BMB
Reports

Invited Mini Review

Development of CRISPR technology for precise single-base genome editing: a brief review

Hyomin K. Lee[1],[#], Yeounsun Oh[2],[3],[#], Juyoung Hong[4], Seung Hwan Lee[5],[*] & Junho K. Hur[4],[6],[*]

[1]Department of Medicine, Graduate School, Hanyang University, Seoul 04763, [2]Futuristic Animal Resource & Research Center (FARRC), Korea Research Institute of Bioscience and Biotechnology (KRIBB), Cheongju 28116, [3]Department of Biotechnology, College of Life Sciences and Biotechnology, Korea University, Seoul 28116, [4]Graduate School of Biomedical Science & Engineering, Hanyang University, Seoul 04763, [5]National Primate Research Center (NPRC), Korea Research Institute of Bioscience and Biotechnology (KRIBB), Cheongju 28116, [6]Department of Genetics, College of Medicine, Hanyang University, Seoul 04763, Korea

The clustered regularly interspaced short palindromic repeats (CRISPR) system is a family of DNA sequences originally discovered as a type of acquired immunity in prokaryotes such as bacteria and archaea. In many CRISPR systems, the functional ribonucleoproteins (RNPs) are composed of CRISPR protein and guide RNAs. They selectively bind and cleave specific target DNAs or RNAs, based on sequences complementary to the guide RNA. The specific targeted cleavage of the nucleic acids by CRISPR has been broadly utilized in genome editing methods. In the process of genome editing of eukaryotic cells, CRISPR-mediated DNA double-strand breaks (DSB) at specific genomic loci activate the endogenous DNA repair systems and induce mutations at the target sites with high efficiencies. Two of the major endogenous DNA repair machineries are non-homologous end joining (NHEJ) and homology-directed repair (HDR). In case of DSB, the two repair pathways operate in competition, resulting in several possible outcomes including deletions, insertions, and substitutions. Due to the inherent stochasticity of DSB-based genome editing methods, it was difficult to achieve defined single-base changes without unanticipated random mutation patterns. In order to overcome the heterogeneity in DSB-mediated genome editing, novel methods have been developed to incorporate precise single-base level changes without inducing DSB. The approaches utilized catalytically compromised CRISPR in conjunction with base-modifying enzymes and DNA polymerases, to accomplish highly efficient and precise genome editing of single and multiple bases. In this review, we introduce some of the advances in single-base level CRISPR genome editing methods and their applications. [BMB Reports 2021; 54(2): 98-105]

INTRODUCTION

The recent development of target-specific genome editing technology has enabled various applications in diverse biological systems. Among the genome editing tools available, the CRISPR system has heavily contributed to improving the efficiency and accuracy of genome editing (1-3). The CRISPR modules, which were originally identified as the immune systems of bacteria and archaea, have been widely applied as genome editing tools in mammalian systems as well as plants and microorganisms (4). The natural CRISPR molecules composed of CRISPR effector proteins and guide RNAs induce double-strand breaks (DSBs) in DNA at specific sites where the target DNA sequences are recognized by base complementarity to the guide RNA (5). Based on a mechanistic perspective, double helix cleavages by the CRISPR modules induce efficient genome editing in concert with intracellular operating DNA repair systems including non-homologous end joining (NHEJ), homology-directed repair (HDR) and microhomology-mediated end joining (MMEJ) (6). The mechanisms of the intracellular repair systems are distinct, and the systems function competitively upon the incidences of DNA breaks. Therefore, the results of DNA repair could differ depending on which repair systems were utilized: NHEJ is a non-templated DNA joining process and could introduce insertions and deletions, HDR is template-dependent and uses endogenous or exogenous DNA with homology, MMEJ uses the short homologous sequences present near the DSB points to join the DNA. The CRISPR genome editing tools that function by precisely cutting the target DNA are developed in concert with the intracellular DNA repair systems to demonstrate high genome editing efficiencies in several biological systems.

Nonetheless, the accuracy and fidelity of the DSB-mediated CRISPR genome editing methods may be suboptimal for specific applications such as developing gene therapies for human diseases. A major challenge is the off-target events that introduce unintended

mutations at loci where the DNA sequences are partially complementary to the guide RNAs (7). Another concern is the variability of alterations in DNA sequences at the target sites following CRISPR genome editing (8-11). The non-uniform DNA sequences at the edited loci are caused at least in part by simultaneous action of several distinct intracellular repair systems that induce heterogenous insertions and deletions (12). Therefore, DSB-based CRISPR genome editing methods were particularly inappropriate for applications that require precise substitution of select single bases, such as correcting pathogenic single nucleotide variations (SNVs) (13). To address the problem, several methods were developed to achieve the desired genome correction via a scarless HDR pathway (6, 14-17). However, these methods still relied on DSB and therefore the approaches were unable to effectively eliminate the intrinsic challenges associated with the repair of DNA DSB (18).

In an effort to overcome the indel-by-DSB issues, recent advances in CRISPR demonstrated that precise genome editing could be conducted without DSB (19-21). Among the approaches, base editing methods adopted utilizing base-modifying enzymes in combination with CRISPR systems to substitute single bases at the target sites (19, 20). These methods enabled targeted DNA base substitution at a defined window, generally less than 10 bases, and the DNA base changes such as cytosine to thymine or adenine to guanine were controlled. The windows of base substitutions could be deliberately widened or constricted, depending on the purpose of the base editing. Another approach, called prime editing, combined CRISPR with RNA-dependent DNA polymerase to overcome the limitations of single-base substitutions by base editing (21). In this brief review, we introduce some of the recent developments and applications of DSB-free CRISPR genome editing methods that enable genome editing at single base level with enhanced accuracy.
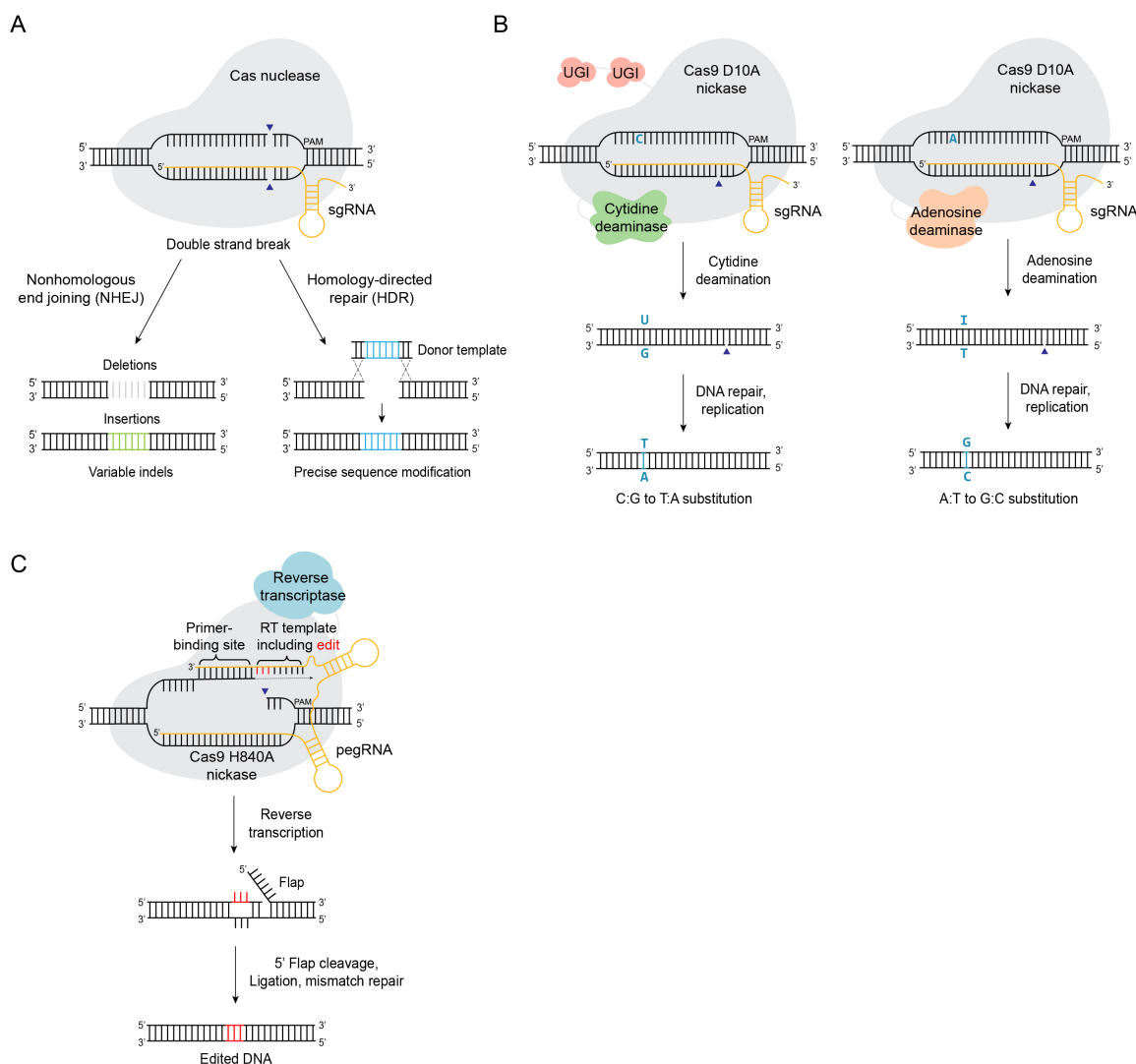
CYTOSINE BASE EDITORS ENABLE SINGLE-BASE LEVEL GENOME EDITING

The original form of CRISPR-based genome editing method induces double strand DNA breaks (DSB) as an initiation step. The DSB activates the intracellular DNA repair systems and the intended genome editing occurs during the process (Fig. 1A). The DSB repair inherently involves multiple pathways, such as NHEJ and HDR, and results in stochastic variations of repaired DNA sequences. As an alternative approach, a 'base editing' method attempted DSB-free genome editing via targeted DNA changes using a cytidine deaminase to modify cytidine to uridine (19). In the base editing method (BE1), a fusion construct that consists of *Streptococcus pyogenes* Cas9 (SpCas9) and cytidine deaminase was generated to select target sites by accompanying sgRNAs and induce targeted C to T (or G to A) conversions within a window of ∼5 nucleotides (Fig. 1B). In the initial step of base editing, the C-to-T conversions by cytidine deaminase generated intermediate wobble G:U base pairs. Analysis of the outcomes of G:U pair resolution in cells revealed

that the action of Uracil DNA glycosylase (UDG) negatively affects the base editing efficiency. UDG recognizes the intermittent G:U pairs and trigger base-excision repair (BER) to catalytically remove the uracil, resulting in reversion of G:U to G:C pair. The mechanism of BER suggests that the inhibition of UDG may increase base editing efficiencies. Consistently, fusing a Uracil DNA glycosylase inhibitor (UGI) with BE1 resulted in the second-generation base editor (BE2). As anticipated, BE2 showed ∼3-fold increase in the base editing efficiencies compared to BE1. Notably, as the steps of base editing are independent of NHEJ or HDR, low indel rates (below < 0.1%) were observed for both BE1 and BE2.

Theoretically, even in the absence of BER, the maximum efficiencies of base editing via G:U intermediate were limited to 50% as both strands can be used as templates for DNA replication. Nonetheless, a higher conversion rate could be achieved by deliberately introducing single-strand DNA breaks in the non-edited DNA strand containing the guanine base of the G:U wobble pair. The single-strand break activated mismatch repair (MMR) that actively removed the unedited guanines as they were recognized as damaged DNA. Based on the approach, the third-generation base editor (BE3) was prepared via fusion of Cas9 nickase (D10A), APOBEC1 cytidine deaminase and UGI. Accordingly, in base editing via BE3, the wobble G:U pairs are preferentially resolved to A:U (A:T) products to yield 2-6-fold and 6-18-fold higher efficiencies compared to BE2 and BE1 respectively. The highly efficient base editing by BE3 was, however, accompanied by dual DNA nicking that could induce rare but detectable undesired indel mutations. Nevertheless, the observed indel rates of BE3 were significantly lower than that of conventional CRISPR genome editing.

Analyses of base editing data showed that the undesired by-products of UDG were more prominent in target DNA sequences carrying single cytidines within the windows of base editing (22). As an effort to increase the base editing efficiencies, other cytidine deaminases were utilized: cytidine deaminase 1 (CDA1) to generate CDA-BE3, activation-induced cytidine deaminase (AID) to generate AID-BE3, and apolipoprotein B mRNA editing enzyme catalytic subunit 3G (APOBEC3G) to generate APOBEC3G-BE3 (22). Among the variants, CDA-BE3 and AID-BE3 showed higher editing efficiencies compared to BE3 at specific targets containing "GC" sequences. Based on the analyses of BE3 variants, an enhanced version of BE3, called BE4, was prepared by modulating the length of the linkers between APOBEC1, Cas9 nickase and UGI, and incorporating an additional UGI. Following the optimization, BE4 showed ∼1.5-fold increase in base editing efficiencies compared to BE3 (up to 27-fold compared to BE1). BE4 also showed ∼2-fold decrease in formation of undesired non-T products. The strategy could also be applied for enhanced base editing using *Streptococcus aureus* Cas9 (SaCas9) to develop SaBE4 (22). Also, undesired indel formation during base editing was decreased by fusing a bacteriophage Mu protein, called Gam, which binds DNA at DSB sites (22).

Fig. 1. Schematics of gene editing methods at single-base levels. (A) CRISPR genome editing via double-strand DNA break (B) cytosine base editor (CBE), and adenine base editor (ABE), (C) prime editor (PE).

The efficiency and applicability of the cytosine base editing methods was further enhanced by changing the enzyme modules, optimizing the codon usage, and modifying the nuclear localization signal sequences (23, 24). Ancestral reconstruction of the deaminase component of BE4max, an engineered base editor, resulted in AncBE4max with highly efficient base editing even with the delivery of significantly reduced levels of base editor plasmids (24). Interestingly, an altered base editing method that converts cytosine to guanine, instead of thymine, was also developed (25). The targeted C-to-G substitution was accomplished with a fusion construct composed of Cas9 nickase, a uracil DNA N-glycosylase derived from *Escherichia coli* and a variant of rat APOBEC1 cytidine deaminase (R33A).

A-TO-G AND T-TO-C BASE CONVERSIONS MEDIATED BY ADENINE BASE EDITOR EDITING

The cytosine base editors provide precise editing to convert C to T and G to A, but the method is not suitable for base conversion in the reverse direction. In order to address the issue, an adenine base editing (ABE) method was developed to enable conversion of A to G and T to C (Fig. 1B) (20). The first-generation ABE (ABE1.2) was generated by fusing CRISPR-Cas9 nickase proteins with an engineered variant of *Escherichia coli* (*E.coli*) transfer RNA adenosine deaminase (ecTadA). In its original form, the *E.coli* TadA enzymes are homologous to APOBEC and converts adenine (A) within tRNAs to inosine (I)

(26). Analyses of unbiased libraries of ecTadA showed that a variant with two point mutations (A106V and D108N), termed TadA*, exhibited DNA editing capabilities (20). A more stringent protein screen showed that a second-generation ABE (ABE2.1) with additional mutations (D147Y and E155V) into TadA* resulted in 2- to 7-fold increase in base editing efficiency compared to ABE1.2. In addition, the base editing activities of ABE1 and ABE2 were further enhanced via multiple rounds of protein evolutions. The evolution process first identified modified ABEs (ABE3, ABE4, and ABE5) with higher editing efficiencies at limited subsets of target DNA. Further iterations of protein evolution found ABE6 and ABE7 that exhibited higher activities across a broad range of target DNA sequences. The ABE7.10 variant showed A-to-G conversion rates up to 68%, which were significantly higher compared to the HDR-mediated A-to-G conversion rates of up to 10.6% (20). Subsequently, a more efficient version of ABE, ABE8e, was generated by applying phage-assisted evolution to ABE7.10 (27). In addition, optimizing the codon usage and modifying the nuclear localization signal sequences were applied to ABE7.10 to generate ABEmax (24).

The windows of base conversion by ABEs were generally 4-6 nucleotides wide, similar to cytosine base editing methods. For example, ABE7.10 showed high activities at sgRNA positions 4 to 7 (20). Notably, adenine base editing at positions upstream of the typical 4-6 nucleotide windows could be facilitated by utilizing longer sgRNAs (28). The off-target effects of ABE7.10 were significantly lower at DNA levels compared to both CRISPR nucleases and cytosine base editors (20, 29). However, analyses of RNA modification showed that ABEmax induced low but detectable adenine-to-inosine conversions in mRNA (30). Both native TadA and TadA* components in ABEmax contributed to the transcriptome-wide A-to-I RNA conversions. Analyzing the protein variants of TadA and TadA* resulted in ABEmaxAW with two point mutations (TadA E59A, TadA* V106W) that showed substantially suppressed the RNA editing, almost comparable to background level detected when Cas9 nickase alone was applied.

ENHANCEMENT OF CYTOSINE AND ADENINE BASE EDITORS FOR WIDER TARGET COVERAGE AND IMPROVED PRECISION

The target sequences of the base editor methods were constrained by the PAM sequence recognized by the CRISPR system. For instance, BE3 is a base editor system using SpCas9, and requires "NGG" PAM sequences adjacent to the 3' end of the target site. In order to overcome the PAM sequence limitations, various natural and engineered Cas9 variants were used to expand the repertoire of target sequences (31, 32). *Staphylococcus aureus* Cas9 (SaCas9) is a Cas9 homolog with "NNGRRT" PAM sequence, and its nickase form (SaCas9c) is utilized in base editing (31). The fusion construct of APOBEC1-SaCas9-UGI named SaBE3 showed 50-75% base editing efficiency in human cells. Moreover, the engineered variants of SpCas9 with extended

PAM (NGA, NGAG, NGCG) and a SaCas9 variant with "NNNRRT" PAM sequence were also used to increase the target sequences of base editing (31). From a clinical point of view, while base editing by SpCas9 alone covers 27% of pathogenic point mutations targetable via C to T (or G to A), the expansion of target sequences by the Cas9 variants increases the range to 66%. The PAM restrictions are further lowered by a SpCas9 variant with relaxed PAM constraints (32). A Cas9 variant (xCas9) that was developed via phage-assisted continuous evolution can recognize "NG, GAA, and GAT" PAM, facilitating application to wider range of target sequences. Notably, xCas9 showed less off-target activity than SpCas9 despite xCas9 potentially recognizing increased variations of off-target sequences. Applying the Cas9 variants with the "non-NGG" PAM sequences widened the range of target DNA sequences of CBE and ABEs beyond the "NGG" restriction (33). Further protein engineering of SpCas9 showed that a variant Cas9 known as SpRY can be used for base editing of almost all PAM sequences (34). As an alternative to Cas9-mediated base editing, another CRISPR family of proteins known as Cas12a that recognize T-rich PAM are used for base editing (35). Since the PAM sequences of Cas12a variants are thymine-rich without any guanine, Cas12a enables base editing of some DNA sequences that cannot be targeted by Cas9. The base-editing windows of BE and ABE are also modulated via circular permutation to the Cas9 component (33). Base editing by circularly permutated SpCas9 showed an increased editing window from ∼4-5 to ∼8-9 nucleotides.

While base editors enable genome editing without random indels, erroneous C-to-T conversions at off-target site still remain as potential safety concerns in biological and medical applications. Studies have adopted different approaches to address this issue. Applying a high-fidelity CRISPR-Cas9 (HF-Cas9) (36) to BE3 reduced off-target effects (37). HF-Cas9 is an engineered variant of SpCas9 containing four point mutations (N497A, R661A, Q695A, and Q926A) that result in decreased non-specific interactions with the phosphate backbone of DNA target strand. Base editing using HF-Cas9 (HF-BE3) demonstrated a markedly decreased off-target base editing activity in human cells. Off-target effects are further reduced by ribonucleoprotein (RNP) delivery (37). RNP delivery of BE3 and HF-BE3 resulted in editing efficiencies comparable to conventional plasmid delivery. Notably, higher on-target editing efficiencies in base editing of human cells via plasmid delivery were generally accompanied by increased off-target editing. However, base editing via RNP delivery led to efficient on-target editing without detectable off-target effects, similar to a previous study (38). Such decoupling of the linear relationship between on- and off-target editing rates facilitates RNP delivery base editing for enhanced specificity.

Precision of base editor could also be enhanced by modifying the cytidine deaminase (31, 39). In BE3, a five-base window exists, which increases the likelihood of substitution in the included cytidines. Occasionally, undesired substitutions may occur in nearby cytidines. The editing window could be modulated by inducing mutations in APOBEC1 that are involved in

substrate binding (31). Combining three amino acid mutations in APOBEC1 resulted in a base editor that induced C-to-T conversion in a window of 1-2 nucleotides. Applying a human cytidine deaminase enzyme (APOBEC3A) also generates a cytosine base editor with reduced bystander and off-target activities (39). In the study, an engineered human APOBEC3A that characteristically recognizes a "TC" motif enabled a 40-fold increase in the specificity of cytidine substitution.

BIOLOGICAL AND MEDICAL APPLICATIONS OF BASE EDITORS

Diverse applications of base editing were demonstrated in various biological systems (Table 1). Corrections of pathogenic single-base substitutionsin mammalian cells could be conducted via cytosine base editing (BE3) (19). In the study, two point mutations in APOE4 gene (C158R), associated with late-onset Alzheimer's

Table 1. A list of applications of base editing and prime editing

Gene	Target sequence	Base changes	Editor Type	Delivery method	Target organism	Efficiency (%)	Ref
Cytosine base editing							
APOE4	GAAG**C**GCCTGGCAGTGTACC<u>A</u>GG	C to T	BE3	Plasmid delivery	Mouse	74.9	19
TP53	GCTTG**C**AGATGGCCATGGCG**C**GG		BE3	Plasmid delivery	Human	7.6	
HBB promoter	CTGACTT**C**TATGCCCAGCCC<u>T</u>GG		eA3A-BE3	RNP delivery	Human	22.48	39
CTNNB1	CTGGA**C**T**C**TGGAATCCATTC<u>T</u>GG		BE3	RNP delivery	Mouse	13	40
casein	GAGTT**C**AAAGAAGGCAGGAAA<u>G</u>AG		VQR-BE3	RNA microinjection	Mouse	56	41
casein	CTT**CC**TTGTTCACACCCTTT<u>G</u>GG		BE4	RNA microinjection	Mouse	66	
OsCDC48	GA**CC**AG**CC**AGCGTCTGGCGC**C**GG		PBE (APOBEC1-XTEN-n/dCas9-UGI)	Agrobacterium-mediated transformation	Plant (rice)	1.61-8.35	42
OsNRT1.1B	CGG**C**G**A**CGGCGAGCAAGTGG<u>A</u>GG					4.58	
TPA1 (OsSPL14)	CT**C**TT**C**TGTCAACCCAGCCA<u>T</u>GG					3.51	
TaLOX2 / S1	GT**C**GA**C**AT**C**AACAACCTCGA**C**GG				Plant (Wheat)	3.74-6.90	
TaLOX2 / S2	CTT**CC**TGGGCTACACGCTCA<u>A</u>GG					3.42	
TaLOX2 / S3	AAGGA**CC**T**C**ATCCCCATGGG**C**GG					8.27-12.48	
ZmCENH3	AG**CCC**T**CC**TTGCGCTGCAAG<u>A</u>GG				Plant (maize)	0.31-4.47	
Adenine base editing							
HBG	GTGGGG**A**AGGGGCCCCCAAG<u>A</u>GG	A to G	ABE7.10	Plasmid delivery	Human	29.4-30.1	20
HFE C282Y	ACGT**A**CCAGGTGGAGCACCC<u>A</u>GG		ABE7.10	Plasmid delivery	Human	28.4	
Tyr	CC**A**TAACAGAGACTCTTACA<u>T</u>GG		ABE7.10	mRNA microinjection	Mouse	8.5-20	28
Dmd (KO)	AACT**A**GCTTTTAATTGCTGT<u>T</u>GG		ABE7.10	AAV	Mouse	3.3	
OsSPL14	AGAG**A**GAGCACAGCTCGAGT**C**GG		ABE-P1	Agrobacterium-mediated	Plant (rice)	26	43
Prime editing							
HBB (E6V)	CATGGTGCACCTGACTCCTG**T**<u>GG</u>	T to A	PE3	Plasmid delivery	Human	26-52	21
HEXA[1278+TATC]	ATCCTTCCAGTCAGGGCCATA**GGAT**AGA	4bp deletion	PE3b	Plasmid delivery	Human	33	
PRNP	CAGTGGTGGGGGGCCTTGGC**C**<u>GG</u>	G to A	PE3	Plasmid delivery	Human	53	
OsAAT	CAAGGATCCCAGCCCCGT**GA**<u>A</u>GG	GA to CC	PPE3	PEG-mediated transfection	Plant (rice)	< 0.5	46
OsALS	GTGCTGCCTATGATCCCAA**G**<u>T</u>GG	G to T	PPE3		Plant (wheat)	1.8	
OsCDC48	GACCAGCCAGCGTCTGGCGC**C**<u>GG</u>	G to A	PPE3b			2-3	
OsEPSPS	TACTAAATATACAATCCC**T**T<u>GGG</u>	T to A	PPE3			2-3	
GFP	CTCGTGACCACCCTGACCTA**C**<u>GG</u>	AC to GG	PE3	RNA delivery	Human (AAVS1-eGFP hiPS cells)	7.5	47
Hoxd13	GAGGCATACATCTCCATGGA**G**<u>GG</u>	G to C/T	PE3	mRNA microinjection	Mouse embryo	1.1-18.5	48

The table shows the target genes, target sequence with desired gene editing, and utilized genome editing technique. Also shown are the additional information such as delivery method, model systems, and the efficiencies. The list includes applications of base editor (BE), plant base editor (PBE), prime editor (PE) and plant prime editor (PPE).

disease, were corrected in mouse astrocytes with efficiencies up to 74.9%. The BE3 method was also used to correct an oncogenic point mutation in TP53 gene (Y163C) in human breast cancer cells with a rate of 7.6%. The frequeny of indels using BE3 was significantly lower than that of conventional CRISPR genome editing mediated by DSB. Using mouse models, base editing of post-mitotic cells was achieved via *in vivo* delivery of BE3 ribonucleoproteins to cochlea (40). The *in vivo* base editing installed a S33F mutation in the beta-catenin gene resulting in inhibition of protein phosphorylation and subsequent activation of Wnt signalling. In contrast, HDR-mediated genome editing of beta-catenin did not effectively induce Wnt upregulation. In another study, precision-enhanced base editing using human APOBEC3A resulted in highly specific correction of a promoter mutation causing beta-thalassemia (39). The method also showed lower off-target mutation frequencies compared to BE3. Simultaneous base editing of two linked loci separated by 9 kb was also demonstrated (41). In this study, mouse zygotes were injected with two sgRNAs targeting distinct transcription-binding sites and mRNAs encoding VQR-BE3, which recognized "NGA" PAM, and BE4. As a result, 47% of the mice carried C-to-T conversions in both sites. The results were in contrast to attempts to generate double-mutant mice via HDR genome editing, which often yielded large deletions due to two simultaneous DSB events. Application of cytosine base editing in plants was also demonstrated in rice, wheat and maize (42). To this end, a plant base editor (PBE) was prepared by fusion of catalytically dead or nickase form of CRISPR Cas9, rat APOBEC, and uracil glycosylase inhibitor. The PBE was expressed in plant cells via maize ubiquitin-1 promoter to achieve C-to-T substitution efficacy of up to 43.48%.

In addition to cytosine base editors, adenine base editors were also effectively used in several organisms. As a demonstration of a potential therapeutic approach, ABE7.10 was applied to human cells to install T > C base corrections into the promoters of HBG1 and HBG2 genes that encode fetal hemoglobin (20). The T-to-C point mutations were clinically reported to induce a benign condition called hereditary persistence of fetal hemoglobin (HPFH) that confers resistance to specific beta-globin related diseases. ABE7.10 is also utilized to correct a pathogenic point mutation associated with hereditary haemochromatosis (HHC), a genetic disorderrelated to iron storage (20). In HHC, a G-to-A mutation causes C282Y mutation in human HFE gene, which in turn results in a serious condition via excess iron absorption. Application of ABE7.10 to immortalized lymphoblastoid cell line resulted in correction of pathogenic tyrosine at position 282 to cysteine with a rate of 28%. Delivery of ABE via AAV was used to generate albino mice and in therapy for Duchenne muscular dystrophy (DMD) (28). In this study, a 2-vector split AAV delivery method of ABE efficiently corrected a pathogenic premature stop codon in a DMD mouse model. ABE was also used in plant genome editing (43). A rice genome editing system (ABE-P1) utilized a previously reported 32-amino-acid linker (20), and a VirD2 nuclear localization signal.

The protein component of the ABE-P1 system was expressed in rice via maize ubiquitin promoter, and the sgRNA was produced using rice U6 promoter. By introducing the ABE-P1 into rice via agrobacterium-mediated transformation, transgenic lines were generated with efficiencies up to 26%.

Screening applications were developed using alternative base editing methods to generate diverse libraries by deliberately installing near-random base substitutions within the target windows (44, 45). These methods utilized activation-induced cytidine deaminase (AID) enzymes to induce base substitutions at the target sites with only little bias towards C-to-T and G-to-A. A targeted AID-mediated mutagenesis (TAM) method used a fusion construct of dCas9 and human activation-induced cytidine deaminase (AID) involved in somatic hypermutation (44). Another method called CRISPR-X utilized an engineered and truncated variant of AID protein (AID*Δ) fused to MS2 proteins (45). In this method, the AID*Δ proteins were localized to the target loci by the fused MS2 protein recognizing the MS2 RNA hairpins, which were inserted into the sgRNA sequences. The CRISPR-X method demonstrated induction of near-random DNA substitutions within a wide window of −50 to +50 bp positions relative to PAM.

PRIME EDITING LOWERS THE RESTRICTIONS OF GENOME EDITING

While advances in base editing provide a wide repertoire of single-base editing techniques, the base editors are capable of changing bases in a predetermined direction. Therefore, application of base editors are unsuitable for mutagenesis of multiple bases within defined sequences that requires concomitant combinations of C-to-T and A-to-G conversions. To address the issue, an alternative DSB-free editing method, known as prime editing, was developed (21, 46-48). Although prime editing is similar to base editing in that no DSBs are involved, a distinct molecular mechanism is involved (Fig. 1C). Prime editing methods utilize fusion constructs that are composed of reverse transcriptases (RT) and Cas9 nickase proteins. In prime editing, elongated guide RNAs called prime editing guide RNAs (pegRNA) play a dual role as both sgRNAs for target sequence recognition and RNA template for reverse transcription by RT. In terms of achieving desired DNA sequences beyond single bases, prime editing scores over base editing in that transversion changes (A to C, T or G to C, T) in DNA sequences can be induced by designing the pegRNA sequences (21).

The process of prime editing occurs in three steps including DNA nick, DNA polymerization, and repair (21). First, the Cas9nickase (H840A) within the prime editing fusion protein recognizes the target DNA and introduce a single-strand break at the non-target DNA strand at the designated locus. Next, the 3′ end of the prime editing guide RNA (pegRNA) containing ~13 nucleotides, with sequence complementarity to the nicked DNA strand, invades the target DNA and forms RNA-DNA heteroduplex. The RNA-DNA hybrid then serves as a template

for DNA polymerization via reverse transcriptase derived from Moloney murine leukemia virus (M-MLV). The extended DNA fragment contains the desired mutant sequences that were designed in the RT template region of pegRNA. Finally, the DNA repair process incorporates the 3'-end DNA flaps with the desired mutant sequences into the genomic DNA.

The original prime editing scheme, called PE1, enabled precise genome editing with moderate efficiency. In PE2, modifying a number of amino acid residues within the reverse transcriptase increased the genome editing efficiency. The efficiencies of prime editing are further increased in PE3 or PE3b versions by installing additional nicks at the non-edited DNA strand near the prime editing target sites. The nicks at the non-edited strand facilitate the intracellular DNA repair system (base excision repair) to preferentially incorporate the newly synthesized mutant DNA flaps into the genomic DNA. As a result, prime editing enables multiple-base mutations as continuous stretches of DNA sequences as the polymerase-based method incorporates consecutive DNA bases, and the range is not restricted by the editing windows.

Application of prime editing was demonstrated in several biological systems. In human cells, prime editing facilitated conversion of multiple consecutive bases in genomic DNA (21). Primer editing of plant systems (PPE) was demonstrated in rice and wheat (46). The delivery of primer editing molecules in mRNA forms was also shown in human iPS cells (47) and in mouse embryos (48).

CONCLUSION

Recent DSB-free genome editing methods have improved the accuracy of genome editing compared to conventional techniques. Base editing methods utilize novel approaches and open new possibilities via precise base-by-base corrections. Precise single-base genome editing is particularly useful in addressing point mutations that are associated with phenotypic outcomes. Along with advances in base editing technologies, improved tools for bioinformatic analyses are also being developed (49). Prime editing provides precise and versatile genome editing tools with virtually no constraints in inducing the desired sequence changes: single or multiple base substitutions and defined indels with low rates of NHEJ-mediated random mutations. Analyses showed that prime editing resulted in efficient genome editing outcomes with low rates of unintended indels at the on-target and off-target loci. Notably, prime editing has been shown to result in successful genome editing of relatively short stretches of DNAs, measuring less than 100 bp. In some cases, such as transgene insertions, DNA fragments of several thousand bps are required. Currently, inserting large DNAs have been often conducted by delivering donor DNA and inducing double strand DNA to incorporate the donor DNA into genomic DNA via HDR. However, the DNA insertion by HDR is somewhat less efficient and is prone to unanticipated mutations at the DNA cleavage sites. Hence, it would be of interest to assess whether the efficiencies and precision of prime editing for large DNA with that of conventional HDR-mediated genome editing. In summary, the advances in base-level CRISPR technologies have facilitated unprecedented accuracy and freedom of genome editing that are anticipated to widen the scope of applications in biology and medicine.

ACKNOWLEDGEMENTS

CONFLICTS OF INTEREST

The authors have no conflicting interests.

REFERENCES

1. Kim H and Kim JS (2014) A guide to genome engineering with programmable nucleases. Nat Rev Genet 15, 321-334
2. Lee SH, Park YH, Jin YB, Kim SU and Hur JK (2020) CRISPR diagnosis and therapeutics with single base pair precision. Trends Mol Med 26, 337-350
3. Wright AV, Nunez JK and Doudna JA (2016) Biology and applications of CRISPR systems: harnessing nature's toolbox for genome engineering. Cell 164, 29-44
4. Hsu PD, Lander ES and Zhang F (2014) Development and applications of CRISPR-Cas9 for genome engineering. Cell 157, 1262-1278
5. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA and Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816-821
6. Lee SH, Kim S and Hur JK (2018) CRISPR and Target-specific DNA endonucleases for efficient DNA knock-in in eukaryotic genomes. Mol Cells 41, 943-952
7. Tsai SQ and Joung JK (2016) Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases. Nat Rev Genet 17, 300-312
8. Yen ST, Zhang M, Deng JM et al (2014) Somatic mosaicism and allele complexity induced by CRISPR/Cas9 RNA injections in mouse zygotes. Dev Biol 393, 3-9
9. Midic U, Hung PH, Vincent KA et al (2017) Quantitative assessment of timing, efficiency, specificity and genetic mosaicism of CRISPR/Cas9-mediated gene editing of hemoglobin beta gene in rhesus monkey embryos. Hum Mol Genet 26, 2678-2689
10. Lamas-Toranzo I, Galiano-Cogolludo B, Cornudella-Ardiaca F, Cobos-Figueroa J, Ousinde O and Bermejo-Alvarez P (2019) Strategies to reduce genetic mosaicism following CRISPR-mediated genome edition in bovine embryos. Sci Rep 9, 14900
11. Mehravar M, Shirazi A, Nazari M and Banan M (2019)

Mosaicism in CRISPR/Cas9-mediated genome editing. Dev Biol 445, 156-162

12. Yeh CD, Richardson CD and Corn JE (2019) Advances in genome editing through control of DNA repair pathways. Nat Cell Biol 21, 1468-1478

13. Landrum MJ, Lee JM, Benson M et al (2016) ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res 44, D862-868

14. Rees HA, Yeh WH and Liu DR (2019) Development of hRad51-Cas9 nickase fusions that mediate HDR without double-stranded breaks. Nat Commun 10, 2212

15. Ling X, Xie B, Gao X et al (2020) Improving the efficiency of precise genome editing with site-specific Cas9-oligonucleotide conjugates. Sci Adv 6, eaaz0051

16. Yu Y, Guo Y, Tian Q et al (2020) An efficient gene knock-n strategy using 5'odified double-stranded DNA donors with short homology arms. Nat Chem Biol 16, 387-390

17. Nguyen DN, Roth TL, Li PJ et al (2020) Polymer-stabilized Cas9 nanoparticles and modified repair templates increase genome editing efficiency. Nat Biotechnol 38, 44-49

18. Egli D, Zuccaro MV, Kosicki M, Church GM, Bradley A and Jasin M (2018) Inter-homologue repair in fertilized human eggs? Nature 560, E5-E7

19. Komor AC, Kim YB, Packer MS, Zuris JA and Liu DR (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature 533, 420-424

20. Gaudelli NM, Komor AC, Rees HA et al (2017) Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. Nature 551, 464-471

21. Anzalone AV, Randolph PB, Davis JR et al (2019) Search-nd-replace genome editing without double-strand breaks or donor DNA. Nature 576, 149-157

22. Komor AC, Zhao KT, Packer MS et al (2017) Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity. Sci Adv 3, eaao4774

23. Zafra MP, Schatoff EM, Katti A et al (2018) Optimized base editors enable efficient editing in cells, organoids and mice. Nat Biotechnol 36, 888-893

24. Koblan LW, Doman JL, Wilson C et al (2018) Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. Nat Biotechnol 36, 843-846

25. Kurt IC, Zhou R, Iyer S et al (2021) CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells. Nat Biotechnol 39, 41-46

26. Kim J, Malashkevich V, Roday S, Lisbin M, Schramm VL and Almo SC (2006) Structural and kinetic characterization of Escherichia coli TadA, the wobble-specific tRNA deaminase. Biochemistry 45, 6407-6416

27. Richter MF, Zhao KT, Eton E et al (2020) Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. Nat Biotechnol 38, 883-891

28. Ryu SM, Koo T, Kim K et al (2018) Adenine base editing in mouse embryos and an adult mouse model of Duchenne muscular dystrophy. Nat Biotechnol 36, 536-539

29. Lee HK, Willi M, Miller SM et al (2018) Targeting fidelity of adenine and cytosine base editors in mouse embryos. Nat Commun 9, 4804

30. Rees HA, Wilson C, Doman JL and Liu DR (2019) Analysis and minimization of cellular RNA editing by DNA adenine

31. Kim YB, Komor AC, Levy JM, Packer MS, Zhao KT and Liu DR (2017) Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. Nat Biotechnol 35, 371-376

32. Hu JH, Miller SM, Geurts MH et al (2018) Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. Nature 556, 57-63

33. Huang TP, Zhao KT, Miller SM et al (2019) Circularly permuted and PAM-modified Cas9 variants broaden the targeting scope of base editors. Nat Biotechnol 37, 626-631

34. Walton RT, Christie KA, Whittaker MN and Kleinstiver BP (2020) Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. Science 368, 290-296

35. Li X, Wang Y, Liu Y et al (2018) Base editing with a Cpf1-cytidine deaminase fusion. Nat Biotechnol 36, 324-327

36. Kleinstiver BP, Pattanayak V, Prew MS et al (2016) High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. Nature 529, 490-495

37. Rees HA, Komor AC, Yeh WH et al (2017) Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery. Nat Commun 8, 15790

38. Kim S, Kim D, Cho SW, Kim J and Kim JS (2014) Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. Genome Res 24, 1012-1019

39. Gehrke JM, Cervantes O, Clement MK et al (2018) An APOBEC3A-Cas9 base editor with minimized bystander and off-target activities. Nat Biotechnol 36, 977-982

40. Yeh WH, Chiang H, Rees HA, Edge ASB and Liu DR (2018) In vivo base editing of post-mitotic sensory cells. Nat Commun 9, 2184

41. Lee HK, Willi M, Smith HE et al (2019) Simultaneous targeting of linked loci in mouse embryos using base editing. Sci Rep 9, 1662

42. Zong Y, Wang Y, Li C et al (2017) Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion. Nat Biotechnol 35, 438-440

43. Hua K, Tao X, Yuan F, Wang D and Zhu JK (2018) Precise A.T to G.C base editing in the rice genome. Mol Plant 11, 627-630

44. Ma Y, Zhang J, Yin W, Zhang Z, Song Y and Chang X (2016) Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in mammalian cells. Nat Methods 13, 1029-1035

45. Hess GT, Fresard L, Han K et al (2016) Directed evolution using dCas9-targeted somatic hypermutation in mammalian cells. Nat Methods 13, 1036-1042

46. Lin Q, Zong Y, Xue C et al (2020) Prime genome editing in rice and wheat. Nat Biotechnol 38, 582-585

47. Surun D, Schneider A, Mircetic J et al (2020) Efficient generation and correction of mutations in human iPS cells utilizing mRNAs of CRISPR base editors and prime editors. Genes (Basel) 11, 511

48. Liu Y, Li X, He S et al (2020) Efficient generation of mouse models with the prime editing system. Cell Discov 6, 27

49. Clement K, Rees H, Canver MC et al (2019) CRISPResso2 provides accurate and rapid genome editing sequence analysis. Nat Biotechnol 37, 224-226