Current trends in gene recovery mediated by the CRISPR-Cas system

Hyeon-Ki Jang, Beomjong Song, Gue-Ho Hwang and Sangsu Bae

Abstract
The CRISPR-Cas system has undoubtedly revolutionized the genome editing field, enabling targeted gene disruption, regulation, and recovery in a guide RNA-specific manner. In this review, we focus on currently available gene recovery strategies that use CRISPR nucleases, particularly for the treatment of genetic disorders. Through the action of DNA repair mechanisms, CRISPR-mediated DNA cleavage at a genomic target can shift the reading frame to correct abnormal frameshifts, whereas DNA cleavage at two sites, which can induce large deletions or inversions, can correct structural abnormalities in DNA. Homology-mediated or homology-independent gene recovery strategies that require donor DNAs have been developed and widely applied to precisely correct mutated sequences in genes of interest. In contrast to the DNA cleavage-mediated gene correction methods listed above, base-editing tools enable base conversion in the absence of donor DNAs. In addition, CRISPR-associated transposases have been harnessed to generate a targeted knockin, and prime editors have been developed to edit tens of nucleotides in cells. Here, we introduce currently developed gene recovery strategies and discuss the pros and cons of each.

Introduction
Human genetic disorders, often associated with severe pathological phenotypes, are caused by genomic aberrations such as gene mutations and chromosomal abnormalities. Therefore, a reliable therapeutic method for gene recovery would be quite valuable. Previously, exogenous delivery of therapeutic normal genes via viral vehicles, such as adenoviruses, adeno-associated viruses (AAVs), and lentiviruses, has been tried as a means of providing the normal function of the inactivated/disrupted gene. Although such gene therapy methods have produced successful therapeutic results, this general approach has potential limitations. For example, the exogenous gene is constitutively expressed, unaffected by the native chromatin structure of the endogenous locus, at a level that differs from that of the endogenous gene. Furthermore, the mutated endogenous gene, which is malfunctional and potentially cytotoxic, might still be transcribed.

Precise correction of the endogenous gene of interest is a strongly desirable alternative for gene recovery. Programmable nucleases, which include zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and clustered regularly interspaced short palindromic repeat (CRISPR)-CRISPR-associated (Cas) endonucleases, enable target-specific DNA cleavage and gene editing. CRISPR-mediated gene-editing technologies are now overwhelmingly the method of choice because of their ease of handling and low cost. Since CRISPR nucleases were first harnessed for generating site-specific DNA cleavage in the human genome, new CRISPR-based gene-editing tools, including base-editing technologies, have been developed. The ability to correct endogenous genes in a targeted and predictable manner using such tools has undoubtedly revolutionized gene-based drug development as well as basic research. In this review, we introduce current trends in CRISPR-mediated gene recovery.
gene correction and rescue strategies and describe the pros and cons of each tool.

**DNA repair pathways induced by CRISPR-mediated DNA cleavage in eukaryotic cells**

The type II CRISPR-Cas9 and type V CRISPR-Cas12a (also known as Cpf1) endonucleases are targeted to specific genomic sites by associated guide RNAs and can be used to generate site-specific DNA cleavage in various cell types and organisms, including humans. Typically, researchers use one piece of single guide RNA (sgRNA), which includes a spacer region complementary to the target DNA and a region that binds to the endonuclease. The target DNA sequence recognized by the guide RNA must be associated with a nuclease-specific protospacer adjacent motif (PAM), which is recognized directly by the endonuclease.

The chromosomal double-strand breaks (DSBs) produced by these nucleases are typically repaired by the cell’s own repair processes, such as the non-homologous end joining (NHEJ) pathway, the homology-directed repair (HDR) pathway, or an alternative KU-independent process such as the microhomology-mediated end joining (MMEJ) pathway. DSBs are ligated without a homologous template during the NHEJ process, which frequently leads to small nucleotide insertions and deletions (indels) at the cleavage site. In the presence of a donor DNA template, HDR precisely rejoins the DSB ends based on the donor DNA sequence, which results in precise gene corrections or knockins. The MMEJ pathway involves annealing between identical microhomologous sequences (>2 bp) flanking the DSB. Hence, MMEJ causes sequence-dependent deletions according to the microhomologous sequences that flank the cleavage site. On the basis of these various repair pathways, researchers have established precise endogenous gene recovery strategies in human cells for treating different genetic diseases.

**Gene recovery strategies in the absence of donor DNA**

**Frameshift- and deletion-mediated gene recovery involving one guide RNA**

NHEJ, a dominant repair pathway in mammalian cells that is active throughout the cell cycle, is typically used for gene disruption or knockout via induction of indels. Alternatively, however, NHEJ-mediated indels can be effectively used for genetic disease treatment if they induce a desired frameshift or delete a point mutation (Fig. 2a). For example, in Duchenne muscular dystrophy (DMD) models, premature stop codons induced by deletion of exon 44 were corrected by the introduction of Cas9-mediated frameshifting indels at a nearby location. Additionally, point mutations that lead to aberrant splicing in the DMD gene or in the Hemoglobin B (HBB) gene, which cause β-thalassemia, were removed by the introduction of Cas9- or Cas12a-mediated indels.

Similar to the NHEJ-mediated gene recovery strategy, MMEJ-mediated deletions can also be used to remove disease-causing mutations (Fig. 2a). An 8-bp duplication in exon 1 of the TCAP gene that causes limb-girdle muscular dystrophy type 2G (LGMD2G) was deleted precisely in patient-derived induced pluripotent stem cells (iPSCs) and myoblasts differentiated from the iPSCs via the MMEJ pathway. Likewise, MMEJ was utilized to remove a 16-bp microduplication in exon 15 of the HPS1 gene in B lymphocytes that causes Hermansky–Pudlak syndrome type-1 (HPS1).

In addition, several programs, such as Microhomology predictor, inDelPhi, DeepSpCas9, and DeepCpf1, have been developed to predict gene-editing efficiencies and/or editing outcomes after CRISPR treatment. These
resources should accelerate the use of frameshift- and small deletion-mediated gene recovery strategies.

Large deletion- or inversion-mediated gene recovery involving two guide RNAs

CRISPR nuclease target sites can be changed simply by altering the sgRNA sequence; multiple DNA cleavages are easily obtained by using two or more sgRNAs simultaneously. Therefore, researchers can generate a large deletion or inversion of a gene of interest by using dual sgRNAs (Fig. 2b). Because CRISPR-mediated DNA cleavage is typically accompanied by indel formation at target sites, introns may be more suitable than exons as target sites in such approaches. For example, the abnormal CTG repeat in the 3′ untranslated region of the DMPK gene that results in myotonic dystrophy type-1 was successfully excised in patient-derived iPSCs and muscle cells by treatment with Cas9 and dual sgRNAs. Additionally, mutation-carrying exons were excised by using dual sgRNAs in myoblasts derived from a DMD mouse model and in keratinocytes derived from patients with recessive dystrophic epidermolysis bullosa (RDEB); this approach restored gene function. Other diseases associated with single base substitutions, such as DMD (DMD gene), achondroplasia (FGFR3 gene), Alzheimer’s disease (PSEN2 gene), retinitis pigmentosa

Gene recovery strategies involving donor DNAs

Precise HDR-mediated gene correction

HDR-mediated gene correction is the most popular strategy for gene recovery because the genetic defect is corrected to exactly match the DNA donor template. To date, HDR-mediated gene correction has been widely harnessed to treat various genetic diseases, including sickle cell disease, β-thalassemia, hemophilia A/B, and DMD in cells or organisms. The DNA donors for HDR have been provided in various forms, including single-stranded DNA oligonucleotides (ssODNs), DNA plasmids or viral vectors, as discussed below (Fig. 3a).

For correction of one or a few mutations, an ssODN donor template is preferable because of its low cost and the relatively high associated editing efficacy. The relatively small size of ssODNs (90 ~ 200 nt) is advantageous for the synthesis and delivery of the donor template. Furthermore, ssODNs show a low level of chromosomal integration compared to double-stranded DNA oligonucleotides (dsODNs). A dominant cataract disorder in mice was the first genetic disease for which the ssODN strategy was used to correct the causative mutation, a 1-bp deletion in exon 3 of the Crygc gene that leads to the production of truncated γC-crystallin. Other diseases associated with single base substitutions, such as DMD (DMD gene), achondroplasia (FGFR3 gene), Alzheimer’s disease (PSEN2 gene), retinitis pigmentosa
(Pde6b gene)\textsuperscript{47}, sickle cell disease (HBB gene)\textsuperscript{48}, and hemophilia B (HBB gene)\textsuperscript{37}, have also been targeted with ssODN-templated gene correction. Additionally, a 4-bp deletion in the HBB gene (causing \(\beta\)-thalassemia)\textsuperscript{32,33} and a 4-bp insertion in the FANCF gene (causing Fanconi anemia)\textsuperscript{49} were corrected exactly by the ssODN-templated HDR strategy.

The ssODN strategy is convenient for gene correction at the cellular level but is not easily applied for in vivo corrections in organisms such as mice due to the lack of a reliable delivery method. For example, ssODNs cannot be carried by viral vehicles, in contrast with DNA plasmids. Hence, donor templates for in vivo applications are typically prepared as plasmids carrying the desired sequence flanked by long homology arms (~800 bp) (Fig. 3b). The donor plasmid template resembles an endogenous template (i.e., the opposite allele) in the homologous recombination process. HDR-mediated genome editing with a plasmid donor can precisely replace a long stretch of nucleotides or insert large constructs, such as a sequence encoding a tagged protein, at a specific locus. Cystic fibrosis was the first genetic disease for which HDR-mediated gene correction with a donor plasmid was performed in patient-derived intestinal stem cells carrying a homozygous 3-bp deletion in exon 11 of the CFTR gene\textsuperscript{50}. The donor plasmid contained a puromycin resistance gene in addition to the wild-type CFTR sequence for effective selection of knockin cells. Indeed, the availability of selection markers, including drug resistance genes or genes encoding fluorescent proteins, is one of the benefits of using a plasmid donor versus an ssODN donor. By using HDR with a plasmid carrying donor DNA, various mutation patterns have been corrected, including small-sized mutations in the HBB gene (associated with \(\beta\)-thalassemia)\textsuperscript{30,31,34} and the RPGR gene (associated with retinitis pigmentosa)\textsuperscript{51}, as well as large-sized mutations corrected by the precise knockin of a large DNA fragment from the FVIII gene (associated with hemophilia A)\textsuperscript{38} and from the DMD gene (associated with DMD)\textsuperscript{13}.

However, the efficiency of plasmid-templated HDR is typically lower than that of ssODN-templated HDR. To enhance the editing efficiency of plasmid-templated HDR, some modifications have been adopted. A minicircle plasmid (~1 kb), the size of which was minimized by removal of the bacterial backbone sequence, was utilized as a donor template to increase transfection efficiency\textsuperscript{52}. Additionally, linearized templates that were prepared by PCR amplification or restriction enzyme-mediated digestion were associated with increased editing efficiencies in various cell types; this method was termed targeted integration with linearized dsDNA-CRISPR (Tild-CRISPR)\textsuperscript{53}.

To maximize donor plasmid delivery into the cell nucleus, recent studies used viral vectors, including an integrase-defective lentiviral vector, adenoviral vector, or adeno-associated viral vector, as the donor template. Genetic mutations associated with SCD\textsuperscript{27–29}–29, hemophilia B\textsuperscript{35,36}, or RDEB\textsuperscript{54} were corrected by Cas9-mediated HDR using a viral vector donor. Delivery of an AAV6 donor...
along with Cas9 RNPs to SCD patient-derived hematopoietic stem and progenitor cells (HSPCs) resulted in successful targeted gene correction at 19% efficiency. The combination of Cas9 RNPs and an AAV6 donor template has been reported to be a powerful tool for gene correction in HSPCs, T cells, and iPSCs, resulting in precise editing efficiencies of ~60%. Notably, AAV donors can be used as an alternative to plasmid donors, especially in some cell types that show poor HDR editing efficiency or cytotoxicity when transfected with plasmids.

**Homology-independent gene recovery**

Despite the precision of HDR-mediated gene correction, it is not the method of choice in all circumstances. This method frequently shows low editing efficacy, and its utility is limited in non-dividing or fully differentiated cells, because HDR is active only in the late S and G2 phases of the cell cycle. In contrast, gene expression cassettes can be integrated via the NHEJ-mediated knockin method regardless of the cell cycle phase (Fig. 3b). The representative example of this method is homology-independent targeted integration (HIITI). HIITI employs donor plasmids that lack homology arms but include Cas9 cleavage sites flanking the donor sequence. Therefore, Cas9 nucleases cut both the genomic target sequence and the donor plasmid, after which the cleaved donor DNA can be incorporated into the target gene. One remarkable property of HIITI is its high accuracy even in fully differentiated cells such as neurons, which can be achieved by repeating Cas9-mediated cleavage until the donor DNA is inserted in the desired orientation. This property of HIITI makes it a promising method for curing genetic diseases by gene replacement. Royal College of Surgeons (RCS) rats, an animal model of retinitis pigmentosa caused by deletion of the Mertk gene, have morphological changes in the degenerating photoreceptor outer nuclear layer (ONL). Injection of HIITI-AAV vectors in the subretinal space in the eyes of RCS rats significantly increased Mertk mRNA levels and preserved ONL thickness. Consistent with this, the MERTK protein was observed in the eyes, and electroretinography tests showed improved eye function.

Similar to NHEJ, MMEJ can also mediate knockin of a large gene construct. In a technique known as CRISPR-PITCh, CRISPR nucleases are used to cleave both genomic and donor DNA at sites with microhomology, resulting in precise integration into the target chromosome (Fig. 3b). CRISPR-PITCh requires three sgRNAs, and the CRISPR-PITCh vector must include two different sgRNA target sites. A recent study showed the potential for CRISPR-PITCh in gene replacement therapy. Hydrodynamic injection of Cas9-expressing and Fah-MMEJ constructs into Fah−/− mice resulted in greater proliferation of normal hepatocytes than the MMEJ strategy.

**Gene recovery by base editors without DNA DSB generation**

Mutation of a single nucleotide in a gene, which can induce an amino acid substitution in the encoded protein (missense mutation) or truncation of the protein (nonsense mutation), is the main cause of genetic diseases (>58% of the entries in the ClinVar database). Although the single mutated nucleotide can be precisely repaired through HDR, the low editing efficiency and the restriction to non-dividing cells of this method obstruct its therapeutic application. In addition, recent studies have warned that DNA DSBs can lead to a p53-mediated DNA damage response and can frequently cause unexpected large deletions. Therefore, an alternative gene correction method that does not generate DNA DSBs is required. CRISPR-based base-editing tools, including cytosine and adenine base editors (CBEs and ABEs), have recently been developed for highly efficient single nucleotide correction, which occurs without donor DNA but does not generate DNA DSBs.
variant that exhibits nickase activity (nCas9) fused to a cytidine deaminase such as rAPOBEC1 or pmCDA1 and a uracil DNA glycosylase inhibitor (UGI)\textsuperscript{71,73} (Fig. 4). The CBE initially replaces a cytosine in the nontarget strand with uracil, after which the guanine in the target strand that was previously paired with cytosine is replaced with adenine by the cell’s repair mechanism. Finally, the uracil is replaced with thymine, thereby generating a T–A pair in place of the previous C–G pair\textsuperscript{71}. UGI prevents the base excision repair process that removes uracil from the nontarget strand, enhancing the C-to-T conversion efficiency.

Subsequent studies led to the development of several CBE variants with improved editing efficiency. The
editing window of BE3 was narrowed by altering the sequences linking nCas9 to the cytidine deaminase or by APOBEC1 mutation, and BE3 expression was enhanced in mammalian cells by optimizing the nCas9-encoding sequences. Adding a second UGI on BE3 (BE4) decreased the formation of undesired byproducts of editing. BE4 was further improved by adopting nuclear localization signal (NLS) sequences and optimized codons (BE4max) and by employing APOBEC homologs (AncBE4max).

CBEs show great promise for the treatment of genetic diseases because many preliminary studies have shown tangible results. For example, nucleofection of astrocytes with BE3 constructs converted the apolipoprotein E gene in human breast cancer cells (HCC1954). A nonsense mutation in the superoxide dismutase 1 (SOD1) gene was induced by BE3 in amyotrophic lateral sclerosis (ALS) model mice, prolonging survival. The HBB gene mutation that causes β-thalassemia was corrected by BE3 or its variant YEE-BE3 in skin fibroblasts, human embryos, and blastomes. The FBN1 mutation, a cause of Marfan syndrome, was repaired in human embryos using YE1-BE3 or YEE-BE3.

A successful CRISPR-based gene correction strategy requires that an appropriate PAM be located near the mutation of interest. The most commonly used Cas9 nuclease, SpCas9 from Streptococcus pyogenes, recognizes a 5′-NGG-3′ PAM downstream of the target, but such motifs are not always present. This limitation can be overcome by swapping the Cas9 nickase with other nickases preferring different PAM sequences. nSaKKH-BE3, which adopts mutated SaCas9 derived from Staphylococcus aureus, recognizes alternative PAM sequences (5′-NNNRRT-3′) and was used to correct the Fah gene mutation in the hepatocytes of phenylketonuria model mice.

**Adenine base editors (ABEs)**

Analogous to CBEs, ABEs replace adenines in the nontarget strand with inosines, generating I-T pairs. Because inosine preferentially base pairs with cytosine over thymine, DNA repair machinery converts the I-T pairs to I-C pairs and finally to G-C pairs. ABEs differ from CBEs in that they contain deoxyadenosine deaminase instead of cytidine deaminase. Because a natural deoxyadenosine deaminase is not known to exist, directed evolution for protein engineering was used to convert a transfer RNA adenosine deaminase, TadA, to a DNA deoxyadenosine deaminase, referred to as TadA*, which was then fused to a nCas9 to create AB7.10.

This initially optimized ABE was further improved by optimizing codons and adding an NLS sequence (ABEmax). Recently, various ABEmax variants, based on different Cas9 variants, were developed (VRQR-ABEmax, VRER-ABEmax, xABEmax, NG-ABEmax, SaABEmax, and SaKKH-ABEmax); these versions expand the target sites of ABEs.

Similar to the situation with CBEs, there have been many efforts to apply ABEs to gene therapy. For example, the expression of the HBG1 and HBG2 genes in adults is thought to alleviate the symptoms of β-globin-related blood diseases, and AB7.10 and ABEmax were successfully used to induce the desired mutation in the HBG1 and HBG2 promoters in HEK293T cells. The mutation in the HFE gene causing hereditary hemochromatosis was corrected by AB7.10 in an immortalized lymphoblastoid cell line. In addition, by converting a premature stop codon in DMD to a Glu codon, AB7.10 restored dystrophin expression in myofibers of DMD model mice. In HT1 model mice, targeting RA6.3 (an improved AB6.3 variant) to the Fah gene corrected the HT1-associated mutation and alleviated disease symptoms. In another example, a mutation in the COL7A1 gene causing recessive RDEB was corrected by ABEmax in primary fibroblasts and iPSCs. Furthermore, CjABE, an ABE that contains catalytically impaired Cas9 from Campylobacter jejuni, was used to correct a brain tumor-associated mutation in the TRET gene in primary glioblastomas.

**New gene recovery strategies: CRISPR-associated transposase and prime editors**

Recently, it was reported that several Tn7-like transposons are associated with CRISPR-Cas systems. This finding has led to the development of new gene-editing tools. One approach is to use a CRISPR-associated transposase from the cyanobacterium Syctonema hofmanni (ShCAST); this transposase is made up of Tn7-like transposase subunits and a type V Cas12k. The ShCAST complex is recruited to the target site in an RNA-guided manner due to the type V Cas12k and unidirectionally inserts the cargo genes of the donor plasmid into the target site via its transposase subunits. The other tool is insertion of transposable elements by guide RNA-assisted targeting (INTEGRATE) based on a type I transposon-associated CRISPR-Cas system. INTEGRATE consists of a CRISPR-associated complex for antiviral defense (Cascade) complex (Cas6, Cas7, and Cas8), which directs the editing machinery to the target site, and Tn7-like
| Disease                                | Target gene | Type of mutation | Gene recovery strategy                          |
|----------------------------------------|-------------|------------------|-------------------------------------------------|
| Achondroplasia                         | Fgfr3       | c.1120 G > A (p.G374R) | HDR (ssODN)15                                    |
| Alzheimer’s disease (AD)               | PSEN2       | c.422 A > T (p.N141I) | HDR (ssODN)16                                    |
|                                        | APOE        | ApoE4 Arg158     | CBE71                                           |
| Amyotrophic lateral sclerosis (ALS)    | SOD1        | c.281 G > C (G94A) (G93A transgenic mouse) | CBE (KO of mutant SOD1)19                       |
| β-thalassemia                          | HBB         | c.93-21 G > A or c.316-197 C > T (inducing an aberrant splice) | Indels to remove a point mutation16 |
|                                        |             | c.126_129delCTTT (CD 41/42 (-CTTT)) | HDR (ssODN)12,33,35 HDR (donor plasmid)33,34 |
|                                        |             | c.654 C > T      | HDR (donor plasmid)15                          |
|                                        |             | g.-28A > G       | CBE80                                           |
| β-globin-related disease               | HBG1        | g.-175T or g.-198T* | ABE72,78                                        |
| Cancer                                 | TPS3        | c.488 A > G (p.Y163C) | HDR (donor plasmid)51                           |
| Cataract                               | Crgc        | c.461delG (at exon 3) | HDR (ssODN or w/o donor)43b                      |
| Cystic fibrosis                        | CFTR        | c.1521_1523delCTT (p.F508del) | HDR (donor plasmid)50                           |
| Congenital disorder of glycosylation   | MPDU1       | c.356T > C (p.L119P) | CBE78                                           |
| type-1 f                               |             |                  |                                                 |
| Chronic pain                           | Scn9a       | c.689-1 C        | CBE78                                           |
| Duchenne muscular dystrophy (DMD)      | DMD         | c.2983 C > T (at exon 23, p.Q995X, mdx mouse) | Large deletion using dual sgRNAs24 HDR (ssODN)14,40 HDR (Adv donor)39 |
|                                        |             |                  | Frame-fitting indels13,14                       |
|                                        |             |                  | Indels for skipping exon 4513,14                |
|                                        |             |                  | HDR (donor plasmid)14                           |
|                                        |             |                  | Indels to disrupt a splice acceptor at exon 5115 |
|                                        |             |                  | Indels to remove the cryptic splice acceptor15  |
|                                        |             |                  | Large deletion using dual sgRNAs15 ABE85       |
| Epidermolysis bullosa simplex (EBS)    | KRT14       | c.1231 G > A (at exon 6) | HDR (donor plasmid)52                           |
| Fanconi anemia                         | FANCF       | c.828insTAAA     | CBE19                                           |
| Hemophilia A (HA)                      | FVIII       | Gross chromosomal inversion of 140-kbp or 600-kbp involving introns 1 and 22 | Large inversion using dual sgRNAs26 |
|                                        |             | Gross deletion of 94,172 bp from exon 8 to intron 22 | HDR (donor plasmid)38 |
| Hemophilia B (HB)                      | FIX         | c.1111 T > G (p.Y371D, mouse) | HDR (ssODN or donor plasmid)37 |
|                                        |             | c.1477 G > A (p.Q418G, dog) | HDR (AAV donor)35 HDR (Adv donor)35 |
| Hereditary tyrosinemia type I (HTI)    | Fah         | c.1136 G > A (p.R337Q) (R333Q transgenic mouse) | HDR (Adv donor)36 |
|                                        |             |                  | MMEJ-mediated KI63                             |
transposase subunits for cargo gene insertion into the target site. INTEGRATE differs from ShCAST in that it additionally requires the Tn7 transposition protein tnsA and can insert the cargo gene into the target site bidirectionally. Although both ShCAST and INTEGRATE have been tested only in Escherichia coli to date94,95, it is expected that they can be employed as great alternatives to NHEJ for gene knockin therapy through subsequent improvement, which will enable them to function in eukaryotic cells.

Another new approach involves the use of a prime editor (PE), which is composed of a reverse transcriptase (RT) and a Cas9 nickase, and a unique “prime editing” guide RNA (pegRNA)96 (Fig. 4). A PE that is recruited to the target site by a pegRNA generates a nick in the PAM-containing strand. Because the pegRNA has a template sequence at its 3’ end, the RT enzyme of a PE can copy the information from the template to the 3’ end of the nicked strand of DNA. This in situ synthesized donor-templated repair of DNA enables PEs to precisely induce substitutions, including transversion and transition, as well as indels, which cannot be achieved by conventional base editors71–73,96. In addition, PEs showed low off-target editing effects and generated negligible byproduct mutations, suggesting that PEs shows promise in medical applications96. Indeed, PEs effectively corrected the mutations in the HBB gene and HEXA gene in HEK293T cells, which cause sickle cell disease and Tay-Sachs disease, respectively96. Furthermore, a recent study revealed that a plant prime editor (PPE), an optimized version of PEs for plant cells, can induce various substitutions and indels efficiently in rice and wheat protoplasts97, suggesting that PEs can be applied to diverse systems, including plants and animals.

### Conclusion

In this review, we have focused on CRISPR nuclease-mediated gene recovery strategies rather than methods for

---

Table 1 continued

| Disease | Target gene | Type of mutation | Gene recovery strategy |
|---------|-------------|------------------|------------------------|
| Hermansky–Pudlak syndrome | HPS1 | c.1472_1487dup (16-bp duplication) | HDR (donor plasmid) |
| Hereditary hemochromatosis (HHC) | HFE | c.845 G > A (p.C282Y) | HDR (donor plasmid) |
| Limb-girdle muscular dystrophy (LGMD) | TCAP | c.26_33dup (8-bp duplication) | HDR (donor plasmid) |
| Marfan syndrome | FBN1 | c.7498 T > C | HDR (donor plasmid) |
| Menkes disease | ATP7A | c.6913_6917dupCTTAT | HDR (donor plasmid) |
| Myotonic dystrophy type-1 (DM1) | DMPK | CTG repeat expansion in the 3’UTR | HDR (donor plasmid) |
| Phenylketonuria (PKU) | Pah | c.835 T > C (p.F263S) | HDR (donor plasmid) |
| Retinitis pigmentosa | Pde6b | c.1041 C > A (p.Y347X) | HDR (donor plasmid) |
| | RPGR | c.1685_1686del (at exon 14) | HDR (donor plasmid) |
| | Merk | Gross deletion of 1.9 kbp from intron 1 to exon 2 | HDR (donor plasmid) |
| Recessive dystrophic epidermolysis bullosa (RDEB) | COL7A1 | c.189delG (at exon 2) | HDR (donor plasmid) |
| | | c.652T>c (at exon 80) | HDR (donor plasmid) |
| | | c.553 C > T (p.R185X) or c.1573 C > T (p.R525X) | HDR (donor plasmid) |
| Sickle cell disease (SCD) | HBB | c.20 A > T (p.E6V or p.E7V) | HDR (donor plasmid) |
| | | HDR (AAV donor) | HDR (donor plasmid) |
| Tay-Sachs disease | HEXA | c.1274_1278dup | HDR (donor plasmid) |
| Xeroderma pigmentosum, complementation group C (XPC) | XPC | c.1840C > T (p.R614X) | HDR (donor plasmid) |

Adv adeno viral vector, IDLV integrase-defective lentiviral vector, AAV adeno-associated viral vector, PTC premature termination codon.

*Inducing mutations that alleviate disease symptoms.

*Using the normal allele on the homologous chromosome as a template.
gene knockout or gene regulation such as gene inhibition (CRISPRi) or activation (CRISPRa)\textsuperscript{98}. Gene correction technologies have been rapidly developed and widely tested for the treatment of many genetic diseases (Table 1). Gene correction based on DNA cleavage in the absence of donor DNA is relatively simple if frameshifting will lead to gene recovery but is of limited value for correcting many types of mutations, such as substitutions. For such purposes, HDR-mediated gene correction is the most precise and well-characterized method, but its usefulness is limited because of its low efficiency, lack of activity in non-dividing cells, and need for donor DNA. Although NHEJ- or MMEJ-mediated gene correction enables knockin even in non-dividing cells with higher efficacy than HDR, these methods still require donor DNA and are typically less precise than HDR. Furthermore, recent studies have reported unexpected outcomes, including large deletions and chromosomal translocations, after DNA cleavage\textsuperscript{69,70}, such results warn against the use of DNA cleavage-mediated gene recovery strategies.

On the other hand, CRISPR/Cas-based base editors, which enable precise, highly efficient base conversion without requiring donor DNA or producing DNA cleavage, represent an alternative approach. However, CBEs and ABes can convert bystander bases positioned within the editing window as well as target bases, generating unwanted mutations\textsuperscript{74,75}, and, in such cases, failing to produce exact gene corrections. Furthermore, these tools are limited in the types of targeted base conversions they induce (C→T or A→T→G→C only). In addition, several recent studies reported genome-wide off-target deaminase effects on DNA (for CBEs)\textsuperscript{99,100} and RNA (for CBEs and ABES)\textsuperscript{101,102} and unwanted effects on DNA cytosines (for ABES)\textsuperscript{103}, necessitating further improvement. The recently developed CAST and INTEGRATE methods represent alternative gene knockin strategies, but their use has not yet been demonstrated in mammalian cells. PE is likewise a potentially attractive tool that requires further characterization.

In summary, all gene-editing tools are associated with some limitations or side effects, such as off-target editing, but intensive research has led to rapid improvement to compensate for these limitations. In the near future, we anticipate that several genetic diseases that lack reliable drugs will be treatable with an appropriate gene recovery strategy on the basis of the underlying genomic abnormality. Such precise endogenous gene recovery will herald a new era in human genetic medicine.

Acknowledgements
This research was supported by grants from the National Research Foundation of Korea (NRF) (no. 2018M3A9H3022412), the Next Generation BioGreen 21 Program (PJ01319301), the Korea Healthcare Technology R&D Project (HI16C1012), and the Technology Innovation Program (no. 20000158) to S.B.

Conflict of interest
The authors declare that they have no conflict of interest.

Publisher’s note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 7 April 2020 Accepted: 21 May 2020
Published online: 10 July 2020

References
1. Kotterman, M. A., Chalberg, T. W. & Schaffer, D. V. Viral vectors for gene therapy: translational and clinical outlook. Annu. Rev. Biomed. Eng. 17, 63–89 (2015).
2. Anguela, X. M. & High, K. A. Entering the modern era of gene therapy. Annu. Rev. Med. 70, 273–288 (2019).
3. 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).
4. Christian, M. et al. Targeting DNA double-strand breaks with TAL effector nucleases. Genetics 186, 757–761 (2010).
5. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. Science 339, 819–823 (2013).
6. Mali, P. et al. RNA-guided human genome engineering via Cas9. Science 339, 823–826 (2013).
7. Cho, S. W., Kim, S., Kim, J. M. & Kim, J.-S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat. Biotechnol. 31, 230–232 (2013).
8. Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816–821 (2012).
9. Zetsche, B. et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell 163, 759–771 (2015).
10. Yeh, C. D., Richardson, C. D. & Corn, J. E. Advances in genome editing through control of DNA repair pathways. Nat. Cell Biol. 21, 1468–1478 (2019).
11. Davis, A. J. & Chen, D. J. DNA double strand break repair via non-homologous end-joining. Transl. Cancer Res. 2, 130–143 (2013).
12. Guiroiulf-Barbat, J. et al. Impact of the KU80 pathway on NHEJ-induced genome rearrangements in mammalian cells. Mol. Cell. 14, 611–624 (2004).
13. Li, H. L. et al. Precise correction of the dystrophin gene in duchenne muscular dystrophy patient-induced pluripotent stem cells by TALEN and CRISPR-Cas9. Stem cell Rep. 4, 143–154 (2015).
14. Min, Y.-L. et al. CRISPR-Cas9 corrects Duchenne muscular dystrophy exon 44 deletion mutations in mice and human cells. Sci. Adv. 5, eaa4324 (2019).
15. Long, C. et al. Correction of diverse muscular dystrophy mutations in human engineered heart muscle by single-site genome editing. Sci. Adv. 4, eaap9004 (2018).
16. Xu, S. et al. Editing aberrant splice sites efficiently restores β-globin expression in β-thalassemia. Blood 133, 2255–2262 (2019).
17. Iyer, S. et al. Precise therapeutic gene correction by a simple nuclease-induced double-stranded break. Nature 568, 561–565 (2019).
18. Bae, S., Kweon, J., Kim, H. S. & Kim, J.-S. Microhomology-based choice of Cas9 nucleus target sites. Nat. Methods 11, 705–706 (2014).
19. Shen, M. W. et al. Predictable and precise template-free CRISPR editing of pathogenic variants. Nature 563, 646–651 (2018).
20. Kim, H. K. et al. SpCas9 activity prediction by DeepSpCas9, a deep learning-based model with high generalization performance. Sci. Adv. 5, eaao9249 (2019).
21. Kim, H. K. et al. Deep learning improves prediction of CRISPR–Cpf1 guide RNA activity. Nat. Biotechnol. 36, 239–241 (2018).
22. Daoudar, S. et al. Efficient CRISPR/Cas9-mediated editing of trinucleotide repeat expansion in myotonic dystrophy patient-derived iPSCs and myogenic cells. Nucleic Acids Res. 46, 8275–8280 (2018).
23. Lo Scrudato, M. et al. Genome editing of expanded CTG repeats within the human DM1 protein sequence restores nuclear RNA foci in the muscle of D1 mouse. Mol. Ther. 27, 1372–1388 (2019).
24. Xu, L. et al. CRISPR-mediated genome editing restores dystrophin expression and function in mdx mice. Mol. Ther. 24, 564–569 (2016).
25. Bonafont, J. et al. Clinically relevant correction of recessive dystrophic epidermolysis bullosa by dual sgRNA CRISPR/Cas9-mediated gene editing. Mol. Ther. 27, 986–998 (2019).
26. Park, C. Y. 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).
27. Dever, D. P. et al. CRISPR/Cas9 β-globin gene targeting in human haematopoietic stem cells. Nature 539, 384–389 (2016).
28. Hoban, M. D. et al. CRISPR/Cas9-mediated correction of the sickle mutation in human CD34+ cells. Mol. Ther. 24, 1561–1569 (2016).
29. Martin, R. M. et al. Highly efficient and marker-free genome editing of human pluripotent stem cells by CRISPR-Cas9 RNP and AAV6 donor-mediated homologous recombination. Cell Stem Cell 24, 821–828 (2019).
30. Yang, Y. et al. Naïve induced pluripotent stem cells generated from β-thalassemia fibroblasts allow efficient Gene correction with CRISPR/Cas9. Stem Cells Transl. Med. 5, 267 (2016).
31. Xu, P. et al. Both TALENs and CRISPR/Cas9 directly target the HBB homologous recombination. Hum. Mol. Genet. 29, 525–533 (2017).
32. Niu, X. et al. Combining single strand oligodeoxynucleotides and CRISPR/Cas9 to correct gene mutations in β-thalassemia-induced pluripotent stem cells. J. Biol. Chem. 291, 16576–16685 (2016).
33. Xie, F. et al. Seamless gene correction of β-thalassemia mutations in patient-specific iPSCs using CRISPR/Cas9 and piggyBac. Genome Res. 24, 1526–1533 (2014).
34. Gao, J. et al. Viral vector-based delivery of CRISPR/Cas9 and donor DNA for homology-directed repair in an in vitro model for canine Hemophilia B. Mol. Ther. Nucleic Acids. 14, 364–376 (2019).
35. Stephens, C. J. et al. Long-term correction of hemophilia B using adeno-associated delivery of CRISPR/Cas9. J. Control Release 298, 128–141 (2019).
36. Guan, Y. et al. CRISPR/Cas9-mediated somatic correction of a novel coagulation factor IX gene mutation ameliorates hemophilia in mice. EMBO Mol. Med. 8, 477–488 (2016).
37. Chen, J. X. et al. Correction of β-thalassemia mutations in patient-derived iPSCs. Exp. Mol. Med. 51, 1–9 (2019).
38. Zhao, L. et al. CRISPR/Cas9-mediated genome editing corrects dystrophin mutation in skeletal muscle stem cells in a model of muscle dystrophy. Mol. Ther. Nucleic Acids. 7, 31–41 (2017).
39. Long, C. et al. Prevention of muscular dystrophy in mice by CRISPR/Cas9-mediated editing of germline DNA. Science 345, 1184–1188 (2014).
40. Ryu, S.-M., Hur, J. W., Kim, K. Evolution of CRISPR towards accurate and efficient mammal genome engineering. BMB Rep. 52, 475–481 (2019).
41. Renaud, J.-B. et al. Improved genome editing efficiency and flexibility using modified deoxyadenosinecleotides with TALEN and CRISPR-Cas9 nucleases. Cell Rep. 14, 2263–2272 (2016).
42. Wu, Y. et al. Correction of a genetic disease in mouse viaUse of CRISPR/Cas9. Cell Stem Cell 13, 659–662 (2013).
43. Fanconi anoema defect by error-prone end joining or templated repair. Sci. Rep. 9, 66 (2019).
44. 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).
78. Koblan, L. W. et al. Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. Nat. Biotechnol. 36, 843–846 (2018).
79. Lim, C. K. W. et al. Treatment of a mouse model of ALS by in vivo base editing. Mol. Ther. https://doi.org/10.1016/j.ymthe.2020.01.005 (2020).
80. Liang, P. et al. Correction of β-thalassemia mutant by base editor in human embryos. Protein Cell 8, 811–822 (2017).
81. Arbustini, E. et al. Identification of sixty-two novel and twelve known FBN1 mutations in eighty-one unrelated probands with Marfan syndrome and other fibrillinopathies. Hum. Mutat. 26, 494–494 (2005).
82. Villiger, L. et al. Treatment of a metabolic liver disease by in vivo genome editing in human cells and heterozygous embryos. Mol. Ther. 37, 626–631 (2019).
83. Zeng, Y. et al. Correction of the marfan syndrome pathogenic FBN1 mutation by base editing in adult mice. Nat. Med. 36, 536–539 (2018).
84. Wu, Q. et al. Prime genome editing in rice and wheat. Nat. Biotechnol. 38, 582–585 (2020).
85. Pickar-Oliver, A. & Gensbach, C. A. The next generation of CRISPR-Cas technologies and applications. Nat. Rev. Mol. Cell Biol. 20, 490–507 (2019).
86. Jin, S. et al. Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. Science 364, 289–292 (2019).
87. Lee, C. et al. CRISPR-Pass: gene rescue of nonsense mutations using adenine base editors. Mol. Ther. 27, 1364–1371 (2019).
88. Peters, J. E., Makarova, K. S., Shinak, S. & Koonin, E. V. Recruitment of CRISPR-Cas systems by TrR-like transposons. Proc. Natl Acad. Sci. USA 114, E7358–E7366 (2017).
89. Faure, G. et al. CRISPR-Cas in mobile genetic elements: counter-defence and beyond. Nat. Rev. Microbiol. 17, 513–525 (2019).
90. Zhou, C. et al. Off-target RNA mutation induced by DNA base editing and its correction in an adult mouse model of Duchenne muscular dystrophy. Hum. Mutat. 40, 470–479 (2019).
91. Grünewald, J. et al. Transcriptome-wide off-target RNA editing induced by CRISPR-Cas systems in human cells and embryos. Mol. Ther. 36, 1525–1534 (2018).
92. Kim, H. S., Jeong, Y. K., Hur, J. K., Kim, J.-S. & Bar, S. Adenine base editors catalyze cytosine conversions in human cells. Nat. Biotechnol. 37, 1145–1148 (2019).