Current progress in therapeutic gene editing for monogenic diseases

Versha Prakash[1], Marc Moore[1] and Rafael J Yáñez-Muñoz[1]

[1]School of Biological Sciences, Royal Holloway, University of London, Egham TW20 0EX, UK

Correspondence should be addressed to R.J.Y.-M. (Email: rafael.yanez@royalholloway.ac.uk)

Work was performed in Egham, Surrey, United Kingdom

Corresponding author address:

Rafael J. Yáñez-Muñoz

School of Biological Sciences, Royal Holloway, University of London, Egham, Surrey, TW20 0EX, United Kingdom

Telephone: +44 (0)1784-443180

Fax: +44 (0)1784-414224

E-mail: rafael.yanez@royalholloway.ac.uk

Short title: Gene editing for inherited disease

Programmable nucleases allow defined alterations in the genome with ease-of-use, efficiency and specificity. Their availability has led to accurate and widespread genome engineering, with multiple applications in basic research, biotechnology and therapy. With regards to human gene therapy, nuclease-based gene editing has facilitated development of a broad range of therapeutic strategies based on both non-homologous end-joining and homology-dependent repair. This review discusses current progress in nuclease-based therapeutic applications for a subset of inherited monogenic diseases including cystic fibrosis, Duchenne muscular dystrophy, diseases of the bone marrow and haemophilia, and highlights associated challenges and future prospects.

Keywords: gene editing, inherited disease, engineered nuclease

INTRODUCTION

There are five-eight thousand monogenic diseases, defined as inherited conditions arising from mutations on a single gene[1]. These often manifest during childhood and lead to morbidity and sometimes premature death. While each monogenic disease is rare, it has been estimated that together they will affect about 6% of people at some point in their lives[1]. Diagnosis and treatment for these diseases remain largely insufficient and the care is primarily palliative, focusing on disease management without addressing the underlying genetic defects. The realisation of the social and economic importance of rare diseases and the acute need for diagnostics and treatments has led to initiatives like the International Rare Disease Research Consortium (IRDiRC, http://www.irdirc.org), the Undiagnosed Diseases Network (UDN, http://www.genome.gov/27562471) and Syndromes Without a Name UK (SWAN UK, http://www.geneticalliance.org.uk/projects/swan.htm).

Gene therapy, which encompasses a range of strategies, aimed from the outset to treat inherited disorders, assuming that monogenic diseases would be the easiest to target. Classical gene therapy approaches have centred on the delivery of DNA to augment endogenous gene expression. Predominantly, these approaches rely on the transfer of functional genes using a variety of viral vectors, due to their intrinsic ability to effectively transduce human cells. Retroviral vectors provided the first clear demonstrations of therapeutic benefit in primary immunodeficiencies, and they also highlighted the risk of adverse events attributable to insertional mutagenesis due to genomic integration of proviruses[2]. Among several, other success stories include Glybera[3], the first clinically approved gene therapy in the EU, which uses

an adeno-associated virus (AAV) vector drug for lipoprotein lipase deficiency; and in the case of cystic fibrosis, repeated nebulisation of liposomes encoding the cystic fibrosis transmembrane conductance regulator (CFTR) gene has shown some therapeutic benefit[4]. Thus gene augmentation shows great therapeutic promise and has set the stage for the gene-editing approaches reviewed here.

Gene editing is a gene therapy approach that relies on designer nucleases to recognise and cut specific DNA sequences, and subsequently exploits innate cellular DNA repair pathways, namely non-homologous end joining (NHEJ) and homology directed repair (HDR), to introduce targeted modifications in the genome (**Fig. 1a**). Four nuclease families have been used in this context: meganucleases (MGNs), zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and clustered regulatory interspaced short palindromic repeats associated RNA-guided Cas9 (CRISPR-Cas9) nucleases[5]. These can be designed to precisely introduce a double stranded break (DSB) at the target locus of interest. The DSB will then be repaired by either NHEJ or HDR. NHEJ involves direct ligation of DNA ends in a highly efficient but error-prone manner, which causes small insertions and/or deletions (Indels) at the break site. In the context of a disease-causing locus, NHEJ can be exploited to excise or disrupt deleterious sequences, or even restore the reading frame of a gene (**Fig. 1b**). In contrast, HDR requires a donor DNA containing sequences homologous to those adjacent to the DSB. The donor DNA can be used to repair a mutation, or to knock-in a block of exons ("superexon") or a full cDNA at either the endogenous locus (reconstituting the wild-type sequence) or at a genomic 'safe harbour' (a region of DNA where transgenes can integrate and express in a predictable manner without insertional mutagenesis or perturbation of gene function)[6] (**Fig.1b**). Gene editing thus

opens up the possibility of permanently modifying a genomic sequence of interest by enabling targeted disruption, insertion, excision, and correction in both *ex vivo* and *in vivo* settings (**Fig. 1c**). Whilst these advances are expected to revolutionise the field at large, current gene editing approaches are limited by efficacy of modification, safety concerns related to the specificity of nucleases and delivery of gene editing tools to target cell types. This review aims to outline prominent gene editing research across a range of monogenic disorders (Table 1), and to highlight recent advancements and current challenges.

GENE EDITING IN MONOGENIC DISORDERS

CYSTIC FIBROSIS

Cystic Fibrosis (CF) is an autosomal recessive disease resulting from mutations in the *CFTR* gene, which encodes an epithelial anion channel. CFTR is distributed across a wide range of organs including pancreas, kidney, liver, lungs, gastrointestinal and reproductive tracts, making CF a multi-organ disease. Mutations in *CFTR* lead to sub-optimal ion transport and fluid retention, causing the prominent clinical manifestations of abnormal thickening of the mucus in lungs and pancreatic insufficiency[7]. In the lung, dysfunctional CFTR hinders mucociliary clearance, rendering the organ susceptible to bacterial infections and inflammation, ultimately leading to airway occlusion, respiratory failure and premature death[8]. CF remains the most common and lethal genetic disease among the Caucasian population with 70,000–100,000 sufferers estimated worldwide, highlighting a real need for the development of better treatments.

One major challenge to the development of a therapeutic strategy for CF is the huge diversity of mutation types. ΔF508 (deletion of phenylalanine at codon 508) mutation, with a prevalence of >80% in CF patients, is by far the most common but more than 1,990 *CFTR*-mutations have been described[9]. They cause premature stop codons, aberrant splicing, incorrect protein folding or trafficking to the cell surface, and dysfunctional CFTRs with limited channel opening capacity[9]. Pharmacological interventions have been targeted to several of these processes, in the form of: (i) read-through therapeutics, which recognise premature stop codons, thereby allowing full-length protein production and a decline in associated nonsense-mediated decay[10]; (ii) correctors, which enable slightly misfolded proteins to evade endoplasmic reticulum quality control and insert at the primary epithelia[11]; and (iii) potentiators, which target gating mutations and increase channel opening[8,12]. One prominent drug, the potentiator Ivacaftor (Kalydeco), has demonstrated significant improvement of numerous clinical endpoints such as forced expiratory volume (FEV), weight gain, reduction of hospital admissions (related to requirement of IV antibiotic administration) and increase in lung clearance index in diverse patient subsets bearing the G551D-CFTR missense mutation[13–16]. Ivacaftor also demonstrated clinical benefit in other gating mutation types: G1244E, G1349D, G178R, G551S, G970R, S1251N, S1255P, S549N and S549R, with G970R being one exception[17]. In the case of ΔF508, ivacaftor has been tested in a Phase 3 clinical trial in combination with lumacaftor[18]. The improvements seen in FEV were modest compared to ivacaftor monotherapy in G551D studies[13,15]. Thus, while drug administration is therapeutic in some gating mutation types, the commonly occurring ΔF508 still requires a more effective treatment. QR-010 is a drug based on modified single stranded RNA and designed to repair ΔF508 mRNA; research in human cell lines

and mice appears promising and is currently progressing to a phase 1b clinical trial[19,20]. However, this still would not address mutations resulting from aberrant splicing; it is in these instances gene editing could prove most beneficial.

The premise of permanent correction by gene-editing, as opposed to drug and recent non-viral gene therapy treatments (previously reviewed[9,21]) which require repeated administration, is promising. Such *CFTR* correction at the genome level has been trialled across many of the gene-editing platforms. ZFNs were first used to achieve HDR-mediated knock-in of a 4.5 kb genomic donor harbouring *CFTR* exons 10-24. The efficiency in patient-derived epithelial cells was modest, with DNA cleavage as measured by NHEJ estimated at 7.8% and the subsequent HDR occurring at <1%, respectively[22]. The ΔF508 mutation was targeted and functional repair obtained in stem cell organoids, using CRISPR-Cas9-mediated HDR and an exon 11 and puromycin resistance-containing DNA donor. In most instances, this resulted in a heterozygous phenotype or mono-allelic correction. Moreover, the forskolin-induced swelling assay, which demonstrates functional correction of the organoids via fluid secretion into the epithelia, showed more improvement than chemical correctors[23,24]. Although for stem cell organoids there are colon engraftment data, extension of these studies to lung would be helpful to assess the therapeutic promise of this approach[25]. Similar HDR strategies have demonstrated *CFTR* correction in induced pluripotent stem cells (iPSCs) derived from CF patients, bearing heterozygous ΔF508/ΔI507 mutations in the ZFN study[26]; and Δ508 homozygous mutations in the TALENs[27] as well as Cas9[28] studies. Genetically modified iPSCs can be selected and amplified at clonal level, allowing the production of pure populations of corrected cells, something unlikely to be achieved with primary cells. Two of these reports demonstrated that corrected

clones could be differentiated successfully into epithelial cells, whilst retaining *CFTR* expression[26,27]. Even with these advances, the multi-organ involvement of CF would limit an *ex vivo* approach as engraftment of corrected cells would only provide localised correction. In addition, there is a lack of consensus regarding which cells of the sinuses and lungs should be targeted in the amelioration of CF.

Given that drug therapy has demonstrated moderate repair, the focus may now be upon reaching a therapeutic threshold. Current threshold estimates are in the range of 10-24% of normal expression[9], similar to improvements in FEV following the use of ivacaftor[9], and perhaps achievable for other mutations by a combination of gene and drug therapies. The development of a catalytically inactive Cas9 in which the DNA cleavage domains harbour point mutations, tethered to a VP64 transcriptional activator, could be used to upregulate *CFTR* expression[29,30]. This transcriptional enhancement of the mutant gene, coupled to corrector and potentiator therapy, could be beneficial as the pool of protein available to be inserted within the epithelial surface would be increased.

DUCHENNE MUSCULAR DYSTROPHY

Duchenne muscular dystrophy (DMD) is an inherited X-linked disease resulting from mutations that disrupt the reading frame in the gene encoding dystrophin. This protein plays a crucial role in stabilisation of muscle sarcolemma and signalling. In the absence of dystrophin, progressive muscle wasting, with concomitant declines in respiratory and cardiac function occur. Over time,

this results in the loss of ambulation, necessitates the use of invasive ventilation, and ultimately leads to premature death[31,32].

Presently, there is no effective treatment for DMD. Current interventions including the use of prednisolone are inadequate, targeting only secondary characteristics of inflammation and muscle loss[33]. Despite being recognised as a prime candidate for gene therapy since the discovery of *DMD* in 1987, advancement has been hindered because *DMD* is the largest naturally occurring gene, spanning 2.5 Mb, with a cDNA of 11.2 Kb. A great deal of work has focused upon mutation-specific strategies using pharmacological and gene therapy approaches. Many of the current approaches, such as exon skipping, aim to restore the reading frame by targeting mRNA, masking splice sites or enhancers of exons that shift the reading frame[34,35]. The aim of these strategies is to produce a truncated but viable dystrophin protein, resulting in a clinically milder Becker rather than Duchenne muscular dystrophy phenotype[36].

Genome editing approaches for DMD include promoting permanent exon removal[37], and HDR-mediated cDNA knock-in[38,39]. Notable demonstrations in primary patient myoblasts have been the permanent excision of exon 51 (which would be applicable to 13% of patients) achieved using ZFNs, leading to restoration of the reading frame in an approach akin to exon skipping[40], and permanent removal of exons 45-55 by multiplexed Cas9 (**Fig. 1b**). The latter is a mutational hotspot which if targeted could be therapeutically applicable for 62% of patients[37]. Both of these approaches result in a truncated but functional dystrophin protein[41]. The mutational hotspot has also been targeted by HDR, with meganuclease-mediated repair of exons 45-52 in

immortalised patient cells[38]. The benefit of the HDR-mediated approach is that the subsequent correction would enable the restoration of full-length dystrophin (**Fig. 1b**).

DMD has also been restored by genome editing in iPSCs from a patient lacking exon 44 by three different strategies: exon 45 skipping by disruption of its splicing acceptor, Indel-mediated frameshift of exon 45, and exon 44 knock-in, utilising both TALEN and CRISPR-Cas9 approaches[39]. Targeted differentiation of such iPSCs could eventually progress to correction of individual mutation types via engraftment of corrected proliferative cells into muscles. Some concerns with this approach are the modulation of cell proliferation, efficiency of the engraftment process and localised intramuscular regeneration[42].

Model systems are also providing data of relevance to human therapy. Direct germline correction of a murine *Dmd* mutation has recently been demonstrated in the *mdx* mouse, in which the exon 23 nonsense mutation was corrected via Cas9 mRNA injection followed by implantation into pseudo pregnant females. This resulted in mosaicism within targeted animals, showing correction ranging from 2-100%; this type of work could establish the level of dystrophin expression required to provide therapeutic benefit, which is currently predicted to be between 15-20%[43]. Two genotypically distinct rat models have also been produced, one with a C>T nonsense mutation in exon 23 analogous to that of the *mdx* mouse, and another with a large deletion spanning exons 6-13[44,45]. Additionally, the use of CRISPR-Cas9 in pigs and non-human primates for the generation of DMD phenotypes, demonstrates that such gene-editing can be easily transitioned into larger animals[46–48]. This ease of generation of new DMD animal models allows for a range of mutation types to be produced, thereby providing a greater

diversity of models for translational research. Moreover, larger animal models such as rats, pigs and primates tend to exhibit muscle phenotypes such as fibrosis, which is more representative of the clinical manifestations seen in patients than those in the commonly used *mdx* mouse[49,50]. Thus, such models could serve to enrich our understanding of DMD at large and produce more robust and reliable end-points to determine efficacy of treatments.

DISORDERS OF THE BONE MARROW

Bone marrow diseases comprise a variety of conditions including severe anaemic haemoglobinopathies and over 250 different primary immunodeficiencies (PIDs). Current treatment modalities include transfusion of blood, or blood-derived products such as erythrocytes, immunoglobulins and platelets. However, haematopoietic stem cell (HSC) transplantation remains the only curative therapy to achieve permanent reconstitution. Despite the growing number of donor depositories, a human-leukocyte-antigen matched donor cannot be found for some patients. In these cases gene therapy using gene addition in autologous patient cells may offer a potentially safe and efficacious strategy. Successful and durable reconstitution using retroviral and/or lentiviral vectors has been achieved in patients suffering with various bone marrow disorders, including X-linked severe combined immunodeficiency (SCID-X1)[2,51,52], adenosine deaminase (ADA) deficiency[53–57], Wiskott-Aldrich syndrome[58,59], and β-thalassemia[60,61].

Haemoglobinopathies: Beta-thalassemia and sickle cell anaemia

Haemoglobinopathies or 'genetic anaemias' result from defects of mature haemoglobin. Both β-thalassemia and sickle cell anaemia (SCA) are caused by mutations on the *HBB* gene which encodes the β-globin chain. β-thalassemias are heterogeneous, with >200 mutation types across the *HBB* locus affecting different steps of β-globin production (initiation of transcription, splicing, and post-translational modification). The subsequent excess of α-globin causes apoptosis of red blood cells (RBCs) resulting in severe anaemia. In contrast, SCA is caused by a missense mutation (A-to-T transversion) at codon 6 of *HBB*, which causes RBCs to distort to a 'sickle' shape. Sickled RBCs constrict small capillaries causing severe tissue damage, acute painful crises, respiratory insufficiency and progressive organ damage[62]. SCA is one of the most common monogenic diseases with over 250,000 cases every year and is a major cause of morbidity and mortality worldwide[63].

Although non-targeted gene augmentation for haemoglobinopathies has made considerable progress[60,61,64], the high levels of gene expression required and the potential risk of insertional mutagenesis associated with uncontrolled viral integration remain challenging. Editing the *HBB* locus using programmable nucleases can instead allow for permanent β-globin correction. Recently, TALENs were used to achieve ~20% HDR-mediated knock-in of therapeutic β-globin full-length cDNA to the endogenous β-globin locus in K562 erythroleukemia cells[65]. Such a strategy would result in expression of β-globin from the cDNA instead of wild-type genomic sequence and would be therapeutic for both SCA and β-thalassemia. Separately, ZFNs were used to repair the SCA point mutation in CD34+ HSC progenitor cells (HSPCs). Co-delivery of ZFN

mRNA and a donor template led to 15% and 18% HDR in CD34+ cells derived from healthy-donor and patients with SCA, respectively[66]. Furthermore, correction of SCA mutation in patient cells led to production of wild-type haemoglobin *in vitro*. As only 10-30% of corrected donor cells are required to generate sufficient RBCs, gene repair in HSCs obtained from patients can reach a therapeutic threshold[66].

Another strategy being explored in the context of haemoglobinopathies is gene editing of disease-causing mutations in iPSCs. These can be differentiated into HSCs, which can then be used for autologous transplantation. Recently, independent groups have demonstrated progress in this area by applying ZFNs[67], TALENs[68] or Cas9[68,69] to correct different β-thalassemia mutations in various patient iPSCs. Corrected iPSCs maintained their ability to differentiate into erythroblasts with increased transcription of β -globin[69]. Similar progress has also been made for SCA where ZFNs[70] and TALENs[71] were used to repair the sickle cell point mutation in patient iPSCs. For treatment of sickle cell disease, it is particularly important to ensure that the highly paralogous genes encoding gamma-globin and delta-globin are not inadvertently mutagenized. Both studies specifically analysed these loci and demonstrated no nuclease-associated off-target activity. These studies demonstrate that human stem cells including HSCs and iPSCs, and nuclease-induced gene editing approaches can be used in combination to create corrected patient-derived cells. Despite the many advantages of iPSC technology, for immune-based therapy potential concerns of immunogenicity in iPSC and its derivatives should be thoroughly examined before clinical translation[72].

Apart from adult haemoglobin, the level of foetal haemoglobin (HbF) is also a key modifier of clinical severity of haemoglobinopathies. In patients with SCA, high HbF is associated with generally milder disease phenotype. This has been attributed to naturally occurring variants in the enhancer regions of *BCL11A*, a transcriptional repressor of HbF production in adult erythroid cells[73]. Recently, TALENs/ZFNs[74] and Cas9-nuclease[75] were used to specifically disrupt enhancer regions in *BCL11A* resulting in substantial HbF induction without the detrimental effects associated with complete loss of *BCL11A*. Such therapeutic gene editing approach could be used to elevate HbF to clinically relevant levels thereby ameliorating β haemoglobin disorders.

Primary immune deficiencies

Primary immune deficiencies comprise a heterogeneous group of rare diseases in which part of the immune system is missing or functions improperly. On the clinical spectrum, severe combined immunodeficiency (SCID) is the most severe form of PID, resulting in a development block in production of T-cells, with additional defects of B- and natural killer cells. The most common form of SCID, SCID-X1, is caused by mutations in the gene encoding the interleukin 2 receptor common gamma chain (*IL2RG*). Several groups have successfully used ZFNs to target and induce HDR in the *IL2RG* locus in human HSCs and embryonic stem cells, albeit with relatively low efficiencies[76–78]. A notable demonstration was the ZFN-mediated insertion of corrective cDNA (exons 5-8 that would correct all SCID-X1 mutations downstream of exon 4) into the *IL2RG* mutational hotspot in long-term repopulating HSCs. This led to the correction of

defective *IL2RG* in HSPCs from a subject with SCID-X1 and multi-lineage differentiation upon transplantation into immunodeficient SCID mice[78]. Separately, TALENs were used to specifically target and induce HDR in the *IL2RG* locus of Jurkat cells[79].

While direct genome editing in HSCs is an attractive alternative to viral gene addition therapy, both approaches are dependent on the capability to efficiently culture and expand HSCs *ex vivo*. In addition, HSC-based transplantation approaches generally involve myeloablative conditioning, which given the young age and immunocompromised state of SCID patients, poses significant risk. An iPSC-based approach could provide an unlimited source of subject-derived, corrected cells from which immune cells could be derived continuously for transfusion, but it would be a cumbersome approach. Issues regarding the efficiency of iPSC differentiation towards the haematopoietic lineage also need addressing. On the other hand, in diseases where the number of HSC is compromised, iPSCs could provide a ready source of corrected cells. ZFNs have been used to correct various types of chronic granulomatous disease (CGD) by introducing five different functional genes into the *AAVS1* 'safe harbour' in iPSCs generated from peripheral HSCs. Using *in vitro* myeloid differentiation, normal granulocytes were generated from the corrected iPSCs[80]. Provirus-free iPSCs resulting from methods in which no transgene integration events are required have been generated[81]. Such iPSCs overcome concerns related to insertional mutagenesis and spurious transgene expression, and are preferred for clinical application. Provirus-free iPSCs have been generated from a SCID-X1 patient, and the genetic defect corrected utilizing TALENs[82]. These iPSCs retained their differentiation potential into NK cells, which expressed mature cell markers and had the correctly spliced *IL2RG*. This provided the first evidence of gene repair of SCID-X1 patient iPSCs

resulting in regeneration of mature lymphoid cells *in vitro*. For radiosensitive SCID, ZFNs were utilized to correct a mutation on *Prkdc* gene in primary mouse fibroblasts and iPSCs[83]. The corrected cells retained their potential to differentiate into functional T-cells *in vitro,* overcoming the developmental block.

Even though gene editing in the context of PIDs is not yet ready to be applied in a clinical setting, it already offers valuable tools to study and model immune disorders at the cellular level. With CRISPR-Cas9, zygote injections can be done in a one-day procedure generating gene modified mice in less than 4 weeks[84]. NSG mice have been efficiently manipulated in this way[85]. Similarly, a rat model of SCID-X1 has been generated using ZFNs[86]. Additionally, endonucleases have been used to generate knockout models in animals previously unamenable to efficient genetic modification. These include: rabbits with *IL2RG*, *RAG1*, or *RAG2* knockout[87–90]; hamsters with *STAT2* knockout[91]; and monkeys with *RAG1* knockout[92].

HAEMOPHILIA

Haemophilia is a group of inherited bleeding disorders that affect the blood clotting process. This deficit is most often caused by mutations in genes coding for either clotting factor VIII (haemophilia A) or factor IX (haemophilia B), two crucial components of the blood coagulation cascade. People with haemophilia often experience internal bleeding into knee, hip, elbow and ankle and subsequent joint disease. Current care for haemophilia involves life-long infusions of clotting factors. Although highly effective at disease management, clotting factors are short-

lived in circulation. This necessitates intravenous delivery at least 2-3 times per week, which is both invasive and expensive[93]. Alternatively gene therapy, via transfer of a normal copy of *F8* or *F9* gene (encoding FVIII and FIX), may enable permanent correction and stable endogenous expression. Estimates suggest that an increase of only 1% in plasma FVIII or FIX levels would be therapeutic and encouraging results have been obtained by AAV-vector mediated gene transfer[94,95].

Genome editing using programmable nucleases has also shown promise by allowing *in situ* targeting of haemophilia A and B. For haemophilia B, where mutations span across the entire coding region of human *F9 (hF9)* gene, HDR has been demonstrated by direct injections of AAV8-ZFNs and a corrective cDNA (promoterless exons 2-8 bearing a splice acceptor and a poly A signal flanked by homology arms) into livers of neonatal[96] and adult humanized haemophilic mice[97]. In neonatal mice the level of gene repair was sufficient to correct clotting times, and partial hepatectomy showed stable genome modification, as levels of FIX were stable in the genome-edited liver. In sharp contrast, episomal AAV *F9* transgene delivery could not overcome partial hepatectomy and FIX levels decreased to almost background levels, highlighting the advantage of gene editing[96]. Analyses in adult hF9 mice showed sustained expression of human FIX, averaging 23% of normal levels at week 60[97].

In a novel strategy, ZFNs were utilised for targeted integration of promoterless *F8* and *F9* therapeutic transgenes within the highly expressed albumin gene[98]. The albumin gene is expressed at high levels in liver cells and loss of its expression from a few percent of cells does not appear to be detrimental. This study demonstrated AAV8-ZFN mediated long-term

expression of both human FVIII and FIX at therapeutic levels in haemophilia A and B mouse models, respectively[98] and is progressing towards clinical application. This highlights that certain genomic sequences, such as the albumin locus, are both permissive and amenable for transgene integration. In this context it is worth mentioning a parallel study, which demonstrated nuclease-free, AAV-vector mediated targeting of the albumin locus to drive expression of FVIII and FIX in new born and adult mice. The animals expressed levels of clotting factor that were between 7-20% of normal[99].

Efforts to re-introduce *F8* in haemophilia A have been more challenging. Such an approach has primarily been limited by the large size (7 Kb) of its cDNA, inefficient protein production and complex mutations, comprising mostly of large inversions and duplications. Among the genotypes that result in haemophilia A, two different types of chromosomal inversions involving a portion of *F8* gene are most frequent, accounting for 50% of cases. An initial study using TALENs led to correction of an inversion mutation in an iPSC model, establishing proof-of-concept[100]. Building up on this, a separate study demonstrated correction of a ~600 Kb inversion using CRISPR-Cas9 system in iPSCs derived from haemophilia A patients[101]. This was achieved by using Cas9 nucleases with target sites on either side of the inversion. Corrected iPSC colonies were clonally expanded before differentiating into epithelial cells, which produced FVIII protein *in vitro*. Transplantation of corrected endothelial cells rescued injury mortality in haemophilic mice in a short-term experiment[101].

Haemophilia gene therapy is a forerunner in the field, with promising clinical development using safer viral vectors. Although the genome editing strategies outlined here are still under

development and currently lack safety validation, they can be combined with existing advancements such as AAV-vector mediated delivery to the liver to further clinical application. Furthermore, manipulation of the albumin locus may allow for this condition to be treated in the absence of nucleases, and the locus could also be used as a 'safe harbour' for expression of other secreted proteins[99].

CONCLUSION

The past few years have seen notable demonstrations of genome editing being applied across a multitude of diseases. Whilst the application of nucleases holds significant therapeutic promise, optimum progress can only be achieved by examining the advancement of gene editing holistically. A number of ubiquitous challenges need to be considered, mostly relating to efficacy of genome editing at the target sequence, safety concerns related to nuclease-associated off-target effects and delivery of gene editing tools.

Editing efficiency is dependent upon the DNA repair process being relied upon. In instances where the desired effect can be achieved by NHEJ, the correction will most likely occur at a relatively high frequency as NHEJ is the major repair pathway in mammalian cells, although the usefulness of this approach may be limited by the stochastic nature of the Indels being formed. As discussed earlier, NHEJ has been used to mediate disruption of coding and regulatory sequences, targeted deletions of exons or large intervening sequences and disruption of splice sites. Methods that can predict and evaluate micro-homology sites can be used to bias the

repair towards frameshift mutations in protein coding sequence[102]. This would partially address the potential reduction in efficiency caused by micro-homology-mediated end joining, a secondary end-joining pathway with a bias for in-frame deletions[103].

Precise HDR-based locus alterations allow targeted insertion or *in situ* correction of mutated DNA sequence, which are suitable for a large subset of disease-causing mutations. However, they are reliant upon homologous recombination, which is normally limited to S and G2 phases of the cell cycle, requires a DNA template and inherently occurs at lower frequencies. HDR-based strategies may also require enrichment and expansion of corrected cells, normally restricted to *ex vivo* approaches. While *ex vivo* manipulation may be possible in diseases like those affecting bone marrow HSCs, it would limit applications in diseases with multi-organ involvement or those where transplantation is not an option. Further progress into enabling HDR with higher efficacy would therefore be beneficial. In this respect, a recent report has demonstrated that it may be possible to transiently activate HDR in G1 cells by restoring BRCA1-PALB2 interaction[104], possibly facilitating HDR genome-editing in quiescent cells.

Specificity of genome editing is one of the major safety concerns for translational research. Owing to sequence similarities within the genome, endonucleases can cleave and modify off-target regions that are distinct from the site of interest. Off-target effects can lead to unwanted genetic modifications causing cellular stress, functional impairment or enhancement, and oncogenicity, all of which could have detrimental effects clinically[105]. Considerable work is being undertaken to increase fidelity through better design of nuclease components, which has led to improved variants such as megaTALs[106], dead Cas9-Fok1 fusion nucleases[107], Cas9 nickases[108]

and Cas9 nucleases with truncated guide RNAs[109]. Furthermore, screens of bacterial strains have led to discovery of several alternative Cas9-nucleases with varying specificities[110–113]. More recently, Cpf1, a prominent CRISPR variant that requires a shorter RNA and generates a staggered cut which could improve HDR, has also been described[114]. All of these variants highlight the progress in the field but still require extensive examination prior to their application in a translational research setting. Specificity of modification can also be helped by careful target site selection and use of delivery methods that would allow for efficient but transient expression of nucleases. New methods, such as GUIDE-seq[115] and BLESS[116], have also been developed for unbiased evaluation of off-target modifications on a genome-wide scale.

The final challenge pertains to the delivery of gene editing reagents including nucleases and a donor template in case of HDR. A variety of delivery approaches are being explored depending on cell types to be targeted. Cells that can be cultured and engrafted under *ex vivo* conditions are amenable to delivery via nucleic acids, proteins and viral vector systems; mRNA and protein delivery of the nucleases are now well-established procedures. However, for *in vivo* gene editing applications the most promising delivery systems are viral vectors. Both integrating and non-integrating viral vector systems have been explored in this context, although the latter are favoured due to their safety profile. In particular, AAV vectors, with a wide range of serotypes and ability to transduce a variety of tissue types are promising candidates. However, AAV vectors are restricted by their small packaging capacity, which poses challenges for large nuclease proteins such as TALENs or Cas9.

Despite the outlined challenges, genome editing is advancing at fast pace, with continued focus on pioneering and improving strategies. The successful use of ZFNs targeting CCR5, the co-receptor necessary for HIV to infect T-cells, to control AIDS, remains the single demonstration of gene-editing in a therapeutic setting[117], but will be quickly followed by others. Our developing understanding of programmable nucleases and DNA repair, coupled to general progress in regenerative medicine and knowledge of inherited disease pathophysiology, should warrant exciting outcomes from therapeutic genome editing, which we could only dream of in the nineties[118] – we look forward to the next twenty years.

ACKNOWLEDGEMENTS

We would like to thank Linda Popplewell for critical reading of the manuscript and for helpful discussions. Rafael J. Yáñez-Muñoz acknowledges financial support from the 7th EU Framework Programme (PERSIST project, grant agreement no. 222878) and Royal Holloway, University of London; his work has also received research support from Sangamo Biosciences Inc. Sangamo has had no role in the current study. Versha Prakash and Marc Moore acknowledge financial support from Royal Holloway, University of London and Muscular Dystrophy UK.

REFERENCES

1. Rodwell C., Aymé S., eds. *2014 Report on the State of the Art of Rare Disease Activities in*

Europe Part Ii: Key Developments in the Field of Rare Diseases in Europe in 2013. 1–90 (2014). http://www.eucerd.eu/upload/file/Reports/2014ReportStateofArtRDActivities.pdf

2. Hacein-Bey-Abina, S., Hauer, J., Lim, A., Picard, C., Wang, G. P., Berry, C. C., *et al.* Efficacy of gene therapy for X-linked severe combined immunodeficiency. *N. Engl. J. Med.* **363,** 355–64 (2010).
3. Ylä-Herttuala, S. Endgame: glybera finally recommended for approval as the first gene therapy drug in the European union. *Mol. Ther.* **20,** 1831–2 (2012).
4. Alton, E. W. F. W., Armstrong, D. K., Ashby, D., Bayfield, K. J., Bilton, D., Bloomfield, E. V, *et al.* Repeated nebulisation of non-viral CFTR gene therapy in patients with cystic fibrosis: a randomised, double-blind, placebo-controlled, phase 2b trial. *Lancet Respir. Med.* **3,** 684–91 (2015).
5. Gaj, T., Gersbach, C. A. & Barbas, C. F. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* **31,** 397–405 (2013).
6. Sadelain, M., Papapetrou, E. P. & Bushman, F. D. Safe harbours for the integration of new DNA in the human genome. *Nat. Rev. Cancer* **12,** 51–58 (2011).
7. O'Sullivan, B. P. & Freedman, S. D. Cystic fibrosis. *Lancet* **373,** 1891–904 (2009).
8. Fanen, P., Wohlhuter-Haddad, A. & Hinzpeter, A. Genetics of cystic fibrosis: CFTR mutation classifications toward genotype-based CF therapies. *Int. J. Biochem. Cell Biol.* **52,** 94–102 (2014).
9. Griesenbach, U., Pytel, K. M. & Alton, E. W. F. W. Cystic Fibrosis Gene Therapy in the UK and Elsewhere. *Hum. Gene Ther.* **26,** 266–75 (2015).

10. Welch, E. M., Barton, E. R., Zhuo, J., Tomizawa, Y., Friesen, W. J., Trifillis, P., *et al.* PTC124 targets genetic disorders caused by nonsense mutations. *Nature* **447,** 87–91 (2007).

11. Ren, H. Y., Grove, D. E., De La Rosa, O., Houck, S. A., Sopha, P., Van Goor, F., *et al.* VX-809 corrects folding defects in cystic fibrosis transmembrane conductance regulator protein through action on membrane-spanning domain 1. *Mol. Biol. Cell* **24,** 3016–24 (2013).

12. Eckford, P. D. W., Li, C., Ramjeesingh, M. & Bear, C. E. Cystic fibrosis transmembrane conductance regulator (CFTR) potentiator VX-770 (ivacaftor) opens the defective channel gate of mutant CFTR in a phosphorylation-dependent but ATP-independent manner. *J. Biol. Chem.* **287,** 36639–49 (2012).

13. Davies, J. C., Wainwright, C. E., Canny, G. J., Chilvers, M. A., Howenstine, M. S., Munck, A., *et al.* Efficacy and safety of ivacaftor in patients aged 6 to 11 years with cystic fibrosis with a G551D mutation. *Am. J. Respir. Crit. Care Med.* **187,** 1219–25 (2013).

14. Davies, J., Sheridan, H., Bell, N., Cunningham, S., Davis, S. D., Elborn, J. S., *et al.* Assessment of clinical response to ivacaftor with lung clearance index in cystic fibrosis patients with a G551D-CFTR mutation and preserved spirometry: a randomised controlled trial. *Lancet. Respir. Med.* **1,** 630–8 (2013).

15. Ramsey, B. W., Davies, J., McElvaney, N. G., Tullis, E., Bell, S. C., Dřevínek, P., *et al.* A CFTR potentiator in patients with cystic fibrosis and the G551D mutation. *N. Engl. J. Med.* **365,** 1663–72 (2011).

16. Barry, P. J., Plant, B. J., Nair, A., Bicknell, S., Simmonds, N. J., Bell, N. J., *et al.* Effects of ivacaftor in patients with cystic fibrosis who carry the G551D mutation and have severe lung disease. *Chest* **146,** 152–8 (2014).

17. Yu, H., Burton, B., Huang, C.-J., Worley, J., Cao, D., Johnson, J. P., *et al.* Ivacaftor potentiation of multiple CFTR channels with gating mutations. *J. Cyst. Fibros.* **11,** 237–45 (2012).

18. Wainwright, C. E., Elborn, J. S., Ramsey, B. W., Marigowda, G., Huang, X., Cipolli, M., *et al.* Lumacaftor-Ivacaftor in Patients with Cystic Fibrosis Homozygous for Phe508del CFTR. *N. Engl. J. Med.* **373,** 220–31 (2015).

19. Beumer, W., Swildens, J., Henig, N., Anthonijsz, H., Biasutto, P., Leal, T., *et al.* WS01.2 QR-010, an RNA therapy, restores CFTR function using in vitro and in vivo models of ΔF508 CFTR. *J. Cyst. Fibros.* **14,** S1 (2015).

20. N. Henig, W. Beumer, H. Anthonijsz, M. Beka, N. Panin, T. Leal, B. Matthee, T. R. QR-010, an RNA Therapy, Restores CFTR Function in the Saliva Secretion Assay. in *American Thoracic Society International Conference Meetings Abstracts* (2015)

21. Griesenbach, U. & Alton, E. W. F. W. Moving forward: Cystic fibrosis gene therapy. *Hum. Mol. Genet.* **22,** 52–58 (2013).

22. Lee, C. M., Flynn, R., Hollywood, J. A., Scallan, M. F. & Harrison, P. T. Correction of the ΔF508 Mutation in the Cystic Fibrosis Transmembrane Conductance Regulator Gene by Zinc-Finger Nuclease Homology-Directed Repair. *Biores. Open Access* **1,** 99–108 (2012).

23. Schwank, G., Koo, B.-K., Sasselli, V., Dekkers, J. F., Heo, I., Demircan, T., *et al. Functional Repair of CFTR by CRISPR/Cas9 in Intestinal Stem Cell Organoids of Cystic Fibrosis Patients*. *Cell Stem Cell* **13,** (2013).

24. Dekkers, J. F., Wiegerinck, C. L., de Jonge, H. R., Bronsveld, I., Janssens, H. M., de Winter-de Groot, K. M., *et al.* A functional CFTR assay using primary cystic fibrosis intestinal

organoids. *Nat. Med.* **19,** 939–45 (2013).

25. Yui, S., Nakamura, T., Sato, T., Nemoto, Y., Mizutani, T., Zheng, X., *et al.* Functional engraftment of colon epithelium expanded in vitro from a single adult $Lgr5^+$ stem cell. *Nat. Med.* **18,** 618–23 (2012).

26. Crane, A. M., Kramer, P., Bui, J. H., Chung, W. J., Li, X. S., Gonzalez-Garay, M. L., *et al.* Targeted Correction and Restored Function of the CFTR Gene in Cystic Fibrosis Induced Pluripotent Stem Cells. *Stem cell reports* **4,** 569–77 (2015).

27. Sargent, R. G., Suzuki, S. & Gruenert, D. C. Nuclease-mediated double-strand break (DSB) enhancement of small fragment homologous recombination (SFHR) gene modification in human-induced pluripotent stem cells (hiPSCs). *Methods Mol. Biol.* **1114,** 279–90 (2014).

28. Firth, A. L., Menon, T., Parker, G. S., Qualls, S. J., Lewis, B. M., Ke, E., *et al.* Functional Gene Correction for Cystic Fibrosis in Lung Epithelial Cells Generated from Patient iPSCs. *Cell Rep.* **12,** 1385–90 (2015).

29. Cheng, A. W., Wang, H., Yang, H., Shi, L., Katz, Y., Theunissen, T. W., *et al.* Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res.* **23,** 1163–71 (2013).

30. Perez-Pinera, P., Kocak, D. D., Vockley, C. M., Adler, A. F., Kabadi, A. M., Polstein, L. R., *et al.* RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods* **10,** 973–6 (2013).

31. Bushby, K., Finkel, R., Birnkrant, D. J., Case, L. E., Clemens, P. R., Cripe, L., *et al.* Diagnosis and management of Duchenne muscular dystrophy, part 1: diagnosis, and pharmacological and psychosocial management. *Lancet Neurol.* **9,** 77–93 (2010).

32. Bushby, K., Finkel, R., Birnkrant, D. J., Case, L. E., Clemens, P. R., Cripe, L., *et al.* Diagnosis and management of Duchenne muscular dystrophy, part 2: implementation of multidisciplinary care. *Lancet Neurol.* **9,** 177–189 (2010).

33. Ricotti, V., Ridout, D. a, Scott, E., Quinlivan, R., Robb, S. a, Manzur, A. Y., *et al.* Long-term benefits and adverse effects of intermittent versus daily glucocorticoids in boys with Duchenne muscular dystrophy. *J. Neurol. Neurosurg. Psychiatry* **84,** 698–705 (2013).

34. Aartsma-Rus, A., Fokkema, I., Verschuuren, J., Ginjaar, I., van Deutekom, J., van Ommen, G.-J., *et al.* Theoretic applicability of antisense-mediated exon skipping for Duchenne muscular dystrophy mutations. *Hum. Mutat.* **30,** 293–9 (2009).

35. Jarmin, S., Kymalainen, H., Popplewell, L. & Dickson, G. New developments in the use of gene therapy to treat Duchenne muscular dystrophy. *Expert Opin. Biol. Ther.* **14,** 209–30 (2014).

36. England, S. B., Nicholson, L. V, Johnson, M. a, Forrest, S. M., Love, D. R., Zubrzycka-Gaarn, E. E., *et al.* Very mild muscular dystrophy associated with the deletion of 46% of dystrophin. *Nature* **343,** 180–182 (1990).

37. Ousterout, D. G., Kabadi, A. M., Thakore, P. I., Majoros, W. H., Reddy, T. E. & Gersbach, C. a. Multiplex CRISPR/Cas9-based genome editing for correction of dystrophin mutations that cause Duchenne muscular dystrophy. *Nat. Commun.* **6,** 6244 (2015).

38. Popplewell, L., Koo, T., Leclerc, X., Duclert, A., Mamchaoui, K., Gouble, A., *et al.* Gene correction of a duchenne muscular dystrophy mutation by meganuclease-enhanced exon knock-in. *Hum. Gene Ther.* **24,** 692–701 (2013).

39. Li, H. L., Fujimoto, N., Sasakawa, N., Shirai, S., Ohkame, T., Sakuma, T., *et al.* Precise

Correction of the Dystrophin Gene in Duchenne Muscular Dystrophy Patient Induced Pluripotent Stem Cells by TALEN and CRISPR-Cas9. *Stem Cell Reports* **4,** (2014).

40. Ousterout, D. G., Kabadi, A. M., Thakore, P. I., Perez-Pinera, P., Brown, M. T., Majoros, W. H., *et al.* Correction of dystrophin expression in cells from duchenne muscular dystrophy patients through genomic excision of exon 51 by zinc finger nucleases. *Mol. Ther.* **23,** 523–32 (2015).

41. Ousterout, D. G., Kabadi, A. M., Thakore, P. I., Majoros, W. H., Reddy, T. E. & Gersbach, C. A. Multiplex CRISPR/Cas9-based genome editing for correction of dystrophin mutations that cause Duchenne muscular dystrophy. *Nat. Commun.* **6,** 6244 (2015).

42. Briggs, D. & Morgan, J. E. Recent progress in satellite cell/myoblast engraftment -- relevance for therapy. *FEBS J.* **280,** 4281–93 (2013).

43. Long, C., McAnally, J. R., Shelton, J. M., Mireault, A. a, Bassel-Duby, R. & Olson, E. N. Prevention of muscular dystrophy in mice by CRISPR/Cas9-mediated editing of germline DNA. *Science* **1184,** 1–34 (2014).

44. Nakamura, K., Fujii, W., Tsuboi, M., Tanihata, J., Teramoto, N., Takeuchi, S., *et al.* Generation of muscular dystrophy model rats with a CRISPR/Cas system. *Sci. Rep.* **4,** 5635 (2014).

45. Larcher, T., Lafoux, A., Tesson, L., Remy, S., Thepenier, V., François, V., *et al.* Characterization of dystrophin deficient rats: a new model for Duchenne muscular dystrophy. *PLoS One* **9,** e110371 (2014).

46. Chen, Y., Zheng, Y., Kang, Y., Yang, W., Niu, Y., Guo, X., *et al.* Functional disruption of the dystrophin gene in Rhesus Monkey Using CRISPR/Cas9. *Hum. Mol. Genet.* **24,** 1–11

(2015).

47. Klymiuk, N., Blutke, A., Graf, A., Krause, S., Burkhardt, K., Wuensch, A., *et al.* Dystrophin-deficient pigs provide new insights into the hierarchy of physiological derangements of dystrophic muscle. *Hum. Mol. Genet.* **22,** 4368–82 (2013).

48. Hollinger, K., Yang, C. X., Montz, R. E., Nonneman, D., Ross, J. W. & Selsby, J. T. Dystrophin insufficiency causes selective muscle histopathology and loss of dystrophin-glycoprotein complex assembly in pig skeletal muscle. *FASEB J.* **28,** 1600–1609 (2014).

49. Whitmore, C. & Morgan, J. What do mouse models of muscular dystrophy tell us about the DAPC and its components? *Int. J. Exp. Pathol.* **95,** 365–77 (2014).

50. McGreevy, J. W., Hakim, C. H., McIntosh, M. A. & Duan, D. Animal models of Duchenne muscular dystrophy: from basic mechanisms to gene therapy. *Dis. Model. Mech.* **8,** 195–213 (2015).

51. Cavazzana-Calvo, M., Hacein-Bey, S., de Saint Basile, G., Gross, F., Yvon, E., Nusbaum, P., *et al.* Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science* **288,** 669–72 (2000).

52. Gaspar, H. B., Cooray, S., Gilmour, K. C., Parsley, K. L., Adams, S., Howe, S. J., *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).

53. Blaese, R. M., Culver, K. W., Miller, A. D., Carter, C. S., Fleisher, T., Clerici, M., *et al.* T lymphocyte-directed gene therapy for ADA- SCID: initial trial results after 4 years. *Science* **270,** 475–80 (1995).

54. Kohn, D. B., Weinberg, K. I., Nolta, J. A., Heiss, L. N., Lenarsky, C., Crooks, G. M., *et al.*

Engraftment of gene-modified umbilical cord blood cells in neonates with adenosine deaminase deficiency. *Nat. Med.* **1,** 1017–23 (1995).

55. Aiuti, A., Slavin, S., Aker, M., Ficara, F., Deola, S., Mortellaro, A., *et al.* Correction of ADA-SCID by stem cell gene therapy combined with nonmyeloablative conditioning. *Science* **296,** 2410–3 (2002).

56. Aiuti, A., Cattaneo, F., Galimberti, S., Benninghoff, U., Cassani, B., Callegaro, L., *et al.* Gene therapy for immunodeficiency due to adenosine deaminase deficiency. *N. Engl. J. Med.* **360,** 447–58 (2009).

57. Gaspar, H. B., Cooray, S., Gilmour, K. C., Parsley, K. L., Zhang, F., Adams, S., *et al.* Hematopoietic stem cell gene therapy for adenosine deaminase-deficient severe combined immunodeficiency leads to long-term immunological recovery and metabolic correction. *Sci. Transl. Med.* **3,** 97ra80 (2011).

58. Boztug, K., Schmidt, M., Schwarzer, A., Banerjee, P. P., Díez, I. A., Dewey, R. A., *et al.* Stem-cell gene therapy for the Wiskott-Aldrich syndrome. *N. Engl. J. Med.* **363,** 1918–27 (2010).

59. Aiuti, A., Biasco, L., Scaramuzza, S., Ferrua, F., Cicalese, M. P., Baricordi, C., *et al.* Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. *Science* **341,** 1233151 (2013).

60. Bank, A., Dorazio, R. & Leboulch, P. A phase I/II clinical trial of beta-globin gene therapy for beta-thalassemia. *Ann. N. Y. Acad. Sci.* **1054,** 308–16 (2005).

61. Cavazzana-Calvo, M., Payen, E., Negre, O., Wang, G., Hehir, K., Fusil, F., *et al.* Transfusion independence and HMGA2 activation after gene therapy of human β-thalassaemia.

Nature **467,** 318–22 (2010).

62. Sankaran, V. G. & Weiss, M. J. Anemia: progress in molecular mechanisms and therapies. *Nat. Med.* **21,** 221–30 (2015).

63. Modell, B. & Darlison, M. Global epidemiology of haemoglobin disorders and derived service indicators. *Bull. World Health Organ.* **86,** 480–7 (2008).

64. Levasseur, D. N., Ryan, T. M., Pawlik, K. M. & Townes, T. M. Correction of a mouse model of sickle cell disease: lentiviral/antisickling beta-globin gene transduction of unmobilized, purified hematopoietic stem cells. *Blood* **102,** 4312–9 (2003).

65. Voit, R. A., Hendel, A., Pruett-Miller, S. M. & Porteus, M. H. Nuclease-mediated gene editing by homologous recombination of the human globin locus. *Nucleic Acids Res.* **42,** 1365–1378 (2013).

66. Hoban, M. D., Cost, G. J., Mendel, M. C., Romero, Z., Kaufman, M. L., Joglekar, A. V., *et al.* Correction of the sickle cell disease mutation in human hematopoietic stem/progenitor cells. *Blood* **125,** 2597–2604 (2015).

67. Ma, N., Shan, Y., Liao, B., Kong, G., Wang, C., Huang, K., *et al.* Factor-induced Reprogramming and Zinc Finger Nuclease-aided Gene Targeting Cause Different Genome Instability in β-Thalassemia Induced Pluripotent Stem Cells (iPSCs). *J. Biol. Chem.* **290,** 12079–89 (2015).

68. Xu, P., Tong, Y., Liu, X., Wang, T., Cheng, L., Wang, B., *et al.* Both TALENs and CRISPR/Cas9 directly target the HBB IVS2-654 (C > T) mutation in β-thalassemia-derived iPSCs. *Sci. Rep.* **5,** 12065 (2015).

69. Xie, F., Ye, L., Chang, J. C., Beyer, A. I., Wang, J., Muench, M. O., *et al.* Seamless gene

correction of β-thalassemia mutations in patient-specific iPSCs using CRISPR/Cas9 and piggyBac. *Genome Res.* **24,** 1526–33 (2014).

70. Sebastiano, V., Maeder, M. L., Angstman, J. F., Haddad, B., Khayter, C., Yeo, D. T., *et al.* In situ genetic correction of the sickle cell anemia mutation in human induced pluripotent stem cells using engineered zinc finger nucleases. *Stem Cells* **29,** 1717–26 (2011).

71. Sun, N. & Zhao, H. Seamless correction of the sickle cell disease mutation of the HBB gene in human induced pluripotent stem cells using TALENs. *Biotechnol. Bioeng.* **111,** 1048–53 (2014).

72. Jiang, Z., Han, Y. & Cao, X. Induced pluripotent stem cell (iPSCs) and their application in immunotherapy. *Cell. Mol. Immunol.* **11,** 17–24 (2014).

73. Bauer, D. E., Kamran, S. C., Lessard, S., Xu, J., Fujiwara, Y., Lin, C., *et al.* An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. *Science* **342,** 253–7 (2013).

74. Vierstra, J., Reik, A., Chang, K.-H., Stehling-Sun, S., Zhou, Y., Hinkley, S. J., *et al.* Functional footprinting of regulatory DNA. *Nat. Methods* **12,** 927–30 (2015).

75. Canver, M. C., Smith, E. C., Sher, F., Pinello, L., Sanjana, N. E., Shalem, O., *et al.* BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. *Nature* **527,** 192–7 (2015).

76. Urnov, F. D., Miller, J. C., Lee, Y.-L., Beausejour, C. M., Rock, J. M., Augustus, S., *et al.* Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* **435,** 646–51 (2005).

77. Lombardo, A., Genovese, P., Beausejour, C. M., Colleoni, S., Lee, Y.-L., Kim, K. A., *et al.*

Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat. Biotechnol.* **25,** 1298–306 (2007).

78. Genovese, P., Schiroli, G., Escobar, G., Di Tomaso, T., Firrito, C., Calabria, A., *et al.* Targeted genome editing in human repopulating haematopoietic stem cells. *Nature* **510,** 235–40 (2014).

79. Matsubara, Y., Chiba, T., Kashimada, K., Morio, T., Takada, S., Mizutani, S., *et al.* Transcription activator-like effector nuclease-mediated transduction of exogenous gene into IL2RG locus. *Sci. Rep.* **4,** 5043 (2014).

80. Merling, R. K., Sweeney, C. L., Chu, J., Bodansky, A., Choi, U., Priel, D. L., *et al.* An AAVS1-targeted minigene platform for correction of iPSCs from all five types of chronic granulomatous disease. *Mol. Ther.* **23,** 147–57 (2015).

81. Li, M. & Izpisua Belmonte, J. C. No factor left behind: generation of transgene-free induced pluripotent stem cells. *Am. J. Stem Cells* **1,** 75–80 (2012).

82. Menon, T., Firth, A. L., Scripture-Adams, D. D., Galic, Z., Qualls, S. J., Gilmore, W. B., *et al.* Lymphoid Regeneration from Gene-Corrected SCID-X1 Subject-Derived iPSCs. *Cell Stem Cell* **16,** 367–72 (2015).

83. Rahman, S. H., Kuehle, J., Reimann, C., Mlambo, T., Alzubi, J., Maeder, M. L., *et al.* Rescue of DNA-PK Signaling and T-Cell Differentiation by Targeted Genome Editing in a prkdc Deficient iPSC Disease Model. *PLoS Genet.* **11,** e1005239 (2015).

84. Yang, H., Wang, H. & Jaenisch, R. Generating genetically modified mice using CRISPR/Cas-mediated genome engineering. *Nat. Protoc.* **9,** 1956–1968 (2014).

85. Li, F., Cowley, D. O., Banner, D., Holle, E., Zhang, L. & Su, L. Efficient genetic manipulation

of the NOD-Rag1-/-IL2RgammaC-null mouse by combining in vitro fertilization and CRISPR/Cas9 technology. *Sci. Rep.* **4,** 5290 (2014).

86. Mashimo, T., Takizawa, A., Voigt, B., Yoshimi, K., Hiai, H., Kuramoto, T., *et al.* Generation of knockout rats with X-linked severe combined immunodeficiency (X-SCID) using zinc-finger nucleases. *PLoS One* **5,** e8870 (2010).

87. Wang, Y., Fan, N., Song, J., Zhong, J., Guo, X., Tian, W., *et al.* Generation of knockout rabbits using transcription activator-like effector nucleases. *Cell Regen. (London, England)* **3,** 3 (2014).

88. Yan, Q., Zhang, Q., Yang, H., Zou, Q., Tang, C., Fan, N., *et al.* Generation of multi-gene knockout rabbits using the Cas9/gRNA system. *Cell Regen. (London, England)* **3,** 12 (2014).

89. Flisikowska, T., Thorey, I. S., Offner, S., Ros, F., Lifke, V., Zeitler, B., *et al.* Efficient immunoglobulin gene disruption and targeted replacement in rabbit using zinc finger nucleases. *PLoS One* **6,** e21045 (2011).

90. Song, J., Zhong, J., Guo, X., Chen, Y., Zou, Q., Huang, J., *et al.* Generation of RAG 1- and 2-deficient rabbits by embryo microinjection of TALENs. *Cell Res.* **23,** 1059–62 (2013).

91. Fan, Z., Li, W., Lee, S. R., Meng, Q., Shi, B., Bunch, T. D., *et al.* Efficient gene targeting in golden Syrian hamsters by the CRISPR/Cas9 system. *PLoS One* **9,** e109755 (2014).

92. Niu, Y., Shen, B., Cui, Y., Chen, Y., Wang, J., Wang, L., *et al.* Generation of gene-modified cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos. *Cell* **156,** 836–43 (2014).

93. High, K. H., Nathwani, a., Spencer, T. & Lillicrap, D. Current status of haemophilia gene

therapy. *Haemophilia* **20,** 43–49 (2014).

94. Nathwani, A. C., Tuddenham, E. G. D., Rangarajan, S., Rosales, C., McIntosh, J., Linch, D. C., *et al.* Adenovirus-associated virus vector-mediated gene transfer in hemophilia B. *N. Engl. J. Med.* **365,** 2357–65 (2011).

95. Nathwani, A. C., Reiss, U. M., Tuddenham, E. G. D., Rosales, C., Chowdary, P., McIntosh, J., *et al.* Long-Term Safety and Efficacy of Factor IX Gene Therapy in Hemophilia B. *N. Engl. J. Med.* **371,** 1994–2004 (2014).

96. Li, H., Haurigot, V., Doyon, Y., Li, T., Wong, S. Y., Bhagwat, A. S., *et al.* In vivo genome editing restores haemostasis in a mouse model of haemophilia. *Nature* **475,** 217–21 (2011).

97. Anguela, X. M., Sharma, R., Doyon, Y., Miller, J. C., Li, H., Haurigot, V., *et al.* Robust ZFN-mediated genome editing in adult hemophilic mice. *Blood* **122,** 3283–3287 (2013).

98. Sharma, R., Anguela, X. M., Doyon, Y., Wechsler, T., DeKelver, R. C., Sproul, S., *et al.* In vivo genome editing of the albumin locus as a platform for protein replacement therapy. *Blood* **4521,** (2015).

99. Barzel, a, Paulk, N. K., Shi, Y., Huang, Y., Chu, K., Zhang, F., *et al.* Promoterless gene targeting without nucleases ameliorates haemophilia B in mice. *Nature* **517,** 360–364 (2014).

100. Park, C.-Y., Kim, J., Kweon, J., Son, J. S., Lee, J. S., Yoo, J.-E., *et al.* Targeted inversion and reversion of the blood coagulation factor 8 gene in human iPS cells using TALENs. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 9253–8 (2014).

101. Park, C.-Y., Kim, D. H., Son, J. S., Sung, J. J., Lee, J., Bae, S., *et al.* Functional Correction of

Large Factor VIII Gene Chromosomal Inversions in Hemophilia A Patient-Derived iPSCs Using CRISPR-Cas9. *Cell Stem Cell* **17,** 213–220 (2015).

102. Bae, S., Kweon, J., Kim, H. S. & Kim, J.-S. Microhomology-based choice of Cas9 nuclease target sites. *Nat. Methods* **11,** 705–6 (2014).

103. Morton, J., Davis, M. W., Jorgensen, E. M. & Carroll, D. Induction and repair of zinc-finger nuclease-targeted double-strand breaks in Caenorhabditis elegans somatic cells. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 16370–5 (2006).

104. Orthwein, A., Noordermeer, S. M., Wilson, M. D., Landry, S., Enchev, R. I., Sherker, A., *et al.* A mechanism for the suppression of homologous recombination in G1 cells. *Nature* **528,** 422–426 (2015).

105. Maggio, I. & Gonçalves, M. A. F. V. Genome editing at the crossroads of delivery, specificity, and fidelity. *Trends Biotechnol.* (2015). doi:10.1016/j.tibtech.2015.02.011

106. Boissel, S., Jarjour, J., Astrakhan, A., Adey, A., Gouble, A., Duchateau, P., *et al.* megaTALs: a rare-cleaving nuclease architecture for therapeutic genome engineering. *Nucleic Acids Res.* **42,** 2591–601 (2014).

107. Tsai, S. Q., Wyvekens, N., Khayter, C., Foden, J. a, Thapar, V., Reyon, D., *et al.* Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat. Biotechnol.* **32,** 569–76 (2014).

108. Ran, F. A., Hsu, P. D., Lin, C.-Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., *et al.* Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154,** 1380–9 (2013).

109. 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,** (2014).

110. Sapranauskas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P. & Siksnys, V. The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. *Nucleic Acids Res.* **39,** 9275–82 (2011).

111. Hou, Z., Zhang, Y., Propson, N. E., Howden, S. E., Chu, L.-F., Sontheimer, E. J., *et al.* Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 15644–9 (2013).

112. Ran, F. A., Cong, L., Yan, W. X., Scott, D. A., Gootenberg, J. S., Kriz, A. J., *et al.* In vivo genome editing using Staphylococcus aureus Cas9. *Nature* **520,** 186–91 (2015).

113. Esvelt, K. M., Mali, P., Braff, J. L., Moosburner, M., Yaung, S. J. & Church, G. M. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat. Methods* **10,** 1116–21 (2013).

114. Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., *et al.* Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell* **163,** 759–771 (2015).

115. Tsai, S. Q., Zheng, Z., Nguyen, N. T., Liebers, M., Topkar, V. V, Thapar, V., *et al.* GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* **33,** 187–197 (2014).

116. Crosetto, N., Mitra, A., Silva, M. J., Bienko, M., Dojer, N., Wang, Q., *et al.* Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. *Nat. Methods* **10,** 361–5 (2013).

117. Tebas, P., Stein, D., Tang, W. W., Frank, I., Wang, S. Q., Lee, G., *et al.* Gene editing of

CCR5 in autologous CD4 T cells of persons infected with HIV. *N. Engl. J. Med.* **370,** 901–10 (2014).

118. Yáñez, R. J. & Porter, A. C. Therapeutic gene targeting. *Gene Ther.* **5,** 149–59 (1998).

Table 1 Therapeutic gene editing approaches applied in selected monogenic diseases

Disease	Nuclease	Gene editing strategy	Nuclease delivery route	Experimental model
Cystic Fibrosis (CF)	ZFNs	HDR-mediated cDNA knock-in (4.5 Kb)	Plasmid transfection	Human bronchial and CF tracheal epithelia[22]
		HDR of ΔF508 mutation using plasmid donor	Plasmid transfection	iPSCs[26]
	TALENs	HDR of ΔF508 mutation using short DNA fragments	Plasmid electroporation (Amaxa)	iPSCs[27]
	Cas9	HDR-mediated cDNA knock-in	Plasmid transfection	Stem cell organoids[23]
		HDR of ΔF508 mutation using *piggyBac* transposon	Plasmid electroporation (Lonza)	iPSCs[28]
Duchenne Muscular Dystrophy (DMD)	MGNs	HDR with 4.2 Kb cDNA	Lentiviral transduction	Immortalised patient myoblasts[38]
	ZFNs	Excision of exon 51 to restore the reading frame (applicable to 13% of patients)	Plasmid electroporation (Gene Pulsar X-Cell)	Immortalised patient myoblasts[40]
	TALENs	Exon 45 skipping by disruption of splice acceptor, NHEJ restoration of reading frame, HDR-mediated exon 44 cDNA knock-in	Plasmid electroporation (Neon Life Technologies)	Patient fibroblasts or iPSCs[39]
	Cas9	Excision of exons 45-55 restoring the reading frame (applicable to 62% of patients)	Plasmid electroporation (Gene Pulsar X-Cell)	Immortalised patient myoblasts[40]
		HDR using a ssODN donor	Cas9 mRNA injection	*Mdx* zygote[43]
		Exon 45 skipping by disruption of splice acceptor, NHEJ restoration of reading frame, HDR-mediated exon 44 cDNA knock-in	Plasmid electroporation (Neon Life Technologies)	Patient fibroblasts or iPSC[39]
Sickle Cell Anaemia (SCA) and β-Thalassemia	ZFNs	HDR using plasmid donor	Plasmid electroporation (Amaxa)	SCA-patient iPSCs[70]
		HDR using IDLV/ssODN donor	ZFN mRNA electroporation (Harvard apparatus)	Healthy donor and SCA-patient CD34+ cells[66]
		HDR using plasmid donor	Plasmid electroporation	β-thalassemia-patient iPSCs[67]
		NHEJ-mediated disruption of *BCL11A* enhancers for upregulation of HbF	mRNA transfection (BTX device/ MaxCyte)	Mobilized human (adult) CD34+ HSCs[74]

	TALENs	HDR-mediated full-length cDNA knock-in	Plasmid electroporation (Lonza)	K562 cell line[65]
		HDR using *piggyBac* transposon	Plasmid electroporation (Amaxa*)	β-thalassemia-patient iPSCs[68], SCA-patient iPSCs*[71]
		NHEJ-mediated disruption of *BCL11A* enhancers for upregulation of HbF	mRNA transfection (BTX device/ MaxCyte)	Mobilized human (adult) CD34+ HSCs[74]
	Cas9	HDR using *piggyBac* transposon	Plasmid electroporation (Lonza)	β-thalassemia-patient iPSCs[68,69]
		NHEJ-mediated disruption of *BCL11A* enhancers for upregulation of HbF	Lentiviral transduction	HUDEP-2 (immortalised human CD34+) and CD34+ HSPCs[75]
Primary immune deficiencies (PIDs)	ZFN	HDR-mediated cDNA knock-in at *IL2RG* using IDLV donor	IDLV ZFN transduction	K562, mouse embryonic stem cells and CD34+ cells[77]
		HDR-mediated knock-in at *AAVS1* using CGD minigene plasmid donor	ZFN mRNA (Lonza)	CGD-patient iPSCs[80]
		HDR of *Prkdc* point mutation using plasmid donor	Plasmid transfection, electroporation (Lonza)	RS-SCID mouse primary fibroblast, iPSCs[83]
	TALENs	HDR-mediated cDNA knock-in (exons 5-8) at *IL2RG* using IDLV donor	ZFN mRNA electroporation (Lonza)	Healthy/SCID-X1 donor CD34+ cells[78]
		HDR-mediated cDNA knock-in at *IL2RG* using plasmid donor	Plasmid electroporation (Nepa Gene)	Jurkat cells[79]
		HDR of a splice-site mutation in *IL2RG* using plasmid donor	Plasmid electroporation (Lonza)	SCID-X1 patient iPSCs[82]
Haemophilia	ZFN	HDR at *hF9* using AAV-8 donor	AAV-8 ZFN transduction	Humanised haemophilia B neonatal[96], adult mice[97]
		HDR-mediated insertion of *F8* and *F9* cDNA within Albumin locus using AAV-8 donor	AAV-8 ZFN transduction	Humanised haemophilia A and B adult mice[98]
	NA	HDR-mediated insertion of *F9* cDNA within Albumin locus using AAV-8 donor	*N/A*	New-born and adult haemophilia B mice[99]

	TALENs	NHEJ-mediated correction of 140 Kb inversion in *F8* gene	Plasmid electroporation	Haemophilia A-patient iPSCs[100]
	Cas9	NHEJ-mediated correction of 140 Kb and 600 Kb inversions in *F8* gene	Cas9 protein and in vitro transcribed gRNA electroporation (Neon Life Technologies)	Haemophilia A-patient iPSCs[101]

Abbreviations: *CGD (chronic granulomatous disease), HDR (homology-directed repair), IDLV (Integration-deficient lentiviral vector), iPSCs (induced pluripotent stem cells), MGNs (Meganucleases), NA (not applicable), NHEJ (non-homologous end-joining), RS-SCID (radiosensitive severe combined immunodeficiency), ssODN (single-stranded oligonucleotide), TALENs (Transcription activator-like effector nucleases), ZFNs (Zinc finger nucleases)*

Figure Legend

Figure 1

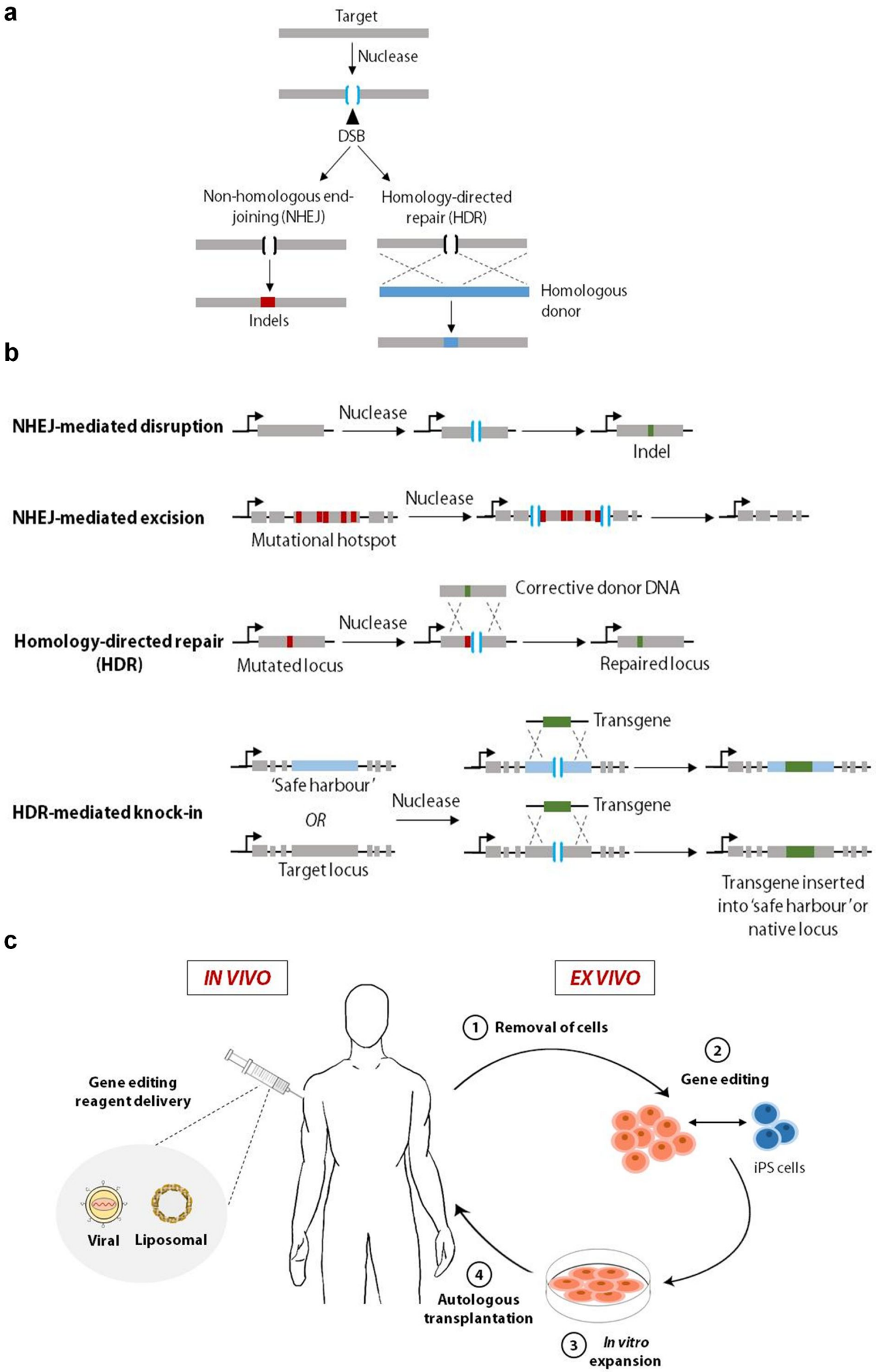
Overview of therapeutic gene editing

(a) DNA repair pathways for the resolution of a double stranded break (DSB). A nuclease is targeted towards a defined genomic locus, introducing a DSB. This may undergo one of two major repair pathways known as non-homologous end joining (NHEJ) or homology directed repair (HDR), depending upon cell cycle stage and availability of a DNA donor template. NHEJ is an error prone mechanism, which causes small insertions or deletions (Indels) upon ligating the ends of the DNA break. HDR is a precise mechanism which repairs the break by using a homologous donor template.

(b) Functional gene editing strategies using DNA repair pathways. 1) NHEJ can be used to disrupt genomic sequences as a consequence of Indels. This can cause frameshift mutations leading to an early stop codon (or restoration of the reading frame by splice site disruption). 2) NHEJ can mediate targeted deletions. This requires generation of DSBs on both sides of the target genomic sequence, which then deletes the intervening sequence while NHEJ re-joins the DNA ends. 3) HDR can be used to correct a specific mutation by introducing a nuclease-mediated DSB (in proximity to the target site) in the presence of a homologous donor DNA containing corrective sequence. Upon recombination, the repair template corrects the mutated locus. 4) Likewise, by supplying exogenous DNA on the donor template flanked between regions of homology, HDR can be used to mediate targeted gene insertion or knock-in.

(c) Schematic diagram comparing *ex vivo* and *in vivo* approaches. *In vivo* approaches involve direct transfer (denoted by the syringe) of genome editing reagents such as programmable nucleases and donor templates to the human body. In this instance, two prominent gene transfer agents, viral vectors and

liposomes are shown. *Ex-vivo* is centred on correction of the genetic defect outside of the body. This is a staged-approach whereby: 1) Patient cells are obtained. 2) Gene editing is performed *in vitro*. This involves delivery of nucleases on their own or concomitantly with repair template. The patient cells can be programmed into induced pluripotent stem cells (iPSCs) before or after gene editing. Once corrected, iPSCs may be differentiated into cell types of interest. 3) The genetically corrected cells are characterized and expanded. 4) The corrected cells are then re-grafted back into the patient through autologous transplantation.


a
Target
Nuclease
DSB
Non-homologous end-joining (NHEJ)
Homology-directed repair (HDR)
Indels
Homologous donor
b
NHEJ-mediated disruption
Nuclease
Indel
NHEJ-mediated excision
Nuclease
Mutational hotspot
Homology-directed repair (HDR)
Corrective donor DNA
Nuclease
Mutated locus
Repaired locus
HDR-mediated knock-in
Transgene
'Safe harbour'
OR
Target locus
Nuclease
Transgene
Transgene inserted into 'safe harbour' or native locus
c
IN VIVO
EX VIVO
1 Removal of cells
2 Gene editing
iPS cells
Gene editing reagent delivery
Viral
Liposomal
4 Autologous transplantation
3 In vitro expansion