Development and Application of Base Editors

Bei Yang,1 Li Yang,2,3 and Jia Chen3,4

Abstract
Base editing is emerging as a potent new strategy to achieve precise gene editing. By combining different nucleobase deaminases with Cas9 or Cpf1 proteins, several base editors have recently been developed to achieve targeted base conversions in different genomic contexts. Importantly, base editors have been successfully applied in animals, plants, and bacteria to induce precise substitutions at the single-base level with high efficiency. In this review, we summarize recent progress in the development and application of base editors and discuss some of the future directions of the technology.

Introduction
Genome editing is a cohort of genetic engineering technologies to insert, delete, or modify the sequences of genome in a living organism.1–3 Utilizing genome editing tools to genetically manipulate the genomic information of cells and living organisms has broad applications in life sciences research, development of biotechnology and agricultural technology, and pharmaceutical and clinical innovation and therapeutics.

Early genome editing tools mainly involved zinc finger nucleases and transcription activator-like effector nucleases.1 Although these early gene editing tools enabled researchers to manipulate genomes programmably, their DNA targeting relies on protein–DNA interactions. For its high efficiency, convenience, and broad application in a vast array of living organisms, the CRISPR-Cas system has been a powerful genome editing tool since its conception.4–6 Directed by a guide RNA (gRNA), a Cas nuclease can generate DNA double-strand breaks (DSBs) at targeted genomic sites. These DSBs are then repaired by the endogenous DNA repair system, which could be employed to perform desired genome editing. In general, two major DNA repair pathways can be activated by DSBs: nonhomologous end joining (NHEJ) and homology-directed repair (HDR).7

NHEJ can introduce random insertions or deletions (indels) in the genomic DNA regions surrounding these DSBs, thereby leading to open reading frame shifts and ultimately gene inactivation. In contrast, when HDR is triggered, the genomic DNA sequence at the target site can be replaced by the sequence of the exogenous donor DNA, resulting in precise editing and in principle the correction of genetic mutations. Although NHEJ-mediated gene knockouts are highly efficient, the efficiency of HDR-mediated precise editing in practice is generally low because the occurrence of homologous recombination requires more complicated machinery and is cell cycle-dependent. As a consequence, NHEJ is triggered much more frequently than HDR.5

Base editing is a recently developed gene-editing system that has been successfully applied in many species to induce targeted base substitutions in DNA and RNA with high precision and efficiency. By combining the APOBEC (apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like)/AID (activation-induced cytidine deaminase) family of cytidine deaminases9,10 with the CRISPR-Cas gene editing system,5,6 Alexis Komor and David Liu first developed a series of cytosine base editors (BEs/CBEs) to induce targeted base editing. Linking different CRISPR-Cas proteins with different nucleobase deaminases further leads to a variety of CBEs or adenine base editors (ABEs), the latter developed first by Nicole Gaudelli in the Liu lab, which achieve conversions of...
cytosine to thymine (C to T)\textsuperscript{11} or adenine to guanine (A to G)\textsuperscript{12} in various cells, animals, plants, and bacteria.

Here, we discuss the early studies related to base editing, the development and the application of base editors, the comparison of base editing with Cas-mediated gene editing and the potential directions of base editing.

**Unexpected Link Between APOBEC and CRISPR-Cas**

Even before the development of BE, it was known that APOBECs induce C-to-T mutations in complementary DNA that is reverse transcribed from viral RNA genome to restrict viral activity.\textsuperscript{13,14} Recent studies also indicate that APOBEC/AID can actually cause a broader array of mutations, which involves not only the genome of large DNA viruses\textsuperscript{15} but also genomic DNA during replication,\textsuperscript{16,17} repair,\textsuperscript{18–21} carcinogenesis,\textsuperscript{22–24} and antibody diversification.\textsuperscript{25–27} As cytidine deaminases, APOBECs catalyze the deamination of cytosines in single-stranded (ss) nucleic acids including ssRNA and ssDNA.\textsuperscript{9,28} Notably, most APOBECs prefer to deaminate the C in TpC dinucleotides except for APOBEC3G, which prefers the latter C of CpC.\textsuperscript{10,29}

Indeed, a linkage between APOBECs and CRISPR-Cas9–mediated gene editing was appreciated in early studies that were attempting to reduce off-target (OT) indel formation. To reduce the unwanted DSBs and indels at OT sites, a double-nicking strategy was exploited by using a pair of gRNA and either Cas9 nickase (nCas9) or the catalytically dead Cas9 (dCas9) fused with a single FokI nuclease domain.\textsuperscript{30,31} Theoretically, the DNA single-strand break (SSB) at a particular OT site would be readily sealed by DNA ligases and thus not create indels. Nevertheless, unwanted indels were still induced by nCas9 monomer at some OT sites, albeit at lower frequency.\textsuperscript{32,33} Surprisingly, nCas9 monomer was also found to induce unexpected C-to-T base substitutions at on-target sites.\textsuperscript{34} By analyzing the sequence context, Tsai et al. found that most of the mutated cytosines were in TpC dinucleotides, manifesting a typical APOBEC mutational signature. This suggested that APOBECs might be involved in nCas9-triggered mutagenic process\textsuperscript{34} (Fig. 1A).

Around the same time, Chen et al. independently found that endogenous human APOBEC family members, including hAPOBEC3B (hA3B), hAPOBEC3C (hA3C) and hAPOBEC3F (hA3F), can induce C-to-T base substitutions during the repair of a preexisting DNA SSB in an episomal shuttle vector\textsuperscript{20} (Fig. 1B). These early studies implied that endogenous APOBECs were likely involved in the DNA repair process triggered by the nCas9-generated SSB and responsible for the base substitutions observed therein (Fig. 1). Notably, such notion was later proved by Lei et al. experimentally\textsuperscript{35} (see below).
Development of Cytosine Base Editing

Editing efficiency

In a landmark series of experiments that began by fusing rat APOBEC1 (rA1) and dCas9, Komor et al. developed the first generation of base editor (BE1). Although BE1 can induce efficient C-to-T base editing in vitro, it yielded only low levels of base substitution in mammalian cells. As uracil DNA glycosylase (UDG) recognizes and removes uracils from genomic DNA, Komor et al. fused a uracil DNA glycosylase inhibitor (UGI) into BE1 to inhibit endogenous base excision repair (BER) at the target site, producing a second-generation base editor (BE2), although the editing efficiency of BE2 was still relatively low at most loci.

To further improve the editing efficiency, Komor et al. replaced the dCas9 with nCas9 (D10A), which nicks the gRNA-complementary DNA strand (target strand, T-strand), to develop the third generation of base editor (BE3) (Fig. 2A). BE3 generates a U/G mismatch with a flanking nick in the T-strand, which serves as a preferred substrate for cellular mismatch repair (MMR). Being an endogenous DNA repair system, MMR recognizes mismatched bases during DNA replication and then excises the newly synthesized DNA strands according to the existence of nicks. Thus, MMR can also recognize the U/G mismatch and the flanking nick generated by BE3 in the T-strand, and excise the T-strand that contains the G of U/G mismatch. As a consequence, subsequent DNA re-synthesis will use the remaining U-containing nontarget strand (NT-strand) as a template to install a U/A pair, which will be converted to a T/A pair after DNA replication or repair. Hence, by taking advantage of the endogenous MMR system, BE3 leads to C-to-T base substitution at higher frequency than BE2 in mammalian cells.

As another cellular DNA repair system, BER starts by removing damaged bases with DNA glycosylases. Although one copy of UGI is fused in BE3 to prevent the abasic site (AP site) formation catalyzed by UDG, unintended by-products derived from AP sites (e.g., C-to-A or C-to-G conversions) were still detected at some loci. These results suggested that the U in the editing intermediate (U/G pair) can still be excised by UDG somehow, thereby leading to compromised C-to-T editing efficiency. To keep the U in place, Komor et al. fused another copy of UGI into BE3 to develop BE4, which manifested higher editing efficiency than BE3.

In an alternative approach, Wang et al. fused 2A (self-cleaving peptide)-UGI sequences into BE3 to develop an enhanced base editor (eBE) (Fig. 2B), which can express free UGI to intensively inhibit BER and thus improve base editing efficiency. Of note, while some studies showed that expressing free UGI could increase the mutation frequency of a shuttle vector plasmid in thymine DNA glycosylase–deficient cells and shift the mutational pattern of a hypermutating chicken cell line, DT40, free UGI expression was reported not to generate either spontaneous or induced mutations in the mitochondrial DNA of human cells. Considering the redundancy of mammalian DNA repair system for uracil repair, for example, MMR or alternative uracil glycosylases that are insensitive to UGI (such as thymine DNA glycosylase), whether expressing free UGI increases random C to T mutations in the genome of normal mammalian cells needs to be further investigated.

More recently, Koblan et al. and Zafra et al. sought to increase the editing efficiency of base editors by optimizing the codon of BEs to improve their expression in mammalian cells. These codon-optimized BEs (e.g., BE4max) appreciably increased base editing frequency (~1.7- to 9-fold) especially when transfection efficiency is limited. Furthermore, Koblan et al. used ancestral sequence reconstruction to develop AncBE4max, which showed even higher editing efficiency than BE4max at some loci.
DNA modification is another factor that might restrain base editing efficiency. DNA methylation at CpG sites greatly suppresses the cytosine deamination catalyzed by mouse APOBEC1 (mA1). By modulating local DNA methylation, Wang et al. showed that CpG methylation indeed has a generally negative effect on the C-to-T editing efficiency mediated by rA1-based BE3. To develop a BE that can efficiently induce base editing in highly-methylated regions, Wang et al. screened a dozen of BEs, each containing a distinct APOBEC/AID cytidine deaminase family member. Among those tested, the human hA3A-derived BE (hA3A-BE3) induced the highest editing efficiency,48 which is consistent with reports that hA3A can catalyze the deamination of methylated cytosine efficiently.49,50 Through fusing three copies of 2A-UGI sequences to hA3A-BE, Wang et al. also developed hA3A-eBE to further enhance base editing efficiency.48

Thus, via manipulating cellular DNA repair systems, codon optimization, ancestral-sequence reconstruction and screening different APOBECs family members, the efficiency of CBEs has been significantly improved in various contexts in the relatively short time since their initial description.

Targeting scope
In base editing, Cas9 or Cpf1 (also known as Cas12a) proteins generally work as locators to guide cytidine deaminase family members. Among those tested, the human hA3A-derived BE (hA3A-BE3) induced the highest editing efficiency,48 which is consistent with reports that hA3A can catalyze the deamination of methylated cytosine efficiently.49,50 Through fusing three copies of 2A-UGI sequences to hA3A-BE, Wang et al. also developed hA3A-eBE to further enhance base editing efficiency.48

In parallel, Kim et al. replaced the SpCas9 in BE3 with engineered SpCas9 proteins that recognize altered PAM sequences, Staphylococcus aureus Cas9 (SaCas9) or an engineered SaCas9 with altered PAM specificity to develop VQR-BE3, EQR-BE3, VRER-BE3, SaBE3, and SaKKH-BE3.61 These new BEs bypass the requirement of an NGG PAM sequence. Recently, Hu et al. took advantage of phage-assisted continuous evolution (PACE)62 to evolve a SpCas9 variant (xCas9) that can recognize a much broader range of PAM sequences than wild-type SpCas9.63 Alternatively, Nishimasu et al. used rational design to engineer SpCas9 and also obtained a variant (SpCas9-NG) that recognizes an NG instead of an NGG PAM sequence.64 xCas9 and SpCas9-NG were then respectively fused with rA1 and AID to induce base editing at expanded target sites.63,64 Hence, by making use of Cas9 and Cpf1 proteins that exhibit different PAM preferences, the editing repertoire of BEs have been greatly expanded.

Editing precision
Cas9 and Cpf1 proteins are guided by gRNAs to the appropriate target sites. However, the binding or editing at OT sites, where the sequences are similar to that of the on-target (ON) site, is always one of the major concerns in CRISPR-related research.65,66 The OT effects of Cas9 or Cpf1 have been determined in many studies.55,67,68 Similar to Cas9 and Cpf1 nucleases, BEs are also reported to cause OT effects, which raise concerns about the precision of BEs.69–72 Interestingly, in a genome-wide study about the OT effects of BE, Kim et al. reported that Cas9 and the Cas9-driven BE did not edit the same OT sites,73 suggesting that the Cas9-mediated DNA double-strand cleavage and the BE-mediated cytidine deamination may exhibit a different tolerance of the mismatches between gRNA and target genomic DNA. By fusing rA1 with high fidelity (HF)-Cas9, which has an improved targeting specificity, Rees et al. developed HF-BE3 to reduce the unwanted C-to-T mutations at OT sites.70 In the same study, the delivery of ribonucleoprotein (RNP) complex instead of plasmid DNA further decreased the unwanted mutations at OT sites while maintaining the editing efficiency at ON site.

In addition to editing at OT sites, indels and non-C-to-T conversions (i.e., C to A or C to G) at ON sites are also unexpected byproducts of base editing. Replacing dCas9 with Cas9 nickase D10A significantly enhanced the editing efficiency, yet also increased the formation of unintended indels and non-C-to-T conversions.71,74 During BE3-induced editing, MMR was triggered to resect the nicked T-strand, thereby leaving the ‘U’-containing NT-strand as ssDNA (Fig. 3). The cleavage of T-strand will
lead to the disassociation of BE3 and thus the BE3-fused UGI from the target site, which in turn renders the ‘U’ in the NT-strand accessible by UDG. Then endogenous UDG will be able to excise the ‘U’ and generate an AP site in the single-stranded NT-strand. The cleavage by AP endonuclease or a spontaneous breakage at AP site of the NT-strand will produce a DSB and induce indels around ON site35 (Fig. 3). Alternatively, copying AP site by translesion DNA synthesis polymerase75,76 during the resynthesis of T-strand will result in non-C-to-T conversions (Fig. 3). Through fusing 2A-UGI sequences to the C-terminus of BE3, Wang et al. developed eBE, which expresses free UGI in addition to the BE3-fused one, to induce higher editing efficiency. Moreover, eBE also reduced the formation of unwanted indels and non-C-to-T conversions, owing to its better preservation of the ‘U’ in NT-strand. Because of a similar mechanism, dCpf1-eBE (Fig. 2D) also induces purer editing products than does dCpf1-BE.59

**Editing window**
The editing window is the range of nucleotides in the gRNA target region, wherein all the cytosines have the chance in theory to be converted to thymines. If an editing window is too big, multiple cytosines in the same editing window would be edited simultaneously, resulting in the generation of undesired “bystander” byproducts.77 Hence, the editing window is an important parameter to consider when the goal is to edit a particular cytosine by BE.

---

**FIG. 3.** Unintended products of nCas9-containing cytosine base editors. The pathways by which the third generation of base editor generates indels (left) and non-C-to-T substitutions (right) are shown.
In order to narrow the 5-bp editing window of BE3, Kim et al. engineered rA1 by introducing amino acid changes in the deaminase domain to develop YE1-BE3, YE2-BE3, EE-BE3, and YEE-BE3, all of which showed constricted editing windows compared to BE3.64 Using the same strategy, Li et al. developed dCpf1-BE-YE and dCpf1-BE-YEE, which manifested narrower editing windows as well (albeit compromising editing efficiency).59 After analyzing the structural information of hA3A, Wang et al. substituted amino acid changes into the substrate recognition domain of hA3A to develop hA3A-BE3-Y130F and hA3A-BE3-Y132D, both of which showed narrowed editing windows and retained high editing efficiency.60 Recently, Tan et al. truncated the C-terminus of sea lamprey cytidine deaminase (pmCDA1) and changed the length of rigid linkers between pmCDA1 and nCas9, which leads to new BEs with narrowed editing windows.78 Alternatively, Gehrke et al. screened several hA3A mutants and developed A3A-BE3-N57G (engineered A3A-BE3, eA3A-BE3) to stringently edit the C within a context of TpC dinucleotides.79

Although considerable effort has been made to narrow the editing window, a larger window is sometimes desirable, especially for applications such as making premature stop codons to knock out genes (iSTOP or CRISPR-STOP)80,81 or identifying the effects of new point mutations.82 Ma et al. fused AIDx (an engineered AID with enhanced activity) with dCas9 to develop targeted AID-mediated mutagenesis,83 which leads to an editing window of around 25 bp and has been effectively applied to modulate RNA splicing.84 Meanwhile, Hess et al. used the MS2 loop-containing gRNAs to recruit multiple copies of AID*Δ (a hyperactive AID mutant) fused with MS2-loop binding protein, and the resulting CRISPR-X expands the editing window to around 100 bp.85 Moreover, Jiang et al. fused SunTag with dCas9 to recruit multiple copies of APOBEC-UGI fusion proteins to develop BE-PLUS,86 which has an editing window around 13 bp. Recently, Liu et al. also took advantage of MS2 loop–containing gRNA to recruit stable monomers of AID or hA3A, and developed diversifying base editors to achieve antibody affinity maturation ex vivo.87

Taken together, the multifaceted improvement of BEs (Table 1) greatly enriches the genome editing toolbox.

**Cytosine Base Editing Applications**

**Animals**

Base editing has been successfully applied in various animals (Table 2). Kim et al. started by microinjecting or electroporating gRNAs and the mRNA of BE3 into mouse embryos to make nonsense mutations in two genes, Dmd (dystrophin) and Tyr (tyrosinase),88 mutations that result in Duchenne muscular dystrophy and albinism, respectively. The corresponding phenotypes (e.g., loss of dystrophin expression and ocular albinism) were observed in the edited mice as expected. Liu et al. fused rA1 with dCas9-HF2 to develop HF2-BE2 and then successfully applied HF2-BE2 in mouse embryos to induce missense mutations in Tyr.89 Similarly, Yang et al. used the iSTOP strategy to introduce stop codons into V-set immunoregulatory receptor (VISTA) and CD160 antigen (CD160) in mouse embryos, and obtained homozygous mutants in F1 mice.90 Yang et al. also utilized BE3 to make point mutations (R17H in Hist1H3 and Hist2H3) or premature stop codons (Carni) in genes involved in chromatin modification to study the epigenetic regulation in mouse embryo development.91

Meanwhile, Sasaguri et al. compared the efficiencies of BE3 and Target-AID at Psen1 (presenilin 1) in mice and found that BE3 induced higher editing frequency than Target-AID.92 In addition, they also found that VQR-BE3 can generate efficient base editing at the target sites with altered PAM sequences.92 Recently, Li et al. compared the editing of BE3 and hA3A-BE3-Y130F at eight genomic loci, i.e., Tyr, Hoxd13 (homeobox D13), Ar (androgen receptor), Gfap (gliarial fibrillary acidic protein), Dmd, Lmma (Lamin A/C), Mecp2 (methyl CpG binding protein 2), Tnni3 (troponin I3) and Abcd1 (ATP Binding Cassette Subfamily D Member 1) in mice. They found that hA3A-BE3-Y130F induced higher editing efficiency in G/C-rich regions,93 consistent with the finding in mammalian cells.94 The mice with Ar mutation displayed the AIS (androgen insensitivity syndrome)-like sex reversal phenotype. Of note, Li et al. also did whole-genome sequencing for the Ar-mutant mice to determine OT effects but no significant base substitution was found at the potentials OT sites, which share similar sequence with ON site.93

Instead of editing zygotes, Chadwick et al. packaged BE3 into an adenoviral vector and delivered it into the livers of adult mice to generate premature stop codons in Pcsk9 (proprotein convertase subtilisin/kexin type 9) and found that both plasma PCSK9 and cholesterol levels were significantly reduced in the edited mice.94 Moreover, Chadwick et al. used the same strategy to mutate ANGPTL3, another gene involved in lipid metabolism, and reduced blood lipid levels in the mutant mice.95 In utero gene editing is another strategy to potentially treat genetic diseases that manifest significant morbidity or mortality. By using BE3 in utero, Rossidis et al. mounted an early stop codon in Pcsk9 in wild-type mice and a nonsense mutation in Hpd (4-hydroxyphenylpyruvate dioxygenase) in a mouse model of hereditary tyrosinemia type 1.96 Successful base editing of these genes reduced the cholesterol level...
and rescued the lethal phenotype of hereditary tyrosinemia type 1 in mice. Another BE, SaKKH-BE3, has also been used to treat a metabolic liver disease in adult mice. Importantly, Villiger et al. leveraged a split-intein moiety to split SaKKH-BE3 into two parts, both of which can be packaged into a single adeno-associated virus (AAV) vector, and the split-BE was applied to correct the Pah mutation in adult Pahenu2 mice.60 As expected, the 65% edited embryos.98 A premature stop codon in was amplification induced by BE3 and BE4-Gam at the loci of double-muscle, albino, or Hutchinson–Gilford progeria syndrome, respectively. Liu et al. also compared the editing efficacy induced by BE3 and BE4-Gam at the loci of Dmd and Tia1 and found that BE4-Gam induced higher editing frequency and product purity.97 In pigs, Li et al. used BE3 to mount a pathogenic E75K mutation in Be3 to induce mutations in Mstn (Myostatin), Dmd, Tia1 (cytotoxic granule-associated RNA binding protein 1), Tyr, and Lmna.97 The rabbits with Mstn, Tyr, or Lmna mutations showed corresponding phenotypes of double-muscle, albino, or Hutchinson–Gilford progeria syndrome, respectively. Liu et al. also compared the editing efficacy induced by BE3 and BE4-Gam at the loci of Dmd and Tia1 and found that BE4-Gam induced higher editing frequency and product purity.97 In pigs, Li et al. used BE3 to mount a pathogenic E75K mutation in twist-related protein 2 (Twist2) in porcine fetal fibroblast and then performed somatic cell nuclear transfer to generate edited embryos.98 A premature stop codon in Tyr was also induced in the similar way. The mutations in Twist2

Table 1. Base Editors with Various Characteristics

<table>
<thead>
<tr>
<th>Base editor</th>
<th>Deaminase</th>
<th>Cas</th>
<th>PAM</th>
<th>NLS</th>
<th>UGI</th>
<th>Codon optimization</th>
<th>Editing window (nt)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>BE1</td>
<td>rA1</td>
<td>dCas9</td>
<td>NGG</td>
<td>1</td>
<td>x</td>
<td>4–8</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>BE2</td>
<td>rA1</td>
<td>dCas9</td>
<td>NGG</td>
<td>1</td>
<td>x</td>
<td>4–8</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>BE3</td>
<td>rA1</td>
<td>nCas9</td>
<td>NGG</td>
<td>1</td>
<td>x</td>
<td>4–8</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>HF-BE3</td>
<td>rA1</td>
<td>HF-nCas9</td>
<td>NGG</td>
<td>1</td>
<td>x</td>
<td>4–8</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>HF2-BE2</td>
<td>rA1</td>
<td>dCas9-HF2</td>
<td>NGG</td>
<td>1</td>
<td>x</td>
<td>4–8</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td>BE4 &amp; BE4-Gam</td>
<td>rA1</td>
<td>nCas9</td>
<td>NGG</td>
<td>2</td>
<td>x</td>
<td>4–8</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>eBE-S1</td>
<td>rA1</td>
<td>nCas9</td>
<td>NGG</td>
<td>1</td>
<td>x</td>
<td>4–8</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>eBE-S3</td>
<td>rA1</td>
<td>nCas9</td>
<td>NGG</td>
<td>1</td>
<td>x</td>
<td>4–8</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>BE4max</td>
<td>rA1</td>
<td>nCas9</td>
<td>NGG</td>
<td>2</td>
<td>x</td>
<td>4–8</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>AncBe4max</td>
<td>rA1</td>
<td>nCas9</td>
<td>NGG</td>
<td>2</td>
<td>x</td>
<td>4–8</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>YE1-BE3</td>
<td>rA1-YE1</td>
<td>nCas9</td>
<td>NGG</td>
<td>1</td>
<td>x</td>
<td>4–7</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>YE2-BE3</td>
<td>rA1-YE2</td>
<td>nCas9</td>
<td>NGG</td>
<td>1</td>
<td>x</td>
<td>5–6</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>EE-BE3</td>
<td>rA1-EE</td>
<td>nCas9</td>
<td>NGG</td>
<td>1</td>
<td>x</td>
<td>5–6</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>YEE-B3</td>
<td>rA1-YEE</td>
<td>nCas9</td>
<td>NGG</td>
<td>1</td>
<td>x</td>
<td>5–6</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>VQR-BE3</td>
<td>rA1</td>
<td>nVQRCas9</td>
<td>NGA</td>
<td>1</td>
<td>x</td>
<td>4–11</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>VRER-BE3</td>
<td>rA1</td>
<td>nVRERCas9</td>
<td>NCG</td>
<td>1</td>
<td>x</td>
<td>3–10</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>SaBe3</td>
<td>rA1</td>
<td>nSaCas9</td>
<td>NNGRRT</td>
<td>1</td>
<td>x</td>
<td>3–12</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>SaBe4 &amp; SaBe4-Gam</td>
<td>rA1</td>
<td>nSaCas9</td>
<td>NNGRRT</td>
<td>1</td>
<td>x</td>
<td>3–12</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>SaKKH-BE3</td>
<td>rA1</td>
<td>nSaKKHCas9</td>
<td>NNNRRT</td>
<td>1</td>
<td>x</td>
<td>3–12</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>dCpf1-BE</td>
<td>rA1</td>
<td>dCpf1</td>
<td>TTTV</td>
<td>3</td>
<td>x</td>
<td>8–13</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>dCpf1-BE-YE</td>
<td>rA1-YE1</td>
<td>dCpf1</td>
<td>TTTV</td>
<td>3</td>
<td>x</td>
<td>10–12</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>dCpf1-eBE</td>
<td>rA1</td>
<td>dCpf1</td>
<td>TTTV</td>
<td>3</td>
<td>x</td>
<td>8–13</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>dCpf1-eBE-YE</td>
<td>rA1-YE1</td>
<td>dCpf1</td>
<td>TTTV</td>
<td>3</td>
<td>x</td>
<td>10–12</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>xBE3</td>
<td>rA1</td>
<td>nxcas9</td>
<td>NG, GAA, GAT</td>
<td>1</td>
<td>x</td>
<td>4–8</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>BE-PLUS</td>
<td>rA1-scFv</td>
<td>nCas9</td>
<td>NGG</td>
<td>1</td>
<td>x</td>
<td>4–14</td>
<td>86</td>
<td></td>
</tr>
<tr>
<td>hA3A-BE3</td>
<td>rA1</td>
<td>nA3A</td>
<td>nCas9</td>
<td>1</td>
<td>x</td>
<td>2–13</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>hA3A-BE3-Y130F</td>
<td>rA1</td>
<td>nA3A-Y130F</td>
<td>nCas9</td>
<td>1</td>
<td>x</td>
<td>3–8</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>hA3A-BE3-Y132D</td>
<td>rA1</td>
<td>nA3A-Y132D</td>
<td>nCas9</td>
<td>1</td>
<td>x</td>
<td>3–7</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>hA3A-eBE-Y130F</td>
<td>rA1</td>
<td>nA3A-Y130F</td>
<td>nCas9</td>
<td>1</td>
<td>x</td>
<td>3–8</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>hA3A-eBE-Y132D</td>
<td>rA1</td>
<td>nA3A-Y132D</td>
<td>nCas9</td>
<td>1</td>
<td>x</td>
<td>3–7</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>eA3A-BE3</td>
<td>rA1</td>
<td>nA3A-N575G</td>
<td>nCas9</td>
<td>1</td>
<td>x</td>
<td>4–8 (TpC)</td>
<td>79</td>
<td></td>
</tr>
<tr>
<td>eA3A-HF1-BE3-2xUGI</td>
<td>rA1</td>
<td>nA3A-N575G</td>
<td>nCas9-HF1</td>
<td>NNG</td>
<td>1</td>
<td>4–8 (TpC)</td>
<td>79</td>
<td></td>
</tr>
<tr>
<td>eA3A-Hypa-BE3-2xUGI</td>
<td>rA1</td>
<td>nA3A-N575G</td>
<td>nHypaCas9</td>
<td>NNG</td>
<td>1</td>
<td>4–8 (TpC)</td>
<td>79</td>
<td></td>
</tr>
<tr>
<td>DBE-A3A</td>
<td>rA1</td>
<td>nA3A-M2S</td>
<td>nCas9</td>
<td>NGG</td>
<td>1</td>
<td>x</td>
<td>2–17</td>
<td>87</td>
</tr>
<tr>
<td>Target-AID</td>
<td>pmCDA1</td>
<td>nCas9</td>
<td>NGG</td>
<td>1</td>
<td>x</td>
<td>2–8</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>Target-AID-NG</td>
<td>pmCDA1</td>
<td>nCas9-N9G</td>
<td>NG</td>
<td>1</td>
<td>x</td>
<td>2–8</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>TAM</td>
<td>rA1</td>
<td>hAIDx</td>
<td>nCas9</td>
<td>1</td>
<td>x</td>
<td>4–8</td>
<td>83</td>
<td></td>
</tr>
<tr>
<td>CRISPR-X</td>
<td>rA1</td>
<td>hAIDM2</td>
<td>nCas9</td>
<td>NGG</td>
<td>1</td>
<td>x</td>
<td>~50–50</td>
<td>85</td>
</tr>
<tr>
<td>DBE-AIDmono</td>
<td>rA1</td>
<td>hAID-mono2</td>
<td>nCas9</td>
<td>NGG</td>
<td>1</td>
<td>x</td>
<td>2–17</td>
<td>87</td>
</tr>
<tr>
<td>ABET7.9</td>
<td>rA1</td>
<td>TadA-TadA*</td>
<td>nCas9</td>
<td>NGG</td>
<td>1</td>
<td>x</td>
<td>5–9</td>
<td>12</td>
</tr>
<tr>
<td>ABET7.10</td>
<td>rA1</td>
<td>TadA-TadA*</td>
<td>nCas9</td>
<td>NGG</td>
<td>1</td>
<td>x</td>
<td>4–7</td>
<td>12</td>
</tr>
<tr>
<td>ABExmax</td>
<td>rA1</td>
<td>TadA-TadA*</td>
<td>nCas9</td>
<td>NGG</td>
<td>1</td>
<td>x</td>
<td>4–7</td>
<td>45</td>
</tr>
<tr>
<td>xABE</td>
<td>rA1</td>
<td>TadA-TadA*</td>
<td>nCas9</td>
<td>NG, GAA, GAT</td>
<td>1</td>
<td>x</td>
<td>4–7</td>
<td>63</td>
</tr>
<tr>
<td>VQR-ABE</td>
<td>rA1</td>
<td>TadA-TadA*</td>
<td>nVQRCas9</td>
<td>NGA</td>
<td>1</td>
<td>x</td>
<td>4–8</td>
<td>113, 106</td>
</tr>
<tr>
<td>VRER-ABE</td>
<td>rA1</td>
<td>TadA-TadA*</td>
<td>nVRERCas9</td>
<td>NCGG</td>
<td>1</td>
<td>x</td>
<td>4–8</td>
<td>106</td>
</tr>
<tr>
<td>ABEsa</td>
<td>rA1</td>
<td>TadA-TadA*</td>
<td>nSaCas9</td>
<td>NNGRRT</td>
<td>1</td>
<td>x</td>
<td>8–14</td>
<td>114</td>
</tr>
<tr>
<td>SaKKH-ABE</td>
<td>rA1</td>
<td>TadA-TadA*</td>
<td>nSaKKHCas9</td>
<td>NNNRRT</td>
<td>1</td>
<td>x</td>
<td>8–13</td>
<td>113, 106</td>
</tr>
</tbody>
</table>
and Tyr led to phenotypes of ablepharon macrostomia syndrome and albinism, respectively. These studies thus served as principles of proof that BEs can also induce efficient base editing in big animals such as pigs and rabbits.

BEs have also been successfully applied in nonmammals. Zhang et al. injected gRNA and the codon-optimized mRNA of BE3 into the zebrafish embryos at one-cell stage to induce mutations in Twist2, Gdf6 (growth differentiation factor 6), Ntl (no tail), and Tyr. The amino-acid change resulted from the Tyr mutation led to a phenotype that is similar to human ocular albinism. They also used VQR-BE3 to induce mutations in Twist2, Tial1, Tyr, Mtn, Dmd, Tia1, Tyr, and Lmna, respectively. These studies thus expanded the editing scope in vivo.

### Plants
As gene editing can help to improve the traits of crops, base editors have also been applied in various plants.
(Table 2). Zong et al. transfected the plasmids expressing codon-optimized BE3 (plant base editor, PBE) and individual gRNAs into protoplasts to mutate three rice genes (OsCDC48, OsNRT1.1B, and OsSPL14), a wheat gene (TaLOX2) and a maize gene (ZmCENH3). The editing of these genes succeeded, albeit with relatively low frequencies. The OsCDC48-edited rice, the TaLOX2-edited wheat and the ZmCENH3-edited maize were obtained by Agrobacterium-mediated transformation or plasmid delivery into immature plant embryos. More recently, Zong et al. also optimized hA3A-BE3 to develop the corresponding plant version (A3A-PBE). A3A-PBE induced very efficient base editing in wheat (TaALS, TaMTL, TaLOX2, TaDEP1, TaHPPD and TaVRN1-A1), rice (OsAAT, OsCDC48, OsDEP1, OsNRT1.1B, OsTOD, OsEV, and OsHPPD) and potato (StALS and StGBPSS) genes. In particular, A3A-PBE induced efficient editing at GpC sites across seven genomic regions in wheat and rice, whereas PBE induced virtually no editing events. The editing window of A3A-PBE spans from the first base to the seventeenth in protospacer regions in wheat and rice, whereas PBE induced virtually no editing events. The editing window of A3A-PBE (17-nt width) is suitable for mutagenesis-oriented editing such as iSTOP but may lead to more bystander mutations when precise editing is needed. Thus, it is interesting to know whether the hA3A-derived BEs that have narrowed editing windows (e.g., hA3A-BE3-Y130F and hA3A-BE3-Y132D) can also induce precise editing while maintaining efficiency in plants.

Independently, Li et al. and Lu et al. used BE3 and BE3-ΔUGI (rA1-ncas9 fusion) to induce base editing in rice genes (OsPDS, OsSBEIIb, NRT1.1B, and SLR1). By using AID as the deaminating module, Shimatani et al. and Ren et al. respectively developed the plant version of Target-AID (PmcDA-ncas9 fusion) and rBE9 (a BE containing ncas9, UGI and AID*Δ, a hyperactive AID mutant), which can induce base editing in rice genes (ALS and FTIP1e, OsAOS1, OsJAR1, OsJAR2, and OsCOII). Hua et al. also expanded the base editing scope in plants by replacing the Cas9 nickase of BE3 (CBE-P1 as the plant version) with VQR-Cas9 nickase and the latter BE (CBE-P3) induced base editing at the expanded targets site (PMS3).

**Bacteria**

Being a part of a prokaryotic immune system, CRISPR-Cas was first found to cleave targeted DNA in bacteria. Whether BEs can also be applied in bacteria has also been tested (Table 2). Banno et al. adopted the Target-AID system in bacteria to generate point mutations in various genes (galK, rpoB, xylB, manA, pta, adhE, and tpiA) in E. coli. Notably, Target-AID simultaneously induced 41 targeted mutations with four gRNAs against multicopy transposable elements. By using BE3, Zheng et al. also induced stop codons and missense mutations in rppH. In addition to E. coli, three reports generated early stop codons in bacterial strains *Brucella melitensis* (virB10), *S. aureus* (agrA, cntA, and esdA) and *Pseudomonas* (rhiR, rhiB, cadR, ompR, per, aspC, gacA, and hrpL) by using BE3 and BE3-UGI, and showed that base editing system can be widely applied in various prokaryotic species.

**Development and Application of Adenine Base Editors**

**Development and improvement of adenine base editors**

In addition to CBEs, ABEs have been developed recently to induce A-to-G base editing. After seven rounds of directed evolution in vitro, Gaudelli, Komor, and colleagues evolved a tRNA adenosine deaminase (TadA) of *E. coli* into a deoxyadenosine deaminase (TadA*) that can induce adenine to inosine (A-to-I) deamination on ssDNA. In order to improve the binding of TadA* to ssDNA substrate, a wild-type TadA monomer was fused at the N-terminus of TadA*. Then the TadA-TadA* heterodimer was further linked with nCas9 (D10A) to develop a series of ABEs (Table 1). Compared to CBEs, ABEs do not need a DNA glycosylase inhibitor, probably because inosines in DNA cannot be removed efficiently by any known mammalian DNA glycosylase. The subsequent MMR or DNA replication will use the inosine-containing strand as the template and then insert a cytosine opposite to the inosine, which eventually installs an A/T-pair to G/C-pair conversion at the target site.

Several versions of ABE have been developed since the milestone report by Gaudelli et al. (Table 1). By replacing the nCas9 with nxCas9 in ABE7.10, Hu et al. developed xABE, which expands the targeting scope of ABEs to the genomic regions containing NG, GAA or GAT PAM sequences. Other Cas9 variants with altered PAM specificity (i.e., VQR and VRER) have also been used to develop new ABEs, which recognize the PAM sequences NG and NGCG respectively. In addition to the SpCas9-containing ABEs, the SaCas9-containing ABE and its engineered form SaKKHCas9 have also been constructed to induce editing in the regions containing NNGRRT and NNNRRT PAM sequences. Recently, Koblan et al. also applied the codon-optimization and additional NLS strategy on ABE7.10, and the resulting ABEmax has an even higher editing efficiency compared to ABE7.10.
Application of adenine base editors

Several versions of ABEs have been applied in animals successfully (Table 2). Ryu et al. first used ABE7.10 to induce a missense mutation in Tyr gene in mice, which modeled the phenotype of the Himalayan mouse. Importantly, Ryu et al. also packaged ABE7.10 into a dual trans-splicing AAV system to correct a nonsense mutation (CAG to TAG) of Dmd gene in a mouse model of Duchenne muscular dystrophy, which they generated with BE3 previously. The expression of dystrophin was partially restored in the Dmd-corrected mice, suggesting the therapeutic potential of ABE. Liu et al. also used ABE7.10 to install mutations in AR and Hoxd13 genes in mice and the relevant phenotypes were observed. Of note, Liu et al. demonstrated that the SpCas9-derived ABE and the SaCas9-derived SaBE3 can induce efficient base editing at different target sites orthogonally. Liang et al. employed ABE7.10 to introduce mutations at the splicing sites of Tyr and Dmd genes and Duchenne muscular dystrophy-like phenotypes were observed in the Dmd-edited mice. Independently, Yang et al. generated mutations in Hbb-bs and Fah genes with ABE7.10 and the phenotypes of tyrosinemia type I were correspondingly observed in the mice containing the Fah mutation. Furthermore, Yang et al. used two new ABEs (SaKKH-ABE and VQR-ABE) they developed to induce mutations in Octc and Hbb-bs genes in mice.

In rats, Yang et al. installed mutations in Gaa with ABE7.10 and the abnormal accumulation of large lysosomes filled with glycogen in multiple tissues, a typical phenotype of Pompe disease, was found in multiple tissues of Gaa<sup>Del435Gly636Val</sup> F<sub>1</sub> offspring. In rabbits, Liu et al. utilized ABE7.10 to efficiently induce mutations in Dmd, Octc, and Sod1 genes, and the rabbits with T279A mutant of Dmd displayed typical clinical symptoms similar to human X-linked dilated cardiomyopathy. Together, these studies demonstrated that ABEs are efficient in editing mammals.

In addition to mammals, ABEs have been utilized in zebrafish. Qin et al. used the zebrafish-compatible zABE7.10 to introduce mutations in musk, rps14, atp5b, and wu:fc01d11 genes and the mutation of rps14 recapitulate the typical mutant phenotypes. And through applying the strategy succeeded in ABEmax (i.e. further codon-optimization and introducing more NLS) Qin et al. also developed zABE7.10max.

Meanwhile, many plant-compatible ABEs have been rapidly developed and applied as well (Table 2). Hua et al. generated the ABE-P1 (ABE7.10) and ABE-P2 (ABEesa) to induce mutations in six rice genes (SLR1, OsSPL14, OsSPL16, OsSPL17, OsSPL18, and LOC_Os02g24720) and Yan et al. constructed rBE14 (ABE7.10) to introduce mutations in four rice genes (OsMPK6, OsMPK13, OsSERK2 and OsWRKY45). Around the same time, Li et al. developed seven versions of PABE that have different TadA-TadA*/nCas9/NLS configurations and found that PABE-7, which has the same configuration of ABE7.10 but with three copies of NLS at C-terminus, induced the highest editing efficiency. Then, Li et al. used PABE-7 to introduce mutations in eight rice genes (OsACC, OsALS, OsCDC48, OsAAAT, OsEV, OsOD, OsDEP1, OsNRT1.1B) and three wheat genes (TaDEP1, TaEPSPS, TaGW2), and one resulted rice strain bearing C2186R substitution in OsACC is resistant to herbicide. More recently, Kang et al. used the plant-compatible ABE7.10 (pcABE7.10) to induce mutations in four Arabidopsis genes (AtALS, AtPDS, AtFT, and AtLFY) and two rapeseed genes (BnALS and BnPDS). The AtFT-targeted plants displayed a late-flowering phenotype and the AtPDS3-targeted plants showed a range of dwarfism and mosaic albino phenotypes, both confirming that ABE can be used to alter the phenotype of Arabidopsis. In order to expand the editing scope in plants, Hua et al. developed ABE-P3 (VQR-ABE), ABE-P4 (VRER-ABE), and ABE-P5 (SaKKH-ABE) and utilized these plant versions of ABEs to install mutations in seven rice genes (OsSPL14, OsSPL16, OsSPL17, OsSPL18, OsTOE1, OsIDS1, and SNB).

### Base Editing or CRISPR-Cas?

With both serving as precise and efficient gene editing tools, base editing overlaps with CRISPR-Cas in certain applications. Choosing between them in such circumstances becomes a sweet burden. Base substitution is by definition the kind of gene editing that is well suited to BE. As BE directly catalyzes the deamination of cytosine to install C-to-T conversions in targeted DNA, the process of base editing is generally independent of DSB or a DNA donor. With recent improvements, the editing efficiencies of the latest versions of BE can reach ~50–70% in various mammalian cell lines. In contrast, the base substitutions installed by CRISPR-Cas9-mediated HDR rely on homologous recombination, which is cell-cycle/type dependent and generally manifests low efficiency in most mammalian cells. Thus, BE is an appropriate tool to be chosen when targeted base substitution is desired.

Another type of gene editing, gene knockout, can be performed with either Cas or BE. When combined with one or more gRNA, Cas can generate DSB to trigger NHEJ, which will result in indels at DSB sites. This method can be broadly used to knockout various genes and DNA fragments. Yet, it is worth noting that in some studies, the generation of DSB triggered a p53-mediated DNA damage response and/or cause...
unintended large deletions, translocations or DNA rearrangements.\textsuperscript{126,127} In contrast, BEs can be used to terminate the production of a functional protein by creating premature stop codons (iSTOP or CRISPR-STOP), which avoids the generation of DSB.\textsuperscript{80,81} However, the iSTOP or CRISPR-STOP strategy is hard to be used in knocking out genes that do not express protein products (e.g., non-coding RNA)\textsuperscript{128} and Cas9 or Cpf1 would therefore be a better choice in these situations.

Insertion or deletion of a specific DNA fragment at a specific target site is another common type of gene editing, which can be achieved with Cas though HDR, but not with BE. The NHEJ repair outcome of DSB can be predicted or even designed with high accuracy, depending on the sequence of cleavage site.\textsuperscript{129–131} Thus, some specific deletions and insertions can be directly generated by Cas9 without the need of a donor DNA, which may be used to overcome the limitation of HDR efficiency in some cases.

Delivery is another factor to consider when weighing BE and Cas for \textit{in vivo} editing. Though various viral systems have been used to deliver Cas,\textsuperscript{132–134} the size of most Cas is reaching the packaging limit of AAV, a popular vector for \textit{in vivo} editing.\textsuperscript{135} Because a nucleobase deaminase module needs to be fused to a Cas in BE, the size of BE is even bigger than that of Cas. Although BEs can be split and efficiently packaged into two separate AAVs,\textsuperscript{80} the nonviral delivery system seems more promising for BEs, especially when considering the future addition of more functional modules into BEs. Recently, Yeh \textit{et al.} packaged the gRNA/BE3 RNP in cationic lipid nanoparticles and injected the RNP-containing nanoparticles into the inner ears of mouse pups. The post-mitotic editing installed an S33F mutation in $\beta$-catenin (\textit{CTNNB1}), which induces cell reprogramming.\textsuperscript{136} In the future, the specificity and efficiency of nonviral delivering system\textsuperscript{137} can be further improved to potentiate clinic-related applications.\textsuperscript{138}

**Perspective**

Since the development of the first BE just a few years ago, base editing has undergone rapid expansion.\textsuperscript{77,139,140} Current BE systems perform the transition of bases (i.e., a purine to a purine or a pyrimidine to a pyrimidine).\textsuperscript{11,12} In the future, the development of new BEs that can induce transversions—a purine to a pyrimidine or vice versa—will enrich the repertoire of base editing. Currently, base editing at the “single” base level has not been realized in most cases. The development of new base editors with a 1-bp editing window, while still keeping high editing efficiency, will further improve editing precision. Similar to the OT effects of Cas-mediated gene editing, unintended base substitutions induced by BEs at OT sites will remain a concern for basic research and potential therapeutic applications. Besides, mutations at other unexpected sites (e.g., the sites that gRNAs do not bind to) may also be induced by the nucleobase deaminase fused in BEs.\textsuperscript{71,72,141} Thus, the development of new methods for the genome-wide detection of BE-induced mutations in human somatic cells and new BEs with even higher editing specificity will further expand the application of base editing system, particularly in therapeutics.

**Author Disclosure Statement**

J.C. is supported by funding from Ministry of Science and Technology (grant number 2018YFC1004602) and National Natural Science Foundation of China (NSFC) (grant numbers 31822016, 81872305, and 31600654). B.Y. is supported by NSFC (grant number 31600619). L. Y. is supported by NSFC (grant number 31730111).

**References**

2012.03.030.


70. Rees HA, Komor AC, Yeh WH, et al. Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery. Nat Commun 2017;8:15790. DOI: 10.1038/ncomms15790.


129. Yang AT AL.