseudotyping) of their capsids or envelopes. Importantly, viral vectors have been under preclinical and clinical development for many years as 'classical' gene therapy, oncolytic, or recombinant vaccine agents. These efforts encompass vigorous testing and regulatory history build-up. Such track-records bode well for the implementation of viral vectors in GE-based translational research.

Lentiviral vectors (LVs), integrating or otherwise, are mostly based on HIV-1 (*Retroviridae* family). This virus contains two ssRNA molecules (~9.5 kb) packaged in a nucleocapsid shell surrounded by a phospholipid envelope (~120 nm). After host cell entry, triggered by envelope glycoprotein–cell surface receptor interactions, the RNA genome is converted *en route* to the nucleus into linear free-ended double-stranded complementary DNA (cDNA) via virion-associated reverse transcriptases. Next, the cDNA actively enters the nucleus via its association with a karyophylic pre-integration complex. Finally, proviruses are formed through integrase-dependent semi-random host chromosomal DNA insertion [102].

Adenoviral vectors (AdVs) are derived from *Adenoviridae* family members. Adenoviruses contain a linear, protein-capped, double-stranded DNA genome (~35–40 kb) packaged in an icosahedral nonenveloped capsid (~90–100 nm) that displays protruding receptor-interacting fiber proteins [119].

Recombinant adeno-associated viral vectors (rAAVs) are derived from *Parvoviridae* family members. Because AAVs depend on an unrelated virus for replication (e.g., adenovirus), they are assigned to the *Dependovirus* genus. These viruses have a linear ssDNA genome (~4.7 kb) with two ORFs (*rep* and *cap*) packaged in a fiberless icosahedral non-enveloped capsid (~20 nm). The coding sequences are flanked by *cis*-acting inverted terminal repeats (ITRs), which can form T-shaped secondary structures [120].

TALE repeats suffer rearrangements, presumably due to the expression of recombination-enhancing factors in the producer insect cells [32].

AdVs provide for an efficient and general platform for the delivery of the main DN classes, namely, ZFNs [5,26,28,33,42,43], TALENs [34,36,37,40,41,43], and RNA-guided nucleases (RGNs) [35,37–40] both *in vitro* [5,26,28,33,34,36,37,39,40,42,43] and *in vivo* [35,38,41]. The former studies included targeted mutagenesis in $CD4^+$ T lymphocytes [5,33], mesenchymal stromal cells [39], and keratinocytes [28]; the latter encompassed modeling therapeutic and oncogenic loss-of-function phenotypes in the liver [35,41] and oncogenic chromosomal rearrangements in the lung [38]. The versatility and relatively high genetic stability of AdVs are also underscored by their capacity to co-transduce dimers of ZFNs [5,26,33,42,43] and, in their helper-dependent version (Box 4), TALENs [40,43]. The co-transduction of RGN elements in single AdV particles has also been demonstrated [35,38]. In addition, similarly to the aforementioned viral vectors, AdVs are also being explored for introducing donor HR substrates into DN-exposed target cells [28,37,40,42].

rAAVs have served for delivering ZFNs and donor DNA to target cells *in vitro* and *in vivo* [44–49]. The former studies encompassed gene repair experiments at recombinant target alleles in human cell lines and ESCs [45,46,48,49]; the latter involved a proof-of-principle for *in vivo* therapeutic GE by using hemophilia B murine models [44,47]. More recently, rAAV technology was also deployed for introducing RGN components into mice. These experiments included loss-of-function studies to investigate gene function in the mammalian brain and to model lung cancer development [50,51]. Of note, the versatility of the rAAV-RGN system is set to profit from Cas9 variants (i.e., engineered or derived from orthogonal species) whose smaller sizes bypass the limited AAV capsid capacity issue.

Delivering predictability: improving DN specificity

DN specificity can vary greatly depending on the selected set of reagents and experimental conditions. For instance, there is evidence that short-term exposure of cells to DNs reduces off-target activity and cytotoxicity [11,14,16,18,56], with off-target activity correlating with cytotoxicity [57]. Acute cytotoxic effects may possibly result from ATM (ataxia telangiectasia mutated)- or ATR (ATM- and RAD3-related)-dependent protein kinase signaling pathway activation leading to cell cycle arrest and apoptosis. This, together with the potential involvement of off-target DSBs in generating unintended mutations, 'illegitimate' recombination events, and translocations, makes DN specificity a paramount parameter for advancing GE (Box 6). Therefore, the development of DNs is often associated with considerable efforts directed towards not only determining but also maximizing their specificity.

Improved ZFN specificity has been achieved through optimizing their DNA-binding domains (DBDs), nuclease motifs, and intervening linker sequences. Regardless of the DN platform, the DBD is clearly the major determinant conferring on-target selectivity. In the case of ZFNs, DBD assembly is complicated by the fact that interaction between each zinc-finger unit and its nucleotide triplet is often affected by the surrounding context. Although some of the various assembly platforms in use [58] take into account this context-dependency, ZFN construction remains difficult and time-consuming for non-experts. Despite this, zinc-finger assembly methods can be combined with refinements to the *Fok*I-derived nuclease domain. Through cooperative dimerization, the native *Fok*I enzyme forms a catalytically active site at the nuclease domain interface. In the context of ZFNs, dimerization involving the same ZFN subunit can generate DSBs at sequences recognized by such homodimers. To minimize this issue, the dimerization interface of *Fok*I cleavage domains has been redesigned by site-directed mutagenesis to generate obligate heterodimeric ZFN variants [59,60]. Rewardingly, these variant pairs lead to a significant reduction in the number of off-target cleavage events. Furthermore, the inter-domain linker can also be optimized because its composition and length influences ZFN activity and target-site selectivity [61,62].

The construction of functional TALENs is more straightforward and predictable than that of ZFNs owing to their direct repeat-to-nucleotide correspondence and lower context-dependency. Moreover, besides displaying wider genomic space coverage, TALENs present lower

Box 6. The GE landscape: a relationship among multiple factors

Exogenous template-independent targeted gene knockout (mutagenesis) strategies are based on the repair of site-specific DSBs by the NHEJ pathway. The specificity of gene knockout strategies depends on the ratio between on-target versus off-target DSB-forming events resulting from DN activity. Off-target DSBs are undesirable in that they induce unpredictable mutagenesis and can serve as deleterious translocation-initiating substrates. On-target DSBs lead to the intended targeted mutagenesis but can also serve as translocation-initiating lesions. Exogenous template-dependent gene knock-in (targeting) strategies are based on the repair of site-specific DSBs by the HR pathway. The exogenously added donor DNA templates bear sequence identity with the target site (HR substrates) and are used as the source of new genetic information. The fidelity of gene knock-in strategies depends on the generation of integrants whose endogenous–exogenous DNA junctions are the result of HR events at the target site. Both of these strategies rely on delivery systems for introducing the different sets of GE tools (i.e., DNs or DNs plus donor DNA) into the target cells. Regardless of the GE strategy, there are sometimes, in addition to the intended gene-modifying event, stochastic AGEs either in the same cell or in other cells of the population (Figure IA).

The 'good' integrants are on-target and contain exogenous DNA-target site junctions formed by HR events; the 'bad' correspond to off-target integrants; the 'ugly' contain one junction derived from homology-directed gene targeting and another junction formed by non-HR or 'illegitimate' recombination processes (e.g., NHEJ) involving on-target or off-target DSBs. The 'bad' and the 'ugly' integrant forms can also include undefined tandem repeats formed by exogenous DNA end-to-end ligations (concatemers) and delivery vehicle-derived sequences (e.g., prokaryotic DNA and/or viral *cis*-acting elements). These integrant forms contribute to reducing the uniformity of transgene expression in genome-modified cell populations and hindering the restoration of endogenous reading frames following gene targeting and gene repair approaches (Figure IB).

The relative weight given to the 'efficiency', 'specificity', and 'fidelity' of GE procedures is in a trend towards equilibrium (Figure IC).

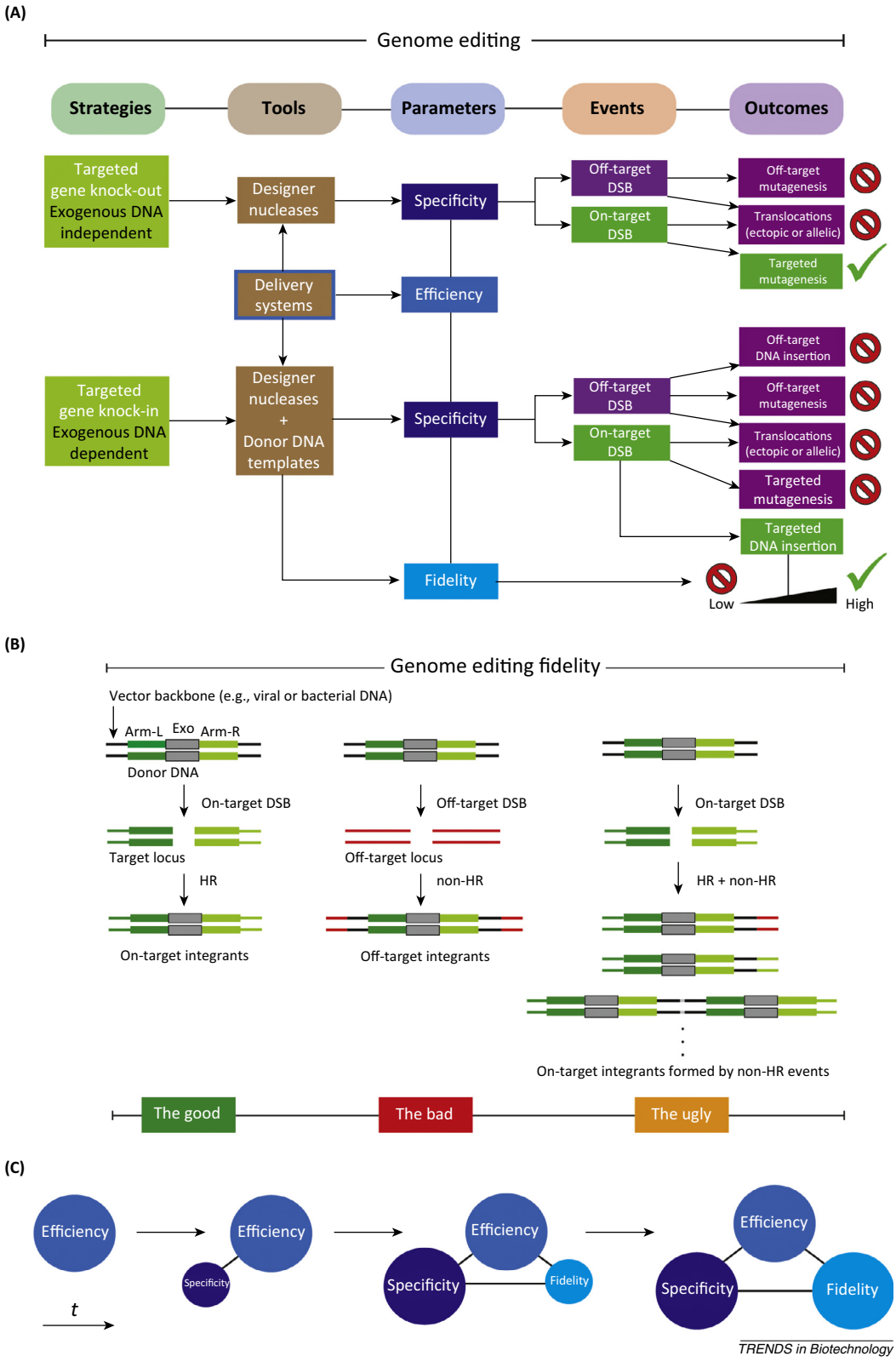


Figure I. The GE landscape. **(A)** From strategies to outcomes. **(B)** Intended and deviant GE-derived integrants. **(C)** The evolving *Zeitgeist* regarding the three principal GE parameters.

off-target activity and cytotoxicity in human cells, as revealed in side-by-side comparisons with ZFNs [63]. Because TALENs and ZFNs share a similar general architecture (Box 3), approaches to improve the already high specificity profile of TALENs are in some cases reminiscent of those applied to ZFNs. These strategies include coupling the DBDs of TALEN pairs to obligate heterodimeric *Fok*I motifs [64] and hybrid monomeric DNA-binding proteins

consisting of TALE repeats fused to homing endonuclease domains [65,66]. The specificity of TALENs is also being improved via optimizing the TALEN terminal domains as well as the length and composition of TALE arrays [67–69].

Initial studies on the specificity profile of RGNs in target cell populations indicated substantial rates of off-target mutagenesis [70–73]. Of note, more recent experiments based on whole-genome sequencing of individual clones derived from RGN-modified pluripotent stem cells (PSCs) point towards rare RGN off-target activity at the single cell level [40,74,75]. Hitherto, the specificity of RGNs has been enhanced in three principal ways: (i) using paired Cas9 'nickases' to make targeted DSBs through cooperative offset nicks on opposite DNA strands [76,77], (ii) using truncated gRNAs for fine-tuning gRNA–DNA binding energies [78], and (iii) using catalytically inactive Cas9 fused to *Fok*I nuclease domains to induce DSBs through cooperative binding of such fusion products [79,80]. Furthermore, deploying orthogonal Cas9 variants [81] that recognize different protospacer adjacent motif (PAM) sequences is yet another promising strategy for enhancing the RGN specificity profile.

Regardless of the DN platform, readily available and plentiful bioinformatics tools can greatly aid target site selection in the context of complex genomic sequences to judiciously reduce off-target DNA cleavage. Of note, beyond primary target sequence selection, the impact epigenetically regulated chromatin conformations have on the overall specificity and activity of DNs remains poorly understood. Finally, the in-depth characterization of the specificity profile of the various DN platforms will require the implementation of unbiased genome-wide and sensitive methods for tracing off-target events in treated cells [24,82–84].

Delivering precision: improving GE fidelity

The depicted sought-after GE outcomes (Figure 1) are often compounded by collateral adverse genome-modifying events (AGEs) (Box 6). In this regard, the DN-induced gene knock-in strategies are particularly multifaceted because their ultimate performance depends not only on DN efficiency and specificity but also on the frequency and fidelity with which the exogenous DNA is inserted at the target site. The deviant GE outcomes impacting the fidelity of DN-induced gene knock-in procedures are manifold. For instance, AGEs can result from the engagement of donor DNA not only with DN-induced off-target DSBs but also with breakage-prone fragile sites and DSBs created by DNA metabolic processes or environmental mutagens. Most notably, next to precise genome-modifying events, random insertions and targeted single-copy or concatemeric insertions generated through non-homologous recombination processes (e.g., NHEJ) can occur. Although targeted, the 'ugly' integrant fraction (Box 6) introduces disruptive delivery vehicle-derived backbone sequences (e.g., viral or bacterial DNA) into the cellular DNA. Moreover, in DN-exposed cell populations, a sizable fraction of target alleles will undergo gene disruption instead of homology-directed gene targeting. This stems from the fact that NHEJ occurs throughout the cell cycle and competes with HR for DSB repair [85]. Finally, targeted DSBs can lead to translocations involving not only ectopic but also allelic chromosomal sequences, as recently demonstrated [82].

The discovery that ssDNA breaks (nicks), albeit to a lesser extent than DSBs, trigger HR at recombinant and endogenous loci in mammalian cells [86–88] has provided a strong rationale for developing nicking versions of DNs. Indeed, the fact that nicks are normally not engaged by NHEJ makes such 'nickases' promising tools for increasing the overall precision of GE. Examples include developing mutant DN scaffolds [89–91] and engineered DBDs fused to heterologous sequence- and strand-specific cleaving motifs (e.g., Tev-I or MutH) [92,93].

The GE 'fidelity' parameter is expected to start receiving higher scrutiny and emphasis in the near future (Box 6). Indeed, high-fidelity GE will be instrumental for 'genome surgery' interventions and basic research (e.g., repairing defective genes and generating isogenic lines to unambiguously establish genotype–phenotype relationships, respectively). Therefore, the nature and structure of the donor DNA component is set to play a paramount role.

Early studies showed that both linear and supercoiled plasmid conformations are prone to 'illegitimate' recombination processes resulting in random chromosomal insertion and concatemerization [94]. These events can further be compounded by the incorporation of immunogenic and silencing-prone prokaryotic sequences [37]. In line with this, recent experiments targeting different loci in PSCs subjected to TALEN and donor plasmid transfections revealed that cellular fractions harboring random integrants were large, constituting in some cases the majority of the genome-modified population [95].

In settings compatible with small exogenous DNA insertions (e.g., point-mutation addition or removal), delivery of ssODNs offers a simpler and faster alternative to plasmid- and viral vector-borne donor DNA [17,18,55,96,97]. The first experiments combining DN and ssODNs resulted in 0.16% faithful repair of a defective *EGFP* allele in 293 cells following ZFN delivery. Similar levels were obtained with a control double-stranded plasmid [97]. Unfortunately, the majority of ssODN-modified cells harbored low-fidelity integrants represented by donor DNA capture and internal mutations. A subsequent study using panels of ZFNs and target cell lines demonstrated high-efficiency ssODN-based GE, with a ssODN yielding twofold higher GE levels than a corresponding plasmid substrate. Furthermore, this work also demonstrated that the single-stranded character of these templates reduces donor DNA capture at site-specific DSBs. Of note, however, DNA sequencing did reveal unintended mutations within ssODN-derived integrants [96]. Whether these mutations are the result of ssODN synthesis errors or are instead caused by intracellular DNA repairing processes requires further investigation. More recently, ssODN-based GE has been expanded to include the use of TALENs and RGNs [17,18,55].

Although IDLV genomes serve as efficient substrates for DN-assisted GE [23–25,27–29], a sizable proportion of the chromosomally integrated exogenous DNA represents inaccurate integrants which, among others, include off-targeted and concatemeric DNA forms. These fractions, which in some cases can be as high as 5–40% [24,37], are presumably generated by non-HR events

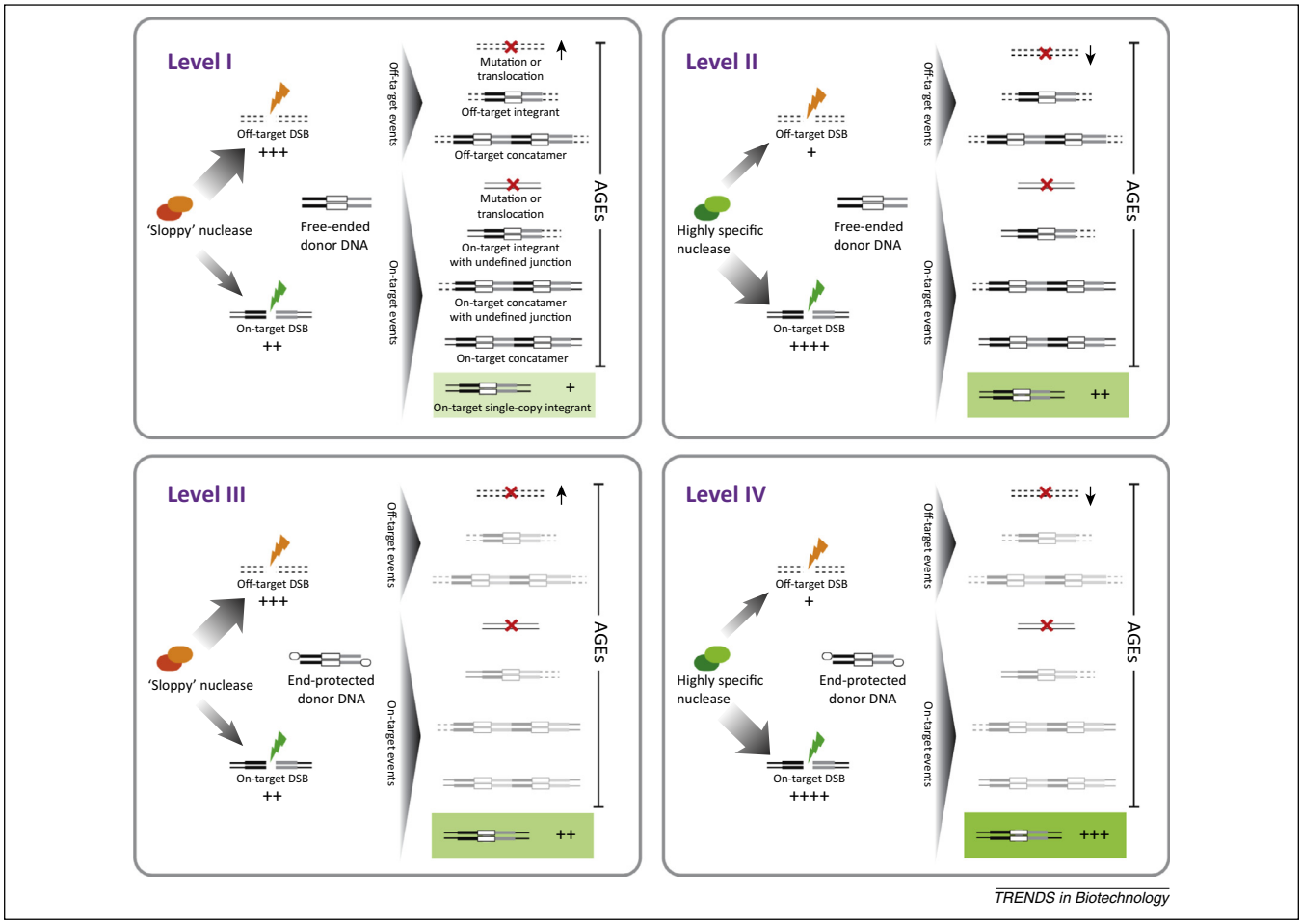


Figure 2. Grading of GE precision resulting from DNs and donor DNA templates with different specificity- and fidelity-imparting features. 'Sloppy' nucleases inducing high rates of off-target cleavage and donor DNA templates with free-ended termini yield the highest frequencies of adverse genome-modifying events (AGEs) (Level I). AGEs include mutagenesis and translocations caused by 'illegitimate' recombination-mediated repair of induced chromosomal breaks, concatemeric integrant forms, off-target integrants, and on-target integrants with undefined structures (e.g., formed by NHEJ) and/or with unwarranted composition (e.g., delivery vehicle-derived sequences). Conversely, highly-specific nucleases generating low rates of off-target cleavage and end-protected donor DNA templates yield the lowest frequencies of AGEs (Level IV). Intermediate levels of AGEs follow from using free-ended donor DNA and highly-specific nucleases (Level II), and from deploying end-protected donor DNA together with 'sloppy' nucleases (Level III). Shaded DNA forms indicate their relative lower frequencies in genome-modified cell populations. Upward and downward vertical arrows represent high and low rates, respectively, of mutations and translocations involving off-target DSBs. Abbreviations: DSB, double-stranded DNA break; GE, genome editing; NHEJ, non-homologous end-joining.

involving end-to-end ligation of IDLV genomes to each other or to site-specific and sporadic chromosomal DSBs. This picture is generally recapitulated in DSB-induced gene targeting of rAAV donor DNA, which shares with IDLV free-ended linear genomes [47,49].

Conversely, although the absolute frequencies of AdV donor DNA-modified cells are significantly lower than those obtained with IDLV, detailed side-by-side clonal analyses revealed that the vast majority of AdV-modified cells undergo proper GE [37] – that is, they harbor targeted and single-copy integrants whose chromosomal junctions result from *bona fide* HR events at both ends (Box 6). This finding of scarless DN-induced AdV DNA editing ('Ad.iting', in short) could be attributed to the capping of linear AdV genomes by covalently attached terminal proteins [37]. These protein-DNA structures presumably hinder, either in a steric or biochemical manner, 'illegitimate' recombination-mediated joining of vector genomes with each other and with spontaneous or off-target chromosomal DSBs. It is enticing to speculate a parallel operational role between the sheltered telomeric ends of linear eukaryotic chromosomes [98] and the end-protected linear AdV genomes in preventing their DNA termini from degradation and inappropriate recombination. Hence, DNs and donor DNA structures both contribute decisively to the ultimate precision of the GE process (Figure 2). Whether protein–DNA structures other than those of AdV genomes will also yield high-fidelity GE awaits further investigation.

Concluding remarks and future perspectives

The relative emphases given to the efficiency, specificity, and fidelity of GE are changing. The earlier focus on achieving meaningful levels of GE tool delivery and ensuing GE efficiency is being complemented by the relatively more recent efforts in improving the specificity of DNs and the precision with which exogenous DNA becomes incorporated at genomic target sequences. The prolific and rapidly emerging research lines reviewed in this work are systematically addressing these complementary parameters, which together underlie robust and accurate

GE. This trend is expected to continue gathering momentum and, considering the multifaceted nature of GE, be grounded on a wide range of disciplines and scientific insights.

Acknowledgments

We are grateful to our previous and current team members as well as collaborators for supporting our research in this field. This work was in part funded by grants from the Dutch Prinses Beatrix Spierfonds (W.OR11-18) and the French AFMTéléthon (16621).

References

1 Gaj, T. *et al.* (2013) ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* 31, 397–405
2 Segal, D.J. and Meckler, J.F. (2013) Genome engineering at the dawn of the golden age. *Annu. Rev. Genomics Hum. Genet.* 14, 135–158
3 Lukacsovich, T. *et al.* (1994) Repair of a specific double-strand break generated within a mammalian chromosome by yeast endonuclease I-SceI. *Nucleic Acids Res.* 22, 5649–5657
4 Rouet, P. *et al.* (1994) Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 91, 6064–6068
5 Tebas, P. *et al.* (2014) Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *N. Engl. J. Med.* 370, 901–910
6 Porteus, M.H. and Baltimore, D. (2003) Chimeric nucleases stimulate gene targeting in human cells. *Science* 300, 763
7 Urnov, F.D. *et al.* (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* 435, 646–651
8 Shi, N.Q. *et al.* (2014) A survey on 'Trojan Horse' peptides: opportunities, issues and controlled entry to 'Troy'. *J. Control Release* 194C, 53–70
9 Kuhn, A.N. *et al.* (2012) mRNA as a versatile tool for exogenous protein expression. *Curr. Gene Ther.* 12, 347–361
10 Wadia, J.S. *et al.* (2004) Transducible TAT-HA fusogenic peptide enhances escape of TAT-fusion proteins after lipid raft macropinocytosis. *Nat. Med.* 10, 310–315
11 Gaj, T. *et al.* (2012) Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. *Nat. Methods* 9, 805–807
12 Ru, R. *et al.* (2013) Targeted genome engineering in human induced pluripotent stem cells by penetrating TALENs. *Cell. Regen. (Lond.)* 2, 5
13 Liu, J. *et al.* (2014) Cell-penetrating peptide-mediated delivery of TALEN proteins via bioconjugation for genome engineering. *PLoS ONE* 9, e85755
14 Ramakrishna, S. *et al.* (2014) Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. *Genome Res.* 24, 1020–1027
15 Chen, Z. *et al.* (2013) Receptor-mediated delivery of engineered nucleases for genome modification. *Nucleic Acids Res.* 41, e182
16 Kim, S. *et al.* (2014) Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res.* 24, 1012–1019
17 Lin, S. *et al.* (2014) Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *Elife* 15, 4
18 Zuris, J.A. *et al.* (2014) Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing *in vitro* and *in vivo*. *Nat. Biotechnol.* 33, 73–80
19 Sahin, U. *et al.* (2014) mRNA-based therapeutics-developing a new class of drugs. *Nat. Rev. Drug Discov.* 13, 759–780
20 Genovese, P. *et al.* (2014) Targeted genome editing in human repopulating haematopoietic stem cells. *Nature* 510, 235–240
21 Cai, Y. *et al.* (2014) Targeted genome editing by lentiviral protein transduction of zinc-finger and TAL-effector nucleases. *Elife* 3, e01911
22 Mock, U. *et al.* (2014) Novel lentiviral vectors with mutated reverse transcriptase for mRNA delivery of TALE nucleases. *Sci. Rep.* 4, 6409
23 Lombardo, A. *et al.* (2007) Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat. Biotechnol.* 25, 1298–1306
24 Gabriel, R. *et al.* (2011) An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat. Biotechnol.* 29, 816–823
25 Lombardo, A. *et al.* (2011) Site-specific integration and tailoring of cassette design for sustainable gene transfer. *Nat. Methods* 8, 861–869
26 Provasi, E. *et al.* (2012) Editing T cell specificity towards leukemia by zinc finger nucleases and lentiviral gene transfer. *Nat. Med.* 18, 807–815
27 Benabdallah, B.F. *et al.* (2013) Targeted gene addition of microdystrophin in mice skeletal muscle via human myoblast transplantation. *Mol. Ther. Nucleic Acids* 2, e68
28 Coluccio, A. *et al.* (2013) Targeted gene addition in human epithelial stem cells by zinc-finger nuclease-mediated homologous recombination. *Mol. Ther.* 21, 1695–1704
29 Joglekar, A.V. *et al.* (2013) Integrase-defective lentiviral vectors as a delivery platform for targeted modification of adenosine deaminase locus. *Mol. Ther.* 21, 1705–1717
30 Lei, Y. *et al.* (2011) Gene editing of human embryonic stem cells via an engineered baculoviral vector carrying zinc-finger nucleases. *Mol. Ther.* 19, 942–950
31 Zhu, H. *et al.* (2013) Baculoviral transduction facilitates TALEN-mediated targeted transgene integration and Cre/LoxP cassette exchange in human-induced pluripotent stem cells. *Nucleic Acids Res.* 41, e180
32 Lau, C-H. *et al.* (2014) Genetic rearrangements of variable di-residue (RVD)-containing repeat arrays in a baculoviral TALEN system. *Mol. Ther. Methods Clin. Dev.* 1, 14050
33 Perez, E.E. *et al.* (2008) Establishment of HIV-1 resistance in $CD4^+$ T cells by genome editing using zinc-finger nucleases. *Nat. Biotechnol.* 26, 808–816
34 Holkers, M. *et al.* (2013) Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. *Nucleic Acids Res.* 41, e63
35 Ding, Q. *et al.* (2014) Permanent alteration of PCSK9 with in vivo CRISPR-Cas9 genome editing. *Circ. Res.* 115, 488–492
36 Holkers, M. *et al.* (2014) Construction and characterization of adenoviral vectors for the delivery of TALENs into human cells. *Methods* 69, 179–187
37 Holkers, M. *et al.* (2014) Adenoviral vector DNA for accurate genome editing with engineered nucleases. *Nat. Methods* 11, 1051–1057
38 Maddalo, D. *et al.* (2014) In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. *Nature* 516, 423–427
39 Maggio, I. *et al.* (2014) Adenoviral vector delivery of RNA-guided CRISPR/Cas9 nuclease complexes induces targeted mutagenesis in a diverse array of human cells. *Sci. Rep.* 4, 5105
40 Suzuki, K. *et al.* (2014) Targeted gene correction minimally impacts whole-genome mutational load in human-disease-specific induced pluripotent stem cell clones. *Cell Stem Cell* 15, 31–36
41 Zhang, S. *et al.* (2014) TALEN-mediated somatic mutagenesis in murine models of cancer. *Cancer Res.* 74, 5311–5321
42 Zhang, W. *et al.* (2014) Targeted genome correction by a single adenoviral vector simultaneously carrying an inducible zinc finger nuclease and a donor template. *J. Biotechnol.* 188C, 1–6
43 Saydaminova, K. *et al.* (2015) Efficient genome editing in hematopoietic stem cells with helper-dependent Ad5/35 vectors expressing site-specific endonucleases under microRNA regulation. *Mol. Ther. Methods Clin. Dev.* 1, 14057
44 Li, H. *et al.* (2011) In vivo genome editing restores haemostasis in a mouse model of haemophilia. *Nature* 475, 217–221
45 Händel, E.M. *et al.* (2012) Versatile and efficient genome editing in human cells by combining zinc-finger nucleases with adeno-associated viral vectors. *Hum. Gene Ther.* 23, 321–329
46 Asuri, P. *et al.* (2012) Directed evolution of adeno-associated virus for enhanced gene delivery and gene targeting in human pluripotent stem cells. *Mol. Ther.* 20, 329–338
47 Anguela, X.M. *et al.* (2013) Robust ZFN-mediated genome editing in adult hemophilic mice. *Blood* 122, 3283–3287
48 Ellis, B.L. *et al.* (2013) Zinc-finger nuclease-mediated gene correction using single AAV vector transduction and enhancement by Food and Drug Administration-approved drugs. *Gene Ther.* 20, 35–42
49 Rahman, S.H. *et al.* (2013) The nontoxic cell cycle modulator indirubin augments transduction of adeno-associated viral vectors and zinc-finger nuclease-mediated gene targeting. *Hum. Gene Ther.* 24, 67–77
50 Platt, R.J. *et al.* (2014) CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell* 159, 440–455

51 Swiech, L. *et al.* (2014) In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9. *Nat. Biotechnol.* 33, 102–106
52 Pelascini, L.P. *et al.* (2013) Histone deacetylase inhibition rescues gene knockout levels achieved with integrase-defective lentiviral vectors encoding zinc-finger nucleases. *Hum. Gene Ther. Methods* 24, 399–411
53 Kantor, B. *et al.* (2009) Epigenetic activation of unintegrated HIV-1 genomes by gut-associated short chain fatty acids and its implications for HIV infection. *Proc. Natl. Acad. Sci. U.S.A.* 106, 18786–18791
54 Pelascini, L.P. *et al.* (2013) Histone deacetylase inhibition activates transgene expression from integration-defective lentiviral vectors in dividing and non-dividing cells. *Hum. Gene Ther.* 24, 78–96
55 Yang, L. *et al.* (2013) Optimization of scarless human stem cell genome editing. *Nucleic Acids Res.* 41, 9049–9061
56 Pruett-Miller, S.M. *et al.* (2009) Attenuation of zinc finger nuclease toxicity by small-molecule regulation of protein levels. *PLoS Genet.* 5, e1000376
57 Cornu, T.I. *et al.* (2008) DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. *Mol. Ther.* 16, 352–358
58 Urnov, F.D. *et al.* (2010) Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* 11, 636–646
59 Doyon, Y. *et al.* (2011) Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat. Methods* 8, 74–79
60 Szczepek, M. *et al.* (2007) Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat. Biotechnol.* 25, 786–793
61 Händel, E.M. *et al.* (2009) Expanding or restricting the target site repertoire of zinc-finger nucleases: the inter-domain linker as a major determinant of target site selectivity. *Mol Ther.* 17, 104–111
62 Wilson, K.A. *et al.* (2013) Expanding the repertoire of target sites for zinc finger nuclease-mediated genome modification. *Mol. Ther. Nucleic Acids* 2, e88
63 Mussolino, C. *et al.* (2014) TALENs facilitate targeted genome editing in human cells with high specificity and low cytotoxicity. *Nucleic Acids Res.* 2, 6762–6773
64 Cade, L. *et al.* (2012) Highly efficient generation of heritable zebrafish gene mutations using homo- and heterodimeric TALENs. *Nucleic Acids Res.* 40, 8001–8010
65 Boissel, S. *et al.* (2014) megaTALs: a rare-cleaving nuclease architecture for therapeutic genome engineering. *Nucleic Acids Res.* 42, 2591–2601
66 Lin, J. *et al.* (2014) Creating a monomeric endonuclease TALE-I-SceI with high specificity and low genotoxicity in human cells. *Nucleic Acids Res.* 43, 1112–1122
67 Reyon, D. *et al.* (2012) FLASH assembly of TALENs for high-throughput genome editing. *Nat. Biotechnol.* 30, 460–465
68 Guilinger, J.P. *et al.* (2014) Broad specificity profiling of TALENs results in engineered nucleases with improved DNA-cleavage specificity. *Nat. Methods* 11, 429–435
69 Juillerat, A. *et al.* (2014) Comprehensive analysis of the specificity of transcription activator-like effector nucleases. *Nucleic Acids Res.* 42, 5390–5402
70 Cradick, T.J. *et al.* (2013) CRISPR/Cas9 systems targeting β-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res.* 41, 9584–9592
71 Fu, Y. *et al.* (2013) High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* 31, 822–826
72 Hsu, P.D. *et al.* (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* 31, 827–832
73 Pattanayak, V. *et al.* (2013) High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat. Biotechnol.* 31, 839–843
74 Smith, C. *et al.* (2014) Whole-genome sequencing analysis reveals high specificity of CRISPR/Cas9 and TALEN-based genome editing in human iPSCs. *Cell Stem Cell* 15, 12–13
75 Veres, A. *et al.* (2014) Low incidence of off-target mutations in individual CRISPR-Cas9 and TALEN targeted human stem cell clones detected by whole-genome sequencing. *Cell Stem Cell* 15, 27–30
76 Mali, P. *et al.* (2013) CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* 31, 833–838
77 Ran, F.A. *et al.* (2013) Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154, 1380–1389
78 Fu, Y. *et al.* (2014) Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* 32, 279–284
79 Guilinger, J.P. *et al.* (2014) Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat. Biotechnol.* 32, 577–582
80 Tsai, S.Q. *et al.* (2014) Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat. Biotechnol.* 32, 569–576
81 Esvelt, K.M. *et al.* (2013) Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat. Methods* 10, 1116–1121
82 Frock, R.L. *et al.* (2014) Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat. Biotechnol.* 33, 179–186
83 Gabriel, R. *et al.* (2015) Mapping the precision of genome editing. *Nat. Biotechnol.* 33, 150–152
84 Hendel, A. *et al.* (2015) Quantifying on- and off-target genome editing. *Trends Biotechnol.* 33, 132–140
85 Kass, E.M. and Jasin, M. (2010) Collaboration and competition between DNA double-strand break repair pathways. *FEBS Lett.* 584, 3703–3708
86 Lee, G.S. *et al.* (2004) RAG proteins shepherd double-strand breaks to a specific pathway, suppressing error-prone repair, but RAG nicking initiates homologous recombination. *Cell* 117, 171–184
87 McConnell Smith, A. *et al.* (2009) 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.* 106, 5099–5104
88 van Nierop, G.P. *et al.* (2009) Stimulation of homology-directed gene targeting at an endogenous human locus by a nicking endonuclease. *Nucleic Acids Res.* 37, 5725–5736
89 Mali, P. *et al.* (2013) RNA-guided human genome engineering via Cas9. *Science* 339, 823–826
90 Wang, J. *et al.* (2012) Targeted gene addition to a predetermined site in the human genome using a ZFN-based nicking enzyme. *Genome Res.* 22, 1316–1326
91 Ramirez, C.L. *et al.* (2012) Engineered zinc finger nickases induce homology-directed repair with reduced mutagenic effects. *Nucleic Acids Res.* 40, 5560–5568
92 Beurdeley, M. *et al.* (2013) Compact designer TALENs for efficient genome engineering. *Nat. Commun.* 4, 1762
93 Gabsalilow, L. *et al.* (2013) Site- and strand-specific nicking of DNA by fusion proteins derived from MutH and I-SceI or TALE repeats. *Nucleic Acids Res.* 41, e83
94 Bishop, J.O. and Smith, P. (1989) Mechanism of chromosomal integration of microinjected DNA. *Mol. Biol. Med.* 6, 283–298
95 Hockemeyer, D. *et al.* (2011) Genetic engineering of human pluripotent cells using TALE nucleases. *Nat. Biotechnol.* 29, 731–734
96 Chen, F. *et al.* (2011) High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nat. Methods* 8, 753–755
97 Radecke, S. *et al.* (2010) Zinc-finger nuclease-induced gene repair with oligodeoxynucleotides: wanted and unwanted target locus modifications. *Mol. Ther.* 18, 743–753
98 Stewart, J.A. *et al.* (2012) Maintaining the end: roles of telomere proteins in end-protection, telomere replication and length regulation. *Mutat. Res.* 730, 12–19
99 Capecchi, M.R. (2005) Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. *Nat. Rev. Genet.* 6, 507–512
100 Maetzig, T. *et al.* (2011) Gammaretroviral vectors: biology, technology and application. *Viruses* 3, 677–713
101 Biasco, L. *et al.* (2012) Retroviral integrations in gene therapy trials. *Mol. Ther.* 20, 709–716
102 Schambach, A. *et al.* (2013) Biosafety features of lentiviral vectors. *Hum. Gene Ther.* 24, 132–142
103 Schröder, A.R. *et al.* (2002) HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* 110, 521–529
104 Wu, X. *et al.* (2003) Transcription start regions in the human genome are favored targets for MLV integration. *Science* 300, 1749–1751
105 Matreyek, K.A. and Engelman, A. (2013) Viral and cellular requirements for the nuclear entry of retroviral preintegration nucleoprotein complexes. *Viruses* 5, 2483–2511
106 Baum, C. (2014) Data vs. dogma: HIV-1 integrations driving clonal selection. *Mol Ther.* 22, 1557–1558

107 Aiuti, A. *et al.* (2013) Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. *Science* 341, 1233151
108 Biffi, A. *et al.* (2013) Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. *Science* 341, 1233158
109 Ivics, Z. *et al.* (2009) Transposon-mediated genome manipulation in vertebrates. *Nat. Methods* 6, 415–422
110 Turan, S. and Bode, J. (2011) Site-specific recombinases: from tag-and-target- to tag-and-exchange-based genomic modifications. *FASEB J.* 25, 4088–4107
111 Recchia, A. and Mavilio, F. (2011) Site-specific integration by the adeno-associated virus rep protein. *Curr. Gene Ther.* 11, 399–405
112 González-Prieto, C. *et al.* (2013) HUH site-specific recombinases for targeted modification of the human genome. *Trends Biotechnol.* 31, 305–312
113 Marcaida, M.J. *et al.* (2010) Homing endonucleases: from basics to therapeutic applications. *Cell. Mol. Life Sci.* 67, 727–748
114 Gonçalves, M.A.F.V. *et al.* (2010) Rapid and sensitive lentivirus vector-based conditional gene expression assay to monitor and quantify cell fusion activity. *PLoS ONE* 5, e10954
115 Kim, Y.G. *et al.* (1996) Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. U.S.A.* 93, 1156–1160
116 Gaj, T. and Barbas, C.F., 3rd (2014) Genome engineering with custom recombinases. *Methods Enzymol.* 546, 79–91
117 Redondo, P. *et al.* (2008) Molecular basis of xeroderma pigmentosum group C DNA recognition by engineered meganucleases. *Nature* 456, 107–111
118 Philpott, N.J. and Thrasher, A.J. (2007) Use of nonintegrating lentiviral vectors for gene therapy. *Hum. Gene Ther.* 18, 483–489
119 Gonçalves, M.A.F.V. and de Vries, A.A.F. (2006) Adenovirus: from foe to friend. *Rev. Med. Virol.* 16, 167–186
120 Gonçalves, M.A.F.V. (2005) Adeno-associated virus: from defective virus to effective vector. *Virol. J.* 2, 43