nature medicine

REVIEW

Therapeutic genome editing: prospects and challenges

David Benjamin Turitz Cox[1–4], Randall Jeffrey Platt[1,4–6] & Feng Zhang[1,4–6]

Recent advances in the development of genome editing technologies based on programmable nucleases have substantially improved our ability to make precise changes in the genomes of eukaryotic cells. Genome editing is already broadening our ability to elucidate the contribution of genetics to disease by facilitating the creation of more accurate cellular and animal models of pathological processes. A particularly tantalizing application of programmable nucleases is the potential to directly correct genetic mutations in affected tissues and cells to treat diseases that are refractory to traditional therapies. Here we discuss current progress toward developing programmable nuclease–based therapies as well as future prospects and challenges.

Among the approximately 25,000 annotated genes in the human genome, mutations in over 3,000 have already been linked to disease phenotypes (www.omim.org/statistics/geneMap), and more disease-relevant genetic variations are being uncovered at a staggeringly rapid pace. Now, because of sharp drops in sequencing costs, the completion of the human genome project, and the exponential growth of human genome sequencing data from diseased individuals, the role of genetics in human health has become a major focus of research, clinical medicine and the development of targeted therapeutics[1]. These advances in our knowledge of the genetic basis of disease have improved our understanding of disease mechanisms and pointed toward potential therapeutic strategies. However, despite valid therapeutic hypotheses and strong efforts in drug development, there have been only a limited number of successes in using small molecules to treat diseases with strong genetic contributions[2]. Emerging therapeutic strategies that can modify nucleic acids within disease-affected cells and tissues have potential for the treatment of monogenic, highly penetrant diseases, such as severe combined immunodeficiency (SCID), hemophilia and certain enzyme deficiencies, owing to their well-defined genetics and often lack of safe, effective alternative treatments.

Two of the most powerful genetic therapeutic technologies developed thus far are gene therapy, which enables restoration of missing gene function by viral transgene expression, and RNA interference (RNAi), which mediates repression of defective genes by knockdown of the target mRNA (reviewed in refs. 3,4). Gene therapy has been used to successfully treat monogenic recessive disorders affecting the hematopoietic system, such as SCID and Wiskott-Aldrich syndrome, by semi-randomly integrating functional genes into the genome of hematopoietic stem/progenitor cells[5–7]. RNAi has been used to repress the function of genes implicated in cancer, age-related macular degeneration and transthyretin (TTR)-related amyloidosis, among others in clinical trials (www.clinicaltrials.gov trial numbers NCT00882180, NCT01961921 and NCT00259753). Despite promise and recent success, gene therapy and RNAi have limitations that preclude their utility for a large number of diseases. For example, viral gene therapy may cause mutagenesis at the insertion site and result in dysregulated transgene expression[6]. Meanwhile, the use of RNAi is limited to targets for which gene knockdown is beneficial. Also, RNAi often cannot fully repress gene expression and is therefore unlikely to provide a benefit for diseases in which complete ablation of gene function is necessary for therapy. RNAi may also have poor specificity, posing potential safety concerns and sometimes decreasing the effectiveness of treatment[8–10].

Genome editing technologies based on programmable nucleases such as meganucleases (reviewed in ref. 11), zinc finger nucleases (reviewed in ref. 12), transcription activator–like effector nucleases (reviewed in refs. 13,14) and the clustered regularly interspaced short palindromic repeat (CRISPR)-associated nuclease Cas9 (reviewed in ref. 15) are opening up the possibility of achieving therapeutic genome editing in diseased cells and tissues, resulting in the removal or correction of deleterious mutations or the insertion of protective mutations.

In this Review, we will describe the different nuclease-based genome editing technologies, the mechanisms by which they produce genetic changes, considerations for their uses in therapeutic settings and major challenges that will need to be addressed to realize their clinical translation. Although many genome editing therapeutic efforts have focused on the treatment of monogenic, highly penetrant disorders, we also discuss intriguing treatment strategies to apply this class of therapy to diseases whose genetic underpinnings are more complex, such as viral infections and polygenic diseases (**Box 1**).

[1]Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA. [2]Department of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. [3]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. [4]McGovern Institute for Brain Research at MIT, Cambridge, Massachusetts, USA. [5]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. [6]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. Correspondence should be addressed to F.Z. (zhang@broadinstitute.org).

Genome editing technologies

Programmable nucleases enable precise genome editing by introducing DNA double-strand breaks (DSBs) at specific genomic loci. DSBs subsequently recruit endogenous repair machinery for either non-homologous end-joining (NHEJ) or homology-directed repair (HDR) to the DSB site to mediate genome editing[16].

To date, four major classes of nucleases—meganucleases and their derivatives[17–20], zinc finger nucleases (ZFNs)[21–29], transcription activator–like effector nucleases (TALENs)[30–35] and CRISPR-associated nuclease Cas9 (refs. 36–44)—have been developed to enable site-specific genome editing (**Table 1**). These nuclease systems can be broadly classified into two categories based on their mode of DNA recognition: ZFNs, TALENs and meganucleases achieve specific DNA binding via protein-DNA interactions, whereas Cas9 is targeted to specific DNA sequences by a short RNA guide molecule that base-pairs directly with the target DNA and by protein-DNA interactions.

Meganucleases are endonucleases with large (>14-bp) recognition sites, the DNA binding domains of which are also responsible for cleavage of target sequences[19]. ZFNs and TALENs are chimeric enzymes consisting of a DNA binding domain fused to the sequence-agnostic FokI nuclease domain[21,32]. Re-targeting of ZFNs and meganucleases requires protein engineering, whereas re-targeting of TALENs requires complex molecular cloning[19,45,46]. In contrast, the Cas9 protein is invariant and can be easily re-targeted to new DNA sequences by changing a small portion of the sequence of an accompanying RNA guide that base-pairs directly with target DNA. Another potential advantage of Cas9 is its ability to introduce multiple DSBs in the same cell (also referred to as multiplexing) via expression of distinct guide RNAs. All four types of nucleases have been shown to achieve efficient genome editing in a wide range of model organisms and mammalian cells, and efforts are now underway in both industry and academia to develop these tools as therapeutics[47–50].

Once the DSB has been made, the lesion may be repaired by either NHEJ or HDR depending on the cell state and the presence of a repair template (**Fig. 1**). NHEJ may repair the lesion by directly rejoining the two DSB ends in a process that does not require a repair template. Although NHEJ-mediated DSB repair can be accurate, repeated repair of the same DSB by NHEJ machinery eventually results in the formation of small insertion or deletion mutations (indels) bridging the break site. Indels introduced into the coding sequence of a gene can cause frameshift mutations that lead to mRNA degradation by nonsense-mediated decay or result in the production of nonfunctional truncated proteins[51]. Thus, NHEJ may be used to suppress gene function similarly to RNAi, but it may lead to permanent inactivation by introducing loss-of-function mutations into the gene in targeted cells.

In comparison, HDR allows researchers to use an exogenous DNA template to specify the outcome of the DSB repair[16,23,52–56]. Upon introduction of a targeted DSB, HDR machinery may use exogenously provided single- or double-stranded DNA templates with sequence similarity to the break site to synthesize DNA that is used to repair the lesion, incorporating any changes encoded in the template DNA. For example, HDR may be used along with an appropriately designed

Box 1 Using genome editing to treat non-monogenic diseases

Introduction of protective mutations for complex diseases treatment. The abundance of genetic information has made it possible to identify naturally occurring mutations that confer resistance to disease phenotypes. These mutations occur in both coding and noncoding regions of the genome and have received attention as therapeutic targets for complex, non-monogenic diseases such as cardiovascular disease[58,136], HIV[97], Alzheimer disease[59] and hemoglobinopathies[137]. Genome editing provides the possibility of introducing these protective mutations into affected individuals to reverse illness.

Many known protective mutations involve loss-of-function alleles, which can be introduced via NHEJ-mediated gene disruption. This approach has rapidly gained traction as a result of the high efficiency of NHEJ in therapeutically accessible cells, and this strategy is currently in clinical trials for the treatment of HIV[47].

Mutations that protect against disease also lie hidden outside the coding region of the genome. Recently, genome-wide association studies (GWAS) and chromatin immunoprecipitation sequencing (ChIP-seq) have been used together to identify noncoding mutations that may be important targets for genome editing therapy[137,138]. A major factor controlling the severity of sickle cell disease is the expression level of fetal hemoglobin (HbF), with increased HbF levels decreasing disease severity. A GWAS for regions controlling HbF expression identified variation within the *BCL11A* gene, the product of which is known to negatively regulate HbF expression[138,139]. This variation promotes transcription factor binding within an intron that enhances BCL11A expression in the erythroid lineage, thereby decreasing HbF expression levels in red blood cells. TALENs were used to directly remove this intron from erythroid cells, and this resulted in an increase in HbF levels[137]. However, this study was not carried out in HSCs, the cell population most therapeutically relevant to this disease. It will be interesting to see whether this approach can be extended to these cells to provide a clinical benefit for affected patients. Furthermore, it is worth noting that noncoding regions will likely hold other therapeutically important regions: data from the ENCODE project suggests that 93% of GWAS hits, disease- and trait-associated are found within noncoding regions[140].

Programmable nucleases as antivirals. In addition, programmable nucleases may be developed as antiviral therapies. In principle, nucleases may be used to target viral sequences for cleavage and subsequent destruction. Additionally, NHEJ-based mutagenesis of elements critical for viral fitness could render latent viruses incapable of propagating infection. Alternatively, multiplexed nucleases like Cas9 could be used to excise proviruses from the genomes of infected cells, leading to their degradation by cellular nucleases.

Efforts to develop genome editing nucleases for antiviral therapy have focused primarily on HIV, where large reservoirs of latent provirus can persist in the presence of anti-retroviral therapies and serve to reactivate infection once treatment is stopped. The long terminal repeats (LTRs) of HIV drive viral gene expression and are critical for viral fitness. One study recently demonstrated the possibility of mutating the proviral LTR by targeting Cas9 to cleave LTR sequences, significantly reducing the expression of HIV genes in T cells[64]. Although this is an exciting discovery, there are several additional challenges to translating these *in vitro* results to the clinic. Likely the greatest will be delivering nucleases to all HIV-carrying cells in an infected individual so as to eliminate all of the latent provirus. Currently, there are no therapeutic platforms capable of delivering genome editing nucleases to the majority of T cells. Similar strategies have shown promise with human papillomavirus (HPV)[141] and hepatitis B virus (HBV)[65,66], but most infectious diseases face the same problem as HIV: extremely efficient delivery of genome editing tools is likely to be needed to achieve complete removal of viral infection.

Table 1 Comparison of different programmable nuclease platforms

	Zinc finger nuclease	TALEN	Cas9	Meganuclease
Recognition site	Typically 9–18 bp per ZFN monomer, 18–36 bp per ZFN pair	Typically 14–20 bp per TALEN monomer, 28–40 bp per TALEN pair	22 bp (20-bp guide sequence + 2-bp protospacer adjacent motif (PAM) for *Streptococcus pyogenes* Cas9); up to 44 bp for double nicking	Between 14 and 40 bp
Specificity	Small number of positional mismatches tolerated	Small number of positional mismatches tolerated	Positional and multiple consecutive mismatches tolerated	Small number of positional mismatches tolerated
Targeting constraints	Difficult to target non-G-rich sequences	5′ targeted base must be a T for each TALEN monomer	Targeted sequence must precede a PAM	Targeting novel sequences often results in low efficiency
Ease of engineering	Difficult; may require substantial protein engineering	Moderate; requires complex molecular cloning methods	Easily re-targeted using standard cloning procedures and oligo synthesis	Difficult; may require substantial protein engineering
Immunogenicity	Likely low, as zinc fingers are based on human protein scaffold; FokI is derived from bacteria and may be immunogenic	Unknown; protein derived from *Xanthamonas* sp.	Unknown; protein derived from various bacterial species	Unknown; meganucleases may be derived from many organisms, including eukaryotes
Ease of *ex vivo* delivery	Relatively easy through methods such as electroporation and viral transduction	Relatively easy through methods such as electroporation and viral transduction	Relatively easy through methods such as electroporation and viral transduction	Relatively easy through methods such as electroporation and viral transduction
Ease of *in vivo* delivery	Relatively easy as small size of ZFN expression cassettes allows use in a variety of viral vectors	Difficult due to the large size of each TALEN and repetitive nature of DNA encoding TALENs, leading to unwanted recombination events when packaged into lentiviral vectors	Moderate: the commonly used Cas9 from *S. pyogenes* is large and may impose packaging problems for viral vectors such as AAV, but smaller orthologs exist	Relatively easy as small size of meganucleases allows use in a variety of viral vectors
Ease of multiplexing	Low	Low	High	Low

repair template to replace a mutated gene directly, thereby restoring gene function while preserving physiological regulation of gene expression.

Therapeutic genome editing strategies

Genome editing based therapy can be achieved through a number of approaches including correction or inactivation of deleterious mutations, introduction of protective mutations, addition of therapeutic transgenes and disruption of viral DNA.

Pathogenic mutations can be broadly classified as causing either gain or loss of function in a gene product. A gain-of-function mutation, such as those found in the *HTT* gene in Huntington disease (http://omim.org/entry/143100) and in *FGFR3* in achondroplasia (http://omim.org/entry/100800), results in the expression of a pathogenic gene product and may be treated by using NHEJ-mediated mutations to specifically inactivate the mutant allele while leaving the wild-type allele intact on the homologous chromosome (**Fig. 1a**). Additionally, it may be possible to treat nucleotide expansion disorders, such as spinocerebellar ataxia (http://www.omim.org/entry/164400), Huntington disease and Friedriech ataxia (http://omim.org/entry/229300), by NHEJ-based deletion of the pathogenic insertion via the creation of two DSBs on both sides of the expansion (**Fig. 1b**). A combination of DSBs may also be used to edit multiple loci to achieve a therapeutic effect.

However, some gain-of-function mutations, such as the *SOD1* G93A mutation found in some individuals with amyotrophic lateral sclerosis (ALS) (http://omim.org/entry/147450), are point mutations, which may not be sufficiently different from the functioning allele on the homologous chromosome to be distinguished by the current generation of programmable nucleases, potentially leading to an undesirable complete loss of protein function if the mutation is targeted using NHEJ. In such cases HDR could instead be used to change the gain-of-function allele to the wild-type sequence, restoring gene function and eliminating pathogenic activity while preserving physiological levels of gene expression (**Fig. 1c**). Similarly, loss-of-function mutations, such as those found in Tay-Sachs disease (http://omim.org/entry/272800), would necessitate precise sequence changes to eliminate pathogenicity, requiring HDR gene correction to revert the loss-of-function mutation to the wild-type sequence. This same logic can also be extended to mutations that protect against infectious or genetic disease, which may be loss of function—as in the case of CCR5 mutations in HIV[48,57] (**Fig. 1a**) or PCSK9 mutations in hypercholesterolemia[58]—and therefore require inactivation by NHEJ, or be change of function as for *APP* (p.A673T) in Alzheimer disease[59] and therefore require correction by HDR.

For deleterious loss-of-function mutations and protective gain-of-function mutations, a therapeutic effect may also be achieved by introducing a copy of the wild-type gene or gain-of-function mutant, respectively (**Fig. 1d**). The therapeutic transgene may be inserted into a new locus, including identified 'safe harbor' loci—regions of the genome whose disruption does not lead to discernible phenotypic effects—to restore missing gene function[60–62]. Gene insertion may also be used to stably confer on cells novel functions that protect against disease, as with the insertion of chimeric-antigen receptors (CAR) into T cells to target certain leukemias[63]. Such gene insertion strategies are similar to viral-mediated gene therapy but provide better control over transgene copy number and expression levels, which may be important for gene targets whose function is sensitive to dosage.

Programmable nucleases may also be targeted to foreign DNA, such as viral genomes that are either integrated as proviruses or maintained extrachromosomally[64–68]. Targeting of extrachromosomal DNA may lead to depletion of viral genomes, while mutagenesis of the provirus genome at important coding sequences or regulatory regions may inactivate viral replication. Additionally, multiple DSBs might be used to excise proviral genomes[64]. As viral sequences may bear little sequence similarity to the host genome, this class of treatments may produce fewer off-target effects than editing therapies targeting endogenous loci.

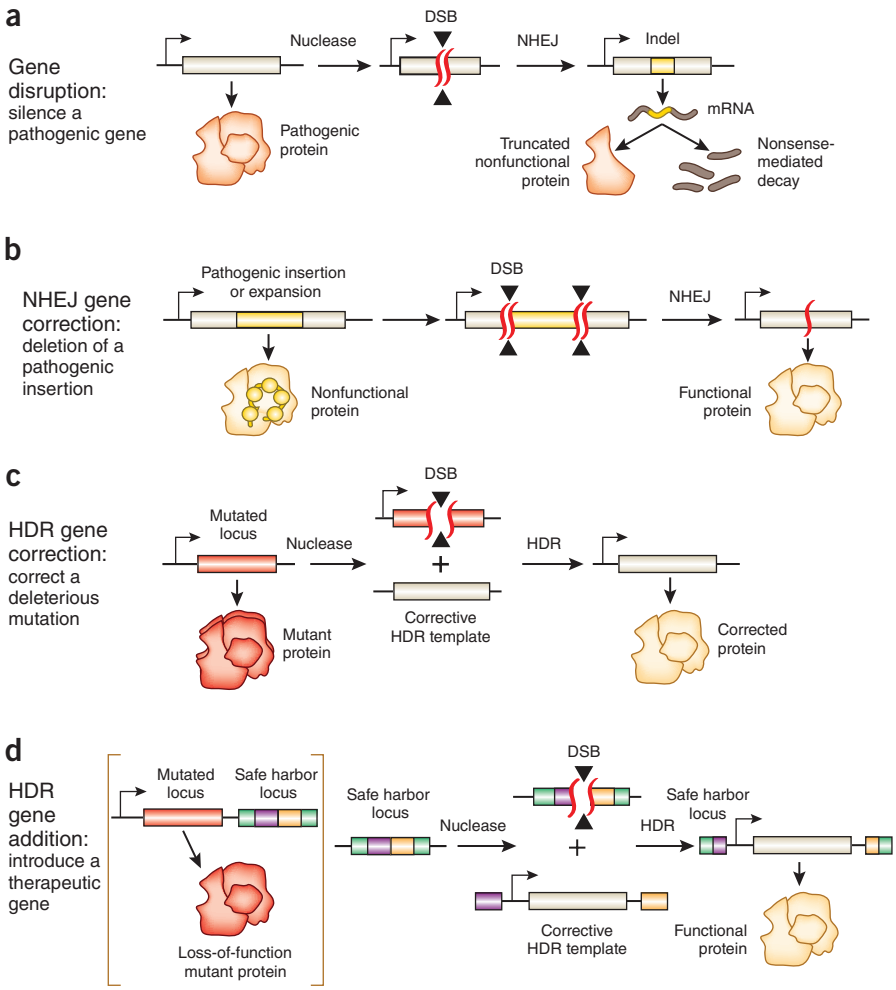

Figure 1 Types of therapeutic genome modifications. The specific type of genome editing therapy depends on the nature of the mutation causing disease. (**a**) In gene disruption, the pathogenic function of a protein is silenced by targeting the locus with NHEJ. Formation of indels in the gene of interest often results in frameshift mutations that create premature stop codons resulting in a nonfunctional protein product or nonsense-mediate decay of transcripts, suppressing gene function. Gene disruption may also be used to introduce protective loss-of-function mutations into wild-type genes to generate a therapeutic effect (**Box 1**). (**b**) In NHEJ gene correction, two DSBs targeted to both sides of a pathogenic expansion or insertion may be resolved by NHEJ, causing a therapeutic deletion of the intervening sequences. This form of treatment would require multiplexed targeting of disease-causing mutations. (**c**) HDR gene correction can be used to correct a deleterious mutation. A DSB is induced near the mutation site in the presence of an exogenously provided, corrective HDR template. HDR repair of the break site with the exogenous template corrects the mutation, restoring gene function. (**d**) An alternative to gene correction is HDR gene addition, which introduces a therapeutic transgene into a predetermined locus. This may be the native locus, a safe harbor locus or a non-native locus. A DSB is induced at the desired locus, and an HDR template containing sequence similarity to the break site, a promoter, a transgene and a polyadenylation sequence is introduced to the nucleus. HDR repair restores gene function in the target locus, albeit without true physiological control over gene expression.

Factors influencing therapeutic efficacy

Genome editing has been successfully applied to a number of diseases at the preclinical level as well as in a phase 1 clinical trial (**Table 2**)[47,49,50,69,70]. In evaluating the feasibility of a genome editing–based therapy, the therapeutic effect of the desired genetic change should first be clearly established. The subsequent success of a given strategy will depend on the ease with which a therapeutic modification 'threshold' is achieved, a criterion that is governed by the fitness of edited cells; the DSB repair pathway used to edit the genome; and the efficiency of delivery of genome editing molecules to target cell types.

Fitness of edited cells. If edited cells have an increased fitness relative to unedited cells, this will result in a selective advantage for edited cells, reducing the number of cells that initially needs to be edited to reverse disease symptoms (**Fig. 2**). For example, SCID-X1 is caused by mutations in the *IL2RG* gene, the function of which is required for proper development of the hematopoietic lymphocyte lineage (http://www.omim.org/entry/300400). Hematopoietic progenitor cells with a functional *IL2RG* gene selectively expand relative to their unedited counterparts. For example, in people with SCID-X1 who received viral gene therapy for SCID-X1 (refs. 71,72), as well as in a rare affected individual who had a spontaneous correction of a SCID-X1 mutation in a

Table 2 Examples of applications of genome editing to therapeutic models

Disease type	Nuclease platform	Therapeutic strategy	Reference(s)
Hemophilia B	ZFN	HDR-mediated insertion of correct gene sequence	49
HIV	ZFN and CRISPR	NHEJ-mediated inactivation of CCR5	47,69,70,131
Duchenne muscular dystrophy (DMD)	CRISPR and TALEN	NHEJ-mediated removal of stop codon, and HDR-mediated gene correction	132,133
Hepatitis B virus (HBV)	TALEN and CRISPR	NHEJ-mediated depletion of viral DNA	65,66
SCID	ZFN	HDR-mediated insertion of correct gene sequence	48
Cataracts	CRISPR	HDR-mediated correction of mutation in mouse zygote	134
Cystic fibrosis	CRISPR	HDR-mediated correction of CFTR in intestinal stem cell organoid	135
Hereditary tyrosinemia	CRISPR	HDR-mediated correction of mutation in liver	50

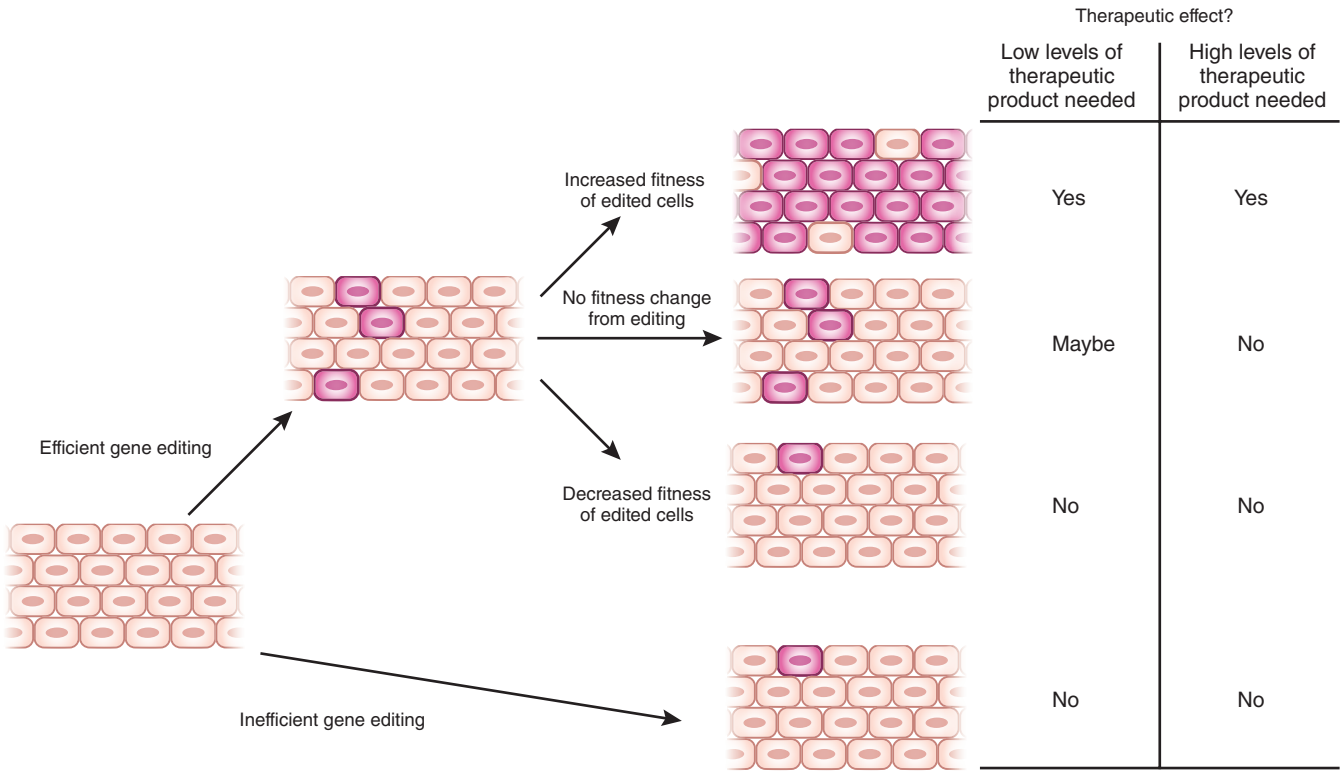

Figure 2 Factors influencing therapeutic efficacy. For a genome editing therapy to be efficacious, enough cells carrying the desired genome modification must exist in a tissue to reverse disease. If editing is efficient, treatment will create a population of cells carrying the desired genomic modification (depicted in pink). Depending on whether the editing event creates a fitness change in target cells, edited cells will proportionally increase or decrease relative to unedited cells (depicted in brown) over time in tissues. Proportionally high levels of cells carrying therapeutic genome modifications in a disease-affected tissue are likely to produce a therapeutic effect. However, if low levels of a secreted gene product are needed to reverse disease, then successfully editing only a small number of cells may be therapeutically efficacious.

T cell progenitor[73], corrected hematopoietic progenitor cells were able to overcome the lymphoid development block and expand relative to their diseased counterparts to mediate a therapeutic effect.

In contrast, in diseases in which edited cells do not exhibit a change in fitness, the number of cells that must be modified to achieve a therapeutic effect is higher. For example, chronic granulomatous disorder (CGD) is caused by mutations in genes encoding phagocytic oxidase proteins that are involved in the generation of reactive oxygen species by neutrophils to kill pathogens (http://www.omim.org/entry/306400). Dysfunction of phagocytic oxidase proteins does not influence the fitness or development of hematopoietic progenitor cells, and thus there would probably be no preferential expansion of cells edited to treat this disease. Indeed, no selective advantage for gene-corrected cells in CGD has been observed in gene therapy trials, leading to difficulties with long-term cell engraftment[74,75].

In some cases in which edited cells do not confer a change in fitness, it is still possible to reverse diseases symptoms with low numbers of therapeutically modified cells. For example, for genes that function in a non-cell-autonomous fashion, only a small number of functioning alleles may be enough to produce enough gene product to treat disease. For instance, hemophilia B is caused by mutations in the gene encoding the secreted factor IX protein involved in the blood clotting cascade; severe disease is associated with the presence of less than 1% of normal activity, and restoration of at least 1% of factor IX activity prevents the most severe bleeding conditions, while greater levels of restoration will further improve other clinically relevant complications in patients with hemophilia B[76,77]. This suggests that small changes in the amount of factor IX activity achieved through correction of mutant alleles in even a small percentage of liver cells may be therapeutic. Indeed, a study using ZFNs to correct a mouse model of hemophilia B shortly after birth demonstrated that correction of 3–7% of mutated factor IX alleles was sufficient to reverse disease symptoms, providing preclinical evidence for this hypothesis[49].

In the case where editing imposes a fitness disadvantage, such as the correction of mutated tumor suppressor genes in cancer cells, modified cells would be outcompeted by their diseased counterparts, causing the benefit of treatment to be low. The modification threshold of this final class of diseases would be extremely high, requiring many cells to be directly modified, and these diseases may not be suited for genome editing therapy. Therefore, given the current state of technology, genome editing therapies are most ideally suited for cases where editing confers a fitness advantage or where a small change in gene product levels can influence clinical outcomes.

Efficiency of genome editing. The efficiency of NHEJ- and HDR-mediated DSB repair varies substantially by cell type and cell state; in most cases, however, NHEJ is more active than HDR. This difference in activity makes it more challenging to treat diseases that require gene correction or gene insertion than those requiring gene inactivation. NHEJ is thought to be active throughout the cell cycle and has been observed in a variety of cell types, including dividing and post-mitotic cells[78,79]. NHEJ may therefore be used to facilitate high levels of gene disruption in target cell populations. In contrast, HDR acts primarily during the S/G2 phase and is therefore largely restricted to cells that are actively

dividing, limiting treatments that require precise genome modifications to mitotic cells[80,81].

The efficiency of correction by HDR may be controlled by a number of factors. First, the nature of genome modification may influence editing rates, as large HDR-mediated insertions have been found to occur at a lower rate than HDR-mediated small deletions, insertions or substitutions[61,82]. Second, the exact sequence changes made through HDR may influence therapeutic efficacy, as editing events that do not destroy the nuclease recognition site may be subject to further mutagenesis by NHEJ, potentially reducing therapeutic editing rates. Third, increasing the extent of similarity between the repair template and the DSB site may increase HDR rates, possibly by promoting the stability of D-loop intermediates formed during synthesis from a template[82–84]. Fourth, the topology of the HDR template may influence editing efficiency, as single-stranded DNA oligonucleotides and viruses may yield higher HDR rates than double-stranded substrates[85,86]. Last, suppressing competing DNA repair pathways such as NHEJ has also been shown to increase HDR rates moderately[87], although the safety of this strategy is not known and should be carefully assessed before it is implemented in a therapeutic context.

In addition to these approaches, further investigations aimed at improving HDR efficiency will be necessary to address a broader range of diseases with genome editing. Furthermore, many of these approaches may be synergistic and can be implemented in combination to increase the rate of HDR past the therapeutic editing threshold needed to treat many diseases. Despite the challenges associated with HDR, proof-of-concept preclinical HDR treatments have now been described for mouse models of hemophilia B and hereditary tyrosinemia[49,50].

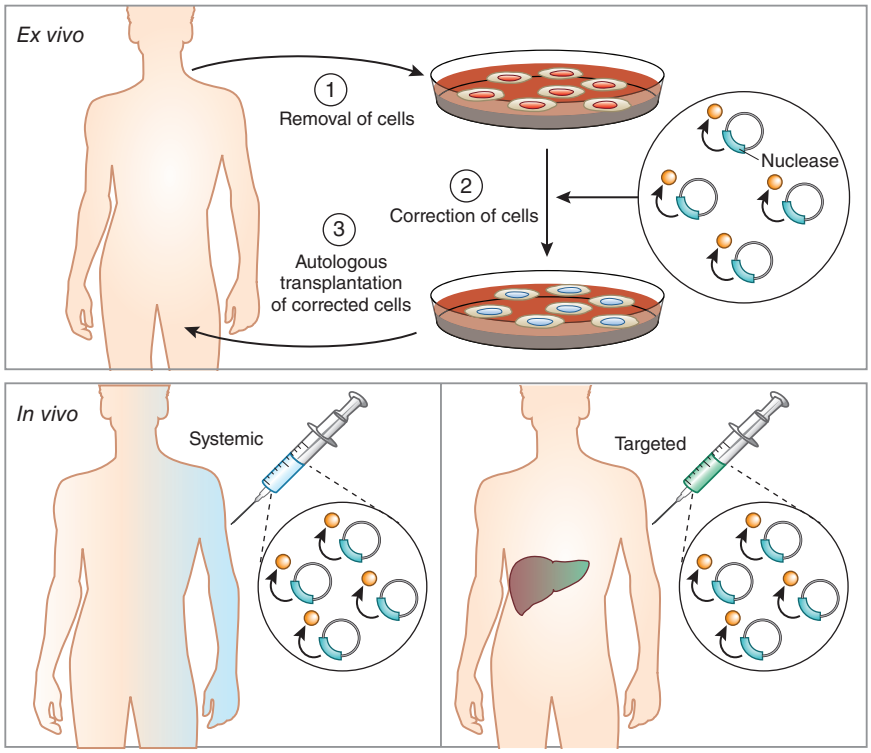

Figure 3 *Ex vivo* versus *in vivo* editing therapy. Top: in *ex vivo* editing therapy, cells are removed from a patient being treated, edited and then re-engrafted. For this mode of therapy to be successful, the target cells must be capable of surviving outside the body and homing back to target tissues after transplantation. Bottom: *in vivo* therapy involves genome editing of cells *in situ*. For *in vivo* systemic therapy (left), delivery agents that are relatively agnostic to cell identity or state would be used to effect editing in a wide range of tissue types. Alternatively, targeted *in vivo* therapy may also be achieved through targeted local injection (right) of viral vectors to the affected tissue or through the systemic injection of viral vectors with inherent tropism for specific diseased tissues, such as the eye brain, or muscle.

Modes of delivery: *ex vivo* versus *in vivo* editing. Achieving therapeutic editing requires delivery of programmable nucleases to target cells, which can be achieved either *ex vivo*, by modification and autologous transplantation of cells, or *in vivo*, by direct application of nucleases to diseased cells in the body (**Fig. 3**). Given that nucleases can potentially be mutagenic, the ideal delivery system would permit transient nuclease activity. Currently, nucleases can be delivered either as nucleic acids encoding the desired editing system or directly as proteins. While delivery of nucleic acids and proteins are both capable of achieving transient expression in target cell types, protein delivery is likely to provide the best control over nuclease dosage, since there is no signal amplification. Another important consideration is that DNA-based nuclease expression systems pose risks of insertional mutagenesis by the vector itself. So far, a variety of delivery methods have been developed.

In *ex vivo* editing therapy, the target cell population is removed from the body, modified with programmable nucleases and then transplanted back into the original host (**Fig. 3**, top). This mode of therapy allows the target cell population to be manipulated with a wide range of delivery platforms, such as electroporation, cationic lipids, cell-penetrating peptides, carbon nanowires and viral vectors. *Ex vivo* therapies are likely to achieve high editing rates as a result of the extensive development of these delivery systems for research and gene therapy applications. Moreover, many *ex vivo* therapies afford control over the specific dosage of therapeutic molecules delivered to cells. This may be particularly important when off-target modifications are a concern, as limiting the amount of nuclease may decrease such mutations[88].

However, there are two large drawbacks with *ex vivo* editing. First, target cells must be capable of surviving manipulation outside the body, which is a challenge with many tissues because cells either fail to survive or lose properties necessary for their function *in vivo*. Thus, *ex vivo* therapy is largely limited to tissues with adult stem cell populations amenable to culture and manipulation, such as the hematopoietic system. Second, cultured cells often engraft poorly upon reintroduction into the patient, decreasing the effectiveness of treatment. However, engraftment may be enhanced for hematopoietic cells by ablative conditioning regimens that deplete unedited cells before transplantation. This is clinically feasible but introduces substantial risks to patients[89].

In vivo genome editing involves direct delivery of programmable nucleases to disease-affected cells in their native tissues (**Fig. 3**, bottom). There are two advantages to *in vivo* editing therapy over *ex vivo* approaches. First, *in vivo* editing can be applied to diseases in which the affected cell population is not amenable to *ex vivo* manipulation. Second, *in vivo* delivery has the potential to target multiple tissue types, potentially allowing for the treatment of diseases that affect multiple organ systems. These properties will probably allow

in vivo treatment to be applied to a wider range of diseases than *ex vivo* therapies.

To date, *in vivo* editing has largely been achieved through the use of viral vectors with defined, tissue-specific tropism. Such vectors are currently limited in their cargo-carrying capacity and tropism, restricting this mode of therapy to organ systems where transduction with clinically useful vectors is efficient, such as the liver, muscle and eye[90–92]. Another major potential barrier to the development of *in vivo* delivery is the immune response that may be raised in response to the large amounts of virus necessary for treatment—a phenomenon that is not unique to genome editing but is observed with other virus-based gene therapies[93]. It is also possible that peptides from editing nucleases themselves could be presented on MHC class I molecules to stimulate an immune response, although there is little evidence to support this happening at the preclinical level. Another major challenge with this mode of therapy is the difficulty of controlling the distribution and consequently the dosage of genome editing nucleases *in vivo*, which can lead to off-target mutation profiles that may be difficult to predict. To address some of these concerns, nonviral delivery systems are under active development to reduce the potential risks currently associated with the use of viral vectors and expand the range of targetable tissues (reviewed in ref. 94).

The potential clinical complications faced by therapeutic genome editing overlap considerably with those of gene therapy, which makes use of similar delivery agents and results in the expression of novel gene products in the host. For a more in-depth discussion of the safety concerns regarding transgene expression and viral vectors for therapy, the reader is referred to recent reviews and studies on gene therapy[95,96].

Examples of successful genome editing therapeutic strategies

***Ex vivo* editing therapy.** The longstanding clinical expertise with the purification, culture and transplantation of hematopoietic cells has made diseases affecting the blood system such as SCID, Fanconi anemia, Wiskott-Aldrich syndrome and sickle cell anemia the focus of *ex vivo* genome editing therapy. Another reason to focus on hematopoietic cells is that, thanks to previous efforts to design gene therapy for blood disorders, delivery systems of relatively high efficiency already exist. Despite these advantages, the often low efficiency of cell engraftment upon transplantation requires that this mode of therapy be applied to diseases in which edited cells possess a fitness advantage, so that a small number of engrafted, edited cells can expand and treat disease. One such disease is HIV, as HIV infection results in a fitness disadvantage to $CD4^+$ T cells.

The rationale for using genome editing in HIV treatment originates from the observation that individuals homozygous for loss-of-function mutations in CCR5, a cellular co-receptor for the virus, are highly resistant to infection and otherwise healthy, suggesting that mimicking this mutation with genome editing could be a safe and effective therapeutic strategy[57]. This idea was clinically validated when an HIV-infected patient was given an allogeneic bone marrow transplant from a donor homozygous for a loss-of-function CCR5 mutation resulting in undetectable levels of HIV and restoration of normal $CD4^+$ T-cell counts[97]. Although bone marrow transplantation is not a realistic treatment strategy for most HIV patients, owing to the limited number of CCR5-null donors and the potential for graft-versus-host disease, HIV therapies that convert an individual's own T cells into CCR5-null cells are.

Early studies using ZFNs and NHEJ to knock out CCR5 in humanized mouse models of HIV showed that transplantation of CCR5-edited $CD4^+$ T cells improved viral load and $CD4^+$ T cell counts[70]. Importantly, these models also showed that HIV infection resulted in selection for cells not expressing CCR5. As a result of this promising study, genome editing therapy that knocks out CCR5 in the T cells of humans with HIV has now been tested. In a recent phase 1 clinical trial, $CD4^+$ T cells from patients with HIV were removed, edited with ZFNs designed to knock out the *CCR5* gene, and autologously transplanted back into the patients[47]. Early results from this trial suggest that genome editing through ZFNs of the CCR5 locus is safe, although the follow-up time has been too short to provide a full understanding of the risks and efficacy of treatment.

Gene correction strategies have also been successfully demonstrated in a recent study in which a mutated *IL2RG* gene was targeted for correction with ZFNs in hematopoietic stem cells (HSCs) obtained from a patient suffering from SCID-X1 (ref. 48). First, HSCs were transduced using an integration-deficient lentivirus containing an HDR template encoding a therapeutic cDNA for *IL2RG*. Following transduction, cells were electroporated with mRNA encoding ZFNs targeting a mutational hotspot in *IL2RG* to stimulate HDR-based gene correction. To increase HDR rates, culture conditions were optimized with small molecules to encourage HSC division. This strategy resulted in gene-corrected HSCs from the SCID-X1 patient being obtained in culture at therapeutically relevant rates. HSCs from unaffected individuals that underwent the same gene correction procedure could sustain long-term hematopoiesis in mice. HSCs are capable of giving rise to all hematopoietic cell types and can be autologously transplanted, making them an extremely valuable cell population for all hematopoietic genetic disorders[98]. Gene-corrected HSCs could, in principle, be used to treat a wide range of genetic blood disorders, making this study an exciting breakthrough for therapeutic genome editing.

***In vivo* genome editing therapy.** *In vivo* genome editing therapy faces similar challenges to *ex vivo* strategies and is also limited by the small number of efficient delivery systems. Inefficient modification of target loci will be compounded by any inefficiencies in delivery, making tissues lacking robust delivery platforms particularly difficult to treat with this mode of therapy. For organ systems where delivery is efficient, however, there have already been a number of exciting preclinical therapeutic successes.

The first example of successful *in vivo* editing therapy was demonstrated in a mouse model of hemophilia B[49]. Restoring factor IX activity to above 1% of normal levels in severely affected individuals can transform the disease into a milder form, as infusion of recombinant factor IX into such individuals prophylactically from a young age to achieve such levels largely ameliorates the most severe bleeding complications[76]. In addition, factor IX is synthesized and secreted by the liver, an organ that can be transduced efficiently by viral vectors encoding editing systems.

Using hepatotropic adeno-associated viral (AAV) serotypes encoding ZFNs and a corrective HDR template, up to 7% of mutated, humanized factor IX alleles could be genetically corrected in murine liver tissue[49]. This resulted in improvement of clot formation kinetics, a measure of the function of the clotting cascade, demonstrating for the first time that *in vivo* editing therapy is not only feasible but also efficacious in treating this condition.

Building on this study, other groups have recently used *in vivo* genome editing of the liver with Cas9 to successfully treat a mouse model of hereditary tyrosinemia and to create mutations that provide protection against cardiovascular disease[50,99]. These two distinct applications demonstrate the versatility of this approach for treating disorders that involve hepatic dysfunction. Application of *in vivo* editing to other organ systems will be necessary to prove that this strategy is widely applicable. Currently, efforts to optimize both viral and nonviral vectors are underway, with the goal of expanding the range of disorders that can be treated with this mode of therapy[90,94]. Although most preclinical studies

have focused on the treatment of monogenic diseases, novel strategies using genome editing to treat polygenic diseases and as an antiviral have recently begun to show promise (**Box 1**).

Challenges to clinical translation

Translating genome editing technologies to the clinic involves major challenges, primarily in terms of the safety and efficacy of these treatments. Owing to the distinctly different molecular nature of these therapies compared to small-molecule and biologic therapies, engineering developments in several areas will be needed for these tools to be brought to bear on clinical medicine.

Increasing efficiency of gene correction. Although the amount of genome modification in a target cell population required to create a therapeutic effect differs depending on the disease, the efficacy of most editing treatments will be improved with increased editing rates. As previously noted, editing rates are controlled by the activity of DSB repair pathways.

Since NHEJ-mediated DSB repair is active in most cell types and is relatively efficient, the primary challenge to date has been to increase the efficiency of HDR. So far, applications of HDR in genome editing have been limited primarily to dividing cells because of the selective expression of HDR machinery during cell division and its downregulation in slowly cycling or post-mitotic cells. Cell cycle regulation can now be somewhat bypassed for slowly cycling cell types through stimulation of mitosis with pharmacologic agents *ex vivo*[48]. However, truly post-mitotic cells are unlikely to be amenable to such manipulation, limiting the applicability of this strategy. Nevertheless, further work to enable precise gene correction in post-mitotic cells such as neurons is critical to developing therapeutic strategies for the numerous neurological disorders that are currently untreatable. The solution to improved HDR in neurons will likely surface as we improve our understanding of DNA damage repair mechanisms in the brain and are able to harness heterologous systems. For example, the neurotrophic herpes simplex virus (HSV), which depends on single-strand annealing (SSA)—a form of HDR—to replicate, expresses viral proteins to facilitate SSA, might provide answers to achieving efficient gene correction in post-mitotic cells[100].

Additional, non-HDR-based strategies could also facilitate precise gene correction in post-mitotic cells. For example, attempts have been made to completely circumvent the need for HDR through NHEJ-based ligation of DNA templates containing therapeutic transgenes into targeted DSBs. Such ligation events have been successfully demonstrated using ZFNs[101], but with Cas9 the ligation rates are low[102,103]. This difference might be due to differences in cleavage patterns between ZFNs and Cas9: ZFNs generate a predictable 4-bp overhang, whereas Cas9 generates a blunt cut. Future structure-guided engineering may be able to alter the cleavage pattern of Cas9 to generate sticky ends.

Understanding and improving specificity of editing nucleases. The specificity of genome editing tools is one of the main safety concerns for clinical application. Genetic modifications are permanent, and deleterious off-target mutations could create cells with oncogenic potential, reduced fitness or functional impairment. Furthermore, oncogenic mutations resulting from off-target editing may lead to expansion of edited cells, and thus even low levels of off-target mutagenesis may have devastating consequences.

Two issues remain outstanding: evaluating and reducing off-target effects. A number of studies have attempted to evaluate the targeting specificities of ZFN, TALEN and Cas9 nucleases. The limited number of studies characterizing ZFN[104,105] and TALEN[106] specificity have only highlighted the challenges of detecting ZFN and TALEN off-target activity. Of note, the two independent studies attempting to characterize the off-target profile of the same pair of CCR5-targeting ZFNs have returned distinct and non-overlapping lists of off-target sites, which highlights the challenges associated with analysis of nuclease specificity.

Many studies have attempted to evaluate the specificity of Cas9, partly because the simplicity of the RNA-guided DNA targeting mechanism of Cas9 makes it considerably easier to establish hypotheses regarding possible off-targeting mechanisms based on Watson-Crick base-pairing rules relative to the protein-DNA interactions that mediate ZFN and TALEN targeting. While initial bacterial[40], biochemical[41,42] and mammalian[43] experiments have suggested that the 8–12-bp 3′ seed region of the guide sequence can be sensitive to single base mismatches, further work has shown that this rule of thumb is not necessarily accurate, especially in situations where there are high concentrations of Cas9 and guide RNA[88,107–110]. Many of these studies were carried out in cell lines and examined Cas9-mediated mutagenesis at genomic sites with high similarity to the target sequence, and they found that subsets of off-target sites with high sequence similarity to the target were statistically significantly mutated by the nuclease. However, the scope of possible off-target sites evaluated by these studies was limited to computationally predicted sites.

More recently, whole-genome sequencing of Cas9-edited cell lines revealed a low incidence of off-target mutation, which suggests that Cas9-mediated genome editing may be specific[111–113]. Despite these studies, unbiased assessment of genome-wide off-target editing using more advanced methods such as direct capture of DSBs[114], labeling of DSBs with oligo captures[115], and techniques that can detect larger structural variations (such as translocations) potentially imposed by nuclease treatment[116] will help us further understand the true risk of mutagenesis imposed by programmable nucleases. It is worth noting that off-target effects may be cell type specific: for example, off-target effects in transformed cell lines with dysregulated DSB repair pathways may provide an overestimate for the off-target effects that would be observed in healthy primary cells.

In order to reduce the frequency of off-target effects, many groups are rapidly improving the targeting specificity of Cas9. For example, transformation of Cas9 into a single-strand DNA nickase that primarily generates DSBs by creating two separate single-strand breaks on opposite DNA strands, via the expression of two separate guide RNAs, reduces off-target indel formation at computationally predicted off-target sites[102,109]. Additionally, truncation of the guide RNA, or the use of an RNA-guided FokI nuclease based on fusion between catalytically inactive Cas9 and the FokI nuclease domain, can also improve targeting specificity[117–119]. It is worth noting that the specificity requirements for each editing therapy will also depend on the total number of cells that are being exposed to the nuclease. For example, a nuclease with an off-target rate of 1 out of 1 million cells will have a significantly lower off-targeting risk when applied, under identical conditions, to 10,000 cells than to 1,000,000 cells.

Alternative genome editing strategies not involving nucleases have also been explored and may pose a lower mutagenic risk[62]. AAV genomes containing transgenes flanked by homology to target loci are capable of stimulating HDR in the absence of a nuclease, albeit at lower rates[49,86,120–122]. Using this strategy, one group targeted a factor IX cDNA to the highly expressed albumin locus and thereby corrected the bleeding diathesis phenotype in factor IX–deficient mice[62]. By targeting a highly expressed locus, the authors were able to achieve 7–20% of wild-type factor IX protein levels, despite an HDR rate of only 0.5%. Although this strategy may not be widely applicable owing to the low absolute targeting rate, this and future improved nuclease strategies should also be considered for therapeutic applications.

Delivery. Another major challenge for clinical translation is the delivery of editing systems to target cell types. A variety of nucleic acid or protein delivery methods may be used to introduce genome editing nucleases into target cells *ex vivo* or *in vivo* (**Fig. 3**). Depending on the choice of delivery method, the nucleases may be either transiently or permanently expressed in the target cell. Given that nucleases may exhibit off-target cleavage activity or trigger immune responses, the delivery system should be carefully selected.

For *ex vivo* applications, such as editing of hematopoietic stem cells, electroporation may be used to achieve transient nuclease expression through delivery of DNA-based nuclease expression vectors, mRNA or protein. Both integration-competent and integration-deficient lentiviral vectors have also been successfully used to drive nuclease expression. However, integrating lentiviral vectors may be less desirable because they drive constitutive expression and may result in more off-target activity. In addition, all three nuclease platform are amenable to modifications allowing proteins to be directly delivered into cells either through engineered cell-penetrating peptides or chemical conjugation[106,123,124].

For *in vivo* applications, the most promising delivery systems are viral vectors, particularly AAV vectors, which have recently been approved for clinical use[125]. AAVs come in many serotypes and have high delivery efficacy for a variety of tissue types including the eye, brain, liver and muscle[126]. However, the relatively small packaging capacity of AAV vectors poses some challenges for nuclease delivery. Whereas ZFNs are relatively small, and a dimeric ZFN pair can be packaged into a single AAV, a dimeric TALEN pair is much larger and will likely need to be packaged into two separate AAV vectors. For Cas9, short orthologs may be packaged along with guide RNAs into a single AAV. So far, AAV-mediated nuclease expression has been successful in several tissue types, including liver and brain[49,127]. In the case of viral-mediated Cas9 delivery, which may result in constitutive expression of nuclease proteins and cause genome instability and toxicity, self-cleaving mechanisms may be used to inactivate the nuclease transgene on the delivery vector[128].

Notwithstanding the potential of AAV-mediated *in vivo* nuclease expression, AAV-mediated nuclease expression also poses several challenges that will require further work. First, AAV-mediated nuclease expression is often constitutive, whereas it would be desirable to be able to shut down nuclease expression after a successful genome editing event has occurred in the target cell. Second, people who have already been naturally exposed to AAV are likely to have developed immunity against specific serotypes, so that AAV may not be an appropriate delivery vehicle for such patients.

To overcome these challenges, nanoparticle- and lipid-based *in vivo* mRNA or protein delivery systems may provide attractive alternatives to viral vectors[123,129]. Delivery of nuclease mRNA, via nanoparticle conjugation, or of nuclease proteins will permit more precise dosage control, which has been shown to affect the level of off-target mutation rate[88,124]. mRNA or protein delivery will also be transient, thereby minimizing nuclease-induced toxicity. Finally, for delivery of nuclease proteins, especially the microbially derived TALENs and Cas9, exposure of the proteins may stimulate immune reactions. Potential strategies for circumventing immunotoxicity resulting from protein delivery may include limiting dosage and humanizing the proteins to reduce their immunogenicity[130].

Conclusion

The enormous excitement surrounding genome editing needs to be coupled with strategic planning and rigorous but enabling regulatory processes to ensure the successful development of this class of potentially life-changing medicines. The technology will require a number of iterations to systematically optimize its efficacy, safety and specificity. Although still in its infancy, genome editing presents tantalizing opportunities for tackling a number of diseases that are beyond the reach of previous therapies. Given the accelerating pace of technological advances and broad range of basic science and clinical applications, the road ahead will undoubtedly be an exciting one.

ACKNOWLEDGMENTS

The authors would like to thank J. Gootenberg, O. Abudayyeh, F. Ran and C. Men for critical reading of the manuscript, and all members of the Zhang lab for helpful discussions. D.B.T.C. is supported by award number T32GM007753 from the National Institute of General Medical Sciences. R.J.P. is supported by a National Science Foundation (NSF) Graduate Research Fellowship under grant number 1122374. F.Z. is supported by the National Institute of Mental Health through a US National Institutes of Health (NIH) Director's Pioneer Award (DP1-MH100706); the National Institute of Neurological Disorders and Stroke through an NIH Transformative R01 grant (R01-NS 07312401); an NSF Waterman Award; and the Keck, Damon Runyon, Searle Scholars, Klingenstein, Vallee, Merkin, and Simons Foundations. F.Z. is also supported by Bob Metcalfe. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the NIH.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

1. Lander, E.S. Initial impact of the sequencing of the human genome. *Nature* **470**, 187–197 (2011).
2. Thoene, J.G. *Small Molecule Therapy for Genetic Disease* (Cambridge University Press, 2010).
3. Kay, M.A. State-of-the-art gene-based therapies: the road ahead. *Nat. Rev. Genet.* **12**, 316–328 (2011).
4. Vaishnaw, A.K. *et al.* A status report on RNAi therapeutics. *Silence* **1**, 14 (2010).
5. Gaspar, H.B. *et al.* Long-term persistence of a polyclonal T cell repertoire after gene therapy for X-linked severe combined immunodeficiency. *Sci. Transl. Med.* **3**, 97ra79 (2011).
6. Howe, S.J. *et al.* Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *J. Clin. Invest.* **118**, 3143–3150 (2008).
7. Aiuti, A. *et al.* Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. *Science* **341**, 1233151 (2013).
8. Castanotto, D. & Rossi, J.J. The promises and pitfalls of RNA-interference-based therapeutics. *Nature* **457**, 426–433 (2009).
9. Tiemann, K. & Rossi, J.J. RNAi-based therapeutics—current status, challenges and prospects. *EMBO Mol. Med.* **1**, 142–151 (2009).
10. Jackson, A.L. & Linsley, P.S. Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application. *Nat. Rev. Drug Discov.* **9**, 57–67 (2010).
11. Stoddard, B.L. Homing endonucleases: from microbial genetic invaders to reagents for targeted DNA modification. *Structure* **19**, 7–15 (2011).
12. Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S. & Gregory, P.D. Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* **11**, 636–646 (2010).
13. Bogdanove, A.J. & Voytas, D.F. TAL effectors: customizable proteins for DNA targeting. *Science* **333**, 1843–1846 (2011).
14. Scharenberg, A.M., Duchateau, P. & Smith, J. Genome engineering with TAL-effector nucleases and alternative modular nuclease technologies. *Curr. Gene Ther.* **13**, 291–303 (2013).
15. Hsu, P.D., Lander, E.S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262–1278 (2014).
16. Rouet, P., Smih, F. & Jasin, M. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol. Cell. Biol.* **14**, 8096–8106 (1994).
17. Thierry, A. & Dujon, B. Nested chromosomal fragmentation in yeast using the meganuclease I-Sce I: a new method for physical mapping of eukaryotic genomes. *Nucleic Acids Res.* **20**, 5625–5631 (1992).
18. Thierry, A. *et al.* Cleavage of yeast and bacteriophage T7 genomes at a single site using the rare cutter endonuclease I-Sce I. *Nucleic Acids Res.* **19**, 189–190 (1991).
19. Smith, J. *et al.* A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences. *Nucleic Acids Res.* **34**, e149 (2006).
20. Boissel, S. *et al.* megaTALs: a rare-cleaving nuclease architecture for therapeutic genome engineering. *Nucleic Acids Res.* **42**, 2591–2601 (2014).
21. Kim, Y.G., Cha, J. & Chandrasegaran, S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. USA* **93**, 1156–1160 (1996).
22. Wolfe, S.A., Nekludova, L. & Pabo, C.O. DNA recognition by Cys_2His_2 zinc finger proteins. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 183–212 (2000).
23. Bibikova, M., Beumer, K., Trautman, J.K. & Carroll, D. Enhancing gene targeting with designed zinc finger nucleases. *Science* **300**, 764 (2003).

24. Bibikova, M., Golic, M., Golic, K.G. & Carroll, D. Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. *Genetics* **161**, 1169–1175 (2002).
25. Miller, J., McLachlan, A.D. & Klug, A. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *EMBO J.* **4**, 1609–1614 (1985).
26. Miller, J.C. *et al.* An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat. Biotechnol.* **25**, 778–785 (2007).
27. Urnov, F.D. *et al.* Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* **435**, 646–651 (2005).
28. Porteus, M.H. & Baltimore, D. Chimeric nucleases stimulate gene targeting in human cells. *Science* **300**, 763 (2003).
29. Smith, J. *et al.* Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. *Nucleic Acids Res.* **28**, 3361–3369 (2000).
30. Boch, J. *et al.* Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509–1512 (2009).
31. Moscou, M.J. & Bogdanove, A.J. A simple cipher governs DNA recognition by TAL effectors. *Science* **326**, 1501 (2009).
32. Christian, M. *et al.* Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* **186**, 757–761 (2010).
33. Miller, J.C. *et al.* A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.* **29**, 143–148 (2011).
34. Mahfouz, M.M. *et al.* De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. *Proc. Natl. Acad. Sci. USA* **108**, 2623–2628 (2011).
35. Li, T. *et al.* TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Res.* **39**, 359–372 (2011).
36. Bolotin, A., Quinquis, B., Sorokin, A. & Ehrlich, S.D. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151**, 2551–2561 (2005).
37. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712 (2007).
38. Garneau, J.E. *et al.* The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**, 67–71 (2010).
39. Deltcheva, E. *et al.* CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471**, 602–607 (2011).
40. Sapranauskas, R. *et al.* The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res.* **39**, 9275–9282 (2011).
41. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
42. Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. USA* **109**, E2579–E2586 (2012).
43. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
44. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).
45. Isalan, M. Zinc-finger nucleases: how to play two good hands. *Nat. Methods* **9**, 32–34 (2012).
46. Sun, N. & Zhao, H. Transcription activator-like effector nucleases (TALENs): a highly efficient and versatile tool for genome editing. *Biotechnol. Bioeng.* **110**, 1811–1821 (2013).
47. Tebas, P. *et al.* Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *N. Engl. J. Med.* **370**, 901–910 (2014).
48. Genovese, P. *et al.* Targeted genome editing in human repopulating haematopoietic stem cells. *Nature* **510**, 235–240 (2014).
49. Li, H. *et al. In vivo* genome editing restores haemostasis in a mouse model of haemophilia. *Nature* **475**, 217–221 (2011).
50. Yin, H. *et al.* Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nat. Biotechnol.* **32**, 551–553 (2014).
51. Hentze, M.W. & Kulozik, A.E. A perfect message: RNA surveillance and nonsense-mediated decay. *Cell* **96**, 307–310 (1999).
52. Choulika, A., Perrin, A., Dujon, B. & Nicolas, J.F. Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **15**, 1968–1973 (1995).
53. Bibikova, M. *et al.* Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Mol. Cell. Biol.* **21**, 289–297 (2001).
54. Krejci, L., Altmannova, V., Spirek, M. & Zhao, X. Homologous recombination and its regulation. *Nucleic Acids Res.* **40**, 5795–5818 (2012).
55. Plessis, A., Perrin, A., Haber, J.E. & Dujon, B. Site-specific recombination determined by I-SceI, a mitochondrial group I intron-encoded endonuclease expressed in the yeast nucleus. *Genetics* **130**, 451–460 (1992).
56. Rudin, N., Sugarman, E. & Haber, J.E. Genetic and physical analysis of double-strand break repair and recombination in *Saccharomyces cerevisiae*. *Genetics* **122**, 519–534 (1989).
57. Liu, R. *et al.* Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* **86**, 367–377 (1996).
58. Cohen, J. *et al.* Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. *Nat. Genet.* **37**, 161–165 (2005).
59. Jonsson, T. *et al.* A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. *Nature* **488**, 96–99 (2012).
60. Lombardo, A. *et al.* Site-specific integration and tailoring of cassette design for sustainable gene transfer. *Nat. Methods* **8**, 861–869 (2011).
61. Moehle, E.A. *et al.* Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. *Proc. Natl. Acad. Sci. USA* **104**, 3055–3060 (2007).
62. Barzel, A. *et al.* Promoterless gene targeting without nucleases ameliorates haemophilia B in mice. *Nature* doi:10.1038/nature13864 (2014).
63. Maude, S.L. *et al.* Chimeric antigen receptor T cells for sustained remissions in leukemia. *N. Engl. J. Med.* **371**, 1507–1517 (2014).
64. Hu, W. *et al.* RNA-directed gene editing specifically eradicates latent and prevents new HIV-1 infection. *Proc. Natl. Acad. Sci. USA* **111**, 11461–11466 (2014).
65. Lin, S.R. *et al.* The CRISPR/Cas9 system facilitates clearance of the intrahepatic HBV templates in vivo. *Mol. Ther. Nucleic Acids* **3**, e186 (2014).
66. Bloom, K., Ely, A., Mussolino, C., Cathomen, T. & Arbuthnot, P. Inactivation of hepatitis B virus replication in cultured cells and in vivo with engineered transcription activator-like effector nucleases. *Mol. Ther.* **21**, 1889–1897 (2013).
67. Cradick, T.J., Keck, K., Bradshaw, S., Jamieson, A.C. & McCaffrey, A.P. Zinc-finger nucleases as a novel therapeutic strategy for targeting hepatitis B virus DNAs. *Mol. Ther.* **18**, 947–954 (2010).
68. Weber, N.D. *et al.* AAV-mediated delivery of zinc finger nucleases targeting hepatitis B virus inhibits active replication. *PLoS ONE* **9**, e97579 (2014).
69. Holt, N. *et al.* Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 *in vivo*. *Nat. Biotechnol.* **28**, 839–847 (2010).
70. Perez, E.E. *et al.* Establishment of HIV-1 resistance in $CD4^+$ T cells by genome editing using zinc-finger nucleases. *Nat. Biotechnol.* **26**, 808–816 (2008).
71. Hacein-Bey-Abina, S. *et al.* Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy. *N. Engl. J. Med.* **346**, 1185–1193 (2002).
72. Gaspar, H.B. *et al.* Gene therapy of X-linked severe combined immunodeficiency by use of a pseudotyped gammaretroviral vector. *Lancet* **364**, 2181–2187 (2004).
73. Bousso, P. *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. USA* **97**, 274–278 (2000).
74. Malech, H.L. *et al.* Prolonged production of NADPH oxidase-corrected granulocytes after gene therapy of chronic granulomatous disease. *Proc. Natl. Acad. Sci. USA* **94**, 12133–12138 (1997).
75. Kang, H.J. *et al.* Retroviral gene therapy for X-linked chronic granulomatous disease: results from phase I/II trial. *Mol. Ther.* **19**, 2092–2101 (2011).
76. Löfqvist, T., Nilsson, I.M., Berntorp, E. & Pettersson, H. Haemophilia prophylaxis in young patients—a long-term follow-up. *J. Intern. Med.* **241**, 395–400 (1997).
77. Kaushansky, K. & Williams, W.J. *Williams Hematology* (McGraw-Hill Medical, New York, 2010).
78. Rothkamm, K., Kruger, I., Thompson, L.H. & Lobrich, M. Pathways of DNA double-strand break repair during the mammalian cell cycle. *Mol. Cell. Biol.* **23**, 5706–5715 (2003).
79. Sharma, S. Age-related nonhomologous end joining activity in rat neurons. *Brain Res. Bull.* **73**, 48–54 (2007).
80. Ciccia, A. & Elledge, S.J. The DNA damage response: making it safe to play with knives. *Mol. Cell* **40**, 179–204 (2010).
81. Chapman, J.R., Taylor, M.R. & Boulton, S.J. Playing the end game: DNA double-strand break repair pathway choice. *Mol. Cell* **47**, 497–510 (2012).
82. Beumer, K.J., Trautman, J.K., Mukherjee, K. & Carroll, D. Donor DNA utilization during gene targeting with zinc-finger nucleases. *G3* http://dx.doi.org/10.1534/g3.112.005439 (2013).
83. Moynahan, M.E. & Jasin, M. Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis. *Nat. Rev. Mol. Cell Biol.* **11**, 196–207 (2010).
84. Orlando, S.J. *et al.* Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. *Nucleic Acids Res.* **38**, e152 (2010).
85. Chen, F. *et al.* High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nat. Methods* **8**, 753–755 (2011).
86. Miller, D.G. *et al.* Gene targeting *in vivo* by adeno-associated virus vectors. *Nat. Biotechnol.* **24**, 1022–1026 (2006).
87. Beumer, K.J. *et al.* Efficient gene targeting in *Drosophila* by direct embryo injection with zinc-finger nucleases. *Proc. Natl. Acad. Sci. USA* **105**, 19821–19826 (2008).
88. Hsu, P.D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–832 (2013).
89. Bunn, H.F. & Aster, J. *Pathophysiology of Blood Disorders* (McGraw-Hill, New York, 2011).
90. Kotterman, M.A. & Schaffer, D.V. Engineering adeno-associated viruses for clinical gene therapy. *Nat. Rev. Genet.* **15**, 445–451 (2014).
91. Nguyen, T.H. & Ferry, N. Liver gene therapy: advances and hurdles. *Gene Ther.* **11** (suppl. 1), S76–S84 (2004).
92. Boye, S.E., Boye, S.L., Lewin, A.S. & Hauswirth, W.W. A comprehensive review of retinal gene therapy. *Mol. Ther.* **21**, 509–519 (2013).
93. Bessis, N., GarciaCozar, F.J. & Boissier, M.C. Immune responses to gene therapy vectors: influence on vector function and effector mechanisms. *Gene Ther.* **11** (suppl. 1), S10–S17 (2004).
94. Yin, H. *et al.* Non-viral vectors for gene-based therapy. *Nat. Rev. Genet.* **15**, 541–555 (2014).
95. Mingozzi, F. & High, K.A. Immune responses to AAV vectors: overcoming barriers to successful gene therapy. *Blood* **122**, 23–36 (2013).
96. Lamers, C.H. *et al.* Immune responses to transgene and retroviral vector in patients treated with ex vivo-engineered T cells. *Blood* **117**, 72–82 (2011).

97. Hütter, G. *et al.* Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *N. Engl. J. Med.* **360**, 692–698 (2009).
98. Weissman, I.L. & Shizuru, J.A. The origins of the identification and isolation of hematopoietic stem cells, and their capability to induce donor-specific transplantation tolerance and treat autoimmune diseases. *Blood* **112**, 3543–3553 (2008).
99. Ding, Q. *et al.* Permanent alteration of PCSK9 with in vivo CRISPR-Cas9 genome editing. *Circ. Res.* **115**, 488–492 (2014).
100. Weller, S.K. & Sawitzke, J.A. Recombination promoted by DNA viruses: phage lambda to herpes simplex virus. *Annu. Rev. Microbiol.* **68**, 237–258 (2014).
101. Maresca, M., Lin, V.G., Guo, N. & Yang, Y. Obligate ligation-gated recombination (ObLiGaRe): custom-designed nuclease-mediated targeted integration through non-homologous end joining. *Genome Res.* **23**, 539–546 (2013).
102. Ran, F.A. *et al.* Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**, 1380–1389 (2013).
103. Nakade, S. *et al.* Microhomology-mediated end-joining-dependent integration of donor DNA in cells and animals using TALENs and CRISPR/Cas9. *Nat. Commun.* **5**, 5560 (2014).
104. Pattanayak, V., Ramirez, C.L., Joung, J.K. & Liu, D.R. Revealing off-target cleavage specificities of zinc-finger nucleases by *in vitro* selection. *Nat. Methods* **8**, 765–770 (2011).
105. Gabriel, R. *et al.* An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat. Biotechnol.* **29**, 816–823 (2011).
106. Guilinger, J.P. *et al.* Broad specificity profiling of TALENs results in engineered nucleases with improved DNA-cleavage specificity. *Nat. Methods* **11**, 429–435 (2014).
107. Fu, Y. *et al.* High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* **31**, 822–826 (2013).
108. Cho, S.W. *et al.* Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res.* **24**, 132–141 (2014).
109. Mali, P. *et al.* Barcoding cells using cell-surface programmable DNA-binding domains. *Nat. Methods* **10**, 403–406 (2013).
110. Pattanayak, V. *et al.* High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat. Biotechnol.* **31**, 839–843 (2013).
111. Veres, A. *et al.* 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 (2014).
112. Tan, E.P., Li, Y., Del Castillo Velasco-Herrera, M., Yusa, K. & Bradley, A. Off-target assessment of CRISPR-Cas9 guiding RNAs in human iPS and mouse ES cells. *Genesis* doi:10.1002/dvg.22835 (2014).
113. Yang, L. *et al.* Targeted and genome-wide sequencing reveal single nucleotide variations impacting specificity of Cas9 in human stem cells. *Nat. Commun.* **5**, 5507 (2014).
114. Crosetto, N. *et al.* Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. *Nat. Methods* **10**, 361–365 (2013).
115. Tsai, S.Q. *et al.* GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* doi:10.1038/nbt.3117 (2014).
116. Frock, R.L. *et al.* Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat. Biotechnol.* doi:10.1038/nbt.3101 (2014).
117. Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M. & Joung, J.K. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* **32**, 279–284 (2014).
118. Guilinger, J.P., Thompson, D.B. & Liu, D.R. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat. Biotechnol.* **32**, 577–582 (2014).
119. Tsai, S.Q. *et al.* Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat. Biotechnol.* **32**, 569–576 (2014).
120. Lisowski, L. *et al.* Ribosomal DNA integrating rAAV-rDNA vectors allow for stable transgene expression. *Mol. Ther.* **20**, 1912–1923 (2012).
121. Shi, Y., Falahati, R., Zhang, J., Flebbe-Rehwaldt, L. & Gaensler, K.M. Role of antigen-specific regulatory $CD4^+CD25^+$ T cells in tolerance induction after neonatal IP administration of AAV-hF.IX. *Gene Ther.* **20**, 987–996 (2013).
122. Paulk, N.K., Loza, L.M., Finegold, M.J. & Grompe, M. AAV-mediated gene targeting is significantly enhanced by transient inhibition of nonhomologous end joining or the proteasome in vivo. *Hum. Gene Ther.* **23**, 658–665 (2012).
123. Zuris, J.A. *et al.* Cationic lipid–mediated delivery of proteins enables efficient protein-based genome editing *in vitro* and *in vivo*. *Nat. Biotechnol.* doi:10.1038/nbt.3081 (2014).
124. Gaj, T., Guo, J., Kato, Y., Sirk, S.J. & Barbas, C.F. III. Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. *Nat. Methods* **9**, 805–807 (2012).
125. Wirth, T., Parker, N. & Yla-Herttuala, S. History of gene therapy. *Gene* **525**, 162–169 (2013).
126. Samulski, R.J. & Muzyczka, N. AAV-mediated gene therapy for research and therapeutic purposes. *Annu. Rev. Virol.* **1**, 427–451 (2014).
127. Swiech, L. *et al. In vivo* interrogation of gene function in the mammalian brain using CRISPR-Cas9. *Nat. Biotechnol.* doi:10.1038/nbt.3055 (2014).
128. Moore, R. *et al.* CRISPR-based self-cleaving mechanism for controllable gene delivery in human cells. *Nucleic Acids Res.* doi:10.1093/nar/gku1326 (2014).
129. Kormann, M.S. *et al.* Expression of therapeutic proteins after delivery of chemically modified mRNA in mice. *Nat. Biotechnol.* **29**, 154–157 (2011).
130. Riechmann, L., Clark, M., Waldmann, H. & Winter, G. Reshaping human antibodies for therapy. *Nature* **332**, 323–327 (1988).
131. Ye, L. *et al.* Seamless modification of wild-type induced pluripotent stem cells to the natural CCR5Delta32 mutation confers resistance to HIV infection. *Proc. Natl. Acad. Sci. USA* **111**, 9591–9596 (2014).
132. Ousterout, D.G. *et al.* Reading frame correction by targeted genome editing restores dystrophin expression in cells from Duchenne muscular dystrophy patients. *Mol. Ther.* **21**, 1718–1726 (2013).
133. Long, C. *et al.* Prevention of muscular dystrophy in mice by CRISPR/Cas9-mediated editing of germline DNA. *Science* **345**, 1184–1188 (2014).
134. Wu, Y. *et al.* Correction of a genetic disease in mouse via use of CRISPR-Cas9. *Cell Stem Cell* **13**, 659–662 (2013).
135. Schwank, G. *et al.* Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell* **13**, 653–658 (2013).
136. Crosby, T. *et al.* Loss-of-function mutations in APOC3, triglycerides, and coronary disease. *N. Engl. J. Med.* **371**, 22–31 (2014).
137. Bauer, D.E. *et al.* An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. *Science* **342**, 253–257 (2013).
138. Menzel, S. *et al.* A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. *Nat. Genet.* **39**, 1197–1199 (2007).
139. Xu, J. *et al.* Correction of sickle cell disease in adult mice by interference with fetal hemoglobin silencing. *Science* **334**, 993–996 (2011).
140. Pennisi, E. The Biology of Genomes. Disease risk links to gene regulation. *Science* **332**, 1031 (2011).
141. Kennedy, E.M. *et al.* Inactivation of the human papillomavirus E6 or E7 gene in cervical carcinoma cells using a bacterial CRISPR/Cas RNA-guided endonuclease. *J. Virol.* doi:10.1128/JVI.01879-14 (2014).